Overview of runs-type signalling rules

by

Rauldo Celliers

Submitted in partial fulfillment of
the requirements for the degree

Magister Scientiae (Mathematical Statistics)

in the
Faculty of Natural & Agricultural Sciences
University of Pretoria
Pretoria

April 2016

Declaration

I declare that the thesis, which I hereby submit for the degree Magister Scientiae (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

Thanks go to the Statistics Department at UP for granting me the opportunity to complete my master's degree there. I would want to thank all the lecturers which I came into contact with during my undergraduate as well as postgraduate studies. Their contribution is often taken for granted and deserves acknowledgement.

I would like to sincerely thank my supervisor and co-supervisor, Dr. S.W. Human and Dr. M.A. Graham for their contribution towards the thesis. Their feedback and timely responses were much appreciated. I am also grateful for all their insights and knowledge that they were willing to share as well as the time they put aside to guide me.

I am also grateful for the support and understanding of my family in completing the thesis. They provided me with motivation and strength in order to complete the tasks at hand.

Summary

Runs-type signalling rules and sensitizing rules (Page (1955)) were originally introduced to detect non-random patterns in the process being monitored (Montgomery (2013)). Further development led to runs-type signalling rules being applied in order to monitor small shifts or a combination of small and large shifts (Khoo et al. 2006) in the process parameter(s) of interest; the detection of small process shifts in location and spread being a drawback of the traditional Shewhart control charts.

Chapter 1 provides an introduction to the topic and discusses the Shewhart control chart as well as the design of the control chart. The problem statement as well as the scope of the study is provided. Runs-type signalling rules are defined and an overview of the research on this subject is given. This is followed by a discussion on the run-length random variable as well as the Markov-chain approach to evaluating/analysing the run-length distribution; this approach is emphasized (as opposed to simulation and/or closed form expressions) as it is well suited to derive/describe the run-length distribution in case of Shewhart control charts supplemented with runs-type signalling rules.

Chapter 2 provides an overview and in-depth discussion regarding parametric runs-type signalling rules followed by an overview and discussion on nonparametric rules in **Chapter3**.

Chapter 4 provides an overview of multivariate runs-type signalling rules.

Chapter 5 provides a general discussion on the design and implementation of Shewhart-type control charts supplemented with runs-type signalling rules.

Chapter 6 concludes this thesis with a summary of the research done. Concluding remarks concerning unanswered questions and proposals for future research are also given.

Table of Contents

Chapter 1

Introduction

1.1. Notation

Notation is important to ensure consistency and avoid ambiguity throughout the thesis. Hence the table below lists some of the abbreviations and notation that will be used frequently throughout the thesis. Where deviation from the notation occurs clarification will be provided.

Table 1.1. Abbreviations and notation

Notation/Abbreviation	Definition/Description
SPC	Statistical process control
cdf	Cumulative distribution function
n	Sample size / rational subgroup size
$X_1, X_2, \dots, X_n$	Random variables in a sample
$x_1, x_2, \dots, x_n$	Observations in a sample
ARL	Average run-length
ARL_0	In-control average run-length
ARL_ξ	Out-of-control average run-length
ξ	Shift in the process parameter
$SDRL$	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
IC	In-control
OOC	Out-of-control
TPM	Transition probability matrix
A	Absorbent
NA	Non-absorbent
UCL_2	Outer upper control limit
UCL_1	Inner upper control limit
LCL_2	Outer lower control limit
LCL_1	Inner lower control limit
LWL	Lower warning limit
UWL	Upper warning limit
T	The run-length random variable of the control chart

1.2. Terminology and problem statement

Two important problems in normal SPC are monitoring the process mean and / or the process standard deviation. More generally, monitoring the center or the location (or a shift) parameter and/or a scale parameter of a process is of interest. The location parameter represents a typical value and could be the mean or some percentile, such as the median of the distribution; the latter is especially attractive when the underlying distribution is expected to be skewed. Let $F(x)$ denote the unknown cumulative distribution function (cdf) of the monitored continuous variable X. It is assumed that F follows either

i. a location model, with a cdf $F(x-\theta)$, where $x\epsilon(-\infty,\infty)$ and $\theta\epsilon(-\infty,\infty)$ is the location parameter, or,

ii. a scale model, with a cdf $F\left(\frac{x}{\tau}\right)$, where $x\epsilon(-\infty,\infty)$ and $\tau>0$ is the scale parameter, or,

iii. a location-scale model with cdf $F\left(\frac{x-\theta}{\tau}\right)$, where $x\epsilon(-\infty,\infty)$ and $\theta\epsilon(-\infty,\infty)$ and $\tau>0$ are the location and the scale parameter, respectively.

Thus, the usual problem is to track θ or τ or both, under these model assumptions, based on random samples or rational subgroups of data usually taken at equally spaced time points.

1.3. The Shewhart control chart

One of the main aims of statistical process control (SPC) is to detect and reduce assignable causes of variation; this type of variation typically reduces the quality of a product/process and should thus ideally be controlled and removed. A control chart is a visual representation of the charting statistic on the vertical axis versus time or sample (subgroup) number on the horizontal axis as well as the associated control limits. A typical two-sided Shewhart-type control chart (Walter A. Shewhart developed the statistical control chart concept in 1924) is shown in Figure 1.1 consisting of a center line (CL) and two horizontal lines, one on each side of the CL as well as the charting statistic. The line above the CL is called the upper control limit (UCL) whereas the line below the CL is called the lower control limit (LCL).

For the standard Shewhart $\bar{X}$ control chart these control limits are symmetric since the control chart is developed for an IC ARL (ARL_0) of between 370 and 500 resulting in a typical *FAR* of 0.0027 and 0.002 respectively. The process can either be in-control (IC) or out-of-control (OOC) depending on where the charting statistic(s) plots. For the standard Shewhart control chart the chart signals and the process is declared OOC if the charting statistic plots on or above the UCL or on or below the LCL. On the contrary, the process is declared IC if all the charting statistics plot between the control limits and there is no non-random pattern present in the charting statistics. Once a signal has occurred a search for assignable causes is started and corrective action is initiated.

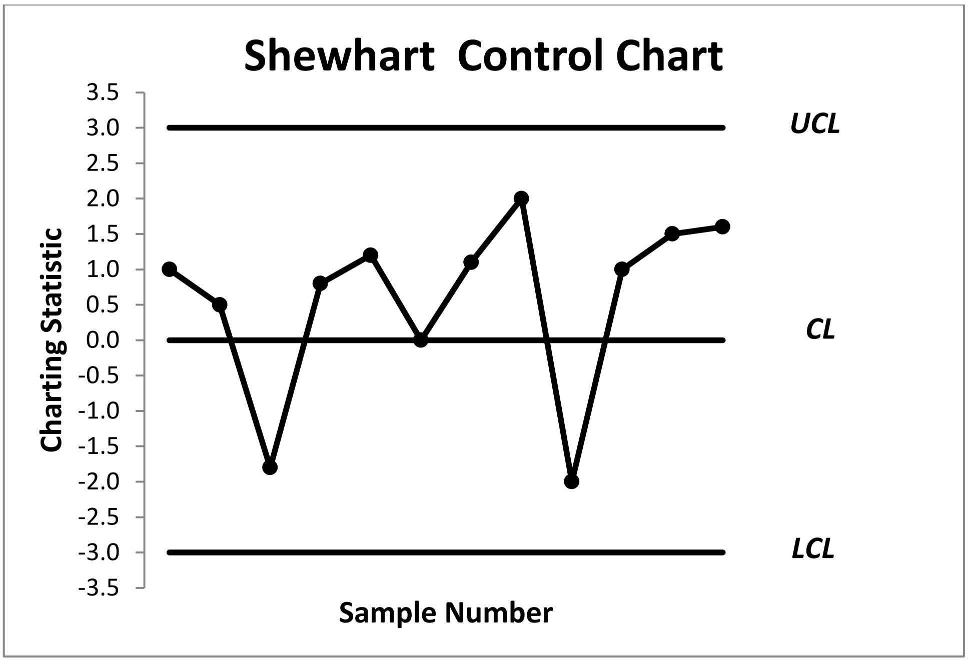

Figure 1.1: A two-sided Shewhart control chart.

To describe the Shewhart control chart in more detail, assume that $X_{t1}, X_{t2}, \dots, X_{tn}$ denote a random sample (i.e. measurements from some quality characteristic) of size $n \geq 1$ from the process at time $t = 1,2,3, \dots$. Suppose the sample statistic that is used to monitor the quality of the process is denoted by V. So, for a sample $t = 1,2,3$ .... the statistic would be denoted by V_t. In the case of the Shewhart chart, the sample statistic is also the charting statistic.

For the charting statistic V calculated, let μ_v and σ_v denote the known mean and standard deviation, respectively. The control limits and CL are then equal to the following equations:

$$\begin{aligned} UCL &= \mu_v + L\sigma_v \\ CL &= \mu_v \\ LCL &= \mu_v - L\sigma_v \end{aligned} \tag{1.1}$$

where L represents the distance of the upper-and-lower control limits from the CL expressed in standard deviation units and is called the charting constant. This value will typically be chosen to be 3 and if so, the control limits are referred to as the "three sigma" control limits.

If the process parameters are unknown they are usually calculated using at least $m > 1$ (e.g. 20 to 25) samples of size $n > 1$ (e.g. 4 to 5) each, see e.g. Montgomery (2013). The dataset available for designing the control chart i.e. estimating the control limits, has $N = m \times n$ observations in total. The estimated control limits and center line are then given by:

$$\begin{aligned} \widehat{UCL} &= \hat{\mu}_v + L\hat{\sigma}_v \\ \widehat{CL} &= \hat{\mu}_v \\ \widehat{LCL} &= \hat{\mu}_v - L\hat{\sigma}_v \end{aligned} \tag{1.2}$$

where $\hat{\mu}_v$ and $\hat{\sigma}_v$ denote unbiased point estimators of the process parameters. Note the difference between Equation (1.1) and Equation (1.2) – in Equation (1.2) we use the “hat” notation to indicate that the control limits and center line are random variables as they are functions of the point estimators $\hat{\mu}_v$ and $\hat{\sigma}_v$ i.e. they are all functions of the observed data.

If the process parameters are unknown we distinguish between phase I and phase II analysis. The aim of a phase I analysis is to establish whether the process is IC. The control limits are adjusted continually until no more assignable causes of variation are present. Thus during a phase I analysis an estimate for the unbiased point estimators denoted in Equaion (1.2) would be used in setting up the control limits. On the other hand during a phase II analysis the control limits are fixed and the aim is to monitor the process. In the case where both the mean and standard deviation are unknown, point estimators are used to calculate the control limits as shown in Equation (1.2).

1.4. Scope of this study

To clarify the scope of this study and to position the research carried-out in this thesis, the table below provides a broad classification of the different scenarios and different types of control charts in the SPC environment. The cells marked with an 'x' in Table 1.2 indicate the various control charts being discussed for the purposes of this study.

Table 1.2: Broad classification of control charts in the SPC environment.

<table>
<tr><td colspan="2" rowspan="3"></td><th colspan="6">Univariate</th><th>Multivariate</th></tr>
<tr><th colspan="3">Parametric</th><th colspan="3">Nonparametric</th><td rowspan="5">x</td></tr>
<tr><th>Shewhart</th><th>CUSUM</th><th>EWMA</th><th>Shewhart</th><th>CUSUM</th><th>EWMA</th></tr>
<tr><th colspan="2">Case K</th><td>x</td><td></td><td></td><td>x</td><td></td><td></td></tr>
<tr><th rowspan="2">Case U</th><td>Phase I</td><td>x</td><td></td><td></td><td>x</td><td></td><td></td></tr>
<tr><td>Phase II</td><td>x</td><td></td><td></td><td>x</td><td></td><td></td></tr>
</table>

Table 1.2 shows the following:

(i) The overall split between univariate and multivariate control charts. Univariate charts are used when only one process parameter or characteristic is of interest and monitored. Multivariate charts are used to simultaneously monitor multiple i.e. two or more, characteristics with one control chart.

(ii) Whether the sample statistic (and hence the control chart) is parametric or non-parametric in nature. A nonparametric control chart procedure is defined as a control chart which has the same IC run-length distribution for every continuous distribution; see for example Chakraborti et al. (2004, 2007). These charts are typically based on nonparametric statistics such as the Sign test or Wilcoxon Signed-Rank test etc. (see e.g. Gibbons and Chakraborti (2011)).

A parametric control chart, on the other hand, assumes a particular underlying process distribution (such as the normal distribution) and is based on sample statistics such as the sample mean ($\bar{X}$) and the sample variance (S^2). Hence, as the underlying process distribution changes so does the distribution of the sample statistic and that of the charting statistic and ultimately, the control chart's properties change.

(iii) Whether the control chart is a Shewhart-, EWMA- or CUSUM-type control chart. EWMA- or CUSUM-type control charts supplemented with runs-type signalling rules exists in the literature, see e.g. Riaz et al. (2011) or Abbas et al. (2011), but won't be discussed since the focus of this study is Shewhart-type control charts supplemented with runs-type signalling rules.

(iv) Whether the process parameters are known (indicated as Case K) or unknown (indicated as Case U); this is important because in case the parameters are unknown there are two phases i.e. Phase I and Phase II (see below).

Other/additional considerations which are not included in Table 1.2 are:

(i) The sample size, i.e. whether individual observations ($n = 1$) are used and/or if rational subgroups of size ($n \geq 1$) are used. In this study we focus on both $n = 1$ and $n \geq 1$.

(ii) In case of the parametric charts one has to assume a particular process distribution, e.g. continuous or discrete. In this thesis we specifically focus on univariate and multivariate Shewhart-type control charts with runs-type signalling rules used to monitor the location or scale of a distribution; this includes procedures/control charts that are used to monitor univariate and multivariate distributions. To simplify matters it is assumed that in case of parametric charts the process follows a normal distribution. However, in case of the nonparametric charts, it is only assumed that the underlying process is continous unless the additional requirement of symmetry is needed, e.g. in case of the Wilcoxon Signed-Rank test.

(iii) The correlation structure (i.e. dependence or independence) of the samples and observations within each sample. In this study, for the parametric charts we assume that the individual observations i.e. X_{tj}, are independent and identically distributed (iid), e.g. $X_{tj} \sim N(\mu, \sigma^2)$ $t = 1,2, \ldots$ $j = 1, \ldots, n$ with μ and σ^2 being known (case K) or unkown (case U) as mentioned previously.

(iv) The process characteristic or process parameter being estimated, e.g. the location and/or the spread or any other important characteristic as indicated by the quality practitioner, e.g. a percentile.

(v) The sample statistic used to estimate and monitor the process parameter (mentioned in point (iv) above). For example, when the location is monitored one can use sample statistics such as the mean, the median or the trimmed mean or nonparametric statistics such as the Sign-test or the Wilcoxon Signed-Rank test. Each of the sample statistics has its own pros and cons and justification to be used in a particular setting.

(vi) The signalling rules used to declare a process OOC:

a. the chart signals when a single charting statistic plots on or outside the control limits.
b. the chart signals when $k \geq 2$ consecutive charting statistics all plot on or outside the control limits.
c. the chart signals when *k-of-w* or at least *k-of-w* (where $k \leq w, k \geq 2$) charting statistics plot on or outside the control limits.

As can be seen from the above discussion, there are clearly a multitude of considerations when designing a control chart with each combination of the above-mentioned choices leading to a unique chart.

Other types of control charts such as the Synthetic control charts, Average Run-length unbiased control charts, Bayesian control charts, control charts based on a change-point model and/or charts used for profile monitoring are described in the literature but fall outside the scope of this study.

1.5. General steps to design a control chart

This Section focuses on the general steps followed when designing a univariate parametric control chart. These steps are closely linked to the considerations mentioned in Section 1.4 and are illustrated in Figure 1.2 below.

1
- Identify or choose the underlying process distribution (say $F(x)$).
- Identify the correlation structure (iid or non-iid observations).
- Example: Suppose $X_{tj} \sim iid\ F(x)$.

2
- Decide on the process characteristic that has to be monitored e.g. the location or spread.
- Decide on the associated sample statistic to be used e.g. $\bar{X}_t$, S_t , R_t etc.
- In general, denote this sample statistic by $L_t = f(x_{t1}, \dots, x_{tn})$ where f is some function of the sample observations.
- Example: Given the assumption in Step 1 i.e. $X_{tj} \sim iid\ F(x)$, it follows that $L_t \sim iid\ G(x)$ if the process is in control; where $G(x)$ denotes the underlying distribution function of the sample statistic. In general, it will be the case that $F(x) \neq G(x)$ i.e. the distribution of X_{tj} is not the same as that of the sample statistic.

3
- Decide on the type of control chart e.g. Shewhart, CUSUM or EWMA. This will identify the charting statistic to be calculated.
- Example: In the case of the Shewhart-type chart, $C_t = L_t$ i.e. the sample statistic is the charting statistic. In the case of the EWMA- and CUSUM-type charts, $C_t = h(L_1, L_2, \dots, L_t)$ i.e. C_t is a function of all the sample statistics from time period t= 1 (L_1) until the most recent or latest sample statistic (L_t).
- It follows that $C_t \sim H(x)$ where in general, $H(x) \neq G(x) \neq F(x)$; $H(x)$ denotes the underlying distribution function of the charting statistic. It should be clear that $H(x)$ will depend on $G(x)$ and $G(x)$ will depend on $F(x)$ i.e. the distribution of the charting statistic depends on the distribution of the sample statistic and the distribution of the sample statistic depends on the underlying process distribution.

4
- Decide on the signalling rule.
- Example: Process is said to be out of control if *k-of-w* points plot on or outside the control limits.

5
- Obtain the run-length distribution of the control chart and find or solve for the charting constants so as to obtain the desired performance. The design criteria are generally based on the properties of the run-length distribution.
- Example: The chart is designed to obtain a specific nominal false alarm rate (*FAR*) or specific in-control average run-length (ARL_0).

Figure 1.2: The Steps in designing a control chart.

1.6. Signalling rules

Sensitizing rules have been proposed to increase the sensitivity of control charts (see, for example, Page (1955), the Western Electric Company (1956), Roberts (1958), Bissell (1978), Wheeler (1983) and Nelson (1984)). Sensitizing rules are signalling rules designed to detect some improbable and/or non-random pattern of the charting statistics on the control chart (see e.g. Montgomery (2013)).

These sensitizing rules typically include upper/lower warning/control limits or a combination of upper/lower warning/control limits. The aim is to monitor a sequence of points in order to establish whether the control chart is IC or OOC. Control limits, also referred to as action limits, signal if a point plots on or outside the limit or in a sequence of consecutive points a certain number plot on or above the warning (control) limits; note that when considering inner and outer warning (control) limits a signal would be produced if in a sequence of consecutive points a certain number, e.g. ‘x’ plot between the inner and outer limits or on the inner limit, where ‘x’ depends on the specific signalling rule. Once a signal is produced a search for assignable causes is undertaken and possible corrective action taken. Warning limits on the other hand are used with control limits to establish whether a process is possibly OOC. Figure 1.3 provides an example of a sensitizing rule with warning and control limits, see Western Electric rule number two.

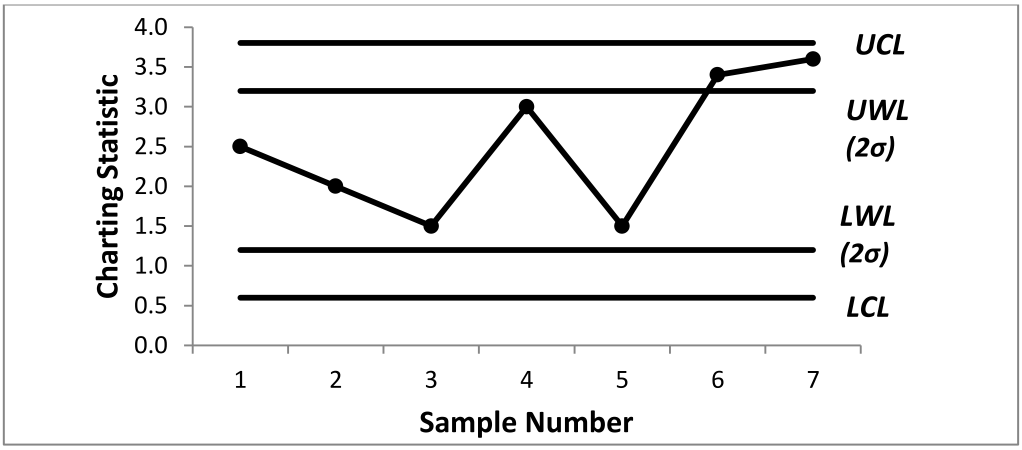

Figure 1.3: Sensitizing rule with two control and warning limits.

As an example the chart would provide a warning signal if two out of three consecutive points fall within the range (UWL, UCL) or (LWL, LCL). Later on in the study when discussing runs-type signalling rules these warning limits are replaced by lower and upper inner control limits (LCL_1,UCL_1). To provide clarity on control limits used later in the thesis Figure 1.4 provides a visual representation of the inner and outer control limits of a two-sided control chart; in Figure 1.3 the UCL/LCL is replaced by UCL_2 /LCL_2 and the $LWL/$ UWL being replaced by LCL_1/UCL_1. A signal, as indicated by the star, is then produced at sample number 7 for the rule evaluating the last two out three charting statistics, see Definition 1.1 and 1.2.

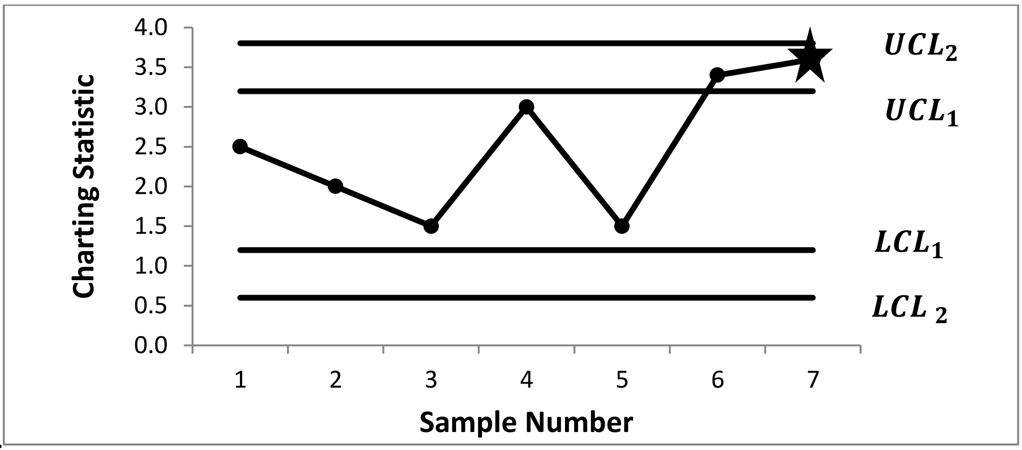

Figure 1.4: Two-sided control chart with outer and inner control limits.

The Western Electric rules make use of these warning limits as mentioned previously to determine the state of the process (IC/OOC). The Western Electric rules (Western Electric Company (1956)) are provided below; these rules consist of warning limits, typically in the range of one or two-sigma and also control limits, see Figure 1.5 and 1.6.

1. One or more points outside the control limits.
2. Two of three consecutive points outside the two-sigma warning limits but still inside the control limits.
3. Four of five consecutive points above the one-sigma limits.
4. A run of eight consecutive points on one side of the CL.
5. Six points in a row steadily increasing or decreasing.
6. Fifteen points in a row between the CL and the one-sigma limits.
7. Fourteen points in a row alternating up and down.
8. Eight points in a row on both sides of the center line but none in the area between the CL and the one-sigma upper and lower warning limits.
9. An unusual or non-random pattern in the data.

10. One or more points near a warning or control limit.

A visual representation of the first four Western Electric rules are provided in Figure 1.5 and 1.6.

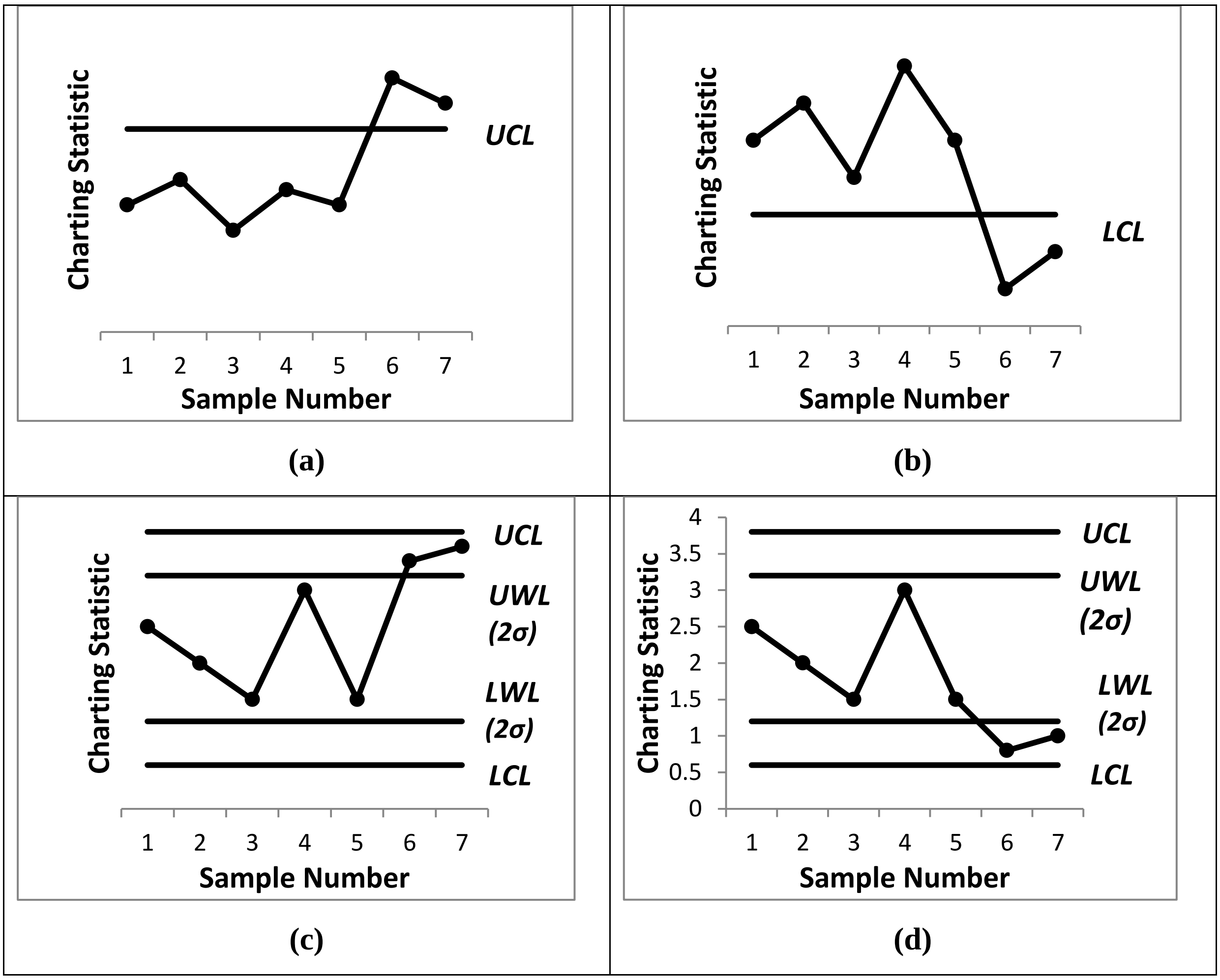

Figure 1.5: The first two Western Electric rules.

For the Western Electric rule number two (Figure 1.5 (c) and (d)) the warning limits function as inner upper and lower control limits producing a warning if two out of the last three points are observed in the following ranges: (UWL, UCL) or (LWL, LCL).

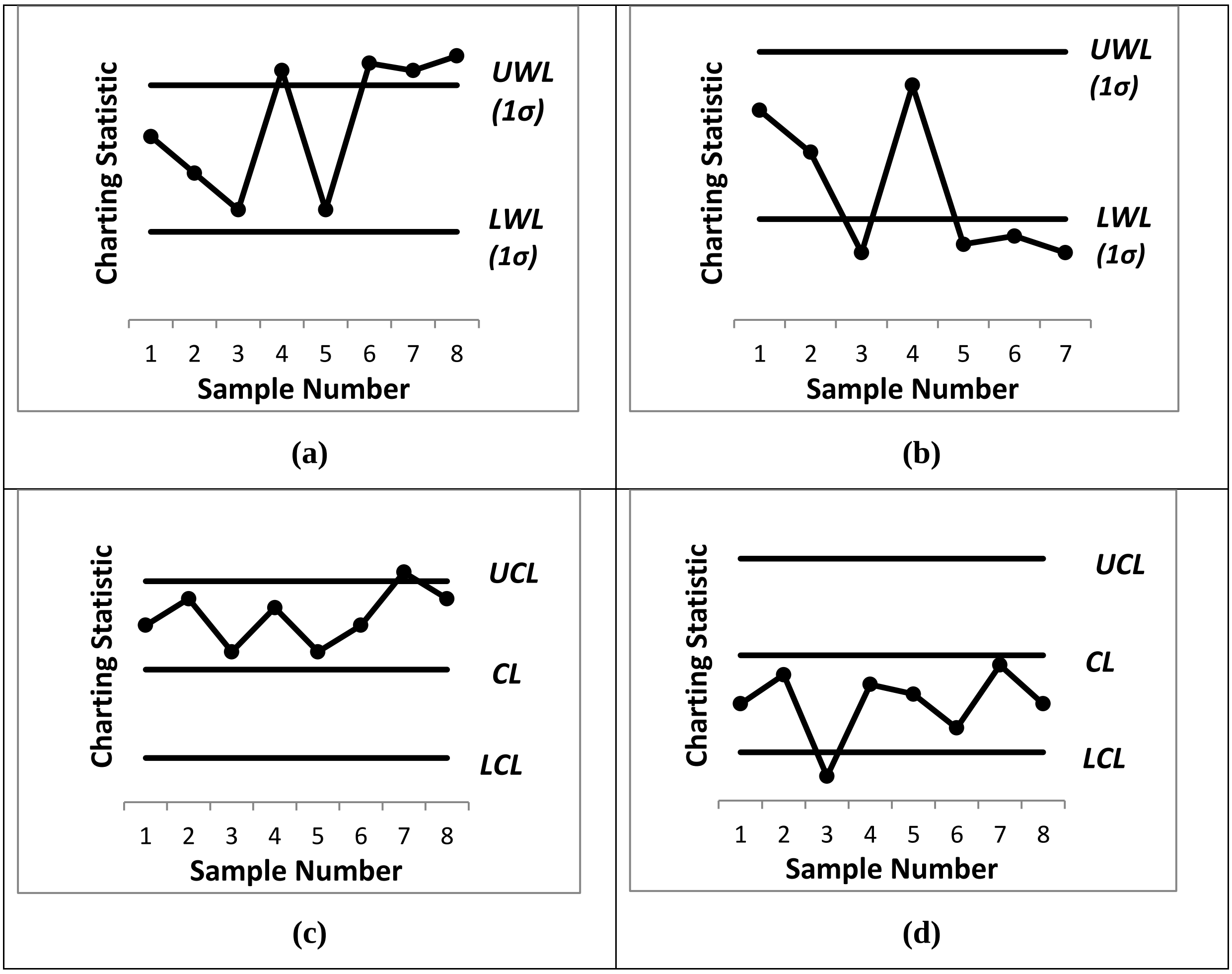

Figure 1.6: Western Electric rule number 3 and 4.

In general, control charts could either be one or two-sided depending on the presence or absence of certain control limits with one-sided control charts subdivided further into upper or lower control charts. One-sided charts are identified by the presence of either an upper or lower control/warning limit while for two-sided charts both upper and lower control (warning) limits should be present. From a practical perspective the engineer might only be interested whether a process breaches some limit value which would lead to an upper control chart or when monitoring the variance of a process only positive variances are considered. Figure 1.5 (a) and (b) are typical one-sided control charts while the graphs in Figure 1.5 (c)-(d) as well as Figure 1.6 are two-sided control charts.

Different signal criteria apply to the graphs illustrated in Figure 1.5 and 1.6. Figure 1.5 (a) and (b) give a warning signal if one or more points plot on or outside the UCL or LCL respectively - both Figure 1.5 (a) and (b) provide a warning signal at sample number 6. Figure 1.5 (c) and (d) give a warning signal if two of three consecutive points plot on or outside the two-sigma warning limits but still inside the UCL in case (a) or LCL for (b) – both provide a warning signal at sample number 7. Figure 1.6 (a) and (b) give the warning signal if four out of five consecutive points plot on or above/below the one-sigma limits (UWL/LWL) – both providing the warning signal at sample number 8. Figure 1.6 (c) and (d) give the warning signal if a run of eight consecutive points plot on one side of the CL producing a warning signal at sample number 8 for both.

Building on the sensitizing rules, runs-type signalling rules were developed to increase the sensitivity of a Shewhart-type control chart in detecting either a small shift in location or spread. A "run" is generally used to describe a sequence or succession of objects or observations of the same type in an ordered sequence (Chakraborti and Graham (2007)). In this case a run refers to a sequence of charting statistics. Note that the signalling rules available in the literature have been incorporated in both parametric and nonparametric control charts with parametric control charts subdivided into univariate and multivariate control charts (see e.g. Table 1.2).

Different types of signalling rules

The simplest or most basic signalling rule uses information only from the last sample to declare whether the process is IC or OOC (known as the *1-of-1* signalling rule), and consequently, the *1-of-1* signalling rule is relatively insensitive to small shifts in the process (see Klein (2000a)) and accordingly, other signalling rules were developed. These additional rules include, for example, the rule that declares the process OOC if the last two charting statistics plot on or outside the control limits (i.e. the *2-of-2* signalling rule) and the rule that declares the process OOC if two of the last three charting statistics plot on or outside the control limits (i.e. the *2-of-3* signalling rule). The *2-of-2* and *2-of-3* signalling rules are also known as runs-type signalling rules (see e.g. Klein (2000a)).

The *1-of-1* signalling rule is a member of the *k-of-k* signalling rules class. The *k-of-w*, the improved *k-of-w* and the revised *k-of-w* are all runs-type signalling rules developed to determine if the process is OOC. Figure 1.4 is an example of the standard *k-of-w* runs-type signalling rule where $k = 2$ and $w = 3$; the definitions and discussions follow in the following pages. The improved *k-of-w* is a combination of the *1-of-1* and *k-of-w* runs-type signalling rule. This provides the control chart with both large (*1-of-1*) and small (*k-of-w*) shift detection capability. The revised *k-of-w* runs-type signalling rule is similar to the improved *k-of-w* in the sense that it uses both the *1-of-1* and the *k-of-w* runs-type signalling rules in order to determine whether a process is OOC. The difference between the improved and revised *k-of-w* chart is that the actual control limits differs; it is claimed that the revised *k-of-w* improved on the signal detection capability of the improved *k-of-w*. Note that when considering the revised *k-of-w* control charts all the charting statistics in the run of consecutive points being evaluated should plot on one side of the CL, while for the improved *k-of-w* control charts this condition is not mandatory.

Definitions 1.1 and 1.2 provide the signalling rules for the *k-of-k*, the *k-of-w*, the improved *k-of-w* and the revised *k-of-w* signalling rules, respectively (Note that the *k-of-k* is a variant of the *k-of-w* rule).

Definition 1.1

OOC conditions for *k-of-w* runs-type signalling rule
In a run of w consecutive points, k of these points must plot on or outside the control limits. For the upper-and-lower one-sided charts k points must plot on or outside the UCL or LCL, respectively.

Definition 1.2

OOC conditions for the improved/revised *k-of-w* runs-type signalling rules
A point plots on or beyond either the UCL_2 or LCL_2. In a run of w consecutive points, k of these points must plot between $UCL_1(LCL_1)$ and $UCL_2(LCL_2)$; note that if a point plots on the $UCL_1(LCL_1)$ it is considered inside this region , see e.g. Figure 1.4.

The disadvantage of the *k-of-k* and the *k-of-w* signalling rules are that least k points must be evaluated to produce an OOC signal which leads to poor large shift detection. For this reason the improved *k-of-w* and the revised *k-of-w* signalling rules were developed.

A visual representation of runs-type signaling rules are provided in Figure 1.7 to 1.9.

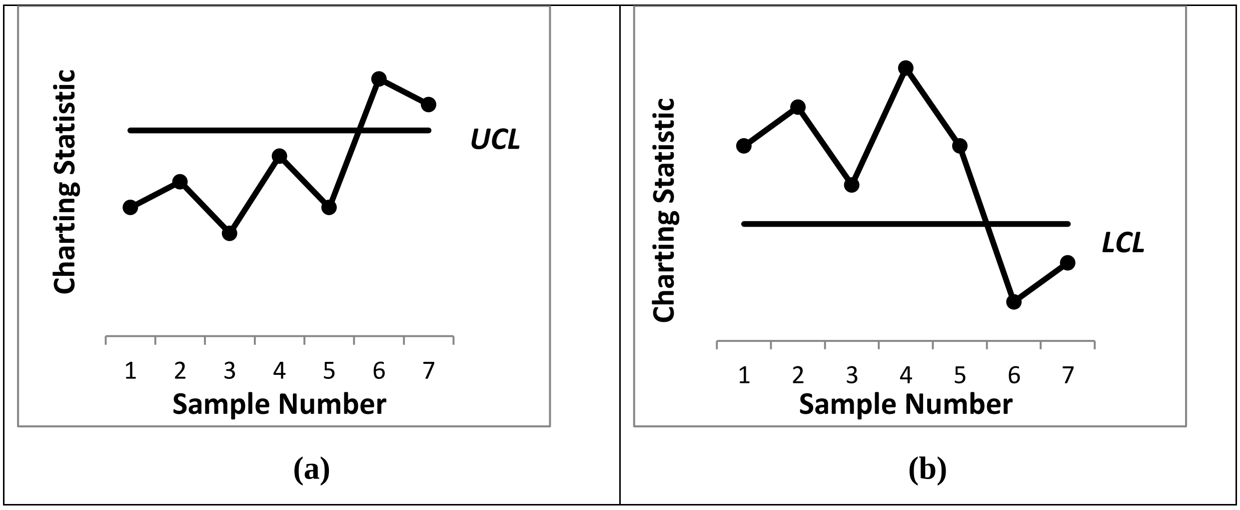

Figure 1.7: One-sided control charts supplemented with the *2-of-2* runs-type signalling rule.

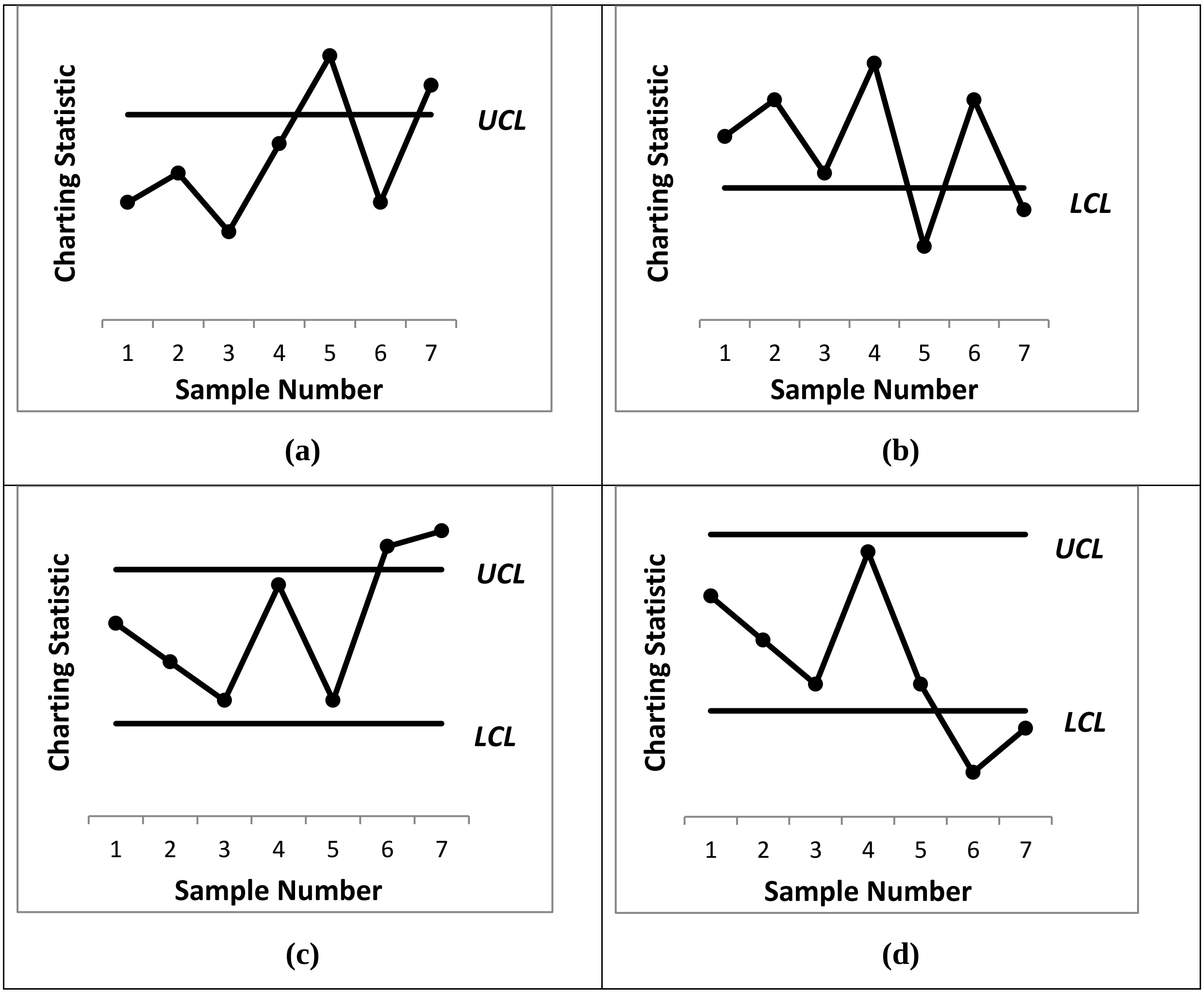

Figure 1.8: One and two-sided control charts supplemented with the *2-of-2* and *2-of-3* runs-type signalling rules.

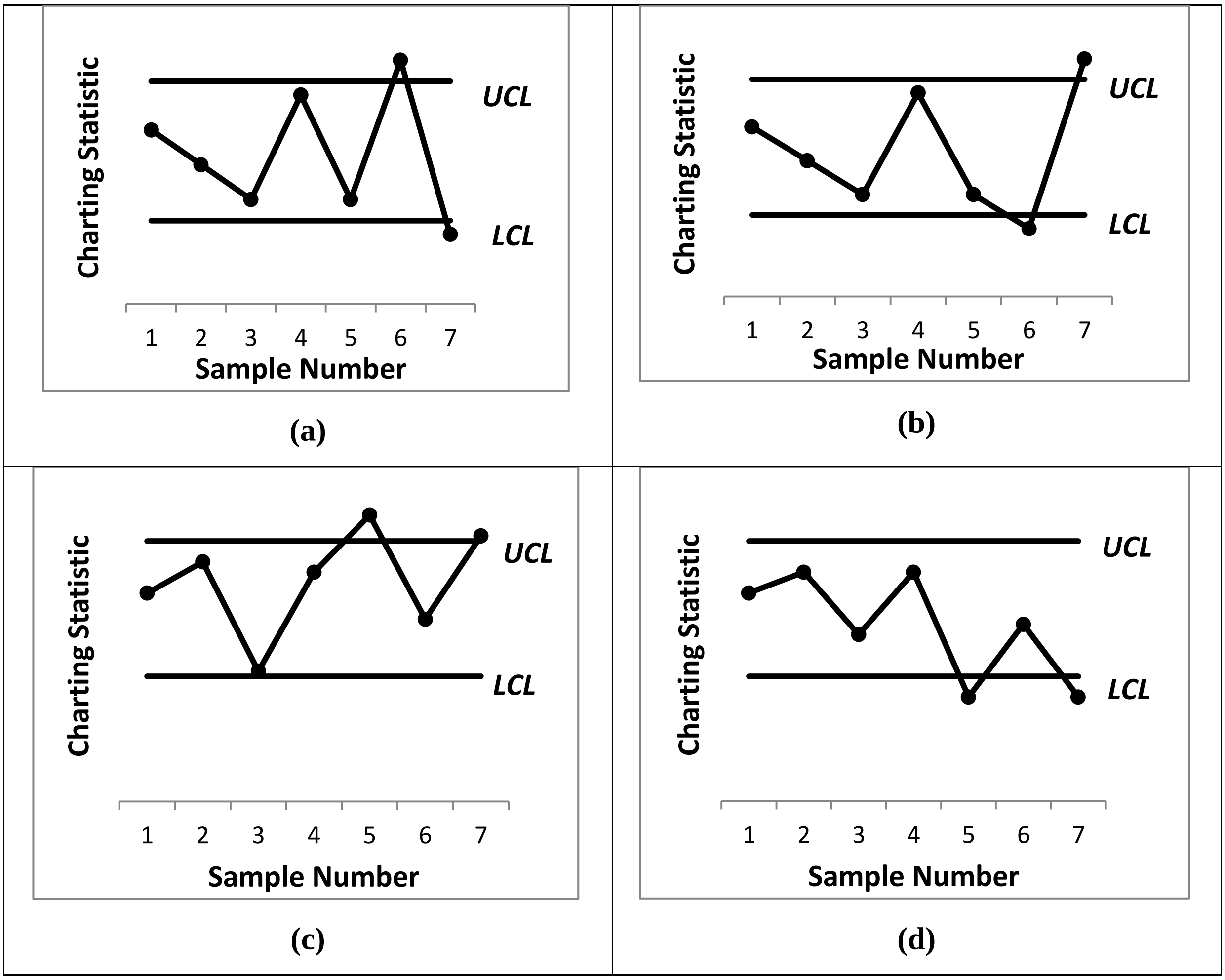

Figure 1.9: Two-sided control charts supplemented with the *2-of-2* and *2-of-3* runs-type signalling rules.

Figure 1.7 (a) and (b) refers to the one-sided *2-of-2* runs-type signalling rule. Figure 1.8 (a) and (b) refers to the upper-and lower one-side *2-of-3* rule, while (c) and (d) refers to the two-sided control chart monitoring an upward and downward shift when considering the *2-of-2* runs-type signalling rule. Figure 1.9 (a) and (b) also refer to the two-sided control chart incorporating the *2-of-2* runs-type signalling rule monitoring a swing in the process, thus a signal can be produced if the charting statistic plots above the UCL followed by another below the LCL or the charting statistic plotting below the LCL followed by another plotting above the UCL; this approach is referred to as the Derman and Ross (DR) approach further on in the study. Figure 1.9 (c) and (d) refers to the two-sided *2-of-3* runs-type signalling rule monitoring both an upward (c) or downward (d) shift. A signal is produced for all the runs-type signalling rules mentioned above at sample number seven.

Literature overview of runs-type signalling rules

Firstly we consider a broad overview of parametric control charts for monitoring the location of a process, followed by parametric control charts for monitoring the spread. The same order is also considered for nonparametric control charts. The diagram in Figure 1.10 provides an outline of the thesis that follows.

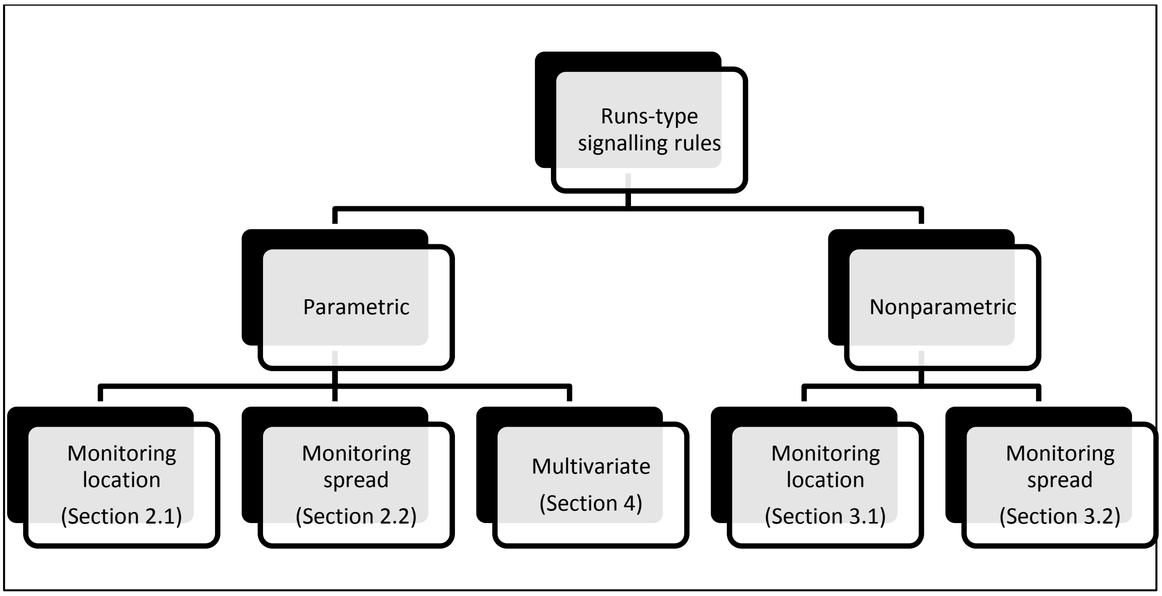

Figure 1.10: Classifications of the runs-type signalling rules control chart.

The articles discussed from Chapter 2-4 are listed below according to the description in Figure 1.10.

Parametric control charts with runs-type signalling rules for monitoring location

i. Statistical process control using Shewhart control charts with supplementary runs rules (Koutras et al. (2007)).
ii. Two alternatives to the Shewhart $\bar{X}$ control chart (Klein (2000a)).
iii. Two improved runs rules for the Shewhart $\bar{X}$ control chart (Khoo et al. (2006)).
iv. Design of runs rules schemes (Khoo (2003)).
v. The revised *m-of-k* runs rule (Antzoulakos and Rakitzis (2008a)).
vi. The revised *m-of-k* runs rule based on median run length (Low et al. (2012)).
vii. Two sets of runs rules for the chart (Acosta-Mejia (2007)).
viii. Modified r out of m control chart (Antzoulakos and Rakitzis (2008b)).

ix. Runs rules $\overline{X}$ charts when process parameters are unknown (Zhang et al. (2010).

Parametric control charts with runs-type signalling rules for monitoring spread

i. Modified S charts for controlling process variability (Klein (2000b)).
ii. The performance of control charts for monitoring process variation (Lowry et al. (1995)).
iii. Runs rules schemes for monitoring process variability (Antzoulakos and Rakitzis (2010)).
iv. ARL design of S charts with *k-of-k* rules (Acosta-Mejia et al. (2009)).
v. Modified R charts for improved performance (Acosta-Mejia et al. (2008)).
vi. Control charts with switching and sensitizing runs rules for monitoring process variation (Antzoulakos and Rakitzis (2014)).

Nonparametric control charts with runs-type signalling rules for monitoring location

i. A nonparametric Shewhart-type signed rank control chart based on runs (Chakraborti et al. (2007)).
ii. A distribution-free Shewhart quality control chart based on signed-ranks (Bakir (2004)).
iii. Nonparametric Shewhart-type sign control charts based on runs (Human et al. (2010)).
iv. A distribution free control chart based on order statistics (Balakrishnan et al. (2010)).
v. A phase II nonparametric control chart based on precedence statistics with runs-type signalling rules (Chakraborti et al. (2009)).

Nonparametric control charts with runs-type signalling rules for monitoring spread

i. Nonparametric control chart for controlling variability based on rank test (Das (2008b)).

Runs-type signalling rules based on multivariate sample statistics

i. A performance analysis of Hotellings χ^2 control chart with supplementary runs rules (Aparisi et al. (2004)).
ii. Improving the performance of the Chi-square control chart via runs-rules (Koutras et al. (2006)).

iii. Adaptive Hotelling T^2 control chart with run rules (Lee (2013)).
iv. Incorporating runs rules into Hotellings χ^2 control charts (Khoo et al. (2003)).
v. Powerfull rules for Hotellings χ^2 control chart (Khoo et al. (2005)).
vi. Alternative to the multivariate control chart for process dispersion (Khoo et al. (2004)).

1.7. Run-length distribution

In order to compare the performance of two control charts some quantifiable measure is needed. The run-length variable along with its associated characteristics fulfils this purpose.
"The number of rational subgroups to be collected or the number of charting statistics to be plotted on a control chart before the first OOC signal is observed is the run-length of a chart", see Human and Graham (2007). The run-length is a random variable, denoted by T, with a mean and variance. The average run-length (ARL) is the average number of points that should be plotted before an OOC condition is observed. Thus for the *k-of-w* runs-type signalling rule the run-length would be the number of charting statistics plotted until *k* out of *w* consecutive points plot between the UCL_1 and UCL_2 in the case of the upper control chart. It should be noted that a limitation of the *k-of-w* runs-type signalling rule is that a signal will only be produced if the last charting statistic plots in the range providing the signal, e.g. between the UCL_1 and UCL_2 for the upper control chart mentioned above; thus a situation can arise where k out of w charting statistics plot inside the signal range with the last charting statistic plotting IC, leading to no signal. Since the run-length distribution is significantly right-skewed other measures are also used to evaluate the performance of the chart, namely the standard deviation of the run-length ($SDRL$) and the median run-length (MRL).

The run-length distribution can be evaluated using the following four methods:

i. An exact approach (for Shewhart-type, EWMA and CUSUM charts).
ii. A Markov chain approach.
iii. The integral equation approach.
iv. Computer simulations (Monte Carlo simulations).

When considering the Markov chain approach for a Shewhart-type control chart supplemented with runs-type signalling rules the results are exact; thus, the Markov chain approach is an exact approach. The computer simulation approach consists of generating a random sample and then testing whether a signal is produced. This process is repeated a large number of times, e.g. $10\,000$ samples are generated. The advantage of this approach is that no matter how complicated the run-length distribution is, computer simulations can almost always be used with relative ease to calculate the run-length distribution and its associated characteristics fairly accurately, provided the simulation size is large enough. For this study the Markov chain approach is followed.

When using the Markov chain approach (Fu et al. (2003)) the run-length variable T is viewed as the waiting time until the first signal. Define a Markov chain $\{Y_t, t > 1\}$, completely characterized by a state space S and a transition probability matrix M. The state space consists of non-absorbing and absorbing states with the latter consisting of states that are entered if an OOC signal is observed. If no signal is observed the process is in a transient state. The transition probability matrix is written in a partitioned form,

$$\boldsymbol{M}_{(\boldsymbol{v}+\mathbf{1})\times(\boldsymbol{v}+\mathbf{1})} = \begin{pmatrix} \boldsymbol{Q}_{\boldsymbol{v}\times\boldsymbol{v}} & | & \underline{\boldsymbol{p}}_{\boldsymbol{v}\times\mathbf{1}} \\ - & - & - \\ \underline{\mathbf{0}}'_{1\times v} & | & 1_{1\times 1} \end{pmatrix}$$

where the submatrix $\boldsymbol{Q}_{\boldsymbol{v}\times\boldsymbol{v}}$ (containing v rows and v columns), called the essential transition probability matrix (TPM), contains all the probabilities of going from one transient state to another; the column vector $\underline{\boldsymbol{p}}_{\boldsymbol{v}\times\mathbf{1}}$ contains the probabilities of going from each transient to an absorbing state; $\underline{\mathbf{0}}'_{\mathbf{1}\times\boldsymbol{v}}$ is a row vector of zeros containing all the probabilities of going from an absorbing (A) to a non-absorbing (NA) state and the scalar value one is the probability of going from an absorbing state to the absorbing state ($\underline{\mathbf{0}}'_{\mathbf{1}\times\boldsymbol{v}}$ is the column vector $\underline{\mathbf{0}}_{\boldsymbol{v}\times\mathbf{1}}$ transposed).

The run-length random variable T of the Shewhart-type control chart can be considered the waiting time for the Markov chain to enter the absorbing state for the first time. Fu and Lou (2003) derived the various characteristics of T as follows.

$$P(T = t) = \underline{\boldsymbol{\omega}}\boldsymbol{Q}^{t-1}(\boldsymbol{I} - \boldsymbol{Q})\underline{\mathbf{1}} \qquad \text{for} \qquad t = 1,2,3, \ldots \tag{1.3}$$

$$E(T) = \underline{\boldsymbol{\omega}}(\mathbf{I} - \boldsymbol{Q})^{-1}\underline{\mathbf{1}} \tag{1.4}$$

$$SDRL(T) = \sqrt{\underline{\boldsymbol{\omega}}(\boldsymbol{I}+\boldsymbol{Q})(\boldsymbol{I}-\boldsymbol{Q})^{-2}\underline{\mathbf{1}} - (E(T))^2} \tag{1.5}$$

$$P(T \leq t) = 1 - \underline{\boldsymbol{\omega}}\boldsymbol{Q}^t\underline{\mathbf{1}} \qquad \text{for} \qquad t = 1,2,3,\dots \tag{1.6}$$

Note that $\boldsymbol{I} = \boldsymbol{I}_{\boldsymbol{v}\times\boldsymbol{v}}$ is the identity matrix, $\boldsymbol{Q} = \boldsymbol{Q}_{\boldsymbol{v}\times\boldsymbol{v}}$ is the essential TPM, $\underline{\mathbf{1}} = \underline{\mathbf{1}}_{\boldsymbol{v}\times\mathbf{1}}$ is a column vector with all elements equal to one and $\underline{\boldsymbol{\omega}} = \underline{\boldsymbol{\omega}}_{\mathbf{1}\times\boldsymbol{v}}$ is a row vector called the initial probability vector which contains the probabilities that the Markov chain starts in a given state. The vector $\underline{\boldsymbol{\omega}} = (\omega_{-s},\dots,\omega_s)$ is typically chosen with $2s+1 = v$ such that $\sum_{i=-s}^{s}\omega_t = 1$; thus the process starts in a given state with probability one. Assume $\omega_0 = 1$ and $\omega_t = 0$ for all $t \neq 0$ with $t \in [-s, s]$. This implies that the Markov chain starts in state zero with probability one; with the other states consisting of values in the following range, $(-s, s)$ excluding 0. This implies that the process is IC when we start monitoring the process.

To derive the expected waiting time until the first signal is observed also referred to as the ARL in Equation (1.4) the probability generating function is differentiated. For the complete derivation the reader is referred to "Distribution Theory in runs and Patterns and its applications" (Fu et al. 2003) or the appendix.

In order to calculate the ARL the absorbing states and non-absorbing states have to be identified, the control limits and associated probabilities calculated, the essential TPM derived and Equation (1.4) applied.

1.8. Nonparametric control charts

Nonparametric control charts (i.e. distribution free charts) have been developed for situations where normality or any other specific parametric model assumption about the underlying distribution is not met, see Definition 1.4.

Definition 1.4

Distribution-free or nonparametric control chart

If the IC run-length distribution is the same for every continuous distribution, then the chart is called distribution-free; note that for the Wilcoxon Signed Rank test statistic the additional assumption of symmetry is needed.

Chakraborti et al. (2001) summarized the advantages and disadvantages of nonparametric control charts as follows:

Advantages:

i. They are easy to implement,
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. Higher 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.

Disadvantages:

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.

Chapter 2

Parametric charts: Runs-type signalling rules

2.1 Runs-type signalling rules for monitoring location

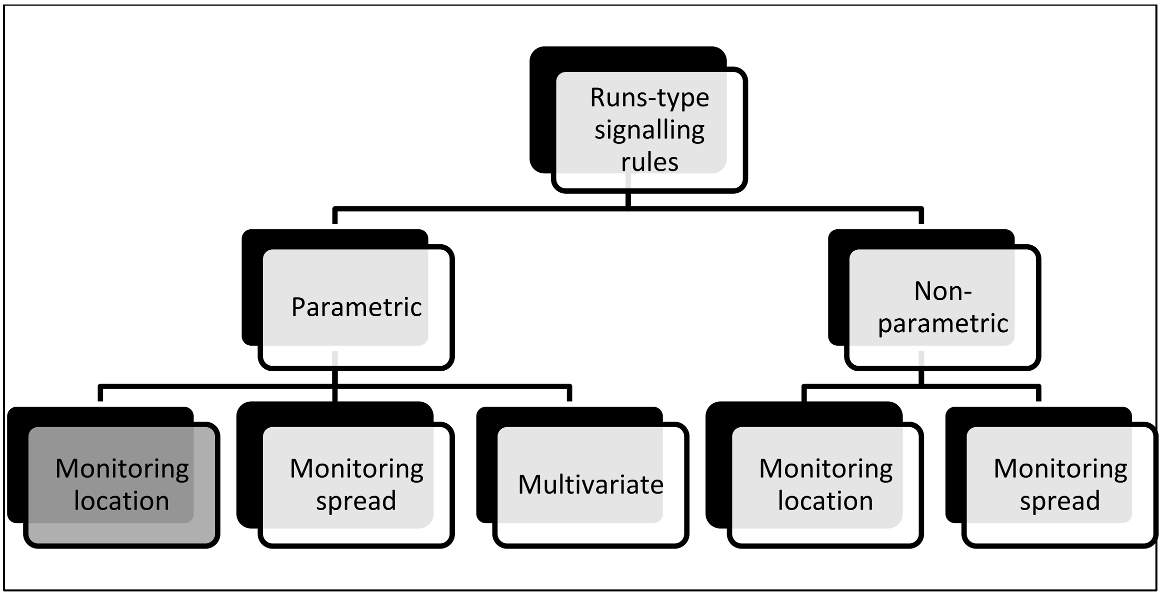

Figure 2.1: Classifications of the runs-type signalling rules control chart.

In general, control charts supplemented with runs-type signalling rules could be classified as being parametric or nonparametric. For a parametric control chart the distribution is known but the underlying parameters could be known (case K) or unknown (case U) as mentioned in the previous chapter. In the nonparametric case the underlying distribution is unknown but assumed to be continuous. Runs-type signalling rules in the parametric environment are classified as charts for monitoring location, spread or multivariate characteristics. On the other hand runs-type signalling rules in the nonparametric environment are classified as charts for monitoring location or spread.

This section provides a detailed review of the articles based on runs-type signalling rules detecting a shift in location using a Shewhart-type control chart.

The article by Koutras et al. (2007) provides an introduction to runs-type signalling rules for monitoring a shift in the location.

2.1.1 Statistical process control using Shewhart control charts with supplementary runs rules (Koutras et al. (2007))

Koutras et al. (2007) provided a review on Shewhart control charts supplemented with runs-type signalling rules, deriving the run-length distribution and calculating the ARL. (Koutras et al. (2007) suggests using runs-type signalling rules for Phase I analysis and a EWMA/CUSUM approach for Phase II analysis to monitor small shifts). During a Phase I analysis the purpose is to establish whether the sampled data comes from an IC process or not, while Phase II analysis aims to monitor the process for a shift due to a change in one of the parameters of the underlying distribution or for non-random patterns. When designing a control chart there is a trade off between the α and β risk (i.e. the type and type II error). In practice the ARL_0 is fixed and the control limits are adjusted to minimize the ARL_ξ.

Koutras et al. (2007) discussed the run-length T in detail. In general it is desirable to have a large IC ARL denoted by ARL_0 and a small OOC ARL denoted by ARL_ξ,where ξ refers to the process shift. The perfect control chart would be a chart that gives zero false alarms, i.e. having an in-control average run-length of infinity, and to immediately signal when the process moves out-of-control, i.e. the out-of-control average run-length equals one with probability 1. The scenarios above are not feasible in practice and some trade-offs have to be made. In practice control charts are typically designed for an ARL_0 of 370 or 500 with corresponding false alarm rates denoted by FAR of 0.0027 or 0.002. The distribution of T is geometric with $E(T) = ARL_0 = 1/p_{in}$ since the observations are distributed iid $N(0,1)$ and the signalling events independent; p_{in} denotes the probability of the charting statistic plotting outside the control limits given the process is IC. In order to calculate the ARL_0 and ARL_ξ the probability of the charting statistic C plotting outside the control limits needs to be calculated. If the random variable $C = f(X)$ follows a normal distribution with parameters μ and σ^2 and the cdf of C equals $F(x)$, then $F(x) \neq K(x)$ if the process is OOC, where $K(x)$ denotes the cdf of C if a shift in the process mean occured. On the contrary if the process is IC $F(x) = K(x)$. Given the status of the process (IC/OOC) the probability of plotting outside the control limits (UCL/LCL) is denoted by p_{in} or p_{out}; where p_{in} refers to the probability of plotting outside the control limits given no shift occurred and p_{out} to the probability of plotting outside the control limits given a shift, denoted by ξ occurred. It should be noted that no signal is produced if the charting statistic plots on the control limits; however, for most of the other control charts discussed further on in the thesis a signal is produced if the charting statistic plots on or above/below either the UCL/LCL.

$$\begin{aligned} p_{in} &= 1 - P(LCL \leq C \leq UCL | C \sim N(\mu, \sigma^2)) \\ &= 1 - \phi(L - \xi) + \phi(-L - \xi). \\ &= 1 - \phi(L) + \phi(-L) \text{ where } \xi = 0\,. \end{aligned}$$

$$\begin{aligned} p_{out} &= 1 - P(LCL \leq C \leq UCL | C \sim N(\mu + \xi\sigma, \sigma^2)) \\ &= 1 - \phi(L - \xi) + \phi(-L - \xi). \end{aligned}$$

where $UCL = L$ and $LCL = -L$, $\phi(.)$ denotes the cdf of a standard normal variable and ξ the shift in the sample mean. Since the distribution of the run length variable T is geometric the $ARL_{\xi} = 1/p_{out}$. Note that the parameters μ and σ mentioned above are known.

In order to calculate the ARL Koutras et al. (2007) first calculated the probability generating function and derived the ARL from this equation. Alternatively the Markov chain approach as discussed by Fu and Lou (2003) could be applied, see Equation (1.4). The aim of this article is to establish the probabilities associated with the different process states and derive the essential TPM. An OOC signal is observed if a single point plots in the two regions $(-\infty, -3)$ or $(3, \infty)$, or if two consecutive points plot between $(-3, -2)$ or $(2,3)$. The essential TPM is provided in Table 2.1 for the combined *1-of-1* and *2-of-2* runs-type signalling rule and Equation (1.4) can now be applied to calculate the ARL.

Z_1 is defined as the state entered if the charting statistic plots in the region defined as $(2,3)$ while Z_2 is the state entered if the charting statistic plots inside $(-3, -2)$ and Z_0 the state entered if the charting statistic plots inside $[-2,2]$, see e.g. Figure 2.2. The probabilities provided in Table 2.1 correspond to the regions in Figure 2.2, e.g. p_0 denotes the probability of plotting in Region 0.

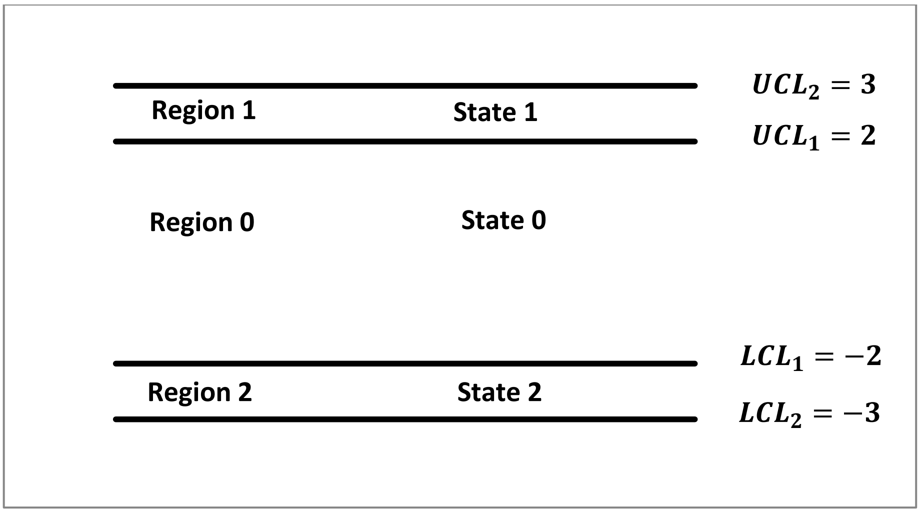

Figure 2.2: Charting regions and state space for the *1-of-1* and *2-of-2* runs-type signalling rules.

Table 2.1 Essential TPM combining the *1-of-1* and *2-of-2* runs-type signalling rules.

State	Z_0	Z_1	Z_2
Z_0	p_0	p_1	p_2
Z_1	p_0	0	p_2
Z_2	p_0	p_1	0

Note that Table 2.1 refers to the essential TPM, thus the probability of moving from a single state to all possible other states won't always sum to 1 since the absorbent states are excluded from the essential TPM. If it is assumed that the process starts in state Z_1 then either state Z_0, Z_1, or Z_2 could be entered when considering the combined *1-of-1* and *2-of-2* runs-type signalling rules; the state entered if the process plots below the LCL_2 or above the UCL_2 is not considered here since the process immediately signals if the charting statistics plots in either of those regions which leads to an absorbent state not considered in the essential TPM. Similarly the state entered if two consecutive charting statistics plot in Region 1 or 2 is excluded from the essential TPM since a signal is produced and the absorbent state entered. The probability of the process moving from state Z_1 to Z_0 is p_0, from state Z_1 to Z_1 0 (since absorbent state is entered) and from state Z_1 to Z_2 p_2. The probability of moving from state Z_1 to Z_1 in the essential TPM is 0 since an absorbent state is entered where the probability $p_1$would then be allocated to the TPM (Section 1.7), where the rows would some to one.

When considering the essential TPM the rows would thus in some cases not sum to one since the absorbent states are excluded from the matrix.
Care should be taken when combining different runs-type signalling rules since there exists a possibility of increased false alarms. This increase (false alarms or Type-I error) was noted by Montgomery (2013). The following article explores this concept further.

2.1.2 Two alternatives to the Shewhart $\bar{X}$ control chart (Klein (2000a))

Klein (2000a) proposed the *2-of-2* and *2-of-3* runs-type signalling rules since an increase was noted in the *FAR* of the rules developed by Nelson (1984) and Montgomery (2013). Derman and Ross (1997) developed two schemes considering the last two and three charting statistics, similar to Klein but his signalling criteria differs. The *2-of-2* and *2-of-*3 rules are members of the k-*of*-w runs-type signalling rule class where a signal is produced if k of the last w charting statistics fall either on or above, or on or below the UCL or LCL respectively; thus if k out of w charting statistics plot in Region U or L in Figure 2.3. The rules proposed by Derman and Ross differ from Klein's since a signal is produced if a charting statistic plots above the UCL followed by a charting statistic plotting below the LCL or a charting statistic plotting below LCL followed by a charting statistic plotting above the UCL for his *2-of-2* rule. Klein's approach would only produce a signal if both charting statistics plot either above the UCL or both below the LCL. The *2-of-3* rule similarly only considers the last 3 consecutive points. For clarification purposes, if the charting statistic plots in Region U, State U is entered and likewise if the charting statistic plots in Region L, State L is entered.

The distributional assumption is that the random variables are iid normally distributed. The control limits were set up for an ARL_0 of 370 and the reader is referred to Chapter 5 for a detailed discussion on establishing the limits.

The state space is defined as the region above the UCL, below the LCL and a region between the UCL and LCL, as depicted in Figure 2.3. The associated transition probabilities are p_u, p_l and p as shown in Figure 2.3, where $p_u = p(X > UCL) = 1 - \phi(UCL)$, $p_l = p(X \leq LCL) = \phi(LCL)$ and $p = 1 - p_u - p_l$.

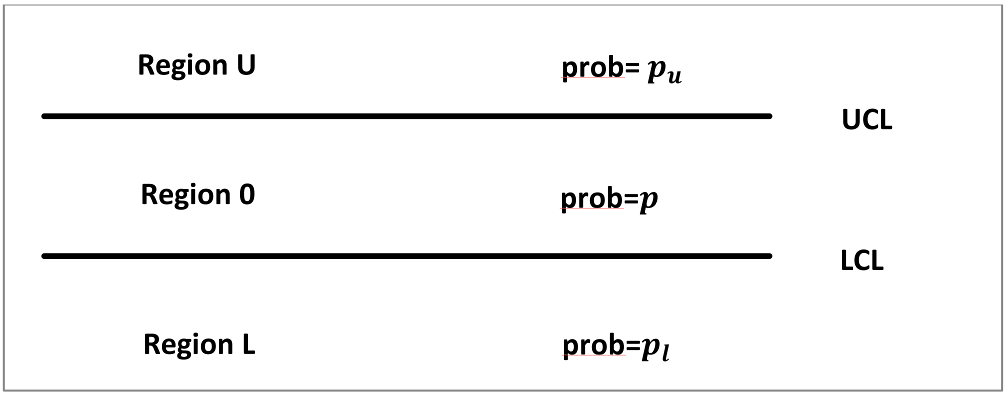

Figure 2.3: Charting regions as probabilities for the *2-of-2* and *2-of-3* runs-type signalling rules.

In order to calculate the ARL the Markov chain approach could be used by setting up the essential TPM and solving Equation (1.4). The essential TPM's for the *2-of-2* and *2-of-3* runs-type signalling rules are provided in Tables 2.2 and 2.3 respectively. State S_U refers to the charting statistic plotting above the UCL while State S_{0U} refers to the first charting statistic plotting between the control limits in Region 0 and the second charting statistic plotting above the UCL in Region U. It should be noted that the first and second charting statistics refer to two consecutive charting statistics.

Table 2.2: Essential TPM: *2-of-2* runs-type signalling rule.

State	S_0	S_U	S_L
S_0	p	p_u	p_l
S_U	p	0	p_l
S_L	p	p_u	0

Table 2.3: Essential TPM: *2-of-3* runs-type signalling rule.

States	S_{00}	S_{0U}	S_{0L}	S_{UL}	S_{U0}	S_{L0}	S_{LU}
S_{00}	p	p_u	p_l	0	0	0	0
S_{0U}	0	0	0	p_l	p	0	0
S_{0L}	0	0	0	0	0	p	p_u
S_{UL}	0	0	0	0	0	p	0
S_{U0}	p	0	p_l	0	0	0	0
S_{L0}	p	p_u	0	0	0	0	0
S_{LU}	0	0	0	0	p	0	0

Klein (2000a) uses first passage time equations to calculate the ARL_0; this approach is similar to the Markov chain approach. In order to calculate the ARL_0 the first passage time from the starting state to the absorbing state needs to be established; this is also referred to as the expected waiting time until the first OOC signal. For the *2-of-2* signalling rules the first passage time equations are as follows:

$$M_{14} = 1 + pM_{14} + p_u M_{24} + p_l M_{34.}$$
$$M_{24} = 1 + pM_{14} + p_l M_{34.}$$
$$M_{34} = 1 + pM_{14} + p_u M_{34}.$$

Where M_{24} refers to the first passage time from State 2 to State 4.

Hurwitz and Marthur (1992) developed a solution to the system. By applying symmetric control limits the equation reduces to: $M_{14} = (1 + p^*)/2(p^{*2}) = ARL$. Since using symmetric control limits, $p_l = p_u = p^*$. For an ARL_ξ the value of p_u and p_l have to be recalculated. If a shift of b units occur $p_l = p(X + b < LCL)$ and $p_u = p(X + b > UCL)$. A similar approach could be followed for the *2-of-3* runs-type signalling rules.

Klein compared the performance of the *2-of-2* and *2-of-3* rule to the Shewhart $\bar{X}$ chart, Derman and Ross's *2-of-2* and *2-of-3* as well as the EWMA (Exponentially Weighted Moving Average) chart. He found that the proposed *2-of-2* and *2-of-3* control charts outperform all of the above mentioned except the EWMA($L = 2.9, \lambda = 0.25$), but on the flipside is simpler to implement than the EWMA; L and λ are design parameters for the EWMA chart.

2.1.3 Two improved runs rules for the Shewhart $\bar{X}$ control chart (Khoo et al. (2006))

To overcome the problem of an increased type-I error Klein (2000a) developed the *2-of-2* and the *2-of-3* runs-type signalling rules. These rules provided better performance for small to moderate changes in the process mean compared to the standard Shewhart $\bar{\text{X}}$ control chart but inferior performance was noted in detecting large shifts. Shifts up to 0.8σ units from the mean are considered small, a shift between 0.8σ and 2.8σ as moderate and shifts larger than 2.8σ as large.

The aim of the improved *2-of-2* and *2-of-3* rules developed by Khoo et al. (2006) is to provide improved small and large shift detection capabilities by considering both at the same time. The improved *2-of-2* Shewhart-type chart signals if a charting statistic plots on or above the UCL_2 or on or below the LCL_2 , or if two consecutive charting statistics plot on or beyond the limits UCL_1 and LCL_1 respectively. The improved *2-of-3* rule signals if a charting statistic plots on or above the UCL_2, or on or below the LCL_2, or if two out of three consecutive charting statistics plot on or beyond the limits UCL_1 and LCL_1 respectively. If a charting statistic plots on the UCL_1 (LCL_1) it is considered as plotting between (UCL_1 /LCL_1) and (UCL_2 /LCL_2). It should be noted that the inner control limits (UCL_1 /LCL_1) for these two rules differ but the outer control limits (UCL_2 /LCL_2) are the same.

The distributional assumptions are that observations are distributed iid $N(0,1)$ and the control limits are designed for an ARL_0 of 370. In order to establish the control limits the outer control limits are fixed (typically wider than $\pm 3\sigma$), while the inner control limits are determined for a desired ARL_0.

The state space for the *2-of-2* rule is set up as follows:
State 1 (S_1): no point beyond the control limits,
State 2 (S_2): a point between the upper outer and inner control limits (UCL_1 and UCL_2),
State 3 (S_3): a point between the lower outer and inner control limits (LCL_1 and LCL_2),
State 4 (S_4): the absorbing state, where a point is either beyond the outer control limits or if two consecutive points fall between the upper outer and inner control limits or between the lower outer and inner control limits.

A graphical representation of the associated regions and probabilities are provided in Figure 2.4. The reader is referred to the Khoo et al. (2006) for the definition of the state space of the *2-of-3* rule.

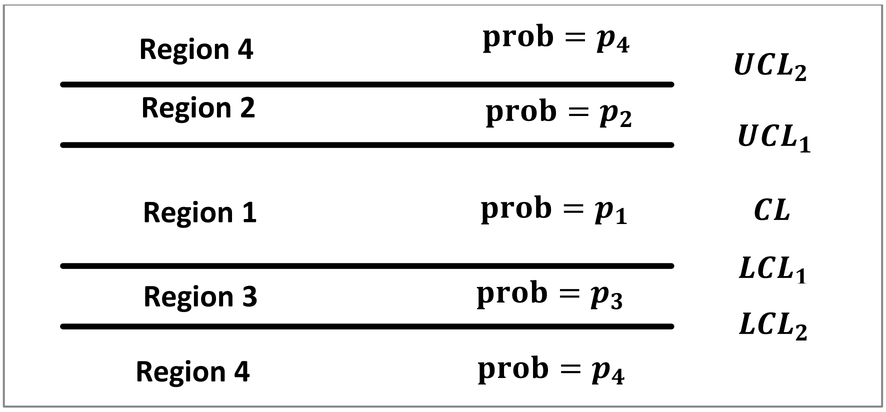

Figure 2.4: Charting regions and probabilities for the improved *2-of-2* and *2-of-3* runs-type signalling rules.

In order to calculate the ARL the Markov chain approach can be implemented, whereby the essential TPM is set up and Equation (1.4) applied. The essential TPM for the improved *2-of-2* and *2-of-3* runs-type signalling rules are provided in Tables 2.4 and 2.5 where S_1 refers to State 1; thus if the charting statistic plots in Region 1 as shown in Figure 2.4 State 1 is entered. S_{11} refers to the first charting statistic plotting in Region 1, followed by the second charting statistic also plotting in Region 1.

Table 2.4: Essential TPM for the improved *2-of-2* runs-type signalling rule.

State	S_1	S_2	S_3
S_1	p_1	p_2	p_3
S_2	p_1	0	p_3
S_3	p_1	p_2	0

Table 2.5 Essential TPM for the improved *2-of-3* runs-type signalling rule.

State	S_{11}	S_{12}	S_{13}	S_{23}	S_{21}	S_{31}	S_{32}
S_{11}	p_1	p_2	p_3	0	0	0	0
S_{12}	0	0	0	p_3	p_1	0	0
S_{13}	0	0	0	0	0	p_1	p_2
S_{23}	0	0	0	0	0	p_1	0
S_{21}	p_1	0	p_3	0	0	0	0
S_{31}	p_1	p_2	0	0	0	0	0
S_{32}	0	0	0	0	p_1	0	0

A simulation study of 5000 trials was conducted by Khoo et al. (2006) and it was found that the ARL performance of the improved *2-of-2* and *2-of-3* rules are similar to Klein's (2000a) results, but has an improved capability to detect large shifts. An increase in the type I error is generally noted if two rules are used simultaneously but since the control limits are chosen for a specified ARL_0 (type I error) there is no such concern regarding these two rules.

2.1.4 Design of runs rules schemes (Khoo (2003))

The *2-of-3* runs-type signalling rule proposed by Klein (2000a) is complicated to set up since it involves a linear system of seven equations. Khoo (2003) proposed a different approach and considers the *2-of-3*, *2-of-4*, *3-of-3* and *3-of-4* runs-type signalling rules. From a practical point of view Khoo (2003) proposed having a standard lookup table derived i.e. by simulation matching combinations of $p_u$or p_l and ARL_0. Note that p_u refers to the probability of the charting statistic plotting above the UCL and p_l the probability of the charting statistic plotting below the LCL, see Figure 2.3. In practice this could be benefitial since the linear system of seven equations doesn't have to be solved each time a different ARL_0 is considered. Khoo (2003) proposed the *3-of-4* runs-type signalling rule for small shifts (up to 1.8σ), the *2-of-2* runs-type signalling rule for larger shifts (2σ) and for vary large shifts (>3σ) the Shewhart $\overline{\mathrm{X}}$ control chart. The steps to set up the control chart are as follows:

i. Decide on the magnitude of shift to monitor.
ii. Find p_u and p_l, see Figure 2.3 and Klein (2000a).
iii. Determine the control limits from standard normal tables, or a software package, e.g. Excel.

The distributional assumption is that the observations are distributed iid $N(0,1)$. The regions considered in setting up the control limits are similar to Figure 2.3. To establish the ARL the Markov chain approach can be followed where the essential TPM is set up and the ARL's calculated using Equation (1.4). The essential TPM and associated probabilities for the *2-of-3* runs-type signalling rule is provided in Table 2.3 (Section 2.1.2). The essential TPM for the *3-of-3* runs-type signalling rule is provided in Table 2.6. State S_{0L} refers to the first charting statistic plotting IC in Region O, followed by the second charting statistic plotting in Region L (below LCL). The reader is referred to Figure 2.3 for a graphical illustration of the regions.

Table 2.6: Essential TPM for the *3-of-3* runs-type signalling rule.

States	S_{00}	S_{0U}	S_{0L}	S_{UL}	S_{LU}	S_{U0}	S_{L0}	S_{UU}	S_{LL}
S_{00}	p	p_u	p_l	0	0	0	0	0	0
S_{0U}	0	0	0	p_l	0	p	0	p_u	0
S_{0L}	0	0	0	0	p_u	0	p	0	p_l
S_{UL}	0	0	0	0	p_u	0	p	0	p_l
S_{LU}	0	0	0	p_l	0	p	0	p_u	0
S_{U0}	p	p_u	p_l	0	0	0	0	0	0
S_{L0}	p	p_u	p_l	0	0	0	0	0	0
S_{UU}	0	0	0	p_l	0	p	0	0	0
S_{LL}	0	0	0	0	p_u	0	p	0	0

The transient states for the *3-of-4* runs-type signalling rule are as follows: OOO, OOU, $OUO, UOO, OOL, OLO, LOO, OUU, UOU, UUO, OLL, LOL, LLO, OUL, UOL, LOU,$ OLU, $ULO, LUO, UUL, ULU, LUU, LLU, LUL, ULL$. The essential TPM could be set up incorporating the transient states and the ARL calculated by applying Equation (1.4).

The performance of the above mentioned runs-type signalling rules were measured against the *2-of-2* runs-type signalling rule and Shewhart $\overline{\mathrm{X}}$ control chart. In general, the *3-of-4* runs-type signalling rule had the best ARL_{ξ} performance for a shift in the mean of up to 1.8σ, for a shift of $(2-2.6)\sigma$ the *2-of-3* and *2-of-2* runs-type signalling rules performed best and for shifts larger than 3σ the Shewhart $\overline{\mathrm{X}}$ chart outperformed the rest.

2.1.5 The revised *m-of-k* runs rule (Antzoulakos and Rakitzis (2008a))

The revised *m-of-k* (similar to the *k-of-w*) runs-type signalling rules were developed as an alternative to the improved runs-type signalling rule of Khoo and Ariffin (2006). The improved *m-of-k* uses wider inner control limits than the revised *m-of-k* if the same outer control limit (K) is considered, thus decreasing the amount of false alarms. Another difference is that for the improved chart a signal can be produced if the charting statistic plots on the opposite side of the CL in the sequence of charting statistics being monitored, while the rest of the charting statistics plot between either the two upper control limits or the two lower control limits. For the revised *m-of-k* runs-type signalling rule a signal is only produced if the charting statistics being monitored plots on one side of the CL; for the improved chart the charting statistics are allowed to be on either sides of the CL.

A signal is observed in the following two instances:

i. a charting statistic plots on or above/below the UCL_2/LCL_2.

ii. *m-of-k* (*k-of-w* previously) charting statistics plot between $UCL_1(LCL_1)$ and the $UCL_2(LCL_2)$ in a cluster of points between the CL and the $UCL_2(LCL_2)$; if the charting statistic plots on either the UCL_1 or the LCL_1 it is considered to plot in the range between the $UCL_1(LCL_1)$ and the $UCL_2(LCL_2)$.

The distributional assumption is that observations are distributed iid $N(0,1)$ and the control limits are determined for a specified ARL_0. The value of the outer control limit constant (K) is fixed and the inner control limit constant (d) is calculated. The reader is referred to Chapter 5 for a detailed discussion on establishing the control limits. The regions and associated probabilities for the revised *m-of-k* control chart supplemented with runs-type signalling rules are provided in Figure 2.5. Symmetric control limits are used with $UCL_2 = K$, $LCL_2 = -K$ and $UCL_1 = d$, $LCL_1 = -d$. The probabilities are then derived as follows:

$p_1 = \phi(K - \xi) - \phi(d - \xi).$

$p_2 = \phi(d - \xi) + \phi(\xi) - 1.$

$p_3 = \phi(d + \xi) - \phi(\xi).$

$p_4 = \phi(K + \xi) - \phi(d + \xi).$

$p_5 = 1 - p_1 - p_2 - p_3 - p_4.$

with ξ the process shift and $\phi(.)$ the cdf of a standard normal variable.

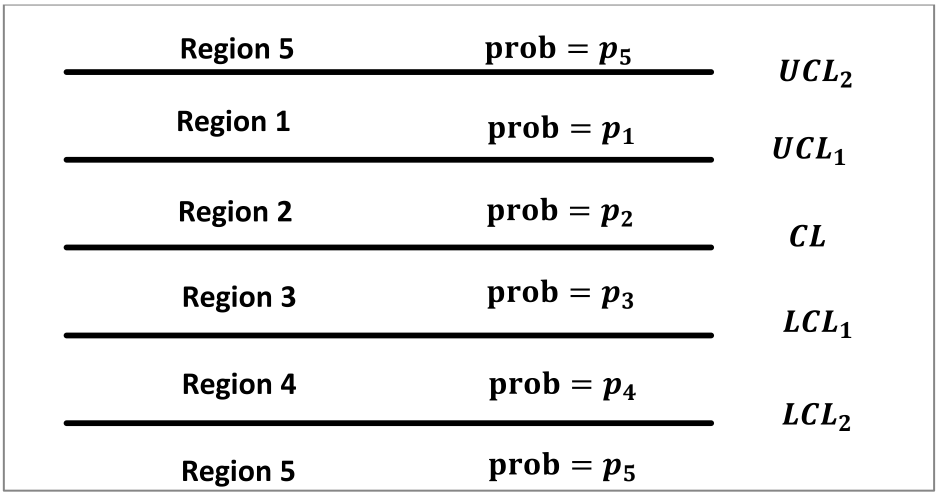

Figure 2.5: Charting regions and probabilities for the revised *m-of-k* runs-type signalling rule.

For the revised *2-of-3* runs-type signalling rule a signal is observed if the following events occur: $E = \{11,121,44,434,5\}$ where 121 denotes the first charting statistic plotting in Region 1 (State1), the second in Region 2 followed by the third plotting in Region 1 again. The essential TPM for the revised *2-of-3* runs-type signalling rule is provided in Table 2.7 and the ARL could be calculated by applying Equation (1.4). A more detailed breakdown of the signalling events for the *2-of-3* runs-type signalling rule would be $\{211,11,121,344,44,434,5\}$; note that Antzoulako and Rakitzis (2008a) groups 211 and 11 together as well as 344 and 44.

Table 2.7: Essential TPM for the revised *2-of-3* rule.

States	S_2 or S_3	S_1	S_{12}	S_4	S_{43}
S_2 or S_3	$p_2 + p_3$	p_1	0	p_4	0
S_1	p_3	0	p_2	p_4	0
S_{12}	$p_2 + p_3$	0	0	p_4	0
S_4	p_2	p_1	0	0	p_3
S_{43}	$p_2 + p_3$	p_1	0	0	0

In conclusion, the revised *4-of-5* runs-type signalling rule is recommended for shifts up to 1.6σ units ($K = 3.8$, $d = 0.96$). For shifts between 1.6σ and 3σ the revised *2-of-3* runs-type signalling rule is preferred ($K = 3.4$, $d = 1.926$). Both these rules give slightly better ARL performance than the improved *2-of-3* and *4-of-5* runs-type signalling rules.

2.1.6 The revised *m-of-k* runs rule based on median run length (Low et al. (2012))

The revised runs-type signalling rule based on the median run-length was developed since the shape of the run-length distribution changes with the magnitude of the shift in the process mean ranging from highly skewed when the process is in control to almost symmetric when a large shift occurs. The median run-length is considered as a measure of performance compared to the average run length for the revised *m-of-k* runs-type signalling rule.

Signal detection, setting up the control limits and the Markov chain approach is similar to the discussion regarding the revised *m-of-k* runs-type signalling rule (Antzoulakos and Rakitzis (2008a)). The control limits are determined for a specified ARL_0 or MRL_0. For the calculation of the charting constants e.g. K and d the reader is referred to Anzoulakos and Rakitzis (2008a).

The Markov chain approach for the *3-of-4* runs-type signalling rule will be considered next. The reader is referred to Figure 2.5 for the regions and associated probabilities used further in the discussion. Deriving the probabilities are exactly the same as in Section 2.15. For the *3-of-4* runs-type signalling rule an OOC signal is observed for the following compound event $E = \{111,1211,1121,444,4344,4434,5\}$ where 111 denotes three charting statistics in a row plotting in Region 1.

In order to calculate the ARL Equation (1.4) could again be applied. Brooke and Evans (1972) derived the cdf of the run-length for the *3-of-4* runs-type signalling rule as $P(T \leq t) = \underline{e}(I - Q^t)1'$. Note that $I = I_{\nu\times\nu}$ is the identity matrix, $Q = Q_{\nu\times\nu}$ is the essential TPM, $\underline{1} = \underline{1}_{\nu\times 1}$ is a column vector with all elements equal to one and $\underline{e} = \underline{e}_{1\times\nu}$ is a row vector called the initial probability vector which contains the probabilities that the Markov chain starts in a given state. The 100γth percentile of the run-length is determined such that: $P\left(T \leq t_\gamma - 1\right) \leq \gamma$ and $P\left(T \leq t_\gamma\right) > \gamma$, see Gan (1993); where the 100γth percentile of the run-length distribution corresponding to a standardized mean shift of ξ equal to t_γ. Thus to calculate the *MRL*, set $\gamma = 0.5$.

In order to calculate the control limits the outer control limits are fixed and the inner control limits are calculated for a given MRL_0 by applying the cdf equation derived by Brooke and Evans (1972); as provided in the previous paragraph. The aim is to establish t if $P(T \leq t) = 0.5$.

The corresponding essential TPM (denoted by Q) is used in calculating the MRL and the ARL for the *3-of-4* runs-type signalling rule, see Table 2.8. To establish the MRL t needs to be established in Equation (1.6) where $P(T \leq t) = 0.5$. As an example State S_{11} refers to the first charting statistic plotting in Region 1 followed by the second also plotting in Region 1. State S_{12} refers to the first charting statistic plotting in Region 1 and the second plotting in Region 2. The same line of thought is applied to establish the other states as well.

Table 2.8: Essential TPM for the revised *3-of-4* runs-type signalling rule.

States	$S_{2/3}$	S_1	S_{11}	S_{12}	S_{121}	S_{112}	S_4	S_{44}	S_{43}	S_{434}	S_{443}
$S_{2/3}$	p_2+p_3	p_1	0	0	0	0	p_4	0	0	0	0
S_1	p_3	0	p_1	p_2	0	0	p_4	0	0	0	0
S_{11}	p_3	0	0	0	0	p2	p4	0	0	0	0
S_{12}	p_2+p_3	0	0	0	p_1	0	p_4	0	0	0	0
S_{121}	p_3	0	0	p_2	0	0	p_4	0	0	0	0
S_{112}	p_2+p_3	0	0	0	0	0	p_4	0	0	0	0
S_4	p_2	p_1	0	0	0	0	0	p_4	p_3	0	0
S_{44}	p_2	p_1	0	0	0	0	0	0	0	0	p_3
S_{43}	p_2+p_3	p_1	0	0	0	0	0	0	0	p_4	0
S_{434}	p_2	p_1	0	0	0	0	0	0	p_3	0	0
S_{443}	p_2+p_3	p_1	0	0	0	0	0	0	0	0	0

2.1.7 Two sets of runs rules for the chart (Acosta-Mejia (2007))

Klein (2000a) proposed the *2-of-2* and *2-of-3* runs type signalling rules as a solution to the increased FAR noted when using multiple signalling rules at once. Acosta-Mejia (2007) considered a generalisation of the above rules namely the *k-of-k* and *k-of-k+1* runs type signalling rules. A disadvantage of the *k-of-k* and *k-of-k+1* runs-type signalling rules are that a signal is only produced after k and $k+1$ observations respectively. It is suggested to use the Shewhart $\bar{X}$ control chart (*1-of-1* runs-type signalling rule) combined with the *k-of-k* or *k-of-k+1* runs-type signalling rules in order to monitor both large and small process shifts in location simultaneously.

The following runs-type signalling rules are proposed in conjunction with the Shewhart $\bar{\mathrm{X}}$ control chart (*1-of-1* runs-type signalling rule), with $R(k, k+1, a, b)$ denoting a runs-type signalling rule where k of the last *k+1* observations are evaluated in the interval (a, b) and similarly $R(k, k, a, b)$ denoting a runs-type signalling rule where k of the last k observations are evaluated in the interval (a, b).

1) $R(5,5,-3.3,-0.66)$ and $R(5,5,0.66,3.3)$.
2) $R(7,8,-3.3,-0.5)$ and $R(7,8,0.5,3.3)$.
3) $R(12,13,-3.3,0)$ and $R(12,13,0,3.3)$.

4) $R(9,9,-3.3,0)$ and $R(9,9,0,3.3)$.

The combination of the Shewhart $\overline{X}$ control chart (*1-of-1* runs-type signalling rule) and the *k-of-k* or *k-of-k+1* runs-type signalling rules, signals if a charting statistic plots on or above/below the UCL_2/LCL_2 or between the UCL_2 and UCL_1 or LCL_2 and LCL_1 respectively. The regions and control limits considered are provided in Figure 2.6.

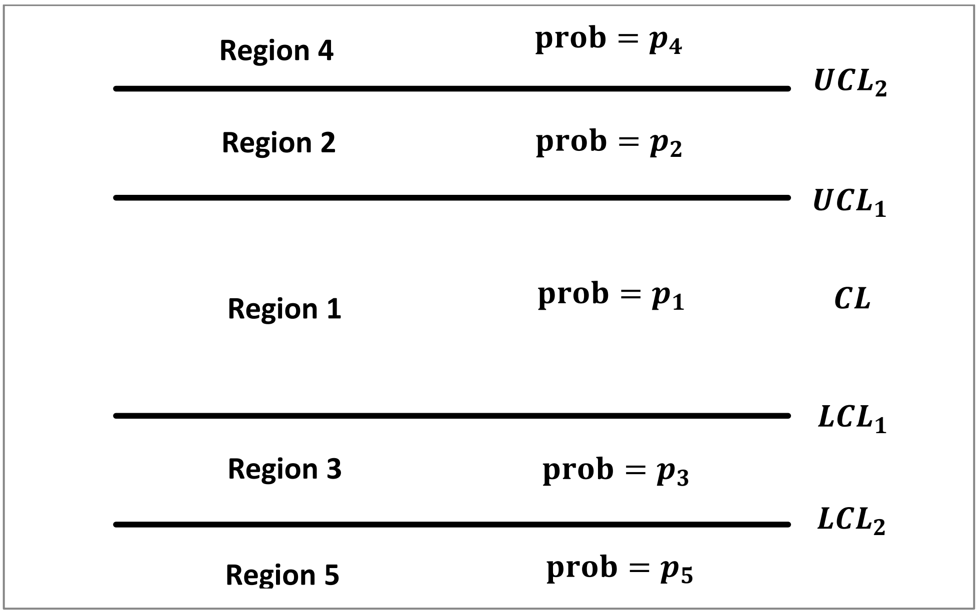

Figure 2.6: Charting regions and probabilities for the *k-of-k* and *k-of-k+1* runs-type signalling rules.

For a process shift of ξ standard deviations the probabilities associated with the regions in Figure 2.6 are as follows:

$p_1 = \phi(UCL_1 - \xi) - \phi(LCL_1 - \xi).$
$p_2 = \phi(UCL_2 - \xi) - \phi(UCL_1 - \xi).$
$p_3 = \phi(LCL_1 - \xi) - \phi(LCL_2 - \xi).$
$p_4 = 1 - \phi(UCL_2 - \xi).$
$p_5 = 1 - p_1 - p_2 - p_3 - p_4.$

In order to calculate the control limits it is suggested to set the UCL_2 and LCL_2 for a desired ARL_0 and then calculate the inner control limits, denoted as UCL_1 and LCL_1 for a minimum ARL_ξ.

The ARL of the control chart consisting of a combination of the Shewhart $\overline{X}$ control (*1-of-1* runs-type signalling rule) chart and the *k-of-k* or *k-of-k+1* runs-type signalling rules was derived by Page (1955) as $ARL = {1 - p_2^k}\Big/{1 - p_1 - p_2 + p_1 p_2^k}$. Alternatively to calculate the ARL the Markov chain approach can be applied by solving Equation (1.4). The essential TPM used in the calculation of the ARL for a combination of the *1-of-1* and the *4-of-4* runs-type signalling rules is provided in Table 2.9.

Table 2.9: Essential TPM for the combined *1-of-1* and *4-of-4* runs-type signalling rules.

States	S_1	S_2	S_3	S_{22}	S_{33}	S_{222}	S_{333}
S_1	p_1	p_2	p_3	0	0	0	0
S_2	p_1	0	p_3	p_2	0	0	0
S_3	p_1	p_2	0	0	p_3	0	0
S_{22}	p_1	0	p_3	0	0	p_2	0
S_{33}	p_1	p_2	0	0	0	0	p_3
S_{222}	p_1	0	p_3	0	0	0	0
S_{333}	p_1	p_2	0	0	0	0	0

The following article explores a variation of the *k-of-w* runs-type signalling rule with some modification applied.

2.1.8 Modified r out of m control chart (Antzoulakos and Rakitzis (2008b))

A modified version of the *r-of-m* control chart studied by Klein (2000a) and Khoo et al. (2004) is proposed by Antzoulakos and Rakitzis (2008b).

The modified *r-of-m* (previously *k-of-w*) control chart outperforms the general *r-of-m* (*k-of-w*) control chart as well as the Shewhart $\bar{X}$ control chart for shifts up to 2.6σ from the mean, where the distributional assumption is that the observations are iid $N(0,1)$. An OOC signal is observed if *r* out of *m* charting statistics plot either above the UCL or below the LCL. The *m-r* charting statistics completing the run of *m* points (charting statistics) being evaluated should plot between the CL and UCL/LCL depending on whether the pattern lies above or below the CL as indicated in Figure 2.7, when for example the *2-of-3* runs-type signalling rule is considered; this is a similar approach for the signalling events applied by the revised *m-of-k* (*k-of-w*) runs-type signalling rule (Antzoula and Rakitzis (2008a)).

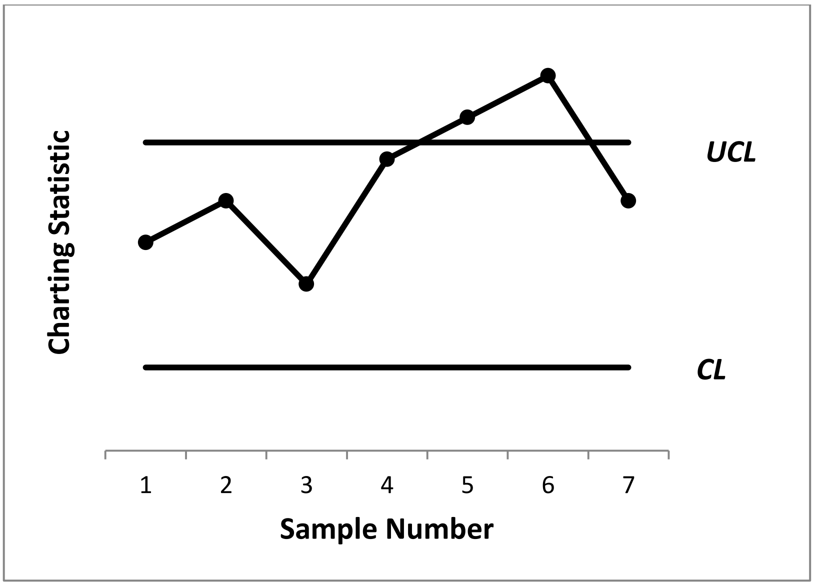

Figure 2.7: The *2-of-3* runs-type signalling rule.

The Regions (state space) used further in the discussion is provided in Figure 2.8. The state space is directly related to the plotting regions. State 1 is associated with Region 1, State 2 with Region 2 etc. The associated probabilities are then given by:

$p_u = 1 - \phi(L - \xi).$

$p_l = 1 - \phi(L + \xi).$

$q_u = \phi(L + \xi) + \phi(\xi) - 1.$

$q_l = \phi(L + \xi) - \phi(\xi).$

where $UCL = L, LCL = -L$ and $CL = 0.$

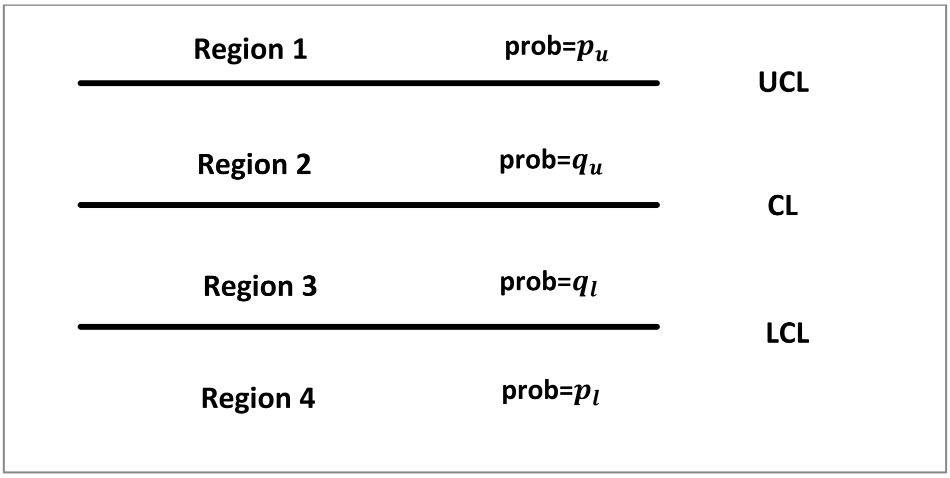

Figure 2.8: Charting regions and probabilities for the *2-of-3* runs-type signalling rule.

As stated previously the ARL can be established by setting up the essential TPM and applying Equation (1.4). The essential TPM for the *2-of-3* runs-type signalling rule is provided in Table 2.10. S_0 refers to the charting statistic plotting in Region 0 and S_{01} to the first charting statistic plotting in Region 0 followed by the second charting statistic plotting in Region 1. The absorbent states for the *2-of-3* and *3-of-4* runs-type signalling rules are provided by the following compound patterns:

$E_{23} = \{11,121,211,44,434,344\}$.

$E_{34} = \{111,1211,1121,444,4344,4434\}$.

Note that E_{23} and E_{34} refers to compound patterns (grouping) which denotes the absorbent states that can be entered for either the *2-of-3* or *3-of-4* runs-type signalling rules. Antzoulakos and Rakitzis (2008b) defines 111 and 444 as signalling events for the *3-of-4* runs-type signalling rule. Similarly 11 and 44 was included in the compound pattern for the *2-of-3* runs-type signalling rule; it could be argued that the since two of three charting statistics are evaluated, the last point has to plot in the signalling region, since only two charting statistics have been plotted the last point could plot IC. Following this line of thought the following signalling events in the compound patterns E_{23} and E_{34}would be excluded since the last charting statistic could still plot IC: 11, 44, 111 and 444. The absorbent states of the *2-of-3* rule reduces then to the following compound pattern with the essential TPM provided in Table 2.10:

$E_{23} = \{121,211,434,344\}$.

Table 2.10: The essential TPM for the modified *2-of-3* runs-type signalling rule.

States	S_2/S_3	S_1	S_4	S_{12}	S_{21}	S_{34}	S_{43}
S_2/S_3	$q_u + q_l$	0	0	0	p_u	p_l	0
S_1	q_l	0	p_l	q_u	0	0	0
S_4	q_u	p_u	0	0	0	0	q_l
S_{12}	$q_u + q_l$	0	p_l	0	0	0	0
S_{21}	q_l	0	p_l	q_u	0	0	0
S_{34}	q_u	p_u	0	0	0	0	q_l
S_{43}	$q_u + q_l$	p_u	0	p_u	0	0	0

Table 2.11: The essential TPM for the modified *2-of-3* runs-type signalling rule excluding previous absorbent States S_{11} and S_{44}.

States	S_2/S_3	S_1	S_4	S_{12}	S_{21}	S_{34}	S_{43}
S_2/S_3	$q_u + q_l$	0	0	0	p_u	p_l	0
S_1	q_l	p_u	p_l	q_u	0	0	0
S_4	q_u	p_u	p_l	0	0	0	q_l
S_{12}	$q_u + q_l$	0	p_l	0	0	0	0
S_{21}	q_l	0	p_l	q_u	0	0	0
S_{34}	q_u	p_u	0	0	0	0	q_l
S_{43}	$q_u + q_l$	p_u	0	p_u	0	0	0

As an alternative to the approach followed in Section 5 to calculate the control limits whereby the essential TPM is established and Equation (1.4) applied, the following steps can be followed to set up the control limits for the *r-of-m* (*k-of-w*) runs-type signalling rule.

1) Choose an integer value r (dependant on the runs-type signalling rule).
2) Fix the desired ARL_0 and calculate p ($p = p_u = p_l$) from Equation (2.1).
3) Calculate the control limit ($UCL = \phi^{-1}(1-p)$).

$$ARL_0 = \mathrm{E}(\mathrm{T}|\xi = 0) = {(1-p^r)}\big/{(2p^r(1-p))} \tag{2.1}$$

For the derivation of Equation (2.1) the reader is referred to Antzoulakos and Rakitzis (2008b).

2.1.9 Run rules $\overline{\mathrm{X}}$ charts when process parameters are unkown (Zhang et al. (2010)

In practise the process parameters are rarely known and have to be estimated. Zhang et al. (2010) introduces a control chart supplemented with the *2-of-3* and *3-of-4* runs-type signalling rules if the process parameters are unknown; a comparison is drawn to cases where the parameters are known and the run-length distribution evaluated.

Zhang et al. (2010) firstly considers the case where the process parameters are known also referred to as case K in the thesis. When the process parameters are known and the observed data (quality characteristic) is distributed iid $N(\mu_0, \sigma_0)$ the charting statistic denoted by $\overline{\mathrm{X}}$ is distributed $N(\mu_0, \frac{\sigma_0}{n})$. The *2-of-3* control chart signals if 2 out of 3 charting statistics plot either above the UCL or below the LCL, where the $UCL = \mu_0 + L\sigma_0$ and $LCL = \mu_0 - L\sigma_0$ (see Section 1.3 and Figure 2.9); note that L denotes a constant value in the equations above. The *2-of-2* control chart signals if two consecutive charting statistics plot either above the UCL or below the LCL, while the *3-of-4* control chart provides a signal if 3 out of 4 charting statistics plot either above the UCL or below the LCL. In order to calculate the ARL the essential TPM has to be set up and Equation (1.4) applied. The transient states related to the essential TPM (see Table 2.12) for the *2-of-3* runs-type signalling rule are:

i. State 00 (S_{00}): Both charting statistics plotting between the UCL and LCL.
ii. State 01 (S_{01}): The first charting statistic plotting between the UCL and LCL and the second plotting above the UCL.
iii. State 02 (S_{02}): The first charting statistic plotting between the UCL and LCL and the second plotting below the LCL.
iv. State 10 (S_{10}): The first charting statistic plotting above the UCL and the second plotting between the UCL and LCL.
v. State 12 (S_{12}): The first charting statistic plotting above the UCL and the second plotting below the LCL.
vi. State 20 (S_{20}): The first charting statistic plotting below the LCL and the second plotting above the UCL.
vii. State 21 (S_{21}): The first charting statistic plotting above the UCL and the second plotting below the LCL.

Table 2.12: Essential TPM for the $\bar{\mathbf{X}}$ control chart supplemented with the *2-of-3* runs-type signalling rule.

State	$\mathbf{S_{00}}$	$\mathbf{S_{01}}$	$\mathbf{S_{02}}$	$\mathbf{S_{10}}$	$\mathbf{S_{12}}$	$\mathbf{S_{20}}$	$\mathbf{S_{21}}$
$\mathbf{S_{00}}$	p_0	p_u	p_l	0	0	0	0
$\mathbf{S_{01}}$	0	0	0	p_0	p_l	0	0
$\mathbf{S_{02}}$	0	0	0	0	0	p_0	p_u
$\mathbf{S_{10}}$	p_0	0	p_l	0	0	0	0
$\mathbf{S_{12}}$	0	0	0	0	0	p_0	p_u
$\mathbf{S_{20}}$	p_0	p_u	0	0	0	0	0
$\mathbf{S_{21}}$	0	0	0	p_0	0	0	0

When considering the *2-of-2* control chart the associated transient states would be:

i. State 0: The charting statistic plotting between the UCL and LCL (Region 0).
ii. State 1: The charting statistic plotting above the UCL (Region 1).

iii. State 2: The charting statistic plotting below the LCL (Region 2).

Table 2.13: Essential TPM for the $\bar{\text{X}}$ control chart supplemented with the *2-of-2* runs-type signalling rule.

State	S_0	S_1	S_2
S_0	p_0	p_u	p_l
S_1	p_0	0	p_l
S_2	p_0	p_u	0

Note that the probabilities for the essential TPM is defined as:

$p_u = \phi(-(L-\xi)\sqrt{n})$.

$p_l = \phi(-(L+\xi)\sqrt{n})$.

$p_0 = 1 - p_u - p_l$.

where $\xi = \frac{(\mu_0 - \mu_1)}{\sigma_0}$.

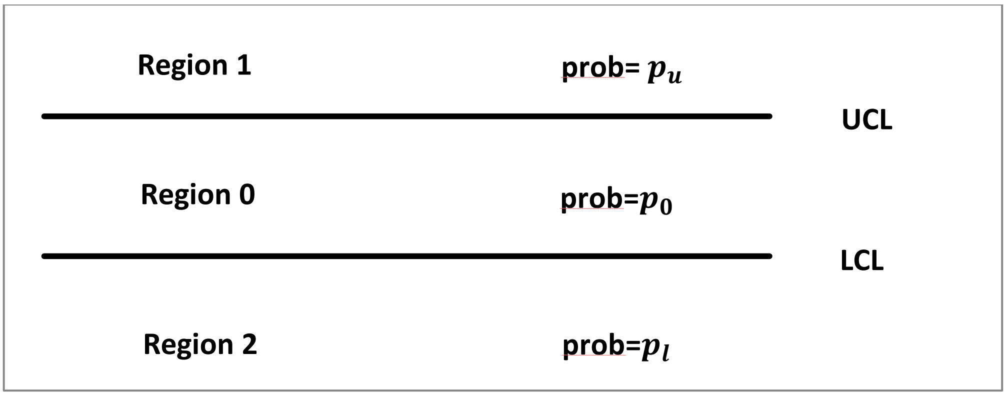

Figure 2.9: Charting regions and probabilities for the $\bar{\text{X}}$ control chart supplemented with the *2-of-3* and *3-of-4* runs-type signalling rules.

When the process parameters (μ and σ) are unknown they are estimated from a phase I process with i=1,2,…m subgroups of size n where a subgroup is denoted as $\{X_{i,1}, X_{i,2}, .. X_{i,n}\}$. Note that the observations are assumed to be distributed iid $N(\mu_0, \sigma_0)$ where μ_0 and σ_0 are estimated as $\widehat{\mu_0} = \frac{1}{mn}\sum_{i=1}^{m}\sum_{j=1}^{n} X_{i,j}$ and $\widehat{\sigma_0} = \sqrt{\frac{1}{m(n-1)}\sum_{i=1}^{m}\sum_{j=1}^{n}\left(X_{i,j} - \bar{X}_i\right)^2}$ where $\bar{X}_i = \frac{1}{n}\sum_{j=1}^{m} X_{i,j}$; $\widehat{\sigma}_0$ is the unbiased estimator for σ_0. Zhang et al. (2010) introduced a phase II sample $\{Y_{i,1}, Y_{i,2}, .. Y_{i,n}\}$ where $Y_{i,j} \sim N(\mu_0 + \xi\sigma_0, \sigma_0)$; note that for an IC process $\xi = 0$. If the process parameters are unknown the control limits are random variables where

$\widehat{LCL} = \hat{\mu}_0 - L\hat{\sigma}_0$ and $\widehat{UCL} = \hat{\mu}_0 + L\hat{\sigma}_0$. The transition probabilities then equals: $p_l = P\left(\bar{Y}_i < \widehat{LCL} | \hat{\mu}_0, \hat{\sigma}_0\right)$, $p_u = P\left(\bar{Y}_i > \widehat{UCL} | \hat{\mu}_0, \hat{\sigma}_0\right)$ and $p_0 = 1 - p_l - p_u$. By applying some algebra, replacing $\widehat{LCL}$ and $\widehat{UCL}$ and subtracting $\mu_0 + \xi\sigma_0$ as well as multiplying by $\frac{\sqrt{n}}{\sigma_0}$ transforms the transition probabilities to the following equations:

$$p_l = P\left((\bar{Y}_i - \mu_0 - \xi\sigma_0)\frac{\sqrt{n}}{\sigma_0} < (\hat{\mu}_0 - L\hat{\sigma}_0 - \mu_0 - \xi\sigma_0)\frac{\sqrt{n}}{\sigma_0} | \hat{\mu}_0, \hat{\sigma}_0 \right).$$

$$= \phi\left((\hat{\mu}_0 - \mu_0)\frac{\sqrt{n}}{\sigma_0} - L\frac{\hat{\sigma}_0\sqrt{n}}{\sigma_0} - \xi\sqrt{n} \right).$$

$$= \phi\left(U - LV - \xi\sqrt{n}\right).$$

$$p_u = P\left((\bar{Y}_i - \mu_0 - \xi\sigma_0)\frac{\sqrt{n}}{\sigma_0} < (\hat{\mu}_0 + L\hat{\sigma}_0 - \mu_0 - \xi\sigma_0)\frac{\sqrt{n}}{\sigma_0} | \hat{\mu}_0, \hat{\sigma}_0 \right).$$

$$= \phi\left(-(\hat{\mu}_0 - \mu_0)\frac{\sqrt{n}}{\sigma_0} - L\frac{\hat{\sigma}_0\sqrt{n}}{\sigma_0} + \xi\sqrt{n} \right).$$

$$= \phi\left(-U - LV + \xi\sqrt{n}\right).$$

where $U = (\hat{\mu}_0 - \mu_0)\frac{\sqrt{n}}{\sigma_0}$, $V = \frac{\hat{\sigma}_0\sqrt{n}}{\sigma_0}$ and L denotes the constant associated with the control limits.

$p_0 = 1 - p_u - p_l$.

Since $\hat{\mu}_0 \sim N(\mu_0, \frac{\sigma_0}{\sqrt{mn}})$, $U \sim N(0, \frac{1}{\sqrt{m}})$ and $f_u(u|m) = f(u|\mu = 0, \sigma = \frac{1}{\sqrt{m}})$. To establish the distribution of V consider $\frac{\hat{\sigma}_0{}^2}{\sigma_0{}^2} \sim \Upsilon(\frac{m(n-1)}{2}, \frac{2}{m(n-1)})$, thus $V^2 \sim \Upsilon(\frac{m(n-1)}{2}, \frac{2n}{m(n-1)})$ and $f_v(v|m,n) = 2v f_\Upsilon(v^2 | \frac{m(n-1)}{2}, \frac{2n}{m(n-1)})$ where f_Υ denotes a gamma distribution with parameters $\frac{m(n-1)}{2}$ and $\frac{2n}{m(n-1)}$. By following the conditioning unconditioning approach the ARL can be calculated as $ARL = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} ARL(1.4) f_u(u|m)\, f_v(v|m,n)\, dv du$ where $ARL(1.4)$ denotes Equation (1.4).

When comparing the ARL_ξ when a shift occurred Zhang et al. (2010) found that the control charts based on estimated parameters supplemented with runs-type signalling rules provided slightly reduced ARL_ξ performance compared to control charts based on known parameters supplemented with runs-type signalling rules.

The next section focusses on parametric control charts monitoring spread supplemented with runs-type signalling rules.

2.2 Runs-type signalling rules for monitoring spread

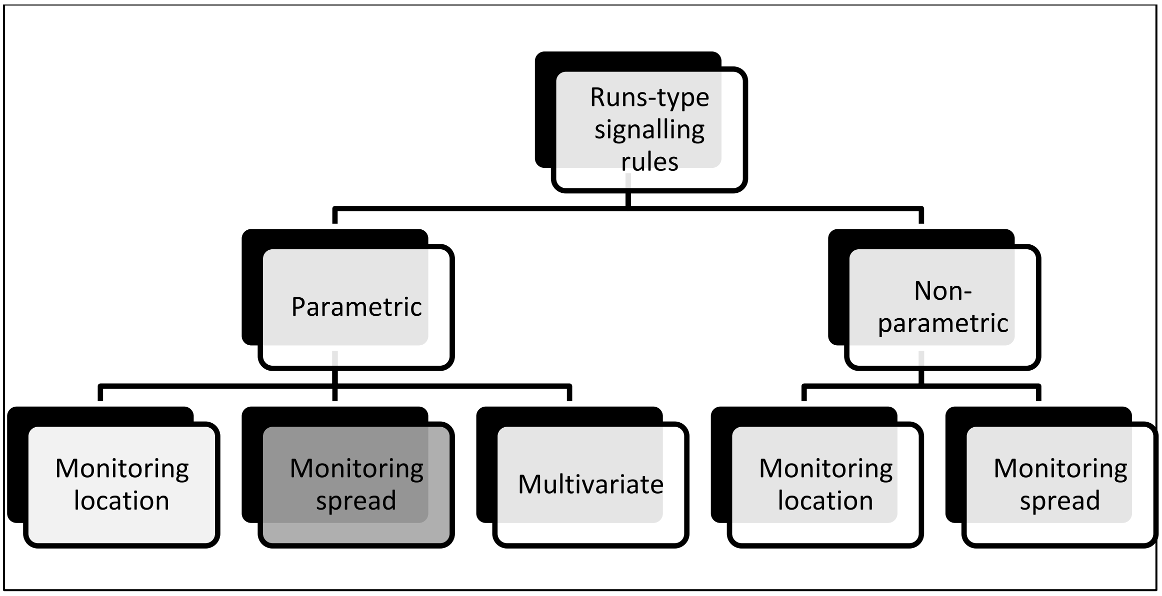

Figure 2.10: Classifications of runs-type signalling rules control chart.

This Section provides a detailed review of the articles based on runs-type signalling rules detecting a shift in spread using a Shewhart-type control chart.

2.2.1 Modified *S* charts for controlling process variability (Klein (2000b))

The aim of the article is to explore additional control charts to help in the detection of variances increases and decreases. The following control charts are evaluated in the article: the standard *S* control chart, the *S* control chart with unequal chi-square tail probabilities, the *S* control chart with equal chi-square tail probabilities and the *2-of-2* control chart based on the standard deviation as charting statistic. The discussion that follows will focus more on the *2-of-2* rule since the scope of the thesis is runs-type signalling rules.

Klein (2000b) mentioned that when considering process increases and decreases for *S*-type control charts the maximum *ARL* is found for a decrease in standard deviation, since the run-length distribution (*ARL* profile) is not symmetrical as can be seen in Lowry et al. (1995). It should be noted that in general the ARL_0 for Shewhart-type S charts are less than 370 which is a typical ARL_0 used for the Shewhart $\bar{X}$ control chart.

For subgroups of size 2 and 5 the standard Shewhart S control chart outperforms all the charts mentioned above. Klein (2000b) stresses the fact that, as in the case of monitoring a shift in location, if the *2-of-2* chart is used in combination with standard S charts an increase would be noted for the IC *FAR* as mentioned by Montgomery (2013). From a practical point of view a low IC *FAR* is desirable.

Three states are considered in order to set up the essential TPM for the *2-of-2* runs-type signalling rule, see Figure 2.11:

i. State 1: The charting statistic plotting between the UWL and LWL.
ii. State 2: The charting statistic plotting on or above the UWL.
iii. State 3: The charting statistic plotting on or below the LWL.

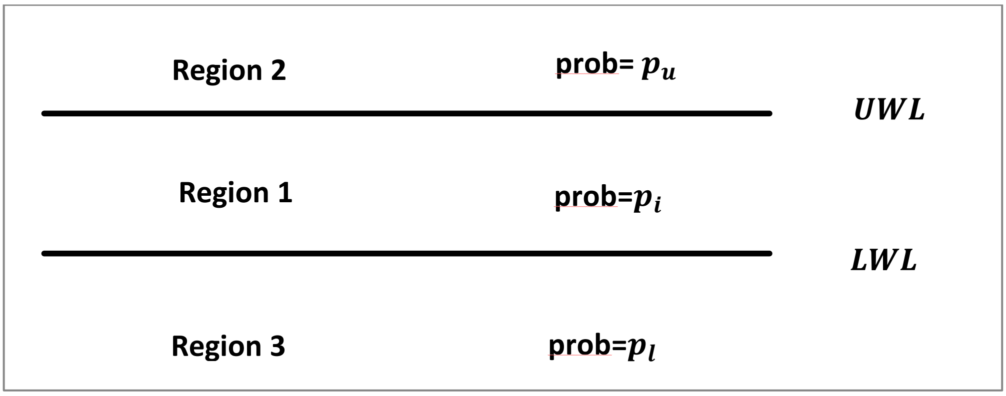

Figure 2.11: Charting regions and probabilities for the *2-of-2* runs-type signalling rule.

As mentioned in previous discussions when applying the Markov-chain approach, the essential TPM (Table 2.14) is required. The ARL can then be calculated by applying Equation (1.4). It should be noted that equal tailed probabilities are used for the *2-of-2* runs-type signalling rule; thus $p_u = p_l$ and $p_u + p_l + p_i = 1$. When applying Equation (1.4) for a given ARL_0 the probabilities can be established since only 1 unknown variable exists due to the constraints mentioned.

Table 2.14: Essential TPM: *2-of-2* runs-type signalling rule.

State	State 1	State 2	State 3
State 1	p_i	p_u	p_l
State 2	p_i	0	p_l
State 3	p_i	p_u	0

Alternatively when calculating the ARL the first passage time equations for the linear system can be solved (Derman et al. 1973). States 1 to 4 are considered, where state 4 is classified as the absorbent state. The first passage time equations are provided below, with M_{14} being the first passage time from State 1 to State 4:

$M_{14} = 1 + p_i M_{14} + p_u M_{24} + p_l M_{34}.$

$M_{24} = 1 + p_i M_{14} + p_l M_{34}.$

$M_{34} = 1 + p_i M_{14} + p_u M_{24}.$

Hurwitz and Mathur (1992) calculated the solution as:

$$M_{14} = \left(\frac{1 - p_i - p_u}{1 + p_u} - \frac{p_l}{1 + p_l}\right)^{-1}.$$

To solve the equation above some additional constraints must be taken into consideration. These constraints are provided below:

$p_u + p_l + p_i = 1.$

$p_u = p_l.$

$M_{14} = ARL_0$.

Calculating the tail probabilities and setting up the control limits

The following Section explores setting up the control limits and calculating the associated probabilities. The distributional assumptions are that observations are distributed iid $N(0,1)$ and that the subgroups of observations collected are independent. In the discussion that follows it should be noted that p refers to the probability of plotting outside the control limits and n denotes the subgroup size.

Consider the following random variable ${(n-1)S^2}/{\sigma^2} \sim \chi^2(n-1)$.

$$p = 1 - P(\chi^2_{(n-1,1-p_l)} < {(n-1)S^2}/{\sigma^2} \leq \chi^2_{(n-1,p_u)})$$

$$= 1 - (1 - p_l - p_u).$$

where $p = p_u + p_l$.

After applying some algebra to the equation above $1 - p$ is calculated as:

$$1 - p = P(LCL < S \leq UCL) = P(\left(\sqrt{\frac{\sigma^2\left(\chi^2_{(n-1,1-p_l)}\right)}{n-1}}\right) < S \leq \left(\sqrt{\frac{\sigma^2\left(\chi^2_{(n-1,p_u)}\right)}{n-1}}\right))$$

Then:

$$LCL = \sqrt{(\sigma^2\left(\chi^2_{(n-1,1-p_l)}\right)/(n-1))}$$

and $UCL = \sqrt{(\sigma^2\left(\chi^2_{(n-1,p_u)}\right)/(n-1))}$.

In order to calculate the average run-length p must first be solved. For the standard Shewhart-type *S* chart the probability and associated *ARL* can be calculated as follows:

$$1 - p = P(LCL < S \leq UCL).$$

$$= P\left({(n-1)LCL^2}/{\sigma^2} < {(n-1)S^2}/{\sigma^2} \leq {(n-1)UCL^2}/{\sigma^2}\right).$$

$$ARL = {1}/{p}.$$

To calculate the tail probabilities for the *2-of-2* runs-type signalling rule consider the equations above.

$$p_i = 1 - (p_u + p_l).$$

$$= P\left({(n-1)LCL^2}/{\sigma^2} < {(n-1)S^2}/{\sigma^2} \leq {(n-1)UCL^2}/{\sigma^2}\right).$$

where $p_l = P\left({(n-1)S^2}/{\sigma^2} < {(n-1)LCL^2}/{\sigma^2}\right)$

and $p_u = P\left({(n-1)S^2}/{\sigma^2} > {(n-1)UCL^2}/{\sigma^2}\right)$.

with σ^2 indicating the shift in process standard deviation; $\sigma^2 = 1$ indicates no shift, $\sigma^2 > 1$ indicates an increase in standard deviation and $\sigma^2 < 1$ indicates a decrease in standard deviation.

From the equations above the control limits and tail probabilities can now be calculated, which provides the capability of deriving the essential TPM and calculating the ARL by applying Equation (1.4).

2.2.2 The performance of control charts for monitoring process variation (Lowry et al. (1995))

The purpose of this article is to give a careful analysis on previously recommended runs-type signalling rules for R-and S charts. The Shewhart R-and S charts are not sensitive to small shifts in the variance of a process, similarly to the Shewhart $\bar{X}$ control chart, thus runs-type signalling rules have been developed. Nelson (1990) emphasizes the need to monitor increases and decreases in process variation in conjunction with the monitoring of the mean, since the variance plays a large part in establishing whether a process is in control.

Lowry et al. (1995) examines various runs-type signalling rules for the R-and S chart, firstly considering the Western electric rules, or combinations of the rules, then proposing a set of runs-type signalling rules with the same IC ARL values as when these rules are used with the Shewhart $\bar{\mathrm{X}}$ control control chart. These runs-type signalling rules' ARL performance is then also compared to CUSUM and EWMA charts.

It was found that runs-type signalling rules based on the S chart outperform rules based on the R chart; Montgomery suggests switching from a R to an S chart for sample sizes larger than 10. The EWMA and CUSUM charts provide slightly better performance than the R-and S charts with runs-type signalling rules, but is more difficult to implement in practice. A drawback of the R-and S chart is that the ARL_0 values are lower than the ARL_0 value of the Shewhart $\bar{X}$ control control chart supplemented with runs-type signalling rules, e.g. the ARL_0 of the S chart based on the *1-of-1* rule is 256.28 compared to 370 when considering the same rule for the Shewhart $\bar{\mathrm{X}}$ control control chart.

To address this issue Lowry et al. (1995) developed runs-type signalling rules with the same ARL_0 value as their Shewhart $\bar{X}$ control chart counterparts. It should be noted that the control limits for the proposed charts are not symmetrical. Some of the proposed rules for the R-and S charts are given below when a sample size of 5 units is considered. The following notation will be implemented below, $R(k, w, a, b)$, where an OOC signal is produced if k of the last w charting statistics plot in the interval (a, b).

Suggested runs-type signalling rules for the *R* chart

$R(1,1,-\infty,-2.23310)$ and $R(1,1,3.53206,\infty)$.

$R(2,3,-2.23310,-1.73414)$ and $R(2,3,2.21759,3.53206)$.

$R(4,5,-2.23341,-1.00527)$ and $R(4,5,1.00405,3.53206)$.

Suggested runs-type signalling rules for the *S* chart

$R(1,1,-\infty,-2.2785)$ and $R(1,1,3.42787,\infty)$.

$R(2,3,-2.2785,-1.76154)$ and $R(2,3,2.1867,3.42787)$.

$R(4,5,-2.2785,-1.01325)$ and $R(4,5,1.00939,3.42787)$.

Lowry et al. (1995) highlights the assumption made in the Western Electric Handbook that if the false alarm probabilities for an R chart matches closely with the $\bar{X}$ control chart then the run length properties are the same. It was found that this is not the case and it is suggested not to use the *FAP* as a measure of the charts performance as shown by Adams, Lowry, and Woodall (1992). The *ARL* should instead be used as a measure of a chart's performance.

To conclude, S charts with runs-type signalling rules outperform R charts with runs-type signalling rules and have slightly decreased performance than the EWMA and CUSUM charts, but there ease of implementation is a big advantage in practice.

2.2.3 Runs rules schemes for monitoring process variability (Anzoulakos et al. (2010))

Anzoulakos et al. (2010) introduces various runs-type signalling rules supplementing the S control chart. These runs-type signalling rules were developed to overcome the high *FAR* noted when using runs-type signalling rules with Shewhart-type control charts. Upper-and lower one-sided control charts were developed to monitor both increases and decreases in the process variation.

Anzoulakos et al. (2010) employs the MDL (median limit) as the CL for the various charts since the distribution of the range and standard deviation is asymmetrical (If the median is used the asymmetrical distribution doesn't affect the number of runs above or below the MDL). The regions considered in order to construct the upper- and lower one-sided control charts are provided in Figure 2.12 and 2.13. If the charting statistic plots in Region 2 State 2 is entered.

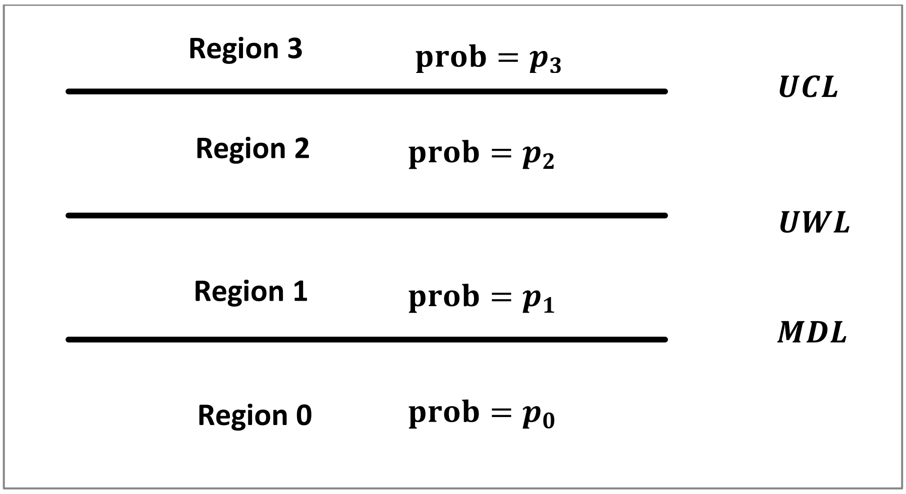

Figure 2.12: Regions relating to the upper one-sided control chart with runs-type signalling rules.

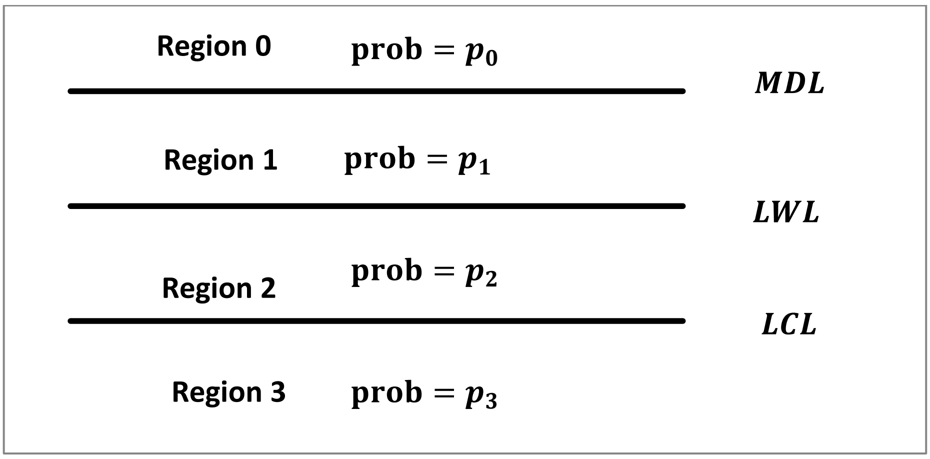

Figure 2.13: Regions relating to the lower one-sided control chart with runs-type signalling rules.

The upper one-sided chart signals if one charting statistic plots on or above the UCL, or if k out of w consecutive charting statistics plot between the UWL and the UCL. On the other hand, the lower one-sided chart signals if a single charting statistic plots on or below the LCL, or if k out of w charting statistics plot between the LWL and the LCL. The two-sided chart combines both conditions listed above in order for a signal to be observed.

The reader is referred to Section 2.2.1 for a discussion on setting up the control limits for the upper-and lower one-sided charts. The UCL and LCL for the upper-and lower one-sided charts are provided below, where the UCL is associated with the upper one-sided chart and the LCL with the lower one-sided chart.

$$UCL = \sigma_0 \sqrt{\frac{\chi^2_{n-1,\alpha}}{n-1}}.$$

$$LCL = \sigma_0 \sqrt{\frac{\chi^2_{n-1,1-\alpha}}{n-1}}.$$

with σ_0 the IC standard deviation, α the *FAR* and n the sample size.

To establish the warning limits the steps are:

i. Firstly determine the UCL or LCL. Then for a given ARL_0 calculate the UWL or LWL.
ii. Various combinations of (UWL, UCL) or (LWL, LCL) could yield the same ARL_0. Choose the pair that minimizes the ARL_ξ.

Two approaches exist in determining the control and warning limits for one-sided charts. In the first approach the outer limit is fixed and the warning limit is calculated for a specified ARL_0. For the second approach the inner limit (warning limit) is fixed and the control limit established for a given ARL_0. Various combinations of warning and control limits could provide the same ARL_0, the dilemma is solved by choosing the pair minimizing the ARL_ξ.

The associated probabilities and regions considered in setting up the TPM for the upper one-sided control charts is provided in Figure 2.12; the TPM consideres states but there is a direct relationship between the region the charting static plots in and the state entered, e.g. if the charting statistic plots in Region 1 State 1 is entered. The equations for the various probabilities are provided below, see Figure 2.12.

$$p_0 = F_{\chi^2_{n-1}}(\frac{(n-1)MDL^2}{(\lambda\sigma_0)^2}.$$

$$p_1 = F_{\chi^2_{n-1}}(\frac{(n-1)UWL^2}{(\lambda\sigma_0)^2} - F_{\chi^2_{n-1}}(\frac{(n-1)MDL^2}{(\lambda\sigma_0)^2}.$$

$$p_2 = F_{\chi^2_{n-1}}(\frac{(n-1)UCL^2}{(\lambda\sigma_0)^2} - F_{\chi^2_{n-1}}(\frac{(n-1)UWL^2}{(\lambda\sigma_0)^2}.$$

$$p_3 = 1 - p_0 - p_1 - p_2.$$

where $F_{\chi^2_{n-1}}(.)$ denotes the cdf of a $\chi^2(n-1)$ random variable.

In order to calculate the ARL the essential TPM needs to be set up and Equation (1.4) applied. The essential TPM's are provided in Table 2.15 and Table 2.16 for the *2-of-3* and *2-of-5* runs-type signalling rules.

Table 2.15: Essential TPM for the upper one-sided chart based on the *2-of-3* runs-type signalling rule.

States	$S_{0,1}$	S_2	S_{21}
$S_{0,1}$	$p_0 + p_1$	p_2	0
S_2	p_0	0	p_1
S_{21}	$p_0 + p_1$	0	0

Table 2.16: Essential TPM for the upper one-sided chart based on the *2-of-5* runs-type signalling rule.

States	$S_{0,1}$	S_2	S_{21}	S_{211}	S_{2111}
$S_{0,1}$	$p_0 + p_1$	p_2	0	0	0
S_2	p_0	0	p_1	0	0
S_{21}	p_0	0	0	p_1	0
S_{211}	p_0	0	0	0	p_1
S_{2111}	$p_0 + p_1$	0	0	0	0

$S_{0,1}$ refers to the charting statistic being either in State 0 or 1, which implies plotting in either Region 0 or 1. S_{21} refers to the first charting statistic plotting in Region 2, followed by the second plotting in Region 1.

S_{211} refers similarly to the first charting statistic plotting in Region 2 followed by two consecutive charting statistics consecutively plotting in Region 1.

Calculating the ARL for the lower one-sided chart is similar, but the probabilities would differ. The probabilities associated with the lower one-sided chart are provided in Figure 2.13. The associated probabilities are:

$p_0 = 1 - F_{\chi^2_{n-1}}(\frac{(n-1)MDL^2}{(\lambda\sigma_0)^2}.$

$p_1 = F_{\chi^2_{n-1}}(\frac{(n-1)MDL^2}{(\lambda\sigma_0)^2} - F_{\chi^2_{n-1}}(\frac{(n-1)LWL^2}{(\lambda\sigma_0)^2}.$

$p_2 = F_{\chi^2_{n-1}}(\frac{(n-1)LWL^2}{(\lambda\sigma_0)^2} - F_{\chi^2_{n-1}}(\frac{(n-1)LCL^2}{(\lambda\sigma_0)^2}.$

$p_3 = 1 - p_0 - p_1 - p_2.$

where $F_{\chi^2_{n-1}}(.)$ denotes the cdf of a $\chi^2(n-1)$ random variable.

In conclusion, Anzoulakos et al. (2010**)** recommends the upper one-sided *2-of-5* chart for small to moderate increases ($1 < \lambda \leq 1.5$). For small to moderate decreases two lower one-sided charts are recommended depending on the shift in standard deviation. For a shift parameter in the following range, $0.8 \leq \lambda < 1$ the *8-of-8* chart is recommended, while if the shift parameter's range is in the interval $0.5 \leq \lambda < 0.8$ the *4-of-5* chart is recommended. The two-sided *k-of-w* chart outperforms the standard two-sided S chart as well as the S control chart with unequal chi-square tail probabilities (Klein 2000b) and the *2-of-2* control chart based on the standard deviation proposed by Klein.

2.2.4 *ARL*-design of *S* charts with *k-of-k* runs rules (Acosta-Mejia et al. (2009))

The aim of the article is to investigate the ARL performance of the S chart supplemented with the *k-of-k* runs-type signalling rule where both increases and decreases in standard deviation are considered.

Three control charts are evaluated: the *k-of-k*, the *1-of-1* in combination with the *k-of-k* and the modified S chart. In all three cases the S chart is supplemented with the *k-of-k* runs-type signalling rule. For the modified S chart the warning limits are taken as the CL. Thus a signal will be produced if a charting statistic plots either on or above the UCL, or on or below the LCL, or if k consecutive points plot in either the interval (CL, UCL) or (LCL, CL). The *k-of-k* chart would signal if k of the last k charting statistics plot in the interval $[UWL, UCL)$ or $[LWL, LCL)$. For the combined *1-of-1* and the *k-of-k* chart a signal is produced if a point plots either on or above the UCL or on or below the LCL, or if the signalling conditions for the *k-of-k* chart are met. The regions considered for the various control charts are provided in Figure 2.14.

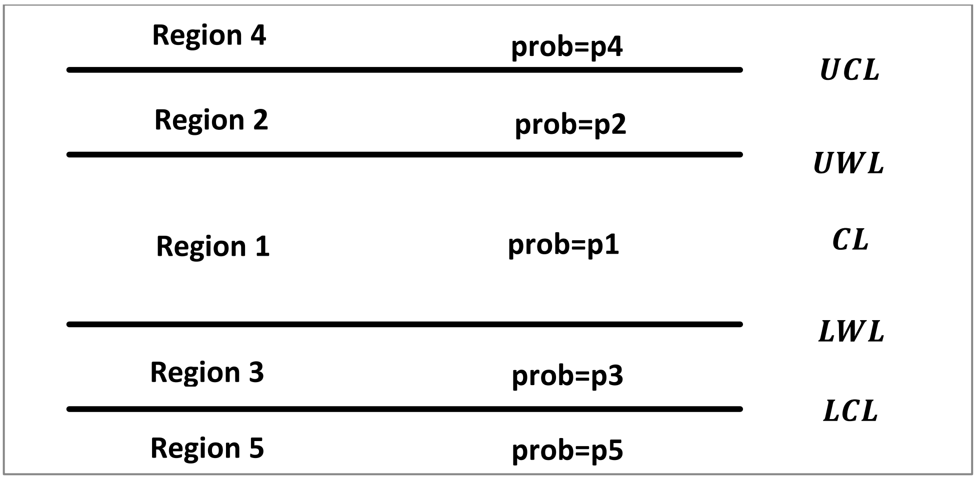

Figure 2.14: Charting regions for *S* chart supplemented with *k-of-k* runs-type signalling rules.

For a detailed discussion on the calculation of the probabilities shown in Figure 2.14 the reader is referred to Section 2.2.1. The probabilities and control limits associated with the S chart supplemented with the combination of the *1-of-1* and *k-of-k* runs-type signalling rules are provided below.

$$p_1 = P(LWL < S < UWL).$$

$$p_1 = F_{\chi^2_{n-1}}\left(\frac{(n-1)UWL^2}{(\lambda\sigma_0)^2}\right) - F_{\chi^2_{n-1}}\left(\frac{(n-1)LWL^2}{(\lambda\sigma_0)^2}\right).$$

$$p_2 = P(UWL < S < UCL).$$

$$p_2 = F_{\chi^2_{n-1}}\left(\frac{(n-1)UCL^2}{(\lambda\sigma_0)^2}\right) - F_{\chi^2_{n-1}}\left(\frac{(n-1)UWL^2}{(\lambda\sigma_0)^2}\right).$$

$$p_3 = P(LCL < S < LWL).$$

$$p_3 = F_{\chi^2_{n-1}}\left(\frac{(n-1)LWL^2}{(\lambda\sigma_0)^2}\right) - F_{\chi^2_{n-1}}\left(\frac{(n-1)LCL^2}{(\lambda\sigma_0)^2}\right).$$

$p_4 = P(S > UCL).$

$$p_4 = 1 - F_{\chi^2_{n-1}}\left(\frac{(n-1)UCL^2}{(\lambda\sigma_0)^2}\right).$$

$p_5 = P(S < LCL).$

$$p_5 = F_{\chi^2_{n-1}}\left(\frac{(n-1)LCL^2}{(\lambda\sigma_0)^2}\right).$$

where $F_{\chi^2_{n-1}}(.)$ denotes the cdf of a $\chi^2(n-1)$ random variable.

$$UCL = \sigma_0\sqrt{\frac{\chi^2_{n-1,1-p_4}}{n-1}}.$$

$$UWL = \sigma_0\sqrt{\frac{\chi^2_{n-1,1-p_2-p_4}}{n-1}}.$$

$$LCL = \sigma_0\sqrt{\frac{\chi^2_{n-1,p_5}}{n-1}}.$$

$$LWL = \sigma_0\sqrt{\frac{\chi^2_{n-1,p_3+p_5}}{n-1}}.$$

For the modified S chart $p_1 = 0$ and $p = p_2 = p_3$ since the chart signals if k out of k consecutive points plot in the interval (CL, UCL) or (LCL, CL), as well as if a point is either above the UCL or below the LCL. The corresponding control limits would then be:

$$UCL = \sigma_0\sqrt{\frac{\chi^2_{n-1,1-p_4}}{n-1}}.$$

$$= \sigma_0\sqrt{\frac{\chi^2_{n-1,0.5+p_2}}{n-1}}.$$

$$LCL = \sigma_0\sqrt{\frac{\chi^2_{n-1,p_5}}{n-1}}.$$

$$= \sigma_0\sqrt{\frac{\chi^2_{n-1,0.5-p_3}}{n-1}}.$$

$$CL = \sigma_0 \sqrt{\frac{\chi^2_{n-1,0.5}}{n-1}}.$$

To set up the essential TPM for the S chart supplemented with the *k-of-k* runs-type signalling rule the following states are considered:

State 1 (S_1): The charting statistic plotting between the warning limits.

State 2 (S_2): The charting statistic plotting between the UWL and UCL.

State 3 (S_3): The charting statistic plotting between the LCL and LWL.

State 4 (S_4): Two consecutive charting statistics plotting between the UWL and UCL.

State 5 (S_5): Two consecutive charting statistics plotting between the LCL and LWL.

State 6 (S_6): Three consecutive charting statistics plotting between the UWL and UCL.

State 7 (S_7): Three consecutive charting statistics plotting between the LCL and LWL.

………

State $2k-2$ (S_{2k-2}): $k-1$ consecutive charting statistics plotting between the UWL and UCL.

State $2k-1$ (S_{2k-1}): $k-1$ consecutive charting statistics plotting between the LCL and LWL.

Table 2.17: Essential TPM for the *S* chart supplemented with *k-of-k* runs-type signalling rules.

State	S_1	S_2	S_3	S_4	S_5	S_6	S_7	….	S_{2k-2}	S_{2k-1}
S_1	p_1	p_2	p_3	0	0	0	0	…	0	0
S_2	p_1	0	p_3	p_2	0	0	0	….	0	0
S_3	p_1	p_2	0	0	p_3	0	0	…	0	0
S_4	p_1	0	p_3	0	0	p_2	0	…	0	0
S_5	p_1	p_2	0	0	0	0	p_3	…	0	0
S_6	p_1	0	p_3	0	0	0	0	…	0	0
S_7	p_1	p_2	0	0	0	0	0	…	0	0
…	….	…	…	…	…	…	…	…	…	…
S_{2k-2}	p_1	0	p_3	0	0	0	0	…	0	0
S_{2k-1}	p_1	p_2	0	0	0	0	0	…	0	0

In order to calculate the ARL Equation (1.4) can be applied. To conclude, combining the *1-of-1* and *k-of-k* runs-type signalling rules for the S chart provides both large and small shift detection capabilities. When considering the *k-of-k* runs-type signalling rule in isolation (not in combination with the *1-of-1*) and looking at a decrease in standard deviation, the *2-of-2* and *3-of-3* runs-type signalling rules provide the smallest ARL_{ξ}, while for an increase larger than $1.1\sigma_0$, the *1-of-1* runs-type signalling rule provide the best performance. When combining the *1-of-1* with the *k-of-k* rules and considering decreases in standard deviation, the *7-of-7* and *10-of-10* runs-type signalling rules provide the best results (optimized for process shift of $1.1\sigma_0$). For increases in standard deviation, the *3-of-3* ($k = 3$) rule in combination with the *1-of-1* runs-type signalling rule provides the best results. Optimization for a shift of $2\sigma_0$ results in the *3-of-3* and *7-of-7* runs-type signalling rules providing the best performance when considering both increases and decreases in standard deviation. For the modified S chart the *1-of-1* rule in combination with the *8-of-8* runs-type signalling rule provides the best results for process decreases from $0.6\sigma_0$ to 0.95, while the *1-of-1* rule in combination with the *10-of-10* runs-type signalling rule provides the best ARL performance for increases in standard deviation larger than $1.2\sigma_0$.

2.2.5 Modified R charts for improved performance (Acosta-Mejia et al. (2008))

Acosta-Mejia et al. (2008) aims to analyse traditional R charts and introduce runs-type signalling rules as modifications in order to monitor increases or decreases in process dispersion more effectively.

Various control charts are evaluated: the traditional R chart, the R chart with asymmetric limits, the R chart with the *k-of-k* runs-type signalling rule (asymmetric and symmetric control limits) and the improved R chart, which is a combination of the *1-of-1* and the *k-of-k* runs-type signalling rules. Asymmetric control limits are considered since the distribution of the range is highly skewed. Acosta-Mejia et al. (2008) focused on small subgroup sizes since the performance of the R-and S charts would be relatively similar. It is well known that for larger sample sizes the S chart outperforms the R chart.

The distribution is assumed to be iid $N(0,1)$ and the standard deviation of the IC process (σ_0) is assumed to be 1 without loss of generality. The run length would then have a generalized geometric distribution with parameter p (see Philippou et al. 1983).

Traditionally the R chart signals if $R_i > D_2$ or if $R_i < D_1$ where $D_1 = \max(0, \sigma_0(d_2 - k_l d_3))$ and $D_2 = \sigma_0(d_2 + k_u d_3)$; d_2 and d_3 are factors associated with the sample size while k_u and

k_l are charting constants. For Shewhart-type charts these values as typically chosen as $k_u = k_l = 3$. Nelson (1990) suggested a signalling rule that produces a signal when a number of consecutive charting statistics plot below the median line. For the R chart supplemented with the *k-of-k* runs-type signalling rule a signal is observed if k consecutive charting statistics plot either on or above or on or below the control limits (When asymmetric/symmetric control limits are used). For the improved R chart, which consists of a combination of the *1-of-1* and the *k-of-k* runs-type signalling rule a signal is produced if the charting statistic plots on or above or on or below the UCL/LCL, or if k consecutive points plot between the MDL(median) and the UCL/LCL. The regions considered in the discussion above as well as the associated probabilities are provided in Figure 2.15, with the state space being defined further on.

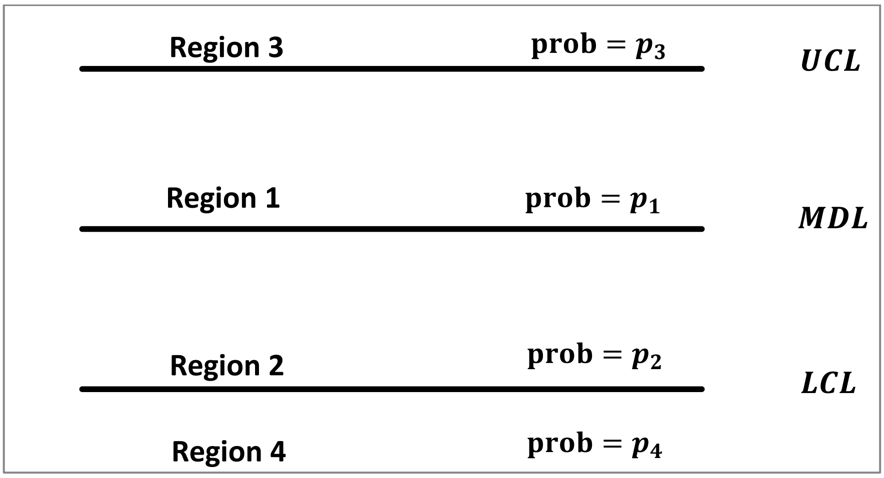

Figure 2.15: Regions considered for the R chart supplemented with *k-of-k* runs-type signalling rules.

$p_1 = P(MDL < X < UCL)$.
$p_2 = P(LCL < X < MDL)$.
$p_3 = P(X \geq UCL)$.
$p_4 = P(X \leq LCL)$.
where X refers to the charting statistic.

When setting up the control limits Acosta-Mejia et al. (2008) employed a general approach. Various pairs of control limits produce a ARL_0 of 370. In order to find the optimum control limit the pair that minimizes the ARL_{ξ} is chosen. As an example an extract of the control limits used for the different control charts are provided in Table 2.18.

Table 2.18: Control limits for various runs-type signalling rules used with the *R* chart.

Chart-type	Signalling rule	Sample size	UCL	LCL
***R* chart with *k*-of-*k* rule**	*2-of-2*	$n = 3$	1.7869	-1.7869
***R* chart with *k*-of-*k* rule**	*3-of-3*	$n = 3$	1.1983	-1.1983
***R* chart with *k*-of-*k* rule (Asymmetric limits)**	*2-of-2*	$n = 3$	2.265	1.434
***R* chart with *k*-of-*k* rule (Asymmetric limits)**	*3-of-3*	$n = 4$	1.405	1.103
Improved *R* chart	*1-of-1* and *9-of-9*	$n = 3$	5.399	0.036
Improved *R* chart	*1-of-1* and *9-of-9*	$n = 4$	5.540	0.123

The Markov chain procedure was used to calculate the ARL and Acosta-Mejia et al. (2008) derived the approximate probability distribution by applying the algorithm AS 126 formulated by Barnard (1978). Alternatively to calculate the ARL the essential TPM can be set up and Equation (1.4) applied. The essential TPM is set up similarly to the previous articles where the different non-absorbent states are considered along with their associated probabilities. Definitions of the various non-absorbent states and the essential TPM for the improved *R* chart incorporating the *9-of-9* runs-type signalling rule are provided below.

State 1 (S_1): A charting statistic plotting between (UCL, LCL).
State 2 (S_2): 2 Consecutive charting statistics plotting between (MDL, UCL).
State 3 (S_3): 2 Consecutive charting statistics plotting between (MDL, LCL).
State 4 (S_4): 3 Consecutive charting statistics plotting between (MDL, UCL).
State 5 (S_5): 3 Consecutive charting statistics plotting between (MDL, LCL).

State 6 (S_6): 4 Consecutive charting statistics plotting between (MDL, UCL).

State 7 (S_7): 4 Consecutive charting statistics plotting between (MDL, LCL).

State 8 (S_8): 5 Consecutive charting statistics plotting between (MDL, UCL).

State 9 (S_9): 5 Consecutive charting statistics plotting between (MDL, LCL).

State 10 (S_{10}): 6 Consecutive charting statistics plotting between (MDL, UCL).

State 11 (S_{11}): 6 Consecutive charting statistics plotting between (MDL, LCL).

State 12 (S_{12}): 7 Consecutive charting statistics plotting between (MDL, UCL).

State 13 (S_{13}): 7 Consecutive charting statistics plotting between (MDL, LCL).

State 14 (S_{14}): 8 Consecutive charting statistics plotting between (MDL, UCL).

State 15 (S_{15}): 8 Consecutive charting statistics plotting between (MDL, LCL).

Table 2.19: Essential TPM for the improved *R* chart.

State	S_1	S_2	S_3	S_4	S_5	S_6	S_7	**....**	S_{14}	S_{15}
S_1	0	p_1	p_2	0	0	0	0	…	0	0
S_2	p_2	0	0	p_1	0	0	0	….	0	0
S_3	p_1	0	0	0	p_2	0	0	…	0	0
S_4	p_2	0	0	0	0	p_1	0	…	0	0
S_5	p_1	0	0	0	0	0	p_2	…	0	0
S_6	p_2	0	0	0	0	0	0	…	0	0
S_7	p_1	0	0	0	0	0	0	…	0	0
…	….	…	…	…	…	…	…	…	…	…
S_{12}	p_2	0	0	0	0	0	0	…	p_1	0
S_{13}	p_1	0	0	0	0	0	0	…	0	p_2
S_{14}	p_2	0	0	0	0	0	0	…	0	0
S_{15}	p_1	0	0	0	0	0	0	…	0	0

In conclusion the R charts with asymmetric control limits outperform their counterparts with symmetric control limits since the distribution of the range is skewed. Improved performance was noted for runs-type signalling rules with asymmetric control limits. The improved R chart has the best ARL performance and Acosta-Mejia et al. (2008) suggested using the *1-of-1* in combination with the *9-of-9* runs-type signalling rule. The advantage is that both large and small increases/decreases in process deviation can be monitored simultaneously. It was also noted that the $ARL_\xi > ARL_0$ for small decreases where $k < 4$ as well as for small increases where $k > 6$ (k refers to the number of consecutive points).

2.2.6 Control charts with switching and sensitizing runs rules for monitoring process variation (Rakitzis and Antzoulakos (2014))

This study involves a one-sided variable sampling interval S control chart supplemented with runs-type signalling rules. Rakitzis and Antzoulakos (2014) aims to improve the performance in detecting small and moderate shifts in standard deviation; secondary the aim is to reduce the frequency of switches without seriously affecting the performance of the chart. Previous work of Rakitzis and Anzoulakos (2011) focuses on S control charts supplemented with sensitizing rules.

The proposed chart combines the *1-of-1* rule with the *k-of-k* runs-type signalling rule where $k \geq 2$. For the choice of sampling interval length the *k*-run switching rule is applied (see Bai and Lee 2002).

Variable sampling intervals (VSI) are used as opposed to the conventional fixed sampling intervals (FSI). When considering VSI and there is an indication of a change in the charting statistic, samples are taken at shorter intervals denoted by d_s, while if no change is noted samples are taken at longer intervals denoted by d_l. For *FSI* all samples are taken at fixed intervals denoted by d_f. The advantage of VSI is that a process shift could be detected quicker.

The *k*-run switching rule (Sw_k) denotes the number of switches between sample interval lengths (how many changes occurred between d_s and d_l). If a random variable D_i is defined as the sampling interval between the i^{th} and $(i+1)^{th}$ sample with $i = 1,2, \ldots$ then:

$D_i = \{d_s, UWL \leq S_i < UCL,$

$\{d_l, S_j < UWL, j = i - k + 1, i - k + 2, \ldots., i$

To evaluate the performance of the charts various performance measures are used: the ARL, ATS (average time to signal, which is the expected time from the start of the process to an OOC condition), $AATS$ (Adjusted average time to signal, which is the expected time from a process shift to a signal) and $ANSW$ (average number of switches, denoting the number of switches between sampling interval lengths).

For a discussion on calculating the control limits and probabilities denoted in Figure 2.16 the reader is referred to Section 2.2.1. The warning and control limits for the upper S chart supplemented with the *1-of-1* and *k-of-k* runs-type signalling rules are provided below.

$$UCL = \sigma_0 \sqrt{\frac{\chi^2_{n-1,1/c}}{n-1}}.$$

$$UWL = \sigma_0 \sqrt{\frac{\chi^2_{n-1,1-b}}{n-1}}.$$

where $b = \left(\frac{d_f - d_s}{d_l - d_s}\right)^{1/k}$ and $c = ARL_0$.

The different regions with their associated probabilities are provided in Figure 2.16.

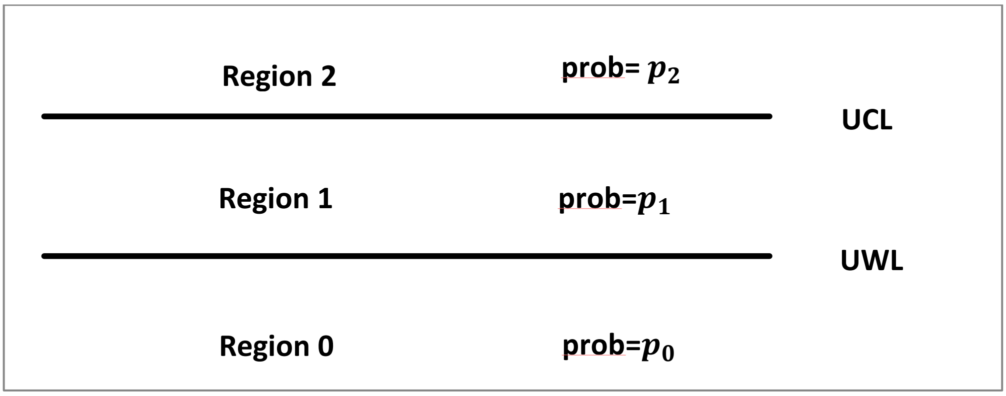

Figure 2.16: Charting regions considered for upper *S*-chart supplemented with the *1-of-1* and *k-of-k* runs-type signalling rules.

$$p_0 = F_{\chi^2_{n-1}}\left(\frac{(n-1)UWL^2}{(\lambda\sigma_0)^2}\right).$$

$$p_1 = F_{\chi^2_{n-1}}\left(\frac{(n-1)UCL^2}{(\lambda\sigma_0)^2}\right) - F_{\chi^2_{n-1}}\left(\frac{(n-1)UWL^2}{(\lambda\sigma_0)^2}\right).$$

$$p_2 = F_{\chi^2_{n-1}}\left(\frac{(n-1)UCL^2}{(\lambda\sigma_0)^2}\right).$$

In order to calculate the ARL the essential TPM (Table 2.20) is set up and Equation (1.4) applied. For the S chart supplemented with the *1-of-1* and *4-of-4* runs-type signalling rules, the following states were considered:

State 0 (S_0): The charting statistic plotting on or below the UWL, in Region 0.

State 1 (S_1): The charting statistic plotting between the UWL and UCL, in Region 1.

State 2 (S_2): Two consecutive charting statistics plotting between the UWL and UCL, in Region 1.
State 3 (S_3): Three consecutive charting statistics plotting between the UWL and UCL, in Region 1.

Table 2.20: Essential TPM for the S chart supplemented with the *1-of-1* and *4-of-4* runs-type signalling rule.

States	S_0	S_1	S_2	S_3
S_0	p_0	p_1	0	0
S_1	p_0	0	p_1	0
S_2	p_0	0	0	p_1
S_3	p_0	0	0	0

To conclude it was seen that the upper sided chart with signalling and switching rules provide superior ARL performance to a chart with only switching rules for a shift of $1 < \lambda < 1.5$; for a decreasing shift of $0.7 < \lambda < 1$ the chart with only switching rules' performance is superior. Using swiching rules enchances the performance since a signal can be detected earlier with smaller sampling intervals being utilised. It should be noted that excessive amount of switches between sampling intervals might lead to practical problems (Amin and Letsinger 1991).

2.3 Conclusion of Chapter 2

Chapter 2 provides an overview of control charts supplemented with runs-type signalling rules in the parametric environment when considering either a shift in location or spread. The articles discussed in Chapter 2 are listed below.

Parametric control charts with runs-type signalling rules for monitoring location

i. Statistical process control using Shewhart control charts with supplementary runs rules (Koutras et al. (2007)).
ii. Two alternatives to the Shewhart $\bar{X}$ control chart (Klein (2000a)).
iii. Two improved runs rules for the Shewhart $\bar{X}$ control chart (Khoo et al. (2006)).
iv. Design of runs rules schemes (Khoo (2003)).
v. The revised *m-of-k* runs rule (Antzoulakos and Rakitzis (2008a)).
vi. The revised *m-of-k* runs rule based on median run length (Low et al. (2012)).

vii. Two sets of runs rules for the chart (Acosta-Mejia (2007)).
viii. Modified r out of m control chart (Antzoulakos and Rakitzis (2008b)).

Parametric control charts with runs-type signalling rules for monitoring spread

i. Modified S charts for controlling process variability (Klein (2000b)).
ii. The performance of control charts for monitoring process variation (Lowry et al. (1995)).
iii. Runs rules schemes for monitoring process variability (Antzoulakos and Rakitzis (2010)).
iv. ARL design of S charts with *k-of-k* rules (Acosta-Mejia et al. (2009)).
v. Modified R charts for improved performance (Acosta-Mejia et al. (2008)).
vi. Control charts with switching and sensitizing runs rules for monitoring process variation (Antzoulakos and Rakitzis (2014)).

Koutras et al. (2007) provided a review on Shewhart control charts supplemented with runs-type signalling rules, deriving the run-length distribution and calculating the ARL. Khoo et al. (2006) designed a control chart, denoted in the literature as the improved runs-type signalling rules, in order to detect both large and small shifts simultaneously. Combining different runs-type signalling rules simultaneously however lead to an increased FAR as noted by Montgomery (2013). To counter the increased FAR the control limits can be adjusted see e.g. Klein (2000a**)** or Acosta-Mejia (2007). Antzoulakos and Rakitzis (2008a) proposed the revised *k-of-w* control chart where a signal can only be detected for the *k-of-w* runs-type signalling rule if all the charting statistics considered either plot above or below the CL when considering an upward or downward shift. Low et al. (2012) proposed a chart based on the MRL since the shape of the run-length distribution changes with the magnitude of the shift in the process mean ranging from highly skewed when the process is IC to almost symmetric when a large shift occurs. Antzoulakos and Rakitzis (2008b) proposed a modified version of the *r-of-m* control chart studied by Klein (2000) and Khoo (2004).

Lowry et al. (1995) proposed S and R charts supplemented with runs-type signalling rules similar to some of the sensitizing rules of the Western Electric company that have similar ARL_0 performance compared to the standard Shewhart $\bar{X}$ control chart. Klein (2000b) proposed various control charts monitoring spread: the standard S control chart, the S control chart with unequal chi-square tail probabilities, the S control chart with equal chi-square tail probabilities and the *2-of-2* control chart. Similarly to the parametric control charts monitoring location, an increase in *FAR* was noted for parametric control charts monitoring spread when multiple runs-type signalling rules were used simultaneously. To address the increased *FAR* Anzoulakos et al. (2010) proposed various signalling rules monitoring spread in order to monitor both increases and decreases in process standard deviation. Acosta-Mejia et al. (2008, 2009) proposed R and S control charts supplemented with the *k-of-k* runs-type signalling rules in order to monitor both increases and decreases effectively. Rakitzis and Antzoulakos (2014) proposed a control chart supplemented with the *k-of-k* runs-type signalling rule and applied VSI as well.

The next chapter focuses on nonparametric control charts for monitoring either location or spread.

Chapter 3

Nonparametric charts: Runs-type signalling rules

3.1 Runs-type signalling rules for monitoring location

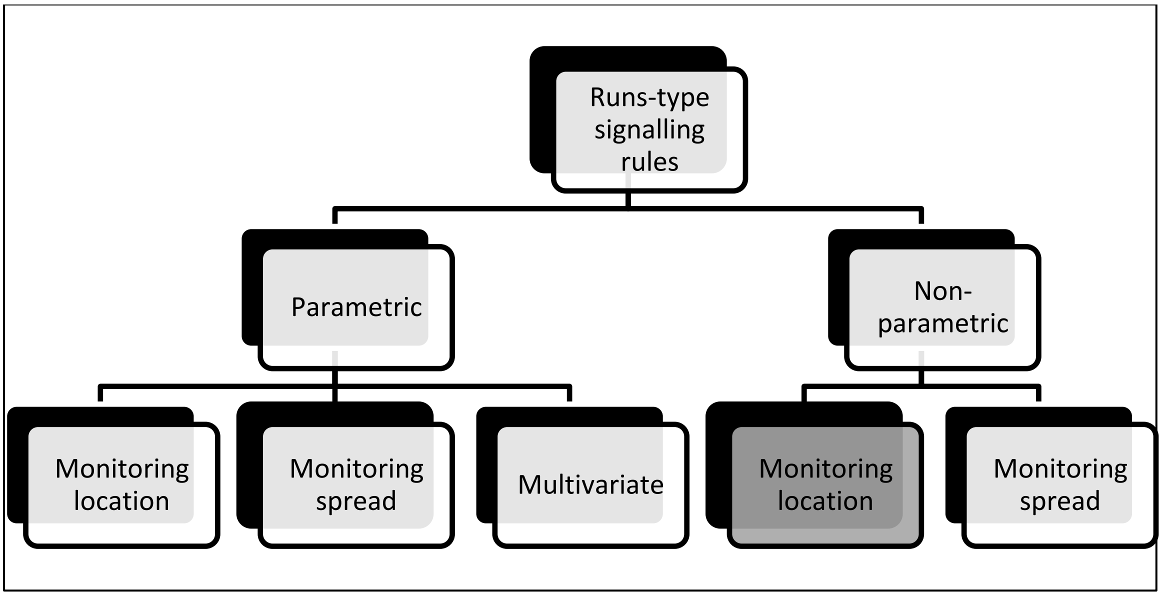

Figure 3.1: Classifications of the runs-type signalling rules control chart.

This section provides a detailed review of the articles based on runs-type signalling rules detecting a shift in location using a Shewhart-type control chart in the nonparametric setting. For the runs-type signalling rules considered in Chapter 2 the distributional assumption was that the observations are distributed iid $N(0,1)$ without loss of generality. For nonparametric charts this assumption doesn't have to hold. The observations could be from any distribution as long as the distributions are continuous. However, if the signed-rank statistic is considered the additional assumption of symmetry is also needed. The review of the articles that follow will provide a more detailed discussion on nonparametric charts supplemented with runs-type signalling rules and there run-length properties.

3.1.1 A nonparametric Shewhart-type signed-rank control chart based on runs (Chakraborti and Eryilmaz (2007))

Bakir (2004) proposed a nonparametric control chart based on signed ranks incorporating the *1-of-1* runs-type signalling rule. Chakraborti and Eryilmaz (2007) built on the work of Bakir (2004) by proposing a nonparametric control chart based on signed ranks supplemented with the k-*of*-k and k-*of*-w runs-type signalling rules that offers smaller FAR*'s* and larger $ARL_0's$; in essence the idea was to design a control chart with a wider range of FAR and ARL values in order to obtain design parameters resulting in FAR and ARL values closer to the nominal/specified values. Both one- and two-sided charts are considered. The median is monitored using the well-known Wilcoxon's signed rank statistic denoted by W_t^+, see p75 and p76.

In the parametric setting, if the assumption of normality is not satisfied, a decrease is noted in the chart's performance. However for non-normal distributions nonparametric charts provide good ARL_0 performance, but not always superior ARL_0 performance. The only condition that needs to be satisfied when considering nonparametric charts is that the distribution should be continuous; if the signed-rank statistic is considered the additional assumption of symmetry is also needed. Some advantages of nonparametric charts are their robustness towards outliers as well as the fact that the IC run-length distribution is the same for every continuous distribution; thus the ARL_0 would be similar for various continuous distributions.

The chart proposed by Bakir (2004) produces an OOC signal if 1 charting statistic plots on or outside the control limits. For the proposed upper one-sided charts a signal is produced if k consecutive charting statistics plot on or above the UCL, when considering the k-*of*-k rule, or if k out of the last w charting statistics plot on or above the UCL for the k-*of*-w rule; thus in a run of w charting statistics k should plot on or above the UCL and the rest then below the UCL. For the k-*of*-k rule all the charting statistics should plot on or above the UCL. For illustrative purposes the reader is referred to Figure 3.2 regarding the k-*of*-k and k-*of*-w runs-type signalling rules where $k = 2$ and $w = 3$. Note that Figure 3.2 (a) refers to the *2-of-2* runs-type signalling rule while Figure 3.2 (b) refers to the *2-of-3* runs-type signalling rule. For the two-sided charts the same decision rule is applicable with k out of k or k out of w charting statistics plotting on or outside the control limits. Two approaches are considered when evaluating the signalling criteria, the Derman and Ross (*DR*) approach and the approach followed by Klein (*KL*). The main difference between the two approaches is that, in the *DR* case, a signal can be produced if a charting statistic plots on or above the UCL followed by another on or below the LCL when considering the *2-of-2* runs-type signalling rule. The Klein method signals if two consecutive charting statistics plot on or above either control limit; thus, both charting statistics need to plot either on or above the UCL or on or below the LCL. Klein's approach is more suitable in detecting an upward or downward shift in the process while the *DR* approach could be used to detect a swing in the median. For the purpose of this study the *KL* approach is followed.

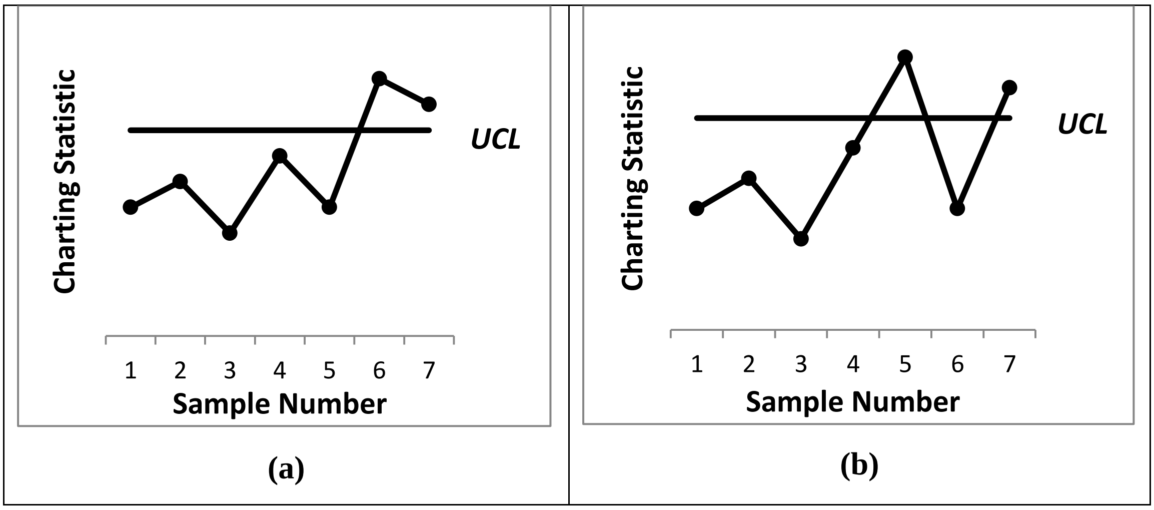

Figure 3.2: The upper one-sided control chart supplemented with the *k-of-k* and *k-of-*w runs-type signalling rules.

The regions and associated probabilities for the upper one-sided and two-sided charts are provided in Figure 3.3; thus, State 1 is entered if the charting statistic plots in Region 1. Thus if $\psi_t \geq UCL$ the State =1 and Region =1, if $\psi_t \leq LCL$ the State = 2 and Region =2, if $LCL < \psi_t < UCL$ the State =0 and Region =0. The reader is referred to p76 regarding the definition of the charting statistic denoted by ψ_t.

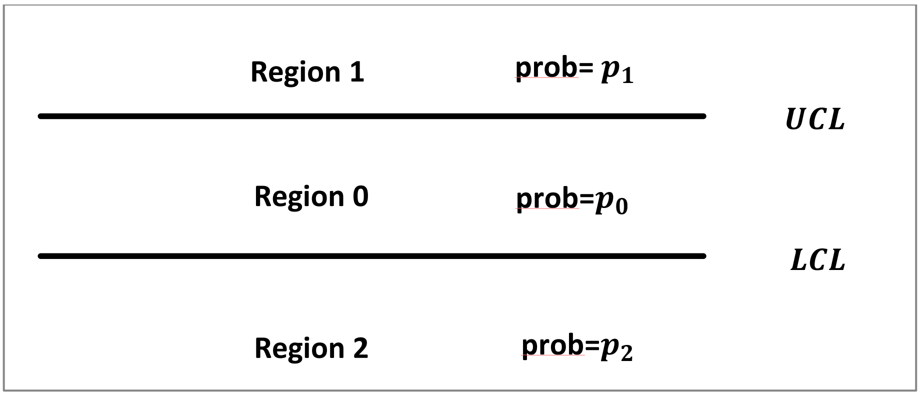

Figure 3.3: Regions considered for the upper one-sided and two-sided nonparametric chart supplemented with the *k-of-k* and *k-of-w* runs-type signalling rules

The charting statistic considered is $\psi_t = 2W_t^+ - {}^{n(n+1)}/_{2}$ where W_t^+ refers to the well-known Wilcoxon signed rank statistic measuring the sum of ranks of absolute values corresponding to positive deviations from the sample median; ψ_t refers to the charting statistic implemented by Bakir (2004), where $\psi_t = \sum_{j=1}^{n} sign(x_{tj} - \Theta_0)R_{tj}^+ \; t = 1,2, \dots\dots,$ $R_{tj}^+ = 1 + \sum_{i=1}^{n} I(|x_{ti} - \Theta_0| < |x_{tj} - \Theta_0|)$ and $sign(x) = 1$ if $x > 0$, 0 if $x = 0$ and -1 if $x < 0$. Note that Θ_0 denotes the IC median and $I(.)$ the indicator function. Bakir (2004) derived the following equation for the IC ARL:

$ARL_0 = {}^{1-p_1^k}/_{(1-p_1)p_1^k}$, where $p_1 = P(\psi_t \geq UCL)$.

An alternative approach to calculate the ARL is to consider the Markov chain approach with the associated essential TPM provided in Tables 3.1 and 3.2 for the *k-of-k* runs-type signalling rule ($k = 2$) and the *k-of-w* runs-type signalling rule ($k = 2$, $w = 3$), respectively. The ARL_0 can then be calculated by applying Equation (1.4). In order to calculate the ARL_ξ Chaktraborti et al. (2007) followed a simulation approach.

Table 3.1: Essential TPM for the upper one-sided nonparametric chart supplemented with the *2-of-2* signalling rule

States	S_0	S_1
S_0	p_0	p_1
S_1	p_0	0

State S_0 refers to the charting statistic plotting below the UCL (Region 0), while S_1 refers to the charting statistic plotting on or above the UCL (Region 1). The control limits are calculated for a specific ARL_0. If $k = 2$, $p_1 = P(\psi_t \geq UCL) = P\left(W_t^+ \geq \frac{UCL}{2} + \frac{n(n+1)}{2}\right)$ and $p_0 = 1 - p_1$.

Table 3.2: Essential TPM for the nonparametric chart supplemented with the *2-of-3* runs-type signalling rule

States	S_0	S_1	S_{01}	S_{10}
S_0	p_0	0	p_1	0
S_1	0	0	0	p_0
S_{01}	p_0	0	0	0
S_{10}	p_0	0	0	0

State S_{01} refers to the first charting statistic plotting below the UCL and the next charting statistic plotting on or above the UCL. On the other hand S_{10} refers to the first charting statistic plotting on or above the UCL followed by the next charting statistic plotting below the UCL. For illustrative purposes the reader is referred to Figure 3.2 regarding the *2-of-3* runs-type signalling rule discussed above. If the third charting statistic plots on or above the UCL a signal is observed for both cases and the absorbent state is entered.

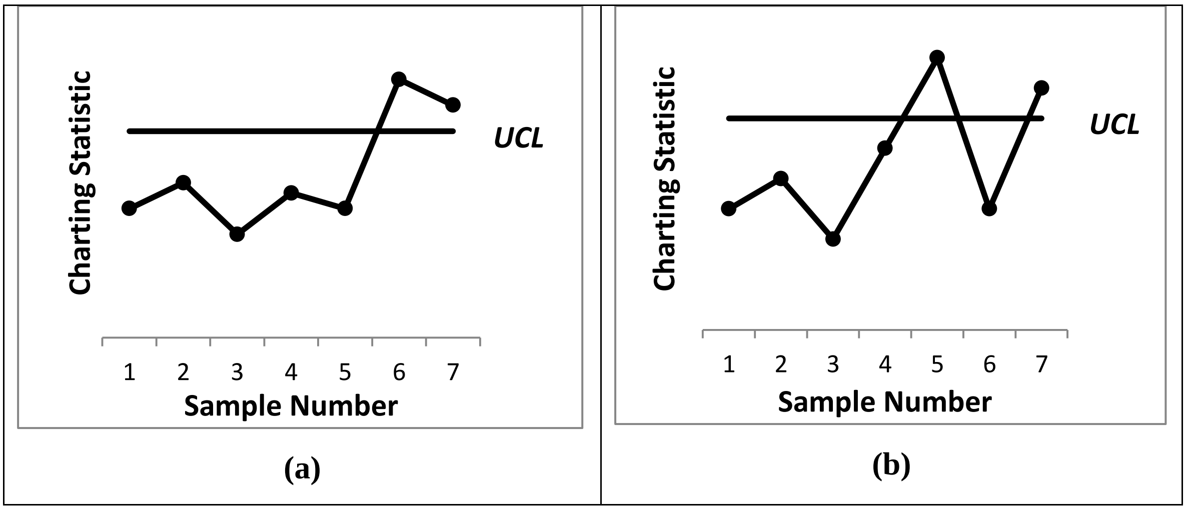

Figure 3.4: Upper one-sided control charts supplemented with the *2-of-3* runs-type signalling rule.

For a visual representation of the *2-of-3* runs-type signalling rule Figure 3.4 (a) and (b) were constructed where state S_{01} is entered at sample number 6 and a signal observed at sample number 7 for both graphs.

The following states are considered for the two-sided nonparametric chart supplemented with *k-of-k* runs-type signalling rules where $k = 2$; S_0, S_1 and S_2. S_0 refers to the charting statistic plotting between the UCL and the LCL. S_1 refers to the charting statistic plotting on or above the UCL and S_2 refers to the charting statistic plotting on or below the LCL. The essential TPM for the nonparametric chart supplemented with the *k-of-k* runs-type signalling rule ($k = 2$) is provided in Table 3.3.

Table 3.3: Essential TPM for the two-sided nonparametric chart supplemented with the *2-of-2* runs-type signalling rule.

States	$\boldsymbol{S_0}$	$\boldsymbol{S_1}$	$\boldsymbol{S_2}$
$\boldsymbol{S_0}$	p_0	p_1	p_2
$\boldsymbol{S_1}$	p_0	0	0
$\boldsymbol{S_2}$	p_0	0	0

When comparing the ARL performance of the upper one-sided charts and considering observations from a normal distribution, it is seen that the proposed *2-of-2* signed rank chart outperforms the *1-of-1* chart proposed by Bakir (2004) as well as the standard Shewhart $\bar{X}$ chart for shifts up to 0.6σ. When considering observations from a double exponential and Cauchy distribution, which is symmetric having heavier tails than the normal distribution, the *2-of-2* signed rank chart outperforms the *1-of-1* chart proposed by Bakir (2004) as well as the standard Shewhart $\bar{X}$ chart for shifts between 0.4σ and 0.8σ deviations. Similar performance is noted for the two-sided chart where the *2-of-2* chart under Klein's methodology outperforms the *1-of-1* chart proposed by Bakir (2004), the standard Shewhart $\bar{X}$ chart as well as the *2-of-2* chart using Derman and Ross' criteria for shifts up to 0.8σ.

3.1.2 A distribution-free Shewhart-type quality control chart based on signed-ranks (Bakir (2004))

Two drawbacks of traditional parametric control charts are that their IC run-length characteristics differ for distributions other than the normal and that their performance deteriorates when outliers are present (Hackl and Ledolter (1991)). To address these shortcomings Bakir (2004) proposed a distribution-free (nonparametric) chart based on signed-ranks; which in effect is a distribution-free control chart considering the *1-of-1* runs-type signalling rule.

The distributional assumptions are that the observations are independent identically distributed and have a continuous symmetric probability distribution. For the upper- and lower one-sided control charts a signal is produced if the charting statistic plots on or above the UCL or on or below the LCL. (The two-sided chart produces a signal if either of the conditions listed for the upper or lower one-sided charts are satisfied). It should be noted that symmetric control limits are used for the two-sided control chart. The regions and probabilities considered for the upper one-sided and two-sided control charts are similar to Figure 3.3 in Section 3.1.1.

For the proposed control chart the characteristic to be monitored is the median, denoted by Θ (Θ_0 refers to the in-control median). The within-group absolute rank of the deviations from the median is defined by $R_{tj}^{+} = 1 + \sum_{i=1}^{n} I\left(|x_{ti} - \Theta_0| < \left|x_{tj} - \Theta_0\right|\right)$ where $I(v)$ denotes the indicator function taking on the values 0 and 1. The charting statistic is then defined as $\psi_t = \sum_{j=1}^{n} sign(x_{tj} - \Theta_0)R_{tj}^{+}$, where $sign(v) = -1,0,1$ if $v < 0, v = 0, v > 0$. The charting statistic is linearly related to the well-known Wilcoxon signed-rank statistic (W_t^{+}) through $\psi_t = 2W_t^{+} - n(n+1)/2$.

Various distributions are considered for the ARL performance comparison. The normal distribution and some light-tailed distributions are evaluated, e.g. the uniform distribution while the heavy-tailed distributions considered are the double exponential and Cauchy distributions. For the normal distribution the exact distributions are calculated for the ARL_{ξ} while a simulation approach is used by Bakir (2004) for the double exponential, uniform and Cauchy distributions. The chart outperforms the traditional Shewhart $\bar{X}$ chart for heavy-tailed distributions but is less efficient than traditional Shewhart $\bar{X}$ chart for light-tailed distributions.

The ARL for the upper- and lower one-sided as well as the two-sided charts equals $ARL = 1/p$, where p refers to the probability of plotting on or outside the control limits for the various charts. The signalling criteria for the upper one-sided, lower one-sided and two-sided charts are given by $p = P(\psi_t \geq UCL)$, $p = P(\psi_t \leq LCL)$ and $p = P(\psi_t \geq UCL) + P(\psi_t \leq LCL)$, respectively.

An alternative approach would be to set up the Essential TPM and use Equation (1.4) to calculate the ARL. For the charts only 1 non-absorbent state is present since a signal is produced if a point plots on or outside the control limits. When considering Figures 3.3 the essential TPM would reduce to a scalar value consisting of p_0, since only 1 transient state exists. The essential TPM is given in Table 3.4.

Table 3.4: Essential TPM for the one-sided and two-sided nonparametric charts supplemented with the *1-of-1* runs-type signalling rule.

States	$\boldsymbol{S_0}$
$\boldsymbol{S_0}$	p_0

To conclude, a nonparametric control chart using the *1-of-1* signalling rule is proposed which outperforms the traditional Shewhart $\bar{X}$ chart for heavy-tailed distributions but shows inferior performance when considering light-tailed distributions. The chart has the added advantage of robustness towards outliers. A drawback of the signed-rank (*SR*) chart is that the underlying distribution has to be symmetric; no such assumptions need to be satisfied when using the sign statistic as will be seen in the following article by Human et al. (2010). Another disadvantage is that the signed-rank chart can only employ the median, whereas the sign statistic can be used to monitor any percentile of interest.

3.1.3 Nonparametric Shewhart-type sign control charts based on runs (Human et al. (2010))

Human et al. (2010) proposed three charts using the sign test statistic; the *1-of-1*, the *k-of-k* and the *k-of-w*. The sign statistic is preferred above the signed rank statistic since the distributional assumption of symmetry is bypassed and any percentile of interest could be monitored which is not the case for the signed rank statistic. The only assumptions that need to be satisfied are that the observations are independent and have a continuous distribution.

Amin (1995) proposed a control chart based on the charting statistic $SN_i = \sum_{j=1}^{n} sign(X_j > \theta_0)$ where $sign(x) = 1$ if $x > 0$, -1 $if\ x < 0$ and 0 if $x = 0$. The charting statistic proposed by Human et al. (2010) is the classic sign statistic $T_i = \sum_{j=1}^{n} I(X_{ij} > \theta_0)$, where θ_0 denotes the target value (median or percentile) and $I(X_{ij} > \theta_0)$ the indicator function. When the process is IC $\theta = \theta_0$ and $p_0 = P\left(X_{ij} > \theta_0\right)$, indicating the probability of an observation being larger than the median. T_i follows a binomial distribution with parameters n and p_0. The charting statistic T_i then indicates the number of observations in the i^{th} sample larger than some target value θ_0.

The upper and lower control limits as well as the center line are given by $UCL = n - b, LCL = a$ and $CL = n\theta_0$, respectively, where a and b are values between 0 and n, typically chosen to obtain a large ARL_0. The signal criteria for the upper and lower one-sided charts using the *1-of-1* runs-type signalling rule, is to produce a signal if $T_i \leq LCL$ (lower one-sided chart) or $T_i \geq UCL$ (upper one-sided chart). For the two-sided chart a signal is produced if either one of the one-sided signalling conditions are satisfied. The *k-of-k* (w) chart produces a signal if k out of the last k (w) charting statistics plot either on or above the UCL or on or below the LCL. Human et al. (2010) distinguishes between the Derman and Ross (*DR*) and Klein (*KL*) approach to signalling. As an example, when considering Figure 3.5 and following the *DR* approach a signal would be produced for both (a) and (b) for the *2-of-2* runs-type signalling rule. However when following the KL approach a signal would only be produced for (a), since both the charting statistics needs to plot above the UCL or both below the LCL.

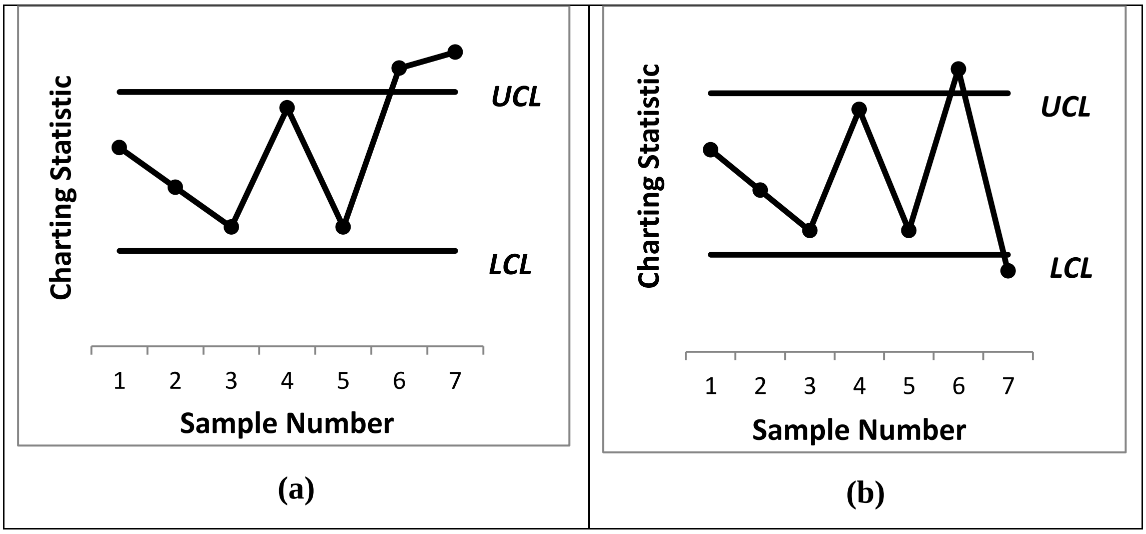

Figure 3.5: Two-sided control charts supplemented with the *2-of-2* runs-type signalling rule.

For the purposes of this study the Klein approach would be followed. The associated probabilities for the one- and two-sided *1-of-1* charts are given by:

$p_1 = P(T_i \geq UCL) = P(T_i \geq n - b) = I_p(n - b, b + 1).$

$p_2 = P(T_i \leq LCL) = P(T_i \leq a) = I_p(a + 1, n - a).$

$p_{two-sided} = 1 - P(LCL < T_i < UCL) = 1 - I_p(a + 1, n - a) + I_p(n - b, b + 1).$

where $I_p(u, v) = [B(u, v)]^{-1} \int_0^p w^{u-1} (1 - w)^{v-1}$ is the cdf of a $Beta(u, v)$ distribution with $0 < w < 1$ and $0 < p < 1$. Figure 3.6 gives a graphical representation of the probabilities and associated regions.

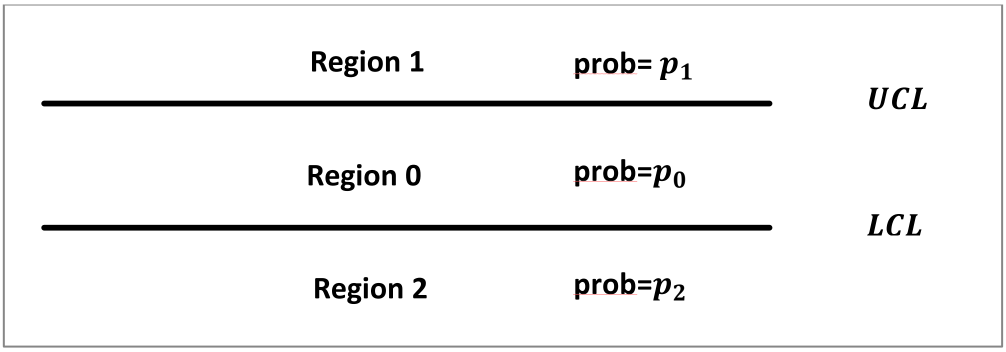

Figure 3.6: Probabilities and regions considered for two-sided *1-of-1*, *k-of-k* and *k-of-w* control charts.

In order to calculate the ARL the essential TPM is set up and Equation (1.4) applied. The essential TPM for the upper and lower one-sided as well as the two-sided charts supplemented with the *k-of-w* ($k = 2$, $w = 3$) runs-type signalling rule are provided in Tables 3.5 to 3.7. S_0 (State 0) refers to the charting statistic plotting in Region 0, while S_{01} refers to the first charting statistic plotting in Region 0 followed by another plotting in Region 1. S_ϕ refers to the dummy state.

Table 3.5: Essential TPM for the upper one-sided control chart supplemented with the *2-of-3* runs-type signalling rule.

States	S_ϕ	S_0	S_1	S_{01}	S_{10}
S_ϕ	0	$1-p_1$	p_1	0	0
S_0	0	$1-p_1$	0	p_1	0
S_1	0	0	0	0	$1-p_1$
S_{01}	0	0	0	0	$1-p_1$
S_{10}	0	$1-p_1$	0	0	0

Table 3.6: Essential TPM for the lower one-sided control chart supplemented with the *2-of-3* runs-type signalling rule.

States	S_ϕ	S_0	S_2	S_{02}	S_{20}
S_ϕ	0	$1-p_2$	p_2	0	0
S_0	0	$1-p_2$	0	p_2	0
S_2	0	0	0	0	$1-p_2$
S_{02}	0	0	0	0	$1-p_2$
S_{20}	0	$1-p_2$	0	0	0

Table 3.7: Essential TPM for the two-sided control chart supplemented with the *2-of-3* runs-type signalling rule.

States	S_ϕ	S_0	S_1	S_2	S_{01}	S_{10}	S_{02}	S_{20}
S_ϕ	0	p_0	p_1	p_2	0	0	0	0
S_0	0	p_0	0	0	p_1	0	p_2	0
S_1	0	0	0	p_2	0	p_0	0	0
S_2	0	0	p_1	0	0	0	0	p_0
S_{01}	0	0	0	p_2	0	p_0	0	0
S_{10}	0	p_0	0	0	0	0	p_2	0
S_{02}	0	0	p_1	0	0	0	0	p_0
S_{20}	0	p_0	0	0	p_1	0	0	0

The ARL performance of the proposed sign chart supplemented with the *1-of-1*, *k-of-k* and *k-of-w* runs-type signalling rules are compared to the *1-of-1* and the *2-of-2 SR* charts proposed by Bakir (2004) and Chakraborti et al. (2007), under the normal, double exponential and Cauchy distributions. The chart performing the best, under the normal distribution when considering one-sided charts, was the *2-of-2 SR* chart closely followed by the *2-of-2* and *2-of-3* sign charts. Under the double exponential and Cauchy distributions the *2-of-2* and *2-of-*3 sign charts outperform the rest. Similar performance is noted for two-sided control charts with the *2-of-2* and *2-of-3* sign charts outperforming the rest under the double exponential and Cauchy distributions, only to be beaten by the *2-of-2 SR* chart under the normal distribution.

3.1.4 A distribution-free control chart based on order statistics (Balakrishnan et al. (2010))

A disadvantage of the median control chart, proposed by Chakraborti et al. (2004), is that it only considers a single order statistic to determine whether a process is IC or OOC. It may be the case that this order statistic lies within the control limits while the rest of the observations are outside the control limits; resulting in an IC condition while in fact the process could well be OOC. Balakrishnan et al. (2010) proposed a control chart for phase II applications taking into account the location of the single order statistic (*1-of-1* runs-type signalling rule) from the test sample as well as the number of observations between the control limits. The general idea of the control chart was to increase sensitivity for small shifts in the process.

In order to construct the control limits for the proposed chart a reference sample of size m is chosen. It is assumed that the reference sample $X_1, X_2 \dots., X_m$ was collected from an IC distribution with cdf $F(x)$ and a test sample $Y_1, Y_2, \dots, Y_n$ exits with cdf $G(x)$. Another assumption is that the observations of the test sample are drawn independently from one another as well as from the reference sample. The aim of the chart is then to check whether the observed process is IC or OOC. The null and alternative hypotheses would then be: $H_0: F(x) = G(x)$ and $H_1: F(x) \neq G(x)$; thus under H_0 the process would be IC and under H_1 OOC. The control limits would then be two order statistics where $LCL = X_{a:m}$ and $UCL = X_{b:m}$ (if symmetrical control limits are used $b = m - a + 1$); note that $X_{a:m}$ and $X_{b:m}$ are defined as two order statistics where $1 \leq a \leq b \leq m$. In order to determine the design parameters m, n, a, b, j and r two approaches are considered. The first approach considers a specific FAR while the second considers a specific ARL_0. The first approach involves solving the equation, $FAR = 1 - P(LCL \leq Y_{j:r} \leq UCL \text{ and } L \geq r)$, where L is the number of observations of the test sample lying between the control limits established from the reference sample. The second approach involves a conditioning approach (see Chakraborti (2000)) where, given the control limits $X_{a:m} = x_a$ and $X_{b:m} = x_b$, the run-length follows a geometric distribution with a success probability $p(x_a, x_b) = P(x_a \leq Y_{j:n} \leq x_b \text{ and } L(Y_1, Y_2, \dots . Y_n; x_a, x_b) \geq r)$ where $L(Y_1, Y_2, \dots . Y_n; X_{a:m}, X_{b:m}) = |(i \in (1,2, \dots, n): X_{a:m} \leq Y_i \leq X_{b:m})|$. The probability mass function equals $(1 - p(x_a, x_b))(p(x_a, x_b))^{k-1} = p(x_a, x_b)^{k-1} - p(x_a, x_b)^k$. It follows that the unconditional distribution of the run-length is $P(T = k) = D(k - 1) - D(k)$ where $D(0) = 1$ and $D(k) = E_{X_{a:m}, X_{b:m}}(p(X_{a:m}, X_{b:m})^k), k = 1,2, \dots$. The ARL can then be calculated as follows: $ARL = E(T) = \sum_{k=0}^{\infty} E_{X_{a:m}, X_{b:m}} (p(X_{a:m}, X_{b:m})^k) = \sum_{k=0}^{\infty} D(k)$.

In conclusion, the proposed chart outperforms the Shewhart $\bar{X}$ chart for non-normal data, and the median control chart proposed by Chakraborti (2004).

3.1.5 A Phase II nonparametric control chart based on precedence statistics with runs-type signalling rules (Chakraborti et al. (2009))

Two control charts are proposed for monitoring the unknown location parameter of a continuous population in phase II applications. Exact run length distributions and equations for the ARL are derived for the proposed charts using a conditioning approach and some results from the theory of runs. The proposed charts, the *2-of-2 DR* and *2-of-2 KL*, implement the *DR* (Derman and Ross) approach as well as the *KL* (Klein) approach. To provide insight into the two approaches the following situation is discussed: Consider two consecutive charting statistics, the first plotting above the UCL followed by the second plotting below the LCL (when considering a two-sided control chart). The *DR* chart will provide a signal when the *2-of-2* runs-type signalling rules is considered while the *KL* chart would not signal since both charting statistics must either plot above the UCL or below the LCL, respectively. The reader is referred to Figure 3.4 (b) for a graphical illustration of two consecutive charting statistics when distinguishing between the *DR* and *KL* approach. The charting statistic could be any other statistic but the median is considered for simplicity and robustness.

It is assumed that a reference sample of size m is available from an IC process with an unknown continuous distribution function F. Let $X_{1:m} < X_{2:m} < \cdots < X_{m:m}$ denote order statistics from the reference sample with LCL=$X_{a:m}$ and the UCL=$X_{b:m}$ where $1 \leq a < b \leq m$. In phase II, test samples of size n are drawn sequentially and independently from each other as well as the reference sample. The charting statistic $Y_{j:n}^{h}$ can be any order statistic but the median is chosen as mentioned in the introduction. The following indicator variables were defined for the h^{th} test sample: $Z_h = 1, if\ Y_{j:n}^{h} \notin (LCL, UCL)$ or $Z_h = 0, if\ Y_{j:n}^{h} \in (LCL, UCL)$ with $h = 1,2,3$ ... Since the control limits are order statistics from the reference sample the signalling indicators $Z_1, Z_2,$ .... are dependant binary variables. It is also assumed that the test sample originates from a continuous distribution with cdf G. As an example for the *1-of-1* runs-type signalling rule the probability of no signal given $X_{a:m} = x_1$ and $X_{b:m} = x_2$ is $P\left(x_1 < Y_{j:n} < x_2 \middle| X_{a:m} = x_1, X_{b:m} = x_2\right) = G_j(x_2) - G_j(x_1)$ where G_j denotes the cdf of the jth order statistic in a sample of size n from a distribution with cdf G. Since $G_j(x) = I_{G(x)}(i, n-j+1)$ where $I_x(y,z)$ denotes the incomplete beta function the probability of a signal reduces to $G_j(x_2) - G_j(x_1) = I_{G(x_2)}(i, n-j+1) - I_{G(x_1)}(i, n-j+1)$.

The unconditional probability of no signal can be found by averaging over the joint distribution of $X_{a:m}$ and $X_{b:m}$. Transforming the results to $(0,1)$ results in the following equation:

$$P(Z_h = 0) = \int_0^1 \int_0^y I_{GF^{-1}(y)}(j, n-j+1) - I_{GF^{-1}(x)}(j, n-j+1) f_{a,b}(x,y)\,dxdy.$$

Note that the outside integral transforms the equation onto the interval $(0,1)$, while the integral denoted by $\int_0^y I_{GF^{-1}(y)} \ldots f_{a,b}(x,y)\,dxdy$ is responsible for the averaging.

The process is declared OOC when:

a) A single charting statistic plots on or outside the control limits.
b) Two consecutive charting statistics, both plot on or above the UCL or, both plot on or below the LCL, or the first charting statistic plotting on or above the UCL with the second one plotting on or below the LCL or, the first charting statistic plotting below the LCL with the second plotting on or above the UCL. (*DR* approach).
c) Both charting statistics plotting on or above the UCL or on or below the LCL (*KL* approach).

The performance of the control charts are evaluated by considering their run-length distribution, which can be viewed as the waiting time until the first signal. The waiting times are defined as follows (refer to the OOC conditions a,b and c previously mentioned):

a) $T_1 = \min(t: Z_t = 1)$.
b) $T_2 = \min(t: Z_{t-1} = 1, Z_t = 1)$.
c) $T_2' = \min(T_2^{(1)}, T_2^{(2)})$
   where $T_2^{(1)} = \min(t: Z_{t-1}' = 1, Z_t' = 1)$, $T_2^{(2)} = \min(t: Z_{t-1}' = 1, Z_t' = 2)$
   and $Z_h' = 0 \; if \; Y_{j:n} \in (X_{a:m}, X_{b:m})$, $Z_h' = 1 \; if \; Y_{j:n} \geq X_{a:m}$, $Z_h' = 2 \; if \; Y_{j:n} \leq X_{a:m}$.

George and Bowman (1995) derived, as an alternative to the Markov chain approach, the distribution for the total number of successes in a sequence of n binary trials as:

$P(S_n = s) = \binom{n}{s} \sum_{i=0}^{n-s} (-1)^i \binom{n-s}{i} \lambda_{s+i}$.

where $\lambda_t = P(Z_1 = 1, \ldots\ldots, Z_t) = 1$ for $t = 1,2 \ldots\ldots n$.

This result is used in deriving the unconditional run-length distribution. Alternatively the Markov chain approach can be used whereby the essential TPM is established and equation (1.4) applied.

The unconditional distribution of T_2 (the waiting time for the first signal of the *2-of-2 DR* chart) is given by the following equations with the relevant proof provided in the article's appendix.

$P(T_2 = x) = 0$ if $0 \leq x < 2$ and λ_2 if $x = 2$.

For $x = 3$,

$$P(T_2 = x) = \sum_{y=1}^{x-2} \sum_{j=0}^{\min\left(y, \left|\frac{x-y-2}{2}\right|\right)} \sum_{i=0}^{y} (-1)^j (-1)^i \binom{y}{j} \binom{y}{i} \binom{x-2(j+1)-1}{y-1} \lambda_{x-y+i}.$$

When considering the *2-of-2 KL* chart the random variables $Z_1', Z_2', \ldots, Z_t'$ are distributed iid, conditionally on $X_{a:m}$ and $X_{b:m}$, with the probability plotting above and below the UCL and LCL given by:

$p_L = P\left(Y_{j:n} \leq X_{a:m} | X_{a:m} = x_1\right) = I_{G(x_1)}(j, n-j+1)$

and

$p_U = P(Y_{j:n} \geq X_{b:m} | X_{b:m} = x_2) = 1 - I_{G(2)}(j, n-j+1)$.

Note that the random variables $Z'_1, Z'_2, \ldots, Z'_t$ mentioned above are defined as indicator variables when applying the KL approach.

The distribution of the T_2^*, defined as the waiting time for the first signal when considering the *2-of-2* KL control chart is obtained by applying Theorem 5.2 of Fu and Lou (2003, p68):

$$P(T_2^* = x | X_{a:m}, X_{b:m}) = \underline{\xi} Q^{x-1}(I - Q)\underline{1}, \quad x \geq 2.$$

where

$\underline{\xi} = [1 \quad 0 \quad 0 \quad 0]$ and Q is provided in Table 3.8.

Table 3.8: Essential TPM (Q) for the two-sided control chart supplemented with the *2-of-2* runs-type signalling rule.

States	S_ϕ	S_0	S_U	S_L
S_ϕ	0	$1 - p_L - p_U$	p_U	p_L
S_0	0	$1 - p_L - p_U$	p_U	p_L
S_U	0	$1 - p_L - p_U$	0	p_L
S_L	0	$1 - p_L - p_U$	p_U	0

Note that the transition states in Table 3.8 are denoted by S_ϕ, S_0, S_U and S_L, while the definition of the George and Bowman (1995) equation on p87 defines S_n as the total number of successes in a sequence of n binary trials. The transition states referred to in Table 3.8 are defined as the empty state denoted by S_ϕ, the state denoted by S_0, defined as the state entered if the charting statistic plots between the UCL and the LCL, the state denoted by S_U and defined as the state entered if the charting statistic plots on or above the UCL and lastly the state denoted by S_L defined as the state entered if the charting statistic plots on or below the LCL. The reader is referred to Figure 3.7 regarding the charting regions and the transition states entered.

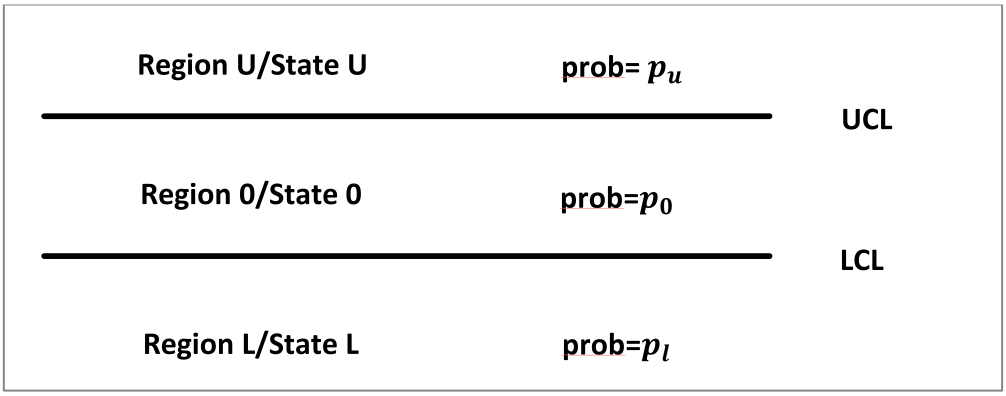

Figure 3.7: Charting regions/states and probabilities for the *2-of-2* runs-type signalling rules.

The unconditional distribution of T_2' is obtained from T_2^* (by averaging over the joint distribution of $X_{a:m}$ and $X_{b:m}$)

$$P(T_2' = x) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} P(T_2^* = x | X_{a:m} = x_1, X_{b:m} = x_2) h_{a,b}\,(x_1, x_2) d_{x1} d_{x2}, \quad x > 2.$$

where

$$P(T_2^* = x | X_{a:m}, X_{b:m}) = \underline{\xi} Q^{x-1}(I - Q)\underline{1}, \qquad \mathrm{x} \geq 2.$$

and $h_{a,b}(x_1, x_2)$ refers to the joint pdf of $X_{a:m}$ and $X_{b:m}$.

Chakraborti et al. (2009) derived exact *ARL* distributions for both the *DR* and *KL* charts. Given $X_{a:m}$ and $X_{b:m}$ the random variable T_2 has a geometric distribution of order 2, see Balakrishnan and Koutras (2002). The conditional expected value is then equal to the following equation: $E(T_2 | X_{a:m} = x_1, X_{b:m} = x_2) = \frac{1+p}{p^2}$ where $p = 1 - \left(G_j(x_2) - G_j(x_1)\right) = 1 - (I_{G(x_2)}(j, n - j + 1) - I_{G(x_1)}(j, n - j + 1))$. The unconditional *ARL* for the *DR* chart uses the previous result and is given by the following equation:

$$ARL_{DR} = E_{X_{a:m}, X_{b:m}} E(T_2 | X_{a:m}, X_{b:m})$$

$$= \int_0^1 \int_0^y \left(\frac{2 - (G_j(F^{-1}(y)) - G_j(F^{-1}(x)))}{\left(1 - (G_j(F^{-1}(y)) - G_j(F^{-1}(x)))\right)^2} \right)$$

To establish an exact equation for the *ARL* of the *2-of-2 KL* chart the derivation of the *ARL* in Klein (2000a) can be used. Thus, given $X_{a:m}$ and $X_{b:m}$ the conditional expected value of T_2' is given by the following equation: $E(T_2' | X_{a:m}, X_{b:m}) = \frac{1}{p - \frac{p_U}{1+p_U} - \frac{p_L}{1+p_L}}$ where p_U and p_L were defined earlier. Using this result the unconditional *ARL* is then established as:

$$ARL_{KL} = \int_0^1 \int_0^y \left[\frac{(1 + \left(1 - G_j\left(F^{-1}(y)\right)\right))(1 + G_j\left(F^{-1}(x)\right))}{(G_j\left(F^{-1}(x)\right))^2 (1 + \left(1 - G_j\left(F^{-1}(y)\right)\right) + \left(1 - G_j\left(F^{-1}(y)\right)\right)^2 (1 + G_j\left(F^{-1}(x)\right))} \right] * f_{a,b}(x, y) dx dy$$

The next section focuses on nonparametric charts for monitoring spread using runs-type signalling rules.

3.2 Runs-type signalling rules for monitoring spread

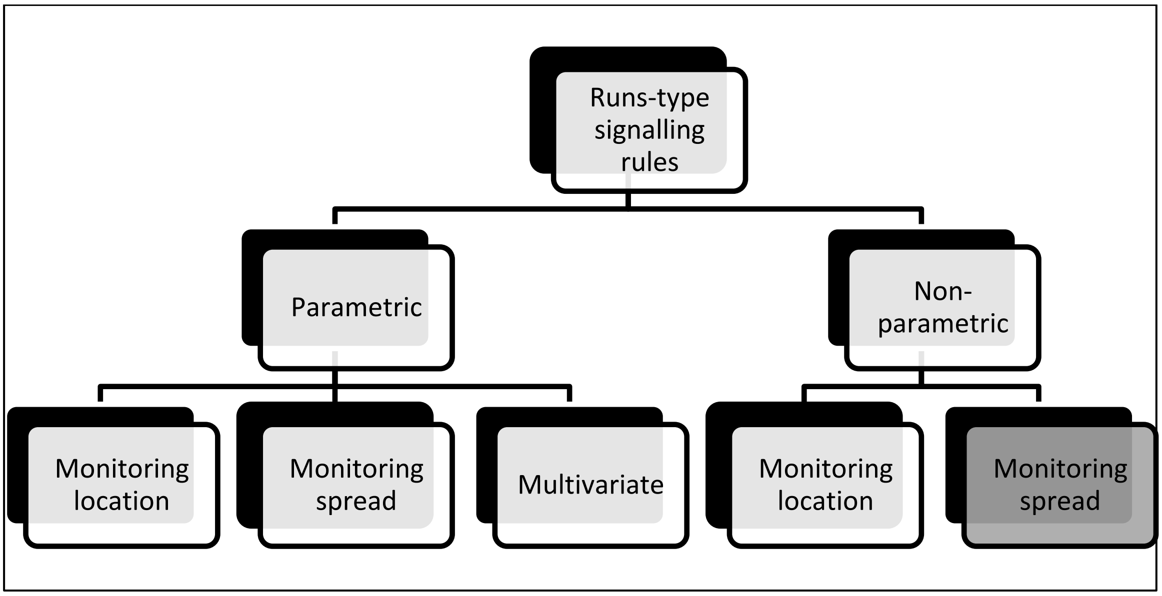

Figure 3.8: Classifications of the runs-type signalling rules control chart.

Most research in the nonparametric environment supplemented with runs-type signalling rules focus on detecting a shift in location, very few focus on spread. The following article by Das (2008b) proposes a nonparametric control chart for controlling the variability of a process.

3.2.1 Nonparametric control chart for controlling variability based on rank test (Das (2008b))

Lehman (1975), Bradley (1968) and Das (2008b) are a few of the authors that have done research in the nonparametric setting based on spread. Both Lehman and Bradley tested for the equality of variances using nonparametric tests. The approach followed by Bradley is based on the Westenberg's Two-Sample Interquartile Range Test. Das (2008b) proposed a nonparametric chart for monitoring variability based on the two sample variability test proposed by Ansari and Bradley (1960).

Das (2008b) proposed two nonparametric control charts based on spread using the *1-of-1* signalling rule. The Rank test is implemented using Mood and Siegel & Tukey's approach.

The assumption of normality being violated in traditional control charts generally lead to decreased ARL performance, see e.g. Shewhart (1986), Ferrel(1953) , Tukey (1960), Lagenberg and Iglewicz (1986), Jacobs (1990), Alloway and Raghavachari (1991), Yourstone and Zimmer (1992), Woodall and Montgomery (1999), Woodall (2000). Distribution free control charts, as the one proposed solves the dilemma since the underlying distribution is not of concern. The effects of non-normality have a bigger impact on control charts monitoring spread than on the control charts monitoring location. Wetherill and Brown (1991) highlighted the effect that the kurtosis of the original distribution has on the variance of S^2 and noted that the effect doesn't disappear with an increase in sample size.

The two proposed charts are based on the approach followed by Mood and Siegel & Tukey. Mood's test for dispersion is based on the following assumptions:

i. The observations considered are from two random populations.
ii. The probability distributions of the two populations are continuous.
iii. The two populations are independent.
iv. The two distributions differ only in spread.

The null and alternative hypotheses for the two-sided and one-sided charts would then be:

i. $H_0(two\ sided)\colon \sigma_x = \sigma_y, H_1\colon \sigma_x \neq \sigma_y$
ii. $H_0(upper\ one\ sided)\colon \sigma_x \geq \sigma_y, H_1\colon \sigma_x < \sigma_y$
iii. $H_0(lower\ one\ sided)\colon \sigma_x \leq \sigma_y, H_1\colon \sigma_x > \sigma_y$

The test statistic is denoted by $M = \sum_{i=1}^{n_1}(r_i - \frac{N+1}{2})^2$ where $N = n_1 + n_2$ and r_i is the rank of the i^{th} sample element (in population X, where X denotes one of the two populations) when the pooled sample is considered. Note that n_1 refers to the sample size of population 1 and n_2 to the sample size of population 2. The charting statistic for samples with a sample size greater than 30 would be $Z = \frac{M-E(M)}{\sqrt{V(M)}}$ with $E(M)$ the expected value of M and $V(M)$ the variance of M, where $Z \sim N(0,1)$. For sample sizes less than 30 Laubsher suggested the correction for continuity resulting in the following charting statistic: $Z = \frac{M-E(M)}{V(M)+\frac{1}{2V(M)}}$. The expected value and variance of the test statistic M is:

$E(M) = \frac{n_1}{12}(N-1)(N+1)$ and $V(M) = \frac{n_1 n_2}{180}(N+1)(N-2)(N+2)$

where n_1 refers to the sample size of population 1 and n_2 to the sample size of population 2.

Tukey's test is also based on a test for equality of variances, with the null hypothesis that variances are equal. The procedure involves aligning the two populations if there is a difference in location by subtracting or adding a constant value. Then the two populations are pooled and ranked. Rank 1 is assigned to the smallest value, rank 2 to the largest, 3 to the next largest, 4 and 5 to the next smallest, 6 and 7 to the next largest and so on. The population with the smallest sum of ranks are tested for significance; a smaller sum of ranks is associated with larger spread.

Let R_x denote the ranks of sample X. The charting statistic is calculated similarly to Mood's method by establishing the expected value and variance of R denoted by $E(R)$ and $V(R)$ and standardizing. The standardized value would then be $Z = \frac{2R_x - n_1(n_1+n_2+1) \pm 1}{\sqrt{(n_1(n_1+n_2+1)n_2/3}}$. It should be noted that the normal approximation can be used for a sample size larger than 20 and that there are no assumptions that need to be considered when applying this test.

Since the charting statistic involves standardized values for both tests, the control limits are: $UCL = 3$, $LCL = -3$ and $CL = 0$. An OOC signal would then be given if the charting statistic plots on or above the UCL, or on or below the LCL.

To measure the chart's performance Das (2008b) suggested evaluating the percentage of correct classifications given a particular shift. It is assumed that if the shift is large enough a signal should be produced and the process classified as OOC. The percentage correct classifications are directly related to a type II error (β), while the ARL is directly related to a type I and type II error. A type I error would be the probability for a chart to signal when the process is actually IC, while a type II error is the probability of detecting no signal when the process is OOC. The ARL is directly related to a type II error as follows: $ARL_\xi = {}^1/_{(1-\beta)} = {}^1/_{(Power\ of\ Test)}$. The second equation holds since the $Power\ of\ test = 1 - \beta$. A simulation study was conducted to calculate the ARL for the normal, uniform and Laplace distributions. Das (2008b) noted larger ARL_0 values when implementing Tukey's method while Mood's method provides better OOC performance up to a shift of $\lambda = 5$ where $\lambda = {}^{\sigma_1}/_{\sigma_0}$ for the normal and Laplace distributions when a sample size of 10 is considered;

note that σ_1 denotes the dispersion if a shift occured and σ_0 the dispersion if no shift occurs. For the uniform distribution Tukey's chart outperforms Mood's from a shift of $\lambda = 3$ to $\lambda = 7$ considering a sample size of 10. For sample sizes 15 and 20 Tukey's method generally outperforms Mood's when considering the ARL_{ξ}.

3.3 Conclusion of Chapter 3

Chapter 3 provides an overview of control charts supplemented with runs-type signalling rules in the nonparametric environment when considering either a shift in location or spread. The articles discussed in Chapter 3 are listed below.

Nonparametric control charts with runs-type signalling rules for monitoring location

i. A nonparametric Shewhart-type signed rank control chart based on runs (Chakraborti et al. (2007)).
ii. A distribution-free Shewhart quality control chart based on signed-ranks (Bakir (2004)).
iii. Nonparametric Shewhart-type sign control charts based on runs (Human et al. (2010)).
iv. A distribution free control chart based on order statistics (Balakrishnan et al. (2010)).
v. A phase II nonparametric control chart based on precedence statistics with runs-type signalling rules (Chakraborti et al. (2009)).

Nonparametric control charts with runs-type signalling rules for monitoring spread

ii. Nonparametric control chart for controlling variability based on rank test (Das (2008b)).

Bakir (2004) proposed a nonparametric control chart based on signed ranks incorporating the *1-of-1* runs-type signalling rule. Chakraborti and Eryilmaz (2007) proposed a nonparametric control chart based on signed ranks supplemented with the *k-of-k* and *k-of-w* runs-type signalling rules that offers smaller *FAR's* and larger $ARL_0's$. Human et al. (2010) proposed three charts using the sign test statistic; the *1-of-1*, the *k-of-k* and the *k-of-w*. Balakrishnan et al. (2010) proposed a control chart for phase II applications taking into account the location of the single order statistic (*1-of-1* runs-type signalling rule) from the test sample as well as the number of observations between the control limits. Chakraborti et al. (2009) proposed two control charts for monitoring the unknown location parameter of a continuous population in phase II applications. Das (2008b) proposed two nonparametric control charts based on spread using the *1-of-1* signalling rule.

Chapter 4 provides an overview of the research done in the multivariate environment.

Chapter 4

Multivariate charts: Runs-type signalling rules

4.1 Runs-type signalling rules for monitoring location and spread

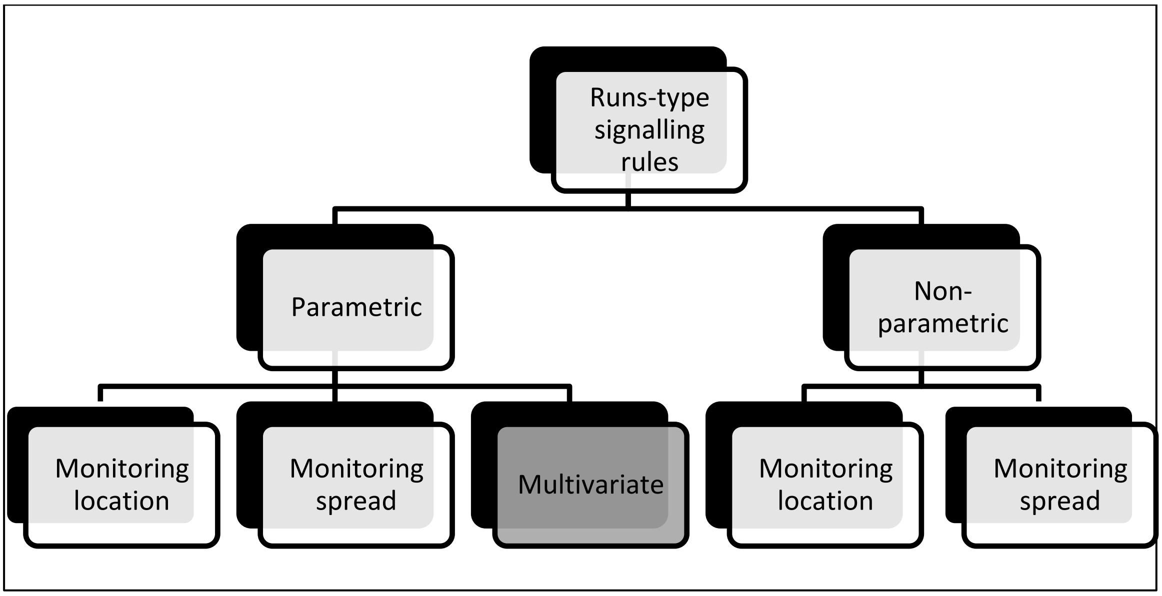

Figure 4.1: Classifications of the runs-type signalling rules control chart.

This Section provides a detailed review of the articles based on multivariate control charts supplemented with runs-type signalling rules. The most popular multivariate control procedure is the chi-square control chart. One major disadvantage of this control chart is that it uses only data of the most recently inspected sample and monitors large shifts in the process. In order to detect small shifts in the process the control charts are supplemented with runs-type signalling rules. The articles that follow provide insight into various runs-type signalling rules for multivariate control charts.

4.1.1 A performance analysis of Hotelling's χ^2 control chart with supplementary runs rules (Aparisi et al. (2004))

Champ (1986) suggested using separate $\bar{X}$ charts in order to monitor various quality characteristics simultaneously. The drawback of this approach is that as the quality characteristics become large so do the number of control charts used for monitoring the process. Multivariate control charts have been introduced for that reason, thus only one control chart is necessary to monitor various characteristics simultaneously.

Various runs-type signalling rules are considered: the *1-of-1*, the *2-of-3*, one considering eight consecutive points above the median and then another considering seven consecutive rising points. For the purpose of this study only the first three rules mentioned above will be considered since the last rule didn't provide good ARL performance. The charting statistic is the well known Mahalanobis distance denoted by $T_i^2 = n(\,\bar{\boldsymbol{X}}_{\boldsymbol{i}} - \mu_{\boldsymbol{0}})^{\boldsymbol{t}}\boldsymbol{\Sigma}^{-\boldsymbol{1}}(\bar{\boldsymbol{X}}_{\boldsymbol{i}} - \mu_{\boldsymbol{0}})$ where $i = 1,2 \ldots \ldots$ and $T_i^2 \sim \chi_p^2$; where n denotes the sample size, $\bar{\boldsymbol{X}}_{\boldsymbol{i}}$the mean vector of sample i, $\mu_{\boldsymbol{0}}$ the IC mean vector and Σ the variance covariance matrix. If a process shift has occured T_i^2 has a non-central chi-square distribution with p degrees of freedom. If the IC mean vector ($\mu_{\boldsymbol{0}}$) and the variance covariance matrix (Σ) is unknown they can be estimated by the vector of averages of the sample means vector ($\bar{\bar{X}}$) and the average of the average of the sample of covariance matrices ($\bar{S}$). The regions considered for the control chart are provided in Figure 4.2.

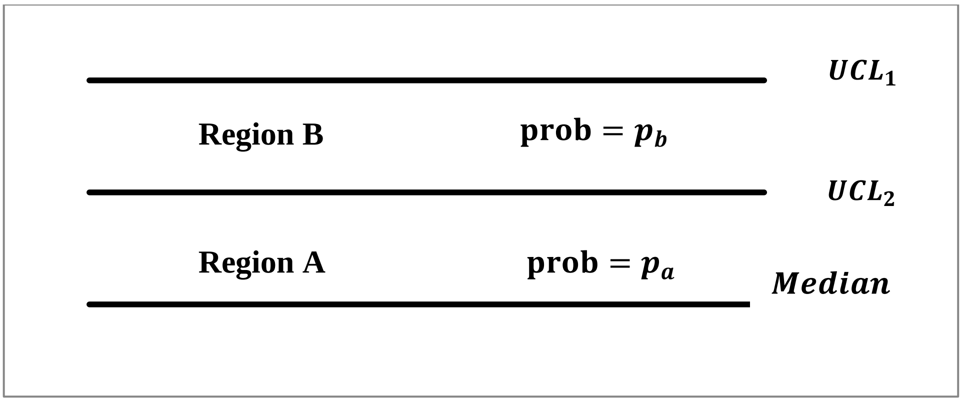

Figure 4.2: Charting regions considered for the χ^2 control chart supplemented with runs-type signalling rules.

For the *1-of-1* runs-type signalling rule a signal is produced if a single charting statistic plots on or above UCL_1. The *2-of-3* runs-type signalling rule produces a signal if two out of the last three charting statistics plot between UCL_2 and UCL_1, while for the last rule a signal is produced if eight consecutive charting statistics plot above the median. Aparisi et al. (2004**)** denotes the UCL_2 as a warning limit but for the purposes of this study it is seen as a secondary upper control limit. Note that UCL_2 refers to the inner upper control limit and UCL_1 to the outer upper control limit when considering this article. In the rest of the thesis UCL_2 refers to the outer upper control limit and UCL_1 to the inner upper control limit.

The control limits were calculated similarly to the previous articles discussed, by first establishing the outer upper control limit (UCL_1) and then calculating the inner upper control limit (UCL_2) for a given ARL_0 of 200. The outer upper control limit (UCL_1) was calculated as the 99.5th percentile of a Chi-square distribution with p degrees of freedom.

In order to calculate the ARL Aparisi et al. (2004**)** implemented a simulation approach. An alternative approach could be to implement the Markov Chain approach as depicted in Section 1.7 where the essential TPM is set up and the ARL calculated by using Equation (1.4). As an example the essential TPM for the χ^2 control chart supplemented with the *2-of-3* runs-type signalling rule is provided in Table 4.1. S_A refers to the state entered if the charting statistic plots below the UCL_2 while S_B refers to the state entered if the charting statistic plots between the UCL_1 and the UCL_2. The probabilities p_a and p_b refer to the probabilities associated with S_A and S_B, where S_A and S_B refer to the states entered if the charting statistic plots in Region A and B respectively.

Table 4.1: Essential TPM of a multivariate control chart supplemented with the *1-of-1* runs-type signalling rules combined with the *2-of-3*.

State	S_A	S_B	S_{AB}	S_{BA}
S_A	p_a	0	p_b	0
S_B	0	p_b	0	p_a
S_{AB}	0	0	0	p_a
S_{BA}	p_a	0	0	0

Note that Aparisi et al. (2004**)** compared χ^2 control charts with and without runs-type signalling rules for various shifts in the process mean. It was seen that the charts with runs-type signalling rules outperform their counterparts without the additional rules for shifts ranging up to 3σ. Aparisi et al. (2004) suggests implementing the *1-of-1* rule to monitor large shifts, the *2-of-3* rule for moderate shifts and the rule stating that a signal is produced if 8 charting statistics plot above the median to monitor small shifts in the process mean.

4.1.2 Improving the performance of the Chi-square control chart via runs rules (Koutras et al. (2006))

Aparisi et al. (2004) investigated a chi-square control chart supplemented with runs-type signalling rules. Koutras et al. (2006) follows this lead and proposes three chi-square control charts supplemented with runs-type signalling rules; the first is supplemented with the *1-of-1* runs-type signalling rule, the second with the *k-of-k* and the third applies a combination of the *k-of-k* and *r-of-r* runs-type signalling rules.

Multivariate control charts involve monitoring various quality characteristics simultaneously in a phase II process. The quality characteristics are denoted by $x_1, x_2, x_3, \dots x_m$. It is assumed that the IC joint probability distribution of the vector $x = (x_1, x_2, x_3, \dots, x_m) \sim N_m(\mu_{\mathbf{0}}, \boldsymbol{\Sigma}_{\mathbf{0}})$ where N_m denotes a m-variate normal distribution with mean vector $\mu_{\mathbf{0}}$ and variance-covariance matrix $\boldsymbol{\Sigma}_{\mathbf{0}}$.

The charting statistic considered is $D_i^2 = n(\overline{\boldsymbol{x}_i} - \mu_{\mathbf{0}})^{\boldsymbol{t}} \boldsymbol{\Sigma}_{\mathbf{0}}^{-\mathbf{1}} (\overline{\boldsymbol{x}_i} - \mu_{\mathbf{0}})$ where D_i^2 represents the weighted distance (Mahalanobis) between $\overline{\boldsymbol{x}_i}$ and $\mu_{\mathbf{0}}$ ($\overline{\boldsymbol{x}_i}$ denotes the mean vector of sample i). If the process is IC $D_i^2 \sim \chi_m^2$ while on the other hand if the process is OOC then $D_i^2 \sim \chi_m^2(\lambda)$, where $\chi_m^2(\lambda)$ denotes a non-central Chi-square distribution with m degrees of freedom. The non-centrality parameter is equal to $\lambda = \lambda(\mu_1) = n(\mu_{\mathbf{1}} - \mu_{\mathbf{0}})^t \Sigma_0^{-1} (\mu_{\mathbf{1}} - \mu_{\mathbf{0}}) = n\xi^t \Sigma_0^{-1} \xi$. If a shift occurred then the mean vector would be $E(x) = \mu_1 = \mu_0 + \xi$ where μ_0 indicates the IC mean vector.

The upper control limit then equals $\chi_{m,a}^2$ where $P\left(D_i^2 > \chi_{m,a}^2\right) = a$. It should be noted that no lower control limit is considered since the test statistic considers the distance between the IC mean vector and the i^{th} sample vector. A large distance would indicate an OOC condition while a small distance would indicate an IC condition; thus no lower control limit is present.

To establish the control limits for the *k-of-k* rule the unique root of $\frac{1-P^k}{p^k(1-p^k)} = ARL_0$ should be calculated where $p = P(W_i > UCL)$ and W_i refers to the charting statistic, see Klein (2000a) and Balakrishnan et al. (2002). The *1-of-1* runs-type signalling rule would signal if the charting statistic plots on or outside the UCL; the *k-of-k* runs-type signalling rule signals if k of the last k charting statistics plot on or outside the UCL; and the combination of the *r-of-r* and *k-of-k* runs-type signalling rules signal if either k out of k charting statistics or r out of r plot on or outside UCL_k or UCL_r respectively; where the UCL_k is the UCL associated with the *k-of-k* rule. The reader is referred to Figure 4.3 and Figure 4.4 for a graphical representation of the control limits. If a single characteristic is considered with both $\boldsymbol{\mu_0}$ and Σ_0 unknown then the control chart reduces to Hotelling's T^2 control chart. The reader is refered to Alt and Smith (1988) or Lowry and Montgomery (1995).

Since the run length distribution follows a geometric distribution, the ARL_0 for the chi-square chart supplemented with the *1-of-1* signalling rule reduces to: $ARL_0 = \frac{1}{P(D_i^2 > UCL)} = \frac{1}{1-F_m(UCL)}$. The $ARL_\xi = \frac{1}{P(D_i^2(\lambda) > UCL)} = \frac{1}{1-F_m(UCL,\lambda)}$. For the *k-of-k* chart the waiting time until the first signal occurs (ARL) equals: $E(T) = \frac{1-P^k}{p^k(1-p^k)}$ where $p = P(W_i > UCL)$ and W_i refers to the charting statistic. The reader is referred to Balakrishnan et al. (2002) and Fu and Lou (2003) for the derivation of $E(T)$. The OOC ARL then reduces to $ARL_\xi = \frac{1-(P^k(\lambda))}{p^k(\lambda)(1-p^k(\lambda))} = \frac{1-(1-F_m(UCL,\lambda))^k}{(1-F_m(UCL,\lambda))^k F_m(UCL,\lambda)}$.

An alternative approach to calculate the ARL would be to derive the essential TPM and apply Equation (1.4) as previously. For the *1-of-1* rule only 1 transient state needs to be considered; thus the matrix reduces to a scalar value which equals the probability of plotting IC. When considering the *3-of-3* runs-type signalling rule the transient states would be $S_{0,}$ S_1, S_{01}, S_{10} and S_{11}; where S_{10} indicates the first charting statistic plotting in Region 1 and the second in Region 0. Figure 4.3 illustrates the regions graphically and Table 4.2 provides the essential TPM for the *k-of-k* runs-type signalling rule, where $k = 3$.

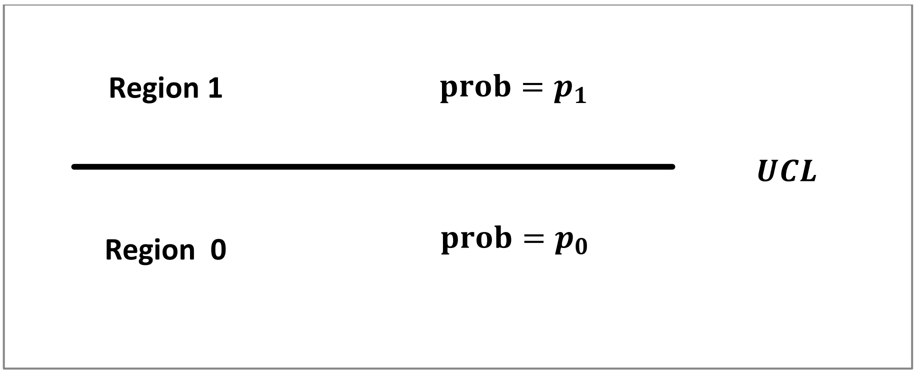

Figure 4.3: Charting probabilities and regions considered for the χ^2 control chart supplemented with the *k-of-k* runs-type signalling rules.

Table 4.2: Essential TPM of a multivariate control chart supplemented with the *3-of-3* runs-type signalling rule.

State	S_0	S_1	S_{01}	S_{10}	S_{11}
S_0	p_0	0	p_1	0	0
S_1	0	0	0	p_0	p_1
S_{01}	0	0	0	p_0	p_1
S_{10}	p_0	0	p_1	0	0
S_{11}	0	0	0	p_0	0

The third proposed chart implements a combination of the *r-of-r* and *k-of-k* runs-type signalling rules, but two separate upper control limits are considered as shown in Figure 4.4. A signal would be produced if either r out of the last r charting statistics plot outside the UCL_r or if k out of the last k charting statistics plot outside UCL_k. The associated probabilities are p_k and p_r where $p_k = P(D_i^2 > UCL_k)$ and $p_r = P(D_i^2 > UCL_r)$. The $ARL = \left(p_1 + p_2 - \frac{(p_1 - p_k)}{(1-p_k)}\right)^{-1}$.

The reader is referred to Page (1955), Aki (1992), Antzoulakos (2001) and Fu and Chang (2003) with regards to the derivation of the ARL above when considering a combination of different runs-type signalling rules on the same control chart.

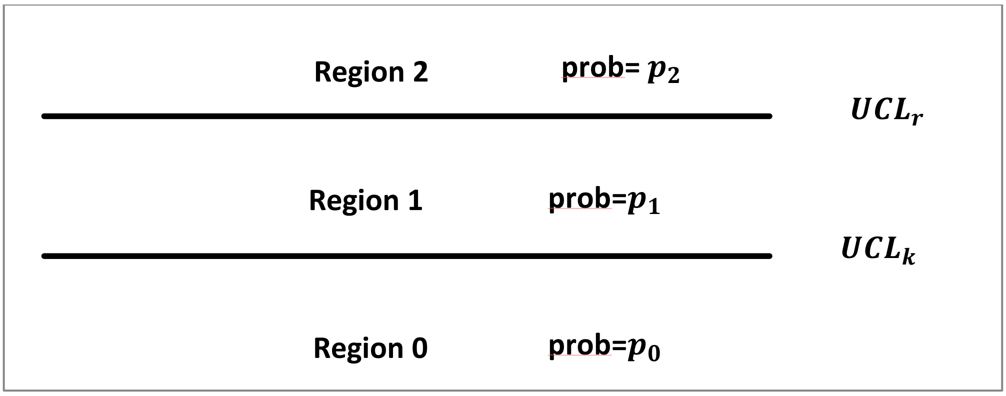

Figure 4.4: Charting probabilities and regions considered for the chi-square control chart with the combined *k-of-k* and *r-of-r* runs-type signalling rules.

Overall the chi-square control chart supplemented the *1-of-1* and *k-of-k* runs-type-signalling rules provide the best ARL performance. Another advantage of the proposed charts is it's robustness if the assumption of normality is not satisfied. In the following article Lee proposes various Hotelling's T^2 control charts supplemented with runs-type signalling rules as well as implementing variable sample sizes and intervals.

4.1.3 Adaptive Hotelling's T^2 control charts with run rules (Lee (2013))

Various research have been done to improve hotelling's T^2 chart, see Champ and Aparisi (2008), Aparisi and de Luna (2009) and Khoo et al. (2009). The proposed charts implement Hotelling's T^2 statistic supplemented by runs-type signalling rules (*1-of-1*, *2-of-3* and *4-of-5*) and implements variable sample size (VSS), variable sample intervals (VSI) or both ($VSSI$). These runs-type signalling rules are combined, thus increasing the *FAR* but resulting in overall better signal detection capability.

The charting statistic is similar to the previous article and is denoted by $T_i^2 = n(\bar{\boldsymbol{x}}_i - \mu_0)^{\boldsymbol{t}} \boldsymbol{\Sigma}_{\boldsymbol{0}}^{-\boldsymbol{1}} (\bar{\boldsymbol{x}}_i - \mu_0)$ with $i = 1,2,3..$ and μ_0 being the IC mean vector, $\boldsymbol{\Sigma}_{\boldsymbol{0}}$ the variance-covariance matrix and $\bar{\boldsymbol{x}}_{\boldsymbol{i}}$ the mean vector of sample *i*. If the process is IC $T_i^2 \sim \chi_m^2$ while if the process is OOC $D_i^2 \sim \chi_m^2(\lambda)$, where λ is the non-centrality parameter, see Koutras et al. (2006).

The regions considered in the control chart are provided below in Figure 4.5 along with the associated probabilities, where p_a denotes the probability of plotting in Region A denoted here as zone A. These zones are similar to regions discussed elsewhere in the thesis. A signal is produced if the charting statistic plots in Zone O, 2 of the last 3 charting statistics plot in Zone A and lastly 4 of the last 5 charting statistics plotting in either Zone A or B. The probabilities of plotting in the regions provided below are: $p_a = P(UCL_1 < \chi^2_{p,d} \leq UCL)$, $p_b = P(UCL_2 < \chi^2_{p,d} \leq UCL_1)$, $p_c = P(\chi^2_{p,d} \leq UCL_2)$ where $d = n_i \xi^2$ and p refers to the degrees of freedom of the chi-square variable.

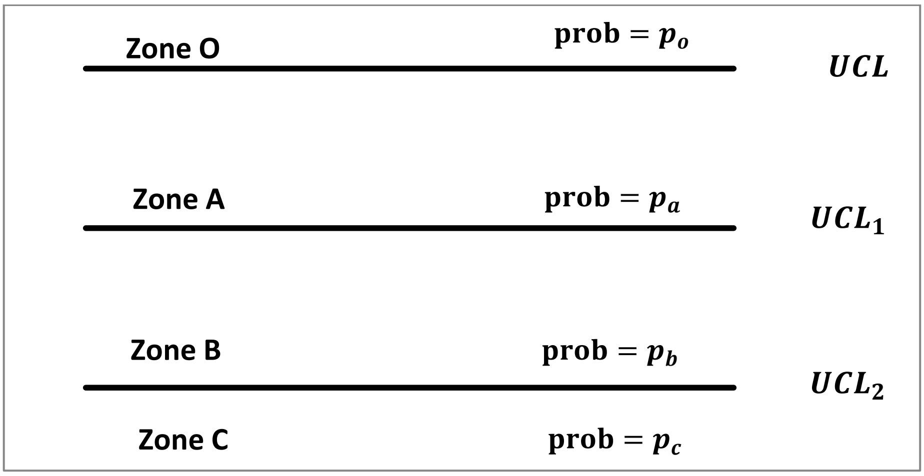

Figure 4.5: Charting zones considered for the adaptive Hotelling's T^2 charts.

All three runs-type signalling rules are considered simultaneously in order to detect an OOC condition. The three performance measures implemented by Lee (2013) for an IC process are the average time to signal (ATS_0), the in-control average number of observations to signal ($ANOS_0$) and the average number of samples to signal ($ANSS_0$). The equations for these performance measures are: $ATS_0 = \boldsymbol{r}'(\boldsymbol{I} - \boldsymbol{Q_0})^{-1}\mathbf{h}$, $ANOS_0 = r'(\boldsymbol{I} - \boldsymbol{Q_0})^{-1}\mathbf{n}$ and $ANSS_0 = r'(\boldsymbol{I} - \boldsymbol{Q_0})^{-1}\mathbf{1}$ where $\boldsymbol{r}$ denotes the initial probability vector, $\boldsymbol{I}$ the identity matrix, $\boldsymbol{Q_0}$ the IC essential TPM and $\mathbf{1}$ indicates a column vector all elements equal to one; $\boldsymbol{h}$ denotes the vector containing the sampling interval lengths and n the vector of sample sizes for the different points in time that the sample is taken, see Lee (2013). The performance measure to monitor an OOC process is the steady state average time to signal ($AATS$) and is given by: $AATS = \boldsymbol{b}'((\boldsymbol{I} - \boldsymbol{Q_1})^{-1}\boldsymbol{h}) - \frac{1}{2}\boldsymbol{h}$ where $\boldsymbol{Q_1}$ indicates the OOC essential TPM and $\boldsymbol{b}'$ the row vector of modified steady-state probabilities given by $b_i = \frac{e_i h_i}{\boldsymbol{e}'\boldsymbol{h}}$ where $\boldsymbol{e}'$ refers to the vector of steady-state probabilities.

An alternative performance measure could be the ARL where the essential TPM is established and Equation (1.4) applied. For practical purposes Zone B and C will be grouped together as Zone D when considering the transient states for the *2-of-3* runs-type signalling rule. Likewise Zone A and B will be grouped together as Zone E when considering the *4-of-5* runs-type signalling rule. The transient states for the *2-of-3* and *4-of-5* runs-type signalling rules are $A, D, AD, DA, C, E, EE, EC, CE, CC, EEE, EEC, ECE, ECC, CEE, CEC, CCE, CCC, EEEC, EECE, EECC, ECEE, ECEC, ECCE, ECCC, CEEE, CEEC, CECE, CECC, CCEE, CCEC, CCCE, CCCC$. The essential TPM for the multivariate control chart supplemented with the *2-of-3* runs-type signalling rule is provided in Table 4.3. Note that S_d and S_a refer to the states entered if the charting statistic plots in Zone D and Zone A, see Figure 4.4; where Zone D refers to the combination of Zone B and C.

Table 4.3: Essential TPM of a multivariate control chart supplemented with the *2-of-3* runs-type signalling rule.

State	$\boldsymbol{S_d}$	$\boldsymbol{S_a}$	$\boldsymbol{S_{da}}$	$\boldsymbol{S_{ad}}$
$\boldsymbol{S_d}$	p_d	0	p_a	0
$\boldsymbol{S_a}$	0	0	0	p_d
$\boldsymbol{S_{da}}$	0	0	0	p_d
$\boldsymbol{S_{ad}}$	p_d	0	0	0

Combining these runs-type signalling rules results in a greater *FAR*, but adjusting the control limits reduces the negative effect. Variable sample size and sampling intervals improve the charts performance since samples are taken at shorter intervals if the process seems OOC. When considering two quality characteristics the charts implementing *VSS* outperform the charts with *VSI* for shifts ranging between 0.25 and 1; for shifts between 1.5 and 3 it was seen that the charts implementing VSI outperform the *VSS* charts. For shifts between 0.25 and 0.5 the control chart supplemented with the *1-of-1* and *3-of-4* runs-type signalling rules as well as implementing *VSS* provided the best performance. The control chart supplemented with the *1-of-1* and *3-of-4* runs-type signalling rules also performed best for a shift of 1.5 when implementing *VSSI*. For shifts between 2 and 2.5 the control chart supplemented with the *1-of-1* and *2-of-3* runs-type signalling rules implementing *VSI* provided the best ARL performance. Khoo and Quah (2003) propose Hotelling's control chart supplemented with additional runs-type signalling rules in the following article.

4.1.4 Incorporating runs rules into Hotelling's χ^2 control chart (Khoo and Quah (2003))

This article is one of the pioneers when considering multivariate control charts supplemented with runs-type signalling rules. Khoo and Quah (2003) proposed Hotelling's control chart supplemented with the *1-of-1*, *2-of-2*, *2-of-3* and *2-of-4* runs-type signalling rules. These charts deliver superior performance for small shifts in the process mean compared to the standard Hotelling's χ^2 control chart.

The signalling criteria for the proposed charts are as follows: for the *1-of-1* chart a signal is produced if the charting statistic plots on or above the UCL; the *2-of-2* chart produce a signal if two charting statistics out of the last two plot on or above the UCL; the *2-of-3* chart if two charting statistics out of the last three plot on or above the UCL and lastly the *2-of-4* chart if two charting statistics out of the last four plot on or above the UCL. A graphical illustration of the control limits and probabilities are provided in Figure 4.6 where State A is entered if the charting statistic plots in Region A.

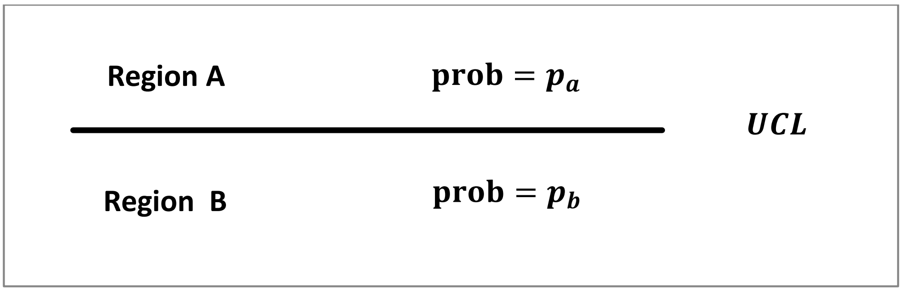

Figure 4.6: Charting probabilities and regions considered for the χ^2 control chart supplemented with runs-type signalling rules.

It is assumed that observations are distributed iid multivariate normally with $X_i \sim N_p(\mu, \boldsymbol{\Sigma})$ where $i = 1,2, \dots n$ and μ denotes the mean vector, $\boldsymbol{\Sigma}$ the variance-covariance matrix and p the number of quality characteristics monitored simultaneously. The test statistic equals $T_n^2 = (\boldsymbol{X_n} - \mu_0)^t \boldsymbol{\Sigma_0}^{-1} (\boldsymbol{X_n} - \mu_0)$ where $T_n^2 \sim \chi_p^2$.

To obtain the charting statistic a transformation is applied to T_n^2. Khoo and Quah (2003) noted that in general if X represents a random variable then the transformation $F(X)$ has a uniform distribution on the interval $(0,1)$. Now if U represents a random variable on the above mentioned interval then $\phi^{-1}(U) \sim N(0,1)$ where $\phi^{-1}(.)$ represents the inverse of the standard normal cdf function. The following transformation is applied to T_n^2: $Y_n = \phi^{-1}(H_p(T_n^2))$ where $H_p(.)$ refers to the chi-square distribution function with p degrees of freedom. Tracy et al. (1992) derived an expression for T_n^2 in the case where μ and $\boldsymbol{\Sigma}$ are unknown: $T_f^2 = \left(\boldsymbol{X_f} - \boldsymbol{\bar{X}_m}\right)^{\boldsymbol{t}} \boldsymbol{S_m}^{-1}\left(\boldsymbol{X_f} - \boldsymbol{\bar{X}_m}\right)$ where $f = 1,2,3, \ldots$ and m refers to the reference sample size. Khoo and Quah (2003) derived the exact distribution of T_f^2 as $T_f^2 \sim \frac{p(m-1)(m+1)}{m(m-p)} F_{p,m-p}$ where p refers to the number of quality characteristics being studied. The transformation $Y_f = \phi^{-1}(F_{p,m-p}\left(\frac{m(m-p)}{p(m-1)(m+1)}\right) T_f^2)$ can then be applied resulting in iid $N(0,1)$ random variables.

In order to measure the ARL performance the noncentrality parameter λ has to be established where $\lambda = n(\boldsymbol{\mu_1} - \boldsymbol{\mu_0})^t \Sigma_0^{-1}(\boldsymbol{\mu_1} - \boldsymbol{\mu_0})$ as mentioned in previous articles; $\boldsymbol{\mu_1}$ refers to the shifted mean vector, $\boldsymbol{\mu_0}$ the IC mean vector and $\boldsymbol{\Sigma_0}$ the variance covariance matrix.

The expected number of transitions from a particular state to the absorbent state, in other words the ARL, can be found by solving the following first passage time (waiting time) equations when considering the *2-of-2* runs-type signalling rule. $N_{AC} = 1 + p_b N_{AC} + p_a N_{BC}$, $N_{BC} = 1 + p_b N_{AC}$ where p_a and p_b is the probability of the charting statistic plotting above or below the UCL as shown in Figure 4.4; N_{AC} is the expected number of transitions from State A to State C, where State A refers to the charting statistic plotting in Region A, State B to the charting statistic plotting in Region B and State C the absorbent state being entered if a signal is produced. The ARL_0 is then equal to N_{AC}.

An alternative way to find the ARL would be to set up the essential TPM and apply Equation (1.4). The essential TPM for the *2-of-2* and *2-of-3* runs-type signalling rules are provided in Table 4.4 and Table 4.5. S_A refers to State A, S_B to State B and S_{AB} to the charting statistic first plotting in Region A followed by the next plotting in Region B.

Table 4.4: Essential TPM for the χ^2 Control Chart supplemented with the *2-of-2* runs-type signalling rule.

State	S_A	S_B
S_A	0	p_b
S_B	p_a	p_b

Table 4.5: Essential TPM for the χ^2 Control Chart supplemented with the *2-of-3* runs-type signalling rule.

State	S_A	S_B	S_{AB}	S_{BA}
S_A	0	0	p_b	0
S_B	0	p_b	0	p_a
S_{AB}	0	p_b	0	0
S_{BA}	0	0	p_b	0

To conclude, the control chart supplemented with the *2-of-4* runs-type signalling rules outperforms the rest for shifts up to $\xi = 3$. In the next article Khoo proposes Hotelling's χ^2 control chart supplemented with runs-type signalling rules in detecting large and small shifts in the process mean simultanously.

4.1.5 Powerful rules for Hotelling's χ^2 control chart (Khoo et al. (2005))

Runs-type signalling rules typically require several consecutive points in order to detect a small process shift. This leads to a delay in order to identify an OOC process. To monitor both small and large shifts simultaneously Khoo et al. (2005) proposed a control chart based on Hotelling's χ^2 control chart supplemented with combined runs-type signalling rules.

A phase I analysis as well as phase II analysis is performed for the proposed chart. It is assumed that the observations are iid multivariate normally distributed with $X_i \sim N_p(\mu, \boldsymbol{\Sigma})$ where $i = 1,2,\ldots n$ and μ defined as the mean vector, $\boldsymbol{\Sigma}$ the variance-covariance matrix and p the number of quality characteristics monitored simultaneously. The charting statistic is defined as $T_n^2 = (\boldsymbol{x_n} - \mu_0)^{\boldsymbol{t}} \boldsymbol{\Sigma}^{-1} (\boldsymbol{x_n} - \mu_0)$ where $n = 1,2 \ldots \ldots$ and μ_0 denotes the IC mean vector; then $T_n^2 \sim \chi_p^2$. If the mean vector and the variance-covariance matrix are unknown they can be estimated by $\bar{X}_m$ and S_m respectively, with m refering to the reference sample size. The charting statistics with the estimated parameters are $T_i^2 = (\boldsymbol{X_i} - \bar{\boldsymbol{X}}_{\boldsymbol{m}})^{\boldsymbol{t}} \boldsymbol{S}_{\boldsymbol{m}}^{-1} (\boldsymbol{X_i} - \bar{\boldsymbol{X}}_{\boldsymbol{m}})$ where $i = 1,2 \ldots \ldots$m. Tracy et al. (1992) noted that if single observations are considered

$(n = 1)$ then the control limits of the phase I process should be based on a beta distribution with the UCL equal to $UCL = \frac{(m-1)^2}{m} B_{\alpha,p/2,(m-p-1)/2}$ where the parameters of the beta distribution are $\frac{p}{2}$ and $(m-p-1)/2$ and $B_{\alpha,p/2,(m-p-1)/2}$ refers to the uppper α percentage point.

The charting statistic for a phase II process equals $T_f^2 = \left(\boldsymbol{X_f} - \boldsymbol{\bar{X}_m}\right)^{\boldsymbol{t}} \boldsymbol{S_m^{-1}}\left(\boldsymbol{X_f} - \boldsymbol{\bar{X}_m}\right)$ where $f = 1,2 \dots, \bar{X}_m$ and S_m are both obtained from the phase I reference sample. The UCL for the phase II analises is then $UCL = \frac{p(m+1)(m-1)}{m(m-p)} F_{\alpha,p,m-p}$ where $F_{\alpha,p,m-p}$ refers to the F distribution with parameters p and $m-p$. For the combined runs-type signaling rules the charting statistics $T_n^2, T_i^2\ T_f^2$ is transformed into standard normal variables as follows: $Y_n = \phi^{-1}(H_p(T_n^2))$, $Y_i = \phi^{-1}(B_{\frac{p}{2},\frac{m-p-1}{2}}\left(\frac{m}{(m-1)^2} T_i^2\right))$, $Y_f = \phi^{-1}(F_{p,m-p}\left(\frac{m(m-p)}{p(m+1)(m-1)} T_f^2\right))$ where $H_p(.)$ refers to the chi-square distribution function, $\phi^{-1}(.)$ to the inverse of the standard normal distribution function and $n = 1,2,..,\ f = 1,2,...$ and $i = 1,2,..m$.

Three runs-type signalling rules are proposed; the *1-of-1* combined with the *2-of-2*, the *1-of-1* combined with the *2-of-3* and the *1-of-1* in combination with the *2-of-4*. The signalling criteria for the three rules are to signal if a charting statistic plots on or above the outer upper (UCL_1) control limit, or if 2 of the last 2(3)(4) charting statistics (*2-of-2(3)(4)* chart) plot within the outer and inner upper control limit (UCL_2) , thus between the UCL_1 and UCL_2. A graphical representation of the control limits and associated probabilities are provided in Figure 4.7. Note that UCL_1 refers to the outer upper control limit and UCL_2 to the inner upper control limit. In the rest of the thesis UCL_1 refers to the inner upper control limit and UCL_2 to the outer upper control limit.

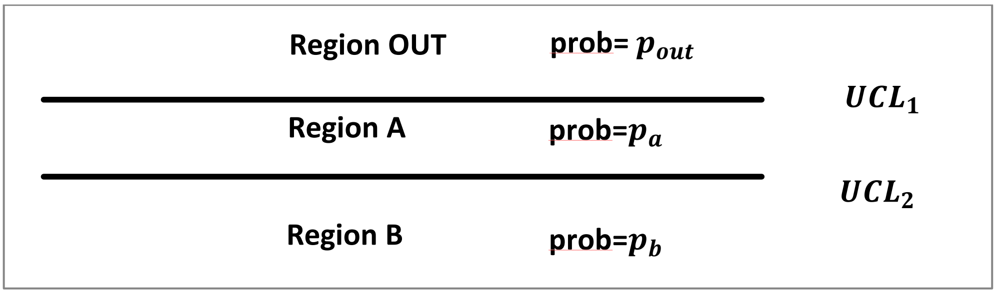

Figure 4.7: Charting probabilities and regions considered for the χ^2 control chart supplemented with runs-type signalling rules.

In order to calculate the UCL_1 and UCL_2, the value of the UCL_1 is fixed and the UCL_2 calculated for a specified ARL_0. The expected number of transitions from a certain transient state to the absorbent state for the combined *1-of-1* and *2-of-2* runs-type signalling rules are calculated by solving the following equations known as the first passage time or waiting time equations: $N_{13} = 1 + p_b N_{13} + p_a N_{23}$, $N_{23} = 1 + p_b N_{13}$ and $p_a + p_b + p_{out} = 1$ where $N_{13} = ARL_0$. Khoo et al. (2005) solved the equations in order to obtain the ARL_0: $ARL_0 = \frac{(1+p_a)}{(p_a^2+p_{out}+p_a p_{out})}$. When designing the control chart the UCL_1 is fixed implying that p_{out} is known, then p_a can be calculated from the ARL_0 equation. The transition equations relating to the combined *1-of-1* and *2-of-3* runs-type signalling rules are: $N_{14} = 1 + p_b N_{14} + p_a N_{24}$, $N_{24} = 1 + p_b N_{34}$, $N_{34} = 1 + p_b N_{14}$ with additional constraint $p_a + p_b + p_{out} = 1$.

An alternative approach to calculate the ARL would be to set up the essential TPM and apply Equation (1.4). The essential TPM's for two of the combined runs-type signalling rules are provided below in Tables 4.6 and 4.7. S_A and S_B refers to State A and B respectively, where S_A is entered if the charting statistic plots in Region A with probability p_a.

Table 4.6: Essential TPM of a multivariate control chart using the *1-of-1* runs-type signalling rules combined with the *2-of-2*.

State	S_A	S_B
S_A	0	p_b
S_B	p_a	p_b

The transient states considered for the combined *1-of-1* and *2-of-3* runs-type signalling rules are: S_A, S_B, S_{AB}, S_{BA} where S_{AB} refers to the state with the first charting statistic plotting in Region A and the second in Region B. The essential TPM for the multivariate control chart supplemented with the *1-of-1* and *2-of-3* runs-type signalling rules are provided in Table 4.7.

Table 4.7: Essential TPM of a multivariate control chart using the *1-of-1* runs-type signalling rules combined with the *2-of-3*.

State	S_A	S_B	S_{AB}	S_{BA}
S_B	0	p_b	0	p_a
S_A	p_a	0	p_b	0
S_{AB}	0	p_b	0	0
S_{BA}	0	p_b	0	0

The χ^2 control chart supplemented with runs-type signalling rules provide superior performance in detecting a combination of large and small shifts in the process. The three proposed charts were compared to the standard *1-of-1* and *2-of-2* charts. For shifts (λ) of 0.4 and 3 the proposed charts outperform these standard charts. For a shift where $0.4 \leq \lambda \leq 2.6$, the combined *1-of-1* and *2-of-4* chart provided the best ARL performance. The combined *1-of-1* and *2-of-3* provided superior performance only for a shift in the range $2.8 \leq \lambda \leq 3$.

4.1.6 Alternatives to the multivariate control chart for process dispersion (Khoo and Quah (2004))

Most of the articles studied in the multivariate field of SPC focus on a shift in the mean of the process based on a number of quality characteristics. This article focuses on detecting a shift in process dispersion while incorporating runs-type signalling rules. Khoo and Quah (2004) proposed six control charts incorporating runs-type signalling rules in order to detect small shifts in the process quicker. The proposed control charts implement the following runs-type signalling rules, the *1-of-1*, *2-of-2*, *2-of-3*, *2-of-4*, *3-of-3* and *3-of-4*.

Various approaches exist in monitoring the variability of a multivariate process, see e.g. Alt (1985) or Alt and Smith (1988). Another approach is based on the sample generalized variance denoted by $|\boldsymbol{S}|$. The methodology used is based on the fact that most of the probability distribution of $|\boldsymbol{S}|$ falls in the interval $E(|\boldsymbol{S}|) \pm 3\sqrt{Var(|\boldsymbol{S}|)}$.

The charting statistic implemented is also based on the generalized variance $|S|$. The following distributional properties and transformation is used in deriving the charting statistic: $2(n-1)|\boldsymbol{S}|^{\frac{1}{2}}/|\Sigma_0|^{1/2} \sim \chi^2_{2n-4}$ and $\phi^{-1}\left[G_{2n-4}\left\{2(n-1)|\boldsymbol{S_i}|^{\frac{1}{2}}/|\Sigma_0|^{1/2}\right\}\right]$ where $i = 1,2,3 \ldots \ldots$ and G_{2n-4} denotes the chi-square distribution function with $2n-4$ degrees of freedom, $\phi^{-1}(.)$ the inverse of the standard normal distribution function, $|\Sigma_0|$ the covariance marix of constants and $|\boldsymbol{S_i}|$ the generalized sample covariance matrix. The charting statistic is denoted by $C_i = \phi^{-1}\left[G_{2n-4}\left\{2(n-1)|\boldsymbol{S_i}|^{\frac{1}{2}}/|\Sigma_0|^{1/2}\right\}\right]$ whereby the $C_i{}'s$ are a sequence of iid standard normal variables with i denoting the subgroup number, see Khoo et al. (2003) and Khoo et al. (2005). The control limits, regions and probabilities considered in designing the chart is provided in Figure 4.8.

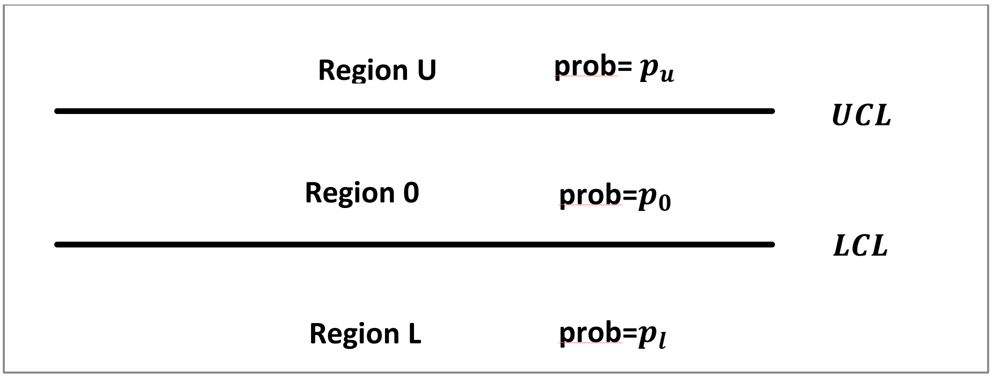

Figure 4.8: Charting probabilities and regions considered for the multivariate control chart for process dispersion.

A signal is produced if a sequence of charting statistics plots either on or above the UCL, or on or below the LCL. For the *1-of-1* the chart signals if a single charting statistic plots on or above the UCL or on or below the LCL. The *2-of-2* runs-type signalling rule produces a signal if two of the last two charting statistics plot on or above or on or below the control limits (UCL/LCL). Similarly the *2-of-3* chart signals if two of the last three charting statistics plot on or above, or on or below the control limits (UCL/LCL). The same line of thought can be applied in deriving the signalling conditions for the *2-of-4*, the *3-of-3* and the *3-of-4* charts.

The control limits can be established by using standard normal tables or a software package, e.g. Excel. Note that $P(C_i > UCL) + P(C_i < LCL) = \alpha = p_u + p_l$. Due to the symmetrical properties of the normal distribution $2P(C_i > UCL) = 2P(C_i < LCL) = \alpha = 2p_u = 2p_l$. Klein (2000a) discussed a Markov chain approach in obtaining the control limits for the *2-of-2* and *2-of-3* charts, also see Khoo et al. (2004). It involves solving the following equation $ARL_0 = (1 + p_*)/2p_*^2$ with $p_* = p_u = p_l$.

In establishing the ARL the essential TPM is established and Equation (1.4) applied. The essential TPM for the chart implementing the *2-of-3* runs-type signalling rule is provided in Table 4.8. State S_0, S_u and S_u refers to the charting statistic plotting in Region 0, U and L respectively as depicted in Figure 4.8. State S_{u0} refers to the first charting statistic plotting in Region U and the second plotting in Region 0.

Table 4.8: Essential TPM of the *2-of-3* runs-type signalling for a multivariate process monitoring a shift in process dispersion.

State	S_0	S_u	S_l	S_{uo}	S_{ou}	S_{lo}	S_{ol}
S_0	p_0	p_u	p_l	0	0	0	0
S_u	0	0	0	p_0	0	0	0
S_l	0	0	0	0	0	p_0	0
S_{uo}	p_0	0	0	0	0	0	0
S_{ou}	0	0	0	p_0	0	0	0
S_{lo}	p_0	0	0	0	0	0	0
S_{ol}	p_0	0	0	0	0	0	0

When considering the ARL performance of the various charts two quality characteristics were considered. Two scenarios arose, case one where the standard devation of one quality characteristic increases or decreases while the other remains constant and case two, where the standard deviations of both quality characteristics either increase or decrease. For case one, the best ARL performance for a shift (ξ) between 0.7 and 1.5 units was the *3-of-4* rule, while for a shift between 2 and 4 units the *1-of-1* rule provided superior performance. For case two the *3-of-4* rule provided the best ARL performance for shifts between 0.8 and 0.9 units as well as between 1.03 and 1.1 units. The *1-of-1* provided the best performance for shifts between 1.5 and 4 units as well as between 0.2 and 0.6 units, while the *2-of-4* rule provided the best performance for a shift between 1.15 and 2 units. The proposed charts also outperformed the chart based on the generalized sample variance $|\boldsymbol{S}|$ which has been mentioned earlier.

4.2 Conclusion of Chapter 4

Chapter 4 provides an overview of control charts supplemented with runs-type signalling rules in the multivariate environment when considering either a shift in location or spread. The articles discussed in Chapter 4 are listed below.

Runs-type signalling rules based on multivariate sample statistics

i. A performance analysis of Hotellings χ^2 control chart with supplementary runs rules (Aparisi et al. (2004)).

ii. Improving the performance of the Chi-square control chart via runs-rules (Koutras et al. (2006)).
iii. Adaptive Hotelling T^2 control chart with run rules (Lee (2013)).
iv. Incorporating runs rules into Hotellings χ^2 control charts (Khoo et al. (2003)).
v. Powerfull rules for Hotellings χ^2 control chart (Khoo et al. (2005)).
vi. Alternative to the multivariate control chart for process dispersion (Khoo et al. (2004)).

Aparisi et al. (2004) proposed the *1-of-1* and the *2-of-3* runs-type signalling rules in order to monitor various quality characteristics simultaneously. Koutras et al. (2006) proposes three chi-square control charts supplemented with runs-type signalling rules; the first is supplemented with the *1-of-1* runs-type signalling rule, the second with the *k-of-k* and the third applies a combination of the *k-of-k* and *r-of-r* runs-type signalling rules. Lee (2013) proposed charts that implement Hotelling's T^2 statistic supplemented by runs-type signalling rules differentiating between variable sample size (VSS), variable sample intervals (VSI) or both ($VSSI$). Khoo and Quah (2003) proposed Hotelling's control chart supplemented with the *1-of-1*, *2-of-2*, *2-of-3* and *2-of-4* runs-type signalling rules. To monitor both small and large shifts simultaneously Khoo et al. (2005) proposed a control chart based on Hotelling's χ^2 control chart supplemented with combined runs-type signalling rules. Khoo and Quah (2004) proposed six control charts incorporating runs-type signalling rules in order to detect small shifts in the process quicker. The proposed control charts implement the following runs-type signalling rules, the *1-of-1*, *2-of-2*, *2-of-3*, *2-of-4*, *3-of-3* and *3-of-4*.

Chapter 5 discusses the design and implementation of control charts supplemented with runs-type signalling rules.

Chapter 5

Design and implementation of control charts supplemented with runs-type signalling rules

Introduction

This Chapter is divided into two parts, Section 5.1 covering the design process while Section 5.2 focusses on the implementation of various control charts supplemented with runs-type signalling rules. Section 5.1 discusses the design process for univariate parametric, univariate nonparametric and multivariate control charts monitoring location and spread supplemented with runs-type signalling rules. Section 5.2 focusses on the implementation of the various control charts mentioned above considering both the IC and OOC scenarios. In order to establish various properties of the run-length distribution the simulation, Markov chain and exact approaches are considered in Section 5.2.

5.1. Design of control charts supplemented with runs-type signalling rules

When designing a control chart one of the main aims is to establish the control limits given some pre-determined criteria; for example a given ARL_0 or FAR. This leads to solving a general equation $ARL_0(\underline{\Theta})$ with k unknown parameters denoted by $\underline{\Theta} = (\Theta_1, \Theta_2, \Theta_3, \dots, \Theta_k)$. The parameters are a combination of the type of runs-type signalling rule used, the sample size as well as the control limits that need to be established; thus $ARL_0 = f(k, w, n, UCL_2, UCL_1, LCL_2, LCL_1)$ where k and w refer to the *k-of-w* runs-type signalling rule, n the sample size and the rest of the parameters to the control limits illustrated in Figure 2.5.

For the Shewhart $\bar{X}$ control chart the $ARL = 1/FAR$ and the run-length distribution follows a geometric distribution. Fu and Lou (2003) derived the ARL as a function of the essential TPM denoted by $\boldsymbol{Q}$ with

$E(T) = \underline{\boldsymbol{\omega}}(\boldsymbol{I}-\boldsymbol{Q})^{-1}\underline{\mathbf{1}} = f(k, w, n, UCL_2, UCL_1, LCL_2, LCL_1)$; see Equation (1.4) with the probabilities denoted in $\boldsymbol{Q}$ expressed as a function of the control limits and ω denoting the initial probability vector.

The design process is simplified by choosing the type of runs-type signalling rule as well as the sample size; typical runs type signalling rules would be, e.g. the *2-of-3*, *2-of-2*, *3-of-3* or *3-of-4* rules. It is well-known that a larger sample size leads to quicker OOC detection, provided all other parameters of interest are the same. The sample size is typically chosen to be 5 or 10. If a large sample size is chosen there could be a delay in signalling since e.g. 30 observations are needed compared to 5 or 10 . This could be impractical especially if say the *3-of-4* rule is implemented whereby three out of four plotting statistics need to plot above or below the control limits where each plotting statistic is determined from a sample of 30 values.

Depending on the control chart being implemented the unknown parameters could be further reduced, e.g. for the upper one-sided control chart the unknown parameters would be UCL_2 or UCL_1 or a combination of both depending on the control chart. When considering two-sided parametric or nonparametric control charts monitoring location the unknown parameters could be further reduced by implementing symmetrically placed control limits. For parametric control charts for monitoring spread the reduction could be established by considering equal tailed probabilities.

The design parameters are established for a given constraint which is typically an ARL_0 value of 370 or 500. When designing the control chart it is important to note that the given ARL_0 is not always attainable, since the distribution of the charting statistic is not always continuous. If the charting statistic is continuous then $ARL^* = ARL_0$, where ARL^* denotes the attained ARL value. This would be the case for parametric charts where the values are distributed iid $N(0,1)$.

On the other hand if the charting statistic is discrete, e.g. $Bin(n, p)$, then $ARL^* \neq ARL_0$. The control chart is then designed where $ARL^* \geq ARL_0$. An example of the scenario mentioned would be a nonparametric control chart with runs-type signalling rules where the control limits are integers resulting in a certain ARL_0 not being attained.

If it occurs that two competing control charts have the same ARL_0, the control chart with the smallest ARL_ξ is chosen as this control chart will be the quickest in detecting a shift in the process. Figure 5.1 provides a graphical illustration of two competing control charts having the same ARL_0 but different ARL_ξ values (say ARL_1 and ARL_2) for various shifts; the control chart 1, denoted by the diamond pattern, would be chosen since it provides the smallest ARL_ξ; where ξ denotes a process shift with $\xi > 0$.

Another scenario to consider would be for a certain shift, e.g. a small shift that control chart 2 provides the smallest ARL_ξ while for e.g. a large shift the control chart 1 provides the smallest ARL_ξ, see Figure 5.2. For control charts with the same ARL_0 the various shift intervals should then be considered with the control chart providing the smallest ARL_ξ for the interval chosen as the winner. When considering Figure 5.2 for a shift of $\xi \leq 1$, chart 2 (distinguished by the block pattern) provides the best ARL performance while for shifts where $\xi > 1$, chart 1 (distinguished by diamond pattern) provides the best ARL performance.

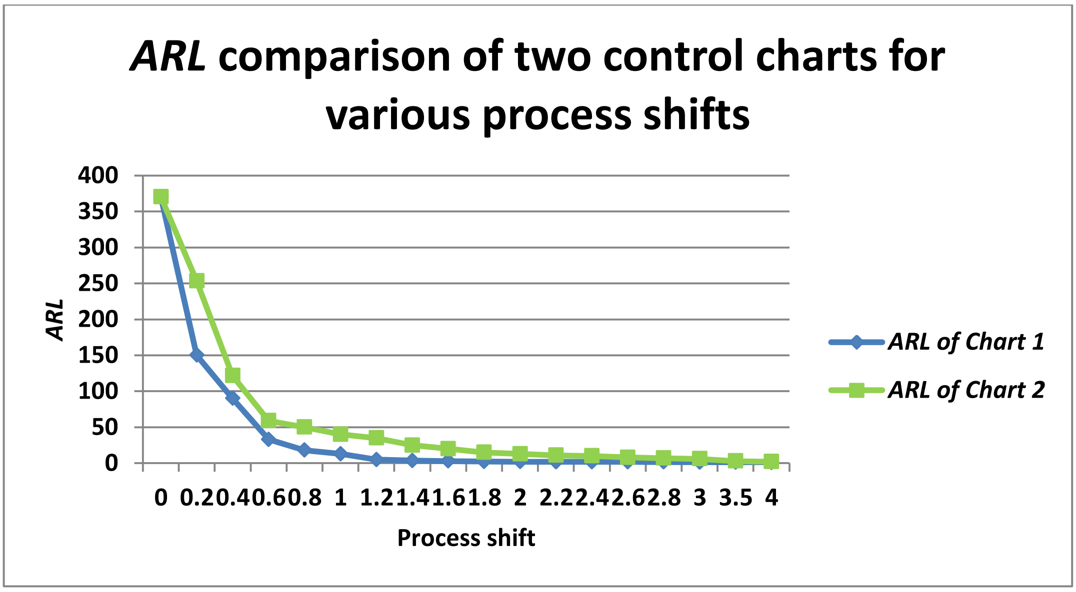

Figure 5.1: *ARL* comparison for various process shifts.

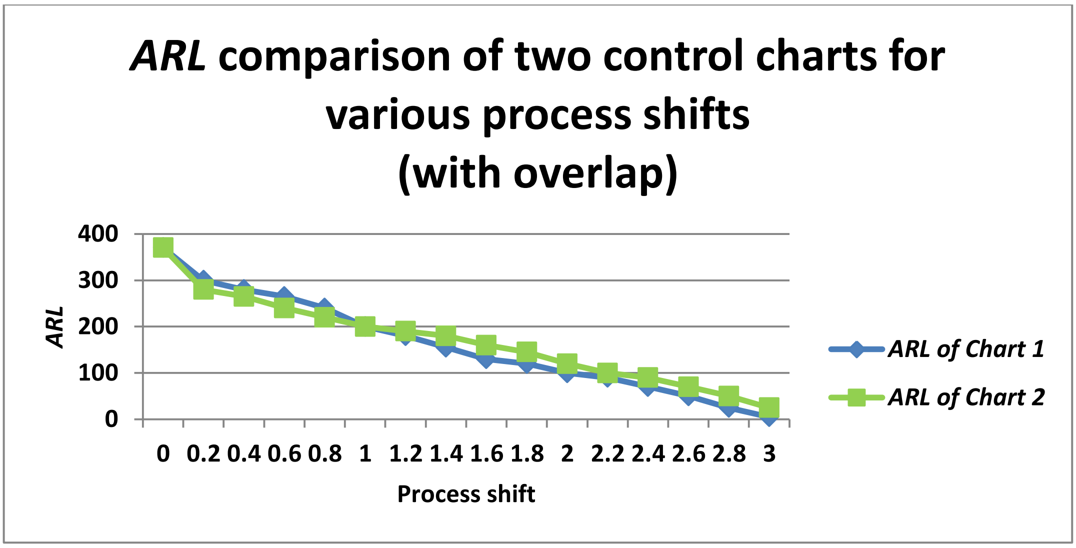

Figure 5.2 ***ARL*** **comparison for various process shifts with different** ARL_ξ **performance.**

In general the ARL is a function of the charting probabilities as can be seen from Equation (1.4), where Q is the essential TPM. Note that the charting (transition) probabilities are themselves usually functions of the control limits for both parametric and nonparametric control charts. Thus in order to establish the control limits Equation (1.4) needs to be solved for a given ARL_0 . For parametric charts monitoring spread, the probabilities are derived from the essential TPM given some criteria, e.g. ARL_0, and the control limits are established from the probabilities as shown in Section 2.2.1.

For the two-sided control charts, a minimum of two or a maximum of four control limits need to be found depending on whether both inner and outer control limits are present. To simplify the design process, when considering control charts monitoring location the control limits are symmetrically placed, reducing the number of unknown parameters from 4 to 2 when considering inner and outer control limits. Thus by applying symmetrically placed control limits it is assumed that increases and decreases are equally likely. Either the inner or outer control limits are then fixed further reducing the unknown parameters to only 1 value which can be solved by applying Equation (1.4) for a given ARL_0. When deriving the control limits for control charts monitoring spread supplemented with runs-type signalling rules either the inner or outer control limits are fixed but in order to simplify the process it is assumed that equal tailed probabilities are used reducing the unknown probabilities by two.

The steps in establishing the control limits are then as follows:

i. Define the regions and associated probabilities of the charting statistic.
ii. Establish the transition probabilities in terms of the control limits or as a function of the unknown parameters.
iii. Set up the essential TPM.
iv. Solve Equation (1.4) for a given ARL_0.

The steps outlined above can be followed for the one and two-sided control charts. An extra parameter is added when considering inner and outer control limits, e.g. the inner control limit. Two approaches can be followed in order to establish the control limits. Either the inner or outer control limits can be fixed while obtaining the unknown control limit. If the outer control limits are fixed a value is usually chosen larger than 3σ units, typically between 3.2σ and 3.7σ. The inner control limits are then established for a given ARL_0. The approach above is followed in order to monitor both small and large process shifts simultaneously. When applying the second approach the inner control limits are fixed and the outer control limits calculated for a given ARL_0.

In the Sections that follow the design of parametric and nonparametric control charts monitoring location and spread as well as multivariate control charts will be discussed.

5.1.1 Design of control charts for monitoring location supplemented with runs-type signalling rules

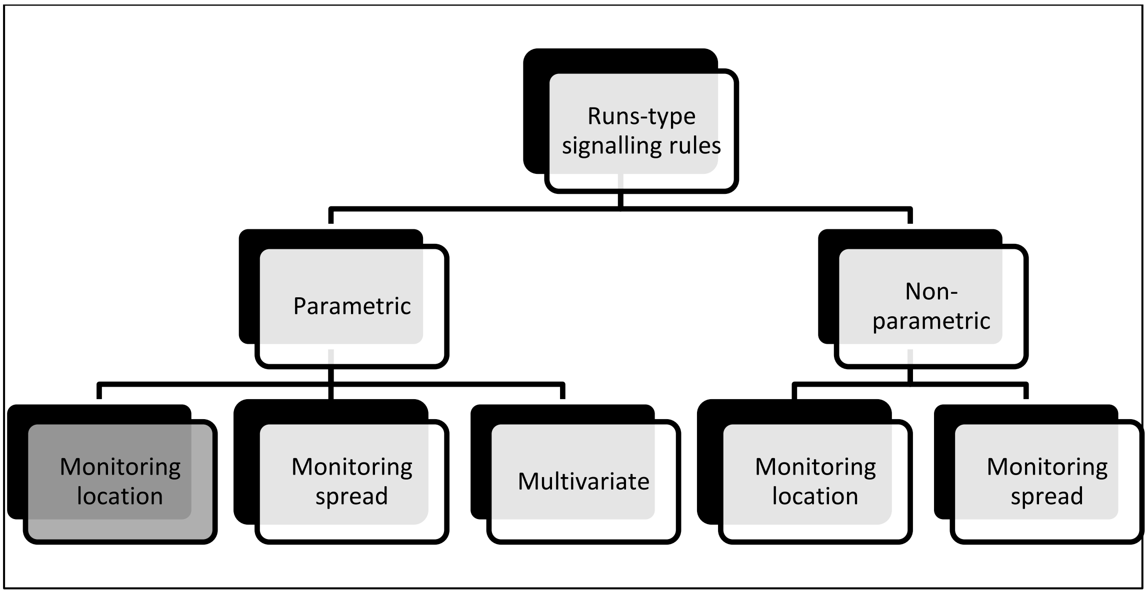

Figure 5.3: Classifications of the runs-type signalling rules control chart.

The following control charts are considered in the discussion that follows: the upper or lower one-sided control chart, the two-sided control chart as well as the two-sided control chart with inner and outer control limits.

5.1.1.1 Designing one-sided control charts for monitoring location supplemented with runs-type signalling rules

The upper one-sided chart will be discussed next. The process in designing the lower one-sided chart is similar. Figure 5.4 graphically illustrates the regions and associated probabilities considered in the design process in order to obtain the UCL denoted as d in the discussion below.

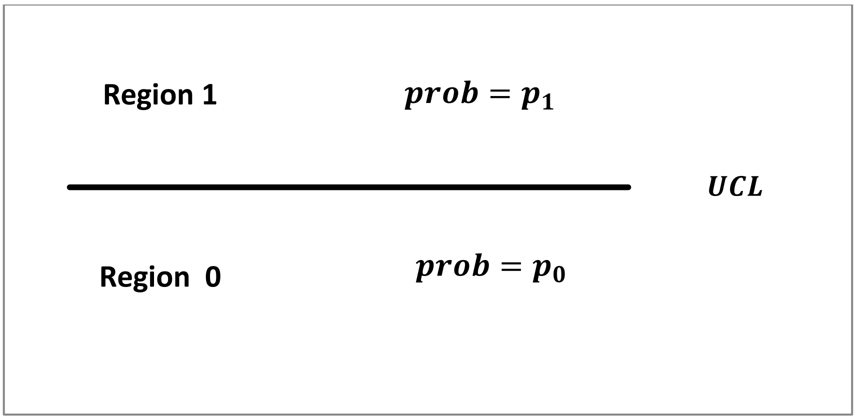

Figure 5.4: Charting (transition) probabilities and regions considered for the *2-of-3* runs-type signalling rule.

The steps in order to solve d (UCL) are as follows:

i. Derive the transition probabilities as a function of the control limit.
ii. Set up the essential TPM.
iii. Solve Equation (1.4) with respect to the UCL denoted as d.

The transition probabilities in terms of the control limit equals $p_1 = 1 - \phi(d)$ and $p_0 = \phi(d)$. As an example, when considering the *2-of-3* runs-type signalling rule the essential TPM reduces to Table 5.1.

Table 5.1: Essential TPM for the one-sided control chart supplemented with the *2-of-3* runs-type signalling rule.

State	S_0	S_1	S_{01}	S_{10}
S_0	p_0	0	p_1	0
S_1	0	p_1	0	p_0
S_{01}	0	0	0	p_0
S_{10}	p_0	0	0	0

Equation (1.4) should then be applied and solved for the $UCL = d$ as follows.

$$ARL_0 = (1 \quad 0 \quad 0 \quad 0)\begin{pmatrix} 1-p_0 & 0 & -p_1 & 0 \\ 0 & 1-p_1 & 0 & -p_0 \\ 0 & 0 & 1 & -p_0 \\ -p_0 & 0 & 0 & 1 \end{pmatrix}^{-1}\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$
$$= \frac{1 - p_1 + p_1(1-p_1) + p_1 p_0 (1-p_1)}{(1-p_0)(1-p_1) - p_1(1-p_1)p_0^2}.$$

The only unknown in the equation above is d which can be solved since the ARL_0 usually equals 370 or 500. The reader is referred to the appendix regarding the derivation of $(\mathrm{I}-Q)^{-1}$; the closed form expression was then obtained by applying matrix algebra. The procedure to obtain the LCL for the lower one-sided chart is similar whereby the transition probabilities are derived as a function of the control limit, the essential TPM set up and equation (1.4) applied.

5.1.1.2 Designing two-sided control charts for monitoring location supplemented with runs-type signalling rules

For the two-sided control chart supplemented with runs-type signalling rules the reader is referred to Klein (2000). The steps in designing the control chart are similar to the upper one-sided chart:

i. Derive the probabilities as a function of the control limits.
ii. Set up the essential TPM.
iii. Solve Equation (1.4) with respect to the UCL denoted as d.

For the discussion that follows the reader is referred to Figure 2.3 regarding the charting probabilities and regions; the probability plotting in Region 0 referred to in Section 2.1.1 is denoted by p_0 below. Note that the definitions of the transition probabilities p_u and p_l are provided in Section 2.1.2. For the purposes of this discussion a signal would be produced if two out of the last two charting statistics plot on or above the UCL or on or below the LCL. Not all the signalling criteria discussed in Klein (2000a) would be incorporated. The signalling events would then be UU and LL (see Figure 2.3) with the essential TPM provided in Table 5.2. By assuming symmetric control limits the only unknown in the essential TPM is the value of the $UCL = d$. Note that $p_0 = 1 - p_l - p_u$, $p_u = 1 - \phi(d)$, $p_l = \phi(-d)$.

Table 5.2: Essential TPM for the two-sided control chart supplemented with the *2-of-2* runs-type signalling rule

State	$\boldsymbol{S_0}$	$\boldsymbol{S_U}$	$\boldsymbol{S_L}$
$\boldsymbol{S_0}$	p_0	p_U	p_L
$\boldsymbol{S_U}$	p_0	0	p_L
$\boldsymbol{S_L}$	p_0	p_U	0

The $UCL = d$ can be solved by applying Equation (1.4) and using the ARL_0 value of 370 or 500. Equation (1.4) then reduces to:

$$ARL_0 = \begin{pmatrix} 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1-p_0 & -p_U & -p_L \\ -p_0 & 1 & -p_L \\ -p_0 & -p_U & 1 \end{pmatrix}^{-1} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$
$$= \frac{1 - p_L p_U + (1-p_L)p_U + p_U + p_U p_L + p_L}{1 - p_0 + p_0 p_U(1-p_L) - p_0 p_U p_L + (1-p_0)p_U(1-p_L) - p_0 p_U}.$$

The reader is referred to the appendix regarding the derivation of $(\mathrm{I} - Q)^{-1}$ in the equation above; the final closed form expression was obtained by applying matrix algebra.

5.1.1.3 Designing two-sided control charts with inner and outer control limits for monitoring location and supplemented with runs-type signalling rules

The reader is referred to the "Revised *m-of-k* Runs rule" (Antzoulakos and Rakitzis (2008)) as an example of two-sided control charts implementing both inner and outer control limits.

The regions and associated probabilities are illustrated in Figure 2.5. The control limits are calculated by fixing the outer control limits denoted by K and $-K$ and then calculating the inner control limits denoted by d and $-d$, for a specific ARL_0; note that $UCL_2 = K, LCL_2 = -K, UCL_1 = d, LCL_1 = -d$. The steps in designing the control chart are then as follows:

i. Derive the probabilities as a function of the control limits.
ii. Set up the essential TPM.
iii. Solve Equation (1.4) with respect to the inner upper and lower control limits, denoted as d and $-d$ (assuming symmetrically placed control limits).

The associated transition probabilities are then provided by the following equations, where the shift $\xi = 0$ if the process is IC and the outer control limit K is fixed.

$p1 = \phi(K-\xi) - \phi(d-\xi)$.

$p2 = \phi(d-\xi) + \phi(\xi) - 1$.

$p3 = \phi(d+\xi) - \phi(\xi)$.

$p4 = \phi(K+\xi) - \phi(d+\xi)$.

$p5 = 1 - p1 - p2 - p3 - p4$.

As an example the essential TPM for the revised *2-of-2* runs-type signalling rule is provided in Table 5.3.

Table 5.3: Essential TPM for the two-sided control chart with inner and outer control limits supplemented with the *2-of-2* runs-type signalling rule.

State	$\boldsymbol{S_{2\text{ or }3}}$	$\boldsymbol{S_1}$	$\boldsymbol{S_4}$
$\boldsymbol{S_{2\text{ or }3}}$	$p_2 + p_3$	p_1	p_4
$\boldsymbol{S_1}$	$p_2 + p_3$	0	p_4
$\boldsymbol{S_4}$	$p_2 + p_3$	p_1	0

Now by applying Equation (1.4) d can be solved as follows, replacing the probability equations provided above:

$$ARL_0 = \begin{pmatrix} 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 - p_2 - p_3 & -p_1 & -p_4 \\ -p_2 - p_3 & 1 & -p_4 \\ -p_2 - p_3 & -p_1 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

$$= \frac{1 + p_1 p_4 + p_1 + p_4}{1 - p_2 - p_3 - 2(p_2 + p_3)p_1 p_4 - (1 - p_2 - p_3)p_1 p_4 + (p_2 + p_3)p_4 - (p_2 + p_3)p_1}.$$

The reader is referred to the appendix regarding the derivation of $(\mathrm{I} - Q)^{-1}$ in the equation above.

5.1.2 Design of control charts for monitoring spread supplemented with runs-type signalling rules

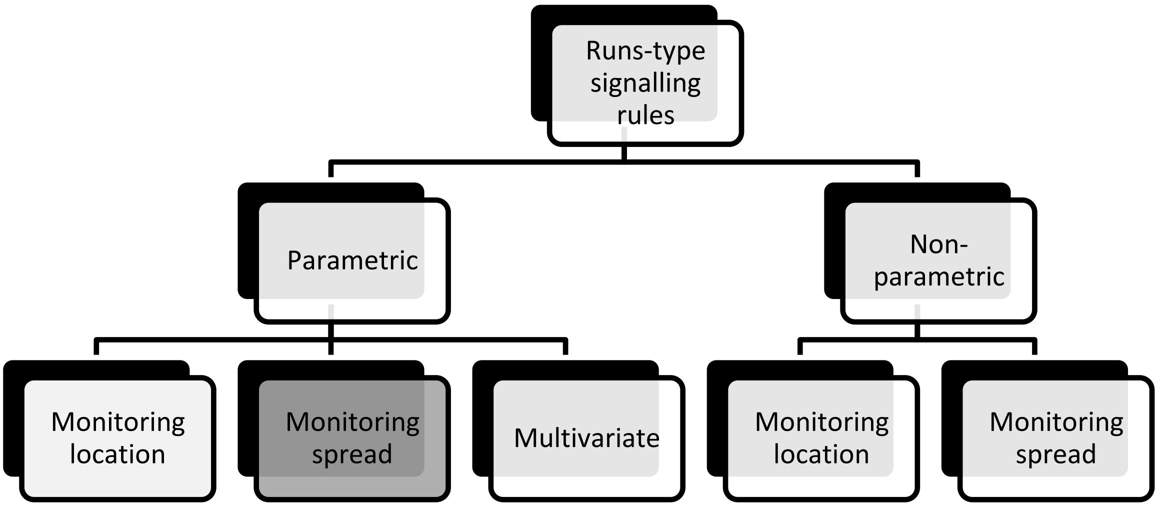

Figure 5.5: Classifications of the runs-type signalling rules control chart.

5.1.2.1 Designing one-sided control charts for monitoring spread supplemented with runs-type signalling rules

Setting up the control limits for the *k-of-w* $(k = 2, w = 3)$ runs-type signalling rule is discussed below for a control chart monitoring spread. The same methodology could be followed in order to derive the control limits for other variations of the *k-of-w* runs-type signalling rules class. Establishing the control limits involves finding the probabilities p_1 and p_0 denoted in Figure 5.6 by applying the following steps:

i. Setting up the essential TPM, see Table 5.4.
ii. Solving Equation (1.4).
iii. Finding the control limits by using the transition probabilities.

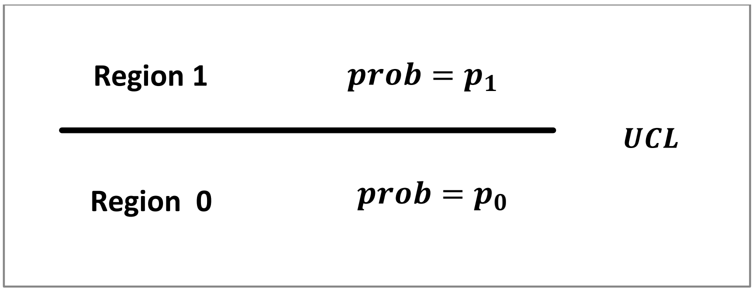

Figure 5.6: Transition probabilities and regions considered for the *2-of-3* runs-type signalling rule.

Table 5.4: Essential TPM for the one-sided control chart supplemented with the *2-of-3* runs-type signalling rule.

State	S_0	S_1	S_{01}	S_{10}
S_0	$1-p_1$	0	p_1	0
S_1	0	p_1	0	$1-p_1$
S_{01}	$1-p_1$	0	0	0
S_{10}	$1-p_1$	0	0	0

Equation (1.4) should then be applied and the transition probabilities solved as follows:

$$ARL_0 = (1 \quad 0 \quad 0 \quad 0)\begin{pmatrix} p_1 & 0 & -p_1 & 0 \\ 0 & 1-p_1 & 0 & 1-p_1 \\ 1-p_1 & 0 & 1 & 0 \\ 1-p_1 & 0 & 0 & 1 \end{pmatrix}^{-1}\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$

$$= \frac{1-p_1+p_1(1-p_1)}{p_1(1-p_1)+p_1(1-p_1)^2}.$$

The only unknown in the equation above is p_1 which can be solved since the ARL_0 usually equals 370 or 500. The reader is referred to the appendix regarding the derivation of $(\mathrm{I}-Q)^{-1}$ in the equation above. The control limits can be established as in Section 2.2.1 by applying the following equation: $UCL = \sqrt{(\sigma^2(\chi^2_{(n-1,1-p_1)})/(n-1))}$.

5.1.2.2 Designing two-sided control charts for monitoring spread and supplemented with runs-type signalling rules

The reader is referred to Klein (2000b) for a discussion on two sided control charts monitoring spread supplemented with runs-type signalling rules. The discussion that follows will focus on deriving control limits based on equal tail probabilities. In order to derive the control limits for control charts supplemented with runs-type signalling rules using unequal tail probabilities a simulation approach is suggested. The approach to establish the control limits are similar to the design of one-sided charts (Section 5.2.1) whereby the probabilities are found and the control limits derived from them.

The reader is referred to Figure 2.11 for a graphical representation of the regions and associated transition probabilities discussed further; with State 2 and 3 referring to State U and L in Table 5.5 and S_0 referring to the state entered if the charting statistic plots in Region 1. Since equal tail probabilities are used, $p_u = p_l$. The essential TPM for the *2-of-3* and *2-of-2* runs-type signalling rules are provided as an example, in Table 5.4 and Table 5.5.

Table 5.5: Essential TPM for the two-sided control chart supplemented with the *2-of-2* runs-type signalling rule.

State	$\boldsymbol{S_0}$	$\boldsymbol{S_U}$	$\boldsymbol{S_L}$
$\boldsymbol{S_0}$	p_0	p_U	p_L
$\boldsymbol{S_U}$	p_0	0	p_L
$\boldsymbol{S_L}$	p_0	p_U	0

The transition probabilities can be solved by applying Equation (1.4) and applying the ARL_0 value of 370 or 500. Equation (1.4) then reduces to:

$$ARL_0 = (1 \quad 0 \quad 0)\begin{pmatrix} 1-p_0 & -p_U & -p_L \\ -p_0 & 1 & -p_L \\ -p_0 & -p_U & 1 \end{pmatrix}^{-1}\begin{pmatrix}1\\1\\1\end{pmatrix}$$
$$= \frac{1+p_Up_L+p_U+p_L}{1-p_0-2p_0p_Up_L-(1-p_0)p_Up_L-p_0p_L-p_0p_U}.$$

Only one unknown exists in the equation above since $p_U = p_L$ and $p_U + p_L + p_0 = 1$. The equation can be solved by substituting an ARL_0 of either 370 or 500 in the equation above. The reader is referred to the appendix regarding the derivation of $(\mathrm{I}-Q)^{-1}$ in the equation above. The control limits can then be established by applying the following equations, see Section 2.2.1:

$$LCL = \sqrt{(\sigma^2(\chi^2_{(n-1,p_l)})/(n-1))}$$

and $UCL = \sqrt{(\sigma^2\chi^2_{(n-1,1-p_u)}/(n-1))}$.

5.1.2.3 Designing two-sided control charts with inner and outer control limits for monitoring spread and supplemented with runs-type signalling rules

The reader is referred to Acosta-Mejia et al. (2009) as an example of control charts monitoring spread supplemented with the *k-of-k* runs-type signalling rules. For a graphical representation of the regions and transition probabilities that follow the reader is referred to Figure 2.14; the regions and probabilities shown in Figure 2.14 are used in the discussion below. The following Section illustrates how to design an equal tailed control chart supplemented with *k-of-k* runs type signalling rules.

In order to determine the inner control limits the outer control limits are fixed; thus p_4 and p_5 are known. Since the chart is equal tailed, $p_2 = p_3$ and $p_4 = p_5$ where the exact definitions for the transition probabilities are as follows:

$p_1 = P(LWL < S < UWL)$.

$$p_1 = F_{\chi^2_{n-1}}\left(\frac{1}{\xi^2}\chi^2_{n-1,p_1+p_3+p_5}\right) - F_{\chi^2_{n-1}}\left(\frac{1}{\xi^2}\chi^2_{n-1,p_3+p_5}\right).$$

$p_2 = P(UWL < S < UCL)$.

$$p_2 = F_{\chi^2_{n-1}}\left(\frac{1}{\xi^2}\chi^2_{n-1,1-p_4}\right) - F_{\chi^2_{n-1}}\left(\frac{1}{\xi^2}\chi^2_{n-1,p_1+p_3+p_5}\right).$$

$p_3 = P(LCL < S < LWL)$.

$$p_3 = F_{\chi^2_{n-1}}\left(\frac{1}{\xi^2}\chi^2_{n-1,p_3+p_5}\right) - F_{\chi^2_{n-1}}\left(\frac{1}{\xi^2}\chi^2_{n-1,p_5}\right).$$

$p_4 = P(S > UCL)$.

$$p_4 = 1 - F_{\chi^2_{n-1}}\left(\frac{1}{\xi^2}\chi^2_{n-1,1-p_4}\right).$$

$p_5 = P(S < LCL)$.

$$p_5 = F_{\chi^2_{n-1}}\left(\frac{1}{\xi^2}\chi^2_{n-1,p_5}\right).$$

where $\xi^2 = \frac{\sigma_t^2}{\sigma_0^2}$ and $\chi^2_{n,p}$ denotes a Chi-square random variable with parameters (n,p).

The essential TPM can be set up and solved for p_2 since the other probabilities would be known or be expressed in terms of p_2. As an example, the essential TPM for the *2-of-2* runs-type signalling rule is provided in Table 5.6.

Table 5.6: Essential TPM for the two-sided control chart supplemented with the *2-of-2* runs-type signalling rule.

State	S_1	S_2	S_3
S_1	p_1	p_2	p_3
S_2	p_1	0	p_3
S_3	p_1	p_2	0

Equation (1.4) can be solved in terms of p_2 as follows:

$$ARL_0 = (1 \quad 0 \quad 0)\begin{pmatrix} 1-p_1 & -p_2 & -p_3 \\ -p_1 & 1 & -p_3 \\ -p_1 & -p_2 & 1 \end{pmatrix}^{-1}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$
$$= \frac{1+p_3}{1-p_1-2p_1p_2p_3-(1-p_1)p_2p_3-p_1p_3-p_1p_2}.$$

The reader is referred to the appendix regarding the derivation of $(\mathrm{I}-Q)^{-1}$ in the equation above. Depending on whether the inner or outer control limits are fixed the unknown control limits can be established by applying the following equations given that the probabilities were already obtained, see Section 2.1.2:

$$LCL = \sqrt{(\sigma^2\chi^2_{(n-1,p_5)}/(n-1))}$$

and $UCL = \sqrt{(\sigma^2\chi^2_{(n-1,1-p_4)}/(n-1))}$.

5.1.3 Design of nonparametric control charts for monitoring location supplemented with runs-type signalling rules

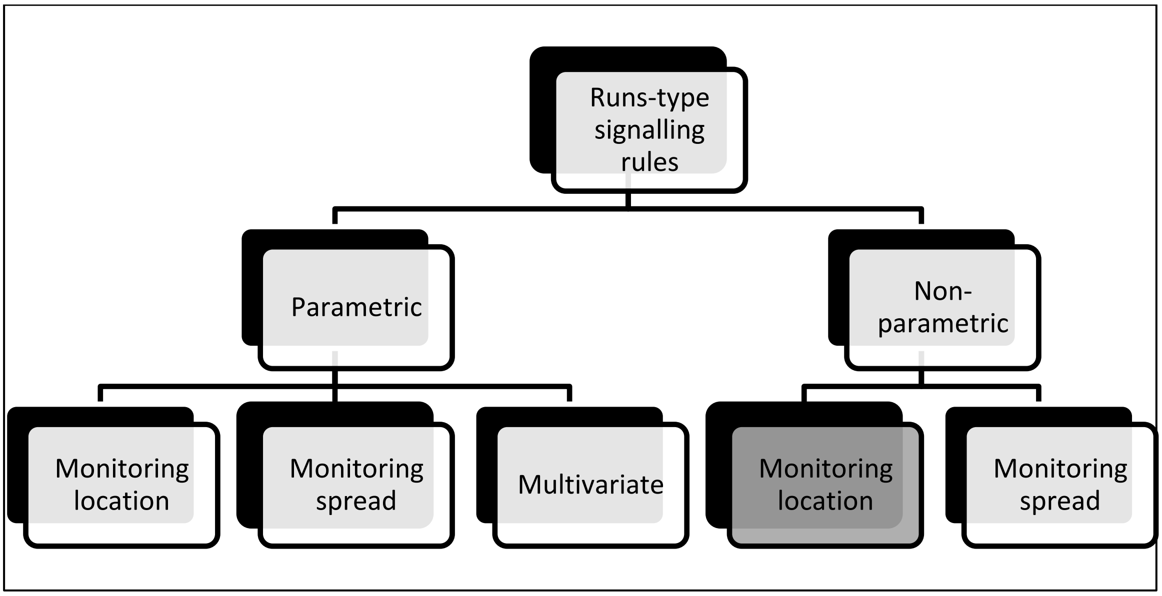

Figure 5.7: Classifications of the runs-type signalling rules control chart.

5.1.3.1 Designing nonparametric upper one-sided control charts for monitoring location supplemented with runs-type signalling rules

In the nonparametric setting two scenarios will be discussed namely the upper one-sided chart and the two-sided chart. The design process differs slightly from the parametric control charts. Since the charting statistic is discrete, i.e.$\sim Bin(n,p)$, the $ARL^{*} \neq ARL_0$, where ARL^{*} denotes the attainable ARL value. The control chart is then designed such that $ARL^{*} \geq ARL_0$. As in the parametric setting the probabilities are expressed as a function of the control limits, the essential TPM set up and Equation (1.4) solved for a given ARL_0.

For a graphical representation of the transition probabilities and regions the reader is referred to Figure 3.6. (The UCL equals $n-b$ while the LCL equals a. As an example the associated probability of a charting statistic plotting on or above the control limit for the upper one-sided *2-of-3* control chart is given by: $p_1 = P(T_i \geq UCL) = P(T_i \geq n-b) = I_p(n-b, b+1)$). The reader is referred to Table 3.5 for the essential TPM. The probability p_1 can be solved by applying Equation (1.4) for a given ARL_0 as follows:

$$ARL_0 = (1 \quad 0 \quad 0 \quad 0)\begin{pmatrix} p_1 & 0 & -p_1 & 0 \\ 0 & 1 & 0 & -(1-p_1) \\ 0 & 0 & 1 & -(1-p_1) \\ -(1-p_1) & 0 & 0 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$

$$= \frac{1+2p_1-{p_1}^2}{p_1}.$$

Note that the empty state referred to in Table 3.5 was excluded in the essential TPM above. In order to calculate the ARL_0 above, the inverse of a $4x4$ matrix needs to be established; the reader is referred to the appendix regarding the calculation of the inverse of a $4x4$ matrix.

The procedure followed in order to calculate the control limit is as follows:

i. Consider all possible discrete upper control limits available.
ii. For each of these control limits, calculate p_1 and the ARL_0.
iii. The control limit chosen is the limit providing an attained ARL larger or equal to the ARL_0 ($ARL^* \geq ARL_0$).

The properties of a cdf function of a $Bin(n,p)$ distribution are used in order to calculate p_1 and then establish the ARL_0; Note that $p_1 = P(T_i \geq n-b) = 1 - P(T_i \leq n-b-1) = \sum_{i=1}^{n-b-1}\binom{n}{i}p^i(1-p)^n$. If the median is monitored and the process is IC then $p = 0.5$, while for any other percentile of interest $p \neq 0.5$, even if the proces is IC. Note that the charting statistic follows a $Bin(n,p)$ distribution as mentioned previously.

When considering lower one sided charts the procedure is exactly the same in order to establish the control limit(s); all possible lower control limits are considered, p_2 and the ARL_0 calculated and the limit chosen such that $ARL^* \geq ARL_0$; where $p_2 = P(T_i \leq LCL)$= $P(T_i \leq a)$.

5.1.3.2 Designing nonparametric two-sided control charts for monitoring location supplemented with runs-type signalling rules

Exactly the same approach as for the one-sided control chart (Section 5.3.1) can be followed for the two-sided control chart with the following associated transition probabilities applied in the essential TPM (Table 5.7):

$p_1 = P(T_i \geq UCL) = P(T_i \geq n-b) = I_p(n-b, b+1) = 1 - P(T_i \leq n-b-1)$.

$p_2 = P(T_i \leq LCL) = P(T_i \leq a) = I_p(a+1, n-a)$.

$p_0 = 1 - p_1 - p_2$.

The reader is referred to Figure 3.3 for a graphical representation of the associated regions and transition probabilities and to Table 5.7 for the essential TPM; as an example the *2-of-2* runs-type signalling rule is considered.

Table 5.7: Essential TPM for the nonparametric two-sided control chart supplemented with the *2-of-2* runs-type signalling rule.

State	S_0	S_1	S_2
S_0	p_0	p_1	p_2
S_1	p_0	0	p_2
S_2	p_0	p_1	0

The associated probabilities and control limits can then be solved by applying Equation (1.4) for a given ARL_0 as follows:

$$ARL_0 = (1 \quad 0 \quad 0)\begin{pmatrix} 1-p_0 & -p_1 & -p_2 \\ -p_0 & 1 & -p_2 \\ -p_0 & -p_1 & 1 \end{pmatrix}^{-1}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

$$= \frac{1 + p_1p_2 + p_1 + p_2}{1 - 2p_0 - p_0p_1p_2 - (1-p_0)p_1p_2 - p_0p_2 + p_0p_1}.$$

In order to calculate the ARL_0 above, the inverse of a $3x3$ matrix needs to be established; the reader is referred to the appendix regarding the calculation of the inverse of a $3x3$ matrix. The final closed form expression is obtained by applying matrix algebra.

The $ARL^* \neq ARL_0$ since the charting statistic is discrete. The approach followed to establish the control limits is similar to the one sided chart, where various combinations of the control limits are established and the combination resulting in an ARL^* larger or equal to ARL_0 being chosen as the control limit.

The reader is referred to Human et al. (2010) with respect to the *2-of-2* chart based on precedence statistics. The transition probabilities associated with plotting either above or below the UCL or LCL given the control limits are then as follows:

$p_L = P\left(Y_{j:n} \leq X_{a:m} | X_{a:m} = x_1\right) = I_{G(x_1)}(j, n-j+1).$

and

$p_U = P\left(Y_{j:n} \geq X_{b:m} | X_{b:m} = x_2\right) = 1 - I_{G(2x_2)}(j, n-j+1).$

where $I_{G(x_1)}$ and $I_{G(2x_2)}$ are defined in Section 3.1.5.

The essential TPM is provided in Section 3.1.5; the empty state is included but for the calculation below the empty state would be excluded. These probabilities can be solved by applying Equation (1.4) for a given ARL_0 as follows:

$$ARL_0 = (1 \quad 0 \quad 0)\begin{pmatrix} p_L + p_U & -p_U & -p_L \\ p_L + p_U & 1 & -p_L \\ p_L + p_U & -p_U & 1 \end{pmatrix}^{-1}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$
$$= \frac{p_L + p_U + p_L}{p_L + p_U + (p_L + p_U)p_U p_L + (p_L + p_U)p_U + (p_L + p_U)p_L}.$$

In order to calculate the ARL_0 above, the inverse of a $3x3$ matrix needs to be established; the reader is referred to the appendix regarding the calculation of the inverse of a $3x3$ matrix. The final closed form expression is obtained by applying matrix algebra.

5.1.4 Designing multivariate charts supplemented with runs-type signalling rules

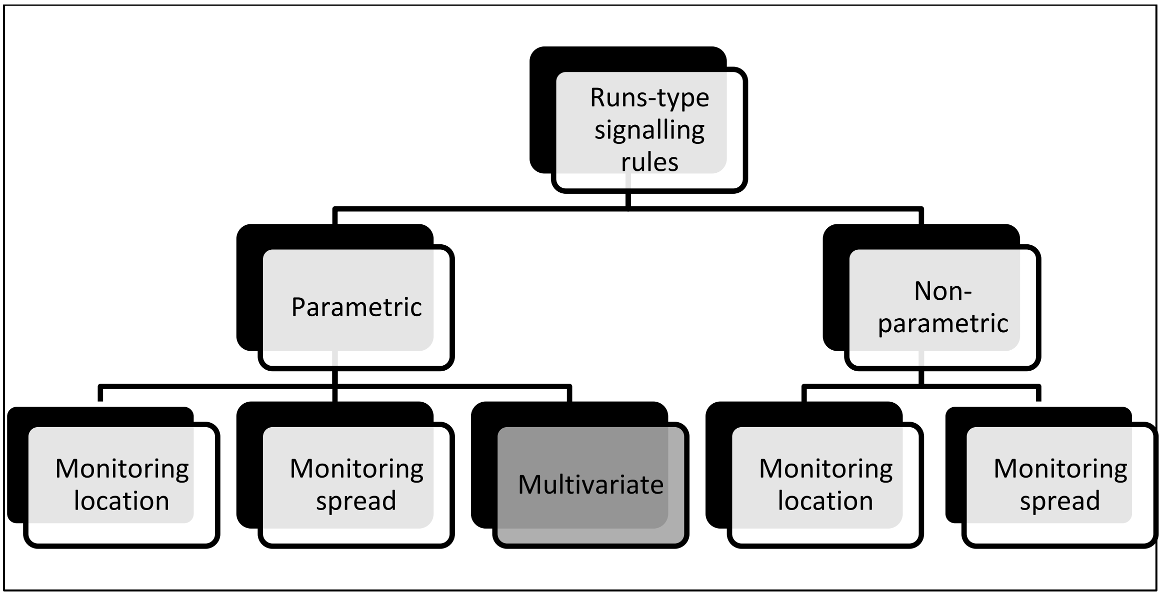

Figure 5.8: Classifications of the runs-type signalling rules control chart.

5.1.4.1 Designing one-sided multivariate control charts for monitoring location supplemented with runs-type signalling rules

The reader is referred to Khoo and Quah (2003) in Section 4.1.4 as an example of a one-sided control chart in the multivariate environment. The process is similar to the previous discussions where the design procedure is:

i. Deriving the probability of the charting statistic plotting in a specific region.
ii. Setting up the essential TPM.
iii. Solving Equation (1.4) for a given ARL_0.

The reader is referred to Figure 4.6 for a graphical representation of the regions and probabilities discussed further. The essential TPM is provided in Table 5.8 with the probabilities p_A and p_B defined in Section 4.1.4.

Table 5.8: Essential TPM for the multivariate control chart supplemented with the *2-of-3* runs-type signalling rule.

State	S_B	S_A	S_{BA}	S_{AB}
S_B	p_B	0	p_A	0
S_A	0	p_A	0	p_B
S_{BA}	0	0	0	p_B
S_{AB}	p_B	0	0	0

The transition probabilities can be found by solving Equation (1.4) for a given ARL_0 as follows:

$$ARL_0 = (1 \quad 0 \quad 0 \quad 0)\begin{pmatrix} 1-p_B & 0 & -p_A & 0 \\ 0 & 1-p_A & 0 & -p_B \\ 0 & 0 & 1 & -p_B \\ -p_B & 0 & 0 & 1 \end{pmatrix}^{-1}\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$

$$= \frac{1}{(1-p_B)(1-p_A) - p_A(1-p_A){p_B}^2}.$$

where $p_A + p_B = 1$ and $ARL_0 = 370$ or 500. The control limits can then be established from standard normal tables or by using a software package such as Excel. The reader is referred to the appendix regarding the derivation of $(\mathrm{I} - Q)^{-1}$ in the equation above.

5.1.4.2 Designing one-sided multivariate control charts with inner and outer control limits for monitoring location supplemented with runs-type signalling rules

The reader is referred to Khoo et al. (2005) in Section 4.1.5 as an example of a one-sided multivariate control chart with inner and outer upper control limits supplemented with runs-type signalling rules. In order to design the chart the following steps are followed:

i. Deriving the probability of the charting statistic plotting in a specific region.
ii. Setting up the essential TPM.
iii. Solving Equation (1.4) for a given ARL_0.

The reader is referred to Figure 4.7 for a graphical representation and Table 4.7 for the essential TPM for *1-of-1* combined with *2-of-3* runs-type signalling rule. The probabilities can be solved by applying Equation (1.4) for a given ARL_0 as follows:

$$ARL_0 = (1 \quad 0 \quad 0 \quad 0)\begin{pmatrix} 1 & -p_B & 0 & -p_A \\ -p_A & 1 & -p_B & 0 \\ 0 & -p_B & 1 & 0 \\ 0 & -p_B & 0 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$

$$= \frac{1 - {p_B}^2 - (1 - p_B) + p_A p_B + p_A}{1 - {p_B}^2 - {p_A}^2 {p_B}^2}.$$

where $p_A + p_B + p_{OUT} = 1$ and $ARL_0 = 370$ or 500. The reader is referred to the appendix regarding the derivation of $(\mathrm{I} - Q)^{-1}$ in the equation above.

5.1.4.3 Designing two-sided multivariate control charts for monitoring spread supplemented with runs-type signalling rules

The reader is referred to Khoo and Quah (2004) in Section 4.1.6 as an example of a two-sided multivariate control charts monitoring dispersion supplemented with runs-type signalling rules. Figure 4.8 illustrates the regions and probabilities discussed in the following section and Table 5.9 the essential TPM for the *2-of-2* runs-type signalling rule.

Table 5.9: Essential TPM for the two-sided multivariate control chart supplemented with the *2-of-2* runs-type signalling rule.

State	S_0	S_u	S_l
S_0	p_0	p_u	p_l
S_u	p_0	0	p_l
S_l	p_0	p_u	0

The steps in setting up the control limits are similar to the previous discussions. The probabilities can then be solved by applying Equation (1.4) for a given ARL_0 as follows:

$$ARL_0 = (1 \quad 0 \quad 0)\begin{pmatrix} 1 - p_0 & -p_u & -p_l \\ -p_0 & 1 & -p_l \\ -p_0 & -p_u & 1 \end{pmatrix}^{-1} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \frac{p_u p_l + p_l}{(1-p_0) - 2p_0 p_u p_l - (1-p_0) p_u p_l - 2 p_u p_l}.$$

where $p_U = P(C_i > UCL)$ and $p_L = P(C_i < LCL)$ and $C_i \sim N(0,1)$. The reader is referred to the appendix regarding the derivation of $(\mathrm{I} - Q)^{-1}$ when applying Equation (1.4).

5.2. Implementation of control charts supplemented with runs-type signalling rules

Section 5.2 focusses on the implementation of univariate parametric and nonparametric control charts monitoring location and spread supplemented with runs-type signalling rules. The ARL is evaluated for both the IC and OOC states in order to compare the results to control charts available in the literature; this is achieved by considering tables, box and whisker like plots as well as CDF graphs. Note that the implementation of multivariate control charts supplemented with runs-type signalling rules were excluded from Section 5.2.

The 6 control charts chosen for implementation purposes were selected in order to be representative of the thesis, covering the parametric and nonparametric cases as well as shifts in location and spread for various runs-type signalling rules; note that case K (known process parameters) and case U (unknown process parameters) were included in the following discussion. The univariate parametric control charts monitoring both location and spread as well as nonparametric control charts monitoring both location and spread as well as non-parametric control charts monitoring location that were chosen are as follows:

i. The Shewhart $\overline{\mathrm{X}}$ control chart.
ii. The improved *2-of-2* control chart (Khoo et al. (2006)).
iii. The revised *2-of-3* control chart (Antzoulakos and Rakitzis (2008a)).
iv. *S*-chart supplemented with the *2-of-2* runs-type signalling rule (Anzoulakos et al. (2010)).
v. Nonparametric control chart supplemented with the *2-of-2* runs-type signalling rule (Human et al. (2010)).
vi. Run rules $\overline{\mathrm{X}}$ charts when process parameters are unknown (Zhang et al. (2010).

In order to evaluate various essential TPM sizes, the *2-of-2* as well as the *2-of-3* runs-type signalling rules was considered. The Shewhart $\overline{\mathrm{X}}$ control chart was considered since it is very well known in the literature. Note that normally for box and whisker plots the minimum and maximum end points are considered but for illustrative purposes these values were substituted with the 1^{st} and 99^{th} percentiles. It should also be noted that a sample size $(\mathrm{n} = 5)$ was generally considered in order to obtain the results unless otherwise specified.

5.2.1 Implementation of the Shewhart $\bar{X}$ control chart.

The following results depict the simulation, Markov chain and exact approach for the Shewhart $\bar{X}$ control chart when considering observations from a $N(0,1)$ distribution. Note that the Shewhart $\bar{X}$ control chart considers the *1-of-1* runs-type signalling rule as mentioned in Section 1.6. An OOC signal is produced if the charting statistics plots either on or above the UCL or on or below the LCL, see Figure 1.1.

Note that the Shewhart $\bar{X}$ control chart considered was designed with an $ARL_0 = 370$. It is well known that the advantage of the Shewhart $\bar{X}$ control chart is their large shift detecting capabilities, the detection of small shifts in location being a drawback. The results for control charts with enhanced small shift detection capabilities are discussed in Section 5.2.2 to Section 5.2.5.

Table 5.10: Results obtained from a simulation approach for the Shewhart $\bar{X}$ control chart.

Shift	ARL	$SDRL$	5th Percentile	25th Percentile	50th Percentile	75th Percentile	95th Percentile
0	369.55	368.01	19	106	256	513	1105
0.2	309.46	307.52	16	90	215	429	919
0.4	198.51	197.07	11	58	138	276	592
0.6	119.56	119.52	7	35	83	165	359
0.8	71.72	71.08	4	21	50	99	214
1	43.8	43.32	3	13	30	61	131
1.2	27.87	27.42	2	8	20	38	82
1.4	18.25	17.75	1	6	13	25	54
1.6	12.39	11.87	1	4	9	17	36
1.8	8.67	8.20	1	3	6	12	25
2	6.34	5.78	1	2	5	9	18
2.2	4.72	4.19	1	2	3	6	13
2.4	3.66	3.11	1	1	3	5	10
2.6	2.90	2.34	1	1	2	4	8
2.8	2.38	1.81	1	1	2	3	6
3	2.00	1.41	1	1	1	2	5
4	1.19	0.47	1	1	1	1	2
5	1.02	0.15	1	1	1	1	1
6	1.00	0.04	1	1	1	1	1

It is clear that as the shift in location becomes larger the ARL decreases; for a shift of 2 units in location an OOC signal is produced after 6.34 samples, see Table 5.10. The simulation, Markov chain and exact approach provide very similar results as can be seen from Tables 5.10-5.12.

Table 5.11: Results obtained from the Markov chain approach for the Shewhart $\bar{X}$ control chart.

Shift	ARL	$SDRL$	5th Percentile	25th Percentile	50th Percentile	75th Percentile	95th Percentile
0	370.40	369.90	19	107	257	513	1109
0.2	308.43	307.93	16	89	214	427	923
0.4	200.08	199.57	11	58	139	277	598
0.6	119.67	119.16	7	35	83	166	357
0.8	71.55	71.05	4	21	50	99	213
1	43.89	43.39	3	13	31	61	130
1.2	27.82	27.32	2	8	19	38	82
1.4	18.24	17.74	1	6	13	25	54
1.6	12.38	11.87	1	4	9	17	36
1.8	8.690	8.18	1	3	6	12	25
2	6.30	5.78	1	2	5	9	18
2.2	4.72	4.19	1	2	3	6	13
2.4	3.64	3.11	1	1	3	5	10
2.6	2.90	2.35	1	1	2	4	8
2.8	2.37	1.81	1	1	2	3	6
3	2.00	1.41	1	1	1	2	5
4	1.19	0.47	1	1	1	1	2
5	1.02	0.1	1	1	1	1	1
6	1.00	0.04	1	1	1	1	1

Table 5.12: Results obtained from the Exact approach for the Shewhart $\overline{X}$ control chart.

Shift	ARL	$SDRL$	5th Percentile	25th Percentile	50th Percentile	75th Percentile	95th Percentile
0	370.39	369.90	19	107	257	513	1109
0.2	308.43	307.93	16	89	214	427	923
0.4	200.08	199.57	11	58	139	277	598
0.6	119.67	119.16	7	35	83	166	357
0.8	71.55	71.05	4	21	50	99	213
1	43.89	43.39	3	13	31	61	130
1.2	27.82	27.32	2	8	19	38	82
1.4	18.25	17.74	1	6	13	25	54
1.6	12.38	11.87	1	4	9	17	36
1.8	8.69	8.18	1	3	6	12	25
2	6.30	5.78	1	2	5	9	18
2.2	4.72	4.19	1	2	3	6	13
2.4	3.65	3.11	1	1	3	5	10
2.6	2.90	2.35	1	1	2	4	8
2.8	2.38	1.81	1	1	2	3	6
3	2.00	1.41	1	1	1	2	5
4	1.19	0.47	1	1	1	1	2
5	1.02	0.1	1	1	1	1	0
6	1.00	0.04	1	1	1	1	0

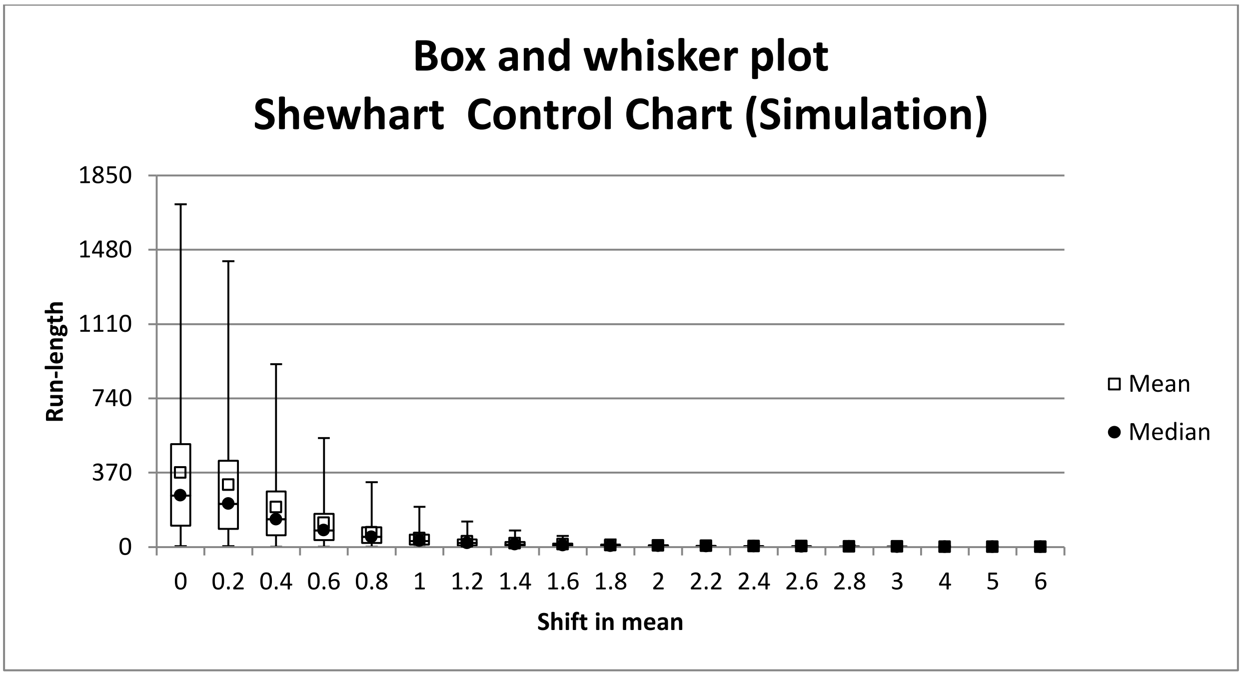

Figure 5.9: Box and whisker like plot for the Shewhart $\overline{X}$ control chart.

The large shift detection capability can be seen from Table 5.10 to Table 5.12 and Figure 5.9 when considering a shift size larger than 2.

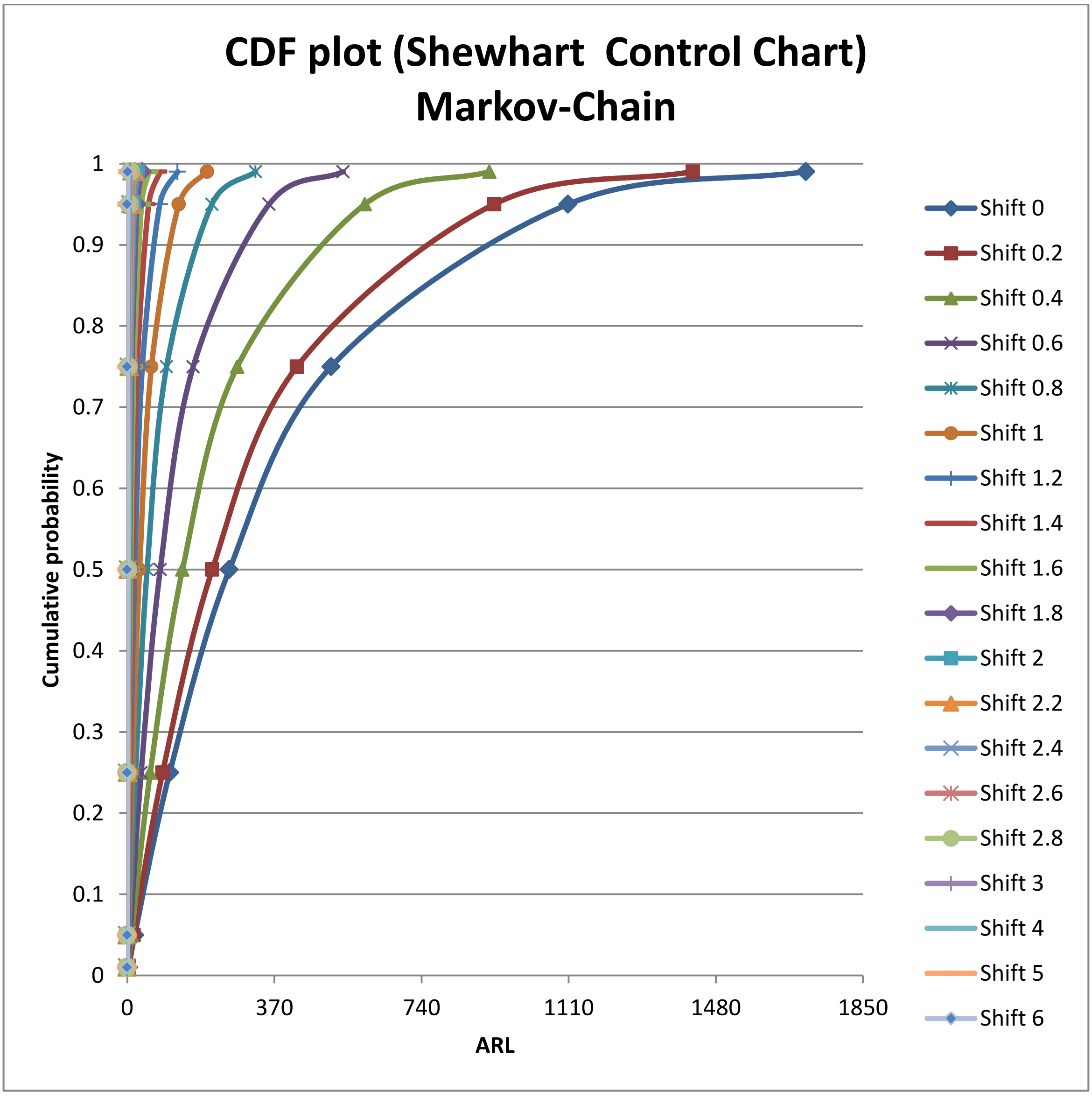

Figure 5.10: CDF plot for the Shewhart $\bar{X}$ control chart.

The purpose of the CDF plot is to provide a sense of stochastic ordering. From Figure 5.10 it is clear that the ARL/MRL decreases as the shift in location increases.

5.2.2 Implementation of the improved *2-of-2* control chart (Khoo (2006)).

The reader is referred to Section 2.1.3 for a summary of the article as well as a visual representation of the charting regions and associated transition probabilities as well as the construction of the essential TPM for the improved *2-of-2* control chart. The advantage of the improved *2-of-2* runs-type signalling rule is that both large and small shift detection capabilities are present. Results for the simulation, Markov chain and exact approaches are provided in Table 5.13 to Table 5.15 when considering observations from a $N(0,1)$ distribution.

Table 5.13: Results obtained from a simulation approach for the improved *2-of-2* control chart.

Shift	ARL	$SDRL$	5th Percentile	25th Percentile	50th Percentile	75th Percentile	95th Percentile
0	370.15	367.50	20	108	258	512	1107
0.2	278.56	276.64	15	81	193	385	831
0.4	151.94	149.95	9	45	106	210	453
0.6	80.41	79.31	5	24	56	111	238
0.8	44	42.71	3	14	31	61	130
1	25.78	25.54	2	8	18	35	74
1.2	16.05	14.92	2	5	11	22	46
1.4	10.61	9.38	2	4	8	14	29
1.6	7.4	6.25	2	3	5	10	20
1.8	5.45	4.35	1	2	4	7	14
2	4.22	3.12	1	2	3	5	10
2.2	3.39	2.32	1	2	3	4	8
2.4	2.82	1.77	1	2	2	4	6
2.6	2.41	1.38	1	2	2	3	5
2.8	2.11	1.1	1	1	2	2	4
3	1.89	0.9	1	1	2	2	4
4	1.28	0.48	1	1	1	2	2
5	1.05	0.23	1	1	1	1	2
6	1	0.07	1	1	1	1	1

Table 5.14: Results obtained from a Markov chain approach for the improved *2-of-2* control chart.

Shift	ARL	$SDRL$	5th Percentile	25th Percentile	50th Percentile	75th Percentile	95th Percentile
0	370.60	369.30	20	107	257	513	1108
0.2	278.70	277.50	15	81	194	386	832
0.4	152.40	151.20	9	45	106	211	454
0.6	79.97	78.75	5	24	56	110	237
0.8	43.91	42.7	3	13	31	60	129
1	25.67	24.48	2	8	18	35	75
1.2	15.99	14.82	2	5	11	22	46
1.4	10.58	9.44	2	4	8	14	29
1.6	7.41	6.29	2	3	5	10	20
1.8	5.46	4.36	1	2	4	7	14
2	4.21	3.13	1	2	3	5	10
2.2	3.39	2.32	1	2	3	4	8
2.4	2.81	1.76	1	2	3	4	6
2.6	2.41	1.37	1	2	3	3	5
2.8	2.11	1.10	1	1	2	3	4
3	1.89	0.91	1	1	2	3	4
4	1.28	0.47	1	1	1	2	4
5	1.06	0.23	1	1	1	1	2
6	1.01	0.07	1	1	1	1	1

When comparing the ARL performance for the improved *2-of-2* to the Shewhart $\overline{\mathrm{X}}$ control chart for a small shift, e.g. 0.4, an improvement is noted; $ARL_{0.4} = 152$ compared to the Shewhart $\overline{\mathrm{X}}$ control chart with $ARL_{0.4} = 198.51$. Note that similar results are noted for shifts up to 5 units from the mean.

Table 5.15: Results obtained from the Exact approach for the improved *2-of-2* control chart.

Shift	*ARL*	*SDRL*	5th Percentile	25th Percentile	50th Percentile	75th Percentile	95th Percentile
0	370.6	369.3	20	107	257	513	1108
0.2	278.7	277.5	15	81	194	386	832
0.4	152.4	151.2	9	45	106	211	454
0.6	79.97	78.75	5	24	56	110	237
0.8	43.91	42.7	3	13	31	60	129
1	25.67	24.48	0	8	18	35	75
1.2	15.99	14.82	2	5	11	22	46
1.4	10.58	9.439	2	4	8	14	29
1.6	7.41	6.289	2	3	5	10	20
1.8	5.461	4.359	1	2	4	7	14
2	4.214	3.129	1	2	3	5	10
2.2	3.385	2.315	1	2	3	4	8
2.4	2.814	1.761	1	2	3	4	6
2.6	2.41	1.374	1	2	3	3	5
2.8	2.113	1.1	1	1	2	3	4
3	1.89	0.906	1	1	2	3	4
4	1.284	0.474	1	1	1	2	4
5	1.055	0.228	1	1	1	1	2
6	1.005	0.068	1	1	1	1	1

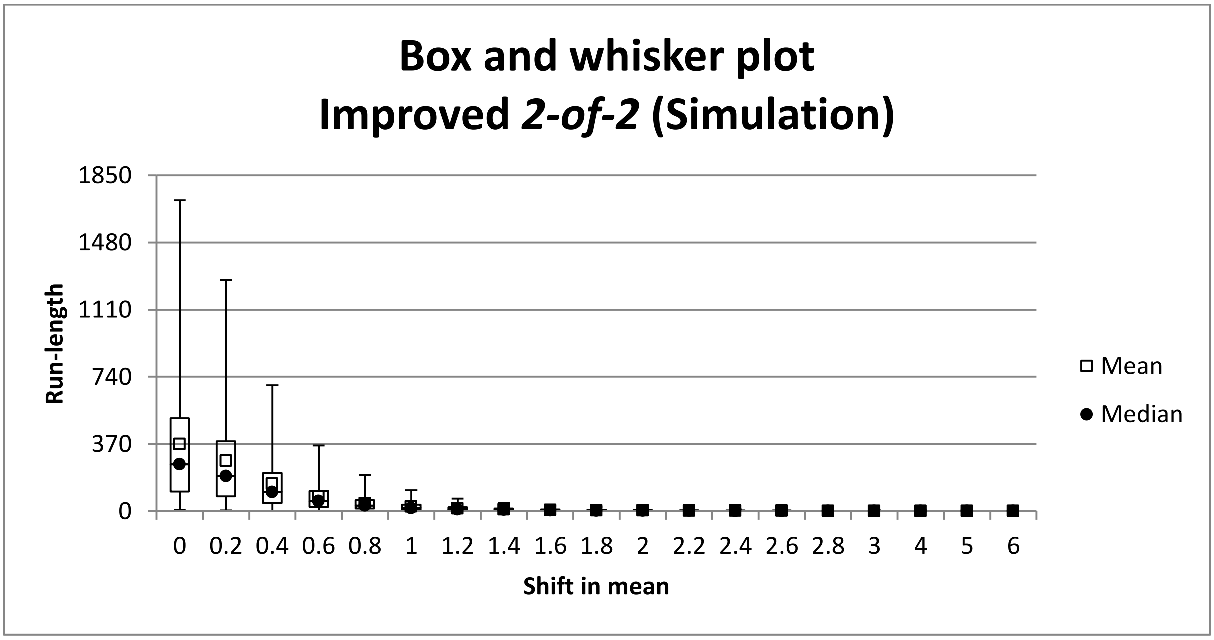

Figure 5.11: Box and whisker like plot for the improved *2-of-2* control chart.

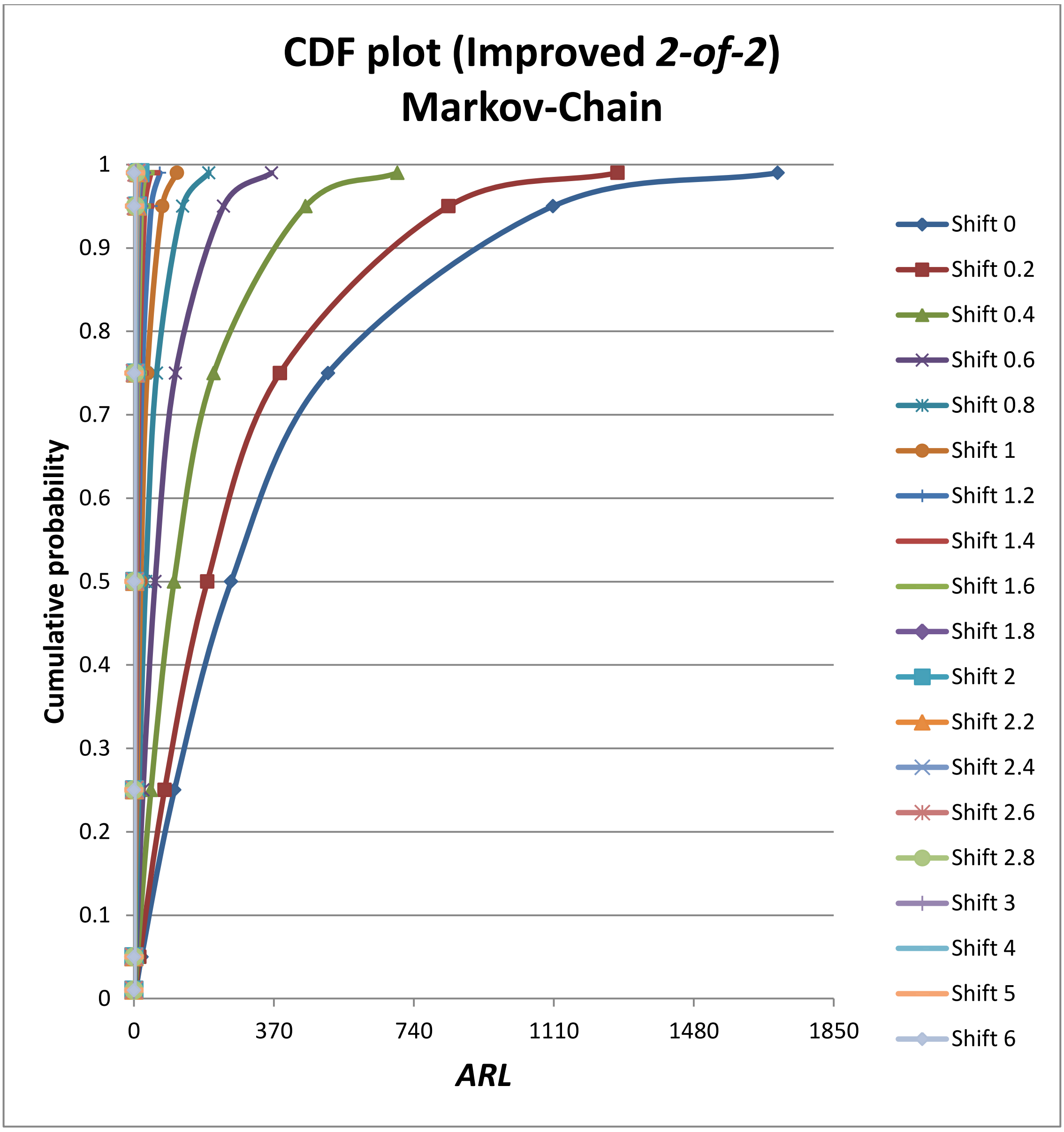

Figure 5.12: CDF plot for the improved *2-of-2* control chart.

Stochastic ordering is present as was the case for the Shewhart $\overline{\text{X}}$ control chart in Figure 5.10.

5.2.3 Implementation of the revised *2-of-3* control chart (Antzoulakos and Rakitzis (2008a)).

The reader is referred to Section 2.1.5 for a summary of the article as well as a visual representation of the charting regions and associated transition probabilities as well as the construction of the essential TPM for the revised *2-of-3* control chart. For the revised *2-of-3* control chart, both small and large shift detection capabilities are present, similar to the improved *2-of-2* control chart. Note that improved OOC performance is noted compared to the improved *2-of-2* control chart, see Table 5.16 to Table 5.18. Note that observations from a $N(0,1)$ distribution were considered in calculating the results noted in Table 5.16 to Table 5.18.

Table 5.16: Results obtained from a simulation approach for the revised *2-of-3* control chart.

Shift	ARL	$SDRL$	5^{th} Percentile	25^{th} Percentile	50^{th} Percentile	75^{th} Percentile	95^{th} Percentile
0	367.39	360.42	27	111	254	509	1097.5
0.2	269.14	274.66	17	82.5	196.5	367	790
0.4	135.13	138.95	8	36.5	89.5	187	400
0.6	68.3	67.75	5	20	47.5	93	199
0.8	39.34	37.27	4	12	28	55	107
1	21.79	19.78	3	7	15	30	62
1.2	14.07	12.44	3	5	10	19	40
1.4	9.57	8.62	3	4	6	12	26
1.6	7.17	5.6	3	3	5	9	18
1.8	5.14	3.51	2	3	4	6	12
2	4.12	2.47	1	3	3	5	9
2.2	3.48	1.81	1	3	3	4	7
2.4	3.03	1.42	1	2	3	3	6
2.6	2.7	1.18	1	2	3	3	5
2.8	2.45	1.04	1	2	3	3	4
3	2.22	0.97	1	1	2	3	3
4	1.4	0.66	1	1	1	2	3
5	1.07	0.27	1	1	1	1	2
6	1.01	0.08	1	1	1	1	1

Table 5.17: Results obtained from a Markov chain approach for the revised *2-of-3* control chart.

Shift	ARL	$SDRL$	5th Percentile	25th Percentile	50th Percentile	75th Percentile	95th Percentile
0	370.93	369.38	20	108	258	514	1108
0.2	267.66	266.09	15	78	186	370	799
0.4	137.76	136.18	9	41	96	190	410
0.6	69.45	67.88	5	21	49	96	205
0.8	37.34	35.81	3	12	26	51	109
1	21.69	20.21	3	7	15	29	62
1.2	13.59	12.16	2	5	10	18	38
1.4	9.13	7.74	2	4	7	12	25
1.6	6.53	5.17	2	3	5	9	17
1.8	4.92	3.60	2	3	4	6	12
2	3.89	2.60	1	2	3	5	9
2.2	3.20	1.94	1	2	3	4	7
2.4	2.72	1.49	1	2	3	3	6
2.6	2.37	1.18	1	2	3	3	5
2.8	2.11	0.97	1	2	3	3	4
3	1.91	0.82	1	1	2	3	3
4	1.32	0.49	1	1	1	2	3
5	1.07	0.25	1	1	1	1	2
6	1.01	0.08	1	1	1	1	1

Table 5.18: Results obtained from the Exact approach for the revised *2-of-3* control chart.

Shift	*ARL*	*SDRL*	5th Percentile	25th Percentile	50th Percentile	75th Percentile	95th Percentile
0	370.93	369.38	20	108	258	514	1108
0.2	267.66	266.09	15	78	186	370	799
0.4	137.76	136.18	9	41	96	190	410
0.6	69.45	67.88	5	21	49	96	205
0.8	37.34	35.81	3	12	26	51	109
1	21.69	20.21	3	7	15	29	62
1.2	13.59	12.16	2	5	10	18	38
1.4	9.13	7.74	2	4	7	12	25
1.6	6.53	5.17	2	3	5	9	17
1.8	4.92	3.60	2	3	4	6	12
2	3.89	2.60	1	2	3	5	9
2.2	3.20	1.94	1	2	3	4	7
2.4	2.72	1.49	1	2	3	3	6
2.6	2.37	1.18	1	2	3	3	5
2.8	2.11	0.97	1	2	3	3	4
3	1.91	0.82	1	1	2	3	3
4	1.32	0.49	1	1	1	2	3
5	1.07	0.25	1	1	1	1	2
6	1.01	0.08	1	1	1	1	1

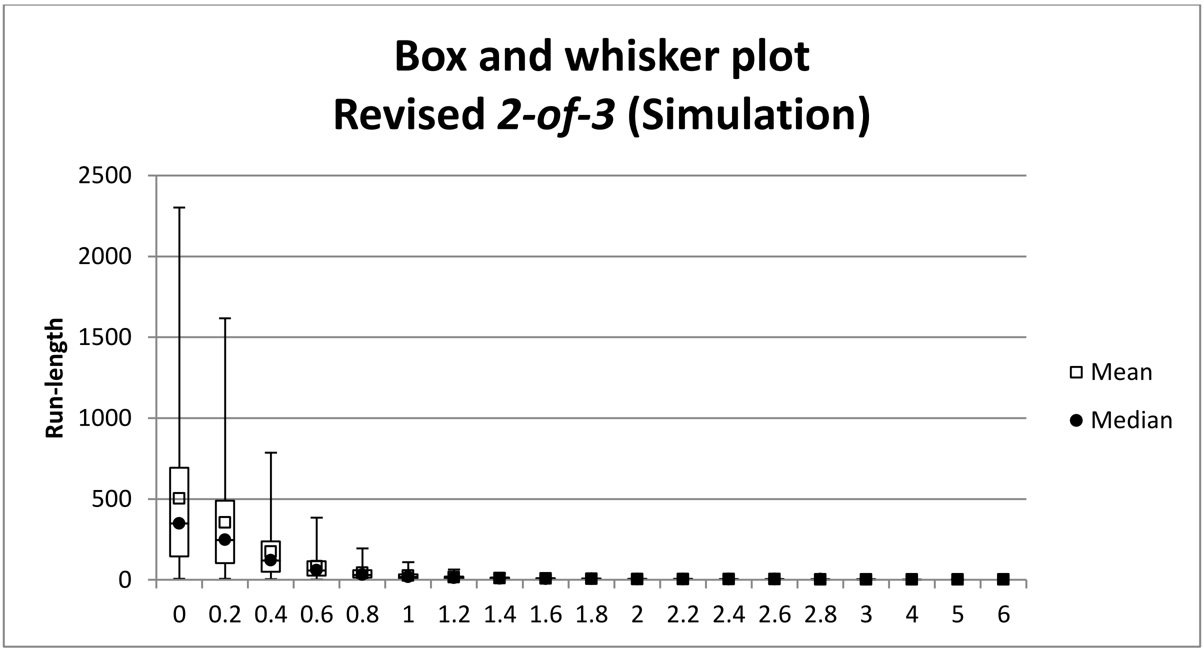

Figure 5.15: Box and whisker like plot for the revised *2-of-3* control chart.

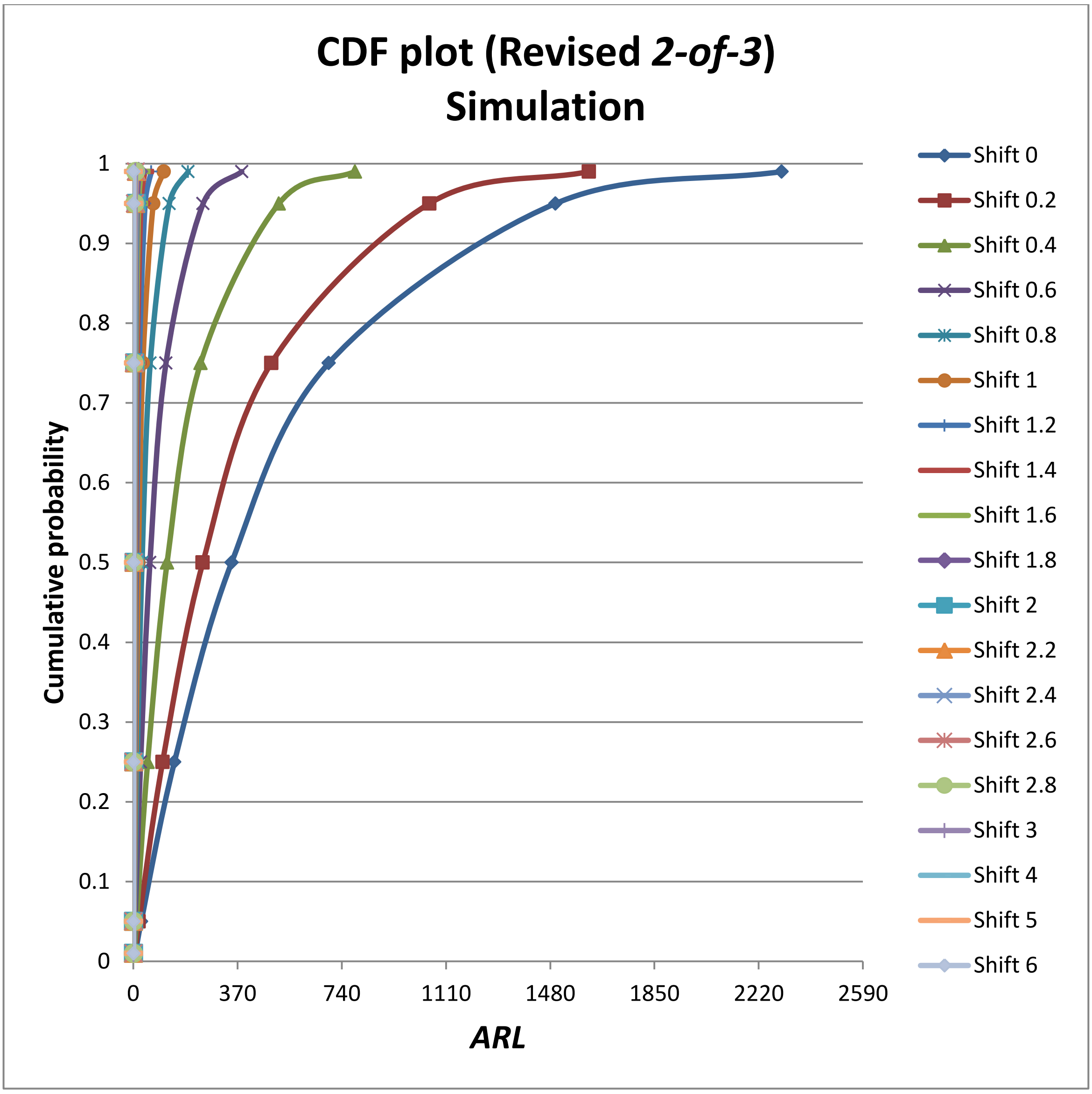

Figure 5.16: CDF plot for the revised *2-of-3* control chart.

Stochastic ordering is again present for the revised *2-of-3* control chart, see Figure 5.16.

5.2.4 Implementation of the *S*-chart supplemented with the *2-of-2* runs-type signalling rule (Antzoulakos and Rakitzis (2010)).

The reader is referred to Section 2.2.4 for a summary of the article discussed below as well as a visual representation of the charting regions and associated probabilities as well as the construction of the essential TPM for the S chart supplemented with the *2-of-2* runs-type signalling rule; note that for discussion purposes the two-sided control chart would be evaluated. Section 5.2.4 evaluates the S chart supplemented with the *2-of-2* runs-type rule. It is noted that the ARL_0 for the chart is 226.24, which is less than the standard ARL_0 noted for control charts based on location. When considering control charts based on location the ARL_0 is usually in the region of 370. Note that the distribution of the standard deviation is skewed to the right leading to a heavier tail and increased signalling events. Note that a decreased ARL_0 for control charts based on spread is noted in the literature. Results for the simulation, Markov chain and exact approaches are provided in Table 5.19 to Table 5.21 when considering observations from a $N(0,1)$ distribution.

Table 5.19: Results obtained from a simulation approach for the *2-of-2* S-chart.

Shift	ARL	$SDRL$	5^{th} Percentile	25^{th} Percentile	50^{th} Percentile	75^{th} Percentile	95^{th} Percentile
0.2	2.00	0.09	2	2	2	2	2
0.4	4.03	2.68	2	2	3	5	9
0.6	19.73	18.52	2	7	14	27	56
0.8	102.75	101.57	7	30	72	142	306
1	226.24	224.57	13	66	158	313	674
1.2	39.87	38.94	3	12	28	55	117
1.4	11.22	10.32	1	4	8	15	32
1.6	5.35	4.51	1	2	4	7	14
1.8	3.38	2.60	1	2	3	4	9
2	2.50	1.76	1	1	2	3	6
2.2	2.03	1.30	1	1	2	2	5
2.4	1.73	1.01	1	1	1	2	4
2.6	1.55	0.84	1	1	1	2	3
2.8	1.43	0.72	1	1	1	2	3
3	1.33	0.61	1	1	1	2	2
4	1.12	0.35	1	1	1	1	2
5	1.05	0.23	1	1	1	1	2
6	1.03	0.17	1	1	1	1	1

The shift referred to in Table 5.19 to Table 5.21 is defined as shift $= {}^{\sigma_1}/_{\sigma_0}$; the IC state would be where the shift $= 1$. Note that both increases and decreases in spread are being monitored where shift >1 or shift <1.

Table 5.20: Results obtained from a Markov chain approach for the *2-of-2* S-chart.

Shift	ARL	$SDRL$	5th Percentile	25th Percentile	50th Percentile	75th Percentile	95th Percentile
0.2	2.00	0.09	2	2	2	2	2
0.4	4.01	2.68	2	2	3	5	9
0.6	19.75	18.41	2	7	14	27	56
0.8	102.56	101.15	7	31	72	142	304
1	226.28	225.04	13	66	157	313	675
1.2	39.78	38.82	3	12	28	55	117
1.4	11.23	10.35	1	4	8	15	32
1.6	5.36	4.54	1	2	4	7	14
1.8	3.38	2.61	1	2	3	4	9
2	2.50	1.75	1	1	2	3	6
2.2	2.02	1.29	1	1	2	3	5
2.4	1.74	1.02	1	1	1	2	4
2.6	1.55	0.84	1	1	1	2	3
2.8	1.43	0.71	1	1	1	2	3
3	1.33	0.61	1	1	1	2	2
4	1.12	0.35	1	1	1	1	2
5	1.05	0.24	1	1	1	1	2
6	1.03	0.17	1	1	1	1	1

Table 5.21: Results obtained from the exact approach for the *2-of-2* S-chart.

Shift	ARL	$SDRL$	5th Percentile	25th Percentile	50th Percentile	75th Percentile	95th Percentile
0.2	2.00	0.09	2	2	2	2	2
0.4	4.01	2.68	2	2	3	5	9
0.6	19.75	18.41	2	7	14	27	56
0.8	102.56	101.15	7	31	72	142	304
1	226.28	225.04	13	66	157	313	675
1.2	39.78	38.82	3	12	28	55	117
1.4	11.23	10.35	1	4	8	15	32
1.6	5.36	4.54	1	2	4	7	14
1.8	3.38	2.61	1	2	3	4	9
2	2.50	1.75	1	1	2	3	6
2.2	2.02	1.29	1	1	2	3	5
2.4	1.74	1.02	1	1	1	2	4
2.6	1.55	0.84	1	1	1	2	3
2.8	1.43	0.71	1	1	1	2	3
3	1.33	0.61	1	1	1	2	2
4	1.12	0.35	1	1	1	1	2
5	1.05	0.24	1	1	1	1	2
6	1.03	0.17	1	1	1	1	1

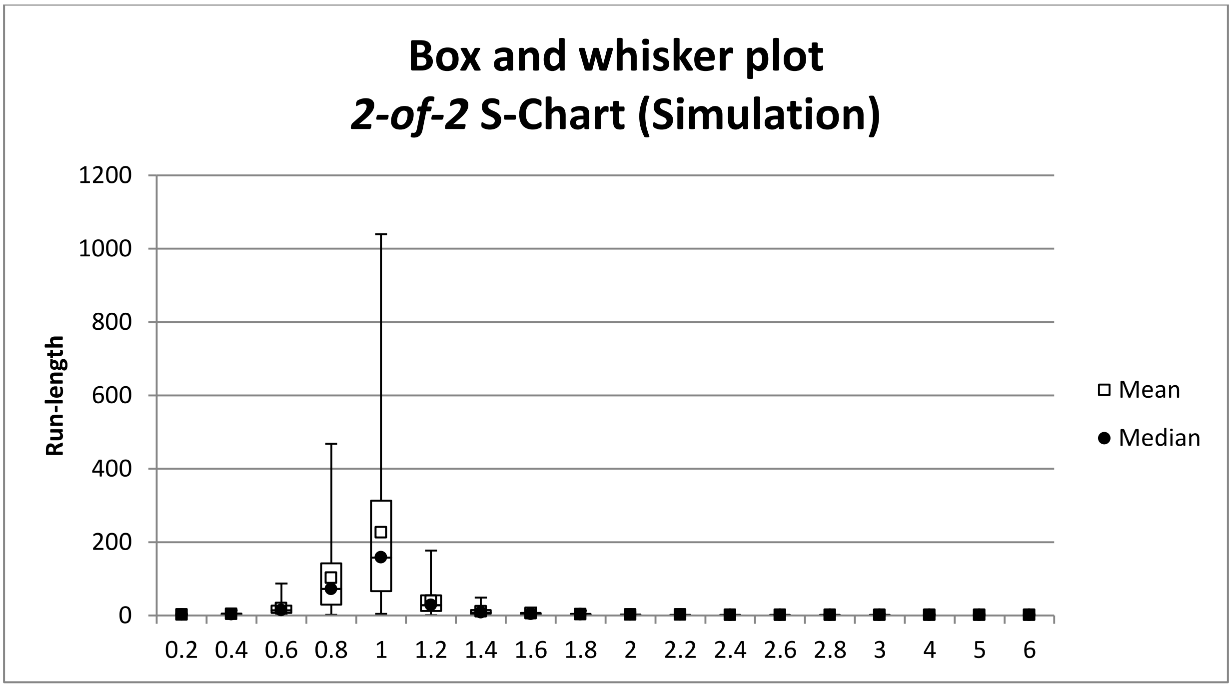

Figure 5.17: Box and whisker like plot for the *2-of-2* S-chart.

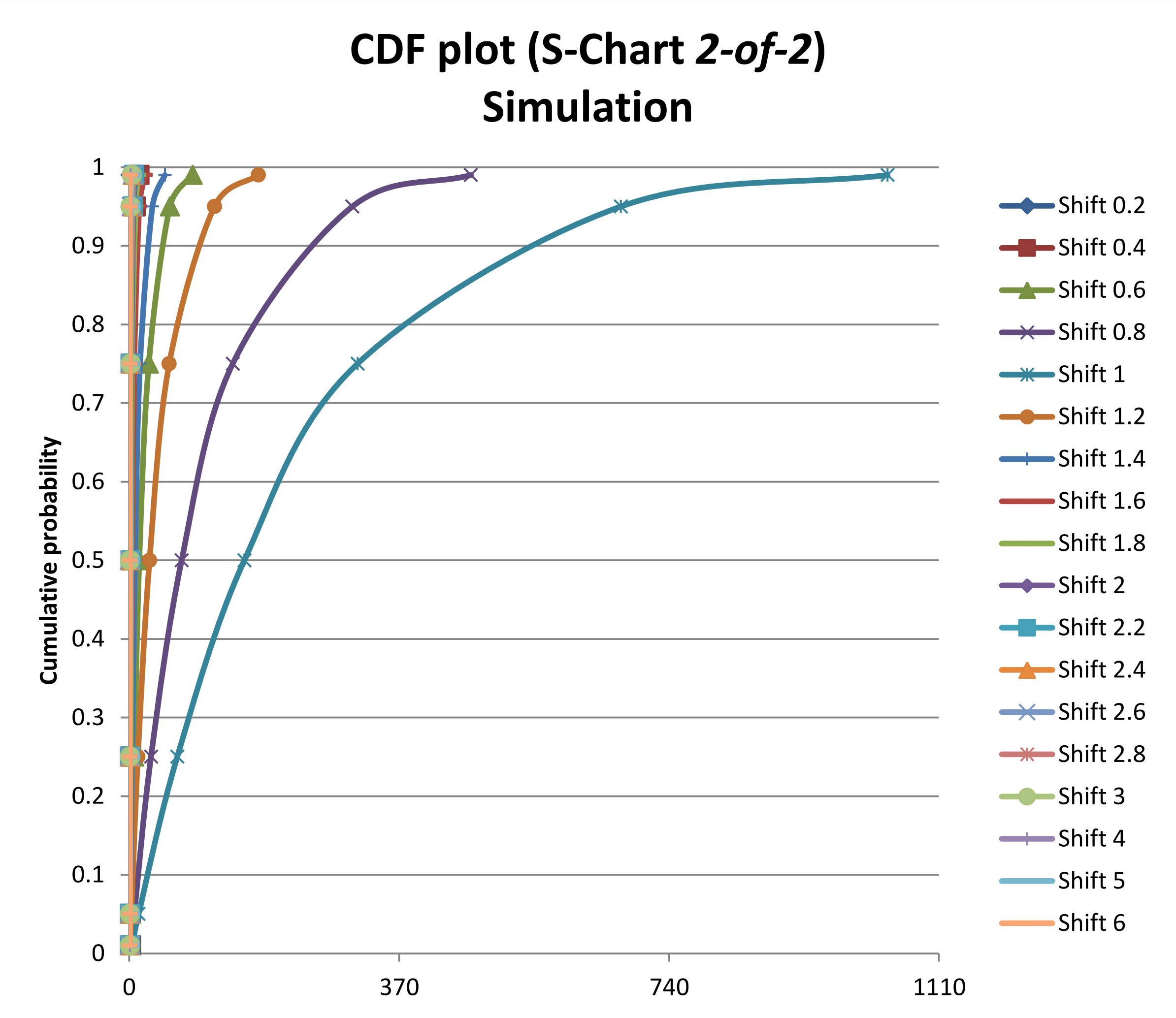

Figure 5.18: CDF plot for the revised *2-of-2* S-chart.

5.2.5 Implementation of the nonparametric *2-of-2* control chart (Human et al. (2010)).

The reader is referred to Section 3.1.3 for a summary of the article discussed below as well as a visual representation of the charting regions and associated probabilities as well as the construction of the essential TPM for the nonparametric control chart based on the sign statistic supplemented with the *2-of-2* runs-type signalling rule.

By definition nonparametric charts have the same ARL_0 for multiple distributions, this being a drawback of parametric charts where the ARL_0 could decrease when considering non normal distributions. The *2-of-2* nonparametric chart was designed with an $ARL_0 = 528$ with superior OOC signalling capabilities compared to the Shewhart $\overline{\text{X}}$ control chart. The ARL

results and distributional properties of the run-length are provided in Table 5.22 to Table 5.24 when considering observations from a $N(0,1)$ distribution.

Table 5.22: Results obtained from a simulation approach for the nonparametric *2-of-2* control chart.

Shift	ARL	$SDRL$	5th Percentile	25th Percentile	50th Percentile	75th Percentile	95th Percentile
0	526.01	523.82	29	152	364	729	1575
0.2	240.72	239.51	14	70	167	334	719
0.4	76.46	75.16	5	23	53	106	227
0.6	29.57	28.25	3	9	21	40	86
0.8	14.11	12.74	2	5	10	19	40
1	7.99	6.65	2	3	6	11	21
1.2	5.21	3.89	2	2	4	7	13
1.4	3.84	2.51	2	2	3	5	9
1.6	3.09	1.71	2	2	2	4	7
1.8	2.64	1.21	2	2	2	3	5
2	2.38	0.87	2	2	2	2	4
2.2	2.23	0.65	2	2	2	2	4
2.4	2.13	0.48	2	2	2	2	3
2.6	2.07	0.35	2	2	2	2	2
2.8	2.04	0.26	2	2	2	2	2
3	2.02	0.19	2	2	2	2	2
4	2.00	0.03	2	2	2	2	2
5	2.00	-	2	2	2	2	2
6	2.00	-	2	2	2	2	2

Table 5.23: Results obtained from a Markov-chain approach for the nonparametric *2-of-2* control chart.

Shift	ARL	$SDRL$	5th Percentile	25th Percentile	50th Percentile	75th Percentile	95th Percentile
0	528.00	526.53	28	153	366	731	1579
0.2	240.12	238.68	14	70	167	332	716
0.4	76.49	75.08	5	23	53	105	226
0.6	29.63	28.27	3	9	21	41	86
0.8	14.10	12.77	2	5	10	19	40
1	8.00	6.69	2	3	6	11	21
1.2	5.24	3.93	2	2	4	7	13
1.4	3.84	2.51	2	2	3	5	9
1.6	3.08	1.70	2	2	2	4	7
1.8	2.64	1.21	2	2	2	3	5

2	2.38	0.88	2	2	2	2	4
2.2	2.22	0.65	2	2	2	2	4
2.4	2.13	0.48	2	2	2	2	3
2.6	2.07	0.35	2	2	2	2	2
2.8	2.04	0.26	2	2	2	2	2
3	2.02	0.19	2	2	2	2	2
4	2.00	0.03	2	2	2	2	2
5	2.00	0.00	2	2	2	2	2
6	2.00	0.00	2	2	2	2	2

Table 5.24: Results obtained from the exact approach for the nonparametric *2-of-2* control chart.

Shift	ARL	$SDRL$	5th Percentile	25th Percentile	50th Percentile	75th Percentile	95th Percentile
0	528.00	526.53	28	153	366	731	1579
0.2	240.12	238.68	14	70	167	332	716
0.4	76.49	75.08	5	23	53	105	226
0.6	29.63	28.27	3	9	21	41	86
0.8	14.10	12.77	2	5	10	19	40
1	8.00	6.69	2	3	6	11	21
1.2	5.24	3.93	2	2	4	7	13
1.4	3.84	2.51	2	2	3	5	9
1.6	3.08	1.70	2	2	2	4	7
1.8	2.64	1.21	2	2	2	3	5
2	2.38	0.88	2	2	2	2	4
2.2	2.22	0.65	2	2	2	2	4
2.4	2.13	0.48	2	2	2	2	3
2.6	2.07	0.35	2	2	2	2	2
2.8	2.04	0.26	2	2	2	2	2
3	2.02	0.19	2	2	2	2	2
4	2.00	0.03	2	2	2	2	2
5	2.00	0.00	2	2	2	2	2
6	2.00	0.00	2	2	2	2	2

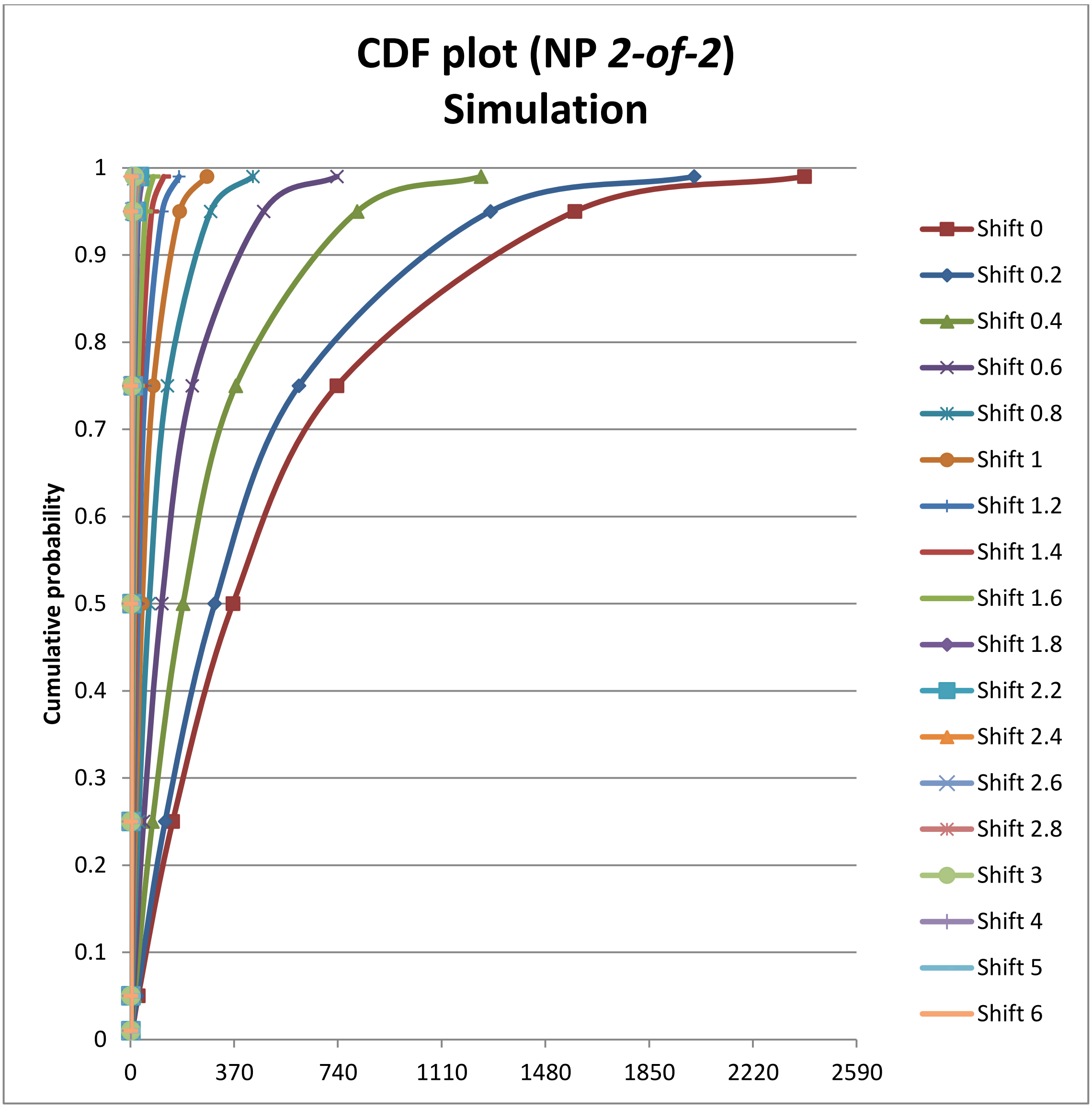

Figure 5.19: CDF plot for the nonparametric *2-of-2* chart.

Stochastic ordering is again noted in Figure 5.19; note that as the shift in location increases the ARL/MRL decreases.

5.2.6 Implementation of the parametric *2-of-2* and *2-of-3* control chart when process parameters are unknown (Zhang et al. (2010)).

The reader is referred to Section 2.1.9 for a summary of the article discussed below as well as a visual representation of the charting regions and associated probabilities as well as the construction of the essential TPM for the parametric control chart supplemented with the *2-of-2* and *2-of-3* runs-type signalling rules when the process parameters are unknown. The process parameters are estimated in a phase I analysis after which a phase II analysis is conducted in order to establish whether the process is IC or OOC. Note that the observations are assumed to be distributed $N(\mu, \sigma)$, where μ and σ are unknown parameters that are estimated from a phase I analysis, see Section 2.1.9. The ARL results and other distributional properties of the run-length are provided in Table 5.25 to Table 5.28.

Table 5.25: Results obtained from the Markov chain approach for the parametric *2-of-3* control chart with unknown process parameters.

Shift	ARL	$SDRL$	5^{th} Percentile	25^{th} Percentile	50^{th} Percentile	75^{th} Percentile	95^{th} Percentile
0	330.68	328.78	19	96	230	458	987
0.2	98.14	96.36	7	30	69	135	290
0.4	26.81	25.14	3	9	19	37	77
0.6	10.15	8.58	2	4	7	13	27
0.8	5.19	3.69	2	3	4	7	13
1	3.37	1.85	2	2	3	4	7
1.2	2.60	1.01	2	2	2	3	5
1.4	2.25	0.58	2	2	2	2	3
1.6	2.10	0.33	2	2	2	2	3
1.8	2.03	0.19	2	2	2	2	2
2	2.01	0.10	2	2	2	2	2
2.2	2.00	0.05	2	2	2	2	2
2.4	2.00	0.02	2	2	2	2	2
2.6	2.00	0.01	2	2	2	2	2
2.8	2.00	0.00	2	2	2	2	2
3	2.00	0.00	2	2	2	2	2
4	2.00	0.00	2	2	2	2	2
5	2.00	-	2	2	2	2	2
6	2.00	-	2	2	2	2	2

Table 5.26: Results obtained from the Simulation approach for the parametric *2-of-3* control chart with unknown process parameters.

Shift	ARL	$SDRL$	5^{th} Percentile	25^{th} Percentile	50^{th} Percentile	75^{th} Percentile	95^{th} Percentile
0	376.80	380.80	21	108	258	519	1131
0.2	272.80	270.90	16	81	192	375	813
0.4	141.98	140.16	10	42	99	196	422
0.6	71.76	70.61	5	22	50	98	210
0.8	39.22	37.25	4	12	27	54	114
1	23.39	21.48	3	8	17	32	67

1.2	14.55	12.78	3	5	11	20	40
1.4	10.04	8.26	3	4	7	13	27
1.6	7.37	5.64	3	3	5	9	19
1.8	5.65	3.74	3	3	4	7	13
2	4.65	2.64	3	3	3	5	10
2.2	3.99	1.87	3	3	3	4	8
2.4	3.56	1.29	3	3	3	3	6
2.6	3.34	0.94	3	3	3	3	5
2.8	3.17	0.63	3	3	3	3	4
3	3.09	0.43	3	3	3	3	4
4	3.00	0.05	3	3	3	3	3
5	3.00	-	3	3	3	3	3
6	3.00	-	3	3	3	3	3

Table 5.27: Results obtained from the Markov chain approach for the parametric *2-of-2* control chart with unknown process parameters.

Shift	ARL	$SDRL$	5th Percentile	25th Percentile	50th Percentile	75th Percentile	95th Percentile
0	207.78	203.74	12	61	144	287	620
0.2	89.92	88.51	6	27	63	124	267
0.4	25.93	24.57	3	8	18	35	75
0.6	10.10	8.78	2	4	7	13	28
0.8	5.22	3.91	2	2	4	7	13
1	3.38	2.02	2	2	2	4	7
1.2	2.60	1.15	2	2	2	3	5
1.4	2.24	0.68	2	2	2	2	4
1.6	2.09	0.40	2	2	2	2	3
1.8	2.03	0.22	2	2	2	2	2
2	2.01	0.11	2	2	2	2	2
2.2	2.00	0.05	2	2	2	2	2
2.4	2.00	0.02	2	2	2	2	2
2.6	2.00	0.01	2	2	2	2	2
2.8	2.00	0.00	2	2	2	2	2
3	2.00	0.00	2	2	2	2	2
4	2.00	0.00	2	2	2	2	2
5	2.00	-	2	2	2	2	2
6	2.00	-	2	2	2	2	2

Table 5.28: Results obtained from the Simulation approach for the parametric *2-of-2* control chart with unknown process parameters.

Shift	ARL	$SDRL$	5th Percentile	25th Percentile	50th Percentile	75th Percentile	95th Percentile
0	206.61	207.91	12	60	142	283	633
0.2	174.29	170.73	11	52	123	244	510
0.4	104.16	102.65	7	31	73	143	311
0.6	57.26	56.18	4	17	40	80	167
0.8	33.01	31.22	3	10	24	46	96
1	20.43	19.01	2	7	15	28	59
1.2	13.15	11.96	2	5	9	18	37
1.4	9.15	7.67	2	4	7	12	25
1.6	6.70	5.41	2	3	5	9	17
1.8	5.14	3.85	2	2	4	7	13
2	4.13	2.73	2	2	3	5	10
2.2	3.46	2.07	2	2	3	4	8
2.4	3.03	1.64	2	2	2	4	6
2.6	2.69	1.28	2	2	2	3	5
2.8	2.47	1.00	2	2	2	2	4
3	2.31	0.77	2	2	2	2	4
4	2.03	0.22	2	2	2	2	2
5	2.00	0.06	2	2	2	2	2
6	2.00	0.01	2	2	2	2	2

Decreased ARL performance is noted when the process parameters are estimated compared to the cases where the process parameters are known, see Table 5.25 to Table 5.28.

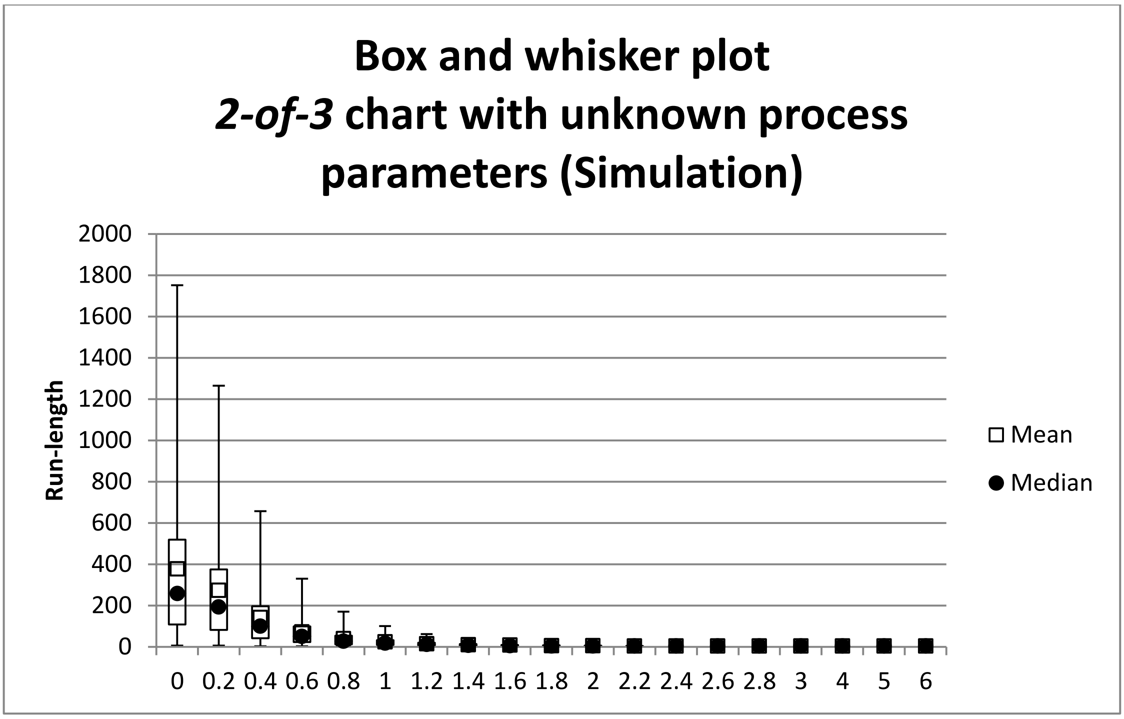

Figure 5.20: Box and whisker like plot for the *2-of-3* control chart with unknown process parameters.

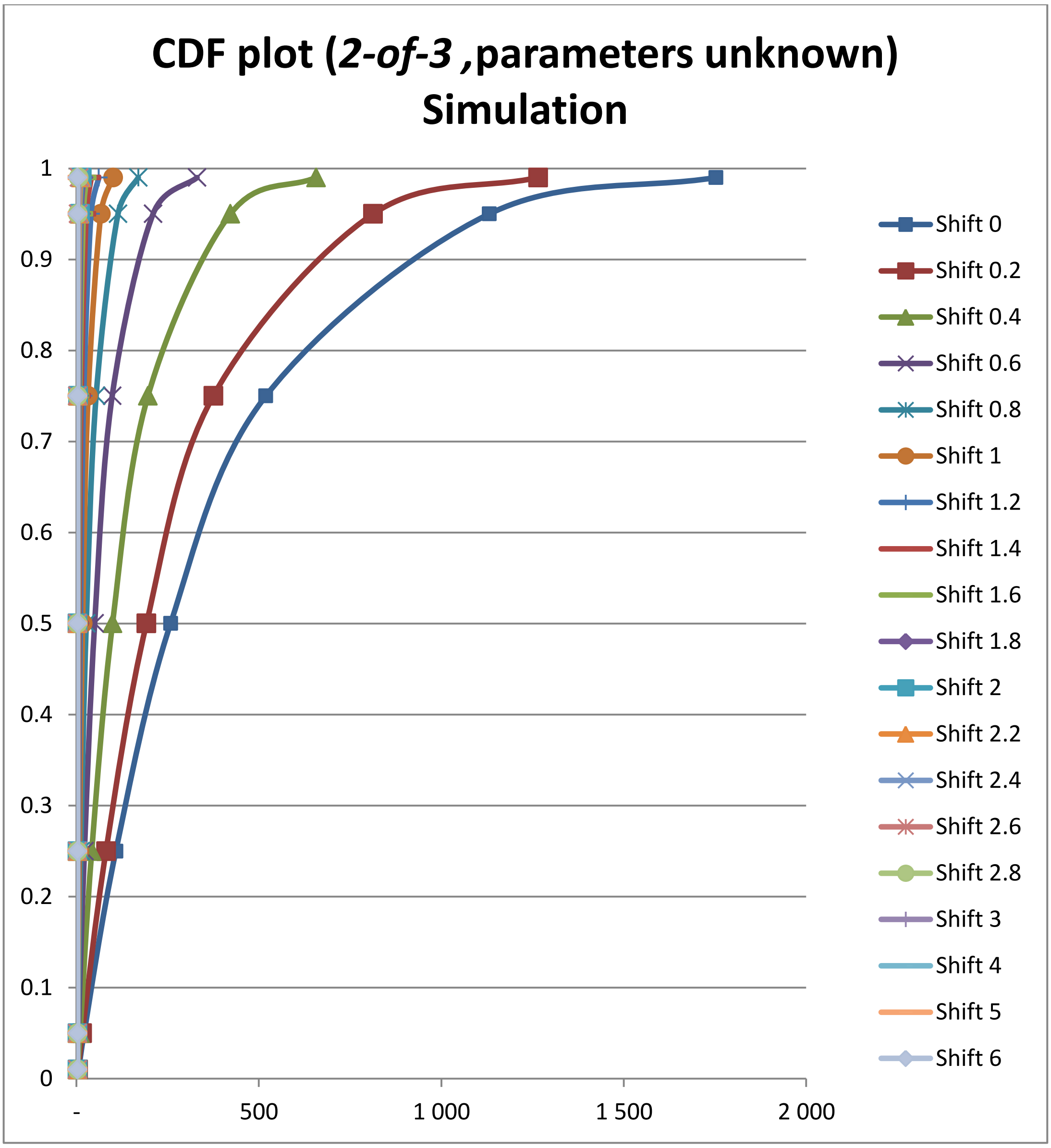

Figure 5.20: CDF plot for the parametric *2-of-3* chart where the parameters of the distribution is unknown.

Stochastic ordering is again present in Figure 5.20, with a decrease in the ARL/MRL noted as the shift in the process being monitored increases.

To conclude, Section 5.2 focused on a simulation, Markov chain and exact approach for the Shewhart $\overline{\mathrm{X}}$ control chart, the improved *2-of-2* control chart, the revised *2-of-3* control chart, the *S* chart supplemented with the *2-of-2* runs-type signalling rule, the nonparametric control chart supplemented with the *2-of-2* runs-type signalling rule as well as the *2-of-3* (*2-of-2*) control chart with unknown process parameters. Distributional properties of the run-length are provided as well as various graphical illustrations to highlight ARL performance. Chapter 6 concludes the thesis and provides a summary regarding runs-type signalling rules.

Chapter 6

Conclusion

One of the main aims of statistical process control (SPC) is to detect and reduce assignable causes of variation. Shewhart $\bar{X}$ and S charts were developed to detect large shifts in either location or spread but one drawback was the inefficiency to detect small shifts in location or spread. The aim of runs-type signalling rules is to detect these small shifts or a combination of small and large shifts (Khoo et al. (2006)) in either location or spread in both the parametric and nonparametric settings effectively. It should be noted that runs-type signalling rules were initially developed to detect non-random patterns in the process (Montgomery (2013)). Although EWMA and CUSUM control charts supplemented with runs-type signalling rules are available in the research, see e.g. Riaz et al. (2011) or Abbas et al. (2011), these type of control charts are not considered since the focus of the thesis is Shewhart-type control charts supplemented with runs-type signalling rules.

Sensitizing rules were proposed to increase the sensitivity of control charts (see, for example, Page (1955), the Western Electric Company (1956), Roberts (1958), Bissell (1978), Wheeler (1983) and Nelson (1984)). The Western Electric sensitizing rules used warning limits and considered, in most cases, a sequence of points in order to produce a warning signal. Runs-type signalling rules took the lead and control charts were developed whereby an OOC signal is produced if a sequence of charting statistics plots within or on or above/below certain control limits. Some examples of control charts incorporating runs-type signalling rules are: Lowry et al. (1995), Klein (2000a), Khoo (2003), Khoo et al. (2004), Aparisi et al. (2004), Koutras (2006), Khoo (2007), Koutras (2007), Acosta-Mejia (2007), Bakir (2004), Chen (2007), Chakraborti and Eryilmaz (2007), Antzoulakos and Rakitzis (2008a), Antzoulakos and Rakitzis (2008b), Chakraborti et al. (2009), Acosta-Mejia (2008), Das (2008b), Antzoulakos and Rakitzis (2010), Balakrishnan et al. (2010), Lee (2013), Antzoulakos and Rakitzis (2014).

Parametric and nonparametric control charts monitoring location and spread supplemented with runs-type signalling rules were discussed as well as univariate and multivariate control charts. A graphical illustration of the various chapters is provided in Figure 6.1.

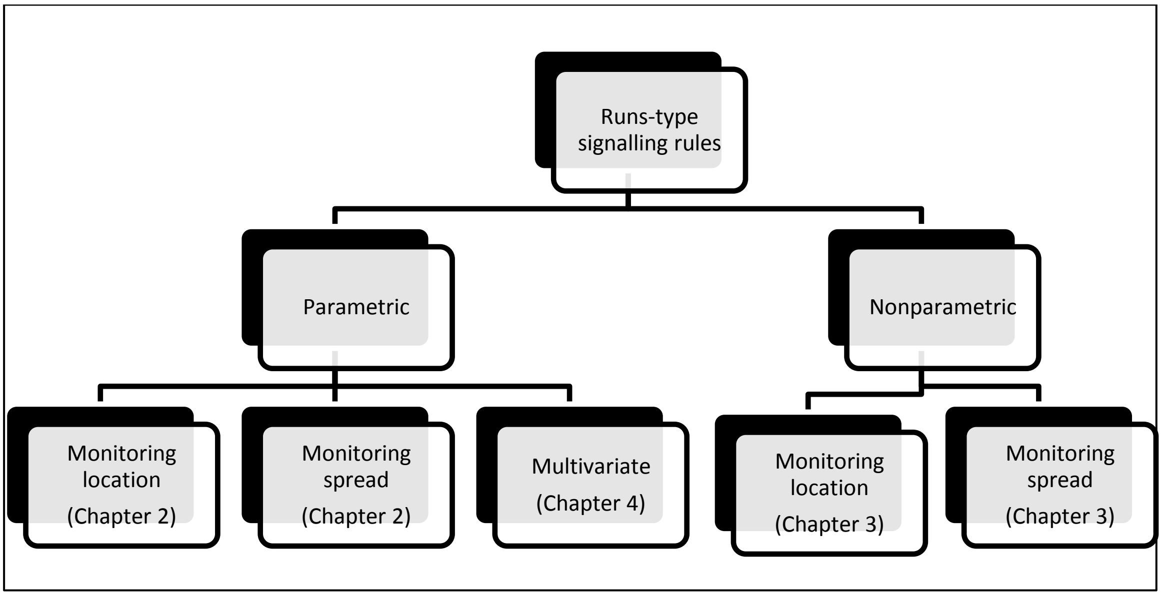

Figure 6.1: Classifications of the runs-type signalling rules control chart.

General variations of control charts discussed in the thesis are the upper one-sided, lower one-sided and two-sided control charts as depicted in Figures 1.5 and 1.6. Depending on whether a process increase, decrease or both an increase and decrease are monitored the appropriate chart has to be selected. For a process increase the upper one-sided chart is sufficient while if both an increase and decrease has to be monitored the two-sided control chart will do. The presence or absence of inner and outer control limits also influence the signal criteria; a signal is produced if a sequence of charting statistics plots within the inner and outer control limits or on or outside the control limit in the case of a single control limit. Different variations of runs-type signalling rules discussed in the thesis are the *k-of-k* and *k-of-w* or a combination of the *k-of-k* and the *k-of-w* as shown in Section 1.6.

Some of the earliest work on runs-type signalling rules were done by Moore (1958), Champ and Woodall (1987), Lowry et al. (1995) and Klein (2000a). The advantage of combining different runs-type signalling rules, e.g. "Two Improved Runs Rules for the Shewhart $\bar{X}$ Control Chart" (Khoo et al. 2006), is that both large and small shifts are detected. It should however be kept in mind that combining different runs-type signalling rules lead to an increased *FAR* as noted by Montgomery (2013). To counter the increased *FAR* the control limits can be adjusted (Klein (2000a**)**, Acosta-Mejia (2007)). An additional improvement to the performance was provided by Antzoulakos and Rakitzis (2008a) who proposed the revised *k-of-w* control chart where a signal can only be detected for the *k-of-w* runs-type signalling rule if all the charting statistics considered either plot above or below the CL when considering an upward or downward shift; note that the improved *k-of-w* runs-type signalling rule provides a signal regardless of the spread of the charting statistics above or below the CL. No signal would be provided for the revised *k-of-w* control chart if one charting statistic plots in the upper half of the control chart (above CL) followed by the next one plotting in the lower half (below CL); thus in the sequence of charting statistics being evaluated all should either plot above or below the CL.

Low et al. (2012) proposed a chart based on the MRL since the shape of the run-length distribution changes with the magnitude of the shift in the process mean ranging from highly skewed when the process is in-control (IC) to almost symmetric when a large shift occurs. Increased performance was noted for the chart proposed by Low et al. (2012).

Compared to parametric charts monitoring location supplemented with runs-type signalling rules, not that much work has been done on parametric charts monitoring spread supplemented with runs-type signalling rules. Lowry et al. (1995) proposed S and R charts supplemented with runs-type signalling rules similar to some of the sensitizing rules of the Western Electric company that have similar ARL_0 performance compared to the standard Shewhart $\bar{X}$ control chart. Klein (2000b) proposed various control charts monitoring spread: the standard S control chart, the S control chart with unequal chi-square tail probabilities, the S control chart with equal chi-square tail probabilities and the *2-of-2* control chart. Similarly to the parametric control charts monitoring location, an increase in FAR was noted for parametric control charts monitoring spread when multiple runs-type signalling rules were used simultaneously. To address the increased FAR Antzoulakos and Rakitzis (2010) proposed various signalling rules monitoring spread in order to monitor both increases and decreases in process standard deviation. Acosta-Mejia et al. (2008, 2009) proposed R-and S control charts supplemented with the *k-of-k* runs-type signalling rules in order to monitor both increases and decreases effectively. Antzoulakos and Rakitzis (2014) proposed a control chart supplemented with the *k-of-k* runs-type signalling rule and applied VSI as well.

Nonparametric control charts supplemented with runs-type signalling rules is an area for possible new research since not a lot of work has been done in this field. Bakir (2004) proposed a nonparametric control chart based on signed ranks incorporating the *1-of-1* runs-type signalling rule. Chakraborti and Eryilmaz (2007) followed by proposing a nonparametric control chart based on signed ranks supplemented with the *k-of-k* and *k-of-w* runs-type signalling rules in order to decrease the FAR and increase ARL_0 performance. Human et al. (2010) proposed three charts implementing the sign test statistic; the *1-of-1*, the *k-of-k* and the *k-of-w*. Balakrishnan et al. (2010) proposed a control chart taking into account the location of the single order statistic (*1-of-1* runs-type signalling rule) from the test sample as well as the number of observations between the control limits. Chakraborti et al. (2009) proposed a control chart based on the precedence statistics supplemented with runs-type signalling rules. It should be noted that, for some nonparametric control charts, e.g. the nonparametric control chart based on the sign-rank statistic, the additional assumption of symmetry is needed; this is additional to the assumption that observations are from a continuous distribution.

Multivariate control charts supplemented with runs-type signalling rules is another area where extra research is needed. Multivariate control charts were developed as an alternative to having different charts monitoring various quality characteristics of a process. Aparisi et al. (2004) investigated a chi-square control chart supplemented with various runs-type signalling rules, e.g. the *1-of-1*, the *2-of-3* and a control chart monitoring the last eight points above the median. Koutras et al. (2006) developed control charts supplemented with runs-type signalling rules implementing a combination of the *k-of-k* and *r-of-r* rules simultaneously. Various other authors contributed to the field of multivariate control charts supplemented with runs-type signalling rules, e.g. Khoo and Quah (2003), Khoo et al. (2005), Champ and Aparisi (2008), Aparisi and de Luna (2009), Lee (2013). Khoo and Quah (2004) also studied multivariate control charts for dispersion supplemented with runs-type signalling rules focussing on the *1-of-1*, *2-of-2*, *2-of-3*, *2-of-4*, *3-of-3* and *3-of-4* control charts.

The two main performance measures proposed are the ARL and the MRL. In order to calculate the ARL the Markov chain approach is applied with the different transient and absorbent states being identified, the essential TPM set up and Equation (1.4) applied. Various other techniques exist to evaluate the run-length distribution; these are the exact approach, the integral equation approach and computer simulations. The Markov chain approach is the approach being followed in the thesis. The reader is referred to Section 1.6 for a detailed discussion on the run-length distribution as well as the Markov chain approach. To calculate the median run-length Equation (1.6) has to be solved for t. Since the distribution of the run-length is skewed the MRL is applied as a performance measure.

The distributional assumptions for parametric control charts monitoring location and spread are that observations are distributed iid $N(0,1)$. This aids in the design process since if the probability of plotting above or below certain control limits is known, the corresponding control limit(s) can be calculated. For nonparametric control charts the distribution is unknown but continuous. The case where process parameters are unknown and a phase I and phase II analysis needs to be completed is referred to in Section 2.1.9 and Section 3.1.5. For multivariate control charts the observations are also assumed to distributed iid $N(0,1)$.

In the design process the aim is to establish the control limits for a predetermined criteria, e.g. some ARL_0 or FAR. Typically the ARL_0 would be in the region of 370 or 500. For control charts having the same ARL_0 the chart chosen would be the one with the smallest ARL_ξ as depicted in Figure 5.1. In order to design a control chart supplemented with runs-type signalling rules the following information is needed:

i. The runs-type signalling rule used, e.g. *k-of-w* or the *1-of-1* combined with the *k-of-w* runs-type signalling rule; also establishing k and w.
ii. Sample size m or n (Depending on a phase I or II analysis).
iii. Control limits.

The steps in designing the control chart are then as follows:

i. Define the regions and associated probabilities of the charting statistic.
ii. Establish the transition probabilities in terms of the control limits or as a function of the unknown parameters.
iii. Set up the essential TPM.
iv. Solve Equation (1.4) for a given ARL_0.

In order to simplify the design process symmetric control limits are used for charts monitoring location. This reduces the unknown parameters when both inner and outer control limits are considered, see e.g. Section 5.1.3. For control charts monitoring spread when both increases and decreases are monitored the process is simplified by considering equal tail probabilities, see e.g. Section 5.1.2. In order to solve the control limits Equation (1.4) has to be applied for a given ARL_0. The reader is referred to Chapter 5 for a detailed discussion on the design process for Shewhart-type control charts supplemented with runs-type signalling rules.

As discussed in Chapter 5 the attainable ARL, denoted by ARL^*, doesn't always equal the ARL_0 in the design process. If the charting statistic is continous, e.g. iid $N(0,1)$, then the $ARL^* = ARL_0$. If on the other hand the charting statistic is discrete, e.g. $Bin(n,p)$, then the $ARL^* \neq ARL_0$ as is the case for nonparametric control charts. Where the $ARL^* \neq ARL_0$ the control charts are designed such that the $ARL^* \geq ARL_0$; this being a conservative approach.

Section 5.2 focusses on the implementation of univariate parametric and nonparametric control charts monitoring location and spread supplemented with runs-type signalling rules. The properties of the run-length distribution are evaluated by considering the Markov chain, simulation and exact approach. Similar results were noted for the various type of approaches mentioned before, see Section 5.2.

Scope for further research include nonparametric charts, especially nonparametric charts monitoring spread as well as parametric charts implementing VSS, VSI and $VSSI$ monitoring location and spread. EWMA charts supplemented with runs-type signalling rules in order to counter the inertia problem is another research possibility; note that only Shewhart-type charts are considered for this study.

Various control charts supplemented with runs-type signalling rules were discussed in the thesis, parametric and nonparametric control charts monitoring location and spread as well as multivariate control charts. To conclude, control charts supplemented with runs-type signalling rules effectively detect small shifts or a combination of small and large shifts in the process characteristic being monitored and are a viable alternative to EWMA and CUSUM charts with the added benefit of design simplicity.

Appendix

Introduction

The appendix is a reference to the results in the thesis. For clarification A1 refers to the results specified in Chapter 1.

A1 (Chapter 1)

Proof 1.1

Let:

$$\boldsymbol{W}_j = \boldsymbol{I} + \boldsymbol{Q} + \boldsymbol{Q}^2 + \boldsymbol{Q}^3 + \cdots + \boldsymbol{Q}^{j-1} \qquad \text{(A1.1)}$$

$$\boldsymbol{Q}\boldsymbol{W}_j = \boldsymbol{Q} + \boldsymbol{Q}^2 + \boldsymbol{Q}^3 + \cdots + \boldsymbol{Q}^{j} \qquad \text{(A1.2)}$$

((A1.1)-(A1.2)): $W_j - \boldsymbol{Q}\boldsymbol{W}_j = \left(\boldsymbol{I} + \boldsymbol{Q} + \boldsymbol{Q}^2 + \cdots + \boldsymbol{Q}^{j-1}\right) - \left(\boldsymbol{Q} + \boldsymbol{Q}^2 + \cdots + \boldsymbol{Q}^{j}\right)$

$$\boldsymbol{W}_j(\boldsymbol{I} - \boldsymbol{Q}) = \boldsymbol{I} - \boldsymbol{Q}^j$$

$$\therefore \boldsymbol{W}_j = (\boldsymbol{I} - \boldsymbol{Q}^j)(\boldsymbol{I} - \boldsymbol{Q})^{-1} \text{ since } (\boldsymbol{I} - \boldsymbol{Q}) \text{ is invertible}$$

Since all the elements of Q lie in the interval [0,1) $\lim_{j\to\infty} \boldsymbol{Q}^j = 0$

$$\begin{aligned}
\therefore \lim_{j\to\infty} \boldsymbol{W}^j &= \lim_{j\to\infty} (\boldsymbol{I} - \boldsymbol{Q}^j)(\boldsymbol{I} - \boldsymbol{Q})^{-1} \\
&= \left(\boldsymbol{I} - \left(\lim_{j\to\infty} \boldsymbol{Q}^j\right)\right)(\boldsymbol{I} - \boldsymbol{Q})^{-1} \\
&= (\boldsymbol{I} - (\boldsymbol{0}))(\boldsymbol{I} - \boldsymbol{Q})^{-1} \\
&= (\boldsymbol{I})(\boldsymbol{I} - \boldsymbol{Q})^{-1} \\
&= (\boldsymbol{I} - \boldsymbol{Q})^{-1}
\end{aligned}$$

$$\therefore \boldsymbol{I} + \boldsymbol{Q} + \boldsymbol{Q}^2 + \boldsymbol{Q}^3 + \cdots = (\boldsymbol{I} - \boldsymbol{Q})^{-1}$$

$$\lim_{j\to\infty} \boldsymbol{W}^j = (\boldsymbol{I} - \boldsymbol{Q})^{-1} \qquad \text{(A1.3)}$$

Proof 1.2

Since $P\left(Y_j < w_{k+1}\middle|Y_0 = \phi\right) = \sum_{i=0}^{h} P(Y_j = w_i\,|Y_0 = \phi)$ and $P(Y_0 = \phi) = 1$ for a given $j = 1,2,3,\ldots$, the result follows from the matrix version of the Chapman-Kolmogorov equation.

$$P\left(Y_j < w_{k+1}\middle|Y_0 = \phi\right) = \sum_{i=0}^{k} P(Y_j = w_i\,|Y_0 = \phi)$$

$$= (1,0,0,\ldots,0)\boldsymbol{M}^j(1,1,\ldots 1_k, 0_{k+1}, 0, \ldots, 0_m)$$

$$= \boldsymbol{\omega}\boldsymbol{Q}^j\boldsymbol{1} \text{ where } \boldsymbol{1}_{\boldsymbol{kx1}} = (1,1,1,\ldots 1)^{\boldsymbol{T}}$$

$$\text{and } \boldsymbol{\omega}_{\boldsymbol{1}xk} = (1,0,0,\ldots,0)$$

$$\therefore P\left(Y_j < w_{k+1}\middle|Y_0 = \phi\right) = \boldsymbol{\omega}\boldsymbol{Q}^j\boldsymbol{1} \tag{A1.4}$$

Proof 1.3

For a given $j = 1,2,3,\ldots$, since $k\ (k \geq 1)$ and $m - k\ (m - k \geq 1)$ are the number of transient and absorbing states respectively, it follows from the definitions of $\{Y_j\}$ and $\{N\}$, that $\left(Y_j < w_{k+1}\right) \leftrightarrow (N > j)$.

Hence, from A1.4

$$P(N = j|Y_0 = \phi) = P(N > j - 1|Y_0 = \phi) - P(N > j|Y_0 = \phi)$$

$$= P\left(Y_{j-1} < w_{k+1}\middle|Y_0 = \phi\right) - P\left(Y_j < w_{k+1}\middle|Y_0 = \phi\right)$$

$$= (\boldsymbol{\omega}\boldsymbol{Q}^{\boldsymbol{j}-1}\boldsymbol{1}) - (\boldsymbol{\omega}\boldsymbol{Q}^j\boldsymbol{1})$$

$$= \boldsymbol{\omega}\boldsymbol{Q}^{\boldsymbol{j}-1}\boldsymbol{1}(\boldsymbol{I} - \boldsymbol{Q})\boldsymbol{1}$$

$$P(N = j|Y_0 = \phi) = \boldsymbol{\omega}\boldsymbol{Q}^{\boldsymbol{j}-1}(\boldsymbol{I} - \boldsymbol{Q})\boldsymbol{1} \tag{A1.5}$$

Proof 1.4

From the definition of the moment generating function (Definition 2.5.1 Bain & Engelhardt (1992)) the moment generating function of the random variable T can be derived as follows:

$$M_T(n) = E(e^{nT})$$

$$= \sum_{t=1}^{\infty} e^{nt} P(T = t | Y_0 = \phi)$$

$$= \sum_{t=1}^{\infty} e^{nt} (P(T > t-1 | Y_0 = \phi) - P(T > t | Y_0 = \phi))$$

$$= \sum_{t=1}^{\infty} e^{nt} P(T > t-1 | Y_0 = \phi) - e^{nt} P(T > t | Y_0 = \phi)$$

$$= \left(\sum_{t=1}^{\infty} e^{nt} P(Y_{t-1} < w_{h+1} | Y_0 = \phi\right) - \left(\sum_{t=1}^{\infty} e^{nt} P(Y_t < w_{h+1} | Y_0 = \phi\right) \text{ (from A1.5)}$$

$$= \left(\sum_{t=1}^{\infty} e^{nt} \boldsymbol{\omega} Q^{t-1} \mathbf{1}\right) - \left(\sum_{t=1}^{\infty} e^{nt} \boldsymbol{\omega} \boldsymbol{Q}^{t} \mathbf{1}\right) \text{ (from A2.4)}$$

$$= \boldsymbol{\omega} \left(\sum_{t=1}^{\infty} e^{nt} \boldsymbol{Q}^{t-1} - \sum_{t=1}^{\infty} e^{nt} \boldsymbol{Q}^{t}\right) \mathbf{1}$$

$$= \boldsymbol{\omega}(e^n \boldsymbol{I} + e^{2n} \boldsymbol{Q} + e^{3n} \boldsymbol{Q}^2 + \cdots - e^n \boldsymbol{Q} - e^{2n} \boldsymbol{Q}^2 - e^{3n} \boldsymbol{Q}^3 -) \mathbf{1}$$

$$= \boldsymbol{\omega}(e^n \boldsymbol{Q}(e^n - 1) + e^{2n} \boldsymbol{Q}^2 (e^n - 1) + \cdots) 1 + \boldsymbol{\omega} e^n \boldsymbol{I} \mathbf{1}$$

$$= (e^n - 1) \boldsymbol{\omega} (e^n \boldsymbol{Q} + e^{2n} \boldsymbol{Q}^2 + \cdots) \mathbf{1} + \boldsymbol{\omega} e^n \boldsymbol{I} \mathbf{1}$$

$$= (e^n - 1) \boldsymbol{\omega} (\boldsymbol{I} - \boldsymbol{I} + e^n \boldsymbol{Q} + e^{2n} \boldsymbol{Q}^2 + \cdots) \mathbf{1} + \boldsymbol{\omega} e^n \boldsymbol{I} \mathbf{1}$$

$$= (e^n - 1) \boldsymbol{\omega} (\boldsymbol{I} + e^n \boldsymbol{Q} + e^{2n} \boldsymbol{Q}^2 + \cdots) \mathbf{1} + \boldsymbol{\omega} e^n \boldsymbol{I} \mathbf{1} - (e^n - 1) \boldsymbol{\omega} \boldsymbol{I} \mathbf{1}$$

$$= (e^n - 1) \boldsymbol{\omega} (\textstyle\sum_{t=0}^{\infty} e^{nt} \boldsymbol{Q}^t) \mathbf{1} + e^n \boldsymbol{\omega} \boldsymbol{I} \mathbf{1} - e^n \boldsymbol{\omega} \boldsymbol{I} \mathbf{1} + \boldsymbol{\omega} \boldsymbol{I} \mathbf{1}$$

$$= (e^n - 1) \boldsymbol{\omega} \left(\sum_{t=0}^{\infty} e^{nt} \boldsymbol{Q}^t\right) \mathbf{1} + \boldsymbol{\omega} \boldsymbol{I} \mathbf{1}$$

$$= (e^n - 1) \boldsymbol{\omega} \left(\sum_{t=0}^{\infty} e^{nt} \boldsymbol{Q}^t\right) \mathbf{1} + 1$$

From A1.2 and the equation above it follows that:

$$M_{\boldsymbol{T}}(\mathrm{n}) = (e^n - 1) \boldsymbol{\omega} (\boldsymbol{I} - e^n \boldsymbol{Q})^{-\mathbf{1}} \mathbf{1} + 1 \tag{A1.6}$$

Fu et al. (2003)

Proof 1.5

By differentiating the moment generating function with respect to n and setting $n = 0$, the moments of the run-length variable T can be found i.e.:

$$\frac{d}{dn} M_T(n)||_{n=0} = M'_T(0) = E(T)$$

$$\frac{d^2}{dn^2} M_T(n)||_{n=0} = M''_T(0) = E(T^2)$$

$$M_T(n) = (e^n - 1)\boldsymbol{\omega}(\boldsymbol{I} - e^n\boldsymbol{Q})^{-\mathbf{1}}\mathbf{1} + 1$$

$$\begin{aligned}
M'_T(n) &= \frac{d}{dn} M_T(n) \\
&= (1)(e^n - 1)\boldsymbol{\omega}(\boldsymbol{I} - e^n\boldsymbol{Q})^{-\mathbf{1}}\mathbf{1} + (e^n - 1)\boldsymbol{\xi}(-1)(\boldsymbol{I} - e^n\boldsymbol{Q})^{-\mathbf{2}}(-e^n\mathbf{Q})\mathbf{1} \\
&= e^n\boldsymbol{\omega}(\boldsymbol{I} - e^n\boldsymbol{Q})^{-\mathbf{1}}\mathbf{1} + (e^n - 1)\boldsymbol{\omega}(\boldsymbol{I} - e^n\boldsymbol{Q})^{-\mathbf{2}}(e^n\mathbf{Q})\mathbf{1} \\
&= e^n\boldsymbol{\omega}(\boldsymbol{I} - e^n\boldsymbol{Q})^{-\mathbf{1}}\mathbf{1} + (e^{2n} - e^n)\boldsymbol{\omega}(\boldsymbol{I} - e^n\boldsymbol{Q})^{-\mathbf{2}}\mathbf{1}
\end{aligned}$$

$$\begin{aligned}
M''_T(n) &= \frac{d^2}{dn^2} M_T(n) \\
&= \frac{d}{dn} M'_T(n) \\
&= e^n\boldsymbol{\omega}(\boldsymbol{I} - e^n\boldsymbol{Q})^{-\mathbf{1}}\mathbf{1} + (e^n)\boldsymbol{\omega}(-\mathbf{1})(\boldsymbol{I} - e^n\boldsymbol{Q})^{-\mathbf{2}}(-e^n\mathbf{Q})\mathbf{1} \\
&\qquad + (1)(e^{2n} - e^n)^0(e^{2n}2 - e^n)\boldsymbol{\omega}(\boldsymbol{I} - e^n\boldsymbol{Q})^{-\mathbf{2}}\mathbf{Q}\mathbf{1} \\
&\qquad + (e^{2n} - e^n)^0\boldsymbol{\omega}(-\mathbf{2})(\boldsymbol{I} - e^n\boldsymbol{Q})^{-\mathbf{3}}\mathbf{Q}(-e^n\boldsymbol{Q})\boldsymbol{Q}\mathbf{1}
\end{aligned}$$

$$\begin{aligned}
M'_T(0) &= e^0\boldsymbol{\omega}(\boldsymbol{I} - e^0\boldsymbol{Q})^{-\mathbf{1}}\mathbf{1} + \left(e^{2(0)} - e^0\right)\boldsymbol{\omega}(\boldsymbol{I} - e^0\boldsymbol{Q})^{-\mathbf{2}}\boldsymbol{Q}\mathbf{1} \\
&= 1\boldsymbol{\omega}(\boldsymbol{I} - 1\boldsymbol{Q})^{-\mathbf{1}}\mathbf{1} + (1 - 1)\boldsymbol{\omega}(\boldsymbol{I} - 1\boldsymbol{Q})^{-\mathbf{2}}\boldsymbol{Q}\mathbf{1} \\
&= \boldsymbol{\omega}(\boldsymbol{I} - \boldsymbol{Q})^{-\mathbf{1}}\mathbf{1} + (0)\boldsymbol{\omega}(\boldsymbol{I} - \boldsymbol{Q})^{-\mathbf{2}}\boldsymbol{Q}\mathbf{1} \\
&= \boldsymbol{\omega}(\boldsymbol{I} - \boldsymbol{Q})^{-\mathbf{1}}\mathbf{1}
\end{aligned}$$

$$\begin{aligned}
M''_T(0) &= e^0\boldsymbol{\omega}(\boldsymbol{I} - e^0\boldsymbol{Q})^{-\mathbf{1}}\mathbf{1} + \left(e^{2(0)}\right)\boldsymbol{\omega}(\boldsymbol{I} - e^0\boldsymbol{Q})^{-\mathbf{2}}\boldsymbol{Q}\mathbf{1} \\
&\qquad + \left(\mathbf{2}e^{2(0)} - e^0\right)\boldsymbol{\omega}(\boldsymbol{I} - e^0\boldsymbol{Q})^{-\mathbf{2}}\boldsymbol{Q}\mathbf{1} + \mathbf{2}(e^{3(0)} - e^{2(0)})\boldsymbol{\omega}(\boldsymbol{I} - e^0\boldsymbol{Q})^{-\mathbf{3}}\boldsymbol{Q}^{\mathbf{2}}\mathbf{1} \\
&= \mathbf{1}\,\boldsymbol{\omega}(\boldsymbol{I} - 1\boldsymbol{Q})^{-\mathbf{1}}\mathbf{1} + (1)\boldsymbol{\omega}(\boldsymbol{I} - 1\boldsymbol{Q})^{-\mathbf{2}}\boldsymbol{Q}\mathbf{1} + (\mathbf{2}(1) - 1)\boldsymbol{\omega}(\boldsymbol{I} - 1\boldsymbol{Q})^{-\mathbf{2}}\boldsymbol{Q}\mathbf{1} + \mathbf{2}(1 \\
&\qquad - 1)\boldsymbol{\omega}(\boldsymbol{I} - 1\boldsymbol{Q})^{-\mathbf{3}}\boldsymbol{Q}^{\mathbf{2}}\mathbf{1} \\
&= \boldsymbol{\omega}(\boldsymbol{I} - \boldsymbol{Q})^{-\mathbf{1}}\mathbf{1} + \boldsymbol{\omega}(\boldsymbol{I} - \boldsymbol{Q})^{-\mathbf{2}}\boldsymbol{Q}\mathbf{1} + \boldsymbol{\omega}(\boldsymbol{I} - \boldsymbol{Q})^{-\mathbf{2}}\boldsymbol{Q}\mathbf{1} \\
&= \boldsymbol{\omega}(\boldsymbol{I} + (\boldsymbol{I} - \boldsymbol{Q})^{-\mathbf{1}}\boldsymbol{Q} + (\boldsymbol{I} - \boldsymbol{Q})^{-\mathbf{1}}\boldsymbol{Q})(\boldsymbol{I} - \boldsymbol{Q})^{-\mathbf{1}}\mathbf{1}
\end{aligned}$$

$$= \boldsymbol{\omega}((\boldsymbol{I}-\boldsymbol{Q})\boldsymbol{I}+\boldsymbol{Q}+\boldsymbol{Q})(\boldsymbol{I}-\boldsymbol{Q})^{-\mathbf{1}}(\boldsymbol{I}-\boldsymbol{Q})^{-\mathbf{1}}\mathbf{1}$$
$$= \boldsymbol{\omega}(\boldsymbol{I}-\boldsymbol{Q}+\boldsymbol{Q}+\boldsymbol{Q})(\boldsymbol{I}-\boldsymbol{Q})^{-\mathbf{2}}\mathbf{1}$$
$$= \boldsymbol{\omega}(\boldsymbol{I}+\boldsymbol{Q})\,(\boldsymbol{I}-\boldsymbol{Q})^{-\mathbf{2}}\mathbf{1}$$

$$E(T) = \mathrm{M}_T'\,(0) = \boldsymbol{\omega}(\boldsymbol{I}-\boldsymbol{Q})^{-\mathbf{1}}\mathbf{1} \tag{A1.7}$$

$$E(T^2) = \mathrm{M}_T''\,(0) = \boldsymbol{\omega}(\boldsymbol{I}+\boldsymbol{Q})\,(\boldsymbol{I}-\boldsymbol{Q})^{-\mathbf{2}}\mathbf{1} \tag{A1.8}$$

$$Var(T) = \boldsymbol{E}(T^2) - (E(T))^2$$
$$= \boldsymbol{\omega}(\boldsymbol{I}+\boldsymbol{Q})(\boldsymbol{I}-\boldsymbol{Q})^{-\mathbf{2}}\mathbf{1} - (\boldsymbol{\xi}(\boldsymbol{I}-\boldsymbol{Q})^{-\mathbf{1}}\mathbf{1})^{\mathbf{2}}$$

$$\therefore\ Var(T) = \boldsymbol{\omega}(\boldsymbol{I}+\boldsymbol{Q})(\boldsymbol{I}-\boldsymbol{Q})^{-\mathbf{2}}\mathbf{1} - (\boldsymbol{\xi}(\boldsymbol{I}-\boldsymbol{Q})^{-\mathbf{1}}\mathbf{1})^{\mathbf{2}} \tag{A1.9}$$

$$Stdv(T) = \sqrt{Var(T)} = \sqrt{\boldsymbol{\omega}(\boldsymbol{I}+\boldsymbol{Q})(\boldsymbol{I}-\boldsymbol{Q})^{-\mathbf{2}}\mathbf{1} - (\boldsymbol{\omega}(\boldsymbol{I}-\boldsymbol{Q})^{-\mathbf{1}}\mathbf{1})^{\mathbf{2}}} \tag{A1.10}$$

Proof 1.6

It follows from the definition of the probability generating function (Definition 2.5.2 Bain & Engelhardt (1992)) that the random variable T has probability generating function:

$$\begin{aligned}
G_T(n) &= E(n^T)\\
&= \sum_{t=1}^{\infty} n^t\, P(T=t|Y_0=\phi)\\
&= \sum_{t=1}^{\infty} n^t\, (P(T>t-1|Y_0=\phi) - P(T>t|Y_0=\phi))\\
&= \left(\sum_{t=1}^{\infty} n^t\, P(T>t-1|Y_0=\phi)\right) - \left(\sum_{t=1}^{\infty} n^t\, P(T>t|Y_0=\phi)\right)\\
&= \left(\sum_{t=1}^{\infty} n^t\, P(Y_{n-1}<w_{k+1}|Y_0=\phi)\right)\\
&\qquad - \left(\sum_{t=1}^{\infty} n^t\, P(Y_n<w_{k+1}|Y_0=\phi)\right) \text{ (from A1.4)}\\
&= \left(\sum_{t=1}^{\infty} n^t \boldsymbol{\omega}\boldsymbol{Q}^{\boldsymbol{n}-\mathbf{1}}\mathbf{1}\right) - \left(\sum_{t=1}^{\infty} n^t \boldsymbol{\omega}\boldsymbol{Q}^{\boldsymbol{n}}\mathbf{1}\right) \text{ (from A1.3)}
\end{aligned}$$

$$
\begin{aligned}
&= \boldsymbol{\omega}\left(\sum_{t=1}^{\infty} n^t \boldsymbol{Q^{t-1}} - \sum_{t=1}^{\infty} n^t \boldsymbol{Q^t}\right)\mathbf{1} \\
&= \boldsymbol{\omega}\Big((n\boldsymbol{I} + n^2\boldsymbol{Q} + n^3\boldsymbol{Q^2} + \cdots) - (n\boldsymbol{Q} + n^2\boldsymbol{Q^2} + n^3\boldsymbol{Q^3} + \cdots)\Big)\mathbf{1} \\
&= \boldsymbol{\omega}(nQ(n-1) + n^2Q^2(n-1) + \cdots)1 + \boldsymbol{\omega n I 1} \\
&= (n-1)\boldsymbol{\omega}(n\boldsymbol{Q} + n^2\boldsymbol{Q^2} + n^3\boldsymbol{Q^3} + \cdots)\mathbf{1} + \boldsymbol{\omega n I 1} \\
&= (n-1)\boldsymbol{\omega}(\boldsymbol{I} - \boldsymbol{I} + \boldsymbol{nQ} + n^2\boldsymbol{Q^2} + \cdots)\mathbf{1} + \boldsymbol{\omega n I 1} \\
&= (n-1)\boldsymbol{\omega}(\boldsymbol{I} + \boldsymbol{nQ} + n^2\boldsymbol{Q^2} + \cdots)\mathbf{1} + \boldsymbol{\omega n I 1} - (n-1)\boldsymbol{\omega I 1} \\
&= (n-1)\boldsymbol{\omega}\left(\sum_{\boldsymbol{t=0}}^{\infty} \boldsymbol{n^t\, Q^n}\right)\mathbf{1} + \boldsymbol{n\omega I 1} - \boldsymbol{n\omega I 1} + \boldsymbol{\omega I 1} \\
&= (n-1)\boldsymbol{\omega}\left(\sum_{\boldsymbol{t=0}}^{\infty} \boldsymbol{n^t\, Q^n}\right)\mathbf{1} + \boldsymbol{\omega I 1} \\
&= (n-1)\boldsymbol{\omega}\left(\sum_{\boldsymbol{t=0}}^{\infty} \boldsymbol{n^t\, Q^n}\right)\mathbf{1} + 1
\end{aligned}
$$

From **A1.2** and the equation above it follows that:

$$G_T(n) = (\mathrm{n}-1)\boldsymbol{\xi}(\boldsymbol{I} - n\boldsymbol{Q})^{-\mathbf{1}}\mathbf{1} + 1 \qquad \text{(A1.11)}$$

Proof 1.7

$$
\begin{aligned}
G_T'(n) &= \frac{d}{dt}(\mathrm{n}-1)\boldsymbol{\omega}(\boldsymbol{I} - n\boldsymbol{Q})^{-\mathbf{1}}\mathbf{1} + 1 \\
&= \boldsymbol{\omega}(\boldsymbol{I} - n\boldsymbol{Q})^{-\mathbf{1}}\mathbf{1} + (\mathrm{n}-1)\boldsymbol{\omega}(-\mathbf{1})(-\boldsymbol{Q})(\boldsymbol{I} - n\boldsymbol{Q})^{-\mathbf{2}}\mathbf{1} \\
&= \boldsymbol{\omega}(\boldsymbol{I} - n\boldsymbol{Q})^{-\mathbf{1}}\mathbf{1} + (\mathrm{n}-1)\boldsymbol{Q}(\boldsymbol{I} - n\boldsymbol{Q})^{-\mathbf{2}}\mathbf{1}
\end{aligned}
$$

$$
\begin{aligned}
G_T''(n) &= \frac{d}{dt}G_T'(n) \\
&= \frac{d}{dt}(\boldsymbol{\omega}(\boldsymbol{I} - n\boldsymbol{Q})^{-\mathbf{1}}\mathbf{1} + (\mathrm{n}-1)\boldsymbol{\omega}(-\mathbf{1})(-\boldsymbol{Q})(\boldsymbol{I} - n\boldsymbol{Q})^{-\mathbf{2}}\mathbf{1}) \\
&= (\boldsymbol{\omega}(-\mathbf{1})(-\boldsymbol{Q})(\boldsymbol{I} - n\boldsymbol{Q})^{-\mathbf{2}}\mathbf{1} + \boldsymbol{\omega}(-\mathbf{1})(-\boldsymbol{Q})(\boldsymbol{I} - n\boldsymbol{Q})^{-\mathbf{2}}\mathbf{1}) + \\
&\quad (n-1)\boldsymbol{\omega}(-\mathbf{1})(-\mathbf{2})(-\boldsymbol{Q})(-\boldsymbol{Q})(\boldsymbol{I} - \boldsymbol{nQ})^{-\mathbf{3}}\mathbf{1} \\
&= (\boldsymbol{\omega Q})(\boldsymbol{I} - n\boldsymbol{Q})^{-\mathbf{2}}\mathbf{1} + \boldsymbol{\omega Q}(\boldsymbol{I} - n\boldsymbol{Q})^{-\mathbf{2}}\mathbf{1}) + 2(n-1)\boldsymbol{\omega}(\boldsymbol{Q^2})(\boldsymbol{I} - \boldsymbol{nQ})^{-\mathbf{3}}\mathbf{1}
\end{aligned}
$$

$$
\begin{aligned}
G_T'(1) &= \boldsymbol{\omega}(\boldsymbol{I} - (\mathbf{1})\boldsymbol{Q})^{-\mathbf{1}}\mathbf{1} + (1-1)\boldsymbol{Q}(\boldsymbol{I} - \mathbf{1}\boldsymbol{Q})^{-\mathbf{2}}\mathbf{1} \\
&= \boldsymbol{\omega}(\boldsymbol{I} - (\mathbf{1})\boldsymbol{Q})^{-\mathbf{1}}\mathbf{1} + (0)\boldsymbol{Q}(\boldsymbol{I} - \mathbf{1}\boldsymbol{Q})^{-\mathbf{2}}\mathbf{1} \\
&= \boldsymbol{\omega}(\boldsymbol{I} - (\mathbf{1})\boldsymbol{Q})^{-\mathbf{1}}\mathbf{1}
\end{aligned}
$$

$$
\begin{aligned}
G_T''(1) &= (\boldsymbol{\omega Q})(\boldsymbol{I} - (1)\boldsymbol{Q})^{-\mathbf{2}}\mathbf{1} + \boldsymbol{\omega Q}(\boldsymbol{I} - (1)\boldsymbol{Q})^{-\mathbf{2}}\mathbf{1}) + 2((1)-1)\boldsymbol{\omega}(\boldsymbol{Q^2})(\boldsymbol{I} - (\mathbf{1})\boldsymbol{Q})^{-\mathbf{3}}\mathbf{1} \\
&= (\boldsymbol{\omega Q})(\boldsymbol{I} - \boldsymbol{Q})^{-\mathbf{2}}\mathbf{1} + \boldsymbol{\omega Q}(\boldsymbol{I} - \boldsymbol{Q})^{-\mathbf{2}}\mathbf{1}) + 2(0)\boldsymbol{\omega}(\boldsymbol{Q^2})(\boldsymbol{I} - \boldsymbol{Q})^{-\mathbf{3}}\mathbf{1} \\
&= \mathbf{2}\boldsymbol{\omega Q}(\boldsymbol{I} - \boldsymbol{Q})^{-\mathbf{2}}\mathbf{1}
\end{aligned}
$$

From Theorem 2.5.4 of Bain & Engelhardt (1992) it follows that:

$$
\begin{aligned}
&G_T'(1) = E(T) \\
&G_T''(1) = E\big(T(T-1)\big) = E(T^2) - E(T)
\end{aligned}
$$

It then follows that:

$$\begin{aligned} E(T^2) &= G_T''(1) + E(T) \\ &= G_T''(1) + G_T'(1) \end{aligned}$$

From the equations above the variance formulae can then be derived as follows:

$$\begin{aligned} Var(T) &= E(T^2) - (E(T))^2 \\ &= E(T^2) - E(T) + E(T) - (E(T))^2 \\ &= (E(T^2) - E(T)) + E(T) - (E(T))^2 \\ &= G_T''(1) + G_T'(1) - (G_T'(1))^2 \end{aligned}$$

$$\therefore E(T) = G_T'(1) = \boldsymbol{\omega}(\boldsymbol{I} - \boldsymbol{Q})^{-\mathbf{1}}\mathbf{1} \qquad \text{(A1.12)}$$

(Fu and Lou (2003), p. 73)

$$\begin{aligned} E(T^2) &= G_T''(1) + G_T'(1) \\ &= 2\boldsymbol{\omega}\boldsymbol{Q}(\boldsymbol{I} - \boldsymbol{Q})^{-\mathbf{2}}\mathbf{1} + \boldsymbol{\xi}(\boldsymbol{I} - \boldsymbol{Q})^{-\mathbf{1}}\mathbf{1} \\ &= \boldsymbol{\omega}(2\boldsymbol{Q}(\boldsymbol{I} - \boldsymbol{Q})^{-\mathbf{1}} + \boldsymbol{I})(\boldsymbol{I} - \boldsymbol{Q})^{-\mathbf{1}}\mathbf{1} \\ &= \boldsymbol{\omega}(\boldsymbol{I} + \mathbf{2}\boldsymbol{Q}(\boldsymbol{I} - \boldsymbol{Q})^{-\mathbf{1}}(\boldsymbol{I} - \boldsymbol{Q})^{-\mathbf{1}}\mathbf{1} \\ &= \boldsymbol{\omega}(\boldsymbol{I} + \mathbf{2}\boldsymbol{Q}(\boldsymbol{I} - \boldsymbol{Q})^{-\mathbf{1}}(\boldsymbol{I} - \boldsymbol{Q})^{-\mathbf{1}}\mathbf{1} \\ &= \boldsymbol{\omega}(\boldsymbol{I} + \mathbf{2}\boldsymbol{Q}(\boldsymbol{I} + \boldsymbol{Q} + \boldsymbol{Q}^{\mathbf{2}} + \boldsymbol{Q}^{\mathbf{3}} + \cdots))(\boldsymbol{I} - \boldsymbol{Q})^{-\mathbf{1}}\mathbf{1} \\ &= \boldsymbol{\omega}(\boldsymbol{I} + \mathbf{2}\boldsymbol{Q} + \mathbf{2}\boldsymbol{Q}^{\mathbf{2}} + \mathbf{2}\boldsymbol{Q}^{\mathbf{3}} + \cdots))(\boldsymbol{I} - \boldsymbol{Q})^{-\mathbf{1}}\mathbf{1} \\ &= \boldsymbol{\omega}(\boldsymbol{I} + \boldsymbol{Q} + \boldsymbol{Q}^{\mathbf{2}} + \boldsymbol{Q}^{\mathbf{3}} + \cdots + \boldsymbol{Q} + \boldsymbol{Q}^{\mathbf{2}} + \boldsymbol{Q}^{\mathbf{3}} + \cdots))(\boldsymbol{I} - \boldsymbol{Q})^{-\mathbf{1}}\mathbf{1} \\ &= \boldsymbol{\omega}(\boldsymbol{I} + \boldsymbol{Q} + \boldsymbol{Q}^{\mathbf{2}} + \boldsymbol{Q}^{\mathbf{3}} + \cdots + \boldsymbol{Q} + \boldsymbol{Q}^{\mathbf{2}} + \boldsymbol{Q}^{\mathbf{3}} + \cdots))(\boldsymbol{I} - \boldsymbol{Q})^{-\mathbf{1}}\mathbf{1} \\ &= \boldsymbol{\omega}((\boldsymbol{I} + \boldsymbol{Q} + \boldsymbol{Q}^{\mathbf{2}} + \boldsymbol{Q}^{\mathbf{3}} + \cdots) + \boldsymbol{Q}(\boldsymbol{I} + \boldsymbol{Q} + \boldsymbol{Q}^{\mathbf{2}} + \boldsymbol{Q}^{\mathbf{3}} + \cdots))(\boldsymbol{I} - \boldsymbol{Q})^{-\mathbf{1}}\mathbf{1} \\ &= \boldsymbol{\omega}((\boldsymbol{I} - \boldsymbol{Q})^{-\mathbf{1}} + \boldsymbol{Q}((\boldsymbol{I} - \boldsymbol{Q})^{-\mathbf{1}}))\,(\boldsymbol{I} - \boldsymbol{Q})^{-\mathbf{1}}\mathbf{1} \\ &= \boldsymbol{\omega}((\boldsymbol{I} + \boldsymbol{Q})(\boldsymbol{I} - \boldsymbol{Q})^{-\mathbf{1}})\,(\boldsymbol{I} - \boldsymbol{Q})^{-\mathbf{1}}\mathbf{1} \\ &= \boldsymbol{\omega}((\boldsymbol{I} + \boldsymbol{Q})(\boldsymbol{I} - \boldsymbol{Q})^{-\mathbf{2}}\mathbf{1} \end{aligned}$$

$$\therefore E(T^2) = \boldsymbol{\omega}(\boldsymbol{I} + \boldsymbol{Q})(\boldsymbol{I} - \boldsymbol{Q})^{-\mathbf{2}}\mathbf{1} \qquad \text{(A1.13)}$$

$$\begin{aligned} Var(T) &= G_T''(1) + G_T'(1) - (G_T'(1))^2 \\ &= (G_T''(1) + G_T'(1)) - (G_T'(1))^2 \\ &= E(T^2) - (E(T))^2 \\ &= \boldsymbol{\omega}(\boldsymbol{I} + \boldsymbol{Q})(\boldsymbol{I} - \boldsymbol{Q})^{-\mathbf{2}}\mathbf{1} - (\boldsymbol{\omega}(\boldsymbol{I} - \boldsymbol{Q})^{-\mathbf{1}}\mathbf{1})^{\mathbf{2}} \end{aligned}$$

$$\therefore Var(T) = \boldsymbol{\omega}(\boldsymbol{I} + \boldsymbol{Q})(\boldsymbol{I} - \boldsymbol{Q})^{-\mathbf{2}}\mathbf{1} - (\boldsymbol{\omega}(\boldsymbol{I} - \boldsymbol{Q})^{-\mathbf{1}}\mathbf{1})^{\mathbf{2}} \qquad \text{(A1.14)}$$

$$\therefore Stdv(T) = \sqrt{\boldsymbol{\omega}(\boldsymbol{I} + \boldsymbol{Q})(\boldsymbol{I} - \boldsymbol{Q})^{-\mathbf{2}}\mathbf{1} - (\boldsymbol{\omega}(\boldsymbol{I} - \boldsymbol{Q})^{-\mathbf{1}}\mathbf{1})^{\mathbf{2}}} \qquad \text{(A1.15)}$$

Proof 1.8

Consider (A1.4):

$$P\left(Y_j < w_{k+1} \middle| Y_0 = \phi\right) = P(N > j|Y_0 = \phi) \\ = \boldsymbol{\omega Q^j 1}, \forall j = 1,2,3 \ldots \quad \text{(A1.4)}$$

$$\therefore P(N > j|Y_0 = \phi) = \boldsymbol{\omega Q^j 1}, \forall j = 1,2,3 \ldots \quad \text{(A1.16)}$$

Using the above result:

$$P(N \leq j|Y_0 = \phi) = 1 - P(N > j|Y_0 = \phi) \\ = 1 - \ \boldsymbol{\omega Q^j 1}, \forall j = 1,2,3 \ldots$$

$$\therefore P(N \leq j|Y_0 = \phi) = 1 - \boldsymbol{\omega Q^j 1}, \forall j = 1,2,3 \ldots \quad \text{(A1.17)}$$

A5 (Chapter 5)

The following results are used to calculate the inverse of $3x3$ and $4x4$ matrices.

Consider the following matrix A ($3x3$):

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

A^{-1} exists if $\det(A) \neq 0$

$$A^{-1} = \frac{1}{\det(A)} \begin{pmatrix} a_{22}a_{33} - a_{23}a_{32} & a_{13}a_{32} - a_{12}a_{33} & a_{12}a_{23} - a_{13}a_{22} \\ a_{23}a_{31} - a_{21}a_{33} & a_{11}a_{33} - a_{13}a_{31} & a_{13}a_{21} - a_{11}a_{23} \\ a_{21}a_{32} - a_{22}a_{31} & a_{12}a_{31} - a_{11}a_{32} & a_{11}a_{22} - a_{11}a_{23} \end{pmatrix}$$

where

$$\det(A) = \mathrm{a}_{11}\mathrm{a}_{22}\mathrm{a}_{33} + \mathrm{a}_{21}\mathrm{a}_{32}\mathrm{a}_{13} + \mathrm{a}_{31}\mathrm{a}_{12}\mathrm{a}_{23} - \mathrm{a}_{11}\mathrm{a}_{32}\mathrm{a}_{23} - \mathrm{a}_{31}\mathrm{a}_{22}\mathrm{a}_{13} - \mathrm{a}_{21}\mathrm{a}_{12}\mathrm{a}_{33}$$

Consider the following matrix C ($4x4$):

$$C = \begin{pmatrix} c_{11} & c_{12} & c_{13} & c_{14} \\ c_{21} & c_{22} & c_{23} & c_{24} \\ c_{31} & c_{32} & c_{33} & c_{34} \\ c_{41} & c_{42} & c_{43} & c_{44} \end{pmatrix}$$

C^{-1} exists if $\det(C) \neq 0$

$$C^{-1} = \frac{1}{\det(C)} \begin{pmatrix} b_{11} & b_{12} & b_{13} & b_{14} \\ b_{21} & b_{22} & b_{23} & b_{24} \\ b_{31} & b_{32} & b_{33} & b_{34} \\ b_{41} & b_{42} & b_{343} & b_{44} \end{pmatrix}$$

$where$

$$b_{11} = c_{22}c_{33}c_{44} + c_{23}c_{34}c_{42} + c_{24}c_{32}c_{43} - c_{22}c_{34}c_{43} - c_{23}c_{32}c_{44} - c_{24}c_{33}c_{42}$$
$$b_{12} = c_{12}c_{34}c_{43} + c_{13}c_{32}c_{44} + c_{14}c_{33}c_{42} - c_{12}c_{33}c_{44} - c_{13}c_{34}c_{42} - c_{14}c_{32}c_{43}$$
$$b_{13} = c_{12}c_{23}c_{44} + c_{13}c_{32}c_{44} + c_{14}c_{22}c_{43} - c_{12}c_{24}c_{43} - c_{13}c_{22}c_{44} - c_{14}c_{23}c_{42}$$
$$b_{14} = c_{12}c_{24}c_{33} + c_{13}c_{22}c_{34} + c_{14}c_{23}c_{32} - c_{12}c_{23}c_{34} - c_{13}c_{24}c_{32} - c_{14}c_{22}c_{33}$$
$$b_{21} = c_{21}c_{34}c_{43} + c_{23}c_{31}c_{44} + c_{24}c_{33}c_{41} - c_{21}c_{33}c_{44} - c_{23}c_{34}c_{41} - c_{24}c_{31}c_{43}$$
$$b_{22} = c_{11}c_{33}c_{44} + c_{13}c_{34}c_{41} + c_{14}c_{31}c_{43} - c_{11}c_{34}c_{43} - c_{13}c_{31}c_{44} - c_{14}c_{33}c_{41}$$
$$b_{23} = c_{11}c_{24}c_{43} + c_{13}c_{21}c_{44} + c_{14}c_{23}c_{41} - c_{11}c_{23}c_{44} - c_{13}c_{24}c_{41} - c_{14}c_{21}c_{43}$$
$$b_{24} = c_{11}c_{23}c_{34} + c_{13}c_{24}c_{31} + c_{14}c_{21}c_{33} - c_{11}c_{24}c_{33} - c_{13}c_{21}c_{34} - c_{14}c_{23}c_{31}$$
$$b_{31} = c_{21}c_{32}c_{44} + c_{22}c_{34}c_{41} + c_{24}c_{31}c_{42} - c_{21}c_{34}c_{42} - c_{22}c_{31}c_{44} - c_{24}c_{32}c_{41}$$
$$b_{32} = c_{11}c_{32}c_{42} + c_{12}c_{31}c_{44} + c_{12}c_{34}c_{41} - c_{11}c_{32}c_{44} - c_{12}c_{34}c_{41} - c_{14}c_{31}c_{42}$$
$$b_{33} = c_{11}c_{22}c_{44} + c_{12}c_{24}c_{41} + c_{14}c_{21}c_{42} - c_{11}c_{24}c_{42} - c_{12}c_{21}c_{44} - c_{14}c_{22}c_{41}$$
$$b_{34} = c_{11}c_{22}c_{34} + c_{12}c_{21}c_{34} + c_{14}c_{22}c_{31} - c_{11}c_{22}c_{34} - c_{12}c_{24}c_{31} - c_{14}c_{21}c_{32}$$
$$b_{41} = c_{21}c_{33}c_{42} + c_{22}c_{31}c_{43} + c_{23}c_{32}c_{41} - c_{21}c_{32}c_{43} - c_{22}c_{33}c_{41} - c_{23}c_{31}c_{42}$$
$$b_{42} = c_{11}c_{23}c_{43} + c_{12}c_{33}c_{41} + c_{13}c_{31}c_{42} - c_{11}c_{33}c_{42} - c_{12}c_{31}c_{43} - c_{13}c_{32}c_{41}$$
$$b_{43} = c_{11}c_{23}c_{42} + c_{12}c_{21}c_{43} + c_{13}c_{22}c_{41} - c_{11}c_{22}c_{43} - c_{12}c_{23}c_{41} - c_{13}c_{21}c_{42}$$
$$b_{44} = c_{11}c_{22}c_{33} + c_{12}c_{23}c_{31} + c_{13}c_{21}c_{32} - c_{11}c_{23}c_{32} - c_{12}c_{21}c_{33} - c_{13}c_{22}c_{31}$$

and

$$\begin{aligned}\det(C) = {} & c_{11}c_{22}c_{33}c_{44} + c_{11}c_{23}c_{34}c_{42} + c_{11}c_{24}c_{32}c_{43} \\ & + c_{11}c_{21}c_{34} + c_{12}c_{23}c_{31}c_{44} + c_{12}c_{24}c_{33}c_{41} + c_{13}c_{21}c_{32}c_{44} + c_{13}c_{22}c_{34}c_{41} \\ & + c_{13}c_{24}c_{31}c_{42} + c_{14}c_{21}c_{33}c_{42} + c_{14}c_{22}c_{31}c_{43} + c_{14}c_{23}c_{32}c_{41} - c_{11}c_{22}c_{34}c_{43} \\ & - c_{11}c_{23}c_{32}c_{44} - c_{11}c_{24}c_{33}c_{42} - c_{12}c_{21}c_{33}c_{44} - c_{12}c_{23}c_{34}c_{41} - c_{12}c_{24}c_{31}c_{43} \\ & - c_{13}c_{21}c_{34}c_{42} - c_{13}c_{22}c_{31}c_{44} - c_{13}c_{24}c_{32}c_{41} - c_{14}c_{21}c_{32}c_{43} - c_{14}c_{22}c_{33}c_{41} \\ & - c_{14}c_{23}c_{31}c_{42}\end{aligned}$$

SAS programs

Shewhart $\bar{X}$ control chart - Markov chain approach

```
proc iml;
reset nolog;
/*Assumption is that observations distributed IID N(0,1)*/
/*Define UCL and LCL for Xbar chart*/
UCL=3;
LCL=-3;
/*p1 probability plotting on or above UCL*/
/*p2 probability plotting on or below LCL*/
/*p0 probability plotting between UCL and LCL*/

/*Process IC*/
/*calculate transitional probabilities*/
p1=1-cdf('normal',UCL,0,1);
p2=cdf('normal',LCL,0,1);
p0=1-p1-p2;
print p0;
/*set up essential TPM*/
/*ARL=w*(inv(I-Q))*1 Fu and Lou (2003)*/
w=1;
I=1; /*identity matrix Fu and Lou (2003)*/
Q=p0; /*only one transient state*/
one=1;
ARL0=w*(inv(I-Q))*one;

/*SDRL=sqrt(w(I+Q)((I-Q)^-2)*1 Fu and Lou (2003)*/
SDRL=sqrt((w*(I+Q)*((I-Q)**-2)*one)-ARL0##2);
print SDRL;

/*Cdf of run-length variable T : P(T=t)=1-Q^t*1 Fu and Lou (2003)*/
/*75th percentile P(T=t)=0.75 etc. */
/*calculate distribution of run length*/
cdf=J(2,2000,0); /*first row denotes the run length variable t, second row
the cdf value*/
perc=J(1,7,0); /*1st, 5th, 25th, 50th, 75th, 95th, 99th percentile*/

do b=1 to 2000;
 cdf[1,b]=b;
 cdf[2,b]=1-w*(Q**b)*one;
  if b>1 then do;
     if cdf[2,b-1]<0.01 & cdf[2,b]>=0.01 then perc[1,1]=b;
     else if cdf[2,b-1]<0.05 & cdf[2,b]>=0.05 then perc[1,2]=b;
     else if cdf[2,b-1]<0.25 & cdf[2,b]>=0.25 then perc[1,3]=b;
       else if cdf[2,b-1]<0.5 & cdf[2,b]>=0.5 then perc[1,4]=b;
       else if cdf[2,b-1]<0.75 & cdf[2,b]>=0.75 then perc[1,5]=b;
     else if cdf[2,b-1]<0.95 & cdf[2,b]>=0.95 then perc[1,6]=b;
     else if cdf[2,b-1]<0.99 & cdf[2,b]>=0.99 then perc[1,7]=b;
 end;
end;

print cdf;
print '1st 5th, 25th, 50th, 75th, 95th, 99th  percentile of run-length for
IC process';
print perc;

/*Process OOC*/
Shift=J(1,30,0);
```

```
do i=1 to 30;
 Shift[1,i]=0.2*i;
 end;
/*the distribution shifted - thus P1, P2 and P0 would differ to IC state*/
/*calcU is vector associated with UCL, calcL vector associated with LCL,
after shift occured*/
calcU=J(1,30,0);
calcL=J(1,30,0);
p1s=J(1,30,0);
p2s=J(1,30,0);
p0s=J(1,30,0);

/*calculate transition probabilities for different shifts*/
/*calculate OOC ARL*/
w=1;
I=1;
one=1;
ARL1=J(1,30,0);
SDRL1=J(1,30,0);

/*Cdf of run-length variable T : P(T=t)=1-Q^t*1 Fu and Lou (2003)*/
/*75th percentile P(T=t)=0.75 etc. */
/*calculate distribution of run length*/
/*For each shift in the process the percentiles have to be calculated*/
cdf1=J(31,2000,0); /*first row denotes the run length variable t, second-21
row denotes the cdf value of a process shift*/
perc1=J(30,7,0); /*25th, 50th, 75th percentiles for each process shift -
various rows denotes the shift*/

do j=1 to 30;
calcU[1,j]=(UCL-Shift[1,j])/1;
calcL[1,j]=(LCL-Shift[1,j])/1;
p1s[1,j]=1-cdf('normal',calcU[1,j],0,1);
p2s[1,j]=cdf('normal',calcL[1,j],0,1);
p0s[1,j]=1-p1s[1,j]-p2s[1,j];
ARL1[1,j]=w*(inv(I-p0s[1,j]))*one;
SDRL1[1,j]=sqrt((w*(I+p0s[1,j])*((I-p0s[1,j])**-2)*one)-ARL1[1,j]##2);

 do a=1 to 2000;
 cdf1[1,a]=a;
 cdf1[j+1,a]=1-w*(p0s[1,j]**a)*one;
  if a>1 then do;
     if cdf1[j+1,a-1]<0.01 & cdf1[j+1,a]>=0.01 then perc1[j,1]=a;
     if cdf1[j+1,a-1]<0.05 & cdf1[j+1,a]>=0.05 then perc1[j,2]=a;
     else if cdf1[j+1,a-1]<0.25 & cdf1[j+1,a]>=0.25 then perc1[j,3]=a;
       else if cdf1[j+1,a-1]<0.5 & cdf1[j+1,a]>=0.5 then perc1[j,4]=a;
       else if cdf1[j+1,a-1]<0.75 & cdf1[j+1,a]>=0.75 then perc1[j,5]=a;
     else if cdf1[j+1,a-1]<0.95 & cdf1[j+1,a]>=0.95 then perc1[j,6]=a;
     else if cdf1[j+1,a-1]<0.99 & cdf1[j+1,a]>=0.99 then perc1[j,7]=a;
 end;
end;
end;

print calcU, calcL, p1s, p2s, p0s, ARL0, ARL1;
print SDRL1;

print '1st 5th, 25th, 50th, 75th, 95th, 99th percentile of run-length for
various shift in the process mean';
print perc1;
quit;
run;
```

Shewhart $\overline{X}$ control chart – Simulation approach

```
proc iml;
reset nolog;
/*Assumption is that observations are distributed N(0,1)*/
/*Define UCL and LCL for Xbar chart*/
n=1;
UCL=0+3/sqrt(n);
LCL=0-3/sqrt(n);
/*Process*/
/*Generate charting statistics until a signal is observed*/
/*Charting statistic equals the average of 5 observations from a N(0,1)
process*/
/*Capture run-length*/
/*Repeat simulation 100000 times*/

x=J(n,1,.); /*define sample vector*/
mean_var=6; /*IC mean*/


/*rl= runlength variable*/
simnum=100000;
rl=J(simnum,1,0);

do i=1 to simnum;

chartstat=J(1,500000,0);
signal=0;
count=0;

do until (signal=1);
count=count+1;

call randgen(x,'normal', mean_var,1);
chartstat[1,count]=mean(x);  /*charting statistic is the mean of the
sample*/

if (chartstat[1,count]>=UCL) | (chartstat[1,count]<=LCL) then signal=1;
/*produces a signal if charting statistic plots on or above UCL or on or
below LCL*/

if count=500000 then signal=1; /*stops infinite loop*/

end;

rl[i,1]=count;

end;

create rldata from rl[colname={RL}];
append from rl;

proc univariate data=rldata noprint;
histogram;
inset mean std p1 p5 q1 median q3 p95 p99 /format=10.2;
run;
```

Shewhart $\overline{X}$ control chart – Exact approach

```
proc iml;
reset nolog;
/*Assumption is that observations distributed IID N(0,1)*/
/*Define UCL and LCL for Xbar chart*/
UCL=3;
LCL=-3;
/*p1 probability plotting on or above UCL*/
/*p2 probability plotting on or below LCL*/
/*p0 probability plotting between UCL and LCL*/

/*IC state*/
/*define transition probabilities*/
p1=1-cdf('normal',UCL,0,1);
p2=cdf('normal',LCL,0,1);
p0=1-p1-p2;
print p1,p2,p0;

/*run-length distribution follows a geometric distribution*/
/*ARL=1/FAR, FAR=p1+p2*/
ARL0=1/(p1+p2);
print ARL0;

SDRL0=sqrt(p0/(p1+p2));
print SDRL0;

cdf0=J(2,2000,0); /*first row denotes the run length variable t, second-21
row denotes the cdf value of a process shift*/
perc0=J(1,7,0); /*1st 5th, 25th, 50th, 75th, 95th, 99th percentiles for
each process shift - various rows denotes the shift*/
pdf0=J(1,2000,0); /*calculate pdf for each shift for various variations of
run-length*/

sum=0;
do a=1 to 2000;
 cdf0[1,a]=a;
 pdf0[1,a]=(1-p0)#(p0)##(a-1); /*pdf of geometric distribution*/
 sum=sum+pdf0[1,a];

cdf0[2,a]=sum;
*print cdf0, pdf0;
if a>1 then do;
     if cdf0[2,a-1]<0.01 & cdf0[2,a]>=0.01 then perc0[1,1]=a;
     else if cdf0[2,a-1]<0.05 & cdf0[2,a]>=0.05 then perc0[1,2]=a;
     else if cdf0[2,a-1]<0.25 & cdf0[2,a]>=0.25 then perc0[1,3]=a;
     else if cdf0[2,a-1]<0.5 & cdf0[2,a]>=0.5 then perc0[1,4]=a;
       else if cdf0[2,a-1]<0.75 & cdf0[2,a]>=0.75 then perc0[1,5]=a;
     else if cdf0[2,a-1]<0.95 & cdf0[2,a]>=0.95 then perc0[1,6]=a;
       else if cdf0[2,a-1]<0.99 & cdf0[2,a]>=0.99 then perc0[1,7]=a;
 end;
end;

print perc0;

/*OOC State*/
Shift=J(1,30,0);
do i=1 to 30;
Shift[1,i]=0.2*i;
end;
```

```
/*the distribution shifted - thus P1, P2 and P0 would differ for each
shift*/
/*calcU is vector associated with UCL, calcL vector associated with LCL,
after shift occured*/
calcU=J(1,30,0);
calcL=J(1,30,0);

p1s=J(1,30,0);
p2s=J(1,30,0);
p0s=J(1,30,0);

/*ARL1 denotes the OOC ARL for each process shift*/
ARL1=J(1,30,0);
SDRL1=J(1,30,0);
/*Since run-length follows Geometric distribution pfd  equals
f(x)=pq^(x-1)*/
/*cdf can be calculated from pdf*/
cdf1=J(31,2000,0); /*first row denotes the run length variable t, second-21
row denotes the cdf value of a process shift*/
perc1=J(30,7,0); /*25th, 50th, 75th percentiles for each process shift -
various rows denotes the shift*/
pdf1=J(30,2000,0); /*calculate pdf for each shift for various variations of
run-length*/

do j=1 to 30;
calcU[1,j]=(UCL-Shift[1,j]);
calcL[1,j]=(LCL-Shift[1,j]);
p1s[1,j]=1-cdf('normal',calcU[1,j],0,1);
p2s[1,j]=cdf('normal',calcL[1,j],0,1);
p0s[1,j]=1-p1s[1,j]-p2s[1,j];
ARL1[1,j]=1/(p1s[1,j]+p2s[1,j]);
SDRL1[1,j]=sqrt(p0s[1,j]/(p1s[1,j]+p2s[1,j]));
 sum=0;
do a=1 to 2000;
 cdf1[1,a]=a;
 pdf1[j,a]=(1-p0s[1,j])#(p0s[1,j])##(a-1); /*pdf of geometric
distribution*/
 sum=sum+pdf1[j,a];
 cdf1[j+1,a]=sum;
  if a>1 then do;
     if cdf1[j+1,a-1]<0.01 & cdf1[j+1,a]>=0.01 then perc1[j,1]=a;
     else if cdf1[j+1,a-1]<0.05 & cdf1[j+1,a]>=0.05 then perc1[j,2]=a;
     else if cdf1[j+1,a-1]<0.25 & cdf1[j+1,a]>=0.25 then perc1[j,3]=a;
     else if cdf1[j+1,a-1]<0.5 & cdf1[j+1,a]>=0.5 then perc1[j,4]=a;
       else if cdf1[j+1,a-1]<0.75 & cdf1[j+1,a]>=0.75 then perc1[j,5]=a;
     else if cdf1[j+1,a-1]<0.95 & cdf1[j+1,a]>=0.95 then perc1[j,6]=a;
       else if cdf1[j+1,a-1]<0.99 & cdf1[j+1,a]>=0.99 then perc1[j,7]=a;
 end;
end;

end;

print 'ARL1,SDRL1 calculated for shifts 0.1 to 2 in intervals of 0.1';
print ARL1, SDRL1;
print 'For each shift in process mean the 1st 5th, 25th, 50th, 75th, 95th,
99th percentile is calculated (denoted by the columns in perc1)';
print 'The shift ranges from 0.1 to 2 (denoted by the various rows)';
print perc1;

quit;
run;
```

The improved *2-of-2* control chart (Khoo et al. (2006)) - Markov chain approach

```
proc iml;
reset nolog;
/*Assumption is that observations distributed IID N(0,1)*/
/*Define UCL1, UCL2, LCL1, LCL2 for the chart*/
/*Variations of control limits (UCL2=3.4, UCL1=1.843), (UCL2=3.5,
UCL1=1.823), (UCL2=3.6, UCL1=1.81), (UCL2=3.7, UCL1=1.798), (UCL2=3.8,
UCL1=1.792) */
/*2-of-2 runs-type signalling rule*/
UCL2=3.4;
UCL1=1.843;
LCL2=-3.4;
LCL1=-1.843;

/*p1 probability plotting IC - between LCL1 ad UCL1*/
/*p2 probability plotting between UCL1 and UCL2*/
/*p3 probability plotting between LCL1 and LCL2*/
/*p4 probability plotting outside UCL2 or LCL2*/

/*Process IC*/
/*calculate transitional probabilities*/

p1=cdf('normal',UCL1,0,1)-cdf('normal',LCL1,0,1);
p2=cdf('normal',UCL2,0,1)-cdf('normal',UCL1,0,1);
p3=cdf('normal',-LCL2,0,1)-cdf('normal',-LCL1,0,1);

p4=1-p1-p2-p3;

print p1,p2, p3, p4;

/*set up essential TPM*/
/*ARL=w*(inv(I-Q))*1 Fu and Lou (2003)*/
w=J(1,3,0); w[1,1]=1;
I=I(3); /*identity matrix Fu and Lou (2003)*/
Q=J(3,3,0);

/*setting up essential TPM for IC process*/
Q[1,1]=p1; Q[2,1]=p1; Q[3,1]=p1;
Q[1,2]=p2; Q[3,2]=p2;
Q[1,3]=p3; Q[2,3]=p3;

one=J(3,1,1);
print w, I, Q, one;
ARL0=w*(inv(I-Q))*one;
print ARL0;
/*quit;*/
/*run;*/

SDRL=sqrt((w*(I+Q)*(inv(I-Q)**2)*one)-ARL0**2);
print SDRL;

/*Cdf of run-length variable T : P(T=t)=1-Q^t*1 Fu and Lou (2003)*/
/*75th percentile P(T=t)=0.75 etc. */
/*calculate distribution of run length*/
cdf=J(2,2000,0); /*first row denotes the run length variable t, second row
the cdf value*/
perc=J(1,7,0); /*25th, 50th, 75th percentile*/

do b=1 to 2000;
 cdf[1,b]=b;
```

```
 cdf[2,b]=1-w*(Q**b)*one;
  if b>1 then do;
     if cdf[2,b-1]<0.01 & cdf[2,b]>=0.01 then perc[1,1]=b;
     else if cdf[2,b-1]<0.05 & cdf[2,b]>=0.05 then perc[1,2]=b;
     else if cdf[2,b-1]<0.25 & cdf[2,b]>=0.25 then perc[1,3]=b;
       else if cdf[2,b-1]<0.5 & cdf[2,b]>=0.5 then perc[1,4]=b;
       else if cdf[2,b-1]<0.75 & cdf[2,b]>=0.75 then perc[1,5]=b;
     else if cdf[2,b-1]<0.95 & cdf[2,b]>=0.95 then perc[1,6]=b;
     else if cdf[2,b-1]<0.99 & cdf[2,b]>=0.99 then perc[1,7]=b;
end;
end;


print '1st 5th, 25th, 50th, 75th, 95th, 99th percentile of run-length for IC
process';
print perc;

/*Process OOC*/
/*calculate transitional probabilities*/
shift=J(30,1,0);
/*shift vector - various shifts used - in multiples of 0.2*/
do r=1 to 30;
 shift[r,1]=0.2*r;
end;

ARL1=J(30,1,0);
SDRL1=J(30,1,0);

/*Cdf of run-length variable T : P(T=t)=1-Q^t*1 Fu and Lou (2003)*/
/*75th percentile P(T=t)=0.75 etc. */
/*calculate distribution of run length*/
/*For each shift in the process the percentiles have to be calculated*/
cdf1=J(31,2000,0); /*first row denotes the run length variable t, second-21
row denotes the cdf value of a process shift*/
perc1=J(30,7,0); /*25th, 50th, 75th percentiles for each process shift -
various rows denotes the shift*/

do j=1 to 30;
p1=cdf('normal',UCL1-shift[j,1],0,1)-cdf('normal',LCL1-shift[j,1],0,1);
p2=cdf('normal',UCL2-shift[j,1],0,1)-cdf('normal',UCL1-shift[j,1],0,1);
p3=cdf('normal',-LCL2+shift[j,1],0,1)-cdf('normal',-LCL1+shift[j,1],0,1);

p4=1-p1-p2-p3;
/*set up essential TPM*/
/*ARL=w*(inv(I-Q))*1 Fu and Lou (2003)*/
/*w, I, and one defined as for IC state*/
/*setting up essential TPM for OOC process*/
Q[1,1]=p1; Q[2,1]=p1; Q[3,1]=p1;
Q[1,2]=p2; Q[3,2]=p2;
Q[1,3]=p3; Q[2,3]=p3;

/*print w, I, Q, one;*/
ARL1[j,1]=w*(inv(I-Q))*one;
SDRL1[j,1]=sqrt((w*(I+Q)*(inv(I-Q)**2)*one)-ARL1[j,1]##2);

do a=1 to 2000;
 cdf1[1,a]=a;
 cdf1[j+1,a]=1-w*(Q**a)*one;
  if a>1 then do;
     if cdf1[j+1,a-1]<0.01 & cdf1[j+1,a]>=0.01 then perc1[j,1]=a;
     else if cdf1[j+1,a-1]<0.05 & cdf1[j+1,a]>=0.05 then perc1[j,2]=a;
```

```
     else if cdf1[j+1,a-1]<0.25 & cdf1[j+1,a]>=0.25 then perc1[j,3]=a;
       else if cdf1[j+1,a-1]<0.5 & cdf1[j+1,a]>=0.5 then perc1[j,4]=a;
       else if cdf1[j+1,a-1]<0.75 & cdf1[j+1,a]>=0.75 then perc1[j,5]=a;
     else if cdf1[j+1,a-1]<0.95 & cdf1[j+1,a]>=0.95 then perc1[j,6]=a;
       else if cdf1[j+1,a-1]<0.99 & cdf1[j+1,a]>=0.99 then perc1[j,7]=a;
 end;
end;

end;
print ARL1;

print SDRL1;

print '1st 5th, 25th, 50th, 75th, 95th, 99th percentile of run-length for
OOC process';
print perc1;

quit;
run;
```

The improved *2-of-2* control chart (Khoo et al. (2006)) – Simulation approach

```
proc iml;
reset nolog;
/*Assumption is that observations are distributed N(0,1), for shift
N(shift,1)*/
/*signal is produced if 1 charting statistic plots above/below UCL2/LCL2 or
two consecutive charting statistics plot above/below UCL1/LCL1*/
n=5;
stdev=1;
/*Inner upper and lower control limits*/
UCL1=0+(1.843*stdev)/sqrt(n);
LCL1=0-(1.843*stdev)/sqrt(n);
/*Outer upper and lower control limits*/
UCL2=0+(3.4*stdev)/sqrt(n);
LCL2=0-(3.4*stdev)/sqrt(n);
print UCL1, LCL1, UCL2, LCL2;

x=J(n,1,.); /*define sample vector*/
eta=6;
mean_var=eta*stdev/sqrt(n); /*OOC mean*/

/*rl= runlength variable*/
simnum=100000;
rl=J(simnum,1,0);

do i=1 to simnum;

chartstat=J(1,500000,0);
signal=0;
count=0;

do until (signal=1);
count=count+1;

call randgen(x,'normal', mean_var,1);
chartstat[1,count]=mean(x);  /*charting statistic is the mean of the
sample*/
```

```
if (chartstat[1,count]>=UCL2) | (chartstat[1,count]<=LCL2) then signal=1;
/*produces a signal if charting statistic plots on or above UCL2 or on or
below LCL2*/
/*2-of-2 rule, check if above or below UCL1/LCL1*/
/*no signal yet*/
if count>1  & signal=0 then do;
 if ((chartstat[1,count-1]>=UCL1) & (chartstat[1,count]>=UCL1))  |
((chartstat[1,count-1]<=LCL1) & (chartstat[1,count]<=LCL1)) then signal=1;
end;

if count=500000 then signal=1; /*stops infinite loop*/
end;

rl[i,1]=count;

end;

create rldata from rl[colname={RL}];
append from rl;

proc univariate data=rldata noprint;
histogram;
inset mean std p1 p5 q1 median q3 p95 p99/format=10.2;
run;

quit;
run;
```

The improved *2-of-2* control chart (Khoo et al. (2006)) – Exact approach

```
proc iml;
reset nolog;
/*Define UCL1, UCL2, LCL1, LCL2 for the chart*/
/*Variations of control limits (UCL2=3.4, UCL1=1.843), (UCL2=3.5,
UCL1=1.823), (UCL2=3.6, UCL1=1.81), (UCL2=3.7, UCL1=1.798), (UCL2=3.8,
UCL1=1.792) */
/*2-of-2 runs-type signalling rule*/
UCL2=3.4;
UCL1=1.843;
LCL2=-3.4;
LCL1=-1.843;

/*p1 probability plotting IC - between LCL1 ad UCL1*/
/*p2 probability plotting between UCL1 and UCL2*/
/*p3 probability plotting between LCL1 and LCL2*/
/*p4 probability plotting outside UCL2 or LCL2*/

/*Process IC*/
/*calculate transitional probabilities*/

p1=cdf('normal',UCL1,0,1)-cdf('normal',LCL1,0,1);
p2=cdf('normal',UCL2,0,1)-cdf('normal',UCL1,0,1);
p3=cdf('normal',-LCL2,0,1)-cdf('normal',-LCL1,0,1);

p4=1-p1-p2-p3;

print p1,p2, p3, p4;

ARL0=(1+p2+p2#p3+p3)/(1-p2#p3-p1-p1#p2-p1#p2#p3-p3#p1);
```

```
print ARL0;

/*OOC*/
shift=J(18,1,0);
/*shift vector - various shifts used - in multiples of 0.2*/

ARL1=J(18,1,0);

do r=1 to 15;
 shift[r,1]=0.2*r;
end;
shift[16,1]=4; shift[17,1]=5; shift[18,1]=6;

do j=1 to 18;
p1=cdf('normal',UCL1-shift[j,1],0,1)-cdf('normal',LCL1-shift[j,1],0,1);
p2=cdf('normal',UCL2-shift[j,1],0,1)-cdf('normal',UCL1-shift[j,1],0,1);
p3=cdf('normal',-LCL2+shift[j,1],0,1)-cdf('normal',-LCL1+shift[j,1],0,1);

p4=1-p1-p2-p3;

ARL1[j,1]=(1+p2+p2#p3+p3)/(1-p2#p3-p1-p1#p2-p1#p2#p3-p3#p1);

end;

print ARL1;

quit;
run;
```

The revised *2-of-3* control chart (Antzoulakos and Rakitzis (2008a)) - Markov chain approach

```
proc iml;
reset nolog;
/*Assumption is that observations distributed IID N(0,1)*/
/*Define UCL1, UCL2, LCL1, LCL2 for the chart*/
/*Variations of control limits (UCL2=3.4, UCL1=1.926), (UCL2=3.5,
UCL1=1.906), (UCL2=3.6, UCL1=1.892), (UCL2=3.7, UCL1=1.884), (UCL2=3.8,
UCL1=1.878) */
/*2-of-3 runs-type signalling rule*/
UCL2=3.5;
UCL1=1.906;
LCL2=-3.5;
LCL1=-1.906;

/*p1 probability plotting betwee UCL1 and UCL2*/
/*p2 probability plotting between CL and UCL1*/
/*p3 probability plotting between CL and LCL1*/
/*p4 probability plotting betwee LCL1 and LCL2*/
/*p5 probability plotting outside UCL2 or LCL2*/

/*Process IC*/
/*calculate transitional probabilities*/
p1=cdf('normal',UCL2,0,1)-cdf('normal',UCL1,0,1);
p2=cdf('normal',UCL1,0,1)+cdf('normal',0,0,1)-1;
p3=cdf('normal',-LCL1,0,1)-cdf('normal',0,0,1);
p4=cdf('normal',-LCL2,0,1)-cdf('normal',-LCL1,0,1);

/*p4=p1;
```

```
p3=p2;*/
p5=1-p1-p2-p3-p4;
print p1,p2, p3, p4, p5;

/*set up essential TPM*/
/*ARL=w*(inv(I-Q))*1 Fu and Lou (2003)*/
w=J(1,5,0); w[1,1]=1;
I=I(5); /*identity matrix Fu and Lou (2003)*/
Q=J(5,5,0);
/*setting up essential TPM for IC process*/
Q[1,1]=p2+p3; Q[3,1]=p2+p3; Q[5,1]=p2+p3;
Q[2,1]=p3; Q[4,5]=p3;
Q[4,1]=p2; Q[2,3]=p2;
Q[1,4]=p4;  Q[2,4]=p4; Q[3,4]=p4;
Q[1,2]=p1; Q[4,2]=p1; Q[5,2]=p1;

one=J(5,1,1);
print w, I, Q, one;
ARL0=w*(inv(I-Q))*one;
print ARL0;

inverse=J(5,5,0);
inverse=inv(I-Q);
print inverse;

SDRL=sqrt((w*(I+Q)*(inv(I-Q)**2)*one)-ARL0**2);
print SDRL;

/*Cdf of run-length variable T : P(T=t)=1-Q^t*1 Fu and Lou (2003)*/
/*75th percentile P(T=t)=0.75 etc. */
/*calculate distribution of run length*/
cdf=J(2,2000,0); /*first row denotes the run length variable t, second row
the cdf value*/
perc=J(1,7,0); /*1st 5th, 25th, 50th, 75th, 95th, 99th percentile */

do b=1 to 2000;
 cdf[1,b]=b;
 cdf[2,b]=1-w*(Q**b)*one;
  if b>1 then do;
     if cdf[2,b-1]<0.01 & cdf[2,b]>=0.01 then perc[1,1]=b;
     else if cdf[2,b-1]<0.05 & cdf[2,b]>=0.05 then perc[1,2]=b;
     else if cdf[2,b-1]<0.25 & cdf[2,b]>=0.25 then perc[1,3]=b;
       else if cdf[2,b-1]<0.5 & cdf[2,b]>=0.5 then perc[1,4]=b;
       else if cdf[2,b-1]<0.75 & cdf[2,b]>=0.75 then perc[1,5]=b;
     else if cdf[2,b-1]<0.95 & cdf[2,b]>=0.95 then perc[1,6]=b;
     else if cdf[2,b-1]<0.99 & cdf[2,b]>=0.99 then perc[1,7]=b;
end;
end;

print '5th, 25th, 50th, 75th and 95th percentile of run-length for IC
process';
print perc;

/*Process OOC*/
/*calculate transitional probabilities*/
/*test for shift of 0.2*/
/*shf=0.2;*//*shift in mean=0.2 units*/
shift=J(30,1,0);
/*shift vector - various shifts used - in multiples of 0.2*/
do r=1 to 30;
 shift[r,1]=0.2*r;
```

```
end;

ARL1=J(30,1,0);
SDRL1=J(30,1,0);

/*Cdf of run-length variable T : P(T=t)=1-Q^t*1 Fu and Lou (2003)*/
/*75th percentile P(T=t)=0.75 etc. */
/*calculate distribution of run length*/
/*For each shift in the process the percentiles have to be calculated*/
cdf1=J(31,2000,0); /*first row denotes the run length variable t, second-21
row denotes the cdf value of a process shift*/
perc1=J(30,7,0); /*25th, 50th, 75th percentiles for each process shift -
various rows denotes the shift*/

do j=1 to 30;
p1=cdf('normal',UCL2-shift[j,1],0,1)-cdf('normal',UCL1-shift[j,1],0,1);
p2=cdf('normal',UCL1-shift[j,1],0,1)+cdf('normal',shift[j,1],0,1)-1;
p3=cdf('normal',-LCL1+shift[j,1],0,1)-cdf('normal',shift[j,1],0,1);
p4=cdf('normal',-LCL2+shift[j,1],0,1)-cdf('normal',-LCL1+shift[j,1],0,1);

p5=1-p1-p2-p3-p4;
/*set up essential TPM*/
/*ARL=w*(inv(I-Q))*1 Fu and Lou (2003)*/
/*w, I, and one defined as for IC state*/
/*setting up essential TPM for OOC process*/
Q[1,1]=p2+p3; Q[3,1]=p2+p3; Q[5,1]=p2+p3;
Q[2,1]=p3; Q[4,5]=p3;
Q[4,1]=p2; Q[2,3]=p2;
Q[1,4]=p4;  Q[2,4]=p4; Q[3,4]=p4;
Q[1,2]=p1; Q[4,2]=p1; Q[5,2]=p1;

/*print w, I, Q, one;*/
ARL1[j,1]=w*(inv(I-Q))*one;
SDRL1[j,1]=sqrt((w*(I+Q)*(inv(I-Q)**2)*one)-ARL1[j,1]##2);

do a=1 to 2000;
 cdf1[1,a]=a;
 cdf1[j+1,a]=1-w*(Q**a)*one;
  if a>1 then do;
     if cdf1[j+1,a-1]<0.01 & cdf1[j+1,a]>=0.01 then perc1[j,1]=a;
     else if cdf1[j+1,a-1]<0.05 & cdf1[j+1,a]>=0.05 then perc1[j,2]=a;
     else if cdf1[j+1,a-1]<0.25 & cdf1[j+1,a]>=0.25 then perc1[j,3]=a;
       else if cdf1[j+1,a-1]<0.5 & cdf1[j+1,a]>=0.5 then perc1[j,4]=a;
       else if cdf1[j+1,a-1]<0.75 & cdf1[j+1,a]>=0.75 then perc1[j,5]=a;
     else if cdf1[j+1,a-1]<0.95 & cdf1[j+1,a]>=0.95 then perc1[j,6]=a;
       else if cdf1[j+1,a-1]<0.99 & cdf1[j+1,a]>=0.99 then perc1[j,7]=a;
 end;
end;

end;
print ARL1;

print SDRL1;

print '1st 5th, 25th, 50th, 75th, 95th, 99th percentile of run-length for
OOC process';
print perc1;

quit;
run;
```

The revised *2-of-3* control chart (Antzoulakos and Rakitzis (2008a)) – Simulation approach

```
/*Simulation approach Improved 2-of-3 runs-type signalling rule -
Rakitzis*/
/*Assumption is that observations are distributed N(0,1), for shift
N(shift,1)*/
proc iml;
reset nolog;
/*signal is produced if 1 charting statistic plots above/below UCL2/LCL2 or
two consecutive charting statistics of three plot above/below UCL1/LCL1 and
below/above UCL2/LCL2*/
/*in cluster of points between CL and UCL2/LCL2*/
n=5;
stdev=1;
/*Inner upper and lower control limits*/
UCL1=0+(1.906*stdev)/sqrt(n);
LCL1=0-(1.906*stdev)/sqrt(n);
/*Outer upper and lower control limits*/
UCL2=0+(3.5*stdev)/sqrt(n);
LCL2=0-(3.5*stdev)/sqrt(n);
CL=0;
print UCL1, LCL1, UCL2, LCL2;

x=J(n,1,.); /*define sample vector*/
eta=0;
mean_var=eta*stdev/sqrt(n); /*OOC mean*/

/*rl= runlength variable*/
simnum=100000;
rl=J(simnum,1,0);

do i=1 to simnum;

chartstat=J(1,500000,0);
signal=0;
count=0;

do until (signal=1);
count=count+1;

call randgen(x,'normal', mean_var,1);
chartstat[1,count]=mean(x);  /*charting statistic is the mean of the
sample*/

if (chartstat[1,count]>=UCL2) | (chartstat[1,count]<=LCL2) then signal=1;
/*produces a signal if charting statistic plots on or above UCL2 or on or
below LCL2*/

/*2-of-3 rule, check if above or below UCL1/LCL1*/
/*no signal yet*/
if count>2  & signal=0 then do;
/*signal could be 101 or 011 for 2-of-3 rule*/
/*for 101 signal*/
if ( ((chartstat[1,count-2]>=UCL1) &(chartstat[1,count-2]<UCL2)) &
((chartstat[1,count-1]>=CL) & (chartstat[1,count-1]<=UCL1)) &
((chartstat[1,count]>=UCL1) &(chartstat[1,count]<UCL2)) )
```

```
| ( ((chartstat[1,count-2]<=LCL1) &(chartstat[1,count-2]>LCL2))&
((chartstat[1,count-1]<=CL) & (chartstat[1,count-1]>=LCL1)) &
((chartstat[1,count]<=LCL1) &(chartstat[1,count]>LCL2)) )
then signal=1;
/*for 011 signal*/
if ( ((chartstat[1,count-2]>=CL) &(chartstat[1,count-2]<=UCL1)) &
((chartstat[1,count-1]>=UCL1) & (chartstat[1,count-1]<UCL2)) &
((chartstat[1,count]>=UCL1) &(chartstat[1,count]<UCL2)) )
| ( ((chartstat[1,count-2]<=CL) &(chartstat[1,count-2]>=LCL1))&
((chartstat[1,count-1]<=LCL1) & (chartstat[1,count-1]>LCL2)) &
((chartstat[1,count]<=LCL1) &(chartstat[1,count]>LCL2)) )
then signal=1;

if ( ((chartstat[1,count-2]>=UCL1) &(chartstat[1,count-2]<=UCL2))
&((chartstat[1,count-1]>=UCL1) &(chartstat[1,count-1]<=UCL2))  )
| ( ((chartstat[1,count-2]<=LCL1) &(chartstat[1,count-2]>=LCL2))&
((chartstat[1,count-1]<=LCL1) &(chartstat[1,count-1]>=LCL2))  )
then signal=1;

end;


if count=500000 then signal=1; /*stops infinite loop*/

end; /*end of do until*/

rl[i,1]=count;
end;

create rldata from rl[colname={RL}];
append from rl;

proc univariate data=rldata noprint;
histogram;
inset mean std p1 p5 q1 median q3 p95 p99 min max/format=10.2;
run;

quit;
run;
```

The revised *2-of-3* control chart (Antzoulakos and Rakitzis (2008a)) – Exact approach

```
proc iml;
reset nolog;
/*Define UCL1, UCL2, LCL1, LCL2 for the chart*/
/*Variations of control limits (UCL2=3.4, UCL1=1.926), (UCL2=3.5,
UCL1=1.906), (UCL2=3.6, UCL1=1.892), (UCL2=3.7, UCL1=1.884), (UCL2=3.8,
UCL1=1.878) */
/*2-of-3 runs-type signalling rule*/
UCL2=3.5;
UCL1=1.906;
LCL2=-3.5;
LCL1=-1.906;

/*p1 probability plotting betwee UCL1 and UCL2*/
/*p2 probability plotting between CL and UCL1*/
/*p3 probability plotting between CL and LCL1*/
/*p4 probability plotting betwee LCL1 and LCL2*/
/*p5 probability plotting outside UCL2 or LCL2*/
```

```
/*Process IC*/
/*calculate transitional probabilities*/
p1=cdf('normal',UCL2,0,1)-cdf('normal',UCL1,0,1);
p2=cdf('normal',UCL1,0,1)+cdf('normal',0,0,1)-1;
p3=cdf('normal',-LCL1,0,1)-cdf('normal',0,0,1);
p4=cdf('normal',-LCL2,0,1)-cdf('normal',-LCL1,0,1);
p5=1-p1-p2-p3-p4;
print p1,p2,p3,p4,p5;
/*set up essential TPM*/
/*ARL=w*(inv(I-Q))*1 Fu and Lou (2003)*/
w=J(1,5,0); w[1,1]=1;
I=I(5); /*identity matrix Fu and Lou (2003)*/
Q=J(5,5,0);
/*setting up essential TPM for IC process*/
Q[1,1]=p2+p3; Q[3,1]=p2+p3; Q[5,1]=p2+p3;
Q[2,1]=p3; Q[4,5]=p3;
Q[4,1]=p2; Q[2,3]=p2;
Q[1,4]=p4;  Q[2,4]=p4; Q[3,4]=p4;
Q[1,2]=p1; Q[4,2]=p1; Q[5,2]=p1;

inverse=J(5,5,0);
inverse=inv(I-Q);
print inverse;

ARL0=inverse[1,+];
print ARL0;

/*OOC*/
shift=J(30,1,0);
/*shift vector - various shifts used - in multiples of 0.2*/
do r=1 to 30;
 shift[r,1]=0.2*r;
end;

ARL1=J(30,1,0);
SDRL1=J(30,1,0);

do j=1 to 30;
p1=cdf('normal',UCL2-shift[j,1],0,1)-cdf('normal',UCL1-shift[j,1],0,1);
p2=cdf('normal',UCL1-shift[j,1],0,1)+cdf('normal',shift[j,1],0,1)-1;
p3=cdf('normal',-LCL1+shift[j,1],0,1)-cdf('normal',shift[j,1],0,1);
p4=cdf('normal',-LCL2+shift[j,1],0,1)-cdf('normal',-LCL1+shift[j,1],0,1);
p5=1-p1-p2-p3-p4;

/*set up essential TPM*/
/*ARL=w*(inv(I-Q))*1 Fu and Lou (2003)*/
/*w, I, and one defined as for IC state*/
/*setting up essential TPM for OOC process*/
Q[1,1]=p2+p3; Q[3,1]=p2+p3; Q[5,1]=p2+p3;
Q[2,1]=p3; Q[4,5]=p3;
Q[4,1]=p2; Q[2,3]=p2;
Q[1,4]=p4;  Q[2,4]=p4; Q[3,4]=p4;
Q[1,2]=p1; Q[4,2]=p1; Q[5,2]=p1;

inverse1=J(5,5,0);
inverse1=inv(I-Q);
print inverse1;

ARL1[j,1]=inverse1[1,+];
end;
```

```
print ARL1;

quit;
run;
```

The *S*-chart supplemented with the *2-of-2* runs-type signalling rule (Anzoulakos et al. (2010) - Markov chain approach

```
proc iml;
reset nolog;
/*Assumption is that observations distributed IID N(0,1)*/
/*Define UCL1, UCL2, LCL1, LCL2 for the chart*/
UCL=2.145;
UWL=1.603;
LWL=0.417;
LCL=0.0009;/*<0.001*/

/*p1 probability plotting IC - between LWL ad UWL*/
/*p2 probability plotting between UWL and UCL*/
/*p3 probability plotting between LWL and LCL*/
/*p4 probability plotting outside UCL*/
/*p5 probability plotting outside LCL*/

/*Process IC*/
/*calculate transitional probabilities*/
/*cdf with 4 degrees of freedom*/
n=5;
var=1;  /*1 indicates no shift*/
T_UWL=(n-1)*(UWL##2)/(var##2);
T_LWL=(n-1)*(LWL##2)/(var##2);
T_UCL=(n-1)*(UCL##2)/(var##2);
T_LCL=(n-1)*(LCL##2)/(var##2);


p1=cdf('CHISQUARE',T_UWL,4)-cdf('CHISQUARE',T_LWL,4);
p2=cdf('CHISQUARE',T_UCL,4)-cdf('CHISQUARE',T_UWL,4);
p3=cdf('CHISQUARE',T_LWL,4)-cdf('CHISQUARE',T_LCL,4);
p4=1-cdf('CHISQUARE',T_UCL,4);

p5=1-p1-p2-p3-p4;
print p1,p2, p3, p4, p5;

/*set up essential TPM*/
/*ARL=w*(inv(I-Q))*1 Fu and Lou (2003)*/
w=J(1,3,0); w[1,1]=1;
I=I(3); /*identity matrix Fu and Lou (2003)*/
Q=J(3,3,0);

/*setting up essential TPM for IC process*/
Q[1,1]=p1; Q[2,1]=p1; Q[3,1]=p1;
Q[1,2]=p2; Q[3,2]=p2;
Q[1,3]=p3; Q[2,3]=p3;

one=J(3,1,1);
print w, I, Q, one;
ARL0=w*(inv(I-Q))*one;
print ARL0;

SDRL=sqrt((w*(I+Q)*(inv(I-Q)**2)*one)-ARL0**2);
```

```
print SDRL;

/*Cdf of run-length variable T : P(T=t)=1-Q^t*1 Fu and Lou (2003)*/
/*75th percentile P(T=t)=0.75 etc. */
/*calculate distribution of run length*/
cdf=J(2,2000,0); /*first row denotes the run length variable t, second row
the cdf value*/
perc=J(1,7,0); /*25th, 50th, 75th percentile*/

do b=1 to 2000;
 cdf[1,b]=b;
 cdf[2,b]=1-w*(Q**b)*one;
  if b>1 then do;
     if cdf[2,b-1]<0.01 & cdf[2,b]>=0.01 then perc[1,1]=b;
     else if cdf[2,b-1]<0.05 & cdf[2,b]>=0.05 then perc[1,2]=b;
     else if cdf[2,b-1]<0.25 & cdf[2,b]>=0.25 then perc[1,3]=b;
       else if cdf[2,b-1]<0.5 & cdf[2,b]>=0.5 then perc[1,4]=b;
       else if cdf[2,b-1]<0.75 & cdf[2,b]>=0.75 then perc[1,5]=b;
     else if cdf[2,b-1]<0.95 & cdf[2,b]>=0.95 then perc[1,6]=b;
       else if cdf[2,b-1]<0.99 & cdf[2,b]>=0.99 then perc[1,7]=b;
 end;
end;


print '1st 5th, 25th, 50th, 75th, 95th, 99th percentile of run-length for IC
process';
print perc;

/*Process OOC*/
/*calculate transitional probabilities*/
shift=J(30,1,0);
/*shift vector - various shifts used - in multiples of 0.2*/
do r=1 to 30;
 shift[r,1]=0.2*r;
end;

ARL1=J(30,1,0);
SDRL1=J(30,1,0);

/*Cdf of run-length variable T : P(T=t)=1-Q^t*1 Fu and Lou (2003)*/
/*75th percentile P(T=t)=0.75 etc. */
/*calculate distribution of run length*/
/*For each shift in the process the percentiles have to be calculated*/
cdf1=J(31,2000,0); /*first row denotes the run length variable t, second-21
row denotes the cdf value of a process shift*/
perc1=J(30,7,0); /*25th, 50th, 75th percentiles for each process shift -
various rows denotes the shift*/

do j=1 to 30;

var=shift[j,1];
T_UWL=(n-1)*(UWL##2)/(var##2);
T_LWL=(n-1)*(LWL##2)/(var##2);
T_UCL=(n-1)*(UCL##2)/(var##2);
T_LCL=(n-1)*(LCL##2)/(var##2);

p1=cdf('CHISQUARE',T_UWL,4)-cdf('CHISQUARE',T_LWL,4);
p2=cdf('CHISQUARE',T_UCL,4)-cdf('CHISQUARE',T_UWL,4);
p3=cdf('CHISQUARE',T_LWL,4)-cdf('CHISQUARE',T_LCL,4);
p4=1-cdf('CHISQUARE',T_UCL,4);
p5=1-p1-p2-p3-p4;
```

```
/*set up essential TPM*/
/*ARL=w*(inv(I-Q))*1 Fu and Lou (2003)*/
/*w, I, and one defined as for IC state*/

/*setting up essential TPM for OOC process*/
Q[1,1]=p1; Q[2,1]=p1; Q[3,1]=p1;
Q[1,2]=p2; Q[3,2]=p2;
Q[1,3]=p3; Q[2,3]=p3;

ARL1[j,1]=w*(inv(I-Q))*one;
SDRL1[j,1]=sqrt((w*(I+Q)*(inv(I-Q)**2)*one)-ARL1[j,1]##2);

do a=1 to 2000;
 cdf1[1,a]=a;
 cdf1[j+1,a]=1-w*(Q**a)*one;
  if a>1 then do;
     if cdf1[j+1,a-1]<0.01 & cdf1[j+1,a]>=0.01 then perc1[j,1]=a;
     else if cdf1[j+1,a-1]<0.05 & cdf1[j+1,a]>=0.05 then perc1[j,2]=a;
     else if cdf1[j+1,a-1]<0.25 & cdf1[j+1,a]>=0.25 then perc1[j,3]=a;
       else if cdf1[j+1,a-1]<0.5 & cdf1[j+1,a]>=0.5 then perc1[j,4]=a;
       else if cdf1[j+1,a-1]<0.75 & cdf1[j+1,a]>=0.75 then perc1[j,5]=a;
     else if cdf1[j+1,a-1]<0.95 & cdf1[j+1,a]>=0.95 then perc1[j,6]=a;
       else if cdf1[j+1,a-1]<0.99 & cdf1[j+1,a]>=0.99 then perc1[j,7]=a;
 end;
end;

end;
print ARL1;
print SDRL1;

print '1st 5th, 25th, 50th, 75th, 95th, 99th percentile of run-length for
OOC process';
print perc1;

quit;
run;
```

The *S*-chart supplemented with the *2-of-2* runs-type signalling rule (Anzoulakos et al. (2010) - Simulation approach

```
proc iml;
reset nolog;
/*Assumption is that observations are distributed N(0,1), for shift
N(shift,1)*/
/*signal is produced if 1 charting statistic plots above/below UCL/LCL or
two consecutive charting statistics plot between UWL/UCL or between
LWL/LCL*/
n=5;
/*Control and Warning limits*/
UCL=2.145;
UWL=1.603;
LWL=0.417;
LCL=0.00099; /*Less than 0.001*/
print UCL, UWL, LWL, LCL;
x=J(n,1,.); /*define sample vector*/
/*charting statistic the standard deviation*/
mean_var=0;
```

```
/*rl= runlength variable*/
simnum=100000;
rl=J(simnum,1,0);

do i=1 to simnum;

chartstat=J(1,500000,0);
signal=0;
count=0;

do until (signal=1);
count=count+1;

call randgen(x,'normal', mean_var,6); /*shifts in standard deviations are
monitored ranging from 0.2 to 6*/
chartstat[1,count]=STD(x);  /*charting statistic is the standard deviation
of the sample*/

if (chartstat[1,count]>=UCL) | (chartstat[1,count]<=LCL) then signal=1;
/*produces a signal if charting statistic plots on or above UCL or on or
below LCL*/

/*2-of-2 rule, check if above or below UWL/LWL*/
/*no signal yet*/
if count>1  & signal=0 then do;
 if ((chartstat[1,count-1]>=UWL) & (chartstat[1,count]>=UWL))  |
((chartstat[1,count-1]<=LWL) & (chartstat[1,count]<=LWL)) then signal=1;
end;

if count=500000 then signal=1; /*stops infinite loop*/

end;

rl[i,1]=count;

end;

create rldata from rl[colname={RL}];
append from rl;

proc univariate data=rldata noprint;
histogram;
inset mean std p1 p5 q1 median q3 p95 p99 /format=10.2;
run;

quit;
run;
```

The *S*-chart supplemented with the *2-of-2* runs-type signalling rule (Anzoulakos et al. (2010) - Exact approach

```
proc iml;
reset nolog;
/*Define UCL1, UCL2, LCL1, LCL2 for the chart*/
/*2-of-2 runs-type signalling rule*/
UCL=2.145;
UWL=1.603;
LWL=0.417;
LCL=0.0009;/*<0.001*/
```

```
/*p1 probability plotting IC - between LWL ad UWL*/
/*p2 probability plotting between UWL and UCL*/
/*p3 probability plotting between LWL and LCL*/
/*p4 probability plotting outside UCL*/
/*p5 probability plotting outside LCL*/

/*Process IC*/
/*calculate transitional probabilities*/
/*cdf with 4 degrees of freedom*/
n=5;
var=1;  /*1 indicates no shift*/
T_UWL=(n-1)*(UWL##2)/(var##2);
T_LWL=(n-1)*(LWL##2)/(var##2);
T_UCL=(n-1)*(UCL##2)/(var##2);
T_LCL=(n-1)*(LCL##2)/(var##2);

p1=cdf('CHISQUARE',T_UWL,4)-cdf('CHISQUARE',T_LWL,4);
p2=cdf('CHISQUARE',T_UCL,4)-cdf('CHISQUARE',T_UWL,4);
p3=cdf('CHISQUARE',T_LWL,4)-cdf('CHISQUARE',T_LCL,4);
p4=1-cdf('CHISQUARE',T_UCL,4);
p5=1-p1-p2-p3-p4;
print p1,p2, p3, p4, p5;
ARL0=(1+p2+p2#p3+p3)/(1-p2#p3-p1-p1#p2-p1#p2#p3-p3#p1);
print ARL0;

/*OOC*/
shift=J(18,1,0);
/*shift vector - various shifts used - in multiples of 0.2*/
ARL1=J(18,1,0);

do r=1 to 15;
 shift[r,1]=0.2*r;
end;
shift[16,1]=4; shift[17,1]=5; shift[18,1]=6;


do j=1 to 18;
var=shift[j,1];
T_UWL=(n-1)*(UWL##2)/(var##2);
T_LWL=(n-1)*(LWL##2)/(var##2);
T_UCL=(n-1)*(UCL##2)/(var##2);
T_LCL=(n-1)*(LCL##2)/(var##2);

/*n-1=4*/
p1=cdf('CHISQUARE',T_UWL,4)-cdf('CHISQUARE',T_LWL,4);
p2=cdf('CHISQUARE',T_UCL,4)-cdf('CHISQUARE',T_UWL,4);
p3=cdf('CHISQUARE',T_LWL,4)-cdf('CHISQUARE',T_LCL,4);
p4=1-cdf('CHISQUARE',T_UCL,4);

p5=1-p1-p2-p3-p4;
ARL1[j,1]=(1+p2+p2#p3+p3)/(1-p2#p3-p1-p1#p2-p1#p2#p3-p3#p1);

end;

print ARL1;

quit;
run;
```

Nonparametric control chart supplemented with the *2-of-2* runs-type signalling rule (Human et al. (2010)) - Markov chain approach

```
proc iml;
reset nolog;
/*n=5, b=0, a=0*/
UCL=5;
LCL=0;
n=5; b=0; a=0;

p1=1-cdf('binom',n-b-1,0.5,n);
print p1;
p2=cdf('binom',a,0.5,n);
print p2;
p0=1-p1-p2;

/*set up essential TPM*/
/*ARL=w*(inv(I-Q))*1 Fu and Lou (2003)*/
w=J(1,3,0); w[1,1]=1;
I=I(3); /*identity matrix Fu and Lou (2003)*/
Q=J(3,3,0);

/*setting up essential TPM for IC process*/
Q[1,1]=p0; Q[2,1]=p0; Q[3,1]=p0;
Q[1,2]=p1; Q[3,2]=p1;
Q[1,3]=p2; Q[2,3]=p2;

one=J(3,1,1);
print w, I, Q, one;
ARL0=w*(inv(I-Q))*one;
print ARL0;

SDRL=sqrt((w*(I+Q)*(inv(I-Q)**2)*one)-ARL0**2);
print SDRL;

/*Cdf of run-length variable T : P(T=t)=1-Q^t*1 Fu and Lou (2003)*/
/*75th percentile P(T=t)=0.75 etc. */
/*calculate distribution of run length*/
cdf=J(2,3000,0); /*first row denotes the run length variable t, second row
the cdf value*/
perc=J(1,7,0); /*1st 5th, 25th, 50th, 75th, 95th, 99th percentile */

do b=1 to 3000;
 cdf[1,b]=b;
 cdf[2,b]=1-w*(Q**b)*one;
  if b>1 then do;
     if cdf[2,b-1]<0.01 & cdf[2,b]>=0.01 then perc[1,1]=b;
     else if cdf[2,b-1]<0.05 & cdf[2,b]>=0.05 then perc[1,2]=b;
     else if cdf[2,b-1]<0.25 & cdf[2,b]>=0.25 then perc[1,3]=b;
       else if cdf[2,b-1]<0.5 & cdf[2,b]>=0.5 then perc[1,4]=b;
       else if cdf[2,b-1]<0.75 & cdf[2,b]>=0.75 then perc[1,5]=b;
     else if cdf[2,b-1]<0.95 & cdf[2,b]>=0.95 then perc[1,6]=b;
     else if cdf[2,b-1]<0.99 & cdf[2,b]>=0.99 then perc[1,7]=b;
end;
end;


print '1st 5th, 25th, 50th, 75th, 95th, 99th percentile of run-length for IC
process';
print perc;
```

```
/*OOC*/
shift=J(18,1,0);

do r=1 to 15;
 shift[r,1]=0.2*r;
end;
shift[16,1]=4; shift[17,1]=5; shift[18,1]=6;

ARL1=J(18,1,0);
SDRL1=J(18,1,0);

n=5; b=0; a=0;
/*calculate probability p for the charting statistic that follows a
binomial (n,p) distribution*/
binomp=J(18,1,0);
p1=J(18,1,0);
p2=J(18,1,0);
p0=J(18,1,0);
perc1=J(18,7,0);

   do z=1 to 18;
    binomp[z,1]=1-cdf('normal',0,shift[z,1],1);
    p1[z,1]=1-cdf('binom',n-b-1, binomp[z,1],n);
    p2[z,1]=cdf('binom',a, binomp[z,1],n);
    p0[z,1]=1-p1[z,1]-p2[z,1];

    Q=J(3,3,0);
    /*setting up essential TPM for OOC process*/
    Q[1,1]=p0[z,1]; Q[2,1]=p0[z,1]; Q[3,1]=p0[z,1];
    Q[1,2]=p1[z,1]; Q[3,2]=p1[z,1];
    Q[1,3]=p2[z,1]; Q[2,3]=p2[z,1];

    ARL1[z,1]=w*(inv(I-Q))*one;
    SDRL1[z,1]=sqrt((w*(I+Q)*(inv(I-Q)**2)*one)-ARL1[z,1]##2);
/*Cdf of run-length variable T : P(T=t)=1-Q^t*1 Fu and Lou (2003)*/
     /*calculate distribution of run length*/
   cdf1=J(2,5000,0); /*first row denotes the run length variable t, second
row the cdf value*/
    /*1st 5th, 25th, 50th, 75th, 95th, 99th percentile*/

do c=1 to 5000;
 cdf1[1,c]=c;
 cdf1[2,c]=1-w*(Q**c)*one;
  if c>1 then do;
     if cdf1[2,c-1]<0.01 & cdf1[2,c]>=0.01 then perc1[z,1]=c;
     else if cdf1[2,c-1]<0.05 & cdf1[2,c]>=0.05 then perc1[z,2]=c;
     else if cdf1[2,c-1]<0.25 & cdf1[2,c]>=0.25 then perc1[z,3]=c;
       else if cdf1[2,c-1]<0.5 & cdf1[2,c]>=0.5 then perc1[z,4]=c;
       else if cdf1[2,c-1]<0.75 & cdf1[2,c]>=0.75 then perc1[z,5]=c;
     else if cdf1[2,c-1]<0.95 & cdf1[2,c]>=0.95 then perc1[z,6]=c;
     else if cdf1[2,c-1]<0.99 & cdf1[2,c]>=0.99 then perc1[z,7]=c;
end;
end;

  end;

print ARL1;
print SDRL1;
print perc1;
quit;
run;
```

Nonparametric control chart supplemented with the *2-of-2* runs-type signalling rule (Human et al. (2010)) – Simulation approach

```
proc iml;
reset nolog;
/*Assumption is that observations are distributed N(0,1), for shift
N(shift,1)*/
/*signal is produced if  two consecutive charting statistics plot
above/below UCL/LCL*/
n=5;
stdev=1;
/*Upper and lower control limits*/
UCL=4;
LCL=0;
IC_med=0;
print UCL, LCL;
x=J(n,1,.); /*define sample vector*/
eta=2;
mean_var=eta*stdev/sqrt(n); /*OOC mean*/
 /*rl= runlength variable*/
simnum=100000;
rl=J(simnum,1,0);

do i=1 to simnum;

chartstat=J(1,500000,0);
signal=0;
count=0;

do until (signal=1);
count=count+1;

call randgen(x,'normal', mean_var,1);

count2=0;
/*calc charting statistic*/
do k=1 to 5;
 if x[k,1]>0 then count2=count2+1;/*charting statistic denotes the number
of values in the sample greater than 0*/
end;

chartstat[1,count]=count2;

/*2-of-2 rule, check if above or below UCL/LCL*/
if count>1  & signal=0 then do;
 if ((chartstat[1,count-1]>=UCL) & (chartstat[1,count]>=UCL))  |
((chartstat[1,count-1]<=LCL) & (chartstat[1,count]<=LCL)) then signal=1;
end;


if count=500000 then signal=1; /*stops infinite loop*/

end;

rl[i,1]=count;

end;

create rldata from rl[colname={RL}];
append from rl;
```

```
proc univariate data=rldata noprint;
histogram;
inset mean std p1 p5 q1 median q3 p95 p99 /format=10.2;
run;

quit;
run;
```

Nonparametric control chart supplemented with the *2-of-2* runs-type signalling rule (Human et al. (2010)) – Exact approach

```
proc iml;
reset nolog;
/*IC*/
n=5; b=0; a=0;
binomp0=1-cdf('normal',0,0,1);
p1i=1-cdf('binom',n-b-1, binomp0,n);
p2i=cdf('binom',a,binomp0 ,n);
ARL0=1/((p1i##2)/(p1i+1)+(p2i##2)/(1+p2i));
print ARL0;

/*OOC*/
shift=J(18,1,0);

do r=1 to 15;
 shift[r,1]=0.2*r;
end;
shift[16,1]=4; shift[17,1]=5; shift[18,1]=6;

/*calculate probability p for the charting statistic that follows a
binomial (n,p) distribution*/
binomp=J(18,1,0);
p1=J(18,1,0);
p2=J(18,1,0);
p0=J(18,1,0);
perc1=J(18,7,0);
ARL1=J(18,1,0);


   do z=1 to 18;
    binomp[z,1]=1-cdf('normal',0,shift[z,1],1);
    p1[z,1]=1-cdf('binom',n-b-1, binomp[z,1],n);
    p2[z,1]=cdf('binom',a, binomp[z,1],n);
    p0[z,1]=1-p1[z,1]-p2[z,1];
    ARL1[z,1]=1/((p1[z,1]##2)/(p1[z,1]+1)+(p2[z,1]##2)/(1+p2[z,1]));

end;

print p1, p2, p0;
print ARL1;

quit;
run;
```

Run rules $\overline{X}$ charts when process parameters are unkown (Zhang et al. (2010) – Markov chain approach (*2-of-2* runs-type signalling rule).

```
/*apply Markov chain approach to the 2-of-2 control chart (new) - Zhang
(2010) proposed the 2-of-3 and 3-of-4 for unknown parameters*/
/*Assumption is that observations distributed IID N(mu,sigma), where mu and
sigma are unknown*/
proc iml;
reset nolog;
/*2-of-2 runs-type signalling rule*/
/*phase I and II analysis*/

/*phase I analysis -estimating mu and sigma, observations distributed N(mu,
sigma)*/
n=5;
m=10;
x=J(n,1,.);
sum=0;
sumt=0;
meanvt=J(n,1,0);
dif=J(n,1,0);
difsqsum=0;
dif=J(n,1,0);
mean=J(m,1,0);
difsum=0;

do a=1 to m;
call randgen(x,'normal', 0,1);
print x;
sum=x[+,];
sumt=sumt+sum;
mean[a,1]=sum/n;

 do b=1 to n;
   difsum=difsum+(x[b,1]-mean[a,1])##2;
 end;
print sum;
print sumt;
print difsum;

end;

Mu_est=sumt/(n*m);
print Mu_est;

print mean;
stdev=sqrt(difsum/(m*n-1));
print stdev;

print Mu_est, stdev;
/*Define UCL and LCL for the chart*/
L=0.7966666;

UCL=Mu_est+L#stdev;
LCL=Mu_est-L#stdev;

/*phase II analysis*/
/*calculate transitional probabilities*/
Mu_std=0; sdevst=1;
```

```
temp1=(Mu_est-0)*sqrt(n)/1-L*stdev*sqrt(n)/1;
temp2=-(Mu_est-0)*sqrt(n)/1-L*stdev*sqrt(n)/1;

p2=cdf('normal',temp1,0,1);
p1=cdf('normal',temp2,0,1);
p0=1-p1-p2;

print p0, p1, p2;

/*set up essential TPM*/
/*ARL=w*(inv(I-Q))*1 Fu and Lou (2003)*/
w=J(1,3,0); w[1,1]=1;
I=I(3); /*identity matrix Fu and Lou (2003)*/
Q=J(3,3,0);

/*setting up essential TPM for IC process*/
Q[1,1]=p0; Q[2,1]=p0; Q[3,1]=p0;
Q[1,2]=p1; Q[3,2]=p1;
Q[1,3]=p2; Q[2,3]=p2;

one=J(3,1,1);
print w, I, Q, one;
ARL0=w*(inv(I-Q))*one;
print ARL0;

/*quit;
run;*/

SDRL=sqrt((w*(I+Q)*(inv(I-Q)**2)*one)-ARL0**2);
print SDRL;

/*Cdf of run-length variable T : P(T=t)=1-Q^t*1 Fu and Lou (2003)*/
/*75th percentile P(T=t)=0.75 etc. */
/*calculate distribution of run length*/
cdf=J(2,10000,0); /*first row denotes the run length variable t, second row
the cdf value*/
perc=J(1,7,0); /*25th, 50th, 75th percentile*/

do b=1 to 10000;
 cdf[1,b]=b;
 cdf[2,b]=1-w*(Q**b)*one;
  if b>1 then do;
     if cdf[2,b-1]<0.01 & cdf[2,b]>=0.01 then perc[1,1]=b;
     else if cdf[2,b-1]<0.05 & cdf[2,b]>=0.05 then perc[1,2]=b;
     else if cdf[2,b-1]<0.25 & cdf[2,b]>=0.25 then perc[1,3]=b;
       else if cdf[2,b-1]<0.5 & cdf[2,b]>=0.5 then perc[1,4]=b;
       else if cdf[2,b-1]<0.75 & cdf[2,b]>=0.75 then perc[1,5]=b;
     else if cdf[2,b-1]<0.95 & cdf[2,b]>=0.95 then perc[1,6]=b;
     else if cdf[2,b-1]<0.99 & cdf[2,b]>=0.99 then perc[1,7]=b;
end;
end;

print '5th, 25th, 50th, 75th and 95th percentile of run-length for IC
process';
print perc;

/*Process OOC*/
/*calculate transitional probabilities*/
shift=J(30,1,0);
/*shift vector - various shifts used - in multiples of 0.2*/
do r=1 to 30;
```

```
 shift[r,1]=0.2*r;
end;

ARL1=J(30,1,0);
SDRL1=J(30,1,0);

/*Cdf of run-length variable T : P(T=t)=1-Q^t*1 Fu and Lou (2003)*/
/*75th percentile P(T=t)=0.75 etc. */
/*calculate distribution of run length*/
/*For each shift in the process the percentiles have to be calculated*/
cdf1=J(31,10000,0); /*first row denotes the run length variable t, second-
21 row denotes the cdf value of a process shift*/
perc1=J(30,7,0); /*25th, 50th, 75th percentiles for each process shift -
various rows denotes the shift*/

/*p1b is probability plotting above the UCL if the process is OOC*/
/*p2b is the probability plotting below the LCL if the process is OOC*/
/*p0b is the probability plotting between the UCL and LCL if the process is
OOC*/
do j=1 to 30;
temp1b=(Mu_est-0)*sqrt(n)/1-L*stdev*sqrt(n)/1-shift[j,1]*sqrt(n);
temp2b=-(Mu_est-0)*sqrt(n)/1-L*stdev*sqrt(n)/1+shift[j,1]*sqrt(n);

p2b=cdf('normal',temp1b,0,1);
p1b=cdf('normal',temp2b,0,1);
p0b=1-p1b-p2b;

/*print p0b, p1b, p2b;*/

/*set up essential TPM*/
/*ARL=w*(inv(I-Q))*1 Fu and Lou (2003)*/
/*w, I, and one defined as for IC state*/

/*setting up essential TPM for OOC process*/
Q[1,1]=p0b; Q[2,1]=p0b; Q[3,1]=p0b;
Q[1,2]=p1b; Q[3,2]=p1b;
Q[1,3]=p2b; Q[2,3]=p2b;

/*print w, I, Q, one;*/
ARL1[j,1]=w*(inv(I-Q))*one;
SDRL1[j,1]=sqrt((w*(I+Q)*(inv(I-Q)**2)*one)-ARL1[j,1]##2);

do a=1 to 10000;
 cdf1[1,a]=a;
 cdf1[j+1,a]=1-w*(Q**a)*one;
  if a>1 then do;
     if cdf1[j+1,a-1]<0.01 & cdf1[j+1,a]>=0.01 then perc1[j,1]=a;
     else if cdf1[j+1,a-1]<0.05 & cdf1[j+1,a]>=0.05 then perc1[j,2]=a;
     else if cdf1[j+1,a-1]<0.25 & cdf1[j+1,a]>=0.25 then perc1[j,3]=a;
       else if cdf1[j+1,a-1]<0.5 & cdf1[j+1,a]>=0.5 then perc1[j,4]=a;
       else if cdf1[j+1,a-1]<0.75 & cdf1[j+1,a]>=0.75 then perc1[j,5]=a;
     else if cdf1[j+1,a-1]<0.95 & cdf1[j+1,a]>=0.95 then perc1[j,6]=a;
       else if cdf1[j+1,a-1]<0.99 & cdf1[j+1,a]>=0.99 then perc1[j,7]=a;
 end;
end;


end;
print ARL1;
print SDRL1;
```

```
print '5th, 25th, 50th, 75th and 95th percentile of run-length for OOC
process';
print perc1;

quit;
run;
```

Run rules $\bar{X}$ charts when process parameters are unkown (Zhang et al. (2010) – Markov chain approach (*2-of-3* runs-type signalling rule).

```
/*apply Markov chain approach to the 2-of-3 control chart (Zhang 2010)*/
/*Assumption is that observations distributed IID N(mu,sigma), mu and sigma
unknown*/
proc iml;
reset nolog;
/*phase I and II analysis*/
/*2-of-3 runs-type signalling rule*/

/*phase I analysis -estimating mu and sigma, observations distributed N(mu,
sigma)*/
n=5;
m=10;
x=J(n,1,.);
sum=0;
sumt=0;
meanvt=J(n,1,0);
dif=J(n,1,0);
difsqsum=0;
dif=J(n,1,0);
mean=J(m,1,0);
difsum=0;

do a=1 to m;
call randgen(x,'normal', 0,1);
print x;
sum=x[+,];
sumt=sumt+sum;
mean[a,1]=sum/n;

 do b=1 to n;
   difsum=difsum+(x[b,1]-mean[a,1])##2;
 end;
print sum;
print sumt;
print difsum;

end;

Mu_est=sumt/(n*m);
print Mu_est;

print mean;
stdev=sqrt(difsum/(m*n-1));
print stdev;

print Mu_est, stdev;
/*Define UCL and LCL for the chart*/
L=0.8628;
```

```
UCL=Mu_est+L#stdev;
LCL=Mu_est-L#stdev;

/*p1 probability plotting above UCL*/
/*p2 probability plotting below LCL*/
/*p0 probability plotting between UCL and LCL*/

/*Process IC*/
/*calculate transitional probabilities*/
Mu_std=0; sdevst=1;
temp1=(Mu_est-0)*sqrt(n)/1-L*stdev*sqrt(n)/1;
temp2=-(Mu_est-0)*sqrt(n)/1-L*stdev*sqrt(n)/1;

p2=cdf('normal',temp1,0,1);
p1=cdf('normal',temp2,0,1);
p0=1-p1-p2;

print p0, p1, p2;


/*set up essential TPM*/
/*ARL=w*(inv(I-Q))*1 Fu and Lou (2003)*/
w=J(1,7,0); w[1,1]=1;
I=I(7); /*identity matrix Fu and Lou (2003)*/
Q=J(7,7,0);
/*setting up essential TPM for IC process*/
Q[1,1]=p0; Q[2,4]=p0; Q[3,6]=p0; Q[4,1]=p0; Q[5,6]=p0; Q[6,1]=p0;
Q[7,4]=p0;
Q[1,2]=p1; Q[5,7]=p1; Q[6,2]=p1;
Q[1,3]=p2; Q[2,5]=p2; Q[3,7]=p2; Q[4,3]=p2;

one=J(7,1,1);
print w, I, Q, one;
ARL0=w*(inv(I-Q))*one;
print ARL0;

inverse=J(5,5,0);
inverse=inv(I-Q);
print inverse;

SDRL=sqrt((w*(I+Q)*(inv(I-Q)**2)*one)-ARL0**2);
print SDRL;


/*Cdf of run-length variable T : P(T=t)=1-Q^t*1 Fu and Lou (2003)*/
/*75th percentile P(T=t)=0.75 etc. */
/*calculate distribution of run length*/
cdf=J(2,10000,0); /*first row denotes the run length variable t, second row
the cdf value*/
perc=J(1,7,0); /*25th, 50th, 75th percentile*/

do b=1 to 10000;
 cdf[1,b]=b;
 cdf[2,b]=1-w*(Q**b)*one;
  if b>1 then do;
     if cdf[2,b-1]<0.01 & cdf[2,b]>=0.01 then perc[1,1]=b;
     else if cdf[2,b-1]<0.05 & cdf[2,b]>=0.05 then perc[1,2]=b;
     else if cdf[2,b-1]<0.25 & cdf[2,b]>=0.25 then perc[1,3]=b;
       else if cdf[2,b-1]<0.5 & cdf[2,b]>=0.5 then perc[1,4]=b;
       else if cdf[2,b-1]<0.75 & cdf[2,b]>=0.75 then perc[1,5]=b;
     else if cdf[2,b-1]<0.95 & cdf[2,b]>=0.95 then perc[1,6]=b;
```

```
     else if cdf[2,b-1]<0.99 & cdf[2,b]>=0.99 then perc[1,7]=b;
end;
end;


print '5th, 25th, 50th, 75th and 95th percentile of run-length for IC
process';
print perc;

/*Process OOC*/
/*calculate transitional probabilities*/
/*test for shift of 0.2*/
/*shf=0.2;*//*shift in mean=0.2 units*/
shift=J(30,1,0);
/*shift vector - various shifts used - in multiples of 0.2*/
do r=1 to 30;
 shift[r,1]=0.2*r;
end;

ARL1=J(30,1,0);
SDRL1=J(30,1,0);

/*Cdf of run-length variable T : P(T=t)=1-Q^t*1 Fu and Lou (2003)*/
/*75th percentile P(T=t)=0.75 etc. */
/*calculate distribution of run length*/
/*For each shift in the process the percentiles have to be calculated*/
cdf1=J(31,10000,0); /*first row denotes the run length variable t, second-
21 row denotes the cdf value of a process shift*/
perc1=J(30,7,0); /*25th, 50th, 75th percentiles for each process shift -
various rows denotes the shift*/

do j=1 to 30;

temp1b=(Mu_est-0)*sqrt(n)/1-L*stdev*sqrt(n)/1-shift[j,1]*sqrt(n);
temp2b=-(Mu_est-0)*sqrt(n)/1-L*stdev*sqrt(n)/1+shift[j,1]*sqrt(n);

p2b=cdf('normal',temp1b,0,1);
p1b=cdf('normal',temp2b,0,1);
p0b=1-p1b-p2b;

/*set up essential TPM*/
/*ARL=w*(inv(I-Q))*1 Fu and Lou (2003)*/
/*w, I, and one defined as for IC state*/

/*setting up essential TPM for OOC process*/
Q[1,1]=p0b; Q[2,4]=p0b; Q[3,6]=p0b; Q[4,1]=p0b; Q[5,6]=p0b; Q[6,1]=p0b;
Q[7,4]=p0b;
Q[1,2]=p1b; Q[5,7]=p1b; Q[6,2]=p1b;
Q[1,3]=p2b; Q[2,5]=p2b; Q[3,7]=p2b; Q[4,3]=p2b;

/*print w, I, Q, one;*/
ARL1[j,1]=w*(inv(I-Q))*one;
SDRL1[j,1]=sqrt((w*(I+Q)*(inv(I-Q)**2)*one)-ARL1[j,1]##2);

do a=1 to 10000;
 cdf1[1,a]=a;
 cdf1[j+1,a]=1-w*(Q**a)*one;
  if a>1 then do;
     if cdf1[j+1,a-1]<0.01 & cdf1[j+1,a]>=0.01 then perc1[j,1]=a;
     else if cdf1[j+1,a-1]<0.05 & cdf1[j+1,a]>=0.05 then perc1[j,2]=a;
     else if cdf1[j+1,a-1]<0.25 & cdf1[j+1,a]>=0.25 then perc1[j,3]=a;
```

```
        else if cdf1[j+1,a-1]<0.5 & cdf1[j+1,a]>=0.5 then perc1[j,4]=a;
        else if cdf1[j+1,a-1]<0.75 & cdf1[j+1,a]>=0.75 then perc1[j,5]=a;
      else if cdf1[j+1,a-1]<0.95 & cdf1[j+1,a]>=0.95 then perc1[j,6]=a;
        else if cdf1[j+1,a-1]<0.99 & cdf1[j+1,a]>=0.99 then perc1[j,7]=a;
 end;
end;



end;
print ARL1;

print SDRL1;

*print cdf1;
print '5th, 25th, 50th, 75th and 95th percentile of run-length for OOC
process';
print perc1;

quit;
run;
```

Run rules $\bar{X}$ charts when process parameters are unkown (Zhang et al. (2010) – Simulation approach (*2-of-3* runs-type signalling rule).

```
/*Simulation approach 2-of-3 runs-type signalling rule with estimated
parameters */
proc iml;
reset nolog;
/*signal is produced if two consecutive charting statistics of three plot
above/below UCL/LCL */

/*phase I*/
n=5;
m=10;
x=J(n,1,.);
sum=0;
sumt=0;
meanvt=J(n,1,0);
dif=J(n,1,0);
difsqsum=0;
dif=J(n,1,0);
mean=J(m,1,0);
difsum=0;

do a=1 to m;
call randgen(x,'normal', 0,1);
print x;
sum=x[+,];
sumt=sumt+sum;
mean[a,1]=sum/n;

 do b=1 to n;
   difsum=difsum+(x[b,1]-mean[a,1])##2;
 end;
print sum;
print sumt;
```

```
print difsum;

end;

Mu_est=sumt/(n*m);
print Mu_est;

print mean;
stdev=sqrt(difsum/(m*n-1));
print stdev;

/*Inner upper and lower control limits*/
UCL= Mu_est +(1.7814*stdev)/sqrt(n);
LCL= Mu_est -(1.7814*stdev)/sqrt(n);

/*UCL=0.841519235;
LCL=-0.866931235; */

/*stdev=0.9900617; /*estimated*/
Mu_est=-0.012706; */

print UCL, LCL;

/*phase II*/

x=J(n,1,.); /*define sample vector*/
/*adjust eta according to shift from 0 to 6, see article*/
eta=0;
mean_var=Mu_est+eta*stdev/sqrt(n); /*OOC mean*/

/*rl= runlength variable*/
simnum=10000;
rl=J(simnum,1,0);

do i=1 to simnum;

chartstat=J(1,10000,0);
signal=0;
count=0;

do until (signal=1);
count=count+1;

call randgen(x,'normal', mean_var,stdev);
chartstat[1,count]=mean(x);  /*charting statistic is the mean of the
sample*/

/*2-of-3 rule, check if above or below UCL/LCL*/
/*no signal yet*/
if count>2  & signal=0 then do;
/*signal could be 101 or 011 for 2-of-3 rule*/
/*for 101 signal*/
if ( (chartstat[1,count-2]>=UCL)  & ((chartstat[1,count-1]>LCL) &
(chartstat[1,count-1]<UCL)) & (chartstat[1,count]>=UCL)  )
| ( (chartstat[1,count-2]<=LCL) &  ((chartstat[1,count-1]>LCL) &
(chartstat[1,count-1]<UCL)) & (chartstat[1,count]<=LCL))
then signal=1;

/*for 011 signal*/
if ( ((chartstat[1,count-2]>LCL) &(chartstat[1,count-2]<UCL)) &
(chartstat[1,count-1]>=UCL)  & (chartstat[1,count]>=UCL) )
```

```
| ( ((chartstat[1,count-2]>LCL) &(chartstat[1,count-2]<UCL))&
(chartstat[1,count-1]<=LCL)  & (chartstat[1,count]<=LCL)  )
then signal=1;

if ( (chartstat[1,count-2]>=UCL)  &(chartstat[1,count-1]>=UCL) )
| ( (chartstat[1,count-2]<=LCL) & (chartstat[1,count-1]<=LCL)  )
then signal=1;

end;

if count=10000 then signal=1; /*stops infinite loop*/

end; /*end of do until*/

*print chartstat;
*print count;

rl[i,1]=count;

end;

create rldata from rl[colname={RL}];
append from rl;

proc univariate data=rldata noprint;
histogram;
inset mean std p1 p5 q1 median q3 p95 p99 min max/format=10.2;
run;

quit;
run;
```

Run rules $\bar{X}$ charts when process parameters are unkown (Zhang et al. (2010) – Simulation approach (*2-of-2* runs-type signalling rule).

```
/*Simulation approach 2-of-2 rule with estimated parameters*/
proc iml;
reset nolog;
/*signal is produced if two consecutive charting statistics of three plot
above/below UCL/LCL */

/*phase I*/
n=5;
m=10;
x=J(n,1,.);
sum=0;
sumt=0;
meanvt=J(n,1,0);
dif=J(n,1,0);
difsqsum=0;
dif=J(n,1,0);
mean=J(m,1,0);
difsum=0;

do a=1 to m;
call randgen(x,'normal', 0,1);
print x;
sum=x[+,];
```

```
sumt=sumt+sum;
mean[a,1]=sum/n;

 do b=1 to n;
   difsum=difsum+(x[b,1]-mean[a,1])##2;
 end;
print sum;
print sumt;
print difsum;

end;

Mu_est=sumt/(n*m);
print Mu_est;

print mean;
stdev=sqrt(difsum/(m*n-1));
print stdev;
print Mu_est, stdev;

L=0.7966666;
/*Inner upper and lower control limits*/
UCL= Mu_est +(L *stdev)/sqrt(n); /*UCL_hat=mu_est+L*stdev_est, UCL here
refers to UCL_hat a random variable*/
LCL= Mu_est -(L *stdev)/sqrt(n); /*LCL_hat=mu_est-L*stdev_est, LCL refers
to LCL_hat also a random variable*/
stdev=1;
/*UCL=0.750627022;
LCL=-0.720062222; */
print UCL, LCL;

/*phase II*/
x=J(n,1,.); /*define sample vector*/
/*adjust eta according to shift from 0 to 6, see article*/
eta=6;
mean_var= Mu_est+eta*stdev/sqrt(n); /*OOC mean*/

/*rl= runlength variable*/
simnum=10000;
rl=J(simnum,1,0);

do i=1 to simnum;

chartstat=J(1,10000,0);
signal=0;
count=0;

do until (signal=1);
count=count+1;

call randgen(x,'normal', mean_var,1);
chartstat[1,count]=mean(x);  /*charting statistic is the mean of the
sample*/
/*2-of-2 rule, check if above or below UCL/LCL*/
/*no signal yet*/
if count>1  & signal=0 then do;
 if ((chartstat[1,count-1]>=UCL) & (chartstat[1,count]>=UCL))  |
((chartstat[1,count-1]<=LCL) & (chartstat[1,count]<=LCL)) then signal=1;
end;

if count=10000 then signal=1; /*stops infinite loop*/
```

```
end; /*end of do until*/

rl[i,1]=count;

end;

create rldata from rl[colname={RL}];
append from rl;

proc univariate data=rldata noprint;
histogram;
inset mean std p1 p5 q1 median q3 p95 p99 min max/format=10.2;
run;


quit;
run;
```

References

Abbas, N., Riaz, M., Does, R. (2011). "Enhancing the performance of EWMA Charts". *Quality and Reliability Engineering International*, **27**, 821-833.

Acosta-Mejia, C.A. (2007). "Two sets of Runs Rules for the Chart". *Quality Engineering*, **19 (2),** 129-136.

Acosta-Mejia, C.A., Pignatiello Jr., J.J. (2008). "Modified R Charts for Improved Performance". *Quality Engineering*, **20 (3),** 361-369.

Acosta-Mejia, C.A., Pignatiello Jr., J.J. (2009). "*ARL*-Design of S Charts with *k-of-k* Runs Rules". *Communications in Statistics - Simulation and Computation*, **38 (8),** 1625-1639.

Adams, B.M., Lowry, C., Woodall, W.H. (1992). "The use (and misuse) of false alarm probabilities in control chart design". *Frontiers in Statistical Quality Control*, **4**, Heidelberg, Physica-Verlag, 155-168.

Aerne, L.A., Champ, C.W., Rigdon, S.E. (1991). "Evaluation of control charts under linear trend". *Communications in Statistics*: *Theory and Methods*, **20**, 3341–3349.

Aki., S. (1992). "Waiting time problems for a sequence of discrete random variables". *Annals of the Institute of Statistical Mathematics*, **44**, 363–378.

Alloway, J.A., Raghavachari, M. (1991). "Control chart based on Hodges-Lehman Estimator". *Journal of Quality Technology*, **23**, 336-347.

Alt, F.B. (1985). "Multivariate quality control". *Encyclopedia of Statistical Sciences* (John Wiley and Sons), **6**, 110-122.

Alt, F.B., Smith, N.D. (1988). "Multivariate process control". Handbook of Statistics: Quality Control and Reliability, Amsterdam, North-Holland, **7**.

Amin, R.W., Letsinger, W.C. (1991). "Improved switching rules in control procedures using variable sampling intervals". *Communications in Statistics: Simulation and Computation*, **20**, 205-230.

Amin, R.W., Reynolds Jr., M.R., Bakir, S.T. (1995). "Nonparametric quality control charts based on the sign statistic". *Communications in Statistics: Theory and Methods*, **24 (6),** 1597-1623.

Ansari, A.R., Bradley, R.A. (1960). "Rank Sum Test for Dispersion''. *Annals of Mathematical Statistics*, **31**, 1174-1189.

Antzoulakos, D.L. (2001). "Waiting times for patterns in a sequence of multistate trials". *Journal of Applied Probability*, **38**, 508–518.

Antzoulakos, D.L., Rakitzis, A.C. (2008a). "The Revised *m-of-k* Runs Rule". *Quality Engineering* , **20 (1)**, 75-81.

Antzoulakos, D.L., Rakitzis, A.C. (2008b). "The modified r out of m control chart". *Communications in Statistics: Simulation and Computation*, **37 (2)**, 396-408.

Antzoulakos, D.L., Rakitzis, A.C. (2010). "Runs rules schemes for monitoring process variability". *Journal of Applied Statistics*, **37 (7)**, 1231-1247.

Antzoulakos, D.L., Rakitzis, A.C. (2014). "Control charts with switching and sensitizing runs rules for monitoring process variation". *Journal of Statistical Computation and Simulation*, **84 (1)**, 37-56.

Aparisi, F., Champ, C.W., Garcia Diaz, J.C. (2004). "A performance Analysis of Hotelling's T^2 control chart with supplementary run rules". *Quality Engineering*, **16**, 359-368.

Aparisi, F., de Luna, M.A. (2009). "The design and performance of the multivariate synthetic T^2 control chart". *Communications in Statistics, Theory and Methods*, **38 (2)**, 173–192.

Bai, D.S., Lee, K.T. (2002). "Variable sampling interval $\overline{X}$ control charts with an improved switching rule". *International Journal of Production Economics*, **76**, 189–199.

Bakir, S.T. (2004). "A distribution-free Shewhart quality control chart based on signed-ranks". *Quality Engineering*, **16 (4)**, 613-623.

Balakrishnan, N., Koutras, M.V. (2002). "Runs and Scans with Applications". Wiley, New York.

Balakrishnan, N., Triantafyllou, I.S., Koutras, M.V. (2010). "A Distribution-Free Control Chart Based on Order Statistics". *Communications in Statistics, Theory and Methods*, **39 (20)**, 3652-3677.

Barnard, J. (1978). "Probability integral of the normal range". *Applied Statistics*, **28**, 197-198.

Bissell, A.F. (1969). "CUSUM techniques for quality control". *Journal of the Royal Statistical Society*, Series C, **18 (1)**, 1-30.

Bradly, J.V. (1968). "Distribution-Free Statistical Tests". Prentice-Hall, New Jersey.

Brook, D., Evans, D. A. (1972). "An approach to the probability distribution of CUSUM run length". *Biometrika*, **59 (3)**, 539–549.

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

Chakraborti, S., van der Laan, P., van de Wiel, M.A. (2004). "A class of distribution-free control charts". *Journal of Royal Statistical Society*, **53 (3)**, 443–462.

Chakraborti, S., Eryilmaz, S. (2007). "A Nonparametric Shewhart-type signed-rank control chart based on runs". *Communications in Statistics: Simulation and Computation*, **36 (2)**, 335-356.

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

Chakraborti, S., Human, S., Eryilmaz, S. (2009). "A phase II nonparametric control chart based on precedence statistics with runs-type signalling rules". *Computational Statistics and Data analysis*, **53**, 1054-1065.

Champ, C.W. (1986). "Exact results for shewhart control charts with supplementary runs rules". Ph.D. dissertation, Department of Statistics, University of Louisiana, Lafayette, LA.

Champ, C.W. (1992). "Steady-state run length analysis of a Shewhart quality control chart with supplementary runs rules". *Communications in Statistics - Theory and Methods*, **21**, 765–777.

Champ, C.W., Aparisi, F. (2008). "Double sampling Hotelling's T^2 charts". *Quality and Reliability Engineering International*, **24 (2)**, 153-166.

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

Champ, C.W., Woodall, W.H. (1990). "A program to evaluate the run length distribution of a Shewhart control chart with supplementary runs rules". *Journal of Quality Technology*, **22**, 68-73.

Champ, C.W., Woodall, W.H. (1997). "Signal probabilities of runs rules supplementing a Shewhart control chart". *Communications in Statistics. Simulation and Computation*, **26**, 1347–1360.

Chen, Y. K. (2007). "Adaptive sampling enhancement for Hotelling's T^2 charts". *European Journal of Operational Research*, **178**, 841–857.

Davis, R.B., Woodall, W.H. (1990). "Performance of the control chart trend rule under linear shift". *Journal of Quality Technology*, **20**, 260-262.

Davis, R.B. Homer, A., Woodall, W.H. (1990). "Performance of the zone control chart". *Communications in Statistics - Theory and Methods*, **19**, 1581-1587.

Davis, R.B., Chun, J., Guo, Y.Y. (1994). "Improving the performance of the zone control chart." *Communications in Statistics - Theory and Methods*, **23**, 3557-3565.

Das, N. (2008a). "A Note on Efficiency of Nonparametric Control chart for Monitoring Process variability". *Economic Quality Control* **23 (1)**, 85-93.

Das, N. (2008b). "Nonparametric Control Chart for Controlling Variability Based on Rank Test". *Economic Quality Control* **23 (2)**, 227-242.

Derman, C., Gleser, L.J., Olkin, I. (1973). "A Guide to Probability theory and application". Holt Rhinehart and Winston, New York, N.Y.

Derman, C., Ross, S.M. (1997). "Statistical Aspects of Quality Control", San Diego: Academic Press.

Divoky, J.J., Taylor, R.W. (1995). "Detecting process drift with combinations of trend and zonal supplementary runs rules". *International Journal of Quality & Reliability Management*, **12 (2)**, 60-71.

Ferrel, E.B. (1953). "Control Charts using Midranges and Medians". *Industrial Quality Control*, **9**, 30-34.

Fu, J.C., Chang, Y.M. (2003). "On ordered series and later waiting time distributions in a sequence of Markov dependent multistate trials". *Journal of Applied Probability*, **40 (3)**, 623–642.

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

Fu, J.C., Shmueli, G., Chang, Y.M. (2003). "A unified Markov chain approach for computing the run length distribution in control charts with simple or compound rules". *Statistics and Probability Letters*, **65 (4)**, 457-466.

Fu, J.C., Spiring, F.A., 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 EWMA control charts based on median run length". *Journal of Statistical Computation and Simulation*, **45 (3, 4)**, 169-184.

George, E.O., Bowman, D. (1995). "A full likelihood procedure for analysing exchangeable binary data". *Biometrics*, **51**, 512-523.

Gibbons, J.D., Chakraborti, S. (2011). "Nonparametric Statistical Inference". Fourth edition, Marcel Dekker, New York.

Hackl, P., Ledolter, J. (1992). "A control chart based on ranks". *Journal of Quality Technology*, **23 (2)**, 117-124.

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

Human, S.W., Chakraborti, S., Smit, C.F. (2010). "Nonparametric Shewhart-type sign control charts based on runs." *Communications in Statistics: Theory and Methods*, **39 (11)**, 2046 – 2062.

Hurwitz, A.M., Mathur, M.A. (1992). "Very simple set of process control rules". *Quality Engineering*, **1**, 21–29.

Jacobs, D.C. (1990). "Statistical Process Control: Watch for Non-normal Distributions". *Chemical Engineering Progress*, **86**, 19-27.

Jaehn, A.H. (1987). "Zone control charts: A new tool for quality control". *Technical Association Pulp and Paper Industry*, **87**, 159-161.

Jaehn, A.H. (1989). "Zone control charts find new applications". *ASQC Quality Congress Transactions*, 890-895.

Khoo, M.B.C. (2003). "Design of runs rules schemes". *Quality Engineering*, **16 (1)**, 27-43.

Khoo, M.B.C., Quah, S.H. (2003), "Incorporating runs rules into Hotelling's χ2 control charts". *Quality Engineering*, **15,** 671–675.

Khoo, M.B.C., Quah, S.H. (2004), "Alternatives to the Multivariate Control Chart for Process Dispersion". *Quality Engineering*, **16 (3)**, 423-435.

Khoo, M.B.C., Quah, S.H., Low, H.C. (2005). "Powerful Rules for the Hotelling's χ 2 Control Chart". *Quality Engineering*, **17 (1)**, 139-149.

Khoo, M.B.C., Ariffin, N.K. (2006), "Two Improved Runs Rules for the Shewhart $\bar{X}$ Control Chart". *Quality Engineering*, **18**, 173-178.

Klein, M. (1997). "Modified Shewhart-EWMA control charts". *IIE Transactions*, **29**, 1051–1056.

Klein, M. (2000a). "Two alternatives to the Shewhart $\overline{X}$ control chart". *Journal of Quality Technology*, **32 (4)**, 427-431.

Klein, M. (2000b). "Modified $S-$charts for controlling process variability". *Communications in Statistics -Simulation and Computation*, **29(3),** 919-940.

Koutras, M.V., Bersimis, S., Maravelakis, P.E. (2006). "Improving the Performance of the Chi-square Control Chart via Runs Rules". *Methodology and Computing in Applied Probability*, **8**, 409-426.

Koutras, M.V., Bersimis, S., Maravelakis, P.E. (2007). "Statistical process control using Shewhart control charts with supplementary runs rules". *Methodology and Computing in Applied Probability*, **9**, 207-224.

Lee, M.H. (2013). "Adaptive Hotelling's T^2 Control Charts with Run Rules". *Communications in Statistics*, **42**, 883-897.

Lagenberg, P., Iglewicz, B. (1986). "Trimmed Mean and R charts". *Journal of Quality Technology*, **18**, 152-161.

Lehman, E.L. (1975). "Nonparametric statistical Methods Based on Ranks". Holden-Day. San Fransisco, California.

Low, C.K., Khoo, M.B.C., Teoh, W.L., Wu, Z. (2012). "The Revised m-of-k runs rule based on the Median Run Length". *Communications in Statistics – Simulation and Computation*, **41 (8)**, 1463–1477.

Lowry, C.A., Champ, C.W., Woodall, W.H. (1995). "The performance of control charts for monitoring process variation". *Communications in Statistics – Simulation and Computation*, **24 (2)**, 409–437.

Lowry, C.A., Montgomery, D.C. (1995). "A review of multivariate control charts". *IIE Transactions*, **27**, 800-810.

Montgomery, D.C. (2013). "*Statistical Quality Control: A Modern Introduction*, 7th ed.". John Wiley & Sons, Singapore Pty. Ltd.

Mood, A.M. (1954). "On the asymptotic efficiency of certain non-parametric two sample test". *The Annal of Mathematical Statistics*, **25**, 514-22.

Moore, P.G. (1958). "Some properties of runs in quality control procedures". *Biometrika*, **45**, 89-95.

Mosteller, F. (1941). "Note on application of runs to quality control charts". *Annals of Mathematical Statistics*, **12**, 228–231.

Nelson, L.S. (1985). "Interpreting Shewharts X-bar chart". *Journal of Quality Technology*, **17**, 114–116.

Nelson, L.S. (1984). "The Shewhart control chart—test for special causes". *Journal of Quality Technology*, **16**, 237–239.

Nelson, L.S (1990). "Monitoring reduction in variation with a range chart". *Journal of Quality Technology*, **22**, 163–165.

Page, E.S. (1954). "Continuous inspection schemes". *Biometrika* ,**41**, 100–115.

Page, E.S. (1955). "Control charts with warning lines". *Biometrika,* **42** (12), 243-257.

Palm, A. (1990). "Tables of run length percentiles for determining the sensitivity of Shewhart control charts for averages with supplementary runs rules". *Journal of Quality Technology*, **22 (4)**, 289–298.

Philippou, A.N., Georghiou, C., Philippou, G.N. (1983). "A generalized geometric distribution and some of its properties". *Statistics & Probability*, **1**, 171–175.

Rakitzis, A.C., Antzoulakos, D.L. (2011). "On the improvement of one-sided S control charts". *Journal of Applied Statistics*, **38**, 2839-2858.

Rakitzis, A.C., Antzoulakos, D.L. (2014). "Control charts with switching and sensitizing runs rules for monitoring process variation". *Journal of Statistical Computation and Simulation*, **84 (1)**, 37-56.

Riaz, M., Abbas, N., Does, R. (2011). "Improving the performance of Cusum Charts". *Quality and Reliability Engineering International*, **27**, 415-424.

Roberts, S.W. (1958**).** "Properties of control chart zone tests". *The Bell System Technical Journal*, **37,** 83-114.

Roberts, S.W. (1959). "Control charts based on geometric moving averages". *Technometrics*, **1**, 239–250.

Shewhart, W.A. (1939). "Statistical Methods from the Viewpoint Of Quality Control". Republished in 1986, Dover Publication, New York.

Shmueli, G., Cohen, A. (2003). "Run length distribution for control charts with runs and scans rules." *Communications in Statistics: Theory and Methods***, 32 (2),** 475-495.

Siegel, S., Tukey, J.W. (1960). "Nonparametric Sum of ranks procedure for relative spread in unpaired samples". *Journal of the American Statistical Association* **55**, 429-445.

Tracy, N.D., Young, J.C., Mason, R.L. (1992). "Multivariate control charts for individual observations". *Journal of Quality Technology*, **24**, 88-95.

Tukey, J.W. (1960). "A survey of sampling from contaminated Distributions". *Contributions of Probability and Statistics*. Essay in Honour of Harold Hotelling, Stanford University Press, Stanford CA.

Walker, E., Philpot, J.W., Clement, J. (1991). "False signal rates for the Shewhart control chart with supplementary runs tests". *Journal of Quality Technology*, **23**, 247–252.

Weiler, H. (1953). "The use of runs to control the mean in quality control". *Annals of Mathematical Statistics*, **48**, 816–825.

Western Electric (1956). "Statistical Quality Control Handbook". Western Electric Corporation, Indianapolis, IN.

Westgard, J.A., Groth, T. (1979). "Power functions for statistical control rules". *Clinical Chemistry*, **25**, 863–869.

Westgard, J.A. Barr, P.L., Hunt, M.R., Groth, T. (1982). "A multi-run Shewhart chart for quality control in clinical chemistry". *Clinical Chemistry*, **27**, 493–501.

Wetherill, G.B., Brown, D.W. (1991). "Statistical Process Control". Chapman and Hall, New York.

Wheeler, D.J. (1983). "Detecting a shift in process average: Tables of the power function for X charts". *Journal of Quality Technology*, **15**, 155-170.

Woodall, W.H., Montgomery, D.C. (1999). "Research Issues and Ideas in Statistical Process Control". *Journal of Quality Technology*, **31**, 376-386.

Woodall, W.H. (2000). "Controversies and contradictions in statistical process control". *Journal of Quality Technology*, **32**, 341–350.

Yourstone, S.A., Zimmer, W.J. (1992). "Non-normality and the design of Control Charts for averages". *Decision Sciences*, **23**, 1099-1113.

Zhang, Y., Castagliola, P. (2010). "Run rules $\overline{X}$ charts when process parameters are unknown". *International Journal of Reliability, Quality and Safety Engineering"*, **17 (4),** 381-399.