Chapter 4

Nonparametric Shewhart-type control charts with runs-type signaling rules: Case K and Case U

4.0 Chapter overview

Introduction

The commonly used control charts in the statistical process control (SPC) environment with data that can be measured on a continuous numerical scale are generally designed and used with a specific parametric distribution, such as the normal distribution, in mind. It is well-known that if the underlying process is not as assumed, the performance of the parametric charts can be significantly degraded. In this context, one key problem is the lack of in-control robustness of some of the well-known parametric charts; this, for example, implies that there can be too many false alarms than what is typically expected and obviously this could mean considerable loss of time and resources (see e.g. Chakraborti et al., (2004)). Thus it is desirable, from a practical point of view, to develop and apply a set of control charts that are not designed under the assumption of normality (or any other parametric distribution). Such charts can expected to be more flexible in that they require no or little assumption regarding the underlying process distribution and would be useful in some applications.

To this end, nonparametric control charts are helpful. Nonparametric control charts have received considerable attention over the last few years. See, for example, Bakir (2004), Albers and Kallenberg (2004), Chakraborti et al. (2004) and Albers et al. (2006), where various nonparametric alternatives to the classical Shewhart-type charts are proposed, by adapting (for example) the corresponding nonparametric test for the process parameter, and have been shown to outperform the Shewhart $\overline{X}$ chart (and some other well-known charts) in terms of in-control robustness and efficiency, particularly

for heavy-tailed distributions. A thorough review of the literature on nonparametric charts can be found in Chakraborti et al. (2001, 2007).

The main advantage of the nonparametric charts is that they have known in-control properties that remain unchanged for all continuous distributions. Thus, for example, while the false alarm rate (FAR) of a Shewhart or Cumulative Sum (CUSUM) or Exponentially Weighted Moving Average (EWMA) chart for the mean will fluctuate depending on the underlying distribution, the FAR of a nonparametric chart can be calculated exactly and will remain the same (for in-control conditions) no matter what the distribution; this is a very useful feature for the practitioner.

A formal definition of a nonparametric or distribution-free control chart is given in terms of its run-length distribution. The number of samples that needs to be collected before the first out-of-control signal is given by a chart is a random variable called the run-length; the probability distribution of the run-length is referred to as the run-length distribution. If the in-control run-length distribution is the same for every continuous probability distribution the chart is called distribution-free or nonparametric (see e.g. Chakraborti et al., (2004)).

Note that, the term nonparametric is not intended to imply that there are no parameters involved, quite to the contrary. While the term distribution-free seems to be a better description of what one expects these charts to accomplish, nonparametric is perhaps the term more often used. In this chapter, both terms (distribution-free and nonparametric) are used to emphasize the fact that they mean the same.

Motivation

To construct a nonparametric control chart for the specified (or known or target) median of a distribution that is continuous and symmetric Bakir (2004) considered a Shewhart-type chart based on the Wilcoxon signed-rank (SR) test statistic. This chart, referred to as the *1-of-1* SR chart, signals when a single charting statistic falls outside the control limits, was shown to compete well with the Shewhart $\overline{X}$ chart in the case of the normal distribution and it performed better in the case of a heavier-tailed (than the normal) distribution such as the double exponential and the Cauchy. However, the false alarm rates for the *1-of-1* SR chart were just too high (i.e. the in-control average run-lengths were too short) for an application in practice, unless the subgroup size n was in the neighborhood of 20, which is not typical in SPC.

Having realized the potential and yet the practical shortcoming of the *1-of-1* SR chart, Chakraborti and Eryilmaz (2007) extended the idea of Bakir (2004) and considered an alternative class of nonparametric charts using the same SR statistic as charting statistic but incorporating some signaling rules based on runs; these charts are called runs-rule enhanced signed-rank charts. The new SR charts were shown to be more appealing from a practitioner's point of view in that they have much larger attainable in-control average run-lengths (ARL_0 's), much smaller attainable *FAR*'s and have better out-of-control (OOC) performance than the *1-of-1* SR chart.

Although the SR charts are useful, the requirement that the underlying distribution be symmetric is an additional assumption to be verified and may in fact not be satisfied in some situations in practice. If not much knowledge is available about the shape of the distribution, an alternative nonparametric test called the sign test can be used to make inference about any percentile including the median, whereas the SR test applies only to the median. The sign test does not require the assumption of symmetry of the underlying continuous population distribution (see e.g. Gibbons and Chakraborti, (2003)) and is easy to apply. Another advantage is that one does not require the actual measurements to be available to apply the sign test; all one needs to know is how many of the observations within each sample are larger (or smaller) than the specified value of the parameter (percentile) of interest. Thus the sign test can be applied with binary data, when the only information available, for each unit tested, is whether or not the measurement was higher (or lower) than the target value of the percentile of interest. Neither the normal theory chart nor the SR charts can be applied with just the dichotomized data.

A further requirement for applying the SR charts (and charts based on the sign test) is that the in-control process median (or percentile) must be specified; a situation commonly referred to as Case K. This may not be the case in some applications and could limit the application of the charts, with or without signaling rules. For example, when a new product is being developed not much information or expert knowledge may be initially available to specify the distribution and/or the in-control value of the percentile of interest. Hence there is a need to also develop control charts when the in-control process percentile of interest (or, in general, the location) is unknown. This is the scenario where the process distribution is continuous and unknown (no symmetry necessary) and the in-control percentile (or the location parameter) is unknown or unspecified (unlike in Case K); this situation is referred to as Case U.

Our objective is to overcome the drawbacks of the SR charts by studying and developing a new class of nonparametric control charts with runs-type signaling rules for the scenario where the

percentile (or location parameter) of interest of the process distribution is known and then, second, when it is unknown, without having to assume symmetry of the underlying process distribution. In the former situation (or Case K) the control charts are based on the well-known sign test statistic while in the latter scenario (or Case U) the charts are based on the two-sample median test statistic.

It will be seen that the charts we consider provide a new class of flexible, yet powerful, nonparametric charts to be used in practice.

Although one can consider other types of nonparametric charts such as the CUSUM and the EWMA (see e.g. the reviews by Chakraborti et al. (2001, 2007)), in this chapter, we keep the discussion focused on the Shewhart-type charts because of their inherent practical appeal and global effectiveness (see e.g. Montgomery, (2005) p. 385). The development of nonparametric CUSUM and EWMA charts will be a topic for future research.

Methodology

We use a Markov chain approach (see e.g. Fu and Lou, (2003)) to derive the necessary results (such as the run-length distributions, average run-lengths etc.) for our runs-rule enhanced charts because this approach provides a more compact and unified view of the derivations, and as stated by Balakrishnan and Koutras (2002), p.14, “The Markov techniques possess a great advantage (over the classical combinatory methods) as they are easily adjustable to many run-related problems; they often simplify the solutions to specific problems they are applied on and remain valid even for cases involving non-identical or dependent trials”. In some cases, however, we draw on the results of the geometric distribution of order k (see e.g. Balakrishnan and Koutras, (2002), Chapter 2) to obtain closed form and explicit expressions for the run-length distributions and/or their associated performance characteristics.

In Case U, like in Chapter 3, we use a two-step approach to derive the run-length distributions which involve the method of conditioning (see e.g. Chakraborti, (2000)). First we derive the conditional run-length distributions i.e. conditioned on two order statistics (control limits), which lets one focus on specific values of the control limits. Second we derive the unconditional (or marginal) run-length distributions by averaging over the joint distribution of the two order statistics. The unconditional run-length distributions reflect the bigger picture and reveal the overall performance of the charts taking into account that the control limits are estimated.

Layout of Chapter 4

In Section 4.1 we define and describe in detail (using graphs) the runs-type signaling rules for the one-sided and two-sided charts. In Section 4.2 we derive the run-length distributions of our new nonparametric control charts with signaling rules for Case K and then, in Section 4.3, we derive the run-length distributions of the charts for Case U. Section 4.4 gives a summary and some concluding remarks.

4.1 Runs-type signaling rules

Introduction

We consider a class of nonparametric Shewhart-type control charts for monitoring the percentile (or location parameter) of a process of which the distribution is assumed to be continuous but not necessarily symmetric. First we study the scenario where the π^{th} percentile of the process distribution is known and then, second, when it is unknown. In the former situation (or Case K), studied in Section 4.2, the control charts are based on the well-known sign test statistic while in the latter scenario (or Case U), which is looked at in Section 4.3, the charts are based on the two-sample median test statistic.

Signaling rules

The new control charts are "runs-rule enhanced" charts in which a process is declared out-of-control (OOC) when either

(i) A single charting statistic plots outside the control limits (*1-of-1* chart), or

(ii) k consecutive charting statistics all plot outside the control limits (*k-of-k* chart), or

(iii) exactly k of the last w charting statistics plot outside the control limits (*k-of-w* chart).

It is clear that rules (i) and (ii) are special cases of rule (iii). Rule (i) is the simplest and most frequently used whilst rules (ii) and (iii) have been used in the context of the parametric Shewhart-type charts such as the well-known $\overline{X}$ chart (see e.g. Derman and Ross, (1997) and Klein, (2000)).

One-sided and two-sided charts

We consider one-sided and two-sided control charts. Amongst the one-sided charts two situations can arise: (i) when just upward shifts are of interest so that an upper control limit is adequate, and (ii) when only detecting downward shifts are of interest so that a lower control limit is sufficient. The former is called the positive-sided (or upper one-sided) chart whereas the latter is labeled the negative-sided (or lower one-sided) chart. We study both the upper and the lower one-sided charts.

When a shift in any direction (up or down) is of concern a two-sided chart is used which has an upper and a lower control limit.

Charting statistic and control limits

We denote the charting statistic for the i^{th} subgroup, in general, by Q_i for $i = 1,2,3,...$ and the upper and the lower control limits by UCL and LCL, respectively; this allows us to simultaneously deal with Case K and Case U when we define and describe the runs-type signaling rules. Later, when we individually discuss the control charts of Case K and Case U we define and replace Q_i, UCL and LCL by their appropriate and representative counterparts.

Signaling indicators

Let

$$\xi_i^+ = I(Q_i \geq UCL) = \begin{cases} 1 & \text{if} \quad Q_i \geq UCL \\ 0 & \text{if} \quad Q_i < UCL \end{cases} \tag{4-1}$$

and

$$\xi_i^- = I(Q_i \leq LCL) = \begin{cases} 1 & \text{if} \quad Q_i \leq LCL \\ 0 & \text{if} \quad Q_i > LCL \end{cases} \tag{4-2}$$

for $i = 1,2,3,...$ denote the indicator functions for the one-sided charts corresponding to the events $\{Q_i \geq UCL\}$ and $\{Q_i \leq LCL\}$, respectively. In other words, ξ_i^+ (ξ_i^-) denotes the signaling indicator for the event when Q_i plots on or above (below) the upper (lower) control limit of the positive-sided (negative-sided) chart. Likewise, we let

$$\xi_i = \begin{cases} 1 & \text{if} \quad Q_i \geq UCL \\ 0 & \text{if} \quad LCL < Q_i < UCL \\ 2 & \text{if} \quad Q_i \leq LCL \end{cases} \tag{4-3}$$

denote the signaling indicator for the two-sided chart so that ξ_i indicates whether Q_i plots on or above the UCL (in which case $\xi_i = 1$), between the LCL and the UCL (so that $\xi_i = 0$), or on or below the LCL ($\xi_i = 2$).

Remark 1

Not only do the signaling indicators in (4-1), (4-2) and (4-3) allow us to clearly define and describe signaling rules (i), (ii) and (iii), but their statistical properties (e.g. whether they are independent, their "success" probabilities such as $\Pr(\xi_i^+ = 1)$, $\Pr(\xi_i^- = 1)$ and $\Pr(\xi_i^\pm = 1)$ etc.) are also important since they play a key role in deriving the run-length distributions of the new class of proposed runs-rule enhanced charts.

4.1.1 The *1-of-1* charts

The *1-of-1* charts are the least complicated and most frequently used charts. The signaling rules for the ***1-of-1*** charts, as defined by the signaling indicators in (4-1), (4-2) and (4-3), are given by:

The ***1-of-1* upper one-sided** chart signals when event A_1 occurs where:

$$A_1 : \{Q_i \geq UCL\} \Leftrightarrow \{\xi_i^+ = 1\}^*,$$

The ***1-of-1* lower one-sided** chart signals when event A_2 occurs where:

$$A_2 : \{Q_i \leq LCL\} \Leftrightarrow \{\xi_i^- = 1\}, \text{ and}$$

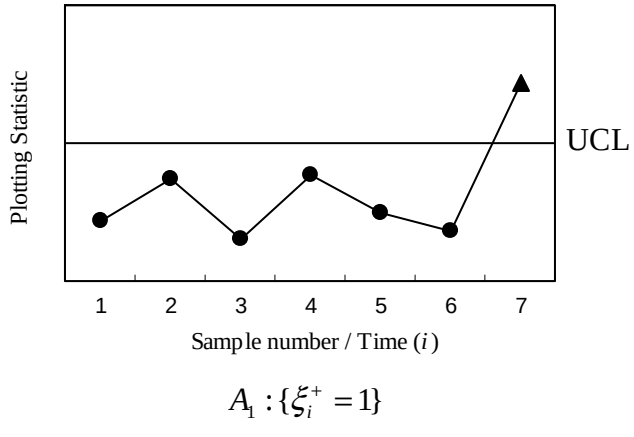
The ***1-of-1* two-sided** chart signals when event A occurs where:

$$A : \{Q_i \geq UCL \text{ or } Q_i \leq LCL\} \Leftrightarrow \{A_1 \text{ or } A_2\} \Leftrightarrow \{\xi_i^+ = 1 \text{ or } \xi_i^- = 1\} \Leftrightarrow \{\xi_i = 1 \text{ or } \xi_i = 2\}.$$

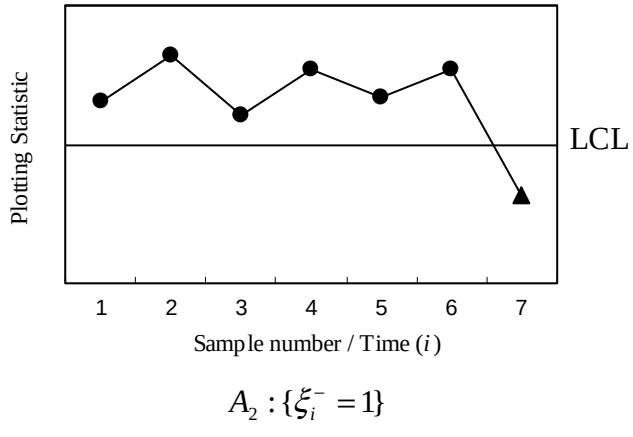
* The symbol "$\Leftrightarrow$" in an expression such as $P \Leftrightarrow Q$ is read as 'P is true if and only if Q is true'.

For illustration, panels (a) and (b) of Figure 4.1 show examples of the *1-of-1* upper and lower one-sided charts whilst Figure 4.2 displays examples of the *1-of-1* two-sided chart.

The *1-of-1* upper (lower) one-sided charts signals, for illustration only, at time $i = 7$ when Q_7 plots above (below) the UCL (LCL). The process is therefore declared OOC with the conclusion of an upward (downward) shift in the process location. Similarly, both of the *1-of-1* two-sided charts signal at time $i = 7$; the first chart signals when Q_7 plots above the upper control limit (indicative of an upward shift) whereas the second chart signals when Q_7 plots below the LCL (indicative of a downward shift).



(a) The *1-of-1* rule for the upper one-sided chart (upward shift detected)



(b) The *1-of-1* rule for the lower one-sided chart (downward shift detected)

Figure 4.1: The *1-of-1* rule for the upper and the lower one-sided charts

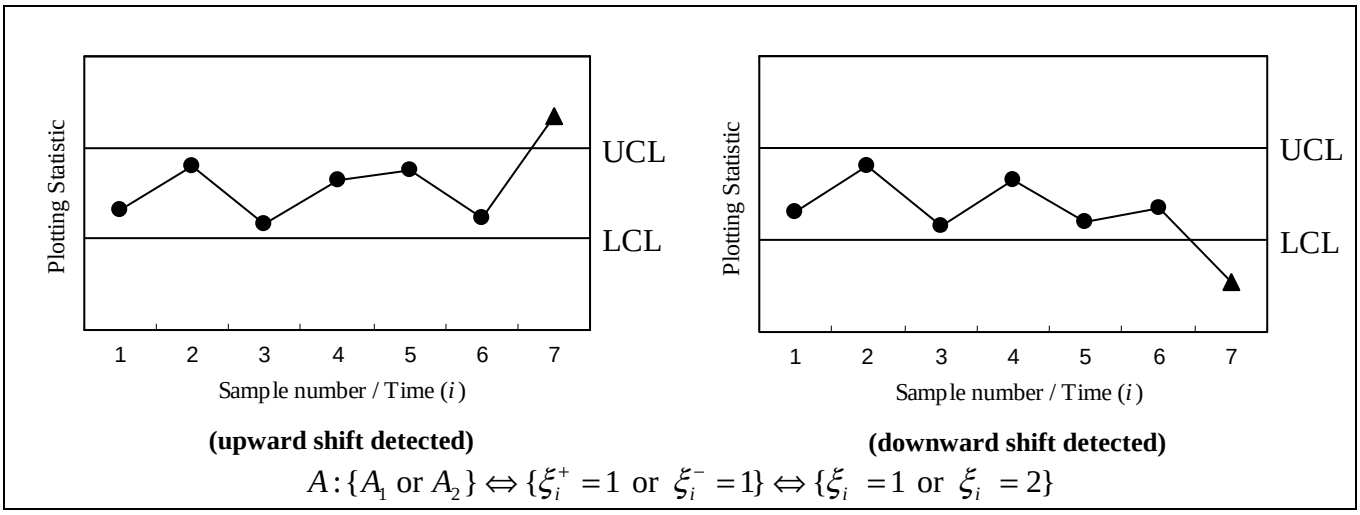


Figure 4.2: The *1-of-1* rule for the two-sided chart

Because the *1-of-1* charts are based on signaling rule (i) which uses only the information from the most recent (last) sample to make a decision whether or not the process is in-control (IC), one feels these charts can be improved upon by using rules (ii) and (iii); this idea is discussed in the next section.

4.1.2 The *k-of-k* and *k-of-w* charts

The runs-rule Shewhart-type charts we consider adopt a sequential approach and use the information from multiple samples including the most recent one to signal. Unlike the CUSUM and EWMA charts, the sequential approach we study is carried out by considering runs of the charting statistic outside the control limit(s) which includes the charting statistic from the current sample and one or more charting statistics from past samples. The resulting charts are easy to use and it will be seen that they offer the user greater practical flexibility in designing a chart so that more (practically) attractive, attainable, false alarm rates are available. Moreover, it will be shown that the new charts have higher efficiency (i.e. smaller OOC *ARL*'s) compared to the *1-of-1* charts.

According to the *k-of-k* $(k \geq 2)$ chart the control chart signals at any point in time when k consecutive charting statistics (from k consecutive samples), of which the last one is the most recent one, all plot outside the control limit(s). A generalization of the *k-of-k* chart is the *k-of-w* chart which signals when exactly k of the last w charting statistics all plot outside the limit(s), of which the last one plots outside the control limits.

It is clear that we can consider charts for any pair of positive integers k and w where $1 \leq k \leq w$ and $w \geq 2$. Although various values of k and w can be considered in theory, from a practical point of view, it is important that the resulting charts are easy to apply; so we focus on the *2-of-2* $(k = w = 2)$ and the *2-of-3* $(k = 2, w = 3)$ charts. A user is therefore required to keep track of only the last two or three of the most recent charting statistics.

4.1.2.1 One-sided *k-of-k* and *k-of-w* charts

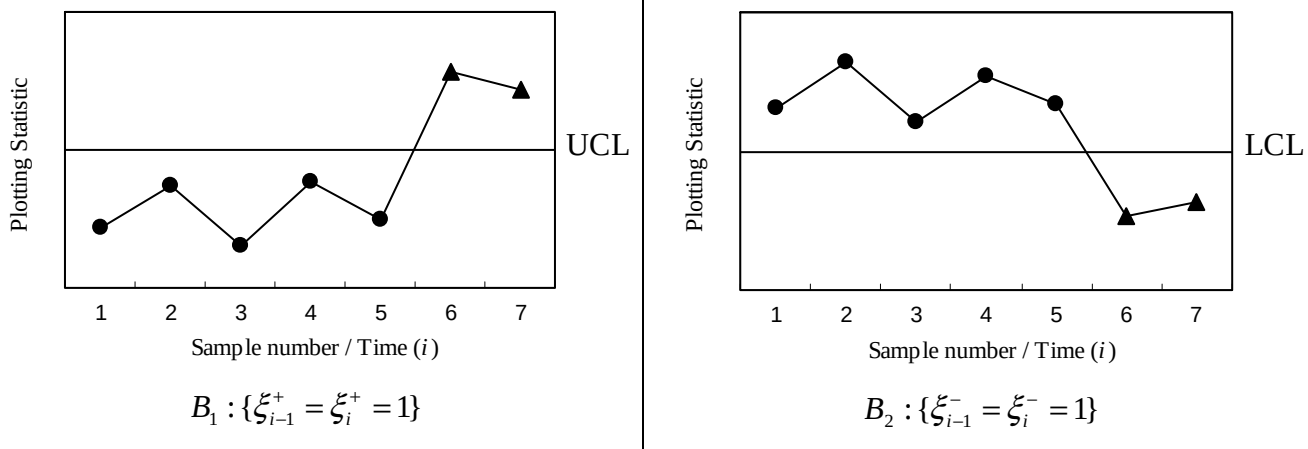
As noted earlier, the upper (lower) one-sided chart has only an upper (lower) control limit and is typically used when only an upward (downward) shift is of concern. We first describe the *2-of-2* upper (lower) one-sided chart and then the *2-of-3* upper (lower) one-sided chart.

One-sided *2-of-2* charts

The *2-of-2* chart requires the user to keep track of only the last two charting statistics Q_{i-1} and Q_i at any given point in time, $i \geq 2$. The upper one-sided *2-of-2* chart signals (declares the process OOC) if both Q_{i-1} and Q_i plot on or outside the upper control limit; otherwise no signal is given and we declare the process IC. Likewise, the lower one-sided *2-of-2* chart signals if both Q_{i-1} and Q_i plot on or outside the lower control limit. Thus the ***2-of-2*** one-sided charts are:

The ***2-of-2* upper one-sided** chart signals when the event B_1 occurs
where $B_1 : \{\xi_{i-1}^{+} = \xi_i^{+} = 1\}$, and
The ***2-of-2* lower one-sided** chart signals when the event B_2 occurs
where $B_2 : \{\xi_{i-1}^{-} = \xi_i^{-} = 1\}$.

For illustration, panels (a) and (b) of Figure 4.3 show examples of the *2-of-2* upper and the *2-of-2* lower one-sided charts. Both of the charts signal, again for illustration only, at time $i = 7$ i.e. on the first occurrence of a run of length two of the charting statistic above (below) the UCL (LCL) at time or sample number 7. Hence, the process is declared OOC with the conclusion of an upward (downward) shift in the process location.



(a) The *2-of-2* rule for the upper one-sided chart (upward shift detected)

(b) The *2-of-2* rule for the lower one-sided chart (downward shift detected)

Figure 4.3: The *2-of-2* rule for the upper and the lower one-sided charts

Remark 2

The *2-of-2* one-sided charts can be alternatively defined in terms of the minimum and/or the maximum of the statistics Q_{i-1} and Q_i. For example, the *2-of-2* upper one-sided chart signals if $\min(Q_{i-1}, Q_i)$ plots on or above the *UCL*. Similarly the *2-of-2* lower one-sided chart signals if $\max(Q_{i-1}, Q_i)$ plots on or below the *LCL*. We next consider a generalization of the *2-of-2* chart.

One-sided *2-of-3* charts

The *2-of-2* charts utilize moving (over time) blocks of only two charting statistics. It is therefore natural to investigate if there can be any sizeable gain in efficiency when moving blocks of three charting statistics are utilized. Thus we consider *2-of-3* charts for which at any time point $i \geq 3$ we need to keep track of Q_{i-2}, Q_{i-1} and Q_i; the upper (lower) one-sided *2-of-3* chart signals if two of these three statistics plot on or above (below) the upper (lower) control limit.

A similar chart was considered by Klein (2000) for the Shewhart $\overline{X}$ chart. However, note that although there are three ways for exactly two of the last three charting statistics to plot on or above (below) the upper (lower) control limit, we take only two of the three ways, namely where the last charting statistic plots on or above (below) the upper (lower) control limit to define a signal.

This is unlike the chart of Klein (2000) which can signal in any of the three ways; our charts eliminate the possibility that the process is declared OOC when both the first and the second charting statistics plot outside the control limits but the third (last) one plots between the control limits (see e.g. Figure 4.5), which we feel is somewhat undesirable in practice. Thus our ***2-of-3*** one-sided charts are:

The ***2-of-3* upper one-sided** chart signals when the event $C(+)=\{C_1 \text{ or } C_2\}$ occurs

where $C_1:\{Q_{i-2}<UCL \text{ and } Q_{i-1}\geq UCL \text{ and } Q_i \geq UCL\} \Leftrightarrow \{\xi^+_{i-2}=0, \xi^+_{i-1}=1, \xi^+_i=1\}$

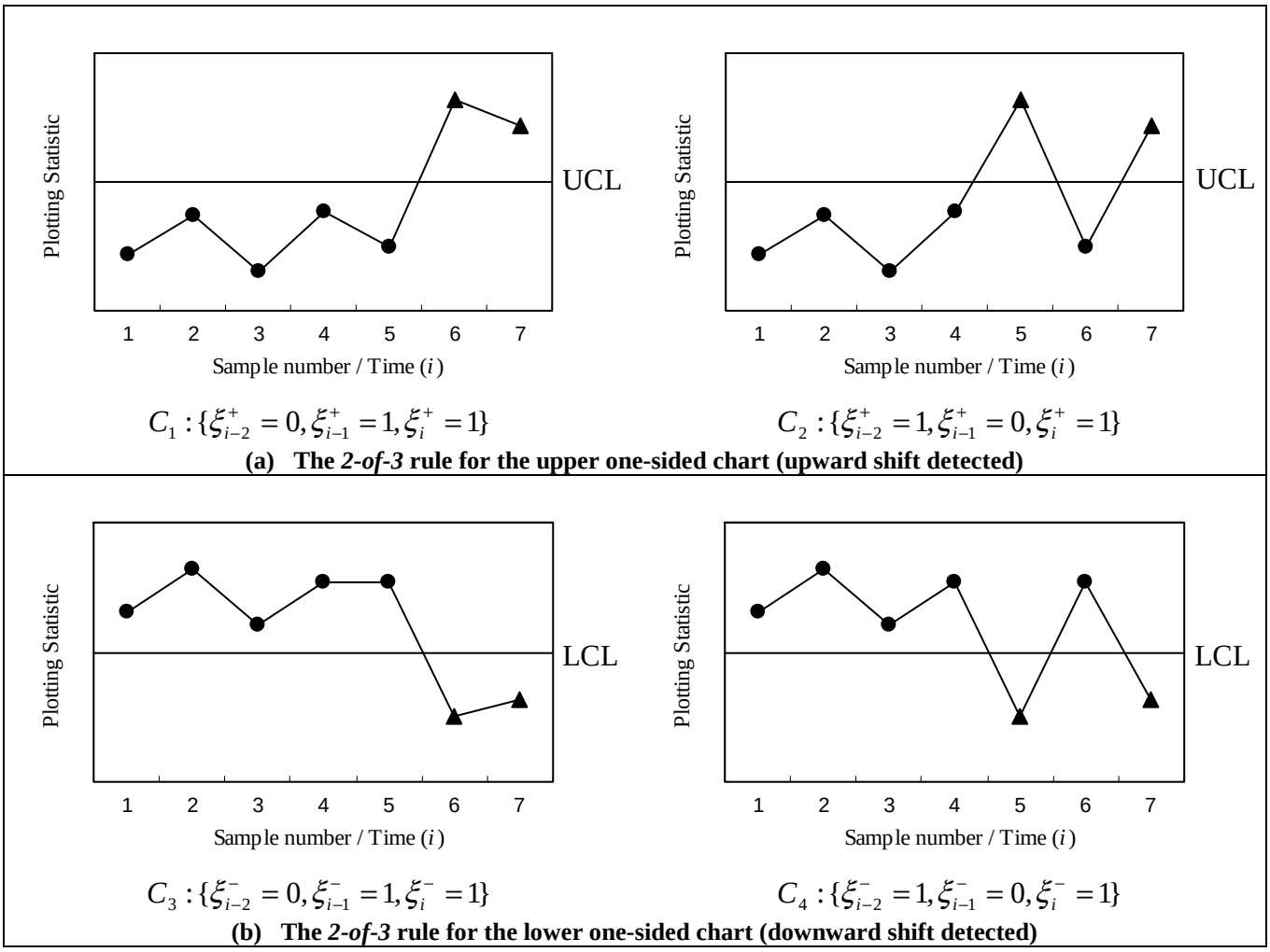
$C_2:\{Q_{i-2}\geq UCL \text{ and } Q_{i-1}<UCL \text{ and } Q_i \geq UCL\} \Leftrightarrow \{\xi^+_{i-2}=1, \xi^+_{i-1}=0, \xi^+_i=1\}$.

The ***2-of-3* lower one-sided** chart signals when the event $C(-)=\{C_3 \text{ or } C_4\}$ occurs

where $C_3:\{Q_{i-2}>LCL \text{ and } Q_{i-1}\leq LCL \text{ and } Q_i \leq LCL\} \Leftrightarrow \{\xi^-_{i-2}=0, \xi^-_{i-1}=1, \xi^-_i=1\}$

$C_4:\{Q_{i-2}\leq LCL \text{ and } Q_{i-1}>LCL \text{ and } Q_i \leq LCL\} \Leftrightarrow \{\xi^-_{i-2}=1, \xi^-_{i-1}=0, \xi^-_i=1\}$.

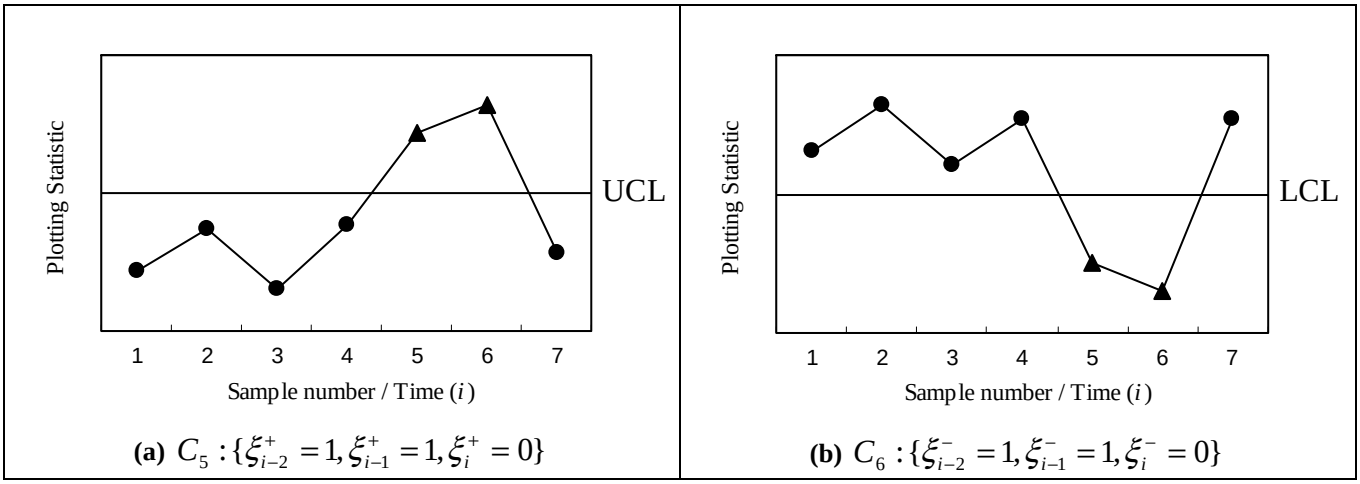
Panels (a) and (b) of Figure 4.4 show examples of what the signaling events C_1, C_2, C_3 and C_4 might look like in case of the *2-of-3* upper and lower one-sided charts. For example, both of the *2-of-3* upper one-sided charts, shown in panel (a), signal at time $i=7$ and the signals are interpreted to be indicative of an upward shift since both the charting statistics fall above the upper control limit.



(a) The *2-of-3* rule for the upper one-sided chart (upward shift detected)

(b) The *2-of-3* rule for the lower one-sided chart (downward shift detected)

Figure 4.4: The *2-of-3* rule for the upper and the lower one-sided charts



(a) $C_5 : \{\xi^+_{i-2} = 1, \xi^+_{i-1} = 1, \xi^+_i = 0\}$

(b) $C_6 : \{\xi^-_{i-2} = 1, \xi^-_{i-1} = 1, \xi^-_i = 0\}$

Figure 4.5: The *2-of-3* events excluded as signaling events for the upper and the lower one-sided charts

4.1.2.2 Two-sided *k-of-k* and *k-of-w* charts

Two-sided charts are typically used to detect either an upward or a downward shift and thus have both an upper and a lower control limit.

Two-sided *2-of-2* charts

Like the *2-of-2* one-sided charts, we also need to keep track of both Q_{i-1} and Q_i at any time point $i \geq 2$, but for the two-sided chart there are two control limits and so there are four ways for Q_{i-1} and Q_i to plot outside the limits. Any one (or more) of the four scenarios may be used to define a signal. We consider two *2-of-2* two-sided charts; both capable of detecting an upward or a downward shift in the location parameter.

The first *2-of-2* two-sided chart signals when any two successive charting statistics both plot on or outside the control limits. In other words, a signal is given when:

(i) both charting statistics plot on or above the UCL, or
(ii) both charting statistics plot on or below the LCL, or
(iii) the first charting statistic plots on or above the UCL and the second charting statistic plots on or below the LCL, or
(iv) the first charting statistic plots on or below the LCL and the second charting statistic plots on or above the UCL.

This signaling rule was proposed by Derman and Ross (1997) in the context of the Shewhart $\overline{X}$ chart; we refer to this chart as the ***2-of-2* DR** two-sided chart.

The second *2-of-2* two-sided chart signals when two successive charting statistics:

(i) both plot on or above the UCL, or
(ii) both plot on or below the LCL.

This signaling rule was considered by Klein (2000) in the context of the Shewhart $\overline{X}$ chart; we refer to this chart as the ***2-of-2* KL** two-sided chart.

More specifically, the ***2-of-2*** two-sided charts are:

The ***2-of-2*** **DR two-sided** chart signals when the event $D(DR):\{D_1 \text{ or } D_2 \text{ or } D_3 \text{ or } D_4\}$ occurs, and

The ***2-of-2*** **KL two-sided** chart signals when the event $D(KL)=\{D_1 \text{ or } D_2\}$ occurs

where

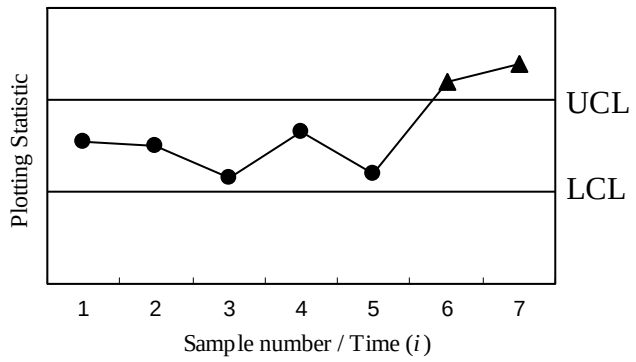
$$D_1:\{Q_{i-1}\geq UCL \text{ and } Q_i \geq UCL\} \Leftrightarrow \{\xi_{i-1}=1,\xi_i=1\},$$

$$D_2:\{Q_{i-1}\leq LCL \text{ and } Q_i \leq LCL\} \Leftrightarrow \{\xi_{i-1}=2,\xi_i=2\},$$

$$D_3:\{Q_{i-1}\geq UCL \text{ and } Q_i \leq LCL\} \Leftrightarrow \{\xi_{i-1}=1,\xi_i=2\}, \text{ and}$$

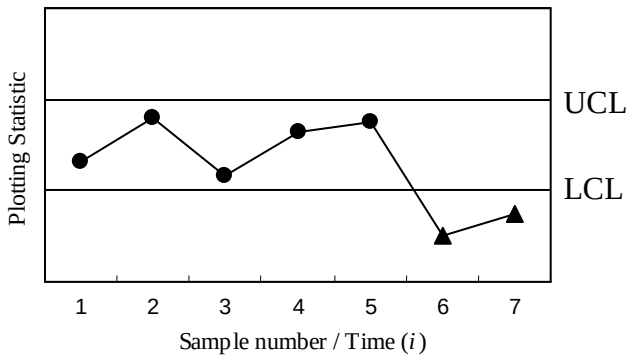
$$D_4:\{Q_{i-1}\leq LCL \text{ and } Q_i \geq UCL\} \Leftrightarrow \{\xi_{i-1}=2,\xi_i=1\}.$$

Figure 4.6 shows some examples of the events D_i, $i=1,2,3,4$. It is clear that the *2-of-2* DR chart signals on the seventh sample in each of the four panels of Figure 4.6 whereas the *2-of-2* KL chart signals only in panels (a) and (b). Thus, whenever the *2-of-2* KL chart signals so does the *2-of-2* DR chart, but the converse may not always happen. Furthermore, it seems that the *2-of-2* DR and KL charts are both suitable for detecting an upward or a downward shift, but the *2-of-2* DR chart can also detect a possible swing; this is when an upward shift is immediately followed by a downward shift or vice versa (Chakraborti and Eryilmaz, (2007)).



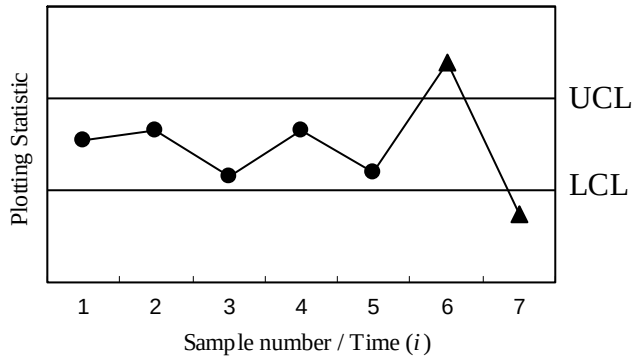
$D_1 : \{\xi_{i-1} = 1, \xi_i = 1\}$

(a) (signal above / upward shift detected)



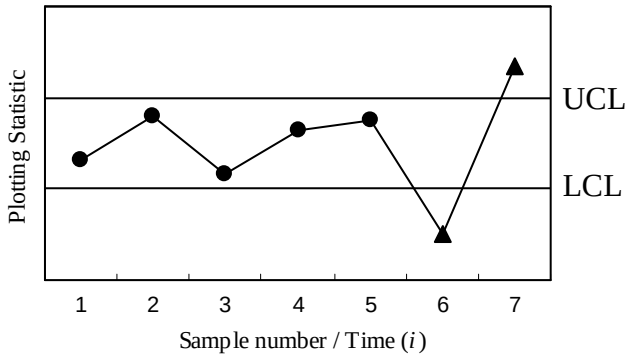
$D_2 : \{\xi_{i-1} = 2, \xi_i = 2\}$

(b) (signal below / downward shift detected)



$D_3 : \{\xi_{i-1} = 1, \xi_i = 2\}$

(c) (signal above followed by a signal below / possible swing detected)



$D_4 : \{\xi_{i-1} = 2, \xi_i = 1\}$

(d) (signal below followed by a signal above / possible swing detected)

Figure 4.6: The *2-of-2* rule for the two-sided charts

Two-sided *2-of-3* chart

Analogous to the *2-of-3* one-sided charts the *2-of-3* two-sided chart signals when:

(i) exactly two of the last three charting statistics both plot on or above the UCL, or

(ii) exactly two of the last three charting statistics both plot on or below the LCL.

Hence, the ***2-of-3*** two-sided chart defined in terms of the signaling indicator ξ_i is given by:

The ***2-of-3* two-sided** chart signals when event $E = \{E_1 \text{ or } E_2 \text{ or } E_3 \text{ or } E_4\}$ occurs where

$E_1 : \{LCL < Q_{i-2} < UCL \text{ and } Q_{i-1} \geq UCL \text{ and } Q_i \geq UCL\} \Leftrightarrow \{\xi_{i-2} = 0, \xi_{i-1} = 1, \xi_i = 1\}$,

$E_2 : \{Q_{i-2} \geq UCL \text{ and } LCL < Q_{i-1} < UCL \text{ and } Q_i \geq UCL\} \Leftrightarrow \{\xi_{i-2} = 1, \xi_{i-1} = 0, \xi_i = 1\}$,

$E_3 : \{LCL < Q_{i-2} < UCL \text{ and } Q_{i-1} \leq LCL \text{ and } Q_i \leq LCL\} \Leftrightarrow \{\xi_{i-2} = 0, \xi_{i-1} = 2, \xi_i = 2\}$, and

$E_4 : \{Q_{i-2} \leq LCL \text{ and } LCL < Q_{i-1} < UCL \text{ and } Q_i \leq LCL\} \Leftrightarrow \{\xi_{i-2} = 2, \xi_{i-1} = 0, \xi_i = 2\}$.

Figure 4.7 displays examples of events E_i for $i = 1,2,3,4$ and shows that when there is a signal, the proposed *2-of-3* two-sided chart offers a practical interpretation for the signal. For example, when either event E_1 or E_2 occurs (shown in panels (a) and (b)) the signal is interpreted as an upward shift. Similarly, if either event E_3 or E_4 occurs (displayed in panels (c) and (d)) a downward shift is inferred.

Remark 3

Apart from events E_1, E_2, E_3 and E_4 there are a further eight scenarios in case of the *2-of-3* two-sided chart where exactly two of the last three charting statistics can plot outside the control limits. We, however, exclude these events as signaling events when we calculate the statistical characteristics or properties of the *2-of-3* two-sided control charts; even though four of the events may possibly be linked to genuine or tangible changes in the process.

Figure 4.8 shows the eight events together with the practical interpretations (if any). For example, panels (a) and (b) show events E_5 and E_6 that could be considered a swing in the process, whereas panels (c) and (d) show events E_7 and E_8 that could be interpreted as trends (up or down) in the process. The events in panels (e), (f), (g) and (h) are excluded as signaling events because, as mentioned earlier, the last point plots between the control limits.

Most importantly, by excluding events E_5, E_6, …, E_{12} we are left with events E_1, E_2, E_3 and E_4, which makes the signaling events of the *2-of-3* one-sided charts and that of the *2-of-3* two-sided chart more alike (compare, for example, the signaling events shown in Figure 4.4 with that of Figure 4.7).

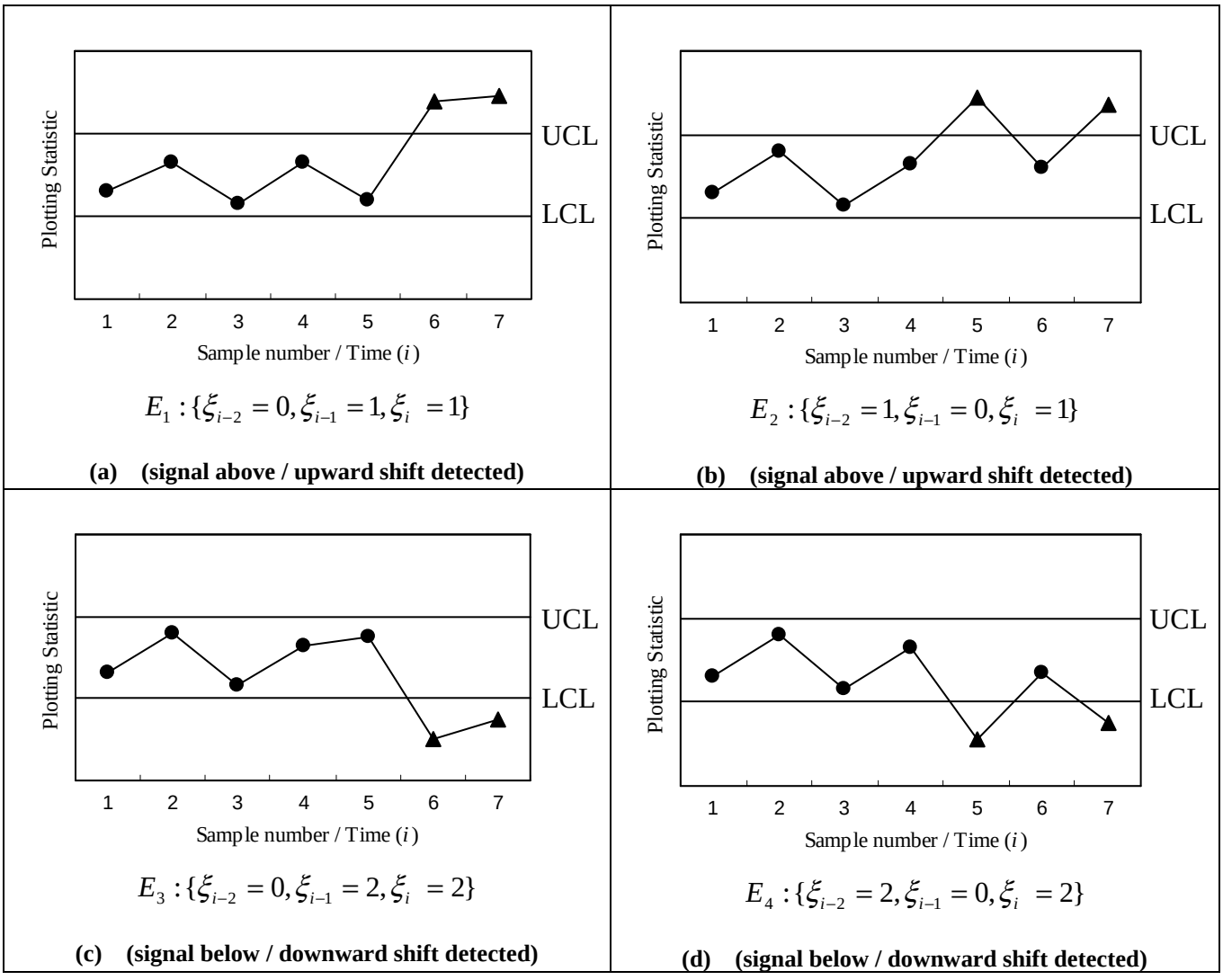
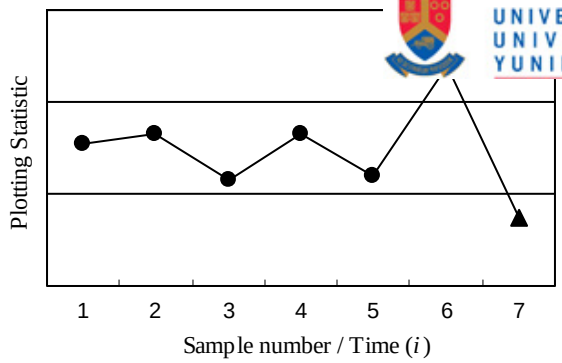
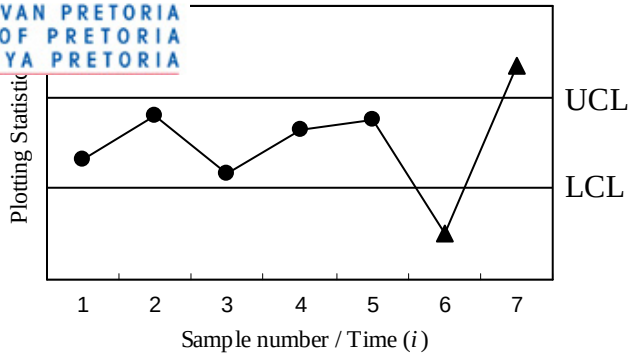


Figure 4.7: The *2-of-3* rule for the two-sided chart



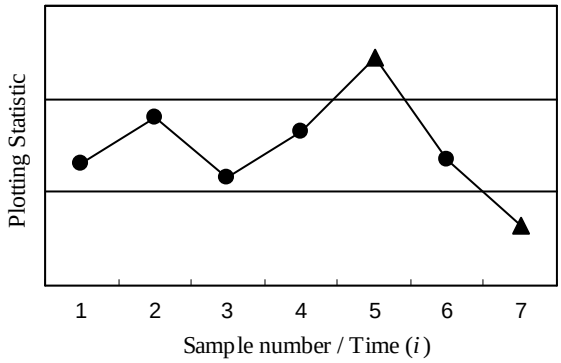
$E_5 : \{\xi_{i-2} = 0, \xi_{i-1} = 1, \xi_i = 2\}$

(a) (possible swing detected)



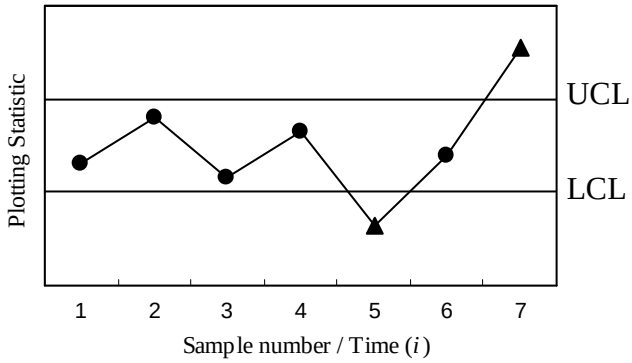
$E_6 : \{\xi_{i-2} = 0, \xi_{i-1} = 2, \xi_i = 1\}$

(b) (possible swing detected)



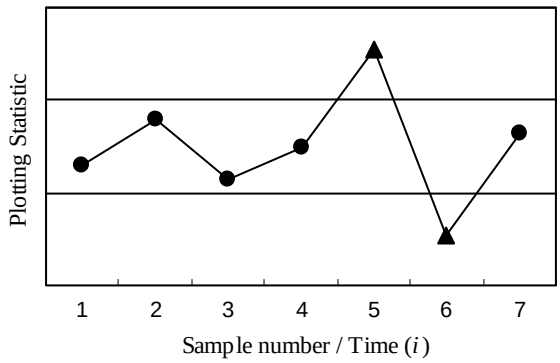
$E_7 : \{\xi_{i-2} = 1, \xi_{i-1} = 0, \xi_i = 2\}$

(c) (downward trend detected)



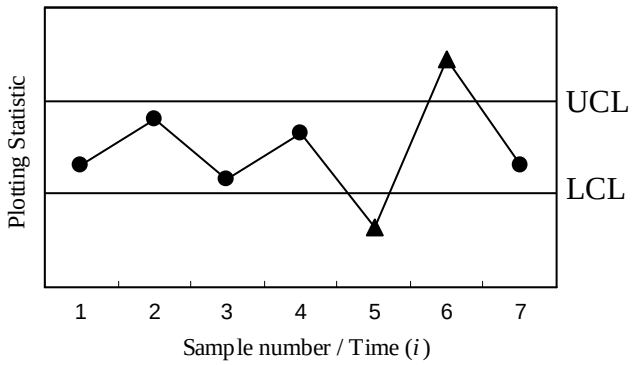
$E_8 : \{\xi_{i-2} = 2, \xi_{i-1} = 0, \xi_i = 1\}$

(d) (upward trend detected)



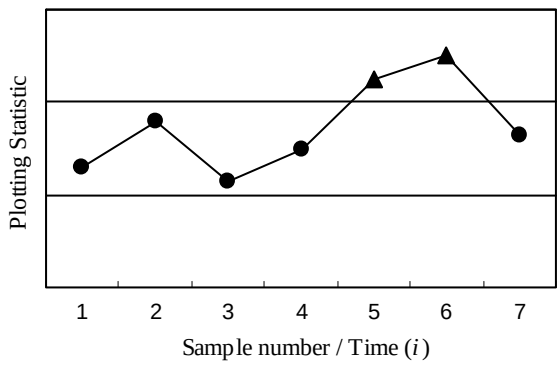
$E_9 : \{\xi_{i-2} = 1, \xi_{i-1} = 2, \xi_i = 0\}$

(e) (last point plots between *LCL* and *UCL*)



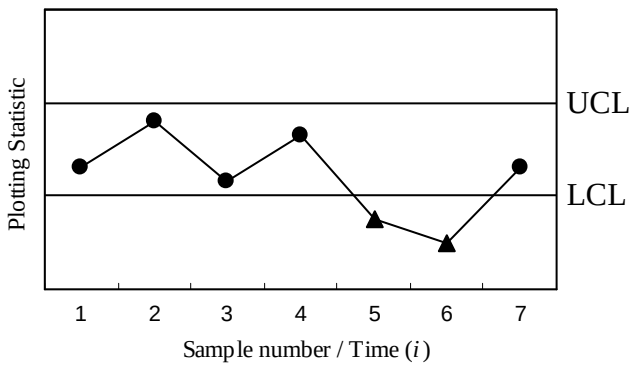
$E_{10} : \{\xi_{i-2} = 2, \xi_{i-1} = 1, \xi_i = 0\}$

(f) (last point plots between *LCL* and *UCL*)



$E_{11} : \{\xi_{i-2} = 1, \xi_{i-1} = 1, \xi_i = 0\}$

(g) (last point plots between *LCL* and *UCL*)



$E_{12} : \{\xi_{i-2} = 2, \xi_{i-1} = 2, \xi_i = 0\}$

(h) (last point plots between *LCL* and *UCL*)

Figure 4.8: The *2-of-3* events excluded as signaling events for the two-sided charts

Section 4.1 outlined, in general, the operation of the runs-rule enhanced charts. However, we are yet to define (choose) the charting statistic Q_i and the control limits (UCL and LCL). In Sections 4.2 and 4.3 we do just this and show, in particular, how to obtain the run-length distributions and how to design and implement the runs-rule enhanced charts in case the π^{th} percentile of the process distribution is known (Case K) and unknown (Case U). The performance of the charts is then further examined via properties of their run-length distributions such as the average run-length (ARL), the false alarm rate (FAR), the standard deviation of the run-length ($SDRL$) and some of the percentiles of the run-length distributions.

4.2 Sign charts for the known π^{th} quantile (Case K)

Introduction

Case K refers to the situation when the π^{th} quantile (or percentile) of the process distribution is known or specified. The new control charts we design for Case K are based on the well-known sign (test) statistic.

Assumptions

Let $(X_{i1}, X_{i2}, ..., X_{in})$ denote a random sample of size $n > 1$ taken at sampling stage (time) $i = 1,2,3,....$ Assume that the samples are independent and the observations come from a continuous distribution with cumulative distribution function (c.d.f) $F_X(x)$ with the unique $100\pi^{th}$ percentile denoted by $\theta = F_X^{-1}(\pi)$, $0 < \pi < 1$.

In many cases the percentile of interest is the median because it is a robust measure of central tendency so that $\pi = 0.5$ and $\theta = F_X^{-1}(0.5)$, however this is not necessary for our developments as the new sign charts can be applied for any percentile of interest.

Charting statistics

Amin et al. (1995) considered a *1-of-1* Shewhart-type sign chart for monitoring the median of a distribution based on the charting statistic

$$SN_i = \sum_{j=1}^{n} sign(X_{ij} - \theta_0) \quad \text{for} \quad i = 1,2,... \tag{4-4}$$

where $sign(x) = 1$ if $x > 0$, 0 if $x = 0$ and -1 if $x < 0$ and θ_0 denotes the specified value of the median.

We consider any percentile $\theta = F_X^{-1}(\pi)$ for $0 < \pi < 1$ and the charting statistic for our sign charts is the classical sign statistic

$$T_i = \sum_{j=1}^{n} I(X_{ij} > \theta_0) \quad \text{for} \quad i = 1,2,... \tag{4-5}$$

where θ_0 denotes the known (or the specified or the target) value of the percentile of interest and $I(X_{ij} > \theta_0)$ denotes the indicator function for the event $\{X_{ij} > \theta_0\}$. Thus T_i denotes the number of observations larger than θ_0 in the i^{th} sample and it is easily seen that T_i follows a binomial distribution with parameters n and probability of success $p = \Pr(X_{ij} > \theta_0)$. When $\theta = \theta_0$ the percentile of interest is equal to its specified value, the process is said to be in-control (IC) and then p is denoted by p_0 and equals

$$p_0 = \Pr(X_{ij} > \theta_0 \mid IC) = 1 - \pi . \qquad (4\text{-}6)$$

Thus, for example, when the percentile of interest is the median ($\pi = 0.5$), the process is IC when $\theta = \theta_0$ (the specified value of the median) and then $p = p_0 = 0.5$; similarly when θ is the first quartile ($\pi = 0.25$), the process is IC when $\theta = \theta_0$ (the specified value of the first quartile) and $p = p_0 = 0.75$ and so on.

Control limits

The upper and lower control limits of our sign charts are

$$UCL = n - b \qquad \text{and} \qquad LCL = a \qquad (4\text{-}7)$$

where the charting constants a and b are integers between (and including) zero and n, that is, $a, b \in \{0,1,2,...,n\}$, and selected so that the UCL is greater than the LCL; determination of a and b will be discussed later. Note that, the new sign charts do not have a centerline.

4.2.1 Run-length distributions of the sign charts

The run-length distribution and its associated characteristics (such as the mean, the standard deviation, the median etc.) reveal important information regarding the performance of a control chart (see e.g. Human and Graham, (2007)).

There are various approaches to finding the run-length distribution. We use (for the most part) a Markov chain approach (see e.g. Fu and Lou, (2003)) to derive the necessary results for our runs-rule enhanced charts because this approach provides a more compact and unified view of the derivations, and as stated by Balakrishnan and Koutras (2002), p.14, "The Markov techniques possess a great advantage (over the classical combinatory methods) as they are easily adjustable to many run-related problems; they often simplify the solutions to specific problems they are applied on and remain valid even for cases involving non-identical or dependent trials".

The Markov chain approach entails that we:

(a) classify each charting statistic T_i (based on its value) into one of two categories (for a one-sided chart) or into one of three categories (for a two-sided chart) depending on whether T_i plots on or above the UCL, on or below the LCL and/or between the LCL and UCL,

(b) define a new sequence of random variables $Y_1, Y_2, Y_3, \ldots$ (say) that keeps track of the classification of the $T_i's$, and then

(c) construct a Markov Chain $\{Z_i : i \geq 0\}$ to find the run-length distribution.

For example, consider the upper one-sided sign chart. Each T_i can be either on or above the UCL or below. Let $Y_i = 1$ (a success) in the former case and $Y_i = 0$ (a failure) in the latter case. Thus, corresponding to a sequence of $T_i's$ we get a sequence of $Y_i's$ that are all binary; for example, if $(T_1, T_2, T_3, T_4) = (4, 3, 8, 7)$ and $UCL = 5$, we get $(Y_1, Y_2, Y_3, Y_4) = (0, 0, 1, 1)$.

Thus, the run-length of the *1-of-1* upper one-sided chart i.e. the time when for the first time a T_i plots on or above the UCL, is "3" for our example and can be equivalently expressed as the time when for the first time we obtain a "1" (a success) among the four $Y_i's$.

Similarly, the run-length of the *2-of-2* upper one-sided chart i.e. the time when for the first time two consecutive $T_i's$ plot out-of-control, which equals "4" in our example, can be equivalently viewed as the time when for the first time we obtain two successive "1's" (or successes) among the four $Y_i's$.

Hence the run-length for the $T_i's$ can be equivalently defined as the waiting time for the first success (or, more generally, the first particular run or pattern of successes) among the $Y_i's$; it is this correspondence that makes the study of the statistical properties of the run-length random variable more amenable using results about the waiting time distributions in a sequence of Bernoulli (binary or two-state) and other types (three or more states) of random variables.

There is a rather vast literature on waiting time distributions. A detailed discussion about general results on the distribution theory of runs and patterns with various applications can be found in Balakrishnan and Koutras (2002) and Fu and Lou (2003). Some of these results pertain to the exact probability distribution of the waiting time for the first occurrence of a simple or a compound pattern in a sequence of i.i.d. (or homogeneous Markov dependent) 2-state (binary) or 3 or more-state trials (see e.g. Fu and Lou, (2003); Chapters 3, 4 and 5). The approach is to "properly imbed" (see e.g. Fu and Lou (2003), page 64; Definition 2.6) the random variable of interest (the run-length in our case) into a finite Markov chain which means constructing a "proper" Markov chain so that the probability that the run-length random variable N takes on some specific value is expressed in terms of the probability that the imbedded Markov chain $\{Z_i : i \geq 0\}$ resides in a specific subset S of the state space Ω.

The latter probability can be more easily computed using results about the transition probability matrix of the Markov chain. For example, given the $m \times m$ transition probability matrix of the Markov chain

$$\mathbf{M}_{m\times m} = \left[\begin{array}{c|c} \mathbf{Q}_{h\times h} & \mathbf{C}_{h\times(m-h)} \\ \hline \mathbf{0}_{(m-h)\times h} & \mathbf{I}_{(m-h)\times(m-h)} \end{array}\right]$$

(written in a partitioned form), the probability mass function (p.m.f), the expected value (ARL) and the variance ($VARL$) of the run-length random variable N can be directly obtained, using Theorems 5.2 and 7.4 of Fu and Lou (2003), as

$$P(N = j \mid n, a, b, \theta) = \boldsymbol{\xi}\mathbf{Q}^{j-1}(\mathbf{I} - \mathbf{Q})\mathbf{1} \qquad \text{for} \qquad j = 1,2,3,... \tag{4-8}$$

$$E(N \mid n, a, b, \theta) = \boldsymbol{\xi}(\mathbf{I} - \mathbf{Q})^{-1}\mathbf{1} \tag{4-9}$$

and

$$\text{var}(N \mid n, a, b, \theta) = \boldsymbol{\xi}(\mathbf{I} + \mathbf{Q})(\mathbf{I} - \mathbf{Q})^{-2}\mathbf{1} - (E(N))^2 \tag{4-10}$$

where the sub-matrix matrix $\mathbf{Q} = \mathbf{Q}_{h\times h}$ is called the essential transition probability sub-matrix, $\mathbf{I} = \mathbf{I}_{h\times h}$ (used in (4-8) and (4-10)) and $\mathbf{I}_{(m-h)\times(m-h)}$ (used in the definition of $\mathbf{M}_{m\times m}$) are identity matrices, $\xi_{1\times h} = (1, 0, 0, ..., 0)$ is the initial distribution, $\mathbf{1}_{h\times 1} = (1, 1, ..., 1)^{\mathrm{T}}$ is the unit vector, m denotes the number of states in the state space Ω and $m-h$ is the number of unique simple patterns that defines a signal; the (non-essential) matrix $\mathbf{C}_{h\times(m-h)}$ will be illustrated later.

The point is that we only need to construct the state space Ω and the essential transition probability sub-matrix $\mathbf{Q}_{h\times h}$ of the Markov chain in order to be capable to calculate the entire run-length distribution.

Signaling probabilities

Whilst the key to construct the state space Ω depends on the particular signaling rule and whether a one-sided or two-sided chart is looked at, the building blocks of the transition probability matrix are the one-step transition probabilities (i.e. the elements of the transition probability matrix).

The one-step transition probabilities are denoted by

$$p_{k,j} = \Pr(Z_i = j \mid Z_{i-1} = k)$$

and interpreted as the conditional probability given that at any specific time $i-1$ the system was in state k, the system will be in state j at time i for $i \geq 1$ and $j, k \in \Omega$. The transition probabilities $p_{k,j}$ are all functions of and depend on the signaling probabilities i.e. the probability for a single charting statistic to plot outside the control limit(s), and therefore play a key role in the derivation of the run-length distributions of the runs-rules enhanced charts. In case of the upper and lower one-sided charts the signaling probabilities are

$$p^{+}(n, b, \theta) = \Pr(T_i \geq UCL) = \Pr(T_i \geq n-b) = \Pr(\xi_i^{+} = 1) = I_p(n-b, b+1) \qquad (4\text{-}11)$$

and

$$p^{-}(n, a, \theta) = \Pr(T_i \leq LCL) = \Pr(T_i \leq a) = \Pr(\xi_i^{-} = 1) = 1 - I_p(a+1, n-a), \qquad (4\text{-}12)$$

respectively; for the two-sided chart the probability for any of the charting statistics to plot outside either the UCL or the LCL is

$$p^{\pm}(n, a, b, \theta) = 1 - \Pr(LCL < T_i < UCL) = 1 - I_p(a+1, n-a) + I_p(n-b, b+1) \qquad (4\text{-}13)$$

where $I_p(u,v) = [\beta(u,v)]^{-1}\int_0^p w^{u-1}(1-w)^{v-1}dw$ with $0 < p = \Pr(X_{ij} > \theta_0) < 1$, is the c.d.f of the $Beta(u,v)$ distribution, also known as the incomplete beta function, which helps us write various expressions in a more compact form.

Note that, for notational simplicity and brevity we denote the probabilities $p^+(n,b,\theta)$, $p^-(n,a,\theta)$, and $p^\pm(n,a,b,\theta)$ simply by p^+, p^- and $p^\pm$, respectively.

Remark 4

(i) If $\theta = \theta_0$ the signaling probabilities (and hence the distribution of T_i and the in-control run-length distributions and their associated characteristics) depend only on

a. the sample size n,
b. the charting constants a and/or b, and
c. the percentile of interest $\theta_0 = F_X^{-1}(\pi)$ where π is specified.

Any decision rule (signaling rule) based on the T_i's will therefore be distribution-free as long as the underlying distributions (at each point in time) are continuous and identical. It follows that the in-control run-length distributions of the runs-rules enhanced sign charts are distribution-free and therefore charts based on the T_i's will be distribution-free.

(ii) To obtain the in-control run-length distribution and its mean and variance one substitutes $\theta = \theta_0$ in expressions (4-8), (4-9) and (4-10); by substituting $\theta \neq \theta_0$ one obtains the corresponding results for the out-of-control situation which depends on the underlying process distribution.

4.2.2 Transition probability matrices of the sign charts

To illustrate the derivation of the transition probability matrices and the run-length distributions, we begin with the one-sided upper (lower) control charts and then proceed to study the two-sided charts.

In case of the one-sided and the two-sided charts, we first look at the run-length distribution of the *1-of-1* chart (which uses the least complicated signaling rule) before we study the run-length distributions of the run rules enhanced charts, that is, the *k-of-k* ($k \geq 2$) and the *k-of-w* ($1 \leq k \leq w$ and $w \geq 2$) sign charts.

For each chart the key is to imbed the run-length into a proper homogenous Markov chain and obtain the essential transition probability sub-matrix $\mathbf{Q}_{h\times h}$ associated with the particular Markov chain.

Note that, we discuss the derivation of the transition probability matrices of the one-sided and the two-sided sign charts in detail so that later, in Case U, we can merely make use of these results.

4.2.2.1 One-sided sign charts

For the upper one-sided sign chart we view the series of signaling indicators $\xi_1^+, \xi_2^+, \xi_3^+, \ldots$ associated with the charting statistics $T_1, T_2, T_3, \ldots$ and the *UCL* as a series of independent binary random variables, each being either "a success or 1" (T_i plotting on or above the *UCL*) or "a failure or 0" (T_i plotting below the *UCL*) with probabilities p^+ and $1-p^+$, respectively (see e.g. Figure 4.9).

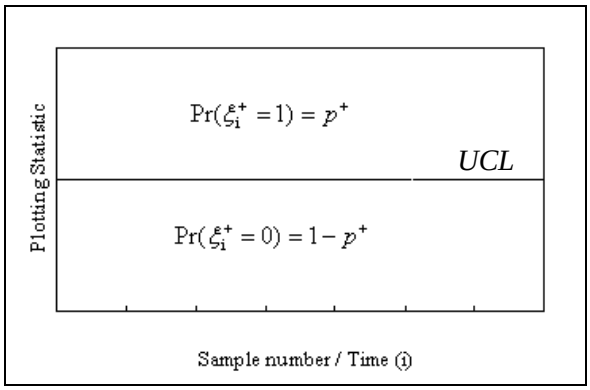


Figure 4.9: The two regions on the upper one-sided control chart ('0' and '1') and their associated probabilities used to classify the charting statistic

Upper one-sided *1-of-1* sign chart

The run-length N_{1of1}^{+} of the *1-of-1* upper one-sided chart is the waiting time until event A_1 (see panel (a) of Figure 4.1) occurs, which can be viewed as the waiting time for the first occurrence of a "1" (success) in the series of ξ_i^{+}'s (i.e. of 0's and 1's).

The imbedded Markov chain associated with N_{1of1}^{+} is a homogeneous Markov chain defined on the finite state space $\Omega = \{\phi, 0, 1\}$ with $m = 3$ states, where

(a) the state $\{1\}$ is called the "absorbing" state (when the process is declared out-of-control),

(b) the state $\{0\}$ is called the "transient" state (i.e. the process can remain in state $\{0\}$, which means that the process is IC and the charting procedure continues, or the process can move from state $\{0\}$ into state $\{1\}$, which implies that the process goes OOC and the charting procedure stops), and

(c) the state ϕ is the "dummy" state introduced for convenience. The dummy state ϕ is in fact added to the state space Ω so that with probability one the process is assumed to begin in-control with the intention that the corresponding initial probability distribution is taken as $\boldsymbol{\xi}_{1\times 2} = (1, 0)$.

The 3×3 transition probability matrix of $\{Z_i : i \geq 0\}$ associated with N_{1of1}^{+} is given by

$$\mathbf{M}_{3\times 3} = \left[\begin{array}{c|c} \mathbf{Q}_{2\times 2} & \mathbf{C}_{2\times 1} \\ \hline \mathbf{0}_{1\times 2} & \mathbf{I}_{1\times 1} \end{array}\right] = \left[\begin{array}{cc|c} p_{\phi,\phi} & p_{\phi,0} & p_{\phi,1} \\ p_{0,\phi} & p_{0,0} & p_{0,1} \\ \hline p_{1,\phi} & p_{1,0} & p_{1,1} \end{array}\right] = \left[\begin{array}{cc|c} 0 & 1-p^{+} & p^{+} \\ 0 & 1-p^{+} & p^{+} \\ \hline 0 & 0 & 1 \end{array}\right] \qquad (4\text{-}14)$$

where, for example, $p_{0,1}$ (the entry in the 2nd row and the 3rd column of $\mathbf{M}_{3\times 3}$) is the probability that the system goes from state $\{0\}$ (that is where T_{i-1} plots IC) at time $i-1$, to state $\{1\}$ (that is where T_i plot OOC) at time i; this probability is simply the probability that T_i plots at or above the *UCL* at time i, which is $\Pr(\xi_i^{+} = 1) = p^{+} = p^{+}(n, b, \theta)$. The rest of the elements of $\mathbf{M}_{3\times 3}$ in (4-14) can be calculated in a similar way.

Upper one-sided *2-of-2* sign chart

The run-length N^{+}_{2of2} of the *2-of-2* upper one-sided chart is the waiting time until event B_1 (see panel (a) of Figure 4.3) occurs, which can be viewed as the waiting time for the first occurrence of two consecutive 1's (i.e. successes) in the series of ξ_i^{+}'s (i.e. 0's and 1's).

The imbedded Markov chain associated with N^{+}_{2of2} is a homogeneous Markov chain defined on the finite state space $\Omega = \{\phi, 0, 1, 11\}$ with $m = 4$ states, where

(a) the last state $\{11\}$ is the absorbing state,

(b) the two states $\{0\}$ and $\{1\}$ are the transient (non-absorbent) states,

(c) and ϕ is the dummy state, which is again added to Ω so that (with probability one) the process is assumed to begin in-control and with the intention that the corresponding initial probability distribution is taken as $\boldsymbol{\xi}_{1\times3} = (1, 0, 0)$.

The transition probability matrix of $\{Z_i : i \geq 0\}$ associated with N^{+}_{2of2} is given by

$$\mathbf{M}_{4\times4} = \left[\begin{array}{c|c} \mathbf{Q}_{3\times3} & \mathbf{C}_{3\times1} \\ \hline \mathbf{0}_{1\times3} & \mathbf{I}_{1\times1} \end{array}\right] = \left[\begin{array}{ccc|c} p_{\phi,\phi} & p_{\phi,0} & p_{\phi,1} & p_{\phi,11} \\ p_{0,\phi} & p_{0,0} & p_{0,1} & p_{0,11} \\ p_{1,\phi} & p_{1,0} & p_{1,1} & p_{1,11} \\ \hline p_{11,\phi} & p_{11,0} & p_{11,1} & p_{11,11} \end{array}\right] = \left[\begin{array}{ccc|c} 0 & 1-p^{+} & p^{+} & 0 \\ 0 & 1-p^{+} & p^{+} & 0 \\ 0 & 1-p^{+} & 0 & p^{+} \\ \hline 0 & 0 & 0 & 1 \end{array}\right] \qquad (4\text{-}15)$$

where, for example, the probability that the system goes from state $\{1\}$ (where T_{i-1} plots OOC) at time $i-1$, to state $\{11\}$ (where both T_{i-1} and T_i plot OOC) at time i, denoted by $p_{1,11}$, is the entry in the 3rd row and the 4th column of $\mathbf{M}_{4\times4}$; this is simply the probability that T_i plots at or above the *UCL* at time *i*, which is, as earlier, $\Pr(\xi_i^{+} = 1) = p^{+} = p^{+}(n, b, \theta)$.

Remark 5

Because

(i) the signaling indicators ξ_i^+'s are a sequence of i.i.d. Bernoulli random variables each with probability of success p^+, and

(ii) the run-length random variable N_{1of1}^+ (N_{2of2}^+) is the waiting time for the first success (the first occurrence of two consecutive successes),

one can equivalently obtain the distribution (i.e. the p.m.f, the mean, the variance etc.) of N_{1of1}^+ (N_{2of2}^+) from the distribution of the variable T_k where $k=1$ (or 2).

The stopping time variable T_k ($k \geq 1$) is, in general, the waiting time to observe a sequence of k consecutive successes for the first time in a sequence of i.i.d. Bernoulli random variables with success probability α and should not be confused with the plotting statistic T_i defined in (4-5).

The distribution of T_k is known to be the geometric distribution of order k (see e.g. Chapter 2 of Balakrishnan and Koutras, (2002)) with p.m.f, expected value and variance given by

$$\Pr(T_k = j) = \begin{cases} 0 & \text{if} \quad 0 \leq j < k \\ \alpha^k & \text{if} \quad j = k \\ \sum_{i=1}^{[\frac{j+1}{k+1}]} (-1)^{i-1} \frac{\alpha^{ik}}{(1-\alpha)^{1-i}} \left\{ \binom{j-ik-1}{i-2} + (1-\alpha)\binom{j-ik-1}{i-1} \right\} & \text{if} \quad j \geq k+1 \end{cases} \tag{4-16}$$

$$E(T_k) = \frac{1-\alpha^k}{(1-\alpha)\alpha^k} \qquad \text{and} \qquad \text{var}(T_k) = \frac{1-(2k+1)(1-\alpha)\alpha^k - \alpha^{2k+1}}{(1-\alpha)^2 \alpha^{2k}} \tag{4-17}$$

respectively.

The equivalence between the distribution of T_k (i.e. the geometric distribution of order k) and the distribution of N^+_{kofk} (i.e. the waiting until for the first time k consecutive T_i's plot on or above the UCL) can be verified by substituting the essential transition probability sub-matrix $\mathbf{Q}_{h\times h}$ of $\{Z_i : i \geq 0\}$ associated with N^+_{kofk} in expressions (4-8), (4-9) and (4-10), and then simplifying symbolically (using, for example, computer software with matrix manipulations capabilities such as Scientific Workplace®); upon doing so one obtains explicit and closed form expressions, via the Markov chain approach, for the p.m.f, the ARL and the $VARL$ of the run-length random variable N^+_{kofk}.

For the upper one-sided *1-of-1* sign chart, for example, we substitute

$$\mathbf{Q}_{2\times 2} = \begin{bmatrix} 0 & 1-p^+ \\ 0 & 1-p^+ \end{bmatrix}$$

in expressions (4-8), (4-9) and (4-10) so that upon simplifying we obtain

$$\Pr(N^+_{1of1} = j \mid n,b,\theta) = (1-p^+)^{j-1}\,p^+ \quad \text{for} \quad j = 1,2,3,... \tag{4-18}$$

$$E(N^+_{1of1} \mid n,b,\theta) = 1/p^+ \quad \text{and} \quad \text{var}(N^+_{1of1} \mid n,b,\theta) = (1-p^+)/(p^+)^2. \tag{4-19}$$

Expressions (4-18) and (4-19) are identical to expressions (4-16) and (4-17) with $k = 1$ and $\alpha = p^+$, respectively.

Upper one-sided *2-of-3* sign chart

Like the upper one-sided *1-of-1* and *2-of-2* charts we can find the run-length distribution of the *2-of-3* upper one-sided chart using a Markov chain.

The run-length N^{+}_{2of3} of the *2-of-3* upper one-sided chart is the waiting time until event C_1 or C_2 (see panel (a) of Figure 4.4) occurs, which can be viewed as the waiting time for the first occurrence of the pattern $\Lambda = \{011 \text{ or } 101\}$ in the series of ξ_i^{+}'s (i.e. of 0's and 1's). The pattern Λ is called a "compound pattern" and written as: $\Lambda = \Lambda_1 \cup \Lambda_2$ where $\Lambda_1 = 011$ and $\Lambda_2 = 101$ are two so-called distinct "simple patterns".

The imbedded Markov chain associated with N^{+}_{2of3} is a homogeneous Markov chain defined on the state space $\Omega = \{\phi, 0, 1, 01, 10, \alpha_1, \alpha_2\}$ with $m = 7$ states, where

(a) the two states $\alpha_1 = \{011\}$ and $\alpha_2 = \{101\}$ are the absorbing states (when the process is declared OOC),

(b) the four states $\{0, 1, 01, 10\}$ are the transient states (i.e. the process can move from one of these states to another, which means that the charting procedure continues), and

(c) ϕ is the dummy state introduced for convenience.

The transient states are the sequential sub-patterns of $\Lambda_1 = 011$ and $\Lambda_2 = 101$, respectively. For example, the state $\{0\}$ is the sub-pattern of the state $\{01\}$, whereas the two states $\{0\}$ and $\{01\}$ are the sub-patterns of $\Lambda_1 = 011$, and the states $\{1\}$ and $\{10\}$ are the sub-patterns of $\Lambda_2 = 101$.

As earlier, the dummy state ϕ is again added to Ω so that (with probability one) the process is assumed to begin in-control, and the corresponding initial probability distribution is taken as $\xi_{1\times5} = (1, 0, 0, 0, 0)$.

The transition probability matrix of $\{\xi_i^+, i \geq 0\}$ associated with N_{2of3}^+ is given by

$$\mathbf{M}_{7\times7} = \left[\begin{array}{c:c} \mathbf{Q}_{5\times5} & \mathbf{C}_{5\times2} \\ \hdashline \mathbf{0}_{2\times5} & \mathbf{I}_{2\times2} \end{array}\right] = \left[\begin{array}{ccccc:cc} 0 & 1-p^+ & p^+ & 0 & 0 & 0 & 0 \\ 0 & 1-p^+ & 0 & p^+ & 0 & 0 & 0 \\ 0 & 0 & p^+ & 0 & 1-p^+ & 0 & 0 \\ 0 & 0 & 0 & 0 & 1-p^+ & p^+ & 0 \\ 0 & 1-p^+ & 0 & 0 & 0 & 0 & p^+ \\ \hdashline 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{array}\right] \quad \text{(4-20)}$$

where, for example, the probability that the system goes from state {01} (that is where T_{i-2} plots IC and T_{i-1} plots OOC) at time $i-1$, to state $\alpha_1 = \{011\}$ (that is where T_{i-2} plots IC and both T_{i-1} and T_i plot OOC) at time i is the entry in the 4th row and the 6th column of $\mathbf{M}_{7\times7}$.

Remark 6

A few comments concerning the application and the implementation of the upper one-sided *2-of-3* sign chart are in order:

(i) To declare a process out-of-control (OOC) we need at least three charting statistics (T_{i-2}, T_{i-1} and T_i, say) of which exactly one should plot in-control (IC) i.e. either T_{i-2} plots below the UCL with T_{i-1} and T_i plotting on or above the UCL or, T_{i-1} plots below the UCL with T_{i-2} and T_i plotting on or above the UCL (see e.g. events C_1 and C_2 in panel (a) of Figure 4.4). Thus, we can only declare the process OOC beginning from time $i \geq 3$ and, we need at least one charting statistic to plot below the upper control limit before we can declare the process OOC.

(ii) Because of these two build-in conditions of the upper one-sided *2-of-3* sign chart, the chart has a hitch at start-up: If $T_i \geq UCL$ for $i = 1,2,3,...,r$, that is, if all the charting statistics plot on or above the upper control limit from the time that the chart is implemented until time r, the chart would not immediately signal that the process is OOC even though the pattern of the points on the chart suggests otherwise. The chart would most likely give a "delayed" or a "late" OOC signal instead.

While this glitch is possible, we need to stress an important assumption:

> The design and the implementation of all the charts that are proposed in this chapter are based on an IC process at start-up as well as the trade-off between minimizing the probability that a charting statistic plots on or outside the control limit(s) when the process is actually IC and quickly detecting an OOC process.

This assumption means two things:

a. The process is IC at start-up; hence, we must ensure (to the extent that it is possible) that the process is IC *before* we start monitoring it.

b. We typically choose the UCL such that the probability that a T_i plots on or above the UCL when the process is IC i.e. $p_0^+ = \Pr(T_i \geq UCL \mid IC)$, is small, which automatically implies that the probability that a T_i plots below the UCL when the process is IC i.e. $1 - p_0^+ = \Pr(T_i < UCL \mid IC)$, is large.

The latter implies that the probability that all the charting statistics up to and including the r^{th} one plot on or above the *UCL* when the process is IC i.e. $(p_0^+)^r$, would decrease rapidly as r increases. But, most importantly, it also implies that as we continue to monitor the process, the probability that the r^{th} charting statistic plots below the *UCL* when the process is IC stays constant and equal to $1-p_0^+$; this is so because we assume that successive samples (or charting statistics) are independent.

Hence, what is of importance to the practitioner is to know what the risk is that this hitch occurs. This risk can be measured by calculating and studying the *odds* that a T_i plots below the *UCL* when the process is IC i.e. $(1-p_0^+)/(p_0^+)$, at any time $i=1,2,3,\ldots$.

To investigate the effect of p_0^+ on the above *odds*, Table 4.1 shows values of $(1-p_0^+)/(p_0^+)$ for values of $p_0^+ = 0.001(0.001)0.005$ and $0.01(0.01)0.20$. The values of p_0^+ that we use to construct Table 4.1 are representative of the typical values that one would consider when designing the proposed upper one-sided *2-of-3* chart (see e.g. Tables 4.6. and 4.7).

From Table 4.1 we observe that:

(i) The ratio $(1-p_0^+)/(p_0^+)$ is larger than or equal to 4.0 for all values of p_0^+ that we consider. This implies that, for a process that is IC at start-up (which is a fundamental assumption of our earlier theoretical developments and the reason for adding the dummy state, ϕ, to the state spaces off all the proposed charts) it is *at least* four times more likely for any new incoming T_i to plot below the *UCL* than for any new incoming T_i to plot on or above the *UCL*.

(ii) For $p_0^+ = 0.01$, which is a very reasonable choice considering all the values of p_0^+ in Tables 4.6 and 4.7, the ratio $(1-p_0^+)/(p_0^+)$ is equal to 99.0; this is relatively large.

(iii) The largest value for $(1-p_0^+)/(p_0^+)$ is 999.0 (when $p_0^+ = 0.001$) and will increase even further as p_0^+ decreases; this is good because smaller values of p_0^+ are typically preferred and also recommended in practice.

The above-mentioned observations are all relevant for the practitioner because, for a process that is IC at start-up, (which is a key assumption when implementing any of the charts that are proposed in this chapter) they show that the risk associated with the proposed *2-of-3* sign chart at start-up is: (a) almost negligible, and (b) decreases rapidly as we continue to monitor the process because the probability that all the charting statistics up to and including the r^{th} one plot on or above the *UCL* when the process is IC i.e. $(p_0^+)^r$, would decrease towards zero quickly as r increases. This should be reassuring for the practitioner.

Table 4.1: The ratio $(1-p_0^+)/(p_0^+)$ as a function of p_0^+

p_0^+	$(1-p_0^+)/(p_0^+)$	p_0^+	$(1-p_0^+)/(p_0^+)$
0.001	999.0	**0.09**	10.1
0.002	499.0	**0.10**	9.0
0.003	332.3	**0.11**	8.1
0.004	249.0	**0.12**	7.3
0.005	199.0	**0.13**	6.7
0.01	99.0	**0.14**	6.1
0.02	49.0	**0.15**	5.7
0.03	32.3	**0.16**	5.3
0.04	24.0	**0.17**	4.9
0.05	19.0	**0.18**	4.6
0.06	15.7	**0.19**	4.3
0.07	13.3	**0.20**	4.0
0.08	11.5		

To overcome the imperfection of the upper one-sided *2-of-3* sign chart at start-up, we could use the event $C_7 = \{T_1 \geq UCL, T_2 \geq UCL\} \Leftrightarrow \{\xi_1^+ = 1, \xi_2^+ = 1\}$, in addition to the events C_1 and C_2 shown in panel (a) of Figure 4.4, as a third signaling event. The event C_7 is special in two ways: (a) it prevents the hitch at start-up by enabling the chart to signal at time $i = 2$, and (b) it occurs if and only if the first two charting statistics, T_1 and T_2, both plot on or above the upper control limit; hence, event C_7 can not occur from time $i \geq 3$.

The resultant chart is an augmented upper one-sided *2-of-3* sign chart. Adding the extra event leads to an augmented state space i.e. $\Omega = \{\phi, 0, 1, 01, 10, 11, 011, 101\}$, where the three states $\{11\}$, $\{011\}$ and $\{101\}$ are the absorbent states and implies that the transition probability matrix in (4-20) be altered slightly to become

$$\mathbf{M}_{8\times 8} = \left[\begin{array}{c|c} \mathbf{Q}_{5\times 5} & \mathbf{C}_{5\times 3} \\ \hline \mathbf{0}_{3\times 5} & \mathbf{I}_{3\times 3} \end{array}\right] = \left[\begin{array}{ccccc|ccc} 0 & 1-p^+ & p^+ & 0 & 0 & 0 & 0 & 0 \\ 0 & 1-p^+ & 0 & p^+ & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1-p^+ & p^+ & 0 & 0 \\ 0 & 0 & 0 & 0 & 1-p^+ & 0 & p^+ & 0 \\ 0 & 1-p^+ & 0 & 0 & 0 & 0 & 0 & p^+ \\ \hline 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{array}\right].$$

To investigate the impact (i.e. gain or loss) of augmenting the transition probability matrix on the in-control performance of the chart, we calculated the in-control average run-lengths and the false alarm rates of the proposed upper one-sided *2-of-3* sign chart and that of the augmented upper one-sided *2-of-3* sign chart (when it is of interest to monitor the median of the process) for different combinations of the sample size, n, and the upper control limit, UCL.

The values of the in-control average run-lengths (denoted by ARL^{+}_{2of3} and ARL^{A+}_{2of3}, respectively) were calculated according to expression (4-9) using the transition probability matrix in (4-20) and the augmented transition probability matrix given above, respectively.

The false alarm rate of the proposed upper one-sided *2-of-3* sign chart (denoted FAR^{+}_{2of3}) and that of the augmented upper one-sided *2-of-3* sign chart (denoted FAR^{A+}_{2of3}) can be easily obtained from the definitions of the signaling events that are used by each chart and are given by

$$FAR^{+}_{2of3} = \begin{cases} 0 & \text{if} \quad i = 1 \text{ or } 2 \\ 2(1-p_0^+)(p_0^+)^2 & \text{if} \quad i \geq 3 \end{cases} \quad \text{and} \quad FAR^{A+}_{2of3} = \begin{cases} 0 & \text{if} \quad i = 1 \\ (p_0^+)^2 & \text{if} \quad i = 2 \\ 2(1-p_0^+)(p_0^+)^2 & \text{if} \quad i \geq 3 \end{cases}$$

respectively.

There is only a slight modification of the expression for FAR^{+}_{2of3} to obtain FAR^{A+}_{2of3}; this leads to the following similarities and/or differences in the false alarm rates of the charts:

(i) At time $i = 1$: $FAR^{+}_{2of3} = FAR^{A+}_{2of3} = 0$,

(ii) At time $i = 2$: $FAR^{+}_{2of3} = 0$ but $FAR^{A+}_{2of3} = (p_0^+)^2$, and

(iii) At time $i \geq 3$: $FAR^{+}_{2of3} = FAR^{A+}_{2of3} = 2(1-p_0^+)(p_0^+)^2$.

These similarities and/or differences are a direct consequence of the signaling events used by each chart i.e.

(i) Neither one of the charts can signal at time $i=1$ because, the proposed upper one-sided *2-of-3* sign chart needs at least three charting statistics to signal whereas the augmented upper one-sided *2-of-3* sign chart needs at least two charting statistics to signal.

(ii) It is only the augmented upper one-sided *2-of-3* sign chart that can give a false alarm at time $i=2$ and, it can do so if and only if event C_7 occurs.

(iii) From time $i \geq 3$, both the charts can signal if and only if event C_1 or event C_2 occurs. The event C_7, as mentioned earlier, can only occur at time $i=2$ and therefore does not influence the false alarm rate of the augmented upper one-sided *2-of-3* sign chart at or beyond time $i=3$.

Based on our calculations, we found that:

(i) The in-control average run-lengths of the two charts were almost identical; ARL^{+}_{2of3} is only slightly larger than ARL^{A+}_{2of3}.

(ii) Depending on the combination of n and UCL, the $FAR^{A+}_{2of3}=(p_0^{+})^2$ at time $i=2$ can be reasonably large, which might be a concern for the practitioner.

To further compare the impact of augmenting the proposed upper one-sided *2-of-3* sign chart, Table 4.2 shows values of the in-control probability mass functions (p.m.f's) and the in-control cumulative distribution functions (c.d.f's) of the run-length random variables, N^{+}_{2of3} and N^{A+}_{2of3}, associated with the proposed and the augmented charts; these values are denoted by $\Pr(N^{+}_{2of3}=i \mid IC)$, $\Pr(N^{A+}_{2of3}=i \mid IC)$, $\Pr(N^{+}_{2of3} \leq i \mid IC)$ and $\Pr(N^{A+}_{2of3} \leq i \mid IC)$, respectively and are calculated for values of $i=1,2,...,15$.

The calculations in Table 4.2 assume that we monitor the process median using samples of size $n = 5$ (which is a very popular choice in practice) and that the upper control limit is $UCL = 5$. For this particular combination of n and UCL, it is calculated that $p_0^+ = \Pr(T_i \geq 5 \mid IC) = 0.03125$ where $T_i \sim Bin(5, 0.5)$, and it was found that ($ARL_{2of3}^{+} = 552.65$; $FAR_{2of3}^{+} = 0.00189$ for $i \geq 3$) while ($ARL_{2of3}^{A+} = 552.13$; $FAR_{2of3}^{A+} = 0.00098$ at $i = 2$ and $FAR_{2of3}^{A+} = 0.00189$ for $i \geq 3$).

From Table 4.2 we see that there are two key differences with respect to the in-control characteristics and the in-control performance of the charts:

(i) The augmented upper one-sided *2-of-3* sign chart can signal incorrectly (with probability 0.00098) after having observed only two charting statistics whereas the proposed upper one-sided *2-of-3* sign chart cannot.

(ii) The ratio $\Pr(N_{2of3}^{A+} \leq i \mid IC) / \Pr(N_{2of3}^{+} \leq i \mid IC)$, decreases to 1 as i increases; this observation is supported by the fact that ARL_{2of3}^{A+} is only slightly less than ARL_{2of3}^{+} i.e. $ARL_{2of3}^{A+} / ARL_{2of3}^{+} \approx 1$. These observations imply that, from start-up (when the process is IC) the augmented upper one-sided *2-of-3* sign chart always has a higher cumulative probability for a shorter run-length than the proposed upper one-sided *2-of-3* sign chart..

Table 4.2: The in-control probability mass functions (p.m.f's) and the in-control cumulative distribution functions (c.d.f's) of the proposed *2-of-3* and the augmented *2-of-3* sign charts when $n = 5$ and $UCL = 5$

	***2-of-3* sign chart**		**Augmented *2-of-3* sign chart**	
i	$\Pr(N^{+}_{2of3} = i \mid IC)$	$\Pr(N^{+}_{2of3} \leq i \mid IC)$	$\Pr(N^{A+}_{2of3} = i \mid IC)$	$\Pr(N^{A+}_{2of3} \leq i \mid IC)$
1	0	0	0	0
2	0	0	0.00098	0.00098
3	0.00189	0.00189	0.00189	0.00287
4	0.00186	0.00375	0.00183	0.00470
5	0.00181	0.00556	0.00180	0.00651
6	0.00180	0.00736	0.00180	0.00831
7	0.00180	0.00917	0.00180	0.01011
8	0.00180	0.01097	0.00180	0.01191
9	0.00180	0.01276	0.00179	0.01370
10	0.00179	0.01455	0.00179	0.01549
11	0.00179	0.01634	0.00179	0.01728
12	0.00179	0.01813	0.00178	0.01906
13	0.00178	0.01991	0.00178	0.02085
14	0.00178	0.02169	0.00178	0.02262
15	0.00178	0.02347	0.00177	0.02440

To summarize the above discussion and our findings based on the analysis, we can state that:

(i) The proposed upper one-sided *2-of-3* sign chart has a hitch at start-up but, the odds that this problem occurs are typically small; this should be reassuring for the practitioner.

(ii) It is possible to fix the imperfection of the proposed upper one-sided *2-of-3* sign chart by adding a third signaling event but, even this modification has a drawback: the performance of the augmented chart is degraded at start-up i.e. its false alarm rate is nonzero at time $i = 2$ (unlike the proposed chart) and the cumulative probability for a shorter run-length is higher than that of the proposed chart.

(iii) Neither the proposed nor the augmented upper one-sided *2-of-3* sign chart can be implemented without taking a risk i.e. there is a trade-off between having a hitch at start-up and the possibility of a false alarm at time $i = 2$.

(iv) The inherent risk of each chart cannot be completely eliminated but, these risks can be minimized (or at least reduced) by ensuring that the process is IC at start-up and/or by choosing p_0^+ to be small.

(v) The in-control performance of the charts are almost identical: there is only a bit of a difference in their in-control ARL's and, at time $i = 2$ we have that $FAR_{2of3}^{A+} = (p_0^+)^2$ whereas $FAR_{2of3}^{+} = 0$.

(vi) If a shift/change in the process occurs *after* start-up i.e. from time $i \geq 3$, both the charts can signal only on the occurrence of events C_1 or C_2. So, the OOC performance of these two charts would be the same.

We recommend that practitioners use either the proposed or the augmented upper one-sided *2-of-3* sign chart but, we suggest that they familiarize themselves with the inherent risk associated with the selected chart. If the practitioner is not willing to accept the risk(s) associated with the *2-of-3* charts, he/she should use another chart e.g. the new proposed upper one-sided *2-of-2* sign chart or the original upper one-sided *1-of-1* sign chart.

Based on the above analysis and the fact that the augmented chart can signal after having observed only two charting statistics instead of the proposed three charting statistics (which implies that the augmented chart is not a "true" *2-of-3* chart), it was decided to focus on the proposed upper one-sided *2-of-3* sign chart and not to investigate the statistical properties of the augmented upper one-sided *2-of-3* sign chart any further in this thesis.

Furthermore, although the above discussion focussed specifically on the *upper* one-sided *2-of-3* sign chart, these comments also apply to the *lower* one-sided *2-of-3* chart and the *two-sided 2-of-3* chart. In fact, these comments are relevant for any one-sided or two-sided *k-of-w* chart whenever $k < w$. This is so, because we need at least w charting statistics before we can declare the process OOC and we need at least $w - k$ charting statistics to plot IC (i.e. below the UCL or, above the LCL or, between the LCL and UCL, depending on the chart that is used).

Remark 7

If the upper one-sided *2-of-3* sign chart were to signal upon any one of the three events in which two of the last three charting statistics can plot on or above the *UCL* i.e. the occurrence of either event C_1 or C_2 or C_5 (see e.g. panel (a) of Figure 4.4 and panel (a) of Figure 4.5), the p.m.f as well as the mean (ARL) and the variance ($VARL$) of N^{+}_{2of3} would be obtainable from the distribution and the associated characteristics of the random variable $T_k^{(w)}$.

The random variable $T_k^{(w)}$ is the waiting time for the first occurrence of a scan or run of type k/w, where the term scan or generalized run of type k/w refers to sub sequences $\xi_i^{+}, \xi_{i+1}^{+}, ..., \xi_{i+j-1}^{+}$ of length $j \le w$ such that the number of successes contained therein is at least k, that is, $\sum_{s=i}^{i+j-1} \xi_s^{+} \ge k$ (see e.g. Chapter 9 of Balakrishnan and Koutras, (2002)); the probability distribution of $T_k^{(w)}$ is known as the geometric distribution of order k/w and derived via combinatorial methods.

Because we exclude event C_5 as a signaling event in case of the upper one-sided *2-of-3* chart (because the possibility of declaring a process out-of-control when the first and second charting statistic plot OOC but the third one plots IC is undesirable in practice), we cannot make use of the p.m.f or the associated properties of the geometric distribution of order k/w, that is, $T_k^{(w)}$; this supports the statement in the beginning of section 4.2.1 that the Markov chain technique has a great advantage over the classical combinatory techniques for finding the distribution(s) of run-related problems.

Example 1

Consider the upper one-sided *2-of-3* sign chart for monitoring the median $\theta = F_X^{-1}(\pi = 0.5)$ and suppose that the subgroup size $n = 5$, the charting constant $b = 0$ (so that $UCL = 5$) and θ_0 denotes the target (IC) value for the median.

As noted earlier, when the process is in-control $p_0 = \Pr(X_{ij} > \theta_0 \mid IC) = 0.5$ and therefore

$$p_0^+ = p_0^+(n = 5, b = 0, \theta = \theta_0) = I_{p_0=0.5}(5,1) = 0.03125.$$

Substituting $p_0^+ = 0.03125$ for p^+ in (4-20) and using (4-9) we get

$$ARL_0 = E(N^+_{2of3,0} \mid n = 5, b = 0, \theta = \theta_0) = 552.65.$$

Similarly, using (4-10), the $SDRL_0 = 550.218$.

The in-control c.d.f of N^+_{2of3} can be obtained using the p.m.f in (4-8) and is given by

$$\Pr(N^+_{2of3,0} \le j) = \sum_{i=1}^{j} \boldsymbol{\xi} \mathbf{Q}^{i-1}_{5\times5,0}(\mathbf{I} - \mathbf{Q}_{5\times5,0})\mathbf{1}$$

where $N^+_{2of3,0}$ denotes the in-control run-length random variable and $\mathbf{Q}_{5\times5,0}$ is found from (4-20) by substituting p_0^+ for p^+. For illustration, we calculate and show the in-control p.m.f and the in-control c.d.f values for $j = 1,2,3,4,5,6$ in Table 4.3.

Table 4.3: The in-control probability mass function (p.m.f) and the in-control cumulative distribution function (c.d.f) for the upper one-sided *2-of-3* sign chart

j	**1**	**2**	**3**	**4**	**5**	**6**
$\Pr(N^+_{2of3,0} = j)$	0	0	0.00189	0.00186	0.00181	0.00180
$\Pr(N^+_{2of3,0} \le j)$	0	0	0.00189	0.00375	0.00556	0.00736

Given the c.d.f we can find the $100\pi^{\text{th}}$ percentile of the run-length distribution, which is the smallest integer j so that $\Pr(N^+_{2of3,0} \le j) \ge \pi$. For example, the second quartile (the median run-length, denoted $MDRL$) is found to be $Q_2 = 384$. The percentiles provide useful information regarding the efficacy of the control chart in addition to the moments such as the ARL_0 and the $SDRL_0$.

Lower one-sided *1-of-1*, *2-of-2* and *2-of-3* sign charts

By substituting $p^- = p^-(n,a,\theta)$, which is defined in (4-12), for $p^+ = p^+(n,b,\theta)$, which is defined in (4-11), in the transition probability matrices of (4-14), (4-15) and (4-20), the distributions of the run-length random variables N^-_{1of1}, N^-_{2of2} and N^-_{2of3} of the lower (negative) one-sided *1-of-1*, *2-of-2* and *2-of-3* charts, respectively can be straightforwardly obtained. This is so because each lower one-sided chart is a mirror image of the corresponding upper one-sided chart.

Also, note that, when we monitor the median, the in-control distribution of the plotting statistic is symmetric i.e. $T_i \sim Bin(n, 0.5)$. In this case, it makes practical sense to use symmetrically placed control limits and set $b = a$ so that $UCL = n - a$ and $LCL = a$; this implies that the control limits are equidistant from both ends. For this specific choice of the control limits we have that $\xi_i^+ \overset{d}{=} \xi_i^-$ i.e. the signaling indicators used to define the upper one-sided charts have the same distribution as the signaling indicators used to define the lower one-sided charts, and implies that the in-control performance of the lower and the upper one-sided sign charts, for monitoring the median, are identical. The performance of the upper and the lower one-sided sign charts will be further discussed in section 4.2.4 when we study their design.

Lastly, note that, the distributions of N^-_{1of1} and N^-_{2of2} can also be obtained from those of T_k (see Remark 5) by setting $\alpha = p^-$ and substituting $k = 1$ or $k = 2$ in (4-16) and (4-17), respectively.

Two-sided sign charts

The derivation of the transition probability matrices and the run-length distributions of the two-sided charts (via the Markov chain approach) parallel those of the one-sided charts.

For the two-sided charts, the signaling indicators ξ_i's are defined by a series of values 0, 1 or 2, depending on whether the corresponding charting statistic T_i plots between the two control limits, on or above the *UCL*, or on or below the *LCL*, respectively; the probabilities for these three events are $1-p^{\pm}=1-p^{+}-p^{-}$, p^{+} and p^{-}, respectively (see e.g. Figure 4.10 below).

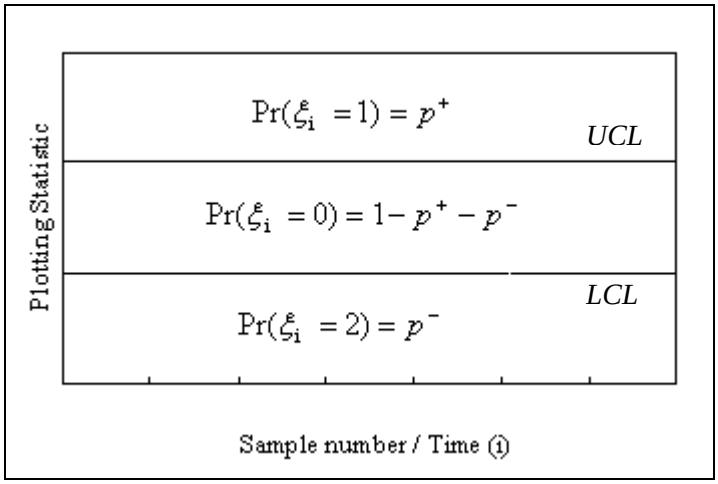


Figure 4.10: The three regions on the two-sided control chart ('0' , '1' and '2') and their associated probabilities used to classify the charting statistic

Two-sided *1-of-1* sign chart

For the *1-of-1* two-sided chart the run-length N_{1of1} is the waiting time for the first occurrence of the event $A = A_1 \cup A_2$ (see e.g. Figure 4.2), which is the first occurrence of the compound pattern $\Lambda = \Lambda_1 \cup \Lambda_2$ in the series of ξ_i's (i.e. among the 0's, 1's and 2's) where $\Lambda_1 = 1$ and $\Lambda_2 = 2$ are two distinct simple patterns in this situation.

The state space for the imbedded Markov chain associated with the variable N_{1of1} is $\Omega = \{\phi, 0, 1, 2\}$, which has $m = 4$ states. The absorbing states are $\{1\}$ and $\{2\}$ whereas $\{0\}$ is the transient state and $\{\phi\}$ is the dummy state.

The transition probability matrix $\mathbf{M}_{4\times4}$ is given by

$$\mathbf{M}_{4\times4} = \left[\begin{array}{c|c} \mathbf{Q}_{2\times2} & \mathbf{C}_{2\times2} \\ \hline \mathbf{0}_{2\times2} & \mathbf{I}_{2\times2} \end{array}\right] = \left[\begin{array}{cc|cc} p_{\phi,\phi} & p_{\phi,0} & p_{\phi,1} & p_{\phi,2} \\ p_{0,\phi} & p_{0,0} & p_{0,1} & p_{0,2} \\ \hline p_{1,\phi} & p_{1,0} & p_{1,1} & p_{1,2} \\ p_{2,\phi} & p_{2,0} & p_{2,1} & p_{2,2} \end{array}\right] = \left[\begin{array}{cc|cc} 0 & 1-p^+-p^- & p^+ & p^- \\ 0 & 1-p^+-p^- & p^+ & p^- \\ \hline 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right] \qquad (4\text{-}21)$$

where, for example, the entry in the 2nd row and 2nd column of $\mathbf{M}_{4\times4}$, denoted by $p_{0,0}$, is the probability that the system remains in state $\{0\}$, that is, where T_{i-1} plots IC at time $i-1$ and T_i plots IC at time i.

Two-sided *2-of-2* DR sign chart

The run-length N_{2of2}^{DR} of the *2-of-2* DR two-sided chart is the waiting time for the first occurrence of the event D_1 or D_2 or D_3 or D_4 (see e.g. Figure 4.6), which is the first occurrence of the compound pattern $\Lambda = \bigcup_{i=1}^{4} \Lambda_i$ in the series of ξ_i's, where $\Lambda_1 = 11$, $\Lambda_2 = 22$, $\Lambda_3 = 12$ and $\Lambda_4 = 21$ are the four distinct simple patterns.

The imbedded Markov chain, in this case, is defined on the state space $\Omega = \{\phi, 0, 1, 2, 11, 22, 12, 21\}$, where ϕ is the dummy state, the three states $\{0\}$, $\{1\}$ and $\{2\}$ are the transient states and the four states $\{11\}, \{22\}$, $\{12\}$ and $\{21\}$ are the absorbing states.

The transition probability matrix $\mathbf{M}_{8\times8}$ is given by

$$\mathbf{M}_{8\times8} = \left[\begin{array}{c:c} \mathbf{Q}_{4\times4} & \mathbf{C}_{4\times4} \\ \hdashline \mathbf{0}_{4\times4} & \mathbf{I}_{4\times4} \end{array}\right] = \left[\begin{array}{cccc:cccc} 0 & 1-p^+-p^- & p^+ & p^- & 0 & 0 & 0 & 0 \\ 0 & 1-p^+-p^- & p^+ & p^- & 0 & 0 & 0 & 0 \\ 0 & 1-p^+-p^- & 0 & 0 & p^+ & 0 & p^- & 0 \\ 0 & 1-p^+-p^- & 0 & 0 & 0 & p^- & 0 & p^+ \\ \hdashline 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{array}\right]. \qquad (4\text{-}22)$$

Remark 8

If we let

$$\xi_i^{\pm} = I(T_i \notin (LCL,UCL)) = \begin{cases} 1 & \text{if} \quad T_i \notin (LCL,UCL) \\ 0 & \text{if} \quad T_i \in (LCL,UCL) \end{cases}$$

where $I(T_i \notin (LCL,UCL))$ denotes the indicator function for the event $\{T_i \notin (LCL,UCL)\}$ then

$$\{\xi_i^{\pm} = 1\} \text{ if and only if } \{\xi_i = 1\} \cup \{\xi_i = 2\} \tag{4-23}$$

so that

$$\Pr(\xi_i^{\pm} = 1) = \Pr(\xi_i = 1) + \Pr(\xi_i = 2) \tag{4-24}$$

where the $\xi_i^{\pm}$'s is a sequence of i.i.d. Bernoulli random variables each with probability of success $\Pr(\xi_i^{\pm} = 1) = \Pr(T_i \notin (LCL,UCL)) = p^{\pm} = p^{+} + p^{-}$ and the ξ_i's is a sequence of i.i.d tri-variate random variables with probabilities $\Pr(\xi_i = 1) = p^{+}$, $\Pr(\xi_i = 0) = 1 - p^{\pm} = 1 - p^{+} - p^{-}$ and $\Pr(\xi_i = 2) = p^{-}$, respectively.

Expressions (4-23) and (4-24) permit us to define the signaling events and obtain the run-length distributions of the two-sided *1-of-1* and the two-sided *2-of-2* DR sign charts using the $\xi_i^{\pm}$'s instead of using the ξ_i's. This means that, instead of using the Markov chain approach, we can find the distributions of N_{1of1} and N_{2of2}^{DR} using the results (or properties) of the geometric distribution of order k.

In particular, it follows from (4-23) and (4-24) that the run-length N_{1of1} of the two-sided *1-of-1* chart, which is the waiting time for the first occurrence of the event $A = A_1 \cup A_2 = \{\xi_i = 1\} \cup \{\xi_i = 2\}$, is equivalent to the waiting time for the first success (i.e. 1) among the $\xi_i^{\pm}$'s, that is, $A = \{\xi_i^{\pm} = 1\}$. Likewise, the run-length N_{2of2}^{DR} of the two-sided 2-*of-2* chart, which is the waiting time for the first occurrence of the event D_1 or D_2 or D_3 or D_4 (see e.g. Figure 4.6), is the same as the waiting time for the first occurrence of two consecutive successes (two successive 1's) among the $\xi_i^{\pm}$'s, that is, $D(DR) = \{\xi_{i-1}^{\pm} = \xi_i^{\pm} = 1\}$ so that

$$\Pr(D(DR)) = \Pr(D_1 \cup D_2 \cup D_3 \cup D_4) = \Pr(\{\xi_{i-1}^{\pm} = \xi_i^{\pm} = 1\}) = (p^{\pm})^2.$$

The distributions of N_{1of1} and N^{DR}_{2of2} are therefore both geometric distributions of order k so that closed form expressions for the p.m.f's of N_{1of1} and N^{DR}_{2of2} can be conveniently obtained from (4-16) by setting

$$k=1 \quad \text{with} \quad \alpha = \Pr(A) = p^{\pm}$$

and

$$k=2 \quad \text{with} \quad \alpha = \Pr(D(DR)) = (p^{\pm})^2$$

instead of symbolically simplifying expression (4-8).

For example, upon substituting the essential transition probability sub-matrix of the two-sided *1-of-1* sign chart

$$\mathbf{Q}_{2\times 2} = \begin{bmatrix} 0 & 1-p^{+}-p^{-} \\ 0 & 1-p^{+}-p^{-} \end{bmatrix} = \begin{bmatrix} 0 & 1-p^{\pm} \\ 0 & 1-p^{\pm} \end{bmatrix}$$

into (4-8) and simplifying symbolically, we get an explicit formula for the p.m.f of N_{1of1} (via the Markov chain approach) that corresponds to the already available p.m.f one obtains after substituting $p^{\pm}$ and $k=1$ into (4-16) i.e.

$$\Pr(N_{1of1}=j \mid n,a,b,\theta) = \Pr(T_{k=1}=j \mid n,a,b,\theta) = (1-p^{\pm})^{j-1}p^{\pm} \quad \text{for} \quad j=1,2,3,\ldots. \qquad (4\text{-}25)$$

Two-sided *2-of-2* KL sign chart

The run-length N_{2of2}^{KL} of the *2-of-2* KL two-sided chart is the waiting time for the first occurrence of the event D_1 or D_2 (see e.g. panels (a) and (b) of Figure 4.6), which is the first occurrence of the compound pattern $\Lambda = \Lambda_1 \cup \Lambda_2$ in the series of ξ_i's, where $\Lambda_1 = 11$ and $\Lambda_2 = 22$ are the two distinct simple patterns in this case.

The imbedded Markov chain associated with the run-length variable N_{2of2}^{KL} is defined on the state space $\Omega = \{\phi, 0, 1, 2, 11, 22\}$, which has $m = 6$ states, where $\{11\}$ and $\{22\}$ are the two absorbing states.

The transition probability matrix $\mathbf{M}_{6\times6}$ of the Markov chain is given by

$$\mathbf{M}_{6\times6} = \left[\begin{array}{c|c} \mathbf{Q}_{4\times4} & \mathbf{C}_{4\times2} \\ \hline \mathbf{0}_{2\times4} & \mathbf{I}_{2\times2} \end{array}\right] = \left[\begin{array}{cccc|cc} 0 & 1-p^+-p^- & p^+ & p^- & 0 & 0 \\ 0 & 1-p^+-p^- & p^+ & p^- & 0 & 0 \\ 0 & 1-p^+-p^- & 0 & p^- & p^+ & 0 \\ 0 & 1-p^+-p^- & p^+ & 0 & 0 & p^- \\ \hline 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{array}\right]. \qquad (4\text{-}26)$$

Two-sided *2-of-3* sign chart

The run-length N_{2of3} of the *2-of-3* two-sided chart is the waiting time for the first occurrence of the event E_1 or E_2 or E_3 or E_4 (see e.g. Figure 4.7), which is the first occurrence of the compound pattern $\Lambda = \bigcup_{i=1}^{4} \Lambda_i$ in the series of ξ_i's (i.e. among the 0's, 1's and 2's), where $\Lambda_1 = 011$, $\Lambda_2 = 101$, $\Lambda_3 = 022$ and $\Lambda_4 = 202$ are the four distinct simple patterns.

The imbedded Markov chain, in this case, is defined on the finite state space $\Omega = \{\phi, 0, 1, 2, 01, 10, 02, 20, \alpha_1, \alpha_2, \alpha_3, \alpha_4\}$ with $m = 12$ states, where the four states $\alpha_1 = \{011\}, \alpha_2 = \{101\}$, $\alpha_3 = \{022\}$ and $\alpha_4 = \{202\}$ are the absorbing states, ϕ is the dummy state, and the eight transient states are all the sequential sub-patterns of $\Lambda_1 = 011$, $\Lambda_2 = 101$, $\Lambda_3 = 022$ and $\Lambda_4 = 202$, respectively. In this case, the essential transition probability sub-matrix $\mathbf{Q}_{8\times8}$ of the transition probability matrix

$$\mathbf{M}_{12\times12} = \left[\begin{array}{c|c} \mathbf{Q}_{8\times8} & \mathbf{C}_{8\times4} \\ \hline \mathbf{0}_{4\times8} & \mathbf{I}_{4\times4} \end{array}\right]$$

is given by

$$\mathbf{Q}_{8\times8} = \begin{bmatrix} 0 & 1-p^+-p^- & p^+ & p^- & 0 & 0 & 0 & 0 \\ 0 & 1-p^+-p^- & 0 & 0 & p^+ & 0 & p^- & 0 \\ 0 & 0 & p^+ & p^- & 0 & 1-p^+-p^- & 0 & 0 \\ 0 & 0 & p^+ & p^- & 0 & 0 & 0 & 1-p^+-p^- \\ 0 & 0 & 0 & p^- & 0 & 1-p^+-p^- & 0 & 0 \\ 0 & 1-p^+-p^- & 0 & 0 & 0 & 0 & p^- & 0 \\ 0 & 0 & p^+ & 0 & 0 & 0 & 0 & 1-p^+-p^- \\ 0 & 1-p^+-p^- & 0 & 0 & p^+ & 0 & 0 & 0 \end{bmatrix} \quad (4\text{-}27)$$

whilst the non-essential transition probability sub-matrix $\mathbf{C}_{8\times4}$ is given by

$$\mathbf{C}_{4\times8} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & p^+ & 0 & 0 \\ 0 & 0 & 0 & 0 & p^+ & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & p^- \\ 0 & 0 & 0 & 0 & 0 & 0 & p^- & 0 \end{bmatrix}^T.$$

Remark 9

In general, if $F_1, F_2, ..., F_r$, $r \geq 1$ are the set of *all possible events* in which (a) k consecutive charting statistics, or (b) exactly k of the last w charting statistics) can plot OOC, one can design a chart that signals on the first occurrence of the event $F = \bigcup_{i=1}^{r} F_i$. The run-length of such a chart would be (a) geometric of order k, or (b) geometric of order k/w.

However, we prefer, due to practical considerations, to exclude some of the F_i's ; in doing so the distribution of the run-length is not necessarily geometric of order k or geometric of order k/w and we then use the Markov chain approach to find the run-length distribution.

For example, as mentioned earlier, because the two-sided *2-of-2* KL chart signals only if event D_1 or D_2 occurs for the first time and does not signal (unlike the two-sided *2-of-2* DR chart) in case event D_3 or event D_4 occurs (see Figure 4.6), the distribution of N_{2of2}^{KL}, in general, is not a geometric distribution of order $k = 2$.

Likewise, because the two-sided *2-of-3* sign chart signals only on the first occurrence of event E_1 or E_2 or E_3 or E_4 (see e.g. Figure 4.7) and excludes the remaining eight events in which exactly two of the last three charting statistics can plot on or outside the control limits i.e. events E_5, E_6, E_7, E_8, E_9, E_{10}, E_{11} and E_{12} (see e.g. Figure 4.8), as signaling events the distribution of N_{2of3}, in general, is not a geometric distribution of order $2/3$.

If, however, we were to design a two-sided *2-of-3* sign chart that signals on the first occurrence of either one of the events E_i for $i = 1,2,...,12$ the distribution of the run-length random variable associated with such a chart would be a geometric distribution of order $2/3$ with probability of success $\Pr(T_i \notin (LCL, UCL)) = p^{\pm} = p^{+} + p^{-}$ (also see Remark 7).

4.2.3 The in-control run-length characteristics of the one-sided and two-sided sign charts

The characteristics of the in-control (IC) run-length distributions are essential in the design of a control chart. Furthermore, for out-of-control (OOC) performance comparisons their in-control average run-length (ARL_0) and/or false alarm rate (FAR) should be equal or, at least, approximately so.

Tables 4.4 and 4.5 summarize the expressions for the ARL and the FAR of the various sign charts. The ARL expressions, in general, follow from having written the corresponding essential transition probability sub-matrix $\mathbf{Q}_{h\times h}$, substituting it in (4-9) and simplifying symbolically.

For example, for the *1-of-1* two-sided chart with state space $\Omega = \{\phi, 0, 1, 2\}$, it was shown that

$$\mathbf{Q}_{2\times 2} = \begin{bmatrix} 0 & 1-p^+-p^- \\ 0 & 1-p^+-p^- \end{bmatrix}$$

so that upon substitution in (4-9) and simplifying we get an explicit formula for the ARL given by

$$ARL_{1of1} = E(N_{1of1}) = \begin{bmatrix} 1 & 0 \end{bmatrix} \left[\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} 0 & 1-p^+-p^- \\ 0 & 1-p^+-p^- \end{bmatrix} \right]^{-1} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \frac{1}{p^+ + p^-}.$$

Alternatively, in some cases (such as the upper and the lower one-sided *1-of-1* and *2-of-2* charts as well as the two-sided *1-of-1* and *2-of-2* DR charts) one can obtain closed form expressions by using available results of the geometric distribution of order k. For example, for the two-sided *2-of-2* DR chart, one obtains the ARL upon substituting $p^{\pm} = p^+ + p^-$ for α in (4-17); this gives

$$ARL^{DR}_{2of2} = E(N^{DR}_{2of2}) = E(T_2) = \frac{1-(p^+ + p^-)^2}{(1-p^+-p^-)(p^+ + p^-)^2} = \frac{p^+ + p^- + 1}{(p^+ + p^-)^2}.$$

Note that, for the in-control average run-length

$$p_0^+ = p_0^+(n, b, \theta = \theta_0) = I_{p_0}(n-b, b+1)$$

and

$$p_0^- = p_0^-(n, a, \theta = \theta_0) = 1 - I_{p_0}(a+1, n-a)$$

where p_0 is defined in (4-6), are to be substituted for p^+ and p^-, respectively, in the ARL expressions of Tables 4.4 and 4.5.

The expressions for the FAR can be obtained from the definitions of the charts in a straightforward manner. For example, for the *2-of-2* upper one-sided chart, the false alarm rate is

$$FAR^{+}_{2of2} = \Pr(B_1 \mid IC) = \Pr(T_{i-1} \geq UCL \mid IC) \times \Pr(T_i \geq UCL \mid IC) = (p_0^{+})^2$$

where B_1 is defined in panel (a) of Figure 4.3, whereas the false alarm rate for the *2-of-2* KL chart is

$$\begin{aligned} FAR^{KL}_{2of2} &= \Pr(D_1 \mid IC) + \Pr(D_2 \mid IC) \\ &= \Pr(T_{i-1} \geq UCL, T_i \geq UCL \mid IC) + \Pr(T_{i-1} \leq LCL, T_i \leq LCL \mid IC) \\ &= (p_0^{+})^2 + (p_0^{-})^2 \end{aligned}$$

where D_1 and D_2 are defined in panels (a) and (b) of Figure 4.6.

Table 4.4: Average run-lengths (*ARL*'s) and false alarm rates (*FAR*'s) of the upper one-sided sign charts

***1-of-1* upper**	***2-of-2* upper**	***2-of-3* upper**
$ARL^{+}_{1of1} = \dfrac{1}{p^{+}}$	$ARL^{+}_{2of2} = \dfrac{1+p^{+}}{(p^{+})^2}$	$ARL^{+}_{2of3} = \dfrac{(p^{+})^3 - 2(p^{+})^2 + p^{+} + 1}{(p^{+})^2[(p^{+})^2 - 3p^{+} + 2)]}$
$FAR^{+}_{1of1} = p_0^{+}$	$FAR^{+}_{2of2} = (p_0^{+})^2$	$FAR^{+}_{2of3} = 2(1-p_0^{+})(p_0^{+})^2$

Table 4.5: Average run-lengths (*ARL*'s) and false alarm rates (*FAR*'s) of the two-sided sign charts

1-of-1	***2-of-2* DR**	***2-of-2* KL**
$ARL_{1of1} = \dfrac{1}{p^{+}+p^{-}}$	$ARL^{DR}_{2of2} = \dfrac{p^{+}+p^{-}+1}{(p^{+}+p^{-})^2}$	$ARL^{KL}_{2of2} = \left(\dfrac{(p^{+})^2}{(p^{+}+1)} + \dfrac{(p^{-})^2}{(p^{-}+1)}\right)^{-1}$
$FAR_{1of1} = p_0^{+} + p_0^{-}$	$FAR^{DR}_{2of2} = (p_0^{+})^2 + (p_0^{-})^2 + 2(p_0^{+})(p_0^{-})$	$FAR^{KL}_{2of2} = (p_0^{+})^2 + (p_0^{-})^2$

2-of-3

$$ARL_{2of3} = \frac{\left[p^{-} + p^{+}(p^{-})^2 - p^{+}p^{-} - 2(p^{-})^2 + (p^{-})^3 + 1\right]^2 \left[p^{+} + (p^{+})^2 p^{-} - p^{+}p^{-} - 2(p^{+})^2 + (p^{+})^3 + 1\right]^2}{\left[(p^{+}+p^{-}-1)\begin{pmatrix} -2p^{+}(p^{-})^2 - 2(p^{+})^2 p^{-} + 3p^{+}(p^{-})^3 + 3(p^{+})^3 p^{-} - p^{+}(p^{-})^4 \\ -(p^{+})^4 p^{-} + 8(p^{+})^2(p^{-})^2 - 6(p^{+})^2(p^{-})^3 - 6(p^{+})^3(p^{-})^2 \\ +(p^{+})^2(p^{-})^4 + 2(p^{+})^3(p^{-})^3 + (p^{+})^4(p^{-})^2 - 2(p^{+})^2 + (p^{+})^3 \\ -2(p^{-})^2 + (p^{-})^3 \end{pmatrix}\right]^2}$$

$$= \frac{2q^3 - 3q^2 + q + 1}{2q^2(2q^2 - 5q + 2)} \quad \text{if} \quad p^{+} = p^{-} = q \text{ (say) i.e. symmetrically placed control limits (equidistant from both ends)}$$

$$FAR_{2of3} = 2(p_0^{+})^2(1 - p_0^{+} - p_0^{-}) + 2(p_0^{-})^2(1 - p_0^{+} - p_0^{-})$$

Remark 10

As mentioned earlier, the FAR and the ARL_0 of all the sign charts depend only on the probabilities p_0^+ and/or p_0^-, which in turn depend only on the sample size n and the charting constants a and/or b and not on the underlying process distribution. The in-control run-length distributions therefore remain the same for all continuous process distributions, and hence the proposed sign charts are nonparametric or distribution-free.

4.2.4 Design of the upper (lower) one-sided *1-of-1*, *2-of-2* and *2-of-3* sign charts

In order to design the proposed charts and assess their in-control performance the design parameters need to be chosen. The design parameters include

(i) the sample size n,

(ii) the charting constants a and b, and

(iii) the target value θ_0.

Because rational subgroups in SPC are small, we focus on $n = 4(1)15$, 20 and 25.

To monitor the center of a process, one typically chooses θ to be the process median and this is the case we study here; hence $\pi = 0.5$ so that $\theta = F_X^{-1}(0.5)$ and $p = p_0 = \Pr(X_{ij} > \theta_0 \mid IC) = 0.5$. However, other choices of θ might be desirable in some situations. For example, to monitor the 25th percentile of a processes' distribution we would set $\pi = 0.25$ so that $\theta = F_X^{-1}(0.25)$ and then $p_0 = \Pr(X_{ij} > \theta_0 \mid IC) = 0.75$. The sign charts are flexible enough to allow one to do that.

The charting constants a and b can be any integer between and including 0 and n. However, a and/or b are typically chosen so that the ARL_0 is reasonably large.

Tables 4.6 and 4.7 display the in-control characteristics of the *1-of-1* sign chart of Amin et al. (1995) and the new proposed runs-rule enhanced *2-of-2* and *2-of-3* one-sided sign charts.

Note that, Tables 4.6 and 4.7 apply to both the lower and the upper one-sided sign charts because in case of the median we have that $p = p_0 = 0.5$ when the process is in-control, which implies that the charting statistic T_i has a binomial (n, 0.5) distribution which is symmetric. Hence, when $b = a$ (as in Tables 4.6 and 4.7), the in-control performance of the lower and the upper one-sided sign charts are identical.

For example, if $n = 6$ and the LCL of the lower one-sided chart is $a = 1$, the UCL for the upper one-sided chart is simply $n - b = 6 - 1 = 5$, and both of these charts have an in-control ARL of 9.14 and a FAR of 0.10938.

An examination of the *ARL* and *FAR* values in Tables 4.6 and 4.7 reveal the advantages of the new sign charts

(i) They offer more practically attractive ARL_0's and *FAR*'s in that, for any particular combination of n and a, the attained ARL_0 values of the *2-of-2* and the *2-of-3* charts are much higher than those of the *1-of-1* chart with a corresponding decrease in the *FAR*.

For example, for $n = 5$ and $a = 0$ the ARL_0 of the *1-of-1* chart is 32.00 with a fairly large *FAR* of 0.03125, but for the *2-of-3* chart the ARL_0 increases to a more reasonable 552.65 and the *FAR* decreases to 0.00189, whereas for the *2-of-2* chart, the ARL_0 equals 1056.00 with a *FAR* of 0.00098.

(ii) Most importantly, when using the *1-of-1* chart the industry standard ARL_0 value of 370 and *FAR* of 0.0027 is far from being attainable, but with the proposed *2-of-2* chart, for example, when $n = 10$ and $a = 2$, we can be almost on target e.g. the ARL_0 and *FAR* values are 352.65 and 0.00299, respectively.

Thus, by carefully choosing the sample size n, the charting constants a and/or *b*, and the values of k and w, we can attain more familiar and recommended values for the ARL_0 and the *FAR* for the proposed nonparametric sign charts. Even for a sample size as small as $n = 4$, an ARL_0 of 272.00 with a *FAR* of 0.00391 is possible when the *2-of-2* chart is used with $a = 0$.

Amin et al. (1995) noted that the largest possible ARL_0 of their *1-of-1* one-sided sign chart for the median is 2^n. However, our runs-rules based sign charts provide a wider range of attainable ARL_0 values and false alarm rates. For instance, for $n = 15$ the sign charts can attain an ARL_0 (*FAR*) as low (high) as 6.00 (0.25) and an ARL_0 (*FAR*) as high (low) as 3293.23 (0.00031) with the *2-of-2* chart.

Table 4.6: The in-control characteristics (*ARL* and *FAR*) of the one-sided *1-of-1*, *2-of-2* and *2-of-3* sign charts for the median (for samples of size $n = 4(1)11$)*

Sample size	*LCL*	*UCL*	*1-of-1*		*2-of-2*		*2-of-3*	
n	***a***	***n-b=n-a***	ARL_0	***FAR***	ARL_0	***FAR***	ARL_0	***FAR***
4	**0**	**4**	16.00	0.06250	272.00	0.00391	148.68	0.00732
	1	**3**			13.44	0.09766	10.13	0.13428
5	**0**	**5**	32.00	0.03125	1056.00	0.00098	552.65	0.00189
	1	**4**	5.33	0.18750	33.78	0.03516	21.71	0.05713
	2	**3**			6.00	0.25000		
6	**0**	**6**	64.00	0.01563	4160.00	0.00024	2128.64	0.00048
	1	**5**	9.14	0.10938	92.73	0.01196	53.95	0.02131
	2	**4**			11.37	0.11816	8.94	0.15509
7	**0**	**7**	128.00	0.00781			8352.63	0.00012
	1	**6**	16.00	0.06250	272.00	0.00391	148.68	0.00732
	2	**5**			23.90	0.05133	16.13	0.07940
	3	**4**			6.00	0.25000		
8	**0**	**8**	256.00	0.00391				
	1	**7**	28.44	0.03516	837.53	0.00124	440.75	0.00239
	2	**6**	6.92	0.14453	54.79	0.02089	33.35	0.03574
	3	**5**			10.33	0.13197	8.34	0.16806
9	**0**	**9**	512.00	0.00195				
	1	**8**	51.20	0.01953	2672.64	0.00038	1375.36	0.00075
	2	**7**	11.13	0.08984	135.02	0.00807	76.56	0.01469
	3	**6**			19.45	0.06447	13.59	0.09620
	4	**5**			6.00	0.25000		
10	**0**	**10**	1024.00	0.00098				
	1	**9**	93.09	0.01074	8759.01	0.00012	4449.96	0.00023
	2	**8**	18.29	0.05469	352.65	0.00299	190.71	0.00565
	3	**7**	5.82	0.17188	39.67	0.02954	25.00	0.04893
	4	**6**			9.69	0.14209	7.98	0.17706
11	**0**	**11**	2048.00	0.00049				
	1	**10**	170.67	0.00586				
	2	**9**	30.57	0.03271	964.92	0.00107	506.04	0.00207
	3	**8**	8.83	0.11328	86.75	0.01283	50.73	0.02276
	4	**7**			16.92	0.07530	12.14	0.10928
	5	**6**			6.00	0.25000		

**Note*: Only ARL_0 values greater than 5 and less than 10 000 are shown.

Table 4.7: The in-control characteristics (*ARL* and *FAR*) of the one-sided *1-of-1*, *2-of-2* and *2-of-3* sign charts for the median (for samples of size $n = 12(1)15(5)25$)*

Sample size	*LCL*	*UCL*	*1-of-1*		*2-of-2*		*2-of-3*	
n	***a***	***n-b=n-a***	ARL_0	***FAR***	ARL_0	***FAR***	ARL_0	***FAR***
12	**0**	**12**	4096.00	0.00024				
	1	**11**	315.08	0.00317				
	2	**10**	51.85	0.01929	2740.07	0.00037	1409.56	0.00073
	3	**9**	13.70	0.07300	201.36	0.00533	111.64	0.00988
	4	**8**	5.16	0.19385	31.77	0.03758	20.58	0.06059
	5	**7**			9.25	0.14993	7.73	0.18375
13	**0**	**13**	8192.00	0.00012				
	1	**12**	585.14	0.00171				
	2	**11**	89.04	0.01123	8017.78	0.00013	4076.31	0.00025
	3	**10**	21.67	0.04614	491.35	0.00213	262.59	0.00406
	4	**9**	7.49	0.13342	63.67	0.01780	38.21	0.03085
	5	**8**			15.29	0.08441	11.20	0.11977
	6	**7**			6.00	0.25000		
14	**1**	**13**	1092.27	0.00092				
	2	**12**	154.57	0.00647				
	3	**11**	34.86	0.02869	1250.05	0.00082	651.82	0.00160
	4	**10**	11.14	0.08978	135.19	0.00806	76.66	0.01467
	5	**9**			26.97	0.04493	17.87	0.07082
	6	**8**			8.93	0.15623	7.55	0.18896
15	**1**	**14**	2048.00	0.00049				
	2	**13**	270.81	0.00369				
	3	**12**	56.89	0.01758	3293.23	0.00031	1689.92	0.00061
	4	**11**	16.88	0.05923	301.88	0.00351	164.28	0.00660
	5	**10**	6.63	0.15088	50.56	0.02276	31.02	0.03866
	6	**9**			14.14	0.09218	10.53	0.12839
	7	**8**			6.00	0.25000		
20	**2**	**18**	4969.55	0.00020				
	3	**17**	776.15	0.00129				
	4	**16**	169.23	0.00591				
	5	**15**	48.32	0.02069	2383.29	0.00043	1228.53	0.00084
	6	**14**	17.34	0.05766	318.13	0.00332	172.75	0.00627
	7	**13**	7.60	0.13159	65.35	0.01732	39.13	0.03007
	8	**12**			19.75	0.06336	13.76	0.09483
	9	**11**			8.32	0.16966	7.21	0.19956
25	**4**	**21**	2196.55	0.00046				
	5	**20**	490.52	0.00204				
	6	**19**	136.67	0.00732			9511.45	0.00011
	7	**18**	46.21	0.02164	2181.12	0.00047	1125.86	0.00092
	8	**17**	18.56	0.05388	363.08	0.00290	196.13	0.00549
	9	**16**	8.71	0.11476	84.64	0.01317	49.59	0.02332
	10	**15**			26.93	0.04502	17.85	0.07094
	11	**14**			11.30	0.11904	8.90	0.15594
	12	**13**			6.00	0.25000		

**Note*: Only ARL_0 values greater than 5 and less than 10 000 are shown.

4.2.5 Performance comparison of the one-sided sign charts

We compare the performance of the *1-of-1*, the *2-of-2* and the *2-of-3* sign charts to that of the competing *1-of-1* SR chart of Bakir (2004) and the *2-of-2* SR charts of Chakraborti and Eryilmaz (2007) under the normal, the double exponential (or Laplace) and the Cauchy distributions; for completeness, we also include the well-known Shewhart $\overline{X}$ chart.

The double exponential and the Cauchy distributions are normal like with different tail behavior (see e.g. Figure 4.11). For the double exponential distribution the scale parameter was set equal to $1/\sqrt{2}$ for a standard deviation of 1; for the Cauchy distribution the scale parameter was set equal to 0.2605 in order to achieve a tail probability of 0.05 above $\theta_0 + 1.645$ - the same as for a $N(\theta_0,1)$ distribution.

Without loss of generality, we take the in-control median to be $\theta_0 = 0$. All three distributions are symmetric and a shift refers to a shift in the mean (median). The amount of shift in the median was taken over the range $\delta = 0(0.2)1.2$.

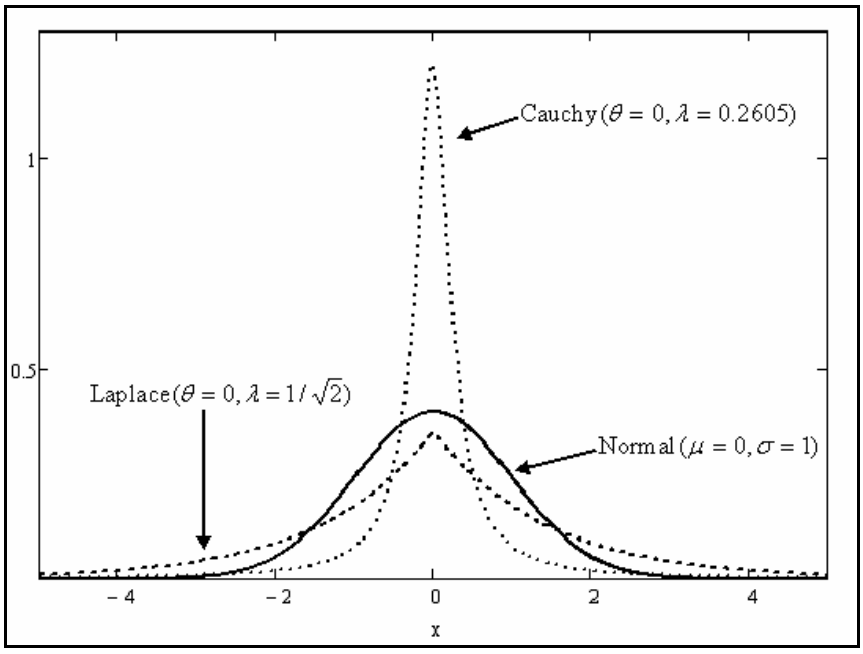


Figure 4.11: Probability distributions used for the performance comparison of the sign control charts

For comparison purposes the control charts are designed so that the ARL_0 values are high and are approximately equal. However, because the nonparametric charts are based on charting statistics that have discrete distributions, it is not possible to straightforwardly design the charts such that their ARL_0 values are all equal, and equal to some desired value such as 370 for a given sample size n (see e.g. Tables 4.6 and 4.7) .

Randomization was therefore used to ensure that the charts all have the same ARL_0 for a selected sample size. The technique is mainly used in the testing literature to compare the power of tests based on discrete test statistics so that they have identical nominal Type I error probability such as 0.05 (see e.g. Gibbons and Chakraborti, (2003)). We provide an example for illustration with the *1-of-1* sign chart; randomization for the other nonparametric charts can be handled in a similar way.

Example 2

Consider constructing a *1-of-1* upper one-sided sign chart with ARL_0 of 370 when samples of size $n = 10$ are used. From Table 4.6 we see that for this chart, exact in-control ARL values of 1024 (when $UCL = 10$) and 93.09 (when $UCL = 9$) are attainable that trap the target value 370.

The following randomized decision rule has an exact ARL_0 of 370:

"Declare the process OOC if $T_i \geq UCL = 10$ (with probability 1) and with probability q if $T_i = UCL - 1 = 9$, where $0 < q < 1$ is chosen such that $\Pr(T_i \geq 10 \mid IC) + q.\Pr(T_i = 9 \mid IC) = 1/370$".

Assuming that the median is the parameter of interest and the process is IC, $T_i \sim Bin(10,0.5)$ and therefore

$$q = [1/370 - \Pr(T_i \geq 10 \mid IC)]/\Pr(T_i = 9 \mid IC) = 0.1768 \approx 0.18 .$$

Hence, if we declare the process OOC every time the charting statistic is greater than or equal to 10 and declare the process OOC in 17.68% of the cases the charting statistic equals 9, we would have a *1-of-1* upper one-sided sign chart with an in-control ARL of 370.

In practice, we could use a random number generator to make a decision; for example, if the charting statistic equals 9, we could draw a random number between 1 and 100; if the drawn number is between 1 and 18, the process is declared OOC, otherwise it is not.

Note that, randomization is used to ensure that the in-control ARL values of the competing charts are equal so that their OOC performance can be fairly compared. The implementation and application of the charts remain as defined earlier and require no randomization.

The in-control (when $\delta = 0$) and the out-of-control (when $\delta \neq 0$) characteristics of the various charts for samples of size $n = 10$ are shown in Tables 4.8, 4.9 and 4.10 under the normal, the double exponential and the Cauchy distribution, respectively. The characteristics include the ARL, the $SDRL$ as well as the 5th, the 25th (the first quartile, Q_1), the 50th (the median run-length, $MDRL$), the 75th (the third quartile, Q_3) and the 95th percentiles of the run-length distribution.

Note that, since randomization was used and the out-of-control distribution for the SR statistic is unavailable for most distributions, we used simulations (100 000 samples each of $n = 10$) to estimate these characteristics in SAS®9.1; these programs can be found in Appendix 4A.

Table 4.8: In-control and out-of-control characteristics of the run-length distributions of the one-sided *1-of-1* sign, the *2-of-2* sign, the *2-of-3* sign, the *1-of-1* SR and the *2-of-2* SR chart for the median under the normal distribution

	***1-of-1* sign (*UCL*=10)**							***1-of-1* SR (*UCL*=53)**						
Shift	***ARL***	***SDRL***	**5th**	**Q_1**	***MDRL***	**Q_3**	**95th**	***ARL***	***SDRL***	**5th**	**Q_1**	***MDRL***	**Q_3**	**95th**
0.0	370	370.4	19	107	259	514	1107	370	369.7	19	106	255	509	1099
0.2	104	103.7	6	30	72	144	310	88.7	88.2	5	26	62	123	265
0.4	35.6	35.2	2	11	25	49	106	27.3	26.8	2	8	19	38	81
0.6	14.7	14.2	1	5	10	20	43	10.6	10.1	1	3	7	14	31
0.8	7.3	6.8	1	2	5	10	21	5	4.5	1	2	4	7	14
1.0	4.2	3.7	1	2	3	6	12	2.9	2.3	1	1	2	4	7
1.2	2.8	2.2	1	1	2	4	7	1.9	1.3	1	1	1	2	5
	***2-of-2* sign (*UCL*=9)**							***2-of-2* SR (*UCL*=33)**						
Shift	***ARL***	***SDRL***	**5th**	**Q_1**	***MDRL***	**Q_3**	**95th**	***ARL***	***SDRL***	**5th**	**Q_1**	***MDRL***	**Q_3**	**95th**
0.0	370	368.9	20	108	258	509	1107	370	371.9	21	108	256	514	1122
0.2	64	62.6	5	19	45	88	189	50.9	49.4	4	16	36	70	150
0.4	17.5	16.2	2	6	13	24	50	12.7	11.4	2	5	9	17	35
0.6	7.2	5.9	2	3	5	10	19	5.2	3.9	2	2	4	7	13
0.8	4	2.7	2	2	3	5	9	3.1	1.7	2	2	2	4	7
1.0	2.9	1.5	2	2	2	3	6	2.4	0.9	2	2	2	2	4
1.2	2.4	0.8	2	2	2	2	4	2.1	0.4	2	2	2	2	3
	***2-of-3* sign (*UCL*=9)**							**X-bar (*UCL* = 0.8797)**						
Shift	***ARL***	***SDRL***	**5th**	**Q_1**	***MDRL***	**Q_3**	**95th**	***ARL***	***SDRL***	**5th**	**Q_1**	***MDRL***	**Q_3**	**95th**
0.0	370	362.8	21	108	255	510	1094	370	370.3	20	107	257	513	1109
0.2	62.2	60.3	5	19	44	86	183	63.4	62.5	4	19	44	88	187
0.4	17.2	15.5	3	6	12	23	48	15.5	14.9	1	5	11	21	45
0.6	7.2	5.6	2	3	5	9	18	5.3	4.8	1	2	4	7	15
0.8	4.1	2.6	2	2	3	5	9	2.5	1.9	1	1	2	3	6
1.0	3	1.4	2	2	3	3	6	1.5	0.9	1	1	1	2	3
1.2	2.5	0.8	2	2	2	3	4	1.2	0.5	1	1	1	1	2

Table 4.9: In-control and out-of-control characteristics of the run-length distributions of the one-sided *1-of-1* sign, the *2-of-2* sign, the *2-of-3* sign, the *1-of-1* SR and the *2-of-2* SR chart for the median under the double exponential distribution

	***1-of-1* sign (*UCL*=10)**							***1-of-1* SR (*UCL*=53)**						
Shift	***ARL***	***SDRL***	**5^{th}**	**Q_1**	***MDRL***	**Q_3**	**95^{th}**	***ARL***	***SDRL***	**5^{th}**	**Q_1**	***MDRL***	**Q_3**	**95^{th}**
0.0	370	371.2	19	107	257	514	1110	370	370.4	19	106	256	511	1111
0.2	54.6	53.8	3	16	38	75	162	48.7	48.2	3	14	34	68	145
0.4	16.6	16.1	1	5	12	23	49	13.3	12.8	1	4	9	18	39
0.6	7.5	7	1	3	5	10	21	5.7	5.1	1	2	4	8	16
0.8	4.3	3.8	1	2	3	6	12	3.2	2.6	1	1	2	4	8
1.0	3	2.4	1	1	2	4	8	2.2	1.6	1	1	2	3	5
1.2	2.2	1.7	1	1	2	3	6	1.7	1	1	1	1	2	4
	***2-of-2* sign (*UCL*=9)**							***2-of-2* SR (*UCL*=33)**						
Shift	***ARL***	***SDRL***	**5^{th}**	**Q_1**	***MDRL***	**Q_3**	**95^{th}**	***ARL***	***SDRL***	**5^{th}**	**Q_1**	***MDRL***	**Q_3**	**95^{th}**
0.0	370	368.5	21	107	256	512	1099	370	369	21	107	255	511	1102
0.2	29	27.5	3	9	20	40	84	31.8	30.4	3	10	22	44	92
0.4	8	6.7	2	3	6	11	21	8.1	6.8	2	3	6	11	22
0.6	4.1	2.8	2	2	3	5	10	4	2.7	2	2	3	5	9
0.8	2.9	1.5	2	2	2	4	6	2.8	1.4	2	2	2	3	6
1.0	2.4	0.9	2	2	2	2	4	2.3	0.8	2	2	2	2	4
1.2	2.2	0.6	2	2	2	2	4	2.1	0.5	2	2	2	2	3
	***2-of-3* sign (*UCL*=9)**							**X-bar (*UCL*=0.9267)**						
Shift	***ARL***	***SDRL***	**5^{th}**	**Q_1**	***MDRL***	**Q_3**	**95^{th}**	***ARL***	***SDRL***	**5^{th}**	**Q_1**	***MDRL***	**Q_3**	**95^{th}**
0.0	370	361	20	105	251	505	1089	370	368.1	20	107	257	510	1103
0.2	28.2	26.4	3	9	20	38	81	80.1	79.9	5	23	56	111	240
0.4	8	6.4	2	3	6	10	21	20.9	20.3	2	6	15	29	61
0.6	4.2	2.6	2	2	3	5	10	6.9	6.4	1	2	5	9	20
0.8	3	1.4	2	2	3	3	6	3	2.4	1	1	2	4	8
1.0	2.5	0.9	2	2	2	3	4	1.7	1.1	1	1	1	2	4
1.2	2.3	0.6	2	2	2	2	3	1.2	0.5	1	1	1	1	2

Table 4.10: In-control and out-of-control characteristics of the run-length distributions of the one-sided *1-of-1* sign, the *2-of-2* sign, the *2-of-3* sign, the *1-of-1* SR and the *2-of-2* SR chart for the median under the Cauchy distribution

	***1-of-1* sign (*UCL*=10)**							***1-of-1* SR (*UCL*=53)**						
Shift	***ARL***	***SDRL***	**5^{th}**	**Q_1**	***MDRL***	**Q_3**	**95^{th}**	***ARL***	***SDRL***	**5^{th}**	**Q_1**	***MDRL***	**Q_3**	**95^{th}**
0.0	370	370.9	19	106	259	516	1113	370	367.7	19	105	255	511	1101
0.2	18.2	17.7	1	6	13	25	53	15.6	15.1	1	5	11	21	46
0.4	5.4	4.9	1	2	4	7	15	4.5	4	1	2	3	6	12
0.6	3.2	2.6	1	1	2	4	8	2.7	2.1	1	1	2	4	7
0.8	2.4	1.8	1	1	2	3	6	2.1	1.5	1	1	2	3	5
1.0	2	1.4	1	1	2	3	5	1.7	1.1	1	1	1	2	4
1.2	1.8	1.2	1	1	1	2	4	1.6	1	1	1	1	2	3
	***2-of-2* sign (*UCL*=9)**							***2-of-2* SR (*UCL*=33)**						
Shift	***ARL***	***SDRL***	**5^{th}**	**Q_1**	***MDRL***	**Q_3**	**95^{th}**	***ARL***	***SDRL***	**5^{th}**	**Q_1**	***MDRL***	**Q_3**	**95^{th}**
0.0	370	368.4	21	107	257	511	1103	370	370.4	20	108	258	512	1111
0.2	8.8	7.4	2	3	6	12	24	11.4	10.1	2	4	8	15	31
0.4	3.3	1.9	2	2	2	4	7	4.1	2.8	2	2	3	5	10
0.6	2.5	1	2	2	2	3	5	2.9	1.5	2	2	2	4	6
0.8	2.2	0.7	2	2	2	2	4	2.5	1	2	2	2	3	5
1.0	2.1	0.5	2	2	2	2	3	2.3	0.8	2	2	2	2	4
1.2	2.1	0.4	2	2	2	2	3	2.2	0.6	2	2	2	2	4
	***2-of-3* sign (*UCL*=9)**							**X-bar (*UCL* = 30.6802)**						
Shift	***ARL***	***SDRL***	**5^{th}**	**Q_1**	***MDRL***	**Q_3**	**95^{th}**	***ARL***	***SDRL***	**5^{th}**	**Q_1**	***MDRL***	**Q_3**	**95^{th}**
0.0	370	360.5	21	106	252	503	1085	370	368	19	107	258	513	1101
0.2	8.7	7.1	2	4	7	11	23	367	366.5	19	106	255	507	1096
0.4	3.4	1.8	2	2	3	4	7	367	367.2	19	106	254	507	1097
0.6	2.6	1	2	2	2	3	5	364	365.5	19	105	251	504	1090
0.8	2.3	0.7	2	2	2	3	4	361	361.7	19	104	248	500	1084
1.0	2.2	0.5	2	2	2	2	3	360	360	19	104	249	499	1077
1.2	2.2	0.4	2	2	2	2	3	355	353.8	19	103	247	492	1066

Table 4.11 summarizes our findings from Tables 4.8, 4.9 and 4.10 and ranks the charts (from the most to the least favorable) under each of the three distributions. The ranking was based primarily on their ARL (the current norm in the SPC literature), but since the run-length distributions are right (positive) skewed, we also looked at the median run-length ($MDRL$), the first and third quartiles (i.e. Q_1 and Q_3), as well as the 5^{th} and the 95^{th} percentiles.

Table 4.11: Ranking (from most to least favorable) of the one-sided nonparametric charts for the median under the normal, the double exponential and the Cauchy distributions based on out-of-control ARL and run-length percentiles. The $ARL_0 = 370$

Normal	Double Exponential	Cauchy
2-of-2 SR	*2-of-2* sign / *2-of-3* sign	*2-of-2* sign / *2-of-3* sign
2-of-2 sign / *2-of-3* sign	*2-of-2* SR	*2-of-2* SR
1-of-1 $\overline{X}$	*1-of-1* SR	*1-of-1* SR
1-of-1 SR	*1-of-1* sign	*1-of-1* sign
1-of-1 sign	*1-of-1* $\overline{X}$	*1-of-1* $\overline{X}$

Overall, it is concluded that the proposed sign charts

(i) have substantially better out-of-control performance (i.e. shorter ARL values) than the *1-of-1* sign chart of Amin et al. (1995),

(ii) compete well with the SR charts of Bakir (2004) and Chakraborti and Eryilmaz (2007), and

(iii) outperform the Shewhart $\overline{X}$ chart in case of the heavier tailed distributions.

4.2.6 Design of the two-sided *2-of-2* DR, the *2-of-2* KL and the *2-of-3* sign charts

The characteristics of the in-control run-length distribution are typically used in the design and/or the implementation of a chart. As noted before, the ARL_0 should be high so that the time and/or effort spent on searching for nonexistent out-of-control conditions is not wasted.

Tables 4.12 and 4.13 display the ARL_0 and the FAR values of the two-sided *1-of-1*, *2-of-2* DR, *2-of-2* KL and *2-of-3* sign charts, respectively. For simplicity we only consider symmetrically placed control limits for the median i.e. $LCL = a$ and $UCL = n - a$, so that $p_0^+ = \Pr(T_i \geq UCL \mid IC) = \Pr(T_i \leq LCL \mid IC) = p_0^-$. Asymmetric control limits may of course be used when necessary, say for monitoring percentiles other than the median.

To attain the desired ARL_0 and/or FAR (for any one of the four charting procedures) the practitioner may use Tables 4.12 and 4.13 to select the suitable charting constant a (hence the control limits) for the sample size n at hand. Note that, as pointed out by Amin et al. (1995), the largest possible in-control ARL for the two-sided *1-of-1* sign chart is 2^{n-1} when $p = 0.5$, and thus unless n is sufficiently large, it is not possible to get close (even approximately) to an ARL_0 such as 370; this makes the *1-of-1* charts somewhat unattractive from a practical point of view.

However, for any combination of n and a values the ARL_0 (or FAR) values of the *2-of-2* DR, the *2-of-2* KL and the *2-of-3* sign charts are higher (or smaller) than that of the *1-of-1* sign chart.

For example, if $n = 5$ and $a = 0$ the $LCL = 0$ and the $UCL = 5$; the *1-of-1* sign chart has an ARL_0 of 16.00 (with a FAR of 0.06250), whereas both the *2-of-2* DR and the *2-of-2* KL charts have much higher ARL_0 values, 272.00 and 528.00, respectively (and much smaller FAR values, 0.00391 and 0.00195, respectively). Therefore, the new two-sided sign charts with signaling rules are more useful to the practitioner.

Table 4.12: The in-control characteristics (*ARL* and *FAR*) of the two-sided *1-of-1*, the *2-of-2* DR, the *2-of-2* KL and the *2-of-3* sign charts for the median (for samples of size $n = 4(1)14$)*

Sample size	*LCL*	*UCL*	*1-of-1*		*2-of-2* DR		*2-of-2* KL		*2-of-3*	
n	***a***	***n-a***	ARL_0	*FAR*	ARL_0	*FAR*	ARL_0	*FAR*	ARL_0	*FAR*
4	**0**	**4**	8.00	0.12500	72.00	0.01563	136.00	0.00781	79.37	0.01367
	1	**3**					6.72	0.19531	8.74	0.14648
5	**0**	**5**	16.00	0.06250	272.00	0.00391	528.00	0.00195	285.27	0.00366
	1	**4**			9.78	0.14063	16.89	0.07031	13.75	0.08789
6	**0**	**6**	32.00	0.03125	1056.00	0.00098	2080.00	0.00049	1081.23	0.00095
	1	**5**			25.47	0.04785	46.37	0.02393	30.45	0.03738
	2	**4**					5.69	0.23633	8.75	0.14771
7	**0**	**7**	64.00	0.01563	4160.00	0.00024	8256.00	0.00012	4209.21	0.00024
	1	**6**	8.00	0.12500	72.00	0.01563	136.00	0.00781	79.37	0.01367
	2	**5**			7.08	0.20532	11.95	0.10266	11.01	0.11229
8	**0**	**8**	128.00	0.00781						
	1	**7**	14.22	0.07031	216.49	0.00494	418.77	0.00247	228.45	0.00460
	2	**6**			15.43	0.08356	27.40	0.04178	19.74	0.05940
	3	**5**					5.16	0.26395	9.00	0.14435
9	**0**	**9**	256.00	0.00391						
	1	**8**	25.60	0.03906	680.96	0.00153	1336.32	0.00076	701.4	0.00147
	2	**7**	5.57	0.17969	36.54	0.03229	67.51	0.01614	42.18	0.02649
	3	**6**			5.85	0.25787	9.72	0.12894	9.87	0.12692
10	**0**	**10**	512.00	0.00195						
	1	**9**	46.55	0.02148	2213.02	0.00046	4379.50	0.00023	2249.15	0.00045
	2	**8**	9.14	0.10938	92.73	0.01196	176.33	0.00598	100.94	0.01065
	3	**7**			11.37	0.11816	19.83	0.05908	15.43	0.07755
	4	**6**							9.32	0.13987
11	**0**	**11**	1024.00	0.00098						
	1	**10**	85.33	0.01172	7367.11	0.00014			7432.31	0.00014
	2	**9**	15.28	0.06543	248.87	0.00428	482.46	0.00214	261.61	0.00400
	3	**8**	4.41	0.22656	23.90	0.05133	43.38	0.02567	28.78	0.03970
	4	**7**			5.14	0.30121	8.46	0.15061	9.29	0.13590
12	**0**	**12**	2048.00	0.00049						
	1	**11**	157.54	0.00635						
	2	**10**	25.92	0.03857	697.98	0.00149	1370.04	0.00074	718.66	0.00143
	3	**9**	6.85	0.14600	53.77	0.02131	100.68	0.01066	60.31	0.01820
	4	**8**			9.23	0.15031	15.89	0.07515	13.18	0.09203
									9.66	0.13529
13	**0**	**13**	4096.00	0.00024						
	1	**12**	292.57	0.00342						
	2	**11**	44.52	0.02246	2026.71	0.00050	4008.89	0.00025	2061.32	0.00049
	3	**10**	10.84	0.09229	128.25	0.00852	245.67	0.00426	137.7	0.00773
	4	**9**			17.79	0.07121	31.83	0.03560	22.26	0.05221
	5	**8**					7.64	0.16881	8.99	0.14145
14	**0**	**14**	8192.00	0.00012						
	1	**13**	546.13	0.00183						
	2	**12**	77.28	0.01294	6049.95	0.00017			6109.11	0.00017
	3	**11**	17.43	0.05737	321.23	0.00329	625.02	0.00165	335.57	0.00310
	4	**10**	5.57	0.17957	36.58	0.03224	67.60	0.01612	42.23	0.02645
	5	**9**			7.92	0.17973	13.49	0.08987	11.84	0.10354
	6	**8**							10.00	0.13091

**Note*: Only ARL_0 values greater than 5 and less than 10 000 are shown.

Table 4.13: The in-control characteristics (*ARL* and *FAR*) of the two-sided *1-of-1*, the *2-of-2* DR, the *2-of-2* KL and the *2-of-3* sign charts for the median (for samples of size $n = 15(5)25$)*

Sample size	*LCL*	*UCL*	*1-of-1*		*2-of-2* DR		*2-of-2* KL		*2-of-3*	
n	***a***	***n-a***	ARL_0	***FAR***	ARL_0	***FAR***	ARL_0	***FAR***	ARL_0	***FAR***
15	**1**	**14**	1024.00	0.00098						
	2	**13**	135.40	0.00739						
	3	**12**	28.44	0.03516	837.53	0.00124	1646.62	0.00062	860.1	0.00119
	4	**11**	8.44	0.11847	79.69	0.01403	150.94	0.00702	87.38	0.01237
	5	**10**			14.30	0.09106	25.28	0.04553	18.53	0.06358
	6	**9**					7.07	0.18437	8.82	0.14483
20	**2**	**18**	2484.78	0.00040						
	3	**17**	388.07	0.00258						
	4	**16**	84.62	0.01182	7244.68	0.00014			7309.35	0.00014
	5	**15**	24.16	0.04139	607.90	0.00171	1191.64	0.00086	627.27	0.00164
	6	**14**	8.67	0.11532	83.87	0.01330	159.07	0.00665	91.73	0.01176
	7	**13**			18.24	0.06926	32.68	0.03463	22.74	0.05103
	8	**12**			5.93	0.25346	9.88	0.12673	9.94	0.12586
25	**3**	**22**	6388.89	0.00016						
	4	**21**	1098.27	0.00091						
	5	**20**	245.26	0.00408						
	6	**19**	68.34	0.01463	4738.32	0.00021	9408.31	0.00011	4790.78	0.00021
	7	**18**	23.10	0.04329	556.83	0.00187	1090.56	0.00094	575.4	0.00179
	8	**17**	9.28	0.10775	95.41	0.01161	181.54	0.00581	103.71	0.01036
	9	**16**			23.34	0.05268	42.32	0.02634	28.18	0.04059
	10	**15**			7.91	0.18008	13.46	0.09004	11.83	0.10366
	11	**14**					5.65	0.23808	8.76	0.14759

**Note*: Only ARL_0 values greater than 5 and less than 10 000 are shown.

4.2.7 Performance comparison of the two-sided sign charts

The out-of-control performance of the two-sided sign control charts were compared amongst one another and with that of the two-sided SR charts under the normal, the double exponential and the Cauchy distributions; again we included the Shewhart $\overline{X}$ chart for completeness. The design parameters of the charts were chosen (coupled with randomization) so that the in-control *ARL* values were all equal to 370. As for the one-sided charts various characteristics of the run-length distributions were obtained using simulations and shown in Tables 4.14, 4.15 and 4.16, respectively with a summary of our findings given in Table 4.17.

Table 4.14: In-control and out-of-control properties of the run-length distributions of the two-sided *1-of-1* sign, the *2-of-2* DR sign, the *2-of-2* KL sign, the *2-of-3* sign, the *1-of-1* SR, *2-of-2* DR SR, *2-of-2* KL SR and the *1-of-1* X-bar charts under the Normal distribution

	***1-of-1* sign (*LCL*=0 , *UCL* = 10)**							***1-of-1* SR (*LCL* = -*UCL* = -55)**						
Shift	***ARL***	***SDRL***	**5^{th}**	**Q_1**	***MDRL***	**Q_3**	**95^{th}**	***ARL***	***SDRL***	**5^{th}**	**Q_1**	***MDRL***	**Q_3**	**95^{th}**
0.0	370	369.78	19	106	255	507	1106	370	377.3	20	110	264	528	1138
0.2	175.00	173.96	10	51	122	242	523	170.00	169.03	9	49	118	237	506
0.4	56.42	55.87	3	17	39	78	167	51.81	51.31	3	15	36	72	154
0.6	21.57	21.14	2	7	15	30	64	19.15	18.55	1	6	14	26	56
0.8	9.82	9.33	1	3	7	13	29	8.60	8.14	1	3	6	12	25
1.0	5.23	4.71	1	2	4	7	15	4.61	4.09	1	2	3	6	13
1.2	3.24	2.70	1	1	2	4	9	2.85	2.29	1	1	2	4	7
	***2-of-2* sign DR (*LCL*=1 , *UCL*=9)**							***2-of-2* SR DR (*LCL* = -*UCL* = -39)**						
Shift	***ARL***	***SDRL***	**5^{th}**	**Q_1**	***MDRL***	**Q_3**	**95^{th}**	***ARL***	***SDRL***	**5^{th}**	**Q_1**	***MDRL***	**Q_3**	**95^{th}**
0.0	370	363.53	20	106	253	501	1086	370	367.32	21	108	257	508	1103
0.2	166.30	164.54	10	49	116	230	496	129.10	127.36	8	38	90	179	385
0.4	43.75	42.3	4	14	31	60	128	27.42	26.04	3	9	19	38	80
0.6	14.68	13.25	2	5	11	20	41	8.77	7.42	2	3	6	12	24
0.8	6.83	5.47	2	3	5	9	18	4.25	2.92	2	2	3	5	10
1.0	4.08	2.71	2	2	3	5	10	2.82	1.40	2	2	2	3	6
1.2	2.97	1.54	2	2	2	4	6	2.29	0.74	2	2	2	2	4
	***2-of-2* sign KL (*LCL*=1, *UCL* = 9)**							***2-of-2* SR KL (*LCL* = -*UCL* = -37)**						
Shift	***ARL***	***SDRL***	**5^{th}**	**Q_1**	***MDRL***	**Q_3**	**95^{th}**	***ARL***	***SDRL***	**5^{th}**	**Q_1**	***MDRL***	**Q_3**	**95^{th}**
0.0	370	364.24	21	107	256	512	1099	370	368.73	21	108	256	514	1102
0.2	113.10	111.56	7	34	79	156	337	86.09	85.13	6	26	60	119	257
0.4	28.28	26.81	3	9	20	39	82	18.69	17.48	2	6	13	25	53
0.6	10.42	9.09	2	4	8	14	29	6.68	5.34	2	3	5	9	17
0.8	5.29	3.92	2	2	4	7	13	3.60	2.23	2	2	3	4	8
1.0	3.44	2.06	2	2	3	4	8	2.57	1.11	2	2	2	3	5
1.2	2.65	1.17	2	2	2	3	5	2.18	0.58	2	2	2	2	4
	***2-of-3* sign (*LCL*=1 , *UCL*=9)**							***1-of-1* X-bar (*LCL* = -*UCL* = -0.94858)**						
Shift	***ARL***	***SDRL***	**5^{th}**	**Q_1**	***MDRL***	**Q_3**	**95^{th}**	***ARL***	***SDRL***	**5^{th}**	**Q_1**	***MDRL***	**Q_3**	**95^{th}**
0.0	370	363.37	20	106	252	504	1088	370	368.45	19	106	256	510	1105
0.2	108.50	106.69	7	33	76	150	318	110.50	109.73	6	32	77	153	329
0.4	26.72	25.04	3	9	19	36	77	24.04	23.46	2	7	17	33	71
0.6	9.94	8.33	2	4	7	13	27	7.39	6.87	1	2	5	10	21
0.8	5.18	3.58	2	3	4	7	12	3.12	2.58	1	1	2	4	8
1.0	3.43	1.84	2	2	3	4	7	1.77	1.17	1	1	1	2	4
1.2	2.68	1.07	2	2	2	3	5	1.27	0.59	1	1	1	1	2

Table 4.15: In-control and out-of-control properties of the run-length distributions of the two-sided *1-of-1* sign, the *2-of-2* DR sign, the *2-of-2* KL sign, the *2-of-3* sign, the *1-of-1* SR, *2-of-2* DR SR, *2-of-2* KL SR and the *1-of-1* X-bar charts under the double exponential distribution

	***1-of-1* sign (*LCL*=0 , *UCL* = 10)**							***1-of-1* SR (*LCL* = -*UCL* = -55)**						
Shift	***ARL***	***SDRL***	**5^{th}**	**Q_1**	***MDRL***	**Q_3**	**95^{th}**	***ARL***	***SDRL***	**5^{th}**	**Q_1**	***MDRL***	**Q_3**	**95^{th}**
0.0	370	369.58	20	107	255	515	1106	370	377.57	21	113	263	528	1132
0.2	91.20	90.31	5	27	63	126	272	88.55	88.69	5	25	61	123	263
0.4	24.44	23.92	2	7	17	34	72	22.9	22.20	2	7	16	32	67.5
0.6	10.04	9.49	1	3	7	14	29	9.19	8.73	1	3	6	13	27
0.8	5.41	4.88	1	2	4	7	15	4.85	4.26	1	2	3	7	13
1.0	3.47	2.92	1	1	3	5	9	3.06	2.48	1	1	2	4	8
1.2	2.52	1.95	1	1	2	3	6	2.29	1.73	1	1	2	3	6
	***2-of-2* sign DR (*LCL*=1 , *UCL*=9)**							***2-of-2* SR DR (*LCL* = -*UCL* = -39)**						
Shift	***ARL***	***SDRL***	**5^{th}**	**Q_1**	***MDRL***	**Q_3**	**95^{th}**	***ARL***	***SDRL***	**5^{th}**	**Q_1**	***MDRL***	**Q_3**	**95^{th}**
0.0	370	361.71	20	106	253	503	1086	370	370.37	20	107	257	511	1106
0.2	77.17	75.93	5	23	54	106	229	78.94	77.11	5	24	55	109	233
0.4	16.89	15.57	2	6	12	23	48	15.76	14.52	2	5	11	21	45
0.6	6.93	5.56	2	3	5	9	18	6.17	4.85	2	3	5	8	16
0.8	4.17	2.79	2	2	3	5	10	3.65	2.30	2	2	3	4	8
1.0	3.09	1.67	2	2	2	4	6	2.76	1.33	2	2	2	3	6
1.2	2.59	1.11	2	2	2	3	5	2.36	0.85	2	2	2	2	4
	***2-of-2* sign KL (*LCL*=1, *UCL* = 9)**							***2-of-2* SR KL (*LCL* = -*UCL* = -37)**						
Shift	***ARL***	***SDRL***	**5^{th}**	**Q_1**	***MDRL***	**Q_3**	**95^{th}**	***ARL***	***SDRL***	**5^{th}**	**Q_1**	***MDRL***	**Q_3**	**95^{th}**
0.0	370	372.82	20	108	260	512	1114	370	368.05	21	107	257	513	1103
0.2	48.75	47.51	4	15	34	66	143	51.43	49.96	4	16	36	71	151
0.4	11.73	10.26	2	4	9	16	32	11.18	9.82	2	4	8	15	31
0.6	5.45	4.01	2	2	4	7	13.5	4.89	3.56	2	2	4	6	12
0.8	3.55	2.16	2	2	3	4	8	3.13	1.76	2	2	2	4	7
1.0	2.76	1.28	2	2	2	3	5	2.49	1.02	2	2	2	3	5
1.2	2.40	0.86	2	2	2	2	4	2.22	0.63	2	2	2	2	4
	***2-of-3* sign (*LCL*=1 , *UCL*=9)**							***1-of-1* X-bar (*LCL* = -*UCL* = -1.011335)**						
Shift	***ARL***	***SDRL***	**5^{th}**	**Q_1**	***MDRL***	**Q_3**	**95^{th}**	***ARL***	***SDRL***	**5^{th}**	**Q_1**	***MDRL***	**Q_3**	**95^{th}**
0	370	363.27	20	106	253	504	1089	370	342.30	23	141	258	504	1058
0.2	46.66	44.73	4	15	33	64	136	159.80	180.02	5.5	46	107	210	416
0.4	11.21	9.55	2	4	8	15	30	42.08	36.77	2.5	13	29	64.5	115
0.6	5.26	3.68	2	3	4	7	13	10.44	8.83	1	3	8	16	24.5
0.8	3.49	1.9	2	2	3	4	7	4.02	3.65	1	2	3	6	9.5
1.0	2.77	1.16	2	2	2	3	5	2.01	1.59	1	1	1	2	5.5
1.2	2.42	0.77	2	2	2	3	4	1.44	0.74	1	1	1	2	3

Table 4.16: In-control and out-of-control properties of the run-length distributions of the two-sided 1-of-1 sign, the *2-of-2* DR sign, the *2-of-2* KL sign, the *2-of-3* sign, the *1-of-1* SR, *2-of-2* DR SR, *2-of-2* KL SR and the *1-of-1* X-bar charts under the Cauchy distribution

	***1-of-1* sign (*LCL*=0 , *UCL* = 10)**							***1-of-1* SR (*LCL* = -*UCL* = -55)**						
Shift	***ARL***	***SDRL***	**5^{th}**	**Q_1**	***MDRL***	**Q_3**	**95^{th}**	***ARL***	***SDRL***	**5^{th}**	**Q_1**	***MDRL***	**Q_3**	**95^{th}**
0.0	370	369.19	19	107	255	513	1111	370	381.92	20	113	268	525	1143
0.2	27.06	26.62	2	8	19	37	80	25.38	24.72	2	8	18	35	75
0.4	7.00	6.43	1	2	5	10	20	6.55	6.06	1	2	5	9	18
0.6	3.83	3.31	1	1	3	5	10	3.61	3.08	1	1	3	5	10
0.8	2.76	2.21	1	1	2	4	7	2.59	1.99	1	1	2	3	7
1.0	2.25	1.68	1	1	2	3	6	2.13	1.54	1	1	2	3	5
1.2	1.97	1.38	1	1	1	2	5	1.88	1.30	1	1	1	2	5
	***2-of-2* sign DR (*LCL*=1 , *UCL*=9)**							***2-of-2* SR DR (*LCL* = -*UCL* = -39)**						
Shift	***ARL***	***SDRL***	**5^{th}**	**Q_1**	***MDRL***	**Q_3**	**95^{th}**	***ARL***	***SDRL***	**5^{th}**	**Q_1**	***MDRL***	**Q_3**	**95^{th}**
0.0	370	361.28	20	106	252	504	1089	370	368.00	20	107	257	512	1099
0.2	18.92	17.59	2	6	14	26	54	24.99	23.54	3	8	18	34	72
0.4	5.08	3.70	2	2	4	7	12	7.28	5.98	2	3	5	10	19
0.6	3.29	1.87	2	2	2	4	7	4.66	3.34	2	2	4	6	11
0.8	2.71	1.25	2	2	2	3	5	3.77	2.42	2	2	3	5	9
1.0	2.46	0.95	2	2	2	3	4	3.33	1.97	2	2	2	4	7
1.2	2.32	0.76	2	2	2	2	4	3.07	1.69	2	2	2	4	6
	***2-of-2* sign KL (*LCL*=1, *UCL* = 9)**							***2-of-2* SR KL (*LCL* = -*UCL* = -37)**						
Shift	***ARL***	***SDRL***	**5^{th}**	**Q_1**	***MDRL***	**Q_3**	**95^{th}**	***ARL***	***SDRL***	**5^{th}**	**Q_1**	***MDRL***	**Q_3**	**95^{th}**
0.0	370	363.78	21	108	257	500	1085	370	369.31	20	107	257	511	1099
0.2	13.04	11.65	2	5	9	18	36	16.61	15.22	2	6	12	22	47
0.4	4.14	2.79	2	2	3	5	10	5.34	4.00	2	2	4	7	13
0.6	2.88	1.45	2	2	2	3	6	3.55	2.20	2	2	3	4	8
0.8	2.46	0.94	2	2	2	3	4	2.96	1.54	2	2	2	4	6
1.0	2.30	0.72	2	2	2	2	4	2.69	1.24	2	2	2	3	5
1.2	2.21	0.58	2	2	2	2	4	2.53	1.04	2	2	2	3	5
	***2-of-3* sign (*LCL*=1 , *UCL*=9)**							***1-of-1* X-bar (*LCL* = -*UCL* = -61.36038)**						
Shift	***ARL***	***SDRL***	**5^{th}**	**Q_1**	***MDRL***	**Q_3**	**95^{th}**	***ARL***	***SDRL***	**5^{th}**	**Q_1**	***MDRL***	**Q_3**	**95^{th}**
0.0	370	361.35	20	106	253	507	1084	370	368.97	19	106	257	510	1103
0.2	12.34	10.65	2	5	9	16	34	371.00	371.53	20	107	257	515	1109
0.4	4.09	2.51	2	2	3	5	9	369.2	368.89	20	107	257	510	1103
0.6	2.90	1.30	2	2	2	3	6	373.1	372.22	19	107	259	517	1116
0.8	2.50	0.86	2	2	2	3	4	369.8	368.44	20	107	258	513	1103
1.0	2.33	0.66	2	2	2	3	4	370.3	371.86	19	106	256	514	1110
1.2	2.23	0.53	2	2	2	2	3	372.7	371.80	20	107	258	517	1114

Table 4.17: Ranking (from most to least favorable) of the two-sided nonparametric charts under the normal, the double exponential and the Cauchy distributions based on out-of-control *ARL* and run-length percentiles. The ARL_0 = 370

Normal	Double Exponential	Cauchy
2-of-2 KL SR	*2-of-2* KL sign / *2-of-3* sign	*2-of-2* KL sign / *2-of-3* sign
2-of-2 KL sign / *2-of-3* sign	*2-of-2* KL SR	*2-of-2* KL SR
1-of-1 $\overline{X}$	*2-of-2* DR sign	*2-of-2* DR sign
2-of-2 DR SR	*2-of-2* DR SR	*2-of-2* DR SR
2-of-2 DR sign	*1-of-1* SR	*1-of-1* SR
1-of-1 SR	*1-of-1* sign	*1-of-1* sign
1-of-1 sign	*1-of-1* $\overline{X}$	*1-of-1* $\overline{X}$

In general, we observe that:

(i) the *2-of-2* DR sign, the *2-of-2* KL sign and the *2-of-3* sign charts all outperform the original *1-of-1* sign chart under all three the distributions, and

(ii) the *2-of-2* KL sign chart and the *2-of-3* sign chart are best overall; only outperformed by the *2-of-2* KL SR chart in case of the normal distribution. (Note: the *2-of-2* KL charts generally outperform the *2-of-2* DR charts; whether the chart is based on the sign test or the SR test).

More specifically, we note that:

(i) under the normal distribution, the two-sided *2-of-2* KL SR chart performs the best (this was also the case for the one-sided charts), but the *2-of-2* KL sign and the *2-of-3* sign charts are good/close competitors, whereas

(ii) under the double exponential distribution and the Cauchy distribution:

(a) the *2-of-2* KL sign and the *2-of-3* sign charts are the top performers,

(b) the sign charts generally perform better than the SR charts except in case of the *1-of-1* chart (i.e. the *1-of-1* SR chart is better than the *1-of-1* sign chart),

(c) the *2-of-2* KL sign and the *2-of-3* sign charts are both better than the *2-of-2* KL SR chart, and

(d) the *2-of-2* DR sign chart is better than the *2-of-2* DR SR chart.

4.3 Precedence charts for the unknown π^{th} quantile (Case U)

Introduction

Case U is the scenario when the π^{th} percentile of the process distribution is unknown or unspecified; this is unlike Case K and as a consequence the control limits are unknown.

To estimate the control limits a reference sample is obtained; this reference sample is also called the preliminary sample or the calibration sample or the Phase I sample. Once the control limits are estimated, Phase II starts. In Phase II the estimated control limits are used for future monitoring of the process using new incoming samples taken sequentially from the process; this is the prospective monitoring phase.

The new control charts we consider here, in Case U, are based on the median test, which is essentially a modified sign test for two independent samples and is a member of a more general class of nonparametric two-sample tests referred to as precedence tests or precedence statistics (see e.g. Gibbons and Chakraborti, (2003)). We therefore refer to the charts of Case U as precedence charts.

Assumptions

We assume that

(i) the reference sample $(X_1, X_2, ..., X_m)$ is a random sample of size m available from an in-control (IC) distribution with an unknown continuous cumulative distribution function (c.d.f) $F_X(x) = F(x-\theta)$ where θ is the location parameter and F is some continuous c.d.f with median zero,

(ii) each Phase II test sample $(Y_{i1}, Y_{i2}, ..., Y_{in})$ taken at sampling stage (time) $i = 1,2,3,...$ is a random sample (rational subgroup) of size $n > 1$ from an unknown continuous distribution with c.d.f $G_Y(y) = F(y-\theta_i)$ where θ_i is the location parameter of the i^{th} test sample, and

(iii) the Phase II test samples are drawn sequentially and independently of one another and of the reference sample.

Charting statistics and control limits

The control limits, as mentioned before, are estimated from the Phase I reference sample and then used for prospective monitoring of the process. In Phase II, one charting statistic is calculated from each new incoming sample and then compared to the estimated control limits.

The estimated control limits are found by arranging the Phase I observations in ascending order, that is,

$$X_{1:m} < X_{2:m} < \ldots < X_{m:m}$$

where $X_{j:m}$ denotes the j^{th} order statistic of the reference sample of size m, and selecting two order statistics $X_{a:m}$ and $X_{b:m}$ (for a given $1 \le a < b \le m$) so that the estimated control limits for the two-sided precedence charts are given by

$$\hat{LCL} = X_{a:m} \qquad \text{and} \qquad \hat{UCL} = X_{b:m} \tag{4-28}$$

respectively, where a and b are labeled the charting constants; determination of the charting constants will be discussed later. Note that, like the sign charts of Case K, the precedence charts do not have a centerline.

The charting statistic at time $i = 1,2,3,\ldots$ is an order statistic $Y_{j:n}^{i}$ for $1 < j \le n$ from each of the Phase II test samples.

The operation and the signaling rules (i.e. when a process is declared OOC) of the runs-rule enhanced precedence charts is similar to that of the sign charts; however, instead of comparing T_i (the sign statistic) with the known control limits UCL and LCL of (4-7) we now compare $Y_{j:n}^{i}$ with the estimated control limits $\hat{UCL}$ and $\hat{LCL}$ of (4-28) at each sampling stage $i = 1,2,3,\ldots$.

Remark 11

(i) The median is a robust and flexible estimator of location in the sense that it is preferred in situations where large measurement errors are expected and is applicable in more diverse situations (unlike the mean). Thus, although we develop and discuss the theory of the *1-of-1* and the runs-rule enhanced precedence charts so that any order statistic can be used as charting statistic, the median is a popular choice in practice and we therefore focus mainly on the median chart, that is, the case where the charting statistic is taken as the test sample median.

Furthermore, to simplify matters we assume that the sample size $n = 2s+1$ is odd so that the median of the test sample $Y^i_{j:n}$ is uniquely defined with $j = s+1$. Thus, for example, when the subgroup size n is equal to 5, as is fairly common in SPC applications, the charting statistic is the 3rd smallest value in the test sample.

(ii) Only two-sided precedence charts are studied. The required modifications for the one-sided precedence charts are simple and briefly indicated in section 4.3.3.

(iii) The proposed precedence charts do not signal unless the charting statistic $Y^i_{j:n}$ is less than or equal to the estimated lower control limit $X_{a:m}$ or is greater than or equal to the estimated upper control limit $X_{b:m}$. Although this is theoretically negligible as the underling process distributions are assumed to be continuous, in practice, one needs to apply the charts in a correct manner as ties might occur when it is found that the Y order statistic (i.e. charting statistic) is equal to one of the control limits.

(iv) The precedence charts can be applied as soon as the necessary order statistic is available and can be a practical advantage in some applications. We comment more on this point later.

4.3.1 Run-length distributions of the two-sided precedence charts

The run-length distributions and the statistical characteristics of the precedence charts (such as the ARL, $VARL$ etc.) are required to design the charts and reveal important information regarding their performance.

We again use a Markov chain approach to derive the run-length distributions and in some cases draw on the results of the geometric distribution of order k to obtain closed form expressions.

Even though the operation (i.e. the signaling events, when a process is declared OOC etc.) of the runs-rule enhanced precedence charts of Case U are similar to that of the runs-rule enhanced sign charts of Case K, there is a fundamental difference in deriving the run-length distributions of the precedence charts compared to that of the sign charts.

In particular, because the control limits are estimated they are random variables (as indicated by the ^ - notation in (4-28)) and, consequently, the signaling indicators of the (two-sided) precedence charts i.e.

$$\hat{\xi}_i = \begin{cases} 1 & \text{if} \quad Y_{j:n}^i \geq U\hat{C}L \\ 0 & \text{if} \quad L\hat{C}L < Y_{j:n}^i < U\hat{C}L \\ 2 & \text{if} \quad Y_{j:n}^i \leq L\hat{C}L \end{cases} \tag{4-29}$$

and

$$\hat{\xi}_i^{\pm} = I(Y_{j:n}^i \notin (L\hat{C}L, U\hat{C}L)) = \begin{cases} 1 & \text{if} \quad Y_{j:n}^i \notin (L\hat{C}L, U\hat{C}L) \\ 0 & \text{if} \quad Y_{j:n}^i \in (L\hat{C}L, U\hat{C}L) \end{cases} \tag{4-30}$$

for $i = 1,2,3,...$ are dependent tri-variate (or binary) random variables.

The design, analysis and performance of the charts must therefore take account of the additional variability introduced as a result of estimating the control limits; this is the main stumbling block in calculating the run-length distributions, here, in Case U, particularly for the charts that use signaling rules (ii) and (iii) defined in the beginning of section 4.1 on page 258. Like in Chapter 3 we use a two-step approach to derive the run-length distribution which involves the method of conditioning (see e.g. Chakraborti, (2000)).

First we derive the *conditional* run-length distributions i.e. conditioned on the two order statistics (control limits), which lets us focus on specific values of the control limits. The performance of the charts as measured by their *conditional* run-length distributions are therefore different for each user as each user has his/her own control limits based on his/her own Phase I data (sample).

Second we derive the *unconditional* (or marginal) run-length distributions by averaging over the joint distribution of the two order statistics. The *unconditional* run-length distributions reflect the bigger picture or the overall performance of the charts and take into account that the control limits are estimated. The performance of the charts as measured by their *unconditional* run-length distributions are therefore the same for each user.

Signaling probabilities

The key ingredients to the *conditional* run-length distributions are

(i) the one-step transition probabilities $p_{i,k}$ and

(ii) the success probability α.

The one-step transition probabilities are the elements of the transition probability matrix and are required in case one uses the Markov chain approach. The success probability, on the other hand, is a prerequisite if one wishes to use the properties of the geometric distribution of order k as it is a parameter of the distribution (see e.g. expressions (4-16) and (4-17)).

The one-step transition probabilities and the success probability all depend on and are functions of the conditional probability of a signal i.e. the probability for a charting statistic to plot OOC given that (or conditionally on having observed) $X_{a:m} = x_{a:m}$ and $X_{b:m} = x_{b:m}$, which is given by

$$\begin{aligned} p_C^{\pm}(X_{a:m}, X_{b:m}, F, G) &= \Pr(\hat{\xi}_i^{\pm} = 1 \mid \hat{LCL} = x_{a:m}, \hat{UCL} = x_{b:m}) \\ &= 1 - \Pr(X_{a:m} < Y_{j:n}^i < X_{b:m} \mid X_{a:m} = x_{a:m}, X_{b:m} = x_{b:m}) \\ &= 1 - G_j(X_{b:m}) + G_j(X_{a:m}) \end{aligned} \qquad (4\text{-}31)$$

where G_j denotes the c.d.f of the j^{th} order statistic in a sample of size n from a distribution with c.d.f G and the subscript "C" in $p_C^{\pm}(X_{a:m}, X_{b:m}, F, G)$ indicates that (4-31) is a conditional probability.

Using the probability integral transformation (PIT) and the fact that the j^{th} order statistic from a *uniform*(0,1) distribution follows a beta distribution with parameters j and $n - j + 1$ (see e.g. Gibbons and Chakraborti, (2003)) it follows, for example, that

$$\begin{aligned} G_j(X_{a:m}) &= \Pr(Y_{j:n}^i \le X_{a:m} \mid X_{a:m} = x_{a:m}) \\ &= \Pr(G(Y_{j:n}^i) \le G(X_{a:m}) \mid X_{a:m} = x_{a:m}) \\ &= \Pr(U_{j:n}^i \le G(X_{a:m})) = I_{G(X_{a:m})}(j, n - j + 1) \end{aligned}$$

where $U_{j:n}^i$ is the j^{th} order statistic from a *uniform*(0,1) distribution and

$$I_p(u,v) = [\beta(u,v)]^{-1} \int_0^p w^{u-1}(1-w)^{v-1} dw \quad \text{for} \quad u, v > 0$$

is the c.d.f of the $Beta(u,v)$ distribution, also known as the incomplete beta function.

Thus, the conditional probability of a signal in (4-31) can be expressed as

$$p_C^{\pm}(X_{a:m}, X_{b:m}, F, G) = 1 - I_{G(X_{b:m})}(j, n-j+1) + I_{G(X_{a:m})}(j, n-j+1). \quad (4\text{-}32)$$

With the conditional probability of a signal in (4-32) we can without difficulty find the *conditional* run-length distributions of the two-sided *1-of-1*, the *2-of-2* DR, the *2-of-2* KL and the *2-of-3* precedence charts. The *unconditional* run-length distributions, in general, follow straightforwardly from the *conditional* run-length distributions.

Remark 12

(i) We denote, without loss of generality, the two order statistics $(X_{a:m}, X_{b:m})$ by $\mathbf{Z} = (X, Y)$ and their observed values $(x_{a:m}, x_{b:m})$ by $\mathbf{z} = (x, y)$. Thus when writing $\mathbf{Z} = \mathbf{z}$ it means $(X_{a:m}, X_{b:m}) = (x_{a:m}, x_{b:m})$ or $(X, Y) = (x, y)$.

In particular, this notation permits us to write (4-32) as

$$p_C^{\pm}(X, Y, F, G) = 1 - I_{G(Y)}(j, n-j+1) + I_{G(X)}(j, n-j+1). \tag{4-33}$$

(ii) It is instructive to compare the signaling probability of the two-sided sign chart of Case K with that of the two-sided precedence chart of Case U.

Specifically, we note that by substituting $G(X)$ and $G(Y)$ for p, replacing a with $j-1$ and swapping b for $n-j$ in (4-13) we obtain (4-33).

(iii) Because $\{\hat{\xi}_i^{\pm} = 1\}$ if and only if $\{\hat{\xi}_i = 1\} \cup \{\hat{\xi}_i = 2\}$ we can re-express the conditional probability of a signal of the two-sided precedence chart in terms of that of the upper and the lower one-sided charts i.e.

$$p_C^{\pm}(X, Y, F, G) = p_C^{-}(X, F, G) + p_C^{+}(Y, F, G) \tag{4-34}$$

where

$$p_C^{-}(X, F, G) = \Pr(\hat{\xi}_i = 2 \mid X_{a:m} = x_{a:m}) = I_{G(X)}(j, n-j+1) \tag{4-35}$$

and

$$p_C^{+}(Y, F, G) = \Pr(\hat{\xi}_i = 1 \mid X_{b:m} = x_{b:m}) = 1 - I_{G(Y)}(j, n-j+1). \tag{4-36}$$

Expression (4-34) will be particularly useful when deriving the run-length distributions of the two-sided precedence charts via the Markov chain approach.

(iv) For notational simplicity and brevity we denote $p_C^{\pm}(X, Y, F, G)$, $p_C^{-}(X, F, G)$ and $p_C^{+}(Y, F, G)$ simply by $p_C^{\pm}$, p_C^{-} and p_C^{+}, respectively.

4.3.1.1 Distribution of N_{1of1}: Run-length distribution for the *1-of-1* precedence chart

The two-sided *1-of-1* precedence chart was studied in detail by Chakraborti et al. (2004); this chart is called the "basic" precedence chart. The authors derived explicit formulae for both the conditional and the unconditional run-length distributions and their associated statistical characteristics (such as the ARL, $VARL$, FAR etc.) by applying, amongst numerous other techniques, results of the geometric distribution of order $k=1$ coupled with the method of conditioning (expectation by conditioning); doing so they have taken proper account of the dependency between the Phase II signaling events. In the paragraphs that follow, we simply review the most important statistical characteristics of the *1-of-1* precedence chart; for complete details on the derivations of the results, see the original article by Chakraborti et al. (2004).

Conditional run-length distribution

In particular, Chakraborti et al. (2004) showed that given $\mathbf{Z}=\mathbf{z}$ the *conditional* distribution of the run-length N_{1of1} is geometric with parameter (success probability) $p_C^{\pm}=p_C^{\pm}(X,Y,F,G)$. Accordingly, all properties and characteristics of the conditional run-length distribution follow conveniently from the properties of the geometric distribution of order $k=1$.

For example, the conditional p.m.f of N_{1of1} is

$$\Pr(N_{1of1}=t\mid\mathbf{Z}=\mathbf{z})=(1-p_C^{\pm})^{t-1}p_C^{\pm}\quad\text{for}\quad t=1,2,3,...$$

whereas the conditional average run-length ($CARL$) and the conditional variance of the run-length ($CVARL$) are given by

$$CARL_{1of1}=E(N_{1of1}\mid\mathbf{Z}=\mathbf{z})=1/p_C^{\pm}\quad\text{and}\quad CVARL_{1of1}=\text{var}(N_{1of1}\mid\mathbf{Z}=\mathbf{z})=(1-p_C^{\pm})/(p_C^{\pm})^2$$

respectively.

The conditional false alarm rate also follows straightforwardly as it is found by substituting $F=G$ in $p_C^{\pm}$, that is,

$$CFAR_{1of1}=p_C^{\pm}(X,Y,F,F)=1-I_{F(Y)}(j,n-j+1)+I_{F(X)}(j,n-j+1).$$

Unconditional run-length distribution

Most importantly Chakraborti et al. (2004) showed that by averaging over the joint distribution of $\mathbf{Z}$ one obtains the *unconditional* or marginal run-length distribution and its associated characteristics.

The unconditional p.m.f of N_{1of1}, in particular, is given by

$$\Pr(N_{1of1}=t)=E_{\mathbf{Z}}(\Pr(N_{1of1}=t\mid\mathbf{Z}))=E_{\mathbf{Z}}\left((q_C^{\pm})^{t-1}(1-q_C^{\pm})\right)=D^*(t-1)-D^*(t) \quad (4\text{-}37)$$

for $t=1,2,3,...$ and $D^*(0)=1$ where

$$D^*(t)=E_{\mathbf{Z}}\left((q_C^{\pm})^t\right)=\int_0^1\int_0^y\left(\frac{1}{\beta(j,n-j+1)}\sum_{h=0}^{n-j}\frac{(-1)^h}{j+h}\binom{n-j}{h}\left(GF^{-1}(y)^{j+h}-GF^{-1}(x)^{j+h}\right)\right)^t f_{a,b}(x,y)dxdy$$

with $q_C^{\pm}=1-p_C^{\pm}$ and where $f_{a,b}(x,y)$ denotes the joint p.d.f of the a^{th} and the b^{th} order statistics in a reference sample of size m from the *uniform*(0,1) distribution, given by

$$f_{a,b}(x,y)=\frac{m!}{(a-1)!(b-a-1)!(m-b)!}x^{a-1}(y-x)^{b-a-1}(1-y)^{m-b} \qquad 0<x<y<1.$$

Likewise, by writing the conditional ARL as

$$CARL_{1of1}=\sum_{t=0}^{\infty}\Pr(N_{1of1}>t\mid\mathbf{Z}=\mathbf{z})=\sum_{t=0}^{\infty}(q_C^{\pm})^t$$

the unconditional average run-length ($UARL_{1of1}$) follows by averaging over the joint distribution of $\mathbf{Z}$ and then simplifying i.e.

$$\begin{aligned}UARL_{1of1}&=E_{\mathbf{Z}}(CARL_{1of1})=\sum_{t=0}^{\infty}E_{\mathbf{Z}}\left((q_C^{\pm})^t\right)=\sum_{t=0}^{\infty}D^*(t)\\&=\int_0^1\int_0^y\left(1-\frac{1}{\beta(j,n-j+1)}\sum_{h=0}^{n-j}\frac{(-1)^h}{j+h}\binom{n-j}{h}\left(GF^{-1}(y)^{j+h}-GF^{-1}(x)^{j+h}\right)\right)^{-1}f_{a,b}(x,y)dxdy.\end{aligned} \quad (4\text{-}38)$$

The unconditional probability of a signal follows in the same manner and is given by

$$p^{\pm}(F,G) = \Pr(\hat{\xi}_i^{\pm} = 1) = E_{\mathbf{Z}}(\Pr(\hat{\xi}_i^{\pm} = 1 \mid \mathbf{Z})) = E_{\mathbf{Z}}(p_C^{\pm}(x,y,F,G))$$
$$= \int_0^1 \int_0^y (1 - I_{GF^{-1}(y)}(j, n-j+1) + I_{GF^{-1}(x)}(j, n-j+1))\, f_{a,b}(x,y)dxdy. \quad (4\text{-}39)$$

For complete details on the derivation of expressions (4-37), (4-38) and (4-39) see Chakraborti et al. (2004).

Remark 13

Chakraborti et al. (2004) noted that, from equations (4-37), (4-38) and (4-39), "… it is evident that in general the run length distribution depends on the distribution functions F and G through the composite function $\psi = GF^{-1}$. For example, when $F = G$, the process is in control, so $\psi(u) = u$, and the in-control run length distribution follows …".

In particular, the in-control unconditional p.m.f is given by

$$\Pr(N_{1of1,0} = t) = \Pr_{F=G}(N_{1of1} = t) = D(t-1) - D(t) \quad \text{for} \quad t = 1,2,3,... \quad \text{and} \quad D(0) = 1$$

where $N_{1of1,0}$ denotes the in-control run-length random variable and

$$D(t) = \int_0^1 \int_0^y \left(\frac{1}{\beta(j, n-j+1)} \sum_{h=0}^{n-j} \frac{(-1)^h}{j+h} \binom{n-j}{h} \left(y^{j+h} - x^{j+h}\right) \right)^t f_{a,b}(x,y)dxdy$$

whereas the unconditional in-control average run-length ($UARL_{1of1,0}$) and the unconditional false alarm rate ($UFAR_{1of1}$) follows from (4-38) and (4-39) and given by

$$UARL_{1of1,0} = \int_0^1 \int_0^y \left(1 - \frac{1}{\beta(j, n-j+1)} \sum_{h=0}^{n-j} \frac{(-1)^h}{j+h} \binom{n-j}{h} (y^{j+h} - x^{j+h}) \right)^{-1} f_{a,b}(x,y)dxdy$$

and

$$UFAR_{1of1} = p^{\pm}(F,F) = \int_0^1 \int_0^y \left(1 - I_y(j, n-j+1) + I_x(j, n-j+1)\right) f_{a,b}(x,y)dxdy,$$

respectively.

4.3.1.2 Distribution of N_{2of2}^{DR}: Run-length distribution for the *2-of-2* DR precedence chart

As pointed out in section 4.3.1.1, in Chakraborti et al. (2004) the idea of conditioning on the reference sample order statistics $\mathbf{Z}$ was effectively used to derive the distribution of the run-length N_{1of1} of the two-sided *1-of-1* precedence chart and to study various properties of their chart in a convenient way. Using the same conditioning idea we derive the *conditional* and the *unconditional* run-length distributions of the two-sided *2-of-2* DR chart.

Conditional run-length distribution

Given that $\mathbf{Z}=\mathbf{z}$ the sequence of signaling indicators $\hat{\xi}_1^{\pm},\hat{\xi}_2^{\pm},\hat{\xi}_3^{\pm},...$ in (4-30) are i.i.d. Bernoulli random variables with success probability $p_C^{\pm}=\Pr(I(Y_{j:n}^{i}\notin(L\hat{C}L,U\hat{C}L))=1\,|\,\mathbf{Z}=\mathbf{z})$. Thus, conditionally on the order statistics $\mathbf{Z}$ the run-length N_{2of2}^{DR} of the two-sided *2-of-2* DR chart follows a geometric distribution of order two.

The conditional p.m.f of N_{2of2}^{DR} is therefore given by (4-16) with $\alpha=p_C^{\pm}$ and $k=2$ i.e.

$$\Pr(N_{2of2}^{DR}=t\,|\,\mathbf{Z}=\mathbf{z})=\begin{cases}0 & \text{if} \quad 0\le t<2\\ (p_C^{\pm})^2 & \text{if} \quad t=2\end{cases} \tag{4-40a}$$

and for $t\ge 3$

$$\Pr(N_{2of2}^{DR}=t\,|\,\mathbf{Z}=\mathbf{z})=\sum_{i=1}^{[\frac{t+1}{3}]}(-1)^{i-1}\frac{(p_C^{\pm})^{2i}}{(1-p_C^{\pm})^{1-i}}\left\{\binom{t-2i-1}{i-2}+(1-p_C^{\pm})\binom{t-2i-1}{i-1}\right\} \tag{4-40b}$$

whereas the conditional average run-length (expected value or mean) and the conditional variance of the run-length can be found from (4-17) and given by

$$CARL_{2of2}^{DR}=E(N_{2of2}^{DR}\,|\,\mathbf{Z}=\mathbf{z})=\frac{1+p_C^{\pm}}{(p_C^{\pm})^2} \tag{4-41}$$

and

$$CVARL_{2of2}^{DR}=\mathrm{var}(N_{2of2}^{DR}\,|\,\mathbf{Z}=\mathbf{z})=\frac{1-5(1-p_C^{\pm})(p_C^{\pm})^2-(p_C^{\pm})^5}{(1-p_C^{\pm})^2(p_C^{\pm})^4}, \tag{4-42}$$

respectively (see e.g. Remark 5).

Unconditional run-length distribution

The complexity of the *conditional* distribution in (4-40), particularly for $t \geq 3$, makes a direct application of conditioning to derive a closed form expression for the *unconditional* distribution of N_{2of2}^{DR} unattractive. Instead, we find the *unconditional* distribution of N_{2of2}^{DR} by first conditioning on the total number of successes $S_n = \sum_{i=1}^{n} \hat{\xi}_i^{\pm}$ in the sequence of *n* random variables $\hat{\xi}_1^{\pm}, \hat{\xi}_2^{\pm}, ..., \hat{\xi}_n^{\pm}$. (Note that, here, n is the number of random variables and not the sample size.)

To this end, note that, although $\hat{\xi}_1^{\pm}, \hat{\xi}_2^{\pm}, ..., \hat{\xi}_n^{\pm}$ is a sequence of dependent binary random variables they are exchangeable or symmetrically dependent; this means that any permutation of any subset of these random variables has the same distribution; this can be written as

$$\Pr(\hat{\xi}_{\pi(1)}^{\pm} = 1, ..., \hat{\xi}_{\pi(u)}^{\pm} = 1) = \Pr(\hat{\xi}_1^{\pm} = 1, ..., \hat{\xi}_u^{\pm} = 1) \tag{4-43}$$

for any permutation $\pi(1), ..., \pi(u)$ of $1,2,...,u \leq n$. Using (4-43) we can derive an exact closed form expression for the *unconditional* p.m.f of N_{2of2}^{DR}.

George and Bowman (1995) derived the distribution of the total number of successes S_n in a sequence of n exchangeable binary trials. According to their result

$$\Pr(S_n = s) = \binom{n}{s} \sum_{i=0}^{n-s} (-1)^i \binom{n-s}{i} \lambda_{s+i} \quad \text{for} \quad s = 1,2,...,n \tag{4-44}$$

where

$$\lambda_u = \Pr(\hat{\xi}_1^{\pm} = 1, ..., \hat{\xi}_u^{\pm} = 1) \quad \text{for} \quad u = 1,2,..,n. \tag{4-45}$$

Using (4-44) the unconditional distribution of N_{2of2}^{DR} is given by

$$\Pr(N_{2of2}^{DR} = t) = \begin{cases} 0 & \text{if } 0 \leq t < 2 \\ \lambda_2 & \text{if } t = 2 \end{cases} \tag{4-46a}$$

and for $t \geq 3$

$$\Pr(N_{2of2}^{DR} = t) = \sum_{y=1}^{t-2} \sum_{k=0}^{\min\left(y, \left[\frac{t-y-2}{2}\right]\right)} \sum_{i=0}^{y} (-1)^k (-1)^i \binom{y}{k} \binom{y}{i} \binom{t-2(k+1)-1}{y-1} \lambda_{t-y+i} \tag{4-46b}$$

The proof of (4-46) is straightforward for $t \le 2$ i.e

$$\Pr(N_{2of2}^{DR} = 2) = \Pr(\hat{\xi}_1^{\pm} = 1, \hat{\xi}_2^{\pm} = 1) = \lambda_2$$

where λ_2 is defined in (4-45).

For $t \ge 3,$ we write the unconditional distribution of N_{2of2}^{DR} as

$$\Pr(N_{2of2}^{DR} = t) = \sum_{y=1}^{t-2} \Pr(N_{2of2}^{DR} = t \mid S_t = t - y) \Pr(S_t = t - y) \qquad (4\text{-}47)$$

and then first consider the conditional probability $\Pr(N_{2of2}^{DR} = t \mid S_t = t - y)$.

By de Finetti's theorem a sequence of exchangeable random variables is conditionally i.i.d.. Hence, the conditional distribution of N_{2of2}^{DR} given the number of successes in exchangeable binary random variables is the same as that for a sequence of i.i.d. binary variables; this latter distribution has been worked out in the literature (see e.g. Balakrishnan and Koutras (2002), p 56; note a typo) and is given by

$$\Pr(N_{2of2}^{DR} = t \mid S_t = t - y) = \binom{t}{y}^{-1} \sum_{k=0}^{[\frac{t-y-2}{2}]} (-1)^k \binom{y}{k} \binom{t - 2(k+1) - 1}{y - 1}. \qquad (4\text{-}48)$$

Now, using (4-44) we have

$$\Pr(S_t = t - y) = \binom{t}{y} \sum_{i=0}^{y} (-1)^i \binom{y}{i} \lambda_{t-y+i},$$

so that (4-46) follows by substituting $\Pr(S_t = t - y)$ and (4-48) in (4-47).

Remark 14

(i) Conditionally on the reference sample order statistics that define the Phase II control limits we have that

$$\lambda_u^C = \Pr(\hat{\xi}_1^{\pm} = 1,\ldots,\hat{\xi}_u^{\pm} = 1 \mid \mathbf{Z} = \mathbf{z}) = \prod_{i=1}^{u} \Pr(\hat{\xi}_i^{\pm} = 1 \mid \mathbf{Z} = \mathbf{z}) = (p_C^{\pm})^u$$

so that the unconditional probability λ_u in (4-45) equals

$$\begin{aligned}\lambda_u &= \Pr(\hat{\xi}_1^{\pm} = 1,\ldots,\hat{\xi}_u^{\pm} = 1) = E_{\mathbf{Z}}(\lambda_u^C) = E_{\mathbf{Z}}((p_C^{\pm})^u) \\ &= \int_0^1 \int_0^y \left(1 - I_{GF^{-1}(y)}(j, n-j+1) + I_{GF^{-1}(x)}(j, n-j+1)\right)^u f_{a,b}(x,y)dxdy.\end{aligned} \tag{4-49}$$

(ii) The run-length distribution of the two-sided *2-of-2* DR chart depends on the distribution functions F and G through the composite function $\psi = GF^{-1}$ present in λ_u (see expression (4-49)). Thus, the in-control run-length distribution is obtained by substituting $\lambda_{u,0}$ in (4-46a) and (4-46b) where

$$\lambda_{u,0} = \int_0^1 \int_0^y \left(1 - I_y(j, n-j+1) + I_x(j, n-j+1)\right)^u f_{a,b}(x,y)dxdy \tag{4-50}$$

and is found from (4-49) by substituting $F = G$. It is evident from (4-50) that the in-control run-length distribution of the two-sided *2-of-2* DR (like that of the *1-of-1* chart) is free from either *F* or *G* and that the *2-of-2* DR chart is thus distribution-free.

(iii) The unconditional false alarm rate, the unconditional average run-length etc of the two-sided *2-of-2* DR chart is calculated later in section 4.3.2.

4.3.1.3 Distribution of N_{2of2}^{KL} and N_{2of3}: Run-length distribution for the two-sided *2-of-2* KL and *2-of-3* precedence charts

In sections 4.3.1.1 and 4.3.1.2 we illustrated how to find the *conditional* and *unconditional* distributions of N_{1of1} and N_{2of2}^{DR} via the geometric distribution of order k (= 1 or 2). Here, in section 4.3.1.3, we illustrate how to find the *conditional* and *unconditional* distributions of N_{2of2}^{KL} and N_{2of3} via the Markov chain approach. The conditional and unconditional distributions of N_{1of1} and N_{2of2}^{DR}, via the Markov chain approach, can be found in a similar manner and is not shown here.

The Markov chain approach for finding the *conditional* run-length distributions of the (two-sided) precedence charts in Case U is similar to those of the sign charts in Case K. In particular, the state spaces are identical so that we merely substitute:

(i) $p_C^+(Y,F,G)$ (defined in (4-36)) for $p^+(n,b,\theta)$, and

(ii) $p_C^-(X,F,G)$ (defined in (4-35)) for $p^-(n,a,\theta)$

in any one of the essential transition probability matrices of the two-sided sign charts (i.e. the $\mathbf{Q}_{h\times h}$'s given in (4-21), (4-22), (4-26) etc.) to obtain the *conditional* essential transition probability matrices $\mathbf{Q}_{h\times h}^C$ (say) of the precedence charts. Note that, here, in Case U, the superscript "C" in $\mathbf{Q}_{h\times h}^C$ indicates that we work with a *conditional* essential transition probability matrix i.e. conditioned on the order statistics $\mathbf{Z}$.

Upon substituting $\mathbf{Q}_{h\times h}^C$ into (4-8), (4-9) and (4-10) we obtain the *conditional* p.m.f, the *conditional ARL* (*CARL*) and the *conditional VARL* (*CVARL*), respectively. The *unconditional* run-length distributions and the associated *unconditional* characteristics of the precedence chart is then found by averaging over the distribution of $\mathbf{Z}$.

Conditional distributions of N_{2of2}^{KL} and N_{2of3}

Conditional on $\mathbf{Z}$ the sequence of signaling indicators $\hat{\xi}_1, \hat{\xi}_2, \hat{\xi}_3, \ldots$ in (4-29) are i.i.d. tri-variate random variables with:

(i) $\Pr(\hat{\xi}_i = 2 \mid \mathbf{Z} = \mathbf{z}) = p_C^-$,

(ii) $\Pr(\hat{\xi}_i = 1 \mid \mathbf{Z} = \mathbf{z}) = p_C^+$ and

(iii) $\Pr(\hat{\xi}_i = 0 \mid \mathbf{Z} = \mathbf{z}) = 1 - p_C^- - p_C^+$

respectively.

Thus, *conditional* on $\mathbf{Z}$ the run-length distribution of the two-sided *2-of-2* KL precedence chart is

$$\Pr(N_{2of2}^{KL} = t \mid \mathbf{Z} = \mathbf{z}) = \boldsymbol{\xi}(\mathbf{Q}_{4\times4}^{C})^{t-1}(\mathbf{I} - \mathbf{Q}_{4\times4}^{C})\mathbf{1} \quad \text{for} \quad t = 1,2,3,\ldots \tag{4-51}$$

with the *conditional* average run-length and the *conditional* variance of the run-length given by

$$CARL_{2of2}^{KL} = E(N_{2of2}^{KL} \mid \mathbf{Z} = \mathbf{z}) = \boldsymbol{\xi}(\mathbf{I} - \mathbf{Q}_{4\times4}^{C})^{-1}\mathbf{1} \tag{4-52}$$

and

$$CVARL_{2of2}^{KL} = \operatorname{var}(N_{2of2}^{KL} \mid \mathbf{Z} = \mathbf{z}) = \boldsymbol{\xi}(\mathbf{I} + \mathbf{Q}_{4\times4}^{C})(\mathbf{I} - \mathbf{Q}_{4\times4}^{C})^{-2}\mathbf{1} - (CARL_{2of2}^{KL})^2 \tag{4-53}$$

respectively, where

$$\mathbf{Q}_{4\times4}^{C} = \begin{bmatrix} 0 & 1-p_C^- - p_C^+ & p_C^+ & p_C^- \\ 0 & 1-p_C^- - p_C^+ & p_C^+ & p_C^- \\ 0 & 1-p_C^- - p_C^+ & 0 & p_C^- \\ 0 & 1-p_C^- - p_C^+ & p_C^+ & 0 \end{bmatrix} \tag{4-54}$$

denotes the *conditional* essential transition probability matrix of the *2-of-2* KL chart and follows from (4-26) having substituted the conditional probabilities p_C^- and p_C^+ (defined in (4-35) and (4-36)) for p^- and p^+, respectively.

Likewise, the *conditional* p.m.f, the *conditional* average run-length and the *conditional* variance of the two-sided *2-of-3* precedence chart are

$$\Pr(N_{2of3} = t \mid \mathbf{Z} = \mathbf{z}) = \boldsymbol{\xi}(\mathbf{Q}_{8\times8}^{C})^{t-1}(\mathbf{I} - \mathbf{Q}_{8\times8}^{C})\mathbf{1} \quad \text{for} \quad t = 1,2,3,\ldots, \tag{4-55}$$

$$CARL_{2of3} = E(N_{2of3} \mid \mathbf{Z} = \mathbf{z}) = \boldsymbol{\xi}(\mathbf{I} - \mathbf{Q}_{8\times8}^{C})^{-1}\mathbf{1}, \tag{4-56}$$

and

$$CVARL_{2of3} = \text{var}(N_{2of3} \mid \mathbf{Z} = \mathbf{z}) = \boldsymbol{\xi}(\mathbf{I} + \mathbf{Q}_{8\times8}^{C})(\mathbf{I} - \mathbf{Q}_{8\times8}^{C})^{-2}\mathbf{1} - (CARL_{2of3})^{2} \tag{4-57}$$

respectively, where the *conditional* essential transition probability matrix of the *2-of-3* chart is given by

$$\mathbf{Q}_{8\times8}^{C} = \begin{bmatrix}
0 & 1-p_C^- - p_C^+ & p_C^+ & p_C^- & 0 & 0 & 0 & 0 \\
0 & 1-p_C^- - p_C^+ & 0 & 0 & p_C^+ & 0 & p_C^- & 0 \\
0 & 0 & p_C^+ & p_C^- & 0 & 1-p_C^- - p_C^+ & 0 & 0 \\
0 & 0 & p_C^+ & p_C^- & 0 & 0 & 0 & 1-p_C^- - p_C^+ \\
0 & 0 & 0 & p_C^- & 0 & 1-p_C^- - p_C^+ & 0 & 0 \\
0 & 1-p_C^- - p_C^+ & 0 & 0 & 0 & 0 & p_C^- & 0 \\
0 & 0 & p_C^+ & 0 & 0 & 0 & 0 & 1-p_C^- - p_C^+ \\
0 & 1-p_C^- - p_C^+ & 0 & 0 & p_C^+ & 0 & 0 & 0
\end{bmatrix} \tag{4-58}$$

and follows from (4-27). In particular, expressions (4-55), (4-56) and (4-57) follow from having substituted $\mathbf{Q}_{8\times8}^{C}$ in (4-8), (4-9) and (4-10), respectively.

Remark 15

The *conditional ARL* expressions in (4-52) and (4-56) have been symbolically simplified and closed form expressions are given in Table 4.5; however, here, in Case U, we substitute p_C^+ for p^+ and p_C^- for p^-, respectively. Closed form expressions of the *conditional VARL*'s in (4-53) and (4-57) can be obtained in a similar manner (i.e. simplifying the expressions symbolically).

Unconditional distributions of N_{2of2}^{KL} and N_{2of3}

The unconditional distributions of N_{2of2}^{KL} and N_{2of3} are obtained by averaging the conditional distributions given in (4-51) and (4-55) over the distribution of $\mathbf{Z}$ i.e.

$$\begin{aligned}\Pr(N_{2of2}^{KL}=t) &= \int_0^1\int_0^y \Pr(N_{2of2}^{KL}=t\mid \mathbf{Z})\, f_{a,b}(x,y)dxdy \\ &= \int_0^1\int_0^y \xi(\mathbf{Q}_{4\times4}^{C})^{t-1}(\mathbf{I}-\mathbf{Q}_{4\times4}^{C})\mathbf{1}\, f_{a,b}(x,y)dxdy\end{aligned} \tag{4-59}$$

and

$$\begin{aligned}\Pr(N_{2of3}=t) &= \int_0^1\int_0^y \Pr(N_{2of3}=t\mid \mathbf{Z})\, f_{a,b}(x,y)dxdy \\ &= \int_0^1\int_0^y \xi(\mathbf{Q}_{8\times8}^{C})^{t-1}(\mathbf{I}-\mathbf{Q}_{8\times8}^{C})\mathbf{1}\, f_{a,b}(x,y)dxdy\end{aligned} \tag{4-60}$$

for $t=1,2,3,\ldots.$

Remark 16

The *unconditional* or marginal distributions of N_{2of2}^{KL} and N_{2of3} in (4-59) and (4-60) depend on the distribution functions F and G through the composite function $\psi = GF^{-1}$ present in both

$$p_C^- = I_{GF^{-1}(X)}(j, n-j+1) \quad \text{and} \quad p_C^+ = 1 - I_{GF^{-1}(Y)}(j, n-j+1),$$

which form part of the *conditional* essential transition probability matrices in (4-54) and (4-58).

The in-control unconditional run-length distributions follow by substituting $F = G$ in p_C^- and p_C^+ so that $GF^{-1}(u) = u$.

For example, the in-control marginal (or unconditional) run-length distribution of the two-sided *2-of-2* KL chart is

$$\Pr(N_{2of2,0}^{KL} = t) = \Pr_{F=G}(N_{2of2}^{KL} = t) = \int_0^1 \int_0^y \Pr_{F=G}(N_{2of2}^{KL} = t \mid \mathbf{Z})\, f_{a,b}(x,y)\,dxdy$$
$$= \int_0^1 \int_0^y \boldsymbol{\xi}(\mathbf{Q}_{4\times4,0}^C)^{t-1}(\mathbf{I} - \mathbf{Q}_{4\times4,0}^C)\mathbf{1}\, f_{a,b}(x,y)\,dxdy$$

for $t = 1,2,3,\ldots$ with

$$\mathbf{Q}_{4\times4,0}^C = \begin{bmatrix} 0 & I_Y(j,n-j+1) - I_X(j,n-j+1) & 1 - I_Y(j,n-j+1) & I_X(j,n-j+1) \\ 0 & I_Y(j,n-j+1) - I_X(j,n-j+1) & 1 - I_Y(j,n-j+1) & I_X(j,n-j+1) \\ 0 & I_Y(j,n-j+1) - I_X(j,n-j+1) & 0 & I_X(j,n-j+1) \\ 0 & I_Y(j,n-j+1) - I_X(j,n-j+1) & 1 - I_Y(j,n-j+1) & 0 \end{bmatrix}.$$

Once again, the in-control run-length distribution is seen to be free from both F and G, and thus the *2-of-2* KL chart is distribution-free; the same being true for *2-of-3* precedence chart.

4.3.2 Unconditional *ARL*, *VARL* and *FAR* calculations

In order to design and study the performance of a Phase II control chart one typically examines the average and the variance of the *unconditional* run-length distribution ($UARL$ & $UVARL$) along with the *unconditional* false alarm rate ($UFAR$).

For the proposed nonparametric runs-rule enhanced precedence charts, expressions for the average and the variance of the *unconditional* run-length distribution can be obtained exactly and most conveniently derived by using the characteristics of the *conditional* run-length distributions coupled with conditional expectation.

To this end, note that,

$$E(N) = E_{\mathbf{Z}}(E(N \mid \mathbf{Z})) \qquad \text{and} \qquad \text{var}(N) = E_{\mathbf{Z}}(\text{var}(N \mid \mathbf{Z})) + \text{var}_{\mathbf{Z}}(E(N \mid \mathbf{Z})) \qquad (4\text{-}61)$$

where $E(N)$ and $\text{var}(N)$ denote the *unconditional* characteristics whilst $E(N \mid \mathbf{Z})$ and $\text{var}(N \mid \mathbf{Z})$ denote the *conditional* characteristics (i.e. conditioned on $\mathbf{Z}$). The *unconditional* false alarm rate ($UFAR$) can be obtained in a similar manner and is shown below.

4.3.2.1 Unconditional *ARL*, *VARL* and *FAR* of the *2-of-2* DR chart

Unconditional *ARL*

The *unconditional ARL* for the *2-of-2* DR chart is computed by averaging expression (4-41) over the joint distribution of the order statistics.

Thus

$$UARL_{2of2}^{DR} = E_{\mathbf{Z}}(E(N_{2of2}^{DR} \mid \mathbf{Z})) = E_{\mathbf{Z}}(CARL_{2of2}^{DR})$$

$$= \int_0^1 \int_0^y \left(\frac{1 + p_C^{\pm}}{(p_C^{\pm})^2} \right) f_{a,b}(x,y)dxdy \tag{4-62}$$

$$= \int_0^1 \int_0^y \left(\frac{2 - I_{GF^{-1}(y)}(j, n-j+1) + I_{GF^{-1}(x)}(j, n-j+1)}{(1 - I_{GF^{-1}(y)}(j, n-j+1) + I_{GF^{-1}(x)}(j, n-j+1))^2} \right) f_{a,b}(x,y)dxdy.$$

The in-control unconditional average run-length ($UARL_{2of2,0}^{DR}$) is obtained by substituting $F = G$ in (4-62) and given by

$$UARL_{2of2,0}^{DR} = \int_0^1 \int_0^y \left(\frac{2 - I_y(j, n-j+1) + I_x(j, n-j+1)}{(1 - I_y(j, n-j+1) + I_x(j, n-j+1))^2} \right) f_{a,b}(x,y)dxdy. \tag{4-63}$$

Unconditional *VARL*

The *unconditional* variance of the *2-of-2* DR chart is obtained by noting that, in general, the unconditional variance in (4-61) can be re-written as

$$\begin{aligned} \text{var}(N) &= E_{\mathbf{Z}}(\text{var}(N \mid \mathbf{Z})) + \text{var}_{\mathbf{Z}}(E(N \mid \mathbf{Z})) \\ &= E_{\mathbf{Z}}(\text{var}(N \mid \mathbf{Z})) + \{E_{\mathbf{Z}}[(E(N \mid \mathbf{Z}))^2] - [E_{\mathbf{Z}}(E(N \mid \mathbf{Z}))]^2\}. \end{aligned} \tag{4-64}$$

For the *2-of-2* DR chart, in particular, we have that $E(N_{2of2}^{DR} \mid \mathbf{Z})$ and $\text{var}(N_{2of2}^{DR} \mid \mathbf{Z})$ are given by (4-41) and (4-42), respectively so that the unconditional variance of the run-length of the *2-of-2* DR chart is given by

$$UVARL_{2of2}^{DR} = \text{var}(N_{2of2}^{DR}) = E_{\mathbf{z}}\left(\frac{1-5(1-p_C^{\pm})(p_C^{\pm})^2-(p_C^{\pm})^5}{(1-p_C^{\pm})^2(p_C^{\pm})^4}\right)+E_{\mathbf{z}}\left(\left(\frac{1+p_C^{\pm}}{(p_C^{\pm})^2}\right)^2\right)-\left(E_{\mathbf{z}}\left(\frac{1+p_C^{\pm}}{(p_C^{\pm})^2}\right)\right)^2. \quad (4\text{-}65)$$

The in-control variance of the unconditional run-length distribution of the *2-of-2* DR chart is found by substituting $1-I_Y(j,n-j+1)+I_X(j,n-j+1)$ for $p_C^{\pm}$ in expression (4-65), where $p_C^{\pm}$ is defined in (4-33).

Unconditional *FAR*

The *conditional* false alarm rate of the *2-of-2* DR chart follows from Table 4.5 by substituting $p_C^{+} = p_C^{+}(Y,F,G)$ (with $F=G$) for p_0^{+} and $p_C^{-} = p_C^{-}(X,F,G)$ (with $F=G$) for p_0^{-} and is given by

$$\begin{aligned} CFAR_{2of2}^{DR} &= (p_C^{+}(Y,F,F))^2+(p_C^{+}(X,F,F))^2+2p_C^{+}(Y,F,F)p_C^{+}(X,F,F) \\ &= (1-I_Y(j,n-j+1))^2+(I_X(j,n-j+1))^2+2(1-I_Y(j,n-j+1))(I_X(j,n-j+1)). \end{aligned}$$

By averaging over the joint distribution of the order statistics we obtain the *unconditional* false alarm rate of the *2-of-2* DR chart i.e.

$$\begin{aligned} UFAR_{2of2}^{DR} = E_{\mathbf{z}}(CFAR_{2of2}^{DR}) &= \int_0^1 (1-I_y(j,n-j+1))^2 f_b(y)dy+\int_0^1 (I_x(j,n-j+1))^2 f_a(x)dx \\ &\quad +2\int_0^1\int_0^y (1-I_y(j,n-j+1))(I_x(j,n-j+1)) f_{a,b}(x,y)dxdy \end{aligned} \quad (4\text{-}66)$$

where $f_a(x)$ and $f_b(x)$ denotes the marginal p.d.f's of the a^{th} and the b^{th} order statistics in a random sample (the reference or Phase I sample) of size m from a *uniform*(0,1) distribution, which are known to be a $Beta(a,m-a+1)$ distribution and a $Beta(b,m-b+1)$ distribution, respectively.

4.3.2.2 Unconditional *ARL*, *VARL* and *FAR* of the *2-of-2* KL chart

Unconditional *ARL*

The *conditional ARL* (or the conditional expected value) of the *2-of-2* KL chart follows from (4-52), with a symbolically simplified version given in Table 4.5, i.e.

$$CARL_{2of2}^{KL} = E(N_{2of2}^{KL} \mid \mathbf{Z}) = \boldsymbol{\xi}(\mathbf{I} - \mathbf{Q}_{4\times4}^{C})^{-1}\mathbf{1} = \left(\frac{(p_C^+)^2}{(p_C^+ + 1)} + \frac{(p_C^-)^2}{(p_C^- + 1)}\right)^{-1}$$

so that by averaging over $\mathbf{Z}$, the *unconditional ARL* of the *2-of-2* KL chart is found to be

$$\begin{aligned} UARL_{2of2}^{KL} &= E_{\mathbf{Z}}(E(N_{2of2}^{KL} \mid \mathbf{Z})) = E_{\mathbf{Z}}(CARL_{2of2}^{KL}) = \int_0^1\int_0^y \left(\frac{(p_C^+)^2}{(p_C^+ + 1)} + \frac{(p_C^-)^2}{(p_C^- + 1)}\right)^{-1} f_{a,b}(x,y)dxdy \\ &= \int_0^1\int_0^y \left(\frac{(1 - I_{GF^{-1}(y)}(j, n-j+1))^2}{(2 - I_{GF^{-1}(y)}(j, n-j+1))} + \frac{(I_{GF^{-1}(x)}(j, n-j+1))^2}{(I_{GF^{-1}(x)}(j, n-j+1)+1)}\right)^{-1} f_{a,b}(x,y)dxdy. \end{aligned} \tag{4-67}$$

The in-control unconditional average run-length is again obtained by substituting $F = G$ in (4-67)

$$UARL_{2of2,0}^{KL} = \int_0^1\int_0^y \left(\frac{(1 - I_y(j, n-j+1))^2}{(2 - I_y(j, n-j+1))} + \frac{(I_x(j, n-j+1))^2}{(I_x(j, n-j+1)+1)}\right)^{-1} f_{a,b}(x,y)dxdy \tag{4-68}$$

which is distribution-free.

Unconditional *VARL*

Substituting $\text{var}(N_{2of2}^{KL} \mid \mathbf{Z})$ (given in (4-53)) and $E(N_{2of2}^{KL} \mid \mathbf{Z})$ (given in (4-52)) in (4-64) we find that the *unconditional* variance of N_{2of2}^{KL} is given by

$$\begin{aligned} UVARL_{2of2}^{KL} = \text{var}(N_{2of2}^{KL}) = E_{\mathbf{Z}}\left(\boldsymbol{\xi}(\mathbf{I} + \mathbf{Q}_{4\times4}^{C})(\mathbf{I} - \mathbf{Q}_{4\times4}^{C})^{-2}\mathbf{1} - (CARL_{2of2}^{KL})^2\right) \\ + E_{\mathbf{Z}}\left(\left(\boldsymbol{\xi}(\mathbf{I} - \mathbf{Q}_{4\times4}^{C})^{-1}\mathbf{1}\right)^2\right) - \left(E_{\mathbf{Z}}\left(\boldsymbol{\xi}(\mathbf{I} - \mathbf{Q}_{4\times4}^{C})^{-1}\mathbf{1}\right)\right)^2. \end{aligned} \tag{4-69}$$

Unconditional *FAR*

The *conditional* false alarm rate of the *2-of-2* KL chart follows from Table 4.5 by substituting $p_C^+ = p_C^+(Y,F,G)$ (with $F=G$) for p_0^+ and $p_C^- = p_C^-(X,F,G)$ (with $F=G$) for p_0^- and is given by

$$\begin{aligned} CFAR_{2of2}^{KL} &= (p_C^+(Y,F,F))^2 + (p_C^+(X,F,F))^2 \\ &= (1-I_Y(j,n-j+1))^2 + (I_X(j,n-j+1))^2. \end{aligned} \tag{4-70}$$

By averaging over the joint distribution of the order statistics the *unconditional* false alarm rate is obtain as

$$UFAR_{2of2}^{KL} = E_{\mathbf{Z}}(CFAR_{2of2}^{KL}) = \int_0^1 (1-I_y(j,n-j+1))^2 f_b(y)dy + \int_0^1 (I_x(j,n-j+1))^2 f_a(x)dx \tag{4-71}$$

which is again distribution-free.

4.3.2.3 Unconditional *ARL*, *VARL* and *FAR* of the *2-of-3* chart

Unconditional *ARL* and *VARL*

The *unconditional ARL* and unconditional *VARL* of the *2-of-3* chart are obtained in the same manner as that of the *2-of-2* KL chart; that is, we use the *conditional* counterparts derived via the Markov chain approach and find that

$$UARL_{2of3} = E_Z(CARL_{2of3}) = \int_0^1 \int_0^y \boldsymbol{\xi}(\mathbf{I} - \mathbf{Q}_{8\times8}^C)\mathbf{1}\, f_{a,b}(x,y)dxdy$$

and

$$\begin{aligned} UVARL_{2of3} &= E_{\mathbf{Z}}(\text{var}(N_{2of3} \mid \mathbf{Z})) + \{E_{\mathbf{Z}}[(E(N_{2of3} \mid \mathbf{Z}))^2] - [E_{\mathbf{Z}}(E(N_{2of3} \mid \mathbf{Z}))]^2\} \\ &= E_{\mathbf{Z}}\left(\boldsymbol{\xi}(\mathbf{I} + \mathbf{Q}_{8\times8}^C)(\mathbf{I} - \mathbf{Q}_{8\times8}^C)^{-2}\mathbf{1} - (CARL_{2of3})^2\right) + E_{\mathbf{Z}}\left(\left(\boldsymbol{\xi}(\mathbf{I} - \mathbf{Q}_{8\times8}^C)^{-1}\mathbf{1}\right)^2\right) - \left(E_{\mathbf{Z}}\left(\boldsymbol{\xi}(\mathbf{I} - \mathbf{Q}_{8\times8}^C)^{-1}\mathbf{1}\right)\right)^2, \end{aligned}$$

respectively.

Unconditional *FAR*

The *conditional FAR* of the *2-of-3* chart is found (like that of the *2-of-2* DR chart and the *2-of-2* KL chart) from Table 4.5 by substituting the conditional probabilities p_C^- and p_C^+ (defined in (4-35) and (4-36) with $F = G$) for p^- and p^+, respectively and is given by

$$\begin{aligned} UFAR_{2of3} = {} & 2(1 - I_Y(j, n-j+1))^2(1 - I_Y(j, n-j+1) + I_X(j, n-j+1)) \\ & + 2(I_X(j, n-j+1))^2(1 - I_Y(j, n-j+1) + I_X(j, n-j+1)). \end{aligned}$$

The *unconditional* FAR is thus given by

$$\begin{aligned} UFAR_{2of3} = {} & 2\int_0^1 \int_0^y (1 - I_y(j, n-j+1))^2(1 - I_y(j, n-j+1) + I_x(j, n-j+1))\, f_{a,b}(x,y)dxdy \\ & + 2\int_0^1 \int_0^y (I_x(j, n-j+1))^2(1 - I_y(j, n-j+1) + I_x(j, n-j+1))\, f_{a,b}(x,y)dxdy. \end{aligned}$$

4.3.3 Run-length distributions of the one-sided precedence charts

If detecting higher (lower) values is of interest, that is, whether the parameter or percentile of interest has shifted to the right (left), we can use a one-sided upper (lower) control chart with an estimated upper (lower) control limit $U\hat{C}L = X_{b:m}$ ($L\hat{C}L = X_{a:m}$) only.

The operation of the one-sided upper and lower runs-rules enhanced precedence charts of Case U is similar to that of the one-sided upper and lower runs-rules enhanced sign charts of Case K. For example, the *2-of-2* one-sided upper (lower) precedence chart signals on the first occurrence of a run of length two of the charting statistic $Y_{j:n}^{i}$ on or above (below) the estimated upper (lower) control limit.

The derivation of the run-length distributions of the one-sided runs-rules enhanced precedence charts parallels that of the two-sided precedence charts. In particular, we let

$$\hat{\xi}_i^{+} = I(Y_{j:n}^{i} \geq U\hat{C}L) = \begin{cases} 1 & \text{if} \quad Y_{j:n}^{i} \geq U\hat{C}L \\ 0 & \text{if} \quad Y_{j:n}^{i} < U\hat{C}L \end{cases}$$

and

$$\hat{\xi}_i^{-} = I(Y_{j:n}^{i} \leq L\hat{C}L) = \begin{cases} 1 & \text{if} \quad Y_{j:n}^{i} \leq L\hat{C}L \\ 0 & \text{if} \quad Y_{j:n}^{i} > L\hat{C}L \end{cases}$$

for $i = 1,2,3,...$ denote the indicator functions for the one-sided precedence charts corresponding to the events $\{Y_{j:n}^{i} \geq U\hat{C}L\}$ and $\{Y_{j:n}^{i} \leq L\hat{C}L\}$, respectively. Then, we can again use a two-step approach to derive the run-length distribution. In other words, we first derive the *conditional* run-length distribution i.e. conditioned on the particular order statistic (control limit) and then, second, we derive the *unconditional* or marginal run-length distribution by "averaging over" the distribution of the order statistic that constitutes the Phase II control limit.

In particular, given $X_{b:m} = x_{b:m}$ the sequence of signaling indicators $\hat{\xi}_1^{+}, \hat{\xi}_2^{+}, \hat{\xi}_3^{+}, ...$ are i.i.d. Bernoulli random variables with success probability

$$p_C^{+} = \Pr(\hat{\xi}_i^{\pm} = 1 \mid X_{b:m} = x_{b:m}) = \Pr(\hat{\xi}_i^{+} = 1 \mid X_{b:m} = x_{b:m}) = p_C^{+}(Y, F, G)$$

so that the *conditional* distribution of the run-length variable N^{+}_{2of2} of the *2-of-2* upper one-sided precedence chart, for example, is geometric of order $k=2$ with parameter (success probability) $p^{+}_{C}=p^{+}(y,F,G)$. Consequently, all the properties and the characteristics of the *conditional* run-length distribution follow conveniently from the properties of the geometric distribution of order $k=2$ by substituting $\alpha=p^{+}_{C}$ and $k=2$ in expressions (4-16) and (4-17).

Alternatively, we can use a Markov chain approach; doing so we find that

$$\Pr(N^{+}_{2of2}=t \mid X_{b:m}=x_{b:m})=\boldsymbol{\xi}(\mathbf{Q}^{C}_{3\times3})^{t-1}(\mathbf{I}-\mathbf{Q}^{C}_{3\times3})\mathbf{1} \qquad \text{for} \qquad t=1,2,3,...$$

with the *conditional* average run-length and the *conditional* variance of the run-length given by

$$CARL^{+}_{2of2}=E(N^{+}_{2of2} \mid X_{b:m}=x_{b:m})=\boldsymbol{\xi}(\mathbf{I}-\mathbf{Q}^{C}_{3\times3})^{-1}\mathbf{1}$$

and

$$CVARL^{+}_{2of2}=\text{var}(N^{+}_{2of2} \mid X_{b:m}=x_{b:m})=\boldsymbol{\xi}(\mathbf{I}+\mathbf{Q}^{C}_{3\times3})(\mathbf{I}-\mathbf{Q}^{C}_{3\times3})^{-2}\mathbf{1}-(CARL^{+}_{2of2})^{2}$$

respectively, where

$$\mathbf{Q}^{C}_{3\times3}=\begin{bmatrix}0 & 1-p^{+}_{C} & p^{+}_{C}\\ 0 & 1-p^{+}_{C} & p^{+}_{C}\\ 0 & 1-p^{+}_{C} & 0\end{bmatrix}$$

denotes the *conditional* essential transition probability matrix of the *2-of-2* upper one-sided precedence chart and follows from (4-15) having substituted p^{+}_{C} for p^{+}.

The *unconditional* p.m.f of N^{+}_{2of2}, for example, is obtained by averaging the *conditional* run-length distribution over the distribution of $X_{b:m}$ i.e.

$$\Pr(N^{+}_{2of2}=t)=\int_{0}^{1}\Pr(N^{+}_{2of2}=t \mid Y=y)\,f_{b}(y)dy=\int_{0}^{1}\boldsymbol{\xi}(\mathbf{Q}^{C}_{3\times3})^{t-1}(\mathbf{I}-\mathbf{Q}^{C}_{3\times3})\mathbf{1}\,f_{b}(y)dy.$$

To obtain a closed form expression of the *unconditional* p.m.f of N^{+}_{2of2} requires the same steps as carried out in case of the *2-of-2* DR precedence chart of section 4.3.1.2 and therefore not shown here.

4.3.4 Design and implementation of the two-sided precedence charts

In order to implement the proposed precedence charts in practice we need the upper and the lower control limits. This means that we need to find the indices (charting constants) a and b that specify the reference sample order statistics, which constitute the lower and the upper control limit, respectively.

Determination of charting constants

In Phase II applications one typically determines the charting constants a and b so that a specified in-control unconditional average run-length (say, $UARL_0^*$ equal to 370 or 500) is obtained. This means that we have to solve

$$UARL_{1of1,0}^{*} = \int_0^1 \int_0^y \left(1 - \frac{1}{\beta(j,n-j+1)} \sum_{h=0}^{n-j} \frac{(-1)^h}{j+h} \binom{n-j}{h} \left(y^{j+h} - x^{j+h}\right)\right)^{-1} f_{a,b}(x,y)dxdy \qquad (4\text{-}72)$$

for the *1-of-1* chart,

$$UARL_{2of2,0}^{DR*} = \int_0^1 \int_0^y \left(\frac{2 - I_y(j,n-j+1) + I_x(j,n-j+1)}{(1 - I_y(j,n-j+1) + I_x(j,n-j+1))^2}\right) f_{a,b}(x,y)dxdy \qquad (4\text{-}73)$$

for the *2-of-2* DR chart,

$$UARL_{2of2,0}^{KL*} = \int_0^1 \int_0^y \left(\frac{(1 - I_y(j,n-j+1))^2}{(2 - I_y(j,n-j+1))} + \frac{(I_x(j,n-j+1))^2}{(I_x(j,n-j+1)+1)}\right)^{-1} f_{a,b}(x,y)dxdy \qquad (4\text{-}74)$$

for the *2-of-2* KL chart and

$$UARL_{2of3,0}^{*} = \int_0^1 \int_0^y \left(\xi(\mathbf{I} - \mathbf{Q}_{8\times8,0}^{C})^{-1}\mathbf{1}\right) f_{a,b}(x,y)dxdy \qquad (4\text{-}75)$$

for the *2-of-3* chart where $\mathbf{Q}_{8\times8,0}^{C}$ follows from (4-58) by substituting $1 - I_Y(j,n-j+1)$ for p_C^+ and $I_Y(j,n-j+1)$ for p_C^-.

4.3.4.1 Charting constants of the *1-of-1* chart

Chakraborti et al. (2004) provided values for the charting constants a and b for the two-sided *1-of-1* precedence chart for a number of different choices (combinations) of the size m of the Phase I reference sample, the size n of Phase II samples and j (the selected order statistic) so that the in-control unconditional average run-length (i.e. $UARL_{1of1,0}$) is close to 370 , 500 and 1000, respectively.

4.3.4.2 Charting constants of the *2-of-2* DR, *2-of-2* KL and *2-of-3* charts

Tables 4.18, 4.19 and 4.20 display various choices (combinations) of the charting constants a and b for the two-sided *2-of-2* DR, the two-sided *2-of-2* KL and the two-sided *2-of-3* charts, for a given or specified in-control unconditional ARL in the neighborhood of 300 and 500, when reference samples of size $m = 50$, 100, 200 and 500 are used to estimate the control limits in Phase I and these limits are used to monitor the location (center) of a process using the medians of Phase II (test) samples of size $n = 5$, 7, or 9, respectively. Thus, j equals 3, 4, and 5, respectively in the tables.

Note that for each combination of values of *n*, *j* and *m* the tables display (in each cell) the $UARL_0$, the $UFAR$ and (a,b) values, where the $UARL_0$ values are in the neighborhood of 300 to 500.

Since the Phase II (test) sample median is used as the charting statistic and the Phase II sample size n is odd, it seems reasonable to use symmetric control limits, and thus we take $b = m - a + 1$, so that only a needs to be determined. However, this needs not be the case when the chart constants are to be determined for a charting statistic other than the median that might be of interest.

In addition, note that, in general, it is rare to achieve an $UARL_0$ (or an $UFAR$) exactly as specified (i.e. 300 or 500) with the nonparametric charts because the in-control distribution of the run-length distribution is discrete. However, as can be seen, one can get reasonably close to the values typically used in practice.

For example, from Table 4.18 for $m = 500$, $n = 5$ and $j = 3$, one set of constants for the *2-of-2* DR chart are given by $a = 72$ and $b = 500 - 72 + 1 = 429$ so that $\hat{LCL} = X_{72:500}$ and $\hat{UCL} = X_{429:500}$. In this case the achieved (or attained) $UARL_0$ and the attained unconditional FAR of the chart are 496.90 and 0.0025, respectively. Moreover, these are the exact values and remain the same for all continuous distributions. If instead we took $a = 71$ and $b = 500 - 71 + 1 = 430$, so that $\hat{LCL} = X_{71:500}$

and $\hat{UCL} = X_{430:500}$, the achieved $UARL_0$ increases to 536.72 and the attained FAR decreases to 0.0023. For a more moderate reference sample size, such as $m = 50$, Table 4.18 shows that it is possible to obtain $UARL_0$ values such as 275.30 or 605.44; the latter which may be deemed reasonably large in practice. Obviously as *m* and/or *n* increase, the available choices for the $UARL_0$ values also increase.

Similar behavior is observed in the case of the two-sided *2-of-2* KL and the *2-of-3* charts shown in Tables 4.19 and 4.20, respectively. For instance, in case of the *2-of-2* KL chart, when $m = 500$ and one uses $a = 80$ and $b = 500 - 80 + 1 = 421$, so that $\hat{LCL} = X_{80:500}$ and $\hat{UCL} = X_{421:500}$, the ARL_0 of the *2-of-2* KL chart (when $n = 5$ and $j = 3$) is 524.39, whereas the FAR is 0.0023. However, if instead one chooses to use $a = 81$ and $b = 500 - 81 + 1 = 420$, so that $\hat{LCL} = X_{81:500}$ and $\hat{UCL} = X_{420:500}$, the $UARL_0$ decreases to 490.21, whereas the $UFAR$ slightly increases to 0.0024. Although for $m = 500$, $n = 5$ and $j = 3$ a specified $UARL_0$ such as 500 cannot be obtained exactly, by increasing the size of the reference sample m and/or the test sample size *n*, the range of possible $UARL_0$ and $UFAR$ values that can be attained increases.

All equations (i.e. (4-73), (4-74) and (4-75)) are solved using the software package Mathcad®14.0.

Table 4.18: Unconditional in-control average run-length ($UARL_0$), unconditional false alarm rate ($UFAR$) and chart constants (a,b) [1] for the *2-of-2* DR nonparametric chart for m = 50,100,200,500 and $(n, j) = (5, 3), (7, 4), (9, 5)$

n=5 , j=3				n=7 , j=4				n=9 , j=5			
m=50	100	200	500	m=50	100	200	500	m=50	100	200	500
605.44	548.99	537.62	536.72	597.80	509.54	597.72	526.08	976.53	739.47	558.51	528.95
0.0072	0.0040	0.0029	0.0023	0.0090	0.0048	0.0027	0.0024	0.0084	0.0040	0.0031	0.0024
(8,43)	(15,86)	(29,172)	(71,430)	(10,41)	(19,82)	(36,165)	(90,411)	(11,40)	(21,80)	(42,159)	(104,397)
275.30	373.31	443.56	496.90	264.91	345.93	490.44	487.01	383.92	481.18	456.18	488.41
0.0121	0.0055	0.0034	0.0025	0.0150	0.0065	0.0033	0.0026	0.0144	0.0056	0.0037	0.0026
(9,42)	(16,85)	(30,171)	(72,429)	(11,40)	(20,81)	(37,164)	(91,410)	(12,39)	(22,79)	(43,158)	(105,396)
	261.69	368.80	460.60		241.21	405.20	451.33	172.47	322.26	375.04	451.43
	0.0074	0.0040	0.0026		0.0088	0.0039	0.0028	0.0236	0.0077	0.0044	0.0028
	(17,84)	(31,170)	(73,428)		(21,80)	(38,163)	(92,409)	(13,38)	(23,78)	(44,157)	(106,395)
		308.82	427.48			336.97	418.70		221.57	310.28	417.68
		0.0047	0.0028			0.0046	0.0030		0.0104	0.0053	0.0030
		(32,169)	(74,427)			(39,162)	(93,408)		(24,77)	(45,156)	(107,394)
		260.37	397.20			281.98	388.83			258.24	386.83
		0.0056	0.0031			0.0054	0.0032			0.0062	0.0033
		(33,168)	(75,426)			(40,161)	(94,407)			(46,155)	(108,393)
			369.50				361.45				358.60
			0.0033				0.0034				0.0035
			(76,425)				(95,406)				(109,392)
			344.12				336.33				332.75
			0.0035				0.0037				0.0038
			(77,424)				(96,405)				(110,391)
			320.83				313.25				309.06
			0.0037				0.0039				0.0041
			(78,423)				(97,404)				(111,390)
			299.44				292.03				287.31
			0.0040				0.0042				0.0044
			(79,422)				(98,403)				(112,389)

[1]The three rows of each cell shows the achieved (attained) $UARL_0$, the $UFAR$ and the charting constants (a , b), respectively

Table 4.19: Unconditional in-control average run-length ($UARL_0$), unconditional false alarm rate ($UFAR$) and chart constants (a,b) [1] for the *2-of-2* KL nonparametric chart for m = 50,100,200,500 and $(n, j) = (5, 3), (7, 4), (9, 5)$

n=5 , j=3				n=7 , j=4				n=9 , j=5			
m=50	100	200	500	m=50	100	200	500	m=50	100	200	500
1010.37	650.75	559.01	524.39	985.39	594.56	504.01	506.61	1591.68	547.12	548.41	530.19
0.0048	0.0033	0.0026	0.0023	0.0063	0.0041	0.0031	0.0024	0.0062	0.0049	0.0031	0.0023
(8,43)	(16,85)	(32,169)	(80,421)	(10,41)	(20,81)	(40,161)	(99,402)	(11,40)	(23,78)	(45,156)	(112,389)
460.89	456.52	471.18	490.21	437.32	414.67	424.10	472.95	626.67	376.11	456.29	493.12
0.0079	0.0044	0.0031	0.0024	0.0102	0.0054	0.0036	0.0026	0.0103	0.0066	0.0036	0.0025
(9,42)	(17,84)	(33,168)	(81,420)	(11,40)	(21,80)	(41,160)	(100,401)	(12,39)	(24,77)	(46,155)	(113,388)
237.00	328.69	399.60	458.70	217.33	296.08	358.81	441.90	281.29	264.69	381.78	459.05
0.0123	0.0057	0.0036	0.0026	0.0160	0.0070	0.0042	0.0027	0.0165	0.0086	0.0042	0.0027
(10,41)	(18,83)	(34,167)	(82,419)	(12,39)	(22,79)	(42,159)	(101,400)	(13,38)	(25,76)	(47,154)	(114,387)
	242.15	340.87	429.62			305.16	413.24			321.15	427.69
	0.0074	0.0041	0.0027			0.0048	0.0029			0.0049	0.0029
	(19,82)	(35,166)	(83,418)			(43,158)	(102,399)			(48,153)	(115,386)
		292.37	402.76			260.82	386.77			271.54	398.81
		0.0047	0.0029			0.0056	0.0031			0.0057	0.0031
		(36,165)	(84,417)			(44,157)	(103,398)			(49,152)	(116,385)
			377.91				362.28				372.18
			0.0031				0.0033				0.0033
			(85,416)				(104,397)				(117,384)
			354.91				339.62				347.61
			0.0033				0.0035				0.0035
			(86,415)				(105,396)				(118,383)
			333.60				318.62				324.92
			0.0035				0.0037				0.0038
			(87,414)				(106,395)				(119,382)
			313.83				299.16				303.95
			0.0037				0.0040				0.0040
			(88,413)				(107,394)				(120,381)
			295.48								284.55
			0.0039								0.0043
			(89,412)								(121,380)

[1]The three rows of each cell shows the achieved (attained) $UARL_0$, the $UFAR$ and the charting constants (a , b), respectively

Table 4.20: Unconditional in-control average run-length ($UARL_0$), unconditional false alarm rate ($UFAR$) and chart constants (a,b) [1] for the *2-of-3* nonparametric chart for m = 50,100,200,500 and $(n, j) = (5, 3), (7, 4), (9, 5)$

n=5 , j=3				n=7 , j=4				n=9 , j=5			
m=50	100	200	500	m=50	100	200	500	m=50	100	200	500
1336.27	755.74	623.35	672.76	1313.99	690.80	685.53	653.58	2423.24	1025.16	786.51	712.63
0.0049	0.0033	0.0025	0.0018	0.0065	0.0040	0.0025	0.0019	0.0062	0.0035	0.0024	0.0018
(7,44)	(14,87)	(28,73)	(68,433)	(9,42)	(18,83)	(35,166)	(87,414)	(10,41)	(20,81)	(40,161)	(100,401)
527.33	502.46	513.03	621.56	513.55	460.68	561.53	604.19	819.75	653.26	636.94	656.49
0.0084	0.0045	0.0030	0.0020	0.0107	0.0055	0.0030	0.0021	0.0105	0.0049	0.0029	0.0020
(8,43)	(15,86)	(29,172)	(69,432)	(10,41)	(19,82)	(36,165)	(88,413)	(11,40)	(21,80)	(41,160)	(101,400)
246.51	346.18	425.78	575.05	233.72	316.82	463.41	559.19	329.94	430.22	519.65	605.46
0.0134	0.0060	0.0035	0.0021	0.0169	0.0073	0.0035	0.0022	0.0169	0.0066	0.0034	0.0021
(9,42)	(16,85)	(30,171)	(70,431)	(11,40)	(20,81)	(37,164)	(89,412)	(12,39)	(22,79)	(42,159)	(102,399)
130.77	246.05	356.16	532.74	120.09	224.02	385.16	518.13	152.81	291.98	426.99	559.02
0.0201	0.0078	0.0042	0.0023	0.0252	0.0095	0.0041	0.0024	0.0257	0.0087	0.0041	0.0023
(10,41)	(17,84)	(31,170)	(71,430)	(12,39)	(21,80)	(38,163)	(90,411)	(13,38)	(23,78)	(43,158)	(103,398)
	179.74	300.11	494.18		162.45	322.29	480.62	79.60	203.71	353.24	516.70
	0.0101	0.0048	0.0024		0.0122	0.0048	0.0026	0.0374	0.0114	0.0048	0.0025
	(18,83)	(32,169)	(72,429)		(22,79)	(39,162)	(91,410)	(14,37)	(24,77)	(44,157)	(104,397)
		254.64	459.00			271.43	446.32		145.81	294.15	478.10
		0.0056	0.0026			0.0056	0.0028		0.0147	0.0056	0.0027
		(33,168)	(73,428)			(40,161)	(92,409)		(25,76)	(45,156)	(105,396)
		217.47	426.85			230.00	414.91			246.49	442.85
		0.0065	0.0028			0.0065	0.0030			0.0065	0.0029
		(34,167)	(74,427)			(41,160)	(93,408)			(46,155)	(106,395)
			397.43				386.11			207.81	410.63
			0.0030				0.0032			0.0076	0.0031
			(75,426)				(94,407)			(47,154)	(107,394)
			370.48				359.69				381.14
			0.0031				0.0034				0.0033
			(76,425)				(95,406)				(108,393)
			345.75				335.41				354.12
			0.0034				0.0030				0.0035
			(77,424)				(96,405)				(109,392)
			323.04				313.09				329.35
			0.0037				0.0039				0.0038
			(78,423)				(97,404)				(110,391)
			302.15				292.53				306.61
			0.0039				0.0042				0.0041
			(79,422)				(98,403)				(111,390)
			282.91				273.59				285.71
			0.0042				0.0044				0.0044
			(80,421)				(99,402)				(112,389)
			265.18				256.12				266.50
			0.0045				0.0047				0.0047
			(81,420)				(100,401)				(113,388)

[1]The three rows of each cell shows the achieved (attained) $UARL_0$, the $UFAR$ and the charting constants (a , b), respectively

Example 3

In order to illustrate the runs-rule enhanced nonparametric precedence charts we use the data given in Table 5.1 on p. 213 and Table 5.2 on p. 219 of Montgomery (2001).

The goal of this study was to establish statistical control of the inside diameter of the piston rings for an automotive engine manufactured in a forging process. Twenty-five retrospective or Phase I samples, each of size five, were collected when the process was thought to be in-control. As shown in Example 5.1 on p. 213 of Montgomery (2001), the traditional Shewhart $\bar{X}$ and R charts provide no indication of an out-of-control condition, so these data are considered to be Phase I reference data and these "trial" limits were adopted for use in on-line process control.

In order to implement the nonparametric control charts the charting constants are needed. Possible symmetric control limits ($b = m - a + 1$) for the four charts are shown in Table 4.21, for $m = 125$, $n = 5$ and $j = 3$, along with the corresponding $UARL_0$ and $UFAR$ values.

Table 4.21: Unconditional in-control average run-length ($UARL_0$), unconditional false alarm rate ($UFAR$) and chart constants (a, b) for the two-sided *1-of-1*, *2-of-2* DR, *2-of-2* KL and *2-of-3* precedence charts when $m = 125$, $n = 5$ and $j = 3$

1-of-1				*2-of-2* DR				*2-of-2* KL				*2-of-3*			
a	***b***	$UARL_0$	*UFAR*	***a***	***b***	$UARL_0$	*UFAR*	***a***	***b***	$UARL_0$	*UFAR*	***a***	***b***	$UARL_0$	*UFAR*
5	121	1315.98	0.0019	17	109	898.74	0.0023	18	108	1125.44	0.0018	17	109	822.40	0.0026
6	120	695.09	0.0029	18	108	638.60	0.0031	19	107	819.47	0.0024	18	108	590.03	0.0034
7	119	413.80	0.0044	19	107	464.38	0.0040	20	106	608.81	0.0030	19	107	433.39	0.0043
8	118	267.40	0.0062	20	106	344.73	0.0052	21	105	460.54	0.0038	20	106	325.09	0.0055
9	117	183.47	0.0084	21	105	260.69	0.0066	22	104	354.09	0.0048	21	105	248.51	0.0069
				22	104	200.46	0.0084	23	103	276.28	0.0059	22	104	193.27	0.0086

Using Table 4.21, for an $UARL_0$ of 500, one can take $a = 7$ so that $b = 119$, and therefore the control limits for the *1-of-1* precedence chart are the 7th and the 119th ordered values of the reference sample. Thus $\hat{LCL} = X_{7:125} = 73.984$ and $\hat{UCL} = X_{119:125} = 74.017$, which yield an in-control unconditional ARL of 413.80 and an unconditional FAR of 0.0044.

A plot of the medians for the *1-of-1* chart is shown in Figure 4.12 for all forty samples, the first twenty five of which are from Phase I. It is seen that the 37th median is outside the control limits and so the *1-of-1* precedence chart signals on the 12th (i.e. 37th – 25th) sample in the prospective phase.

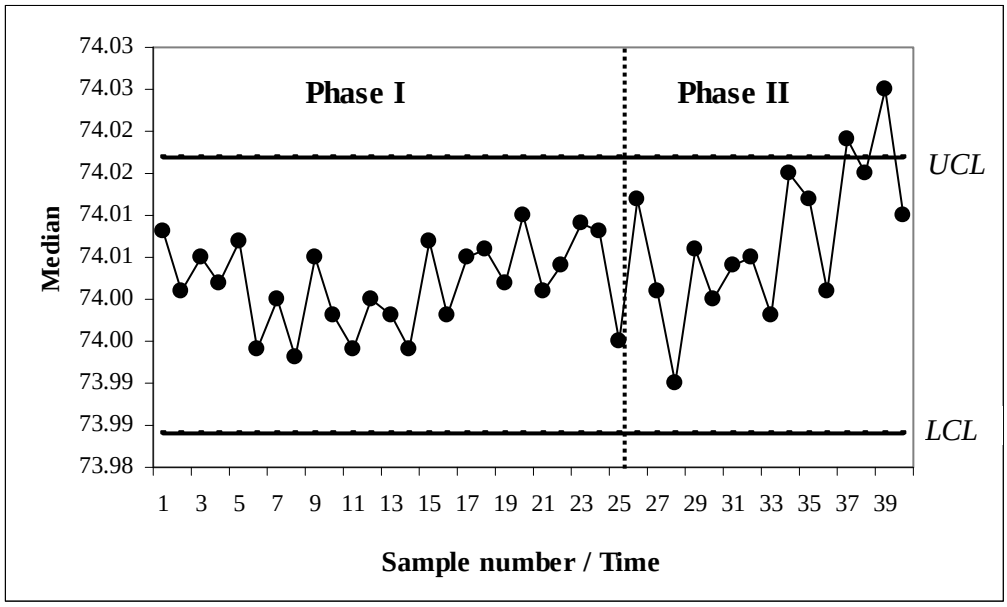


Figure 4.12: The *1-of-1* Phase II Precedence chart for the Montgomery (2001) piston-ring data

For the *2-of-2* DR chart, we take $a = 19$ so that $b = 125 - 19 + 1 = 107$ and the resulting limits, $\hat{LCL} = X_{19:125} = 73.990$ and $\hat{UCL} = X_{107:125} = 74.012$, yield an $UARL_0$ and $UFAR$ of 464.38 and 0.0040, respectively. Note, however, that if one chooses $a = 20$ so that $b = 106$, the control limits become $\hat{LCL} = X_{20:125}$ and $\hat{UCL} = X_{106:125}$ and the corresponding $UARL_0$ decreases to 344.73, whereas the $UFAR$ slightly increases to 0.0052. The *2-of-2* DR chart is shown in Figure 4.13.

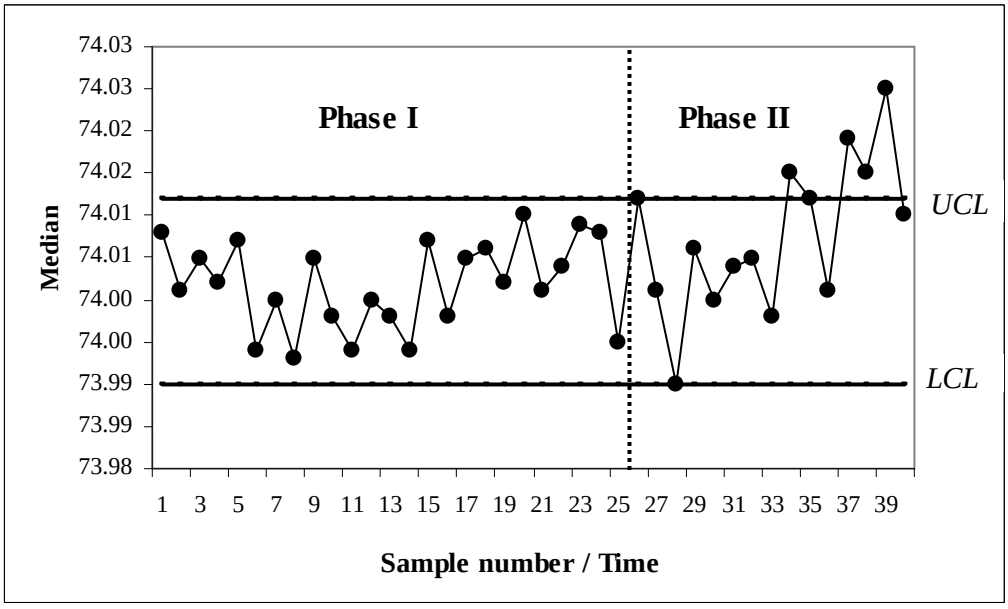


Figure 4.13: The *2-of-2* Phase II DR chart for the Montgomery (2001) piston-ring data

For the *2-of-2* KL chart we take $a = 21$ so that $b = 125 - 21 + 1 = 105$ and thus $L\hat{C}L = X_{21:125} = 73.992$ and $U\hat{C}L = X_{105:125} = 74.011$; this yields an $UARL_0$ of 460.54 and an *UFAR* of 0.0038, respectively. The *2-of-2* KL chart is shown in Figure 4.14.

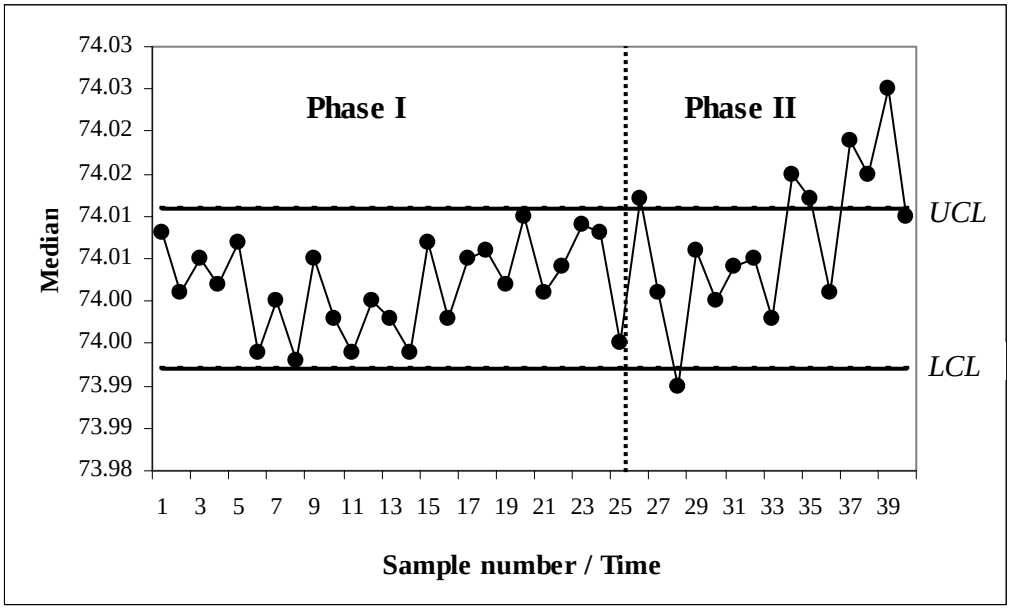


Figure 4.14: The *2-of-2* Phase II KL chart for the Montgomery (2001) piston ring data

Finally, for the *2-of-3* chart we take $a = 19$ so that $b = 125 - 19 + 1 = 107$ and thus $L\hat{C}L = X_{19:125} = 73.990$ and $U\hat{C}L = X_{107:125} = 74.012$ which yields an $UARL_0$ of 433.39 and an *UFAR* of 0.0043, respectively. The *2-of-3* chart is identical to the *2-of-2* DR chart shown in Figure 4.14 and is thus omitted; this is so because the control limits (in this example) of the *2-of-3* chart are exactly the same as that of the *2-of-2* DR chart.

The *2-of-2* DR charts signals on the 3rd sample whereas both the *2-of-2* KL and the *2-of-3* charts signal on the 10th sample in the prospective phase. Note, however, that the achieved *UFAR* values for the four charts are much larger (63%, 48%, 41% and 59%, respectively) than the nominal *FAR* of 0.0027.

4.3.5 Performance comparison of the two-sided precedence charts

The performance of Phase II control charts is typically compared by first designing each control chart to (roughly) have the same in-control *unconditional* average run-length ($UARL_0$) and then examining their out-of-control *unconditional* average run-length ($UARL_1$) values at some out-of-control value(s) of the parameter of interest. The control chart with the shorter (or smaller) out-of-control average run-length is usually preferred. Since the proposed run rules enhanced Phase II charts are nonparametric Shewhart-type charts applicable in Case U, their main competitor is the basic *1-of-1* precedence control chart of Chakraborti et al. (2004).

To study robustness, three different underlying process distributions i.e. the normal distribution, the *t*-distribution and the gamma distribution, were used in a simulation study with 100 000 repetitions for each distribution investigated. Because the shape of the *t*-distribution is very similar to that of the normal distribution (it is symmetric, but with more probability in the tails) it was used to study the effect of heavier tails. The gamma distribution was used to study the effect of skewness (see e.g. Figure 4.15). In order for the results of the three distributions to be comparable, the t and gamma distributions were scaled so that they also had a mean of zero and a variance of one. Thus, the $n(0,1)$, the $\frac{1}{\sqrt{2}}t(4)$ and the $Gamma(1,1)-1$ distributions were used.

The parametric Shewhart $\bar{X}$ chart was included in the comparison for the normal distribution but not for the *t* and the gamma distribution since the $\bar{X}$ chart is well-known to be non-robust under non-normality (see e.g. Chakraborti et al. (2004)).

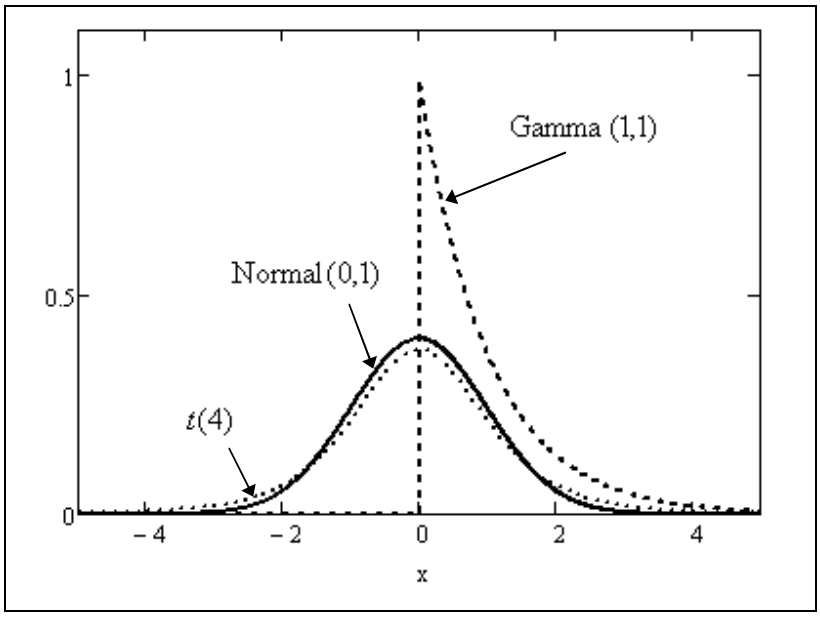


Figure 4.15: Probability distributions used for the performance comparison of the two-sided precedence control charts

Tables 4.22, 4.23 and 4.24 display the performance comparison results when a reference sample of size $m = 500$ is used to estimate the control limits to monitor location in Phase II (future or test) samples of size $n = 5$ and using $Y_{3:5}$, the median, as the charting statistic. The charts were designed so that an $UARL_0$ value close to 500 was achieved.

Instead of using randomization to get an $UARL_0$ exactly equal to 500, two combinations of chart constants are used for each nonparametric chart for which the $UARL_0$ was the nearest below and the nearest above the target $UARL_0^*$ of 500. The tables show the *unconditional* average run-length ($UARL_0$) along with the *unconditional* standard deviation of the run-length ($SDRL_0$). The shift refers to a shift in the mean of the distribution.

From Table 4.22 it is seen that even under the normal distribution, the nonparametric charts can be quite efficient i.e. good at detecting shifts. The *2-of-2* KL and the *2-of-3* charts are both almost as efficient as the $\overline{X}$ chart, with shorter ARL's but a slightly higher $SDRL$'s when the process is OOC, especially for small shifts.

When the distribution is $t(4)$, that is symmetric yet with heavier tails than the normal, Table 4.23 shows that the *2-of-2* DR, the *2-of-2* KL and the *2-of-3* schemes perform better than the basic precedence *1-of-1* chart in detecting small shifts, with the *2-of-2* KL chart being the best and is closely followed by the *2-of-3* chart. Thus, the three new nonparametric Shewhart-type charts with signaling rules provide better alternatives than the basic precedence *1-of-1* chart and the $\overline{X}$ chart, especially for smaller shifts. The same observation applies in the case of a right-skewed distribution such as the $Gamma(1,1)$ as shown in Table 4.24 but with the *2-of-2* KL chart doing the best. So the runs-type signaling rules enhance the nonparametric chart's sensitivity to a location shift.

Moreover, the gain in efficiency (relative to the *1-of-1* chart) can be substantial; for example, for the $t(4)$ distribution for a shift of 0.5, the OOC ARL values of the *1-of-1*, *2-of-2* DR, *2-of-3* and the *2-of-2* KL charts are 117.63, 40.98, 26.64 and 26.28, respectively when the corresponding ARL_0 values are very comparable, 520.27, 536.72, 532.74 and 524.39, respectively. Note that in Table 4.24 for the $Gamma(1,1)$ distribution the basic precedence chart display somewhat of a strange behavior in that both the ARL and $SDRL$ values first increase from their corresponding values for the in-control case for a shift of 0.25; thereafter the ARL and $SDRL$ values decrease for increasing shifts as it might be expected. We have not been able to explain this phenomenon. A repeat of the simulations produced similar results.

Table 4.22: Unconditional *ARL* and *SDRL* values for the *normal* distribution for the *2-of-2* DR, *2-of-2* KL, *2-of-3*, Basic (*1-of-1*) Precedence chart and the Shewhart X-bar chart when $m = 500$, $n = 5$, $j = 3$

	2-of-2 DR				*2-of-2* KL				*2-of-3*			
Shift	***ARL***	***SDRL***	***ARL***	***SDRL***	***ARL***	***SDRL***	***ARL***	***SDRL***	***ARL***	***SDRL***	***ARL***	***SDRL***
0.00	496.90	573.05	536.72	621.20	490.21	554.18	524.39	594.55	494.18	569.01	532.74	615.81
0.25	233.82	278.56	250.23	299.59	170.07	203.00	180.06	215.88	161.67	196.04	171.95	209.89
0.50	58.22	66.10	61.33	69.99	39.37	43.17	41.11	45.28	36.62	39.34	38.26	41.33
0.75	17.55	17.85	18.23	18.64	12.99	12.60	13.39	13.06	12.75	11.06	13.15	11.58
1.00	7.36	6.41	7.56	6.63	5.99	4.90	6.12	5.04	6.54	4.23	6.64	4.36
1.25	4.12	2.90	4.19	2.97	3.62	2.34	3.67	2.39	4.43	1.98	4.47	2.03
1.50	2.88	1.52	2.91	1.55	2.67	1.26	2.69	1.29	3.60	1.04	3.62	1.06
1.75	2.36	0.86	2.37	0.87	2.27	0.72	2.28	0.74	3.24	0.58	3.25	0.58
2.00	2.13	0.49	2.14	0.50	2.10	0.41	2.10	0.42	3.09	0.32	3.09	0.33
2.25	2.04	0.28	2.05	0.28	2.03	0.23	2.03	0.23	3.03	0.18	3.03	0.18
2.50	2.01	0.14	2.01	0.15	2.01	0.12	2.01	0.12	3.01	0.09	3.01	0.09
2.75	2.00	0.07	2.00	0.07	2.00	0.06	2.00	0.06	3.00	0.05	3.00	0.05
3.00	2.00	0.03	2.00	0.03	2.00	0.02	2.00	0.02	3.00	0.02	3.00	0.02
	(a=72,b=429)		**(a=71,b=430)**		**(a=81,b=420)**		**(a=80,b=421)**		**(a=72,b=429)**		**(a=71,b=430)**	

	1-of-1				X-bar	
Shift	***ARL***	***SDRL***	***ARL***	***SDRL***	***ARL***	***SDRL***
0.00	460.22	538.61	520.27	613.67	500.00	571.14
0.25	233.27	290.26	261.60	329.17	184.12	216.66
0.50	70.42	85.43	77.73	95.38	43.38	48.51
0.75	23.74	27.01	25.79	29.64	13.12	13.71
1.00	9.58	10.11	10.26	10.93	5.19	4.93
1.25	4.63	4.43	4.88	4.72	2.63	2.15
1.50	2.66	2.21	2.76	2.34	1.67	1.08
1.75	1.78	1.22	1.83	1.28	1.26	0.58
2.00	1.36	0.72	1.39	0.75	1.09	0.32
2.25	1.16	0.44	1.17	0.45	1.03	0.17
2.50	1.06	0.26	1.07	0.27	1.01	0.08
2.75	1.02	0.15	1.02	0.16	1.00	0.03
3.00	1.01	0.08	1.01	0.09	1.00	0.01
	(a=25,b=476)		**(a=24,b=477)**		**3.084500892**	

Table 4.23: Unconditional *ARL* and *SDRL* values for the *2-of-2* DR, *2-of-2* KL, *2-of-3* and the Basic (*1-of-1*) Precedence chart for the *t*(4) distribution when $m = 500$, $n = 5$, $j = 3$

	***2-of-2* DR**				***2-of-2* KL**			
Shift	***ARL***	***SDRL***	***ARL***	***SDRL***	***ARL***	***SDRL***	***ARL***	***SDRL***
0.00	496.90	573.05	536.72	621.20	490.21	554.18	524.39	594.55
0.25	200.92	248.00	215.95	268.18	138.19	170.25	146.90	182.03
0.50	38.68	45.31	40.98	48.41	25.09	27.66	26.28	29.19
0.75	10.01	9.77	10.41	10.26	7.43	6.66	7.65	6.92
1.00	4.26	3.11	4.35	3.22	3.61	2.36	3.67	2.43
1.25	2.72	1.33	2.75	1.37	2.52	1.08	2.54	1.10
1.50	2.23	0.67	2.24	0.68	2.17	0.55	2.17	0.56
1.75	2.07	0.36	2.08	0.36	2.05	0.30	2.05	0.30
2.00	2.02	0.19	2.02	0.20	2.02	0.16	2.02	0.16
2.25	2.01	0.11	2.01	0.11	2.00	0.09	2.00	0.09
2.50	2.00	0.06	2.00	0.06	2.00	0.05	2.00	0.05
2.75	2.00	0.03	2.00	0.03	2.00	0.03	2.00	0.03
3.00	2.00	0.02	2.00	0.02	2.00	0.02	2.00	0.02
	(*a*=72,*b*=429)		**(*a*=71,*b*=430)**		**(*a*=81,*b*=420)**		**(*a*=80,*b*=421)**	
	2-of-3				***1-of-1***			
Shift	***ARL***	***SDRL***	***ARL***	***SDRL***	***ARL***	***SDRL***	***ARL***	***SDRL***
0.00	494.18	569.01	532.74	615.81	460.22	538.61	520.27	613.67
0.25	138.65	175.16	147.99	189.23	288.43	370.47	328.18	426.13
0.50	25.36	27.21	26.64	28.92	102.82	143.88	117.63	167.75
0.75	8.19	6.28	8.43	6.57	32.84	45.71	37.43	53.44
1.00	4.53	2.14	4.59	2.20	11.19	14.51	12.58	16.84
1.25	3.49	0.92	3.51	0.94	4.47	5.01	4.91	5.71
1.50	3.16	0.53	3.16	0.46	2.25	1.97	2.40	2.20
1.75	3.05	0.35	3.05	0.38	1.46	0.90	1.51	0.98
2.00	3.02	0.18	3.01	0.13	1.16	0.46	1.18	0.49
2.25	3.01	0.60	3.00	0.07	1.05	0.24	1.06	0.26
2.50	3.00	0.04	3.00	0.05	1.02	0.13	1.02	0.14
2.75	3.00	0.02	3.00	0.02	1.01	0.07	1.01	0.08
3.00	3.00	0.01	3.00	0.01	1.00	0.04	1.00	0.04
	(*a*=72,*b*=429)		**(*a*=71,*b*=430)**		**(*a*=25,*b*=476)**		**(*a*=24,*b*=477)**	

Table 4.24: Unconditional *ARL* and *SDRL* values for the *2-of-2* DR, *2-of-2* KL *2-of-3*and the Basic (*1-of-1*) Precedence chart for the *gamma*(1,1) distribution when $m = 500$, $n = 5$, $j = 3$

	2-of-2 DR				*2-of-2* KL			
Shift	***ARL***	***SDRL***	***ARL***	***SDRL***	***ARL***	***SDRL***	***ARL***	***SDRL***
0.00	496.90	573.05	536.72	621.20	490.21	554.18	524.39	594.55
0.25	233.82	815.92	250.23	887.98	310.12	405.49	331.64	436.02
0.50	58.22	216.59	61.33	234.84	88.52	111.41	94.24	119.33
0.75	17.55	61.43	18.23	66.27	28.03	33.05	29.65	35.23
1.00	7.36	18.96	7.56	20.33	10.26	10.74	10.75	11.39
1.25	4.12	6.42	4.19	6.84	4.55	3.79	4.72	4.00
1.50	2.88	2.30	2.91	2.45	2.61	1.34	2.66	1.42
1.75	2.36	0.74	2.37	0.80	2.05	0.34	2.06	0.37
2.00	2.13	0.13	2.14	0.15	2.00	0.03	2.00	0.04
2.25	2.04	0.00	2.05	0.00	2.00	0.00	2.00	0.00
2.50	2.01	0.00	2.01	0.00	2.00	0.00	2.00	0.00
2.75	2.00	0.00	2.00	0.00	2.00	0.00	2.00	0.00
3.00	2.00	0.00	2.00	0.00	2.00	0.00	2.00	0.00
	(*a*=72,*b*=429)		**(*a*=71,*b*=430)**		**(*a*=81,*b*=420)**		**(*a*=80,*b*=421)**	
	2-of-3				***1-of-1***			
Shift	***ARL***	***SDRL***	***ARL***	***SDRL***	***ARL***	***SDRL***	***ARL***	***SDRL***
0.00	494.18	569.01	532.74	615.81	460.22	538.61	520.27	613.67
0.25	314.31	425.10	336.97	459.01	527.27	730.48	600.16	844.59
0.50	90.55	115.92	98.04	126.21	255.49	351.96	290.46	406.50
0.75	30.22	35.15	31.84	37.14	124.76	170.53	141.61	196.68
1.00	11.98	11.68	12.57	12.35	61.56	83.20	69.72	95.80
1.25	6.01	4.18	6.23	4.48	30.80	40.94	34.78	47.05
1.50	3.88	1.58	3.96	1.67	15.70	20.35	17.67	23.33
1.75	3.15	0.51	3.17	0.54	8.22	10.23	9.20	11.69
2.00	3.01	0.08	3.01	0.09	4.47	5.19	4.96	5.92
2.25	3.00	0.01	3.00	0.01	2.58	2.64	2.83	3.02
2.50	3.00	0.00	3.00	0.00	1.63	1.32	1.75	1.52
2.75	3.00	0.00	3.00	0.00	1.19	0.61	1.24	0.72
3.00	3.00	0.00	3.00	0.00	1.03	0.23	1.05	0.29
	(*a*=72,*b*=429)		**(*a*=71,*b*=430)**		**(*a*=25,*b*=476)**		**(*a*=24,*b*=477)**	

Because the run-length distribution is highly right-skewed, exclusive use of the *ARL* (and the *SDRL*) to characterize chart performance has been criticized in the literature and some researchers have strongly suggested an examination of the percentiles too (see e.g. Radson and Boyd, (2005) and Chakraborti, (2007)).

To this end, the three quartiles (Q_1, Q_2 and Q_3) along with the 5^{th} and the 95^{th} percentiles are shown in Tables 4.25, 4.26, 4.27 and 4.28, for the *2-of-2* DR, the *2-of-2* KL, the *2-of-3* and the *1-of-1* precedence chart, for the normal, $t(4)$ and $Gamma(1,1)$ distributions, respectively for $n = 5$ and $j = 3$. Note that (i) these values are all *unconditional* i.e. being averaged over the joint distribution of the order statistics $X_{a:m}$ and $X_{b:m}$, and (ii) these values were obtained via simulations (200 000 repetitions) using SAS®9.1 since exact calculations, via enumeration of the c.d.f, took too long for the upper percentiles. The SAS®-programs used in the simulations are provided in Appendix 4A.

A comparison of the quartiles lead to the same general observation that the newly proposed nonparametric charts are more efficient than the basic precedence chart, with the *2-of-2* KL and the *2-of-3* charts having a slight edge.

For example, in the in-control case and with the *t*-distribution, for the *2-of-2* DR chart (with $a = 72$ & $b = 429$) the three quartiles are 127, 313 and 658, respectively, which are very close to those for the *2-of-2* KL chart (with $a = 81$ & $b = 420$) and the *2-of-3* chart (with $a = 72$ & $b = 429$): 126, 312 & 650 and 127, 312 & 653, respectively. By contrast, for the *1-of-1* precedence chart (with $a = 25$ & $b = 476$) the three quartiles are 116, 287 and 603, respectively, which are all smaller. Since we want the in-control percentiles to be larger, the new charts are better. On the other hand, in the out-of-control case, for a shift of 0.50 in the mean, the quartiles for the *2-of-2* KL chart (with $a = 81$ & $b = 420$) and *2-of-3* chart (with $a = 72$ & $b = 429$) are all shorter: 7, 16 & 33 and 8, 17 & 32, respectively, compared to both the *2-of-2* DR chart (with $a = 72$ & $b = 429$): 11, 24 and 50 and the *1-of-1* precedence chart (with $a = 25$ & $b = 476$): 23, 57 and 127. This shows that the *2-of-2* KL and the *2-of-3* charts are superior.

Table 4.25: The three quartiles (Q_1, Q_2 and Q_3) and the 5th & 95th percentiles of the run-length distribution of the *2-of-2* DR chart; charting constants (a=72 , b=429)

	normal					*t*(4)					*gamma*(1,1)				
Shift	**5th**	**Q_1**	**Q_2**	**Q_3**	**95th**	**5th**	**Q_1**	**Q_2**	**Q_3**	**95th**	**5th**	**Q_1**	**Q_2**	**Q_3**	**95th**
0.00	23	128	314	657	1587	23	127	313	658	1588	23	127	313	653	1586
0.25	11	58	144	306	758	10	49	121	260	662	23	129	329	734	2002
0.50	4	16	37	76	183	3	11	24	50	122	8	37	93	205	547
0.75	2	6	12	23	52	2	4	7	13	29	3	12	29	61	158
1.00	2	3	5	10	20	2	2	3	5	10	2	5	10	21	51
1.25	2	2	3	5	10	2	2	2	3	5	2	2	4	8	18
1.50	2	2	2	3	6	2	2	2	2	4	2	2	2	4	8
1.75	2	2	2	2	4	2	2	2	2	2	2	2	2	2	4
2.00	2	2	2	2	3	2	2	2	2	2	2	2	2	2	2
2.25	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
2.50	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
2.75	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
3.00	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2

Table 4.26: The three quartiles (Q_1, Q_2 and Q_3) and the 5th & 95th percentiles of the run-length distribution of the *2-of-2* KL chart; charting constants (a=81 , b=420)

	normal					*t*(4)					*gamma*(1,1)				
Shift	**5th**	**Q_1**	**Q_2**	**Q_3**	**95th**	**5th**	**Q_1**	**Q_2**	**Q_3**	**95th**	**5th**	**Q_1**	**Q_2**	**Q_3**	**95th**
0.00	23	128	313	652	1554	24	126	312	650	1549	24	128	314	655	1552
0.25	9	43	105	223	549	7	34	84	179	449	14	72	181	394	1037
0.50	3	11	26	52	122	2	7	16	33	77	5	22	53	114	292
0.75	2	4	9	17	38	2	3	5	10	20	2	8	17	36	88
1.00	2	2	4	8	16	2	2	3	4	8	2	4	7	13	30
1.25	2	2	3	4	8	2	2	2	3	5	2	2	3	6	12
1.50	2	2	2	3	5	2	2	2	2	4	2	2	2	3	5
1.75	2	2	2	2	4	2	2	2	2	2	2	2	2	2	2
2.00	2	2	2	2	3	2	2	2	2	2	2	2	2	2	2
2.25	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
2.50	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
2.75	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
3.00	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2

Table 4.27: The three quartiles (Q_1, Q_2 and Q_3) and the 5th & 95th percentiles of the run-length distribution of the *2-of-3* chart; charting constants (a=72 , b=429)

	normal					*t*(4)					*gamma*(1,1)				
Shift	**5th**	**Q_1**	**Q_2**	**Q_3**	**95th**	**5th**	**Q_1**	**Q_2**	**Q_3**	**95th**	**5th**	**Q_1**	**Q_2**	**Q_3**	**95th**
0.00	25	127	312	653	1576	24	127	312	653	1580	25	128	313	655	1582
0.25	10	41	99	208	524	8	34	82	176	457	15	72	179	392	1051
0.50	4	11	24	48	111	4	8	17	32	76	6	23	54	114	295
0.75	3	5	9	16	34	3	4	6	10	20	4	9	19	38	94
1.00	3	4	5	8	15	3	3	4	5	9	3	5	8	15	33
1.25	3	3	4	5	8	3	3	3	4	5	3	3	4	7	14
1.50	3	3	3	4	6	3	3	3	3	4	3	3	3	4	7
1.75	3	3	3	3	4	3	3	3	3	3	3	3	3	3	4
2.00	3	3	3	3	4	3	3	3	3	3	3	3	3	3	3
2.25	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
2.50	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
2.75	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
3.00	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3

Table 4.28: The three quartiles (Q_1, Q_2 and Q_3) and the 5th & 95th of the run-length distribution of the Basic (*1-of-1*) precedence chart; charting constants (a=25 , b=476)

	normal					*t*(4)					*gamma*(1,1)				
Shift	**5th**	**Q_1**	**Q_2**	**Q_3**	**95th**	**5th**	**Q_1**	**Q_2**	**Q_3**	**95th**	**5th**	**Q_1**	**Q_2**	**Q_3**	**95th**
0.00	21	116	288	608	1489	21	116	287	603	1479	21	115	287	606	1477
0.25	10	56	140	300	765	12	66	168	368	956	20	115	294	655	1786
0.50	3	18	43	92	228	4	23	57	127	350	10	57	144	319	864
0.75	2	6	15	31	75	2	8	19	41	110	5	28	71	157	420
1.00	1	3	6	13	29	1	3	7	14	36	3	14	35	77	206
1.25	1	2	3	6	13	1	1	3	6	13	2	7	18	39	103
1.50	1	1	2	3	7	1	1	2	3	6	1	4	9	20	51
1.75	1	1	1	2	4	1	1	1	2	3	1	2	5	10	26
2.00	1	1	1	2	3	1	1	1	1	2	1	1	3	5	14
2.25	1	1	1	1	2	1	1	1	1	2	1	1	2	3	7
2.50	1	1	1	1	2	1	1	1	1	1	1	1	1	2	4
2.75	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2
3.00	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

4.4 Concluding remarks: Summary and recommendations

The new class of runs-rule enhanced nonparametric sign charts of Case K and the distribution-free precedence charts of Case U can be useful for the quality practitioner in that they

(i) enhance the in-control and the out-of-control performance of the *1-of-1* sign chart of Amin et al. (1995) and the basic precedence chart of Chakraborti et al. (2004), respectively, and

(ii) outperform the classical and well-known Shewhart $\overline{X}$ chart (especially for heavy-tailed or skewed distributions).

In particular, the charts based on the *k-of-k* and the *k-of-w* signaling rules facilitate larger ARL_0 and smaller FAR values which allow practitioners greater flexibility while designing charts to best suit their needs.

The key advantage and main benefits of the nonparametric charts are:

(i) their in-control run-length distributions (and all associated performance characteristics such as the ARL_0 and FAR, for example) are the same for all continuous distributions, and

(ii) one does not have to assume symmetry of the underlying distribution (unlike the SR charts). Thus, practitioners need not worry about what the underlying distribution is (and the serious consequences/ramifications/costs if it is not normal, for example) as far as implementing and understanding the charts' properties are concerned.

The sign charts have an added advantage as they can be applied in situations where the data are just dichotomous.

A further practical advantage of the precedence charts is their potential to save time and resources in situations where the data are naturally collected in an ordered fashion, as is common in "life-testing" type situations, where one observes the "time to failure" of some item and it is costly and time consuming to wait for all units to fail. Because the control limits and the charting statistic of the precedence charts are based on order statistics, they can be applied as soon as the required order statistics are observed, whereas the Shewhart or CUSUM or EWMA $\overline{X}$ charts can not be applied since one needs the full dataset to calculate the average.

Also, the precedence charts can be adapted to and applied in the case of ordinal data. The charting statistic can be chosen to be any order statistic of the Phase II sample suitable in a specific application. The median, used in this chapter, of course enjoys the robustness property and is therefore less affected by the presence of outliers (very small or large observations) than the $\overline{X}$ chart, for example.

Finally, the implementation and application of the sign and precedence charts are easy using the tables with the charting constants (and attained ARL_0 and FAR values) and it is recommend that they be used more frequently in practice.

4.5 Appendix 4A: SAS® programs

4.5.1 SAS® programs to simulate the run-length distributions of the upper one-sided X-bar, sign and SR charts in Case K

4.5.1.1 The *1-of-1* X-bar, sign and SR charts

```
*1-of-1 upper one-sided X-bar chart;

proc iml;
ARL = 370;
sim = 100000;
n = 10;
UCL = probit(1-1/ARL)/sqrt(n);
simrl = j(sim,13,.);
do d = 1 to 2.2 by 0.2;
do j = 1 to sim;
ct = 0;
do k = 1 to 10000000 while ( ^((ct>=UCL))  );
x = j(n,1,.);
call randgen(x,'NORMAL',d-1,1);
ct = sum(x)/n;
rl = k;
end;
simrl[j,d*5+1]=rl;
end;
end;
create RL1of1_Xbar from simrl[colname={delta000
delta020 delta040 delta060 delta080 delta100 delta120}];
append from simrl;
quit;
proc univariate data=RL1of1_Xbar;
var delta000 delta020 delta040 delta060 delta080 delta100 delta120;
run;
```

```
*1-of-1 upper one-sided sign chart;

proc iml;
ARL = 370;
sim = 100000;
a = 0;
n = 10;
UCL = n-a;
med = j(n,1,0);
simrl = j(sim,13,.);
q = (1/ARL - 1 + probbnml(0.5,n,n-a-1) ) / (probbnml(0.5,n,n-a-1) - probbnml(0.5,n,n-
a-2)) ;
do d = 1 to 2.2 by 0.2;
do j = 1 to sim;
ct = 0;
random = 0;
do k = 1 to 10000000 while ( ^((ct>=UCL)|((ct=UCL-1)&(random<=q)))  );
x = j(n,1,.);
call randgen(x,'NORMAL',d-1,1);
vec = x > med;
ct = sum(vec);
random = ranuni(0);
rl = k;
end;
simrl[j,d*5+1]=rl;
end;
end;
*print simrl;
create RL1of1_sign from simrl[colname={delta000
delta020 delta040 delta060 delta080 delta100 delta120}];
append from simrl;
quit;
proc univariate data=RL1of1_sign;
var delta000 delta020 delta040 delta060 delta080 delta100 delta120;
run;
```

```
*1-of-1 upper one-sided SR chart;

proc iml;
ARL = 370;
sim = 100000;
n = 10;
UCL = 53;
UCL1 = 51;
cdfUCL = 0.002;
pmfUCL1 = 0.0009;
med = j(n,1,0);
simrl = j(sim,13,.);
q = (1/ARL - cdfUCL) / ( pmfUCL1 ) ;
do d = 1 to 2.2 by 0.2;
do j = 1 to sim;
random = 0;
ct = 0;
do k = 1 to 10000000 while ( ^(   (ct>=UCL)  |  ((ct=UCL1)&(random<=q))   )  );
x = j(n,1,.);
call randgen(x,'NORMAL',d-1,1);
vec   = x > med;
wplus = (vec`)*rank(abs(x));
ct = 2*wplus - n*(n+1)/2;
random = ranuni(0);
rl = k;
end;
simrl[j,d*5+1]=rl;
end;
end;
create RL1of1_SR from simrl[colname={delta000
delta020 delta040 delta060 delta080 delta100 delta120}];
append from simrl;
quit;
proc univariate data=RL1of1_SR;
var delta000 delta020 delta040 delta060 delta080 delta100 delta120;
run;
```

4.5.1.2 The *2-of-2* sign and SR charts

```
*2-of-2 upper one-sided sign chart;

proc iml;
ARL = 370;
sim = 100000;
a = 1;
n = 10;
UCL = n-a;
med = j(n,1,0);
simrl = j(sim,13,.);
q = ( (sqrt(4*ARL+1)+1)/(2*ARL) - 1 + probbnml(0.5,n,n-a-1) ) /
    (probbnml(0.5,n,n-a-1) - probbnml(0.5,n,n-a-2)) ;
do d = 1 to 2.2 by 0.2;
do j = 1 to sim;
x = j(n,1,.);
ct1 = 0;
ct = 0;
random1 = 0;
random = 0;
do k = 1 to 10000000 while (
^( (  (ct1>=UCL)&(ct>=UCL)  )                        |
   (  ((ct1=UCL-1)&(random1<=q))&(ct>=ucl)  )        |
   (  (ct1>=UCL)&((ct=UCL-1)&(random<=q))  )         |
   (  ((ct1=UCL-1)&(random1<=q))&((ct=UCL-1)&(random<=q)))  )  );
ct1 = ct;
call randgen(x,'NORMAL',d-1,1);
vec = x > med;
ct = sum(vec);
random = ranuni(0);
random1 = ranuni(0);
rl = k;
end;
simrl[j,d*5+1]=rl;
end;
end;
create RL2of2_sign from simrl[colname={delta000
delta020 delta040 delta060 delta080 delta100 delta120}];
append from simrl;
quit;
proc univariate data=RL2of2_sign;
var delta000 delta020 delta040 delta060 delta080 delta100 delta120;
run;
```

```
*2-of-2 upper one-sided SR chart;

proc iml;
ARL = 370;
sim = 100000;
n = 10;
UCL = 33;
UCL1 = UCL - 2;
cdfUCL = 0.0527;
pmfUCL1 = 0.0127;
med = j(n,1,0);
simrl = j(sim,13,.);
q = (  (sqrt(4*ARL+1)+1)/(2*ARL) -  cdfUCL ) / ( pmfUCL1 ) ;
do d = 1 to 2.2 by 0.2;
do j = 1 to sim;
x = j(n,1,.);
ct1 = 0;
ct = 0;
random1 = 0;
random = 0;
do k = 1 to 10000000 while (
^( (  (ct1>=UCL)&(ct>=UCL)  )                        |
 (  ((ct1=UCL1)&(random1<=q))&(ct>=ucl)  )     |
 (  (ct1>=UCL)&((ct=UCL1)&(random<=q))  )      |
 (  ((ct1=UCL1)&(random1<=q))&((ct=UCL1)&(random<=q)))  )  );
ct1 = ct;
call randgen(x,'NORMAL',d-1,1);
vec   = x > med;
wplus = (vec`)*rank(abs(x));
ct = 2*wplus - n*(n+1)/2;
random = ranuni(0);
random1 = ranuni(0);
rl = k;
end;
simrl[j,d*5+1]=rl;
end;
end;
create RL2of2_SR from simrl[colname={delta000
delta020 delta040 delta060 delta080 delta100 delta120}];
append from simrl;
quit;
proc univariate data=RL2of2_SR;
var delta000 delta020 delta040 delta060 delta080 delta100 delta120;
run;
```

4.5.1.3 The *2-of-3* sign chart

```
*2-of-3 upper-sided sign chart;

proc iml;
ARL = 370;
sim = 100000;
a = 1;
n = 10;
UCL = n-a;
med = j(n,1,0);
simrl = j(sim,13,.);
q = 0.632202808;
do d = 1 to 2.2 by 0.2;
do j = 1 to sim;
x = j(n,1,.);
ct2 = 0;
ct1 = 0;
ct = 0;
random2 = 0;
random1 = 0;
random = 0;
do k = 1 to 10000000 while (
^(
 ( (ct2>=UCL)                 & (ct1<UCL) &  (ct>=UCL)               )     |
 ( ((ct2=UCL-1)&(random2<=q)) & (ct1<UCL) &  (ct>=UCL)               )     |
 ( (ct2>=UCL)                 & (ct1<UCL) & ((ct=UCL-1)&(random<=q)) )     |
 ( ((ct2=UCL-1)&(random2<=q)) & (ct1<UCL) & ((ct=UCL-1)&(random<=q)) )     |
 ( (ct1>=UCL)                 & (ct2<UCL) &  (ct>=UCL)               )     |
 ( ((ct1=UCL-1)&(random1<=q)) & (ct2<UCL) &  (ct>=UCL)               )     |
 ( (ct1>=UCL)                 & (ct2<UCL) & ((ct=UCL-1)&(random<=q)) )     |
 ( ((ct1=UCL-1)&(random1<=q)) & (ct2<UCL) & ((ct=UCL-1)&(random<=q)) )      )   );
ct2 = ct1;
ct1 = ct;
call randgen(x,'NORMAL',d-1,1);
vec = x > med;
ct = sum(vec);
random = ranuni(0);
random1 = ranuni(0);
random2 = ranuni(0);
rl = k;
end;
simrl[j,d*5+1]=rl;
end;
end;
ARL = sum(simrl)/sim;
create RL2of3_sign from simrl[colname={delta000
delta020 delta040 delta060 delta080 delta100 delta120}];
append from simrl;
quit;
proc univariate data=RL2of3_sign;
var delta000 delta020 delta040 delta060 delta080 delta100 delta120;
run;
```

4.5.2 SAS® programs to simulate the run-length distributions of the two-sided precedence charts in Case U

4.5.2.1 The *1-of-1* precedence chart

```
proc iml;
m = 500;
n = 5;
j = (n+1)/2;
sim = 100000;
a = 25;
b = 476;
rlvec = j(sim,13,.);
do delta = 1 to 3 by 0.25;
shift = 4*delta-3;
do k = 1 to sim;
xref = j(m,1,0);
call randgen(xref,'NORMAL'); yref = xref;
call sort(yref, {1});
lcl = yref[a,1];
ucl = yref[b,1];
count = 1;
signal = 0;
above = j(2,1,0);
below = j(2,1,1);
do while (signal = 0);
xfut = j(n,1,0);
call randgen(xfut,'NORMAL',delta-1,1); yfut = xfut;
call sort(yfut, {1});
plotstat = yfut[j,1];
cl=j(2,1,0);
cl[1,1]=ucl;
cl[2,1]=lcl;
plotstatvec=j(2,1,plotstat);
check = plotstatvec <= cl;
if check = above then signal = 1;
else if check = below then signal = 1;
else count = count + 1;
count1 = count;
rlvec[k,shift] = count1;
end;
end;
end;
create RL1of1_Precedence_normal from rlvec[colname={delta000 delta025 delta050
delta075 delta100 delta125 delta150 delta175 delta200 delta225 delta250 delta275
delta300}];
append from rlvec;
quit;
proc univariate data=RL1of1_Precedence_normal;
var delta000 delta025 delta050 delta075 delta100 delta125 delta150 delta175
    delta200 delta225 delta250 delta275 delta300;
run;
```

4.5.2.2 The *2-of-2* DR and the *2-of-2* KL precedence charts

```
proc iml;
m = 500; n = 5; j = (n+1)/2;
sim = 100000;
a = 81; b = 420;
rlvec = j(sim,13,.);
do delta = 1 to 3 by 0.25;
shift = 4*delta-3;
do k = 1 to sim;
xref = j(m,1,0);
call randgen(xref,'NORMAL'); yref = xref;
call sort(yref, {1});
lcl = yref[a,1];
ucl = yref[b,1];
count = 1;
signal = 0;
dummy = j(2,1,0);
check = {1,0};
above = j(2,2,0);
below = j(2,2,1);
abovebelow = {0 1 , 0 1};
belowabove = {1 0 , 1 0};
matrix = j(2,2,0);
do while (signal = 0);
dummy = check;
xfut = j(n,1,0);
call randgen(xfut,'NORMAL',delta-1,1); yfut = xfut;
call sort(yfut, {1});
plotstat = yfut[j,1];
cl=j(2,1,0);
cl[1,1]=ucl;
cl[2,1]=lcl;
plotstatvec=j(2,1,plotstat);
check = plotstatvec <= cl;
matrix = dummy||check;
if matrix = above then signal = 1; * DR and KL;
else if matrix = below then signal = 1; * DR and KL;
else if matrix = abovebelow then signal = 1; * DR only;
else if matrix = belowabove then signal = 1; * DR only;
else count = count + 1;
count1 = count;
rlvec[k,shift] = count1;
end;
end;
end;
create RL2of2_Precedence_normal from rlvec[colname={delta000 delta025 delta050
delta075 delta100 delta125 delta150 delta175 delta200 delta225 delta250 delta275
delta300}];
append from rlvec;
quit;
proc univariate data= RL2of2_Precedence_normal;
var delta000 delta025 delta050 delta075 delta100 delta125 delta150 delta175
    delta200 delta225 delta250 delta275 delta300;
run;
```

4.5.2.3 The *2-of-3* precedence chart

```
proc iml;
m = 500; n = 5; j = (n+1)/2;
sim = 100000;
a = 72; b = 429;
rlvec = j(sim,13,.);
do delta = 1 to 3 by 0.25;
shift = 4*delta-3;
do k = 1 to sim;
xref = j(m,1,0);
call randgen(xref,'NORMAL'); yref = xref;
call sort(yref, {1});
lcl = yref[a,1];
ucl = yref[b,1];
count = 2;
signal = 0;
dummy1 = {1,0}; dummy2 = j(2,1,.); check = {1,0};
between_above_above = {1 0 0,
                       0 0 0};
between_below_below = {1 1 1,
                        0 1 1};
below_between_below = {1 1 1,
                       1 0 1};
above_between_above = {0 1 0,
                       0 0 0};
matrix = j(2,3,.);
do while (signal = 0);
dummy2 = dummy1;
dummy1 = check;
xfut = j(n,1,0);
call randgen(xfut,'NORMAL',delta-1,1); yfut = xfut;
call sort(yfut, {1});
plotstat = yfut[j,1];
cl=j(2,1,0);
cl[1,1]=ucl;
cl[2,1]=lcl;
plotstatvec=j(2,1,plotstat);
check = plotstatvec <= cl;
matrix = dummy2||dummy1||check;
if matrix = between_above_above then signal = 1;
else if matrix = between_below_below then signal = 1;
else if matrix = below_between_below then signal = 1;
else if matrix = above_between_above then signal = 1;
else count = count + 1;
count1 = count;
rlvec[k,shift] = count1;
end;
end;
end;
create RL2of3_Precedence_normal from rlvec[colname={delta000 delta025 delta050
delta075 delta100 delta125 delta150 delta175 delta200 delta225 delta250 delta275
delta300}];
append from rlvec;
quit;
proc univariate data= RL2of3_Precedence_normal;
var delta000 delta025 delta050 delta075 delta100 delta125 delta150 delta175
    delta200 delta225 delta250 delta275 delta300;
run;
```

4.5 Appendix 4A: SAS® programs

4.5.1 SAS® programs to simulate the run-length distributions of the upper one-sided X-bar, sign and SR charts in Case K

4.5.1.1 The *1-of-1* X-bar, sign and SR charts

```
*1-of-1 upper one-sided X-bar chart;

proc iml;
ARL = 370;
sim = 100000;
n = 10;
UCL = probit(1-1/ARL)/sqrt(n);
simrl = j(sim,13,.);
do d = 1 to 2.2 by 0.2;
do j = 1 to sim;
ct = 0;
do k = 1 to 10000000 while ( ^((ct>=UCL))  );
x = j(n,1,.);
call randgen(x,'NORMAL',d-1,1);
ct = sum(x)/n;
rl = k;
end;
simrl[j,d*5+1]=rl;
end;
end;
create RL1of1_Xbar from simrl[colname={delta000
delta020 delta040 delta060 delta080 delta100 delta120}];
append from simrl;
quit;
proc univariate data=RL1of1_Xbar;
var delta000 delta020 delta040 delta060 delta080 delta100 delta120;
run;
```

```
*1-of-1 upper one-sided sign chart;

proc iml;
ARL = 370;
sim = 100000;
a = 0;
n = 10;
UCL = n-a;
med = j(n,1,0);
simrl = j(sim,13,.);
q = (1/ARL - 1 + probbnml(0.5,n,n-a-1) ) / (probbnml(0.5,n,n-a-1) - probbnml(0.5,n,n-
a-2)) ;
do d = 1 to 2.2 by 0.2;
do j = 1 to sim;
ct = 0;
random = 0;
do k = 1 to 10000000 while ( ^((ct>=UCL)|((ct=UCL-1)&(random<=q)))  );
x = j(n,1,.);
call randgen(x,'NORMAL',d-1,1);
vec = x > med;
ct = sum(vec);
random = ranuni(0);
rl = k;
end;
simrl[j,d*5+1]=rl;
end;
end;
*print simrl;
create RL1of1_sign from simrl[colname={delta000
delta020 delta040 delta060 delta080 delta100 delta120}];
append from simrl;
quit;
proc univariate data=RL1of1_sign;
var delta000 delta020 delta040 delta060 delta080 delta100 delta120;
run;
```

```
*1-of-1 upper one-sided SR chart;

proc iml;
ARL = 370;
sim = 100000;
n = 10;
UCL = 53;
UCL1 = 51;
cdfUCL = 0.002;
pmfUCL1 = 0.0009;
med = j(n,1,0);
simrl = j(sim,13,.);
q = (1/ARL - cdfUCL) / ( pmfUCL1 ) ;
do d = 1 to 2.2 by 0.2;
do j = 1 to sim;
random = 0;
ct = 0;
do k = 1 to 10000000 while ( ^(   (ct>=UCL)  |  ((ct=UCL1)&(random<=q))   )  );
x = j(n,1,.);
call randgen(x,'NORMAL',d-1,1);
vec   = x > med;
wplus = (vec`)*rank(abs(x));
ct = 2*wplus - n*(n+1)/2;
random = ranuni(0);
rl = k;
end;
simrl[j,d*5+1]=rl;
end;
end;
create RL1of1_SR from simrl[colname={delta000
delta020 delta040 delta060 delta080 delta100 delta120}];
append from simrl;
quit;
proc univariate data=RL1of1_SR;
var delta000 delta020 delta040 delta060 delta080 delta100 delta120;
run;
```

4.5.1.2 The *2-of-2* sign and SR charts

```
*2-of-2 upper one-sided sign chart;

proc iml;
ARL = 370;
sim = 100000;
a = 1;
n = 10;
UCL = n-a;
med = j(n,1,0);
simrl = j(sim,13,.);
q = ( (sqrt(4*ARL+1)+1)/(2*ARL) - 1 + probbnml(0.5,n,n-a-1) ) /
    (probbnml(0.5,n,n-a-1) - probbnml(0.5,n,n-a-2)) ;
do d = 1 to 2.2 by 0.2;
do j = 1 to sim;
x = j(n,1,.);
ct1 = 0;
ct = 0;
random1 = 0;
random = 0;
do k = 1 to 10000000 while (
^( (  (ct1>=UCL)&(ct>=UCL)  )                        |
   (  ((ct1=UCL-1)&(random1<=q))&(ct>=ucl)  )        |
   (  (ct1>=UCL)&((ct=UCL-1)&(random<=q))  )         |
   (  ((ct1=UCL-1)&(random1<=q))&((ct=UCL-1)&(random<=q)))  )  );
ct1 = ct;
call randgen(x,'NORMAL',d-1,1);
vec = x > med;
ct = sum(vec);
random = ranuni(0);
random1 = ranuni(0);
rl = k;
end;
simrl[j,d*5+1]=rl;
end;
end;
create RL2of2_sign from simrl[colname={delta000
delta020 delta040 delta060 delta080 delta100 delta120}];
append from simrl;
quit;
proc univariate data=RL2of2_sign;
var delta000 delta020 delta040 delta060 delta080 delta100 delta120;
run;
```

```
*2-of-2 upper one-sided SR chart;

proc iml;
ARL = 370;
sim = 100000;
n = 10;
UCL = 33;
UCL1 = UCL - 2;
cdfUCL = 0.0527;
pmfUCL1 = 0.0127;
med = j(n,1,0);
simrl = j(sim,13,.);
q = (  (sqrt(4*ARL+1)+1)/(2*ARL) -  cdfUCL ) / ( pmfUCL1 ) ;
do d = 1 to 2.2 by 0.2;
do j = 1 to sim;
x = j(n,1,.);
ct1 = 0;
ct = 0;
random1 = 0;
random = 0;
do k = 1 to 10000000 while (
^( (  (ct1>=UCL)&(ct>=UCL)  )                        |
 (  ((ct1=UCL1)&(random1<=q))&(ct>=ucl)  )     |
 (  (ct1>=UCL)&((ct=UCL1)&(random<=q))  )      |
 (  ((ct1=UCL1)&(random1<=q))&((ct=UCL1)&(random<=q)))  )  );
ct1 = ct;
call randgen(x,'NORMAL',d-1,1);
vec   = x > med;
wplus = (vec`)*rank(abs(x));
ct = 2*wplus - n*(n+1)/2;
random = ranuni(0);
random1 = ranuni(0);
rl = k;
end;
simrl[j,d*5+1]=rl;
end;
end;
create RL2of2_SR from simrl[colname={delta000
delta020 delta040 delta060 delta080 delta100 delta120}];
append from simrl;
quit;
proc univariate data=RL2of2_SR;
var delta000 delta020 delta040 delta060 delta080 delta100 delta120;
run;
```

4.5.1.3 The *2-of-3* sign chart

```
*2-of-3 upper-sided sign chart;

proc iml;
ARL = 370;
sim = 100000;
a = 1;
n = 10;
UCL = n-a;
med = j(n,1,0);
simrl = j(sim,13,.);
q = 0.632202808;
do d = 1 to 2.2 by 0.2;
do j = 1 to sim;
x = j(n,1,.);
ct2 = 0;
ct1 = 0;
ct = 0;
random2 = 0;
random1 = 0;
random = 0;
do k = 1 to 10000000 while (
^(
 ( (ct2>=UCL)                 & (ct1<UCL) &  (ct>=UCL)               )     |
 ( ((ct2=UCL-1)&(random2<=q)) & (ct1<UCL) &  (ct>=UCL)               )     |
 ( (ct2>=UCL)                 & (ct1<UCL) & ((ct=UCL-1)&(random<=q)) )     |
 ( ((ct2=UCL-1)&(random2<=q)) & (ct1<UCL) & ((ct=UCL-1)&(random<=q)) )     |
 ( (ct1>=UCL)                 & (ct2<UCL) &  (ct>=UCL)               )     |
 ( ((ct1=UCL-1)&(random1<=q)) & (ct2<UCL) &  (ct>=UCL)               )     |
 ( (ct1>=UCL)                 & (ct2<UCL) & ((ct=UCL-1)&(random<=q)) )     |
 ( ((ct1=UCL-1)&(random1<=q)) & (ct2<UCL) & ((ct=UCL-1)&(random<=q)) )      )   );
ct2 = ct1;
ct1 = ct;
call randgen(x,'NORMAL',d-1,1);
vec = x > med;
ct = sum(vec);
random = ranuni(0);
random1 = ranuni(0);
random2 = ranuni(0);
rl = k;
end;
simrl[j,d*5+1]=rl;
end;
end;
ARL = sum(simrl)/sim;
create RL2of3_sign from simrl[colname={delta000
delta020 delta040 delta060 delta080 delta100 delta120}];
append from simrl;
quit;
proc univariate data=RL2of3_sign;
var delta000 delta020 delta040 delta060 delta080 delta100 delta120;
run;
```

4.5.2 SAS® programs to simulate the run-length distributions of the two-sided precedence charts in Case U

4.5.2.1 The *1-of-1* precedence chart

```
proc iml;
m = 500;
n = 5;
j = (n+1)/2;
sim = 100000;
a = 25;
b = 476;
rlvec = j(sim,13,.);
do delta = 1 to 3 by 0.25;
shift = 4*delta-3;
do k = 1 to sim;
xref = j(m,1,0);
call randgen(xref,'NORMAL'); yref = xref;
call sort(yref, {1});
lcl = yref[a,1];
ucl = yref[b,1];
count = 1;
signal = 0;
above = j(2,1,0);
below = j(2,1,1);
do while (signal = 0);
xfut = j(n,1,0);
call randgen(xfut,'NORMAL',delta-1,1); yfut = xfut;
call sort(yfut, {1});
plotstat = yfut[j,1];
cl=j(2,1,0);
cl[1,1]=ucl;
cl[2,1]=lcl;
plotstatvec=j(2,1,plotstat);
check = plotstatvec <= cl;
if check = above then signal = 1;
else if check = below then signal = 1;
else count = count + 1;
count1 = count;
rlvec[k,shift] = count1;
end;
end;
end;
create RL1of1_Precedence_normal from rlvec[colname={delta000 delta025 delta050
delta075 delta100 delta125 delta150 delta175 delta200 delta225 delta250 delta275
delta300}];
append from rlvec;
quit;
proc univariate data=RL1of1_Precedence_normal;
var delta000 delta025 delta050 delta075 delta100 delta125 delta150 delta175
    delta200 delta225 delta250 delta275 delta300;
run;
```

4.5.2.2 The *2-of-2* DR and the *2-of-2* KL precedence charts

```
proc iml;
m = 500; n = 5; j = (n+1)/2;
sim = 100000;
a = 81; b = 420;
rlvec = j(sim,13,.);
do delta = 1 to 3 by 0.25;
shift = 4*delta-3;
do k = 1 to sim;
xref = j(m,1,0);
call randgen(xref,'NORMAL'); yref = xref;
call sort(yref, {1});
lcl = yref[a,1];
ucl = yref[b,1];
count = 1;
signal = 0;
dummy = j(2,1,0);
check = {1,0};
above = j(2,2,0);
below = j(2,2,1);
abovebelow = {0 1 , 0 1};
belowabove = {1 0 , 1 0};
matrix = j(2,2,0);
do while (signal = 0);
dummy = check;
xfut = j(n,1,0);
call randgen(xfut,'NORMAL',delta-1,1); yfut = xfut;
call sort(yfut, {1});
plotstat = yfut[j,1];
cl=j(2,1,0);
cl[1,1]=ucl;
cl[2,1]=lcl;
plotstatvec=j(2,1,plotstat);
check = plotstatvec <= cl;
matrix = dummy||check;
if matrix = above then signal = 1; * DR and KL;
else if matrix = below then signal = 1; * DR and KL;
else if matrix = abovebelow then signal = 1; * DR only;
else if matrix = belowabove then signal = 1; * DR only;
else count = count + 1;
count1 = count;
rlvec[k,shift] = count1;
end;
end;
end;
create RL2of2_Precedence_normal from rlvec[colname={delta000 delta025 delta050
delta075 delta100 delta125 delta150 delta175 delta200 delta225 delta250 delta275
delta300}];
append from rlvec;
quit;
proc univariate data= RL2of2_Precedence_normal;
var delta000 delta025 delta050 delta075 delta100 delta125 delta150 delta175
    delta200 delta225 delta250 delta275 delta300;
run;
```

4.5.2.3 The *2-of-3* precedence chart

```
proc iml;
m = 500; n = 5; j = (n+1)/2;
sim = 100000;
a = 72; b = 429;
rlvec = j(sim,13,.);
do delta = 1 to 3 by 0.25;
shift = 4*delta-3;
do k = 1 to sim;
xref = j(m,1,0);
call randgen(xref,'NORMAL'); yref = xref;
call sort(yref, {1});
lcl = yref[a,1];
ucl = yref[b,1];
count = 2;
signal = 0;
dummy1 = {1,0}; dummy2 = j(2,1,.); check = {1,0};
between_above_above = {1 0 0,
                       0 0 0};
between_below_below = {1 1 1,
                        0 1 1};
below_between_below = {1 1 1,
                       1 0 1};
above_between_above = {0 1 0,
                       0 0 0};
matrix = j(2,3,.);
do while (signal = 0);
dummy2 = dummy1;
dummy1 = check;
xfut = j(n,1,0);
call randgen(xfut,'NORMAL',delta-1,1); yfut = xfut;
call sort(yfut, {1});
plotstat = yfut[j,1];
cl=j(2,1,0);
cl[1,1]=ucl;
cl[2,1]=lcl;
plotstatvec=j(2,1,plotstat);
check = plotstatvec <= cl;
matrix = dummy2||dummy1||check;
if matrix = between_above_above then signal = 1;
else if matrix = between_below_below then signal = 1;
else if matrix = below_between_below then signal = 1;
else if matrix = above_between_above then signal = 1;
else count = count + 1;
count1 = count;
rlvec[k,shift] = count1;
end;
end;
end;
create RL2of3_Precedence_normal from rlvec[colname={delta000 delta025 delta050
delta075 delta100 delta125 delta150 delta175 delta200 delta225 delta250 delta275
delta300}];
append from rlvec;
quit;
proc univariate data= RL2of3_Precedence_normal;
var delta000 delta025 delta050 delta075 delta100 delta125 delta150 delta175
    delta200 delta225 delta250 delta275 delta300;
run;
```

Chapter 5

Concluding remarks: Summary and recommendations for future research

To finish-off this thesis, we give here a brief summary of the research conducted in the thesis and offer concluding remarks concerning unanswered questions and/or future research opportunities.

In this thesis, in general, we focused on a variety of aspects related to the basic (yet powerful) statistical tool often used in quality improvement efforts within the realm of statistical quality control, that is, the Shewhart-type of control chart. First, we looked at Shewhart-type Phase I variables charts for the variance, the standard deviation and the range; this was followed by an overview of the literature on Shewhart-type Phase I variables charts for the location and the spread of a process. Second, we studied the Shewhart-type Phase II *p*-chart and the Shewhart-type Phase II *c*-chart in Case U (when the parameters are unknown) and assessed the influence when the parameters are estimated from a Phase I sample on the performance of these charts; both these charts are attributes charts and are widely used in practice. Lastly, we developed a new class of nonparametric Shewhart-type Phase I and Phase II control charts, for monitoring or controlling a certain quantile of the underlying probability distribution of a process, based on runs-type signaling rules using the well-known sign test and the two-sample median test statistic as plotting statistics. In the next few paragraphs we point out the highlights of the research carried out in this thesis and state some research ideas to be pursued in the near future. We also list the research outputs related to this thesis; this includes a list of technical reports and peer-reviewed articles that were published in international journals, contributions to local and international conferences where the author of this thesis presented papers and some draft articles that were submitted and are currently under review.

Variables control charts

Assuming that the underlying process distribution follows a normal distribution with an unknown mean and an unknown variance, in Chapter 2 we studied the design of the well-known Shewhart-type

S^2, S and R charts for Phase I applications based on the availability of m independent rational subgroups each of size $n > 1$. We showed that, because multiple plotting statistics are simultaneously compared to the same set of estimated control limits, the signaling events (i.e. the event when a plotting statistic plots outside the control limits) are mutually dependent. We further argued (with reference to the article of Champ and Jones, (2004)) that the correct design criterion of Shewhart-type Phase I charts is the false alarm probability (FAP), which is the probability of at least one false alarm, and not the false alarm rate (FAR), which deals with only one plotting statistic at a time and is defined as the probability of a signal at any particular sampling stage. Accordingly, we found the appropriate charting constants for a variety of (m, n)-combinations for each of the three charts (using intensive computer simulation experiments) so that the FAP of each chart does not exceed 0.01, 0.05 and 0.10, respectively.

The literature overview, in Chapter 2, regarding univariate parametric Phase I Shewhart-type charts for the location and the spread of a process not only presented the current state of the art of constructing these charts, but also brought several important points under our attention:

(i) There is a lack of proper guidance to the practitioner on the correct statistical design and implementation of Phase I charts. In a search of the standard statistical process control textbooks on the market, none to very little material was found, including the standard book of Montgomery (2005), who discusses the topic without the necessary statistical theory.

(ii) Some of the authors that studied the Phase I problem (especially when the process parameters are estimated) ignore the dependency between the Phase I plotting statistics and incorrectly used the FAR (which only deals with a single plotting statistic at a time) to design the charts as apposed to the FAP (which takes into account that multiple charting statistics are simultaneously compared to the estimated control limits). This would certainly deteriorate the performance of these charts. Our methodology provides the correct control limits for the applications studied.

(iii) There seems to be no consensus on exactly how one should compare the performance of competing Phase I charts. This boils down to the question of how to formulate and define an out-of-control situation in Phase I. One current proposal is to adopt the scenario that one of the Phase I samples is out-of-control and that the remaining $m-1$ samples are in-control and then (via computer simulation) compare the empirical

probability that at least one point plots outside the estimated control limits. The chart with the highest empirical probability of detecting the out-of-control sample is then declared the winner; this can be investigated further.

(iv) There is a genuine need to develop a Phase I control chart for the case when $n = 1$, that is, for individuals data. Admittedly there are some articles available in the literature that address the problem (see e.g. Nelson (1982), Roes, Does and Schurink (1993), Rigdon, Cruthis and Champ (1994) and Bryce, Gaudard and Joiner (1997)) but the problem has not yet been solved satisfactorily. The main stumbling block appears to be finding a suitable point estimator for the variance or the standard deviation and deriving the exact joint distribution of the standardized plotting statistics. Since individuals data is so common nowadays in many industries, this problem is important and will be studied using methods similar to the ones in this thesis.

(v) The design of Phase I control charts for correlated data needs to be looked at. A good starting point is the articles by Boyles (2000) and Maragah and Woodall (1992).

(vi) Except for the study by Borror and Champ (2001), there is apparently no other published work regarding the design of Phase I Shewhart-type attributes charts. This is an important aspect because the study of the Phase II *p*-chart and the Phase II *c*-chart is based on the availability of an in-control reference sample, which is usually obtained at the end of a successful Phase I study.

(vii) It would definitely be helpful and beneficial to the quality practitioner if a unified approach to the design and implementation of Phase I variables and attributes charts is available; this is a topic currently under investigation by the author of this thesis and his supervisors.

Attributes control charts

The Phase II Shewhart-type *p*-chart and *c*-chart were studied in detail in Chapter 3. The aim was to determine the effect of estimating the unknown process parameters from a Phase I reference sample on the performance of the charts in their Phase II application. The methodology that we used was based on the two-step procedure which was introduced in the statistical process control arena by Chakraborti (2000). The procedure entails that we first condition on a particular observed value of the point

estimator from Phase I (in order to obtain the conditional Phase II run-length distribution and the associated characteristics of the conditional run-length distribution) and then calculate the unconditional Phase II run-length distribution and the associated properties of the unconditional run-length distribution by averaging over all the values of the point estimator. We numerically investigated the various properties of the conditional and the unconditional run-length distributions, for the in-control and the out-of-control scenarios, and compared the results to the benchmark values of Case K (i.e. when the parameters are known). It was found that the widely-followed guidelines regarding the number of Phase I rational subgroups, m, and the sample size, n, is not adequate to control the average run-length and/or the false alarm rate at acceptable levels. The cause of the discrepancy between the attained ARL and the attained FAR values and the industry standards of 370 and 0.0027 (respectively) is twofold. The discrepancy is partly due to the fact that the underlying process distributions are discrete and to some extent it is caused by the fact that the standard formula, i.e. mean $\pm$ 3 $\times$ standard deviation, for calculating the control limits, is not 100% correct; this is so because the normal approximation to the binomial and the Poisson distributions is not very good for all values of the parameters p and c (especially p close to 0 or 1 and c close to 0).

The question of how we can correctly design the Phase II Shewhart-type *p*-chart and *c*-chart remains, in some way, unanswered. As pointed out in an earlier section, the formulae for the characteristics of the unconditional run-length distribution can be helpful in this regard and there are two possible routes to follow:

(i) The usual approach is to specify a certain attribute of the unconditional Phase II run-length distribution (such as the unconditional average run-length, which is common in routine applications (see e.g. Chakraborti, (2006))) and then solve for the charting constant(s).

Even though this approach is viable, it would only be successful insofar it is possible to accurately specify the unknown parameters p and c. The reason for this drawback is the fact that the unconditional properties of the charts are unconditional only with respect to the point estimators and not with respect to the unknown parameters.

(ii) A second approach one can pursue is to also uncondition on the properties of the charts with respect to the parameters p and c. This approach, which is closely linked to a

Bayes approach, entails that we treat p and c as random variables and that we choose appropriate (prior) distributions to model the uncertainty in the parameters.

As suggested earlier, the standard beta distribution (with support on the interval (0,1)) and the gamma distribution (with the positive real numbers as support) would work. However, the dilemma in this approach is that we still require expert knowledge and guidance when choosing the parameters of the beta and the gamma distributions.

Currently, the topic of finding and comparing suitable charting constants for the application of the Shewhart-type Phase II *p*-chart and *c*-chart is underway by the author of this thesis and co-workers.

Nonparametric Shewhart-type control charts with runs-type signaling rules

Lastly, in Chapter 4 we designed new nonparametric control charts based on runs-type signaling rules using the well-known sign test statistic and the two-sample median test statistic as plotting statistics. The sign test was used in the design of the charts when the percentile under investigation of the underlying process distribution was known (or specified) whereas the two-sample median test was used to construct the charts when the percentile was unknown. The main advantages of the nonparametric charts are:

(i) The fact that the underlying distribution needs not be specified (as we only require continuity of the distribution);

(ii) The precise numerical measurements need not be available (because we only count the number of observations greater or smaller than a specified value or simply rank the observations within each sample). Neither the counting nor the ranking procedure requires exact measurements;

(iii) The sign charts have the added advantage that they can be applied in scenarios where the data are just dichotomous (e.g. yes/no); and

(iv) The precedence charts give us the flexibility to apply the chart in situations where the data is naturally collected in an ordered fashion (e.g. time to failure).

We derived the run-length distributions of this new class of distribution-free control charts using a Markov chain approach and, where possible, we also used the results related to the geometric

distribution of order k. Where necessary we again used the two-step conditioning and unconditioning idea by Chakraborti (2000) to obtain the Phase II run-length distributions.

Having derived the run-length distributions and the associated characteristics of the new charts, extensive tables were provided with the suitable charting constants for each chart which should help the practitioner in the setting up of the charts. A numerical example was also given to illustrate the implementation and operation of the charts. However, having pointed out the benefits of the new runs rules enhanced charts, there are two important aspects concerning nonparametric control charts (in general) that are worth mentioning:

(i) There is a major shortcoming regarding the application of the nonparametric charts in industry because there is a lack of a proper understanding (and perhaps an appreciation) of the topic nonparametrics and consequently the important role these charts can play in practice.

The main reason for this limitation seems to be that distribution-free (nonparametric) methods are typically only touched on in undergraduate statistics courses in most programs and are not necessarily taught at a post-graduate level and, in most cases, not even taught to the engineers and/or the operator personnel who have to deal with the monitoring of the processes. What is more, is the fact that none of the available (standard) textbooks on statistical process control covers the topic of nonparametric control charting procedures in any detail.

(ii) It would be a great improvement and definitely to the advantage of the quality practitioner if software developers were to include the nonparametric control charts that are already available, as standard options or procedures in their statistical computer packages. Currently, these nonparametric control chart procedures are not available for practitioners and they simply resort to the standard parametric control chart methodologies.

Research outputs

A number of research outputs related to and based on this thesis have seen the light. Below we provide a list with the details of the technical reports and the peer-reviewed articles that were published, the articles that were accepted for publication, the local and the international conferences where papers were presented and the draft articles that were submitted and currently under review.

Published articles

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

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

(iii) Chakraborti, S., Human, S.W., Graham, M.A. (2008). "Phase I statistical process control charts: An overview and some results". *Quality Engineering*, 21(1):52-62; and

(iv) Chakraborti, S., Eryilmaz, S., Human, S. W. (2009). "A Phase II nonparametric control chart based on precedence statistics with runs-type signaling rules". *Computational Statistics and Data Analysis*, 53(1):1054-1065.

Articles accepted for publication

(i) Human, S. W., Chakraborti, S., Smit, C. F. "Nonparametric Shewhart-type sign control charts based on runs". *Communications in Statistics – Theory and Methods.*

Articles under review

(ii) Human, S. W., Chakraborti, S., Smit, C. F. "Control charts for variation in Phase I applications", Submitted to *Computational Statistics and Data Analysis*.

Technical reports

(i) Human, S. W., Chakraborti, S., Smit, C. F. (2009). "Shewhart-type S^2, S and R control charts for Phase I applications". Technical Report 09/01, Department of Statistics, University of Pretoria, ISBN: 978-1-86854-735-7.

(ii) Human, S. W., Chakraborti, S., Smit, C. F. (2009). "Nonparametric Shewhart-type control charts with runs-type signaling rules". Technical Report 09/02, Department of Statistics, University of Pretoria, ISBN: 978-1-86854-738-8.

International conference

(i) The 7th World Congress in Probability and Statistics in Singapore jointly sponsored by the Bernoulli Society and the Institute of Mathematical Statistics (2008) where the results related to the nonparametric control charts of Chapter 4 was presented.

Local conferences

(i) The annual conference of the South African Statistical Association (SASA) hosted by the Department of Statistics of the Rhodes University in Grahamstown (2005) where the results related to the Phase II *p*-chart of Chapter 3 was presented;

(ii) The annual conference of the South African Statistical Association (SASA) hosted by the Department Statistics and Actuarial Science of the University of Stellenbosch (2006) where the results related to the Phase II *c*-chart of Chapter 3 was presented;

(iii) The annual conference of the South African Statistical Association (SASA) hosted by the Department of Statistics and Actuarial Science of the University of Witwatersrand (2007) where the results related to the Phase I S^2, S and R control charts of Chapter 2 was presented; and

(iv) The annual conference of the South African Statistical Association (SASA) hosted by the Department of Statistics of the University of Pretoria (2008) where the results related to the nonparametric control charts of Chapter 4 was presented.

The end.