ORDENAÇÃO ESTOCÁSTICA NA ANÁLISE DE DESEMPENHO DE ESQUEMAS DE CONTROLO DE QUALIDADE

Nome: Manuel João Cabral Morais
Doutoramento em: Matemática
Orientador: Professor Doutor António Manuel Pacheco Pires
Provas concluídas em:

Resumo

A detecção de desvios do estado de controlo estatístico pode ser efectuada através do registo e da observação dos dados naquilo que usualmente se designa de esquema (ou carta) de controlo.

Subjacente à análise de um esquema, está aquela que é, indiscutivelmente, a mais popular de todas as medidas de desempenho, o "run length" (*RL*) ou o número de amostras recolhidas até à emissão de um sinal. O conhecimento da respectiva distribuição desempenha um papel preponderante na avaliação da capacidade de um esquema de controlo para assegurar a qualidade de um processo.

Propõe-se o uso da ordenação estocástica com o objectivo de avaliar — de um modo qualitativo e mais objectivo — a forma como a capacidade de detecção de um esquema se altera face a modificações nos parâmetros de planeamento e intrínsecos ao modelo dos dados, ou esta se compara com a de um outro esquema de controlo. É dada especial atenção ao *RL* de esquemas de controlo do tipo *CUSUM* e *EWMA*, sendo o *RL* encarado como um tempo de primeira passagem associado a uma cadeia de Markov absorvente, logo com distribuição "phase-type".

Ao tirar partido do facto de, em certas condições, estas cadeias de Markov serem regidas por matrizes de probabilidades de transição estocasticamente monótonas (Cap. 2), foi possível estabelecer resultados gerais de monotonia estocástica para o *RL* no que diz respeito ao valor inicial da estatística sumária e outros parâmetros (Cap. 3). Estes resultados são aplicados ao *RL* de esquemas combinados do tipo *CUSUM–Shewhart* para dados *binomiais* (Cap. 4) e ao de esquemas para μ na presença de alterações em σ (Cap. 5). A isto segue-se a introdução (e estudo do comportamento monótono) de duas medidas de desempenho referentes ao fenómeno dos sinais erróneos ("misleading signals"): a probabilidade de um sinal erróneo e o número de amostras recolhidas até à emissão de um sinal erróneo. Estas medidas, a acrescentar ao *RL*, são cruciais na avaliação de esquemas de controlo simultâneo de μ e de σ (Cap. 6). Avalia-se também a influência em termos estocásticos dos parâmetros autoregressivos do modelo dos dados no *RL* de esquemas residuais para o valor esperado de dados autocorrelacionados (Cap. 7).

Finalmente, concentramo-nos na investigação em curso, nas contribuições mais importantes desta tese e nalgumas recomendações para trabalho futuro.

Palavras chave: Controlo de Qualidade, Abordagem Markoviana, Tempos de Primeira Passagem, Distribuições "Phase-type", Ordenação Estocástica, Matrizes Estocasticamente Monótonas.

STOCHASTIC ORDERING IN THE PERFORMANCE ANALYSIS OF QUALITY CONTROL SCHEMES

Abstract

Departures from a state of statistical control can be detected by plotting and viewing the process data in what is usually called a control scheme (or chart).

Underlying the scheme analysis, there is an indisputedly popular performance measure, the run length (RL). The knowledge of its distribution plays a major role in helping understand the ability of the control scheme to monitor process quality.

The use of stochastic ordering is proposed to assess — in a qualitative and more objective way — how the scheme monitoring ability is changed by modifications in design or model parameters, or how its performance compares with the one of another control scheme. Special attention is given to the RL of control schemes, such as *CUSUM* and *EWMA*, which is viewed as a first passage time associated to an absorbing Markov chain, thus, with a phase-type distribution.

By capitalizing on the fact that under certain conditions those Markov chains are governed by stochastically monotone transition matrices (Chap. 2), we were able to establish general stochastic monotonicity results concerning the RL in terms of the initial value of the scheme summary statistic and other parameters (Chap. 3). These results are applied to the RL of combined *CUSUM–Shewhart* schemes for binomial data (Chap. 4) and schemes for μ in the presence of shifts in σ (Chap. 5). This is followed by the introduction (and study of the monotone behaviour) of two performance measures referring to the phenomenon of misleading signals: the probability of misleading signal and the run length to a misleading signal, which add up to the RL and are crucial in the joint monitoring of μ and σ (Chap. 6). The influence of the autoregressive parameters of the data model in the RL of residual schemes for the mean of autocorrelated data is also evaluated (Chap. 7).

Finally, we focus on our current research, the major contributions of this thesis and on a few recommendations for future work.

Keywords: Quality Control, Markovian Approach, First Passage Times, Phase-type Distributions, Stochastic Ordering, Stochastically Monotone Matrices.

Acknowledgements

I would like to express my sincere thanks to my supervisor, Professor António Pacheco, for his invaluable help, crucial guidance and pertinent suggestions throughout this project.

Special thanks should be given to Professor Wolfgang Schmid for the helpful discussions and feedback while he was visiting Centro de Matemática Aplicada.

My thanks are also extensive to:

- The anonymous referees of the joint papers written along this project for their feedback and thoughtful comments which lead to improved versions of the original drafts;
- Professors Amílcar and Cristina Sernadas, for the incentive and for reminding me that I should conclude this thesis in time to include it in the final report of PROBLOG;
- Professor Ana Pires, for having encouraged me to submit a paper for publication in an international journal back in 1997, and for the computer facilities;
- Professor F. Miguel Dionísio, for a crucial hint in functional programming used in the numerical evaluation of the percentage points of the *RLMS*s;
- Professor João Branco, Leader of our research group in Centro de Matemática Aplicada, for all his support;
- Professor José Natário, for a fruitful pre-lunch discussion that led to the proof of result (5.52);
- Professor M. Fernanda Ramalhoto, for introducing me to stochastic ordering and quality control;
- Professor Sven Knoth, for the productive coffee break discussions during the VIIth International Workshop on Intelligent Statistical Quality Control;
- The ever inspiring Aretha Franklin ("Queen of not only Soul Music"), Nusrat Fateh Ali Khan ("The Brightest Star in Qawwali — Shahen-Shah-e-Qawwali") and Ravi Shankar ("Master of the Sitar and India's Crown Jewel in the Coffers of World Music"), for their intoxicating music; and
- All my colleagues from "Piso 5", in particular, to my "sisters-in-crime", Rosário Oliveira, Cláudia Nunes, Conceição Amado and Isabel Rodrigues for "being there", and to Carlos Caleiro and Jaime Ramos for the "x-files" and for putting up with all my babbling.

This work could not have been undertaken without the support and care of my parents, Carol and João, my sister Carmen and my boyfriend Nuno. I sincerely hope your infinite patience and kindness is well rewarded.

This thesis was written with the partial support of Centro de Matemática Aplicada and grants Praxis PCEX/P/MAT/41/96 (Analysis and Computer Implementation of Selected Methods in Applied Statistics) and Praxis PCEX/P/MAT/10002/98 (PROBabilistic methods in LOGic of reactive systems). I must gratefully acknowledge the financial support they provided for acquiring new computer facilities and for attending several national and international conferences, a summer school and several seminars given by visiting professors, during the preparation of this thesis.

Contents

List of Tables

List of Figures

Basic Notation

All vectors and matrices are in boldface type. Vectors are additionally underlined, and unless specified otherwise, all vectors are column vectors. "Increasing" and "decreasing" are used in the nonstrict sense, i.e., they mean "nondecreasing" and "nonincreasing", respectively.

Moreover, the following notation is used throughout this thesis:

Spaces

C	decision region
$\mathbb{N} = \{1, 2, \ldots\}$	set of all positive integers
$\mathbb{N}_0 = \{0, 1, 2, \ldots\}$	set of all nonnegative integers
$\mathbb{R}$	real line
$\mathbb{R}_0^+$	set of all nonnegative real numbers
$\mathbb{R}^{x+1}$	Euclidean coordinate space of all $(x+1)-$ dimensional vectors
Ξ	parameter space

General

$\underline{\mathbf{1}}$ ($\underline{\mathbf{0}}$)	vector of ones (zeroes); in most cases a $(x+1)-$dimensional vector
$\underline{\mathbf{a}}$	a vector; usually $(x+1)-$dimensional
$\underline{\mathbf{a}}^\top$	transpose of $\underline{\mathbf{a}}$
$\mathbf{A} = [a_{ij}]_{i,j=k}^{l}$	matrix with entries $a_{ij}, i, j = k, k+1, \ldots, l$
$bin(n, p)$	binomial distribution with parameters n and p
χ_k^2	chi-square distribution with k degrees of freedom
$\underline{\mathbf{e}}_u$	$(u+1)^{th}$ vector of the orthonormal basis for $\mathbb{R}^{x+1}$
$geo(p)$	geometric distribution with parameter p
$\mathbf{I}$	identity matrix; with rank $x+1$ in most cases
$I_A(x)$	indicator function of set A; it is equal to 1, if x belongs to the set A, and equal to 0, otherwise
i.i.d.	independent and identically distributed
μ	mean (of normal data)

n	sample size
N	sample number
$N(\mu, \sigma^2)$	univariate normal distribution with expected value μ and variance σ^2
$\mathbf{O}$	matrix with all entries equal to zero; in general a $(x+2) \times (x+2)$ matrix
σ	standard deviation (of normal data)
$\lfloor x \rfloor$	integer part of a real number x

Functions

The following functions refer to the random variable X:

d_X	relative decrease of the probability function
F_X	distribution function
$\overline{F}_X$	survival function
λ_X	hazard rate function
$\overline{\lambda}_X$	reversed hazard rate function
P_X	probability function
r_X	equilibrium rate function

Stochastic orders

ev	expected value sense
hr	hazard rate sense
lr	likelihood ratio sense
M	majorization sense
rh	reversed hazard rate sense
st	usual sense

For $* = ev, hr, lr, hr, M, rh, st$, we define:

$\leq_*$	stochastically smaller in the $*-$sense
$\downarrow_*$	stochastically decreasing in the $*-$sense
$\uparrow_*$	stochastically increasing in the $*-$sense
$\mathcal{M}_*$	class of stochastically monotone matrices in the $*-$ sense

Acronyms

ARL	average run length
CKRL	coefficient of kurtosis of the run length
CSRL	coefficient of skewness of the run length
CUSUM	cumulative sum (scheme)
CVRL	coefficient of variation of the run length
EWMA	exponentially weighted moving average (scheme)
FPT	first passage time
FSI	fixed sampling intervals
LCL	lower control limit
MS	misleading signal
OQC	on-line quality control
PMS	probability of a misleading signal
RL	run length
RLMS	run length to a misleading signal
SDRL	standard deviation of the run length
SO	stochastic ordering
SPC	statistical process control
UCL	upper control limit
VSI	variable sampling intervals

The acronyms of several control schemes can be found in Appendix A and throughout the dissertation.

The symbol • denotes the end of an Example, a Proof or a Remark.

Each chapter is divided into sections, with consecutive labelling of Corollaries, Definitions, Equations, Examples, Lemmas, Propositions, Remarks and Theorems within each chapter. Please note that results which are of general interest are given as Propositions, and those results which only concern the run length are usually stated as Theorems, Lemmas or Corolaries.

Chapter 1

Introduction

1.1 Background

Manufacturing processes are typically monitored using data: a random sample of n items is usually taken on a regular basis and the observed values of a summary statistic are sequentially plotted together with appropriate control limits. The resulting graphical device is grandly termed as quality control scheme (or chart). It is used to track process performance over time and identify the presence of assignable (or special) causes that may affect the quality of the output, in order to deploy corrective actions to bring the process back to target as quickly as possible.

This statistical tool was developed by Walter A. Shewhart of the Bell Telephone Laboratories in 1924 (Montgomery (1985, p. 13)). It has gained widespread acceptance in industry, and is among the most important and widely used devices in statistics (Stoumbos *et al.* (2000)). In fact, its use is not confined to manufacturing processes. Some of the current applications of quality control schemes include

- administration (Hawkins and Olwell (1998, p. v) — document misfilling),
- clinical chemestry (Westgard *et al.* (1977) — detection of systematic changes in an analytical process),
- epidemiology (Blacksell *et al.* (1994) — veterinary disease diagnosis),
- fraud detection (Johnson (1984) — systematic stealing by till understriking),
- health care (Hawkins and Olwell (1998, pp. v-vi) — control of the time taken for a blood sample to be turned around),
- safety (Lucas (1985) — monitoring of accident data),
- staff management (Olwell (1997) — misconduct monitoring),
- water monitoring (Gibbons (1999) — detection programs at waste disposal facilities),

and also athletics, biology, environmental science, genetics, finance and law enforcement (see Stoumbos *et al.* (2000)).

The best-known control schemes are those pioneered by Walter A. Shewhart (1931) and rightly called *Shewhart* schemes: the $\bar{X}$ and R charts for monitoring possible changes in the process mean and variance, respectively. The simplicity and the one-size-fits-all character of these and other *Shewhart* schemes, such as the S^2 chart, are responsible for their sheer popularity between practitioners. However, the fact that the *Shewhart* schemes only use the information about the process given by the last observed value of their summary statistics, and completely ignore any other information contained in the previous observed samples, is responsible for a serious limitation: they are not effective in the detection of assignable causes that lead to small and moderate shifts in the parameter being monitored.

The cumulative sum (or *CUSUM*) and the exponentially weighted moving average (or *EWMA*) schemes for the process mean — introduced by Page (1954) and Roberts (1959), respectively — incorporate all the information in the sequence of observed values of a simple summary statistic (such as the sample mean) and prove to be more effective than *Shewhart* schemes for detecting small and moderate shifts in the process mean.[1] Although the *CUSUM* and *EWMA* schemes date from the 50's, the usage of these efficient schemes was very infrequent for many years. Nowadays their use is still relatively low — but slowly and steadily increasing — when compared to the *Shewhart* schemes (Stoumbos *et al.* (2000)).

The process monitoring problem can be briefly described as follows. Let X denote a quality characteristic and $F_\varrho(x)$ its distribution function. This function is indexed by ϱ, a vector of at least one parameter. The production process is stable or in-control when the (quality) parameter ϱ is equal to some target value ϱ_0. The quality control operators and engineers are alerted to the possible presence of assignable causes that affect the value of ϱ by a signal, given when the value of the summary statistic of one of the schemes described above is beyond its control limits.

The control limits are set in such a way that the summary statistic is very unlikely to fall outside them when $\varrho = \varrho_0$. This choice is due to the fact that a control scheme should trigger no signal for a period as long as possible if the process is in-control, contributing to a reduction of the frequency of false alarms. In opposition, a control scheme should give a signal with minimum delay if the process is out-of-control.

Thus, it comes as no surprise that the performance of any quality control scheme is usually assessed in terms of characteristics of the run length (RL) — the number of samples taken before a signal is triggered by the scheme —, assuming that the quality parameter has remained constantly at ϱ. The average run length (ARL) is by far the most popular of those characteristics and has been — throughly and incorrectly — used to describe the likely performance of a control scheme.[2]

[1] Similar results are reported for the combined *CUSUM–Shewhart* and *EWMA* schemes for the variance by Yashchin (1985) and Crowder and Hamilton (1992), respectively.

[2] According to Lai (1974), it was Aroian and Levene (1950) who proposed the average efficiency number, now more commonly known as the *average run length*, in the assessment of the performance of quality control schemes.

1.2 Aim and methodology

In the (on-line) quality control literature, the description and comparison of the performance of control schemes is often tackled *numerically*. Tables and graphs, usually referring to the ARL, are provided to make the descriptions and comparisons possible, but analytical proofs are rarely presented.

This thesis mainly focus on providing *analytical* answers to the following question:

- *What* is the impact of a change in a design or model parameter in the properties of a control scheme?

By establishing stochastic ordering results, we are able to

- *Tell how* the control scheme performance is affected when a design or model parameter changes, and

- *Determine which* one of two or more competing schemes has the best performance in a specific sense,

without the need to compute the performance of the control schemes themselves.

Stochastic order relations between the performance measures of control schemes provide a qualitative basis for a more effective comparison of the control schemes we are dealing with than the comparisons based entirely on numerical results. They can induce monotonicity properties that are important theoretically and in practice because they lead to various insights about the performance of the schemes.

Next we describe some stochastic order relations between the run length measures of two control schemes 1 and 2:[3] RL_1 and RL_2, respectively.

- *Stochastically smaller in the expected value sense* ($\leq_{ev}$)

$$RL_1 \leq_{ev} RL_2 \Leftrightarrow ARL_1 \leq ARL_2, \tag{1.1}$$

  where ARL_1 and ARL_2 are the average run lengths of schemes 1 and 2, respectively. Since $1/ARL$ may be regarded as the average detection speed, $RL_1 \leq_{ev} RL_2$ means that scheme 1 is in average faster than scheme 2 in giving signals (i.e., signals towards the detection of an assignable cause, or false alarms).

  However, the ARL provides an unidimensional and possibly misleading snapshot of the scheme performance, specially if the summary statistic has a Markovian structure (as in the case of *EWMA* and *CUSUM* schemes) because the out-of-control RL distribution may deviate considerably from the geometric distribution (Luceño and Puig-Pey (2000)). For this reason, some authors have argued that the percentage points of the RL provide a second and more appropriate performance measure when

[3] It should be pointed out that schemes 1 and 2 can correspond to the same control scheme with different design or model parameters — e.g. two *CUSUM* control schemes that only differ in the reference value or an *EWMA* scheme in the in-control and out-of-control situations.

Shewhart control schemes are not used (see, e.g., Yashchin (1985)). As a result, stronger stochastic order relations, such as the ones described below, should be established.

- *Stochastically smaller in the usual stochastic sense* ($\leq_{st}$)

$$RL_1 \leq_{st} RL_2 \Leftrightarrow P(RL_1 > m) \leq P(RL_2 > m),\ m \in I\!R. \tag{1.2}$$

$RL_1 \leq_{st} RL_2$ means that scheme 1 signals within the first m samples more frequently than scheme 2, for any value of m. Thus, the probability that we obtain a signal at the first sample is greater in scheme 1 than in scheme 2.

Since $RL_1 \leq_{st} RL_2 \Rightarrow P(RL_1 \leq RL_2) \geq 1/2$, it follows that if $RL_1 \leq_{st} RL_2$ then scheme 1 signals before scheme 2 most of the time. In addition, $RL_1 \leq_{st} RL_2 \Rightarrow RL_1 \leq_{ev} RL_2$, and, furthermore, $RL_1 \leq_{st} RL_2 \Leftrightarrow E[f(RL_1)] \leq E[f(RL_2)]$ for all increasing (cost) functions for which the expectations exist (see Shaked and Shanthikumar (1994, p. 4)). Therefore, the stochastic order relation $RL_1 \leq_{st} RL_2$ provides more information than $RL_1 \leq_{ev} RL_2$, and can be thought as providing a bidimensional snapshot of the performance comparison.

- *Stochastically smaller in the hazard rate sense* ($\leq_{hr}$)

$$RL_1 \leq_{hr} RL_2 \Leftrightarrow \lambda_{RL_1}(m) \geq \lambda_{RL_2}(m),\ m \in I\!N, \tag{1.3}$$

where $\lambda_{RL}(m) = P(RL = m)/P(RL \geq m)$ represents the hazard rate function of the RL and was proposed by Margavio *et al.* (1995) as the alarm rate at sample m.

If $RL_1 \leq_{hr} RL_2$, we can state that, for any value of m, scheme 1 is more likely to signal at sample m than scheme 2, given that the previous $m-1$ samples were not responsible for triggering a signal in any of the two schemes. Moreover, following Marvagio *et al.* (1995), we can add that scheme 1 has a larger alarm rate than scheme 2, regardless of the value of the sample number.

The stochastic comparison based on hazard rate functions is far more meaningful than the one based on survival functions because: $RL_1 \leq_{hr} RL_2 \Rightarrow RL_1 \leq_{st} RL_2$ and the hazard rate comparison provides a performance confrontation in a specific conditional setting, thus, it gives a conditional snapshot of the performance comparison.

- *Stochastically smaller in the reversed hazard rate sense* ($\leq_{rh}$)

$$RL_1 \leq_{rh} RL_2 \Leftrightarrow \overline{\lambda}_{RL_1}(m) \leq \overline{\lambda}_{RL_2}(m),\ m \in I\!N, \tag{1.4}$$

where $\overline{\lambda}_{RL}(m) = P(RL = m)/P(RL \leq m)$ denotes the reversed hazard rate function (see Kijima (1997, p. 113) or Shanthikumar, Yamazaki and Sakasegawa (1991)) of the RL.

In case the inequality above holds, we can interpret $\leq_{rh}$ similarly to $\leq_{hr}$: for any value of m, scheme 1 is less likely to signal at sample m than scheme 2, given that both control schemes triggered at least one signal up to sample m.

- *Stochastically smaller in the likelihood ratio sense* ($\leq_{lr}$)

$$RL_1 \leq_{lr} RL_2 \Leftrightarrow P(RL_1 = m)/P(RL_2 = m) \downarrow_m \text{ over the set } I\!N. \tag{1.5}$$

  Thus, the odds of scheme 1 signalling by the m^{th} sample against scheme 2 triggering a signal at the same sample decreases as m increases.

  Note that $RL_1 \leq_{lr} RL_2 \Rightarrow RL_1 \leq_{hr} RL_2$ and $RL_1 \leq_{lr} RL_2 \Leftrightarrow (RL_1|\{a \leq RL_1 \leq b\}) \leq_{st} (RL_2|\{a \leq RL_2 \leq b\})$ whenever $a \leq b$ (Shaked and Shanthikumar (1994, p. 29)). Therefore, establishing the order relation $RL_1 \leq_{lr} RL_2$ provides a comparison between the detection speed of schemes 1 and 2 in a stricter conditional setting than the one imposed by $RL_1 \leq_{hr} RL_2$.

  Furthermore,

$$RL_1 \leq_{lr} RL_2 \Leftrightarrow r_{RL_1}(m) \geq r_{RL_2}(m),\ m \in I\!N, \tag{1.6}$$

  where r_{RL} denotes the equilibrium rate (see Shaked and Shanthikumar (1994, p. 436)) of RL; i.e., $r_{RL}(1) = 0$ and $r_{RL}(m) = P(RL = m-1)/P(RL = m), m = 2, 3, \ldots$.

  It is worth adding that if we denote by $d_{RL}(m)$ the relative decrease in the probability that the m^{th} sample triggers a signal relative to the probability of the signal being given by the previous sample — that is,

$$d_{RL}(m) = \frac{P(RL = m-1) - P(RL = m)}{P(RL = m-1)} = 1 - \frac{1}{r_{RL}(m)},\ m = 2, 3, \ldots \tag{1.7}$$

  — then we have:

$$RL_1 \leq_{lr} RL_2 \Leftrightarrow d_{RL_1}(m) \geq d_{RL_2}(m),\ m = 2, 3, \ldots. \tag{1.8}$$

  Thus, $RL_1 \leq_{lr} RL_2$ allows us to assert that the relative sequential decrease in the probability of emission of a signal is larger in scheme 1 than in scheme 2.

Clearly (see Theorems 1.C.1 and 1.B.1 of Shaked and Shanthikumar (1994, pp. 28 and 14)),

$$RL_1 \leq_{lr} RL_2 \Rightarrow \left\{ \begin{array}{l} RL_1 \leq_{hr} RL_2 \\ RL_1 \leq_{rh} RL_2 \end{array} \right. \Rightarrow RL_1 \leq_{st} RL_2 \Rightarrow RL_1 \leq_{ev} RL_2. \tag{1.9}$$

Establishing stochastic order relations involving two RLs is rather trivial when we are dealing with *Shewhart* schemes for a single parameter, based on independent data from an univariate quality control characteristic. This follows by virtue of the fact that in this case the RL has a geometric distribution, as reported in Section 2.1.

When *CUSUM* or *EWMA* control schemes are at use, the RL is (or can be approximated by) a first passage time of a Markov chain. In this case, stochastic ordering results do not follow in a straightforward manner, although the Markov approach provides a convenient way of obtaining detailed properties of the *CUSUM* and *EWMA* control schemes, as described in Section 2.3, including the exact (or the approximate) average detection speed and other moments of RL, and the RL survival, hazard rate and probability functions.

1.3 A brief literature review

The desire to confront what is random is probably as old as probability itself. Consider, for instance, the problem of comparing two games in order to decide which one is more favorable to a gambler. For this reason, Bawa (1982) traces the origins of stochastic dominance in the work of J. Bernoulli published in 1713, *Ars Conjectardi*.[4]

A landmark in the history of stochastic ordering is the pioneering work by Lorenz (1905) in the assessment of income inequality in a population of n individuals. Lorenz — feeling that all of the summary measures then under consideration constituted too much condensation of the data (Arnold (1987, p. 2)) — proposed what is now known as the Lorenz curve. Given the incomes of n individuals, $x_1, \ldots, x_n$, the Lorenz curve unites the points $(k/n, S_k/S_n),\ k = 0, \ldots, n$ with $S_0 = 0$ and $S_k = \sum_{i=1}^{k} x_{(i)}$ where $x_{(i)}$ denotes the i^{th} smallest income, so that S_k represents the total income of the poorest k individuals. If the total income is equally distributed among the n individuals the Lorenz curve is nothing but a 45^o line.

Lorenz suggested the following rule of interpretation: a high level of income inequality is associated to a severely bent curve. This rule was subsequently clarified by Dalton (1920) and mathematically formulated by Hardy, Littlewood and Pólya (1929): the vector $\underline{\mathbf{x}} = (x_1, \ldots, x_n)$ represents a less unequal income distribution than $\underline{\mathbf{y}} = (y_1, \ldots, y_n)$, in notation $\underline{\mathbf{x}} \leq_M \underline{\mathbf{y}}$, if and only if

$$\sum_{i=1}^{k} x_{(i)} \geq \sum_{i=1}^{k} y_{(i)},\ k = 1, \ldots, n-1 \tag{1.10}$$

and $\sum_{i=1}^{n} x_{(i)} = \sum_{i=1}^{n} y_{(i)}$ (see Marshall and Olkin (1979, p.5)).

For a comprehensive historical perspective of stochastic ordering see Mosler and Scarsini (1993, p. 1–5). Extensive listings (but by no means exhaustive) of references on stochastic ordering, stochastic dominance, and their applications can be found in Bawa (1982), Levy (1992) and Mosler and Scarsini (1993).

Shaked and Shanthikumar (1994) is another extremely important reference in stochastic ordering. The first part of this book (corresponding to six chapters) is an exposition of many results on a large number of stochastic orders. The second part (ten chapters) consists of material written by leading researchers in various fields in which stochastic ordering is applied — including Statistical Inference, Risk Theory, Biology, Scheduling, Queueing Theory, and Reliability Theory.

Papers or other material devoted simultaneously to on-line quality control (OQC) and stochastic ordering (SO) are not as numerous as we would expect, although the advantages of the combination of these two fields would seem to follow straightforwardly.[5] Here are some examples and a brief reference to the relevant results of each paper to provide a setting for the contributions of this thesis.

[4] For the definition of stochastic dominance of degree k see Arnold (1987, p. 81–87). Note that the stochastic dominance of degree $k = 1$ corresponds to the stochastic ordering in the usual sense.

[5] There seems to exist even fewer references combining off-line quality control (or acceptance sampling) and stochastic ordering. Kirmani and Peddada (1993) and Yao and Zheng (1995, 1999) are some of the exceptions.

- *Lai (1974).* As far as we have investigated, this is the oldest reference relating OQC and SO, and provides an inequality which can be rephrased as a stochastic ordering result involving two RLs.

  Lemma 2 of Lai (1974) provides an upper bound for the survival function of the RL of an upper one-sided moving average chart for the mean of *i.i.d.* data: under certain conditions involving the distribution of the data and the weights of the moving average, there exist $\alpha \in (0,1)$ and $\lambda > 0$ such that $P(RL > n) \leq \lambda\alpha^n$ for all n. This relation can be rewritten as $RL \leq_{st} N$ where $P(N = 1) = 1 - \lambda\alpha$ and $P(N = n) = \lambda(1-\alpha)\alpha^{n-1},\ n = 2, 3, \ldots$, with $0 < \alpha < 1$ and $0 < \lambda \leq 1/\alpha$.[6]

- *Ghosh, Reynolds Jr. and Hui (1981).* These authors consider an upper one-sided $\bar{X}$ scheme for the case where the process variance, σ^2, is unknown and the upper control limit depends on an estimator of σ^2, S_0^2, the variance of an auxiliary and independent random sample. Since the same S_0 is permanently used in the monitoring process, the distribution of the RL is no longer geometric but can be stochastically related to a geometric random variable:

  $$geo(p(\delta)) \leq_{st} RL(\delta), \tag{1.11}$$

  where $\delta = \sqrt{n}(\mu - \mu_0)/\sigma$ and $p(\delta) = P(\bar{X}_i \geq \mu_0 + kS_0/\sqrt{n} \mid \mu = \mu_0 + \delta\sigma/\sqrt{n})$ denotes the probability that the procedure signals at sample i, with μ_0 being the target value for the mean μ and k a positive constant.

- *Jensen (1984, p. 32).* This reference devotes its Theorem 8 to the effects of a reduction in variability on the run length of a Hotelling T^2 scheme for monitoring the mean vector of a multivariate normal distribution $N_p(\underline{\mu}, \Sigma)$ where the nominal value of $\underline{\mu}$ is known and equal to $\underline{\mu}_0$ and the covariance matrix Σ is also known (see, e.g. Woodall and Ncube (1985) or Runger and Prabhu (1996)). The result presented in Theorem 8 is an application of the order preservation under Loewner ordering $\leq_L$ (Loewner (1934)) of positive semidefinite Hermitian $p \times p$ matrices, and states that if efforts are made to tighten the process variability in such a way that $\Sigma_2 \leq_L \Sigma_1$ then the associated run lengths verify $RL_2 \leq RL_1$ for all $\underline{\mu} \in I\!R^p$.

- *Jensen and Hui (1990a,b).* Theorem 1 of Jensen and Hui (1990a) reports a similar result to the one derived by Ghosh, Reynolds Jr. and Van Hui (1981) for the non-standard R and S^2 schemes, whose decision rule depends on a prior estimator of the unknown variance σ^2. Hence, these schemes tend to signal less frequently (in the usual sense) than standard R and S^2 schemes that assume a known σ_0^2 and share the false alarm rate of the former schemes, whether the process variance is or not on-target.

[6] Böhm and Hackl (1990) derive a lower bound for the ARL that does not depend on normality, and is much sharper than the one obtained by Lai (1974) in many situations. This bound and the remaining ones in this reference cannot be easily stated as an useful stochastic relation order (in the expected value sense) between two RLs.

Jensen and Hui (1990b) consider a more general situation comprising a non-sustained shift in the process variance. Let $\{\sigma_0, \sigma_1, \sigma_2, \ldots\}$ be the scale parameters corresponding to the initial and subsequent samples, where $\sigma_i \geq \sigma_0, i = 1, 2, \ldots$. Then, according to Theorem 1 of this reference, the run lengths of non-standard schemes R and S^2 stochastically decrease in the usual sense with $\underline{\gamma} = \{\sigma_1/\sigma_0, \sigma_2/\sigma_0, \ldots\}$ under the componentwise ordering of vectors.[7]

Now, we proceed to describe the work of W. Schmid and his collaborators on control schemes for correlated data. Among other merits, these references prove that, under mild conditions, the in-control *RL*s of several modified schemes[8] for the mean of a correlated process are stochastically larger than the corresponding *RL*s in the case of independent variables.

- *Schmid (1995).* The author proves in this paper that the result mentioned above holds for a two-sided modified $\bar{X}$ scheme for the mean, μ, of rather general stochastic processes $\{Y_N\}$, like the ones with elliptically contoured marginal distributions (see Tong (1990, p. 62–63)) or positively lower orthant-dependent random variables (Tong (1990, p. 93)). Moreover, if all marginal distributions of $Y_N - \mu_0$ (where μ_0 denotes the known target value of μ) have a continuous, unimodal and symmetric (about the origin) density, Schmid (1995) proves that the RL stochastically decreases in the usual sense with the absolute value of the magnitude of the shift in μ.

- *Schmid (1997a,b).* These two references provide similar results to Schmid (1995) for the modified schemes of types *EWMA* and *CUSUM* for the mean of stationary Gaussian processes with particular autocovariance structure.

- *Schmid and Schöne (1997).* The main result of this reference can be translated in stochastic ordering terms as:

$$RL_{iid} \leq_{st} RL_G, \tag{1.12}$$

  where RL_G (RL_{iid}) represents the in-control RL of an upper one-sided modified *EWMA* scheme for the mean of a stationary Gaussian process with autocovariance function $\{\gamma_\nu\}$ such that $\gamma_0 > 0$ and $\gamma_\nu \geq 0,\ \nu \in I\!N$ (an upper one-sided *EWMA* scheme for the mean of Gaussian *i.i.d.* data), with initial value equal to μ_0 and upper control limit dependent on the exact variance of the *EWMA* summary statistic.

- *Schöne, Schmid and Knoth (1999).* This paper presents a corollary which is an extension of (1.12). Consider two stationary Gaussian processes with autocovariance functions $\{\gamma_\nu, \nu \in I\!N_0\}$ and $\{\delta_\nu, \nu \in I\!N_0\}$, such that: $\gamma_0, \delta_0 > 0,\ \gamma_\nu, \delta_\nu \geq 0,\ \nu = 1, \ldots, k$ and $\gamma_\nu \times \delta_\xi \geq \gamma_\xi \times \delta_\nu$ for $0 \leq \xi < \nu \leq k-1, k \in I\!N$. Let RL_γ and RL_δ be the in-control *RL*s of the associated upper one-sided *EWMA* charts for the process mean, defined as in Schmid and Schöne (1997). Then $P(RL_\gamma > k) \geq P(RL_\delta > k)$.

[7] $\underline{\gamma} = (\gamma_1, \gamma_2, \ldots) \leq \underline{\gamma}^* = (\gamma_1^*, \gamma_2^*, \ldots)$ if and only if $\gamma_i \leq \gamma_i^*, i = 1, 2, \ldots$.
[8] This sort of scheme for autocorrelated data is described in Chapter 7.

If the autocovariances monotonicity condition is valid for every nonnegative integer k this corollary can be phrased as a stochastic ordering result: $RL_\gamma \geq_{st} RL_\delta$.

The proof of (1.12) involves the use of Slepian's inequality (Theorem 5.1.7 of Tong (1990, p. 103)) for the multivariate normal distribution. Furthermore, the theorems of Schmid and Schöne (1997) and Schöne, Schmid and Knoth (1999) still hold for stationary processes having elliptically contoured marginal distributions because of a generalization of Slepian's inequality for these distributions (Theorem 4.3.6 of Tong (1980, p. 74)).

- *Kramer and Schmid (2000).* Theorems 1 and 2 cast some light in the comparison of the in-control *RL*s, in the stationary and *i.i.d.* situations, of modified and residual Shewhart schemes that make use of an initial estimate of the process variance.

The author of this thesis is the co-author of some papers that are concerned with the control of *i.i.d.* data and relate SO and OQC. We conclude this section with a brief reference to the main results in some of those references.

- *Morais and Natário (1998).* Let ATS_{FSI} and ATS_{VSI} be the average times to signal of the upper one-sided *c*-scheme with fixed sampling intervals (*FSI*), for controlling the expected value of Poisson data, and of the corresponding matched in-control upper one-sided *c*-scheme with variable sampling intervals (*VSI*).[9] Using the fact that the χ^2_ν random variable stochastically increases in the likelihood sense with ν, the authors show that the *VSI* feature improves the average detection time of any increase in the expected value of magnitude θ, i.e.,

$$ATS_{VSI}(\theta) \leq ATS_{FSI}(\theta), \theta > 0. \tag{1.13}$$

  Morais (1995) and Ramalhoto and Morais (1995, 1999) obtain similar results for upper and lower one-sided *FSI* and *VSI Shewhart* schemes for the scale parameter of data with three-parameter Weibull distribution, using the increasing failure rate character of the summary statistic which has a χ^2_ν distribution.

- *Morais and Pacheco (1998b).* This paper provides sufficient conditions and a proof by induction of two intuitive stochastic results concerning the *RL* of upper one-sided *EWMA* and *CUSUM* schemes, using a representation of such control schemes as Markov chains and the special features of their associated transition probability matrices. In particular, they show that:

  - the run length stochastically decreases (in the usual sense) with the shift in the parameter being controlled, implying that these schemes become progressively more sensitive to the shift as it grows in magnitude; and
  - if the process is out-of-control when the control scheme is started (or is restarted following an ineffective corrective action), the *RL* is stochastically reduced (also

[9]For more details on *VSI* schemes refer to Reynolds Jr. *et al.* (1988), Reynolds Jr., Amin and Arnold (1990) or the introduction of Chapter 4.

in the usual sense) when a head start value is used. If the process is in-control, there is also a stochastic reduction of the RL when a head start value is given to the chart, producing short RLs more frequently.

Alternative proofs for both properties can be also found in Chapter 3.

- *Morais and Pacheco (2001b).* By using the stochastic monotonicity properties of the RLs of the individual upper one-sided *EWMA* schemes for μ and σ (see Chapter 5) of normally distributed data, this paper establishes monotonicity properties for the probability of misleading signal (*PMS*) and stochastic monotonicity properties for the run length to a misleading signal (*RLMS*) — which are both performance measures of joint schemes for μ and σ, introduced by Morais and Pacheco (2000a) and Morais and Pacheco (2001b), respectively. The extension of all these properties for other joint schemes for μ and σ can be found in Chapter 6.

1.4 Organization and summary of contributions

We now summarize the content, organization and what we believe are some of the main contributions of this thesis. Here is what lies ahead in the next seven chapters and Appendix A. (The number of every chapter is put between parentheses.)

(2) *First passage times and RL* — RL is viewed as a first passage time and we single out two possible distributions: the geometric, obtained when basic Shewhart schemes are at use; and the discrete phase-type distribution, for some Markov-type schemes.

Associating stochastically monotone matrices — in the usual, the hazard rate, the reversed hazard rate and the likelihood ratio senses — to control schemes whose RL has a discrete phase-type distribution is a vital contribution of this chapter. Some ageing properties of the RL are also reviewed.

(3) *Monotonicity and RL* — We take advantage of the stochastic monotonicities of the transition matrices to obtain stochastic monotonicity results concerning the behaviour of the RL of Markov-type schemes.

The first lot of results concerns the decreasing monotonicity of the RL in terms of the initial value of the summary statistic, in the usual, hazard rate and likelihood ratio senses. A monotonicity result in the usual sense is also obtained for the RL in terms of any other parameter; the problems that arise when we try to strengthen this result are also discussed.

In addition, we explore the validity of all these results for the exact RL of schemes such as *CUSUM* and *EWMA* for continuous data.

(4) *Combined CUSUM–Shewhart schemes for binomial data* — Upper one-sided combined *CUSUM–Shewhart* schemes for binomial data are discussed in this chapter; we stress an analogy between the RL of these schemes and the first passage time introduced by Li and Shaked (1995, 1997).

This type of scheme illustrates some of the stochastic properties alluded in the previous chapter. It also draws our attention to the fact that the associated transition matrix is not stochastically monotone in the hazard rate and likelihood ratio senses, and that the *RL* does not decrease with the head start in the likelihood ratio sense.

We assess the stochastic impact of supplementing an upper one-sided *CUSUM* scheme with a *Shewhart* upper control limit. An extended example numerically illustrates the magnitude of this impact, and numerical comparisons between upper one-sided combined *CUSUM–Shewhart* schemes and upper one-sided *CUSUM* schemes with a 50% head start are carried out too.

(5) *Upper one-sided schemes for μ in the presence of shifts in σ* — This chapter explores some monotonicity properties referring to the stochastic influence of changes in the mean and variance of a normally distributed quality characteristic. For instance, we prove that under certain conditions the discriminating effect of the upper one-sided schemes for μ of *Shewhart*, *CUSUM*, combined *CUSUM–Shewhart*, *EWMA* and combined *EWMA–Shewhart* types stochastically decreases with the magnitude of the shift in σ.

We conclude the chapter with an investigation of the decreasing monotonicity of the *RL* in terms of the initial value of the summary statistic of all four Markov-type upper one-sided schemes mentioned earlier.

(6) *Misleading signals in joint schemes for μ and σ* — We introduce two performance measures in this chapter, adding up to the *RL*. They refer to the phenomenon of misleading signals — which can possibly send the user of the joint scheme for the mean and variance to try to diagnose and correct a nonexistent assignable cause — and are the probability of a misleading signal (*PMS*) and the run length to a misleading signal (*RLMS*).

Monotonicity properties of *PMS*s and stochastic ordering results involving *RLMS*s are proved. To complete the chapter we present some striking examples showing that the ocurrence of misleading signals is surely a cause of concern in practice and, thus, *PMS* and *RLMS* should be used in the assessment of the performance of joint schemes for the process mean and standard deviation.

(7) *Assessing the impact of autocorrelation in the performance of residual schemes for μ* — This is the only chapter of this thesis addressing control schemes involving autocorrelated data.

Sufficient conditions are obtained to guarantee that the *RL* of standard Shewhart residual schemes for the mean of a stationary Gaussian $AR(1)$ ($AR(2)$) model stochastically increases with the (first) autoregressive parameter in the hazard rate sense. A similar monotonicity result in the usual sense also holds for the *RL* of upper one-sided *CUSUM* and *EWMA* residual schemes for the mean of stationary Gaussian $AR(1)$ processes.

(8) *Concluding remarks* — The first part of the chapter focus on our current research on the comparison between phase-type and geometric *RL*s, and on our joint work

with W. Schmid concerning the extension of some of the stochastic ordering results in Chapter 5 to modified *EWMA* schemes for autocorrelated data.

Finally, the major contributions of this thesis are briefly reviewed and recommendations for further work are outlined.

Appendix A — This appendix comprises the summary statistics, control limits and *RL* distributions of several individual control schemes for μ and σ for a normally distributed quality characteristic. A few monotonicity properties of the individual schemes for σ are stated and proved in the last section of this appendix. This appendix provides a working basis for Chapters 5 and 6.

We end this chapter with the following remarks. A detailed or brief review generally preceeds the main results of each chapter to provide a setting for our contributions. Several programs for the package *Mathematica 4.0* (Wolfram (1996)) have been written to produce all the graphs and tables in this thesis; these programs will be made available to those who are interested and request them from the author.

Chapter 2

First passage times and run lengths

First passage times (*FPT*s) arise naturally in level-crossing problems in:

- *Reliability theory* — *FPT*s of appropriate stochastic processes often represent the time to failure of a device subjected to shocks (and wear), which fails when its damage level crosses a threshold or the maximal increment of the damage level exceeds a critical value (Li and Shaked (1995, 1997));

- *Queueing systems* — The identity of the first customer whose waiting time exceeds a critical threshold is a *FPT* (Greenberg (1997)). This is an important performance measure of single-server (multiserver) queues with impatient customers, in paralel with the number of lost customers and the number of sucessful departures over given intervals as discussed by Bhattacharya and Ephremides (1991);

and certainly in

- *Quality control* — For the basic control schemes (i.e., with no supplementary run rules) the out-of-control signal is given as soon as the summary statistic falls outside the control limits. So it follows that the *RL* is the first passage time

  $$RL = \min\{N : Z_N \notin C\}, \tag{2.1}$$

  where Z_N and C represent the summary statistic at the sampling period N and the decision region, respectively.

For further examples of *FPT*s see, e.g., Wasan (1994) and Greenberg (1997). Wasan (1994) covers detailed techniques to find first passage probabilities and *FPT* distributions for several stochastic processes. Greenberg (1997) reviews approximate computational methods, using Markov chains, to determine expected *FPT*s in *CUSUM* schemes, and extends them to more general stochastic processes.

Let the underlying distribution function of the quality characteristic X be denoted by

$$F_\varrho(x),\ \ \varrho \in \Xi, \tag{2.2}$$

where Ξ is the set of possible values of ϱ, and

$$\theta = g(\varrho, \varrho_0) \tag{2.3}$$

be the parameter which relates the quality level ϱ to its target value ϱ_0. If

$$\Theta = \{g(\varrho, \varrho_0) : \varrho \in \Xi\} \tag{2.4}$$

represents the set of possible values of θ (Θ is usually a convex set), it can be added that the production process is operating in-control as long as $\theta = \theta_0$, where

$$\theta_0 = g(\varrho_0, \varrho_0), \tag{2.5}$$

and it will be out-of-control as soon as an assignable cause is responsible for a change to a value $\theta \neq \theta_0$, so that

$$\theta = g(\varrho, \varrho_0), \text{ for some } \varrho \in \Xi \setminus \{\varrho_0\}. \tag{2.6}$$

As described above, the basic control charting amounts to a sequential yes/no decision procedure based on the successive observed values of the summary statistic and the decision region. Therefore, the process monitoring procedure can be viewed as the repeated testing of hypotheses

$$H_0 : \theta = \theta_0 \text{ (in control)} \quad \text{vs.} \quad H_1 : \theta \in \Theta \setminus \{\theta_0\} \text{ (out-of-control)}, \tag{2.7}$$

if we assume that an assignable cause results in a sustained shift in the parameter of interest ϱ.[1]

The monitoring is performed by: updating the value of the summary statistic calculated from the incoming data each time a new sample is collected; and triggering a signal as soon as the value of the summary statistic is beyond the decision region. This signal is called: a valid alarm, in case the process is out-of-control; and a false alarm, otherwise. The scheme parameters should be suitably selected in order to provide as few false alarms as possible and, simultaneously, swift detection when the process is operating out-of-control.

The sensitivity of the control scheme is usually measured by the RL — the number of samples taken up to and including the one which is responsible for the signal. The RL depends upon the form of the density of the observations and upon the values of scheme parameters such as the decision region or the initial value of the summary statistic (for example, when *EWMA* or *CUSUM* schemes are at use). Nevertheless, we will usually suppress the dependence on the parameters whenever no ambiguities can arise, except for θ. Henceforth, the run length is denoted by $RL(\theta)$. Moreover, we will not suppress in general the notational dependence on θ of the N^{th} random sample $\underline{X}_N = (X_{1N}, \ldots, X_{nN})$ or the summary statistic Z_N, from now on denoted $\underline{X}_N(\theta) = (X_{1N}(\theta), \ldots, X_{nN}(\theta))$ and $Z_N(\theta)$, respectively.

[1] This resemblance "better reflects statistical thinking in showing ties between two important areas of statistics" and "provides a formal basis for evaluating properties of control charts", as put by Woodall (2000). However, there are some disagreements regarding the similarity between control charting and repeated hypothesis testing; for further details please refer to the third section of Woodall (2000).

In the next section, we discuss and give examples of geometric RLs, the simplest RLs, in the setting described above.

In order to provide RL related mesures we need some preparatory (and rather trivial) results. Let us consider the probability generating function (PG), and the integer factorial moments (FM) and central moments (CM) of an arbitrary positive integer random variable X, respectively:

$$PG_X(z) = E\left(z^X\right),\ 0 \leq z \leq 1; \tag{2.8}$$

$$FM_X(s) = E\left[X^{(s)}\right] = E[X(X-1)\dots(X-s+1)],\ s \in I\!N; \tag{2.9}$$

$$CM_X(s) = E\left[X^{[s]}\right] = E[\{X - E(X)\}^s],\ s \in I\!N. \tag{2.10}$$

Recall that $E(X) = E\left[X^{(1)}\right]$, and that the standard deviation (SD) and the coefficients of variation (CV), skewness (CS) and kurtosis (CK) of X can be written in terms of its central moments as follows:

$$SD(X) = \sqrt{E\left[X^{[2]}\right]} \tag{2.11}$$

$$CV(X) = \frac{SD(X)}{E(X)} \tag{2.12}$$

$$CS(X) = \frac{E\left[X^{[3]}\right]}{[SD(X)]^3} \tag{2.13}$$

$$CK(X) = \frac{E\left[X^{[4]}\right]}{[SD(X)]^4} - 3. \tag{2.14}$$

Also note that the central moments of order 2, 3 and 4 in the expressions above are the following functions of the factorial moments:

$$\begin{aligned} E\left[X^{[2]}\right] = {} & E\left[X^{(2)}\right] \\ & + E\left[X^{(1)}\right] \times \left\{1 - E\left[X^{(1)}\right]\right\} \end{aligned} \tag{2.15}$$

$$\begin{aligned} E\left[X^{[3]}\right] = {} & E\left[X^{(3)}\right] \\ & + E\left[X^{(2)}\right]\left\{3 - 3E\left[X^{(1)}\right]\right\} \\ & + E\left[X^{(1)}\right]\left(1 - 3E\left[X^{(1)}\right] + 2\left\{E\left[X^{(1)}\right]\right\}^2\right) \end{aligned} \tag{2.16}$$

$$\begin{aligned} E\left[X^{[4]}\right] = {} & E\left[X^{(4)}\right] \\ & + E\left[X^{(3)}\right]\left\{6 - 4E\left[X^{(1)}\right]\right\} \\ & + E\left[X^{(2)}\right]\left(7 - 12E\left[X^{(1)}\right] + 6\left\{E\left[X^{(1)}\right]\right\}^2\right) \\ & + E\left[X^{(1)}\right]\left(1 - 4E\left[X^{(1)}\right] + 6\left\{E\left[X^{(1)}\right]\right\}^2 - 3\left\{E\left[X^{(1)}\right]\right\}^3\right). \end{aligned} \tag{2.17}$$

2.1 Geometric RLs

In case the control scheme operates as a sequence of *independent* hypotheses tests (as in the basic *Shewhart* schemes) and the quality parameter remains constant at ϱ, the summary statistics $Z_N(\theta), N \in I\!N$, are also *i.i.d.* to some statistic $Z(\theta)$, and $RL(\theta)$ has a very simple probability function:

$$P_{RL(\theta)}(m) = [\pi(\theta)]^{m-1}[1 - \pi(\theta)], \; m \in I\!N, \tag{2.18}$$

where

$$1 - \pi(\theta) = P[Z(\theta) \notin C]. \tag{2.19}$$

That is, $RL(\theta)$ is a geometric random variable with parameter $1 - \pi(\theta)$.

The parameter $1 - \pi(\theta)$ denotes the probability that, given θ, a signal is triggered by the scheme each time a new sample is collected. It can be thought as the power function of the hypothesis test (2.7) which is sequentially repeated after drawing each sample. In addition, in the quality control literature,

$$\pi(\theta), \; \theta \in \Theta, \tag{2.20}$$

represents the well known operating-characteristic function.

A summary of twelve $RL(\theta)$ related measures — namely: the probability, survival, hazard rate and equilibrium rate functions, $p \times 100\%$ percentage point, probability generating function; and the integer factorial moment of order s, expected value, standard deviation, and coefficients of variation, skewness and kurtosis — can be found in Table 2.1, where $\lfloor x \rfloor$ denotes the integer part of the real number x. Also note that the $p \times 100\%$ percentage point can be written in terms of the generalized inverse function of F_{RL}, as defined in Szekli (1995, p. 3), $F_{RL}^{-1}(p) = \inf\{m \in I\!R : F_{RL}(m) \geq p\}$.

In what follows we examine the stochastic monotone behaviour of geometric RLs with regard to the parameter $1 - \pi(\theta)$.

Lemma 2.1 — *Let* $RL(\theta)$ *and* $RL(\theta')$ *be two* RL*s with geometric distribution with parameters* $1 - \pi(\theta)$ *and* $1 - \pi(\theta')$. *If* $1 - \pi(\theta) \geq 1 - \pi(\theta')$ *then* $RL(\theta) \leq_{lr} RL(\theta')$.

Proof — When $1 - \pi(\theta) \geq 1 - \pi(\theta')$, the likelihood ratio $\frac{P[RL(\theta)=m]}{P[RL(\theta')=m]}$ equals $\left[\frac{\pi(\theta)}{\pi(\theta')}\right]^{m-1} \times \frac{1-\pi(\theta)}{1-\pi(\theta')}$ and decreases with m over the set of positive integers. Thus, $RL(\theta) \leq_{lr} RL(\theta')$.

We can provide an alternative proof: according to Table 2.1, we can assert that $1 - \pi(\theta) \geq 1 - \pi(\theta') \Rightarrow r_{RL(\theta)}(m) \geq r_{RL(\theta')}(m), \; m = 2, 3, \ldots \Leftrightarrow RL(\theta) \leq_{lr} RL(\theta')$. •

Remark 2.2 — The stochastic monotone behaviour of geometric RLs is completely defined by the parameter $1-\pi(\theta)$. In fact $RL(\theta)$ stochastically decreases as $1-\pi(\theta)$ increases — in the likelihood ration sense, thus, in the hazard rate, the usual and the expected value senses.[2]

[2]And, since $FM_{RL(\theta)}(s)$ also decreases with $1 - \pi(\theta)$, $RL(\theta)$ stochastically decreases in the factorial moments order sense (see Shaked and Shanthikumar (1994, p. 100)). Note that this behaviour immediately follows from the decreasing monotonicity of $RL(\theta)$ with $1 - \pi(\theta)$ in the usual sense (see Theorem 3.B.11 in Shaked and Shanthikumar (1994, p. 102)).

Table 2.1: Run length related measures — geometric case.

Probability function	$P_{RL(\theta)}(m) = [\pi(\theta)]^{m-1} \times [1-\pi(\theta)],\ m \in I\!N$
Survival function	$\overline{F}_{RL(\theta)}(m) = \begin{cases} 1, & m < 1 \\ [\pi(\theta)]^{\lfloor m \rfloor}, & m \geq 1 \end{cases}$
Hazard Rate function	$\lambda_{RL(\theta)}(m) = 1 - \pi(\theta),\ m \in I\!N$
Equilibrium Rate function	$r_{RL(\theta)}(m) = [\pi(\theta)]^{-1},\ m = 2, 3, \ldots$
$p \times 100\%$ Percentage point	$F^{-1}_{RL(\theta)}(p) = \inf\{m \in I\!R : F_{RL(\theta)}(m) \geq p\},\ 0 < p < 1$
Probability Generating function	$PG_{RL(\theta)}(z) = z[1 - z\pi(\theta)]^{-1}[1-\pi(\theta)],\ 0 \leq z < [\pi(\theta)]^{-1}$
Factorial Moment of order s	$FM_{RL(\theta)}(s) = s! \times [\pi(\theta)]^{s-1}[1-\pi(\theta)]^{-s},\ s \in I\!N$
Expected Value	$ARL(\theta) = [1-\pi(\theta)]^{-1}$
Standard Deviation	$SD[RL(\theta)] = [\pi(\theta)]^{1/2}[1-\pi(\theta)]^{-1}$
Coefficient of Variation	$CV[RL(\theta)] = [\pi(\theta)]^{1/2}$
Coefficient of Skewness	$CS[RL(\theta)] = [1+\pi(\theta)][\pi(\theta)]^{-1/2}$
Coefficient of Kurtosis	$CK[RL(\theta)] = 4 + [\pi(\theta)]^{-1} + \pi(\theta)$

In addition, a close study of the remaining RL related measures in Table 2.1 allows us to assert that the standard deviation and coefficient of variation (coefficients of skewness and kurtosis) of $RL(\theta)$ decrease (increase) with $1 - \pi(\theta)$. •

For simplicity, let us assume that $\varrho \in \Xi$ represents a real parameter, whose increases (increases or decreases) are controlled by an upper one-sided scheme (a standard scheme) with decision region — now a decision interval — C. This assumption is made so that any technical details will not obfuscate the essence of the interpretation of Lemma 2.1 in Remark 2.3, and the illustrations in Example 2.4 and Example 2.5 of Section 2.2.

Remark 2.3 — Assume that the quality characteristic has a distribution belonging to an increasing monotone likelihood ratio family.[3] If we are dealing with an upper one-sided uniformly most powerful (UMP) test (Casella and Berger (1990, p. 365)) for ϱ and $\theta = g(\varrho, \varrho_0)$ increases with ϱ, for $\varrho \geq \varrho_0$, then $1 - \pi(\theta)$ increases with $\theta \geq \theta_0$. As a consequence

$$RL(\theta) \downarrow_{lr} \text{ with } \theta, \text{ for } \theta \geq \theta_0. \tag{2.21}$$

[3] A family of random variables $\{X(\theta), \theta \in [a,b]\}$ is said to be an increasing monotone likelihood ratio family if $X(\theta_1) \leq_{lr} X(\theta_2), a \leq \theta_1 \leq \theta_2 \leq b$, i.e. $X(\theta) \uparrow_{lr}$ with θ.

Analogously, if we are dealing with an uniformly most powerful unbiased (UMPU) test (Casella and Berger (1990, p. 374)) for ϱ then $1-\pi(\theta)$ decreases (increases) with θ for $\theta \leq \theta_0$ ($\theta \geq \theta_0$). Therefore

$$RL(\theta) \uparrow_{lr} \ (\downarrow_{lr}) \quad \text{with } \theta, \text{ for } \theta \leq \theta_0 \ (\theta \geq \theta_0). \tag{2.22}$$

Thus, in the first (second) situation, the larger the increase (increase or decrease) in ϱ, the smaller the number of samples taken until the detection of such a change in the likelihood ratio sense. These properties enhance what Ramalhoto and Morais (1999) called the "primordial criterion", which can be phrased as follows: the in-control ARL should be larger than any out-of-control ARL. •

2.2 Examples of geometric RLs

When we deal with nonconformities or defects such as surface flaws on a sheet metal panel, weaving irregularities in bolts of cloth or color inconsistencies on a painted surface, the quality assessment of a running production process is made using the so-called control charts for attribute data.

One of the main concerns, when dealing with attribute data, is to alert quality control operators and engineers of possible increases in the total expected number of

- defects (e.g., number of surface flaws on a panel) or
- defectives items[4] (e.g., number of cans which leak)

per random sample of constant size n. The underlying distributions of these totals are usually taken as *Poisson* and *binomial*, respectively, and arise from counting and categorization.

A comprehensive bibliography and review on control charting using attribute data can be found in Woodall (1997).

Example 2.4 — Suppose that in the final phase of Hi-Fi decks production, a deck is considered defective if it has more than two color inconsistencies in its front panel surface. Furthermore, the expected number of defective decks in each sample of $n = 100$ decks should not exceed two. That is, while operating in-control the quality characteristic has a $Bernoulli(p_0)$ distribution with $p_0 = 0.02$. The presence of an assignable cause yields an increase in the expected number of defectives per random sample — from np_0 to $n(p_0+\theta)$, where $0 < np_0 < n(p_0+\theta) < n$.[5] Although these defects do not affect the deck functioning, they are perceptible; thus, they can affect its price.

To control increases in the expected number of defective decks in a sample of n items, np, an upper one-sided np-scheme is adopted with the defectives count as summary statistic,

$$Z_N(\theta) = \sum_{i=1}^{n} X_{iN}(\theta), \tag{2.23}$$

[4]Number of items with at least one disqualifying defect.

[5]In this particular case $\Xi = [p_0, 1)$, $\theta = g(p, p_0) = p - p_0$ and $\Theta = [0, 1-p_0)$.

where $X_{iN}(\theta)$ is equal to 1, if the i^{th} deck of sample N is defective, and equal to 0, otherwise. The decision interval is given by

$$C = [LCL, UCL] = [0, \lfloor np_0 + \gamma\sqrt{np_0(1-p_0)}\rfloor] \tag{2.24}$$

where γ is a positive constant that should be chosen in such a way that the probability of a false alarm takes a small value. For example, if $\gamma = 5/\sqrt{1.96}$ then $UCL = 7$ and a false alarm occurs with probability

$$1 - \pi(0) = \overline{F}_{Z_N(0)}(UCL) = 1 - F_{bin(100,0.02)}(7) \simeq 0.000932. \tag{2.25}$$

Note that, since the underlying data distribution is Bernoulli$(0.02 + \theta)$, the RL of this control scheme has a geometric distribution with parameter

$$1 - \pi(\theta) = 1 - F_{bin(100,0.02+\theta)}(7), \tag{2.26}$$

regardless of the value of γ in the interval $[5/\sqrt{1.96}, 6/\sqrt{1.96})$.

For any positive integer n and $0 < p \leq p' < 1$, the likelihood ratio

$$\frac{P_{bin(n,p)}(m)}{P_{bin(n,p')}(m)} = \left[\frac{p(1-p')}{p'(1-p)}\right]^m \times \left(\frac{1-p}{1-p'}\right)^n \tag{2.27}$$

is a decreasing function of m over the set $\{0, 1, \ldots, n\}$. Thus, it follows that the random variable $bin(n,p)$ is stochastically smaller than $bin(n,p')$ in the likelihood ratio sense, and therefore

$$1 - F_{bin(n,p)}(x) \leq 1 - F_{bin(n,p')}(x), \quad -\infty < x < \infty. \tag{2.28}$$

For this reason, $1 - \pi(\theta)$ increases with θ and so, using Lemma 2.1,

$$RL(\theta) \downarrow_{lr} \text{ with } \theta. \tag{2.29}$$

Thus, the larger the increase in the expected number defective items per random sample, the smaller (in the likelihood ratio sense) the number of samples taken until the detection of such a change. This is an expected result and a direct consequence of dealing with a sequence of independent UMP tests for np. •

Many quality characteristics — such as weight, temperature, blood pressure, fuel consumption and insurance claims — are expressed in terms of numerical measurements. Schemes for those quality characteristics are usually called control schemes for variables. These schemes usually provide more information about the production process and are more effective than control schemes for attributes.

The next example (taken from Morais and Pacheco (1998a)) refers to the standard $\bar{X}$ scheme for the expected value and provides what we think is a useful complementary illustration of Lemma 2.1 and Remark 2.3.

Example 2.5 — Assume that it is possible to specify the target values of the process mean (μ) and variance (σ^2) of a normally distributed quality characteristic so that an analysis of past data is not required. Also assume that the process variance remains constant at the target level. Then the control limits of a standard $\bar{X}$ scheme are given by

$$LCL = \mu_0 - \gamma\sigma_0/\sqrt{n} \quad \text{and} \quad UCL = \mu_0 + \gamma\sigma_0/\sqrt{n}, \tag{2.30}$$

where: μ_0 and σ_0 are the target values of μ and σ, and the process mean relates to these target values as follows, $\mu = \mu_0 + \theta\sigma_0/\sqrt{n}$;[6] and γ is chosen in such way that the in-control ARL, $ARL(0)$, takes a fixed large value and using the fact that $\gamma = \Phi^{-1}(1-[2ARL(0)]^{-1})$.

Conditioned on the fact that the process mean equals $\mu = \mu_0 + \theta\sigma_0/\sqrt{n}$ $(-\infty < \theta < +\infty)$, the RL of this control scheme, $RL(\theta)$, has a geometric distribution with parameter

$$1 - \pi(\theta) = 1 - [\Phi(\gamma - \theta) - \Phi(-\gamma - \theta)]. \tag{2.31}$$

This is a continuous and even function of θ, and its first derivative is an odd function equal to

$$\frac{d[1 - \pi(\theta)]}{d\theta} = \sqrt{2/\pi} e^{-(\gamma^2+\theta^2)/2} \times \sinh(\gamma\theta), \tag{2.32}$$

which is nonpositive for $\theta \in (-\infty, 0]$ and nonnegative for $\theta \in [0, +\infty)$. Therefore we can assert that $1 - \pi(\theta)$ takes the minimum value at the origin and increases as $|\theta|$ increases. As a consequence of Lemma 2.1,

$$RL(\theta) \downarrow_{lr} \quad \text{with} \quad |\theta|, \tag{2.33}$$

i.e., the time to detect a change in μ with magnitude $\theta\sigma_0/\sqrt{n}$ stochastically decreases (in the likelihood ratio sense) as $|\theta|$ increases. Note that it is a well known fact that this control scheme operates as a sequence of independent UMPU hypotheses tests. •

In Chapter 3 we prove that stochastic monotonicity results such as (2.21) — which are intuitive and obviously imply

$$RL(\theta) \downarrow_{ev} \text{ with } \theta \tag{2.34}$$

— can be extended to one-sided control schemes whose summary statistics are governed by Markov chains, such as $EWMA$ and $CUSUM$ schemes. The stochastic implication of setting such control schemes to an initial head start value is also studied in Chapter 3, leading to other stochastic monotonicity properties.

We close this section with a final example. It illustrates that a well established scheme in practice can have undesired (stochastic) properties.

Example 2.6 — The use of a S^2 scheme is recommended to practitioners for controlling the variance σ^2 of normally distributed data. The summary statistic at the N^{th} sample, $(X_{1N}(\theta), \ldots, X_{nN}(\theta))$, is the sample variance $S^2_N(\theta) = (n-1)^{-1} \sum_{i=1}^{n} [X_{iN}(\theta) - \bar{X}_N(\theta)]^2$ and the control limits are

$$LCL = \frac{\sigma_0^2}{n-1} \times F^{-1}_{\chi^2_{n-1}}(\alpha/2) \tag{2.35}$$

[6] $\Xi = (-\infty, +\infty)$, $\theta = g(\mu, \mu_0) = \sqrt{n}(\mu - \mu_0)/\sigma_0$ and $\Theta = (-\infty, +\infty)$.

$$UCL = \frac{\sigma_0^2}{n-1} \times F^{-1}_{\chi^2_{n-1}}(1-\alpha/2), \tag{2.36}$$

where σ_0^2 is the target value, related to σ^2 by $\theta = \sigma/\sigma_0$,[7] and α is such that the in-control ARL of this scheme, $ARL(1)$, is equal to $1/\alpha$.

The RL of this scheme has a geometric distribution with parameter

$$1-\pi(\theta) = 1 - \left\{ F_{\chi^2_{n-1}}\left[\frac{F^{-1}_{\chi^2_{n-1}}(1-\alpha/2)}{\theta^2}\right] - F_{\chi^2_{n-1}}\left[\frac{F^{-1}_{\chi^2_{n-1}}(\alpha/2)}{\theta^2}\right]\right\}. \tag{2.37}$$

In order to establish stochastic order relations involving $RL(\theta)$ we have to study the behaviour of the function $1-\pi(\theta)$. The first derivative of $1-\pi(\theta)$ has the same sign as

$$F^{-1}_{\chi^2_{n-1}}(1-\alpha/2) \times f_{\chi^2_{n-1}}\left[\frac{F^{-1}_{\chi^2_{n-1}}(1-\alpha/2)}{\theta^2}\right] - F^{-1}_{\chi^2_{n-1}}(\alpha/2) \times f_{\chi^2_{n-1}}\left[\frac{F^{-1}_{\chi^2_{n-1}}(\alpha/2)}{\theta^2}\right] \tag{2.38}$$

which in turn has the same sign as

$$k(\theta) = \left[\frac{F^{-1}_{\chi^2_{n-1}}(1-\alpha/2)}{F^{-1}_{\chi^2_{n-1}}(\alpha/2)}\right]^{\frac{n-1}{2}} \times \exp\left\{-\frac{1}{2\theta^2}\left[F^{-1}_{\chi^2_{n-1}}(1-\alpha/2) - F^{-1}_{\chi^2_{n-1}}(\alpha/2)\right]\right\} - 1. \tag{2.39}$$

Since $k(\theta)$ is continuous and strictly increasing function of θ in $(0,+\infty)$, such that

$$\lim_{\theta\to 0^+} k(\theta) = -1 \quad \text{and} \quad \lim_{\theta\to+\infty} k(\theta) = \left[\frac{F^{-1}_{\chi^2_{n-1}}(1-\alpha/2)}{F^{-1}_{\chi^2_{n-1}}(\alpha/2)}\right]^{\frac{n-1}{2}} - 1 > 0, \tag{2.40}$$

the first derivative of $1-\pi(\theta)$ changes sign (from negative to positive) only once in $(0,+\infty)$. As a consequence, $1-\pi(\theta)$ takes its minimum value at the unique root of equation $k(\theta)=0$ in $(0,+\infty)$, $\theta^*(\alpha,n)$, given by

$$\theta^*(\alpha,n) = \sqrt{\frac{F^{-1}_{\chi^2_{n-1}}(1-\alpha/2) - F^{-1}_{\chi^2_{n-1}}(\alpha/2)}{(n-1)\left\{\ln\left[F^{-1}_{\chi^2_{n-1}}(1-\alpha/2)\right] - \ln\left[F^{-1}_{\chi^2_{n-1}}(\alpha/2)\right]\right\}}}. \tag{2.41}$$

Thus, we conclude that $1-\pi(\theta)$ decreases (increases) with θ for $\theta \le \theta^*(\alpha,n)$ ($\theta \ge \theta^*(\alpha,n)$).

This particular behaviour of $1-\pi(\theta)$ is illustrated for $\alpha = 0.02$ and $n = 5$ in Figure 2.1, and for other values of n in Table 2.2. This figure and this table also suggest that $0 < \theta^*(\alpha,n) < 1$, which holds in general, since

$$k(1) = \left[\frac{F^{-1}_{\chi^2_{n-1}}(1-\alpha/2)}{F^{-1}_{\chi^2_{n-1}}(\alpha/2)}\right]^{\frac{n-1}{2}} \times \exp\left\{-\frac{1}{2}\left[F^{-1}_{\chi^2_{n-1}}(1-\alpha/2) - F^{-1}_{\chi^2_{n-1}}(\alpha/2)\right]\right\} - 1 > 0. \tag{2.42}$$

Table 2.2: Values of $1-\pi(\theta)$ for S^2-schemes with $\sigma_0^2=1$ and $\alpha=0.002$ (i.e., $ARL(1)=500$).

	n					
θ	4	5	7	10	15	100
0.50	0.007828	0.014624	0.042134	0.132929	0.406761	1.000000
0.75	0.002359	0.003089	0.005036	0.009313	0.020672	0.762450
0.80	**0.001958**	0.002409	0.003528	0.005751	0.011016	0.419837
0.90	**0.001533**	**0.001652**	**0.001926**	0.002391	0.003274	0.037724
0.95	**0.001600**	**0.001628**	**0.001699**	**0.001819**	0.002035	0.006949
1.00	0.002000	0.002000	0.002000	0.002000	0.002000	0.002000
1.10	0.004522	0.004874	0.005553	0.006569	0.008323	0.054761
1.20	0.010808	0.012654	0.016447	0.022530	0.033848	0.373172

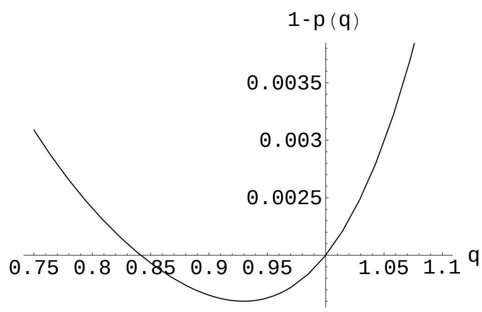

Figure 2.1: Values of $1-\pi(\theta)$ for a S^2-scheme with $\sigma_0^2=1$, $\alpha=0.002$ and $n=5$.

which implies that $0<\theta^*(\alpha,n)<1$, for all $n\geq 2$ and $0<\alpha<1$.

By virtue of the monotonous behaviour of $1-\pi(\theta)$, the stochastic order relations

$$RL(\theta)\downarrow_{lr} \text{ with } \theta, \text{ for } 0<\theta\leq\theta^*(\alpha,n)<1 \tag{2.43}$$

$$RL(\theta)\uparrow_{lr} \text{ with } \theta, \text{ for } \theta\geq\theta^*(\alpha,n) \tag{2.44}$$

hold, for $n\geq 2$ and $0<\alpha<1$. Hence the number of samples taken until a false alarm occurs can be stochastically smaller (in the likelihood ratio sense) than the one to detect certain small and moderate decreases in the process variance. For instance, the number of samples taken until the detection of a 5% decrease in σ ($\theta=0.95$) is stochastically larger (in the likelihood ratio sense) than the one until the ocurrence of a false alarm, for small to moderate sample sizes and $ARL(1)=500$. (For further examples, please refer to the entries in **bold** in Table 2.2.)

It is worth adding that this sort of performance behaviour is apparent in standard control schemes introduced by some authors, namely Kaminsky *et al.* (1992) — see, for instance, its Figure 4 which refers to the operating-characteristic curve of a $g-$scheme for geometric data. •

[7] $\Xi=(0,+\infty)$, $g(\sigma,\sigma_0)=\sigma/\sigma_0$ and $\Theta=(0,+\infty)$.

2.3 Discrete phase-type RLs

One way of increasing the sensitivity to permanent shifts in parameter(s) involves the accumulation of information across successive observations. The *CUSUM* scheme (Page (1954)) is an example and, undoubtedly, a very informative graphical device as it also provides simple graphical estimates of the time of occurrence of the shift and of its magnitude (Hawkins and Olwell (1998, p. 20)). The *EWMA* scheme (Roberts (1959)) is another important example.

The oldest *CUSUM* and *EWMA* schemes are the ones for detecting changes in the expected value of normal data. These schemes are also defined for the parameters of all standard distributions used in control charting, such as the expected value of the total of defective items in a sample (np), as illustrated in Example 2.7 ahead.

In addition, the *CUSUM* and *EWMA* schemes have *dependent* summary statistics and can be regarded as forming *Markov* chains in discrete time and discrete or continuous state space, leading to what is usually called the Markov approach.

This approach, originally proposed by Brook and Evans (1972), hardly fails to provide the exact (or approximate) *RL* distribution and any other *RL* related performance measures of that sort of scheme if taylored for discrete (or continuous) data.

Example 2.7 — Consider the same setting as in Example 2.4. However, suppose that the detection of increases in np is done by using an upper one-sided *CUSUM* scheme (without head start). Following Hawkins and Olwell (1998, pp. 122-123), the summary statistic of this scheme is given by

$$\mathrm{Z}_N(\theta) = \begin{cases} 0, \ N = 0 \\ \max\{0, Z_{N-1}(\theta) + Y_N(\theta) - k\}, \ N \in I\!N, \end{cases} \tag{2.45}$$

where $Y_N(\theta) = \sum_{i=1}^{n} X_{iN}(\theta)$ denotes the defectives count for the N^{th} sample and k, usually called reference value, is a positive constant, smaller than n for this scheme.

If k is a positive integer, then the summary statistic is governed by a discrete time Markov chain with infinite state space $I\!N_0$, null initial state, and transition matrix, dependent on the parameter θ,

$$\begin{bmatrix} F_\theta(k) & P_\theta(k+1) & P_\theta(k+2) & \cdots \\ F_\theta(k-1) & P_\theta(k) & P_\theta(k+1) & \cdots \\ F_\theta(k-2) & P_\theta(k-1) & P_\theta(k) & \cdots \\ F_\theta(k-3) & P_\theta(k-2) & P_\theta(k-1) & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}, \tag{2.46}$$

where $F_\theta(i) = F_{bin(100,0.02+\theta)}(i)$ and $P_\theta(i) = P_{bin(100,0.02+\theta)}(i)$ represent the distribution and probability functions of $Y_N(\theta)$ for any nonnegative integer i.

Now assume that a signal is triggered as soon as the summary statistic exceeds the upper control limit $UCL = x$, where x is a positive integer.[8] Thus, the run length of this upper one-sided *CUSUM* scheme is defined as the following *FPT*:

$$\min\{N : \ Z_N(\theta) > x \mid Z_0(\theta) = 0\}. \tag{2.47}$$

[8]The suitable choice of the reference value and UCL, namely the possibility of nonintegers values for these two parameters, will be discussed later.

This random variable is related to a FPT of the absorbing discrete time Markov chain $\{S_N(\theta), N \in I\!N_0\}$, where: $S_0(\theta) = Z_0(\theta)$; and, for $N \in I\!N$,

$$S_N(\theta) = \begin{cases} Z_N(\theta), \text{ if } Z_N(\theta) \leq x \text{ and } S_{N-1}(\theta) \leq x \\ x+1, \text{ otherwise.} \end{cases} \tag{2.48}$$

This Markov chain has finite state space $\{0, 1, \ldots, x+1\}$ and absorbing state $x+1$. Furthermore, its transitions are ruled by the transition matrix

$$\begin{bmatrix} F_\theta(k) & P_\theta(k+1) & P_\theta(k+2) & \cdots & P_\theta(k+x) & 1-F_\theta(k+x) \\ F_\theta(k-1) & P_\theta(k) & P_\theta(k+1) & \cdots & P_\theta(k+x-1) & 1-F_\theta(k+x-1) \\ F_\theta(k-2) & P_\theta(k-1) & P_\theta(k) & \cdots & P_\theta(k+x-2) & 1-F_\theta(k+x-2) \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ F_\theta(k-x) & P_\theta(k-x+1) & P_\theta(k-x+2) & \cdots & P_\theta(k) & 1-F_\theta(k) \\ 0 & 0 & 0 & \cdots & 0 & 1 \end{bmatrix}. \tag{2.49}$$

In fact, the run length of this $CUSUM$ scheme and the FPT

$$\min\{N: \; S_N(\theta) = x+1 \mid S_0(\theta) = 0\} \tag{2.50}$$

have the same distribution. •

For the sake of clarity, we restrict our attention to a control scheme whose sequence of summary statistics $\{Z_N(\theta), N \in I\!N_0\}$ form an irreducible discrete time Markov chain with infinite state space $I\!N_0$, initial (possibly random) state $Z_0(\theta)$, and transition matrix

$$\tilde{\mathbf{P}}(\theta) = [\tilde{p}_{ij}(\theta)]_{i,j \in I\!N_0}, \tag{2.51}$$

which depends on the model parameter $\theta \in \Theta$. In addition, we assume that the scheme does not trigger a signal while $Z_N(\theta) \leq x$ (i.e. $C = [0, x]$). Then, for each pair (x, θ) belonging to $I\!N \times \Theta$, we can construct an absorbing Markov chain $\{S_N(\theta), N \in I\!N_0\}$ defined as follows: $S_0(\theta) = Z_0(\theta)$; and, for $N \in I\!N$,

$$S_N(\theta) = \begin{cases} Z_N(\theta), \text{ if } Z_N(\theta) \in C \text{ and } S_{N-1}(\theta) \in C \\ x+1, \text{ otherwise.} \end{cases} \tag{2.52}$$

This Markov chain has finite state space $\{0, 1, \ldots, x+1\}$, with state $x+1$ being an absorbing state and states $0, 1, \ldots, x$ being transient states, and transition matrix, represented in partitioned form,

$$\mathbf{P}(\theta) = \begin{bmatrix} \mathbf{Q}(\theta) & [\mathbf{I} - \mathbf{Q}(\theta)]\,\underline{\mathbf{1}} \\ \underline{\mathbf{0}}^\top & 1 \end{bmatrix} \tag{2.53}$$

where:

- $\mathbf{Q}(\theta) = [\tilde{p}_{ij}(\theta)]_{i,j=0}^{x}$, i.e., this sub-stochastic $(x+1) \times (x+1)$ matrix is obtained from $\tilde{\mathbf{P}}(\theta)$ by deleting its final row and column;
- $\underline{\mathbf{1}}$ ($\underline{\mathbf{0}}^\top$) is a column vector (row vector) of $x+1$ ones (zeros); and
- $\mathbf{I}$ is the identity matrix with rank $x+1$.

As shown in Example 2.7, the RL of a control scheme can be related to the number of transitions needed for absorption to occur in the Markov chain $\{S_N(\theta), N \geq 0\}$ described in the previous paragraph. In other words, the RL is governed by a discrete phase-type distribution, depending on θ.[9]

[9] See Neuts (1981, Chap.2) for a discussion on continuous and discrete phase-type distributions.

Proposition 2.8 — *Let* $(\alpha_0(\theta), \ldots, \alpha_x(\theta), \alpha_{x+1}(\theta)) = (\underline{\alpha}(\theta), \alpha_{x+1}(\theta))$ *be the initial probability vector of the absorbing Markov chain* $\{S_N(\theta), N \in I\!N_0\}$*, where* $\underline{\alpha}(\theta)$ *ranges over all the sub-stochastic* $(x+1)-$*vectors, and* $\alpha_{x+1}(\theta) < 1$ *to avoid a trivial process. Then* $\mathbf{Q}(\theta)$ *is a sub-stochastic matrix such that* $\mathbf{I} - \mathbf{Q}(\theta)$ *is nonsingular, and the distribution of the number transitions until absorption is said to have a* $\theta-$*parameterized discrete phase-type distribution with representation* $(\underline{\alpha}(\theta), \mathbf{Q}(\theta))$.

As far as its probability law is concerned, we are dealing with a nonnegative integer-valued random value $T(\theta)$ defined as follows:

$$P_{T(\theta)}(m) = \begin{cases} \alpha_{x+1}(\theta), \ m = 0 \\ \underline{\alpha}^\top(\theta)\,[\mathbf{Q}(\theta)]^{m-1}\,[\mathbf{I} - \mathbf{Q}(\theta)]\,\underline{\mathbf{1}}, \ m \in I\!N. \end{cases} \tag{2.54}$$

Its probability generating function and factorial moment of order s may be derived from (2.54) and, following Neuts (1981, p. 46), are equal to

$$PG_{T(\theta)}(z) = \alpha_{x+1}(\theta) + z \times \underline{\alpha}^\top(\theta)\,[\mathbf{I} - z\mathbf{Q}(\theta)]^{-1}\,[\mathbf{I} - \mathbf{Q}(\theta)]\,\underline{\mathbf{1}} \tag{2.55}$$

and

$$FM_{T(\theta)}(s) = s! \times \underline{\alpha}^\top(\theta)\,[\mathbf{Q}(\theta)]^{s-1}\,[\mathbf{I} - \mathbf{Q}(\theta)]^{-s}\,\underline{\mathbf{1}}, \ s \in I\!N, \tag{2.56}$$

respectively.

One of the purposes of this thesis is to investigate closely the stochastic properties of the *RL* and, in particular, to what extent we can establish stochastic order relations involving such a performance measure. In light of the fact that the *RL* is a positive random variable, we only consider $\theta-$parameterized discrete phase-type random variables with representation $(\underline{\alpha}(\theta), \mathbf{Q}(\theta))$ such that $\underline{\alpha}^\top(\theta)\,\underline{\mathbf{1}} = 1$ (i.e. $\alpha_{x+1}(\theta) = 0$).

Corollary 2.9 — *Let* $RL^u(\theta)$ *represent the* RL *of a control scheme whose summary statistic has initial value* u*,* $u \in \{0, 1, .., x\}$*. Then* $RL^u(\theta)$ *has a* θ*-parameterized discrete phase-type distribution with representation* $(\underline{\mathbf{e}}_u, \mathbf{Q}(\theta))$*, where* $\underline{\mathbf{e}}_u$ *denotes the* $(u+1)^{th}$ *vector of the orthonormal basis for* $I\!R^{x+1}$.

Originally, Page (1954) set the initial value of the *CUSUM* statistic equal to 0. However, if $Z_0(\theta) = u > 0$ then a head start (Lucas and Crosier (1982)) value has been given to the control scheme and $RL^u(\theta)$ is the corresponding *RL*.

Another possibility (but fairly unusual in the quality control literature) is to assign a random initial value to the summary statistic, belonging to the decision interval, whose probability vector is denoted by $\underline{\alpha}(\theta)$ with $\underline{\alpha}^\top(\theta)\underline{\mathbf{1}} = 1$. In this case, $RL(\theta)$ has a θ-parameterized discrete phase-type distribution with representation $(\underline{\alpha}(\theta), \mathbf{Q}(\theta))$.[10]

For future reference we list in Table 2.3 all but the last four *RL* related measures in Table 2.1 (which concerns geometric *RL*s) for the discrete phase-type $RL(\theta)$. Note that

[10] Instead of representing the *RL* by $RL^{\underline{\alpha}(\theta)}(\theta)$, we simplify the notation and denote it by $RL(\theta)$ whenever no ambiguities arise.

from the higher order factorial moments in Table 2.3 we can derive the central moments of orders 2-4 of RL using Equations (2.15)–(2.17) and, then compute the standard deviation, and the coefficients of variation, skewness and kurtosis of $RL(\theta)$.

There are some similarities between the entries of Table 2.3 and those of Table 2.1; for that matter $\mathbf{Q}(\theta)$ can be thought as the matrix analogue of $\pi(\theta)$. This is somehow expected since the discrete phase-type distribution corresponds to a generalization of the geometric distribution used in the matrix-analytic methodology.[11]

The fact that phase-type distributions are very appealing in computational work is also apparent in Table 2.3. First, the $RL(\theta)$ related measures can be represented only in terms of two parameters ($\underline{\alpha}(\theta)$ and $\mathbf{Q}(\theta)$). Second, the evaluation of all these measures only involves the use of trivial operations such as:

- matrix multiplication (to obtain, for instance, the probability and survival functions);
- matrix inversion (to evaluate the factorial moments and the ARL).

Third, some of these measures may be computed in a recursive way. As an illustration, we recall that the probability function of $RL(\theta)$ may be computed recursively (Champ and Rigdon (1991)).

Let $\underline{RL}(\theta) = [RL^u(\theta)]_{u=0}^{x}$ be the vector of the RLs associated with the $(x+1)$ possible initial values of the summary statistic, and let $P_{\underline{RL}(\theta)}(m) = \left[P_{RL^u(\theta)}(m)\right]_{u=0}^{x}$ be the vector of the corresponding probability functions. Then

$$P_{\underline{RL}(\theta)}(m) = \mathbf{Q}(\theta)\, P_{\underline{RL}(\theta)}(m-1),\ m = 2,3,\ldots \tag{2.57}$$

and, thus,

$$\begin{aligned} P_{RL(\theta)}(m) &= \underline{\alpha}^\top(\theta)\, P_{\underline{RL}(\theta)}(m) \\ &= \underline{\alpha}^\top(\theta)\, \mathbf{Q}(\theta)\, P_{\underline{RL}(\theta)}(m-1),\ m = 2,3,\ldots. \end{aligned} \tag{2.58}$$

2.4 Example of a phase-type RL

To design of a control scheme we have to make a trade-off between a large in-control RL and a quick detection of a specific change in the process parameter. Bearing this in mind, Gan (1993) suggests, for instance, that the reference value for the upper one-sided $CUSUM$ scheme for binomial data should be selected to be close to

$$n \times \frac{ln[(1-p_0)/(1-p_1)]}{ln[(1-p_0)p_1/(1-p_1)p_0]}. \tag{2.59}$$

Recall that np_0 is the nominal expected number of defectives per random sample of size n, and np_1 denotes the corresponding out-of-control value that we want to quickly detect.[12] Gan (1993) alleged that extensive numerical results suggest that the reference value in (2.59) leads to upper one-sided $CUSUM$ schemes for *binomial* data which are optimal — in the ARL sense — in detecting an upward shift of magnitude $p_1 - p_0$.

[11] Recall that the discrete phase-type distribution is a particular case of a matrix-geometric distribution (Asmussen and O'Cinneide (1999, p. 436)).

[12] This reference value corresponds to the one used in the sequential probability ratio test (SPRT) of $H_0:\ p = p_0$ against $H_1:\ p = p_1$ $(p_1 > p_0)$.

Table 2.3: Run length related measures — discrete phase-type case ($\alpha_{x+1}(\theta) = 0$).

Probability function	$P_{RL(\theta)}(m) = \underline{\alpha}^\top(\theta)\,[\mathbf{Q}(\theta)]^{m-1}\,[\mathbf{I} - \mathbf{Q}(\theta)]\,\underline{\mathbf{1}},\; m \in I\!N$
Survival function	$\overline{F}_{RL(\theta)}(m) = \begin{cases} 1,\; m < 1 \\ \underline{\alpha}^\top(\theta)\,[\mathbf{Q}(\theta)]^{\lfloor m \rfloor}\,\underline{\mathbf{1}},\; m \geq 1 \end{cases}$
Hazard Rate function	$\lambda_{RL(\theta)}(m) = 1 - \frac{\underline{\alpha}^\top(\theta)\,[\mathbf{Q}(\theta)]^{m}\,\underline{\mathbf{1}}}{\underline{\alpha}^\top(\theta)\,[\mathbf{Q}(\theta)]^{m-1}\,\underline{\mathbf{1}}},\; m \in I\!N$
Equilibrium Rate function	$r_{RL(\theta)}(m) = \frac{\underline{\alpha}^\top(\theta)\,[\mathbf{Q}(\theta)]^{m-2}\,[\mathbf{I}-\mathbf{Q}(\theta)]\,\underline{\mathbf{1}}}{\underline{\alpha}^\top(\theta)\,[\mathbf{Q}(\theta)]^{m-1}\,[\mathbf{I}-\mathbf{Q}(\theta)]\,\underline{\mathbf{1}}},\; m = 2, 3, \ldots$
$p \times 100\%$ Percentage point	$F^{-1}_{RL(\theta)}(p) = \inf\{m \in I\!N :\; F_{RL(\theta)}(m) \geq p\},\; 0 < p < 1$
Probability Generating f.	$PG_{RL(\theta)}(z) = z \times \underline{\alpha}^\top(\theta)\,[\mathbf{I} - z\mathbf{Q}(\theta)]^{-1}\,[\mathbf{I} - \mathbf{Q}(\theta)]\,\underline{\mathbf{1}},\; 0 \leq z \leq 1$
Factorial Moment of order s	$FM_{RL(\theta)}(s) = s! \times \underline{\alpha}^\top(\theta)\,[\mathbf{Q}(\theta)]^{s-1}\,[\mathbf{I} - \mathbf{Q}(\theta)]^{-s}\,\underline{\mathbf{1}},\; s \in I\!N$
Expected Value	$E[RL(\theta)] = \underline{\alpha}^\top(\theta)\,[\mathbf{I} - \mathbf{Q}(\theta)]^{-1}\,\underline{\mathbf{1}}$

Example 2.10 — Consider the upper one-sided *CUSUM* scheme with no head start described in Example 2.7 with reference value $k = 3$ — which corresponds to $np_1 = 4.27685$ according to Equation (2.59) — and upper control limit $x = 6$. In this case, the in-control RL, $RL^0(0)$ has a discrete phase-type distribution represented by $(\underline{\mathbf{e}}_0, \mathbf{Q}(0))$, where the in-control substochastic matrix $\mathbf{Q}(0)$ equals:

$$\begin{bmatrix} 0.8590 & 0.0902 & 0.0353 & 0.0114 & 0.0031 & 0.0007 & 0.0002 \\ 0.6767 & 0.1823 & 0.0902 & 0.0353 & 0.0114 & 0.0031 & 0.0007 \\ 0.4033 & 0.2734 & 0.1823 & 0.0902 & 0.0353 & 0.0114 & 0.0031 \\ 0.1326 & 0.2707 & 0.2734 & 0.1823 & 0.0902 & 0.0353 & 0.0114 \\ 0 & 0.1326 & 0.2707 & 0.2734 & 0.1823 & 0.0902 & 0.0353 \\ 0 & 0 & 0.1326 & 0.2707 & 0.2734 & 0.1823 & 0.0902 \\ 0 & 0 & 0 & 0.1326 & 0.2707 & 0.2734 & 0.1823 \end{bmatrix}. \qquad (2.60)$$

The parameters yield to a scheme with ARL at the nominal and out-of-control values np_0 and np_1 equal to $ARL^0(0) = 1015.71$ — which is close to the in-control ARL of the $np-$scheme in Example 2.4, 1073.03 — and $ARL^0(p_1 - p_0) = 5.932$, as reported in Table 2.4.

This table pictures the stochastic behaviour of $RL^0(\theta)$, through the inclusion of several RL related measures, for $\theta = 0, 0.001, 0.0025, 0.005, 0.0075, 0.01, 0.02, p_1 - p_0, 0.03$. It also illustrates how unreliable the ARL can be as a performance measure of a control scheme, in the on-target situation; for instance, the probability of a signal being triggered within the first 295 samples is of at least 0.25, although the in-control ARL barely exceeds 1015 samples. Besides, in the absence of a shift in p, the $SDRL$ is also about 1000 samples, therefore it is possible to have observations beyond the control limits much sooner or much later than expected.

Table 2.4: Some *RL* percentage points, *ARL*, *SDRL*, *CVRL*, *CSRL* and *CKRL* values for an upper one-sided *binomial CUSUM* scheme ($n = 100, p_0 = 0.02, p_1 = 0.0427685$) and the upper one-sided *np*–scheme from Example 2.4.

Upper one-sided *binomial CUSUM* scheme									
RL perc. points	$\theta = p - p_0$								
	0	0.001	0.0025	0.005	0.0075	0.01	0.02	$p_1 - p_0$	0.03
5%	55	34	18	9	6	4	2	2	2
25%	295	173	85	32	16	10	4	4	3
Median	705	411	198	72	33	19	6	5	4
75%	1407	819	392	140	63	34	9	7	5
90%	2334	1358	649	230	101	53	13	10	7
95%	3036	1765	843	297	130	68	16	12	8
ARL	1015.71	591.724	284.121	102.081	46.227	25.458	7.194	5.932	4.095
SDRL	1012.18	588.012	280.175	97.895	42.022	21.419	4.320	3.322	1.998
CVRL	0.997	0.994	0.986	0.959	0.909	0.841	0.600	0.560	0.488
CSRL	2.000	2.000	2.000	1.998	1.989	1.961	1.627	1.523	1.303
CKRL	6.000	6.000	5.999	5.992	5.953	5.833	4.296	3.814	2.853
Upper one-sided *np*–scheme									
RL perc. points	$\theta = p - p_0$								
	0	0.001	0.0025	0.005	0.0075	0.01	0.02	$p_1 - p_0$	0.03
5%	56	41	27	14	8	5	2	1	1
25%	309	227	148	78	45	27	6	5	3
Median	744	546	355	187	107	65	15	11	6
75%	1487	1092	710	374	214	130	29	21	11
90%	2470	1813	1179	621	355	216	48	35	17
95%	3214	2359	1534	808	461	281	62	45	22
ARL	1073.030	787.737	512.346	270.112	154.275	94.128	21.047	15.369	7.815
SDRL	1072.530	787.237	511.846	269.611	153.774	93.627	20.541	14.861	7.298
CVRL	1.000	0.999	0.999	0.998	0.997	0.995	0.976	0.967	0.934
CSRL	2.000	2.000	2.000	2.000	2.000	2.000	2.001	2.001	2.005
CKRL	6.000	6.000	6.000	6.000	6.000	6.000	6.002	6.005	6.019

We can also add that the well-known *RL* distribution skewness to the right of the *RL* of Markov-type schemes (see Woodall (2000)) starts to steadily decrease only for moderate values of the magnitude of the shift; the same behaviour holds for the coefficient of kurtosis of this scheme. However, these two coefficients increase with θ when we make use of the upper one-sided *np*–scheme.

Moreover, replacing the upper one-sided *np*–scheme in Example 2.4 by the upper one-sided *binomial CUSUM* scheme yields a reduction in both the *ARL* and *SDRL* and in most of percentage points, as illustrated by Table 2.4. •

Before we proceed into the special features of the stochastic matrix $\mathbf{P}(\theta)$ we present three brief notes.

Any standard Markov-type scheme, with discrete summary statistics and a null (non-null) lower control limit, can also be associated to an absorbing Markov chain with state space $\{0, 1, \ldots, x, x+1\}$ (or $\{-x, \ldots, -1, 0, 1, \ldots, x, x+1\}$ as in Chapter 6); this results from a suitable left or right shifting of the original Markov chain, and from associating both the values below LCL and above UCL to the absorbing state $x+1$.

In case the summary statistic takes fractional values the Markov approach can be applied after covering those values by suitable rescaling, as suggested by Brook and Evans (1972), Lucas (1985) and Gan (1993). See Chapter 4 for further details.

The application of the Markov chain approach to schemes for continuous data is addressed in Section 3.3 and Chapters 5-6.

2.5 Special features of the transition matrix

A few features are apparent in the matrix $\mathbf{Q}(\theta)$ given in Equation (2.60) from Example 2.10. As suggested by Brook and Evans (1972), since we are dealing with a nonnegative random variable, the substochastic matrix $\mathbf{Q}(\theta)$ has a triangular block of zeros in the lower left hand corner corresponding to states $i = k+1, \ldots, x$. We can also add that for the last x columns all values along a line parallel to the main diagonal are equal. In fact, the absorbing Markov chain is space homogeneous with two boundaries: a reflecting one (0), and an absorbing one $(x+1)$.

The matrices $\mathbf{Q}(\theta)$ and, in particular, $\mathbf{P}(\theta)$ that usually arise in the quality control literature have far more important features than these.

In what follows, we shall omit the argument θ and consider it fixed. Moreover, we consider a transition matrix $\mathbf{P} = [p_{ij}]_{i,j=0}^{x+1}$; it depends on θ and governs the behaviour of a Markov chain $\{S_N, N \in I\!N_0\}$ with state space $\{0, 1, \ldots, x+1\}$.[13] Moreover, $\leq_*$ denotes some fixed but unspecified stochastic order relation.

Definition 2.11 — *The Markov chain $\{S_N, N \in I\!N_0\}$ is said to be stochastically monotone in the $*-$sense if, for $N \in I\!N_0$,*

$$(S_{N+1}|S_N = i) \leq_* (S_{N+1}|S_N = m)\,,\ 0 \leq i \leq m \leq x+1. \tag{2.61}$$

In this case we write $\{S_N, N \geq 0\} \in \mathcal{M}_*$ or $\mathbf{P} \in \mathcal{M}_*$.

Remark 2.12 — The constraint (2.61) is equivalent to

$$(S_{N+1}|S_N = i) \leq_* (S_{N+1}|S_N = i+1)\,,\ 0 \leq i \leq x \tag{2.62}$$

or, in other words,

$$(S_{N+1}|S_N = i) \uparrow_* \text{ with } i \tag{2.63}$$

over the set $\{0, 1, \ldots, x+1\}$, for $N \in I\!N_0$. Furthermore $\mathcal{M}_*, * = st, hr, rh, lr$, stands here for the class of all matrices that are

[13] This finite state space setting is chosen for convenience and has no bearing on most of the auxiliary notions and results stated ahead.

- stochastically monotone in the usual sense ($\mathcal{M}_{st}$),
- stochastically monotone in the hazard rate sense ($\mathcal{M}_{hr}$),
- stochastically monotone in the reversed hazard rate sense ($\mathcal{M}_{rh}$), and
- stochastically monotone in the likelihood ratio sense ($\mathcal{M}_{lr}$).

Actually, Definition 2.11 is nothing but Corollary 3.5 in Kijima (1997, pp. 130-131) restated. •

Remark 2.13 — The notion of stochastically monotone matrices in the usual sense was introduced by Daley (1968) for real-valued discrete time Markov chains. We were not able to trace back the origin of the notion of stochastically monotone Markov chains in the hazard rate sense ($\mathcal{M}_{hr}$) or in the reversed hazard rate sense ($\mathcal{M}_{rh}$). As for the notion of stochastically monotone Markov chains in the likelihood ratio sense ($\mathcal{M}_{lr}$): a reading of Kijima (1998) misleadingly suggests that they can be traced back to Karlin and McGregor (1959);[14] however, Karlin (1964) implicitely states that totally positive of order 2 [15] transition matrices possess a monotone likelihood ratio property and, thus, gets very close to defining stochastically monotone Markov chains in the likelihood ratio sense.

Recalling that the i^{th} row of $\mathbf{P}$ corresponds to the probability (row vector) of the random variable $(S_{N+1}|S_N = i)$, and taking advantage of the notions of $\le_{st}$, $\le_{hr}$, $\le_{rh}$ and $\le_{lr}$, we can conclude that the easiest way of investigating whether $\mathbf{P} \in \mathcal{M}_*$, $* = st, hr, rh, lr$, is to check if (respectively):

- $\sum_{l=j}^{x+1} p_{il} = P(S_{N+1} \ge j|S_N = i) \uparrow_i$ over the set $\{0, 1, \dots, x+1\}$, for any fixed j;
- $p_{ij} / \sum_{l=j}^{x+1} p_{il} \downarrow_i$ over the set $\{0, 1, \dots, x+1\}$, for any fixed j;
- $p_{ij} / \sum_{l=0}^{j} p_{il} \uparrow_i$ over the set $\{0, 1, \dots, x+1\}$, for any fixed j;
- $p_{ij}/p_{i+1\,j} \downarrow_j$ over the set $\{0, 1, \dots, x+1\}$, for any fixed i such that $0 \le i \le x$. •

In order to rephrase once more the four notions mentioned in Remark 2.12, we clearly need a few preparatory definitions.

First, we consider the square matrix $\mathbf{U}$, introduced by Keilson and Kester (1977), with ones on and below the diagonal and zeroes elsewhere. That is,

$$\mathbf{U} = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ 1 & 1 & 0 & \dots & 0 \\ 1 & 1 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & 1 & \dots & 1 \end{bmatrix} \quad \text{and} \quad \mathbf{U}^{-1} = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ -1 & 1 & 0 & \dots & 0 \\ 0 & -1 & 1 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & -1 & 1 \end{bmatrix}. \tag{2.64}$$

[14] In fact, Karlin and McGregor (1959) deal with continuous time Markov chains governed by totally positive of order 2 probability transition matrices (see definition ahead). However, these matrices are not even called as stochastically monotone.

[15] For the definition of totally positivity of order 2 see Definition 2.14 below, and for the relation between monotone Markov chains in the likelihood ration sense and Markov chains possessing a totally positive of order 2 transition matrix please refer to Proposition 2.15.

The rank of the matrix $\mathbf{U}$ will not be explicitly mentioned in the text, so that it should be inferred as the appropriate one in each instance of $\mathbf{U}$.

Finally, we recall the notions of total positivity of order 2 (abbreviated TP_2) matrices, and, for the sake of completeness and for future reference, of sign-regular of order 2 (RR_2) matrices.[16]

Definition 2.14 — *The nonnegative matrix* $\mathbf{A} = [a_{ij}]_{i,j=0}^{x+1}$ *is said to be totally positive of order* 2*, denoted here by* $\mathbf{A} \in TP_2$*, iff all the* 2×2 *minors of* $\mathbf{A}$ *are nonnegative; i.e.,*

$$a_{ij} \times a_{i'j'} \geq a_{i'j} \times a_{ij'},\ 0 \leq i \leq i' \leq x+1,\ 0 \leq j \leq j' \leq x+1. \tag{2.65}$$

If the inequality (2.65) is reversed, $\mathbf{A}$ *is called a sign-regular of order 2 matrix, i.e.* $\mathbf{A} \in RR_2$*.*

Note that $\mathbf{U}$ is a TP_2 matrix and in what follows the inequality $\mathbf{A} \geq \mathbf{B}$, for matrices $\mathbf{A} = [a_{ij}]$ and $\mathbf{B} = [b_{ij}]$ with the same dimension, is in the componentwise sense; i.e., $\mathbf{A} \geq \mathbf{B}$ if and only if $a_{ij} \geq b_{ij}$, for all i, j.

Proposition 2.15 — *As pointed out in Definition 3.11 from Kijima (1997, p. 129) the following results are valid:*

$$\mathbf{P} \in \mathcal{M}_{st} \Leftrightarrow \mathbf{U}^{-1}\mathbf{P}\mathbf{U} \geq \mathbf{O} \Leftrightarrow (\mathbf{U}^{\top})^{-1}\mathbf{P}\mathbf{U}^{\top} \geq \mathbf{O} \tag{2.66}$$

$$\mathbf{P} \in \mathcal{M}_{hr} \Leftrightarrow \mathbf{P}\mathbf{U} \in TP_2 \tag{2.67}$$

$$\mathbf{P} \in \mathcal{M}_{rh} \Leftrightarrow \mathbf{P}\mathbf{U}^{\top} \in TP_2 \tag{2.68}$$

$$\mathbf{P} \in \mathcal{M}_{lr} \Leftrightarrow \mathbf{P} \in TP_2. \tag{2.69}$$

Clearly,

$$\mathbf{P} \in \mathcal{M}_{lr} \Rightarrow \begin{cases} \mathbf{P} \in \mathcal{M}_{hr} \\ \mathbf{P} \in \mathcal{M}_{rh} \end{cases} \Rightarrow \mathbf{P} \in \mathcal{M}_{st}. \tag{2.70}$$

Remark 2.16 — Let $(\mathbf{B})_{ij}$ denote the entry in row i and column j of a matrix $\mathbf{B}$ and let $\mathbf{A} = [a_{ij}]_{i,j=0}^{x+1}$ be an $(x+2) \times (x+2)$ matrix. Then, for $i, j = 0, 1, \ldots, x+1$:

$$(\mathbf{A}\mathbf{U})_{ij} = \sum_{l=j}^{x+1} a_{i\,l} \text{ and } (\mathbf{A}\mathbf{U}^{\top})_{ij} = \sum_{l=0}^{j} a_{i\,l}, \tag{2.71}$$

$$(\mathbf{U}^{-1}\mathbf{A}\mathbf{U})_{ij} = \sum_{l=j}^{x+1} a_{i\,l} - \sum_{l=j}^{x+1} a_{i-1\,l} \text{ and}$$

$$((\mathbf{U}^{\top})^{-1}\mathbf{A}\mathbf{U}^{\top})_{ij} = \textstyle\sum_{l=0}^{j} a_{i\,l} - \sum_{l=0}^{j} a_{i+1\,l}, \tag{2.72}$$

[16] For the definitions of totally positive of order r (TP_r) and sign-regular of order r (RR_r) functions see Karlin (1964) or Karlin (1968, Chap.1). This last reference is an authorative and comprehensive treatment of total positivity theory and its applications in a wide variety of fields.

where $a_{-1\,l} = a_{x+2\,l} = 0$ for $l = 0, 1, \ldots, x+1$.

Also note that the second equivalence in (2.66) does not hold in general for non-stochastic matrices. For example, if

$$\mathbf{A} = \begin{bmatrix} 0.5 & 0 \\ 0 & 1 \end{bmatrix} \quad \text{and} \quad \mathbf{B} = \begin{bmatrix} 1 & 0 \\ 0 & 0.5 \end{bmatrix} \tag{2.73}$$

then $\mathbf{U}^{-1}\mathbf{A}\mathbf{U} \geq \mathbf{O}$ and $(\mathbf{U}^\top)^{-1}\mathbf{B}\mathbf{U}^\top \geq \mathbf{O}$ whereas $(\mathbf{U}^\top)^{-1}\mathbf{A}\mathbf{U}^\top \ngeq \mathbf{O}$ and $\mathbf{U}^{-1}\mathbf{B}\mathbf{U} \ngeq \mathbf{O}$. •

When we are dealing with an absorbing Markov chain with transition matrix $\mathbf{P}$ as in (2.53) we can infer several properties of the substochastic matrix $\mathbf{Q}$ (which governs the transition between the transient states) from the features of $\mathbf{P}$, and vice-versa, as stated in the following lemma.

Lemma 2.17 — *Let $\mathbf{P} = [p_{ij}]_{i,j=0}^{x+1}$ be the transition matrix of a Markov chain with absorbing state $x+1$, so that $\mathbf{P}$ can be represented in the partitioned form*

$$\mathbf{P} = \begin{bmatrix} \mathbf{Q} & (\mathbf{I}-\mathbf{Q})\,\underline{\mathbf{1}} \\ \underline{\mathbf{0}}^\top & 1 \end{bmatrix}. \tag{2.74}$$

Then the following relations hold:

$$(\mathbf{U}^\top)^{-1}\mathbf{Q}\mathbf{U}^\top \geq \mathbf{O} \Leftrightarrow (\mathbf{U}^\top)^{-1}\mathbf{P}\mathbf{U}^\top \geq \mathbf{O} \Leftrightarrow \mathbf{U}^{-1}\mathbf{P}\mathbf{U} \geq \mathbf{O} \tag{2.75}$$

$$\mathbf{P} \in TP_2 \Rightarrow \mathbf{Q} \in TP_2 \tag{2.76}$$

$$\mathbf{P}\mathbf{U}^\top \in TP_2 \Rightarrow \mathbf{Q}\mathbf{U}^\top \in TP_2 \tag{2.77}$$

However, in general,

$$\mathbf{U}^{-1}\mathbf{P}\mathbf{U} \geq \mathbf{O} \not\Rightarrow \mathbf{U}^{-1}\mathbf{Q}\mathbf{U} \geq \mathbf{O} \quad \textit{and} \quad \mathbf{U}^{-1}\mathbf{Q}\mathbf{U} \geq \mathbf{O} \not\Rightarrow \mathbf{U}^{-1}\mathbf{P}\mathbf{U} \geq \mathbf{O} \tag{2.78}$$

$$\mathbf{Q} \in TP_2 \not\Rightarrow \mathbf{P} \in TP_2 \tag{2.79}$$

$$\mathbf{Q}\mathbf{U}^\top \in TP_2 \not\Rightarrow \mathbf{P}\mathbf{U}^\top \in TP_2 \tag{2.80}$$

$$\mathbf{P}\mathbf{U} \in TP_2 \not\Rightarrow \mathbf{Q}\mathbf{U} \in TP_2 \quad \textit{and} \quad \mathbf{Q}\mathbf{U} \in TP_2 \not\Rightarrow \mathbf{P}\mathbf{U} \in TP_2. \tag{2.81}$$

Proof — Equation (2.75) follows from the second equivalence in (2.66) and the fact that

$$(\mathbf{U}^\top)^{-1}\mathbf{P}\mathbf{U}^\top = \begin{bmatrix} (\mathbf{U}^\top)^{-1}\mathbf{Q}\mathbf{U}^\top & \underline{\mathbf{0}} \\ \underline{\mathbf{0}}^\top & 1 \end{bmatrix}. \tag{2.82}$$

Results (2.76) and (2.77) are immediate consequences of the definition of total positivity of order 2 and the form of $\mathbf{P}$ and $\mathbf{P}\mathbf{U}^\top$, respectively, since

$$\mathbf{P}\mathbf{U}^\top = \begin{bmatrix} \mathbf{Q}\mathbf{U}^\top & \underline{\mathbf{1}} \\ \underline{\mathbf{0}}^\top & 1 \end{bmatrix}. \tag{2.83}$$

The relations (2.78)–(2.81) follow quite easily if we take, e.g.,

$$\mathbf{P}_1 = \begin{bmatrix} \mathbf{Q}_1 & (\mathbf{I}-\mathbf{Q}_1)\underline{\mathbf{1}} \\ \underline{\mathbf{0}}^\top & 1 \end{bmatrix} = \begin{bmatrix} 0.2 & 0 & 0.8 \\ 0.1 & 0.3 & 0.6 \\ 0 & 0 & 1 \end{bmatrix} \tag{2.84}$$

and

$$\mathbf{P}_2 = \begin{bmatrix} \mathbf{Q}_2 & (\mathbf{I}-\mathbf{Q}_2)\underline{\mathbf{1}} \\ \underline{\mathbf{0}}^\top & 1 \end{bmatrix} = \begin{bmatrix} 0.4 & 0.1 & 0.5 \\ 0.3 & 0 & 0.7 \\ 0 & 0 & 1 \end{bmatrix}. \tag{2.85}$$

In this case, we have

$$\mathbf{Q}_1 \in TP_2,\ \mathbf{Q}_1\mathbf{U} \in TP_2,\ \mathbf{Q}_1\mathbf{U}^\top \in TP_2,\ \mathbf{U}^{-1}\mathbf{Q}_1\mathbf{U} \geq \mathbf{O},\ \text{and}$$
$$\mathbf{P}_2\mathbf{U} \in TP_2 \text{ and } \mathbf{U}^{-1}\mathbf{P}_2\mathbf{U} \geq \mathbf{O} \tag{2.86}$$

whereas

$$\mathbf{P}_1 \notin TP_2,\ \mathbf{P}_1\mathbf{U} \notin TP_2,\ \mathbf{P}_1\mathbf{U}^\top \notin TP_2,\ \mathbf{U}^{-1}\mathbf{P}_1\mathbf{U} \not\geq \mathbf{O},\ \text{and}$$
$$\mathbf{Q}_2\mathbf{U} \notin TP_2 \text{ and } \mathbf{U}^{-1}\mathbf{Q}_2\mathbf{U} \not\geq \mathbf{O}. \tag{2.87}$$

•

Since the classes of stochastic matrices $\mathcal{M}_*, * = st, hr, rh, lr$, include many of the common Markov models of applied probability theory (see, e.g., Keilson and Kester (1977) and Bäuerle and Rolsky (2000)) it is natural to inquire whether there are control schemes (with dependent statistics) associated to these four classes of matrices.

Let us take an upper one-sided *CUSUM* scheme for discrete data with summary statistic

$$\mathrm{Z}_N = \begin{cases} 0, & N = 0 \\ \max\{0, Z_{N-1} + Y_N - k\}, & N \in I\!N, \end{cases} \tag{2.88}$$

such as the one described in Example 2.7, with θ omitted. That is, a control scheme whose RL has a discrete phase-type distribution related to a finite, space homogeneous, absorbing Markov chain with two boundaries — a reflecting one (0), and an absorbing one $(x+1)$ — and governed by the transition matrix

$$\mathbf{P} = \begin{bmatrix} F(k) & P(k+1) & P(k+2) & \cdots & P(k+x) & \overline{F}(k+x) \\ F(k-1) & P(k) & P(k+1) & \cdots & P(k+x-1) & \overline{F}(k+x-1) \\ F(k-2) & P(k-1) & P(k) & \cdots & P(k+x-2) & \overline{F}(k+x-2) \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ F(k-x) & P(k-x+1) & P(k-x+2) & \cdots & P(k) & \overline{F}(k) \\ 0 & 0 & 0 & \cdots & 0 & 1 \end{bmatrix} \tag{2.89}$$

where $F(.)$, $P(.)$ and $\overline{F}(.)$ refer to the distribution, the probability and the survival functions of the nonnegative integer random variable Y_N. This matrix can be written in a condensed way that makes the spatial homogeneity far more obvious: $\mathbf{P} = [p_{ij}]_{i,j=0}^{x+1}$, where

$$p_{i\,j} = \begin{cases} F[k+(j-i)], \ 0 \le i \le x+1, \ j=0 \\ P[k+(j-i)], \ 0 \le i \le x, \ 1 \le j \le x \\ \overline{F}[k+(j-i)-1], \ 0 \le i \le x, \ j=x+1 \\ I_{\{0\}}(j-i), \ i=x+1, \ 0 \le j \le x+1. \end{cases} \tag{2.90}$$

The next result, Proposition 2.20, provides sufficient conditions that guarantee that the transition matrix $\mathbf{P}$ belongs to the classes $\mathcal{M}_{st}$, $\mathcal{M}_{hr}$, $\mathcal{M}_{rh}$ and $\mathcal{M}_{lr}$. However, before we present the result, we recall a few important ageing notions taken from Kijima (1997, Section 3.2). We use the convention $0/0 = 0$.

Definition 2.18 — *Consider a nonnegative integer random variable Y with probability function P and distribution function F. Then*

- *Y is new better than used, $Y \in NBU$, if*

$$\overline{F}(i+j) \le \overline{F}(i)\overline{F}(j), \ i,j \in I\!N_0; \tag{2.91}$$

- *Y has increasing hazard rate in average, $Y \in IHRA$, if*

$$\overline{F}^{i+1}(i-1) \ge \overline{F}^{i}(i), \ i \in I\!N. \tag{2.92}$$

- *Y has increasing hazard rate, $Y \in IHR$, if*

$$\overline{F}^{2}(i) \ge \overline{F}(i-1) \times \overline{F}(i+1), \ i \in I\!N; \tag{2.93}$$

- *Y has decreasing reversed hazard rate, $Y \in DRHR$, if*

$$F^{2}(i) \ge F(i-1) \times F(i+1), \ i \in I\!N; \tag{2.94}$$

- *Y has decreasing likelihood ratio (i.e. Y has a Pólya frequency character of order 2), $Y \in DLR$, if*

$$P^{2}(i+1) \ge P(i) \times P(i+2), \ i \in I\!N_0. \tag{2.95}$$

Remark 2.19 — Note that by raising both member of (2.92) to the power $\frac{1}{i(i+1)}$ we can conclude that

$$Y \in IHRA \Leftrightarrow \overline{F}^{\frac{1}{i+1}}(i) \downarrow_i \ \text{over } I\!N_0. \tag{2.96}$$

In addition, the following equivalent characterizations of properties *IHR*, *DRHR* and *DLR* [17] hold, as reported by Kijima (1997, p. 113-115):

$$Y \in IHR \quad \Leftrightarrow \quad \frac{P(i)}{\overline{F}(i-1)} \uparrow_i \ \text{over } I\!N_0 \tag{2.97}$$

$$\Leftrightarrow \quad \begin{bmatrix} \overline{F}(1) & \overline{F}(2) & \overline{F}(3) & \dots \\ \overline{F}(0) & \overline{F}(1) & \overline{F}(2) & \dots \end{bmatrix} \in TP_2 \tag{2.98}$$

[17] If we had defined the likelihood ratio as equal to the equilibrium rate, $P(i-1)/P(i)$, as Shanthikumar and Yao did in Shaked and Shanthikumar (1994, p. 437), then Y would have an increasing lihelihood ratio.

$$Y \in DRHR \quad \Leftrightarrow \quad \frac{P(i)}{F(i)} \downarrow_i \text{ over } \mathbb{N}_0 \tag{2.99}$$

$$\Leftrightarrow \quad \begin{bmatrix} F(1) & F(2) & F(3) & \dots \\ F(0) & F(1) & F(2) & \dots \end{bmatrix} \in TP_2 \tag{2.100}$$

$$Y \in DLR \quad \Leftrightarrow \quad \frac{P(i+1)}{P(i)} \downarrow_i \text{ over } \mathbb{N}_0 \tag{2.101}$$

$$\Leftrightarrow \quad \begin{bmatrix} P(0) & P(1) & P(2) & \dots \\ 0 & P(0) & P(1) & \dots \end{bmatrix} \in TP_2. \tag{2.102}$$

It is worth mentioning that (2.97), (2.99) and (2.101) rightly justify the names given to the ageing notions *IHR*, *DRHR* and *DLR*. •

Kijima (1997, pp. 120-121) provides some illustrative examples of random variables with decreasing likelihood ratio, namely the *binomial* and *Poisson* distributions, and adds, in page 118, a well known fact:

$$Y \in DLR \Rightarrow \begin{cases} Y \in IHR \\ Y \in DRHR \end{cases} \Rightarrow Y \in IHRA \Rightarrow Y \in NBU. \tag{2.103}$$

We now present Proposition 2.20, which is inspired by Example 3.11 from Kijima (1997, p. 131). This example only provides proofs of results analogous to (2.104) and (2.105).

Proposition 2.20 — *Let Y be a random variable with the same distribution as Y_N defined in (2.88) and* $\mathbf{P}$ *be the transition matrix defined by (2.89). Then the following results hold:*

$$\mathbf{P} \in \mathcal{M}_{st} \tag{2.104}$$

$$Y \in IHR \Rightarrow \mathbf{P} \in \mathcal{M}_{hr} \tag{2.105}$$

$$Y \in DRHR \Rightarrow \mathbf{P} \in \mathcal{M}_{rh} \tag{2.106}$$

$$Y \in DLR \Rightarrow \mathbf{P} \in \mathcal{M}_{lr} \tag{2.107}$$

Proof — A sketch of the proof of results (2.106) and (2.107) immediately follows.

If $Y \in DRHR$ then we get from (2.100)

$$F(i)F(j) \geq F(i-1) \times F(j+1), \ 1 \leq i \leq j. \tag{2.108}$$

This result in turn implies that

$$\mathbf{PU}^\top = \begin{bmatrix} F(k) & F(k+1) & F(k+2) & \cdots & F(k+x) & 1 \\ F(k-1) & F(k) & F(k+1) & \cdots & F(k+x-1) & 1 \\ F(k-2) & F(k-1) & F(k) & \cdots & F(k+x-2) & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ F(k-x) & F(k-x+1) & F(k-x+2) & \cdots & F(k) & 1 \\ 0 & 0 & 0 & \cdots & 0 & 1 \end{bmatrix} \in TP_2, \tag{2.109}$$

i.e., $\mathbf{P} \in \mathcal{M}_{rh}$.

In case $Y \in DLR$, it follows from (2.102) that

$$P(i)\,P(j) \geq P(i-1)\,P(j+1),\ 1 \leq i \leq j. \tag{2.110}$$

Moreover, due to (2.65), we conclude that $\mathbf{P} = [p_{ij}]_{i,j=0}^{x+1} \in \mathcal{M}_{lr}$ if

$$\frac{p_{i0}}{p_{i+1\,0}} \geq \frac{p_{i1}}{p_{i+1\,1}} \geq \frac{p_{i2}}{p_{i+1\,2}} \geq \ldots \geq \frac{p_{ix}}{p_{i+1\,x}} \geq \frac{p_{i\,x+1}}{p_{i+1\,x+1}} \tag{2.111}$$

holds for $0 \leq i \leq x-1$. We shall now prove (2.111) using (2.110). The first inequality of (2.111) is valid since

$$p_{i0}\,p_{i+1\,1} - p_{i+1\,0}\,p_{i\,1} \quad = \quad F(k-i)\,P(k-i) - F(k-i-1)\,P(k-i+1) \tag{2.112}$$

equals the nonnegative sum

$$\begin{aligned} P(0)\,P(k-i) \quad &+ \quad [P(1)\,P(k-i) - P(0)\,P(k-i+1)] \\ &+ \quad [P(2)\,P(k-i) - P(1)\,P(k-i+1)] \\ &+ \quad \ldots \\ &+ \quad [P(k-i)\,P(k-i) - P(k-i-1)\,P(k-i+1)]. \end{aligned} \tag{2.113}$$

As for the last inequality, we can state that

$$\begin{aligned} p_{ix}\,p_{i+1\,x+1} - p_{i+1\,x}\,p_{i\,x+1} \quad = \quad & P(k+x-i)\,[1 - F(k+x-i-1)] \\ & - P(k+x-i-1)\,[1 - F(k+x-i)] \end{aligned} \tag{2.114}$$

equals

$$\sum_{j=0}^{+\infty} [P(k+x-i)\,P(k+x-i+j) - P(k+x-i-1)\,P(k+x-i+j+1)] \tag{2.115}$$

which is nonnegative as well, by (2.110).

The remaining inequalities in (2.111) are particular cases of (2.110). •

Features such as the total positivity of order 2 of the transition matrix play a vital role is establishing monotonicity properties concerning the RL distribution.[18] This issue is discussed in more detail in the next chapter.

2.6 A few ageing properties of RL

The immense literature on FPTs has extensively devoted its attention to the study of the ageing properties of these random variables (see Shaked and Li (1997) and references therein). It shows that such properties are closely related to the stochastic monotonicity properties of the underlying process, as noted by Li and Shaked (1995).

This observation takes form in a series of results stated in Shaked and Li (1997); a few of these results can be found in the next lemma conveniently restated for the RL.

[18]Such features also influence the monotonicity and unimodality of the transition probabilities. For related results see Karlin (1964) and Section 3.5 of Kijima (1997).

Lemma 2.21 — *Let $\{S_N, N \in I\!N_0\}$ be a discrete time Markov chain with state space $I\!N_0$, governed by the transition matrix $\mathbf{P} = [p_{ij}]_{i,j \in I\!N_0}$. Furthermore, consider the following* FPT*:*

$$RL^u = \min\{N : \; S_N > x \mid S_0 = u\}, \; u = 0, 1, \ldots, x. \tag{2.116}$$

That is, RL^u has a discrete phase-type distribution with parameter $(\underline{\mathbf{e}}_u, \mathbf{Q})$ where $\mathbf{Q} = [p_{ij}]_{i,j=0}^x$ Then:

$$\mathbf{P} \in \mathcal{M}_{st} \Rightarrow RL^u \in NBU, \; u = 0, 1, \ldots, x \tag{2.117}$$

$$\mathbf{P} \in \mathcal{M}_{st} \text{ and } \mathbf{P} \text{ is upper triangular } \Rightarrow RL^0 \in IHRA \tag{2.118}$$

$$\mathbf{P} \in \mathcal{M}_{rh} \Rightarrow RL^0 \in IHR \tag{2.119}$$

$$\mathbf{P} \in \mathcal{M}_{lr} \Rightarrow RL^0 \in DLR. \tag{2.120}$$

Results (2.117), (2.119) and (2.120) are due to Brown and Chaganty (1983), Durham, Lynch and Padgett (1990) and Assaf, Shaked and Shanthikumar (1985), respectively. We can also add that Shaked and Li (1997) obtained (2.118) from a general result of Shaked and Shanthikumar (1987).

We were unable to prove the discrete analogue of Theorem 4.1 of Kijima (1998):

$$\mathbf{P} \in \mathcal{M}_{hr} \Rightarrow RL^0 \in DRHR. \tag{2.121}$$

However, Lemma 3.1 of Kijima (1992) suggests that such a ageing character cannot be derived by merely using $\mathbf{P}$. In fact this lemma asserts that:

$$\left.\begin{array}{l} \mathbf{QU} \in TP_2 \\ \mathbf{U}^{-1}\mathbf{QU} \geq \mathbf{O} \\ \begin{bmatrix} \underline{\alpha} \\ \underline{\alpha}\mathbf{Q} \end{bmatrix} \mathbf{U} \in TP_2 \; (RR_2) \end{array}\right\} \Rightarrow RL \in DHR \; (IHR) \tag{2.122}$$

$$\left.\begin{array}{l} \mathbf{QU}^\top \in TP_2 \\ (\mathbf{U}^\top)^{-1}\mathbf{QU}^\top \geq \mathbf{O} \\ \begin{bmatrix} \underline{\alpha} \\ \underline{\alpha}\mathbf{Q} \end{bmatrix} \mathbf{U}^\top \in TP_2 \; (RR_2) \end{array}\right\} \Rightarrow RL \in IHR \; (DHR), \tag{2.123}$$

where RL stands for the run length considering a random initial value with probability vector $(\underline{\alpha}, 0)$. The proof of this lemma follows from four auxiliary and useful results, presented below and taken from pages 122, 107 (Theorem 3.1), 108 (Theorem 3.2) and 110 (Theorem 3.3) of Kijima (1997), respectively.

Proposition 2.22 — *Let X and Y be discrete random variables defined on $I\!N$, with probability vectors $\underline{\mathbf{a}}$ and $\underline{\mathbf{b}}$, respectively. Then:*

$$X \leq_{st} Y \; (\text{or } \underline{\mathbf{a}} \leq_{st} \underline{\mathbf{b}}) \Leftrightarrow \underline{\mathbf{a}}^\top \mathbf{U} \leq \underline{\mathbf{b}}^\top \mathbf{U} \; (\text{or } \underline{\mathbf{a}}^\top \mathbf{U}^\top \geq \underline{\mathbf{b}}^\top \mathbf{U}^\top) \tag{2.124}$$

$$X \leq_{hr} Y \; (\text{or } \underline{\mathbf{a}} \leq_{hr} \underline{\mathbf{b}}) \Leftrightarrow \begin{bmatrix} \underline{\mathbf{a}}^\top \\ \underline{\mathbf{b}}^\top \end{bmatrix} \mathbf{U} \in TP_2 \tag{2.125}$$

$$X \leq_{rh} Y \Leftrightarrow \ (or\ \underline{\mathbf{a}} \leq_{rh} \underline{\mathbf{b}}) \Leftrightarrow \begin{bmatrix} \underline{\mathbf{a}}^\top \\ \underline{\mathbf{b}}^\top \end{bmatrix} \mathbf{U}^\top \in TP_2 \tag{2.126}$$

$$X \leq_{lr} Y \ (or\ \underline{\mathbf{a}} \leq_{lr} \underline{\mathbf{b}}) \Leftrightarrow \begin{bmatrix} \underline{\mathbf{a}}^\top \\ \underline{\mathbf{b}}^\top \end{bmatrix} \in TP_2. \tag{2.127}$$

The result below is also known as the *basic composition formula.*

Proposition 2.23 — *Suppose* $\mathbf{A} = [a_{ij}]_{i,j \in I\!N}$ *and* $\mathbf{B} = [b_{ij}]_{i,j \in I\!N}$ *are both* TP_2 *matrices. If* $\mathbf{C} = \mathbf{AB}$ *is well defined then* $\mathbf{C} \in TP_2$*.*

The next proposition is due to Shanthikumar (1988, Lemma 2.1), and is referred by Kijima (1997, p. 108) as the *composition law.* See also Lemma 1.2 of Kijima (1992).

Proposition 2.24 — *Let* $\mathbf{A} = [a_{ij}]_{i,j \in I\!N}$ *and* $\mathbf{B} = [b_{ij}]_{i,j \in I\!N}$ *be nonnegative matrices. If* $\mathbf{AU} \in TP_2$*,* $\mathbf{BU} \in TP_2$ *and* $\mathbf{U}^{-1}\mathbf{BU} \geq \mathbf{O}$*, then* $\mathbf{ABU} \in TP_2$*.*

The *dual composition law* also follows similarly to Kijima (1997, p. 110).

Proposition 2.25 — *Let* $\mathbf{A} = [a_{ij}]_{i,j \in I\!N}$ *and* $\mathbf{B} = [b_{ij}]_{i,j \in I\!N}$ *be nonnegative matrices. If* $\mathbf{AU}^\top \in TP_2$*,* $\mathbf{BU}^\top \in TP_2$ *and* $(\mathbf{U}^\top)^{-1}\mathbf{BU}^\top \geq \mathbf{O}$*, then* $\mathbf{ABU}^\top \in TP_2$*.*

In Chapter 8 we briefly discuss the role that some of these ageing properties of the run length may have in the stochastic comparison of geometric and phase-type run lengths.

Chapter 3

Monotonicity and run length

Woodall (2000) notes that, according to Pearson (1967), the more mathematical treatment of the statistical performance of control schemes began in Great Britain after a visit there by Walter A. Shewhart in 1932. In this thesis we are clearly pursuing this line of thought by aiming to obtain qualitative statements about quantities which, for a given control scheme(s), describe its (their) behaviour as completely as possible. The quantities considered so far are related to a *FPT*, the run length.

In the previous chapter we realized that, for the simplest schemes (Shewhart-type), we are able to obtain simple expressions for the *RL* related measures and — through the behaviour of the parameter $1-\pi(\theta)$ of the geometric-type *RL*, i.e., the probability that a signal is triggered at each sample — give straight answers for questions concerning, e.g., the assessment of the impact of a change in a design or model parameter on the performance of a control scheme.

However, Markov-type control schemes lead to formulae that, although easy to handle numerically, require some mathematical work on the special features of the matrix analogue of $\pi(\theta)$, $\mathbf{Q}(\theta)$, to provide answers for questions such as the one in the previous paragraph. These answers will be stated as monotonicity properties, which are among the most important qualitative properties of stochastic models, as noted by Stoyan (1983, p. 39).

Early results concerning the monotonicity behaviour of *FPT*s in terms of the initial value can be found in two papers, and immediately derived from a third one:[1]

- *Kamae, Krengel and O'Brien (1977).* These authors prove in Corollary 1 (ii) of Theorem 2 that, for a monotone Markov chain (in the usual sense) $\{S_N, N \in I\!N_0\}$, with either discrete or continuous state space, the first passage time $T^u = min\{N : S_N > x | S_0 = u\}$ stochastically decreases with u in the usual sense.

  Particular cases of this result are given in Morais and Pacheco (1998b). This reference provides an induction-based proof of a result that implies that the *RL* of some upper one-sided *EWMA* and *CUSUM* control schemes is stochastically decreasing in

[1] Some of the original results are stated in a more general form than the one presented here.

the head-start value.[2]

Brown and Chaganty (1983) used Theorem 2 of Kamae, Krengel and O'Brien (1977)[3] to prove that if the state space of $\{S_N, N \in I\!N_0\}$ has a minimum u_0, then T^{u_0} is a NBU random variable.

- *Lee and Lynch (1997).* To prove the *IHR* ageing character of the continuous *FPT*s, $T^u = min\{t : S_t > x | S_0 = u\}, u = 0, 1, \ldots, x$, these authors state and prove that this random variable decreases with u, in the hazard rate sense, in case the underlying discrete-state non-homogeneous continuous-time Markov chain, $\{S_t, t \in I\!R_0^+\}$, is stochastically monotone in the reversed hazard rate sense.

- *Karlin (1964).* This is an account of the connections between totally positive of order r transition probability matrices/kernels and the TP_r and RR_r characters of several *FPT*s of one-dimensional Markov chains in discrete/continuous time with discrete/continuous state space.

  In the light of result (i) of Theorem 2.1 of this paper we sucessfully manage to prove a stochastic decreasing behaviour of T^u in the likelihood ratio sense.

Kalmykov (1969) also deserves a few comments although it does not provide a stochastic monotonicity result in the initial state in any of the senses we mentioned earlier. This paper provides estimates to

$$P(\hat{\tau}_\xi[s, \{a(t), t \in T\}] < min\{z, \hat{\tau}_\xi[s, \{b(t), t \in T\}]\} | \xi_s = x) \tag{3.1}$$

where: $\{\xi_t, t \in T\}$ is a real one-dimensional Markov process with $T \subset I\!R$; $\hat{\tau}_\xi[s, \{c(t), t \in T\}]$ denotes the first time after s ($s \in T$) at which a sample trajectory of the process intersects the curve $\{c(t), t \in T\}$. This quantity arises in hypothesis testing. In fact, if $\{a(t), t \in T\}$ is the boundary separating the domain of acceptance of hypothesis H_0 from the domain of indifference and $\{b(t), t \in T\}$ ($a(t) < b(t)$, for all $t \in T$) does exactly the same for hypothesis H_1, then (3.1) represents the probability that H_0 will be accepted up to time z considering observations from time s on at point x. (Thus, if we consider $b(t) = +\infty$, (3.1) would be the distribution function of the run length of an one-sided scheme at time z^-.) Theorem 2.1 of this reference asserts that the probability in (3.1) can be bounded by a similar probability involving another Markov process

$$P(\hat{\tau}_{\xi'}[s, \{a(t), t \in T\}] < min\{z, \hat{\tau}_{\xi'}[s, \{b(t), t \in T\}]\} | \xi'_s = x') \tag{3.2}$$

with $x \le x'$. This holds under certain conditions which reminds us of a restricted version — in time and in the state space — of the stochastic order in the usual sense between two probability matrices. For further details see Kalmykov (1969).

Effort has also been made to find stochastic monotonicity results with respect to other parameters of *FPT*s (e.g. θ). Our investigations seem to indicate a lack of results probably due to the fact that the *FPT*s considered in the literature are (usually) parameter-free.

[2]That result also leads to a monotone behaviour of the *RL* in terms of the magnitude of the shift in the parameter being monitored by the control scheme.

[3]Brown and Chaganty (1983) incorrectly mention that they use Theorem 1 of Kamae, Krengel and O'Brien (1977).

In order to provide monotonicity results concerning the run length we have to tackle important issues like ordering Markov chains. This is what lies in the next section.

3.1 Ordering Markov chains

We begin this section with an abbreviated presentation of some useful ordering notions for stochastic processes like Markov chains.

Let $\underline{\mathbf{x}} = (x_1, \ldots, x_m)$ and $\underline{\mathbf{x}}' = (x'_1, \ldots, x'_m)$ be two vectors in $I\!R^m$; then we write $\underline{\mathbf{x}} \leq \underline{\mathbf{x}}'$ if $x_i \leq x'_i, i = 1, \ldots, m$. Additionally, recall that $U \subseteq I\!R^m$ is called an upper set if $\underline{\mathbf{x}}' \in U$ whenever $\underline{\mathbf{x}} \leq \underline{\mathbf{x}}'$ and $\underline{\mathbf{x}} \in U$.

Definition 3.1 — *Let $\underline{\mathbf{X}}$ and $\underline{\mathbf{X}}'$ be two $m-$dimensional random vectors. Then $\underline{\mathbf{X}}$ is said to be smaller than $\underline{\mathbf{X}}'$ in the usual sense — $\underline{\mathbf{X}} \leq_{st} \underline{\mathbf{X}}'$ — in case*

$$P(\underline{\mathbf{X}} \in U) \leq P(\underline{\mathbf{X}}' \in U) \tag{3.3}$$

for every upper set U in $I\!R^m$.

Shaked and Shanthikumar (1994, pp. 114-115) also state that $\underline{\mathbf{X}} \leq_{st} \underline{\mathbf{X}}'$ iff the inequality

$$E[\phi(\underline{\mathbf{X}})] \leq E[\phi(\underline{\mathbf{X}}')] \tag{3.4}$$

holds for every increasing function ϕ on $I\!R^m$ for which the two expectations exist.

Now we are able to introduce a notion of stochastic ordering between two univariate stochastic processes in discrete time.

Definition 3.2 — *Let $\{X_N, N \in I\!N_0\}$ and $\{X'_N, N \in I\!N_0\}$ be two stochastic processes with state space E. Then $\{X_N, N \in I\!N_0\}$ is said to be smaller than $\{X'_N, N \in I\!N_0\}$ in the usual sense if*

$$(X_{N_1}, \ldots, X_{N_m}) \leq_{st} (X'_{N_1}, \ldots, X'_{N_m}) \tag{3.5}$$

for every $m \in I\!N$ and $(N_1, \ldots, N_m) \in I\!N_0^m$.

Definition 4.1.2 from Stoyan (1983, p. 59) provides a weaker order relation between two stochastic processes, as it only requires that $X_N \leq_{st} X'_N$ for all $N \in I\!N_0$.

The next proposition gives a condition that leads to an ordering of stochastic processes.

Proposition 3.3 — *The stochastic order relation in the usual sense in Definition 3.2 is equivalent to*

$$E[g(\underline{\mathbf{X}})] \leq E[g(\underline{\mathbf{X}}')] \tag{3.6}$$

for every increasing functional g for which the expectations in (3.6) exist.[4]

[4]Recall that a functional g is called increasing if $g(\{x_N, N \in I\!N_0\}) \leq g(\{x'_N, N \in I\!N_0\})$ whenever $x_N \leq x'_N$, for $N \in I\!N_0$ (see Shaked and Shanthikumar (1994, p. 124)).

The next theorem, whose simple proof is omitted, relates the survival function of the run length with an increasing functional of a Markov chain and proves to be crucial in the proof of some monotonicity properties in the usual sense.

Theorem 3.4 — *Let $\{S_N, N \in I\!N_0\}$ be a discrete time Markov chain with transient states $\{0, \ldots, x\}$ and absorbing state $x+1$. The survival function of the first passage time*

$$RL^u = \min\{N \in I\!N_0 : S_N = x+1 \mid S_0 = u\}, \tag{3.7}$$

evaluated at m, can be written in terms of the number of visits made by the Markov chain $\{S_N, N \in I\!N_0\}$ to the absorbing state $x+1$ until time m, $\sum_{N=0}^{m} I_{\{x+1\}}(S_N)$. Namely,

$$\begin{aligned} P[RL^u > m] &= 1 - E\left(\min\left\{1, \sum_{N=0}^{m} I_{\{x+1\}}(S_N)\right\}\right) \\ &= 1 - E\left[g(\{S_N, N \in I\!N_0\})\right], \end{aligned} \tag{3.8}$$

where $g(s_0, \ldots, s_m) = \min\left\{1, \sum_{N=0}^{m} I_{\{x+1\}}(s_N)\right\}$ is an increasing functional in $\{0, \ldots, x+1\}$, for $m \in I\!N_0$.

Theorem 2 by Kalmykov (1962) suggests a stochastic ordering between the matrices that govern two discrete time Markov chains with the same state space; it corresponds to what Kulkarni (1995, pp. 148-149) calls the Kalmykov-dominance or Kalmykov order.

Definition 3.5 — *Let $\{S_N, N \in I\!N_0\}$ and $\{S'_N, N \in I\!N_0\}$ be two discrete time Markov chains on $\{0, 1, ..., x+1\}$ with transition probability matrices, $\mathbf{P} = [p_{ij}]_{i,j=0}^{x+1}$ and $\mathbf{P'} = [p'_{ij}]_{i,j=0}^{x+1}$, respectively. Then, $\mathbf{P}$ is said to be smaller than $\mathbf{P'}$ in the usual sense (or in the Kalmykov sense) — $\mathbf{P} \leq_{st} \mathbf{P'}$ — if*

$$\sum_{l=j}^{x+1} p_{il} \leq \sum_{l=j}^{x+1} p'_{ml}, \;\; 0 \leq i \leq m \leq x+1, \;\; 0 \leq j \leq x+1. \tag{3.9}$$

That is, for $N \in I\!N_0$,

$$(S_{N+1} | S_N = i) \leq_{st} (S'_{N+1} | S'_N = m), \;\; 0 \leq i \leq m \leq x+1. \tag{3.10}$$

Remark 3.6 — The inequality (3.10) rightly clarifies the meaning of the Kalmykov order in terms of the conditional distributions of the two Markov chains. It is worth mentioning here that if $\mathbf{P}, \mathbf{P'} \in \mathcal{M}_{st}$ then (3.9) is equivalent to

$$\sum_{l=j}^{x+1} p_{il} \leq \sum_{l=j}^{x+1} p'_{i\,l}, \;\; 0 \leq i \leq x+1, \;\; 0 \leq j \leq x+1. \tag{3.11}$$

It is easy to check that condition (3.11) is equivalent to

$$(S_{N+1} | S_N = i) \leq_{st} (S'_{N+1} | S'_N = i), \;\; 0 \leq i \leq x+1. \tag{3.12}$$

Moreover, $\mathbf{P} \in \mathcal{M}_{st}$ iff $\mathbf{P} \leq_{st} \mathbf{P}$. •

By virtue of Definition 3.5, Theorem 4.B.16 in Shaked and Shanthikumar (1994, p. 125) can be rephrased as in Kulkarni (1995, p. 149) or as in the next proposition.

Proposition 3.7 — *Let* $\mathbf{P}$ *and* $\mathbf{P}'$ *be two transition probability matrices governing the discrete time Markov chains on* $\{0,1,...,x+1\}$, $\{S_N, N \in I\!N_0\}$ *and* $\{S'_N, N \in I\!N_0\}$, *respectively. If* $S_0 \leq_{st} S'_0$ *and* $\mathbf{P} \leq_{st} \mathbf{P}'$, *then the following stochastic order relation holds*

$$\{S_N, N \in I\!N_0\} \leq_{st} \{S'_N, N \in I\!N_0\}. \tag{3.13}$$

This proposition proves to be useful to assess the monotonicity behaviour of RL^u in u, in the usual sense.

3.2 Monotonicities in the initial value and other parameters

Lucas and Crosier (1982) recommended the use of head start values, that is, a non-zero value for the initial value of (Markov-type) summary statistics.

The rationale of it is as follows: if the process is operating in-control, the summary statistic of the scheme is soon brought to zero (for instance, by the reference value in case a *CUSUM* scheme is at use), so that the expected effect of the head start is minimal; otherwise, the operator is alerted to the out-of-control situation much sooner, which may prevent start-up problems.

These authors only assess the influence of the adoption of a head start numerically. In this section we assess it stochastically, in the usual, hazard rate and likelihood ratio senses.

The next theorem yields new comparisons and also unifies some existing results. It should be noted that this theorem could also have been casted in the framework of more general *FPT*s and, therefore, provide more generality to the results. In addition, it shows that the stochastic monotone character of the Markov chain is crucial to guarantee the decreasing behaviour of the RL with respect to the initial value of the summary statistic of the control scheme.

Theorem 3.8 — *Let* $\{S_N, N \in I\!N_0\}$ *be an absorbing Markov chain on* $\{0,1,\ldots,x+1\}$ *with transition matrix* $\mathbf{P}$ *and absorbing state* $x+1$, *and define* $RL^u = \min\{N : S_N = x+1 \mid S_0 = u\}$, $u = 0,1,\ldots,x$. *Then, the following stochastic implications of the adoption of a head start hold:*

$$\mathbf{P} \in \mathcal{M}_{st} \Rightarrow RL^u \downarrow_{st} \text{ with } u \tag{3.14}$$

$$\mathbf{P} \in \mathcal{M}_{rh} \Rightarrow RL^u \downarrow_{hr} \text{ with } u \tag{3.15}$$

$$\mathbf{P} \in \mathcal{M}_{lr} \Rightarrow RL^u \downarrow_{lr} \text{ with } u. \tag{3.16}$$

We prove the stochastic monotonicity results (3.14)–(3.16) separately.

Proof (3.14) — We provide an alternative proof to the one presented by Morais and Pacheco (1998b); it closely follows the one presented in Morais and Pacheco (2000b).

Let $\{S^*_N, N \in I\!N_0\}$ and $\{S'_N, N \in I\!N_0\}$ be two absorbing Markov chains, with initial values $S^*_0 = u^*$ and $S'_0 = u'$ such that $0 \leq u^* \leq u' \leq x$, both governed by the stochastically monotone matrix $\mathbf{P}$. Since $\mathbf{P} \in \mathcal{M}_{st}$ it follows that $\mathbf{P} \leq_{st} \mathbf{P}$. If we add to this the relation order $S_0^* = u^* \le_{st} u' = S_0'$ and use Proposition 3.7, we conclude that $\{S_N^*, N \in I\!N_0\} \le_{st} \{S_N', N \in I\!N_0\}$.[5] Thus, Proposition 3.3 and Theorem 3.4 let us assert that, with

$$g(s_0, \ldots, s_m) = \min\left\{1, \sum_{N=0}^{m} I_{\{x+1\}}(s_N)\right\} \tag{3.17}$$

for $m \in I\!N$ and $s_0, s_1, \ldots, s_m \in I\!R$,

$$\begin{aligned} P(RL^u > m) &= 1 - E\left[g(\{S_N^*, N \in I\!N_0\})\right] \\ &\ge 1 - E\left[g(\{S_N', N \in I\!N_0\})\right] \\ &= P(RL^{u'} > m), \end{aligned} \tag{3.18}$$

which proves the result. •

Proof (3.15) — Let $\{S_N^*, N \in I\!N_0\}$ and $\{S_N', N \in I\!N_0\}$, as well as u^* and u', be as in the previous proof but now with $\mathbf{P} \in \mathcal{M}_{rh}$. In addition, consider: $\underline{\mathbf{u}}_0^* = \underline{\mathbf{e}}_{u^*}$, $\underline{\mathbf{u}}_0' = \underline{\mathbf{e}}_{u'}$; $\underline{\mathbf{u}}_m^* = \underline{\mathbf{u}}_{m-1}^* \mathbf{Q}$ and $\underline{\mathbf{u}}_m' = \underline{\mathbf{u}}_{m-1}' \mathbf{Q}$ for $m \in I\!N$, where $\mathbf{Q}$ is the substochastic matrix in the block representation (2.74) of $\mathbf{P}$.

Note that, for $m \in I\!N$, $\underline{\mathbf{u}}_{m-1}^* \underline{\mathbf{1}} = P(RL^{u^*} \ge m)$, $\underline{\mathbf{u}}_{m-1}' \underline{\mathbf{1}} = P(RL^{u'} \ge m)$, $\underline{\mathbf{u}}_{m-1}^*(\mathbf{I} - \mathbf{Q})\underline{\mathbf{1}} = P(RL^{u^*} = m)$ and $\underline{\mathbf{u}}_{m-1}'(\mathbf{I} - \mathbf{Q})\underline{\mathbf{1}} = P(RL^{u'} = m)$.

Now, remember that since $\mathbf{P} \in \mathcal{M}_{rh}$ we have $\mathbf{P}\mathbf{U}^\top \in TP_2$, which in view of Proposition 3.7 implies $(\mathbf{U}^\top)^{-1}\mathbf{P}\mathbf{U}^\top \ge \mathbf{O}$. As a consequence we get $\mathbf{Q}\mathbf{U}^\top \in TP_2$ and $(\mathbf{U}^\top)^{-1}\mathbf{Q}\mathbf{U}^\top \ge \mathbf{O}$ according to Lemma 2.17. Moreover, by virtue of the fact that $0 \le u^* \le u' \le x$, we have $\underline{\mathbf{u}}_0^* \le_{rh} \underline{\mathbf{u}}_0'$, i.e.,

$$\begin{bmatrix} \underline{\mathbf{u}}^{*\top}_0 \\ \underline{\mathbf{u}}'^{\top}_0 \end{bmatrix} \mathbf{U}^\top \in TP_2, \tag{3.19}$$

according to result (2.126) from Proposition 2.22. By the dual of the composition law (Proposition 2.25) and using induction, we conclude that

$$\begin{bmatrix} \underline{\mathbf{u}}^{*\top}_{m-1} \\ \underline{\mathbf{u}}'^{\top}_{m-1} \end{bmatrix} \mathbf{U}^\top \in TP_2, \tag{3.20}$$

for all $m \in I\!N$. Applying the fact that the entries of $(\mathbf{I} - \mathbf{Q})\,\underline{\mathbf{1}}$ increase with i (i.e. $(\mathbf{U}^\top)^{-1}(\mathbf{I} - \mathbf{Q})\,\underline{\mathbf{1}} \le \underline{\mathbf{O}}$) to result (3.20) we get, for any $m \in I\!N$,

$$\begin{aligned} \begin{bmatrix} \underline{\mathbf{u}}^{*\top}_{m-1} \\ \underline{\mathbf{u}}'^{\top}_{m-1} \end{bmatrix} \mathbf{U}^\top \in TP_2 \quad &\Rightarrow \quad \frac{\underline{\mathbf{u}}^{*\top}_{m-1}}{\underline{\mathbf{u}}^{*\top}_{m-1}\underline{\mathbf{1}}} \le_{rh} \frac{\underline{\mathbf{u}}'^{\top}_{m-1}}{\underline{\mathbf{u}}'^{\top}_{m-1}\underline{\mathbf{1}}} \\ &\Rightarrow \quad \frac{\underline{\mathbf{u}}^{*\top}_{m-1}}{\underline{\mathbf{u}}^{*\top}_{m-1}\underline{\mathbf{1}}} \le_{st} \frac{\underline{\mathbf{u}}'^{\top}_{m-1}}{\underline{\mathbf{u}}'^{\top}_{m-1}\underline{\mathbf{1}}} \\ &\Rightarrow \quad \frac{\underline{\mathbf{u}}^{*\top}_{m-1}\mathbf{U}^\top}{\underline{\mathbf{u}}^{*\top}_{m-1}\underline{\mathbf{1}}} \ge \frac{\underline{\mathbf{u}}'^{\top}_{m-1}\mathbf{U}^\top}{\underline{\mathbf{u}}'^{\top}_{m-1}\underline{\mathbf{1}}} \\ &\Rightarrow \quad \frac{\underline{\mathbf{u}}^{*\top}_{m-1}\mathbf{U}^\top(\mathbf{U}^\top)^{-1}(\mathbf{I} - \mathbf{Q})\,\underline{\mathbf{1}}}{\underline{\mathbf{u}}^{*\top}_{m-1}\underline{\mathbf{1}}} \end{aligned}$$

[5]This result corresponds to Theorem 4.B.17 in Shaked and Shanthikumar (1994, p. 126), which is stated without a proof.

$$\leq \frac{\underline{\mathbf{u}}'^{\top}_{m-1}\mathbf{U}^{\top}(\mathbf{U}^{\top})^{-1}(\mathbf{I}-\mathbf{Q})\,\underline{\mathbf{1}}}{\underline{\mathbf{u}}'^{\top}_{m-1}\underline{\mathbf{1}}}$$

$$\Leftrightarrow \quad \frac{P(RL^{u^*}=m)}{P(RL^{u^*}\geq m)} \leq \frac{P(RL^{u'}=m)}{P(RL^{u'}\geq m)}$$

$$\Leftrightarrow \quad \lambda_{RL^{u^*}}(m) \leq \lambda_{RL^{u'}}(m). \tag{3.21}$$

That is, $RL^{u^*} \geq_{hr} RL^{u'}$, $0 \leq u^* \leq u' \leq x$. •

Proof (3.16) — A simplified version of result (i) of Theorem 2.1 of Karlin (1964) reads as follows: if $\mathbf{P} \in TP_2$ (equivalently, $\mathbf{P} \in \mathcal{M}_{lr}$) then $P(RL^u = m)$ is RR_2 in u and m,[6] i.e.,

$$P(RL^{u^*}=m-1)\times P(RL^{u'}=m) \leq P(RL^{u^*}=m)\times P(RL^{u'}=m-1) \tag{3.22}$$

for $0 \leq u^* \leq u' \leq x$ and $m \in I\!N$. This inequality is equivalent to

$$\frac{P(RL^{u^*}=m-1)}{P(RL^{u^*}=m)} \leq \frac{P(RL^{u'}=m-1)}{P(RL^{u'}=m)},\ m \in I\!N \Leftrightarrow$$

$$r_{RL^{u^*}}(m) \leq r_{RL^{u'}}(m),\ m \in I\!N \Leftrightarrow$$

$$RL^{u^*} \geq_{lr} RL^{u'}, \tag{3.23}$$

for $0 \leq u^* \leq u' \leq x$. That is, the run length decreases with u in the likelihood ratio sense. •

Remark 3.9 — An induction-based proof of the following result

$$(\mathbf{U}^{\top})^{-1}\mathbf{Q}\mathbf{U}^{\top} \geq \mathbf{O} \Rightarrow RL^u \downarrow_{st} \text{ with } u \tag{3.24}$$

can be found in Morais and Pacheco (1998b). Please note that, in this reference, instead of condition $(\mathbf{U}^{\top})^{-1}\mathbf{Q}\mathbf{U}^{\top} \geq \mathbf{O}$ we have the equivalent condition (see (2.72)): $\sum_{l=0}^{j} p_{i\,l}$ decreases with i in $\{0,1,\ldots,x\}$, for fixed $j = 0,1,\ldots,x$.

Capitalizing on the closure property of the usual stochastic order under mixtures[7] we can generalize result (3.14) to the (fairly unusual) case of a random head start.

Corollary 3.10 — *Let $\underline{\alpha}$ and $\underline{\alpha}'$ represent the probability vectors of two random initial values (which eventually depend on θ), and $RL^{\underline{\alpha}}$ and $RL^{\underline{\alpha}'}$ be the corresponding run lengths. If $\mathbf{P} \in \mathcal{M}_{st}$ and $\underline{\alpha} \leq_{st} \underline{\alpha}'$ then*

$$RL^{\underline{\alpha}} \geq_{st} RL^{\underline{\alpha}'}. \tag{3.25}$$

Remark 3.11 — We are not able to strengthen this result to the $\leq_{hr}$ and $\leq_{lr}$ orderings, due to a limitation of these two orderings: they have not the property of being simply closed under mixtures, according to pages 18 and 33 of Shaked and Shanthikumar (1994).

[6] The original result refers to a TP_r discrete time Markov chain.

[7] Theorem 1.A.3 (d) of Shaked and Shanthikumar (1994, p. 6) reads as follows. Let X, Y and U be random variables such that $(X|U=u) \leq_{st} (Y|U=u)$ for all u in the support of U. Then $X \leq_{st} Y$.

It is worth noting that both orders have the closure property under quite strong conditions that do not hold in the RL setting, namely: given the random variables X, Y — both conditioned on U —, $X \leq_* Y$ if $(X|U = u) \leq_* (Y|U = u')$, for all u and u' in the support of U, where $* = hr, lr$ (see Theorems 1.B.8 and 1.C.10 of Shaked and Shanthikumar (1994, pp. 18 and 33)). •

So far we have assessed the stochastic consequences of adopting a head start. A few other possibilities include studying the stochastic implication of:

- an increase of the magnitude of the shift;
- a change in parameters, such as, the reference value in a *CUSUM* scheme, the range of the decision interval, etc.

To proceed with the discussion of these issues we have to resume the initial notation that includes an argument in the Markov chain and its transition matrix, and so on. Assume ρ is one of the parameters mentioned in the previous paragraph, whose influence in the run length we are concerned with.

Theorem 3.12 — *Let $\{S_N(\rho), N \in I\!N_0\}$ and $\{S_N(\rho'), N \in I\!N_0\}$ be two absorbing Markov chains characterized as follows: they have the same state space; their initial values are $S_0(\rho) = u$ and $S_0(\rho') = u'$ with $u \leq u'$; and they are ruled by the transition matrices $\mathbf{P}(\rho)$ and $\mathbf{P}(\rho')$, respectively. Then*

$$\mathbf{P}(\rho) \leq_{st} \mathbf{P}(\rho') \Rightarrow RL^u(\rho) \geq_{st} RL^{u'}(\rho'). \tag{3.26}$$

Proof — Since $S_0(\rho) \leq_{st} S_0(\rho')$ and $\mathbf{P}(\rho) \leq_{st} \mathbf{P}(\rho')$, Proposition 3.7 allows us to conclude that $\{S_N(\rho), N \in I\!N_0\} \leq_{st} \{S_N(\rho'), N \in I\!N_0\}$. Thus,

$$\begin{aligned} P[RL^u(\rho) > m] &= 1 - E\left[g(\{S_N(\rho), N \in I\!N_0\})\right] \\ &\geq 1 - E\left[g(\{S_N(\rho'), N \in I\!N_0\})\right] \\ &= P[RL^{u'}(\rho') > m], \end{aligned} \tag{3.27}$$

where g is the increasing functional defined in Theorem 3.4. •

As a consequence of the previous theorem and the closure property of the $\leq_{st}$ ordering under mixtures, we obtain the following result.

Corollary 3.13 — *If $S_0(\rho) \leq_{st} S_0(\rho')$ and $\mathbf{P}(\rho) \leq_{st} \mathbf{P}(\rho')$ then $RL(\rho) \geq_{st} RL(\rho')$.*

Another consequence of Theorem 3.12 is the following corollary, which can be found in Morais and Pacheco (1998b) along with an induction-based proof.

Corollary 3.14 — *If $\mathbf{P}(\rho) \in \mathcal{M}_{st}$ for all ρ and the left partial sums of $\mathbf{Q}(\rho)$, $\sum_{l=0}^{j} p_{i\,l}(\rho)$, decrease with ρ for every $i, j = 0, \ldots, x$ — i.e., all the entries of $\mathbf{Q}(\rho)\mathbf{U}^\top$ decrease with ρ — then $RL^u(\rho) \downarrow_{st}$ with ρ.*

A few problems that arise when we try to strengthen stochastic property (3.26) to the $\leq_{hr}$ and $\leq_{lr}$ orderings ought to be discussed.

Multivariate stochastic orders are throughly discussed by Shaked and Shanthikumar (1994, pp. 113-152), namely: the usual multivariate stochastic order (Definition 3.2); the cumulative hazard order; the multivariate hazard order (Shaked and Shanthikumar (1987)); the multivariate likelihood ratio order (which is closely related to multivariate totally positivity of order 2 introduced and studied by Karlin and Rinott (1980)); the multivariate mean residual life order; and a few other multivariate stochastic orders.

However, as far as we have investigated, the orderings $\leq_{hr}$, $\leq_{rh}$ and $\leq_{lr}$ have not been proposed to Markov chains or transition matrices. This is a major drawback since a strong ordering of performance measures like RL would "naturally" require a strong ordering of the Markov chains.

Following the relation between the usual multivariate stochastic order and the ordering of transition matrices in the usual sense (Definition 3.5), we would consider ordering two univariate Markov chains in discrete time in a similar manner: $\{X_N, N \in I\!N_0\}$ is said to be smaller than $\{X'_N, N \in I\!N_0\}$ in the $* - sense$ if

$$(X_{N_1}, \ldots, X_{N_m}) \leq_* (X'_{N_1}, \ldots, X'_{N_m}) \tag{3.28}$$

for every $m \in I\!N$ and $(N_1, \ldots, N_m) \in I\!N_0^m$. A totally different possibility would be to define

$$\mathbf{P} \leq_* \mathbf{P'} \Leftrightarrow (X_N \mid X_{N-1} = i) \leq_* (X'_N \mid X'_{N-1} = m),\ i \leq m, \tag{3.29}$$

and state that $\{X_N, N \in I\!N_0\}$ is stochastically smaller than $\{X'_N, N \in I\!N_0\}$ in the $*-$sense if $\mathbf{P} \leq_* \mathbf{P'}$.

Nevertheless, effort have been made to set sufficient conditions under which $RL^u(\rho) \downarrow_*$ with ρ, for $* = hr, lr$, using mathematical induction and two other approaches described as follows.

The multivariate likelihood ratio ordering is preserved under strictly monotone transformations of each individual coordinate of the underlying random vectors, and is closed under marginalization and conjunctions (Shaked and Shanthikumar (1994, p. 133)). These consequences have no bearing whatsoever in the proof of the result $RL^u(\rho) \downarrow_{lr}$ with ρ, under the assumption of stochastically monotone transition probability matrices (in the likelihood ratio sense), related as in Equation (3.29) with $* = lr$. A similar problem holds for the multivariate hazard rate order which has even less properties.

Effort has also been made to establish the results using a technique similar to the one used to prove result (3.15). This approach requires an analogue of the composition law (and its dual) involving two substochastic matrices, $\mathbf{Q}(\rho)$ and $\mathbf{Q}(\rho')$.

As a consequence, stronger versions of (3.26) are not stated.

3.3 Results in the continuous state space case

The Markov approach was originally introduced by Brook and Evans (1972) as an alternative to the integral equation method proposed by Van Dobben de Bruyn (1968) (see also Crowder (1987a)) for approximating the *RL* related measures of *CUSUM* schemes designed to control any continuous quality characteristic. The publication of Lucas and Saccucci (1990) had a tremendous impact in the widespread application of the Markov approach. Its popularility stems from the way that it leads to very elegant recurrence relations for some *RL* related measures (such as the probability and distribution functions), and therefore may be computationally implemented using only few operations.

When we are dealing with continuous data the Markov chain approach begins by approximating the problem — through the discretization of the summary statistic state space (see Example 3.15) —, and then by obtaining the exact solution to the approximate problem. On the other hand, the integral equation approach starts with the exact problem and finds an approximate solution to it.[8]

More recently, Luceño and Puig-Pey (2000) proposed a fast and accurate approximation algorithm for obtaining *RL* related quantities for any continuous data.

The next example shows how the Markov approach can be used to produce approximations to *RL* related measures of an upper one-sided *CUSUM* scheme for the standard deviation (σ) of a normally distributed quality characteristic.

Example 3.15 — Let $(X_{1N}, \ldots, X_{nN})$ be the random sample of size n at the sampling period N taken from a normal distribution with nominal mean and standard deviation μ_0 and σ_0. An increase in σ is measured by $\theta = \sigma/\sigma_0 \geq 1$.

The summary statistic of the upper one-sided *CUSUM* scheme for σ depends on θ and can be found, such as presented below, in Gan (1995):

$$Z_N = \begin{cases} u, N = 0 \\ \max\{0, Z_{N-1} + \ln(S_N^2) - k\}, N \in I\!N \end{cases} \tag{3.30}$$

where: u belong to the decision interval $[LCL, UCL) = [0, h)$; $S_N^2 = (n-1)^{-1} \sum_{i=1}^{n} (X_{iN} - \overline{X}_N)^2$ represents the sample variance at the sampling period N; and, given θ, $(n-1)S_N^2/(\theta\sigma_0)^2$ has a χ^2_{n-1} distribution.

This summary statistic is associated to $\{V_N, N \in I\!N_0\}$, a discrete time, absorbing and homogeneous Markov chain with continuous state space $[LCL, UCL]$, defined as follows: $V_0 = u$ and

$$V_N = \begin{cases} UCL, \text{ if } V_{N-1} + \ln(S_N^2) - k \geq UCL \text{ or } V_{N-1} = UCL \\ \max\{0, V_{N-1} + \ln(S_N^2) - k\}, \text{ otherwise.} \end{cases} \tag{3.31}$$

The initial distribution $P(V_0 \leq z) = I_{[u,\infty)}(z)$ and the transition probability function,

$$K(z; y) = P(V_N \leq z \mid V_{N-1} = y), \tag{3.32}$$

of this homogeneous Markov chain uniquely determine its behaviour. This function equals

$$K(z; y) = \begin{cases} 0, & z < LCL \\ F_{\chi^2_{n-1}} \left[(n-1) \exp(z - y + k)/(\theta\sigma_0)^2 \right], & LCL \leq z < UCL \\ \overline{F}_{\chi^2_{n-1}} \left[(n-1) \exp(UCL - y + k)/(\theta\sigma_0)^2 \right], & z \geq UCL, \end{cases} \tag{3.33}$$

[8] We note in passing that Champ and Rigdon (1991) show the following result: if the product midpoint rule is used to approximate the integral equation, then both approaches yield the same approximations for the *ARL*.

for $y \in [LCL, UCL)$, and is given by $K(z; y) = I_{[UCL,\infty)}(z)$, for $y = UCL$.

To discretize the continuous state space of the Markov chain $\{V_N, N \in I\!N_0\}$ it suffices to divide the decision interval in $(x+1)$ sub-intervals with equal range $\Delta = (UCL - LCL)/(x+1)$, $[e_i, e_{i+1}),\ i = 0, 1, \ldots, x$ where x is a positive integer and

$$e_i = LCL + i \times \Delta,\ i = 0, 1, \ldots, x. \tag{3.34}$$

Each sub-interval $[e_i, e_{i+1})$ is then associated to transient state i of an approximating absorbing Markov chain with *discrete state space* $\{0, 1, \ldots, x+1\}$, $\{V_N(x), N \in I\!N_0\}$. According to this discretization the real value y, $y \in [LCL, UCL)$, is associated to state

$$e(y; x) = \left\lfloor \frac{y - LCL}{UCL - LCL} \times (x+1) \right\rfloor; \tag{3.35}$$

and $y = UCL$ is associated to the absorbing state $x+1$. Furthermore, the initial value of the approximating Markov chain $\{V_N(x), N \in I\!N_0\}$ is equal to $e(u; x)$, which is the state related to $V_0 = u,\ u \in [LCL, UCL)$.

This approximating Markov chain is ruled by the transition matrix $\mathbf{P}(x) = [p_{ij}(x)]_{i,j=0}^{x+1}$ where:

$$\begin{aligned} p_{ij}(x) &= P\left(e_j \leq V_N < e_{j+1} \mid V_{N-1} = \frac{e_i + e_{i+1}}{2}\right) \\ &= K\left(e_{j+1}; \frac{e_i + e_{i+1}}{2}\right) - K\left(e_j; \frac{e_i + e_{i+1}}{2}\right) \end{aligned} \tag{3.36}$$

for $i, j = 0, 1, \ldots, x$, with

$$K\left(e_l; \frac{e_i + e_{i+1}}{2}\right) = F_{\chi^2_{n-1}}\left(\frac{n-1}{\theta^2 \sigma_0^2} \times \exp\left\{k + \frac{h[l - (i+1/2)]}{x+1}\right\}\right), \tag{3.37}$$

for $l = 0, 1, \ldots, x$. Since $\mathbf{P}(x)$ can be partitioned as in Equation (2.53) we have:

$$p_{i\ x+1}(x) = 1 - \sum_{l=0}^{x} p_{i\,l}(x),\ i = 0, 1, \ldots, x \tag{3.38}$$

$$p_{x+1\,j}(x) = I_{\{x+1\}}(j),\ i = 0, 1, \ldots, x+1. \tag{3.39}$$

Thus, the exact run length of this scheme is approximated by a phase-type distribution involving the substochastic matrix $\mathbf{Q}(x) = [p_{i\,j}(x)]_{i,j=0}^{x}$; and the approximations of all the RL related measures in Table 2.3 are obtained by replacing $\mathbf{Q}(\theta)$ with $\mathbf{Q}(x)$. •

Note that, although having proposed the Markov approach in the case of continuous quality characteristics, Brook and Evans (1972) did not prove that $\{V_N(x), N \in I\!N_0\}$ converges in law to $\{V_N, N \in I\!N_0\}$ as $x \to \infty$, or that the approximating run length,

$$RL^{e(u;x)}(x) = \min\{N : V_N(x) = UCL \mid V_0(x) = e(u; x)\}, \tag{3.40}$$

converges in law to the exact run length,

$$RL^u = \min\{N : Z_N \geq UCL \mid Z_0 = u\}. \tag{3.41}$$

For the sake of completeness these two results are stated and proved below. (Analogous convergence results for all schemes in Chapters 5 and 6 follow quite similarly.)

Lemma 3.16 — The following convergence results hold, as $x \to \infty$:

$$\{V_N(x), N \in I\!N_0\} \longrightarrow_{st} \{V_N, N \in I\!N_0\} \tag{3.42}$$

$$RL^{e(u;x)}(x) \longrightarrow_{st} RL^u, \tag{3.43}$$

for all $u \in [LCL, UCL)$, where "$\longrightarrow_{st}$" denotes convergence in distribution.

Proof — Before proving (3.42) and (3.43), recall the following result involving the integer part of a specific real number:

$$\lim_{x\to\infty} \lfloor c(x+1)/d \rfloor \times d/(x+1) = c, \tag{3.44}$$

for all $c \in I\!R$ and $d \neq 0$.

Since the initial distribution and the transition probability function uniquely define the Markov chain $\{V_N, N \in I\!N_0\}$, to show (3.42) we must prove that

$$\lim_{x\to\infty} P[V_0(x) \leq e(z;x)] = P(V_0 \leq z), \tag{3.45}$$

$$\lim_{x\to\infty} P[V_N(x) \leq e(z;x) | V_{N-1}(x) = e(y;x)] = K(z;y), N \in I\!N, \tag{3.46}$$

for every continuity point $z \in [LCL, UCL]$. Thus, (3.42) follows since, using (3.44),

$$\begin{aligned} &\lim_{x\to\infty} P[V_0(x) \leq e(z;x)] \\ &= \lim_{x\to\infty} P\left[V_0 \leq LCL + \frac{e(z;x)(UCL-LCL)}{x+1}\right] \\ &= \lim_{x\to\infty} P\left[V_0 \leq LCL + \left\lfloor \frac{(z-LCL)\times(x+1)}{UCL-LCL} \right\rfloor \times \frac{UCL-LCL}{x+1}\right] \\ &= P(V_0 \leq z), \end{aligned} \tag{3.47}$$

and, in addition, for $N \in I\!N$,

$$\begin{aligned} &\lim_{x\to\infty} P[V_N(x) \leq e(z;x) | V_{N-1}(x) = e(y;x)] \\ &= \lim_{x\to\infty} P\left[V_N \leq LCL + \frac{e(z;x)(UCL-LCL)}{x+1} \right| \\ &\qquad\qquad\qquad \left. V_{N-1} = LCL + \frac{[e(y;x)+1/2](UCL-LCL)}{x+1}\right] \\ &= \lim_{x\to\infty} P\left[V_N \leq LCL + \left\lfloor \frac{(z-LCL)\times(x+1)}{UCL-LCL} \right\rfloor \times \frac{UCL-LCL}{x+1} \right| \\ &\qquad \left. V_{N-1} = LCL + \left\{ \left\lfloor \frac{(y-LCL)\times(x+1)}{UCL-LCL} \right\rfloor + \frac{1}{2} \right\} \times \frac{UCL-LCL}{x+1}\right] \\ &= P(V_N \leq z | V_{N-1} = y) \\ &= K(z;y). \end{aligned} \tag{3.48}$$

The convergence result (3.43) concerning the approximation of the run length follows immediately from (3.42) and Theorem 3.4. •

A question that immediately springs to mind is: "Do the monotonicity properties of the approximating run length $RL^{e(u;x)}(x)$ still hold for the exact run length RL^u when we are dealing with schemes for continuous data?" The answer is affirmative because the stochastic orderings $\leq_{st}$, $\leq_{hr}$, $\leq_{rh}$ and $\leq_{lr}$ have what Stoyan (1983, p. 2) calls the weak convergence property — i.e. these orderings are closed under the limit operation as referred by the next proposition.

Proposition 3.17 — *Let $\{X_i, i \in I\!N\}$ and $\{Y_i, i \in I\!N\}$ be two sequences of positive integer value random variables such that:*

- $X_i \leq_* Y_i, i \in I\!N$, *where* $*$ *stands for the orderings* $\leq_{st}$, $\leq_{hr}$, $\leq_{rh}$ *and* $\leq_{lr}$;
- $X_i \longrightarrow_{st} X$ *and* $Y_i \longrightarrow_{st} Y$, *as* $i \to \infty$.

Then $X \leq_* Y$.

Proof — Proposition 3.17 corresponds to Proposition 1.2.3 of Stoyan (1983, p. 6) and Theorem 1.A.3c) of Shaked and Shanthikumar (1994, p. 6) when $* = st$. The three remaining cases are not stated in those two references. The proofs of all these results are quite similar and are inspired by the one presented by Stoyan (1983, p. 6). Thus, we only present the proof corresponding to $* = hr$.

Let $\{F_i(x), i \in I\!N\}$ and $\{G_i(x), i \in I\!N\}$ be the sequences of distribution functions associated with $\{X_i, i \in I\!N\}$ and $\{Y_i, i \in I\!N\}$. Under the conditions of the proposition we have, for every $i \in I\!N$:

$$\frac{1 - F_i(y)}{1 - F_i(y-1)} \leq \frac{1 - G_i(y)}{1 - G_i(y-1)}, y \in I\!N, \tag{3.49}$$

for the ordering $\leq_{hr}$. Moreover, $F_i(y)$ and $G_i(y)$ converge to $F(y)$ and $G(y)$, for all continuity points of $F(y)$ and $G(y)$.

By virtue of these facts we conclude that

$$\frac{1 - F(y)}{1 - F(y-1)} = \lim_{i\to\infty} \frac{1 - F_i(y)}{1 - F_i(y-1)} \leq \lim_{i\to\infty} \frac{1 - G_i(y)}{1 - G_i(y-1)} = \frac{1 - G(y)}{1 - G(y-1)}, \tag{3.50}$$

for every continuity point y of F and G.[9]

Now take a discontinuity point y of F and G and admit that

$$\frac{1 - F(y)}{1 - F(y-1)} > \frac{1 - G(y)}{1 - G(y-1)} \tag{3.51}$$

for the ordering hr. Since the discontinuity points are at most countable there exists a sequence of continuity points $\{w_m(y), m \in I\!N\}$ of F and G for which $w_m(y) \downarrow y$ as $m \to \infty$.

[9]Note that in this case $y - 1$ is also a continuity point.

Therefore, by the right-continuity property of F and G:

$$\begin{aligned}\frac{1-F(y)}{1-F(y-1)} &= \lim_{m\to\infty} \frac{1-F[w_m(y)]}{1-F[w_m(y-1)]} \\ &\leq \lim_{m\to\infty} \frac{1-G[w_m(y)]}{1-G[w_m(y-1)]} = \frac{1-G(y)}{1-G(y-1)},\end{aligned} \tag{3.52}$$

whence a contradiction. •

Lemma 3.18 — If $RL^{e(u;x)}(x)$ satisfies a stochastic monotonicity property (in the st, hr, rh or lr sense) so will RL^u.

Proof — The results can be inferred from a direct application of the stochastic convergence of the Markov approximation and Proposition 3.17. •

As a final remark we would like to add that Morais and Pacheco (1998c, 1999b) applied result (3.14) and Corollary 3.13 to first passage times/performance measures arising in reliability theory, queueing systems and quality control.

Chapter 4

Combined CUSUM–Shewhart schemes for binomial data

Improving the detection speed of a small, a moderate or a large upward shift in the expected number of defective items can be done in several ways. For instance:

- *By changing the summary statistic* — Upper one-sided *CUSUM* and modified[1] *EWMA* schemes tend to give earlier indication of small and moderate increases in process parameters when compared to the classical upper one-sided *Shewhart* control schemes. See, for example, Gan (1990a, 1993) for *binomial* data, and Lucas (1985), Gan (1990b) and White, Keats and Stanley (1997) in the case of *Poisson* data.

- *By adopting variable sampling intervals* (*VSI*) — The efficiency of a control scheme may be enhanced by relating the sampling interval to the observed values of the summary statistic: for instance, using a longer sampling interval if the sample statistic is close to the target and a shorter one otherwise. Results on the advantages of using *VSI* policies have been presented, for example, by: Reynolds Jr. *et al.* (1988) for the $\bar{X}$-scheme; Shobe (1988) for count data control schemes; Reynolds Jr., Amin and Arnold (1990) for *CUSUM* control schemes; Vaughan (1993) for *binomial* data; Morais and Natário (1998) for upper one-sided c-schemes; and Ramalhoto and Morais (1998, 1999) for the scale parameter of a Weibull control variable.[2]

- *By changing the decision rule* — We can also use two or more observations of a summary statistic to come up with a non standard decision rule that is better (in some sense) than the standard decision rule which is based on only one observation. Several supplementary runs rules have been suggested in the literature. They may

[1] This adjective has here quite a different meaning from the one in the papers by W. Schmid and his collaborators.

[2] Note that Vaughan proposes another version of the *VSI* policy for *binomial* data, which is far more complex in terms of operationality than the *VSI* policy described in the previously mentioned papers: a) if a sample reveals more than c defective items the state of the process is investigated and possibly corrected; the sampling interval in this case is equal to the time required for the process to produce k_R units of output; b) if d ($d \leq c$) defectives are observed, production is allowed to continue, and the next sample is begun after k_d units of output have been produced.

be stated in the following way: an out-of-control signal is triggered if k of the last m observed values of the summary statistic fall in a specific interval defined by what is commonly known as the warning lines.[3] As shown by Champ and Woodall (1987) for normal data, these non standard decision rules lead to *Shewhart* schemes capable of providing a quick detection of a change in the process parameter (although they do not bring their performance up to *CUSUM* levels) at the cost of increasing the false alarm rate.

Nelson (1997) suggests a runs test for a $np-$scheme to signal the occurrence of a special cause indicating that p has decreased. This test is particularly advantageous when there is no lower control limit.

- *By adopting combined upper one-sided* CUSUM–Shewhart *control schemes* — Adding *Shewhart* limits is probably the simplest possible modification of a *CUSUM* scheme. Also considerable advantage is to be gained by using the combination of these two types of schemes since it takes advantage of two well known facts: the *Shewhart* schemes are favored when a large shift has occured, and the *CUSUM* schemes allows a fast detection of small and moderate shifts. Numerical results concerning the combined upper one-sided *CUSUM–Shewhart* schemes for *Poisson* data can be found in Yashchin (1985), Abel (1990) and Morais and Pacheco (1999a).

The statistical research in the area of combined *CUSUM–Shewhart* control schemes seems to have been initiated by

- *Westgard et al. (1977).* These authors propose the use of this sort of scheme for singular observations as an easier alternative to a standard *CUSUM* control scheme with a *V-mask*, which had not been readily accepted in clinical chemistry laboratories.

This paper was soon followed by:

- *Lucas (1982).* In this paper Lucas proposes a combined two-sided *CUSUM–Shewhart* scheme. In any case this work is restricted to the control of increases and decreases of the mean of a normally distributed quality control characteristic, and to the comparison of non-matched *CUSUM* and combined *CUSUM–Shewhart* schemes.

- *Yashchin (1985).* This reference not only considers combined one-sided and two-sided *CUSUM–Shewhart* schemes for normal data, but also addresses the use of combined *CUSUM–Shewhart* control schemes for count data as in Abel (1990).[4] The count data considered had, in both cases, a *Poisson* distribution.

As for applications (other than industrial quality control) of these combined schemes, we mention the following papers:

[3]According to Lai (1974), the use of warning lines was proposed by Dudding and Jennett (1944).

[4]One of the major contributions of Yashchin (1985) is the proposed design procedure: it is not based in the *ARL* behaviour as in Lucas (1982). This procedure is based, for instance, in the *RL* survival function behaviour in the on-target situation.

- *Blacksell et al. (1994).* It is a mere application of a combined *CUSUM–Shewhart* scheme for the mean of normally distributed data — as described by Westgard *et al.* (1977) — to the control of results from an enzyme-linked immunoabsorbent assay (*ELISA*) for the typing of foot and mouth disease virus (*FMVD*) antigen in the routine laboratory diagnosis of *FMVD*.

- *Gibbons (1999).* This author starts his presentation by referring that the combined *CUSUM–Shewhart* schemes have been routinely applied to detection monitoring programs at waste disposal facilities. According to the author, this follows by virtue of the fact that this combined scheme is the only statistical procedure that is directly recommended for use in intra-well monitoring by U.S. EPA (1989, 1992).

As a final note we mention two other papers referring to combined *Shewhart–Cumulative Score* schemes for controlling the process mean of a continuous process production: Ncube and Woodall (1984) and Amin and Ncube (1991).

This review of closely related work gives the distinct idea that combined *CUSUM–Shewhart* quality control schemes have mainly been concerned with monitoring continuous variables and *Poisson* data. In fact, as far as we have investigated, we have not seen any work on combined *CUSUM–Shewhart* control scheme for *binomial* data, although the advantages of this combined approach would seem to follow straightforwardly for this kind of data, as reported in the following sections.

4.1 Description of the combined scheme

The upper one-sided *binomial CUSUM* scheme is planned to detect increases from the nominal value np_0 to a larger expected value of defectives count $n(p_0+\theta)$, $\theta \in \Theta\backslash\{0\}$ where $\Theta = [0, 1-p_0)$. This control scheme makes use of a summary statistic similar to the one considered in Example 2.7:

$$\mathrm{Z}_N(\theta) = \begin{cases} u, \ N = 0 \\ \max\{0, Z_{N-1}(\theta) + Y_N(\theta) - k\}, \ N > 0, \end{cases} \tag{4.1}$$

Recall that: u represents the initial value given to the summary statistic; $Y_N(\theta) = \sum_{i=1}^{N} X_{iN}(\theta)$ is of course the total number of defectives in the N^{th} sample; given θ, the $Y_N(\theta)$'s are *i.i.d.* to $Y(\theta) \sim bin(n, p_0+\theta)$; k is the so called reference value.

The usual practice in maintaining an upper one-sided *CUSUM* control scheme for this type of data (and certainly for other types of atribute and for continuous data) is to trigger a signal as soon as $Z_N(\theta)$ exceeds the upper control limit x. However, it seems reasonable that the upper one-sided *CUSUM* scheme gives a signal at the first time N such that

$$Z_N(\theta) > x \quad \text{or} \quad Z_N(\theta) - Z_{N-1}(\theta) > y. \tag{4.2}$$

The reason for this stems from the fact that a large increment in the summary statistic, say exceeding y $(0 \leq y \leq x)$, should also be taken as an indication of a shift in the process parameter, even if the summary statistic does not exceed the upper control limit.

The second condition in (4.2) will be responsible for a signal at time N when

$$Z_N(\theta) - Z_{N-1}(\theta) = Y_N(\theta) - k > y, \tag{4.3}$$

or when

$$Y_N(\theta) > y + k \ \text{ if } \ Z_{N-1}(\theta) = 0. \tag{4.4}$$

Therefore, the combined decision rule (4.2) corresponds to the simultaneous use of standard upper one-sided control schemes of types *CUSUM* and *Shewhart*, with summary statistics $Z_N(\theta)$ and $Y_N(\theta)$, respectively. This combined scheme states that the process is out of control at time N if

$$Z_N(\theta) > x \ \text{ or } \ Y_N(\theta) > y + k. \tag{4.5}$$

Now we illustrate the performance of an upper one-sided *binomial CUSUM* control scheme before and after combining it with an upper one-sided *Shewhart* scheme (i.e. an upper one-sided $np-$scheme), using a simulated data set and a control scheme, both taken from Gan (1993, pp. 453-4).

Example 4.1 — Table 4.2 has the observed defective counts, that is, the observed values of the summary statistic of an upper one-sided $np-$scheme. The first 50 observations were drawn when the process was operating at the nominal mean level $np_0 = 100 \times 0.05$. The next 20 observations were taken from the same process after a shift to $n(p_0 + \theta) = 100 \times (0.05 + 0.006)$.

Table 4.1: Observed values of the *binomial CUSUM* statistic with: $n = 100$; $p = p_0 = 0.05$, for $N = 1, \dots, 50$; and $p = p_0 + \theta = 0.056$, for $N = 51, \dots, 70$.

N	z_N	N	z_N	N	z_N	N	z_N	N	z_N	N	z_N	N	z_N
1	0	11	8.1	21	0.20	31	5.30	41	12.40	51	**7.50**	61	**19.60**
2	4.71	12	7.81	22	0.91	32	5.01	42	9.11	52	**7.21**	62	**23.31**
3	4.42	13	7.52	23	2.62	33	4.72	43	11.82	53	**8.92**	63	**23.02**
4	10.13	14	5.23	24	2.33	34	6.43	44	10.53	54	**12.63**	64	**20.73**
5	6.84	15	3.94	25	3.04	35	10.14	45	10.24	55	**11.34**	65	**21.44**
6	7.55	16	2.65	26	4.75	36	9.85	46	12.95	56	**12.05**	66	**24.15**
7	4.26	17	5.36	27	7.46	37	12.56	47	13.66	57	**15.76**	67	**22.86**
8	6.97	18	4.07	28	5.17	38	13.27	48	14.37	58	**17.47**	68	**23.57**
9	9.68	19	5.78	29	5.88	39	13.98	49	10.08	59	**18.18**	69	**22.28**
10	8.39	20	1.49	30	4.59	40	13.69	50	7.79	60	**18.89***	70	**22.99**

* first out-of-control signal

The observed values of the *CUSUM* statistic $Z_N(\theta)$ can be found in Table 4.1, for the reference value $k = 5.29$ and the initial value $u = 0$ (that is, no head start ($0\%HS$) has been given to the scheme). The upper control limit of the *CUSUM* scheme is equal to $UCL_C = x = 18.3$. Moreover, if the critical level for the increments of the *CUSUM* statistic is equal to $y = 3.5$ then the upper control limit of the *Shewhart* scheme is equal to $UCL_S = y + k = 8.79$.

Table 4.2: Observed defective counts y_N with: $n = 100$; $p = p_0 = 0.05$, for $N = 1, \ldots, 50$; and $p = p_0 + \theta = 0.056$, for $N = 51, \ldots, 70$.

N	y_N	N	y_N	N	y_N	N	y_N	N	y_N	N	y_N	N	y_N
1	4	11	5	21	4	31	6	41	4	51	**5**	61	**6**
2	10†	12	5	22	6	32	5	42	2	52	**5**	62	**9**
3	5	13	5	23	7	33	5	43	8	53	**7**	63	**5**
4	11‡	14	3	24	5	34	7	44	4	54	**9***	64	**3**
5	2	15	4	25	6	35	9††	45	5	55	**4**	65	**6**
6	6	16	4	26	7	36	5	46	8	56	**6**	66	**8**
7	2	17	8	27	8	37	8	47	6	57	**9**	67	**4**
8	8	18	4	28	3	38	6	48	6	58	**7**	68	**6**
9	8	19	7	29	6	39	6	49	1	59	**6**	69	**4**
10	4	20	1	30	4	40	5	50	3	60	**6**	70	**6**

† first false alarm; ‡ second false alarm; †† third false alarm

* first out-of-control signal

The upper one-sided *binomial CUSUM* scheme gives an out-of-control signal at observation 60. However, if the combined upper one-sided *binomial CUSUM–Shewhart* scheme described above had been active at the time this data was being collected, the operators or engineers would have been immediately alerted to this out-of-control situation by the 54^{th} observation, i.e. only 4 observations (instead of 10) after the shift had ocurred. Then, they could have been able to assign a cause and improve the control system or remove the source of disturbance earlier.

It is worth mentioning that the upper one-sided *Shewhart* scheme from the combined control scheme was also responsible for 3 false alarms (prior to the ocurrence of the shift).

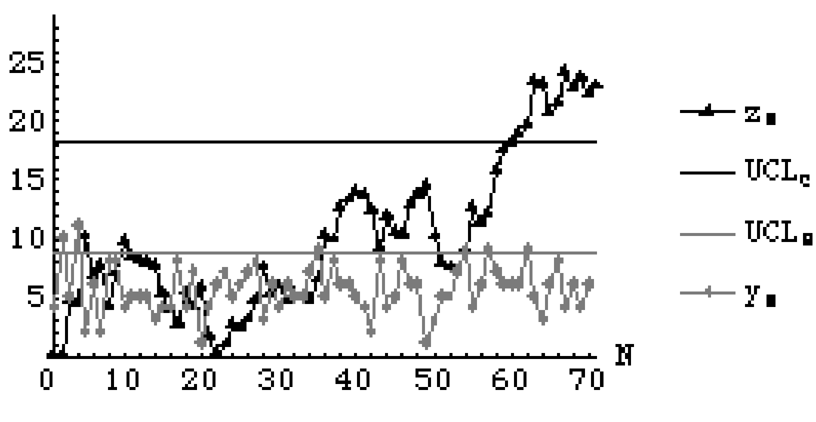

Figure 4.1: Observed values of the *CUSUM* (z_N) and *Shewhart* (y_N) summary statistics.

Both *CUSUM* and *Shewhart* summary statistics can be plotted simultaneously on a single chart and both statistics are interpreted similarly against upper control limits x

and $y+k$ (respectively), thus, minimizing the effort in introducing and maintaining two separate charts, as suggested by Westgard *et al.* (1977) and illustrated in Figure 4.1. •

We close this section remarking that Li and Shaked (1997) stress the importance in reliability of first passage times associated with a discrete time Markovian failure model $\{X_N, N \geq 0\}$ that describes situations where a device fails once its damage level X_N crosses a threshold x or the increment of the damage level process $X_N - X_{N-1}$ exceeds a certain critical value y. This first passage time immediately reminds us the one that follows from the decision rule (4.2) of the combined upper one-sided *CUSUM–Shewhart* schemes. These authors present sufficient conditions (involving the transition matrix of the underlying Markov chain) under which the first passage time described above possesses ageing properties such as *IHR* and *IHRA*.

4.2 RL distribution

Let the *RL*s of the upper one-sided *CUSUM* and the combined upper one-sided *CUSUM–Shewhart* schemes be denoted by $RL^u(x;\theta)$ and $RL^u(x,y;\theta)$. Morever, let $\{Z_N(\theta), N \in I\!N_0\}$ denote the discrete time Markov chain formed by the evolution of the summary statistic of the upper one-sided *CUSUM* scheme.

Now note that both Abel (1990) and Yashchin (1985) consider integer values for the reference value of the combined *CUSUM–Shewhart* scheme for *Poisson* data. However, following Lucas (1985) and Gan (1993), we can consider *CUSUM* schemes for count data with rational reference value k because the Markov approach can still give an exact solution for the *RL* distribution. Thus, if the reference value, the decision interval, the initial value and the critical increment are rational numbers in the reduced form $k = a/b$, $x = c/b$, $u = d/b$, and $y = e/b$ (respectively), $\{Z_N(\theta), N \in I\!N_0\}$ has, after this rescaling, a discrete state space $\{0, 1/b, 2/b, \ldots, c/b, c/b+1, \ldots\}$.

In this setting

$$RL^u(x;\theta) = \min\{N : Z_N(\theta) > x \mid Z_0(\theta) = u\} \tag{4.6}$$

$$\begin{aligned} RL^u(x,y;\theta) &= \min\{N : Z_N(\theta) > x \text{ or } Z_N(\theta) - Z_{N-1}(\theta) > y \mid Z_0(\theta) = u\} \\ &= \min\{N : Z_N(\theta) > x \text{ or } Y_N(\theta) > y+k \mid Z_0(\theta) = u\}. \end{aligned} \tag{4.7}$$

The associated absorbing discrete time Markov chains have discrete state space $\{0, 1/b, 2/b, \ldots, (c-1)/b, c/b, c/b+1\}$ and are denoted by $\{S_N(\theta), N \in I\!N_0\}$ and $\{\overline{S}_N(\theta), N \in I\!N_0\}$, for the upper one-sided *CUSUM* and the upper one-sided combined *CUSUM–Shewhart* schemes, respectively. $\{S_N(\theta), N \in I\!N_0\}$ is defined as in Equation (2.52); and $\{\overline{S}_N(\theta), N \in I\!N_0\}$ is defined as follows: $\overline{S}_0(\theta) = S_0(\theta) = Z_0(\theta)$; and, for $N \in I\!N$

$$\overline{S}_N(\theta) = \begin{cases} Z_N(\theta), \text{ while } \overline{S}_N(\theta) \leq x \text{ and } \overline{S}_N(\theta) - \overline{S}_{N-1}(\theta) \leq y \\ \overline{S}_N(\theta) = x+1, \text{ otherwise.} \end{cases} \tag{4.8}$$

Furthermore $RL^u(x;\theta)$ and $RL^u(x,y;\theta)$ have $\theta-$parameterized phase-type distributions with parameters $(\underline{\mathbf{e}}_u, \mathbf{Q}(\theta))$ and $(\underline{\mathbf{e}}_u, \overline{\mathbf{Q}}(\theta))$, respectively.

For the sake of simplicity the entries of these two substochastic matrices are indexed on $\{0, 1, \ldots, xb\}$, that is, $\mathbf{Q}(\theta) = [p_{ij}(\theta)]_{i,j=0}^{xb}$ and $\overline{\mathbf{Q}}(\theta) = [\overline{p}_{ij}(\theta)]_{i,j=0}^{xb}$. The entries of $\mathbf{Q}(\theta)$ can be written in an alternative way:

$$p_{ij}(\theta) = \sum_{l=0}^{j} p_{il}(\theta) - \sum_{l=0}^{j-1} p_{il}(\theta),\ i, j = 0, 1, \ldots, xb, \tag{4.9}$$

where

$$\sum_{l=0}^{-1} p_{il}(\theta) \quad = \quad 0,\ i = 0, 1, \ldots, xb \tag{4.10}$$

$$\begin{aligned} \sum_{l=0}^{j} p_{il}(\theta) &= P[Z_{N+1}(\theta) \leq j/b \mid Z_N(\theta) = i/b] \\ &= F_{Y(\theta)}[(j-i)/b + k],\ i, j = 0, 1, \ldots, xb. \end{aligned} \tag{4.11}$$

Additionally, the non-zero entries of $\overline{\mathbf{Q}}(\theta)$ are equal to

$$\overline{p}_{ij}(\theta) \quad = \quad \sum_{l=0}^{j} p_{il}(\theta) - \sum_{l=0}^{j-1} p_{il}(\theta), \qquad i = 0, 1, \ldots, xb, \quad j = 0, 1, \ldots, \min\{i + yb, xb\}, \tag{4.12}$$

since any transition corresponding to an increment $Z_{N+1}(\theta) - Z_N(\theta)$ larger than $y = e/b$ is not allowed.[5]

So far we have not discussed the possible values for the critical increment y. This issue is somehow related to the following topic: the relationship between the *RL*s of the upper one-sided *Shewhart* and *CUSUM* schemes, and the combined *CUSUM–Shewhart* scheme.

The upper one-sided *Shewhart* and *CUSUM* schemes are obviously special cases of the combined upper one-sided *CUSUM–Shewhart* scheme. In fact, if $RL_S(y + k; \theta)$ and $RL^u(x; \theta)$ represent the run lengths of the upper one-sided *Shewhart* and *CUSUM* schemes, we have:

$$RL_S(y + k; \theta) =_{st} \lim_{x \to +\infty} RL^0(x, y; \theta) =_{st} RL^0(+\infty, y; \theta) \tag{4.13}$$

$$RL^u(x; \theta) =_{st} RL^u(x, y; \theta),\ y \geq x. \tag{4.14}$$

We can, thus, conclude that y should be always smaller than x, otherwise the performance of the resulting scheme will not be any different from the one of the upper one-sided *CUSUM* scheme.

[5] Equivalently, the entries of the matrix $\overline{\mathbf{Q}}(\theta)$ can be defined in terms of an improper distribution function as in Yashchin (1985): $\overline{p}_{ij}(\theta) = \sum_{l=0}^{j} \overline{p}_{il}(\theta) - \sum_{l=0}^{j-1} \overline{p}_{il}(\theta)$, $i, j = 0, 1, \ldots, xb$, where $\sum_{l=0}^{-1} \overline{p}_{il}(\theta) = 0$, $i = 0, 1, \ldots, xb$, and, for $i, j = 0, 1, \ldots, xb$, $\sum_{l=0}^{j} \overline{p}_{il}(\theta) = F^{+}_{Y(\theta)}[(j-i)/b + k] = F_{Y(\theta)}[(j-i)/b + k]$ for $(j-i)/b \leq y$ and $F_{Y(\theta)}(y + k)$ when $(j-i)/b > y$.

Example 4.2 — Since Example 4.1 makes use of the reference value $k = 5.29$ and the upper control limit $x = 18.3$, and therefore requires the use of a huge number of transient states, $1 + xb = 1831$, we consider a combined upper one-sided *CUSUM–Shewhart* scheme whose *CUSUM* and *Shewhart* components correspond to the schemes described in Example 2.10 and Example 2.4, respectively. Recall that the parameters are: $n = 100$; $p_0 = 0.02$ and $p_1 = 0.0427685$; the reference value and the upper control limit of the *CUSUM* component are equal to $k = 3$ and $x = 6$; and the critical value for the increment is $y = 4$ because the upper control limit of the *Shewhart* component equals $y + k = 7$.

Then $RL^0(6, 4; 0)$ has a discrete phase-type distribution represented by $(\underline{\mathbf{e}}_0, \overline{\mathbf{Q}}(0))$, where

$$\overline{\mathbf{Q}}(0) = \begin{bmatrix} 0.8590 & 0.0902 & 0.0353 & 0.0114 & 0.0031 & 0 & 0 \\ 0.6767 & 0.1823 & 0.0902 & 0.0353 & 0.0114 & 0.0031 & 0 \\ 0.4033 & 0.2734 & 0.1823 & 0.0902 & 0.0353 & 0.0114 & 0.0031 \\ 0.1326 & 0.2707 & 0.2734 & 0.1823 & 0.0902 & 0.0353 & 0.0114 \\ 0 & 0.1326 & 0.2707 & 0.2734 & 0.1823 & 0.0902 & 0.0353 \\ 0 & 0 & 0.1326 & 0.2707 & 0.2734 & 0.1823 & 0.0902 \\ 0 & 0 & 0 & 0.1326 & 0.2707 & 0.2734 & 0.1823 \end{bmatrix}. \tag{4.15}$$

This set of parameters yields to $ARL^0(6, 4; 0) = 603.743$ which represents a 40.6% reduction in the in-control *ARL* of the *CUSUM* component of the combined scheme, $ARL^0(6; 0) = 1015.71$, and a 43.7% reduction in the in-control *ARL* of its *Shewhart* component, $ARL_S(4+3; 0) = 1073.03$. Additionally, $ARL^0(6, 4; p_1 - p_0) = 3.99$, $ARL^0(6; p_1 - p_0) = 5.93$ and $ARL_S(4 + 3; p_1 - p_0) = 15.37$. These results lead us to assert that an upward shift in p will remain unnoticed for less time in average when the combined scheme is adopted at the cost of a higher false alarm rate, as previously noted in Example 4.1. •

In the next sections we shall make a clarification for such matters as the special features of the stochastic matrix $\overline{\mathbf{P}}(\theta)$ of $\{\overline{S}_N(\theta), N \in I\!N_0\}$ in order to investigate some of the monotonicity properties of $RL^u(x, y; \theta)$ and properly answer the following question: "Can we go beyond the numerical comparisons concerning the *ARL* and stochastically compare the performance of the combined upper one-sided *CUSUM–Shewhart* with the two constituent schemes?"

4.3 Some features of the transition matrix and stochastic monotonicities in the initial state

The stochastic matrix $\overline{\mathbf{P}}(\theta)$ results from a simple modification of $\mathbf{P}(\theta)$ and is equal to[6]

$$\begin{bmatrix} p_{0\,0} & p_{0\,1} & \cdots & p_{0\,y} & 0 & \cdots & \cdots & 0 & 1-\sum_{l=0}^{y} p_{0\,l} \\ p_{1\,0} & p_{1\,1} & \cdots & \cdots & p_{1\,y+1} & 0 & \cdots & 0 & 1-\sum_{l=0}^{y+1} p_{1\,l} \\ \vdots & \vdots & \cdots & \cdots & \cdots & \cdots & \cdots & \vdots & \vdots \\ p_{x-y-1\,0} & p_{x-y-1\,1} & \cdots & \cdots & \cdots & \cdots & p_{x-y-1\,x-1} & \mathbf{0} & \mathbf{1}-\sum_{\mathbf{l=0}}^{\mathbf{x-1}} \mathbf{p_{x-y-1\,l}} \\ p_{x-y\,0} & p_{x-y\,1} & \cdots & \cdots & \cdots & \cdots & \cdots & \mathbf{p_{x-y\,x}} & \mathbf{1}-\sum_{\mathbf{l=0}}^{\mathbf{x}} \mathbf{p_{x-y\,l}} \\ p_{x-y+1\,0} & p_{x-y+1\,1} & \cdots & \cdots & \cdots & \cdots & \cdots & p_{x-y+1\,x} & 1-\sum_{l=0}^{x} p_{x-y+1\,l} \\ \vdots & \vdots & \cdots & \cdots & \cdots & \cdots & \cdots & \vdots & \vdots \\ p_{x\,0} & p_{x\,1} & \cdots & \cdots & \cdots & \cdots & \cdots & p_{x\,x} & 1-\sum_{l=0}^{x} p_{x\,l} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}. \tag{4.16}$$

[6]We shall omit θ for space reasons whenever we write down all the entries of $\overline{\mathbf{P}}(\theta)$, $\overline{\mathbf{P}}(\theta)\mathbf{U}^\top$ or $\overline{\mathbf{P}}(\theta)\mathbf{U}$.

As mentioned in Section 2.5 the *binomial* distribution has a decreasing likelihood ratio (DLR). Thus, $\mathbf{P}(\theta) \in \mathcal{M}_{lr}$ according to Proposition 2.20. The next example help us realize that $\overline{\mathbf{P}}(\theta)$ does not inherit this and another special feature of $\mathbf{P}(\theta)$.

Example 4.3 — Take once again the combined scheme described in Example 4.2. The in-control transition matrix that governs the associated absorbing Markov chain equals

$$\overline{\mathbf{P}}(0) = \begin{bmatrix} 0.8590 & 0.0902 & 0.0353 & 0.0114 & 0.0031 & 0 & 0 & 0.0010 \\ 0.6767 & 0.1823 & 0.0902 & 0.0353 & 0.0114 & 0.0031 & \mathbf{0} & \mathbf{0.0010} \\ 0.4033 & 0.2734 & 0.1823 & 0.0902 & 0.0353 & 0.0114 & \mathbf{0.0031} & \mathbf{0.0010} \\ 0.1326 & 0.2707 & 0.2734 & 0.1823 & 0.0902 & 0.0353 & 0.0114 & 0.0041 \\ 0 & 0.1326 & 0.2707 & 0.2734 & 0.1823 & 0.0902 & 0.0353 & 0.0155 \\ 0 & 0 & 0.1326 & 0.2707 & 0.2734 & 0.1823 & 0.0902 & 0.0508 \\ 0 & 0 & 0 & 0.1326 & 0.2707 & 0.2734 & 0.1823 & 0.1410 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}. \tag{4.17}$$

It is clear that $\overline{\mathbf{P}}(0) \notin TP_2$: this rightly follows from the fact that the 2×2 minor in **bold** in the previous equation is negative. Thus, $\overline{\mathbf{P}}(0) \notin \mathcal{M}_{lr}$. Writing down

$$\overline{\mathbf{P}}(0)\mathbf{U} = \begin{bmatrix} 1 & 0.1410 & 0.0508 & 0.0155 & 0.0041 & 0.0010 & 0.0010 & 0.0010 \\ 1 & 0.3233 & 0.1410 & 0.0508 & 0.0155 & 0.0041 & \mathbf{0.0010} & \mathbf{0.0010} \\ 1 & 0.5967 & 0.3233 & 0.1410 & 0.0508 & 0.0155 & \mathbf{0.0041} & \mathbf{0.0010} \\ 1 & 0.8674 & 0.5967 & 0.3233 & 0.1410 & 0.0508 & 0.0155 & 0.0041 \\ 1 & 1 & 0.8674 & 0.5967 & 0.3233 & 0.1410 & 0.0508 & 0.0155 \\ 1 & 1 & 1 & 0.8674 & 0.5967 & 0.3233 & 0.1410 & 0.0508 \\ 1 & 1 & 1 & 1 & 0.8674 & 0.5967 & 0.3233 & 0.1410 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix} \tag{4.18}$$

leads to a similar conclusion: $\overline{\mathbf{P}}(0) \notin \mathcal{M}_{hr}$.

Despite the fact that $\overline{\mathbf{P}}(0) \notin \mathcal{M}_{hr}, \mathcal{M}_{lr}$, all the 2×2 minors of

$$\overline{\mathbf{P}}(0)\mathbf{U}^\top = \begin{bmatrix} 0.8590 & 0.9492 & 0.9845 & 0.9959 & 0.9990 & 0.9990 & 0.9990 & 1 \\ 0.6767 & 0.8590 & 0.9492 & 0.9845 & 0.9959 & 0.9990 & 0.9990 & 1 \\ 0.4033 & 0.6767 & 0.8590 & 0.9492 & 0.9845 & 0.9959 & 0.9990 & 1 \\ 0.1326 & 0.4033 & 0.6767 & 0.8590 & 0.9492 & 0.9845 & 0.9959 & 1 \\ 0 & 0.1326 & 0.4033 & 0.6767 & 0.8590 & 0.9492 & 0.9845 & 1 \\ 0 & 0 & 0.1326 & 0.4033 & 0.6767 & 0.8590 & 0.9492 & 1 \\ 0 & 0 & 0 & 0.1326 & 0.4033 & 0.6767 & 0.8590 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \tag{4.19}$$

are nonnegative, according to our calculations using *Mathematica*. Therefore $\overline{\mathbf{P}}(0) \in \mathcal{M}_{rh}$.
•

Theorem 4.4 — *The matrix* $\overline{\mathbf{P}}(\theta)$ *has the following properties*

$$\overline{\mathbf{P}}(\theta) \in \mathcal{M}_{st} \tag{4.20}$$

$$Y(\theta) \in DLR \Rightarrow \overline{\mathbf{P}}(\theta) \in \mathcal{M}_{rh} \tag{4.21}$$

$$\overline{\mathbf{P}}(\theta) \notin \mathcal{M}_{hr} \tag{4.22}$$

$$\overline{\mathbf{P}}(\theta) \notin \mathcal{M}_{lr}, \tag{4.23}$$

for $0 < y < x$.

These special features of $\overline{\mathbf{P}}(\theta)$ essentially mean that, if we associate probability functions of discrete random variables to each row of $\overline{\mathbf{P}}(\theta)$, the associated random variables increase stochastically — in the usual sense and in the reversed hazard rate sense in some cases, but not in the hazard rate and likelihood ratio senses —, as we progress in the rows of this stochastic matrix.

Theorem 4.4 holds for *binomial* and *Poisson* data and any other discrete *DLR* distribution. Without loss of generality we consider $b = 1$ in the proof of this theorem.

Proof (4.20) — A direct substitution yields that

$$\begin{aligned} \sum_{l=0}^{j} \overline{p}_{il}(\theta) &= \sum_{l=0}^{\min\{j,\min\{i+y,x\}\}} p_{il}(\theta) \\ &= F_{Y(\theta)}(\min\{j-i, \min\{y, x-i\}\} + k) \end{aligned} \tag{4.24}$$

is a decreasing function of i, for all $j = 0, 1, \ldots, x$, regardless of the distribution of $Y(\theta)$. In the light of this result we can conclude that $\overline{\mathbf{P}}(\theta)$ is stochastically monotone in the usual sense. •

Proof (4.21) — Taking into consideration that $\mathbf{P}(\theta) \in TP_2$ for $Y(\theta) \in DLR$ and that $\overline{\mathbf{P}}(\theta)\mathbf{U}^\top$ equals

$$\begin{bmatrix} p_{0\,0} & \cdots & \sum_{l=0}^{y-1} p_{0\,l} & \sum_{l=0}^{y} p_{0\,l} & \cdots & \cdots & \cdots & \sum_{l=0}^{y} p_{0\,l} & 1 \\ p_{1\,0} & \cdots & \cdots & \sum_{l=0}^{y} p_{1\,l} & \sum_{l=0}^{y+1} p_{1\,l} & \cdots & \cdots & \sum_{l=0}^{y+1} p_{1\,l} & 1 \\ \vdots & \vdots & \cdots & \cdots & \cdots & \cdots & \cdots & \vdots & \vdots \\ p_{x-y-1\,0} & \cdots & \cdots & \cdots & \cdots & \cdots & \sum_{l=0}^{x-1} p_{x-y-1\,l} & \sum_{l=0}^{x-1} p_{x-y-1\,l} & 1 \\ p_{x-y\,0} & \cdots & \cdots & \cdots & \cdots & \cdots & \sum_{l=0}^{x-1} p_{x-y\,l} & \sum_{l=0}^{x} p_{x-y\,l} & 1 \\ p_{x-y+1\,0} & \cdots & \cdots & \cdots & \cdots & \cdots & \sum_{l=0}^{x-1} p_{x-y+1\,l} & \sum_{l=0}^{x} p_{x-y+1\,l} & 1 \\ \vdots & \vdots & \cdots & \cdots & \cdots & \cdots & \cdots & \vdots & \vdots \\ \vdots & \vdots & \cdots & \cdots & \cdots & \cdots & \cdots & \vdots & \vdots \\ p_{x\,0} & \cdots & \cdots & \cdots & \cdots & \cdots & \sum_{l=0}^{x-1} p_{x\,l} & \sum_{l=0}^{x} p_{x\,l} & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \tag{4.25}$$

we only need to prove that

$$\begin{bmatrix} \sum_{l=0}^{y+h} p_{h\,l} & 1 \\ \sum_{l=0}^{y+h+1} p_{h+1\,l} & 1 \end{bmatrix} \in TP_2, h = 0, 1, \ldots, x-y-1 \tag{4.26}$$

and

$$\begin{bmatrix} \sum_{l=0}^{x} p_{h\,l} & 1 \\ \sum_{l=0}^{x} p_{h+1\,l} & 1 \end{bmatrix} \in TP_2, h = x-y, x-y+1 \ldots, x-1 \tag{4.27}$$

(see the proof of Theorem 2.8 of Li and Shaked (1997)). Statement (4.26) follows from

$$\sum_{l=0}^{y+h} p_{h\,l} = \sum_{l=0}^{y+h+1} p_{h+1\,l} = F_Y(y+k), \tag{4.28}$$

due to the spatial homogeneity of the Markov chain ruled by $\mathbf{P}(\theta)$ and to Equation (4.11).[7] On the other hand (4.27) is a consequence of the stochastic monotonicity of $\mathbf{P}(\theta)$. •

Proof (4.22) — Visualizing $\overline{\mathbf{P}}(\theta)\mathbf{U}$,

$$\begin{bmatrix}
1 & 1-p_{0\,0} & \cdots & 1-\sum_{l=0}^{y} p_{0\,l} & \cdots & \cdots & 1-\sum_{l=0}^{y} p_{0\,l} & 1-\sum_{l=0}^{y} p_{0\,l} \\
1 & 1-p_{1\,0} & \cdots & 1-\sum_{l=0}^{y} p_{1\,l} & 1-\sum_{l=0}^{y+1} p_{1\,l} & \cdots & 1-\sum_{l=0}^{y+1} p_{1\,l} & 1-\sum_{l=0}^{y+1} p_{1\,l} \\
\vdots & \vdots & \cdots & \cdots & \cdots & \cdots & \vdots & \vdots \\
1 & 1-p_{x-y-1\,0} & \cdots & \cdots & \cdots & \cdots & \mathbf{1-\sum_{l=0}^{x-1} p_{x-y-1\,l}} & \mathbf{1-\sum_{l=0}^{x-1} p_{x-y-1\,l}} \\
1 & 1-p_{x-y\,0} & \cdots & \cdots & \cdots & \cdots & \mathbf{1-\sum_{l=0}^{x-1} p_{x-y\,l}} & \mathbf{1-\sum_{l=0}^{x} p_{x-y\,l}} \\
1 & 1-p_{x-y+1\,0} & \cdots & \cdots & \cdots & \cdots & 1-\sum_{l=0}^{x-1} p_{x-y+1\,l} & 1-\sum_{l=0}^{x} p_{x-y+1\,l} \\
\vdots & \vdots & \cdots & \cdots & \cdots & \cdots & \vdots & \vdots \\
\vdots & \vdots & \cdots & \cdots & \cdots & \cdots & \vdots & \vdots \\
1 & 1-p_{x\,0} & \cdots & \cdots & \cdots & \cdots & 1-\sum_{l=0}^{x-1} p_{x\,l} & 1-\sum_{l=0}^{x} p_{x\,l} \\
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1
\end{bmatrix}, \quad (4.29)$$

shall help us prove that $\overline{\mathbf{P}}(\theta)\mathbf{U}$ is not a TP_2 matrix.

This result can be immediately inferred if we observe that, from the spatial homogeneity of $\mathbf{P}(\theta)$, the 2×2 minor in **bold**, concerning the two last columns of $\overline{\mathbf{P}}(\theta)\mathbf{U}$ and states $(x-y-1)$ and $(x-y)$, is equal to

$$-p_{x-y\,x} \times \left(1 - \sum_{l=0}^{x-1} p_{x-y-1\,l}\right) < 0. \quad (4.30)$$

•

Proof (4.23) — (4.23) is a consequence of (4.22) in view of (2.70); and alternative proof of (4.23) is as follows.

From (4.16), and also as suggested by Example 4.3, it follows that any 2×2 minor that includes the two last columns of $\overline{\mathbf{P}}(\theta)$ and involves the line associated with state $(x-y)$ and one of the first $(x-y)$ rows of $\overline{\mathbf{P}}(\theta)$ is negative. In fact, if we pick the rows corresponding to states $x-y$ and h $(0 \leq h \leq x-y-1)$ and the last two columns of $\overline{\mathbf{P}}(\theta)$, given by (4.16), the corresponding 2×2 minor is equal to

$$-p_{x-y\,x} \times \left(1 - \sum_{l=0}^{y+h} p_{h\,l}\right) < 0. \quad (4.31)$$

•

Remark 4.5 — We note that (4.23) is also a consequence of Theorem 2.1 of Keilson and Kester (1978), which describes the zero-structure of TP_2 matrices as follows. Let $\mathbf{B}_{m\,n}^{NE}$ be the "northeast" quadrant of a matrix $\mathbf{B} = [b_{ij}]_{i,j=1}^{N,M}$, that is, the submatrix with component indices (i,j) with $i \leq m$, $j \geq n$. Define the "southwest" quadrant of $\mathbf{B}$, $\mathbf{B}_{m\,n}^{SW}$, similarly.

[7] Li and Shaked (1997) used instead the assumption of stochastic convexity of the transition matrix. Recall that a matrix $\mathbf{R} = [r_{i\,j}]_{i,j \in \mathbb{N}_0}$ is stochastically monotone convex if $\sum_{j=k}^{+\infty} r_{i\,j} \leq \sum_{j=k+1}^{+\infty} r_{i+1\,j}, i,k \in \mathbb{N}_0$. According to Remark 2.3 in Li and Shaked (1997) it easy to see that $\mathbf{R}$ is upper triangular. Clearly, we are not dealing with a stochastic matrix with such a property, in the quality control setting.

Theorem 2.1 of Keilson and Kester (1978) asserts that if $\mathbf{B} \in TP_2$ and $b_{m\ n}$ is a zero element in a non-null row and non-null column, then either $\mathbf{B}^{NE}_{m\,n} = \mathbf{O}$, or $\mathbf{B}^{SW}_{m\,n} = \mathbf{O}$. Needless to say that every zero entry above the diagonal of $\mathbf{P}(\theta)$ — take for instance $(m, n) = (h, x), h = 0, 1, \ldots, x - y - 1$ — belongs to a non-null row and a non-null column and has no "northeast" or "southwest" quadrants equal to $\mathbf{O}$. Thus, $\overline{\mathbf{P}}(\theta)$ is not TP_2. •

Under the conditions of Theorem 4.4, a direct application of Theorem 3.8 enables us to state the next theorem.

Theorem 4.6 — *The run length of the upper one-sided combined CUSUM–Shewhart has the following stochastic behaviour in terms of the initial value u of the summary statistic, for discrete data and a critical increment $y < x$:*

$$RL^u(x, y; \theta) \downarrow_{st} \text{ with } u \tag{4.32}$$

$$Y(\theta) \in DLR \Rightarrow RL^u(x, y; \theta) \downarrow_{hr} \text{ with } u. \tag{4.33}$$

Proof — The first result of Theorem 4.6 follows from result (4.20) of Theorem 4.4 and the stochastic implication of the adoption of a head start condensed in result (3.14) of Theorem 3.8. Similarly, (4.33) is a direct consequence of results (4.21) and (3.15), which also refer to Theorem 4.4 and Theorem 3.8, respectively. •

Remark 4.7 — Giving a head start to a combined upper one-sided *CUSUM–Shewhart* scheme leads to:

- more frequent signals within the first m samples (for any value of m), namely, it yields an increase in the chance that the combined scheme will detect a shift immediately after it occurred; and
- an increase of the alarm rate of this scheme at sample m, $\lambda_{RL^u}(m)$, in case the discrete data has a *DLR* distribution.

In practice, false alarms will be more likely to happen as the initial value of the summary statistic grows. In addition, the number of samples taken until detection of an increase in the parameter is stochastically reduced by increasing the initial value of the summary statistic. •

As seen in Theorem 4.4 the transition matrix $\overline{\mathbf{P}}(\theta)$ can be stochastically monotone in reversed hazard rate sense. Therefore, if we apply the result (2.119) of Lemma 2.21, we get the following corollary concerning the monotone character of the hazard rate of $RL^0(x, y; \theta)$.

Corollary 4.8 — *If $Y(\theta) \in DLR$ then $RL^0(x, y; \theta) \in IHR$.*

Since $\overline{\mathbf{P}}(\theta) \notin M_{lr}$, $RL^u(x, y; \theta)$ is bound to have no decreasing behaviour in the likelihood ratio sense. The next example illustrates not only this fact but also the monotone character of the alarm rate function of $RL^0(x, y; \theta)$.

Example 4.9 — Consider $RL^u(6,4;\theta)$, the RL of the upper one-sided combined *CUSUM–Shewhart* scheme for *binomial* data described in Example 4.2.

In Table 4.3 we can find some values of the hazard (or alarm) rate and equilibrium rate functions of four distinct RLs associated with the combined scheme without head start ($RL^0(6,4;0)$ and $RL^0(6,4;p_1-p_0)$), and with a 50% head start ($RL^3(6,4;0)$ and $RL^3(6,4;p_1-p_0)$).

Table 4.3: ARLs, alarm rates and equilibrium rates of $RL^u(6,4;\theta)$, for $u=0,3$.

	$\lambda_{RL}(m)$				$r_{RL}(m)$	
m	$RL^0(6,4;0)$	$RL^3(6,4;0)$	$RL^0(6,4;p_1-p_0)$	$RL^3(6,4;p_1-p_0)$	$RL^0(6,4;0)$	$RL^3(6,4;0)$
1	0.000932	0.004062	0.065065	0.136973	—	—
2	0.001013	0.006201	0.089031	0.272616	0.920552	0.657783
3	0.001206	0.005159	0.157957	0.300351	0.841304	1.209509
4	0.001372	0.003938	0.208567	0.303617	0.880166	1.316624
5	0.001484	0.003063	0.239501	0.301431	0.925239	1.290967
10	0.001648	0.001764	0.284820	0.292443	0.996186	1.042719
20	0.001661	0.001661	0.290638	0.290772	1.001636	1.001887
30	0.001661	0.001661	0.290740	0.290742	1.001663	1.001664
100	0.001661	0.001661	0.290741	0.290741	1.001663	1.001663
ARL	603.743	592.559	193.813	184.854	—	—

If we compare columns 2 and 3, and 4 and 5, we can see that the alarm rate function increases with the initial value of the summary statistic, as stated in Theorem 4.6. These values also show how unlikely (likely) is the emission of a false alarm (of a correct signal) at sample m, given that no previous signal has been triggered.

The use of a 50% head start yields a mild reduction in the in-control ARL and a slightly larger relative reduction in the out-control ARLs, because the right tail behaviour of the in-control RL distribution is practically independent of the head start (see Lucas and Crosier (1982)). However, the alarm rate at the first samples can increase considerably, specially in out-of-control situations, as we can see by comparing columns 4 and 5 of Table 4.3.

Columns 2 and 4 of Table 4.3 illustrate the increasing behaviour of the alarm rate function of $RL^0(6,4;\theta)$ (Corollary 4.8). Thus, signalling, given that no observation has previously exceeded the upper control limit, becomes more likely, as we proceed with the sampling procedure. However, this table and Figure 4.2 enable us to also add that $RL^3(6,4;0), RL^3(6,4;p_1-p_0) \notin IHR$ (see columns 3 and 5).

We should also add that the limiting form of the probability function of the RL is geometric-like with parameter $1-\xi(\theta)$, where $\xi(\theta)$ is the maximum real eigenvalue of $\overline{\mathbf{Q}}(\theta)$ (see Brook and Evans (1972)), regardless of the initial value u of the summary statistic. Thus, it comes as no surprise that the values of the alarm rate functions of $RL^0(6,4;0)$

and $RL^3(6,4;0)$, and $RL^0(6,4;p_1-p_0)$ and $RL^3(6,4;p_1-p_0)$ rapidly converge to

$$\lim_{m\to+\infty} \lambda_{RL^u(6,4;0)}(m) = 1-\xi(0) = 0.001661 \tag{4.34}$$

$$\lim_{m\to+\infty} \lambda_{RL^u(6,4;p_1-p_0)}(m) = 1-\xi(p_1-p_0) = 0.290741, \tag{4.35}$$

respectively, as seen in Table 4.3 and Figure 4.2.

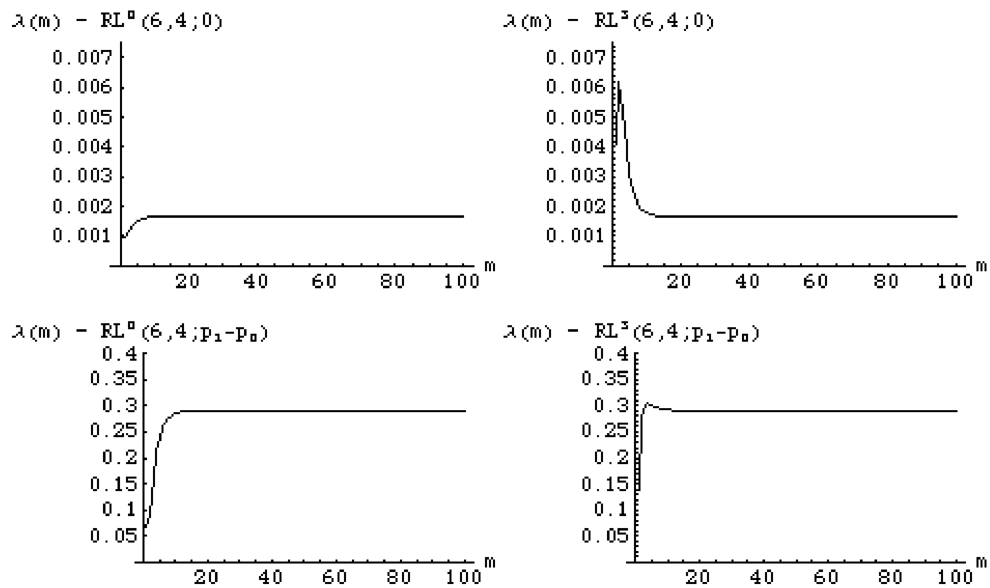

Figure 4.2: Alarm rates of $RL^u(6,4;0)$ and $RL^u(6,4;p_1-p_0)$, for $u=0$ (on the left) and $u=3$ (on the right).

The fact that the equilibrium rate of $RL^3(6,4;0)$, $r_{RL^3(6,4;0)}(m) = P[RL^3(6,4;0) = m-1]/P[RL^3(6,4;0) = m], m = 2,3\ldots$, is not always pointwise above $r_{RL^0(6,4;0)}(m)$ is apparent from the two last columns of Table 4.3. As a consequence we can infer that $RL^3(6,4;0) \not\leq_{lr} RL^0(6,4;0)$. •

4.4 Other stochastic monotonicities

As we have seen in the previous section, adopting a head start HS can lead to a stochastic reduction (in the usual and hazard rate senses) of the run length of the upper one-sided combined *CUSUM–Shewhart* scheme for discrete data.

In this section, we present, interpret (as remarks) and prove several other stochastic properties of combined upper one-sided *CUSUM–Shewhart* schemes, namely the stochastic consequences of the adoption of such a scheme in the performance of upper one-sided *Shewhart* and *CUSUM* control charts.

In the following consider *binomial* data (without any loss of generality), $k \geq 0$, $x > 0$, $0 \leq u \leq x$, $0 \leq y \leq x$ and $\theta \in \Theta = [0, 1-p_0)$.

In addition, we adopt a specific notation: for instance, $RL^u(x, y; \theta, n)$ ($\overline{\mathbf{P}}(\theta; n)$) is used to explicitly declare that the run length (transition matrix) depends on the sample size n and that we are interested in knowing the stochastic implications of changes in n. This also holds for other parameters like k (or x and y, when we refer to the transition matrix).

Theorem 4.10 — *$RL^u(x, y; \theta)$ has the following decreasing stochastic monotonicity properties in the usual sense:*

$$RL^u(x, y; \theta) \downarrow_{st} \text{ with } \theta \tag{4.36}$$

$$RL^u(x, y; \theta, n) \downarrow_{st} \text{ with } n. \tag{4.37}$$

Furthermore, $RL^u(x, y; \theta)$ increases stochastically in the usual sense with k, x and y:

$$RL^u(x, y; \theta, k) \uparrow_{st} \text{ with } k \tag{4.38}$$

$$RL^u(x, y; \theta) \uparrow_{st} \text{ with } x \tag{4.39}$$

$$RL^u(x, y; \theta) \uparrow_{st} \text{ with } y. \tag{4.40}$$

Remark 4.11 — Statements (4.36)-(4.40) are very intuitive; they essentially (and respectively) mean that:

- The control chart stochastically increases its ability to trigger an out-of-control signal as the increase in the parameter becomes more severe.

- In the case of *binomial* data, increasing the sample size corresponds to colecting more information about the process, thus, the chart is more sensitive to upward shifts at the cost of increasing the frequency of false alarms. However, recall that "unnecessarily large samples sizes may result in reaction to small effects of no practical significance", as mentioned by Woodall (2000) when this author refers to Woodall and Faltin (1996).[8]

- An increase in the reference value k implies values of the summary statistic closer to the origin, yielding a control chart that stochastically produces fewer false alarms and larger detection times of upward shifts in the parameter.

- Adopting a larger decision interval leads to stochastically larger RLs. Therefore, the false alarm rate decreases, but an increase of the process mean from its target value will stochastically remain unnoticed for more time.

- Increasing y means a less stringent monitoring program and corresponds to weakning the impact of the *Shewhart* part of the combined scheme. Thus, a stochastic increase of the RL occurs if the critical increment for the summary statistic increases. •

[8] According to Lai (1974), Weiler (1952) discussed the choice of the sample size and pointed out that larger samples than those usually taken in industry should be used when it is important to detect small shifts in the mean, and that small but frequent samples may be adequate for detecting large shifts in the mean.

The stochastic improvement over the performance of the upper one-sided *CUSUM* scheme, derived from the adoption of the combined upper one-sided *CUSUM–Shewhart* scheme, and two other stochastic properties in the usual sense are stated in detail in the next theorem.

Recall that $RL_S(y+k;\theta)$ and $RL^u(x;\theta)$ denote the run lengths of the upper one-sided Shewhart and CUSUM schemes, respectively.

Theorem 4.12 — *The run lengths of the upper one-sided* Shewhart *and* CUSUM *schemes have the following properties:*

$$RL^u(x,y;\theta) \leq_{st} RL_S(y+k;\theta) \tag{4.41}$$

$$RL^u(x,y;\theta) \leq_{st} RL^u(x;\theta), y \leq x \tag{4.42}$$

$$RL^u(x;\theta) \leq_{st} RL^u(x',y;\theta), x' \geq x,\ x \leq y \leq x'. \tag{4.43}$$

Remark 4.13 — The combined upper one-sided *CUSUM–Shewhart* scheme certainly has a larger ability of discriminating increases in the expected value of the defects/defectives count than both its *Shewhart* and *CUSUM* components, leading for instance to smaller *ARL*s. However, the adoption of the combined scheme has an unpleasant disadvantage: it also leads to a stochastic reduction of the *RL* when the production process is in-control, therefore false alarms occur more frequently.

If we desire to compensate the stochastic decrease of the in-control *RL* of the upper one-sided *CUSUM* scheme due to the adoption of the combined upper one-sided *CUSUM–Shewhart* scheme, by using an upper control limit x', larger than x, then we should not forget that we can never improve, stochastically, an upper one-sided *CUSUM* scheme (with upper control limit equal to x), by adopting a combined upper one-sided *CUSUM–Shewhart* scheme with larger decision interval x' and, simultaneously, a critical increment y such that $x \leq y \leq x'$. •

The proofs of Theorems 4.10 and 4.12 are presented below, now with $b \in I\!N$. We need to always have in mind Theorem 3.12 and that we are dealing with stochastically monotone matrices in the usual sense.

Proof ((4.36) and (4.37): monotonicities in θ and n) — These two properties follow trivially from the fact that the random variable $bin(n,p)$ stochastically increases with n and p, in the likelihood ratio sense. Therefore, for $i,j = 0,1,\ldots,xb$,

$$\sum_{l=0}^{j} \overline{p}_{il}(\theta;n) = F_{bin(n,p_0+\theta)}(\min\{j-i,yb\}+kb) \tag{4.44}$$

decreases with θ and n, for *binomial* data. Thus, we have: $\overline{\mathbf{P}}(\theta) \leq_{st} \overline{\mathbf{P}}(\theta')$, $\theta \leq \theta'$; $\overline{\mathbf{P}}(\theta;n) \leq_{st} \overline{\mathbf{P}}(\theta;n')$, $n \leq n'$; and, according to Theorem 3.12, $RL^u(x,y;\theta) \geq_{st} RL^u(x,y;\theta')$, $\theta \leq \theta'$, and $RL^u(x,y;n) \geq_{st} RL^u(x,y;n')$, $n \leq n'$, respectively. •

Proof ((4.38): monotonicity in k) — We can add that $\sum_{l=0}^{j} \overline{p}_{i\,l}(\theta;k)$ increases with the reference value k, that is $\overline{\mathbf{P}}(\theta;k') \le_{st} \overline{\mathbf{P}}(\theta;k), k \le k'$. Applying Theorem 3.12, we get $RL^u(x,y;k') \ge_{st} RL^u(x,y;k)$, $k \le k'$. •

Proof ((4.39): monotonicity in x) — The proof of this stochastic result is slightly more delicate since this property regards two Markov chains with different state spaces, one for each upper control limit.

We start by considering the absorbing Markov chains $\{\overline{S}_N(\theta;x+1/b), N \in I\!N_0\}$ and $\{\overline{S}'_N(\theta), N \in I\!N_0\}$. In one hand, the first absorbing Markov chain is associated with a combined upper one-sided *CUSUM–Shewhart* scheme whose *CUSUM* component has upper control limit $x+1/b$, $\{0,1/b,\ldots,x,x+1/b,x+2/b\}$ as state space, and is governed by $\overline{\mathbf{P}}(\theta;x+1/b)$. On the other hand, $\{\overline{S}'_N(\theta), N \in I\!N_0\}$ results from a slight variation of the absorbing Markov chain $\{\overline{S}_N(\theta;x), N \in I\!N_0\}$; it is also defined on $\{0,1/b,\ldots,x,x+1/b,x+2/b\}$, however, it is ruled by the transition matrix

$$\overline{\mathbf{P}}'(\theta) = \begin{bmatrix} \overline{\mathbf{Q}}(\theta;x) & \underline{\mathbf{0}} & [\mathbf{I}-\overline{\mathbf{Q}}(\theta;x)]\,\underline{\mathbf{1}} \\ \underline{\mathbf{0}}^\top & 0 & 1 \\ \underline{\mathbf{0}}^\top & 0 & 1 \end{bmatrix}. \tag{4.45}$$

Recall that the entries of $\overline{\mathbf{P}}(\theta;x+1/b)$ and $\overline{\mathbf{P}}'(\theta)$ are indexed on $\{0,1,\ldots,xb,xb+1,xb+2\}$. Then note that, for $0 \le u \le x$, $RL^u(x,y;\theta) = \min\{N : \overline{S}_N(\theta;x) > x \text{ or } \overline{S}_N(\theta;x) - \overline{S}_{N-1}(\theta;x) > y\}$ has the same distribution as

$$\min\{N : \overline{S}'_N(\theta) > x+1/b \text{ or } \overline{S}'_N(\theta) - \overline{S}'_{N-1}(\theta) > y\}. \tag{4.46}$$

The simple inspection of $\overline{\mathbf{P}}(\theta;x+1/b)$ and $\overline{\mathbf{P}}'(\theta)$, and the fact that $\overline{\mathbf{P}}'(\theta) \in \mathcal{M}_{st}$ allow us to assert that

$$\sum_{l=0}^{j} \overline{p}_{i\,l}(\theta;x+1/b) \ge \sum_{l=0}^{j} \overline{p}'_{i\,l}(\theta) \ge \sum_{l=0}^{j} \overline{p}'_{m\,l}(\theta), \tag{4.47}$$

for $0 \le i \le m \le xb+2,\ 0 \le j \le xb+2$. Therefore, we conclude that

$$\overline{\mathbf{P}}(\theta;x+1/b) \le_{st} \overline{\mathbf{P}}'(\theta). \tag{4.48}$$

Using Theorem 3.12, we get

$$RL^u(x,y;\theta) \le_{st} RL^u(x+1/b,y;\theta), \tag{4.49}$$

which leads to the stated increasing behaviour in x. •

Proof ((4.40): monotonicity in y) — Consider two absorbing Markov chains governed by $\overline{\mathbf{P}}(\theta) = \overline{\mathbf{P}}(\theta;y)$ and $\overline{\mathbf{P}}'(\theta) = \overline{\mathbf{P}}(\theta;y')$. Due to the definition of the block matrices $\overline{\mathbf{Q}}(\theta) = \overline{\mathbf{Q}}(\theta;y)$ and $\overline{\mathbf{Q}}'(\theta) = \overline{\mathbf{Q}}(\theta;y')$ and the fact that $\overline{\mathbf{P}}(\theta), \overline{\mathbf{P}}'(\theta) \in \mathcal{M}_{st}$, we immediately conclude that, for $0 < y \le y' < xb$,

$$\begin{aligned} \sum_{l=0}^{j} \overline{p}'_{il}(\theta) &= \sum_{l=0}^{j} \overline{p}_{il}(\theta;y') \ge \sum_{l=0}^{j} \overline{p}_{il}(\theta;y) \ge \sum_{l=0}^{j} \overline{p}_{ml}(\theta;y) \\ &= \sum_{l=0}^{j} \overline{p}_{ml}(\theta), 0 \le i \le m \le xb+1,\ 0 \le j \le xb+1. \end{aligned} \tag{4.50}$$

Thus, for $0 < y \leq y' < xb$,

$$\overline{\mathbf{P}}(\theta; y') \leq_{st} \overline{\mathbf{P}}(\theta; y). \tag{4.51}$$

Finally, taking into account Theorem 3.12, we conclude that

$$RL^u(x, y; \theta) \leq_{st} RL^u(x, y'; \theta), \text{ for } 0 < y \leq y' < xb, \tag{4.52}$$

which proves the monotonicity in y. •

Proof ((4.41) and (4.42)) — These results follow by virtue of the monotonicity in terms of x and y (see (4.39) and (4.40), respectively) and the facts $RL_S(y+k; \theta) =_{st} RL^0(+\infty, y; \theta)$ and $RL^u(x; \theta) =_{st} RL^u(x, y; \theta)$, $y \geq x$. •

Proof (4.43) — Since $RL^u(x; \theta) =_{st} RL^u(x, x; \theta)$, we have

$$RL^u(x; \theta) =_{st} RL^u(x, y; \theta) \leq_{st} RL^u(x', y; \theta), \ x \leq y \leq x' \tag{4.53}$$

by using the monotonicity property in the decision interval x, stated in (4.39). •

Remark 4.14 — The stochastic property (4.43) could have been proved in a similar way to property (4.39). In fact, let $\overline{S}(\theta; x', y)$ and $\overline{S}'(\theta)$ be two absorbing Markov chains, both on $\{0, 1/b, \ldots, x, x+1/b, \ldots, x', x'+1/b\}$, where $x' = c'/b \geq x$. These two Markov chains are governed by the probability transition matrices $\overline{\mathbf{P}}(\theta) = \overline{\mathbf{P}}(\theta; x', y)$ and $\overline{\mathbf{P}}'(\theta)$, with this latter matrix given by

$$\begin{bmatrix} \overline{\mathbf{Q}}(\theta; x) & \underline{\mathbf{0}}\ \underline{\mathbf{0}}_{c'-c}^\top & [\mathbf{I} - \overline{\mathbf{Q}}(\theta; x)]\,\underline{\mathbf{1}} \\ \underline{\mathbf{0}}_{c'-c}\ \underline{\mathbf{0}}^\top & \underline{\mathbf{0}}_{c'-c}\ \underline{\mathbf{0}}_{c'-c}^\top & \underline{\mathbf{1}}_{c'-c} \\ \underline{\mathbf{0}}^\top & \underline{\mathbf{0}}_{c'-c}^\top & 1 \end{bmatrix} \tag{4.54}$$

where $\underline{\mathbf{1}}$ and $\underline{\mathbf{1}}'_{c'-c}$ ($\underline{\mathbf{0}}$ and $\underline{\mathbf{0}}'_{c'-c}$) are column vectors of $xb+1$ and $c'-c$ ones (zeroes), respectively. Now, noting that, for $0 \leq u \leq xb$, $RL^u(x, y; \theta)$ and the first passage time

$$\min\{N : \overline{S}'_N(\theta) > x' \text{ or } \overline{S}'_N(\theta) - \overline{S}'_{N-1}(\theta) > y\} \tag{4.55}$$

are equally distributed and that, for $x \leq y \leq x'$, we have

$$\overline{\mathbf{P}}(\theta; x', y) \leq_{st} \overline{\mathbf{P}}'(\theta), \tag{4.56}$$

the stochastic behaviour (4.43) follows from Theorem 3.12.

It is interesting to notice that a stochastic order relation in the usual sense can never be established between the matrices $\overline{\mathbf{P}}(\theta; x', y)$, with $y < x$, and the matrix $\overline{\mathbf{P}}'(\theta)$ defined in (4.54). Therefore, the associated run lengths cannot be related in the usual sense.

The intuition on most of these properties previously discussed is so strong that stating them as stochastic ordering results and proving them could be thought as excessive. However, as seen earlier, the proof of these results can be non trivial and depend — crucially — on the stochastic monotonicity character of the transition matrix. In addition, the extension of all the results to the hazard rate and likelihood ratio senses is still an open problem.

Nevertheless, Table 4.3 and Figure 4.2 suggest that $RL^u(6,4;0) \leq_{hr} RL^u(6,4;p_1-p_0)$, $u=0,3$. Additional numerical results not only suggest that this stochastic order relation holds for other constelations of parameters but also that the result holds in the likelihood ratio sense.

4.5 Comparative assessment of Shewhart, CUSUM and combined schemes: an example

Instead of illustrating the gist of some of the stochastic properties stated earlier on and complement the few existing illustrations in the quality control literature,[9] the foregoing tables and graphs provide an extended example of a comparative assessment of the performance of the following three schemes for *binomial* counts (with $n=100$, $p_0=0.02$, $p_1=0.0427685$):

- upper one-sided *CUSUM* scheme without head start (C^0),
- upper one-sided *Shewhart* scheme (S), and
- combined upper one-sided *CUSUM–Shewhart* scheme with no head start (CS^0),

as described in Example 4.2.

In Table 4.4, there are several values of *ARL*, *SDRL*, *CVRL*, *CSRL*, *CKRL* and *RL* percentage points; this table refers to the three schemes mentioned above.

Table 4.5 display values of the percentage reduction in these *RL* related measures, due to the adoption of the combined scheme (CS^0). For example, the percentage reductions in the *ARL* are

$$\left[1-\frac{ARL^0(x,y;\theta)}{ARL^0(x;\theta)}\right]\times 100\%, \quad \text{and} \quad \left[1-\frac{ARL^0(x,y;\theta)}{ARL_S(y+k;\theta)}\right]\times 100\% \tag{4.57}$$

when we compare the average detection time of the combined scheme CS^0 with the corresponding measure of its *CUSUM* and *Shewhart* constituent charts (respectively). These reductions are defined similarly for the five remaining *RL* related measures.

These two tables illustrate the impact of the adoption of the combined scheme in the detection of small and moderate shifts with magnitude $\theta=0.001, 0.0025, 0.005, 0.0075, 0.01, 0.02, p_1-p_0, 0.03$. The range of values includes the magnitude of the shift in p we want to detect as quickly as possible, $p_1-p_0=0.0227685$.

Using the combined scheme CS^0 virtually yields a 40% reduction in the *RL* of schemes S and C^0, when the production process is in-control. Thus, false alarms are more frequent, as suggested by Table 4.5. For instance, a false alarm occurs in the combined scheme CS^0

[9] In the case of *Poisson* data we note that: Tables IV and VI (V) in Abel (1990) illustrate the monotonicities of $ARL^u(x,y;\theta)$ in terms of x, y and θ (u); and Figure 9 of Yashchin (1985) illustrates the monotonicity of some percentage points of $RL^u(x,y;\theta)$ in terms of θ. Lucas (1982) illustrates through Table 2 (in four parts) the monotonicities of *ARL* of combined schemes for the mean of *normal* data, in terms of x, k, u, $y+k$ and θ (it denotes the positive displacement of the population mean).

Table 4.4: *ARL*, *SDRL*, *CVRL*, *CSRL*, *CKRL* and *RL* percentage points values — listed in order corresponding to schemes C^0, S and CS^0.

	θ								
	0	0.001	0.0025	0.005	0.0075	0.01	0.02	$p_1 - p_0$	0.03
ARL	1015.710	591.724	284.121	102.081	46.227	25.458	7.194	5.932	4.095
	1073.030	787.737	512.346	270.112	154.275	94.128	21.047	15.369	7.815
	603.743	394.192	214.972	87.704	42.275	23.973	6.882	5.648	3.824
SDRL	1012.18	588.012	280.175	97.895	42.022	21.419	4.320	3.322	1.998
	1072.53	787.237	511.846	269.611	153.774	93.627	20.541	14.861	7.298
	601.712	391.853	212.199	84.390	38.724	20.452	4.384	3.418	2.134
CVRL	0.997	0.994	0.986	0.959	0.909	0.841	0.600	0.560	0.488
	1.000	0.999	0.999	0.998	0.997	0.995	0.976	0.967	0.934
	0.997	0.994	0.987	0.962	0.916	0.853	0.637	0.605	0.558
CSRL	2.000	2.000	2.000	1.998	1.989	1.961	1.627	1.523	1.303
	2.000	2.000	2.000	2.000	2.000	2.000	2.001	2.001	2.005
	2.000	2.000	2.000	1.997	1.983	1.947	1.515	1.376	1.087
CKRL	6.000	6.000	5.999	5.992	5.953	5.833	4.296	3.814	2.853
	6.000	6.000	6.000	6.000	6.000	6.000	6.002	6.005	6.019
	6.000	6.000	5.998	5.986	5.931	5.775	3.880	3.289	2.087

RL perc. points	θ								
	0	0.001	0.0025	0.005	0.0075	0.01	0.02	$p_1 - p_0$	0.03
5%	55	34	18	9	6	4	2	2	2
	56	41	27	14	8	5	2	1	1
	33	22	14	8	5	4	2	1	1
25%	295	173	85	32	16	10	4	4	3
	309	227	148	78	45	27	6	5	3
	175	115	64	28	15	10	4	3	2
Median	705	411	198	72	33	19	6	5	4
	744	546	355	187	107	65	15	11	6
	419	274	150	62	31	18	6	5	4
75%	1407	819	392	140	63	34	9	7	5
	1487	1092	710	374	214	130	29	21	11
	836	546	297	120	57	32	9	7	5
90%	2334	1358	649	230	101	53	13	10	7
	2470	1813	1179	621	355	216	48	35	17
	1388	905	491	198	93	51	13	10	7
95%	3036	1765	843	297	130	68	26	12	8
	3214	2359	1534	808	461	281	62	45	22
	1805	1176	638	256	119	65	15	12	8

within the first 419 samples with probability of at least 50%, whereas the median of the run length of scheme C^0 equals 705.

Table 4.5: Percentage reduction in the *ARL*, *SDRL*, *CVRL*, *CSRL*, *CKRL* and *RL* percentage points values due to the adoption of the CS^0 scheme — listed in order corresponding to $(1 - CS^0/C^0) \times 100\%$ and $(1 - CS^0/S) \times 100\%$.

Percentage reduction	θ = 0	0.001	0.0025	0.005	0.0075	0.01	0.02	$p_1 - p_0$	0.03
ARL	40.6	33.4	24.3	14.1	8.5	5.8	4.3	4.8	6.6
	43.7	50.0	58.0	67.5	72.6	74.5	67.3	63.2	51.1
SDRL	40.6	33.4	24.3	13.8	7.8	4.5	-1.5	-2.9	-6.8
	43.9	50.2	58.5	68.7	74.8	78.2	78.7	77.0	70.8
CVRL	0.0	0.0	-0.1	-0.3	-0.8	-1.4	-6.1	-8.1	-14.4
	0.3	0.5	1.2	3.6	8.1	14.2	34.7	37.4	40.2
CSRL	0.0	0.0	0.0	0.1	0.3	0.7	6.9	9.7	16.6
	0.0	0.0	0.0	0.2	0.8	2.7	24.3	31.2	45.8
CKRL	0.0	0.0	0.0	0.1	0.4	1.0	9.7	13.8	26.9
	0.0	0.0	0.0	0.2	1.1	3.8	35.4	45.2	65.3
Percentage reduction	θ = 0	0.001	0.0025	0.005	0.0075	0.01	0.02	$p_1 - p_0$	0.03
5%	40.0	35.3	22.2	11.1	16.7	0.0	0.0	50.0	50.0
	41.1	46.3	48.1	42.9	37.5	20.0	0.0	0.0	0.0
25%	40.7	33.5	24.7	12.5	6.2	0.0	0.0	25.0	33.3
	43.4	49.3	56.8	64.1	66.7	63.0	33.3	40.0	33.3
Median	40.6	33.3	24.2	13.9	6.1	5.3	0.0	0.0	0.0
	43.7	49.8	57.7	66.8	71.0	72.3	60.0	54.5	33.3
75%	40.6	33.3	24.2	14.3	9.5	5.9	0.0	0.0	0.0
	43.8	50.0	58.2	67.9	73.4	75.4	69.0	66.7	54.5
90%	40.5	33.4	24.3	13.9	7.9	3.8	0.0	0.0	0.0
	43.8	50.1	58.4	68.1	73.8	76.4	72.9	71.4	58.8
95%	40.5	33.4	24.3	13.8	8.5	4.4	6.2	0.0	0.0
	43.8	50.1	58.4	68.3	74.2	76.9	75.8	73.3	63.6

Moreover, the adoption of the combined scheme CS^0 yields a benefit of at least 10% — in *ARL*, *SDRL* and the *RL* percentage points considered here — of both constituent charts of such a scheme, for small shifts up to an 25% increase in p (i.e. for $0 < \theta \leq 0.005$). These improvements can be larger than the 40% reduction in the in-control *RL* measures, when we add a *CUSUM* upper control limit to the upper one-sided *Shewhart* scheme — thus, obtaining once again the combined upper one-sided *CUSUM–Shewhart* scheme — for values of θ ranging from 0.001 to 0.01. When we do the opposite, the benefit of supplementing an upper one-sided *CUSUM* scheme with a *Shewhart* upper control limit is always smaller than the 40% reduction we have just mentioned.

The skewness and kurtosis coefficients of the three upper one-sided schemes only differ for moderate shifts, as shown in Table 4.5. Also, according to additional graphs (for

$\theta \in [0, 0.2]$) ommitted here, substituting scheme S by scheme CS^0 always leads to a reduction of these two coefficients. However, they can increase if we replace scheme C^0 with scheme CS^0.

When it comes to $\theta = p_1 - p_0$, we fail to see a reason for advocating the addition of the *Shewhart* upper control limit $y+k = 7$ to the upper one-sided *CUSUM* scheme: the changes in *ARL* and the *RL* percentage points considered here seem to be irrelevant. In fact the results obtained here give the distinct impression that the performance of the *CUSUM* scheme that makes use of the optimal reference value for the detection of an upward shift with magnitude $p_1 - p_0$ (as suggested by Gan (1993)) can be hardly improved.

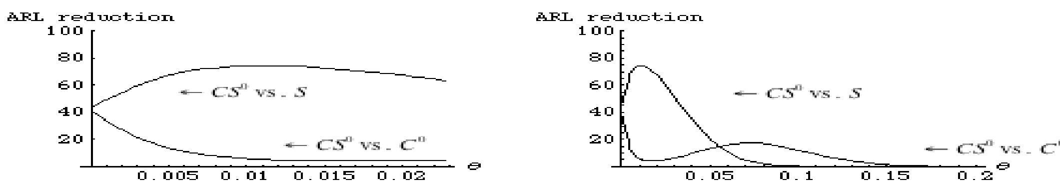

Figure 4.3: *ARL* percentage reductions due to the adoption of the combined scheme CS^0 in the intervals $[0, p_1 - p_0]$ and $[0, 0.2]$.

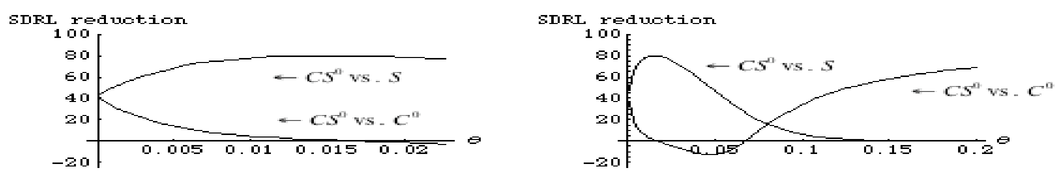

Figure 4.4: *SDRL* percentage reductions due to the adoption of the combined scheme CS^0 in the intervals $[0, p_1 - p_0]$ and $[0, 0.2]$.

Figure 4.3 and Figure 4.4 portray the overall behaviour of the *ARL* and *SDRL* reductions (respectively) in the interval $[0, 0.2]$.

Figure 4.3 confirms what we expected from the combined scheme CS^0. Its average detection speed outperforms substantially the one of its *Shewhart* component in the case of small and moderate shifts. Additionally, the *CUSUM* component — which is unable to detect large shifts faster than the *Shewhart* scheme — is outperformed in the detection of large shifts by the combined scheme.

Figure 4.4 adds that the *SDRL* can increase after replacing scheme C^0 by the combined scheme in the detection of moderate shifts. On the other hand *SDRL* never increases when we use the combined scheme CS^0 by supplementing scheme S with scheme C^0.

Table 4.6: *ARL*, *SDRL*, *CVRL*, *CSRL*, *CKRL* and *RL* percentage points values — listed in order corresponding to schemes C^3, CS^0 and CS^3.

	θ								
	0	0.001	0.0025	0.005	0.0075	0.01	0.02	$p_1 - p_0$	0.03
ARL	995.070	574.634	270.937	93.044	39.670	20.475	4.920	3.991	2.710
	603.743	394.192	214.972	87.704	42.275	23.973	6.882	5.648	3.824
	592.559	383.773	205.790	80.490	36.661	19.560	4.870	3.960	2.699
SDRL	1011.980	587.776	279.885	97.517	41.574	20.930	3.879	2.918	1.684
	601.712	391.853	212.199	84.390	38.724	20.452	4.384	3.418	2.134
	601.585	391.686	211.963	84.032	38.260	19.916	3.826	2.885	1.671
CVRL	1.017	1.023	1.033	1.048	1.048	1.022	0.788	0.731	0.621
	0.997	0.994	0.987	0.962	0.916	0.853	0.637	0.605	0.558
	1.015	1.021	1.030	1.044	1.044	1.018	0.786	0.729	0.619
CSRL	2.001	2.002	2.006	2.020	2.047	2.081	2.026	1.949	1.721
	2.000	2.000	2.000	1.997	1.983	1.947	1.515	1.376	1.087
	2.001	2.002	2.006	2.021	2.049	2.085	2.032	1.954	1.723
CKRL	6.005	6.009	6.024	6.085	6.210	6.384	6.290	5.936	4.816
	6.000	6.000	5.998	5.986	5.931	5.775	3.880	3.289	2.087
	6.005	6.010	6.025	6.088	6.221	6.406	6.339	5.980	4.841

RL perc. points	θ								
	0	0.001	0.0025	0.005	0.0075	0.01	0.02	$p_1 - p_0$	0.03
5%	35	17	6	3	2	2	1	1	1
	33	22	14	8	5	4	2	1	1
	22	12	6	3	2	2	1	1	1
25%	274	156	71	23	10	6	2	2	2
	175	115	64	28	15	10	4	3	2
	164	104	54	20	9	5	2	2	2
Median	684	394	185	63	26	14	4	3	2
	419	274	150	62	31	18	6	5	4
	408	263	140	54	24	13	4	3	2
75%	1386	802	379	130	55	28	6	5	3
	836	546	297	120	57	32	9	7	5
	825	535	288	113	51	27	6	5	3
90%	2313	1340	636	220	94	48	10	8	5
	1388	905	491	198	93	51	13	10	7
	1376	894	482	190	87	46	10	8	5
95%	3015	1748	830	288	123	63	13	10	6
	1805	1176	638	256	119	65	15	12	8
	1793	1166	629	248	113	60	12	10	6

Table 4.7: Percentage reduction in the *ARL*, *SDRL*, *CVRL*, *CSRL*, *CKRL* and *RL* percentage points due to the adoption of the *CS* scheme — listed in order corresponding to $(1 - CS^0/C^3) \times 100\%$ and $(1 - CS^3/C^3) \times 100\%$.

Percentage reduction	θ								
	0	0.001	0.0025	0.005	0.0075	0.01	0.02	$p_1 - p_0$	0.03
ARL	39.3	31.4	20.7	5.7	-6.6	-17.1	-39.9	-41.5	-41.1
	40.5	33.2	24.0	13.5	7.6	4.5	1.0	0.8	0.4
SDRL	40.5	33.3	24.2	13.5	6.9	2.3	-13.0	-17.1	-26.7
	40.6	33.4	24.3	13.8	8.0	4.8	1.4	1.1	0.8
CVRL	2.0	2.8	4.4	8.2	12.6	16.5	19.2	17.2	10.2
	0.2	0.2	0.3	0.4	0.4	0.4	0.3	0.3	0.3
CSRL	0.1	0.1	0.3	1.2	3.1	6.5	25.2	29.4	36.9
	0.0	0.0	0.0	0.0	-0.1	-0.2	-0.3	-0.3	-0.1
CKRL	0.1	0.2	0.4	1.6	4.5	9.5	38.3	44.6	56.7
	0.0	0.0	0.0	-0.1	-0.2	-0.3	-0.8	-0.7	-0.5

Percentage reduction	θ								
	0	0.001	0.0025	0.005	0.0075	0.01	0.02	$p_1 - p_0$	0.03
5%	5.7	-29.4	-133.3	-166.7	-150.0	-100.0	-100.0	0.0	0.0
	37.1	29.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0
25%	36.1	26.3	9.9	-21.7	-50.0	-66.7	-100.0	-50.0	0.0
	40.1	33.3	23.9	13.0	10.0	16.7	0.0	0.0	0.0
Median	38.7	30.5	18.9	1.6	-19.2	-28.6	-50.0	-66.7	-100.0
	40.4	33.2	24.3	14.3	7.7	7.1	0.0	0.0	0.0
75%	39.7	31.9	21.6	7.7	-3.6	-14.3	-50.0	-40.0	-66.7
	40.5	33.3	24.0	13.1	7.3	3.6	0.0	0.0	0.0
90%	40.0	32.5	22.8	10.0	1.1	-6.2	-30.0	-25.0	-40.0
	40.5	33.3	24.2	13.6	7.4	4.2	0.0	0.0	0.0
95%	40.1	32.7	23.1	11.1	3.3	-3.2	-15.4	-20.0	-33.3
	40.5	33.3	24.2	13.9	8.1	4.8	7.7	0.0	0.0

Another sort of comparison that springs to mind involves the very well known head start technique. This confrontation will tell us if a practitioner should go to the trouble of adopting a combined scheme or just add a head start to the upper one-sided *CUSUM* scheme.

The Tables 4.6 and 4.7 and the Figures 4.5 and 4.6 have numerical results that are crucial to make comparisons possible between the

- upper one-sided *CUSUM* scheme with a 50% head start (C^3),

the CS^0 scheme, and finally with the

- upper one-sided combined *CUSUM–Shewhart* scheme with a 50% head start (CS^3).

We begin with a brief comment concerning schemes CS^0 and CS^3 and the consequences of adding a 50% head start. It was proved that the alarm rate at sample m increases after having given the scheme a head start. In addition, the numerical results in Table 4.6 suggest that the standard deviation of the combined scheme is virtually independent of the 50% head start, for small and moderate shifts in p.

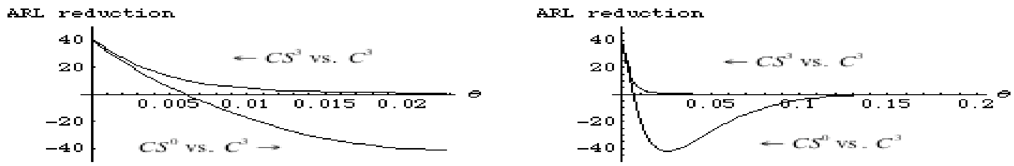

Figure 4.5: *ARL* percentage reductions relative to the upper one-sided *CUSUM* scheme with a 50% headstart C^3 in the intervals $[0, p_1 - p_0]$ and $[0, 0.2]$.

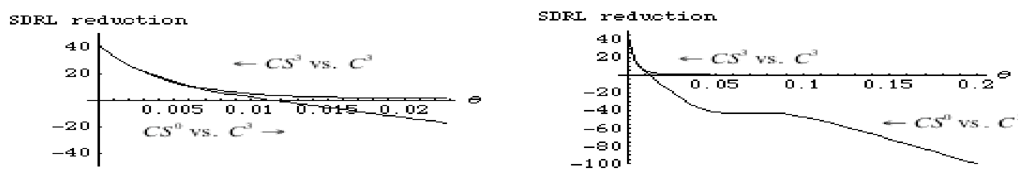

Figure 4.6: *SDRL* percentage reductions relative to the upper one-sided *CUSUM* scheme with a 50% headstart C^3 in the intervals $[0, p_1 - p_0]$ and $[0, 0.2]$.

Adding a *Shewhart* upper control limit to scheme C^3 — thus, yielding the scheme CS^3 — has not the same consequences we reported previously when we compared schemes C^0 and CS^0. For instance, the schemes C^3 and CS^3 give approximately the same protection to large shifts, as we can see from Figure 4.5; in addition, the *SDRL* of scheme CS^3 is larger than the *SDRL* of scheme C^3, not only for moderate but also for large shifts (see Figure 4.6).

Results in Tables 4.6 and 4.7 show that an upper one-sided *CUSUM* chart with a 50% head start gives most of the improvement of the slightly more elaborate upper one-sided combined *CUSUM–Shewhart* scheme without a head start, for moderate and large shifts. In fact, scheme C^3 has larger in-control *ARL*, but substantial increases occur in *ARL* and *SDRL*, when this scheme is replaced by the scheme CS^0 for moderate and large shifts. Furthermore, note that, although the in-control $ARL^0(6,4;\theta)$ is smaller than $ARL^3(6;\theta)$ for small shifts, the 5% percentage points of $RL^3(6;\theta)$ may be remarkably smaller than those of $RL^0(6,4;\theta)$, meaning that a correct signals can occur earlier when we use scheme C^3.

However, scheme CS^0 seems to outperform the scheme C^3 for small shifts, in the interval $(0, 0.005]$, when it comes to average time detection, standard deviation, and the variation, skewness and kurtosis coefficients, as suggested by Table 4.7, Figure 4.5 and Figure 4.6. Thus, the combined upper one-sided *CUSUM–Shewhart* scheme without head start does not seem to be completely superseded by the upper one-sided *CUSUM* scheme with the well know 50% head start. We would like to remind the reader that we are dealing with a very small nominal value p_0, therefore small changes are more likely to occur than moderate or large shifts (despite the value of p_1, which doubles p_0).

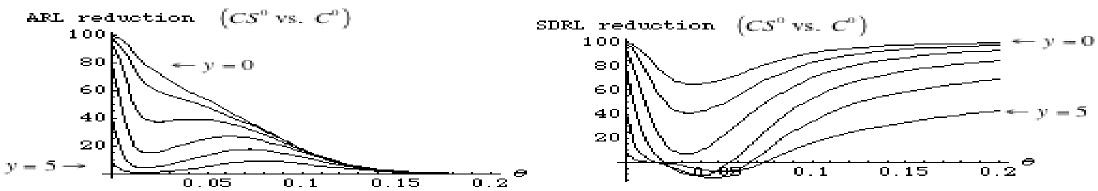

Figure 4.7: *ARL* and *SDRL* percentage reductions due to the adoption of the combined scheme CS^0 for $y = 0, \dots, 5$ (from top to bottom).

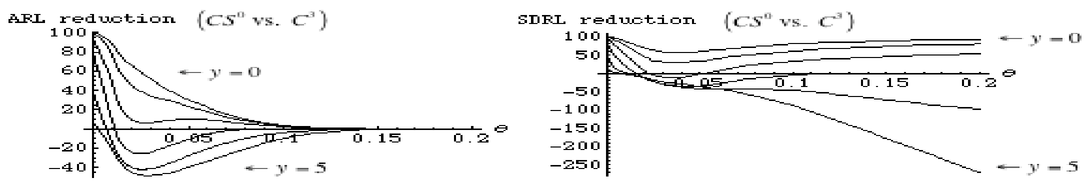

Figure 4.8: *ARL* and *SDRL* percentage reductions relative to the upper one-sided *CUSUM* scheme with a 50% headstart C^3 for $y = 0, \dots, 5$ (from top to bottom).

Some additional investigations on the impact of the choice of y in the *ARL* and *SDRL* of the combined scheme (when compared to schemes C^0 and C^3) are gathered next.

Although Figure 4.7 and Figure 4.8 have self explanatory captions, we note that they include several graphs of the reduction in *ARL* and *SDRL* induced by the adoption of a CS^0 scheme as an alternative to scheme C^0 (Figure 4.7) and yielded by the replacement of scheme C^3 by scheme CS^0 (Figure 4.8), for distinct values of the critical increment, y. The graphs correspond to $y = 0, 1, \dots, x - 1 = 5$, from top to bottom.

We add to this observation that the average performance of scheme CS^0 appears to show reasonable improvement over scheme C^0 — without a tremendous decrease in the in-control *ARL* — only for $y = 4$ (see Figure 4.7). Additionally, the *SDRL* generally decreases after that substitution. Therefore $y = 4$ seems to be the natural choice for the critical increment.

As far as the comparison of schemes C^3 and CS^0 is concerned, only values $y = 0, 1, 2$ yield an overall decrease in *ARL* (and in *SDRL* for most of the values of θ in the interval $[0, 0.2]$); however the increase in the in-control *ARL* is rather unreasonable, as Figure 4.7 clearly suggests also for $y = 3$.

As suggested by Lucas (1982) when a combined scheme is implemented it may be advantageous to increase the value of the *CUSUM* upper control limit x slightly to get a larger in-control *ARL*. However, our numerical investigations with $x = 7$ and $y = 3, 4$ lead to worse results than those achieved previously with $x = 6$ and $y = 4$. With $x = 7$ and $y = 4$ the obtained improvements were far less significant (due to the stochastic property (4.39)); and with $x = 7$ when $y = 3$ there was a dramatic 75% reduction in the in-control *ARL*.

As a concluding remark, we would like to recall that supplementing the upper one-sided *CUSUM* scheme (with or without head start) with a *Shewhart* upper control limit can provide meaningful improvement in the *RL* related measures for small changes in the process parameter with little additional computational effort. But this is achieved with an undesirable reduction of the in-control *ARL*. Therefore adopting the combined upper one-sided *CUSUM–Shewhart* scheme would depend on whether we are more concerned with detecting small shifts as fast as possible or with false alarms.

Furthermore, the best way to compare control chart performance is probably to match the in-control *ARL* and then compare the out-of-control performance; barring that, comparisons are difficult as seen in this section. However, the discrete character of *binomial* data affects *ARL*s and causes unique problems in the matching procedure of *CUSUM* and combined *CUSUM–Shewhart* schemes — it can never be achieved for this sort of data (or any other discrete data) — and, thus, prevents a "fair" numerical comparison between different schemes. Matters such as this are going to be addressed in the next two chapters, while making use of combined schemes for continuous data.

Chapter 5

Upper one-sided schemes for μ in the presence of shifts in σ

One of the standard assumptions, when control schemes for the process mean μ are at use, is that the process data is a collection of observed values of independent random variables with *constant* standard deviation σ equal to a *known* target value σ_0.

Some attention has been given in the statistical process control literature to the *RL* distribution of schemes for the expected value μ, when the target of the process standard deviation is unknown (although constant), and therefore the control limits have to be estimated (see, e.g., Bagshaw and Johnson (1975), Ghosh, Reynolds Jr. and Hui (1981) and Chen (1997)). In any case it may be quite misleading to assume that the behaviour of the *RL* for σ known can be carried over to draw conclusions about the case where σ is unknown, specially in the in-control situation.

More recently, Jones, Champ and Rigdon (2001) derived the run length distribution of the *EWMA* scheme with estimated parameters, and discussed the effect of estimation on the performance of the scheme in several practical scenarios.

Another compelling question concerning schemes for μ is to know what happens to the *RL* distribution in the presence of shifts in the process standard deviation σ, if the operator falsely assumes that σ is constant and equal to a known target value σ_0, and designs the control scheme for μ accordingly.

- *Hawkins and Olwell (1998, pp. 66-67).* These authors briefly discuss this problem. They note that standard deviation shifts have severe effects on *Shewhart* and *CUSUM* schemes for the process mean, and give a numerical illustration for the in-control *ARL* of an upper one-sided *CUSUM* scheme for μ.

- *Gan (1989).* This paper provides *ARL* tables for two-sided (*Shewhart*, *CUSUM* and *EWMA*) schemes for μ in such a setting. These tables cast some light on the performance of these schemes. For example, Tables 6–8 show that, for some large but fixed shifts in μ, the *ARL* of these two-sided schemes can be an increasing or even a nonmonotone function of σ. This behaviour — although not commented by the author — is apparent in the last rows of Tables 6–8, and it implies that these schemes become progressively less sensitive to some shifts in the process mean, as the process standard deviation grows.

- *Morais and Pacheco (2001a).* In this reference we establish sufficient conditions for the RL of the upper one-sided $\bar{X}$ ($S^+ - \mu$) and $EWMA$ ($E^+ - \mu$) schemes for the process mean to stochastically increase (or decrease) with an increase in the process standard deviation.

Such conditions and their analogue for the upper one-sided $CUSUM$ ($C^+ - \mu$), combined *CUSUM-Shewhart* ($CS^+ - \mu$) and combined *EWMA-Shewhart* ($ES^+ - \mu$) schemes are stated and proved in this chapter.

These preparatory results and several other stochastic properties of the RL cast interesting light on the performance of the five upper one-sided schemes for μ in Table A.1, namely on their ability to detect shifts in process standard deviation, and prove to be crucial for Chapter 6, devoted to misleading signals in joint schemes for μ and σ.

Throughout the remainder of this chapter we shall consider that the N^{th} random sample $X_N = (X_{1N}, \ldots, X_{nN})$ is drawn from a distribution belonging to the normal family $\{N(\mu, \sigma^2), -\infty < \mu < +\infty, \sigma^2 > 0\}$.

The shift in μ is represented in terms of the nominal value of the sample mean standard deviation $\delta = \sqrt{n}(\mu - \mu_0)/\sigma_0$, and the increase of the process standard deviation will be measured by $\theta = \sigma/\sigma_0$ where μ_0 and σ_0 represent the nominal values of μ and σ, $\delta \geq 0$ and $\theta \geq 1$. When the process is in control we have $(\delta, \theta) = (0, 1)$, whereas out-of-control (δ, θ) takes a constant value (assumed to be known) in $[0, +\infty) \times [1, +\infty) \setminus \{(0, 1)\}$. And in the absence of shifts in σ we have $(\delta, \theta) \in [0, +\infty) \times \{1\}$.

Finally, note that the summary statistics, the control limits and the RL distribution of the five upper one-sided schemes for μ considered here can be found in Tables A.2, A.3 and A.8.

5.1 The Shewhart scheme

The production is declared out-of-control at time N by the upper one-sided $\bar{X}$ scheme, $S^+ - \mu$, if

$$z_N^+ = \max\{0, z_N\} = \max\{0, \sqrt{n} \times (\bar{x}_N - \mu_0)/\sigma_0\}, \tag{5.1}$$

the observed value of the summary statistic Z_N^+, equals or is above the control limit

$$UCL_{S^+ - \mu} = \xi_\mu^+. \tag{5.2}$$

$\sigma_0/\sqrt{n}$ denotes the nominal value of the sample mean standard deviation. ξ_μ^+ belongs to $(0, +\infty)$ and is usually selected by fixing the ARL for the $S^+ - \mu$ scheme in two situations: one being when the quality level is acceptable ($\mu = \mu_0$ and $\sigma = \sigma_0$), and one when it is rejectable ($\mu > \mu_0$ or $\sigma > \sigma_0$),

Conditioned on the fact that the process mean and standard deviation equal $\mu = \mu_0 + \delta\sigma_0/\sqrt{n}$ and $\sigma = \theta \times \sigma_0$ (respectively), the RL of the upper one-sided scheme $S^+ - \mu$

— $RL_{S^+-\mu}(\delta,\theta)$ — has a geometric distribution with survival function given by

$$\overline{F}_{RL_{S^+-\mu}(\delta,\theta)}(m) = \begin{cases} 1, & m < 1 \\ \{\Phi[(\xi_\mu^+ - \delta)/\theta]\}^{\lfloor m \rfloor}, & m \geq 1. \end{cases} \tag{5.3}$$

Theorem 5.1 — *$RL_{S^+-\mu}(\delta,\theta)$ has the following stochastic properties:*

- *for fixed $\theta \geq 1$,*

$$RL_{S^+-\mu}(\delta,\theta) \downarrow_{lr} \text{ with } \delta; \tag{5.4}$$

- *for fixed $\delta < \xi_\mu^+$,*

$$RL_{S^+-\mu}(\delta,\theta) \downarrow_{lr} \text{ with } \theta; \tag{5.5}$$

- *for fixed $\delta = \xi_\mu^+$,*

$$RL_{S^+-\mu}(\delta,\theta) =_{st} RL_{S^+-\mu}(\delta,\theta'), \text{ for } \theta,\theta' \geq 1; \tag{5.6}$$

- *for fixed $\delta > \xi_\mu^+$,*

$$RL_{S^+-\mu}(\delta,\theta) \uparrow_{lr} \text{ with } \theta. \tag{5.7}$$

Proof — Recall that $RL_{S^+-\mu}(\delta,\theta)$ has a geometric distribution with parameter $\{1 - \Phi[(\xi_\mu^+ - \delta)/\theta]\}$, and that this distribution stochastically decreases in the likelihood ratio sense, as its parameter increases, according to Remark 2.2. Now, noting that, for $\delta \geq 0$ and $\theta \geq 1$, $\{1 - \Phi[(\xi_\mu^+ - \delta)/\theta]\}$ is an increasing function of δ and an increasing, constant and decreasing function of θ when $\delta < \xi_\mu^+$, $\delta = \xi_\mu^+$ and $\delta > \xi_\mu^+$ (respectively), statements (5.4)–(5.7) follow immediately. •

It can be seen from result (5.4) ((5.5)) that, for fixed $\theta \geq 1$ (for fixed $\delta < \xi_\mu^+$), the larger the increase in the process mean (standard deviation), the smaller (in the likelihood ratio sense) the number of samples taken until the detection of such a change. Thus, $d_{RL_{S^+-\mu}(\delta,\theta)}(m)$ — the (relative) decrease in the probability that the m^{th} sample triggers a signal relative to the probability of the signal being given by the previous sample — increases with δ. In addition, since $\leq_{lr} \Rightarrow \leq_{hr}$ we can also assert that the alarm rate at any fixed sample m increases as the shift in μ becomes more severe.

It is worth mentioning that result (5.6) means that when $\delta = \xi_\mu^+$ the RL has a distribution function which does not depend on θ and it is a geometric random variable with parameter 0.5 (see Figure 5.1). This fact is not surprising because scheme $S^+ - \mu$ (as well as the other upper one-sided schemes for μ) is not designed to detect changes in σ.

Remark 5.2 — Result (5.7) can be phrased more clearly by the following statement. For fixed $\delta > \xi_\mu^+$, the more severe is the shift in σ the smaller is the ability (in the likelihood ratio sense) of the scheme $S^+ - \mu$ to discriminate effectively changes in the two process parameters, and in particular, in the process mean.

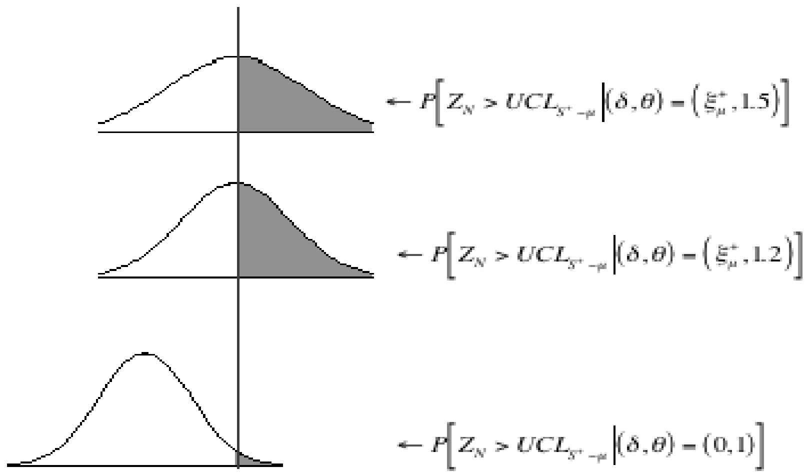

Figure 5.1: Two out-of-control sample mean distributions ($(\delta, \theta) = (\xi_\mu^+, 1.5)$, $(\delta, \theta) = (\xi_\mu^+, 1.2)$), illustrating result (5.6), and in-control distribution ($(\delta, \theta) = (0, 1)$).

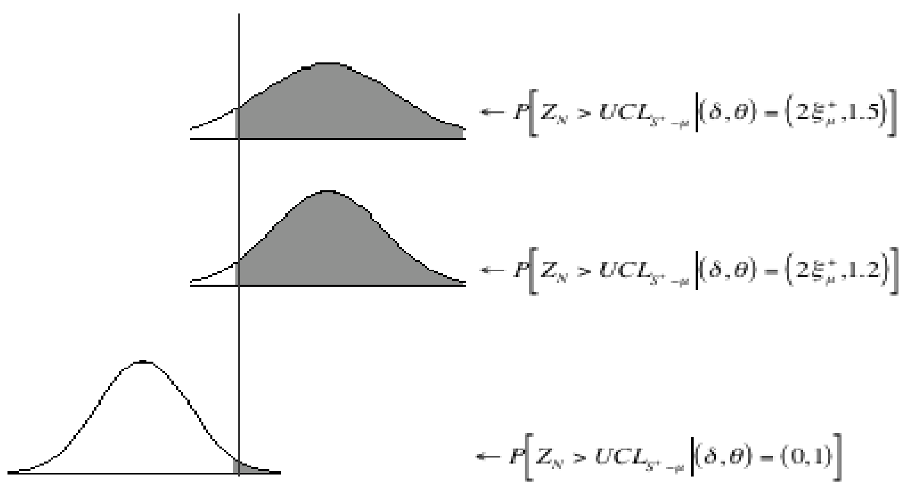

Figure 5.2: Two out-of-control sample mean distributions ($(\delta, \theta) = (2\xi_\mu^+, 1.5)$, $(\delta, \theta) = (2\xi_\mu^+, 1.2)$), illustrating result (5.7), and in-control distribution ($(\delta, \theta) = (0, 1)$).

As a consequence, $RL_{S+-\mu}(\delta, \theta)$ also increases with θ in the hazard rate sense; i.e., the signalling rate at sample m decreases with θ, regardless of the value of m. An intuitive explanation was mentioned in the previous paragraph.

A further justification for this increasing behaviour of the RL stems from the fact that, when $\delta > \xi_\mu^+$, the true process mean is above the upper control limit $UCL_{S+-\mu}$, and the probability that the nominal standardized sample mean exceeds $UCL_{S+-\mu}$ is equal to $\Phi[(\delta - \xi_\mu^+)/\theta]$ and larger than 0.5; thus, a decrease of σ will make it more likely that the nominal standardized sample mean will be responsible for a signal at each sample than it would be if σ would increase (see Figure (5.2)). •

5.2 The EWMA scheme

The summary statistic of the upper one-sided *EWMA* scheme can be expressed as follows:

$$W^+_{\mu,N} = \begin{cases} w^+_{\mu,0}, \; N = 0 \\ max\left\{0, \; (1-\lambda^+_\mu) \times W^+_{\mu,N-1} + \lambda^+_\mu \times Z_N\right\}, \; N > 0. \end{cases} \tag{5.8}$$

λ^+_μ belongs to $(0,1]$ and corresponds to the weight given to the most recent sample mean and $w^+_{\mu,0}$ is the initial value given to the $E^+ - \mu$ control scheme statistic.

This scheme gives an out-of-control signal at the sampling period N if the observed value of $W^+_{\mu,N}$ equals or exceeds the control limit

$$UCL_{E^+-\mu} = \gamma^+_\mu \sqrt{\lambda^+_\mu/(2-\lambda^+_\mu)}, \tag{5.9}$$

where γ^+_μ belongs to $(0,+\infty)$ and is selected in the same manner as ξ^+_μ.

Set

$$w^+_{\mu,0} = \alpha\, \gamma^+_\mu \sqrt{\lambda^+_\mu/(2-\lambda^+_\mu)}, \tag{5.10}$$

where $\alpha \in [0,1)$; if $\alpha \in (0,1)$ $(\alpha = 0)$ a $\alpha \times 100\%$ head start (no head start) has been given to the chart.

Let $RL^\alpha_{E^+-\mu}(\delta,\theta;x^+_\mu) = RL^{\lfloor\alpha(x^+_\mu+1)\rfloor}_{E^+-\mu}(\delta,\theta;x^+_\mu)$ be the Markov approximation for the run length of the scheme $E^+ - \mu$ with an $\alpha \times 100\%$ head start, based on an absorbing Markov chain with discrete state space $\{0,1,\ldots,x^+_\mu+1\}$ and absorbing state $(x^+_\mu+1)$.[1]

In this case $RL^\alpha_{E^+-\mu}(\delta,\theta;x^+_\mu)$ has survival function given by

$$\overline{F}_{RL^\alpha_{E^+-\mu}(\delta,\theta;x^+_\mu)}(m) = \begin{cases} 1, \; m < 1 \\ \underline{\mathbf{e}}^\top_{\lfloor\alpha(x^+_\mu+1)\rfloor} \times \left[\mathbf{Q}(\delta,\theta;x^+_\mu)\right]^{\lfloor m\rfloor} \\ \qquad\qquad \times \underline{\mathbf{1}}, \; \alpha \in [0,1), \; m \geq 1, \end{cases} \tag{5.11}$$

where the $(x^+_\mu+1) \times (x^+_\mu+1)$ sub-stochastic matrix which rules the transitions between the transient states is defined as follows

$$\mathbf{Q}(\delta,\theta;x^+_\mu) = \left[p_{ij}(\delta,\theta;x^+_\mu)\right]^{x^+_\mu}_{i,j=0} = \left[\sum_{l=0}^{j} p_{i\,l}(\delta,\theta;x^+_\mu) - \sum_{l=0}^{j-1} p_{i\,l}(\delta,\theta;x^+_\mu)\right]^{x^+_\mu}_{i,j=0} \tag{5.12}$$

with

$$\sum_{l=0}^{j} p_{i\,l}(\delta,\theta;x^+_\mu) \;=\; \Phi\left(\frac{1}{\theta} \times \left\{\frac{\gamma^+_\mu \times [(j+1)-(1-\lambda^+_\mu)(i+1/2)]}{(x^+_\mu+1)\sqrt{\lambda^+_\mu(2-\lambda^+_\mu)}} - \delta\right\}\right), \tag{5.13}$$

for $i,j = 0,\ldots,x^+_\mu$, and $\sum_{l=0}^{-1} p_{i\,l}(\delta,\theta;x^+_\mu) = 0, \; i = 0,\ldots,x^+_\mu$. (For the underlying details of this Markov chain refer to Appendix A.)

The next three theorems concern the stochastic monotone character of the transition matrix $\mathbf{P}(\delta,\theta;x^+_\mu)$ and the monotone behaviour of *RL* with respect to the head start and

[1]Note that a $\alpha \times 100\%$ head start corresponds to the initial state $\lfloor\alpha(x^+_\mu+1)\rfloor$ in the Markov approximation.

other parameters. A further feature of this transition matrix is discussed in detail in Section 5.5 because we are not dealing with a space homogeneous Markov chain (with a reflecting and an absorbing barrier) like the one associated with upper one-sided *CUSUM* schemes.

Theorem 5.3 — *Let* $\mathbf{P}(\delta, \theta; x^+_\mu)$ *be the stochastic matrix ruling the approximating Markov chain associated with scheme* $E^+ - \mu$*. Then*

$$\mathbf{P}(\delta, \theta; x^+_\mu) \in \mathcal{M}_{st}. \tag{5.14}$$

Proof — The decreasing behaviour of $\sum_{l=0}^{j} p_{i\,l}(\delta, \theta; x^+_\mu)$ in terms of i is responsible for this result. •

Theorem 5.4 — *For fixed* $\delta \geq 0$ *and* $\theta \geq 1$*,*

$$RL^\alpha_{E^+-\mu}(\delta, \theta; x^+_\mu) \downarrow_{st} \text{ with } \alpha. \tag{5.15}$$

Proof — This stochastic property follows from Theorem 5.3 and result (3.14) from Theorem 3.8. •

Theorem 5.5 — *For fixed* $0 \leq \alpha < 1$*, results (5.4), (5.5) and (5.7) have an analogue for the scheme* $E^+ - \mu$*:*

- *for fixed* $\theta \geq 1$*,*

$$RL^\alpha_{E^+-\mu}(\delta, \theta; x^+_\mu) \downarrow_{st} \text{ with } \delta; \tag{5.16}$$

- *for fixed* $\delta \leq \underline{\delta}_{E^+-\mu}(x^+_\mu)$*,*

$$RL^\alpha_{E^+-\mu}(\delta, \theta; x^+_\mu) \downarrow_{st} \text{ with } \theta, \tag{5.17}$$

- *for fixed* $\delta \geq \overline{\delta}_{E^+-\mu}(x^+_\mu)$*,*

$$RL^\alpha_{E^+-\mu}(\delta, \theta; x^+_\mu) \uparrow_{st} \text{ with } \theta, \tag{5.18}$$

where

$$\underline{\delta}_{E^+-\mu}(x^+_\mu) = \frac{\gamma^+_\mu[1 - (1 - \lambda^+_\mu)(x^+_\mu + 1/2)]}{(x^+_\mu + 1)\sqrt{\lambda^+_\mu(2 - \lambda^+_\mu)}} \tag{5.19}$$

and

$$\overline{\delta}_{E^+-\mu}(x^+_\mu) = \frac{\gamma^+_\mu[(x^+_\mu + 1) - (1 - \lambda^+_\mu)/2]}{(x^+_\mu + 1)\sqrt{\lambda^+_\mu(2 - \lambda^+_\mu)}} \tag{5.20}$$

depend on the number of transient states of the approximating Markov chain.

Proof ((5.16): monotonicity in δ) — The approximating RL of schemes $E^+\!-\mu$ is related to a stochastically monotone matrix $\mathbf{P}(\delta,\theta;x_\mu^+)$ and to the left partial sums $\sum_{l=0}^{j} p_{i\,l}(\delta,\theta;x_\mu^+)$ that decrease with δ (see (5.13)), for $i,j=0,\ldots,x_\mu^+$. Hence, by Corollary 3.14, $RL^{\alpha}_{E^+-\mu}$ $(\delta,\theta;x_\mu^+)$ stochastically decrease with δ. •

Proof ((5.17) and (5.18): monotonicities in θ) — The identification of the stochastic monotone behaviour of $RL^{\alpha}_{E^+-\mu}(\delta,\,\theta;x_\mu^+)$ with respect to θ is far less obvious and implies the calculation of the following sign:

$$sign\left\{\frac{\partial\sum_{l=0}^{j} p_{i\,l}(\delta,\theta;x_\mu^+)}{\partial\theta}\right\}$$
$$= sign\left\{-\frac{\gamma_\mu^+\times[(j+1)-(1-\lambda_\mu^+)(i+1/2)]}{(x_\mu^+ +1)\sqrt{\lambda_\mu^+(2-\lambda_\mu^+)}}+\delta\right\} \tag{5.21}$$

for $i,j=0,\ldots,x_\mu^+$. This sign is negative if

$$\delta \;<\; \min_{i,j=0,\ldots,x_\mu^+}\left\{\frac{\gamma_\mu^+\times[(j+1)-(1-\lambda_\mu^+)(i+1/2)]}{(x_\mu^+ +1)\sqrt{\lambda_\mu^+(2-\lambda_\mu^+)}}\right\}\Leftrightarrow$$
$$\delta \;<\; \frac{\gamma_\mu^+\times[1-(1-\lambda_\mu^+)(x_\mu^+ +1/2)]}{(x_\mu^+ +1)\sqrt{\lambda_\mu^+(2-\lambda_\mu^+)}}=\underline{\delta}_{E^+-\mu}(x_\mu^+), \tag{5.22}$$

and it is positive if

$$\delta \;>\; \max_{i,j=0,\ldots,x_\mu^+}\left\{\frac{\gamma_\mu^+\times[(j+1)-(1-\lambda_\mu^+)(i+1/2)]}{(x_\mu^+ +1)\sqrt{\lambda_\mu^+(2-\lambda_\mu^+)}}\right\}\Leftrightarrow$$
$$\delta \;>\; \frac{\gamma_\mu^+\times[(x_\mu^+ +1)-(1-\lambda_\mu^+)/2]}{(x_\mu^+ +1)\sqrt{\lambda_\mu^+(2-\lambda_\mu^+)}}=\overline{\delta}_{E^+-\mu}(x_\mu^+). \tag{5.23}$$

Also note that, for $\delta=\underline{\delta}_{E^+-\mu}(x_\mu^+)$ ($\delta=\overline{\delta}_{E^+-\mu}(x_\mu^+)$), all the derivatives of $\sum_{l=0}^{j} p_{i\,l}(\delta,\theta;x_\mu^+)$ with respect to θ are surely nonpositive (nonnegative). Thus, using Corollary 3.14, we can conclude that

$$RL^{\alpha}_{E^+-\mu}(\delta,\theta;x_\mu^+)\downarrow_{st}\ (\uparrow_{st})\text{ with }\theta, \tag{5.24}$$

for fixed $\delta\le\underline{\delta}_{E^+-\mu}(x_\mu^+)$ ($\delta\ge\overline{\delta}_{E^+-\mu}(x_\mu^+)$). •

Remark 5.6 — We note that as far as its monotone behaviour is concerned, $\sum_{l=0}^{j} p_{i\,l}(\delta,\theta;$ $x_\mu^+)$ is an increasing function of λ_μ^+ if

$$sign\left\{\frac{\partial\sum_{l=0}^{j} p_{i\,l}(\delta,\theta;x_\mu^+)}{\partial\lambda_\mu^+}\right\}= sign\left\{-\frac{\gamma_\mu^+\times[(j+1)(1-\lambda_\mu^+)-(i+1/2)]}{(x_\mu^+ +1)\left[\lambda_\mu^+(2-\lambda_\mu^+)\right]^{3/2}}\right\} \tag{5.25}$$

is nonnegative, for $i,j=0,\ldots,x_\mu^+$. This is necessarily the case for

$$\lambda_\mu^+\ge\max_{i,j=0,\ldots,x_\mu^+}\left(1-\frac{i+1/2}{j+1}\right)\Leftrightarrow\lambda_\mu^+\ge 1-\frac{1}{2(x_\mu^+ +1)}. \tag{5.26}$$

Recalling that $\lambda \in (0,1]$ and using Corollary 3.13, we conclude that, for fixed $\delta \geq 0$, $\theta \geq 1$ and $0 \leq \alpha < 1$,

$$RL^{\alpha}_{E^+-\mu}(\delta,\theta;x^+_\mu) \uparrow_{st} \text{ with } \lambda^+_\mu, \text{ for } 1-\frac{1}{2(x^+_\mu+1)} \leq \lambda^+_\mu \leq 1. \tag{5.27}$$

Two additional remarks on the stochastic monotone behaviour with respect to λ^+_μ. Regarding the sufficient condition (5.27), please note that we get $\lambda^+_\mu = 1$ when $x^+_\mu \to \infty$; thus, statement (5.27) is of no practical use. Moreover, (5.25) is nonpositive if

$$\lambda \leq \min_{i,j=0,\ldots,x^+_\mu} \left(1-\frac{i+1/2}{j+1}\right) \Leftrightarrow \lambda \leq 1-\frac{x^+_\mu+1/2}{2}, \tag{5.28}$$

where $1-\frac{x^+_\mu+1/2}{2}$ is negative for all positive integers but $x^+_\mu = 1$, and particularly large for the usual values of x^+_μ; therefore it comes as no surprise that we did not state a sufficient condition for the decreasing stochastic behaviour of $RL^{\alpha}_{E^+-\mu}(\delta,\theta;x^+_\mu)$ in terms of λ^+_μ. •

Corollary 5.7 — *Let $RL^{\alpha}_{E^+-\mu}(\delta,\theta)$ be the exact* RL *of the $E^+-\mu$ scheme. Then properties (5.15)–(5.18) remain valid for $RL^{\alpha}_{E^+-\mu}(\delta,\theta)$, with $\underline{\delta}_{E^+-\mu}(x^+_\mu)$ and $\overline{\delta}_{E^+-\mu}(x^+_\mu)$ replaced by*

$$\underline{\delta}_{E^+-\mu} = \lim_{x^+_\mu\to\infty} \underline{\delta}_{E^+-\mu}(x^+_\mu) = \frac{-\gamma^+_\mu(1-\lambda^+_\mu)}{\sqrt{\lambda^+_\mu(2-\lambda^+_\mu)}} \tag{5.29}$$

and

$$\overline{\delta}_{E^+-\mu} = \lim_{x^+_\mu\to\infty} \overline{\delta}_{E^+-\mu}(x^+_\mu) = \frac{\gamma^+_\mu}{\sqrt{\lambda^+_\mu(2-\lambda^+_\mu)}} \tag{5.30}$$

in (5.17) and (5.18).

Proof — This is a consequence of Theorems 5.4 and 5.5, and Lemmas 3.16 and 3.18, which concern the convergence of the approximating RL and the closure of $\leq_{st}$ under the limit operation. •

Result (15) of Table 5.1 — which has a summary of the stochastic properties of $RL^{\alpha}_{E^+-\mu}(\delta,\theta)$ — reflects the stochastic implication of setting the $E^+-\mu$ scheme to an initial head start value: it will imply a decrease of the run length, which is reasonable enough because, for $0 \leq \alpha < \alpha' < 1$, the scheme with a $\alpha \times 100\%$ head start tends to signal less frequently than the same chart with a $\alpha' \times 100\%$ head start.

Result (16) of Table 5.1 essentially means that the run length of the $E^+-\mu$ scheme is stochastically decreasing in δ. Thus, the control chart increases its ability to detect an increase in δ as this change becomes more severe; result (16) of Table 5.1 is somehow expected since there is some similarity between the survival functions of $RL_{S^+-\mu}(\delta,\theta)$ and $RL_{E^+-\mu}(\delta,\theta;x^+_\mu)$. Morais and Pacheco (1998), p. 950, briefly mention these two stochastic properties.

Remark 5.8 — Given the usual values of ξ_μ^+, $RL_{S^+-\mu}(\delta,\theta)$ decreases stochastically with θ for most of the likely values of δ (i.e., small and moderate values of δ). But, notice that $\underline{\delta}_{E^+-\mu}(x_\mu^+)$ is rarely positive (in fact its limit $\underline{\delta}_{E^+-\mu}$ is negative) because we tend to use large values of x_μ^+ and small values of λ_μ^+, to provide a fine approximation to the properties of the true run length of the $E^+ - \mu$ scheme, and to yield a good performance in comparison to $RL_{S^+-\mu}$, respectively. •

Remark 5.9 — We cannot predict the monotone behaviour of $RL^\alpha_{E^+-\mu}(\delta,\theta;x_\mu^+)$ in terms of θ, for any fixed δ in the interval $(\underline{\delta}_{E^+-\mu},\overline{\delta}_{E^+-\mu})$. If we recall the considerations about the sign of $\underline{\delta}_{E^+-\mu}$ and the fact that $\overline{\delta}_{E^+-\mu}$ is rather large for the same reasons pointed out previously, we could assert that the interval $(\underline{\delta}_{E^+-\mu},\overline{\delta}_{E^+-\mu})$ includes all the relevant values of δ — the small and moderate ones. This is a severe disadvantage of not having the $E^+ - \mu$ scheme associated to a unique turning point in terms of stochastic monotonous behaviour (similarly to the upper one-sided scheme $S^+ - \mu$) as a consequence of dealing with a summary statistic with a far more complex structure — a Markov chain with continuous state space. •

The stochastic monotonicity result (18) of Table 5.1 is valid for "very large" shifts in μ, in particular, for $\delta \geq \overline{\delta}_{E^+-\mu}$. Thus, the associated ARL values are quite small and it seems at first glance that the increasing behaviour of the RL has no practical significance. However, this behaviour has an impact in some performance measures of the joint schemes for μ and σ, as we shall see in the next chapter.

After all these comments and remarks, we conclude that results (17) and (18) of Table 5.1 are not perfect analogues of (5.5) and (5.7).

5.3 The remaining schemes

Results of the same flavour of those in Theorems 5.4 and 5.5 and Corollary 5.7 can be stated for the three remaining upper one-sided schemes for μ, $C^+ - \mu$, $CS^+ - \mu$ and $ES^+ - \mu$. Table 5.1 gives an overview of such stochastic monotonicity properties.

The proof of results (5)–(9) of Table 5.1 concerning scheme $C^+ - \mu$ are sketched next; the proof and a remark on results (10)-(14) and (19)–(22), which concern the combined schemes and are somewhat more delicate, close this section.

Proof ((5) in Table 5.1: monotonicity in α, scheme $C^+ - \mu$) — The transition matrix associated to scheme $C^+ - \mu$ is stochastically monotone in the usual sense because the corresponding left partial sums in Table A.7 decrease with the row index i. Thus, according to result (3.14) from Theorem 3.8 and Lemma 3.18 the approximating and exact RLs stochastically decrease with the initial state, i.e., with α. •

Proof ((6) in Table 5.1: monotonicity in δ, scheme $C^+ - \mu$) — The stochastic monotonicity in δ of the approximating (and, thus, of the exact) RL of scheme $C^+ - \mu$ follow suit. In fact, the associated left partial sums in Table A.7 decrease with δ. Hence, the stochastic decreasing behaviour with regard to δ of this RL. •

Table 5.1: Stochastic properties of the exact RLs of the upper one-sided schemes for μ.

Scheme for μ	Stochastic properties
$S^+ - \mu$	(1) $RL_{S+-\mu}(\delta,\theta) \downarrow_{lr}$ with δ (2) $RL_{S+-\mu}(\delta,\theta) \downarrow_{lr}$ with θ, if $\delta < \underline{\delta}_{S+-\mu} = \overline{\delta}_{S+-\mu} = \xi_\mu^+$ (3) $RL_{S+-\mu}(\delta,\theta) =_{st} RL_{S+-\mu}(\delta,\theta')$ for $\theta,\theta' \in [1,+\infty)$, if $\delta = \underline{\delta}_{S+-\mu} = \overline{\delta}_{S+-\mu} = \xi_\mu^+$ (4) $RL_{S+-\mu}(\delta,\theta) \uparrow_{lr}$ with θ, if $\delta > \underline{\delta}_{S+-\mu} = \overline{\delta}_{S+-\mu} = \xi_\mu^+$
$C^+ - \mu$	(5) $RL^\alpha_{C+-\mu}(\delta,\theta) \downarrow_{st}$ with α (6) $RL^\alpha_{C+-\mu}(\delta,\theta) \downarrow_{st}$ with δ (7) $RL^\alpha_{C+-\mu}(\delta,\theta) \downarrow_{st}$ with θ, if $\delta \leq \underline{\delta}_{C+-\mu} = \lim_{x_\mu^+ \to +\infty} \left\{ k_\mu^+ + \frac{h_\mu^+[1-(x_\mu^+ + 1/2)]}{(x_\mu^+ + 1)} \right\} = k_\mu^+ - h_\mu^+$ (8) $RL^\alpha_{C+-\mu}(\delta,\theta) \uparrow_{st}$ with θ, if $\delta \geq \overline{\delta}_{C+-\mu} = \lim_{x_\mu^+ \to +\infty} \left\{ k_\mu^+ + \frac{h_\mu^+[(x_\mu^+ + 1) - 1/2]}{(x_\mu^+ + 1)} \right\} = k_\mu^+ + h_\mu^+$ (9) $RL^\alpha_{C+-\mu}(\delta,\theta;k_\mu^+) \uparrow_{st}$ with k_μ^+
$CS^+ - \mu$	(10) $RL^\alpha_{CS+-\mu}(\delta,\theta) \downarrow_{st}$ with α (11) $RL^\alpha_{CS+-\mu}(\delta,\theta) \downarrow_{st}$ with δ (12) $RL^\alpha_{CS+-\mu}(\delta,\theta) \downarrow_{st}$ with θ, if $\overline{\delta}_{C+-\mu} - \theta\xi_\mu^+ < \delta \leq \underline{\delta}_{C+-\mu}$ and $\theta > \frac{2h_\mu^+}{\xi_\mu^+} = \lim_{x_\mu^+ \to +\infty} \frac{\overline{\delta}_{C+-\mu}(x_\mu^+) - \underline{\delta}_{C+-\mu}(x_\mu^+)}{\xi_\mu^+}$ (13) $RL^\alpha_{CS+-\mu}(\delta,\theta) \uparrow_{st}$ with θ, if $\delta \geq \overline{\delta}_{C+-\mu}$ (14) $RL^\alpha_{CS+-\mu}(\delta,\theta;k_\mu^+) \uparrow_{st}$ with k_μ^+
$E^+ - \mu$	(15) $RL^\alpha_{E+-\mu}(\delta,\theta) \downarrow_{st}$ with α (16) $RL^\alpha_{E+-\mu}(\delta,\theta) \downarrow_{st}$ with δ (17) $RL^\alpha_{E+-\mu}(\delta,\theta) \downarrow_{st}$ with θ, if $\delta \leq \underline{\delta}_{E+-\mu} = \lim_{x_\mu^+ \to +\infty} \frac{\gamma_\mu^+[1-(1-\lambda_\mu^+)(x_\mu^+ + 1/2)]}{(x_\mu^+ + 1)\sqrt{\lambda_\mu^+(2-\lambda_\mu^+)}} = \frac{-\gamma_\mu^+(1-\lambda_\mu^+)}{\sqrt{\lambda_\mu^+(2-\lambda_\mu^+)}}$ (18) $RL^\alpha_{E+-\mu}(\delta,\theta) \uparrow_{st}$ with θ, if $\delta \geq \overline{\delta}_{E+-\mu} = \lim_{x_\mu^+ \to +\infty} \frac{\gamma_\mu^+[(x_\mu^+ + 1) - (1-\lambda_\mu^+)/2]}{(x_\mu^+ + 1)\sqrt{\lambda_\mu^+(2-\lambda_\mu^+)}} = \frac{\gamma_\mu^+}{\sqrt{\lambda_\mu^+(2-\lambda_\mu^+)}}$
$ES^+ - \mu$	(19) $RL^\alpha_{ES+-\mu}(\delta,\theta) \downarrow_{st}$ with α (20) $RL^\alpha_{ES+-\mu}(\delta,\theta) \downarrow_{st}$ with δ (21) $RL^\alpha_{ES+-\mu}(\delta,\theta) \downarrow_{st}$ with θ, if $\overline{\delta}_{E+-\mu} - \theta\xi_\mu^+ < \delta \leq \underline{\delta}_{E+-\mu}$ and $\theta > \frac{\gamma_\mu^+}{\xi_\mu^+}\sqrt{(2-\lambda_\mu^+)(\lambda_\mu^+)^{-1}} = \lim_{x_\mu^+ \to +\infty} \frac{\overline{\delta}_{E+-\mu}(x_\mu^+) - \underline{\delta}_{E+-\mu}(x_\mu^+)}{\xi_\mu^+}$ (22) $RL^\alpha_{ES+-\mu}(\delta,\theta) \uparrow_{st}$ with θ, if $\delta \geq \overline{\delta}_{E+-\mu}$

Proof ((7) and (8) in Table 5.1: monotonicities in θ, scheme $C^+ - \mu$) — The stochastic monotone behaviours of the RLs in terms of θ are proved below and follow quite similarly to the proofs for scheme $E^+ - \mu$. When it comes to scheme $C^+ - \mu$ we are dealing with

$$sign\left\{-k_\mu^+ - \frac{h_\mu^+ \times [(j+1)-(i+1/2)]}{x_\mu^+ + 1} + \delta\right\}, \tag{5.31}$$

for $i, j = 0, \ldots, x_\mu^+$. This sign is negative if

$$\begin{aligned} \delta &< \min_{i,j=0,\ldots,x_\mu^+} \left\{k_\mu^+ + \frac{h_\mu^+ \times [(j+1)-(i+1/2)]}{x_\mu^+ + 1}\right\} \Leftrightarrow \\ \delta &< k_\mu^+ - \frac{h_\mu^+ \times [1-(x_\mu^+ + 1/2)]}{x_\mu^+ + 1} = \underline{\delta}_{C^+-\mu}(x_\mu^+), \end{aligned} \tag{5.32}$$

and positive if

$$\begin{aligned} \delta &> \max_{0 \le i,j \le x_\mu^+} \left\{k_\mu^+ + \frac{h_\mu^+ \times [(j+1)-(i+1/2)]}{x_\mu^+ + 1}\right\} \Leftrightarrow \\ \delta &> k_\mu^+ + \frac{h_\mu^+ \times [(x_\mu^+ + 1) - 1/2]}{x_\mu^+ + 1} = \overline{\delta}_{C^+-\mu}(x_\mu^+). \end{aligned} \tag{5.33}$$

Also, for $\delta = \underline{\delta}_{C^+-\mu}(x_\mu^+)$ ($\delta = \overline{\delta}_{C^+-\mu}(x_\mu^+)$), all the derivatives of the left partial sums of the transition matrix with respect to θ are nonpositive (nonnegative). Thus, we get result (7) ((8)) in Table 5.1 by using Corollary 3.13 and Lemma 3.18. •

Proof ((9) in Table 5.1: monotonicity in k_μ^+, scheme $C^+ - \mu$) — The increasing stochastic behaviour in k_μ^+ of $RL^\alpha_{C^+-\mu}(\delta, \theta)$ follows from the fact that the approximating RL, $RL^\alpha_{C^+-\mu}(\delta, \theta; x_\mu^+)$, is associated to left partial sums that increase with k_μ^+ (see Table A.7), along with Lemma 3.18. •

Proof ((10) and (19) in Table 5.1: monotonicity in α, combined schemes $CS^+ - \mu$ and $ES^+ - \mu$) — The transition matrices associated to the combined schemes $CS^+ - \mu$ and $ES^+ - \mu$ are all stochastically monotone in the usual sense because the corresponding left partial sums in Table A.7 decrease with the row index i, whether we use the distribution function Φ or its truncated version Φ^+. Thus, the stochastic monotonicity in α of the run length of both combined schemes follows from result (3.14) from Theorem 3.8 and Lemma 3.18. •

Proof ((11) and (20) in Table 5.1: monotonicity in δ, combined schemes $CS^+ - \mu$ and $ES^+ - \mu$) — Results (6) and (16) in Table 5.1 still hold for the RL of the combined scheme $CS^+ - \mu$ and $ES^+ - \mu$ because the left partial sums are defined in terms of a truncation that does not affect their decreasing behaviour with regard to δ. Hence, the stochastic monotone decreasing behaviour in terms of δ of the run length of schemes $CS^+ - \mu$ and $ES^+ - \mu$ follows from Lemma 3.18. •

Proof ((14) in Table 5.1: monotonicity in k_μ^+, combined scheme $CS^+ - \mu$) — An analogue of result (9) in Table 5.1 is also valid for the RL of the combined scheme $CS^+ - \mu$ for the same reason pointed in the previous proof. In fact, the truncation that does not affect the sign of the derivative of the left partial sums with respect to k_μ^+: they remain positive. Thus, we conclude that the run length of scheme $CS^+ - \mu$ stochastically increases with k_μ^+, by using once again Lemma 3.18. •

Proof ((12) and (13) in Table 5.1: monotonicities in θ, combined scheme $CS^+ - \mu$) — To obtain the analogues of results (7) and (8) in Table 5.1, for scheme $CS^+ - \mu$, we must add that the sign of the derivatives of the left partial sum with respect to θ is non-null and equal to (5.31) if

$$\frac{1}{\theta}\left\{k_\mu^+ + \frac{h_\mu^+ \times [(j+1)-(i+1/2)]}{x_\mu^+ + 1} - \delta\right\} < \xi_\mu^+, \tag{5.34}$$

due to the truncation in ξ_μ^+ (see Equation (A.2)).

The sign in (5.31) is nonpositive, for $i, j = 0, \ldots, x_\mu^+$, if

$$\begin{aligned} \delta &\leq \min_{i,j=0,\ldots,x_\mu^+}\left\{k_\mu^+ + \frac{h_\mu^+ \times [(j+1)-(i+1/2)]}{x_\mu^+ + 1}\right\} = \underline{\delta}_{C^+-\mu}(x_\mu^+) \text{ and} \\ \delta &> \max_{i,j=0,\ldots,x_\mu^+}\left\{k_\mu^+ + \frac{h_\mu^+ \times [(j+1)-(i+1/2)]}{x_\mu^+ + 1}\right\} - \theta\xi_\mu^+ \\ &= \overline{\delta}_{C^+-\mu}(x_\mu^+) - \theta\xi_\mu^+. \end{aligned} \tag{5.35}$$

These two conditions only make sense when $\theta > \frac{\overline{\delta}_{C^+-\mu} - \underline{\delta}_{C^+-\mu}}{\xi_\mu^+}$, and in that case we get $\overline{\delta}_{C^+-\mu} - \theta\xi_\mu^+ < \delta \leq \underline{\delta}_{C^+-\mu}$.

Similarly, the sign in (5.31) is nonnegative under the constraint (5.34), for $i, j = 0, \ldots, x_\mu^+$, if

$$\begin{aligned} \delta &\geq \max_{0\leq i,j\leq x_\mu^+}\left\{k_\mu^+ + \frac{h_\mu^+ \times [(j+1)-(i+1/2)]}{x_\mu^+ + 1}\right\} = \overline{\delta}_{C^+-\mu}(x_\mu^+) \text{ and} \\ \delta &> \overline{\delta}_{C^+-\mu}(x_\mu^+) - \theta\xi_\mu^+, \end{aligned} \tag{5.36}$$

i.e., if $\delta \geq \overline{\delta}_{C^+-\mu}(x_\mu^+)$. Thus, the stochastic monotone behaviours (12) and (13) in Table 5.1 follow using Corollary 3.13 and Lemma 3.18. •

Proof ((21) and (22) in Table 5.1: monotonicities in θ, combined scheme $ES^+ - \mu$) — Finally, we ought to obtain the analogues of (17) and (18) in Table 5.1 for the combined scheme $ES^+ - \mu$. They follow quite easily. This time the sign of derivatives of the left partial sum with respect to θ is non-null and equal to (5.21) if

$$\frac{1}{\theta}\left\{\frac{\gamma_\mu^+ \times [(j+1)-(1-\lambda_\mu^+)(i+1/2)]}{(x_\mu^+ + 1)\sqrt{\lambda_\mu^+(2-\lambda_\mu^+)}} - \delta\right\} < \xi_\mu^+, \tag{5.37}$$

due to the truncation in ξ_μ^+.

The sign in (5.21) is nonpositive under the constraint (5.37), for $i,j=0,\dots,x_\mu^+$, if

$$\begin{aligned}
\delta &\le \min_{i,j=0,\dots,x_\mu^+}\left\{\frac{\gamma_\mu^+\times[(j+1)-(1-\lambda_\mu^+)(i+1/2)]}{(x_\mu^+ +1)\sqrt{\lambda_\mu^+(2-\lambda_\mu^+)}}\right\} = \underline{\delta}_{E^+-\mu}(x_\mu^+) \quad \text{and}\\
\delta &> \max_{i,j=0,\dots,x_\mu^+}\left\{\frac{\gamma_\mu^+\times[(j+1)-(1-\lambda_\mu^+)(i+1/2)]}{(x_\mu^+ +1)\sqrt{\lambda_\mu^+(2-\lambda_\mu^+)}}\right\} - \theta\xi_\mu^+\\
&= \overline{\delta}_{E^+-\mu}(x_\mu^+) - \theta\xi_\mu^+. \qquad (5.38)
\end{aligned}$$

Once again, these two conditions only make sense when $\theta > \frac{\overline{\delta}_{E^+-\mu}-\underline{\delta}_{E^+-\mu}}{\xi_\mu^+}$, and in that case we get $\overline{\delta}_{E^+-\mu} - \theta\xi_\mu^+ < \delta \le \underline{\delta}_{E^+-\mu}$.

Analogously, the sign in (5.21) is nonnegative under the constraint (5.37) in case

$$\begin{aligned}
\delta &\ge \max_{i,j=0,\dots,x_\mu^+}\left\{\frac{\gamma_\mu^+\times[(j+1)-(1-\lambda_\mu^+)(i+1/2)]}{(x_\mu^+ +1)\sqrt{\lambda_\mu^+(2-\lambda_\mu^+)}}\right\} = \overline{\delta}_{E^+-\mu} \quad \text{and}\\
\delta &> \overline{\delta}_{E^+-\mu}(x_\mu^+) - \theta\xi_\mu^+, \qquad (5.39)
\end{aligned}$$

that is, if $\delta > \overline{\delta}_{E^+-\mu}(x_\mu^+)$.

Thus, the stochastic monotone behaviours of the RL of scheme $ES^+-\mu$ in terms of θ, summarized by the expressions (21) and (22) in Table 5.1, follow from Corollary 3.13 and Lemma 3.18. •

Remark 5.10 — As we have demonstrated, the exact RL of scheme $C^+-\mu$ stochastically increases with θ if $\delta \ge \overline{\delta}_{C^+-\mu}$; this is also valid for the combined scheme $CS^+-\mu$. However, the extension of the stochastic decreasing behaviour of $RL^{\alpha}_{C^+-\mu}(\delta,\theta)$ to the combined scheme $CS^+-\mu$ requires a slight modification — it only holds for $\underline{\delta}_{C^+-\mu}-\theta\xi_\mu^+ < \delta \le \underline{\delta}_{C^+-\mu}$ and in case $\theta > (\overline{\delta}_{C^+-\mu}-\underline{\delta}_{C^+-\mu})/\xi_\mu^+ = 2h_\mu^+/\xi_\mu^+$, that is, for θ usually larger than the unity (see Table 5.2).

Similar extensions hold for the combined scheme $ES^+-\mu$, for which the stochastic decreasing behaviour occurs for rather large values of θ (also refer to Table 5.2). •

5.4 Examples

We shall not illustrate the properties alluded to in the previous sections. We proceed with a numerical investigation over the stochastic monotone behaviour with respect to θ of the RLs of the five upper one-sided schemes, in the interval $[0,\overline{\delta})$.

These examples refer to schemes $S^+-\mu$, $C^+-\mu$, $CS^+-\mu$, $E^+-\mu$ and $ES^+-\mu$, obtained by considering $n=5$, $\mu_0=0$, $\sigma_0=1$, $x_\mu^+=40$ (i.e. 41 transient states in the Markov chain approximations) and the remaining parameters equal to those in Table 5.2.

The in-control $ARL_{S^+-\mu}(0,1)$ is fixed at 500 samples which leads to $\xi_\mu^+ = \Phi^{-1}(1-1/500) = 2.87816$. The range of the decision intervals $[LCL, UCL)$ of the remaining schemes for μ was chosen in such way that, when no head start has been adopted, all five schemes are approximately matched in-control, that is, all schemes require the same average number of samples to a false alarm — thus, $ARL^0_{C+-\mu}(0,1)$, $ARL^0_{E+-\mu}(0,1)$, $ARL^0_{CS+-\mu}(0,1)$ and $ARL^0_{ES+-\mu}(0,1)$ are close to 500. In both combined schemes for μ, the *Shewhart* type constituent charts have in-control ARL equal to 1000 samples. For further details on the choice of the parameters of these schemes please refer to Chapter 6, where these and other schemes are used in the joint monitoring of μ and σ.[2]

Regarding the process parameter θ, we have to note that the set of parameters used here lead to values of $\underline{\delta}(x^+_\mu)$, $\overline{\delta}(x^+_\mu)$, $\underline{\delta}$ and $\overline{\delta}$, also in Table 5.2.

Table 5.2: Parameters, in-control ARLs, and values of $\underline{\delta}(x^+_\mu)$, $\overline{\delta}(x^+_\mu)$, $\underline{\delta}$ and $\overline{\delta}$, for the upper one-sided schemes for μ.

Scheme	Parameters	$ARL(0,1)$	$\underline{\delta}(x^+_\mu)$	$\overline{\delta}(x^+_\mu)$	$\underline{\delta}$	$\overline{\delta}$
$S^+-\mu$	$\xi^+_\mu = 2.87816$	500.000	—	—	2.87816	2.87816
$C^+-\mu$	$h^+_\mu = 4.4456$, $k^+_\mu = 0.50$	500.021	-3.782956	4.891385	-3.9456	4.9456
$CS^+-\mu$	$\xi^+_\mu = 3.09023$, $h^+_\mu = 4.9854$, $k^+_\mu = 0.50$	500.020	-4.303007*	5.424602	-4.4854**	5.4854
$E^+-\mu$	$\gamma^+_\mu = 2.8116$, $\lambda^+_\mu = 0.134$	500.097	-4.672742	5.563324	-4.869263	5.622705
$ES^+-\mu$	$\xi^+_\mu = 3.09023$, $\gamma^+_\mu = 3.0016$, $\lambda^+_\mu = 0.134$	500.094	-4.988513$^{(*)}$	5.939278	-5.198314$^{(**)}$	6.00267

for θ larger than: * 3.14786; ** 3.22655; (*) 3.536236; and (**) 3.624642.

Figure 5.3 presents three ARL curves for values of θ ranging from the in-control situation to a 500% increase in the process standard deviation.

These plots show that, for $\delta = \overline{\delta}_{S+-\mu} - 0.1$, $\delta = \overline{\delta}_{S+-\mu}$ and $\delta = \overline{\delta}_{S+-\mu} + 0.1$ (recall that $\overline{\delta}_{S+-\mu} = \underline{\delta}_{S+-\mu} = \xi^+_\mu$), $ARL_{S+-\mu}(\delta,\theta)$ is (respectively) a decreasing, constant and increasing function of θ, as a consequence of the stochastic monotonicity results (5.5)–(5.7).

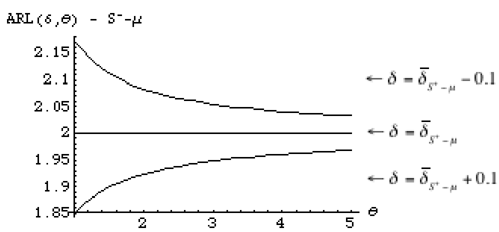

Figure 5.3: Plots of $ARL_{S+-\mu}(\delta,\theta)$, for $\delta = \overline{\delta}_{S+-\mu} - 0.1$, $\delta = \overline{\delta}_{S+-\mu}$, $\delta = \overline{\delta}_{S+-\mu} + 0.1$ and $\theta \in [1,5]$.

The plots on the right of Figure 5.4 portray a completely different scenario for $\delta = \overline{\delta} - 0.1$, $\delta = \overline{\delta}$ and $\delta = \overline{\delta} + 0.1$ and schemes $C^+-\mu$, $CS^+-\mu$, $E^+-\mu$ and $ES^+-\mu$.

The function $ARL^0(\delta,\theta)$ of all these four schemes increases with θ when $\delta = \overline{\delta} + 0.1$ and $\delta = \overline{\delta}$; recall that both monotone behaviours are a direct consequence of results (8), (13), (18) and (22) in Table 5.1. However, $ARL^0(\delta,\theta)$ has no monotone behaviour in terms of θ, for $\delta = \overline{\delta}_{E+-\mu} - 0.1$.

[2]Note that in Morais and Pacheco (2000c) the examples only refer to a pair of matched upper one-sided schemes $S^+-\mu$ and $E^+-\mu$ with in-control ARL equal to 500.567 samples.

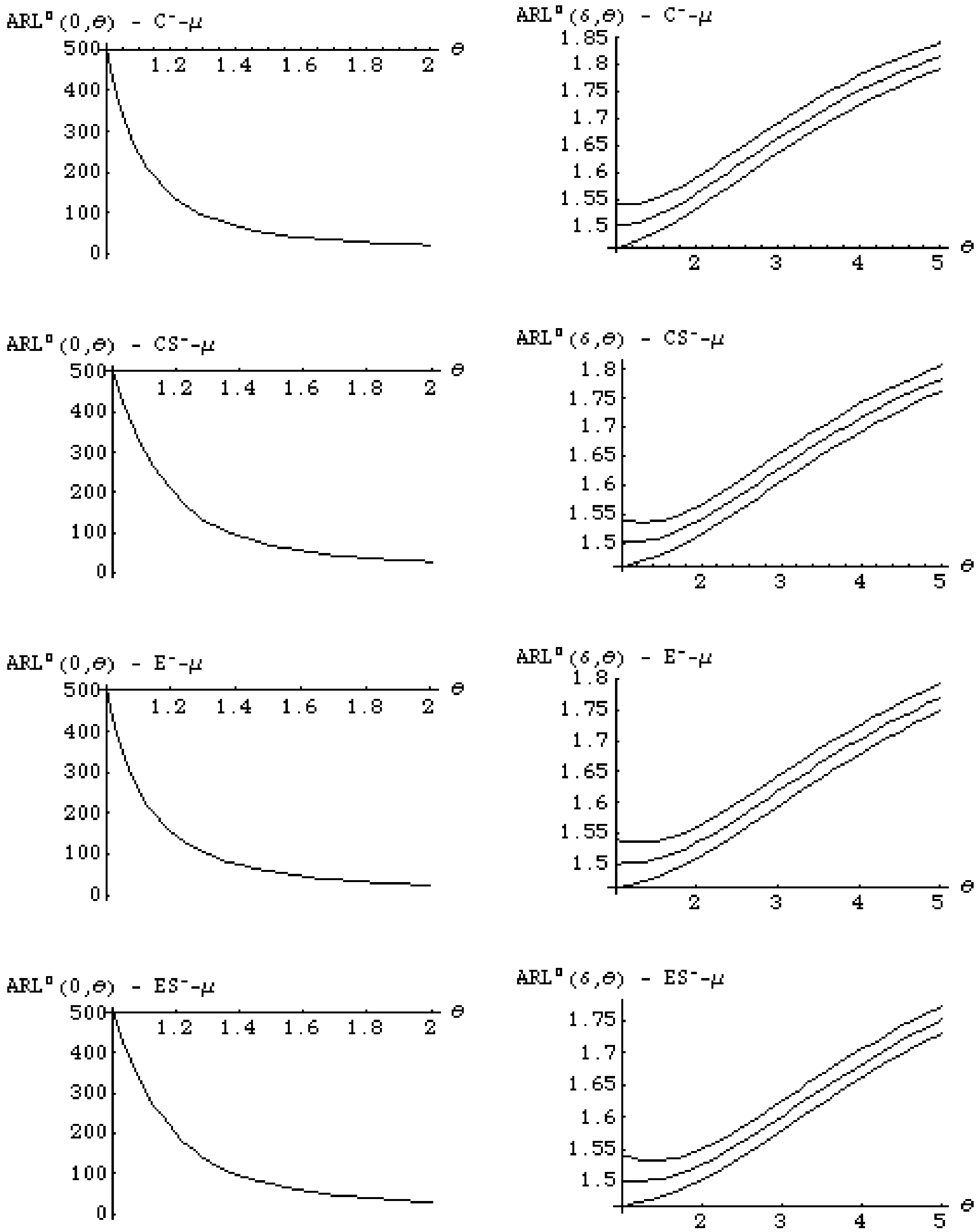

Figure 5.4: Plots of $ARL^0(\delta,\theta)$, for schemes $C^+ - \mu$, $E^+ - \mu$, $CS^+ - \mu$ and $ES^+ - \mu$. On the left: $\delta = 0$ and $\theta \in [1,2]$. On the right: $\delta = \overline{\delta} - 0.1$, $\delta = \overline{\delta}$ and $\delta = \overline{\delta} + 0.1$ from top to bottom, and $\theta \in [1,5]$.

A plausible explanation for this nonmonotone behaviour as the level of the standard deviation increases when δ is fixed and belongs to $(\underline{\delta}, \overline{\delta})$ is the fact that $\delta \leq \underline{\delta}$ ($\delta \geq \overline{\delta}$) is not a necessary condition for a stochastic decreasing (increasing) behaviour of $RL^{\alpha}(\delta, \theta)$ in terms of θ.

The plots from the left hand side of Figure 5.4 suggest that $ARL^0(0, \theta)$ decreases with θ for any of the four control schemes — regardless of the fact that the stochastic behaviour of $RL^{\alpha}(\delta, \theta)$ is unknown, for any fixed δ in the interval $[0, \overline{\delta})$.

Table 5.3: Percentage points of $RL^0(\delta, \theta)$ and $ARL^0(\delta, \theta)$ values for schemes $C^+ - \mu$, $CS^+ - \mu$, $E^+ - \mu$ and $ES^+ - \mu$.

	$C^+ - \mu$						$CS^+ - \mu$					
	$\delta = 0.1$			$\delta = 2$			$\delta = 0.1$			$\delta = 2$		
RL perc. points	θ			θ			θ			θ		
	1	2	3	1	2	3	1	2	3	1	2	3
5%	18	3	1	2	1	1	20	3	2	2	2	1
25%	75	7	3	3	2	2	90	8	4	3	2	2
Median	174	14	7	3	3	3	210	17	8	4	3	3
75%	342	25	12	**4**	**5**	**4**	415	30	14	**5**	**5**	**5**
90%	563	40	19	**5**	**6**	**6**	686	49	22	**6**	**7**	**7**
95%	731	51	24	**6**	**8**	**8**	891	63	28	**6**	**8**	**9**
ARL	247.9	18.6	8.8	**3.6**	**3.7**	**3.4**	300.9	22.7	10.2	**4.0**	**4.0**	**3.8**

	$E^+ - \mu$						$ES^+ - \mu$					
	$\delta = 0.1$			$\delta = 2$			$\delta = 0.1$			$\delta = 2$		
RL perc. points	θ			θ			θ			θ		
	1	2	3	1	2	3	1	2	3	1	2	3
5%	19	3	2	2	2	1	21	4	2	3	2	1
25%	74	8	4	3	2	2	90	9	4	3	3	2
Median	168	16	8	4	3	3	208	18	8	4	4	3
75%	329	28	14	**5**	**5**	**5**	410	33	15	**5**	**5**	**5**
90%	541	45	21	**6**	**7**	**7**	677	53	24	**6**	**7**	**7**
95%	702	58	27	**6**	**8**	**9**	878	68	31	**7**	**9**	**9**
ARL	239.0	21.2	10.1	9.6	7.6	5.9	297.1	24.8	11.3	10.7	8.3	6.4

Our additional numerical investigations yield, in any case, similar conclusions when a 50% head start has been given to all of the four schemes with phase-type distributed *RL*s. Nevertheless, we were able to obtain increasing functions $ARL^0_{C^+-\mu}(\overline{\delta}_{C^+-\mu} - 0.1, \theta)$ and $ARL^0_{CS^+-\mu}(\overline{\delta}_{C^+-\mu} - 0.1, \theta)$ for a different constelation of parameters: $h^+_\mu = 2.3807$, $k^+_\mu = 1.00$, and $\xi^+_\mu = 3.09023$, $h^+_\mu = 2.2515$, $k^+_\mu = 1.00$.

All these numerical results concerning $ARL^0(\delta,\theta)$, for $\delta = 0, \overline{\delta} - 0.1$, suggest a close investigation of the percentage points of $RL^0(\delta,\theta)$ for some values of θ, and of δ in the interval $[0, \overline{\delta})$.

The results in Table 5.3 portray an interesting situation: most of the $ARL^0(\delta,\theta)$ values decrease with θ but the same does not occur with the percentage points in **bold**. In fact, the percentage points of $RL^0(\delta,\theta)$, for $\theta = 1, 2, 3$ and for $\delta = 2$ for schemes $C^+ - \mu$, $CS^+ - \mu$, $E^+ - \mu$ and $ES^+ - \mu$, exhibit decreasing, constant, increasing and nonmonotone behaviours. Thus, the phase-type distributed RLs do not seem to make a sharp transition from being stochastically decreasing in θ (for $\delta \leq \underline{\delta}$) to being stochastically increasing in θ (when $\delta \geq \overline{\delta}$), as the run length of the upper one-sided $S^+ - \mu$ scheme, $RL_{S+-\mu}(\delta,\theta)$.

Table 5.4: Percentage points of $RL^0(0,\theta)$ and $ARL^0(0,\theta)$ values for schemes $C^+ - \mu$, $CS^+ - \mu$, $E^+ - \mu$ and $ES^+ - \mu$.

RL perc. points	$C^+ - \mu$						$CS^+ - \mu$					
	θ						θ					
	1.01	1.1	1.25	1.5	2	3	1.01	1.1	1.25	1.5	2	3
5%	29	17	10	6	3	1	28	20	12	7	3	2
25%	137	75	36	17	8	4	140	91	48	22	9	4
Median	322	173	82	37	16	7	331	213	109	48	19	8
75%	638	341	160	71	29	13	659	421	213	93	36	15
90%	1056	564	263	116	47	20	1092	696	351	152	58	23
95%	1372	732	341	149	60	26	1420	905	455	196	74	30
ARL	461.5	247.5	116.5	52.0	21.4	9.4	476.5	304.8	154.7	68.0	26.5	11.0

RL perc. points	$E^+ - \mu$						$ES^+ - \mu$					
	θ						θ					
	1.01	1.1	1.25	1.5	2	3	1.01	1.1	1.25	1.5	2	3
5%	30	19	11	7	4	2	29	21	13	8	4	2
25%	138	79	40	20	9	4	140	93	50	24	11	5
Median	324	180	89	42	18	8	332	217	114	52	21	9
75%	640	355	173	79	33	15	660	429	222	100	39	16
90%	1059	585	283	129	53	23	1093	709	366	163	63	26
95%	1375	759	367	166	69	30	1421	920	475	210	82	33
ARL	463.7	257.6	126.0	58.2	24.7	10.9	477.3	310.6	161.7	73.0	29.2	12.2

We close this section with the analysis of a particularly important scenario — $(\delta,\theta) \in \{0\} \times (1, +\infty)$ — that is closely related to misleading signals of type III, defined in the next chapter.

Table 5.4 has several percentage points of the $RL^0(0,\theta)$ of all four charts, for $\theta = 1.01, 1.1, 1.25, 1.5, 2, 3$. According to these tables $RL^0(0,\theta)$ seems to stochastically decrease as the process standard deviation increases.[3]

5.5 Further stochastic properties

In this section we shall try to sharpen the stochastic monotonicity properties with regard to the initial state, derived in Sections 5.2 and 5.3. This naturally requires the investigation of the features of the transition matrix, which will be summarized in Theorem 5.14. For that matter we shall need the following lemma concerning the decreasing likelihood ratio (or PF_2) character of the standard normal distribution.

Proposition 5.11 — *Let* $X =_{st} N(0,1)$. *Then* $X \in DLR$, *and therefore* $X \in DRHR$.

Proof — For the lack of a reference, we include a proof of both results. The density function and the distribution function of X are represented by $\phi(x) = (2\pi)^{-1/2}\exp^{-x^2/2}$ and $\Phi(x)$, respectively. Therefore $\frac{d^2\ log\phi(x)}{dx^2} = -1 \leq 0$, that is, $\phi(x)$ is log-concave; in other words $\phi(x) \in PF_2$ (Barlow and Proschan (1975, p. 76)) or $X \in DLR$.

As mentioned by Shaked and Shanthikumar (1994, p. 12) the notion of hazard rate (and certainly the one of reversed hazard rate) may be also defined for general random variables although usually applied to nonnegative random variables. Let $\lambda(x) = \phi(x)/[1 - \Phi(x)]$, $\overline{\lambda}(x) = \phi(x)/\Phi(x)$, be the hazard rate function and the reversed hazard rate function of X, respectively.

As a consequence of Theorems 1.C.22, 1.C.1 and 1.B.19 of Shaked and Shanthikumar (1994, pp. 40, 28 and 23) and the symmetry of $\phi(x)$, we successively get

$$\begin{aligned} X \in DLR \quad &\Leftrightarrow \quad (X - t \,|\, X > t) \geq_{lr} (X - t' \,|\, X > t'), \text{ for all } t \leq t' \\ &\Rightarrow \quad (X - t \,|\, X > t) \geq_{hr} (X - t' \,|\, X > t'), \text{ for all } t \leq t' \\ &\Leftrightarrow \quad X \in IHR \\ &\Leftrightarrow \quad \lambda(x) \uparrow_x \text{ over } I\!R \\ &\Leftrightarrow \quad \overline{\lambda}(x) = \lambda(-x) \downarrow_x \text{ over } I\!R. \end{aligned} \tag{5.40}$$

Thus, the reversed hazard rate of X, $\phi(x)/\Phi(x)$, decreases with x. •

We have just proved that the normal distribution has *DRHR*. If we add to this property result (2.106), Proposition 2.20 would immediately suggest that the matrix $\mathbf{P}_{C^+-\mu}(\delta,\theta)$ belongs to the class of stochastically monotone matrices in the reversed hazard rate sense. This is in fact the case although we are dealing now with the discretized version of normal data.

Lemma 5.12 — *Let* $\mathbf{P}_{C^+-\mu}(\delta,\theta;x_\mu^+)$ *be the transition matrix ruling the approximating Markov chain associated to the* $C^+ - \mu$ *scheme. Then*

$$\mathbf{P}_{C^+-\mu}(\delta,\theta;x_\mu^+) \in \mathcal{M}_{rh}. \tag{5.41}$$

[3]This result will provide an interesting explanation for the nonmonotonous character of the probability of a misleading signal of type III, introduced in the following chapter.

Proof — From the decreasing behaviour of $\overline{\lambda}(x)$ for $X =_{st} N(0,1)$, we conclude that, for any $x \in I\!R$ and $\xi > 0$,

$$\frac{d\ \Phi(x+\xi)/\Phi(x)}{dx} = [\overline{\lambda}(x+\xi) - \overline{\lambda}(x)] \times \frac{\Phi(x+\xi)}{\Phi(x)} \leq 0, \tag{5.42}$$

that is,

$$\Phi(x+\xi)/\Phi(x) \downarrow_x \text{ over } I\!R, \text{ for any } \xi > 0. \tag{5.43}$$

Thus,

$$\begin{bmatrix} \Phi(x_1) & \Phi(x_2) & \Phi(x_3) & \ldots \\ \Phi(x_0) & \Phi(x_1) & \Phi(x_2) & \ldots \end{bmatrix} \in TP_2, \tag{5.44}$$

where $x_i = x_0 + i\xi, i = 0, 1, \ldots$. By capitalizing on the fact that each pair of consecutive rows of $\mathbf{P}_{C^+-\mu}(\delta, \theta; x_\mu^+)\mathbf{U}^\top$ corresponds to a right-truncated version of the matrix in (5.44) supplemented by a column of two ones, we conclude that $\mathbf{P}_{C^+-\mu}(\delta, \theta; x_\mu^+)\mathbf{U}^\top \in TP_2$, and, thus, prove the result. •

We shall need the next lemma in order to strenghten or extend result (5.41) to schemes $CS^+ - \mu$, $E^+ - \mu$ and $ES^+ - \mu$.

Lemma 5.13 — *The following conditions involving* 2×2 *minors hold, for any* $x \in I\!R$, $\Delta > 0$ *and* $0 \leq \epsilon \leq \Delta$*:*

$$\begin{vmatrix} \Phi(x) & \Phi(x+\Delta) - \Phi(x) \\ \Phi(x-\Delta+\epsilon) & \Phi(x+\epsilon) - \Phi(x-\Delta+\epsilon) \end{vmatrix} \geq 0, \tag{5.45}$$

$$\begin{vmatrix} \Phi(x) - \Phi(x-\Delta) & \Phi(x+\Delta) - \Phi(x) \\ \Phi(x-\Delta+\epsilon) - \Phi(x-2\Delta+\epsilon) & \Phi(x+\epsilon) - \Phi(x-\Delta+\epsilon) \end{vmatrix} \geq 0. \tag{5.46}$$

Proof — Condition (5.45) is equivalent to

$$\Phi(x)\Phi(x+\epsilon) - \Phi(x+\Delta)\Phi(x-\Delta+\epsilon) \geq 0, \tag{5.47}$$

which proves to be true since the decreasing behaviour of $\Phi(x+\xi)/\Phi(x)$ in x (for any $\xi > 0$) guarantees that, for $0 \leq \epsilon \leq \Delta$,

$$\frac{\Phi(x+\epsilon)}{\Phi(x+\epsilon-\Delta)} \geq \frac{\Phi(x+\Delta)}{\Phi(x)}. \tag{5.48}$$

On the other hand, the 2×2 minor in (5.46), is equal to

$$\int_{x-\Delta}^{x} \int_{x-\Delta+\epsilon}^{x+\epsilon} \phi(u)\phi(v)dvdu - \int_{x-2\Delta+\epsilon}^{x-\Delta+\epsilon} \int_{x}^{x+\Delta} \phi(u)\phi(v)dvdu$$

$$= \int_0^\Delta \int_0^\Delta [\phi(x-\Delta+a)\phi(x-\Delta+\epsilon+b) - \phi(x-2\Delta+\epsilon+a)\phi(x+b)]dbda$$

$$= \int_0^\Delta \int_0^\Delta [\phi(x-\Delta+a)\phi(x-\Delta+\epsilon+b)\left\{1 - \exp[-(\epsilon-\Delta)(a-b-\Delta)]\right\} dbda. \tag{5.49}$$

Under the conditions of Lemma 5.13 this 2×2 minor is nonnegative since $-(\epsilon-\Delta)(a-b-\Delta) \leq 0$ and, thus, the integrand is nonnegative; this proves (5.46). •

Theorem 5.14 — *Let $\mathbf{P}_{*^+-\mu}(\delta,\theta;x_\mu^+)$ be the approximating stochastic matrix ruling the Markov chain associated to the upper one-sided scheme "$*^+-\mu$". Then*

$$\mathbf{P}_{*^+-\mu}(\delta,\theta;x_\mu^+)\in\mathcal{M}_{lr},\ \textit{for}\ *^+-\mu=C^+-\mu,E^+-\mu \tag{5.50}$$

$$\mathbf{P}_{*^+-\mu}(\delta,\theta;x_\mu^+)\in\mathcal{M}_{rh},\ \textit{for}\ *^+-\mu=CS^+-\mu,ES^+-\mu. \tag{5.51}$$

Proof ($C^+-\mu$) — By using a similar argumentation as in the proof of Lemma 5.12, $\mathbf{P}_{C^+-\mu}(\delta,\theta;x_\mu^+)\in\mathcal{M}_{lr}$ if we guarantee that

$$\begin{bmatrix}\Phi(x_1) & \Phi(x_2)-\Phi(x_1) & \Phi(x_3)-\Phi(x_2) & \dots\\ \Phi(x_0) & \Phi(x_1)-\Phi(x_0) & \Phi(x_2)-\Phi(x_1) & \dots\end{bmatrix}\in TP_2, \tag{5.52}$$

where $x_i=x_{i-1}+\Delta$, $i=1,2,\ldots$.

This result immediately follows since the 2×2 minors of the matrix in (5.52)

$$\begin{vmatrix}\Phi(x_1) & \Phi(x_2)-\Phi(x_1)\\ \Phi(x_0) & \Phi(x_1)-\Phi(x_0)\end{vmatrix} \tag{5.53}$$

$$\begin{vmatrix}\Phi(x_{i+1})-\Phi(x_i) & \Phi(x_{i+2})-\Phi(x_{i+1})\\ \Phi(x_i)-\Phi(x_{i-1}) & \Phi(x_{i+1})-\Phi(x_i)\end{vmatrix},\ i=1,2,\ldots, \tag{5.54}$$

are all nonnegative by applying Lemma 5.13 and considering the range of the intervals used in the discretization procedure equal to $\Delta=h_\mu^+/(x_\mu^++1)$, and $\epsilon=0$. •

Proof ($CS^+-\mu$) — By benefiting from the proof of result (4.21) of Theorem 4.4 and from the fact that $\mathbf{P}_{C^+-\mu}(\delta,\theta;x_\mu^+)\in\mathcal{M}_{lr}$, we similarly conclude that the stochastic matrix $\mathbf{P}_{CS^+-\mu}(\delta,\theta;x_\mu^+)$ is stochastically monotone in the reversed hazard rate sense. •

Proof ($E^+-\mu$) — The nontruncated generic form of two consecutive rows of $\mathbf{P}_{E^+-\mu}(\delta,\theta;x_\mu^+)$ is as follows

$$\begin{bmatrix}\Phi(x_1) & \Phi(x_2)-\Phi(x_1) & \Phi(x_3)-\Phi(x_2) & \dots\\ \Phi(x_1-\Delta+\epsilon) & \Phi(x_2-\Delta+\epsilon)-\Phi(x_1-\Delta+\epsilon) & \Phi(x_3-\Delta+\epsilon)-\Phi(x_2-\Delta+\epsilon) & \dots\end{bmatrix}, \tag{5.55}$$

where $x_i=x_{i-1}+\Delta$, $i=1,2,\ldots$, with $\Delta=\frac{\gamma_\mu^+}{(x_\mu^++1)\sqrt{\lambda_\mu^+(2-\lambda_\mu^+)}}$ and $\epsilon=\lambda_\mu^+\Delta$. Thus, by considering Δ and ϵ as stated and applying result (5.46) from Lemma 5.13, we prove that $\mathbf{P}_{E^+-\mu}(\delta,\theta;x_\mu^+)\in TP_2$. •

Proof ($ES^+-\mu$) — This proof is similar to the one of $\mathbf{P}_{CS^+-\mu}(\delta,\theta;x_\mu^+)\in\mathcal{M}_{rh}$. •

Theorem 5.15 — *The approximating* RL*s of schemes $C^+-\mu$, $E^+-\mu$, $CS^+-\mu$ and $ES^+-\mu$ have the following stochastic properties:*

$$RL^\alpha_{C^+-\mu}(\delta,\theta;x_\mu^+),RL^\alpha_{E^+-\mu}(\delta,\theta;x_\mu^+)\downarrow_{lr}\ \textit{with}\ \alpha \tag{5.56}$$

$$RL^\alpha_{CS^+-\mu}(\delta,\theta;x_\mu^+),RL^\alpha_{ES^+-\mu}(\delta,\theta;x_\mu^+)\downarrow_{hr}\ \textit{with}\ \alpha. \tag{5.57}$$

Proof — The mere application of Theorem 3.8 and Theorem 5.14 leads to all the results. •

Corollary 5.16 — *Results (5.56) and (5.57) hold for the exact* RL*s of schemes* $C^+ - \mu$, $E^+ - \mu$, $CS^+ - \mu$ *and* $ES^+ - \mu$.

Proof — Apply the closure property under the limit operation stated in Lemma 3.18. •

The next example provides a few values of the alarm rate and equilibrium rate functions to illustrate the results in Theorem 5.15.

Example 5.17 — Let $RL^{\alpha}_{C^+-\mu}(\delta,\theta)$, $RL^{\alpha}_{CS^+-\mu}(\delta,\theta)$, $RL^{\alpha}_{E^+-\mu}(\delta,\theta)$ and $RL^{\alpha}_{ES^+-\mu}(\delta,\theta)$ now denote the approximate *RL*s of the Markov-type upper one-sided schemes $C^+ - \mu$, $CS^+ - \mu$, $E^+ - \mu$ and $ES^+ - \mu$ (respectively), with a $\alpha \times 100\%$ head start. Moreover consider the parameters given at the beginning of Section 5.4.

The values in Table 5.5 and Table 5.6 illustrate the increases in the alarm rates of the four schemes after the adoption of a 50% head start, in agreement with Theorem 5.15. As in Example 4.9, this increase steadily converges to zero since

$$\lim_{m\to+\infty} \lambda_{RL^{\alpha}_{*^+-\mu}(\delta,\theta)}(m) = \xi(\delta,\theta), \tag{5.58}$$

for any $\alpha \in [0,1)$, where $\xi(\delta,\theta)$ is the maximal real eigenvalue of the substochastic matrix $\mathbf{Q}_{*^+-\mu}(\delta,\theta;x^+_{\mu})$ concerning scheme $*^+ - \mu$, as seen in Figures 5.5–5.8.

We can also add some comments on the monotone behaviour of the alarm rate functions of all these *RL*s as we did in Example 4.9. For example, when no head start has been given to schemes $C^+ - \mu$, $CS^+ - \mu$, $E^+ - \mu$, and $ES^+ - \mu$, the alarm rate functions of the associated *RL*s increase as we collect more samples. Moreover, the *RL*s of most schemes with a 50% head start are not *IHR*, as apparent in the graphs on the right hand side of Figures 5.5–5.8.

The numerical values in Table 5.5 and Table 5.6 also show that after having given such a head start the equilibrium rate functions of the *RL*s of schemes $C^+ - \mu$ and $E^+ - \mu$ increase, a consequence of the decreasing behaviour of these *RL*s with regard to α, in the likelihood ratio sense. This increase essentially mean that giving a head start yields an increase in the relative sequential decrease in the probability that the schemes triggers a signal, $1 - 1/r_{RL}(m)$. However, this property does not hold for schemes $CS^+ - \mu$ and $ES^+ - \mu$, as illustrated by the two last columns of Table 5.5 and Table 5.6; thus, we can assert that both $RL^{\alpha}_{CS^+-\mu}(\delta,\theta)$ and $RL^{\alpha}_{ES^+-\mu}(\delta,\theta)$ do not decrease with α in the likelihood ratio sense. •

An increase in the process standard deviation is certainly a cause of concern. Quite apart from the relevance of a standard deviation change in its own right, a dilation in this scale parameter changes the stochastic behaviour of the *RL* of a scheme for the process mean.

In fact, it was proved that the run length of the upper one-sided $\bar{X}$ control scheme $(S^+ - \mu)$ can be a decreasing, a constant, or an increasing function of θ (in the usual sense), depending on the fixed value of δ.

Table 5.5: ARLs, alarm rates and equilibrium rates of the RLs of schemes $C^+ - \mu$ and $CS^+ - \mu$, for $\alpha \times 100\% = 0\%, 50\%$ head starts.

	$\lambda_{RL}(m)$				$r_{RL}(m)$	
m	$RL^{0\%}_{C^+-\mu}(0,1)$	$RL^{50\%}_{C^+-\mu}(0,1)$	$RL^{0\%}_{C^+-\mu}(1,1)$	$RL^{50\%}_{C^+-\mu}(1,1)$	$RL^{0\%}_{C^+-\mu}(0,1)$	$RL^{50\%}_{C^+-\mu}(0,1)$
1	0.000001	0.003237	0.000050	0.042462	—	—
2	0.000069	0.010014	0.008228	0.166623	0.007232	0.324255
3	0.000346	0.010196	0.042735	0.209693	0.200194	0.992045
4	0.000731	0.008403	0.085672	0.218102	0.473420	1.225965
5	0.001088	0.006614	0.120316	0.216824	0.671807	1.281159
10	0.001881	0.002701	0.185221	0.202120	0.966266	1.124782
20	0.002017	0.002033	0.196988	0.197540	1.001380	1.005169
30	0.002020	0.002020	0.197364	0.197382	1.002012	1.002084
100	0.002020	0.002020	0.197376	0.197376	1.002024	1.002024
ARL	500.021	476.580	9.164	5.761	—	—

	$\lambda_{RL}(m)$				$r_{RL}(m)$	
m	$RL^{0\%}_{CS^+-\mu}(0,1)$	$RL^{50\%}_{CS^+-\mu}(0,1)$	$RL^{0\%}_{CS^+-\mu}(1,1)$	$RL^{50\%}_{CS^+-\mu}(1,1)$	$RL^{0\%}_{CS^+-\mu}(0,1)$	$RL^{50\%}_{CS^+-\mu}(0,1)$
1	0.001000	0.001383	0.001000	0.023147	—	—
2	0.001001	0.006759	0.003181	0.130277	1.000347	0.204853
3	0.001056	0.007922	0.022413	0.184327	0.948952	0.858976
4	0.001195	0.007136	0.056353	0.200748	0.884502	1.118928
5	0.001365	0.005998	0.089745	0.204036	0.876235	1.198345
10	0.001879	0.002807	0.166770	0.193255	0.973301	1.117519
20	0.002008	0.002038	0.185496	0.186951	1.001127	1.007224
30	0.002012	0.002013	0.186485	0.186568	1.001988	1.002184
100	0.00201	0.002012	0.186545	0.186545	1.002016	1.002016
ARL	500.020	482.516	10.216	6.328	—	—

As for the run length of the remaining upper one-sided schemes — $C^+ - \mu$, $E^+ - \mu$, $CS^+ - \mu$ and $ES^+ - \mu$ —, sufficient conditions were established for this random variable to increase (decrease) stochastically as the process standard deviation grows, implying that this chart becomes progressively less (more) sensitive to the same shift in the process mean.

Some numerical investigations carried on this chapter show that the phase-type run length of the upper one-sided schemes $C^+ - \mu$, $E^+ - \mu$, $CS^+ - \mu$ and $ES^+ - \mu$ can exhibit a nonmonotonic stochastic behaviour in θ for certain values of δ.

These properties are due to the fact that these individual schemes for μ are not tailored for the detection of shifts in σ. And the differences between the behaviours of the geometric RL and the phase-type RLs, in terms of θ, are a natural consequence of dealing, respectively, with independent and dependent summary statistics.

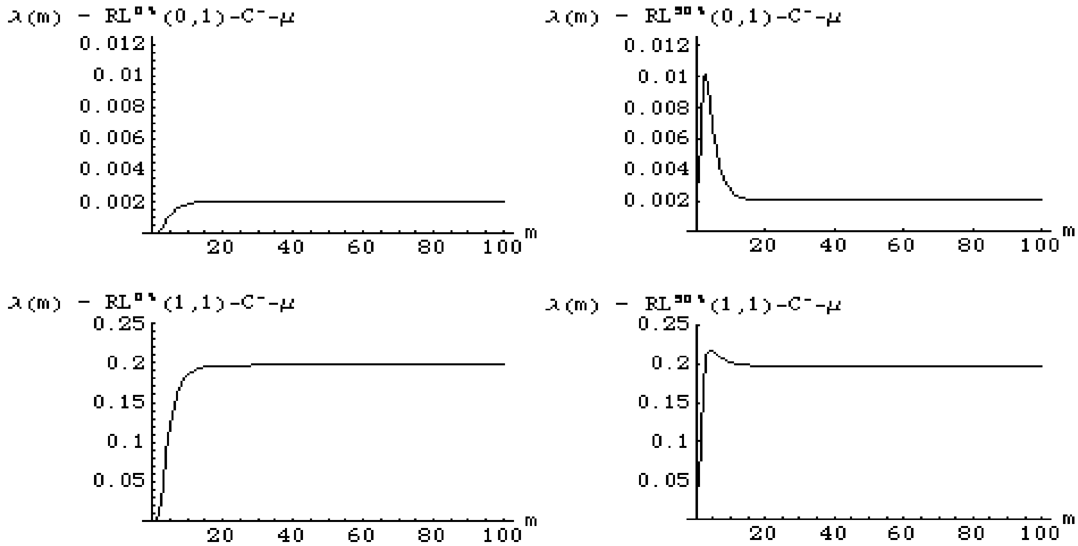

Figure 5.5: Alarm rates of $RL^{\alpha}_{C^{+}-\mu}(0,1)$ and $RL^{\alpha}_{C^{+}-\mu}(1,1)$, for $\alpha \times 100\% = 0\%$ (on the left) and $\alpha \times 100\% = 50\%$ (on the right).

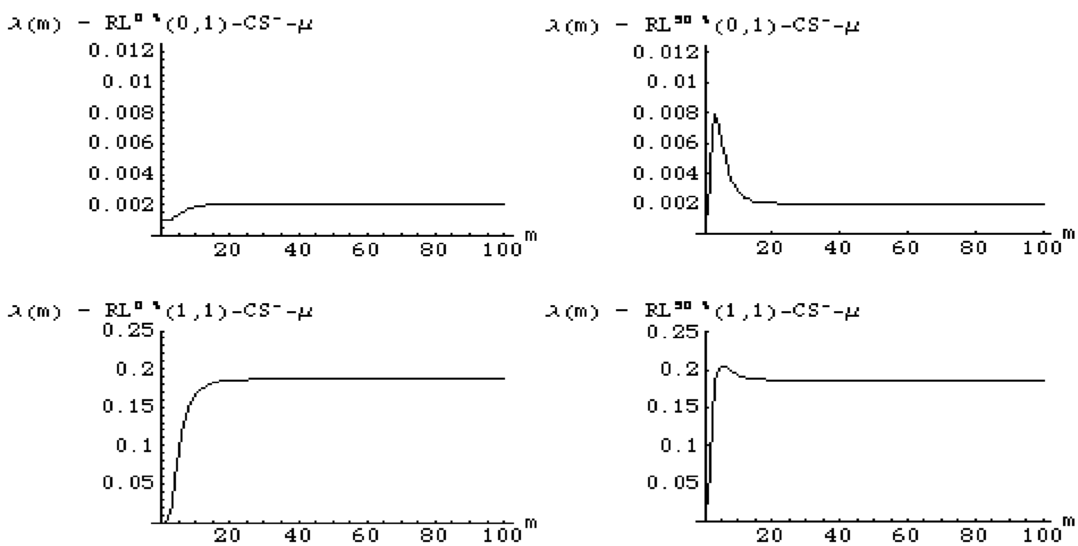

Figure 5.6: Alarm rates of $RL^{\alpha}_{CS^{+}-\mu}(0,1)$ and $RL^{\alpha}_{CS^{+}-\mu}(1,1)$, for $\alpha \times 100\% = 0\%$ (on the left) and $\alpha \times 100\% = 50\%$ (on the right).

Table 5.6: ARLs, alarm rates and equilibrium rates of the RLs of schemes $E^+ - \mu$ and $ES^+ - \mu$, for $\alpha \times 100\% = 0\%, 50\%$ head starts.

	$\lambda_{RL}(m)$				$r_{RL}(m)$	
m	$RL^{0\%}_{E^+-\mu}(0,1)$	$RL^{50\%}_{E^+-\mu}(0,1)$	$RL^{0\%}_{E^+-\mu}(1,1)$	$RL^{50\%}_{E^+-\mu}(1,1)$	$RL^{0\%}_{E^+-\mu}(0,1)$	$RL^{50\%}_{E^+-\mu}(0,1)$
1	0.000000	0.000716	0.000002	0.014332	—	—
2	0.000013	0.003740	0.002554	0.097783	0.001037	0.191608
3	0.000124	0.004882	0.024212	0.152818	0.103295	0.768992
4	0.000367	0.004744	0.064038	0.176760	0.336491	1.034228
5	0.000672	0.004263	0.103045	0.187671	0.546769	1.117923
10	0.001693	0.002568	0.185644	0.199420	0.932163	1.077627
20	0.002011	0.002051	0.200405	0.200781	0.999269	1.006436
30	0.002025	0.002027	0.200807	0.200818	1.001907	1.002230
100	0.002026	0.002026	0.200819	0.200819	1.002030	1.002030
ARL	500.047	486.277	9.610	6.798	—	—

	$\lambda_{RL}(m)$				$r_{RL}(m)$	
m	$RL^{0\%}_{ES^+-\mu}(0,1)$	$RL^{50\%}_{ES^+-\mu}(0,1)$	$RL^{0\%}_{ES^+-\mu}(1,1)$	$RL^{50\%}_{ES^+-\mu}(1,1)$	$RL^{0\%}_{ES^+-\mu}(0,1)$	$RL^{50\%}_{ES^+-\mu}(0,1)$
1	0.001000	0.001000	0.001000	0.008119	—	—
2	0.001000	0.002754	0.001681	0.072191	1.001001	0.363471
3	0.001019	0.003744	0.013588	0.124273	0.982640	0.737562
4	0.001100	0.003793	0.042615	0.149588	0.927100	0.990758
5	0.001231	0.003540	0.076083	0.161943	0.894356	1.075691
10	0.001791	0.002428	0.160007	0.176447	0.961360	1.058632
20	0.002003	0.002034	0.177804	0.178359	0.999928	1.005800
30	0.002014	0.002016	0.178402	0.178421	1.001908	1.002222
100	0.002015	0.002015	0.178424	0.178424	1.002019	1.002019
ARL	500.044	491.560	10.740	7.694	—	—

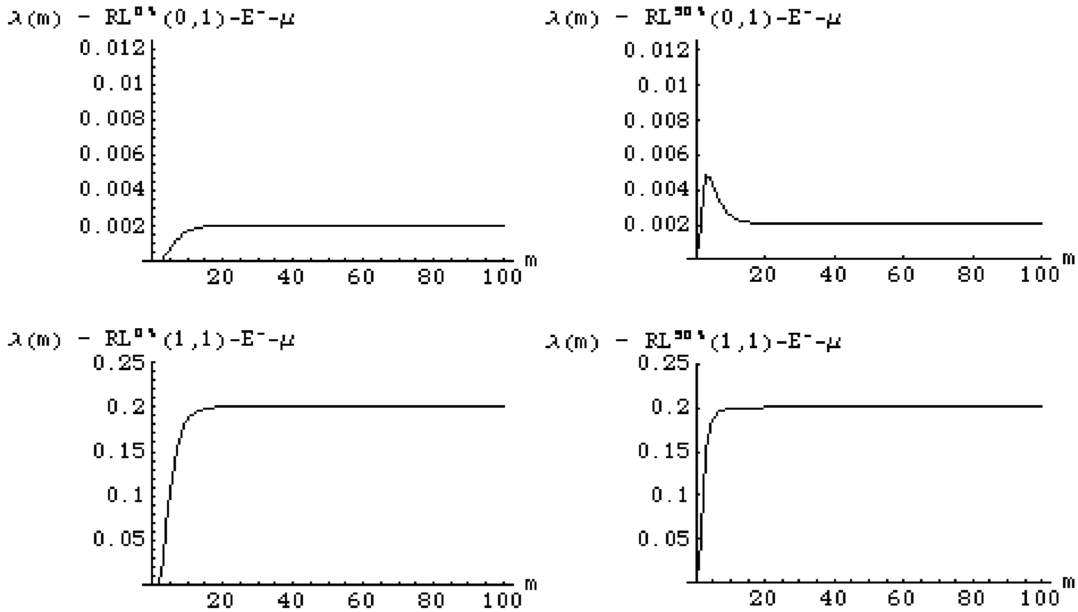

Figure 5.7: Alarm rates of $RL^{\alpha}_{E+-\mu}(0,1)$ and $RL^{\alpha}_{E+-\mu}(1,1)$, for $\alpha \times 100\% = 0\%$ (on the left) and $\alpha \times 100\% = 50\%$ (on the right).

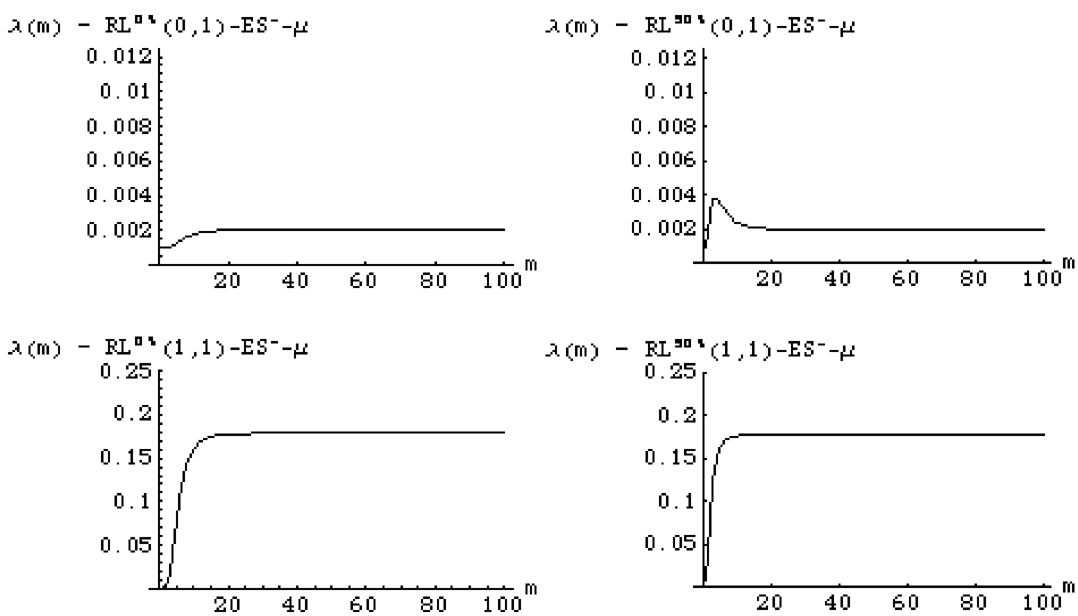

Figure 5.8: Alarm rates of $RL^{\alpha}_{ES+-\mu}(0,1)$ and $RL^{\alpha}_{ES+-\mu}(1,1)$, for $\alpha \times 100\% = 0\%$ (on the left) and $\alpha \times 100\% = 50\%$ (on the right).

Chapter 6

Misleading signals in joint schemes for μ and σ

Control schemes are also widely used as process monitoring tools to detect simultaneous changes in the process mean μ and in its standard deviation σ which can indicate a deterioration in quality. The joint monitoring of these two parameters can be achieved by running what is grandly termed a joint (or combined) scheme.

Joint schemes for μ and σ — when the quality characteristic has a normal distribution and the data is independent — have received a great deal of attention in the quality control literature. The several joint schemes that have been proposed and studied can be divided in two broad and distinct categories:

- the joint schemes which make use of one control chart for an univariate summary statistic (Chengalur, Arnold and Reynolds Jr. (1989), Domangue and Patch (1991)) or a bivariate summary statistic (Takahashi (1989));
- and the popular joint schemes that result from running simultaneously two individual control charts — a chart for μ and another one for σ (Crowder (1987b), Saniga (1989), Gan (1989, 1995), St. John and Bragg (1991), Morais (1998), Morais and Pacheco (2000a)).

This last category comprises the joint schemes in Table 6.1 that are briefly described in the next paragraphs.

Primary interest is usually in detecting increases or decreases in the process mean, and yet we consider both standard charts for μ and upper one-sided charts for μ. The former individual charts have acronyms $S-\mu$, $C-\mu$, $CS-\mu$, $E-\mu$ and $ES-\mu$ and the latter are denoted by $S^+-\mu$, $C^+-\mu$, $CS^+-\mu$, $E^+-\mu$ and $ES^+-\mu$; these charts can be found in Table 6.1.

Moreover, we only consider the problem of detecting inflations in the process standard deviation, because an increase in σ corresponds to a reduction in quality and, as put by Reynolds Jr. and Stoumbos (2001), in most processes, an assignable cause that influences the standard deviation is more likely to result in an increase in σ.[1] Thus, only upper

[1] Recall that if reductions in σ occur they are more likely to be the result of an attempt to improve process quality (Reynolds Jr. and Stoumbos (2001)).

one-sided charts for σ are considered in this chapter. Their acronyms are $S^+ - \sigma, C^+ - \sigma, CS^+ - \sigma, E^+ - \sigma$ and $ES^+ - \sigma$ and can be also found in Table 6.1.

Table 6.1: Some individual and joint schemes for μ and σ.

Acronym	Individual scheme for μ	
$S - \mu$	$\bar{X}$	
$C - \mu$	*CUSUM*	
$CS - \mu$	Combined *CUSUM-Shewhart*	
$E - \mu$	*EWMA*	
$ES - \mu$	Combined *EWMA-Shewhart*	
$S^+ - \mu$	Upper one-sided $\bar{X}$	
$C^+ - \mu$	Upper one-sided *CUSUM*	
$CS^+ - \mu$	Combined upper one-sided *CUSUM-Shewhart*	
$E^+ - \mu$	Upper one-sided *EWMA*	
$ES^+ - \mu$	Combined upper one-sided *EWMA-Shewhart*	
	Individual scheme for σ	
$S^+ - \sigma$	Upper one-sided S^2	
$C^+ - \sigma$	Upper one-sided *CUSUM*	
$CS^+ - \sigma$	Combined upper one-sided *CUSUM-Shewhart*	
$E^+ - \sigma$	Upper one-sided *EWMA*	
$ES^+ - \sigma$	Combined upper one-sided *EWMA-Shewhart*	
Joint scheme	Scheme for μ	Scheme for σ
SS	$S - \mu$	$S^+ - \sigma$
CC	$C - \mu$	$C^+ - \sigma$
CCS	$CS - \mu$	$CS^+ - \sigma$
EE	$E - \mu$	$E^+ - \sigma$
CES	$ES - \mu$	$ES^+ - \sigma$
SS^+	$S^+ - \mu$	$S^+ - \sigma$
CC^+	$C^+ - \mu$	$C^+ - \sigma$
CCS^+	$CS^+ - \mu$	$CS^+ - \sigma$
EE^+	$E^+ - \mu$	$E^+ - \sigma$
CES^+	$ES^+ - \mu$	$ES^+ - \sigma$

As a result we shall deal with the joint schemes denoted by SS, CC, CCS, EE, CES, SS^+, CC^+, CCS^+, EE^+ and CES^+ in Table 6.1. These joint schemes involve the individual charts for μ and σ whose summary statistics and control limits can be found in Tables A.2–A.3 and Tables A.5–A.6 (respectively), in Appendix A.

The process is deemed out-of-control whenever a signal is observed on either individual chart of any of these joint schemes for μ and σ. Thus, when dealing with a joint scheme a signal from any of the individual schemes could indicate a possible change in the process mean or in its standard deviation or both. Also, the following events are likely to happen:

- a signal is triggered by the chart for μ although μ is on-target and σ is off-target;

- μ is out-of-control and σ is in-control, however, a signal is given by the chart for σ.

These are some instances of what St. John and Bragg (1991) called "misleading signals" (MS).

Keeping this in mind, this chapter provides striking and instructive examples that alert the user — namely of the ten joint schemes for μ and σ referred above — to the phenomenon of misleading signals.

6.1 Misleading signals

Diagnostic procedures that follow a signal can differ depending on whether the signal is given by the chart for the mean or the chart for the standard deviation. Moreover, they can be influenced by the fact that the signal is given by the positive or negative side of the chart for μ (that is, the observed value of the summary statistic is above the upper control limit or below the lower control limit, respectively). Therefore a misleading signal can possibly send the user of a joint scheme in the wrong direction in the attempt to diagnose and correct a nonexistent assignable cause (St. John and Bragg (1991)). These misleading results suggest inappropriate corrective action, aggravating unnecessarily process variability and increasing production (inspection) costs.

St. John and Bragg (1991) identified the following types of misleading signals arising in joint schemes for μ and σ:

I. the process mean increases but the signal is given by the chart for σ, or the signal is observed on the negative side of the chart for μ;

II. μ shifts down but the signal is observed on the chart for σ, or the chart for μ gives a signal on the positive side;

III. an inflation of the process standard deviation occurs but the signal is given by the chart for μ.

Only type III correspond to what is called a "pure misleading signal" by Morais and Pacheco (2000a) because it is associated to a change in the value of one of the two parameters that is followed by an out-of-control signal by the chart for the other parameter — it corresponds to misinterpreting a standard deviation change as a shift in the mean. However, there is a situation that also leads to a "pure misleading signal" and is related to both misleading signals of Types I and II:

IV. a shift occurs in μ but the out-of-control signal is observed on the chart for σ.

This is called a mislealing signal of Type IV (although it is a sub-type of Types I or II) by Morais and Pacheco (2000a) and it corresponds to misinterpretating a mean change as a shift in the process standard deviation.

In this chapter we shall also consider that the shift in μ is represented in terms of the nominal value of the sample mean standard deviation $\delta = \sqrt{n}(\mu - \mu_0)/\sigma_0$ and the inflation of the process standard deviation will be measured by $\theta = \sigma/\sigma_0$ with: $-\infty < \delta < +\infty$ and $\theta \geq 1$, for the joint schemes SS, CC, EE, CCS and CES; and $\delta \geq 0$ and $\theta \geq 1$, for schemes SS^+, CC^+, EE^+, CCS^+ and CES^+.

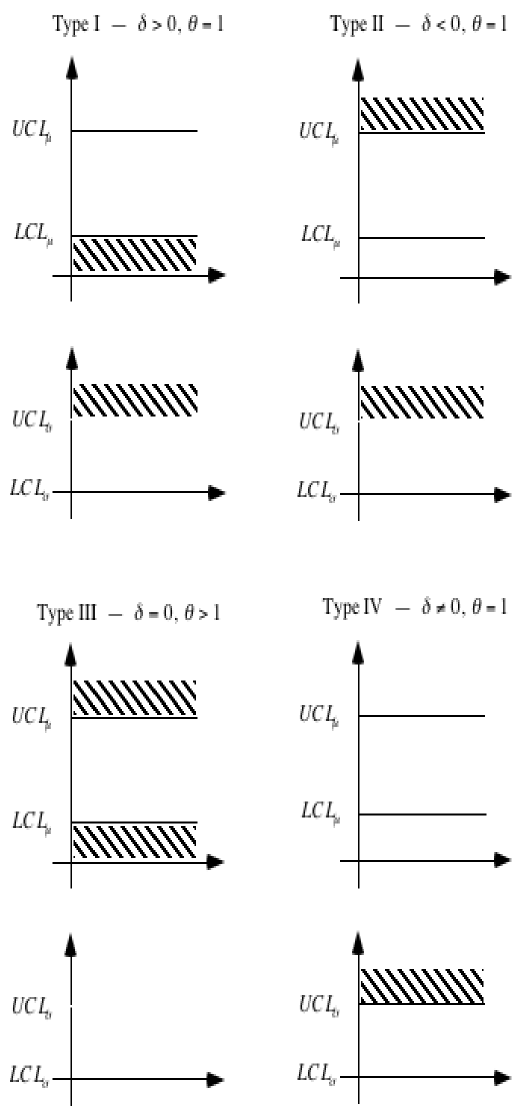

Figure 6.1: Misleading signals of Types I–IV for the schemes SS, CC, EE, CCS and CES.

In this setting misleading signals of Type III occur when $\delta = 0$ and $\theta > 1$. On the other hand, misleading signals of Type IV occur when: $\delta \neq 0$ and $\theta = 1$, for schemes SS, CC, EE, CCS and CES; and $\delta > 0$ and $\theta = 1$, for schemes SS^+, CC^+, EE^+, CCS^+ and CES^+.

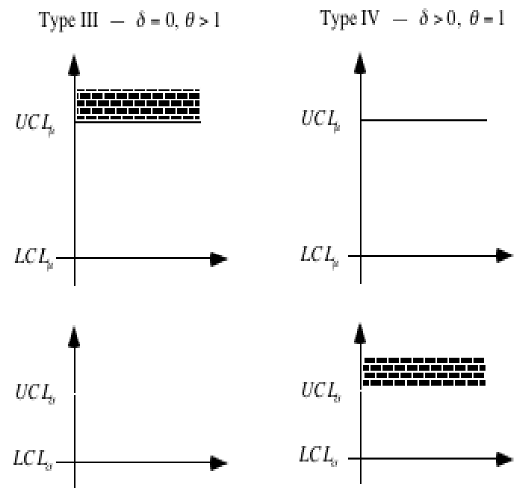

Figure 6.2: Misleading signals of Types III and IV for the schemes SS^+, CC^+, EE^+, CCS^+ and CES^+.

The misleading signals of Types I-IV are graphically described in Figure 6.1 for the joint schemes that make use of a standard chart for μ. Since the joint schemes comprising upper one-sided charts for μ have no negative side only misleading signals of Types III and IV are illustrated in Figure 6.2. A misleading signal corresponds to an observation(s) in the shadowed area of the each graph in Figure 6.1 and Figure 6.2.

Following Morais and Pacheco (2000a) we shall only focus on these two last types of misleading signals because they both correspond to "pure misleading signals". To illustrate the occurrence of misleading signals of Types III and IV, we consider an example with two simulated data sets and a joint SS scheme.

Example 6.1 — The nominal value for the mean and the standard deviation of a chemical reactant temperature are $\mu_0 = 100^oC$ and $\sigma_0 = 1^oC$, respectively. Suppose groups of $n = 9$ temperatures of the reactant are recorded every hour for ten consecutive hours. In the first case the process only has its standard deviation level initially off-target and equal to $\sigma = 1.2^oC$. In the second case only the process mean is initially out-of-control at level $\mu = 100.05^oC$.

The SS scheme was run using a $\bar{X}$ chart for μ ($S-\mu$) and an upper one-sided S^2 chart for σ ($S^+ -\sigma$), whose summary statistics are $\bar{X}$ and $\max\{\sigma_0^2, S^2\}$, respectively. The control limits of these two individual charts are $(LCL_{S-\mu}, UCL_{S-\mu}) = (99.064, 100.936)$ and $(LCL_{S^+-\sigma}, UCL_{S^+-\sigma}) = (1, 2.744)$, and both charts have in-control ARL equal to 200. The observed values of these two summary statistics and the sample standard deviation for two simulated temperature data sets are in Table 6.2.

The joint scheme produced a misleading signal of Type III by the 9^{th} observation of the 1^{st} data set as shown by Table 6.2. Similarly, the 1^{th} observation of the 2^{nd} data set fell outside the control limits of the scheme for σ, indicating that the process standard deviation was seemingly out of control, thus, producing a misleading signal of Type IV.

It is worth mentioning that the 10^{th} sample of the 1^{st} data set was responsible for a correct signal (i.e., a non misleading signal), triggered by the $S^+ - \sigma$ chart. However, the

Table 6.2: Mean ($\bar{x}$), variance (s^2) and $\max\{\sigma_0^2, s^2\}$ of the simulated temperatures.

	$(\mu,\sigma) = (100^oC, 1.2^oC)$			$(\mu,\sigma) = (100.05^oC, 1^oC)$		
N	$\bar{x}$	s^2	$\max\{\sigma_0^2, s^2\}$	$\bar{x}$	s^2	$\max\{\sigma_0^2, s^2\}$
1	99.887	0.437	1.000	99.980	3.295	**3.295***
2	99.429	1.085	1.085	100.478	0.922	1.000
3	100.807	0.610	1.000	99.962	0.963	1.000
4	99.992	1.497	1.497	99.878	0.978	1.000
5	100.025	0.761	1.000	100.130	0.904	1.000
6	100.380	1.113	1.113	99.589	1.402	1.402
7	100.702	1.861	1.861	99.776	0.943	1.000
8	99.897	0.512	1.000	100.093	1.819	1.819
9	**101.015***	1.343	1.343	100.408	1.507	1.507
10	100.139	4.779	**4.779****	100.116	1.281	1.281

* Misleading signal of **Type III**   *** Misleading signal of **Type IV**

** Non misleading signal

$\mu_0 = 100^oC$; $\sigma_0 = 1^oC$; $n = 9$;
$(LCL_{S-\mu}, UCL_{S-\mu}) = (99.064, 100.936)$; $(LCL_{S^+-\sigma}, UCL_{S^+-\sigma}) = (1, 2.744)$.

$S-\mu$ chart did not give any out-of-control signals within the first 10 observations of the 2^{nd} data set. •

We strongly believe that no quality control operator or engineer with proper training would be so naïve to think that a signal from the scheme for the mean only indicates possible shifts in the mean. However, based on the independence between μ and the RL distributions of the schemes for σ, signals given by the scheme for σ are more likely to be associated to an eventual shift in this parameter. Nevertheless, the main question here is not whether there will be misleading signals but rather

- the "probability of a misleading signal" (PMS) and the
- the number of sampling periods before a misleading signal is given by a joint scheme, the "run length to a misleading signal" ($RLMS$).

St. John and Bragg (1991) believed that the phenomenon of misleading signals had not been previously reported. The fact that no such studies had been made is rather curious because misleading signals can arise in any joint scheme for multiple parameters (as, e.g., the multivariate *CUSUM* quality control schemes proposed by Woodall and Ncube (1985)) and also in any two-sided control scheme for a single parameter. But, in fact, as far as we have investigated, there are few references devoting attention to misleading signals or even realizing that they are confronted with such signals.

- *St. John and Bragg (1991).* Figure 3 of this reference illustrates the frequency of misleading signals for various shifts in the process mean when a joint scheme of type $(\bar{X}, R)$ is used; the results were obtained considering 5000 simulated runs of subgroups of five observations from a normal process with mean μ and standard deviation σ.

- *Yashchin (1985).* In Figure 10 of Yashchin (1985) we can find three values of the probability of a signal being given by the upper one-sided chart for μ when there is a decrease in this parameter; but the author does not mention that these values refer in fact to the probability of a misleading signal for a two-sided control scheme for μ which comprises an upper one-sided chart and a lower one-sided chart for μ.

- *Morais and Pacheco (2000a).* These authors provide formulae for the probability of misleading signals of Types III and IV for joint schemes for μ and σ. Based on those expressions these probabilities are evaluated for the joint schemes SS and EE. This paper also accounts for the comparison of these two joint schemes, not only in terms of conventional performance measures such as ARL and RL percentage points, but also with regard to the probabilities of misleading signals.

- *Morais and Pacheco (2001b).* As mentioned in the initial chapter, this paper introduces the notion of run length to a misleading signal and provides monotonicity properties to both PMS and $RLMS$ of the joint $EWMA$ scheme EE^+.

- *Reynolds Jr. and Stoumbos (2001).* This paper refers to the joint monitoring of μ and σ using individuals observations and also discusses the phenomenon of misleading signals (although not referred as such). Table 3 provides simulation-based values not only of the probability of misleading signals of Types III and IV but also of the probability that correct signals (i.e. non misleading signals) occur and of the probability of a simultaneous signal in both individual charts, when μ is on-target and σ is out-of-control, and when σ is in-control and μ is off-target. The authors claim that these probabilities can provide guidelines in the diagnosis of the type of parameter(s) shift(s) that have ocurred.

The monotonicity behaviours of *PMS*s and the stochastic monotonicity properties of *RLMS*s of some of these joint schemes are addressed in this chapter.

Comparisons between the joint schemes are also carried out. They are based on *PMS*s and *RLMS*s. The numerical study that we conduct is designed with careful thought into the appropriate selection of individual chart parameters to ensure common *ARL*s for these charts and hence fair comparisons among the joint schemes.

Based on this extensive study (that closely follows Morais and Pacheco (2000b) though adding other aspects to the investigations done by these authors), we believe that the *PMS*s and the *RLMS*s of Types III and IV should also be taken in consideration as additional performance measures in the design of joint schemes for μ and σ or any joint scheme for more than one parameter. Recall that Morais and Pacheco (2000a) proposed a design strategy for the joint schemes SS and EE, whose Step 4 involves the investigation of the overall behaviour of the designed combined scheme, in terms of not only the in-control and out-of-control *ARL*s and RL percentage points but also in terms of the *PMS*s of Types III and IV.

This completes the introduction on misleading signals. Some considerations on the evaluation of conventional performance measures of the joint schemes such as the RL survival function, the RL percentage points and the ARL is what lies ahead.

6.2 Approximations to the RL distribution and ARL

For convenience, let $RL_\mu(\delta,\theta)$, $RL_\sigma(\theta)$ and $RL_{\mu,\sigma}(\delta,\theta)$ denote, in general and respectively, the RLs of an individual chart for μ, an individual chart for σ and a joint scheme for μ and σ. When we want to specifically address the Markov approximations to the RLs of those individual charts and schemes we write $RL^{\alpha}_{\mu}(\delta,\theta;x_\mu \text{ or } x^+_\mu)$, $RL^{\beta}_{\sigma}(\theta;x^+_\sigma)$ and $RL^{\alpha,\beta}_{\mu,\sigma}(\delta,\theta;x_\mu \text{ or } x^+_\mu, x^+_\sigma)$. Similarly, the exact RLs are denoted by $RL^{\alpha}_{\mu}(\delta,\theta)$, $RL^{\beta}_{\sigma}(\theta)$ and $RL^{\alpha,\beta}_{\mu,\sigma}(\delta,\theta)$.

Since we are dealing with a normally distributed quality characteristic, the summary statistics of the two individual charts for μ and σ are independent, given (δ,θ). Thus, we can provide simple expressions for the performance measures of these joint schemes, such as the RL survival function, RL percentage points and ARL, which depend exclusively on the performance of the two individual charts.

A change in the process production must be detected quickly so that a corrective action can be taken, and, as mentioned before, a signal is given by the joint scheme at the sampling period N if such a signal is observed on either individual scheme at time N. Thus, following Gan (1995) (who proposed the use of a joint scheme which comprises individual charts of type $EWMA$ for μ and σ without head starts) and Woodall and Ncube (1985) (who considered the simultaneous use of m univariate charts for m expected values), the RL of the joint schemes studied here is given by

$$RL_{\mu,\sigma}(\delta,\theta) = \min\{RL_\mu(\delta,\theta), RL_\sigma(\theta)\}. \tag{6.1}$$

And since $RL_\mu(\delta,\theta)$ and $RL_\sigma(\theta)$ are independent, conditioned on (δ,θ), the survival function of $RL_{\mu,\sigma}(\delta,\theta)$ equals the product of the survival functions of the RL of the two individual schemes:

$$\overline{F}_{RL_{\mu,\sigma}(\delta,\theta)}(m) = \overline{F}_{RL_\mu(\delta,\theta)}(m) \times \overline{F}_{RL_\sigma(\theta)}(m). \tag{6.2}$$

As a consequence, the RLs of the joint schemes SS and SS^+ are also geometric random variables with survival functions with parameters

$$1 - \{\Phi[(\xi_\mu-\delta)/\theta] - \Phi[(-\xi_\mu-\delta)/\theta]\} \times F_{\chi^2_{n-1}}(\xi^+_\sigma/\theta^2) \tag{6.3}$$

$$1 - \Phi[(\xi^+_\mu-\delta)/\theta] \times F_{\chi^2_{n-1}}(\xi^+_\sigma/\theta^2), \tag{6.4}$$

respectively.

When we are dealing with Markov-type summary statistics we must use the Markov approximations to the survival functions of the individual charts for μ and σ in Equation

(6.2), given in Tables A.8 and A.9 in Appendix A.[2]

The $p \times 100\%$ $(0 < p < 1)$ percentage point of $RL_{\mu,\sigma}(\delta,\theta)$, $F^{-1}_{RL_{\mu,\sigma}(\delta,\theta)}(p)$, is defined by the smallest integer m satisfying

$$\overline{F}_{RL_{\mu,\sigma}(\delta,\theta)}(m) \leq 1 - p \tag{6.5}$$

Exact or approximate values to this percentage point can be easily obtained plugging the exact or approximate RL survival function.

The ARLs of the joint schemes SS and SS^+ follow quite trivially because the RLs have geometric distribution. The ARL of the remaining schemes can be aproximated using any one of the three procedures described below.

In the first approximation procedure we take into account that the ARL is found by simply summing the survival function of $RL_{\mu,\sigma}(\delta,\theta)$:

$$ARL^{\alpha,\beta}_{\mu,\sigma}(\delta,\theta) = \sum_{m=0}^{+\infty} \overline{F}_{RL^{\alpha}_{\mu}(\delta,\theta)}(m) \times \overline{F}_{RL^{\beta}_{\sigma}(\theta)}(m). \tag{6.6}$$

Hence, $ARL^{\alpha,\beta}_{\mu,\sigma}(\delta,\theta)$ can be computed approximately obtained by truncating this series after having replaced $\overline{F}_{RL^{\alpha}_{\mu}(\delta,\theta)}(m)$ and $\overline{F}_{RL^{\beta}_{\sigma}(\theta)}(m)$ by their Markov approximation in Tables A.8 and A.9. Moreover, the approximation procedure used here (and by Morais and Pacheco (2000b)) to compute ARL values considers 41 transient states and assumes that the convergence of the series is attained as soon as the relative error is less than 10^{-6}.

Denote the matrix of all ARLs of the joint scheme CC^+ by

$$\mathbf{U}(\delta,\theta;x^+_\mu,x^+_\sigma) = [ARL^{u,v}(\delta,\theta;x^+_\mu,x^+_\sigma)]^{x^+_\mu+1;\,x^+_\sigma+1}_{u=0;\,v=0} \tag{6.7}$$

where $ARL^{u,v}(\delta,\theta;x^+_\mu,x^+_\sigma)$ represents the ARL of this scheme assuming that the initial states of the summary statistics of the individual schemes for μ and σ are associated with the in-control or out-of-control states u and v in the Markov approximation.

In the second approximation procedure, the matrix $\mathbf{U}(\delta,\theta;x^+_\mu,x^+_\sigma)$ is computed by adapting the iterative procedure proposed by Prabhu and Runger (1996) to this particular control scheme[3]

$$\mathbf{U}^{(k+1)}(\delta,\theta;x^+_\mu,x^+_\sigma) = \mathbf{T} \bullet [\mathbf{T} + \mathbf{P}_\mu(\delta,\theta;x^+_\mu)\,\mathbf{U}^{(k)}(\delta,\theta;x^+_\mu,x^+_\sigma)\,\mathbf{P}_\sigma(\theta;x^+_\sigma)], \tag{6.8}$$

for $k = 0, 1, \ldots$, where:

- $\mathbf{U}^{(k)}(\delta,\theta;x^+_\mu,x^+_\sigma)$ represents the approximation for $\mathbf{U}(\delta,\theta;x^+_\mu,x^+_\sigma)$ in the k^{th} iteration;

[2]It is worth noticing that an approximation to the distribution function of $RL_{\mu,\sigma}(\delta,\theta)$ could have been obtained considering a two-dimensional Markov chain. However, the independence between the horizontal and vertical transitions of this approximating two-dimensional Markov chain makes it possible to avoid the computation of a transition matrix with unusual dimensions: $(x_\mu+2)^2 \times (x_\sigma+2)^2$.

[3]The original procedure was used by Runger and Prabhu (1996) to obtain approximate values to the ARL of a multivariate $EWMA$ scheme for the control of a multivariate normal mean vector.

- $\mathbf{U}^{(0)}(\delta,\theta;x_\mu^+,x_\sigma^+) = \underline{\mathbf{0}}_\mu^\top\,\underline{\mathbf{0}}_\sigma$ where $\underline{\mathbf{0}}_\mu$ and $\underline{\mathbf{0}}_\sigma$ denote a vector of $(x_\mu^+ + 2)$ and $(x_\sigma^+ + 2)$ zeroes, respectively;

- $\mathbf{P}_\mu(\delta,\theta;x_\mu^+)$ and $\mathbf{P}_\sigma(\theta;x_\sigma^+)$ are probability transition matrices defined in terms of the sub-stochastic matrices $\mathbf{Q}_\mu(\delta,\theta;x_\mu^+)$ and $\mathbf{Q}_\sigma(\theta;x_\sigma^+)$;

- $\mathbf{T}$ is a $(x_\mu^+ + 2)\times(x_\sigma^+ + 2)$ matrix which indicates the in-control or transient states of the approximating two-dimensional Markov chain, defined as

$$\mathbf{T} = \begin{bmatrix} \underline{\mathbf{1}}_\mu\,\underline{\mathbf{1}}_\sigma^\top & \underline{\mathbf{0}}_\mu \\ \underline{\mathbf{0}}_\sigma^\top & 0 \end{bmatrix}, \tag{6.9}$$

  where $\underline{\mathbf{1}}_\mu$ and $\underline{\mathbf{1}}_\sigma$ (and now $\underline{\mathbf{0}}_\mu$ and $\underline{\mathbf{0}}_\sigma$) represent a vector of $(x_\mu^+ + 1)$ and $(x_\sigma^+ + 1)$ ones (zeroes);

- and the symbol "$\bullet$" indicates elementwise multiplication of the matrices.

This iterative procedure, which was used by Morais (1998) to approximate the ARL of a joint scheme EE without head starts, requires a large number of iterations until convergence is attained for the in-control ARL of this joint scheme, and a steadily decreasing number of iterations when $|\delta|$ or θ increases. The results obtained here and by Morais and Pacheco (2000b) using the first approximation procedure lead to the common conclusion that the number of iterations is virtually independent of the head starts given to the individual control charts.

We could have also used a third approximation procedure proposed by Knoth and Schmid (1999) for the determination of the ARL of joint residual $EWMA$ schemes for the process mean and standard deviation of autocorrelated data.[4]

Following these two authors, the Markov approximation to $ARL_{\mu,\sigma}^{\alpha,\beta}(\delta,\theta)$, with $2x_\mu+1 = x_\mu^+ + 1 = x_\sigma^+ + 1 = x^+ + 1$ transient states, equals

$$\begin{aligned} ARL_{\mu,\sigma}^{\alpha,\beta}(\delta,\theta;x^+,x^+) &= \sum_{m=0}^{+\infty} \overline{F}_{RL_\mu^\alpha(\delta,\theta;x^+)}(m) \times \overline{F}_{RL_\sigma^\beta(\theta;x^+)}(m) \\ &= \sum_{m=0}^{+\infty} \Big\{ \underline{\mathbf{e}}_{\lfloor\alpha(x^++1)\rfloor}^\top \left[\mathbf{Q}_\mu(\delta,\theta;x^+)\right]^m \underline{\mathbf{1}} \\ &\qquad\qquad \times \underline{\mathbf{e}}_{\lfloor\beta(x^++1)\rfloor}^\top \left[\mathbf{Q}_\sigma(\theta;x^+)\right]^m \underline{\mathbf{1}} \Big\} \\ &= \underline{\mathbf{e}}_{\lfloor\alpha(x^++1)\rfloor}^\top\,\mathbf{Z}(\delta,\theta;x^+)\,\underline{\mathbf{e}}_{\lfloor\beta(x^++1)\rfloor}^\top \end{aligned} \tag{6.10}$$

where

$$\mathbf{Z}(\delta,\theta;x^+) = \sum_{m=0}^{+\infty} \left(\left[\mathbf{Q}_\mu(\delta,\theta;x^+)\right]^m \underline{\mathbf{1}} \times \underline{\mathbf{1}}^\top \left\{ \left[\mathbf{Q}_\sigma(\theta;x^+)\right]^\top \right\}^m \right) \tag{6.11}$$

is the solution of the matrix equation

$$\mathbf{Z}(\delta,\theta;x^+) = \underline{\mathbf{1}}\,\underline{\mathbf{1}}^\top + \mathbf{Q}_\mu(\delta,\theta;x^+)\,\mathbf{Z}(\delta,\theta;x^+)\,\left[\mathbf{Q}_\sigma(\theta;x^+)\right]^\top \tag{6.12}$$

[4]Residual schemes are addressed in Chapter 7.

or equivalently

$$\begin{aligned}
&\left[\mathbf{I} - \mathbf{Q}_\mu(\delta,\theta;x^+)\right]^{-1} \mathbf{Z}(\delta,\theta;x^+) \\
&\qquad + \mathbf{Z}(\delta,\theta;x^+)\ \left[\mathbf{Q}_\sigma(\theta;x^+)\right]^\top \left\{\mathbf{I} - \left[\mathbf{Q}_\sigma(\theta;x^+)\right]^\top\right\}^{-1} \\
&\qquad = \left[\mathbf{I} - \mathbf{Q}_\mu(\delta,\theta;x^+)\right]^{-1} \underline{\mathbf{1}} \times \underline{\mathbf{1}}^\top \left\{\mathbf{I} - \left[\mathbf{Q}_\sigma(\theta;x^+)\right]^\top\right\}^{-1}, \qquad (6.13)
\end{aligned}$$

a Sylvester equation that can be solved with the NAG function *sylv()*.

6.3 Probability of a misleading signal (PMS)

Once again the independence between the summary statistics of the individual charts for μ and σ plays a major role in providing plain expressions for the probabilities of misleading signals of Types III and IV, denoted in general by $PMS_{III}(\theta)$ and $PMS_{IV}(\delta)$.

Lemma 6.2 — *The expressions of the* PMS*s of Types III and IV for joint schemes involving individual schemes with independent summary statistics (such as the ten joint schemes in Table 6.1) are*

$$\begin{aligned}
PMS_{III}(\theta) &= P[RL_\sigma(\theta) > RL_\mu(0,\theta)] \\
&= \sum_{i=2}^{+\infty} F_{RL_\mu(0,\theta)}(i-1) \times P_{RL_\sigma(\theta)}(i) \qquad (6.14) \\
&= \sum_{i=1}^{+\infty} P_{RL_\mu(0,\theta)}(i) \times \overline{F}_{RL_\sigma(\theta)}(i),\ \theta > 1 \qquad (6.15)
\end{aligned}$$

$$\begin{aligned}
PMS_{IV}(\delta) &= P[RL_\mu(\delta,1) > RL_\sigma(1)] \\
&= \sum_{i=1}^{+\infty} \overline{F}_{RL_\mu(\delta,1)}(i) \times P_{RL_\sigma(1)}(i) \qquad (6.16) \\
&= \sum_{i=2}^{+\infty} P_{RL_\mu(\delta,1)}(i) \times F_{RL_\sigma(1)}(i-1),\ \delta \neq 0 \qquad (6.17)
\end{aligned}$$

(or $\delta > 0$ when we are using upper one-sided schemes for μ), where $RL_\mu(\delta,\theta)$ and $RL_\sigma(\theta)$ represent the run lengths of the individual schemes for μ and σ, as previously mentioned.

The exact expressions of the *PMS*s of the joint schemes SS and SS^+ can be found in Table 6.3 and follow immediately by pluging in the survival functions of the run lengths RL_μ and RL_σ (see Tables A.8 and A.9, respectively) into equations (6.15) and (6.16).

The approximations to *PMS*s of the remaining joint schemes are found by using the Markov approximations to the survival function of RL_μ and RL_σ (also in Tables A.8 and A.9) and truncating the series (6.15) and (6.16). The approximate values of the *PMS*s converge to the true values due to the convergence in law of the approximate *RL*s involved in the definition of the *PMS*s.

Table 6.3: Exact *PMS* of Types III and IV for the joint schemes SS and SS^+.

Joint scheme	$PMS_{III}(\theta), \theta > 1$	$PMS_{IV}(\delta), \delta \neq 0\ (\delta > 0)$
SS	$\frac{1-[\Phi(\xi_\mu/\theta)-\Phi(-\xi_\mu/\theta)]}{[F_{\chi^2_{n-1}}(\xi_\sigma^+/\theta^2)]^{-1}-[\Phi(\xi_\mu/\theta)-\Phi(-\xi_\mu/\theta)]}$	$\frac{1-F_{\chi^2_{n-1}}(\xi_\sigma^+)}{[\Phi(\xi_\mu-\delta)-\Phi(-\xi_\mu-\delta)]^{-1}-F_{\chi^2_{n-1}}(\xi_\sigma^+)}$
SS^+	$\frac{1-\Phi(\xi_\mu^+/\theta)}{[F_{\chi^2_{n-1}}(\xi_\sigma^+/\theta^2)]^{-1}-\Phi(\xi_\mu^+/\theta)}$	$\frac{1-F_{\chi^2_{n-1}}(\xi_\sigma^+)}{[\Phi(\xi_\mu^+-\delta)]^{-1}-F_{\chi^2_{n-1}}(\xi_\sigma^+)}$

Theorem 6.3 — *The monotonicity properties (1)–(16) in the table below are valid for the (exact)* PMSs *of Types III and IV of the joint schemes in Table 6.1 based exclusively on upper one-sided individual charts: SS^+, CC^+, CCS^+, EE^+ and CES^+.*

Joint scheme	Type III ($\delta = 0, \theta > 1$)	Type IV ($\delta > 0, \theta = 1$)
SS^+	(1) $PMS_{III}(\theta) \downarrow$ with θ	(8) $PMS_{IV}(\delta) \downarrow$ with δ
CC^+, CCS^+	(2) $PMS_{III}(\theta) \uparrow$ with α	(9) $PMS_{IV}(\delta) \downarrow$ with α
	(3) $PMS_{III}(\theta) \downarrow$ with β	(10) $PMS_{IV}(\delta) \uparrow$ with β
	(C1) $PMS_{III}(\theta)$*	(11) $PMS_{IV}(\delta) \downarrow$ with δ
	(4) $PMS_{III}(\theta) \downarrow$ with k_μ^+	(12) $PMS_{IV}(\theta) \uparrow$ with k_μ^+
	(5) $PMS_{III}(\theta) \uparrow$ with k_σ^+	(13) $PMS_{IV}(\theta) \downarrow$ with k_σ^+
EE^+, CES^+	(6) $PMS_{III}(\theta) \uparrow$ with α	(14) $PMS_{IV}(\delta) \downarrow$ with α
	(7) $PMS_{III}(\theta) \downarrow$ with β	(15) $PMS_{IV}(\delta) \uparrow$ with β
	(C2) $PMS_{III}(\theta)$*	(16) $PMS_{IV}(\delta) \downarrow$ with δ

*(C1, C2) Conjecture of no monotone behaviour in terms of θ

Proof — The monotonicity properties of *PMS*s of Types III and IV given in Theorem 6.3 are intuitive and have an analytical justification — most of them follow directly from expressions (6.14)–(6.17) from Lemma 6.2, and from the stochastic monotone properties of the *RL*s of the individual schemes for μ (see Table 5.1) and of those for σ (refer to Table A.10).

Take for instance the *PMS* of Type IV of the joint scheme SS^+,

$$PMS_{IV,SS^+}(\delta) = P[RL_{S^+-\mu}(\delta,1) > RL_{S^+-\sigma}(1)]. \tag{6.18}$$

It is a decreasing function of δ because: the distribution of $RL_{S^+-\sigma}(1)$ does not depend on δ; $RL_\mu(\delta,1)$ stochastically decreases with δ and so does its survival function for any i and $PMS_{IV,SS^+}(\delta)$ defined by Equation (6.16). As a consequence the joint scheme SS^+ tends to trigger less misleading signals of Type IV as δ increases.

As for the monotone behaviour of the *PMS* of Type III of the joint scheme SS^+, $PMS_{III,SS^+}(\theta)$, it follows immediately that it is a decreasing function of θ if we note that, for $\theta > 1$,

$$\begin{aligned} PMS_{III,SS^+}(\theta) &= \frac{1-\Phi(\xi_\mu^+/\theta)}{[F_{\chi^2_{n-1}}(\xi_\sigma^+/\theta^2)]^{-1}-\Phi(\xi_\mu^+/\theta)} \\ &= \left[1-\frac{1-[F_{\chi^2_{n-1}}(\xi_\sigma^+/\theta^2)]^{-1}}{1-\Phi(\xi_\mu^+/\theta)}\right]^{-1}. \end{aligned} \tag{6.19}$$

The results concerning the remaining schemes follow quite similarly. For example, the approximation to the PMS of Type III of the joint scheme CC^+,

$$\begin{aligned} PMS^{\alpha,\beta}_{III,CC^+}(\theta;x_\mu^+,x_\sigma^+) &= P[RL^{\beta}_{C^+-\sigma}(\theta;x_\sigma^+) > RL^{\alpha}_{C^+-\mu}(0,\theta;x_\mu^+)] \\ &= \sum_{i=1}^{+\infty} P_{RL^{\alpha}_{C^+-\mu}(0,\theta;x_\mu^+)}(i)\times \overline{F}_{RL^{\beta}_{C^+-\sigma}(\theta;x_\sigma^+)}(i), \end{aligned} \tag{6.20}$$

decreases with β because a change in the head-start of the individual scheme for σ does not influence the probability function of $RL^{\alpha}_{C^+-\mu}(0,\theta;x_\mu^+)$ but is responsible for a decrease in the survival function of $RL^{\beta}_{C^+-\sigma}(\theta;x_\sigma^+)$, as a consequence of property (2) in Table A.10. Hence, it leads to a decrease in $PMS^{\alpha,\beta}_{CC^+}(\theta;x_\mu^+,x_\sigma^+)$, given by Equation (6.20), and also in

$$\begin{aligned} PMS^{\alpha,\beta}_{III,CC^+}(\theta) &= P[RL^{\beta}_{C^+-\sigma}(\theta) > RL^{\alpha}_{C^+-\mu}(0,\theta)] \\ &= \lim_{x_\mu^+,x_\sigma^+\to+\infty} PMS^{\alpha,\beta}_{III,CC^+}(\theta;x_\mu^+,x_\sigma^+). \end{aligned} \tag{6.21}$$

•

Remark 6.4 — The conjectures (C1) and (C2) included in the table of Theorem 6.3 concern the *PMS*s of Type III of the joint schemes CC^+ (CCS^+) and EE^+ (CES^+), respectively, and surely deserve a comment.

$PMS_{III}(\theta)$ involves in its definition $RL_\mu(0,\theta)$ and $RL_\sigma(\theta)$. If, in one hand, this latter random variable stochastically decreases with θ (see Table A.10), on the other hand, we cannot tell what is the stochastic monotone behaviour of $RL_\mu(0,\theta)$ in terms of θ,[5] since $\delta = 0$ belongs to the intervals

$$(\underline{\delta}_{C^+-\mu},\overline{\delta}_{C^+-\mu}) = (k_\mu^+ - h_\mu^+, k_\mu^+ + h_\mu^+) \tag{6.22}$$

$$(\underline{\delta}_{E^+-\mu},\overline{\delta}_{E^+-\mu}) = \left(-\frac{\gamma_\mu^+(1-\lambda_\mu^+)}{\sqrt{\lambda_\mu^+(2-\lambda_\mu^+)}}, \frac{\gamma_\mu^+}{\sqrt{\lambda_\mu^+(2-\lambda_\mu^+)}}\right). \tag{6.23}$$

Thus, establishing a monotonic behaviour of $PMS_{III}(\theta)$ in terms of θ seems to be non trivial.[6]

[5]However, note that the percentage points in Table 5.4 and those reported by Morais and Pacheco (2000c) suggest that $RL_\mu(0,\theta)$ decreases with θ.

[6]In view of the numerical results from Section 5.4 and additional investigations we dare to hint that $RL_\mu(0,\theta)$ stochastically decreases with θ. This fact is a possible explanation for the nonmonotonous behaviour of $PMS_{III}(\theta)$.

Note that this fact also holds for the joint schemes based on upper one-sided combined schemes, CCS^+ and CES^+, according to the results (12)–(13) and (21)–(22) in Table 5.1.

As we shall see, the numerical results in the next section support conjectures (C1) and (C2): values of $PMS_{III}(\theta)$ seem to decrease and then increase with θ for the joint schemes CC^+, CCS^+, EE^+ and CES^+. The practical significance of this nonmonotonous behaviour is as follows: the joint scheme ability to misidentify a shift in σ can increase as the displacement in σ becomes more severe. •

6.4 PMS: numerical illustrations

As an illustration, the *PMS*s of Types III and IV were obtained for the ten joint schemes considered in Table 6.1, whose individual schemes for μ and σ have the parameters given in Table 6.4. We provide values for those probabilities considering:

- sample size equal to $n = 5$;
- nominal values $\mu_0 = 0$ and $\sigma_0 = 1$; and
- $\delta = 0.05, 0.10, 0.20, 0.30, 0.40, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0, 1.5, 2.0, 3.0$ and $\theta = 1.01, 1.03, 1.05, 1.10, 1.20, 1.30, 1.40, 1.50, 1.60, 1.70, 1.80, 1.90, 2.00, 3.00$,

which practically cover the same range that was found useful previously in Gan (1989).

With the exception of the joint schemes SS and SS^+, the values of these probabilities are approximate and based on the Markov approach using 41 transient states and considering a relative error of 10^{-6} in the truncation of the series in (6.15) and (6.16) as mentioned earlier.

The range of the decision intervals $[LCL, UCL)$ of all the individual schemes for μ and for σ has been chosen in such way that, when no head start has been adopted, these schemes are approximately matched in-control and all the corresponding in-control ARLs are close to 500 samples (see Table 6.4). Take for instance the *Shewhart*–type schemes:

- $S - \mu$: $\xi_\mu = \Phi^{-1}[1 - 1/(2 \times 500)] = 3.09023$;
- $S^+ - \mu$: $\xi_\mu^+ = \Phi^{-1}(1 - 1/500) = 2.87816$;
- $S^+ - \sigma$: $\xi_\sigma^+ = F^{-1}_{\chi^2_4}(1 - 1/500) = 16.9238$.

It should be also added that, in the case of the individual combined *CUSUM–Shewhart* and *EWMA–Shewhart* schemes for μ and for σ, all the *Shewhart*–type constituent charts have in-control ARL equal to 1000. For example:

- $CS - \mu$, $ES - \mu$: $\xi_\mu = \Phi^{-1}[1 - 1/(2 \times 1000)] = 3.29053$;
- $CS^+ - \mu$, $ES^+ - \mu$: $\xi_\mu^+ = \Phi^{-1}(1 - 1/1000) = 3.09023$;
- $CS^+ - \sigma$, $ES^+ - \sigma$: $\xi_\sigma^+ = F^{-1}_{\chi^2_4}(1 - 1/1000) = 18.4668$.

Table 6.4: Parameters and in-control ARLs of individual and joint schemes for μ and σ.

Scheme for μ	Parameters	$ARL_\mu(0,1)$
$S-\mu$	$\xi_\mu = 3.09023$	500.000
$C-\mu$	$h_\mu = 22.7610$	500.001
$CS-\mu$	$\xi_\mu = 3.29053,\ h_\mu = 31.1810$	500.001
$E-\mu$	$\gamma_\mu = 2.8891,\ \lambda_\mu = 0.134$	499.988
$ES-\mu$	$\xi_\mu = 3.29053,\ \gamma_\mu = 3.0934,\ \lambda_\mu = 0.134$	499.999
$S^+-\mu$	$\xi_\mu^+ = 2.87816$	500.000
$C^+-\mu$	$h_\mu^+ = 4.4456,\ k_\mu^+ = 0.5$	500.021
$CS^+-\mu$	$\xi_\mu^+ = 3.09023,\ h_\mu^+ = 4.9854,\ k_\mu^+ = 0.5$	500.020
$E^+-\mu$	$\gamma_\mu^+ = 2.8116,\ \lambda_\mu^+ = 0.134$	500.047
$ES^+-\mu$	$\xi_\mu^+ = 3.09023,\ \gamma_\mu^+ = 3.0016,\ \lambda_\mu^+ = 0.134$	500.044
Scheme for σ	Parameters	$ARL_\sigma(1)$
$S^+-\sigma$	$\xi_\sigma^+ = 16.9238$	500.000
$C^+-\sigma$	$h_\sigma^+ = 3.5069,\ k_\sigma^+ = 0.055$	499.993
$CS^+-\sigma$	$\xi_\sigma^+ = 18.4668,\ h_\sigma^+ = 3.9897,\ k_\sigma^+ = 0.055$	500.002
$E^+-\sigma$	$\gamma_\sigma^+ = 1.2198,\ \lambda_\sigma^+ = 0.043$	500.027
$ES^+-\sigma$	$\xi_\sigma^+ = 18.4668,\ \gamma_\sigma^+ = 1.3510,\ \lambda_\sigma^+ = 0.043$	500.033
Joint scheme	No. of Iterations for $ARL_{\mu,\sigma}(0,1)$	$ARL_{\mu,\sigma}(0,1)$
$SS-\sigma$	——	250.250
$CC-\sigma$	1882	272.653
$CCS-\sigma$	1909	267.309
$EE-\sigma$	2050	253.318
$CES-\sigma$	2060	252.003
$SS^+-\sigma$	——	250.250
$CC^+-\sigma$	2051	253.230
$CCS^+-\sigma$	2059	252.071
$EE^+-\sigma$	2049	253.475
$CES^+-\sigma$	2059	252.093

A few more remarks on the choice of the individual charts parameters ought to be made, namely that most of them were taken from the literature in order to optimize the detection of a shift in μ from the nominal value μ_0 to $\mu_0 + 1.0 \times \sigma_0/\sqrt{n}$ and a shift in σ from σ_0 to $1.25 \times \sigma_0$. "Optimality" here means that extensive numerical results suggest that with such reference values and smoothing constants will produce an individual chart for μ (σ) with the smallest possible out-of-control ARL, $ARL_\mu(1.0, 1.0)$ ($ARL_\sigma(1.25)$), for a fixed in-control ARL of 500 samples. The upper control limits of the individual charts are always searched in such way that, when no head start has been adopted, they have in-control ARLs close to 500.

- $C - \mu$: we adopted a null reference value because we are dealing with a standard *CUSUM*;[7]

- $E - \mu$: the smoothing constant $\lambda_\mu = 0.134$ was taken from Gan (1995) and aggrees with Figure 4 from Crowder (1989);

- $C^+ - \mu$: the reference value $k^+_\mu = 0.5$ is suggested by Gan (1991) for a two-sided chart for μ;

- $E^+ - \mu$: we adopt the same smoothing constant $\lambda^+_\mu = 0.134$ as $E - \mu$;

- $C^+ - \sigma$, $E^+ - \sigma$: Gan (1995) suggests a reference value $k^+_\sigma = 0.055$ and a smoothing constant $\lambda^+_\sigma = 0.043$, respectively.

Finally, note that the individual charts $CS - \mu$, $ES - \mu$, $CS^+ - \mu$, $ES^+ - \mu$, $CS^+ - \sigma$, $ES^+ - \sigma$ have the same reference values and smoothing constants as $C - \mu$, $E - \mu$, $C^+ - \mu$, $E^+ - \mu$, $C^+ - \sigma$, $E^+ - \sigma$. The control limits are larger than the corresponding "non-combined" individual schemes and are obtained taking into account the supplementary *Shewhart* control limits and the in-control matching of the ARLs.

We proceed to illustrate the monotonicity properties stated in the previous section with the joint scheme EE^+. For that purpose some head starts have been given to this joint scheme: $HS_\mu = 0\%, 50\%$ and $HS_\sigma = 0\%, 50\%$.[8]

The results in Table 6.5 not only show that PMSs of Types III and IV can be as high as 0.47 but also remind us of some of the monotonicity properties in Theorem 6.3, namely that:

- giving head starts to the individual chart for μ, $E^+ - \mu$, leads to an increase of PMSs of Type III and a decrease of the PMSs of Type IV;

- adopting a head start to the individual chart for σ, $E^+ - \sigma$, yields a decrease of the values of PMS of Type III and an increase of the ones of Type IV; and

- underestimating the magnitude of the changes in μ results in an overestimation of the PMS values of Type IV.

The numerical results also suggest that underestimating the magnitude of the changes in σ also leads to an overestimation of the values of the PMSs of Type III (recall Conjecture (C1)), for most of the values we considered for θ. However, note that $PMS_{III}(\theta)$ changes

[7] A positive (negative) reference value would suggest that positive (negative) shifts are more likely to occur. Also note that this standard individual chart is different from the two-sided scheme which makes use of an upper and a lower one-sided chart for μ.

[8] It is worth recalling that we did not consider $HS_\mu, HS_\sigma = -50\%$ because both individual schemes are upper one-sided.

Table 6.5: PMSs of Types III and IV for the joint scheme EE^+ and number of iterations until convergence.

	$PMS_{III}(\theta)$					$PMS_{IV}(\delta)$			
	HS_σ					HS_σ			
	0%		50%			0%		50%	
	HS_μ		HS_μ			HS_μ		HS_μ	
θ	0%	50%	0%	50%	δ	0%	50%	0%	50%
1.01	.455274	.471146	.432002	.447743	0.05	.403991	.389425	.429989	.415191
1.03	.377092	.397075	.351439	.371084	0.10	.319232	.304463	.348887	.333675
1.05	.313194	.337278	.285745	.309181	0.20	.191651	.177663	.226708	.211665
1.10	.206131	.239939	.176070	.207955	0.30	.114210	.101826	.152328	.138045
1.20	.114615	.164954	.083117	.126973	0.40	.069767	.059179	.109302	.095759
1.30	.081130	.144553	.049742	.100509	0.50	.044152	.035268	.084026	.070964
1.40	.065605	.139364	.034609	.089256	0.60	.028898	.021542	.068364	.055500
1.50	.057295	**.139322**	.026587	.083500	0.70	.019432	.013413	.057926	.045048
1.60	.052531	**.141293**	.021890	.080240	0.80	.013327	.008462	.050401	.037394
1.70	.049768	**.144098**	.018961	.078290	0.90	.009262	.005378	.044563	.031405
1.80	.048249	**.147214**	.017070	.077056	1.00	.006491	.003430	.039755	.026497
1.90	.047556	**.150386**	.015842	.076181	1.50	.001126	.000351	.022935	.010964
2.00	**.047439**	**.153486**	.015069	.075427	2.00	.000185	.000032	.012581	.004034
3.00	**.059958**	**.181832**	.015744	.058932	3.00	.000004	.000000	.003124	.000368

	No. of iterations for $PMS_{III}(\theta)$					No. of iterations for $PMS_{IV}(\delta)$			
	HS_σ					HS_σ			
	0%		50%			0%		50%	
	HS_μ		HS_μ			HS_μ		HS_μ	
θ	0%	50%	0%	50%	δ	0%	50%	0%	50%
1.01	1772	1759	1772	1758	0.05	1707	1707	1686	1685
1.03	1320	1307	1320	1307	0.10	1388	1388	1367	1366
1.05	991	979	991	979	0.20	885	885	865	864
1.10	517	507	517	506	0.30	561	561	542	541
1.20	194	187	194	185	0.40	364	364	347	346
1.30	101	95	100	94	0.50	246	246	231	230
1.40	64	60	63	58	0.60	174	174	160	159
1.50	45	42	45	41	0.70	128	127	116	114
1.60	35	32	34	31	0.80	97	97	87	85
1.70	28	26	28	25	0.90	76	76	67	66
1.80	24	22	23	21	1.00	62	61	54	52
1.90	20	19	20	18	1.50	28	28	23	22
2.00	18	17	18	16	2.00	17	17	14	13
3.00	9	8	8	8	3.00	9	9	7	7

its monotonous behaviour, for large values of θ, according to the values in **bold** in Table 6.5.

As mentioned earlier, when we addressed the approximation procedures to the ARL of a joint scheme, the evaluation of approximate values of *PMS*s of Types III and IV also requires a large number of iterations until convergence is attained for small values of θ and δ, as suggested by Table 6.5. However, this number steadily decreases with both θ and δ. Moreover, according to Table 6.5, the number of iterations needed for the convergence of $PMS_{III}(\theta)$ ($PMS_{IV}(\delta)$) is virtually independent of HS_σ (HS_μ) and varies slightly with HS_μ (HS_σ).

The second illustration accounts for the comparative assessment of the schemes SS, CC, CCS, EE and CES, and of the schemes SS^+, CC^+, CCS^+, EE^+ and CES^+, with regard to probabilities of misleading signals of Types III and IV.

Tables 6.6 and 6.7 provide values of these probabilities for the former and the latter groups of five joint schemes, respectively. Also, no head starts have been given to any of the individual schemes whose approximate *RL* has a phase-type distribution. Values of $PMS_{IV}(\delta)$, for $\delta < 0$, were ommitted from Table 6.6 by virtue of the fact that the run length $RL_\mu(\delta, 1)$ is identically distributed to $RL_\mu(-\delta, 1)$ for symmetric values of HS_μ, hence $PMS_{IV}(-\delta)$ for $HS_\mu = -\alpha \times 100\%$ equals $PMS_{IV}(\delta)$ for $HS_\mu = \alpha \times 100\%$, where $\alpha \in [0, 1)$.

Tables 6.6 and 6.7 and Figures 6.3 and 6.4 show how the use of joint schemes based on *CUSUM* and *EWMA* summary statistics can offer substantial improvement with regard to the (non)emission of *MS*s of Type III: schemes SS and SS^+ tend to produce this type of *MS*s more frequently for a wide range of the values of θ.

We can add that the joint schemes EE and CES are outperformed by schemes CC and CCS (respectively) in terms of *PMS*s of Type III. However, the joint schemes EE^+ and CES^+ appear to offer a slightly better performance than CC^+ and CCS^+, having in general lower *PMS*s of Type III.

The numerical results in Tables 6.6 and 6.7 and Figures 6.3 and 6.4 also suggest that the use of joint schemes CCS, CES, CCS^+ and CES^+ instead of CC, EE, CC^+ and EE^+ (respectively), causes in general an increase in *PMS*s of Type III. This is probably due to the fact that a combined scheme has two constituent charts, thus, two possible sources of *MS*s. However, note that there are a few instances where combined joint schemes appear to trigger slightly less *PMS*s of Type III than their “non combined” counterparts, for moderate and large values of θ.

Remark 6.4 gives a plausible justification for the nonmonotonicity of $PMS_{III}(\theta)$ for joint schemes based on upper one-sided individual charts for μ. In fact, $PMS_{III}(\theta)$ appears to be a nonmonotonous function of θ for the joint schemes CC^+ CCS^+, EE^+ and CES^+, according to the numerical results in **bold** in Table 6.7 and in Figure 6.4, thus, supporting the Conjectures (C1) and (C2).

Note that, with the exception of schemes EE and CES, the joint schemes that make use of a standard individual chart for μ have decreasing $PMS_{III}(\theta)$, i.e., the misinterpretation of a change in σ for a shift in μ becomes less likely as the inflation in σ increases.

Table 6.6: *PMS*s of Types III and IV for the joint schemes SS, CC, CCS, EE and CES (standard case).

	$PMS_{III}(\theta)$						$PMS_{IV}(\delta)$				
θ	SS	CC	CCS	EE	CES	δ	SS	CC	CCS	EE	CES
1.01	.487829	.392328	.426881	.456768	.472459	0.05	.496258	.458452	.470651	.471953	.481438
1.03	.465842	.272037	.335857	.380735	.412206	0.10	.486730	.335180	.371602	.404501	.433744
1.05	.445584	.180287	.249729	.318577	.351238	0.20	.451344	.193887	.239985	.249228	.299705
1.10	.401783	.057484	.107222	.214222	.230228	0.30	.400673	.131095	.172396	.143296	.184767
1.20	.337471	.005523	.029520	.124961	.122136	0.40	.343289	.096875	.132907	.084406	.112475
1.30	.294136	.000727	.015075	.092832	.086656	0.50	.286308	.075459	.107255	.052103	.071164
1.40	.263400	.000145	.009973	.078522	.072009	0.60	.234262	.060823	.089280	.033605	.047440
1.50	.240238	.000042	.007442	.071400	.065085	0.70	.189271	.050199	.076024	.022426	.033238
1.60	.221722	.000016	.005945	.067838	.061766	0.80	.151773	.042144	.065858	.015332	.024304
1.70	.206146	.000008	.004961	**.066312**	**.060427**	0.90	.121258	.035833	.057821	.010654	.018418
1.80	.192512	.000005	.004265	**.066071**	**.060288**	1.00	.096797	.030763	.051311	.007479	.014388
1.90	.180230	.000003	.003748	**.066698**	**.060987**	1.50	.032678	.015611	.031452	.001329	.005883
2.00	.168950	.000003	.003349	**.067936**	**.062554**	2.00	.012359	.008445	.021488	.000226	.003465
3.00	.088310	.000003	.001663	**.097349**	**.092315**	3.00	.002305	.002570	.012065	.000005	.001818

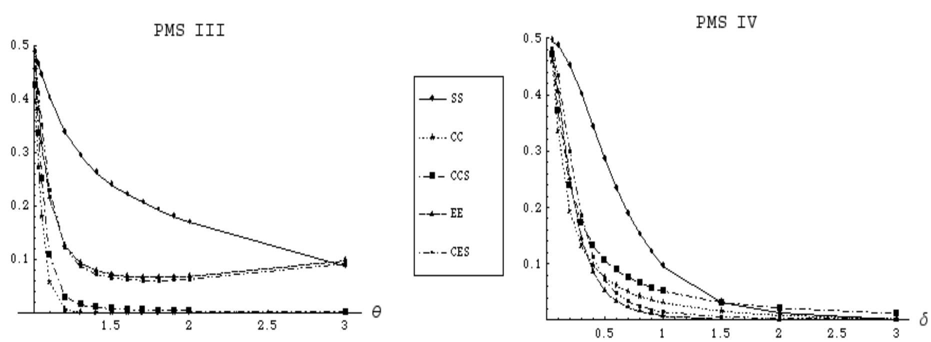

Figure 6.3: *PMS*s of Types III and IV for the joint schemes SS, CC, CCS, EE and CES (standard case).

Tables 6.6 and 6.7 and Figures 6.3 and 6.4 show that *MS*s of Type IV are more likely to happen in schemes SS and SS^+ than while using the remaining joint schemes for μ and σ.

All these latter joint schemes seem to have a similar behaviour in terms of the frequency of *PMS*s of Type IV, in particular when the individual charts for μ are upper one-sided, as shown by Figures 6.3 and 6.4.

For additional values of *PMS*s of Type III referring to $HS_\mu, HS_\sigma = 0\%, 50\%$[9] and of Type IV considering $HS_\mu = -50\%, 0\%, 50\%$, $HS_\sigma = 0\%, 50\%$, but a smaller range of values of θ and δ ($\theta = 1.05, 1.10, 1.20, 1.30, 1.40, 1.50, 2.00$ and $\delta = 0.05, 0.10, 0.25, 0.50, 0.75,$

[9]Please note that the run length $RL_{\mu,\sigma}(0,\theta)$ associated with the joint schemes when $\delta = 0$ has the same distribution for symmetric HS_μs. Therefore the *PMS*s of Type III are equal for symmetric HS_μs and, thus, values of $PMS_{III}(\theta)$ were not considered for $HS_\mu = -50\%$.

1.00, 1.50, 2.00[10]), please refer to Morais and Pacheco (2000b).

Table 6.7: *PMS*s of Types III and IV for the joint schemes SS^+, CC^+, EE^+, CCS^+ and CES^+ (upper one-sided case).

	$PMS_{III}(\theta)$						$PMS_{IV}(\delta)$				
θ	SS^+	CC^+	CCS^+	EE^+	CES^+	δ	SS^+	CC^+	CCS^+	EE^+	CES^+
1.01	.484676	.456547	.471625	.455274	.471752	0.05	.460162	.407942	.438859	.403991	.437911
1.03	.456701	.379127	.408637	.377092	.410277	0.10	.421864	.326254	.373455	.319232	.371239
1.05	.430911	.316409	.345546	.313194	.348189	0.20	.349949	.199310	.247793	.191651	.245285
1.10	.375334	.212859	.223748	.206131	.225089	0.30	.286075	.119085	.153533	.114210	.152790
1.20	.295048	.126933	.119744	.114615	.114963	0.40	.231295	.071855	.094389	.069767	.095342
1.30	.242637	.096917	.086868	.081130	.078150	0.50	.185599	.044411	.059826	.044152	.061665
1.40	.206805	.083728	.073449	.065605	.062268	0.60	.148269	.028177	.039686	.028898	.041781
1.50	.180893	.077166	.067059	.057295	.054102	0.70	.118230	.018281	.027593	.019432	.029617
1.60	.161108	.073836	.063830	.052531	.049520	0.80	.094298	.012058	.020030	.013327	.021845
1.70	.145270	**.072337**	**.062265**	.049768	.046871	0.90	.075349	.008040	.015103	.009262	.016671
1.80	.132095	**.071984**	**.062058**	.048249	**.045348**	1.00	.060389	.005395	.011776	.006491	.013103
1.90	.120806	**.072392**	**.063020**	**.047556**	**.044512**	1.50	.021323	.000736	.004941	.001126	.005488
2.00	.110920	**.073315**	**.064443**	**.047439**	**.044674**	2.00	.008458	.000092	.002980	.000185	.003272
3.00	.051170	**.092146**	**.086259**	**.059958**	**.058549**	3.00	.001644	.000001	.001544	.000004	.001717

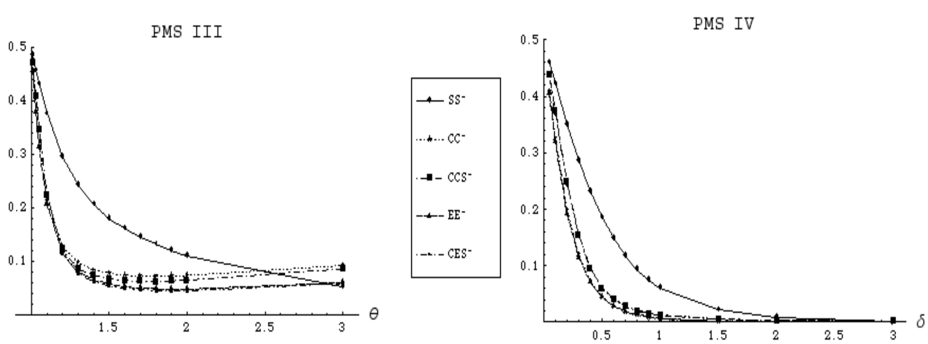

Figure 6.4: *PMS*s of Types III and IV for the joint schemes SS^+, CC^+, CCS^+, EE^+ and CES^+ (upper one-sided case).

Further justification for the behaviour of the *PMS*s of Types III and IV can be partially found in the $\log(ARL)$ profiles of the individual schemes for μ in Figures 6.5 and 6.6 or in the corresponding values in Tables 6.8 and 6.9.

For example, since $ARL_\sigma(\theta)$ does not depend on δ and all the individual schemes for σ have in-control ARL close to 500 samples we can roughly say that the larger the values of $RL_\mu(\delta,1)$ the larger those of $PMS_{IV}(\delta) = P[RL_\mu(\delta,1) > RL_\sigma(1)]$. This would explain why the curves of $ARL_\mu(\delta,1)$ and $PMS_{IV}(\delta)$ appear in the same order in the top right hand side graphs of Figures 6.5 and 6.6 and Figures 6.3 and 6.4.

[10] *PMS*s of Type IV for negative values of δ were omitted for the same reason pointed previously when we referred to Tables 6.6 and 6.7.

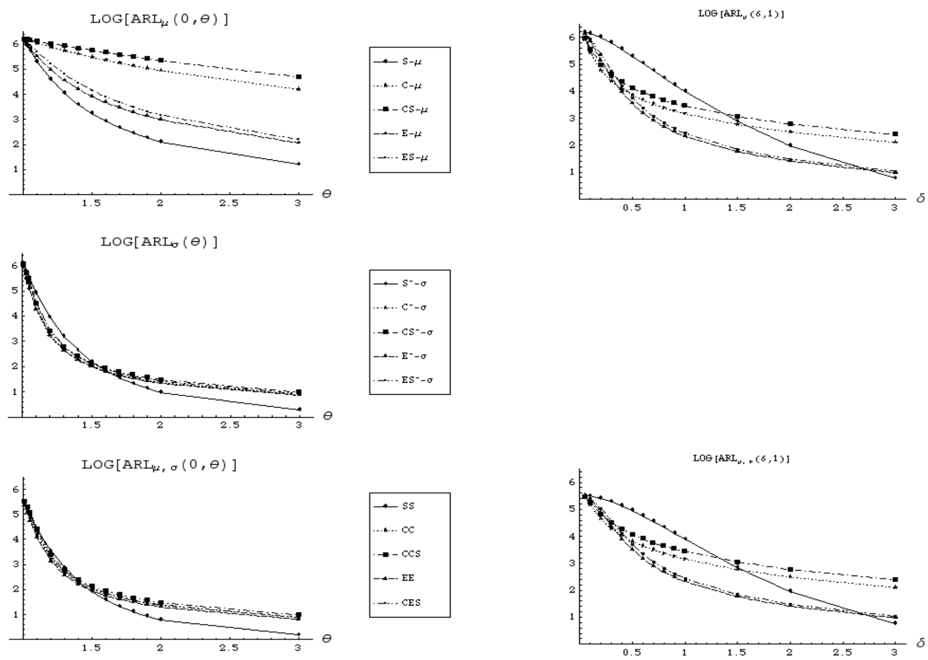

Figure 6.5: *ARL*s of individual and joint schemes for μ and σ (standard case).

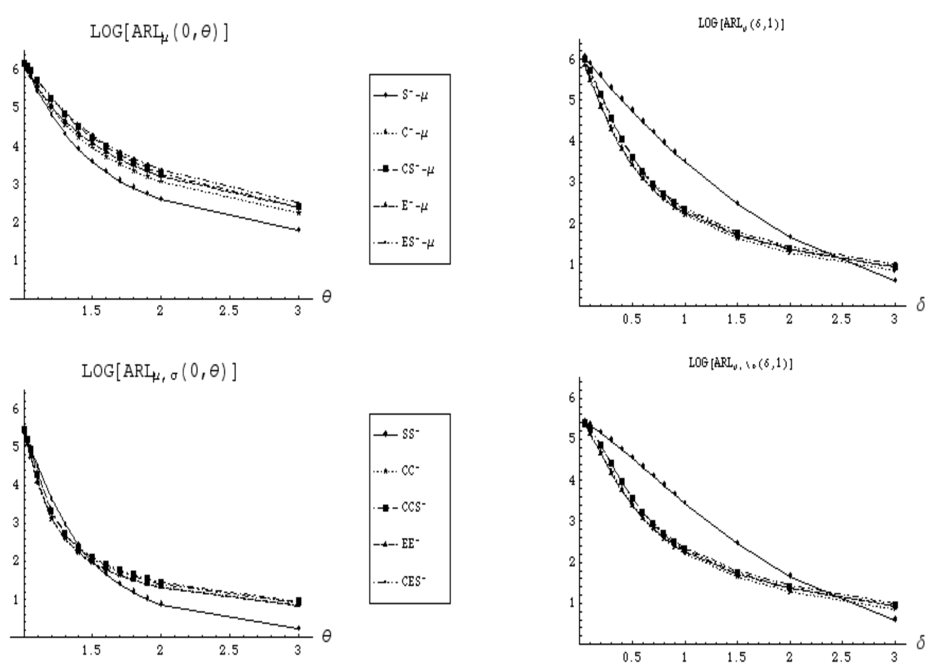

Figure 6.6: *ARL*s of individual schemes for μ and joint schemes for μ and σ (upper one-sided case).

In addition, the curves of $PMS_{III}(\theta)$ are practically in the inverse order of those of $ARL_\mu(0,\theta)$, as shown by the top graphs on the left hand side of Figures 6.5 and 6.6 and Figures 6.3 and 6.4. A crude explanation could be: a large value of $ARL_\mu(0,\theta)$ will tend to imply a small value of $PMS_{III}(\theta) = P[RL_\sigma(\theta) > RL_\mu(0,\theta)]$; note, however, that $ARL_\sigma(\theta)$ does depend on θ and surely influences $PMS_{III}(\theta)$.

It is worth noting that the graphs of $\ln[ARL_\mu(0,\theta)]$ in Figures 6.5 and 6.6 also give considerable evidence that there is no benefit in replacing *Shewhart*-type individual charts for μ by *CUSUM* or *EWMA*-based ones, when the process mean is on-target and there is a displacement in σ. As a matter of fact charts $S - \mu$ and $S^+ - \mu$ are far less sensitive to a shift in σ than the remaining individual charts for μ. However, supplementing the *CUSUM* and *EWMA*-based individual charts for μ with another one for σ gives extra protection and yields joint schemes with better performance than SS and SS^+, at least for small and moderate shifts in μ and σ, as suggested by Morais (1998) and Morais and Pacheco (2000a), and by Figures 6.5 and 6.6.

We would like to note in passing that additional numerical investigation suggest that the adoption of matched in-control combined *CUSUM-Shewhart* or *EWMA-Shewhart* schemes in place of *CUSUM* or *EMWA* did not yield a significant improvement in the *ARL*, as shown by Tables 6.8 and 6.9. It is worth mentioning that Lucas (1982) was only able to propose combined *CUSUM-Shewhart* schemes for μ which led to improved *ARL* curve for large shifts in μ, with only small changes both in the detection speed with regard to small values of δ and in the in-control *ARL* values, hence, this author did not compare matched in-control schemes. Based on all these and other computations reported in Morais and Pacheco (2000b), we conclude that the use of combined schemes, matched in-control with the corresponding "non combined" schemes, deserves more careful scrutiny in application.

6.5 Run length to a misleading signal (RLMS)

Another performance measure that also springs to mind is the number of sampling periods until a misleading signal is given by the joint scheme, the run length to a misleading signal (*RLMS*) of Type III, $RLMS_{III}(\theta)$, and of Type IV, $RLMS_{IV}(\delta)$. $RLMS_{III}(\theta)$ and $RLMS_{IV}(\delta)$ are improper random variables with an atom in $+\infty$ because the non occurrence of a misleading signal is an event with non-zero probability:

$$P[RLMS_{III}(\theta) = +\infty] = 1 - PMS_{III}(\theta) \qquad (6.24)$$

$$P[RLMS_{IV}(\delta) = +\infty] = 1 - PMS_{IV}(\delta). \qquad (6.25)$$

The next lemma not only adds the survival functions of the *RLMS*s of Types III and IV but also provides alternative expressions that prove to be useful in the investigation of the stochastic monotonicity properties of this performance measure.

Table 6.8: *ARL*s of individual and joint schemes for μ and σ (standard case).

	$ARL_{\mu}(0,\theta)$						$ARL_{\mu}(\delta,1)$				
θ	$S-\mu$	$C-\mu$	$CS-\mu$	$E-\mu$	$ES-\mu$	δ	$S-\mu$	$C-\mu$	$CS-\mu$	$E-\mu$	$ES-\mu$
1.01	451.251	491.245	495.290	461.639	476.311	0.05	493.572	369.878	389.250	449.206	465.876
1.03	370.673	474.400	486.031	396.282	431.224	0.10	475.145	228.388	260.233	342.792	384.446
1.05	307.724	458.395	476.986	343.131	389.922	0.20	412.318	117.510	146.124	171.110	216.088
1.10	201.414	421.720	455.309	247.542	303.007	0.30	335.269	78.433	100.758	89.933	116.518
1.20	99.816	360.362	414.947	144.613	187.984	0.40	262.370	58.875	76.808	52.937	67.274
1.30	57.309	311.456	377.915	94.629	124.044	0.50	201.582	47.142	62.048	34.545	42.581
1.40	36.640	271.891	344.923	67.090	87.151	0.60	153.964	39.323	52.031	24.490	29.328
1.50	25.391	239.455	315.665	50.423	64.578	0.70	117.729	33.739	44.800	18.509	21.648
1.60	18.715	212.547	289.820	39.586	49.966	0.80	90.465	29.553	39.339	14.687	16.866
1.70	14.472	189.986	266.312	32.136	40.021	0.90	69.995	26.298	35.00	12.094	13.695
1.80	11.626	170.889	245.250	26.780	32.958	1.00	54.585	23.696	31.642	10.248	11.480
1.90	9.629	154.584	226.493	22.790	27.755	1.50	17.891	15.900	21.310	5.789	6.307
2.00	8.175	140.552	209.781	19.728	23.802	2.00	7.257	12.015	16.113	4.078	4.392
3.00	3.301	66.473	109.184	7.810	8.940	3.00	2.155	8.148	10.904	2.644	2.819
	$ARL_{\sigma}(\theta)$						$ARL_{\sigma}(1)$				
θ	$S^{+}-\sigma$	$C^{+}-\sigma$	$CS^{+}-\sigma$	$E^{+}-\sigma$	$ES^{+}-\sigma$		$S^{+}-\sigma$	$C^{+}-\sigma$	$CS^{+}-\sigma$	$E^{+}-\sigma$	$ES^{+}-\sigma$
1.01	430.804	387.589	425.080	389.535	427.167		500.000	499.993	500.02	500.027	500.033
1.03	324.266	243.185	298.499	246.280	303.307						
1.05	248.318	161.017	207.283	163.868	212.635						
1.10	136.277	70.418	90.421	71.873	93.6764						
1.20	51.843	25.035	29.872	25.228	30.372						
1.30	24.881	14.007	16.150	13.917	16.035						
1.40	14.102	9.702	11.027	9.548	10.778						
1.50	9.029	7.508	8.473	7.343	8.201						
1.60	6.332	6.198	6.964	6.037	6.698						
1.70	4.758	5.331	5.973	5.178	5.721						
1.80	3.772	4.717	5.274	4.571	5.036						
1.90	3.117	4.258	4.755	4.120	4.530						
2.00	2.662	3.903	4.354	3.772	4.140						
3.00	1.320	2.436	2.664	2.364	2.534						
	$ARL_{\mu,\sigma}(0,\theta)$						$ARL_{\mu,\sigma}(\delta,1)$				
θ	SS	CC	CCS	EE	CES	δ	SS	CC	CCS	EE	CES
1.01	220.646	238.318	245.735	214.552	227.373	0.05	248.633	233.719	238.225	240.010	243.276
1.03	173.209	178.971	200.300	154.981	180.47	0.10	243.878	173.108	189.248	206.795	219.663
1.05	137.671	133.266	157.237	113.715	140.057	0.20	226.221	103.637	124.167	130.327	153.296
1.10	81.523	66.758	81.519	57.788	73.630	0.30	200.936	72.763	90.746	78.138	96.377
1.20	34.348	24.928	29.170	22.720	27.382	0.40	172.301	55.938	71.220	49.090	60.560
1.30	17.563	14.000	15.978	13.021	15.068	0.50	143.868	45.407	58.536	33.105	40.062
1.40	10.388	9.701	10.955	9.080	10.296	0.60	117.897	38.210	49.647	23.883	28.250
1.50	6.860	7.508	8.433	7.039	7.895	0.70	95.446	32.983	43.092	18.227	21.127
1.60	4.928	6.198	6.939	5.811	6.474	0.80	76.735	29.018	38.065	14.545	16.587
1.70	3.777	5.331	5.956	4.994	5.541	0.90	61.508	25.909	34.089	12.019	13.532
1.80	3.046	4.717	5.261	4.413	4.882	1.00	49.302	23.405	30.867	10.206	11.378
1.90	2.555	4.258	4.745	3.978	4.392	1.50	17.307	15.815	21.004	5.786	6.286
2.00	2.212	3.903	4.345	3.640	4.011	2.00	7.167	11.984	15.957	4.078	4.384
3.00	1.203	2.436	2.661	2.238	2.409	3.00	2.150	8.143	10.843	2.644	2.816

Table 6.9: ARLs of individual schemes for μ and joint schemes for μ and σ (upper one-sided case).

	$ARL_\mu(0,\theta)$						$ARL_\mu(\delta,1)$				
θ	$S^+-\mu$	$C^+-\mu$	$CS^+-\mu$	$E^+-\mu$	$ES^+-\mu$	δ	$S^+-\mu$	$C^+-\mu$	$CS^+-\mu$	$E^+-\mu$	$ES^+-\mu$
1.01	456.983	461.525	476.531	463.720	477.296	0.05	427.203	349.407	393.685	342.199	391.100
1.03	384.563	396.027	431.831	401.451	434.155	0.10	365.849	247.914	300.905	239.007	297.105
1.05	326.623	342.871	390.994	350.429	394.466	0.20	270.170	131.341	168.372	124.546	165.376
1.10	225.140	247.531	304.779	257.588	310.552	0.30	201.354	74.991	95.017	71.156	93.846
1.20	121.479	145.263	190.941	155.421	198.070	0.40	151.445	46.318	56.825	44.494	56.835
1.30	74.541	95.735	127.404	104.501	134.139	0.50	114.948	30.871	36.693	30.193	37.228
1.40	50.253	68.46	90.590	75.849	96.453	0.60	88.040	22.032	25.556	21.955	26.253
1.50	36.355	51.950	67.963	58.204	72.981	0.70	68.041	16.656	19.000	16.880	19.700
1.60	27.762	41.199	53.260	46.561	57.549	0.80	53.058	13.192	14.884	13.557	15.531
1.70	22.112	33.795	43.215	38.451	46.907	0.90	41.745	10.838	12.141	11.265	12.724
1.80	18.211	28.462	35.985	32.556	39.266	1.00	33.135	9.164	10.216	9.610	10.740
1.90	15.406	24.480	30.583	28.118	33.591	1.50	11.894	5.139	5.670	5.529	6.009
2.00	13.322	21.418	26.486	24.679	29.185	2.00	5.265	3.603	3.958	3.923	4.214
3.00	5.928	9.4160	10.977	10.886	12.174	3.00	1.823	2.342	2.544	2.558	2.718
	$ARL_{\mu,\sigma}(0,\theta)$						$ARL_{\mu,\sigma}(\delta,1)$				
θ	SS^+	CC^+	CCS^+	EE^+	CES^+		SS^+	CC^+	CCS^+	EE^+	CES^+
1.01	222.004	213.870	226.934	215.125	227.672	0.05	230.621	208.849	222.491	206.544	221.726
1.03	176.173	153.681	179.023	155.856	181.045	0.10	211.510	168.670	190.144	164.803	188.715
1.05	141.315	112.297	138.040	114.564	140.688	0.20	175.624	106.212	127.987	101.961	126.346
1.10	85.127	56.840	71.863	58.322	74.078	0.30	143.752	66.700	81.341	63.789	80.533
1.20	36.547	22.553	27.072	22.930	27.559	0.40	116.416	43.374	52.034	41.840	52.056
1.30	18.844	13.083	15.206	13.136	15.164	0.50	93.614	29.732	34.850	29.131	35.328
1.40	11.186	9.202	10.539	9.158	10.362	0.60	74.986	21.552	24.762	21.487	25.405
1.50	7.396	7.175	8.155	7.100	7.947	0.70	59.997	16.438	18.616	16.656	19.278
1.60	5.312	5.946	6.729	5.863	6.519	0.80	48.055	13.086	14.678	13.444	15.300
1.70	4.067	5.125	5.783	5.041	5.582	0.90	38.599	10.785	12.021	11.204	12.587
1.80	3.273	4.539	5.111	4.457	4.921	1.00	31.134	9.136	10.140	9.576	10.652
1.90	2.740	4.098	4.607	4.019	4.430	1.50	11.640	5.137	5.654	5.527	5.990
2.00	2.367	3.754	4.216	3.680	4.049	2.00	5.221	3.603	3.951	3.922	4.206
3.00	1.252	2.314	2.542	2.287	2.455	3.00	1.820	2.342	2.542	2.558	2.715

Lemma 6.5 — *Let $RLMS_{III}(\theta)$ and $RLMS_{IV}(\delta)$ denote the* RLMS*s of Types III and IV of any of the joints schemes in Table 6.1. Then*

$$\overline{F}_{RLMS_{III}(\theta)}(m) = 1 - \sum_{i=1}^{m} P_{RL_\mu(0,\theta)}(i) \times \overline{F}_{RL_\sigma(\theta)}(i) \quad (6.26)$$

$$= 1 - \sum_{i=1}^{m} F_{RL_\mu(0,\theta)}(i) \times P_{RL_\sigma(\theta)}(i+1) - F_{RL_\mu(0,\theta)}(m) \times \overline{F}_{RL_\sigma(\theta)}(m+1), \ \theta > 1 \quad (6.27)$$

$$\overline{F}_{RLMS_{IV}(\delta)}(m) = 1 - \sum_{i=1}^{m} \overline{F}_{RL_\mu(\delta,1)}(i) \times P_{RL_\sigma(1)}(i) \quad (6.28)$$

$$= 1 - \sum_{i=1}^{m} P_{RL_\mu(\delta,1)}(i+1) \times F_{RL_\sigma(1)}(i) - \overline{F}_{RL_\mu(\delta,1)}(m+1) \times F_{RL_\sigma(1)}(m), \ \delta \neq 0 \quad (6.29)$$

(or $\delta > 0$ when upper one-sided schemes for μ are at use), for any positive integer m.

Exact expressions for the survival functions of the $RLMS$ of Types III and IV of the schemes SS and SS^+ can be found in Table 6.10.

Table 6.10: Exact survival functions of *RLMS*s of Types III and IV for schemes SS and SS^+.

Joint scheme	$\overline{F}_{RLMS_{III}(\theta)}(m)$	$\overline{F}_{RLMS_{IV}(\theta)}(m)$
SS	$1 - \frac{(1-a)b[1-(ab)^m]}{1-ab}$	$1 - \frac{(1-d)c[1-(cd)^m]}{1-cd}$
SS^+	$1 - \frac{(1-e)b[1-(eb)^m]}{1-eb}$	$1 - \frac{(1-d)f[1-(fd)^m]}{1-fd}$

$a = \Phi(\xi_\mu/\theta) - \Phi(-\xi_\mu/\theta)$; $b = F_{\chi^2_{n-1}}(\xi^+_\sigma/\theta^2)$; $c = \Phi(\xi_\mu - \delta) - \Phi(-\xi_\mu - \delta)$; $d = F_{\chi^2_{n-1}}(\xi^+_\sigma)$; $e = \Phi(\xi^+_\mu/\theta)$; $f = \Phi(\xi^+_\mu - \delta)$

Additionally, if we recall the exact expressions of the *PMS*s of the joint schemes SS and SS^+ we get

$$F_{RLMS_{III}(\theta)}(m) = PMS_{III}(\theta) \times F_{RL_{\mu,\sigma}(0,\theta)}(m) \quad (6.30)$$

$$F_{RLMS_{IV}(\delta)}(m) = PMS_{IV}(\delta) \times F_{RL_{\mu,\sigma}(\delta,1)}(m) \quad (6.31)$$

where $F_{RL_{\mu,\sigma}(\delta,\theta)}(m) = 1 - \overline{F}_{RL_\mu(\delta,\theta)}(m) \times \overline{F}_{RL_\sigma(\theta)}(m)$. These alternative expressions of the distribution function of the *RLMS*s are due to the fact that the constituent *Shewhart* charts of schemes SS and SS^+ deal in any case with (time–)independent summary statistics.

Approximations to these performance measures for the remaining joint schemes are obviously obtained by replacing the survival functions in Equations (6.26) and (6.28) by the corresponding Markov approximations and Equations (6.30) and (6.31) do not hold in this case because of the (time–)dependence structure of their summary statistics.

Theorem 6.6 — *The stochastic monotonicity properties (1)–(16) below hold for the (exact)* RLMS*s of Types III and IV of the joint schemes SS^+, CC^+, CCS^+, EE^+ and CES^+.*

Joint scheme	Type III ($\delta = 0, \theta > 1$)	Type IV ($\delta > 0, \theta = 1$)
SS^+	(C3) $RLMS_{III}(\theta)$*	(7) $RLMS_{IV}(\delta) \uparrow_{st}$ with δ
CC^+, CCS^+	(1) $RLMS_{III}(\theta) \downarrow_{st}$ with α	(8) $RLMS_{IV}(\delta) \uparrow_{st}$ with α
	(2) $RLMS_{III}(\theta) \uparrow_{st}$ with β	(9) $RLMS_{IV}(\delta) \downarrow_{st}$ with β
	(C4) $RLMS_{III}(\theta)$*	(10) $RLMS_{IV}(\delta) \uparrow_{st}$ with δ
	(3) $RLMS_{III}(\theta) \uparrow_{st}$ with k_μ^+	(11) $RLMS_{IV}(\theta) \downarrow_{st}$ with k_μ^+
	(4) $RLMS_{III}(\theta) \downarrow_{st}$ with k_σ^+	(12) $RLMS_{IV}(\theta) \uparrow_{st}$ with k_σ^+
EE^+, CES^+	(5) $RLMS_{III}(\theta) \downarrow_{st}$ with α	(14) $RLMS_{IV}(\delta) \uparrow_{st}$ with α
	(6) $RLMS_{III}(\theta) \uparrow_{st}$ with β	(15) $RLMS_{IV}(\delta) \downarrow_{st}$ with β
	(C5) $RLMS_{III}(\theta)$*	(16) $RLMS_{IV}(\delta) \uparrow_{st}$ with δ

* (C3, C4, C5): without stochastic monotonous behaviour, in the usual sense, regarding θ

Proof — The stochastic monotonicity properties described in Theorem 6.6 come as no surprise — they point in the opposite direction of the monotone behaviour of the corresponding *PMS*s.[11] These properties are ensured by the stochastic monotonicity properties of $RL_\mu(\delta, \theta)$ and $RL_\sigma(\theta)$ and Equations (6.26)–(6.29).

For example, the increasing behaviour of the survival function of $RLMS_{IV,SS^+}(\delta)$ follows from (6.28) and the fact that $RL_\mu(\delta, 1)$ stochastically decreases with δ.

The run length to a misleading signal of Type III of the joint schemes CC^+, CCS^+, EE^+ and CES^+, $RLMS_{III}(\theta)$, stochastically decreases with α. This conclusion can be immediately drawn from (6.27) because: $RL_\sigma^\beta(1)$ does not depend on the head start α; and $RL_\mu^\alpha(0, \theta)$ stochastically decreases with α, thus $F_{RL_\mu(0,\theta)}(i)$ increases with α for any i. However, this result could not be drawn from (6.26) because $P_{RL_\mu^\alpha(0,\theta)}(i)$ is not an increasing function of α, although $RL_\mu^\alpha(0, \theta)$ stochastically decreases with α.

Similarly, to prove that $RLMS_{IV}(\delta)$ stochastically decreases with β, we have to use (6.29) instead of (6.28). •

We ought to add that, based on the percentage points of $RLMS_{III}(\theta)$ numerically obtained in the next section and by Morais and Pacheco (2000b), we conjecture that $RLMS_{III}(\theta)$ has no stochastic behaviour, in terms of θ, for none of the five upper one-sided joint schemes under investigation.

[11] Except for Conjecture (C3) which refers to the joint scheme SS^+ that we proved it has decreasing $PMS_{III}(\theta)$.

6.6 RLMS: numerical illustrations

The percentage points of *RLMS* are crucial because this random variable has no expected value or any other moment; and note that any $p \times 100\%$ percentage point of probability, for p equal or larger than *PMS*, is equal to $+\infty$, as illustrated by Tables 6.11–6.13.

The presentation of the numerical results concerning *RLMS* follows closely the one in Section 6.4. Thus, we use the same constelation of parameters for the ten joint schemes and we begin with an illustration of some stochastic monotonicity properties of the *RLMS* of the joint scheme EE^+.

Table 6.11: Percentage points of *RLMS* of Type III and Type IV of the joint scheme EE^+ (listed in order corresponding to $p \times 100\% = 1\%, 5\%, 10\%, 15\%, 20\%$ percentage points, for each θ and each δ); percentage points of $RL_{\mu,\sigma}(0,\theta)$ and $RL_{\mu,\sigma}(\delta,1)$ are in parenthesis.

	$F^{-1}_{RLMS_{III}(\theta)}(p)$ $(F^{-1}_{RL_{\mu,\sigma}(0,\theta)}(p))$					$F^{-1}_{RLMS_{IV}(\delta)}(p)$ $(F^{-1}_{RL_{\mu,\sigma}(\delta,1)}(p))$			
	HS_σ					HS_σ			
	0%		50%			0%		50%	
	HS_μ		HS_μ			HS_μ		HS_μ	
θ	0%	50%	0%	50%	δ	0%	50%	0%	50%
1.01	11 (8)	3 (3)	11 (3)	4 (2)	0.05	12 (8)	12 (3)	3 (3)	3 (2)
	31 (18)	17 (12)	32 (8)	18 (5)		34 (17)	35 (10)	11 (9)	12 (5)
	59 (29)	43 (23)	62 (18)	46 (12)		64 (28)	66 (21)	39 (19)	40 (12)
	90 (41)	73 (35)	96 (30)	78 (24)		100 (40)	104 (32)	71 (31)	75 (23)
	127 (53)	109 (47)	136 (43)	116 (36)		143 (52)	151 (44)	111 (43)	116 (35)
1.03	10 (7)	3 (3)	11 (3)	3 (2)	0.10	12 (8)	12 (3)	3 (3)	3 (2)
	28 (15)	14 (10)	30 (6)	15 (5)		34 (15)	35 (8)	11 (8)	12 (5)
	53 (23)	37 (18)	57 (12)	40 (9)		66 (24)	70 (16)	39 (17)	42 (10)
	82 (31)	65 (26)	90 (21)	71 (16)		107 (33)	114 (25)	75 (26)	80 (18)
	119 (40)	98 (35)	**132** (30)	109 (25)		162 (42)	176 (35)	120 (35)	130 (28)
1.05	10 (7)	3 (3)	10 (2)	3 (2)	0.20	12 (7)	12 (3)	3 (3)	3 (2)
	25 (12)	12 (8)	27 (5)	13 (4)		36 (12)	38 (6)	11 (7)	12 (4)
	48 (18)	33 (14)	53 (9)	36 (7)		77 (17)	85 (10)	42 (13)	47 (7)
	77 (24)	58 (21)	**87** (15)	66 (11)		152 (23)	183 (16)	90 (18)	103 (12)
	116 (31)	92 (27)	**136** (21)	**106** (17)		$+\infty$ (28)	$+\infty$ (21)	189 (24)	261 (17)
1.10	8 (5)	3 (3)	9 (2)	3 (2)	0.30	12 (6)	13 (2)	3 (3)	3 (2)
	21 (9)	9 (6)	23 (4)	10 (3)		40 (9)	45 (4)	12 (7)	13 (4)
	41 (12)	24 (10)	49 (5)	29 (4)		125 (13)	234 (7)	49 (10)	60 (6)
	73 (15)	47 (13)	**105** (8)	61 (6)		$+\infty$ (16)	$+\infty$ (10)	225 (14)	$+\infty$ (8)
	188 (18)	89 (16)	$+\infty$ (10)	**164** (8)		$+\infty$ (20)	$+\infty$ (13)	$+\infty$ (17)	$+\infty$ (11)

Table 6.11 and the following ones include $1\%, 5\%, 10\%, 15\%, 20\%$ percentage points of $RLMS_{III}(\theta)$ and $RLMS_{IV}(\delta)$ for a smaller range of θ and δ values: $\theta = 1.01, 1.03, 1.05,$ 1.10 and $\delta = 0.05, 0.10, 0.20, 0.30$. Also, in order to give the user an idea of how quick a misleading signal of Type III and Type IV is triggered by a joint scheme, the correspond-

ing percentage points of $RL_{\mu,\sigma}(0,\theta)$ and $RL_{\mu,\sigma}(\delta,1)$ have been added in parenthesis to Tables 6.11–6.13 and, thus, can be compared to the corresponding percentage points of $RLMS_{III}(\theta)$ and $RLMS_{IV}(\delta)$.

The results in Table 6.11 illustrate the findings concerning the *RLMS*s stochastic monotonicity properties. The emission of *MS*s of Type III is indeed speeded up by the adoption of a head start HS_μ; giving a head start to the individual chart for σ has exactly the opposite effect. Besides this $RLMS_{IV}(\delta)$ stochastically increases with δ, which means that *MS* of Type IV will tend to occur later as the increase in μ becomes more severe. However, the entries in **bold** in Table 6.11 show that a few percentage points of $RLMS_{III}(\theta)$ do not decrease with θ.

Table 6.11 also gives the reader an idea of how soon misleading signals can occur. For instance, the probability of triggering a misleading signal of Type III within the first 59 samples is of at least 0.10 when there is a shift of 1% in the process standard deviation and no head start has been adopted for scheme EE^+.

Tables 6.12 and 6.13 in this section allow us to assess and compare the performance of all the joint schemes in terms of number of sampling periods until the emission of misleading signals. The examination of these two tables leads to overall conclusions similar to those referring to the *PMS*s.

According to the percentage points in Table 6.12, the joint scheme SS tends to produce *MS*s of Type III considerably sooner than scheme CC, and the schemes EE, CCS and CES appear to trigger them later than scheme SS. This probably comes by virtue of the fact that $PMS_{III}(\theta)$ of scheme SS is larger than the corresponding *PMS*s of the remaining joint schemes, as mentioned in Section 6.4. Note, however, that the percentage points of $RLMS_{III}(\theta)$ for schemes SS^+, CC^+, CCS^+, EE^+ and CES^+ differ much less than when standard individual charts for μ are at use, as apparent in Table 6.13.

The schemes CC and CCS tend to require more samples to trigger *MS*s of Type III than joint schemes EE and CES, respectively. On the other hand, scheme CC^+ appears to give such *MS*s almost as late as scheme EE^+. The same seems to hold for schemes CCS^+ and CES^+.

Supplementing *Shewhart* upper control limits to the individual *CUSUM* charts for μ and σ of the joint CC scheme seems to substantially speed up the emission of *MS* of Type III. The same comments do not stand for schemes EE and CES, or CC^+ and CCS^+, or even EE^+ and CES^+.

A brief remark on the percentage points of $RLMS_{III,SS^+}(\theta)$: all the values suggest that this random variable stochastically decreases as the shift in the process standard deviation increases. However, recall that we proved that $PMS_{III,SS^+}(\theta)$ decreases with θ. Thus, if we had considered larger values of θ, as we did in the additional investigations or in Table G1.b in Morais and Pacheco (2000b), we would soon get percentage points equal to $+\infty$ instead of smaller ones. In addition, note that Conjectures (C4) and (C5) concerning the remaining schemes CC^+, CCS^+, EE^+ and CES^+ are conveniently supported by the numerical results in **bold** in Table 6.13.

As for *RLMS*s of Type IV, it is interesting to notice that the joint schemes SS and SS^+ do not show such a poor performance when compared to their Markov-type counterparts

Table 6.12: Percentage points of *RLMS*s of Type III and Type IV (listed in order corresponding, to $p \times 100\% = 1\%, 5\%, 10\%, 15\%, 20\%$ percentage points, for each θ and each δ); percentage points of $RL_{\mu,\sigma}(0,\theta)$ and $RL_{\mu,\sigma}(\delta,1)$ are in parenthesis.

	$F^{-1}_{RLMS_{III}(\theta)}(p)$ $(F^{-1}_{RL_{\mu,\sigma}(0,\theta)}(p))$						$F^{-1}_{RLMS_{IV}(\delta)}(p)$ $(F^{-1}_{RL_{\mu,\sigma}(\delta,1)}(p))$				
θ	*SS*	*CC*	*CCS*	*EE*	*CES*	δ	*SS*	*CC*	*CCS*	*EE*	*CES*
1.01	5 (3)	64 (11)	11 (5)	11 (8)	8 (5)	0.05	6 (3)	12 (12)	9 (5)	12 (9)	9 (5)
	24 (12)	105 (27)	54 (19)	30 (17)	29 (16)		27 (13)	33 (33)	31 (20)	33 (19)	31 (17)
	51 (24)	144 (47)	110 (35)	58 (29)	57 (28)		56 (27)	60 (56)	59 (38)	63 (31)	60 (30)
	81 (36)	184 (65)	159 (52)	89 (41)	89 (41)		90 (41)	89 (74)	90 (57)	96 (45)	94 (43)
	117 (50)	231 (82)	208 (70)	126 (53)	127 (54)		129 (56)	122 (90)	124 (76)	136 (59)	133 (58)
1.03	4 (2)	63 (9)	11 (5)	10 (7)	8 (5)	0.10	6 (3)	12 (12)	9 (5)	12 (8)	9 (5)
	20 (9)	106 (20)	56 (17)	27 (14)	27 (14)		27 (13)	33 (33)	31 (20)	34 (17)	31 (16)
	42 (19)	151 (33)	114 (29)	51 (22)	53 (24)		56 (26)	60 (53)	59 (38)	64 (28)	61 (28)
	68 (29)	204 (45)	168 (42)	80 (31)	83 (34)		90 (40)	90 (67)	90 (57)	100 (40)	96 (40)
	97 (39)	282 (58)	228 (56)	117 (40)	121 (44)		129 (55)	127 (79)	126 (75)	143 (52)	137 (53)
1.05	4 (2)	62 (8)	11 (5)	9 (7)	8 (5)	0.20	6 (3)	12 (12)	9 (5)	12 (7)	9 (5)
	17 (8)	109 (15)	57 (15)	24 (12)	25 (13)		27 (12)	33 (32)	31 (20)	35 (14)	31 (14)
	35 (15)	169 (24)	121 (24)	46 (18)	49 (20)		57 (24)	61 (45)	59 (38)	70 (21)	64 (22)
	57 (23)	277 (33)	188 (34)	74 (24)	79 (28)		92 (37)	101 (53)	92 (55)	120 (28)	106 (30)
	82 (31)	$+\infty$ (42)	287 (44)	112 (31)	117 (36)		133 (51)	$+\infty$ (60)	140 (68)	206 (35)	166 (39)
1.10	3 (1)	64 (6)	11 (5)	8 (5)	7 (5)	0.30	6 (3)	12 (12)	9 (5)	12 (6)	9 (5)
	11 (5)	179 (10)	69 (11)	19 (9)	21 (10)		27 (11)	33 (29)	31 (20)	37 (11)	33 (11)
	24 (9)	$+\infty$ (14)	236 (16)	38 (12)	43 (14)		58 (22)	66 (38)	59 (38)	91 (15)	74 (17)
	38 (14)	$+\infty$ (18)	$+\infty$ (20)	67 (15)	75 (18)		95 (33)	$+\infty$ (43)	102 (51)	$+\infty$ (19)	153 (22)
	56 (19)	$+\infty$ (21)	$+\infty$ (25)	145 (18)	141 (22)		139 (45)	$+\infty$ (48)	$+\infty$ (59)	$+\infty$ (24)	$+\infty$ (28)

CC, EE, CCS, CES and CC^+, EE^+, CCS^+, CES^+, respectively.

Also, joint schemes CC, EE, CCS and CES seem to have similar performance as far as $RLMS_{IV}(\delta)$ is concerned, as suggested earlier by the $PMS_{IV}(\delta)$ values in Table 6.6 and by Figure 6.3. The same appears to happen with schemes CC^+, EE^+, CCS^+, CES^+.

In addition, the use of combined schemes yields in general smaller values for the percentage points of $RLMS_{IV}(\delta)$; that is, the MSs of Type IV tend to be triggered sooner.

It is worth mentioning that the *RLMS* is important to assess the performance (in terms of misleading signals) of schemes for μ and σ based on univariate summary statistics, such as the Shewhart type scheme proposed by Chengalur, Arnold and Reynolds Jr. (1989) whose statistic is $n^{-1}\sum_{i=1}^{n}(\frac{X_i-\mu_0}{\sigma_0})^2$. This comes by virtue of the fact that in such cases we cannot define *PMS*. The *RLMS*s of those schemes are particular cases of the run length of the joint scheme itself:

$$RLMS_{III}(\theta) =_{st} RL_{\mu,\sigma}(0,\theta) \tag{6.32}$$

$$RLMS_{IV}(\delta) =_{st} RL_{\mu,\sigma}(\delta,1). \tag{6.33}$$

Table 6.13: Percentage points of *RLMS*s of Type III and Type IV (listed in order corresponding to $p \times 100\% = 1\%, 5\%, 10\%, 15\%, 20\%$ percentage points, for each θ and each δ); percentage points of $RL_{\mu,\sigma}(0,\theta)$ and $RL_{\mu,\sigma}(\delta,1)$ are in parenthesis.

	$F^{-1}_{RLMS_{III}(\theta)}(p)$ $(F^{-1}_{RL_{\mu,\sigma}(0,\theta)}(p))$						$F^{-1}_{RLMS_{IV}(\delta)}(p)$ $(F^{-1}_{RL_{\mu,\sigma}(\delta,1)}(p))$				
θ	SS^+	CC^+	CCS^+	EE^+	CES^+	δ	SS^+	CC^+	CCS^+	EE^+	CES^+
1.01	5 (3)	10 (8)	8 (5)	11 (8)	9 (5)	0.05	6 (3)	13 (8)	9 (5)	12 (8)	9 (5)
	25 (12)	29 (17)	29 (16)	31 (18)	29 (16)		27 (12)	34 (17)	31 (16)	34 (17)	31 (16)
	52 (24)	57 (29)	57 (28)	59 (29)	57 (28)		57 (25)	65 (28)	61 (28)	64 (28)	61 (28)
	83 (36)	88 (40)	89 (41)	90 (41)	90 (41)		91 (38)	101 (39)	96 (40)	100 (40)	95 (40)
	118 (50)	125 (53)	126 (54)	127 (53)	127 (54)		132 (52)	144 (52)	138 (53)	143 (52)	137 (53)
1.03	4 (2)	9 (7)	8 (5)	10 (7)	8 (5)	0.10	6 (3)	13 (7)	9 (5)	12 (8)	9 (5)
	21 (10)	26 (14)	27 (14)	28 (15)	27 (14)		27 (11)	35 (15)	32 (14)	34 (15)	31 (15)
	44 (19)	50 (22)	53 (24)	53 (23)	54 (24)		58 (23)	67 (23)	63 (25)	66 (24)	62 (25)
	70 (29)	79 (31)	83 (34)	82 (31)	84 (34)		93 (35)	108 (33)	100 (35)	107 (33)	99 (35)
	102 (40)	115 (40)	120 (44)	119 (40)	122 (45)		136 (48)	162 (43)	147 (46)	162 (42)	146 (46)
1.05	4 (2)	8 (6)	7 (5)	10 (7)	8 (5)	0.20	6 (2)	13 (6)	9 (5)	12 (7)	9 (5)
	18 (8)	23 (12)	25 (13)	25 (12)	26 (13)		27 (9)	37 (11)	33 (12)	36 (12)	32 (12)
	38 (15)	45 (18)	49 (20)	48 (18)	50 (20)		59 (19)	77 (17)	68 (19)	77 (17)	67 (19)
	61 (23)	**73** (24)	**79** (28)	**77** (24)	**81** (28)		99 (29)	147 (23)	119 (26)	152 (23)	118 (26)
	88 (32)	**111** (30)	**118** (36)	**116** (31)	**120** (36)		149 (40)	$+\infty$ (29)	206 (33)	$+\infty$ (28)	207 (33)
1.10	3 (1)	7 (5)	7 (5)	8 (5)	7 (5)	0.30	6 (2)	13 (5)	9 (4)	12 (6)	9 (5)
	13 (5)	18 (9)	21 (10)	21 (9)	22 (10)		28 (8)	41 (9)	35 (10)	40 (9)	34 (10)
	27 (9)	37 (12)	43 (14)	41 (12)	45 (14)		62 (16)	119 (12)	84 (14)	125 (13)	83 (15)
	44 (14)	**66** (15)	**76** (18)	**73** (15)	**79** (18)		107 (24)	$+\infty$ (16)	288 (19)	$+\infty$ (16)	299 (19)
	65 (19)	**145** (18)	**149** (22)	**188** (18)	**153** (22)		173 (32)	$+\infty$ (20)	$+\infty$ (23)	$+\infty$ (20)	$+\infty$ (24)

The numerical results obtained along this chapter suggest that the schemes SS and SS^+ compare unfavorably to the more sophisticated joint schemes CC, CC^+, etc., in terms of *MS*s of both types, in most cases. Thus, the SS and SS^+ schemes are far from being reliable in identifying which parameter has changed. This is the answer for St. John and Bragg's (1991) concluding question: — (Misleading signals can be a serious problem for the user of joint charts.) *Would alternatives (EWMA or CUSUM) perform better in this regard?*.

Tables 6.6 and 6.7 give the distinct impression that joint schemes for μ and σ can be very sensitive to *MS*s of both types: the values of *PMS*s are far from negligible, especially for small and moderate shifts in μ and σ, thus, misidentification of signals is likely to occur.

The practical significance of all these results will depend on the amount of time and money that is spent in attempting to identify and correct nonexisting causes of variation in μ (σ), i.e., when a *MS* of Type III (IV) occurs.

No monotonicity properties results have been added whatsoever to the *RL*s of standard individual schemes for μ in Chapter 5, and therefore to the *PMS*s and the *RLMS*s of the joint schemes CC, CCS, EE and CES; scheme SS followed suit. This is due to the fact that the constituent individual schemes for μ are not associated with stochastically monotone matrices, an absolutely crucial characteristic to prove the stochastic monotonicity results of the *RL*s and, thus, the monotonicity properties of the *PMS*s and the stochastic monotonicity properties of the *RLMS*s.

Finally, we would like to refer that Reynolds Jr. and Stoumbos (2001) advocate that these probabilities are useful to diagnose which parameter(s) has(have) changed, and suggest the use of the pattern of the points beyond the control limits of the constituent charts in the identification of the parameter that has effectively changed. A plausible justification for this diagnostic aid stems from the fact that changes in μ and σ have different impacts in those patters.

Chapter 7

Assessing the impact of autocorrelation in the performance of residual schemes for μ

Assuming that the observations from the process output are independent is a standard assumption when developing a control scheme. However, this assumption can be totally irrealistic and significantly affect the performance of standard control schemes — mistakenly designed to detect departures from in-control parameter values of independent data.

The effects and implications of autocorrelation have been frequently addressed in the Statistical Process Control (SPC) literature. For a detailed review please refer to Knoth and Schmid (2001). These issues are usually tackled only numerically: take for instance the investigations by Johnson and Bagshaw (1974), Alwan (1992), Maragah and Woodall (1992), Wardell, Moskowitz and Plante (1994), Runger, Willemain and Prabhu (1995), VanBrackle III and Reynolds Jr. (1997) and Lu and Reynolds Jr. (1999a). These and several other papers have tables and graphs, usually referring to the ARL, to provide evidence that the performance of the appealing traditional control schemes is severely compromised by the presence of serial correlation.

Analytical investigations on the effects and implications of autocorrelation have been so far presented in a few papers, e.g., Schmid (1995, 1997a, 1997b), Schmid and Schöne (1997), Schöne, Schmid and Knoth (1999) and Kramer and Schmid (2000), already reviewed in Chapter 1. These papers provide in general a comparison between the survival functions of the RLs of two modified schemes for monitoring the mean of two (weakly) stationary Gaussian processes with different autocorrelation functions. Most of these stochastic order relations are established by using the fact that the joint distribution of any finite collection of X_t's from a stationary Gaussian process $\{X_t\}$ is a multivariate normal distribution.

In these latter investigations, some emphasis is given to ARL based interpretations of the results because the ARL is by far the most popular of the RL related performance measures and has been extensively used to describe the likely performance of a control scheme. However, we should have in mind that confronting two ARLs essentially means comparing unidimensional and possibly misleading snapshots of the performances of the two schemes and ignores detailed information about the probabilistic behaviour of the *RL*s. On the other hand, establishing stochastic order relations, in the line of work pioneered by W. Schmid, provides a qualitative and a rather more objective assessment of the impact of serial correlation in the performance of quality control schemes we are dealing with than the comparisons based entirely on *ARL*s.

We focus in this latter aspect and special attention is given to *Shewhart* (*CUSUM* and *EWMA*) residual schemes for the mean of stationary autoregressive models of order 1 and 2 (of order 1), described in the following sections and in Morais and Pacheco (2001c).

7.1 Shewhart residual schemes

One of the two following approaches is usually adopted to build control schemes for the mean of autocorrelated data. In the first approach, the original data is plotted in a traditional scheme (*Shewhart*, *CUSUM*, *EWMA*, etc.), however, with readjusted control limits to account for the autocorrelation; the resulting monitoring tool is the so-called *modified scheme* (Vasilopoulos and Stamboulis (1978), Schmid (1995, 1997a) and Zhang (1998)).

The second approach also makes use of a traditional scheme, but the residuals of a time-series model are plotted instead of the original data; this sort of scheme is termed a *special-cause control scheme* (Alwan and Roberts (1988) and Wardell, Moskowitz and Plante (1992, 1994)) or, more commonly, a *residual scheme* (Runger, Willemain and Prabhu (1995) and Zhang (1997)).[1] The rationale behind residual schemes is that the residuals are independent in case the time-series model is valid, thus they meet one of the standard assumptions of traditional schemes, which facilitates the evaluation of *RL* related measures.

There are a few points in favour of residual schemes. According to the comparison studies in Schmid (1995) and Schmid (1997b), and as mentioned by Kramer and Schmid (2000) residual schemes tend to be better than modified schemes (in the *ARL* sense) in the detection of shifts in the mean of a stationary Gaussian autoregressive of order 1 model, when the autoregressive parameter is negative. In addition, the critical values required to implement the residuals schemes do not depend on the underlying in-control process, as those needed by modified schemes. Finally, the *ARL* of modified schemes are usually obtained using simulations (see Schmid (1995, 1997a, 1997b)) and can only have closed expressions for very special cases like exchangeable normal variables (Schmid (1995)).

Let $\{X_t\}$ be a Gaussian autoregressive model of order 2 ($AR(2)$) satisfying the equation

$$(1 - \phi_1 B - \phi_2 B^2)(X_t - \mu) = a_t \tag{7.1}$$

where: μ represents the process mean; $\{a_t\}$ corresponds to Gaussian white noise with variance σ_a^2; and B represents the backshift operator defined as $BX_t = X_{t-1}$ and $B^j X_t = X_{t-j}$. Also, recall that $\{X_t\}$ is a stationary process if $\phi_1 + \phi_2 < 1$, $\phi_2 - \phi_1 < 1$ and

[1] According to Knoth and Schmid (2001), Harrison and Davies (1964) were the first authors to use control charts for residuals.

$-1 < \phi_2 < 1$ (Box, Jenkins and Reinsel (1994, p. 60)). Needless to say that we are dealing with a stationary autoregressive model of order 1 ($AR(1)$) in case $\phi_2 = 0$ and $-1 < \phi_1 = \phi < 1$.

The residual scheme proposed by Zhang (1997) for $AR(2)$ data makes use of the residuals

$$e_t = (1 - \phi_1 B - \phi_2 B^2)(X_t - \mu_0) \tag{7.2}$$

where μ_0 is the nominal value of the process mean, $E(X_t)$. It is worth noticing that these residuals are summary statistics since they only depend on the nominal value of the process mean. Thus, they do not vary with the true value of the process mean as the residuals used in Wardell, Moskowitz and Plante (1994).

The main purpose of the *Shewhart* residual scheme we consider is to detect a single step change in the process mean from the nominal value $E(X_t) = \mu_0$, for $t = \ldots, -1, 0$, to $E(X_t) = \mu_0 + \delta\sigma_x$, for $t = 1, 2, \ldots$, where $\delta \neq 0$ and $\sigma_x^2 = V(X_t)$ remains constant.

In the absence of an assignable cause, the residuals given by (7.2) verify $e_t \sim_{iid} N(0, \sigma_a^2)$, where $\sigma_a^2 = (1 + \phi_2)\{(1 - \phi_2)^2 - \phi_1^2\}/(1 - \phi_2) \times \sigma_x^2$ (Box, Jenkins and Reinsel (1994, p. 62)). As a consequence, the control scheme ought to trigger a signal at time t if

$$e_t < -\xi \times \sigma_a \quad \text{or} \quad e_t > \xi \times \sigma_a \tag{7.3}$$

where ξ is a positive constant usually selected by fixing the ARL in two situations: one being when the quality level is acceptable — i.e., $\delta = 0$ — and one when it is rejectable — that is, δ is equal to some fixed nonzero value. Moreover, the RLs of the *Shewhart* residual schemes — sharing the same ξ — are all matched in-control and have $geo(2[1 - \Phi(\xi)])$ in-control distribution. Thus, these RLs are identically distributed to the one of a standard *Shewhart* X-scheme for *i.i.d.* data, with control limits $\mu_0 \pm \xi\sigma_x$, as mentioned by Wardell, Moskowitz and Plante (1994).

In the out-of-control situation, we still get independent residuals with variance σ_a^2. However: $E(e_1) = \delta\sigma_x$ and $E(e_t) = \delta(1 - \phi)\sigma_x, t = 2, 3, \ldots$, for the $AR(1)$ model; and $E(e_1) = \delta\sigma_x$, $E(e_2) = \delta(1 - \phi_1)\sigma_x$, $E(e_t) = \delta(1 - \phi_1 - \phi_2)\sigma_x, t = 3, 4, \ldots$, for the $AR(2)$ model. All these properties enable us to independently determine the probability that the residual e_t is beyond the control limits and, therefore, to assess the detection speed of the *Shewhart* residual scheme in a straighforward manner, as we shall see below.

7.1.1 AR(1) model

Let $RL(\phi, \delta)$ be the run length of the *Shewhart* residual scheme for an $AR(1)$ process, conditioned on the fact that the autoregressive parameter is equal to ϕ and the mean μ equals μ_0, for $t = \ldots, -1, 0$, and $\mu = \mu_0 + \delta\sigma_x$ ($-\infty < \delta < +\infty$), for $t = 1, 2, \ldots$. Then, the survival function, the hazard rate (or alarm rate) function and the equilibrium rate function of $RL(\phi, \delta)$ are given, respectively, by

$$P[RL(\phi, \delta) > t] = \beta(1; \phi, \delta) \times [\beta(2; \phi, \delta)]^{t-1}, \; t = 1, 2, \ldots \tag{7.4}$$

$$\lambda_{RL(\phi,\delta)}(t) = \begin{cases} 1 - \beta(1; \phi, \delta), & t = 1 \\ 1 - \beta(2; \phi, \delta), & t = 2, 3, \ldots \end{cases} \tag{7.5}$$

$$
r_{RL(\phi,\delta)}(t) \;=\; \begin{cases} 0, & t=1 \\ \frac{1-\beta(1;\phi,\delta)}{\beta(1;\phi,\delta)[1-\beta(2;\phi,\delta)]}, & t=2 \\ \frac{1}{\beta(2;\phi,\delta)}, & t=3,4,\ldots \end{cases} \tag{7.6}
$$

where

$$
1-\beta(t;\phi,\delta) = 1 - \{\Phi\left[\xi - \delta \times f(t;\phi)\right] - \Phi\left[-\xi - \delta \times f(t;\phi)\right]\} \tag{7.7}
$$

with

$$
f(t;\phi) = \frac{|E(e_t)|}{\delta\sigma_a} = \begin{cases} \frac{1}{\sqrt{1-\phi^2}}, & t=1 \\ \sqrt{\frac{1-\phi}{1+\phi}}, & t=2,3,\ldots. \end{cases} \tag{7.8}
$$

Given that no signal has been triggered before time t, $[1-\beta(\min\{t,2\};\phi,\delta)]$ represents the probability that sample t is responsible for triggering a signal, and according to Equation (7.5) it also denotes the alarm rate at sample t. The function $f(t;\phi)$ is what Zhang (1997) called the *detection capability index.*

According to Equation (7.5), the alarm rate takes at most two distinct values: $1-\beta(1;\phi,\delta)$ at sample 1, and $1-\beta(2;\phi,\delta)$ at the following samples. As a consequence, $RL(\phi,\delta)$ has either increasing, constant or decreasing hazard rate function depending on whether the alarm rate at sample 1 is smaller, equal or greater than the alarm rate at the subsequent samples. Therefore the monotone character of $\lambda_{RL(\phi,\delta)}(t)$ allows the comparison of the scheme ability to trigger a signal at sample 1 and at the following samples.

Theorem 7.1 — *$RL(\phi,\delta)$ has an increasing (constant, decreasing) hazard rate function, if and only if $-1<\phi<0$ ($\phi=0, 0<\phi<1$).*

Proof — The detection capability index verifies:

- $f(t;0)=1$
- $f(t;\phi)>1$, if $-1<\phi<0$ (Corollary 1, Zhang (1997, p. 479)).

Moreover, since the sign of the derivative $\frac{d\,\beta(t;\phi,\delta)}{d\,f(t;\phi)}$ equals the sign of $-2\delta\times\sinh[\xi\delta f(t;\phi)]$, we can assert that $\beta(t;\phi,\delta)$ is a decreasing function of the detection capability index, for any $-\infty<\delta<+\infty$. Then, by noting that $f(t;\phi)\geq f(1;\phi)$, for $-1<\phi<0$ and $t=2,3,\ldots$, we immediately conclude that $\lambda_{RL(\phi,\delta)}(t) = 1-\beta(t;\phi,\delta)\geq 1-\beta(1;\phi,\delta) = \lambda_{RL(\phi,\delta)}(1)$. In view of (7.5), we conclude that $\lambda_{RL(\phi,\delta)}$ is an increasing hazard rate function.

Analogously, we get, for $0<\phi<1$ and $t=2,3,\ldots$, $f(t;\phi)\leq f(1;\phi)$. Thus, $\lambda_{RL(\phi,\delta)}(t) = 1-\beta(t;\phi,\delta)\leq 1-\beta(1;\phi,\delta) = \lambda_{RL(\phi,\delta)}(1)$, and $\lambda_{RL(\phi,\delta)}(t)$ is a decreasing hazard rate function for $0<\phi<1$.

For $\phi=0$ we have a constant alarm rate: $\lambda_{RL(0,\delta)}(t) = 1-\beta(1;0,\delta), t=1,2,\ldots$. •

The increasing behaviour of the alarm rate function for $-1<\phi<0$ is consistent with the fact that when the process is negatively correlated "the one step-ahead forecast moves in the opposite direction of the shift" (Wardell, Moskowitz and Plante (1994)). This yields increasingly large residuals and hence increasing hazard rates. This fact is illustrated by the numerical results in Table 7.1. Conversely, for positively correlated data, the alarm rate decreases because residuals tend to get smaller; thus, early detection is more likely to happen.

The next theorem concerns the increasing stochastic behaviour of $RL(\phi, \delta)$ in terms of the autoregressive parameter and can be found in Morais and Pacheco (2000d). Let $RL_{iid}(\delta) = RL(0, \delta)$, so that $RL_{iid}(\delta)$ represents the RL of a standard *Shewhart* X-scheme whose summary statistic and control limits are given by Equations (7.2) and (7.3) with $\phi_1 = \phi_2 = 0$.

Theorem 7.2 — *If* $-1 < \phi \leq 0$ *then*

$$RL(\phi, \delta) \uparrow_{hr} \text{ with } \phi. \tag{7.9}$$

As a consequence $RL(\phi, \delta) \leq_{hr} RL_{iid}(\delta)$, *for* $-1 < \phi < 0$.

Proof — Recall that $\beta(t; \phi, \delta)$ is a decreasing function of $f(t; \phi)$, for $-1 < \phi < 0$, and in this case $f(1; \phi)$ and $f(2; \phi)$ are both decreasing functions of ϕ. The result follows immediately since $1 - \beta(t; \phi, \delta)$ turns out to be a decreasing function of ϕ in the interval $(-1, 0]$. •

Theorem 7.2 can be phrased more clearly by noting that the detection ability (or the alarm rate) of the *Shewhart* residual scheme for any stationary Gaussian $AR(1)$ model decreases with a nonpositive autoregressive parameter ϕ. For the numerical illustration of these properties please refer to Table 7.1.

Remark 7.3 — a) $\beta(1; \phi, \delta)$ decreases with ϕ whereas $\beta(2; \phi, \delta)$ increases with ϕ, for positively autocorrelated $AR(1)$ data. Therefore, it comes as no surprise that we are unable to establish the stochastic monotone behaviour of $RL(\phi, \delta)$, in the usual, the hazard rate or likelihood ratio senses, for $0 < \phi < 1$.

b) It is worth mentioning in passing that $RL(\phi, \delta) \not\leq_{lr} RL_{iid}(\delta)$, for $-1 < \phi < 0$, because the equilibrium rate can be a nonmonotonous function of ϕ (for fixed t). Furthermore, when $r_{RL(\phi,\delta)}(2)$ and $r_{RL(\phi,\delta)}(3)$ are monotone functions of ϕ, $-1 < \phi < 0$, they usually have distinct behaviour and we cannot assert that the odds of a signal at sample t has a monotone behaviour with regard to ϕ. The mentioned properties are apparent in Table 7.1.

c) Theorem 7.2 strenghtens Conclusion 2(a) of Zhang (1997, p. 484):

$$ARL(\phi, \delta) \leq ARL_{iid}(\delta), \text{ for } -1 < \phi < 0. \tag{7.10}$$

d) Theorem 1 in Schmid and Schöne (1997) refers to modified *EWMA* schemes and reads as follows: $RL_{iid}(0) \leq_{st} RL_G(0)$, where $RL_{iid}(0)$ and $RL_G(0)$ refer here to the in-control RLs of *EWMA* modified schemes for *i.i.d.* data and a stationary Gaussian model with nonnegative autocovariance function. $RL(\phi, \delta) \leq_{hr} RL_{iid}(\delta)$, for $-1 < \phi < 0$, is a consequence of result (7.9) and resembles Schmid and Schöne's result; however, the stochastic order

Table 7.1: *ARL*s, alarm rates and equilibrium rates of some *RL*s of the *Shewhart* residual scheme for $AR(1)$ data ($\xi = 3$).

			Alarm rate		Equilibrium rate	
δ	ϕ	*ARL*	$t = 1$	$t \geq 2$	$t = 2$	$t \geq 3$
0	$\phi \in$ (-1,1)	370.4	0.002700	0.002700	1.002707	1.002707
0.10	-0.75	273.2	0.003007	0.003663	0.823405	1.003677
	-0.50	322.2	0.002878	0.003105	0.929754	1.003114
	-0.25	342.1	0.002842	0.002923	0.975109	1.002932
	0	352.9	0.002833	0.002833	1.002842	1.002842
0.50	-0.75	22.1	0.012500	0.046767	0.270673	1.049061
	-0.50	61.2	0.007877	0.016478	0.481861	1.016754
	-0.25	106.6	0.006722	0.009407	0.719371	1.009497
	0	155.2	0.006442	0.006442	1.006484	1.006484
0.75	-0.75	7.3	0.031031	0.154890	0.206759	1.183277
	-0.50	23.1	0.016478	0.044484	0.376628	1.046555
	-0.25	47.7	0.013107	0.021126	0.628686	1.021581
	0	81.2	0.012313	0.012313	1.012466	1.012466
1.00	-0.75	3.6	0.068360	0.361576	0.202933	1.566358
	-0.50	10.5	0.032513	0.102409	0.328153	1.114093
	-0.25	23.3	0.024607	0.043734	0.576857	1.045734
	0	43.9	0.022782	0.022782	1.023313	1.023313
2.00	-0.75	1.5	0.509460	0.989033	1.050088	91.181141
	-0.50	2.1	0.244909	0.678713	0.477880	3.112477
	-0.25	3.4	0.175047	0.337970	0.627837	1.510504
	0	6.3	0.158656	0.158656	1.188574	1.188574
3.00	-0.75	1.1	0.937679	1.000000	15.045836	2.524×10^6
	-0.50	1.3	0.678713	0.985959	0.985959	71.222445
	-0.25	1.6	0.539187	0.808664	0.808664	5.226406
	0	2.0	0.500000	0.500000	2.000000	2.000000

relation we obtained is not only reversed but also stronger, and refers to both in-control and out-of-control situations. These differences probably stem from the fact that *Shewhart* residual schemes have simpler *RL* characteristics than the modified *EWMA* schemes considered by Schmid and Schöne (1997), and that we are dealing with a stricter class of stationary processes (autoregressive models) with simpler autocorrelation behaviour than the class considered by those authors.

e) Theorem 7.2 is also in agreement with the notes in the last paragraph of page 182 of Kramer and Schmid (2000). •

7.1.2 AR(2) model

The survival, hazard rate (or alarm rate) and equilibrium rate functions of $RL(\phi_1,\phi_2,\delta)$, the run length of the *Shewhart* residual scheme for stationary $AR(2)$ data conditioned on ϕ_1, ϕ_2 and δ, are equal, respectively, to

$$P[RL(\phi_1,\phi_2,\delta) > t] = \begin{cases} \beta(1;\phi_1,\phi_2,\delta), & t=1 \\ \beta(1;\phi_1,\phi_2,\delta)\beta(2;\phi_1,\phi_2,\delta) & \\ \times[\beta(3;\phi_1,\phi_2,\delta)]^{t-2}, & t=2,3,\ldots \end{cases} \tag{7.11}$$

$$\lambda_{RL(\phi_1,\phi_2,\delta)}(t) = \begin{cases} 1-\beta(1;\phi_1,\phi_2,\delta), & t=1 \\ 1-\beta(2;\phi_1,\phi_2,\delta), & t=2 \\ 1-\beta(3;\phi_1,\phi_2,\delta), & t=3,4,\ldots \end{cases} \tag{7.12}$$

$$r_{RL(\phi_1,\phi_2,\delta)}(t) = \begin{cases} 0, & t=1 \\ \frac{1-\beta(t-1;\phi_1,\phi_2,\delta)}{\beta(t-1;\phi_1,\phi_2,\delta)[1-\beta(t;\phi_1,\phi_2,\delta)]}, & t=2,3 \\ \frac{1}{\beta(3;\phi_1,\phi_2,\delta)}, & t=4,5,\ldots \end{cases} \tag{7.13}$$

where

$$\beta(t;\phi_1,\phi_2,\delta) = \Phi\left[k-\delta f(t;\phi_1,\phi_2)\right] - \Phi\left[-k-\delta f(t;\phi_1,\phi_2)\right],\ t=1,2,\ldots \tag{7.14}$$

with the detection capability index given by

$$f(t;\phi_1,\phi_2) = \begin{cases} \sqrt{\frac{1-\phi_2}{1+\phi_2}\times\frac{1}{(1-\phi_2)^2-\phi_1^2}}, & t=1 \\ |1-\phi_1|\times f(1;\phi_1,\phi_2), & t=2 \\ (1-\phi_1-\phi_2)\times f(1;\phi_1,\phi_2), & t=3,4,\ldots, \end{cases} \tag{7.15}$$

following the expressions (6)–(8) and (A1)–(A2) of Zhang (1997, p. 478 and p. 489).

Let us define the following sets in $I\!R^2$:

$$\begin{aligned} A &= \{(\phi_1,\phi_2): \phi_1+\phi_2<1, \phi_2-\phi_1<1, -1<\phi_2<1\} \\ B &= \{(\phi_1,\phi_2): \phi_1<0, \phi_2<0\} \\ C &= \{(\phi_1,\phi_2): \phi_1>0, (\phi_2>0 \text{ or } 2\phi_1+\phi_2>2)\} \end{aligned}$$

$$\begin{aligned} D &= \Big\{(\phi_1,\phi_2): \phi_1 < \left[(1-\phi_2)-\sqrt{(1-\phi_2)(1-\phi_2-2\phi_2^2)}\right]/2 \quad \text{or} \\ &\qquad \phi_1 > \left[(1-\phi_2)+\sqrt{(1-\phi_2)(1-\phi_2-2\phi_2^2)}\right]/2\Big\}; \\ E &= \Big\{(\phi_1,\phi_2): \phi_1+\phi_2-\phi_2^2<0\Big\}. \end{aligned}$$

Recall that A corresponds to the stationarity region for the $AR(2)$ model. The meaning of the remaining sets will be discussed next.

The following result corresponds to the analogue of Theorem 7.1 for $AR(2)$ models and can be also found in Morais and Pacheco (2000d).

Theorem 7.4 — *$RL(\phi_1,\phi_2,\delta)$ has increasing (constant, decreasing) hazard rate if $(\phi_1,\phi_2)\in A\cap B$, ($(\phi_1,\phi_2)=(0,0)$, $(\phi_1,\phi_2)\in A\cap C$).*

Proof — The alarm rate function $\lambda_{RL(\phi_1,\phi_2,\delta)}(t)$ increases with t if $f(1;\phi_1,\phi_2) < |1-\phi_1| \times f(1;\phi_1,\phi_2) < (1-\phi_1-\phi_2) \times f(1;\phi_1,\phi_2)$. Taking into account that (ϕ_1,ϕ_2) must belong to set A to guarantee the process stationarity, the double inequality $1 < |1-\phi_1| < 1-\phi_1-\phi_2$ holds if

$$\begin{aligned}(\phi_1,\phi_2) &\in \{(\phi_1',\phi_2') \in A : (\phi_1' < 0 \text{ or } \phi_1' > 2), (\phi_2' < 0 \text{ and } 2\phi_1' + \phi_2' < 2)\} \\ &= \{(\phi_1',\phi_2') \in A : \phi_1' < 0, \phi_2' < 0\} = A \cap B.\end{aligned}$$

The inequalities are reversed for the decreasing behaviour and the result also follows by adding the stationarity condition:

$$\begin{aligned}(\phi_1,\phi_2) &\in \{(\phi_1',\phi_2') \in A : (\phi_1' > 0 \text{ and } \phi_1' < 2), (\phi_2' > 0 \text{ or } 2\phi_1' + \phi_2' > 2)\} \\ &= \{(\phi_1',\phi_2') \in A : \phi_1' > 0, (\phi_2' > 0 \text{ or } 2\phi_1' + \phi_2' > 2)\} = A \cap C.\end{aligned}$$

The increasing behaviour of the alarm rate for $\phi_1, \phi_2 < 0$ is illustrated in Table 7.2, for $\delta = 0.10, 1$.

Table 7.2: *ARL*s, alarm rates and equilibrium rates of some *RL*s of the *Shewhart* residual scheme for $AR(2)$ data ($\xi = 3$).

				Alarm rate			Equilibrium rate		
δ	ϕ_1	ϕ_2	*ARL*	$t=1$	$t=2$	$t\geq 3$	$t=2$	$t=3$	$t\geq 4$
0	$(\phi_1,\phi_2)\in A$		370.4	0.002700	0.002700	0.002700	1.002707	1.002707	1.002707
0.10	-0.50	-0.50	284.3	0.002901	0.003156	0.003522	0.921759	0.899073	1.003534
		-0.25	310.1	0.002870	0.003085	0.003227	0.932858	0.959012	1.003237
		0	322.2	0.002878	0.003105	0.003105	0.929754	1.003114	1.003114
	-0.25	-0.50	306.0	0.002883	0.002988	0.003270	0.967858	0.916438	1.003281
		-0.25	329.4	0.002848	0.002933	0.003036	0.974035	0.968682	1.003046
		0	342.1	0.002842	0.002923	0.002923	0.975109	1.002932	1.002932
	0	-0.50	322.2	0.002878	0.002878	0.003105	1.002887	0.929754	1.003114
		-0.25	342.1	0.002842	0.002842	0.002923	1.002850	0.975109	1.002932
		0	352.9	0.002833	0.002833	0.002833	1.002841	1.002841	1.002841
1.00	-0.50	-0.40	5.9	0.033501	0.106052	0.217532	0.326839	0.545359	1.278007
	-0.25		9.6	0.029326	0.053285	0.120942	0.566996	0.465379	1.137582
	0		15.4	0.028158	0.028158	0.070449	1.028974	0.411276	1.075789
	-0.50	0	10.5	0.032513	0.102409	0.102409	0.328153	1.114093	1.114093
	-0.25		23.3	0.024607	0.043734	0.043734	0.576857	1.045734	1.045734
	0		43.9	0.022782	0.022782	0.022782	1.023313	1.023313	1.023313
	-0.50	0.40	4.0	0.152412	0.484359	0.203620	0.371251	4.613153	1.255683
	-0.25		39.6	0.035963	0.066850	0.023892	0.558032	2.998391	1.024477
	0		100.0	0.028158	0.028158	0.009634	1.028974	3.007654	1.009727

Before we proceed, we would like to remind the reader that the detection capability index of a standard *Shewhart* X-scheme is unitary, as observed by Zhang (1997); also, the alarm rate function is constant. Moreover, note that, as in the $AR(1)$ model, the alarm rate, $1-\beta(t;\phi_1,\phi_2,\delta)$, is also an increasing function of the detection capability index. Summing up these facts, we can assert that the alarm rate of the *Shewhart* residual scheme for $AR(2)$ data is larger than the one of the standard *Shewhart* X-scheme in case $f(t;\phi_1,\phi_2)>1, t=1,2,3,\ldots$. Hence, it is very important to identify the sets where these inequalities hold, as we do in the next lemma that corresponds to Theorem 1 of Zhang (1997).

Lemma 7.5 — *For fixed t, the detection capability index satisfies*

$$f(t;\phi_1,\phi_2,\delta)>1 \Leftrightarrow (\phi_1,\phi_2)\in \begin{cases} A, & t=1 \\ A\cap D, & t=2 \\ A\cap E, & t=3,4,\ldots \end{cases} \tag{7.16}$$

Let $RL_{iid}(\delta)=RL(0,0,\delta)$. Then an analogue of Theorem 7.2 holds for the stationary $AR(2)$ model.

Theorem 7.6 — *For $(\phi_1,\phi_2)\in A\cap D\cap E$,*

$$RL(\phi_1,\phi_2,\delta)\leq_{hr} RL_{iid}(\delta). \tag{7.17}$$

Furthermore, for fixed ϕ_2 $(-1<\phi_2<1)$,

$$RL(\phi_1,\phi_2,\delta)\uparrow_{hr} \text{ with } \phi_1 \text{ over the interval } (\phi_2-1,0]. \tag{7.18}$$

Proof — First, note that in the stationarity region A the two following conditions hold: $1-\phi_1-\phi_2>0$ and $1-\phi_2+\phi_1>0$. As the sign of $\frac{\partial f(1;\phi_1,\phi_2)}{\partial\phi_1}$ equals the one of $\phi_1[(1-\phi_2)^2-\phi_1^2]$, $\frac{\partial f(1;\phi_1,\phi_2)}{\partial\phi_1}$ is nonpositive if $\phi_1\leq 0$. Moreover, since

$$\frac{\partial f(2;\phi_1,\phi_2)}{\partial\phi_1}=-f(1;\phi_1,\phi_2)+(1-\phi_1)\frac{\partial f(1;\phi_1,\phi_2)}{\partial\phi_1}, \tag{7.19}$$

and

$$\frac{\partial f(3;\phi_1,\phi_2)}{\partial\phi_1}=-f(1;\phi_1,\phi_2)+(1-\phi_1-\phi_2)\frac{\partial f(1;\phi_1,\phi_2)}{\partial\phi_1}, \tag{7.20}$$

a sufficient condition for $\frac{\partial f(t;\phi_1,\phi_2)}{\partial\phi_1}\leq 0, t=1,2,3,\ldots$ is to have the pair of parameters (ϕ_1,ϕ_2) belonging to the set $\{(\phi_1,\phi_2)\in A:\phi_1\leq 0\}=\{(\phi_1,\phi_2):-1<\phi_2<1,\phi_2-1<\phi_1\leq 0\}$. In this last set, the alarm rate function decreases with ϕ_1 since $\lambda_{RL(\phi_1,\phi_2,\delta)}(t)$ is an increasing function of $f(t;\phi_1,\phi_2)$. •

The stochastic increasing behaviour in the hazard rate sense of $RL(\phi_1,\phi_2,\delta)$ with respect to ϕ_1 is apparent in Table 7.2, for $\delta=1$. Although the ARL increases with ϕ_2, Table 7.2 illustrates the nonmonotonous behaviour of the alarm rate in terms of ϕ_2 (for $\delta=0.10$) — thus the ARL gives a misleading idea of the scheme performance.

Remark 7.7 — a) Corollary 2 of Zhang (1997, p. 481) can be read as follows: $ARL(\phi_1,\phi_2,\delta)\leq ARL_{iid}(\delta)$, $-1<\phi_1<-0.25$. Equation (7.17) clearly strengthens this corollary.

b) Once again, we cannot compare the RL of the *Shewhart* residual scheme and the RL of the standard X-scheme for uncorrelated data in the likelihood ratio sense. Actually, for the same reasons pointed out for the $AR(1)$ model, $RL(\phi_1,\phi_2,\delta)\not\leq_{lr} RL_{iid}(\delta)$, as we can see from the numerical results regarding the equilibrium rate function in Table 7.2. •

7.2 CUSUM and EWMA residual schemes

The improvement of the sensitivity of Shewhart residual schemes to small and moderate shifts through the adoption of *CUSUM* and *EWMA* residuals schemes has been also addressed in the SPC literature. For example, Harris and Ross (1991), Runger, Willemain and Prabhu (1995) and Lu and Reynolds Jr. (1999a) discuss the application of the *CUSUM* and the *EWMA* techniques and conclude that the resulting residual schemes have superior performance than their Shewhart counterparts, in *ARL* terms.

In this section we assess the impact of the serial autocorrelation in the performance of upper one-sided *CUSUM* and *EWMA* schemes for residuals of a stationary $AR(1)$ model, by means of stochastic ordering. The main purpose of the upper one-sided *CUSUM* and *EWMA* residual schemes, described in the next two subsections, is the detection of upward shifts in the process mean — from μ_0 to $\mu_0 + \delta\sigma_x$, where $\delta > 0$.

7.2.1 CUSUM residual schemes

The detection of upward shifts is done by the upper one-sided *CUSUM* scheme which uses the following summary statistic

$$V_t = \begin{cases} v, & t = 0 \\ \max\{0,\ V_{t-1} + (e_t - k \times \sigma_a)\}, & t = 1, 2, \ldots \end{cases} \tag{7.21}$$

and lower and upper control limits

$$LCL_C = 0 \text{ and } UCL_C = h \times \sigma_a. \tag{7.22}$$

v denotes the initial value given to the upper one-sided *CUSUM* statistic. Let $v = LCL_C + \beta(UCL_C - LCL_C)$ for some $\beta \in [0, 1)$. If $\beta \in (0, 1)$ (or $\beta = 0$) a $\beta \times 100\%$ head start (or no head start) has been given to the chart.

By virtue of the fact that the reference value and the upper control limit are both multiples of σ_a, the *RL*s of the *CUSUM* residual schemes are identically distributed in the absence of an assignable cause ($\delta = 0$) for all values of the parameter ϕ. Thus, all these *CUSUM* residual schemes have matched in-control performance, as well as the Shewhart residual schemes considered in the previous section.

Let $RL_C^{\lfloor\beta(x^+ +1)\rfloor}(\phi, \delta; x)$ be the Markov approximation for the run length of the upper one-sided *CUSUM* residual scheme for stationary $AR(1)$ data with an $\beta \times 100\%$ head start, based on an absorbing Markov chain with discrete state space $\{0, 1, \ldots, x^+ + 1\}$ and absorbing state $(x^+ + 1)$. Note that a $\beta \times 100\%$ head start corresponds to the initial state $\lfloor\beta(x^+ + 1)\rfloor$ in the Markov approximation.

As mentioned earlier, the Markov approach starts with the division of the decision interval $[LCL, UCL)$ in $(x^+ + 1)$ sub-intervals, $[e_i, e_{i+1})$, with equal length $\Delta = (UCL - LCL)/(x^+ + 1)$. These sub-intervals are then associated with the $(x^+ + 1)$ transient states of an absorbing Markov chain, say $\{S_t(\phi, \delta; x^+), t = 0, 1, \ldots\}$, with discrete state space $\{0, 1, \ldots, x^+ + 1\}$ and absorbing state $(x^+ + 1)$.

Since, in the presence of a shift in the process mean, the residuals e_t have their expected value equal to $\delta\sigma_x$ (for $t = 1$) and $\delta(1 - \phi)\sigma_x$ (for $t = 2, 3, \ldots$), the transitions between the transient states of $\{S_t(\delta, \phi; x^+), t = 0, 1, \ldots\}$ are governed by the substochastic matrices $\mathbf{Q}(\phi, \delta; x^+)$, for the first transition, and $\mathbf{Q}(\phi, \delta(1-\phi); x^+)$, for the subsequent transitions. The matrices $\mathbf{Q}$ have the following generic form

$$\mathbf{Q}(\phi, \delta'; x^+) = [q_{i\ j}(\phi, \delta'; x^+)]_{i,j=0}^{x^+} = \sum_{l=0}^{j} q_{i\ l}(\phi, \delta'; x^+) - \sum_{l=0}^{j-1} q_{i\ l}(\phi, \delta'; x^+). \tag{7.23}$$

where

$$\sum_{l=0}^{j} q_{i\ l}(\phi, \delta'; x^+) = \Phi\left(k + \frac{h \times [(j+1) - (i+1/2)]}{x+1} - \frac{\delta'}{\sqrt{1-\phi^2}}\right), \tag{7.24}$$

for $i, j = 0, \ldots, x^+$, and $\sum_{l=0}^{-1} q_{i\ l}(\phi, \delta'; x^+) = 0$, for $i = 0, \ldots, x^+$.

On account of the first transition, $RL_C^{\lfloor\beta(x^+ +1)\rfloor}(\phi, \delta; x^+)$ is associated to a time-non-homogeneous Markov chain. In this setting, the approximating RL is related to a first passage time: $T_C^{\underline{\alpha}(\phi,\delta;x^+)}(\phi, \delta(1-\phi); x^+)$, that takes values in the set $I\!N_0$ and concerns a time-homogeneous absorbing Markov chain whose random initial state equals the discretized version of V_1 and whose transitons between transient states are governed by the substochastic matrix $\mathbf{Q}(\phi, \delta(1-\phi); x^+)$. Thus,

$$P[RL_C^{\lfloor\beta(x^+ +1)\rfloor}(\phi, \delta; x^+) = m] = P[T_C^{\underline{\alpha}(\phi,\delta;x^+)}(\phi, \delta(1-\phi); x^+) = m-1] \tag{7.25}$$

for $m = 1, 2, \ldots$.

$T_C^{\underline{\alpha}(\phi,\delta;x^+)}(\phi, \delta(1-\phi); x^+)$ has a discrete phase-type distribution with parameters

$$(\underline{\alpha}(\phi, \delta; x^+), \mathbf{Q}(\phi, \delta(1-\phi); x^+)) \tag{7.26}$$

where

$$\underline{\alpha}^\top(\phi, \delta; x^+) = \underline{\mathbf{e}}_{\lfloor\beta(x^+ +1)\rfloor}^\top \times \mathbf{Q}(\phi, \delta; x^+). \tag{7.27}$$

Considering $\underline{\mathbf{1}}$ a $(x^+ + 1)-$dimensional vector of ones,

$$(\underline{\alpha}^\top(\phi, \delta; x^+), 1 - \underline{\alpha}^\top(\phi, \delta; x^+) \times \underline{\mathbf{1}}) \tag{7.28}$$

corresponds to the probability vector of the discretized version of V_1, i.e., the initial state of the absorbing Markov chain concerning $T_C^{\underline{\alpha}(\phi,\delta;x^+)}(\phi, \delta(1-\phi); x^+)$.

As far as its survival function is concerned, $RL_C^{[\beta(x+1)]}(\phi, \delta; x^+)$ is defined as follows:

$$P[RL_C^{\lfloor\beta(x^+ +1)\rfloor}(0, \delta; x^+) > m] = \underline{\mathbf{e}}_{\lfloor\beta(x^+ +1)\rfloor}^\top \times [\mathbf{Q}(0, \delta; x^+)]^m \times \underline{\mathbf{1}}, m = 1, 2, \ldots, \tag{7.29}$$

for *i.i.d.* data; and

$$\begin{aligned} P[RL_C^{\lfloor\beta(x^+ +1)\rfloor}(\phi, \delta; x^+) > m] &= \\ &= \begin{cases} \underline{\alpha}^\top(\phi, \delta; x^+) \times \underline{\mathbf{1}}, & m = 1 \\ \underline{\alpha}^\top(\phi, \delta; x^+) \times [\mathbf{Q}(\phi, \delta(1-\phi); x^+)]^{m-1} \times \underline{\mathbf{1}}, & m = 2, \ldots, \end{cases} \end{aligned} \tag{7.30}$$

for $AR(1)$ data. Moreover, the Markov approximation to the ARL, $ARL_C^{\lfloor\beta(x^+ +1)\rfloor}(\phi, \delta; x^+)$, equals:

$$\underline{\mathbf{e}}_{\lfloor\beta(x^+ +1)\rfloor}^\top \times [\mathbf{I} - \mathbf{Q}(0, \delta; x^+)]^{-1} \times \underline{\mathbf{1}} \tag{7.31}$$

for *i.i.d.* data; and

$$1 + \underline{\alpha}^\top(\phi, \delta; x^+) \times [\mathbf{I} - \mathbf{Q}(\phi, \delta(1-\phi); x^+)]^{-1} \times \underline{\mathbf{1}} \tag{7.32}$$

otherwise. Also the alarm rate function is obviously given by

$$\lambda_{RL_C^{\lfloor\beta(x^++1)\rfloor}(\phi,\delta;x^+)}(m) = 1 - \frac{P[RL_C^{\lfloor\beta(x^++1)\rfloor}(\phi, \delta; x^+) > m]}{P[RL_C^{\lfloor\beta(x^++1)\rfloor}(\phi, \delta; x^+) > m-1]} \tag{7.33}$$

for $m = 1, 2, \ldots$.

In Subsection 7.1.1 we were able to obtain simple expressions for the *RL* related measures of the *Shewhart* residual schemes and — through the behaviour of the detection capability index — give straight answers for questions concerning the assessment of the impact of a change in the autogressive parameter ϕ on the performance of a control scheme. So far we have been able to assess the stochastic behaviour of $RL_C^{\lfloor\beta(x^++1)\rfloor}$, with regard to ϕ, in the usual sense.

Furthermore, since $RL_C^{\lfloor\beta(x^++1)\rfloor}(\phi, \delta; x^+)$ converges in law to the exact *RL* of this scheme, $RL_C^{\beta}(\phi, \delta)$, as $x^+ \to \infty$, the next theorem still holds for the exact run length $RL_C^{\beta}(\phi, \delta)$.

Theorem 7.8 — *If* $-1 < \phi \leq 0$ *then, for fixed* k *and* h,

$$RL_C^{\lfloor\beta(x^++1)\rfloor}(\phi, \delta; x^+) \uparrow_{st} \quad \text{with} \quad \phi. \tag{7.34}$$

Namely, $RL_C^{\lfloor\beta(x^++1)\rfloor}(\phi, \delta; x^+) \leq_{st} RL_C^{\lfloor\beta(x^++1)\rfloor}(0, \delta; x^+)$, *for* $-1 < \phi < 0$.

Proof — Let $\{S_t(\phi, \delta; x^+), t = 0, 1, \ldots\}$ be the absorbing chain used to obtain the approximation of the run length $RL_C^{\lfloor\beta(x^++1)\rfloor}(\phi, \delta; x)$. Analogously, $\{U_t(\phi, \delta(1-\phi); x^+), t = 1, 2, \ldots\}$ denotes the absorbing Markov chain whose absorption time is represented by $T^{\underline{\alpha}(\phi,\delta;x^+)}(\phi, \delta(1-\phi); x^+)$.

The initial state of this last Markov chain, $U_1(\phi, \delta(1-\phi); x^+)$, has probability vector $(\underline{\alpha}^\top(\phi, \delta; x^+), 1 - \underline{\alpha}^\top(\phi, \delta; x^+) \times \underline{\mathbf{1}})$ where, as mentioned earlier, $\underline{\alpha}^\top(\phi, \delta; x^+) = \underline{\mathbf{e}}^\top_{\lfloor\beta(x^++1)\rfloor} \times \mathbf{Q}(\phi, \delta; x^+)$. In addition, the state transitions are governed by the stochastic matrix

$$\begin{aligned} \mathbf{P}(\phi, \delta(1-\phi); x^+) &= [p_{i\ j}(\phi, \delta(1-\phi); x^+)]_{i,j=0}^{x+1} \\ &= \begin{bmatrix} \mathbf{Q}(\phi, \delta(1-\phi); x^+) & [\mathbf{I} - \mathbf{Q}(\phi, \delta(1-\phi); x^+)] \times \underline{\mathbf{1}} \\ \underline{\mathbf{0}}^\top & 1 \end{bmatrix}. \end{aligned} \tag{7.35}$$

In view of (7.24), $a_{i\ j}(\phi, \delta; x^+) = \sum_{l=0}^{j} p_{i\ l}(\phi, \delta; x^+)$ increases with $\phi \in (-1, 0]$, the initial state $U_1(\phi, \delta(1-\phi); x^+)$ stochastically decreases with ϕ:

$$U_1(\phi, \delta(1-\phi); x^+) \geq_{st} U_1(\phi', \delta(1-\phi'); x^+), \text{ for } -1 < \phi \leq \phi' \leq 0. \tag{7.36}$$

Another consequence (7.24) is:

$$\sum_{l=0}^{j} p_{i\ l}(\phi, \delta(1-\phi); x^+) \leq \sum_{l=0}^{j} p_{i\ l}(\phi', \delta(1-\phi'); x^+), \text{ for } -1 < \phi \leq \phi' \leq 0; \tag{7.37}$$

i.e., the probability transition matrices $\mathbf{P}(\phi, \delta(1-\phi); x^+)$ and $\mathbf{P}(\phi', \delta(1-\phi'); x^+)$ can be ordered in the usual sense:

$$\mathbf{P}(\phi, \delta(1-\phi); x^+) \geq_{st} \mathbf{P}(\phi', \delta(1-\phi'); x^+), -1 < \phi \leq \phi' \leq 0. \tag{7.38}$$

Then, from Proposition 3.7, we conclude that the Markov chains $\{U_t(\phi, \delta(1-\phi); x^+), t = 1, 2, \ldots\}$ and $\{U_t(\phi', \delta(1-\phi')), t = 1, 2, \ldots\}$ can be also ordered in the usual sense:

$$\{U_t(\phi, \delta(1-\phi); x^+), t = 1, 2, \ldots\} \geq_{st} \{U_t(\phi', \delta(1-\phi'); x^+), t = 1, 2, \ldots\}, \tag{7.39}$$

for $-1 < \phi \leq \phi' \leq 0$. Therefore, according to Collorary 3.13,

$$P[T^{\underline{\alpha}(\phi,\delta;x^+)}(\phi, \delta(1-\phi); x^+) > m-1] \leq P[T^{\underline{\alpha}(\phi',\delta;x^+)}(\phi', \delta(1-\phi'); x^+) > m-1], \tag{7.40}$$

for all $m = 1, 2, \ldots$ and $-1 < \phi \leq \phi' \leq 0$. That is,

$$RL^{\lfloor\beta(x^++1)\rfloor}(\phi, \delta; x^+) \leq_{st} RL^{\lfloor\beta(x^++1)\rfloor}(\phi', \delta; x^+), -1 < \phi \leq \phi' \leq 0. \tag{7.41}$$

This concludes the proof. •

Theorem 7.8 allows us to assert that, for any stationary Gaussian $AR(1)$ model with nonpositive parameter, the detection speed of the *CUSUM* residual scheme decreases with the autoregressive parameter ϕ.

7.2.2 EWMA residual schemes

Now we provide a brief description of the upper one-sided *EWMA* residual scheme for the stationary $AR(1)$ model.

This residual scheme uses the summary statistic

$$W_t = \begin{cases} w, & t = 0 \\ \max\{0,\ (1-\lambda) \times W_{t-1} + \lambda \times e_t\}, & t = 1, 2, \ldots \end{cases} \tag{7.42}$$

and the lower and upper control limits

$$LCL_E = 0 \text{ and } UCL_E = \gamma\sqrt{\lambda(2-\lambda)^{-1}} \times \sigma_a, \tag{7.43}$$

where $\lambda \in (0, 1]$ corresponds to the weight given to the most recent sample residual and w is the initial value of the summary statistic.

To avoid repetition, we merely mention that the Markov approximation to the RL of this residual scheme, $RL_E^{\lfloor\beta(x^++1)\rfloor}(\phi, \delta; x^+)$, has a similar distribution to the one of the upper one-sided *CUSUM* residual scheme previously described — the two substochastic matrices, $\mathbf{Q}(\phi, \delta; x^+)$ and $\mathbf{Q}(\phi, \delta(1-\phi); x^+)$ have to be conveniently replaced. The generic form of the left partial sums of the entries of these matrices is

$$\sum_{l=0}^{j} q_{i\ l}(\phi, \delta'; x^+) = \Phi\left(\frac{\gamma \times [(j+1) - (1-\lambda)(i+1/2)]}{(x^++1)\sqrt{\lambda(2-\lambda)}} - \frac{\delta'}{\sqrt{1-\phi^2}}\right), \tag{7.44}$$

for $i, j = 0, \ldots, x^+$, and once more $\sum_{l=0}^{-1} q_{i\ l}(\phi, \delta'; x^+) = 0$, for $i = 0, \ldots, x^+$.

In addition, $RL_E^{\lfloor\beta(x^++1)\rfloor}(\phi, \delta; x)$ and the exact RL of this *EWMA* residual scheme, $RL_E^{\beta}(\phi, \delta)$, also have a stochastic increasing behaviour in terms of ϕ $(-1 < \phi \leq 0)$.

Theorem 7.9 — *If* $-1 < \phi \leq 0$ *then, for fixed* λ *and* γ,

$$RL_E^{\lfloor\beta(x^++1)\rfloor}(\phi,\delta;x^+) \uparrow_{st} \text{ with } \phi. \tag{7.45}$$

Thus, $RL_E^{\lfloor\beta(x^++1)\rfloor}(\phi,\delta;x^+) \leq_{st} RL_E^{\lfloor\beta(x^++1)\rfloor}(0,\delta;x^+)$, *for* $-1 < \phi < 0$.

Proof — The proof of this result is similar to the one of Theorem 7.8 since $\sum_{l=0}^{j} q_{i\ l}(\phi,\delta;x^+)$ and $\sum_{l=0}^{j} q_{i\ l}(\phi,\delta(1-\phi);x^+)$ both increase with ϕ. •

7.2.3 Numerical illustrations

Now we proceed into the numerical illustration of the properties stated in Theorems 7.8 and 7.9 and other properties of the upper one-sided *CUSUM* and *EWMA* residual schemes.

The numerical results in Table 7.3 refer to the *ARL* and alarm rate function of an upper one-sided *CUSUM* residual scheme with no head start, $k = 0.125$, $h = 9.9609$ and $x^+ + 1 = 30$ transient states, yielding to an in-control *ARL* of approximately 370.409 samples. Similarly, Table 7.4 comprises the same performance measures for an upper one-sided *EWMA* residual scheme with $\lambda = 0.05$, $\gamma = 2.4432$ and, once again, with 30 transient states; with this constelation of parameters we got an in-control *ARL* of 370.382 samples.

We note that our *ARL* values for the upper one-sided *CUSUM* residual scheme differ from those obtained by Runger, Willemain and Prabhu (1995) not only because of the constelation of parameters but also for other reasons: these authors considered the discretization proposed by Brook and Evans (1972), whereas we followed the one in Morais and Pacheco (2001a); the magnitude of the upward shift equals δ in the former paper instead of $\delta\sigma_x$; and, in addition, σ_a was taken equal to 1.

The first apparent feature from both tables is the magnitude of the alarm rate values for $t = 1, 2$: both schemes often lead to values that up to six decimal places are equal to 0.000000, which are then ommited. This fact happens with stronger emphasis in the *CUSUM* case and leads, in particular, to the omission of all the alarm rate values referring to $t = 1$ from Tables 7.3 and 7.4. This is essentially due to the well known initial inertia of *CUSUM* and *EWMA* summary statistics which causes the skewness to the right of the *RL* of these schemes.

These two tables also suggest that the alarm rate of both upper one-sided *CUSUM* and *EWMA* residual schemes increases as we collect more samples for fixed ϕ ($\phi \in (-1, 0]$) and δ and it decreases with ϕ ($\phi \in (-1, 0]$), as in the *Shewhart* case. Thus, the *RL*s of both upper one-sided *CUSUM* and *EWMA* residual schemes seem to also increase with ϕ ($\phi \in (-1, 0]$) in the hazard rate sense.

In addition, the numerical results concerning *ARL*, in both tables, misleadingly suggest that the *RL* stochastically increases with ϕ in the interval $[0, 1)$. However, additional investigations made us realize that the same does not occur to some *RL* percentage points concerning the upper one-sided *EWMA* residual scheme, leading to the conclusion that the increasing behaviour does not hold in the usual sense.

The upper one-sided *EWMA* residual scheme leads to slightly larger alarm rates at the first samples than the matched in-control upper one-sided *CUSUM* residual scheme, thus suggesting a greater ability to detect upward shifts at the first samples.

Table 7.3: *ARL*s and alarm rates of some *RL*s of the *CUSUM* residual scheme for $AR(1)$ data ($k = 0.125, h = 9.9609; x^+ = 29$).*

			Alarm rate				
δ	ϕ	*ARL*	$t=2$	$t=3$	$t=4$	$t=5$	$t=10$
0	$\phi \in$ (-1,1)	370.4				0.000002	0.000194
0.10	-0.75	54.2			0.000002	0.000020	0.002472
	-0.50	86.8			0.000001	0.000009	0.001110
	-0.25	116.1			0.000001	0.000006	0.000737
	0	144.5				0.000005	0.000560
0.50	-0.75	9.3		0.000049	0.002737	0.023931	0.341026
	-0.50	14.2		0.000003	0.000193	0.002132	0.096229
	-0.25	19.4		0.000001	0.000046	0.000527	0.035666
	0	25.7			0.000018	0.000199	0.015614
0.75	-0.75	6.3	0.000001	0.001769	0.053872	0.238243	0.736166
	-0.50	9.4		0.000057	0.002889	0.023974	0.328881
	-0.25	12.5		0.000009	0.000484	0.004711	0.142409
	0	16.3		0.000003	0.000137	0.001382	0.062660
1.00	-0.75	4.9	0.000016	0.026329	0.321621	0.666664	0.934758
	-0.50	7.1		0.000699	0.024186	0.130327	0.605919
	-0.25	9.3		0.000081	0.003563	0.027027	0.328620
	0	12.0		0.000018	0.000840	0.007180	0.163284
2.00	-0.75	2.9	0.111184	0.973210	0.999664	0.999964	0.999997
	-0.50	3.8	0.001292	0.294935	0.837998	0.956693	0.993862
	-0.25	4.8	0.000070	0.045220	0.377360	0.684913	0.925994
	0	5.9	0.000010	0.008185	0.119386	0.344764	0.755253
	0.75	13.1	0.000005	0.000587	0.005833	0.020255	0.122137
	0.90	18.8	0.000211	0.003474	0.012649	0.024945	0.065966
3.00	-0.75	2.0	0.956633	1.000000	1.000000	1.000000	1.000000
	-0.50	2.9	0.164324	0.979785	0.999641	0.999951	0.999995
	-0.25	3.3	0.015067	0.645757	0.960742	0.990412	0.998560
	0	4.0	0.002161	0.248735	0.736428	0.896450	0.976054

*The empty cells in the table correspond to alarm rate values that up to six decimal places are equal to 0.000000.

Replacing the *Shewhart* residual scheme described in the previous section for any of the two upper one-sided *CUSUM* and *EWMA* residual schemes leads to smaller *ARL* values for small and moderate values of δ. Besides this well known feature of the *CUSUM* and *EWMA* schemes, we got smaller alarm rate values for $t = 1, 2$ and significantly larger values for $t \geq 3$ with small or moderate values of δ ($\delta = 0.10, 0.50, 0.75, 1.00$). These results would suggest an occasional early detection of upward shifts with the *Shewhart* residual schemes and a more persistent early detection with the upper one-sided *CUSUM*

Table 7.4: *ARL*s and alarm rates of some *RL*s of the *EWMA* residual scheme for $AR(1)$ data ($\lambda = 0.05, h = 2.4432; x^+ = 29$).*

δ	ϕ	ARL	Alarm rate $t=2$	$t=3$	$t=4$	$t=5$	$t=10$
0	$\phi \in$ (-1,1)	370.4		0.000002	0.000017	0.000070	0.000944
0.10	-0.75	57.2		0.000010	0.000119	0.000530	0.007756
	-0.50	95.1		0.000005	0.000064	0.000277	0.004012
	-0.25	127.7		0.000004	0.000047	0.000202	0.002855
	0	158.1		0.000003	0.000039	0.000164	0.002266
0.50	-0.75	7.8	0.000021	0.003512	0.036154	0.112973	0.391210
	-0.50	12.4	0.000003	0.000430	0.005117	0.020052	0.141798
	-0.25	17.6	0.000001	0.000143	0.001685	0.006865	0.064591
	0	24.3	0.000001	0.000070	0.000784	0.003161	0.033304
0.75	-0.75	5.3	0.000385	0.043672	0.252497	0.472803	0.751732
	-0.50	7.9	0.000026	0.003856	0.036830	0.111089	0.378395
	-0.25	10.8	0.000007	0.000951	0.009941	0.034961	0.191760
	0	14.5	0.000003	0.000367	0.003775	0.013827	0.099256
1.00	-0.75	4.1	0.004340	0.233316	0.650779	0.823683	0.935195
	-0.50	5.9	0.000207	0.022930	0.151943	0.327040	0.632614
	-0.25	7.9	0.000042	0.004887	0.041574	0.117537	0.376109
	0	10.3	0.000015	0.001613	0.014442	0.045754	0.211409
2.00	-0.75	2.4	0.629697	0.998391	0.999941	0.999985	0.999996
	-0.50	3.2	0.068988	0.734780	0.949054	0.979868	0.993286
	-0.25	3.9	0.011018	0.303690	0.679317	0.823076	0.925811
	0	4.8	0.002924	0.109529	0.368147	0.552753	0.763822
	0.75	11.4	0.001573	0.015601	0.042530	0.070282	0.139912
	0.90	17.8	0.017660	0.040531	0.054688	0.062002	0.067155
3.00	-0.75	2.0	0.999493	1.000000	1.000000	1.000000	1.000000
	-0.50	2.3	0.713931	0.998649	0.999926	0.999979	0.999994
	-0.25	2.8	0.259578	0.925537	0.988458	0.995323	0.998320
	0	3.3	0.084119	0.665812	0.891416	0.942356	0.974353

*The empty cells in the table correspond to alarm rate values that up to six decimal places are equal to 0.000000.

and *EWMA* residual schemes, for small and moderate shifts.

Additional numerical investigations with an upper one-sided *Shewhart* residual scheme with an in-control *ARL* of 370.4 samples lead to the same conclusions.

Special attention has been given to the impact of serial autocorrelation in the alarm rate function of a few residual schemes for $AR(1)$ and $AR(2)$ data. Sufficient conditions have been established to guarantee that the run length stochastically increases (in the hazard rate sense, in the Shewhart case; and in the usual sense, in the *CUSUM* and *EWMA* cases) with the autoregressive parameter(s).

The monotonicity of the alarm rate, in terms of the autoregressive parameter(s), has practical importance if the quality engineers design the control scheme following the recommendations of Margavio *et al.* (1995); that is, by choosing the control limits according to a desired in-control ARL and a desired pattern of false alarm rate.

As a final remark, we would like to add that stochastic ordering provides insight into how residual schemes work in practice, and we strongly believe that it leads to an effective assessment of the impact of autocorrelation in an objective and informative manner.

Chapter 8

Concluding remarks

There is yet other directions in which we can study the interplay between stochastic ordering and the performance analysis of control schemes, as shown for instance by our current research on

- the comparison between phase-type and geometric *RL*s,

and by the ongoing joint work with W. Schmid concerning

- the stochastic behaviour of the *RL* of modified *EWMA* schemes for the mean of autocorrelated data in the presence of shifts in the process variance.

Both issues are discussed in the next section, followed by another section devoted to a summary of what we firmly believe are the main contributions of this thesis.

We close this concluding chapter with a few recommendations for future work.

8.1 Ongoing research

As seen earlier, the *RL* of Markov-type schemes that use sequentially accumulated information in order to detect out-of-control situations have (or can be approximated) by phase-type distributions. Assessing the improvement due to the replacement of a *Shewhart* scheme by a Markov-type scheme requires two sorts of comparisons between geometric and discrete phase-type *RL*s:

- numerical comparisons which have been extensively discussed in the quality control literature; and

- inherently difficult stochastic confrontations,

as we shall shortly see.

The rationale behind these stochastic results is the discrete analogue of a well known fact: the *exponential* distribution can be stochastically compared to other nonnegative continuous random variables with *NBU*, *IHRA* and *IHR* ageing properties, as stated in Theorem 3.C.5 of Shaked and Shanthikumar (1994, pp. 107–108). This theorem involves transform orders, such as the superadditive ($\leq_{su}$), the star ($\leq_*$) and the convex ($\leq_c$) orders (see, for example, Shaked and Shanthikumar (1994, pp. 105–109)).[1]

Therefore, obtaining similar stochastic order relations between geometric and discrete phase-type *RL*s demands for a discrete version of these three transform orders for positive integer-valued random variables.

If these discrete orderings satisfy $\leq_c \Rightarrow \leq_* \Rightarrow \leq_{su}$ and if, in addition, $\leq_*$ implies the Lorenz order ($\leq_{Lorenz}$) (see the definition of this latter ordering in Shaked and Shanthikumar (1994, p. 59)) then we would conclude that:

- if we take two matched in-control schemes with run lengths $RL^u(\theta_0)$ and $RL(\theta_0) =_{st} geo(1-\pi(\theta_0))$ — i.e., $ARL^u(\theta_0) = 1/[1-\pi(\theta_0)]$ — and $RL^u(\theta_0) \in IHRA$ or $RL^u(\theta_0) \in IHR$ then this phase-type *RL* has a smaller variance than the one with the geometrically distributed *RL*, $RL(\theta_0)$.

As seen in Chapter 5, as the process variance grows, upper one-sided schemes for the process mean of independent data can become either progressively more sensitive or less sensitive to a shift in μ with magnitude δ, depending on the value of δ. Sufficient conditions (involving bounds for δ) for this sort of behaviour were established for the upper one-sided *EWMA* schemes (and also for the upper one-sided *CUSUM*, combined *CUSUM–Shewhart* and combined *EWMA–Shewhart* schemes) using the Markov approach, and correspond to Equations (17) and (18) in Table 5.1.

Schmid, Morais and Pacheco (2001) obtained analogous sufficient conditions — thus, bounds for δ — concerning the upper one-sided modified *EWMA* scheme for the process mean of a stationary Gaussian process. These results are proved using a different approach because the presence of correlation destroys the Markov property of the summary statistic (Yashchin (1993)).[2]

It is worth mentioning that these bounds for δ are often difficult to obtain. However, we prove that, under mild conditions involving the in-control variance of the *EWMA* statistic, these bounds can have very simple expressions: one of the bounds is related to the assymptotic value of such variance and the other one is a trivial function of the range of the decision interval. We also prove that these simpler bounds hold for the stationary $AR(1)$ and $ARMA(1,1)$ models.

8.2 Claim of contributions

As far as contributions are concerned this thesis documents several aspects, described below in order of appearance.

[1] Let X and Exp denote a nonnegative continuous random variable and an exponential random variable (with arbitrary mean), respectively. Then: $X \in NBU \Leftrightarrow X \leq_{su} Exp$; $X \in IHRA \Leftrightarrow X \leq_* Exp$; $X \in IHR \Leftrightarrow X \leq_c Exp$.

[2] Recall that modified schemes plot the original *EWMA* observations and have control limits adjusted to account for the autocorrelation inherent to the data. Moreover, their summary statistic is different from the one used by Morais and Pacheco (2001a) which immediately resets any value below the target mean μ_0 and turns the analysis of the *RL* distribution practically infeasible in the case of autocorrelated data.

1. We provide useful interpretations of several stochastic order relations in a *RL* setting, in Chapter 1.

   The orders $\leq_{ev}$ (expected value) and $\leq_{st}$ (usual sense) can be found (implicitely or explicitely) in the *SPC* literature.

   However, the stochastic orders involving the notions of hazard rate (or alarm rate), reversed hazard rate and likelihood ratio of *RL*s — $\leq_{hr}$, $\leq_{rh}$ and $\leq_{lr}$, respectively — are not as common as that. Indeed, the notion of alarm rate at sample N of the *RL* was seemingly introduced by Margavio *et al.* (1995) but the associated order was not proposed in the paper. In addition, as far as we have investigated the remaining orders have been not used so far by those or other authors in the *SPC* literature.

2. The *RL*s of Markov-type schemes are viewed as random variables with discrete phase-type distributions in Chapter 2, with related stochastic matrices $\mathbf{P}(\theta)$ with special features, adding up to a few properties mentioned by Brook and Evans (1972).

   We sincerely believe that we brought the issue of stochastically monotone matrices into focus in the *SPC* setting. We proved that the stochastic matrices $\mathbf{P}(\theta)$ in the partitioned form (2.53) that usually arise from upper one-sided control schemes (and, analogously, in lower one-sided schemes, as observed by Morais and Pacheco (1998b)) are stochastically monotone in the usual sense. If, in addition, the (discrete) quality characteristic is $IHR/DRHR/DLR$ then $\mathbf{P}(\theta)$ is stochastically monotone in the hazard rate/reversed hazard rate/likelihood ratio senses, respectively.

   By trivial capitalization on results from the first passage times literature we were able to characterize some *RL*s in terms of ageing properties.

3. In Chapter 3, we give some stochastic flavour to the influence of the adoption of a head start in the *RL* distribution: this random variable stochastically decreases as we increase the initial value u of the summary statistic of the control scheme. The implications of giving a scheme a head start are surely guaranteed by the stochastic monotone character of the associated Markov chain, devised in the previous chapter.

   The decreasing behaviour of the *RL* in the hazard rate sense corresponds to the discrete time analogue of a result by Lee and Lynch (1997), whereas the one in the likelihood ratio sense is a direct consequence of the RR_2 character of the probability function of a specific first passage time, established by Karlin (1964).

   We also established a decreasing monotonicity property of the *RL* in terms of any other parameter ρ; however, we have proved so far that it holds in the usual sense. If ρ represents the magnitude of the shift in the parameter being monitored then the stochastic decreasing behaviour of the *RL* paralels with the notion of a sequentially repeated *UMP* test as in *Shewhart*–type schemes.

   For the sake of completeness, we also proved that all these monotonicity results are still valid for the exact *RL* of schemes for continuous data, in case they hold for the Markov approximation of the *RL*. This is ensured by the convergence in law of the Markov approximation to the exact *RL* of schemes for continuous data (which was not proved by Brook and Evans (1972) when they introduced this sort of approximation) and the closure property under the limit operation of all the stochastic order relations involved here ensures that both stochastic monotonicity properties hold for the exact *RL* if they are valid for the approximating *RL*.

4. The validity of the stochastic monotonicity properties, alluded in the previous chapter, is investigated in Chapter 4 for upper one-sided combined *CUSUM–Shewhart* schemes which seem to have not been applied so far to binomial data.

   By virtue of the presence of a few null entries above the main diagonal of the associated stochastic matrix, these schemes are not governed by a stochastically monotone matrix in the hazard rate or likelihood ratio senses. As a consequence, the adoption of a head start does not yield a stochastic decrease in the *RL* in the likelihood sense although it implies an increase in the alarm rate function of the *RL*.

   Besides proving that supplementing an upper one-sided *CUSUM* scheme with a *Shewhart* upper control limit yields a stochastically smaller *RL* we numerically assess the impact of replacing the former scheme (with or without a 50% head start) by the combined scheme.

   In view of the numerical results we fail to see strong reasons for advocating the addition of the *Shewhart* upper control limit to the upper one-sided *CUSUM* scheme, considering the fact that this is followed by a significant decrease in the in-control *ARL* and a smaller decrease in out-of-control *ARL*s. Moreover, the numerical confrontations suggest that a practitioner should add a head start to the upper one-sided *CUSUM* scheme instead of going to the trouble of adopting a combined scheme, unless he/she wants to privilege the detection of very small shifts in the percentage of defectives per sample.

5. Chapter 5 generalizes some results published by Morais and Pacheco (2001a) for the *RL* of upper one-sided *Shewhart* and *EWMA* schemes for the mean μ of a normally distributed quality characteristic — in the presence of shifts in σ — to upper one-sided *CUSUM*, combined *CUSUM–Shewhart* and combined *EWMA–Shewhart* schemes.

   These results refer to an underreaction to a wide range of increases in μ when the process standard deviation increases and gets off-target.

   Bearing in mind the ageing character of normal data we demonstrated that a decreasing monotonicity property holds for the *RL*, in terms of the head start given to any of the four Markov-type upper one-sided schemes mentioned earlier. The strongest results (in the likelihood ratio sense) are not valid for the combined schemes, as reported in Chapter 4.

6. The phenomenon of misleading signals was brought to our attention in 1998 by St. John and Bragg (1991). Since then we proposed the misleading signals of Type III and Type IV, and two related performance measures: the probability of a misleading signal (in Morais and Pacheco (2000a)) and, more recently, the run length to a misleading signal (Morais and Pacheco (2001b)).

The monotonicity properties of the *PMS*s and the stochastic ordering results referring to the *RLMS*s, established in Chapter 6, prove to be vital, e.g., in the assessment of the frequency of such signals when we underestimate changes in σ (*MS* of Type III) and in μ (*MS* of Type IV).

We are convinced that the paper by Reynolds Jr. and Stoumbos (2001) will finally draw the attention of quality control practitioners to the issue of misleading signals, whose ocurrence is certainly a cause of concern in the joint monitoring of the process mean and variance or any other vector of parameters.

7. Finally, in Chapter 7 we address residual schemes for the mean of autocorrelated data and the impact of the autoregressive parameters in the *RL*.

   Results such as the various ageing characters of the *RL* of standard Shewhart residual schemes for the mean of stationary Gaussian $AR(1)$ and $AR(2)$ models provide information on how the alarm rate behaves as we proceed on collecting more samples. Sufficient conditions regarding the autoregressive parameters guarantee an increasing, constant or decreasing behaviour of the alarm rate function. We also prove that the alarm rate function of the associated *RL* can decrease with the first autoregressive parameter in case it is nonpositive.

   The *RL* of upper one-sided *CUSUM* and *EWMA* residual schemes for the mean of $AR(1)$ processes also increases — however, in the usual sense — with the autoregressive parameter also for nonpositive values.

8.3 Recommendations for further work

The limitations of some stochastic monotonicity results concerning phase-type *RL*s were the starting point of many wanderings and naturally suggested further work.

Take for instance the effort towards stronger versions of Theorem 3.12 concerning the monotonicity in the usual sense of the *RL* of Markov-type schemes, in terms of a parameter other than the initial value of the summary statistic. This issue was already addressed in Chapter 3. However, we note in passing that extensive numerical investigations suggest that when the *RL* stochastically decreases in the usual sense with the magnitude of the shift, it also decreases in the likelihood ratio sense. Moreover, the same seems to hold, e.g., for the range of the decision interval with the exception of those schemes combining *CUSUM* or *EWMA* and *Shewhart* schemes probably because the associated transition matrices are not TP_2.

Another possibility of further work, that certainly deserves some consideration, is the assessment of the impact of the autoregressive parameters in the *PMS*s and the stochastic properties of *RLMS*s of Types III and IV of joint residual schemes for the mean and variance of autocorrelated data. Knoth, Schmid and Schöne (1998) and Knoth, Schmid and Schöne (1999) propose several residual and modified joint schemes for μ and σ, assuming that the samples are independent but the observations within each sample follow an $AR(1)$ model, whereas Lu and Reynolds (1999b) consider the joint residual *EWMA* scheme based on individual observations governed by an $AR(1)$ model. We believe that capitalizing on the results therein we shall be able to obtain numerical values for the *PMS*s and *RLMS*s of Types III and IV, and establish some monotonicity results concerning these two performance measures, in particular for residual schemes.

Control schemes are often used to monitor a process for the sole purpose of detecting assignable causes that result in changes in parameters which in turn may result in lower-quality output.

Unlike some authors who argue that consideration of the theoretical properties of the control schemes (the "probabilistic" approach) reduces the usefulness of the techniques (as observed by Woodall (2000)), we believe that the knowledge of the *RL* distribution and its stochastic monotone behaviour in terms of design or model parameters provides decisive insights into how schemes work in practice and helps practitioners better understand the ability of the control scheme to monitor process quality and the way its performance can change or be improved. Moreover, this knowledge provides guidelines to design schemes that have a desired *RL* performance by adjusting design parameters like initial values, control limits, etc.

Stochastic ordering has been largely excluded from the quality control literature despite the fact that it is a vital and powerful tool in the study of the statistical performance of control schemes. This thesis essentially proposes the use of such a tool in quality control, thus, giving a stimulus and contributing to fill the huge gap between these two areas.

We also hope this thesis can make some of the contributions of the leading authors in the *FPT* and stochastic ordering literature more accessible to quality control practicioners.

Appendix A

Individual control schemes for μ and σ

A.1 Summary statistics and control limits

The joint schemes for μ and σ considered in this thesis involve the individual schemes for μ and σ with the acronyms in Table A.1 and Table A.4, respectively.

Table A.1: Individual schemes for μ.

Acronym	Scheme for μ
$S-\mu$	$\bar{X}$
$C-\mu$	*CUSUM*
$CS-\mu$	Combined *CUSUM-Shewhart*
$E-\mu$	*EWMA*
$ES-\mu$	Combined *EWMA-Shewhart*
$S^+-\mu$	Upper one-sided $\bar{X}$
$C^+-\mu$	Upper one-sided *CUSUM*
$CS^+-\mu$	Combined upper one-sided *CUSUM-Shewhart*
$E^+-\mu$	Upper one-sided *EWMA*
$ES^+-\mu$	Combined upper one-sided *EWMA-Shewhart*

Let $(X_{1N},\dots,X_{nN})$ denote the random sample of size n at the sampling period N ($N \in I\!N$). $\bar{X}_N = n^{-1}\sum_{i=1}^{n} X_{iN}$, $S_N^2 = (n-1)^{-1}\sum_{i=1}^{n}(X_{iN}-\bar{X}_N)^2$ and $Z_N = \sqrt{n}\times(\bar{X}_N-\mu_0)/\sigma_0$ represent the sample mean, sample variance and nominal standardized sample mean, respectively. According to this, we define the summary statistics of the individual schemes for μ and σ in Table A.2 and Table A.5 (respectively), preceeded by the corresponding acronym.

Some of the summary statistics of these individual control schemes are trivial ($S-\mu$, $S^+-\mu$, $S^+-\sigma$) or can be found, such as presented here or with a slight variation, in Montgomery and Runger (1994, p. 875) ($C-\mu$), Lucas and Saccucci (1990) ($E-\mu$ and $ES-\mu$), Lucas and Crosier (1982) ($C^+-\mu$), Yashchin (1985) ($CS^+-\mu$), Crowder and

Table A.2: Summary statistics of the individual schemes for μ.

Chart for μ	Summary statistic
$S-\mu$	Z_N
$C-\mu$	$V_{\mu,N} = \begin{cases} v_{\mu,0}, & N=0 \\ V_{\mu,N-1}+Z_N, & N>0 \end{cases}$
$CS-\mu$	$V_{\mu,N}$ and Z_N
$E-\mu$	$W_{\mu,N} = \begin{cases} w_{\mu,0}, & N=0 \\ (1-\lambda_\mu)\times W_{\mu,N-1}+\lambda_\mu \times Z_N, & N>0 \end{cases}$
$ES-\mu$	$W_{\mu,N}$ and Z_N
$S^+-\mu$	$Z_N^+ = \max\{0,\ Z_N\}$
$C^+-\mu$	$V^+_{\mu,N} = \begin{cases} v^+_{\mu,0}, & N=0 \\ \max\left\{0,\ V^+_{\mu,N-1}+(Z_N-k^+_\mu)\right\}, & N>0 \end{cases}$
$CS^+-\mu$	$V^+_{\mu,N}$ and Z^+_N
$E^+-\mu$	$W^+_{\mu,N} = \begin{cases} w^+_{\mu,0}, & N=0 \\ \max\left\{0,\ (1-\lambda^+_\mu)\times W^+_{\mu,N-1}+\lambda^+_\mu \times Z_N\right\}, & N>0 \end{cases}$
$ES^+-\mu$	$W^+_{\mu,N}$ and Z^+_N

Table A.3: Control limits of the individual schemes for μ.

Scheme for μ	Control limits
$S-\mu$	$C_{S-\mu} = [-\xi_\mu,\ \xi_\mu)$
$C-\mu$	$C_{C-\mu} = [-h_\mu,\ h_\mu)$
$CS-\mu$	$C_{C-\mu}$ and $C_{S-\mu}$
$E-\mu$	$C_{E-\mu} = \left[-\gamma_\mu \times \sqrt{\lambda_\mu/(2-\lambda_\mu)},\ \gamma_\mu \times \sqrt{\lambda_\mu/(2-\lambda_\mu)}\right)$
$ES-\mu$	$C_{E-\mu}$ and $C_{S-\mu}$
$S^+-\mu$	$C_{S^+-\mu} = [0,\ \xi^+_\mu)$
$C^+-\mu$	$C_{C^+-\mu} = [0,\ h^+_\mu)$
$CS^+-\mu$	$C_{C^+-\mu}$ and $C_{S^+-\mu}$
$E^+-\mu$	$C_{E^+-\mu} = \left[0,\ \gamma^+_\mu \times \sqrt{\lambda^+_\mu/(2-\lambda^+_\mu)}\right)$
$ES^+-\mu$	$C_{E^+-\mu}$ and $C_{S^+-\mu}$

Hamilton (1992) ($E^+-\sigma$), Gan(1995) ($C^+-\sigma$), or are natural extensions of existing ones ($CS-\mu$, $E^+-\mu$, $ES^+-\mu$, $CS^+-\sigma$, $ES^+-\sigma$).

Note that, with the exception of the charts $S-\mu$, $S^+-\mu$ and $S^+-\sigma$, an initial value has to be considered for the summary statistic of the standard schemes for μ ($C-\mu$,

Table A.4: Individual schemes for σ.

Initials	Scheme for σ
$S^+ - \sigma$	Upper one-sided S^2
$C^+ - \sigma$	Upper one-sided *CUSUM*
$CS^+ - \sigma$	Combined upper one-sided *CUSUM-Shewhart*
$E^+ - \sigma$	Upper one-sided *EWMA*
$ES^+ - \sigma$	Combined upper one-sided *EWMA-Shewhart*

Table A.5: Summary statistics of the individual schemes for σ.

Scheme for σ	Summary statistic
$S^+ - \sigma$	$S_N^{2+} = \max\left\{\sigma_0^2, S_N^2\right\}$
$C^+ - \sigma$	$V_{\sigma,N}^+ = \begin{cases} v_{\sigma,0}^+, \; N = 0 \\ \max\left\{0, \; V_{\sigma,N-1}^+ + [ln(S_N^2) - k_\sigma^+]\right\}, \; N > 0 \end{cases}$
$CS^+ - \sigma$	$V_{\sigma,N}^+$ and S_N^{2+}
$E^+ - \sigma$	$W_{\sigma,N}^+ = \begin{cases} w_{\sigma,0}^+, \; N = 0 \\ \max\left\{\ln(\sigma_0^2), \; (1-\lambda_\sigma^+) \times W_{\sigma,N-1}^+ + \lambda_\sigma^+ \times \ln(S_N^2)\right\}, \; N > 0 \end{cases}$
$ES^+ - \sigma$	$W_{\sigma,N}^+$ and S_N^{2+}

Table A.6: Control limits of the individual schemes for σ.

Scheme for σ	Control limits
$S^+ - \sigma$	$C_{S^+-\sigma} = \left[\sigma_0^2, \; \xi_\sigma^+ \times \sigma_0^2/(n-1)\right)$
$C^+ - \sigma$	$C_{C^+-\sigma} = [0, \; h_\sigma^+)$
$CS^+ - \sigma$	$C_{C^+-\sigma}$ and $C_{S^+-\sigma}$
$E^+ - \sigma$	$C_{E^+-\sigma} = \left[\ln(\sigma_0^2), \; \ln(\sigma_0^2) + \gamma_\sigma^+ \times \sqrt{\psi'\left(\frac{n-1}{2}\right) \times \lambda_\sigma^+/(2-\lambda_\sigma^+)}\right)$
$ES^+ - \sigma$	$C_{E^+-\sigma}$ and $C_{S^+-\sigma}$

$CS-\mu$, $E-\mu$, $ES-\mu$) and the individual upper one-sided schemes for μ and σ ($C^+-\mu$, $CS^+-\mu$, $E^+-\mu$, $ES^+-\mu$, and $C^+-\sigma$, $CS^+-\sigma$, $E^+-\sigma$, $ES^+-\sigma$).

Let $(LCL+UCL)/2+\alpha\times(UCL-LCL)/2 \;\; (LCL+\alpha\times(UCL-LCL))$ be the initial value of the summary statistic of the standard schemes for μ (upper one-sided schemes for μ and σ), with $\alpha \in (-1,1) \;\; (\alpha \in [0,1))$. If $\alpha \in (-1,0)\cup(0,1)$ $(\alpha \in (0,1))$ an $\alpha \times 100\%$ head start (*HS*) has been given to the standard schemes for μ (to the individual upper one-sided schemes for μ and σ); no head start has been adopted, otherwise.

The adoption of a head start may speed up the detection of shifts by the control scheme at start-up and also after a restart following a (possibly ineffective) control action (see Lucas (1982), Lucas and Crosier (1982), Yashchin (1985) and Lucas and Saccucci (1990)).

The control limits of the individual charts for μ and σ are in Table A.3 and Table A.6, respectively. It is worth adding that the control limits of the individual *EWMA* charts for μ and σ are all specified in terms of the exact asymptotic standard deviation of $W_{\mu,N}$ and $W_{\sigma,N} = (1-\lambda_\sigma) \times W_{\sigma,N-1} + \lambda_\sigma \times \ln(S_N^2)$ (see Lucas and Saccucci (1990) and Crowder and Hamilton (1992), and recall that these latter authors used an infinite series expansion to approximate the trigamma function $\psi'(.)$).

A.2 Run length distribution

Conditioned on the values of δ and θ, where $\delta = \sqrt{n}(\mu - \mu_0)/\sigma_0$ and $\theta = \sigma/\sigma_0$, the *RL*s of the *Shewhart* type charts $S-\mu$, $S^+-\mu$ and $S^+-\sigma$ are geometric random variables with survival functions given in Table A.8 and Table A.9.

As for schemes $C-\mu$, $E-\mu$, $CS-\mu$ and $ES-\mu$ ($C^+-\mu$, $E^+-\mu$, $CS^+-\mu$, $ES^+-\mu$, $C^+-\sigma$, $E^+-\sigma$, $CS^+-\sigma$ and $ES^+-\sigma$) we have to divide the decision interval $[LCL, UCL)$ in $2x+1$ ($x+1$) sub-intervals with equal range Δ, $[e_i, e_{i+1})$, $i=-x,\ldots,x$ ($i=0,\ldots,x$) where x is a positive integer and $e_i = LCL + (i+x) \times \Delta$ ($e_i = LCL + i \times \Delta$), $i = -x,\ldots,x$ ($i=0,\ldots,x$). These sub-intervals are associated to the $2x+1$ ($x+1$) transient states of a Markov chain with one absorbing state.

The transitions between the transient states are governed by a sub-stochastic matrix $\mathbf{Q} = [p_{i\,j}]$, where $p_{ij} = a_{i\,j} - a_{i\,j-1}$ with the left partial sums $a_{i\,j}$ defined in Table A.7.

It is important to notice that, when dealing with combined schemes, a_{ij} depends on improper distribution functions as suggested by Yashchin (1985). In this case they are:

$$\Phi^*(y) = \begin{cases} \Phi(-\xi_\mu), \ y < -\xi_\mu \\ \Phi(y), \ -\xi_\mu \le y < \xi_\mu \\ \Phi(\xi_\mu), \ y \ge \xi_\mu \end{cases} \tag{A.1}$$

$$\Phi^+(y) = \begin{cases} \Phi(y), \ y < \xi_\mu^+ \\ \Phi(\xi_\mu^+), \ y \ge \xi_\mu^+ \end{cases} \tag{A.2}$$

$$F^+_{\chi^2_{n-1}}(y) = \begin{cases} F_{\chi^2_{n-1}}(y), \ y < \xi_\sigma^+ \\ F_{\chi^2_{n-1}}(\xi_\sigma^+), \ y \ge \xi_\sigma^+. \end{cases} \tag{A.3}$$

Table A.7: Left partial sums of the sub-stochastic matrices **Q** of the individual schemes for μ and σ (Markov approach).

Scheme for μ	Left partial sums
$C-\mu$	$a_{\mu,\ i\ j}(\delta,\theta;x_\mu)=\Phi\left(\frac{1}{\theta}\times\left\{\frac{2h_\mu\times[(j+1+x_\mu)-(i+1/2+x_\mu)]}{2x_\mu+1}-\delta\right\}\right),$ $i=-x_\mu,\dots,0,\dots,x_\mu,\ j=-x_\mu-1,\dots,0,\dots,x_\mu$
$CS-\mu$	$a_{\mu,\ i\ j}(\delta,\theta;x_\mu)=\Phi^*\left(\frac{1}{\theta}\times\left\{\frac{2h_\mu\times[(j+1+x_\mu)-(i+1/2+x_\mu)]}{2x_\mu+1}-\delta\right\}\right),$ $i=-x_\mu,\dots,0,\dots,x_\mu,\ j=-x_\mu-1,\dots,0,\dots,x_\mu$
$E-\mu$	$a_{\mu,\ i\ j}(\delta,\theta;x_\mu)=\Phi\left(\frac{1}{\theta}\times\left\{-\frac{\gamma_\mu}{\sqrt{\lambda_\mu^{-1}(2-\lambda_\mu)}}+\frac{2\gamma_\mu\times[(j+1+x_\mu)-(1-\lambda_\mu)(i+1/2+x_\mu)]}{(2x_\mu+1)\sqrt{\lambda_\mu(2-\lambda_\mu)}}-\delta\right\}\right),$ $i=-x_\mu,\dots,0,\dots,x_\mu,\ j=-x_\mu-1,\dots,0,\dots,x_\mu$
$ES-\mu$	$a_{\mu,\ i\ j}(\delta,\theta;x_\mu)=\Phi^*\left(\frac{1}{\theta}\times\left\{-\frac{\gamma_\mu}{\sqrt{\lambda_\mu^{-1}(2-\lambda_\mu)}}+\frac{2\gamma_\mu\times[(j+1+x_\mu)-(1-\lambda_\mu)(i+1/2+x_\mu)]}{(2x_\mu+1)\sqrt{\lambda_\mu(2-\lambda_\mu)}}-\delta\right\}\right),$ $i=-x_\mu,\dots,0,\dots,x_\mu,\ j=-x_\mu-1,\dots,0,\dots,x_\mu$
$C^+-\mu$	$a_{\mu,\ i\ j}(\delta,\theta;x_\mu^+)=\Phi\left(\frac{1}{\theta}\times\left\{k_\mu^+ + \frac{h_\mu^+\times[(j+1)-(i+1/2)]}{x_\mu^++1}-\delta\right\}\right),\ i,j=0,\dots,x_\mu^+$ $a_{\mu,\ i\ -1}(\delta,\theta;x_\mu^+)=0,\ i=0,\dots,x_\mu^+$
$CS^+-\mu$	$a_{\mu,\ i\ j}(\delta,\theta;x_\mu^+)=\Phi^+\left(\frac{1}{\theta}\times\left\{k_\mu^+ + \frac{h_\mu^+\times[(j+1)-(i+1/2)]}{x_\mu^++1}-\delta\right\}\right),\ i,j=0,\dots,x_\mu^+$ $a_{\mu,\ i\ -1}(\delta,\theta;x_\mu^+)=0,\ i=0,\dots,x_\mu^+$
$E^+-\mu$	$a_{\mu,\ i\ j}(\delta,\theta;x_\mu^+)=\Phi\left(\frac{1}{\theta}\times\left\{\frac{\gamma_\mu^+\times[(j+1)-(1-\lambda_\mu^+)(i+1/2)]}{(x_\mu^++1)\sqrt{\lambda_\mu^+(2-\lambda_\mu^+)}}-\delta\right\}\right),\ i,j=0,\dots,x_\mu^+$ $a_{\mu,\ i\ -1}(\delta,\theta;x_\mu^+)=0,\ i=0,\dots,x_\mu^+$
$ES^+-\mu$	$a_{\mu,\ i\ j}(\delta,\theta;x_\mu^+)=\Phi^+\left(\frac{1}{\theta}\times\left\{\frac{\gamma_\mu^+\times[(j+1)-(1-\lambda_\mu^+)(i+1/2)]}{(x_\mu^++1)\sqrt{\lambda_\mu^+(2-\lambda_\mu^+)}}-\delta\right\}\right),\ i,j=0,\dots,x_\mu^+$ $a_{\mu,\ i\ -1}(\delta,\theta;x_\mu^+)=0,\ i=0,\dots,x_\mu^+$

Scheme for σ	Left partial sums
$C^+-\sigma$	$a_{\sigma,\ i\ j}(\theta;x_\sigma^+)=F_{\chi^2_{n-1}}\left(\frac{n-1}{\theta^2\sigma_0^2}\times\exp\left\{k_\sigma^+ + \frac{h_\sigma^+\times[(j+1)-(i+1/2)]}{x_\sigma^++1}\right\}\right),\ i,j=0,\dots,x_\sigma^+$ $a_{\sigma,\ i\ -1}(\theta;x_\sigma^+)=0,\ i=0,\dots,x_\sigma^+$
$CS^+-\sigma$	$a_{\sigma,\ i\ j}(\theta;x_\sigma^+)=F^+_{\chi^2_{n-1}}\left(\frac{n-1}{\theta^2\sigma_0^2}\times\exp\left\{k_\sigma^+ + \frac{h_\sigma^+\times[(j+1)-(i+1/2)]}{x_\sigma^++1}\right\}\right),$ $i,j=0,\dots,x_\sigma^+$ $a_{\sigma,\ i\ -1}(\theta;x_\sigma^+)=0,\ i=0,\dots,x_\sigma^+$
$E^+-\sigma$	$a_{\sigma,\ i\ j}(\theta;x_\sigma^+)=F_{\chi^2_{n-1}}\left(\frac{n-1}{\theta^2}\times\exp\left\{\frac{\gamma_\sigma^+\times\sqrt{\psi\prime[(n-1)/2]}\times[(j+1)-(1-\lambda_\sigma^+)(i+1/2)]}{(x_\sigma^++1)\sqrt{\lambda_\sigma^+(2-\lambda_\sigma^+)}}\right\}\right),$ $i,j=0,\dots,x_\sigma^+$ $a_{\sigma,\ i\ -1}(\theta;x_\sigma^+)=0,\ i=0,\dots,x_\sigma^+$
$ES^+-\sigma$	$a_{\sigma,\ i\ j}(\theta;x_\sigma^+)=F^+_{\chi^2_{n-1}}\left(\frac{n-1}{\theta^2}\times\exp\left\{\frac{\gamma_\sigma^+\times\sqrt{\psi\prime[(n-1)/2]}\times[(j+1)-(1-\lambda_\sigma^+)(i+1/2)]}{(x_\sigma^++1)\sqrt{\lambda_\sigma^+(2-\lambda_\sigma^+)}}\right\}\right),$ $i,j=0,\dots,x_\sigma^+$ $a_{\sigma,\ i\ -1}(\theta;x_\sigma^+)=0,\ i=0,\dots,x_\sigma^+$

The time to absorption of this Markov chain approximates the RL of the corresponding control scheme. Conditional to δ, θ — and to the adoption of an $\alpha \times 100\% HS$, with $\alpha \in (-1,1)$ ($\alpha \in [0,1)$), which corresponds to the initial state $\lfloor \alpha(x+1/2)+1/2 \rfloor$ ($\lfloor \alpha(x_\mu+1) \rfloor$) —, the Markov approximation of the survival function of the time to absorption is given in Table A.8 and Table A.9, where: $\lfloor y \rfloor$ represents the integer part of the real number y; $\underline{\mathbf{e}}_u$ denotes the $(u+1)^{th}$ vector of the orthonormal basis for $I\!R^{2x+1}$ ($I\!R^{x+1}$); and $\underline{\mathbf{1}}$ is a column vector of $2x+1$ ($x+1$) ones.

Table A.8: Survival functions of the RL of the individual schemes for μ.

Scheme for μ	Survival function of RL, for $i = 0, 1, \ldots$
$S-\mu$	$\{\Phi[(\xi_\mu - \delta)/\theta] - \Phi[(-\xi_\mu - \delta)/\theta]\}^i$
$C-\mu$, $CS-\mu$, $E-\mu$, $ES-\mu$	$\underline{e}^\top_{\lfloor \alpha(x_\mu+1/2)+1/2 \rfloor + x_\mu} \times [\mathbf{Q}_\mu(\delta,\theta;x_\mu)]^i \times \underline{\mathbf{1}}$, $\alpha \in (-1,1)$ *
$S^+-\mu$	$\left\{\Phi[(\xi^+_\mu - \delta)/\theta]\right\}^i$
$C^+-\mu$, $CS^+-\mu$, $E^+-\mu$, $ES^+-\mu$	$\underline{e}^\top_{\lfloor \alpha(x^+_\mu+1) \rfloor} \times \left[\mathbf{Q}_\mu(\delta,\theta;x^+_\mu)\right]^i \times \underline{\mathbf{1}}$, $\alpha \in [0,1)$ *

* Markov approximation

Table A.9: Survival functions of the RL of the individual schemes for σ.

Scheme for σ	Survival function of RL, for $i = 0, 1, \ldots$
$S^+-\sigma$	$\left[F_{\chi^2_{n-1}}(\xi^+_\sigma/\theta^2)\right]^i$
$C^+-\sigma$, $CS^+-\sigma$, $E^+-\sigma$, $ES^+-\sigma$	$\underline{e}^\top_{\lfloor \beta(x^+_\sigma+1) \rfloor} \times [\mathbf{Q}_\sigma(\theta;x^+_\sigma)]^i \times \underline{\mathbf{1}}$, $\beta \in [0,1)$ *

* Markov approximation

A.3 Stochastic properties of RL_σ

The RL of the upper one-sided individual schemes for σ have some stochastic monotonicity properties which prove to be fundamental to demonstrate some monotonicity behaviours of the corresponding *PMS*s (and some stochastic monotonicity properties of their run lengths to misleading signals, *RLMS*s).

These results are condensed in Table A.10 and proved in this section. As we did with the upper one-sided schemes for μ (see Chapter 5) we first prove the property for the approximating RL and then we apply the closure property under the limit operation.

Let $v = LCL + \beta \times (UCL - LCL)$ be the initial value of the summary statistic of an upper one-sided control scheme for σ and $RL^\beta(\theta)$ represent the associated exact run length.

Also consider $RL^{\lfloor\beta(x+1)\rfloor}(\theta; x_\sigma^+)$ the Markov approximation of $RL^\beta(\theta)$, based on $(x_\sigma^+ + 1)$ transient states and initial value of the summary statistic $\lfloor\beta(x_\sigma^+ + 1)\rfloor$. Let $f_{\chi^2_{n-1}}(z)$ be the probability density function of the chi-square distribution with $n-1$ degrees of freedom.

Table A.10: Stochastic properties of the exact RL of the individual schemes for σ.

Scheme for σ	Stochastic properties of RL
$S^+ - \sigma$	(1) $RL_{S^+-\sigma}(\theta) \downarrow_{lr}$ with θ
$C^+ - \sigma$	(2) $RL^\beta_{C^+-\sigma}(\theta) \downarrow_{st}$ with β (3) $RL^\beta_{C^+-\sigma}(\theta) \downarrow_{st}$ with θ (4) $RL^\beta_{C^+-\sigma}(\theta; k_\sigma^+) \uparrow_{st}$ with k_σ^+
$CS^+ - \sigma$	(5) $RL^\beta_{CS^+-\sigma}(\theta) \downarrow_{st}$ with β (6) $RL^\beta_{CS^+-\sigma}(\theta) \downarrow_{st}$ with θ (7) $RL^\beta_{CS^+-\sigma}(\theta; k_\sigma^+) \uparrow_{st}$ with k_σ^+
$E^+ - \sigma$	(8) $RL^\beta_{E^+-\sigma}(\theta) \downarrow_{st}$ with β (9) $RL^\beta_{E^+-\sigma}(\theta) \downarrow_{st}$ with θ
$ES^+ - \sigma$	(10) $RL^\beta_{ES^+-\sigma}(\theta) \downarrow_{st}$ with β (11) $RL^\beta_{ES^+-\sigma}(\theta) \downarrow_{st}$ with θ

Proof ((1) in Table A.10: stochastic property of $RL_{S^+-\sigma}(\theta)$) — The parameter of the geometric random variable $RL_{S^+-\sigma}(\theta)$ verifies

$$\frac{\partial}{\partial\theta}\left\{1 - F_{\chi^2_{n-1}}(\xi_\sigma^+/\theta^2)\right\} = \frac{2\xi_\sigma^+}{\theta^3} \times f_{\chi^2_{n-1}}(\xi_\sigma^+/\theta^2) \geq 0. \tag{A.4}$$

Thus, $RL_{S^+-\sigma}(\theta)$ stochastically decreases with θ in the likelihood sense according to Lemma 2.1. •

Proof ((2), (5), (8) and (10) in Table A.10: stochastic monotonicity in the head start of the phase-type RL_σs) — The control schemes $C^+ - \sigma$, $CS^+ - \sigma$, $E^+ - \sigma$, and $ES^+ - \sigma$ have approximating RLs associated to $a_{\sigma,\ i\ -1} = 0,\ i = 0, ..., x_\sigma^+$, and $a_{\sigma,\ ij},\ i, j = 0, ..., x_\sigma^+$, which decrease with i (see Table A.7), whether we use the distribution function $F_{\chi^2_{n-1}}$ or its truncated version $F^+_{\chi^2_{n-1}}$. Thus, the approximating chains are stochastically monotone in the usual sense and we conclude by using Theorem 3.8 that the respective RLs stochastically decrease with the initial state. Hence, by the closure property under the limit operation (Lemma 3.18), the exact RLs, $RL^\beta_{C^+-\sigma}(\theta)$, $RL^\beta_{CS^+-\sigma}(\theta)$, $RL^\beta_{E^+-\sigma}(\theta)$ and $RL^\beta_{ES^+-\sigma}(\theta)$ stochastically decrease in the usual sense with β. •

Proof ((3), (6), (9) and (11) in Table A.10: stochastic monotonicity in θ of the phase-type RL_σs) — The approximating RLs of these schemes for σ are related to stochastically monotone matrices $\mathbf{P}(\theta; x_\sigma^+)$, and, according to Table A.7, to $a_{ij}(\theta) = a_{ij}(\theta; x_\sigma^+)$ which decrease with θ. Thus, by applying Theorem 3.12 and Lemma 3.18, we can assert that $RL^{\beta}_{C^{+}-\sigma}(\theta)$, $RL^{\beta}_{CS^{+}-\sigma}(\theta)$, $RL^{\beta}_{E^{+}-\sigma}(\theta)$ and $RL^{\beta}_{ES^{+}-\sigma}(\theta)$ stochastically decrease with θ. •

Proof ((4) and (7) in Table A.10: stochastic monotonicity in k^{+}_{σ} of the phase-type RL_{σ}s) — The increasing stochastic behaviour in k^{+}_{σ} of $RL^{\beta}_{C^{+}-\sigma}(\theta)$ and $RL^{\beta}_{CS^{+}-\sigma}(\theta)$ follows from the fact that the approximating RLs are associated to a_{ij} which are in any case increasing functions of k^{+}_{σ} (see Table A.7), and the application of Theorem 3.12 and Lemma 3.18. •

Bibliography

Abel, V. (1990). On one-sided combined Shewhart-CUSUM quality control schemes for Poisson counts. *Computational Statistics Quaterly* **6**, 31–39.

Alwan, L.C. (1992). Effects of autocorrelation on control chart performance. *Communications in Statistics — Theory and Methods* **21**, 1025–1049.

Alwan, L.C. and Roberts, H.V. (1988). Time-series modeling for statistical process control. *Journal of Business and Economic Statistics* **6**, 87–95.

Amin, R.W. and Ncube, M.M. (1991). Variable sampling interval combined Shewhart-cumulative score quality control procedure. *Applied Statistics* **40**, 1–12.

Arnold, B. (1987). *Majorization and the Lorenz Order: A Brief Introduction*. Springer-Verlag, Berlin.

Aroian, L.A. and Levene, H. (1950). The effectiveness of quality control procedures. *Journal of the American Statistical Association* **45**, 520–529.

Asmussen, S. and O'Cinneide, C.A. (1999). Matrix-exponential distributions. In *Encyclopedia of Statistical Sciences: Update Volume 3*, 435–440. Kotz, S., Read, C.B. and Banks, D.L. (eds.). John Wiley & Sons, New York.

Assaf, D., Shaked, M. and Shanthikumar, J.G. (1985). First-passage times with PF_r densities. *Journal of Applied Probability* **22**, 185–196.

Bagshaw, M. and Johnson, R.A. (1975). The influence of reference values and estimated variance on the ARL of CUSUM tests. *Journal of the Royal Statistical Society (Series B — Statistical Methodology)* **37**, 413–420.

Barlow, R.E. and Proschan, F. (1975). *Statistical Theory of Reliability and Life Testing: Probability Models*. Holt, Rinehart and Winston, Inc.

Bawa, V.S. (1982). Stochastic dominance: a research bibliography. *Management Science* **28**, 698–712.

Bäuerle, N. and Rolsky, T. (2000). A monotonicity result for the workload in Markov-modulated queues. *Journal of Applied Probability* **35**, 741–747.

Bhattacharya, P.P. and Ephremides, A. (1991). Stochastic monotonicity properties of multiserver queues with impatient customers. *Journal of Applied Probability* **28**, 673–682.

Blacksell, S.D., Gleeson, L.J., Lunt, R.A. and Chamnanpood, C. (1994). Use of combined Shewhart-CUSUM control charts in internal quality control of enzyme-linked immunosorbent assays for the typing of foot and mouth disease virus antigen. *Revue Scientifique et Technique* **13**, 687–699.

Box, G.E.P., Jenkins, G.M. and Reinsel, G.C. (1994). *Time Series Analysis: Forecasting and Control*. Prentice-Hall, Englewood Cliffs, New Jersey.

Böhm, W. and Hackl, P. (1990). Improved bounds for the average run length of control charts based on finite weighted sums. *The Annals of Statistics* **18**, 1895–1899.

Brook, D. and Evans, D.A. (1972). An approach to the probability distribution of CUSUM run length. *Biometrika* **59**, 539–549.

Brown, M. and Chaganty, N.R. (1983). On the first passage time distribution for a class of Markov chains. *The Annals of Probability* **11**, 1000–1008.

Casella, G. and Berger, R.L. (1990). *Statistical Inference.* Duxberry Press, Belmont, California.

Champ, C.W. and Rigdon, S.E. (1991). A comparison of the Markov chain and the integral equation approaches for evaluating the run length distribution of quality control charts. *Communications in Statistics — Simulation and Computation* **20**, 191–204.

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

Chen, G. (1997). The mean and standard deviation of the run length distribution of $\bar{X}$ charts when control limits are estimated. *Statistica Sinica* **7**, 789–798.

Chengalur, I.N., Arnold, J.C. and Reynolds Jr., M.R. (1989). Variable sampling intervals for multiparameter Shewhart charts. *Communications in Statistics — Theory and Methods* **18**, 1769–1792.

Crowder, S.V. (1987a). A simple method for studying run-length distributions of exponentially weighted moving average charts. *Technometrics* **29**, 401–407.

Crowder, S.V. (1987b). Computation of ARL for combined individual measurements and moving range charts. *Journal of Quality Technology* **19**, 98–102.

Crowder, S.V. (1989). Design of exponentially weighted moving average schemes. *Journal of Quality Technology* **21**, 155–162.

Crowder, S.V. and Hamilton, M.D. (1992). An EWMA for monitoring a process standard deviation. *Journal of Quality Technology* **24**, 12–21.

Daley, D.J. (1968). Stochastically monotone Markov chains. *Zeitschrift Wahrscheinlichkeitstheorie werwandte Gebiete* **10**, 305–317.

Dalton, H. (1920). The measurement of the inequality of incomes. *Economic Journal* **30**, 348–361.

Domangue, R. and Patch, S.C. (1991). Some omnibus exponentially weighted moving average statistical process monitoring schemes. *Technometrics* **32**, 299–314.

Dudding, B.P. and Jennett, W.J. (1944). *Quality Control Chart Technique.* General Electric, London.

Durham, S., Lynch, J. and Padgett, W.J. (1990). TP_2–orderings and the IFR property with applications. *Probabability in the Engineering and Informational Sciences* **4**, 73–88.

Gan, F.F. (1989). Combined mean and variance control charts. *ASQC Quality Congress Transactions - Toronto*, 129–139.

Gan, F.F. (1990a). Monitoring observations generated from a binomial distribution using modified exponentially weighted moving average control chart. *Journal of Statistical Computation and Simulation* **37**, 45–60.

Gan, F.F. (1990b). Monitoring Poisson observations using modified Exponentially Weighted Moving Average control charts. *Communications in Statistics — Simulation and Computation* **19**, 103–124.

Gan, F.F. (1991). Computing the percentage points of the run length distribution of an exponentially weighted moving average control chart. *Journal of Quality Technology* **23**, 359–365.

Gan, F.F. (1993). An optimal design of CUSUM control charts for binomial counts. *Journal of Applied Statistics* **20**, 445–460.

Gan, F.F. (1995). Joint monitoring of process mean and variance using exponentially weighted moving average control charts. *Technometrics* **37**, 446–453.

Ghosh, B.K., Reynolds Jr., M.R. and Hui, Y.V. (1981). Shewhart $\bar{X}$ with estimated process variance. *Communications in Statistics — Theory and Methods* **10**, 1797–1822.

Gibbons, R.D. (1999). Use of combined Shewhart-CUSUM control charts for ground water monitoring applications. *Ground Water* **37**, 682–691.

Greenberg, I. (1997). Markov chain approximation methods in a class of level-crossing problems. *Operations Research Letters* **21**, 153–158.

Hardy, G.H., Littlewood, J.E. and Pólya, G. (1929). Some simple inequalities satisfied by convex functions. *Messenger of Mathematics* **58**, 145–152.

Harris, T.J. and Ross, W.H. (1991). Statistical process control procedures for correlated observations. *The Canadian Journal of Chemical Engineering* **69**, 48–57.

Harrison, P.J. and Davies, O.L. (1964). The use of cumulative (CUSUM) techniques for the control of routine forecasts of product demand. *Operations Research* **12**, 325–333.

Hawkins, D.M and Olwell, D.H. (1998). *Cumulative Sum Charts and Charting for Quality Improvement.* Springer-Verlag, New York.

Jensen, D.R. (1984). Invariant ordering and order preservation. In *Inequalities in Statistics and Probability* (Lecture Notes — Monograph Series Vol. 5), 26–34. Tong, Y.L. (ed.). Institute of Mathematical Statistics, Hayward, California.

Jensen, D.R. and Hui, Y.V. (1990a). Properties of R and S^2 charts. I. Dependent statistics. *Journal of Statistical Computation and Simulation* **34**, 168–172.

Jensen, D.R. and Hui, Y.V. (1990b). Properties of R and S^2 charts. II. Nonstationarity. *Journal of Statistical Computation and Simulation* **34**, 172–175.

Johnson, D.G. (1984). Trial by computer — A case study of the use of simple statistical techniques in the detection of a fraud. *Journal of the Operational Research Society* **35**, 811–820.

Johnson, R.A. and Bagshaw, M. (1974). The effect of serial correlation on the performance of CUSUM test. *Technometrics* **16**, 103–112.

Jones, L.A., Champ, C.W. and Rigdon, S.E. (2001). The performance of exponentially weighted moving average charts with estimated parameters. *Technometrics* **43**, 156–167.

Kalmykov, G.I. (1962). On the partial ordering of one-dimensional Markov processes. *Theory of Probability and its Applications* **7**, 456–459.

Kalmykov, G.I. (1969). Semiordering of the probabilities of the first passage time of Markov processes. *Theory of Probability and its Applications* **14**, 704–710.

Kamae, T., Krengel, U. and O'Brien, G.L. (1977). Stochastic inequalities on partially ordered spaces. *The Annals of Probability* **5**, 899–912.

Kaminsky, F.C., Benneyan, J.C., Davis, R.D. and Burke, R.J. (1992). Statistical control charts based on a geometric distribution. *Journal of Quality Technology* **24**, 63–69.

Karlin, S. (1964). Total positivity, absorption probabilities and applications. *Transactions of the American Mathematical Society* **11**, 33–107.

Karlin, S. (1968). *Total Positivity* — Vol. I. Stanford University Press, California.

Karlin, S. and McGregor, J. (1959). A characterization of birth and death processes. *Proceedings of the National Academy of Sciences of the United States of America* **45**, 375–379.

Karlin, S. and Rinott, Y. (1980). Classes of orderings of measures and related correlation inequalities. I. Multivariate totally positive distributions. *Journal of Multivariate Analysis* **10**, 467–498.

Keilson, J. and Kester, A. (1977). Monotone matrices and monotone Markov processes. *Stochastic Processes and their Applications* **5**, 231–241.

Keilson, J. and Kester, A. (1978). Unimodality preservation in Markov chains. *Stochastic Processes and their Applications* **7**, 179–190.

Kijima, M. (1992). Further monotonicity properties of renewal processes. *Advances in Applied Probability* **25**, 575–588.

Kijima, M. (1997). *Markov Processes for Stochastic Modeling.* Chapman and Hall, London.

Kijima, M. (1998). Hazard rate and reversed hazard rate monotonicities in continuous-time Markov chains. *Journal of Applied Probability* **35**, 545–556.

Kirmani, S.N.U.A. and Peddada, S. (1993) Stochastic ordering approach to off-line quality control, *Journal of the Royal Statistical Society (Series C — Applied Statistics)* **42**, 271-281.

Knoth, S., Schmid, W. and Schöne, A. (1998). *Simultaneous Shewhart-type charts for the mean and the variance of a time series.* Arbeitsbericht 107, Lehrstuhl für Quatitative Methoden insbesondere Statistik, Europa–Universität Viadrina, Frankfurt (Oder), Germany.

Knoth, S. and Schmid, W. (1999). *Monitoring the mean and the variance of a stationary process.* Arbeitsbericht 130, Lehrstuhl für Quatitative Methoden insbesondere Statistik, Europa–Universität Viadrina, Frankfurt (Oder), Germany.

Knoth, S. and Schmid, W. (2001). Control charts for time series: a review. In *Proceedings of the VIIth International Workshop on Intelligent Statistical Quality Control* (Session: Statistical Product & Process Control II), University of Waterloo, Canada, September 5–7, 2001.

Kramer, H. and Schmid, W. (2000). The influence of parameter estimation on the ARL of Shewhart type charts for time series. *Statistical Papers* **41**, 173–196.

Kulkarni, V.G. (1995). *Modeling and Analysis of Stochastic Systems.* Chapman and Hall, London.

Lai, T.L. (1974). Control charts based on weighted sums. *The Annals of Statistics* **2**, 134–147.

Lee, S. and Lynch, J. (1997). Total positivity of Markov chains and the failure rate character of some first passage times. *Advances in Applied Probability* **29**, 713–732.

Levy, H. (1992). Stochastic dominance and expected utility: survey and analysis. *Management Science* **38**, 555–593.

Li, H. and Shaked, M. (1995). On the first passage times for Markov processes with monotone convex transition kernels. *Stochastic Processes and their Applications* **58**, 205–216.
Li, H. and Shaked, M. (1997). Ageing first-passage times of Markov processes: a matrix approach. *Journal of Applied Probability* **34**, 1–13.
Lorenz, M.O. (1905). Methods of measuring the concentration of wealth. *Publication of the American Statistical Association* **9**, 209–219.
Loewner, C. (1934). Über monotone Matrixfunktionen. *Mathematische Zeitschrift* **38**, 177–216.
Lu, C.-W. and Reynolds Jr., M.R. (1999a). EWMA control charts for monitoring the mean of autocorrelated processes. *Journal of Quality Technology* **31**, 166–188.
Lu, C.-W. and Reynolds Jr., M.R. (1999b). Control charts for monitoring the mean and variance of autocorrelated processes. *Journal of Quality Technology* **31**, 259–274.
Lucas, J.M. (1982). Combined Shewhart-CUSUM quality control schemes. *Journal of Quality Technology* **14**, 51–59.
Lucas, J.M. (1985). Counted data CUSUMs. *Technometrics* **27**, 129–144.
Lucas, J.M. and Crosier, R.B. (1982). Fast initial response for CUSUM quality-control schemes: give your CUSUM a head start. *Technometrics* **24**, 199–205.
Lucas, J.M. and Saccucci, M.S. (1990). Exponentially weighted moving average control schemes: properties and enhancements. *Technometrics* **32**, 1–12.
Luceño, A. and Puig-Pey, J. (2000). Evaluation of the run-length probability distribution for CUSUM charts: assessing chart performance. *Technometrics* **42**, 411–416.
Maragah, H.D. and Woodall, W.H. (1992). The effect of autocorrelation on the retrospective $\bar{X} - chart$. *Journal of Statistical Simulation and Computation* **40**, 29-42.
Margavio, T.M., Conerly, M.D., Woodall, W.H. and Drake, L.G. (1995). Alarm rates for quality control charts. *Statistics & Probability Letters* **24**, 219-224.
Marshall, A.W. and Olkin, I. (1979). *Inequalities: Theory of Majorization and Its Applications*. Academic Press, New York.
Montgomery, D.C. (1985). *Introduction to Statistical Quality Control*. John Wiley & Sons, New York.
Montgomery, D.C. and Runger, J.C. (1994). *Applied Statistics and Probability for Engineers*. John Wiley & Sons, New York.
Morais, M.J.C. (1995). *Cartas de controlo FSI e VSI para o parâmetro de escala da população Weibull tri-paramétrica*. (FSI and VSI control charts for the scale parameter of a three parameter Weibull distribution.) M.Sc. thesis, IST. Supervised by M.F. Ramalhoto. In Portuguese.
Morais, M.J.C. (1998). Esquemas CEWMA para o controlo simultâneo de μ e de σ^2: uma abordagem markoviana (CEWMA control schemes for the joint monitoring of μ and σ^2: a Markov approach). In *Estatística: A Diversidade na Unidade*, 207–219. Souto de Miranda, M. and Pereira, I. (eds.). Edições Salamandra, Lisboa. In Portuguese.
Morais, M. and Natário, I. (1998). Improving an upper one-sided c-chart. *Communications in Statistics — Theory and Methods* **27**, 353–364.
Morais, M.J.C. and Pacheco, A. (1998a). Ordenação estocástica: um pouco de história e aplicações (On the History and applications of stochastic ordering). In *Estatística: A*

Diversidade na Unidade, 221–235. Souto de Miranda, M. and Pereira, I. (eds.). Edições Salamandra, Lisboa. In Portuguese.

Morais, M.C. and Pacheco, A. (1998b). Two stochastic properties of one-sided exponentially weighted moving average control charts. *Communications in Statistics — Simulation and Computation* **27**, 937–952.

Morais, M.C. and Pacheco, A. (1998c). Comparing first passage times of Markovian processes. In *Proceedings of the 2nd International Symposium on Semi-Markov Models: Theory and Applications* (Session 4: Statistical Estimation II). Janssen, J. and Limnios, N. (eds.).

Morais, M.C. and Pacheco, A. (1999a). Counted data control charts sensitive to sudden increases of the summary statistic. *Revista de Estatística* **1/99**, 25–45.

Morais, M.C. and Pacheco, A. (1999b). Comparação estocástica de tempos de primeira passagem: uma aplicação a controlo de qualidade, fiabilidade e filas de espera (Stochastic comparison of first passage times: an application to quality control, reliability and queues). In *Afirmar a Estatística: Um Desafio para o Século XXI*, 303–314. Paulino, D., Pacheco, A., Pires, A. and Ferreira da Cunha, A. (eds.). Edições SPE, Lisboa. In Portuguese.

Morais, M.C. and Pacheco, A. (2000a). On the performance of combined EWMA schemes for μ and σ: A Markovian approach. *Communications in Statistics — Simulation and Computation* **29**, 153–174.

Morais, M.C. and Pacheco, A. (2000b). *Misleading signals in joint schemes for μ and σ.* Technical Report 17/2000, DM–IST, Lisboa, Portugal. Submitted for publication.

Morais, M.C. and Pacheco, A. (2000c). *Some stochastic properties of upper one-sided $\bar{X}$ and EWMA charts for μ in the presence of shifts in σ.* Technical Report 38/2000, DM–IST, Lisboa, Portugal.

Morais, M.C. and Pacheco, A. (2000d). *Ordenação estocástica na análise de desempenho de esquemas de controlo de qualidade* (Stochastic ordering in the performance analysis of quality control schemes). To appear in the Proceedings of the 8th Annual Meeting of the Portuguese Statistical Society, Peniche, Portugal, October 14–16, 2001. In Portuguese.

Morais, M.C. and Pacheco, A. (2001a). Some stochastic properties of upper one-sided $\bar{X}$ and $EWMA$ charts for μ in the presence of shifts in σ. *Sequential Analysis* **20**(1/2), 1–12.

Morais, M.C. and Pacheco, A. (2001b). Misleading signals em esquemas combinados EWMA para μ e σ (Misleading signals in joint EWMA schemes for μ and σ). In *Um Olhar sobre a Estatística*, 334–348. Oliveira, P. and Athayde, E.(eds.). Edições SPE. In Portuguese.

Morais, M.C. and Pacheco, A. (2001c). Assessing the impact of correlation in the performance of residual schemes: a stochastic ordering approach. In *Proceedings of the VIIth International Workshop on Intelligent Statistical Quality Control* (Session: Statistical Product & Process Control II).

Mosler, K. Scarsini, M. (1993). *Stochastic Orders and Applications: A Classified Bibliography.* Springer-Verlag, Berlin Heidelberg.

Ncube, M.M. and Woodall, W.H. (1984). A combined Shewhart-cumulative score quality control chart. *Journal of the Royal Statistical Society (Series C — Applied Statistics)* **33**, 259–265.

Nelson, L.S. (1997). Supplementary runs tests for *np* control charts. *Journal of Quality Technology* **29**, 225–227.

Neuts, M.F. (1981). *Matrix-Geometric Solutions in Stochastic Models.* Johns Hopkins University Press, Baltimore, MD.

Olwell, D.H. (1997). Managing misconduct: statistical process control applied to sexual harassment. *1997 Proceedings of the Section on Quality and Productivity.* Alexandria, VA: American Statistical Association.

Page, E.S. (1954). Continuous inspection schemes. *Biometrika* **41**, 100–115.

Pearson, E.S. (1967). Some notes on W.A. Shewhart's influence on the application of statistial methods in Great Britain. *Industrial Quality Control*, 81–83.

Prabhu, S.S. and Runger, G.C. (1996). Analysis of a two-dimensional Markov chain. *Communications in Statistics – Simulation and Computation* **25**, 75–79.

Ramalhoto, M.F. and Morais, M. (1995). *Algumas cartas do tipo Shewhart para o parâmetro de localização da população Weibull tri-paramétrica.* (Some simple control charts for the location parameter of the three-parameter Weibull distribution.) Technical Report 19/1995, DM–IST, Lisboa, Portugal.

Ramalhoto, M.F. and Morais, M. (1998). EWMA control charts for the scale parameter of a Weibull control variable with fixed and variable sampling intervals. *Economic Quality Control — Journal and Newsletter for Quality and Reliability* **13**, 23–46.

Ramalhoto, M.F. and Morais, M. (1999). Shewhart control charts for the scale parameter of a Weibull control variable with fixed and variable sampling intervals. *Journal of Applied Statistics* **26**, 129–160.

Reynolds Jr., M.R., Amin, R.W. and Arnold, J.C. (1990). CUSUM charts with variable sampling intervals. *Technometrics* **32**, 371–384.

Reynolds Jr., M.R., Amin, R.W., Arnold, J.C. and Nachlas, Joel A. (1988). $\bar{X}$-charts with variable sampling intervals. *Technometrics* **30**, 181–192.

Reynolds Jr., M.R. and Stoumbos, Z.G. (2001). Monitoring the process mean and variance using individual observations and variable sampling intervals. *Journal of Quality Technology* **33**, 181–205.

Roberts, S.W. (1959). Control charts tests based on geometric moving averages. *Technometrics* **1**, 239–250.

Runger, G.C. and Prabhu, S.S. (1996). A Markov chain model for the multivariate exponentially weighted moving average control chart. *Journal of the American Statistical Association* **91**, 1701–1706.

Runger, G.C., Willemain, T.R. and Prabhu, S. (1995). Average run lengths for CUSUM control charts applied to residuals. *Communications in Statistics — Theory and Methods* **24**, 273–282.

Saniga, E.M. (1989). Economic statistical control chart design with an application to $\bar{X}$ and R charts. *Technometrics* **31**, 313–320.

Schmid, W. (1995). On the run length of a Shewhart chart for correlated data. *Statistical Papers* **36**, 111–130.

Schmid, W. (1997a). On EWMA charts for time series. In: Lenz, H.J., Wilrich, P.-Th. (Eds.), *Frontiers of Statistical Quality Control*, Vol.5, p. 115-137. Physica-Verlag, Heidelberg.

Schmid, W. (1997b). CUSUM control schemes for Gaussian processes. *Statistical Papers* **38**, 191–217.
Schmid, W., Morais, M.C. and Pacheco, A. (2001). On the stochastic behaviour of the RL of control schemes for the mean of correlated data in the presence of shifts in σ. *Working paper.*
Schmid, W. and Schöne, A. (1997). Some properties of the EWMA control chart in the presence of autocorrelation. *The Annals of Statistics* **25**, 1277–1283.
Schöne, A., Schmid, W. and Knoth, S. (1999). On the run length of the EWMA scheme: a monotonicity result for normal variables. *Journal of Statistical Planning and Inference* **79**, 289–297.
Shaked, M. and Li, H. (1997). Aging first-passage times. In *Encyclopedia of Statistical Sciences: Update Volume 1*, 11–20. Kotz, S., Read, C.B. and Banks, D.L. (eds.). John Wiley & Sons, New York.
Shaked, M. and Shanthikumar, J.G. (1987). Multivariate hazard rate and stochastic ordering. *Advances in Applied Probability* **19**, 123–137.
Shaked, M. and Shanthikumar, J.G. (1994). *Stochastic Orders and Their Applications.* Academic Press, London.
Shanthikumar, J.G. (1988). DFR property of first-passage times and its preservation under geometric compounding. *The Annals of Probability* **16**, 397–406.
Shanthikumar, J.G., Yamazaki, G. and Sakasegawa, H. (1991). Characterization of optimal order of services in a tandem queue with blocking. *Operations Research Letters* **10**, 17–22.
Shewhart, W. (1931). *Economic Control of Quality of Manufactured Product.* Van Notrand, Princeton.
Shobe, K.N. (1988). *Variable sampling intervals for control charts using count data.* M.Sc. thesis, Department of Statistics, Virginia Tech, Blacksburg, VA.
St. John, R.C. and Bragg, D.J. (1991). Joint X-bar R charts under shift in mu or sigma. *ASQC Quality Congress Transactions – Milwaukee*, 547–550.
Stoumbos, Z.G., Reynolds, Jr., M., Ryan, T.P. and Woodall, W.H. (2000). The state of statistical process control as we proceed into the 21st century. *Journal of the American Statistical Association* **95**, 992–998.
Stoyan, D. (1983). *Comparison Methods for Queues and Other Stochastic Models.* John Wiley & Sons, Chichester.
Szekli, R. (1995). *Stochastic Ordering and Dependence in Applied Probability.* Springer-Verlag, New York.
Takahashi, T. (1989). Simultaneous control charts based on $(\bar{x}, s)$ and $(\bar{x}, r)$ for multiple decision on process change. *Reports of Statistical Application Research* **36**, 1–20.
Tong, Y.L. (1980). *Probability Inequalities in Multivariate Distributions.* Academic Press, New York.
Tong, Y.L. (1990). *The Multivariate Normal Distribution.* Springer-Verlag, New York.
Van Dobben de Bruyn, C.S. (1968). *Cumulative Sum Tests: Theory and Practice.* Griffin, London.
VanBrackle III, L.N. and Reynolds Jr., M.R. (1997). EWMA and CUSUM control charts in the presence of correlation. *Communications in Statistics — Simulation and Computation*

26, 979–1008.
Vasilopoulos, A.V. and Stamboulis, A.P. (1978). Modification of control chart limits in the presence of data correlation. *Journal of Quality Technology* **10**, 20–30.
Vaughan, T. (1993). Variable sampling interval *np* process control chart. *Communications in Statistics — Theory and Methods* **22**, 147–167.
Wardell, D.G., Moskowitz, H. and Plante, R.D. (1992). Control charts in the presence of data correlation *Management Science* **38**, 1084–1104.
Wardell, D.G., Moskowitz, H. and Plante, R.D. (1994). Run-length distributions of special-cause control charts for correlated processes. *Technometrics* **36**, 3–27.
Wasan, M.T. (1994). *Stochastic Processes and Their First Passage Times.* Queen's Papers in Pure and Applied Mathematics, No. 96. Queen's University, Kingston, Ontario.
Weiler, H. (1952). On the most economical sample size for controlling the mean of a population. *Annals of Mathematical Statistics* **23**, 247–254.
Westgard, J.O., Groth, T., Aronsson, T. and de Verdier, C.-H. (1977). Combined Shewhart-CUSUM control chart for improved quality control in clinical chemistry. *Clinical Chemistry* **23**, 1881–1887.
White, C.H., Keats, J.B. and Stanley, J. (1997). Poisson CUSUM versus c chart for defect data. *Quality Engineering* **9**, 673–679.
Wolfram, S. (1996). *The Mathematica Book* — 3rd edition. Addison–Wesley Publishing Company, Inc.
Woodall, W.H. (1997). Control charts based on attribute data: Bibliography and review. *Journal of Quality Technology* **29**, 172–183.
Woodall, W.H. (2000). Controversies and contradictions in Statistical Process Control (with discussion and response). *Journal of Quality Technology* **32**, 341–378.
Woodall, W.H. and Faltin, F.W. (1996). An overview and perspective on control charting. In *Statistical Applications in Process Control*, 7–20. Keats, J.B. and Montgomery, D.C. (eds.). Marcel Dekker, New York.
Woodall, W.H. and Ncube, M.M. (1985). Multivariate CUSUM quality-control procedures. *Technometrics* **27**, 285–292.
Yao, D.D. and Zheng, S. (1995). Sequential quality control of machining process, *Proc. IEEE 34th Conference on Decison and Control (New Orleans, LA)*, 3128–3133.
Yao, D.D. and Zheng, S. (1999) Coordinated quality control in a two-stage system, *IEEE Transactions on Automatic control* **44**, 1166–1179.
Yashchin, E. (1985). On the analysis and design of CUSUM-Shewhart control schemes. *IBM Journal of Research and Development* **29**, 377–391.
Yashchin, E. (1993). Performance of CUSUM control schemes for serially correlated observations. *Technometrics* **35**, 37–52.
Zhang, N.F. (1997). Detection capability of residual control chart for stationary process data. *Journal of Applied Statistics* **24**, 475–492.
Zhang, N.F. (1998). A statistical control chart for stationary process data. *Technometrics* **40**, 24–38.