Statistical Process Control in Non-Standard Situations

Omslagontwerp: Brechje Vermaat
b.vermaat@studioiets.nl

NUR: 916
ISBN-10: 90-6464-047-5
ISBN-13: 978-90-6464-047-6

Statistical Process Control in Non-Standard Situations

Academisch Proefschrift

ter verkrijging van de graad van doctor
aan de Universiteit van Amsterdam
op gezag van de Rector Magnificus
prof. mr. P.F. van der Heijden
ten overstaan van een door het college voor promoties ingestelde
commissie, in het openbaar te verdedigen in de Aula der Universiteit
op woensdag 6 december 2006, te 14.00 uur

door

Mattheus Bertus Vermaat
geboren te Leeuwarden

Promotiecommissie

Promotor:	Prof.dr. R.J.M.M. Does
Overige leden:	Prof.dr. S. Bisgaard
	Prof.dr.ir. J.G. de Gooijer
	Prof.dr. C.A.J. Klaassen
	Prof.dr. M.R.H. Mandjes
	Dr. J. de Mast
	Prof.dr. W. Schaafsma
	Dr. A. Trip

Faculteit der Natuurwetenschappen, Wiskunde en Informatica.

Es mejor ser loado de los pocos sabios que burlado de los muchos necios.

vrij vertaald: Hecht meer aan de lof van enkele wijzen
dan aan de spot van vele dwazen.

Miguel de Cervantes

IBIS UvA BV
INSTITUUT VOOR BEDRIJFS- EN INDUSTRIËLE STATISTIEK

Dit proefschrift is mede mogelijk gemaakt door een financiële bijdrage van het Instituut voor Bedrijfs- en Industriële Statistiek van de Universiteit van Amsterdam (IBIS UvA).

Preface

As I started my internship at IBIS UvA in 2002, I was at the beginning of my research in control charts in non-standard situations. This research field has many opportunities. It contains the problems which form the basis for this thesis. Most of my time during the last four years was well spent: I delved into a problem, unravelled it and ended up with a paper worthwhile reading. This thesis is based on five of those problems. Nevertheless, months were lost trying to solve problems that could not be solved - or at least I was not able to solve. To some extent, it is all a matter of luck. The challenge is to increase your chances on good luck.

Over the years, I was fortunate enough to have people around me that helped me doing just that, through their useful comments or moral support. First of all I would like to thank Ronald Does. He gave me the opportunity to work on my PhD thesis. He supported and motivated me, especially when I was out of luck. He increased the chance that the problem I worked on had a successful outcome.

Furthermore, I would like to thank the people who co-wrote the papers with me: Søren Bisgaard, Chris Klaassen, Roxana Ion, Frank van der Meulen and the late Ton Steerneman. Discussions with them increased my knowledge of the research field and kept me sharp and focussed. Thanks are also due to the dissertation committee for carefully reading the manuscript and their useful comments for improvement.

My colleagues at IBIS UvA never failed to make me laugh, entertain me with anecdotes and stimulate me through discussions during lunch hours and breaks. Thank you all for that.

Finally, I would like to thank my family and friends. They corrected draft versions of my thesis on English spelling and grammar, but far more importantly they encouraged me from start to finish and cheered me up when I needed it. I am lucky to have them in my life.

Thijs Vermaat
Amsterdam, September 2006

Contents

Preface **vii**

1 Introduction **1**

1.1 Development of the control chart 1

1.2 Control charts and non-normality 5

1.3 Control charts and autocorrelation 9

1.4 Outline of the thesis 13

2 A comparison of Shewhart individual control charts based on normal, non-parametric, and extreme-value theory **15**

2.1 Introduction 15

2.2 Description of the control charts 16

2.2.1 Control chart based on the average of the moving ranges 16

2.2.2 Control chart based on empirical quantiles (bootstrap) 17

2.2.3 Control chart based on kernel estimators 18

2.2.4 Control chart based on extreme-value theory 19

2.3 Simulations 21

2.4 Conclusion 31

Appendix 2A: Convergence of control limits obtained with extreme-value theory . . 32

3 A semi-Bayesian method for Shewhart individual control charts **39**

3.1 Introduction 39

3.2 Description of the control charts 40

3.2.1 The traditional individual control chart based on moving ranges 40

3.2.2 The exact average moving range control chart 41

3.2.3 Empirical quantile control chart 41

3.2.4 A control chart based on a Bernstein approximation 42

3.3 Real life example 43

3.4 Design of the simulation study 45

3.4.1 Simulation design for the Bernstein approximation control chart 46

3.4.2 Discussion of the simulation results 47
3.5 Conclusion 58
Appendix 3A: The Bernstein approximation 59
Appendix 3B: Illustration of the Bernstein density estimation method 61

4 The variance of the EWMA statistic for autoregressive processes 65
4.1 Introduction 65
4.2 Preliminaries on second order autoregressive processes 67
4.3 The variance of the EWMA statistic 68
4.4 Estimators for the variance of the EWMA statistic 70
4.5 Asymptotics of the estimators 72
4.5.1 Asymptotic distributions 72
4.5.2 Asymptotic relative efficiencies 74
4.6 Simulations 77
4.7 An elaborated example 78
4.7.1 Control limits for the exact EWMA control chart 78
4.7.2 Sensitivity analysis of the inflation factor 79
4.7.3 Conclusions from the example 82
4.8 Conclusion 82
Appendix 4A: Derivation of the variance of the EWMA statistic 83
Appendix 4B: Properties of the variance of the EWMA statistic 85

References 87

Samenvatting 95

Curriculum vitae 99

Chapter 1

Introduction

The discipline of Industrial Statistics is a relatively young field of research. It began a century ago with William Gosset, who worked for the Guinness breweries. In 1908 he published his famous paper "The probable error of the mean". In this paper, which Gosset was forced to publish under the pseudonym "Student", he laid down the basis of the familiar t-test. His work would start a revolution in what was then an infant discipline of statistics and it inspired men like Sir Ronald Fisher, Egon Pearson, and Jerzy Neyman to develop much of the now standard machinery for analyzing the results of comparative experiments (hypothesis testing, the analysis of variance, regression analysis, et cetera). See De Mast (2006) for a more elaborate description. Nowadays Industrial Statistics is a full-grown research field. One of the topics that is studied, is Statistical Process Control. This chapter gives a short introduction in the control chart as the most important statistical tool in Statistical Process Control (SPC). Section 1.1 concerns the origins of the control chart. Two problems which show up in the application of the control chart are discussed in Sections 1.2 and 1.3. We finish this chapter with an outline of this thesis.

1.1 Development of the control chart

For as long as people have made products, they have been concerned with quality. The Egyptians already carried out inspections of the building blocks used in their buildings. This assured a solid construction of temples and pyramids. In the Middle Ages craftsmen were organized into guilds. Within these guilds, boys learnt a trade from a master, while serving as his apprentice. Once the apprentice had enough experience, he would become a member of the guild. This was how constant quality of products was maintained.

The industrial revolution at the end of the nineteenth century put the notion of quality in a different perspective. Masters and apprentices became foremen and workers. Mechanization offered huge possibilities for new products, but above all for large volumes and high productivity. Inspection assured that bad products were filtered out. Quality departments were set up

to assure product quality. Nevertheless, quality departments were more engaged in inspection of process output than in improvements in the production processes.

In the 1920s at Western Electric, a company which has become specialized in the telephone market, Walter A. Shewhart introduced the control chart to monitor process output. The first sketch of a control chart appeared in a memorandum from Shewhart in 1924. With the famous book "Economic Control of Quality of Manufactured Product", published in 1931, Shewhart grounded the research of control charts. Duncan (1986) has written a short historical introduction to quality control. Other historical overviews of Statistical Quality Control can be found in Pearson (1973), Mann (1988), and Stratton (1991).

The control chart

Shewhart introduced two kinds of variation, i.e. variation due to assignable causes and due to chance causes. An assignable causes refers to patterns in process measurements that are to be interpreted as a signal that something in the process has changed. All other patterns in the data are to be interpreted as random noise, which operators should ignore. Shewhart developed a practical tool that helps operators discern between the two: the control chart already mentioned. The control chart is an outstanding instrument to facilitate process control. It is a chart indicating the time at which a quality characteristic is measured on the horizontal axis and the value of the quality characteristic on the vertical axis. By charting a quality characteristic as a function of time we can judge whether or not a process is statistically in-control. Figure 1.1 shows an example of a control chart.

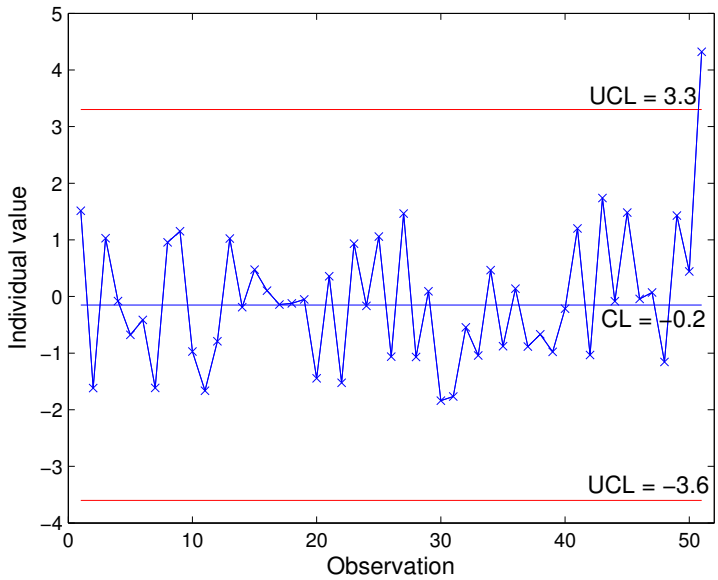

Figure 1.1: An example of a control chart.

This can be described more formally as follows. Shewhart describes a "constant cause

system" – meaning a system of chance causes independent over time – as opposed to a non-constant cause system, a system which changes over time. From time to time, there may be assignable causes which change the constant cause system. A phenomenon that can be predicted, at least within limits associated with a given probability, is said to be controlled. In such a case we say that a quality characteristic is statistically in-control. We call a non-constant cause system, i.e. a system with assignable causes, statistically out-of-control.

If the process is in control and samples of size n are drawn and characterized by some statistic (e.g. the mean, the median, or the maximum) then these values can be regarded as outcomes of random variables which are more or less independent if the duration of time between moments of sampling is sufficiently large. Suppose that they have the same density f_n. If one knows f_n (from past experience) then one can compute

$$P = \int_{\theta_1}^{\theta_2} f_n(x)dx. \tag{1.1}$$

Shewhart proposes to choose the values θ_1 and θ_2 so that the associate value P is *economic*. Shewhart (1931) describes economic as: "Even when no trouble exists, we shall look for trouble $(1-P)N$ times on the average after inspecting N samples of size n. On the other hand, the smaller the probability P the more often in the long run may we expect to catch trouble if it exists. We must try to strike a balance between the advantages to be gained by increasing the value P through reduction in the cost of looking for trouble when it does not exist" (p. 276). Based on experience, Shewhart sets θ_1 and θ_2 equal to $\bar{\mu} \pm 3\sigma$ to obtain an acceptable economic value, where $\bar{\mu}$ is the statistic's mean and σ the statistic's standard deviation of the observations. We get a signal for an assignable cause if an observed value is outside the interval (θ_1, θ_2).

Drawn control limits provide easy checks on the stability of the process. Control charts are usually constructed using 20 to 30 initial samples of about 5 items each, which in general are supposed to result from purely random sampling. A treatment of statistical aspects of control charts in a typical textbook situation was given by Does and Schriever (1992). Much larger data sets (at least 300 observations) are needed if the sample size consists of one item (cf. Quesenberry (1993)). For this situation the individual control chart has been developed (cf. Duncan (1986)). In order to estimate the parameters (μ and σ) of the control limits, Duncan (1986) uses the sample mean and the average of the moving ranges respectively. In the literature other estimators for the control limits of individual measurements are proposed, cf. Cryer and Ryan (1990) and Roes et al. (1993). Up to date books on Statistical Process Control incorporating more advanced methods useful in practice are Wheeler (1995), Quesenberry (1997), Does et al. (1999), and Montgomery (2004).

In this thesis we only consider individual control charts. In Figure 1.1 an individual control chart is drawn, where the Lower Control Limit (LCL) is equal to θ_1, the Central Line (CL) is equal to the mean of the sample, and the Upper Control Limit (UCL) is equal to θ_2.

Woodall and Montgomery (1999) distinguish the retrospective phase (Phase I) and the pro-

cess monitoring phase (Phase II). In Phase I data are collected on the process to check whether or not the process was statistically in-control. If the process is statistically in-control, the obtained control limits can be used in Phase II to check the stability of the process over time. For a discussion about the differences between Phase I and Phase II we refer to De Mast and Roes (2005).

Figure 1.1 illustrates an example in which the control limits are calculated based on the first 50 individual observations. These observations stem from a process which is statistically in-control (Phase I). The last observation (observation denoted as 51) is above the Upper Control Limit (UCL), so the control chart signals for an out-of-control situation, i.e. an assignable cause in Phase II. Even if a process is statistically in-control, a signal for an assignable cause occurs with probability $1 - P$; we call such a signal a false alarm.

Using the Shewhart control chart for individual measurements, the practitioner discovers some drawbacks of this technique. For example, it turns out that, small shifts in the mean of the quality characteristic are hard to detect. Furthermore, distributional aspects and autocorrelation of the observations are issues for discussion. The first issue, the distribution of the process output, is discussed in Section 1.2. Autocorrelation of the process output will be discussed in Section 1.3.

Control charts and persistent causes

Shewhart control charts are effective in detecting outliers and large shifts in the mean. However, a Shewhart chart only uses the information about the process contained in the observation plotted last and ignores any information given by the former observations. This feature makes the Shewhart control chart relatively insensitive to small shifts in the process. Hence, there is a need to make the control chart more sensitive to small shifts in the mean, say in the order of about 1.5σ or less. Three methods for a faster detection of these kinds of assignable causes are as follows.

The first method is introduced by Western Electric Company (1959). A set of decision rules is added to the control chart to make it more sensitive to changes in the cause system. Typical "run rules" attached to the Shewhart control chart are:
- two out of three observations more than two standard deviations from the central line,
- six observations in a row, all increasing or all decreasing, and
- four out of five observations more than one standard deviation from the central line.

These run rules are very effective improvements over the classical Shewhart control chart for the detection of persistent assignable causes.

The second method is due to Page (1954). He proposed a control chart named the CUmulative SUM (CUSUM) control chart. The CUSUM chart incorporates all the information of the observations by plotting the cumulative sums of the deviations of the observations from a target value. If the target value is the process mean this chart is efficient in detecting small shifts in

the process mean. A formal procedure to detect visually small shifts in the mean is introduced by Barnard (1959) with the development of the "V-mask".

The third method is the Exponentially Weighted Moving Average (EWMA) control chart. This method for checking the stability of the mean was introduced by Roberts (1959). The EWMA control chart utilizes all previous observations and attaches weights to the data. These weights are exponentially decreasing as the observations get older, as opposed to the CUSUM control chart where all past observations are equally weighted. The EWMA control chart will be used in the study for autocorrelation in Section 1.3 and Chapter 4.

These three methods – the extra run rules, the CUSUM control chart, and the EWMA control chart – are all developed to make a control chart more sensitive to persistent assignable causes. In the following two sections distributional aspects and autocorrelations of the observations are discussed.

1.2 Control charts and non-normality

In this section we discuss how to deal with observations from a constant cause system where the statistic under study is not normally distributed. This section gives an introduction to Chapters 2 and 3.

Shewhart states that we have to detect assignable causes. Usually, the underlying distribution function is assumed to be normal for the evaluation of the statistical performance. Statistical aspects of individual charts have been studied in Roes et al. (1993), Reynolds and Stoumbos (2001a, 2001b). A number of authors have pointed out that Shewhart charts for subgroup means work well irrespective whether the measurements are normally distributed or not (Schilling and Nelson (1976) and Wheeler (1991)). However, the behavior of the traditional control chart for individual measurements is seriously affected by departures from normality: P in (1.1) is no longer valid where f_n is supposed to be normal and hence the values for θ_1 and θ_2 are no longer *economic* (see e.g. Yourstone and Zimmer (1992), Borror et al. (1999), and Stoumbos and Reynolds (2000)). The normality assumption is risky particularly individual measurements are used. Non-normal data are quite common in SPC applications. In situations with a large number of observations it may be possible to subgroup the data to avoid an individual chart.

In Borror et al. (1999) an EWMA control chart is designed that is robust to non-normality. In a recent paper, Woodall and Montgomery (1999) have stated that: "There would appear to be an increasing role for non-parametric methods, particularly as data availability increases. Abundant data would cause the loss of power associated with non-parametric methods to become less of an issue." If large data sets are available, then an attractive approach to parametric statistical inference would be to use these large data sets to study distributional form.

A motivating example: depth of a groove

To motivate the work presented in Chapters 2 and 3, consider collected data concerning the part of a printer that squirts the ink. The ink is spurted through a groove. A relevant quality characteristic to be measured is the groove's depth. We will study 535 observations of this depth. For reasons of confidentiality, the measurements are multiplied by a constant. In Figure 1.2 the observations are plotted in a time series plot. As indicated in Figure 1.2 the data contain several obvious

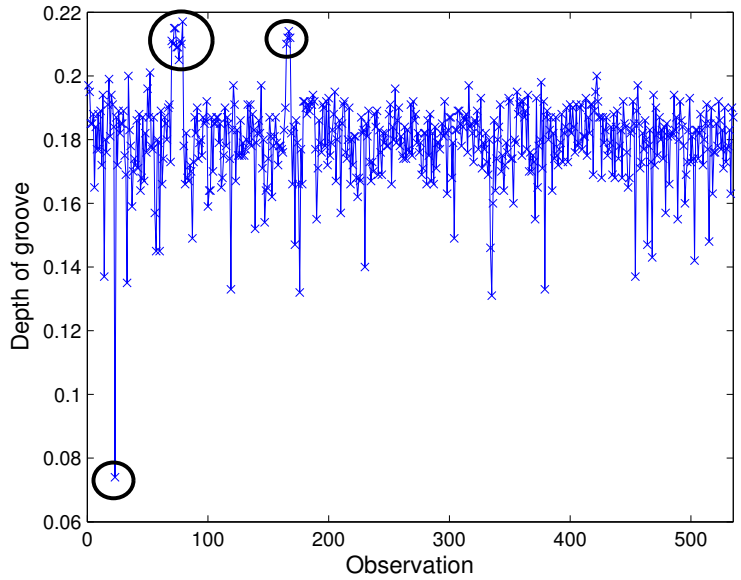

Figure 1.2: Time series plot of the depth of a groove.

outliers. By studying the data set, we found out, that for the measurements 70 up to 79 and 165 up to 168, the grooves where cut by a supplier who uses a different method to cut the grooves. Observation 23 is clearly an outlier, probably caused by contamination or dirt. These problems were detected and eliminated. Deleting these outliers we get a data set of 520 observations. Based on these observations a Shewhart individual control chart is drawn to get the appropriate control limits for Phase II. The Upper Control Limit is defined by

$$\mathrm{UCL} = \mu + 3\sigma$$

and the Lower Control Limit is defined by

$$\mathrm{LCL} = \mu - 3\sigma.$$

Natural choices to estimate μ and σ are

$$\hat{\mu} = \bar{X}_n = \frac{1}{n}\sum_{t=1}^{n} X_t \tag{1.2}$$

and

$$\hat{\sigma}_X = \frac{\overline{\mathrm{MR}}_n}{d_2(2)} = \frac{1}{(n-1)d_2(2)} \sum_{t=2}^{n} |X_t - X_{t-1}| \tag{1.3}$$

respectively, where $d_2(2) = 2/\sqrt{\pi}$ to obtain an estimator for σ_X that is unbiased under normality. The fluctuations of the process displayed in Figure 1.3 are common for these kinds of processes. So these observations are from a process which

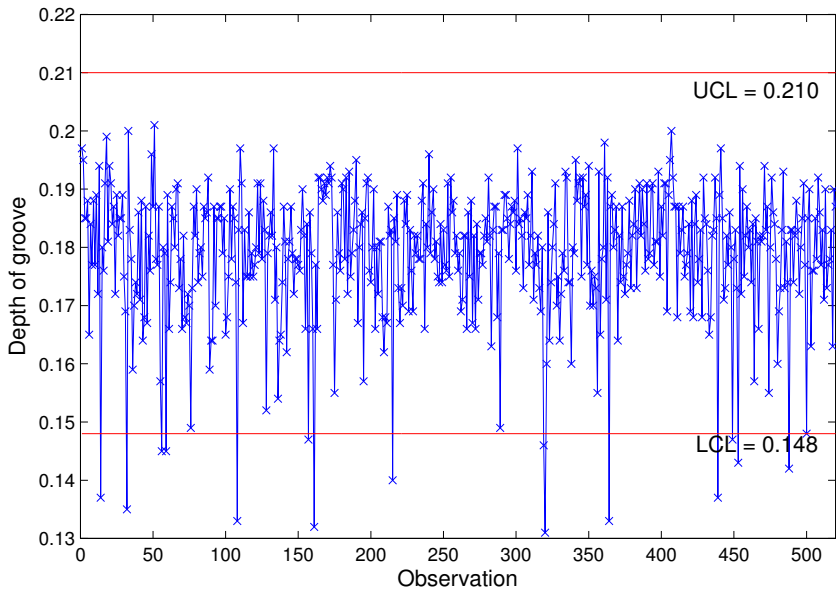

Figure 1.3: Shewhart individual control chart for the depth of a groove.

is in a state of statistical control. Looking at the control limits, we see that 15 out-of-control signals are given. In this case these are false alarms. Hence, based on the traditional control chart, the operators would look for trouble where there is none to be found, as the nature of this constant cause system is not in line with the normality assumption.

Approaches for non-normality in a constant cause system

If subgrouping is not possible and non-normality is evident, Shore (1998) distinguishes two categories of procedures in analyzing non-normal data. The first category attempts to make the procedure robust against deviations from normality. Burr (1967) studied the impact of the non-normality on the performance of control charts and corrected certain constants to obtain a better probability P in (1.1). Others who studied the effects of deviations from normality for control charts are Schilling and Nelson (1976), Borror et al. (1999), and Stoumbos and Reynolds (2000).

The second category focuses on the underlying distribution. The study of this category can be split into four approaches (cf. Shore (1998)):

1. Attempts to transform the data into a sequence of variables with a normal distribution. Possibilities are to use the Box-Cox transformation (cf. Box and Cox (1964)) or Johnson transformation (cf. Chou et al. (1998)). Statistical software programs like Minitab may be used to carry out such a transformation almost automatically.

2. Development of a control chart which is valid under certain mild distributional assumptions. For example, Ferrel (1958) has proposed a control chart based on log-normal distributions.

3. If large data sets are available, an attractive alternative to parametric statistical inference is to use these large data sets to study the distributional form, see for example Pearn and Kotz (1994). For statistical literature in the areas of goodness-of-fit and parametric modelling see D'Agostino and Stephens (1986).

4. Make use of non-parametric methods to calculate the control limits of the control charts. See for example Bai and Choi (1995) and Vermaat et al. (2003).

In Chapters 2 and 3 we will concentrate on the fourth alternative. If large data sets are available, i.e. at least 1,000 observations, non-parametric methods become attractive. In Chapter 2 we discuss individual control charts based on empirical quantiles, kernel estimators, and extreme-value theory. It turns out that these alternative individual control charts are quite robust against deviations from normality if indeed the number of observations are at least 1,000. Control charts for non-normal data are studied by Chou et al. (1998), Shore (1998), Borror et al. (1999), and Stoumbos and Reynolds (2000), among others.

If data sets contain less than 1,000 observations, Bayesian methods become attractive. In Chapter 3 we describe a semi-Bayesian method to calculate the control limits. This chapter is based on Vermaat and Does (2006). It turns out that this method works quite well for data sets with at least 250 observations. The estimation of control limits based on Bayesian methods is also studied in the literature. In Sturm et al. (1991) a discussion of the application of empirical Bayes for the estimation of a distribution of the characteristic under study is given (see also Woodward and Naylor (1993), Tagaras and Nikolaidis (2002), and Menzefricke (2002)). Other papers that use Bayesian methods in quality control are Girshick (1952), Carter (1972), Feltz and Shiau (2001), Girard and Parent (2001), and Hamada (2002).

Summarizing, if we have moderately large data sets, i.e. between 250 up to 1,000 observations, we can use the semi-Bayesian method to estimate the control limits (Chapter 3). If data sets are even larger, non-parametric methods can be applied (Chapter 2). In Chapter 3 the use of the control chart adapted for non-normality is illustrated with the example of the groove's depth data. If we have less than 250 observations not much can be done to obtain the necessary precision of the control chart, unless we make strong assumptions on the underlying distribution, see e.g. Quesenberry (1993), and Roes et al. (1999).

1.3 Control charts and autocorrelation

In this section we discuss the problem of autocorrelation of the observations from a constant cause system. This section is based on Vermaat et al. (2006a).

Today's manufacturing environment hardly resembles that in which control chart methods were originally introduced. Due to automated sensors, the sampling rate is often very high today. Hence, the assumption of independence of successive observations is likely to be violated. Serial correlation may seriously affect the performance of traditional control charts because these are based on the assumption of independent and identically distributed data. Thus serial correlation should not be ignored.

This raises a more fundamental question of what is required for a process to be statistically in-control. According to ANSI/ISO/ASQC Standard A3534-2-1993, statistical control is defined as: "[A] state in which the variations among the observed sampling results can be attributed to a system of chance causes that does not appear to change with time." There is general agreement that this definition implies that when in a state of statistical control, the mean and standard deviation of the process remain constant over time. The definition is sometimes also interpreted to imply that consecutive observations are statistically independent. However, this latter condition is unnecessarily restrictive. A weakly stationary time series exhibits a probability distribution with the first two moments constant. Hence, we would consider it to be in statistical control. See also Alwan and Roberts (1988) for a general discussion.

MacGregor (1991) argued that some or all autocorrelation should be eliminated by feedback control. Though situations may exist where such approach is useful, we believe that it is usually better to accept autocorrelation in the sense that a process can very well said to be in-control if we have a stationary process. The function of a control chart will then be to monitor the process to see if it remains stationary and otherwise to sound an alarm, see Alwan and Roberts (1988).

When working in industry and elsewhere, e.g. services industries, one is often required to proceed slowly and with caution. Trying to introduce tools which are too sophisticated may backfire. Indeed, it is typically advisable to recommend the use of tools that are already familiar to the operators. A sub-optimal tool in active use is better than an optimal one ignores. Thus, it will often be a good practical approach to use a familiar standard control chart, possibly with simple modifications, even when the process is known to violate some of the basic assumptions.

A motivating example: temperature readings from a ceramic furnace

To motivate the work presented in Chapter 4, consider the hourly temperature readings from a large ceramic furnace. When fully loaded, the furnace holds around 1,000 metric tons of material, and the total mass of the furnace itself is well beyond that. Furthermore, the control, exercised primarily via manual changes to the input fuel, has a relatively weak relationship to the actual temperature. The dead-time,

the time from when a change in fuel is made until the change takes effect, is in the order of 5 to 10 hours and the full effect occurs only after several more hours of delay. Ceramic furnaces of this size are typically operated 24 hours a day, 365 days a year. The operators work in three shifts and are typically well-trained in ceramic engineering. However, they often have little or no training in statistical process control theory. It is therefore important that control charts are relatively simple and easy to interpret. In particular, given the operator's intuitive understanding of the chemistry of the process, it is important that the displayed quantities are presented in the original metric, i.e. degree centigrade. Individuals Moving Range and EWMA control charts, although not necessarily the best charts, have therefore practical appeal in this environment.

Now consider the 80 consecutive hourly temperature readings from the ceramic furnace originally discussed by Bisgaard and Kulahci (2005). A time series plot of the data is shown in Figure 1.4. The process seems to vary around a fixed mean and shows limited variability. However it exhibits periodicities. Indeed it can be hard to distinguish this process from one that is not in statistical control.

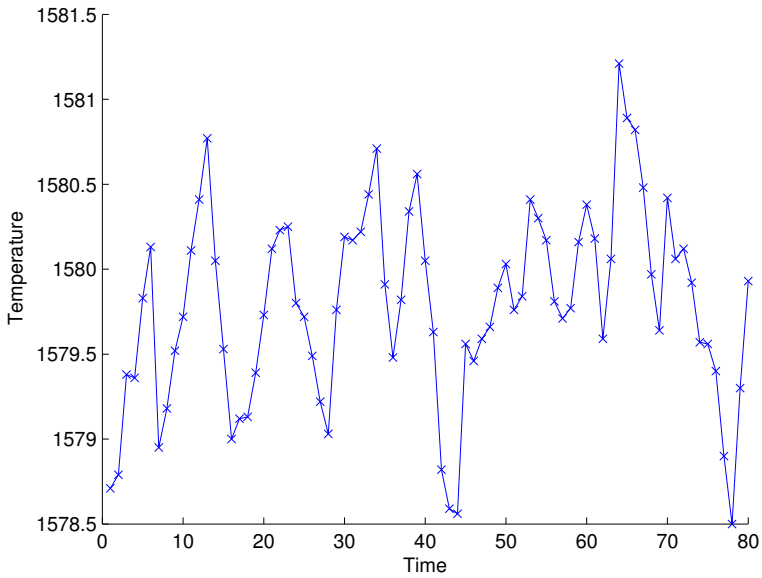

Figure 1.4: The time series plot of the furnace temperature data.

Bisgaard and Kulahci (2005) modelled these data as a stationary autoregressive process of order 2, AR(2). The parameter estimations were based on these 80 hourly observations. It may, of course be argued that 80 observations are not enough to secure good statistical estimates. However, in the present case the process had been refurbished and after 80 hours of operation since the process overhaul, there was an urgent need for establishing a new control chart for controlling

the process going forward, i.e. a Phase II application.

Now suppose we ignore the autocorrelation and apply a Exponentially Weighted Moving Average (EWMA) chart to this data. For a review of the basic theory of EWMA control charts see e.g. Box and Luceño (1997) and Montgomery (2004). Let $\{X_t\}$ for $t = 1, 2, \ldots$ denote observations of a process of independent and identically distributed (IID) random variables with $\mathrm{E}X_t = \mu$. The EWMA statistic at time t is defined as

$$W_t = \lambda X_t + (1 - \lambda) W_{t-1},$$

where λ is a constant, such that $0 < \lambda \leq 1$. Note, that if $\lambda \to 1$, then $W_t \to X_t$. Now let $W_0 = \mu$, then the variance of W_t is

$$\sigma^2_{W_t} = \sigma^2_X \left(\frac{\lambda}{2 - \lambda} \right) \left[1 - (1 - \lambda)^{2t} \right] \tag{1.4}$$

where σ^2_X is the variance of X, see Montgomery (2004). If t is sufficiently large, then it follows from (1.4) that the control limits of the EMWA control chart are approximately given by

$$\begin{aligned} \mathrm{UCL} &= \mu + c\sigma_X \sqrt{\frac{\lambda}{2 - \lambda}}, \\ \mathrm{LCL} &= \mu - c\sigma_X \sqrt{\frac{\lambda}{2 - \lambda}}, \end{aligned}$$

where $c > 0$ is a user specified constant, see Montgomery (2004). The parameters μ and σ_X have to be estimated. Natural choices are given in (1.2) and (1.3). For the furnace example we use $c = 2.86$ and $\lambda = 0.2$. With these choices the Average Run Length (ARL) – the average number of observations before a signal is given – are approximately 370, see Crowder (1989).

Figure 1.5 shows the resulting EWMA control chart with these standard parameters for the furnace data. Note that there are many out-of-control signals suggesting that the temperature is not in statistical control. Of course, it should be noted that when we would have applied the EWMA control chart retrospectively to the same data that are used to calculate the control limits, we engage in a Phase I application. However, if used prospectively, and if the process remains the same, we should expect similar out-of-control signals for new data.

Now an engineer familiar with ceramic processes will notice that the temperature over this 80 hour period covers a range of only 3 degrees centigrade. That is indeed a narrow range and unlikely to be of concern from a technical point of view. Thus anyone suggesting that this process is out-of-control would not earn the respect

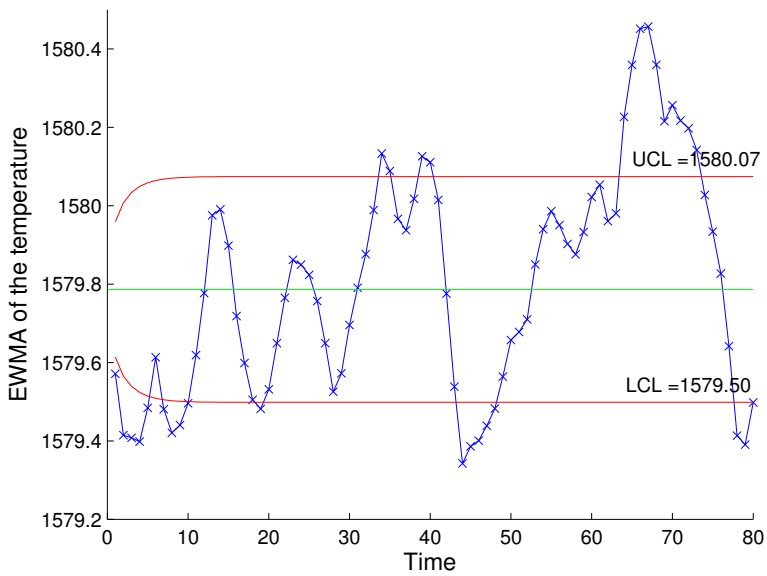

Figure 1.5: The EWMA control chart of the furnace temperature data.

of a ceramic engineer. Moreover, the size of the furnace suggests that it would be difficult and certainly not economically worthwhile to try to reduce this small variation. Thus we did not find it unreasonable that the process engineers in this case despairingly declared that traditional statistical process control methods gave ridiculous answers.

Approaches for autocorrelation in a constant cause system

Alwan and Roberts (1988) and Faltin et al. (1997) provided a comprehensive overview and discussion of modern process monitoring methods for processes that follow a linear discrete time series model. Lu and Reynolds (1999) provide a more recent overview of control charts for processes with autocorrelated data. Essentially two approaches have been promoted in the past.

The first approach is to use standard control charts with suitably modified control limits taking into account the autocorrelation. Zhang (1998) and (2000), Jiang et al. (2000), and Apley and Lee (2003), among others, adopted this approach. The practical appeal is that it is relatively simple to explain; the familiar "good old methods" can be used avoiding retraining and fear of new methods, and it allows operators to retain their physical intuition about the process. However, it may lead to charts that are less sensitive to detecting process changes.

The second approach is to fit an appropriate time series model to the data, and to monitor the residuals with standard control charts, for example individual charts, $\bar{X}$ charts, or EWMA charts. An early reference is Berthouex et al. (1978), but see also Alwan and Roberts (1988), Alwan (1991), Montgomery and Mastrangelo (1991), Wardell et al. (1994), and Koehler et al.

(2001). For the use of residual charting for the furnace data, see Bisgaard and Kulahci (2005). This approach is less intuitive, because the residuals can be difficult to interpret and are not presented in the original scale of the measurements.

In industrial practice, stationary processes are often be modelled as an autoregressive process of order one or two. Higher order processes, or processes better modelled with additional moving average terms may of course occur but are less common. Thus a closed form expression for the variance of the EWMA statistic under an AR(2) process, with the AR(1) process as a special case, will cover a significant number of practical applications. The EWMA control chart is frequently used without any reference to the underlying process. In other cases it is implied that the model is IID normal. The EWMA chart, when used for control, is based on the assumption that the observations of the process are IID. In Chapter 4 we present an explicit expression for the variance of the EWMA statistic for an AR(2) process. Using our closed form expression of the asymptotic variance we can relatively easily establish control limits for a Phase II application of an EWMA control chart. The use of this modified EWMA control chart will be illustrated in Chapter 4 with the ceramic furnace example.

1.4 Outline of the thesis

In the next two chapters we examine the problem of non-normality of the constant cause system. In Chapter 2 we discuss the non-parametric solution to construct control charts. This chapter is based on Vermaat et al. (2003) and the appendix is based on Vermaat et al. (2005). It seems that this method works quite well for large data sets, i.e. for at least 1,000 observations. In Chapter 3, which is based on Vermaat and Does (2006), the gap between 250 observations and 1,000 observation is closed by the introduction of a semi-Bayesian method. The last chapter discusses the issue of autocorrelation in the constant cause system. This chapter is based on Vermaat et al. (2006a) and Vermaat et al. (2006b).

Each of the following chapters stands on its own. Thus, some modest overlap in the chapters is unavoidable.

Chapter 2

A comparison of Shewhart individual control charts based on normal, non-parametric, and extreme-value theory

In this chapter several control charts for individual observations are compared. Traditional ones are the well-known Shewhart individual control charts based on the average of the moving ranges. Alternative ones are non-parametric control charts based on empirical quantiles, on kernel estimators, and on extreme-value theory. Their in-control and out-of-control performance are studied by simulation combined with computation. It turns out that the alternative control charts are not only quite robust against deviations from normality but also perform reasonably well under normality of the observations. The performance of the Empirical Quantile control chart is excellent for all distributions considered, if the Phase I sample is sufficiently large. This chapter is based on Vermaat et al. (2003), the Appendix is based on Vermaat et al. (2005).

2.1 Introduction

In the present chapter, we consider the non-parametric situation for individual measurements in which the underlying distribution function, denoted by F, is assumed to be unimodal, but otherwise unknown. This means that we include unimodal distributions, such as the normal, Student, uniform, exponential, Laplace, and logistic distribution. It has been shown by Ion and Klaassen (2005) that for densities with more than one mode, any Shewhart control chart for individual measurements is inappropriate, or at least suboptimal.

In our comparison we will use the standard, essentially parametric Shewhart control chart with control limits based on the average of the moving ranges of the individual measurements. The alternative control charts we will compare, are charts based on empirical quantiles (which are related to the bootstrap method), charts based on kernel estimators, and charts based on

extreme-value theory. The availability of modern computing power in statistical process control, enables one to apply these computationally intensive techniques.

The determination of the control limits is based on the observations obtained in the so-called Phase I, in which the data are collected from the production process and parameters are estimated (cf. Woodall and Montgomery (1999)). In the present chapter, we consider the monitoring phase which is usually called Phase II. In most evaluations and comparisons of performance of control charts in Phase II, it is assumed, as noted in Woodall and Montgomery (1999), that the in-control parameters are known, which is not the case in practice. For this reason, the statistical performance of the classical and newly proposed control charts will be studied by simulating the average and standard deviation of the in-control and out-of-control run length in Phase II of the control charts with the control limits determined by in-control observations from Phase I.

In Quesenberry (1993) it has been shown by simulation that in order to estimate control limits for individual measurements sufficiently accurately, one needs rather large samples of 300 or more observations. Our simulation results support this for all control charts considered.

It turns out that the control chart based on the average of the moving ranges is suboptimal compared to the newly proposed control charts, except for independent and normally distributed random variables. However, even under normality, the alternative charts have quite good performance, especially when sufficiently many data are available. The control chart based on empirical quantiles is excellent for all distributions considered.

This chapter is organized as follows. In the next section, the four control charts are defined. Next, results are given of an extensive Monte Carlo study based on 10,000 simulations for six different distributions and for several Phase I sample sizes. This number of simulations turned out to be large enough for our simulation method. Finally, conclusions are given.

2.2 Description of the control charts

We will consider the usual control charts with a lower control limit (LCL) and an upper control limit (UCL). This means that if the measurement value X is lower than LCL or higher than UCL, then the process is called out-of-control.

2.2.1 Control chart based on the average of the moving ranges

If the distribution function F is assumed to be normal, then the traditional Shewhart individual control chart has limits defined by

$$\begin{aligned} \mathrm{UCL} &= \mu + \Phi^{-1}(1 - \frac{\alpha}{2})\sigma, \\ \mathrm{LCL} &= \mu + \Phi^{-1}(\frac{\alpha}{2})\sigma, \end{aligned}$$

where Φ^{-1} is the standard normal quantile function, and where μ is the mean and σ is the standard deviation of the normal distribution function F. Level α is the false alarm rate. In practice normality is questionable and μ and σ are unknown. However, we assume that a Phase I sample $X_1, X_2, \cdots, X_k$ of independent and identically distributed random variables is available such that μ and σ can be estimated. Classical estimators of μ and σ are the sample mean $\bar{X}_k = \sum_{i=1}^{k} X_i/k$ and the sample standard deviation $S_k = \sqrt{\sum_{i=1}^{k} (X_i - \bar{X}_k)^2/(k-1)}$. The sample standard deviation is asymptotically efficient for independent and identically distributed normal random variables, but it has the disadvantage that it is sensitive to trends and oscillations. Consequently, when such phenomena might occur, we have to use estimators of the standard deviation that are less sensitive to these deviations, cf. Kamat (1953). The average of the moving ranges can be scaled by $2/\sqrt{\pi}$ to obtain such an estimator, and is defined by

$$\overline{\mathrm{MR}}_k = \frac{1}{k-1} \sum_{i=2}^{k} |X_i - X_{i-1}|.$$

In Duncan (1986) the individual control chart is mentioned with control limits based on the average of the moving ranges (AMR) defined by

$$\begin{aligned} \mathrm{UCL}_{\mathrm{AMR}} &= \bar{X}_k + \Phi^{-1}(1 - \frac{\alpha}{2}) \frac{\sqrt{\pi}}{2} \overline{\mathrm{MR}}_k, \\ \mathrm{LCL}_{\mathrm{AMR}} &= \bar{X}_k + \Phi^{-1}(\frac{\alpha}{2}) \frac{\sqrt{\pi}}{2} \overline{\mathrm{MR}}_k. \end{aligned}$$

In practice, this AMR control chart is the standard chart for individual observations. The reasons for this are: First, the constant by which the average of the moving ranges has to be multiplied in order to obtain an unbiased estimator of σ, is quite similar for probability distributions having density curves of very different shapes, cf. Burr (1967). Secondly, the AMR control chart tends to perform reasonably well for moderate Phase I sample sizes more or less independently of the probability distribution the observations stem from. This is thoroughly studied in Wheeler (1995). In this chapter we compare the performance of the AMR control chart to three competitors.

A more exact version of the individual control chart is developed in Roes et al. (1993). In this chapter we will restrict attention to the control chart described above. In Chapter 3 the more exact version of the individual control chart is studied.

2.2.2 Control chart based on empirical quantiles (bootstrap)

A natural estimator of the q-quantile of the distribution function F is the empirical quantile $\hat{F}_k^{-1}(q)$, which is defined as

$$\hat{F}_k^{-1}(q) = \inf\{x |\ \hat{F}_k(x) \geq q\}, \ \ 0 < q < 1,$$

where $\hat{F}_k$ is the empirical distribution function that puts mass $1/k$ at each X_i, $1 \leq i \leq k$, i.e.

$$\hat{F}_k(x) = \frac{1}{k}\sum_{i=1}^{k} I_{\{X_i \leq x\}}, \quad -\infty < x < \infty,$$

with I the indicator function, i.e. $I_{\{x \leq y\}}$ equals 1 if $x \leq y$ holds and 0 otherwise. Now, an obvious estimator of the upper control limit is based on the empirical quantile (EQ)

$$\mathrm{UCL_{EQ}} = \hat{F}_k^{-1}(1 - \frac{\alpha}{2}) = X_{(\lceil (1-\frac{\alpha}{2})k \rceil)},$$

with $X_{(1)} \leq X_{(2)} \leq \cdots \leq X_{(k)}$ denoting the order statistics of the initial sample $X_1, X_2, \cdots, X_k$ and $\lceil y \rceil$ the smallest integer not smaller than y. The lower control limit estimated by the empirical quantile is defined by

$$\mathrm{LCL_{EQ}} = X_{(\lfloor \frac{\alpha}{2}k+1 \rfloor)},$$

where $\lfloor y \rfloor$ denotes the largest integer not larger than y. The exceedance probability for this estimation method is not always smaller than α. To be more specific:

$$\mathrm{P}(X > \mathrm{UCL_{EQ}} \text{ or } X < \mathrm{LCL_{EQ}}) = 2\frac{\lfloor \frac{\alpha}{2}k \rfloor + 1}{k+1} \in \left(\alpha - \frac{\alpha}{k+1}, \alpha + \frac{2-\alpha}{k+1}\right].$$

To overcome this problem another estimation method is proposed in Ion and Klaassen (2005), which guarantees that the exceedance probability equals at most the significance level α. However, the resulting alternative does not have a better performance than the EQ control chart for shifts in the mean. Therefore, we will not consider this alternative.

It can be shown that the Empirical Quantile control chart is a special case of the bootstrap control chart defined by Willemain and Runger (1996). The philosophy of the bootstrap approach to statistical problems is to replace the unknown distribution function F of a random variable X by an empirical distribution function. In control charts for individual observations one would like to use the $\mathrm{UCL_{EQ}} = F^{-1}(1 - \alpha/2)$. Consequently, the bootstrap approach applied to control charts for individual observations, with a Phase I sample $X_1, X_2, \cdots, X_k$, yields $\hat{F}_k^{-1}(1-\alpha/2)$ as an estimate of the upper control limit. For the more general situation of control charts for averages the bootstrap method is studied in Liu and Tang (1996). Note that in Jones and Woodall (1998) several control charts based on the bootstrap are compared.

2.2.3 Control chart based on kernel estimators

In Rosenblatt (1956) and Parzen (1962) kernel estimators of the density function f are defined as

$$\hat{f}_w(x) = \frac{1}{k}\sum_{i=1}^{k} \frac{1}{h} w(\frac{x - X_i}{h}), \quad -\infty < x < \infty,$$

where the kernel w is non-negative such that $\int_{-\infty}^{\infty} w(x)dx = 1$, and where the bandwidth h is positive and small. Consequently, with $W(x) = \int_{-\infty}^{x} w(y)dy$ the distribution function corresponding to the density w,

$$\hat{F}_w(x) = \frac{1}{k}\sum_{i=1}^{k} W(\frac{x - X_i}{h}), \quad -\infty < x < \infty,$$

is a kernel estimator of the distribution function F. As in Reiss (1989) a smooth alternative to the conventional sample quantile function may be defined by

$$\hat{F}_w^{-1}(q) = \inf\{x \mid \frac{1}{k}\sum_{i=1}^{k} W(\frac{x - X_i}{h}) \geq q\}\ , \quad 0 < q < 1.$$

The choice of the bandwidth h is more important than the choice of the kernel w. A relatively large value of h gives too much smoothness, and a relatively small value of h gives big fluctuations. Several choices of the kernel are possible, e.g. the Gaussian kernel and the so-called Epanechnikov kernel, i.e.

$$w(x) = \begin{cases} \frac{3}{4\sqrt{5}}(1 - \frac{x^2}{5}), & \text{if} \quad |x| < \sqrt{5}, \\ \\ 0, & \text{if} \quad |x| \geq \sqrt{5}. \end{cases}$$

In this chapter we restrict attention to the Epanechnikov kernel because it gives better results than the Gaussian kernel (cf. Ion (2001)).

In view of Azzalini (1981), the optimal choice of the bandwidth h is $h = Ck^{-1/3}$, where C is a constant that depends on σ, the standard deviation of F. Azzalini did some numerical work using the Epanechnikov kernel, and he concluded that good values for the constant C are between σ and 2σ for a large number of distributions. Based on an extensive simulation study we found that the best choice in our situation is $C = 2\sigma$. Since σ is unknown, we estimate it by the sample standard deviation $S_k = \sqrt{\sum_{i=1}^{k}(X_i - \bar{X}_k)^2/(k-1)}$. Note, that we could also use other estimators for σ like the one based on the average of the moving ranges. So for the Epanechnikov kernel (EK) the control limits are:

$$\begin{aligned} \text{UCL}_{\text{EK}} &= \inf\{x \mid \frac{1}{k}\sum_{i=1}^{k} W(\frac{x - X_i}{2k^{-1/3}S_k}) \geq 1 - \alpha/2\}, \\ \text{LCL}_{\text{EK}} &= \sup\{x \mid \frac{1}{k}\sum_{i=1}^{k} W(\frac{x - X_i}{2k^{-1/3}S_k}) \leq \alpha/2\}. \end{aligned}$$

2.2.4 Control chart based on extreme-value theory

Extreme-value theory yields another method to estimate the control limits of a Shewhart control chart. To the best of our knowledge, Shewhart control charts based on this theory have not been considered before.

Define

$$M_k^{(r)} = \frac{1}{m}\sum_{j=1}^{m}(\log X_{(k-j+1)} - \log X_{(k-m)})^r, \tag{2.1}$$

$$\bar{M}_k^{(r)} = \frac{1}{m}\sum_{j=1}^{m}(\log X_{(j)} - \log X_{(m+1)})^r,$$

where the integer r takes the values $r = 1, 2$ and m is the number of upper respectively lower order statistics used in the estimation of the control limits.

Extreme-value theory deals with the tail behavior of distributions. These tails can be modelled by an extreme-value distribution, which is determined by an extreme-value index γ (cf. Dekkers et al. (1989)). If we do not make any assumptions on γ, we may use the moment estimator of it, defined by

$$\hat{\gamma}_k = M_k^{(1)} + 1 - \frac{1}{2}\left\{1 - \frac{(M_k^{(1)})^2}{M_k^{(2)}}\right\}^{-1}. \tag{2.2}$$

Furthermore, the $1-q$-quantile of the distribution function F is estimated in Dekkers et al. (1989) as

$$\hat{F}_k^{-1}(1-q;\hat{\gamma}_k) = X_{(k-m)} + \frac{(m/(kq))^{\hat{\gamma}_k} - 1}{\hat{\gamma}_k}(1 - (\hat{\gamma}_k \wedge 0))X_{(k-m)}M_k^{(1)} \tag{2.3}$$

with $0 < q < \frac{1}{2}$. Recall that $x \wedge y$ and $x \vee y$ denote the minimum and maximum respectively, of x and y. Consequently, the upper control limit of the Extreme-Value theory (EV) control chart, with $0 < \frac{\alpha}{2} < \frac{1}{2}$, may be defined by

$$\mathrm{UCL_{EV}} = \hat{F}_k^{-1}(1 - \frac{\alpha}{2};\hat{\gamma}_k). \tag{2.4}$$

In a similar way, the lower control limit may be taken to be

$$\mathrm{LCL_{EV}} = X_{(m+1)} + \frac{(m/(k\alpha/2))^{\bar{\gamma}_k} - 1}{\bar{\gamma}_k}(1 - (\bar{\gamma}_k \wedge 0))X_{(m+1)}\bar{M}_k^{(1)}, \tag{2.5}$$

with $\bar{\gamma}_k$ defined as

$$\bar{\gamma}_k = \bar{M}_k^{(1)} + 1 - \frac{1}{2}\left\{1 - \frac{(\bar{M}_k^{(1)})^2}{\bar{M}_k^{(2)}}\right\}^{-1}.$$

Of course, the sequence $m = m(k)$ has to be chosen appropriately in order to obtain good performance of $\mathrm{UCL_{EV}}$ and $\mathrm{LCL_{EV}}$. In the numerical evaluation of the Extreme-Value theory control chart we have considered several values of m. We found that $5 \vee (k/500)$ is a reasonable choice.

Many estimators based on extreme-value theory are not invariant under a shift of location. In Appendix 2A modified estimators for the quantiles in (2.4) and (2.5) are derived, which are

invariant under a shift of location. Appendix 2A is based on Vermaat et al. (2005). Moreover, these modified quantile estimators are also symmetric with respect to the mean for symmetric distributions, which is not the case for some of the estimators based on EVT. The modified quantile estimators are a limiting result of the estimators proposed by Dekkers et al. (1989). The results can be used in establishing control limits for Shewhart control charts.

2.3 Simulations

In order to study the performance of the control charts introduced in Section 2.2 we have conducted a simulation experiment, for some choices of the distribution with distribution function F and of the Phase I sample size k. To describe our simulation procedure let us assume as above that we have a sample $X_1, X_2, \ldots, X_k$ of size k from a distribution function F. This training sample is used to estimate the LCL and UCL by $\widehat{\mathrm{LCL}}$ and $\widehat{\mathrm{UCL}}$, respectively. Note, that the process generating this training sample is in statistical control. We do not study robustness properties, regarding moderately out-of-control situations during Phase I. We denote by $\hat{p}_k(X_1, \ldots, X_k)$ the conditional probability given the training sample, that a new independent random variable X from the same distribution F exceeds this upper control limit or is below the lower control limit, i.e.

$$\hat{p}_k(X_1, \cdots, X_k) = \mathrm{P}\big(X > \widehat{\mathrm{UCL}} \text{ or } X < \widehat{\mathrm{LCL}} \,|\, X_1, \cdots, X_k\big).$$

Given the training sample and hence the control limits, the run length RL of the resulting control chart is a random variable with a geometric distribution with parameter $\hat{p}_k(X_1, \ldots, X_k)$ and consequently with average run length (ARL)

$$\mathrm{E}(\mathrm{RL}|X_1, \ldots, X_k) = 1/\hat{p}_k(X_1, \ldots, X_k).$$

Note that this is a conditional average run length and hence a random variable. Taking the expectation over the training sample $X_1, \ldots, X_k$ we get the unconditional average run length

$$\mathrm{ARL} = \mathrm{E}(\mathrm{RL}) = \mathrm{E}(\mathrm{E}(\mathrm{RL}|X_1, \ldots, X_k)) = \mathrm{E}\frac{1}{\hat{p}_k(X_1, \ldots, X_k)}.$$

Since this expectation cannot be computed directly, we have simulated it by generating at each instance 10,000 training samples, by computing for each training sample $(x_1, \ldots, x_k)$ the average run length $1/\hat{p}_k(x_1, \ldots, x_k)$, and by averaging these average run lengths over the 10,000 training samples. In the same way it can be seen that the unconditional standard deviation of

the average run length equals

$$\begin{aligned}\mathrm{SDRL} &= \sqrt{\mathrm{Var}(\mathrm{RL})} \\ &= \sqrt{\mathrm{E}(\mathrm{Var}(\mathrm{RL}|X_1,\ldots,X_k)) + \mathrm{Var}(\mathrm{E}(\mathrm{RL}|X_1,\ldots,X_k))} \\ &= \sqrt{2\mathrm{E}(\frac{1}{\hat{p}_k(X_1,\ldots,X_k)})^2 - (\mathrm{E}\frac{1}{\hat{p}_k(X_1,\ldots,X_k)})^2 - \mathrm{E}\frac{1}{\hat{p}_k(X_1,\ldots,X_k)}}.\end{aligned}$$

Moreover, we have simulated this SDRL by computing for each training sample also $(1/\hat{p}_k(x_1,\ldots,x_k))^2$, by averaging these squares over the 10,000 training samples, and by substituting the appropriate averages into the above formula for the SDRL.

To study the performance of the control chart in an out-of-control situation we have studied shifts in the mean. We have considered shifts $\delta\sigma$ with δ ranging from 0.25 through 5 where σ is the standard deviation of the studied distribution. Of course, we also have considered $\delta = 0$ in order to study the in-control properties of the control charts. Given the 10,000 training samples $(x_1,\ldots,x_k)$, we have calculated for each shift the average run length ARL and the standard deviation of the run length SDRL as described above.

This procedure of simulating the performance of control charts differs apparently from the one used by e.g. Quesenberry (1993), where once the $\widehat{\mathrm{LCL}}$ and $\widehat{\mathrm{UCL}}$ have been obtained from the training sample, one realization of the run length is obtained by again simulation. However, this second simulation step is not necessary, as we have seen above, since the conditional average run length can be computed exactly. Of course, there is no point in simulation when exact computation is possible.

In our simulations the false alarm rate α is chosen to be equal to 0.0027, since this value yields the traditional 3σ limits in the classical Shewhart control chart. Consequently, if the variation of $\hat{p}_k(X_1,\ldots,X_k)$ would have been small, the mean (ARL) and the standard deviation (SDRL) of the run length should both have been close to 370 under $\delta = 0$.

The simulations have been done for 6 different choices of distribution function F, namely the normal, Student's t with 4 degrees of freedom, uniform, exponential, Laplace, and logistic distribution and for sample sizes k equal to 250; 500; 1,000; 2,500; 5,000; and 10,000. The shift in the mean of size $\delta\sigma$ is done for 17 different values of δ, namely 0 (0.25) 3.5, 4, and 5.

The results of the simulations are presented for each distribution by means of two complementary figures indicating the ARL and SDRL. Although we have simulated at six different sample sizes k, we present here the results only for $k = 1{,}000$ and under normal F also for $k = 250$. A complete survey of the simulation results may be found in the master thesis of Vermaat (2003). The simulation results are valid for all values of σ, because all methods are scale invariant.

Under normality the control chart based on the average of the moving ranges has to perform

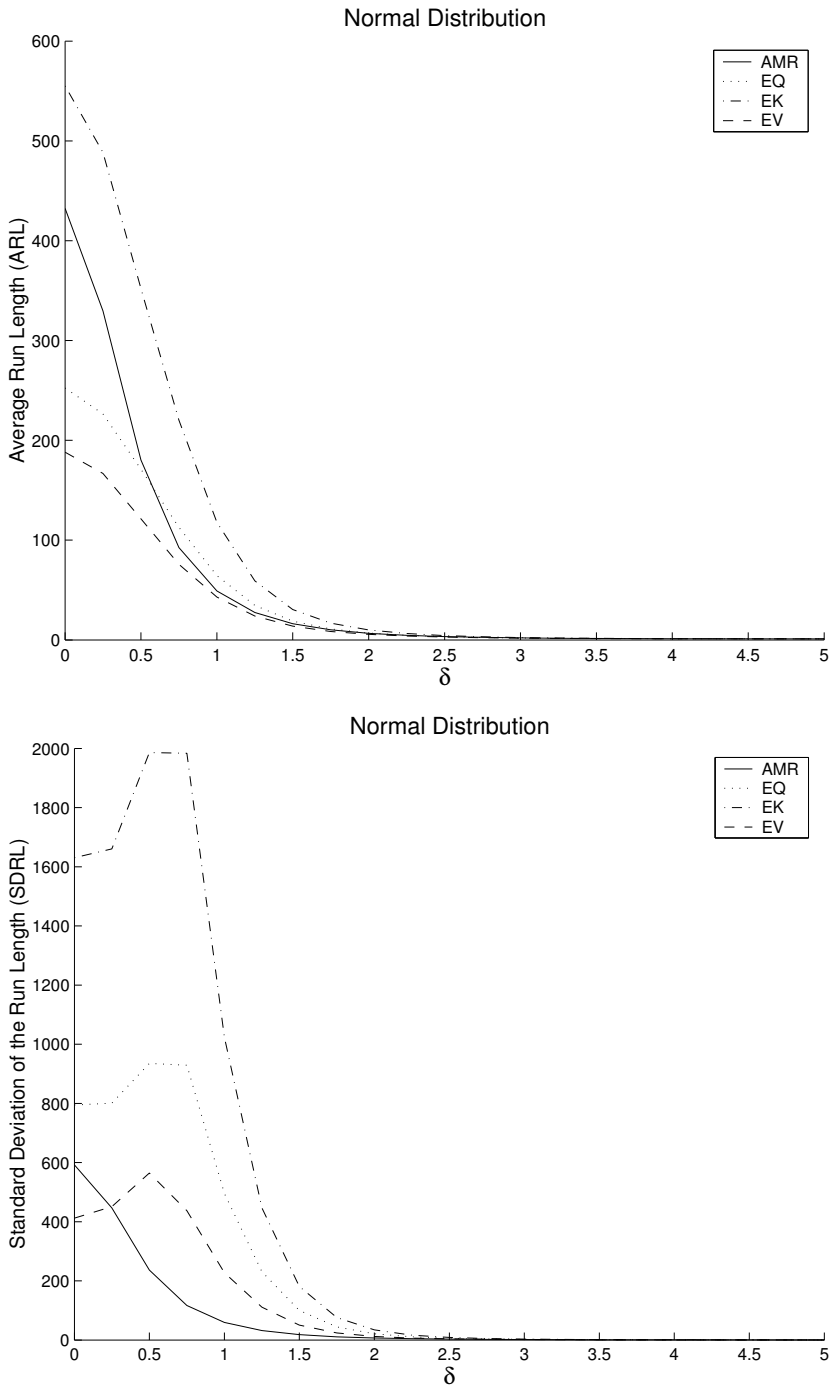

Figure 2.1: The ARL and the SDRL of the four control charts under normality for $k = 250$ and standardized shifts δ.

closely to optimal. This is confirmed by Figure 2.1, which also shows that the other control charts perform rather badly. Apparently, the sample size $k = 250$ is too small for them. The behavior of all four control charts under the non-normal distributions is rather bad overall. The same conclusions hold for $k = 500$. Note, that for these 'small' values of k the EQ control chart has control limits equal to the smallest and largest observation. From $k = 1{,}000$ onwards the differences between the control charts become clearer. In Figure 2.2 we see that all control charts behave quite similarly under normality.

In Figure 2.3 the results for a t_4 distribution are given. Indeed, the Average Moving Range control chart generates a lot of false alarms. The behavior of the other three control charts is quite reasonable for the in-control situation. However, few false alarms tend to create few real alarms under shifts.

In Figure 2.4 the results for uniform distributions are given. Only the Empirical Quantile control chart has finite ARL and SDRL. The other control charts are not even applicable when the underlying distribution is uniform.

Figure 2.5 presents the results for exponential distributions. We see, that the Empirical Quantile control chart has a maximum in the ARL around $\delta = 0.25$. This maximum is due to the fact, that the Empirical Quantile control chart estimates an LCL within the support of the exponential distribution. If the process shifts towards the UCL when δ increases, the LCL is of course harder to violate and the UCL easier. Since the density of the exponential is larger near the LCL than near the UCL, this causes the probability of an alarm to decrease and hence the ARL to increase for small δ. Because the estimates of the LCL in the Epanechnikov Kernel and the Extreme-Value control charts typically fall outside the support of the exponential distribution, these control charts do not show this phenomenon. The EK control chart and the EQ control chart show almost identical performance for $\delta > 0.25$. Note, that the behavior of the AMR control chart for $\delta = 0$ is really bad.

In Figure 2.6 and in Figure 2.7 the results are given for logistic and Laplace distributions respectively. We see again that the AMR control chart performs very poorly for $\delta = 0$. The other control charts perform quite similarly. Note, that although the shape of a logistic distribution is comparable with a normal distribution, the tails are completely different.

For all studied distributions it holds that if the sample size increases the results of the Empirical Quantile, the Epanechnikov Kernel, and the Extreme-Value theory control charts get closer to each other. The EQ control chart is the only one which behaves reasonably for a uniform distribution. The AMR control chart generates a lot of false alarms for non-normal distributions.

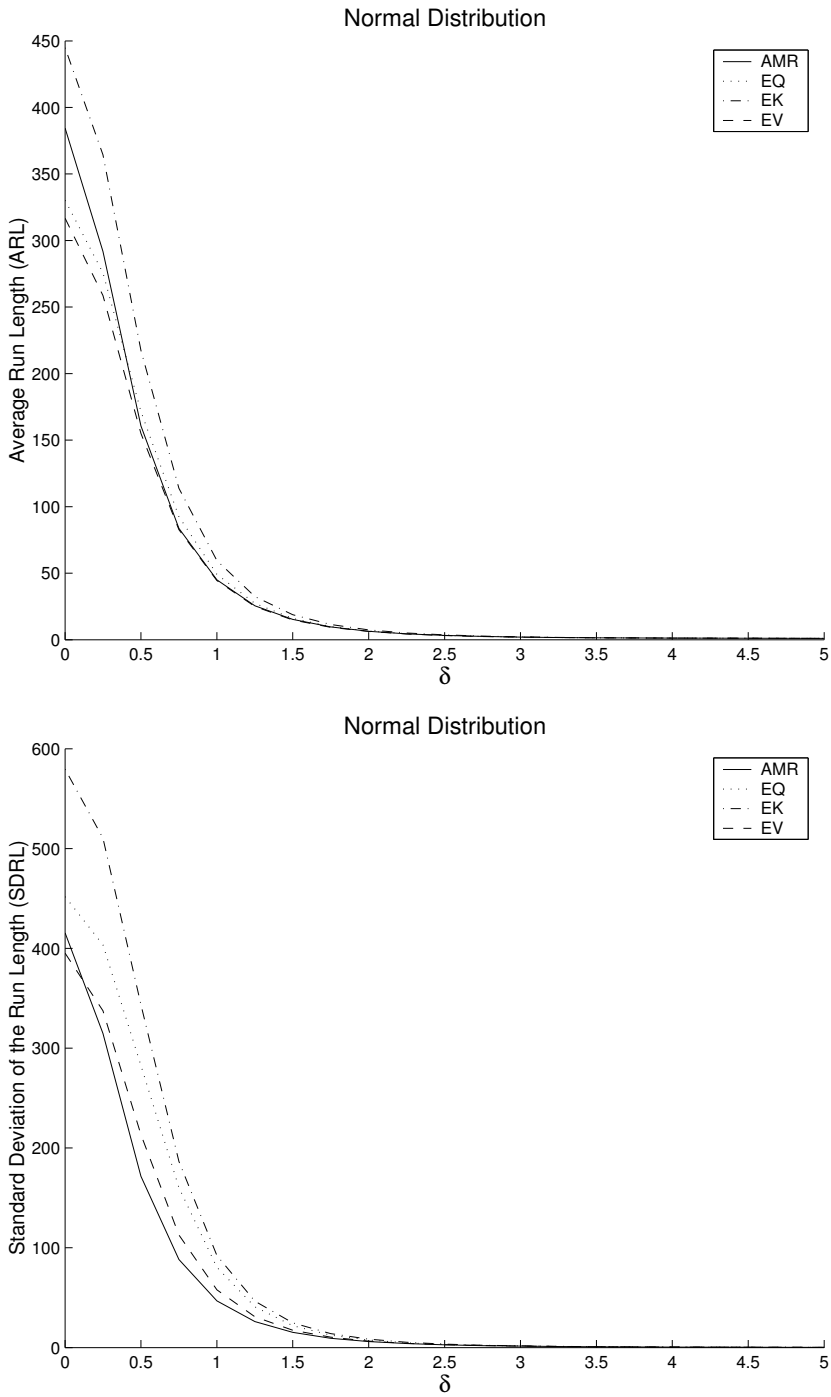

Figure 2.2: The ARL and the SDRL of the four control charts under normality for $k = 1{,}000$ and standardized shifts δ.

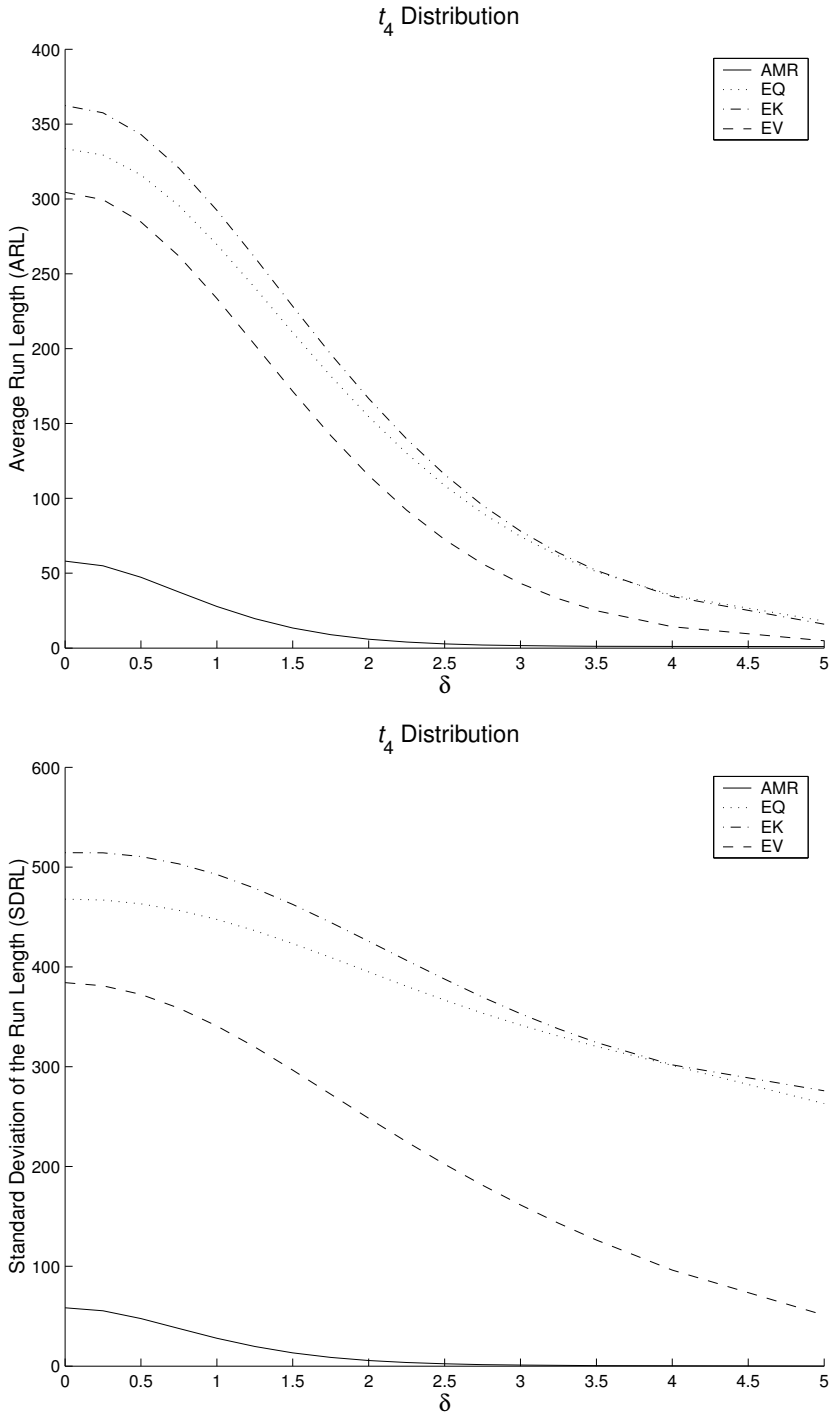

Figure 2.3: The ARL and the SDRL of the four control charts under a Student t_4 distibution for $k = 1{,}000$ and standardized shifts δ.

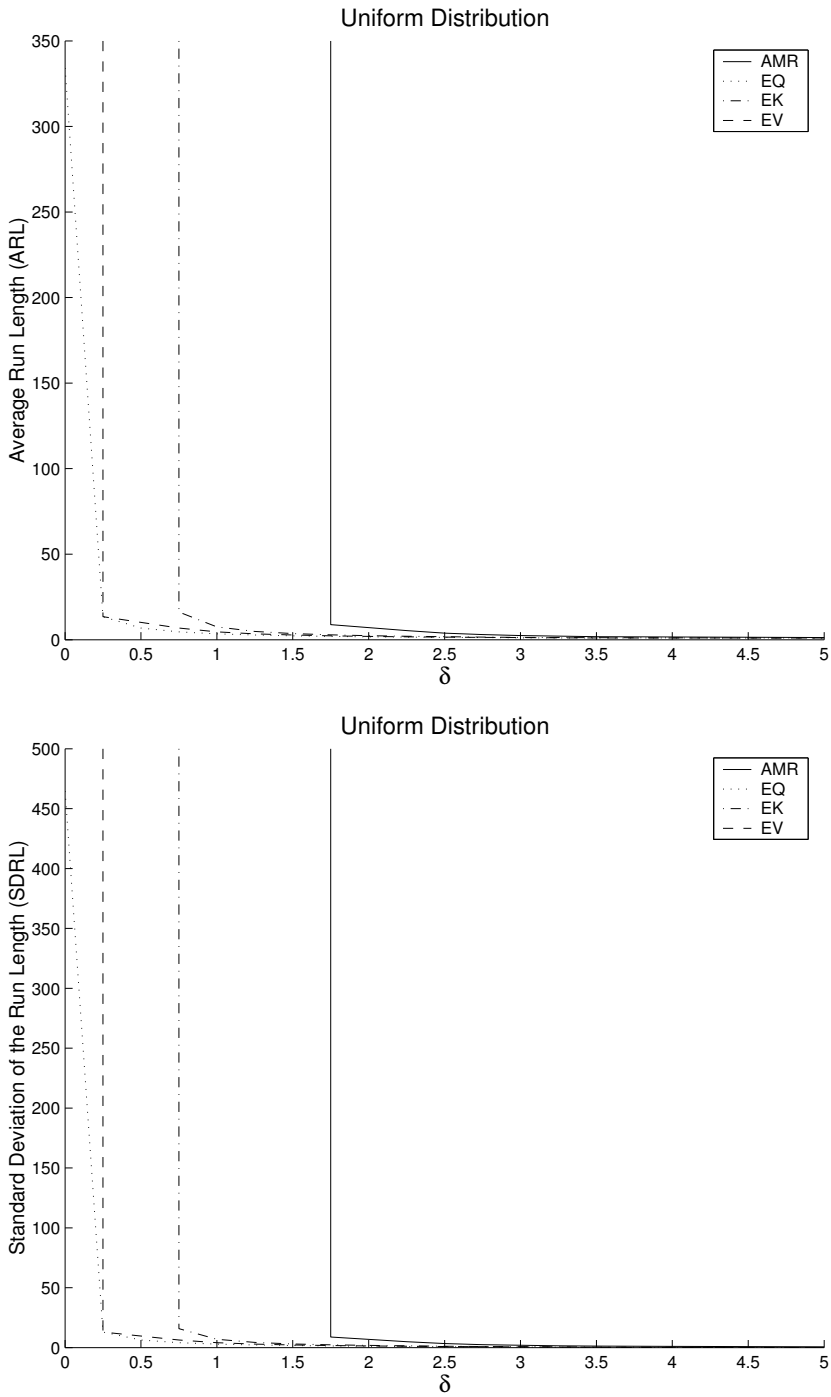

Figure 2.4: The ARL and the SDRL of the four control charts under uniformity for $k = 1{,}000$ and standardized shifts δ.

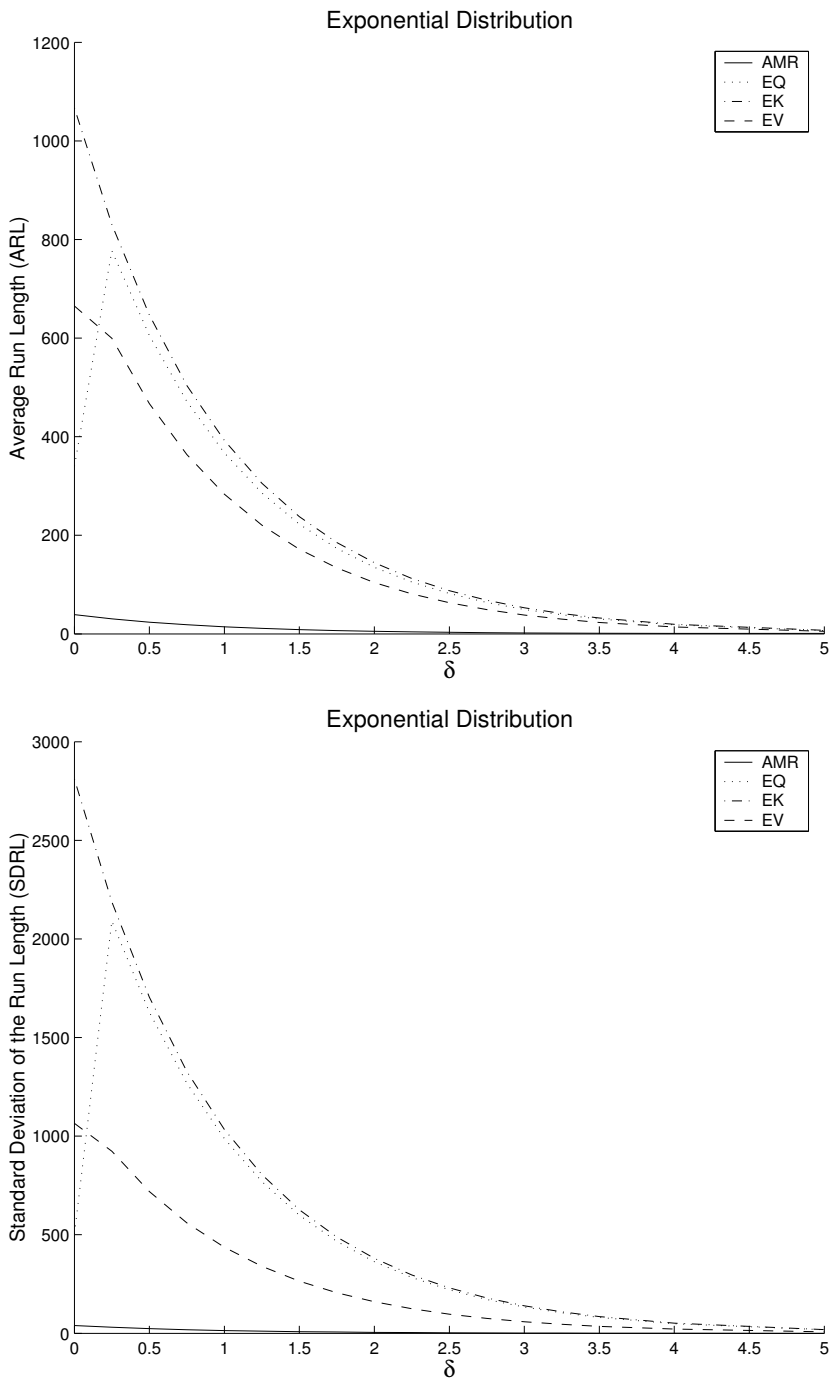

Figure 2.5: The ARL and the SDRL of the four control charts under an exponential distribution for $k = 1{,}000$ and standardized shifts δ.

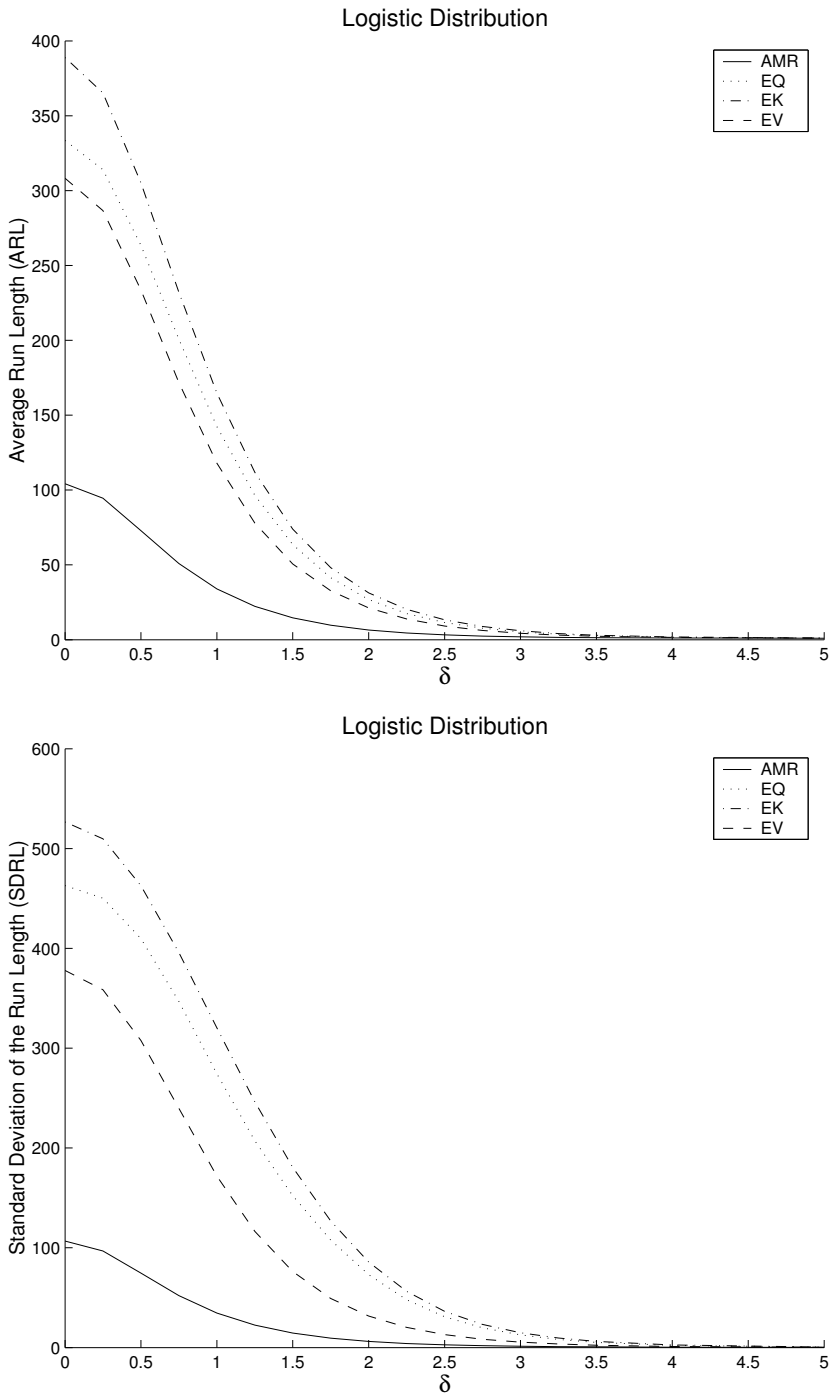

Figure 2.6: The ARL and the SDRL of the four control charts under a logistic distribution for $k = 1{,}000$ and standardized shifts δ.

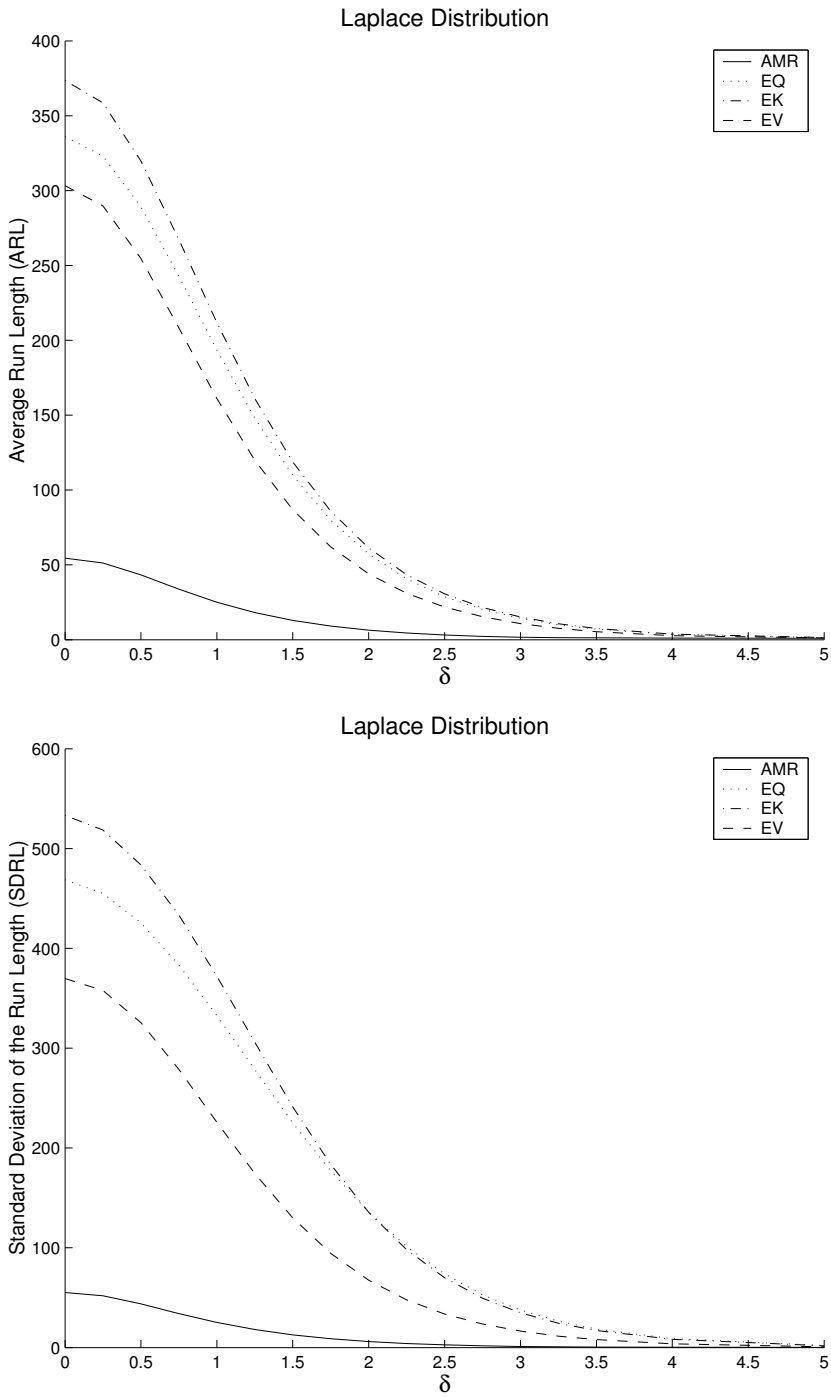

Figure 2.7: The ARL and the SDRL of the four control charts under a Laplace distribution for $k = 1{,}000$ and standardized shifts δ.

2.4 Conclusion

If the distribution function F is normal, then the Average Moving Range control chart based on the average of the moving ranges behaves quite well, but under non-normal distributions its performance is extremely bad in the in-control situation. Hence it is reasonable to consider other control charts based on e.g. non-parametric and extreme-value theory.

The Empirical Quantile control chart has the advantage that it is easy to compute and distribution-free in the in-control situation. Simulations show that for a broad range of distributions we get almost the same results. When we look to the control charts based on the Epanechnikov kernel and extreme-value theory respectively, we may conclude that their behavior is quite similar to that of the Empirical Quantile control chart, except for the uniform distribution. The Empirical Quantile control chart turns out to be the best of all proposed control charts. It can be improved only partially for some specific situations, e.g. when normality holds.

It should be noted that the non-parametric and the Extreme-Value theory control charts need more than 500 and preferably at least 1,000 observations from Phase I in order to attain reasonable performance. Such a limitation is not present for the AMR control chart, which may work well in some situations even for a training set of size much smaller than 250, cf. Wheeler (1995).

Appendix 2A: Convergence of control limits obtained with extreme-value theory

The control limits based on the quantile estimators of Dekkers et al. (1989) are for certain distributions, e.g. the uniform distribution, not quite satisfactory. It appears that for the uniform distribution on $[0, 1]$ the upper and lower control limits are not symmetric with respect to the average (i.e. 0.5). The reason for this is as follows. The calculated lower control limit, i.e. the small quantile, is very close to zero. Dekkers et al. (1989) use a log-transformation to estimate the quantiles. The log-function is very steep in the neighborhood of zero, which causes asymmetry in the estimated small and large quantiles. To tackle this problem, a large shift of location is performed. We recalculate the quantiles for the transformed data. To get the quantiles for the original data, an inverse transformation of the calculated quantiles is executed. The question is how large this shift of location has to be in order to get appropriate results. In other words, what happens if the shift of location goes to infinity? In this appendix we derive the limiting values for large and small quantile estimators. These quantiles can be used as control limits for designing a Shewhart control chart. As an illustration we calculate the control limits of a uniform distribution on $[0, 1]$ based on the work of Dekkers et al. (1989) as well as the modified estimator.

Convergence of the quantiles

First we shift the original observations with a constant K. For the transformed observations we estimate the quantiles with Dekkers' et al. (1989) method. Applying the inverse transformation we obtain estimators for the quantiles of the original data.

Define $X^* := X + K$, where X is a random variable with an absolutely continuous probability distribution function F and K is a constant. By replacing X with X^* the definitions of $\hat{\gamma}_k$ in (2.2) and $M_k^{(r)}$ in (2.1) are modified in $\hat{\gamma}_k^*$ and $M_k^{(r)^*}$. The $(1-q)$-quantile of the distribution function of X^* can be easily obtained from (2.3) by replacing $X_{(k-m)}, \hat{\gamma}_k$ and $M_k^{(1)}$ by $X_{(k-m)}^*, \hat{\gamma}_k^*$ and $M_k^{(r)^*}$, respectively. A modified estimator of the $(1-q)$-quantile of the distribution function F of the original random variable X is the limiting value for $K \to \infty$ of $\hat{F}_k^{-1}(1-q; \hat{\gamma}_k^*; K)$.

We have executed equivalent replacements mutatis mutandis for the q-quantile and we obtained the modified q-quantile. These results are formulated in Theorem 2.1.

Theorem 2.1 *Suppose that $\frac{m}{kq} \geq 1$ and suppose that the largest and smallest m order statistics are all either on the positive reals or on the negative reals.*
a) By definition the modified $(1-q)$-quantile estimator is

$$\lim_{K\to\infty} \hat{F}_k^{-1}(1-q; \hat{\gamma}_k^*; K) = \hat{F}_{k,\mathrm{MDEH}}^{-1}(1-q),$$

where

$$\hat{F}^{-1}_{k,\mathrm{MDEH}}(1-q) := X_{(k-m)} + D\left(\frac{m}{kq}\right)\frac{1}{m}\sum_{j=1}^{m}(X_{(k-j+1)} - X_{(k-m)}), \tag{2.6}$$

with $0 < q < \frac{1}{2}$,

$$D\Big(\frac{m}{kq}\Big) = \frac{(\frac{m}{kq})^{G_k} - 1}{G_k}(1 - (G_k \wedge 0)), \tag{2.7}$$

$$G_k = 1 - \frac{1}{2(1-Q_k)},$$

and

$$Q_k = \frac{\left(\frac{1}{m}\sum_{j=1}^{m}(X_{(k-j+1)} - X_{(k-m)})\right)^2}{\frac{1}{m}\sum_{j=1}^{m}(X_{(k-j+1)} - X_{(k-m)})^2}.$$

b) By definition the modified q-quantile estimator is

$$\lim_{K\to\infty} \hat{F}^{-1}_{k}(q;\bar{\gamma}^*_k;K) = \hat{F}^{-1}_{k,\mathrm{MDEH}}(q),$$

where

$$\hat{F}^{-1}_{k,\mathrm{MDEH}}(q) := X_{(m+1)} + \bar{D}\left(\frac{m}{kq}\right)\frac{1}{m}\sum_{j=1}^{m}(X_{(j)} - X_{(m+1)}), \tag{2.8}$$

with $0 < q < \frac{1}{2}$,

$$\bar{D}\Big(\frac{m}{kq}\Big) = \frac{(\frac{m}{kq})^{\bar{G}_k} - 1}{\bar{G}_k}(1 - (\bar{G}_k \wedge 0)), \tag{2.9}$$

$$\bar{G}_k = 1 - \frac{1}{2(1-\bar{Q}_k)},$$

and

$$\bar{Q}_k = \frac{\left(\frac{1}{m}\sum_{j=1}^{m}(X_{(j)} - X_{(m+1)})\right)^2}{\frac{1}{m}\sum_{j=1}^{m}(X_{(j)} - X_{(m+1)})^2}.$$

Proof a) A modified estimator of the $(1-q)$-quantile of the distribution function F of the original random variable X is derived as the limiting value of

$$\hat{F}^{-1}_{k}(1-q;\hat{\gamma}^*_k;K) = X^*_{(k-m)} + \frac{(m/(kq))^{\hat{\gamma}^*_k} - 1}{\hat{\gamma}^*_k}(1 - (\hat{\gamma}^*_k \wedge 0))X^*_{(k-m)}M_k^{(1)^*} - K$$

for $K \to \infty$. First we calculate $\lim_{K\to\infty} \hat{\gamma}^*_n$:

$$\lim_{K\to\infty} \hat{\gamma}^*_k = \lim_{K\to\infty}\left\{M_k^{(1)^*} + 1 - \frac{1}{2}\{1 - \frac{(M_k^{(1)^*})^2}{M_k^{(2)^*}}\}^{-1}\right\}. \tag{2.10}$$

So, we have to consider

$$\frac{(M_k^{(1)^*})^2}{M_k^{(2)^*}} = \frac{\left(\frac{1}{m}\sum_{j=1}^m \log(X^*_{(k-j+1)}/X^*_{(k-m)})\right)^2}{\frac{1}{m}\sum_{j=1}^m \left(\log(X^*_{(k-j+1)}/X^*_{(k-m)})\right)^2}.$$

Now we write

$$\begin{aligned}
\log X^*_{(k-j+1)}/X^*_{(k-m)} &= \log \frac{X_{(k-j+1)}+K}{X_{(k-m)}+K} \\
&= \log(1+\frac{X_{(k-j+1)}-X_{(k-m)}}{X_{(k-m)}+K}) \\
&= \log(1+C_jH),
\end{aligned}$$

where $C_j = X_{(k-j+1)} - X_{(k-m)}$ and $H = 1/(X_{(k-m)}+K)$. Since $K \to \infty$ we have $H \to 0$, and hence

$$\lim_{H\to 0} \frac{\log(1+C_jH)}{H} = C_j.$$

This implies that

$$\lim_{H\to 0} \frac{\left\{\frac{1}{m}\sum_{j=1}^m \log(1+C_jH)\right\}}{H} = \frac{1}{m}\sum_{j=1}^m C_j$$

and

$$\lim_{H\to 0} \frac{\left\{\frac{1}{m}\sum_{j=1}^m \left(\log(1+C_jH)\right)^2\right\}}{H^2} = \frac{1}{m}\sum_{j=1}^m C_j^2.$$

Hence, we find that

$$\lim_{K\to\infty} \frac{(M_k^{(1)^*})^2}{M_k^{(2)^*}} = \frac{(\frac{1}{m}\sum_{j=1}^m C_j)^2}{\frac{1}{m}\sum_{j=1}^m C_j^2}.$$

By substitution of the C_j we get

$$\lim_{K\to\infty} \frac{(M_k^{(1)^*})^2}{M_k^{(2)^*}} = \frac{\left(\frac{1}{m}\sum_{j=1}^m (X_{(k-j+1)}-X_{(k-m)})\right)^2}{\frac{1}{m}\sum_{j=1}^m (X_{(k-j+1)}-X_{(k-m)})^2} = Q_k.$$

Note that $0 < Q_k < 1$ with probability one. Further, we know that

$$\begin{aligned}
\lim_{K\to\infty} M_k^{(1)^*} &= \lim_{K\to\infty} \frac{1}{m}\sum_{j=1}^m \log(X^*_{(k-j+1)}/X^*_{(k-m)}) \\
&= \lim_{K\to\infty} \frac{1}{m}\sum_{j=1}^m \log\left(\frac{X_{(k-j+1)}+K}{X_{(k-m)}+K}\right) = 0.
\end{aligned}$$

Based on these findings we conclude from (2.10) that

$$\lim_{K\to\infty} \hat{\gamma}_k^* = 1 - \frac{1}{2(1-Q_k)} = G_k, \tag{2.11}$$

with $G_k \in (-\infty, \frac{1}{2})$ with probability one. Using (2.11) we obtain that

$$\frac{(m/(kq))^{\hat{\gamma}_k^*} - 1}{\hat{\gamma}_k^*}(1 - (\hat{\gamma}_k^* \wedge 0)) \to D\Big(\frac{m}{kq}\Big),$$

where (cf. (2.7))

$$D\Big(\frac{m}{kq}\Big) = \frac{(\frac{m}{kq})^{G_k} - 1}{G_k}(1 - (G_k \wedge 0)).$$

It is supposed that $\frac{m}{kq} \geq 1$, so $D(\frac{m}{kq}) \geq 0$ and bounded, with probability one. Now we have

$$\begin{aligned}
&\lim_{K\to\infty} \hat{F}_k^{-1}(1-q; \hat{\gamma}_k^*; K) \\
&\quad = \lim_{K\to\infty} \left\{ X_{(k-m)}^* + \frac{(m/(kq))^{\hat{\gamma}_k^*} - 1}{\hat{\gamma}_k^*}(1 - (\hat{\gamma}_k^* \wedge 0)) X_{(k-m)}^* M_k^{(1)^*} - K \right\} \\
&\quad = X_{(k-m)} + \lim_{K\to\infty} \frac{(m/(kq))^{\hat{\gamma}_k^*} - 1}{\hat{\gamma}_k^*}(1 - (\hat{\gamma}_k^* \wedge 0)) X_{(k-m)}^* M_k^{(1)^*}.
\end{aligned} \tag{2.12}$$

In order to evaluate (2.12), we focus on $X_{(k-m)}^* M_n^{(1)^*}$. Using the same notation for C_j and H as before, we derive that

$$\begin{aligned}
\lim_{K\to\infty} X_{(k-m)}^* M_k^{(1)^*} &= \lim_{K\to\infty} \frac{1}{m} \sum_{j=1}^{m} (X_{(k-m)} + K) \log\left(\frac{X_{(k-j+1)} + K}{X_{(k-m)} + K}\right) \\
&= \lim_{H\to 0} \frac{1}{m} \sum_{j=1}^{m} \frac{\log(1 + C_j H)}{H} \\
&= \frac{1}{m} \sum_{j=1}^{m} C_j.
\end{aligned}$$

Hence, if $K \to \infty$ then $\hat{F}_k^{-1}(1-q; \hat{\gamma}_k^*; K)$ converges to

$$\hat{F}_{k,\mathrm{MDEH}}^{-1}(1-q) := X_{(k-m)} + D\left(\frac{m}{kq}\right) \frac{1}{m} \sum_{j=1}^{m} (X_{(k-j+1)} - X_{(k-m)}),$$

with $D(\frac{m}{kq})$ as in (2.7).
b) The proof for the quantiles of small order runs analogously. □

Remark 2.1 A necessary restriction in Theorem 2.1 is that $\frac{m}{kq} \geq 1$, otherwise $D(\frac{m}{kq})$ and $\bar{D}(\frac{m}{kq})$ in (2.7) and (2.9) do not converge for all G_k and $\bar{G}_k$ in $(-\infty, \frac{1}{2})$, respectively. This restriction follows to a certain extent from Dekkers et al. (1989). They distinguish two situations:
1. finite case (Theorem 4.1 in Dekkers et al. (1989)): the number m of order statistics is fixed, $m > c$ with $c \in (0, \infty)$ and suppose that $q = q_k \to 0$, $kq_k \to c$, for $k \to \infty$;
2. infinite case (Theorem 5.1 in Dekkers et al. (1989)): the number m of order statistics is not fixed. Suppose $q_k \to 0,\ kq_k \to \infty\ (k \to \infty)$ and $m(k) := kq_k$.
Further De Haan and Rootzen (1993) reported that if $q < \frac{1}{k}$ nothing can be done unless one imposes extra conditions on F. So $kq > 1$. This restriction in combination with the two situations mentioned implies the prerequisite $\frac{m}{kq} \geq 1$. De Haan and Rootzen (1993) made extra assumptions to analyse large quantile estimators if $q < \frac{1}{k}$.

Remark 2.2 In Section 2.2.4 the EVT is used to estimate control limits. The upper and lower control limit can be regarded as a large and a small quantile, respectively. Suppose we have $X_1, \ldots, X_k$ from a uniform distribution on $[0, 1]$. The control limits based on the estimators by Dekkers et al. (1989) are defined by (see (2.5) respectively (2.4))

$$\begin{aligned} \mathrm{LCL_{EV}} &= \hat{F}_k(q, \hat{\gamma}_k), \\ \mathrm{UCL_{EV}} &= \hat{F}_k(1-q, \hat{\gamma}_k). \end{aligned}$$

The control limits are defined for a false alarm risk $q = 0.00135$. This false alarm risk gives an Average Run Length (ARL) of around 370. The control limits based on the modified estimators are defined by (see (2.8) respectively (2.6))

$$\begin{aligned} \mathrm{LCL_{MDEH}} &= \hat{F}_{n,\mathrm{MDEH}}(q, \hat{\gamma}_n), \\ \mathrm{UCL_{MDEH}} &= \hat{F}_{n,\mathrm{MDEH}}(1-q, \hat{\gamma}_n). \end{aligned}$$

To illustrate the difference for both methods, we sample 10,000 observations from a uniform distribution on $[0, 1]$. Then we calculate the control limits based on both methods. We repeat this procedure 10,000 times and calculate the average of the control limits. For the control limits based on the estimators of Dekkers et al. (1989) we get $\mathrm{LCL_{EV}} = 0.000675$ and $\mathrm{UCL_{EV}} = 0.998617$. For the modified method we get $\mathrm{LCL_{MDEH}} = 0.001388$ and $\mathrm{UCL_{MDEH}} = 0.998617$. We see that the results for both UCLs are the same and are close to the theoretical value 0.99865. However, we see the $\mathrm{LCL_{EV}}$ differs substantially from the theoretical value of 0.00135. The $\mathrm{LCL_{MDEH}}$ is close to this value. Moreover, the $\mathrm{LCL_{MDEH}}$ is symmetric around 0.5 compared with the $\mathrm{UCL_{MDEH}}$.

Conclusion

In this appendix the large and small quantile estimators of Dekkers et al. (1989) are evaluated when a shift of location is applied. If the data are translated and the large and small quantiles

are calculated, the reverse translated quantiles converge, and modified quantile estimators arise. The advantage of these new quantile estimators is that these are symmetric around the mean for a symmetric distribution, e.g. the uniform distribution. This is not the case for the estimator proposed by Dekkers et al. (1989). In Theorem 2.1 one can also calculate quantiles of negative values. Dekkers et al. (1989) considered only positive values and, if not, they advise to apply a simple shift. However, a shift does change the outcomes of the estimator: the estimator is not location invariant. Our modified estimator is resistant in this respect; it is location equivariant. The modified estimator may be applied for establishing control limits.

Chapter 3

A semi-Bayesian method for Shewhart individual control charts

Shewhart control limits for individual observations are traditionally based on the average of the moving ranges. The performance of this control chart behaves quite well if the underlying distribution is normal and the sample size is greater than 250. Under non-normality it is recommended to use control charts based on non-parametric statistics (cf. Chapter 2). The drawback of these individual control charts is that at least 1,000 observations are needed to obtain appropriate results. In this chapter we propose an alternative individual control chart which behaves quite well under non-normality for moderate sample sizes in the range of 250 through 1,000 observations. To apply this control chart one starts with an initial guess for the density function of the characteristic under study. Based on this initial guess and the observed data a density function can be derived by means of an approximation with Bernstein polynomials. The in-control and out-of-control performance of the proposed control chart and the traditional control charts are studied by simulation.

If the initial guess is appropriate, then for non-normal data and moderate sample sizes in the order of 250 through 1,000 observations, the new method performs better than the individual control charts based on the average of the moving ranges or based on non-parametric statistics. So for these sample sizes we have tried to close the gap. This chapter is based on Vermaat and Does (2006).

3.1 Introduction

For sample sizes till 1,000 there is no good alternative under non-normality (cf. Quesenberry (1993) and Vermaat et al. (2003)), unless a sufficient amount of additional information is available of the distribution function F. In this chapter we propose a new method for the estimation of the control limits for moderate sample sizes (i.e. for sample sizes in the range of 250 through 1,000). This method makes use of an initial guess of the underlying distribution

function and Bernstein polynomials. The Bernstein polynomial is a linear function of order statistics with smooth weight functions and may be used to estimate quantile functions (cf. Perez and Palacin (1987)). In De Bruin et al. (1999) Bernstein polynomials are used for non-parametric density estimation. In Albers and Schaafsma (2003) some improvements are studied. We compare the individual control charts based on Bernstein polynomials with the traditional Shewhart control chart for individuals with control limits based on the average of the moving ranges and with the empirical quantile control chart studied in Chapter 2, Section 2.2.2, and in Vermaat et al. (2003). The statistical performance of these methods will be studied by an extensive simulation study. Much depends on the reliability of the initial guess, especially in the tails.

The chapter is organized as follows. In the next section the different methods for the estimation of the control limits are given. We illustrate the different methods by means of the real life example about the groove's depth introduced in Chapter 1. To compare the different methods a simulation study is presented. Finally, we conclude with our findings.

3.2 Description of the control charts

3.2.1 The traditional individual control chart based on moving ranges

The traditional Shewhart individual control chart has control limits defined by

$$\mathrm{UCL} = \mu + \Phi^{-1}(1 - \frac{\alpha}{2})\sigma,$$

$$\mathrm{LCL} = \mu + \Phi^{-1}(\frac{\alpha}{2})\sigma,$$

where Φ^{-1} is the standard normal quantile function, and where μ is the mean and σ is the standard deviation of the normal distribution function F. Level α is the false alarm rate (e.g. for $\alpha = 0.0027$ we obtain $\Phi^{-1}(1 - \frac{\alpha}{2}) = 3$). In practice μ and σ are unknown. However, we assume that a Phase I sample $X_1, \ldots, X_k$ of independent and identically distributed random variables is available to estimate μ and σ. Recall from Woodall and Montgomery (1999), that Phase I refers to the retrospective analysis phase and Phase II refers to the monitoring phase. Estimators of μ and σ are usually the sample mean $\bar{X}_k = \sum_{i=1}^{k} X_i/k$ and the average of the moving ranges

$$\overline{\mathrm{MR}}_k = \frac{1}{k-1}\sum_{i=2}^{k} |X_i - X_{i-1}|$$

respectively, see Duncan (1986). The last estimator is scaled by $d_2(2) = 2/\sqrt{\pi}$ to obtain an unbiased estimator for σ under normality. With these estimations for μ and σ we obtain the

traditional Shewhart individual control chart with control limits

$$\widehat{\mathrm{UCL}} = \bar{X}_k + \Phi^{-1}(1-\frac{\alpha}{2})\frac{\overline{\mathrm{MR}}_k}{d_2(2)},$$
$$\widehat{\mathrm{LCL}} = \bar{X}_k + \Phi^{-1}(\frac{\alpha}{2})\frac{\overline{\mathrm{MR}}_k}{d_2(2)}. \tag{3.1}$$

3.2.2 The exact average moving range control chart

In Roes et al. (1993) a more exact version of the traditional Shewhart individual control chart has derived. With $V_k = \overline{\mathrm{MR}}_k/d_2(2)$ they approximate the distribution of $W_k^2 = (V_k/\sigma)^2$ by $\tau^2\chi^2(\nu)/\nu$, where

$$\tau = \sqrt{\mathrm{Var}(W_k)+1},$$
$$\nu = \frac{1}{2}\Big(1+\frac{1}{\mathrm{Var}(W_k)}\Big),$$

and $\chi^2(\nu)$ is the chi-square distribution with ν degrees of freedom. Using this approximation it follows that for $t > k$, $(X_t - \bar{X}_k)/(\overline{\mathrm{MR}}_k/d_2(2))$ is approximately distributed as $\frac{\sqrt{1+1/k}}{\tau}$ times a Student's t-distribution with ν degrees of freedom. The Average Moving Range (AMR) control chart has control limits defined by

$$\widehat{\mathrm{UCL}}_{\mathrm{AMR}} = \bar{X}_k + \frac{\sqrt{1+1/k}}{\tau}t(1-\frac{\alpha}{2};\nu)\frac{\overline{\mathrm{MR}}_k}{d_2(2)}, \tag{3.2}$$

$$\widehat{\mathrm{LCL}}_{\mathrm{AMR}} = \bar{X}_k + \frac{\sqrt{1+1/k}}{\tau}t(\frac{\alpha}{2};\nu)\frac{\overline{\mathrm{MR}}_k}{d_2(2)}, \tag{3.3}$$

where $t(p,\nu)$ denotes the p-quantile of a t-distribution with ν degrees freedom. This AMR control chart is more accurate than the traditional Shewhart individual control chart in (3.1): i.e. the rates of the traditional Shewhart individual control chart of falsely signalling an out-of-control situation are much larger than intended, see Roes (1995).

3.2.3 Empirical quantile control chart

In this subsection the control limits of the empirical quantile control chart are defined as in Vermaat et al. (2003). A natural estimator of the q-quantile of the distribution function F is the empirical quantile $\hat{F}_k^{-1}(q)$, which is defined as

$$\hat{F}_k^{-1}(q) = \inf\{x|\ \hat{F}_k(x) \geq q\},\ \ 0 < q < 1,$$

where $\hat{F}_k$ is the empirical distribution function that puts mass $1/k$ at each X_i, $1 \leq i \leq k$, i.e.

$$\hat{F}_k(x) = \frac{1}{k}\sum_{i=1}^{k} I_{\{X_i \leq x\}}, \quad -\infty < x < \infty,$$

with I the indicator function, i.e. $I_{\{x \leq y\}}$ equals 1 if $x \leq y$ holds and 0 otherwise. Hence, an obvious estimator of the upper control limit based on the empirical quantile (EQ) is

$$\widehat{\mathrm{UCL}}_{\mathrm{EQ}} = \hat{F}_k^{-1}(1 - \frac{\alpha}{2}) = X_{(\lceil (1-\frac{\alpha}{2})k \rceil)}, \tag{3.4}$$

with $X_{(1)} \leq X_{(2)} \leq \cdots \leq X_{(k)}$ denoting the order statistics of the initial sample $X_1, \ldots, X_k$ and $\lceil y \rceil$ the smallest integer not smaller than y. The lower control limit estimated by the empirical quantile is defined by

$$\widehat{\mathrm{LCL}}_{\mathrm{EQ}} = X_{(\lfloor \frac{\alpha}{2}k+1 \rfloor)}, \tag{3.5}$$

where $\lfloor y \rfloor$ denotes the largest integer not larger than y.

3.2.4 A control chart based on a Bernstein approximation

In this subsection we introduce a semi-Bayesian method for estimating a density function f. From this we derive the corresponding cumulative distribution function. This method is introduced in De Bruin et al. (1999) and Albers and Schaafsma (2003). For technical details we refer to Appendix 3A. Moreover, in Appendix 3B you will find an example how the method works in practice.

Assume that the density function f is continuous and strictly positive on the interval (a, b) and that F is the corresponding cumulative distribution function. We choose an a priori density function ψ as an initial guess for the density function f. Based on the corresponding continuous cumulative distribution function Ψ, the order statistics $X_{(1)}, \ldots, X_{(k)}$ are transformed on $[0, 1]$, by $Y = \Psi(X)$, such that

$$Y_{(0)} = 0, \quad Y_{(i)} = \Psi(X_{(i)}), \quad i = 1, \cdots, k, \quad Y_{(k+1)} = 1.$$

Because $\mathrm{P}(Y \leq y) = \mathrm{P}(\Psi(X) \leq y) = \mathrm{P}(X \leq \Psi^{-1}(y)) = F(\Psi^{-1}(y))$ we have that $B = (F(\Psi^{-1}))^{-1}$ is the quantile function of the random variable $Y = \Psi(X)$. This quantile function is estimated by the so-called Bernstein polynomials $B_k^{(m)}(p)$, where m is a smoothing parameter and $0 < p < 1$, see Appendix 3A. It follows that $\hat{F}_{k,BA}^{(m)} = B_k^{(m)^{-1}}(\Psi)$ is an estimate for F.

As described in De Bruin et al. (1999), this method is not a full-Bayesian approach, where a initial guess ψ would be a prior for one value of X, whereas we use ψ for all values of X.

Other polynomials to estimate the quantile function can be used as well. In Perez and Palacin (1987) both the Kantorowitz polynomials and the Bernstein polynomials were discussed for estimating the quantile function. In this chapter we restrict ourselves to the Bernstein polynomials.

The estimators for the upper and lower control limits based on the Bernstein approximation (BA) with a false alarm rate α are now given by

$$\widehat{\mathrm{UCL}}_{\mathrm{BA}} = \hat{F}_{k,BA}^{(m)^{-1}}(1-\frac{\alpha}{2}), \tag{3.6}$$

$$\widehat{\mathrm{LCL}}_{\mathrm{BA}} = \hat{F}_{k,BA}^{(m)^{-1}}(\frac{\alpha}{2}), \tag{3.7}$$

where $\hat{F}_{k,BA}^{(m)^{-1}}$ is defined by

$$\hat{F}_{k,BA}^{(m)^{-1}} = \Psi^{-1}(B_k^{(m)}(p)),$$

cf. Equations (3.10) and (3.11) in Appendix 3A. The control chart based on these control limits is denoted as the BA control chart.

The method described above needs an initial guess ψ. In principle, the initial guess has to be chosen a priori. This assumption is not necessarily unrealistic, cf. Albers (2003). Based on process knowledge we can make an initial guess for the distribution of the process characteristic. The described method is to 'fine tune' the initial guess.

For the Bernstein approximation the smoothing parameter m has to be chosen. In Appendix 3A it is shown that we may take $m = 5.2\sqrt{k}$ and round this off to the nearest integer.

3.3 Real life example

The Bernstein approximation control chart seems to be rather cumbersome. However, with the use of computers and software, methods like this can be used quite easily by practitioners. To show this we apply the Bernstein approximation control chart to the real life example introduced in Chapter 1.

The collected data concern the part of a printer that squirts the ink. The ink is spurted through a groove. The relevant quality characteristic to be measured is the depth of the groove. We will study 535 observations of this depth. For reasons of confidentiality the measurements are multiplied by a constant. By studying the data, we found out that for 15 measurements the grooves were cut by a supplier who uses a different method to cut the grooves. These problems were detected and eliminated. Eliminating these data points results in a stable process. We will study the remaining data set, which consists of 520 measurements. For the calculation of the control limits we use the estimators described in the previous section. We found that the control limits based on the average of the moving ranges are (cf. Equations (3.2) and (3.3))

$$\widehat{\mathrm{UCL}}_{\mathrm{AMR}} = 0.210,$$

$$\widehat{\mathrm{LCL}}_{\mathrm{AMR}} = 0.148,$$

and for the empirical quantiles method we obtain (cf. Equations (3.4) and (3.5))

$$\widehat{\mathrm{UCL}}_{\mathrm{EQ}} = 0.201,$$

$$\widehat{\mathrm{LCL}}_{\mathrm{EQ}} = 0.131.$$

To obtain the control chart based on Bernstein polynomials we use the normal distribution as initial guess to estimate the control limits (cf. Equations (3.6) and (3.7)). From Equations (3.10) and (3.11) it is clear that we have to estimate the parameters of the initial guess and the smoothing parameter m. Because we have chosen the normal distribution as initial guess its parameters (μ and σ) can be easily estimated by the sample mean and sample standard deviation of the 520 observations. Because $5.2\sqrt{k} = 118.6$, we take $m = 119$. The computations provide that the BA control limits are

$$\widehat{\mathrm{UCL}}_{\mathrm{BA}} = 0.207,$$

$$\widehat{\mathrm{LCL}}_{\mathrm{BA}} = 0.136.$$

Figure 3.1 shows that the distribution of the data is slightly skewed to the left. The autocorrelogram of the data in Figure 3.2 indicates that the data are not significantly correlated. In Figure 3.3 the data with the different control limits have been drawn in one graph.

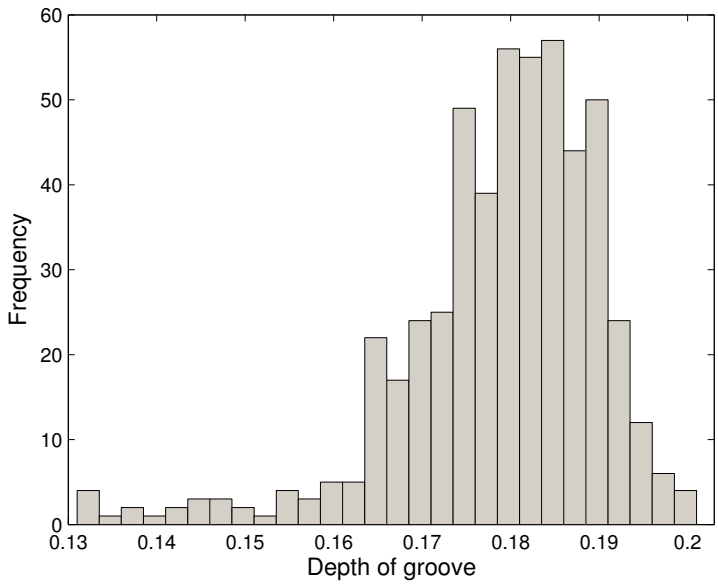

Figure 3.1: The histogram for the depth of groove data.

Note that because we use $\alpha = 0.0027$ and $n = 520$ the EQ control chart takes the largest and smallest observation as upper control limit and lower control limit, respectively. Hence, all observations are between these two control limits. The AMR control chart shows 15 out-of-control signals and its control limits are symmetric with respect to the center line. The BA control chart signals only 5 out-of-control signals.

In Phase II (monitoring phase) the control limits may be updated regularly as new data become available. This is done in the same way as in Phase I. In Chapter 5 of Does et al. (1999) an activity plan is given for the implementation and maintenance of a control chart.

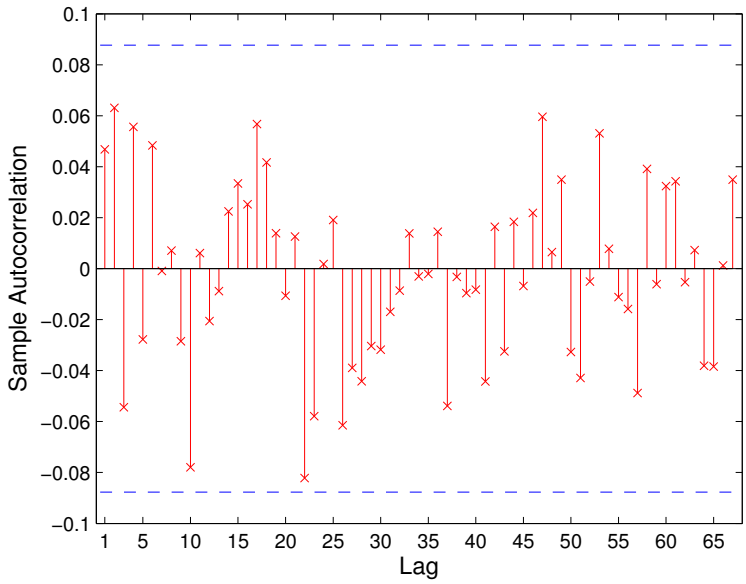

Figure 3.2: The autocorrelogram for the depth of groove series (with 5% significance limits for the autocorrelation).

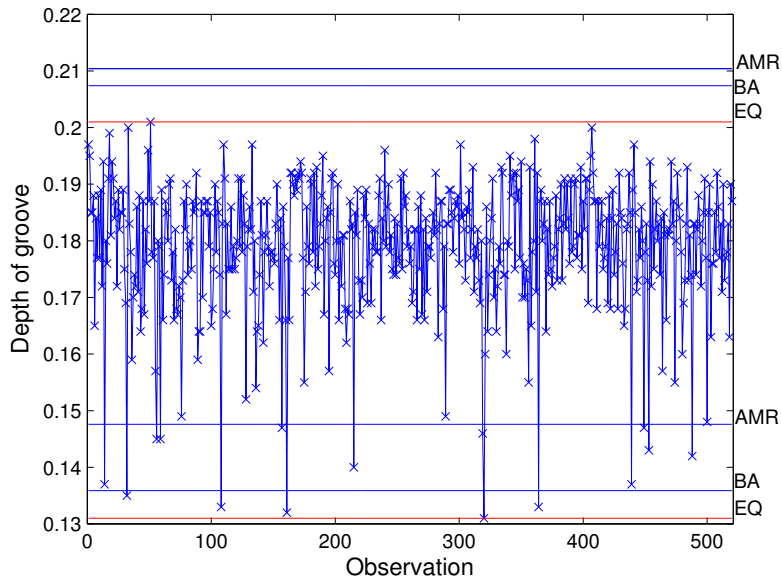

Figure 3.3: Depth of groove data and control limits based on the average moving range (AMR), the empirical quantiles (EQ), and the Bernstein approximation (BA). CL is the center line and represents the process average.

3.4 Design of the simulation study

In order to study the performance of the control charts introduced in the previous sections we have conducted an extensive simulation study for some choices of the distribution function F

and the sample size k. Based on the Phase I training sample $X_1, \ldots, X_k$ the UCL and LCL are estimated. The performance of each control chart is measured by calculating the average and standard deviation of the run length – ARL and SDRL respectively – in Phase II. We follow Vermaat et al. (2003) in the calculations of the ARL and SDRL. They denote by $\hat{p}_k(X_1, \ldots, X_k)$ the conditional probability given the training sample, that a new independent random variable X from the same distribution F exceeds the upper control limit or is below the lower control limit. The average run length is

$$\mathrm{ARL} = \mathrm{E}\frac{1}{\hat{p}_k(X_1, \ldots, X_k)}$$

and the standard deviation of the average run length

$$\mathrm{SDRL} = \sqrt{2\mathrm{E}\left(\frac{1}{\hat{p}_k(X_1, \ldots, X_k)}\right)^2 - \left(\mathrm{E}\frac{1}{\hat{p}_k(X_1, \ldots, X_k)}\right)^2 - \mathrm{E}\frac{1}{\hat{p}_k(X_1, \ldots, X_k)}}. \quad (3.8)$$

Since these ARL and SDRL cannot be computed directly, 10,000 training samples $x_1, \ldots, x_k$ are drawn. For each training sample we calculate $1/\hat{p}(x_1, \ldots, x_k)$ and $1/\hat{p}(x_1, \ldots, x_k)^2$. Averaging $1/\hat{p}(x_1, \ldots, x_k)$ over the 10,000 training samples gives the ARL. In more or less the same way we calculate the SDRL.

We study the control charts in an in-control and several out-of-control situations. For the comparison we distinguish different shifts in the mean of size $\delta\sigma$, where δ ranges from 0 through 5 and σ is the standard deviation of the studied distribution. With $\delta = 0$ we have of course the in-control situation. Given the 10,000 training samples $x_1, \cdots, x_k$ we calculate for each shift the ARL and SDRL. Based on a false alarm rate α equal to 0.0027, which corresponds with the traditional 3σ-limits of Shewhart individual control charts, we expect that the ARL and SDRL are around 370.

The simulations are carried out for six different distributions, i.e. normal, Student's t, logistic, exponential, chi-square, and Weibull distribution and for sample sizes k equal to 250, 500, and 1,000. Seventeen shifts in the mean of size $\delta\sigma$ are used in the range of 0 (0.25) 3.5, 4, and 5.

3.4.1 Simulation design for the Bernstein approximation control chart

As mentioned before, the BA control chart has been developed to 'fine tune' an initial guess for the underlying density function. Based on physical properties of a product and/or on experience it is generally known whether the characteristic under study has approximately a normal distribution or some skewed distribution. In our simulation we use this fact in order to differentiate between a normal and a skewed distribution.

To distinguish two moderate deviations from a normal distribution, we use Student's t-distribution with 30 degrees of freedom and the logistic distribution with scale parameter equal

to 1. The BA control chart performance will be illustrated by using these two distributions and a normal distribution. The initial guess will be in these three cases the normal distribution. Based on the training sample $x_1, \ldots, x_k$ the μ and σ of the normal distribution (i.e. the initial guess distribution) are estimated by the sample quantities of the parameters.

Secondly, we study deviations from a gamma distribution. The initial guess ψ will be a gamma density. The two parameters of this gamma density are estimated from the first two moments of the training sample (cf. Johnson et al. (1994)). For the comparison we use the standard exponential, the chi-square distribution with one degree of freedom, and the Weibull distribution with shape parameter equal to 1 and scale parameter equal to 2.

3.4.2 Discussion of the simulation results

The results of the simulations are given in two complementary figures. In the first figure the ARL is depicted for the different estimation methods. In the second figure the SDRL is depicted. A complete survey of the simulation results may be found in the master thesis of Vermaat (2003).

The results of the simulations are given in Figures 3.4 through 3.12 showing the ARL and SDRL respectively. The shift in the mean is put on the horizontal axes and the ARL (SDRL) on the vertical axes. For sample sizes 250 and 500 only the results of the AMR and the BA control charts are given. For sample size equal to 1,000 the graphs are added supplemented the results of the EQ control chart. For smaller sample sizes its behavior is not satisfactory (cf. Vermaat et al. (2003)). Furthermore, we have added in all figures the theoretical (TH) ARL and SDRL (based on the exact exceedance probability p for the given distribution). This was easy because the exact distribution was known in the simulation study.

In Figure 3.4 the results for the normal distribution function are given for sample size $k = 250$ and with the normal density as initial guess. We see that the BA control chart is closer to the theoretical ARL and SDRL than the AMR control chart. If the sample size k increases, we observe that the performance lines of both control charts converge to the theoretical line (see Figure 3.5). Note that, in Figure 3.5 the control limits of the EQ control chart are equal to the second largest respectively second smallest observation. The performance of all charts is comparable.

In Figure 3.6 the results for the t-distribution with 30 degrees of freedom and sample size $k = 500$ are given. We see that the BA control chart is slightly better than the AMR control chart.

Figure 3.7 shows the results for the logistic distribution. Although the logistic distribution looks like a normal distribution the performance of the AMR control chart and BA control chart are very bad for $k = 250$. This is due to the fact that the tail behavior of the logistic distribution is very different from the tail behavior of a normal distribution. Even for large values of k the behavior of both control charts is bad compared with the theoretical one. For $k = 1{,}000$ we

observe a more regular behavior of the ARL (and SDRL) for the EQ control chart, see Figure 3.8.

In Figures 3.9 and 3.10 we show the results for the exponential distribution function for sample sizes $k = 250$ and $k = 1{,}000$. We see that for a training sample with $k = 1{,}000$ the performance of the BA control chart is much better than for a training sample with $k = 250$. The same pattern holds for a chi-square distribution. This can be explained by the fact that the estimators of the parameters of the initial guess, i.e. the gamma distribution, need a lot of data to be accurate and unbiased, (cf. Johnson et al. (1994), p. 358). More data lead to a more accurate initial guess, which implies a better performance of the BA control chart. We also observe in Figures 3.9 and 3.10, that the BA and theoretical control chart have a maximum in the ARL and SDRL around $\delta = 0.25$. This maximum is due to the fact, that these control charts estimate an LCL within the support of the exponential distribution. If the process shifts towards the UCL when δ increases, the LCL is of course harder to violate and the UCL easier. Since the density of the exponential is larger near the LCL than near the UCL, this causes that the probability of an alarm decreases and hence the ARL increases for small δ. Because the estimates of the LCL in the AMR control chart typically fall outside the support of the exponential distribution, this control chart does not show this phenomenon. Note that the behavior of the AMR control chart for $\delta = 0$ is really bad.

In Figure 3.11 and 3.12 the results for a Weibull distribution have been drawn for sample sizes $k = 250$ and $k = 1{,}000$. We see that the results for both the BA and AMR control charts for this distribution are bad. For the BA control chart this is due to the fact that the initial guess (i.e. a gamma distribution) differs too much from the Weibull distribution as theoretical distribution. In this case only the performance of the EQ control chart is satisfying for the ARL with $k = 1{,}000$.

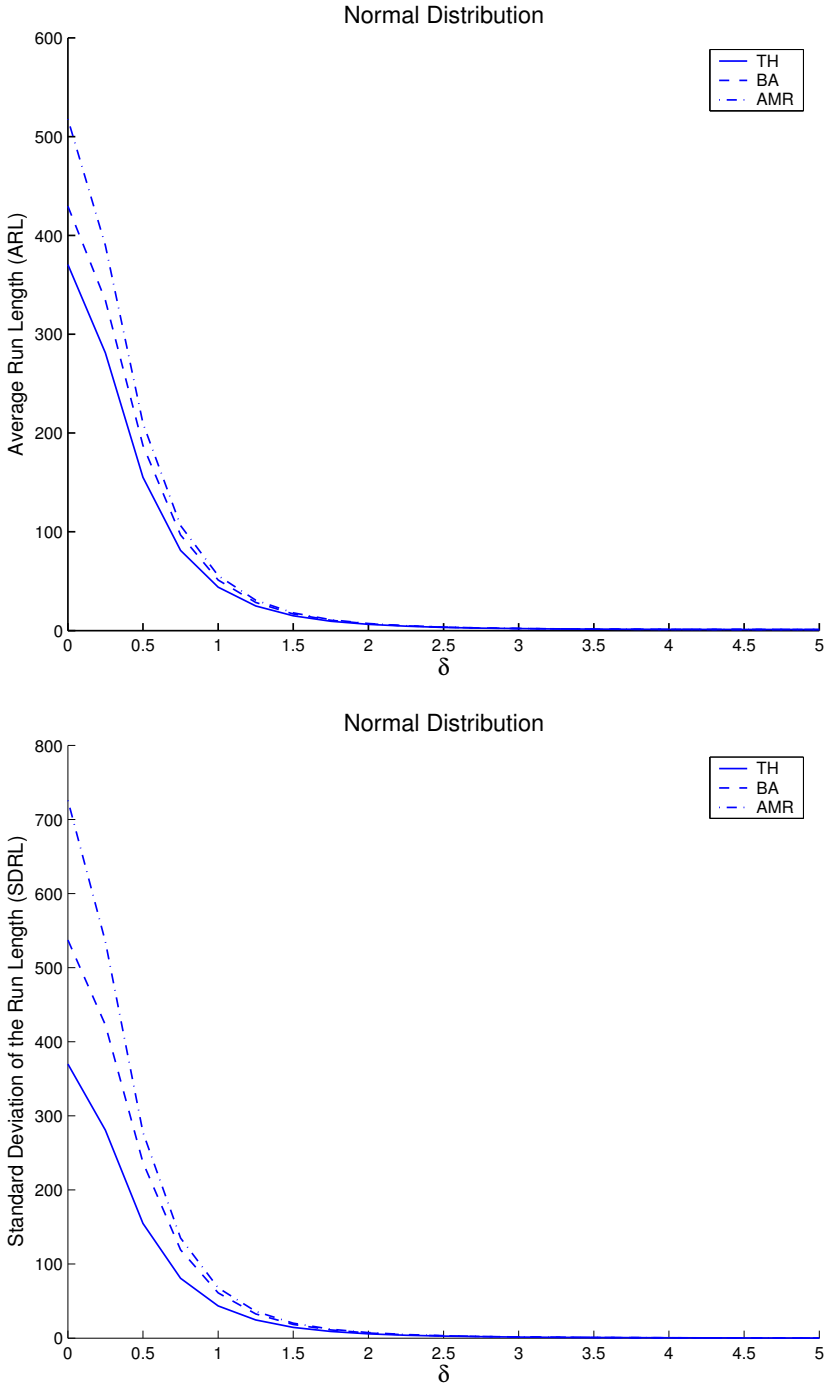

Figure 3.4: The ARL and the SDRL for $k = 250$ based on an underlying normal distribution.

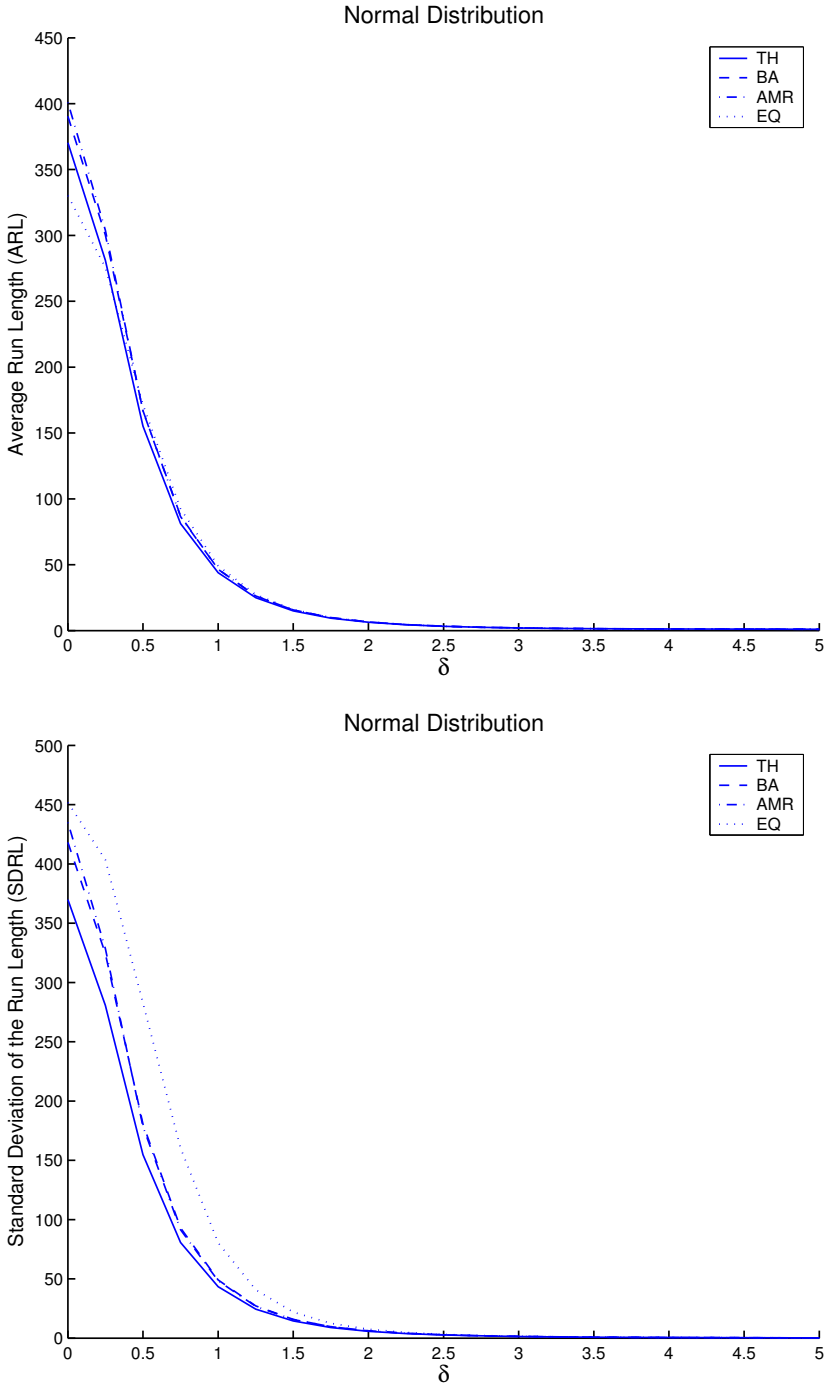

Figure 3.5: The ARL and the SDRL for $k = 1{,}000$ based on an underlying normal distribution.

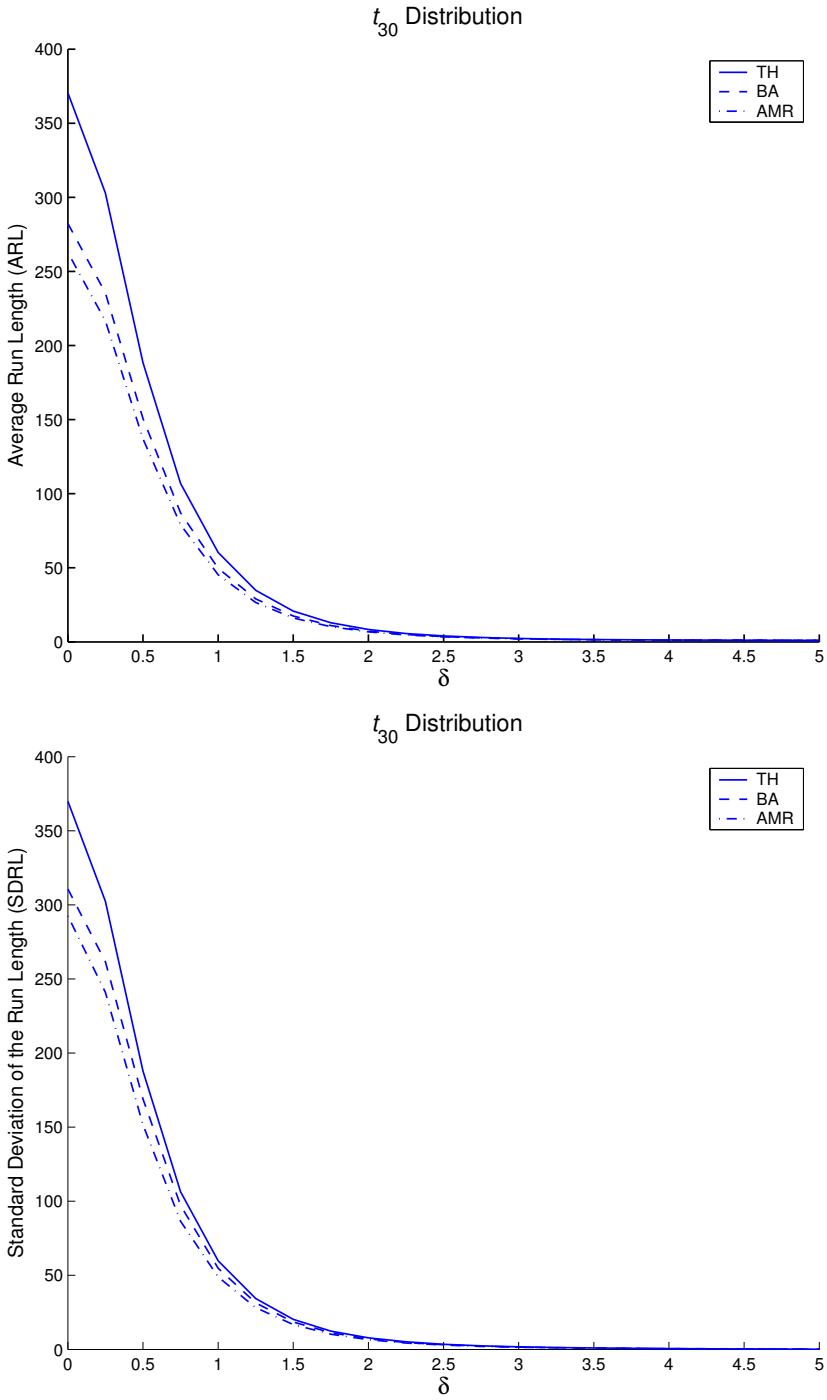

Figure 3.6: The ARL and the SDRL for $k = 500$ based on an underlying t-distribution with 30 degrees of freedom.

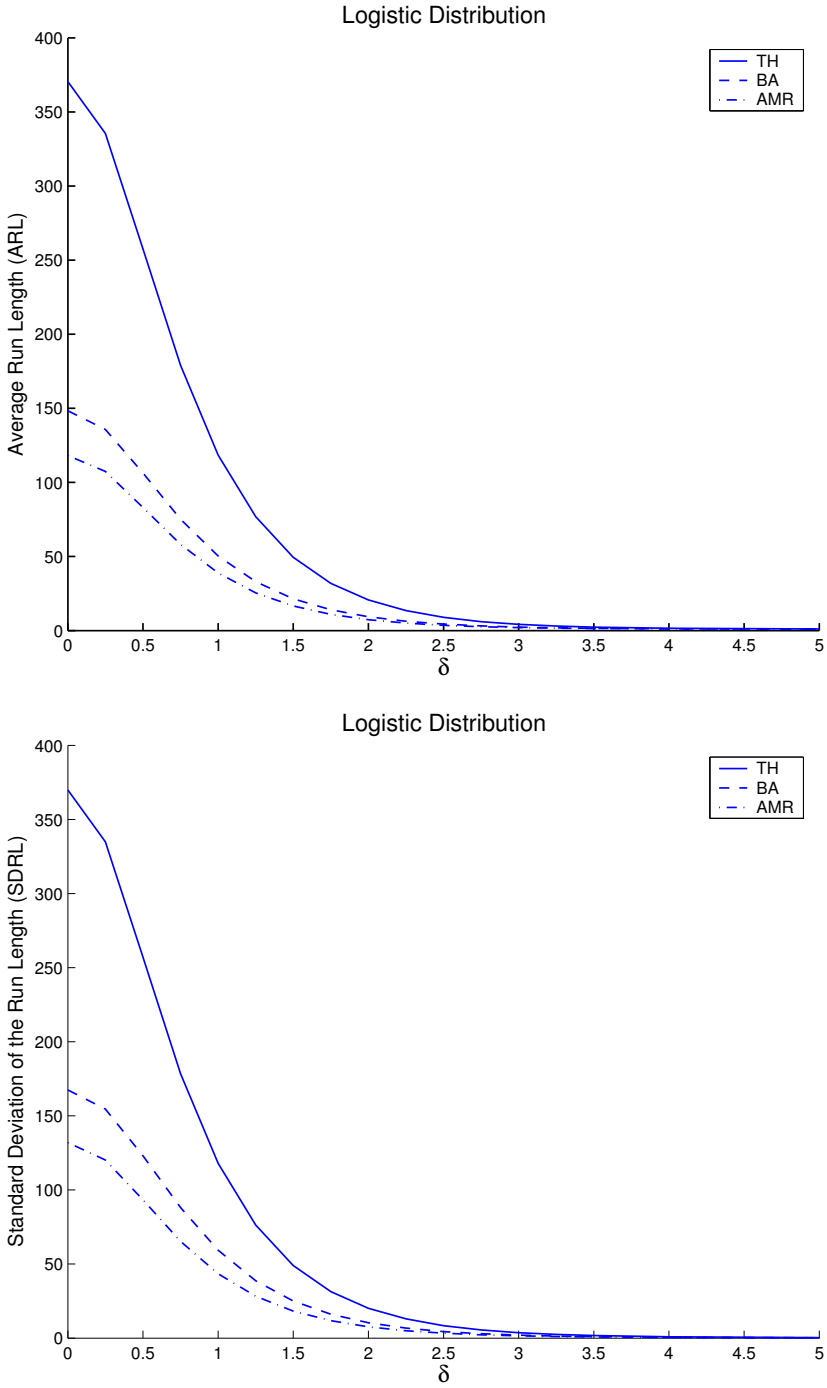

Figure 3.7: The ARL and the SDRL for $k = 250$ based on an underlying logistic distribution.

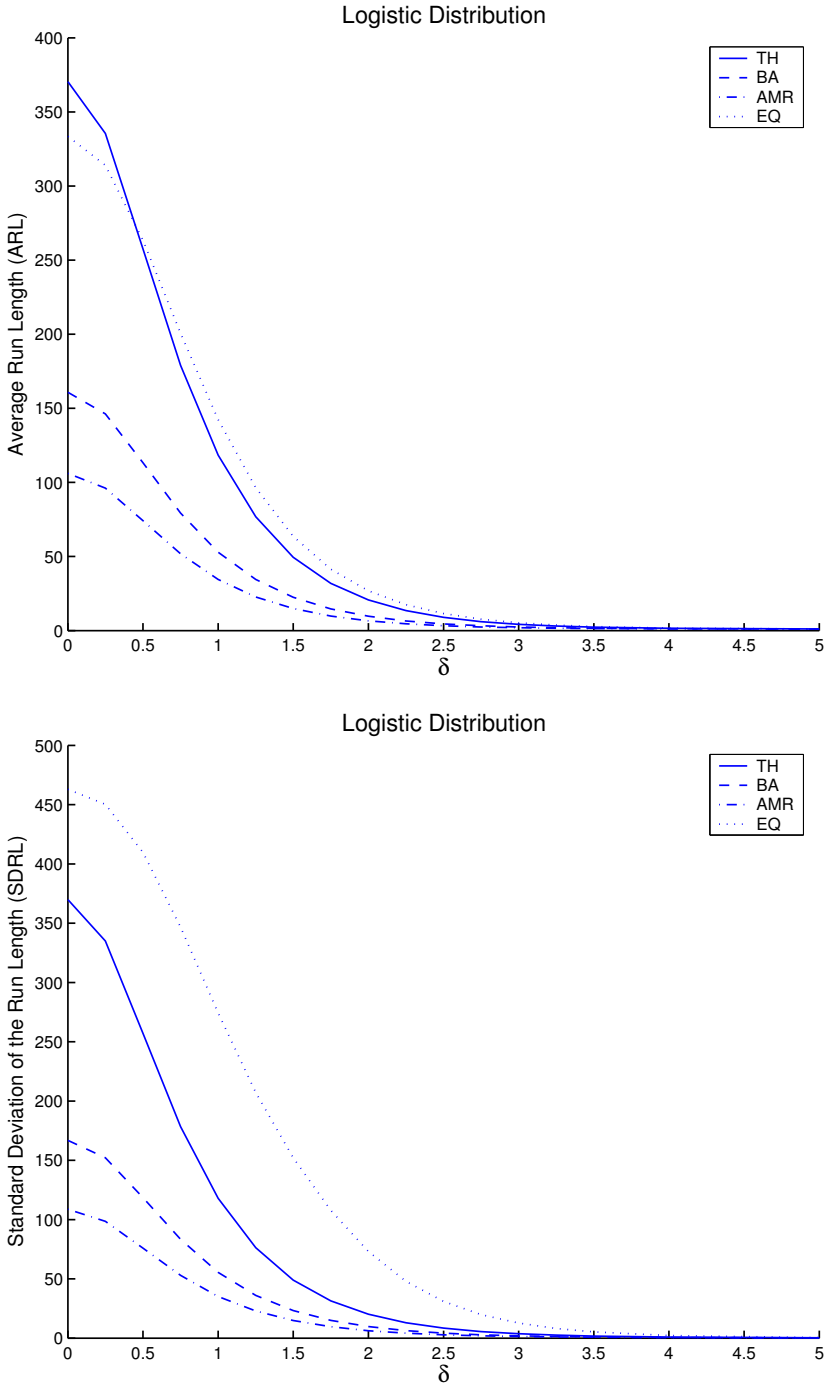

Figure 3.8: The ARL and the SDRL for $k = 1{,}000$ based on an underlying logistic distribution.

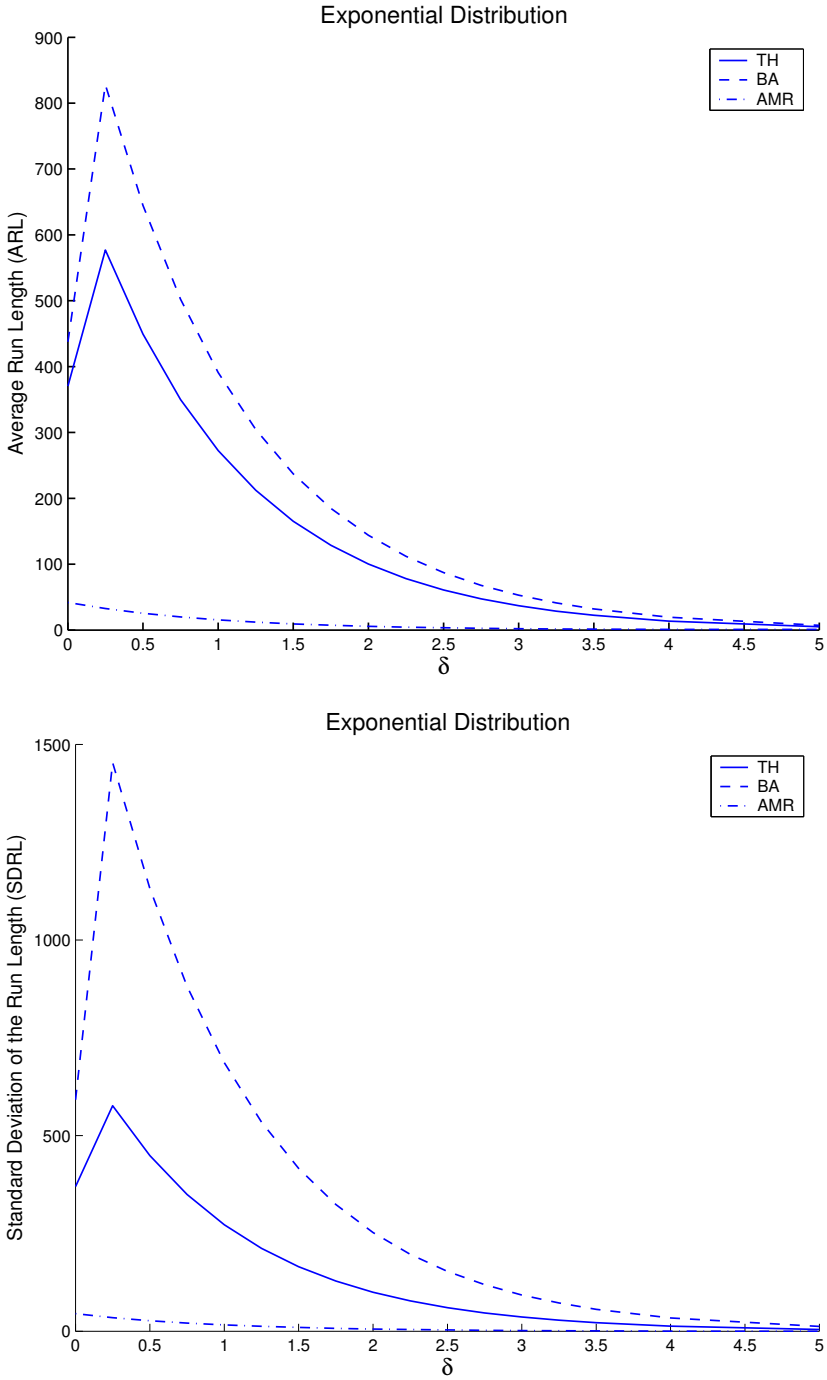

Figure 3.9: The ARL and the SDRL for $k = 250$ based on an underlying exponential distribution.

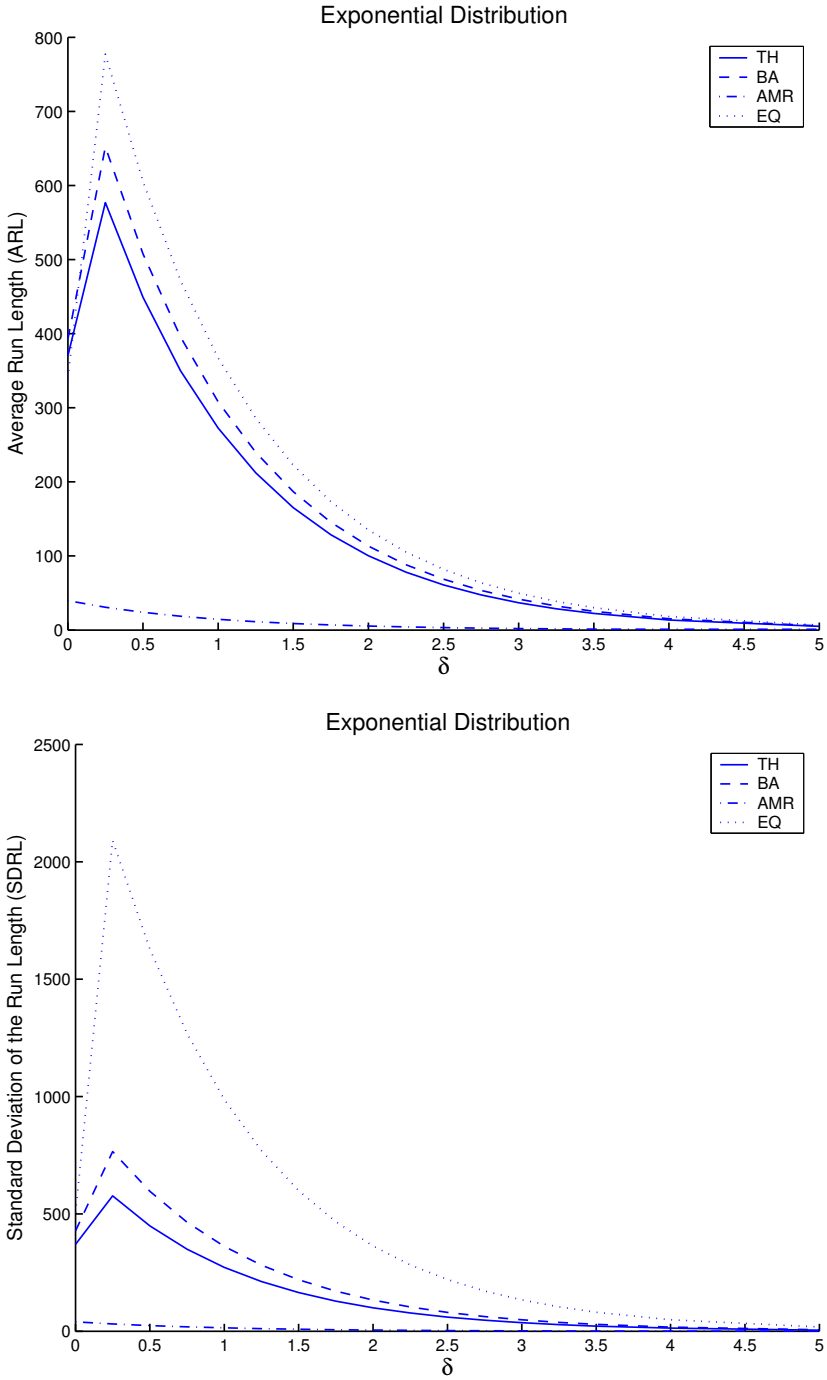

Figure 3.10: The ARL and the SDRL for $k = 1{,}000$ based on an underlying exponential distribution.

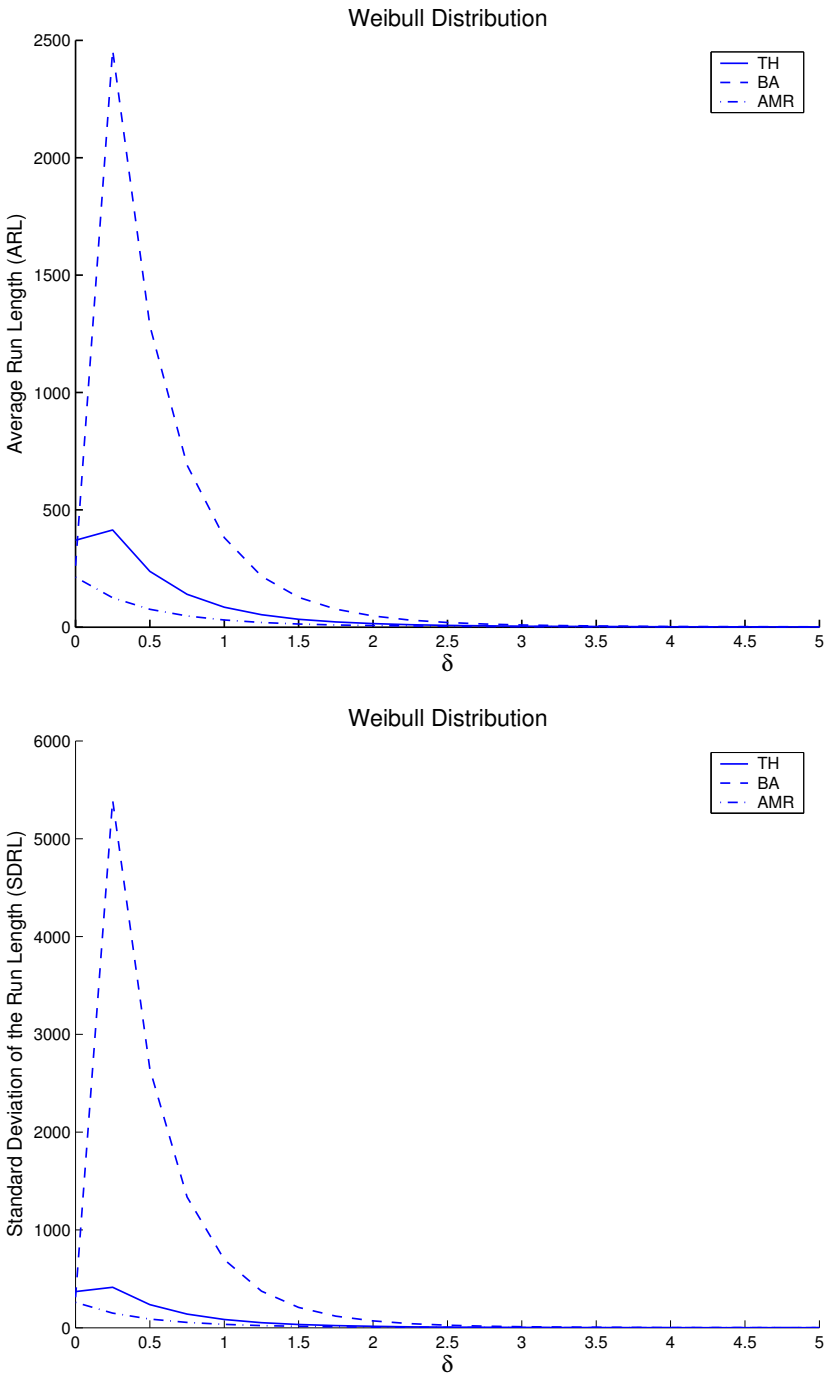

Figure 3.11: The ARL and the SDRL for $k = 250$ based on an underlying Weibull distribution.

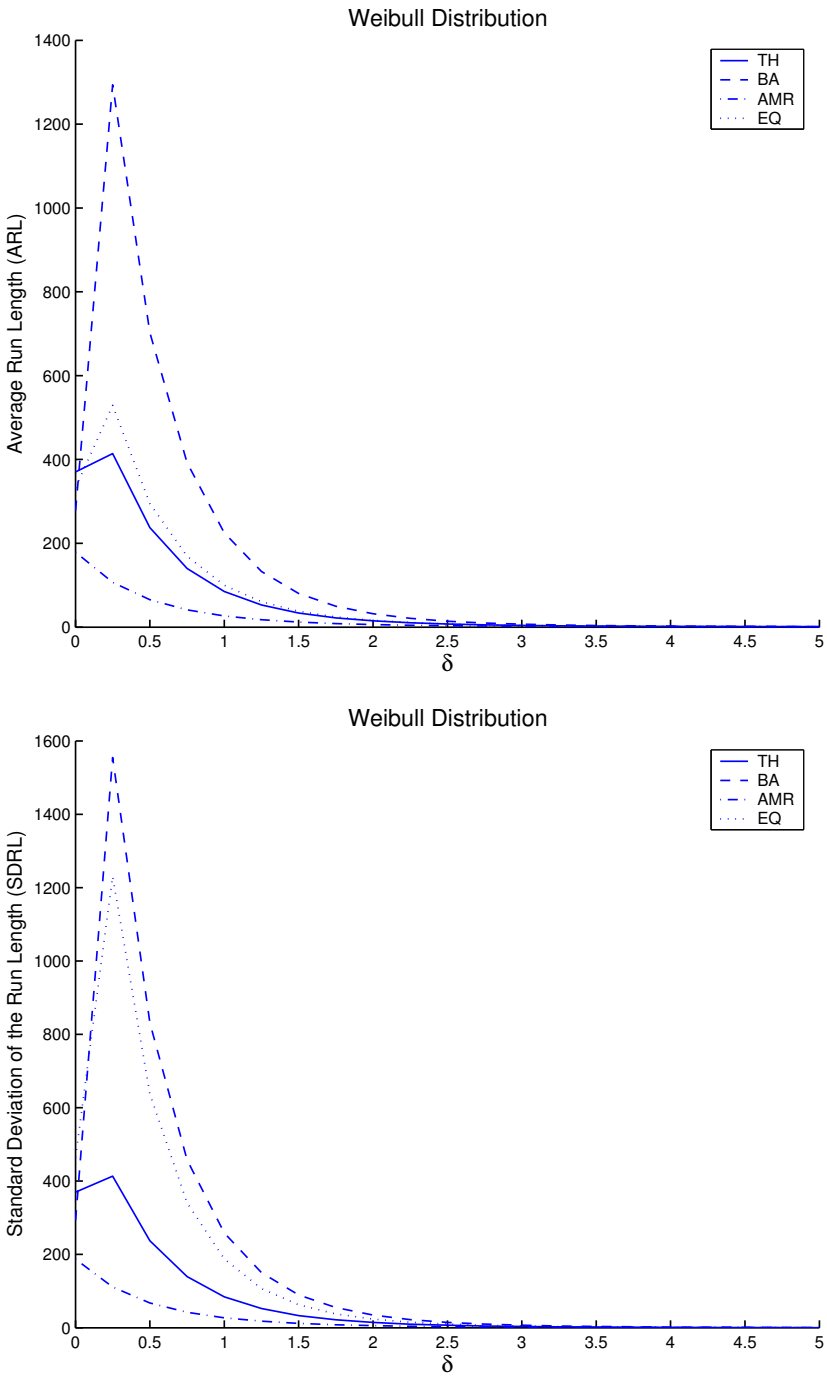

Figure 3.12: The ARL and the SDRL for $k = 1{,}000$ based on an underlying Weibull distribution.

3.5 Conclusion

In this study we have applied a semi-Bayesian method for the estimation of the control limits of a Shewhart individual control chart. The control chart based on the average of the moving ranges performs well for normally distributed observations. If the observations are not normally distributed the behavior becomes quite bad. In this situation non-parametric (e.g. empirical quantiles) control charts perform quite well if they are based on a Phase I sample of at least 1,000 observations. In situations of moderate sample sizes (i.e. 250 to 1,000) we suggest to use the Bernstein approximation method. The purpose of this method is to 'fine tune' an initial guess of the underlying density function. With this method the control chart is robust against small violations of normality. We have also shown that if the initial guess is a reasonable approximation of the underlying distribution, the performance of the new control chart is better than the AMR control chart. If on the other hand the initial guess is a bad approximation of the underlying distribution the results of the new control chart are not good enough, but nevertheless better than the AMR control chart.

Appendix 3A: The Bernstein approximation

Suppose we have a sample $X_1, \ldots, X_k$. Denote the sample order statistics by $X_{(1)} < X_{(2)} < \cdots < X_{(k)}$. Suppose ψ is an a priori density function for f and Ψ the corresponding cumulative a priori distribution function for the cumulative distribution F. Assume that f is continuous and strictly positive on the interval (a, b). Transform the order statistics $X_{(1)}, \ldots, X_{(k)}$ by the transformation $Y = \Psi(X)$, such that

$$Y_{(0)} = 0, \quad Y_{(i)} = \Psi(X_{(i)}), \quad i = 1, \cdots, k, \quad Y_{(k+1)} = 1. \tag{3.9}$$

Because $\mathrm{P}(Y \leq y) = \mathrm{P}(\Psi(X) \leq y) = \mathrm{P}(X \leq \Psi^{-1}(y)) = F(\Psi^{-1}(y))$ we have that $B = (F(\Psi^{-1}))^{-1}$ is the quantile function of the random variable $Y = \Psi(X)$. This quantile function is estimated by the so-called Bernstein polynomial approximation of degree $(k+1)$ defined by

$$B_{k+1}(p|Y_0, \ldots, Y_{k+1}) = \sum_{i=0}^{k+1} Y_{(i)} \binom{k+1}{i} p^i (1-p)^{k+1-i}$$

and $0 < p < 1$. Let $H_k = \Psi^{-1}(B_k)$ then $\hat{F}_{k,BA} = H_k^{-1} = B_k^{-1}(\Psi)$ is used to estimate F.

An improvement of the Bernstein approximation can be established by introducing a smoothing parameter m, cf. Albers and Schaafsma (2003). First, split the original sample into $\binom{k}{m}$ subsamples of size m each. Then, the approximation of B of the quantile function of Y is defined by

$$B_k^{(m)}(p) = \binom{k}{m}^{-1} \sum_{1 \leq \alpha_1 \leq \cdots \leq \alpha_m \leq k} B_m(p|Y_{\alpha_1}, \ldots, Y_{\alpha_m}). \tag{3.10}$$

This can be rewritten as an L-statistic

$$B_k^{(m)}(p) = p^{m+1} + \sum_{j=1}^{m} \binom{m+1}{j} p^j (1-p)^{m+1-j} \sum_{i=j}^{k-m+j} \frac{\binom{i-1}{j-1}\binom{k-i}{m-j}}{\binom{k}{m}} Y_{(i)}. \tag{3.11}$$

From this we obtain $\hat{F}_{k,BA}^{(m)} = {B_k^{(m)}}^{-1}(\Psi)$, as an estimator of F. Differentiating $\hat{F}_{k,BA}^{(m)}$ gives $\hat{f}_{k,BA}^{(m)}$ which is an estimator for f.

For the choice of the smoothing parameter m, we use the rule of thumb determined in Albers and Schaafsma (2003)

$$m = 2.6 k^{1/2} v w \tag{3.12}$$

and round this off to the nearest integer, where $v = \|\psi - f\|_1$ with $\|\cdot\|_1$ the L_1-distance, $w = 1$ if the number of sign changes of $\psi - f$ are equal to 1 and $w = 2$ if the number of sign changes are greater than 1. A sign change occurs if $\psi - f > 0$ changes in $\psi - f < 0$ or vice versa on the support of $\psi - f$. Note that $0 \leq v \leq 2$.

For practical purposes we suggest to take $w = 2$ (because most of the time we do not know the number of sign changes of $\psi - f$) and $v = 1$ (which puts a lot of weight on the data). This means that we choose $m = 5.2\sqrt{k}$ and round this off to the nearest integer (cf. Equation (3.12)).

In Appendix 3B an example is given how the Bernstein approximation works in practice.

Appendix 3B: Illustration of the Bernstein density estimation method

To illustrate the method we describe an example from Albers and Schaafsma (2003). Take a sample of size $k = 100$ from a distribution on $(0, 1)$ with density function $f(x) = 1/2 + x$. In order to estimate the density function f with the method described in Appendix 3A we have to carry out the following steps:

1. Take an initial guess ψ. We start in this case with the uniform density on $(0, 1)$.
2. Take the order statistics of the sample: $x_{(1)}, \ldots, x_{(100)}$ and apply (3.9) to obtain $y_{(0)}, y_{(1)}, \ldots, y_{(100)}, y_{(101)}$.
3. Based on $y_{(0)}, y_{(1)}, \ldots, y_{(100)}, y_{(101)}$ and the choice of m, (3.11) can be calculated. The choice for m can be determined by the rule of thumb of Albers and Schaafsma (2003), see Equation (3.12). For this example we have $k = 100$, $v = \|\psi - f\|_1 = 0.25$ and the number of sign changes of $\psi - f$, $w = 1$. Thus m equals 6 or 7. In Figure 3.13 the result for $m = 6$ has been drawn. In this figure the stepwise function $G_{100}^{-1}((i-0.5)/100)$ is the inverse sample distribution function of Y, $i = 1, \ldots, 100$. Function B is used to smooth and to approximate this function. With $m = 6$ the approximation is based on $B_{100}^{(6)}(p)$ which is the average of $\binom{100}{6}$ functions $B_6(p)$, cf. Equation (3.10). In Figure 3.14 a sample of five $B_6(p)$'s and the resulting $B_{100}^{(6)}(p)$ have been drawn.
4. An estimate $\hat{f}$ (cf. the dotted line in Figure 3.15) of the density function f is obtained by inverting the function $B_{100}^{(6)}(p)$ such that $F_{100}^{(6)} = {B_{100}^{(6)}}^{-1}(\Psi)$ and subsequently by differentiating $F_{100}^{(6)}$. In Figure 3.15 also density functions with $m = 1$ (dashed line), $m = 100$ (dashed dotted line), and the original density function (solid line) are drawn.

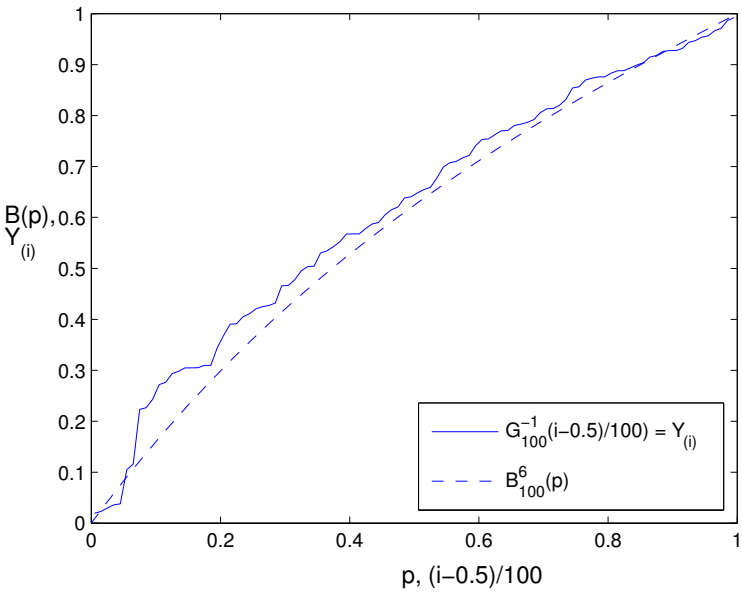

Figure 3.13: Comparison of the empirical quantiles and the function $B_{100}^{(6)}(p)$.

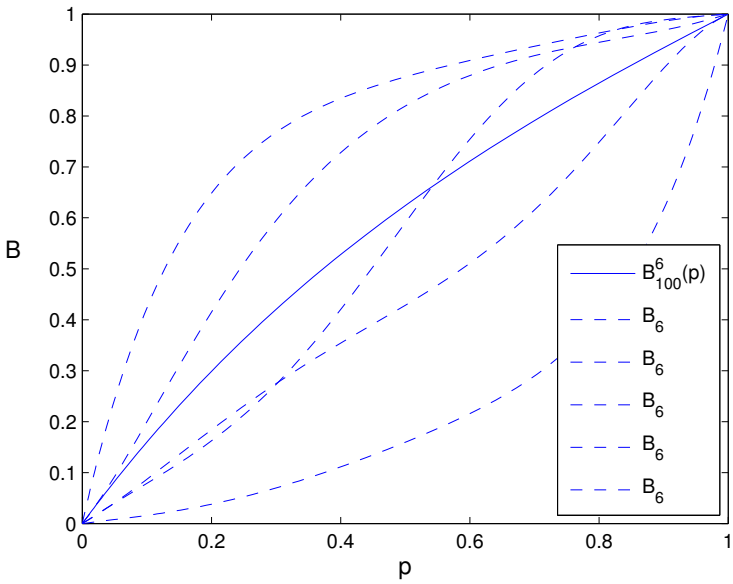

Figure 3.14: The function $B_{100}^{(6)}(p)$ and a sample from the $\binom{100}{6}$ possibilities of B_6.

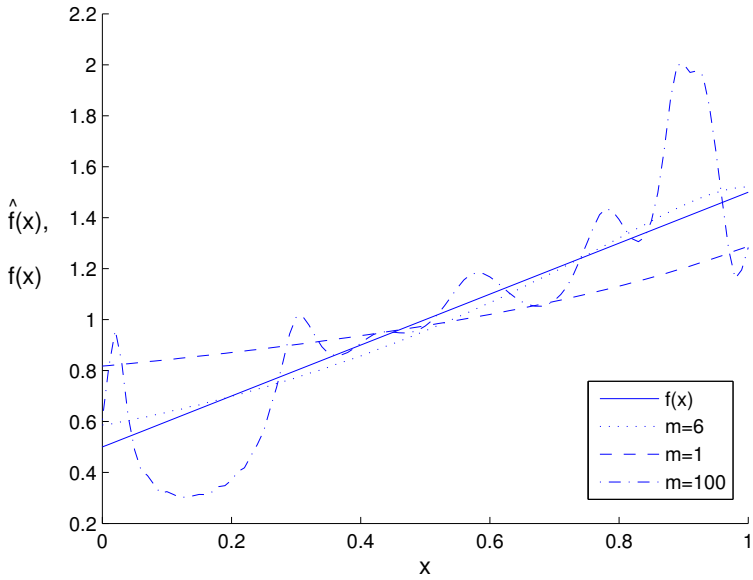

Figure 3.15: Comparison of the density function f and its approximations based on $m = 1$, $m = 6$, and $m = 100$.

Chapter 4

Asymptotic behavior of the variance of the EWMA statistic for autoregressive processes

Today's manufacturing environment is different from the time when control chart methods were originally introduced. Automated sensors and the high sampling rates often challenge the assumption of independence of successive observations. Serial correlation can seriously affect the performance of the traditional control charts. In this chapter we provide explicit easy-to-use expressions that can be used to modify the well-known and commonly used EWMA control chart for Phase II applications when the underlying process is autoregressive of order one or two, the most common type of serial correlation encountered in practice. The resulting control chart has the advantage that (a) it is a familiar tool to industrial practitioners and (b) the data are plotted on the original scale making the chart easier to interpret for practitioners than charts based on residuals. We compare three different estimation methods for the variance of the EWMA statistic adapted for autocorrelated data. This comparison is based on the asymptotic relative efficiency of the estimators. This chapter is based on two papers: Vermaat et al. (2006a) and Vermaat et al. (2006b).

4.1 Introduction

Control charts that take autocorrelation into account, get a lot of attention in the literature. Two approaches are common. The first approach is to use standard control charts with adjusted control limits taking into account the autocorrelation. The second approach is to fit an appropriate time series model to the data and monitor the residuals with standard control charts. In this chapter we focus on the first approach.

Vasilopoulos and Stamboulis (1978) calculated the exact variance for the $\bar{X}$ statistic assuming an AR(2) process. Based on this variance, appropriate control limits can be derived. Schmid

(1997), VanBrackle and Reynolds (1997), and Wieringa (1999) derived the exact variance for the EWMA statistic under the assumption of an AR(1) process.

In Zhang (1998) the variance of the EWMA statistic is studied for an AR(p) process. Zhang expressed the variance as an infinite sum of autocorrelation coefficients of the AR(p) process. However, the expressions are not in a closed form as are those derived by Schmid (1997), VanBrackle and Reynolds (1997), and Wieringa (1999). In this chapter we derive and study a closed form expression for the variance of the EWMA statistic to modify the standard EWMA control chart when the underlying process is either autoregressive of order one or two (AR(1) or AR(2) processes), the most commonly encountered types of serial correlation, see Box et al. (1994), p. 98.

The control limits obtained with the use of our closed form expression are asymptotically the same as those obtained from Zhang's (1998) expression. The difference between these two expressions is the estimation method. In Section 4.5 it is shown that our closed form expression is statistically more efficient than Zhang's (1998) expression. Moreover, Zhang (1998) compared the average run lengths (ARLs) of his control chart with the residual chart and the individual chart for AR(2) (see Table 4.2 in Zhang (1998)). Zhang (1998) notes that if the process is not nearly non-stationary then his control chart has a better performance. However, if the process is nearly non-stationary then the residual chart performs better. Hence, it seems reasonable to assume that the control chart proposed in this chapter will be comparable in performance to that of Zhang (1998) but will have an even better efficiency.

Consider observations from a stationary process $X = (X_t, t = 0, \pm 1, \pm 2, \ldots)$. The EWMA statistic at time t is defined by

$$W_t = \lambda X_t + (1 - \lambda) W_{t-1}, \qquad t = 1, 2, \ldots \tag{4.1}$$

where λ is a constant satisfying $0 < \lambda \leq 1$. Usually W_0 is set equal to a target value. The control limits of the EWMA chart converge to $\mu \pm L\sqrt{\alpha}$, where L is an appropriately chosen constant (often 3), $\mu = \mathrm{E}X_1$ and α is given by (cf. (1.4))

$$\alpha = \lim_{t \to \infty} \mathrm{Var}(W_t).$$

Therefore, after an initial period, the control limits are determined by α. Zhang (1998) derives an expression for α in terms of all autocorrelations, the EWMA parameter λ, and the variance of X. We shall denote the latter by σ_X^2. Schmid (1997) already derived the expression of Zhang (1998), see also the comments on Zhang (1998) in Schmid (1998). In case of a stationary first-order autoregressive process, Schmid (1997) has shown that Zhang's expression for α can be simplified in terms of the model parameter of an AR(1) process, σ_X^2, and λ. Similarly, in case X is a stationary second-order autoregressive process, we show that Zhang's expression can be simplified to the two model parameters of an AR(2) process, σ_X^2, and λ.

We are interested in estimating α. Let $\rho(k)$ denote the autocorrelation of X at lag $k \in \mathbb{N}$. Zhang (1998) approximates α by $\mathrm{Var}(W_M)$ for M sufficiently large. The expression ob-

tained in this way depends on the autocorrelations at lags up till M, λ, and σ_X^2. Subsequently, Zhang (1998) uses sample autocorrelations to estimate the first M autocorrelation coefficients $\rho(1), \ldots, \rho(M)$. Knowing that X is an AR(1) or AR(2) process, we can estimate α by a plug-in estimator, using the Yule-Walker estimators for the model parameters (cf. Brockwell and Davis (1991)). It is a natural question to ask which of these estimators should be preferred. It is the goal of this chapter to answer this question. To analyze the performance of the estimators, we study their asymptotic properties as the number of observations tend to infinity. Their asymptotic distributions are used to study the relative efficiency of the estimators. It turns out that in case X is an autoregressive process, the estimators based on plugging in the Yule-Walker estimators outperform the estimator of Zhang (1998) in terms of relative efficiency. Intuitively, this was to be expected since fewer parameters have to be estimated and more information is available. The analysis is checked by a simulation study.

This chapter is organized as follows. In the next section we present a few preliminary results on autoregressive processes. In Section 4.3 we present the variances of the EWMA statistic as in Schmid (1997), Zhang (1998), and Vermaat et al. (2006a). In Section 4.4 we introduce the estimators for these variances. In the subsequent section we will study their asymptotic distributions and their relative efficiencies. Section 4.6 contains a simulation study which is used to confirm the derived asymptotic results. In Section 4.7 the modified EWMA control chart, using the asymptotic expression for the variance, is applied to the data of the ceramic furnace example, introduced in Chapter 1. We finish this chapter with some concluding remarks.

4.2 Preliminaries on second order autoregressive processes

In this section we summarize some results on autoregressive processes that we will need in the following. Let $A = (A_t, t = 0, \pm 1, \pm 2, \ldots)$ be a white-noise series, i.e. a series of zero mean, uncorrelated random variables. A second order autoregressive process (an AR(2) process) $X = (X_t, t = 0, \pm 1, \pm 2, \ldots)$ is defined via the recursive relation

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + A_t, \quad t = \ldots, -2, -1, 0, 1, 2, \ldots. \tag{4.2}$$

Denote the variance of A_1 by σ_A^2. Define the quadratic function ϕ by $\phi(z) = 1 - \phi_1 z - \phi_2 z^2$. If

$$\phi(z) \neq 0 \text{ for all } z \in \mathbb{C} \text{ with } |z| \leq 1, \tag{4.3}$$

then the unique stationary AR(2) process exists and is given by $X_t = \sum_{j=0}^{\infty} \psi_j A_{t-j}$ where $\{\psi_j\}_{j=0}^{\infty}$ is a sequence of numbers that is absolutely summable and which can be determined from ϕ by the relation $\psi(z) = 1/\phi(z) = \sum_{j=0}^{\infty} \psi_j z^j$ ($|z| \leq 1$), see Brockwell and Davis (1991), Theorem 3.1.1. The latter implies that the relationship between A and X is *causal* (cf.

Brockwell and Davis (1991), Definition 3.1.3). The set of values of (ϕ_1, ϕ_2) for which $\phi(z) \neq 0$ for all $z \in \mathbb{C}$ with $|z| \leq 1$ equals

$$\{(\phi_1, \phi_2) \in \mathbb{R}^2 \,:\, \phi_1 + \phi_2 < 1,\ \phi_2 - \phi_1 < 1 \text{ and } -1 < \phi_2 < 1\}.$$

Of course, the case that $\phi_2 = 0$ corresponds to a first order autoregressive process.

We recall that the autocovariance function of a stationary time series is given by $\gamma(k) = \text{cov}(X_{t+k}, X_t)$, $k = 0, 1, \ldots$. The autocorrelation function is given by $\rho(k) = \gamma(k)/\gamma(0)$. The first two autocovariances of the AR(2) process are

$$\gamma(0) = \frac{(\phi_2 - 1)\sigma_A^2}{(1 + \phi_2)(\phi_1^2 - (1 - \phi_2)^2)} \tag{4.4}$$

and

$$\gamma(1) = \frac{-\phi_1 \sigma_A^2}{(1 + \phi_2)(\phi_1^2 - (1 - \phi_2)^2)}. \tag{4.5}$$

the covariance function of the AR(2) process is

$$\gamma(k) = \begin{cases} \dfrac{\nu_1^k(\gamma(1) - \nu_2\gamma(0))}{\nu_1 - \nu_2} - \dfrac{\nu_2^k(\gamma(1) - \nu_1\gamma(0))}{\nu_1 - \nu_2} & \text{if} \quad \nu_1 \neq \nu_2 \\ \\ \nu^{k-1}[k(\gamma(1) - \nu\gamma(0)) + \nu\gamma(0)] & \text{if} \quad \nu_1 = \nu_2 = \nu. \end{cases} \tag{4.6}$$

where ν_1 and ν_2 are the roots of the characteristic functions $\pi^2 - \phi_1\pi - \phi_2 = 0$, see Fuller (1996), pp. 54–56. The AR(2) process is stationary if the roots ν_1 and ν_2 are within the unit circle, i.e. $|\nu_1| < 1, |\nu_2| < 1$. Note, that the unit roots of the characteristic equation are complex if $\phi_1^2 + 4\phi_2 < 0$. The autocovariance function (4.6) can in that case be written as

$$\gamma(k) = \frac{r^{k-1}[\gamma(1)\sin k\theta - \gamma(0) r \sin(k-1)\theta]}{\sin\theta},$$

where $\nu_1 = re^{i\theta}$ and $\nu_2 = re^{-i\theta}$.

4.3 The variance of the EWMA statistic

In this section we give expressions for the variance of the EWMA statistic. These expressions are used to determine the control limits for the EWMA control chart. Zhang (1998) proposed the EWMAST chart (EWMA chart for general stationary processes) to monitor stationary processes. Zhang (1998) derived the variance of the EWMA statistic W_t at time t as

$$\alpha_t = \text{Var}(W_t) = \frac{\lambda}{2 - \lambda}\sigma_X^2 \left(1 - (1 - \lambda)^{2t} + 2\sum_{k=1}^{t-1} \rho(k)(1 - \lambda)^k (1 - (1 - \lambda)^{2(t-k)})\right). \tag{4.7}$$

If $t \to \infty$, this becomes

$$\alpha := \lim_{t\to\infty} \alpha_t = \frac{\lambda}{2-\lambda}\left(2\sum_{k=0}^{\infty}\rho(k)(1-\lambda)^k - 1\right)\sigma_X^2. \tag{4.8}$$

The following lemma shows that if we assume that X is a causal stationary AR(2) process, α can be expressed in the AR(2) model parameters.

Lemma 1 *Suppose X is a causal stationary AR(2) process as defined in (4.2). Then*

$$\alpha = \frac{\lambda}{2-\lambda}\left(\frac{\phi_1(1+\phi_2)(\lambda-1)+(\phi_2-1)(1+\phi_2(\lambda-1)^2)}{(1-\phi_2)(-1+\phi_1(1-\lambda)+\phi_2(\lambda-1)^2)}\right)\sigma_X^2. \tag{4.9}$$

In particular; if $\phi_2 = 0$ (i.e. an AR(1) process), then

$$\alpha = \frac{\lambda}{2-\lambda}\left(\frac{1+\phi_1(1-\lambda)}{1-\phi_1(1-\lambda)}\right)\sigma_X^2. \tag{4.10}$$

Proof The autocorrelation function $\rho(k)$ for an AR(2) process satisfies the second order difference equation

$$\rho(k) = \phi_1\rho(k-1) + \phi_2\rho(k-2), \qquad k > 0, \tag{4.11}$$

with starting values $\rho(0) = 1$ and $\rho(1) = \phi_1/(1-\phi_2)$. The general solution of (4.11) can be found in Brockwell and Davis (1991), p. 108. As it turns out, this solution depends on the zeros of the characteristic equation corresponding to (4.11): $\pi^2 - \phi_1\pi - \phi_2 = 0$. These zeros are given by

$$\nu_1 = \left(\phi_1 + \sqrt{\phi_1^2 + 4\phi_2}\right)/2, \quad \nu_2 = \left(\phi_1 - \sqrt{\phi_1^2 + 4\phi_2}\right)/2. \tag{4.12}$$

Hence, we can distinguish three cases: *(i)* both zeros are real, *(ii)* there is one real zero with multiplicity two, *(iii)* both zeros are conjugate complex numbers. Some tedious computations yield that the solution to (4.11) is given by

$$\rho(k) = \begin{cases} \dfrac{\nu_2^k(\phi_1+\nu_1(\phi_2-1)) - \nu_1^k(\phi_1+\nu_2(\phi_2-1))}{(\nu_1-\nu_2)(\phi_2-1)} & \text{if} \quad \nu_1 \neq \nu_2 \quad \text{are real} \\[2ex] \dfrac{\nu^{k-1}(-k\phi_1+\nu(k-1+\phi_2(1-k)))}{\phi_2-1} & \text{if} \quad \nu_1 = \nu_2 = \nu \quad \text{is real,} \\[2ex] c\nu_1^k + \bar{c}\,\overline{\nu_1}^k & \text{if} \quad \nu_1 = \overline{\nu_2} \quad \text{are complex,} \end{cases} \tag{4.13}$$

where $c = (1 - i\phi_1(1+\phi_2)(1-\phi_2)^{-1}|\phi_1^2 + 4\phi_2|^{-1/2})/2$. Note that condition (4.3) ensures that $\max\{|\nu_1|, |\nu_2|\} < 1$. Hence, substituting (4.13) into (4.8) yields a convergent series. After some tedious calculations we find that

$$\sum_{k=0}^{\infty}\rho(k)(1-\lambda)^k = \frac{-1+\phi_2(1+(\lambda-1)\phi_1)}{(\phi_2-1)(1+(\lambda-1)\phi_1-(1-\lambda)^2\phi_2)}.$$

Substituting this expression into (4.8) gives the result. □

Remark 4.1 (i) The variance of the marginal distribution of X is related to σ_A^2 by (cf. Box et al. (1994), p. 62)

$$\sigma_X^2 = \frac{1-\phi_2}{1+\phi_2}\left(\frac{\sigma_A^2}{(1-\phi_2)^2-\phi_1^2}\right). \tag{4.14}$$

Substituting this expression into (4.10), we recover the result of Schmid (1997): for a causal stationary AR(1) process α is given by

$$\alpha = \frac{\lambda}{2-\lambda}\left(\frac{\sigma_A^2}{1-\phi_1^2}\right)\frac{1+\phi_1(1-\lambda)}{1-\phi_1(1-\lambda)}. \tag{4.15}$$

(ii) The resulting variance in (4.9) is also derived in Vermaat et al. (2006a). Moreover, in their paper an expression for α_t in the AR(2) model parameters is derived. This derivation can also be found in Appendix 4A.

(iii) In Appendix 4B three special cases of the variance α in (4.9) are given.

4.4 Estimators for the variance of the EWMA statistic

In this section we define three estimators for α. Suppose we observe the values $X_1, \ldots, X_n$ from the stationary process $X = (X_t, t = 0, \pm 1, \pm 2)$, then the autocorrelation function $\rho(k)$ is estimated by the sample autocorrelation at lag k defined by $\hat{\rho}_n(k) = \hat{\gamma}_n(k)/\hat{\gamma}_n(0)$, where

$$\hat{\gamma}_n(k) = \frac{1}{n}\sum_{t=1}^{n-k}(X_t - \bar{X}_n)(X_{t+k} - \bar{X}_n), \qquad k = 0, 1, 2, \ldots, n-1, \tag{4.16}$$

$$\bar{X}_n = \frac{1}{n}\sum_{t=1}^{n} X_t.$$

Zhang's estimator: For a stationary time series Zhang (1998) approximates α by α_M (cf. (4.7)), where M is an integer greater than 25, setting $(1-\lambda)^{2M} \approx 0$. This gives the following approximation for α (cf. (4.8))

$$\alpha^{(Z)} = \frac{\lambda}{2-\lambda}\sigma_X^2\left(1 + 2\sum_{k=1}^{M}\rho(k)(1-\lambda)^k(1-(1-\lambda)^{2(M-k)})\right). \tag{4.17}$$

We suppress the dependence on M in the notation. We estimate $\alpha^{(Z)}$ by a plug-in estimator $\hat{\alpha}_n^{(Z)}$, that is defined by

$$\hat{\alpha}_n^{(Z)} = \frac{\lambda}{2-\lambda}\hat{\gamma}_n(0)\left(1 + 2\sum_{k=1}^{M}\hat{\rho}_n(k)(1-\lambda)^k(1-(1-\lambda)^{2(M-k)})\right).$$

Schmid's estimator: For a causal stationary AR(1) process, Schmid (1997) estimates α by (cf. (4.15))

$$\hat{\alpha}_n^{(S)} = \frac{\lambda}{2-\lambda}\left(\frac{1+\hat{\phi}_1(1-\lambda)}{1-\hat{\phi}_1(1-\lambda)}\right)\hat{\gamma}_n(0), \tag{4.18}$$

where $\hat{\phi}_1 = \hat{\gamma}_n(1)/\hat{\gamma}_n(0)$ is the Yule-Walker estimator for ϕ_1.

Estimator based on Vermaat et al. (2006a): For a causal stationary AR(2) process, Vermaat et al. (2006a) estimate α by (cf. (4.9))

$$\hat{\alpha}_n^{(V)} = \frac{\lambda}{2-\lambda}\left(\frac{\hat{\phi}_1(1+\hat{\phi}_2)(\lambda-1)+(\hat{\phi}_2-1)(1+\hat{\phi}_2(\lambda-1)^2)}{(1-\hat{\phi}_2)(-1+\hat{\phi}_1(1-\lambda)+\hat{\phi}_2(\lambda-1)^2)}\right)\hat{\gamma}_n(0). \tag{4.19}$$

Here $\hat{\phi}_1$ and $\hat{\phi}_2$ are the Yule-Walker estimators for ϕ_1 and ϕ_2 respectively. The latter are given by

$$\hat{\phi}_1 = \frac{\hat{\gamma}_n(1)(\hat{\gamma}_n(0)-\hat{\gamma}_n(2))}{\hat{\gamma}_n(0)^2-\hat{\gamma}_n(1)^2} \quad \text{and} \quad \hat{\phi}_2 = \frac{\hat{\gamma}_n(0)\hat{\gamma}_n(2)-\hat{\gamma}_n(1)^2}{\hat{\gamma}_n(0)^2-\hat{\gamma}_n(1)^2}. \tag{4.20}$$

Hence, for a first order autoregressive process we have two estimators that can be used to derive the control limits of the EWMA-chart: $\hat{\alpha}_n^{(Z)}$ and $\hat{\alpha}_n^{(S)}$. In the next section we compare these estimators by studying their asymptotic distribution. Similarly, we will compare $\hat{\alpha}_n^{(Z)}$ and $\hat{\alpha}_n^{(V)}$ for second order autoregressive processes.

As Zhang (1998) reported, the advantage of using the approximate variance as in (4.17) hinges on the fact that no modelling efforts are required. Only a number of autocorrelations have to be estimated to obtain an estimator for the variance of the EWMA statistic. However:

1. In practice it is valuable to know whether the process under study exhibits autocorrelation. Modelling the autocorrelation structure of a process gives insight in the working of the process. This may generate improvement actions besides the monitoring purpose.
2. With modern software it is relatively easy to fit a time series model to a given series.
3. Zhang (1998) reported that for certain values of the parameters of the time series model other control charts perform better than the EWMAST, in the sense that they have better average run length (ARL) properties for signalling out-of-control situations, like a shift in the mean. So, to choose the most appropriate control chart we need to estimate the time series parameters.
4. As we can see the estimator $\hat{\alpha}_n^{(V)}$ uses only estimators for ϕ_1, ϕ_2, and σ_X^2, while the estimator $\hat{\alpha}_n^{(Z)}$ uses up to M estimates for the autocorrelations and an estimate for the process variance σ_X^2. Box et al. (1994), p. 44 pointed out that each correlation is a parameter to be estimated. Hence, Zhang's approach might be very prodigal with parameters, whereas the approach via the estimated model parameters could be parsimonious.

5. The estimators $\hat{\alpha}_n^{(S)}$ and $\hat{\alpha}_n^{(V)}$ are asymptotically more efficient for AR(1) respectively AR(2) processes than the estimator $\hat{\alpha}_n^{(Z)}$, which is shown in Section 4.5 of this chapter.

We illustrate the difference in the two approaches by analyzing the asymptotic behavior of the estimators for the control limits.

4.5 Asymptotics of the estimators

In this section we analyze the asymptotic behavior of the estimators $\hat{\alpha}_n^{(Z)}$,$\hat{\alpha}_n^{(S)}$, and $\hat{\alpha}_n^{(V)}$ as defined in (4.17), (4.18), and (4.19) respectively, when n grows large.

4.5.1 Asymptotic distributions

The following result will be used in the sequel. We will use the abbreviation IID to denote Independent and Identically Distributed.

Proposition 4.1 *Suppose X is a causal AR(p) process generated by the IID sequence A. Assume $\mathrm{E}(A_1^4) < \infty$ and let $\eta = \mathrm{E}(A_1^4)/\sigma_A^4$. For any non-negative integer k and some $(k+1) \times (k+1)$-matrix Ω_{k+1},*

$$\sqrt{n}\left(\begin{bmatrix} \hat{\gamma}_n(0) \\ \hat{\gamma}_n(1) \\ \vdots \\ \hat{\gamma}_n(k) \end{bmatrix} - \begin{bmatrix} \gamma(0) \\ \gamma(1) \\ \vdots \\ \gamma(k) \end{bmatrix}\right) \xrightarrow{d} N_{k+1}(\mathbf{0}, \Omega_{k+1}),$$

where $\xrightarrow{d}$ denotes convergence in distribution. The estimator $\hat{\gamma}_n(k)$ is given in (4.16) and

$$\Omega_{k+1} = \left[(\eta - 3)\gamma(q)\gamma(r) + \sum_i (\gamma(i)\gamma(i-q+r) + \gamma(i+r)\gamma(i-q))\right]_{q,r=0,\ldots,k}. \quad (4.21)$$

Proof A causal AR(p) process X is a moving average of A with $X_t = \sum_{j=0}^{\infty} \phi_j A_{t-j}$ (see Chapter 3 in Brockwell and Davis (1991)). Proposition 4.1 is proved in Section 7.3 of Brockwell and Davis (1991) for an MA(∞) process and hence for a causal AR(p) process. □

Now we will give in three theorems the asymptotic distributions of the three estimators $\hat{\alpha}_n^{(Z)}$,$\hat{\alpha}_n^{(S)}$, and $\hat{\alpha}_n^{(V)}$ as defined in (4.17), (4.18), and (4.19) respectively.

Theorem 4.1 *Suppose X is a causal AR(p) process generated by the IID sequence A. Assume $\mathrm{E}(A_1^4) < \infty$ and let $\eta = \mathrm{E}(A_1^4)/\sigma_A^4$, then*

$$\sqrt{n}(\hat{\alpha}_n^{(Z)} - \alpha^{(Z)}) \xrightarrow{d} N(0, V^{(Z)}),$$

where $V^{(Z)}$ is given by $V^{(Z)} = \mathbf{c}'\Omega_{M+1}\mathbf{c}$ for $\mathbf{c}$ as in (4.22) and Ω_{M+1} as in (4.21).

Proof Recall the definition of $\hat{\alpha}_n^{(Z)}$ in (4.17). Write $\hat{\alpha}_n^{(Z)} = \boldsymbol{c}'\hat{\boldsymbol{\gamma}}_{\boldsymbol{n}} := \boldsymbol{c}'(\hat{\gamma}_n(0), \ldots, \hat{\gamma}_n(M))'$ for $\boldsymbol{c} \in \mathbb{R}^{M+1}$ defined by

$$\boldsymbol{c} = \frac{\lambda}{2-\lambda} \begin{bmatrix} 1 \\ 2(1-\lambda)(1-(1-\lambda)^{2(M-1)}) \\ 2(1-\lambda)^2(1-(1-\lambda)^{2(M-2)}) \\ \vdots \\ 2(1-\lambda)^{M-1}(1-(1-\lambda)^2) \\ 0 \end{bmatrix}. \tag{4.22}$$

Let $\boldsymbol{\gamma} = (\gamma(0), \ldots, \gamma(M))'$. Using Proposition 4.1, an application of the Cramér-Wold device (cf. Pollard (2002), p. 202) gives

$$\sqrt{n}(\hat{\alpha}_n^{(Z)} - \alpha^{(Z)}) = \sqrt{n}(\boldsymbol{c}'\hat{\boldsymbol{\gamma}}_{\boldsymbol{n}} - \boldsymbol{c}'\boldsymbol{\gamma}) \xrightarrow{d} N(0, \boldsymbol{c}'\Omega_{M+1}\boldsymbol{c}).$$

□

Theorem 4.2 *Suppose X is a causal AR(1) process generated by the IID sequence A. Assume $\mathrm{E}(A_1^4) < \infty$ and let $\eta = \mathrm{E}(A_1^4)/\sigma_A^4$, then*

$$\sqrt{n}(\hat{\alpha}_n^{(S)} - \alpha) \xrightarrow{d} N(0, V^{(S)}).$$

where $V^{(S)}$ is given by $V^{(S)} = \boldsymbol{b}'\Omega_2\boldsymbol{b}$ for $\boldsymbol{b}$ as

$$\boldsymbol{b} = \frac{\lambda}{2-\lambda}\left(\frac{2\gamma(0)+\gamma(1)(1-\lambda)}{\gamma(0)-\gamma(1)(1-\lambda)} - \frac{\gamma(0)^2+\gamma(0)\gamma(1)(1-\lambda)}{(\gamma(0)-\gamma(1)(1-\lambda))^2}, \frac{2\gamma(0)^2(1-\lambda)}{(\gamma(0)-\gamma(1)(1-\lambda))^2}\right)'$$

and Ω_2 as in (4.21).

Proof Defining $f : \mathbb{R}^2 \to \mathbb{R}$ by

$$f(x,y) = \frac{\lambda}{2-\lambda}\left(\frac{x^2+xy(1-\lambda)}{x-y(1-\lambda)}\right),$$

we have that

$$\sqrt{n}(\hat{\alpha}_n^{(S)} - \alpha) = \sqrt{n}(f(\hat{\gamma}_n(0), \hat{\gamma}_n(1)) - f(\gamma(0), \gamma(1))).$$

By the Delta-method (cf. Pollard (2002), p. 184), together with Proposition 4.1, this converges weakly to a normal distribution with variance $\nabla f(\gamma(0), \gamma(1))\Omega_1\nabla f(\gamma(0), \gamma(1))'$. One can verify that the gradient $\nabla f(\gamma(0), \gamma(1))$ equals the vector b. □

Theorem 4.3 *Suppose X is a causal AR(2) process generated by the IID sequence A. Assume $\mathrm{E}(A_1^4) < \infty$ and let $\eta = \mathrm{E}(A_1^4)/\sigma_A^4$, then*

$$\sqrt{n}(\hat{\alpha}_n^{(V)} - \alpha) \xrightarrow{d} N(0, V^{(V)}),$$

where $V^{(V)}$ is given by $V^{(V)} = \boldsymbol{d}'\Omega_3\boldsymbol{d}$ for $\boldsymbol{d} = \nabla h(\gamma(0), \gamma(1), \gamma(2))'$ as in (4.23) and Ω_3 as in (4.21) .

Proof Substituting (4.20) into (4.19) we obtain $\hat{\alpha}_n^{(V)} = h(\hat{\gamma}_n(0), \hat{\gamma}_n(1), \hat{\gamma}_n(2))$, with $h : \mathbb{R}^3 \to \mathbb{R}$ defined by

$$\begin{aligned} h(x,y,z) &= \frac{\lambda}{2-\lambda}\left(\frac{x^3+2y^3(\lambda-1)-x^2(\lambda-1)(y+z(1-\lambda))}{x^2+x(\lambda-1)(y+z(1-\lambda))+y(z(1-\lambda)+y(\lambda-2)\lambda))}\right. \\ &\qquad \left. -\frac{xy(z(\lambda-1)+y(2+(\lambda-2)\lambda))}{x^2+x(\lambda-1)(y+z(1-\lambda))+y(z(1-\lambda)+y(\lambda-2)\lambda))}\right). \end{aligned}$$

An application of Proposition 1, combined with the Delta method, gives

$$\sqrt{n}(\hat{\alpha}_n^{(V)} - \alpha) \xrightarrow{d} N(0, \nabla h(\gamma(0), \gamma(1), \gamma(2))\Omega_3 \nabla h(\gamma(0), \gamma(1), \gamma(2))').$$

Here, the gradient $\nabla h : \mathbb{R}^3 \to \mathbb{R}^3$ of h has elements

$$\begin{aligned} \frac{\partial h(x,y,z)}{\partial x} &= u_\lambda(x,y,z)(x^4+2x^3(\lambda-1)(y+z-z\lambda)+2xy\{-z^2(\lambda-1)^3 \\ &\quad +yz(\lambda-1)^4-y^2(\lambda-1)(2+(\lambda-2)\lambda)\} \\ &\quad +x^2(y+z(\lambda-1))\{-z(\lambda-1)^3+y(1+3(\lambda-2)\lambda)\} \\ &\quad +y^2\{z^2(\lambda-1)^2+2yz(\lambda-1)(2+(\lambda-2)\lambda) \\ &\quad -y^2(2+(\lambda-2)\lambda(4+(\lambda-2)\lambda))\} \\ \frac{\partial h(x,y,z)}{\partial y} &= u_\lambda(x,y,z)\,(2(\lambda-1)\{-x^4+x^2(2y^2+2yz(\lambda-1)+z^2(\lambda-1)^2) \\ &\quad -2x^3y(\lambda-1)+2xy^2(\lambda-1)(y+z-z\lambda) \\ &\quad +y^3(-2z(\lambda-1)+y(\lambda-2)\lambda)\}) \\ \frac{\partial h(x,y,z)}{\partial z} &= u_\lambda(x,y,z)2(\lambda-1)^2(x^2-y^2)^2, \end{aligned} \tag{4.23}$$

where,

$$u_\lambda(x,y,z) = \lambda(2-\lambda)^{-1}\left(x^2+x(\lambda-1)(y+z-z\lambda)+y(z-z\lambda+y(\lambda-2)\lambda)\right)^{-2}.$$

□

Remark 4.2 For autoregressive processes the Yule-Walker estimators are asymptotically equivalent to the maximum likelihood estimators based on a Gaussian likelihood (cf. Brockwell and Davis (1991), p. 240). That is, rescaled differences of the estimators subtracted by their true value converge weakly to a multivariate normal distribution with the same variance. This implies efficiency of the Yule-Walker estimators. Our estimators are obtained by plugging in these efficient estimators.

4.5.2 Asymptotic relative efficiencies

In this section we study the efficiency of $\hat{\alpha}_n^{(S)}$ over $\hat{\alpha}_n^{(Z)}$ for an AR(1) process and $\hat{\alpha}_n^{(V)}$ and $\hat{\alpha}_n^{(Z)}$ for an AR(2) process. Note that $\alpha_n^{(S)}$ and $\hat{\alpha}_n^{(V)}$ are consistent estimators for α while $\hat{\alpha}_n^{(Z)}$ is not.

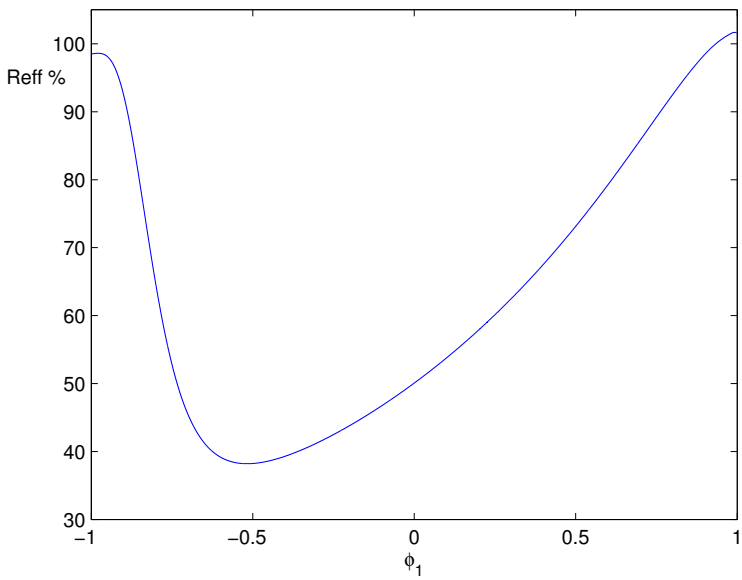

Figure 4.1: The asymptotic relative efficiency of the estimators $\hat{\alpha}_n^{(Z)}$ (cf. (4.17)) and $\hat{\alpha}_n^{(S)}$ (cf. (4.18)) for different causal stationary AR(1) processes.

First suppose that the process under study is a causal stationary first order autoregressive process (this is equivalent to $|\phi_1| < 1$). We define the asymptotic relative efficiency for the estimators $\hat{\alpha}_n^{(Z)}$ and $\hat{\alpha}_n^{(S)}$ by

$$Reff(\hat{\alpha}_n^{(Z)}, \hat{\alpha}_n^{(S)}) = \frac{V^{(S)}}{V^{(Z)}} = \frac{\boldsymbol{b}'\Omega_2\boldsymbol{b}}{\boldsymbol{c}'\Omega_{M+1}\boldsymbol{c}}.$$

In this definition we neglect the fact that $\hat{\alpha}_n^{(Z)}$ estimates $\alpha^{(Z)}$, whereas $\hat{\alpha}_n^{(S)}$ estimates α. We use the current definition of relative efficiency purely to compare the asymptotic variances of the estimators. In Figure 4.1 we have plotted the asymptotic relative efficiency for different ϕ_1s in the stationary region of an AR(1) process with $M = 25$. As we can see $Reff(\hat{\alpha}_n^{(Z)}, \hat{\alpha}_n^{(S)})$ varies from 0.4 to 1. Only at the boundaries of the stationary region $Reff(\hat{\alpha}_n^{(Z)}, \hat{\alpha}_n^{(S)}) > 1$, here the estimator of Zhang (1998)(cf. (4.17)) is more efficient.

For a causal stationary AR(2) process we define the asymptotic relative efficiency for the estimators $\hat{\alpha}_n^{(Z)}$ and $\hat{\alpha}_n^{(V)}$ by

$$Reff(\hat{\alpha}_n^{(Z)}, \hat{\alpha}_n^{(V)}) = \frac{V^{(V)}}{V^{(Z)}} = \frac{\boldsymbol{d}'\Omega_3\boldsymbol{d}}{\boldsymbol{c}'\Omega_{M+1}\boldsymbol{c}}.$$

In Figure 4.2 we have drawn contour lines of $Reff(\hat{\alpha}_n^{(Z)}, \hat{\alpha}_n^{(V)})$ for different ϕ_1 and ϕ_2 in the stationary region of an AR(2) process, i.e. $\phi_1+\phi_2 < 1$, $\phi_1-\phi_2 < 1$, and $-1 < \phi_2 < 1$ and with $M = 25$. The asymptotic relative efficiency varies from .56 to 1. Again $Reff(\hat{\alpha}_n^{(Z)}, \hat{\alpha}_n^{(V)}) > 1$ at the boundaries of the stationary region.

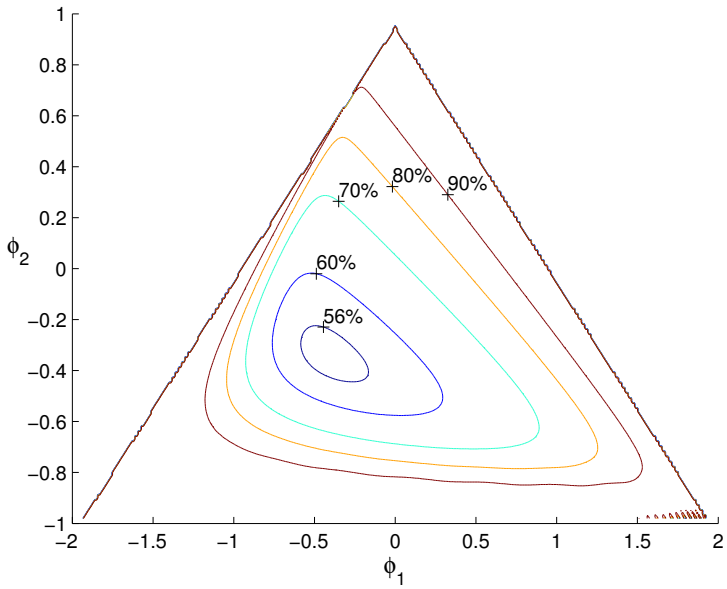

Figure 4.2: Contour plot of the asymptotic relative efficiency of the estimators $\hat{\alpha}_n^{(Z)}$ (cf. (4.17)) and $\hat{\alpha}_n^{(V)}$ (cf. (4.19)) for different causal stationary AR(2) processes.

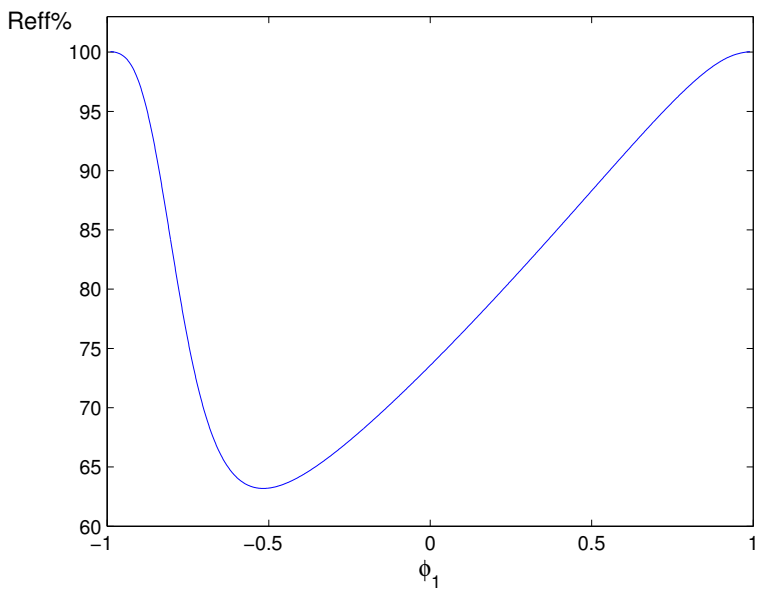

Figure 4.3: The asymptotic relative efficiency of the estimators $\hat{\alpha}_n^{(S)}$ (cf. (4.18)) and $\hat{\alpha}_n^{(V)}$ (cf. (4.19)) for different causal stationary AR(1) processes.

Remark 4.3 Suppose X is a (stationary, causal) first order autoregressive time series, but we estimate α by $\hat{\alpha}_n^{(V)}$, which is "designed" for a second order autoregressive process. In Figure 4.3 we have plotted

$$Reff(\hat{\alpha}_n^{(V)}, \hat{\alpha}_n^{(S)}) = \frac{V^{(S)}}{V^{(V)}} = \frac{\boldsymbol{b}'\Omega_2\boldsymbol{b}}{\boldsymbol{d}'\Omega_3\boldsymbol{d}}.$$

for $\phi_1 \in (-1,1)$. As we can see the $Reff(\hat{\alpha}_n^{(V)}, \hat{\alpha}_n^{(S)}) < 1$ for all $\phi_1 \in (-1,1)$. Thus, overfitting by one degree results in a loss of efficiency, which can be up to 35%.

4.6 Simulations

In this section the asymptotic properties out of the previous sections are confirmed. The simulation study was done as follows. We simulate $\sqrt{n}(\hat{\alpha}_n - \alpha)$ 10,000 times for all three estimators. The length of the times series n varies from 75, 500, to 3,000. For the estimation of α we have chosen to set $\lambda = 0.2$ and the noise to be Gaussian with zero mean and standard deviation one. The sample means of the 10,000 simulations of $\sqrt{n}(\alpha_n^{(Z)} - \alpha^{(Z)})$, $\sqrt{n}(\alpha_n^{(S)} - \alpha)$, and $\sqrt{n}(\alpha_n^{(V)} - \alpha)$ are denoted by $\overline{\sqrt{n}(\alpha_n^{(Z)} - \alpha^{(Z)})}$, $\overline{\sqrt{n}(\alpha_n^{(S)} - \alpha)}$, and $\overline{\sqrt{n}(\alpha_n^{(V)} - \alpha)}$ respectively and their sample variances are denoted by $V_n^{(Z)}$, $V_n^{(S)}$, and $V_n^{(V)}$ respectively.

Table 4.1: Simulated results for AR(1) processes.

	n	$\overline{\sqrt{n}(\hat{\alpha}_n^{(Z)} - \alpha^{(Z)})}$	$V^{(Z)}$	$\overline{\sqrt{n}(\hat{\alpha}_n^{(S)} - \alpha)}$	$V_n^{(S)}$	$Reff_n$
ϕ_1=-0.8	75	-0.0199	0.0205	0.0128	0.0148	0.7237
	500	-0.0089	0.0211	0.0041	0.0139	0.6588
	3,000	-0.0037	0.0219	0.0022	0.0142	0.6478
	∞	0	0.0219	0	0.0142	0.6455
ϕ_1=-0.4	75	-0.0496	0.0238	0.0009	0.0123	0.5177
	500	-0.0226	0.0295	-0.0006	0.0120	0.4068
	3,000	-0.0071	0.0305	0.0016	0.0120	0.3919
	∞	0	0.0305	0	0.0120	0.3930
ϕ_1= 0.5	75	-0.4754	1.1651	-0.2010	1.0459	0.8998
	500	-0.1797	1.5169	-0.0703	1.1607	0.7652
	3,000	-0.0990	1.6129	-0.0386	1.1716	0.7264
	∞	0	1.6116	0	1.1787	0.7314

In Table 4.1 the results of the simulation study for an AR(1) process are given. In the last column of Table 4.1 we have calculated the empirical relative efficiency, based on the 10,000 simulations. The results are given for $\phi_1 = -0.8$, $\phi_1 = -0.5$ and $\phi_1 = 0.4$. The empirical relative efficiency between Zhang's and Schmid's estimator is defined by

$$Reff_n(\hat{\alpha}_n^{(Z)}, \hat{\alpha}_n^{(S)}) = \frac{V_n^{(S)}}{V_n^{(Z)}}. \tag{4.24}$$

The empirical relative efficiency converges to the asymptotic relative efficiency. From Table 4.1 we observe for all ϕ_1 that the variances of the sequences $V_n^{(Z)}$ and $V_n^{(S)}$ converge to their derived asymptotic values (which are in Table 4.1 given in the rows with $n = \infty$).

Table 4.2: Simulated results for AR(2) processes.

	n	$\overline{\sqrt{n}(\hat{\alpha}_n^{(Z)} - \alpha^{(Z)})}$	$V^{(Z)}$	$\overline{\sqrt{n}(\hat{\alpha}_n^{(V)} - \alpha)}$	$V_n^{(V)}$	$Reff_n$
ϕ_1=-1.0	75	0.0003	0.0100	0.0113	0.0087	0.8706
ϕ_2=-0.6	500	0.0002	0.0099	0.0041	0.0084	0.8425
	3,000	0.0005	0.0096	0.0026	0.0081	0.8449
	∞	0	0.0096	0	0.0081	0.8359
ϕ_1=-0.4	75	-0.0302	0.0127	-0.0021	0.0087	0.6857
ϕ_2=-0.3	500	-0.0136	0.0150	-0.0017	0.0087	0.5779
	3,000	-0.0042	0.0152	0.0004	0.0087	0.5711
	∞	0	0.0154	0	0.0085	0.5540
ϕ_1= 0	75	-0.3358	0.4772	-0.1993	0.4752	0.9979
ϕ_2= 0.4	500	-0.1442	0.6934	-0.0889	0.5961	0.8597
	3,000	-0.0533	0.7146	-0.0312	0.5929	0.8297
	∞	0	0.7174	0	0.6004	0.8370
ϕ_1= 0.5	75	-0.2197	0.3653	-0.0925	0.3227	0.8849
ϕ_2=-0.2	500	-0.0929	0.4550	-0.0425	0.3427	0.7531
	3,000	-0.0469	0.4764	-0.0176	0.3506	0.7359
	∞	0	0.4786	0	0.3507	0.7328
ϕ_1= 1.2	75	-0.3995	2.5065	-0.0109	2.3954	0.9565
ϕ_2=-0.7	500	-0.1458	2.8557	0.0100	2.2892	0.8016
	3,000	-0.0322	2.9126	0.0264	2.2758	0.7814
	∞	0	2.9200	0	2.2867	0.7831

We have chosen to vary ϕ_1 and ϕ_2 from (-1 -0.6; -0.4 -.3; 0.0 0.4; 0.5 -0.2; 1.2 -0.7). In Table 4.2 we see the results of the simulation for different AR(2) processes. We observe comparable results for the AR(2) process as for the AR(1) process. The empirical relative efficiency between Zhang's estimator and the estimator based on Vermaat et al. (2006a) is defined mutatis mutandis as in (4.24). As before, the empirical relative efficiency converges to the asymptotic relative efficiency.

4.7 An elaborated example

4.7.1 Control limits for the exact EWMA control chart

The EWMA statistic is defined in (4.1). Assume the sequence $\{X_t\}$ consists of AR(2) observations. If t is sufficiently large, the variance of W_t is given by (4.9). Hence, the control limits

for the EWMA control chart for AR(2) data are given by

$$\begin{aligned} \text{UCL} &= \mu + c\sqrt{\alpha}, \\ \text{LCL} &= \mu - c\sqrt{\alpha}, \end{aligned}$$

where c is a constant, e.g. $c = 2.86$, to be chosen by the designer of the control chart. Furthermore, we have to choose λ and provide estimates for ϕ_1, ϕ_2, μ, and σ_A^2. The estimator for μ is defined by

$$\hat{\mu} = \bar{X}_n = \frac{1}{n}\sum_{t=1}^{n} X_t.$$

The estimation of the parameters of the AR(2) process ϕ_1, ϕ_2, and σ_A^2 can be done by maximum likelihood estimation. In Appendix 4B we provide additional properties of the variance (4.9). The designer of the control chart has to choose the average run length in the in-control situation (ARL(0)) and the magnitude of the shift (δ) in the mean that it is desired to detect. Based on the ARL(0), δ, and for the given ϕ_1, and ϕ_2, the optimal λ and c can be obtained. We refer to Crowder (1987) for a discussion in the IID case. For an AR(1) process with an ARL(0) of 370 and for λ equal to 0.2 the constant c ranges from 2.43 till 2.96 for all ϕ_1 in the stationary region, see Zhang (2000). Developing a table or function for these values for an AR(2) process will be the subject for further research. For practical applications using the AR(2) processes with ARL(0)=370, $\lambda = 0.2$ and $c = 2.86$ seems to work quite well, especially if we take the uncertainty of the estimation of ϕ_1, and ϕ_2 into account.

Bisgaard and Kulahci (2005) showed that the furnace data fits an AR(2) model well. The maximum likelihood estimates of the parameters are $\hat{\phi}_1 = 0.97$, $\hat{\phi}_2 = -0.36$, and $\hat{\sigma}_A = 0.37$. To obtain the control limits of the modified EWMA control chart assuming the underlying process is AR(2) we substitute the standard maximum likelihood estimates into (4.9). Figure 4.4 shows the resulting modified EWMA control chart using $\lambda = 0.2$ and $c = 2.86$.

We see that the control limits now are inflated so that the EWMA at any point in time is clearly within the control limits, as opposite to the classical EWMA control chart applied in Chapter 1. Thus these limits can be used prospectively to monitor the process going forward, a Phase II application.

4.7.2 Sensitivity analysis of the inflation factor

We define the variance inflation factor IF for large t as the ratio of the asymptotic variance of the EWMA statistic for an AR(2) process relative to the variance of the classical EWMA statistic based on IID observations. Thus, the variance inflation factor is defined by (cf. (4.9), (4.14), and (4.29))

$$\text{IF} = \frac{\phi_1(1+\phi_2)(\lambda-1) + (\phi_2-1)(1+\phi_2(\lambda-1)^2)}{(1+\phi_2)((1-\phi_2)^2-\phi_1^2)(-1+\phi_1(1-\lambda)+\phi_2(\lambda-1)^2)}.$$

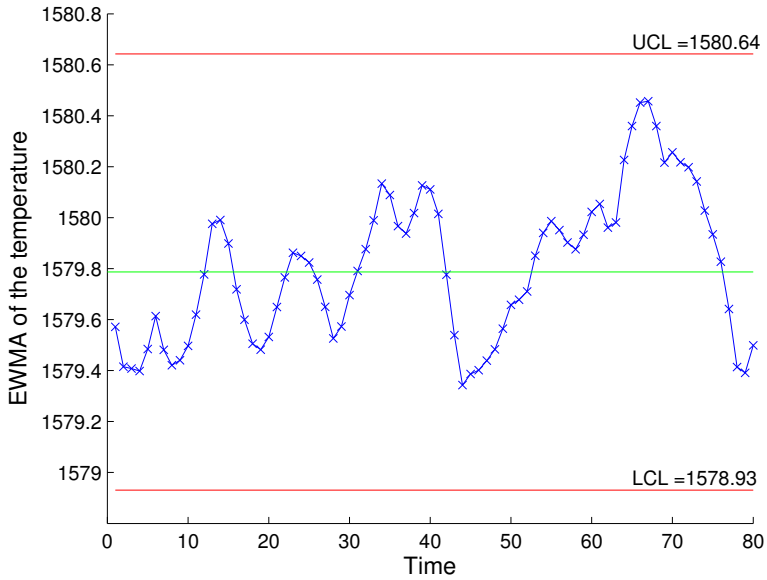

Figure 4.4: The exact EWMA control chart of the furnace temperature data.

The stationarity conditions for an AR(2) process are

$$\phi_2 + \phi_1 < 1$$
$$\phi_2 - \phi_1 < 1$$
$$-1 < \phi_2 < 1$$

Figure 4.5 shows a contour plot of $\sqrt{\mathrm{IF}}$. For $\phi_1 = 0$ and $\phi_2 = 0$ the $\sqrt{\mathrm{IF}} = 1$. As we can see from Figure 4.5, the square root of the inflation factor increases dramatically at the boundaries of the stationarity region. Thus for certain combinations of ϕ_1 and ϕ_2 the inflation factor differs substantially from 1. For these combinations it is particularly important to use the modified EWMA control chart that takes into account the serial correlation.

To study the sensitivity of the inflation factor in the example of the furnace data , we have in Figure 4.6 expanded a part of the contour plot from Figure 4.5. Figure 4.6 shows a 95% confidence region for the parameter estimates for ϕ_1 and ϕ_2. The confidence region is approximately bounded by the contour on the sum of squares surface, see Box et al. (1994), p. 245,

$$\mathrm{SS}(\phi_1, \phi_2) = \mathrm{SS}(\hat{\phi}_1, \hat{\phi}_2)\left[1 + \frac{\chi^2_\varepsilon(\mathrm{df})}{n}\right],$$

where SS the sum of squares of the residuals given the choice of the parameters ϕ_1 and ϕ_2, $\chi^2_\varepsilon(\mathrm{df})$ is the significance point exceeded by a proportion ε of the chi-square distribution with df degrees of freedom, and n is the number of observations. For the present example, $\mathrm{SS}(\hat{\phi}_1, \hat{\phi}_2) = 10.5$, $\chi^2_{0.05}(2) = 5.99$, and $n = 78$ and hence $SS(\phi_1, \phi_2) = 11.3$. Note that $n = 78$ and not 80, because for the first two observations, there are no residuals.

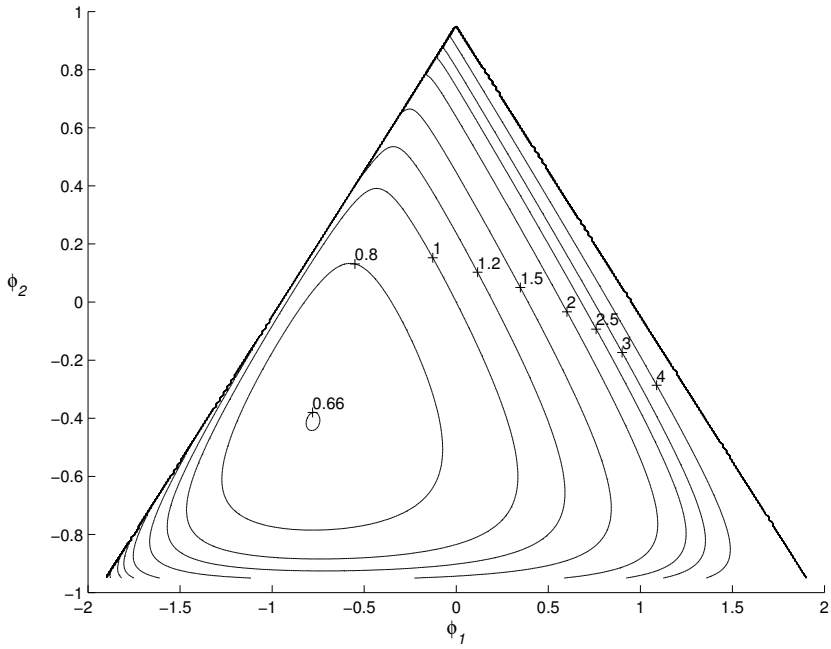

Figure 4.5: The contour plot of the square root of the inflation factor IF

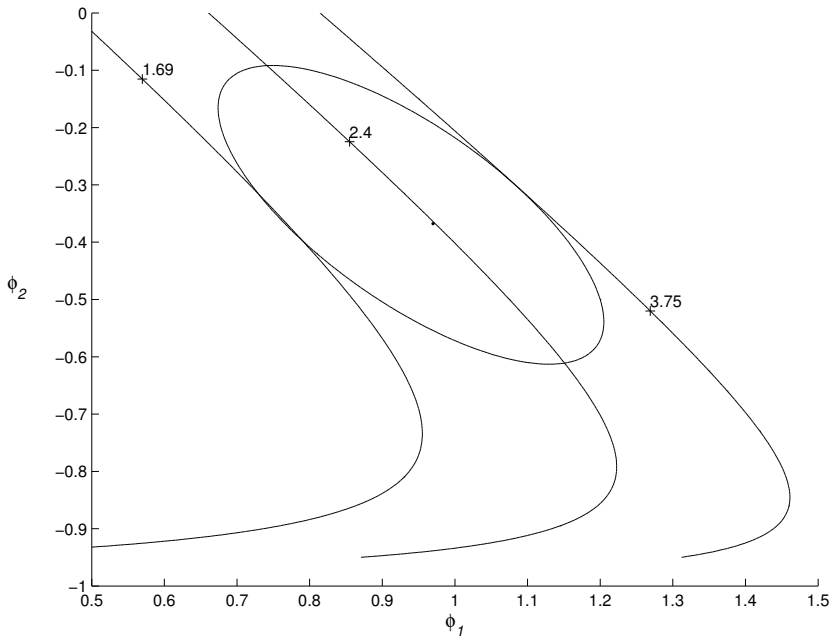

Figure 4.6: A part of the contour plot of the square root of the inflation factor IF with the 95% confidence region of the parameters ϕ_1 and ϕ_2.

As can be seen from this confidence region, the square root of the estimated inflation factor ranges from 1.69 to 3.75. Hence, the number of false alarms using the IF decrease dramatically, but keeps the risk low that a special cause is not observed.

Note that if the roots are outside the unit circle, the process is non-stationary. Consequently

the variance of the AR(2) process and the EWMA statistic will be infinite. In such cases process control is not suitable. An alternative would be to use feedback control, see e.g. Box and Luceño (1997).

Furthermore, note that the special case of an AR(1) process is treated in Schmid (1997), VanBrackle and Reynolds (1997), and Wieringa (1999). In particular note that the formula of the variance of α in (4.9) coincides with the formula of the variance of α for an AR(1) model when we set $\phi_2 = 0$, and coincides with the formula of the variance of α for the IID model when we set $\phi_1 = \phi_2 = 0$, see also Appendix 4B.

4.7.3 Conclusions from the example

The classical EWMA control chart is extremely non-robust to serial correlation. This is illustrated in Chapter 1 with an example of temperature data from a ceramic furnace which can be modelled by an AR(2) process. Autocorrelated observations are common in industry, especially when data are sampled at a high frequency from processes with a lot of inertia.

To construct control charts for autocorrelated processes we may either use specially designed control charts based on time series modeling of the process or use standard control charts with appropriately modified control limits. The latter is not necessarily optimal, but has the appeal of being simple to introduce to engineers already familiar with standard Statistical Process Control (SPC) tools.

4.8 Conclusion

We have derived the asymptotic distributions for the three estimators for the variance of the EWMA statistic introduced by Zhang (1998), Schmid (1997) and Vermaat et al. (2006a), where the estimator of Schmid is a special case of the estimator of Vermaat et al. (2006a) for an AR(1) process. The asymptotic relative efficiency of the estimator by Zhang (1998) is compared to those by Schmid (1997) and Vermaat et al. (2006a). It appears that the estimators of Schmid (1997) and Vermaat et al. (2006a) are more efficient for AR(1) respectively AR(2) processes. Hence, for processes that can be modelled by an AR(1) or an AR(2) process, we would advocate the EWMA control chart based on Schmid's estimator and the estimator based on Vermaat et al. (2006a) respectively.

Appendix 4A: Derivation of the variance of the EWMA statistic

From Box et al. (1994), p. 27, we have that

$$\begin{aligned}
\alpha_t = \mathrm{Var}(W_t) &= \lambda^2 \sum_{i=0}^{t-1}\sum_{j=0}^{t-1}(1-\lambda)^{i+j}\gamma(|j-i|). \\
&= 2\lambda^2 \sum_{i=0}^{t-2}\sum_{j=i+1}^{t-1}(1-\lambda)^{i+j}\gamma(j-i) + \lambda^2\sum_{i=0}^{t-1}(1-\lambda)^{2i}\gamma(0) \\
&= 2\lambda^2 \sum_{i=0}^{t-2}\sum_{j=i+1}^{t-1}(1-\lambda)^{i+j}\gamma(j-i) + \gamma(0)\left(\frac{\lambda}{2-\lambda}\right)[1-(1-\lambda)^{2t}].
\end{aligned}$$

As noted by Zhang (1998) and Schmid (1997) this expression can easily be rewritten as

$$\alpha_t = \frac{\lambda}{2-\lambda}\left[\Gamma_t + \mathrm{E}_t\right],$$

where

$$\begin{aligned}
\Gamma_t &= 2\sum_{k=1}^{t-1}\gamma(k)(1-\lambda)^k(1-(1-\lambda)^{2(t-k)}) \\
&= 2\left[\sum_{k=1}^{t-1}\gamma(k)(1-\lambda)^k - \sum_{k=1}^{t-1}\gamma(k)(1-\lambda)^{2t-k}\right], \\
\mathrm{E}_t &= \gamma(0)(1-(1-\lambda)^{2t}).
\end{aligned}$$

By substituting the covariance from (4.6) into Γ_t and after some tedious but simple algebra we get

$$\begin{aligned}
\Gamma_t &= 2(\nu_1-\nu_2)^{-1}\Bigg\{\sum_{k=1}^{t-1}(1-\lambda)^k\nu_1^k\gamma(1) - \sum_{k=1}^{t-1}(1-\lambda)^k\nu_1^k\nu_2\gamma(0) \\
&\quad - \sum_{k=1}^{t-1}(1-\lambda)^k\nu_2^k\gamma(1) + \sum_{k=1}^{t-1}(1-\lambda)^k\nu_2^k\nu_1\gamma(0) - \sum_{k=1}^{t-1}(1-\lambda)^{2t-k}\nu_1^k\gamma(1) \\
&\quad + \sum_{k=1}^{t-1}(1-\lambda)^{2t-k}\nu_1^k\nu_2\gamma(0) + \sum_{k=1}^{t-1}(1-\lambda)^{2t-k}\nu_2^k\gamma(1) - \sum_{k=1}^{t-1}(1-\lambda)^{2t-k}\nu_2^k\nu_1\gamma(0)\Bigg\} \\
&= 2(\nu_1-\nu_2)^{-1}\left\{(\gamma(1)-\nu_2\gamma(0))f(\nu_1) - (\gamma(1)-\nu_1\gamma(0))f(\nu_2)\right. \\
&\qquad \left. -(\gamma(1)-\nu_2\gamma(0))g(\nu_1) + (\gamma(1)-\nu_1\gamma(0))g(\nu_2)\right\},
\end{aligned}$$

where the functions $f(x) = \sum_{k=1}^{t-1}(1-\lambda)^k x^k$ and $g(x) = \sum_{k=1}^{t-1}(1-\lambda)^{2t-k}x^k$. Now using the fact that $\sum_{k=1}^{n} r^k = r(1-r^n)/(1-r)$, $f(x)$ and $g(x)$ can be rewritten as

$$f(x) = \frac{(1-\lambda)^t x^t - (1-\lambda)x}{(1-\lambda)x - 1}$$

and

$$g(x) = \frac{(1-\lambda)^{2t}x - (1-\lambda)^{t+1}x^t}{1-\lambda-x}.$$

Substituting these expressions into Γ_t we get after some straightforward simplifications that

$$\Gamma_t = 2(\nu_1-\nu_2)^{-1}\left\{\lambda(\lambda-2)(\mathrm{A}_t-\mathrm{B}_t)+\mathrm{C}_t+\mathrm{D}\right\},$$

where

$$\begin{aligned}
\mathrm{A}_t &= \frac{\nu_2^{t+1}(1-\lambda)^t(\gamma(1)-\nu_1\gamma(0))}{(\lambda+\nu_2-1)((1-\lambda)\nu_2-1)},\\
\mathrm{B}_t &= \frac{\nu_1^{t}(1-\lambda)^t(\gamma(1)-\nu_2\gamma(0))}{(\lambda+\nu_1-1)((1-\lambda)\nu_1-1)},\\
\mathrm{C}_t &= (1-\lambda)^{2t+1}(\nu_1-\nu_2)\left(\frac{-\gamma(1)+\gamma(0)\nu_1\nu_2(1-\lambda)^{-1}}{(\lambda+\nu_1-1)(\lambda+\nu_2-1)}\right),\\
\mathrm{D} &= (1-\lambda)(\nu_1-\nu_2)\left(\frac{\gamma(1)-\gamma(0)\nu_1\nu_2(1-\lambda)}{((1-\lambda)\nu_1-1)((1-\lambda)\nu_2-1}\right).
\end{aligned}$$

From this it follows that

$$\begin{aligned}
\alpha_t &= \frac{\lambda}{2-\lambda}\left[\Gamma_t+\mathrm{E}_t\right]\\
&= \frac{\lambda}{2-\lambda}\left[2(\nu_1-\nu_2)^{-1}(\lambda(\lambda-2)(\mathrm{A}_t-\mathrm{B}_t)+\mathrm{C}_t+\mathrm{D})+\mathrm{E}_t\right].
\end{aligned}$$

Now letting $t\to\infty$ it can be seen that $\mathrm{A}_t, \mathrm{B}_t$ and C_t goes to zero. We are then left with,

$$\begin{aligned}
\alpha = \lim_{t\to\infty}\alpha_t &= \frac{2\lambda(\lambda-1)(\gamma(1)+\gamma(0)\nu_1\nu_2(\lambda-1))}{(\lambda-2)(1+\nu_1(\lambda-1))(1+\nu_2(\lambda-1))}+\frac{\gamma(0)\lambda}{2-\lambda}\\
&= \frac{\lambda}{2-\lambda}\left(\frac{2(1-\lambda)(\gamma(1)+\gamma(0)\nu_1\nu_2(\lambda-1))}{(1+\nu_1(\lambda-1))(1+\nu_2(\lambda-1))}+\gamma(0)\right).
\end{aligned}$$

Finally substituting the expressions for $\gamma(0)$ and $\gamma(1)$ from (4.4) and (4.5) and ν_1 and ν_2 from (4.12) we get the desired asymptotic expression for the variance as in (4.9),

$$\alpha = \frac{\lambda}{2-\lambda}\left(\frac{\phi_1(1+\phi_2)(\lambda-1)+(\phi_2-1)(1+\phi_2(\lambda-1)^2)}{(1+\phi_2)((1-\phi_2)^2-\phi_1^2)(-1+\phi_1(1-\lambda)+\phi_2(\lambda-1)^2)}\right)\sigma_A^2 \quad (4.25)$$

$$(4.26)$$

$$= \frac{\lambda}{2-\lambda}\left(\frac{\phi_1(1+\phi_2)(\lambda-1)+(\phi_2-1)(1+\phi_2(\lambda-1)^2)}{(1-\phi_2)(-1+\phi_1(1-\lambda)+\phi_2(\lambda-1)^2)}\right)\sigma_X^2. \quad (4.27)$$

Appendix 4B: Properties of the variance of the EWMA statistic

In this appendix we provide several properties of the variance of the EWMA statistic in (4.25) and in (4.9).

Property 4.1 *If $\phi_2 = 0$ we have an AR(1) process. The variance of the EWMA statistic, α_t given by (4.9) for large t reduces to*

$$\alpha = \frac{\lambda}{2-\lambda}\left(\frac{\sigma_A^2}{1-\phi_1^2}\right)\frac{1+\phi_1(1-\lambda)}{1-\phi_1(1-\lambda)}. \tag{4.28}$$

See Schmid (1997), VanBrackle and Reynolds (1997), and Wieringa (1999).

Property 4.2 *If $\phi_1 = \phi_2 = 0$ the AR(2) process reduces to an independent and identically distributed random noise process and the variance of the EWMA statistic α_t given by (4.9) for large t reduces to the well known standard expression*

$$\alpha = \frac{\lambda}{2-\lambda}\sigma_A^2. \tag{4.29}$$

Property 4.3 *If $\lambda = 1$ in the expression for the EWMA statistic, then the EWMA statistic reduces to X_t, simply the AR(2) process itself. In other words, the EWMA has no memory and the variance is the same as the process variance*

$$\alpha = \frac{1-\phi_2}{1+\phi_2}\left(\frac{\sigma_A^2}{(1-\phi_2)^2-\phi_1^2}\right). \tag{4.30}$$

This result can also be found in Box et al. (1994), p. 62.

References

Albers, C.J. (2003). *Distributional Inference: The Limits of Reason*, Ph.D. Thesis, University of Groningen, The Netherlands.

Albers, C.J. and Schaafsma, W. (2003). Estimating a density by adapting an initial guess. *Computational Statistics & Data Analysis*, Vol. 42, 27-36.

Alwan, L.C. (1991). Autocorrelation: fixed versus variable control limits. *Quality Engineering*, Vol. 4(2), 167-188.

Alwan, L.C. and Roberts, H.V. (1988). Time-series modeling for statistical process control. *Journal of Business & Economic Statistics*, Vol. 6(1), 87-95.

Apley, D.W. and Lee, H.C. (2003). Design of exponentially weighted moving average control charts for autocorrelated processes with model uncertainty. *Technometrics*, Vol. 45(3), 187-198.

Azzalini, A. (1981). A note on the estimation of a distribution function and quantiles by a kernel method. *Biometrika*, Vol. 68, 326-328.

Bai, D.S. and Choi, I.S. (1995). $\bar{X}$ and R control charts for skewed populations. *Journal of Quality Technology*, Vol. 27(2), 120-131.

Barnard, G.A. (1959). Control charts and stochastic processes. *Journal of the Royal Statistical Society*, Vol. 21(2), 239-271.

Berthouex, P.M., Hunter, W.G., and Pallesen, L. (1978). Monitoring sewage treatment plants: some quality control aspects. *Journal of Quality Technology*, Vol. 10(4), 139-149.

Bisgaard, S. and Kulahci, M. (2005). Quality Quandaries: The effect of autocorrelation on statistical process control procedures. *Quality Engineering*, Vol. 17(3), 481-489.

Borror, C.M., Montgomery, D.C., and Runger, G.C. (1999). Robustness of the EWMA control chart to non-normality. *Journal of Quality Technology*, Vol. 31(3), 309-316.

Box, G.E.P. and Cox, D.R. (1964). An analysis of transformations. *Journal of the Statistical Society. Series B (Methodological)*, Vol. 26(2), 211-252.

Box, G.E.P., Jenkins, G.M., and Reinsel, G.C. (1994). *Time Series Analysis*, 3rd ed. Prentice Hall, Englewood Cliffs, New Jersey.

Box, G.E.P. and Luceño, A. (1997). *Statistical Control by Monitoring and Feedback Adjustment*, John Wiley & Sons, New York.

Brockwell, P.J. and Davis, R.A. (1991). *Time Series: Theory and Methods*, Springer-Verlag, New York.

Burr, IW (1967). The effect of non-normality on constants for X and R charts. *Industrial Quality Control,* Vol. 24, 563-569.

Carter, P.L. (1972). A Bayesian approach to quality control. *Management Science*, Vol. 18(11), 647-656.

Chou, Y-M., Polansky, A.M., and Mason, R.L. (1998). Transforming non-normal data to normality in statistical process control. *Journal of Quality Technology*, Vol. 30(2), 133-141.

Crowder, S.V. (1987). A simple method for studying run-length distributions of exponentially weighted moving average charts. *Technometrics*, Vol. 29(4), 481-489.

Crowder, S.V. (1989). Design of exponentially weighted moving average schemes. *Journal of Quality Technology*, Vol. 21(3), 155-162.

Cryer, J.D. and Ryan, T.P. (1990). The estimation of sigma for an X chart: $\overline{\mathrm{MR}}/d_2$ or S/c_4? *Journal of Quality Technology*, Vol. 22(3), 187-192.

D'Agostino, R.B. and Stephens, M.A. (1986). *Goodness-of-Fit Techniques*, Marcel Dekker Inc., New York.

De Bruin, R., Salomé, D., and Schaafsma, W. (1999). A semi-Bayesian method for non-parametric density estimation. *Computational Statistics & Data Analysis*, Vol. 30, 19-30.

Dekkers, A. L. M., Einmahl, J. H. J., and De Haan, L. (1989). A moment estimator for the index of an extreme-value distribution. *The Annals of Statistics*, Vol. 17(4), 1833-1855.

De Mast, J. (2006). A history of industrial statistics and quality and efficiency improvement. In *Statistical Practice in Business and Industry*, Coleman, S., Montgomery, D., Greenfield, T., and Stewardson, D. eds., John Wiley & Sons, (to appear).

De Mast, J. and Roes, C.B. (2005). Robust individuals control chart for exploratory analysis. *Quality Engineering*, Vol. 16(3), 407-421.

Does, R.J.M.M., Roes, K.C.B., and Trip, A. (1999). *Statistical Process Control in Industry*, Kluwer Academic, Dordrecht, The Netherlands.

Does, R.J.M.M. and Schriever, B.F. (1992). Variables control chart limits and tests for special causes. *Statistica Neerlandica*, Vol. 46(4), 229-245.

Duncan, A.J. (1986). *Quality Control and Industrial Statistics,* 5th ed. Irwin Homewood, Illinois.

Faltin, F.W., Mastrangelo, C.M., Runger, G.C., and Ryan T.P. (1997). Considerations in the monitoring of autocorrelated and independent data. *Journal of Quality Technology,* Vol. 29(2), 131-133.

Feltz, C.J. and Shiau, J-J. H. (2001). Statistical process monitoring using an empirical bayes multivariate process control chart. *Quality and Reliability Engineering International*, Vol. 17, 119-124.

Ferrel, E.B. (1958). Control charts for lognormal universe. *Industrial Quality Control,* Vol. 15, 4-6.

Fuller, W.A. (1996). *Introduciton to Statistical Time Series*, 2nd ed. John Wiley & Sons, New York.

Girard, P. and Parent, E. (2001). Bayesian analysis of autocorrelated ordered categorical data for industrial quality monitoring. *Technometrics*, Vol. 43(2), 180-191.

Girshick, M.A. and Rubin, H. (1952). A Bayes approach to a quality control model. *The Annals of Mathematical Statistics*, Vol. 23(1), 114-125.

Hamada, M. (2002). Bayesian tolerance interval control limits for attributes. *Quality and Reliability Engineering International*, Vol. 18, 45-52.

Ion, R.A. (2001). *Nonparametric Statistical Process Control*, Ph.D. thesis, University of Amsterdam, Amsterdam, The Netherlands.

Ion, R.A. and Klaassen, C.A.J. (2005). Non-parametric Shewhart control charts. *Journal of Nonparametric Statistics*, Vol. 17(8), 971-988.

Jiang, W., Tsui, K-L., and Woodall, W.H. (2000). A new SPC monitoring method: the ARMA chart. *Technometrics*, Vol. 42(4), 399-410.

Johnson, N.L., Kotz, S., and Balakrishnan, N. (1994). *Continuous Univariate Distributions*, Vol. 1, 2nd ed. John Wiley & Sons, New York.

Jones, L.A. and Woodall, W.H. (1998). The performance of bootstrap control charts. *Journal of Quality Technology*, Vol. 30(4), 362-375.

Kamat, A.R. (1953). On the mean successive difference and its ratio to the root mean square. *Biometrika*, Vol. 40, 116-127.

Koehler, A.B., Marks, N.B., and O'Connell, R.T. (2001). EWMA control charts for autoregressive processes. *Journal of the Operational Research Society*, Vol. 52, 699-707.

Liu, R.Y. and Tang, J. (1996). Control charts for dependent and independent measurements based on bootstrap methods. *Journal of the American Statistical Association*, Vol. 91, 1694-1700.

Lu, C-W. and Reynolds, M.R. (1999). EWMA control charts for monitoring the mean of autocorrelated processes. *Journal of Quality Technology*, Vol. 31(2), 166-188.

MacGregor, J.F. (1991). Discussion on the paper of Montgomery, D.C. and Mastrangelo, C.M.. *Journal of Quality Technology*, Vol. 23(3), 198-199.

Mann, N.R. (1988). Why it happened in Japan and not in the U.S.. *Chance,* Vol. 1(3), 3-15.

Menzefricke, U. (2002). On the evaluation of control chart limits based on predictive distributions. *Communications in Statistics: Theory and Methods*, Vol. 31(8), 1423-1440.

Montgomery, D.C. (2004). *Introduction to Statistical Quality Control*, 5th ed. John Wiley & Sons, New York.

Montgomery, D.C. and Mastrangelo, C.M. (1991). Some statistical process control methods for autocorrelated data. *Journal of Quality Technology*, Vol. 23(3), 179-193.

Page, E.S. (1954). Continouos Inspections Schemes. *Biometrika*, Vol. 41(1/2), 100-115.

Parzen, E. (1962). On estimation of a probability density function and mode. *Annals of Mathematical Statistics*, Vol. 33, 1065-1076.

Pearn, W.L. and Kotz, S. (1994). Application of Clemtents' method for calculating second- and third-generation process capability indices for non-normal Pearsonian populations. *Quality Engineering*, Vol. 7(1), 139-145.

Pearson, E.S. (1973). Some historical reflections on the introduction of statistical methods in industry. *The Statistician,* Vol. 22(3), 165-179.

Perez, J.M. and Palacin, A.F. (1987). Estimating the quantile function by Bernstein polynomials. *Computational Statistics & Data Analysis*, Vol. 5, 391-397.

Pollard, D. (2002). *A User's Guide to Measure Theoretic Probability*, Cambridge University Press, United Kingdom.

Quesenberry, C.P. (1993). The effect of sample size on estimated limits for $\bar{X}$ and X control charts. *Journal of Quality Technology*, Vol. 25(4), 237-247.

Quesenberry, C.P. (1997). *SPC Methods for Quality Improvement*, John Wiley & Sons, New York.

Reiss, R.D. (1989). *Approximate Distributions of Order Statistics*, Springer-Verlag, New York.

Reynolds, M.R. Jr and Stoumbos, Z.G. (2001a). Monitoring the process mean and variance using individual observations and variable sampling intervals. *Journal of Quality Technology*, Vol. 33(2), 181-205.

Reynolds, M.R. Jr and Stoumbos, Z.G. (2001b). Individuals control schemes for monitoring the mean and variance of processes subject to drifts. *Stochastic Analysis and Applications*, Vol. 19, 863-892.

Roberts, S.W. (1959). Control charts tests based on geometric moving averages. *Technometrics*, Vol. 1(3), 239-250.

Roes, C.B. (1995). *Shewhart-Type Charts in Statistical Process Control*, Ph.D. Thesis, University of Amsterdam, The Netherlands.

Roes, K.C.B., Does, R.J.M.M., and Jonkers, B. (1999). Effective application of $Q(R)$ charts in low-volume manufaturing. *Quality and Reliability Engineering International*, Vol. 15, 175-190.

Roes, K.C.B., Does, R.J.M.M., and Schurink, Y. (1993). Shewhart-type control charts for individual observations. *Journal of Quality Technology*, Vol. 25(3), 188-198.

Rosenblatt, M. (1956). Remarks on some nonparametric estimates of a density function. *Annals of Mathematical Statistics*, Vol. 27(3), 832-837.

Schilling, E.G. and Nelson, P.R. (1976). The effect of non-normality on the control limits of $\bar{X}$ charts. *Journal of Quality Technology*, Vol. 8(4), 183-188.

Schmid, W. (1997). On EWMA charts for time series. In *Frontiers in Statistical Quality Control*, Lenz, H.-J., Wilrich, P.-Th., eds., Vol. 5, Physica-Verlag, Heidelberg, 115-137.

Schmid, W. (1998). Comments on Zhang (1998). *Technometrics*, Vol. 40(4), 362.

Shewhart, W.A. (1931). *Economic Control of Quality of Manufactured Product*, Van Nostrand, Princeton, New York.

Shore, H. (1998). A new approach to analysing non-normal quality data with application to process capability analysis. *International Journal of Production Research,* Vol. 36(7), 1917-1933.

Stoumbos, Z.G. andReynolds, M.R. Jr. (2000). Robustness to non-normality and autocorrelation of individuals control charts. *Journal of Statistical Computation and Simulation*, Vol. 66, 145-187.

Stratton, B. (editor) (1991). Quality goes to war: an overview. *Quality Progress,* Vol. 24(12), 18-42.

Sturm, G.W., Feltz, C.J., and Yousry, M.A. (1991). An empirical Bayes strategy for analysing manufacuring data in real time. *Quality and Reliability Engineering International*, Vol. 7, 159-167.

Tagaras, G. and Nikolaidis, Y. (2002). Comparing the effectiveness of various Bayesian $\bar{X}$ control charts. *Operations Research*, Vol. 50(5), 878-888.

VanBrackle, L.N. and Reynolds, M.R. Jr. (1997). EWMA and CUSUM control charts in the presence of correlation. *Communications in Statistics: Simulation and Computation*, Vol. 26(3), 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*, Vol. 10(1), 20-30.

Vermaat, M.B. (2003). *A Comparison of Shewhart Individuals Control Charts Based on Normal, Nonparametric and Extreme-Value Theory*, M.Sc. Thesis, University of Groningen, The Netherlands.

Vermaat, M.B. and Does, R.J.M.M. (2006). A semi-Bayesian method for Shewhart individual control charts. *Quality Technology and Quantitative Management*, Vol. 3(1), 111-125.

Vermaat, M.B., Does, R.J.M.M., and Bisgaard, S. (2006a). EWMA control chart limits fo first and second-order autoregressive processes. *Submitted for publication*.

Vermaat, M.B., Does, R.J.M.M., and Steerneman, A.G.M. (2005). A modified quantile estimator using extreme-value theory with applications. *Economic Quality Control*, Vol. 20(1), 31-39.

Vermaat, M.B., Ion, R.A., Does, R.J.M.M., and Klaassen, C.A.J. (2003). A comparison of Shewhart individuals control charts based on normal, non-parametric, and extreme-value theory. *Quality and Reliability Engineering International*, Vol. 19, 337-353.

Vermaat, M.B., Van der Meulen, F.H., and Does, R.J.M.M. (2006b). Asymptotic behavior of the variance of the EWMA statistic for autoregressive processes. *Submitted for publication*.

Wardell, D.G., Moskowitz, H., and Plante, R.D. (1994). Run-length distributions of special-cause control charts for correlated processes. *Technometrics*, Vol. 36(1), 3-17.

Western Electric Company (1956). *Statistical quality control handbook*, Author, Indianapolis.

Wheeler, D.J. (1991). Shewhart's charts: myths, facts, and competitors. *45th annual quality congress transactions ASQC*, 533-538.

Wheeler, D.J. (1995). *Advanced Topics in Statistical Process Control*, SPC Press, Knoxville, Tennessee.

Wieringa, J.E. (1999). *Statistical Process Control for Serially Correlated Data*, Ph.D. Thesis, University of Groningen, The Netherlands.

Willemain, T.R. and Runger, G.C. (1996). Designing control charts using an empirical reference distribution. *Journal of Quality Technology*, Vol. 28(1), 31-38.

Woodall, W.H. and Montgomery, D.C. (1999). Research issues and ideas in statistical process control. *Journal of Quality Technology*, Vol. 31(4), 376-386.

Woodward, P.W. and Naylor, J.C. (1993). An application of Bayesian methods in SPC. *The Statistician*, Vol. 42, 461-469.

Yourstone, S.A. and Zimmer, W.J. (1992). Non-normality and the design of control charts for averages. *Decision Sciences*, Vol. 23, 1099-1113.

Zhang, N.F. (1998). A statistical control chart for stationary process data. *Technometrics*, Vol. 40(1), 24-38.

Zhang, N.F. (2000). Statistical control charts for monitoring the mean of a stationary process. *Journal of Statistical Computation and Simulation*, Vol. 66, 249-258.

Samenvatting

De titel van dit proefschrift luidt “Statistical Process Control in Non-Standard Situations”. Statistical Process Control (SPC) wordt vertaald als statistische procesbeheersing. Dit is een methode om bestaande processen te beheersen. Een belangrijke techniek binnen SPC is de zogenoemde regelkaart (control chart). Sir Walter Shewhart introduceerde de regelkaart in het begin van de 20ste eeuw.

De regelkaart

Productieprocessen worden voortdurend beïnvloed door allerlei factoren die niet of nauwelijks zijn op te sporen en/of weg te nemen. Dit wordt procesinherente variatie genoemd. Een regelkaart wordt gebruikt om variatie door speciale oorzaken te onderscheiden van proces inherente variatie.

Een regelkaart is een grafiek waarin een meetbare karakteristiek van een product of proces op de verticale as wordt uitgezet tegen de tijd op de horizontale as. Als de karakteristiek boven of onder een bepaalde grens komt dan geeft de regelkaart een signaal: het proces is niet langer beheerst. Deze grenzen worden de Upper Control Limit (UCL) en de Lower Control Limit (LCL) genoemd. Het vaststellen van deze grenzen vindt plaats op basis steekproeven die afkomstig zijn uit een beheerst proces. Vervolgens wordt het proces in de tijd gemonitord op basis van deze grenzen. Dit wordt Fase II genoemd, de monitor fase. Regelkaarten kunnen ook gebruikt worden als instrument bij het opsporen van variatie door speciale oorzaken uit het verleden. Deze retrospectieve analyse wordt Fase I genoemd. In dit proefschrift worden alleen de Fase II toepassingen bestudeerd.

Indien een observatie boven de UCL komt of onder de LCL, geeft de regelkaart een signaal: er is een indicatie voor een speciale oorzaak van deze variatie. Als de variatie procesinherente variatie is, en de regelkaart geeft toch een signaal, dan noemen we zo’n signaal een vals alarm. Het aantal observaties totdat een signaal gegeven wordt, wordt de “runlengte”genoemd. Als het proces beheerst is dan willen we dat de runlengte zolang mogelijk is – immers elk signaal is in dat geval een vals alarm – terwijl de runlengte zo kort mogelijk moet zijn bij het optreden van een speciale oorzaak. De runlengtes – die afhangen van de UCL en de LCL – bepalen de prestatie van een regelkaart.

Niet-standaard situaties

De klassieke regelkaarttechnieken veronderstellen dat de procesinherente variatie van de observaties onderling onafhankelijk en identiek normaal verdeeld is. In de praktijk zien we echter veel niet-standaard situaties: de verdeling van de observaties is niet-normaal of de observaties zijn niet onafhankelijk. De klassieke methoden presteren dan niet meer naar behoren in deze situaties. Er is derhalve een noodzaak de technieken te verbeteren voor deze niet-standaard situaties. Dit proefschrift behandelt regelkaarten voor twee niet-standaard situaties

1. de verdeling van de observaties is niet-normaal; regelkaarten voor deze situatie worden besproken in Hoofdstukken 2 en 3.
2. de observaties zijn serieel gecorreleerd; regelkaarten voor deze situatie worden besproken in Hoofdstuk 4.

Niet-normaliteit

In de praktijk blijkt dat observaties vaak een niet-normale verdeling hebben. In Hoofdstukken 2 en 3 worden verschillende regelkaarten voor individuele waarnemingen met elkaar vergeleken waarbij de onderliggende verdelingen variëren van normaal en niet-normaal. De traditionele regelkaarten zijn gebaseerd op het werk van Shewhart uit het begin van de vorige eeuw. Het betreft veelal de individuele regelkaart gebaseerd op het gemiddelde van de “moving ranges”. In Hoofdstuk 2 worden de volgende alternatieven bekeken: niet-parametrische regelkaarten gebaseerd op empirische kwantielen, kernschatters en extreme-waarde theorie. In Hoofdstuk 3 wordt een semi-Bayesiaanse methode bekeken. De semi-Bayesiaanse methode start met een initiële gok over de verdeling van de karakteristiek die beheerst moet gaan worden. Gebaseerd op deze intiële gok en de geobserveerde gegevens kan er een dichtheidsfunctie opgesteld worden door gebruik te maken van een benadering met zogenaamde Bernstein polynomen.

Het gedrag van de regelkaarten wordt bestudeerd met behulp van een simulatiestudie. De klassieke regelkaarten presteert goed voor normaal verdeelde data evenals de alternatieve regelkaarten voor steekproefomvangen van voldoende grootte.

Als de data echter niet meer normaal verdeeld zijn, dan moeten we opzoek naar alternatieven. Als de steekproefomvang in Fase I voldoende groot is (dat wil zeggen meer dan 1000 waarnemingen), dan is de prestatie van de empirische regelkaart uitstekend voor alle soorten verdelingen. Veelal hebben we niet de beschikking over zulke grote data sets en met minder dan 1000 waarnemingen presteert de empirische regelkaart minder goed tot slecht. In de situatie met minder dan 1000, maar wel met meer dan 250 waarnemingen in Fase I, is de prestatie van de semi-Bayesiaanse methode uit Hoofdstuk 3 een goed alternatief. De prestatie van deze methode is het beste van de vergeleken regelkaarten onder niet-normaliteit voor steekproefgroottes van 250 tot 1000. Voor steekproefgroottes groter dan 1000 is de empirische regelkaart de beste. Voor datasets uit Fase I met minder dan 250 waarnemingen is geen van de bestudeerde regelkaarten een goede keus.

Seriële correlatie

Automatische sensoren zorgen voor snel op elkaar volgende metingen van een proces. Hierdoor staat de veronderstelling van onafhankelijke metingen onder druk. Seriële correlatie kan de huidige regelkaarttechnieken enorm beïnvloeden in hun prestatie. In Hoofdstuk 4 geven we uitdrukking die gebruikt kan worden om de bekende "Exponential Weighted Moving Average"(EWMA) regelkaart aan te passen voor Fase II toepassingen, wanneer het onderliggende proces onderhevig is aan autocorrelatie van de order een of twee. In de praktijk blijkt dat we het merendeel van de autogecorreleerde processen met deze twee modellen kunnen modelleren. De nieuwe methode voor het maken van een regelkaart heeft het voordeel (a) dat het een bekende techniek voor de industriële gebruikers en (b) de data worden weergegeven in de originele schaal. In Hoofdstuk 4 wordt deze nieuwe regelkaart vergeleken met drie verschillende schattingsprocedures. Deze vergelijking wordt gedaan op basis van asymptotische relatieve efficiëntie van de schatters. Het blijkt dat de nieuwe methode efficiënter is dan de bestaande methoden voor autogecorreleerde processen.

Curriculum vitae

Thijs Vermaat werd geboren op 6 december 1977 te Leeuwarden. Na zijn middelbare school te Leeuwarden is hij naar Groningen gegaan om daar Econometrie en Operationele Research te studeren. Tijdens deze opleiding had het vakgebied van de statistiek zijn buitengewone interesse. Dit was de reden om in zijn derde studiejaar ook te starten met de opleiding Statistiek. In 1999 studeerde hij een half jaar aan de Université Paul Sabatier in Toulouse, Frankrijk. Na een stage bij het Instituut voor Bedrijfs- en Industriële Statistiek aan de Universiteit van Amsterdam (IBIS UvA) studeerde de auteur van dit proefschrift in 2002 af in de Econometrie en Operationele Research. Daarna trad hij in dienst als adviseur/onderzoeker bij het Instituut voor Bedrijfs- en Industriële Statistiek aan de Universiteit van Amsterdam (IBIS UvA). In 2003 behaalde hij zijn doctoraal in de Statistiek, eveneens aan de Universiteit van Groningen. De afstudeerscriptie van de opleiding Statistiek vormde het begin van het proefschrift dat nu voor u ligt.

Bij het IBIS UvA houdt de auteur zich bezig met onderzoek en advisering in de industriële statistiek. Naast deze twee taken heeft de auteur in het academische jaar 2004/2005 het Mastervak "Industrial Statistics" aan de Universiteit van Amsterdam verzorgd. Het onderzoek bij het IBIS UvA mondde uit in dit proefschrift, alsmede een aantal publicaties over Lean Six Sigma. Advisering in de industriële statistiek bij het IBIS UvA concentreert zich met name tot de begeleiding bij de invoering van verbetermethodiek Lean Six Sigma. De betroffene gaf training en ondersteuning aan meer dan 300 projectleiders/ medewerkers bij het uitvoeren van hun Lean Six Sigma projecten. De auteur is in dit kader actief geweest voor onder meer de volgende bedrijven: het Rode Kruis Ziekenhuis te Beverwijk, het Virga Jesseziekenhuis te Hasselt (België), het Canisius Wilhelmina Ziekenhuis te Nijmegen, de ABN AMRO bank (Business Unit Nederland en Business Unit Services Operations Transaction Banking), Getronics te Amsterdam, Burgers Ergon te Eindhoven en Apeldoorn en Noviant te Nijmegen.

De auteur is lid van de Vereniging voor Statistiek en Operationele Research (VVS-OR) en het European Network for Business and Industrial Statistics (ENBIS).