



UvA-DARE (Digital Academic Repository)

Statistical control of Shewhart control charts

Goedhart, R.

Publication date

2018

Document Version

Final published version

License

Other

[Link to publication](#)

Citation for published version (APA):

Goedhart, R. (2018). *Statistical control of Shewhart control charts*. [Thesis, fully internal, Universiteit van Amsterdam].

General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.



Statistical Control of Shewhart Control Charts

Rob Goedhart

Statistical Control of Shewhart Control Charts

Rob Goedhart



Statistical Control of Shewhart Control Charts

Rob Goedhart



Instituut voor Bedrijfs- en Industriële Statistiek

Publisher IBIS UvA, Amsterdam
Printed by Gildeprint, Enschede
ISBN 978-94-9301-463-3

Statistical Control of Shewhart Control Charts

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Universiteit van Amsterdam
op gezag van de Rector Magnificus
prof. dr. ir. K.I.J. Maex
ten overstaan van een door het College voor Promoties
ingestelde commissie, in het openbaar te verdedigen
in de Agnietenkapel
op woensdag 17 oktober 2018, te 12:00 uur

door

Rob Goedhart

geboren te Eindhoven

Promotiecommissie

Promotor

Prof. dr. R. J. M. M. Does

Universiteit van Amsterdam

Co-promotor

Dr. M. Schoonhoven

Universiteit van Amsterdam

Overige leden

Prof. dr. S. Chakraborti

University of Alabama

Prof. dr. C. G. H. Diks

Universiteit van Amsterdam

Prof. dr. C. A. J. Klaassen

Universiteit van Amsterdam

Prof. dr. M. R. H. Mandjes

Universiteit van Amsterdam

Prof. dr. M. Salomon

Universiteit van Amsterdam

Prof. dr. J. Tuinstra

Universiteit van Amsterdam

Dr. I. M. Zwetsloot

City University of Hong Kong

To my beloved sister Caroline

CONTENTS

1 Introduction	1
1.1 Statistical Process Monitoring	2
1.2 Shewhart Control Charts	4
1.3 The Effect of Parameter Estimation	6
1.4 Design of Control Charts	7
1.5 Outline and Scientific Contribution	9
2 Bias Criterion for Control Charts for Location	11
2.1 Introduction	11
2.2 Model and Approach	13
2.3 Albers and Kallenberg's Correction Factors	15
2.4 New Correction Factors	19
2.5 Implications of Proposed Corrected Charts	25
2.6 Application of the Proposed Control Chart	32
2.7 Concluding Remarks	35
2.8 Appendix: Albers and Kallenberg Correction	36
2.9 Appendix: Correction Factor	37
2.10 Appendix: Subgroup Control Chart	38
2.11 Appendix: Individuals Control Chart	39
3 Exceedance Probability Criterion for Control Charts for Location	41
3.1 Introduction	41
3.2 Model and Approach	42
3.3 Correction Terms	46
3.4 Out-of-Control Performance	56
3.5 Comparison with Existing Methods	59
3.6 Extension of Tolerance Interval Theory	66
3.7 Concluding Remarks	69
3.8 Appendix A	70
3.9 Appendix B	72

4 Control Charts for Dispersion	75
4.1 Introduction	75
4.2 Determination of the Adjusted Control Limit	77
4.3 Out-of-Control Performance	88
4.4 Balancing In-Control and Out-of-Control Performance	95
4.5 Practical Example	97
4.6 Concluding Remarks	98
5 Nonparametric Control Charts	101
5.1 Introduction	101
5.2 Nonparametric Control Limits	103
5.3 Alternative Parametric Methods	111
5.4 Concluding Remarks	119
6 Summary	121
References	126
Samenvatting	133
Acknowledgements	137
Curriculum Vitae	139

1. Introduction

Data availability has increased immensely in the past years, and so has the need for data analysis techniques. A key point of interest is often to use process data to detect changes in the underlying process. This applies to numerous environments, ranging from standard manufacturing processes to intelligence agencies using complex network data to detect possible terrorist cells, or even our own body. Even though many processes may appear constant at first glance, the corresponding process data will vary over time. If we look at our own body for example, our heart rate, blood glucose, and many other characteristics are not constant values. Certain variation in these values is inherent to the process under consideration, and pinpointing the exact cause of these differences is often very difficult, if not impossible.

However, special events or disturbances can change the underlying process, bringing a different source of variation. For example, one can think of machine wear or defects in a manufacturing environment, or a virus or disease affecting ourselves, each influencing certain processes and the corresponding data. If no corrective actions are taken, this may lead to undesirable and potentially harmful consequences, depending on the circumstances. Moreover, detection of these changes and the underlying causes may provide valuable information that can be used for process improvement. This is common practice in many improvement methodologies, such as Lean Six Sigma. The field of statistical process monitoring (SPM) provides tools to detect process changes by monitoring data streams.

This dissertation revolves around the design of one of such tools, namely the Shewhart control chart. In this chapter, an overview is given of the concepts and motivation for this dissertation. In particular, Section 1.1 provides the general concept and applications of SPM, after which the Shewhart control chart is discussed in Section 1.2. Next, in Section 1.3, the effect of parameter estimation on control chart performance is discussed and illustrated. Section

1.4 provides some details on the directions and tradeoffs to be made to take this effect into account. Finally, Section 1.5 contains the outline of this dissertation.

1.1 Statistical Process Monitoring

In SPM, statistical techniques are applied to monitor processes for the purposes of process control as well as improvement. The origins of SPM date from the early 1920s, when Walter A. Shewhart worked at Bell Laboratories. Shewhart recognized the importance of reducing variation in manufacturing processes, and pointed out that variation actually increased when process adjustments were made based on the number of non-conforming products. While every process shows variation, he made a distinction between two sources of variation, non-assignable (chance cause) and assignable cause variation. Nowadays, these are mostly referred to as *common cause* and *special cause* variation, respectively.

Common cause variation is inherent to the process under consideration. This variation is often a combination of many small influences, where it is nearly impossible to determine and control all the exact causes. Examples of common cause variation are the outcomes of a regular dice roll, or a roulette wheel at the casino. Many factors influence these outcomes, such as the starting position and strength of the roll or spin. Yet, even when we feel like we are repeating the same actions, we (or at least most of us) are not able to produce the same outcome every time. It is an aggregation of minor factors, constantly active within the process, of which the outcome is probabilistically predictable. In statistics, this variation is often referred to as *noise*. A process that displays only common cause variation is addressed in SPM as being *statistically in control*, henceforth *in-control*.

Special cause variation is variation induced by a specific, assignable cause. This variation can be observed in different forms, such as trends, drifts, shifts, or outliers, and indicates some change(s) or disturbances in the process. A process that incorporates this type of variation is referred to as *statistically out of control*, henceforth *out-of-control*. For example, in a manufacturing process one can think of defects, changes in operator(s), machines, or software, but also of changes that may not be as visible at first glance, such as machine wear leading to a slow deterioration of output, or problems with input mate-

rials. These changes can have substantial consequences for the process under consideration, and it is therefore important to detect them so that adequate counter measures can be taken. Even when such changes do not immediately impact the process output or performance, detecting this form of variability yields valuable information as they indicate influence factors of the outcome variable under consideration.

To this end, Shewhart introduced a statistical tool, the *control chart*, to monitor process data. The control chart is a commonly used tool in quality control, and is designed to provide a *signal* when it suspects an out-of-control situation. The foundation for the control chart, and for statistical quality control in general, was laid down by Shewhart in his earlier works, Shewhart (1926) and Shewhart (1931). Originally, control charts were mainly applied in manufacturing processes. As an example, consider a factory that fills bottles with liquid. Due to common causes of variation, different bottles will not contain exactly the same volume of the corresponding liquid. If the volume is too low, this may result in rework because the bottle has to be processed again. If the volume is too high, this may be favorable to the customer, but actually means a waste of the liquid for the selling company. For this company, it is therefore of major importance to control the contents of the bottles. If any of the influence factors for the process output changes (e.g. the operator, input material, machine settings, etc.), the distribution of the output data will change. If these changes can be detected before they actually lead to non-conforming products, corrective actions can be taken in a timely manner to prevent rework or waste of products.

SPM provides the statistical tools and properties for monitoring data streams. The entire control loop, from monitoring to acting upon out-of-control signals, is generally referred to as statistical process control (SPC). While originally mainly used for quality control in manufacturing processes, the (potential) areas of application of SPM and SPC have increased in recent times. In health care, a well known example is *Google Flu Trends*, which was designed to detect, or even predict, outbreaks of flu by monitoring search queries of Google users. Using this information can aid in early detection of such outbreaks, allowing for a quick response that may limit their impact. The underlying principle of detecting changes from data streams is not limited to specific environments. With the rapid increase in data availability in general, the tools to turn these data into information become more and more important

in many fields.

1.2 Shewhart Control Charts

As described in the previous section, the goal of SPM is to detect changes in the underlying process. In order to detect these changes, a distinction between common and special causes of variation has to be made. Generally, it is unknown in advance which variation is common. To this end, various statistical techniques have been developed, including the control chart. A control chart consists of a *plotting statistic*, also referred to as monitoring statistic or charting statistic, as well as *control limits*. The plotting statistic is some variable of interest of the process under consideration, while the control limits determine for which values of this variable an out-of-control signal is produced. These control limits are determined based on the statistical properties of the underlying process data.

The most common control chart is the Shewhart control chart, which uses an upper control limit (UCL) and a lower control limit (LCL). These limits are based on the process mean (μ_0) and standard deviation (σ_0) when in-control, and are originally of the simple form

$$\begin{aligned}UCL &= \mu_0 + 3\sigma_0 \\LCL &= \mu_0 - 3\sigma_0.\end{aligned}\tag{1.1}$$

Under the assumption of normally distributed data, this set of control limits yields a 0.27% chance of a signal for each plotting statistic when the process is actually in-control. The likeliness of producing a signal when no assignable causes are present is referred to as the *false alarm rate* (FAR), and is denoted by α_0 in general. As an illustration, Figure 1.1 shows the UCL and LCL in relation to the normal distribution. Because the monitoring of a plotting statistic with these control limits is an ongoing process, an interesting property of the control chart is the *run length*, which is the number of plotting statistics until a signal is obtained. When the process is in-control, every plotting statistic has a probability of α_0 of producing a signal. The run length of the control chart then follows a geometric distribution with parameter α_0 . Consequently, the in-control average run length (ARL) of the control chart equals $1/\alpha_0$ in general, so in this case 370.4. The FAR and ARL are the most commonly used

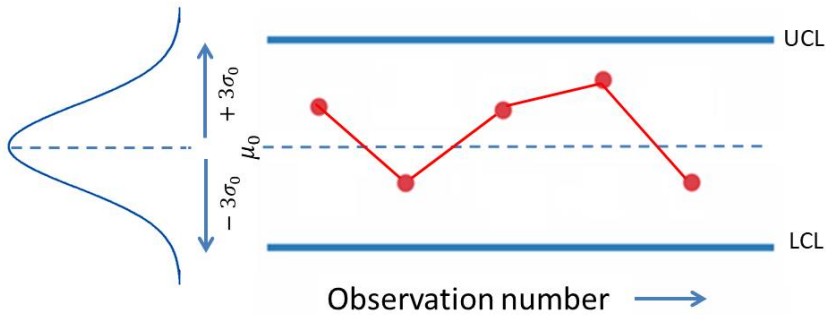


Figure 1.1: *Shewhart control chart and the normal distribution*

characteristics to evaluate the performance of a control chart.

Although the control chart is designed to detect out-of-control situations, the in-control performance should not be overlooked. The in-control and out-of-control performance are related in a similar way as the type I and type II error are in statistical hypothesis tests. One could increase the out-of-control detection rate by narrowing down the control limits, and thus increasing the probability of a signal, but this increases the FAR. On the other hand, lowering the FAR by widening the control limits leads to worse detection rates. A high FAR can lead to a waste of time and resources looking for assignable causes that aren't present. Note that when a process is out-of-control, the actual probability of a signal (which is no longer a false signal), is dependent on the type and size of the change in the underlying process, and can therefore not be determined without further prior information on the change. For this reason, control chart performance is generally evaluated in terms of in-control properties, such as the FAR and in-control ARL.

In practice, the in-control parameters μ_0 and σ_0 are generally unknown, and have to be estimated. Therefore, SPM distinguishes between two phases, Phase I and Phase II. In Phase I the aim is to determine the in-control behavior of the data, including the required process parameters. This involves a Phase I reference sample, consisting of m subgroups of size n each. Here $n = 1$ means that data are collected as individual observations, while $n > 1$ means that multiple observations are collected at each time instance. Because the estimation in Phase I affects the control chart performance in Phase II, a thorough investigation of possible out-of-control situations or data contamination

is required for the obtained sample. Details on robust estimation methods to deal with this problem can be found in Schoonhoven (2011). Phase II is the monitoring phase, where the plotting statistic is prospectively monitored using the control limits obtained in Phase I.

1.3 The Effect of Parameter Estimation

Although robust and efficient estimation is an important first step, this alone does not solve all of the problems associated with parameter estimation in Phase I. An influential paper in the field of SPM to this end is Quesenberry (1993), who pointed out the dependency between consecutive false alarm probabilities for plotting statistics when parameters are estimated. This means that, in contrast to the known parameters case, the run length distribution is not geometric. He suggests the use of at least $400/(n - 1)$ Phase I subgroups in order for the control charts to behave like the known parameters situation when $n > 1$, and around 300 observations when $n = 1$. In Jensen et al. (2006), these and other findings are discussed in an overview paper on the effects of parameter estimation in SPM.

However, Quesenberry (1993) neglected an important effect of parameter estimation, namely the *practitioner-to-practitioner* variation. This represents the fact that, even when different practitioners sample from the same distribution, they will obtain different samples and therefore different parameter estimates. Consequently, their estimated control limits and the corresponding control chart performance will differ as well. As an illustration, we have simulated 1,000 Phase I samples from a standard normal distribution, consisting of 50 subgroups of size 5 each. For each of these samples, we have constructed the control limits using estimated parameters. Since the actual distribution of the underlying process is known, it is possible to calculate the FAR and the ARL for each chart, conditionally on the corresponding estimates. Figure 1.2 shows a histogram of the 1,000 obtained conditional ARL values. As can be seen, the differences between these values are substantial, and the values may be very different from the design (nominal) value of 370.4. Note that the variation becomes smaller as more data are collected, as this increases the estimation accuracy. A detailed paper on the effect of parameter estimation for Shewhart \bar{X} control charts is Saleh et al. (2015b). They illustrate the conditional performance of Shewhart \bar{X} control charts and show that the earlier

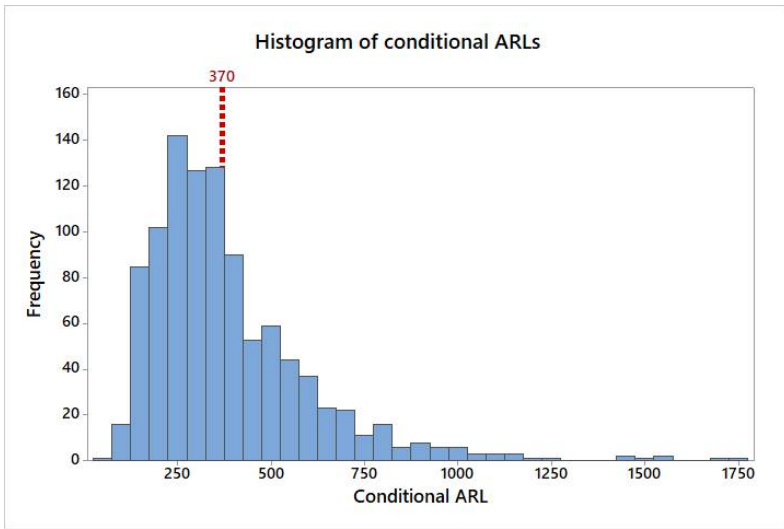


Figure 1.2: *Effect of parameter estimation*

sample size requirements proposed by Quesenberry (1993) are not sufficient.

1.4 Design of Control Charts

In many occasions the required sample sizes for a sufficient control chart performance are not available, which should be taken into consideration in the design and evaluation of control charts. In order to compensate or correct for the effect of parameter estimation, a solution is to adjust the control limits such that a certain performance criterion is matched. Although many specific criteria are thinkable, there are two general directions, namely unconditional and conditional performance. Albers and Kallenberg (2004a,b, 2005) denote the criteria corresponding to these directions as the *bias criterion* and the *exceedance probability criterion* respectively.

Originally, performance measures to evaluate control chart performance consisted of unconditional performance measures such as the unconditional FAR or the unconditional ARL. These are the expected values of the FAR and ARL. When adjusting the control limits to match a certain unconditional performance, the most logical choice is to provide a specified nominal performance in expectation. The performance measure is generally some function $g(\alpha_0)$ of the nominal FAR value α_0 , see also Albers and Kallenberg (2004b).

Relating to Figure 1.2, the function under consideration is the ARL (e.g. $g(\alpha_0) = 1/\alpha_0$), such that the expectation of $g(\alpha_0)$ can be thought of as the average of all the conditional ARL values. In that case, using the bias criterion means that control limits should be derived such that the expectation of the ARL (EARL) equals a specified nominal value $EARL_0 = 1/\alpha_0$. In Chapter 2 we elaborate further on the bias criterion and the required adjustments to the control limits to satisfy it.

When focusing on the expected performance, there may still be a large proportion of the practitioners with an unsatisfactory control chart performance. The second option is then to focus on conditional performance by means of the exceedance probability criterion. The idea is to guarantee a minimum conditional in-control control chart performance (often in terms of FAR or ARL) to practitioners with a pre-specified probability. In Chapters 3 to 5 we elaborate further on this criterion.

Another aspect to consider in the design of control limits is their functional form, and the corresponding underlying distributional assumptions. The limits described in (1.1) are derived for normally distributed data, and it is obvious that this form of control limits will not work as well for other distributions. Especially for skewed data, the symmetrical character of these control limits will cause issues with control chart performance. A first way to overcome this is by introducing probability limits, as is done for example for control charts for dispersion when the data itself are normally distributed (e.g. Montgomery, 2013). These limits are based on the true distribution of the characteristic of interest. Such parametric control limits can also be derived for other distributions in combination with the bias or the exceedance probability criterion. However, although these parametric control limits provide a solution for various data distributions, in practice the problem is that the true distribution is generally unknown, and has to be estimated along with its parameters. This leads to an extra source of estimation error, and introduces an extra tradeoff in the design of control charts.

As mentioned also in Albers et al. (2004), the total estimation error can be split up in two different distinct errors, the *model error* (ME) and the *stochastic error* (SE). The first is caused by incorrect assumptions on the distributional form, while the latter is the error resulting from parameter estimation. In order to reduce the reliance on distributional assumptions, one can consider the use of nonparametric methods. In that case, the ME will vanish and the

variation will be caused by the SE only. While the first is dependent on the distribution under consideration, the latter can be reduced by collecting larger samples.

These tradeoffs and adjustments for the design of Shewhart control charts when parameters are estimated are the topic of this dissertation. However, the findings of this dissertation are more generally applicable, as the tradeoffs are also relevant in many other statistical applications.

1.5 Outline and Scientific Contribution

In this dissertation we elaborate on the Shewhart control chart, and provide various new designs to account for the effect of parameter estimation.

In Chapter 2 control limits for Shewhart control charts for location are derived that provide a specified control chart performance in expectation, following the bias criterion. In this chapter the main focus is placed on the ARL, but the results can be generalized to other characteristics easily, because the derivations contain a general function for performance measures. The chapter also contains a practical application as an illustration.

This chapter has been published under the title “Correction Factors for Shewhart X and \bar{X} Control Charts to Achieve Desired Unconditional ARL” in the *International Journal of Production Research*. This paper, Goedhart et al. (2016), was combined work with dr. M. Schoonhoven and prof. dr. R.J.M.M. Does, in which I took the lead regarding the mathematical derivations, M. Schoonhoven took the lead in writing, and R.J.M.M. Does provided supervision.

In Chapter 3, control limits for Shewhart control charts for location are derived following the exceedance probability criterion. Moreover, a comparison of these limits with other existing methods based on a similar criterion is provided. Among these methods are self-starting control charts, a bootstrap procedure, and tolerance intervals.

This chapter is based on two papers, Goedhart et al. (2017b) and Goedhart et al. (2018a), both published in the *Journal of Quality Technology*. The first paper, entitled “Guaranteed In-Control Performance for the Shewhart X and \bar{X} Control Charts”, lies at the basis of this chapter. The second paper

is a follow-up of this paper, and is titled accordingly as “On Guaranteed In-Control Performance for the Shewhart X and \bar{X} Control Charts”. Both papers are combined work with dr. M. Schoonhoven and prof. dr. R.J.M.M. Does in which I took the lead.

Chapter 4 provides adjusted control charts for monitoring process dispersion, following the exceedance probability criterion. The results are made quite general, so that they can be applied to many different estimators of dispersion and their corresponding control charts.

This chapter is based on the paper “Shewhart Control Charts for Dispersion Adjusted for Parameter Estimation”, published in *IISE Transactions*. This paper, Goedhart et al. (2017a), is a combined work with ms. M.M. da Silva, dr. M. Schoonhoven, dr. E.K. Epprecht, prof. dr. S. Chakraborti, prof. dr. R.J.M.M. Does and dr. A. Veiga. Originally, it started as a combined work with M. Schoonhoven and R.J.M.M. Does in which I took the lead, after which it turned out that another group of researchers was working on a similar paper. We then decided to merge our results into a single paper, which became a combined effort consisting primarily of matching notation and terminology. The most important part of my personal contribution is the general derivation of the control limit coefficients, which makes the proposed limits applicable for many different estimators.

In Chapter 5 the use of nonparametric control limits is described and evaluated. The performance of these limits is compared with other general methods that aim to relax the model assumptions in Phase I, or limit the impact of model errors, such as a general bootstrap procedure and transformations to normality. Moreover, the tradeoffs between more general models and the data demands accompanying them are discussed in this chapter.

This chapter is based on the paper “Tolerance Intervals in Statistical Process Monitoring”, which is submitted for publication. This paper, Goedhart et al. (2018b), is combined work with dr. M. Schoonhoven and prof. dr. R.J.M.M. Does, in which I took the lead.

Finally, in Chapter 6, a summary of this dissertation is provided along with some concluding remarks.

2. Bias Criterion for Control Charts for Location

In this chapter we derive correction factors for Shewhart control charts that guarantee a certain unconditional in-control ARL. In practice, the distribution parameters of the process characteristic of interest are unknown and have to be estimated, which may impact the control chart's performance (see Section 1.3). A well-known performance measure within SPM is the expectation of the ARL. A practitioner may want to design a control chart such that, in the in-control situation, it has a certain expected ARL. We use approximations to derive the required factors and show their accuracy and performance in out-of-control situations. We also evaluate the variation between the ARLs of the individually estimated control charts. This chapter is based on Goedhart et al. (2016).

2.1 Introduction

Shewhart X and \bar{X} control charts are common tools to monitor process means. For the X control chart, individual observations are used as plotting statistic in Phase II, while the \bar{X} control chart uses subgroup averages instead. In practice, the true process parameters are often unknown, and need to be estimated using a Phase I reference sample. However, since different practitioners have different Phase I data, the estimates of the process parameters will vary across practitioners. Because of this the estimated control limits, and consequently the control chart performance, are actually random variables. This variation in control chart performance has been addressed by several researchers, and is of great interest in current SPC literature. An overview of this literature is given by Jensen et al. (2006) and Psarakis et al. (2014).

The performance of a control chart is generally measured in terms of the FAR or the ARL. However, as shown by for example Quesenberry (1993) and Chen (1997), the amount of Phase I data required for the control chart to perform properly is substantially larger than initially thought. Although increasing the amount of Phase I data generally improves the accuracy of the parameter estimates, the required amounts are not always feasible in practice. Therefore, another solution is to apply corrections to the control chart limits to take parameter estimation into account. For example, Hillier (1969), Does and Schriever (1992) and Schoonhoven et al. (2009) provide corrections leading to a desired value of the FAR in expectation. Next to that, more recent works, for example Albers and Kallenberg (2004b, 2005), Chakraborti (2006), Faraz et al. (2015) and Saleh et al. (2015b), also consider other performance measures such as the ARL.

Performance of control charts based on estimated parameters can be evaluated for a given set of parameter values (known as conditional performance), or in expectation (known as unconditional performance). In this chapter we derive new corrections for the Shewhart X and \bar{X} control charts to achieve a desired unconditional ARL. Although Albers and Kallenberg (2005) derived correction factors for the same objective function, they have only considered the individuals chart. Therefore, we extend their approach for the Shewhart \bar{X} control chart. In addition, we derive new correction factors and compare their performance. It will be shown that our newly proposed correction factors outperform those of Albers and Kallenberg (2005).

This chapter is organized as follows. The next section discusses the considered model and notation. In Section 2.3, the performance of the correction factors of Albers and Kallenberg (2005) is assessed. Next, in Section 2.4 we describe the derivation of the newly proposed correction factors, and compare their performance. In Section 2.5 we discuss the implications of using the proposed corrections in in-control and out-of-control situations. In Section 2.6 the implementation of the control chart is illustrated with a real life example. Finally, in Section 2.7 we provide concluding remarks.

2.2 Model and Approach

The Shewhart control chart for monitoring the mean with known parameters has control limits

$$\begin{aligned} UCL &= \mu_0 + K \frac{\sigma_0}{\sqrt{n}} \\ LCL &= \mu_0 - K \frac{\sigma_0}{\sqrt{n}} \end{aligned} \quad (2.1)$$

where μ_0 and σ_0 are the in-control mean and standard deviation of the process characteristic of interest, respectively, and where K is the constant used to achieve the desired probability of a false signal (α_0). Let Y_{ij} denote the j -th observation in Phase II subgroup i ($i = 1, 2, \dots$ and $j = 1, 2, \dots, n$) and let Y_{ij} be i.i.d. $N(\mu_0 + \delta\sigma_0/\sqrt{n}, \sigma_0)$ random variables, where $\delta = 0$ corresponds to the in-control situation, and where the process is considered out-of-control for $\delta \neq 0$. Then, $K = \Phi^{-1}(1 - \alpha_0/2)$ with Φ^{-1} the inverse of the standard normal cumulative distribution function Φ . The run length (RL), i.e. the number of Phase II subgroups before the control chart gives a signal, has a geometric distribution with parameter α_0 . As a consequence, the in-control ARL is given by $1/\alpha_0$.

In practice, μ_0 and σ_0 are not known and therefore have to be estimated. To this end, m subgroups of n measurements on the process characteristic are collected when the process is considered to be in-control. Let X_{ij} (the j -th observation in the i -th subgroup in Phase I) be i.i.d. $N(\mu_0, \sigma_0)$ random variables ($i = 1, 2, \dots, m$ and $j = 1, 2, \dots, n$). Note that the X_{ij} 's correspond to Phase I, and the Y_{ij} 's correspond to Phase II.

For the subgroup control chart, μ_0 is usually estimated by the grand sample mean

$$\bar{\bar{X}} = \frac{1}{m} \sum_{i=1}^m \left(\frac{1}{n} \sum_{j=1}^n X_{ij} \right), \quad (2.2)$$

and σ_0 is usually estimated by the pooled standard deviation

$$S_{pooled} = \left(\frac{1}{m} \sum_{i=1}^m S_i^2 \right)^{1/2}, \quad (2.3)$$

where S_i is the i -th subgroup standard deviation defined by

$$S_i = \left(\frac{1}{n-1} \sum_{j=1}^n (X_{ij} - \bar{X}_i)^2 \right)^{1/2}.$$

An unbiased estimator of σ_0 is $S_{pooled}/c_4(m(n-1)+1)$, see for example Schoonhoven and Does (2012) or Saleh et al. (2015b), where $c_4(k)$ is defined by

$$c_4(k) = \left(\frac{2}{k-1} \right)^{1/2} \frac{\Gamma(k/2)}{\Gamma((k-1)/2)}.$$

For the individuals chart ($n = 1$), μ_0 is usually estimated by the sample average

$$\bar{X} = \frac{1}{m} \sum_{i=1}^m X_i, \tag{2.4}$$

where X_i is the i -th observation ($i = 1, 2, \dots, m$). When $n = 1$, σ_0 is usually estimated by the average moving range

$$\overline{MR} = \frac{1}{m-1} \sum_{i=1}^{m-1} |X_{i+1} - X_i|. \tag{2.5}$$

An unbiased estimator of σ_0 is $\overline{MR}/d_2(2)$ where $d_2(2) = \frac{2}{\sqrt{\pi}}$. Hence we use the most common estimators of σ_0 in both situations. However, it is possible to use any other desired estimator in our approach.

The performance of a control chart design based on estimated parameters can be evaluated for a given set of parameter values, known as conditional performance, or in expectation, known as unconditional performance. The conditional ARL (henceforth denoted as CARL) of a control chart with estimated parameters can be determined as follows. We define E_i ($i = 1, 2, \dots$) as the event that \bar{Y}_i falls outside the control limits. Then

$$P(E_i | \hat{\mu}_0, \hat{\sigma}_0) = P(\bar{Y}_i < \widehat{LCL}) + P(\bar{Y}_i > \widehat{UCL}), \tag{2.6}$$

where \widehat{LCL} and \widehat{UCL} are the control limits as defined in (2.1) but with μ_0 and σ_0 replaced by their unbiased estimates $\hat{\mu}_0$ and $\hat{\sigma}_0$, respectively. Given a pair of control limits, the events E_i are independent so that the conditional RL is geometrically distributed with parameter $P(E_i | \hat{\mu}_0, \hat{\sigma}_0)$. The CARL is then

given by $1/P(E_i|\hat{\mu}_0, \hat{\sigma}_0)$ and the unconditional ARL (henceforth denoted as EARL) by $E(1/P(E_i|\hat{\mu}_0, \hat{\sigma}_0))$. The unconditional probability of a signal can be determined by $E(P(E_i|\hat{\mu}_0, \hat{\sigma}_0))$.

For the Shewhart X and \bar{X} control charts with estimated parameters, we derive control limits of the form

$$\begin{aligned}\widehat{UCL} &= \hat{\mu}_0 + \tilde{K} \frac{\hat{\sigma}_0}{\sqrt{n}} \\ \widehat{LCL} &= \hat{\mu}_0 - \tilde{K} \frac{\hat{\sigma}_0}{\sqrt{n}}\end{aligned}\tag{2.7}$$

with \widehat{UCL} and \widehat{LCL} the respective upper and lower control chart limits based on the unbiased estimators $\hat{\mu}_0$ and $\hat{\sigma}_0$ of μ_0 and σ_0 , respectively, and $\tilde{K} = K + c$ the factor used to achieve a desired in-control EARL, which is denoted by $EARL_0$.

2.3 Albers and Kallenberg's Correction Factors

In this section, we assess the performance of the correction factors derived by Albers and Kallenberg (2005). First, we extend their work to subgroup control charts and then, based on simulations, we assess the accuracy of the proposed correction factors. The final subsection gives an overview of the accuracy of the correction factors.

2.3.1 Derivation of the Correction Factors

Albers and Kallenberg (2005) derived correction factors for the individuals X chart in a multiplicative form of $K(1 + \tilde{c})$ rather than the additive form $K + c$ that we use. However, their corrections can easily be transformed into a comparable form since $K(1 + \tilde{c}) = K + K\tilde{c} = K + c_{AK}$. Hence, we can simply multiply their proposed corrections c by K to make a comparison. For the individuals X chart with $n = 1$, their proposed correction c_{AK} is equal to

$$c_{AK} = -\frac{K + K^3\tau^2}{2m},\tag{2.8}$$

where $\tau^2 = \lim_{m \rightarrow \infty} \left[m \text{var} \left(\frac{\hat{\sigma}_0}{E\hat{\sigma}_0} \right) \right]$. Although Albers and Kallenberg (2005) considered estimators of σ_0 based on grouped observations (such as the pooled

standard deviation), they did not derive corrections for the Shewhart \bar{X} chart when $n > 1$. However, following their derivations (see Appendix 2.8) we arrive at the following corrections for the \bar{X} control chart

$$c_{AK} = -\frac{nK + K^3\tau^2}{2mn} \tag{2.9}$$

where $\tau^2 = \lim_{mn \rightarrow \infty} \left[mn \text{var} \left(\frac{\hat{\sigma}_0}{E\hat{\sigma}_0} \right) \right]$. Note that this more general correction is equal to their proposed correction for $n = 1$. The resulting correction factors are given in Table 2.1.

		$EARL_0 = 1000$	$EARL_0 = 370$	$EARL_0 = 200$	$EARL_0 = 100$
		$\alpha_0 = 0.001$	$\alpha_0 = 0.0027$	$\alpha_0 = 0.005$	$\alpha_0 = 0.01$
n	m	($K = 3.29$)	($K = 3.00$)	($K = 2.81$)	($K = 2.58$)
1	20	-0.8180	-0.6325	-0.5269	-0.4173
	30	-0.5453	-0.4217	-0.3513	-0.2782
	40	-0.4090	-0.3163	-0.2635	-0.2087
	50	-0.3272	-0.2530	-0.2108	-0.1669
	75	-0.2181	-0.1687	-0.1405	-0.1113
	100	-0.1636	-0.1265	-0.1054	-0.0835
3	20	-0.3049	-0.2437	-0.2084	-0.1712
	30	-0.2033	-0.1625	-0.1389	-0.1141
	40	-0.1525	-0.1219	-0.1042	-0.0856
	50	-0.1220	-0.0975	-0.0834	-0.0685
	75	-0.0813	-0.0650	-0.0556	-0.0457
	100	-0.0610	-0.0487	-0.0417	-0.0342
5	20	-0.1936	-0.1594	-0.1393	-0.1178
	30	-0.1291	-0.1062	-0.0929	-0.0785
	40	-0.0968	-0.0797	-0.0696	-0.0589
	50	-0.0774	-0.0637	-0.0557	-0.0471
	75	-0.0516	-0.0425	-0.0371	-0.0314
	100	-0.0387	-0.0319	-0.0279	-0.0236
7	20	-0.1565	-0.1312	-0.1163	-0.1000
	30	-0.1043	-0.0875	-0.0775	-0.0667
	40	-0.0782	-0.0656	-0.0581	-0.0500
	50	-0.0626	-0.0525	-0.0465	-0.0400
	75	-0.0417	-0.0350	-0.0310	-0.0267
	100	-0.0313	-0.0262	-0.0233	-0.0200

Table 2.1: Correction factors c_{AK} for various values of α_0 , m and n

2.3.2 Simulation Procedure

The accuracy of the correction factor c_{AK} is assessed for several combinations of parameter values, namely $n \in \{1, 3, 5, 7\}$, $m \in \{20, 30, 40, 50, 75, 100\}$ and $\alpha_0 \in \{0.001, 0.0027, 0.005, 0.01\}$.

For each combination of parameter values, we calculate the EARL as follows. Without loss of generality, we generate numerous Phase I datasets (at least 1,000,000 datasets consisting of m subgroups of size n) from a $N(0, 1)$ distribution in order to obtain the CARL values. For each Phase I dataset, we determine the control limits \widehat{UCL} and \widehat{LCL} according to (2.7). Recall that $P(E_i|\hat{\mu}_0, \hat{\sigma}_0)$ is defined as the probability that subgroup i generates a signal conditional on $\hat{\mu}_0$ and $\hat{\sigma}_0$, i.e.

$$P(E_i|\hat{\mu}_0, \hat{\sigma}_0) = P(\bar{Y}_i < \widehat{LCL}) + P(\bar{Y}_i > \widehat{UCL}).$$

Then, conditional on $\hat{\mu}_0$ and $\hat{\sigma}_0$, the distribution of the run length is geometric with parameter $P(E_i|\hat{\mu}_0, \hat{\sigma}_0)$. Hence, the CARL is given by

$$CARL = E(RL|\hat{\mu}_0, \hat{\sigma}_0) = \frac{1}{P(E_i|\hat{\mu}_0, \hat{\sigma}_0)}. \quad (2.10)$$

When we take the expectation over all Phase I samples, we obtain the EARL

$$EARL = E\left(\frac{1}{P(E_i|\hat{\mu}_0, \hat{\sigma}_0)}\right).$$

This expectation is obtained by simulation: for each Phase I dataset, $E(RL|\hat{\mu}_0, \hat{\sigma}_0)$ is computed. The number of Phase I datasets was at least 1,000,000 for each parameter combination. For a few combinations, additional datasets are generated so that the maximum relative standard error of the EARL is below 5% for $n = 1$, and below 1% for $n > 1$ for all cases.

BIAS CRITERION FOR CONTROL CHARTS FOR LOCATION

		$EARL_0 = 1000$	$EARL_0 = 370$	$EARL_0 = 200$	$EARL_0 = 100$
		$\alpha_0 = 0.001$	$\alpha_0 = 0.0027$	$\alpha_0 = 0.005$	$\alpha_0 = 0.01$
n	m	($K = 3.29$)	($K = 3.00$)	($K = 2.81$)	($K = 2.58$)
1	20	554 (1151800)	269 (191160)	160 (7858)	85 (968)
	30	807 (45575)	321 (3632)	178 (1079)	91 (326)
	40	878 (11698)	337 (1828)	185 (589)	94 (210)
	50	914 (5458)	346 (1084)	188 (442)	95 (173)
	75	948 (2532)	355 (699)	192 (320)	97 (139)
	100	963 (1945)	359 (581)	194 (280)	98 (128)
3	20	616 (2378)	251 (648)	142 (301)	75 (134)
	30	734 (1693)	287 (520)	160 (259)	83 (117)
	40	796 (1444)	306 (472)	169 (238)	87 (112)
	50	835 (1335)	318 (447)	175 (231)	89 (110)
	75	887 (1201)	335 (415)	183 (219)	93 (106)
	100	913 (1148)	343 (405)	187 (213)	94 (105)
5	20	624 (1328)	249 (433)	141 (224)	74 (106)
	30	727 (1184)	283 (406)	158 (214)	81 (103)
	40	786 (1127)	302 (391)	167 (209)	86 (101)
	50	822 (1095)	314 (389)	173 (204)	88 (101)
	75	876 (1053)	331 (384)	181 (202)	92 (100)
	100	898 (1036)	338 (377)	184 (201)	93 (100)
7	20	624 (1102)	248 (387)	140 (200)	74 (98)
	30	724 (1055)	282 (379)	157 (199)	81 (98)
	40	781 (1038)	300 (374)	166 (200)	85 (98)
	50	819 (1018)	312 (371)	172 (199)	88 (99)
	75	867 (1005)	328 (369)	180 (197)	91 (99)
	100	897 (1004)	338 (370)	184 (199)	93 (98)

Table 2.2: EARL of corrected (based on c_{AK}) and, between brackets, uncorrected control charts

2.3.3 Accuracy of Correction Factors

For a wide range of parameter values, the realized values of the EARL are given in Table 2.2. It can be observed that these values are still quite different from the desired values $EARL_0$, especially for small sample sizes. This difference appears to be larger for larger values of $EARL_0$. For larger sample sizes, the uncorrected control limits even perform better than the corrected control limits, as the EARL of the uncorrected chart are then closer to the desired value. It is clear that the proposed correction factors by Albers and Kallenberg (2005) are not sufficient: the EARL still deviates heavily from its desired value. For this reason, we derive new correction factors in the next section.

2.4 New Correction Factors

In this section, we derive new correction factors, give a numerical example and present the accuracy of the proposed correction factors.

2.4.1 Derivation of New Correction Factor

The idea behind the derivation of the factor is as follows. We want to derive a correction factor c such that the EARL equals a desired value, namely $EARL_0 = 1/\alpha_0$. In order to do this, we solve c from the equation $E(1/P(E_i|\hat{\mu}_0, \hat{\sigma}_0)) = 1/\alpha_0$. Note that the conditional probability of signaling ($P(E_i|\hat{\mu}_0, \hat{\sigma}_0)$) is given by

$$\begin{aligned}
 P(E_i|\hat{\mu}_0, \hat{\sigma}_0) &= 1 - P(\widehat{LCL} < \bar{Y}_i < \widehat{UCL}) \\
 &= 1 - \left[P\left(\frac{\bar{Y}_i - \mu_0}{\sigma_0/\sqrt{n}} < \frac{\hat{\mu}_0 - \mu_0}{\sigma_0/\sqrt{n}} + \tilde{K} \frac{\hat{\sigma}_0}{\sigma_0}\right) - P\left(\frac{\bar{Y}_i - \mu_0}{\sigma_0/\sqrt{n}} < \frac{\hat{\mu}_0 - \mu_0}{\sigma_0/\sqrt{n}} - \tilde{K} \frac{\hat{\sigma}_0}{\sigma_0}\right) \right] \\
 &= 1 - \left[\Phi\left(\frac{\hat{\mu}_0 - \mu_0}{\sigma_0/\sqrt{n}} + \tilde{K} \frac{\hat{\sigma}_0}{\sigma_0}\right) - \Phi\left(\frac{\hat{\mu}_0 - \mu_0}{\sigma_0/\sqrt{n}} - \tilde{K} \frac{\hat{\sigma}_0}{\sigma_0}\right) \right] \\
 &= [1 - \Phi(\tilde{K} + \Delta_1(\tilde{K}))] + [1 - \Phi(\tilde{K} + \Delta_2(\tilde{K}))] \\
 &= \bar{\Phi}(\tilde{K} + \Delta_1(\tilde{K})) + \bar{\Phi}(\tilde{K} + \Delta_2(\tilde{K}))
 \end{aligned} \tag{2.11}$$

with $\tilde{K} = K + c$, $\bar{\Phi}(x) = 1 - \Phi(x)$ and

$$\Delta_1(\tilde{K}) = \frac{\hat{\mu}_0 - \mu_0}{\sigma_0/\sqrt{n}} + \tilde{K} \left(\frac{\hat{\sigma}_0}{\sigma_0} - 1 \right)$$

and

$$\Delta_2(\tilde{K}) = -\frac{\hat{\mu}_0 - \mu_0}{\sigma_0/\sqrt{n}} + \tilde{K} \left(\frac{\hat{\sigma}_0}{\sigma_0} - 1 \right).$$

This notation is similar to that of Albers and Kallenberg (2004b), who derived corrections for the one-sided individuals control chart. From (2.11), we can write for any function g of $P(E_i|\hat{\mu}_0, \hat{\sigma}_0)$

$$g(P(E_i|\hat{\mu}_0, \hat{\sigma}_0)) = h(\tilde{K} + \Delta_1(\tilde{K}), \tilde{K} + \Delta_2(\tilde{K})) = h(x, y).$$

where we use $h(x, y)$ as general notation, which is required for the next steps in the derivation. Thus, for a specific value of \tilde{K} we can write $h(\tilde{K} + \Delta_1(\tilde{K}), \tilde{K} + \Delta_2(\tilde{K})) = h(x_0 + \Delta x, y_0 + \Delta y)$ where $(x_0, y_0) = (\tilde{K}, \tilde{K})$ and $(\Delta x, \Delta y) = (\Delta_1(\tilde{K}), \Delta_2(\tilde{K}))$. Next, we approximate the probability $P(E_i|\hat{\mu}, \hat{\sigma})$ by using a two-step Taylor expansion of $h(x, y) = h(\tilde{K} + \Delta_1(\tilde{K}), \tilde{K} + \Delta_2(\tilde{K}))$ around $(x_0, y_0) = (\tilde{K}, \tilde{K})$.

$$\begin{aligned} & h(\tilde{K} + \Delta_1(\tilde{K}), \tilde{K} + \Delta_2(\tilde{K})) \\ & \approx h(\tilde{K}, \tilde{K}) + h_x(\tilde{K}, \tilde{K})\Delta_1(\tilde{K}) + h_y(\tilde{K}, \tilde{K})\Delta_2(\tilde{K}) \\ & + \frac{1}{2}[h_{xx}(\tilde{K}, \tilde{K})\Delta_1^2(\tilde{K}, \tilde{K}) + 2h_{xy}(\tilde{K}, \tilde{K})\Delta_1(\tilde{K})\Delta_2(\tilde{K}) + h_{yy}(\tilde{K}, \tilde{K})\Delta_2^2(\tilde{K})], \end{aligned} \tag{2.12}$$

where $h_x(x, y)$ denotes the first-order partial derivative with respect to x , $h_{xx}(x, y)$ denotes the second-order partial derivative with respect to x (and likewise for y) and $h_{xy}(x, y)$ denotes the cross-partial derivative with respect to x and y .

When parameters are estimated, $Eg(P(E_i|\hat{\mu}_0, \hat{\sigma}_0)) \neq g(\alpha_0)$. Note that in (2.12) we have $h_x(\tilde{K}, \tilde{K}) = h_y(\tilde{K}, \tilde{K})$ and $h_{xx}(\tilde{K}, \tilde{K}) = h_{yy}(\tilde{K}, \tilde{K})$. To ensure that $Eg(P(E_i|\hat{\mu}_0, \hat{\sigma}_0)) = g(\alpha_0)$, we want to find a correction factor c such that

Point	h	h_x	h_{xx}	h_{xy}
(x, y)	$\frac{1}{\bar{\Phi}(x) + \bar{\Phi}(y)}$	$\frac{\phi(x)}{(\bar{\Phi}(x) + \bar{\Phi}(y))^2}$	$\frac{2\phi^2(x)}{(\bar{\Phi}(x) + \bar{\Phi}(y))^3} - \frac{x\phi(x)}{(\bar{\Phi}(x) + \bar{\Phi}(y))^2}$	$\frac{2\phi(x)\phi(y)}{(\bar{\Phi}(x) + \bar{\Phi}(y))^3}$
(K, K)	$\frac{1}{2\bar{\Phi}(K)}$	$\frac{\phi(K)}{4\bar{\Phi}^2(K)}$	$\frac{\phi^2(K)}{4\bar{\Phi}^3(K)} - \frac{K\phi(K)}{4\bar{\Phi}^2(K)}$	$\frac{\phi^2(K)}{4\bar{\Phi}^3(K)}$

Table 2.3: Derivatives of the function h , general at point (x, y) or specifically at point (K, K) for $g(\alpha_0) = 1/\alpha_0$

for $\tilde{K} = K + c$ the following holds (cf. (2.12)):

$$\begin{aligned}
 Eg(P(E_i|\hat{\mu}_0, \hat{\sigma}_0)) &= Eh(\tilde{K} + \Delta_1(\tilde{K}), \tilde{K} + \Delta_2(\tilde{K})) \\
 &\approx h(\tilde{K}, \tilde{K}) + h_x(\tilde{K}, \tilde{K}) \left[E\Delta_1(\tilde{K}) + E\Delta_2(\tilde{K}) \right] \\
 &\quad + \frac{1}{2}h_{xx}(\tilde{K}, \tilde{K}) \left[E\Delta_1^2(\tilde{K}) + E\Delta_2^2(\tilde{K}) \right] + h_{xy}(\tilde{K}, \tilde{K})E \left[\Delta_1(\tilde{K})\Delta_2(\tilde{K}) \right] \\
 &= g(\alpha_0) = h(K, K).
 \end{aligned} \tag{2.13}$$

In Appendix 2.9 we show that, for unbiased estimators $\hat{\mu}_0$ and $\hat{\sigma}_0$, we can rewrite this formula as

$$c = -\frac{h_{xx}(K, K)E\Delta_1^2(K) + h_{xy}(K, K)E[\Delta_1(K)\Delta_2(K)]}{2h_x(K, K)}. \tag{2.14}$$

The functions $h(x, y)$, $h_x(x, y)$, $h_{xx}(x, y)$ and $h_{xy}(x, y)$ depend on the function $g(\alpha_0)$ that is being considered. For the ARL, $g(\alpha_0) = 1/\alpha_0$. The derivatives for this function at point (K, K) are given in Table 2.3. Note that the approach can easily be adapted for any other function $g(\alpha_0)$ by adapting the corresponding derivatives. Commonly considered functions, apart from the ARL, are the false alarm rate ($g(\alpha_0) = \alpha_0$) and the probability that the run length is at most a specified value j ($g(\alpha_0) = 1 - (1 - \alpha_0)^j$), see e.g. Tsai et al. (2005).

The derivations given in (2.11)-(2.14) are still in general form with regards to the used estimators. For $\hat{\mu}_0 = \bar{X}$ and $\hat{\sigma}_0 = S_{pooled}/c_4(m(n-1)+1)$ (cf. (2.2)

and (2.3)) we have (see Appendix 2.10)

$$E[\Delta_1(K)\Delta_2(K)] \approx \frac{K^2}{2(m(n-1)+1)} - \frac{1}{m} \quad (2.15)$$

and

$$E\Delta_1^2(K) = E\Delta_2^2(K) \approx \frac{K^2}{2(m(n-1)+1)} + \frac{1}{m}. \quad (2.16)$$

Besides the application to the Shewhart \bar{X} control chart it can also be applied to the individuals control chart. Since a different estimator for the standard deviation is used in this case, namely $\overline{MR}/d_2(2)$, the explicit calculation of $E[\Delta_1(K)\Delta_2(K)]$, $E\Delta_1^2(K)$ and $E\Delta_2^2(K)$ will be different here. In Appendix 2.11 we show that in this case

$$E[\Delta_1(K)\Delta_2(K)] \approx K^2 \left[\frac{0.8264m - 1.082}{(m-1)^2} \right] - \frac{1}{m} \quad (2.17)$$

and

$$E\Delta_1^2(K) = E\Delta_2^2(K) \approx K^2 \left[\frac{0.8264m - 1.082}{(m-1)^2} \right] + \frac{1}{m}. \quad (2.18)$$

2.4.2 Examples of Computation of the Correction Factors

In this section, we demonstrate how to calculate the correction factor for both a subgroup control chart and an individuals control chart. In addition, we give an overview of correction factors for different values of α_0 , n , and m .

First, we give an example of the calculation of a correction factor for a subgroup control chart with $\alpha_0 = 0.0027$, $n = 5$ and $m = 50$. In this case, $K = \Phi^{-1}(1 - \alpha_0/2) = 3$. We need to find the correction factor c in $\tilde{K} = K + c$. For the subgroup control chart, c is given by formula (2.14). To calculate c , we need to find $h_{xx}(K, K)$, $h_{xy}(K, K)$, $h_x(K, K)$, $E\Delta_1^2(K)$ and $E[\Delta_1(K)\Delta_2(K)]$. These are as follows.

$$h_x(K, K) = \frac{\phi(3)}{4\Phi^2(3)} = 608.03,$$

$$h_{xx}(K, K) = \frac{\phi^2(3)}{4\bar{\Phi}^3(3)} - \frac{3\phi(3)}{4\bar{\Phi}^2(3)} = 172.13,$$

$$h_{xy}(K, K) = \frac{\phi^2(3)}{4\bar{\Phi}^3(3)} = 1996.2,$$

$$E\Delta_1^2(K) = \frac{3^2}{2(50(5-1)+1)} + \frac{1}{50} = 0.042,$$

$$E[\Delta_1(K)\Delta_2(K)] = \frac{3^2}{2(50(5-1)+1)} - \frac{1}{50} = 0.0024.$$

From this, we can derive c as

$$c = -\frac{172.13 \cdot 0.042 + 1996.2 \cdot 0.0024}{2 \cdot 608.03} = -0.0099.$$

Next, we give an example of the calculation of a correction factor for an individuals control chart with $\alpha_0 = 0.005$, $n = 1$ and $m = 100$. Note that in this case σ_0 is estimated by $\overline{MR}/d_2(2)$. We now have $K = \Phi^{-1}(1 - \alpha_0/2) = 2.81$. Again, we need to find the correction factor c in $\tilde{K} = K + c$ using (2.14). To calculate c , we need to find $h_{xx}(K, K)$, $h_{xy}(K, K)$, $h_x(K, K)$, $E\Delta_1^2(K, K)$ and $E[\Delta_1(K)\Delta_2(K)]$. These are as follows.

$$h_x(K, K) = \frac{\phi(2.81)}{4\bar{\Phi}^2(2.81)} = 310.44,$$

$$h_{xx}(K, K) = \frac{\phi^2(2.81)}{4\bar{\Phi}^3(2.81)} - \frac{3\phi(2.81)}{4\bar{\Phi}^2(2.81)} = 92.30,$$

$$h_{xy}(K, K) = \frac{\phi^2(2.81)}{4\bar{\Phi}^3(2.81)} = 963.70,$$

$$E\Delta_1^2(K) = 2.81^2 \left[\frac{0.8264 \cdot 100 - 1.082}{(100-1)^2} \right] + \frac{1}{100} = 0.076,$$

$$E [\Delta_1(K)\Delta_2(K)] = 2.81^2 \left[\frac{0.8264 \cdot 100 - 1.082}{(100 - 1)^2} \right] - \frac{1}{100} = 0.056.$$

The desired value of c is now given by

$$c = -\frac{92.30 \cdot 0.076 + 963.70 \cdot 0.076}{2 \cdot 310.44} = -0.098.$$

In Table 2.4, we present the correction factors for $\alpha_0 = 0.001$, $\alpha = 0.0027$, $\alpha_0 = 0.005$ and $\alpha_0 = 0.01$ for various values of n and m . As can be seen from the table, a smaller n and m result in a larger correction. Similarly, the smaller the value of α_0 , the larger the correction needed. Finally, smaller values of α_0 and n result in a negative correction factor. The reason is that there is a larger probability of extreme limits resulting in high CARLs. These high CARLs have an impact on the EARL - calculated as the average of the CARLs - so that the correction factor has to be negative. This is not the case for higher values of n and m or higher values of α_0 .

2.4.3 Accuracy of Correction Factors

In this section, we assess the accuracy of the correction factors. In order to do so, we simulate the performance of control charts with these correction factors. The simulation procedure is described in Section 2.3.2.

The in-control EARL values are presented in Table 2.5. For comparison purposes, the in-control EARLs of the uncorrected charts are presented in brackets. The EARLs of the uncorrected charts deviate from the desired value $EARL_0$. The smaller the value of α_0 and the smaller the values of m and n , the larger this deviation is. As we can conclude from Table 2.5, the correction factor performs relatively well: for sample sizes greater than $n = 1$ and $m = 40$, the correction factor leads to an EARL close to the desired value. As we will see in the next section, this leads to better out-of-control performance as well: when the in-control ARL is exceptionally high, the control chart is not able to detect changes in the mean very quickly.

2.5 Implications of Proposed Corrected Charts

In this section we discuss the implications of using the proposed correct control charts. First, we discuss the out-of-control performance compared to uncorrected charts. Afterwards, we evaluate the variation between CARLs in both the in-control and out-of-control situation.

2.5.1 Out-of-Control Performance of Corrected Charts

To evaluate the out-of-control performance of the charts with the new correction factors, we determine the EARLs for two out-of-control situations. We use the same simulation procedure as described earlier but with the adjustment that the process mean has changed, i.e. $Y_{ij} \sim N(\mu_0 + \delta\sigma_0/\sqrt{n}, \sigma_0)$ with $\delta \neq 0$. Again, without loss of generality we set $\mu_0 = 0$ and $\sigma_0 = 1$ in the simulations.

Table 2.6 shows the results for $\delta = 0.5$ and Table 2.7 for $\delta = 1$. We can conclude from the tables that the corrected control charts are more powerful in detecting shifts in the process mean than the uncorrected charts, the reason being the high in-control CARLs of some uncorrected charts. The positive effect of the correction factor shows up mainly for small sample sizes: for the individuals control chart and $n = 3$ for the subgroup control chart. In conclusion, for small sample sizes the corrected control charts lead to more stable in-control as well as out-of-control performance compared to the uncorrected control charts.

2.5.2 Variation in CARLs

We have derived correction factors that lead to the desired value of the in-control EARL. Note that a given correction factor leads only to a desired average value: the CARL values will vary across the conditional control charts used by each individual practitioner. To illustrate this, we have drawn boxplots for the CARL values obtained in the simulations. For each simulation run, we can calculate $E(RL|\hat{\mu}_0, \hat{\sigma}_0)$ from (2.10) and we can calculate the percentiles of the CARLs obtained by the simulations runs.

Figures 2.1a and 2.1b show the 10th, 25th, 50th, 75th and 90th percentiles of the in-control CARLs when $\alpha_0 = 0.0027$, for $n = 3, m = 50$ and $n = 1, m = 100$ respectively. As the figures show, both the corrected and the uncorrected control charts display quite some variation within the CARL values. However, this variation is much less for the corrected charts than for the uncorrected charts. The same results are found in the out-of-control situation, as can be seen in Figures 2.2a and 2.2b. These figures show the 10th, 25th, 50th, 75th and 90th percentiles of the out-of-control ($\delta = 0.5$) CARLs when $\alpha_0 = 0.0027$, for $n = 3, m = 50$ and $n = 1, m = 100$ respectively.

2.5 IMPLICATIONS OF PROPOSED CORRECTED CHARTS

		$EARL_0 = 1000$	$EARL_0 = 370$	$EARL_0 = 200$	$EARL_0 = 100$
		$\alpha_0 = 0.001$	$\alpha_0 = 0.0027$	$\alpha_0 = 0.005$	$\alpha_0 = 0.01$
n	m	($K = 3.29$)	($K = 3.00$)	($K = 2.81$)	($K = 2.58$)
<hr/>					
1	20	-0.8022	-0.6116	-0.5032	-0.3910
	30	-0.5279	-0.4024	-0.3310	-0.2571
	40	-0.3934	-0.2998	-0.2466	-0.1915
	50	-0.3135	-0.2389	-0.1965	-0.1526
	75	-0.2079	-0.1585	-0.1303	-0.1012
	100	-0.1556	-0.1185	-0.0975	-0.0757
<hr/>					
3	20	-0.1698	-0.1207	-0.0933	-0.0654
	30	-0.1146	-0.0815	-0.0631	-0.0443
	40	-0.0865	-0.0616	-0.0476	-0.0335
	50	-0.0694	-0.0494	-0.0383	-0.0269
	75	-0.0465	-0.0331	-0.0257	-0.0181
	100	-0.0350	-0.0249	-0.0193	-0.0136
<hr/>					
5	20	-0.0453	-0.0241	-0.0126	-0.0013
	30	-0.0306	-0.0163	-0.0086	-0.0010
	40	-0.0231	-0.0123	-0.0065	-0.00084
	50	-0.0185	-0.0099	-0.0053	-0.00072
	75	-0.0124	-0.0067	-0.0035	-0.00051
	100	-0.0093	-0.0050	-0.0027	-0.00039
<hr/>					
7	20	-0.0032	0.0087	0.0148	0.0204
	30	-0.0023	0.0057	0.0098	0.0135
	40	-0.0018	0.0042	0.0073	0.0101
	50	-0.0014	0.0033	0.0058	0.0081
	75	-0.00098	0.0022	0.0039	0.0054
	100	-0.00074	0.0016	0.0029	0.0040

Table 2.4: Correction factors c for various values of α_0 , n and m

BIAS CRITERION FOR CONTROL CHARTS FOR LOCATION

		$EARL_0 = 1000$	$EARL_0 = 370$	$EARL_0 = 200$	$EARL_0 = 100$
		$\alpha_0 = 0.001$	$\alpha_0 = 0.0027$	$\alpha_0 = 0.005$	$\alpha_0 = 0.01$
n	m	($K = 3.29$)	($K = 3.00$)	($K = 2.81$)	($K = 2.58$)
1	20	534 (1151800)	307 (191160)	179 (7858)	96 (968)
	30	864 (45575)	337 (3632)	192 (1079)	101 (326)
	40	962 (11698)	375 (1828)	197 (589)	100 (210)
	50	993 (5458)	367 (1084)	198 (442)	99 (173)
	75	986 (2532)	369 (699)	198 (320)	99 (139)
	100	1002 (1945)	370 (581)	199 (280)	100 (128)
3	20	1090 (2378)	398 (648)	212 (301)	107 (134)
	30	1046 (1693)	383 (520)	208 (259)	102 (117)
	40	1022 (1444)	378 (472)	203 (238)	101 (112)
	50	1017 (1335)	375 (447)	204 (231)	101 (110)
	75	1008 (1201)	371 (415)	202 (219)	101 (106)
	100	1008 (1148)	372 (405)	200 (213)	101 (105)
5	20	1111 (1328)	397 (433)	214 (224)	106 (106)
	30	1053 (1184)	383 (406)	208 (214)	103 (103)
	40	1034 (1127)	374 (391)	204 (209)	101 (101)
	50	1022 (1095)	376 (389)	201 (204)	101 (101)
	75	1007 (1053)	371 (384)	200 (202)	100 (100)
	100	1002 (1036)	370 (377)	199 (201)	100 (100)
7	20	1089 (1102)	399 (387)	210 (200)	105 (98)
	30	1047 (1055)	387 (379)	205 (199)	102 (98)
	40	1031 (1038)	379 (374)	205 (200)	101 (98)
	50	1013 (1018)	375 (371)	203 (199)	101 (99)
	75	1002 (1005)	370 (369)	200 (197)	100 (99)
	100	1001 (1004)	372 (370)	201 (199)	100 (98)

Table 2.5: EARL of corrected and, between brackets, uncorrected control charts

2.5 IMPLICATIONS OF PROPOSED CORRECTED CHARTS

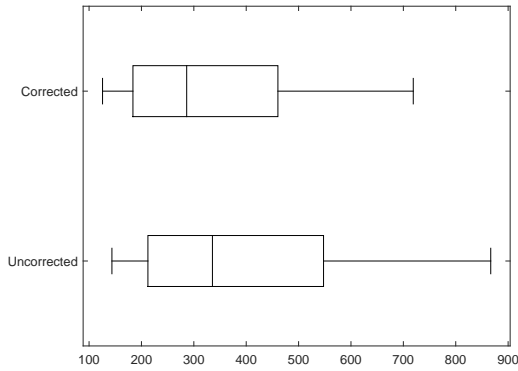
n	m	$\alpha_0 = 0.001$	$\alpha_0 = 0.0027$	$\alpha_0 = 0.005$	$\alpha_0 = 0.01$
		($K = 3.29$)	($K = 3.00$)	($K = 2.81$)	($K = 2.58$)
1	20	276 (1759500)	138 (16449)	97 (7172)	52 (441)
	30	339 (11595)	157 (1301)	97 (474)	53 (145)
	40	373 (3091)	160 (589)	95 (250)	52 (100)
	50	379 (1636)	162 (460)	95 (193)	52 (85)
	75	379 (871)	159 (276)	93 (142)	52 (70)
	100	377 (672)	160 (237)	93 (125)	51 (63)
3	20	511 (1112)	208 (326)	117 (161)	61 (75)
	30	473 (734)	191 (252)	107 (131)	57 (65)
	40	444 (608)	178 (217)	104 (120)	55 (60)
	50	428 (546)	175 (204)	101 (113)	55 (59)
	75	405 (474)	168 (186)	99 (106)	53 (55)
	100	399 (448)	165 (178)	96 (101)	52 (54)
5	20	546 (646)	210 (227)	119 (124)	62 (62)
	30	491 (547)	190 (201)	110 (113)	59 (59)
	40	452 (489)	184 (191)	104 (106)	56 (56)
	50	427 (454)	176 (182)	101 (103)	55 (55)
	75	412 (429)	169 (173)	98 (99)	53 (53)
	100	401 (413)	165 (167)	95 (96)	52 (52)
7	20	544 (550)	215 (209)	119 (114)	62 (59)
	30	486 (490)	191 (188)	111 (108)	58 (56)
	40	453 (455)	184 (181)	105 (103)	56 (55)
	50	431 (433)	178 (176)	103 (101)	55 (54)
	75	410 (412)	169 (168)	98 (97)	53 (52)
	100	410 (402)	165 (164)	96 (96)	52 (52)

Table 2.6: Corrected and, between brackets, uncorrected out-of-control EARL for $\delta = 0.5$

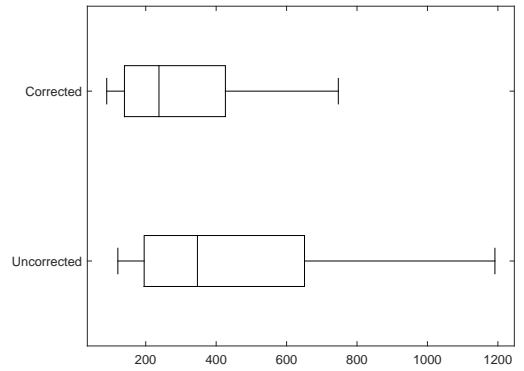
n	m	$\alpha_0 = 0.001$	$\alpha_0 = 0.0027$	$\alpha_0 = 0.005$	$\alpha_0 = 0.01$
		($K = 3.29$)	($K = 3.00$)	($K = 2.81$)	($K = 2.58$)
1	20	60 (122320)	35 (2451)	25 (307)	17 (94)
	30	74 (12232)	40 (204)	27 (97)	17 (38)
	40	80 (410)	42 (121)	27 (58)	17 (29)
	50	84 (281)	42 (90)	28 (49)	17 (25)
	75	87 (171)	43 (68)	28 (39)	17 (22)
	100	89 (145)	43 (60)	28 (36)	17 (21)
3	20	122 (224)	57 (84)	35 (46)	21 (25)
	30	110 (158)	52 (66)	33 (39)	19 (21)
	40	105 (136)	50 (59)	31 (35)	19 (20)
	50	102 (125)	48 (54)	31 (34)	19 (20)
	75	97 (110)	47 (51)	30 (32)	18 (19)
	100	96 (106)	47 (50)	30 (31)	18 (19)
5	20	131 (151)	59 (63)	37 (38)	22 (22)
	30	115 (126)	53 (56)	34 (34)	20 (20)
	40	110 (117)	51 (53)	32 (33)	19 (19)
	50	105 (110)	49 (51)	31 (32)	19 (19)
	75	99 (103)	48 (48)	30 (30)	18 (18)
	100	96 (99)	46 (47)	29 (30)	18 (18)
7	20	135 (137)	62 (60)	37 (36)	22 (21)
	30	118 (119)	54 (53)	34 (33)	20 (20)
	40	110 (110)	51 (51)	33 (32)	19 (19)
	50	105 (106)	50 (49)	32 (31)	19 (19)
	75	100 (100)	48 (48)	30 (30)	18 (18)
	100	98 (98)	46 (46)	30 (30)	18 (18)

Table 2.7: Corrected and, between brackets, uncorrected out-of-control EARL for $\delta = 1$

2.5 IMPLICATIONS OF PROPOSED CORRECTED CHARTS

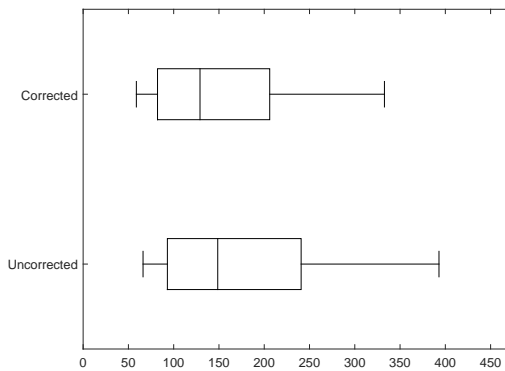


(a) $n = 3, m = 50$ and $\alpha_0 = 0.0027$

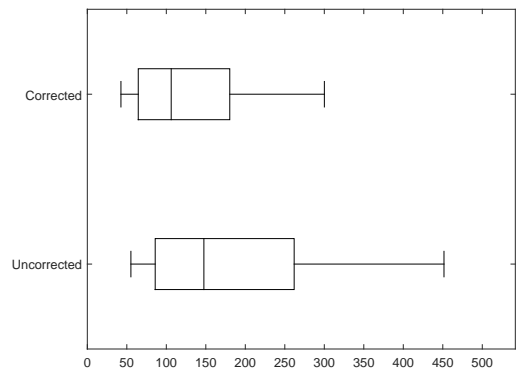


(b) $n = 1, m = 100$ and $\alpha_0 = 0.0027$

Figure 2.1: 10th, 25th, 50th, 75th and 90th percentiles of the in-control CARLs for different parameter combinations.



(a) $n = 3, m = 50$ and $\alpha_0 = 0.0027$



(b) $n = 1, m = 100$ and $\alpha_0 = 0.0027$

Figure 2.2: 10th, 25th, 50th, 75th and 90th percentiles of the out-of-control ($\delta = 0.5$) CARLs for different parameter combinations.

2.6 Application of the Proposed Control Chart

In this section we demonstrate the application of the proposed control chart through a real life example. Data is collected of the torque of Torque-to-Yield bolts that are used as fasteners in engines at a subsidiary of Paccar. The bolts are tightened at several different positions of the engines using a specific fastening procedure. At the end of this procedure, the torque is measured (in Newton-meter) for bolts at two specific positions. The measurements are performed by a process engineer for the purpose of process monitoring. Clearly, detecting trends or anomalies in the applied torque is very important, as this indicates problems with either the bolt(s) or the fastening procedure. For example, the performance of the used wrenches can deteriorate over time, which can result in fasteners being too tight or too loose. It is thus of major importance to detect such out-of-control situations.

An initial Phase I dataset of 40 observations, consisting of $m = 20$ subgroups of size $n = 2$, is collected by the process engineer to construct the control limits (see Table 2.8). First, we have checked our data for normality. We found no reason to reject the normality assumption from the Shapiro-Wilk test for normality. Afterwards, the control limits are constructed and used to monitor the process mean. The Phase II dataset consists of 31 subgroups of size 2 (see also Table 2.8). The control limits are obtained using the following steps:

1. First, we determine our parameters. We have $m = 20$, $n = 2$ and choose $\alpha_0 = 0.0027$. The latter is conventional for Shewhart control charts, and means that we have $K = 3$. As estimators of μ_0 and σ_0 we use $\hat{\mu}_0 = \bar{\bar{X}}$ and $\hat{\sigma}_0 = S_{pooled}/c_4(m(n-1)+1)$ respectively. As parameter estimates we find $\hat{\mu}_0 = 164.08$ and $\hat{\sigma}_0 = 0.0508$ from our Phase I dataset.
2. Next, we calculate the required correction c . This requires the calculations of $h_{xx}(K, K)$, $h_{xy}(K, K)$, $h_x(K, K)$, $E\Delta_1^2(K)$ and $E[\Delta_1(K)\Delta_2(K)]$. These are the same as calculated in Section 2.4.2, except for

$$E\Delta_1^2(K) = \frac{3^2}{2(20(2-1)+1)} + \frac{1}{20} = 0.2643,$$

$$E[\Delta_1(K)\Delta_2(K)] = \frac{3^2}{2(20(2-1)+1)} - \frac{1}{20} = 0.1643.$$

From this we can calculate c as

$$c = -\frac{172.13 \cdot 0.2643 + 1996.21 \cdot 0.1643}{2 \cdot 608.03} = -0.3071.$$

3. Finally, we calculate the control limits using $\hat{\mu}_0$, $\hat{\sigma}_0$ and $\tilde{K} = 3 - 0.3071 = 2.6929$, such that:

$$\widehat{UCL} = 164.08 + 2.6929 \cdot 0.0508/\sqrt{2} = 164.1767,$$

$$\widehat{LCL} = 164.08 - 2.6929 \cdot 0.0508/\sqrt{2} = 163.9833.$$

The Phase II data is now monitored using these control limits, as is illustrated in Figure 2.3. Note that the width of the control limits is quite small. As can be observed, an out-of-control signal is given by observation 7 in the dataset. Usually, when a signal is given, the wrenches should be recalibrated. However, after further investigation no action was required in this case, because the out-of-control signal turned out to be caused by a measurement error. The control chart is used in practice to detect substantial anomalies and clear trends.

Subgroup	X_{i1}	X_{i2}	Y_{i1}	Y_{i2}
1	164.06	163.98	164.13	164.19
2	164.11	164.05	164.18	164.02
3	164.03	164.09	164.17	164.02
4	164.10	164.13	164.10	164.07
5	164.04	164.15	163.95	164.04
6	164.06	164.22	164.15	164.03
7	163.98	164.11	163.92	164.02
8	164.06	164.09	164.08	164.15
9	164.10	164.08	164.06	163.96
10	164.03	164.03	163.97	164.05
11	164.12	164.09	164.11	164.15
12	164.13	164.04	164.10	164.15
13	164.03	164.10	163.98	164.02
14	164.17	164.05	164.08	164.08
15	164.00	164.06	164.02	164.16
16	164.15	163.98	164.02	164.18
17	163.96	164.02	164.11	164.03
18	164.02	164.08	164.03	164.05
19	164.17	164.23	163.98	164.00
20	164.05	164.07	164.09	163.99
21	-	-	164.14	164.04
22	-	-	163.94	164.03
23	-	-	164.12	164.02
24	-	-	164.03	164.12
25	-	-	164.15	164.18
26	-	-	164.13	164.11
27	-	-	164.00	164.05
28	-	-	164.10	164.15
29	-	-	164.15	164.16
30	-	-	164.33	164.02
31	-	-	164.07	164.28

Table 2.8: Phase I (X_{ij} 's) and Phase II (Y_{ij} 's) data of torque values

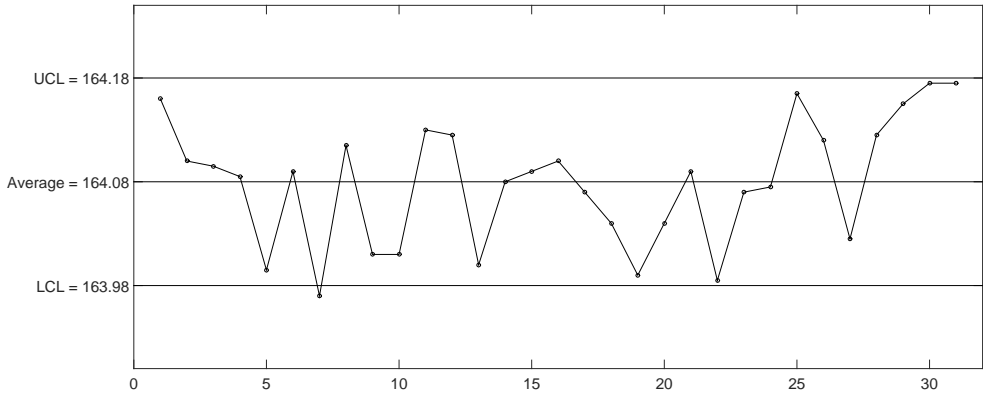


Figure 2.3: Control chart of the monitoring of torque values with the proposed control limits. The monitored Phase II sample consists of $m = 31$ subgroups of size $n = 2$ each.

2.7 Concluding Remarks

The ARL is a well-known metric to evaluate control chart performance. However, when the parameters of the process characteristic are estimated from a limited amount of data, the expected value of the in-control ARL differs substantially from the desired value.

In this chapter, we have derived correction factors for the Shewhart X and \bar{X} control charts that can be used to obtain the desired value of the in-control ARL in expectation. As is shown, the correction factor is accurate: it leads to the desired value of the in-control EARL for a broad range of parameter values (α_0 , n and m). The added value of the correction factor shows up predominantly when sample sizes are small, namely the individuals chart ($n = 1$) and the subgroup control chart when $n = 3$ or 5 and $m = 20$. Control limit adjustments for control charts for dispersion based on the bias criterion can be found in Diko et al. (2017).

2.8 Appendix: Albers and Kallenberg Correction

Although Albers and Kallenberg (2005) derived corrections for the individuals X chart only, it is possible to apply their approach to the \bar{X} chart as well. First note that they derive correction factors \tilde{c} in the form $K(1 + \tilde{c})$ rather than $K + c$ as we do. Their approach for the individuals chart states that in the one-sided case, the bias is best removed when

$$\tilde{c} = \frac{1}{2} \left(Var \left(\frac{\bar{X}}{\sigma_0} \right) + K^2 Var \left(\frac{\hat{\sigma}_0}{\sigma_0} \right) \right). \quad (2.19)$$

Next, note that \bar{X} can essentially be treated as a $N(\mu_0, \sigma_0/\sqrt{n})$ variable. Consequently, replacing σ_0 and $\hat{\sigma}_0$ in (2.19) by σ_0/\sqrt{n} and $\hat{\sigma}_0/\sqrt{n}$ respectively leads to the required corrections for the \bar{X} chart. This gives us

$$\tilde{c} = \frac{1}{2} \left(Var \left(\frac{\bar{X}}{\sigma_0/\sqrt{n}} \right) + K^2 Var \left(\frac{\hat{\sigma}_0}{\sigma_0} \right) \right) = \frac{1}{2} \left(\frac{1}{m} + \frac{K^2 \tau^2}{mn} \right) = \frac{n + K^2 \tau^2}{2mn}.$$

In order to obtain the required correction for the two-sided case, Albers and Kallenberg (2005) instruct to add a minus sign in front of the obtained correction, and to adjust the value of K from $\Phi^{-1}(1 - \alpha_0)$ to $\Phi^{-1}(1 - \alpha_0/2)$. Finally, since their correction is in the form of $K(1 + \tilde{c}) = K + K\tilde{c} = K + c_{AK}$, we then construct $c_{AK} = -\frac{nK + K^3 \tau^2}{2mn}$ to make our correction factors comparable. Note that for $n = 1$ this result corresponds with the original proposal of Albers and Kallenberg (2005).

2.9 Appendix: Correction Factor

Here, we show how to derive (2.14) from (2.13). As given in (2.13), we want to have

$$\begin{aligned}
 Eh(\tilde{K} + \Delta_1(\tilde{K}), \tilde{K} + \Delta_2(\tilde{K})) &\approx h(\tilde{K}, \tilde{K}) + h_x(\tilde{K}, \tilde{K}) [E\Delta_1(\tilde{K}) + E\Delta_2(\tilde{K})] \\
 &+ \frac{1}{2}h_{xx}(\tilde{K}, \tilde{K}) [E\Delta_1^2(\tilde{K}) + E\Delta_2^2(\tilde{K})] + h_{xy}(\tilde{K}, \tilde{K})E [\Delta_1(\tilde{K})\Delta_2(\tilde{K})] = h(K, K).
 \end{aligned}$$

Next, we apply a two-step Taylor approximation around K to the middle part and ignore relatively small factors such as c^2 , $cE\Delta_1(K)$ and $cE\Delta_2(K)$. This essentially boils down to approximating $h(\tilde{K}, \tilde{K})$ by $h(K, K) + c[h_x(K, K) + h_y(K, K)]$ and replacing \tilde{K} by K in the remaining parts. We then find that this equation is approximately equal to

$$\begin{aligned}
 Eh(\tilde{K} + \Delta_1(\tilde{K}), \tilde{K} + \Delta_2(\tilde{K})) &\approx \\
 h(K, K) + c[h_x(K, K) + h_y(K, K)] + h_x(K, K) [E\Delta_1(K) + E\Delta_2(K)] \\
 &+ \frac{1}{2}h_{xx}(K, K) [E\Delta_1^2(K) + E\Delta_2^2(K)] + h_{xy}(K, K)E [\Delta_1(K)\Delta_2(K)] = h(K, K).
 \end{aligned}$$

Note that $h_x(K, K) = h_y(K, K)$, and that for unbiased estimators $\hat{\mu}$ and $\hat{\sigma}$ there holds $E\Delta_1(K) = E\Delta_2(K) = 0$ and $E\Delta_1^2(K) = E\Delta_2^2(K)$. Then, solving this equation to c leads to

$$c = -\frac{h_{xx}(K, K)E\Delta_1^2(K) + h_{xy}(K, K)E [\Delta_1(K)\Delta_2(K)]}{2h_x(K, K)}.$$

2.10 Appendix: Subgroup Control Chart

For $E[\Delta_1(K)\Delta_2(K)]$, we have

$$E[\Delta_1(K)\Delta_2(K)] = -E\left(\frac{\hat{\mu}_0 - \mu_0}{\sigma_0/\sqrt{n}}\right)^2 + K^2 E\left(\frac{\hat{\sigma}_0}{\sigma_0} - 1\right)^2,$$

where for $\hat{\mu}_0 = \bar{X}$ and $\hat{\sigma}_0 = S_{pooled}/c_4(m(n-1)+1)$ (cf. (2.2) and (2.3)) we have

$$-E\left(\frac{\hat{\mu}_0 - \mu_0}{\sigma_0/\sqrt{n}}\right)^2 = -\frac{1}{m},$$

and

$$\begin{aligned} & E\left(\frac{\hat{\sigma}_0}{\sigma_0} - 1\right)^2 \\ &= \frac{1}{c_4^2(m(n-1)+1)} E\left(\frac{S_{pooled}^2}{\sigma_0^2}\right) - 1 \\ &\approx \frac{1}{2(m(n-1)+1)}, \end{aligned}$$

so that

$$E[\Delta_1(K)\Delta_2(K)] \approx \frac{K^2}{2(m(n-1)+1)} - \frac{1}{m}$$

and

$$E\Delta_1^2(K) = E\Delta_2^2(K) \approx \frac{K^2}{2(m(n-1)+1)} + \frac{1}{m}.$$

2.11 Appendix: Individuals Control Chart

For the individuals control chart, the calculation of $E[\Delta_1(K)\Delta_2(K)]$, $E\Delta_1^2(K)$ and $E\Delta_2^2(K)$ is different. For $E[\Delta_1(K)\Delta_2(K)]$, we have

$$\begin{aligned} E[\Delta_1(K)\Delta_2(K)] &= -E\left(\frac{\hat{\mu}_0 - \mu_0}{\sigma_0}\right)^2 + K^2 E\left(\frac{\hat{\sigma}_0}{\sigma_0} - 1\right)^2 \\ &\approx -\frac{1}{m} + K^2 \text{var}\left(\frac{\hat{\sigma}_0}{\sigma_0}\right). \end{aligned}$$

Cryer and Ryan (1990) showed that

$$\text{Var}(\overline{MR}/\sigma_0) \approx d_2^2(2) \left[\frac{0.8264m - 1.082}{(m-1)^2} \right]$$

and thus

$$\text{Var}\left(\frac{\hat{\sigma}_0}{\sigma_0}\right) \approx \left[\frac{0.8264m - 1.082}{(m-1)^2} \right].$$

Hence,

$$E[\Delta_1(K)\Delta_2(K)] = -\frac{1}{m} + K^2 \left[\frac{0.8264m - 1.082}{(m-1)^2} \right]$$

and

$$E\Delta_1^2(K) = E\Delta_2^2(K) = \frac{1}{m} + K^2 \left[\frac{0.8264m - 1.082}{(m-1)^2} \right].$$

3. Exceedance Probability Criterion for Control Charts for Location

To prevent low in-control ARLs, new corrections for Shewhart \bar{X} and \bar{X} control charts are given in this chapter that guarantee a minimum in-control performance with a specified probability. To balance the tradeoff between in-control and out-of-control performance, the minimum performance threshold and specified probability can be adjusted as desired. Furthermore, a comparison is made with tolerance intervals, self-starting control charts, and bootstrap. This chapter is based on Goedhart et al. (2017b) and Goedhart et al. (2018a).

3.1 Introduction

In the previous chapter we have derived corrections for the Shewhart \bar{X} and \bar{X} control charts based on the bias criterion. This provides a specified control chart performance in expectation. However, as discussed in Section 1.4, there may still be a large proportion of practitioners with unsatisfactory control chart performance when this criterion is used. To this end, the exceedance probability criterion, which has been of great interest in current SPM literature, provides an alternative design criterion. While Albers and Kallenberg (2005) give relatively simple closed form correction terms for different performance measures, a bootstrap approach to obtain more accurate correction factors is becoming more common, see for example Jones and Steiner (2012),

Gandy and Kvaløy (2013) and Saleh et al. (2015a,b).

In this chapter we determine an analytical correction term for the Shewhart \bar{X} and \bar{X} control charts under normal theory in a general setup. This term is applicable to both one- and two-sided control charts, and can make use of different estimators for location and spread. Also, the specifications of the correction can be adjusted by specifying the parameters. This gives the freedom to adapt the correction to the circumstances as is desired by the practitioner. It turns out that our corrections perform better than the correction terms of Albers and Kallenberg (2005). Next to that, since our correction terms are analytical expressions, the computation time is negligible.

Furthermore, we compare our approach with the construction of two-sided tolerance intervals for normal populations. This problem has a long history of academic interest, see for example Wald and Wolfowitz (1946), Weissberg and Beatty (1960), Gardiner and Hull (1966), Howe (1969) and Krishnamoorthy and Mathew (2009). In those publications various approximations and practical guidelines were developed along with rigorous justification for the guarantee of tolerance probability. Finally, we compare our limits with self-starting control charts introduced by Hawkins (1987) and Quesenberry (1991), which also guarantee the conditional and unconditional in-control performance in an exact way. In all comparisons we show that our corrections perform better and/or are more general than the existing ones.

This chapter is organized as follows. In the next section we present our model and approach. In Section 3.3 we derive the required correction term, illustrate its performance and compare the results with the case of uncorrected limits. In Section 3.4 we address the out-of-control performance of the proposed method. In Section 3.5 we illustrate the performance and consequences of the proposed correction terms, compared to existing methods. Concluding remarks and recommendations are given in Section 3.7.

3.2 Model and Approach

Similarly as in the previous chapter, we consider estimated control limits as described in (2.7), where \tilde{K} again represents an adjusted control limit factor, this time based on the exceedance probability criterion. Recall from Section

2.2 that the true probability of a false alarm is a function of the estimated control limits, and is thus dependent on \tilde{K} , m , n , $\hat{\mu}_0$ and $\hat{\sigma}_0$. Given the estimators $\hat{\mu}_0$ and $\hat{\sigma}_0$ it is possible to calculate the FAR conditional on these estimators. Consider Y_i to be the Phase II observations as in Section 2.2, and consider the case that the process is in-control (i.e. $\delta = 0$). The conditional FAR (CFAR), in the derivations in this chapter denoted as $P_{mn}(\tilde{K}; \hat{\mu}_0, \hat{\sigma}_0)$ following Goedhart et al. (2017b), can then be written as

$$\begin{aligned}
 P_{mn}(\tilde{K}; \hat{\mu}_0, \hat{\sigma}_0) &= 1 - P\left(\widehat{LCL} < \bar{Y}_i < \widehat{UCL}\right) \\
 &= 1 - P\left(\bar{Y}_i < \hat{\mu}_0 + \tilde{K} \frac{\hat{\sigma}_0}{\sqrt{n}}\right) + P\left(\bar{Y}_i < \hat{\mu}_0 - \tilde{K} \frac{\hat{\sigma}_0}{\sqrt{n}}\right) \\
 &= 1 - \Phi\left(\frac{\hat{\mu}_0 - \mu_0}{\sigma_0/\sqrt{n}} + \tilde{K} \frac{\hat{\sigma}_0}{\sigma_0}\right) + \Phi\left(\frac{\hat{\mu}_0 - \mu_0}{\sigma_0/\sqrt{n}} - \tilde{K} \frac{\hat{\sigma}_0}{\sigma_0}\right)
 \end{aligned} \tag{3.1}$$

which can in turn be rewritten as

$$P_{mn}(\tilde{K}; Z, W) = 1 - \Phi\left(\frac{Z}{\sqrt{m}} + \tilde{K}W\right) + \Phi\left(\frac{Z}{\sqrt{m}} - \tilde{K}W\right) \tag{3.2}$$

where $Z = (\hat{\mu}_0 - \mu_0)/(\sigma_0/\sqrt{mn})$ and $W = \hat{\sigma}_0/\sigma_0$. Note that Z and W are random variables, whose distributions depend on the estimators $\hat{\mu}_0$ and $\hat{\sigma}_0$ respectively. Consequently, the unconditional FAR (i.e. before $\hat{\mu}_0$ and $\hat{\sigma}_0$ are obtained) is a random variable as well. Because practitioners use different Phase I samples, their estimations will vary, resulting in a different CFAR for their control charts. This variation is known as practitioner-to-practitioner variability, as described in Section 1.3.

If the distribution of $\hat{\mu}_0$ is symmetric (such as is the case for $\hat{\mu}_0 = \bar{\bar{X}}$), the corrections for one-sided control charts with either a UCL or an LCL are identical. However, one can not simply apply these corrections separately to two-sided control charts. For the equivalence of the one-sided corrections, note that the CFAR when using only a UCL is equal to the first part of (3.2),

and can be written as

$$1 - \Phi\left(\frac{Z}{\sqrt{m}} + \tilde{K}W\right) = \Phi\left(-\frac{Z}{\sqrt{m}} - \tilde{K}W\right). \quad (3.3)$$

If $\hat{\mu}_0$ follows a symmetric distribution around μ_0 , consequently this also holds for Z around 0. That means that the distribution of (3.3) is in that case equal to that of $\Phi\left(\frac{Z}{\sqrt{m}} - \tilde{K}W\right)$, which is the CFAR for a control chart using an LCL only. Therefore, both one-sided charts require the same correction. However, applying this one-sided correction to both sides of a two-sided chart is essentially the same as correcting for twice the CFAR of one side, which is distributed as

$$2\Phi\left(\frac{Z}{\sqrt{m}} - \tilde{K}W\right) = 1 - \Phi\left(-\frac{Z}{\sqrt{m}} + \tilde{K}W\right) + \Phi\left(\frac{Z}{\sqrt{m}} - \tilde{K}W\right) \quad (3.4)$$

Note that the expression in (3.4) is different from the actual two-sided CFAR as defined in (3.2). The intuition behind it is that in the two-sided case an underestimation of the mean increases $P\left(\bar{Y}_i > \widehat{UCL}\right)$, but at the same time decreases $P\left(\bar{Y}_i < \widehat{LCL}\right)$ (and vice versa for an overestimation of the mean). This effect is not taken into account when one simply applies one-sided corrections to a two-sided control chart. For this reason, one-sided and two-sided control charts should be treated separately.

As mentioned, in this chapter we derive a correction term that is based on the exceedance probability criterion. This approach aims to correct the control chart limits such that at least a specified minimum in-control performance is obtained with a specified probability. More specifically, when considering the FAR as performance measure, the correction term aims to obtain a CFAR that is equal to $(1+\epsilon)\alpha_0$ or smaller with a probability of $1-p$. Here ϵ determines our minimum threshold, proportional to the nominal level, while p represents the probability of obtaining a control chart performance lower than the specified minimum. In our view, p should be small (e.g. 0.05 or 0.10) and ϵ may be slightly larger (e.g. 0.2 to 0.5). This is because $1-p$ represents the probability for the practitioners with which the minimum performance is satisfied, while ϵ determines our *minimum* performance threshold. Our choices of p and ϵ

are the same as in recent research, see for example Albers and Kallenberg (2005), and for $\epsilon = 0$ also Saleh et al. (2015a). Hence, for the FAR we derive a correction term such that

$$P\left(P_{mn}(\tilde{K}; Z, W) < (1 + \epsilon)\alpha_0\right) = 1 - p. \quad (3.5)$$

Note that for the ARL in the in-control situation large CARLs are preferred. The criterion would then be

$$P\left(\frac{1}{P_{mn}(\tilde{K}; Z, W)} > (1 - \epsilon)\frac{1}{\alpha_0}\right) = 1 - p. \quad (3.6)$$

The required correction term for the ARL can be obtained through solving the term for the FAR, whilst replacing ϵ by $\tilde{\epsilon} = \frac{\epsilon}{1-\epsilon}$. This is because the left hand side of (3.6) is equivalent to

$$\begin{aligned} &P\left(P_{mn}(\tilde{K}; Z, W) < \frac{1}{1 - \epsilon}\alpha_0\right) \\ &= P\left(P_{mn}(\tilde{K}; Z, W) < \left(1 + \frac{\epsilon}{1 - \epsilon}\right)\alpha_0\right) \\ &= P\left(P_{mn}(\tilde{K}; Z, W) < (1 + \tilde{\epsilon})\alpha_0\right) \end{aligned}$$

Because of this equivalence, our further derivations in this chapter are based on the FAR as performance measure.

In order to determine the correction term c , we require information on the distribution of $P_{mn}(\tilde{K}; Z, W)$, as we need to find the value c for which the $1 - p$ percentile of the distribution of $P_{mn}(\tilde{K}; Z, W)$ equals $(1 + \epsilon)\alpha_0$. Although the exact distribution of $P_{mn}(\tilde{K}; Z, W)$ for an arbitrary \tilde{K} is unknown, it is possible to calculate its moments using integrals, similar to Chen (1997). The first and second moment can be calculated by

$$E\left(P_{mn}(\tilde{K}; Z, W)\right) = \int_{-\infty}^{\infty} \int_0^{\infty} P_{mn}(\tilde{K}; z, w) f(z) f(w) dw dz \quad (3.7)$$

and

$$E(P_{mn}^2(\tilde{K}; Z, W)) = \int_{-\infty}^{\infty} \int_0^{\infty} P_{mn}^2(\tilde{K}; z, w) f(z) f(w) dw dz \quad (3.8)$$

respectively. Here $f(z)$ equals the probability density of $Z = (\hat{\mu}_0 - \mu_0) / (\sigma_0 / \sqrt{mn})$, and $f(w)$ is the probability density of $W = \hat{\sigma}_0 / \sigma_0$. The variance of $P_{mn}(\tilde{K}; Z, W)$ can be calculated through

$$Var(P_{mn}(\tilde{K}; Z, W)) = E(P_{mn}^2(\tilde{K}; Z, W)) - E^2(P_{mn}(\tilde{K}; Z, W)) \quad (3.9)$$

Using these moments, we approximate the distribution of $P_{mn}(\tilde{K}; Z, W)$ by a $A\chi_B^2/B$ distribution (for a detailed motivation we refer to Appendix 3.8). Since $E\left(\frac{A\chi_B^2}{B}\right) = A$ and $Var\left(\frac{A\chi_B^2}{B}\right) = \frac{2A^2}{B}$, we have

$$A = E(P_{mn}(\tilde{K}; Z, W)) \quad (3.10)$$

and

$$B = \frac{2A^2}{Var(P_{mn}(\tilde{K}; Z, W))} = \frac{2E^2(P_{mn}(\tilde{K}; Z, W))}{Var(P_{mn}(\tilde{K}; Z, W))} \quad (3.11)$$

Note that both the expectation and variance depend on $\tilde{K} = K + c$.

3.3 Correction Terms

3.3.1 Two-sided Control Limits

As mentioned in the previous section, both A and B as in (3.10) and (3.11) depend on \tilde{K} . This means that changing K to $\tilde{K} = K + c$ changes the $A\chi_B^2/B$ distribution as well. The correction factor c has to be determined such that for $\tilde{K} = K + c$, the $(1 - p)$ 'th percentile of this distribution lies at $(1 + \epsilon)\alpha_0$.

If we denote the expectation and variance of $P_{mn}(\tilde{K}; Z, W)$ in (3.7) and (3.9), evaluated at K , by E and V respectively, and their derivatives with respect to \tilde{K} , evaluated at K , by $\frac{dE}{d\tilde{K}}$ and $\frac{dV}{d\tilde{K}}$ respectively, then the correction

term can be expressed by

$$c = \frac{\Phi^{-1}(1-p) - Y(K)}{Y'(K)} \quad (3.12)$$

where

$$Y(K) = \sqrt[3]{(1+\epsilon)\alpha_0} \frac{3E^{2/3}}{\sqrt{V}} - \frac{3E}{\sqrt{V}} + \frac{\sqrt{V}}{3E} \quad (3.13)$$

$$Y'(K) = \sqrt[3]{(1+\epsilon)\alpha_0} \frac{2E^{-1/3} \sqrt{V} \frac{dE}{dK} - \frac{3E^{2/3}}{2\sqrt{V}} \frac{dV}{dK}}{V} - \frac{3 \frac{dE}{dK} \sqrt{V} - \frac{3E}{2\sqrt{V}} \frac{dV}{dK}}{V} + \frac{\frac{3E}{2\sqrt{V}} \frac{dV}{dK} - 3 \frac{dE}{dK} \sqrt{V}}{9E^2} \quad (3.14)$$

and where

$$\begin{aligned} \frac{dE}{dK} &= \int_{-\infty}^{\infty} \int_0^{\infty} \frac{dP_{mn}(K; z, w)}{d\tilde{K}} f(z) f(w) dw dz \\ &= \int_{-\infty}^{\infty} \int_0^{\infty} -w \left[\phi\left(\frac{z}{\sqrt{m}} + Kw\right) + \phi\left(\frac{z}{\sqrt{m}} - Kw\right) \right] f(z) f(w) dw dz \end{aligned} \quad (3.15)$$

$$\frac{dV}{dK} = \frac{d(E^2)}{dK} - 2E \frac{dE}{dK} \quad (3.16)$$

with

$$\begin{aligned} \frac{d(E^2)}{dK} &= \int_{-\infty}^{\infty} \int_0^{\infty} \frac{dP_{mn}^2(K; z, w)}{d\tilde{K}} f(z) f(w) dw dz \\ &= -2 \int_{-\infty}^{\infty} \int_0^{\infty} w P_{mn}(K; z, w) \left[\phi\left(\frac{z}{\sqrt{m}} + Kw\right) + \phi\left(\frac{z}{\sqrt{m}} - Kw\right) \right] f(z) f(w) dw dz \end{aligned} \quad (3.17)$$

For a detailed derivation of the correction term we refer to Appendix 3.9.

Note that, instead of a minimum performance threshold that is relative (through ϵ) to α_0 (or $1/\alpha_0$), one could also chose to specify an absolute minimum performance threshold. This can be done by setting $\alpha_0 = CFAR_0 = 1/CARL_0$ and $\epsilon = 0$, where $CFAR_0$ and $CARL_0$ are the desired threshold values for the CFAR or CARL respectively. However, in order to make a better comparison between the corrected and uncorrected charts, we consider relative thresholds. The same approach has been used by other authors (e.g. Albers and Kallenberg, 2005).

We use simulation to evaluate the performance of the proposed correction term. In order to do this we need to define our estimators, and determine the corresponding distributions $f(z)$ and $f(w)$ of $Z = (\hat{\mu}_0 - \mu_0)/(\sigma_0/\sqrt{mn})$ and $W = \hat{\sigma}_0/\sigma_0$ respectively. We consider $\hat{\mu}_0 = \bar{X}$ as the estimator for location, in which case $f(z)$ equals the standard normal probability density. For the spread we consider two estimators. The first estimator, which we use in the case that $n > 1$ (groups), is based on the pooled standard deviation (S_{pooled}), and is equal to

$$\hat{\sigma}_1 = \frac{S_{pooled}}{c_4(m(n-1) + 1)} \quad (3.18)$$

where $c_4(m(n-1) + 1)$ is such that $\hat{\sigma}_1$ is an unbiased estimator of σ_0 . For this estimator $W = \hat{\sigma}_1/\sigma_0$ is distributed as $\tau\chi_\nu/\sqrt{\nu}$ with $\tau = 1/c_4(m(n-1) + 1)$ and $\nu = m(n-1)$. The probability density of $\tau\chi_\nu/\sqrt{\nu}$ equals

$$f(w; \tau, \nu) = \left(\frac{2}{\tau}\right) \frac{(\nu/2)^{\nu/2}}{\Gamma(\nu/2)} \left(\frac{w}{\tau}\right)^{\nu-1} \exp\left(-\frac{\nu}{2} \left(\frac{w}{\tau}\right)^2\right) \quad (3.19)$$

(see also Chen, 1997). The second estimator, which is used in the case $n = 1$ (individuals), is based on the average moving range \overline{MR} and is equal to

$$\hat{\sigma}_2 = \frac{\overline{MR}}{d_2(2)} \quad (3.20)$$

where $d_2(2) = \frac{2}{\sqrt{\pi}}$ which yields that $\hat{\sigma}_2$ is an unbiased estimator of σ_0 . Although the exact distribution of \overline{MR} is not easy to obtain, the distribution of

$W = \hat{\sigma}_2/\sigma_0$ can be approximated by $\beta\chi_\gamma/\sqrt{\gamma}$, where

$$\begin{aligned}\beta &= \sqrt{\text{Var}(W) + 1} \\ \gamma &= \frac{1}{2} \left(1 + \frac{1}{\text{Var}(W)} \right).\end{aligned}\tag{3.21}$$

See for example Roes et al. (1993) for this approximation, or Patnaik (1950) for a similar one. The variance of $W = \hat{\sigma}_2/\sigma_0$ is investigated by Cryer and Ryan (1990) and can be approximated by

$$\text{Var} \left(\frac{\hat{\sigma}_2}{\sigma_0} \right) = \left[\frac{0.8264m - 1.082}{(m - 1)^2} \right]\tag{3.22}$$

Note that it is possible to use any other estimator for location and spread, by applying their corresponding probability functions $f(z)$ and $f(w)$. In Table 3.1 and Table 3.2 several estimators for μ_0 and σ_0 respectively are listed. For the estimators of μ_0 the (approximated) probability density function $f(z)$ is tabulated. For the mean we have an exact distribution and for the median the distribution has been approximated (cf. Johnson and Kotz, 1970). For estimators of σ_0 , the distribution of $W = \hat{\sigma}_0/\sigma_0$ is either exact or approximated by a $\zeta\chi_\lambda/\sqrt{\lambda}$ distribution, so that $f(w)$ is given by (3.19). The approximation is similar to that of \overline{MR} , by determining the corresponding values of ζ and λ . The required values of ζ and λ for the considered estimators are listed in Table 3.2. We refer to Roes et al. (1993) for the approximations, and to Albers and Kallenberg (2005) for explicit expressions of the listed estimators of σ_0 .

As the ARL is the most commonly used performance measure of control charts, we evaluate the performance of our proposed correction terms based on that. The corresponding criterion function is given in (3.6). To calculate the required correction terms, we use the model as described in the previous section, thus with replacing ϵ by $\tilde{\epsilon} = \frac{\epsilon}{1-\epsilon}$ in the correction terms for the FAR. We have calculated the correction terms and simulated their performance for a wide range of parameter values. For each combination of parameter values 1,000,000 simulation runs are performed. The relative standard errors of all reported EARL values in Tables 3.3-3.8 are less than 1%. Table 3.3 and Table

3.4 illustrate, for multiple combinations of n and m , the correction term and its performance compared to an uncorrected chart, each for a different set of p , ϵ and α_0 . The performance is measured by the exceedance probability as in (3.6), while for the comparisons the EARLs are also given.

The results illustrate that after implementing our suggested correction term, the exceedance probability is very close to the desired level. Especially for small sample sizes the differences between the corrected and the uncorrected chart are large. Note that this is in agreement with the sample size recommendations by for example Quesenberry (1993), which states that about $400/(n - 1)$ subgroups of size n are required in Phase I for the \bar{X} chart to behave properly on average. Furthermore, Saleh et al. (2015b) show that far larger amounts of Phase I data are required when also taking the practitioner-to-practitioner variability into account. As more data is available in Phase I, the required correction becomes smaller. At some point the exceedance probability of the uncorrected chart is already below the desired level, meaning that the correction term becomes negative to increase the exceedance probability. The use of a negative correction term in this case is questionable, as it will decrease the in-control performance in that situation. However, the out-of-control performance improves, whilst keeping the in-control performance at a desired level.

As mentioned earlier, this correction term is applicable for any estimator. Previously we have shown the results for the \bar{X} chart using an estimator for σ_0 based on the pooled standard deviation (see Equation (3.18)). In this section we also show the results for the X chart, by using an unbiased version of the average moving range as estimator of σ , as in (3.20). Similarly as for the \bar{X} chart, Table 3.5 and Table 3.6 indicate the correction term and its performance compared to the uncorrected chart, for multiple values of m (and $n = 1$), each for a different set of p , ϵ and α_0 . Even though we use an approximation of the density of $\hat{\sigma}_2$, the correction term still performs well, with the exceedance probabilities close to the desired level.

Estimator	$f(z)$
Average (\bar{X})	$\frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}z^2\right)$
Median (\tilde{X})	$\frac{1}{\pi} \exp\left(-\frac{z^2}{\pi}\right)$

Table 3.1: Estimators for location, including the (approximated) probability density function $f(z)$ of $Z = (\hat{\mu}_0 - \mu_0)/(\sigma_0/\sqrt{mn})$.

Estimator	ζ and λ	var(W)
<i>Individuals (n = 1)</i>		
Average moving range, $\hat{\sigma}_0 = \frac{MR(\mathbf{X})}{d_2(2)}$	$\zeta = \sqrt{Var(W) + 1},$ $\lambda = \frac{1}{2} \left(1 + \frac{1}{Var(W)}\right)$	$\frac{0.8264m-1.082}{(m-1)^2}$
Interquartile range $\hat{\sigma}_0 = \frac{IQR(\mathbf{X})}{1.349}$	$\zeta = \sqrt{Var(W) + 1},$ $\lambda = \frac{1}{2} \left(1 + \frac{1}{Var(W)}\right)$	$\frac{2.46}{1.820m}$
Sample standard deviation $\hat{\sigma}_0 = \frac{s}{c_4(m)}$	$\zeta = 1/c_4(m),$ $\lambda = m - 1$	$\frac{1-c_4^2(m)}{c_4^2(m)}$
<i>Groups (n > 1)</i>		
Average standard deviation $\hat{\sigma}_0 = \frac{\bar{s}}{c_4(n)}$	$\zeta = \sqrt{Var(W) + 1},$ $\lambda = \frac{1}{2} \left(1 + \frac{1}{Var(W)}\right)$	$\frac{1-c_4^2(n)}{mc_4^2(n)}$
Pooled standard deviation $\hat{\sigma}_0 = \frac{S_{pooled}}{c_4(m(n-1)+1)}$	$\zeta = 1/c_4(m(n-1)+1),$ $\lambda = m(n-1)$	$\frac{1-c_4^2(m(n-1)+1)}{c_4^2(m(n-1)+1)}$

Table 3.2: Estimators for spread, including the (approximated) probability density function $f(w)$ of $W = \hat{\sigma}_0/\sigma_0$.

EXCEEDANCE PROBABILITY CRITERION FOR CONTROL CHARTS FOR LOCATION

m	n	Correction c	Exc. Pr. (cor)	Exc. Pr. (unc)	EARL (cor)	EARL (unc)
25	3	0.5687	0.0516	0.4836	7261	569
	5	0.3970	0.0478	0.4715	1890	418
	9	0.2822	0.0473	0.4534	984	364
50	3	0.3532	0.0483	0.4275	1721	449
	5	0.2311	0.0494	0.3956	879	389
	9	0.1541	0.0501	0.3595	610	362
75	3	0.2615	0.0495	0.3908	1078	418
	5	0.1651	0.0507	0.3451	674	382
	9	0.1034	0.0512	0.2922	514	364
100	3	0.2097	0.0501	0.3617	850	405
	5	0.1302	0.0511	0.3132	585	376
	9	0.0756	0.0519	0.2434	469	365
150	3	0.1540	0.0504	0.3239	662	389
	5	0.0884	0.0521	0.2555	503	374
	9	0.0447	0.0519	0.1752	425	366
200	3	0.1202	0.0512	0.2893	579	384
	5	0.0647	0.0520	0.2120	463	373
	9	0.0274	0.0522	0.1289	402	367
250	3	0.0980	0.0516	0.2600	531	381
	5	0.0491	0.0521	0.1792	438	372
	9	0.0160	0.0523	0.0971	388	368

Table 3.3: Correction terms c for the \bar{X} chart, including the corresponding exceedance probability and EARL for the corrected (cor) and uncorrected (unc) control limits. Parameter values are $\alpha_0 = 0.0027$, $p = 0.05$ and $\epsilon = 0.2$.

m	n	c	Exc. Pr. (cor)	Exc. Pr. (unc)	EARL (cor)	EARL (unc)
25	3	0.2325	0.0975	0.3049	275	123
	5	0.1216	0.0967	0.2352	151	104
	9	0.0507	0.0946	0.1615	111	96
50	3	0.0875	0.0975	0.1956	144	110
	5	0.0124	0.0987	0.1150	105	101
	9	-0.0349	0.0988	0.0518	88	97
75	3	0.0289	0.0991	0.1347	116	106
	5	-0.0305	0.0994	0.0601	92	101
	9	-0.0688	0.1004	0.0172	80	98
100	3	-0.0040	0.0996	0.0946	103	104
	5	-0.0529	0.1005	0.0337	86	100
	9	-0.0875	0.1017	0.0059	77	98
150	3	-0.0390	0.1005	0.0511	91	102
	5	-0.0802	0.1008	0.0103	79	100
	9	-0.1081	0.1027	0.0007	73	99
200	3	-0.0606	0.1005	0.0272	85	102
	5	-0.0957	0.1016	0.0033	76	100
	9	-0.1197	0.1032	0.0001	71	99
250	3	-0.0749	0.1012	0.0152	82	101
	5	-0.1060	0.1020	0.0011	74	100
	9	-0.1273	0.1035	0.0000	69	99

Table 3.4: Correction terms c for the \bar{X} chart, including the corresponding exceedance probability and EARL for the corrected (cor) and uncorrected (unc) control limits. Parameter values are $\alpha_0 = 0.01$, $p = 0.1$ and $\epsilon = 0.4$.

EXCEEDANCE PROBABILITY CRITERION FOR CONTROL CHARTS FOR LOCATION

m	c	Exc. Pr. (cor)	Exc. Pr. (unc)	EARL (cor)	EARL (unc)
50	0.6930	0.0563	0.4723	78131	1086
75	0.5510	0.0492	0.4492	9902	697
100	0.4596	0.0471	0.4308	4156	580
150	0.3495	0.0470	0.4041	1916	491
200	0.2852	0.0475	0.3819	1317	455
250	0.2425	0.0483	0.3633	1053	436
500	0.1419	0.0502	0.2979	655	401
1000	0.0760	0.0516	0.2191	498	385

Table 3.5: Correction terms c for the individuals X chart, with the corresponding exceedance probability and EARL for the corrected (cor) and uncorrected (unc) control limits. Parameter values are $\alpha_0 = 0.0027$, $p = 0.05$ and $\epsilon = 0.2$.

m	c	Exc. Pr. (cor)	Exc. Pr. (unc)	EARL (cor)	EARL (unc)
50	0.3176	0.0979	0.3279	658	174
75	0.2127	0.0959	0.2773	301	140
100	0.1512	0.0959	0.2393	212	127
150	0.0808	0.0966	0.1837	151	117
200	0.0407	0.0970	0.1440	127	112
250	0.0142	0.0984	0.1156	114	109
500	-0.0483	0.0993	0.0409	91	105
1000	-0.0898	0.1011	0.0066	79	102

Table 3.6: Correction terms c for the individuals X chart, with the corresponding exceedance probability and EARL for the corrected (cor) and uncorrected (unc) control limits. Parameter values are $\alpha_0 = 0.01$, $p = 0.1$ and $\epsilon = 0.4$.

3.3.2 One-sided Control Limits

A fairly straightforward change in the previous derivations leads to the required correction term in case we are dealing with either a UCL or LCL only (one-sided). As the correction term for the UCL and LCL is the same, we consider the case of the UCL. The main difference lies in the change of (3.2) to

$$P_{mn}(\tilde{K}; Z, W) = 1 - \Phi\left(\frac{z}{\sqrt{m}} + \tilde{K}w\right) \quad (3.23)$$

where K in $\tilde{K} = K + c$ is now equal to $\Phi^{-1}(1 - \alpha_0)$. Then this formula should be used to calculate the expectation and variance of $P_{mn}(\tilde{K}; Z, W)$ in (3.7) and (3.9). The expressions in (3.15) and (3.17) will change in

$$\begin{aligned} \frac{dE}{d\tilde{K}} &= \int_{-\infty}^{\infty} \int_0^{\infty} \frac{dP_{mn}(K; z, w)}{d\tilde{K}} f(z) f(w) dw dz \\ &= \int_{-\infty}^{\infty} \int_0^{\infty} -w\phi\left(\frac{z}{\sqrt{m}} + Kw\right) f(z) f(w) dw dz \end{aligned} \quad (3.24)$$

$$\begin{aligned} \frac{d(E^2)}{dK} &= \int_{-\infty}^{\infty} \int_0^{\infty} \frac{dP_{mn}^2(K; z, w)}{d\tilde{K}} f(z) f(w) dw dz \\ &= -2 \int_{-\infty}^{\infty} \int_0^{\infty} wP_{mn}(K; z, w)\phi\left(\frac{z}{\sqrt{m}} + Kw\right) f(z) f(w) dw dz \end{aligned} \quad (3.25)$$

which can be used to obtain the corresponding expression in (3.16):

$$\frac{dV}{dK} = \frac{d(E^2)}{dK} - 2E \frac{dE}{dK} \quad (3.26)$$

The corresponding correction term c in (3.12) for the one-sided control limit can be obtained by implementing the resulting expressions in (3.13) and (3.14).

3.4 Out-of-Control Performance

Correcting the control chart limits in order to guarantee a minimum performance clearly has an advantage in the in-control situation. However, this inevitably leads to a deterioration of the out-of-control performance. In the previous section, more specifically in Tables 3.3, 3.4, 3.5 and 3.6, the EARLs of the corrected and uncorrected charts are given. It can already be seen that there is a large difference between the corrected and uncorrected control limits for small sample sizes. As more information becomes available, and as a consequence the correction term becomes smaller, this difference becomes smaller. For the out-of-control situation, this behavior is very similar. To illustrate this we have simulated the EARL of the corrected and uncorrected \bar{X} charts for different sizes of shifts. We consider $p = 0.05$, $\epsilon = 0.2$ and $\alpha_0 = 0.0027$, and have again simulated for a wide range of values for m and n . For the out-of-control situation we consider a shift in the mean, such that the Phase II subgroup average \bar{Y}_i comes from a $N(\mu_0 + \delta\sigma_0/\sqrt{n}, \sigma_0/\sqrt{n})$ distribution, with δ equal to 0.5, 1, or 2. The results are listed in Table 3.7.

We find that in this situation, for small sample sizes, the differences in EARLs between the corrected and uncorrected charts are substantial. Note that we have chosen a rather strict set of parameters, as we guarantee an in-control CARL of at least 296 with a probability of 90%. The out-of-control performance becomes better as α_0 , ϵ and p increase. More specifically, increasing the value of ϵ and/or α_0 results in a lower minimum in-control performance threshold, and consequently in a lower EARL. On the other hand, a larger value of p means that we allow a larger proportion of the in-control CARLs to be below the minimum performance threshold. This has the consequence that the EARL will be smaller. Thus, increasing any of the parameters α_0 , ϵ and/or p leads to lower EARL values, which is beneficial for the out-of-control situation, but of course not for the in-control situation. This trade-off between in-control and out-of-control performance is inherent to control charts. The advantage of our proposed method is that the parameters can be easily adjusted, in order to balance the performance of the chart as is desired by the practitioner. As addition, Table 3.8 illustrates the corrections, exceedance probabilities and EARL values of the corrected and uncorrected charts for various combinations of m and n , for $\alpha_0 = 0.01$, $\epsilon = 0.2$ and $p = 0.1$.

		δ					
		0.5		1		2	
m	n	EARL (cor)	EARL (unc)	EARL (cor)	EARL (unc)	EARL (cor)	EARL (unc)
25	3	2906	278	509	72	28	8
	5	856	213	190	59	15	7
	9	473	189	117	54	12	7
50	3	685	205	151	55	13	7
	5	379	182	93	51	10	7
	9	274	172	72	48	9	7
75	3	431	186	102	51	11	7
	5	287	173	74	48	9	7
	9	226	166	61	47	8	7
100	3	342	178	85	49	9	7
	5	248	168	65	47	8	7
	9	204	163	55	46	7	6
150	3	269	169	69	47	8	7
	5	212	163	57	46	7	6
	9	183	161	51	45	7	6
200	3	236	165	62	46	8	6
	5	195	161	53	45	7	6
	9	172	159	48	45	7	6
250	3	218	163	58	46	8	6
	5	185	160	51	45	7	6
	9	166	158	46	45	7	6

Table 3.7: Out-of-control performance of the \bar{X} chart for shifts in the mean of size $\frac{\delta}{\sqrt{n}}$, for both the corrected (cor) and uncorrected (unc) control limits. Parameter values are $\alpha_0 = 0.0027$, $p = 0.05$ and $\epsilon = 0.2$.

EXCEEDANCE PROBABILITY CRITERION FOR CONTROL CHARTS FOR LOCATION

		δ					
		0.5		1		2	
m	n	EARL (cor)	EARL (unc)	EARL (cor)	EARL (unc)	EARL (cor)	EARL (unc)
25	3	143	69	42	23	6	4
	5	84	60	27	21	5	4
	9	64	56	22	20	4	4
50	3	74	58	24	20	4	4
	5	57	55	19	19	4	4
	9	48	53	17	18	3	4
75	3	60	56	20	19	4	4
	5	49	53	17	18	4	4
	9	44	52	16	18	3	4
100	3	53	54	18	19	4	4
	5	46	52	16	18	3	4
	9	41	52	15	18	3	4
150	3	47	52	17	18	3	4
	5	42	51	15	18	3	4
	9	39	51	14	18	3	4
200	3	44	52	16	18	3	4
	5	40	51	15	18	3	4
	9	38	51	14	18	3	4
250	3	43	51	15	18	3	4
	5	39	51	14	18	3	4
	9	37	51	14	18	3	4

Table 3.8: Out-of-control performance of the \bar{X} chart for shifts in the mean of size $\frac{\delta}{\sqrt{n}}$, for both the corrected (cor) and uncorrected (unc) control limits. Parameter values are $\alpha_0 = 0.01$, $p = 0.1$ and $\epsilon = 0.4$.

3.5 Comparison with Existing Methods

In order to illustrate the performance of the proposed correction term, a comparison is made with the existing methods. First, we make a comparison with the methods of Albers and Kallenberg (2005) and Gandy and Kvaløy (2013). Next, we compare the proposed control chart with the self-starting Q chart of Quesenberry (1993). Finally, we compare the proposed X chart with tolerance intervals for a normal distribution, since these use an equivalent criterion for $n = 1$.

3.5.1 Comparison of Shewhart X and \bar{X} Control Charts

We consider the two-sided case, with $\hat{\mu} = \bar{\bar{X}}$ and $\hat{\sigma}_0$ as in (3.18) for $n > 1$, and (3.20) for $n = 1$. For this situation Albers and Kallenberg (2005) proposed control limits of the form

$$\begin{aligned}\widehat{UCL}_{AK} &= \bar{\bar{X}} + K \frac{\hat{\sigma}_0}{\sqrt{n}}(1 + c_{AK}) \\ \widehat{LCL}_{AK} &= \bar{\bar{X}} - K \frac{\hat{\sigma}_0}{\sqrt{n}}(1 + c_{AK})\end{aligned}\tag{3.27}$$

where

$$c_{AK} = \frac{\Phi^{-1}(1 - p)\theta}{\sqrt{mn}} - \frac{\epsilon}{K^2}\tag{3.28}$$

and where $\theta^2 = \lim_{mn \rightarrow \infty} \left[mn \text{var} \left(\frac{\hat{\sigma}_0}{E\hat{\sigma}_0} \right) \right]$. For the estimators $\hat{\sigma}_0$ as in (3.18) and (3.20) the value of θ^2 equals $\frac{n}{2(n-1)}$ and 0.826 respectively. Note that their proposed correction (c_{AK}) only depends on the initial Phase I sample through its size (m and n). This is in line with our correction. For the bootstrap approach of Gandy and Kvaløy (2013) this is different, as the actual correction depends on the sample estimates. Therefore, in the comparison only the realized exceedance probabilities are shown. For the explicit bootstrapping procedure we refer to Saleh et al. (2015a), who provide a simplification of the computations in Gandy and Kvaløy (2013) approach for the Shewhart control

EXCEEDANCE PROBABILITY CRITERION FOR CONTROL CHARTS FOR LOCATION

m	Proposed	AK	Bootstrap	m	Proposed	AK	Bootstrap
25	0.0916	0.1822	0.1046	25	0.0921	0.1918	0.1044
50	0.0934	0.1580	0.0993	50	0.0936	0.1739	0.0992
75	0.0957	0.1462	0.0989	75	0.0973	0.1649	0.0989
100	0.0947	0.1440	0.1010	100	0.0982	0.1673	0.1010
150	0.0967	0.1350	0.1054	150	0.1008	0.1665	0.1054
200	0.0989	0.1318	0.0983	200	0.0999	0.1662	0.0983
250	0.0998	0.1263	0.1021	250	0.1004	0.1708	0.1022

(a) $n = 5, \alpha_0 = 0.0027, p = 0.1$ and $\epsilon = 0$. (b) $n = 5, \alpha_0 = 0.0027, p = 0.1$ and $\epsilon = 0.2$.

Table 3.9: Exceedance probabilities of the proposed correction, Albers and Kallenberg (2005) (AK) correction and Gandy and Kvaløy (2013) (bootstrap) method for different values of m .

chart. We performed the bootstrap procedure (based on 1001 bootstraps) for 10,000 simulated Phase I samples, in order to calculate the exceedance probability. The results of the simulations are listed in Table 3.9a for $p = 0.1, \alpha_0 = 0.0027$, and $\epsilon = 0$, and in 3.9b for $p = 0.1, \alpha_0 = 0.0027$, and $\epsilon = 0.2$, each for different values of m , and $n = 5$. Results for other values of n are similar. It is clear that our correction performs much better than the correction of Albers and Kallenberg (2005), as it is closer to the desired level of $p = 0.1$. The bootstrap procedure of Gandy and Kvaløy (2013) also has a good performance. There is no real difference with our method in the sense of performance. Also, the performance of our method and the bootstrap method appears to be less sensitive to the value of ϵ , as becomes clear when changing its value. This is shown in Table 3.9b, which indicates the performance of the three methods when implementing $\epsilon = 0.2$, whilst leaving the other parameters as in Table 3.9a.

To illustrate the performance and consequences of the proposed methods graphically, the distributions of the CARLs of the different methods in the in-control ($\delta = 0$) and out-of-control situation ($\delta = 1$) are shown in Figures 3.1 and 3.2, respectively, for $m = 50, n = 5, p = 0.1, \epsilon = 0.2$ and $\alpha_0 = 0.0027$. The vertical line represents the desired threshold of the $100p$ 'th (in this case 10th) percentile of the in-control CARL distribution. The desired threshold with $\alpha_0 = 0.0027$ and $\epsilon = 0.2$ is equal to 296. As could also be seen in Table 3.9b our proposed correction and the bootstrap method perform best, with the 10th percentile close to the desired level. It is gratifying to note that with our correction term no extensive bootstrapping is needed for the Shewhart

control charts for location.

The tradeoff between in-control and out-of-control performance of the control chart also becomes clear from these figures, if we compare the proposed methods with the uncorrected chart. The corrected charts correspond with better in-control performance, but have a slower detection in the out-of-control situation.

3.5.2 Self-Starting Control Charts

There are also other control chart designs that lead to a desirable in-control performance. In particular, for normally distributed data, self-starting control charts by Hawkins (1987) and Quesenberry (1991) can guarantee a good in-control performance in the long run as well. However, the major drawback is that, because of the continuous updating of the control chart limits, there is a risk that out-of-control data influences the in-control process estimates. A small change in the process mean can therefore slowly change the control limits, making the out-of-control situations harder to detect. This can result in larger out-of-control CARLs. To illustrate this, we have simulated CARLs for both the self-starting Q chart in Quesenberry (1991) and our corrections, for both the in-control and out-of-control situation (with $\delta = 1$). As can be observed in Figure 3.3, the self-starting Q chart indeed leads to large out-of-control CARLs, much larger than CARLs of the proposed corrections. Next to that, although the in-control performance of the Q chart is very stable, our proposed correction yields much larger in-control CARLs, which is highly beneficial for the practitioners.

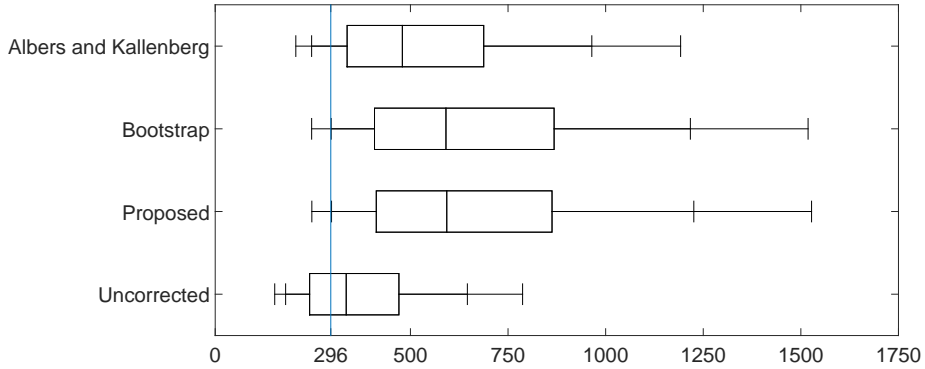


Figure 3.1: In-control ($\delta = 0$) CARL distributions for $m = 50$, $n = 5$, $\alpha_0 = 0.0027$, $p = 0.1$ and $\epsilon = 0.2$. The boxplots indicate the 5th, 10th, 25th, 50th, 75th, 90th and 95th percentiles of the distributions. The vertical line represents the desired threshold level of the CARL (296).

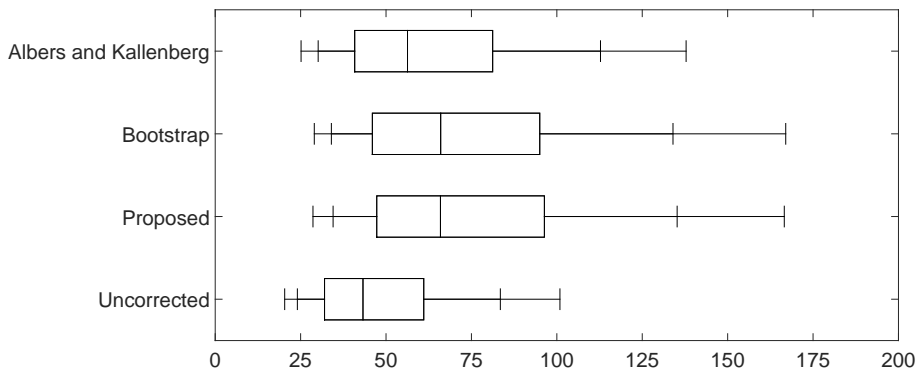


Figure 3.2: Out-of-control ($\delta = 1$) CARL distributions for $m = 50$, $n = 5$, $\alpha_0 = 0.0027$, $p = 0.1$ and $\epsilon = 0.2$. The boxplots indicate the 5th, 10th, 25th, 50th, 75th, 90th and 95th percentiles of the distributions.

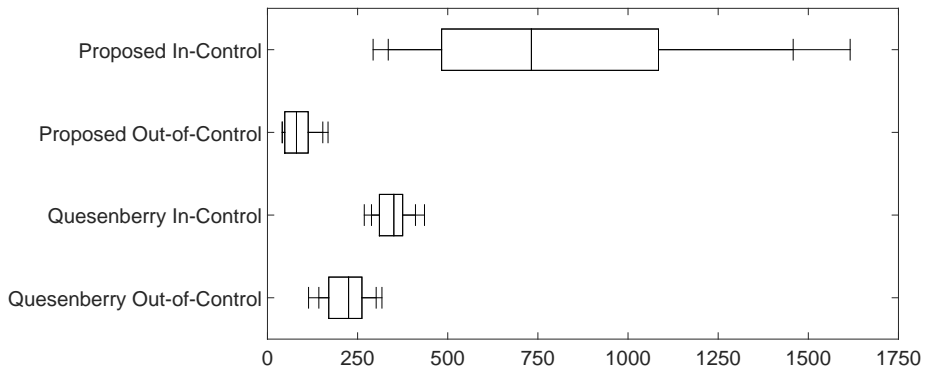


Figure 3.3: In-control ($\delta = 0$) and out-of-control ($\delta = 1$) CARL distribution of the self-starting Q chart in Quesenberry (1991) and our proposed correction when $m = 50$, $n = 5$, $\alpha_0 = 0.0027$, $p = 0.1$ and $\epsilon = 0$. The boxplots indicate the 5th, 10th, 25th, 50th, 75th, 90th and 95th percentiles of the distributions.

3.5.3 Tolerance Intervals

The literature of tolerance intervals considers a criterion that is closely related to the proposed corrections in this chapter. From Krishnamoorthy and Mathew (2009) we cite “a tolerance interval is expected to capture a certain proportion or more of the population, with a given confidence”. The tolerance intervals are based on the sample average (\bar{X}) and the sample standard deviation (S), and are in the form of $\bar{X} \pm k_2 S$, where k_2 is determined such that

$$P\left(P\left(\bar{X} - k_2 S \leq X \leq \bar{X} + k_2 S \mid \bar{X}, S\right) \geq 1 - \alpha_0\right) = 1 - p \quad (3.29)$$

Using the tolerance limits as control limits guarantees an in-control CFAR (CARL) that is smaller (larger) than α_0 ($1/\alpha_0$) with probability $1 - p$. Thus, k_2 is equivalent to our \tilde{K} in the case that $\epsilon = 0$ and $n = 1$, when using \bar{X} and S as Phase I estimators. Note that one can also incorporate ϵ by replacing α_0 by $\alpha_{tol} = (1 + \epsilon)\alpha_0$ in Equation (3.29). Note also that the tolerance limit approach can be applied when $n > 1$, by treating \bar{X} as an individual variable. However, this would mean that standard deviation of \bar{X} should be used to estimate σ_0 , rather than the standard deviation of X . This means that the within-subgroup variation is ignored in the estimation. The performance of the tolerance interval approach is then equal to the performance when $n = 1$. Because our approach does consider the within subgroup-variation, it has less uncertainty in parameter estimation, leading to less variation in ARLs for $n > 1$. However, it is possible to extend their underlying derivations in order to make it applicable for $n > 1$. We elaborate further on this in Section 3.6.

Because of the arguments mentioned above, we have compared our proposed corrections with the approximated tolerance factors as in Krishnamoorthy and Mathew (2009) for the case $n = 1$. Although they provide tolerance factors K_{KM} instead of corrections c_{KM} , we can determine their ‘correction’ by subtracting K from K_{KM} . We consider \bar{X} as estimator of μ_0 and S as estimator of σ_0 . Their proposed correction c_{KM} then equals

$$c_{KM} = \left(\frac{(n-1)\chi_{1;1-\alpha_0}^2(1/n)}{\chi_{n-1;p}^2} \right)^{\frac{1}{2}} - K, \quad (3.30)$$

where $\chi_{d;q}^2$ represents the q -quantile of a chi-square distribution with d degrees of freedom, and $\chi_{d;q}^2(\theta)$ represents the q -quantile of a noncentral chi-square distribution with d degrees of freedom and noncentrality parameter θ .

For different combinations of α_0 and p we have calculated the corrections and their performance. The results are shown in Table 3.10a for $\alpha_0 = 0.0027$ and $p = 0.05$, and in Table 3.10b for $\alpha_0 = 0.01$ and $p = 0.1$. As can be seen, there appears to be no significant difference in performance, since the resulting exceedance probabilities are close to the desired value p for both approximations.

m	c	Exc. Pr. (c)	c_{KM}	Exc. Pr. (c_{KM})	m	c	Exc. Pr. (c)	c_{KM}	Exc. Pr. (c_{KM})
50	0.6286	0.0507	0.6403	0.0484	50	0.4130	0.1013	0.4249	0.0957
75	0.4990	0.0514	0.4966	0.0521	75	0.3257	0.1071	0.3304	0.1032
100	0.4215	0.0503	0.4174	0.0519	100	0.2753	0.1026	0.2781	0.1006
150	0.3322	0.0473	0.3293	0.0486	150	0.2179	0.0997	0.2196	0.0984
200	0.2812	0.0458	0.2794	0.0469	200	0.1851	0.1010	0.1865	0.0982
250	0.2475	0.0505	0.2465	0.0514	250	0.1634	0.1067	0.1646	0.1044
500	0.1683	0.0515	0.1686	0.0512	500	0.1118	0.1000	0.1127	0.0981
1000	0.1159	0.0529	0.1165	0.0523	1000	0.0773	0.0996	0.0779	0.0979

(a) $\alpha_0 = 0.0027, p = 0.05$ and $\epsilon = 0$.

(b) $\alpha_0 = 0.01, p = 0.1$ and $\epsilon = 0$.

Table 3.10: Corrections and exceedance probabilities of the proposed correction and the approximated tolerance factors from Krishnamoorthy and Mathew (2009) for different parameter combinations, different values of m , and $n = 1$ (individuals).

3.6 Extension of Tolerance Interval Theory

Although tolerance intervals are intended for use on individual observations, their theory can be extended to also be applicable to the \bar{X} chart. From equation (3.1) we know that the probability of not obtaining a signal when the process is in-control can be written as

$$\Phi\left(\frac{\hat{\mu}_0 - \mu_0}{\sigma_0/\sqrt{n}} + \tilde{K} \frac{\hat{\sigma}_0}{\sigma_0}\right) - \Phi\left(\frac{\hat{\mu}_0 - \mu_0}{\sigma_0/\sqrt{n}} - \tilde{K} \frac{\hat{\sigma}_0}{\sigma_0}\right) \quad (3.31)$$

Consider a general unbiased estimator for location that follows, either exactly or approximately, a normal distribution when the data are normally distributed, such as the grand sample average or the grand sample median. In that case, we have $\hat{\mu}_0 \sim N(\mu_0, \sigma_{\hat{\mu}_0})$. This means that $\tilde{Z} = \frac{\hat{\mu}_0 - \mu_0}{\sigma_0/\sqrt{n}} \sim N(0, \frac{\sigma_{\hat{\mu}_0}}{\sigma_0/\sqrt{n}})$. Also, consider an estimator $\hat{\sigma}_0$ such that $W = \frac{\hat{\sigma}_0}{\sigma_0} \sim \frac{a\chi_b}{\sqrt{b}}$, either exactly or approximately, with a and b some constants whose values depend on the estimator $\hat{\sigma}_0$. Then we can rewrite equation (3.31) as

$$\Phi(\tilde{Z} + \tilde{K}W) - \Phi(\tilde{Z} - \tilde{K}W). \quad (3.32)$$

The goal is to determine the value \tilde{K} that provides an in-control conditional CFAR of at most $\alpha_{tol} = (1 + \epsilon)\alpha_0$ with probability $1 - p$, which is in turn equivalent to having a probability of no signal of at least $1 - \alpha_{tol}$ with probability $1 - p$. Mathematically, this can be written as

$$P_{\tilde{Z}, W}(\Phi(\tilde{Z} + \tilde{K}W) - \Phi(\tilde{Z} - \tilde{K}W) \geq 1 - \alpha_{tol}) = 1 - p. \quad (3.33)$$

Result 1.2.1 in Krishnamoorthy and Mathew (2009) states that for some $\tilde{Z} \sim N(0, \sigma_{\tilde{Z}})$ independently of $Q \sim \frac{\chi_\nu^2}{\nu}$, where χ_ν^2 denotes a chi-square random variable with degrees of freedom ν , an approximate solution for the value k_2 such that

$$P_{\tilde{Z},Q} \left(\Phi(\tilde{Z} + k_2\sqrt{Q}) - \Phi(\tilde{Z} - k_2\sqrt{Q}) \geq 1 - \alpha_{tol} \right) = 1 - p \quad (3.34)$$

is given by

$$k_2 = \left(\frac{\nu\chi_{1;1-\alpha_{tol}}^2(\sigma_{\tilde{Z}})}{\chi_{\nu;p}^2} \right)^{1/2} \quad (3.35)$$

where $\chi_{d;q}^2(\theta)$ again denotes the q -quantile of a noncentral chi-square distribution with d degrees of freedom and noncentrality parameter θ , and where $\chi_{d;q}^2$ again denotes the q -quantile of a chi-square distribution with d degrees of freedom. To this end, we can write (3.33) as

$$P_{\tilde{Z},W} \left(\Phi \left(\tilde{Z} + \tilde{K}a\frac{W}{a} \right) - \Phi \left(\tilde{Z} - \tilde{K}a\frac{W}{a} \right) \geq 1 - \alpha_{tol} \right) = 1 - p \quad (3.36)$$

Next, since $\frac{W}{a} \sim \frac{\chi_b}{\sqrt{b}}$, we use the result derived by Krishnamoorthy and Mathew (2009) to find the desired value \tilde{K} , namely

$$\tilde{K} = \left(\frac{b\chi_{1;1-\alpha_{tol}}^2(\sigma_{\tilde{Z}}^2)}{a^2\chi_{b;p}^2} \right)^{1/2} \quad (3.37)$$

where $\sigma_{\tilde{Z}} = \frac{\sigma_{\mu_0}}{\sigma_0/\sqrt{n}}$.

As an example, consider $\hat{\mu}_0 = \bar{X}$ and $\hat{\sigma}_0 = S_{pooled}$. This means that we have $\sigma_{\tilde{Z}} = \frac{\sigma_{\bar{\mu}}}{\sigma_0/\sqrt{n}} = \frac{\sigma_0/\sqrt{mn}}{\sigma_0/\sqrt{n}} = 1/\sqrt{m}$, $a = 1$, and $b = m(n-1)$. We then find the desired control charting constant to be

$$\tilde{K}_{\bar{X},S_{pooled}} = \left(\frac{m(n-1)\chi_{1;1-\alpha_{tol}}^2(1/m)}{\chi_{m(n-1);p}^2} \right)^{1/2}. \quad (3.38)$$

Note that the required quantiles, and consequently this formula, can easily be calculated in common statistical software programs such as R or Matlab.

Note also that similar expressions for \tilde{K} can easily be obtained for different estimators. The only restrictions are that the estimator for location (approximately) follows a normal distribution and that the estimator for dispersion (approximately) follows a scaled chi distribution, and that these estimators are independent. This holds for many other commonly used estimators, such as the grand sample median for location, and the average sample standard deviation, the average sample range and the grand sample standard deviation for dispersion.

Control Limit Performance

In Tables 3.11a and 3.11b we illustrate the adjusted control limit factors and their resulting exceedance probability (the fraction of the control charts with a conditional in-control ARL below the pre-specified threshold ARL_0) for two different sets of parameters. All exceedance probabilities are determined based on 100,000 simulated Phase I samples. Note that the values should by design be close to p . As can be observed from the tables, both approximations provide good results. We obtain similar performances for other parameter values and common standard deviation estimators.

m	\tilde{K}_G	EP_G	\tilde{K}_{KM}	EP_{KM}
25	3.3827	0.0969	3.3687	0.1060
50	3.2473	0.0966	3.2399	0.1042
100	3.1640	0.0962	3.1595	0.1030
150	3.1291	0.0985	3.1266	0.1030
200	3.1094	0.0974	3.1077	0.1009
300	3.0870	0.0986	3.0862	0.1006
500	3.0657	0.0995	3.0654	0.1007
1000	3.0454	0.0993	3.0453	0.1000

(a) $\alpha_{tol} = 0.0027, p = 0.1, n = 5$

m	\tilde{K}_G	EP_G	\tilde{K}_{KM}	EP_{KM}
25	2.9665	0.0607	2.9743	0.0568
50	2.8315	0.0575	2.8357	0.0547
100	2.7478	0.0547	2.7492	0.0535
150	2.7124	0.0539	2.7137	0.0526
200	2.6922	0.0543	2.6933	0.0525
300	2.6691	0.0537	2.6700	0.0519
500	2.6468	0.0505	2.6474	0.0489
1000	2.6251	0.0509	2.6255	0.0497

(b) $\alpha_{tol} = 0.01, p = 0.05, n = 5$

Table 3.11: Adjusted control limit factors \tilde{K} and exceedance probabilities (EP) for the correction terms of Section 3.3 (subscript G) and this section (subscript KM) for different parameter combinations and various values of m indicated in the table.

3.7 Concluding Remarks

In this chapter we derived new correction terms for the Shewhart X and \bar{X} control charts in order to account for the effect of parameter estimation. The newly proposed correction terms are in line with the idea introduced by Albers and Kallenberg (2004a), which corrects the control limits to guarantee a specified minimum performance of the control chart with a specified probability. The correction terms are shown to be very accurate in achieving this. The performance of the proposed method is much better than Albers and Kallenberg (2005), and similar to the bootstrap method of Gandy and Kvaløy (2013). However, no bootstrapping is required, as the proposed correction only depends on the initial Phase I sample through its size, rather than its parameter estimates. Moreover, we made a comparison with tolerance intervals and self-starting control charts. The conclusions are that our corrections behave very well for the individuals Shewhart X control chart and outperform the self-starting control chart in both in-control and out-of-control situations.

Because of the guarantee of minimum performance, the corrected chart performs better than an uncorrected chart in the in-control situation. This inevitably leads to a deterioration of the out-of-control performance. However, the strictness of the correction can be easily adapted by changing α_0 , ϵ and p as desired. The choice of parameters should be based on the context. As costs of a false alarm and costs of running a process out-of-control are very dependent on the application, it is recommended to take this into account when making the tradeoff between in-control and out-of-control performance.

3.8 Appendix A

First note that $P_{mn}(\tilde{K}; Z, W)$ can be written according to (3.2). This can be rewritten as

$$P_{mn}(\tilde{K}; \hat{\mu}_0, \hat{\sigma}_0) = \bar{\Phi}(\tilde{K} + \Delta_1(\tilde{K})) + \bar{\Phi}(\tilde{K} + \Delta_2(\tilde{K})) \quad (3.39)$$

with $\bar{\Phi}(x) = 1 - \Phi(x)$ and

$$\Delta_1(\tilde{K}) = \frac{\hat{\mu}_0 - \mu_0}{\sigma_0/\sqrt{n}} + \tilde{K} \left(\frac{\hat{\sigma}_0}{\sigma_0} - 1 \right) \quad (3.40)$$

and

$$\Delta_2(\tilde{K}) = -\frac{\hat{\mu}_0 - \mu_0}{\sigma_0/\sqrt{n}} + \tilde{K} \left(\frac{\hat{\sigma}_0}{\sigma_0} - 1 \right) \quad (3.41)$$

Hence, for any function $g(\alpha_0)$ we can write

$$g(P_{mn}(\tilde{K}; \hat{\mu}_0, \hat{\sigma}_0)) = h(\tilde{K} + \Delta_1(\tilde{K}), \tilde{K} + \Delta_2(\tilde{K})) = h(x, y) \quad (3.42)$$

Using a two-step Taylor expansion, and abbreviating $P_{mn}(\tilde{K}; Z, W)$ as P_{mn} , this is approximately equal to

$$\begin{aligned} g(P_{mn}) &= h(\tilde{K} + \Delta_1(\tilde{K}), \tilde{K} + \Delta_2(\tilde{K})) \\ &\approx h(\tilde{K}, \tilde{K}) + h_x(\tilde{K}, \tilde{K})\Delta_1(\tilde{K}) + h_y(\tilde{K}, \tilde{K})\Delta_2(\tilde{K}) \\ &\quad + \frac{1}{2}[h_{xx}(\tilde{K}, \tilde{K})\Delta_1^2(\tilde{K}) + 2h_{xy}(\tilde{K}, \tilde{K})\Delta_1(\tilde{K})\Delta_2(\tilde{K}) + h_{yy}(\tilde{K}, \tilde{K})\Delta_2^2(\tilde{K})] \end{aligned} \quad (3.43)$$

where $h_x(x, y)$ and $h_y(x, y)$ are the first order partial derivatives of $h(x, y)$ with respect to x and y respectively, $h_{xx}(x, y)$ and $h_{yy}(x, y)$ are the second order partial derivatives of $h(x, y)$ with respect to x and y respectively, and $h_{xy}(x, y)$

equals the cross partial derivative of $h(x, y)$ with respect to x and y . Note that $h_x(\tilde{K}, \tilde{K}) = h_y(\tilde{K}, \tilde{K})$ and $h_{xx}(\tilde{K}, \tilde{K}) = h_{yy}(\tilde{K}, \tilde{K})$. Taking this into account, we can simplify (3.43) into

$$\begin{aligned} g(P_{mn}) &\approx h(\tilde{K}, \tilde{K}) + h_x(\tilde{K}, \tilde{K}) \left[\Delta_1(\tilde{K}) + \Delta_2(\tilde{K}) \right] \\ &+ \frac{1}{2} h_{xx}(\tilde{K}, \tilde{K}) \left[\Delta_1^2(\tilde{K}) + \Delta_2^2(\tilde{K}) \right] + h_{xy}(\tilde{K}, \tilde{K}) \Delta_1(\tilde{K}) \Delta_2(\tilde{K}) \end{aligned} \quad (3.44)$$

Using (3.40) and (3.41) we can rewrite this into

$$\begin{aligned} g(P_{mn}) &\approx h(\tilde{K}, \tilde{K}) + h_x(\tilde{K}, \tilde{K}) 2\tilde{K} \left(\frac{\hat{\sigma}_0}{\sigma_0} - 1 \right) + (h_{xx}(\tilde{K}, \tilde{K}) - h_{xy}(\tilde{K}, \tilde{K})) \left(\frac{\hat{\mu}_0 - \mu_0}{\sigma_0/\sqrt{n}} \right)^2 \\ &+ (h_{xx}(\tilde{K}, \tilde{K}) + h_{xy}(\tilde{K}, \tilde{K})) \tilde{K}^2 \left(\frac{\hat{\sigma}_0}{\sigma_0} - 1 \right)^2 \end{aligned} \quad (3.45)$$

For X_{ij} i.i.d. $N(\mu_0, \sigma_0)$ -distributed random variables and $\hat{\mu}_0 = \bar{X}$, we know that $\left(\frac{\hat{\mu}_0 - \mu_0}{\sigma_0/\sqrt{mn}} \right)^2$ follows a chi-square distribution. Common estimators of the standard deviation, such as the (pooled) sample standard deviation, follow a scaled chi distribution, while even the distribution of the average moving range can be approximated as such (see e.g. Roes et al., 1993). This means that $\hat{\sigma}_0/\sigma_0$ and $\hat{\sigma}_0^2/\sigma_0^2$ generally follow a scaled chi and scaled chi-square distribution respectively. Hence, we may conclude that $P_{mn}(\tilde{K}; Z, W)$ is approximately a combination of scaled chi and chi-square distributed random variables. Note that the distribution of $P_{mn}(\tilde{K}; Z, W)$ should be approximated such that it not only gives an accurate description, but also such that it is possible to obtain the required correction term. The difficulty in obtaining this correction term lies in the sense that the correction term c not only changes the expectation, but also the variance of the distribution. As the currently obtained approximation is still of rather complicated form, and because the scaled chi-square part appears to be dominant, we approximate the distribution of $P_{mn}(\tilde{K}; Z, W)$ by a $A\chi_B^2/B$ distribution, where we use the first two central moments of $P_{mn}(\tilde{K}; Z, W)$ to identify A and B .

3.9 Appendix B

In order to obtain the required correction term we use the Wilson-Hilferty transformation (Wilson and Hilferty, 1931), which states that for $X \sim \chi_B^2$ we have $\sqrt[3]{\frac{X}{B}} \overset{\text{approx}}{\sim} N(1 - 2/(9B), \sqrt{2/(9B)})$. This transformation is quite accurate, which was shown recently by Inglot (2010). Henceforth we abbreviate $P_{mn}(\tilde{K}; Z, W)$ as P_{mn} . Then in our case, when $P_{mn} \overset{\text{approx}}{\sim} \frac{A\chi_B^2}{B}$ or similarly $\frac{B}{A}P_{mn} \overset{\text{approx}}{\sim} \chi_B^2$, we obtain

$$\sqrt[3]{\frac{P_{mn}}{A}} \overset{\text{approx}}{\sim} N(1 - 2/(9B), \sqrt{2/(9B)}).$$

This is equivalent to

$$\frac{\sqrt[3]{\frac{P_{mn}}{A}} - 1 + \frac{2}{9B}}{\sqrt{\frac{2}{9B}}} \overset{\text{approx}}{\sim} N(0, 1).$$

We want to have $P(P_{mn} < (1 + \epsilon)\alpha_0) = 1 - p$ (cf. (3.5)). This is equivalent to

$$\begin{aligned} P(P_{mn} < (1 + \epsilon)\alpha_0) &= P\left(\frac{\sqrt[3]{\frac{P_{mn}}{A}} - 1 + \frac{2}{9B}}{\sqrt{\frac{2}{9B}}} < \frac{\sqrt[3]{\frac{(1+\epsilon)\alpha_0}{A}} - 1 + \frac{2}{9B}}{\sqrt{\frac{2}{9B}}}\right) \\ &\approx \Phi\left(\frac{\sqrt[3]{\frac{(1+\epsilon)\alpha_0}{A}} - 1 + \frac{2}{9B}}{\sqrt{\frac{2}{9B}}}\right) = 1 - p \end{aligned} \quad (3.46)$$

which in turn, by using the inverse of the standard normal cdf (denoted Φ^{-1}), leads to the following equation that needs to be solved

$$\frac{\sqrt[3]{\frac{(1+\epsilon)\alpha_0}{A}} - 1 + \frac{2}{9B}}{\sqrt{\frac{2}{9B}}} = \Phi^{-1}(1 - p) \quad (3.47)$$

Note again that both A and B are functions of \tilde{K} . Given the values of m, n, α_0 and ϵ , the left-hand side of (3.47) is a function of \tilde{K} only, say $Y(\tilde{K})$. Using (3.10) and (3.11) we can write $Y(\tilde{K})$ as

$$Y(\tilde{K}) = \sqrt[3]{(1 + \epsilon)\alpha_0} \frac{3E(P_{mn})^{2/3}}{\sqrt{\text{Var}(P_{mn})}} - \frac{3E(P_{mn})}{\sqrt{\text{Var}(P_{mn})}} + \frac{\sqrt{\text{Var}(P_{mn})}}{3E(P_{mn})} \quad (3.48)$$

In order to solve (3.47) we need to find c such that for $\tilde{K} = K + c$ there holds $Y(\tilde{K}) = \Phi^{-1}(1 - p)$. This value of c is found by a linear approximation of $Y(\tilde{K})$ as $Y(\tilde{K}) \approx Y(K) + c \frac{dY(K)}{d\tilde{K}}$.

If we denote the derivatives of $E(P_{mn})$ and $V(P_{mn})$ with respect to \tilde{K} , evaluated at K , by $\frac{dE}{d\tilde{K}}$ and $\frac{dV}{d\tilde{K}}$ respectively, and $E(P_{mn})$ and $V(P_{mn})$, evaluated at K , by E and V respectively, we have

$$Y'(K) = \frac{dY(K)}{d\tilde{K}} = \frac{2E^{-1/3}\sqrt{V}\frac{dE}{d\tilde{K}} - \frac{3E^{2/3}}{2\sqrt{V}}\frac{dV}{d\tilde{K}}}{V} - \frac{3\frac{dE}{d\tilde{K}}\sqrt{V} - \frac{3E}{2\sqrt{V}}\frac{dV}{d\tilde{K}}}{V} + \frac{\frac{3E}{2\sqrt{V}}\frac{dV}{d\tilde{K}} - 3\frac{dE}{d\tilde{K}}\sqrt{V}}{9E^2} \quad (3.49)$$

The values of $\frac{dE}{d\tilde{K}}$ and $\frac{dV}{d\tilde{K}}$ can be calculated (as can be seen from (3.7), (3.9) and (3.8)) as

$$\begin{aligned} \frac{dE}{d\tilde{K}} &= \int_{-\infty}^{\infty} \int_0^{\infty} \frac{dP_{mn}(z, w)}{d\tilde{K}} f(z)f(w)dwdz \\ &= \int_{-\infty}^{\infty} \int_0^{\infty} -w \left[\phi\left(\frac{z}{\sqrt{m}} + Kw\right) + \phi\left(\frac{z}{\sqrt{m}} - Kw\right) \right] f(z)f(w)dwdz \end{aligned} \quad (3.50)$$

and

$$\frac{dV}{d\tilde{K}} = \frac{d(E^2)}{d\tilde{K}} - 2E\frac{dE}{d\tilde{K}} \quad (3.51)$$

where

$$\begin{aligned} \frac{d(E^2)}{dK} &= \int_{-\infty}^{\infty} \int_0^{\infty} \frac{dP_{mn}^2(K; z, w)}{d\tilde{K}} f(z)f(w)dwdz \\ &= -2 \int_{-\infty}^{\infty} \int_0^{\infty} wP_{mn}(K; z, w) \left[\phi\left(\frac{z}{\sqrt{m}} + Kw\right) + \phi\left(\frac{z}{\sqrt{m}} - Kw\right) \right] f(z)f(w)dwdz \end{aligned} \quad (3.52)$$

Getting back to the approximation $Y(\tilde{K}) \approx Y(K) + c \frac{dY(K)}{d\tilde{K}} = Y(K) + cY'(K)$, we thus obtain

$$c = \frac{\Phi^{-1}(1-p) - Y(K)}{Y'(K)} \quad (3.53)$$

4. Control Charts for Dispersion

In this chapter, we derive adjusted control limits for control charts for dispersion, using the exceedance probability criterion. The adjustments are given in a quite general form, so that they can be used with a wide range of estimators for dispersion. This chapter is based on Goedhart et al. (2017a).

4.1 Introduction

In the previous chapters we have focused on the Shewhart X and \bar{X} control charts. However, control charts for monitoring dispersion suffer from similar estimation issues as the previously discussed control charts. Epprecht et al. (2015), for the case of the S chart, evaluated the effects of parameter estimation on the in-control conditional run length distribution. Because the conditional run length distribution is geometric, it can be characterized by the probability of success, the conditional false alarm rate (CFAR), or its reciprocal, the in-control conditional average run length (CARL). Since the CFAR (CARL) is a random variable, Epprecht et al. (2015) proposed a prediction bound formulation to determine the number of Phase I subgroups required such that a minimum in-control performance based on the CFAR or CARL of the one-sided S chart is guaranteed with a pre-specified probability. Note that this is equivalent to the exceedance probability criterion. However, it was seen that according to this approach, the number of subgroups required to guarantee practically attractive in-control chart performance is often in the order of several thousands, depending on the choice of the parameters. These numbers are substantially higher than the values recommended by based on the unconditional in-control ARL. These findings are similar to the findings for the

X and \bar{X} control charts in Saleh et al. (2015b).

To this end, we derive adjusted control limits for control charts for dispersion in this chapter. Since an increase in the process dispersion is deemed more important to detect than a decrease (as it indicates process degradation), we consider the one-sided $\hat{\sigma}$ chart with an UCL only (where $\hat{\sigma}$ is the sample charting statistic of dispersion), equivalent to Epprecht et al. (2015), and derive the corresponding adjusted upper control limits. These adjusted limits guarantee that the CFAR exceeds a pre-specified tolerated bound with only a small probability, for a given number of Phase I samples of a given size. This adjusted control limit coefficient is obtained analytically, as opposed to using the bootstrap approach as in Faraz et al. (2015), and its efficacy depends also on the type of estimator used to estimate the in-control process dispersion from the given Phase I data. Our formulation and derivations allow the use of different estimators for the process dispersion. Note further that the adjusted control limit can also be obtained for control charts based on monotone increasing functions of the process dispersion $\hat{\sigma}$, such as $\hat{\sigma}^2$ or $\log(\hat{\sigma})$, by taking the corresponding monotone increasing function of the adjusted limit for the $\hat{\sigma}$ chart. Furthermore, our framework is also applicable for the range charts. See also the tutorial on estimating the standard deviation written by Vardeman (1999).

In this chapter, we provide an analytical solution that enables a straightforward calculation of the adjusted control limit for any combination of parameters, which include the number of available Phase I subgroups m , subgroup size n , the desired performance threshold (e.g. α_{tol}) and a specified probability p of this threshold being exceeded. We give tables of the adjustment coefficients for a wide range of these parameter values. Also, we provide a formula for computing the out-of-control CARL of the resulting chart for any increase in the process dispersion. This is compared with the out-of-control CARL for the chart with the unadjusted limits in order to assess the impact of the adjustment on the out-of-control performance. A practitioner has to balance two things: controlling the in-control performance versus the deterioration in the out-of-control situation. We provide some guidelines on how a practitioner can balance this tradeoff. Our derivations are based on Shewhart type control charts for dispersion.

This chapter is structured as follows. We present the analytical derivations of the adjusted limits in Section 4.2 and of the resulting chart power (the

probability that the chart gives a true signal in case of an increase in process dispersion) in Section 4.3. Then in Section 4.4, the adjusted limits as well as the resulting power and the out-of-control ARL of the resulting chart are tabulated for a comprehensive spectrum of cases and the results are discussed. In Section 4.5 we present and discuss a practical example as illustration. Finally, the general conclusions are summarized in Section 4.6.

4.2 Determination of the Adjusted Control Limit

Assume that in Phase I a reference sample from a $N(\mu_0, \sigma_0)$ distribution is available, and that Phase II observations come from a $N(\mu_0, \sigma)$ distribution, where parameters are assumed unknown. In Phase II, a monitoring statistic $\hat{\sigma}_i$ is obtained each time period i . Note that σ may differ from the in-control value σ_0 . Our methodology works for all monitoring statistics $\hat{\sigma}_i$ that have a distribution either exactly or approximately proportional to a chi distribution (where the approximation is based on Patnaik, 1950). For example, we can take $\hat{\sigma}_i$ (i.e., the i -th Phase II sample standard deviation) and since it is well-known that $(n-1)S_i^2/\sigma_i^2 \sim \chi_{n-1}^2$ exactly, it follows that $S_i \sim \sigma\chi_{n-1}/\sqrt{n-1}$ exactly, under the normal distribution. On the other hand, if we consider $\hat{\sigma}_i = R_i/d_2(n)$ then we approximate the distribution of $\hat{\sigma}_i/\sigma$ by $a\chi_b/\sqrt{b}$ based on the Patnaik (1950) approach. We elaborate on this point later, in Section 4.2.4.

4.2.1 Unadjusted Control Limit

Consider a monitoring statistic $\hat{\sigma}_i$ such that $\hat{\sigma}_i/\sigma \sim a\chi_b/\sqrt{b}$ for suitable values $a > 0$ and $b > 0$. Obviously, the UCL for the \bar{X} and \bar{X} control chart as in (2.1) will not work in this case, as it is designed for a normally distributed plotting statistic. Therefore, probability limits are generally used in control charts for dispersion (see e.g. Montgomery, 2013). We discuss the traditional (unadjusted) α_0 -probability limits for Phase II control charts for dispersion in this section.

Given a desired nominal false alarm rate α_0 , the traditional estimated upper control limit of the chart is set at

$$\widehat{UCL} = L\hat{\sigma}_0 \tag{4.1}$$

with

$$L = \sqrt{a^2\chi_{b;1-\alpha_0}^2/b} \tag{4.2}$$

where $\chi_{v;p}^2$ denotes the 100p-percentile of a chi-square distribution with v degrees of freedom and $\hat{\sigma}_0$ is the estimator of σ_0 . Note that, for the specific case that $\hat{\sigma}$ is equal to the sample standard deviation S we know that $a = 1$ and $b = n - 1$, so that Equation (4.2) corresponds to Montgomery (2013), page 267. This standard control limit does not account for either parameter estimation or practitioner-to-practitioner variability and is referred to as the unadjusted control limit in this chapter.

Note that, as mentioned before, one can consider various estimators for $\hat{\sigma}_0$ that have a distribution proportional to the chi distribution. In order to find an expression for the probability of signal, define the standard deviation ratio as

$$\gamma = \sigma/\sigma_0 \tag{4.3}$$

where σ is the current (Phase II) process standard deviation. Note that when the process is in control, $\sigma = \sigma_0$, so that $\gamma = 1$. When special causes result in an increase of the process standard deviation, $\sigma > \sigma_0$ and consequently $\gamma > 1$. Similarly, with a reduction in the process standard deviation we have $\gamma < 1$. However, we don't consider this latter case here, because the meaning and usefulness of detecting decreases in the process dispersion are totally different. Of course, one may consider a lower control limit using the same methodology as in our approach to study this behavior. Next, define the error factor of the estimate $\hat{\sigma}_0$ as the ratio:

$$W = \hat{\sigma}_0/\sigma_0. \tag{4.4}$$

Since it is assumed in this chapter that the data come from a normal distribution, it is easy to see (Epprecht et al., 2015) that in Phase II, the conditional probability of an alarm (denoted CPA in this chapter) of the upper one-sided

$\hat{\sigma}$ chart is given by

$$\begin{aligned} CPA(\gamma, L) &= P(\hat{\sigma}_i > \widehat{UCL}) = P\left(\frac{b\hat{\sigma}_i^2}{a^2\sigma^2} > \frac{W^2 bL^2}{\gamma^2 a^2}\right) \\ &= P\left(\frac{b\hat{\sigma}_i^2}{\sigma^2} > \frac{W^2}{\gamma^2} \chi_{b;1-\alpha_0}^2\right) = 1 - F_{\chi_b^2}\left(\frac{W^2}{\gamma^2} \chi_{b;1-\alpha_0}^2\right) \end{aligned} \quad (4.5)$$

since $bL^2 = \chi_{b;1-\alpha_0}^2$ from (4.2) and where $F_{\chi_b^2}$ denotes the cdf of the chi-square distribution with b degrees of freedom. Note that expression (4.5) holds for both the in-control case ($\gamma = 1$), in which case it corresponds to the conditional false alarm rate (denoted CFAR), and in the out-of-control case ($\gamma > 1$), where it represents the conditional probability of an alarm (CPA) of the chart, both in Phase II. Thus,

$$CFAR = CPA(1, L) = 1 - F_{\chi_b^2}\left(W^2 \chi_{b;1-\alpha_0}^2\right) \quad (4.6)$$

Note that the CPA of the $\hat{\sigma}$ chart in Equation (4.5) is also the CPA of the $\hat{\sigma}^2$ chart with \widehat{UCL} equal to the square of the $\hat{\sigma}$ chart's \widehat{UCL} in (4.1). These two charts are equivalent: one will signal if and only if the other will signal. In this chapter we use the $\hat{\sigma}$ chart for illustration, but all the analyses, numerical results and conclusions apply to the $\hat{\sigma}^2$ chart as well. In fact, these observations hold for any monotone increasing function of $\hat{\sigma}$ such as $\log(\hat{\sigma})$ which is sometimes used in practice. In summary, for any monotone increasing function $g(\hat{\sigma})$ of $\hat{\sigma}$, $P(\hat{\sigma} > \widehat{UCL})$ is equivalent to $P(g(\hat{\sigma}) > g(\widehat{UCL}))$. Therefore, the adjusted limits proposed for the $\hat{\sigma}$ chart can be applied to any monotone increasing function $g(\hat{\sigma})$ of $\hat{\sigma}$ by applying the same transformation $g(\widehat{UCL})$ to the \widehat{UCL} obtained for the $\hat{\sigma}$ chart.

The CFAR shown in Equation (4.6) is a function of the error factor of the estimate W and, as a result, is also a random variable. Thus, the value of the CFAR will be different for different values of W , corresponding to different values of the estimator, from different Phase I samples obtained by practitioners. This is true unless $W = 1$, which is the case for consistent estimators when the number of reference samples tends to infinity, which implies the known parameter case. The finite sample distribution of W depends on

the distribution of the estimator $\hat{\sigma}_0$ used for σ_0 . In the case of normally distributed data, the most common estimators of the process standard deviation available in the literature follow, either exactly or approximately, a scaled chi distribution. Thus, in a general formulation, we consider estimators $\hat{\sigma}$ such that $\hat{\sigma}/\sigma \sim a\chi_b/\sqrt{b}$ either exactly or approximately, for suitable a and b . To distinguish between Phase I and Phase II estimators, we use a_0 and b_0 for Phase I and a and b for Phase II. This formulation allows a more general and comprehensive treatment of Phase II monitoring of σ , covering most common estimators of standard deviation used (i) in the Phase I control limit and (ii) as a plotting statistic in Phase II.

For example, a commonly used estimator of σ_0 in the Phase I control limit is the square root of pooled variances (also recommended by Mahmoud et al., 2010), as given in (2.3). Since it is well-known that $m(n-1)S_{pooled}^2/\sigma_0^2$ follows a chi-square distribution with $m(n-1)$ degrees of freedom, it follows that $W = S_{pooled}/\sigma_0 \sim a_0\chi_{b_0}/\sqrt{b_0}$ exactly, where $a_0 = 1$ and $b_0 = m(n-1)$. The plotting statistic for the Phase II S chart is the sample standard deviation S_i , and since it is well-known that $((n-1)S_i^2)/\sigma^2 \sim \chi_{n-1}^2$, it follows that $S_i/\sigma \sim a\chi_b/\sqrt{b}$ exactly where $a = 1$ and $b = n-1$. Other Phase I estimators of σ_0 such as \bar{S}/c_4 , where $\bar{S} = \sum S_i/m$ and c_4 is an unbiasing constant (see Montgomery, 2013) can be considered, but we don't pursue this here and use the estimator S_{pooled} for illustration throughout. Note further that other monitoring statistics and charts, such as the R chart, can also be considered under this framework. We make some comments about these points in Section 4.2.3.

As in Epprecht et al. (2015), the cdf of the CFAR (denoted as F_{CFAR}) of the upper one-sided $\hat{\sigma}$ chart with the traditional \widehat{UCL} as in Equations (4.1) and (4.2), set for a nominal false alarm rate α_0 , can be shown to be equal to:

$$F_{CFAR}(t; L) = P\left(1 - F_{\chi_b^2}\left(W^2\chi_{b;1-\alpha_0}^2\right) \leq t\right) = P\left(W^2 > \frac{\chi_{b;1-t}^2}{\chi_{b;1-\alpha_0}^2}\right) \quad (4.7)$$

for $0 < t < 1$. Epprecht et al. (2015) showed that this CFAR has a non-negligible probability of being much larger than the specified nominal value α_0 , unless the number of Phase I samples is prohibitively high. Motivated by this with a practical point of view, we derive the adjusted control limit based on the exceedance probability criterion, for given values of m and n .

4.2.2 Adjusted Control Limits

We are interested in finding an adjustment to the traditional UCL in Equation (4.1) such that the probability that CFAR exceeds a tolerated upper bound (e.g. α_{tol}) is controlled at a small value p . Formally, we want to determine an adjusted coefficient L^* to be used to calculate the Phase I upper control limit $\widehat{UCL}^* = L^* \hat{\sigma}_0$ such that

$$1 - F_{CFAR}(\alpha_{tol}; L^*) = p \quad (4.8)$$

where $\alpha_{tol} = (1 + \epsilon)\alpha_0$ as before. Note that Equation (4.8) can be rewritten as $P(CFAR(t, L^*) \leq \alpha_{tol}) = F_{CFAR}(\alpha_{tol}; L^*) = 1 - p$, so that the interval $[0, \alpha_{tol}]$ can be interpreted as a $100(1 - p)\%$ prediction interval for the CFAR. Note also that when the process is in-control, Equation (4.8) is equivalent to writing $P(CARL < 1/\alpha_{tol}) = p$. Thus, solving (4.8) is equivalent to finding L^* that guarantees a minimum in-control CARL equal to $CARL_{tol} = 1/\alpha_{tol}$ with a specified probability $1 - p$. Because of this equivalence, our further derivations are based on CFAR. Note that the interval $[1/\alpha_{tol}, \infty)$ can be interpreted as a $100(1 - p)\%$ prediction interval for the in control CARL.

In order to determine L^* from (4.8) we need the cdf of CFAR. Similar to Equation (4.5) we can write the CFAR when using the adjusted limits as

$$CFAR(L^*) = CPA(1, L^*) = P\left(\frac{b\hat{\sigma}_i^2}{a^2\sigma_0^2} > W^2 \frac{bL^{*2}}{a^2}\right) = 1 - F_{\chi_b^2}\left(W^2 \frac{bL^{*2}}{a^2}\right) \quad (4.9)$$

Thus, similar to Equation (4.7), the cdf is obtained as

$$\begin{aligned} F_{CFAR}(t; L^*) &= P\left(1 - F_{\chi_b^2}\left(W^2 \frac{bL^{*2}}{a^2}\right) \leq t\right) = P\left(W^2 > \frac{a^2 \chi_{b;1-t}^2}{bL^{*2}}\right) \\ &= P\left(\frac{b_0}{a_0^2} W^2 > \frac{b_0}{a_0^2} \frac{a^2 \chi_{b;1-t}^2}{bL^{*2}}\right) = 1 - F_{\chi_{b_0}^2}\left(\frac{b_0}{a_0^2} \frac{a^2 \chi_{b;1-t}^2}{bL^{*2}}\right) \end{aligned} \quad (4.10)$$

for $0 < t < 1$, since we assume that $W \sim a_0 \chi_{b_0} / \sqrt{b_0}$ so that $b_0 W^2 / a_0^2 \sim \chi_{b_0}^2$.

Now, L^* is determined such that

$$F_{CFAR}(\alpha_{tol}; L^*) = 1 - p. \quad (4.11)$$

Hence, using Equation (4.10), we obtain

$$\chi_{b_0;p}^2 = \frac{b_0}{a_0^2} \frac{a^2 \chi_{b;1-\alpha_{tol}}^2}{bL^{*2}} \quad (4.12)$$

which leads to the general solution

$$L^* = \sqrt{\frac{b_0}{b} \frac{a^2}{a_0^2} \frac{\chi_{b;1-\alpha_{tol}}^2}{\chi_{b_0;p}^2}} \quad (4.13)$$

We emphasize that Equation (4.13) is a general expression for the adjusted control limit coefficient that can be applied with any Phase I estimator $\hat{\sigma}_0$ for which $W = \hat{\sigma}_0/\sigma_0 \sim a_0\chi_{b_0}/\sqrt{b_0}$ and for which the Phase II monitoring statistic satisfies $\hat{\sigma}/\sigma \sim a\chi_b/\sqrt{b}$. Note that, if one would be interested in the equivalent correction for the \widehat{LCL}^* , the only changes required are to substitute α_{tol} for $1 - \alpha_{tol}$ and $1 - p$ for p in equation (4.13).

4.2.3 Use of Different Estimators in Phase II

Until now we have considered general Phase I and Phase II estimators. However, it is possible to use a wide range of estimators in both Phase I and Phase II. Consider for example the use of the pooled standard deviation $\hat{\sigma}_0 = S_{pooled}$ in Phase I and the standard deviation $\hat{\sigma} = S$ in Phase II. We then have $a_0 = 1$, $b_0 = m(n - 1)$, $a = 1$ and $b = n - 1$. Implementing these values in Equation (4.13) then gives us the required control limit for this special case as

$$L^* = \sqrt{m \frac{\chi_{n-1;1-\alpha_{tol}}^2}{\chi_{m(n-1);p}^2}} \quad (4.14)$$

Note that this special case is equal to the result of Tietjen and Johnson (1979), who determine tolerance intervals for this specific example. However, our approach is more generally applicable as we allow a wide range of estimators

in both Phase I and Phase II. For more information on estimators of dispersion we refer to Vardeman (1999).

To illustrate the application and the consequences of implementing the adjusted control limits, Figures 4.1 and 4.2 show boxplots of 1,000 simulated CARL values for the in-control ($\gamma = 1$) and an out-of-control ($\gamma = 1.5$) situation, respectively, with $m = 50$, $n = 5$, $\alpha_0 = 0.005$, $\epsilon = 0.1$ and $p = 0.1$, so that $\alpha_{tol} = 0.0055$. We have used S_{pooled} as estimator of Phase I standard deviation, and S as estimator in Phase II. In the in-control situation (Figure 4.1), $CARL_{tol} = 1/\alpha_{tol} = 1/0.0055 = 182$. For comparison purposes, the boxplots also show the results of using the bootstrap method of Gandy and Kvaløy (2013) and Faraz et al. (2015), and of the chart with unadjusted limits. The minimum tolerated CARL of 182 is indicated with a vertical dashed line, and the p -quantile of the CARL distribution is indicated in each boxplot with an added short vertical line. As can be seen, the p -quantile coincides with $CARL_{tol} = 182$ for the adjusted control limits. Note that the difference between the proposed method and the bootstrap approach is negligible. However, the proposed control limits are analytical expressions that are easier to implement, and the required adjustments are found more directly through the statistical distribution theory. For additional insight into the effect of the choices of ϵ and p , Figures 4.3 and 4.4 illustrate, equivalent to Figures 4.1 and 4.2 respectively, this effect for $\epsilon = 0$ and $p = 0.05$. There the p -quantile coincides with $CARL_{tol} = 200$.

4.2.4 Distributions of Different Estimators of Standard Deviation

We have noted that in general we can consider any estimator $\hat{\sigma}_0$ such that $W = \hat{\sigma}_0/\sigma_0 \sim a_0\chi_{b_0}/\sqrt{b_0}$ either exactly or approximately. We make a few comments here in this direction. First, the estimator used in Phase I usually dictates the chart to be used in Phase II. So if $\hat{\sigma}_0$ is estimated by a function of the Phase I sample standard deviations, we use an S chart in Phase II for consistency, whereas if $\hat{\sigma}_0$ is estimated by a function of the Phase I sample ranges, we use an R chart in Phase II for monitoring the standard deviation. For some estimators such as S_{pooled} , as shown above, this distribution theory is exact but that is not the case for all estimators that have been proposed in the literature for the standard deviation of a normal distribution. However,

for many of the available estimators, their distribution can be approximated. One common approach to do this is by equating the first two moments of W with those of $a_0\chi_{b_0}/\sqrt{b_0}$ (see for example Patnaik, 1950). We consider such an approximation based on Patnaik (1950), as used in Roes et al. (1993). For an estimator $\hat{\sigma}_0$ such that $E[W] = 1$, so that $\hat{\sigma}_0$ is an unbiased estimator of σ_0 , Roes et al. (1993) showed that the required values of a_0 and b_0 must equal

$$\begin{aligned} a_0 &= \sqrt{V[W] + 1} \\ b_0 &= \frac{1}{2} \left(1 + \frac{1}{V[W]} \right) \end{aligned} \tag{4.15}$$

where $V[W]$ denotes the variance of W . Recognizing the fact that practitioners might use other estimators of the standard deviation, we summarize three popular estimators of the Phase I standard deviation σ_0 along with the corresponding values of a_0 and b_0 in Table 4.1. Of course, the same approach can be used for the Phase II estimators. Note that we have also indicated the corresponding Phase II plotting statistic and its corresponding values of a and b in Table 4.1. Note also that the constants $c_4(n)$, $d_2(n)$, and $d_3(n)$ indicated in this table can be found in Appendix VI in Montgomery (2013).

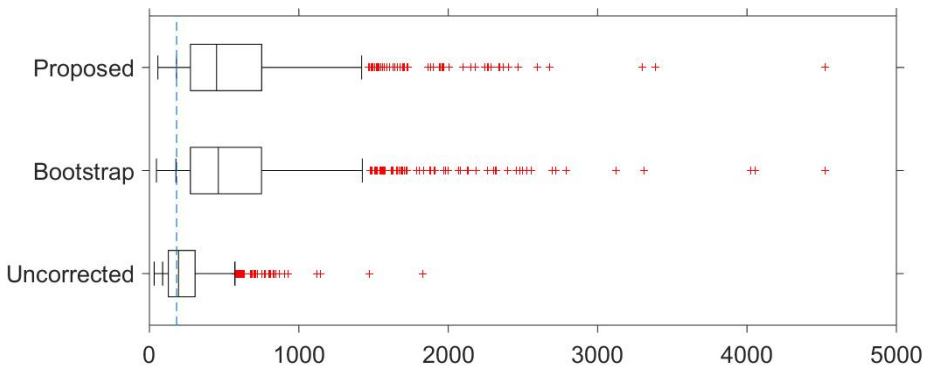


Figure 4.1: Boxplots of 1,000 simulated values of in-control ($\gamma = 1$) CARL. Parameter values are $m = 50$, $n = 5$, $\alpha_0 = 0.005$, $\epsilon = 0.1$ and $p = 0.1$. The dashed vertical line indicates $CARL_{tol} = 1/\alpha_{tol} = 182$, and the p -quantiles are indicated with an added vertical line in the boxplots.

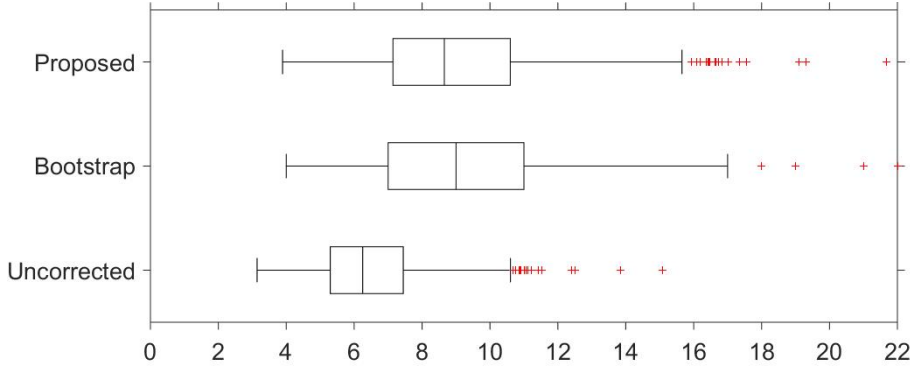


Figure 4.2: Boxplots of 1,000 simulated values of out-of-control ($\gamma = 1.5$) CARL. Parameter values are $m = 50$, $n = 5$, $\alpha_0 = 0.005$, $\epsilon = 0.1$ and $p = 0.1$.

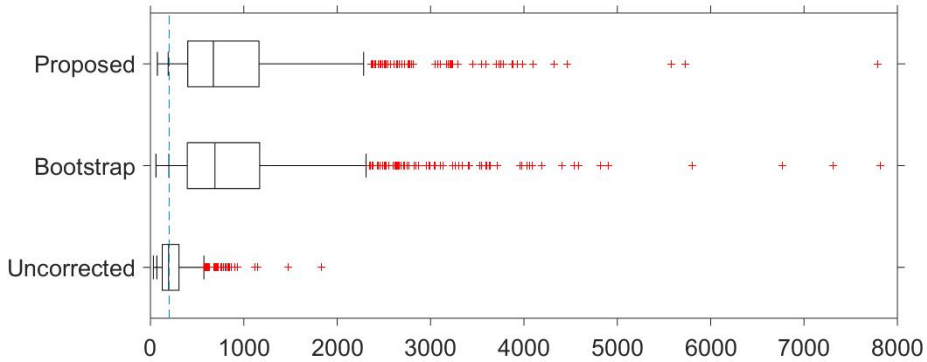


Figure 4.3: Boxplots of 1,000 simulated values of in-control ($\gamma = 1$) CARL. Parameter values are $m = 50$, $n = 5$, $\alpha_0 = 0.005$, $\epsilon = 0$ and $p = 0.05$. The dashed vertical line indicates $CARL_{tol} = 1/\alpha_{tol} = 200$, and the p -quantiles are indicated with an added vertical line in the boxplots.

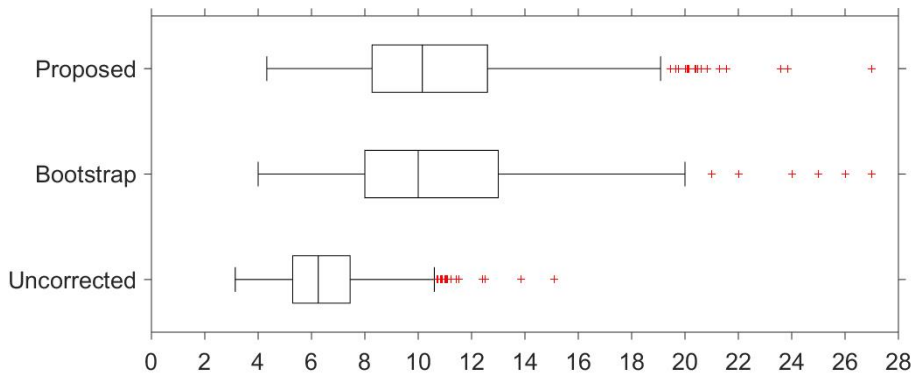


Figure 4.4: Boxplots of 1,000 simulated values of out-of-control ($\gamma = 1.5$) CARL. Parameter values are $m = 50$, $n = 5$, $\alpha_0 = 0.005$, $\epsilon = 0$ and $p = 0.05$.

Phase II plotting statistic	Values of a and b	Phase I estimator	Values of a_0 and b_0	$V[W]$
$\hat{\sigma}_i = \bar{S}_i$	$a = 1,$ $b = 1$	Pooled std. dev. $\hat{\sigma}_0 = S_{pooled}$	$a_0 = 1,$ $b_0 = m(n - 1)$	$1 - c_4^2(m(n - 1) + 1)$
$\hat{\sigma}_i = S_i$	$a = 1,$ $b = 1$	Average std. dev. $\hat{\sigma}_0 = \bar{S}/c_4(n)$	$a_0 = \sqrt{V[W] + 1},$ $b_0 = \frac{1}{2} \left(1 + \frac{1}{V[W]} \right)$	$\frac{1 - c_4^2(n)}{mc_4^2(n)}$
$\hat{\sigma}_i = R_i/d_2(n)$	$a = \sqrt{mV[W] + 1},$ $b = \frac{1}{2} \left(1 + \frac{1}{mV[W]} \right)$	Average range $\hat{\sigma}_0 = \bar{R}/d_2(n)$	$a_0 = \sqrt{V[W] + 1},$ $b_0 = \frac{1}{2} \left(1 + \frac{1}{V[W]} \right)$	$\frac{d_3^2(n)}{md_2^2(n)}$

Table 4.1: Required values of a and b (and a_0 and b_0) for different estimators used in the calculation of the adjusted control limit.

4.2.5 Deviations from the Proposed Model

The derivations as shown in this section are based on one-sided control charts for dispersion under normality. In the case that the normality assumption is violated, the adjustments are less accurate in terms of providing a specified in-control performance. However, a suitable alternative is to apply a Box-Cox transformation (see Box and Cox, 1964) or a Johnson type transformation (see Chou et al., 1998) to the data first, and determine the control limits afterwards using this transformed data. Of course, the Phase II data needs to be transformed in the same way during the monitoring stage. Another way to investigate deviations from normality is to use robust estimators, in order to deal with contaminations. However, this means an investigation as in Schoonhoven and Does (2012) is required, which may be the subject of a future study.

Whereas the problem with deviations from normality can be addressed in practice by applying a Box-Cox transformation, the proposed adjustments for the one-sided charts cannot be generalized directly to the two-sided control charts for dispersion. A common first guess when implementing the two-sided control chart is to calculate the upper and lower control limits using $\alpha_0/2$ instead of α_0 . However, in that case the guaranteed in-control performance results provided for the one-sided charts will no longer hold (see also Chapter 3), so the two-sided case need to be analyzed separately. However, note that in practice, it is most important to detect increases in process dispersion. By focusing on control charts with an upper control limit only, we provide control limits that are more sensitive to detecting these increases. Two-sided control charts would require a different approach, and their ability to detect increases would suffer from the addition of the lower control limit, while the benefits (detecting decreases in variation) are of minor importance in practice.

4.3 Out-of-Control Performance

In analogy with the power of a hypothesis test, the power of a control chart can be defined to be the probability that the chart gives a true signal. Since the conditional run length distribution of the chart is geometric, the reciprocal of the conditional probability of a true signal, the CPA (or the conditional power) is the conditional out-of-control ARL, denoted CARL. As with the CFAR, the power (and the reciprocal of this power, the out-of-control CARL) of the $\hat{\sigma}$ chart with the adjusted limit, depends on the realization (w) of the (unknown) error factor of the estimate, the random variable W . Note that the same is also true for the $\hat{\sigma}$ (or some other similar) chart with the unadjusted limit (see Equation 4.5). The power of the chart in the unknown parameter case will be larger (smaller) than the power for the known parameter case if $w < 1$, that is when σ_0 is under-estimated (or if $w > 1$, that is, when σ_0 is overestimated). The gain or loss in the out-of-control performance of the chart due to the use of the adjusted limit (relative to that of the chart using the unadjusted limit) will thus vary according to the realization of W . That is, the gain/loss of out-of-control performance is also a random variable. See Zwetsloot (2016) for a more detailed comparison of the effect of estimation error in control charts for dispersion.

In order to examine the implications of using the adjusted limits on the out-of-control performance (power) of the chart, we evaluate the chart performance for the case when the estimation error is zero ($w = 1$; the known parameter case), and for increases of 50% and 100% in the standard deviation σ ($\gamma = 1.5$ and $\gamma = 2$, respectively). This is because the conditional probability of an alarm depends only on the ratio between w and γ , and not on their absolute values (as can be seen from Equation (4.5)).

To calculate the out-of-control CARL, recall from Equation (4.5) that the conditional probability of an alarm (CPA) with the adjusted limit is given by

$$CPA(\gamma, L^*) = P\left(\frac{b\hat{\sigma}_i^2}{a^2\sigma^2} > \frac{W^2}{\gamma^2} \frac{bL^{*2}}{a^2}\right) \quad (4.16)$$

which, for the estimators $\hat{\sigma}_0 = S_{pooled}$ and $\hat{\sigma} = S$ equals, using Equation (4.14),

$$\begin{aligned}
 CPA(\gamma, L^*) &= P\left(\frac{(n-1)S_i^2}{\sigma^2} > \frac{\tilde{Q}}{\gamma^2} \frac{\chi_{n-1;1-\alpha_{tol}}^2}{\chi_{m(n-1);p}^2}\right) \\
 &= 1 - F_{\chi_{n-1}^2}\left(\frac{\tilde{Q}}{\gamma^2} \frac{\chi_{n-1;1-\alpha_{tol}}^2}{\chi_{m(n-1);p}^2}\right)
 \end{aligned}
 \tag{4.17}$$

where $\tilde{Q} = m(n-1)S_{pooled}^2/\sigma_0^2 \sim \chi_{m(n-1)}^2$. The out-of-control ($\gamma \neq 1$) CARL is the reciprocal of $CPA(\gamma, L^*)$ in Equation (4.16) in general, and of Equation (4.17) in particular for the estimators $\hat{\sigma}_0 = S_{pooled}$ and $\hat{\sigma} = S$. The probability in Equation (4.16) can be calculated for various values (quantiles) from the distribution of Y and the out-of-control performance can be examined. Thus, from equation (4.16) it is also possible to determine the distribution of the CPA for given values of γ and L^* .

Tables 4.2 and 4.3 show, for several values of m and n , the values of the adjusted control limit coefficient L^* for the estimator S_{pooled} , for a nominal false-alarm rate of $\alpha_0 = 0.005$, and the combinations of two values of ϵ (0.10 and 0.20) and two values of p (0.05 and 0.10). Table 4.2 corresponds to $\epsilon = 0.10$ and 4.3 corresponds to $\epsilon = 0.20$. For each value of n , the first row gives the values of L^* and the second row gives in italics the adjustment factor, which is the ratio between these values and the unadjusted control limit coefficient L . Other estimators of σ_0 can be considered in a similar manner. As it can be seen and as it might be expected, the adjusted limit converges to the unadjusted limit when n or m (or both) increase. This happens since the variance of the estimators decreases with the increase in the number of observations. Also, keeping m and n fixed, the adjusted limit decreases (becoming closer to the unadjusted limit) with increases in ϵ and p , which mean greater tolerance to CFAR values larger (or to CARL values smaller) than the nominal.

Tables 4.4 and 4.5 give the out-of-control CARLs of the charts for $\gamma = 1.5$ and $\gamma = 2.0$, respectively. They correspond, as mentioned earlier, to the case $w = 1$. In these tables, the column "Unadjusted" gives the out-of-control CARL of the chart with the unadjusted control limit. The values indicate that, with the traditionally recommended numbers of preliminary samples (m between 25 and 75 samples) and the usual sample sizes (less than 10 observations), the adjustment made for the guaranteed in-control performance

entails substantial deterioration of the out-of-control performance. For example, with less than 50 Phase I subgroups each with a sample size of 5, the out-of-control CARLs of the chart with adjusted limits are 40% to 100% larger than the ones of the chart with unadjusted limits. Even with $m = 100$, the out-of-control CARLs for $\gamma = 1.5$ and $n = 3$ are 30% to 50% larger than the ones for the chart with the unadjusted control limit. However, if the increase in the process standard deviation that is relevant to detect quickly is a little bigger, $\gamma = 2.0$, the increases in the out-of-control CARLs due to the adjustment become less substantial, except for the smaller values of n (3 and 5) and m (25 and 50) (see Table 4.5). With larger sample sizes, the CARL values are already close to 1 with or without the adjustment. However, if the user considers for example 100 or 200 Phase I subgroups (a number much smaller and far more practical than the several hundreds to thousands found to be required in Epprecht et al. (2015)), the increase in the CARL (relative to the chart with unadjusted limits) is of the order of 25% or less; that is, the adjustment yields in this case a good compromise between the guaranteed in-control performance, the out-of-control performance and the number of Phase I subgroups.

Note that Faraz et al. (2015) concluded that “adjusting the S^2 control limit did not have too much an effect on the out-of-control performance of the chart.” By contrast, as we have seen, this effect can be significant for small values of m and n . Clearly, the user should seek a combination of a number of preliminary samples, m , and a level of adjustment (through the specification of the values of α_{tol} and p) that result in an appropriate compromise between the risk of a large false-alarm rate and a poor out-of-control performance. The ideal compromise depends on the particular situation. The tables in this chapter provide guidance to the user’s decision, and the formulae in Section 4.2 and this section enable the calculation of the adjusted control limit and the out-of-control CARL for any particular situation (i.e., for specified values of m , n , ϵ , p and α_0 , and for known γ).

p	n	L	L^*				
			L^*/L				
			$m = 25$	50	100	200	500
0.05	3	2.302	2.736	2.584	2.487	2.422	2.368
			<i>1.188</i>	<i>1.123</i>	<i>1.080</i>	<i>1.052</i>	<i>1.029</i>
	5	1.927	2.167	2.086	2.032	1.996	1.965
			<i>1.125</i>	<i>1.082</i>	<i>1.054</i>	<i>1.035</i>	<i>1.019</i>
	10	1.619	1.746	1.704	1.675	1.655	1.638
			<i>1.079</i>	<i>1.052</i>	<i>1.035</i>	<i>1.023</i>	<i>1.012</i>
	15	1.496	1.588	1.557	1.537	1.522	1.510
			<i>1.062</i>	<i>1.041</i>	<i>1.027</i>	<i>1.018</i>	<i>1.009</i>
	20	1.425	1.499	1.475	1.458	1.446	1.436
			<i>1.052</i>	<i>1.035</i>	<i>1.023</i>	<i>1.015</i>	<i>1.008</i>
	25	1.378	1.441	1.420	1.406	1.396	1.387
			<i>1.046</i>	<i>1.031</i>	<i>1.020</i>	<i>1.013</i>	<i>1.007</i>
	30	1.343	1.399	1.381	1.368	1.359	1.352
			<i>1.042</i>	<i>1.028</i>	<i>1.018</i>	<i>1.012</i>	<i>1.006</i>
0.10	3	2.302	2.627	2.513	2.440	2.390	2.349
			<i>1.141</i>	<i>1.092</i>	<i>1.060</i>	<i>1.039</i>	<i>1.020</i>
	5	1.927	2.108	2.046	2.005	1.977	1.953
			<i>1.094</i>	<i>1.062</i>	<i>1.040</i>	<i>1.026</i>	<i>1.013</i>
	10	1.619	1.715	1.683	1.660	1.645	1.632
			<i>1.059</i>	<i>1.039</i>	<i>1.026</i>	<i>1.016</i>	<i>1.008</i>
	15	1.496	1.565	1.542	1.526	1.515	1.505
			<i>1.047</i>	<i>1.031</i>	<i>1.020</i>	<i>1.013</i>	<i>1.006</i>
	20	1.425	1.481	1.462	1.449	1.440	1.432
			<i>1.039</i>	<i>1.026</i>	<i>1.017</i>	<i>1.011</i>	<i>1.005</i>
	25	1.378	1.426	1.410	1.398	1.391	1.384
			<i>1.035</i>	<i>1.023</i>	<i>1.015</i>	<i>1.009</i>	<i>1.004</i>
	30	1.343	1.386	1.371	1.362	1.355	1.349
			<i>1.031</i>	<i>1.021</i>	<i>1.014</i>	<i>1.008</i>	<i>1.004</i>

Table 4.2: Control limit coefficients, unadjusted and adjusted, for S_{pooled} with $\epsilon = 0.10$ and $\alpha_0 = 0.005$. Every second row (in italics) gives the ratio between the adjusted and the unadjusted limit.

CONTROL CHARTS FOR DISPERSION

p	n	L	L^*				
			L^*/L				
			$m = 25$	50	100	200	500
0.05	3	2.302	2.713	2.562	2.501	2.466	2.402
			<i>1.178</i>	<i>1.113</i>	<i>1.087</i>	<i>1.071</i>	<i>1.044</i>
	5	1.927	2.153	2.072	2.018	1.982	1.951
			<i>1.117</i>	<i>1.075</i>	<i>1.047</i>	<i>1.028</i>	<i>1.012</i>
	10	1.619	1.737	1.695	1.666	1.647	1.630
			<i>1.073</i>	<i>1.047</i>	<i>1.029</i>	<i>1.017</i>	<i>1.007</i>
	15	1.496	1.581	1.551	1.530	1.516	1.503
			<i>1.057</i>	<i>1.037</i>	<i>1.023</i>	<i>1.013</i>	<i>1.005</i>
	20	1.425	1.494	1.469	1.452	1.441	1.431
			<i>1.048</i>	<i>1.031</i>	<i>1.019</i>	<i>1.011</i>	<i>1.004</i>
	25	1.378	1.436	1.415	1.401	1.391	1.382
			<i>1.042</i>	<i>1.027</i>	<i>1.017</i>	<i>1.010</i>	<i>1.003</i>
	30	1.343	1.395	1.376	1.364	1.355	1.347
			<i>1.038</i>	<i>1.025</i>	<i>1.015</i>	<i>1.009</i>	<i>1.003</i>
0.10	3	2.302	2.605	2.492	2.446	2.419	2.370
			<i>1.132</i>	<i>1.083</i>	<i>1.063</i>	<i>1.051</i>	<i>1.030</i>
	5	1.927	2.094	2.033	1.992	1.964	1.940
			<i>1.086</i>	<i>1.055</i>	<i>1.033</i>	<i>1.019</i>	<i>1.006</i>
	10	1.619	1.706	1.674	1.652	1.637	1.624
			<i>1.054</i>	<i>1.034</i>	<i>1.020</i>	<i>1.011</i>	<i>1.003</i>
	15	1.496	1.558	1.535	1.519	1.508	1.498
			<i>1.042</i>	<i>1.026</i>	<i>1.016</i>	<i>1.008</i>	<i>1.002</i>
	20	1.425	1.475	1.457	1.444	1.435	1.427
			<i>1.035</i>	<i>1.022</i>	<i>1.013</i>	<i>1.007</i>	<i>1.001</i>
	25	1.378	1.421	1.405	1.394	1.386	1.379
			<i>1.031</i>	<i>1.019</i>	<i>1.011</i>	<i>1.006</i>	<i>1.001</i>
	30	1.343	1.381	1.367	1.357	1.350	1.344
			<i>1.028</i>	<i>1.018</i>	<i>1.010</i>	<i>1.005</i>	<i>1.001</i>

Table 4.3: Control limit coefficients, unadjusted and adjusted, for S_{pooled} with $\epsilon = 0.20$ and $\alpha_0 = 0.005$. Every second row (in italics) gives the ratio between the adjusted and the unadjusted limit.

ϵ	p	n	CARL	CARL				
			Unadjusted	Adjusted with estimator S_{pooled}				
				m=25	50	100	200	500
0.10	0.05	3	10.5	27.8	19.4	15.6	13.6	12.1
		5	6.3	12.6	9.8	8.4	7.6	7.0
		10	3.2	4.9	4.2	3.8	3.6	3.4
		15	2.2	3.0	2.7	2.5	2.4	2.3
		20	1.7	2.2	2.0	1.9	1.8	1.8
		25	1.5	1.8	1.6	1.6	1.5	1.5
		30	1.3	1.5	1.4	1.4	1.4	1.3
0.10	0.10	3	10.5	21.5	16.6	14.1	12.7	11.6
		5	6.3	10.5	8.8	7.8	7.2	6.8
		10	3.2	4.4	3.9	3.7	3.5	3.3
		15	2.2	2.8	2.5	2.4	2.3	2.3
		20	1.7	2.1	1.9	1.9	1.8	1.8
		25	1.5	1.7	1.6	1.5	1.5	1.5
		30	1.3	1.4	1.4	1.4	1.3	1.3
0.20	0.05	3	10.5	26.3	18.5	14.9	13.0	11.6
		5	6.3	12.0	9.4	8.1	7.3	6.7
		10	3.2	4.8	4.1	3.7	3.5	3.3
		15	2.2	2.9	2.6	2.4	2.3	2.2
		20	1.7	2.1	2.0	1.9	1.8	1.8
		25	1.5	1.7	1.6	1.6	1.5	1.5
		30	1.3	1.5	1.4	1.4	1.3	1.3
0.20	0.10	3	10.5	20.4	15.8	13.5	12.1	11.1
		5	6.3	10.1	8.4	7.5	7.0	6.5
		10	3.2	4.3	3.8	3.6	3.4	3.2
		15	2.2	2.7	2.5	2.4	2.3	2.2
		20	1.7	2.0	1.9	1.8	1.8	1.7
		25	1.5	1.6	1.6	1.5	1.5	1.5
		30	1.3	1.4	1.4	1.3	1.3	1.3

Table 4.4: Out-of-control CARL for $\gamma = 1.5$ of the S chart with and without the adjusted limits, when $W = 1$ and $\alpha_0 = 0.005$.

ϵ	p	n	CARL	CARL				
			Unadjusted	Adjusted with estimator S_{pooled}				
				m=25	50	100	200	500
0.10	0.05	3	3.8	6.5	5.3	4.7	4.3	4.1
		5	2.2	3.1	2.8	2.6	2.4	2.4
		10	1.3	1.5	1.5	1.4	1.4	1.4
		15	1.1	1.2	1.2	1.1	1.1	1.1
		20	1.0	1.1	1.1	1.1	1.0	1.0
		25	1.0	1.0	1.0	1.0	1.0	1.0
		30	1.0	1.0	1.0	1.0	1.0	1.0
0.10	0.10	3	3.8	5.6	4.9	4.4	4.2	4.0
		5	2.2	2.9	2.6	2.5	2.4	2.3
		10	1.3	1.5	1.4	1.4	1.4	1.4
		15	1.1	1.3	1.3	1.3	1.2	1.2
		20	1.0	1.1	1.1	1.0	1.0	1.0
		25	1.0	1.0	1.0	1.0	1.0	1.0
		30	1.0	1.0	1.0	1.0	1.0	1.0
0.20	0.05	3	3.8	6.3	5.2	4.6	4.2	4.0
		5	2.2	3.1	2.7	2.5	2.4	2.3
		10	1.3	1.5	1.4	1.4	1.4	1.3
		15	1.1	1.2	1.2	1.1	1.1	1.1
		20	1.0	1.1	1.1	1.0	1.0	1.0
		25	1.0	1.0	1.0	1.0	1.0	1.0
		30	1.0	1.0	1.0	1.0	1.0	1.0
0.20	0.10	3	3.8	5.5	4.7	4.3	4.1	3.9
		5	2.2	2.8	2.6	2.4	2.3	2.3
		10	1.3	1.5	1.4	1.4	1.4	1.3
		15	1.1	1.2	1.1	1.1	1.1	1.1
		20	1.0	2.0	1.9	1.8	1.8	1.7
		25	1.0	1.0	1.0	1.0	1.0	1.0
		30	1.0	1.0	1.0	1.0	1.0	1.0

Table 4.5: Out-of-control CARL for $\gamma = 2$ of the S chart with and without the adjusted limits, when $W = 1$ and $\alpha_0 = 0.005$.

4.4 Balancing In-Control and Out-of-Control Performance

The aim of the proposed correction is to guarantee a minimum in-control performance. However, the resulting out-of-control performance should not be ignored, as detecting out-of-control situations is still the main purpose behind using control charts. To examine this issue, we define the cdf of the CPA, which can be used to evaluate the out-of-control performance more closely. First, recall that the CPA as in (4.16) can be rewritten as

$$CPA(\gamma, L^*) = P\left(\frac{b\hat{\sigma}_i^2}{a^2\sigma^2} > \frac{W^2 bL^{*2}}{\gamma^2 a^2}\right) = 1 - F_{\chi_b^2}\left(\frac{W^2 bL^{*2}}{\gamma^2 a^2}\right).$$

Consequently, in a similar way as equation (4.9), we can write the cdf of $CPA(\gamma, L^*)$ as

$$\begin{aligned} F_{CPA}(t; \gamma, L^*) &= P\left(1 - F_{\chi_b^2}\left(\frac{W^2 bL^{*2}}{\gamma^2 a^2}\right) \leq t\right) = P\left(W^2 > \frac{\gamma^2 a^2 \chi_{b;1-t}^2}{bL^{*2}}\right) \\ &= P\left(\frac{b_0}{a_0^2} W^2 > \frac{b_0 \gamma^2 a^2 \chi_{b;1-t}^2}{a_0^2 bL^{*2}}\right) = 1 - F_{\chi_{b_0}^2}\left(\frac{b_0 \gamma^2 a^2 \chi_{b;1-t}^2}{a_0^2 bL^{*2}}\right) \end{aligned}$$

Note that $F_{CPA}(t; 1, L^*)$ is equal to $F_{CFAR}(t; L^*)$ from equation (4.10). In the out-of-control situation, it is desired to obtain a signal as quickly as possible. This means that in the out-of-control situation the CPA (CARL) is preferred to be large (small). For a given increase in standard deviation γ and adjusted control limit coefficient L^* one can use $F_{CPA}(t; \gamma, L^*)$ to determine the probability of obtaining a CPA smaller than a specified value t . Of course, $F_{CPA}(t; \gamma, L^*)$ is desired to be small for $\gamma > 1$, since this means that a low probability of an alarm (high CARL) is unlikely in the out-of-control situation. Based on the situation (e.g. what sizes of shifts may be desired to be detected), one may argue whether or not such an out-of-control performance is sufficient.

For example, consider the case that $m = 50$ samples of size $n = 5$ each

are available in Phase I, and that we use S_{pooled} as Phase I estimator and S_i as plotting statistic in Phase II. Moreover, consider $\alpha_0 = 0.005$, $\epsilon = 0.1$ (such that $\alpha_{tol} = 0.0055$) and $p = 0.05$. Finally, suppose that we are interested in detecting changes in process dispersion such that $\gamma = 1.5$. The required value of L^* can be found to be equal to 2.086 (from Table 4.2 or Equation (4.14)). This control limit coefficient guarantees that the CFAR is less than $\alpha_{tol} = 0.0055$ (or the CARL is greater than $1/\alpha_{tol} = 182$ when $\gamma = 1$) with probability $1 - p = 0.95$. From Table 4.4 we conclude that, in the absence of estimation error, the CARL of this chart will be 9.8 when $\gamma = 1.5$. Of course, the out-of-control CARL for practitioners depends on the estimation error. Suppose that we are interested in the probability of having CFAR less than 0.067 (or equivalently CARL greater than 15) when $\gamma = 1.5$. For that specific case, one can calculate this probability to be $F_{CPA}(0.067; 1.5, 2.086) = 0.091$, which is an indication of the out-of-control performance.

If the performance is not deemed sufficient, then two alternatives are possible. The first is to increase the amount of Phase I data, and the second is to be more lenient on the guaranteed in-control performance. Gathering more data will provide more accurate estimates of the in-control process in Phase I, which is incorporated in the adjusted control limits. As there is less uncertainty to account for, the adjusted limit coefficient will be smaller, resulting in lower out-of-control CARL values. However, often this amount of data is not available, or costly to collect. If increasing the amount of used Phase I data is not an option, one may have to choose to be more lenient on the adjustment. Using larger values of α_{tol} means that we choose a less strict minimum performance threshold, while using a larger value of p means that there is a larger probability obtaining CARL values below the minimum threshold. Both of these lead to lower adjusted control limit coefficients, and consequently to lower CARL values in both the in-control and out-of-control situation. In the example discussed in the previous paragraph, consider adjusting our values of p and ϵ to $p = 0.1$ and $\epsilon = 0.2$, while leaving the rest as before. This changes α_{tol} to 0.006 and $CARL_{tol}$ to $1/0.006=167$, which indicates a deterioration of the out-of-control performance. We would then find (from 4.2 or equation (4.14)) that $L^* = 2.033$, which results in $F_{CPA}(0.067; 1.5, 2.033) = 0.030$. This is substantially smaller than the 0.091 obtained previously, which illustrates the balancing of the in-control and out-of-control performance.

4.5 Practical Example

In this section we illustrate how the proposed control limits should be implemented in practice, by means of a practical example. We consider a dataset provided in Montgomery (2013) that contains the inside diameters of forged automobile engine piston rings. In Phase I, the control limits are constructed based on $m = 25$ samples of size $n = 5$ each. These obtained limits are then used for monitoring the process standard deviation in Phase II. As an example, the application of the control chart is illustrated using $m = 15$ samples of size $n = 5$ each, obtained in Phase II. The corresponding Phase I and Phase II data can be found in Table 6.3 and Table 6E.8 of Montgomery (2013), pages 260 and 283, respectively.

Before the control limits are constructed, we check if the Phase I data used to calculate them follow a normal distribution. From the Shapiro-Wilk test for normality we find no reason to reject the normality assumption, as the p-value is close to 0.9. This means that we can continue with the construction of the control limits, which is done through the following steps:

1. First, we determine our parameters. We have $m = 25$ and $n = 5$, and we choose $\alpha_0 = 0.005$, $p = 0.1$ and $\epsilon = 0$. Note that for this case $\alpha_{tol} = (1 + \epsilon)\alpha_0 = 0.005$, and consequently $CARL_{tol} = 1/\alpha_{tol} = 200$. Combined with $p = 0.1$, this means that a minimum in-control CARL of at least 200 is guaranteed with $100(1 - p)\% = 90\%$ probability.
2. Secondly, we need to determine our estimator and plotting statistic, and their corresponding values of a , b , a_0 and b_0 . We consider $\hat{\sigma}_0 = S_{pooled}$ as Phase I estimator, and $\hat{\sigma}_i = S_i$ as plotting statistic, which means that we have $a = 1$, $b = n - 1 = 4$, $a_0 = 1$, and $b_0 = m(n - 1) = 100$. As Phase I estimate we find $S_{pooled} = 0.0100$.
3. The values from steps 1 and 2 are implemented in equation (4.13) to obtain $L^* = 2.124$. As a comparison, using equation (4.2) we find $L = 1.928$. We can now calculate the adjusted and unadjusted control limits as $\widehat{UCL}^* = L^*\hat{\sigma}_0 = 2.124 \cdot 0.0100 = 0.0212$ and $\widehat{UCL} = L\hat{\sigma}_0 = 1.928 \cdot 0.0100 = 0.0193$, respectively.

The obtained limit \widehat{UCL}^* can now be used to monitor the process standard deviation, as is illustrated in Figure 4.5 for the provided Phase II dataset. For comparison purposes, we have also added the unadjusted control limit to this figure. As can be seen in the figure, the adjusted limit is larger than the unadjusted limit to prevent low in-control CARL values. Note that, for this specific data, neither of the control limits indicate an out-of-control situation, since all observations are below both limits.

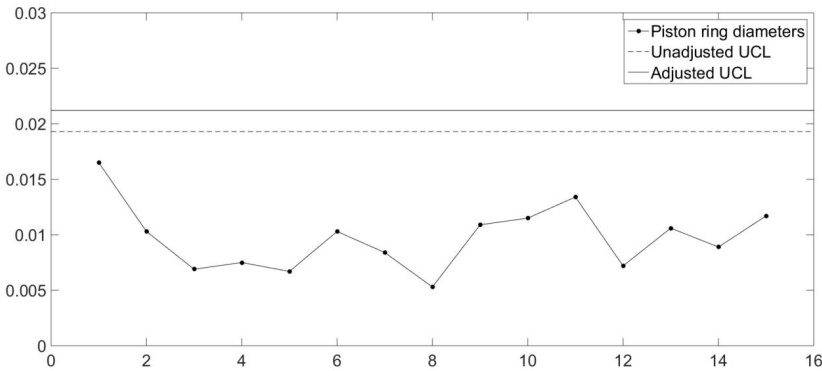


Figure 4.5: Application of the proposed control chart with the piston ring dataset. Both the unadjusted and adjusted control limits are indicated.

4.6 Concluding Remarks

Not accounting for the effects of parameter estimation can cause substantial deterioration in control chart performance, both in the in-control and out-of-control cases. The in-control performance at some nominal value is very important when it comes to implementing and using a control chart that can be relied upon. A Phase I dataset is typically required to estimate the unknown parameters and as the Phase I dataset will differ across practitioners, so will the control chart limits, and consequently the control chart performance. Increasing the amount of available Phase I data will improve the control chart performance and decrease the performance variation among practitioners.

However, the required amount of Phase I data to achieve good (close to nominal) in-control control chart performance is typically infeasible (see Eprecht et al., 2015). Thus, for a practical implementation, we propose the use of adjusted control chart limits for Shewhart control charts for dispersion under normality adopting an alternative point of view. The idea is in line with recent research, for example Faraz et al. (2015), Saleh et al. (2015a,b), and with the results in Chapter 3. The adjusted control limits are determined such that a minimum in-control chart performance is guaranteed with a pre-specified probability. However, whereas Faraz et al. (2015) use a bootstrap approach following Gandy and Kvaløy (2013), we derive analytical expressions to determine the adjusted limits that are easier to implement, and give more insight in the required adjustment. Our adjusted limits allow different available estimators of standard deviation to be used in the analysis, and are easily applicable to any monotone increasing transformation of S .

Because of the formulation based on the CFAR, the adjusted control chart (limit) accounts for parameter estimation and yields better in-control performance compared to the unadjusted chart which does not account for parameter estimation. The in-control performance of the unadjusted chart suffers from parameter estimation and the practitioner-to-practitioner variability and therefore is unreliable. However, note that this gain in the in-control performance of the adjusted chart is seen to lead to a deterioration of the out-of-control performance for smaller number of Phase I subgroups. Thus, a trade-off has to be made in practice to balance the in-control and out-of-control performance properties of the control chart depending on the amount of Phase I data at hand. We recommend using the adjusted limits since the in-control (stability) performance of a control chart is deemed more important in practice and our results show that the loss of some out-of-control performance can be tolerable, particularly for larger amounts of Phase I data.

5. Nonparametric Control Charts

In the previous chapters, control limit adjustments are provided for when data are normally distributed. However, in practice, the distribution of the data is generally unknown. Therefore, we provide alternative methods in this chapter that are less dependent on this normality assumption. In particular, we provide nonparametric control limits, as well as some intermediate parametric methods such as transformations to normality and a bootstrap procedure. All methods described in this chapter are based on the exceedance probability criterion. This chapter is based on Goedhart et al. (2018b).

5.1 Introduction

When the data are normally distributed, the methods described in the previous chapters provide suitable adjustments to compensate for the effect of parameter estimation. Although it is possible to derive adjusted control limits for various distributions, the true distribution of a dataset under consideration is in general unknown and has to be estimated along with its parameters. When the distributional assumptions are violated these adjustments are no longer appropriate, and yield unsatisfactory control chart performance.

As mentioned also in Section 1.4, the total estimation error can be split up in two distinct errors: the model error (ME) and the stochastic error (SE). In order to remove the ME entirely, one can consider the use of nonparametric methods. Most nonparametric methods revolve around the use of order statistics, such as the nonparametric tolerance intervals described in Krishnamoorthy and Mathew (2009). Other methods, such as the bootstrap procedure of Gandy and Kvaløy (2013), are often only partially nonparametric.

Although bootstrapping the Phase I sample may be performed in a nonparametric way, the key of this method is to determine the required limits for each bootstrap replication. This, in turn, requires some distributional assumptions in order to be accurate for small samples. For this reason, Gandy and Kvaløy (2013) themselves already advised to use a parametric instead of nonparametric bootstrap procedure for Shewhart type control charts. We will elaborate on this further in Section 5.3.2. For some general information on nonparametric control charts we refer to Chakraborti et al. (2015).

In this chapter we estimate the control limits based on nonparametric tolerance intervals, and determine the corresponding exceedance probabilities. A major advantage of nonparametric control charts is that they can be applied to individual observations as well as subgroup statistics. For example, when treating subgroup standard deviations as individual observations, one can apply a nonparametric control chart to monitor the standard deviation. The same can be done for various other test statistics as well (e.g. the average, range, or other robust estimators for location or dispersion), regardless of the distribution under consideration. Because the nonparametric setup does not give severe restrictions for the charting statistic Y_i , this makes the approach very general.

However, nonparametric methods generally yield a larger stochastic error compared to parametric methods, which means that larger sample sizes are required in order to obtain accurate results. To this end, we compare the results of the nonparametric methods with several options that make use of adjustments developed for normally distributed data and transformations to normality, as well as a bootstrap procedure. We would also like to emphasize that the choices and consequences regarding sample sizes, model assumptions and accuracy of results, are not limited to SPM. The findings and trade-offs discussed in this chapter also apply to many other statistical applications.

This chapter is organized as follows. In Section 5.2 we illustrate the proposed nonparametric control chart design. Next, in Section 5.3 we discuss some alternative methods and make a comparison. Finally, in Section 5.4 we provide some concluding remarks.

5.2 Nonparametric Control Limits

We want to determine the nonparametric estimated control limits \widehat{LCL} and \widehat{UCL} that satisfy the exceedance probability criterion. Consider again Y_i as the Phase II monitoring statistic, and consider the case that the process is in-control. Note that, in this chapter, Y_i is not assumed to be normally distributed, and can be an individual observation as well as a subgroup statistic. Because of the analogy between the CFAR and the CARL, we again focus on CFAR in this chapter. In this case, the exceedance probability criterion is equal to

$$P(CFAR > \alpha_{tol}) = P\left(1 - P\left(\widehat{LCL} \leq Y_i \leq \widehat{UCL}\right) > \alpha_{tol}\right) = p. \quad (5.1)$$

For a random variable Y_i , this criterion is equivalent to the construction of a tolerance interval with a coverage of at least $1 - \alpha_{tol}$ with probability $1 - p$. For this reason, we consider nonparametric tolerance interval theory.

The nonparametric tolerance intervals as described in Krishnamoorthy and Mathew (2009) are constructed using order statistics. In particular, each limit is determined by a single order statistic. Although this method entirely rules out the ME caused by model misspecification, it is obvious that this increases the SE compared to parametric methods. Other disadvantages of using only a single order statistic for each limit is that this approach generally leads to rather conservative estimates, or to cases where the desired coverage probability can't be guaranteed because of small sample sizes. These issues are addressed in Young and Mathew (2014), who suggest to construct tolerance limits based on interpolated and extrapolated order statistics, depending on the sample size. In this section we explore their procedure for determining the required limits for completeness and evaluate its performance.

5.2.1 Estimation of Nonparametric Control Limits

The first step is to determine whether one should interpolate or extrapolate. This is done by determining the minimum sample size requirement for the construction of a two-sided nonparametric tolerance interval based on unweighted order statistics, as in Krishnamoorthy and Mathew (2009). If the sample size available is large enough, one should interpolate to make the estimated interval less conservative. If the sample size is not large enough, one has to extrapolate to reach a desired exceedance probability. Note that in this chapter the sample size and corresponding requirements relate to m , the number of observations of Y_i , and not to the subgroup size n . The sample size m is sufficient when the following equation holds

$$(m - 1) (1 - \alpha_{tol})^m - m (1 - \alpha_{tol})^{m-1} + 1 \geq 1 - p. \tag{5.2}$$

For a derivation of Equation (5.2) we refer to Section 8.6.1 of Krishnamoorthy and Mathew (2009). Solutions to this equation can for example be determined relatively easily with the function `distree.est()` included in the R-package `tolerance` (see also Young and Mathew, 2014). In Table 5.1 we provide an overview of $m_2(\alpha_{tol}, p)$, the minimum sample size requirement for a two-sided tolerance interval based on unweighted order statistics, for several values of α_{tol} and p . We now make a distinction between the limits when interpolating ($m \geq m_2(\alpha_{tol}, p)$) and when extrapolating ($m < m_2(\alpha_{tol}, p)$).

α_{tol}	p		
	0.2	0.1	0.05
0.05	59	77	93
0.01	299	388	473
0.005	598	777	947
0.0027	1109	1440	1756

Table 5.1: Minimum required sample size $m_2(\alpha_{tol}, p)$ for two-sided tolerance intervals based on unweighted order statistics.

5.2.2 Interpolated Control Limits

When $m \geq m_2(\alpha_{tol}, p)$, the starting point is the two-sided tolerance interval $[X_{(r)}, X_{(s)}]$ according to Krishnamoorthy and Mathew (2009), where $X_{(j)}$ denotes the j -th order statistic from a Phase I sample X_i , with $i = 1, \dots, m$. Note that X_i may be an individual observation or a subgroup statistic. This interval yields a coverage probability of $P(B \leq k - 1) \geq 1 - p$ where $B \sim Bin(m, 1 - \alpha_{tol})$, and where $k = s - r$. In terms of equation (5.1) this is equivalent to $P(CFAR > \alpha_{tol}) \leq p$. Next, the two intervals $[X_{(r+1)}, X_{(s)}]$ and $[X_{(r)}, X_{(s-1)}]$ are considered, which yield a coverage probability of $P(B \leq k - 2) < 1 - p$. Linear interpolation is then used at both sides of the original tolerance interval to obtain

$$\lambda_1 = \frac{(1 - p) - P(B \leq k - 2)}{P(B \leq k - 1) - P(B \leq k - 2)} = \frac{(1 - p) - P(B \leq k - 2)}{P(B = k - 1)} \quad (5.3)$$

$$X_{(r^*)} = \lambda_1 X_{(r)} + (1 - \lambda_1) X_{(r+1)}$$

$$X_{(s^*)} = \lambda_1 X_{(s)} + (1 - \lambda_1) X_{(s-1)}$$

The proposed two-sided nonparametric tolerance interval is then given by the shortest interval of $[X_{(r^*)}, X_{(s)}]$ and $[X_{(r)}, X_{(s^*)}]$. \widehat{LCCL} and \widehat{UCCL} are then set equal to the lower and upper limit of this interval, respectively.

5.2.3 Extrapolated Control Limits

When $m < m_2(\alpha_{tol}, p)$, the coverage probability of $[X_{(1)}, X_{(m)}]$ is not sufficient, and one has to extrapolate. Note that the coverage probability of $[X_{(1)}, X_{(m)}]$ equals $P(B \leq m - 2)$, and that the coverage probability of the intervals $[X_{(1)}, X_{(m-1)}]$ and $[X_{(2)}, X_{(m)}]$ equals $P(B \leq m - 3)$. Then, with linear extrapolation, the following results are obtained

$$\lambda_2 = -\frac{(1 - p) - P(B \leq m - 2)}{P(B \leq m - 2) - P(B \leq m - 3)} = -\frac{(1 - p) - P(B \leq m - 2)}{P(B = m - 2)} \quad (5.4)$$

$$X_{(1^*)} = \lambda_2 X_{(2)} + (1 - \lambda_2) X_{(1)}$$

$$X_{(m^*)} = \lambda_2 X_{(m-1)} + (1 - \lambda_2) X_{(m)}$$

The two-sided nonparametric tolerance interval proposed by Young and Mathew (2014) is then equal to $[X_{(1^*)}, X_{(m^*)}]$. For a nonparametric control chart, one can thus set the \widehat{LCL} and \widehat{UCL} equal to the lower and upper limit of this interval, respectively.

5.2.4 Performance of the Proposed Limits

In this section we evaluate the performance of these methods for various distributions. Although Young and Mathew (2014) already provide an extensive evaluation, their results are focused on theory on tolerance intervals. In that setting, values for α_{tol} of interest are generally in the order of 0.05 to 0.5, while in statistical process monitoring the interest lies much further in the tails. For that reason, we focus mainly on the results for $\alpha_{tol} = 0.0027$, but we have included the values 0.01 and 0.05 as well for completeness.

The evaluation of the proposed control limits is done for various distributions. In particular we consider the standard normal distribution, its exponent (i.e. lognormal with $\mu_0 = 0$ and $\sigma_0 = 1$), a chi-square distribution with 4 degrees of freedom (χ_4^2), and a t -distribution with 4 degrees of freedom (t_4). The standard normal distribution gives a good representation of the performance when normal theory is applicable. The lognormal distribution provides a common skewed alternative. The χ_4^2 provides another skewed alternative, and gives a good indication of the performance of these control limits for various estimators of dispersion. Note that if the data itself are normally distributed, most estimators of dispersion are distributed according to a chi-square distribution, as also discussed in Chapter 4. The chosen degrees of freedom ($df = 4$) corresponds to standard deviations of subgroups of size 5. The t_4 provides a symmetrical alternative to the normal distribution, but with wider tails. Also, similarly to the χ_4^2 , various test statistics for location are based on the t -distribution, such that the chosen $df = 4$ could represent test statistics in subgroups of size 5.

In order to assess the performance of the proposed limits, we have performed a simulation study. In particular, for various combinations of m and α_{tol} together with $p = 0.1$ and $p = 0.2$, we have applied the following procedure

1. A dataset consisting of m observations is drawn from the specified distribution.
2. Estimate the control limits \widehat{LCL} and \widehat{UCL} according to formulas (5.3) and (5.4), depending on the value of m , as described in the previous section. This method is available in the R-package `tolerance` with the function `nptol.int()`.
3. Determine $P(\widehat{LCL} \leq Y_i \leq \widehat{UCL})$ using the in-control distribution of the data.
4. Repeat steps 1 to 3 for 10,000 different Phase I samples, and calculate the proportion for which $1 - P(\widehat{LCL} \leq Y_i \leq \widehat{UCL}) > \alpha_{tol}$. This proportion should be approximately equal to p according to the criterion described in (5.1).

The results of the described procedure are given in Table 5.2 for $p = 0.1$ and Table 5.3 for $p = 0.2$. Recall that the values displayed should be approximately equal to p . As can be seen in the tables, the obtained values are in general close to their desired value. However, one can still observe a deterioration of the performance when α_{tol} becomes smaller, given the sample sizes considered. While we observe only small deviations from p when $\alpha_{tol} = 0.05$, these deviations are substantially larger for $\alpha_{tol} = 0.0027$. Also, for $\alpha_{tol} = 0.0027$ the performance is not entirely consistent for the smallest sample sizes under consideration, as we observe quite some different values for the sample sizes $m = 100, 250$ and 500 .

As we increase the sample size to larger values such as $m = 1500$, the performance moves closer to the desired level. This is, of course, not entirely unexpected because of the sample sizes and α_{tol} values under consideration. For example, estimating a tolerance interval with a coverage of $1 - 0.0027 = 0.9973$ based on order statistics from a sample of $m = 250$ observations is not as accurate as we would like. Note also that $m_2(\alpha_{tol}, p)$, the minimum sample size requirement when using Krishnamoorthy and Mathew (2009), equals 1440 for $\alpha_{tol} = 0.0027$ and $p = 0.1$, as can be seen in Table 5.1. When the interpolation Formula (5.3) is applied, the results for $p = 0.1$ are almost perfect. Moreover, contrary to other parametric methods, the distribution under consideration only has a small impact on the performance. This difference only becomes

slightly larger when the difference in coverage probability between the intervals that are used for the interpolation increases, as can be seen for example when $p = 0.2$ and $\alpha_{tol} = 0.0027$

As observed, the proposed control limits provide better performance in terms of exceedance probability when sample sizes are larger. Moreover, the variation between the resulting CFAR values becomes smaller as sample sizes are increased, due to a smaller SE. This is illustrated as well in Figure 5.1. For $\alpha_{tol} = 0.0027$, $p = 0.1$ and various sample sizes under consideration, this figure shows the variation between CFAR values obtained from the different simulated Phase I samples by means of boxplots. We have used the results from the normally distributed data, but the results for other distributions are similar. The horizontal dotted line indicates the location of α_{tol} , which should be close to the $(1-p)$ -quantile of the boxplots (the actual value can be obtained from Table 5.2). We have left out the case $m = 100$ due to the large variation present there, which would make the rest of the boxplots more difficult to compare.

As can be observed, increasing the sample sizes results in less variation, and less extreme CFAR values. Although perhaps more difficult to detect directly, these extreme CFAR values occur close to zero as well, which can be seen from the location of the bottom side of the box and/or whisker. For $m = 250$, it is obvious from the box that there are many low CFAR values. For $m = 100$, these values are even more frequent. This is mainly caused by the fact that the extrapolations will go further beyond the furthest order statistics obtained when sample sizes are small. Since the extrapolation is linear, this may result in rather extreme estimates of the control limits. This is undesirable, as it may lead to a substantial deterioration in control limit performance in out-of-control situations. When sample sizes are sufficient so that the control limits can be derived by interpolation, the estimated limits will never go beyond the smallest/largest order statistics obtained from the sample, resulting in fewer extreme estimates. Therefore, when sample sizes are small (e.g. $m = 100$), it might be better to step away from the exceedance probability criterion or choose more lenient parameter values (e.g. larger α_{tol} and/or p).

	m	α_{tol}		
		0.05	0.01	0.0027
Normal	100	0.0932	0.2374	0.0927
	250	0.1124	0.1612	0.2146
	500	0.0733	0.0906	0.2477
	1000	0.0820	0.0972	0.1526
	1500	0.1127	0.1089	0.0988
	2500	0.0915	0.0802	0.1034
	χ_4^2	100	0.0836	0.2236
250		0.1028	0.1630	0.1862
500		0.0990	0.0808	0.2299
1000		0.0903	0.0871	0.1494
1500		0.1037	0.0973	0.0884
2500		0.0966	0.0939	0.0987
Lognormal		100	0.0728	0.2355
	250	0.0942	0.1705	0.2200
	500	0.0958	0.0738	0.2401
	1000	0.1000	0.0900	0.1613
	1500	0.0983	0.0900	0.0912
	2500	0.1032	0.1002	0.0943
	t_4	100	0.0904	0.2756
250		0.1013	0.1794	0.2617
500		0.0769	0.0803	0.2744
1000		0.0745	0.1022	0.1705
1500		0.1094	0.1082	0.0909
2500		0.0845	0.0781	0.0999

Table 5.2: Exceedance probabilities for proposed control limits for $p = 0.1$.

NONPARAMETRIC CONTROL CHARTS

	m	α_{tol}		
		0.05	0.01	0.0027
Normal	100	0.1124	0.2664	0.1056
	250	0.1954	0.2192	0.2368
	500	0.1612	0.1239	0.2665
	1000	0.2146	0.1325	0.2129
	1500	0.2066	0.2039	0.2015
	2500	0.2028	0.1732	0.1435
	χ_4^2	100	0.1654	0.2484
250		0.1973	0.2127	0.2179
500		0.1887	0.1703	0.2613
1000		0.1958	0.1856	0.2014
1500		0.1931	0.1920	0.1677
2500		0.1919	0.1890	0.1996
Lognormal		100	0.1835	0.2747
	250	0.1979	0.2235	0.2379
	500	0.1964	0.1765	0.2725
	1000	0.2016	0.1944	0.2089
	1500	0.2016	0.1928	0.1582
	2500	0.1967	0.1989	0.2015
	t_4	100	0.1338	0.3115
250		0.1944	0.2350	0.2883
500		0.1657	0.1331	0.3114
1000		0.2138	0.1432	0.2075
1500		0.2022	0.1998	0.1775
2500		0.2037	0.1748	0.1538

Table 5.3: Exceedance probabilities for proposed control limits for $p = 0.2$.

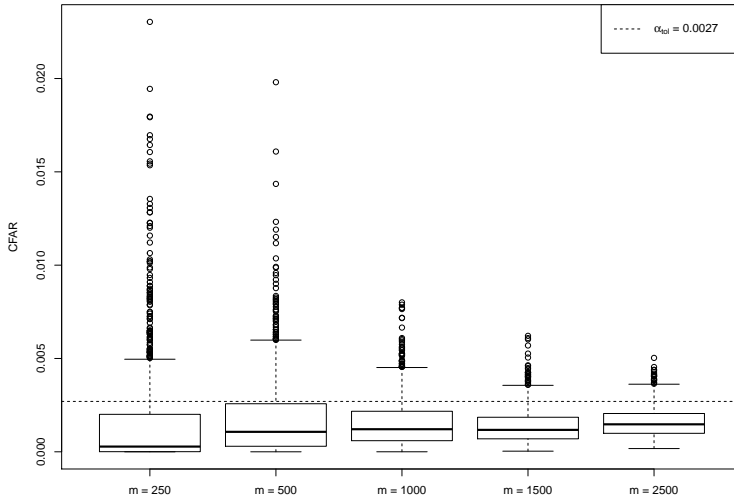


Figure 5.1: Boxplots of CFAR for normally distributed data with $\alpha_{tol} = 0.0027$ and $p = 0.1$, when applying nonparametric control limits. Horizontal dashed line indicates α_{tol} .

5.3 Alternative Parametric Methods

In this section we discuss alternative methods that aim to satisfy equation (5.1). First, we discuss methods based on normal theory, after which we elaborate more on the bootstrap procedure proposed by Gandy and Kvaløy (2013). In parametric methods, SE will be smaller compared to nonparametric methods due to better use of available information. However, ME plays a bigger role there since deviations from the model assumptions might be larger than the actual decrease in SE. Moreover, the errors can differ in size for different distributions. At the end of this section, we include a performance comparison using a numerical study.

5.3.1 Normal Tolerance Interval Methods and Extensions

Various publications have been devoted to tolerance intervals for normal populations, see also Krishnamoorthy and Mathew (2009). Moreover, in Chapter 3 we also discuss control limits based on this theory. However, in practice the problem is that data often aren't normally distributed. To this end, we evaluate some techniques that aim to make normal theory applicable. In particular, we discuss the application of the Central Limit Theorem (CLT) on subgroup averages to create an approximately normally distributed dataset, as well as transformations to normality as proposed for SPM by Chou et al. (1998). The intention of such techniques is to retain the relatively small SE compared to completely nonparametric models, but reducing the accompanying ME caused by deviations from normality.

Application of CLT in SPM

The use of subgroups is a common practice in SPM during data collection. In such a sampling strategy, one could argue to use the subgroup averages as individual observations and apply CLT. Often, subgroups of size 5 are recommended in SPM, while a sample size of 30 is generally deemed enough for CLT to apply. Especially with the recent developments in increasing data availability, one could argue that collecting such amounts of data should not be a problem.

Although CLT works well for the major (middle) part of the data, it is of less use when the far tails are of interest. As recently shown by Huberts et al. (2018), in some cases it might even require subgroups of more than 1,000 observations before the control chart performance is satisfactory for application in SPM. This is primarily caused by the fact that SPM is interested in the far tail of the distribution. While common statistical tests are generally based on a 5% significance level, the original Shewhart control chart based on 3σ -limits goes as far as 0.27%. Also, distributions with wider tails (such as the t -distribution) and highly skewed distributions (lognormal in particular) require much larger subgroup sizes before the control chart behaves as it would under normal theory. While using subgroup averages would thus be appropriate for statistical tests or tolerance intervals based on a 5% significance level or coverage, its benefits in SPM are only marginal.

Transformations to Normality

Another option to make the use of techniques developed for normally distributed data viable, is to transform the data such that a normal distribution is appropriate. One of the most applied transformations in statistics is the Box-Cox transformation, originating from Box and Cox (1964). This method is used to transform data into a symmetrical distribution. Although this generally lowers the deviation from normality, the resulting transformed data are often far from normally distributed. Similar conclusions can be found in Sakia (1992) among others.

A method that specifically intends to transform data to normality in the field of SPM is proposed by Chou et al. (1998). They propose transformations based on the Johnson system of distributions, and provide a step-by-step procedure revolving around the Shapiro-Wilk test for normality. Transformations are only applied when normality of the data is rejected in the first step. One can use this procedure in combination with the suggested control limits in Equation (3.37), which satisfy the exceedance probability criterion for normally distributed data. This results in the following procedure

1. First, test the data for normality using the Shapiro-Wilk test. If normality is not rejected, proceed to step 3. Otherwise, continue with step 2.
2. Transform the data as described in Chou et al. (1998). This can be done for example with the R-package RE.Johnson.
3. Compute the average and standard deviation of the (transformed) data, and use these to construct the control limits using Equation (3.37) with the input parameters m , α_{tol} and p . Note that for the average and standard deviation we have $a = 1$, $b = m - 1$ and $c = 1/m$ in this Equation. Of course, other estimators can also be considered.
4. If the data were transformed, transform the control limits back to their original scale by inverting the transformation used in step 2. If not, no further actions are required.

5.3.2 Bootstrap Procedure

In the previous Subsection 5.3.1 we have described several methods that aim to make the normal tolerance interval theory applicable. Another option is to derive parametric limits for various other specific distributions, as is done in Krishnamoorthy and Mathew (2009). In their book they considered the lognormal, Gamma, two-parameter exponential and the Weibull distribution. However, in practice the true distribution of the data is unknown and still has to be estimated. To this end, we consider methods that aim to provide a more general parametric model in order to find the required limits. Such methods have to be flexible with regard to the location, shape, skewness and kurtosis of the distribution under consideration. A well known system of distributions that allows for such flexibility is the Pearson system, which consists of different types of distributions. An alternative that has one single distribution type based on four parameters to be estimated is given in Low (2013).

The next step is to derive tolerance limits for these general models. To do this, they can be combined with the Gandy and Kvaløy (2013) bootstrap procedure. The idea of the bootstrap procedure is that first, a distribution is fitted to the data. This fit is then assumed to be the true distribution, after which bootstrap samples are drawn from the assumed distribution. These bootstrap samples provide an estimate of the estimation uncertainty accompanying the Phase I data. Then, in each bootstrap replicate, one has to determine the required limits in order to provide a desired result in the assumed distribution of the Phase I sample. This however, requires some distributional assumptions to be accurate, as it requires estimation of small (e.g. $\alpha_{tol}/2$ and $1 - \alpha_{tol}/2$) quantiles of the assumed Phase I distribution. Thus, even when the bootstrap samples are drawn in a nonparametric way (using the empirical cdf), determining the required limits for each bootstrap sample is not accurate when done nonparametrically. To still remain general with the distributional assumptions, one can combine the bootstrap method with the Pearson system of distributions. This procedure is as follows

1. A Pearson distribution is fitted on the dataset based on the first four central moments. Procedures to do this are available in various statistical software programs, such as the function `pearsonFitM` from the R-package `PearsonDS`. We denote this fitted distribution as \hat{F} .
2. From the fitted distribution \hat{F} , draw m_B (e.g. 500 or 1,000) bootstrap

samples of size m each. For each bootstrap sample, fit a Pearson distribution, similarly to step 1. Denote the fitted distribution of the bootstrap sample as \tilde{F} .

3. For each bootstrap sample ($i = 1, \dots, m_B$), determine the required value $\alpha_{B,i}$ such that $\hat{F}(\tilde{F}^{-1}(\alpha_{B,i}/2)) + 1 - \hat{F}(\tilde{F}^{-1}(1 - \alpha_{B,i}/2)) = \alpha_{tol}$. Note that, for two-sided control charts, determining the limits through adjusting the chosen quantiles $\alpha/2$ and $1 - \alpha/2$ seems to be the most reasonable choice for correcting equally on both sides, as treating the LCL and UCL separately may lead to an infinite number of solutions. Note also that, due to the bounded character of some of the Pearson distribution types, this equation does not always have a solution. This happens in particular for highly skewed distributions. When this is the case, set $\alpha_{B,i} = \alpha_{tol}/100$, or some other small value. The reason to choose such a rather small value is that when it occurs rarely, it won't impact the outcome of the bootstrap procedure (see the next step). However, when it's not rare, it means that such small quantiles are actually required to control the performance, but going for even smaller values would result in control limits converging to infinity on one side.
4. Determine the required adjusted value α_{adj} as the p -quantile of the vector α_B obtained in the previous step. Then construct control limits according to $\widehat{LCL} = \hat{F}^{-1}(\alpha_{adj}/2)$ and $\widehat{UCL} = \hat{F}^{-1}(1 - \alpha_{adj}/2)$.

Note that, instead of the Pearson system of distributions, one could also consider other general parametric methods such as Low (2013). However, our results in that case were similar.

5.3.3 Performance Comparison

In order to assess the performance of the methods proposed in this section, we have evaluated the prescribed procedures for various settings. We illustrate the results for the limits from Equation (3.37), both with (denoted Chou) and without (denoted Normal) the transformation proposed in Chou et al. (1998), as well as the Gandy and Kvaløy (2013) bootstrap procedure (denoted GK) described in 5.3.2. Also, we consider the same distributions as in Section 5.2.4.

For various sample sizes m together with $p = 0.1$ and $\alpha_{tol} = 0.0027$, we have assessed the exceedance probability criterion as described in (5.1). Note that in this comparison we only consider $\alpha_{tol} = 0.0027$ since we are interested in the application of these methods in SPM. We have applied the following procedure for each parameter combination

1. A dataset consisting of m observations is drawn from the specified distribution.
2. The methods under consideration are applied to determine the estimated control limits \widehat{LCL} and \widehat{UCL} for each method.
3. Determine $P(\widehat{LCL} \leq Y_i \leq \widehat{UCL})$ using the original in-control distribution of the data.
4. Repeat steps 1 to 3 for 1,000 different Phase I samples, and calculate the proportion for which $1 - P(\widehat{LCL} \leq X \leq \widehat{UCL}) > \alpha_{tol}$. This proportion should be approximately equal to p according to the criterion described in (5.1).

The results of the described procedure can be found in Table 5.4. We have also included the results from Table 5.2 (denoted by YM) for comparison purposes. Recall that the values displayed in the tables should be approximately equal to p . As can be observed from the table, the performance for the parametric methods is not very stable for this small value of α_{tol} , as the outcomes vary substantially. It can also be seen that the nonparametric control limits from Section 5.2 perform best for every distribution except the normal distribution, as the values are in general closest to p . It is of course not unexpected that the other methods perform better for normally distributed data. The control limits of (3.37) are derived specifically for this case, while the Johnson transformation from Chou et al. (1998) is only applied when the normality assumption is rejected. The Pearson system also subsumes the normal distribution as special case in various types. Although the Pearson system also subsumes the t -distribution and the chi-square distribution as special cases, these are each only incorporated in a single type of the Pearson system. Since the type has to be estimated as well in the bootstrap procedure, this yields some extra estimation uncertainty, resulting in an unsatisfactory performance for these distributions. The lognormal distribution is not incorporated in the

Pearson system, and the corresponding results are very unsatisfactory for the bootstrap procedure. Due to the highly skewed character of the lognormal distribution, small deviations in the estimation of the limits lead to large changes in the CFAR or CARL.

We also observe some influence of the sample size m on the performance. However, increasing the sample size to values such as 2,500 does not seem to overcome the performance issues corresponding to small values of α_{tol} for the parametric methods. Taking these aspects in consideration, together with the fact that the true distribution of the data is generally unknown, we argue that the proposed limits in Section 5.2 yield a more stable and satisfying performance when moderately large sample sizes are available. When sample sizes are small, one has to lower the demands placed on the control chart performance. To do this, there are two general directions. The first option is to evaluate which assumptions one can reasonably make about the data, and use these assumptions to determine the control limits. For example, if one is confident that an underlying dataset is (approximately) normally distributed, one can use the control limits from Equation (3.37) for more satisfying results.

The second option is to implement less strict performance criteria, which can be done through parameter settings (e.g. p and α_{tol}). Especially when combined with subgroup or aggregated statistics, the latter is not that strange. For example, consider a process for which an observation is collected once every day. For a control chart with individual observations, $\alpha_{tol} = 0.0027$ corresponds to about one false signal per year. When using some weekly (aggregated) statistic, such as an average, instead of the individual daily observations, the same value of $\alpha_{tol} = 0.0027$ would lead to a false signal about once every seven years. In order to remain with the one false signal per year, one has to increase α_{tol} accordingly. This holds for every situation where observations are aggregated, regardless of the actual time difference between observations. The aggregation itself may help in transforming the data to a known (such as normal) distribution, while somewhat larger values of α_{tol} (in the direction of 0.05) improve the performance of the described methods due to a lower dependency on the far tail behavior. Thus, when aggregating observations or when using subgroup statistics, considering an increased value of α_{tol} is actually a plausible idea.

NONPARAMETRIC CONTROL CHARTS

	m	Normal	Chou	GK	YM
Normal	100	0.0990	0.0970	0.3830	0.0927
	250	0.0960	0.1180	0.1830	0.2146
	500	0.1090	0.1210	0.0900	0.2477
	1000	0.1000	0.1150	0.0880	0.1526
	1500	0.1000	0.1200	0.0890	0.0988
	2500	0.0960	0.1310	0.0870	0.1034
χ_4^2	100	0.9910	0.4590	0.8630	0.0814
	250	1.0000	0.4030	0.7900	0.1862
	500	1.0000	0.3700	0.6870	0.2299
	1000	1.0000	0.3010	0.6260	0.1494
	1500	1.0000	0.2900	0.5500	0.0884
	2500	1.0000	0.2590	0.4650	0.0987
Lognormal	100	0.9800	0.3860	1.0000	0.0963
	250	0.9960	0.3730	1.0000	0.2200
	500	0.9990	0.3520	1.0000	0.2401
	1000	0.9990	0.3190	1.0000	0.1613
	1500	1.0000	0.3220	1.0000	0.0912
	2500	0.9990	0.3270	1.0000	0.0943
t_4	100	0.9880	0.4910	0.5550	0.1243
	250	0.9940	0.3950	0.2650	0.2617
	500	0.9980	0.4590	0.2240	0.2744
	1000	0.9990	0.5660	0.2310	0.1705
	1500	0.9980	0.6420	0.2610	0.0909
	2500	1.0000	0.7360	0.2990	0.0999

Table 5.4: Comparison of exceedance probabilities for $\alpha_{tol} = 0.0027$ and $p = 0.1$.

5.4 Concluding Remarks

Since the data distribution and its corresponding parameters are generally unknown, they have to be estimated using a Phase I reference sample. Because of the variability induced by sampling (stochastic error), the estimated control limits will vary for different Phase I samples. This leads to different control chart performance for different practitioners. To correct for the effect of parameter estimation, we proposed the application of nonparametric tolerance intervals in statistical process monitoring in this chapter, and compare them with some general parametric models.

A major advantage of nonparametric control charts is that they can be applied to any monitoring statistic of interest, by treating subgroup statistics as individual observations. In that way, the nonparametric control chart can be applied to X , \bar{X} , R , and S control charts, or even other test statistics, regardless of the distribution under consideration. The proposed control limits in this chapter perform well in satisfying this criterion in general, especially when moderately large samples are available.

When available sample sizes are small, or when data collection is costly, one may have to be more lenient towards the desired control chart performance. This means that one might have to place some additional restrictions on the used model, such as distributional assumptions, or lower the demands through the choices of parameters p and α_{tol} . The latter is also a feasible choice when sample sizes are made small artificially due to aggregation of individual observations.

6. Summary

In this dissertation we discuss the effects of parameter estimation in statistical process monitoring (SPM), and provide adjusted and new control chart designs to take these effects into account. As data availability is steadily increasing in recent times, the need for tools to monitor large data streams increases with it. The field of SPM revolves around detecting possible changes in a process through monitoring its process data. An important tool that aids in this detection is the control chart. In this dissertation, we focus on a specific type of control chart, namely the Shewhart control chart.

Every process displays variation. For example, if we think of our own body, our heart rate, blood glucose, and many other characteristics vary over time. It is often very difficult to pinpoint the exact cause of these differences, as they are an aggregation of many minor influences. However, certain special events or disturbances can change the underlying process, leading to another source of variation. For example, a virus or disease can affect certain characteristics in our body. Detecting these special events or disturbances early makes it possible to take adequate counter measures, limiting the impact of the special causes.

In order to detect possible changes in an underlying process, a distinction is made between two sources of variation, common cause and special cause variation. The first is variation that is inherent to the process, while the latter is caused by special events. When a process only contains common causes of variation, it is referred to as in-control. When special causes are present, the process is referred to as out-of-control. The first step in order to detect possible out-of-control situations, is to determine the in-control behavior of the process. To this end, a Shewhart control chart consists of two control limits, an upper and a lower control limit (UCL and LCL, respectively) that indicate whether or not a process is in-control.

Originally, the in-control behavior of a process was assumed known. More

specifically, normally distributed data with known mean and standard deviation were assumed in the determination of the control limits. These lead to a known in-control performance in terms of the false alarm rate (FAR) or the average run length (ARL). The FAR indicates the probability that the control chart signals a possible out-of-control situation, while the process is actually in-control. The ARL indicates the average number of observations until a possible out-of-control situation is detected. In practice however, the true mean and standard deviation, as well as the actual distribution of the data, are generally unknown and have to be estimated. The estimation of the in-control behavior is referred to as Phase I.

Estimation of the in-control behavior in Phase I is generally done by means of a reference sample. As different practitioners obtain different samples, this leads to different parameter estimates. Consequently, the estimated control limits and their corresponding performance will differ between practitioners. This variation is often referred to as practitioner-to-practitioner variation. Increasing the sample size reduces this variation, but the sample sizes required for sufficient control chart performance are often not available. Therefore, we consider adjusted control limits in order to compensate for the effect of parameter estimation in Phase I in this dissertation.

In Chapter 2 we consider Shewhart control charts for location for normally distributed data. We provide corrections such that a specified control chart performance is obtained in expectation. The objective considered is referred to as the bias criterion, and relates to the unconditional (expected) performance of the control chart. The reason behind these corrections is that, when parameters are estimated, the expected control chart performance is not equal to the nominal (design) value. The derivations allow different performance measures to be used, such as the FAR or ARL. The related control limit adjustments for Shewhart control charts for dispersion using the bias criterion can be found in Diko et al. (2017). However, a disadvantage of the bias criterion is that there may still be a large proportion of practitioners with an unsatisfactory control chart performance.

In Chapter 3 we again consider Shewhart control charts for location for normally distributed data, but this time in combination with the exceedance probability criterion. This criterion focuses on the conditional performance, by providing a specified conditional in-control performance with a specified probability. We provide control limit adjustments to satisfy this criterion. In

this chapter we also make a comparison with other related methods, such as a bootstrap procedure, self-starting control charts, and tolerance intervals. We also illustrate the out-of-control implications of the proposed adjustments. Since we emphasize the importance of the conditional performance, we remain with the exceedance probability criterion as design criterion in the following chapters.

In Chapter 4 we derive adjusted control limits for control charts for dispersion. When data itself are normally distributed, most estimators of dispersion follow a chi-square distribution. Obviously, the symmetrical form of the control limits described in the previous chapters would not work well for monitoring dispersion. Thus, adjusted probability limits are considered here that take the chi-square distribution of the monitoring statistic into account. Following a similar procedure, one could derive adjusted control limits for monitoring location of non-normally distributed data. However, the true in-control distribution is generally unknown, and has to be estimated as well, along with its parameters.

Therefore, in Chapter 5 we consider nonparametric control limits, as well as methods that aim to make normal (or other parametric) theory applicable. A major advantage of nonparametric control charts is that they are not dependent on the distribution of the data, and can therefore also be applied to many different monitoring statistics. Thus, they can also serve as a control chart for both location and dispersion. The disadvantage is that the sample size requirements to obtain a specified control chart performance increase. When larger sample sizes are not available, one has to put more restrictions on the data under consideration, or be more lenient in the desired performance. Especially when monitoring aggregated statistics rather than individual observations, having a higher FAR value as minimum performance threshold is a suitable solution.

In summary, we have derived several adjusted control limits for Shewhart control charts to take the effect of parameter estimation into account. We have considered the bias criterion and the exceedance probability criterion, and have derived limits for normal theory as well as nonparametric limits. We recommend to focus on the conditional performance by means of the exceedance probability criterion when determining the control limits. Moreover, when sufficient data is available, we recommend the use of nonparametric control limits as they provide a better overall performance for differ-

SUMMARY

ent distributions. When sample sizes are small, one has to choose between placing more demanding assumptions on the dataset, or setting a lower in-control performance threshold. The findings and tradeoffs discussed in this dissertation are also relevant in many other statistical applications. In Figure 6.1 a schematic overview is given of the contributions and recommendations given in this dissertation.

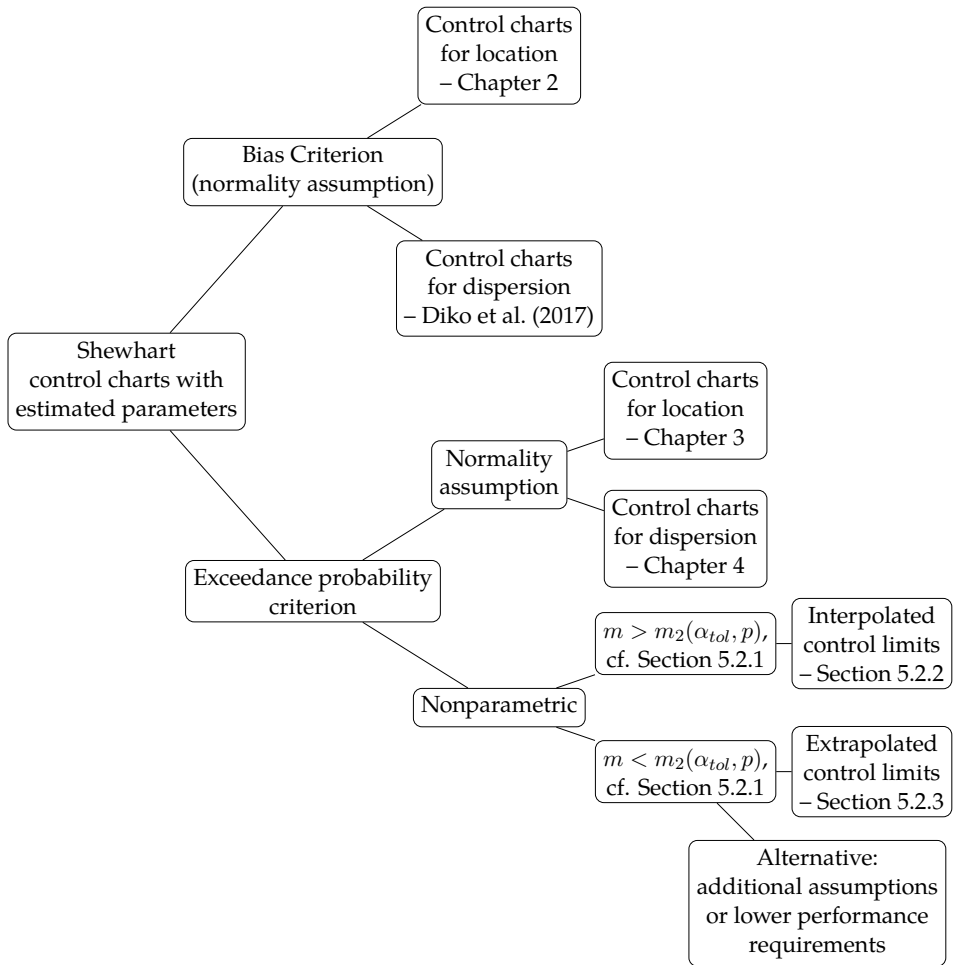


Figure 6.1: Summary and recommendations of this dissertation.

References

- Albers, W. and Kallenberg, W. C. M. (2004a). Are Estimated Control Charts in Control? *Statistics*, 38(1):67–79.
- Albers, W. and Kallenberg, W. C. M. (2004b). Estimation in Shewhart Control Charts: Effects and Corrections. *Metrika*, 59(3):207–234.
- Albers, W. and Kallenberg, W. C. M. (2005). New Corrections for Old Control Charts. *Quality Engineering*, 17(3):467–473.
- Albers, W., Kallenberg, W. C. M., and Nurdati, S. (2004). Parametric Control Charts. *Journal of Statistical Planning and Inference*, 124(1):159–184.
- Box, G. E. P. and Cox, D. R. (1964). An Analysis of Transformations. *Journal of the Royal Statistical Society. Series B (Methodological)*, 26(2):211–252.
- Chakraborti, S. (2006). Parameter Estimation and Design Considerations in Prospective Applications of the \bar{X} Chart. *Journal of Applied Statistics*, 33(4):439–459.
- Chakraborti, S., Qiu, P., and Mukherjee, A. (2015). Editorial to the Special Issue: Nonparametric Statistical Process Control Charts. *Quality and Reliability Engineering International*, 31(1):1–2.
- Chen, G. (1997). The Mean and Standard Deviation of the Run Length Distribution of \bar{X} Charts when Control Limits are Estimated. *Statistica Sinica*, 7(3):789–798.
- 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*, 30(2):133–141.

REFERENCES

- Cryer, J. D. and Ryan, T. P. (1990). The Estimation of Sigma for an X Chart: MR/d_2 or S/c_4 ? *Journal of Quality Technology*, 22(3):187–192.
- Diko, M. D., Goedhart, R., Chakraborti, S., Does, R. J. M. M., and Epprecht, E. K. (2017). Phase II Control Charts for Monitoring Dispersion when Parameters are Estimated. *Quality Engineering*, 29(4):605–622.
- Does, R. J. M. M. and Schriever, B. F. (1992). Variables Control Chart Limits and Tests for Special Causes. *Statistica Neerlandica*, 46(4):229–245.
- Epprecht, E. K., Loureiro, L. D., and Chakraborti, S. (2015). Effect of the Amount of Phase I Data on the Phase II Performance of S^2 and S Control Charts. *Journal of Quality Technology*, 47(2):139–155.
- Faraz, A., Woodall, W. H., and Heuchenne, C. (2015). Guaranteed Conditional Performance of the S^2 Control Chart with Estimated Parameters. *International Journal of Production Research*, 53(14):4405–4413.
- Gandy, A. and Kvaløy, J. T. (2013). Guaranteed Conditional Performance of Control Charts via Bootstrap Methods. *Scandinavian Journal of Statistics*, 40(4):647–668.
- Gardiner, D. A. and Hull, N. C. (1966). An Approximation to Two-Sided Tolerance Limits for Normal Populations. *Technometrics*, 8(1):115–122.
- Goedhart, R., da Silva, M. M., Schoonhoven, M., Epprecht, E. K., Chakraborti, S., Does, R. J. M. M., and Veiga, A. (2017a). Shewhart Control Charts for Dispersion Adjusted for Parameter Estimation. *IIE Transactions*, 49(8):838–848.
- Goedhart, R., Schoonhoven, M., and Does, R. J. M. M. (2016). Correction Factors for Shewhart X and \bar{X} Control Charts to Achieve Desired Unconditional ARL. *International Journal of Production Research*, 54(24):7464–7479.
- Goedhart, R., Schoonhoven, M., and Does, R. J. M. M. (2017b). Guaranteed In-Control Performance for the Shewhart X and \bar{X} Control Charts. *Journal of Quality Technology*, 49(2):155–171.
- Goedhart, R., Schoonhoven, M., and Does, R. J. M. M. (2018a). On Guaranteed In-Control Performance for the Shewhart X and \bar{X} Control Charts. *Journal of Quality Technology*, 50(1):130–132.

- Goedhart, R., Schoonhoven, M., and Does, R. J. M. M. (2018b). Tolerance Intervals in Statistical Process Monitoring. *Submitted for publication*.
- Hawkins, D. M. (1987). Self-Starting CUSUM Charts for Location and Scale. *The Statistician*, 36(4):299–316.
- Hillier, F. S. (1969). \bar{X} - and R -Chart Control Limits Based on a Small Number of Subgroups. *Journal of Quality Technology*, 1(1):17–26.
- Howe, W. G. (1969). Two-Sided Tolerance Limits for Normal Populations – Some Improvements. *Journal of the American Statistical Association*, 64(326):610–620.
- Huberts, L. C. E., Schoonhoven, M., Goedhart, R., Diko, M. D., and Does, R. J. M. M. (2018). The Performance of Control Charts for Large Non-Normally Distributed Datasets. *Quality and Reliability Engineering International*, <https://doi.org/10.1002/qre.2287>.
- Inglot, T. (2010). Inequalities for Quantiles of the Chi-Square Distribution. *Probability and Mathematical Statistics*, 30(2):339–351.
- Jensen, W. A., Jones-Farmer, L. A., Champ, C. W., and Woodall, W. H. (2006). Effects of Parameter Estimation on Control Chart Properties: a Literature Review. *Journal of Quality Technology*, 38(4):349–364.
- Johnson, N. L. and Kotz, S. (1970). *Distributions in Statistics: Continuous Univariate Distributions*. John Wiley and Sons, New York, NY, 1st edition.
- Jones, M. A. and Steiner, S. H. (2012). Assessing the Effect of Estimation Error on Risk-Adjusted CUSUM Chart Performance. *International Journal for Quality in Health Care*, 24(2):176–181.
- Krishnamoorthy, K. and Mathew, T. (2009). *Statistical Tolerance Regions: Theory, Applications, and Computation*. John Wiley and Sons, New York, NY.
- Low, Y. M. (2013). A New Distribution for Fitting Four Moments and its Applications to Reliability Analysis. *Structural Safety*, 42:12–25.
- Mahmoud, M. A., Henderson, G. R., Epprecht, E. K., and Woodall, W. H. (2010). Estimating the Standard Deviation in Quality-Control Applications. *Journal of Quality Technology*, 42(4):348–357.

REFERENCES

- Montgomery, D. C. (2013). *Introduction to Statistical Quality Control*. John Wiley & Sons, Hoboken, NJ., 7th edition.
- Patnaik, P. B. (1950). The Use of Mean Range as an Estimator in Statistical Tests. *Biometrika*, 37(1/2):78–87.
- Psarakis, S., Vyniou, A. K., and Castagliola, P. (2014). Some Recent Developments on the Effects of Parameter Estimation on Control Charts. *Quality and Reliability Engineering International*, 30(8):1113–1129.
- Quesenberry, C. P. (1991). SPC Q Charts for Start-Up Processes and Short or Long Runs. *Journal of Quality Technology*, 23(3):213–224.
- Quesenberry, C. P. (1993). The Effect of Sample Size on Estimated Limits for \bar{X} and X Control Charts. *Journal of Quality Technology*, 25(4):237–247.
- Roes, K. C. B., Does, R. J. M. M., and Schurink, Y. (1993). Shewhart-Type Control Charts for Individual Observations. *Journal of Quality Technology*, 25(3):188–198.
- Sakia, R. M. (1992). The Box–Cox Transformation Technique: a Review. *The Statistician*, 41(2):169–178.
- Saleh, N. A., Mahmoud, M. A., Jones-Farmer, L. A., Zwetsloot, I. M., and Woodall, W. H. (2015a). Another Look at the EWMA Control Chart with Estimated Parameters. *Journal of Quality Technology*, 47(4):363–382.
- Saleh, N. A., Mahmoud, M. A., Keefe, M. J., and Woodall, W. H. (2015b). The Difficulty in Designing Shewhart \bar{X} and X Control Charts with Estimated Parameters. *Journal of Quality Technology*, 47(2):127–138.
- Schoonhoven, M. (2011). *Estimation Methods for Statistical Process Control*. Doctoral Dissertation, University of Amsterdam.
- Schoonhoven, M. and Does, R. J. M. M. (2012). A Robust Standard Deviation Control Chart. *Technometrics*, 54(1):73–82.
- Schoonhoven, M., Riaz, M., and Does, R. J. M. M. (2009). Design Schemes for the \bar{X} Control Chart. *Quality and Reliability Engineering International*, 25(5):581–594.

-
- Shewhart, W. A. (1926). Quality Control Charts. *Bell System Technical Journal*, 5(4):593–603.
- Shewhart, W. A. (1931). *Economic Control of Quality of Manufactured Product*. Van Nostrand Company, New York.
- Tietjen, G. L. and Johnson, M. E. (1979). Exact Statistical Tolerance Limits for Sample Variances. *Technometrics*, 21(1):107–110.
- Tsai, T. R., Lin, J. J., Wu, S. J., and Lin, H. C. (2005). On Estimating Control Limits of \bar{X} Chart when the Number of Subgroups is Small. *International Journal of Advanced Manufacturing Technology*, 26(11-12):1312–1316.
- Vardeman, S. B. (1999). A Brief Tutorial on the Estimation of the Process Standard Deviation. *IIE Transactions*, 31(6):503–507.
- Wald, A. and Wolfowitz, J. (1946). Tolerance Limits for a Normal Distribution. *Annals of Mathematical Statistics*, 17(2):208–215.
- Weissberg, A. and Beatty, G. H. (1960). Tables of Tolerance–Limit Factors for a Normal Distribution. *Technometrics*, 2(4):483–500.
- Wilson, E. B. and Hilferty, M. M. (1931). The Distribution of Chi–Square. *Proceedings of the National Academy of Sciences of the United States of America*, 17(12):684–688.
- Young, D. S. and Mathew, T. (2014). Improved Nonparametric Tolerance Intervals Based on Interpolated and Extrapolated Order Statistics. *Journal of Nonparametric Statistics*, 26(3):415–432.
- Zwetsloot, I. M. (2016). *EWMA Control Charts in Statistical Process Monitoring*. Doctoral Dissertation, University of Amsterdam.

Samenvatting

De effecten van het schatten van parameters in statistische procesbeheersing en de bijbehorende correcties voor regelkaarten (*control charts*) zijn het onderwerp van dit proefschrift. Aangezien de beschikbaarheid van data toeneemt in de huidige tijden, neemt ook de behoefte toe voor methoden om grote datastromen te monitoren. In het vakgebied SPM (*statistical process monitoring*, statistische procesbeheersing) staat het detecteren van mogelijke veranderingen binnen een proces door het monitoren van procesdata centraal. Een belangrijk hulpmiddel voor deze detectie is de regelkaart. Dit proefschrift richt zich op een specifiek type regelkaart, namelijk de *Shewhart* regelkaart.

Elk proces is onderhevig aan variatie. Denk bijvoorbeeld aan je eigen lichaam, waarin je hartslag, bloeddruk, en vele andere karakteristieken variëren over de tijd. Het is vaak erg lastig om de precieze oorzaak van deze verschillen aan te wijzen, aangezien het om combinaties van vele kleine invloeden gaat. Specifieke gebeurtenissen of verstoringen kunnen echter het onderliggende proces veranderen, wat een ander soort variatie tot gevolg heeft. Een virus of ziekte kan bijvoorbeeld invloed hebben op bepaalde karakteristieken in ons lichaam. Het vroeg detecteren van deze gebeurtenissen of verstoringen maakt het mogelijk om tijdig tegenmaatregelen te nemen om de invloed te beperken.

Om het detecteren van veranderingen in een onderliggend proces mogelijk te maken, moet eerst een onderscheid gemaakt worden tussen twee soorten variatie, variatie door normale oorzaken (*common cause variation*) en variatie door bijzondere oorzaken (*special cause variation*). Eerstgenoemde is variatie die inherent is aan het proces, terwijl laatstgenoemde veroorzaakt wordt door speciale gebeurtenissen. Wanneer een proces slechts variatie door normale oorzaken bevat, wordt er gesproken van een *in-control* (beheerst) proces. Als er speciale oorzaken aanwezig zijn, dan wordt gesproken van een *out-of-control* proces. De eerste stap voor het detecteren van mogelijke out-of-control situaties, is het bepalen van de in-control situatie van het proces. Ver-

volgens worden voor de Shewhart regelkaart twee regelgrenzen (*control limits*) geconstrueerd, de UCL (*upper control limit*) en de LCL (*lower control limit*), waarmee bepaald kan worden of een proces in-control is of niet.

Oorspronkelijk werd de in-control situatie van een proces bekend verondersteld. Dit gebeurde door de aanname dat het gemiddelde en de standaardafwijking van een proces bekend zijn bij het opstellen van de UCL en LCL. In dat geval is de prestatie van de regelkaart, in termen van de FAR (*false alarm rate*, de kans op een onterecht out-of-control signaal) of de ARL (*average run length*, het gemiddelde aantal waarnemingen tot een out-of-control signaal), een bekende constante waarde. In de praktijk zijn het gemiddelde en de standaardafwijking van een proces echter onbekend, evenals de werkelijke verdeling van de data, en moeten deze worden geschat. Het schatten van de in-control situatie wordt *Phase I* genoemd.

Het schatten van de in-control situatie in Phase I wordt over het algemeen gedaan met behulp van een representatieve steekproef. Aangezien verschillende gebruikers verschillende steekproeven zullen verzamelen, betekent dit dat zij ook verschillende schattingen verkrijgen. Dit heeft tot gevolg dat de geschatte UCL en LCL, en de bijbehorende regelkaart prestatie, ook verschillen tussen de gebruikers. Het verzamelen van grotere steekproeven reduceert deze variatie tussen verschillende gebruikers, maar de benodigde steekproef omvangen voor een adequate regelkaart prestatie zijn vaak niet beschikbaar. Om die reden richt dit proefschrift zich op aanpassingen van de regelgrenzen ter compensatie van het effect van het schatten van parameters in Phase I.

Hoofdstuk 2 gaat over Shewhart regelkaarten voor locatie (bijvoorbeeld gemiddelde) voor normaal verdeelde data. Hier worden correcties gegeven zodat een gewenste regelkaart prestatie wordt verkregen door gebruikers in verwachting. Dit criterium wordt ook wel het *bias criterion* genoemd, en relateert zich aan de onconditionele (verwachte) prestatie van de regelkaart. De motivatie achter deze correcties is dat wanneer parameters worden geschat, de verwachte regelkaart prestatie niet gelijk is aan de nominale (ontwerp) waarde. De afleidingen in dit hoofdstuk zijn te gebruiken voor verscheidene prestatie indicatoren, zoals de FAR of de ARL. De gerelateerde correcties voor regelkaarten voor spreiding (zoals de standaardafwijking) zijn te vinden in Diko et al. (2017). Een nadeel van sturen op verwachting is echter dat er nog steeds een relatief groot aantal gebruikers kan zijn met een ontoereikende regelkaart prestatie.

Hoofdstuk 3 gaat ook over Shewhart regelkaarten voor locatie voor normaal verdeelde data, maar hier in combinatie met het zogeheten *exceedance probability criterion*. Dit criterium richt zich op de conditionele prestatie van de regelkaart, en houdt in dat tenminste een specifieke conditionele prestatie bereikt wordt met een vooraf gedefinieerde kans. In dit hoofdstuk worden de benodigde aanpassingen aan de regelgrenzen gegeven om dit te bereiken. Ook wordt er een vergelijk gemaakt met andere vergelijkbare methoden. Daarnaast wordt de out-of-control prestatie van de regelkaart onderzocht voor verschillende scenario's. Gezien het belang van de conditionele prestatie in de evaluatie van regelkaarten, richten ook de volgende hoofdstukken zich op dit criterium.

Hoofdstuk 4 gaat over aangepaste regelgrenzen voor regelkaarten voor spreiding. Wanneer data normaal verdeeld zijn, dan volgen veel schatters van spreiding een chi-kwadraat verdeling. In dat geval zijn de symmetrische regelgrenzen uit de voorgaande hoofdstukken niet geschikt. Daarom worden in dit hoofdstuk aangepaste regelgrenzen afgeleid die rekening houden met de chi-kwadraat verdeling van de schatter voor spreiding. Via een zelfde aanpak is het mogelijk om aangepaste regelgrenzen te bepalen voor het monitoren van de locatie van data die niet normaal verdeeld zijn. Echter, de in-control verdeling van de data is over het algemeen onbekend, en moet geschat worden samen met de bijbehorende parameters.

Hoofdstuk 5 gaat daarom over nonparametrische regelgrenzen, evenals methoden die de theorie voor normale data toepasbaar maken. Een voordeel van nonparametrische regelkaarten is dat ze niet afhankelijk zijn van de verdeling van de data, en daardoor ook kunnen worden toegepast voor het monitoren van verschillende statistieken. Ze kunnen dus toegepast worden bij regelkaarten voor zowel locatie als spreiding. Het nadeel is dat de benodigde steekproef omvang om een specifieke regelkaart prestatie te behalen groter is. Als deze steekproef omvang niet beschikbaar is, moeten er meer restricties op de dataset geplaatst worden, of moet een mindere regelkaart prestatie worden geaccepteerd. Wanneer geaggregeerde steekproef statistieken gebruikt worden in plaats van individuele waarnemingen, is het een redelijk alternatief om een hogere FAR als minimum prestatie te accepteren.

Samenvattend bevat dit proefschrift verschillende aangepaste regelgrenzen voor Shewhart regelkaarten om te corrigeren voor het effect van het schatten van parameters. De correcties zijn afgeleid voor zowel onconditionele (in

verwachting) als conditonele prestatie indicatoren. We raden aan om bij het opstellen van de regelgrenzen te concentreren op de conditionele prestaties. Indien genoeg data beschikbaar zijn, adviseren wij om nonparametrische regelgrenzen te gebruiken, aangezien deze een betere algemene prestatie leveren voor verschillende verdelingen. Wanneer de steekproef omvang beperkt is, moet er gekozen worden tussen het aanscherpen van de restricties op de data, of het accepteren van een mindere regelkaart prestatie. De bevindingen en afwegingen in dit proefschrift zijn ook relevant in vele andere statistische toepassingen.

Acknowledgements

The completion of my dissertation has been a milestone in my career. Of course, it wouldn't have been possible without the help of others. I am grateful to many people for the support and encouragement they provided during this period, but I would like to thank a few persons in particular.

First of all, I would like to thank my supervisors, *Ronald Does* and *Marit Schoonhoven*, who have not only provided great guidance during the research in this dissertation, but also contributed significantly to my personal development as a researcher. Marit, I would like to thank you for your always detailed, precise and constructive feedback and research discussions. Ronald, I thank you for your eye for detail and the ability to always ask the right questions. There has always been room for a laugh when working on research. I am also thankful for your confidence in me, not only as a supervisor for the dissertation, but also as director of IBIS UvA. I would like to thank both of you for your time, effort, and pleasant cooperation in the past years.

I would also like to thank my BSc and MSc thesis supervisor *Jan Tuinstra*, who introduced me to IBIS UvA. IBIS UvA is a motivating and pleasant work environment, which is mainly thanks to the colleagues. I would like to thank *Alex Kuiper*, *Inez Zwetsloot* and *Thomas Akkerhuis* for their guidance from the start of my PhD, and the inspiration they have provided for all teaching activities. Furthermore, I would like to thank my IBIS UvA colleagues *Atie*, *Bart*, *Leo*, *Mandla*, *Reza*, *Robert*, *Stevan*, and *Yannik*, as well as the colleagues from Executive Education, fellow PhD candidates, and other UvA colleagues. I would also like to thank my co-authors *Subha Chakraborti*, *Eugenio Epprecht*, *Michele da Silva*, and *Alvaro Veiga*, for their contribution to this dissertation.

On a more personal note, I would like to thank my family. In particular, my parents *Janita* and *Geert Jan*, for their loving care and everlasting support. From my birth I have been blessed with the best people around me that I could

Acknowledgements

have ever wished for. Finally, I want to end with a special word of thanks to the person that I dedicate this dissertation to, my sister *Caroline*, for the love and support she has given me during her life. She has always been an example for me, and it was therefore no coincidence that I followed her with my choice of studying Econometrics. She has been, and will always remain, one of my greatest sources of inspiration.

Rob Goedhart
Amsterdam, August 2018

Curriculum Vitae

Rob Goedhart was born on October 26th, 1993 in Eindhoven, The Netherlands. He obtained his BSc in Econometrics and Operational Research (*cum laude*) in 2013 at the University of Amsterdam (UvA). In 2014 he obtained his MSc in Econometrics with specialization in Mathematical Economics (*cum laude*) at the UvA.

In January 2015 he started his position as consultant and researcher at the Institute of Business and Industrial Statistics (IBIS UvA), which is part of the department of Operations Management of the Amsterdam Business School of the UvA. Here he worked on this dissertation under the supervision of Prof. dr. Ronald Does and Dr. Marit Schoonhoven. His research interests are in industrial statistics, with a specific focus on statistical process monitoring.

As a statistical consultant Rob delivered Lean Six Sigma trainings, and has supervised (Lean Six Sigma) improvement projects. At the Amsterdam Business School, he taught various courses in applied statistics and process improvement.

In his free time Rob likes to do and watch sports, in particular soccer and tennis. He also enjoys listening to music and playing on his guitar or piano.