Multi-aspect NPC control charts: A new class of Shewhart-type control charts

Ca' Foscari University of Venice

Department of Environmental Sciences, Informatics and Statistics

Computer Science Master's Thesis
Academic Year 2021-2022

Graduand	Filippo Maganza
Supervisor	Marco Marozzi
Cosupervisor	Amitava Mukherjee

Acknowledgments

I start by thanking my supervisor, professor M. Marozzi, who has me introduced to the topics discussed in this Thesis and who has always made himself available for clarifying my doubts. I would also like to thank A. Mukherjee, who has shared with me some important ideas and algorithms.

This thesis, as well as my studies, would not have been possible without the help of the people close to me. I would like to thank L. Billotto because she has inspired and helped me to give my best in these years of study. I would also like to thank my family who listened to me and made an effort to understand and discuss my ideas.

Abstract

In the last decades, control charts have proven to be an effective tool to improve the quality of production processes. In recent years, as the Industry 4.0 revolution is progressing, also the interest in Statistical Process Control (SPC) and control charts is growing.

Distribution-free control charts are an interesting family of control charts that has recently been receiving increasing attention from both industry and science. The defining characteristic of distribution-free control charts is that their in-control performances are not affected by the underlying process distribution, so no knowledge regarding the process distribution is required for their practical implementation. This feature makes distribution-free control particularly useful in the early stages of the SPC monitoring phase.

Several Shewhart-type distribution-free control charts for different kinds of monitoring problems have been proposed in the last two decades. The majority of these charts are based on linear rank statistics. In this thesis, we explore a new scheme for devising distribution-free control charts which is based on the Nonparametric Combination of Dependent Tests framework (NPC), borrowed from hypothesis testing and permutation testing theory. The NPC combination idea is very flexible and has already been used for designing multivariate control charts. In this work, we analyze the application of NPC to combine several tests in order to address different aspects of the problem or viewpoints of the data. Studying these multi-aspect NPC charts, we have discovered that combining only linear rank tests leads to a distribution-free NPC chart, whereas combining at least one non linear rank test may lead to a non distribution-free NPC chart.

From a practical perspective, we also propose and evaluate several distribution-free multi-aspect NPC charts for the univariate location and joint location-scale monitoring problems in the standards unknown case. To provide calibration and evaluation routines for multi-aspect NPC charts, a high-performance R package has been developed. Using the implemented algorithms a simulation study has been conducted to compare the proposed charts with the competitors. One of the most interesting results has been obtained with the NPC Wilcoxon-Klotz chart for the location-scale problem. This chart takes full advantage of the flexibility of the NPC framework and has really good performance under a wide variety of interesting distributions.

Keywords   Statistical Process Control, SPC, Control charts, Nonparametric Combination of Dependent Tests, NPC, NPC control charts, multi-aspect control charts, multi-aspect NPC control charts

Contents

Chapter 1

Introduction

In the last decades, control charts have proven to be an effective tool to improve the quality of production processes [26, p. 189]. In recent years, as the Industry 4.0 revolution is progressing, also the interest in Statistical Process Control (SPC) and control charts is growing. Contributing to the spread of SPC methods, there is also the discovery of several new possible applications like health care monitoring [1], financial fraud detection [33] and service quality monitoring [25].

Distribution-free control charts are an interesting family of control charts that has recently been receiving increasing attention from both industry and science. The defining characteristic of distribution-free control charts is that their in-control performances are not affected by the underlying process distribution, so no knowledge regarding the process distribution is required for their practical implementation [6] [4, p. 34]. This feature makes distribution-free control particularly useful in the early stages of the SPC monitoring phase.

Several distribution-free control charts for different kinds of monitoring problems have been proposed in the last two decades. The majority of these charts are based on linear rank statistics. In this thesis, we explore a new scheme for devising distribution-free control charts which is based on the Nonparametric Combination of Dependent Tests framework (NPC), borrowed from hypothesis testing and permutation testing theory. The NPC combination idea is very flexible and has already been used for designing multivariate control charts [11]. In this work, we analyze the application of NPC method to combine several tests in order to address different aspects of the problem or viewpoints of the data.

1.1 Reading guidelines

Chapter 2 and 3 of this Thesis covers the theoretical background required to understand the methods and results presented later on. In particular:

- Chapter 2 forms a short introduction to the principles of statistical hypothesis testing, linear rank statistics, the permutation testing approach and the NPC framework. This Chapter presents also some statistics that will be then experimented in the context of Statistical Process Control.

- Chapter 3 covers the theoretical background of Statistical Process Control and control charts.

After the introductory part of this Thesis, Chapter 4 is entirely dedicated to the proposed multi-aspect NPC scheme for control charts. Starting by introducing

the general concept of multi-aspect NPC charts, we will then describe our software implementation that helped us answer some important research questions concerning these charts. We end up with the presentation and a comparative simulation study of some interesting multi-aspect NPC charts for the location and joint location-scale monitoring problems.

Chapter 2

Theory: Hypothesis testing and the permutation testing approach

This Chapter forms a short introduction to the principles of statistical hypothesis testing, the permutation testing approach and the NPC framework. The notions and algorithms discussed here are useful to understand the key ideas behind the multi-aspect NPC control charts presented in Chapter 4.

2.1   Brief introduction to hypothesis testing

In many scientific studies, a fundamental problem is testing theories in presence of uncertainty [3]. In the field of Statistics, this problem has been a central topic of study for a long time. Statisticians, over the last two centuries, have developed several mathematical-based methods and frameworks that address the hypothesis testing problem. The most popular and widely accepted approach is significance-based hypothesis testing, which, given its importance, is commonly referred to as statistical hypothesis testing.

In statistical hypothesis testing, a *null hypothesis* H_0 is tested against an *alternative hypothesis* H_A and there are two possible outcomes for a test: rejecting or not the null hypothesis. The fundamental problem addressed by the significance-based approach is the maximization of the probability of rejecting a false null[1] (*type 2 error* rate) subject to a bound for the probability of rejecting a true null (*type 1 error* rate) [20, p.56]. This bound to the type 1 error rate, denoted with α, is called *significance level* and is an arbitrary parameter of statistical tests. Instead, the probability of rejecting a false null, which takes the name of *power* of a test, is an intrinsic characteristic of a statistical test and depends on several factors:

- The adopted statistical test procedure

- The degree of falseness of the null

- The adopted significance level α

- The number of observations

Designing a statistical test for a given null and alternative requires three ingredients: a *test statistic*, the specification of the *null distribution* and, optionally, a set

---

[1]For the sake of brevity, from this point on the null hypothesis and the alternative hypothesis may simply be called the null and alternative

of additional assumptions. The test statistic T is a function of the observed data $\mathbf{X}$ and should be designed to be "sensible" to deviations from the null towards the alternative. The null reference distribution is the distribution that the test statistic would assume if the null was true and provided that the assumptions of the test are satisfied. This distribution can be deducted analytically independently from the data or, as explained later in Section 2.5, can be obtained using computational methods using the observed data.

When a statistical test is conducted, two equivalent and closely related methods can be used in order to decide whether reject or not the null.

The first method consists in calculating the *p-value*, i.e. the probability, under the null hypothesis, of obtaining a test statistic value at least as extreme as the value actually observed. The notion of extreme value is tightly connected to the specification of the alternative: A value that provides stronger support in favor of the alternative is considered more extreme. After the p-value has been obtained, if it is below the significance level of the test, then the null hypothesis is rejected, otherwise it is not.

The second method consists in dividing the sample space associated to the test statistic T into a *rejection rejection* and a *complementary region*. These two subsets are constructed such that:

- The values in the rejection region provide stronger support to the alternative than the values in the complementary region.

- Under the null, the test statistic will fall inside the rejection region with probability α and inside the complementary region with probability $1 - \alpha$.

After observing the data, if the observed value of the test static falls inside the rejection region then the null is rejected, otherwise it is not.

An hypothesis (null or alternative) is called *simple* if the statement completely specifies the populations of interests, otherwise it is called *composite* [17, p.18]. When a true null hypothesis is composite, then the type 1 error rate can vary depending on the actual populations. A fundamental requirement of all statistical tests is that the least upper bound (or supremum) of the probability of rejecting a true null, i.e. the *size of the test*, doesn't exceed the significance level α. If this condition is satisfied, then the test is said *conservative* at level α. For a composite or simple null, a test is said *exact* if the size of the test is equal to α for any possible instance of the null and for any α.

An important qualitative criterion used to choose between statistical tests is consistency. A test is said *consistent* for an alternative hypothesis if, when any instance of the alternative is true, the power of the test approaches 1 as the number of observations approaches infinity. Being a qualitative property of statistical tests, consistency does not provide a rule for selecting the best test. However, the consistency criterion is often useful for eliminating some options from the pool of available tests.

Let T_a and T_b be two tests of a null hypothesis $H_0 : \theta = \theta_0$ and alternative hypothesis $H_A : \theta > \theta_0$, where θ is a real population parameter. Assume that T_a and T_b are conservative for any significance level α. We denote with $N_T(x, \alpha, \beta)$ the sample size necessary for the test T in order to attain the power β under significance level α and with underlying parameter $\theta = x$. The *relative efficiency*

of T_a relative to T_b is defined as:

$$RE_{T_a, T_b}(x, \alpha, \beta) = \frac{N_{T_b}(x, \alpha, \beta)}{N_{T_a}(x, \alpha, \beta)}$$

The *asymptotic relative efficiency* (ARE) of T_a relative to T_b is the limit of $RE_{T_a, T_b}(x, \alpha, \beta)$ as $x \to \theta_0$ and while α and β are held fixed. In some cases, the ARE doesn't depend on α or β and can be derived using analytical tools.

2.2   Two-sample problems

Two-sample problems play a central role in hypothesis testing. Within the well-known population model, this type of statistical problem is concerned with finding distributional differences between two populations of interest. Based on this general notion, there are many different concrete formulations that focus on specific kinds of distributional differences and use different sets of assumptions. Among these formulations, the two-sample location and scale problems are two of the most popular.

Assuming an univariate setting, let $\mathbf{X_1} = (X_{11}, ..., X_{1n_1})$ be a random sample drawn from population X_1 with cumulative distribution function (CDF) F_1 and $\mathbf{X_2} = (X_{21}, ..., X_{2n_2})$ be a random sample drawn, independently from $\mathbf{X_1}$, from population X_2 with CDF F_2. The symbols X_1 and X_2 will also be used interchangeably to denote arbitrary observations from the first and second sample respectively, and $N = n_1 + n_2$ will be used to denote the size of the *pooled sample* $\mathbf{X} = (\mathbf{X_1}, \mathbf{X_2})$.

Usually, the null hypothesis being tested in two sample problems is that $\mathbf{X_1}$ and $\mathbf{X_2}$ are drawn from identical populations [17, p.227]:

$$H_0 : F_1(x) = F_2(x) \quad \forall x \in \mathbb{R} \tag{2.1}$$

Which means that X_1 and X_2 have identical distributions. The alternative hypothesis in two sample problems depends on the possible distributional differences between the two populations about which the experimenter is uncertain.

Two-sample location problem   When uncertainty is only about a location difference, a common assumption is that the CDFs of X_1 and X_2 may differ only through a location shift of an unknown amount θ:

$$F_2(x) = F_1(x - \theta) \quad \theta \in \mathbb{R} \quad \forall x \in \mathbb{R}$$

Which means that X_2 has the same distribution of $X_1 + \theta$. Under this *location model* (or *shift model*), the one-sided and two-sided alternatives hypothesis of the location problem can be formulated as follows:

$H_A : \theta < 0$ the distribution of X_2 is shifted to the left (one-sided)

$H_A : \theta > 0$ the distribution of X_2 is shifted to the right (one-sided)

$H_A : \theta \neq 0$ two-sided alternative

Also, the null hypothesis 2.1 can be equivalently stated as:

$$H_0 : \theta = 0$$

Under the location model assumption, the difference between the population means, $E[X_2] - E[X_1]$, and between population quantiles of the same order, $F_2^{-1}(p) - F_1^{-1}(p)$, must be equal to θ [17, p.228].

Two-sample scale problem   When uncertainty is only about a scale difference, the *scale model* can be adopted:

$$F_2(x) = F_1\left(\frac{x}{\delta}\right) \quad \delta \in \mathbb{R} \quad \forall x \in \mathbb{R}$$

Which means that X_2 is distributed as δX_1. Under this model, it assumed that the two populations are identical in all aspects with the exception that the scale of X_2 may be enlarged or compressed according as $\delta > 1$ or $\delta < 1$, respectively. The one-sided and two-sided alternatives hypothesis of the two-sample scale problem can be formulated as follows:

$$H_A : \delta > 1 \text{ scale enlargement of } X_2 \text{ (one-sided)}$$
$$H_A : \delta < 1 \text{ scale compression of } X_2 \text{ (one-sided)}$$
$$H_A : \delta \neq 1 \text{ two-sided alternative}$$

And the null hypothesis 2.1 can be equivalently stated as:

$$H_0 : \delta = 1$$

It is important to consider both the location model and scale models do not use distributional assumptions concerning the shape of the populations.

2.3   Parametric and non-parametric tests

Statistical tests can be distiguished in two broad classes: *parametric* and *non-parametric* tests.

Parametric tests typically assume that the data comes from a specific family of parametric distributions and use mathematical analysis (or simulation) to derive the null reference distribution of the test statistic. The assumption of a parametric model allows for very interesting theoretical results. Indeed, under a parametric model, it is often possible to use the Neyman-Pearson framework in order to find the optimal tests and to derive their performance [17, p.227]. For example, when the populations are normal, it is well known that the Student's t test is the best test for location differences.

A fundamental practical problem of parametric tests is that any conclusion reached using a parametric model may be baseless if the adopted statistical procedure relies on a distributional assumption that is not satisfied in reality. Furthermore, even when large sample approximations are used, the accuracy of these approximations is often difficult to assess [31, p.1].

Non-parametric tests overcome these problems by avoiding assumptions on the distribution of the data. This means that the type 1 error rate of non-parametric tests is kept at the desired (nominal) significance level no matter the underlying distribution. A central consequence of this property is that non-parametric tests can be applied even when there is lack or incomplete knowledge of the underlying distribution [24]. Typically, non-parametric tests require that the null hypothesis is of identical populations (see 2.1), which is sometimes called strong null hypothesis.

2.4   Linear rank statistics

Many popular non-parametric tests for two-sample problems are based on *linear rank statistics*. Linear rank statistics are a family of statistics given by a linear

function of the indicator random variables Z_i ($i = 1...N$), where $Z_i = 1$ if the ith random variable in the pooled ordered sample is from the first sample and $Z_i = 0$ otherwise. In mathematical language, linear rank statistics have the following expression:

$$T(\mathbf{X}) = \sum_i^N c_i Z_i \tag{2.2}$$

Where the c_i are given constants called weights or scores.

Linear rank statistics are addressed extensively in the book [17], which provides the proofs for a series of important general properties for their distribution. For instance, the book gives general formulas for the expectation, variance and covariance of these statistics.

Nonparametric tests are not necessarily bound to linear-rank statistics, indeed a nonparametric test can be developed starting from an arbitrary statistic using the so-called permutation testing framework. The permutation framework is more general of tests based on linear rank statistics in the sense that linear rank tests can be thought of as a subclass of permutation tests. An important feature of linear rank statistics is that their null distribution depends only on the sample sizes n_1 and n_2. As we will see soon, this is not true in general within the permutation testing framework.

2.5   Permutation testing framework

Permutation testing is a non-parametric approach to hypothesis testing which is commonly used in the context of two-sample hypothesis testing problems and works with both the population and randomization models. The permutation testing concept dates back to the 1930's, the earliest contributions were *Fisher's statistical methods for research workers* (1934) [14] and *The design of experiments* (1935) [15].

Permutation inference is based on the key assumption that observations from different groups are *exchangeable* under the null hypothesis [3, p.7-8]. A set of observations is said to be exchangeable if their joint distribution is invariant under permutations of the observations. Random sampling is a sufficient but not a necessary condition for exchangeability [3, p.7] [31, p.36].

Provided that response variables are exchangeable under the null, the permutation testing approach compares the test statistic against its permutation distribution in order to quantify how extreme the observed sample would be if the null was true. The permutation distribution of a test statistic, which takes the role of the null reference distribution, is obtained by considering observed data permutations as equally likely outcomes. This distribution is, in the general case, dependent on the observed data, therefore it can only be calculated after the data have been collected.

2.5.1   Theory of univariate permutation tests

Formally, permutation tests can be derived from the notion of conditioning on a set of sufficient statistics in H_0 for the underlying distribution which is typically unknown [31, pg. 33] [30]. Assume that the pooled sample $\mathbf{X} = (\mathbf{X_1}, \mathbf{X_2}) = $

$(X_0, ..., X_N)$ has an associated N-dimensional probability space $(\mathcal{X}, \mathcal{A}, P)$. Furthermore, let $\mathbf{u}^* = (u_1^*, ..., u_n^*)$ denote an arbitrary permutation of unit labels $(1, ..., N)$, then $\mathbf{X}^* = (X_0^*, ..., X_N^*) = (X_{u_0^*}, ..., X_{u_N^*})$ is the corresponding permutation of $\mathbf{X}$. Statistical inference within the permutation framework is based on the conditioning on the event $\mathcal{X}_{/\mathbf{X}}$, which is the set containing all permutations of $\mathbf{X}$ and taking the name of *permutation sample space*.

It follows from the exchangeability assumption that each permutation $\mathbf{X}^* \in \mathcal{X}_{/\mathbf{X}}$ of $\mathbf{X}$ is a set of sufficient statistic for the underlying distribution of the data, therefore, for any event $A \in \mathcal{A}$:

$$Pr(A; P|\mathcal{X}_{/\mathbf{X}}) = Pr(A|\mathcal{X}_{/\mathbf{X}}) \tag{2.3}$$

Which means that the conditional distribution of $\mathbf{X}$ (and of any of its permutation $\mathbf{X}^*$), given $\mathcal{X}_{/\mathbf{X}}$, is independent of the underlying distribution.

Another fundamental consequence of excheangeability is that the probability of any event $A \in \mathcal{A}$ conditional to $\mathcal{X}_{/\mathbf{X}}$ has the following expression:

$$P(A|\mathcal{X}_{/\mathbf{X}}) = \frac{\sum_{\mathbf{X}^* \in \mathcal{X}_{/\mathbf{X}}} I(\mathbf{X}^* \in A)}{|\mathcal{X}_{/\mathbf{X}}|} \tag{2.4}$$

Where $|\mathcal{X}_{/\mathbf{X}}|$ denotes the cardinality of the permutation sample space and $I(c)$ denotes the indicator function which is equal to 1 when the condition c is true and is equal to 0 otherwise. Consider now a test statistic $T : \mathcal{X}^n \to \mathbb{R}$. From 2.3 it follows that the conditional distribution of T given $\mathcal{X}_{/\mathbf{X}}$ is also independent of P. Then, from 2.4 we may write:

$$P(T \le y|\mathcal{X}_{/\mathbf{X}}) = F_T^*(y) = \frac{\sum_{\mathbf{X}^* \in \mathcal{X}_{/\mathbf{X}}} I(T(\mathbf{X}^*) \le y)}{|\mathcal{X}_{/\mathbf{X}}|} \tag{2.5}$$

$F_T^*(y)$ is known as the permutation CDF of T and, as anticipated before, takes the role of the null reference distribution in permutation tests. For a left tail alternative, the permutation p-value λ_T^L associated with the statistic T is given by:

$$\lambda_T^L = F_T^*(t) = \frac{\sum_{\mathbf{X}^* \in \mathcal{X}_{/\mathbf{X}}} I(T(\mathbf{X}^*) \le t)}{|\mathcal{X}_{/\mathbf{X}}|} \tag{2.6}$$

Where t denotes the observed value of the test statistic, i.e. $t = T(\mathbf{X})$. The right tail permutation p-value instead is given by:

$$\lambda_T^R = \frac{\sum_{\mathbf{X}^* \in \mathcal{X}_{/\mathbf{X}}} I(T(\mathbf{X}^*) \ge t)}{|\mathcal{X}_{/\mathbf{X}}|} \tag{2.7}$$

Obtaining and interpreting the two-sided permutation pvalue λ_T^{RL} requires some additional considerations. The typical definition of λ_T^{RL} is the following:

$$\lambda_T^{RL} = \frac{\sum_{\mathbf{X}^* \in \mathcal{X}_{/\mathbf{X}}} I(|T(\mathbf{X}^*)| \ge |t|)}{|\mathcal{X}_{/\mathbf{X}}|} \tag{2.8}$$

$$= \frac{\sum_{\mathbf{X}^* \in \mathcal{X}_{/\mathbf{X}}} I(T(\mathbf{X}^*) \ge |t|)}{|\mathcal{X}_{/\mathbf{X}}|} + \frac{\sum_{\mathbf{X}^* \in \mathcal{X}_{/\mathbf{X}}} I(T(\mathbf{X}^*) \le -|t|)}{|\mathcal{X}_{/\mathbf{X}}|} \tag{2.9}$$

It is very important to consider that the two terms in 2.9 coincide only if the permutation distribution of T is symmetric about 0. When a two-sided alternative is used, it is generally required that the permutation distribution is symmetric about 0. In [10] it has been shown that an asymmetric permutation distribution may lead to a two-sided permutation test with power equal to 0, which means that the test is also not consistent.

2.5.2 Permutationally equivalent statistics

The notion of permutational equivalence is useful to discover groups of statistics which yields the same permutation inference for any given dataset. If permutation equivalence relation is found between two statistic, then the statistic which is less expensive to compute can be used to optimize the permutation testing procedure.

Proposition 2.5.1. *If two statistics T_1 and T_2 are one-to-one increasingly related given $\mathcal{X}_{/\mathbf{X}}$ for every possible $\mathbf{X}$, then they always induce the same one sided permutation p-values on the same data, symbolically :*

$$\lambda_{T_1}^L = \lambda_{T_2}^L, \ \lambda_{T_1}^R = \lambda_{T_2}^R \quad \forall \, \mathbf{X}$$

Proof. Let T_1 and T_2 be two statistics that are one-to-one increasingly related given $\mathcal{X}_{/\mathbf{X}}$ for every possible $\mathbf{X}$ by a function $g_{\mathcal{X}_{/\mathbf{X}}}$. Then, given $\mathcal{X}_{/\mathbf{X}}$ and for every possible $\mathbf{X}$, it holds that $T_1 = g_{\mathcal{X}_{/\mathbf{X}}}(T_2)$ and, since $g_{\mathcal{X}_{/\mathbf{X}}}$ is one-to-one increasing, that:

$$T_1(\mathbf{X}^*) \leq T_1(\mathbf{X}) \Leftrightarrow T_2(\mathbf{X}^*) \leq T_2(\mathbf{X})$$
$$T_1(\mathbf{X}^*) \geq T_1(\mathbf{X}) \Leftrightarrow T_2(\mathbf{X}^*) \geq T_2(\mathbf{X})$$

Which proves that $\lambda_{T_1}^L = \lambda_{T_2}^L$ and that $\lambda_{T_1}^R = \lambda_{T_2}^R$ since the cardinality of the permutation sample space $|\mathcal{X}_{/\mathbf{X}}|$ doesn't depend on the test statistic. $\square$

Proposition 2.5.2. *If two statistics T_1 and T_2 are one-to-one decreasingly related given $\mathcal{X}_{/\mathbf{X}}$ for every possible $\mathbf{X}$, then they always induce the same one sided permutation p-values on the same data when alternatives are inverted, symbolically:*

$$\lambda_{T_1}^L = \lambda_{T_2}^R, \ \lambda_{T_1}^R = \lambda_{T_2}^L \quad \forall \, \mathbf{X}$$

Proof. The proof follows the same reasoning of that given for Proposition 2.5.1 $\square$

If statistics T_1 and T_2 are one-to-one increasingly or decreasingly related unconditionally, then the same relation holds also when we condition on the event $\mathcal{X}_{/\mathbf{X}}$. If two statistics are one-to-one increasingly or decreasingly related given $\mathcal{X}_{/\mathbf{X}}$ for every possible $\mathbf{X}$, then they are said *permutationally equivalent*.

Proposition 2.5.3. *If, given $\mathcal{X}_{/\mathbf{X}}$ and for every possible $\mathbf{X}$, two statistics T_1 and T_2 are one-to-one related by a function $g_{\mathcal{X}_{/\mathbf{X}}}(z) = c_{\mathcal{X}_{/\mathbf{X}}} z$ where $c_{\mathcal{X}_{/\mathbf{X}}}$ is a constant that may depend only on $\mathcal{X}_{/\mathbf{X}}$, then they always induce the same two-sided permutation p-value on the same data, symbolically:*

$$\lambda_{T_1}^{LR} = \lambda_{T_2}^{LR} \quad \forall \, \mathbf{X}$$

Proof. Let T_1 and T_2 be two statistics that, given $\mathcal{X}_{/\mathbf{X}}$ for every possible $\mathbf{X}$, are related by a function $g_{\mathcal{X}_{/\mathbf{X}}}(z) = c_{\mathcal{X}_{/\mathbf{X}}} z$, then:

$$
\begin{aligned}
\lambda_{T_1}^{RL} &= \frac{\sum_{\mathbf{X}^* \in \mathcal{X}_{/\mathbf{X}}} I(|T_1(\mathbf{X}^*)| \geq |t_1|)}{|\mathcal{X}_{/\mathbf{X}}|} \\
&= \frac{\sum_{\mathbf{X}^* \in \mathcal{X}_{/\mathbf{X}}} I(|c_{\mathcal{X}_{/\mathbf{X}}} T_2(\mathbf{X}^*)| \geq |c_{\mathcal{X}_{/\mathbf{X}}} t_2|)}{|\mathcal{X}_{/\mathbf{X}}|} \\
&= \frac{\sum_{\mathbf{X}^* \in \mathcal{X}_{/\mathbf{X}}} I(|c_{\mathcal{X}_{/\mathbf{X}}}||T_2(\mathbf{X}^*)| \geq |c_{\mathcal{X}_{/\mathbf{X}}}||t_2|)}{|\mathcal{X}_{/\mathbf{X}}|} \\
&= \frac{\sum_{\mathbf{X}^* \in \mathcal{X}_{/\mathbf{X}}} I(|T_2(\mathbf{X}^*)| \geq |t_2|)}{|\mathcal{X}_{/\mathbf{X}}|} = \lambda_{T_2}^{RL}
\end{aligned}
$$

□

Example 2.5.4. Consider the following statistics:

$$T_{DM} = \frac{1}{n_1} \sum_{i=1}^{n_1} X_{1i} - \frac{1}{n_2} \sum_{i=1}^{n_2} X_{2i}$$

$$T_{DS} = \sum_{i=1}^{n_1} X_{1i} - \sum_{i=1}^{n_2} X_{2i}$$

$$T_M = \frac{1}{n_1} \sum_{i=1}^{n_1} X_{1i}$$

$$T_S = \sum_{i=1}^{n_1} X_{1i}$$

It can be shown that, in a one-sided permutation test, they always yield the same inference. First, we let $k_{\mathcal{X}_{/\mathbf{X}}} = \sum_{i=1}^{n_1} X_{1i}^* + \sum_{i=1}^{n_2} X_{2i}^*$, i.e. the sample sum of an arbitrary $\mathbf{X}^* \in \mathcal{X}_{/\mathbf{X}}$ which is constant given $\mathcal{X}_{/\mathbf{X}}$. Thus, given $\mathcal{X}_{/\mathbf{X}}$:

$$T_{DS} = T_S - (k_{\mathcal{X}_{/\mathbf{X}}} - T_S) = 2T_S - k_{\mathcal{X}_{/\mathbf{X}}}$$

$$T_M = \frac{1}{n_1} T_S$$

$$T_{DM} = \frac{1}{n_1} T_S - \frac{1}{n_2}(k_{\mathcal{X}_{/\mathbf{X}}} - T_S) = \frac{N}{n_1 n_2} T_S - \frac{k_{\mathcal{X}_{/\mathbf{X}}}}{n_2}$$

For every possible $\mathbf{X}$, which proves that T_{DM}, T_{DS}, T_M and T_S are one-to-one increasingly related. Therefore, by Proposition 2.5.1 we know that the one-sided permutation p-values will be the same for all these statistics. Moreover, by Proposition 2.5.3 we have that T_M and T_S yield also the same two-sided permutation p-values.

2.5.3 Computational considerations and p-value estimation

Linear rank statistics (Section 2.4) are an interesting subclass of permutation tests because their null distribution, which corresponds to their permutation distribution, doesn't depend on the observed data. To see this, we note that the observed value of linear rank statistics depends only on the order of the observations, therefore the permutation distribution is completely determined by the set of all possible orderings, which is independent of $\mathbf{X}$. The main consequence of this property of linear rank statistics is that their null/permutation distribution can be calculated (using analytical tools or computers) just one time for each relevant combination of n_1 and n_2.

When the permutation distribution of a test statistic depends on the observed data not only through the ranks, the computational cost for obtaining the null reference distribution and the p-value has to be paid each time a new experiment is conducted. For this reason, before the advent of computers, linear rank statistics were the only permutation tests that could be applied easily in practice. Nowadays, with modern computers, this limitation has been partially overcome in the sense that, for sample sizes of practical interest, the permutation distribution can be

computed very quickly. The total number of permutations of unit labels $(1, ..., N)$ for a pooled sample of size N is $N!$. If the test statistic doesn't depend on the order within group samples, then group labels permutations can be considered instead of unit labels permutations. The total number of possible group label rearrangements for a two-sample experiment is $\binom{N}{n_1} = \frac{N!}{n_1! n_2!}$.

Both with group labels and unit labels, the computational cost for evaluating all possible permutations may become impractical when sample sizes are too large or when a permutation test has to be repeatedly performed within a Monte Carlo framework. To reduce the computational complexity of permutation tests, a common approach is to estimate the permutation p-values of the given statistic T by taking a random sample of arbitrary size from its permutation distribution F_T^*. Algorithm 1 shows a pseudo-code implementation to obtain such a random sample. The algorithm used at line 6 must select permutations of $\mathbf{X}$ by assigning to each possible permutation equal probability. There are several possible implementations of the permutation random sampling algorithm which has linear complexity in the size of $\mathbf{X}$, for instance, the C++ standard library offers the `std::shuffle` procedure and the R `base` package offers the `sample` function.

Algorithm 1 Estimation of the permutation distribution of a given statistic T

Input

$\mathbf{x_1}$: the first observed sample

$\mathbf{x_2}$: the second observed sample

T: the statistic. Takes two numerical array arguments representing the first and the second sample

B: the number of random permutations

Output

$permutedValues$: a random sample of size $nPerm$ from the permutation distribution of T

1: $permutedValues \leftarrow$ an uninitialized array of size B
2: $\mathbf{x} \leftarrow$ concatenation of $\mathbf{x_1}$ with $\mathbf{x_2}$
3: $n_1 \leftarrow size(\mathbf{x_1})$
4: $n_2 \leftarrow size(\mathbf{x_2})$
5: **for** $i = 0$ to $B - 1$ **do**
6:     $\mathbf{x}^* \leftarrow$ A random permutation of $\mathbf{x}$
7:     $\mathbf{x_1^*} \leftarrow$ First n_1 values of $\mathbf{x}^*$
8:     $\mathbf{x_2^*} \leftarrow$ Last n_2 values of $\mathbf{x}^*$
9:     $permutedValues[i] \leftarrow T(\mathbf{x_1^*}, \mathbf{x_2^*})$
10: **return** $permutedValues$

After a random sample Q of size B has been obtained from the permutation distribution of T, permutation p-values can be estimated. For a permutation test with a left tail alternative, the estimator $\hat{\lambda}_T^L$ of λ_T^L calculated using B random permutations is given by:

$$\hat{\lambda}_T^L = \hat{F}_T^*(t) = \frac{\sum_{\mathbf{t}^* \in Q} I(t^* \leq t)}{B} \tag{2.10}$$

Where $\hat{F}_T^*$ denotes the permutation empirical cumulative distribution function (ECDF) which estimates the permutation CDF F_T^*. Similarly, the right tail and

two-sided permutation p-value λ_T^R and λ_T^{LR} can be estimated as:

$$\hat{\lambda}_T^R = \frac{\sum_{t^* \in Q} I(t^* \geq t)}{B} \tag{2.11}$$

$$\hat{\lambda}_T^{LR} = \frac{\sum_{t^* \in Q} I(|t^*| \geq |t|)}{B} \tag{2.12}$$

It follows from the Glivenko–Cantelli theorem, that the estimators $\hat{\lambda}_T^L$, $\hat{\lambda}_T^R$, $\hat{\lambda}_T^{LR}$ and $\hat{F}_T^*$ are strongly consistent [31, p.45].

Figure 2.1 shows the estimated permutation distributions and observed values (vertical blue lines) of $T = \overline{\mathbf{X_1}} - \overline{\mathbf{X_2}}$ in two simulated two-sample location tests. For both tests, the sample sizes are $n_1 = 25$ and $n_2 = 9$. In the first test, both samples have been drawn from Exponential($\lambda = 1$). In the second test, the distribution of the second sample has been shifted to the left of one standard deviation unit. The estimated right tail p-values (ratios between the light blue area and total area) are approximately 0.44 for the first test and 0.002 for the second test. These values suggest that the null should be rejected only in the second case. By simulation design, we know that this is the right decision to take.
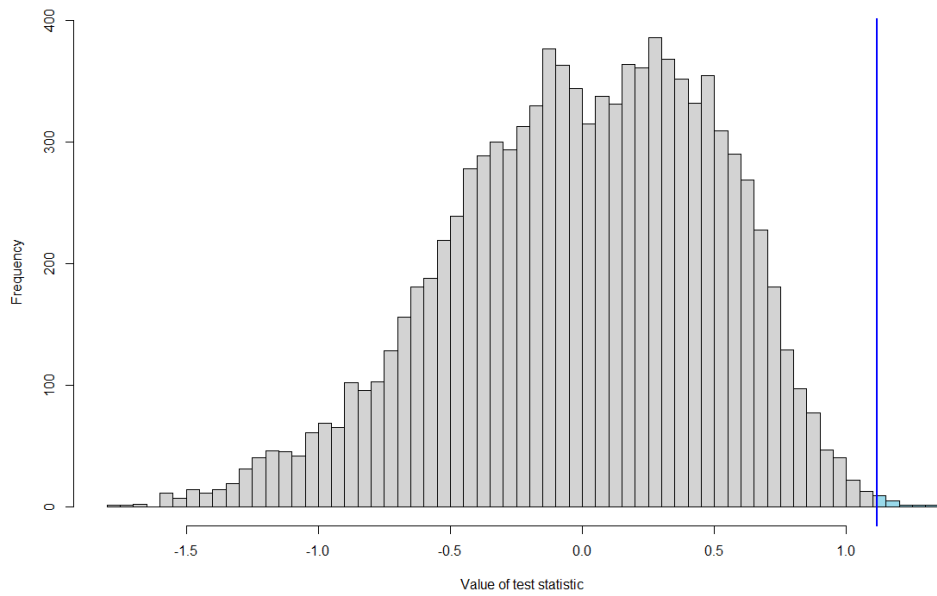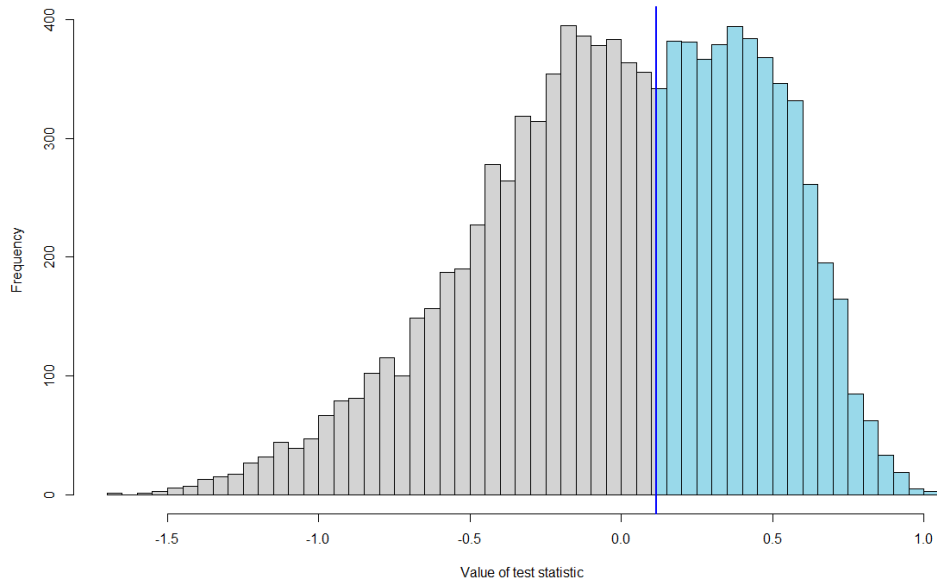
Figure 2.1: Estimated permutation distributions and observed values (vertical blue lines) of the difference of means statistic in two simulated two-sample location tests.

2.5.4 NPC framework

In his *The design of experiments* (1935) [15] Fisher proposed the notion that "different tests of significance are appropriate to test different features of the same null hypothesis". This idea of using several statistical tests within a single statistical problem is very powerful in practice and has many applications like testing elaborate theories in social and medical studies [3].

In the literature can be found three main developments of the general idea of using several tests for testing different "features of the same null hypothesis" proposed by Fisher:

- Multivariate problems: In a multivariate hypothesis testing problem, a feature may be related to the study of a specific response variable or a subset of response variables, see for example [31, p.117] or [23].

- Multiple aspects of a complex hypothesis: When a complex hypothesis can be decomposed into several sub-hypothesis (i.e. a multi-aspect problem), each of the sub-hypothesis can be thought of as a different feature (aspect) of the global hypothesis. For example, the well-known location-scale problem can be decomposed into a location aspect and a scale aspect, and each of these aspects can be addressed by an appropriate statistical test. This is the case of the Lepage test [21], which combines the Wilcoxon test for the location with the Ansari-Bradley test for the scale. The decomposition approach for testing a complex hypothesis is especially useful when a single test for the global hypothesis is not directly available, or difficult to find and/or justify.

- Multiple viewpoints of the same data: A feature of a hypothesis may be also be thought a viewpoint (or aspect) of the data that highlights information useful for detecting shifts towards the alternative under specific classes of distributions. If there is a lack of knowledge regarding the distribution of the data, then it may be useful to combine different viewpoints together in order to obtain a statistical test that exhibits a good power with a large variety of distributions. For the two sample location problem, two concrete applications of this idea will be described in Section 2.6.3 and Section 2.6.4.

The implementation of the idea of using several tests within a single hypothesis testing problem requires a strategy for the combination of chosen tests which takes into account their possible dependency. The Nonparametric Combination of Dependent Tests (NPC) [2] is a framework based on permutation testing which address this problem. NPC offers a template procedure for conducting tests that combine several partial tests, each of which can address a different feature of the same global hypothesis. Being based on permutation testing, NPC does not require distributional assumptions or asymptotic justifications, but only that observations are exchangeable under the null [3].

The NPC algorithm The high level abstract NPC algorithm to obtain the combined test p-value with p partial subtests is the following:

1. Compute the p-dimensional vector of partial test statistics, each one related to the corresponding partial test:

$$(t_1, ..., t_p)$$

---

[2]The NPC framework was originally proposed by Fortunato Pesarin in [29]

2. Recompute each partial test statistic for each of the possible $k = |\mathcal{X}_{/\mathbf{x}}|$ permutations of the data. The entries of the resulting $p \times k$ dimensional matrix are denoted with t^*_{ij} ($i = 1...p$, $j = 1...k$). In each column, the partial test statistics values correspond to the same permutation.

3. Compute the observed value of the combined test as:

$$t_{NPC} = \psi(\lambda_1, ..., \lambda_p)$$

Where ψ is a suitable combining function and λ_i is the permutation p-value associated with partial test i. Permutation p-values can be calculated using the matrix of permuted values obtained from step 2 and the vector of observed values from step 1. It should be noted that each λ_i may be a left, right or two-sided p-value.

4. Compute the permuted values of the combined test as:

$$t^*_{NPCj} = \psi(\lambda_1(t^*_{1j}), ..., \lambda_p(t^*_{pj})), \ j = 1...k$$

Where $\lambda_i(t^*_{ij})$ denotes the permutation p-value of partial test i calculated as if t^*_{ij} was the observed value.

5. Using results from step 3 and 4, compute the combined test permutation p-value λ_{NPC} as:

$$\lambda_{NPC} = F^*_{NPC}(t_{NPC}) = \frac{\sum_{j=1}^{k} t^*_{NPCj} \leq t_{NPC}}{k}$$

Similarly to the permutation testing with a single test statistic, also for NPC tests the p-values can be estimated using an arbitrary number of random permutations. These random permutations should be computed just one time in step 2 to obtain the $p \times k$ dimensional matrix (when p-value estimation is used, in the algorithm above, k becomes equal to the number of random permutations B used for the estimation). Using the estimation procedure just described leads to a strongly consistent estimate of the true combined p-value λ_{NPC} [31, p.127].

In order to obtain a combined statistical test that is consistent and unbiased with respect to the global hypothesis, both the selected partial tests and the combining function have to satisfy some reasonable properties. These properties will not be discussed here because they are quite technical, if the reader is interested in learning more, they can find them explained in [31]. Some frequently used combing functions that satisfy the aforementioned properties are the following:

- Liptak's normal combining function

$$\psi_L(\lambda_1, ..., \lambda_p) = \sum_{i=1}^{p} \phi(\lambda_i) \tag{2.13}$$

- Tippet's combination

$$\psi_T(\lambda_1, ..., \lambda_p) = min_{i=1...p}(\lambda_i) \tag{2.14}$$

- Fisher's combination

$$\psi_F(\lambda_1, ..., \lambda_p) = \sum_{i=1}^{p} log(\lambda_i) \tag{2.15}$$

Where ϕ denotes the CDF of the standard normal distribution.

In multi-aspect problems, the Liptak combining function has the greatest relative power when all sub-alternatives are jointly true. Instead, the Tippet's combination has the greatest relative power when only one of the sub-alternatives is true. The Fisher combining function has an intermediate behavior and then it is useful when nothing is expected about the partial alternative hypotheses.

2.6   Non-parametric tests for location

In this section are presented some permutation tests for the two-sample location problem (Section 2.2) which have inspired the control charts discussed in Chapter 4.

2.6.1   Wilcoxon's rank sum test

In order to address two sample location problem , in [35, p.80] (1945), Wilcoxon proposed to use ranks of the observations instead of numerical data. Based on this idea, the Wilcoxon's rank sum statistic W is defined as the sum of the ranks of the first sample $\mathbf{X_1}$. In mathematical language, W can be expressed as:

$$W = \sum_{i=1}^{n_1} R_i$$

Where R_i is the rank of the ith observation in the pooled sample. Equivalently, using its linear rank statistic representation, W is given by:

$$W = \sum_{i=1}^{N} i Z_i$$

Assuming no ties and that the null is true, the mean and variance of W are given by:

$$E(W) = \frac{n_1(N+1)}{2} \quad \mathrm{Var}(W) = \frac{n_1 n_2 (N+1)}{12}$$

It can be shown that W (the sum of the ranks of $\mathbf{X_1}$) is permutationally equivalent to the sum of ranks of $\mathbf{X_2}$, to the difference between the sum of ranks of $\mathbf{X_1}$ and $\mathbf{X_2}$, and to the Mann-Whitney U statistic given by:

$$U = W - n_1(n_1 - 1)/2$$

2.6.2   Van der Waerden test

To increase the power with normally distributed data with the respect to the Wilcoxon rank sum test, the van der Waerden test [34] (1953) applies the quantile function of the standard normal distribution ϕ^{-1} to normalized ranks:

$$K = \sum_{i=1}^{n_1} \phi^{-1}\left(\frac{R_i}{N+1}\right)$$

Equivalently, using the linear rank representation, the van der Warden test statistic can be written as:

$$K = \sum_{i=1}^{N} \phi^{-1}\left(\frac{i}{N+1}\right) Z_i$$

The van der Warden test is asymptotically optimal when the populations are both normal distributions. Indeed, its ARE relative to the Student's t test is 1 under the normality assumption [17, p.302].

The null distribution of the van der Warden statistic is symmetric about 0.

2.6.3 Bi-aspect test

Marozzi in [22] proposed a bi-aspect permutation test for the two sample location problem which, using the NPC approach with the Tippet combining function (Section 2.5.4), combines the permutation test based on the sample total T_a with a permutation test based on the comparison of sample observations with the pooled median $\tilde{M}$, T_b:

$$T_a = \sum_{i=1}^{n_1} X_i \quad T_b = \sum_{i=1}^{n_1} I(X_i > \tilde{M})$$

The main goal adressed by the author was to propose a test with good overall performance that may be used without worrying about the distribution underlying the data which is usually unknown. The choice of using the Tippet combining function is important because the stand alone tests show different sensitiveness to location differences under different types of underlying population distributions.

Marozzi in his paper presents also a study based on Monte Carlo simulations which shows that the bi aspect T_{ab} performs similarly to the traditional permutation test based on sample total when the latter is a proper statistic to study the location. Considering instead distributions for which the sample total is not a good statistic for detecting location shifts, like heavy tailed distributions, the author shows that the power of T_{ab} is remarkably higher.

For the methods that will be proposed in Chapter 4, it is important to consider that T_b is a linear rank statistics whereas T_a is not. The linear rank representation of T_b is given by:

$$T_b = \sum_{i=1}^{N} I\left(i \geq \left\lceil \frac{N}{2} \right\rceil\right) Z_i$$

2.6.4 Tri-aspect test

In [23], Marozzi proposed a new combined permutation test for the two sample location problem which extends the idea of the bi-aspect test T_{ab} by including a third aspect in the test, the rank aspect T_c. The rationale for including also T_c in the combined test is to improve the performance of the T_{ab} test, similarly to how T_b has been introduced in T_{ab} in order to overcome the drawbacks of T_a. Using a Monte Carlo simulations, Marozzi showed that the tri-aspect test T_{abc} is more powerful than the bi-aspect test with a large variety of distributions including the Normal, Cauchy and Exponential distributions.

The test statistic of the tri-aspect test can be thought of as a function of the original dataset X and two additional datasets Z and W obtained with the following data augmentation strategy:

$$Z_i = I(X_i > \tilde{M})$$

$$R_i = \sum_{j=1}^{n} I(X_i \geq X_j)$$

The three aspects (or data viewpoints) of the test can then be defined as follows:

$$T_a = \sum_{i=1}^{n_1} X_i$$

$$T_b = \sum_{i=1}^{n_1} Z_i = \sum_{i=1}^{n_1} I(X_i > \tilde{M})$$

$$T_c = \sum_{i=1}^{n_1} R_i = \sum_{i=1}^{n_1} \sum_{j=1}^{n} I(X_i \geq X_j)$$

T_b is the number of observations in the first sample greater than the pooled median while T_c is the sum of the ranks W_i of the first sample. In order to combine the three aspects into a single tri-aspect test statistic T_{abc}, the Tippett combining function is used.

2.6.5 Percentile modified rank tests

In [16] (1965) Gastwirth proposed percentile modified rank tests, a class of linear rank statistics which provides a way to increase the ARE relative to popular nonparametric tests under specific distributions without using complicated scoring systems. The main idea is to select two number s and r, both between 0 and 1, and then score only the data in the upper sth and lower rth percentiles of the combined sample. We let $R = \lfloor Nr \rfloor$ and $S = \lfloor Ns \rfloor$ and we define:

$$B_r = \sum_{i=1}^{R} (R - i + 1) Z_i$$

$$T_r = \sum_{i=N-S+1}^{N} [i - (N - S)] Z_i$$

If N is odd, and:

$$B_r = \sum_{i=1}^{R} \left[R - i + \frac{1}{2} \right] Z_i$$

$$T_r = \sum_{i=N-S+1}^{N} \left[i - (N - S) - \frac{1}{2} \right] Z_i$$

If N is even. Then, $T_s - B_r$ provides a test for location and $T_s + B_r$ provides a test for scale [17, p. 305].

An interesting characteristic of percentile modified rank tests is that a good choice of s and r may lead to a test with an higher power than more popular alternatives. For example, the ARE of $T_s - B_r$ under a normal distribution reaches its maximum value of 0.968 when $s = r = 0.42$, while, when $s = r = 0.5$, the ARE is 0.955, as for the Wilcoxon rank-sum statistic. Clearly, under the assumption of an unknown distribution, we usually cannot find the optimal s and r values. Nevertheless, following the idea described in Section 2.6.3 and 2.6.4, a percentile modified rank test can be thought of as a data viewpoint, and the NPC framework can then be used to combine multiple viewpoints in order to obtain a test with a good power under a wide variety of distributions.

2.7 Non-parametric tests for scale

In this section are presented some permutation tests for the two-sample scale problem (Section 2.2) which have inspired control charts discussed in Chapter 4.

2.7.1 Klotz normal-scores test

The Klotz (1962) normal-scores test [18] for scale uses as test statistic:

$$K = \sum_{i=1}^{n_1} \left[\phi^{-1} \left(\frac{R_i}{N+1} \right) \right]^2$$

where ϕ^{-1} denotes the quantile function of the standard normal distribution. In its equivalent linear rank representation, the Klotz test statistic is given by:

$$K = \sum_{i=1}^{N} \left[\phi^{-1} \left(\frac{i}{N+1} \right) \right]^2 Z_i$$

The ARE of the Klotz test relative to the F test equals 1 when both populations are normal. Therefore, since the F test is an optimal test assuming a normal distribution, the Klotz is an aymptotically optimal test for normal data.

Chapter 3

Theory: Statistical Process Control and Control Charts

This Chapter covers the theoretical background of Statistical Process Control and control charts required to understand the notions and methods proposed in Chapter 4.

3.1 Statistical Process Control

One fundamental aspect of the quality of a product is the conformance of the product's features to the specifications required by the design [26, p.6]. In other words, the quality of products is negatively affected by variability in their characteristics. Within any production process, regardless of how well the process was designed, developed and maintained, a certain amount of natural variability will always be present [26, p.181]. This variability, which is also called *background noise*, is intrinsic in the design of the process and has *chance causes* that cannot be controlled without performing significant changes to the process itself [36].

Besides chance causes, other sources of variation may affect the process output. These sources of variation that are not part of the chance cause pattern are called *assignable causes* of variation. Assignable causes are specific and controllable causes and their effect is generally large when compared to the background noise. Therefore, variation with assignable causes typically represents an unacceptable level of process performance. Within manufacturing processes, some examples of assignable causes are defective production materials, operator errors and malfunctions of the machinery.

In the framework of statistical quality control (SPS), a process is considered in a state of statistical control or *in-control* (IC) if it is operating with only chance causes of variation. Conversely, if the output of the process is also affected by assignable causes of variation, then the process is considered *out-of-control* (OOC). A primary objective in statistical process control (SPC) is to quickly detect when a process becomes out-of-control. By signaling rapidly assignable causes of variation, the responsible engineers and technicians can identify the causes and take corrective actions before many non-conforming units are produced. Quick detection helps also analyze sources of variation and their effects. If the source has a negative effect, measures should be taken to prevent the cause from reoccurring. However, if the effect is considered positive, its source could be instead integrated into the process [19, p.5].

Figure 3.3 shows a visualization of chance and assignable causes of variation. Until time t_2, the process shown in the Figure is in-control. The chance causes of variation are represented by a fixed PDF with mean equal to μ_0 and standard deviation equal to σ_0. Note that, when the process is in-control, it is unlikely to observe a product feature value falling outside the upper and lower specifications limits (USL and LCL, respectively). At time t_3, the distribution of the process shifts to the right ($\mu > \mu_0$). This means that an assignable cause of variation is now present in the process and that it has become more likely to observe products that do not conform to the specification. After returning in-control at time t_4, the process is affected by a new cause of variation at time t_5. This time, however, we observe not only a shift to the right ($\mu > \mu_0$), but also an enlargement of the scale ($\sigma > \sigma_0$).
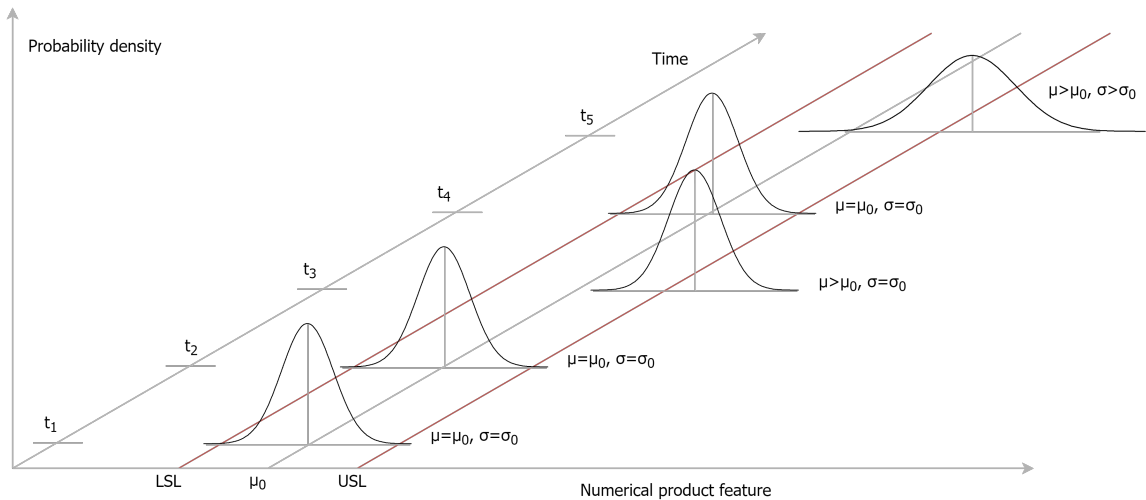
Figure 3.1: A visualization of chance and assignable causes of variation.

3.1.1   Phase I and Phase II

The application of SPC has two main phases: *Phase I* and *Phase II*.

The main goal of Phase I is to obtain a sufficiently large *reference sample* gathered from the process while it is in-control. To obtain the reference sample, data is collected from the process while it is actively monitored to ensure that no assignable causes of variation are present. After gathering the data, Phase I statistical methods can be used to analyze the sample to further ensure that the process has been in-control over the period where the data was collected.

In Phase II, the available information about the IC distribution is used for online monitoring of the process with the main goal of quickly detecting when it becomes out-of-control. Knowledge about the IC distribution may take the form of a reference sample gathered in Phase I or/and may be the result of prior knowledge in the application/domain of interest. An important classification for Phase II scenarios depends on whether the parameters of the IC distribution[1] are known, i.e. the *standards known case* (case K), or unknown, i.e. the *standards unknown case* (Case U) [5]. In case U, the distribution parameters required by the chosen Phase II method are typically estimated from a Phase I sample [26, p. 228]. In

---

[1]Typically, the mean and standard deviation

case K, Phase I is not needed if the known distribution parameters are all the required process information.

3.1.2   Phase II statistical model

After Phase I, suppose that a random sample of size m, denoted with $\mathbf{X_1} = (X_{11}, ..., X_{1n_1})$, has been drawn from the in-control CDF of the process $F_1(x)$. When Phase II starts, a sequence of *test samples* is gathered from the process. Assuming that test samples have all the same size n, let $\mathbf{X_{2h}} = (X_{21h}, ..., X_{2n_2h})$ denote the hth test random sample with underlying cdf $F_{2h}(x)$. The symbols X_1 and X_{2h} will be used to denote arbitrary observations of the reference sample and the test sample at step h, respectively. In its most general form, the Phase II problem consists then in finding distributional deviations between F_1 and F_{2h}.

It is important to consider that, under this general Phase II statistical model, random sampling and independence between the samples are assumed, however, no distributional assumptions are made for F_1 and F_{2h}.

3.1.3   Out-of-control patterns

In practical applications, some kinds of distributional deviations are more important than others. Monitoring changes in the location and in the scale of the process distribution F_1 are typically the two main concerns in Phase II applications [27]. These distributional aspects can be characterized using a suitable distribution parameter, like the mean or median for the location and the standard deviation for the scale. Changes in the process distribution parameters can then be categorized into different common *OOC patterns*:

- *Sustained shift*: occurs when, at some point in time, the parameter being monitored shifts to a new value and remains there.

- *Trend*: happens when the parameter continuously increases or decreases.

- *Cyclic pattern*: occurs when the process parameter changes following a periodic behavior like, for instance, a sinusoidal function.

Figure 3.2 provides some examples of the OOC patterns discussed so far: Figure 3.2a shows a positive sustained shifts of the scale, Figure 3.2b shows an increasing trend in the location and Figure 3.2c displays a cyclic pattern in the location.

Phase II SPC methods usually focus on specific kinds of distributional aspects. For example, some methods are designed to detect changes in location and others in scale. Others more are designed to detect changes in location and scale jointly. Moreover, Phase II methods may also focus on a specific OOC pattern or a group of OOC patterns.
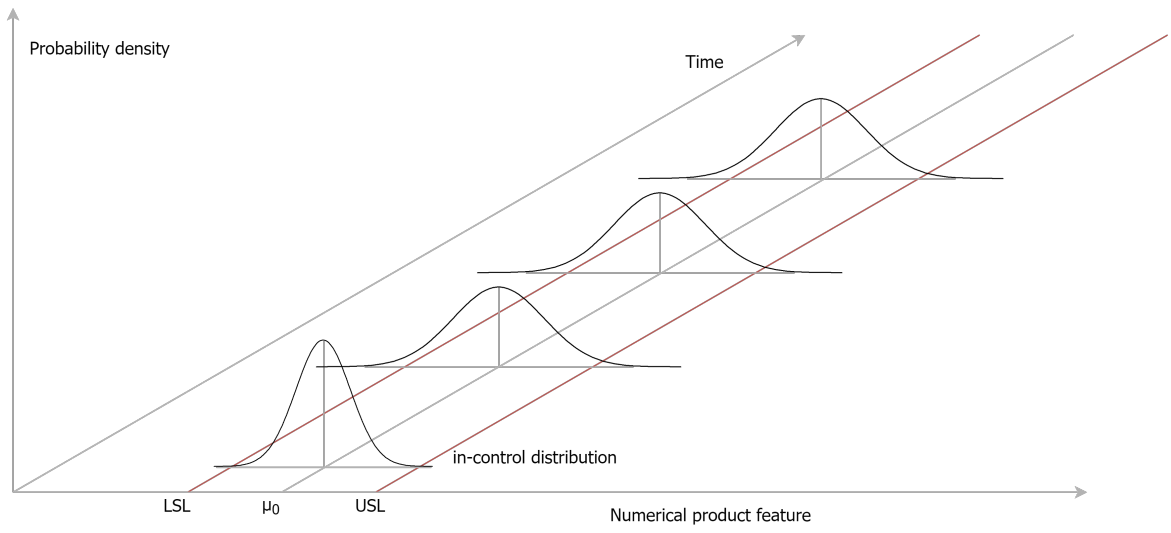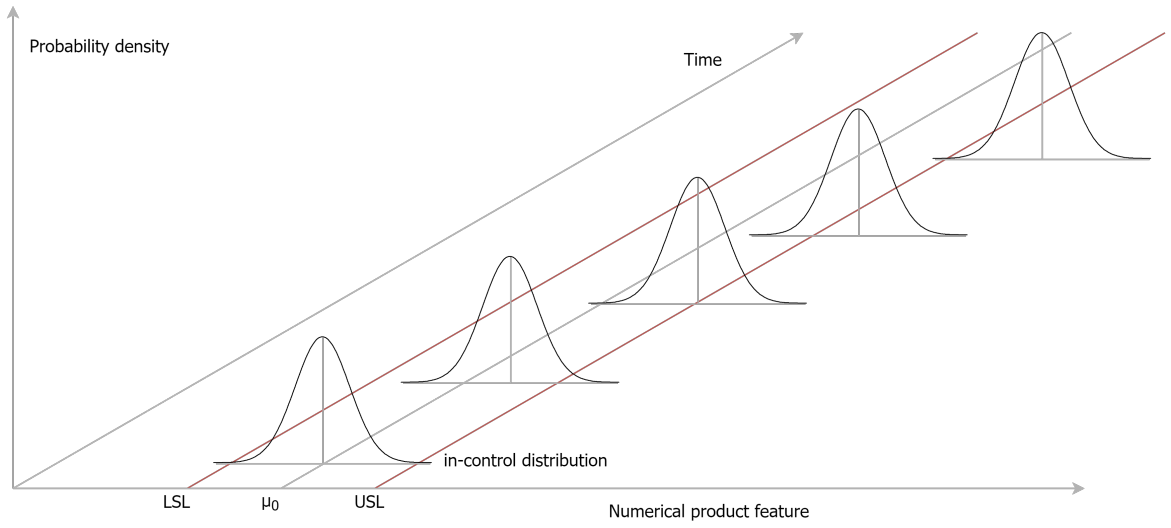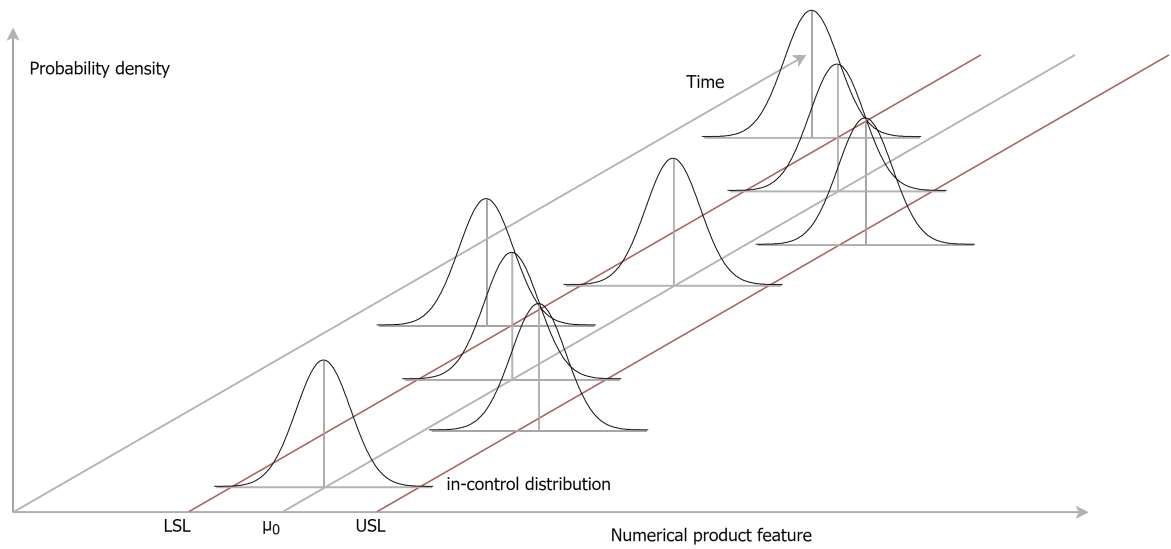
Figure 3.2: (a) A positive sustained shift of the scale. (b) An increasing trend in the location. (c) A cyclic pattern in the location.

3.2   Control Charts

Control charts are an important SPC tool used both in Phase II and in Phase I
[26]). In Phase I they are used to verify if the process has been in control during
the time period in which the data were collected. In Phase II they are used as
monitoring devices with the primary goal of detecting as quickly as possible when
an assignable cause of variation occurs.

The basic building blocks of a control chart are the following:

- A *monitoring statistic* which is plotted versus the sample number or time.

- An *upper and/or lower control limit* (UCL and LCL) representing the bound-
  ary inside which nearly all the points should fall between when the process
  is in control. If a point is found outside of the control limits is interpreted
  as evidence that the process is out of control.

- A *center line* representing the average of the monitoring statistic when the
  process is in-control. This component, unlike the previous two, is not always
  used.

- *Warning limits* complementary to the control limits. Warning limits are
  used to establish fine granular thresholds associated with different levels of
  evidence against the IC state. As the center line, also the warning limits are
  optional.
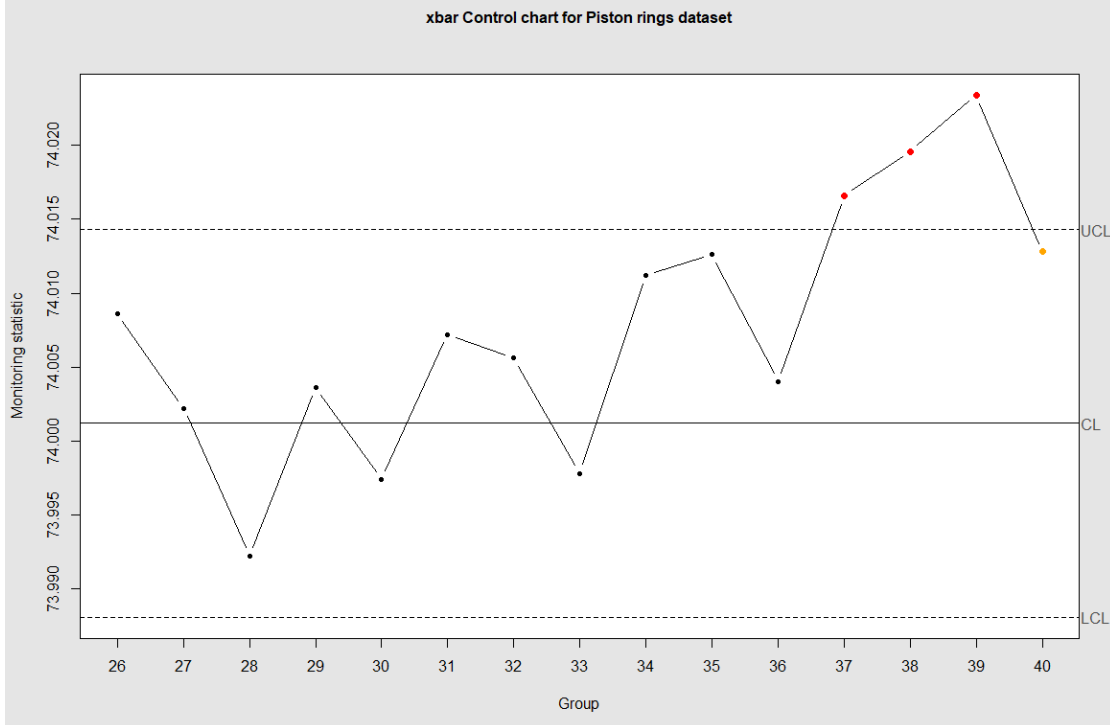
Figure 3.3: A visualization of a control chart. The small dots displays the monitor-
ing statistics values changing over the test samples. The dashed lines correspond
to the upper and lower control limits and the straight line corresponds the center
line.

Figure 3.3 shows an example control chart using a center line and no warning limits. The plotted dots represents the monitoring statistic values at consecutive time steps. The dashed lines correspond to the upper and lower control limits and the solid line is the center line which represents the average value of the monitoring statistic. The control chart shows that, from step 37 to 39, it is likely that the process has gotten out of control.

Based on these building blocks discussed so far, several rules for generating *OOC signals* can be devised. The base signaling rule used by all control charts is emitting an OOC signal whenever the monitoring statistics exceed the control limits (i.e. the monitoring statistic value is lower than LCL or is greater than UCL). This general scheme is useful to detect all the OOC patterns discussed in Section 3.1.3. Additionally to the base signaling rule, other more complex rules can be used to detect faster specific kinds of OOC patterns and to improve the effectiveness of the control chart. For example, the rule of emitting a signal when there is a high number of points in a row steadily decreasing or increasing can be used to detect trend patterns earlier.

In Phase II, when an OOC signal is issued, the process is typically stopped and an investigation begins with the goal of detecting the possible assignable cause. If an assignable cause is detected then measures should be taken in order to eliminate the assignable cause or integrate it in the process if the effect is considered to be positive. There may be situations in which after an OOC signal no assignable causes are detected by operators and engineers. Such situations may arise as a consequence of randomness and their frequency is typically controlled through control chart parameters.

3.2.1 Types of Control Charts

An important classification for control charts is made according to the structure of the monitoring statistic and the control limits design. Concerning this classification, the three main groups are the following:

- *Shewhart-type* control charts, first introduced by W. A. Shewhart in [32] (1926). In Shewhart-type control charts, the monitoring statistic uses the information contained in just the last sample.

- *Cumulative sum* (CUSUM) control charts, introduced by E. S. Page in [28] (1954). The monitoring statistic for this kind of control chart incorporates information contained in both present and past samples using an autoregressive formula.

- *Exponentially weighted moving average* (EWMA) control charts, introduced by S. Roberts in 1959. Following similar principles of CUSUM control charts, also EWMA charts exploit the information contained in both past and present samples.

With respect to CUSUM and EWMA charts, Shewhart-type charts are typically less effective in detecting small sustained shifts in the location and scale parameters of the process distribution. On the other hand, the simplicity of Shewhart control charts makes them easier to implement and interpret. Furthermore, starting from a Shewhart-type is often possible to design its CUSUM or EWMA counterparts.

3.2.2 Performance evaluation

Phase II performance of control charts is often characterized using the average of the run length distribution, i.e. the *average run length* (ARL). The *run length* is the number of test samples that are collected before the first OOC signal is issued [4, p.40]. Ideally, the ARL when the process is OOC (ARL_1), should be as low as possible to increase the probability of quickly identifying OOC conditions. Instead, the ARL when the process is IC (ARL$_0$), can be used as a calibration parameter of the control chart which regulates the control limits. Using an analogy with the hypothesis testing framework, the ARL_0 has the same role of the significance level while ARL_1 has the same role of the test power.

The run length distribution and the ARL of control charts are typically evaluated under a statistical model which extends the general Phase II model described in Section 3.1.2. Two popular models used for evaluating performances under the sustained shift OOC pattern (Section 3.1.3) are described below. These models should not be seen as a set of additional Phase II assumptions. Instead, they should just be considered as a mathematical description for typical OOC scenarios under which the performances of control charts can be evaluated using mathematical tools or computer simulations.

Location model This model assumes that the in-control distribution $F_1(x)$ and the test sample distributions $F_{2h}(x)$ may differ only through a location shift θ. This means that the true process distribution can be shifted to the left or to the right with respect to the IC distribution, but its shape it's assumed to be the same. In mathematical terms, this assumption can be expressed as:

$$F_{2h}(x) = F_1(x - \theta) \quad \theta \in \mathbb{R} \quad \forall h$$

Where θ is the unknown location shift parameter. Under the location model, the process distribution F_{2h} is considered fixed through time, thus the subscript h can be omitted. Moreover, it can be seen that this model is very similar to the hypothesis testing location model described in Section 2.2, the only difference is that in the SPS Phase II location model we have a sequence of samples for the second group. Equivalently, the Phase II location model can be written as:

$$F_1(x) = F(x - \mu_1) \quad \mu_1 \in \mathbb{R}$$
$$F_2(x) = F(x - \mu_2) \quad \mu_2 \in \mathbb{R}$$

Where F is a cdf with location equal to 0, μ_1 is the location parameter of the IC process and μ_2 is the location parameter of the actual process distribution.

Location-scale model A generalization of the location model that considers also a possible change in the scale is the so-called location-scale model. Under this model, it is assumed that the actual process distribution may differ from the IC distribution only through a location shift or a scale enlargement/compression. In mathematical language, this can be written as:

$$F_2(x) = F_1\left(\frac{x - \theta}{\delta}\right) \quad \theta \in \mathbb{R} \quad \delta \in \mathbb{R}$$

Which means that X_2 is distributed as $(X_1 \cdot \delta) + \theta$ and that the process is IC if and only if $\theta = 0$ and $\delta = 1$. An alternative parametrization for the same model

is as follows:

$$F_1(x) = F\left(\frac{x - \mu_1}{\sigma_1}\right) \quad \mu_1 \in \mathbb{R} \quad \sigma_1 \in \mathbb{R}$$

$$F_2(x) = F\left(\frac{x - \mu_2}{\sigma_2}\right) \quad \mu_2 \in \mathbb{R} \quad \sigma_2 \in \mathbb{R}$$

Where F is a CDF with location equal to 0 and scale equal to 1, μ_1 and σ_1 are the unknown location and scale parameters of the IC distribution and μ_2 and σ_2 are the location and scale parameters of the actual process distribution.

Assuming the location-scale model, let S_h denote the monitoring statistic random variable of a Shewhart-type chart at step h and let R denote its run length random variable when just the base signaling rule is used. In the general case, S_h $h = 1, ...$ are not independent, however, given the reference sample $\mathbf{X_1}$, they become independent [9]. Moreover, since under the location-scale model the actual process distribution F_2 is constant over time, S_h $h = 1, ...$ becomes also identically distributed and the symbol S can be used to denote any of these random variables. From these observations we can conclude that, given $\mathbf{X_1}$, R follows a geometric distribution with success probability equal to the conditional signal probability $p(\mathbf{X_1}) = P(S > \text{UCL} \cup S < \text{LCL}|\mathbf{X_1})$, hence:

$$P(R = r|\mathbf{X_1}) = [1 - p(\mathbf{X_1})]^{r-1}p(\mathbf{X_1}) \tag{3.1}$$

$$E(R|\mathbf{X_1}) = \frac{1}{p(\mathbf{X_1})} \tag{3.2}$$

$$\text{Var}(R|\mathbf{X_1}) = \frac{1 - p(\mathbf{X_1})}{p(\mathbf{X_1})^2} \tag{3.3}$$

The distribution of R conditional to $\mathbf{X_1}$ is often referred to as the conditional run length distribution and the conditional expectation of R given $\mathbf{X_1}$ is also known as *conditional average run length* and will be denoted with $\text{CARL}(\mathbf{X_1})$.

In order to obtain the unconditional run length distribution, its average and variance, the laws of total probability, total expectation and total variance can be applied. If we assume that the in-control process admits a probability density function[2] $f_1(x)$, then:

---

[2]Using measure theory and in particular the Radon–Nikodym theorem, we know that a probability density function can be defined if and only if the underlying probability measure is absolutely continuous with respect to the Lebesgue measure.

$$P_R(r) = \int_{-\infty}^{+\infty} ... \int_{-\infty}^{+\infty} P(R = r | \mathbf{X_1} = \mathbf{x_1}) \, f_1(x_{11})dx_{11} \, ... \, f_1(x_{1m})dx_{1m} \quad (3.4)$$

$$= \int_{-\infty}^{+\infty} ... \int_{-\infty}^{+\infty} [1 - p(\mathbf{x_1})]^{r-1} p(\mathbf{x_1}) \, f_1(x_{11})dx_{11} \, ... \, f_1(x_{1m})dx_{1m}$$

$$E(R) = \int_{-\infty}^{+\infty} ... \int_{-\infty}^{+\infty} E(R | \mathbf{X_1} = \mathbf{x_1}) \, f_1(x_{11})dx_{11} \, ... \, f_1(x_{1m})dx_{1m} \quad (3.5)$$

$$= \int_{-\infty}^{+\infty} ... \int_{-\infty}^{+\infty} \frac{1}{p(\mathbf{x_1})} \, f_1(x_{11})dx_{11} \, ... \, f_1(x_{1m})dx_{1m}$$

$$\mathrm{Var}(R) = E(\mathrm{Var}(R|\mathbf{X_1})) + \mathrm{Var}(\mathrm{E}(R|\mathbf{X_1})) \quad (3.6)$$

$E(R)$ is the *unconditional average run length*[3] (ARL) and $\sqrt{\mathrm{Var}(R)}$ is the *unconditional run length standard deviation* (SDRL).

Both the conditional and the unconditional run length distributions can be useful in order to assess the performance of control charts. The conditional distribution might be preferred from a practical point of view since users would most likely assess the performance without "averaging" over all possible reference samples [6]. Instead, they typically want to judge a method based on its performance with the particular reference sample in use. On the other hand, the unconditional distribution is useful for comparing different charts under specific theoretical process distributions (Normal, Laplace, Gamma...). In this kind of comparison, as will be discussed in more detail later on, it is important to study also how the run length distribution is affected by the variability within the reference sample.

Estimating run length distribution using Monte Carlo simulations

Equations 3.4 and 3.5 may be difficult to use in practice because of the m-dimesional integrals involved. Therefore, in order to estimate the run length distribution, Monte Carlo simulations are often used instead. Algorithm 2 shows an implementation of this idea. In each simulation, the reference sample of size m is generated from the IC distribution F_1. Then, a sequence of test samples of size n is generated from the actual process distribution $F_2(x) = F_1\left(\frac{x-\theta}{\delta}\right)$, until the first OOC signal occurs. The number of samples generated before the OOC signal is the simulated run length. By averaging between a large number of simulated run lengths the Monte Carlo estimate of the ARL is obtained.

---

[3]The adjective "unconditional" is typically omitted.

Algorithm 2 Monte Carlo estimation of the run length for a Stewhart-type chart

Input

$nSim$: the number of simulations

F_1: the in-control distribution of the process

T: the monitoring statistic. This function should take as input a reference and a test sample and should produce a real number

UCL, LCL: the upper and lower control limits

n_1, n_2: the reference and test samples sizes

θ, δ: regulates together the location shift and scale enlargement/compression of the actual process distribution. $\theta = 0$ and $\delta = 1$ corresponds to an IC process

$runLengthCap$: a maximum limit for the run length in each simulation

Output

$runLengthArray$: an array of length $nSim$ containing the simulated run lengths

1: $runLengthArray \leftarrow$ an uninitialized array of size $nSim$
2: **for** $i = 0$ to $nSim - 1$ **do**
3:     $\mathbf{x_1} \leftarrow$ a random sample of size n_1 from F_1
4:     $runLength \leftarrow 0$
5:     **repeat**
6:         $runLength \leftarrow runLength + 1$
7:         $\mathbf{x_2} \leftarrow$ a random sample of size n_2 from F_1
8:         $\mathbf{x_2} \leftarrow (\mathbf{x_2} \cdot \delta) + \theta$
9:         $stat \leftarrow T(\mathbf{x_1}, \mathbf{x_2})$
10:     **until** $stat > UCL$ **or** $stat < LCL$ **or** $runLength = runLengthCap$
11:     $runLengthArray[i] \leftarrow runLength$
12: **return** $runLengthArray$

3.2.3   Calibration

Control charts' control limits are typically decided as a function of a desired (target) ARL_0 [6]. The problem of searching the control limits corresponding to a given ARL_0 is also known as control chart *calibration*. In order to calibrate a Shewhart-type control chart, one possible idea is solving Equation 3.5 for UCL and/or LCL with a given ARL_0, however, this strategy is not easy to follow in practice since many integrals are involved [27]. Another possible approach is based on the combination of Monte Carlo simulations with interpolation techniques. This idea is simple to implement using the computational power of modern computers and requires just a fast algorithm for the calculation of the monitoring statistic and a reasonable interpolation strategy. Algorithm 3 and Algorithm 4 illustrates the first and the second part respectively of a possible pseudo-code implementation. This particular implementation assumes a single limit (LCL) and provides only a point estimate for it:

1. Algorithm 3 (first part - simulation): takes as input the IC process distribution F_1, the sample sizes n_1 and n_2, the monitoring statistic T and a sequence of lower control limits $lclSeq$ for which the run length should be recorded in each simulation. This algorithm then simulates the IC process and produces a matrix of run lengths of size $nSim \times length(lclSeq)$. $lclSeq$ is a parameter that is difficult to adjust and typically requires some knowledge of the monitoring statistic distribution or just some trial and error.

2. Algorithm 4 (second part - interpolation): takes as input the result of Algorithm 3, the same sequence of lower control limits $lclSeq$ and a target ARL_0 value. Algorithm 4 then calculates the ARL associated to each control limit and use an interpolation strategy in order to find the control limit corresponding to the target ARL_0 value. Figure 3.4 shows an example of the result of this second part of the calibration (note that a UCL is used instead of a LCL).

Algorithm 3 First part of the unconditional calibration procedure for Stewhart-type control charts using only a LCL

Input

$nSim$: the number of simulations

F_1: the in-control distribution of the process

T: the monitoring statistic. This function should take as input a reference and a test sample and should produce a single real number

n_1, n_2: the reference and test samples sizes

$lclSeq$: the sequence of lower control limits for which the run length should be recorded in each simulation. Values should be sorted in decreasing order

$runLengthCap$: a maximum limit for the run length in each simulation

Output

$resultMatrix$ a matrix of size $nSim \times length(lclSeq)$ with the run lengths associated to each limit in $lclSeq$ obtained by simulation

1:   $resultMatrix \leftarrow$ an uninitialized matrix of size $nSim \times length(lclSeq)$

2:   **for** $simIdx = 0$ to $nSim - 1$ **do**

3:      $\mathbf{x_1} \leftarrow$ a random sample of size n_1 from F_1

4:      $runLength \leftarrow 0$

5:      $lclIdx \leftarrow 0$

6:      **while** true **do**

7:         $runLength \leftarrow runLength + 1$

8:         $\mathbf{x_2} \leftarrow$ a random sample of size n_2 from F_1

9:         $stat \leftarrow T(\mathbf{x_1}, \mathbf{x_2})$

10:        **while** $lclIdx < length(lclSeq)$ & $stat \leq lclSeq[lclIdx]$ **do**

11:           $resultMatrix[simIdx, lclIdx] \leftarrow runLength$

12:           $lclIdx \leftarrow lclIdx + 1$

13:        **if** $lclIdx = length(lclSeq)$ **then**

14:           **break**

15:        **if** $runLength = runLengthCap$ **then**

16:           **while** $lclIdx < length(lclSeq)$ **do**

17:              $resultMatrix[simIdx, lclIdx] \leftarrow runLengthCap$

18:              $lclIdx \leftarrow lclIdx + 1$

19:           **break**

20: **return** $resultMatrix$

Algorithm 4 Second part of the unconditional calibration procedure for Stewhart-type control charts using only a LCL

Input

$runLengthMatrix$ the output of the first part of the procedure (Algorithm 3)
$lclSeq$: the sequence of lower control limits used in in the first part of the procedure (Algorithm 3)
$targetArl$: the nominal ARL_0

Output

LCL : the lower control limit

   1: $arls \leftarrow rowMeans(runLengthMatrix)$ ▷ ARL_0 associated with each control limit
   2: $LCL \leftarrow interpolate(arls, lclSeq, targetArl)$
   3: **return** LCL
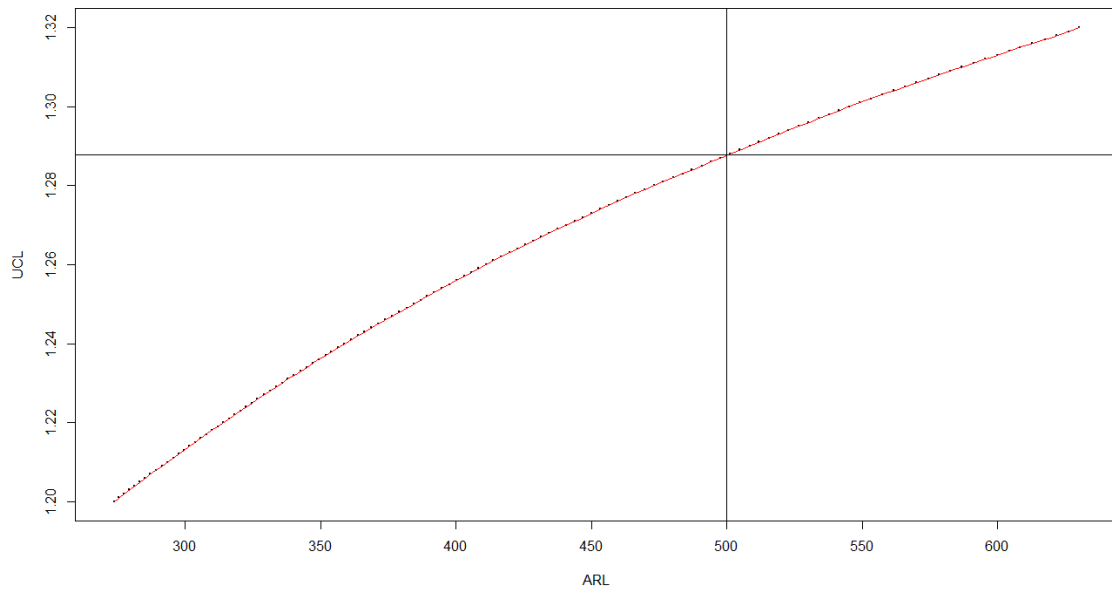
Figure 3.4: Scatter plot of the estimated ARL_0 versus the corresponding UCL for a $\overline{X}$ control chart and a standard normal distribution. The red curve corresponds to a linear interpolation of the points. The horizontal line represents the UCL value associated with $ARL_0 = 500$ (vertical line).

Effect of the variability within the reference sample   It is important to observe that the proposed calibration problem is not conditional to a specific reference sample. A relevant consequence is that, after the calibration, the ARL_0 can vary substantially depending on the reference sample in use. In practical applications, this variability can constitute a problem since the actual ARL_0 shouldn't be too distant from the nominal value. Therefore, when a control chart is analyzed, it is important to study also the conditional run length distribution. This can be done for instance using the approach described in [27], which will also be applied in this Thesis to analyze the control charts presented in Chapter 4.

Effect of the process distribution   Another important aspect of the calibration problem is that the ARL_0 can be affected by the shape of the underlying distribution. Consider for example a control chart that has been calibrated for $\mathrm{ARL}_0 = 500$ under the standard normal distribution. If the true IC process distribution is not standard normal, then the true ARL_0 may be very distant from the target of 500. When there is limited or lack of knowledge about the underlying process distribution, it is desirable to have control charts whose IC performance does not depend on the shape of the distribution [6] [4, p. 34]. These control charts are called *distributon-free control charts* [7, p.305] [8, p.445]. The concept of distribution-free charts is very close to that of non-parametric tests: they are both devised to get rid of questionable distributional assumptions but in two different contexts.

Chapter 4

Multi-aspect NPC control charts

In this Chapter, we describe the multi-aspect NPC scheme for control charts and analyze some practical applications.

Starting by introducing the general concept of multi-aspect NPC charts, we will then describe our software implementation that helped us answer some important research questions concerning these charts. We end up with the presentation and performance analysis of some multi-aspect NPC charts for the location and joint location-scale monitoring problems.

4.1   General concept

This framework described in Section 2.5.4 suggests a Shewhart-type charting procedure where a sequence of dependent two-samples NPC tests is considered. The adopted general statistical model and notation is described in Section 3.1.2. Then, the monitoring statistic at step h for a NPC control chart is given by:

$$t_h = \psi(\lambda_{1h}, ..., \lambda_{ph}) \tag{4.1}$$

Which is the same quantity computed at step 3 of the NPC algorithm described in Section 2.5.4 applied to the samples $\mathbf{X_1}$ and $\mathbf{X_{2h}}$. For performance reasons, the permutation p-value at step h for partial test i, denoted with λ_{ih}, can be estimated using only an arbitrary number of randomly selected permutations as described in Section 2.5.3. When p-value estimation is used, the monitoring statistic becomes:

$$\hat{t}_h = \psi(\hat{\lambda}_{1h}, ..., \hat{\lambda}_{ph}) \tag{4.2}$$

An important feature of NPC charts is that they use only a single control limit regardless of the alternatives used for the partial tests. A relevant consequence is that the calibration procedure described in Section 3.2.3 is generic enough to be used with any NPC control chart.

Similarly to the NPC framework for statistical tests, also NPC control charts are a very flexible tool. They were first introduced in [11] in order to address the multivariate location monitoring problem, but they can be easily adapted to multi-aspect problems, like the joint location-scale monitoring problem. Furthermore, they can also be used to combine together multiple data aspects (viewpoints of the data), each providing important insights under specific classes of distributions.

To clarify the terminology, in this Thesis, with *multi-aspect NPC control charts*, we refer to control charts that use 4.1 or 4.2 in order to combine different tests

each addressing a different aspect of the monitoring problem or a different data viewpoint, so we do not differentiate between the two cases.

4.2   Implementation: the npcharts R package

An important obstacle to overcome in order to implement calibration and evaluation routines for multi-aspect NPC control charts is that the monitoring statistic for these charts is much more computationally intensive than other traditional Shewhart-type charts. Indeed, to obtain 4.2, a sufficiently large number of permutations of the data must be computed and then, for each of these permutations, each partial statistic must be evaluated.

The npcharts R package has been designed with the specific goal of making calibration and evaluation of multi-aspects NPC charts possible in practice. This package leans on different tools, techniques and ideas which are explained below.

Complexity optimization for linear rank statistics   For linear rank statistics, the permuted values of the statistic can be precomputed, sorted and saved in memory. At each time step of the calibration and evaluation routines (Algorithms 3 and 2), using this idea, the calculation of the p-value amounts to an upper bound search[1] in a sorted array (which is similar to a binary search). The upper bound search can be implemented with a temporal complexity of $O(log(B))$ where B is the number of the considered random permutation.

Support of a compiled language   In computer science, it is well known that compilers can be used to obtain a considerable speed advantage making the most out of the underlying computer architecture. R is not a compiled language and thus it may lead to important overheads when used to implement computationally intensive tasks. Fortunately, there exists a well-known R package, Rcpp [13], which makes the integration between R and C++ easy in practice. C++ is a compiled language that can be used to produce a very efficient program. Rcpp allowed us to write efficient C++ functions which can then be easily called from R. The integration with R was very important in our software project design because R is the most popular programming platform for statisticians, who are the main audience of the methods we propose.

A fast PRNG which supports parallel streams   In our software package, we adopted Xoshiro256+ [2], a fast Pseudo Random Number Generator (PRNG) which supports non-overlapping sequences of pseudo random numbers for parallel computing applications. The Xoshiro256+ generator has a period of $2^{256} - 1$ and provides a constant-time function for jumps of size 2^{128} in the state cycle. The jump function can be used to 2^{128} independent sequences each of which can be consumed in a separate execution thread.

Parallelization of Monte Carlo simulations   Parallel computing is an important computer science technique that can be used to improve the performance of a program by executing computations in parallel. The parallel computing approach

---

[1]Searching the upper bound means searching the first number in the sorted array which is greater than the search argument (here the observed value of the test statistic).

takes advantage of modern multi-core processors, GPUs and computer clusters. Using the non-overlapping streams provided by Xoshiro256+, the simulations used in the calibration and evaluation procedures for control charts can be executed in parallel. This idea has been implemented in the `npcharts` package using the multi-core strategy and the OpenMP software framework [12].

Application of the notion of permutational equivalence   As described in Section 2.5.2, studying permutational equivalence of statistics allows writing faster algorithms by selecting the equivalent statistic which requires the minimum number of arithmetic operations.

Currently, we have decided to officially support only NCP combinations of linear rank tests. Besides performance reasons, the rationale for this decision will be addressed extensively in Section 4.3.

Appendix A includes a snapshot of the current R Application Programming Interface (API) of the `npcharts` R package. The included functions are:

- `npcharts.setseed`: should be used to set the seed of the internal PRNG (Xoshiro256+)

- `compute_monitoring_statistic`: should be used to compute the observed value of the monitoring statistic given a reference sample and a test sample

- `compute_permutation_distribution`: should be used to precompute the (estimated) permutation distribution of the linear rank statistics statistics to be combined with NPC

- `evaluation.unconditional`: implements Algorithms 2

- `calibration.unconditional`: implements Algorithm 3

Below, we show an R script that, using the `npcharts` API, performs the calibration and evaluation of the multi-aspect NPC Wilcoxon-Klotz chart, which will be analyzed in detail in Section 4.5. The execution time of the entire script is around 8 seconds on a system with a Ryzen 5600X 6-core processor and 16 GB of DDR4 RAM.

```r
library(npcharts)
calib.seed = 8989
npcharts.setseed(calib.seed)

# Calibration parameters
calib.m = 100
calib.n = 5
calib.monitor_stat.type= "npc"
calib.monitor_stat.params = list(
  "statistics"= list(
    "centered_wilcoxon_rank_sum",
    "klotz_statistic"
  ),
  "permutation_distributions"=list(
    compute_permutation_distribution("centered_wilcoxon_rank_sum", calib.m, calib.n, 10000),
    compute_permutation_distribution("klotz_statistic", calib.m, calib.n, 10000)
  ),
  "tails"=list(
    "two_sided",
    "right"
  ),
  "combining_function"="fisher"
)
calib.is_upper_limit = FALSE
calib.limits_seq = seq(-7, -9.2, -0.001)
calib.ARL0.target = 500
calib.chart_id="npc_wilcoxon_klotz"
calib.dist = "norm"
calib.dist.params = list("mean" = 0 , "sd" =  1)
```

```
calib.nsim = 10000
calib.cap = 50000

start.time = proc.time()

# Calibration
calib.rls = calibrate.unconditional(
  m = calib.m,
  n = calib.n,
  distribution_key = calib.dist,
  distribution_parameters = calib.dist.params,
  monitoring_statistic_type = calib.monitor_stat.type,
  monitoring_statistic_parameters = calib.monitor_stat.params,
  limits_seq = calib.limits_seq,
  is_upper_limit = calib.is_upper_limit,
  nsim = calib.nsim,
  run_length_cap = calib.cap
)
calib.arls = colMeans(calib.rls)
plot(calib.limits_seq ~ calib.arls, pch=20, cex=0.02, xlab="ARL", ylab ="UCL")
afn = approxfun(calib.arls, calib.limits_seq, ties = mean)
curve(afn, add=TRUE, col="red")
calib.limit = afn(calib.ARL0.target)
abline(h = calib.limit)
abline(v = calib.ARL0.target)

# Evaluation parameters
calib.eval.dist = "laplace"
calib.eval.dist.params = list("location"= 0, "scale"= 1)
calib.eval.nsim = calib.nsim

# Evaluation
if (calib.is_upper_limit){
  limits = list("ucl"=calib.limit)
} else {
  limits = list("lcl"=calib.limit)
}
res = evaluate.unconditional(
  m = calib.m,
  n = calib.n,
  limits = limits,
  location_shift = 0,
  scale_multiplier = 1,
  distribution_key = calib.eval.dist,
  distribution_parameters = calib.eval.dist.params,
  monitoring_statistic_type = calib.monitor_stat.type,
  monitoring_statistic_parameters = calib.monitor_stat.params,
  nsim = calib.eval.nsim,
  run_length_cap = calib.cap
)
print(mean(res$run_lengths))

duration.time = proc.time() - start.time
print(duration.time)
```

4.3   Research questions

Considering the practical importance of the distribution-free notion for control
charts presented in Section 3.2.3, two crucial research questions regarding multi-
aspect NPC control charts are the following:

1. Are multi-aspect NPC control charts distribution-free in general?

2. If they are not distribution-free in general, does combining only linear rank
   tests lead to a distribution-free NPC chart?

The first question is motivated by the idea that, since multi-aspect NPC control
charts use only p-values, their IC performance could be unaffected by the under-
lying process distribution. As for the second question, it can be shown that the
answer is "yes". Indeed, the distribution-free-ness proof given in [6] for the Mann-
Whitney chart can be extended also to any monitoring statistic whose conditional
distribution given $\mathbf{X_1}$ depends only on $F_2(x_{11}), ..., F_2(x_{1m})$. This condition is sat-
isfied by linear rank statistics and any deterministic function of them, like the ones
given in 4.1 and 4.2.

  To answer the first question and check the correctness of our intuition for the
second question, we used the simulation approach for studying the IC performance
of multi-aspect NPC charts and how they are affected by the process distribution.

For this experiment, we considered the multi-aspect NPC charts derived from the Tri-aspect test presented in Section 2.6.4. By combining all the three data aspects T_a, T_b and T_c, the C_{abc} chart is obtained, which uses as monitoring statistic:

$$_hC_{abc} = min(\lambda_{T_ah}, \lambda_{T_bh}, \lambda_{T_ch})$$

If only two aspects are used, then the control charts C_{ab}, C_{ac} and C_{bc} are obtained, whose respective monitoring statistics are given by:

$$_hC_{ab} = min(\lambda_{T_ah}, \lambda_{T_bh})$$
$$_hC_{ac} = min(\lambda_{T_ah}, \lambda_{T_ch})$$
$$_hC_{bc} = min(\lambda_{T_bh}, \lambda_{T_ch})$$

Note that C_{bc} combines only linear rank tests, whereas C_{abc}, C_{ab} and C_{ac} use also a statistic which does not belong to the linear rank family (the difference of means, T_a).

Table 4.1 shows the ARL_0 of C_{abc}, C_{ab}, C_{ac} and C_{bc} under several process distributions after being calibrated with the Normal$(0, 1)$ distribution for $ARL_0 = 250$. The choice of using $ARL_0 = 250$ instead of more common values like 370 or 500 is motivated by the fact that, for the charts that incorporates the T_a aspect, the permutation distribution has to be recomputed in each monitoring step, thus the required computational time for the simulations is very high. The sample sizes used in the experiment are $n_1 = 100$ and $n_2 = 10$, the number of simulations is 10000 and for the calculation of λ_{T_ah}, λ_{T_bh} and λ_{T_ch} a positive-shift one-sided alternative and 3500 permutations are used.

To have an additional comparison, Table 4.1 shows also the ARL_0 of the single-aspect charts $\overline{\mathbf{X}}_2$, T_b and T_c which have been calibrated and evaluated using the same parameters adopted for the multi-aspect charts. To ensure a fair comparison, also for the single-aspect charts, the positive-shift one-sided version of the charts is used.

Distribution	$\overline{\mathbf{X}}_2$	T_c	T_b	C_{ab}	C_{ac}	C_{bc}	C_{abc}
Normal$(0, 1)$	250.7	246.6	281.7	253.2	253.7	254.4	238.6
Laplace$(1, \frac{1}{\sqrt{2}})$	185.5	242.7	279.2	256.0	225.8	249.8	211.3
$X_1 = Z - 1, Z \sim$Exp(1)	79.6	247.0	268.8	271.1	221.3	260.0	211.9
$X_1 = \frac{Z}{\sqrt{5}}, Z \sim$Student$(2.5)$	144.8	245.7	266.3	299.8	204.4	249.9	200.0
Cauchy$(0, 1)$	3.6	253.6	276.4	336.3	179.6	253.0	179.4

Table 4.1: ARL_0 of the multi-aspect control charts C_{abc}, C_{ab}, C_{ac} and C_{bc} under different process distributions after a calibration for $ARL_0 = 250$ with the standard normal distribution.

From the results, we observe that ARL_0 is close to 250 for all control charts under the Normal$(0, 1)$ distribution, with T_b as an exception. This was expected since control charts have been calibrated under Normal$(0, 1)$ for $ARL_0 = 250$. As for T_b, it should be considered that, since the statistic assumes only $min(n_1, n_2) + 1$ possible values, an accurate calibration is is very unlikely when the sample sizes are fixed. Turning now our attention to the other distributions, we see that C_{bc} is the only multi-aspect control chart whose performance seems to remain close to the

calibration target. This is particularly visible with the Cauchy distribution. The simulation results then suggest that NPC multi-aspect charts are not distribution-free in general (the answer to research question 1 is "no"). However, for the chart that combines exclusively linear rank tests (C_{bc}), the results show that the distribution-free-ness property is maintained as expected.

4.4 Multi-aspect NPC control charts for location monitoring

In this section we present and analyze some distribution-free multi-aspect NPC control charts for the location monitoring problem in the standards unknown case. All these charts can be easily implemented using the `npcharts` package.

4.4.1 Proposed control charts

The control chart based on the Mann-Whitney statistic [6] is currently the main landmark of distribution-free charts for the location monitoring problem. Since the Mann-Whitney statistic is one-to-one increasingly related to the Wilcoxon rank sum statistic once fixed the sample sizes, the charts based on these two statistics are equivalent in terms of performance. Besides the Mann-Whitney chart, not many other distribution-free control charts have been proposed. In this thesis, we propose three distribution-free control carts for the location monitoring problem:

- Van der Waerden chart: based on a single aspect, the van der Waerden statistic. This statistic provides a test with a power higher than the Mann-Whitney test when data are normally distributed, while the power may be lower for any other distribution. Therefore, we expect the van der Waerden chart to exhibit better OOC performance than the Wilcoxon chart when the data distribution is normal and worse performance for other distributions.

- NPC Wilcoxon-Waerden chart: obtained from the NPC (Fisher) combination of the Wilcoxon rank sum test with the van der Waerden test. With this chart, we would like to achieve better performance than the Wilcoxon chart when data are normally distributed and performance close to the Wilcoxon chart when the distribution is different.

- NPC percentiles(42,46) chart: based on the NPC (Fisher) combination of two different instances of the class of tests described in Section 2.6.5, one with $s_1 = r_1 = 0.42$ and the other with $s_2 = r_2 = 0.46$. The rationale for this combination is that the test corresponding to $s_1 = r_1 = 0.42$ has a greater power than the Wilcoxon rank sum test under a normal distribution, while the test corresponding to $s_2 = r_2 = 0.46$ is useful as an alternative data viewpoint when the distribution is not normal.

Since all these charts are based exclusively on linear rank statistics and functions of them, following the arguments of Section 4.3, we already know that they are distribution-free.

4.4.2 Unconditional IC/OOC performance analysis

In order to compare the proposed charts with the Mann Whitney/Wilcoxon chart, we carried out a comparative simulation study under the location model described in Section 3.2.2 with several distributions and different location shifts.

Methodology

First, we calibrate control charts for $ARL_0 = 370$, $n_1 = 100$ and $n_2 = 5$ using Algorithm 3 and Algorithm 4 under the standard normal distribution. For all control charts, the two-sided version has been considered.

After the calibration, we use Algorithm 2 with 50000 simulations in order to assess the uncoditional run length distribution with different values of the location shift parameter θ, and under the following distributions (also shown in Figure 4.1):

- Standard normal: represents a thin-tailed and symmetric distribution.

- Laplace $(0, 1/\sqrt{2})$: represents a heavy-tailed and symmetric distribution.

- Gamma(shape $= 4$, scale $= 2$) shifted by -2 so that the mean is equal to 0. This distribution, also proposed in [19] for valve leakage monitoring, has been considered to assess performance with asymmetric data. Because of the asymmetry, performance may be different under positive and negative location shifts of the same magnitude. Therefore, with this distribution, both positive and negative location shifts are considered.

All these distributions have mean equal to 0 and standard deviation equal to 1. Having normalized distributions is important since it allows us to compare performance in a fair manner also across different distributions.

For NPC charts, $B = 10000$ permutations are used to estimate the p-values.

Results

The results are shown in Table 4.2, Table 4.3, Table 4.4 and Table 4.5. The result format for all the tables is the following: the first row of each of the cells shows the ARL and [SDRL] values, while the second row shows the 5th, 50th and 95th percentiles of the run length distribution (in this order).

Under the standard normal distribution (Table 4.2), the van der Waerden chart outperforms the Wilcoxon chart as expected. The performance of the NPC Wilcoxon-Waerden chart seems in the middle between the Wilcoxon and the van der Waerden charts, but closer to the van der Waerden. With a location shift of 0.5 standard deviations, the NPC Wilcoxon-Waerden is, on average, around 5.8% faster at detecting the shift than the Wilcoxon chart. With a location shift of 1 standard deviation units this speedup increases to around the 10.5%.

With the Laplace distribution (Table 4.3) the Wilcoxon chart seems to have the best OOC performance. With a location shift of 0.5 standard deviations, the Wilcoxon chart is, on average, around 16.5% faster at detecting the shift than the NPC Wilcoxon-Waerden chart. With a location shift of 1 standard deviation units this speedup decreases to around the 13.2%. In comparison, with the van der Waerden chart, the speedups are 34.1% for $\theta = 0.5$ and 36.8% for $\theta = 1$.

Under the Gamma distribution, for positive location shifts (Table 4.4), we note that all charts seems to have some difficulties with small values of θ. In particular, for $\theta = 0.1$, the ARL of all charts increase instead of decreasing. For $\theta \geq 0.2$ the Wilcoxon chart seems better than the NPC Wilcoxon-Waerden chart, however, for larger location shifts, the speedup becomes smaller and smaller. With negative location shifts (Table 4.5), it is the NPC Wilcoxon-Waerden that outperforms the Wilcoxon chart. However, in contrast to the pattern seen for positive shifts, in this case the speedup seems to remain constant even for larger shifts.

Finally, with all distributions, the performance NPC percentiles(42,46) chart seems very close to those of the Wilcoxon chart if we consider the results of all four charts together.

(a)

(b)

(c)
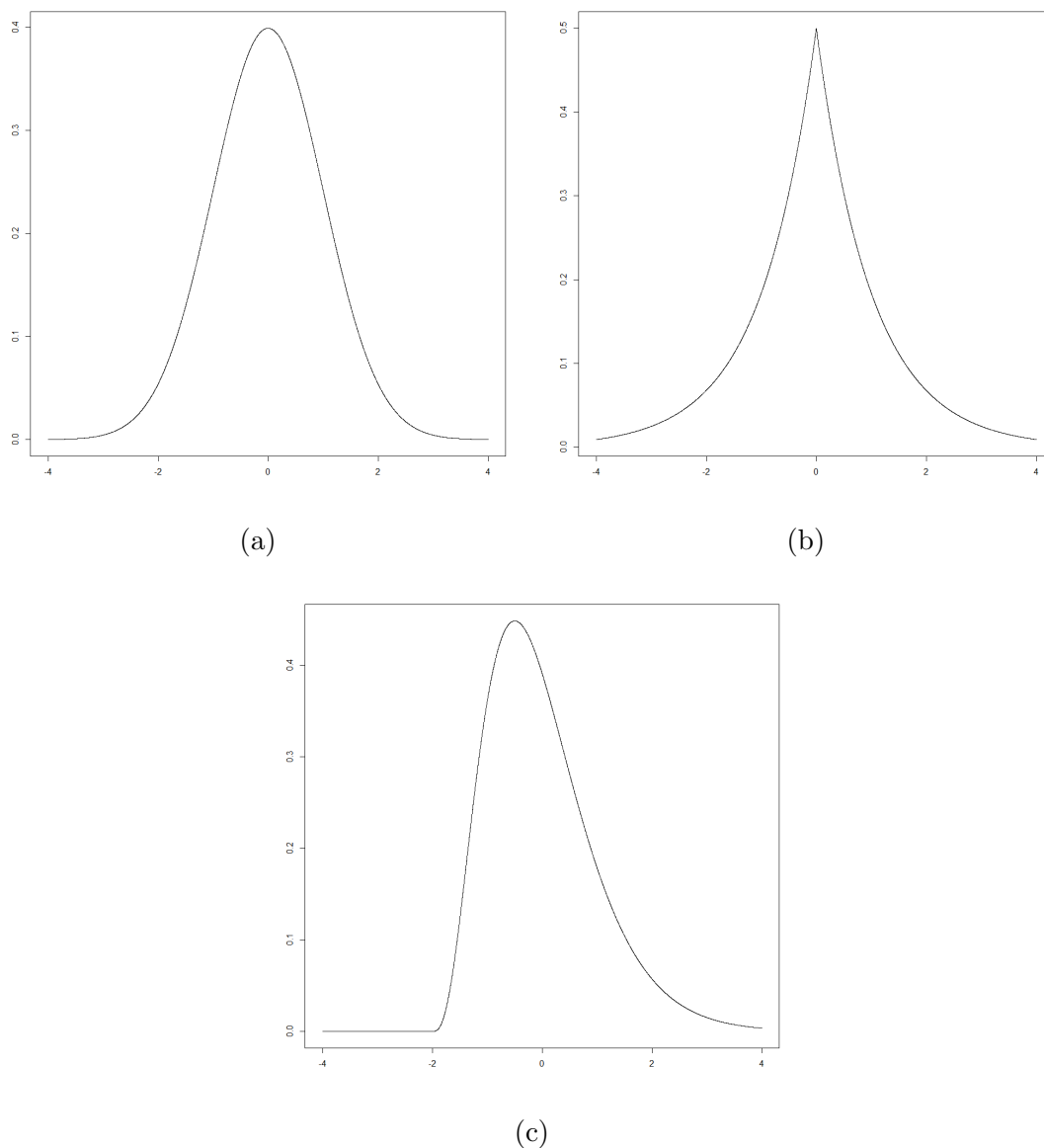
Figure 4.1: PDF of the four distributions used for the performance evaluation. (a) The standard normal distribution. (b) The Laplace$(0, 1/\sqrt{2})$ distribution. (c) The Gamma$(\text{shape} = 4, \text{scale} = 2)$ distribution shifted by -2.

θ	δ	Wilcoxon	Van der Waerden	NPC percentiles(42,46)	NPC Wilcoxon-Waerden
0	1	370.8 [495.8]	367.3 [526.7]	366.4 [484.1]	368.4 [512.8]
		15, 214, 1231	14, 197, 1275	15, 210, 1223	14, 204, 1252
0.1	1	324.6 [453.2]	323.6 [479.7]	327.3 [449.1]	326.1 [475.1]
		12, 183, 1096	11, 170, 1131	12, 182, 1113	12, 177, 1131
0.2	1	234.9 [346.1]	233.3 [388.4]	229.5 [338.6]	225.4 [339.7]
		8, 123, 824	7, 113, 837	8, 121, 792	8, 116, 801
0.3	1	144.1 [216.3]	140.2 [242.4]	142.5 [223.9]	139 [218.7]
		5, 75, 506	4, 66, 503	5, 73, 502	5, 70, 497
0.4	1	83.6 [125.8]	79.5 [133.5]	83 [132.4]	81.9 [131.4]
		3, 44, 289	3, 39, 283	3, 43, 290	3, 42, 286
0.5	1	49.5 [70.1]	46.8 [78.1]	48.8 [69.4]	46.8 [72.6]
		2, 28, 168	2, 24, 165	2, 27, 168	2, 25, 161
0.6	1	30.2 [46.1]	27.7 [42.6]	30.1 [44.1]	28.4 [41.2]
		2, 17, 100	1, 15, 95	2, 17, 101	1, 16, 96
0.7	1	19.2 [24.6]	17.3 [24.8]	19 [24.7]	17.7 [23.3]
		1, 11, 63	1, 10, 57	1, 11, 62	1, 10, 59
0.8	1	12.8 [15.7]	11.2 [14.6]	12.4 [15.4]	11.7 [15]
		1, 8, 41	1, 7, 37	1, 7, 39	1, 7, 37
0.9	1	8.8 [10.3]	7.7 [9.3]	8.5 [10.1]	8 [9.3]
		1, 6, 27	1, 5, 24	1, 5, 27	1, 5, 25
1	1	6.3 [6.8]	5.4 [6.1]	6.1 [6.6]	5.7 [6.2]
		1, 4, 19	1, 3, 16	1, 4, 19	1, 4, 17

Table 4.2: Unconditional performance of the Wilcoxon, Van der Waerden, NPC percentiles(42,46) and NPC Wilcoxon-Waerden charts under the standard normal distribution.

θ	δ	Wilcoxon	Van der Waerden	NPC percentiles(42,46)	NPC Wilcoxon-Waerden
0	1	366.8 [485.9]	364.1 [526.8]	364.6 [485.3]	368.7 [502.5]
		15, 212, 1217	14, 197, 1257	15, 209, 1233	14, 208, 1249
0.1	1	325.2 [451.3]	329.7 [503.4]	324.3 [441.6]	325.4 [462.3]
		12, 179, 1123	11, 169, 1167	12, 178, 1112	12, 175, 1134
0.2	1	226.4 [363.1]	250 [427.7]	230.9 [363.6]	236.1 [373.9]
		7, 112, 817	7, 116, 913	7, 114, 837	7, 116, 853
0.3	1	135.3 [235.3]	158.2 [302.4]	138.8 [242.3]	145.1 [268.1]
		4, 63, 492	4, 68, 591	4, 63, 519	4, 65, 534
0.4	1	74.4 [134.5]	93.7 [184.9]	77.7 [145]	82.3 [162.9]
		3, 34, 271	3, 40, 351	3, 35, 287	3, 36, 303
0.5	1	39.9 [82.6]	53.5 [108.7]	41.8 [77.6]	46.5 [93.1]
		2, 18, 142	2, 23, 194	2, 19, 150	2, 21, 168
0.6	1	22.6 [41.1]	31.4 [67.4]	23.5 [41.2]	25.9 [50.9]
		1, 11, 79	1, 14, 112	1, 11, 83	1, 12, 92
0.7	1	13.1 [22]	18.6 [35]	13.5 [22.4]	14.9 [24.9]
		1, 7, 44	1, 9, 65	1, 7, 45	1, 8, 51
0.8	1	8.2 [12]	11.5 [20.2]	8.4 [13.5]	9.4 [15.1]
		1, 5, 27	1, 6, 40	1, 5, 28	1, 5, 31
0.9	1	5.4 [7.1]	7.5 [11.9]	5.5 [7.3]	6.2 [9.1]
		1, 3, 17	1, 4, 25	1, 3, 17	1, 4, 20
1	1	3.8 [4.3]	5.2 [7.9]	3.9 [4.6]	4.3 [5.6]
		1, 2, 11	1, 3, 16	1, 2, 12	1, 3, 13

Table 4.3: Unconditional performance of the Wilcoxon, Van der Waerden, NPC percentiles(42,46) and NPC Wilcoxon-Waerden charts under the Laplace$(0, 1/\sqrt{2})$ distribution.

θ	δ	Wilcoxon	Van der Waerden	NPC percentiles(42,46)	NPC Wilcoxon-Waerden
0	1	363.7 [476.7]	365.9 [531.1]	366.8 [488.5]	367.3 [487.3]
		15, 211, 1208	13, 197, 1271	15, 209, 1233	14, 207, 1255
0.1	1	400.6 [562.2]	441.4 [748.7]	408 [587.5]	429.3 [675.6]
		15, 217, 1381	13, 214, 1572	15, 219, 1419	14, 221, 1521
0.2	1	335.1 [589.6]	368.4 [735.9]	331 [544.2]	351.9 [630.3]
		10, 161, 1199	10, 157, 1343	10, 161, 1201	10, 161, 1289
0.3	1	217.1 [397.4]	246.2 [524.2]	221.8 [403.2]	227.7 [436.7]
		7, 101, 788	6, 99, 934	7, 102, 809	6, 99, 846
0.4	1	133.7 [253.9]	147.8 [322.7]	134.8 [246.5]	138.8 [261.1]
		4, 62, 486	4, 61, 541	4, 62, 488	4, 61, 513
0.5	1	81.9 [149.5]	90.1 [182.4]	82.7 [154.8]	82.8 [154.8]
		3, 39, 290	3, 38, 334	3, 39, 294	3, 38, 301
0.6	1	50 [87.7]	54.8 [105.8]	50.3 [85.6]	51.9 [90.2]
		2, 24, 176	2, 24, 200	2, 24, 179	2, 24, 187
0.7	1	30.6 [48.3]	34.5 [66.6]	31.9 [54.9]	32.4 [56.4]
		1, 16, 109	1, 16, 123	1, 16, 112	1, 16, 115
0.8	1	20.1 [31.5]	22.5 [50.4]	20.2 [31.2]	20.5 [33.1]
		1, 10, 69	1, 11, 80	1, 10, 70	1, 10, 72
0.9	1	13.2 [19.4]	14.8 [25]	13.5 [22.1]	13.6 [21.5]
		1, 7, 45	1, 7, 51	1, 7, 46	1, 7, 47
1	1	9 [13.1]	10.1 [17.2]	9.1 [13.4]	9.3 [14.6]
		1, 5, 29	1, 5, 34	1, 5, 30	1, 5, 31

Table 4.4: Unconditional performance of the Wilcoxon, Van der Waerden, NPC percentiles(42,46) and NPC Wilcoxon-Waerden charts under the Gamma$(4, 1/2)$ distribution shifted by -2.

θ	δ	Wilcoxon	Van der Waerden	NPC percentiles(42,46)	NPC Wilcoxon-Waerden
0	1	368.1 [485.6]	368.1 [546]	369.5 [497.9]	369.8 [508.2]
		15, 213, 1227	13, 197, 1283	15, 210, 1238	14, 207, 1252
-0.1	1	241 [313.9]	210.6 [298]	238.8 [316.6]	223.7 [293.4]
		10, 140, 813	8, 115, 726	10, 138, 803	9, 128, 754
-0.2	1	131.2 [170.5]	98.2 [131.2]	126.4 [166.5]	112 [148.6]
		5, 76, 444	4, 55, 334	5, 72, 430	5, 64, 380
-0.3	1	67 [84.5]	45.3 [58.4]	63.4 [79.2]	54.2 [69.4]
		3, 40, 220	2, 27, 150	3, 38, 210	3, 32, 181
-0.4	1	36 [43.2]	23 [27.3]	34.2 [41.4]	27.9 [33.1]
		2, 22, 116	1, 14, 74	2, 21, 111	2, 17, 90
-0.5	1	21.2 [24.1]	12.9 [14.2]	19.3 [21.7]	16 [17.6]
		1, 13, 67	1, 8, 40	1, 12, 61	1, 10, 50
-0.6	1	13.3 [14.3]	8.1 [8.4]	12.2 [13.1]	9.8 [10.4]
		1, 9, 41	1, 5, 24	1, 8, 37	1, 7, 30
-0.7	1	9.1 [9.3]	5.5 [5.4]	8.2 [8.3]	6.7 [6.6]
		1, 6, 27	1, 4, 16	1, 6, 24	1, 5, 20
-0.8	1	6.4 [6.3]	4 [3.6]	5.8 [5.7]	4.8 [4.5]
		1, 4, 19	1, 3, 11	1, 4, 17	1, 3, 14
-0.9	1	4.8 [4.6]	3.1 [2.6]	4.4 [4.1]	3.6 [3.2]
		1, 3, 14	1, 2, 8	1, 3, 12	1, 3, 10
-1	1	3.8 [3.4]	2.5 [2]	3.5 [3.1]	2.9 [2.4]
		1, 3, 10	1, 2, 6	1, 2, 10	1, 2, 8

Table 4.5: Unconditional performance (negative shifts) of the Wilcoxon, Van der Waerden, NPC percentiles(42,46) and NPC Wilcoxon-Waerden charts charts under the Gamma$(4, 1/2)$ distribution shifted by -2.

4.4.3 Conditional IC performance analysis

Besides evaluating the unconditional performance, we also studied the performance conditional to specific categories of the reference samples.

Methodology

After a calibration with $\mathrm{ARL}_0 = 370$, the methodology is the same proposed in [27]. We use Algorithm 2 with 50000 run length simulations. Each run length simulation is then classified according to two criteria, the value of the reference sample mean and the value of the reference sample standard deviation. For both of these two criteria, 7 percentile-based bins are used:

- 5th percentile or lower: very high downward bias

- between the 5th and the 25th percentiles: moderately high downward bias

- between the 25th and the 45th percentiles: low downward bias

- between the 45th and the 55th percentiles

- between the 55th and the 75th percentiles: low upward bias

- between the 75th and the 95th percentiles: moderately high upward bias

- beyond the 95th percentiles: high upward bias

Therefore, each of the 50000 run length simulations is categorized in a 7×7 way table depending on the value of the corresponding reference sample mean and SD. For each of the 49 bins, we record the proportion of simulations as well as the sample mean and sample SD of the simulated run lengths that belong to the bin. These last two values are, respectively, an estimate of the conditional ARL_0 and SDRL_0 given the specified percentiles intervals of the reference sample mean and SD with respect to their distributions.

The process distributions and sample sizes adopted in this study are the same used in the unconditional evaluation (Section 4.4.2).

Results

The results are shown in Table 4.6 (Normal), Table 4.7 (Laplace) and Table 4.8 (Gamma). For each cell, the top value (in italics) shows the proportion, the bottom left value shows the conditional ARL_0 and the bottom right value (within the square brackets) shows the conditional SDRL_0. For the colored cells, the conditional ARL_0 value is within the "safe" interval $[250, 900]$ proposed in [9].

In general, for all control charts and distributions, we see the same patterns described in [9]. Given a range of percentiles of the reference sample mean, the larger the SD, the larger the conditional ARL_0. By contrast, given a specific range of percentiles of the SD of the reference sample, the conditional ARL_0 initially increases with an increase in the sample mean, reaches a peak around between the 45 and the 55 percentiles of the sample mean distribution and then decreases.

Comparing the conditional performance between different charts, it seems that, with all distributions, the Wilcoxon chart is the least sensitive to reference samples that have extreme mean or SD. This is a rather surprising result because we have

seen how the Wilcoxon chart is also the most sensitive chart to location shifts under the Laplace distribution. We could therefore have expected that this increased OOC responsiveness under the Laplace distribution translates also to an increased sensitivity to extreme references samples, however, simulation seems to indicate that this is not the case.

Wilcoxon

Sample SD percentiles	Sample mean percentiles						
	0-5	5-25	25-45	45-55	55-75	75-95	95-100
0-5	0.0027 82.4 [97.7]	0.0096 128.3 [137.7]	0.0099 170.5 [178.5]	0.005 169.3 [181.7]	0.0101 151.8 [142]	0.0104 146.2 [160.7]	0.0024 86 [77.7]
5-25	0.0101 120.4 [120.2]	0.0404 186.6 [192.7]	0.0396 230.4 [237.6]	0.02 229.8 [243.4]	0.0399 236.5 [246.5]	0.0395 188.9 [206.1]	0.0105 132.3 [147.3]
25-45	0.0095 142.4 [150.9]	0.0405 264.6 [287]	0.0413 310.5 [321.8]	0.019 311.9 [317.3]	0.0401 311.6 [334.8]	0.0396 262.2 [271.4]	0.0099 168.6 [184.5]
45-55	0.0051 174.4 [171.7]	0.0194 285.9 [310.1]	0.0198 359.3 [395.7]	0.0109 361.2 [367.4]	0.0199 373.5 [388]	0.0202 296.4 [313]	0.0047 166.8 [162.1]
55-75	0.0099 209.5 [236.9]	0.0414 378 [421.2]	0.0394 431 [457]	0.0196 431.9 [439.2]	0.0405 433.3 [470.8]	0.0399 358.9 [377.5]	0.0093 216.2 [247.3]
75-95	0.0101 312.9 [344.9]	0.039 513.8 [551.7]	0.0401 612.9 [684.6]	0.02 682.8 [734.9]	0.0394 628.9 [668.5]	0.0411 511.7 [550.3]	0.0103 317.6 [356]
95-100	0.0027 421.1 [511.7]	0.0098 798.9 [916]	0.0099 1122 [1227.8]	0.0055 1145.3 [1248.8]	0.01 1154.7 [1216.2]	0.0093 815.8 [874.5]	0.0029 536.7 [547]

van der Waerden

Sample SD percentiles	Sample mean percentiles						
	0-5	5-25	25-45	45-55	55-75	75-95	95-100
0-5	0.0022 72.2 [68.1]	0.0098 101.4 [106.9]	0.0104 113.2 [109.3]	0.0053 120.9 [108.3]	0.0093 115.8 [114.4]	0.0098 93.8 [94.1]	0.0032 68 [71.9]
5-25	0.0094 111.2 [110.8]	0.0411 167.6 [166.3]	0.0393 186.3 [195.1]	0.0199 198.4 [210.7]	0.039 184.8 [192.8]	0.0413 158.7 [165.1]	0.01 105.1 [123.1]
25-45	0.01 147.4 [167.7]	0.0403 228 [246.8]	0.0397 278.3 [282.7]	0.0196 288.6 [291.5]	0.0405 280.9 [290.8]	0.0404 229.8 [238.1]	0.0095 146.8 [165.3]
45-55	0.0054 184.4 [188.8]	0.0198 284.9 [285]	0.0199 348.8 [354.6]	0.0095 340.5 [351.6]	0.0198 341.5 [337.3]	0.0204 280.6 [298.4]	0.0054 190.7 [204.7]
55-75	0.0108 231.4 [256.4]	0.0393 360.1 [379.8]	0.0409 433.2 [442.9]	0.0202 442.5 [424.8]	0.0407 425.4 [422.8]	0.0388 354.4 [368]	0.0093 212.2 [242.6]
75-95	0.0099 292.8 [309.1]	0.04 538.2 [575.9]	0.0395 643.6 [677.2]	0.0209 673.4 [708.8]	0.0405 697.6 [714.6]	0.039 550 [597.8]	0.0101 329.9 [387.7]
95-100	0.0022 581.4 [711.8]	0.0096 1043.8 [1072.3]	0.0103 1273.7 [1446.7]	0.0046 1351 [1477.4]	0.0103 1320.8 [1364.9]	0.0104 1000.1 [1554.4]	0.0026 688.9 [800.9]

NPC percentiles(42,46)

Sample SD percentiles	Sample mean percentiles						
	0-5	5-25	25-45	45-55	55-75	75-95	95-100
0-5	0.0023 86.8 [83.1]	0.0101 122.7 [125.4]	0.0096 143.7 [141.8]	0.0048 129 [132.1]	0.0104 160.5 [172.4]	0.0105 126.5 [132.3]	0.0023 84 [72.1]
5-25	0.01 130 [154.7]	0.0401 193.4 [203]	0.0394 228.7 [246.6]	0.0197 244.9 [256.6]	0.0417 221.9 [222]	0.0396 193.6 [203.5]	0.0096 117.9 [140.1]
25-45	0.0094 163.5 [175.1]	0.0392 259.4 [270.3]	0.0399 314.3 [312.9]	0.0201 323.9 [324.2]	0.0405 317.1 [323.2]	0.0407 255 [257.3]	0.0102 172.4 [186.4]
45-55	0.0059 220.1 [236.6]	0.0189 301.2 [310.1]	0.0205 374.6 [387.8]	0.0099 358.6 [348.1]	0.0201 365.4 [382.1]	0.02 312.4 [334.2]	0.0047 195.3 [180.4]
55-75	0.0097 213 [220.1]	0.0413 363.9 [388.7]	0.0405 435.2 [462.7]	0.0202 441.3 [446.3]	0.0377 453.9 [471.9]	0.0402 366.2 [385.3]	0.0105 232.4 [253.6]
75-95	0.0096 303.5 [326]	0.0404 525.4 [572.3]	0.04 634.9 [674.3]	0.0202 632 [648.6]	0.04 627.2 [697.9]	0.0392 496.2 [513]	0.0107 301 [298.1]
95-100	0.0031 551.8 [582.3]	0.01 901 [985]	0.0101 1080.5 [1252.7]	0.0052 1135.4 [1422.1]	0.0096 1073.3 [1248.1]	0.0098 820.2 [945.7]	0.002 587.2 [899.6]

NPC Wilcoxon-Waerden

Sample SD percentiles	Sample mean percentiles						
	0-5	5-25	25-45	45-55	55-75	75-95	95-100
0-5	0.0022 83.7 [77.1]	0.0101 104.5 [106.9]	0.0095 133.4 [131.5]	0.005 120.6 [119.2]	0.0101 143.3 [157]	0.0106 113.1 [122]	0.0025 69 [64.6]
5-25	0.0097 117.1 [130.3]	0.0397 182.9 [192.2]	0.0395 208.1 [207.5]	0.0196 221.1 [228.9]	0.0409 213.3 [221.8]	0.0403 179.6 [190]	0.0104 110.2 [120.8]
25-45	0.0094 162.3 [185.5]	0.039 245.7 [255.8]	0.0405 298.8 [299.5]	0.0199 315.6 [302.7]	0.0404 304.6 [308.7]	0.0409 245.6 [254]	0.0099 175.1 [179.3]
45-55	0.0058 204.8 [214.7]	0.0193 285.2 [290.2]	0.0202 367.8 [374.6]	0.0101 350.4 [335.4]	0.0195 363.5 [377.6]	0.0203 299.5 [305.3]	0.0048 191.3 [184.1]
55-75	0.0103 219.3 [238.4]	0.0415 361 [370.6]	0.0407 434.9 [450.3]	0.02 434 [455.4]	0.0379 428.2 [430.9]	0.0392 356.8 [371]	0.0103 227.5 [236.1]
75-95	0.0096 330.5 [364.6]	0.0406 533.9 [574.4]	0.0396 667.6 [666.4]	0.0201 642.3 [663]	0.041 638.2 [636.4]	0.0388 517.2 [536.1]	0.0103 333.3 [339.4]
95-100	0.003 531.4 [572.5]	0.0099 857.1 [930.1]	0.0101 1178.3 [1377.6]	0.0053 1337.5 [1575.7]	0.01 1279.2 [1416.5]	0.0099 925.1 [1235.4]	0.0019 646.4 [894.5]

Table 4.6: In-control conditional performance of the Wilcoxon, Van der Waerden, NPC percentiles(42,46) and NPC Wilcoxon-Waerden charts under the standard normal distribution.

Wilcoxon

Sample SD percentiles	Sample mean percentiles						
	0-5	5-25	25-45	45-55	55-75	75-95	95-100
0-5	*0.0015* 80.1 [82.7]	*0.0094* 139.5 [148]	*0.0114* 195.3 [205.2]	*0.0056* 174.9 [186.1]	*0.0112* 157.7 [168.7]	*0.01* 138.6 [136.6]	*0.0009* 82.5 [81.8]
5-25	*0.0076* 126.1 [140.9]	*0.0394* 205.5 [225.9]	*0.0416* 251.6 [254.4]	*0.0212* 263.8 [291]	*0.0426* 258.3 [270.7]	*0.0391* 204.7 [211.8]	*0.0085* 124.9 [152.5]
25-45	*0.0088* 165.8 [179]	*0.0406* 272.5 [291.9]	*0.0408* 331.3 [354.1]	*0.0197* 326.3 [342.2]	*0.041* 343.9 [366]	*0.0402* 276.2 [292.7]	*0.0088* 167.8 [178]
45-55	*0.0051* 216.4 [233.7]	*0.0188* 324 [352.7]	*0.0211* 386.9 [398.2]	*0.0098* 418.1 [427.3]	*0.02* 402.1 [450.3]	*0.0197* 341.2 [349.3]	*0.0055* 180.6 [228]
55-75	*0.0105* 236.3 [257.8]	*0.0402* 380.6 [435.5]	*0.0382* 487.7 [551]	*0.0207* 486.9 [550.8]	*0.0384* 462 [516.5]	*0.0413* 368.1 [414]	*0.0107* 222.5 [240.7]
75-95	*0.0124* 298.1 [384.1]	*0.041* 473.5 [547.9]	*0.038* 601.4 [682.6]	*0.0186* 605.9 [675]	*0.0384* 577.5 [656.5]	*0.0394* 497.2 [597.1]	*0.0122* 310.3 [387.1]
95-100	*0.0041* 465.2 [637.5]	*0.0106* 708.2 [1109.6]	*0.0089* 838.7 [965]	*0.0043* 910.8 [936.4]	*0.0083* 844.1 [1072.6]	*0.0103* 758.3 [1134.8]	*0.0035* 620.4 [1106.2]

van der Waerden

Sample SD percentiles	Sample mean percentiles						
	0-5	5-25	25-45	45-55	55-75	75-95	95-100
0-5	*0.0013* 76.4 [66.4]	*0.0091* 104.6 [111.8]	*0.0108* 127 [124.3]	*0.0061* 132.8 [125.4]	*0.0119* 118.5 [119.4]	*0.0095* 101.1 [100.4]	*0.0012* 64 [75.7]
5-25	*0.0077* 92.6 [97.1]	*0.0391* 154.7 [162.4]	*0.0426* 202.4 [205.2]	*0.021* 212.9 [223.8]	*0.0416* 201.9 [201.1]	*0.0397* 168 [181.9]	*0.0083* 100.7 [109.7]
25-45	*0.0092* 135 [146.5]	*0.0395* 235.4 [243.2]	*0.0411* 291.6 [303]	*0.0203* 307.9 [322.5]	*0.0407* 297 [304.7]	*0.0408* 236.4 [245.7]	*0.0084* 147.6 [147.9]
45-55	*0.0049* 190.4 [195.8]	*0.0199* 294.4 [308.8]	*0.0194* 338.7 [341.2]	*0.0101* 376.5 [366.3]	*0.0204* 349.6 [349.7]	*0.0199* 285.9 [303]	*0.0054* 157.9 [155.4]
55-75	*0.0108* 215.9 [238.4]	*0.0406* 355.8 [385.4]	*0.039* 470.9 [476.7]	*0.0195* 479.4 [488.5]	*0.0393* 462 [479.2]	*0.04* 343.4 [360.4]	*0.0108* 202 [221.1]
75-95	*0.0123* 273.3 [318.6]	*0.0412* 523.6 [567.7]	*0.0382* 671.3 [691.3]	*0.0189* 713.6 [796.3]	*0.0381* 658 [750.8]	*0.0395* 533.6 [598.6]	*0.0119* 286.8 [341.5]
95-100	*0.0039* 486 [514.5]	*0.0105* 991 [1315.8]	*0.0089* 1368.4 [1858.2]	*0.0041* 1431.9 [1892.2]	*0.008* 1141.9 [1373.1]	*0.0107* 901.9 [1026.2]	*0.004* 516.1 [627.4]

NPC percentiles(42,46)

Sample SD percentiles	Sample mean percentiles						
	0-5	5-25	25-45	45-55	55-75	75-95	95-100
0-5	*0.0016* 90.2 [121.2]	*0.0094* 134.1 [136.8]	*0.0112* 178.6 [188.5]	*0.0057* 174.6 [187.4]	*0.0109* 159.9 [162.2]	*0.0101* 126.1 [120.5]	*0.001* 77.7 [82.5]
5-25	*0.0074* 115.5 [126.1]	*0.039* 193.4 [205.2]	*0.0418* 245.3 [250.2]	*0.0213* 267.6 [285]	*0.043* 261 [278.5]	*0.039* 197.8 [205]	*0.0085* 129.1 [153.9]
25-45	*0.0087* 159.2 [178.2]	*0.0406* 267.1 [276.8]	*0.0408* 326.5 [350.5]	*0.02* 338.1 [361]	*0.0385* 334.5 [361.4]	*0.0402* 264.5 [282.1]	*0.0088* 160.7 [171]
45-55	*0.0051* 208.3 [219.4]	*0.0194* 324.7 [349.4]	*0.0203* 384 [409.9]	*0.0099* 424.4 [436.4]	*0.0203* 395 [433.3]	*0.0197* 337.1 [348.5]	*0.0054* 195.2 [252.1]
55-75	*0.0106* 222.2 [213.7]	*0.0398* 376 [425]	*0.0387* 479.2 [509.7]	*0.0204* 492 [562.9]	*0.0387* 471.1 [530]	*0.041* 364.9 [401.7]	*0.0108* 231 [244.1]
75-95	*0.0124* 297.3 [352.9]	*0.0413* 477.2 [543.3]	*0.0384* 624.9 [698.7]	*0.0184* 638.6 [677.9]	*0.0379* 566 [636.6]	*0.0398* 492.8 [563.3]	*0.012* 324.3 [428.6]
95-100	*0.0043* 523.1 [796.2]	*0.0104* 725.9 [1034]	*0.0088* 849.7 [968.6]	*0.0043* 986.8 [1109.1]	*0.0084* 887.4 [1128.4]	*0.0102* 818.4 [1302.7]	*0.0036* 570.1 [1066.3]

NPC Wilcoxon-Waerden

Sample SD percentiles	Sample mean percentiles						
	0-5	5-25	25-45	45-55	55-75	75-95	95-100
0-5	*0.0015* 78.6 [103.5]	*0.0091* 117.4 [122.5]	*0.0108* 155.4 [166.9]	*0.0062* 152.3 [159]	*0.0117* 129.5 [134.1]	*0.0097* 113.4 [106.7]	*0.0009* 59.9 [81.5]
5-25	*0.0078* 103.3 [118.1]	*0.0396* 173.6 [177.4]	*0.042* 225.1 [229]	*0.0211* 252.1 [279.4]	*0.0425* 227.5 [230.4]	*0.0384* 184.7 [196]	*0.0086* 111 [134.2]
25-45	*0.0089* 146.4 [151.4]	*0.0398* 249.5 [255.1]	*0.0415* 315.5 [334.1]	*0.0204* 313.5 [314.9]	*0.0401* 327.1 [331.1]	*0.0406* 249.8 [258.3]	*0.0087* 151.3 [167.6]
45-55	*0.0047* 196.5 [215.8]	*0.0196* 294.2 [304.4]	*0.0203* 375.2 [384.7]	*0.0099* 393.2 [385.4]	*0.02* 380.5 [394]	*0.0199* 314.7 [326]	*0.0055* 182.4 [203.2]
55-75	*0.0106* 221.9 [241.4]	*0.0396* 362.1 [391.8]	*0.0384* 488.7 [500.1]	*0.02* 496.8 [524.1]	*0.0392* 465.5 [508.1]	*0.0415* 355.3 [388.9]	*0.0107* 231.3 [254.7]
75-95	*0.0122* 290.7 [334.9]	*0.0419* 507.6 [563.8]	*0.0381* 655.6 [710.8]	*0.0181* 686.5 [764.4]	*0.0379* 615.3 [681.5]	*0.0397* 510.2 [585.6]	*0.012* 321.7 [393.4]
95-100	*0.0042* 511.1 [621.3]	*0.0103* 875.5 [1541.5]	*0.0089* 1073.7 [1424.8]	*0.0043* 1120 [1381.3]	*0.0085* 979.8 [1178.4]	*0.0102* 795.9 [1017.6]	*0.0035* 569 [1037.4]

Table 4.7: In-control conditional performance of the Wilcoxon, Van der Waerden, NPC percentiles(42,46) and NPC Wilcoxon-Waerden charts under the Laplace $(0, 1/\sqrt{2})$ distribution.

Wilcoxon

Sample SD percentiles	Sample mean percentiles						
	0-5	5-25	25-45	45-55	55-75	75-95	95-100
0-5	*0.0133* 120.9 [124.3]	*0.0209* 164 [168.7]	*0.0091* 164.3 [166.5]	*0.0028* 178 [165.3]	*0.0029* 163.3 [173.7]	*0.001* 102.6 [110.6]	*≤0.00006* -
5-25	*0.0217* 156.7 [167.7]	*0.0716* 236.6 [258.1]	*0.0479* 255.7 [291.1]	*0.0182* 249.4 [266.1]	*0.0265* 218.3 [226.1]	*0.0131* 157.7 [185.2]	*0.001* 80.4 [104.3]
25-45	*0.009* 184.3 [191.7]	*0.0489* 308.7 [329.2]	*0.0509* 332.8 [351.1]	*0.0229* 340.4 [358.7]	*0.0384* 296.2 [332.6]	*0.027* 228.9 [255.3]	*0.0029* 118.8 [112.1]
45-55	*0.0024* 220.5 [241.2]	*0.0189* 341.8 [381.9]	*0.0228* 395 [432.4]	*0.0118* 394.1 [403.6]	*0.0227* 346.1 [379.7]	*0.0189* 274.6 [325.9]	*0.0025* 136.7 [142]
55-75	*0.0028* 228.2 [249.3]	*0.0255* 377.1 [417.2]	*0.0402* 467.4 [494.1]	*0.0238* 510.8 [570.5]	*0.0503* 464 [529.9]	*0.0487* 336 [399.2]	*0.0086* 192.3 [245.2]
75-95	*0.0007* 274.8 [281.7]	*0.0131* 508 [614.5]	*0.0262* 572.2 [692.5]	*0.018* 643.9 [783.3]	*0.05* 637.7 [747.4]	*0.0701* 487.8 [589.5]	*0.0218* 273.2 [323.8]
95-100	*≤0.00006* -	*0.0009* 364.4 [385.8]	*0.0028* 764 [910.3]	*0.0025* 807 [882.7]	*0.0092* 832.9 [950.6]	*0.0214* 824.8 [1209.5]	*0.0131* 477.4 [748.4]

van der Waerden

Sample SD percentiles	Sample mean percentiles						
	0-5	5-25	25-45	45-55	55-75	75-95	95-100
0-5	*0.014* 87.6 [94.5]	*0.0206* 117.2 [114.5]	*0.0086* 128.7 [132.1]	*0.0028* 127.3 [129.6]	*0.003* 123.3 [146.9]	*0.0009* 91.3 [92.7]	*≤0.00006* -
5-25	*0.0214* 134.2 [135]	*0.0715* 191.9 [197.7]	*0.0484* 207.9 [213]	*0.0186* 202.6 [227.1]	*0.026* 181 [190.5]	*0.0132* 143.5 [166.1]	*0.0009* 79.6 [71.2]
25-45	*0.0088* 198.6 [208.9]	*0.0511* 272.1 [286.2]	*0.0494* 293.8 [310.4]	*0.0229* 293.7 [298.6]	*0.0379* 271.5 [304.3]	*0.0272* 207.3 [245.7]	*0.0027* 116.8 [127.9]
45-55	*0.0025* 257.5 [315.2]	*0.0175* 352.2 [357.4]	*0.023* 359.1 [373.6]	*0.0119* 350.8 [356.4]	*0.0232* 321.6 [350.5]	*0.0192* 259.9 [276.8]	*0.0027* 161.8 [219.3]
55-75	*0.0024* 250.5 [216.6]	*0.0256* 432.6 [468.9]	*0.0399* 482.2 [505.5]	*0.0237* 477.8 [498.1]	*0.0502* 437 [472.4]	*0.0492* 332.7 [388.1]	*0.009* 181.1 [215.1]
75-95	*0.0008* 340 [362.5]	*0.0128* 642.5 [713.1]	*0.0273* 648.5 [696.3]	*0.0175* 700.8 [817.1]	*0.0507* 701 [840.3]	*0.0689* 522.9 [682.9]	*0.022* 303.1 [431.2]
95-100	*≤0.00006* -	*0.0009* 1029.3 [1192.6]	*0.0035* 1170.6 [1265.4]	*0.0025* 1080.7 [1071.6]	*0.009* 1171.7 [1366.7]	*0.0214* 913.1 [1184.4]	*0.0126* 506.1 [720]

NPC percentiles(42,46)

Sample SD percentiles	Sample mean percentiles						
	0-5	5-25	25-45	45-55	55-75	75-95	95-100
0-5	*0.0135* 111.9 [114.9]	*0.021* 157.5 [163.9]	*0.009* 155.8 [152.7]	*0.0027* 165.1 [162.8]	*0.0027* 152 [168.5]	*0.001* 110.6 [107.4]	*≤0.00006* -
5-25	*0.0217* 155.5 [169.1]	*0.0707* 230.8 [245.8]	*0.0485* 249.8 [276.2]	*0.0188* 236.7 [253.9]	*0.0262* 216.2 [227.4]	*0.0132* 158.3 [188.4]	*0.0009* 65.6 [74.7]
25-45	*0.009* 181.8 [184]	*0.0493* 299.9 [320.3]	*0.0505* 331.8 [352.1]	*0.0226* 326.8 [329.9]	*0.0385* 285.7 [325]	*0.0271* 224.9 [249.8]	*0.0029* 112.5 [111.8]
45-55	*0.0024* 227.9 [255.6]	*0.0186* 344.9 [376.2]	*0.0226* 394.8 [434.1]	*0.0121* 410.1 [414]	*0.023* 345.7 [389.7]	*0.0187* 264.1 [327.8]	*0.0025* 139.3 [155.6]
55-75	*0.0027* 212.9 [211]	*0.0256* 392.9 [446.5]	*0.0402* 468.2 [488.8]	*0.0234* 488.5 [524.1]	*0.0503* 461.3 [523.3]	*0.0489* 334.5 [384.8]	*0.0089* 191.8 [247.9]
75-95	*0.0005* 245 [237.5]	*0.0136* 519.2 [599.1]	*0.0262* 587.7 [694.4]	*0.0181* 670.4 [845.2]	*0.05* 654.7 [773.5]	*0.0699* 486.2 [578.2]	*0.0216* 288.3 [342.4]
95-100	*≤0.00006* -	*0.001* 418.8 [432.1]	*0.003* 791.9 [873.4]	*0.0023* 853.3 [860.4]	*0.0094* 900.6 [1175.7]	*0.0212* 817.4 [1100]	*0.0131* 469.5 [586.4]

NPC Wilcoxon-Waerden

Sample SD percentiles	Sample mean percentiles						
	0-5	5-25	25-45	45-55	55-75	75-95	95-100
0-5	*0.0135* 100.3 [101.9]	*0.0213* 135.2 [132.7]	*0.0087* 145.5 [147.3]	*0.0026* 159.7 [156.8]	*0.0029* 129.2 [137]	*0.0009* 89.7 [86]	*≤0.00006* -
5-25	*0.0217* 148.6 [152.6]	*0.0711* 214.5 [219.5]	*0.0482* 232 [238.6]	*0.0186* 225.7 [243.5]	*0.0263* 207.7 [222.3]	*0.0131* 149.3 [172.5]	*0.0011* 84.5 [93.2]
25-45	*0.0092* 197.9 [204.3]	*0.0502* 281.6 [284.6]	*0.0505* 320.3 [337.9]	*0.0225* 303.8 [296.8]	*0.0381* 281.5 [322]	*0.0265* 210.5 [230.3]	*0.003* 115.9 [126.9]
45-55	*0.0024* 266.1 [302.6]	*0.0184* 344.3 [359]	*0.0229* 375.7 [389.3]	*0.0119* 386.2 [403.2]	*0.0232* 343.6 [384.6]	*0.0185* 263.8 [284.5]	*0.0028* 150.5 [153.6]
55-75	*0.0025* 220.4 [190.6]	*0.0251* 406.4 [434.3]	*0.0398* 484.7 [509.1]	*0.0235* 490.6 [520.2]	*0.0507* 461.9 [497]	*0.0496* 339.5 [419.4]	*0.0088* 178.6 [219.2]
75-95	*0.0007* 306.3 [337.2]	*0.0128* 568.8 [636.1]	*0.0266* 623.9 [673.3]	*0.0182* 690.3 [796.1]	*0.0499* 687.4 [818.8]	*0.0702* 481.5 [591]	*0.0216* 285.3 [354.1]
95-100	*≤0.00006* -	*0.0011* 771 [836.8]	*0.0033* 1061.1 [1188.8]	*0.0026* 927.5 [913.1]	*0.0089* 1046.5 [1364.3]	*0.0212* 862.7 [1128.1]	*0.0128* 494 [639.2]

Table 4.8: In-control conditional performance of the Wilcoxon, Van der Waerden, NPC percentiles(42,46) and NPC Wilcoxon-Waerden charts under the Gamma$(4, 2)$ distribution shifted by -2.

4.5 Multi-aspects NPC control charts for location-scale monitoring

In this section, we present and analyze some distribution-free multi-aspect NPC control charts for the joint location-scale monitoring problem in the standards unknown case. All these charts can be easily implemented using the `npcharts` package.

4.5.1 Proposed control charts

Turning now our attention to the joint location-scale monitoring problem, there are two multi-aspect charts that we propose:

- NPC Lepage-Cucconi chart: obtained from the NPC (Fisher) combination between the Lepage and the Cucconi tests/charts. The control chart based on the Lepage statistic has been proposed in [27] and the control chart based on the Cucconi statistic has been proposed in [9]. Whereas both charts are distribution-free, their OOC performance can vary rather significantly under different distributions. In [9] the authors also show an interesting comparison between the two charts under the Normal and Laplace distributions. The main result was that the Cucconi chart is more effective at detecting scale changes and more effective at detecting location changes under the Normal distribution. On the other hand, the Lepage chart performs better when there is a pure location shift under the Laplace distribution. In order to obtain a control chart that combines the strengths of the Cucconi and Lepage charts, we propose the multi-aspect Cucconi-Lepage chart based on the NPC method with the Fisher combing function. The follow-up strategy after an OOC signal occurs discussed in [9], can still be applied with the NPC Cucconi-Lepage chart.

- NPC Wilcoxon-Klotz chart: obtained from the NPC (Fisher) combination of the Wilcoxon rank sum test with the Klotz test. With this chart, we take advantage of the good performance of the Klotz test under a normal distribution. Furthermore, considering that we are typically interested only in detecting scale changes with $\delta > 1$, a one-sided alternative can be used for the scale test (Klotz test). In the comparison with the Lepage and Cucconi charts, we consider the one-side strategy for the scale test to be "fair" because also the other charts, as shown in Table 4.9, cannot reliably detect a scale compression ($\delta < 1$). Moreover, both Lepage and Cucconi charts cannot be modified to use a single alternative for the scale test, thus the flexibility of NPC here plays an important advantage.

θ	δ	Cucconi	Lepage
0	1	490.6 [834.6]	506 [735.2]
		18, 238, 1718	19, 266, 1761
0	0.9	1607.5 [2917.2]	1246.2 [1858.3]
		36, 690, 5972	41, 658, 4502
0	0.8	6460.6 [10322.4]	3136.6 [5001.6]
		121, 2480, 29457	96, 1532, 10884
0	0.7	22599.8 [19627.9]	6148.9 [9215.8]
		548, 15307, 50001	145, 2772, 25013
0	0.6	42572.2 [14941]	7542.9 [11526.8]
		4996, 50001, 50001	118, 2762, 36463
0	0.5	49166.1 [5378.8]	5164 [10005]
		50001, 50001, 50001	57, 1364, 28525

Table 4.9: Unconditional performance of the Cucconi and Lepage charts under a compression of the scale for a standard normal process distribution. The first row of each of the cells shows the ARL and [SDRL] values, while the second row shows the 5th, 50th and 95th percentiles of the run length distribution (in this order).

4.5.2 Unconditional IC/OOC performance analysis

In order to investigate the performance of proposed multi-aspect charts, we compared them with the Cucconi and Lepage charts under the location-scale model described in Section 3.2.2 with several distributions and at different combinations of θ and δ.

Before introducing the evaluation methodology, it is important to consider that both the Lepage and Cucconi charts have only a single version that is equally sensitive to positive and negative location shifts under symmetric distributions, and, as mentioned earlier, cannot detect scale compressions. By contrast, the NPC Wilcoxon-Klotz chart is more flexible and can be configured to use a one-sided alternative for the location test and/or for the scale test. Using a two-sided alternative for the scale test, however, makes little sense because, even with a two-sided scale test, the NPC Wilcoxon-Klotz chart is not able to detect scale compressions as the other charts.

Methodology

First, we calibrate control charts for $ARL_0 = 500$, $n_1 = 100$ and $n_2 = 5$ using Algorithm 3 and Algorithm 4 under the standard normal distribution. A two-sided alternative is used in the location test of the NPC Wilcoxon-Klotz chart[2].

After the calibration, we use Algorithm 2 with 50000 simulations in order to assess the uncoditional IC/OOC performance with different combinations of θ and δ, and under the following distributions (also shown in Figure 4.1):

- Standard normal: represents a thin-tailed and symmetric distribution.

- Laplace $(0, 1/\sqrt{2})$: represents a heavy-tailed and symmetric distribution.

---

[2]In order to use a two-sided alternative, the Wilcoxon statistic is normalized by subtracting the mean

- Gamma(shape $= 4$, scale $= 2$) shifted by -2 so that the mean is equal to 0. This distribution, also proposed in [19] for valve leakage monitoring, has been considered to assess performance with asymmetric data. Because of the asymmetry, performance may be different under positive and negative location shifts of the same magnitude. Therefore, with this distribution, both positive and negative location shifts are considered.

All these distributions have mean equal to 0 and standard deviation equal to 1. This is important since it allows us to interpret θ as the location shift parameter with respect to the IC distribution independently from the value of the scale enlargement parameter δ. Moreover, having normalized distributions, allows us to compare performance in a fair manner also across different distributions.

For NPC charts, $B = 15000$ permutations are used to estimate the p-values.

Results

The results are shown in Table 4.10, Table 4.11, Table 4.12 and Table 4.5. The result format for all the tables is the following: the first row of each of the cells shows the ARL and [SDRL] values, while the second row shows the 5th, 50th and 95th percentiles of the run length distribution (in this order).

Under the standard normal distribution (Table 4.10) the NPC Wilcoxon-Klotz chart performs considerably better than all the other charts with no exceptions. The performance of the NPC Cucconi-Lepage chart is in the middle between the performance of the Cucconi and the Lepage charts, but closer to the best performing of the two, namely the Cucconi chart.

With the Laplace distribution (Table 4.11), the performance of the NPC Wilcoxon-Klotz chart seems close to that of the Cucconi chart. For all values of δ, the NPC Wilcoxon-Klotz seems slightly more performing than the Cucconi chart when $0 \leq \theta \leq 0.5$ (with just an exception when $\theta = 0$ and $\delta = 1.25$). When instead $1 \leq \theta \leq 2$, the Cucconi chart becomes slightly better than the Wilcoxon-Klotz chart. The performance pattern of the NPC Cucconi-Lepage with the Laplace distribution remains the same as seen with the normal distribution: the ARL is closer to the minimum between Cucconi and the Lepage charts.

Under the Gamma distribution (Table 4.12 and Table 4.13), with both positive and negative location shifts, the NPC Wilcoxon-Klotz charts outperform all the other charts with just one exception: the Laplace chart seems the most sensible to pure positive location shifts. The performance pattern of the NPC Cucconi-Lepage chart is the same as seen with the previous distributions.

θ	δ	Cucconi	Lepage	NPC Cucconi-Lepage	NPC Wilcoxon-Klotz
0	1	501.5 [821.3]	504.6 [728.9]	497.5 [759.3]	504.9 [819.5]
		15, 245, 1800	18, 272, 1747	17, 254, 1749	16, 247, 1823
0.25	1	260.8 [466.4]	257.5 [415.3]	258.2 [446.2]	248.5 [451]
		8, 119, 951	8, 126, 930	8, 122, 929	7, 110, 916
0.5	1	70.6 [121]	67.9 [111]	68.5 [119.9]	62.2 [109.5]
		3, 35, 250	3, 35, 237	3, 34, 240	2, 30, 222
1	1	7.7 [9.3]	7.6 [8.8]	7.6 [9.3]	6.6 [8.2]
		1, 5, 24	1, 5, 24	1, 5, 24	1, 4, 21
1.5	1	2.1 [1.7]	2.2 [1.8]	2.1 [1.7]	1.9 [1.5]
		1, 2, 5	1, 2, 6	1, 2, 5	1, 1, 5
2	1	1.2 [0.5]	1.2 [0.5]	1.2 [0.5]	1.1 [0.4]
		1, 1, 2	1, 1, 2	1, 1, 2	1, 1, 2
0	1.25	74.6 [96.4]	102.2 [124.8]	85.2 [106.1]	71.1 [93.8]
		4, 43, 249	5, 62, 335	4, 51, 281	3, 40, 239
0.25	1.25	54 [69.7]	69.7 [88.9]	60.4 [77.7]	49.4 [66.8]
		3, 31, 180	3, 41, 231	3, 35, 203	2, 28, 167
0.5	1.25	26.1 [32.9]	30.8 [38.6]	27.6 [34.8]	22.9 [30.4]
		2, 15, 86	2, 18, 101	2, 17, 89	1, 13, 76
1	1.25	6.2 [6.6]	6.8 [7.1]	6.4 [6.7]	5.2 [5.4]
		1, 4, 19	1, 4, 20	1, 4, 19	1, 3, 15
1.5	1.25	2.3 [1.9]	2.5 [2.1]	2.4 [1.9]	2 [1.6]
		1, 2, 6	1, 2, 7	1, 2, 6	1, 1, 5
2	1.25	1.4 [0.7]	1.4 [0.8]	1.4 [0.7]	1.2 [0.6]
		1, 1, 3	1, 1, 3	1, 1, 3	1, 1, 2
0	1.5	24.2 [27.3]	37.5 [42.1]	29.4 [33.5]	21 [24.7]
		2, 15, 76	2, 24, 118	2, 19, 93	1, 13, 67
0.25	1.5	20.5 [23]	30.3 [34.1]	24.1 [27.2]	17.6 [20.7]
		1, 13, 65	2, 19, 96	2, 15, 76	1, 11, 56
0.5	1.5	13.6 [15.2]	18 [20.2]	15.2 [16.8]	11.3 [12.7]
		1, 9, 42	1, 12, 57	1, 10, 47	1, 7, 35
1	1.5	5.2 [5.1]	6 [6]	5.5 [5.5]	4.3 [4.2]
		1, 4, 15	1, 4, 18	1, 4, 16	1, 3, 12
1.5	1.5	2.5 [2]	2.7 [2.2]	2.5 [2.1]	2.1 [1.6]
		1, 2, 6	1, 2, 7	1, 2, 7	1, 2, 5
2	1.5	1.5 [0.9]	1.6 [1]	1.5 [0.9]	1.4 [0.7]
		1, 1, 3	1, 1, 4	1, 1, 3	1, 1, 3
0	1.75	11.7 [12.4]	19.1 [20.3]	14.4 [15.1]	9.4 [10.1]
		1, 8, 35	1, 13, 59	1, 10, 44	1, 6, 29
0.25	1.75	10.6 [11]	16.7 [17.6]	13 [13.6]	8.6 [9.1]
		1, 7, 32	1, 11, 51	1, 9, 39	1, 6, 26
0.5	1.75	8.4 [8.6]	12 [12.6]	9.8 [10.2]	6.8 [6.9]
		1, 6, 25	1, 8, 36	1, 7, 30	1, 5, 20
1	1.75	4.4 [4.1]	5.5 [5.3]	4.8 [4.6]	3.6 [3.3]
		1, 3, 12	1, 4, 16	1, 3, 14	1, 3, 10
1.5	1.75	2.5 [2]	2.8 [2.3]	2.6 [2.1]	2.1 [1.6]
		1, 2, 6	1, 2, 7	1, 2, 7	1, 2, 5
2	1.75	1.6 [1]	1.8 [1.2]	1.7 [1.1]	1.4 [0.8]
		1, 1, 4	1, 1, 4	1, 1, 4	1, 1, 3

Table 4.10: Unconditional performance of the Cucconi, Lepage, NPC Cucconi-Lepage and NPC Wilcoxon-Klotz charts under the standard normal distribution.

θ	δ	Cucconi	Lepage	NPC Cucconi-Lepage	NPC Wilcoxon-Klotz
0	1	499.1 [829.9]	498.6 [695.1]	502.7 [802.8]	503.9 [847.1]
		15, 243, 1782	19, 271, 1740	16, 258, 1762	15, 248, 1799
0.25	1	292.1 [566.8]	272.5 [455.8]	278.5 [516.8]	289.6 [590.2]
		8, 125, 1087	8, 123, 1024	7, 121, 1044	7, 120, 1092
0.5	1	82.5 [186.2]	66.5 [134.8]	72.1 [165.4]	79 [178.2]
		2, 32, 310	2, 28, 247	2, 29, 272	2, 30, 297
1	1	6.1 [10.4]	5 [7]	5.4 [8.4]	6.3 [10]
		1, 3, 20	1, 3, 16	1, 3, 17	1, 3, 21
1.5	1	1.6 [1.3]	1.6 [1.1]	1.6 [1.1]	1.7 [1.5]
		1, 1, 4	1, 1, 4	1, 1, 4	1, 1, 4
2	1	1.1 [0.4]	1.1 [0.4]	1.1 [0.4]	1.1 [0.4]
		1, 1, 2	1, 1, 2	1, 1, 2	1, 1, 2
0	1.25	121.3 [167.9]	151.2 [193.3]	131.7 [173]	123.5 [177.9]
		5, 68, 412	7, 88, 507	5, 76, 439	5, 67, 430
0.25	1.25	84.3 [128.9]	98 [136.8]	88 [126.5]	82.1 [124.9]
		3, 44, 295	4, 53, 343	4, 47, 305	3, 42, 290
0.5	1.25	34.8 [55.2]	35.4 [54.5]	34.7 [55.6]	33.6 [55.9]
		2, 18, 124	2, 18, 123	2, 18, 121	2, 17, 117
1	1.25	5.2 [6.9]	4.9 [5.9]	5 [6.2]	5.3 [6.8]
		1, 3, 16	1, 3, 15	1, 3, 15	1, 3, 17
1.5	1.25	1.7 [1.3]	1.8 [1.3]	1.7 [1.3]	1.8 [1.5]
		1, 1, 4	1, 1, 4	1, 1, 4	1, 1, 4
2	1.25	1.2 [0.5]	1.2 [0.5]	1.2 [0.5]	1.2 [0.5]
		1, 1, 2	1, 1, 2	1, 1, 2	1, 1, 2
0	1.5	47.1 [58.3]	66.1 [79.3]	54.5 [68]	46.8 [61.5]
		2, 28, 154	3, 41, 214	3, 33, 177	2, 27, 158
0.25	1.5	37 [47.4]	47.7 [61.4]	40.7 [52.1]	36.1 [49.2]
		2, 22, 122	2, 28, 158	2, 24, 135	2, 20, 123
0.5	1.5	19.4 [26.4]	22.7 [30.9]	20.6 [27.9]	18.9 [26]
		1, 11, 66	1, 13, 76	1, 12, 68	1, 11, 63
1	1.5	4.6 [5.4]	4.8 [5.2]	4.6 [5.1]	4.7 [5.4]
		1, 3, 14	1, 3, 14	1, 3, 14	1, 3, 14
1.5	1.5	1.8 [1.4]	1.9 [1.5]	1.9 [1.4]	1.9 [1.5]
		1, 1, 5	1, 1, 5	1, 1, 5	1, 1, 5
2	1.5	1.2 [0.6]	1.3 [0.6]	1.3 [0.6]	1.3 [0.6]
		1, 1, 2	1, 1, 3	1, 1, 2	1, 1, 2
0	1.75	24 [27.8]	36.4 [41.3]	28.3 [32.3]	23.1 [28.2]
		2, 15, 77	2, 23, 115	2, 18, 90	1, 14, 75
0.25	1.75	20.1 [23.9]	28.7 [33.4]	23.2 [27]	19.4 [24.1]
		1, 12, 65	2, 18, 92	1, 15, 74	1, 12, 64
0.5	1.75	12.7 [15.1]	16.2 [19.4]	14.1 [17]	12.1 [14.6]
		1, 8, 40	1, 10, 52	1, 9, 45	1, 7, 39
1	1.75	4.2 [4.3]	4.6 [4.7]	4.3 [4.4]	4.2 [4.4]
		1, 3, 12	1, 3, 14	1, 3, 12	1, 3, 12
1.5	1.75	1.9 [1.4]	2 [1.6]	1.9 [1.5]	2 [1.5]
		1, 1, 5	1, 1, 5	1, 1, 5	1, 1, 5
2	1.75	1.3 [0.7]	1.4 [0.7]	1.3 [0.7]	1.3 [0.7]
		1, 1, 3	1, 1, 3	1, 1, 3	1, 1, 3

Table 4.11: Unconditional performance of the Cucconi, Lepage, NPC Cucconi-Lepage and NPC Wilcoxon-Klotz charts under the Laplace$(0, 1/\sqrt{2})$ distribution.

θ	δ	Cucconi	Lepage	NPC Cucconi-Lepage	NPC Wilcoxon-Klotz
0	1	494.1 [807.3]	500.7 [717.2]	501.4 [767.2]	504.8 [852.7]
		16, 245, 1768	18, 272, 1732	17, 256, 1782	15, 244, 1811
0.25	1	508.9 [1116.4]	425.4 [763.4]	467.4 [959.1]	476 [1038.7]
		12, 201, 1918	11, 188, 1592	11, 193, 1742	10, 182, 1804
0.5	1	160.9 [334.9]	132.7 [255.6]	142.4 [282.9]	139.7 [312.4]
		4, 66, 598	4, 59, 482	4, 61, 530	4, 55, 518
1	1	16.3 [27.8]	13.9 [22.7]	14.8 [23.6]	14 [25]
		1, 8, 57	1, 7, 47	1, 8, 51	1, 7, 49
1.5	1	2.9 [3.5]	2.6 [2.7]	2.7 [3.1]	2.7 [3.2]
		1, 2, 8	1, 2, 7	1, 2, 8	1, 2, 8
2	1	1.2 [0.6]	1.1 [0.5]	1.2 [0.5]	1.2 [0.6]
		1, 1, 2	1, 1, 2	1, 1, 2	1, 1, 2
0	1.25	54.8 [66.6]	78.1 [92]	64.5 [76.2]	45.3 [54.5]
		3, 33, 176	4, 49, 251	3, 40, 209	2, 27, 147
0.25	1.25	92.2 [129.6]	114.7 [153.9]	100.6 [135.7]	84.2 [117.2]
		4, 51, 316	5, 65, 387	4, 56, 340	3, 46, 289
0.5	1.25	61.7 [94.6]	65.1 [94]	61.1 [91.4]	54.3 [84.7]
		3, 32, 215	3, 35, 224	3, 33, 209	2, 28, 193
1	1.25	13.2 [18.2]	12.8 [16.7]	12.5 [16.8]	10.8 [14.9]
		1, 8, 43	1, 8, 42	1, 7, 41	1, 6, 35
1.5	1.25	3.4 [3.6]	3.3 [3.4]	3.3 [3.5]	3 [3]
		1, 2, 10	1, 2, 9	1, 2, 10	1, 2, 8
2	1.25	1.4 [0.9]	1.4 [0.8]	1.4 [0.9]	1.4 [0.8]
		1, 1, 3	1, 1, 3	1, 1, 3	1, 1, 3
0	1.5	17.1 [18.3]	26.9 [28.8]	20.7 [22.2]	12.9 [13.7]
		1, 11, 53	2, 18, 82	1, 14, 64	1, 9, 40
0.25	1.5	26.6 [30.2]	40.3 [45.9]	31.8 [36.3]	21.6 [25.3]
		2, 17, 84	2, 26, 127	2, 20, 101	1, 13, 69
0.5	1.5	27 [33.5]	34.8 [43.1]	30.2 [37.3]	23.1 [29.3]
		2, 16, 88	2, 21, 115	2, 18, 98	1, 14, 76
1	1.5	10.9 [13.1]	11.6 [13.6]	10.9 [12.8]	9.2 [11]
		1, 7, 35	1, 7, 36	1, 7, 34	1, 6, 29
1.5	1.5	3.8 [3.9]	3.9 [4]	3.8 [3.8]	3.2 [3.2]
		1, 3, 11	1, 3, 11	1, 3, 11	1, 2, 9
2	1.5	1.7 [1.3]	1.8 [1.3]	1.7 [1.2]	1.6 [1.1]
		1, 1, 4	1, 1, 4	1, 1, 4	1, 1, 4
0	1.75	8.7 [8.7]	13.8 [14.1]	10.6 [10.7]	6.3 [6.1]
		1, 6, 26	1, 9, 42	1, 7, 32	1, 4, 18
0.25	1.75	11.7 [12.3]	19.1 [20.1]	14.7 [15.4]	8.8 [9.2]
		1, 8, 36	1, 13, 59	1, 10, 45	1, 6, 27
0.5	1.75	13.2 [14.6]	20 [22]	15.9 [17.7]	10.4 [11.4]
		1, 9, 41	1, 13, 62	1, 10, 50	1, 7, 32
1	1.75	8.8 [9.7]	10.3 [11.3]	9.2 [10]	7.3 [8]
		1, 6, 27	1, 7, 32	1, 6, 28	1, 5, 22
1.5	1.75	4 [3.9]	4.3 [4.2]	4.1 [4]	3.4 [3.3]
		1, 3, 12	1, 3, 12	1, 3, 12	1, 2, 10
2	1.75	2.1 [1.6]	2.1 [1.7]	2.1 [1.6]	1.8 [1.3]
		1, 1, 5	1, 2, 5	1, 1, 5	1, 1, 4

Table 4.12: Unconditional performance of the Cucconi, Lepage, NPC Cucconi-Lepage and NPC Wilcoxon-Klotz charts under the Gamma$(4, 1/2)$ distribution shifted by -2.

θ	δ	Cucconi	Lepage	NPC Cucconi-Lepage	NPC Wilcoxon-Klotz
0	1	505.7 [863.9]	504.6 [714.7]	499.3 [784.8]	503.8 [850.3]
		16, 249, 1798	18, 270, 1766	17, 255, 1775	16, 245, 1828
-0.25	1	88.5 [124.3]	104.2 [142.4]	94.4 [132.4]	73.2 [103.4]
		4, 49, 302	4, 58, 355	4, 52, 320	3, 40, 253
-0.5	1	17.6 [20.4]	20.6 [24.1]	18.6 [21.4]	13.3 [14.9]
		1, 11, 56	1, 13, 65	1, 12, 59	1, 9, 42
-1	1	3 [2.6]	3.3 [2.9]	3.1 [2.7]	2.4 [1.9]
		1, 2, 8	1, 2, 9	1, 2, 8	1, 2, 6
-1.5	1	1.5 [0.8]	1.6 [1]	1.5 [0.9]	1.3 [0.6]
		1, 1, 3	1, 1, 3	1, 1, 3	1, 1, 3
-2	1	1.1 [0.4]	1.2 [0.4]	1.1 [0.4]	1.1 [0.3]
		1, 1, 2	1, 1, 2	1, 1, 2	1, 1, 2
0	1.25	54.9 [66.2]	78.3 [92.9]	64.3 [77.1]	45.3 [55.1]
		3, 34, 177	4, 48, 251	3, 39, 208	2, 27, 148
-0.25	1.25	20.5 [22.7]	28 [30.6]	23.5 [26]	15.5 [16.8]
		1, 13, 64	2, 18, 88	2, 15, 73	1, 10, 48
-0.5	1.25	8.7 [8.7]	10.9 [11.1]	9.5 [9.7]	6.4 [6.3]
		1, 6, 26	1, 7, 33	1, 7, 28	1, 4, 19
-1	1.25	2.8 [2.3]	3.2 [2.7]	3 [2.5]	2.2 [1.7]
		1, 2, 7	1, 2, 9	1, 2, 8	1, 2, 6
-1.5	1.25	1.6 [1]	1.7 [1.1]	1.6 [1]	1.4 [0.7]
		1, 1, 3	1, 1, 4	1, 1, 4	1, 1, 3
-2	1.25	1.2 [0.5]	1.3 [0.6]	1.2 [0.5]	1.1 [0.3]
		1, 1, 2	1, 1, 2	1, 1, 2	1, 1, 2
0	1.5	17 [18.1]	26.9 [28.9]	20.8 [22.1]	13 [13.8]
		1, 11, 52	2, 18, 83	1, 14, 63	1, 9, 40
-0.25	1.5	9.5 [9.6]	13.8 [14.1]	11.5 [11.7]	7.1 [7]
		1, 7, 28	1, 9, 41	1, 8, 34	1, 5, 21
-0.5	1.5	5.6 [5.2]	7.5 [7.3]	6.4 [6]	4.2 [3.8]
		1, 4, 16	1, 5, 22	1, 4, 18	1, 3, 12
-1	1.5	2.6 [2.1]	3.1 [2.6]	2.8 [2.3]	2.1 [1.5]
		1, 2, 7	1, 2, 8	1, 2, 7	1, 2, 5
-1.5	1.5	1.6 [1]	1.8 [1.2]	1.7 [1.1]	1.4 [0.7]
		1, 1, 4	1, 1, 4	1, 1, 4	1, 1, 3
-2	1.5	1.3 [0.6]	1.4 [0.7]	1.3 [0.7]	1.2 [0.4]
		1, 1, 2	1, 1, 3	1, 1, 3	1, 1, 2
0	1.75	8.6 [8.6]	13.8 [14]	10.6 [10.7]	6.3 [6.2]
		1, 6, 25	1, 9, 41	1, 7, 32	1, 4, 19
-0.25	1.75	5.9 [5.6]	8.9 [8.7]	7.2 [6.9]	4.4 [4.1]
		1, 4, 17	1, 6, 26	1, 5, 21	1, 3, 12
-0.5	1.75	4.2 [3.7]	5.8 [5.4]	4.9 [4.5]	3.1 [2.7]
		1, 3, 12	1, 4, 16	1, 3, 14	1, 2, 8
-1	1.75	2.4 [1.8]	3 [2.5]	2.6 [2.1]	1.9 [1.3]
		1, 2, 6	1, 2, 8	1, 2, 7	1, 1, 5
-1.5	1.75	1.6 [1]	1.9 [1.3]	1.8 [1.2]	1.4 [0.8]
		1, 1, 4	1, 1, 5	1, 1, 4	1, 1, 3
-2	1.75	1.3 [0.7]	1.4 [0.8]	1.4 [0.7]	1.2 [0.5]
		1, 1, 3	1, 1, 3	1, 1, 3	1, 1, 2

Table 4.13: Unconditional performance (negative shifts) of the Cucconi, Lepage, NPC Cucconi-Lepage and NPC Wilcoxon-Klotz charts under the Gamma$(4, 1/2)$ distribution shifted by -2.

4.5.3 Conditional IC performance analysis

Besides evaluating the uncoditional performance, we also studied the performance conditional to specific categories of reference samples.

Methodology

After a calibration with $ARL_0 = 500$, the methodology is the same proposed in [27]. We use Algorithm 2 with 50000 run length simulations. Each run length simulation is then classified according to two criteria, the value of the reference sample mean and the value of the reference sample standard deviation. For both of these two criteria, 7 percentile-based bins are used:

- 5th percentile or lower: very high downward bias

- between the 5th and the 25th percentiles: moderately high downward bias

- between the 25th and the 45th percentiles: low downward bias

- between the 45th and the 55th percentiles

- between the 55th and the 75th percentiles: low upward bias

- between the 75th and the 95th percentiles: moderately high upward bias

- beyond the 95th percentiles: high upward bias

Therefore, each of the 50000 run length simulations is categorized in a 7×7 way table depending on the value of the corresponding reference sample mean and SD. For each of the 49 bins, we record the proportion of simulations as well as the sample mean and sample SD of the simulated run lengths that belong to the bin. These last two value are, respectively, an estimate of the conditional ARL_0 and $SDRL_0$ given the specified percentiles intervals of the reference sample mean and SD with respect to their distributions.

The process distributions and sample sizes adopted in this study are the same used in the unconditional evaluation (Section 4.5.2).

Results

The results are shown in Table 4.14 (Normal), Table 4.15 (Laplace) and Table 4.16 (Gamma). For each cell, the top value (in italics) shows the proportion, the bottom left value shows the conditional ARL_0 and the bottom right value (within the square brackets) shows the conditional $SDRL_0$. For the colored cells, the conditional ARL_0 value is within the "safe" interval $[250, 900]$ proposed in [9].

In general, for all control charts, we see the same patterns described in [9]. Given a range of percentiles of the reference sample mean, the larger the SD, the larger the conditional ARL_0. By contrast, given a specific range of percentiles of the SD of the reference sample, the conditional ARL_0 initially increases with an increase in the sample mean, reaches a peak around between the 45 and the 55 percentiles of the sample mean distribution, and then decreases.

Comparing the conditional performance between different charts, the results show that, for all distributions, the Lepage chart is the least sensitive to reference samples that have extreme mean or SD. This is particularly visible if we look at

the number of cells that lies in the safe interval. The NPC Wilcoxon-Klotz chart is instead the most sensitive with all distributions and is closely followed by the Cucconi chart. The conditional performance of the NPC Lepage-Cucconi chart seems in between the Cucconi and Lepage charts.

Cucconi

Sample SD percentiles	\ Sample mean percentiles 0-5	5-25	25-45	45-55	55-75	75-95	95-100
0-5	0.0026 / 72.8 [70.2]	0.0102 / 100 [103]	0.0097 / 114.9 [128.2]	0.0045 / 107.8 [107.8]	0.0102 / 118.5 [118.9]	0.0104 / 111.5 [119.3]	0.0025 / 80.2 [118.4]
5-25	0.0099 / 133 [143.8]	0.0386 / 181.3 [189.2]	0.0401 / 218.7 [223.4]	0.0206 / 210.2 [235.3]	0.04 / 224.9 [231.2]	0.04 / 187.2 [194.6]	0.0107 / 133.1 [150.5]
25-45	0.0098 / 176.8 [179.1]	0.0401 / 272.8 [297.6]	0.04 / 317.1 [324.3]	0.02 / 326.1 [326.4]	0.0406 / 335.7 [353.4]	0.0394 / 299.9 [324.6]	0.01 / 186.4 [210.7]
45-55	0.0048 / 240.7 [270.9]	0.0196 / 368 [392.7]	0.0202 / 421.8 [438.6]	0.01 / 454.1 [431.9]	0.0205 / 410.3 [426.4]	0.0204 / 357.9 [407.6]	0.0045 / 220.3 [227]
55-75	0.0093 / 302.2 [386.1]	0.0412 / 502.4 [525.1]	0.0398 / 562.1 [561.6]	0.0207 / 600.5 [621.5]	0.0392 / 551 [574]	0.0401 / 479 [501.3]	0.0096 / 331.7 [381.7]
75-95	0.0107 / 486.6 [514.6]	0.0404 / 786.6 [873.4]	0.04 / 937.3 [1041.2]	0.0192 / 970.7 [1023.4]	0.0393 / 911.4 [946.2]	0.0404 / 813.8 [922.5]	0.0099 / 503.1 [533.5]
95-100	0.0029 / 831.3 [1034.5]	0.01 / 1610.1 [1863.4]	0.0101 / 2060.6 [2576.1]	0.0049 / 2189.4 [2250.2]	0.0102 / 2093.9 [2201.2]	0.0092 / 1884.4 [2492.3]	0.0027 / 1070.5 [1449.9]

Lepage

Sample SD percentiles	\ Sample mean percentiles 0-5	5-25	25-45	45-55	55-75	75-95	95-100
0-5	0.0023 / 95.9 [107.3]	0.01 / 131.8 [146.5]	0.0096 / 145.6 [150.9]	0.005 / 149.3 [157.1]	0.0101 / 157.3 [170.4]	0.0104 / 142.9 [150.9]	0.0025 / 89 [102]
5-25	0.0102 / 135.5 [149.7]	0.0405 / 216.6 [230.9]	0.04 / 270.6 [276.5]	0.0199 / 270.7 [281.3]	0.0397 / 280.1 [286.5]	0.0394 / 215.5 [223.8]	0.0104 / 155 [161.6]
25-45	0.0096 / 197.9 [199.7]	0.0397 / 309.9 [318.8]	0.0411 / 369 [392.2]	0.0191 / 394.8 [440.9]	0.0404 / 392.3 [399.4]	0.0403 / 328.3 [362.7]	0.0099 / 205 [217.4]
45-55	0.0053 / 255.2 [284]	0.0193 / 386.1 [425.9]	0.0197 / 467.6 [477.3]	0.0107 / 502.6 [513]	0.0205 / 474 [505.9]	0.02 / 395.8 [427.5]	0.0045 / 239.5 [253.2]
55-75	0.0095 / 300.5 [385]	0.0414 / 520.5 [533.2]	0.0398 / 609.1 [624]	0.0203 / 607.9 [642.9]	0.0393 / 599.7 [592.8]	0.0397 / 492.7 [533.7]	0.0099 / 359.4 [401.4]
75-95	0.01 / 448 [485]	0.0389 / 798.8 [835.2]	0.0401 / 888 [966]	0.02 / 972.8 [1111.1]	0.0396 / 886.1 [941.9]	0.0413 / 774.5 [886.1]	0.01 / 462.5 [470.7]
95-100	0.003 / 715.5 [782.5]	0.0102 / 1274.1 [1314.8]	0.0097 / 1747.7 [2228.5]	0.0048 / 1740.6 [1913.3]	0.0105 / 1674.2 [1924]	0.0089 / 1460.7 [1783.8]	0.0028 / 692.1 [709.6]

NPC Lepage-Cucconi

Sample SD percentiles	\ Sample mean percentiles 0-5	5-25	25-45	45-55	55-75	75-95	95-100
0-5	0.0026 / 99.3 [96.2]	0.0102 / 112.3 [123.1]	0.0099 / 123.3 [125.8]	0.0048 / 123.6 [126.2]	0.01 / 131.4 [135.3]	0.0103 / 112 [117.1]	0.0023 / 78.8 [87.3]
5-25	0.0103 / 134.9 [137.9]	0.0409 / 205.4 [214.5]	0.0401 / 226.4 [226.3]	0.019 / 239.9 [260.4]	0.0409 / 222.5 [235.6]	0.0395 / 204.1 [207.2]	0.0094 / 131.7 [146.8]
25-45	0.0101 / 203.4 [203.5]	0.0404 / 292.2 [315.3]	0.0402 / 339.8 [361.1]	0.0205 / 360.6 [372.5]	0.0383 / 345 [370.1]	0.0402 / 284 [291.3]	0.0104 / 210.8 [220.3]
45-55	0.0047 / 236.4 [252.1]	0.0198 / 402.8 [420.5]	0.0202 / 422.8 [438.4]	0.0099 / 459.8 [478.2]	0.0201 / 429.7 [457.8]	0.0207 / 353.6 [376.5]	0.0046 / 254.1 [274.1]
55-75	0.0097 / 310.6 [309.8]	0.0391 / 478.8 [515.2]	0.0395 / 575.3 [585]	0.0209 / 587.3 [582.5]	0.0404 / 572.8 [625.7]	0.0395 / 471.6 [504.2]	0.0109 / 319.6 [344.7]
75-95	0.0096 / 471.9 [525.3]	0.0396 / 758.1 [827.8]	0.0404 / 901.6 [984.4]	0.0198 / 978.1 [1098.8]	0.0401 / 946.5 [1033.6]	0.04 / 750.5 [855.2]	0.0105 / 468.7 [498.2]
95-100	0.0031 / 770.3 [802.6]	0.01 / 1587.3 [1741]	0.0097 / 1862.1 [2217.2]	0.0051 / 1828.3 [2115.6]	0.0103 / 1887.1 [2340.4]	0.0099 / 1616.5 [1990.2]	0.0019 / 796.6 [860.1]

NPC Wilcoxon-Klotz

Sample SD percentiles	\ Sample mean percentiles 0-5	5-25	25-45	45-55	55-75	75-95	95-100
0-5	0.0024 / 76.1 [78.9]	0.0095 / 100.9 [104.3]	0.0109 / 123.6 [131.1]	0.0044 / 112.8 [117.2]	0.0102 / 116.5 [138.4]	0.0101 / 95.6 [105.1]	0.0025 / 69 [69.9]
5-25	0.0096 / 123.4 [138.2]	0.0408 / 177.4 [186.9]	0.0395 / 209 [225]	0.0198 / 213.9 [226.6]	0.0398 / 218.7 [240]	0.04 / 183.8 [200.5]	0.0105 / 120.9 [134.1]
25-45	0.0099 / 182 [225.9]	0.0402 / 287.6 [289.6]	0.0399 / 326.3 [331.1]	0.0192 / 352.2 [360.3]	0.0399 / 337.5 [354.2]	0.0409 / 281.8 [295.5]	0.01 / 172 [206.8]
45-55	0.0048 / 234.1 [232.2]	0.0201 / 353.6 [345.7]	0.0196 / 428.8 [463.9]	0.0103 / 437 [467.4]	0.0207 / 426.8 [437.8]	0.0198 / 358.5 [372.5]	0.0047 / 231.6 [212.7]
55-75	0.0108 / 313.7 [362.5]	0.0406 / 487.4 [529.7]	0.0386 / 570.5 [603]	0.0203 / 595.2 [682.5]	0.04 / 555.7 [562.5]	0.0401 / 466.6 [495]	0.0096 / 305.6 [323.8]
75-95	0.0097 / 455.1 [513.7]	0.0387 / 800.1 [867.3]	0.0413 / 949.2 [1039.6]	0.0215 / 1008.4 [1162.6]	0.0397 / 910.5 [970.8]	0.0392 / 781.8 [854.4]	0.0098 / 478.5 [552.9]
95-100	0.0027 / 995.7 [1118.5]	0.0101 / 1728.5 [2580.3]	0.0102 / 2154.3 [2527.6]	0.0045 / 2226.5 [3112.8]	0.0096 / 2274.5 [2626.8]	0.0099 / 1570.5 [1954.7]	0.0029 / 981.2 [1146.1]

Table 4.14: In-control conditional performance of the Cucconi, Lepage, NPC Lepage-Cucconi and NPC Wilcoxon-Klotz charts under the standard normal distribution.

Cucconi

Sample SD percentiles	Sample mean percentiles						
	0-5	5-25	25-45	45-55	55-75	75-95	95-100
0-5	*0.0016* 83.2 [93]	*0.0093* 103.9 [109.6]	*0.0119* 131 [146.6]	*0.0059* 116.3 [125.3]	*0.0107* 126.9 [142]	*0.0095* 106.6 [113.4]	*0.0013* 84.8 [84.4]
5-25	*0.0082* 128.5 [141.1]	*0.0392* 197.9 [232]	*0.0409* 232.3 [244.5]	*0.0223* 236.4 [258.6]	*0.0423* 231.9 [252.7]	*0.0391* 196.9 [219.9]	*0.0079* 134.3 [166.7]
25-45	*0.0088* 206.1 [218.6]	*0.0393* 293.5 [318.2]	*0.0425* 373.8 [398.2]	*0.0408* 389.5 [426.7]	*0.0408* 372.8 [413.2]	*0.0404* 313.7 [352.5]	*0.0082* 200.1 [233.3]
45-55	*0.005* 265.6 [304.6]	*0.0203* 406.3 [498.8]	*0.0201* 486.5 [554.9]	*0.0094* 481.4 [528.1]	*0.0199* 451.2 [509.4]	*0.0206* 399.1 [436.7]	*0.0047* 224.4 [232.5]
55-75	*0.0107* 280.1 [323.3]	*0.0407* 494.6 [561.3]	*0.0385* 619.3 [712.7]	*0.0198* 617.4 [658.1]	*0.0394* 601.2 [649.6]	*0.0397* 510.5 [579.9]	*0.0112* 305.5 [321.9]
75-95	*0.0121* 468.7 [583.5]	*0.041* 765.8 [878.9]	*0.0372* 933.4 [1089.9]	*0.0188* 947.8 [1143.9]	*0.0394* 965.2 [1188.4]	*0.0403* 750 [854.1]	*0.0129* 479.7 [551.8]
95-100	*0.0035* 715.7 [778.3]	*0.0101* 1459.3 [2155.3]	*0.0088* 1847 [2133]	*0.004* 2057.9 [3088.9]	*0.0093* 1773.6 [2219.2]	*0.0104* 1464.9 [1693.9]	*0.0038* 837.7 [1096.2]

Lepage

Sample SD percentiles	Sample mean percentiles						
	0-5	5-25	25-45	45-55	55-75	75-95	95-100
0-5	*0.0015* 100.2 [110.2]	*0.0093* 142.5 [147]	*0.011* 173.9 [181.9]	*0.006* 189.9 [232.5]	*0.0111* 168.9 [180.9]	*0.0095* 143.9 [171.6]	*0.0015* 90.1 [85.3]
5-25	*0.0079* 131.1 [138.2]	*0.039* 234.7 [250.3]	*0.0416* 294.8 [313.8]	*0.0218* 297.2 [323.3]	*0.0428* 276.6 [293.4]	*0.0394* 237.1 [263.2]	*0.0076* 156.3 [179.7]
25-45	*0.0093* 224.2 [240.9]	*0.0395* 338.9 [374.7]	*0.0421* 434.7 [486.8]	*0.0204* 441 [470.9]	*0.0407* 427.8 [491.8]	*0.0398* 347.6 [379.2]	*0.0082* 213.7 [249.8]
45-55	*0.0049* 249.2 [303.4]	*0.0201* 437.9 [495.2]	*0.0203* 513.8 [592.1]	*0.0093* 608.1 [696.6]	*0.02* 491.9 [542.9]	*0.0207* 424.5 [465.3]	*0.0048* 240.3 [243.5]
55-75	*0.0107* 301.5 [346.6]	*0.0404* 520.1 [591.6]	*0.0393* 617.3 [688.5]	*0.0192* 654.2 [708.2]	*0.0394* 631.6 [698]	*0.0398* 521 [585.6]	*0.0112* 321.5 [387.4]
75-95	*0.012* 444.2 [537.4]	*0.0416* 713.3 [815.5]	*0.037* 872.6 [988.3]	*0.0191* 935 [1086.1]	*0.0373* 912.9 [1104.9]	*0.0405* 732.5 [904.7]	*0.0125* 466.9 [568.9]
95-100	*0.0037* 773.2 [863.7]	*0.0102* 1126.4 [1539.7]	*0.0088* 1442.2 [1797.1]	*0.0042* 1173.8 [1418]	*0.0087* 1449.2 [1764.3]	*0.0102* 1152 [1387.4]	*0.0042* 741.8 [824.9]

NPC Lepage-Cucconi

Sample SD percentiles	Sample mean percentiles						
	0-5	5-25	25-45	45-55	55-75	75-95	95-100
0-5	*0.0015* 103.4 [116.2]	*0.0094* 127.4 [146.7]	*0.0109* 164.7 [179.2]	*0.0061* 154.3 [165.6]	*0.011* 153.7 [171]	*0.0097* 125.2 [132.7]	*0.0014* 84.1 [106]
5-25	*0.0074* 137.3 [135.3]	*0.039* 216.2 [228.2]	*0.0427* 254.9 [284]	*0.0212* 271.7 [307.5]	*0.0423* 273.5 [290]	*0.0395* 215.6 [227.3]	*0.0081* 146.8 [174.2]
25-45	*0.0089* 198.9 [242.6]	*0.0403* 317.5 [348.7]	*0.0422* 382.6 [403.2]	*0.0203* 417.4 [458.3]	*0.0408* 376.4 [409.7]	*0.0393* 318.9 [355.9]	*0.0084* 202.6 [206.2]
45-55	*0.0047* 211.5 [253.7]	*0.0198* 385.1 [406.1]	*0.0204* 480.6 [548.4]	*0.0097* 532 [609.1]	*0.0209* 432.3 [474.8]	*0.0193* 385.2 [441.1]	*0.0052* 230.3 [231.9]
55-75	*0.0105* 328.7 [371.4]	*0.0403* 503.1 [585.8]	*0.0391* 613.8 [686.6]	*0.0196* 633 [687.7]	*0.0394* 615.5 [695.3]	*0.0402* 506.9 [570.4]	*0.011* 291.1 [349]
75-95	*0.0129* 428.1 [567.3]	*0.0361* 746.5 [922.5]	*0.0361* 921.6 [1113.3]	*0.0186* 998.4 [1153.4]	*0.0371* 958.2 [1144.9]	*0.0417* 731.4 [841.8]	*0.0121* 457.3 [575.3]
95-100	*0.0041* 685.2 [809.7]	*0.0097* 1396 [1879.2]	*0.0087* 1717.7 [2116.4]	*0.0045* 1552.8 [1983.2]	*0.0087* 1693.1 [1993.8]	*0.0104* 1253.6 [1815.5]	*0.0039* 669.9 [856.3]

NPC Wilcoxon-Klotz

Sample SD percentiles	Sample mean percentiles						
	0-5	5-25	25-45	45-55	55-75	75-95	95-100
0-5	*0.0017* 54.5 [68]	*0.0097* 88.1 [98.3]	*0.0108* 104.2 [109.4]	*0.0058* 113.3 [125.2]	*0.0111* 109.9 [107]	*0.0093* 99.7 [102.3]	*0.0016* 72.8 [85.5]
5-25	*0.0073* 112.3 [114]	*0.039* 172.9 [179.7]	*0.0423* 214.4 [223.8]	*0.0215* 213.8 [232.7]	*0.0425* 226.9 [236.1]	*0.04* 170 [179.1]	*0.0074* 107.9 [126.2]
25-45	*0.0094* 163.1 [159.8]	*0.0416* 265.7 [277.6]	*0.0407* 347.2 [350.3]	*0.0204* 368.5 [371.2]	*0.0398* 345.5 [373.1]	*0.0386* 287.2 [307.7]	*0.0094* 156.8 [163.6]
45-55	*0.0048* 191.3 [189.9]	*0.0197* 347 [364.6]	*0.0199* 458.9 [467]	*0.0098* 477.8 [483.3]	*0.0205* 461.4 [490.5]	*0.0203* 366.4 [372.9]	*0.0051* 215.3 [257.7]
55-75	*0.0109* 253.3 [285.6]	*0.0403* 465 [509.4]	*0.0392* 620.3 [660.2]	*0.0194* 614.7 [654.7]	*0.0391* 591.6 [629.8]	*0.041* 474.5 [514.3]	*0.0102* 279.8 [297.2]
75-95	*0.0125* 405 [476.4]	*0.04* 755.2 [873.4]	*0.0383* 980.1 [1084.6]	*0.0188* 1004.9 [1134.7]	*0.0378* 967.7 [1090]	*0.0403* 740.6 [812.5]	*0.0123* 375.6 [441.6]
95-100	*0.0034* 744.4 [1104]	*0.0096* 1509 [1763.9]	*0.0088* 2158.5 [2987.1]	*0.0044* 1923.5 [2318.5]	*0.0093* 1892.6 [2089]	*0.0105* 1655 [2344.7]	*0.0041* 735 [808]

Table 4.15: In-control conditional performance of the Cucconi, Lepage, NPC Lepage-Cucconi and NPC Wilcoxon-Klotz charts under the Laplace $(0, 1/\sqrt{2})$ distribution.

Cucconi

Sample SD percentiles	Sample mean percentiles						
	0-5	5-25	25-45	45-55	55-75	75-95	95-100
0-5	*0.0134* 96 [98.2]	*0.0218* 130.8 [142.4]	*0.0091* 118.1 [119.8]	*0.0022* 121.4 [117.1]	*0.0026* 129.9 [143.7]	*0.0009* 87.7 [129.8]	*≤0.00006* -
5-25	*0.0213* 184.2 [206.1]	*0.0703* 236.3 [256.2]	*0.0478* 241.9 [252]	*0.0189* 230.6 [264.1]	*0.0273* 207 [242.1]	*0.0132* 156.5 [177.1]	*0.0012* 118.4 [134.9]
25-45	*0.0092* 302.2 [302]	*0.0493* 382.7 [407.9]	*0.0493* 368.7 [400]	*0.0232* 341 [376.9]	*0.0398* 333.5 [371.3]	*0.0263* 248.6 [302.6]	*0.0029* 144 [169.1]
45-55	*0.0029* 393.2 [417]	*0.018* 474.2 [497.3]	*0.0234* 488.1 [539.4]	*0.0116* 497.5 [620.3]	*0.0232* 486.7 [567.7]	*0.0183* 325.2 [364.3]	*0.0026* 163.5 [183]
55-75	*0.0025* 555.1 [670]	*0.0268* 619.3 [696.6]	*0.0397* 670.6 [745.7]	*0.0228* 632.1 [671]	*0.0495* 609.6 [706.2]	*0.0491* 436.3 [533.8]	*0.0097* 276.3 [322.9]
75-95	*0.0007* 513.5 [584.6]	*0.0126* 837.3 [939.4]	*0.0276* 965.2 [1188.1]	*0.0187* 1017.3 [1255.2]	*0.0491* 986 [1238.6]	*0.0708* 768.9 [1011.8]	*0.0206* 436.6 [638.6]
95-100	*≤0.00006* -	*0.0013* 1744.6 [2242.5]	*0.0031* 1585.4 [2452.6]	*0.0026* 1787.5 [2302.9]	*0.0085* 1666.1 [2250]	*0.0214* 1346.9 [1844.3]	*0.013* 876.7 [1528]

Lepage

Sample SD percentiles	Sample mean percentiles						
	0-5	5-25	25-45	45-55	55-75	75-95	95-100
0-5	*0.0134* 120.3 [134.1]	*0.021* 160.8 [177.1]	*0.0088* 158.5 [163.7]	*0.0027* 160.6 [168.4]	*0.0031* 145.5 [138]	*0.001* 110.1 [111.6]	*≤0.00006* -
5-25	*0.0221* 202 [220.4]	*0.0707* 277.4 [291.9]	*0.0499* 290.9 [316.9]	*0.0179* 285.4 [310.4]	*0.025* 250.4 [276.4]	*0.0134* 185.4 [196.4]	*0.0011* 103.1 [118.3]
25-45	*0.0091* 286.1 [296.4]	*0.0498* 404.6 [430.5]	*0.0498* 436.7 [491.8]	*0.023* 404.2 [432.3]	*0.0397* 387.4 [421.8]	*0.0256* 266.5 [322]	*0.0032* 179.9 [214.4]
45-55	*0.0024* 324.5 [288.3]	*0.0184* 510 [584.9]	*0.023* 530.4 [565.8]	*0.0118* 509.9 [550.1]	*0.0231* 468.9 [555.4]	*0.0189* 360.5 [391.2]	*0.0024* 168.3 [184.5]
55-75	*0.0023* 364.2 [371.3]	*0.0263* 569.1 [608.2]	*0.038* 635 [711.8]	*0.0231* 680.4 [794.4]	*0.0502* 633.6 [799.6]	*0.051* 457.2 [535.2]	*0.009* 270.3 [278.9]
75-95	*0.0008* 465.1 [594.6]	*0.0128* 675.8 [827]	*0.0274* 944.5 [1068]	*0.0185* 901.7 [1030.9]	*0.0502* 951.9 [1152.5]	*0.0691* 746.8 [976.2]	*0.0212* 387.2 [479.5]
95-100	*≤0.00006* -	*0.001* 1112.1 [1512.9]	*0.0031* 1470.8 [1883.5]	*0.003* 1347.9 [1700.4]	*0.0088* 1361 [1579.4]	*0.021* 1147.4 [1432.6]	*0.0131* 693.3 [974.8]

NPC Lepage-Cucconi

Sample SD percentiles	Sample mean percentiles						
	0-5	5-25	25-45	45-55	55-75	75-95	95-100
0-5	*0.0129* 114.3 [120.2]	*0.0213* 142.5 [141.8]	*0.009* 136.8 [135.8]	*0.0027* 139.2 [148.5]	*0.0031* 121.8 [136.7]	*0.0009* 94.6 [90.1]	*≤0.00006* -
5-25	*0.0221* 188.9 [200.7]	*0.0695* 250.3 [276.4]	*0.0483* 263.1 [278.9]	*0.0185* 263.2 [296.6]	*0.0265* 211.8 [249.6]	*0.014* 180 [203.9]	*0.0012* 77.3 [78.2]
25-45	*0.0092* 296.8 [345.5]	*0.0497* 368.2 [390.7]	*0.0502* 393.9 [427.6]	*0.0231* 383.4 [418.9]	*0.0391* 362.4 [401.7]	*0.026* 242 [277.2]	*0.0026* 151.7 [172.9]
45-55	*0.0024* 348.2 [361.1]	*0.0188* 464 [498.6]	*0.0241* 520.9 [552.4]	*0.0112* 469.3 [512]	*0.0225* 437 [506.7]	*0.0182* 356.6 [447.1]	*0.0027* 197.6 [223.7]
55-75	*0.0027* 428.5 [411.3]	*0.0266* 574.9 [627.6]	*0.0391* 653.2 [767.2]	*0.023* 681.2 [760]	*0.0501* 585.5 [728.7]	*0.0499* 462.1 [562.6]	*0.0087* 247.4 [278.2]
75-95	*0.0008* 453.5 [513.5]	*0.0132* 744 [880.8]	*0.0262* 936.1 [1136.9]	*0.0186* 908.1 [1176.7]	*0.0494* 954.7 [1165.6]	*0.0699* 715.9 [975]	*0.0219* 421 [588.4]
95-100	*≤0.00006* -	*0.0009* 1094.5 [1377.5]	*0.003* 1403.1 [1616.6]	*0.0029* 1888.4 [2376.5]	*0.0092* 1588.3 [2032.4]	*0.021* 1259 [1666.4]	*0.0129* 719.7 [977.6]

NPC Wilcoxon-Klotz

Sample SD percentiles	Sample mean percentiles						
	0-5	5-25	25-45	45-55	55-75	75-95	95-100
0-5	*0.0138* 91.5 [89.5]	*0.0207* 109.8 [111.5]	*0.0091* 114 [118.6]	*0.0026* 121.2 [122.3]	*0.0028* 100.1 [107]	*0.001* 63.2 [75.2]	*≤0.00006* -
5-25	*0.0211* 170.6 [176.7]	*0.0712* 213.6 [228.7]	*0.0489* 233 [259.1]	*0.0179* 214.8 [219.7]	*0.0271* 202.3 [223.9]	*0.0131* 151.5 [164.7]	*0.0008* 94.7 [138.6]
25-45	*0.0092* 283.8 [302.3]	*0.0492* 341.7 [349.5]	*0.0488* 362.1 [375.8]	*0.0243* 342.2 [376.5]	*0.0395* 327.2 [375.5]	*0.0261* 240 [274.3]	*0.0029* 192.9 [263]
45-55	*0.0021* 361.8 [422.8]	*0.0181* 468.7 [508]	*0.0235* 483.3 [486.8]	*0.0115* 464.5 [505.4]	*0.023* 438.6 [482.2]	*0.0188* 334.8 [393.4]	*0.003* 195.3 [262.1]
55-75	*0.0025* 462.5 [529.6]	*0.0265* 609.9 [651.6]	*0.0403* 639.9 [666.4]	*0.0232* 608 [629.3]	*0.05* 586.9 [689.7]	*0.0487* 430.9 [540.4]	*0.0087* 253.5 [349.8]
75-95	*0.0012* 563.8 [712.4]	*0.0135* 950.6 [1032.5]	*0.0261* 1047.4 [1138.6]	*0.0179* 1123.4 [1386.5]	*0.0492* 967.9 [1171.4]	*0.0711* 756 [1075.9]	*0.021* 409.8 [573.3]
95-100	*≤0.00006* -	*0.0009* 1412.5 [1917.4]	*0.0034* 1952.1 [2474.7]	*0.0026* 2037 [2230.9]	*0.0085* 1806.5 [2259.3]	*0.0211* 1393.6 [2020.1]	*0.0135* 684.2 [1100.9]

Table 4.16: In-control conditional performance of the Cucconi, Lepage, NPC Lepage-Cucconi and NPC Wilcoxon-Klotz charts under the Gamma$(4, 2)$ distribution shifted by -2.

Chapter 5

Conclusions

In this Thesis, we have achieved several findings and practical results concerning the proposed multi-aspect NPC scheme for Shewhart-type control charts. From a theoretical perspective, we answered important questions regarding the distribution-free-ness of multi-aspect NPC control charts: We have discovered that multi-aspect NPC control charts are not distribution-free in general and that they become distribution-free if they combine only linear rank tests.

From a practical perspective, we proposed and evaluated several distribution-free multi-aspect NPC charts for the univariate location and joint location-scale monitoring problems in case U. To provide calibration and evaluation routines for multi-aspect NPC charts, a high-performance R packet has been developed. Using the implemented algorithms, a comparative simulation study has been carried out to compare the proposed charts with the competitors.

One of the most interesting results has been obtained with the NPC Wilcoxon-Klotz chart for the location-scale problem. This chart takes full advantage of the flexibility of the NPC framework and outperforms the Cucconi and Lepage charts under a wide variety of interesting distributions. It is worth underlying that the NPC Wilcoxon-Klotz chart, as well as the Cucconi and Lepage charts, are not able to detect a compression of the scale of the process distribution. The main disadvantage we found about the NPC Wilcoxon-Klotz chart is that, among the charts we compared, it is the most sensitive to "extreme" reference samples, although the results are fairly close with the Cucconi chart.

As for the location monitoring problem, the NPC Wilcoxon-Waerden chart turned out to be a valid competitor of the popular Mann Whitney chart. The NPC Wilcoxon-Waerden has better OOC performance than the Mann Whitney chart with normal data. Unfortunately, it seems that this performance increase with the normal distribution cannot be achieved without a relatively higher performance degradation with more heavy-tailed distributions, like the Laplace. With the asymmetric Gamma distribution, it seems that the best-performing chart depends on the orientation of the location shift. Finally, the results of the IC conditional performance analysis show that, among all the charts we compared, the Wilcoxon chart is the least affected by "extreme" reference samples.

5.1 Future works

It would be interesting to study further the NPC scheme in the context of multivariate monitoring problems. From our experience, having fast calibration and

evaluation algorithms is essential for this kind of exploration. Therefore, the high-performance R packet that we have developed could be extended to support also multivariate distributions and SPS statistical models.

Another interesting idea to explore is the development of the CUSUM and EWMA versions of NPC charts. CUSUM and EWMA charts have only been briefly discussed in this Thesis. A major advantage of these charts with respect to Shewhart-type charts is that, at each time step, they exploit the information contained both in present and past test samples. Combining the principles of CUSUM and EWMA charts with the flexibility of the NPC framework could lead to the discovery of powerful control charts.

Bibliography

[1] J. Benneyan, R. Lloyd, and P. Plsek. "Statistical process control as a tool for research and healthcare improvement". In: *BMJ Quality & Safety* 12.6 (2003), pp. 458–464.

[2] D. Blackman and S. Vigna. "Scrambled linear pseudorandom number generators". In: *ACM Transactions on Mathematical Software (TOMS)* 47.4 (2021), pp. 1–32.

[3] D. Caughey, A. Dafoe, and J. Seawright. "Nonparametric combination (NPC): A framework for testing elaborate theories". In: *The Journal of Politics* 79.2 (2017), pp. 688–701.

[4] S. Chakraborti and M. Graham. *Nonparametric statistical process control.* John Wiley & Sons, 2019.

[5] S. Chakraborti and M. A. Graham. "Nonparametric (distribution-free) control charts: An updated overview and some results". In: *Quality Engineering* 31.4 (2019), pp. 523–544.

[6] S. Chakraborti and M. A. Van de Wiel. "A nonparametric control chart based on the Mann-Whitney statistic". In: *Beyond parametrics in interdisciplinary research: Festschrift in Honor of Professor Pranab K. Sen.* Institute of Mathematical Statistics, 2008, pp. 156–172.

[7] S. Chakraborti, P. Van der Laan, and S. Bakir. "Nonparametric control charts: an overview and some results". In: *Journal of Quality Technology* 33.3 (2001), pp. 304–315.

[8] S. Chakraborti, P. Van der Laan, and M. Van de Wiel. "A class of distribution-free control charts". In: *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 53.3 (2004), pp. 443–462.

[9] S. Chowdhury, A. Mukherjee, and S. Chakraborti. "A new distribution-free control chart for joint monitoring of unknown location and scale parameters of continuous distributions". In: *Quality and Reliability Engineering International* 30.2 (2014), pp. 191–204.

[10] W. F. Christensen and B. N. Zabriskie. "When your permutation test is doomed to fail". In: *The American Statistician* (2021), pp. 1–11.

[11] L. Corain and L. Salmaso. "Nonparametric permutation and combination-based multivariate control charts with applications in microelectronics". In: *Applied Stochastic Models in Business and Industry* 29.4 (2013), pp. 334–349.

[12] L. Dagum and R. Menon. "OpenMP: an industry standard API for shared-memory programming". In: *IEEE computational science and engineering* 5.1 (1998), pp. 46–55.

[13] D. Eddelbuettel et al. "Rcpp: Seamless R and C++ integration". In: *Journal of statistical software* 40.8 (2011), pp. 1–18.

[14] R. A. Fisher. *Statistical methods for research workers.* Vol. 5th ed.–r. Biological monographs and manuals. Edinburgh: Oliver and Boyd, 1934.

[15] R. A. Fisher. *The design of experiments*. Edinburgh: Oliver and Boyd, 1935.

[16] J. L. Gastwirth. "Percentile modifications of two sample rank tests". In: *Journal of the American Statistical Association* 60.312 (1965), pp. 1127–1141.

[17] J. D. Gibbons and S. Chakraborti. "Nonparametric statistical inference fifth edition". In: *STATISTICS TEXTBOOKS AND MONOGRAPHS* (2010).

[18] J. Klotz. "Nonparametric tests for scale". In: *The Annals of Mathematical Statistics* 33.2 (1962), pp. 498–512.

[19] D. Lanhede. *Non-parametric Statistical Process Control: Evaluation and Implementation of Methods for Statistical Process Control at GE Healthcare, Umeå*. 2015.

[20] E. L. Lehmann and J. P. Romano. *Testing statistical hypotheses*. Springer Science & Business Media, 2006.

[21] Y. Lepage. "A combination of Wilcoxon's and Ansari-Bradley's statistics". In: *Biometrika* 58.1 (1971), pp. 213–217.

[22] M. Marozzi. "A bi-aspect nonparametric test for the two-sample location problem". In: *Computational statistics & data analysis* 44.4 (2004), pp. 639–648.

[23] M. Marozzi. "Multivariate tri-aspect non-parametric testing". In: *Journal of Nonparametric Statistics* 19.6-8 (2007), pp. 269–282.

[24] M. Marozzi. "Some notes on nonparametric inferences and permutation tests". In: *Metron International Journal of Statistics* 60.3-4 (2002), pp. 139–151.

[25] B. Mason and J. Antony. "Statistical process control: an essential ingredient for improving service and manufacuring quality". In: *Managing Service Quality: An International Journal* (2000).

[26] D. C. Montgomery. *Introduction to statistical quality control*. 6th. John Wiley & Sons.

[27] A. Mukherjee and S. Chakraborti. "A distribution-free control chart for the joint monitoring of location and scale". In: *Quality and Reliability Engineering International* 28.3 (2012), pp. 335–352.

[28] E. S. Page. "Continuous inspection schemes". In: *Biometrika* 41.1/2 (1954), pp. 100–115.

[29] F. Pesarin. *Multivariate permutation tests: with applications in biostatistics*. Vol. 240. Wiley Chichester, 2001.

[30] F. Pesarin and L. Salmaso. "A review and some new results on permutation testing for multivariate problems". In: *Statistics and Computing* 22.2 (2012), pp. 639–646.

[31] F. Pesarin and L. Salmaso. *Permutation tests for complex data: theory, applications and software*. John Wiley & Sons, 2010.

[32] W. A. Shewhart. "Quality control charts". In: *The Bell System Technical Journal* 5.4 (1926), pp. 593–603.

[33] P. H. Tran et al. "Real time data-driven approaches for credit card fraud detection". In: *Proceedings of the 2018 international conference on e-business and applications*. 2018, pp. 6–9.

[34] B. Van der Waerden. "Ein neuer Test für das Problem der zwei Stichproben". In: *Mathematische Annalen* 126.1 (1953), pp. 93–107.

[35] F. Wilcoxon. "Individual Comparisons by Ranking Methods". In: *Biometrics Bulletin* 1.6 (1945), pp. 80–83.

[36] W. H. Woodall. "Controversies and contradictions in statistical process control".
      In: *Journal of Quality Technology* 32.4 (2000), pp. 341–350.

Appendix A

npcharts R package documentation

<div style="border:1px solid">

npcharts package

January 30, 2022

---

npcharts.setseed   *Set the seed used by this R packet*

Description

Set the seed used by this R packet

Usage

```
npcharts.setseed(seed)
```

Arguments

seed   The seed. Either an integer scalar or an integer vector of length 2 representing a
            64-bit seed.

---

compute_permutation_distribution
                        Compute permutation distribution

Description

Used for precomputing permutation distribution of linear rank statistics

Usage

```
compute_permutation_distribution(statistic, m, n, n_perm)
```

Arguments

statistic   should be set to a valid statistic key. See details for consulting the available
            statistics and corresponding keys

m   The dimension used for the reference sample

n   The dimension used for the test samples

nperm   the number of random permutations

1

</div>

Details

The supported statistics and the respective keys to be used are:

- Wilcoxon rank sum - "wilcoxon_rank_sum"
- Mann-Whitney - "mann_whitney"
- Mean-normalized wilcoxon rank sum - "centered_wilcoxon_rank_sum"
- Percentile modified rank test with s=r=0.42 - "percentiles_42_42"
- Percentile modified rank test with s=r=0.46 - "percentiles_46_46"
- Klotz - "klotz_statistic"
- Ansari Bradely - "ab_statistic"
- Lepage - "lepage"
- Cucconi - "cucconi"

```
compute_monitoring_statistic
```
Phase II monitoring

Description

Compute the observed value of the monitoring statistic

Usage

```
compute_monitoring_statistic(
  x1,
  x2,
  monitoring_statistic_type,
  monitoring_statistic_parameters
)
```

Arguments

x1	The reference sample
x2	The test sample
monitoring_statistic_type	
	Either "npc" or "simple_statistic"
monitoring_statistic_parameters	
	A list with type-specific parameters of the monitoring statistic. For the "simple_statistic" type the only parameter is "statistic" which should be set to a valid statistic key. For the "npc" type the four (required) parameters are "statistics", "permutation_distributions", "tails" and "combining_function". The "statistics", "permutation_distributions" and "tails" parameters should be three lists of the same length. The "combining_function" parameters should be a string. See details for consulting the available statistics and combining functions and the keys to be used.

66

Details

The supported statistics and the respective keys to be used are:

- Wilcoxon rank sum - "wilcoxon_rank_sum"
- Mann-Whitney - "mann_whitney"
- Mean-normalized wilcoxon rank sum - "centered_wilcoxon_rank_sum"
- Percentile modified rank test with s=r=0.42 - "percentiles_42_42"
- Percentile modified rank test with s=r=0.46 - "percentiles_46_46"
- Klotz - "klotz_statistic"
- Ansari Bradely - "ab_statistic"
- Lepage - "lepage"
- Cucconi - "cucconi"

The three options for the values of the "tails" list parameter of npc monitoring statistics are: "two_sided", "left" and "right" The three options for the "combining_function" parameter of npc monitoring statistics are: "fisher", "liptak" and "tippet"

Value

the monitoring statistic value

---

evaluate.unconditional

Unconditional evaluation

Description

Unconditional evaluation for Stehwart-type charts

Usage

```
evaluate.unconditional(
  m,
  n,
  limits,
  location_shift,
  scale_multiplier,
  distribution_key,
  distribution_parameters,
  monitoring_statistic_type,
  monitoring_statistic_parameters,
  nsim,
  run_length_cap
)
```

Arguments

`m`	The dimension used for the reference sample
`n`	The dimension used for the test samples
`distribution_key`	
	A string which identifies a distribution or a distribution family. Used in combination with the "distribution_parameters" parameter in order to select the IC distribution. See details for consulting the available distributions and the respective keys.
`distribution_parameters`	
	A list with the distribution parameters (an empty list should be supplied if the distribution has not parameters)
`monitoring_statistic_type`	
	Either "npc" or "simple_statistic"
`monitoring_statistic_parameters`	
	A list with type-specific parameters of the monitoring statistic. For the "simple_statistic" type the only parameter is "statistic" which should be set to a valid statistic key. For the "npc" type the four (required) parameters are "statistics", "permutation_distributions", "tails" and "combining_function". The "statistics", "permutation_distributions" and "tails" parameters should be three lists of the same length. The "combining_function" parameters should be a string. See details for consulting the available statistics and combining functions and the keys to be used. the only parameter is "statistic" which should be set to a valid statistic key. For the "npc" type the three (required) parameters are "statistics", "permutation_distributions" and "tails"
`nsim`	The number of simulations
`run_length_cap`	A limit for the run length in the simulations used to guarantee convergence of the algorithm

Details

The supported distributions and the respective parameters are:

- Normal - key:"norm", params:"mean" and "sd"
- Laplace - key:"laplace", params:"location" and "scale"
- Student's T - key:"t", params:"df"
- Cauchy - key:"laplace", params:"location" and "scale"
- Chi squared - key:"chi_squared", params:"df"
- Exp(lambda=1) shifted by -1 - key:"normalized_rate_one_exponential"
- Gamma(scale=4,shape=2) shifted by -2 - key:"centered_gamma_with_shape_equal_four"

The supported statistics and the respective keys to be used are:

- Wilcoxon rank sum - "wilcoxon_rank_sum"
- Mann-Whitney - "mann_whitney"
- Mean-normalized wilcoxon rank sum - "centered_wilcoxon_rank_sum"
- Percentile modified rank test with s=r=0.42 - "percentiles_42_42"
- Percentile modified rank test with s=r=0.46 - "percentiles_46_46"
- Klotz - "klotz_statistic"

- Ansari Bradely - "ab_statistic"
- Lepage - "lepage"
- Cucconi - "cucconi"

The three options for the values of the "tails" list parameter of npc monitoring statistics are: "two_sided", "left" and "right". The three options for the "combining_function" parameter of npc monitoring statistics are: "fisher", "liptak" and "tippet"

Value

A data frame whose columns are "run_lengths", "reference_sample_means" and "reference_sample_sds"

---

calibrate.unconditional

Unconditional calibration

Description

Unconditional calibration for the Stehwart-type charts implemented by this package

Usage

```
calibrate.unconditional(
  m,
  n,
  distribution_key,
  distribution_parameters,
  monitoring_statistic_type,
  monitoring_statistic_parameters,
  limits_seq,
  is_upper_limit,
  nsim,
  run_length_cap
)
```

Arguments

m	The dimension used for the reference sample
n	The dimension used for the test samples
distribution_key	
	A string which identifies a distribution or a distribution family. Used in combination with the "distribution_parameters" parameter in order to select the IC distribution. See details for consulting the available distributions and the respective keys.
distribution_parameters	
	A list with the distribution parameters (an empty list should be supplied if the distribution has not parameters)
monitoring_statistic_type	
	Either "npc" or "simple_statistic"

`monitoring_statistic_parameters`

> A list with type-specific parameters of the monitoring statistic. For the "simple_statistic" type the only parameter is "statistic" which should be set to a valid statistic key. For the "npc" type the four (required) parameters are "statistics", "permutation_distributions", "tails" and "combining_function". The "statistics", "permutation_distributions" and "tails" parameters should be three lists of the same length. The "combining_function" parameters should be a string. See details for consulting the available statistics and combining functions and the keys to be used.

`limits_seq`   The numeric vector of limits for which the run length should be recorded at each simulation

`is_upper_limit`   A boolean parameter used to select whether the chart uses an upper limit or a lower limit

`nsim`   The number of simulations

`run_length_cap`   A limit for the run length in the simulations used to guarantee convergence of the algorithm

Details

The supported distributions and the respective parameters are:

- Normal - key:"norm", params:"mean" and "sd"
- Laplace - key:"laplace", params:"location" and "scale"
- Student's T - key:"t", params:"df"
- Cauchy - key:"laplace", params:"location" and "scale"
- Chi squared - key:"chi_squared", params:"df"
- Exp(lambda=1) shifted by -1 - key:"normalized_rate_one_exponential"
- Gamma(scale=4,shape=2) shifted by -2 - key:"centered_gamma_with_shape_equal_four"

The supported statistics and the respective keys to be used are:

- Wilcoxon rank sum - "wilcoxon_rank_sum"
- Mann-Whitney - "mann_whitney"
- Mean-normalized wilcoxon rank sum - "centered_wilcoxon_rank_sum"
- Percentile modified rank test with s=r=0.42 - "percentiles_42_42"
- Percentile modified rank test with s=r=0.46 - "percentiles_46_46"
- Klotz - "klotz_statistic"
- Ansari Bradely - "ab_statistic"
- Lepage - "lepage"
- Cucconi - "cucconi"

The three options for the values of the "tails" list parameter of npc monitoring statistics are: "two_sided", "left" and "right". The three options for the "combining_function" parameter of npc monitoring statistics are: "fisher", "liptak" and "tippet"

Value

A numeric matrix of size nsim x length(limits_seq)