GEORG-AUGUST-UNIVERSITÄT
GÖTTINGEN

# Adjusting for Covariates in Non-Parametric Simultaneous Inference

Master's Thesis - Institute for Mathematical Stochastic

Thomas Asendorf

First academic advisor: Prof. Dr. Edgar Brunner

Second academic advisor: Prof. Dr. Tatyana Krivobokova

2013

Statutory Declaration

I here with declare that I have completed the presented thesis independently, making use of the specified literature and aids only. The thesis in this form or in any other form has not been submitted to an examination body and has not been published.

..... .....

Date Signature

## Acknowledgements

A special thanks goes out to Prof. Brunner and Dr. Konietschke for their patients and guidance throughout the thesis. I also thank Prof. Krivobokova for kindly taking the role of a second academic advisor.

Thank you Marius for the hours of intense math chat during lunch time and beyond. I also thank brave Philipp for taking the courtesy of correcting my writing style. Last but not least, I thank my beloved Svenja for her unbound support, the best one can wish for.

# Contents

# 1 Introduction

In many fields of research, experiments are conducted to measure differences between subject groups of interest. Experimenters are interested in the behavior of groups of subjects which are influenced by a *factor* and, within a factor, by different *factor levels*. For example, in medical statistics dose finding studies aim at finding a medication dose, at which a certain medication is effective, but not harmful. In such a study, subjects are commonly split into different groups, a placebo group and several groups of different dosage. Then, the medication is referred to as the factor and the different dosage levels as factor levels. After having allocated the subjects into distinct groups, a relevant parameter of interest is measured, the so called *dependent variable*. This can be a blood parameter from a metric scale, like the C-reactive protein (CRP), or a discrete scale in which the subjects describe their own well-being, like the visual analogue scale (VAS). Frequently, more information, such as the age of the subjects, is taken to provide for typical descriptive statistics. These parameters, which are not relevant for the hypothesis of the experiment, but might have an influence on the dependent variable, are referred to as *concomitant variables* or *covariates*.

When trying to verify a difference in the dependent variable between experimental groups, the most common tests are the *t-test* Student (1908), in the case of only two experimental groups, and the *analysis of variance (ANOVA)*, see e.g. Timm (1975) p.359, for inference in more than two groups. These methods require the dependent variable to follow a normal distribution. Should this not be the case, for example when the dependent variable is collected on a discrete scale, non-parametric alternatives like *Wilcoxon-Mann-Whitney-Test* Mann and Whitney (1947), in the case of two experimental groups, and the *Kruskal-Wallis-Test* Kruskal and Wallis (1952), for more than two groups, can be applied to test for differences in the dependent variable between the experimental groups.

While these procedures are adequate when testing the global hypothesis, i.e. "there is no effect between the experimental groups", in the case of more than two groups they do not provide the experimental supervisor with an answer to which groups differ. This question is answered by conducting pairwise-comparisons, comparing all groups or only selected groups of interest with each other. To cope with $\alpha$-inflation, the p-values from the pairwise-comparisons are commonly adjusted using the *Bonferroni correction* or the less conservative *Bonferroni-Holm correction* Holm (1979).

In recent years, procedures have been developed which make the testing of the global hypothesis and afterward sequential testing obsolete. These new procedures are often referred to as *simultaneous test procedures*. For a semi-parametric model, like the case of a normal distributed dependent variable, Hothorn et al. (2008) developed procedures for simultaneous inference, replacing the ANOVA with subsequent p-value correction. In a non-parametric setting, Konietschke et al. (2012a) provide for simultaneous inference in factorial designs.

When additionally taking into account covariates, other procedures have to be applied for statistical inference. In place of the ANOVA, the so called *analysis of covariance (ANCOVA)*, an ANOVA model with an additional regression term (3.1) is used instead. The procedures developed by Hothorn et al. (2008) allow for simultaneous inference in the ANCOVA and ANOVA model alike. In a non-parametric setting, several works provide for statistical analysis, among them Quade (1967), Langer (1998) and Siemer (1999). These merely provide for testing of global hypothesis though, making a pairwise-comparisons and subsequent correction of the p-values necessary when investigating for differences between

the groups. So far, no simultaneous test procedures for the non-parametric setting exist, which are capable of considering covariates. The aim of this thesis will be to provide for a non-parametric simultaneous test procedure which considers covariates.

## 1.1 Motivating Covariates

When taking into account covariates in a parametric setting we assume a linear dependency between the covariate and the dependent variable (3.1). By taking this dependency into account in the statistical inference we hope to reduce the variance of the error term, giving more precise testing results and thus a higher power of the test. Figure 1 illustrates how covariates are supposed to reduce the error term and increase the power of the test statistic.

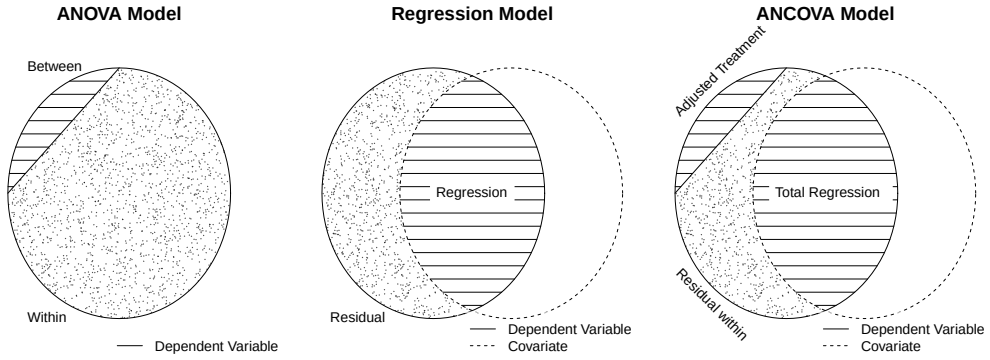


*Figure 1:* Variance partitioning for the ANOVA model (*left*), the regression model *(center)* and the ANCOVA model *(right)*, according to Huitema (1980) p.26.

Figure 1 shows how the ANOVA partitions the variance into between and within variance, where the between variance is explained through the factor levels while the within variance is not. On the other hand, the regression model tries to explain as much of the variance of the residual as possible through the regressor, or in this case, through the covariate, leaving the unexplained variance to the residual. The ANCOVA model combines these two models, and therefore leads to a reduced variation of the error term.

In a non-parametric setting we also aim at reducing the variation of the dependent variable by partially explaining the dependent variable through the covariates, but using the ranks of the observations. More precisely, we assume a linear dependency between the rank of the dependent variable and the rank of the covariate (2.8). Within this setting, we are able to reduce the variance of the dependent variable by adjusting through the covariates.

It is important to understand, that this is the sole purpose of using covariates. Many experimenters wish to correct the dependent variable through covariates, especially when the covariate is unequally distributed among the factor levels and a correction promises more favorable results. Huitema (1980) comments on this issue:

> "If a non-randomized design other than the biased assignment design [a design with sufficiently randomized covariates] is employed and the covariate is measured after treatments are administered, the ANOVA on the covariate will be essentially uninterpretable because treatment effects and pretreatment differences

among the populations will be confounded." (p.190)

In fact, the overwhelming majority of opinions condemns the logic of correcting for covariates as not consistent and an abusive use of covariates, see Miller and Chapman (2001) for a full list of comments out of a broad range of literature. Covariates are not intended to correct a bias caused by bad or unhappy pre-experiment randomization. From this argumentation it immediately follows, that covariates should be chosen to be independent from the factor levels and equally distributed among the factor levels, otherwise interpretation difficulties will arise when conducting statistical analyses.

The thesis is structured as follows. In Section 2 we present the basic notation used throughout the thesis, as well as the underlying non-parametric model with and without covariates. We continue by explaining how hypotheses are formed and how they differ between the models presented. Afterwards, estimators are presented which will be used to conduct statistical analyses. In Section 3 we then continue by introducing parametric and non-parametric procedures for statistical inference with covariates. Then, the simultaneous test procedure for considering covariates in a non-parametric model is derived. In Section 4, the introduced procedures are compared in an extensive simulation study. Section 5 shows how the newly derived procedure can be applied on an example, the Bogalusa Heart Study. Section 6 concludes the thesis, giving an outlook on possible improvements and discussing the results attained. Section 7 contains supplementary material like theoretical results used throughout the thesis, as well as supplementary simulation results and programming code.

# 2 Preliminaries

## 2.1 Basic Notation and Important Definitions

To begin with, we introduce the basic notation which is used throughout the thesis. While some of the notation mentioned is widely spread, it is still listed here to avoid any confusion. Following notation will be used:

- The indicator function will be denoted by $\mathbf{1}_{\{\text{logical request}\}}$. If the logical request is true, then the term $\mathbf{1}_{\{\text{logical request}\}}$ is equal to 1, else it is equal to 0.
- The term $\mathbf{1}_n = (1, 1, \ldots, 1)'$ denotes a vector of ones of length $n$. Although there is a symbolic similarity to the indicator function, the two terms differ in the choice of their index, and therefore in their meaning.
- The unit matrix of dimension $n \times n$ is denoted by $\mathbf{I}_n$.
- The centering matrix of dimension $n \times n$ will be denoted by $\mathbf{P}_n = \mathbf{I}_n - 1/n \cdot \mathbf{1}_n \mathbf{1}_n'$.
- The term $\mathbf{0}$ denotes either a vector $(0, 0, \ldots, 0)'$ or a matrix of 0's of appropriate size for the setting it is used in.
- The operator $\oplus$ denotes the direct sum, i.e. for two matrices $\mathbf{A} \in \mathbb{R}^{n \times m}$ and $\mathbf{B} \in \mathbb{R}^{s \times t}$:

$$\mathbf{A} \oplus \mathbf{B} = \begin{pmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{0} & \mathbf{B} \end{pmatrix} \in \mathbb{R}^{(n+s) \times (m+t)}.$$

- The operator $\otimes$ denotes Kronecker's product, i.e. for two matrices $\mathbf{A} \in \mathbb{R}^{n \times m}$ and $\mathbf{B} \in \mathbb{R}^{s \times t}$:

$$\mathbf{A} \otimes \mathbf{B} = \begin{pmatrix} a_{11}\mathbf{B} & a_{12}\mathbf{B} & \ldots & a_{1m}\mathbf{B} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1}\mathbf{B} & a_{n2}\mathbf{B} & \ldots & a_{nm}\mathbf{B} \end{pmatrix} \in \mathbb{R}^{(ns) \times (mt)}.$$

- For values $X_1, \ldots, X_n$, let $\overline{X}_\cdot = 1/n \sum_{i=1}^n X_i$ denote the mean of these values.
- The rank and trace of a matrix $\mathbf{A} \in \mathbb{R}^{n \times m}$ are denoted by $rank(\mathbf{A})$ and $tr(\mathbf{A})$ respectively.
- The diagonal matrix of a matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$ is denoted as:

$$diag(\mathbf{A}) = \begin{pmatrix} a_{11} & 0 & \ldots & 0 \\ 0 & a_{22} & \ddots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \ldots & a_{nn} \end{pmatrix}.$$

- The diagonal matrix of a set of values $\{a_1, \ldots, a_n\}$ ist denoted as:

$$dia(a_1, \ldots, a_n) = \begin{pmatrix} a_1 & 0 & \ldots & 0 \\ 0 & a_2 & \ddots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \ldots & a_n \end{pmatrix}.$$

**Definition 2.1.1** (Convergence types)**.** *Let $X$ denote a random variable with distribution function $F$ and $X_1, \ldots, X_n$ denote a sequence of random variables with corresponding distribution functions $F_1, \ldots, F_n$. If:*

1. *For all continuity points of $F$ it holds, that $\lim_{n\to\infty} F_n(x) = F(x)$, then $X_n$* ***converges in distribution*** *to $X$ and we write: $X_n \xrightarrow{D} X$.*
2. *For all $\varepsilon > 0$ it holds that $\lim_{n\to\infty} P(|X_n - X| \geq \varepsilon) = 0$, then $X_n$* ***converges in probability*** *to $X$ and we write: $X_n \xrightarrow{P} X$.*
3. *It holds, that $P(\lim_{n\to\infty} X_n = X) = 1$, then $X_n$* ***converges almost surely*** *to $X$ and we write: $X_n \xrightarrow{a.s.} X$.*
4. *It holds, that $\lim_{n\to\infty} \mathrm{E}(|X_n - X|^r) = 0$, then $X_n$* ***converges to $X$ in the $r$th mean*** *and we write: $X_n \xrightarrow{L^r} X$.*

Almost sure convergence and convergence in the $r$th mean imply convergence in probability, and convergence in probability implies convergence in distribution. The opposite is generally not true, see Van der Vaart (1998) p.10.

**Definition 2.1.2** (Asymptotic Equivalence)**.** *Let $(X_n)_{n\in\mathbb{N}}$ and $(Y_n)_{n\in\mathbb{N}}$ denote two sequences of random variables, such that $|X_n - Y_n| \xrightarrow{P} 0$ for $n \to \infty$. Then $X_n$ and $Y_n$ are* ***asymptotically equivalent*** *and we write:*

$$X_i \doteqdot Y_i.$$

This relation is particularly useful when one is interested in the asymptotic distribution of a sequence of random variables. Then, it is sufficient to know the asymptotic distribution of an asymptotically equivalent sequence of random variables.

**Definition 2.1.3** (Asymptotic Distribution)**.** *Let $(X_n)_{n\in\mathbb{N}}$ and $(U_n)_{n\in\mathbb{N}}$ denote two sequences of random variables. Further denote $F$ a distribution function and $U$ a random variable with distribution function $F$ (we write $U \sim F$), such that $U_n \xrightarrow{D} U \sim F$. If $X_n \doteqdot U_n$ we say $X_n$* ***asymptotically follows the distribution*** *$F$ and write:*

$$X_n \dot{\sim} F.$$

Now that basic notations and definitions have been introduced, we will continue by introducing the non-parametric model, both with and without covariates.

## 2.2 The Non-Parametric Model

The statistical inference procedures presented in this master's thesis are supposed to be derived with minimal technical assumptions. Most importantly, we wish to derive statistics in a non-parametric environment, meaning we do not assume that the data has any characteristic structure or follows a certain distribution. The only assumption we make on the data provided is that it is at least ordinal. This allows us to examine data not only from continuous scales, but also from discrete scales. Further, we will be examining a one-factorial design with factor $A$ and factor levels $1, \ldots, a$, also referred to as *groups* of the factor $A$. The results presented will technically allow for unequally distributed covariates between the groups, which is why we do not assume a completely randomized factorial

(CRF) design. However, we strongly advise the user to provide for a CRF design, as unequal distributions of the covariates among the factor levels will inevitably lead to interpretation difficulties, see Miller and Chapman (2001). Although we will only present the calculations for a one-factorial design, the presented results can certainly be used for more complex factorial designs.

### 2.2.1 The Non-Parametric Model without Covariates

Let $X_{ik}$ denote the $k$th observation in the $i$th group, $k = 1, \ldots, n_i$ and $i = 1, \ldots, a$. The total number of observations will be denoted by $N = \sum_{i=1}^{a} n_i$. Further, let:

$$F_i(x) = P(X_{ik} < x) + \frac{1}{2}P(X_{ik} = x) \tag{2.1}$$

denote the *normalized* version of the distribution function, first mentioned in the context of non-parametric models by Lévy (1925). It was later used by numerous personalities, among them Ruymgaart (1980), Akritas et al. (1997), Munzel (1999) and Gao et al. (2008), for deriving asymptotic results for rank statistics. In the following, we will assume that:

$$X_{ik} \overset{i.i.d.}{\sim} F_i, \; i = 1, \ldots, a, \; k = 1, \ldots, n_i.$$

Thus, $X_{ik}$ and $X_{st}$ are independent whenever $i \neq s$ or $k \neq t$. In the non-parametric model, *relative treatment effects* are defined as:

$$p_i = \int H dF_i, \; i = 1, \ldots, a, \tag{2.2}$$

where $H$ denotes the *mean distribution function*, i.e. $H(x) = \sum_{i=1}^{a} \omega_i F_i(x)$ with weights $\omega_i$, fulfilling $\omega_i \geq 0$ for all $i = 1, \ldots, a$ and $\sum_{i=1}^{a} \omega_i = 1$. We will be using the unweighted form, i.e. $\omega_i = 1/a$, discussed by Brunner and Puri (2001), Gao et al. (2008), opposed to the weighted form using $\omega_i = n_i/N$, introduced by Kruskal and Wallis (1952). Unweighted relative treatment effects have *"...the advantage of not being influenced by the allocation of sample sizes in the data."* Gao et al. (2008) p.2575. More specifically, the hypotheses stated on the relative treatment effects are not influenced by the allocation of sample sizes.

Relative treatment effects are of high importance in the non-parametric model, as they can be used to uncover differences between factor levels. They can be interpreted in such a way, that if $p_i < p_j$, then the observations from the $i$th group, i.e. observations from the normalized version of the distribution function $F_i$, tend to be smaller than the observations from the $j$th group. If the relative treatment effect $p_i < 1/2$, then observations from the normalized distribution function $F_i$ tend to be smaller than observations from the mean distribution function $H$, while $p_i > 1/2$ indicates that observations from the normalized distribution function $F_i$ tend to be larger than observations from the mean distribution function $H$.

Figure 2 shows different distribution functions and the resulting properties for the relative treatment effects.

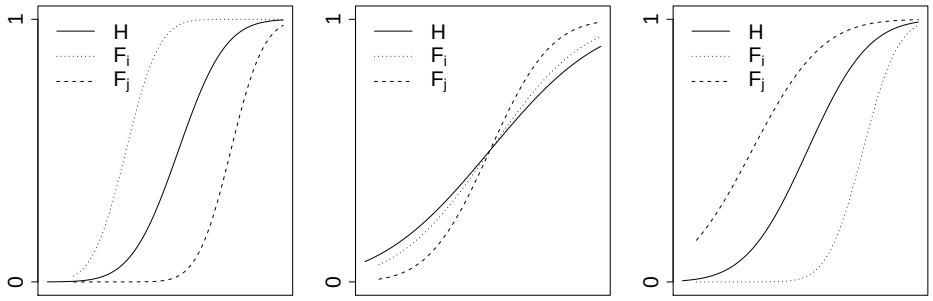


*Figure 2:* Possible properties of relative treatment effects, $p_i < p_j$ (*left*), $p_i = p_j$ (*center*) and $p_i > p_j$ (*right*).

By comparing the normalized distribution functions of the factor levels $F_i$ with the mean distribution function $H$, paradox statements like Efron's Dice cannot occur, see Thangavelu and Brunner (2007). Further we will denote pairwise relative treatment effects with:

$$p_{ij} = \int F_i dF_j, \;\; i = 1, \ldots, a, \;\; j = 1, \ldots, a, \tag{2.3}$$

which can be very helpful when representing relative treatment effects, their corresponding estimators, and which will be used to represent the covariance matrix of derived pivotal quantities in a more elegant way. We observe, that the relative treatment effects $p_i$ (2.2) are in fact a weighted sum of the *pairwise relative treatment effects* $p_{ij}$ (2.3) such that:

$$p_j = \sum_{i=1}^{a} \omega_i p_{ij}, \;\; j = 1, \ldots, a.$$

For more information on relative treatment effects in theory and in practice, we refer to Brunner and Munzel (2013). We will now continue by introducing covariates to the model.

### 2.2.2 The Non-Parametric Model with Covariates

Covariates were introduced into the non-parametric model by Quade (1967) and thoroughly discussed by Langer (1998), Domhof (2001), Siemer (1999), Bathke (1998) and Christophliemk (2001). The aim of using covariates in a non-parametric setting is to improve the estimation of the relative treatment effects of the dependent variable, by reducing the variation of the estimation through consideration of the covariates. As mentioned before, regarding covariates in a non-parametric setting cannot, just as it cannot in a parametric setting, correct the dependent variable for attaining "unbiased" results, see Miller and Chapman (2001) for details.

For denoting covariates we will introduce a further index on the observed data. Let $X_{ik}^{(r)}$ denote the $k$th replicate in the $i$th group of the $r$th covariate, $i = 1, \ldots, a$, $k = 1, \ldots, n_i$ and $r = 0, \ldots, d$, where $r = 0$ refers to the dependent variable and $r = 1, \ldots, d$ to the covariates. Table 1 gives a comprehensive overview of the underlying data.

*Table 1:* The underlying data setting when considering covariates.

| Factor level | Dep. variable | 1st Covariate | ... | $d$th Covariate |
|---|---|---|---|---|
| 1 | $X_{11}^{(0)}$ | $X_{11}^{(1)}$ | ... | $X_{11}^{(d)}$ |
| 1 | $X_{12}^{(0)}$ | $X_{12}^{(1)}$ | ... | $X_{12}^{(d)}$ |
| $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ |
| 1 | $X_{1n_1}^{(0)}$ | $X_{1n_1}^{(1)}$ | ... | $X_{1n_1}^{(d)}$ |
| 2 | $X_{21}^{(0)}$ | $X_{21}^{(1)}$ | ... | $X_{21}^{(d)}$ |
| $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ |
| a | $X_{an_a}^{(0)}$ | $X_{an_a}^{(1)}$ | ... | $X_{an_a}^{(d)}$ |

Analogously to the non-parametric model without covariates, we denote the normalized version of the distribution function by:

$$F_i^{(r)}(x) = P(X_{ik}^{(r)} < x) + \frac{1}{2}P(X_{ik}^{(r)} = x) \tag{2.4}$$

and assume that:

$$(X_{ik}^{(0)}, \ldots, X_{ik}^{(d)})' \overset{i.i.d.}{\sim} \mathbf{F}_i,\ i = 1, \ldots, a,\ k = 1, \ldots, n_i.$$

Now $X_{ik}^{(r)}$ and $X_{st}^{(u)}$ are still independent whenever $i \neq s$ or $k \neq t$, but may be dependent over the indizes $r$ and $u$. The covariates are therefore assumed to be random variables and not fixed constants. In this setting, the relative treatment effects are defined as:

$$p_i^{(r)} = \int H^{(r)} dF_i^{(r)},\ i = 1, \ldots, a,\ r = 0, \ldots d, \tag{2.5}$$

with the mean distribution function $H^{(r)}$, i.e. $H^{(r)}(x) = \sum_{i=1}^{a} \omega_i F_i^{(r)}(x)$ and weights $\omega_i$, fulfilling $\omega_i \geq 0$ for all $i = 1, \ldots, a$ and $\sum_{i=1}^{a} \omega_i = 1$. As in the non-parametric setting without covariates, when using the unweighted form of $H^{(r)}$ with $\omega_i = 1/a$, the relative treatment effects do not depend on the allocation of the sample size and will therefore be favored over the weighted form using $\omega_i = n_i/N$. The pairwise relative treatment effects are denoted by:

$$p_{ij}^{(r)} = \int F_i^{(r)} dF_j^{(r)},\ i, j = 1, \ldots, a,\ r = 1, \ldots, d, \tag{2.6}$$

which can, just as in the case of no covariates, be used for a more elegant representation of the relative treatment effects $p_i^{(r)}$. Further we denote:

$$Y_{ik}^{(r)} = H^{(r)}(X_{ik}^{(r)}),\ i = 1, \ldots, a,\ k = 1, \ldots, n_i,\ r = 0, \ldots, d, \tag{2.7}$$

the *asymptotic rank transformation* of $X_{ik}^{(r)}$.

This notation will help us introduce a model which considers covariates in a non-parametric setting. To adjust the dependent variable for covariates, we need some sort of connection between the dependent variable and the covariates. This connection we will be considering throughout this thesis is the same as can be found in Langer (1998) and Siemer (1999).

The proposed model takes into account covariates by assuming a connection between the covariates and the dependent variable in the asymptotic rank transformation, of the form:

$$Y_{ik}^{(0)} = \sum_{r=1}^{d} \gamma^{(r)} \cdot Y_{ik}^{(r)} + Y_{ik}^{reg}, \tag{2.8}$$

where $Y_{ik}^{reg}$ is an unobservable random variable and $\gamma^{(r)}$, for $r = 1, \ldots, d$, are unknown but fixed regression parameters. This type of connection between the dependent variable and the covariates implies a connection between the relative treatment effects of the dependent variable and the covariate, described through the *adjusted relative treatment effect*, which is given by:

$$p_i^* = p_i^{(0)} - \sum_{r=1}^{d} \gamma^{(r)} \cdot p_i^{(r)}.$$

It is important to state, that $p_i^*$ is not a relative treatment effect, but behaves similarly, in a sense that differences between factor levels will result in differences between the $p_i^*$. This can already be seen in the fact, that the adjusted relative treatment effect is a weighted sum of relative treatment effects which does not have to be within the interval $[0, 1]$. The value $p_i^*$ cannot be interpreted in another meaningful manner, therefore solely being useful for the testing of hypotheses.

## 2.3 Formulation of Hypotheses

### 2.3.1 Formulation of Hypotheses without Covariates

In a one-factorial design, the usual parametric approach lies in comparing the groupwise means, denoted by $\mu_i = \mathrm{E}(X_{ik})$, $i = 1, \ldots, a$, to discover differences between the factor levels. Hypotheses are therefore formulated over $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_a)'$. In a non-parametric model, hypotheses are often formulated over the marginal distributions of the factor levels, $\mathbf{F} = (F_1, \ldots, F_a)'$ (2.1) or over relative treatment effects $\mathbf{p} = (p_1, \ldots, p_a)'$ (2.2). To compare the groupwise means, marginal distributions or relative treatment effects, so called *contrast matrix* are used.

**Definition 2.3.1** (Contrast matrix)**.** *A matrix* $\mathbf{C} \in \mathbb{R}^{q \times a}$ *is called* ***contrast matrix****, when* $\mathbf{C} \neq \mathbf{0}$ *and the row sums of* $\mathbf{C}$ *are equal to* 0*, i.e.* $\mathbf{C}\mathbf{1}_a = \mathbf{0}$.

Using a contrast matrix $\mathbf{C}$, hypotheses in the parametric setting can be formulated by $H_0 : \mathbf{C}\boldsymbol{\mu} = \mathbf{0}$. In a non-parametric setting, contrast matrices are also used, but to formulate hypotheses for the marginal distributions of the data, i.e. $H_0 : \mathbf{CF} = \mathbf{0}$, or for relative treatment effects by $H_0 : \mathbf{Cp} = \mathbf{0}$. It can easily be seen, that if the marginal distributions are equal, i.e. if $H_0 : \mathbf{CF} = \mathbf{0}$ is true, then the other hypotheses over the groupwise means or the relative treatment effects also have to be true, see Brunner and Munzel (2013) p.190. The other direction is generally not true though, see Example 1. Furthermore, the implications $H_0 : \mathbf{Cp} = \mathbf{0} \Rightarrow H_0 : \mathbf{C}\boldsymbol{\mu} = \mathbf{0}$ and $H_0 : \mathbf{C}\boldsymbol{\mu} = \mathbf{0} \Rightarrow H_0 : \mathbf{Cp} = \mathbf{0}$ are generally not true, as can be seen in Examples 2 and 3.

- *Example 1*: Consider two groups, i.e. a=2. Assume that $F_1 = N(0, 1)$ and $F_2 = N(0, 2)$. Then the groups have unequal marginal distributions, but equal means and equal relative treatment effects.
- *Example 2*: Consider two groups and assume that $F_1 = N(1, 1)$ and $F_2 = Exp(1)$. Then $\mu_1 = \mu_2$, but $p_1 \neq p_2$.

– *Example 3*: Again, consider two groups and assume that $F_1 = N(\delta, 1)$ and $F_2 = Exp(1)$. The shift parameter $\delta$ can be chosen such that $p_1 = p_2$, but in that case, $\mu_1 \neq \mu_2$.

When testing hypotheses under consideration of covariates, we find a similar constellation.

### 2.3.2 Formulation of Hypotheses with Covariates

When considering covariates in a parametric setting, hypotheses are formulated through *adjusted population means*, see Section 3.1, denoted by $\mathbf{b}^*$. Hypotheses are therefore formulated by $H_0 : \mathbf{Cb}^* = \mathbf{0}$. In a non-parametric setting, two possibilities of formulating hypotheses have been discussed. For developing hypotheses, Langer (1998) pp.26, 40-41 assumes equal marginal distributions in the covariates, i.e. $F_1^{(r)} = \cdots = F_a^{(r)}$ for all $r = 1, \ldots, d$, and states hypotheses over the marginal distributions of the dependent variable. The developed tests therefore test the hypothesis $H_0 : \mathbf{CF}^{(0)} = \mathbf{0}$, where $\mathbf{F}^{(0)} = (F_1^{(0)}, \ldots, F_a^{(0)})'$. On the other hand, Siemer (1999) p.25 does not assume equal marginal distributions in the covariates and therefore develops tests for the hypothesis $H_0 : \mathbf{Cp}^* = \mathbf{0}$, where $\mathbf{p}^* = (p_1^*, \ldots, p_a^*)'$. In fact, both Langer (1998) and Siemer (1999) see in this hypothesis the only reasonable hypothesis when confronted with unequal marginal distributions in the covariates. The multiple contrast test procedures derived in this thesis will also test for the hypothesis $H_0 : \mathbf{Cp}^* = \mathbf{0}$, but because we would like to provide for confidence intervals for the adjusted relative treatment effects, which can be used for testing hypotheses. This is not possible when formulating the hypothesis over the marginal distributions of the dependent variable. The advantage of allowing for unequal marginal distributions in the covariates when formulating the hypotheses over the adjusted treatment effects will be technically allowed, but treated with caution, as interpretation difficulties will arise when such covariates are used.

The connection between the hypothesis for the parametric setting $H_0 : \mathbf{Cb}^* = \mathbf{0}$ and the hypothesis for the non-parametric setting $H_0 : \mathbf{Cp}^* = \mathbf{0}$ is not as clear as it was in the case without covariates. The most simple connection between the two hypotheses is given, when the covariates have no influence on the dependent variable. In this special case, the regression parameters of the parametric and the non-parametric model are all equal to 0, and the hypotheses are related to each other as if no covariates were involved. As soon as the covariates have an influence on the dependent variable, it becomes very difficult to compare these two hypothesis, because the regression parameters for the linear dependency between the covariates and the dependent variable calculated in the parametric setting and in the non-parametric setting are very different from each other. The relation between the regression parameters from the non-parametric model and the parametric model though, is crucial for deriving implications between the two corresponding hypotheses. The only statement which can surely be made is that if the covariates have equal marginal distributions, then sufficient differences between the factor levels will subsequently lead to a rejection of both hypotheses. For equal marginal distributions of the covariates and if $H_0 : \mathbf{CF} = \mathbf{0}$ is true, then the hypotheses $H_0 : \mathbf{Cb}^* = \mathbf{0}$ and $H_0 : \mathbf{Cp}^* = \mathbf{0}$ are also true, see Langer (1998) p.41. This again shows how important equal marginal distributions of the covariates are for attaining interpretable results.

We will now continue by showing how the adjusted treatment effects can be estimated and derive test statistics for testing the mentioned hypotheses. Estimating $p_i^*$ amounts to two steps. In the first step, the relative treatment effects of the dependent variable and the covariates (2.5) are estimated for $r = 0, \ldots, d$ and $i = 1, \ldots, a$. In a second step, the

unknown regression parameters $\gamma^{(r)}$ are estimated for $r = 1, \ldots, d$.

## 2.4 Estimation of the Relative Treatment Effects

Until now we have been working with non-observable random variables $Y_{ik}^{(r)}$ for $i = 1, \ldots, a$, $k = 1, \ldots, n_i$ and $r = 0, \ldots, d$, from the asymptotic rank transformation (2.7), and unknown regression parameters $\gamma^{(1)}, \ldots, \gamma^{(d)}$ originating from the underlying model (2.8). Now we will introduce estimators for the relative treatment effects when considering covariates, $p_i^{(r)}$, for $r = 0, \ldots, d$ and $i = 1, \ldots, a$, defined in (2.5). The main idea when deriving estimators for relative treatment effects lies in replacing the unknown normalized distribution functions $F_i^{(r)}(x)$, defined in (2.4), with their empirical counterpart:

$$\widehat{F}_i^{(r)}(x) = \frac{1}{n_i} \sum_{k=1}^{n_i} c(x - X_{ik}^{(r)}),\ i = 1, \ldots, a,\ r = 0, \ldots, d,$$

the *empirical cumulative distribution function*, where $c(y) = \mathbf{1}_{\{y>0\}} + \frac{1}{2} \cdot \mathbf{1}_{\{y=0\}}$. Using the empirical cumulative distribution function, we can approximate the non-observable random variables $Y_{ik}^{(r)}$ through the so called *rank transformation*:

$$\widehat{Y}_{ik}^{(r)} = \widehat{H}^{(r)}(X_{ik}^{(r)}) = \sum_{j=1}^{a} \omega_j \widehat{F}_j^{(r)}(X_{ik}^{(r)}),\ i = 1, \ldots, a,\ k = 1, \ldots, n_i, \tag{2.9}$$

with weights $\omega_j$ as described in Section 2.2.2. Apart from the representation through the empirical cumulative distribution function, there also exist rank representations for the values from the rank transformation. In non-parametric literature, results are often derived by using a rank notation for the observations. Therefore, we will briefly present how the rank transformation and ranks of observations are interconnected. To denote the values from the rank transformation using mean ranks, let us introduce a mean rank notation.

– The mean rank of observation $X_{ik}^{(r)}$ within the $r$th covariate is denoted by:

$$R_{ik}^{(r)} = \frac{1}{2} + \sum_{s=1}^{a} \sum_{t=1}^{n_s} c\left(X_{ik}^{(r)} - X_{st}^{(r)}\right).$$

– The mean rank of observation $X_{ik}^{(r)}$ within the $r$th covariate, but without the $j$th factor level is denoted by:

$$R_{ik}^{(r)}(-j) = \frac{1}{2} + \sum_{s \neq j}^{a} \sum_{t=1}^{n_s} c\left(X_{ik}^{(r)} - X_{st}^{(r)}\right).$$

– The mean rank of observation $X_{ik}^{(r)}$ within the $r$th covariate, but only within the $i$th factor level is denoted by:

$$R_{ik}^{(r)}(i) = \frac{1}{2} + \sum_{t=1}^{n_s} c\left(X_{ik}^{(r)} - X_{it}^{(r)}\right).$$

– The mean rank of observation $X_{ik}^{(r)}$ within the $r$th covariate, but only within the $j$th and $i$th factor level is denoted by:

$$R_{ik}^{(r)}(ij) = \frac{1}{2} + \sum_{t=1}^{n_j} c\left(X_{ik}^{(r)} - X_{jt}^{(r)}\right) + \sum_{t=1}^{n_i} c\left(X_{ik}^{(r)} - X_{it}^{(r)}\right).$$

Through this notation of the mean ranks of observations, we attain a mean rank notation for the rank transformation $\widehat{Y}_{ik}^{(r)}$, by:

$$\widehat{Y}_{ik}^{(r)} = \omega_i \frac{1}{n_i}\left(R_{ik}^{(r)}(i) - \frac{1}{2}\right) + \sum_{j\neq i}^{a} \omega_j \frac{1}{n_j}\left(R_{ik}^{(r)} - R_{ik}^{(r)}(-j)\right),\ i = 1,\dots,a,\ k = 1,\dots,n_i.$$

Using the cumulative distribution function, estimators for the relative treatment effects are given by:

$$\widehat{p}_j^{(r)} = \int \widehat{H}^{(r)} d\widehat{F}_j^{(r)} = \sum_{i=1}^{a} \omega_i \cdot \widehat{p}_{ij}^{(r)},\ j = 1,\dots,a,\ r = 0,\dots,d, \tag{2.10}$$

where $\widehat{p}_{ij}^{(r)} = \int \widehat{F}_i^{(r)} d\widehat{F}_j^{(r)} = 1/n_j \sum_{k=1}^{n_j} \widehat{F}_i^{(r)}(X_{jk}^{(r)})$ is an estimator for the pairwise relative treatment effects defined in (2.6). It is also possible to notate the estimators for the relative treatment effects using the mean ranks notation. It holds that:

$$\widehat{p}_j^{(r)} = \sum_{i=1}^{a} \omega_i^{(r)} \cdot \frac{1}{n_i}\left(\overline{R}_{j\cdot}^{(r)}(ij) - \frac{n_j+1}{2}\right) \text{ and } \widehat{p}_{ij}^{(r)} = \frac{1}{n_i}\left(\overline{R}_{j\cdot}^{(r)}(ij) - \frac{n_j+1}{2}\right).$$

The estimator $\widehat{p}_j^{(r)}$ has been shown to be unbiased and consistent for estimating $p_j^{(r)}$.

**Theorem 2.4.1.** *Let $\widehat{p}_j^{(r)}$ and $p_j^{(r)}$ be as defined in (2.10) and (2.5), respectively. If $N \to \infty$ such that $N/n_i \leq N_0 < \infty$ for a $N_0 \in \mathbb{N}$, then:*

$$\mathrm{E}\left(\widehat{p}_j^{(r)}\right) = p_j^{(r)} \quad \textit{and} \quad \mathrm{E}\left((\widehat{p}_j^{(r)} - p_j^{(r)})^2\right) \to 0 \textit{ for } j = 1,\dots,a.$$

*Proof.* The theorem can be proven analogously to Proposition 4.7 from Brunner and Munzel (2013) p.180. We will prove the unbiasedness of the estimator, to demonstrate how the additional index for the covariates is added.

$$\begin{aligned}
\mathrm{E}\left(\widehat{p}_j^{(r)}\right) &= \mathrm{E}\left(\int \widehat{H}^{(r)} d\widehat{F}_j^{(r)}\right) = \mathrm{E}\left(\sum_{i=1}^{a} \int \omega_i \widehat{F}_i^{(r)} d\widehat{F}_j^{(r)}\right) \\
&= \sum_{i=1}^{a} \mathrm{E}\left(\frac{1}{n_j}\sum_{k=1}^{n_j} \omega_i \widehat{F}_i^{(r)}(X_{jk}^{(r)})\right) = \sum_{i=1}^{a} \frac{1}{n_j}\sum_{k=1}^{n_j} \omega_i \mathrm{E}\left(\widehat{F}_i^{(r)}(X_{jk}^{(r)})\right) \\
&= \sum_{i=1}^{a} \frac{1}{n_j}\sum_{k=1}^{n_j} \omega_i \mathrm{E}\left(\frac{1}{n_i}\sum_{l=1}^{n_i} c(X_{jk}^{(r)} - X_{il}^{(r)})\right) \\
&= \sum_{i=1}^{a} \frac{1}{n_j}\sum_{k=1}^{n_j} \omega_i \frac{1}{n_i}\sum_{l=1}^{n_i} \mathrm{E}\left(c(X_{jk}^{(r)} - X_{il}^{(r)})\right) \\
&= \sum_{i=1}^{a} \frac{1}{n_j}\sum_{k=1}^{n_j} \omega_i \int F_i^{(r)} dF_j^{(r)} = \sum_{i=1}^{a} \frac{1}{n_j}\sum_{k=1}^{n_j} \omega_i p_{ij}^{(r)} = \sum_{i=1}^{a} \omega_i p_{ij}^{(r)} = p_j^{(r)}
\end{aligned}$$

□

For further details on the estimation of relative treatment effects, their representation with mean ranks and their properties, as well as the subject of pseudo ranks, we refer to Brunner and Munzel (2013) and Gao et al. (2008). Now that we have appropriate estimators for the relative treatment effects, it remains to find estimators for the regression parameters $\gamma^{(1)},\dots,\gamma^{(r)}$.

## 2.5 Estimation of the Regression Parameters

The basis for our estimation of the regression parameters is equation (2.8). Using this equation, a natural way of estimating the constants $\gamma^{(r)}$, for $r = 1, \ldots, d$, would be a linear regression. The only problem with directly performing a linear regression is that, contrarily to a usual linear regression, no intercept is specifically stated, but instead the intercept is hidden within the terms of the asymptotic rank transformation. Therefore we will perform a linear regression on the mean adjusted form of (2.8). Langer (1998) also uses a mean adjusted form for the estimation of the regression parameters $\gamma^{(r)}$, but assumes equal marginal distributions of the covariates, i.e. $F_i^{(r)} = F^{(r)}$, $i = 1, \ldots, a$, for all $r = 1, \ldots, d$. The validity of this assumption and the consequences for the null hypothesis have been thoroughly discussed by Langer (1998) pp.40-41. A different approach using an adjusted means regression was given by Siemer (1999), who does not assume equal marginal distributions within the covariates. This setting was also shortly discussed by Bathke and Brunner (2003), who give a different possibility from the estimation proposed by Langer (1998) for adjusting the regression model, to cope with unequal marginal distributions within the covariates. We will not assume equal marginal distributions within the covariates and therefore use a mean adjusted regression model as can be found in Bathke and Brunner (2003).

Taking (2.8), we can transfer the model into the mean adjusted form:

$$\begin{aligned} Y_{ik}^{(0)} &= \sum_{r=1}^{d} \gamma^{(r)} \cdot Y_{ik}^{(r)} + Y_{ik}^{reg} \\ \Rightarrow \quad Y_{ik}^{(0)} - \overline{Y}_{i\cdot}^{(0)} &= \sum_{r=1}^{d} \gamma^{(r)} \cdot (Y_{ik}^{(r)} - \overline{Y}_{i\cdot}^{(r)}) + (Y_{ik}^{reg} - \overline{Y}_{i\cdot}^{reg}). \end{aligned}$$

Because $Y_{ik}^{(r)}$ are non-observable random variables, we will approximate the values from the asymptotic rank transformation with corresponding values from the rank transformation $\widehat{Y}_{ik}^{(r)}$ defined in (2.9). Having done this, the least squares estimator for $\boldsymbol{\gamma} = (\gamma^{(1)}, \ldots, \gamma^{(d)})'$ is given by:

$$\widehat{\boldsymbol{\gamma}} = (\widehat{\mathbf{X}}'\widehat{\mathbf{X}})^{-1}\widehat{\mathbf{X}}'\widehat{\mathbf{y}}, \tag{2.11}$$

where $\widehat{\mathbf{X}}$ and $\widehat{\mathbf{y}}$ denote:

$$\widehat{\mathbf{X}} = \begin{pmatrix} \widehat{Y}_{11}^{(1)} - \widehat{\overline{Y}}_{1\cdot}^{(1)} & \ldots & \widehat{Y}_{11}^{(d)} - \widehat{\overline{Y}}_{1\cdot}^{(d)} \\ \widehat{Y}_{12}^{(1)} - \widehat{\overline{Y}}_{1\cdot}^{(1)} & \ldots & \widehat{Y}_{12}^{(d)} - \widehat{\overline{Y}}_{1\cdot}^{(d)} \\ \vdots & \ddots & \vdots \\ \widehat{Y}_{an_a}^{(1)} - \widehat{\overline{Y}}_{a\cdot}^{(1)} & \ldots & \widehat{Y}_{an_a}^{(d)} - \widehat{\overline{Y}}_{a\cdot}^{(d)} \end{pmatrix} \quad \text{and} \quad \widehat{\mathbf{y}} = \begin{pmatrix} \widehat{Y}_{11}^{(0)} - \widehat{\overline{Y}}_{1\cdot}^{(0)} \\ \widehat{Y}_{12}^{(0)} - \widehat{\overline{Y}}_{1\cdot}^{(0)} \\ \vdots \\ \widehat{Y}_{an_a}^{(0)} - \widehat{\overline{Y}}_{a\cdot}^{(0)} \end{pmatrix}. \tag{2.12}$$

In the following, we will further denote the non-empirical counterparts of $\widehat{\mathbf{X}}$ and $\widehat{\mathbf{y}}$ as $\mathbf{X}$ and $\mathbf{y}$, respectively. When introducing the non-parametric model with covariates, we noticed that the covariates are assumed to be random variables and not fixed values. Therefore, the regressors within $\widehat{\mathbf{X}}$ are random variables and we need a conditional assumption on the residuals, i.e. $Y_{ik}^{reg} - \overline{Y}_{i\cdot}^{reg}$, to justify this approach.

For a consistent estimation of $\boldsymbol{\gamma}$ we make the following technical assumptions:

(**A1**) There exists a number $N_0 \in \mathbb{N}$ such that $\frac{N}{n_i} \leq N_0 < \infty$ for all $i = 1, \ldots, a$.

(**A2**) $\mathrm{E}\left(Y_{ik}^{reg} - \overline{Y}_{i\cdot}^{reg} \mid \mathbf{X} = \mathbf{x}\right) = 0$ for all $i = 1, \cdots, a$, $k = 1, \cdots, n_i$.

(**A3**) Let $\lambda_{min}$ be the smallest eigenvalue of the matrix $\frac{1}{N}\mathbf{X}'\mathbf{X}$. Then there exists a constant $\kappa_0$ such that $\lambda_{min} \geq \kappa_0 > 0$ almost surely for all $N \in \mathbb{N}$.

Using these technical assumptions, we are able to prove the consistency of our estimator $\widehat{\boldsymbol{\gamma}}$ for the regression parameters.

**Theorem 2.5.1.** *Under the assumptions (**A1**)-(**A3**) and as $N \to \infty$ it holds that:*

$$\widehat{\boldsymbol{\gamma}} \xrightarrow{P} \boldsymbol{\gamma}$$

*Proof.* For a detailed proof see Langer (1998) Lemma 4.6-4.7 and Theorem 4.8, as well as Siemer (1999) p.59. □

Note that every time $\boldsymbol{\gamma}$ has to be estimated for one of the upcoming test statistics, the assumptions (**A1**)-(**A3**) have to apply. By plugging in the estimator $\widehat{\boldsymbol{\gamma}} = (\widehat{\gamma}^{(1)}, \ldots, \widehat{\gamma}^{(d)})'$ for $\boldsymbol{\gamma}$, we attain estimators for the adjusted relative treatment effects $p_i^*$ given by:

$$\widehat{p}_i^* = \widehat{p}_i^{(0)} - \sum_{r=1}^{d} \widehat{\gamma}^{(r)} \widehat{p}_i^{(r)}, \ i = 1, \ldots, a.$$

These estimators will be the basis for the upcoming simultaneous inference.

## 2.6 Properties of the Adjusted Relative Treatment Effects

From the previous section we have attained an estimator for the adjusted relative treatment effects, given by:

$$\widehat{p}_i^* = \widehat{p}_i^{(0)} - \sum_{r=1}^{d} \widehat{\gamma}^{(r)} \widehat{p}_i^{(r)}, \ i = 1, \ldots, a, \tag{2.13}$$

where definition and properties of the estimators for the regression parameters $\widehat{\gamma}^{(1)}, \ldots, \widehat{\gamma}^{(d)}$ are given in Section 2.5, and definition and properties of the estimators for the relative treatment effects $\widehat{p}_i^{(r)}$, $i = 1, \ldots, a$ and $r = 0, \ldots, d$, are given in Section 2.4.

This estimator, for the adjusted relative treatment effect, can be shown to be asymptotically unbiased and consistent.

**Lemma 2.6.1.** *Under the assumption (**A1**), i.e. $N/n_i \leq N_0 < \infty$ for some $N_0 \in \mathbb{N}$, and for $N \to \infty$ it holds that:*

$$\mathrm{E}(\widehat{p}_i^*) \to \mathrm{E}(p_i^*) \quad \textit{and} \quad \widehat{p}_i^* \xrightarrow{P} p_i^* \quad \textit{for} \quad i = 1, \ldots, a \tag{2.14}$$

*Proof.* From (2.5.1) and (2.4.1) we know that:

$$\widehat{p}_i^{(r)} \xrightarrow{P} p_i^{(r)} \quad \text{and} \quad \widehat{\gamma}^{(r)} \xrightarrow{P} \gamma^{(r)}.$$

By Slutsky's theorem it follows that $\widehat{\gamma}^{(r)}\widehat{p}_i^{(r)} \xrightarrow{D} \gamma^{(r)}p_i^{(r)}$. However, because $p_i^{(r)}$ and $\gamma^{(r)}$ are constants it follows that $\widehat{\gamma}^{(r)}\widehat{p}_i^{(r)} \xrightarrow{P} \gamma^{(r)}p_i^{(r)}$. Then:

$$\widehat{p}_i^* = \widehat{p}_i^{(0)} - \sum_{r=1}^{d} \widehat{\gamma}^{(r)}\widehat{p}_i^{(r)} \xrightarrow{P} p_i^{(0)} - \sum_{r=1}^{d} \gamma^{(r)}p_i^{(r)} = p_i^*$$

With Portmanteau (7.1.2) it immediately follows that $\widehat{p}_i^*$ is asymptotically unbiased. □

The most important remark which has to be made when interpreting $\widehat{p}_i^*$ is, that it is not an estimator for the relative treatment effect of the dependent variable. Instead, it is a consistent and asymptotically unbiased estimator for the relative treatment effect of the dependent variable, shifted by a part influenced through the covariates. The estimator for the adjusted relative treatment effects $\widehat{p}_i^*$ can be interpreted in the same manner as the estimator for the relative treatment effects of the dependent variable when not regarding covariates, $\widehat{p}_i$, in a sense, that the relative asymptotic comparison of the factor levels is given by both estimators. This means that differences between $\widehat{p}_i^*$ and $\widehat{p}_j^*$ can be interpreted as differences between the factor levels $i$ and $j$ of the dependent variable, but it is not possible to interpret the estimator $\widehat{p}_i^*$ as a relative treatment effect. For more details on the interpretation on the estimators for the shifted relative treatment effects we refer to Langer (1998) p.25.

Since an estimator which can be interpreted as a relative treatment effect would be highly desirable, we will give thoughts on a possible estimator further in this thesis, in Section 3.6.

# 3 Simultaneous Inference

In the literature, several procedures for testing with covariates in a non-parametric environment have been developed. Taking an arbitrary contrast matrix $\mathbf{C} \in \mathbb{R}^{s \times a}$ Langer (1998) derives two test statistics for the hypothesis $H_0 : \mathbf{CF}^{(\mathbf{0})} = \mathbf{0}$, a Wald-Type statistic and an ANOVA-Type statistic. Siemer (1999) then uses this as a basis for developing Wald-Type and ANOVA-Type test statistics for the hypothesis $H_0 : \mathbf{Cp}^* = \mathbf{0}$, which becomes necessary, because unlike Langer (1998), Siemer (1999) does not assume equal distributions of the covariates between the groups. Christophliemk (2001) decided to take the test statistics proposed by Siemer (1999) and improve these using transformation methods on the relative treatment effects.

In addition to some of the existing test statistics, we would like to develop a further test statistic based on the work of Konietschke et al. (2012a). This test, for the hypothesis $H_0 : \mathbf{Cp}^* = \mathbf{0}$, will be a multiple contrast test procedure, giving us the ability not only to say if there is a group effect, but additionally also where this group effect can be observed, without having to perform post-hoc analysis.

Before we do so, we briefly introduce the parametric ANCOVA, the parametric multiple contrast test procedures and the methods used by Siemer (1999), as these will be compared to the newly derived multiple contrast test procedure in an extensive simulation study.

## 3.1 Parametric ANCOVA

The ANCOVA is an extension of the well-known ANOVA, which tests for differences between blocks of observations, by a regression parameter. It is well described in Timm (1975), Huitema (1980), Kirk (1982), Seber (1977) and many other works. The described model in Timm (1975) pp.471-474 is given by:

$$\mathbf{X}^{(0)} = \underbrace{\mathbf{M}_1\mathbf{b}}_{\text{ANOVA component}} + \underbrace{\mathbf{M}_2\boldsymbol{\beta}}_{\text{Linear regression component}} + \underbrace{\boldsymbol{\varepsilon}}_{\text{Error term}}, \tag{3.1}$$

where the components of this model are given by:

- The dependent variable $\mathbf{X}^{(0)} = (X_{11}^{(0)}, X_{12}^{(0)}, \cdots, X_{an_a}^{(0)})'$
- The design matrix $\mathbf{M}_1 = \bigoplus_{i=1}^{a} \mathbf{1}_{n_i}$
- The parameter vector of factor level means $\mathbf{b} = (\mu_1, \ldots, \mu_a)'$
- The regressor matrix $\mathbf{M}_2 = \begin{pmatrix} X_{11}^{(1)} & \cdots & X_{11}^{(d)} \\ X_{12}^{(1)} & \cdots & X_{12}^{(d)} \\ \vdots & \ddots & \vdots \\ X_{an_a}^{(1)} & \cdots & X_{an_a}^{(d)} \end{pmatrix}$
- The vector of regression parameters $\boldsymbol{\beta} = (\beta_1, \cdots, \beta_d)'$
- The error term $\boldsymbol{\varepsilon} = (\varepsilon_{11}, \varepsilon_{12}, \cdots, \varepsilon_{an_a})' \sim N(\mathbf{0}, \sigma^2 \cdot \mathbf{I}_N)$.

Contrarily to our model assumptions, Timm (1975) states that the covariates need to be fixed and continuous values. While it is possible to cope with random covariates by stating conditional assumptions on the residuals, i.e. $\mathrm{E}(\boldsymbol{\varepsilon}|\mathbf{M}_2 = \mathbf{m}_2) = 0$ and $\mathrm{Cov}(\boldsymbol{\varepsilon}|\mathbf{M}_2 = \mathbf{m}_2) = \sigma^2 \cdot \mathbf{I}_N$, we will not do so, to avoid unforeseen consequences for the interpretation of the adjusted population means. Therefore the multiple contrast test procedures will be compared to the model as explained in Timm (1975), which assumes continuous fixed observed variables as covariates. The aim of the ANCOVA model described is to test for differences between the levels of the *adjusted population means*, given by $\mathbf{b}^* = \mathbf{b} - (\mathbf{M}_1'\mathbf{M}_1)^{-1}\mathbf{M}_1'\mathbf{M}_2\boldsymbol{\beta}$. As long as the upper assumptions and the conditions that $\mathbf{M}_1'\mathbf{M}_1$ and $\mathbf{M}_2'\mathbf{M}_2$ are regular are fulfilled, we attain estimators for $\boldsymbol{\beta}$, $\mathbf{b}$, the adjusted population means $\mathbf{b}^*$ and $\sigma^2$ through (3.2), (3.3), (3.4) and (3.5), respectively.

$$\widehat{\boldsymbol{\beta}} = \left(\mathbf{M}_2'\left(\mathbf{I}_N - \mathbf{J}_M\right)\mathbf{M}_2\right)^{-1}\mathbf{M}_2'\left(\mathbf{I}_N - \mathbf{J}_M\right)\mathbf{X}^{(0)} \text{ with } \mathbf{J}_M = \mathbf{M}_1(\mathbf{M}_1'\mathbf{M}_1)^{-1}\mathbf{M}_1' \tag{3.2}$$

$$\widehat{\mathbf{b}} = (\mathbf{M}_1'\mathbf{M}_1)^{-1}\mathbf{M}_1'\mathbf{X}^{(0)} \tag{3.3}$$

$$\widehat{\mathbf{b}}^* = \widehat{\mathbf{b}} - (\mathbf{M}_1'\mathbf{M}_1)^{-1}\mathbf{M}_1'\mathbf{M}_2\widehat{\boldsymbol{\beta}} = (\mathbf{M}_1'\mathbf{M}_1)^{-1}\mathbf{M}_1'(\mathbf{X}^{(0)} - \mathbf{M}_2\widehat{\boldsymbol{\beta}}) \tag{3.4}$$

$$\widehat{\sigma}^2 = \left(\mathbf{X}^{(0)\prime}\left(\mathbf{I}_N - \mathbf{J}_M\right)\mathbf{X}^{(0)} - \widehat{\boldsymbol{\beta}}'\mathbf{M}_2'(\mathbf{I}_N - \mathbf{J}_M)\mathbf{M}_2\widehat{\boldsymbol{\beta}}\right)/f \tag{3.5}$$

$$\text{where } f = N - rank(\mathbf{M}_1) - rank(\mathbf{M}_2)$$

Using these estimators, we can attain test statistics for testing the hypothesis $H_0 : \mathbf{C}\mathbf{b}^* = \mathbf{0}$, where $\mathbf{C}$ denotes an arbitrary contrast matrix. Timm (1975) p.474 further shows that with:

$$\begin{aligned}\mathbf{S} &= \mathrm{Var}\left[\widehat{\mathbf{b}}^*\right] \\ &= \sigma^2\left[(\mathbf{M}_1'\mathbf{M}_1)^{-1} + (\mathbf{M}_1'\mathbf{M}_1)^{-1}\mathbf{M}_1'\mathbf{M}_2\left(\mathbf{M}_2'(\mathbf{I}_N - \mathbf{J}_M)\mathbf{M}_2\right)^{-1}\mathbf{M}_2'\mathbf{M}_1\left(\mathbf{M}_1'\mathbf{M}_1\right)^{-1}\right] \\ Q_T &= \widehat{\mathbf{b}}^{*\prime}\mathbf{C}'(\mathbf{C}\mathbf{S}\mathbf{C}')^{-}\mathbf{C}\widehat{\mathbf{b}}^*\end{aligned}$$

Following equation holds:

$$T_T = \frac{Q_T/g}{\widehat{\sigma}^2} \overset{H_0}{\sim} F(g, f), \tag{3.6}$$

where $g = rank(\mathbf{C})$ and $f = N - rank(\mathbf{M}_1) - rank(\mathbf{M}_2)$. The test statistics $T_T$, as one of the most widely spread test statistics for statistical inference with covariates, will be compared to the multiple contrast test procedure in terms of $\alpha$-level and power.

## 3.2 Parametric Multiple Contrast Test Procedures

Additionally to the parametric ANCOVA, multiple contrast test procedures exist for the parametric setting given in (3.1). These procedures allow not only for testing the global hypothesis $H_0 : \mathbf{C}\mathbf{b}^* = \mathbf{0}$, but also for investigating between which groups differences occur, without having to perform post-hoc tests. A simultaneous test procedure for general linear models, and thus for the parametric ANCOVA, was developed by Hothorn et al. (2008). The procedure assumes an underlying semi-parametric model:

$$\mathcal{M} = ((X_1, \ldots, X_n), \boldsymbol{\theta}, \tau),$$

with a fixed but unknown parameter $\boldsymbol{\theta} \in \mathbb{R}^p$ and other random (or nuisance) parameters $\tau$. Further it is required, that if $\widehat{\boldsymbol{\theta}}_n$ is an estimator for $\boldsymbol{\theta}$ and $S_n$ an estimator for $\mathrm{Cov}(\widehat{\boldsymbol{\theta}}_n)$, then following equation has to hold:

$$a_n^{1/2}(\widehat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}) \xrightarrow{D} N(\mathbf{0}, \boldsymbol{\Phi}),$$

for some multivariate central limit theorem and some positive non-decreasing sequence $(a_n)_{n\in\mathbb{N}}$, such that $a_n\mathbf{S}_n \xrightarrow{P} \mathbf{\Phi}$. Under these assumptions, Hothorn et al. (2008) develop simultaneous test procedures for the hypothesis $H_0 : \mathbf{C}\boldsymbol{\theta} = \mathbf{0}$ through a test statistic:

$$\mathbf{T}_H = \mathbf{D}_n^{-1/2}(\widehat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}) \xrightarrow{D} N(\mathbf{0}, \mathbf{R}), \tag{3.7}$$

where $\mathbf{D_n} = diag(\mathbf{C}\mathbf{S}_n\mathbf{C}')$. Then the components of the test statistic $\mathbf{T}_H$ are asymptotically $N(0,1)$ distributed, and a testing procedure is conceived by estimating the unknown correlation matrix $\mathbf{R}$ and using appropriate critical values from the multivariate normal distribution, a similar approach as we will follow when deriving the multiple contrast test procedures in this thesis, derived in Section 3.5. For more information on the technical details of the procedure, we refer to Hothorn et al. (2008).

The implementation developed by Hothorn et al. (2008) is given in statistics program R, with the *multcomp* package. This package contains a function *glht* which can be used as a wrapper for many functions treating general linear models in R, among them the function *aov*, which allows for global inference in the parametric ANCOVA model. Since our upcoming simulation study is conducted in R, we will use the function *aov* for statistical inference, and the wrapper *glht* to allow for simultaneous inference. The procedure can be applied by calling following functions in R:

```
1  ancova<-aov(X0 ~ group+X1,data=datalm)
2  result<-glht(ancova,linfct=mcp(group=''Tukey''))
```

where *datalm* is an appropriate data frame containing the data, *X0* refers to the dependent variable, *group* refers to a group variable employed as a factor and *X1* refers to a covariate. The result from the simultaneous inference will then be compared to the multiple contrast test procedure for the non-parametric setting, derived in Section 3.5. For more details on the implementation of the simultaneous test procedures, as well as examples, we refer to Hothorn et al. (2008).

## 3.3 Wald-Type Statistic

One of the methods we will be comparing the multiple contrast test procedures to is the non-parametric approach using a Wald-type statistic proposed, by Siemer (1999). The reason we do not compare the Wald-type statistic of Langer (1998) with the multiple contrast test procedure, is that Langer (1998) tests a different hypothesis, namely $H_0 : \mathbf{C}\mathbf{F}^{(0)} = \mathbf{0}$ opposed to Siemer (1999), who tests for the hypothesis $H_0 : \mathbf{C}\mathbf{p}^* = \mathbf{0}$, the same hypothesis which is tested by the multiple contrast test procedure. For this denote:

$$\mathbf{\Lambda}_N = \text{Cov}\left(\sqrt{N}\left(\int H^{(r)} d\widehat{\mathbf{F}}^{(r)} - \int \mathbf{F}^{(r)} d\widehat{H}^{(r)}\right)_{r=0,\ldots,d}\right), \tag{3.8}$$

where $\widehat{\mathbf{F}}^{(r)} = (\widehat{F}_1^{(r)}, \ldots, \widehat{F}_a^{(r)})'$ and $\mathbf{F}^{(r)} = (F_1^{(r)}, \ldots, F_a^{(r)})'$. Furthermore denote:

$$\mathbf{V}_N = (\boldsymbol{\gamma}' \otimes \mathbf{I}_a) \cdot \mathbf{\Lambda}_N \cdot (\boldsymbol{\gamma} \otimes \mathbf{I}_a). \tag{3.9}$$

A detailed description on estimating $\mathbf{V}_N$ can be found in in the upcoming Section 3.5.1. For the Wald-type statistic proposed to work properly, following technical assumptions have to hold:

(**A1**) - (**A3**) For estimating the regression coefficients.

(**A4**) The matrix $\mathbf{V}_N \to \mathbf{V}$ such that $rank(\mathbf{V}_N) = rank(\mathbf{V})$ for all $N \geq N_0 \in \mathbb{N}$.

(**A5**) Let $\theta_{N,i}^{(r)}$, $i = 1, \ldots, a$ and $r = 1, \ldots, d$, denote the eigenvalues of $\mathbf{\Lambda}_N$. Then $N\theta_{N,i}^{(r)} \to 0$ or $N\theta_{N,i}^{(r)} \to \infty$ for all $i = 1, \ldots, a$ and $r = 1, \ldots, d$.

Under these assumptions, Siemer (1999) pp.54-55 shows that under $H_0 : \mathbf{Cp}^* = \mathbf{0}$ following equation holds:

$$Q_{\text{WTS}} = N\widehat{\mathbf{p}}^{*\prime}\mathbf{C}'(\mathbf{C}\widehat{\mathbf{V}}_N\mathbf{C})^{-}\mathbf{C}\widehat{\mathbf{p}}^* \stackrel{\cdot}{\sim} \chi^2_{\widehat{f}}, \tag{3.10}$$

with $\widehat{f} = rank(\mathbf{C}\widehat{\mathbf{V}}_N)$. For a full proof we refer to Siemer (1999) Theorem 4.20.

## 3.4 ANOVA-Type Statistic

The following ANOVA-type statistic proposed by Siemer (1999) is supposed to perform better for smaller sample sizes than the Wald-type statistic. For this we define a matrix $\mathbf{K} = \mathbf{C}'(\mathbf{CC}')^{-}\mathbf{C}$. Instead of estimating the complete matrix $(\mathbf{C}\widehat{\mathbf{V}}_N\mathbf{C})^{-}$, the ANOVA-type statistic only requires the traces of the matrices $\mathbf{K}\widehat{\mathbf{V}}_N$ and $\mathbf{K}\widehat{\mathbf{V}}_N\mathbf{K}\widehat{\mathbf{V}}_N$ to be estimated. Using the so called Box-approximation Box (1954), we are then able to construct a test statistic for the hypothesis $H_0 : \mathbf{Cp}^* = \mathbf{0}$.

When constructing the ANOVA-type statistic, we first reformulate our hypothesis.

**Lemma 3.4.1.** *Let* $\mathbf{K} = \mathbf{C}'(\mathbf{CC}')^{-}\mathbf{C}$ *and* $\mathbf{U}$ *denote an arbitrary vector of appropriate size. Then the hypotheses* $H_0 : \mathbf{CU} = \mathbf{0}$ *and* $H_0 : \mathbf{KU} = \mathbf{0}$ *are equivalent.*

*Proof.* For a detailed proof see Langer (1998), Lemma 5.2. □

Under the assumptions (**A1**)-(**A5**) and additionally:

(**A6**) For all $N$ there exists a $c_0 \in \mathbb{R}$ such that: $tr(\mathbf{K}\widehat{\mathbf{V}}_N) = \sum_{i=1}^{a} \lambda_i \geq c_0 > 0$, where $\lambda_1, \ldots, \lambda_a$ are the eigenvalues of $\mathbf{K}\widehat{\mathbf{V}}_N$,

Siemer (1999) pp.56-59 shows, that under $H_0 : \mathbf{Cp}^* = \mathbf{0}$ following equation holds:

$$Q_N = \widehat{f}\frac{N\widehat{\mathbf{p}}^{*\prime}\mathbf{K}\widehat{\mathbf{p}}^*}{tr\left(\mathbf{K}\widehat{\mathbf{V}}_N\right)} \stackrel{\cdot}{\sim} \chi^2_{\widehat{f}}, \tag{3.11}$$

with $\widehat{f} = \frac{tr(\mathbf{K}\widehat{\mathbf{V}}_N)^2}{tr(\mathbf{K}\widehat{\mathbf{V}}_N\mathbf{K}\widehat{\mathbf{V}}_N)}$. Both test procedures were implemented in R and will be compared to the multiple contrast test procedure.

## 3.5 Multiple Contrast Test Procedures

We will now commence with deriving multiple contrast test procedures (MCTP) for the non-parametric setting including covariates. Multiple contrast test procedures are simultaneous test procedures, which allow for simultaneously testing multiple contrasts. As we have already mentioned in the introduction, the aim of MCTP is to not only provide a testing procedure for global hypotheses, but additionally make post-hoc tests, such as the Bonferroni or Bonferroni-Holm correction, unnecessary, by considering dependencies between individual hypotheses and providing for adjusted p-values of the individual hypotheses.

The main motivation for the MCTP in this thesis comes from Konietschke et al. (2012a), where MCTP for relative treatment effects are derived without covariates. The key in providing for MCTP lies in deriving asymptotic distribution properties of the estimator $\widehat{\mathbf{p}}^* = (\widehat{p}_1^*, \ldots, \widehat{p}_a^*)'$. More specifically, we will commence by proving that the pivotal quantity $\sqrt{N}(\widehat{\mathbf{p}}^* - \mathbf{p}^*)$ asymptotically follows a multivariate normal distribution with mean vector $\mathbf{0}$ and covariance matrix $\mathbf{V}_N$, i.e.:

$$\sqrt{N}(\widehat{\mathbf{p}}^* - \mathbf{p}^*) \overset{\cdot}{\sim} N(\mathbf{0}, \mathbf{V}_N).$$

Using this result, we will then derive a simultaneous test procedure. To simplify the calculations in this section, we will calculate asymptotic results by assuming that $\boldsymbol{\gamma}$ is known.

For a more elegant representation of the upcoming equations, we further set $\gamma^{(0)} = -1$. Then it holds, that:

$$\begin{aligned}
\sqrt{N}(\widehat{\mathbf{p}}^* - \mathbf{p}^*) &= \sqrt{N} \begin{pmatrix} \widehat{p}_1^{(0)} - \sum_{r=1}^{d} \widehat{p}_1^{(r)}\gamma^{(r)} - (p_1^{(0)} - \sum_{r=1}^{d} p_1^{(r)}\gamma^{(r)}) \\ \vdots \\ \widehat{p}_a^{(0)} - \sum_{r=1}^{d} \widehat{p}_a^{(r)}\gamma^{(r)} - (p_a^{(0)} - \sum_{r=1}^{d} p_a^{(r)}\gamma^{(r)}) \end{pmatrix} \\
&= \sqrt{N} \begin{pmatrix} \widehat{p}_1^{(0)} - p_1^{(0)} \\ \vdots \\ \widehat{p}_a^{(0)} - p_a^{(0)} \end{pmatrix} - \sqrt{N} \sum_{r=1}^{d} \begin{pmatrix} \widehat{p}_1^{(r)} - p_1^{(r)} \\ \vdots \\ \widehat{p}_a^{(r)} - p_a^{(r)} \end{pmatrix} \gamma^{(r)} \\
&= -\sqrt{N} \sum_{r=0}^{d} \begin{pmatrix} \widehat{p}_1^{(r)} - p_1^{(r)} \\ \vdots \\ \widehat{p}_a^{(r)} - p_a^{(r)} \end{pmatrix} \gamma^{(r)}.
\end{aligned} \tag{3.12}$$

To continue our calculations we need a result from non-parametric theory.

**Theorem 3.5.1** (Asymptotic Equivalence)**.** *Let $X_{i1}^{(r)}, \ldots, X_{in_i}^{(r)}$ be the observations from the $r$th covariate, $r = 0, \ldots, d$ in the $i$th group, $i = 1, \ldots, a$. Then, assuming that $X_{ik}^{(r)}$ and $X_{il}^{(r)}$ are independent for $k \neq l$, following equation holds for $r = 0, \ldots, d$:*

$$\sqrt{N}\left(\widehat{p}_j^{(r)} - p_j^{(r)}\right) \doteq \sqrt{N} \sum_{i=1}^{a} \omega_i Z_{ij}^{(r)}, \tag{3.13}$$

*where the unweighted form is given by $\omega_i = 1/a$, and:*

$$Z_{ij}^{(r)} = \frac{1}{n_j} \sum_{k=1}^{n_j} F_i^{(r)}(X_{jk}^{(r)}) - \frac{1}{n_i} \sum_{k=1}^{n_i} F_j^{(r)}(X_{ik}^{(r)}) + 1 - 2p_{ij}^{(r)}.$$

*Proof.* By adding an index $r$ to the mean distribution functions, Brunner and Munzel (2013) p.192 prove that:

$$\sqrt{N} \int \widehat{H}^{(r)} d\left(\widehat{F}_j^{(r)} - F_j^{(r)}\right) \doteq \sqrt{N} \int H^{(r)} d\left(\widehat{F}_j^{(r)} - F_j^{(r)}\right) \quad j = 1, \ldots, a, \ r = 0, \ldots, d.$$

Following Brunner and Munzel (2013) p.210 and Konietschke et al. (2012a) it follows that:

$$\begin{aligned}
\sqrt{N}\left(\widehat{p}_j^{(r)} - p_j^{(r)}\right) &\doteq \sqrt{N}\left(\int H^{(r)} d\widehat{F}_j^{(r)} + 1 - \int F_j^{(r)} d\widehat{H}^{(r)} - 2\int H^{(r)} dF_j^{(r)}\right) \\
&= \sqrt{N}\left(\int \sum_{i=1}^{a} \omega_i F_i^{(r)} d\widehat{F}_j^{(r)} - \int F_j^{(r)} d\left(\sum_{i=1}^{a} \omega_i \widehat{F}_i^{(r)}\right) + 1 - 2p_j^{(r)}\right) \\
&= \sqrt{N}\sum_{i=1}^{a} \omega_i \left(\int F_i^{(r)} d\widehat{F}_j^{(r)} - \int F_j^{(r)} d\widehat{F}_i^{(r)} + 1 - 2p_{ij}^{(r)}\right) \\
&= \sqrt{N}\sum_{i=1}^{a} \omega_i \left(\frac{1}{n_j}\sum_{k=1}^{n_j} F_i^{(r)}(X_{jk}^{(r)}) - \frac{1}{n_i}\sum_{k=1}^{n_i} F_j^{(r)}(X_{ik}^{(r)}) + 1 - 2p_{ij}^{(r)}\right) \\
&= \sqrt{N}\sum_{i=1}^{a} \omega_i Z_{ij}^{(r)}
\end{aligned}$$

□

Using Theorem (3.5.1) we attain:

$$-\sqrt{N}\sum_{r=0}^{d}\begin{pmatrix}\widehat{p}_1^{(r)} - p_1^{(r)} \\ \vdots \\ \widehat{p}_a^{(r)} - p_a^{(r)}\end{pmatrix}\gamma^{(r)} \doteq -\sqrt{N}\sum_{r=0}^{d}\begin{pmatrix}\sum_{i=1}^{a}\omega_i Z_{i1}^{(r)} \\ \vdots \\ \sum_{i=1}^{a}\omega_i Z_{ia}^{(r)}\end{pmatrix}\gamma^{(r)},$$

so we attain random variables which are asymptotically equivalent to the pivotal quantity we are examining. This approach bears the advantage, that instead of examining the asymptotic properties of the pivotal quantity $\sqrt{N}(\widehat{\mathbf{p}}^* - \mathbf{p}^*)$, we are now able to examine the asymptotic distribution of the asymptotically equivalent term instead, as the asymptotic distributions will be the same. Using the asymptotic equivalence from (3.13), the covariance matrix of the asymptotically equivalent variable is given by:

$$\mathbf{V}_N = \mathrm{Cov}\left[\begin{pmatrix}\sum_{r=0}^{d}\gamma^{(r)}\sqrt{N}\sum_{i=1}^{a}\omega_i Z_{i1}^{(r)} \\ \vdots \\ \sum_{r=0}^{d}\gamma^{(r)}\sqrt{N}\sum_{i=1}^{a}\omega_i Z_{ia}^{(r)}\end{pmatrix}\right] = \mathrm{Cov}\left[\sqrt{N}\cdot I_a \otimes (\gamma^{(0)},\ldots,\gamma^{(d)})\cdot\begin{pmatrix}\sum_{i=1}^{a}\omega_i Z_{i1}^{(0)} \\ \vdots \\ \sum_{i=1}^{a}\omega_i Z_{i1}^{(d)} \\ \sum_{i=1}^{a}\omega_i Z_{i2}^{(0)} \\ \vdots \\ \sum_{i=1}^{a}\omega_i Z_{ia}^{(d)}\end{pmatrix}\right]$$

$$= \text{Cov}\Bigg[\sqrt{N}\cdot\mathbf{I}_a \otimes \underbrace{(\gamma^{(0)},\dots,\gamma^{(d)})}_{=\boldsymbol{\gamma}'}\cdot\underbrace{\mathbf{I}_{ad}\otimes(\omega_1,\dots,\omega_a)}_{=:\mathbf{W}}\cdot \underbrace{\begin{pmatrix} Z_{11}^{(0)}\\ \vdots \\ Z_{a1}^{(0)}\\ Z_{11}^{(1)}\\ \vdots\\ Z_{a1}^{(d)}\\ Z_{12}^{(0)}\\ \vdots\\ Z_{aa}^{(d)}\end{pmatrix}}_{=:\mathbf{Z}}\Bigg]$$

$$= (\mathbf{I}_a\otimes\boldsymbol{\gamma}')\cdot\mathbf{W}\cdot\underbrace{\text{Cov}[\sqrt{N}\mathbf{Z}]}_{=:\boldsymbol{\Sigma}_N}\cdot\mathbf{W}'\cdot(\mathbf{I}_a\otimes\boldsymbol{\gamma}')'. \tag{3.14}$$

Now the main task for calculating $\mathbf{V}_N$ lies in calculating $\text{Cov}(\sqrt{N}\mathbf{Z}) = \boldsymbol{\Sigma}_N$. For this, let us compute the pairwise covariances for all possible indizes, $\text{Cov}\left(Z_{ij}^{(r)}, Z_{st}^{(u)}\right)$. From Section 2.2.2 we remember that $X_{ik}^{(r)}$ and $X_{jl}^{(u)}$ are independent if $i\neq j$ or if $k\neq l$, and $X_{ik}^{(r)} \overset{i.i.d.}{\sim} F_i^{(r)}$ for $k=1,\dots,n_i$. Using these properties, we attain:

$$\begin{aligned}
&\text{Cov}\left(Z_{ij}^{(r)}, Z_{st}^{(u)}\right)\\
=&\text{Cov}\left(\frac{1}{n_j}\sum_{k=1}^{n_j}F_i^{(r)}(X_{jk}^{(r)}) - \frac{1}{n_i}\sum_{k=1}^{n_i}F_j^{(r)}(X_{ik}^{(r)})\ ,\ \frac{1}{n_t}\sum_{k=1}^{n_t}F_s^{(u)}(X_{tk}^{(u)}) - \frac{1}{n_s}\sum_{k=1}^{n_s}F_t^{(u)}(X_{sk}^{(u)})\right)\\
=&\text{Cov}\left(\frac{1}{n_j}\sum_{k=1}^{n_j}F_i^{(r)}(X_{jk}^{(r)})\ ,\ \frac{1}{n_t}\sum_{k=1}^{n_t}F_s^{(u)}(X_{tk}^{(u)})\right)\\
&-\text{Cov}\left(\frac{1}{n_j}\sum_{k=1}^{n_j}F_i^{(r)}(X_{jk}^{(r)})\ ,\ \frac{1}{n_s}\sum_{k=1}^{n_s}F_t^{(u)}(X_{sk}^{(u)})\right)\\
&-\text{Cov}\left(\frac{1}{n_i}\sum_{k=1}^{n_i}F_j^{(r)}(X_{ik}^{(r)})\ ,\ \frac{1}{n_t}\sum_{k=1}^{n_t}F_s^{(u)}(X_{tk}^{(u)})\right)\\
&+\text{Cov}\left(\frac{1}{n_i}\sum_{k=1}^{n_i}F_j^{(r)}(X_{ik}^{(r)})\ ,\ \frac{1}{n_s}\sum_{k=1}^{n_s}F_t^{(u)}(X_{sk}^{(u)})\right)\\
=&\frac{1}{n_j}\sum_{k=1}^{n_j}\frac{1}{n_t}\sum_{l=1}^{n_t}\text{Cov}\left(F_i^{(r)}(X_{jk}^{(r)})\ ,\ F_s^{(u)}(X_{tl}^{(u)})\right)\\
&-\frac{1}{n_j}\sum_{k=1}^{n_j}\frac{1}{n_s}\sum_{l=1}^{n_s}\text{Cov}\left(F_i^{(r)}(X_{jk}^{(r)})\ ,\ F_t^{(u)}(X_{sl}^{(u)})\right)\\
&-\frac{1}{n_i}\sum_{k=1}^{n_i}\frac{1}{n_t}\sum_{l=1}^{n_t}\text{Cov}\left(F_j^{(r)}(X_{ik}^{(r)})\ ,\ F_s^{(u)}(X_{tl}^{(u)})\right)\\
&+\frac{1}{n_i}\sum_{k=1}^{n_i}\frac{1}{n_s}\sum_{l=1}^{n_s}\underbrace{\text{Cov}\left(F_j^{(r)}(X_{ik}^{(r)})\ ,\ F_t^{(u)}(X_{sl}^{(u)})\right)}_{=0\ \text{ for }\ i\neq s\ \text{ or }\ k\neq l}
\end{aligned}$$

$$
\begin{aligned}
=&\frac{1}{n_j^2}\sum_{k=1}^{n_j}\Big[\mathrm{Cov}\left(F_i^{(r)}(X_{jk}^{(r)}),F_s^{(u)}(X_{jk}^{(u)})\right)\cdot\mathbf{1}_{\{j=t\}}\\
&-\mathrm{Cov}\left(F_i^{(r)}(X_{jk}^{(r)}),F_t^{(u)}(X_{jk}^{(u)})\right)\cdot\mathbf{1}_{\{j=s\}}\Big]\\
&-\frac{1}{n_i^2}\sum_{k=1}^{n_i}\Big[\mathrm{Cov}\left(F_j^{(r)}(X_{ik}^{(r)}),F_s^{(u)}(X_{ik}^{(u)})\right)\cdot\mathbf{1}_{\{i=t\}}\\
&-\mathrm{Cov}\left(F_j^{(r)}(X_{ik}^{(r)}),F_t^{(u)}(X_{ik}^{(u)})\right)\cdot\mathbf{1}_{\{i=s\}}\Big]\\
=&\frac{1}{n_j}\Big[\mathrm{Cov}\left(F_i^{(r)}(X_{j1}^{(r)}),F_s^{(u)}(X_{j1}^{(u)})\right)\cdot\mathbf{1}_{\{j=t\}}\\
&-\mathrm{Cov}\left(F_i^{(r)}(X_{j1}^{(r)}),F_t^{(u)}(X_{j1}^{(u)})\right)\cdot\mathbf{1}_{\{j=s\}}\Big]\\
&-\frac{1}{n_i}\Big[\mathrm{Cov}\left(F_j^{(r)}(X_{i1}^{(r)}),F_s^{(u)}(X_{i1}^{(u)})\right)\cdot\mathbf{1}_{\{i=t\}}\\
&-\underbrace{\mathrm{Cov}\left(F_j^{(r)}(X_{i1}^{(r)}),F_t^{(u)}(X_{i1}^{(u)})\right)}_{=:\theta_{jit}^{(r,u)}}\cdot\mathbf{1}_{\{i=s\}}\Big]\\
=&\frac{1}{n_j}\left(\theta_{ijs}^{(r,u)}\cdot\mathbf{1}_{\{j=t\}}-\theta_{ijt}^{(r,u)}\cdot\mathbf{1}_{\{j=s\}}\right)-\frac{1}{n_i}\left(\theta_{jis}^{(r,u)}\cdot\mathbf{1}_{\{i=t\}}-\theta_{jit}^{(r,u)}\cdot\mathbf{1}_{\{i=s\}}\right).
\end{aligned}
$$

Using the upper calculations, all possible index combinations are listed in the following table:

| | |
|---|---|
| For $i=j\vee t=s$: | $\mathrm{Cov}\left(Z_{ij}^{(r)},Z_{st}^{(u)}\right)=0$ |
| For $(i\neq j\wedge t\neq s)\wedge(j=t\wedge i=s)$: | $\mathrm{Cov}\left(Z_{ij}^{(r)},Z_{st}^{(u)}\right)=\frac{1}{n_j}\theta_{ijs}^{(r,u)}+\frac{1}{n_i}\theta_{jit}^{(r,u)}$ |
| For $(i\neq j\wedge t\neq s)\wedge(j=s\wedge i=t)$: | $\mathrm{Cov}\left(Z_{ij}^{(r)},Z_{st}^{(u)}\right)=-\frac{1}{n_j}\theta_{ijt}^{(r,u)}-\frac{1}{n_i}\theta_{jis}^{(r,u)}$ |
| For $(i\neq j\wedge t\neq s)\wedge(j\neq s,t\wedge i=t)$: | $\mathrm{Cov}\left(Z_{ij}^{(r)},Z_{st}^{(u)}\right)=-\frac{1}{n_i}\theta_{jis}^{(r,u)}$ |
| For $(i\neq j\wedge t\neq s)\wedge(j\neq s,t\wedge i=s)$: | $\mathrm{Cov}\left(Z_{ij}^{(r)},Z_{st}^{(u)}\right)=\frac{1}{n_i}\theta_{jit}^{(r,u)}$ |
| For $(i\neq j\wedge t\neq s)\wedge(j=t\wedge i\neq s,t)$: | $\mathrm{Cov}\left(Z_{ij}^{(r)},Z_{st}^{(u)}\right)=\frac{1}{n_j}\theta_{ijs}^{(r,u)}$ |
| For $(i\neq j\wedge t\neq s)\wedge(j=s\wedge i\neq s,t)$: | $\mathrm{Cov}\left(Z_{ij}^{(r)},Z_{st}^{(u)}\right)=-\frac{1}{n_j}\theta_{ijt}^{(r,u)}$ |
| For $(i\neq j\wedge t\neq s)\wedge(j\neq s,t\wedge i\neq s,t)$: | $\mathrm{Cov}\left(Z_{ij}^{(r)},Z_{st}^{(u)}\right)=0.$ |

We have now successfully calculated the covariance matrix $\mathbf{\Sigma}_N$, by calculating the covariances of $Z_{ij}^{(r)}$ and $Z_{st}^{(u)}$ for all possible index combinations, and therefore the covariance matrix $\mathbf{V}_N=(\mathbf{I}_a\otimes\boldsymbol{\gamma}')\cdot\mathbf{W}\cdot\mathbf{\Sigma}_N\cdot\mathbf{W}'\cdot(\mathbf{I}_a\otimes\boldsymbol{\gamma}')'$. It remains to show the asymptotic multivariate normality of the pivotal quantity $\sqrt{N}(\widehat{\mathbf{p}}^*-\mathbf{p}^*)$ and conclude the connection to $\mathbf{V}_N$. This is done in the following theorem.

**Theorem 3.5.2.** *Let $\mathbf{V}_N$ be as given in the upper calculation. Under the assumptions (**A1**), i.e. there exists $N_0 \in \mathbb{N}$ such that $\frac{N}{n_i} \leq N_0 < \infty$ for $i = 1, \ldots, a$, and that $\mathbf{V}_N \to \mathbf{V}$ such that $rank(\mathbf{V}_N) = rank(\mathbf{V}) \geq 1$ for all $N \geq M_0 < \infty$ is fulfilled, the rank statistic $\sqrt{N}(\widehat{\mathbf{p}}^* - \mathbf{p}^*)$ asymptotically follows a multivariate normal distribution with expectation $\mathbf{0}$ and covariance matrix $\mathbf{V}_N$.*

*Proof.* The theorem can be proven analogously to Theorem 2 in Konietschke et al. (2012a). First we take a close look at the covariance matrix $\mathbf{V}_N$. Let $\lambda_{i,N}$, $i = 1, \ldots, a$, denote the eigenvalues of $\mathbf{V}_N$ where $\lambda_N^{\min} = \min\{\lambda_{N,i} | \lambda_{N,i} > 0, i = 1, \ldots, a\}$ is the smallest eigenvalue larger than zero. Then, by the assumptions of this theorem, there exists a constant $c_0 > 0$ such that $\lambda_N^{\min} \geq c_0$ for all $N \geq M_0$. Without loss of generality let $\lambda_{1,N}, \ldots, \lambda_{j,N} \to 0$ and $\lambda_{j+1,N}, \ldots, \lambda_{a,N} \geq c_0$. Since $\mathbf{V}_N$ is a covariance matrix and therefore a symmetric matrix, by the spectral decomposition theorem there exists an invertible matrix $\mathbf{B}$ such that $\mathbf{B}\mathbf{V}_N\mathbf{B}' = \mathbf{D} = \mathbf{D}_1 \oplus \mathbf{D}_2$ where $\mathbf{D}_1 = dia(\lambda_{1,N}, \ldots, \lambda_{j,N})$ a diagonal matrix with the first $j$ eigenvalues and $\mathbf{D}_2 = dia(\lambda_{j+1,N}, \ldots, \lambda_{a,N})$ a diagonal matrix with the remaining eigenvalues. The asymptotic normality of $\sqrt{N}(\mathbf{I}_a \otimes \boldsymbol{\gamma}')\mathbf{W}\mathbf{Z}$ is now established through the Cramer-Wold device (7.1.9). Let $\mathbf{k} = (k_1, \ldots, k_a)'$ denote an arbitrary vector of constants. Since $\mathbf{B}$ is invertible, there exists a vector $\tilde{\mathbf{k}}$ such that $\mathbf{k}' = \tilde{\mathbf{k}}\mathbf{B}$. From Lindeberg-Feller limit theorem it follows that:

$$\frac{\sqrt{N}\mathbf{k}'(\mathbf{I}_a \otimes \boldsymbol{\gamma}')\mathbf{W}\mathbf{Z}}{\sqrt{\mathrm{Var}(\sqrt{N}\mathbf{k}'(\mathbf{I}_a \otimes \boldsymbol{\gamma}')\mathbf{W}\mathbf{Z})}} = \frac{N\mathbf{k}'(\mathbf{I}_a \otimes \boldsymbol{\gamma}')\mathbf{W}\mathbf{Z}}{\sqrt{\mathrm{Var}(N\mathbf{k}'(\mathbf{I}_a \otimes \boldsymbol{\gamma}')\mathbf{W}\mathbf{Z})}} \xrightarrow{D} N(0,1)$$

Because:

$$\begin{aligned}\mathrm{Var}(N\mathbf{k}'(\mathbf{I}_a \otimes \boldsymbol{\gamma}')\mathbf{W}\mathbf{Z}) &= N \cdot \mathrm{Var}(\sqrt{N}\mathbf{k}'(\mathbf{I}_a \otimes \boldsymbol{\gamma}')\mathbf{W}\mathbf{Z}) = N\mathbf{k}'\mathbf{V}_N\mathbf{k} \\ &= N\tilde{\mathbf{k}}'\mathbf{B}\mathbf{V}_N\mathbf{B}'\tilde{\mathbf{k}} = N\tilde{\mathbf{k}}'(\mathbf{D}_1 \oplus \mathbf{D}_2)\tilde{\mathbf{k}} \geq N \cdot \sum_{s=j+1}^{a} \tilde{k}_s c_0 \to \infty.\end{aligned}$$

Therefore the sum of variances of $N\mathbf{k}'(\mathbf{I}_a \otimes \boldsymbol{\gamma}')\mathbf{W}\mathbf{Z}$ diverges for $N \to \infty$ and Lindeberg's condition (7.1.7) is fulfilled, because the random variables $\frac{N}{n_i} \cdot F_i^{(r)}(X_{jk}^{(r)})$ are uniformly bounded, because $N/n_i \leq N_0 < \infty$. Since it might not be clear how the term $N\mathbf{k}'(\mathbf{I}_a \otimes \boldsymbol{\gamma}')\mathbf{W}\mathbf{Z}$ is written as a sum of independent variables, further calculations can be attained from the appendix in Lemma (7.1.11). □

Theorem (3.5.2) is a central result for developing the multiple contrast test procedure. Before calculating test statistics however, it remains to find an appropriate estimator for the covariance matrix $\mathbf{V}_N$, or subsequently $\boldsymbol{\Sigma}_N$, as can be seen through (3.14). An estimator will be provided in the next section.

### 3.5.1 Estimating the Covariance Matrix

Through the upper calculations of the covariance matrix $\boldsymbol{\Sigma}_N$ it becomes obvious that for estimating $\boldsymbol{\Sigma}_N$, and thus $\mathbf{V}_N$, it is sufficient to estimate the parameters $\theta_{ijs}^{(r,u)}$. Normally, we would estimate $\theta_{ijs}^{(r,u)}$ through:

$$\tilde{\theta}_{ijs}^{(r,u)} = \frac{1}{n_j - 1} \sum_{k=1}^{n_j} \left( F_i^{(r)}(X_{jk}^{(r)}) - \frac{1}{n_j} \sum_{l=1}^{n_j} F_i^{(r)}(X_{jl}^{(r)}) \right) \left( F_s^{(u)}(X_{jk}^{(u)}) - \frac{1}{n_j} \sum_{l=1}^{n_j} F_s^{(u)}(X_{jk}^{(u)}) \right),$$

but because the random variables $F_i^{(r)}(X_{jk}^{(r)})$ are non-observable random variables, $\tilde{\theta}_{ijs}^{(r,u)}$ is not a valid estimator for $\theta_{ijs}$. Therefore, we will replace $F_i^{(r)}(X_{jk}^{(r)})$ by $\widehat{F}_i^{(r)}(X_{jk}^{(r)})$, random variables which are observable and presumably close enough to the non-observable random variables for a valid estimation. The resulting estimator for $\theta_{ijs}^{(r,u)}$ is then given by:

$$\widehat{\theta}_{ijs}^{(r,u)} = \frac{1}{n_j - 1} \sum_{k=1}^{n_j} \left( \widehat{F}_i^{(r)}(X_{jk}^{(r)}) - \frac{1}{n_j} \sum_{l=1}^{n_j} \widehat{F}_i^{(r)}(X_{jl}^{(r)}) \right) \cdot \left( \widehat{F}_s^{(u)}(X_{jk}^{(u)}) - \frac{1}{n_j} \sum_{l=1}^{n_j} \widehat{F}_s^{(u)}(X_{jl}^{(u)}) \right).$$

Using the mean rank notation introduced in Section 2.4 enables us to denote the proposed estimator $\widehat{\theta}_{ijs}^{(r,u)}$ using ranks by:

$$\begin{aligned} \widehat{\theta}_{ijs}^{(r,u)} = \frac{1}{n_j - 1} \sum_{k=1}^{n_j} & \frac{1}{n_i} \left( R_{jk}^{(r)}(ij) - R_{jk}^{(r)}(j) - \overline{R}_{j\cdot}^{(r)}(ij) + \frac{n_j + 1}{2} \right) \\ & \cdot \frac{1}{n_s} \left( R_{jk}^{(u)}(sj) - R_{jk}^{(u)}(j) - \overline{R}_{j\cdot}^{(u)}(sj) + \frac{n_j + 1}{2} \right). \end{aligned}$$

With these estimators, we are able to estimate $\mathbf{V}_N$ consistently. Let the proposed estimator, replacing $\theta_{ijs}^{(r,u)}$ by $\widehat{\theta}_{ijs}^{(r,u)}$, be denoted by $\widehat{\mathbf{V}}_N$, and $\widehat{\mathbf{\Sigma}}_N$ the estimator for $\mathbf{\Sigma}_N$ accordingly. The consistency of $\widehat{\mathbf{V}}_N$ is proven in the following theorem.

**Theorem 3.5.3.** *Under the assumption (**A1**), i.e. $N \to \infty$ such that $N/n_i \leq N_0 < \infty$ for $N_0 \in \mathbb{N}$ and all $i = 1, \ldots, a$, it holds that: $\widehat{\mathbf{V}}_N - \mathbf{V}_N \xrightarrow{a.s.} \mathbf{0}$.*

*Proof.* The theorem can be proven analogously to Lemma A.2 and Theorem A.4 from Konietschke et al. (2012a). The key in proving the consistency of $\widehat{\mathbf{V}}_N$ lies in proving the consistency of $\widehat{\mathbf{\Sigma}}_N$. By the strong law of large numbers it holds that $\tilde{\theta}_{ijs}^{(r,u)} - \theta_{ijs}^{(r,u)} \xrightarrow{a.s.} 0$ when $n_j, n_s \to \infty$ . Because the number of groups $a$ and number of covariates $d$ is bounded, it is sufficient to prove the consistency of $\widehat{\mathbf{\Sigma}}_N$ componentwise. Therefore, the proof amounts to showing: $|\tilde{\theta}_{ijs}^{(r,u)} - \widehat{\theta}_{ijs}^{(r,u)}| \xrightarrow{a.s.} 0$. For this denote:

$$\tilde{D}_{ijk}^{(r)} = F_i^{(r)}(X_{jk}^{(r)}) - \frac{1}{n_j} \sum_{l=1}^{n_j} F_i^{(r)}(X_{jl}^{(r)}),$$

and its empirical counterpart by:

$$\widehat{D}_{ijk}^{(r)} = \widehat{F}_i^{(r)}(X_{jk}^{(r)}) - \frac{1}{n_j} \sum_{l=1}^{n_j} \widehat{F}_i^{(r)}(X_{jl}^{(r)}).$$

Then it holds that:

$$\begin{aligned} |\tilde{\theta}_{ijs}^{(r,u)} - \widehat{\theta}_{ijs}^{(r,u)}| &= \left| \frac{1}{n_j - 1} \sum_{k=1}^{n_j} \tilde{D}_{ijk}^{(r)} \tilde{D}_{sjk}^{(u)} - \widehat{D}_{ijk}^{(r)} \widehat{D}_{sjk}^{(u)} \right| \\ &= \left| \frac{1}{n_j - 1} \sum_{k=1}^{n_j} \tilde{D}_{ijk}^{(r)} \tilde{D}_{sjk}^{(u)} - \widehat{D}_{ijk}^{(r)} \widehat{D}_{sjk}^{(u)} \pm \tilde{D}_{ijk}^{(r)} \widehat{D}_{sjk}^{(u)} \right| \\ &= \left| \frac{1}{n_j - 1} \sum_{k=1}^{n_j} \tilde{D}_{ijk}^{(r)} (\tilde{D}_{sjk}^{(u)} - \widehat{D}_{sjk}^{(u)}) - \widehat{D}_{sjk}^{(u)} (\widehat{D}_{ijk}^{(r)} - \tilde{D}_{ijk}^{(r)}) \right| \end{aligned}$$

$$
\begin{aligned}
&\leq \frac{1}{n_j - 1} \sum_{k=1}^{n_j} \left| \tilde{D}_{ijk}^{(r)} (\tilde{D}_{sjk}^{(u)} - \widehat{D}_{sjk}^{(u)}) - \widehat{D}_{sjk}^{(u)} (\widehat{D}_{ijk}^{(r)} - \tilde{D}_{ijk}^{(r)}) \right| \\
&\leq \frac{1}{n_j - 1} \sum_{k=1}^{n_j} \left| \tilde{D}_{ijk}^{(r)} \right| \left| \tilde{D}_{sjk}^{(u)} - \widehat{D}_{sjk}^{(u)} \right| + \left| \widehat{D}_{sjk}^{(u)} \right| \left| \widehat{D}_{ijk}^{(r)} - \tilde{D}_{ijk}^{(r)} \right| \\
&\leq \frac{1}{n_j - 1} \sum_{k=1}^{n_j} \left| \tilde{D}_{sjk}^{(u)} - \widehat{D}_{sjk}^{(u)} \right| + \frac{1}{n_j - 1} \sum_{k=1}^{n_j} \left| \widehat{D}_{ijk}^{(r)} - \tilde{D}_{ijk}^{(r)} \right| \\
&\leq \frac{n_j}{n_j - 1} \max_{k=1,\dots,n_j} \left| \tilde{D}_{sjk}^{(u)} - \widehat{D}_{sjk}^{(u)} \right| + \frac{n_j}{n_j - 1} \max_{k=1,\dots,n_j} \left| \widehat{D}_{ijk}^{(r)} - \tilde{D}_{ijk}^{(r)} \right|, \qquad (3.15)
\end{aligned}
$$

where we used $\left|\tilde{D}_{ijk}^{(r)}\right| \leq 1$ and $\left|\widehat{D}_{sjk}^{(u)}\right| \leq 1$. Considering that:

$$
\begin{aligned}
\left| \widehat{D}_{ijk}^{(r)} - \tilde{D}_{ijk}^{(r)} \right| &= \left| \widehat{F}_i^{(r)}(X_{jk}^{(r)}) - \frac{1}{n_j} \sum_{l=1}^{n_j} \widehat{F}_i^{(r)}(X_{jl}^{(r)}) - F_i^{(r)}(X_{jk}^{(r)}) + \frac{1}{n_j} \sum_{l=1}^{n_j} F_i^{(r)}(X_{jl}^{(r)}) \right| \\
&\leq \left| \widehat{F}_i^{(r)}(X_{jk}^{(r)}) - F_i^{(r)}(X_{jk}^{(r)}) \right| + \left| \frac{1}{n_j} \sum_{l=1}^{n_j} \left( F_i^{(r)}(X_{jk}^{(r)}) - \widehat{F}_i^{(r)}(X_{jk}^{(r)}) \right) \right| \\
&\leq 2 \cdot \max_{k=1,\dots,n_j} \left| F_i^{(r)}(X_{jk}^{(r)}) - \widehat{F}_i^{(r)}(X_{jk}^{(r)}) \right|
\end{aligned}
$$

and denoting the supremum norm by $\|\cdot\|_\infty$, equation (3.15) yields that:

$$
\begin{aligned}
&\frac{n_j}{n_j - 1} \max_{k=1,\dots,n_j} \left| \tilde{D}_{sjk}^{(u)} - \widehat{D}_{sjk}^{(u)} \right| + \frac{n_j}{n_j - 1} \max_{k=1,\dots,n_j} \left| \widehat{D}_{ijk}^{(r)} - \tilde{D}_{ijk}^{(r)} \right| \\
&\leq \frac{2 \cdot n_j}{n_j - 1} \left( \max_{k=1,\dots,n_j} \left| F_s^{(u)}(X_{jk}^{(u)}) - \widehat{F}_s^{(u)}(X_{jk}^{(u)}) \right| + \max_{k=1,\dots,n_j} \left| F_i^{(r)}(X_{jk}^{(r)}) - \widehat{F}_i^{(r)}(X_{jk}^{(r)}) \right| \right) \\
&\leq \frac{2 \cdot n_j}{n_j - 1} \left( \left\| F_s^{(u)} - \widehat{F}_s^{(u)} \right\|_\infty + \left\| F_s^{(r)} - \widehat{F}_s^{(r)} \right\|_\infty \right) \xrightarrow{a.s.} 0,
\end{aligned}
$$

where we used that $2n_j/(n_j - 1) \to 2$ and from the Glivenko-Cantelli theorem (7.1.4) it follows that $\left\| F_s^{(u)} - \widehat{F}_s^{(u)} \right\|_\infty \xrightarrow{a.s.} 0$. Other index combinations of the $\theta_{ijs}^{(r,u)}$ are proven in the same manner, only with different indexing.

From the upper calculations and the assumption **(A1)** it follows, that:

$$
\begin{aligned}
\widehat{\mathbf{V}}_N - \mathbf{V}_N &= (\mathbf{I}_a \otimes \boldsymbol{\gamma}') \cdot \mathbf{W} \cdot \widehat{\boldsymbol{\Sigma}}_N \cdot \mathbf{W}' \cdot (\mathbf{I}_a \otimes \boldsymbol{\gamma}')' - (\mathbf{I}_a \otimes \boldsymbol{\gamma}') \cdot \mathbf{W} \cdot \boldsymbol{\Sigma}_N \cdot \mathbf{W}' \cdot (\mathbf{I}_a \otimes \boldsymbol{\gamma}')' \\
&= (\mathbf{I}_a \otimes \boldsymbol{\gamma}') \cdot \mathbf{W} \cdot \underbrace{(\widehat{\boldsymbol{\Sigma}}_N - \boldsymbol{\Sigma}_N)}_{\xrightarrow{a.s.} \mathbf{0}} \cdot \mathbf{W}' \cdot (\mathbf{I}_a \otimes \boldsymbol{\gamma}')' \xrightarrow{a.s.} \mathbf{0},
\end{aligned}
$$

which proves the claim. Note, that plugging in the estimator $\widehat{\boldsymbol{\gamma}}$ for $\boldsymbol{\gamma}$ as proposed in Section 3.7, will weaken the convergence to $\widehat{\mathbf{V}}_N - \mathbf{V}_N \xrightarrow{P} \mathbf{0}$. □

The results from this section, Theorem (3.5.2) and Theorem (3.5.3), will now be used to derive the MCTP.

### 3.5.2 Derivation of Test Statistics

In Theorem (3.5.2) we have shown that:

$$\sqrt{N}(\widehat{\mathbf{p}}^* - \mathbf{p}^*) \dot{\sim} N(\mathbf{0}, \mathbf{V}_N). \tag{3.16}$$

The main motivation for using MCTP lies in testing hypotheses to uncover differences between factor levels. We will be testing hypotheses using a contrast matrix $\mathbf{C} \in \mathbb{R}^{q\times a}$ (2.3.1) with row vectors $\mathbf{c}_1, \ldots, \mathbf{c}_q$, the desired contrasts, i.e. $\mathbf{c}_l = (c_{l1}, \ldots, c_{la})'$. Our goal will be to derive multiple contrast test procedures for the family of hypotheses:

$$\Omega^{p^*} = \left\{ H_0^{\mathbf{p}^*} : \mathbf{c}_l'\mathbf{p}^* = 0 \mid l = 1, \ldots, q \right\}, \tag{3.17}$$

by providing compatible *simultaneous confidence intervals* for the effects $\delta_l = \mathbf{c}_l'\mathbf{p}^*$. Simultaneous confidence intervals are confidence intervals for the effects $\delta_l$, $l = 1, \ldots, q$, such that if 0 is not contained within at least one of the confidence intervals, then the *global hypothesis*, i.e. $\bigcap_{l=1}^{q}\{H_0^{\mathbf{p}^*} : \mathbf{c}_l'\mathbf{p}^* = 0\}$ or simply $H_0 : \mathbf{C}\mathbf{p}^* = \mathbf{0}$, can be rejected to a pre-specified level $\alpha$. In order to do this, we will begin by deriving test statistics for the *individual hypotheses* $H_0^{\mathbf{p}^*} : \mathbf{c}_l'\mathbf{p}^* = 0$, for $l = 1, \ldots, q$. We define test statistics for the individual hypotheses by:

$$T_l^{\mathbf{p}^*} = \sqrt{N}\mathbf{c}_l'(\widehat{\mathbf{p}}^* - \mathbf{p}^*)/\sqrt{\widehat{v}_{ll}}, \tag{3.18}$$

where $\widehat{v}_{lm} = \mathbf{c}_l'\widehat{\mathbf{V}}_N\mathbf{c}_m$. Then by Theorem (3.5.2), Theorem (3.5.3) and Slutsky's theorem it follows, that $T_l^{\mathbf{p}^*} \xrightarrow{D} N(0,1)$ under the null hypothesis $H_0 : \mathbf{c}_l'\mathbf{p}^* = 0$. For developing MCTP we will require not only the distribution of an individual test statistic, but more importantly the joint distribution of all individual test statistics, i.e. we further need the covariance structure of the vector of test statistics, $\mathbf{T} = (T_1^{\mathbf{p}^*}, \ldots, T_q^{\mathbf{p}^*})$. Denote $\mathbf{R} = \mathrm{Cov}(\mathbf{T})$, and assuming $\mathbf{V}_N$ is known, with:

$$\begin{aligned}
\mathrm{Cov}(T_l^{\mathbf{p}^*}, T_m^{\mathbf{p}^*}) &= \mathrm{Cov}(\sqrt{N}\mathbf{c}_l'(\widehat{\mathbf{p}}^* - \mathbf{p}^*)/\sqrt{v_{ll}}\ ,\ \sqrt{N}\mathbf{c}_m'(\widehat{\mathbf{p}}^* - \mathbf{p}^*)/\sqrt{v_{mm}}) \\
&= \mathrm{Cov}\left(\sqrt{N}\sum_{i=1}^{a} c_{li}(\widehat{p}_i^* - p_i^*), \sqrt{N}\sum_{j=1}^{a} c_{mj}(\widehat{p}_j^* - p_j^*)\right)/\sqrt{v_{ll}v_{mm}} \\
&= \sum_{i=1}^{a} c_{li}\sum_{j=1}^{b} c_{mj}\mathrm{Cov}\left(\sqrt{N}(\widehat{p}_i^* - p_i^*), \sqrt{N}(\widehat{p}_j^* - p_j^*)\right)/\sqrt{v_{ll}v_{mm}} \\
&= \mathbf{c}_l'\mathbf{V}_N\mathbf{c}_m/\sqrt{v_{ll}v_{mm}} = v_{lm}/\sqrt{v_{ll}v_{mm}},
\end{aligned}$$

we can calculate the covariance matrix of $\mathbf{T}$, $\mathrm{Cov}(\mathbf{T}) = \mathbf{R} = (r_{lm})_{l,m=1,\ldots,q}$ where $r_{lm} = v_{lm}/\sqrt{v_{ll}v_{mm}}$. From Theorem (3.5.2), Slutsky's theorem and the upper calculations, it follows that:

$$\mathbf{T} \dot{\sim} N(\mathbf{0}, \mathbf{R}). \tag{3.19}$$

Our goal will be to use this multivariate asymptotic property of the vector of test statistics $\mathbf{T}$ for conceiving simultaneous test procedures. The idea behind a simultaneous test procedure is to control the type I error for the global hypothesis $\bigcap_{l=1}^{q}\{H_0^{\mathbf{p}^*} : \mathbf{c}_l'\mathbf{p}^* = 0\}$ by properly adjusting the individual hypotheses from $\Omega^{\mathbf{p}^*}$ (3.17). The global hypothesis is rejected, when any adjusted individual hypothesis is rejected. To better understand the notion of simultaneous test procedures, let us introduce some important terminology from this field, before theoretically justifying this testing approach. Most of the terminology on simultaneous testing, and the ideas represented, can be found in a similar manner in Gabriel (1969) and Konietschke et al. (2012a).

**Definition 3.5.4** (Joint testing family)**.** *Let $\Pi = \{\pi_i \mid i \in I\}$ be a family of hypotheses. Further let $S_i$ be a real valued statistic corresponding to the hypotheses $\pi_i$ and $S = \{S_i \mid i \in I\}$. The collection $\{\Pi, S\}$ of hypotheses and their corresponding statistics will be called a **testing family**, provided the distribution of $S_i$ is completely specified under $\pi_i$, for all $i \in I$. If, for any subfamily $\tilde{\Pi} = \{\pi_i \mid i \in \tilde{I}\}$ where $\tilde{I} \subseteq I$, the joint distribution of all $S_i$, $i \in \tilde{I}$, is completely specified under $\tilde{\pi}_0 = \bigcap_{\tilde{I}} \pi_i$, the testing family will be called **joint**.*

**Definition 3.5.5** (Simultaneous Test Procedure)**.** *Let $\{\Pi, S\}$ be a joint testing family. Further, let $\xi$ be a critical value, then a **simultaneous test procedure** is defined as the family of tests of all $\pi_i \in \Pi$ which reject any $\pi_i, i \in I$, if $S_i > \xi$, using the same constant $\xi$ for all $S_i \in S$. Such a simultaneous test procedure will be denoted $\{\Pi, S, \xi\}$. The probability:*

$$\alpha = 1 - P_{H_0}(S_1 < \xi, \ldots, S_q < \xi) \tag{3.20}$$

*of falsely rejecting the intersection hypothesis, or global hypothesis, is referred to as the **level** of the simultaneous test procedure.*

Our goal now will be to first prove, that $\{\Omega^{\mathbf{P}^*}, \mathbf{T}\}$ is in fact a joint testing family, before deriving a simultaneous test procedure for testing the global hypothesis through the individual hypotheses within $\Omega^{\mathbf{P}^*}$. This gives way to following Lemma.

**Lemma 3.5.6.** *The family of hypotheses $\Omega^{\mathbf{P}^*}$ and the corresponding test statistics $\mathbf{T}$ asymptotically constitute a joint testing family.*

*Proof.* The proof can be shown analogously to Lemma 1 Konietschke et al. (2012a). We have shown that $\mathbf{T}$ asymptotically follows a multivariate normal distribution with expectation $\mathbf{0}$ and correlation matrix $\mathbf{R}$. Therefore, the joint distribution of $\mathbf{T}$ is completely specified under $H_0^{\mathbf{P}^*} : \mathbf{c}_l'\mathbf{p}^* = 0$ for all $l = 1, \ldots, q$. We have also shown, that each test statistic $T_l^{\mathbf{P}^*}$ converges in distribution to the standard normal distribution. In particular, the asymptotic distribution of $T_l^{\mathbf{P}^*}$ is independent from the distribution of $T_m^{\mathbf{P}^*}$ for $m \neq l$. Therefore, the asymptotic joint distribution of $\mathbf{T}^J = \left(T_j^{\mathbf{P}^*} | j \in J\right)$ is completely specified under arbitrary intersections of the hypotheses $\bigcap_{j \in J} \left\{H_0^{p^*} : \mathbf{c}_j'\mathbf{p}^* = 0\right\}$, where $J \subseteq \{1, \ldots, q\}$ denotes an arbitrary set of indizes. Therefore, $\{\Omega^{\mathbf{P}^*}, \mathbf{T}\}$ is per Definition (3.5.4) a joint testing family. □

Now that we have proven, that $\{\Omega^{\mathbf{P}^*}, \mathbf{T}\}$ constitutes a joint testing family in the sense of Definition (3.5.4), we further need an appropriate critical value for deriving a simultaneous test procedure. Let $z_{1-\alpha,2,\mathbf{R}}$ denote the two-sided equi-coordinate $(1-\alpha)$-quantile of $N(\mathbf{0}, \mathbf{R})$, i.e.:

$$P\left(\bigcap_{l=1}^{q}\{-z_{1-\alpha,2,\mathbf{R}} \leq X_l \leq z_{1-\alpha,2,\mathbf{R}}\}\right) = 1 - \alpha \tag{3.21}$$

for $(X_1, \ldots, X_q) \sim N(0, \mathbf{R})$. Then this is, as we will yet see, an appropriate choice for a critical value.

Figure 3 illustrates the use of equi-coordinate quantiles for a 2-dimensional multivariate normal distribution:

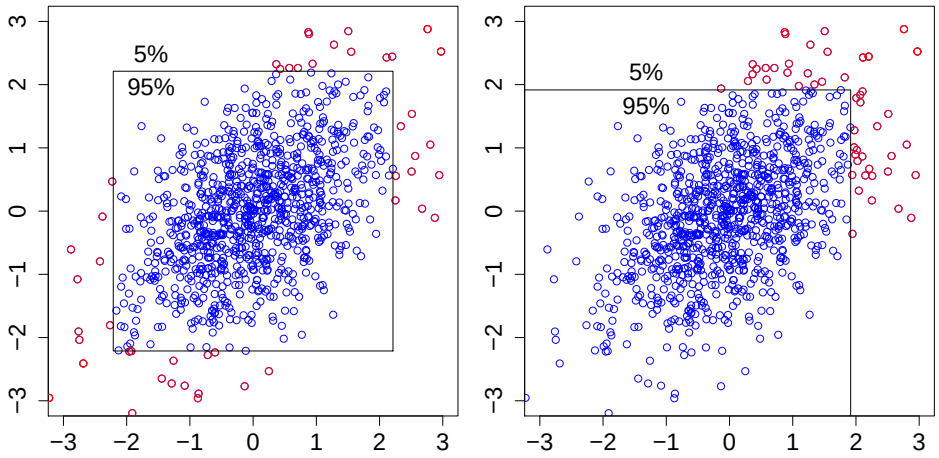


*Figure 3:* Two-sided (*left*) and one-sided (*right*) equi-coordinate 95%-quantiles of a 2-dimensional multivariate standard normal distribution with correlation $\rho = 0.5$.

We write $z_{1-\alpha,2,\mathbf{R}}$ to emphasize that it is the two-sided equi-coordinate quantile; one-sided quantiles are written as $z_{1-\alpha,1,\mathbf{R}}$. For further details on equi-coordinate quantiles and their numerical computation we refer to Bretz et al. (2001) and Genz and Bretz (2009).

The only problem which remains is that the correlation matrix $\mathbf{R}$ is unknown and must be estimated. For this, we can use the estimator $\widehat{\mathbf{V}}_N$ for $\mathbf{V}_N$, given in Section 3.5.1. Then by Slutsky's theorem, an estimator for $\mathbf{R}$ is given by $\widehat{\mathbf{R}} = (\widehat{r}_{lm})_{l,m=1,\ldots,q}$ where $\widehat{r}_{lm} = \widehat{v}_{lm}/\sqrt{\widehat{v}_{ll}\widehat{v}_{mm}}$. Taking the estimator $\widehat{\mathbf{R}}$ into consideration instead of the unknown correlation matrix $\mathbf{R}$, the set $\{\Omega^{p^*}, \mathbf{T}, z_{1-\alpha,2,\widehat{\mathbf{R}}}\}$ asymptotically constitutes a simultaneous test procedure, which controls the family wise error rate in the strong sense. This is shown in the next theorem.

**Theorem 3.5.7.** *The simultaneous test procedure* $\{\Omega^{\mathbf{p}^*}, \mathbf{T}, z_{1-\alpha,2,\widehat{\mathbf{R}}}\}$ *asymptotically controls the family wise error rate in the strong sense.*

*Proof.* The theorem can be proven analogously to Theorem 3 from Konietschke et al. (2012a). From Lemma (3.5.6) we know that $\{\Omega^{\mathbf{p}^*}, \mathbf{T}\}$ asymptotically constitutes a joint testing family. By construction, the simultaneous test procedure $\{\Omega^{\mathbf{p}^*}, \mathbf{T}, z_{1-\alpha,2,\mathbf{R}}\}$ is coherent, i.e. a hypothesis is rejected if any hypothesis implied by it is rejected. Then the assumptions of Theorem 7.1.1 are fulfilled. Now we replace $\mathbf{R}$ by its consistent estimator $\widehat{\mathbf{R}}$. Since the map $f : \mathbb{R}^{q\times q} \to \mathbb{R}$ with $f(\mathbf{R}) = z_{1-\alpha,2,\mathbf{R}}$ is continuous, by the continuous mapping theorem it follows that $f(\mathbf{R}) - f(\widehat{\mathbf{R}}) = z_{1-\alpha,2,\mathbf{R}} - z_{1-\alpha,2,\widehat{\mathbf{R}}} \xrightarrow{P} 0$. Therefore, the assumptions of Theorem 7.1.1 are asymptotically fulfilled and the probability of rejecting one or more true hypotheses $H_0^{\mathbf{p}^*} : \mathbf{c}_l'\mathbf{p}^* = 0$ if all hypotheses are true for $l = 1, \ldots, q$ is equal to the $\alpha$-level of the simultaneous test procedure. Additionally, if not all hypotheses are true, but $H_0^{\mathbf{p}^*} : \mathbf{c}_l'\mathbf{p}^* = 0$ for $l \in J \subseteq \{1, \ldots, q\}$ is true, the probability of rejecting at least one true hypothesis is at most $\alpha$. □

Theorem (3.5.7) justifies a test decision on the global hypothesis $\bigcap_{l=1}^{q}\{H_0^{\mathbf{p}^*} : \mathbf{c}_l'\mathbf{p}^* = 0\}$ by calculating confidence intervals for the individual hypotheses and performing a simultaneous

test procedure. More precisely, Theorem (3.5.7) allows us to construct confidence intervals for the individual hypothesis $H_0 : \mathbf{c}_l'\mathbf{p}^*$ by using the test statistics $T_l^{\mathbf{p}^*}$ with a critical value $z_{1-\alpha,2,\widehat{\mathbf{R}}}$. The simultaneous confidence intervals for a two-sided individual hypotheses are given by:

$$\left[\mathbf{c}_l'\widehat{\mathbf{p}}^* - z_{1-\alpha,2,\widehat{\mathbf{R}}}\sqrt{\widehat{v}_{ll}/N}\ ;\ \mathbf{c}_l'\widehat{\mathbf{p}}^* + z_{1-\alpha,2,\widehat{\mathbf{R}}}\sqrt{\widehat{v}_{ll}/N}\right],\ l = 1,\ldots,q. \tag{3.22}$$

By the construction of these simultaneous confidence intervals, a test decision for the hypothesis $H_0^{\mathbf{p}^*} : \mathbf{c}_l'\mathbf{p}^* = 0$ can be made, i.e. whenever zero is not contained within the corresponding confidence interval, the hypothesis is rejected. In such a case, the global hypothesis $H_0^{\mathbf{p}^*} : \mathbf{C}\mathbf{p}^* = \mathbf{0}$ would also be rejected, at level $\alpha$. One-sided testing is also possible by using the one-sided equi-coordinate quantiles:

$$\left[-\infty\ ;\ \mathbf{c}_l'\widehat{\mathbf{p}}^* + z_{1-\alpha,1,\widehat{\mathbf{R}}}\sqrt{\widehat{v}_{ll}/N}\right]\ \text{ or }\ \left[\mathbf{c}_l'\widehat{\mathbf{p}}^* - z_{1-\alpha,1,\widehat{\mathbf{R}}}\sqrt{\widehat{v}_{ll}/N}\ ;\ \infty\right], \tag{3.23}$$

where the confidence interval on the left corresponds to the hypothesis $H_0 : \mathbf{c}_l'\widehat{\mathbf{p}}^* \geq 0$ and the confidence interval on the right to the hypothesis $H_0 : \mathbf{c}_l'\widehat{\mathbf{p}}^* \leq 0$. The derived simultaneous confidence intervals in (3.22) are the heart and soul of the multiple contrast test procedures and give the possibility not only to decide on the global hypothesis, but more importantly to know between which groups a difference occurs, without having to perform post-hoc procedures. Of course, we could have just as well calculated the $p$-values of the test statistic $\mathbf{T}$ through a multivariate normal distribution (this is done as well in the presented R-script, Section 7.3), but often confidence intervals give the applicant a better feeling for the significance of the result and are therefore required. For example, the European Medicines Agency's guideline on clinical evaluation of diagnostic agents states that:

> "The impact on diagnostic thinking may be presented numerically; the rate of cases where diagnostic uncertainty with a new agent has decreased as compared to pre-test diagnosis should be reported (percentage, and confidence intervals). Positive and negative predictive values may help clinicians modify diagnostic thinking if reasonable thresholds have been reached."(EMEA 2009, Section 7.3, p.15)

Since the multivariate normal distribution of the test statistic $\mathbf{T}$ is only valid asymptotically, it is not clear how good the procedure will perform on finite data. The performance in terms of $\alpha$-level and power on finite data can be attained from a simulation study in Section 4. In the next section, we will try to approximate the distribution of the test statistic $\mathbf{T}$ to improve the test procedure in terms of $\alpha$-level and power.

### 3.5.3 Small Sample Size Approximation

Although the above procedure has been proven to work asymptotically, the convergence to the multivariate normal distribution tends to be rather slow, especially when confronted with small sample sizes or a large number of treatment groups. For this, we propose a further approach which is supposed to approximate the multivariate distribution of the test statistics faster, hence being more suited for small sample sizes.

The small sample size approximation we consider is an adaptation of Gao et al. (2008) and Konietschke et al. (2012a) multivariate t-approximation, to which we introduce covariates. The idea behind this approximation is that instead of taking equi-coordinate quantiles from the multivariate normal distribution, we take equi-coordinate quantiles from

a multivariate t-distribution with $\nu$ degrees of freedom, in hopes of improving the simultaneous confidence intervals (3.22) in terms of coverage probability. The only question which remains is how many degrees of freedom the multivariate t-distribution should have for a useful approximation. For this, let $\mathbf{c}_l = (c_{l1}, \ldots, c_{la})'$, $l = 1, \ldots, q$ denote the $l$th row of the contrast matrix $\mathbf{C}$, of which hypotheses are formed. Further, denote $\mathbf{p}^{(r)} = (p_1^{(r)}, \ldots, p_a^{(r)})'$, its estimator as $\widehat{\mathbf{p}}^{(r)} = (\widehat{p}_1^{(r)}, \ldots, \widehat{p}_a^{(r)})'$ and a further term $A_{lik}^{(r)} = c_{li}\left(H^{(r)}(X_{ik}^{(r)}) - \omega_i F_i^{(r)}(X_{ik}^{(r)})\right) - \sum_{s \neq i} c_{ls}\omega_i F_s^{(r)}(X_{ik}^{(r)})$. Then we can calculate a Satterthwaite approximation for the degrees of freedom by:

$$\begin{aligned}
\sqrt{N}\mathbf{c}_l'(\widehat{\mathbf{p}}^* - \mathbf{p}^*) &= \sqrt{N}\sum_{i=1}^{a} c_{li}(\widehat{p}_i^* - p_i^*) \\
&= \sqrt{N}\sum_{i=1}^{a} c_{li}\sum_{r=0}^{d}\gamma^{(r)}(\widehat{p}_i^{(r)} - p_i^{(r)}) \\
&= \sum_{r=0}^{d}\gamma^{(r)}\sqrt{N}\sum_{i=1}^{a} c_{li}(\widehat{p}_i^{(r)} - p_i^{(r)}) \\
&= \sum_{r=0}^{d}\gamma^{(r)}\sqrt{N}\mathbf{c}_l'(\widehat{\mathbf{p}}^{(r)} - \mathbf{p}^{(r)}) \\
&\doteq \sum_{r=0}^{d}\gamma^{(r)}\sqrt{N}\left[\sum_{i=1}^{a}\frac{1}{n_i}\sum_{k=1}^{n_i} A_{lik}^{(r)} - 2\sum_{i=1}^{a} c_{li}p_i^{(r)}\right] \qquad (3.24) \\
&= \sqrt{N}\sum_{i=1}^{a}\frac{1}{n_i}\sum_{k=1}^{n_i}\sum_{r=0}^{d}\gamma^{(r)}A_{lik}^{(r)} - \sqrt{N}\sum_{r=0}^{d}\gamma^{(r)}\cdot 2\sum_{i=1}^{a} c_{li}p_i^{(r)},
\end{aligned}$$

where the asymptotic equivalence in (3.24) is proven in Lemma (7.1.10). Because the terms $A_{lik}^{(r)}$ and $A_{l'i'k'}^{(r')}$ are independent for $k \neq k'$ and $i \neq i'$, we obtain:

$$\begin{aligned}
\operatorname{Var}\left[\sqrt{N}\sum_{i=1}^{a}\frac{1}{n_i}\sum_{k=1}^{n_i}\sum_{r=0}^{d}\gamma^{(r)}A_{lik}^{(r)}\right] &= N\sum_{i=1}^{a}\frac{1}{n_i}\operatorname{Var}\left[\sum_{r=0}^{d}\gamma^{(r)}A_{li1}^{(r)}\right] \\
&= N\sum_{i=1}^{a}\frac{1}{n_i}\boldsymbol{\gamma}'\underbrace{\operatorname{Cov}\left[\begin{pmatrix}A_{li1}^{(0)}\\ \vdots \\ A_{li1}^{(d)}\end{pmatrix}\right]}_{=:\boldsymbol{\Lambda}_{li}}\boldsymbol{\gamma} = N\sum_{i=1}^{a}\frac{\eta_{li}}{n_i},
\end{aligned}$$

with $\eta_{li} = \boldsymbol{\gamma}'\boldsymbol{\Lambda}_{li}\boldsymbol{\gamma}$. Since the covariance matrices $\boldsymbol{\Lambda}_{li}$ are unknown, they have to be estimated. Using $\widehat{A}_{lik}^{(r)} = c_{li}\left(\widehat{H}^{(r)}(X_{ik}^{(r)}) - \omega_i\widehat{F}_i^{(r)}(X_{ik}^{(r)})\right) - \sum_{s\neq i} c_{ls}\omega_i\widehat{F}_s^{(r)}(X_{ik}^{(r)})$, an estimator for $\boldsymbol{\Lambda}_{li}$ is given by:

$$\widehat{\boldsymbol{\Lambda}}_{li} = \begin{pmatrix}\widehat{A}_{li1}^{(0)} & \ldots & \widehat{A}_{li1}^{(d)} \\ \vdots & \ddots & \vdots \\ \widehat{A}_{lin_i}^{(0)} & \ldots & \widehat{A}_{lin_i}^{(d)}\end{pmatrix}' \cdot \mathbf{P}_{n_i} \cdot \begin{pmatrix}\widehat{A}_{li1}^{(0)} & \ldots & \widehat{A}_{li1}^{(d)} \\ \vdots & \ddots & \vdots \\ \widehat{A}_{lin_i}^{(0)} & \ldots & \widehat{A}_{lin_i}^{(d)}\end{pmatrix}. \qquad (3.25)$$

This enables us to estimate $\eta_{li}$ using $\widehat{\eta}_{li} = \boldsymbol{\gamma}'\widehat{\boldsymbol{\Lambda}}_{li}\boldsymbol{\gamma}$. Following Gao et al. (2008), we can approximate $\mathbf{T}$ by a multivariate $t$-distribution, $t(\nu, \mathbf{0}, \widehat{\mathbf{R}})$ with $\nu = \max\{1, \min\{\nu_1, \ldots, \nu_q\}\}$

degrees of freedom, where:

$$\nu_l = \left( \sum_{i=1}^{a} \widehat{\eta}_{li}^2 / n_i \right)^2 \Big/ \sum_{i=1}^{a} \frac{\widehat{\eta}_{li}^4}{(n_i^2(n_i - 1))}, \; l = 1, \ldots, q. \tag{3.26}$$

This approximation will be especially useful for small sample sizes. How good it performs, in comparison to the MCTP using equi-coordinate quantiles from the multivariate normal distribution can be obtained from the simulation study in Section 4.

## 3.6 Another Approach

In Section 2.6 we shortly mentioned that estimators for the adjusted relative treatment effects, $\widehat{p}_i^*$, which can be interpreted as relative treatment effects would be desirable. We will now give some thoughts on a procedure in which such an estimator is seemingly attained, but which has not been proven to be correct. The reason why we present it here and not only as an outlook for future works is, that it will be part of the simulation study and therefore should be presented before the simulation results are discussed.

The basic idea is to undertake the data a two-step procedure. In a first step, the regression parameters $\gamma^{(r)}$, $r = 1, \ldots, d$, are estimated and the data is adjusted by the covariates within the rank transformation. For this, we take a look at our underlying model (2.8) and take it from there.

The underlying model gave us a presumed connection between the dependent variable and the covariates in the asymptotic rank transformation by:

$$Y_{ik}^{(0)} = \sum_{r=1}^{d} \gamma^{(r)} \cdot Y_{ik}^{(r)} + Y_{ik}^{reg}. \tag{3.27}$$

From here, differences between the factor levels $i = 1, \ldots, a$ were tested for, by testing for differences between the adjusted relative treatment effects. But instead, we could use the adjusted terms $Y_{ik}^{reg}$ to estimate the relative treatment effects of the factor levels. Hopes are, that the resulting estimates for the relative treatment effects correspond to the relative treatment effects of the dependent variable, and are still useful for testing hypotheses. For this, we define the adjusted random variables by:

$$\widehat{Y}_{ik}^{reg} = \widehat{Y}_{ik}^{(0)} - \sum_{r=1}^{d} \widehat{\gamma}^{(r)} \widehat{Y}_{ik}^{(r)}. \tag{3.28}$$

In a second step, the multiple contrast test procedures are performed on the data $\widehat{Y}_{ik}^{reg}$, $i = 1, \ldots, a$, $k = 1, \ldots, n_i$, as described in Konietschke et al. (2012a), i.e. as if no covariates were involved. Under circumstances yet to be studied, the resulting estimates for the adjusted relative treatment effects can be seen as estimates for the relative treatment effects of the dependent variable, and test results using the these estimators remain valid. A similar approach was also discussed by Siemer (1999) p.25, who does not follow it any further, arguing that it is unprovable. Nevertheless, we will take a look as to how good this procedure works, to validate if it is even worth following.

## 3.7 Multiple Contrast Test Procedures with Unknown Regression Parameters

Until now, we assumed that $\boldsymbol{\gamma}$ (2.11), the vector of regression parameters, is known. In practice however, the parameter is unknown and has to be estimated. We have shown a possibility of estimating $\boldsymbol{\gamma}$ in Section 2.5. Further, in Theorem (2.5.1), we have seen that:

$$\widehat{\boldsymbol{\gamma}} \xrightarrow{P} \boldsymbol{\gamma}$$

Since $\boldsymbol{\gamma}$ is an unknown but fixed parameter, using Theorem (2.5.1) and Slutsky's theorem, the results presented, especially the main results Theorem (3.5.2), Theorem (3.5.3) and Theorem (3.5.7), remain valid. That is, the MCTP developed in Section 3.5 is still asymptotically correct when plugging in $\widehat{\boldsymbol{\gamma}}$ for $\boldsymbol{\gamma}$.

# 4 Simulation Study

We will now compare the presented procedures, in terms of power $1-\beta$ and $\alpha$-level, in an extensive simulation study. These settings include altering group numbers ($a = 3, 5$), number of covariates ($d = 1, 2, 5$), changing sample sizes ($n_i = 7, 10, 15, 25, 50$) and different distributions of the given data. Further we will also simulate an unbalanced design and compare four different contrast matrices, the Tukey (*all-pairs*), the Dunnett (*many-to-one*), the centering matrix and the changepoint contrast matrix. Each data setting presented was simulated 10,000 times. We will commence by giving an overview over the different contrast matrices, before presenting the simulation results. For a more convenient readability, we will only present representative simulation results, while further results can be attained from the appendix in Section 7.2.

## 4.1 Contrast Matrices

As mentioned before in Definition (2.3.1), a contrast matrix $\mathbf{C} \neq \mathbf{0} \in \mathbb{R}^{q \times a}$ is a matrix, which row sums are zero, i.e. $\mathbf{C}\mathbf{1}_a = \mathbf{0}$. Although there exist infinitely many possibilities for contrast matrices, some are of higher practical interest than others. Four contrast matrices of high practical interest will be presented here, for more information on contrast matrices we refer to Bretz et al. (2001). The first presented contrast matrix is the Tukey contrast matrix Tukey (1953), given by:

$$\mathbf{C}_{\text{Tukey}} = \begin{pmatrix} -1 & 1 & 0 & \dots & \dots & 0 & 0 \\ -1 & 0 & 1 & 0 & \dots & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ -1 & 0 & 0 & 0 & \dots & 0 & 1 \\ 0 & -1 & 1 & 0 & 0 & \dots & 0 \\ 0 & -1 & 0 & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \dots & \dots & \dots & 0 & -1 & 1 \end{pmatrix}.$$

The Tukey contrast matrix compares all factor levels to each other, and is therefore referred to as the *all-pairs* contrast. The Dunnett contrast matrix Dunnett (1955) is given by:

$$\mathbf{C}_{\text{Dunnett}} = \begin{pmatrix} -1 & 1 & 0 & \dots & \dots & 0 \\ -1 & 0 & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ -1 & 0 & \dots & \dots & 0 & 1 \end{pmatrix}.$$

The Dunnett contrast matrix compares all factor levels to one specific factor level, e.g. the first factor level. It is therefore referred to as the *many-to-one* contrast. The centering matrix is given by:

$$\mathbf{P}_a = \begin{pmatrix} 1-\frac{1}{a} & -\frac{1}{a} & -\frac{1}{a} & \dots & -\frac{1}{a} \\ -\frac{1}{a} & 1-\frac{1}{a} & -\frac{1}{a} & \dots & -\frac{1}{a} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ -\frac{1}{a} & \dots & -\frac{1}{a} & -\frac{1}{a} & 1-\frac{1}{a} \end{pmatrix}.$$

The centering matrix also is a contrast matrix, which compares the relative treatment effect of every factor two the mean of all relative treatment effects. The centering matrix is also

used for testing global hypotheses in the ANOVA and ANCOVA models, and therefore of practical interest. A more complex contrast matrix is the changepoint contrast matrix, or changepoint comparisons, given by:

$$\mathbf{C}_{\text{Changepoint}} = \begin{pmatrix} -1 & \frac{n_2}{N_{2a}} & \cdots & \cdots & \frac{n_{a-1}}{N_{2a}} & \frac{n_a}{N_{2a}} \\ -\frac{n_1}{N_{12}} & -\frac{n_2}{N_{12}} & \frac{n_3}{N_{3a}} & \cdots & \frac{n_{a-1}}{N_{3a}} & \frac{n_a}{N_{3a}} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ -\frac{n_1}{N_{1(a-1)}} & -\frac{n_2}{N_{1(a-1)}} & \cdots & \cdots & -\frac{n_{a-1}}{N_{1(a-1)}} & 1 \end{pmatrix}.$$

where $N_{ij} = \sum_{k=i}^{j} n_k$. The changepoint contrast matrix compares the relative treatment effects in such a way, that change points of the relative treatment effects are uncovered. While these contrasts were used for our simulations, we will only present the results attained using the Tukey or *all-pairs* contrast matrix, the reason being, that the other simulated contrasts behaved very similarly. Therefore, all simulation results are attained using a Tukey contrast matrix, unless stated otherwise. Nevertheless, we will compare the performance of the test procedures for different contrasts for one selected data setting in Section 4.2.5.

## 4.2 Type I Error Simulation

Before taking a look at the power $1 - \beta$ of the test procedures, we will first simulate the $\alpha$-level, or type I error of the test procedures. Table 2 gives a comprehensive overview of the power $1 - \beta$ and the $\alpha$-level.

*Table 2:* Type I and type II errors of statistical inference.

| | $H_0$ is not rejected | $H_0$ is rejected |
|---|---|---|
| $H_0$ is true | $1 - \alpha$ | $\alpha$ (Type I error, level) |
| $H_0$ is false | $\beta$ (Type II error) | $1 - \beta$ (Power) |

The $\alpha$-level, or type I error, refers to the probability of a test procedure rejecting the null hypothesis, although the null hypothesis is true. The power on the other hand, refers to the test procedure rejecting the null hypothesis under alternative. While the nature of statistical inference makes it inevitable for type I errors to occur, the type I error can be freely chosen when testing, provided the corresponding test behaves accordingly. In other words, if the test user decided to test with a certain $\alpha$-level, the test should hold this level. The type II error on the other hand is harder to control and should tend to 0 as differences in the factor levels of the data increase. It is desirable for a test to show a steep increase in power as differences in factor levels increase.

For the upcoming simulation results, an $\alpha$-level of 0.05 was chosen. It is therefore preferable for the compared procedures to have a type I error of exactly 0.05. Should a test show a type I error less than $\alpha$, it is referred to as being *conservative*, because it tends to under reject the corresponding hypothesis. If a test shows a type I error higher than $\alpha$, it is referred to as being *liberal*, because it tends to over reject the corresponding hypothesis.

### 4.2.1 Multivariate Normal Distribution

In the first part of the simulation, the covariates are simulated from a multivariate normal distribution,

$$(X_{ik}^{(1)}, \ldots, X_{ik}^{(d)}) \overset{i.i.d.}{\sim} N(\mathbf{0}, \mathbf{\Psi}),\ i = 1, \ldots, a,\ k = 1, \ldots, n_i, \tag{4.1}$$

where $\mathbf{\Psi}$ denotes a covariance matrix of compound symmetry structure with parameter $\rho$:

$$\mathbf{\Psi} = \rho \cdot \mathbf{1}_d \mathbf{1}_d' + (1 - \rho) \cdot \mathbf{I}_d.$$

The covariates were simulated to be independent, slightly dependent and strongly dependent through $\rho = 0$, $\rho = 0.5$ and $\rho = 0.9$, respectively. In the first simulation result presented, the dependent variable follows a standard normal distribution and is not influenced by the covariates, i.e.:

$$X_{ik}^{(0)} \overset{i.i.d.}{\sim} N(0,1),\ i = 1, \ldots, a,\ k = 1, \ldots, n_i. \tag{4.2}$$

Taking into account covariates in this setting cannot improve the test and will, in the worst case, lead to erroneous results. Table 3 shows the type I error rates of the compared procedures.

*Table 3:* An $\alpha$-level simulation. Design: balanced with $n_i = 7, 10, 15, 25, 50$ observations. Factor levels: $a = 3$. Number of covariates: $d = 1, 2, 5$. Covariates: multivariate normal distributed (4.1), $\rho = 0$. Dependent variable: standard normal distributed (4.2).

| $d$ | $n_i$ | mctp.n | mctp.t | 2-step | sie.ats | sie.wts | mctp.nc | mctp.par | f.par |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 7 | 0.1389 | 0.0824 | 0.0662 | 0.0913 | 0.1458 | 0.0680 | 0.0515 | 0.0502 |
| | 10 | 0.1051 | 0.0723 | 0.0576 | 0.0788 | 0.1100 | 0.0641 | 0.0491 | 0.0506 |
| | 15 | 0.0856 | 0.0652 | 0.0544 | 0.0682 | 0.0886 | 0.0587 | 0.0519 | 0.0519 |
| | 25 | 0.0689 | 0.0560 | 0.0505 | 0.0579 | 0.0705 | 0.0525 | 0.0458 | 0.0469 |
| | 50 | 0.0599 | 0.0557 | 0.0515 | 0.0573 | 0.0623 | 0.0544 | 0.0520 | 0.0537 |
| 2 | 7 | 0.1575 | 0.1001 | 0.0905 | 0.1166 | 0.1656 | 0.0721 | 0.0502 | 0.0509 |
| | 10 | 0.1204 | 0.0816 | 0.0682 | 0.0923 | 0.1250 | 0.0632 | 0.0479 | 0.0490 |
| | 15 | 0.0935 | 0.0712 | 0.0632 | 0.0762 | 0.0991 | 0.0588 | 0.0498 | 0.0494 |
| | 25 | 0.0783 | 0.0687 | 0.0597 | 0.0699 | 0.0812 | 0.0594 | 0.0539 | 0.0533 |
| | 50 | 0.0635 | 0.0571 | 0.0532 | 0.0571 | 0.0629 | 0.0539 | 0.0509 | 0.0500 |
| 5 | 7 | 0.2613 | 0.1804 | 0.1890 | 0.2246 | 0.2699 | 0.0735 | 0.0564 | 0.0570 |
| | 10 | 0.1752 | 0.1272 | 0.1195 | 0.1459 | 0.1776 | 0.0639 | 0.0485 | 0.0508 |
| | 15 | 0.1272 | 0.1029 | 0.0928 | 0.1113 | 0.1327 | 0.0649 | 0.0524 | 0.0534 |
| | 25 | 0.0872 | 0.0734 | 0.0675 | 0.0800 | 0.0890 | 0.0532 | 0.0484 | 0.0480 |
| | 50 | 0.0677 | 0.0620 | 0.0584 | 0.0630 | 0.0684 | 0.0512 | 0.0490 | 0.0486 |

The multiple contrast test procedure using equi-coordinate quantiles from the multivariate normal distribution is denoted by *mctp.n* (Section 3.5), the multiple contrast test procedure using equi-coordinate quantiles from the multivariate t-distribution is denoted by *mctp.t* (Section 3.5.3), the 2-step procedure, which remains unproven, is denoted as *2-step* (Section 3.6), the ANOVA-type statistic (Section 3.4) from Siemer (1999) and the Wald-type statistic (Section 3.3) are denoted by *sie.ats* and *sie.wts* respectively. Furthermore, the multiple contrast test procedure not using covariates from Konietschke et al. (2012a), thus only using the dependent variable for statistical inference, is denoted by *mctp.nc*, the parametric multiple contrast test procedure is denoted by *mctp.par* (Section 3.2) and the parametric f-test is denoted by *f.par* (Section 3.1). Note that the procedure not considering covariates presented here is the procedure using the multivariate t-approximation from Konietschke et al. (2012a). While better procedures have been proposed within the same paper, this procedure is best compared to the mctp.t derived here.

All procedures show asymptotic properties, performing better for high sample sizes ($n_i = 50$) than for low sample sizes ($n_i = 7$). Of the non-parametric approaches, aside from the 2-step procedure, the mctp.t performs best in all displayed settings. However, the mctp.t is very liberal, especially for a large number of covariates and few observations ($d = 5,\ n_i = 7$). The type I error is as large as 0.1804, far higher than the targeted 0.05 and unacceptable for a good statistical inference. For a high number of observations and few covariates ($d = 1, n_i = 50$), the type I error improves to a satisfactory level of 0.0557.

The reason for such a high type I error, especially considering that not using the covariates would result in a drastically improved type I error (mctp.nc), could lie in the estimation of the regression parameters $\gamma^{(1)}, \ldots, \gamma^{(d)},\ d = 1, \ldots, r$. In this setting, an optimal estimation would be $\gamma^{(r)} = 0$ for all $r = 1, \ldots, d$, as it is in the case of mctp.nc. While the estimation in a parametric setting is adequate, in a non-parametric setting the number of combinations in the rank transformation is limited, especially for few observations, and the regression parameters are estimated erroneously, causing the procedure to detect a false coherence between the dependent variable and the covariates, more frequently than in the parametric setting.

This simulation shows that adding covariates to the model in hopes of attaining significant results will work, but yields erroneous results. To attain credible results, the sample size needs to be accordingly high when planning to use covariates. Only then will falsely chosen covariates be ignored by the non-parametric procedures.

In the second simulation result presented, covariates are attained from a multivariate normal distribution as in (4.1). The dependent variable is chosen to be the unweighted sum of the covariates, plus an additional independent random variable, i.e.:

$$X_{ik}^{(0)} = \sum_{r=1}^{d} X_{ik}^{(r)} + U_{ik} \text{ where } U_{ik} \overset{i.i.d.}{\sim} N(0,1),\ i = 1,\dots,a,\ k = 1,\dots,n_i. \tag{4.3}$$

Thus, in this case the dependent variable is influenced by the covariates and taking into account the covariates should yield better results. Table 4 shows the results for $\rho = 0$, i.e. for independent covariates.

*Table 4:* An $\alpha$-level simulation. Design: balanced with $n_i = 7, 10, 15, 25, 50$ observations. Factor levels: $a = 3$. Number of covariates: $d = 1, 2, 5$. Covariates: multivariate normal distributed (4.1), $\rho = 0$. Dependent variable: unweighted sum of covariates (4.3).

| $d$ | $n_i$ | mctp.n | mctp.t | 2-step | sie.ats | sie.wts | mctp.nc | mctp.par | f.par |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 7 | 0.1024 | 0.0596 | 0.0694 | 0.0734 | 0.1078 | 0.0705 | 0.0506 | 0.0503 |
| | 10 | 0.0778 | 0.0501 | 0.0570 | 0.0620 | 0.0810 | 0.0611 | 0.0485 | 0.0483 |
| | 15 | 0.0710 | 0.0531 | 0.0606 | 0.0611 | 0.0728 | 0.0599 | 0.0518 | 0.0516 |
| | 25 | 0.0606 | 0.0514 | 0.0503 | 0.0521 | 0.0607 | 0.0520 | 0.0472 | 0.0487 |
| | 50 | 0.0532 | 0.0486 | 0.0495 | 0.0512 | 0.0550 | 0.0534 | 0.0506 | 0.0517 |
| 2 | 7 | 0.1125 | 0.0645 | 0.0940 | 0.0870 | 0.1170 | 0.0715 | 0.0581 | 0.0580 |
| | 10 | 0.0851 | 0.0555 | 0.0745 | 0.0696 | 0.0887 | 0.0675 | 0.0528 | 0.0522 |
| | 15 | 0.0680 | 0.0493 | 0.0617 | 0.0569 | 0.0707 | 0.0586 | 0.0466 | 0.0467 |
| | 25 | 0.0627 | 0.0523 | 0.0587 | 0.0588 | 0.0638 | 0.0542 | 0.0532 | 0.0536 |
| | 50 | 0.0526 | 0.0475 | 0.0508 | 0.0491 | 0.0532 | 0.0513 | 0.0449 | 0.0449 |
| 5 | 7 | 0.1548 | 0.0935 | 0.1849 | 0.1290 | 0.1609 | 0.0714 | 0.0516 | 0.0518 |
| | 10 | 0.1108 | 0.0754 | 0.1248 | 0.0964 | 0.1142 | 0.0618 | 0.0513 | 0.0544 |
| | 15 | 0.0842 | 0.0641 | 0.0948 | 0.0736 | 0.0860 | 0.0586 | 0.0516 | 0.0527 |
| | 25 | 0.0701 | 0.0584 | 0.0752 | 0.0662 | 0.0705 | 0.0537 | 0.0489 | 0.0496 |
| | 50 | 0.0547 | 0.0500 | 0.0550 | 0.0527 | 0.0545 | 0.0504 | 0.0481 | 0.0490 |

In this setting, the mctp.t again performs better than the other non-parametric alternatives in almost all parameter combinations displayed. The procedure still performs poorly for a high number of covariates and a low number of observations, but not as dramatically as in Table 3. Out of a methodological point of view, taking into account covariates now would be the correct decision and, as long as the ratio between the number of covariates and sample sizes remains healthy, the $\alpha$-level of 0.05 is met more often when considering covariates, opposed to not considering covariates.

In a further simulation study, the covariates were generated in the same way as described in (4.1), with the only difference being, that the dependent variable was chosen to be a weighted sum of the covariates, i.e.:

$$X_{ik}^{(0)} = \sum_{r=1}^{d} \frac{r}{2} \cdot X_{ik}^{(r)} + U_{ik} \text{ where } U_{ik} \overset{i.i.d.}{\sim} N(0,1),\ i = 1, \ldots, a,\ k = 1, \ldots, n_i. \tag{4.4}$$

The corresponding simulation results can be attained from Table 5.

*Table 5:* An $\alpha$-level simulation. Design: balanced with $n_i = 7, 10, 15, 25, 50$ observations. Factor levels: $a = 3$. Number of covariates: $d = 1, 2, 5$. Covariates: multivariate normal distributed (4.1), $\rho = 0$. Dependent variable: weighted sum of covariates (4.4).

| $d$ | $n_i$ | mctp.n | mctp.t | 2-step | sie.ats | sie.wts | mctp.nc | mctp.par | f.par |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 7 | 0.1294 | 0.0765 | 0.0724 | 0.0913 | 0.1368 | 0.0733 | 0.0512 | 0.0521 |
| | 10 | 0.0967 | 0.0663 | 0.0588 | 0.0737 | 0.1035 | 0.0624 | 0.0483 | 0.0496 |
| | 15 | 0.0791 | 0.0594 | 0.0552 | 0.0642 | 0.0822 | 0.0572 | 0.0500 | 0.0509 |
| | 25 | 0.0614 | 0.0502 | 0.0461 | 0.0540 | 0.0620 | 0.0498 | 0.0457 | 0.0455 |
| | 50 | 0.0584 | 0.0523 | 0.0512 | 0.0547 | 0.0588 | 0.0532 | 0.0496 | 0.0511 |
| 2 | 7 | 0.1280 | 0.0706 | 0.0992 | 0.0982 | 0.1341 | 0.0716 | 0.0515 | 0.0520 |
| | 10 | 0.0885 | 0.0579 | 0.0714 | 0.0688 | 0.0933 | 0.0606 | 0.0487 | 0.0487 |
| | 15 | 0.0744 | 0.0563 | 0.0617 | 0.0640 | 0.0772 | 0.0629 | 0.0487 | 0.0490 |
| | 25 | 0.0657 | 0.0543 | 0.0591 | 0.0603 | 0.0666 | 0.0540 | 0.0478 | 0.0491 |
| | 50 | 0.0567 | 0.0526 | 0.0541 | 0.0548 | 0.0579 | 0.0558 | 0.0530 | 0.0530 |
| 5 | 7 | 0.1221 | 0.0714 | 0.1888 | 0.0947 | 0.1278 | 0.0686 | 0.0516 | 0.0513 |
| | 10 | 0.0808 | 0.0531 | 0.1233 | 0.0709 | 0.0864 | 0.0638 | 0.0514 | 0.0526 |
| | 15 | 0.0642 | 0.0459 | 0.0925 | 0.0555 | 0.0665 | 0.0594 | 0.0511 | 0.0508 |
| | 25 | 0.0554 | 0.0466 | 0.0732 | 0.0521 | 0.0560 | 0.0556 | 0.0498 | 0.0503 |
| | 50 | 0.0498 | 0.0457 | 0.0539 | 0.0471 | 0.0494 | 0.0521 | 0.0479 | 0.0485 |

The mctp.t again performs very well compared to the other non-parametric procedures, but not as well as in the case of unweighted covariates (Table 4). Not taking the covariates into account partially promises even better results in terms of $\alpha$-level. It becomes clear though, that the 2-step procedure seems to perform good for a low number of covariates and devastating for small sample sizes and a large number of covariates. As this procedure is unproven, we leave it to the reader to decide on further interpretations. Supplementary results can be attained from the appendix in Section 7.2.1.

### 4.2.2 Multivariate Log-Normal Distribution

After having tested the procedures in different settings involving the multivariate normal distribution, we will now regard simulation results involving the log-normal distribution. For this, the covariates were generated to be multivariate normally distributed as described in (4.1). Then, the components of the vector were transformed to:

$$(exp(X_{ik}^{(1)}), \ldots, exp(X_{ik}^{(d)})),\ i = 1, \ldots, a,\ k = 1, \ldots, n_i, \tag{4.5}$$

where *exp* denotes the exponential function. Thus, the covariates now follow a multivariate log-normal distribution with parameters $\mathbf{0}$ and covariance matrix $\mathbf{\Gamma}$ given by:

$$\mathbf{\Gamma} = (e^1(e^{\Psi_{ij}} - 1))_{i=1,\ldots,d, j=1,\ldots,d}.$$

For the following simulation result, the dependent variable was chosen to be an unweighted sum of the covariates, with a log-normal error term, i.e.:

$$X_{ik}^{(0)} = \sum_{r=1}^{d} exp(X_{ik}^{(r)}) + U_{ik} \text{ where } U_{ik} \overset{i.i.d.}{\sim} logN(0,1),\ i = 1,\dots,a,\ k = 1,\dots,n_i, \quad (4.6)$$

where $logN(0,1)$ denotes the distribution function of $exp(X)$, where $X$ denotes a standard normal random variable. Under these circumstances, neither the dependent variable, nor the covariates are normally distributed. Therefore, the assumptions for using the parametric approaches to considering covariates are no longer fulfilled. Nevertheless, we will compare these procedures to see how well they perform compared to the non-parametric approaches. Table 6 shows the results for this simulation.

*Table 6:* An $\alpha$-level simulation. Design: balanced with $n_i = 7, 10, 15, 25, 50$ observations. Factor levels: $a = 3$. Number of covariates: $d = 1, 2, 5$. Covariates: multivariate log-normal distributed (4.5), $\rho = 0$. Dependent variable: unweighted sum of covariates (4.6).

| $d$ | $n_i$ | mctp.n | mctp.t | 2-step | sie.ats | sie.wts | mctp.nc | mctp.par | f.par |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 7 | 0.1025 | 0.0550 | 0.0689 | 0.0739 | 0.1096 | 0.0716 | 0.0348 | 0.0369 |
| | 10 | 0.0794 | 0.0507 | 0.0572 | 0.0611 | 0.0858 | 0.0627 | 0.0321 | 0.0343 |
| | 15 | 0.0723 | 0.0520 | 0.0566 | 0.0605 | 0.0778 | 0.0632 | 0.0369 | 0.0402 |
| | 25 | 0.0591 | 0.0479 | 0.0502 | 0.0516 | 0.0604 | 0.0554 | 0.0369 | 0.0381 |
| | 50 | 0.0543 | 0.0498 | 0.0505 | 0.0526 | 0.0556 | 0.0518 | 0.0428 | 0.0428 |
| 2 | 7 | 0.1204 | 0.0675 | 0.0962 | 0.0912 | 0.1277 | 0.0790 | 0.0356 | 0.0371 |
| | 10 | 0.0920 | 0.0546 | 0.0718 | 0.0707 | 0.0936 | 0.0587 | 0.0348 | 0.0370 |
| | 15 | 0.0736 | 0.0548 | 0.0643 | 0.0614 | 0.0769 | 0.0594 | 0.0364 | 0.0378 |
| | 25 | 0.0636 | 0.0532 | 0.0578 | 0.0568 | 0.0685 | 0.0549 | 0.0379 | 0.0389 |
| | 50 | 0.0537 | 0.0485 | 0.0531 | 0.0517 | 0.0555 | 0.0531 | 0.0419 | 0.0431 |
| 5 | 7 | 0.1975 | 0.1256 | 0.1856 | 0.1671 | 0.2055 | 0.0663 | 0.0398 | 0.0407 |
| | 10 | 0.1419 | 0.0972 | 0.1275 | 0.1194 | 0.1461 | 0.0624 | 0.0388 | 0.0386 |
| | 15 | 0.1048 | 0.0797 | 0.0929 | 0.0922 | 0.1088 | 0.0616 | 0.0386 | 0.0391 |
| | 25 | 0.0785 | 0.0661 | 0.0727 | 0.0713 | 0.0794 | 0.0573 | 0.0359 | 0.0372 |
| | 50 | 0.0586 | 0.0519 | 0.0543 | 0.0553 | 0.0597 | 0.0514 | 0.0397 | 0.0398 |

As in the simulation results already presented, the mctp.t performs very well. Only for low sample sizes and a high number of covariates the procedure becomes very liberal, still performing better than the non-parametric alternatives. The parametric approaches now turn out to be slightly conservative, which might be because of the underlying skewness in the distribution of the data.

### 4.2.3 Binomial Distribution

One simulation study was conducted using covariates which follow a binomial distribution with parameters $m = 4$ and $q = 0.5$. The dependent variable was chosen to be an unweighted sum of the covariates plus an additional standard normal distributed error term. In other words, the covariates were chosen to be:

$$X_{ik}^{(r)} \overset{i.i.d.}{\sim} Bin(m,q) \text{ for } i = 1,\dots,a,\ k = 1,\dots,n_i,\ r = 1,\dots,d, \quad (4.7)$$

and the dependent variable was calculated as:

$$X_{ik}^{(0)} = \sum_{r=1}^{d} X_{ik}^{(r)} + U_{ik}, \text{ where } U_{ik} \overset{i.i.d.}{\sim} N(0,1),\ i = 1,\dots,a,\ k = 1,\dots,n_i. \tag{4.8}$$

The simulation results are given in Table 7.

*Table 7:* An $\alpha$-level simulation. Design: balanced with $n_i = 7, 10, 15, 25, 50$ observations. Factor levels: $a = 3$. Number of covariates: $d = 1, 2, 5$. Covariates: binomial distributed (4.7). Dependent variable: unweighted sum of covariates (4.8).

| $d$ | $n_i$ | mctp.n | mctp.t | 2-step | sie.ats | sie.wts | mctp.nc | mctp.par | f.par |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 7 | 0.1106 | 0.0606 | 0.0730 | 0.0806 | 0.1180 | 0.0765 | 0.0504 | 0.0501 |
| | 10 | 0.0832 | 0.0541 | 0.0587 | 0.0665 | 0.0858 | 0.0648 | 0.0479 | 0.0486 |
| | 15 | 0.0680 | 0.0491 | 0.0520 | 0.0572 | 0.0703 | 0.0589 | 0.0475 | 0.0457 |
| | 25 | 0.0602 | 0.0504 | 0.0518 | 0.0553 | 0.0638 | 0.0551 | 0.0529 | 0.0517 |
| | 50 | 0.0578 | 0.0531 | 0.0518 | 0.0546 | 0.0576 | 0.0566 | 0.0518 | 0.0503 |
| 2 | 7 | 0.1163 | 0.0646 | 0.0900 | 0.0912 | 0.1202 | 0.0712 | 0.0511 | 0.0510 |
| | 10 | 0.0906 | 0.0601 | 0.0760 | 0.0733 | 0.0962 | 0.0694 | 0.0477 | 0.0484 |
| | 15 | 0.0758 | 0.0582 | 0.0674 | 0.0654 | 0.0772 | 0.0612 | 0.0519 | 0.0510 |
| | 25 | 0.0666 | 0.0539 | 0.0589 | 0.0584 | 0.0663 | 0.0545 | 0.0524 | 0.0511 |
| | 50 | 0.0576 | 0.0516 | 0.0546 | 0.0555 | 0.0574 | 0.0542 | 0.0519 | 0.0513 |
| 5 | 7 | 0.1715 | 0.1046 | 0.1916 | 0.1407 | 0.1762 | 0.0693 | 0.0475 | 0.0494 |
| | 10 | 0.1181 | 0.0801 | 0.1261 | 0.1053 | 0.1242 | 0.0637 | 0.0525 | 0.0507 |
| | 15 | 0.0834 | 0.0632 | 0.0864 | 0.0757 | 0.0862 | 0.0544 | 0.0493 | 0.0506 |
| | 25 | 0.0730 | 0.0609 | 0.0725 | 0.0660 | 0.0723 | 0.0593 | 0.0505 | 0.0505 |
| | 50 | 0.0609 | 0.0563 | 0.0580 | 0.0566 | 0.0602 | 0.0524 | 0.0504 | 0.0501 |

As seen in the previous simulation results, the mctp.t performs very well in comparison to the other non-parametric results, but still very liberal for a high number of covariates and a low sample size. Given the standard normal error term, the parametric approaches perform excellent.

### 4.2.4 Further Distributions

Apart from the multivariate normal, the multivariate log-normal and the binomial distribution, simulation studies were also conducted for covariates from a Poisson, an exponential, a Bernoulli and a Chi-squared distribution. In these settings, the dependent variable was chosen to be an unweighted sum of the covariates, plus a standard normal error term. The results attained are very similar to those of the binomial distribution, and are therefore listed in the appendix, Section 7.2.1. Other simulation results for the multivariate normal and the multivariate log-normal distributions can also be attained from Section 7.2.1.

### 4.2.5 Comparing Contrast Matrices

As mentioned before, we would like to compare the contrast matrices, presented at the beginning of this section, in terms of $\alpha$-level. For comparing contrast matrices, we chose to use covariates generated from a multivariate normal distribution as in (4.1) with dependency

parameter $\rho = 0$ and the dependent variable being an unweighted sum of the covariates, plus an additional standard normal error term, as in (4.3). The results for $d = 2$ covariates can be attained from Table 8.

*Table 8:* An $\alpha$-level simulation. Design: balanced with $n_i = 7, 10, 15, 25, 50$ observations. Factor levels: $a = 3$. Number of covariates: $d = 2$. Covariates: multivariate normal distributed (4.1), $\rho = 0$. Dependent variable: unweighted sum of covariates (4.3). Contrast matrices (*Con*): Centering (*Cen*), Changepoint (*Cha*), Dunnett (*Dun*) and Tukey *(Tuk)*.

| $n_i$ | Con | mctp.n | mctp.t | 2-step | sie.ats | sie.wts | mctp.nc | mctp.par | f.par |
|---|---|---|---|---|---|---|---|---|---|
| 7 | Cen | 0.1023 | 0.0531 | 0.0984 | 0.0792 | 0.1111 | 0.0648 | 0.0494 | 0.0499 |
| | Cha | 0.0960 | 0.0543 | 0.0835 | 0.0803 | 0.1122 | 0.0621 | 0.0503 | 0.0503 |
| | Dun | 0.0928 | 0.0575 | 0.0817 | 0.0803 | 0.1096 | 0.0642 | 0.0500 | 0.0502 |
| | Tuk | 0.1125 | 0.0645 | 0.0940 | 0.0870 | 0.1170 | 0.0715 | 0.0581 | 0.0580 |
| 10 | Cen | 0.0858 | 0.0525 | 0.0805 | 0.0705 | 0.0888 | 0.0611 | 0.0506 | 0.0523 |
| | Cha | 0.0762 | 0.0503 | 0.0729 | 0.0678 | 0.0855 | 0.0567 | 0.0489 | 0.0509 |
| | Dun | 0.0771 | 0.0509 | 0.0654 | 0.0659 | 0.0878 | 0.0576 | 0.0505 | 0.0502 |
| | Tuk | 0.0851 | 0.0555 | 0.0745 | 0.0696 | 0.0887 | 0.0675 | 0.0528 | 0.0522 |
| 15 | Cen | 0.0719 | 0.0513 | 0.0709 | 0.0639 | 0.0731 | 0.0546 | 0.0506 | 0.0505 |
| | Cha | 0.0695 | 0.0531 | 0.0650 | 0.0622 | 0.0738 | 0.0570 | 0.0533 | 0.0525 |
| | Dun | 0.0672 | 0.0513 | 0.0610 | 0.0588 | 0.0717 | 0.0547 | 0.0515 | 0.0500 |
| | Tuk | 0.0680 | 0.0493 | 0.0617 | 0.0569 | 0.0707 | 0.0586 | 0.0466 | 0.0467 |
| 25 | Cen | 0.0616 | 0.0516 | 0.0586 | 0.0541 | 0.0630 | 0.0582 | 0.0497 | 0.0505 |
| | Cha | 0.0577 | 0.0495 | 0.0536 | 0.0542 | 0.0612 | 0.0535 | 0.0502 | 0.0503 |
| | Dun | 0.0574 | 0.0483 | 0.0537 | 0.0518 | 0.0567 | 0.0499 | 0.0488 | 0.0478 |
| | Tuk | 0.0627 | 0.0523 | 0.0587 | 0.0588 | 0.0638 | 0.0542 | 0.0532 | 0.0536 |
| 50 | Cen | 0.0504 | 0.0453 | 0.0496 | 0.0495 | 0.0495 | 0.0491 | 0.0463 | 0.0469 |
| | Cha | 0.0541 | 0.0505 | 0.0538 | 0.0548 | 0.0574 | 0.0508 | 0.0518 | 0.0542 |
| | Dun | 0.0512 | 0.0480 | 0.0519 | 0.0503 | 0.0544 | 0.0509 | 0.0481 | 0.0490 |
| | Tuk | 0.0526 | 0.0475 | 0.0508 | 0.0491 | 0.0532 | 0.0513 | 0.0449 | 0.0449 |

Taking a look at Table 8, it becomes clear that the type of contrast matrix does not make a big difference in terms of $\alpha$-level, independent from the procedure chosen. Further results can be attained from the appendix in Section 7.2.1.

### 4.2.6 Unbalanced Design

Although most experiments aim at having a balanced design, i.e. an equal number of observations per group, only few experiments actually meet this criterion in the end. Unequal group sizes can have many reasons like dropouts, subjects who do not return for an evaluation, or natural restrictions, like in gene datasets, where certain genes simply occur less frequently than others. In statistics, it is therefore very important for a test procedure to allow for unequal sample sizes between the experiment groups, otherwise the test procedure is highly restrictive and can only be applied very seldom.

The derived MCTP (mctp.n, mctp.t) allow for an unbalanced experiment design. In this part of the simulation study, we would like to examine how good the procedures perform in terms of $\alpha$-level when confronted with an unbalanced design. For this, we will examine two

different degrees of imbalance, a low degree of imbalance with $n_1 = 10$, $n_2 = 15$ and $n_3 = 20$ observations in the groups, and a high degree of imbalance with $n_1 = 10$, $n_2 = 50$ and $n_3 = 100$ observations. Further, we examined three different distributions of the data. Here we only present the results from the multivariate normal distribution, with covariates generated as in (4.1) with dependency parameter $\rho = 0$ and the dependent variable generated as in (4.3), an unweighted sum of the covariates with standard normal error term, which results are given in Table 9. The results for the other distributions, namely the log-normal and the binomial distribution, can be attained from the appendix in Section 7.2.1.

*Table 9:* An $\alpha$-level simulation. Design: unbalanced with $n_1 = 10$, $n_2 = 15$, $n_3 = 20$ (*low*) and $n_1 = 10$, $n_2 = 50$, $n_3 = 100$ (*high*) observations. Factor levels: $a = 3$. Number of covariates: $d = 1, 2, 5$. Covariates: multivariate normal distributed (4.1), $\rho = 0$. Dependent variable: unweighted sum of covariates (4.3). The degree of imbalance is abbreviated by *Imba*.

| $d$ | Imba | mctp.n | mctp.t | 2-step | sie.ats | sie.wts | mctp.nc | mctp.par | f.par |
|---|---|---|---|---|---|---|---|---|---|
| 1 | low | 0.0733 | 0.0518 | 0.0528 | 0.0624 | 0.0749 | 0.0499 | 0.0576 | 0.0505 |
| | high | 0.0731 | 0.0623 | 0.0423 | 0.0723 | 0.0744 | 0.0478 | 0.0479 | 0.0490 |
| 2 | low | 0.0792 | 0.0547 | 0.0640 | 0.0672 | 0.0802 | 0.0528 | 0.0608 | 0.0531 |
| | high | 0.0776 | 0.0633 | 0.0465 | 0.0713 | 0.0764 | 0.0490 | 0.0470 | 0.0510 |
| 5 | low | 0.0883 | 0.0662 | 0.0872 | 0.0771 | 0.0906 | 0.0495 | 0.0548 | 0.0492 |
| | high | 0.0734 | 0.0625 | 0.0501 | 0.0735 | 0.0732 | 0.0535 | 0.0481 | 0.0502 |

Comparing the mctp.t in the unbalanced design with the prior balanced designs, the unbalanced observation numbers do not seem to have a strong effect. It seems though, that an increasing imbalance between the factor levels results in slightly liberal test decisions. This is not true, when taking into account 5 covariates, where the estimation of the regression parameters seems to profit more from the higher observation numbers, than the imbalance does damage. As in the prior simulation results, the mctp.t performs better than the non-parametric alternatives, but not as good as the parametric procedures. The mctp.par however, seems to be more conservative when confronted with a higher imbalance between the factor levels.

### 4.2.7 Higher Number of Factor Levels

We shortly mentioned that we would also like to compare the performance of the test procedures when confronted with a higher number of factor levels. Now we would like to compare the test procedures when confronted with $a = 5$ factor levels. Due to the high numerical computation times, we will only present the results for one data setting. The covariates were generated to be multivariate normal distributed as in (4.1), with dependency parameter $\rho = 0$, and the dependent variable was generated to be an unweighted sum of covariates with an additional standard normal error term, as in (4.3). We simulated the case of $d = 2$ covariates. Table 10 yields the results of this simulation.

*Table 10:* An $\alpha$-level simulation. Design: balanced with $n_i = 7, 10, 15, 25, 50$ observations. Factor levels: $a = 5$. Number of covariates: $d = 2$. Covariates: multivariate normal distributed (4.1), $\rho = 0$. Dependent variable: unweighted sum of covariates (4.3).

| $n_i$ | mctp.n | mctp.t | 2-step | sie.ats | sie.wts | mctp.nc | mctp.par | f.par |
|---|---|---|---|---|---|---|---|---|
| 7 | 0.1309 | 0.0510 | 0.0857 | 0.0711 | 0.1397 | 0.0852 | 0.0495 | 0.0487 |
| 10 | 0.1024 | 0.0514 | 0.0651 | 0.0629 | 0.1101 | 0.0716 | 0.0517 | 0.0517 |
| 15 | 0.0798 | 0.0514 | 0.0570 | 0.0553 | 0.0828 | 0.0657 | 0.0501 | 0.0495 |
| 25 | 0.0660 | 0.0499 | 0.0525 | 0.0505 | 0.0652 | 0.0599 | 0.0479 | 0.0510 |
| 50 | 0.0585 | 0.0522 | 0.0504 | 0.0509 | 0.0569 | 0.0530 | 0.0480 | 0.0478 |

Comparing the results with the results attained when considering $a = 3$ factor levels, see Table 4, the mctp.t does not perform worse. In fact, in performs better, which could be because the additional observations lead to a better estimation of the regression parameters $\gamma^{(1)}, \ldots, \gamma^{(d)}$. The sie.ats procedure also performs slightly better, while the sie.wts procedure, the mctp.n and mctp.nc perform worse.

## 4.3 Power Simulation

We will now continue with comparing the power of the simulated procedures. As mentioned in Section 4.2, it is desirable for a procedure to maintain the $\alpha$-level as long as the null hypothesis is fulfilled, and then show a steep increase in power under the alternative. Under the alternative, the best procedure in terms of power is the procedure which power is closest to one.

### 4.3.1 Multivariate Normal Distribution

In a similar manner as for the type I error simulation, we simulated data from the multivariate normal distribution. Analogously to (4.1), the covariates were generated to be multivariate normal distributed, slightly dependent with parameter $\rho = 0.2$. The dependent variable was chosen do be the sum of covariates, with a standard normal distributed error term and a shift term $\delta$ on each factor level, i.e.:

$$X_{ik}^{(0)} = \sum_{r=1}^{d} X_{ik}^{(r)} + U_{ik} + \delta \cdot (i-1), \; i = 1, \ldots, a, \; k = 1, \ldots, n_i, \tag{4.9}$$

where the shift term $\delta$ varied from 0 to 1 in steps of 0.2. All procedures should increase in power, as $\delta$ increases, and thus the differences between the factor levels of the dependent variable increase. Figure 4 shows selected results from the power simulation. For reasons of readability, we have chosen to compare only 4 procedures, the newly developed multiple contrast test procedure considering covariates, *mctp.t*, the current procedure for coping with covariates in a non-parametric setting, *sie.ats*, the parametric alternative to multiple contrast test procedures, *mctp.par*, and finally the case of ignoring covariates, *mctp.nc*. Further comparisons can be attained from the appendix in Section 7.2.2.

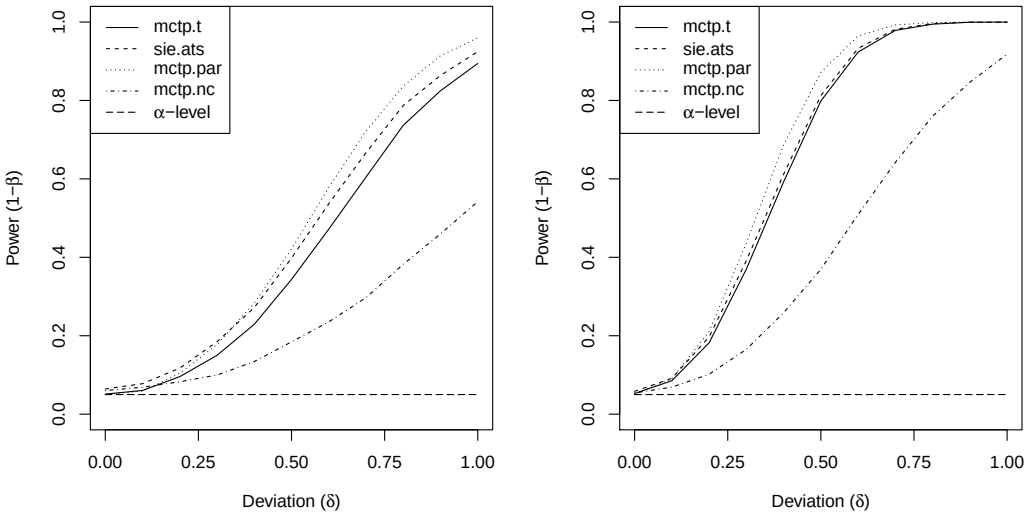


*Figure 4:* Simulation results for the power $(1-\beta)$. Design: balanced with $n_i = 10$ and $n_i = 25$ observations on the *left* and *right* figure, respectively. Factor levels: $a = 3$. Number of covariates: $d = 2$. Covariates: multivariate normal distributed (4.1), $\rho = 0.2$. Dependent variable: unweighted sum of covariates (4.9).

In this power simulation, we have chosen a balanced data set with $n_i = 10$ and $n_i = 20$ observations on the left and on the right figure, respectively. Further, $d = 2$ covariates were observed. On both figures it becomes clear, that considering covariates can dramatically improve the power of the test statistic. Compared to the sie.ats procedure, mctp.t does not perform as good, which could be because the sie.ats procedure is slightly liberal. In this setting, the mctp.par performs better than both non-parametric approaches.

### 4.3.2 Multivariate Log-Normal Distribution

Similar to the type I error simulation, we also conducted a power simulation using the multivariate log-normal distribution. The covariates were generated in the same manner as in (4.1) and then transformed analogously to (4.5). The dependent variable was then given as in (4.6), but with an additional shift parameter $\delta$, i.e.:

$$X_{ik}^{(0)} = \sum_{r=1}^{d} exp(X_{ik}^{(r)}) + U_{ik} + \delta \cdot (i-1), \;\; i = 1, \ldots, a, \;\; k = 1, \ldots, n_i, \tag{4.10}$$

where $U_{ik} \overset{i.i.d.}{\sim} logN(0,1)$. The simulation results can be attained from Figure 5.

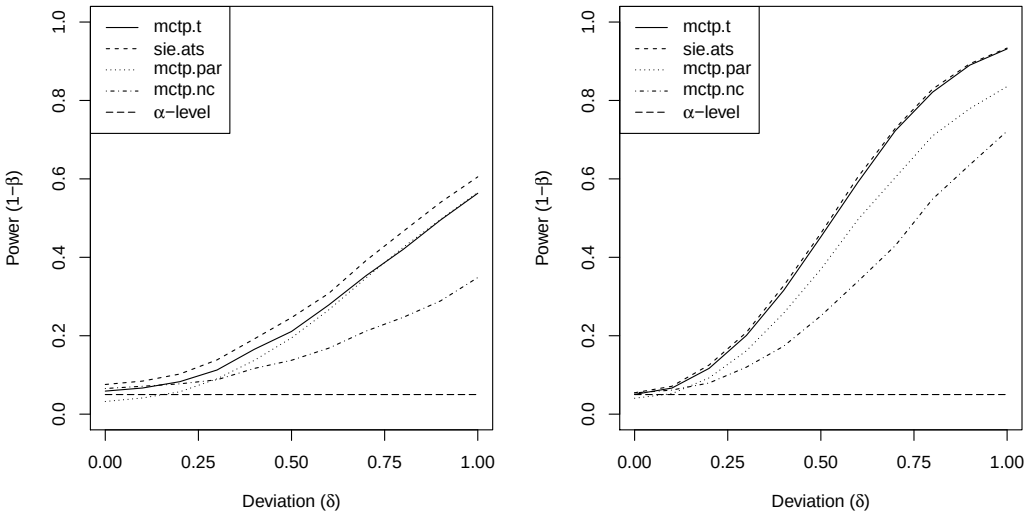


*Figure 5:* Simulation results for the power $(1 - \beta)$. Design: balanced with $n_i = 10$ and $n_i = 25$ observations on the *left* and *right* figure, respectively. Factor levels: $a = 3$. Number of covariates: $d = 2$. Covariates: multivariate log-normal distributed (4.5), $\rho = 0.2$. Dependent variable: unweighted sum of covariates (4.10).

As in the case of the multivariate normal distribution, we chose a balanced design with $n_i = 10$ and $n_i = 25$ observations per group on the left and on the right figure, respectively. The number of covariates was chosen to be $d = 2$ and the number of factor levels is given by $a = 3$. For $n_i = 10$ observations, the power of the statistics is almost equal, with the sie.ats procedure being slightly liberal. For $n_i = 25$ observations, the non-parametric procedures for considering covariates, sie.ats and mctp.t, perform equally well, but better than the parametric alternative mctp.par. As in the case of multivariate normal distributed data, not considering covariates, given this data setting, will lead to a loss in power.

It is important to mention that, while this is a setting in which the non-parametric procedures perform better than the parametric alternatives, this is not always the case when considering the log-normal distribution. As can be seen in the Appendix 7.2.2, raising the number of covariates leads to the parametric procedure being better than the non-parametric procedure.

### 4.3.3 Binomial Distribution

The final power simulation was conducted using binomial distributed covariates, as done for the type I error simulation in (4.7), with parameters $m = 4$ and $q = 0.5$. The dependent variable was calculated analogously to (4.8) with an additional shift term $\delta$, ranging from 0 to 1 in steps of 0.2, i.e.:

$$X_{ik}^{(0)} = \sum_{r=1}^{d} X_{ik}^{(r)} + U_{ik} + \delta \cdot (i-1), \;\; i = 1, \ldots, a, \;\; k = 1, \ldots, n_i, \tag{4.11}$$

where $U_{ik} \overset{i.i.d.}{\sim} N(0, 1)$. The simulation results can be attained from Figure 6.

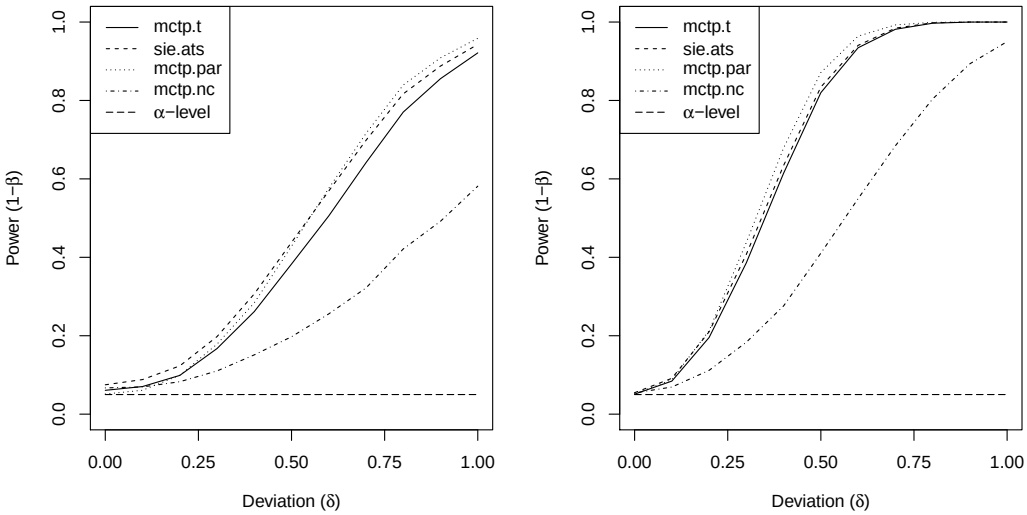


*Figure 6:* Simulation results for the power $(1 - \beta)$. Design: balanced with $n_i = 10$ and $n_i = 25$ observations on the *left* and *right* figure, respectively. Factor levels: $a = 3$. Number of covariates: $d = 2$. Covariates: binomial distributed (4.7). Dependent variable: unweighted sum of covariates (4.11).

The figures show a similar result as for the multivariate normal distributed data. The parametric alternative shows the best results in terms of power, while the sie.ats procedure is the best for non-parametric procedures, which may very well be because it is slightly liberal. The mctp.t still performs adequately.

Concluding the power simulation for balanced designs, we have seen that the slight differences between the non-parametric procedures, i.e. sie.ats and mctp.t, probably result from the sie.ats procedure being slightly liberal, opposed to the mctp.t. The difference is only very small and both procedures perform well in terms of power. Comparing the parametric alternative mctp.par with the mctp.t, we do not have a clearly better procedure, as results vary according to the data setting. Under the presented data settings, an increase in power was always observed when considering covariates, opposed to ignoring these.

### 4.3.4 Unbalanced Design

In Section 4.2.6 we examined the performance of the test procedures in terms of $\alpha$-level, when confronted with an unbalanced design. The aim of this part of the simulation is to verify whether or not the derived procedure mctp.t is still effective in terms of power, when confronted with unequal sample sizes. As in the $\alpha$-level simulation, we chose to compare two different degrees of inequality, a slight inequality between the experiment groups, where sample sizes are $n_1 = 10$, $n_2 = 15$ and $n_3 = 20$, and a grave inequality with sample sizes $n_1 = 10$, $n_2 = 50$ and $n_3 = 100$. Furthermore, we will compare the results for two differently generated datasets. In the first simulation, the underlying data is multivariate normal distributed as described in (4.9), but with dependency parameter $\rho = 0$, i.e. independent covariates, and in the second simulation the data is multivariate log-normal distributed as described in (4.10), again with independent covariates. The results for the first simulation are attained from Figure 7.

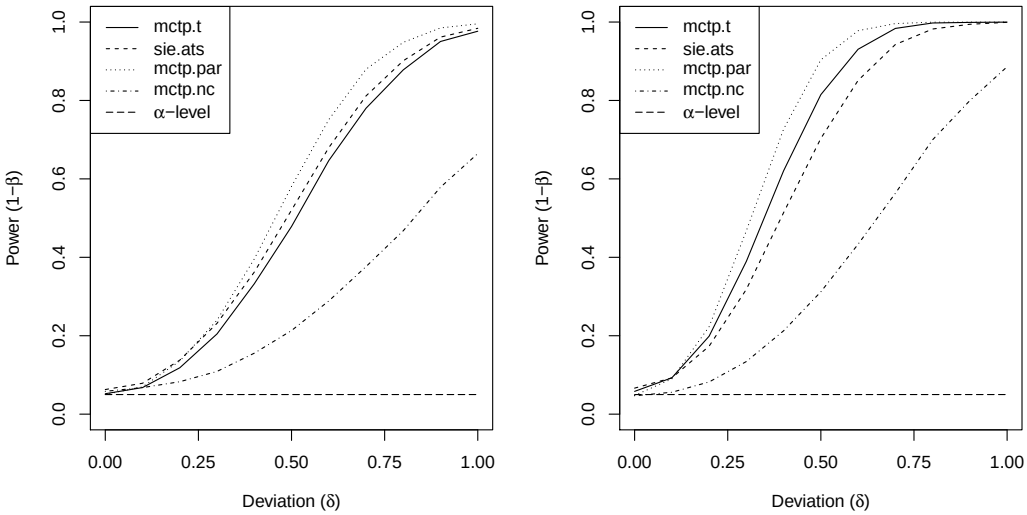


*Figure 7:* Simulation results for the power $(1 - \beta)$. Design: unbalanced with $n_1 = 10$, $n_2 = 15$ and $n_3 = 20$ (*left*) and $n_1 = 10$, $n_2 = 50$ and $n_3 = 100$ (*right*). Factor levels: $a = 3$. Number of covariates: $d = 2$. Covariates: multivariate normal distributed (4.1), $\rho = 0$. Dependent variable: unweighted sum of covariates (4.9).

As in the power simulation for a balanced design and multivariate normal distribution (Figure 4), the best performance is achieved by the parametric approach, mctp.par. Not regarding covariates (mctp.nc) will again yield a lower power. In a slightly unbalanced design, the non-parametric approach sie.ats performs better than the mctp.t, which could be because the sie.ats procedure is more liberal than the mctp.t. This changes though, as the degree of unbalance in the design increases. When increasing the difference of the sample sizes, the mctp.t tends to perform better than the procedure sie.ats.

In the second data setting considered for the power simulation in an unbalanced design, the covariates were chosen to be log-normal distributed as described in (4.5) and the dependent variable was calculated as in (4.10). The results can be attained from Figure 8.

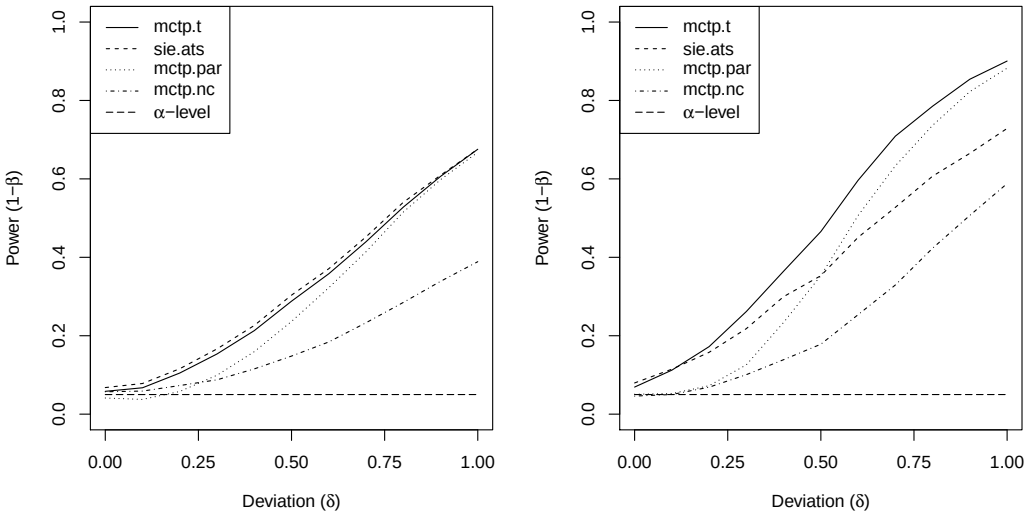


*Figure 8:* Simulation results for the power $(1 - \beta)$. Design: unbalanced with $n_1 = 10$, $n_2 = 15$ and $n_3 = 20$ (*left*) and $n_1 = 10$, $n_2 = 50$ and $n_3 = 100$ (*right*). Factor levels: $a = 3$. Number of covariates: $d = 2$. Covariates: multivariate log-normal distributed (4.5), $\rho = 0$. Dependent variable: unweighted sum of covariates (4.10).

For a small imbalance between the factor levels, the mctp.t and the sie.ats procedure do not seem to differ much. The parametric alternative mctp.par on the other hand, shows an increase in power later than the non-parametric procedures. For a high imbalance between the factor levels, the procedures perform very different from each other. The mctp.t performs best, with a power always higher than that of the other procedures. The sie.ats procedure does not rise as quickly in power, as it did for a low imbalance. The parametric alternative, mcpt.par, again rises in power a little bit later than the non-parametric procedures, this time ending up being better than the sie.ats procedures for $\delta \geq 0.5$. As in the prior simulations, not considering covariates yields the worst power in both degrees of imbalance.

# 5 Example - The Bogalusa Heart Study

Now that the newly developed procedures have been shown to be asymptotically correct and the validity for finite sample sizes has been confirmed through simulations, the procedure will be applied to an example.

The example we will be considering is referred to as the *Bogalusa Heart Study* (BHS), and was previously used as a motivating example in Konietschke et al. (2012b). The aim of this study was to validate a statistical association between a certain genotype and total cholesterol values of participants. For this reason, 525 unrelated individuals of European descent were recruited and 545,821 single-nucleotide polymorphisms (SNP's) of interest were examined on each participant. Additionally to the genomic sequence, 12 clinically relevant traits were examined for each participant.

For the motivating example in Konietschke et al. (2012b), the SNP *rs7738656* in the gene *C6orf170/GJA1* was chosen for statistical analysis. This SNP can have three different genotypes, *GG*, *AG* and *AA*. AA indicates that a subject has an adenine nucleobase on both chromosomes at the relevant SNP. AG indicates that a subject has an adenine nucleobase on one chromosome and a guanine nucleobase on the partner chromosome at the SNP position. GG indicates that a subject has guanine nucleobases on both chromosomes, at the SNP position. The aim of the statistical analysis was to test if carrying a certain genotype, either AA, AG or GG, is responsible for differences in total cholesterol of participants. For the analysis, the multiple contrast test procedure without covariates Konietschke et al. (2012a) was applied. We will now test for differences between the genotypes, but this time regarding some of the 12 additionally examined traits as covariates. More specifically, we will be considering the traits age, heart rate and body mass index as covariates in the statistical analysis, as these traits might also have an influence on the total cholesterol of participants.

## 5.1 Descriptive Statistics

Before going into the analysis, let us take a descriptive look at the given data. For a good interpretation of the analysis, it is important to control the covariates for unequal distributions between the different genotypes. Figure 9 shows boxplots with the covariates split into the genotypes.

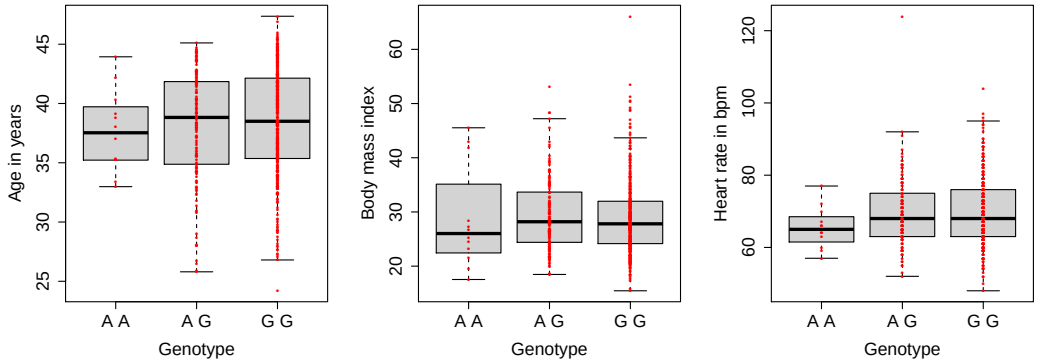


*Figure 9:* Boxplots showing the distribution of the covariates age, body mass index and heart rate.

All covariates seem to be sufficiently equally distributed among the different genotypes and are therefore statistically a good choice, as unequally distributed covariates could results in interpretation difficulties. Figure 10 shows boxplots and histograms for the dependent variable, according to the genotypes.

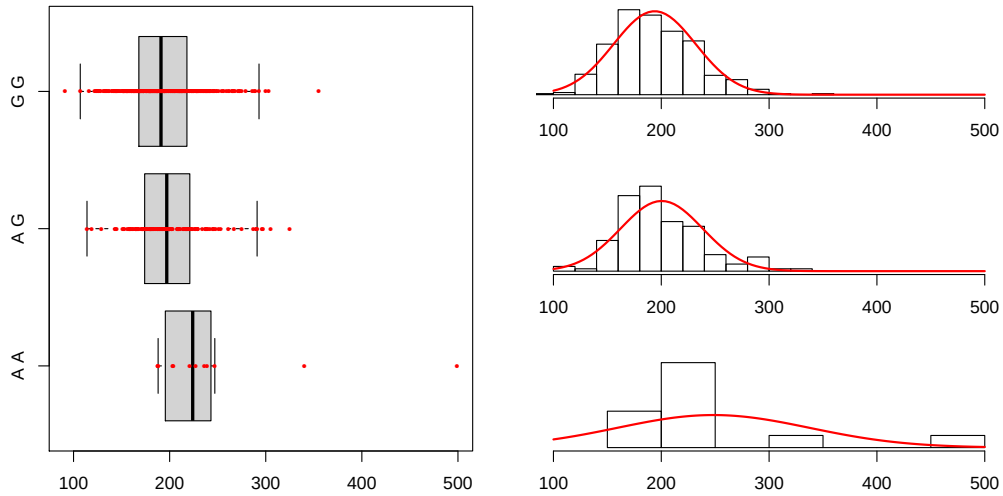


*Figure 10:* Boxplots showing the distribution total cholesterol of subjects with genotypes AA, AG and GG.

While genotypes GG and AG do not seem to show large differences, genotype AA shows higher values for total cholesterol than the other genotypes. Genotype GG shows a symmetric distribution of the total cholesterol, very close to the distribution function of the normal distribution. Genotype AG shows a small amount of skewness to the left. The histogram of genotype AA shows a very skewed distribution. This may very well be because of the two extreme values in only 12 observations for this genotype. Because of this, the assumption of a normal distribution is questionable, and non-parametric procedures should be applied.

## 5.2 Analysis Using R

To allow for the usage of non-parametric multiple contrast test procedures including covariates, an R-script was written (R version 3.0.1). The R-code can be attained from the appendix 7.3 in full detail, while we will explain the application in this section.

The function used for statistical analysis is called *mctp.cov*. The inputs described below can be passed on to the function:

- *formula*: A formula of the form $y \sim group + cov$ has to be passed on, where $y$ refers to the dependent variable, *group* to the group variable and *cov* to a covariate. The number of covariates is technically not limited. The dependent variable and covariates have to be numerical, while the group variable has to be a factor.
- *data*: A data frame containing the data used in the formula.

- *type*: An argument in which the contrast matrix is passed on. It can be either "Tukey", "Dunnett", "Sequen", "Williams", "Changepoint", "AVE", "McDermott", "Marcus", "UmbrellaWilliams" or a contrast matrix, for user defined contrasts. By default, the Tukey matrix is chosen as a contrast matrix.
- *conf.level*: A number between 0 and 1, giving the confidence level to which the global hypothesis is tested. The default confidence level is set to 0.95 (95%).
- *alternative*: The alternative to which the test should be performed. It can be either "two.sided", "less" or "greater". By default, a two sided alternative is tested.
- *asy.method*: The asymptotic method to use. It can be either "mult.t" or "normal", where the multivariate t-approximation is used by default.
- *info*: A logical parameter giving whether or not to post additional information when calling the function.
- *rounds*: The number of digits to which the results should be rounded.
- *effect*: A parameter indicating whether to use unweighted or weighted relative treatment effects. It can be either "unweighted" or "weighted", where unweighted relative treatment effects are chosen by default.

The data set given for the Bogalusa Heart Study was transferred into a data frame suited for the analysis, with following structure:

*Table 11:* Outtake of the Bogalusa Heart Study

| | rs7738656 | tc | AGE | BMI | hr |
|---|---|---|---|---|---|
| 1 | AG | 209 | 38.8 | 42.0 | 63 |
| 2 | GG | 190 | 44.9 | 34.2 | 64 |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| 525 | AG | 144 | 38.0 | 20.5 | 71 |

In genetic models, an allele can either be *dominant*, *additive* or *recessive*. Since it is not known whether the allele is dominant, recessive or additive, we will test for differences between the genotypes by first assuming that G is a dominant allele, secondly by assuming that G is an additive allele and finally by assuming that G is a recessive allele, in a multiple contrast test procedure. A common contrast matrix for this type of testing is the Marcus type contrast matrix, given by:

$$\mathbf{C}_{\text{Marcus}} = \begin{pmatrix} -1 & \frac{n_2}{n_2+n_3} & \frac{n_3}{n_2+n_3} \\ -1 & 0 & 1 \\ -\frac{n_1}{n_1+n_2} & -\frac{n_2}{n_1+n_2} & 1 \end{pmatrix}. \tag{5.1}$$

For testing the hypothesis *no global effect*, i.e. the total cholesterol is not different among the genotypes, or equivalently $H_0 : \mathbf{Cp}^* = \mathbf{0}$, the R-script can be called using following R-code:

```
test<-mctp.cov(tc ~ rs7738656+AGE+BMI+hr,data=BHS,
        info=T,type="Marcus",rounds=6)
```

Executing this code line gives back following result through the console:

```
1  #------Nonparametric Multiple Comparisons for relative effects------#
2
3  - Null Hypothesis: Contrasts of adj. relative treatment effects
4        are unequal 0
5  - Estimation Method: Global pseudo ranks
6  - Type of Contrast: Marcus
7  - Confidence Level: 95 %
8  - Method: Multivariate t-approximation with 66 df
9
10 #-----------------------------------------------------------------------#
11
12   Group Observations
13 1   A A           12
14 2   A G          140
15 3   G G          369
16
17 Number of covariates used: 3 (AGE BMI hr)
18
19 #-----------------------------------------------------------------------#
20
21    Estimator     Lower     Upper Statistic  p.Value
22 C 1 -0.268235 -0.395816 -0.140654 -4.852423 0.000017
23 C 2 -0.277234 -0.405204 -0.149264 -4.999993 0.000009
24 C 3 -0.052023 -0.111314  0.007268 -2.025061 0.095119
25
26 #-----------------------------------------------------------------------#
```

Would the parameter *info* have been set to false, then this output would not have been visible and accessible merely over the newly defined variable *test*.

The visible output gives us some information on the processed data, as well as the test results. The first part (lines 3-8) is intended to confirm the given input by the user. The second part (lines 12-17) helps control for the number of observations processed. In this case for example, not all observations were used, because some covariates were not available on all subjects, and the corresponding subjects were therefore excluded from the analysis. The third part of the visible output (lines 21-24) is the analysis result. Each line corresponds to one individual hypothesis, i.e. one row of the contrast matrix. For example line 22 corresponds to the hypothesis: $H_0 : \mathbf{c}_1\mathbf{p}^* = 0$. The column *Estimator* is equal to $\mathbf{C}\widehat{\mathbf{p}}^*$. The column *Lower* lists the lower limits of the simultaneous confidence intervals for the given contrasts (3.22). The column *Upper* lists the upper limits of the simultaneous confidence intervals for the given contrasts (3.22). The column *Statistic* lists the value of the test statistics for the contrasts (3.18), and finally, the column *p.Value* lists the p-values of the test statistics, calculated by a multivariate t-distribution.

All of the calculated variables can be attained by accessing the newly defined variable *test*. The variable *test* has following structure:

```
1 > str(test)
2 List of 7
3  $ Data.Info  :'data.frame':    3 obs. of  2 variables:
4   ..$ Group       : Factor w/ 3 levels "A_A","A_G","G_G": 1 2 3
5   ..$ Observations: int [1:3] 12 140 369
```

```
$ Contrast    : contrMat [1:3, 1:3] -1 -1 -0.0789 0.275 0 ...
 ..- attr(*, "dimnames")=List of 2
 .. ..$ : chr [1:3] "C_1" "C_2" "C_3"
 .. ..$ : chr [1:3] "A_A" "A_G" "G_G"
 ..- attr(*, "type")= chr "Marcus"
 ..- attr(*, "class")= chr [1:2] "contrMat" "matrix"
$ Analysis    :'data.frame':    3 obs. of  5 variables:
 ..$ Estimator: num [1:3] -0.268 -0.277 -0.052
 ..$ Lower    : num [1:3] -0.396 -0.405 -0.111
 ..$ Upper    : num [1:3] -0.14065 -0.14926 0.00727
 ..$ Statistic: num [1:3] -4.85 -5 -2.03
 ..$ p.Value  : num [1:3] 1.71e-05 9.30e-06 9.51e-02
$ gamma       : num [1:4] -1 0.0799 0.1951 0.0783
$ formula     :Class 'formula' length 3 tc ~ rs7738656 + AGE + BMI + hr
 .. ..- attr(*, ".Environment")=<environment: R_GlobalEnv>
$ data        :'data.frame':    525 obs. of  5 variables:
 ..$ rs7738656: Factor w/ 3 levels "A_A","A_G","G_G": 2 3 3 2 3 3 ...
 ..$ tc       : int [1:525] 209 190 240 219 142 163 197 196 232 233 ...
 ..$ AGE      : num [1:525] 38.8 44.9 43.2 38.3 41.9 ...
 ..$ BMI      : num [1:525] 42 34.2 28.7 31.3 30.2 ...
 ..$ hr       : int [1:525] 63 64 59 53 77 59 71 57 66 65 ...
$ Text.Output:List of 5
 ..$ null.hyp  : chr "Contrasts_of_adj._relative_treatment_effects_..."
 ..$ est.method: chr "Global_pseudo_ranks"
 ..$ contr.type: chr "Marcus"
 ..$ conf.level: num 0.95
 ..$ method    : chr "Multivariate_t-approximation_with_66_df"
```

from which even more estimates, parameters and variables can be accessed, if necessary. For example, the estimated vector of regression parameters $\boldsymbol{\gamma}$ can be accessed through *test$gamma*, where the order corresponds to the order in which the covariates were added through the formula.

## 5.3 Evaluation of the Results

In this particular example, the global hypothesis is rejected to the $\alpha$-level of 0.05, because at least one p-value was lower than 0.05. Konietschke et al. (2012b) come to the same result, only that the p-values for all three individual hypothesis are slightly higher, with $(0.0058, 0.0043, 0.13)$ opposed to $(0.00002, 0.00001, 0.09512)$ when considering covariates. Considering that the first and the second individual hypothesis were rejected, the experimenter can now conclude that, under the assumption that the allele G is dominant or additive, the genotype has an effect on the total cholesterol of the subjects. Subjects with an AA genotype show a significantly higher total cholesterol value, than subjects with an AG or GG genotype.

# 6 Discussion and Outlook

In this thesis we provided for a multiple contrast test procedure for relative treatment effects in a one-factorial design considering covariates. To improve the performance of the test procedure a small sample size approximation using the multivariate t-distribution was introduced. Additionally, an R-script, which can be used for applications of the procedure, was written and demonstrated.

In an extensive simulation study, the validity of the procedure and the small sample size approximation were verified. The simulation results show, that, in terms of $\alpha$-level, the developed small sample size approximation is an improvement over using the asymptotic multivariate normal distribution, and an improvement over using the current non-parametric alternatives for considering covariates, developed by Siemer (1999). It still performs poorly though, especially when the number of covariates is high compared to the number of observations. Simulations showed that to maintain the $\alpha$-level in a balanced design, there should be at least 10 observations per group for each covariate considered. In terms of power, the developed MCTP can compete with current approaches of considering covariates and, in some cases, performs better than parametric alternatives. The developed small sample size approximation should therefore be considered when considering covariates in a non-parametric one-factorial design.

Although the small sample size approximation presented is an improvement over the current non-parametric approaches to considering covariates, it still performs poorly in some data settings. A possible improvement could be considering a transformation of the relative treatment effects. Christophliemk (2001) presented transformation methods for the relative treatment effects when considering covariates, which could still be weaved into this procedure and could yield better results. Another remaining problem is that the resulting estimator for the adjusted relative treatment effects cannot be interpreted as relative treatment effects. The 2-step procedure presented here promises to be a possible alternative, but was not proven to be correct. In fact, it performed poorly in the simulation study, especially when considering a large number of covariates. This could however be, because of a wrong estimation of the degrees of freedom of the multivariate t-distribution which, in the case of the 2-step procedure, does not depend on the number of covariates considered. More insight on this procedure could help understand why it performed so poorly and how it could be improved.

A more elemental point of criticism is the model assumption (2.8) itself. While the assumption of a linear connection between the asymptotic rank transformations of the dependent variable and the covariates is useful for establishing a connection between the two, it is prone to errors. Especially for small sample sizes and a large number of covariates, false coherences between the covariates and the dependent variable are quickly discovered, as the number of permutations a data set can show in the rank transformation is limited. Simulations have shown, that this is a problem which has to be dealt with.

Apart from theoretical improvements, the evaluation program for the user could also be improved. A question which frequently arises when considering covariates, is whether or not covariates even influence the dependent variable. Tests for the influence of covariates on the dependent variable in a non-parametric setting were introduced by Bathke (1998). As we have seen in the simulation study, taking into account covariates which do not influence the dependent variable can lead to a high type I error rate. Testing for effects between the

dependent variable and the covariate, and taking into account only covariates which show a significant influence on the dependent variable might improve the test procedure. Furthermore, the one-sided confidence intervals could be improved by deriving results for upper and lower bounds of the adjusted relative treatment effects. While the relative treatment effects are easily covered, it is more difficult to provide upper and lower bounds for the regression parameters.

Concluding we can say, that although there are many points of criticism and research is not yet finished at this point, the derived simultaneous test procedure for considering covariates in a non-parametric design is applicable and superior to the established non-parametric alternatives, but still performs poorly in some data settings.

# 7 Appendix

## 7.1 Applied Theoretical Results

**Theorem 7.1.1.** *The probability that a coherent simultaneous test procedure* $\{\Pi, \mathbf{S}, \zeta\}$ *of level* $\alpha$ *rejects at least one true* $\pi \in \Pi$ *is* $\alpha$ *if* $\pi_0$ *is true; it is at most* $\alpha$ *irrespective of the truth of* $\pi_0$ *provided* $\{\Pi, \mathbf{S}\}$ *is either closed or joint. The probability of rejecting any particular true* $\pi \in \Pi$ *is no more than the above probability.*

*Proof.* See Gabriel (1969) Theorem 2. □

**Lemma 7.1.2** (Portmanteau)**.** *Let* $X_n$ *and* $X$ *be two random vectors. Then it holds that:*

$$X_n \xrightarrow{D} X \Leftrightarrow \mathrm{E}(f(X_n)) \to \mathrm{E}(f(X))$$

*for all continuous and bounded functions* $f$*.*

*Proof.* See Van der Vaart (1998) p.6. □

**Lemma 7.1.3** (Slutsky)**.** *Let* $X_n$*,* $X$ *and* $Y$ *be random vectors or variables. If* $X_n \xrightarrow{D} X$ *and* $Y_n \xrightarrow{P} c$ *for a constant* $c$*, then:*

1. $X_n + Y_n \xrightarrow{D} X + c$
2. $X_n Y_n \xrightarrow{D} cX$

*Proof.* See Van der Vaart (1998) p.11. □

**Theorem 7.1.4** (Glivenko-Cantelli Theorem)**.** *Let* $X_1, \ldots, X_n$ *denote independently and identically distributed random variables with distribution function* $F$*. Further let* $d_n = \sup_x |\widehat{F}_n(x) - F(x)|$ *denote the supremum of the absolute difference between the distribution function* $F$ *and the empirical distribution function* $\widehat{F}$*. Then:*

$$P\left(\lim_{n\to\infty} d_n = 0\right) = 1 \tag{7.1}$$

*Proof.* See Van der Vaart (1998) p.266. □

**Theorem 7.1.5** ($c_r$-Inequality)**.** *Let* $X$ *and* $Y$ *denote two random variables. Then:*

$$\mathrm{E}\left[|X + Y|^r\right] \leq c_r \cdot \left(\mathrm{E}[|X|^r] + \mathrm{E}[|Y|^r]\right),$$

*where*

$$c_r = \begin{cases} 1 & , \text{ for } 0 < r \leq 1, \\ 2^{r-1} & , \text{ for } r > 1. \end{cases}$$

*Proof.* See Loève (1977) p.157. □

**Theorem 7.1.6** (Jensen's Inequality)**.** *Let* $X$ *denote a random variable with* $\mathrm{E}[X] < \infty$ *and* $g(\cdot)$ *a convex function. Then:*

$$g(\mathrm{E}[X]) \leq \mathrm{E}[g(X)].$$

*Proof.* See Loève (1977) p.161. □

**Theorem 7.1.7** (Lindeberg-Feller). *Let* $X_1, \dots, X_n \overset{i.i.d.}{\sim} F$ *denote random variable with mean* $\mathrm{E}[X_i] = \mu_i$ *and* $\mathrm{Var}[X_i] = \sigma_i^2 > 0$, $i = 1, \dots, n$. *Further denote* $C_n^2 = \sum_{i=1}^n \sigma_i^2$. *Then for it holds that:*

$$\lim_{n\to\infty} \max_{i=1,\dots,n} \frac{\sigma_i}{C_n} = 0 \text{ and } \frac{1}{C_n} \sum_{i=1}^{n} (X_i - \mu_i) \xrightarrow{D} U \sim N(0,1)$$

*if and only if for all* $\varepsilon > 0$ *the Lindeberg condition is fulfilled:*

$$\lim_{n\to\infty} \frac{1}{C_n^2} \sum_{i=1}^{n} \int_{|x-\mu_i|>\varepsilon C_n} (x - \mu_i) dF_i = 0.$$

*While the Lindeberg condition seems unhandy, it can be verified through Corollary (7.1.8).*

*Proof.* See Gnedenko (1962). □

**Corollary 7.1.8.** *Let* $X_1, \dots, X_n$ *be independently distributed and uniformly bounded random variables. Let* $\mathrm{Var}(X_i) = \sigma_i^2 > 0$ *for* $i = 1, \dots, n$. *Then the Lindeberg condition is fulfilled iff* $\sum_{i=1}^n \sigma_i^2 \to \infty$ *for* $n \to \infty$.

*Proof.* See Gnedenko (1962). □

**Theorem 7.1.9** (Cramer-Wold Technique). *Let* $\mathbf{X}_n = (X_{n1}, \dots, X_{nd})'$ *and* $\mathbf{X} = (X_1, \dots, X_d)'$ *denote two d-dimensional random vectors. Then* $\mathbf{X}_n \xrightarrow{D} \mathbf{X}$ *if and only if all linear combinations of* $\mathbf{X}_n$ *converge in distribution against the corresponding linear combinations of* $\mathbf{X}$, *i.e.:*

$$\mathbf{X}_n \xrightarrow{D} \mathbf{X} \Leftrightarrow t'\mathbf{X}_n \xrightarrow{D} t'\mathbf{X} \ \forall t \in \mathbb{R}^d.$$

*Proof.* See Cramer and Wold (1936). □

**Lemma 7.1.10** (Asymptotic Equivalence). *Let* $\mathbf{C} \in \mathbb{R}^{q\times a}$ *be an arbitrary contrast matrix with row vectors* $\mathbf{c}_l = (c_{l1}, \dots, c_{la})$, $l = 1, \dots, q$. *Define:*

$$A_{lik}^{(r)} = c_{li} \left( H^{(r)}(X_{ik}^{(r)}) - \omega_i F_i^{(r)}(X_{ik}^{(r)}) \right) - \sum_{s\neq i} c_{ls}\omega_i F_s^{(r)}(X_{ik}^{(r)})$$

*Then it holds that:*

$$\sqrt{N}\mathbf{c}_l'(\widehat{\mathbf{p}}^{(r)} - \mathbf{p}^{(r)}) \doteq \sqrt{N} \left( \sum_{i=1}^{a} \frac{1}{n_i} \sum_{k=1}^{n_i} A_{lik}^{(r)} - 2 \sum_{i=1}^{a} c_{li} p_i \right)$$

*Proof.* From Brunner and Munzel (2013) p.210 it follows that:

$$\sqrt{N}\left(\widehat{p}_i^{(r)} - p_i^{(r)}\right) \doteq \sqrt{N}\left[\int H^{(r)} d\widehat{F}_i^{(r)} - \int F_i^{(r)} d\widehat{H}^{(r)} + 1 - 2p_i^{(r)}\right]$$

$$= \sqrt{N}\left[\frac{1}{n_i}\sum_{k=1}^{n_i} H^{(r)}(X_{ik}^{(r)}) - \sum_{j=1}^{a} \omega_j \frac{1}{n_j} \sum_{k=1}^{n_j} F_i^{(r)}(X_{jk}^{(r)}) + 1 - 2p_i^{(r)}\right]$$

$$= \sqrt{N}\left[\frac{1}{n_i}\sum_{k=1}^{n_i} \left(H^{(r)}(X_{ik}^{(r)}) - \omega_i F_i^{(r)}(X_{ik}^{(r)})\right) - \sum_{j\neq i}^{a} \omega_j \frac{1}{n_j} \sum_{k=1}^{n_j} F_i^{(r)}(X_{jk}^{(r)}) + 1 - 2p_i^{(r)}\right]$$

$$
\begin{aligned}
&\Rightarrow \sqrt{N}\mathbf{c}_l'(\widehat{\mathbf{p}}^{(r)} - \mathbf{p}^{(r)}) \\
&\doteqdot \sqrt{N}\left[\sum_{i=1}^{a} c_{li}\left(\frac{1}{n_i}\sum_{k=1}^{n_i}\left(H^{(r)}(X_{ik}^{(r)}) - \omega_i F_i^{(r)}(X_{ik}^{(r)})\right) - \sum_{j\neq i}^{a}\omega_j\frac{1}{n_j}\sum_{k=1}^{n_j}F_i^{(r)}(X_{jk}^{(r)}) + 1 - 2p_i^{(r)}\right)\right] \\
&= \sqrt{N}\left[\sum_{i=1}^{a}\frac{1}{n_i}\sum_{k=1}^{n_i} c_{li}\left(H^{(r)}(X_{ik}^{(r)}) - \omega_i F_i^{(r)}(X_{ik}^{(r)})\right) - \sum_{i=1}^{a} c_{li}\sum_{j\neq i}^{a}\omega_j\frac{1}{n_j}\sum_{k=1}^{n_j}F_i^{(r)}(X_{jk}^{(r)}) - 2\sum_{i=1}^{a}c_{li}p_i^{(r)}\right] \\
&= \sqrt{N}\left[\sum_{i=1}^{a}\frac{1}{n_i}\sum_{k=1}^{n_i} c_{li}\left(H^{(r)}(X_{ik}^{(r)}) - \omega_i F_i^{(r)}(X_{ik}^{(r)})\right) - \sum_{i=1}^{a}\frac{1}{n_i}\sum_{k=1}^{n_i}\sum_{s\neq i} c_{ls}\omega_i F_s^{(r)}(X_{ik}^{(r)}) - 2\sum_{i=1}^{a}c_{li}p_i^{(r)}\right] \\
&= \sqrt{N}\left[\sum_{i=1}^{a}\frac{1}{n_i}\sum_{k=1}^{n_i}\left(c_{li}\left(H^{(r)}(X_{ik}^{(r)}) - \omega_i F_i^{(r)}(X_{ik}^{(r)})\right) - \sum_{s\neq i} c_{ls}\omega_i F_s^{(r)}(X_{ik}^{(r)})\right) - 2\sum_{i=1}^{a}c_{li}p_i^{(r)}\right] \\
&= \sqrt{N}\left[\sum_{i=1}^{a}\frac{1}{n_i}\sum_{k=1}^{n_i} A_{lik}^{(r)} - 2\sum_{i=1}^{a}c_{li}p_i^{(r)}\right]
\end{aligned}
$$

□

**Lemma 7.1.11.** *Let the notation be given as in Theorem (3.5.2). In this lemma we will prove that the term $N\mathbf{k}'(\mathbf{I}_a \otimes \boldsymbol{\gamma}')\mathbf{WZ}$ can be written as a sum of independent random variables. For this, denote $\widehat{\mathbf{p}}^{(r)} = (\widehat{p}_1^{(r)}, \ldots, \widehat{p}_a^{(r)})'$ and $\mathbf{p}^{(r)} = (p_1^{(r)}, \ldots, p_a^{(r)})'$.*

*Proof.* Using Theorem (3.5.1) and Lemma (7.1.10) and defining:

$$B_{il}^{(r)} = k_i\left(H^{(r)}(X_{il}^{(r)}) - \omega_i F_i^{(r)}(X_{il}^{(r)})\right) - \sum_{s\neq i} k_s\omega_i F_s^{(r)}(X_{il}^{(r)}),$$

we attain:

$$
\begin{aligned}
&\sqrt{N}\mathbf{k}'(\widehat{\mathbf{p}}^{*} - \mathbf{p}^{*}) = \sum_{r=0}^{d}\gamma^{(r)}\mathbf{k}'\sqrt{N}(\widehat{\mathbf{p}}^{(r)} - \mathbf{p}^{(r)}) \\
&\doteqdot \sum_{r=0}^{d}\gamma^{(r)}\sqrt{N}\left[\sum_{i=1}^{a} k_i\left(\frac{1}{n_i}\sum_{l=1}^{n_i}\left(H^{(r)}(X_{il}^{(r)}) - \omega_i F_i^{(r)}(X_{il}^{(r)})\right) - \sum_{j\neq i}\omega_j\frac{1}{n_j}\sum_{l=1}^{n_j}F_i^{(r)}(X_{jl}^{(r)}) + 1 - 2p_i^{(r)}\right)\right] \\
&= \sum_{r=0}^{d}\gamma^{(r)}\sqrt{N}\left[\sum_{i=1}^{a}\frac{1}{n_i}\sum_{l=1}^{n_i}\left(k_i\left(H^{(r)}(X_{il}^{(r)}) - \omega_i F_i^{(r)}(X_{il}^{(r)})\right) - \sum_{s\neq i} k_s\omega_i F_s^{(r)}(X_{il}^{(r)})\right) + \sum_{i=1}^{a} k_i(1 - 2p_i^{(r)})\right] \\
&= \sum_{r=0}^{d}\gamma^{(r)}\sqrt{N}\left[\sum_{i=1}^{a}\frac{1}{n_i}\sum_{l=1}^{n_i} B_{il}^{(r)} + \sum_{i=1}^{a} k_i(1 - 2p_i^{(r)})\right]
\end{aligned}
$$

Denoting $C_{il} = \sum_{r=0}^{d}\gamma^{(r)}B_{il}^{(r)}$ and $\mu = -\sum_{r=0}^{d}\sum_{i=1}^{a}\gamma^{(r)}k_i(1-2p_i^{(r)})$, we continue our calculations:

$$\sqrt{N}\mathbf{k}'(\widehat{\mathbf{p}}^{*} - \mathbf{p}^{*}) \doteqdot \sqrt{N}\left[\sum_{i=1}^{a}\frac{1}{n_i}\sum_{l=1}^{n_i}\sum_{r=0}^{d}\gamma^{(r)}B_{il}^{(r)} + \sum_{r=0}^{d}\sum_{i=1}^{a}\gamma^{(r)}k_i(1 - 2p_i^{(r)})\right]$$

$$= \sum_{i=1}^{a} \sum_{l=1}^{n_i} \frac{\sqrt{N}}{n_i} C_{il} - \mu$$
$$= \sum_{i=1}^{a} \sum_{l=1}^{n_i} (\frac{\sqrt{N}}{n_i} C_{il} - \mu_{il}),$$

where $\mu_{il}$ are chosen such that $\mathrm{E}(C_{il}) = \mu_{il}$, which implies $\sum_{i=1}^{a} \sum_{l=1}^{n_i} \mu_{il} = \mu$. By construction, $C_{il}$ are independent over the indizes $i$ and $l$. With $\sum_{i=1}^{a} \sum_{l=1}^{n_i} (\frac{N}{n_i} C_{il} - \sqrt{N}\mu_{il})$ we therefore attain a term which is applicable as the enumerator for the Lindeberg-Feller theorem (7.1.7). □

## 7.2 Supplementary Simulation Results

### 7.2.1 Type I Error Simulation

*Table 12:* An $\alpha$-level simulation. Design: balanced with $n_i = 7, 10, 15, 25, 50$ observations. Factor levels: $a = 3$. Number of covariates: $d = 1, 2, 5$. Covariates: multivariate normal distributed (4.1), $\rho = 0.5$. Dependent variable: standard normal distributed (4.2).

| $d$ | $n_i$ | mctp.n | mctp.t | 2-step | sie.ats | sie.wts | mctp.nc | mctp.par | f.par |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 7 | 0.1353 | 0.0842 | 0.0713 | 0.0973 | 0.1441 | 0.0708 | 0.0529 | 0.0525 |
| | 10 | 0.1017 | 0.0725 | 0.0599 | 0.0764 | 0.1070 | 0.0636 | 0.0480 | 0.0496 |
| | 15 | 0.0848 | 0.0641 | 0.0529 | 0.0661 | 0.0874 | 0.0583 | 0.0478 | 0.0476 |
| | 25 | 0.0719 | 0.0615 | 0.0539 | 0.0631 | 0.0739 | 0.0584 | 0.0505 | 0.0508 |
| | 50 | 0.0611 | 0.0565 | 0.0528 | 0.0566 | 0.0622 | 0.0546 | 0.0512 | 0.0507 |
| 2 | 7 | 0.1625 | 0.1001 | 0.0907 | 0.1186 | 0.1697 | 0.0718 | 0.0504 | 0.0512 |
| | 10 | 0.1251 | 0.0883 | 0.0743 | 0.0963 | 0.1310 | 0.0674 | 0.0512 | 0.0516 |
| | 15 | 0.0873 | 0.0676 | 0.0584 | 0.0717 | 0.0917 | 0.0536 | 0.0454 | 0.0451 |
| | 25 | 0.0798 | 0.0673 | 0.0593 | 0.0681 | 0.0816 | 0.0585 | 0.0504 | 0.0507 |
| | 50 | 0.0619 | 0.0569 | 0.0532 | 0.0581 | 0.0626 | 0.0501 | 0.0505 | 0.0509 |
| 5 | 7 | 0.2579 | 0.1742 | 0.1831 | 0.2150 | 0.2641 | 0.0734 | 0.0490 | 0.0512 |
| | 10 | 0.1708 | 0.1265 | 0.1188 | 0.1448 | 0.1772 | 0.0628 | 0.0495 | 0.0485 |
| | 15 | 0.1273 | 0.1029 | 0.0919 | 0.1099 | 0.1317 | 0.0620 | 0.0515 | 0.0504 |
| | 25 | 0.0856 | 0.0736 | 0.0658 | 0.0784 | 0.0882 | 0.0488 | 0.0457 | 0.0475 |
| | 50 | 0.0692 | 0.0645 | 0.0591 | 0.0651 | 0.0690 | 0.0517 | 0.0503 | 0.0510 |

*Table 13:* An $\alpha$-level simulation. Design: balanced with $n_i = 7, 10, 15, 25, 50$ observations. Factor levels: $a = 3$. Number of covariates: $d = 1, 2, 5$. Covariates: multivariate normal distributed (4.1), $\rho = 0.9$. Dependent variable: standard normal distributed (4.2).

| $d$ | $n_i$ | mctp.n | mctp.t | 2-step | sie.ats | sie.wts | mctp.nc | mctp.par | f.par |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 7 | 0.1428 | 0.0878 | 0.0724 | 0.0984 | 0.1525 | 0.0752 | 0.0576 | 0.0557 |
| | 10 | 0.1029 | 0.0703 | 0.0595 | 0.0768 | 0.1068 | 0.0622 | 0.0475 | 0.0473 |
| | 15 | 0.0857 | 0.0662 | 0.0553 | 0.0694 | 0.0886 | 0.0586 | 0.0501 | 0.0505 |
| | 25 | 0.0688 | 0.0596 | 0.0528 | 0.0596 | 0.0716 | 0.0553 | 0.0495 | 0.0497 |
| | 50 | 0.0627 | 0.0566 | 0.0534 | 0.0584 | 0.0632 | 0.0566 | 0.0516 | 0.0519 |
| 2 | 7 | 0.1599 | 0.0980 | 0.0854 | 0.1179 | 0.1690 | 0.0697 | 0.0463 | 0.0483 |
| | 10 | 0.1211 | 0.0841 | 0.0708 | 0.0939 | 0.1259 | 0.0654 | 0.0531 | 0.0533 |
| | 15 | 0.0905 | 0.0679 | 0.0607 | 0.0745 | 0.0922 | 0.0564 | 0.0463 | 0.0457 |
| | 25 | 0.0764 | 0.0655 | 0.0580 | 0.0659 | 0.0787 | 0.0557 | 0.0501 | 0.0488 |
| | 50 | 0.0655 | 0.0608 | 0.0566 | 0.0601 | 0.0657 | 0.0568 | 0.0524 | 0.0527 |
| 5 | 7 | 0.2570 | 0.1731 | 0.1830 | 0.2125 | 0.2638 | 0.0709 | 0.0500 | 0.0500 |
| | 10 | 0.1710 | 0.1263 | 0.1152 | 0.1446 | 0.1780 | 0.0628 | 0.0469 | 0.0460 |
| | 15 | 0.1247 | 0.0991 | 0.0877 | 0.1065 | 0.1254 | 0.0611 | 0.0514 | 0.0506 |
| | 25 | 0.0903 | 0.0771 | 0.0677 | 0.0805 | 0.0926 | 0.0566 | 0.0522 | 0.0538 |
| | 50 | 0.0706 | 0.0645 | 0.0595 | 0.0642 | 0.0691 | 0.0516 | 0.0500 | 0.0490 |

*Table 14:* An $\alpha$-level simulation. Design: balanced with $n_i = 7, 10, 15, 25, 50$ observations. Factor levels: $a = 3$. Number of covariates: $d = 1, 2, 5$. Covariates: multivariate normal distributed (4.1), $\rho = 0$. Dependent variable: unweighted sum of covariates (4.3). Contrast matrix: Centering.

| $d$ | $n_i$ | mctp.n | mctp.t | 2-step | sie.ats | sie.wts | mctp.nc | mctp.par | f.par |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 7 | 0.1020 | 0.0511 | 0.0790 | 0.0737 | 0.1072 | 0.0639 | 0.0477 | 0.0480 |
| | 10 | 0.0784 | 0.0469 | 0.0649 | 0.0621 | 0.0818 | 0.0564 | 0.0480 | 0.0489 |
| | 15 | 0.0688 | 0.0492 | 0.0619 | 0.0565 | 0.0724 | 0.0564 | 0.0509 | 0.0510 |
| | 25 | 0.0623 | 0.0521 | 0.0591 | 0.0562 | 0.0626 | 0.0527 | 0.0543 | 0.0539 |
| | 50 | 0.0558 | 0.0512 | 0.0537 | 0.0536 | 0.0562 | 0.0515 | 0.0532 | 0.0531 |
| 2 | 7 | 0.1023 | 0.0531 | 0.0984 | 0.0792 | 0.1111 | 0.0648 | 0.0494 | 0.0499 |
| | 10 | 0.0858 | 0.0525 | 0.0805 | 0.0705 | 0.0888 | 0.0611 | 0.0506 | 0.0523 |
| | 15 | 0.0719 | 0.0513 | 0.0709 | 0.0639 | 0.0731 | 0.0546 | 0.0506 | 0.0505 |
| | 25 | 0.0616 | 0.0516 | 0.0586 | 0.0541 | 0.0630 | 0.0582 | 0.0497 | 0.0505 |
| | 50 | 0.0504 | 0.0453 | 0.0496 | 0.0495 | 0.0495 | 0.0491 | 0.0463 | 0.0469 |
| 5 | 7 | 0.1563 | 0.0931 | 0.1985 | 0.1305 | 0.1621 | 0.0668 | 0.0487 | 0.0490 |
| | 10 | 0.1097 | 0.0713 | 0.1387 | 0.0974 | 0.1141 | 0.0554 | 0.0530 | 0.0538 |
| | 15 | 0.0837 | 0.0634 | 0.1015 | 0.0758 | 0.0882 | 0.0599 | 0.0492 | 0.0481 |
| | 25 | 0.0667 | 0.0537 | 0.0744 | 0.0608 | 0.0653 | 0.0561 | 0.0494 | 0.0495 |
| | 50 | 0.0586 | 0.0526 | 0.0635 | 0.0561 | 0.0606 | 0.0488 | 0.0497 | 0.0491 |

*Table 15:* An $\alpha$-level simulation. Design: balanced with $n_i = 7, 10, 15, 25, 50$ observations. Factor levels: $a = 3$. Number of covariates: $d = 1, 2, 5$. Covariates: multivariate normal distributed (4.1), $\rho = 0$. Dependent variable: unweighted sum of covariates (4.3). Contrast matrix: Changepoint.

| $d$ | $n_i$ | mctp.n | mctp.t | 2-step | sie.ats | sie.wts | mctp.nc | mctp.par | f.par |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 7 | 0.0931 | 0.0513 | 0.0677 | 0.0750 | 0.1099 | 0.0618 | 0.0518 | 0.0517 |
| | 10 | 0.0718 | 0.0455 | 0.0566 | 0.0598 | 0.0827 | 0.0566 | 0.0468 | 0.0477 |
| | 15 | 0.0647 | 0.0477 | 0.0514 | 0.0598 | 0.0724 | 0.0562 | 0.0535 | 0.0536 |
| | 25 | 0.0605 | 0.0521 | 0.0552 | 0.0543 | 0.0628 | 0.0528 | 0.0518 | 0.0491 |
| | 50 | 0.0546 | 0.0507 | 0.0525 | 0.0538 | 0.0570 | 0.0503 | 0.0513 | 0.0524 |
| 2 | 7 | 0.0960 | 0.0543 | 0.0835 | 0.0803 | 0.1122 | 0.0621 | 0.0503 | 0.0503 |
| | 10 | 0.0762 | 0.0503 | 0.0729 | 0.0678 | 0.0855 | 0.0567 | 0.0489 | 0.0509 |
| | 15 | 0.0695 | 0.0531 | 0.0650 | 0.0622 | 0.0738 | 0.0570 | 0.0533 | 0.0525 |
| | 25 | 0.0577 | 0.0495 | 0.0536 | 0.0542 | 0.0612 | 0.0535 | 0.0502 | 0.0503 |
| | 50 | 0.0541 | 0.0505 | 0.0538 | 0.0548 | 0.0574 | 0.0508 | 0.0518 | 0.0542 |
| 5 | 7 | 0.1437 | 0.0915 | 0.1754 | 0.1318 | 0.1683 | 0.0602 | 0.0489 | 0.0497 |
| | 10 | 0.0996 | 0.0707 | 0.1132 | 0.0900 | 0.1102 | 0.0578 | 0.0496 | 0.0484 |
| | 15 | 0.0747 | 0.0577 | 0.0818 | 0.0705 | 0.0799 | 0.0502 | 0.0508 | 0.0522 |
| | 25 | 0.0627 | 0.0541 | 0.0674 | 0.0616 | 0.0680 | 0.0494 | 0.0503 | 0.0482 |
| | 50 | 0.0585 | 0.0537 | 0.0612 | 0.0605 | 0.0616 | 0.0498 | 0.0528 | 0.0532 |

*Table 16:* An $\alpha$-level simulation. Design: balanced with $n_i = 7, 10, 15, 25, 50$ observations. Factor levels: $a = 3$. Number of covariates: $d = 1, 2, 5$. Covariates: multivariate normal distributed (4.1), $\rho = 0$. Dependent variable: unweighted sum of covariates (4.3). Contrast matrix: Dunnett.

| $d$ | $n_i$ | mctp.n | mctp.t | 2-step | sie.ats | sie.wts | mctp.nc | mctp.par | f.par |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 7 | 0.0877 | 0.0538 | 0.0622 | 0.0722 | 0.1067 | 0.0630 | 0.0479 | 0.0477 |
| | 10 | 0.0805 | 0.0538 | 0.0576 | 0.0657 | 0.0875 | 0.0638 | 0.0534 | 0.0558 |
| | 15 | 0.0650 | 0.0501 | 0.0561 | 0.0577 | 0.0713 | 0.0549 | 0.0491 | 0.0516 |
| | 25 | 0.0558 | 0.0482 | 0.0487 | 0.0515 | 0.0575 | 0.0558 | 0.0465 | 0.0484 |
| | 50 | 0.0528 | 0.0486 | 0.0494 | 0.0505 | 0.0544 | 0.0494 | 0.0510 | 0.0514 |
| 2 | 7 | 0.0928 | 0.0575 | 0.0817 | 0.0803 | 0.1096 | 0.0642 | 0.0500 | 0.0502 |
| | 10 | 0.0771 | 0.0509 | 0.0654 | 0.0659 | 0.0878 | 0.0576 | 0.0505 | 0.0502 |
| | 15 | 0.0672 | 0.0513 | 0.0610 | 0.0588 | 0.0717 | 0.0547 | 0.0515 | 0.0500 |
| | 25 | 0.0574 | 0.0483 | 0.0537 | 0.0518 | 0.0567 | 0.0499 | 0.0488 | 0.0478 |
| | 50 | 0.0512 | 0.0480 | 0.0519 | 0.0503 | 0.0544 | 0.0509 | 0.0481 | 0.0490 |
| 5 | 7 | 0.1461 | 0.0941 | 0.1695 | 0.1314 | 0.1652 | 0.0651 | 0.0461 | 0.0490 |
| | 10 | 0.1039 | 0.0758 | 0.1166 | 0.0950 | 0.1133 | 0.0587 | 0.0522 | 0.0516 |
| | 15 | 0.0797 | 0.0629 | 0.0851 | 0.0736 | 0.0852 | 0.0561 | 0.0487 | 0.0494 |
| | 25 | 0.0632 | 0.0554 | 0.0669 | 0.0639 | 0.0693 | 0.0529 | 0.0512 | 0.0512 |
| | 50 | 0.0547 | 0.0518 | 0.0554 | 0.0545 | 0.0565 | 0.0518 | 0.0484 | 0.0490 |

*Table 17:* An $\alpha$-level simulation. Design: balanced with $n_i = 7, 10, 15, 25, 50$ observations. Factor levels: $a = 3$. Number of covariates: $d = 1, 2, 5$. Covariates: multivariate normal distributed (4.1), $\rho = 0.5$. Dependent variable: unweighted sum of covariates (4.3).

| $d$ | $n_i$ | mctp.n | mctp.t | 2-step | sie.ats | sie.wts | mctp.nc | mctp.par | f.par |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 7 | 0.1428 | 0.0878 | 0.0724 | 0.0984 | 0.1525 | 0.0752 | 0.0576 | 0.0557 |
| | 10 | 0.1029 | 0.0703 | 0.0595 | 0.0768 | 0.1068 | 0.0622 | 0.0475 | 0.0473 |
| | 15 | 0.0857 | 0.0662 | 0.0553 | 0.0694 | 0.0886 | 0.0586 | 0.0501 | 0.0505 |
| | 25 | 0.0688 | 0.0596 | 0.0528 | 0.0596 | 0.0716 | 0.0553 | 0.0495 | 0.0497 |
| | 50 | 0.0627 | 0.0566 | 0.0534 | 0.0584 | 0.0632 | 0.0566 | 0.0516 | 0.0519 |
| 2 | 7 | 0.1599 | 0.0980 | 0.0854 | 0.1179 | 0.1690 | 0.0697 | 0.0463 | 0.0483 |
| | 10 | 0.1211 | 0.0841 | 0.0708 | 0.0939 | 0.1259 | 0.0654 | 0.0531 | 0.0533 |
| | 15 | 0.0905 | 0.0679 | 0.0607 | 0.0745 | 0.0922 | 0.0564 | 0.0463 | 0.0457 |
| | 25 | 0.0764 | 0.0655 | 0.0580 | 0.0659 | 0.0787 | 0.0557 | 0.0501 | 0.0488 |
| | 50 | 0.0655 | 0.0608 | 0.0566 | 0.0601 | 0.0657 | 0.0568 | 0.0524 | 0.0527 |
| 5 | 7 | 0.2570 | 0.1731 | 0.1830 | 0.2125 | 0.2638 | 0.0709 | 0.0500 | 0.0500 |
| | 10 | 0.1710 | 0.1263 | 0.1152 | 0.1446 | 0.1780 | 0.0628 | 0.0469 | 0.0460 |
| | 15 | 0.1247 | 0.0991 | 0.0877 | 0.1065 | 0.1254 | 0.0611 | 0.0514 | 0.0506 |
| | 25 | 0.0903 | 0.0771 | 0.0677 | 0.0805 | 0.0926 | 0.0566 | 0.0522 | 0.0538 |
| | 50 | 0.0706 | 0.0645 | 0.0595 | 0.0642 | 0.0691 | 0.0516 | 0.0500 | 0.0490 |

*Table 18:* An $\alpha$-level simulation. Design: balanced with $n_i = 7, 10, 15, 25, 50$ observations. Factor levels: $a = 3$. Number of covariates: $d = 1, 2, 5$. Covariates: multivariate normal distributed (4.1), $\rho = 0.9$. Dependent variable: unweighted sum of covariates (4.3).

| $d$ | $n_i$ | mctp.n | mctp.t | 2-step | sie.ats | sie.wts | mctp.nc | mctp.par | f.par |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 7 | 0.1078 | 0.0587 | 0.0745 | 0.0761 | 0.1130 | 0.0709 | 0.0515 | 0.0518 |
| | 10 | 0.0809 | 0.0518 | 0.0574 | 0.0646 | 0.0848 | 0.0621 | 0.0488 | 0.0502 |
| | 15 | 0.0681 | 0.0527 | 0.0560 | 0.0587 | 0.0704 | 0.0607 | 0.0512 | 0.0520 |
| | 25 | 0.0569 | 0.0451 | 0.0470 | 0.0492 | 0.0572 | 0.0520 | 0.0465 | 0.0459 |
| | 50 | 0.0551 | 0.0507 | 0.0534 | 0.0519 | 0.0558 | 0.0523 | 0.0493 | 0.0497 |
| 2 | 7 | 0.1081 | 0.0594 | 0.0934 | 0.0797 | 0.1132 | 0.0735 | 0.0520 | 0.0534 |
| | 10 | 0.0810 | 0.0504 | 0.0726 | 0.0648 | 0.0837 | 0.0595 | 0.0482 | 0.0484 |
| | 15 | 0.0690 | 0.0510 | 0.0613 | 0.0574 | 0.0710 | 0.0603 | 0.0495 | 0.0486 |
| | 25 | 0.0605 | 0.0505 | 0.0584 | 0.0561 | 0.0610 | 0.0526 | 0.0534 | 0.0531 |
| | 50 | 0.0517 | 0.0468 | 0.0525 | 0.0512 | 0.0520 | 0.0544 | 0.0497 | 0.0492 |
| 5 | 7 | 0.1535 | 0.0908 | 0.2032 | 0.1268 | 0.1580 | 0.0729 | 0.0516 | 0.0512 |
| | 10 | 0.1075 | 0.0727 | 0.1330 | 0.0937 | 0.1128 | 0.0639 | 0.0547 | 0.0548 |
| | 15 | 0.0765 | 0.0587 | 0.0911 | 0.0675 | 0.0757 | 0.0622 | 0.0505 | 0.0507 |
| | 25 | 0.0656 | 0.0555 | 0.0726 | 0.0612 | 0.0665 | 0.0526 | 0.0502 | 0.0492 |
| | 50 | 0.0526 | 0.0478 | 0.0556 | 0.0504 | 0.0540 | 0.0524 | 0.0491 | 0.0463 |

*Table 19:* An $\alpha$-level simulation. Design: balanced with $n_i = 7, 10, 15, 25, 50$ observations. Factor levels: $a = 3$. Number of covariates: $d = 1, 2, 5$. Covariates: multivariate normal distributed (4.1), $\rho = 0.5$. Dependent variable: weighted sum of covariates (4.4).

| $d$ | $n_i$ | mctp.n | mctp.t | 2-step | sie.ats | sie.wts | mctp.nc | mctp.par | f.par |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 7 | 0.1227 | 0.0726 | 0.0701 | 0.0870 | 0.1304 | 0.0725 | 0.0522 | 0.0518 |
| | 10 | 0.0969 | 0.0658 | 0.0598 | 0.0752 | 0.1013 | 0.0624 | 0.0495 | 0.0506 |
| | 15 | 0.0787 | 0.0601 | 0.0543 | 0.0658 | 0.0819 | 0.0580 | 0.0511 | 0.0504 |
| | 25 | 0.0636 | 0.0530 | 0.0493 | 0.0547 | 0.0632 | 0.0544 | 0.0475 | 0.0464 |
| | 50 | 0.0597 | 0.0549 | 0.0521 | 0.0568 | 0.0595 | 0.0542 | 0.0531 | 0.0543 |
| 2 | 7 | 0.1215 | 0.0654 | 0.0996 | 0.0947 | 0.1266 | 0.0763 | 0.0562 | 0.0552 |
| | 10 | 0.0829 | 0.0506 | 0.0662 | 0.0670 | 0.0855 | 0.0600 | 0.0460 | 0.0444 |
| | 15 | 0.0732 | 0.0539 | 0.0627 | 0.0622 | 0.0738 | 0.0582 | 0.0490 | 0.0481 |
| | 25 | 0.0660 | 0.0560 | 0.0597 | 0.0582 | 0.0654 | 0.0538 | 0.0527 | 0.0518 |
| | 50 | 0.0565 | 0.0512 | 0.0560 | 0.0532 | 0.0555 | 0.0583 | 0.0502 | 0.0504 |
| 5 | 7 | 0.0944 | 0.0513 | 0.1928 | 0.0708 | 0.1001 | 0.0684 | 0.0537 | 0.0530 |
| | 10 | 0.0578 | 0.0356 | 0.1216 | 0.0474 | 0.0600 | 0.0653 | 0.0518 | 0.0532 |
| | 15 | 0.0493 | 0.0339 | 0.0955 | 0.0415 | 0.0490 | 0.0556 | 0.0519 | 0.0515 |
| | 25 | 0.0459 | 0.0377 | 0.0784 | 0.0425 | 0.0451 | 0.0584 | 0.0527 | 0.0516 |
| | 50 | 0.0422 | 0.0385 | 0.0573 | 0.0407 | 0.0407 | 0.0546 | 0.0494 | 0.0478 |

*Table 20:* An $\alpha$-level simulation. Design: balanced with $n_i = 7, 10, 15, 25, 50$ observations. Factor levels: $a = 3$. Number of covariates: $d = 1, 2, 5$. Covariates: multivariate normal distributed (4.1), $\rho = 0.9$. Dependent variable: weighted sum of covariates (4.4).

| $d$ | $n_i$ | mctp.n | mctp.t | 2-step | sie.ats | sie.wts | mctp.nc | mctp.par | f.par |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 7 | 0.1326 | 0.0768 | 0.0739 | 0.0926 | 0.1395 | 0.0742 | 0.0527 | 0.0528 |
| | 10 | 0.0957 | 0.0656 | 0.0581 | 0.0745 | 0.1020 | 0.0637 | 0.0498 | 0.0506 |
| | 15 | 0.0811 | 0.0623 | 0.0585 | 0.0660 | 0.0841 | 0.0603 | 0.0531 | 0.0521 |
| | 25 | 0.0618 | 0.0517 | 0.0486 | 0.0538 | 0.0631 | 0.0518 | 0.0439 | 0.0431 |
| | 50 | 0.0596 | 0.0547 | 0.0524 | 0.0554 | 0.0596 | 0.0537 | 0.0527 | 0.0528 |
| 2 | 7 | 0.1268 | 0.0728 | 0.0981 | 0.0983 | 0.1337 | 0.0773 | 0.0533 | 0.0531 |
| | 10 | 0.0895 | 0.0590 | 0.0740 | 0.0726 | 0.0945 | 0.0603 | 0.0469 | 0.0487 |
| | 15 | 0.0780 | 0.0575 | 0.0661 | 0.0649 | 0.0793 | 0.0619 | 0.0505 | 0.0491 |
| | 25 | 0.0684 | 0.0551 | 0.0596 | 0.0606 | 0.0666 | 0.0575 | 0.0539 | 0.0537 |
| | 50 | 0.0527 | 0.0480 | 0.0512 | 0.0499 | 0.0535 | 0.0522 | 0.0467 | 0.0465 |
| 5 | 7 | 0.1170 | 0.0671 | 0.1875 | 0.0917 | 0.1210 | 0.0733 | 0.0507 | 0.0499 |
| | 10 | 0.0785 | 0.0526 | 0.1219 | 0.0645 | 0.0802 | 0.0678 | 0.0525 | 0.0525 |
| | 15 | 0.0598 | 0.0463 | 0.0879 | 0.0514 | 0.0618 | 0.0626 | 0.0524 | 0.0522 |
| | 25 | 0.0521 | 0.0422 | 0.0683 | 0.0467 | 0.0507 | 0.0543 | 0.0502 | 0.0495 |
| | 50 | 0.0485 | 0.0431 | 0.0570 | 0.0457 | 0.0474 | 0.0547 | 0.0462 | 0.0467 |

*Table 21:* An $\alpha$-level simulation. Design: balanced with $n_i = 7, 10, 15, 25, 50$ observations. Factor levels: $a = 3$. Number of covariates: $d = 1, 2, 5$. Covariates: multivariate log-normal distributed (4.5), $\rho = 0.5$. Dependent variable: unweighted sum of covariates (4.6).

| $d$ | $n_i$ | mctp.n | mctp.t | 2-step | sie.ats | sie.wts | mctp.nc | mctp.par | f.par |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 7 | 0.1042 | 0.0570 | 0.0703 | 0.0736 | 0.1143 | 0.0731 | 0.0343 | 0.0364 |
| | 10 | 0.0831 | 0.0519 | 0.0586 | 0.0630 | 0.0871 | 0.0646 | 0.0317 | 0.0333 |
| | 15 | 0.0749 | 0.0546 | 0.0617 | 0.0629 | 0.0794 | 0.0622 | 0.0363 | 0.0375 |
| | 25 | 0.0617 | 0.0503 | 0.0527 | 0.0544 | 0.0633 | 0.0578 | 0.0351 | 0.0352 |
| | 50 | 0.0583 | 0.0527 | 0.0535 | 0.0546 | 0.0590 | 0.0533 | 0.0450 | 0.0466 |
| 2 | 7 | 0.1049 | 0.0537 | 0.0939 | 0.0784 | 0.1107 | 0.0731 | 0.0355 | 0.0361 |
| | 10 | 0.0815 | 0.0497 | 0.0727 | 0.0652 | 0.0868 | 0.0660 | 0.0331 | 0.0364 |
| | 15 | 0.0677 | 0.0494 | 0.0624 | 0.0552 | 0.0738 | 0.0596 | 0.0372 | 0.0400 |
| | 25 | 0.0597 | 0.0495 | 0.0590 | 0.0565 | 0.0635 | 0.0572 | 0.0395 | 0.0399 |
| | 50 | 0.0565 | 0.0510 | 0.0527 | 0.0528 | 0.0590 | 0.0533 | 0.0422 | 0.0439 |
| 5 | 7 | 0.1514 | 0.0883 | 0.1881 | 0.1270 | 0.1575 | 0.0720 | 0.0375 | 0.0377 |
| | 10 | 0.1080 | 0.0683 | 0.1338 | 0.0932 | 0.1154 | 0.0607 | 0.0352 | 0.0370 |
| | 15 | 0.0764 | 0.0565 | 0.0926 | 0.0702 | 0.0801 | 0.0594 | 0.0363 | 0.0388 |
| | 25 | 0.0658 | 0.0537 | 0.0770 | 0.0613 | 0.0708 | 0.0556 | 0.0399 | 0.0405 |
| | 50 | 0.0558 | 0.0508 | 0.0603 | 0.0546 | 0.0591 | 0.0487 | 0.0436 | 0.0441 |

*Table 22:* An $\alpha$-level simulation. Design: balanced with $n_i = 7, 10, 15, 25, 50$ observations. Factor levels: $a = 3$. Number of covariates: $d = 1, 2, 5$. Covariates: multivariate log-normal distributed (4.5), $\rho = 0.9$. Dependent variable: unweighted sum of covariates (4.6).

| $d$ | $n_i$ | mctp.n | mctp.t | 2-step | sie.ats | sie.wts | mctp.nc | mctp.par | f.par |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 7 | 0.1074 | 0.0563 | 0.0724 | 0.0746 | 0.1157 | 0.0764 | 0.0336 | 0.0356 |
| | 10 | 0.0794 | 0.0524 | 0.0592 | 0.0592 | 0.0850 | 0.0635 | 0.0321 | 0.0349 |
| | 15 | 0.0694 | 0.0522 | 0.0550 | 0.0563 | 0.0733 | 0.0621 | 0.0364 | 0.0374 |
| | 25 | 0.0579 | 0.0469 | 0.0478 | 0.0507 | 0.0597 | 0.0531 | 0.0350 | 0.0365 |
| | 50 | 0.0555 | 0.0515 | 0.0503 | 0.0518 | 0.0553 | 0.0533 | 0.0429 | 0.0434 |
| 2 | 7 | 0.0937 | 0.0457 | 0.0933 | 0.0766 | 0.1011 | 0.0727 | 0.0354 | 0.0371 |
| | 10 | 0.0746 | 0.0443 | 0.0751 | 0.0624 | 0.0809 | 0.0639 | 0.0306 | 0.0330 |
| | 15 | 0.0638 | 0.0429 | 0.0618 | 0.0555 | 0.0685 | 0.0530 | 0.0349 | 0.0369 |
| | 25 | 0.0598 | 0.0491 | 0.0602 | 0.0556 | 0.0635 | 0.0554 | 0.0386 | 0.0393 |
| | 50 | 0.0527 | 0.0482 | 0.0504 | 0.0502 | 0.0564 | 0.0537 | 0.0408 | 0.0411 |
| 5 | 7 | 0.1357 | 0.0738 | 0.1961 | 0.1129 | 0.1357 | 0.0714 | 0.0376 | 0.0376 |
| | 10 | 0.0930 | 0.0563 | 0.1285 | 0.0830 | 0.0975 | 0.0632 | 0.0405 | 0.0415 |
| | 15 | 0.0708 | 0.0492 | 0.0970 | 0.0647 | 0.0735 | 0.0603 | 0.0358 | 0.0373 |
| | 25 | 0.0551 | 0.0424 | 0.0680 | 0.0533 | 0.0595 | 0.0513 | 0.0397 | 0.0420 |
| | 50 | 0.0516 | 0.0452 | 0.0604 | 0.0503 | 0.0572 | 0.0567 | 0.0403 | 0.0407 |

*Table 23:* An $\alpha$-level simulation. Design: balanced with $n_i = 7, 10, 15, 25, 50$ observations. Factor levels: $a = 3$. Number of covariates: $d = 1, 2, 5$. Covariates: Poisson(1) distributed. Dependent variable: unweighted sum of covariates with standard normal error term.

| $d$ | $n_i$ | mctp.n | mctp.t | 2-step | sie.ats | sie.wts | mctp.nc | mctp.par | f.par |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 7 | 0.1118 | 0.0635 | 0.0745 | 0.0788 | 0.1168 | 0.0742 | 0.0536 | 0.0533 |
| | 10 | 0.0891 | 0.0562 | 0.0640 | 0.0700 | 0.0919 | 0.0590 | 0.0497 | 0.0493 |
| | 15 | 0.0687 | 0.0506 | 0.0519 | 0.0549 | 0.0709 | 0.0547 | 0.0453 | 0.0455 |
| | 25 | 0.0637 | 0.0544 | 0.0536 | 0.0598 | 0.0650 | 0.0529 | 0.0527 | 0.0527 |
| | 50 | 0.0590 | 0.0537 | 0.0537 | 0.0550 | 0.0588 | 0.0547 | 0.0505 | 0.0503 |
| 2 | 7 | 0.1194 | 0.0651 | 0.0878 | 0.0922 | 0.1216 | 0.0694 | 0.0488 | 0.0501 |
| | 10 | 0.0877 | 0.0554 | 0.0694 | 0.0703 | 0.0885 | 0.0638 | 0.0478 | 0.0482 |
| | 15 | 0.0728 | 0.0539 | 0.0602 | 0.0629 | 0.0732 | 0.0586 | 0.0486 | 0.0473 |
| | 25 | 0.0618 | 0.0519 | 0.0551 | 0.0548 | 0.0622 | 0.0559 | 0.0488 | 0.0489 |
| | 50 | 0.0574 | 0.0526 | 0.0559 | 0.0525 | 0.0553 | 0.0523 | 0.0524 | 0.0517 |
| 5 | 7 | 0.1882 | 0.1178 | 0.1914 | 0.1599 | 0.1941 | 0.0680 | 0.0512 | 0.0533 |
| | 10 | 0.1318 | 0.0886 | 0.1268 | 0.1126 | 0.1352 | 0.0658 | 0.0470 | 0.0459 |
| | 15 | 0.0930 | 0.0731 | 0.0955 | 0.0827 | 0.0979 | 0.0581 | 0.0509 | 0.0509 |
| | 25 | 0.0725 | 0.0614 | 0.0695 | 0.0659 | 0.0731 | 0.0586 | 0.0519 | 0.0522 |
| | 50 | 0.0592 | 0.0544 | 0.0584 | 0.0555 | 0.0593 | 0.0505 | 0.0531 | 0.0519 |

*Table 24:* An $\alpha$-level simulation. Design: balanced with $n_i = 7, 10, 15, 25, 50$ observations. Factor levels: $a = 3$. Number of covariates: $d = 1, 2, 5$. Covariates: Exp(1) distributed. Dependent variable: unweighted sum of covariates with standard normal error term.

| $d$ | $n_i$ | mctp.n | mctp.t | 2-step | sie.ats | sie.wts | mctp.nc | mctp.par | f.par |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 7 | 0.1072 | 0.0575 | 0.0694 | 0.0759 | 0.1135 | 0.0691 | 0.0491 | 0.0492 |
| | 10 | 0.0887 | 0.0570 | 0.0604 | 0.0677 | 0.0947 | 0.0630 | 0.0507 | 0.0505 |
| | 15 | 0.0693 | 0.0517 | 0.0530 | 0.0582 | 0.0734 | 0.0564 | 0.0492 | 0.0480 |
| | 25 | 0.0609 | 0.0516 | 0.0532 | 0.0568 | 0.0628 | 0.0567 | 0.0503 | 0.0490 |
| | 50 | 0.0556 | 0.0500 | 0.0490 | 0.0531 | 0.0562 | 0.0524 | 0.0497 | 0.0493 |
| 2 | 7 | 0.1201 | 0.0648 | 0.0908 | 0.0903 | 0.1272 | 0.0715 | 0.0491 | 0.0491 |
| | 10 | 0.0928 | 0.0604 | 0.0711 | 0.0774 | 0.0974 | 0.0649 | 0.0499 | 0.0507 |
| | 15 | 0.0767 | 0.0554 | 0.0628 | 0.0635 | 0.0769 | 0.0601 | 0.0495 | 0.0497 |
| | 25 | 0.0656 | 0.0562 | 0.0586 | 0.0609 | 0.0673 | 0.0616 | 0.0503 | 0.0493 |
| | 50 | 0.0576 | 0.0518 | 0.0542 | 0.0536 | 0.0567 | 0.0577 | 0.0508 | 0.0520 |
| 5 | 7 | 0.1897 | 0.1161 | 0.1880 | 0.1641 | 0.1940 | 0.0749 | 0.0519 | 0.0521 |
| | 10 | 0.1323 | 0.0900 | 0.1221 | 0.1116 | 0.1339 | 0.0633 | 0.0509 | 0.0502 |
| | 15 | 0.0972 | 0.0760 | 0.0916 | 0.0893 | 0.1008 | 0.0572 | 0.0487 | 0.0493 |
| | 25 | 0.0705 | 0.0578 | 0.0677 | 0.0638 | 0.0700 | 0.0548 | 0.0491 | 0.0501 |
| | 50 | 0.0640 | 0.0600 | 0.0616 | 0.0597 | 0.0654 | 0.0514 | 0.0507 | 0.0522 |

*Table 25:* An $\alpha$-level simulation. Design: balanced with $n_i = 7, 10, 15, 25, 50$ observations. Factor levels: $a = 3$. Number of covariates: $d = 1, 2, 5$. Covariates: Bernoulli(0.5) distributed. Dependent variable: unweighted sum of covariates with standard normal error term.

| $d$ | $n_i$ | mctp.n | mctp.t | 2-step | sie.ats | sie.wts | mctp.nc | mctp.par | f.par |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 7 | 0.1320 | 0.0809 | 0.0727 | 0.0965 | 0.1371 | 0.0750 | 0.0531 | 0.0526 |
| | 10 | 0.0981 | 0.0671 | 0.0594 | 0.0753 | 0.1032 | 0.0643 | 0.0476 | 0.0478 |
| | 15 | 0.0785 | 0.0611 | 0.0542 | 0.0637 | 0.0811 | 0.0596 | 0.0474 | 0.0495 |
| | 25 | 0.0678 | 0.0559 | 0.0542 | 0.0584 | 0.0710 | 0.0541 | 0.0488 | 0.0495 |
| | 50 | 0.0573 | 0.0529 | 0.0517 | 0.0535 | 0.0585 | 0.0546 | 0.0495 | 0.0512 |
| 2 | 7 | 0.1447 | 0.0856 | 0.0933 | 0.1128 | 0.1482 | 0.0701 | 0.0548 | 0.0535 |
| | 10 | 0.1061 | 0.0712 | 0.0710 | 0.0832 | 0.1102 | 0.0663 | 0.0470 | 0.0468 |
| | 15 | 0.0820 | 0.0636 | 0.0608 | 0.0685 | 0.0853 | 0.0583 | 0.0488 | 0.0482 |
| | 25 | 0.0684 | 0.0572 | 0.0552 | 0.0615 | 0.0701 | 0.0564 | 0.0527 | 0.0517 |
| | 50 | 0.0571 | 0.0527 | 0.0504 | 0.0548 | 0.0588 | 0.0532 | 0.0518 | 0.0509 |
| 5 | 7 | 0.2333 | 0.1497 | 0.1887 | 0.1983 | 0.2390 | 0.0699 | 0.0491 | 0.0498 |
| | 10 | 0.1491 | 0.1048 | 0.1189 | 0.1308 | 0.1546 | 0.0661 | 0.0514 | 0.0504 |
| | 15 | 0.1115 | 0.0874 | 0.0941 | 0.0992 | 0.1171 | 0.0625 | 0.0516 | 0.0500 |
| | 25 | 0.0818 | 0.0708 | 0.0685 | 0.0752 | 0.0826 | 0.0581 | 0.0505 | 0.0486 |
| | 50 | 0.0659 | 0.0596 | 0.0597 | 0.0639 | 0.0673 | 0.0533 | 0.0523 | 0.0544 |

*Table 26:* An $\alpha$-level simulation. Design: balanced with $n_i = 7, 10, 15, 25, 50$ observations. Factor levels: $a = 3$. Number of covariates: $d = 1, 2, 5$. Covariates: $\text{Chi}^2(1)$ distributed. Dependent variable: unweighted sum of covariates with standard normal error term.

| $d$ | $n_i$ | mctp.n | mctp.t | 2-step | sie.ats | sie.wts | mctp.nc | mctp.par | f.par |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 7 | 0.1060 | 0.0587 | 0.0737 | 0.0782 | 0.1150 | 0.0698 | 0.0521 | 0.0527 |
| | 10 | 0.0880 | 0.0556 | 0.0611 | 0.0677 | 0.0933 | 0.0622 | 0.0499 | 0.0511 |
| | 15 | 0.0701 | 0.0526 | 0.0552 | 0.0573 | 0.0726 | 0.0612 | 0.0556 | 0.0555 |
| | 25 | 0.0600 | 0.0510 | 0.0529 | 0.0554 | 0.0617 | 0.0552 | 0.0520 | 0.0526 |
| | 50 | 0.0560 | 0.0518 | 0.0524 | 0.0531 | 0.0559 | 0.0521 | 0.0493 | 0.0476 |
| 2 | 7 | 0.1098 | 0.0583 | 0.0855 | 0.0835 | 0.1159 | 0.0688 | 0.0447 | 0.0450 |
| | 10 | 0.0986 | 0.0630 | 0.0770 | 0.0800 | 0.1027 | 0.0681 | 0.0516 | 0.0516 |
| | 15 | 0.0766 | 0.0579 | 0.0617 | 0.0666 | 0.0769 | 0.0613 | 0.0496 | 0.0483 |
| | 25 | 0.0605 | 0.0517 | 0.0573 | 0.0564 | 0.0642 | 0.0541 | 0.0491 | 0.0469 |
| | 50 | 0.0623 | 0.0583 | 0.0565 | 0.0592 | 0.0632 | 0.0579 | 0.0527 | 0.0524 |
| 5 | 7 | 0.1934 | 0.1227 | 0.1821 | 0.1653 | 0.2016 | 0.0687 | 0.0503 | 0.0493 |
| | 10 | 0.1385 | 0.0945 | 0.1246 | 0.1182 | 0.1435 | 0.0652 | 0.0506 | 0.0515 |
| | 15 | 0.0991 | 0.0753 | 0.0914 | 0.0890 | 0.1010 | 0.0590 | 0.0511 | 0.0510 |
| | 25 | 0.0754 | 0.0621 | 0.0706 | 0.0698 | 0.0777 | 0.0498 | 0.0470 | 0.0452 |
| | 50 | 0.0626 | 0.0587 | 0.0601 | 0.0613 | 0.0653 | 0.0491 | 0.0505 | 0.0505 |

*Table 27:* An $\alpha$-level simulation. Design: unbalanced with $n_1 = 10$, $n_2 = 15$, $n_3 = 20$ (*low*) and $n_1 = 10$, $n_2 = 50$, $n_3 = 100$ (*high*) observations. Factor levels: $a = 3$. Number of covariates: $d = 1, 2, 5$. Covariates: multivariate log-normal distributed (4.5), $\rho = 0$. Dependent variable: unweighted sum of covariates (4.6). The degree of imbalance is abbreviated by *Imba*.

| $d$ | Imba | mctp.n | mctp.t | 2-step | sie.ats | sie.wts | mctp.nc | mctp.par | f.par |
|---|---|---|---|---|---|---|---|---|---|
| 1 | low | 0.0769 | 0.0534 | 0.0551 | 0.0661 | 0.0799 | 0.0549 | 0.0379 | 0.0385 |
| | high | 0.0752 | 0.0618 | 0.0378 | 0.0763 | 0.0754 | 0.0430 | 0.0528 | 0.0517 |
| 2 | low | 0.0811 | 0.0579 | 0.0643 | 0.0675 | 0.0839 | 0.0574 | 0.0388 | 0.0392 |
| | high | 0.0828 | 0.0705 | 0.0452 | 0.0847 | 0.0841 | 0.0444 | 0.0530 | 0.0525 |
| 5 | low | 0.1041 | 0.0756 | 0.0845 | 0.0891 | 0.1088 | 0.0581 | 0.0411 | 0.0401 |
| | high | 0.0837 | 0.0696 | 0.0408 | 0.0824 | 0.0834 | 0.0438 | 0.0528 | 0.0519 |

*Table 28:* An $\alpha$-level simulation. Design: unbalanced with $n_1 = 10$, $n_2 = 15$, $n_3 = 20$ (*low*) and $n_1 = 10$, $n_2 = 50$, $n_3 = 100$ (*high*) observations. Factor levels: $a = 3$. Number of covariates: $d = 1, 2, 5$. Covariates: binomial distributed (4.7). Dependent variable: unweighted sum of covariates (4.8). The degree of imbalance is abbreviated by *Imba*.

| $d$ | Imba | mctp.n | mctp.t | 2-step | sie.ats | sie.wts | mctp.nc | mctp.par | f.par |
|---|---|---|---|---|---|---|---|---|---|
| 1 | low | 0.0758 | 0.0529 | 0.0526 | 0.0660 | 0.0790 | 0.0553 | 0.0494 | 0.0492 |
| | high | 0.0744 | 0.0618 | 0.0423 | 0.0705 | 0.0734 | 0.0487 | 0.0485 | 0.0511 |
| 2 | low | 0.0764 | 0.0559 | 0.0619 | 0.0656 | 0.0780 | 0.0588 | 0.0497 | 0.0493 |
| | high | 0.0725 | 0.0596 | 0.0428 | 0.0746 | 0.0703 | 0.0444 | 0.0505 | 0.0510 |
| 5 | low | 0.0859 | 0.0636 | 0.0824 | 0.0758 | 0.0870 | 0.0562 | 0.0483 | 0.0483 |
| | high | 0.0745 | 0.0612 | 0.0451 | 0.0756 | 0.0743 | 0.0487 | 0.0495 | 0.0520 |

### 7.2.2 Power Simulation

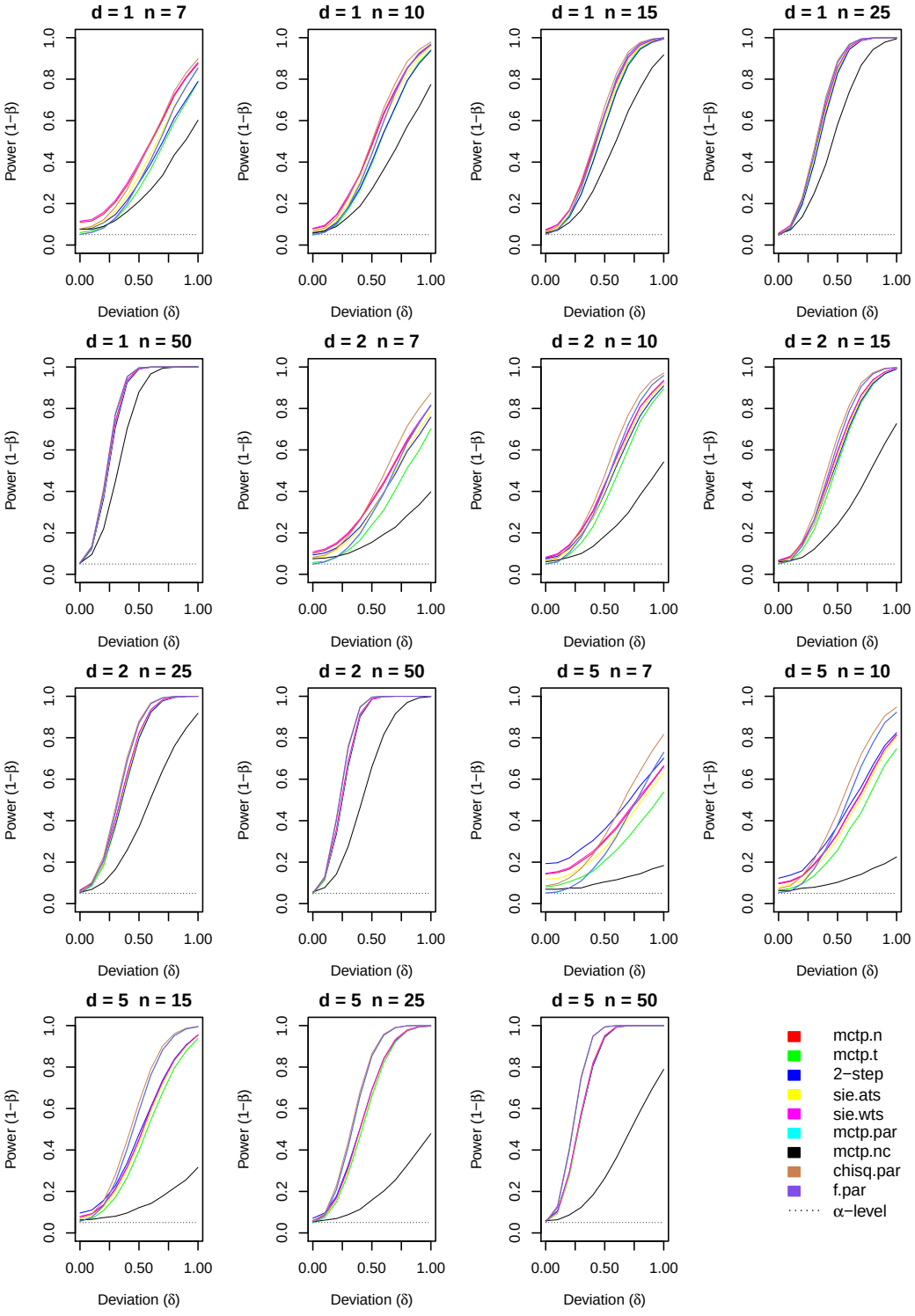


*Figure 11:* Simulation results for the power $(1 - \beta)$. Design: balanced with $n_i = 7, 10, 15, 25, 50$ observations. Factor levels: $a = 3$. Number of covariates: $d = 1, 2, 5$. Covariates: multivariate normal distributed (4.1), $\rho = 0.2$. Dependent variable: unweighted sum of covariates (4.9).

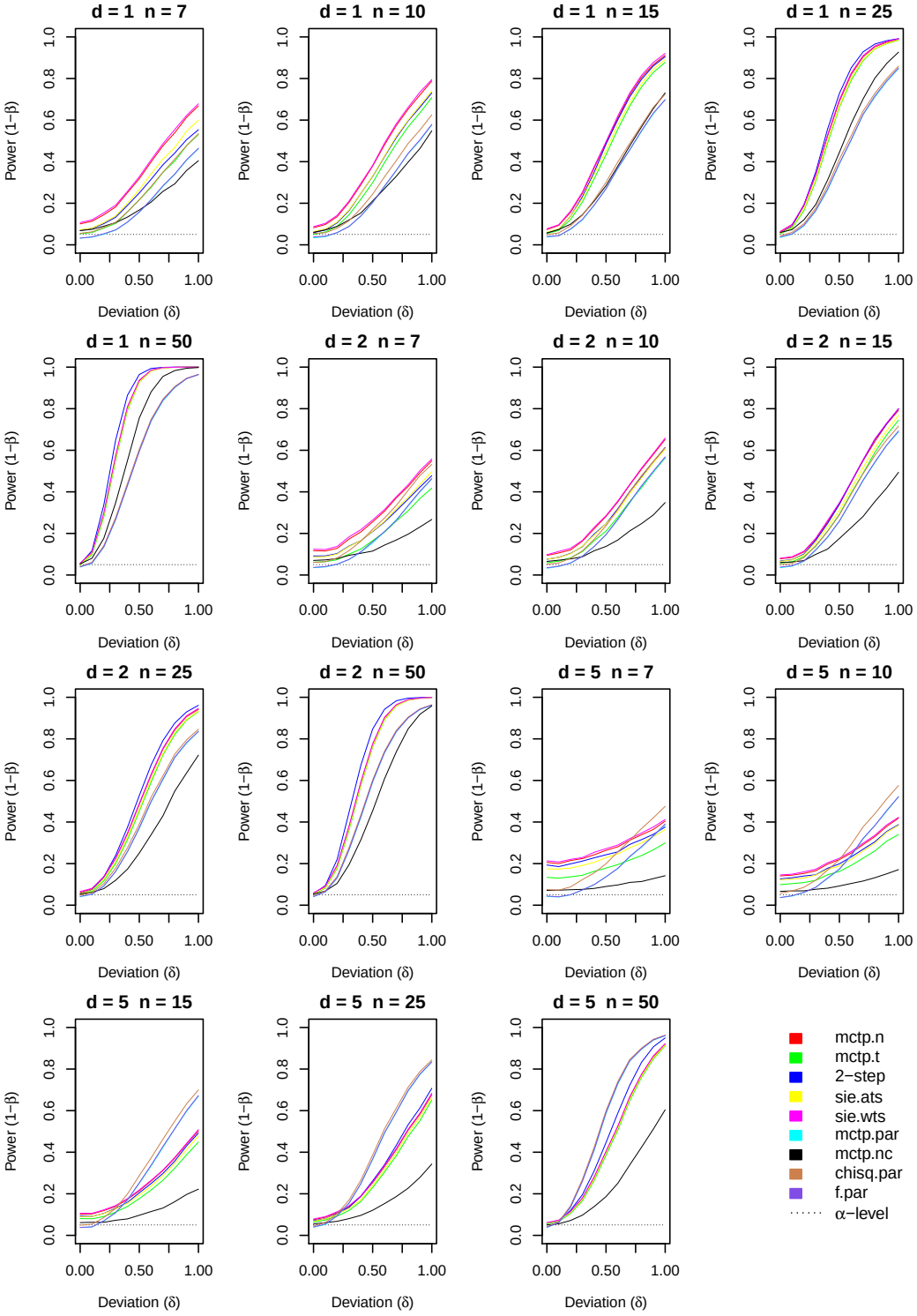


*Figure 12:* Simulation results for the power $(1-\beta)$. Design: balanced with $n_i = 7, 10, 15, 25, 50$ observations. Factor levels: $a = 3$. Number of covariates: $d = 1, 2, 5$. Covariates: multivariate log-normal distributed (4.5), $\rho = 0.2$. Dependent variable: unweighted sum of covariates (4.10).

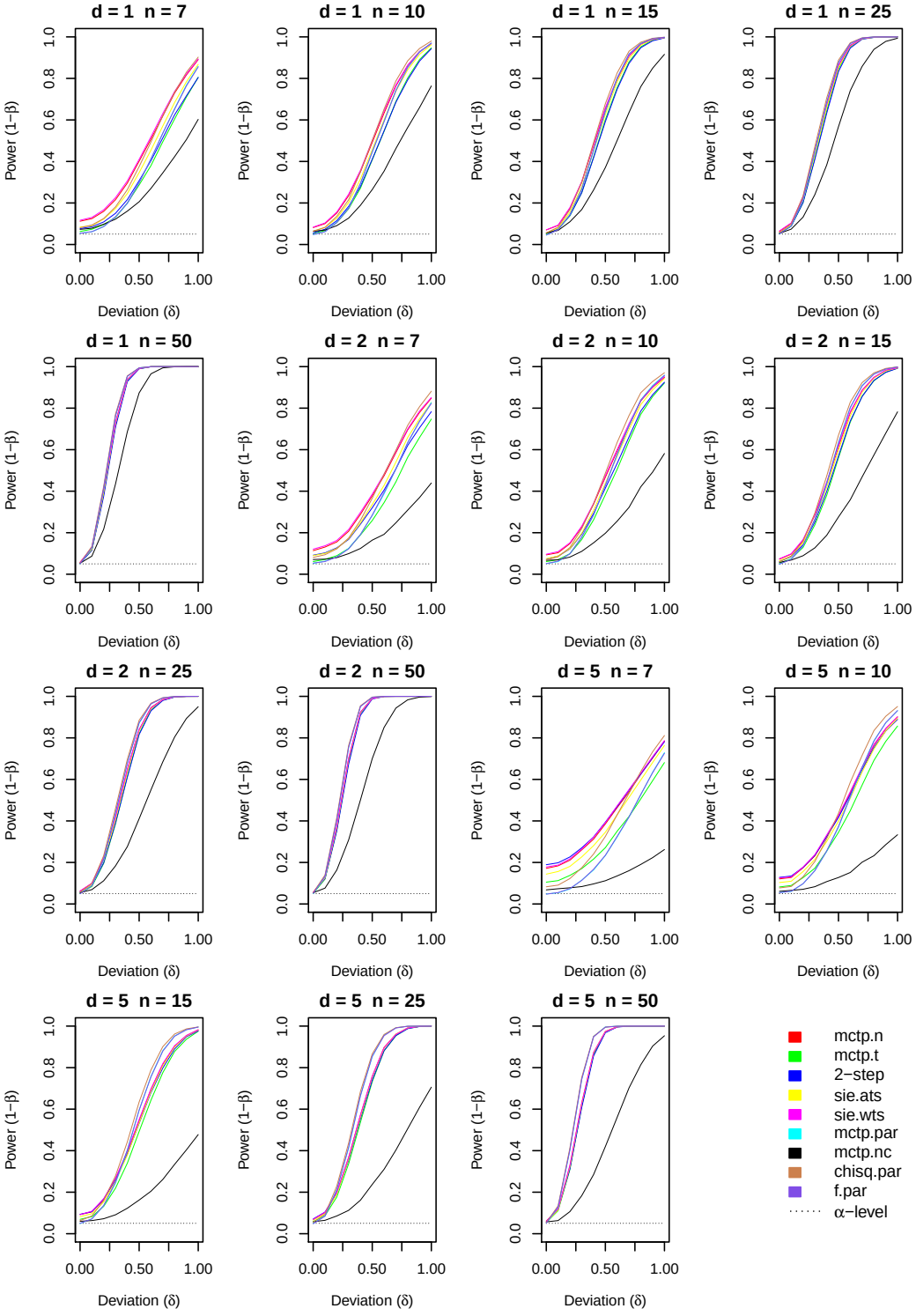


*Figure 13:* Simulation results for the power $(1-\beta)$. Design: balanced with $n_i = 7, 10, 15, 25, 50$ observations. Factor levels: $a = 3$. Number of covariates: $d = 1, 2, 5$. Covariates: binomial distributed (4.7). Dependent variable: unweighted sum of covariates (4.11).

## 7.3 R-Code

In this section we present the R-script written for the user, which was used to evaluate the example in Section 5. The R-code for the simulations and graphics is not listed in this thesis. If you are interested in the R-code for the simulations and graphics, please write me an email: thomas.asendorf@stud.uni-goettingen.de.

```
mctp.cov<-function(formula,data,type=c("Tukey", "Dunnett", "Sequen",
    "Williams", "Changepoint", "AVE", "McDermott", "Marcus",
    "UmbrellaWilliams"),conf.level=0.95,alternative=c("two.sided",
    "less", "greater"),asy.method=c("mult.t","normal"),info=T,rounds=3,
    effect=c("unweighted","weighted")){

input.list<-list(formula=formula,data=data,type=type[1],
        conf.level=conf.level,alternative=alternative[1],
        asy.method=asy.method[1],info=info,rounds=rounds,
        effect=effect[1])

#-----------------------------------------------------------
#Loading required packages
#-----------------------------------------------------------
require(mvtnorm,quietly=T)
require(MASS,quietly=T)
require(nparcomp,quietly=T)
#-----------------------------------------------------------

#-----------------------------------------------------------
#Definition of custom functions
#-----------------------------------------------------------
P<-function(n){
        diag(1,abs(n))-1/n
}

'%+%'<-function(A,B){
        a.1<-nrow(A)
        a.2<-ncol(A)
        b.1<-nrow(B)
        b.2<-ncol(B)

        cbind(rbind(A,matrix(0,ncol=a.2,nrow=b.1)),
        rbind(matrix(0,ncol=b.2,nrow=a.1),B))
}

tr<-function(A){
        sum(diag(A))
}
#-----------------------------------------------------------

#-----------------------------------------------------------
#Possible input errors
#-----------------------------------------------------------
if(conf.level >=1 || conf.level <=0){
        stop("The confidence level must be between 0 and 1")
}
if(length(formula)!=3){
```

```
49          stop("You can only analyse one-way layouts")
50  }
51  type<-match.arg(type)
52  alternative<-match.arg(alternative)
53  asy.method<-match.arg(asy.method)
54  effect<-match.arg(effect)
55  #-------------------------------------------------------------
56
57  #-------------------------------------------------------------
58  #Preparing the data
59  #-------------------------------------------------------------
60  dat<-model.frame(formula,data)
61  data.sav<-data
62
63  if(ncol(dat)==2){
64          return(mctp(formula=formula,data=data,type=type,
65          conf.level=conf.level,
66                  alternative=alternative,asy.method=asy.method,
67                  info=info,rounds=rounds,effect=effect))
68  }
69
70  resp.covs<-dat[,-2]
71  factorx<-as.factor(dat[,2])
72  samples<-split(resp.covs,factorx)
73
74  a<-nlevels(factorx)
75  d<-ncol(resp.covs)
76  n<-sapply(samples,nrow)
77  N<-sum(n)
78  contrast<-type
79  pseudo<-(effect=="unweighted")
80
81  if(any(n<=1)){
82          stop("At least one group has only one observation")
83  }
84
85  data<-as.matrix(samples[[1]])
86  for(i in 2:a){
87          data<-rbind(data,as.matrix(samples[[i]]))
88  }
89  #-------------------------------------------------------------
90
91  #-------------------------------------------------------------
92  #Analyzing the data
93  #-------------------------------------------------------------
94
95  #-------------------------------------------------------------
96  #Calculating inner Ranks
97  #Fs[i,,r] contains F_i^r(X_11^r),..,F_i^r(X_an_a^r)
98  Fs<-array(0,dim=c(a,N,d))
99
100 for(i in 1:a){
101 for(r in 1:d){
102         h<-t(matrix(1,ncol=N,nrow=n[i])*
```

```
                data[(sum(n[0:(i-1)])+1):sum(n[1:i]),r])
        Fs[i,,r]<-rowSums((data[,r]>h)+1/2*(data[,r]==h))*1/n[i]
}}
#-----------------------------------------------------------
#Estimating the p's
M.1<-t(as.matrix(rep(1,n[1]))/n[1])
for(i in 2:a){
        M.1<-M.1%+%t(as.matrix(rep(1,n[i]))/n[i])
}

#p.pairs[i,j,r] contains p_ij^r
p.pairs<-array(0,dim=c(a,a,d))
for(i in 1:a){
        p.pairs[i,,]<-M.1%*%Fs[i,,]
}

p.hat<-array(0,dim=c(a,d))
if(pseudo==T){
        for(r in 1:d){
                p.hat[,r]<-colMeans(p.pairs[,,r])
        }
}else{
        for(r in 1:d){
                p.hat[,r]<-t(n/N)%*%p.pairs[,,r]
        }
}

#-----------------------------------------------------------
#Estimating the gamma's
y.hat<-rep(0,N)

#data.r[j,r] contains Y_ik^r where ik are chosen accordingly
data.r<-array(0,dim=c(N,d))

if(pseudo==T){
        for(r in 1:d){
                data.r[,r]<-t(rep(1/a,a))%*%Fs[,,r]
        }
}else{
        for(r in 1:d){
                data.r[,r]<-t(n/N)%*%Fs[,,r]
        }
}

D<-P(n[1])
for(i in 2:a){
        D<-D%+%P(n[i])
}

y.hat<-D%*%data.r[,1]

X.hat<-D%*%data.r[,-1]
gamma.hat<-solve(t(X.hat)%*%X.hat)%*%t(X.hat)%*%y.hat
gamma<-c(-1,gamma.hat)
```

```
157
158 #-------------------------------------------------------------
159 #Estimating p*
160 p.star<-p.hat%*%gamma
161 #-------------------------------------------------------------
162 #Estimation of VN
163 theta<-function(i,j,s,r,u,Fs,n){
164 n.j<-c(0,n)
165 1/(n[j]-1)*(Fs[i,(sum(n.j[1:j])+1):sum(n.j[1:(j+1)]),r]%*%
166         P(n.j[j+1])%*%Fs[s,(sum(n.j[1:j])+1):sum(n.j[1:(j+1)]),u])
167 }
168
169 Sigma<-matrix(0,ncol=a^2*d,nrow=a^2*d)
170
171 for(t in 1:a){
172 for(u in 1:d){
173 for(s in 1:a){
174 for(j in 1:a){
175 for(r in 1:d){
176 for(i in 1:a){
177         if(i==j | t==s){
178                 Sigma[a*d*(j-1)+a*(r-1)+i,a*d*(t-1)+a*(u-1)+s]<-0
179                 }
180                 if((i!=j & t!=s) & (j==t & i==s)){
181                 Sigma[a*d*(j-1)+a*(r-1)+i,a*d*(t-1)+a*(u-1)+s]<-1/n[j]*
182                 theta(i,j,s,r,u,Fs,n)+1/n[i]*theta(j,i,t,r,u,Fs,n)
183                 }
184                 if((i!=j & t!=s) & (j==s & i==t)){
185                 Sigma[a*d*(j-1)+a*(r-1)+i,a*d*(t-1)+a*(u-1)+s]<-
186                 -1/n[j]*theta(i,j,t,r,u,Fs,n)-1/n[i]*
187                 theta(j,i,s,r,u,Fs,n)
188                 }
189                 if((i!=j & t!=s) & (j!=s & j!=t & i==t)){
190                 Sigma[a*d*(j-1)+a*(r-1)+i,a*d*(t-1)+a*(u-1)+s]<-
191                 -1/n[i]*theta(j,i,s,r,u,Fs,n)
192                 }
193                 if((i!=j & t!=s) & (j!=s & j!=t & i==s)){
194                 Sigma[a*d*(j-1)+a*(r-1)+i,a*d*(t-1)+a*(u-1)+s]<-
195                 1/n[i]*theta(j,i,t,r,u,Fs,n)
196                 }
197                 if((i!=j & t!=s) & (j==t & i!=s & i!=t)){
198                 Sigma[a*d*(j-1)+a*(r-1)+i,a*d*(t-1)+a*(u-1)+s]<-
199                 1/n[j]*theta(i,j,s,r,u,Fs,n)
200                 }
201                 if((i!=j & t!=s) & (j==s & i!=s & i!=t)){
202                 Sigma[a*d*(j-1)+a*(r-1)+i,a*d*(t-1)+a*(u-1)+s]<-
203                 -1/n[j]*theta(i,j,t,r,u,Fs,n)
204                 }
205                 if((i!=j & t!=s) & (j!=s & j!=t & i!=s & i!=t)){
206                 Sigma[a*d*(j-1)+a*(r-1)+i,a*d*(t-1)+a*(u-1)+s]<-0
207                 }
208 }}}}}}
209
210 if(pseudo==T){
```

```
211         W<-diag(a*d)%x%t(rep(1/a,a))
212 }else{
213         W<-diag(a*d)%x%t(n/N)
214 }
215         VN<-N*(diag(a)%x%t(gamma))%*%W%*%Sigma%*%
216                 t(W)%*%(diag(a)%x%gamma)
217
218 #-----------------------------------------------------------------
219 #Making of the contrast matrix
220
221 if(is.matrix(contrast)){
222         CM<-contrast
223 }else{
224         CM<-contrMat(n,contrast)
225 }
226
227 q<-nrow(CM)
228
229 #-----------------------------------------------------------------
230 #MCTP with normal distribution
231 if(asy.method=="normal"){
232
233 CV<-CM%*%VN%*%t(CM)
234 R.hat<-cov2cor(CV)
235
236 switch(alternative,
237         "two.sided"={
238
239         quantile<-qmvnorm(conf.level,tail="both",corr=R.hat)$quantile
240
241         upper.SKI<-CM%*%p.star+quantile*sqrt(diag(CV)/N)
242         lower.SKI<-CM%*%p.star-quantile*sqrt(diag(CV)/N)
243         T.mctp<-(CM%*%p.star)/sqrt(diag(CV)/N)
244
245         mctp.pvalue<-rep(0,nrow(CM))
246         for(i in 1:nrow(CM)){
247         mctp.pvalue[i]<-round(1-pmvnorm(lower=-abs(T.mctp[i]),
248                 upper=abs(T.mctp[i]),mean=rep(0,nrow(CM)),
249                 corr=R.hat)[1],7)
250         }
251         },
252         "less"={
253
254         quantile<-qmvnorm(conf.level,tail="lower.tail",
255                 corr=R.hat)$quantile
256
257         upper.SKI<-Inf
258         lower.SKI<-CM%*%p.star-quantile*sqrt(diag(CV)/N)
259         T.mctp<-(CM%*%p.star)/sqrt(diag(CV)/N)
260
261         mctp.pvalue<-rep(0,nrow(CM))
262         for(i in 1:nrow(CM)){
263                 mctp.pvalue[i]<-round(pmvnorm(lower=T.mctp[i],
264                         upper=Inf,mean=rep(0,nrow(CM)),
```

```
                        corr=R.hat)[1],7)
        }

        },
        "greater"={

        quantile<-qmvnorm(conf.level,tail="lower.tail",
                corr=R.hat)$quantile

        upper.SKI<-CM%*%p.star+quantile*sqrt(diag(CV)/N)
        lower.SKI<--Inf
        T.mctp<-(CM%*%p.star)/sqrt(diag(CV)/N)

        mctp.pvalue<-rep(0,nrow(CM))
        for(i in 1:nrow(CM)){
                mctp.pvalue[i]<-round(pmvnorm(lower=-Inf,
                        upper=T.mctp[i],mean=rep(0,nrow(CM)),
                        corr=R.hat)[1],7)
        }

}
)

h.mctp<-data.frame(row.names=row.names(CM),Estimator=CM%*%p.star,
        Lower=lower.SKI,Upper=upper.SKI,Statistic=T.mctp,
        p.Value=mctp.pvalue)
}

#-----------------------------------------------------------------
#MCTP with student-t distribution
if(asy.method=="mult.t"){

CV<-CM%*%VN%*%t(CM)
R.hat<-cov2cor(CV)

eta<-array(0,dim=c(q,a))
nu<-rep(0,q)

for(l in 1:nrow(CM)){
for(i in 1:a){
        if(pseudo==T){
                W<-matrix(1/a,ncol=d,nrow=a*n[i])
        }else{
                W<-matrix(n/N,ncol=d,nrow=a*n[i])
        }
        h1<-matrix(as.vector(Fs[,(sum(n[0:(i-1)])+1):
                sum(n[1:i]),]),ncol=d,nrow=a*n[i])*W
        h2<-diag(-1,a);h2[i,]<-1;h2[i,i]<-0
        A.li<-(diag(1,n[i])%x%t(CM[l,]))%*%(diag(1,n[i])%x%h2)%*%h1
        eta[l,i]<-t(gamma)%*%t(A.li)%*%P(n[i])%*%A.li%*%gamma
}
nu[l]<-sum(eta[l,]/n)^2/sum(eta[l,]^2/(n^2*(n-1)))
}

```

```
319
320 ft<-round(max(1,min(nu)))
321
322 switch(alternative,
323          "two.sided"={
324          quantile<-qmvt(conf.level,df=ft,tail="both",corr=R.hat)$quantile
325
326          upper.SKI.t<-CM%*%p.star+quantile*sqrt(diag(CV)/N)
327          lower.SKI.t<-CM%*%p.star-quantile*sqrt(diag(CV)/N)
328          T.mctp.t<-(CM%*%p.star)/sqrt(diag(CV)/N)
329
330          mctp.pvalue.t<-rep(0,nrow(CM))
331          for(i in 1:nrow(CM)){
332          mctp.pvalue.t[i]<-round(1-pmvt(lower=
333                   as.numeric(-abs(T.mctp.t[i])),
334                   upper=as.numeric(abs(T.mctp.t[i])),
335                   df=ft,delta=rep(0,q),corr=R.hat)[1],7)
336          }
337 },
338          "greater"={
339          quantile<-qmvt(conf.level,df=ft,tail="lower.tail",
340                   corr=R.hat)$quantile
341
342          upper.SKI.t<-CM%*%p.star+quantile*sqrt(diag(CV)/N)
343          lower.SKI.t<--Inf
344          T.mctp.t<-(CM%*%p.star)/sqrt(diag(CV)/N)
345
346          mctp.pvalue.t<-rep(0,nrow(CM))
347          for(i in 1:nrow(CM)){
348                   mctp.pvalue.t[i]<-round(pmvt(lower=-Inf,
349                            upper=as.numeric(T.mctp.t[i]),df=ft,
350                            delta=rep(0,q),corr=R.hat)[1],7)
351          }
352 },
353          "less"={
354          quantile<-qmvt(conf.level,df=ft,tail="lower.tail",
355                   corr=R.hat)$quantile
356
357          upper.SKI.t<-Inf
358          lower.SKI.t<-CM%*%p.star-quantile*sqrt(diag(CV)/N)
359          T.mctp.t<-(CM%*%p.star)/sqrt(diag(CV)/N)
360
361          mctp.pvalue.t<-rep(0,nrow(CM))
362          for(i in 1:nrow(CM)){
363                   mctp.pvalue.t[i]<-round(pmvt(lower=
364                            as.numeric(T.mctp.t[i]),upper=Inf,
365                            df=ft,delta=rep(0,q),corr=R.hat)[1],7)
366          }
367 }
368 )
369
370 h.mctp<-data.frame(row.names=row.names(CM),Estimator=CM%*%p.star,
371          Lower=lower.SKI.t,Upper=upper.SKI.t,Statistic=T.mctp.t,
372          p.Value=mctp.pvalue.t)
```

```
}

#------------------------------------------------------------------
#Writing the ouput


switch(as.character(pseudo),
         "TRUE"={est.method<-"Global pseudo ranks"},
         "FALSE"={est.method<-"Global ranks"}
)
switch(asy.method,
         "normal"={asy.text<-"Multivariate normal approximation"},
         "mult.t"={asy.text<-paste("Multivariate t-approximation with",
         ft," df",sep="")}
)
switch(alternative,
         "two.sided"={alt.text<-"Contrasts of adj. relative treatment
effects are unequal 0"},
          "less"={alt.text<-"Contrasts of adj. relative treatment
effects are less than 0"},
          "greater"={alt.text<-"Contrasts of adj. relative treatment
effects are greater than 0"}
)
group.info<-data.frame(Group=levels(factorx),Observations=n)
rownames(group.info)<-NULL

if(info==T){
         cat("\n", "#-----------------Nonparametric Multiple Comparisons
for relative effects -----------------#",
                           "\n", "\n", "-", "Null Hypothesis:",alt.text,
                  "\n", "-", "Estimation Method:",est.method,
                           "\n", "-", "Type of Contrast:", type, "\n", "-",
                           "Confidence Level:", conf.level * 100, "%", "\n",
                           "-", "Method:", asy.text, "\n", "\n",
"#-----------------------------------
-------------------------------------------------------------#",
                           "\n", "\n")
         print(group.info)
         cat("\n","Number of covariates used: ")
         cat(d-1," (",sep="")
         cat(colnames(dat)[3:(d+1)])
         cat(") ","\n",sep="")
         cat("\n",
"#-----------------------------------
-------------------------------------------------------------#",
                           "\n", "\n")
         print(round(h.mctp,rounds))
                  cat("\n",
"#-----------------------------------
-------------------------------------------------------------#",)
         cat("\n","\n")
         return(list(Data.Info=group.info,Contrast=CM,
Analysis=h.mctp,gamma=gamma,formula=formula,data=data.sav,
                  Text.Output=list(null.hyp=alt.text,
```

```
427 est.method=est.method,
428 contr.type=type,conf.level=conf.level,method=asy.text)))
429         }else{
430         return(list(Data.Info=group.info,Contrast=CM,Analysis=h.mctp,
431 gamma=gamma,formula=formula,data=data.sav,
432                 Text.Output=list(null.hyp=alt.text,
433 est.method=est.method,
434 contr.type=type,conf.level=conf.level,method=asy.text)))
435 }
436 }
```

# References

M. G. Akritas, S. F. Arnold, and E. Brunner. Nonparametric hypothesis and rank statistics for unbalanced factorial designs. *Journal of the American Statistical Association*, 92: 258–265, 1997.

A. Bathke. *Einfluß von Kovariablen in nichtparametrischen Modellen.* Master's Thesis, Georg-August-University Göttingen, 1998.

Arne Bathke and Edgar Brunner. *Recent advances and trends in nonparametric statistics*, chapter A Nonparametric Alternative to Analysis of Covariance, pages 109–120. Tba, 2003.

G.E.P. Box. Some theorems on quadratic forms applied in the study of analysis of variance problems. I. Effect of inequality of variance in the one-way classification. *Ann. Math. Stat.*, 25:290–302, 1954. ISSN 0003-4851. doi: 10.1214/aoms/1177728786.

F. Bretz, A. Genz, and L. A. Hothorn. On the numerical availability of multiple comparison procedures. *Biometrical Journal*, 43(5):645–656, 2001.

E. Brunner and U. Munzel. *Nichtparametrische Datenanalyse, 2nd Ed.* Springer-Verlag, Heidelberg, 2013.

E. Brunner and M. L. Puri. Nonparametric methods in factorial designs. *Statistical Papers*, 42(1):1–52, 2001.

O. Christophliemk. *Die Transformationsmethode im gemischten nichtparametrischen Modell mit Kovariablen.* Master's Thesis, Georg-August-University Göttingen, 2001.

H. Cramer and H. Wold. Some theorems on distribution functions. *J. Lond. Math. Soc.*, 11: 290–294, 1936. ISSN 0024-6107; 1469-7750/e. doi: 10.1112/jlms/s1-11.4.290.

S. Domhof. *Nichtparametrische relative Effekte.* Dissertation, Georg-August-University Göttingen, 2001.

C. W. Dunnett. A multiple comparison procedure for comparing several treatments with a control. *Journal of the American Statistical Association*, 50(272):1096–1121, 1955.

K. R. Gabriel. Simultaneous test procedures - some theory of multiple comparisons. *Annals of Mathematical Statistics*, 40:224–250, 1969.

X. Gao, M. Alvo, J. Chen, and G. Li. Nonparametric multiple comparison procedures for unbalanced one-way factorial designs. *Journal of Statistical Planning and Inference*, 138 (8):2574–2591, 2008.

A. Genz and F. Bretz. *Computation of Multivariate Normal and t Probabilities.* Springer-Verlag, New York, 2009.

B. V. Gnedenko. *The Theory of Probability.* Chelsea Publishing, New York, 2nd edition, 1962.

Sture Holm. A simple sequentially rejective multiple test procedure. *Scand. J. Stat.*, 6: 65–70, 1979. ISSN 0303-6898; 1467-9469/e.

T. Hothorn, F. Bretz, and P. Westfall. Simultaneous inference in general parametric models. *Biometrical Journal*, 50(3):346–363, June 2008.

B. Huitema. *The analysis of covariance and alternatives.* Wiley, New York, 1980.

R. Kirk. *Experimental Design:Procedures for the Behavioral Sciences.* Brooks/Cole, California, 1982.

F. Konietschke, L. A. Hothorn, and E. Brunner. Rank-based multiple test procedures and simultaneous confidence intervals. *Electronic Journal of Statistics*, 6:1–8, 2012a.

F. Konietschke, O. Libiger, and L. Hothorn. Nonparametric evaluation of quantitative traits in population-based association studies when the genetic model is unknown. *PLoS ONE*, 7(2), February 2012b.

W.H. Kruskal and W.A. Wallis. Use of ranks in one-criterion variance analysis. *J. Amer. Stat. Ass.*, 47:583–621, 1952.

F. Langer. *Berücksichtigung von Kovariablen im nichtparametrischen gemischten Modell.* Dissertation, Georg-August-University Göttingen, 1998.

M. Loève. *Probability Theory I.* Springer Verlag, Berlin, 1977.

P. Lévy. *Calcul des probabilitées.* Gauthiers-Villars, Paris, 1925.

H.B. Mann and D.R. Whitney. On a test of whether one of two random variables is stochastically larger then the other. *Annals of Mathematical Statistics*, 18(2):50–60, June 1947.

G.A. Miller and J.P. Chapman. Misunderstanding analysis of covariance. *American Psychological Association*, 110(1):40–48, 2001.

U. Munzel. Linear rank score statistics when ties are present. *Statistics and Probability Letters*, 41:389–395, 1999.

Dana Quade. Rank analysis of covariance. *Journal of the American Statistical Association*, 62:1187–1200, 1967.

F.H. Ruymgaart. A unified approach to the asymptotic distribution theory of certain midrank statistics. *Statistique non Parametrique Asymptotique, 118, J.P. Raoult. Lecture Notes on Mathematics*, 821, 1980.

Lee A. Seber, G. *Linear Regression Analysis.* Wiley, New Jersey, 1977.

A. Siemer. *Die Berücksichtigung von heterogen verteilten Kovariablen in einem nichtparametrischen Modell.* Master's Thesis, Georg-August-University Göttingen, 1999.

Student. The probable error of a mean. *Biometrika*, 6(1), 1908.

Karthinathan Thangavelu and Edgar Brunner. Wilcoxon–Mann–Whitney test for stratified samples and Efron's paradox dice. *J. Stat. Plann. Inference*, 137(3):720–737, 2007. doi: 10.1016/j.jspi.2006.06.005.

Neil H. Timm. *Multivariate Analysis with Applications in Education and Psychology.* Wadsworth Publishing Company, Inc., 1975. ISBN 0-8185-0096-4.

J. W. Tukey. The problem of multiple comparisons. *Unpublished manuscript reprinted in: The Collected Works of John W. Tukey*, 8:Chapman and Hall, June 1953.

A.W. Van der Vaart. *Asymptotic statistics.* Cambridge: Cambridge Univ. Press, 1998. ISBN 0-521-49603-9/hbk. doi: 10.1017/CBO9780511802256.