The Pennsulvania State University

The Graduate School

Department of Statistics

TESTING IN MULTIFACTOR HETEROSCEDASTIC ANOVA AND REPEATED MEASURES DESIGNS WITH LARGE NUMBER OF LEVELS

A Thesis in

Statistics

by

Haiyan Wang

Submitted in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

August 2004

The thesis of Haiyan Wang was reviewed and approved* by the following:

Michael G. Akritas
Professor of Statistics
Thesis Advisor, Chair of Committee

Steven F. Arnold
Professor of Statistics

Bing Li
Professor of Statistics
Chair of Graduate Study

Vernon M. Chinchilli
Distinguished Professor and Interim Chair of Health Evaluation Sciences

James Rosenberger
Professor of Statistics
Head of the Department of Statistics

*Signatures are on file in the Graduate School.

Abstract

Testing in multifactor heteroscedastic ANOVA and repeated measures designs with large number of levels

Testing methods for factorial designs with independent or dependent observations where some of the factors have a large number of levels have received a lot of attention recently. Most results for independent data in the literature have been restricted to procedures using the original observations for the balanced homoscedastic case, which require strong moment assumptions and are sensitive to outliers. The results in the literature for dependent data were extensively studied in parametric, nonparametric and semiparametric, and Bayesian models but all that do inference require large sample sizes or the normality assumption.

The first part of my thesis considers the use of rank statistics as robust alternatives for testing hypotheses in balanced and unbalanced, homoscedastic and heteroscedastic one-way and two-way ANOVA models when the number of levels of at least one factor is large. The second part of my thesis deals with various testing problems in possibly unbalanced and heteroscedastic multi-factor designs with arbitrary but fixed number of factors when at least one of the factors have large number of factor levels. Procedures based on both original observations and their (mid-)ranks are presented for the same general setting. The first two parts pertain to independent data. The third part of my thesis is focused on testing hypotheses in functional data, a fully nonparametric method for evaluating the effect of several crossed factors on the curve and their interactions with time. The asymptotics, which rely on the large number of measurements per curve (subject) and not on large group sizes, hold under the general assumption of

α-mixing without specifying the covariance structure, and do not require the measurements to be continuous or homoscedastic. A competing set of (mid-)rank procedures is also developed. The procedures in all three parts can be applied to both continuous and discrete ordinal observations. The rank tests are robust to outliers, invariant under monotone transformations, and do not require any restrictive moment conditions. Simulation studies reveal that the (mid-)rank procedures outperform those based on the original observations in all non-normal situations while they do not lose much power when normality holds. Applications to several data sets will be given and potential extensions in several directions will be discussed.

Table of Contents

List of Tables

List of Figures

Acknowledgements

I would like to express my deep appreciation and indebtedness to many people who have supported me to finish this thesis. My top appreciation goes to Dr. Michael G. Akritas, my thesis advisor, for his insightful guidance, enthusiastic comments and consistent encouragement. His support is so crucial in both my thesis writing and future career. I have learnt a lot from his intelligence and his way of getting work done perfectly by seeking elaboration. His non-stop hard-working personality will influence me throughout my life. My special thanks must go to my other committee members, Dr. Steven F. Arnold, Dr. Bing Li and Dr. Vernon M. Chinchilli, for their valuable comments and suggestions at various stages of the research. I am extremely grateful to Dr. Steven F. Arnold whose Linear Models class and textbook ”Mathematical Statistics” used by Dr. Tom Hettmansperger inspired my love and enthusiasm for statistics. I am also specially benifited from the Asymptotic Tools class taught by Dr. David Hunter and the Probability Theory class taught by Dr. Michael G. Akritas.

I would like to thank the group of meritable faculty, hard-working staff and active fellow students in the Department of Statistics at the Penn State University. The positive and friendly environment has made my study joyful and memorable. I owe my thanks to Dr. Steven F. Arnold, Dr. Tom Hettmansperger and Dr. David Hunter who demonstrated me the power of excellent teaching. I am deeply indebted to Dr. William Harkness for his mental support and encouragement for improving my commucation skills and teaching ability, as well as his valuable sacred friendship. I also benefited valuable insight and suggestions from Dr. Bruce G. Lindsay and enjoyed the friendship and discussions with Dr. Laura J. Simon and Dr. Donald St. P. Richards. My fellow students and friends are important part of my life at Penn State.

I want to express my special thank to Efi Antoniou whose love, support, and friendship has encouraged me to survive the hardest time ever in my life so far. Discussion with Lan Wang, Jiping Wang, and Jiawei Liu are often cheerful. In addition, I would like to thank Hong Xu, Weixiu Luo, Hyekyung Jung, and Jane Okou for a data set used in the thesis. I appreciate Bob Zhang's computer and mental support in the past half an year. My thanks also go to Rui Ding, Changxuan Mao, Haihong Li, Mike Kwanisai, Ke Yang, Fengjuan Xuan, Weixin Yao, Yu Wang, Junjia Zhu, Hyunsook Lee, Ryan Elmore, Shaoli Wang, Steve Bai, and Yi-Ju Chen. They all made my life at Penn State happier and unforgettable.

My late father, Xuewang Wang, and mother, Guizhen Jiang, have been encouraging me to pursue knowledge since I was very young. Their unconditional love, support and confidence in me has consistently contributed to my progress and will guide me through my life journey. Finally, I owe a lot to my daughter Luann, and therefore Eura, who has been giving me sun shine in the past years.

Haiyan Wang

The Pennsylvania State University

August, 2004

Chapter 1

Introduction

1.1 Asymptotics for Analysis of Variance

Factorial designs arise very commonly in scientific investigations. The classical ANOVA model assumes that the error terms are iid normal, in which case the F statistics for testing the null hypotheses of no treatment effects or no interaction effects have certain optimality properties (cf. Arnold, 1981, Chapter 7). The study of properties of F-tests under violation of the classical assumptions of normality and homoscedasticity has a long history. See for example Box (1954), Box and Andersen (1955), Scheffé (1959, Chapter 10), Miller (1986, Chapter 4). However, these studies pertain only to the case when the number of treatment levels is small. In this case, Arnold (1980) showed that the classical F-test is robust to the normality assumption if in addition the sample size per treatment level tends to infinity. Portnoy (1984) considered the case where the number of treatment levels also goes to infinity with the sample size. Recently, there has been some interest in investigating the behavior of these classical tests when the number of treatment levels is large but sample size per treatment combination is limited. Motivating examples from agricultural trials are mentioned in Brownie and Boos (1994) and Wang and Akritas (2002) for one- and two-way layouts, respectively.

Boos and Brownie (1995) derived the asymptotic distribution of the classical F-statistic for balanced one-way ANOVA and balanced randomized complete block design (without interaction effects) in this setting. Akritas and Arnold (2000) studied a general class of designs, including balanced and unbalanced, fixed effect and random effect, one- and multi-factor. Independently, and using different asymptotic techniques, Bathke (2002) also generalized the results of Boos and Brownie (1995) to fixed effects balanced multi-factor designs. The aforementioned papers considered only the balanced homoscedastic case and showed that the usual F-procedure (i.e. using the classical F-test statistic with critical points from the F distribution) is asymptotically correct. The unbalanced case, however, is considerably more challenging. When the sample size per treatment combination is small, the assumption of homoscedasticity is difficult to justify. Krutchkoff (1989) found that both the classical and weighted F-tests perform poorly in the heteroscedastic case with a large number of groups and few observations per group. As Akritas and Papadatos (2004) show, even under homoscedasticity the F-procedure in the one-factor design is not asymptotically valid in the unbalanced case unless the group sizes are also large. Moreover, though an asymptotically valid procedure based on the classical F statistic can be constructed for the unbalanced homoscedastic case, the asymptotic and empirical properties of this procedure are less appealing than those of a heteroscedastic procedure they construct for one-way design. Wang and Akritas (2002) study the asymptotics for suitable test statistics in the heteroscedastic two-way design when the number of treatment levels is large, both with small and large sample sizes per cell.

It should be pointed out that F-test statistics typically have no closed form expression in unbalanced designs. This is due to the lack of orthogonality in such designs. Thus, it is challenging to construct suitable test statistics for higherway layouts. In this thesis, we will construct test statistics for the unbalanced heteroscedastic case and consider the asymptotic distributions under various situations. The proposed test statistics are extensions of corresponding test statistics proposed in Akritas and Papadatos (2004) for one-factor. It should be mentioned that this idea is also related to the statistics proposed by Yates (1934) (also called the method of unweighted means, cf. Sahai and Ageel, 2000, p. 220-222) for unbalanced homoscedastic two-way designs. For economy in space, only test procedures for the heteroscedastic case were studied, which of course apply also in the homoscedastic case. In view of the aforementioned findings in Akritas and Papadatos (2004), there is no efficiency loss by applying a heteroscedastic procedure to homoscedastic data which is also justified by simulation studies in this thesis.

The presentation of a general asymptotic testing theory in the case where some or all of the factors may have a large number of levels is considerably more complicated than the case of fixed number of factor levels and large group sizes. In the latter case, all test statistics can be expressed as a quadratic form involving contrasts of the finitely-many group means and thus one general theorem covers all testing problems; see Akritas, Arnold and Brunner (1997). The present test statistics will have either a limiting χ^2 distribution, or a limiting normal distribution, depending on the hypothesis tested and which of the factors have large number of levels. Of course, the arguments also depend on whether the group sizes are large or small. In order to achieve as concise a formulation as possible, only the testing theory for main effects, two- and three-way interactions are presented. The treatment of these testing problems captures all ideas and techniques that would be required for developing test procedures for still higher-way interactions. Moreover, with this limited scope, it is possible to give a formulation that indexes the observations with at most five indices regardless of how many factors are present.

1.2 Asymptotics for Repeated Measures

Studies involving repeated measurements within a subject or stratum arise frequently in scientific investigations, and have attracted considerable attention in the statistical literature. Examples of such data are growth curves, reaction curves in chemical experiments, evaluation of disease progression, biomarkers measured over time, seismic recordings etc. Such data are called longitudinal, curve or functional data, though the latter two terms seem to be more common when the number of observations per subject is large.

Approaches for the analysis of longitudinal data include linear and nonlinear mixed-effects models with parametric (cf. Pinhero and Bates, 1995), semiparametric (Davidian and Gallant, 1993), nonparametric (Mallet, 1986), and Bayesian modeling (cf. Smith and Roberts, 1993), generalized linear models (Liang and Zeger, 1986; Zeger and Liang, 1986) to deal with discrete ordinal data, and the fully nonparametric marginal model for all types of ordinal data (Akritas and Brunner, 1997). For text-book presentations of such material, see Vonesh and Chinchilli (1997), Brunner, Domhof and Langer (2002), Vonesh and Chinchilli (1997), Davidian and Giltinan (1995), Diggle, Liang and Zeger (1994) and Hand and Crowder (1996). These approaches are mainly suitable when the number of within stratum replications is relatively small. Time series models (Brillinger, 1973, 1980), functional ANOVA models (Ramsey and Silverman, 1997), smoothing spline models (Brumback and Rice, 1998; Wang 1998), varying coefficient models with smoothing splines and other nonparametric smoothing techniques (Hoover, Rice, Wu and Yang, 1998, Fan and Zhang, 2000, Chiang, Rice and Wu, 2001, Wang, 2003, Lin, Wang, Welsh, and Carroll, 2004), can also be used for functional or curve data where the number of within stratum replications is large. See also Fan and Lin (1998) who considered a comparison of curves problem, analogous to some lack-of-fit methods with independent observations, and used ideas from adaptive Neyman test and wavelet thresholding to improve the power of their test procedures, and Lin and Ying (2001) who considered a counting processes approach for varying coefficient models using weighted least squares estimators for the (cumulative function of the) coefficient (curve) without smoothing techniques. Asymptotic theory is presented in some of the aforementioned functional data approaches, but always under the assumption of large group sizes, and typically under additional modeling assumptions. As far as we can tell, only Fan and Lin (1998), Fan and Zhang (2000), and Chiang, Rice and Wu(2001) attempt asymptotics by allowing the number of time points to also tend to infinity. Fan and Lin (1998) require stationary errors and the test procedures they provide are only for simple effect of treatment. Chiang, Rice and Wu, (2001) considered componentwise smoothing spline estimator for the coefficient curve of a time varying coefficient model with time independent covariates when the total number of subjects going to infinity and the number of repeated measures is a very small order of the total number of subjects. The asymptotic distribution of their estimator has an asymptotic bias depending on the fourth derivative of the unknow coefficient curve which is barely estimatiable in practice. Reducing the asymptic bias can be achieved if sacrificing the convergence rate. Fan and Zhang (2000) assume no interaction of treatment by time and require stationary errors.

Here we will consider curve or functional data situations with identical sampling points where there is interest in describing and evaluating the effect of several baseline, or time-independent, factors (sex, age, dose level etc) and the interaction of such factors with the time factor. For example, in a dose-response experiment where the response consists of correlated observations over time it might be of interest to evaluate the dose effect and the interaction between dose and time. The marginal model we consider is purely nonparametric. Thus the responses can be measured on continuous or discrete ordinal scale, there is no restriction in the allowed pattern of heteroscedasticity, and no model describing the change in the response distribution from one covariate value to another, or from one time point to another. Moreover, the asymptotic theory does not require large group sizes, and simulations studies confirm this. Thus, the present methods fill a gap in the available data analytic tools for functional data. The present inference procedures can be used in conjunction with existing graphical methods such as principal component analysis and dynamic time warping; see Ramsey and Silverman (1997), Ramsey and Li (1998), Ke and Wang (2001). Some new graphical methods for describing the effects are also discussed here.

Because statistics for the usual normal-based procedures (univariate or multivariate) have no closed form expression in unbalanced designs, and also because they are not designed to accommodate heteroscedasticity, we do not undertake a study of the asymptotic properties of these statistics. See Arnold (1981) for such a study under the condition that the group sample sizes tend to infinity, while the number of groups and repeated measures per subject remain fixed. Here we propose new test statistics falling in two distinct categories. One to test for effects described by a small number of parameters (e.g. group effects), and another to test for effects described by a large number of parameters (e.g. main time effects, and interactions between time and groups). Test statistics in the first category are quadratic forms having asymptotically a χ^2 distribution, while those in the second category are in the form of a difference of two quadratic forms and have asymptotically a normal distribution. The asymptotic theory for both categories of statistics requires novel techniques which rely on the number of time points tending to infinity (the number of subjects can be small or large) under quite general unspecified covariance structure. In particular, all derivations are done under an α-mixing assumption which includes most common time series models.

1.3 Rank Tests and Fully Nonparametric Hypotheses

The aforementioned papers in literature for both independent and repeated measures data focus mainly on procedures based on the original observations which require finite higher order moments. In addition, as is well known, test statistics based on the original observations are sensitive to outliers and can perform poorly away from the normal distribution. Boos and Brownie (1995) also discuss rank tests but in a rather limited context (only for the balanced homoscedastic one-way design and a balanced two-way design with no interaction).

Rank test statistics are constructed by simply replacing the original observations by their (mid-)ranks in the corresponding test statistics. The main tools for our asymptotic development is the asymptotic rank transform (Akritas, 1990). The method of asymptotic rank transform consists in showing that the rank statistic is asymptotically equivalent to another statistic which is based on a non-random transformation of the data. Since the results for the statistics based on the original observations can be used for any non-random transformation of the data, the asymptotic theory for the rank statistics follows. Of course the hypotheses tested by the rank statistics are different from those based on the original observations. When the test statistic is the ratio of two quadratic forms, the rank statistic tests hypotheses about relative treatment effects and are closely related to the *d* statistic used in behavioral psychology (cf. Cliff, 1993). Testing the equality of the relative treatment effects is commonly advocated in the context of rank methods (cf Brunner, Domhof and Langer, 2002, p.37). When the test statistic is a Wald-type quadratic form, the rank statistic tests fully nonparamtric hypotheses which are first introduced by Akritas and Arnold (1994) for multivariate repeated measures designs in which each individual receives each possible treatment combination. Their work is continued for unbalanced factorial designs (Akritas, Arold, and Brunner, 1997) with independent observations. Applying the nonparametric modeling idea and martingal theory, factorial designs are used to analyze independent right-censored data (Akritas and Brunner, 1997a) and repeated measures with right-censored data (O'Gorman and Akritas, 2001). Akritas and Brunner (1997b) worked on mixed models with uncensored data. The common idea of nonparametric hypotheses is to specify the distribution function at each level combination of factors and define the effects from decomposition of the distribution function in the same way as decomposing the expected value of response variable for parametric hypotheses. The hypotheses based on the decomposition of group means, which the statistics based on the original observations test, are implied by the fully nonparametric hypotheses.

In this thesis, we also consider test statistics based on the overall (mid-)ranks of the data for balanced and unbalanced, homoscedastic and heteroscedastic one-way and two-way layouts, possible unbalanced heteroscedastic higherway layout, and heteroscedastic repeated measures data when at least one factors have large number of levels. The hypotheses in one-way and two-way ANOVA are given in Chapter 2. For simplicity, the hypotheses in higherway layout are stated in terms of the fully nonparametric formulation though the actural hypotheses could be on relative effects which are implied by their corresponding nonparametric effects.

The nice feature of rank tests and nonparametric hypotheses are: the methods can be applied to a wide variety of applications including continous and all ordinal responses; both the hypotheses and the test statistics are invariant under monotone transformations of the response variable; the procedures are robust to outliers and do not need any moment assumptions. Simulation studies reported in this thesis indicate that the rank statistics outperform those based on the original observations in both Type I error and Type II error estimate.

1.4 Projection Method for Quadratic Forms

Boos and Brownie (1995) used very complex algebraic arguments in deriving the asymptotic distribution of the F-statistic for some fixed ANOVA models. Their method can hardly extend to other designs. Akritas and Arnold (2000) used different technique by finding the asymptotic joint distribution of MST and MSE and then the asymptotic distribution of F. Note that MST and MSE are typically not even uncorrelated for nonnormal models though they are independent for normal models. Akritas and Papadatos (2004) introduced a novel application of Hájek's projection method for the difference of two quadratic forms. Their method allows for weaker assumptions regarding moments and sample sizes in inference.

To derive the asymptotic distribution of a sequence of statistics T_n, an intuitive idea is to show that it is asymptotically equivalent to a sequence S_n whose asymptotic distribution is easier to establish. For example, projection method is involved in obtaining the asymptotic results for U-statistics. van der Vaart (1998, Chapter 11) provided a nice overview of the projection method.

Definition 1.4.1. *Let $\mathbf{X}_1,\ldots,\mathbf{X}_n$ be independent random vectors and let S be the set of all variables of the form $\sum_{i=1}^n g_i(\mathbf{X}_i)$ for arbitrary measurable functions g_i with $Eg_i^2(\mathbf{X}_i) < \infty$, then the projection of a variable onto S is called its Hájek projection.*

Lemma 1.4.2. *Let $\mathbf{X}_1,\ldots,\mathbf{X}_n$ be independent random vectors. Then the Hájek projection of arbitrary random variable T with finite second moment is given by*

$$\hat{S} = \sum_{i=1}^{n} E(T|\mathbf{X}_i) - (n-1)E(T).$$

The proof of above lemma can be found in, for example, p. 157 of van der Vaart (1998).

In this thesis, a slight modification of Hájek projection is used in the sense that the dependence in some observations is ignored in order to define the class of random variables onto which to project and the asymptotic equivalence were established by further rigid proof.

1.5 Thesis Organization and Some Notations

The rest of the thesis is mainly divided into three parts. The first part, Chapter 2, gives asymptotic results of rank tests for main effects and interaction effects in balanced and unbalanced, homoscedastic and heteroscesastic one-way and two-way designs when at least one factors have large number of levels. The second part, Chapter 3, considering testing problems in possibly unbalanced heteroscedastic higherway layout when there are arbituary but fixed number of factors. Asymptotic results are presented for both tests based on the original observations and their overall (mid-)ranks when at least one factors have large number of levels. The third part, Chapter 4, deals with testing hypotheses in heteroscedastic functional data when the factor, whose levels the subjects are nested in, have fixed number of levels and the number of repeated measurements is large.

In all aforementioned chapters, the sample/group sizes per treatment combination can be either fixed or tend to infinity. Simulation results are presented in each chapter and application to real data sets is given.

Finally Chapter 6 will summarize the work in this thesis and discuss some potential future extensions.

Throughout the thesis we use $c(x,y) = [I(x \le y) + I(x < y)]/2$, where $I(A)$ is the indicator function for the event A, and the following notations. In the case when there are four factors in the model, we define $\overline{X}_{ijkl.} = n_{ijkl}^{-1} \sum_{m=1}^{n_{ijkl}} X_{ijklm}$, $\widetilde{X}_{i....} = (bcd)^{-1} \sum_{j=1}^{b} \sum_{k=1}^{c} \sum_{l=1}^{d} \overline{X}_{ijkl.}$, $\widetilde{X}_{ij...} = (cd)^{-1} \sum_{k=1}^{c} \sum_{l=1}^{d} \overline{X}_{ijkl.}$, $\widetilde{X}_{.j...} = (acd)^{-1} \sum_{i=1}^{a} \sum_{k=1}^{c} \sum_{l=1}^{d} \overline{X}_{ijkl.}$, and $\widetilde{X}_{.....} = (abcd)^{-1} \sum_{i=1}^{a} \sum_{j=1}^{b} \sum_{k=1}^{c} \sum_{l=1}^{d} \overline{X}_{ijkl.}$. Finally, $S^2_{ijkl,X} = (n_{ijkl}-1)^{-1} \sum_{m=1}^{n_{ijkl}} (X_{ijklm} - \overline{X}_{ijkl.})^2$, $N = \sum_{i=1}^{a} \sum_{j=1}^{b} \sum_{k=1}^{c} \sum_{l=1}^{d} n_{ijkl}$ and $n(a,b,c,d) = \min\{n_{ijkl},\ i = 1,\ldots,a, j = 1,\ldots,b, k = 1,\ldots,c, l = 1,\ldots,d\}$. Similar notation applies when X is replaced by e, Y or Z. Analogous versions of this notation will also be used when there are three, two or one factor(s) in the model.

Chapter 2

Rank Tests for ANOVA with Large Number of Factor Levels

Analysis of variance models involve an experiment comparing one or more factors at each of their levels. Usual F-statistics are used to test for the effect of factors and their interactions under the usual assumptions. It is well established that the F-test is robust to the normality assumption if the number of factor levels is fixed and the sample sizes tend to infinity (see Arnold 1980). The theory of weighted least squares statistics is also well known (see Arnold 1981, chapter 13) in this setting. In some disciplines, the design of experiment involves factor(s) with a large number of levels and the replications for each level combination are limited. For example, it is common to see a large number of treatments (cultivars, pesticides, fertilizers, etc) in agricultural study. See Brownie and Boos (1994) and Wang (2003) for one- and two-way layouts in agricultural trials, respectively.

The assumption that a large number of populations are homoscedastic is difficult to justify when the sample size from each population is small. Box (1954a, 1954b) and Scheffé (1959) have systemacically studied the sensitivity of ANOVA methods to deviations from assumptions like normality, equal variances and independent error terms. For example, Box (1954b) calculated the actual α-level for the two-way ANOVA when the error terms are normally distributed and found out that 3 fold column-to-column differences in the variances cause a too liberal behavior of the test for equal column means. Testing for equal row means yields discrepancies of similar order but in the opposite direction (see Box 1954b, p.492). Simulation results in this thesis also revealed extremely unsatisfied type I error estimate and poor power performance in the classical ANOVA in some cases even when the error terms are normally distributed. In this case, the classical ANOVA or the chi-square approximation, which is based on a large number of observations and a small number of levels, would not be appropriate.

Asymptotic results for the case when the number of levels is large and the number of observations for each level is small were first derived by Boos and Brownie (1995), for the null distribution of some balanced fixed effects models. Akritas and Arnold (2000) obtained asymptotic results covering a very general class of designs and fixed local alternatives. However, aforementioned results pertain only to the homoscedastic case with small (fixed) sample sizes. Akritas and Papadatos (2004) investigated test procedures in unbalanced heteroscedastic one-way ANOVA when the number of levels of the factor tends to infinity. They considered both the classical weighted statistic and a new unweighted statistic. Using exact calculations under normality, they demonstrate that the classical weighted statistic is very unstable if the sample sizes are small. They also find that asymptotic approximation to the distribution of the weighted statistic requires that the average sample size tends to infinity faster than $a^{1/2}$, where a is the number of levels. Their new unweighted statistic is applicable with small and large sample sizes and its asymptotic and small sample properties are preferable to those of the procedure based on the F-test statistic even in the homoscedastic case. Wang (2003) considers testing hypothesis in homoscedastic and heteroscedastic two-way ANOVA.

The aforementioned statistics of Akritas and Papadatos (2004) and Wang (2003) are based on the original observations and require finite higher order moments. In addition, as is well known, test statistics based on the original observations are sensitive to outliers and can perform poorly away from the normal distribution. The simulation studies show that the above heteroscedastic procedures based on the original In this chapter, we consider test statistics based on the overall ranks of the data. Rank tests were also considered in Boos and Brownie (1995) but only for the balanced homoscedastic one-way design and a balanced two-way design with no interaction. The main tools for our asymptotic development is the asymptotic rank transform

(Akritas, 1990) and Hájek's projection method for quadratic forms (Akritas and Papadatos, 2002). The reported simulation studies indicate that the rank statistics have more stable Type I error rate and are much more powerful when the error distribution is away from the normal.

The rest of the chapter is organized as the following. In Section 2.1, we present results for the one-way design. The homoscedastic and heteroscedastic cases are treated separately. Section reftwoway pertains to the two-way layout. For reasons explained there we consider only the heteroscedastic case. Simulation results are presented in Section 2.3, while the proofs are given in Section 2.4.

2.1 Main Results on One-Way ANOVA

In one-way analysis of variance, we have independent observations $X_{ij} \sim F_i(x)$, $i = 1, \cdots, r$, $j = 1, \cdots, n_i$. In order to accommodate discrete and continuous data in the same notation, we will define all distribution function, including empirical ones, to be the average of their right and left continuous versions. We are interested in testing the hypothesis of no treatment effect in a setting where $r \to \infty$. An asymptotically equivalent form of the classical Kruskal-Wallis test statistic compares the treatment and error mean squares calculated on the ranks. Here we will consider this form and establish its asymptotic behavior when the number of treatments is large.

Let $H(x) = N^{-1} \sum_{i=1}^{r} \sum_{j=1}^{n_i} F_i(x)$, where $N = \sum_{i=1}^{r} n_i$, and set

$$\widehat{H}(x) = N^{-1} \sum_{i=1}^{r} \sum_{j=1}^{n_i} c(X_{ij}, x),$$

where $c(x,y)$ is defined in §1 , be its empirical version. Then $R_{ij} = 1/2 + N\widehat{H}(X_{ij})$ is the (mid-)rank of X_{ij} among all N observations. Let $MS\alpha_R$ be the mean squares for treatment on mid-rank,

$$MS\alpha_R = \frac{1}{r-1} \sum_{i=1}^{r} \sum_{j=1}^{n_i} \left(\overline{R}_{i.} - \overline{R}_{..}\right)^2. \tag{2.1.1}$$

In all that follows, $Y_{ij} = H(X_{ij})$, $Z_{ij} = \widehat{H}(X_{ij})$, and $\sigma_i^2 = \mathrm{Var}(Y_{ij})$. The results for homoscedastic and heteroscedastic cases are stated separately.

2.1.1 The homoscedastic Case

In the homoscedastic case, the null hypothesis can be stated as $H_{0,1}(\alpha) : F_1 = \cdots = F_r$, which certainly implies homoscedasticity under the null hypothesis. Let

$$MSE_R = \frac{1}{N-r}\sum_{i=1}^{r}\sum_{j=1}^{n_i}\left(R_{ij} - \overline{R}_{i.}\right)^2 \text{ and } F_{R,r} = \frac{MS\alpha_R}{MSE_R}$$

be the error mean squares calculated on the mid-ranks, and the corresponding F-ratio.

For convenience, we first state the two assumptions needed for the next theorem.

1. When the n_i's remain fixed, assume n_i satisfy
$$B_r = \frac{1}{r}\sum_{i=1}^{r} n_i \to b \in (1,\infty), \qquad B_{1r} = \frac{1}{r}\sum_{i=1}^{r}\frac{1}{n_i} \to b_1,$$
as $r \to \infty$.

2. When $n_i = n_i(r) \to \infty$, as $r \to \infty$. Set $n(r) = \min\{n_i(r); i = 1, \ldots, r\}$, $\kappa(r) = \max\{n_i(r); i = 1, \ldots, r\}$, and assume that
$$n(r) \to \infty, \text{ and } \kappa(r) - n(r) \leq C(r), \text{ for all r},$$
where $C(r) = o(n(r))$, as $r \to \infty$.

Theorem 2.1.1. **(Unbalanced homoscedastic Case)** *Let $H_{0,1}(\alpha)$ be satisfied.*

(a) Under assumption 1 and if $\limsup_{r\to\infty} \frac{1}{r}\sum_{i=1}^{r} n_i^{4+\delta} < \infty$, for some $\delta > 0$, then
$$r^{1/2}(F_{R,r} - 1) \to N\left(0, \tau^2\right), \text{ as } r \to \infty,$$
where, letting $\mu_4 = E[H(X_{ij}) - 1/2]^4/\sigma^4$, with $\sigma^2 = Var\{H(H_{1j})\}$,
$$\tau^2 = \frac{2b}{b-1} + (\mu_4 - 3)\frac{b(bb_1 - 1)}{(b-1)^2}.$$

(b) Under assumption 2 listed above,

$$r^{1/2}(F_{R,r}-1) \to N(0,2), \quad \text{as } r \to \infty.$$

Corollary 2.1.2. (Balanced homoscedastic Case) *Let $H_{0,1}(\alpha)$ be satisfied.*

(a) If $n \geq 2$ remains fixed, then

$$r^{1/2}(F_{R,r}-1) \to N\left(0, \frac{2n}{n-1}\right), \quad \text{as } r \to \infty.$$

(b) If $n = n(r) \to \infty$, as $r \to \infty$, then

$$r^{1/2}(F_{R,r}-1) \to N(0,2), \quad \text{as } r \to \infty.$$

2.1.2 The Heteroskedastic Case

In the heteroscedastic case, null hypothesis that will be tested is $H_{0,2}(\alpha) : p_1 = \cdots = p_r$, where $p_i = E\{H(X_{ij})\}$. We remark that the p_i are called relative treatment effects, and are closely related to the d statistic used in behavioral psychology (cf. Cliff, 1993). Testing the equality of the relative treatment effects is commonly advocated in the context of rank methods (cf Brunner, Domhof and Langer, 2002, p.37).

In the balanced case, the expectations of the treatment and error mean squares are equal under the null hypothesis, and we can use the same test statistic as in §2.1.1. However, this does not hold in the unbalanced heteroscedastic case. Akritas and Papadatos (2002) proposed two test statistics for dealing with heteroscedasticity in the unbalanced one-way design. The first alters MSE by considering a different weighted average of the group sample variances in such a way as to have the same expected value as $MS\alpha$, and the other is a weighted Wald-type statistic. The weighted Wald-type statistic requires the sample sizes to also tend to infinity. This is supported by simulations reported in Akritas and Papadatos (2002) which suggest that the asymptotic approximation to the distribution of the statistic becomes satisfactory if the sample sizes are greater than 80. The first approach can be used with small sample sizes and has the advantage that it reduces to the usual statistic in the balanced case. Since it also has a simpler asymptotic theory we will present only a rank version of it, in Theorem 2.1.3. In addition we will consider, in Theorem 2.1.5, a different test statistic, which can be used for both small and large sample sizes.

Let

$$MSE_R^* = \frac{1}{r-1}\sum_{i=1}^{r}\left(1-\frac{n_i}{N}\right)S_{R,i}^2, \tag{2.1.2}$$

where $S_{R,i}^2 = (n_i-1)^{-1}\sum_{j=1}^{n_i}\left(R_{ij}-\overline{R}_{i.}\right)^2$, be the altered MSE applied on the ranks and set $F_R = MS\alpha_R/MSE_R^*$. We will consider the asymptotic distribution of $\sqrt{r}(F_R-1)$ as r gets large.

We have the following result:

Theorem 2.1.3. *(***Unbalanced heteroscedastic case***) Let*

$$\sigma_0^2 = \lim_{r\to\infty}\frac{2}{r}\sum_{i=1}^{r}\left(1-\frac{n_i}{N}\right)\sigma_i^2,\ \ \tau_0 = \lim_{r\to\infty}\frac{2}{r}\sum_{i=1}^{r}\left(1-\frac{n_i}{N}\right)^2\frac{n_i}{n_i-1}\sigma_i^4. \tag{2.1.3}$$

Assume $\sigma_0^2>0$ *and there exists* $\delta>0$ *such that* $r^{-1}\sum_{i=1}^{r}n_i^{2+\delta}$ *is finite as* $r\to\infty$*. Then under* $H_{0,1}(\alpha)$ *or* $H_{0,2}(\alpha)$*,*

$$\sqrt{r}(F_R^{(1)}-1)\xrightarrow{d}N\left(0,\frac{\tau_0}{\sigma_0^4}\right),$$

The proof is given in section 2.4. When the sample sizes are same, $MSE_R^* = MSE_R$. We naturally have the following corollary:

Corollary 2.1.4. *(***Balanced heteroscedastic case***) Let* MSE_R *be the mean square errors calculated on the mid-ranks, and* $F_{R,r} = MS\alpha_R/MSE_R$*. Denote*

$$\tau_0 = \frac{2n}{n-1}\lim_{r\to\infty}\frac{1}{r}\sum_{i=1}^{r}\sigma_i^4 \text{ and } \sigma_0^2 = \lim_{r\to\infty}\frac{2}{r}\sum_{i=1}^{r}\sigma_i^2$$

Assume $\sigma_0^2>0$ *Then*

$$\sqrt{r}(F_{R,r}-1)\xrightarrow{d}N\left(0,\frac{\tau_0}{\sigma_0^4}\right).$$

Note that the test statistic used in the heteroscedastic unbalanced case is valid only if the sample sizes n_i are small. This motivates us to construct another test statistic that is good for both small and large sample sizes.

Define

$$MST = \frac{1}{r-1}\sum_{i=1}^{r}\left(\overline{X}_{i.} - \widetilde{X}_{..}\right)^2, \quad MSE^{(2)} = \frac{1}{r}\sum_{i=1}^{r}\frac{1}{n_i}S_{X,i}^2, \text{ and } F_2 = \frac{MST}{MSE^{(2)}}, \qquad (2.1.4)$$

where $S_{X,i}^2 = (n_i-1)^{-1}\sum_{j=1}^{n_i}\left(X_{ij} - \overline{X}_{i.}\right)^2$. Then under the null hypothesis of no treatment effect,

$$E(MST) = E(MSE^{(2)}) = r^{-1}\sum_{i=1}^{r}\frac{1}{n_i}\sigma_i^2.$$

It is reasonable to compare MST with $MSE^{(2)}$ for the test of no treatment effect. We will consider the rank version of above statistics.

Theorem 2.1.5. *(Unbalanced Heteroscedastic Case) Let $H_{0,1}(\alpha)$ or $H_{0,2}(\alpha)$ be satisfied.*

(a) If $n_i \geq 2$ fixed, assume the following limits exist

$$\nu_2^2 = \lim_{r\to\infty}\frac{1}{r}\sum_{i=1}^{r}\frac{1}{n_i}\sigma_i^2 > 0 \qquad \tau_2 = \lim_{r\to\infty}\frac{1}{r}\sum_{i=1}^{r}\frac{2\sigma_i^4}{n_i(n_i-1)},$$

then $\sqrt{r}(F_{R,2}-1) \xrightarrow{d} N(0,\tau_2/\nu_2^4)$ as $r\to\infty$, where MST_R, $MSE_R^{(2)}$, $F_{R,2}$ are the rank version of MST, $MSE^{(2)}$ and F_2 defined in (2.1.4).

(b) If $n_i = n_i(r)\to\infty$ as $r\to\infty$, assume $n(r) = o(r^{\delta/(4+2\delta)})$ for some $\delta > 0$ and the following limits exist

$$\nu_3^2 = \lim_{r\to\infty}\frac{1}{r}\sum_{i=1}^{r}\frac{n(r)}{n_i}\sigma_i^2 > 0 \qquad \tau_3 = \lim_{r\to\infty}\frac{1}{r}\sum_{i=1}^{r}\frac{2n^2(r)\sigma_i^4}{n_i(n_i-1)},$$

we have $\sqrt{r}(F_{R,2}-1) \xrightarrow{d} N(0,\tau_3/\nu_3^4)$ as $r\to\infty$.

Corollary 2.1.6. *(Unbalanced Homoscedastic Case) Let $H_{0,1}(\alpha)$ or $H_{0,2}(\alpha)$ be satisfied.*

(a) *If* $n_i \geq 2$ *fixed, assume the following limits exist*

$$\nu_2^2 = \lim_{r\to\infty} \frac{1}{r}\sum_{i=1}^{r} \frac{1}{n_i} > 0 \qquad \tau_2 = \lim_{r\to\infty} \frac{1}{r}\sum_{i=1}^{r} \frac{2}{n_i(n_i-1)},$$

then $\sqrt{r}(F_{R,2}-1) \xrightarrow{d} N(0, \tau_2/\nu_2^4)$ *as* $r \to \infty$.

(b) *If* $n_i = n_i(r) \to \infty$ *as* $r \to \infty$, *assume* $n(r) = o(r^{\delta/(4+2\delta)})$ *for some* $\delta > 0$ *and the following limits exist*

$$\nu_3^2 = \lim_{r\to\infty} \frac{1}{r}\sum_{i=1}^{r} \frac{n(r)}{n_i} > 0 \qquad \tau_3 = \lim_{r\to\infty} \frac{1}{r}\sum_{i=1}^{r} \frac{2n^2(r)}{n_i(n_i-1)},$$

we have $\sqrt{r}(F_{R,2}-1) \xrightarrow{d} N(0, \tau_3/\nu_3^4)$ *as* $r \to \infty$.

Under the homoscedastic case with balanced sample sizes, the test statistic reduces to the classical F-statistic introduced in §2.1.1. We naturally have the following corollary.

Corollary 2.1.7. *(Balanced Homoscedastic Case) Let* $H_{0,1}(\alpha)$ *or* $H_{0,2}(\alpha)$ *be satisfied.* $F_{R,a}$ *is defined in* §*2.1.1*

(a) *If* $n \geq 2$ *fixed,*

$$\sqrt{r}(F_{R,r}-1) \xrightarrow{d} N(0, \frac{2n}{n-1}) \text{ as } r \to \infty.$$

(b) *If* $n = n(r) \to \infty$ *as* $r \to \infty$, *assume* $n(r) = o(r^{\delta/(4+2\delta)})$ *for some* $\delta > 0$, *then* $\sqrt{r}(F_{R,r}-1) \xrightarrow{d} N(0,2)$ *as* $r \to \infty$.

Note that Theorem 2.1.1, Corollary 2.1.2, Corollary 2.1.6 and Corollary 2.1.7 are all for homoscedastic case. However, Theorem 2.1.1 and Corollary 2.1.2 are obtained under $H_{0,1}(\alpha)$, while Corollary 2.1.6 and Corollary 2.1.7 hold under both $H_{0,1}(\alpha)$ and $H_{0,2}(\alpha)$.

2.2 Main Results on Two-Way ANOVA

In two-way analysis of variance, we have independent observations $X_{ijk} \sim F_{ij}(x)$, $i = 1, \cdots, r,\ j = 1, \cdots, c, k = 1, \cdots, n_{ij}$. For this section we set $H(x) = N^{-1}\sum_{i=1}^{r}\sum_{j=1}^{c} n_{ij}F_{ij}(x)$ for the average distribution function, and

$$\widehat{H}(x) = N^{-1}\sum_{i=1}^{r}\sum_{j=1}^{c}\sum_{k=1}^{n_{ij}} c(X_{ijk}, x),$$

for its empirical version. Thus $R_{ijk} = 1/2 + N\widehat{H}(X_{ijk}))$ is the (mid-)rank of X_{ijk} among all observations. We also let $Y_{ijk} = H(X_{ijk})$, so the relative treatment effects are $p_{ij} = E(Y_{ijk})$. Consider the decomposition

$$p_{ij} = \mu + \alpha_i + \beta_j + \gamma_{ij}, i = 1, \cdots, r,\ j = 1, \cdots, c, k = 1, \cdots, n_{ij},$$

where $\sum_{i=1}^{r}\alpha_i = \sum_{j=1}^{c}\beta_j = \sum_{i=1}^{r}\gamma_{ij} = \sum_{j=1}^{c}\gamma_{ij} = 0$. We are interested in testing for no main effects and no interaction effects, namely $H_0(\alpha)$: all $\alpha_i = 0$, $H_0(\beta)$: all $\beta_i = 0$, and $H_0(\gamma)$: all $\gamma_{ij} = 0$. Due to the fact that the number of column levels is fixed, testing for column effects requires different techniques and will not be presented here.

Because even under homoscedasticity the rank-transformed data can be heteroscedastic (Akritas, 1990), we will consider only the heteroscedastic case. Motivated by Wang and Akritas (2002), we will consider the test statistics defined from the following mean squares

$$\begin{aligned}
MST_\alpha &= \frac{1}{r-1}\sum_{i=1}^{r}\sum_{j=1}^{c}\left(\widetilde{R}_{i..} - \widetilde{R}_{...}\right)^2 \\
MST_\gamma &= \frac{1}{(r-1)(c-1)}\sum_{i=1}^{r}\sum_{j=1}^{c}\left(\overline{R}_{ij.} - \widetilde{R}_{i..} - \widetilde{R}_{.j.} + \widetilde{R}_{...}\right)^2 \\
MSE_R^{(3)} &= \frac{1}{rc}\sum_{i=1}^{r}\sum_{j=1}^{c}\frac{S_{ij,R}^2}{n_{ij}}.
\end{aligned}$$

Thus we consider test statistics

$$F_{R,\alpha,2} = \frac{MST_\alpha}{MSE_R^{(3)}} \text{ and } F_{R,\gamma,2} = \frac{MST_\gamma}{MSE_R^{(3)}},$$

for the hypotheses $H_0(\alpha)$ and $H_0(\gamma)$, respectively, and study the asymptotic distribution of $\sqrt{r}(F_{R,\alpha,2}-1)$ and $\sqrt{r}(F_{R,\gamma,2}-1)$ when r is large. Note that $F_{R,\alpha,2}$ and $F_{R,\gamma,2}$ reduce to the usual F-ratios in balanced case.

Theorem 2.2.1. *(**Unbalanced case**.) Let $Var(Y_{ijk}) = \sigma_{ij}^2$.*

(a) For $n_{ij} \geq 2$ fixed, we have

$$\text{under} \qquad H_0(\alpha), \quad \sqrt{r}(F_{R,\alpha,2}-1) \xrightarrow{d} N(0, \frac{2(\phi^4+\eta^4)}{c\nu_4^4}) \text{ as } r \to \infty.$$

$$\text{under} \qquad H_0(\gamma), \quad \sqrt{r}(F_{R,\gamma,2}-1) \xrightarrow{d} N(0, \frac{2\phi^4(c-1)^2+2\eta^4}{c\nu_4^4}) \text{ as } r \to \infty,$$

where

$$\nu_4^2 = \lim_{r\to\infty} \frac{1}{rc}\sum_{i=1}^{r}\sum_{j=1}^{c}\frac{\sigma_{ij}^2}{n_{ij}}, \quad \phi^4 = \lim_{r\to\infty} \frac{1}{rc}\sum_{i=1}^{r}\sum_{j=1}^{c}\frac{\sigma_{ij}^4}{n_{ij}(n_{ij}-1)} \quad \eta^4 = \lim_{r\to\infty} \frac{1}{rc}\sum_{i=1}^{r}\sum_{j_1\neq j_2}^{c}\frac{\sigma_{ij_1}^2\sigma_{ij_2}^2}{n_{ij_1}n_{ij_2}},$$

assuming the limits exist.

(b) If $n_{ij} = n_{ij}(r) \to \infty$, set $n(r) = \min\{n_{ij}(r), i = 1,\cdots,r, j = 1,\cdots,c\}$, $\kappa(r) = \max\{n_{ij}(r), i = 1,\cdots,r, j = 1,\cdots,c\}$, and assume that

$$n(r) \to \infty, \text{ and } \kappa(r) - n(r) \leq C(r), \text{ for all } r,$$

where $C(r) = o(n(r))$, as $r \to \infty$. In addition, assume that for some $\delta > 0$, $n(r) = o(r^{\delta/(4+2\delta)})$. Then

$$\text{under} \qquad H_0(\alpha), \quad \sqrt{r}(F_{R,\alpha,2}-1) \xrightarrow{d} N(0, \frac{\tau_4}{\nu_5^4}) \text{ as } r \to \infty.$$

$$\text{under} \qquad H_0(\gamma), \quad \sqrt{r}(F_{R,\gamma,2}-1) \xrightarrow{d} N(0, \frac{\tau_4+\tau_5}{\nu_5^4}) \text{ as } r \to \infty,$$

where

$$\nu_5^2 = \lim_{r\to\infty} \frac{n(r)}{rc}\sum_{i=1}^{r}\sum_{j=1}^{c}\frac{\sigma_{ij}^2}{n_{ij}}, \tau_4 = \lim_{r\to\infty} \frac{2}{rc^2}\sum_{i=1}^{r}\left(\sum_{j=1}^{c}\sigma_{ij}^2\right)^2, \tau_5 = \lim_{r\to\infty} \frac{2(c-2)}{rc}\sum_{i=1}^{r}\sum_{j=1}^{c}\sigma_{ij}^4,$$

assuming the limits exist.

Corollary 2.2.2. *(**Balanced case**.) Assume* $r^{-1}c^{-1}\sum_{i=1}^{r}\sum_{j=1}^{c}\sigma_{ij}^2 \to \nu_{b,4}^2$.

(a) For $n \geq 2$ *fixed we have*

$$\text{under} \qquad H_0(\alpha), \quad \sqrt{r}(F_{R,\alpha,2}-1) \xrightarrow{d} N(0, \frac{2(\phi_b^4+\eta_b^4)}{c\nu_{b,4}^4}) \text{ as } r \to \infty.$$

$$\text{under} \qquad H_0(\gamma), \quad \sqrt{r}(F_{R,\gamma,2}-1) \xrightarrow{d} N(0, \frac{2\phi_b^4(c-1)^2+2\eta_b^4}{c\nu_{b,4}^4}) \text{ as } r \to \infty,$$

where

$$\phi_b^4 = \lim_{r\to\infty} \frac{n}{rc(n-1)} \sum_{i=1}^{r}\sum_{j=1}^{c} \sigma_{ij}^4, \;\; \eta_b^4 = \lim_{r\to\infty} \frac{1}{rc} \sum_{i=1}^{r} \sum_{j_1 \neq j_2}^{c} \sigma_{ij_1}^2 \sigma_{ij_2}^2,$$

assuming the limits exist.

(b) If $n(r) \to \infty$*, As* $r \to \infty$*, and for some* $\delta > 0$*,* $n(r) = o(r^{\delta/(4+2\delta)})$*, we have*

$$\text{under} \qquad H_0(\alpha), \quad \sqrt{r}(F_{R,\alpha,2}-1) \xrightarrow{d} N(0, \frac{\tau_4}{\nu_{b,4}^4}) \text{ as } r \to \infty.$$

$$\text{under} \qquad H_0(\gamma), \quad \sqrt{r}(F_{R,\gamma,2}-1) \xrightarrow{d} N(0, \frac{\tau_4+\tau_5}{\nu_{b,4}^4}) \text{ as } r \to \infty,$$

where τ_4 *and* τ_5 *are defined in Theorem 2.2.1*

2.3 Simulation Results

The simulations reported in this section pertain only to the two-way ANOVA design. We compare the asymptotic tests based on the original observations (Wang and Akritas, 2003) with the present rank tests. Type I error-rate results are reported for both the row main effect and interaction effect, with the number of row factor levels taking values $r = 10$, 15, 20, 25. Results for the achieved power are reported only for testing for row main effects with $r = 20$. The number of column factors is $c = 2$ for all simulations. The simulations are based on 2,000 replications and use the normal, exponential, log-normal and Cauchy distributions. The elements of the asymptotic variance were estimated in the obvious way except for σ_{ij}^4 where we used an unbiased estimator as suggested in Akritas and Papadatos (2002).

Tables 1 and 3 report the achieved α-levels and power, respectively, for the balanced case with $n = 4$. Tables 2 and 4 report the achieved α-levels and power, respectively, for the unbalanced case. When $r = 10$, the sample sizes in the unbalanced case are $\mathbf{n}_1 = c(4,5,4,6,5,6,4,5,4,4)$, $\mathbf{n}_2 = c(4,4,4,4,5,4,4,5,6,5)$; when $r = 15$, $\mathbf{n}_1 = (5,4,4,4,4,4,6,4,4,5,5,4,4,5,4)$, and $\mathbf{n}_2 = (4,4,4,4,4,4,7,4,4,4,4,4,4,5,4)$; when $r = 20$, $\mathbf{n}_1 = (4,4,4,4,4,4,4,4,6,6,4,4,5,4,4,4,4,4,4,5)$, and $\mathbf{n}_2 = (4,4,4,5,4,6,4,5,4,5,4,4,4,4,5,4,4,4,4,4)$; when $r = 25$, $\mathbf{n}_1 = (6,4,5,4,5,4,4,4,5,4,6,4,4,5,4,4,4,4,4,4,4,4,4,4,4)$, and $\mathbf{n}_2 = (6,4,6,4,4,4,4,4,6,6,6,4,4,4,4,4,5,4,4,4,4,4,6,4,5)$. For the power comparison, only $r = 20$ is reported.

In all cases observations under the alternative were generated as $X_{ijk} = i * \tau/r + \text{error}$, where the error term and τ are specified at the tables.

From these tables it is easily seen that the Type I error rate is more stable for the rank test, and the power of the rank test is much better than that based on the original observations away from the normal distribution.

2.4 Proofs

2.4.1 Proofs for the one-way design

Proof of Theorem 2.1.1 Because $MS\alpha_R$ here is the same as in the unbalanced heteroscedastic case, several results/lemmas from the proof of Theorem 2.1.3 will be used.

$$\sqrt{r}(F_{R,r} - 1) = \frac{\sqrt{r}}{MSE_R}(MS\alpha_R - MSE_R)$$

The proof of $MSE_R/N^2 \xrightarrow{p} \sigma^2$ is similar to the proof of Lemma 2.4.2 below, and thus the proof is omitted. So we only need to consider the asymptotic distribution of $\sqrt{r}(MS\alpha_R - MSE_R)/N^2$.

Table 2.1: Estimated level for balanced case, $\alpha = 0.05$, n=4

error	r	WA-test		Rank Test	
		$H_0(\alpha)$	$H_0(\gamma)$	$H_0(\alpha)$	$H_0(\gamma)$
normal(0,1)	10	0.0520	0.0465	0.0600	0.0565
	15	0.0430	0.0405	0.0505	0.0415
	20	0.0415	0.0435	0.0480	0.0460
	25	0.0345	0.0360	0.0375	0.0450
exp(1)	10	0.0340	0.0390	0.0495	0.0590
	15	0.0255	0.0240	0.0510	0.0455
	20	0.0265	0.0280	0.0535	0.0485
	25	0.0300	0.0235	0.0485	0.0425
lognormal(0,1)	10	0.0185	0.0185	0.0580	0.0540
	15	0.0125	0.0135	0.0420	0.0460
	20	0.0085	0.0110	0.0490	0.0455
	25	0.0070	0.0095	0.0375	0.0450
Cauchy	10	0.0040	0.0050	0.0660	0.0600
	15	0.0025	0.0015	0.0435	0.0480
	20	0.0005	0.0015	0.0550	0.0445
	25	0.0000	0.0005	0.0435	0.0440

Table 2.2: Estimated level for unbalanced case, $\alpha = 0.05$

r	error	WA-test		Rank Test	
		$H_0(\alpha)$	$H_0(\gamma)$	$H_0(\alpha)$	$H_0(\gamma)$
10	normal(0,1)	0.0565	0.0545	0.0630	0.0575
	exp(1)	0.0325	0.0325	0.0635	0.0530
	lognormal(0,1)	0.0160	0.0260	0.0540	0.0600
	Cauchy	0.0055	0.0045	0.0505	0.0585
15	normal(0,1)	0.0455	0.0485	0.0485	0.0495
	exp(1)	0.0250	0.0390	0.0530	0.0585
	lognormal(0,1)	0.0060	0.0110	0.0415	0.0570
	Cauchy	0.0000	0.0010	0.0590	0.0495
20	normal(0,1)	0.0430	0.0405	0.0500	0.0425
	exp(1)	0.0245	0.0270	0.0470	0.0585
	lognormal(0,1)	0.0080	0.0120	0.0515	0.0385
	Cauchy	0.0000	0.0000	0.0450	0.0475
25	normal(0,1)	0.0495	0.0335	0.0505	0.0405
	exp(1)	0.0215	0.0275	0.0400	0.0415
	lognormal(0,1)	0.0095	0.0070	0.0505	0.0405
	Cauchy	0.0005	0.0010	0.0405	0.0420

Table 2.3: Achieved power for balanced case, $\alpha = 0.05$, $r = 20$, $n = 4$

error	τ	WA-test	Rank test
normal(0,1)	0	0.0375	0.0460
	0.5	0.1080	0.1190
	1.0	0.4585	0.4580
	1.5	0.8950	0.8870
exp(1)	0	0.0275	0.0505
	0.5	0.0665	0.2490
	1.0	0.4045	0.8280
	1.5	0.8395	0.9915
lognormal(0,1)	0	0.011	0.0460
	1	0.041	0.5835
	2	0.322	0.9895
	3	0.721	1

Table 2.4: Achieved power for unbalanced case, $\alpha = 0.05$, $r = 20$,

error	τ	WA-test	Rank test
normal	0	0.011	0.0460
	0.5	0.041	0.5835
	1.0	0.322	0.9895
	1.5	0.721	1
exp(1)	0	0.0285	0.0505
	0.5	0.089	0.27
	1.0	0.457	0.8545
	1.5	0.8895	0.9955
lognormal(0,1)	0	0.0125	0.043
	1	0.048	0.591
	2	0.3555	0.995
	3	0.751	1

Let $P_{R,\alpha}$ be defined in (2.4.2). Then Lemma 2.4.3 implies that under $H_{0,1}(\alpha)$,

$$\sqrt{r}(MS\alpha_R - P_{R,\alpha})/N^2 \xrightarrow{p} 0 \text{ as } r \to \infty.$$

So $\sqrt{r}(MS\alpha_R - MSE_R)/N^2$ has same asymptotic distribution as $\sqrt{r}(P_{R,\alpha} - MSE_R)/N^2$. Let MSE_Z be similarly defined as MSE_R with R_{ij} replaced by Z_{ij}. Then

$$\begin{aligned} MSE_R/N^2 &= MSE_Z = \frac{1}{N-r}\sum_{i=1}^{r}\sum_{j=1}^{n_i}\left(Z_{ij} - p_i + p_i - \overline{Z}_{i.}\right)^2 \\ &= \frac{1}{N-r}\sum_{i=1}^{r}\sum_{j=1}^{n_i}(Z_{ij} - p_i)^2 - \frac{1}{N-r}\sum_{i=1}^{r} n_i(\overline{Z}_{i.} - p_i)^2 \\ &= \frac{1}{N-r}\sum_{i=1}^{r}\sum_{j=1}^{n_i}\frac{n_i - 1}{n_i}(Z_{ij} - p_i)^2 - \frac{1}{N-r}\sum_{i=1}^{r}\frac{1}{n_i}\sum_{j\neq j'}^{n_i}(Z_{ij} - p_i)(Z_{ij'} - p_i), \end{aligned}$$

and

$$P_{R,\alpha}/N^2 = P_{Z,\alpha} = \frac{1}{r-1}\sum_{i=1}^{r} n_i\left(1 - \frac{n_i}{N}\right)\left(\overline{Z}_{i.} - p_i\right)^2$$

$$= \frac{1}{r-1}\sum_{i=1}^{r}\frac{1}{n_i}\left(1 - \frac{n_i}{N}\right)\sum_{j\neq j'}^{n_i}(Z_{ij} - p_i)(Z_{ij'} - p_i) + \frac{1}{r-1}\sum_{i=1}^{r}\sum_{j=1}^{n_i}\frac{1}{n_i}\left(1 - \frac{n_i}{N}\right)(Z_{ij} - p_i)^2.$$

Hence

$$\sqrt{r}(P_{R,\alpha} - MSE_R)/N^2 = \sqrt{r}(P_{Z,\alpha} - MSE_Z) = T_3(\mathbf{Z}) + T_4(\mathbf{Z}),$$

where

$$\begin{aligned} T_3(\mathbf{Z}) &= \sqrt{r}\sum_{i=1}^{r}\sum_{j\neq j'}^{n_i}\left(d_i + \frac{1}{N-r}\right)(Z_{ij} - p_i)(Z_{ij'} - p_i), \\ T_4(\mathbf{Z}) &= \sqrt{r}\sum_{i=1}^{r}\sum_{j=1}^{n_i} d_i(Z_{ij} - p_i)^2, \end{aligned} \tag{2.4.1}$$

with

$$d_i = \frac{N-1}{(N-r)(r-1)n_i} - \frac{1}{N(r-1)} - \frac{1}{N-r}.$$

$T_3(\mathbf{Y})$ and $T_4(\mathbf{Y})$ are similarly defined with Z_{ij} replaced by Y_{ij}. Following the same procedure as in the proof of Lemma 2.4.4, it can be shown that $T_3(\mathbf{Z}) - T_3(\mathbf{Y}) = o_p(1)$ as $r \to \infty$. Lemma

2.4.1 shows that $T_4(\mathbf{Z}) - T_4(\mathbf{Y}) = o_p(1)$. Therefore, $\sqrt{r}(P_{R,\alpha} - MSE_R)/N^2$ has same asymptotic distribution as $T_3(\mathbf{Y}) + T_4(\mathbf{Y})$. Asymptotic normality of $T_3(\mathbf{Y}) + T_4(\mathbf{Y})$ is shown in the proof of Theorem 3.2 of Akritas and Papadatos (2002), by noting that it can be written as $\sqrt{r}\mathbf{Y}'\mathbf{A}_D\mathbf{Y}$, where $\mathbf{Y} = (Y_{11}, \ldots, Y_{1n_1}, \ldots, Y_{rn_r})$ and $\mathbf{A}_D$ is defined in the aforementioned reference. This completes the proof of the theorem.

Lemma 2.4.1. *Let $T_4(\mathbf{Z})$ be defined in (2.4.1). Under $H_{0,1}(\alpha)$, we have $T_4(\mathbf{Z}) - T_4(\mathbf{Y}) \xrightarrow{p} 0$, as $r \to \infty$.*

Proof. Let $H_{0,1}(\alpha)$ hold, and H denote the common distribution in all categories, note that $p_i = E(H(X_{i1}))$ are all equal to 0.5. Define

$$\begin{aligned} & h(X_{i_1,j_1}, X_{i_2,j_2}, X_{i,j}) \\ = \ & \left[c(X_{i_1,j_1}, X_{i,j}) - p_{i_1}\right]\left[c(X_{i_2,j_2}, X_{i,j}) - p_{i_2}\right] - \left[H(X_{ij}) - p_{i_1}\right]\left[H(X_{ij}) - p_{i_2}\right]. \end{aligned}$$

Then, under $H_{0,1}(\alpha)$,

$$T_4(\mathbf{Z}) - T_4(\mathbf{Y}) = \sqrt{r}\sum_{i=1}^{r}\sum_{j=1}^{n_i} d_i N^{-2} \sum_{i_1=1}^{r}\sum_{j_1=1}^{n_{i_1}}\sum_{i_2=1}^{r}\sum_{j_2=1}^{n_{i_2}} h(X_{i_1,j_1}, X_{i_2,j_2}, X_{i,j}).$$

If any two or all three pairs in $\{(i_1, j_1), (i_2, j_2)\,(i, j)\}$ are same, the summation is $o_p(1)$ since $h(X_{i_1,j_1}, X_{i_2,j_2}, X_{i,j})$ is uniformly bounded and $d_i = O(r^{-1})$. So

$$T_4(\mathbf{Z}) - T_4(\mathbf{Y}) = \sqrt{r}N^{-2} \sum_{(i_1,j_1)\neq(i_2,j_2)\neq(i,j)} d_i h(X_{i_1,j_1}, X_{i_2,j_2}, X_{i,j}) + o_p(1).$$

To show that the above converges to zero, we will consider its projection. Letting

$$\widetilde{h}_1(x) = E(h(X_{i_1,j_1}, X_{i_2,j_2}, X_{i,j})|X_{i_1,j_1} = x) = E[(H(X_{11}) - p_1)(c(x, X_{11}) - H(X_{11}))],$$

we have that for $(i_1, j_1) \neq (i_2, j_2) \neq (i, j)$, $E(h(X_{i_1,j_1}, X_{i_2,j_2}, X_{i,j})|X_{i,j}) = 0$,

$$E(h(X_{i_1,j_1}, X_{i_2,j_2}, X_{i,j})|X_{i_1,j_1}) = \widetilde{h}_1(X_{i_1j_1}), \ \ E(h(X_{i_1,j_1}, X_{i_2,j_2}, X_{i,j})|X_{i_2,j_2}) = \widetilde{h}_1(X_{i_2j_2}),$$

and if (i_3, j_3) is different from these three pairs, $E(h(X_{i_1,j_1}, X_{i_2,j_2}, X_{i,j})|X_{i_3,j_3}) = 0$. So under $H_{0,1}(\alpha)$,

$$E(T_4(\mathbf{Z}) - T_4(\mathbf{Y})|X_{i_3,j_3}) = \frac{2\sqrt{r}}{N} \sum_{i=1}^{r} \sum_{j=1}^{n_i} d_i \widetilde{h}_1(X_{i_3 j_3}),$$

and the projection of $T_4(\mathbf{Z}) - T_4(\mathbf{Y})$ is

$$\begin{aligned}
\widehat{T}_{4ZY} &= \sum_{i_3=1}^{r} \sum_{j=1}^{n_{i_3}} E(T_4(\mathbf{Z}) - T_4(\mathbf{Y})|X_{i_3 j_3}) - (N-1)E(T_4(\mathbf{Z}) - T_4(\mathbf{Y})) \\
&= \frac{2\sqrt{r}}{N} \sum_{i_3=1}^{r} \sum_{j_3=1}^{n_{i_3}} \sum_{i=1}^{r} \sum_{j=1}^{n_i} d_i \widetilde{h}_1(X_{i_3 j_3}) \\
&= \frac{\sqrt{r}}{N^2} \sum_{i_1=1}^{r} \sum_{j_1=1}^{n_{i_1}} \sum_{i_2=1}^{r} \sum_{j_2=1}^{n_{i_2}} \sum_{i=1}^{r} \sum_{j=1}^{n_i} d_i [\widetilde{h}_1(X_{i_1 j_1}) + \widetilde{h}_1(X_{i_2 j_2})].
\end{aligned}$$

Thus,

$$\begin{aligned}
T_4(\mathbf{Z}) - T_4(\mathbf{Y}) - \widehat{T}_{4ZY} &= \frac{\sqrt{r}}{N^2} \sum_{i_1=1}^{r} \sum_{j_1=1}^{n_{i_1}} \sum_{i_2=1}^{r} \sum_{j_2=1}^{n_{i_2}} \sum_{i=1}^{r} \sum_{j=1}^{n_i} d_i h^*(X_{i_1,j_1}, X_{i_2,j_2}, X_{i,j}) \\
&= \frac{\sqrt{r}}{N^2} \sum_{(i_1,j_1) \neq (i_2,j_2) \neq (i,j)} d_i h^*(X_{i_1,j_1}, X_{i_2,j_2}, X_{i,j}) + o_p(1),
\end{aligned}$$

where

$$h^*(X_{i_1,j_1}, X_{i_2,j_2}, X_{i,j}) = h(X_{i_1,j_1}, X_{i_2,j_2}, X_{i,j}) - \widetilde{h}_1(X_{i_1 j_1}) - \widetilde{h}_1(X_{i_2 j_2}).$$

Therefore,

$$E\left(\frac{\sqrt{r}}{N^2} \sum_{(i_1,j_1) \neq (i_2,j_2) \neq (i,j)} d_i h^*(X_{i_1,j_1}, X_{i_2,j_2}, X_{i,j})\right)^2 = \frac{r}{N^4} \sum_{(i_1,j_1) \neq (i_2,j_2) \neq (i,j)}$$
$$\sum_{(i_3,j_3) \neq (i_4,j_4) \neq (i_5,j_5)} d_i d_{i_5} E\left[h^*(X_{i_1,j_1}, X_{i_2,j_2}, X_{i,j}) h^*(X_{i_3,j_3}, X_{i_4,j_4}, X_{i_5,j_5})\right].$$

Note that $E(h^*(X_{i_1,j_1}, X_{i_2,j_2}, X_{i,j})|X_{i_3,j_3}) = 0$, for all indices (i_3, j_3). Thus, if the number of different pairs in $\{(i_1, j_1), (i_2, j_2), (i, j), (i_3, j_3), (i_4, j_4), (i_5, j_5)\}$ is six or five, the expectation under the summation is zero. It follows that $E\left(T_4(\mathbf{Z}) - T_4(\mathbf{Y}) - \widehat{T}_{4ZY}\right)^2 = O(r^{-1})$ and thus

$T_4(\mathbf{Z}) - T_4(\mathbf{Y}) - \widehat{T}_{4ZY} \xrightarrow{p} 0$. However, $\sum_{i=1}^{r}\sum_{j=1}^{n_i} d_i = 0$ implies that $\widehat{T}_{4ZY} = 0$. So $T_4(\mathbf{Z}) - T_4(\mathbf{Y}) \xrightarrow{p} 0$.

Proof of Theorem 2.1.3 The proof uses three lemmas which are stated and proved after the proof of the theorem. By Lemma 2.4.2, $MSE_R^*/N^2 \xrightarrow{p} \sigma_0^2$. So we only need to consider the asymptotic distribution of $\sqrt{r}(MS\alpha_R - MSE_R^*)/N^2$. Set

$$P_{R,\alpha} = (r-1)^{-1}\sum_{i=1}^{r} n_i\left(1-\frac{n_i}{N}\right)\left(\overline{R}_{i.} - \mu_i\right)^2, \tag{2.4.2}$$

where $\mu_i = Np_i + 0.5$. By Lemma 2.4.3, we have that under $H_{0,2}(\alpha)$,

$$\sqrt{r}(MS\alpha_R - P_{R,\alpha})/N^2 \xrightarrow{p} 0 \text{ as } r \to \infty.$$

So $\sqrt{r}(MS\alpha_R - MSE_R^*)/N^2$ has same asymptotic distribution as $\sqrt{r}(P_{R,\alpha} - MSE_R^*)/N^2$. Set

$$MSE_R^*/N^2 = MSE_Z^*,\ \ P_{R,\alpha}/N^2 = P_{Z,\alpha},\ \ S_{R,i}^2/N^2 = S_{Z,i}^2,$$

where MSE_R^*, and $S_{R,i}^2$ are defined in (2.1.2). MSE_Z^* and $S_{Z,i}^2$ are defined similarly. $P_{Z,\alpha}$ is defined similarly as $P_{R,\alpha}$.

For simplicity of notation, denote $T_P(\mathbf{Z}) = \sqrt{r}(P_{Z,\alpha} - MSE_Z^*)$, and use $T_P(\mathbf{Y})$ to denote the quantity similarly defined but with Z_{ij} replaced by Y_{ij}. Lemma 2.4.4 shows that $T_P(\mathbf{Z}) - T_P(\mathbf{Y}) \xrightarrow{p} 0$ as $r \to \infty$. The asymptotic distribution of $T_P(\mathbf{Y})$ can be found by applying results in Akritas and Papadatos (2002) or we can verify Lyapounov's condition easily here because Y_{ij} are uniformly bounded. Indeed,

$$\begin{aligned} T_P(\mathbf{Y}) &= \frac{\sqrt{r}}{r-1}\sum_{i=1}^{r}\left(1-\frac{n_i}{N}\right)\left[n_i\left(\overline{Y}_{i.} - p_i\right)^2 - S_{Y,i}^2\right] \\ &= \frac{\sqrt{r}}{r-1}\sum_{i=1}^{r}\left(1-\frac{n_i}{N}\right)\sum_{j_1\neq j_2}^{n_i}\left(\overline{Y}_{ij_1} - p_i\right)\left(\overline{Y}_{ij_2} - p_i\right) = \frac{\sqrt{r}}{r-1}\sum_{i=1}^{r} W_i^*, \end{aligned}$$

where $W_i^* = (1-n_i/N)\sum_{j\neq j'} e_{ij}e_{ij'}$ and $e_{ij} = Y_{ij} - p_i$. It is easy to see that W_i^*'s are independent and satisfy

$$E(W_i^*) = 0 \text{ and } \mathrm{Var}(W_i^*) = (n_i-1)^{-2}\sum_{j\neq j'}\sum_{j_1\neq j_1'} E(e_{ij}e_{ij'}e_{ij_1}e_{ij_1'}) = \frac{n_i}{n_i-1}2\sigma_i^4$$

Because Y_{ij} are nonnegative and uniformly bounded by 1, using the inequality

$$\left|\sum_{i=1}^{m} z_i\right|^p \leq m^{p-1}\sum_{i=1}^{m}|z_i|^p,\ m \geq 1,\ p \geq 1,$$

which for $p > 1$ follows from Hölder's inequality, we have

$$\begin{aligned} E(W_i^*)^{2+\delta} &= \left(1-\frac{n_i}{N}\right)^{2+\delta} E\left[\left(n_i\left(\overline{Y}_{i.}-\mu_i\right)^2 - S_{Y,i}^2\right)^{2+\delta}\right] \\ &\leq 2^{1+\delta}\left(1-\frac{n_i}{N}\right)^{2+\delta}\left[n_i^{2+\delta}E\left(\overline{Y}_{i.}-\mu_i\right)^{4+2\delta} + E\left(S_{Y,i}^{4+2\delta}\right)\right] \\ &\leq 2^{1+\delta}\left(1-\frac{n_i}{N}\right)^{2+\delta}\left(n_i^{2+\delta} + \left(\frac{n_i}{n_i-1}\right)^{2+\delta}\right) \leq 2^{3+2\delta} n_i^{2+\delta} \end{aligned}$$

By assumption,

$$\frac{1}{\sqrt{\sum_{i=1}^{r}\mathrm{Var}(W_i^*)}^{2+\delta}}\sum_{i=1}^{r}E(W_i^*)^{2+\delta} \leq \frac{1}{\sqrt{\sum_{i=1}^{r}2\sigma_i^4}^{2+\delta}}\sum_{i=1}^{r}(4n_i)^{2+\delta} = O(r^{-\delta/2}) \to 0,$$

as r go to infinity. So the asymptotic distribution of $T_P(\mathbf{Y})$ follows.

Lemma 2.4.2. *For the one-way heteroscedastic model,*

$$MSE_R^*/N^2 \xrightarrow{p} \sigma_0^2 = \lim_{r\to\infty}\frac{1}{r}\sum_{i=1}^{r}\left(1-\frac{n_i}{N}\right)\sigma_i^2,\ \textit{as } r\to\infty.$$

Proof: The proof will follow if we show that

$$MSE_R^*/N^2 - MSE_Y^* \xrightarrow{p} 0,\ \text{ as } r\to\infty, \text{ and} \qquad (2.4.3)$$

$$MSE_Y^* \xrightarrow{p} \sigma_0^2,\ \text{as } r\to\infty. \qquad (2.4.4)$$

We have

$$\begin{aligned} & MSE_R^*/N^2 - MSE_Y^* \\ = & \frac{1}{r-1}\sum_{i=1}^{r}\left(1-\frac{n_i}{N}\right)\frac{1}{n_i-1}\sum_{j=1}^{n_i}\left[(Z_{ij}-\overline{Z}_{i.})^2 - (Y_{ij}-\overline{Y}_{i.})^2\right] \\ = & \frac{1}{r-1}\sum_{i=1}^{r}\left(1-\frac{n_i}{N}\right)\frac{1}{n_i-1}\sum_{j=1}^{n_i}(Z_{ij}-\overline{Z}_{i.}-Y_{ij}+\overline{Y}_{i.})^2 \\ & +\frac{2}{r-1}\sum_{i=1}^{r}\left(1-\frac{n_i}{N}\right)\frac{1}{n_i-1}\sum_{j=1}^{n_i}(Z_{ij}-\overline{Z}_{i.}-Y_{ij}+\overline{Y}_{i.})(Y_{ij}-\overline{Y}_{i.}). \end{aligned}$$

The first summation is $O_p(N^{-1}) = o_p(1)$ since $\sup_x(\widehat{H}(x) - H(x)) = O_p(N^{-1/2})$. The second summation is bounded by

$$\frac{1}{r-1}\sum_{i=1}^{r}\left(1-\frac{n_i}{N}\right)\frac{1}{n_i-1}\sum_{j=1}^{n_i}|Z_{ij}-\overline{Z}_{i.}-Y_{ij}+\overline{Y}_{i.}| = O_p(N^{-1/2}) = o_p(1).$$

So (2.4.3) is shown. Next, note that $0 \le S^2_{Y,i} \le n_i/(n_i-1)$, so $\mathrm{Var}(S^2_{Y,i}) \le E(S^2_{Y,i})^2 \le 4$. Therefore, when $r \to \infty$,

$$E(MSE^*_Y) = \frac{1}{r-1}\sum_{i=1}^{r}\left(1-\frac{n_i}{N}\right)\sigma_i^2 \to \sigma_0^2$$

$$\mathrm{Var}((MSE^*_Y) = \frac{1}{(r-1)^2}\sum_{i=1}^{r}\left(1-\frac{n_i}{N}\right)^2\mathrm{Var}(S^2_{Y,i}) \le \frac{4r}{(r-1)^2} \to 0.$$

Thus (2.4.4) is also true and this finishes the proof.

Lemma 2.4.3. *Let $MS\alpha_R$ be defined in (2.1.1) and $P_{R,\alpha}$ be defined in (2.4.2). Then under $H_{0,1}(\alpha)$ or $H_{0,2}(\alpha)$, $\sqrt{r}(MS\alpha_R - P_{R,\alpha})/N^2 \xrightarrow{p} 0$ as $r \to \infty$.*

Proof: Under $H_{0,1}(\alpha)$ or $H_{0,2}(\alpha)$, $p_i = \overline{p}$,

$$MS\alpha_R/N^2 = \frac{1}{r-1}\sum_{i=1}^{r}\sum_{j=1}^{n_i}(\overline{Z}_{i.}-\overline{Z}_{..})^2 = \frac{1}{r-1}\sum_{i=1}^{r}\sum_{j=1}^{n_i}\left((\overline{Z}_{i.}-p_i)-(\overline{Z}_{..}-\overline{p})\right)^2$$

$$= \frac{1}{r-1}\sum_{i=1}^{r}n_i\left(1-\frac{n_i}{N}\right)\left(\overline{Z}_{i.}-p_i\right)^2 - \frac{1}{N(r-1)}\sum_{i\neq i'}^{r}n_i n_i'\left(\overline{Z}_{i.}-p_i\right)\left(\overline{Z}_{i'.}-p_{i'}\right).$$

So

$$\begin{aligned}\sqrt{r}(MS\alpha_R/N^2 - P_{R,\alpha}) &= -\frac{\sqrt{r}}{N(r-1)}\sum_{i\neq i'}^{r}n_i n_i'\left(\overline{Z}_{i.}-p_i\right)\left(\overline{Z}_{i'.}-p_{i'}\right) \\ &= -\frac{\sqrt{r}}{N(r-1)}\sum_{i\neq i'}^{r}\sum_{j=1}^{c}\sum_{j'=1}^{c}\left(Z_{ij}-p_i\right)\left(Z_{i'j'}-p_{i'}\right).\end{aligned}$$

Denote above quantity as $T(\mathbf{Z})$ and we show that $T(\mathbf{Z}) - T(\mathbf{Y}) = o_p(1)$ as $r \to \infty$, where $T(\mathbf{Y})$ is similarly defined as $T(\mathbf{Z})$ with Z_{ij} replaced by Y_{ij}.

$$\begin{aligned}T(\mathbf{Z}) &= -\frac{\sqrt{r}}{N(r-1)}\sum_{i\neq i_1}^{r}\sum_{j=1}^{n_i}\sum_{j_1=1}^{n_i}\left(Z_{ij}-Y_{ij}+Y_{ij}-p_i\right)\left(Z_{i_1j_1}-Y_{i_1j_1}+Y_{i_1j_1}-p_{i_1}\right) \\ &= D_1 + D_2 + T(\mathbf{Y}),\end{aligned}$$

where

$$D_1 = -\frac{\sqrt{r}}{N(r-1)} \sum_{i \neq i_1}^{r} \sum_{j=1}^{n_i} \sum_{j_1=1}^{n_{i_1}} \left(Z_{ij} - Y_{ij}\right)\left(Z_{i_1 j_1} - Y_{i_1 j_1}\right)$$

$$D_2 = -\frac{2\sqrt{r}}{N(r-1)} \sum_{i \neq i_1}^{r} \sum_{j=1}^{n_i} \sum_{j_1=1}^{n_{i_1}} \left(Z_{ij} - Y_{ij}\right)\left(Y_{i_1 j_1} - p_{i_1}\right).$$

Since $\sup_x(\widehat{H}(x) - H(x)) = O_p(N^{-1/2})$, we have $D_1 = O_p(r^{-1/2}) = o_p(1)$. Next,

$$\begin{aligned}
&\frac{N^2(r-1)^2}{4r} E(D_2^2) \\
= &\sum_{i \neq i_1}^{r} \sum_{j=1}^{n_i} \sum_{j_1=1}^{n_{i_1}} \sum_{i' \neq i'_1}^{r} \sum_{j'=1}^{n_{i'}} \sum_{j'_1=1}^{n_{i'_1}} E\left[\left(Z_{ij} - Y_{ij}\right)\left(Y_{i_1 j_1} - p_{i_1}\right)\left(Z_{i'j'} - Y_{i'j'}\right)\left(Y_{i'_1 j'_1} - p_{i'_1}\right)\right] \\
= &\sum_{i \neq i_1}^{r} \sum_{j=1}^{n_i} \sum_{j_1=1}^{n_i} \sum_{i' \neq i'_1}^{r} \sum_{j'=1}^{n_{i'}} \sum_{j'_1=1}^{n_{i'_1}} \frac{1}{N^2} \sum_{i_2=1}^{r} \sum_{j_2=1}^{n_{i_2}} \sum_{i_3=1}^{r} \sum_{j_3=1}^{n_{i_3}} \\
&E\left[\left(c(X_{i_2 j_2}, X_{ij}) - F_{i_2 j_2}(X_{ij})\right)\left(c(X_{i_3 j_3}, X_{i'j'}) - F_{i_3 j_3}(X_{i'j'})\right)\left(Y_{i_1 j_1} - p_{i_1}\right)\left(Y_{i'_1 j'_1} - p_{i'_1}\right)\right].
\end{aligned}$$

The expectation under the summation is zero if the number of different elements in $\{i, i_1, i', i'_1, i_2, i_3\}$ is five or six or the number of different elements in $\{j, j_1, j', j'_1, j_2, j_3\}$ is five or six. Also note that $c(X_{ij}, X_{i'j'})$, Y_{ij} and $F_{ij}(X)$ are all uniformly bounded by 1, so $E(D_2^2) = O(r^{-1}) \to 0$. Hence $D_2 = o_p(1)$. Therefore it remains to show that $T(\mathbf{Y}) = o_p(1)$. This can be shown easily by verifying $E(T(\mathbf{Y})) = 0$ and $\mathrm{Var}(T(\mathbf{Y})) \to 0$ due to the fact that Y_{ij}'s are independent uniformly bounded random variables.

Lemma 2.4.4. *Assume $r^{-1}\sum_{i=1}^{r} n_{ij}^2$ converges and $T_P(\mathbf{Z})$ and $T_P(\mathbf{Y})$ are defined in the proof of Theorem 2.1.3. We have $T_P(\mathbf{Z}) - T_P(\mathbf{Y}) = o_p(1)$ as $r \to \infty$.*

Proof:

$$\begin{aligned}
T_P(\mathbf{Z}) &= \frac{\sqrt{r}}{r-1} \sum_{i=1}^{r} \left(1 - \frac{n_i}{N}\right) \sum_{j_1 \neq j_2}^{n_i} \left(Z_{ij_1} - p_i\right)\left(Z_{ij_2} - p_i\right) \\
&= \frac{\sqrt{r}}{r-1} \sum_{i=1}^{r} \left(1 - \frac{n_i}{N}\right) \sum_{j_1 \neq j_2}^{n_i} \left(Z_{ij_1} - Y_{ij_1} + Y_{ij_1} - p_i\right)\left(Z_{ij_2} - Y_{ij_2} + Y_{ij_2} - p_i\right) \\
&= D_{p1} + D_{p2} + T_P(\mathbf{Y}),
\end{aligned}$$

where

$$D_{p1} = \frac{\sqrt{r}}{r-1}\sum_{i=1}^{r}\left(1-\frac{n_i}{N}\right)\sum_{j_1\neq j_2}^{n_i}\left(Z_{ij_1}-Y_{ij_1}\right)\left(Z_{ij_2}-Y_{ij_2}\right)$$

$$D_{p2} = \frac{\sqrt{r}}{r-1}\sum_{i=1}^{r}\left(1-\frac{n_i}{N}\right)\sum_{j_1\neq j_2}^{n_i}\left(Z_{ij_1}-Y_{ij_1}\right)\left(Y_{ij_2}-p_i\right).$$

It remains to show that $D_{p1} \xrightarrow{p} 0$ and $D_{p2} \xrightarrow{p} 0$ as $r \to \infty$. $D_{p1} = O_p(r^{-1/2}N^{-1}\sum_{i=1}^{r} n_i^2) = O_p(\sqrt{r}N^{-1}) = o_p(1)$ due to the fact that $\sup_x(\widehat{H}(x) - H(x)) = O_p(N^{-1/2})$. Next,

$$D_{p2} = \frac{\sqrt{r}}{r-1}\sum_{i=1}^{r}\left(1-\frac{n_i}{N}\right)\sum_{j_1\neq j_2}^{n_i}\left(Y_{ij_2}-p_i\right)\frac{1}{N}\sum_{i_3=1}^{r}\sum_{j_3=1}^{n_{i_3}}\left(c(X_{i_3,j_3},X_{ij_1})-F_{i_3}(X_{ij_1})\right)$$

and

$$\begin{aligned} & E(D_{p2}^2) \\ = & \frac{r}{(r-1)^2}\sum_{i=1}^{r}\sum_{i'=1}^{r}\sum_{j_1\neq j_2}^{n_i}\sum_{j_1'\neq j_2'}^{n_{i'}}\frac{1}{N^2}\sum_{i_3=1}^{r}\sum_{j_3=1}^{n_{i_3}}\sum_{i_4=1}^{r}\sum_{j_4=1}^{n_{i_4}}\left(1-\frac{n_i}{N}\right)\left(1-\frac{n_{i'}}{N}\right) \\ & E\left[\left(Y_{ij_2}-p_i\right)\left(Y_{i'j_2'}-p_i'\right)\left(c(X_{i_3,j_3},X_{ij_1})-F_{i_3}(X_{ij_1})\right)\left(c(X_{i_4,j_4},X_{i'j_1'})-F_{i_4}(X_{i'j_1'})\right)\right]. \end{aligned}$$

Note that $j_1 \neq j_2$, $j_1' \neq j_2'$ and

$$E\left[c(X_{i_1,j_1},X_{ij})-F_{i_1}(X_{ij})\right] = E\left\{E[c(X_{i_1,j_1},X_{ij})-F_{i_1}(X_{ij})|X_{ij}]\right\} = 0.$$

So by independence, the expectation under the summation is zero if the number of different elements in $\{i,i',i_3,i_4\}$ is three or four or the number of different elements in $\{j_1,j_2,j_1',j_2',j_3,j_4\}$ is five or six. Also note that $c(X_{ij},X_{i'j'})$, Y_{ij} and $F_i(X)$ are all uniformly bounded by 1, so $E(D_2^2) = O(r^{-1}N^{-2}(\sum_{i=1}^{r} n_i^2)^2) = O(r^{-1}) \to 0$. Hence $D_{p2} = o_p(1)$.

Proof of Theorem 2.1.5 The proof uses three lemmas which are stated and proved after the proof of the theorem. By Lemma 2.4.5, $n(r)MSE_R^{(2)}/N^2$ converges in probability to a constant. So we only need to consider the asymptotic distribution of $\sqrt{r}(MST_R - MSE_R^{(2)})/N^2$ when n_i remain fixed and $n(r)\sqrt{r}(MST_R - MSE_R^{(2)})/N^2$ when $n_i \to \infty$ with r. Under $H_{0,2}(\alpha)$ or

$H_{0,2}(\alpha)$,

$$MST_R/N^2 = \frac{1}{r-1}\sum_{i=1}^{r}\left(\overline{Z}_{i.} - \widetilde{Z}_{..}\right)^2 = \frac{1}{r-1}\sum_{i=1}^{r}\left((\overline{Z}_{i.} - p_i) - (\widetilde{Z}_{..} - \overline{p})\right)^2 = P(\mathbf{Z}) + D_T(\mathbf{Z}),$$

where

$$P(\mathbf{Z}) = \frac{1}{r}\sum_{i=1}^{r}\left(\overline{Z}_{i.} - p_i\right)^2 \tag{2.4.5}$$

$$D_T(\mathbf{Z}) = -\frac{1}{r(r-1)}\sum_{i\neq i'}\left(\overline{Z}_{i.} - p_i\right)\left(\overline{Z}_{i'.} - p_i'\right). \tag{2.4.6}$$

By Lemma 2.4.6, it suffices to consider the asymptotic distribution of

$$T_2(\mathbf{Z}) = \sqrt{r}\left(P(\mathbf{Z}) - MSE_R^{(2)}/N^2\right) = \frac{1}{\sqrt{r}}\sum_{i=1}^{r}\frac{1}{n_i(n_i-1)}\sum_{j_1\neq j_2}^{n_i}(Z_{ij_1} - p_i)(Z_{ij_2} - p_i) \tag{2.4.7}$$

when n_i fixed and that of $n(r)T_2(\mathbf{Z})$ when $n_i \to \infty$ with r. Let $T_2(\mathbf{Y})$ be defined similarly as $T_2(\mathbf{Z})$ with Z_{ij} replaced by Y_{ij}. By Lemma 2.4.7, it remains to find the asymptotic distribution of $T_2(\mathbf{Y})$ and $n(r)T_2(\mathbf{Y})$ for fixed n_i and $n_i \to \infty$ with r, respectively. Let $e_{ij} = Y_{ij} - p_i$ and $W_i = \frac{1}{n_i(n_i-1)}\sum_{j_1\neq j_2}^{n_i} e_{ij_1}e_{ij_2}$, then W_i's are independent with zero mean and $T_2(\mathbf{Y}) = r^{-1/2}\sum_{i=1}^{r} W_i$. We will verify Lyapounov's condition. First,

$$\begin{aligned}
\mathrm{Var}(W_i) &= \frac{1}{n_i^2(n_i-1)^2}\sum_{j\neq j'}^{n_i}\sum_{j_1\neq j_1'}^{n_i} E(e_{ij}e_{ij'}e_{ij_1}e_{ij_1'}) \\
&= \frac{2}{n_i^2(n_i-1)^2}\sum_{j\neq j'} E(e_{ij}^2)E(e_{ij'}^2) \\
&= \frac{2}{n_i(n_i-1)}\sigma_i^4,
\end{aligned}$$

so that

$$E(n(r)T_2(\mathbf{Y})) = 0 \text{ and } \mathrm{Var}(n(r)T_2(\mathbf{Y})) = \frac{2}{r}\sum_{i=1}^{r}\frac{n^2(r)}{n_i(n_i-1)}\sigma_i^4$$

Note that from the original definition of $T_2(\mathbf{Y})$ we can also write W_i as $(\overline{Y}_{i.} - p_i)^2 - n_i^{-1}S_{Y,i}^2$. so by the fact that Y_{ij} are nonnegative and uniformly bounded by 1,

$$\begin{aligned} E(W_i)^{2+\delta} &= E\left[\left(\overline{Y}_{i.}-p_i\right)^2 - n_i^{-1}S_{Y,i}^2\right]^{2+\delta} \\ &\leq 2^{1+\delta}\left[E\left(\overline{Y}_{i.}-p_i\right)^{4+2\delta} + n_i^{-2-\delta}E\left(S_{Y,i}^{4+2\delta}\right)\right] \\ &\leq 2^{1+\delta}\left[1+\left(\frac{4}{n_i}\right)^{2+\delta}\right] \leq 2^{2+\delta}. \end{aligned}$$

By assumption,

$$\begin{aligned} &\frac{1}{\sqrt{\sum_{i=1}^r \mathrm{Var}[n(r)W_i]}^{2+\delta}} \sum_{i=1}^r E[n(r)W_i]^{2+\delta} \\ \leq\ & \frac{1}{\sqrt{\sum_{i=1}^r \frac{2n^2(r)\sigma_i^4}{n_i(n_i-1)}}^{2+\delta}} \sum_{i=1}^r (2n(r))^{2+\delta} \\ =\ & O(n(r)^{2+\delta} r^{-\delta/2}) \to 0, \end{aligned}$$

as r go to infinity. So the asymptotic distribution of $n(r)T_2(\mathbf{Y})$ follows.

If n_i fixed, treat $n(r)$ as a bounded constant, we get the asymptotic distribution of $T_2(\mathbf{Y})$.

Lemma 2.4.5. *Assume $r^{-1}\sum_{i=1}^r \frac{n(r)}{n_i}\sigma_i^2 \to \nu_3^2$, as $r\to\infty$, and let $MSE_R^{(2)}$ be defined in Theorem 2.1.5. Then $MSE_R^{(2)}/N^2$ converges in probability to ν_2^2, if the n_i are fixed, and $n(r)MSE_R^{(2)}/N^2$ converges in probability to ν_3^2, if the n_i tend to infinity.*

Proof: Note that $MSE_R^{(2)}/N^2 = MSE_Z^{(2)}$. First we will show that, in both cases,

$$n(r)(MSE_Z^{(2)} - MSE_Y^{(2)}) \xrightarrow{p} 0, \quad \text{as } r\to\infty. \tag{2.4.8}$$

We have

$$
\begin{aligned}
& n(r)(MSE_R^{(2)}/N^2 - MSE_Y^{(2)}) \\
= \quad & \frac{n(r)}{r}\sum_{i=1}^{r}\frac{1}{n_i(n_i-1)}\sum_{j=1}^{n_i}\left[(Z_{ij}-\overline{Z}_{i.})^2-(Y_{ij}-\overline{Y}_{i.})^2\right] \\
= \quad & \frac{n(r)}{r}\sum_{i=1}^{r}\frac{1}{n_i(n_i-1)}\sum_{j=1}^{n_i}(Z_{ij}-\overline{Z}_{i.}-Y_{ij}+\overline{Y}_{i.})^2 \\
& +\frac{2n(r)}{r}\sum_{i=1}^{r}\frac{1}{n_i(n_i-1)}\sum_{j=1}^{n_i}(Z_{ij}-\overline{Z}_{i.}-Y_{ij}+\overline{Y}_{i.})(Y_{ij}-\overline{Y}_{i.}).
\end{aligned}
$$

The first summation is $O_p(N^{-1}) = o_p(1)$ since $\sup_x(\widehat{H}(x) - H(x)) = O_p(N^{-1/2})$. The second summation is bounded by

$$
\frac{2n(r)}{r}\sum_{i=1}^{r}\frac{1}{n_i(n_i-1)}\sum_{j=1}^{n_i}|Z_{ij}-\overline{Z}_{i.}-Y_{ij}+\overline{Y}_{i.}| = O_p(N^{-1/2}) = o_p(1).
$$

So (2.4.8) holds, in both cases. Next we will show

$$
MSE_Y^{(2)} \xrightarrow{p} \nu_2^2, \text{ as } r\to\infty, \text{ and } n(r)MSE_Y^{(2)} \xrightarrow{p} \nu_3^2, \text{ if } n_i\to\infty \text{ as } r\to\infty. \tag{2.4.9}
$$

Note that $0 \le S_{Y,i}^2 \le n_i/(n_i-1)$, so $\mathrm{Var}(S_{Y,i}^2) \le E(S_{Y,i}^2)^2 \le 4$. Therefore, when $r \to \infty$,

$$
\begin{aligned}
E(n(r)MSE_Y^{(2)}) &= \frac{1}{r}\sum_{i=1}^{r}\frac{n(r)}{n_i}\sigma_i^2 \to \nu_3^2 \\
\mathrm{Var}((MSE_Y^{(2)}) &= \frac{n^2(r)}{r^2}\sum_{i=1}^{r}\frac{1}{n_i^2}\mathrm{Var}(S_{Y,i}^2) \le \frac{4}{r} \to 0.
\end{aligned}
$$

The case that n_i are fixed is similar. So (2.4.9) is true. Combine (2.4.8) and (2.4.9), we finish the proof of the lemma.

Lemma 2.4.6. *$D_T(\mathbf{Z})$ is defined in (2.4.6).*

1. *$\sqrt{r}D_T(\mathbf{Z}) \xrightarrow{p} 0$ as $r\to\infty$ while n_i's remain fixed;*

2. *When $n_i = n_i(r) \to \infty$, as $r\to\infty$, assume $n^2(r) = o(r)$, then $n(r)\sqrt{r}D_T(\mathbf{Z}) \xrightarrow{p} 0$.*

Proof: When $n_i = n_i(r) \to \infty$ as $r \to \infty$, we will show

$$n(r)\sqrt{r}\left[D_T(\mathbf{Z}) - D_T(\mathbf{Y})\right] \xrightarrow{p} 0 \text{ and } n(r)\sqrt{r}D_T(\mathbf{Y}) \xrightarrow{p} 0, \qquad (2.4.10)$$

where $D_T(\mathbf{Y})$ is similarly defined as $D_T(\mathbf{Z})$ with Z_{ij} replaced by Y_{ij}.

$$\begin{aligned} & n(r)\sqrt{r}D_T(\mathbf{Z}) \\ = \; & -\frac{n(r)\sqrt{r}}{r(r-1)} \sum_{i \neq i_1}^{r} \frac{1}{n_i n_{i_1}} \sum_{j=1}^{n_i} \sum_{j_1=1}^{n_{i_1}} \left(Z_{ij} - Y_{ij} + Y_{ij} - p_i\right)\left(Z_{i_1 j_1} - Y_{i_1 j_1} + Y_{i_1 j_1} - p_{i_1}\right) \\ = \; & D_{T1} + D_{T2} + n(r)\sqrt{r}D_T(\mathbf{Y}), \end{aligned}$$

where

$$\begin{aligned} D_{T1} &= -\frac{n(r)}{\sqrt{r}(r-1)} \sum_{i \neq i_1}^{r} \frac{1}{n_i n_{i_1}} \sum_{j=1}^{n_i} \sum_{j_1=1}^{n_{i_1}} \left(Z_{ij} - Y_{ij}\right)\left(Z_{i_1 j_1} - Y_{i_1 j_1}\right) \\ D_{T2} &= -\frac{2n(r)}{\sqrt{r}(r-1)} \sum_{i \neq i_1}^{r} \frac{1}{n_i n_{i_1}} \sum_{j=1}^{n_i} \sum_{j_1=1}^{n_{i_1}} \left(Z_{ij} - Y_{ij}\right)\left(Y_{i_1 j_1} - p_{i_1}\right) \end{aligned}$$

Since $\sup_x(\widehat{H}(x) - H(x)) = O_p(N^{-1/2})$, we have $D_{T1} = O_p(n(r)\sqrt{r}N^{-1}) = o_p(1)$.

$$\begin{aligned} & E(D_{T2}^2) \\ = \; & \frac{4n^2(r)}{r(r-1)^2} \sum_{i \neq i_1}^{r} \frac{1}{n_i n_{i_1}} \sum_{j=1}^{n_i} \sum_{j_1=1}^{n_{i_1}} \sum_{i' \neq i'_1}^{r} \frac{1}{n_{i'} n_{i'_1}} \sum_{j'=1}^{n_{i'}} \sum_{j'_1=1}^{n_{i'_1}} \\ & \qquad E\left[\left(Z_{ij} - Y_{ij}\right)\left(Y_{i_1 j_1} - p_{i_1}\right)\left(Z_{i'j'} - Y_{i'j'}\right)\left(Y_{i'_1 j'_1} - p_{i'_1}\right)\right] \\ = \; & \frac{4n^2(r)}{r(r-1)^2} \sum_{i \neq i_1}^{r} \frac{1}{n_i n_{i_1}} \frac{1}{n_{i'} n_{i'_1}} \sum_{j=1}^{n_i} \sum_{j_1=1}^{n_i} \sum_{i' \neq i'_1}^{r} \sum_{j'=1}^{n_{i'}} \sum_{j'_1=1}^{n_{i'_1}} \frac{1}{N^2} \sum_{i_2=1}^{r} \sum_{j_2=1}^{n_{i_2}} \sum_{i_3=1}^{r} \sum_{j_3=1}^{n_{i_3}} \\ & E\left[\left(c(X_{i_2 j_2}, X_{ij}) - F_{i_2 j_2}(X_{ij})\right)\left(c(X_{i_3 j_3}, X_{i'j'}) - F_{i_3 j_3}(X_{i'j'})\right)\left(Y_{i_1 j_1} - p_{i_1}\right)\left(Y_{i'_1 j'_1} - p_{i'_1}\right)\right] \end{aligned}$$

The expectation under the summation is zero if the number of different elements in $\{i, i_1, i', i'_1, i_2, i_3\}$ is five or six or the number of different elements in $\{j, j_1, j', j'_1, j_2, j_3\}$ is five or six. Also note that $c(X_{ij}, X_{i'j'})$, Y_{ij} and $F_{ij}(X)$ are all uniformly bounded by 1, so

$E(D_{T2}^2) = O_(n^2(r)rN^{-2}) \to 0$. Hence $D_{T2} = o_p(1)$. Therefore $n(r)\sqrt{r}[D_T(\mathbf{Z}) - D_T(\mathbf{Y})] \xrightarrow{p} 0$. It remains to show that $n(r)\sqrt{r}D_T(\mathbf{Y}) = o_p(1)$. This can be shown easily because $E(D_T(\mathbf{Y})) = 0$ and

$$
\begin{aligned}
& \operatorname{Var}\left(n(r)\sqrt{r}D_T(\mathbf{Y})\right) \\
= \;& \frac{n^2(r)}{r(r-1)^2} \sum_{i \neq i'}^{r} \sum_{i_1 \neq i_1'}^{r} E\left[\left(\overline{Z}_{i.} - p_i\right)\left(\overline{Z}_{i'.} - p_{i'}\right)\left(\overline{Z}_{i_1.} - p_{i_1}\right)\left(\overline{Z}_{i_1'.} - p_{i_1'}\right)\right] \\
= \;& \frac{n^2(r)}{r(r-1)^2} \sum_{i \neq i'}^{r} E\left(\overline{Z}_{i.} - p_i\right)^2 E\left(\overline{Z}_{i'.} - p_{i'}\right)^2 \\
= \;& O\left(\frac{n^2(r)}{r}\right) \to 0
\end{aligned}
$$

since Z_{ij}'s are independent uniformly bounded random variables.

Therefore (2.4.10) holds and $n(r)\sqrt{r}D_T(\mathbf{Z}) \xrightarrow{p} 0$ as $r \to \infty$.

When n_i's are fixed, treat $n(r)$ as a bounded constant, then we have $\sqrt{r}D_T(\mathbf{Z}) \xrightarrow{p} 0$ as $r \to \infty$.

Lemma 2.4.7. *$T_2(\mathbf{Z})$ and $T_2(\mathbf{Y})$ are defined in (2.4.7).*

1. *$T_2(\mathbf{Z}) - T_2(\mathbf{Y}) \xrightarrow{p} 0$ as $r \to \infty$ while n_i's remain fixed;*
2. *$n(r)(T_2(\mathbf{Z}) - T_2(\mathbf{Y})) \xrightarrow{p} 0$ when $n_i = n_i(r) \to \infty$, as $r \to \infty$.*

Proof: When $n_i = n_i(r) \to \infty$ as $r \to \infty$,

$$
\begin{aligned}
n(r)T_2(\mathbf{Z}) &= \frac{n(r)}{\sqrt{r}} \sum_{i=1}^{r} \frac{1}{n_i(n_i-1)} \sum_{j_1 \neq j_2}^{n_i} \left(Z_{ij_1} - Y_{ij_1} + Y_{ij_1} - p_i\right)\left(Z_{ij_2} - Y_{ij_2} + Y_{ij_2} - p_i\right) \\
&= D_{21} + D_{22} + n(r)T_2(\mathbf{Y}),
\end{aligned}
$$

where

$$
\begin{aligned}
D_{21} &= \frac{n(r)}{\sqrt{r}} \sum_{i=1}^{r} \frac{1}{n_i(n_i-1)} \sum_{j_1 \neq j_2}^{n_i} \left(Z_{ij_1} - Y_{ij_1}\right)\left(Z_{ij_2} - Y_{ij_2}\right) \\
D_{22} &= \frac{n(r)}{\sqrt{r}} \sum_{i=1}^{r} \frac{1}{n_i(n_i-1)} \sum_{j_1 \neq j_2}^{n_i} \left(Z_{ij_1} - Y_{ij_1}\right)\left(Y_{ij_2} - p_i\right)
\end{aligned}
$$

It remains to show that $D_{21} \xrightarrow{p} 0$ and $D_{22} \xrightarrow{p} 0$ as $r \to \infty$. $D_{21} = O_p(n(r)\sqrt{r}N^{-1}) = o_p(1)$ due to the fact that $\sup_x(\widehat{H}(x) - H(x)) = O_p(N^{-1/2})$.

$$D_{22} = \frac{n(r)}{\sqrt{r}} \sum_{i=1}^{r} \frac{1}{n_i(n_i-1)} \sum_{j_1 \neq j_2}^{n_i} \left(Y_{ij_2} - p_i\right) \frac{1}{N} \sum_{i_3=1}^{r} \sum_{j_3=1}^{n_{i_3}} \left(c(X_{i_3,j_3}, X_{ij_1}) - F_{i_3}(X_{ij_1})\right)$$

and

$$\begin{aligned} & E(D_{22}^2) \\ = \ & \frac{n^2(r)}{r} \sum_{i=1}^{r} \sum_{i'=1}^{r} \sum_{j_1 \neq j_2}^{n_i} \sum_{j_1' \neq j_2'}^{n_{i'}} \frac{1}{N^2} \sum_{i_3=1}^{r} \sum_{j_3=1}^{n_{i_3}} \sum_{i_4=1}^{r} \sum_{j_4=1}^{n_{i_4}} \frac{1}{n_i(n_i-1)} \frac{1}{n_{i'}(n_{i'}-1)} \\ & E\left[\left(Y_{ij_2} - p_i\right)\left(Y_{i'j_2'} - p_{i'}\right)\left(c(X_{i_3,j_3}, X_{ij_1}) - F_{i_3}(X_{ij_1})\right)\left(c(X_{i_4,j_4}, X_{i'j_1'}) - F_{i_4}(X_{i'j_1'})\right)\right] \end{aligned}$$

Note that $j_1 \neq j_2$, $j_1' \neq j_2'$ and

$$E\left[c(X_{i_1,j_1}, X_{ij}) - F_{i_1}(X_{ij})\right] = E\left\{E\left[c(X_{i_1,j_1}, X_{ij}) - F_{i_1}(X_{ij}) | X_{ij}\right]\right\} = 0.$$

So by independence, the expectation under the summation is zero if the number of different elements in $\{i, i', i_3, i_4\}$ is three or four or the number of different elements in $\{j_1, j_2, j_1', j_2', j_3, j_4\}$ is five or six. Also note that $c(X_{ij}, X_{i'j'})$, Y_{ij} and $F_i(X)$ are all uniformly bounded by 1, so $E(D_{22}^2) = O(n^2(r)rN^{-2}) \to 0$. Hence $D_{22} = o_p(1)$. Therefore, $n(r)(T_2(\mathbf{Z}) - T_2(\mathbf{Y})) \xrightarrow{p} 0$ when $n_i = n_i(r) \to \infty$, as $r \to \infty$.

When n_i's are fixed, treat $n(r)$ as a bounded constant, we get $T_2(\mathbf{Z}) - T_2(\mathbf{Y}) \xrightarrow{p} 0$.

2.4.2 Proofs for the two-way design

Proof of Theorem 2.2.1

By lemma 2.4.8, $MSE_R^{(3)}/N^2 \xrightarrow{p} \nu_4^2$ as $r \to \infty$ if n_{ij} *fixed*; $n(r)MSE_R^{(3)}/N^2 \xrightarrow{p} \nu_5^2$ as $r \to \infty$ if $n(r) \to \infty$ as $r \to \infty$. So we only need to consider $\sqrt{r}(MST_\alpha - MSE_R^{(3)})/N^2$. Note that

$R_{ijk} = 1/2 + NZ_{ijk}$, where $Z_{ijk} = \widehat{H}(X_{ijk})$. So

$$\begin{aligned}\sqrt{r}(MST_\alpha - MSE_R^{(3)})/N^2 &= \frac{\sqrt{r}}{r-1}\sum_{i=1}^{r}\sum_{j=1}^{c}\left(\widetilde{Z}_{i..} - \widetilde{Z}_{...}\right)^2 - \frac{\sqrt{r}}{rc}\sum_{i=1}^{r}\sum_{j=1}^{c}\frac{S_{ij,Z}^2}{n_{ij}} \\ &= \frac{\sqrt{r}}{r-1}\sum_{i=1}^{r}\sum_{j=1}^{c}\left[\left(\widetilde{Z}_{i..} - \widetilde{Z}_{...}\right)^2 - \frac{1}{c}\left(1-\frac{1}{r}\right)\frac{S_{ij,Z}^2}{n_{ij}}\right]\end{aligned}$$

Let

$$T_\alpha(\mathbf{Z}) = (rc)^{-1/2}\sum_{i=1}^{r}\sum_{j=1}^{c}\left[\left(\widetilde{Z}_{i..} - \widetilde{Z}_{...}\right)^2 - \frac{1}{c}\left(1-\frac{1}{r}\right)\frac{S_{ij,Z}^2}{n_{ij}}\right] \tag{2.4.11}$$

Then

$$\sqrt{r}(MST_\alpha - MSE_R^{(3)})/N^2 = \frac{r\sqrt{c}}{r-1}T_\alpha(\mathbf{Z})$$

Let $p_{ij} = E(Y_{ijk})$. The projection, under $H_0(\alpha)$, of $T_\alpha(\mathbf{Z})$ onto the class of random variables of the form $\sum_{i=1}^{r} g_i(\mathbf{Z}_i)$, where $\mathbf{Z}_i = (Z_{i11}, \ldots, Z_{i1n_{i1}}, \ldots, Z_{icn_{ic}})$ and g_i are measurable with $Eg_i^2(\mathbf{Z}_i) < \infty$, is given by

$$\widetilde{T}_\alpha(\mathbf{Z}) = (rc)^{-1/2}\sum_{i=1}^{r}\frac{r-1}{rc}\left[\left(\sum_{j=1}^{c}\left(\overline{Z}_{ij.} - p_{ij}\right)\right)^2 - \sum_{j=1}^{c}\frac{S_{ij,Z}^2}{n_{ij}}\right] \tag{2.4.12}$$

By Lemma 2.4.9 and 2.4.10, $T_\alpha(\mathbf{Z})$ has the same asymptotic distribution as $\widetilde{T}_\alpha(\mathbf{Y})$ when n_{ij} fixed, and $n(r)T_\alpha(\mathbf{Z})$ has the same asymptotic distribution as $n(r)\widetilde{T}_\alpha(\mathbf{Z})$ when $n_{ij}(r) \to \infty$ as $r \to \infty$. But the asymptotic distributions of $\widetilde{T}_\alpha(\mathbf{Y})$ when n_{ij} fixed and $n(r)\widetilde{T}_\alpha(\mathbf{Z})$ when $n_{ij}(r) \to \infty$ as $r \to \infty$ are established in Wang and Akritas (2002). Combining these results we establish the convergences under $H_0(\alpha)$ stated in parts (a) and (b) of the theorem.

Similarly, for the test of interaction effect, let

$$T_\gamma(\mathbf{Z}) = (rc)^{-1/2}\sum_{i=1}^{r}\sum_{j=1}^{c}\left[\left(\overline{Z}_{ij.} - \widetilde{Z}_{i..} - \widetilde{Z}_{.j.} + \widetilde{Z}_{...}\right)^2 - \frac{(r-1)(c-1)}{rc}\frac{S_{ij,Z}^2}{n_{ij}}\right], \tag{2.4.13}$$

so

$$\sqrt{r}(MST_\gamma - MSE_R^{(3)})/N^2 = \frac{r\sqrt{c}}{(r-1)(c-1)}T_\gamma(\mathbf{Z}).$$

The projection, under $H_0(\gamma)$, of $T_\gamma(\mathbf{Z})$ onto the same class of random variables defined above is given by

$$\widetilde{T}_\gamma(\mathbf{Z}) = \frac{(r-1)(c-1)}{(rc)^{3/2}} \sum_{i=1}^{r} \sum_{j=1}^{c} \left(\left(\overline{Z}_{ij.} - p_{ij}\right)^2 - \frac{S^2_{ij,Z}}{n_{ij}} \right) - \frac{r-1}{(rc)^{3/2}} \sum_{i=1}^{r} \sum_{j_1 \neq j_2}^{c} \left(\overline{Z}_{ij_1} - p_{ij_1}\right)\left(\overline{Z}_{ij_2} - p_{ij_2}\right). \quad (2.4.14)$$

By Lemma 2.4.9 and 2.4.10, $T_\gamma(\mathbf{Z})$ has the same asymptotic distribution as $\widetilde{T}_\gamma(\mathbf{Y})$ when n_{ij} fixed, and $n(r)T_\gamma(\mathbf{Z})$ has the same asymptotic distribution as $n(r)\widetilde{T}_\gamma(\mathbf{Z})$ when $n_{ij}(r) \to \infty$ as $r \to \infty$. But the asymptotic distributions of $\widetilde{T}_\gamma(\mathbf{Y})$ when n_{ij} fixed and $n(r)\widetilde{T}_\gamma(\mathbf{Z})$ when $n_{ij}(r) \to \infty$ as $r \to \infty$ are established in Wang and Akritas (2002). Combining these results we establish the convergences under $H_0(\gamma)$ stated in parts (a) and (b) of the theorem. This completes the proof.

Lemma 2.4.8. *$MSE_R^{(3)}/N^2 \xrightarrow{p} \nu_4^2$ as $r \to \infty$ if n_{ij} fixed; $n(r)MSE_R^{(3)}/N^2 \xrightarrow{p} \nu_5^2$ as $r \to \infty$ if $n(r) \to \infty$ as $r \to \infty$, where ν_4^2 and ν_5^2 are defined in Theorem 2.2.1.*

Proof: First we will show that, in both cases, $MSE_R^{(3)}/N^2 - MSE_Y^{(3)} = O_p\left(N^{-1/2}n(r)^{-1}\right)$, where $MSE_Y^{(3)}$ is defined similarly as $MSE_R^{(3)}$ with R_{ijk} replaced by Y_{ijk}. We have

$$\begin{aligned} & MSE_R^{(3)}/N^2 - MSE_Y^{(3)} \\ = & \frac{1}{rc}\sum_{i=1}^{r}\sum_{j=1}^{c}\frac{1}{n_{ij}(n_{ij}-1)}\sum_{k=1}^{n_{ij}}\left[(Z_{ijk}-\overline{Z}_{ij.})^2-(Y_{ijk}-\overline{Y}_{ij.})^2\right] \\ = & \frac{1}{rc}\sum_{i=1}^{r}\sum_{j=1}^{c}\frac{1}{n_{ij}(n_{ij}-1)}\sum_{k=1}^{n_{ij}}\left(Z_{ijk}-\overline{Z}_{ij.}-Y_{ijk}+\overline{Y}_{ij.}\right)^2 \\ & +\frac{1}{rc}\sum_{i=1}^{r}\sum_{j=1}^{c}\frac{1}{n_{ij}(n_{ij}-1)}\sum_{k=1}^{n_{ij}}\left(Z_{ijk}-\overline{Z}_{ij.}-Y_{ijk}+\overline{Y}_{ij.}\right)(Y_{ijk}-\overline{Y}_{ij.}) \end{aligned}$$

The first summation is $O_p(N^{-1}n(r)^{-1})$ since $sup_x|\widehat{H}(x) - H(x)| = O_p(N^{-1/2})$. The second summation is bounded by

$$\frac{1}{rc}\sum_{i=1}^{r}\sum_{j=1}^{c}\frac{1}{n_{ij}(n_{ij}-1)}\sum_{k=1}^{n_{ij}}\left|\left(Z_{ijk}-\overline{Z}_{ij.}-Y_{ijk}+\overline{Y}_{ij.}\right)\right| = O_p(N^{-1/2}n(r)^{-1}).$$

So

$$MSE_R^{(3)}/N^2 - MSE_Y^{(3)} = O_p(N^{-1/2}n(r)^{-1}), \tag{2.4.15}$$

as $r \to \infty$ whether n_{ij} is fixed or not. We will be done if we show that

$$MSE_Y^{(3)} \xrightarrow{p} \nu_4^2 \text{ if } n_{ij} \text{ remain fixed;} \qquad n(r)MSE_Y^{(3)} \xrightarrow{p} \nu_5^2 \text{ if } n(r) \to \infty \text{ with } r\ , \tag{2.4.16}$$

as $r \to \infty$.

Denote $V_i = c^{-1}\sum_{j=1}^{c} n_{ij}^{-1}(n_{ij}-1)^{-1}\sum_{k=1}^{n_{ij}}(Y_{ijk} - \overline{Y}_{ij.})^2$, then V_i's are independent random variables uniformly bounded by 1 and $MSE_Y^{(3)} = r^{-1}\sum_{i=1}^{r} V_i$.

If n_{ij} remain fixed,

$$\begin{aligned} E(MSE_Y^{(3)}) &= \frac{1}{rc}\sum_{i=1}^{r}\sum_{j=1}^{c}\frac{\sigma_{ij}^2}{n_{ij}} \to \nu_4^2 \\ \text{Var}(MSE_Y^{(3)}) &= \frac{1}{r^2}\sum_{i=1}^{r}\text{Var}(V_i) \le \frac{1}{r^2}\sum_{i=1}^{r}E(V_i^2) \le \frac{1}{r} \to 0. \end{aligned}$$

If $n(r) \to \infty$ as $a \to \infty$,

$$\begin{aligned} E\left(n(r)MSE_Y^{(3)}\right) &= \frac{n(r)}{rc}\sum_{i=1}^{r}\sum_{j=1}^{c}\frac{\sigma_{ij}^2}{n_{ij}} \to \nu_5^2 \\ \text{Var}\left(n(r)MSE_Y^{(3)}\right) &= \frac{n(r)}{r^2}\sum_{i=1}^{r}\text{Var}(V_i) \le \frac{n(r)}{r^2}\sum_{i=1}^{r}E(V_i^2) \le \frac{n(r)}{r} \to 0. \end{aligned}$$

Thus, (2.4.16) holds and together with (2.4.15) the proof of the lemma is done.

Lemma 2.4.9. *Let $T_\alpha(\mathbf{Z})$, $\widetilde{T}_\gamma(\mathbf{Z})$, $T_\gamma(\mathbf{Z})$ and $\widetilde{T}_\gamma(\mathbf{Z})$ be defined as in the proof of Theorem 2.2.1. Suppose $\frac{1}{rc}\sum_{i=1}^{r}\sum_{j=1}^{c}\sigma_{ij}^2 \to \sigma^2 > 0$, then*

(1) If n_{ij} is fixed,

$$T_\alpha(\mathbf{Z}) - \widetilde{T}_\alpha(\mathbf{Z}) \xrightarrow{p} 0 \text{ under } H_0(\alpha); \qquad T_\gamma(\mathbf{Z}) - \widetilde{T}_\gamma(\mathbf{Z}) \xrightarrow{p} 0 \text{ under } H_0(\gamma).$$

(2) If $n_{ij}(r) \to \infty$*, as* $r \to \infty$*, let* $n(r) = \min\{n_{ij}(r); i = 1, \cdots, r, j = 1, \cdots, c\}$*, assume* $n(r)/r \to 0$*, as* $r \to \infty$*, then*

$$n(r)(T_\alpha(\mathbf{Z}) - \widetilde{T}_\alpha(\mathbf{Z})) \xrightarrow{p} 0 \text{ under } H_0(\alpha); \qquad n(r)(T_\gamma(\mathbf{Z}) - \widetilde{T}_\gamma(\mathbf{Z})) \xrightarrow{p} 0 \text{ under } H_0(\gamma).$$

Proof: It is not hard to prove that

$$\begin{aligned}
& T_\alpha(\mathbf{Z}) - \widetilde{T}_\alpha(\mathbf{Z}) \\
= \; & -(rc)^{-3/2} \sum_{i \neq i'}^{r} \left(\sum_{j=1}^{c} \left(\overline{Z}_{ij.} - \mu_{ij} \right) \right) \left(\sum_{j'=1}^{c} \left(\overline{Z}_{i'j'.} - \mu_{i'j'} \right) \right) \\
= \; & -(rc)^{-3/2} \sum_{i \neq i'}^{r} \sum_{j=1}^{c} \sum_{j'=1}^{c} \left(\overline{Z}_{ij.} - \overline{Y}_{ij.} + \overline{Y}_{ij.} - \mu_{ij} \right) \left(\overline{Z}_{i'j'.} - \overline{Y}_{i'j'.} + \overline{Y}_{i'j'.} - \mu_{i'j'} \right) \\
= \; & D_{\alpha 1} + D_{\alpha 2} + D_{\alpha 3},
\end{aligned}$$

where

$$\begin{aligned}
D_{\alpha 1} &= -(rc)^{-3/2} \sum_{i \neq i'}^{r} \sum_{j=1}^{c} \sum_{j'=1}^{c} \left(\overline{Z}_{ij.} - \overline{Y}_{ij.} \right) \left(\overline{Z}_{i'j'.} - \overline{Y}_{i'j'.} \right) \\
D_{\alpha 2} &= -2(rc)^{-3/2} \sum_{i \neq i'}^{r} \sum_{j=1}^{c} \sum_{j'=1}^{c} \left(\overline{Z}_{ij.} - \overline{Y}_{ij.} \right) \left(\overline{Y}_{i'j'.} - \mu_{i'j'} \right) \\
D_{\alpha 3} &= -(rc)^{-3/2} \sum_{i \neq i'}^{r} \sum_{j=1}^{c} \sum_{j'=1}^{c} \left(\overline{Y}_{ij.} - \mu_{ij} \right) \left(\overline{Y}_{i'j'.} - \mu_{i'j'} \right)
\end{aligned}$$

$$\begin{aligned}
& T_\gamma(\mathbf{Z}) - \widetilde{T}_\gamma(\mathbf{Z}) \\
= \; & -\frac{c-1}{(rc)^{3/2}} \sum_{j=1}^{c} \sum_{i \neq i'}^{r} \left(\overline{Z}_{ij.} - \mu_{ij} \right) \left(\overline{Z}_{i'j.} - \mu_{i'j} \right) + (rc)^{-3/2} \sum_{i \neq i'}^{r} \sum_{j \neq j'}^{c} \left(\overline{Z}_{ij.} - \mu_{ij} \right) \left(\overline{Z}_{i'j'.} - \mu_{i'j'} \right) \\
= \; & -\frac{c-1}{(rc)^{3/2}} \sum_{j=1}^{c} \sum_{i \neq i'}^{r} \left(\overline{Z}_{ij.} - \overline{Y}_{ij.} + \overline{Y}_{ij.} - \mu_{ij} \right) \left(\overline{Z}_{i'j.} - \overline{Y}_{i'j.} + \overline{Y}_{i'j.} - \mu_{i'j} \right) + \\
& (rc)^{-3/2} \sum_{i \neq i'}^{r} \sum_{j \neq j'}^{c} \left(\overline{Z}_{ij.} - \overline{Y}_{ij.} + \overline{Y}_{ij.} - \mu_{ij} \right) \left(\overline{Z}_{i'j'.} - \overline{Y}_{i'j'.} + \overline{Y}_{i'j'.} - \mu_{i'j'} \right) \\
= \; & D_{\gamma 1} + D_{\gamma 2} + D_{\gamma 3},
\end{aligned}$$

where

$$D_{\gamma 1} = -\frac{c-1}{(rc)^{3/2}} \sum_{j=1}^{c} \sum_{i \neq i'}^{r} \left(\overline{Z}_{ij.} - \overline{Y}_{ij.}\right)\left(\overline{Z}_{i'j.} - \overline{Y}_{i'j.}\right) + (rc)^{-3/2} \sum_{i \neq i'}^{r} \sum_{j \neq j'}^{c} \left(\overline{Z}_{ij.} - \overline{Y}_{ij.}\right)\left(\overline{Z}_{i'j'.} - \overline{Y}_{i'j'.}\right)$$

$$D_{\gamma 2} = -2\frac{c-1}{(rc)^{3/2}} \sum_{j=1}^{c} \sum_{i \neq i'}^{r} \left(\overline{Z}_{ij.} - \overline{Y}_{ij.}\right)\left(\overline{Y}_{i'j.} - \mu_{i'j}\right) + 2(rc)^{-3/2} \sum_{i \neq i'}^{r} \sum_{j \neq j'}^{c} \left(\overline{Z}_{ij.} - \overline{Y}_{ij.}\right)\left(\overline{Z}_{i'j'.} - \mu_{i'j'}\right)$$

$$D_{\gamma 3} = -\frac{c-1}{(rc)^{3/2}} \sum_{j=1}^{c} \sum_{i \neq i'}^{r} \left(\overline{Y}_{ij.} - \mu_{ij}\right)\left(\overline{Y}_{i'j.} - \mu_{i'j}\right) + (rc)^{-3/2} \sum_{i \neq i'}^{r} \sum_{j \neq j'}^{c} \left(\overline{Y}_{ij.} - \mu_{ij}\right)\left(\overline{Y}_{i'j'.} - \mu_{i'j'}\right)$$

Because $\sup_x(\widehat{H}(x) - H(x)) = O_p(N^{-1/2})$, we have

$$D_{\alpha 1} = O_p\left(\frac{\sqrt{rc}}{N}\right); \qquad D_{\gamma 1} = O_p\left(\frac{\sqrt{rc}}{N}\right)$$

So if n_{ij} is fixed,

$$D_{\alpha 1} = o_p(1); \qquad D_{\gamma 1} = o_p(1) \text{ as } r \to \infty.$$

If $n_{ij}(r) \to \infty$ as $r \to \infty$,

$$n(r)D_{\alpha 1} = o_p(1) \text{ and } n(r)D_{\alpha 1} = o_p(1).$$

Note that $(Y_{ijk} - \mu_{ij})$'s are independent random variables with zero mean, apply Proposition 3.2 of Wang and Akritas (2002), $n(r)D_{\alpha 3} = o_p(1)$ and $n(r)D_{\gamma 3} = o_p(1)$ if $n_{ij}(r) \to \infty$ as $r \to \infty$; $D_{\alpha 3} = o_p(1)$ and $D_{\gamma 3} = o_p(1)$ if n_{ij} is fixed.

It remains to show $n(r)D_{\alpha 2} = o_p(1)$ and $n(r)D_{\gamma 2} = o_p(1)$. Since $|D_{\gamma 2}| \leq (c-1)|D_{\alpha 2}|$, we only need to show

$$n(r)D_{\alpha 2} = o_p(1) \text{if } n(r) \to \infty \text{ as } r \to \infty; \text{ and } D_{\alpha 2} = o_p(1) \text{ if } n_{ij} \text{ are fixed.} \qquad (2.4.17)$$

$$\begin{aligned}
D_{\alpha 2} &= -2(rc)^{-3/2}\sum_{i\neq i'}^{r}\sum_{j=1}^{c}\sum_{j'=1}^{c}\sum_{k=1}^{n_{ij}}\sum_{k'=1}^{n_{i'j'}}\frac{1}{n_{ij}n_{i'j'}}\left(Z_{ijk}-Y_{ijk}\right)\left(Y_{i'j'k'}-\mu_{i'j'}\right)\\
&= -2(rc)^{-3/2}\sum_{i\neq i'}^{r}\sum_{j=1}^{c}\sum_{j'=1}^{c}\sum_{k=1}^{n_{ij}}\sum_{k'=1}^{n_{i'j'}}\frac{1}{n_{ij}n_{i'j'}N}\sum_{i_1=1}^{r}\sum_{j_1=1}^{c}\sum_{k_1=1}^{n_{i_1j_1}}\\
&\qquad \left[c(X_{i_1j_1k_1},X_{ijk})-F_{i_1j_1}(X_{ijk})\right]\left(Y_{i'j'k'}-\mu_{i'j'}\right),
\end{aligned}$$

where $c(x,y)=[I(x\leq y)+I(x<y)]/2$. Hence,

$$\begin{aligned}
E(D_{\alpha 2}^2) &= \frac{4}{(rc)^3N^2}\sum_{i\neq i'}^{r}\sum_{j=1}^{c}\sum_{j'=1}^{c}\sum_{k=1}^{n_{ij}}\sum_{k'=1}^{n_{i'j'}}\sum_{i_1=1}^{r}\sum_{j_1=1}^{c}\sum_{k_1=1}^{n_{i_1j_1}}\frac{1}{n_{ij}n_{i'j'}}\\
&\quad \sum_{i_2\neq i_3}^{r}\sum_{j_2=1}^{c}\sum_{j_3=1}^{c}\sum_{k_2=1}^{n_{i_2j_2}}\sum_{k_3=1}^{n_{i_3j_3}}\sum_{i_4=1}^{r}\sum_{j_4=1}^{c}\sum_{k_4=1}^{n_{i_4j_4}}\frac{1}{n_{i_2j_2}n_{i_3j_3}}\\
&\quad E\left\{\left[c(X_{i_1j_1k_1},X_{ijk})-F_{i_1j_1}(X_{ijk})\right]\left(Y_{i'j'k'}-p_{i'j'}\right)\right.\\
&\qquad \left.\left[c(X_{i_4j_4k_4},X_{i_2j_2k_2})-F_{i_4j_4}(X_{i_2j_2k_2})\right]\left(Y_{i_3j_3k_3}-p_{i_3j_3}\right)\right\},
\end{aligned}$$

When five or six elements in the set $\{i,i',i_2,i_3,i_1,i_4\}$ are different, the expectation under the summation is zero. Therefore $E(D_{\alpha 2}^2)=O(r^{-1})$ and so $E(n(r)D_{\alpha 2}^2)=O(n(r)/r)$ if $n(r)\to\infty$ as $r\to\infty$.

By given condition, $n(r)/r\to 0$ as $r\to\infty$. So (2.4.17) is proved.

Lemma 2.4.10. *Let $\widetilde{T}_\gamma(\mathbf{Z})$ and $\widetilde{T}_\gamma(\mathbf{Z})$ be defined as in the proof of Theorem 2.2.1. $\widetilde{T}_\gamma(\mathbf{Y})$ and $\widetilde{T}_\gamma(\mathbf{Y})$ are defined by replacing Z_{ijk} with Y_{ijk} in the corresponding functions.*

(1) If n_{ij} is fixed,

$$\widetilde{T}_\alpha(\mathbf{Z})-\widetilde{T}_\alpha(\mathbf{Y})\xrightarrow{p}0;\qquad \widetilde{T}_\gamma(\mathbf{Z})-\widetilde{T}_\gamma(\mathbf{Y})\xrightarrow{p}0.$$

(2) If $n_{ij}(r)\to\infty$, as $r\to\infty$, let $n(r)=\min\{n_{ij}(r);i=1,\cdots,r,j=1,\cdots,c\}$,

$$n(r)(\widetilde{T}_\alpha(\mathbf{Z})-\widetilde{T}_\alpha(\mathbf{Y}))\xrightarrow{p}0;\qquad n(a)(\widetilde{T}_\gamma(\mathbf{Z})-\widetilde{T}_\gamma(\mathbf{Y}))\xrightarrow{p}0.$$

Proof: It can be shown that

$$\begin{aligned}\widetilde{T}_\alpha(\mathbf{Z}) &= \frac{r-1}{(rc)^{3/2}}\sum_{i=1}^{r}\sum_{j=1}^{c}\sum_{k_1\neq k_2}^{n_{ij}}\frac{(Z_{ijk_1}-\mu_{ij})(Z_{ijk_2}-\mu_{ij})}{n_{ij}(n_{ij}-1)} \\ &\quad+\frac{r-1}{(rc)^{3/2}}\sum_{i=1}^{r}\sum_{j_1\neq j_2}^{c}\sum_{k_1=1}^{n_{ij_1}}\sum_{k_2=1}^{n_{ij_2}}\frac{(Z_{ij_1k_1}-\mu_{ij_1})(Z_{ij_2k_2}-\mu_{ij_2})}{n_{ij_1}n_{ij_2}},\end{aligned}$$

and $\widetilde{T}_\alpha(\mathbf{Z})-\widetilde{T}_\alpha(\mathbf{Y})=\widetilde{D}_{\alpha1}+\widetilde{D}_{\alpha2}$, where

$$\begin{aligned}\widetilde{D}_{\alpha1} &= \frac{r-1}{(rc)^{3/2}}\sum_{i=1}^{r}\sum_{j=1}^{c}\sum_{k_1\neq k_2}^{n_{ij}}\frac{(Z_{ijk_1}-Y_{ijk_1})(Z_{ijk_2}-Y_{ijk_2})}{n_{ij}(n_{ij}-1)}+ \\ &\quad \frac{r-1}{(rc)^{3/2}}\sum_{i=1}^{r}\sum_{j_1\neq j_2}^{c}\sum_{k_1=1}^{n_{ij_1}}\sum_{k_2=1}^{n_{ij_2}}\frac{(Z_{ij_1k_1}-Y_{ij_1k_1})(Z_{ij_2k_2}-Y_{ij_2k_2})}{n_{ij_1}n_{ij_2}} \\ \widetilde{D}_{\alpha2} &= \frac{2(r-1)}{(rc)^{3/2}}\sum_{i=1}^{r}\sum_{j=1}^{c}\sum_{k_1\neq k_2}^{n_{ij}}\frac{(Z_{ijk_1}-Y_{ijk_1})(Y_{ijk_2}-\mu_{ij})}{n_{ij}(n_{ij}-1)}+ \\ &\quad \frac{2(r-1)}{(rc)^{3/2}}\sum_{i=1}^{r}\sum_{j_1\neq j_2}^{c}\sum_{k_1=1}^{n_{ij_1}}\sum_{k_2=1}^{n_{ij_2}}\frac{(Z_{ij_1k_1}-Y_{ij_1k_1})(Y_{ij_2k_2}-\mu_{ij_2})}{n_{ij_1}n_{ij_2}}\end{aligned}$$

Follow the same procedure as for $D_{\alpha1}$ and $D_{\alpha2}$ in the proof of lemma 2.4.9, we can show that

$n(r)\widetilde{D}_{\alpha_1}\xrightarrow{p}0$, if $n(r)\to\infty$ with a and $\widetilde{D}_{\alpha_1}\xrightarrow{p}0$ if n_{ij} fixed;

$n(r)\widetilde{D}_{\alpha_2}\xrightarrow{p}0$, if $n(r)\to\infty$ with a and $\widetilde{D}_{\alpha_2}\xrightarrow{p}0$ if n_{ij} fixed.

Then we finished the proof of the lemma.

Chapter 3

Analysis of Heteroscedastic Multifactor Designs with Large Number of Factor Levels

In this chapter, we will first develop testing procedures based on the original observations, for general (i.e. possibly unbalanced and heteroscedastic) multi-factor designs, when at least one of the factors has many levels, and the group sizes are either small or large. Due to the sensitivity of test statistics based on the original observations to outliers and poor performance when the underlying distribution is skewed or heavy-tailed, we also develop test statistics based on the overall (mid)-ranks of the data.

It should be pointed out that F-test statistics typically have no closed form expression in unbalanced designs. This is due to the lack of orthogonality in such designs. Thus, we face the problem of constructing suitable test statistics. In this regard, the proposed test statistics for the unbalanced heteroscedastic case are extensions of corresponding test statistics proposed in Akritas and Papadatos (2003) for one-factor designs. It should be mentioned that this idea is also related to the statistics proposed by Yates (1934) (also called the method of unweighted means, cf. Sahai and Ageel, 2000, p. 220-222) for unbalanced homoscedastic two-way designs. For economy in space, we study directly test procedures for the heteroscedastic case, which of course apply also in the homoscedastic case. In view of the findings in Akritas and Papadatos (2004), we do not compromise efficiency by applying a heteroscedastic procedure to homoscedastic data.

Rank test statistics are constructed by simply replacing the original observations by their (mid-)ranks in the corresponding test statistics. Of course the hypotheses tested by the rank statistics are different. Because of this difference all hypotheses in this chapter are stated in terms of the fully nonparametric formulation of Akritas and Arnold (1994). We remark that the hypotheses based on the decomposition of group means, which the statistics based on the original observations test, are implied by the fully nonparametric hypotheses.

The presentation of a general asymptotic testing theory in the case where some or all of the factors may have a large number of levels is considerably more complicated than the case of fixed number of factor levels and large group sizes. In the latter case, all test statistics can be expressed as a quadratic form involving contrasts of the finitely-many group means and thus one general theorem covers all testing problems; see Akritas, Arnold and Brunner (1997). The present test statistics will have either a limiting χ^2 distribution, or a limiting normal distribution, depending on the hypothesis tested and which of the factors have large number of levels. Of course, the arguments also depend on whether the group sizes are large or small. In order to achieve as concise a formulation as possible, we present the testing theory only for main effects, two- and three-way interactions. The treatment of these testing problems captures all ideas and techniques that would be required for developing test procedures for still higher-way interactions. Moreover, with this limited scope, it is possible to give a formulation that indexes the observations with at most five indices regardless of how many factors are present.

The main tools for our asymptotic development is Hájek's projection method for quadratic forms (Akritas and Papadatos, 2004) and the asymptotic rank transform (Akritas, 1990). The method of asymptotic rank transform consists in showing that the rank statistic is asymptotically equivalent to another statistic which is based on a non-random transformation of the data. Since the results for the statistics based on the original observations can be used for any non-random transformation of the data, the asymptotic theory for the rank statistics follows.

The rest of the chapter is organized as the following. Section 3.1 lays out the notation, describes the non-parametric model and hypotheses and presents the test statistics. Section 3.2 is devoted to results about the test statistics based on the original observations, while Section 3.3 does the same for the (mid)-rank test statistics. Simulation results are presented in Section 3.5.1, followed by the analysis of a real data set from a microarray experiment in Section 3.5.2. Finally the proofs are given in Section 3.6, with some more detailed derivations deferred to the Section 3.7.

3.1 Data representation and test statistics

In this section we want to give a representation of the data, arising from an arbitrary ANOVA design, which is as parsimonious with respect to the indexes used as possible. The fact that we can use less indexes than the number of factors involved is suggested by the fact that, for example, the procedure for testing for no main effects of a factor with many levels depends on whether the total number of levels of all other factors is large or small, and not on the (fixed) number of other factors. Thus for a general (i.e. encompassing all possible designs) testing theory for no main effects of a factor with many levels, it is sufficient to consider a two-way design. Similarly, economy in the number of indexes can be achieved for the other hypotheses we will consider. Below we outline the different testing problems, and give the corresponding parsimonious data representation and test statistics.

In all that follows, we use $c(x,y) = [I(x \le y) + I(x < y)]/2$, where $I(A)$ is the indicator function for the event A, and the following notation. The response variable will be denoted by X while X^* will denote a monotone transformation of X. In the case when there are four factors in the model, we define $\overline{X}_{ijkl.} = n_{ijkl}^{-1} \sum_{m=1}^{n_{ijkl}} X_{ijklm}$, $\widetilde{X}_{i....} = (bcd)^{-1} \sum_{j=1}^{b} \sum_{k=1}^{c} \sum_{l=1}^{d} \overline{X}_{ijkl.}$, $\widetilde{X}_{ij...} = (cd)^{-1} \sum_{k=1}^{c} \sum_{l=1}^{d} \overline{X}_{ijkl.}$, $\widetilde{X}_{.j...} = (acd)^{-1} \sum_{i=1}^{a} \sum_{k=1}^{c} \sum_{l=1}^{d} \overline{X}_{ijkl.}$, and $\widetilde{X}_{.....} = (abcd)^{-1} \sum_{i=1}^{a} \sum_{j=1}^{b} \sum_{k=1}^{c} \sum_{l=1}^{d} \overline{X}_{ijkl.}$. Finally, $S_{ijkl,X}^2 = (n_{ijkl} - 1)^{-1} \sum_{m=1}^{n_{ijkl}} (X_{ijklm} - \overline{X}_{ijkl.})^2$, $N = \sum_{i=1}^{a} \sum_{j=1}^{b} \sum_{k=1}^{c} \sum_{l=1}^{d} n_{ijkl}$ and $n(a,b,c,d) = \min\{n_{ijkl},\ i = 1,\dots,a, j = 1,\dots,b, k = 1,\dots,c, l = 1,\dots,d\}$. Similar notation applies when X is replaced by e, Y or Z. Analogous versions of this notation will also be used when there are three or two factors in the model.

3.1.1 Testing for no main effects

As already mentioned, the data representation for a general theory of testing for no main effects can use only three indexes. Thus, let

$$X_{ijk} \sim F_{ij}, \quad \text{for} \quad k = 1, \ldots, n_{ij}, \tag{3.1.1}$$

denote the k-th observation in cell (i, j), where $i = 1, \ldots, a$ enumerates the levels of factor A, whose main effects we want to test about, and $j = 1, \ldots, b$ enumerates the levels of factor B. As discussed above, factor B in this formulation is a composite factor whose levels consist of the level combinations of all factors other than A. Consider the decomposition

$$F_{ij} = M + A_i + B_j + (AB)_{ij}, \tag{3.1.2}$$

where $\sum_{i=1}^{a} A_i = \sum_{j=1}^{b} B_j = \sum_{i=1}^{a} (AB)_{ij} = \sum_{j=1}^{b} (AB)_{ij} = 0$. We want to test the nonparametric hypothesis of no main factor A effects, i.e. $H_0(A) : A_i = 0$, for all i.

For the case that a is small, we only consider the case that b is large since the case with b also small is the well studied case that requires a large number of replications in each cell. The test statistic is

$$Q_{X^*}(A) = N\mathbf{W}'\mathbf{C}_A'(\mathbf{C}_A\widehat{\mathbf{V}}\mathbf{C}_A')^{-1}\mathbf{C}_A\mathbf{W}, \tag{3.1.3}$$

where $\mathbf{W} = (\widetilde{X}^*_{1..}, \ldots, \widetilde{X}^*_{a..})'$, $\mathbf{C}_A = (\mathbf{1}_{a-1} | -I_{a-1})$, $\widehat{\mathbf{V}} = diag(\hat{\eta}_1, \ldots, \hat{\eta}_a)$, and $\hat{\eta}_i = \dfrac{N}{b^2} \sum_{j=1}^{b} \dfrac{S^2_{ij,X^*}}{n_{ij}}$ with $S^2_{ij,X^*} = (n_{ij} - 1)^{-1} \sum_{k=1}^{n_{ij}} (X^*_{ijk} - \overline{X^*}_{ij.})^2$.

Assume now that a is large. Then the statistic, which is the same regardless of whether b is small or large is given by

$$F_{X^*}(A) = \frac{MST_A}{MSE}, \tag{3.1.4}$$

where

$$MST_A = \frac{1}{a-1}\sum_{i=1}^{a}\sum_{j=1}^{b}\left(\widetilde{X}^*_{i\cdot\cdot} - \widetilde{X}^*_{\cdots}\right)^2, \quad MSE = \frac{1}{ab}\sum_{i=1}^{a}\sum_{j=1}^{b}\frac{S^2_{ij,X^*}}{n_{ij}}. \tag{3.1.5}$$

3.1.2 Testing for no two-way interaction effects

Let A, B denote the two factors whose interaction we are interested in testing. Testing for no A-B interaction involves the following three separate problems:

1. Both A and B have small number of levels. Then we will be interested only in the case that there exists another factor with many levels, for if not we are in the classical case of small number of factor levels.

2. Factor A has large and factor B has small number of levels. This testing procedure depends on whether the total number of levels of all other factors is large or small, and not on the (fixed) number of other factors.

3. Both A and B have large number of levels. Again, this testing procedure depends on whether the total number of levels of all other factors is large or small, and not on the (fixed) number of other factors.

Thus for a general (i.e. encompassing all possible designs, and all of the above cases) testing theory for no A-B interaction effect it is sufficient to consider a three-way design, which uses a total of four indexes. Let

$$X_{ijkm} \sim F_{ijk}, \text{ for } m = 1, \ldots, n_{ijk}, \tag{3.1.6}$$

denote the m-th observation in cell (i, j, k), where $i = 1, \ldots, a,\ j = 1, \ldots, b$ enumerate the levels of factors A and B, respectively, and k enumerates the levels of the composite factor C (i.e. the level combinations of all factors other than A and B). Consider the decomposition of the distribution function of X_{ijkm},

$$F_{ijk} = M + A_i + B_j + C_k + (AB)_{ij} + (AC)_{ik} + (BC)_{jk} + (ABC)_{ijk}, \tag{3.1.7}$$

which is unique under the restrictions $\sum_{i=1}^{a} A_i = \sum_{j=1}^{b} B_j = \sum_{k=1}^{c} C_k = \sum_{i=1}^{a} (AB)_{ij} = \sum_{i=1}^{a} (AC)_{ik} = 0$, $\sum_{j=1}^{b} (AB)_{ij} = \sum_{j=1}^{b} (BC)_{jk} = \sum_{k=1}^{c} (AC)_{ik} = \sum_{k=1}^{c} (BC)_{jk} = 0$ and $\sum_{i=1}^{a} (ABC)_{ijk} = \sum_{j=1}^{b} (ABC)_{ijk} = \sum_{k=1}^{c} (ABC)_{ijk} = 0$. In this decomposition, $(AB)_{ij}$ is the two-way interaction effect of interest.

The statistics for the aforementioned three separate testing problems are, respectively,

1. The test statistic for no A-B interaction when a, b are small and c is large is

$$Q_{X^*}(AB) = N\mathbf{W}'\mathbf{C}'_{AB}\left(\mathbf{C}_{AB}\widehat{\mathbf{V}}\mathbf{C}'_{AB}\right)^{-1}\mathbf{C}_{AB}\mathbf{W}, \tag{3.1.8}$$

   where $\mathbf{W} = (\widetilde{X}^*_{11..}, \ldots, \widetilde{X}^*_{1a..}, \ldots, \widetilde{X}^*_{a1..}, \ldots, \widetilde{X}^*_{ab..})'$, $\mathbf{C}_{AB} = \mathbf{M}_a \otimes \mathbf{M}_b$, where $\mathbf{M}_b = (\mathbf{1}_{b-1}| - I_{b-1})$, and $\widehat{\mathbf{V}} = \text{diag}\{\widehat{\eta}_{11}, \ldots, \widehat{\eta}_{1b}, \ldots, \widehat{\eta}_{b1}, \ldots, \widehat{\eta}_{ab}\}$ with $\widehat{\eta}_{ij} = \frac{N}{c^2}\sum_{k=1}^{c}\frac{S^2_{ijk,X^*}}{n_{ijk}}$, and $S^2_{ijk,X^*} = (n_{ijk}-1)^{-1}\sum_{m=1}^{n_{ijk}}(X^*_{ijkm} - \overline{X^*}_{ijk.})^2$.

2. The test statistic for no A-B interaction when a is large and b is small is, regardless of whether c is small or large is given by $\sqrt{a}(F_{X^*}(AB) - 1)$, where

$$F_{X^*}(AB) = MST_{AB}/MSE, \tag{3.1.9}$$

$$MST_{AB} = \frac{c}{(a-1)(b-1)}\sum_{i,j}\left(\widetilde{X}^*_{ij..} - \widetilde{X}^*_{i...} - \widetilde{X}^*_{.j..} + \widetilde{X}^*_{....}\right)^2, MSE = \frac{1}{abc}\sum_{i,j,k}\frac{S^2_{ijk,X^*}}{n_{ijk}}.$$

3. When both a and b are large, the test statistic for no A-B interaction effects is $\sqrt{ab}(F_{X^*}(AB) - 1)$, where $F_{X^*}(AB)$ is given in item 2 above, regardless of whether c is small or large.

3.1.3 Testing for no three-way interaction effects

The testing procedure for three-way interaction effects involves the following four separate problems:

1. All factors of interest have small number of levels and there is at least one more factor with many levels

2. Two factors of interest have small number of levels and one has large number of levels. The testing procedure would depend on whether the total number of levels of all other factors is large or small (possibly zero), and not on the (fixed) number of other factors.

3. One factor of interest has small number of levels and the other two have large number of levels. Again this testing procedure depends on whether the total number of levels of all other factors is large or small (possibly zero), and not on the (fixed) number of other factors.

4. All factors of interest have large number of levels. The asymptotic results depend on whether the total number of levels of all other factors is large or small (possibly zero), and not on the (fixed) number of other factors.

Therefore it is sufficient to consider a four-way design for all cases to give a general testing theory for no three-way interaction effect. To describe the test statistic, let A, B and C denote the three factors whose interaction is to be tested. Let

$$X_{ijklm} \sim F_{ijkl}(x),\ m = 1,\ldots,n_{ijkl}, \tag{3.1.10}$$

denote the m-th observation in cell (i, j, k, l), where $i = 1,\ldots,a,\ j = 1,\ldots,b$ and $k = 1,\ldots,c$, enumerate the levels of factors A, B, C, respectively, and $l = 1,\ldots,d$, enumerates the levels of the composite factor D. Consider the decomposition

$$\begin{aligned} F_{ijkl} &= M + A_i + B_j + C_k + D_l + (AB)_{ij} + (AC)_{ik} + (AD)_{il} + (BC)_{jk} + (BD)_{jl} \\ &\quad +(CD)_{kl} + (ABC)_{ijk} + (ABD)_{ijl} + (BCD)_{jkl} + (ACD)_{ikl} + (ABCD)_{ijkl} \\ &\quad i = 1,\ldots,r,\ j = 1,\ldots,c,\ k = 1,\ldots,c,\ l = 1,\ldots,d, \end{aligned} \tag{3.1.11}$$

where A_i, B_j, C_k, and D_l are main effects satisfying $\sum_{i=1}^{a} A_i = \sum_{j=1}^{b} B_j = \sum_{k=1}^{c} C_k = \sum_{l=1}^{d} D_l = 0$, $(AB)_{ij}$, $(AC)_{ik}$, $(AD)_{il}$, $(BC)_{jk}$, $(BD)_{jl}$ and $(CD)_{kl}$ are two-way interaction effects satisfying $\sum_{i=1}^{a}(AB)_{ij} = \sum_{j=1}^{b}(AB)_{ij} = 0$, and similar constraints for all the other two-way interaction effects. $(ABC)_{ijk}$, $(ABD)_{ijl}$, $(BCD)_{jkl}$, $(ACD)_{ikl}$ are three-way interaction effect satisfying $\sum_{i=1}^{a}(ABC)_{ijk} = \sum_{j=1}^{b}(ABC)_{ijk} = \sum_{k=1}^{c}(ABC)_{ijk} = 0$, and similar constraints for all the other three-way interaction effects. $(ABCD)_{ijkl}$ is the four-way interaction effect satisfying $\sum_{i=1}^{a}(ABCD)_{ijkl} = \sum_{j=1}^{b}(ABCD)_{ijkl} = \sum_{k=1}^{c}(ABCD)_{ijkl} = \sum_{l=1}^{d}(ABCD)_{ijkl} = 0$.

1. The statistic for testing for no A, B and C interaction effect when a, b, and c are all small and d is large is given by

$$Q_{X^*}(ABC) = N\mathbf{W}'\mathbf{C}'_{ABC}\left(\mathbf{C}_{ABC}\widehat{\mathbf{V}}\mathbf{C}'_{ABC}\right)^{-1}\mathbf{C}_{ABC}\mathbf{W}, \tag{3.1.12}$$

where $\mathbf{W} = (\widetilde{X}^*_{111..},\ldots,\widetilde{X}^*_{11c..},\widetilde{X}^*_{121..},\ldots,\widetilde{X}^*_{12c..},\ldots,\widetilde{X}^*_{ab1..},\ldots,\widetilde{X}^*_{abc..})'$, $\mathbf{C}_{ABC} = \mathbf{M}_a \otimes \mathbf{M}_b \otimes \mathbf{M}_c$ with $\mathbf{M}_a = (\mathbf{1}_{a-1}| - I_{a-1})$, and $\widehat{\mathbf{V}} = diag\{\widehat{\eta}_{111},\ldots,\widehat{\eta}_{11c},\widehat{\eta}_{121},\ldots,\widehat{\eta}_{12c},\ldots,\widehat{\eta}_{ab1},\ldots,\widehat{\eta}_{abc}\}$ with $\widehat{\eta}_{ijk} = \dfrac{N}{d^2}\sum_{l=1}^{d}\dfrac{S^2_{ijkl,X^*}}{n_{ijkl}}$ and $S^2_{ijkl,X^*} = (n_{ijkl}-1)^{-1}\sum_{m=1}^{n_{ijkl}}(X^*_{ijklm} - \overline{X^*}_{ijkl.})^2$.

2. The test statistic when two of the three factors have large number of levels, say a, b are large, and c is small, regardless of whether d is small or large, is $\sqrt{ab}(F_{X^*}(ABC)-1)$, where defining $MSE = \dfrac{1}{abcd}\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{c}\sum_{l=1}^{d}\dfrac{S^2_{ijkl,X^*}}{n_{ijkl}}$ and $MST_{ABC} = [(a-1)(b-1)(c-1)]^{-1}\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{c}\sum_{l=1}^{d}\left(\widetilde{X}^*_{ijk..} - \widetilde{X}^*_{ij...} - \widetilde{X}^*_{i.k..} - \widetilde{X}^*_{.jk..} + \widetilde{X}^*_{i....} + \widetilde{X}^*_{.j...} + \widetilde{X}^*_{..k..} - \widetilde{X}^*_{.....}\right)^2$, $F_{X^*}(ABC)$ is given by

$$F_{X^*}(ABC) = \frac{MST_{ABC}}{MSE}. \tag{3.1.13}$$

3. The test statistic when a is large, b, c are small, and d either large or small, is $\sqrt{a}(F_{X^*}(ABC)-1)$, where $F_{X^*}(ABC)$ is defined in (3.1.13).

4. The test statistic for no A, B and C interaction effect when a, b, and c are all large is $\sqrt{abc}(F_{X^*}(ABC)-1)$, where $F_{X^*}(ABC)$ is defined in (3.1.13), regardless of whether d is small or large.

3.2 Main Results based on original observations

3.2.1 Testing for main effects

As explained in §3.1, we will use two indices for factors to illustrate the testing procedure in this subsection.

Theorem 3.2.1. *For testing $H_0(A):$ all $A_i = 0$ when a is small and b is large, let $Q_X(A)$ be the statistic given in (3.1.3), and assume that for all i,*

$$\left(\frac{1}{b}\sum_{j=1}^{b}\frac{\sigma_{ij}^2}{n_{ij}}\right)^{-2}\frac{1}{b}\sqrt{\frac{1}{b}\sum_{j=1}^{b}\frac{1}{n_{ij}^6}} \to 0, \ \ and \ \ \limsup_{b\to\infty} b^{-1}\sum_{j=1}^{b}E^2(X_{ijm}-E(X_{ijm}))^4 < \infty, \quad (3.2.1)$$

where $\sigma_{ij}^2 = Var(X_{ijk})$. Then

$$Q_X(A) \xrightarrow{p} \chi^2_{a-1} \text{ as } b \to \infty.$$

Remark 3.1: If the fourth moments of the X_{ijk} are uniformly bounded, as is the case where X^*_{ijk} is the asymptotic rank transform, the first part of condition (3.2.1) becomes

$$\left(\frac{1}{b}\sum_{j=1}^{b}\frac{\sigma_{ij}^2}{n_{ij}}\right)^{-2}\frac{1}{b}\frac{1}{b}\sum_{j=1}^{b}\frac{1}{n_{ij}^3} \to 0.$$

When a is large and b is fixed, the test of no main factor A effect based on original observations can be found in Wang and Akritas (2002) and the (mid)-rank test can be found in Wang and Akritas (2003).

Theorem 3.2.2. *For testing $H_0(A):$ all $A_i = 0$ when a, b are large, let $F_X(A)$ be the statistic given in (3.1.4). Suppose X_{ijk} have finite variance σ_{ij}^2 and*
$\limsup(ab)^{-1}\sum_{i=1}^{a}\sum_{j=1}^{b} n_{ij}^{-1}E[X_{ijm}-E(X_{ijm})]^4 < \infty$. *Set*

$$\tau_{1,A} = \frac{2}{ab^2}\sum_{i=1}^{a}\left(\sum_{j=1}^{b}\frac{\sigma_{ij}^2}{n_{ij}}\right)^2, \quad \sigma_A^2 = \frac{1}{ab}\sum_{i=1}^{a}\sum_{j=1}^{b}\frac{\sigma_{ij}^2}{n_{ij}}. \quad (3.2.2)$$

Then under $H_0(A)$, as $a \to \infty$ and $b \to \infty$,

$$\frac{\sqrt{a}(F_X(A)-1)}{\tau_A} \xrightarrow{d} N(0,1), \text{ where } \tau_A = \sqrt{\tau_{1,A}}/\sigma_A^2,$$

holds regardless of whether the $n_{ij} \geq 2$ stay fixed, or tend to infinity with a and/or b.

Remark 3.2: As already mentioned, Theorems 3.2.1, 3.2.2 cover testing for main effects in higher-way designs. To demonstrate how this is done, we give here versions of Theorems 3.2.1, 3.2.2 when there is a third factor whose number of levels c also tending to infinity.

For the case that a is fixed, the test statistic $Q_X(A)$ is given again by (3.1.3), where now $\mathbf{W} = (\widetilde{X}_{1\ldots}, \ldots, \widetilde{X}_{a\ldots})'$, the matrix C_A stays the same, and $\hat{V} = diag(\hat{\eta}_1, \ldots, \hat{\eta}_a)$, where $\hat{\eta}_j = \dfrac{N}{b^2c^2}\sum_{j=1}^{b}\sum_{k=1}^{c}\dfrac{S^2_{ijk}}{n_{ijk}}$. Then the limiting result in Theorem 3.2.1 holds under

$$\left(\frac{1}{bc}\sum_{j,k}\frac{\sigma^2_{ijk}}{n_{ijk}}\right)^{-2}\frac{1}{bc}\sqrt{\frac{1}{bc}\sum_{j,k}\frac{1}{n^6_{ijk}}} \to 0, \text{ and } \limsup_{b,c\to\infty}(bc)^{-1}\sum_{i,j,k}E^2[(X_{ijkm}-E(X_{ijkm}))^4] < \infty.$$

For the case that a tends to infinity, the test statistic $F_X(A)$ is given again by (3.1.4), where now $MST_A = (a-1)^{-1}\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{c}\left(\widetilde{X}_{i\ldots} - \widetilde{X}_{\ldots.}\right)^2$, $MSE = (abc)^{-1}\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{c}(S^2_{ijk}/n_{ijk})$, where $S^2_{ijk} = (n_{ijk}-1)^{-1}\sum_{m=1}^{n_{ijk}}(X_{ijkm} - \overline{X}_{ijk.})^2$. Also let

$$\tau_{1,A} = \frac{2}{ab^2c^2}\sum_{i=1}^{a}\left(\sum_{j=1}^{b}\sum_{k=1}^{c}\frac{\sigma^2_{ijk}}{n_{ijk}}\right)^2, \ \sigma^2_A = \frac{1}{abc}\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{c}\frac{\sigma^2_{ijk}}{n_{ijk}},$$

where σ^2_{ijk} is the variance of observations in cell (i,j,k). Then the limiting result in Theorem 3.2.2 holds under $\limsup(abc)^{-1}\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{c}n^{-1}_{ijk}E[X_{ijkm}-E(X_{ijkm})]^4 < \infty$.

3.2.2 Testing for two-way interaction

The following three theorems correspond to the three situations described in §3.1.2.

Theorem 3.2.3. *For testing $H_0(AB)$: all $(AB)_{ij} = 0$, when a, b are small and c is large, let $Q_X(AB)$ be the statistic in (3.1.8). Let $Var(X_{ijkm}) = \sigma^2_{ijk}$ and assume that for all i, j,*

$$\left(\frac{1}{c}\sum_{k=1}^{c}\frac{\sigma^2_{ijk}}{n_{ijk}}\right)^{-2}\frac{1}{c}\sqrt{\frac{1}{c}\sum_{k=1}^{c}\frac{1}{n^6_{ijk}}} \to 0, \text{ and } \limsup_{c\to\infty}c^{-1}\sum_{k=1}^{c}E^2[(X_{ijkm}-E(X_{ijkm}))^4] < \infty \quad (3.2.3)$$

Then, under $H_0(AB)$: all $(AB)_{ij} = 0$,

$$Q_X(AB) \xrightarrow{d} \chi^2_{(a-1)\times(b-1)}, \text{ as } c\to\infty.$$

Theorem 3.2.4. *For testing $H_0(AB)$: all $(AB)_{ij} = 0$, when b is small and $a \to \infty$, let $F_X(AB)$ is the test statistic defined in (3.1.9). Suppose X_{ijkm} has variance σ^2_{ijk} and assume $\limsup(abc)^{-1}\sum_{i,j,k} n^{-1}_{ijk}E[X_{ijkm} - E(X_{ijkm})]^4 < \infty$. Let $\tau_{AB} = \frac{\sqrt{\tau_1^2+\tau_2^2+\tau_3^2}}{\sigma^2_{AB}}$, where*

$$\tau_1^2 = \frac{2}{b^2c^2a}\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{c}\frac{\sigma^4_{ijk}}{n^2_{ijk}(n_{ijk}-1)}, \quad \tau_2^2 = \frac{2(b-2)}{ab(b-1)^2c^2}\sum_{i=1}^{a}\sum_{j=1}^{b}\left(\sum_{k=1}^{c}\frac{\sigma^2_{ijk}}{n_{ijk}}\right)^2 \tag{3.2.4}$$

$$\tau_3^2 = \frac{2}{ab^2(b-1)^2c^2}\sum_{i=1}^{a}\left(\sum_{j=1}^{b}\sum_{k=1}^{c}\frac{\sigma^2_{ijk}}{n_{ijk}}\right)^2, \quad \sigma^2_{AB} = \frac{1}{abc}\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{c}\frac{\sigma^2_{ijk}}{n_{ijk}}. \tag{3.2.5}$$

then, regardless of whether $n_{ijk} \geq 2$ and c are large or small,

$$\frac{\sqrt{a}(F_X(AB)-1)}{\tau_{AB}} \xrightarrow{d} N(0,1), \ \ as\ a \to \infty.$$

Theorem 3.2.5. *For testing $H_0(AB)$: all $(AB)_{ij} = 0$, when both a and b are large, regardless of whether c large or small, let $F_X(AB)$ is the test statistic defined in (3.1.9). Suppose X_{ijkm} has variance σ^2_{ijk} and assume $\limsup(abc)^{-1}\sum_{i,j,k} n^{-1}_{ijk}E[X_{ijkm} - E(X_{ijkm})]^4 < \infty$. Let τ_1^2 and τ_2^2 be defined in (3.2.4), and σ^2_{AB} defined in(3.2.5), then as $a \to \infty$ and $b \to \infty$, regardless of whether c is large or small, and whether the $n_{ijk} \geq 2$ stay fixed, or tend to infinity with a, b, /or c,*

$$\frac{\sqrt{ab}(F_X(AB)-1)}{\tau_{AB2}} \xrightarrow{d} N(0,1), \ \ where\ \ \tau_{AB2} = \frac{\sqrt{b\tau_1^2 + (b-1)^2/(b-2)\tau_2^2}}{\sigma^2_{AB}}.$$

3.2.3 Testing for three-way interaction

The following four theorems correspond to the four situations described in §3.1.3 for testing $H_0(ABC)$: all $(ABC)_{ijk} = 0$.

Theorem 3.2.6. *For testing $H_0(ABC)$ when a, b, c are small and d is large, let $Q_X(ABC)$ be the statistic defined in (3.1.12) with $X^* = X$. Assume that*

$$\left(\frac{1}{d}\sum_{l=1}^{d}\frac{\sigma^2_{ijkl}}{n_{ijkl}}\right)^{-2}\frac{1}{d}\sqrt{\frac{1}{d}\sum_{l=1}^{d}\frac{1}{n^6_{ijkl}}} \to 0, \ \ \limsup_{d\to\infty} d^{-1}\sum_{l=1}^{d}E^2[(X_{ijklm}-E(X_{ijklm}))^4] < \infty, \tag{3.2.6}$$

hold for all i, j, k. Then under $H_0(ABC)$,

$$Q_X(ABC) \xrightarrow{d} \chi^2_{(a-1)\times(b-1)\times(c-1)}, \text{ as } d \to \infty.$$

In the following three theorems, $F_X(ABC)$ is the test statistic in (3.1.13) with $X^* = X$.

Theorem 3.2.7. *For testing $H_0(ABC)$ when a, c are large and b is small, assume that* $\limsup (abcd)^{-1} \sum_{i,j,k,l} n_{ijkl}^{-1} E[X_{ijklm} - E(X_{ijklm})]^4 < \infty$. *Let*

$$\tau_4 = \frac{2}{ac^2b^2d^2} \sum_{i,j,k,l} \frac{\sigma_{ijkl}^4}{n_{ijkl}^2(n_{ijkl}-1)}, \quad \tau_5 = \frac{2(b-2)}{ac^2d^2b(b-1)^2} \sum_{i,j,k} \left(\sum_{l=1}^{d} \frac{\sigma_{ijkl}^2}{n_{ijkl}} \right)^2 \quad (3.2.7)$$

$$\tau_6 = \frac{2}{ac^2b^2d^2(b-1)^2} \sum_{i,k} \left(\sum_{j,l} \frac{\sigma_{ijkl}^2}{n_{ijkl}} \right)^2, \quad \sigma_{ABC}^2 = \frac{1}{abcd} \sum_{i,j,k,l} \frac{\sigma_{ijkl}^2}{n_{ijkl}}, \quad (3.2.8)$$

where $\sigma_{ijkl}^2 = Var(X_{ijklm})$.Then under $H_0(ABC)$, as $a, c \to \infty$, regardless of whether $d \to \infty$ or remains fixed, and whether the $n_{ijkl} \geq 2$ stay fixed, or tend to infinity with a, c, or d,

$$\frac{\sqrt{ac}(F_X(ABC)-1)}{\tau_{ABC2}} \xrightarrow{d} N(0,1), \qquad \text{where } \tau_{ABC2} = \frac{\sqrt{c(\tau_4+\tau_5+\tau_6)}}{\sigma_{ABC}^2}$$

Theorem 3.2.8. *For testing $H_0(ABC)$ when b, c are small while a is large, assume* $\limsup (abcd)^{-1} \sum_{i,j,k,l} n_{ijkl}^{-1} E[X_{ijklm} - E(X_{ijklm})]^4 < \infty$. *Let*

$$\tau_7 = \frac{2n^2(a,b,c,d)}{ac^2d^2(b-1)^2(c-1)^2} \sum_{i,j} \sum_{k \neq k'}^{c} \sum_{l,l'} \frac{\sigma_{ijkl}^2}{n_{ijkl}} \frac{\sigma_{ijk'l'}^2}{n_{ijk'l'}}, \quad (3.2.9)$$

$$\tau_8 = \frac{2n^2(a,b,c,d)}{ab^2c^2d^2(b-1)^2(c-1)^2} \sum_{i=1}^{a} \sum_{j,j'} \sum_{k \neq k'}^{c} \sum_{l,l'} \frac{\sigma_{ijkl}^2}{n_{ijkl}} \frac{\sigma_{ij'k'l'}^2}{n_{ij'k'l'}}, \quad (3.2.10)$$

where $\sigma_{ijkl}^2 = Var(X_{ijklm})$. Then under $H_0(ABC)$, as $a \to \infty$, regardless of whether $d \to \infty$ or remains fixed, and whether the $n_{ijkl} \geq 2$ stay fixed, or tend to infinity with a or d,

$$\frac{\sqrt{a}(F_X(ABC)-1)}{\tau_{ABC1}} \xrightarrow{d} N(0,1), \quad \text{where } \tau_{ABC1} = \frac{\sqrt{\tau_4+\tau_5+\tau_6+\tau_7+\tau_8}}{\sigma_{ABC}^2},$$

where τ_4, τ_5, τ_6, and σ_{ABC}^2 defined in (3.2.7), (3.2.8).

Theorem 3.2.9. *For testing $H_0(ABC)$ when $a,b,c\to\infty$, regardless of whether $d\to\infty$ or stays fixed, assume* $\limsup(abcd)^{-1}\sum_{i,j,k,l} n_{ijkl}^{-1}E[X_{ijklm}-E(X_{ijklm})]^4<\infty$. *Let τ_4, τ_5 and σ^2_{ABC} be defined in (3.2.7) and (3.2.8) Then under $H_0(ABC)$, as* $\min\{a,b,c\}\to\infty$,

$$\frac{\sqrt{abc}(F_X(ABC)-1)}{\tau_{ABC3}}\xrightarrow{d}N(0,1),\quad \textit{where}\;\; \tau_{ABC3}=\frac{\sqrt{bc\tau_4+c(b-1)^2/(b-2)\tau_5}}{\sigma^2_{ABC}}.$$

holds regardless of whether the $n_{ijkl}\geq 2$ stay fixed, or tend to infinity with a, b, c, or d.

3.3 Main Results Based on Ranks

Throughout this section, the test statistics are calculated on the overall (mid-)ranks of the original observations.

3.4 Testing for main effects

In this sub-section, we will denote $H(x)=N^{-1}\sum_{i=1}^a\sum_{j=1}^b n_{ij}F_{ij}(x)$, the average distribution function, and $Y_{ijk}=H(X_{ijk})$. R_{ijk} is the overall (mid)-rank of X_{ijk} among all the observations.

Theorem 3.4.1. *For testing $H_0(A):$ all $A_i=0$ when a is small and b is large, let $Q_R(A)$ be the statistic given in (3.1.3) calculated with $X^*_{ijk}=R_{ijk}$, and assume that for all i,*

$$\lim_b\frac{N}{b^2}\sum_{j=1}^b\frac{1}{n_{ij}}>0,\qquad \left(\frac{1}{b}\sum_{j=1}^b\frac{\sigma^2_{ij}}{n_{ij}}\right)^{-2}\frac{1}{b}\frac{1}{b}\sum_{j=1}^b\frac{1}{n^3_{ij}}\to 0, \tag{3.4.1}$$

where now $\sigma^2_{ij}=Var(Y_{ijk})$. Then

$$Q_R(A)\xrightarrow{p}\chi^2_{a-1}\ \textit{as}\ b\to\infty.$$

When a is large and b is fixed, the (mid)-rank test of no main factor A effect can be found in Wang and Akritas (2003).

Next we will assume both a and b are large.

Theorem 3.4.2. *For testing $H_0(A):$ all $A_i = 0$ when both a and b are large, let $F_R(A)$ be the statistic given in (3.1.4) with $X^*_{ijk} = R_{ijk}$. Further let $\tau_{1,A}$ and σ^2_A be as defined in 3.2.2 with $\sigma^2_{ij} = Var(Y_{ijk})$. Then under $H_0(A)$, as $a, b \to \infty$,*

$$\frac{\sqrt{a}(F_R(A)-1)}{\tau_A} \xrightarrow{d} N(0,1), \text{ where } \tau_A = \sqrt{\tau_{1,A}}/\sigma^2_A,$$

regardless of whether the $n_{ij} \geq 2$ stay fixed, or tend to infinity with a or b.

3.4.1 Testing for two-way interaction

Here we will give the asymptotic distribution of the test statistics given in §3.1.2 for testing $H_0(AB)$: all $(AB)_{ij} = 0$. In this sub-section, we will denote the average distribution function $H(x) = N^{-1}\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{c} n_{ijk}F_{ijk}(x)$, and $Y_{ijkm} = H(X_{ijkm})$. R_{ijkm} is the overall (mid)-rank of X_{ijkm} among all the observations.

Theorem 3.4.3. *For testing $H_0(AB)$ when both a and b are small, let $Q_R(AB)$ be the statistic given in (3.1.8) with $X^*_{ijkm} = R_{ijkm}$, assume that for all i, j,*

$$\lim_c \frac{N}{c^2}\sum_{k=1}^{c}\frac{1}{n_{ijk}} > 0, \qquad \left(\frac{1}{c}\sum_{k=1}^{c}\frac{\sigma^2_{ijk}}{n_{ijk}}\right)^{-2}\frac{1}{c}\frac{1}{c}\sum_{k=1}^{c}\frac{1}{n^3_{ijk}} \to 0, \tag{3.4.2}$$

where now $\sigma^2_{ijk} = Var(Y_{ijkm})$, then under $H_0(AB)$: all $(AB)_{ij} = 0$,

$$Q_R(AB) \xrightarrow{d} \chi^2_{(a-1)\times(b-1)}, \text{ as } c \to \infty$$

Theorem 3.4.4. *For testing $H_0(AB)$ when a is large and b is small, let $F_R(AB)$ be the statistic given in (3.1.9) with $X^*_{ijkm} = R_{ijkm}$, Let $\tau_{AB} = \sqrt{\tau_1^2 + \tau_2^2 + \tau_3^2}/\sigma^2_{AB}$, where τ_1^2, τ_2^2, are defined in (3.2.4) and τ_3^2 and σ^2_{AB} are as defined in (3.2.5) with $\sigma^2_{ijk} = Var(Y_{ijkm})$. then as $a \to \infty$, regardless of whether c and $n_{ijk} \geq 2$ go to ∞ or stay fixed,*

$$\frac{\sqrt{a}(F_R(AB)-1)}{\tau_{AB}} \xrightarrow{d} N(0,1).$$

Theorem 3.4.5. *For testing $H_0(AB)$ when both a and b are large, let $F_R(AB)$ be the statistic given in (3.1.9) with $X^*_{ijkm} = R_{ijkm}$. then as $a, b \to \infty$, regardless of whether c and $n_{ijk} \geq 2$ tend to ∞ or stay fixed,*

$$\frac{\sqrt{ab}(F_R(AB)-1)}{\tau_{AB2}} \xrightarrow{d} N(0,1), \text{ where } \tau_{AB2} = \frac{\sqrt{b\tau_1^2 + (b-1)^2/(b-2)\tau_2^2}}{\sigma_{AB}^2},$$

where τ_1^2, τ_2^2 and σ_{AB}^2 are as defined in (3.2.4) and (3.2.5) with $\sigma_{ijk}^2 = Var(Y_{ijkm})$.

3.4.2 Testing for three-way interaction

Here we will give the asymptotic results for testing $H_0(ABC)$: all $(ABC)_{ijk} = 0$ in each case using corresponding test statistics with $X^*_{ijklm} = R_{ijklm}$. In this sub-section, we will denote $H(x) = N^{-1}\sum_{i,j,k,l} n_{ijkl}F_{ijkl}(x)$, the average distribution function, and $Y_{ijklm} = H(X_{ijklm})$. R_{ijklm} is the overall (mid)-rank of X_{ijklm} among all the observations.

Theorem 3.4.6. *For testing $H_0(ABC)$ when a, b, c are all small, let $Q_R(ABC)$ be the statistic given in (3.1.12) with $X^*_{ijklm} = R_{ijklm}$. Assume that for all i, j, k,*

$$\lim_d \frac{N}{d^2}\sum_{l=1}^{d}\frac{1}{n_{ijkl}} > 0, \qquad \left(\frac{1}{d}\sum_{l=1}^{d}\frac{\sigma_{ijkl}^2}{n_{ijkl}}\right)^{-2}\frac{1}{d}\frac{1}{d}\sum_{l=1}^{d}\frac{1}{n_{ijkl}^3} \to 0, \tag{3.4.3}$$

where now $\sigma_{ijkl}^2 = Var(Y_{ijklm})$. Then under $H_0(ABC)$: all $(ABC)_{ijk} = 0$,

$$Q_R(ABC) \xrightarrow{d} \chi^2_{(a-1)\times(b-1)\times(c-1)}, \text{ as } d \to \infty.$$

Theorem 3.4.7. *For testing $H_0(ABC)$ when a, c are large and b is small, let $F_R(ABC)$ be the statistic given in (3.1.13) with $X^*_{ijklm} = R_{ijklm}$. Further let τ_4, τ_5, τ_6, and σ_{ABC}^2 be as defined in (3.2.7) and (3.2.8) with $\sigma_{ijkl}^2 = Var(Y_{ijklm})$. Then under $H_0(ABC)$, as $a, c \to \infty$ while b remains fixed, regardless of whether d and $n_{ijkl} \geq 2$ tend to ∞ or remain fixed,*

$$\frac{\sqrt{ac}(F_R(ABC)-1)}{\tau_{ABC2}} \xrightarrow{d} N(0,1), \qquad \text{where } \tau_{ABC2} = \frac{\sqrt{c(\tau_4+\tau_5+\tau_6)}}{\sigma_{ABC}^2}.$$

Theorem 3.4.8. *For testing $H_0(ABC)$ when a is large and b,c are small, let $F_R(ABC)$ be the statistic given in (3.1.13) with $X^*_{ijklm} = R_{ijklm}$. Let τ_4, τ_5, τ_6, and σ^2_{ABC} be given in Theorem 3.4.7 and τ_7, τ_8, be as defined in (3.2.9) and (3.2.10) with $\sigma^2_{ijkl} = Var(Y_{ijklm})$. Then under $H_0(ABC)$, as $a \to \infty$ regardless of whether d and $n_{ijkl} \geq 2$ tend to ∞ or remain fixed,*

$$\frac{\sqrt{a}(F_R(ABC)-1)}{\tau_{ABC1}} \xrightarrow{d} N(0,1), \qquad \text{where } \tau_{ABC1} = \frac{\sqrt{\tau_4+\tau_5+\tau_6+\tau_7+\tau_8}}{\sigma^2_{ABC}}.$$

Theorem 3.4.9. *For testing $H_0(ABC)$ when a,b,c are all large, let $F_R(ABC)$ be the statistic given in (3.1.13) with $X^*_{ijklm} = R_{ijklm}$. Then under $H_0(ABC)$, as $a,b,c \to \infty$, regardless of whether d and $n_{ijkl} \geq 2$ tend to ∞ or remain fixed,*

$$\frac{\sqrt{abc}(F_R(ABC)-1)}{\tau_{ABC3}} \xrightarrow{d} N(0,1), \qquad \text{where } \tau_{ABC3} = \frac{\sqrt{bc\tau_4 + c(b-1)^2/(b-2)\tau_5}}{\sigma^2_{ABC}},$$

where τ_4, τ_5 and σ^2_{ABC} are defined in Theorem 3.4.7.

3.5 Numerical Results

3.5.1 Simulation Results

The simulations reported in this section pertain only to the three-way ANOVA design. We compare the asymptotic heteroscedastic tests based on the original observations (OBS), those based on ranks (RANK) and the classical F-tests (CF). Type I error-rate results are reported for testing $H_0(A)$, $H_0(B)$, $H_0(AC)$, $H_0(AB)$, and $H_0(ABC)$ for the cases where the number of levels of factors A and C are large and that of factor B is small. In particular, the number of levels of factor A is $a = 20$, 30, and 50, and those for factors B and C are $b = 2$ and $c = 20$, respectively. Results for the achieved power are reported only for testing $H_0(A)$, $H_0(AC)$, and $H_0(B)$ with $a = 20$, $b = 2$ and $c = 20$. Both balanced and unbalanced designs are considered. In the balanced cases the group size is always $n = 4$. The group sizes for the unbalanced cases are given below the corresponding tables. The simulations are based on 2,000 replications and use the normal, log-normal and Cauchy distributions. In all tables, Normal(c_1, c_2)

or Lognormal(c_1, c_2) denotes a normal or logarithm of the lognormal, respectively, random variable with (mean, standard deviation) $= (c_1, c_2)$, and Cauchy(c_1, c_2) means that (location, scale)$= (c_1, c_2)$. Tables 3.1-3.5 report the achieved α-levels, while Tables 3.6-3.11 report the achieved power.

Remark 5.1 The group size $n = 4$, chosen for the simulations, is the smallest group size that permits unbiased estimation of σ^4_{ijk} which is needed for consistent estimation of the variance of the test statistic. In particular, we used the unbiased estimator for σ^4_{ijk} suggested in Akritas and Papadatos (2003). In cases where the group size is less than four, as is the case in the data set we analyzed, unbiased estimation of σ^4_{ijk} can be achieved by exploiting patterns of heteroscedasticity, e.g. $\sigma_{ijk} = \sigma_{jk}$, for all i. In absence of such patterns, we recommend a bias correction technique such as jackknife.

Table 3.1: Estimated level for $\alpha = 0.05$, unbalanced heteroscedastic case, b=2, c=20, $X_{ijkm} \sim N(0, 4jk/bc)^{(*)}$

H_0	a=20			a=30			a=50		
	CF	Obs	Rank	CF	Obs	Rank	CF	Obs	Rank
$H_0(A)$	0.0965	0.0702	0.0767	0.0989	0.0661	0.0714	0.1069	0.0658	0.0614
$H_0(AB)$	0.1296	0.0496	0.0536	0.1220	0.0594	0.0603	0.0927	0.0494	0.0525
$H_0(AC)$	0.4911	0.0792	0.0777	0.4693	0.0616	0.0607	0.4754	0.0671	0.0676
$H_0((ABC)$	0.3813	0.0691	0.0709	0.4316	0.0652	0.0638	0.4295	0.0592	0.0583
$H_0(B)$	0.9992	0.0666	0.0662	0.9995	0.0629	0.0705	0.9991	0.0583	0.0565

$^{(*)}$ The group sizes in these simulations are as follows: When $a = 20$, $n_{i1k} = 12$ for $i = 1, \ldots, 10$, and all $k = 1, \ldots, 20$; $n_{i1k} = 10$ for $i = 11$, and $k = 1, \ldots, 20$; and $n_{i1k} = 5$ for $i = 12, \ldots, 20$; $n_{i2k} = 4$, for all i, k. When $a = 30$, $n_{i1k} = 12$ for $i = 1, \ldots, 10$, and all $k = 1, \ldots, 20$; $n_{i1k} = 10$ for $i = 11$, and $k = 1, \ldots, 20$; and $n_{i1k} = 5$ for $i = 12, \ldots, 30$; $n_{i2k} = 4$, for all i, k. When $a = 50$, $n_{i1k} = 12$ for $i = 1, \ldots, 10$, and all $k = 1, \ldots, 20$; $n_{i1k} = 10$ for $i = 11$, and $k = 1, \ldots, 20$; and $n_{i1k} = 5$ for $i = 12, \ldots, 50$; $n_{i2k} = 4$, for all i, k.

Table 3.2: Estimated level, balanced homoscedastic case, $\alpha = 0.05$, b=2, c=20, n=4

a	H_0	Normal(0, 1)			Lognormal(0, 1)			Cauchy(0, 1)		
		CF	Obs	Rank	CF	Obs	Rank	CF	Obs	Rank
20	A	0.0600	0.0755	0.0730	0.050	0.0630	0.0660	0.0160	0.0330	0.0655
	AB	0.0460	0.0630	0.0650	0.059	0.0810	0.0690	0.0195	0.0350	0.0565
	AC	0.0460	0.0595	0.0645	0.057	0.0490	0.0685	0.0105	0.0975	0.0530
	ABC	0.0495	0.0685	0.0715	0.048	0.0565	0.0580	0.0120	0.1000	0.0665
	B	0.0490	0.0490	0.0495	0.050	0.0500	0.0535	0.0175	0.0180	0.0520
30	A	0.0485	0.0640	0.0625	0.0475	0.0560	0.072	0.0210	0.0330	0.0680
	AB	0.0515	0.0660	0.0680	0.0490	0.0625	0.065	0.0165	0.0265	0.0690
	AC	0.0465	0.0610	0.0635	0.0510	0.0440	0.057	0.0090	0.0940	0.0635
	ABC	0.0480	0.0570	0.0610	0.0545	0.0590	0.057	0.0090	0.1100	0.0550
	B	0.0515	0.0515	0.0480	0.0530	0.0530	0.051	0.0245	0.0245	0.0570
50	A	0.0535	0.0635	0.0580	0.0495	0.0575	0.0670	0.0210	0.0240	0.0600
	AB	0.0470	0.0610	0.0705	0.0500	0.0635	0.0640	0.0195	0.0260	0.0585
	AC	0.0460	0.0545	0.0535	0.0530	0.0440	0.0545	0.0100	0.1075	0.0540
	ABC	0.0465	0.0560	0.0605	0.0510	0.0520	0.0565	0.0055	0.0975	0.0530
	B	0.0560	0.0560	0.0545	0.0555	0.0555	0.0505	0.0255	0.0255	0.0545

Table 3.3: Estimated level, unbalanced homoscedastic case, $\alpha = 0.05$, b=2, c=20[(*)]

a	H_0	Normal(0, 1)			Lognormal(0, 1)			Cauchy(0, 1)		
		CF	Obs	Rank	CF	Obs	Rank	CF	Obs	Rank
20	A	0.0485	0.0665	0.0670	0.0480	0.0575	0.0695	0.0215	0.0295	0.0745
	AB	0.0465	0.0640	0.0615	0.0410	0.0580	0.0705	0.0315	0.0310	0.0745
	AC	0.0450	0.0615	0.0625	0.0560	0.0490	0.0690	0.0200	0.1120	0.0700
	ABC	0.0505	0.0665	0.0675	0.0625	0.0570	0.0635	0.1450	0.1110	0.0650
	B	0.0455	0.0450	0.0510	0.0500	0.0490	0.0535	0.0180	0.0190	0.0535
30	A	0.0560	0.0730	0.0745	0.0535	0.0590	0.0710	0.0230	0.0225	0.0645
	AB	0.0490	0.0670	0.0715	0.0430	0.0575	0.0595	0.0435	0.0210	0.0665
	AC	0.0380	0.0500	0.0510	0.0645	0.0490	0.0695	0.2440	0.1095	0.0625
	ABC	0.0560	0.0600	0.0650	0.0730	0.0650	0.0680	0.4085	0.0985	0.0660
	B	0.0415	0.0425	0.0475	0.0465	0.0485	0.0605	0.0280	0.0250	0.0530
50	A	0.0490	0.0615	0.0610	0.0560	0.0645	0.0675	0.0275	0.0215	0.0610
	AB	0.0535	0.0695	0.0655	0.0560	0.0685	0.0620	0.0255	0.0235	0.0705
	AC	0.0445	0.0585	0.0565	0.0560	0.0445	0.0575	0.4145	0.0955	0.0555
	ABC	0.0485	0.0600	0.0600	0.0585	0.0525	0.0590	0.4210	0.1070	0.0645
	B	0.0560	0.0520	0.0535	0.0460	0.0425	0.0470	0.0210	0.0235	0.0510

[(*)] For $a = 20$, 620 of the group sizes are 4, 120 of them are 5 and 60 of them are 6. For $a = 30$, 800 of the group sizes are 4, 160 of them are 5, 220 of them are 6 and 20 of them are 7. For $a = 50$, 1400 of the group sizes are 4, 240 of them are 5, 320 of them are 6 and 40 of them are 7.

Table 3.4: Estimated level, unbalanced heteroscedastic case, $\alpha = 0.05$, b=2, c=20

a	H_0	Normal$(0,4jk/(bc))$			Lognormal(0, 4jk/(bc))			Cauchy(0, 4jk/(bc))		
		CF	Obs	Rank	CF	Obs	Rank	CF	Obs	Rank
20	A	0.0505	0.0695	0.0685	0.0805	0.0030	0.0685	0.0200	0.0315	0.0595
	AB	0.0470	0.0670	0.0705	0.0275	0.0020	0.0585	0.0275	0.0380	0.0570
	AC	0.1120	0.0735	0.0615	0.1475	0.0055	0.0670	0.0385	0.0910	0.0680
	ABC	0.1095	0.0745	0.0725	0.1125	0.0060	0.0550	0.1560	0.0860	0.0580
	B	0.0495	0.0490	0.0580	0.6215	0.6340	0.0570	0.0220	0.0205	0.0575
30	A	0.0525	0.0660	0.0715	0.0990	0.003	0.0610	0.0200	0.0310	0.0675
	AB	0.0605	0.0650	0.0620	0.0355	0.003	0.0745	0.0360	0.0300	0.0585
	AC	0.1075	0.0615	0.0585	0.3855	0.001	0.0585	0.2420	0.1030	0.0610
	ABC	0.1385	0.0620	0.0550	0.4390	0.001	0.0595	0.4355	0.0955	0.0600
	B	0.0530	0.0485	0.0540	0.6880	0.6770	0.0645	0.0265	0.0250	0.0560
50	A	0.044	0.0555	0.0565	0.0340	0.0005	0.0625	0.0275	0.0315	0.0605
	AB	0.044	0.0580	0.0640	0.0345	0.0005	0.0690	0.0305	0.0240	0.0645
	AC	0.112	0.0650	0.0650	0.4515	0.0015	0.0590	0.4560	0.0870	0.0710
	ABC	0.115	0.0590	0.0590	0.4515	0.0015	0.0545	0.4525	0.0875	0.0640
	B	0.059	0.0630	0.0700	0.7065	0.7105	0.0565	0.0245	0.0235	0.0540

(*) For $a = 20$, 620 of the group sizes are 4, 120 of them are 5 and 60 of them are 6. For $a = 30$, 800 of the group sizes are 4, 160 of them are 5, 220 of them are 6 and 20 of them are 7. For $a = 50$, 1400 of the group sizes are 4, 240 of them are 5, 320 of them are 6 and 40 of them are 7.

Table 3.5: Estimated level, unbalanced heteroscedastic case, $\alpha = 0.05$, b=2, $c = 20^{(*)}$.

a	H_0	Normal$(0,4j/b)$			Lognormal(0, 4j/b)$^{(**)}$			Cauchy(0,4j/b)		
		CF	Obs	Rank	CF	Obs	Rank	CF	Obs	Rank
20	A	0.0460	0.0630	0.0580	0.0605	0.0020	0.0755	0.0195	0.0250	0.0595
	AB	0.0445	0.0655	0.0650	0.0270	0.0020	0.0605	0.0260	0.0355	0.0615
	AC	0.0460	0.0535	0.0555	0.0225	0.1070	0.0665	0.0255	0.1050	0.0610
	ABC	0.0530	0.0660	0.0670	0.1480	0.1065	0.0625	0.1465	0.0960	0.0660
	B	0.0460	0.0465	0.0605	0.6475	0.6420	0.0610	0.0215	0.0195	0.0590
30	A	0.0580	0.0685	0.0655	0.0570	0.0005	0.0550	0.0195	0.0190	0.0685
	AB	0.0565	0.0635	0.0640	0.0295	0.0005	0.0605	0.0440	0.0300	0.0605
	AC	0.0515	0.0535	0.0545	0.2670	0.1055	0.0675	0.2350	0.0975	0.0625
	ABC	0.0885	0.0660	0.0605	0.5080	0.1050	0.0635	0.4210	0.1065	0.0530
	B	0.0460	0.0445	0.0580	0.7170	0.7065	0.0590	0.0180	0.0160	0.0565
50	A	0.0495	0.0600	0.0645	0.0445	0.0000	0.0545	0.0260	0.0270	0.0640
	AB	0.0490	0.0595	0.0695	0.0445	0.0000	0.0590	0.0270	0.0250	0.0600
	AC	0.0565	0.0640	0.0675	0.4740	0.1135	0.0465	0.4245	0.0960	0.0605
	ABC	0.0520	0.0635	0.0595	0.4735	0.1130	0.0640	0.4235	0.0975	0.0620
	B	0.0535	0.0545	0.0540	0.0180	0.0175	0.0530	0.0180	0.0175	0.0530

$^{(*)}$ For $a = 20$, 620 of the group sizes are 4, 120 of them are 5 and 60 of them are 6. For $a = 30$, 800 of the group sizes are 4, 160 of them are 5, 220 of them are 6 and 20 of them are 7. For $a = 50$, 1400 of the group sizes are 4, 240 of them are 5, 320 of them are 6 and 40 of them are 7.

$^{(**)}$ For testing $H_0(B)$, the observations are generated as $X - E(X)$, where $ln(X) \sim N(0,4j/b)$.

Table 3.6: Achieved power for testing $H_0(A)$, balanced homoscedastic case, $\alpha = 0.05$, a=20, b=2, c=20, n=4.

τ	Normal$(i\tau/(4a),1)$			Lognormal$(i\tau/(4a),1)$			Cauchy$(i\tau/(4a),1)$		
	CF	Obs	Rank	CF	Obs	Rank	CF	Obs	Rank
0	0.0495	0.0735	0.0660	0.050	0.0630	0.0660	0.0160	0.0330	0.0655
0.5	0.1555	0.1940	0.1945	0.0700	0.0875	0.2975	0.0195	0.0275	0.1070
1	0.6955	0.7460	0.7240	0.1490	0.181	0.9205	0.0245	0.0285	0.2355
1.5	0.9875	0.9940	0.9890	0.3225	0.3745	1.0000	0.0205	0.0320	0.5120
2							0.0160	0.0285	0.8115
4							0.0165	0.0355	1.0000

Table 3.7: Achieved power for testing $H_0(AC)$, balanced homoscedastic case, $\alpha = 0.05$, $a = 20$, $b = 2$, $c = 20$, $n = 4$.

Normal$(ik\tau/(ac),1)$				Lognormal$(ik\tau/(ac),1)$				Cauchy$(ik\tau/(ac),1)$			
CF	Obs	Rank	τ	CF	Obs	Rank	τ	CF	Obs	Rank	τ
0.0450	0.0565	0.0555	0	0.0505	0.0360	0.0660	0	0.0110	0.0970	0.0715	0
0.0765	0.0955	0.1000	0.5	0.0585	0.0505	0.1120	0.5	0.0110	0.1110	0.1555	1.5
0.1765	0.2055	0.1940	1	0.0795	0.0745	0.2360	1	0.0095	0.1030	0.2195	2
0.5000	0.5480	0.4865	1.5	0.0980	0.0985	0.490	1.5	0.0155	0.1065	0.3345	2.5
0.8600	0.8905	0.8190	2	0.1615	0.1740	0.754	2	0.0075	0.0800	0.4745	3
1.0000	1.000	1.0000	3	0.4205	0.4840	0.9750	3	0.0115	0.1280	0.7335	4
				0.7850	0.8505	0.9990	4	0.0135	0.1085	0.8305	4.5
								0.0140	0.1100	0.8955	5
								0.011	0.0990	0.9710	6
								0.0125	0.1120	0.9940	7

Table 3.8: Achieved power for testing $H_0(B)$, balanced homoscedastic case, $\alpha = 0.05$, $a = 20$, $b = 2$, $c = 20$, $n = 4$.

Normal$(j\tau/(192b),1)$				Lognormal$(j\tau/(192b),\ 1\ )$				Cauchy$(j\tau/b,1)$			
CF	Obs	Rank	τ	CF	Obs	Rank	τ	CF	Obs	Rank	τ
0.0505	0.0505	0.0490	0	0.0425	0.0425	0.0505	0	0.0180	0.0180	0.0535	0
0.1450	0.1450	0.1440	12	0.0740	0.0750	0.2195	12	0.0315	0.0315	0.9725	0.5
0.2795	0.2795	0.2635	19.2	0.1105	0.1110	0.4095	19.2	0.0415	0.0415	1.0000	1
0.3290	0.3290	0.3170	20	0.1050	0.1050	0.4430	20	0.0605	0.0605	1.0000	1.5
0.3370	0.3375	0.3365	21	0.1160	0.1160	0.4680	21				
0.3685	0.3690	0.3590	22	0.1160	0.1165	0.5305	22				
0.3955	0.3965	0.3825	23	0.1365	0.1370	0.5555	23				
0.4065	0.4070	0.3940	23.5	0.1440	0.1440	0.6135	23.75				
0.4155	0.4175	0.4000	23.875	0.1565	0.1565	0.7115	27.4285				
0.6470	0.6470	0.6280	32	0.2015	0.2015	0.8285	32				
0.8045	0.8045	0.7940	38.4	0.2595	0.2595	0.9345	38.4				
0.9490	0.9490	0.9395	48	0.3735	0.3735	0.9915	48				

In the homoscedastic case (balanced or unbalanced), all three tests compared here are robust in regard to the type I error rate when the data are normally distributed or even skewed. With heavy-tailed data, however, the achieved α-levels of CF and OBS are either very liberal or very conservative, depending on the hypothesis (Tables 3.2 and 3.3). Moreover, Tables 3.6, 3.7 and 3.8 indicate that CF and OBS have no power at all when the data are heavy-tailed, and only moderate power compared to RANK when the data are skewed.

In the unbalanced heteroscedastic case, both OBS and RANK provide reliable type I error rates under normality, but only RANK does so under non-normality (Tables 3.1, 3.4, 3.5). The poor performance of OBS under non-normality is due to violation of assumptions in the

Table 3.9: Achieved power for testing $H_0(A)$, unbalanced heteroscedastic case, $\alpha = 0.05$, a=20, b=2, c=20, 620 of the group sizes are 4, 120 of them are 5 and 60 of them are 6.

τ	Normal$(i\tau/(2a), 4jk/(bc)))$			Lognormal$(i\tau/(2a), 4jk/(bc))$			Cauchy$(i\tau/(2a), 4jk/(bc))$		
	CF	Obs	Rank	CF	Obs	Rank	CF	Obs	Rank
0	0.0485	0.0640	0.0685	0.0320	0.0025	0.0605	0.0280	0.030	0.0595
0.5	0.1805	0.2235	0.4615	0.0255	0.0025	0.8315	0.0205	0.0320	0.2115
1	0.7635	0.8115	0.9955	0.0340	0.0030	1.0000	0.0205	0.0315	0.7070
1.5	0.9935	0.9950	1.0000	0.0260	0.0000	1.0000	0.0145	0.0270	0.9780

Table 3.10: Achieved power for testing $H_0(AC)$, unbalanced heteroscedastic case, $\alpha = 0.05$, a=20, b=2, c=20, 620 of the group sizes are 4, 120 of them are 5 and 60 of them are 6.

τ	Normal$(4ik\tau/(ac), 4j/b)$			Lognormal$(4ik\tau/(ac), 4j/b)$			Cauchy$(4ik\tau/(ac), 4j/b)$		
	CF	Obs	Rank	CF	Obs	Rank	CF	Obs	Rank
0	0.049	0.0620	0.0555	0.0200	0.1090	0.0670	0.0245	0.1060	0.0605
0.5	0.0900	0.1055	0.1190	0.0205	0.1025	0.716	0.0215	0.1055	0.4465
1	0.3455	0.3695	0.4270	0.0230	0.1050	0.9180	0.0215	0.1245	0.9545
1.5	0.8430	0.8635	0.8795	0.0230	0.1055	0.9835	0.0290	0.1085	1.0000
2	0.9935	0.9950	0.9945						

Table 3.11: Achieved power for testing $H_0(B)$, unbalanced heteroscedastic case, $\alpha = 0.05$, a=20, b=2, c=20, 620 of the group sizes are 4, 120 of them are 5 and 60 of them are 6.

τ	Normal$(j\tau/(4b), 4jk/(bc))$			Lognormal$(j\tau/(4b), 4jk/(bc))$			Cauchy$(j\tau/(4b), 4jk/(bc))$		
	CF	Obs	Rank	CF	Obs	Rank	CF	Obs	Rank
0	0.0600	0.0620	0.0590	0.6385	0.6300	0.0500	0.0170	0.0190	0.0525
0.5	0.1650	0.1635	0.2835	0.6410	0.6425	1.0000	0.040	0.0410	1.0000
1	0.4900	0.4845	0.7585	0.6665	0.6670	1.0000	0.040	0.0410	1.0000
1.5	0.8230	0.8195	0.9800	0.6560	0.6515	1.0000	0.0775	0.0740	1.0000

Cauchy case and the fact that the sample variance is not a reliable estimator of the variance in the lognormal case. The type I error rate of CF can be unacceptable, both under normality (Table 3.1) and non-normality (Table 3.4). It should be pointed out that the unbalancedness in Table 3.4 is rather mild. However, the achieved α-level also depends on the pattern of heteroscedasticity, as seen in Table 3.5, where the achieved α-level of CF is acceptable with identical pattern and degree of unbalancedness as in Table 3.4. Moreover, Tables 3.9, 3.10 and 3.11 indicate that RANK is much more powerful than both CF and OBS under non-normality, while under normality it has comparable power to CF when the latter has acceptable type I error rate.

It is easier to see above summary from the power plots. The following nine figures give the power estimate for testing $H_0(A)$, $H_0(AC)$, $H_0(B)$ in normal, lognormal and Cauchy cases.

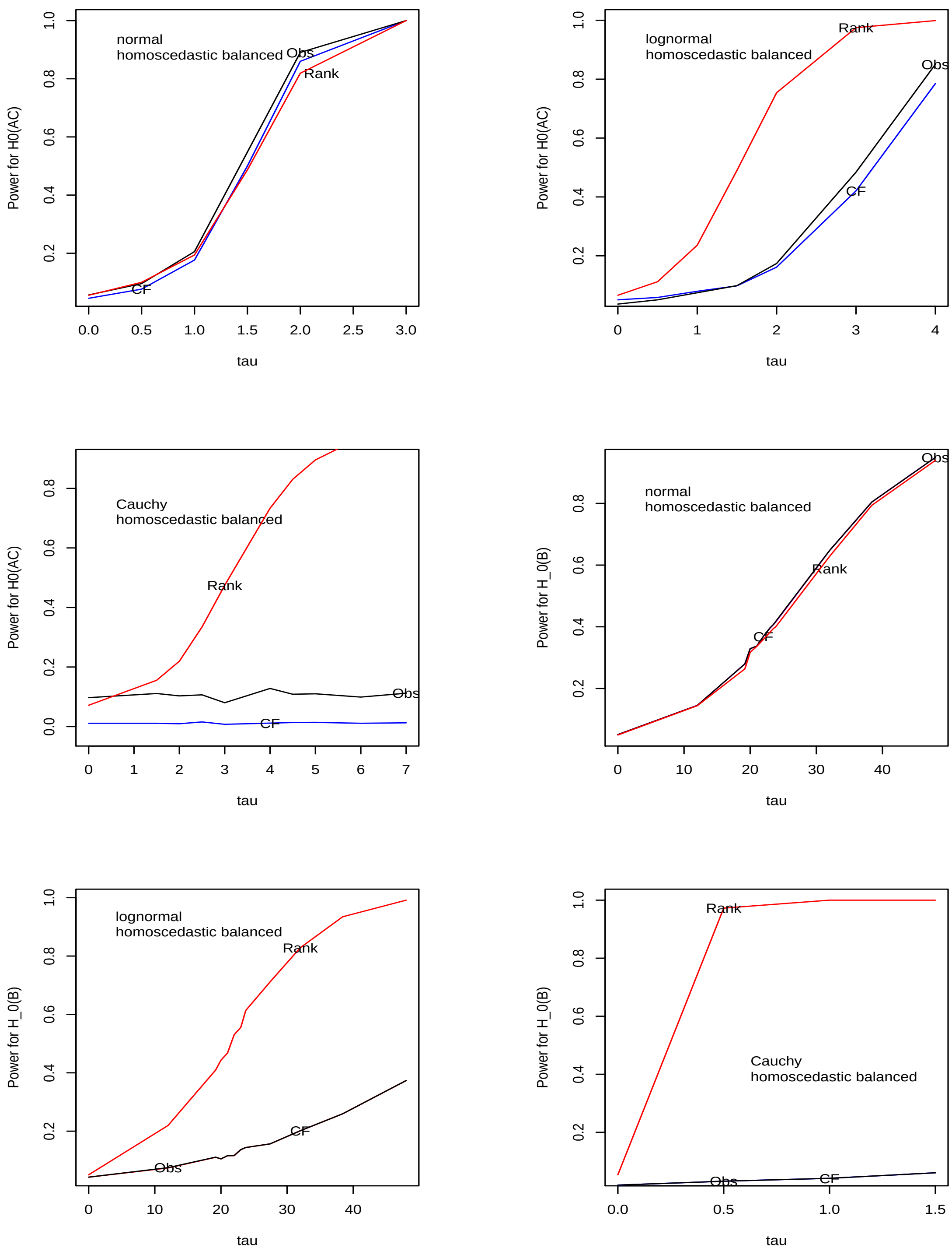

Figure 3.1: *Power function: homoscedastic balanced case*

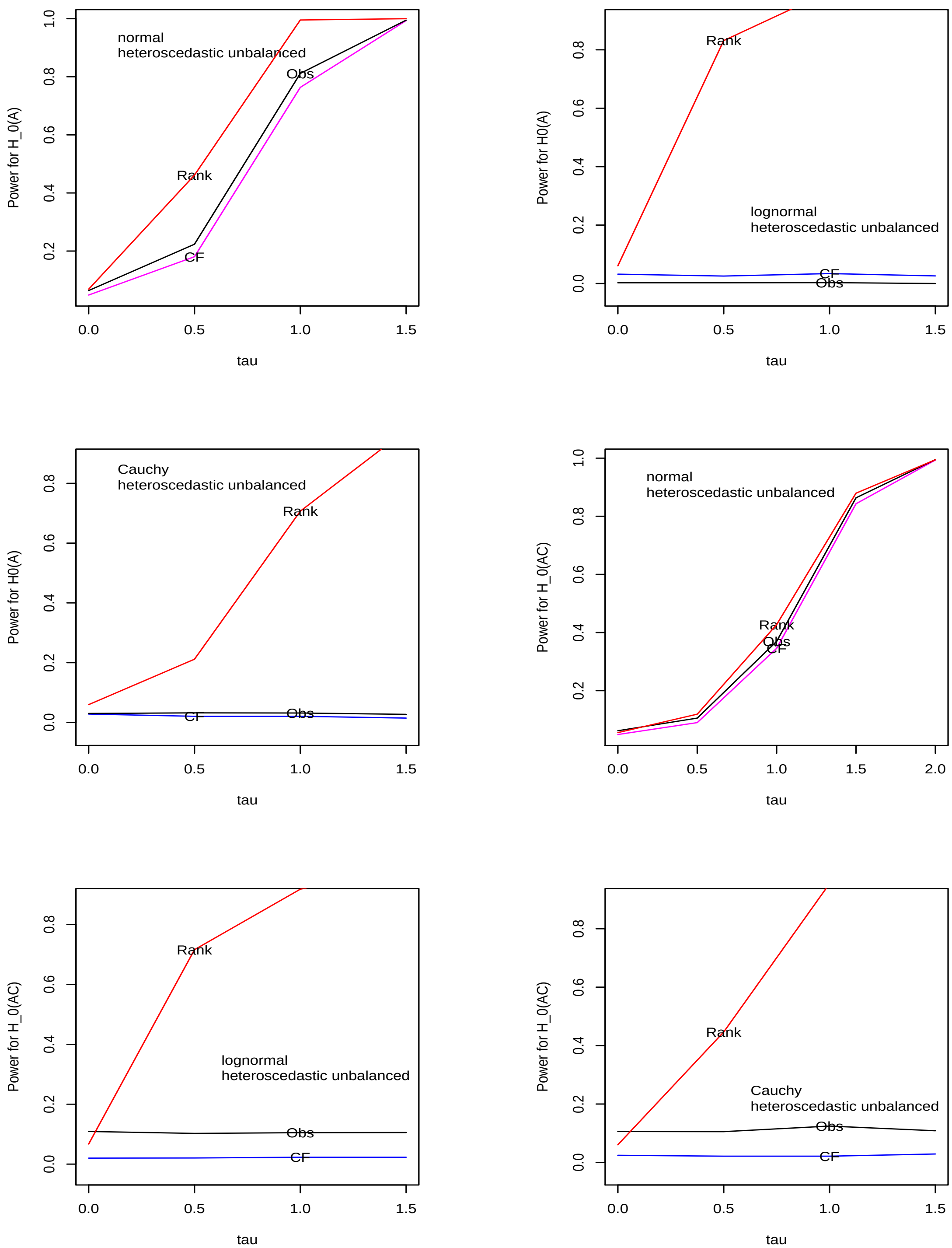

Figure 3.2: *Power function: Heteroscedastic unbalanced case*

In summary, it seems that RANK out-performs both CR and OBS in terms of achieved α-levels and power, and thus it should be preferred in any situation where one of the factors has many levels and the group sizes are small. In fact, the simulations make a compelling case that CF should not be used in such cases in the presence of heteroscedasticity.

3.5.2 Case Study

Spotted cDNA microarrays are a tool for high-throughput analysis of gene expression. In the first step of the technique, DNA is spotted and immobilized on the microarrays. Each spot on an array contains a particular sequence. Next, mRNA from cell populations under study is reverse-transcribed into cDNA and one of two fluorescent dye labels, Cy3 and Cy5, is incorporated. Two pools of differently-labeled cDNA are mixed and washed over the array. Dye-labeled cDNA can hybridize with complementary sequences on the array, and unhybridized cDNA is washed off. The array is then scanned for Cy3 and Cy5 fluorescent intensities. The idea is that the mRNA sample that contained more transcript for a given gene should produce higher fluorescence in the corresponding label in the spot containing that gene. The data consist of Cy3 and Cy5 measurements for every spot on every array.

Recently, ANOVA has been recognized as a powerful tool to analyze data from cDNA microarray experiments. For example, Kerr et al. (2001), referred to as ground-breaking work in Wolfinger, et al. (2001), proposed a general ANOVA model for the logs of the fluorescence measurements. In this section, we reanalyze the data studied in Kerr, et al. (2002), (http://www.jax.org/research/churchill), which come from a microarray experiment to compare gene expression in drug-treated and control cells lines. In this experiment, each array was spotted with the same set of 1920 genes. The cleaned dataset has complete data for 1907 genes. The experimental design included replication, achieved by using six arrays, to control the noise that is associated with microarray data. Control cells are treatment 1 and TCDD-treated cells are treatment 2. As in Kerr, et al. (2002), we refer to the fluorescence Cy3 as dye 1 and Cy5 as dye 2.

In their model for the log-data Kerr, et al. (2002) included effects for Array, Dye, Gene, and Gene×Array, Gene×Dye, Gene×Treatment and Array×Dye interactions. Examination of residual plots revealed only modest heteroscedasticity. In order to remove the interaction effects, they used a loess-based transformation. This was successful in reducing the Array×Dye sum of squares but not those of Gene×Dye. Thus, they did a further 'shift' adjustment to the pre-log data. See Kerr, et al. (2002) for details.

In our analysis, we will treat array as replication rather than a factor. Thus, we consider gene, treatment and dye as factors A, B and C, with factor A having a large number of levels. To confirm the heteroscedasticity found in Kerr, et al. (2002), we did a number of t-tests using the cell sample variance S^2_{ijk}. In particular, to see whether the cell variances depend on Treatment, we conducted t-tests on $S^2_{i1k} - S^2_{i2k},\ i = 1,\ldots,1907$, for $k = 1,2$, and also on $S^2_{i1k}/S^2_{i2k},\ i = 1,\ldots,1907$, for $k = 1,2$. Similarly we tested for the dependence of the cell variances on Dye. The tests were done for both the original observations and their ranks. Fifteen of the sixteen tests were significant, confirming the reported heteroscedasticity. Assessing the dependence of the variances on Gene is more problematic due to the small number of replications, so it is safer to assume such dependence (Kerr et al. did not). Thus, there do not seem to be any patterns in the heteroscedasticity and, according to Remark 5.1, we used the delete-one jackknife bias corrected estimators of σ^4_{ijk}. The p-values of all three tests considered in the simulations (CF, OBS and RANK), based on both the original fluorensence measurements and their log transforms, are reported in Table 1 below.

Table 3.12: P-values of the hypotheses tests

Effect	logs of measurements			Original measurements		
	CF	Obs	Rank	CF	Obs	Rank
Gene	.000	.000	.000	.000	.000	.000
Trt	.000	.000	.000	.000	.000	.000
Dye	.000	.000	.000	.518	.518	.000
Trt×Dye	.000	.000	.000	.000	.000	.000
Gene×Dye	.000	.000	.999	.999	.999	.999
Gene×Trt	.841	.796	.862	.000	.000	.862
Gene×Dye×Trt	.967	.976	.981	.999	.979	.981

We first note that, as expected, the RANK analysis gives the same results for both the original observations and their log scales. However, there are differences between the CF and OBS analyses regarding the Dye effect and the Gene×Dye, Gene×Treatment interactions. These differences are probably due to the fact that the linear hypotheses are not invariant under non-linear transformations; thus a significant (non-significant) interaction or main effect may become non-significant (significant) after a non-linear transformation. Of course, the instability of both CF and OBS to skewness and the sensitivity of CF to heteroscedasticity, all of which were observed in the simulations, might also contribute to the observed discrepancy of the analyses. In particular the difference regarding the Gene×Treatment interaction is critical because it is one of the focal points of the Kerr, et al. (2002) investigation. Using the log scale, the CF and OBS analyses agree with the RANK analysis in all cases except the Gene×Dye interaction. In fact, Kerr, et al. (2002) had put some effort trying to remove the Gene×Dye interaction from their model, and our RANK analysis suggests that this could be done. On the other hand, Kerr, et al. (2002) included neither Treatment nor Treatment×Dye effects in their model, and both are significant.

3.6 Proofs

3.6.1 Proofs for Section 3.2

Tests for main effects

Proof of Theorem 3.2.1 Because under $H_0(A)$, $\mathbf{C}_A E(\mathbf{W}) = \mathbf{0}$, we have $\mathbf{C}_A\mathbf{W} = \mathbf{C}_A(\mathbf{W} - E(\mathbf{W}))$. Thus the result will follow by showing $\sqrt{N}(\mathbf{W} - E(\mathbf{W})) \xrightarrow{d} N_a(\mathbf{0}, \mathbf{V})$, where $\mathbf{V} = diag\{\eta_1, \ldots, \eta_a\}$, and using the Continuous Mapping and Slutsky's Theorems, since $\widehat{\eta}_i$ are consistent estimators of η_i, $i = 1, \ldots, a$. By the independence of the $\widetilde{X}_{i..}$, and since a is fixed, the asymptotic normality of $\mathbf{W}$ will follow from that of each $\widetilde{X}_{i..}$. write $\sqrt{N}(\widetilde{X}_{i..} - E(\widetilde{X}_{i..})) = \sqrt{N}\widetilde{e}_{i..}$, where $e_{ijk} = X_{ijk} - E(X_{ijk})$. By Lyapounov's theorem,

$$\frac{\sqrt{N}\widetilde{e}_{i..}}{\eta_{i,N}} \to N(0,1), \text{ where } \eta_{i,N} = \mathrm{Var}(\sqrt{N}\widetilde{e}_{i..}) = \frac{N}{b^2}\sum_{j=1}^{b}\mathrm{Var}(\bar{e}_{ij.}) = \frac{N}{b^2}\sum_{j=1}^{b}\frac{\sigma_{ij}^2}{n_{ij}},$$

because $L(b) = \frac{1}{\eta_{i,N}^2}\sum_{j=1}^{b}\sum_{k=1}^{n_{ij}} E\left|\frac{\sqrt{N}}{bn_{ij}}e_{ijk}\right|^4 = \left(\sum_{j=1}^{b}\frac{\sigma_{ij}^2}{n_{ij}}\right)^{-2}\sum_{j=1}^{b}\left(\frac{1}{n_{ij}^3}E|e_{ijk}^4|\right) \to 0$, by assumption (3.2.1) and Hölder's inequality. □

Proof of Theorem 3.2.2 By Lemmas 3.7.2, 3.7.3, we only need to consider the asymptotic distribution of $n(a,b)\sqrt{a}(P_A(\mathbf{e}) - MSE)$ under $H_0(A)$, where $n(a,b) = \min_{i,j}\{n_{ij}\}$, and $P_A(\mathbf{e}) = \frac{b}{a}\sum_{i=1}^{a}\widetilde{e}_{i..}^2$, with $e_{ijk} = X_{ijk} - E(X_{ijk})$. Write

$$n(a,b)\sqrt{a}(P_A(\mathbf{e}) - MSE) = T_{1A}(\mathbf{e}) + T_{3A}(\mathbf{e}), \tag{3.6.1}$$

where

$$T_{1A}(\mathbf{e}) = \frac{n(a,b)}{b\sqrt{a}}\sum_{i=1}^{a}\sum_{j\neq j'}^{b}\bar{e}_{ij.}\bar{e}_{ij'.}, \quad T_{3A}(\mathbf{e}) = \frac{n(a,b)}{b\sqrt{a}}\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{m\neq m'}^{n_{ij}}\frac{e_{ijm}e_{ijm'}}{n_{ij}(n_{ij}-1)}. \tag{3.6.2}$$

It is easy to see that $E(T_{1A}(\mathbf{e}))=E(T_{3A}(\mathbf{e}))=0$ and, as $a,b\to\infty$,

$$\mathrm{Var}(T_{3A}(\mathbf{e})) = \frac{2n^2(a,b)}{ab^2}\sum_{i=1}^{a}\sum_{j=1}^{b}\frac{\sigma_{ij}^4}{n_{ij}(n_{ij}-1)} \le \frac{4}{ab^2}\sum_{i=1}^{a}\sum_{j=1}^{b}\left(\frac{n(a,b)}{n_{ij}}\sigma_{ij}^2\right)^2 \to 0, \quad (3.6.3)$$

$$\mathrm{Var}(T_{1A}(\mathbf{e})) = \frac{2n^2(a,b)}{b^2 a}\sum_{i=1}^{a}\sum_{j\neq j'}^{b}\left(\frac{\sigma_{ij}^2}{n_{ij}}\right)\left(\frac{\sigma_{ij'}^2}{n_{ij'}}\right) = n^2(a,b)\tau_{1,A}+o(1). \quad (3.6.4)$$

By (3.6.1), (3.6.3), (3.6.4) it suffices to show the asymptotic distribution of $T_{1A}(\mathbf{e})$. We will use Lyapounov's theorem. Since (3.6.4) is bounded, Lyapounov's condition will be satisfied if $L(a,b)=\sum_{i=1}^{a}E\left|(ab^2)^{-1/2}\sum_{j\neq j'}^{b}n(a,b)\bar{e}_{ij.}\bar{e}_{ij'.}\right|^4\to 0$. We have,

$$\begin{aligned} L(a,b) &= \sum_{i=1}^{a}\sum_{j\neq j'}^{b}\sum_{j_1\neq j_1'}^{b}\sum_{j_2\neq j_2'}^{b}\sum_{j_3\neq j_3'}^{b}\frac{n^4(a,b)}{a^2b^4}E(\bar{e}_{ij.}\bar{e}_{ij'.}\bar{e}_{ij_1.}\bar{e}_{ij_1'.}\bar{e}_{ij_2.}\bar{e}_{ij_2'.}\bar{e}_{ij_3.}\bar{e}_{ij_3'.}) \quad (3.6.5)\\ &= O\left(\sum_{i=1}^{a}\sum_{j\neq j'\neq j_1\neq j_1'}^{b}\frac{n^4(a,b)}{a^2b^4}E(\bar{e}_{ij.})^2E(\bar{e}_{ij'.})^2E(\bar{e}_{ij_1.})^2E(\bar{e}_{ij_1'.})^2\right)\\ &= O\left(\frac{1}{a^2b^4}\sum_{i=1}^{a}\left(\sum_{j=1}^{b}\frac{n(a,b)}{n_{ij}}\sigma_{ij}^2\right)^4\right) = O\left(\frac{b^3}{a^2b^4}\sum_{i=1}^{a}\sum_{j=1}^{b}\sigma_{ij}^8\right) = O(a^{-1}), \end{aligned}$$

where the second equality follows from the fact that when the number of different elements among $\{j,j',j_1,j_1',j_2,j_2',j_3,j_3'\}$ is five or more the expectation on the right hand side of (3.6.5) is zero, and the fourth equality is due to the inequality

$$\left|\sum_{i=1}^{m}z_i\right|^p \le m^{p-1}\sum_{i=1}^{m}|z_i|^p,\ m\geq 1,\ p\geq 1, \quad (3.6.6)$$

which for $p>1$ follows from Hölder's inequality. This completes the proof. □

Tests for two-way interaction

Proof of Theorem 3.2.3 Because under $H_0(AB)$, $\mathbf{C}_{AB}E(\mathbf{W})=\mathbf{0}$, we have $\mathbf{C}_{AB}\mathbf{W}=\mathbf{C}_{AB}(\mathbf{W}-E(\mathbf{W}))$. Thus the result will follow by showing $\sqrt{N}(\mathbf{W}-E(\mathbf{W}))\xrightarrow{d}N_{ab}(\mathbf{0},\mathbf{V})$, where $\mathbf{V}=diag\{\eta_{11},\ldots,\eta_{1b},\ldots,\eta_{a1},\ldots,\eta_{ab}\}$, and using the Continuous Mapping and Slutsky's Theorems, since $\widehat{\eta}_{ij}$ are consistent estimators of η_{ij}, $i=1,\ldots,a$, $j=1,\ldots,b$. By the independence of the $\widetilde{X}_{ij..}$, and since a,b are fixed, the asymptotic normality of $\mathbf{W}$ will follow from that of each $\widetilde{X}_{ij..}$. This is shown by an application of Lyapounov's theorem similar to the one in the proof of Theorem 3.4.1. □

Proof of Theorem 3.2.4 By Lemmas 3.7.4 and 3.7.5, we only need to consider the asymptotic distribution of $n(a,b,c)\sqrt{a}(P_{2,AB}(\mathbf{e})-MSE)$ under $H_0(AB)$, where $e_{ijk}=X_{ijkm}-E(X_{ijkm})$. We will use Lyapounov's Theorem. Write

$$n(a,b,c)\sqrt{a}(P_{2,AB}(\mathbf{e})-MSE)=\frac{1}{\sqrt{a}}\sum_{i=1}^{a}W_{i,AB},$$

where

$$W_{i,AB}=\frac{n(a,b,c)}{bc}\left[-\frac{c^2}{b-1}\sum_{j\neq j'}^{b}\widetilde{e}_{ij..}\widetilde{e}_{ij'..}+\sum_{j=1}^{b}\sum_{k\neq k'}^{c}\widetilde{e}_{ijk.}\widetilde{e}_{ijk'.}+\sum_{j=1}^{b}\sum_{k=1}^{c}\sum_{m\neq m'}^{n_{ijk}}\frac{e_{ijkm}e_{ijkm'}}{n_{ijk}(n_{ijk}-1)}\right],$$

are independent with zero mean. Next, some algebra gives,

$$\begin{aligned}
&\mathrm{Var}\left(n(a,b,c)\sqrt{a}(P_{2,AB}(\mathbf{e})-MSE)\right)\\
=\;&\frac{2n^2(a,b,c)}{b^2c^2a}\sum_{i=1}^{a}\left[\frac{1}{(b-1)^2}\sum_{j\neq j'}^{b}\left(\sum_{k=1}^{c}\frac{\sigma^2_{ijk}}{n_{ijk}}\right)\left(\sum_{k'=1}^{c}\frac{\sigma^2_{ij'k'}}{n_{ij'k'}}\right)+\sum_{j=1}^{b}\sum_{k\neq k'}^{c}\left(\frac{\sigma^2_{ijk}}{n_{ijk}}\right)\left(\frac{\sigma^2_{ijk'}}{n_{ijk'}}\right)\right]\\
&+\frac{2n^2(a,c)}{b^2c^2a}\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{c}\frac{\sigma^4_{ijk}}{n_{ijk}(n_{ijk}-1)}=n^2(a,b,c)[\tau_1^2+\tau_2^2+\tau_3^2],
\end{aligned}$$

where τ_1^2, τ_2^2, τ_3^2 are defined in the statement of the theorem, stays bounded away from zero and ∞. Thus, Lyapounov's condition will be satisfied if $\sum_{i=1}^{a}E\left|\frac{1}{\sqrt{a}}W_{i,AB}\right|^4\to 0$. Using the inequality (3.6.6) we have

$$\begin{aligned}
E(W^4_{i,AB})\;\leq\;&\frac{n^4(a,b,c)}{b^4c^4}3^3\left[E\left|\frac{c^2}{b-1}\sum_{j\neq j'}^{b}\widetilde{e}_{ij..}\widetilde{e}_{ij'..}\right|^4+E\left|\sum_{j=1}^{b}\sum_{k\neq k'}^{c}\overline{e}_{ijk.}\overline{e}_{ijk'.}\right|^4\right.\\
&\left.+\;E\left|\sum_{j=1}^{b}\sum_{k=1}^{c}\sum_{m\neq m'}^{n_{ijk}}\frac{e_{ijkm}e_{ijkm'}}{n_{ijk}(n_{ijk}-1)}\right|^4\right] \qquad (3.6.7)
\end{aligned}$$

Noting that b is fixed, the second term above gives

$$\frac{n^4(a,b,c)}{a^2c^4}\sum_{i=1}^{a}E\left|\sum_{j=1}^{b}\sum_{k\neq k'}^{c}\bar{e}_{ijk.}\bar{e}_{ijk'.}\right|^4$$

$$\begin{aligned}
&= \sum_{i=1}^{a}\sum_{j,j_1,j_2,j_3}^{b}\sum_{k\neq k'}^{c}\sum_{k_1\neq k_1'}^{c}\sum_{k_2\neq k_2'}^{c}\sum_{k_3\neq k_3'}^{c}\frac{n^4(a,b,c)}{a^2c^4}E(\bar{e}_{ijk.}\bar{e}_{ijk'.}\bar{e}_{ij_1k_1.}\bar{e}_{ij_1k_1'.}\bar{e}_{ij_2k_2.}\bar{e}_{ij_2k_2'.}\bar{e}_{ij_3k_3.}\bar{e}_{ij_3k_3'.})\\
&= O\left(\sum_{i=1}^{a}\sum_{j,j_1}^{b}\sum_{k\neq k'\neq k_1\neq k_1'}^{c}\frac{n^4(a,b,c)}{a^2c^4}E(\bar{e}^2_{ijk.})E(\bar{e}^2_{ijk'.})E(\bar{e}^2_{ij_1k_1.})E(\bar{e}^2_{ij_1k_1'.})\right)\\
&= O\left(\sum_{i=1}^{a}\sum_{j,j_1}^{b}\sum_{k\neq k'\neq k_1\neq k_1'}^{c}\frac{n^4(a,b,c)}{a^2c^4}\frac{\sigma^2_{ijk}\sigma^2_{ijk'}\sigma^2_{ij_1k_1}\sigma^2_{ij_1k_1'}}{n_{ijk}n_{ijk'}n_{ij_1k_1}n_{ij_1k_1'}}\right)\\
&= O\left(\frac{1}{a^2c^4}\sum_{i=1}^{a}\left(\sum_{j=1}^{b}\sum_{k=1}^{c}\frac{n(a,b,c)}{n_{ijk}}\sigma^2_{ijk}\right)^4\right) = O\left(\frac{b^3c^3}{a^2c^4}\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{c}\sigma^8_{ijk}\right) = O(a^{-1}), \quad (3.6.8)
\end{aligned}$$

where the second equality is because the expectation under the summation is zero when the number of different elements in $\{j, j_1, j_2, j_3\}$ is more than two or the number of different elements in $\{k, k', k_1, k_1', k_2, k_2', k_3, k_3'\}$ is more than four, and the 5th equality is due to (3.6.6). Similarly, the first term in (3.6.7) gives $\frac{n^4(a,c)}{a^2c^4}\sum_{i=1}^{a}E\left|\sum_{j\neq j'}^{b}\tilde{e}_{ij..}\tilde{e}_{ij'..}\right|^4 = O(a^{-1})$. Finally, the last term in (3.6.7) gives

$$\begin{aligned}
&\frac{n^4(a,b,c)}{a^2c^4}\sum_{i=1}^{a}E\left|\sum_{j,k}\sum_{m\neq m'}^{n_{ijk}}\frac{e_{ijkm}e_{ijkm'}}{n_{ijk}(n_{ijk}-1)}\right|^4 = \frac{n^4(a,b,c)}{a^2c^4}\sum_{i=1}^{a}\sum_{j,j_1,j_2,j_3}^{b}\sum_{k,k_1,k_2,k_3}^{c}\sum_{m\neq m'}^{n_{ijk}}\sum_{m_1\neq m_1'}^{n_{ijk_1}}\\
&\qquad\sum_{m_2\neq m_2'}^{n_{ijk_2}}\sum_{m_3\neq m_3'}^{n_{ijk_3}}\frac{E(e_{ijkm}e_{ijkm'}e_{ij_1k_1m_1}e_{ij_1k_1m_1'}e_{ij_2k_2m_2}e_{ij_2k_2m_2'}e_{ij_3k_3m_3}e_{ij_3k_3m_3'})}{n_{ijk}(n_{ijk}-1)n_{ij_1k_1}(n_{ij_1k_1}-1)n_{ij_2k_2}(n_{ij_2k_2}-1)n_{ij_3k_3}(n_{ij_3k_3}-1)}\\
&= O\left(\frac{n^4(a,b,c)}{a^2c^4}\sum_{i=1}^{a}\sum_{j,j_1}^{b}\sum_{k,k_1}^{c}\sum_{m\neq m'}^{n_{ijk}}\sum_{m_1\neq m_1'}^{n_{ij_1k_1}}\frac{E(e^2_{ijkm})E(e^2_{ijkm'})E(e^2_{ij_1k_1m_1})E(e^2_{ij_1k_1m_1'})}{n^2_{ijk}(n_{ijk}-1)^2n^2_{ij_1k_1}(n_{ij_1k_1}-1)^2}\right)\\
&= O\left(\frac{n^4(a,b,c)}{a^2c^4}\sum_{i=1}^{a}\left(\sum_{j,k}\frac{\sigma^4_{ijk}}{n_{ijk}(n_{ijk}-1)}\right)^2\right) = O\left(\frac{n^4(a,b,c)c}{a^2c^4}\sum_{i,j,k}\frac{\sigma^8_{ijk}}{n^2_{ijk}(n_{ijk}-1)^2}\right)\\
&= O\left(a^{-1}c^{-2}\right). \qquad (3.6.9)
\end{aligned}$$

Thus Lyapounov's condition is satisfied and this completes the proof. □

Proof of Theorem 3.2.5 By Lemmas 3.7.4, 3.7.5 we only need to consider the asymptotic distribution of $n(a,b,c)\sqrt{ab}(P_{1,AB}(\mathbf{e})-MSE)$, under $H_0(AB)$, where $e_{ijkm}=X_{ijkm}-E(X_{ijkm})$. We will apply Lyapounov's Theorem. Write

$$n(a,b,c)\sqrt{ab}(P_{1,AB}(\mathbf{e})-MSE)=(ab)^{-1/2}\sum_{i=1}^{a}\sum_{j=1}^{b}W_{ij,AB},$$

where

$$W_{ij,AB}=\frac{n(a,b,c)}{c}\sum_{k=1}^{c}\sum_{m\neq m'}\frac{e_{ijkm}e_{ijkm'}}{n_{ijk}(n_{ijk}-1)}+\frac{n(a,b,c)}{c}\sum_{k\neq k'}^{c}\bar{e}_{ijk.}\bar{e}_{ijk'.}$$

are independent with zero mean. Thus,

$$\begin{aligned}\text{Var}\left(\frac{1}{\sqrt{ab}}\sum_{i=1}^{a}\sum_{j=1}^{b}W_{ij,AB}\right) &= \frac{2n^2(a,b,c)}{abc^2}\left[\sum_{i,j,k}\frac{\sigma_{ijk}^4}{n_{ijk}(n_{ijk}-1)}+\sum_{i,j}\sum_{k\neq k'}^{c}\left(\frac{\sigma_{ijk}^2}{n_{ijk}}\right)\left(\frac{\sigma_{ijk'}^2}{n_{ijk'}}\right)\right]\\ &= \frac{2n^2(a,b,c)}{abc^2}\sum_{i,j,k}\frac{\sigma_{ijk}^4}{n_{ijk}^2(n_{ijk}-1)}+\frac{2n^2(a,b,c)}{abc^2}\sum_{i,j}\left(\sum_{k=1}^{c}\frac{\sigma_{ijk}^2}{n_{ijk}}\right)^2\\ &= n^2(a,b,c)\left[b\tau_1+\frac{(b-1)^2}{b-2}\tau_2\right],\end{aligned}$$

which converges to a constant, as $a\to\infty$ and $b\to\infty$ regardless of whether c, according to the stated assumptions. So Lyapounov's condition will be satisfied if

$$\sum_{i=1}^{a}\sum_{j=1}^{b}E\left|\frac{1}{\sqrt{ab}}W_{ij,AB}\right|^4=\frac{1}{a^2b^2}\sum_{i=1}^{a}\sum_{j=1}^{b}E(W_{ij,AB}^4)\to 0.$$

To show this, note first that (3.6.6) yields

$$E|W_{ij,AB}^4| \leq 2^3\frac{[n(a,b,c)]^4}{c^4}\left[E\left|\sum_{k=1}^{c}\sum_{m\neq m'}\frac{e_{ijkm}e_{ijkm'}}{n_{ijk}(n_{ijk}-1)}\right|^4+E\left|\sum_{k\neq k'}^{c}\bar{e}_{ijk.}\bar{e}_{ijk'.}\right|^4\right],$$

and then use calculations similar to (3.6.8) and (3.6.9), to obtain

$$\frac{n^4(a,b,c)}{a^2b^2c^4}\sum_{i,j}E\left|\sum_{k=1}^{c}\sum_{m\neq m'}^{n_{ijk}}\frac{e_{ijkm}e_{ijkm'}}{n_{ijk}(n_{ijk}-1)}\right|^4=O\left(a^{-1}b^{-1}c^{-2}\right),$$

$$\frac{n^4(a,b,c)}{a^2b^2c^4}\sum_{i,j}E\left|\sum_{k\neq k'}^{c}\bar{e}_{ijk.}\bar{e}_{ijk'.}\right|^4=O(a^{-1}b^{-1}).$$

□

Proof of Theorem 3.2.6 Because under $H_0(ABC)$, $\mathbf{C}_{ABC}E(\mathbf{W}) = \mathbf{0}$, we have $\mathbf{C}_{ABC}\mathbf{W} = \mathbf{C}_{ABC}(\mathbf{W} - E(\mathbf{W}))$. Thus, by the Continuous Mapping and Slutsky's Theorems, the result follows from $\sqrt{N}(\mathbf{W} - E(\mathbf{W})) \xrightarrow{d} N_{abc}(\mathbf{0}, \mathbf{V})$, where $\mathbf{V}$ is the diagonal matrix which is estimated consistently by $\widehat{\mathbf{V}}$ in (3.1.12). By the independence of the $\widetilde{X}_{ijk..}$, and since a, b, c are fixed, the asymptotic normality of $\mathbf{W}$ will follow from that of each $\widetilde{X}_{ijk..}$. This is shown by Lyapounov's theorem as was done in the proof of Theorem 3.4.1.□

Proof of Theorem 3.2.7 By Slutsky's Theorem and Lemmas 3.7.7 and 3.7.6, it suffices to consider the asymptotic distribution of $n(a,b,c,d)\sqrt{ac}(P_{1,ABC}(\mathbf{e}) - MSE)$ under $H_0(ABC)$, where $e_{ijklm} = X_{ijklm} - E(X_{ijklm})$. To apply Lyapounov's Theorem, write

$$n(a,b,c,d)\sqrt{ac}(P_{1,ABC}(\mathbf{e}) - MSE) = (\sqrt{ac})^{-1} \sum_{i=1}^{a} \sum_{k=1}^{c} W_{ik},$$

where

$$W_{ik} = \frac{n(a,b,c,d)}{bd} \left[\sum_{j,l} \sum_{m \neq m'} \frac{e_{ijklm} e_{ijklm'}}{n_{ijkl}(n_{ijkl} - 1)} + \sum_{j=1}^{b} \sum_{l \neq l'}^{d} \overline{e}_{ijkl.} \overline{e}_{ijkl'.} + \frac{d^2}{(b-1)} \sum_{j \neq j'}^{b} \widetilde{e}_{ijk..} \widetilde{e}_{ij'k..} \right]$$

are independent with zero mean. Thus,

$$\begin{aligned}
& \operatorname{Var}\left(n(a,b,c,d)\sqrt{ac}(P_{1,ABC}(\mathbf{e}) - MSE)\right) \\
= \; & \frac{2n^2(a,b,c,d)}{acb^2d^2} \sum_{i,j,k,l} \frac{\sigma^4_{ijkl}}{n_{ijkl}(n_{ijkl}-1)} + \frac{2n^2(a,b,c,d)}{acb^2d^2} \sum_{i,j,k} \sum_{l \neq l'}^{d} \frac{\sigma^2_{ijkl}}{n_{ijkl}} \frac{\sigma^2_{ijkl'}}{n_{ijkl'}} \\
& + \frac{2n^2(a,b,c,d)}{acb^2d^2(b-1)^2} \sum_{i,k} \sum_{j \neq j'}^{b} \left(\sum_{l=1}^{d} \frac{\sigma^2_{ijkl}}{n_{ijkl}} \right) \left(\sum_{l'=1}^{d} \frac{\sigma^2_{ij'kl'}}{n_{ij'kl'}} \right) \\
= \; & \frac{2n^2(a,b,c,d)}{acb^2d^2} \sum_{i,k} \left[\sum_{j,l} \frac{\sigma^4_{ijkl}}{n^2_{ijkl}(n_{ijkl}-1)} + \frac{b^2-2b}{(b-1)^2} \sum_{j} \left(\sum_{l=1}^{d} \frac{\sigma^2_{ijkl}}{n_{ijkl}} \right)^2 + \frac{1}{(b-1)^2} \left(\sum_{j,l} \frac{\sigma^2_{ijkl}}{n_{ijkl}} \right)^2 \right] \\
= \; & n^2(a,b,c,d)c(\tau_4 + \tau_5 + \tau_6),
\end{aligned}$$

which is bounded away from zero and ∞. So, Lyapounov's condition will hold if

$$\sum_{i=1}^{a} \sum_{k=1}^{c} E \left| \frac{1}{\sqrt{ac}} W_{ik} \right|^4 = \frac{1}{a^2c^2} \sum_{i=1}^{a} \sum_{k=1}^{c} E(W^4_{ik}) \to 0. \tag{3.6.10}$$

By inequality (3.6.6),

$$E|W_{ik}^4| \le 3^3\frac{[n(a,b,c,d)]^4}{(bd)^4}\left[E\left|\sum_{j,l}\sum_{m\neq m'}\frac{e_{ijklm}e_{ijklm'}}{n_{ijkl}(n_{ijkl}-1)}\right|^4+E\left|\sum_{j=1}^{b}\sum_{l\neq l'}^{d}\bar{e}_{ijkl.}\bar{e}_{ijkl'.}\right|^4\right.$$
$$\left.+\left[\frac{1}{(b-1)}\right]^4 E\left|\sum_{j\neq j'}^{b}\sum_{l,l'}\bar{e}_{ijkl.}\bar{e}_{ij'kl'.}\right|^4\right].$$

Similarly to the proof of (3.6.8) and (3.6.9), we obtain that (3.6.10) is bounded by

$$3^3\frac{n^4(a,b,c,d)}{a^2c^2b^4d^4}\sum_{i,k}\left\{E\left|\sum_{j,l}\sum_{m\neq m'}\frac{e_{ijklm}e_{ijklm'}}{n_{ijkl}(n_{ijkl}-1)}\right|^4+E\left|\sum_{j=1}^{b}\sum_{l\neq l'}^{d}\bar{e}_{ijkl.}\bar{e}_{ijkl'.}\right|^4+\frac{1}{(b-1)^4}\right.$$
$$\left.\times\quad E\left|\sum_{j\neq j'}^{b}\sum_{l,l'}\bar{e}_{ijkl.}\bar{e}_{ij'kl'.}\right|^4\right\}=O\left(a^{-1}c^{-1}b^{-2}d^{-2}\right)+O\left(a^{-1}c^{-1}\right)+O\left(a^{-1}c^{-1}(b-1)^{-4}\right).$$

These imply that Lyapounov's condition is satisfied and complte the proof. □

Proof of Theorem 3.2.8 By Lemmas 3.7.6, 3.7.7, and Slutsky's Theorem, it suffices to consider the asymptotic distribution of $n(a,b,c,d)\sqrt{a}(P_{2,ABC}(\mathbf{e})-MSE)$ under $H_0(ABC)$, where $e_{ijklm}=X_{ijklm}-E(X_{ijklm})$. Again we will apply Lyapounov's Theorem. Write $n(a,b,c,d)\sqrt{a}(P_{2,ABC}(\mathbf{e})-MSE)=a^{-1/2}\sum_{i=1}^{a}W_{i,ABC}$, where

$$W_{i,ABC} = \frac{n(a,b,c,d)}{bcd}\left[\sum_{j,k,l}\sum_{m\neq m'}\frac{e_{ijklm}e_{ijklm'}}{n_{ijkl}(n_{ijkl}-1)}+\sum_{j,k}\sum_{l\neq l'}^{d}\bar{e}_{ijkl.}\bar{e}_{ijkl'.}\right]$$
$$+\frac{n(a,b,c,d)d}{c(b-1)(c-1)}\left\{\frac{c-1}{b}\sum_{k=1}^{c}\sum_{j\neq j'}^{b}\widetilde{e}_{ijk..}\widetilde{e}_{ij'k..}+b\sum_{k\neq k'}^{c}\widetilde{e}_{i.k..}\widetilde{e}_{i.k'..}-\sum_{k\neq k'}^{c}\sum_{j=1}^{b}\widetilde{e}_{ijk..}\widetilde{e}_{ijk'..}\right\},$$

are independent with mean zero. Thus

$$
\begin{aligned}
&\mathrm{Var}\left(n(a,b,c,d)\sqrt{a}(P_{2,ABC}(\mathbf{e})-MSE)\right)\\
=\;& \frac{2n^2(a,b,c,d)}{ac^2b^2d^2}\sum_{i,j,k,l}\frac{\sigma^4_{ijkl}}{n_{ijkl}(n_{ijkl}-1)}+\frac{2n^2(a,b,c,d)}{ac^2b^2d^2}\sum_{i,j,k}\sum_{l\neq l'}^{d}\frac{\sigma^2_{ijkl}}{n_{ijkl}}\frac{\sigma^2_{ijkl'}}{n_{ijkl'}}\\
&+\frac{2n^2(a,b,c,d)}{ac^2b^2d^2(b-1)^2}\sum_{i,k}\sum_{j\neq j'}^{b}\left(\sum_{l=1}^{d}\frac{\sigma^2_{ijkl}}{n_{ijkl}}\right)\left(\sum_{l'=1}^{d}\frac{\sigma^2_{ij'kl'}}{n_{ij'kl'}}\right)\\
&+\frac{2n^2(a,b,c,d)}{ac^2d^2(b-1)^2(c-1)^2}\sum_{i,j}\sum_{k\neq k'}^{c}\sum_{l,l'}\left[\frac{\sigma^2_{ijkl}}{n_{ijkl}}\frac{\sigma^2_{ijk'l'}}{n_{ijk'l'}}+\frac{1}{b^2}\sum_{j'=1}^{b}\frac{\sigma^2_{ijkl}}{n_{ijkl}}\frac{\sigma^2_{ij'k'l'}}{n_{ij'k'l'}}\right]\\
=\;& \frac{2n^2(a,c,d)}{ac^2b^2d^2}\sum_{i,k}\left[\sum_{j,l}\frac{\sigma^4_{ijkl}}{n^2_{ijkl}(n_{ijkl}-1)}+\frac{b^2-2b}{(b-1)^2}\sum_{j}\left(\sum_{l=1}^{d}\frac{\sigma^2_{ijkl}}{n_{ijkl}}\right)^2+\frac{1}{(b-1)^2}\left(\sum_{j,l}\frac{\sigma^2_{ijkl}}{n_{ijkl}}\right)^2\right]\\
&+\frac{2n^2(a,b,c,d)}{ac^2d^2(b-1)^2(c-1)^2}\sum_{i,j}\sum_{k\neq k'}^{c}\sum_{l,l'}\left[\frac{\sigma^2_{ijkl}}{n_{ijkl}}\frac{\sigma^2_{ijk'l'}}{n_{ijk'l'}}+\frac{1}{b^2}\sum_{j'=1}^{b}\frac{\sigma^2_{ijkl}}{n_{ijkl}}\frac{\sigma^2_{ij'k'l'}}{n_{ij'k'l'}}\right]\\
=\;& n^2(a,b,c,d)(\tau_4+\tau_5+\tau_6+\tau_7+\tau_8).
\end{aligned}
$$

Because the above variance stays bounded away from zero and ∞ as $a\to\infty$ while b and c remain fixed, regardless of whether d, Lyapounov's condition will be satisfied if , $L(a)=\sum_{i=1}^{a}E\left|a^{-1/2}W_{i,ABC}\right|^4\to 0$. Similarly to the proof for (3.6.10), we can show that $L(a)=O(a^{-1})$. This completes the proof. □

Proof of Theorem 3.2.9 By Lemmas 3.7.6, 3.7.7, and Slutsky's Theorem, it suffices to consider the asymptotic distribution of $n(a,b,c,d)\sqrt{abc}(P_{3,ABC}(\mathbf{e})-MSE)$ under $H_0(ABC)$, where $e_{ijklm}=X_{ijklm}-E(X_{ijklm})$. To apply Lyapounov's Theorem, write $n(a,b,c,d)\sqrt{abc}(P_{3,ABC}(\mathbf{e})-MSE)=\frac{1}{\sqrt{abc}}\sum_{i,j,k}W_{ijk}$, where

$$
W_{ijk}=\frac{n(a,b,c,d)}{d}\left[\sum_{l=1}^{d}\sum_{m\neq m'}^{n_{ijkl}}\frac{e_{ijklm}e_{ij'klm'}}{n_{ijkl}(n_{ijkl}-1)}+\sum_{l\neq l'}^{d}\bar{e}_{ijkl\cdot}\bar{e}_{ijkl'\cdot}\right],
$$

are independent with mean zero. Thus

$$\begin{aligned}
& \mathrm{Var}\left(n(a,b,c,d)\sqrt{abc}(P_{3,ABC}(\mathbf{e})-MSE)\right) \\
= \;& \frac{2n^2(a,b,c,d)}{acbd^2}\sum_{i,j,k,l}\frac{\sigma_{ijkl}^4}{n_{ijkl}(n_{ijkl}-1)}+\frac{2n^2(a,b,c,d)}{acbd^2}\sum_{i,j,k}\sum_{l\neq l'}^{d}\frac{\sigma_{ijkl}^2}{n_{ijkl}}\frac{\sigma_{ijkl'}^2}{n_{ijkl'}} \\
= \;& \frac{2n^2(a,b,c,d)}{acbd^2}\sum_{i,j,k,l}\frac{\sigma_{ijkl}^4}{n_{ijkl}^2(n_{ijkl}-1)}+\frac{2n^2(a,b,c,d)}{acbd^2}\sum_{i,j,k}\left(\sum_{l=1}^{d}\frac{\sigma_{ijkl}^2}{n_{ijkl}}\right)^2 \\
= \;& n^2(a,b,c,d)[bc\tau_4+c(b-1)^2/(b-2)\tau_5],
\end{aligned}$$

and Lyapounov's condition can be easily verified, regardless of whether $d\to\infty$ or not, as in the proof of Theorem 3.2.7. □

3.6.2 Proofs for Section 3.3

In order to include all ordinal data (continuous, discrete, and data with ties) in our formulation, we make use of the representation of the (mid-)ranks via the empirical distribution function (edf). In particular, if $\widehat{H}$ denotes the average of the left and right continuous version of the edf, then the (mid)-rank of an observation with value x is $N\widehat{H}(x)+0.5$. For example, in the case of §3.6.3 below, $\widehat{H}(t)=N^{-1}\sum_i\sum_j\sum_k c(X_{ijk},t)$, where the function c is defined in the last paragraph of Section 1, and the (mid)-rank of X_{ijk} is $R_{ijk}=N\widehat{H}(X_{ijk}))+0.5$. Because all test statistics we consider are invariant under location and scale changes, we will simply set $R_{ijk}=N\widehat{H}(X_{ijk})$.

3.6.3 Tests for main effects

Proof of Theorem 3.4.1 Let $Q_H(A)$, $Q_{\widehat{H}}(A)$ be the $Q_R(A)$ statistic evaluated at $H(X_{ijk})$, $\widehat{H}(X_{ijk})$, respectively, and similarly let $\mathbf{W}_H$, $\widehat{V}_H$, $\mathbf{W}_{\widehat{H}}$, $\widehat{V}_{\widehat{H}}$, be the $\mathbf{W}$, $\widehat{V}$ given below (3.1.3) again with X_{ijk}^* replaced by $H(X_{ijk})$, $\widehat{H}(X_{ijk})$, respectively. Theorem 3.2.1 and Remark 3.1

imply that the $Q_H(A)$ converges in distribution to χ^2_{a-1}. Thus it suffices to establish that

$$Q_{\widehat{H}}(A) - Q_H(A) = \sqrt{N}\left(\mathbf{W}'_{\widehat{H}} - \mathbf{W}'_H\right)\mathbf{C}'_A(\mathbf{C}_A\widehat{\mathbf{V}}_H\mathbf{C}'_A)^{-1}\mathbf{C}_A\sqrt{N}\mathbf{W}_{\widehat{H}}$$
$$+\sqrt{N}\mathbf{W}'_H\mathbf{C}'_A(\mathbf{C}_A\widehat{\mathbf{V}}_H\mathbf{C}'_A)^{-1}\mathbf{C}_A\sqrt{N}\left(\mathbf{W}_{\widehat{H}} - \mathbf{W}_H\right)$$
$$+\sqrt{N}\mathbf{C}'_A\mathbf{W}'_{\widehat{H}}\left[(\mathbf{C}_A\widehat{\mathbf{V}}_{\widehat{H}}\mathbf{C}'_A)^{-1} - (\mathbf{C}_A\widehat{\mathbf{V}}_H\mathbf{C}'_A)^{-1}\right]\mathbf{C}_A\sqrt{N}\mathbf{W}_{\widehat{H}} \to 0, \qquad (3.6.11)$$

since $Q_R(A) = Q_{\widehat{H}}(A)$. Given that the elements of $\widehat{\mathbf{V}}_{\widehat{H}}$ and $\widehat{\mathbf{V}}_H$ stay bounded away from zero and infinity, the first two expressions on the right hand side of (3.6.11) will be shown to converge to zero under $H_0(A)$ if

$$\mathbf{C}_A(\mathbf{W}_{\widehat{H}} - \mathbf{W}_H) = \mathbf{C}_A\int(\widehat{H} - H)\,d\widehat{\mathbf{F}} = \mathbf{C}_A\int(\widehat{H} - H)\,d(\widehat{\mathbf{F}} - \mathbf{F}) = o_p(N^{-1/2}), \qquad (3.6.12)$$

where $\mathbf{F} = (\overline{F}_{1.}, \cdots, \overline{F}_{a.})'$ and $\widehat{\mathbf{F}} = (\widehat{\overline{F}}_{1.}, \cdots, \widehat{\overline{F}}_{a.})'$, with $\overline{F}_{i.}(x) = b^{-1}\sum_{j=1}^{b} F_{ij}(x)$ and $\widehat{\overline{F}}_{i.}(x) = b^{-1}\sum_{j=1}^{b}\widehat{F}_{ij}(x)$, where $\widehat{F}_{ij}(x) = n_{ij}^{-1}\sum_{k=1}^{n_{ij}} c(X_{ijk}, x)$. Note that the second equality in (3.6.12) holds only under $H_0(A)$. For (3.6.12) we will show that for each $i = 1, \ldots, a$,

$$\sqrt{N}\int(\widehat{H} - H)\,d(\widehat{\overline{F}}_{i.} - \overline{F}_{i.}) \xrightarrow{p} 0. \qquad (3.6.13)$$

A similar result is shown in Akritas and Arnold (1994) but under fixed number of factor levels and group sizes tending to infinity, which is not directly applicable to our situation. The proof of (3.6.13) is given in Lemma 3.7.1. Finally, the last expression on the right hand side of (3.6.11) converges to zero in probability by $\widehat{\mathbf{V}}_{\widehat{H}} - \widehat{\mathbf{V}}_H \to 0$ and the fact that $N^{1/2}\mathbf{C}_A\mathbf{W}_{\widehat{H}}$ is bounded in probability, which follows from (3.6.12), and Theorem 3.2.1. □

Proof of Theorem 3.4.2 In this proof, we will keep the notations $Y_{ijm} = H(X_{ijm})$, and R_{ijm} =(mid-)rank of X_{ijm}, and further will denote $Z_{ijm} = \widehat{H}(X_{ijm})$. Set $\mathbf{Y} = (Y_{111}, \cdots, Y_{abn_{ab}})$ and let $\mathbf{Z}$, $\mathbf{R}$ be similarly defined. To be clear, we will use $MST_A(\mathbf{Y})$, $MST_A(\mathbf{Z})$ and $MST_A(\mathbf{R})$ to denote the MST_A statistic defined in connection with (3.1.4) evaluated on $\mathbf{Y}$, $\mathbf{Z}$ and $\mathbf{R}$ respectively. $MSE(\mathbf{R})$, $MSE(\mathbf{Z})$ and $MSE(\mathbf{Y})$ are defined similarly. Note that $MST_A(\mathbf{R})/N^2 = MST_A(\mathbf{Z})$. By Lemmas 3.7.8 and 3.7.9, it suffices to establish the asymptotic distribution of $T(\mathbf{Z}-E(\mathbf{Y})) = n(a,b)\sqrt{a}[P_A(\mathbf{Z}-E(\mathbf{Y})) - MSE(\mathbf{Z}))]$. We will do so by showing $T(\mathbf{Z}-E(\mathbf{Y})) - T(\mathbf{Y}-E(\mathbf{Y})) = o_p(1)$, using the fact that by the proof of Theorem 3.2.2 on $\mathbf{Y}$, we have $T(\mathbf{Y}-E(\mathbf{Y}))/\sqrt{\tau_{1,A}} \xrightarrow{d} N(0,1)$. Write $T(\mathbf{Z}-E(\mathbf{Y})) = T_{1A}(\mathbf{Z}-E(\mathbf{Y})) + T_{3A}(\mathbf{Z}-E(\mathbf{Y}))$, where T_{1A} and T_{3A} are defined in (3.6.2). Using a similar decomposition for $T(\mathbf{Y}-E(\mathbf{Y}))$, it follows that to show $T(\mathbf{Z}-E(\mathbf{Y})) - T(\mathbf{Y}-E(\mathbf{Y})) = o_p(1)$, we only need to show $T_{sA}(\mathbf{Z}-E(\mathbf{Y})) - T_{sA}(\mathbf{Y}-E(\mathbf{Y})) = o_p(1)$, for $s = 1,3$. These proofs are similar to that of $\sqrt{b}[D_6(\mathbf{Z}-E(\mathbf{Y})) - D_6(\mathbf{Y}-E(\mathbf{Y}))] = o_p(1)$ in the proof of Lemma 3.7.13 and thus are omitted. □

Proofs for results of two-way interaction

Proof of Theorem 3.4.3 Let $Q_H(AB)$, $Q_{\widehat{H}}(AB)$ be the $Q_R(AB)$ statistic evaluated at $H(X_{ijkm})$, $\widehat{H}(X_{ijkm})$, respectively, and similarly let $\mathbf{W}_H$, $\widehat{V}_H$, $\mathbf{W}_{\widehat{H}}$, $\widehat{V}_{\widehat{H}}$, be the $\mathbf{W}$, $\widehat{V}$ given below (3.1.8) again with X^*_{ijkm} replaced by $H(X_{ijkm})$, $\widehat{H}(X_{ijkm})$, respectively. By Theorem 3.2.3, and adopting Remark 3.1 in the setting of that theorem, it follows that $Q_H(AB)$ converges in distribution to $\chi^2_{(a-1)(b-1)}$. Thus it suffices to establish that

$$\begin{aligned} Q_{\widehat{H}}(AB) - Q_H(AB) \;=\; & \sqrt{N}\left(\mathbf{W}'_{\widehat{H}} - \mathbf{W}'_H\right)\mathbf{C}'_{AB}(\mathbf{C}_{AB}\widehat{\mathbf{V}}_H\mathbf{C}'_{AB})^{-1}\mathbf{C}_{AB}\sqrt{N}\mathbf{W}_{\widehat{H}} \\ & +\sqrt{N}\mathbf{W}'_H\mathbf{C}'_{AB}(\mathbf{C}_{AB}\widehat{\mathbf{V}}_H\mathbf{C}'_{AB})^{-1}\mathbf{C}_{AB}\sqrt{N}\left(\mathbf{W}_{\widehat{H}} - \mathbf{W}_H\right) \\ & +\sqrt{N}\mathbf{C}'_{AB}\mathbf{W}'_{\widehat{H}}\left[(\mathbf{C}_{AB}\widehat{\mathbf{V}}_{\widehat{H}}\mathbf{C}'_{AB})^{-1} - (\mathbf{C}_{AB}\widehat{\mathbf{V}}_H\mathbf{C}'_{AB})^{-1}\right]\mathbf{C}_{AB}\sqrt{N}\mathbf{W}_{\widehat{H}} \to 0, \qquad (3.6.14) \end{aligned}$$

since $Q_R(AB) = Q_{\widehat{H}}(AB)$. Given that the elements of $\widehat{\mathbf{V}}_{\widehat{H}}$ and $\widehat{\mathbf{V}}_H$ stay bounded away from zero and infinity, the first two expressions on the right hand side of (3.6.14) will be shown to converge to zero under $H_0(AB)$ if

$$\mathbf{C}_{AB}(\mathbf{W}_{\widehat{H}} - \mathbf{W}_H) = \mathbf{C}_{AB}\int(\widehat{H}-H)\,d\widehat{\mathbf{F}} = \mathbf{C}_{AB}\int(\widehat{H}-H)\,d(\widehat{\mathbf{F}}-\mathbf{F}) = o_p(N^{-1/2}), \quad (3.6.15)$$

where $\mathbf{F} = (\overline{F}_{11.},\cdots,\overline{F}_{1b.},\cdots,\overline{F}_{a1.}\cdots,\overline{F}_{ab.})'$ and $\widehat{\mathbf{F}} = (\widehat{\overline{F}}_{11.},\cdots,\widehat{\overline{F}}_{1b.},\cdots,\widehat{\overline{F}}_{a1.}\cdots,\widehat{\overline{F}}_{ab.})'$, with $\overline{F}_{ij.}(x) = c^{-1}\sum_{k=1}^{c}F_{ijk}(x)$ and $\widehat{\overline{F}}_{ij.}(x) = c^{-1}\sum_{k=1}^{c}\widehat{F}_{ijk}(x)$, where $\widehat{F}_{ijk}(x) = n_{ijk}^{-1}\sum_{m=1}^{n_{ijk}} c(X_{ijkm},x)$. Note that the second equality in (3.6.15) holds only under $H_0(AB)$. Relation (3.6.15) follows from the fact that for all $i = 1,\dots,a$, $j = 1,\dots,b$,

$$\sqrt{N}\int(\widehat{H}-H)\,d(\widehat{\overline{F}}_{ij.}-\overline{F}_{ij.}) \xrightarrow{p} 0,$$

which is shown in Lemma 3.7.1. The last expression on the right hand side of (3.6.14) converges to zero in probability by $\widehat{\mathbf{V}}_{\widehat{H}} - \widehat{\mathbf{V}}_H \to 0$, in probability, and the fact that $N^{1/2}\mathbf{C}_{AB}\mathbf{W}_{\widehat{H}}$ is bounded in probability, which follows (3.6.15) and Theorem 3.2.3. □

In the following two proofs, we will keep the notations $Y_{ijkm} = H(X_{ijkm})$, and $R_{ijkm} =$ (mid-)rank of X_{ijkm}, and further will denote $Z_{ijkm} = \widehat{H}(X_{ijkm})$. Note $R_{ijkm} = NZ_{ijkm} + 0.5$. Also denote $\mathbf{Y} = (Y_{1111},\cdots,Y_{abcn_{abc}})$ and let $\mathbf{Z}$, $\mathbf{R}$ be similarly defined. To be clear, we will use $MST_{AB}(\mathbf{Y})$, $MST_{AB}(\mathbf{Z})$ and $MST_{AB}(\mathbf{R})$ to denote the MST_{AB} statistic defined in connection with (3.1.9) evaluated on $\mathbf{Y}$, $\mathbf{Z}$ and $\mathbf{R}$ respectively. $MSE(\mathbf{R})$, $MSE(\mathbf{Z})$ and $MSE(\mathbf{Y})$ are defined similarly. Note that $MST_{ABC}(\mathbf{R})/N^2 = MST_{ABC}(\mathbf{Z})$.

Proof of Theorem 3.4.4 By Lemmas 3.7.10 and 3.7.11, it suffices to establish the asymptotic distribution of $T_2(\mathbf{Z}-E(\mathbf{Y})) = n(a,b,c)\sqrt{a}[P_{2,AB}(\mathbf{Z}-E(\mathbf{Y})) - MSE(\mathbf{Z}))]$. We will do so by showing the asymptotic equivalence of $T_2(\mathbf{Z}-E(\mathbf{Y}))$ and $T_2(\mathbf{Y}-E(\mathbf{Y}))$, using the fact that by the proof of Theorem 3.2.4 we have $T_2(\mathbf{Y}-E(\mathbf{Y}))/\sqrt{\tau_1^2+\tau_2^2+\tau_3^2} \xrightarrow{d} N(0,1)$. Write $T_2(\mathbf{Z}-E(\mathbf{Y})) = T_{1AB}(\mathbf{Z}-E(\mathbf{Y})) + T_{2AB}(\mathbf{Z}-E(\mathbf{Y})) + T_{3AB}(\mathbf{Z}-E(\mathbf{Y}))$, where

$$T_{1AB}(\mathbf{Z}-E(\mathbf{Y})) = -\frac{n(a,b,c)c}{b(b-1)\sqrt{a}}\sum_{i=1}^{a}\sum_{j\neq j'}^{b}(\widetilde{Z}_{ij..}-\overline{p}_{ij.})(\widetilde{Z}_{ij'..}-\overline{p}_{ij'.}),$$

$$T_{2AB}(\mathbf{Z}-E(\mathbf{Y})) = \frac{n(a,b,c)}{bc\sqrt{a}}\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k\neq k'}^{c}(\widetilde{Z}_{ijk.}-p_{ijk})(\widetilde{Z}_{ijk'.}-p_{ijk'}) \qquad (3.6.16)$$

$$T_{3AB}(\mathbf{Z}-E(\mathbf{Y})) = \frac{n(a,b,c)}{bc\sqrt{a}}\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{c}\sum_{m\neq m'}^{n_{ijk}}\frac{(Z_{ijkm}-p_{ijk})(Z_{ijkm'}-p_{ijk})}{n_{ijk}(n_{ijk}-1)}. \qquad (3.6.17)$$

Using a similar decomposition for $T_2(\mathbf{Y}-E(\mathbf{Y}))$, it follows that to show $T_2(\mathbf{Z}-E(\mathbf{Y})) - T_2(\mathbf{Y}-E(\mathbf{Y})) = o_p(1)$, we only need to show $T_{sAB}(\mathbf{Z}-E(\mathbf{Y})) - T_{sAB}(\mathbf{Y}-E(\mathbf{Y})) = o_p(1)$, for $s = 1,2,3$. These proofs are similar to that of $\sqrt{b}[D_6(\mathbf{Z}-E(\mathbf{Y})) - D_6(\mathbf{Y}-E(\mathbf{Y}))] = o_p(1)$ in the proof of Lemma 3.7.13 and thus are omitted. □

Proof of Theorem 3.4.5 By Lemma 3.7.10 By Lemmas 3.7.10 and 3.7.11, it suffices to establish the asymptotic distribution of $T_1(\mathbf{Z}-E(\mathbf{Y})) = n(a,b,c)\sqrt{ab}[P_{1,AB}(\mathbf{Z}-E(\mathbf{Y}))-MSE(\mathbf{Z}))]$. We will do so by showing the asymptotic equivalence of $T_1(\mathbf{Z}-E(\mathbf{Y}))$ and $T_1(\mathbf{Y}-E(\mathbf{Y}))$, using the fact that by the proof of Theorem 3.2.5 we have $T_1(\mathbf{Y}-E(\mathbf{Y}))/\sqrt{b\tau_1^2+(b-1)^2/(b-2)\tau_2^2} \xrightarrow{d} N(0,1)$. Write $T_1(\mathbf{Z}-E(\mathbf{Y})) = \sqrt{b}[T_{2AB}(\mathbf{Z}-E(\mathbf{Y}))+T_{3AB}(\mathbf{Z}-E(\mathbf{Y}))]$, where T_{2AB} and T_{3AB} are given in (3.6.16) and (3.6.17). Using a similar decomposition for $T_1(\mathbf{Y}-E(\mathbf{Y}))$, it follows that to show $T_1(\mathbf{Z}-E(\mathbf{Y}))-T_1(\mathbf{Y}-E(\mathbf{Y})) = o_p(1)$, we only need to show $\sqrt{b}[T_{sAB}(\mathbf{Z}-E(\mathbf{Y}))-T_{sAB}(\mathbf{Y}-E(\mathbf{Y}))] = o_p(1)$, for $s=2,3$, These proofs are again similar to that of $\sqrt{b}[D_6(\mathbf{Z}-E(\mathbf{Y}))-D_6(\mathbf{Y}-E(\mathbf{Y}))] = o_p(1)$ given in the proof of Lemma 3.7.13 and thus are omitted. □

Proofs for results of three-way interaction

Proof of Theorem 3.4.6 Let $Q_H(ABC)$, $Q_{\widehat{H}}(ABC)$ be the $Q_R(ABC)$ statistic evaluated at $H(X_{ijklm})$, $\widehat{H}(X_{ijklm})$, respectively, and similarly let $\mathbf{W}_H$, $\widehat{V}_H$, $\mathbf{W}_{\widehat{H}}$, $\widehat{V}_{\widehat{H}}$, be the $\mathbf{W}$, $\widehat{V}$ given below (3.1.12) again with X^*_{ijklm} replaced by $H(X_{ijklm})$, $\widehat{H}(X_{ijklm})$, respectively. By Theorem 3.2.6, and adopting Remark 3.1 in the setting of that theorem, it follows that $Q_H(ABC)$ converges in distribution to $\chi^2_{(a-1)(b-1)(c-1)}$. Thus it suffices to establish that

$$
\begin{aligned}
Q_{\widehat{H}}(ABC) - Q_H(ABC) \;=\; & \sqrt{N}\left(\mathbf{W}'_{\widehat{H}}-\mathbf{W}'_H\right)\mathbf{C}'_{ABC}(\mathbf{C}_{ABC}\widehat{\mathbf{V}}_H\mathbf{C}'_{ABC})^{-1}\mathbf{C}_{ABC}\sqrt{N}\mathbf{W}_{\widehat{H}} \\
& + \sqrt{N}\mathbf{W}'_H\mathbf{C}'_{ABC}(\mathbf{C}_{ABC}\widehat{\mathbf{V}}_H\mathbf{C}'_{ABC})^{-1}\mathbf{C}_{ABC}\sqrt{N}\left(\mathbf{W}_{\widehat{H}}-\mathbf{W}_H\right) \\
& + \sqrt{N}\mathbf{C}'_{ABC}\mathbf{W}'_{\widehat{H}}\left[(\mathbf{C}_{ABC}\widehat{\mathbf{V}}_{\widehat{H}}\mathbf{C}'_{ABC})^{-1}-(\mathbf{C}_{ABC}\widehat{\mathbf{V}}_H\mathbf{C}'_{ABC})^{-1}\right]\mathbf{C}_{ABC}\sqrt{N}\mathbf{W}_{\widehat{H}} \to 0, \quad (3.6.18)
\end{aligned}
$$

since $Q_R(ABC) = Q_{\widehat{H}}(ABC)$. Given that the elements of $\widehat{\mathbf{V}}_{\widehat{H}}$ and $\widehat{\mathbf{V}}_H$ stay bounded away from zero and infinity, the first two expressions on the right hand side of (3.6.18) will be shown to converge to zero under $H_0(ABC)$ if $\mathbf{C}_{ABC}(\mathbf{W}_{\widehat{H}}-\mathbf{W}_H) = o_p(N^{-1/2})$, or

$$
\mathbf{C}_{ABC}\int(\widehat{H}-H)\,d\widehat{\mathbf{F}} = \mathbf{C}_{ABC}\int(\widehat{H}-H)\,d(\widehat{\mathbf{F}}-\mathbf{F}) = o_p(N^{-1/2}), \qquad (3.6.19)
$$

where $\mathbf{F} = (\overline{F}_{111.}, \cdots, \overline{F}_{11c.}, \cdots, \overline{F}_{1b1.}, \cdots, \overline{F}_{1bc.}, \cdots, \overline{F}_{a11.} \cdots, \overline{F}_{abc.})'$ and $\widehat{\mathbf{F}} = (\widehat{\overline{F}}_{111.}, \cdots, \widehat{\overline{F}}_{11c.}, \cdots, \widehat{\overline{F}}_{1b1.}, \cdots, \widehat{\overline{F}}_{1bc.}, \cdots, \widehat{\overline{F}}_{a11.} \cdots, \widehat{\overline{F}}_{abc.})'$, with $\overline{F}_{ijk.}(x) = d^{-1}\sum_{l=1}^{d} F_{ijkl}(x)$ and $\widehat{\overline{F}}_{ij.}(x) = d^{-1}\sum_{l=1}^{d} \widehat{F}_{ijkl}(x)$, where $\widehat{F}_{ijkl}(x) = n_{ijkl}^{-1}\sum_{m=1}^{n_{ijkl}} c(X_{ijklm}, x)$. Note that the second equality in (3.6.19) holds only under $H_0(ABC)$. Relation (3.6.19) follows from the fact that for all $i = 1, \ldots, a, j = 1, \ldots, b,$

$$\sqrt{N}\int(\widehat{H} - H)\, d(\widehat{\overline{F}}_{ijk.} - \overline{F}_{ijk.}) \xrightarrow{p} 0,$$

which is shown in Lemma 3.7.1. The last expression on the right hand side of (3.6.18) converges to zero in probability by $\widehat{\mathbf{V}}_{\widehat{H}} - \widehat{\mathbf{V}}_H \to 0$, in probability, and the fact that $N^{1/2}\mathbf{C}_{AB}\mathbf{W}_{\widehat{H}}$ is bounded in probability, which follows (3.6.19) and Theorem 3.2.6.

For the remainder part of this section, we will keep the notations $Y_{ijklm} = H(X_{ijklm})$, and R_{ijklm} =(mid-)rank of X_{ijklm}, and further will denote $Z_{ijklm} = \widehat{H}(X_{ijklm})$. Note $R_{ijklm} = NZ_{ijklm} + 0.5$. Also denote $\mathbf{Y} = (Y_{11111}, \cdots, Y_{abcdn_{abcd}})$ and let $\mathbf{Z}$, $\mathbf{R}$ be similarly defined. To be clear, we will use $MST_{ABC}(\mathbf{Y})$, $MST_{ABC}(\mathbf{Z})$ and $MST_{ABC}(\mathbf{R})$ to denote the MST_{ABC} statistic defined in connection with (3.1.13) evaluated on $\mathbf{Y}$, $\mathbf{Z}$ and $\mathbf{R}$ respectively. $MSE(\mathbf{R})$, $MSE(\mathbf{Z})$ and $MSE(\mathbf{Y})$ are defined similarly. Note that $MST_{ABC}(\mathbf{R})/N^2 = MST_{ABC}(\mathbf{Z})$.

Proof of Theorem 3.4.7 By Lemmas 3.7.12 and 3.7.13, it suffices to establish the asymptotic distribution of $T_1(\mathbf{Z} - E(\mathbf{Y})) = n(a,b,c,d)\sqrt{ac}[P_{1,ABC}(\mathbf{Z} - E(\mathbf{Y})) - MSE(\mathbf{Z}))]$. We will do so by showing the asymptotic equivalence of $T_2(\mathbf{Z} - E(\mathbf{Y}))$ and $T_2(\mathbf{Y} - E(\mathbf{Y}))$, using the fact that by the proof of Theorem 3.2.7, we have $T_1(\mathbf{Y} - E(\mathbf{Y}))/\sqrt{c(\tau_4 + \tau_5 + \tau_6)} \xrightarrow{d} N(0,1)$. Thus it suffices to show $T_1(\mathbf{Z} - E(\mathbf{Y})) - T_1(\mathbf{Y} - E(\mathbf{Y})) = o_p(1)$ as $a, c \to \infty$ while b remains fixed regardless of whether d and n_{ijkl} large or small. Write

$$T_1(\mathbf{Y} - E(\mathbf{Y})) = T_{11}(\mathbf{Y} - E(\mathbf{Y})) + T_{12}(\mathbf{Y} - E(\mathbf{Y})) - \frac{T_{13}(\mathbf{Y} - E(\mathbf{Y}))}{b-1},$$

where

$$T_{11}(\mathbf{Y}-E(\mathbf{Y})) = \frac{n(a,b,c,d)}{bd\sqrt{ac}} \sum_{i,j,k,l} \sum_{m\neq m'}^{n_{ijkl}} \frac{(Y_{ijklm}-p_{ijkl})(Y_{ij'klm'}-p_{ij'kl})}{n_{ijkl}(n_{ijkl}-1)} \quad (3.6.20)$$

$$T_{12}(\mathbf{Y}-E(\mathbf{Y})) = \frac{n(a,b,c,d)}{bd\sqrt{ac}} \sum_{i,j,k} \sum_{l\neq l'}^{d} (\overline{Y}_{ijkl.}-p_{ijkl})(\overline{Y}_{ijkl'.}-p_{ijkl'}), \quad (3.6.21)$$

$$T_{13}(\mathbf{Y}-E(\mathbf{Y})) = \frac{dn(a,b,c,d)}{b\sqrt{ac}} \sum_{i,k} \sum_{j\neq j'}^{b} (\widetilde{Y}_{ijk..}-\overline{p}_{ijk.})(\widetilde{Y}_{ij'k..}-\overline{p}_{ij'k.}). \quad (3.6.22)$$

Using a similar decomposition for $T_1(\mathbf{Z}-E(\mathbf{Y}))$, it follows that to show $T_1(\mathbf{Z}-E(\mathbf{Y}))-T_1(\mathbf{Y}-E(\mathbf{Y}))=o_p(1)$, we only need to show $T_{1s}(\mathbf{Z}-E(\mathbf{Y}))-T_{1s}(\mathbf{Y}-E(\mathbf{Y}))=o_p(1)$, for $s=1,2,3$, (note that b is fixed). These proofs are similar to that of $\sqrt{b}[D_6(\mathbf{Z}-E(\mathbf{Y}))-D_6(\mathbf{Y}-E(\mathbf{Y}))]=o_p(1)$ in the proof of Lemma 3.7.13 and thus are omitted. □

Proof of Theorem 3.4.8 By Lemmas 3.7.12 and 3.7.13, it suffices to establish the asymptotic distribution of $T_2(\mathbf{Z}-E(\mathbf{Y}))=n(a,b,c,d)\sqrt{a}[P_{2,ABC}(\mathbf{Z}-E(\mathbf{Y}))-MSE(\mathbf{Z}))]$. We will do so by showing the asymptotic equivalence of $T_2(\mathbf{Z}-E(\mathbf{Y}))$ and $T_2(\mathbf{Y}-E(\mathbf{Y}))$, using the fact that by the proof of Theorem 3.2.8 we have $T_2(\mathbf{Y}-E(\mathbf{Y}))/\sqrt{\tau_4+\tau_5+\tau_6+\tau_7+\tau_8} \xrightarrow{d} N(0,1)$. So it suffices to show $T_2(\mathbf{Z}-E(\mathbf{Y}))-T_2(\mathbf{Y}-E(\mathbf{Y}))=o_p(1)$ as $a\to\infty$ while while b,c remain fixed regardless of whether d and n_{ijkl} large or small. Write

$$\begin{aligned} T_2(\mathbf{Y}-E(\mathbf{Y})) \;\; = \;\; & \frac{T_{11}(\mathbf{Y}-E(\mathbf{Y}))}{\sqrt{c}} + \frac{T_{12}(\mathbf{Y}-E(\mathbf{Y}))}{\sqrt{c}} - \frac{T_{13}(\mathbf{Y}-E(\mathbf{Y}))}{\sqrt{c}(b-1)} \\ & + \frac{D_4(\mathbf{Y}-E(\mathbf{Y}))}{\sqrt{c}} - \frac{D_1(\mathbf{Y}-E(\mathbf{Y}))}{\sqrt{c}(b-1)}, \end{aligned}$$

where $T_{1s}(\mathbf{Y}-E(\mathbf{Y}))$, $s=1,2,3$, are defined in (3.6.20), (3.6.21) and (3.6.22), and D_4, D_1 are defined in (3.7.7) and (3.7.4), respectively. Using a similar decomposition for $T_2(\mathbf{Z}-E(\mathbf{Y}))$, the results from the previous proof, and noting that b,c are fixed, it follows that to show $T_2(\mathbf{Z}-E(\mathbf{Y}))-T_2(\mathbf{Y}-E(\mathbf{Y}))=o_p(1)$, it remains to show

$$D_s(\mathbf{Z}-E(\mathbf{Y})-D_s(\mathbf{Y}-E(\mathbf{Y})=o_p(1), \text{ for } s=1,4 \,.$$

These proofs are again similar to that of $\sqrt{b}[D_6(\mathbf{Z}-E(\mathbf{Y}))-D_6(\mathbf{Y}-E(\mathbf{Y}))]=o_p(1)$ in the proof of Lemma 3.7.13 and thus are omitted. □

Proof of Theorem 3.4.9 By Lemmas 3.7.12 and 3.7.13, it suffices to establish the asymptotic distribution of $T_3(\mathbf{Z}-E(\mathbf{Y}))=n(a,b,c,d)\sqrt{abc}[P_{3,ABC}(\mathbf{Z}-E(\mathbf{Y}))-MSE(\mathbf{Z}))]$. We will do so by showing the asymptotic equivalence of $T_3(\mathbf{Z}-E(\mathbf{Y}))$ and $T_3(\mathbf{Y}-E(\mathbf{Y}))$, using the fact that by the proof of Theorem 3.2.9, $T_3(\mathbf{Y}-E(\mathbf{Y}))/\sqrt{bc\tau_4+c(b-1)^2/(b-2)\tau_5}\xrightarrow{d}N(0,1)$. So it suffices to show that the difference $T_3(\mathbf{Z}-E(\mathbf{Y}))-T_3(\mathbf{Y}-E(\mathbf{Y}))$ converges in probability to zero as $a,b,c\to\infty$ regardless of whether d and n_{ijkl} large or small. Write

$$T_3(\mathbf{Y}-E(\mathbf{Y}))=\sqrt{b}T_{11}(\mathbf{Y}-E(\mathbf{Y}))+\sqrt{b}T_{12}(\mathbf{Y}-E(\mathbf{Y})),$$

where $T_{11}(\cdot)$ and $T_{12}(\cdot)$ are defined in (3.6.20) and (3.6.21), respectively. Using a similar decomposition for $T_3(\mathbf{Z}-E(\mathbf{Y}))$ it follows that $T_3(\mathbf{Z}-E(\mathbf{Y}))-T_3(\mathbf{Y}-E(\mathbf{Y}))=o_p(1)$, will be implied by $\sqrt{b}[T_{1s}(\mathbf{Z}-E(\mathbf{Y})-T_{1s}(\mathbf{Y}-E(\mathbf{Y})]=o_p(1)$, for $s=1,2$, These proofs are similar to that of $\sqrt{b}[D_6(\mathbf{Z}-E(\mathbf{Y}))-D_6(\mathbf{Y}-E(\mathbf{Y}))]=o_p(1)$ in the proof of Lemma 3.7.13 and thus are omitted. □

3.7 Supporting Proofs

3.7.1 Some Auxiliary Results

Lemma 3.7.1. *Let $j=1,\ldots,b$, index the levels of a factor whose number of levels tends to infinity, and let i index the levels of all other factors. Let $\widehat{H}$ denote the (average of the left- and right-continuous versions of the) empirical distribution function of all observations and $H=E\left(\widehat{H}\right)$. Then, regardless of whether or not the cell sizes tend to ∞,*

$$\sqrt{N}\int(\widehat{H}-H)\,d(\widehat{\overline{F}}_{i.}-\overline{F}_{i.})\xrightarrow{p}0,\ \ \text{as } b\to\infty. \tag{3.7.1}$$

Proof The left hand side of (3.7.1) is

$$\frac{1}{b\sqrt{N}}\sum_{i',j}n_{i'j}\sum_{j_1=1}^{b}\int(\widehat{F}_{i'j}-F_{i'j})\,d(\widehat{F}_{ij_1}-F_{ij_1})=\frac{1}{b\sqrt{N}}\sum_{j_1=1}^{b}\sum_{k_1=1}^{n_{ij_1}}\frac{1}{n_{ij_1}}\sum_{i',j}\sum_{k_2=1}^{n_{i'j}}h(X_{i'jk_2},X_{ij_1k_1}),$$

where

$$h(X_{i'jk_2}, X_{ij_1k_1}) = c(X_{i'jk_2}, X_{ij_1k_1}) - F_{i'j}(X_{ij_1k_1}) - \left[1 - F_{ij_1}(X_{i'jk_2}) - \int F_{i'j}\, dF_{ij_1}\right].$$

Thus,

$$\begin{aligned}
& E\left[\sqrt{N}\int (\widehat{H} - H)\, d(\widehat{\overline{F}}_{i.} - \overline{F}_{i.})\right]^2 \\
= \; & \frac{1}{b^2 N} \sum_{j_1=1}^{b} \sum_{k_1=1}^{n_{ij_1}} \sum_{j_3=1}^{b} \sum_{k_3=1}^{n_{ij_3}} \frac{1}{n_{ij_1} n_{ij_3}} \sum_{i',j} \sum_{k_2=1}^{n_{i'j}} \sum_{i_4,j_4} \sum_{k_4=1}^{n_{i_4j_4}} E\left[h(X_{i'jk_2}, X_{ij_1k_1}) h(X_{i_4j_4k_4}, X_{ij_3k_3})\right] \\
= \; & \frac{1}{b^2 N} \sum_{j_1=1}^{b} \sum_{k_1=1}^{n_{ij_1}} \sum_{i',j} \sum_{k_2=1}^{n_{i'j}} \frac{1}{n_{ij_1} n_{ij_1}} E[h^2(X_{i'jk_2}, X_{ij_1k_1})] I(j \neq j_1) I(k_1 \neq k_2) \\
& + \frac{1}{b^2 N} \sum_{j_1=1}^{b} \sum_{k_1=1}^{n_{ij_1}} \sum_{j=1}^{b} \sum_{k_2=1}^{n_{ij}} \frac{1}{n_{ij_1} n_{ij}} E[h(X_{ijk_2}, X_{ij_1k_1}) h(X_{ij_1k_1}, X_{ijk_2})]
\end{aligned}$$

where the second equality holds because $E[h(X_{i_1}, X_{i_2}) h(X_{i_3}, X_{i_4})] = 0$ if the number of different elements in $\{i_1, i_2, i_3, i_4\}$ is four or three due to independence and

$$E[h(X_{i'jk_2}, X_{ij_1k_1})] = E[h(X_{i'jk_2}, X_{ij_1k_1}) | X_{i'jk_2}] = E[h(X_{i'jk_2}, X_{ij_1k_1}) | X_{ij_1k_1}] = 0,$$

Since $|h(\cdot,\cdot)|$ is uniformly bounded by 4, we have

$$E\left[\sqrt{N}\int (\widehat{H} - H)\, d(\widehat{\overline{F}}_{i.} - \overline{F}_{i.})\right]^2 \leq \frac{1}{b^2} \sum_{j_1=1}^{b} \frac{16}{n_{ij_1}} + \frac{16}{N} \to 0,$$

as $b \to \infty$, regardless of whether n_{ij}, which implies that 3.6.13) holds.

Lemma 3.7.2. *Under the settings and assumptions of Theorem 3.2.2, we have*

$$n(a,b)(MSE - \sigma_A^2) \xrightarrow{p} 0, \text{ where } n(a,b) = \min_{i,j}\{n_{ij}\}.$$

Proof Note that $E(MSE) = E\left[\frac{1}{ab}\sum_{i=1}^{a}\sum_{j=1}^{b}\frac{S_{ij,X}^2}{n_{ij}}\right] = \frac{1}{ab}\sum_{i=1}^{a}\sum_{j=1}^{b}\frac{\sigma_{ij}^2}{n_{ij}} = \sigma_A^2$, and

$$
\begin{aligned}
(ab)^2\mathrm{Var}(MSE) &= \sum_{i=1}^{a}\sum_{j=1}^{b}\frac{1}{n_{ij}^2(n_{ij}-1)^2}\mathrm{Var}\left(\frac{n_{ij}-1}{n_{ij}}\sum_{m=1}^{n_{ij}}e_{ijm}^2 - \frac{1}{n_{ij}}\sum_{m\neq m'}^{n_{ij}}e_{ijm}e_{ijm'}\right)\\
&= \sum_{i=1}^{a}\sum_{j=1}^{b}\frac{1}{n_{ij}^4}\left[\sum_{m=1}^{n_{ij}}\left(E(e_{ijm}^4)-\sigma_{ij}^4\right)+\frac{2n_{ij}}{n_{ij}-1}\sigma_{ij}^4\right]\\
&= \sum_{i=1}^{a}\sum_{j=1}^{b}\frac{1}{n_{ij}^4}\left[\sum_{m=1}^{n_{ij}}E(e_{ijm}^4)-\frac{n_{ij}(n_{ij}-3)}{n_{ij}-1}\sigma_{ij}^4\right].
\end{aligned}
$$

So $\mathrm{Var}(n(a,b)MSE) \leq \frac{1}{(ab)^2}\sum_{i=1}^{a}\sum_{j=1}^{b}\frac{n^2(a,b)}{n_{ij}^4}\sum_{m=1}^{n_{ij}}E(e_{ijm}^4) \to 0$ as $a\to\infty, b\to\infty$. Therefore, $n(a,b)[MSE-\sigma_A^2] \xrightarrow{p} 0$.

Lemma 3.7.3. *Define* $P_A(\mathbf{e}) = \frac{b}{a}\sum_{i=1}^{a}\widetilde{e}_{i..}^2$, *where* $e_{ijk} = X_{ijk} - E(X_{ijk})$. *Under the settings and assumptions of Theorem 3.2.2,under* $H_0(A)$, *we have*

$$T_A^*(\mathbf{e}) = n(a,b)\sqrt{a}(MST_A - P_A(\mathbf{e})) \xrightarrow{p} 0, \text{ as } a,b\to\infty, \text{ where } n(a,b) = \min_{i,j}\{n_{ij}\},$$

regardless of whether n_{ij} *remain fixed or go to infinity with* a,b.

Proof Under $H_0(A)$, MST_A can also be expressed in terms of the e_{ijk}'s and we have $T_A^*(\mathbf{e}) = -\frac{bn(a,b)}{\sqrt{a}(a-1)}\sum_{i\neq i'}^{a}\widetilde{e}_{i..}\widetilde{e}_{i'..}$. It remains to show $T_A^*(\mathbf{e}) \xrightarrow{p} 0$. Note that by independence, $E(T_A(\mathbf{e})) = 0$, and

$$
\begin{aligned}
E(T_A^*(\mathbf{e})^2) &= \frac{[n(a,b)b]^2}{a(a-1)^2}\sum_{i\neq i'}^{a}\sum_{i_1\neq i_2}^{a}E(\widetilde{e}_{i..}\widetilde{e}_{i'..}\widetilde{e}_{i_1..}\widetilde{e}_{i_2..}) = \frac{2[n(a,b)b]^2}{a(a-1)^2}\sum_{i\neq i'}^{a}E(\widetilde{e}_{i..}^2)E(\widetilde{e}_{i'..}^2)\\
&= \frac{2n^2(a,b)}{a(a-1)^2b^2}\sum_{i\neq i'}^{a}\sum_{j=1}^{b}\sum_{j'=1}^{b}\frac{\sigma_{ij}^2}{n_{ij}}\frac{\sigma_{i'j'}^2}{n_{i'j'}} \leq \frac{2}{a(a-1)^2b^2}\left(\sum_{i=1}^{a}\sum_{j=1}^{b}\frac{n(a,b)}{n_{ij}}\sigma_{ij}^2\right)^2,
\end{aligned}
$$

which converges to zero. This completes the proof.

Lemma 3.7.4. *Under the settings and assumptions of either Theorem 3.2.4 or 3.2.5 , we have*

$$n(a,b,c)(MSE-\sigma_{AB}^2) \xrightarrow{p} 0, \quad \text{where } n(a,b,c) = \min_{i,j,k}\{n_{ijk}\}.$$

Proof The proof is similar to that of Lemma 3.7.2, and thus it is omitted.

Lemma 3.7.5. *Set* $n(a,b,c) = \min_{i,j,k}\{n_{ijk}\}$, $P_{1,AB}(\mathbf{e}) = \frac{c}{ab}\sum_{i=1}^{a}\sum_{j=1}^{b}\tilde{e}_{ij..}^2$, *where* $e_{ijkm} = X_{ijkm} - E(X_{ijkm})$, *and* $P_{2,AB}(\mathbf{e}) = P_{1,AB}(\mathbf{e}) - \frac{c}{ab(b-1)}\sum_{i=1}^{a}\sum_{j\neq j'}^{b}\tilde{e}_{ij..}\tilde{e}_{ij'..}$. *Then under* $H_0(AB)$, *and regardless of whether* c *and* n_{ijk} *remain fixed or go to infinity,*

1. *under the settings and assumptions of Theorem 3.2.5,*
   $T_{1AB}^*(\mathbf{e}) = n(a,b,c)\sqrt{ab}(MST_{AB} - P_{1,AB}(\mathbf{e})) \xrightarrow{p} 0$, *as* $a,b\to\infty$;

2. *under the settings and assumptions of Theorem 3.2.4,*
   $T_{2AB}^*(\mathbf{e}) = n(a,b,c)\sqrt{a}(MST_{AB} - P_{2,AB}(\mathbf{e})) \xrightarrow{p} 0$, *as* $a\to\infty$, *b remains fixed.*

Proof Under $H_0(AB)$, MST_{AB} can also be expressed in terms of the e_{ijkm}'s and we have

$$n(a,b,c)\sqrt{ab}(MST_{AB} - P_{1,AB}(\mathbf{e})) = D_{1AB}(\mathbf{e}) + D_{2AB}(\mathbf{e}) + D_{3AB}(\mathbf{e}),$$
$$n(a,b,c)\sqrt{a}(MST_{AB} - P_{2,AB}(\mathbf{e})) = \frac{D_{2AB}(\mathbf{e})}{\sqrt{b}} + \frac{D_{3AB}(\mathbf{e})}{\sqrt{b}},$$

where

$$D_{1AB}(\mathbf{e}) = \frac{-cn(a,b,c)}{(b-1)\sqrt{ab}}\sum_{i=1}^{a}\sum_{j\neq j'}^{b}\tilde{e}_{ij..}\tilde{e}_{ij'..},\ D_{2AB}(\mathbf{e}) = \frac{-cn(a,b,c)}{(a-1)\sqrt{ab}}\sum_{i\neq i'}^{a}\sum_{j=1}^{b}\tilde{e}_{ij..}\tilde{e}_{i'j..} \quad (3.7.2)$$

$$D_{3AB}(\mathbf{e}) = \frac{cn(a,b,c)}{(a-1)(b-1)\sqrt{ab}}\sum_{i\neq i'}^{a}\sum_{j\neq j'}^{b}\tilde{e}_{ij..}\tilde{e}_{i'j'..} \quad (3.7.3)$$

The proofs of $D_{1AB}(\mathbf{e}) = o_p(1)$ and $D_{2AB}(\mathbf{e}) = o_p(1)$ are similar. So for both cases, it is enough to give the proof of $D_{2AB}(\mathbf{e}) = o_p(1)$ and $D_{3AB}(\mathbf{e}) = o_p(1)$.

$$\begin{aligned} E[D_{2AB}(\mathbf{e})]^2 &= \frac{2n^2(a,b,c)}{(a-1)^2c^2ab}\sum_{i\neq i'}^{a}\sum_{j=1}^{b}\left(\sum_{k=1}^{c}\frac{\sigma_{ijk}^2}{n_{ijk}}\right)\left(\sum_{k'=1}^{c}\frac{\sigma_{i'jk'}^2}{n_{i'jk'}}\right) \\ &\leq \frac{2n^2(a,b,c)}{(a-1)^2c^2ab}\sum_{j=1}^{b}\left(\sum_{i,k}\frac{\sigma_{ijk}^2}{n_{ijk}}\right)^2 \leq \frac{2n^2(a,b,c)}{(a-1)^2c^2ab}ac\sum_{j=1}^{b}\sum_{i,k}\left(\frac{\sigma_{ijk}^2}{n_{ijk}}\right)^2 = O(a^{-1}), \end{aligned}$$

where for the last inequality we used (3.6.6). Next, using again (3.6.6),

$$
\begin{aligned}
E[D_{3AB}(\mathbf{e})]^2 &= \frac{4n^2(a,b,c)}{(a-1)^2(b-1)^2c^2ab}\sum_{i\neq i'}^{a}\sum_{j\neq j'}^{b}\left(\sum_{k=1}^{c}\frac{\sigma_{ijk}^2}{n_{ijk}}\right)\left(\sum_{k'=1}^{c}\frac{\sigma_{i'jk'}^2}{n_{i'j'k'}}\right)\\
&\leq \frac{4n^2(a,b,c)}{(a-1)^2(b-1)^2c^2ab}\left(\sum_{i,j,k}\frac{\sigma_{ijk}^2}{n_{ijk}}\right)^2 \leq \frac{2n^2(a,b,c)}{(a-1)^2c}\sum_{i,j,k}\left(\frac{\sigma_{ijk}^2}{n_{ijk}}\right)^2 = O(a^{-1}).\square
\end{aligned}
$$

Lemma 3.7.6. *Under the settings and assumptions of either Theorem 3.2.7, or 3.2.8, or 3.2.9,*

$$n(a,b,c,d)(MSE-\sigma_{ABC}^2)\xrightarrow{p}0,\ \ \textit{where } n(a,b,c,d)=\min_{i,j,k,l}\{n_{ijkl}\}.$$

Proof The proof is similar to that of Lemma 3.7.2, and thus it is omitted.

Lemma 3.7.7. *Let* $P_{1,ABC}(\mathbf{e}) = d(abc)^{-1}\sum_{i,k}\left[\sum_{j=1}^{b}\tilde{e}_{ijk..}^2-\frac{1}{b-1}\sum_{j\neq j'}\tilde{e}_{ijk..}\tilde{e}_{ij'k..}\right]$, $P_{2,ABC}(\mathbf{e}) = P_{1,ABC}(\mathbf{e}) + \frac{bd}{(c-1)ac}\sum_{i=1}^{a}\sum_{k\neq k'}^{c}\tilde{e}_{i.k..}\tilde{e}_{i.k'..} - \frac{d}{(b-1)(c-1)ac}\sum_{i,j}\sum_{k\neq k'}^{c}\tilde{e}_{ijk..}\tilde{e}_{ijk'..}$, $P_{3,ABC}(\mathbf{e}) = \frac{d}{abc}\sum_{i,j,k}\tilde{e}_{ijk..}^2$, *where* $e_{ijklm} = X_{ijklm}-E(X_{ijklm})$, *and set* $n(a,b,c,d) = \min_{i,j,k,l}\{n_{ijkl}\}$. *Then under* $H_0(ABC)$,

1. *under the settings and assumptions of Theorem 3.2.7, as* $a,c\to\infty$, *b remains fixed,* $T_1^*(\mathbf{e}) = n(a,b,c,d)\sqrt{ac}(MST_{ABC}-P_{1,ABC}(\mathbf{e}))\xrightarrow{p}0$, *;*
2. *under the settings and assumptions of Theorem 3.2.8, as* $a\to\infty$, b,c *remain fixed,* $T_2^*(\mathbf{e}) = n(a,b,c,d)\sqrt{a}(MST_{ABC}-P_{2,ABC}(\mathbf{e}))\xrightarrow{p}0$ *;*
3. *under the settings and assumptions of Theorem 3.2.9, as* $a,b,c\to\infty$, $T_3^*(\mathbf{e}) = n(a,b,c,d)\sqrt{abc}(MST_{ABC}-P_{3,ABC}(\mathbf{e}))\xrightarrow{p}0$, .

Proof Under $H_0(ABC)$, MST_{ABC} can be expressed in terms of the e_{ijklm}'s and we have

$$
\begin{aligned}
T_1^*(\mathbf{e}) &= D_4(\mathbf{e})+D_5(\mathbf{e})+D_6(\mathbf{e})-\frac{D_1(\mathbf{e})}{b-1}-\frac{D_2(\mathbf{e})}{b-1}-\frac{D_3(\mathbf{e})}{b-1},\\
T_2^*(\mathbf{e}) &= \frac{D_5(\mathbf{e})}{\sqrt{c}}+\frac{D_6(\mathbf{e})}{\sqrt{c}}-\frac{D_2(\mathbf{e})+D_3(\mathbf{e})}{(b-1)\sqrt{c}},\\
T_3^*(\mathbf{e}) &= \sqrt{b}T_1^*(\mathbf{e})+D_7(\mathbf{e}),
\end{aligned}
$$

where

$$D_1(\mathbf{e}) = -\frac{bd\,n(a,b,c,d)}{(c-1)\sqrt{ac}} \sum_{i=1}^{a} \sum_{k\neq k'}^{c} \widetilde{e}_{i.k..}\widetilde{e}_{i.k'..}, \tag{3.7.4}$$

$$D_2(\mathbf{e}) = -\frac{bd\,n(a,b,c,d)}{(a-1)\sqrt{ac}} \sum_{i\neq i'}^{a} \sum_{k=1}^{c} \widetilde{e}_{i.k..}\widetilde{e}_{i'.k..} \tag{3.7.5}$$

$$D_3(\mathbf{e}) = \frac{bd\,n(a,b,c,d)}{(a-1)(c-1)\sqrt{ac}} \sum_{i\neq i'}^{a} \sum_{k\neq k'}^{c} \widetilde{e}_{i.k..}\widetilde{e}_{i'.k'..} \tag{3.7.6}$$

$$D_4(\mathbf{e}) = -\frac{d\,n(a,b,c,d)}{(b-1)(c-1)\sqrt{ac}} \sum_{i,j} \sum_{k\neq k'}^{c} \widetilde{e}_{ijk..}\widetilde{e}_{ijk'..} \tag{3.7.7}$$

$$D_5(\mathbf{e}) = \frac{d\,n(a,b,c,d)}{(a-1)(b-1)(c-1)\sqrt{ac}} \sum_{i\neq i'}^{a} \sum_{j=1}^{b} \sum_{k\neq k'}^{c} \widetilde{e}_{ijk..}\widetilde{e}_{i'jk'..} \tag{3.7.8}$$

$$D_6(\mathbf{e}) = -\frac{d\,n(a,b,c,d)}{(a-1)(b-1)\sqrt{ac}} \sum_{i\neq i'}^{a} \sum_{j,k} \widetilde{e}_{ijk..}\widetilde{e}_{i'jk..} \tag{3.7.9}$$

$$D_7(\mathbf{e}) = \frac{dn(a,b,c,d)}{(b-1)\sqrt{abc}} \sum_{i,k} \sum_{j\neq j'}^{b} \widetilde{e}_{ijk..}\widetilde{e}_{ij'k..} \tag{3.7.10}$$

$$\tag{3.7.11}$$

For all three cases, it suffices to show $D_t(\mathbf{e}) = o_p(1)$, for $t = 1,2,3,7$, and $\sqrt{b}D_s(\mathbf{e}) = o_p(1)$, for $s = 4,5,6$. The proof of $D_7(\mathbf{e}) = o_p(1)$, $\sqrt{b}D_4(\mathbf{e}) = o_p(1)$, and $\sqrt{b}D_6(\mathbf{e}) = o_p(1)$ are similar; the proof of $D_2(\mathbf{e}) = o_p(1)$ and $D_1(\mathbf{e}) = o_p(1)$ are similar; and the proof of $D_3(\mathbf{e}) = o_p(1)$ and $\sqrt{b}D_5(\mathbf{e}) = o_p(1)$ are similar. So it is enough to prove $D_7(\mathbf{e}) = o_p(1)$, $D_2(\mathbf{e}) = o_p(1)$ and

$D_3(\mathbf{e}) = o_p(1)$. Using inequality (3.6.6), we have

$$\begin{aligned}
E[D_7(\mathbf{e})]^2 &= \frac{2n^2(a,b,c,d)}{(b-1)^2d^2abc}\sum_{i,k}\sum_{j\neq j'}^{b}\left(\sum_{l=1}^{d}\frac{\sigma^2_{ijkl}}{n_{ijkl}}\right)\left(\sum_{l'=1}^{d}\frac{\sigma^2_{ij'kl'}}{n_{ij'kl'}}\right)\\
&\leq \frac{2n^2(a,b,c,d)}{(b-1)^2d^2abc}\sum_{i,k}\left(\sum_{j,l}\frac{\sigma^2_{ijkl}}{n_{ijkl}}\right)^2 \leq \frac{2n^2(a,b,c,d)}{(b-1)^2dac}\sum_{i,j,k,l}\frac{\sigma^4_{ijkl}}{n^2_{ijkl}} = O(b^{-1})\\
E[D_2(\mathbf{e})]^2 &= \frac{2n^2(a,b,c,d)}{(a-1)^2b^2d^2ac}\sum_{i\neq i'}^{a}\sum_{k=1}^{c}\left(\sum_{j,l}\frac{\sigma^2_{ijkl}}{n_{ijkl}}\right)\left(\sum_{j',l'}\frac{\sigma^2_{i'j'kl'}}{n_{i'j'kl'}}\right)\\
&\leq \frac{2n^2(a,b,c,d)}{(a-1)^2b^2d^2ac}\sum_{k=1}^{c}\left(\sum_{i,j,l}\frac{\sigma^2_{ijkl}}{n_{ijkl}}\right)^2 \leq \frac{2n^2(a,b,c,d)}{(a-1)^2bdc}\sum_{i,j,k,l}\frac{\sigma^4_{ijkl}}{n^2_{ijkl}} = O(a^{-1})\\
E[D_3(\mathbf{e})]^2 &= \frac{4n^2(a,b,c,d)}{(a-1)^2(c-1)^2b^2d^2ac}\sum_{i\neq i'}^{a}\sum_{k\neq k'}^{c}\left(\sum_{j,l}\frac{\sigma^2_{ijkl}}{n_{ijkl}}\right)\left(\sum_{j',l'}\frac{\sigma^2_{i'j'kl'}}{n_{i'j'k'l'}}\right)\\
&\leq \frac{4n^2(a,b,c,d)}{(a-1)^2(c-1)^2b^2d^2ac}\left(\sum_{i,j,k,l}\frac{\sigma^2_{ijkl}}{n_{ijkl}}\right)^2 \leq \frac{2n^2(a,b,c,d)}{(a-1)^2bd}\sum_{i,j,k,l}\frac{\sigma^4_{ijkl}}{n^2_{ijkl}} = O(a^{-1})
\end{aligned}$$

So $D_7(\mathbf{e}) = o_p(1)$, $D_2(\mathbf{e}) \xrightarrow{p} 0$ and $D_3(\mathbf{e}) \xrightarrow{p} 0$, and we complete the proof. □

3.7.2 Some Lemmas for rank tests

Lemma 3.7.8. *Under the settings and assumptions of Theorem 3.4.2,*

$$n(a,b)[MSE/N^2 - \sigma^2_A] \xrightarrow{p} 0 \text{ as } \max\{a,b\} \to \infty,$$

where $n(a,b) = \min_{i,j}\{n_{ij}\}$, *regardless of whether* n_{ij} *remain fixed or go to infinity.*

ProofBy Lemma 3.7.2, we only need to show that, in both cases, $MSE/N^2 - MSE_Y = o_p\left(n(a,b)^{-1}\right)$, where MSE_Y is similarly defined as MSE with R_{ijk} replaced by $Y_{ijk} = H(X_{ijk})$.

Write

$$
\begin{aligned}
& MSE/N^2 - MSE_Y \\
= & \frac{1}{ab}\sum_{i=1}^{a}\sum_{j=1}^{b}\frac{1}{n_{ij}(n_{ij}-1)}\sum_{k=1}^{n_{ij}}\left[(Z_{ijk}-\overline{Z}_{ij.})^2-(Y_{ijk}-\overline{Y}_{ij.})^2\right] \\
= & \frac{1}{ab}\sum_{i=1}^{a}\sum_{j=1}^{b}\frac{1}{n_{ij}(n_{ij}-1)}\sum_{k=1}^{n_{ij}}\left(Z_{ijk}-\overline{Z}_{ij.}-Y_{ijk}+\overline{Y}_{ij.}\right)^2 \\
+ & \frac{2}{ab}\sum_{i=1}^{a}\sum_{j=1}^{b}\frac{1}{n_{ij}(n_{ij}-1)}\sum_{k=1}^{n_{ij}}\left(Z_{ijk}-\overline{Z}_{ij.}-Y_{ijk}+\overline{Y}_{ij.}\right)(Y_{ijk}-\overline{Y}_{ij.})
\end{aligned}
$$

The first summation is $O_p(N^{-1}n(a,b)^{-1})$ since $sup_x|\widehat{H}(x)-H(x)|=O_p(N^{-1/2})$. Because $Y_{ijk}-\overline{Y}_{ij.}$ is uniformly bounded by 1, the second summation is bounded by

$$
\frac{2}{ab}\sum_{i=1}^{a}\sum_{j=1}^{b}\frac{1}{n_{ij}(n_{ij}-1)}\sum_{k=1}^{n_{ij}}\left|\left(Z_{ijk}-\overline{Z}_{ij.}-Y_{ijk}+\overline{Y}_{ij.}\right)\right| = O_p\left(\frac{1}{\sqrt{N}n(a,b)}\right).
$$

So

$$
MSE/N^2 - MSE_Y = O_p(N^{-1/2}n(a,b)^{-1}) = o_p\left(n(a,b)^{-1}\right),
$$

as $\max\{a,b\}\to\infty$, whether n_{ij} is fixed or not.

Lemma 3.7.9. *Let $P_A(\mathbf{Z}-E(\mathbf{Y}))$ be defined as $P_A(\mathbf{e})$ in Lemma (3.7.3) with e_{ijm} replaced by $Y_{ijm}-E(Y_{ijm})$, and $MST_A(\mathbf{Z})$ be defined as MST_A in (3.1.4) with $X^*_{ijm}=Z_{ijm}$. Then under $H_0(A)$, under the settings and assumptions of Theorem 3.4.2, as $a,b\to\infty$, $T^*_A(\mathbf{Z}-E(\mathbf{Y}))=n(a,b)\sqrt{a}(MST_A(\mathbf{Z})-P_A(\mathbf{Z}-E(\mathbf{Y})))\xrightarrow{p}0$, regardless of whether $n_{ij}\geq 2$ tend to ∞ or stay fixed.*

Proof By Lemma 3.7.3, under $H_0(A)$, $T^*_A(\mathbf{Y}-E(\mathbf{Y}))=o_p(1)$. Therefore, we only need to show $D_{ZY}=T^*_A(\mathbf{Z}-E(\mathbf{Y}))-T^*_A(\mathbf{Y}-E(\mathbf{Y}))=o_p(1)$. Note that under $H_0(A)$, $T^*_A(\mathbf{Y}-E(\mathbf{Y}))$ and $T^*_A(\mathbf{Z}-E(\mathbf{Y}))$ have the same form as the $T^*_A(\mathbf{e})$ in the proof of Lemma 3.7.3 with $\mathbf{e}$ replaced by $\mathbf{Y}-E(\mathbf{Y})$ and $\mathbf{Z}-E(\mathbf{Y})$, respectively. Using this form, write

$$\begin{aligned} T_A^*(\mathbf{Z}-E(\mathbf{Y}))-T_A^*(\mathbf{Y}-E(\mathbf{Y})) &= -\frac{bn(a,b)}{(a-1)\sqrt{a}}\sum_{i\neq i'}^{a}(\widetilde{Z}_{i..}-\widetilde{Y}_{i..})(\widetilde{Z}_{i'..}-\widetilde{Y}_{i'..}) \\ &\quad -\frac{2bn(a,b)}{(a-1)\sqrt{a}}\sum_{i\neq i'}^{a}(\widetilde{Z}_{i..}-\widetilde{Y}_{i..})(\widetilde{Y}_{i'..}-\overline{p}_{i'.}), \end{aligned}$$

where $p_{ij}=E(Y_{ijm})$. The proof that each term of this decomposition is $o_p(1)$ is similar to the proof of Lemma 3.7.13 and thus it is omitted.□

Lemma 3.7.10. *Under the settings and assumptions of Theorem 3.4.4 or 3.4.5,*

$$n(a,b,c)[MSE/N^2-\sigma^2_{AB}]\xrightarrow{p}0 \text{ as } \max\{a,b,c\}\to\infty, \text{ where } n(a,b,c)=\min_{i,j,k}\{n_{ijk}\}\ ,$$

regardless of whether n_{ijk} remain fixed or go to infinity.

Proof By Lemma 3.7.4, we only need to show that, in both cases, $MSE/N^2-MSE_Y=o_p\left(n(a,b,c)^{-1}\right)$, where MSE_Y is similarly defined as MSE with R_{ijkm} replaced by $Y_{ijkm}=H(X_{ijkm})$. The proof of this is similar to that of Lemma 3.7.8 and thus it is omitted.□

Lemma 3.7.11. *Let $P_{1,AB}(\mathbf{Z}-E(\mathbf{Y}))$ and $P_{2,AB}(\mathbf{Z}-E(\mathbf{Y}))$ be defined as $P_{1,AB}(\mathbf{e})$ and $P_{2,AB}(\mathbf{e})$ in Lemma (3.7.5) with e_{ijkm} replaced by $Y_{ijkm}-E(Y_{ijkm})$, and $MST_{AB}(\mathbf{Z})$ is defined as MST_{AB} in (3.1.9) with $X^*_{ijkm}=Z_{ijkm}$. Then under $H_0(AB)$,*

1. *under the settings and assumptions of Theorem 3.4.5, as $a,b\to\infty$,*
   $T_1^*(\mathbf{Z}-E(\mathbf{Y}))=n(a,b,c)\sqrt{ab}(MST_{AB}(\mathbf{Z})-P_{1,AB}(\mathbf{Z}-E(\mathbf{Y})))\xrightarrow{p}0,$

2. *under the settings and assumptions of Theorem 3.4.4, as $a\to\infty$, with b fixed,*
   $T_2^*(\mathbf{Z}-E(\mathbf{Y}))=n(a,b,c)\sqrt{a}(MST_{AB}(\mathbf{Z})-P_{2,AB}(\mathbf{Z}-E(\mathbf{Y})))\xrightarrow{p}0,$

regardless of whether c and $n_{ijk}\geq 2$ tend to ∞ or stay fixed.

Proof By Lemma 3.7.5, under $H_0(AB)$, $T_t^*(\mathbf{Y}-E(\mathbf{Y}))=o_p(1)$, for $t=1,2$. Therefore, we only need to show $D_{t,ZY}=T_t^*(\mathbf{Z}-E(\mathbf{Y}))-T_t^*(\mathbf{Y}-E(\mathbf{Y}))=o_p(1)$. In fact under $H_0(AB)$,

$T_t^*(\mathbf{Y}-E(\mathbf{Y}))$ have the same decompositions as $T_{tAB}(\mathbf{e})$ in the proof of Lemma 3.7.5 with $\mathbf{e}$ replaced by $\mathbf{Y}-E(\mathbf{Y})$, for $t=1,2$. Similar decompositions apply to $T_t^*(\mathbf{Z}-E(\mathbf{Y}))$ with $\mathbf{Z}-E(\mathbf{Y})$ as argument. Thus, to show $D_{t,ZY}=o_p(1)$, $t=1,2$, for both cases, it suffices to prove $D_{tAB}(\mathbf{Z}-E(\mathbf{Y}))-D_{tAB}(\mathbf{Y}-E(\mathbf{Y}))=o_p(1)$, for $t=1,2,3$, where the D_{tAB}, $t=1,2,3$, are defined in (3.7.2), (3.7.3). Thes proofs are similar to those of in the proof of Lemma 3.7.13 and thus they are omitted.□

Lemma 3.7.12. *Under the settings and assumptions of Theorem 3.4.8, 3.4.7, or 3.4.9,*

$$n(a,b,c,d)[MSE/N^2-\sigma^2_{ABC}] \xrightarrow{p} 0 \text{ as } \max\{a,b,c,d\}\to\infty,$$

where $n(a,b,c,d)=\min_{i,j,k,l}\{n_{ijkl}\}$, regardless of whether n_{ijkl} remain fixed or go to infinity.

Proof By Lemma 3.7.4, we only need to show that, in both cases, $MSE/N^2-MSE_Y=O_p\left(N^{-1/2}n(a,b,c,d)^{-1}\right)$, where MSE_Y is similarly defined as MSE with R_{ijklm} replaced by $Y_{ijklm}=H(X_{ijklm})$. The proof of this is similar to that of Lemma 3.7.8 and thus it is omitted.□

Lemma 3.7.13. *Let $P_{1,ABC}(\mathbf{Z}-E(\mathbf{Y}))$, $P_{2,ABC}(\mathbf{Z}-E(\mathbf{Y}))$, and $P_{3,ABC}(\mathbf{Z}-E(\mathbf{Y}))$ be similarly defined as $P_{1,ABC}(\mathbf{e})$, $P_{2,ABC}(\mathbf{e})$, and $P_{3,ABC}(\mathbf{e})$ in Lemma (3.7.7) with e_{ijklm} replaced by $Y_{ijklm}-E(Y_{ijklm})$. Also $MST_{ABC}(\mathbf{Z})$ is similarly defined as MST_{ABC} in (3.1.13) with $X^*_{ijklm}=Z_{ijklm}$. Then under $H_0(ABC)$,*

1. *under the settings and assumptions of Theorem 3.4.7, as $a,c\to\infty$,*
   $T_1^*(\mathbf{Z}-E(\mathbf{Y}))=n(a,b,c,d)\sqrt{ac}(MST_{ABC}(\mathbf{Z})-P_{1,ABC}(\mathbf{Z}-E(\mathbf{Y})))\xrightarrow{p}0$,
2. *under the settings and assumptions of Theorem 3.4.8, as $a\to\infty$,*
   $T_2^*(\mathbf{Z}-E(\mathbf{Y}))=n(a,b,c,d)\sqrt{a}(MST_{ABC}(\mathbf{Z})-P_{2,ABC}(\mathbf{Z}-E(\mathbf{Y})))\xrightarrow{p}0$,
3. *under the settings and assumptions of Theorem 3.4.9, as $a,b,c\to\infty$,*
   $T_3^*(\mathbf{Z}-E(\mathbf{Y}))=n(a,b,c,d)\sqrt{abc}(MST_{ABC}(\mathbf{Z})-P_{3,ABC}(\mathbf{Z}-E(\mathbf{Y})))\xrightarrow{p}0$,

regardless of whether d and n_{ijkl} tend to ∞ or stay fixed as long as $n_{ijkl}\geq 2$.

Proof By Lemma 3.7.7, $T_t^*(\mathbf{Y}-E(\mathbf{Y}))=o_p(1)$, for $t=1,2,3$, under $H_0(ABC)$, where $T_t^*(\mathbf{Y}-E(\mathbf{Y}))$ are similarly defined as $T_t^*(\mathbf{Z}-E(\mathbf{Y}))$. Therefore, we only need to show the difference is $o_p(1)$ under $H_0(ABC)$, i.e., $D_{t,ZY}=T_t^*(\mathbf{Z}-E(\mathbf{Y}))-T_t^*(\mathbf{Y}-E(\mathbf{Y}))=o_p(1)$. In fact under $H_0(ABC)$, $T_t^*(\mathbf{Y}-E(\mathbf{Y}))$ have the same decompositions as $T_t^*(\mathbf{e})$ in the proof of Lemma 3.7.7 with $\mathbf{e}$ replaced by $\mathbf{Y}-E(\mathbf{Y})$, for $t=1,2,3$. Similar decompositions apply to $T_t^*(\mathbf{Z}-E(\mathbf{Y}))$ with $\mathbf{Z}-E(\mathbf{Y})$ as argument. To show $D_{t,ZY}=o_p(1)$, $t=1,2,3$, for all three cases, it suffices to prove $D_t(\mathbf{Z}-E(\mathbf{Y}))-D_t(\mathbf{Y}-E(\mathbf{Y}))=o_p(1)$, for $t=1,2,3,7$, and $\sqrt{b}[D_s(\mathbf{Z}-E(\mathbf{Y}))-D_s(\mathbf{Y}-E(\mathbf{Y}))]=o_p(1)$, for $s=4,5,6$. The proofs are similar and we only give the last one.

$$\sqrt{b}[D_6(\mathbf{Z}-E(\mathbf{Y}))-D_6(\mathbf{Y}-E(\mathbf{Y}))]=D_{31}^*+D_{32}^*,$$

where

$$\begin{aligned}
D_{31}^* &= -\frac{\sqrt{b}d\,n(a,b,c,d)}{(a-1)(b-1)\sqrt{ac}}\sum_{i\neq i'}^{a}\sum_{j,k}(\widetilde{Z}_{ijk..}-\widetilde{Y}_{ijk..})(\widetilde{Z}_{i'jk..}-\widetilde{Y}_{i'jk..})\\
D_{32}^* &= -\frac{2\sqrt{b}d\,n(a,b,c,d)}{(a-1)(b-1)\sqrt{ac}}\sum_{i\neq i'}^{a}\sum_{j,k}(\widetilde{Z}_{ijk..}-\widetilde{Y}_{ijk..})(\widetilde{Y}_{i'jk..}-\overline{p}_{i'jk.}).
\end{aligned}$$

Because $\sup_x(\widehat{H}(x)-H(x))=O_p(N^{-1/2})$, we have $D_{31}^*=O_p\left(d\sqrt{ab}n(a,b,c,d)/N\right)=o_p(1)$.

$$D_{32}^* = -\frac{2\sqrt{b}n(a,b,c,d)}{(a-1)(b-1)d\sqrt{ac}}\sum_{i\neq i'}^{a}\sum_{j,k}\sum_{l,m}\sum_{l',m'}\frac{(Z_{ijklm}-Y_{ijklm})(Y_{i'jkl'm'}-p_{i'jkl'})}{n_{ijkl}n_{i'jkl'}}.$$

Write $Z_{ijklm}-Y_{ijklm}=N^{-1}\sum_{i_4,j_4,k_4,l_4,m_4}(c(X_{i_4,j_4,k_4,l_4,m_4},X_{ijklm})-F_{i_4,j_4,k_4,l_4,m_4}(X_{ijklm}))$, then

$$\begin{aligned}
&E(D_{32}^*)^2\\
=\;& \frac{4bn^2(a,b,c,d)}{(a-1)^2(b-1)^2d^2ac}\sum_{i\neq i'}^{a}\sum_{j,k,l,m}\sum_{l',m'}\sum_{i_1\neq i_1'}^{a}\sum_{j_1,k_1,l_1,m_1}\sum_{l_1',m_1'}\\
&\frac{1}{n_{ijkl}n_{i'jkl'}n_{i_1j_1k_1l_1}n_{i_1'j_1k_1l_1'}}\frac{1}{N^2}\sum_{i_4,j_4,k_4,l_4,m_4}\sum_{i_5,j_5,k_5,l_5,m_5}E\left[\left(Y_{i'jkl'm'}-p_{i'jkl'm'}\right)\right.\\
&\left(Y_{i_1'j_1k_1l_1'm_1'}-p_{i_1'j_1k_1l_1'm_1'}\right)(c(X_{i_4,j_4,k_4,l_4,m_4},X_{ijklm})-F_{i_4,j_4,k_4,l_4,m_4}(X_{ijklm}))\\
&\left.(c(X_{i_5,j_5,k_5,l_5,m_5},X_{i_1j_1k_1l_1m_1})-F_{i_5,j_5,k_5,l_5,m_5}(X_{i_1j_1k_1l_1m_1}))\right].
\end{aligned}$$

By independence the expecation under the summation is zero if the number of different elements in $\{i,i',i_1,i'_1,i_4,i_5\}$ is five or six, or the number of different elements in $\{l,l',l_1,l'_1,l_4,l_5\}$ is five or six, or the number of different elements in $\{m,m',m_1,m'_1,m_4,m_5\}$ is five or six, or the number of different pairs in $\{(j,k),(j_1,k_1),(j_4,k_4),(j_5,k_5)\}$ is three or four. Hence, by the fact that Y_{ijklm} and $c(X_1,X_2)$ are uniformly bounded, $E(D^*_{32})^2 = O(abd^2n^2(a,b,c,d)/N^2) = o(1)$ Therefore, $D^*_{32} = o_p(1)$.

Chapter 4

Inference from Heteroscedastic Functional Data, Part I: Identical Sampling Points

4.1 Introduction

Studies involving repeated measurements within a subject or stratum arise frequently in scientific investigations, and have attracted considerable attention in the statistical literature. Examples are growth curves, reaction curves in chemical experiments, evaluation of disease progression, biomarkers measured over time, seismic recordings etc. Such data are called longitudinal, curve or functional data, though the latter two terms seem to be more common when the number of observations per subject is large. Our discussion below adheres to this distinction.

Approaches to the analysis of longitudinal data include linear and nonlinear mixed-effects models with parametric (cf. Pinhero and Bates, 1995), semiparametric (Davidian and Gallant, 1993), nonparametric (Mallet, 1986), and Bayesian modeling (cf. Smith and Roberts, 1993), generalized linear models (Liang and Zeger, 1986; Zeger and Liang, 1986) to deal with discrete ordinal data, and the fully nonparametric marginal model for all types of ordinal data (Akritas and Brunner, 1997). For text-book presentations of such material, see Vonesh and Chinchilli (1997), Brunner, Domhof and Langer (2002), Davidian and Giltinan (1995), Diggle, Liang and Zeger (1994) and Hand and Crowder (1996). These approaches are mainly suitable when the number of within stratum measurements is relatively small. Time series models

(Brillinger, 1973, 1980), functional ANOVA models (Ramsey and Silverman, 1997), smoothing spline models (Brumback and Rice, 1998; Wang 1998), varying coefficient models with smoothing splines and other nonparametric smoothing techniques (Hoover, Rice, Wu and Yang, 1998, Fan and Zhang, 2000, Chiang, Rice and Wu, 2001), can also be used for functional or curve data where the number of within stratum measurements is large. See also Fan and Lin (1998) who considered a comparison of curves problem, analogous to some lack-of-fit methods with independent observations, and used ideas from adaptive Neyman test and wavelet thresholding to improve the power of their test procedures, and Lin and Ying (2001) who considered a counting processes approach. Asymptotic theory is presented in some of the aforementioned functional data approaches, but always under the assumption of large group sizes, and typically under additional modeling assumptions. Only Fan and Lin (1998), Fan and Zhang (2000) and, in a restricted sense, Chiang, Rice and Wu (2001) attempt asymptotics by allowing the number of time points to also tend to infinity. Hoover, Rice, Wu and Yang (1998) point out that important inferece issues are not always addressed.

Here we will consider curve or functional data situations with identical sampling points where there is interest in describing and evaluating the effect of several baseline, or time-independent, factors (sex, age, dose level etc) and the interaction of such factors with the time factor. For example, in a dose-response experiment where the response consists of correlated observations over time it might be of interest to evaluate the dose effect and the interaction between dose and time. The marginal model we consider is purely nonparametric. Thus the responses can be measured on continuous or discrete ordinal scale, there is no restriction in the allowed pattern of heteroscedasticity, and no model describing the change in the response distribution from one covariate value to another, or from one time point to another. Moreover, the asymptotic theory does not require large group sizes, and simulations studies confirm this. Thus, the present methods fill a gap in the available data analytic tools for functional data. The present inference procedures can be used in conjunction with existing graphical methods such as principal component analysis and dynamic time warping; see Ramsey and Silverman (1997), Ramsey and Li (1998), Ke and Wang (2001). Some new graphical methods for describing the effects are also discussed here.

Because statistics for the usual normal-based procedures (univariate or multivariate) have no closed form expression in unbalanced designs, and also because they are not designed to accommodate heteroscedasticity, we do not undertake a study of the asymptotic properties of these statistics. See Arnold (1981) for such a study under the condition that the group sample sizes tend to infinity, while the number of groups and repeated measures per subject remain fixed. We propose new test statistics falling in two distinct categories. One to test for effects described by a small number of parameters (e.g. group effects), and another to test for effects described by a large number of parameters (e.g. main time effects, and interactions between time and groups). Test statistics in the first category are quadratic forms having asymptotically a χ^2 distribution, while those in the second category are in the form of a difference of two quadratic forms and have asymptotically a normal distribution. The asymptotic theory for both categories of statistics requires novel techniques which rely on the number of time points tending to infinity (the number of subjects can be small or large) under quite general unspecified covariance structure. In particular, all derivations are done under an α-mixing assumption which includes most common time series models.

Because procedures based on the original observations require strong moment assumptions and can perform poorly away from the normal distribution, we also develop a competing set of (mid-)rank procedures. These are constructed by simply replacing the original observations by their (mid-)ranks in the corresponding test statistics. Of course the hypotheses tested by the rank statistics are invariant under monotone transformations and thus are different from the usual parametric hypotheses based on the decomposition of group means. Such invariant hypotheses are described in §4.2.1.

The rest of the chapter is organized as the following. In Section 4.2, we present the model and hypotheses as well as the results based on the original observations and their overall (mid-)ranks. Some simulation results and a real data analysis are presented in Section ref-sec.num.res, while the proofs are given in Section 4.4.

4.2 Main Results

4.2.1 The model and hypotheses

Consider subjects nested within a total of a factor levels such that each subject is measured at b time points $t_1, \ldots, t_b$. Thus, the k-th subject in factor level i generates a time series

$$\mathbf{X}_{ik} = (X_{i1k}, \ldots, X_{ibk})', \; i = 1, \ldots, a, \; k = 1, \ldots, n_i. \tag{4.2.1}$$

We will use a marginal model for making inferences about the effects of the factor levels and time points, in cases where b is large, the n_i may be small, and a is fixed. The a groups can be factor level combinations of several factors, but for the time being we enumerate all with the single index i. We set $n = \sum_{i=1}^{a} n_i$ for the total number of subjects.

The assumptions we make is that time series corresponding to different subjects are independent, and that each time series satisfies an α-mixing condition. Thus, we assume that for some sequence $\alpha_m \to 0$,

$$|P(A \cap B) - P(A)P(B)| \leq \alpha_m,$$

holds for all $A \in \sigma(X_{i1k}, \ldots, X_{i\ell k})$, $B \in \sigma(X_{i,\ell+m,k}, X_{i,\ell+m+1,k}, \ldots)$, and all i, k. The α-mixing assumption is a weak requirement for the derivation of our asymptotic results, as many common time series models have been shown to satisfy this condition. For example, Masry and Tjostheim (1995, 1997) show that under some mild conditions, both ARCH processes and additive

AR processes with exogenous variables are α-mixing. Moreover, we assume that

$$X_{ijk} \sim F_{ij}, \text{ for all } k = 1, \ldots, n_i, \tag{4.2.2}$$

$$(X_{ijk}, X_{i'j'k'}) \sim F_{ij} F_{i'j'}, \text{ for } k \neq k' \text{ or } i \neq i' , \tag{4.2.3}$$

$$(X_{ijk}, X_{ij'k}) \sim F_{ijj'}, \tag{4.2.4}$$

$$\text{Conditionally on the subject, } X_{i1k}, \ldots, X_{ibk} \text{ are independent.} \tag{4.2.5}$$

Assumption (4.2.3) is a restatement of the assumption that time series corresponding to different subjects are independent, (4.2.4) states that the joint distribution of observations in time points t_j, $t_{j'}$ from factor level i have some joint distribution that depends only on i, t_j, $t_{j'}$, and finally (4.2.5) states the dependence of the observations within each time series $\mathbf{X}_{ik}$ is only due to subject k in group i. Note that we do not require continuous observations, place no restrictions in the allowed patterns of heteroscedasticity or the dependency structure (except for the α-mixing assumption) and do not specify how the distribution of the observations changes from one factor level to another or from one time point to another. Thus the above is a purely nonparametric set-up.

Using (4.2.3), (4.2.4) and (4.2.5) we can always (so also for discrete data) write

$$X_{ijk} = \mu_{ij} + d_{jk(i)} + \varepsilon_{ijk}, \; i = 1, \cdots, a, \quad j = 1, \cdots, b, \quad k = 1, \cdots, n_i, \tag{4.2.6}$$

where $\mu_{ij} = E(X_{ijk})$, and

$$d_{jk(i)} = E(X_{ijk}|k(i)) - \mu_{ij}, \quad \varepsilon_{ijk} = X_{ijk} - E(X_{ijk}|k(i)).$$

Note that

$$E(\varepsilon_{ijk}) = 0, \quad E(d_{jk(i)}) = 0, \quad E(\varepsilon_{ijk}\varepsilon_{ij'k}) = 0, \quad E(d_{jk(i)}\varepsilon_{ijk}) = 0,$$

but contrary to the apparent similarity of (4.2.6) to the usual linear mixed effect models, the present is a general nonparametric model which, as mentioned above, holds also for discrete ordinal data. The crucial difference lies in the fact that the usual assumption of identical distribution for the error term (even up to a scale parameter for heteroscedastic linear models) is not made here. A decomposition of the means μ_{ij} yields the usual parametric main and interaction effects:

$$\mu_{ij} = \mu + \alpha_i + \beta_j + \gamma_{ij}, \text{ where } \sum_{i=1}^{a} \alpha_i = \sum_{j=1}^{b} \beta_j = \sum_{i=1}^{a} \gamma_{ij} = \sum_{j=1}^{b} \gamma_{ij} = 0.$$

The usual parametric hypotheses specify that the corresponding parametric effects are zero and will be denoted by

$$H_0(\beta): \text{ all } \beta_j = 0, \ H_0(\gamma): \text{ all } \gamma_{ij} = 0, \ H_0(\phi): \text{ all } \phi_{ij} = \alpha_i + \gamma_{ij} = 0, \text{ and}$$

$$H_0(\alpha): \text{ all } \alpha_i = 0, \text{ or, more generally } \widetilde{H}_0(\alpha): \mathbf{C}\alpha = \mathbf{0},$$

where $\mathbf{C}$ is a contrast matrix and $\alpha = (\alpha_1, \ldots, \alpha_a)'$. We remark that the number of parameters that are involved in the first three of the above hypotheses is large (tends to infinity), while $H_0(\alpha)$ and $\widetilde{H}_0(\alpha)$ involve a small number of parameters. We note that $\widetilde{H}_0(\alpha)$ is useful in cases where we are interested in more specific hypotheses concerning the α_i effects (such as those in Section 4.3, Example 1), or in more specialized hypotheses in case the a groups are factor-level combinations of several factors.

As mentioned in Section 1, however, rank procedures test certain types of nonparametric hypotheses which are invariant under monotone transformations. The first type are those introduced in Akritas and Arnold (1994) by decomposing the distributions F_{ij}, defined in (4.2.2), in a way similar to the decomposition of the μ_{ij}:

$$F_{ij}(x) = M(x) + A_i(x) + B_j(x) + C_{ij}(x), \tag{4.2.7}$$

where $\sum_{i=1}^{a} A_i(x) = \sum_{j=1}^{b} B_j(x) = \sum_{i=1}^{a} C_{ij}(x) = \sum_{j=1}^{b} C_{ij}(x) = 0, \ \forall x$. The functions A_i, B_j, C_{ij} in (4.2.7) are the fully nonparametric effects, and the fully nonparametric hypotheses specify that the corresponding effects are zero. Specifically we denote

$$H_0(B): \text{ all } B_j = 0, \ H_0(C): \text{ all } C_{ij} = 0, \ H_0(D): \text{ all } D_{ij} = A_i + C_{ij} = 0, \text{ and}$$

$$H_0(A): \text{ all } A_i = 0, \text{ or, more generally } \widetilde{H}_0(A): \mathbf{C}\overline{\mathbf{F}} = \mathbf{0}, \text{ where } \overline{\mathbf{F}} = (\overline{F}_{1.}, \cdots, \overline{F}_{a.})'$$

For the meaning and interpretation of the fully nonparametric effects and hypotheses see Akritas and Arnold (1994) and Akritas et. al (1997). The second type are those defined in terms of what Brunner, Domhof and Langer (2002, p.37) call relative treatment effects, and which are closely related to the d statistic defined in Cliff (1993) for use in behavioral psychology. Specifically, write $H(X_{ijk}) = n^{-1}b^{-1}\sum_{i_1=1}^{a}\sum_{j_1=1}^{b} n_{i_1}F_{i_1 j_1}(X_{ijk})$, where $n = \sum_{i=1}^{a} n_i$, and

$$\widetilde{\mu_{ij}} = E(H(X_{ijk})) = \widetilde{\mu} + \widetilde{\alpha}_i + \widetilde{\beta}_j + \widetilde{\gamma}_{ij}, \text{ where } \sum_{i=1}^{a}\widetilde{\alpha}_i = \sum_{j=1}^{b}\widetilde{\beta}_j = \sum_{i=1}^{a}\widetilde{\gamma}_{ij} = \sum_{j=1}^{b}\widetilde{\gamma}_{ij} = 0.$$

The $\widetilde{\alpha}_i$, $\widetilde{\beta}_j$, and $\widetilde{\gamma}_{ij}$ are the relative treatment effects, and the relative treatment null hypotheses specify that the corresponding effects are zero. Specifically we denote

$$H_0(\widetilde{\beta}): \text{ all } \widetilde{\beta}_j = 0,\ H_0(\widetilde{\gamma}): \text{ all } \widetilde{\gamma}_{ij} = 0,\ H_0(\widetilde{\phi}): \text{ all } \widetilde{\phi}_{ij} = \widetilde{\alpha}_i + \widetilde{\gamma}_{ij} = 0.$$

The fully nonparametric hypotheses are stronger than both the usual parametric hypotheses and the relative treatment effects hypotheses (in the sense that they imply but are not implied by them). Thus, all asymptotic results stated in this section hold also under the fully nonparametric hypotheses. We remark that we did not define $H_0(\widetilde{\alpha})$ because none of the proposed test procedures tests this hypothesis.

Section 4.2.2 develops test procedures for the above hypotheses based on the original observations, while in Section 4.2.3 we employ (mid-)rank statistics to test corresponding nonparametric hypotheses. For any transformation Y_{ijk} of X_{ijk} the following notations are used throughout the paper:

$$n = \sum_{i=1}^{a} n_i, \quad N = nb, \quad \widetilde{Y}_{i..} = \overline{Y}_{i..} = \frac{1}{bn_i}\sum_{j=1}^{b}\sum_{k=1}^{n_i} Y_{ijk}, \quad \overline{Y}_{...} = \frac{1}{N}\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n_i} Y_{ijk},$$

$$\widetilde{Y}_{...} = \frac{1}{ab}\sum_{i=1}^{a}\sum_{j=1}^{b}\frac{1}{n_i}\sum_{k=1}^{n_i} Y_{ijk}, \quad \widetilde{Y}_{.j.} = \frac{1}{a}\sum_{i=1}^{a}\frac{1}{n_i}\sum_{k=1}^{n_i} Y_{ijk}, \quad \overline{Y}_{.j.} = \frac{1}{n}\sum_{i=1}^{a}\sum_{k=1}^{n_i} Y_{ijk}.$$

4.2.2 Tests based on the original observations

The classical normal-based test statistics are not appropriate in the present heteroscedastic case, because the expected value of the mean squares for, e.g. main time effects, is not equal to the mean square for error under the null hypothesis. For testing $H_0(\alpha)$ or $\widetilde{H}_0(\alpha)$ we will introduce a Wald-type statistic, while for testing $H_0(\beta)$, $H_0(\gamma)$ and $H_0(\phi)$ we introduce suitable variations of mean squares so that the corresponding expected values match.

The Wald-type test statistic for $H_0(\alpha)$ or $\widetilde{H}_0(\alpha)$ is given later in connection to Theorem 4.2.3. The aforementioned variations of mean squares are

$$
\begin{aligned}
MS\beta &= \frac{1}{b-1}\sum_{i=1}^{a}\sum_{j=1}^{b}\left(\widetilde{X}_{.j.}-\widetilde{X}_{...}\right)^2 \\
MS\gamma &= \frac{1}{(a-1)(b-1)}\sum_{i=1}^{a}\sum_{j=1}^{b}(\overline{X}_{ij.}-\widetilde{X}_{i..}-\widetilde{X}_{.j.}+\widetilde{X}_{...})^2 \\
MSE &= \frac{1}{a(b-1)}\sum_{i=1}^{a}\sum_{j=1}^{b}\frac{1}{n_i(n_i-1)}\sum_{k=1}^{n_i}\left(X_{ijk}-\overline{X}_{ij.}-\overline{X}_{i.k}+\widetilde{X}_{i..}\right)^2 \\
MS\phi &= \frac{1}{(a-1)b}\sum_{i=1}^{a}\sum_{j=1}^{b}\left(\overline{X}_{ij.}-\widetilde{X}_{.j.}\right)^2 \text{ and} \\
MSE_\phi &= \frac{1}{ab}\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n_i}\frac{(X_{ijk}-\overline{X}_{ij.})^2}{n_i(n_i-1)}.
\end{aligned}
$$

It is easily seen that with $\sigma_{ijj'}=\mathrm{Cov}(X_{ijk},X_{ij'k})$ and $\sigma_{ijj}=\sigma_{ij}^2=\mathrm{Var}(X_{ijk})$

$$
\begin{aligned}
E(MSE) &= \frac{1}{a(b-1)}\sum_{i,j}\frac{\sigma_{ij}^2}{n_i}-\frac{1}{ab(b-1)}\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{j'=1}^{b}\frac{\sigma_{ijj'}}{n_i}, \\
E(MS\beta) &= E(MSE), \text{ under } H_0(\beta), \\
E(MS\gamma) &= E(MSE), \text{ under } H_0(\gamma) \\
E(MSE_\phi) &= \frac{1}{ab}\sum_{i,j}\frac{\sigma_{ijj}}{n_i}, \text{ and} \\
E(MS\phi) &= E(MSE_\phi) \text{ under } H_0(\phi).
\end{aligned}
$$

Thus it is reasonable to use

$$
F_\beta=\frac{MS\beta}{MSE},\quad F_\gamma=\frac{MS\gamma}{MSE},\quad \text{and}\quad F_\phi=\frac{MS\phi}{MSE_\phi}
$$

as test statistics for $H_0(\beta)$, $H_0(\gamma)$ and $H_0(\phi)$ respectively.

The asymptotic theory of these test statistics is given next.

Theorem 4.2.1. *Assume that for each group i and subject k, X_{ijk}, $j = 1, 2, \cdots$, is α-mixing with with $\alpha_m = O\left(m^{-5}\right)$. In addition, assume that $\limsup_j E[(X_{ijk} - \mu_{ij})^{32}]) < \infty$, where $\mu_{ij} = E(X_{ijk})$. Then as $b \to \infty$ while a remains fixed, the limits of*

$$\zeta_1 = \frac{2}{a^2 b} \sum_{j=1}^{b} \sum_{j'=1}^{b} \sum_{i=1}^{a} \frac{\sigma^2_{ijj'}}{n_i(n_i - 1)}, \qquad \zeta_2 = \frac{2}{a^2 b} \sum_{j=1}^{b} \sum_{j'=1}^{b} \sum_{i \neq i'}^{a} \frac{\sigma_{ijj'}\sigma_{i'jj'}}{n_i n_{i'}} \tag{4.2.8}$$

exist regardless of whether the n_i are fixed or go to ∞ as $b \to \infty$. Moreover, with $\sigma^2 = \lim_{b\to\infty} E(MSE) = \lim_{b\to\infty} E(MSE_\phi)$, and $\sigma^2_* = \lim_{b\to\infty} E(n(a)MSE) = \lim_{b\to\infty} E(n(a)MSE_\phi)$,

(1) for $n_i \geq 2$ fixed,

under $H_0(\beta)$, $\sqrt{b}(F_\beta - 1) \xrightarrow{d} N\left(0, \frac{\tau^2_\beta}{\sigma^4}\right)$, *where* $\tau^2_\beta = \lim_{b\to\infty}(\zeta_1 + \zeta_2)$,

under $H_0(\gamma)$, $\sqrt{b}(F_\gamma - 1) \xrightarrow{d} N\left(0, \frac{\tau^2_\gamma}{\sigma^4}\right)$, *where* $\tau^2_\gamma = \lim_{b\to\infty}\left(\zeta_1 + \frac{\zeta_2}{(a-1)^2}\right)$,

under $H_0(\phi)$, $\sqrt{b}(F_\phi - 1) \xrightarrow{d} N\left(0, \frac{\tau^2_\gamma}{\sigma^4}\right)$;

(2) if $n_i \to \infty$ as $b \to \infty$, under the additional assumption $\max_i\{n_i\}/\min_i\{n_i\} = O(1)$, we have

under $H_0(\beta)$, $\sqrt{b}(F_\beta - 1) \xrightarrow{d} N\left(0, \frac{\tau^2_{\beta*}}{\sigma^4_*}\right)$, *where* $\tau^2_{\beta*} = \lim_{b\to\infty} n^2(a)(\zeta_1 + \zeta_2)$;

under $H_0(\gamma)$, $\sqrt{b}(F_\gamma - 1) \xrightarrow{d} N\left(0, \frac{\tau^2_{\gamma*}}{\sigma^4_*}\right)$, *where* $\tau^2_{\gamma*} = \lim_{b\to\infty} n^2(a)\left(\zeta_1 + \frac{\zeta_2}{(a-1)^2}\right)$,

under $H_0(\phi)$, $\sqrt{b}(F_\phi - 1) \xrightarrow{d} N\left(0, \frac{\tau^2_{\gamma*}}{\sigma^4_*}\right)$.

Remark 2.1 Under the extra (often unrealistic) assumption of stationarity, the assumption $\limsup_j E((X_{ijk}-\mu_{ij})^{32}) < \infty$ becomes $\limsup_j E((X_{ijk}-\mu_{ij})^{16}) < \infty$.

The next result gives consistent estimators for the limits of ζ_1 and ζ_2, which enter the expressions for the asymptotic variance.

Proposition 4.2.2. *For each $j = 1,\ldots,b$, let $C_u(j,h) = [\min\{b, j+b^h\}]$, $C_l(j,h) = [\max\{0, j-b^h\}]$, for some $0 < h < 1$, where $[x]$ denotes the largest integer less than or equal to x. Set*

$$\widehat{\zeta}_1 = \frac{2}{a^2b}\sum_{j=1}^{b}\sum_{j'=C_l(j,h)}^{C_u(j,h)}\sum_{i=1}^{a}\frac{\widehat{\sigma^2_{ijj'}}}{n_i(n_i-1)}, \qquad \widehat{\zeta}_2 = \frac{2}{a^2b}\sum_{j=1}^{b}\sum_{j'=C_l(j,h)}^{C_u(j,h)}\sum_{i\neq i'}^{a}\frac{\widehat{\sigma}_{ijj'}\widehat{\sigma}_{i'jj'}}{n_in_{i'}}, \text{ where}$$

$$\widehat{\sigma^2_{ijj'}} = \sum_{k_1\neq k_2\neq k_3\neq k_4}^{n_i}\frac{(X_{ijk_1}-X_{ijk_2})(X_{ij'k_1}-X_{ij'k_2})(X_{ijk_3}-X_{ijk_4})(X_{ij'k_3}-X_{ij'k_4})}{4n_i(n_i-1)(n_i-2)q(n_i-3)}, \text{ and}$$

$$\widehat{\sigma}_{ijj'} = \sum_{k=1}^{n_i}\frac{(X_{ijk}-\overline{X}_{ij.})(X_{ij'k}-\overline{X}_{ij'.})}{n_i-1}.$$

Then as $b\to\infty$, $\widehat{\zeta}_1 - \zeta_1 \xrightarrow{p} 0$ and $\widehat{\zeta}_2 - \zeta_2 \xrightarrow{p} 0$.

For testing $\widetilde{H}_0(\alpha) : \mathbf{C}_a\alpha = 0$, where $\mathbf{C}_a$ is a full rank contrast matrix, we need techniques different from those used above. These are detailed in the next result.

Theorem 4.2.3. *Assume $X_{i1k}, X_{i2k}, \cdots,$ is α-mixing with with $\alpha_m = O\left(m^{-5}\right)$ for all i,k, and $\limsup_j E[(X_{ijk}-E(X_{ijk}))^{16}] < \infty$. Let $\mathbf{W} = (\overline{X}_{1..},\cdots,\overline{X}_{a..})'$, and*

$$\widehat{\eta}_i = \frac{n}{bn_i(n_i-1)}\sum_{j=1}^{b}\sum_{j'=C_l(j,h)}^{C_u(j,h)}\sum_{k=1}^{n_i}\left(X_{ijk}-\overline{X}_{ij.}\right)\left(X_{ij'k}-\overline{X}_{ij'.}\right), \tag{4.2.9}$$

where $C_l(j,h)$ and $C_u(j,h)$ are defined in Theorem 4.2.1. Then under $\widetilde{H}_0(\alpha)$,

$$N\mathbf{W}'\mathbf{C}_a'\left[\mathbf{C}_a\, diag(\hat{\eta}_1,\cdots,\hat{\eta}_a)\mathbf{C}_a'\right]^{-1}\mathbf{C}_a\mathbf{W} \xrightarrow{d} \chi^2_r,$$

where r is the rank of $\mathbf{C}_a$, as $b\to\infty$ regardless of whether the n_i remain fixed or tend to ∞ with b provided $\max n_i/\min n_i = O(1)$ and $h < 0.5$.

4.2.3 Rank Tests

For this section we set $H(x) = N^{-1}\sum_{i=1}^{a}\sum_{j=1}^{b} n_i F_{ij}(x)$ for the average distribution function, and

$$\widehat{H}(x) = N^{-1}\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n_i} c(X_{ijk}, x),$$

for its empirical version. Thus $R_{ijk} = 1/2 + N\widehat{H}(X_{ijk})$ is the (mid-)rank of X_{ijk} among all observations. We are interested in constructing (mid-)rank tests for the hypotheses stated in §4.2.1.

Let $MS\beta_R$, $MS\gamma_R$, $MS\phi_R$, MSE_R and $MSE_{\phi,R}$ be the variations of the mean squares defined in §4.2.2 but with X_{ijk} replaced by R_{ijk}, and define

$$F_{R,\beta} = \frac{MS\beta_R}{MSE_R},\ F_{R,\gamma} = \frac{MS\gamma_R}{MSE_R} \text{ and } F_{R,\phi} = \frac{MS\phi_R}{MSE_{\phi,R}}$$

The next theorem gives the asymptotic distribution of test statistics, based on $F_{R,\beta}$, $F_{R,\gamma}$ and $F_{R,\phi}$ under $H_0(\beta)$, $H_0(\gamma)$, and $H_0(\phi)$, respectively.

Theorem 4.2.4. *Assume for each group i and subject k, $X_{ijk}, j = 1, 2, \ldots,$ is α-mixing with $\alpha_m = O\left(m^{-5}\right)$. Set $Y_{ijk} = H(X_{ijk})$, let $\widetilde{\sigma}_{ijj'} = cov(Y_{ijk}, Y_{ij'k})$, and let $\widetilde{\tau}_\beta$, $\widetilde{\tau}_\gamma$, $\widetilde{\sigma}^4$, $\widetilde{\tau}_{\beta*}$, $\widetilde{\tau}_{\gamma*}$, $\widetilde{\sigma}^4_*$ be similarly defined as τ_β, τ_γ, σ^4, $\tau_{\beta*}$, $\tau_{\gamma*}$ and σ^4_* in Theorem 4.2.1 with $\sigma_{ijj'}$ replaced by $\widetilde{\sigma}_{ijj'}$. Then as $b \to \infty$ while a remains fixed,*

(1) for $n_i \geq 2$ fixed,

$$\text{under } H_0(\widetilde{\beta}),\ \sqrt{b}(F_{R,\beta} - 1) \xrightarrow{d} N\left(0, \widetilde{\tau}^2_\beta/\widetilde{\sigma}^4\right);$$
$$\text{under } H_0(\widetilde{\gamma}),\ \sqrt{b}(F_{R,\gamma} - 1) \xrightarrow{d} N\left(0, \widetilde{\tau}^2_\gamma/\widetilde{\sigma}^4\right);$$
$$\text{under } H_0(\widetilde{\phi}),\ \sqrt{b}(F_{R,\phi} - 1) \xrightarrow{d} N\left(0, \widetilde{\tau}^2_\gamma/\widetilde{\sigma}^4\right).$$

(2) if $n_i \to \infty$ as $b \to \infty$, under the additional assumption $\max_i\{n_i\}/\min_i\{n_i\} = O(1)$,

$$\text{under } H_0(\widetilde{\beta}),\ \sqrt{b}(F_{R,\beta}-1) \xrightarrow{d} N\left(0, \widetilde{\tau}^2_{\beta*}/\widetilde{\sigma}^4_*\right);$$

$$\text{under } H_0(\widetilde{\gamma}),\ \sqrt{b}(F_{R,\gamma}-1) \xrightarrow{d} N\left(0, \widetilde{\tau}^2_{\gamma*}/\widetilde{\sigma}^4_*\right);$$

$$\text{under } H_0(\widetilde{\phi}),\ \sqrt{b}(F_{R,\phi}-1) \xrightarrow{d} N\left(0, \widetilde{\tau}^2_{\gamma*}/\widetilde{\sigma}^4_*\right).$$

Proposition 4.2.5. *Let $\tilde{\zeta}_1$ and $\tilde{\zeta}_2$ be defined as ζ_1 and ζ_2 in Theorem 4.2.1 with $\sigma_{ijj'}$ replaced by $\widetilde{\sigma}_{ijj'}$. Also let $\widehat{\tilde{\zeta}}_1$ and $\widehat{\tilde{\zeta}}_2$ be defined as $\widehat{\zeta}_1$ and $\widehat{\zeta}_2$ in Proposition 4.2.2 with X_{ijk} replaced by their overall (mid-)ranks R_{ijk}. Then as $b \to \infty$, $\widehat{\tilde{\zeta}}_1/N^4 - \tilde{\zeta}_1 \xrightarrow{p} 0$ and $\widehat{\tilde{\zeta}}_2/N^4 - \tilde{\zeta}_2 \xrightarrow{p} 0$, provided that $h < 1/2$.*

Theorem 4.2.6. *Let $\mathbf{W}_R = (\overline{R}_{1..}, \cdots, \overline{R}_{a..})'$, where R_{ijk} is the overall (mid-)rank of X_{ijk} among all observations. Assume $X_{i1k}, X_{i2k}, \cdots$, is α-mixing with $\alpha_m = O\left(m^{-5}\right)$ for all i,k. Let $C_l(j,h)$, $C_u(j,h)$ be as defined in Proposition 4.2.2 and set*

$$\widehat{\eta}_{Ri} = \frac{n}{bn_i(n_i-1)} \sum_{j=1}^{b} \sum_{j'=C_l(j,h)}^{C_u(j,h)} \sum_{k=1}^{n_i} \left(R_{ijk} - \overline{R}_{ij.}\right)\left(R_{ijk} - \overline{R}_{ij.}\right)\ i = 1, \ldots, a.$$

Then under $\widetilde{H}_0(A)$,

$$N\mathbf{W}_R'\mathbf{C}_a'\left[\mathbf{C}_a\, diag(\hat{\eta}_{R1}, \cdots, \hat{\eta}_{Ra})\mathbf{C}_a'\right]^{-1}\mathbf{C}_a\mathbf{W}_R \xrightarrow{d} \chi^2_r,$$

where r is the rank of $\mathbf{C}_a$, as $b \to \infty$ regardless of whether the n_i remain fixed or tend to ∞ with b provided $\max n_i/\min n_i = O(1)$ and $h < 0.5$.

4.3 Numerical Results

4.3.1 Simulations

In the following simulations we compare the two proposed tests, based on the original observations (OBS) and on the ranks (RANK), and the univariate F-tests (UF). All results are based on 2000 simulation cycles.

Table 4.1 is the only one comparing the above tests for the hypothesis of no main treatment effect. For the normal and cauchy cases, the data are generated as $X_{ijk} = d_{ijk} + \zeta_{ijk}$, $i = 1,2,\ j = 1,\ldots,b,\ k = 1,\ldots,4$, where $d_{ijk} = (0.4)d_{i,j-1,k} + \varepsilon_{ijk}$, and ζ_{ijk}, ε_{ijk} are iid N(0,1) or Cauchy(0,1); for the lognormal case the data are generated by simply exponentiating data generated according to the normal case. Eventhough the data generation does not conform to the assumptions underlying the UF test, UF performs well under normality, as do the two proposed tests. Both UF and OBS perform poorly under both lognormal and Cauchy, but RANK continues to perform well in all cases and with all choices of b^h, except for the cauchy case with $b = 15$. Simulations for testing for no group effect in unbalanced designs (sizes 4 and 14) gave similar results and thus are not reported.

The type I error-rate for testing no treatment effect over time, no time effect, and no simple treatment effect at $\alpha = 0.05$ is reported in Tables 4.2, 4.3 4.4, and 4.5 for designs with number of time points $b = 20$, 30, 50, and 80. In Tables 4.2 and 4.3 the data are generated exactly as in Table 4.1. Thus we have a balanced design with both groups having vector observations of exactly the same correlation structure. In Tables 4.4 and 4.5 the data are still generated as in Table 4.1 except for the fact that we used an unbalanced design (see the caption of the table) and the error term ζ_{ijk} has scale parameter i. From Tables 4.2 and 4.3 it can be seen that UF is robust to the present type of deviations from the assumptions underlying its properties, provided normality holds. Both UF and OBS perform poorly under both the lognormal and Cauchy distributions. From Tables 4.4 and 4.5 we see that the aforementioned robustness of UF under normality holds only for testing for no main time effect, but not for group-time interaction or simple group effect. OBS continues to perform well under normality, but both UF and OBS perform poorly under the lognormal and Cauchy distributions. On the other hand, RANK performs well in all situations in Tables 4.2, 4.3, 4.4 and 4.5. In a related simulation, not reported here, we used the set-up of Table 4.2, so balanced design, except that the coefficient 0.4 used in the generation of d_{ijk} (see the above description of data generation) was replaced by 0.1 for the second group. With this change the UF for no main time effect continued to be acceptable, but testing for no group-time interaction or simple group effect resulted in the following changes in the permance of UF: a) it became more conservative than the corresponding tests in Tables 4.4 and 4.5 under normality, b) it became very liberal under Cauchy, and c) it was in the acceptable range under lognormality.

Results from power simulations are summarized in Figures 4.1 and 4.2. The simmulations used alternatives to the setting of Table 4.2, only for testing for no treatment effect with $b = 50$ and $b^h = 4$ (Figure 4.1), and for testing for no interaction effect with $b = 20$ and $b^h = 4$ (Figure 4.2). In Figure 1 we see that, though the type I error rate of UF for testing for main group effect is robust to violations of the assumptions of equal correlation, the power of UF suffers relative to that of OBS and RANK even under normality. OBS and UF have almost no power under lognormality and the Cauchy distribution. Same comments apply for Figure 2, except that the power of UF for testing for no interaction is comparable to OBS and RANK under normality.

Table 4.1: Estimated level for testing group effect at $\alpha = 0.05$. The number of groups is $a = 2$, each of size $n = 4$, and the number of time points is b.

		normal			lognormal			Cauchy		
b	b^h	UF	OBS	RANK	UF	OBS	RANK	UF	OBS	RANK
15	2	0.0515	0.0650	0.0625	0.0195	0.0260	0.0625	0.0190	0.0345	0.0930
	3	0.0515	0.0690	0.0670	0.0195	0.0295	0.0670	0.0190	0.0330	0.0910
	4	0.0515	0.0695	0.0665	0.0195	0.0405	0.0665	0.0190	0.0365	0.0880
50	2	0.0475	0.0660	0.0725	0.0205	0.0270	0.0500	0.0195	0.0245	0.0570
	3	0.0475	0.0645	0.0620	0.0205	0.0265	0.0490	0.0195	0.0250	0.0560
	4	0.0475	0.0640	0.0650	0.0205	0.0270	0.0500	0.0195	0.0250	0.0540
	5	0.0475	0.0660	0.0675	0.0205	0.0270	0.0510	0.0195	0.0280	0.0560
	6	0.0475	0.0645	0.0630	0.0205	0.0275	0.0530	0.0195	0.0285	0.0575
80	2	0.0525	0.0570	0.0575	0.0205	0.0270	0.050	0.0250	0.0255	0.0575
	3	0.0525	0.0555	0.0545	0.0205	0.0265	0.0490	0.0250	0.0275	0.0540
	4	0.0525	0.0560	0.0530	0.0205	0.0270	0.0500	0.0250	0.0280	0.0525
	5	0.0525	0.0555	0.0530	0.0205	0.0270	0.0510	0.0250	3e-02	0.0525
	6	0.0525	0.0550	0.0520	0.0205	0.0275	0.0530	0.0250	0.0290	0.0520
	7	0.0525	0.0570	0.0530	0.0205	0.0305	0.0545	0.0250	3e-02	0.0555

Table 4.2: Estimated level for testing time (T), interaction (GT), and simple group effect (G(T)) at $\alpha = 0.05$. The number of groups is $a = 2$, each of size $n = 4$, and the number of time points is b.

			normal			lognormal			Cauchy		
b	b^h	Effect	UF	OBS	RANK	UF	OBS	RANK	UF	OBS	RANK
20	3	T	0.0535	0.0525	0.051	0.0285	0.0015	0.051	0.0210	0.0030	0.0630
		GT	0.0550	0.0555	0.062	0.0185	0.0005	0.062	0.0200	0.0025	0.0535
		G(T)	0.0605	0.0630	0.070	0.0190	0.0010	0.070	0.0195	0.0035	0.0655
20	4	T	0.0535	0.048	0.0500	0.0285	0.0015	0.0500	0.0210	0.0025	0.0590
		GT	0.0550	0.054	0.0595	0.0185	0.0005	0.0595	0.0200	0.0025	0.0520
		G(T)	0.0605	0.058	0.0635	0.0190	0.0010	0.063	0.0195	0.0035	0.0625
30	3	T	0.0645	0.0585	0.0655	0.0270	0.0015	0.0655	0.0205	0.0015	0.0595
		GT	0.0570	0.0510	0.0565	0.0270	0.0005	0.0565	0.0210	0.0020	0.0460
		G(T)	0.0585	0.0555	0.0550	0.0285	0.0005	0.0550	0.0205	0.0020	0.0590
30	4	T	0.0645	0.0550	0.0615	0.0270	0.0015	0.061	0.0205	0.0015	0.0565
		GT	0.0570	0.0495	0.0530	0.0270	0.0005	0.053	0.0210	0.0020	0.0435
		G(T)	0.0585	0.0520	0.0525	0.0285	0.0005	0.0525	0.0205	0.0020	0.0550

Table 4.3: Table 4.2 continued

			normal			lognormal			Cauchy		
b	b^h	Effect	UF	OBS	RANK	UF	OBS	RANK	UF	OBS	RANK
50	3	T	0.0595	0.0490	0.0565	0.0285	0.0010	0.0440	0.0245	0.0005	0.0505
		GT	0.0580	0.0465	0.0490	0.0210	0.0005	0.0500	0.0165	0.0005	0.0585
		G(T)	0.0600	0.0490	0.0520	0.0205	0.0005	0.0510	0.0190	0.0005	0.0655
50	4	T	0.0595	0.0435	0.0535	0.0285	0.0010	0.0395	0.0245	0.0005	0.0475
		GT	0.0580	0.0400	0.0440	0.0210	0.0005	0.0455	0.0165	0.0005	0.0530
		G(T)	0.0600	0.0470	0.0465	0.0205	0.0005	0.0445	0.0190	0.0005	0.0615
80	3	T	0.0640	0.0505	0.0495	0.0275	0.0010	0.050	0.0205	0.0000	0.0445
		GT	0.0530	0.0420	0.0505	0.0170	0.0005	0.046	0.0175	0.0000	0.0500
		G(T)	0.0545	0.0435	0.0520	0.0165	0.0005	0.049	0.0190	0.0005	0.0505
80	4	T	0.0640	0.0440	0.0455	0.0275	0.0010	0.0440	0.0205	0.0000	0.0375
		GT	0.0530	0.0385	0.0460	0.0170	0.0005	0.0430	0.0175	0.0000	0.0445
		G(T)	0.0545	0.0415	0.0470	0.0165	0.0005	0.0440	0.0190	0.0005	0.0460

Table 4.4: Estimated level for testing time (T), interaction (GT), and simple group effect (G(T)) at $\alpha = 0.05$. The number of groups is $a = 2$, with sizes 4, 7, and the number of time points is b.

			normal			lognormal			Cauchy		
b	b^h	Effect	UF	OBS	RANK	UF	OBS	RANK	UF	OBS	RANK
		T	0.056	0.0640	0.0665	0.0230	0.0015	0.0560	0.0225	0.0045	0.059
20	4	GT	0.029	0.0605	0.0565	0.0325	0.0000	0.0535	0.1470	0.0025	0.058
		G(T)	0.028	0.0630	0.0625	0.0305	0.0025	0.0590	0.1520	0.0030	0.075
		T	0.056	0.0615	0.0670	0.0230	0.0015	0.0545	0.0225	0.0045	0.0585
20	5	GT	0.029	0.0590	0.0550	0.0325	0.0000	0.0500	0.1470	0.0025	0.0550
		G(T)	0.028	0.0630	0.0585	0.0305	0.0025	0.0565	0.1520	0.0030	0.0730
		T	0.056	0.0585	0.0635	0.0230	0.0015	0.0510	0.0225	0.0045	0.0560
20	6	GT	0.029	0.0560	0.0530	0.0325	0.0000	0.0485	0.1470	0.0025	0.0530
		G(T)	0.028	0.0605	0.0570	0.0305	0.0025	0.0545	0.1520	0.0030	0.0695
		T	0.056	0.0565	0.0620	0.0230	0.0015	0.0490	0.0225	0.0045	0.0545
20	7	GT	0.029	0.0565	0.0530	0.0325	0.0000	0.0445	0.1470	0.0025	0.0530
		G(T)	0.028	0.0590	0.0550	0.0305	0.0025	0.0540	0.1520	0.0030	0.0680

Table 4.5: Table 4.4 continued

			normal			lognormal			Cauchy		
b	b^h	Effect	UF	OBS	RANK	UF	OBS	RANK	UF	OBS	RANK
30	4	T	0.0515	0.0515	0.0540	0.0230	0.0015	0.0560	0.0170	0.0010	0.0530
		GT	0.0220	0.0485	0.0530	0.0325	0.0000	0.0535	0.1760	0.0010	0.0525
		G(T)	0.0240	0.0560	0.0595	0.0305	0.0025	0.0590	0.1775	0.0005	0.0655
30	5	T	0.0515	0.0500	0.0490	0.0230	0.0015	0.0545	0.0170	0.0010	0.0505
		GT	0.0220	0.0450	0.0490	0.0325	0.0000	0.0500	0.1760	0.0010	0.0505
		G(T)	0.0240	0.0535	0.0580	0.0305	0.0025	0.0565	0.1775	0.0005	0.0615
30	6	T	0.0515	0.0455	0.0460	0.0230	0.0015	0.0510	0.0170	0.0010	0.0490
		GT	0.0220	0.0430	0.0475	0.0325	0.0000	0.0485	0.1760	0.0010	0.0480
		G(T)	0.0240	0.0525	0.0540	0.0305	0.0025	0.0545	0.1775	0.0005	0.0575
30	7	T	0.0515	0.044	0.0430	0.0230	0.0015	0.0490	0.0170	0.0010	0.0470
		GT	0.0220	0.043	0.0440	0.0325	0.0000	0.0445	0.1760	0.0010	0.0445
		G(T)	0.0240	0.050	0.0475	0.0305	0.0025	0.0540	0.1775	0.0005	0.0535

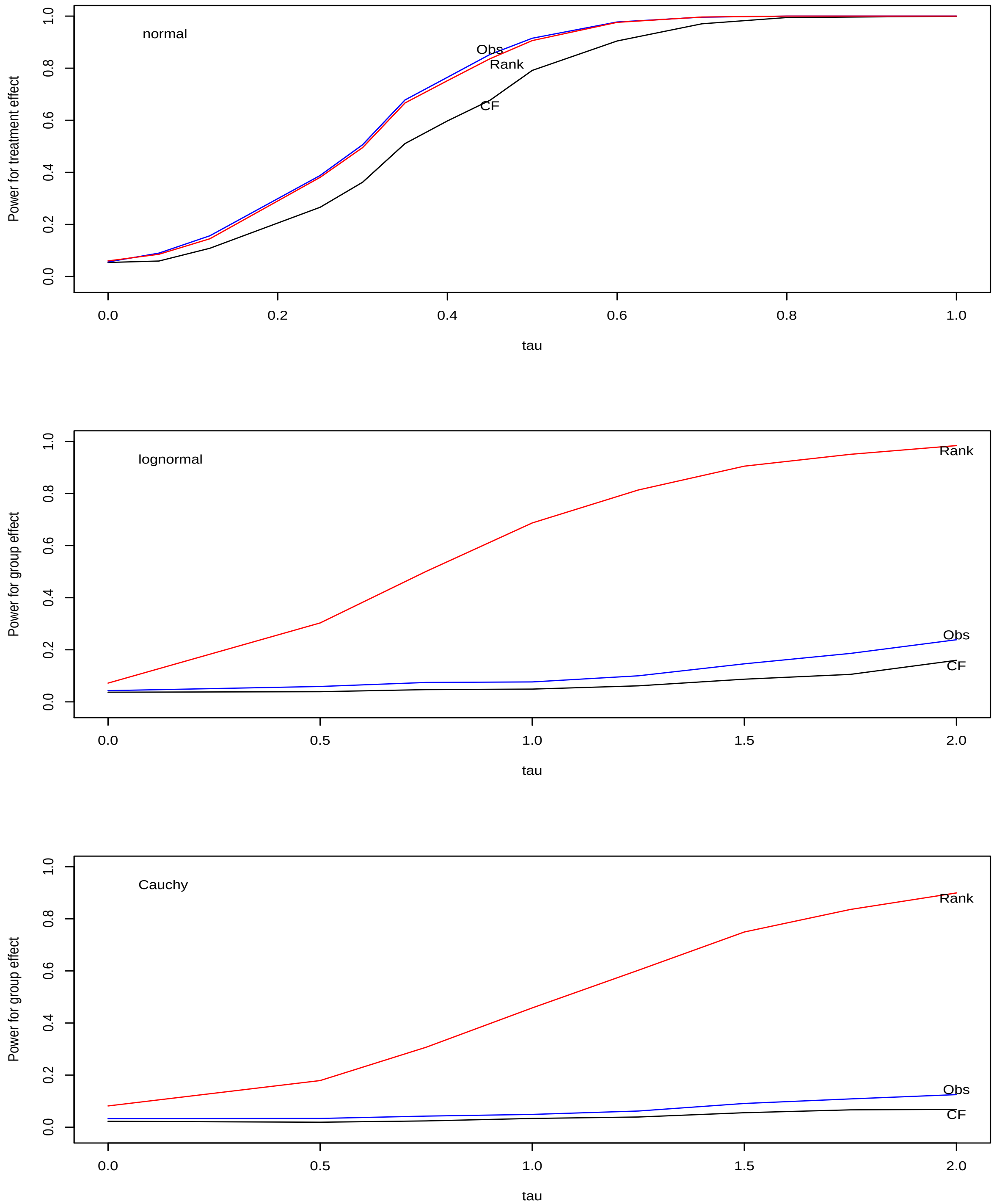

Figure 4.1: *Power function for group effect for alternatives of the form* $d_{ijk} = (-1)^i\tau + 0.4d_{i,j-1,k} + \varepsilon_{ijk}$ *and* $X_{ijk} = d_{ijk} + \zeta_{ijk}$, *as in Table 4.1*

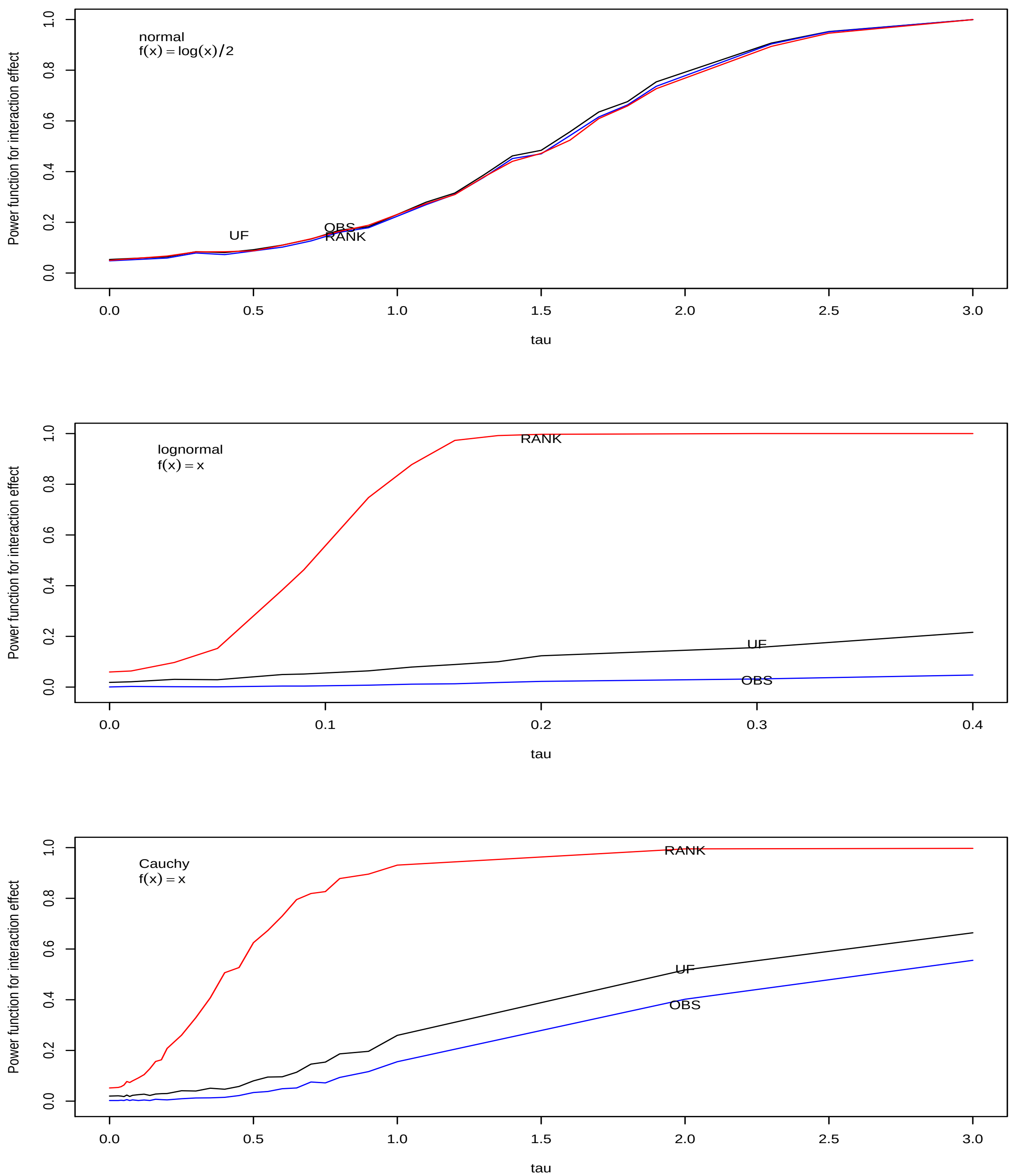

Figure 4.2: *Power function for interaction effect for alternatives of the form* $d_{ijk} = (-1)^i \tau\, f(j-1) + 0.4 d_{i,j-1,k} + \varepsilon_{ijk}$ *and* $X_{ijk} = d_{ijk} + \zeta_{ijk}$*, as in Table 4.2*

4.3.2 Data analysis

Here we apply the proposed methods on two data sets, one with continuous data and one with discrete. In both examples the interest lies in treatment effects and how they vary with time.

Examle 1: *Legumes and beef production.* A four treatment experiment was conducted to examine the effect of legumes on beef production (Verbyla and Cullis 1990). For treatments 2-4 legumes were sown into cultivated native pasture with different amounts of superphosphate applied at the time of sowing and annually. Treatment one used no legumes and the standard amount of superphosphate applied. For a complete description of the treatments see Table 1. The pasture for each treatment was split in four replicate paddocks, and two Hereford heifers were allocated at random to each paddock and replaced after 2 years. The study lasted for 4 years, with 17 measurements of live-weight per paddock taken at equal time intervals for each of the 2-year periods. Data were complete on all paddocks.

The specific questions of interest are a) the effect of legumes, b) the effect of the superphosphate amount at sowing, c) the effect of annual superphosphate amount, and d) how these effects change with time. Questions a), b) and c) are examples where the interest not in the usual hypothesis $H_0(\alpha)$ (respectively $H_0(A)$) but in $\widetilde{H}_0(\alpha)$ (respectively $\widetilde{H}_0(A)$). In particular, question a) is answered by testing that the difference between treatments 3 and 1, the legume contrast, is zero; question b) is answered by testing that the difference between treatments 3 and 2, the sowing contrast, is zero; and question c) is answered by testing that the difference between treatments 3 and 4, the annual contrast, is zero. For our analysis, we performed a simultaneous test for the contrasts involved in questions a), b) and c). Thus we will use Theorems

Table 4.6: Treatment for the experiment on legumes and beef production

Treatment	Pasture treatment	Superphosphate applied (kg/hectare)	
		Sowing	Annual
1	Native pasture	250	125
2	Native + Legumes	125	125
3	Native + Legumes	250	125
4	Native + Legumes	250	0

4.2.3 and 4.2.6 with contrast matrix C_a given by

$$C_a = \begin{pmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 \end{pmatrix}$$

The results are listed in the following table. The b^h which enters the definition of the constants $C_l(j,h), C_u(j,h)$, used in the calculation of the covariance matrix in these theorems, was set to 4 but the values of 3 and 5 produced similar results. The table also reports p-values for interaction and for simple treatment effects, as well as corresponding p-values for the classical (univariate) F-test. A p-value of 0.000 was recorded when the p-value was less than $1e-4$. We remark that in this case we have the same number of units as time points (17 live-weight measurements per two-year period yield 16 weight gain measurements), and thus application of the multivariate (unstructured covariance matrix) F-test is not recommended.

As seen from Table 4.7 at least one of the contrasts of interest is significant in both two-year periods according to our test procedures, but only in the second two-year period ac-

Table 4.7: P-values of the hypotheses tests (with $b^h = 4$)

Effect	Year 1 and 2			Year 3 and 4		
	UF	OBS	RANK	UF	OBS	RANK
Contrasts	0.091	0.000	8.79e-4	1.25e-4	0.000	0.000
Time	0.000	0.000	0.000	0.000	0.000	0.000
Treatment×Time	0.211	0.273	0.509	5.09e-4	2.23e-4	0.015
Simple Treatment	0.126	0.193	0.419	0.000	0.000	0.000

cording to UF. Panels (a) and (c) of Figure 4.3 plot the effects of the individual contrasts of interest for the first and second two-year periods, respectively. For example, the legume contrast effect plot is a plot of the contrast function $\widehat{\overline{F}}_{3.} - \widehat{\overline{F}}_{1.}$, and similarly for the other contrast plots. These panels suggest that the legume effect is significant in the second period, that the initial superphosphate effect is significant in the first period, while the annual superphosphate effect is significant in the second period. (Formal tests for the individual contrast effects can also be done.) For the other tests, the results from all procedures are in agreement. As another illustration of the nonparametric effects, panels (b) and (d) of Figure 4.3 plot $\widehat{\overline{F}}_{\cdot j}(x)$ as a function of time (j) and weight gain (x). Note that $\widehat{\overline{F}}_{\cdot j}(x)$ is the main component of the estimated nonparametric effect, $\widehat{B}_j(x)$, of the time factor. These panels confirm the highly significant test results for the hypothesis of no time effect reported in Table 4.7, i.e. that the distribution functions $\overline{F}_{\cdot j}(x)$, $j = 1, \ldots, b$ are different. Moreover, the plots illustrate the change of $\overline{F}_{\cdot j}(x)$ as a function of j for each fixed x. We note that though panels (b) and (d) visualize nonparametric main effects, just like panels (a) and (c) do, a three-dimentional plot is more appropriate due to the many levels of the time factor.

Examle 2: *Academic and Transitional Experiences of High School Youths at Risk.* The

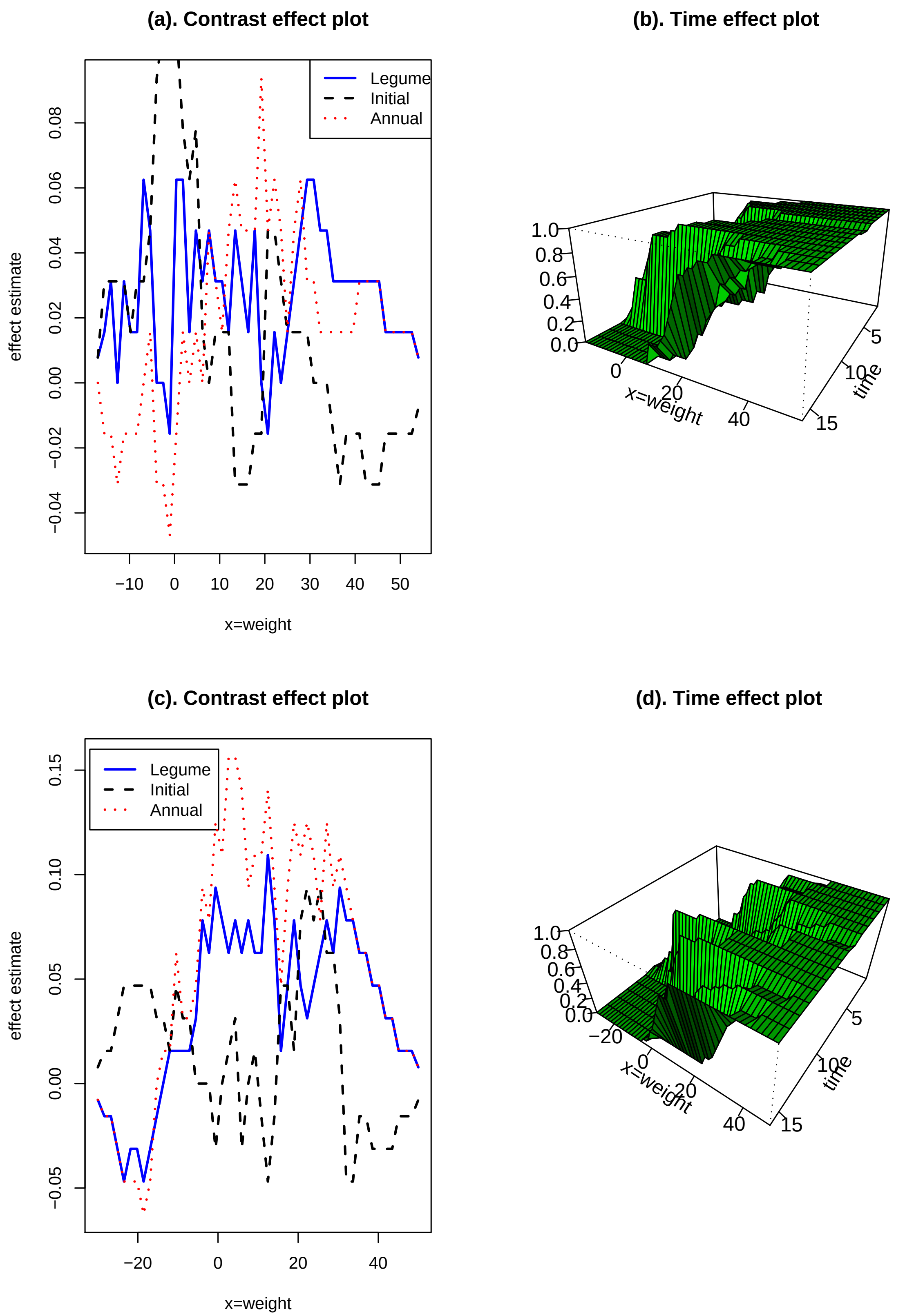

Figure 4.3: *(a) and (b) are Contrast effect and Time effect in the first two-year period; (c) and (d) are Contrast effect and Time effect in the second two-year period*

purpose of this study was to investigate whether employment and unemployment experiences of youth at risk vary by race, gender, and residence. Youth at risk means individuals aged 12-16 years living in households that are below the poverty line. The data were made available to us courtesy of Jane Okou at the Pennsylvania State University. A total of 517 subjects were used, of which 215 are African Americans, 133 are Hispanics, and 169 are whites. The subjects are further classified by gender (230 male, 287 female), and by residence (415 rural and 102 urban). For each subject the study recorded the youth's job status (0 for unemployed, 1 for employed) at each of the 53 weeks.

The group sizes for each of the 12 factor level combinations is given in Table 4.3.2. It is seen that the design is highly unbalanced, with same group sizes being very small. In this table, levels 1, 2 and 3 for race stand for African American, Hispanic and white, respectively; levels 0 and 1 for gender stand for male and female, respectively; and levels 0 and 1 for residence stand for rural and urban, respectively.

	treatment											
	1	2	3	4	5	6	7	8	9	10	11	12
race	1	1	1	1	2	2	2	2	3	3	3	3
gender	0	0	1	1	0	0	1	1	0	0	1	1
residence	0	1	0	1	0	1	0	1	0	1	0	1
group sizes	15	73	18	109	7	59	6	61	21	55	35	58

To test for the effect of interest, we will use Theorems 4.2.3 and 4.2.6 with contrast matrices given as follows. Let $M_d = (\mathbf{1}_{d-1} | -\mathbf{I}_{d-1})$, where $\mathbf{1}_k$ is the k-dimensional vector of ones, and $\mathbf{I}_k$ is the $k \times k$ identity matrix, and set $\mathbf{C}_A = M_3 \otimes (0.5\mathbf{1}_2') \otimes (0.5\mathbf{1}_2')$, $\mathbf{C}_B = (1/3\mathbf{1}_3') \otimes M_2 \otimes (0.5\mathbf{1}_2')$, $\mathbf{C}_C = (1/3\mathbf{1}_3') \otimes 0.5\mathbf{1}_2' \otimes M_2$, $\mathbf{C}_{(AB)} = M_3 \otimes M_2 \otimes (0.5\mathbf{1}_2')$,

$\mathbf{C}_{(AC)} = M_3 \otimes (0.5\mathbf{1}_2') \otimes M_2$, $\mathbf{C}_{(BC)} = (1/3\mathbf{1}_3') \otimes M_2 \otimes M_2$, $\mathbf{C}_{(ABC)} = M_3 \otimes M_2 \otimes M_2$. The matrices $\mathbf{C}_A$, $\mathbf{C}_B$, $\mathbf{C}_C$, $\mathbf{C}_{(AB)}$, $\mathbf{C}_{(AC)}$, $\mathbf{C}_{(BC)}$ and $C_{(ABC)}$ are the contrast matrices for main effect of race, gender, residence, and interaction effect of race×gender, race×residence, gender×residence, and race×gender×residence.

The p-values of the tests are reported in Table 4.8 below (a p-value of 0.000 was recorded when the p-value was less than $1e-4$). OBS and RANK yield exactly the same results, as should be expected with a dichotomous response variable. All tests indicate that residence does not affect employment experience, and this is confirmed by the nonparametric residence effect plotted in Figure 4.4. The plots as well as test results from OBS and RANK indicate that race and gender are significant factors. In addition, the plots reveal that the difference between Hispanics and whites is not as significant as that between African Americans and whites. Table 4.8 also reports the p-values for testing for no time effect and treatment×time interaction. For this last test we combined all factor level combinations into the factor 'treatment' with 12 levels. As an illustration panel (b) of Figure 4.4 plots one of the treatment×time interaction effects, which is basically noise around zero.

4.4 Proofs

4.4.1 Proofs for test based on original observations

The proofs of the theorems will use a basic result which is stated as a lemma and proved here. All other lemmas used in the proof of theorems will be stated and proved following the proof of corresponding Theorem. In this section, we denote $u_{ijk} = X_{ijk} - \mu_{ij}$.

Lemma 4.4.1. *Suppose $X_1, X_2, \cdots,$ is α-mixing with $\alpha_m = O(m^{-5})$, $E(X_i) = 0$ and $\limsup_i E(X_i^{16}) < \infty$, then there exists K, such that*

Table 4.8: P-values of the hypotheses tests (with $b^h = 4$)

Effect	UF	OBS	RANK
Race (A)	0.789	0.000	0.000
Gender (B)	0.524	0.000	0.000
Residence (C)	0.298	0.863	0.863
Race×Gender (AB)	0.160	0.032	0.032
Race×Residence (AC)	0.040	0.377	0.377
Gender×Residence (BC)	0.939	0.000	0.000
Race×Gender×Residence (ABC)	0.397	0.492	0.492
Time	0.000	0.156	0.156
Treatment×Time	0.983	0.548	0.548

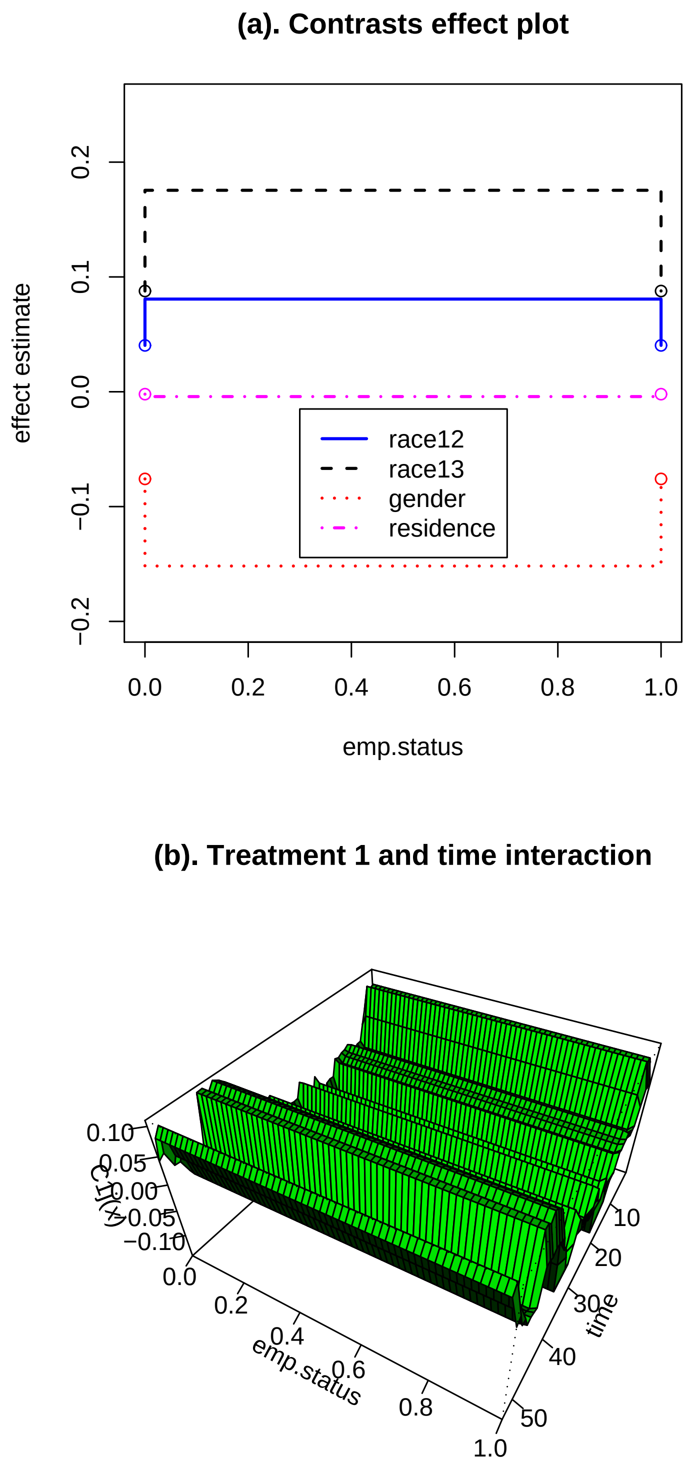

Figure 4.4: *Contrasts effect and Treatment 1 effect over time*

a.

$$\sum_{j<j'}^{b} E(X_j X_{j'}) \leq Kb, \qquad \sum_{j_1<j_2<j_3}^{b} E(X_{j_1}^2 X_{j_2} X_{j_3}) \leq Kb^2, \tag{4.4.1}$$

$$\sum_{j_1<j_2<j_3<j_4}^{b} E(X_{j_1} X_{j_2} X_{j_3} X_{j_4}) \leq Kb^2.$$

b. $S_b = \sum_{j=1}^{b} X_j$ *satisfies* $b^{-1}E(S_b^2) \to \sigma^2$ *for some finite* σ^2 *, and* $S_b/\sqrt{b} \xrightarrow{d} N(0, \sigma^2)$.

Proof. By given condition $\limsup_i E(X_i^{16}) < \infty$, there exists K_1, such that

$$\begin{aligned} \sum_{j<j'}^{b} E(X_j X_{j'}) &= \sum_{j'=2}^{b} \sum_{j=1}^{j'-1} E(X_j X_{j'}) \leq \sum_{j'=2}^{b} \sum_{j=1}^{j'-1} 8[1 + E(X_j^4) + E(X_{j'}^4)] \alpha_{j'-j}^{1/2} \\ &\leq K_1 \sum_{j'=2}^{b} \sum_{j=1}^{j'-1} \alpha_{j'-j}^{1/2} = K_1 \sum_{j'=2}^{b} \sum_{j_1=1}^{j'-1} \alpha_{j_1}^{-5/2} \\ &\leq K_1 \sum_{j'=2}^{b} \sum_{j_1=1}^{\infty} \alpha_{j_1}^{-5/2} \leq Kb \end{aligned}$$

since $\sum_{j_1=1}^{\infty} \alpha_{j_1}^{-5/2} < \infty$.

$$\begin{aligned} \sum_{j_1<j_2<j_3}^{b} E(X_{j_1}^2 X_{j_2} X_{j_3}) &\leq \sum_{j_1<j_2<j_3}^{b} 8[1 + E(X_{j_1}^2)^4 + E(X_{j_2} X_{j_3})^4] \alpha_{j_2-j_1}^{1/2} \\ &\leq \sum_{j_1<j_2<j_3}^{b} 8[1 + E(X_{j_1}^8) + (E(X_{j_2}^8))^{1/2} (E(X_{j_3}^8))^{1/2}] \alpha_{j_2-j_1}^{1/2} \\ &\leq K_1 \sum_{j_1<j_2<j_3}^{b} \alpha_{j_2-j_1}^{1/2} \leq K_1 b \sum_{j_1<j_2}^{b} \alpha_{j_2-j_1}^{1/2} \leq Kb^2 \end{aligned}$$

For $j_1 < j_2 < j_3 < j_4$, by α-mixing condition, there exists K_2, such that

$$\begin{aligned} &|E(X_{j_1} X_{j_2} X_{j_3} X_{j_4})| \leq 8[1 + E(X_{j_1} X_{j_2})^4 + E(X_{j_3} X_{j_4})^4] \alpha_{j_3-j_2}^{1/2} \\ \leq \;& 8\left[1 + [E(X_{j_1}^8)]^{1/2} [E(X_{j_2}^8)]^{1/2} + [E(X_{j_3}^8)]^{1/2} [E(X_{j_4}^8)]^{1/2}\right] \alpha_{j_3-j_2}^{1/2} \leq K_2 \alpha_{j_3-j_2}^{1/2}, \end{aligned}$$

Also,

$$|E(X_{j_1}X_{j_2}X_{j_3}X_{j_4})| \le 8[1+E(X_{j_1}^4)+E(X_{j_2}X_{j_3}X_{j_4})^4]\alpha_{j_2-j_1}^{1/2}$$

$$\le \quad 8\left[1+E(X_{j_1}^4)+[E(X_{j_2}^{12})]^{1/3}[E(X_{j_3}^{12})]^{1/3}[E(X_{j_4}^{12})]^{1/3}\right]\alpha_{j_2-j_1}^{1/2} \le K_2\alpha_{j_2-j_1}^{1/2},$$

So

$$|E(X_{j_1}X_{j_2}X_{j_3}X_{j_4})| \le K_2 \min(\alpha_{j_2-j_1}^{1/2}, \alpha_{j_3-j_2}^{1/2}),$$

and

$$\sum_{j_1<j_2<j_3<j_4}^{b} E(X_{j_1}X_{j_2}X_{j_3}X_{j_4}) \le K_2 \sum_{j_1<j_2<j_3<j_4}^{b} \min(\alpha_{j_2-j_1}^{1/2}, \alpha_{j_3-j_2}^{1/2})$$

$$\le \quad K_2 b \sum_{j_1<j_2<j_3}^{b} \min(\alpha_{j_2-j_1}^{1/2}, \alpha_{j_3-j_2}^{1/2}) = K_2 b \sum_{j_3=3}^{b}\sum_{j_2=2}^{j_3-1}\sum_{j_1=1}^{j_2-1} \min(\alpha_{j_2-j_1}^{1/2}, \alpha_{j_3-j_2}^{1/2})$$

$$= \quad K_2 b \sum_{j_3=3}^{b}\sum_{j_2=2}^{j_3-1}\sum_{m=1}^{j_2-1} \min(\alpha_m^{1/2}, \alpha_{j_3-j_2}^{1/2}) = K_2 b \sum_{j_3=3}^{b}\sum_{t=2}^{j_3-1}\sum_{m=1}^{j_3-t-1} \min(\alpha_m^{1/2}, \alpha_t^{1/2})$$

$$\le \quad K_2 b \sum_{j_3=3}^{b} \sum_{t\ge1, m\ge1, t+m<j_3} \min(\alpha_m^{1/2}, \alpha_t^{1/2}) \le 2K_2 b \sum_{j_3=3}^{b} \sum_{m\le t, t+m<j_3} \alpha_t^{1/2} \quad (4.4.2)$$

$$\le \quad 2K_2 b \sum_{j_3=3}^{b}\sum_{t=1}^{j_3}\sum_{m=1}^{t} = 2K_2 b \sum_{j_3=3}^{b}\sum_{t=1}^{j_3} t\alpha_t^{1/2} \le 2K_2 b^2 \sum_{t=1}^{\infty} t\alpha_t^{1/2} \le Kb^2, \quad (4.4.3)$$

where (4.4.2) is because $\alpha_m^{1/2} \ge \alpha_t^{1/2}$ when $m \le t$, and (4.4.3) is because $\sum_{t=1}^{\infty} t\alpha_t^{1/2} < \infty$. By now, we finished part (a) of the proof.

Next we will show part (b).

$$E(S_m^2) = \sum_{j=1}^{m}\sum_{j'=1}^{m} E(X_j X_{j'}) = \sum_{j=1}^{m} E(X_j^2) + 2\sum_{j<j'}^{m} E(X_j X_{j'}) \le Km. \quad (4.4.4)$$

By (4.4.1), $b^{-1}E(S_b^2)$ converges to some finite number σ^2.

To verify the asymptotic distribution , we split the sum $X_1 + \cdots + X_b$ into alternate blocks of length B_b (the big blocks) and L_b (the small blocks). Namely, let

$$U_{bj} = X_{(j-1)(B_b+L_b)+1} + \cdots + X_{(j-1)(B_b+L_b)+B_b}, 1 \le j \le r_b, \quad (4.4.5)$$

where r_b is the largest integer j for which $(j-1)(B_b+L_b)+B_b<b$. Further, let

$$V_{bj}=X_{(j-1)(B_b+L_b)+B_b+1}+\cdots+X_{j(B_b+L_b)}, 1\leq j<r_b, \tag{4.4.6}$$

$$V_{br_b}=X_{(r_b-1)(B_b+L_b)+B_b+1}+\cdots+X_b. \tag{4.4.7}$$

Then $S_b=\sum_{j=1}^{r_b}U_{bj}+\sum_{j=1}^{r_b}V_{bj}$, and the technique will be to choose the L_b small enough that $\sum_{j=1}^{r_b}V_{bj}$ is small in comparison with $\sum_{j=1}^{r_b}U_{bj}$ but large enough that the U_{bj} are nearly independent, so that Lyapounov's theorem can be adapted to prove $\sum_{j=1}^{r_b}U_{bj}$ asymptotically normal.

Take $B_b=[b^{3/4}]$ and $L_b=[b^{1/4}]$. If r_b is the largest integer j such that $(j-1)(B_b+L_b)+B_b<b$, then

$$B_b\sim b^{3/4},\qquad L_b\sim b^{1/4},\qquad r_b\sim b^{1/4}. \tag{4.4.8}$$

Apply (4.4.4) on V_{bj}, we have $E(V_{bj})^2\leq KL_b$, and so

$$E\left(\frac{1}{\sqrt{b}}\sum_{j=1}^{r_b-1}V_{bj}\right)^2=\frac{1}{b}E\left(\sum_{j=1}^{r_b-1}V_{bj}\right)^2\leq\frac{1}{b}r_b\sum_{j=1}^{r_b-1}E(V_{bj})^2\leq\frac{r_b^2KL_b}{b}=O(b^{-1/4})\to 0$$

$$E\left(\frac{V_{br_b}}{\sqrt{b}}\right)^2=\frac{1}{b}E(V_{br_b}^2)\leq\frac{K(B_b+L_b)}{b}=O(b^{-1/4})\to 0$$

Hence

$$\frac{1}{\sqrt{b}}\sum_{j=1}^{r_b-1}V_{bj}=o_p(1),\qquad \frac{V_{br_b}}{\sqrt{b}}=o_p(1).$$

It remains to show $\frac{1}{\sqrt{b}}\sum_{j=1}^{r_b}U_{bj}\xrightarrow{d}N(0,\sigma^2)$. Let $U'_{bj}, 1\leq j\leq r_b$, be independent random variables having the same distribution as U_{bj}. The characteristic function of $b^{-1/2}\sum_{j=1}^{r_b}U_{bj}$ and of $b^{-1/2}\sum_{j=1}^{r_b}U'_{bj}$ differ by at most $16r_b\alpha_{L_b}=O(b^{-1})$. The asymptotic distribution of $b^{-1/2}\sum_{j=1}^{r_b}U_{bj}$ will be $N(0,\sigma^2)$ if that of $b^{-1/2}\sum_{j=1}^{r_b}U'_{bj}$ is. We will verify Lyapounov's condition for $\delta=2$. Applying (4.4.4) on U_{bj}, we have

$$b^{-1}\mathrm{Var}(\sum_{j=1}^{r_b}U'_{bj})=b^{-1}\sum_{j=1}^{r_b}\mathrm{Var}(U'_{bj})=b^{-1}\sum_{j=1}^{r_b}E(U_{bj}^2)=O(1),$$

Lyapounov's condition will be satisfied if $L(U,b)=\sum_{j=1}^{r_b}E\left(\frac{1}{\sqrt{b}}U'_{bj}\right)^4\to 0$. This is true because

$$L(U,b)=\frac{1}{b^2}\sum_{j=1}^{r_b}E(U'_{bj})^4\le\frac{1}{b^2}\sum_{j=1}^{r_b}KB_b^2=O(b^{-1/4})\to 0,$$

which used the following inequality

$$E(S_m^4)\le K_3m^2,\quad \text{for some finite } K_3 \text{ independent of } m. \tag{4.4.9}$$

To prove (4.4.9),

$$\begin{aligned}E(S_b^4) &= \sum_{j_1=1}^{b}\sum_{j_2=1}^{b}\sum_{j_3=1}^{b}\sum_{j_4=1}^{b}E(X_{j_1}X_{j_2}X_{j_3}X_{j_4})\\ &= \sum_{j=1}^{b}E(X_j^4)+6\sum_{j_1\ne j_2}^{b}E(X_{j_1}^2X_{j_2}^2)+6\sum_{j_1<j_2<j_3}^{b}E(X_{j_1}^2X_{j_2}X_{j_3})\\ &\quad +6\sum_{j_1<j_2<j_3}^{b}E(X_{j_1}X_{j_2}^2X_{j_3})+6\sum_{j_1<j_2<j_3}^{b}E(X_{j_1}X_{j_2}X_{j_3}^2)\\ &\quad +4\sum_{j_1<j_2}^{b}E(X_{j_1}^3X_{j_2})+4\sum_{j_1<j_2}^{b}E(X_{j_1}X_{j_2}^3)+4!\sum_{j_1<j_2<j_3<j_4}^{b}E(X_{j_1}X_{j_2}X_{j_3}X_{j_4})\end{aligned}$$

Similar to the proof of (4.4.1) we can show

$$\sum_{j_1<j_2}^{b}E(X_{j_1}X_{j_2}^3)\le Kb,\ \sum_{j_1<j_2<j_3}^{b}E(X_{j_1}X_{j_2}^2X_{j_3})\le Kb^2,\ \sum_{j_1<j_2<j_3}^{b}E(X_{j_1}X_{j_2}X_{j_3}^2)\le Kb^2$$

Therefore, (4.4.9) is true.

Thus we finished the proof of this lemma.

Lemma 4.4.2. *Assume $X_{ijk}, j=1,2,\ldots,$ is α-mixing with $\alpha_m=O(m^{-5})$ for all i,k. Set $\mathbf{X}_j=(X_{1j1},X_{1j2},\ldots,X_{1jn_1},\ldots,X_{ajn_a})'$, $j=1,\ldots,b$. Then $\mathbf{X}_j, j=1,2,\ldots,$ is α-mixing with $\alpha_m=O(m^{-5})$.*

Proof. We will first show that the difference $P(A\cap B)-P(A)P(B)$ is of order $O(m^{-5})$ for all rectangles $A=\prod_{i=1}^{a}\prod_{k=1}^{r_1}B_{ijk}\in\sigma(\{X_{ij_1k}, i=1,\ldots,a, j_1=1,\ldots,j, k=1,\ldots,r_1\})$,

$B = \prod_{i=1}^{a}\prod_{k=1}^{r_2} C_{i,j+m,k} \in \sigma(\{X_{ij_1k}, i = 1,\ldots,a, j_1 = j+m,\ldots,, k = 1,\ldots,r_2\})$, where r_1, r_2 are some finite numbers, $B_{ijk} \in \sigma(\{X_{ij_1k}, j_1 = 1,\ldots,j\})$ and $C_{i,j+m,k} \in \sigma(\{X_{ij_1k}, j_1 = j+m,\ldots,\})$. Write $\Delta_{ii'jj'kk'} = P\left(B_{ijk} \cap C_{i',j'+m,k'}\right) - P(B_{ijk})P(C_{i',j'+m,k'})$. Then $|\Delta_{ii'jj'kk'}| \leq \alpha_m$. Using the equality

$$\prod_{l=1}^{d}(a_l + b_l) = \prod_{l=1}^{d} a_l + \sum_{l=1}^{d}\prod_{r=l+1}^{d} a_r \prod_{s=1}^{l} b_s,$$

and the fact that probabilities are bounded by one, we have

$$\begin{aligned} & |P(A \cap B) - P(A)P(B)| \qquad (4.4.10) \\ = & \left|\prod_{i,k}\prod_{i',k'} P\left(B_{ijk} \cap C_{i',j'+m,k'}\right) - \prod_{i,k}\prod_{i',k'} P(B_{ijk})P(C_{i',j'+m,k'})\right| \\ = & \left|\prod_{i,k}\prod_{i',k'} [P(B_{ijk})P(C_{i',j'+m,k'}) + \Delta_{ii'jj'kk'}] - \prod_{i,k}\prod_{i',k'} P(B_{ijk})P(C_{i',j'+m,k'})\right| \\ \leq & \sum_{l=1}^{\infty} \alpha_m^l = \frac{\alpha_m}{1-\alpha_m} = O(m^{-5}). \end{aligned}$$

In order to show that $|P(A \cap B) - P(A)P(B)| = O(m^{-5})$ for all $A \in \sigma(\{X_{ij_1k}, i = 1,\ldots,a, j_1 = 1,\ldots,j, k = 1,\ldots,r_1\})$, $B \in \sigma(\{X_{ij_1k}, i = 1,\ldots,a, j_1 = j+m,\ldots,, k = 1,\ldots,r_2\})$, we need to use the good set arguments. Denote $\mathcal{F}_1$, $\mathcal{F}_2$ the class of "good sets" that have the desired property. By above proof, we know that $\mathcal{F}_1$, $\mathcal{F}_2$ contain all rectangles of form $A = \prod_{i=1}^{a}\prod_{k=1}^{r_1} B_{ijk}$, and $B = \prod_{i=1}^{a}\prod_{k=1}^{r_2} C_{i,j+m,k}$, respectively, where $B_{ijk} \in \sigma(\{X_{ij_1k}, j_1 = 1,\ldots,j\})$ and $C_{i,j+m,k} \in \sigma(\{X_{ij_1k}, j_1 = j+m,\ldots,\})$. Then we will show that $\mathcal{F}_1$ and $\mathcal{F}_2$ are closed under the formation of finite disjoint unions and thus $\mathcal{F}_1$ and $\mathcal{F}_2$ are fields. Let $A_l = \prod_{i=1}^{a}\prod_{k=1}^{r_1} B_{ijk}^{(l)} \in \sigma(\{X_{ij_1k}, i = 1,\ldots,a, j_1 = 1,\ldots,j, k = 1,\ldots,r_1\}), l = 1,\ldots,d_A$, be disjoint rectangles in $\mathcal{F}_1$, and $B_l = \prod_{i=1}^{a}\prod_{k=1}^{r_2} C_{i,j+m,k}^{(l)} \in \sigma(\{X_{ij_1k}, i = 1,\ldots,a, j_1 = j+m,\ldots,, k = 1,\ldots,r_2\}), l = 1,\ldots,d_B$, be disjoint rectangles in $\mathcal{F}_2$, where $B_{ijk}^{(l)} \in \sigma(\{X_{ij_1k}, j_1 = 1,\ldots,j\})$, $l = 1,\ldots,d_A$, and $C_{i,j+m,k}^{(l)} \in \sigma(\{X_{ij_1k}, j_1 = j+m,\ldots,\})$, $l = 1,\ldots,d_B$. Note that $\cup_{l=1}^{d_A} B_{ijk}^{(l)} \in \sigma(\{X_{ij_1k}, j_1 = 1,\ldots,j\})$,

$\cup_{l=1}^{d_B} C_{i,j+m,k}^{(l)} \in \sigma(\{X_{ij_1k}, j_1 = j+m,\ldots,\})$, and

$$P\left((\cup_{l=1}^{d_A} A_l) \cap (\cup_{l=1}^{d_B} B_l)\right) = \prod_{i,k}\prod_{i',k'} P\left((\cup_{l=1}^{d_A} B_{ijk}^{(l)}) \cap (\cup_{l=1}^{d_B} C_{i',j'+m,k'}^{(l)})\right),$$

$$P(\cup_{l=1}^{d_A} A_l) P(\cup_{l=1}^{d_B} B_l) = \prod_{i,k}\prod_{i',k'} P\left(\cup_{l=1}^{d_A} B_{ijk}^{(l)}\right) P\left(\cup_{l=1}^{d_B} C_{i',j'+m,k'}^{(l)}\right).$$

By a similar argument as (4.4.10) we get $|P(\cup_{l=1}^{d_A} A_l \cap \cup_{l=1}^{d_B} B_l) - P(\cup_{l=1}^{d_A} A_l) P(\cup_{l=1}^{d_B} B_l)| = O(m^{-5})$. Hence $\cup_{l=1}^{d_A} A_l \in \mathcal{F}_1$, $\cup_{l=1}^{d_B} B_l \in \mathcal{F}_2$ and so $\mathcal{F}_1$ is a field for all finite union of disjoint rectangles in $\mathcal{F}_2$ and vice versa.

Next, let $A_1, A_2, \ldots,$ be an increasing sequence of sets in $\mathcal{F}_1$, where A_l is a rectangle or finite union of disjoint rectangles, $l = 1, 2, \ldots$. Also let B be a rectangle or finite union of disjoint rectangles in $\mathcal{F}_2$. Above argument gives us $|P(A_l \cap B) - P(A_l)P(B)| = O(m^{-5})$. By continuity of probability measure, we have $|P(\cup_{l=1}^{d} A_l \cap B) - P(\cup_{l=1}^{d} A_l)P(B)| = \lim_{l\to\infty} |P(A_l \cap B) - P(A_l)P(B)| = O(m^{-5})$. So $\mathcal{F}_1$ is closed for monotone sequences and therefore $\mathcal{F}_1$ is a σ-field for all rectangles or finite union of disjoint rectangles in $\mathcal{F}_2$ and vice versa.

Now let $B_1, B_2, \ldots,$ be an increasing sequence of sets in $\mathcal{F}_2$, where B_l is a rectangle or finite union of disjoint rectangles, $l = 1, 2, \ldots$. Suppose A be an arbituary set in $\mathcal{F}_1$. By previous argument, we have $|P(A \cap B_l) - P(A)P(B_l)| = O(m^{-5})$. Again by continuity of probability measure, we get $|P\left(A \cap (\cup_{l=1}^{d} B_l)\right) - P(A)P(\cup_{l=1}^{d} B_l)| = \lim_{l\to\infty} |P(A \cap B_l) - P(A)P(B_l)| = O(m^{-5})$. So $\mathcal{F}_2$ is closed for monotone sequences and therefore $\mathcal{F}_2$ is a σ-field for all sets in σ-field $\mathcal{F}_1$ and vice versa. Hence $\mathcal{F}_1 = \sigma(\mathbf{X}_{j_1}, j_1 = 1, 2, \ldots, j)$, $\mathcal{F}_2 = \sigma(\mathbf{X}_{j_1}, j_1 = j+m, \ldots,)$, and we finish the proof.

Lemma 4.4.3. *Assume $X_{ijk}, j = 1, 2, \ldots,$ is α-mixing with $\alpha_m = O\left(m^{-5}\right)$ for all i, k, and $\limsup_j E(u_{ijk}^{32}) < \infty$, then as $b \to \infty$,*

1. *if n_i remain fixed,*

   (a) $\sqrt{b}(MSE - P_{MSE}(\mathbf{u})) = o_p(1)$, where $P_{MSE}(\mathbf{u})$ is defined in (4.4.12).

(b) $MSE \xrightarrow{p} \sigma^2$, *where*

$$\sigma^2 = \lim_{b\to\infty} \frac{1}{ab} \sum_{i,j,k} \frac{\sigma_{ij}^2}{n_i^2}$$

provided that the limit exists.

2. *if* $n(a) \to \infty$ *as* b *goes to* ∞, *then*

   (a) $n(a)\sqrt{b}[MSE - P_{MSE}(\mathbf{u})] = o_p(1)$.

   (b) $n(a)MSE \xrightarrow{p} \sigma_*^2$, *where*

$$\sigma_*^2 = \lim_{b\to\infty} \frac{1}{ab} \sum_{i,j,k} \frac{n^2(a)}{n_i^2} \sigma_{ij}^2$$

provided that the limit exists.

Proof.

Note that $MSE = P_{MSE} + D_1(\mathbf{u}) + D_2(\mathbf{u})$, where

$$D_1(\mathbf{u}) = -\frac{1}{ab(b-1)} \sum_{i=1}^{a} \sum_{k=1}^{n_i} \sum_{j\neq j'}^{b} \frac{u_{ijk}u_{ij'k}}{n_i(n_i-1)}, \quad D_2(\mathbf{u}) = \frac{1}{ab(b-1)} \sum_{i=1}^{a} \sum_{j\neq j'}^{b} \frac{\overline{u}_{ij.}\overline{u}_{ij'.}}{n_i - 1} \quad (4.4.11)$$

To show part (a) in both cases, we only need to show $n(a)D_1(\mathbf{u}) = o_p(b^{-1/2})$ and $n(a)D_2(\mathbf{u}) = o_p(b^{-1/2})$ whether n_i are fixed or not.

$$\begin{aligned} E(n^2(a)D_1^2(\mathbf{u})) &= \frac{n^2(a)}{a^2b^2(b-1)^2} \sum_{i=1}^{a} \sum_{k=1}^{n_i} \sum_{j\neq j'}^{b} \sum_{i_1=1}^{a} \sum_{k_1=1}^{n_{i_1}} \sum_{j_1\neq j_2}^{b} \frac{E(u_{ijk}u_{ij'k}u_{i_1j_1k_1}u_{i_1j_2k_1})}{n_i(n_i-1)n_{i_1}(n_{i_1}-1)} \\ &= \frac{n^2(a)}{a^2b^2(b-1)^2} \sum_{i=1}^{a} \sum_{k=1}^{n_i} \sum_{j\neq j'}^{b} \sum_{j_1\neq j_2}^{b} \frac{E(u_{ijk}u_{ij'k}u_{ij_1k}u_{ij_2k})}{n_i(n_i-1)n_i(n_i-1)} \\ &\quad + \frac{n^2(a)}{a^2b^2(b-1)^2} \sum_{i=1}^{a} \sum_{i_1=1}^{a} \sum_{k\neq k_1}^{n_i} \sum_{j\neq j'}^{b} \sum_{j_1\neq j_2}^{b} \frac{E(u_{ijk}u_{ij'k})E(u_{i_1j_1k_1}u_{i_1j_2k_1})}{n_i(n_i-1)n_{i_1}(n_{i_1}-1)} \end{aligned}$$

By Lemma 4.4.1,

$$\sum_{j\neq j'}^{b} E(u_{ijk}u_{ij'k}) = O(b) \text{ and } \sum_{j\neq j'}^{b} \sum_{j_1\neq j_2}^{b} E(u_{ijk}u_{ij'k}u_{ij_1k}u_{ij_2k}) = O(b^2).$$

So $E(n^2(a)D_1^2(\mathbf{u})) = O(b^{-2}) \to 0$.

Similarly,

$$E(n^2(a)D_2^2(\mathbf{u})) = \frac{n^2(a)}{a^2b^2(b-1)^2}\sum_{i=1}^{a}\sum_{j\neq j'}^{b}\sum_{i_1=1}^{a}\sum_{j_1\neq j_2}^{b}\sum_{k=1}^{n_i}\sum_{k'=1}^{n_i}\sum_{k_1=1}^{n_{i_1}}\sum_{k_2=1}^{n_{i_1}}\frac{E(u_{ijk}u_{ij'k'}u_{i_1j_1k_1}u_{i_1j_2k_2})}{n_i^2n_{i_1}^2(n_i-1)(n_{i_1}-1)} = O(b^{-2})$$

Therefore, $n(a)(MSE - P_{MSE}) = n(a)D_1(\mathbf{u}) + n(a)D_2(\mathbf{u}) = o_p(b^{-1/2})$.

To show part (b) when $n_i \to \infty$ as $b \to \infty$, by the result of part (a), it suffices to show that $n(a)P_{MSE} \xrightarrow{p} \sigma_*^2$, or equivalently,

$$n(a)P_{MSE} - \frac{1}{ab}\sum_{i,j,k}\frac{n(a)}{n_i^2}\sigma_{ij}^2 = \frac{n(a)}{ab}\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n_i}\frac{u_{ijk}^2-\sigma_{ij}^2}{n_i^2} - \frac{n(a)}{ab}\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k\neq k'}^{n_i}\frac{u_{ijk}u_{ijk'}}{n_i^2(n_i-1)} = o_p(1).$$

To show this, we note that

$$E\left(\frac{n(a)}{a}\sum_{k}\frac{u_{ijk}^2-\sigma_{ij}^2}{n_i^2}\right)^{16} \leq \frac{n(a)^{16}n_i^{15}}{a^{16}}\sum_{k}E\left(\frac{u_{ijk}^2-\sigma_{ij}^2}{n_i^2}\right)^{16} \leq 2^{15}[E(u_{ij1}^{32})+\sigma_{ij}^{32}]/a^{16} < \infty$$

Similarly, $E\left(\frac{n(a)}{a}\sum_{k\neq k'}^{n_i}\frac{u_{ijk}u_{ijk'}}{n_i^2(n_i-1)}\right)^{16} < \infty$. By Lemmas 4.4.2 and 4.4.1, we have

$$\frac{n(a)}{a\sqrt{b}}\sum_{j,k}\frac{u_{ijk}^2-\sigma_{ij}^2}{n_i^2} = O_p(1), \quad \frac{n(a)}{a\sqrt{b}}\sum_{j=1}^{b}\sum_{k\neq k'}^{n_i}\frac{u_{ijk}u_{ijk'}}{n_i^2(n_i-1)} = O_p(1).$$

Hence,

$$n(a)P_{MSE} - \frac{1}{ab}\sum_{i,j,k}\frac{n(a)}{n_i^2}\sigma_{ij}^2 = O_p(b^{-1/2}) = o_p(1).$$

When n_i are fixed,the proof can be obtained by simply treating $n(a)$ as a constant in above proof. Thus we finished part (b) of the Lemma for both cases.

Lemma 4.4.4. *Assume $X_{ijk}, j = 1,2,\ldots,$ is α-mixing with $\alpha_m = O\left(m^{-5}\right)$ for all i,k, and* $\limsup_j E(u_{ijk}^{32}) < \infty$. *Let $P_\beta(\mathbf{u})$ and $P_\gamma(\mathbf{u})$ be defined in (4.4.13). Then as $b \to \infty$ while a remains fixed,* *(a) if n_i remain fixed,*

$$\text{under } H_0(\beta),\ \sqrt{b}(MS\beta - P_\beta(\mathbf{u})) \xrightarrow{p} 0; \quad \text{under } H_0(\gamma),\ \sqrt{b}(MS\gamma - P_\gamma(\mathbf{u})) \xrightarrow{p} 0$$

(b) if $n(a) \to \infty$ as b go to ∞,

$$\text{under } H_0(\beta),\ n(a)\sqrt{b}(MS\beta - P_\beta(\mathbf{u})) \xrightarrow{p} 0$$

$$\text{under } H_0(\gamma),\ n(a)\sqrt{b}(MS\gamma - P_\gamma(\mathbf{u})) \xrightarrow{p} 0.$$

Proof.

Note that $MS\beta - P_\beta(\mathbf{u}) = -D_3(\mathbf{u})$, $\quad MS\gamma - P_\gamma(\mathbf{u}) = \dfrac{D_3(\mathbf{u}) - D_4(\mathbf{u})}{a-1}$, where

$$D_3(\mathbf{u}) = \frac{a}{b(b-1)} \sum_{j \neq j'}^{b} \widetilde{u}_{.j.}\widetilde{u}_{.j'.}, \quad D_4(\mathbf{u}) = \frac{1}{b(b-1)} \sum_{i=1}^{a} \sum_{j \neq j'}^{b} \overline{u}_{ij.}\overline{u}_{ij'.}$$

When $n(a) \to \infty$ as $b \to \infty$, similar to the proof of $n(a)D_2(\mathbf{u}) = o_p(b^{-1/2})$, we can show that $E(n^2(a)D_4^2(\mathbf{u})) = O(1/b^2)$. So we only need to show that $n(a)\sqrt{b}D_3(\mathbf{u}) = o_p(1)$. $D_3(\mathbf{u}) = D_{31}(\mathbf{u}) + D_{32}(\mathbf{u}) + D_{33}(\mathbf{u})$, where

$$\begin{aligned}
D_{31}(\mathbf{u}) &= \frac{1}{ab(b-1)} \sum_{i \neq i'}^{a} \sum_{j \neq j'}^{b} \overline{u}_{ij.}\overline{u}_{i'j'.} \\
D_{32}(\mathbf{u}) &= \frac{1}{ab(b-1)} \sum_{i=1}^{a} \sum_{j \neq j'}^{b} \sum_{k \neq k'}^{n_i} \frac{u_{ijk}u_{ij'k'}}{n_i^2} \\
D_{33}(\mathbf{u}) &= \frac{1}{ab(b-1)} \sum_{i=1}^{a} \sum_{j \neq j'}^{b} \sum_{k=1}^{n_i} \frac{u_{ijk}u_{ij'k}}{n_i^2}
\end{aligned}$$

$n(a)\sqrt{b}D_{33}(\mathbf{u}) = o_p(1)$ follows the same argument as $n(a)\sqrt{b}D_1(\mathbf{u}) = o_p(1)$ which is explained in detail in the proof of Lemma 4.4.3. By independence of the error terms from different subjects or groups,

$$\begin{aligned} E(n^2(a)bD_{32}^2(\mathbf{u})) &= \frac{n^2(a)}{a^2b(b-1)^2}\sum_{i=1}^{a}\sum_{j\neq j'}^{b}\sum_{k\neq k'}^{n_i}\sum_{i_1=1}^{a}\sum_{j_1\neq j_2}^{b}\sum_{k_1\neq k_2}^{n_{i_1}}\frac{E(u_{ijk}u_{ij'k'}u_{i_1j_1k_1}u_{i_1j_2k_2})}{n_i^2n_{i_1}^2} \\ &= \frac{n^2(a)}{a^2b(b-1)^2}\sum_{i=1}^{a}\sum_{j\neq j'}^{b}\sum_{k\neq k'}^{n_i}\sum_{j_1\neq j_2}^{b}\sum_{k_1\neq k_2}^{n_i}\frac{E(u_{ijk}u_{ij'k'}u_{ij_1k_1}u_{ij_2k_2})}{n_i^4} \\ &= \frac{4n^2(a)}{a^2b(b-1)^2}\sum_{i=1}^{a}\sum_{j\neq j'}^{b}\sum_{k\neq k'}^{n_i}\sum_{j_1\neq j_2}^{b}\frac{E(u_{ijk}u_{ij_1k})E(u_{ij'k'}u_{ij_2k'})}{n_i^4} = O(b^{-1}), \end{aligned}$$

where the last step used (4.4.1) in Lemma 4.4.1. So $n(a)\sqrt{b}D_{32}(\mathbf{u}) = o_p(1)$. $n(a)\sqrt{b}D_{31}(\mathbf{u}) = o_p(1)$ can be shown similarly. Then we finished the proof of this lemma.

Proof of Theorem 4.2.1 The proof of results under $H_0(\beta)$ and $H_0(\beta)$ are stated first while that under $H_0(\phi)$ is given later. By Lemma 4.4.3 stated below, we only need to consider the asymptotic distribution of $\sqrt{b}n(a)(MS\beta - MSE)$ and $\sqrt{b}n(a)(MS\gamma - MSE)$.

The asymptotic distribution of $\sqrt{b}n(a)(MS\beta - MSE)$ and $\sqrt{b}n(a)(MS\gamma - MSE)$ is obtained by the projection method. $MS\beta$, $MS\gamma$, and MSE will be projected onto the class of random variables of the form $\sum_{j=1}^{b} g_j(\mathbf{u}_j)$, where $\mathbf{u}_j = (u_{1j1}, \ldots, u_{1jn_1}, \ldots, u_{ajn_a})'$ and g_j are measurable with $Eg_j^2(\mathbf{u}_j) < \infty$. We note that the projections are only formal, in the sense that they are computed ignoring the dependence of the data. However, it will be shown that the projections are asymptotically equivalent to the corresponding statistics and thus can be used as usual for deriving the needed asymptotic distributions.

Note that the projection of MSE is P_{MSE} , which is verified by part a in Lemma 4.4.3, where

$$P_{MSE}(\mathbf{u}) = \frac{1}{ab}\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n_i}\frac{u_{ijk}^2}{n_i^2} - \frac{1}{ab}\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k\neq k'}^{n_i}\frac{u_{ijk}u_{ijk'}}{n_i^2(n_i-1)}. \tag{4.4.12}$$

The projections of $MS\beta$ under $H_0(\beta)$ and $MS\gamma$ under $H_0(\gamma)$ are given by

$$P_\beta(\mathbf{u}) = \frac{a}{b}\sum_{j=1}^{b}\tilde{u}_{.j.}^2, \qquad P_\gamma(\mathbf{u}) = \frac{1}{(a-1)b}\sum_{i=1}^{a}\sum_{j=1}^{b}\bar{u}_{ij.}^2 - \frac{a}{b(a-1)}\sum_{j=1}^{b}\tilde{u}_{.j.}^2 \tag{4.4.13}$$

respectively, which is justified by Lemma 4.4.4.

By Slutsky's Theorem, $n(a)\sqrt{b}(MS\beta - MSE)$ and $n(a)\sqrt{b}(MS\gamma - MSE)$ have same asymptotic distribution as their projections $n(a)\sqrt{b}(P_\beta(\mathbf{u}) - P_{MSE})$ and $n(a)\sqrt{b}(P_\gamma(\mathbf{u}) - P_{MSE})$, under $H_0(\beta)$ and $H_0(\gamma)$, respectively, when n_i go to ∞ as $b \to \infty$. Let

$$P_1 = \frac{1}{ab}\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k\neq k'}^{n_i}\frac{u_{ijk}u_{ijk'}}{n_i(n_i-1)}, \quad P_2 = \frac{1}{ab}\sum_{i\neq i'}^{a}\sum_{j=1}^{b}\bar{u}_{ij.}\bar{u}_{i'j.},$$

then $P_\beta(\mathbf{u}) - P_{MSE} = P_1 + P_2$ and $P_\gamma(\mathbf{u}) - P_{MSE} = P_1 - P_2/(a-1)$. We also have that under $H_0(\phi)$, $MS\phi - MSu_\phi = P_1 - P_2/(a-1)$.

Write $n(a)\sqrt{b}(P_\beta(\mathbf{u}) - P_{MSE}) = b^{-1/2}\sum_{j=1}^{b} W_j$, where

$$W_j = \frac{n(a)}{a}\sum_{i=1}^{a}\sum_{k\neq k'}^{n_i}\frac{u_{ijk}u_{ijk'}}{n_i(n_i-1)} + \frac{n(a)}{a}\sum_{i\neq i'}^{a}\bar{e}_{ij.}\bar{e}_{i'j.}.$$

By Lemma 4.4.2, $W_1, W_2, \cdots$, is α-mixing with $\alpha_m = O(m^{-5})$ and $E(W_j) = 0$. We can apply Lemma 4.4.1 if we can prove $\limsup_j E(W_j^{16}) < \infty$. Using the inequality

$$\left|\sum_{i=1}^{m} z_i\right|^p \leq m^{p-1}\sum_{i=1}^{m}|z_i|^p,\ m \geq 1,\ p \geq 1, \tag{4.4.14}$$

which for $p > 1$ follows from Hölder's inequality, we have

$$\begin{aligned}
E(W_j^{16}) &= \frac{n^{16}(a)}{a^{16}} E\left(\sum_{i=1}^{a}\sum_{k\neq k'}^{n_i}\frac{u_{ijk}u_{ijk'}}{n_i(n_i-1)} + \frac{n(a)}{a}\sum_{i\neq i'}^{a}\bar{e}_{ij.}\bar{e}_{i'j.}\right)^{16} \\
&\leq \frac{n^{16}(a)}{a^{16}} 2^{15}\left[E\left(\sum_{i=1}^{a}\sum_{k\neq k'}^{n_i}\frac{u_{ijk}u_{ijk'}}{n_i(n_i-1)}\right)^{16} + E\left(\sum_{i\neq i'}^{a}\bar{e}_{ij.}\bar{e}_{i'j.}\right)^{16}\right] \\
&\leq \frac{n^{16}(a)}{a^{16}} 2^{15}\left[a^{15}\sum_{i=1}^{a}\frac{E\left(\sum_{k\neq k'}^{n_i}u_{ijk}u_{ijk'}\right)^{16}}{n_i^{16}(n_i-1)^{16}} + a^{30}\sum_{i\neq i'}^{a}E\left(\bar{e}_{ij.}\right)^{16}E\left(\bar{e}_{i'j.}\right)^{16}\right] \\
&\leq \frac{n^{16}(a)}{a^{16}} 2^{15}\left[a^{15}\sum_{i=1}^{a}\frac{E\left(\left(\sum_{k=1}^{n_i}u_{ijk}\right)^{2}\right)^{16}}{n_i^{16}(n_i-1)^{16}} + a^{30}\sum_{i\neq i'}^{a}E\left(\bar{e}_{ij.}\right)^{16}E\left(\bar{e}_{i'j.}\right)^{16}\right] \\
&\leq \frac{n^{16}(a)}{a^{16}} 2^{15}\left[a^{15}\sum_{i=1}^{a}\frac{2^{15}\left[E\left(\sum_{k=1}^{n_i}u_{ijk}\right)^{32}\right]}{n_i^{16}(n_i-1)^{16}} + a^{30}\sum_{i\neq i'}^{a}E\left(\bar{e}_{ij.}\right)^{16}E\left(\bar{e}_{i'j.}\right)^{16}\right] \\
&\leq \frac{n^{16}(a)}{a^{16}} 2^{15}\left[(2a)^{15}\sum_{i=1}^{a}\frac{E\left(\sum_{k=1}^{n_i}u_{ijk}\right)^{32}}{n_i^{16}(n_i-1)^{16}} + a^{30}\sum_{i\neq i'}^{a}E\left(\bar{e}_{ij.}\right)^{16}E\left(\bar{e}_{i'j.}\right)^{16}\right].
\end{aligned}$$

Apply inequality: for any $P \geq 2$, there exists a finite positive constant A_P (depending only on P), such that for any iid random variables $Z_1, \cdots, Z_m$ with $E(Z_i) = 0$,

$$E|Z_1 + \cdots + Z_m|^P \leq A_P m^{P/2} E(|Z_1|^P). \tag{4.4.15}$$

This inequality follows if we first use the Marcinkiewich-Zygmund inequality

$$E|Z_1 + \cdots + Z_m|^P \leq A_P E(Z_1^2 + \ldots + Z_n^2)^{P/2}, P \geq 1,$$

and then apply (4.4.14) to the last sum (see Chow and Teicher, 1997, pp. 386-387). Note that $u_{ijk}, k = 1, \cdots, n_i$ are iid, so

$$E\left(\sum_{k=1}^{n_i}u_{ijk}\right)^{32} \leq A_{32}n_i^{16}E(u_{ij1}^{32}); \qquad E\left(\bar{e}_{ij.}\right)^{16} \leq \frac{A_{16}}{n_i^{8}}E(u_{ij1}^{16}).$$

Therefore

$$E(W_j^{16}) \le \frac{2^{15}}{a}\left[2^{15}\sum_{i=1}^{a} A_{32}E(u_{ij1}^{32}) + \frac{\sum_{k=1}^{n_i} E\left(u_{ijk}^{32}\right)}{n_i} + a^{15}\sum_{i\neq i'}^{a} A_{16}^2 E(u_{ij1}^{16})E(u_{i'j1}^{16})\right],$$

and $\limsup_j E(W_j^{16}) < \infty$ by assumption $\limsup_j E(u_{ijk}^{32}) < \infty$.

$$\begin{aligned}
& \frac{1}{b}E\left(\sum_{j=1}^{b} W_j\right)^2 = \frac{1}{b}\mathrm{Var}\left(\sum_{j=1}^{b} W_j\right) \\
= & \frac{n^2(a)}{a^2 b}\sum_{i=1}^{a}\mathrm{Var}\left(\sum_{k\neq k'}^{n_i}\sum_{j=1}^{b}\frac{u_{ijk}u_{ijk'}}{n_i(n_i-1)}\right) + \frac{n^2(a)}{a^2 b}\mathrm{Var}\left(\sum_{i\neq i'}^{a}\sum_{j=1}^{b}\bar{u}_{ij.}\bar{u}_{i'j.}\right) \\
= & \frac{n^2(a)}{a^2 b}\sum_{j=1}^{b}\sum_{j'=1}^{b}\left[\sum_{i=1}^{a}\sum_{k\neq k'}^{n_i}\sum_{k_1\neq k_2}^{n_i}\frac{E(u_{ijk}u_{ijk'}u_{ij'k_1}u_{ij'k_2})}{n_i^2(n_i-1)^2} + \sum_{i\neq i'}^{a}\sum_{i_1\neq i_2}^{a} E\left(\bar{u}_{ij.}\bar{u}_{i'j.}\bar{u}_{i_1j'.}\bar{u}_{i_2j'.}\right)\right] \\
= & \frac{2n^2(a)}{a^2 b}\sum_{j=1}^{b}\sum_{j'=1}^{b}\left[\sum_{i=1}^{a}\sum_{k\neq k'}^{n_i}\frac{E(u_{ijk}u_{ij'k})E(u_{ijk'}u_{ij'k'})}{n_i^2(n_i-1)^2} + \sum_{i\neq i'}^{a} E\left(\bar{u}_{ij.}\bar{u}_{ij'.}\right)E\left(\bar{u}_{i'j.}\bar{u}_{i'j'.}\right)\right] \\
= & \frac{2n^2(a)}{a^2 b}\sum_{j=1}^{b}\sum_{j'=1}^{b}\left[\sum_{i=1}^{a}\frac{\sigma_{ijj'}^2}{n_i(n_i-1)} + \sum_{i\neq i'}^{a}\frac{\sigma_{ijj'}}{n_i}\frac{\sigma_{i'jj'}}{n_{i'}}\right] \to \tilde{\tau}_\beta^2
\end{aligned}$$

Apply Lemma 4.4.1, we finish the proof.

Proof of Proposition 4.2.2 First we will show $\widehat{\zeta}_1 - \zeta_1 \xrightarrow{p} 0$ as $b \to \infty$. By Lemma 3 of Billingsley (1986) p. 377, and note that $\sum_{j=1}^{\infty}\alpha_j < \infty$, so

$$\begin{aligned}
\zeta_1 - E(\widehat{\zeta}_1) &= \frac{2}{a^2 b}\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{|j'-j|>b^h}\frac{\sigma_{ijj'}^2}{n_i(n_i-1)} \\
&\le \frac{2}{a^2 b}\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{|j'-j|>b^h}\frac{\alpha_{|j'-j|}}{n_i(n_i-1)} = \frac{2}{a^2}\sum_{i=1}^{a}\sum_{j_2>b^h}^{b}\frac{\alpha_{j_2}}{n_i(n_i-1)} \to 0.
\end{aligned}$$

$$\mathrm{Var}(\widehat{\zeta}_1) = \frac{4}{a^4 b^2}\sum_{i=1}^{a} n_i^{-2}(n_i-1)^{-2}\sum_{j=1}^{b}\sum_{|j'-j|<b^h}\sum_{j_1=1}^{b}\sum_{|j_1'-j_1|<b^h}\left[E(\widehat{\sigma_{ijj'}^2}\widehat{\sigma_{ij_1j_1'}^2}) - \sigma_{ijj'}^2\sigma_{ij_1j_1'}^2\right]$$

Denote $g_{ijj'j_1j_1'} = E[(X_{ijk}-\mu_{ij})(X_{ij'k}-\mu_{ij'})(X_{ij_1k}-\mu_{ij_1})(X_{ij_1'k}-\mu_{ij_1'})]$. Through some calculation, it can be shown that

$$E(\widehat{\sigma^2_{ijj'}}\widehat{\sigma^2_{ij_1j_1'}}) - \sigma^2_{ijj'}\sigma^2_{ij_1j_1'} = \frac{(n_i-4)(n_i-5)\sigma_{ijj'}\sigma_{ij_1j_1'}}{(n_i-1)(n_i-2)(n_i-3)}\left[g_{ijj'j_1j_1'} - \sigma_{ijj'}\sigma_{ij_1j_1'}\right] + \frac{\left[9g_{ijj'j_1j_1'} + 33\sigma_{ijj'}\sigma_{ij_1j_1'}\right]\left[g_{ijj'j_1j_1'} - \sigma_{ijj'}\sigma_{ij_1j_1'}\right]}{n_i(n_i-1)(n_i-2)(n_i-3)}$$

To show $\mathrm{Var}(\widehat{\zeta}_1) \to 0$, It suffices to show that

$$\sum_{j=1}^{b}\sum_{j'=1}^{b}\sum_{j_1=1}^{b}\sum_{j_1'=1}^{b}\sigma_{ijj'}\sigma_{ij_1j_1'}\left[g_{ijj'j_1j_1'} - \sigma_{ijj'}\sigma_{ij_1j_1'}\right] = o(b^2) \tag{4.4.16}$$

$$\sum_{j,j_1}^{b}\sum_{|j'-j|<b^h}\sum_{|j_1'-j_1|<b^h}\left[9g_{ijj'j_1j_1'} + 33\sigma_{ijj'}\sigma_{ij_1j_1'}\right]\left[g_{ijj'j_1j_1'} - \sigma_{ijj'}\sigma_{ij_1j_1'}\right] = o(b^2) \tag{4.4.17}$$

By Lemma 3 of Billingsley p. 377, for some finite positive K, we have the following facts because $E[(X_{ijk}-\mu_{ij})^8]$ is bounded: if $j < j' < j_1 < j_1'$,

$$|g_{ijj'j_1j_1'} - \sigma_{ijj'}\sigma_{ij_1j_1'}| \le K\alpha^{1/2}_{j_1-j'}$$

$$|9g_{ijj'j_1j_1'} + 33\sigma_{ijj'}\sigma_{ij_1j_1'}| \le K\alpha^{1/2}_{j'-j} + 33|\sigma_{ijj'}\sigma_{ij_1j_1'}|$$

hint: apply Lemma 3 of Billingsley p. 377 for $Y = X_{ijk} - \mu_{ij}$ and

$$Z = (X_{ij'k}-\mu_{ij'})(X_{ij_1k}-\mu_{ij_1})(X_{ij_1'k}-\mu_{ij_1'})$$

if $j < j_1 < j' < j_1'$,

$$|g_{ijj'j_1j_1'} - \sigma_{ijj'}\sigma_{ij_1j_1'}| \le |g_{ijj'j_1j_1'}| + |\sigma_{ijj'}\sigma_{ij_1j_1'}| \le |\sigma_{ijj_1}\sigma_{ij'j_1'}| + K\alpha^{1/2}_{j'-j_1} + |\sigma_{ijj'}\sigma_{ij_1j_1'}|$$

$$|9g_{ijj'j_1j_1'} + 33\sigma_{ijj'}\sigma_{ij_1j_1'}| \le 9K\alpha^{1/2}_{j_1-j} + 33|\sigma_{ijj'};$$

if $j < j_1 < j_1' < j'$,

$$|g_{ijj'j_1j_1'} - \sigma_{ijj'}\sigma_{ij_1j_1'}| \le |g_{ijj'j_1j_1'}| + |\sigma_{ijj'}\sigma_{ij_1j_1'}| \le |\sigma_{ijj_1}\sigma_{ij_1'j'}| + K\alpha^{1/2}_{j_1'-j_1} + |\sigma_{ijj'}\sigma_{ij_1j_1'}|$$

$$|9g_{ijj'j_1j_1'} + 33\sigma_{ijj'}\sigma_{ij_1j_1'}| \le 9K\alpha^{1/2}_{j_1-j} + 33|\sigma_{ijj'}\sigma_{ij_1j_1'}|.$$

The other situations when j, j_1, j_1' and j' are all different are similar to one of above three cases. We will show that the summation on the left hand side of (4.4.16) in one of above cases is $O(b)$ and the proofs for the other cases are similar and are omitted. Apply Lemma 3 of Billingsley p. 377 again, we have $|\sigma_{ijj'}| \le K\alpha_{j'-j}^{1/2}$ and so

$$\sum_{j<j'<j_1<j_1'}^{b} |g_{ijj'j_1j_1'} - \sigma_{ijj'}\sigma_{ij_1j_1'}||\sigma_{ijj'}\sigma_{ij_1j_1'}| \le \sum_{j<j'<j_1<j_1'}^{b} K\alpha_{j_1-j'}^{1/2} K\alpha_{j'-j}^{1/2} K\alpha_{j_1'-j_1}^{1/2}$$

$$= \sum_{j+j_2+j_3+j_4<b} K^3\alpha_{j_2}^{1/2}\alpha_{j_3}^{1/2}\alpha_{j_4}^{1/2} \le \sum_{j=1}^{b}\left(\sum_{j_2=1}^{b}\alpha_{j_2}^{1/2}\right)^3 = O(b) \text{ since } \sum_{j_3=1}^{\infty}\alpha_{j_3}^{1/2} < \infty$$

If $j < j' = j_1 < j_1'$, use the fact that $|g_{ijj'j_1j_1'} - \sigma_{ijj'}\sigma_{ij_1j_1'}|$ is bounded and

$$\sum_{j<j'<j_1'}^{b} |g_{ijj'j'j_1'} - \sigma_{ijj'}\sigma_{ij'j_1'}||\sigma_{ijj'}\sigma_{ij'j_1'}| \le K \sum_{j<j'<j_1'}^{b} \alpha_{j'-j}^{1/2}\alpha_{j_1'-j'}^{1/2}$$

$$= K \sum_{j+j_2+j_3<b}^{b} \alpha_{j_2}^{1/2}\alpha_{j_3}^{1/2} \le Kb\sum_{j_2=1}^{b}\alpha_{j_2}^{1/2}\sum_{j_3=1}^{b}\alpha_{j_3}^{1/2} = O(b)$$

if $j = j' < j_1 < j_1'$, use $|g_{ijj'j_1j_1'} - \sigma_{ijj'}\sigma_{ij_1j_1'}| \le K\alpha_{j_1-j'}^{1/2}$, and so

$$\sum_{j<j_1<j_1'} |g_{ijjj_1j_1'} - \sigma_{ijj}\sigma_{ij_1j_1'}||\sigma_{ijj}\sigma_{ij_1j_1'}| \le K \sum_{j<j_1<j_1'} \alpha_{j_1-j}^{1/2}\sigma_{ijj}K\alpha_{j_1'-j_1}^{1/2} = O(b).$$

The other situations when the number of different elements in $\{j, j', j_1, j_1'\}$ is three are similar to one of the above two cases. The rest of the summations are easier to prove to be $O(b)$ and we omit them.

For (4.4.17), we have similar cases like above and others like the following:

$$\sum_{j,j_1}\sum_{|j'-j|<b^h}\sum_{|j_1'-j_1|<b^h} I(j<j'<j_1<j_1')\,\alpha_{j_1-j'}^{1/2}\alpha_{j'-j}^{1/2} \le \sum_{j=1}^{b}\sum_{j'=j+1}^{j+b^h}\sum_{j_1=j'+1}^{b}\sum_{j_1'=j_1+1}^{j_1+b^h}\alpha_{j_1-j'}^{1/2}\alpha_{j'-j}^{1/2}$$

$$= b^{1+h}\sum_{j_2}^{b}\sum_{j_4=1}^{b^h}\alpha_{j_2}^{1/2}\alpha_{j_4}^{1/2} = O(b^{1+h}) = o(b^2) \qquad \text{since } \sum_{j=1}^{\infty}\alpha_j^{1/2} < \infty$$

Next, we will show $\widehat{\zeta}_2 \xrightarrow{p} 0$. $E(\widehat{\zeta}_2) = \zeta_2$.

$$
\begin{aligned}
&\mathrm{Var}(\widehat{\zeta}_2) \\
= \;& \frac{4}{a^4b^2}\sum_{i\neq i'}^{a}\sum_{i_1\neq i_1'}^{a}\sum_{j=1}^{b}\sum_{|j'-j|<b^h}\sum_{j_1=1}^{b}\sum_{|j_1'-j_1|<b^h}\frac{\left[E(\widehat{\sigma}_{ijj'}\widehat{\sigma}_{i'jj'}\widehat{\sigma}_{i_1j_1j_1'}\widehat{\sigma}_{i_1'j_1j_1'}) - \sigma_{ijj'}\sigma_{i'jj'}\sigma_{i_1j_1j_1'}\sigma_{i_1'j_1j_1'}\right]}{n_in_{i'}n_{i_1}n_{i_1'}} \\
= \;& \frac{8}{a^4b^2}\sum_{i\neq i'\neq i_1}^{a}\sum_{j=1}^{b}\sum_{|j'-j|<b^h}\sum_{j_1=1}^{b}\sum_{|j_1'-j_1|<b^h}\frac{\left[E(\widehat{\sigma}_{ijj'}\widehat{\sigma}_{ij_1j_1'}) - \sigma_{ijj'}\sigma_{ij_1j_1'}\right]\sigma_{i'jj'}\sigma_{i_1j_1j_1'}}{n_i^2n_{i'}n_{i_1}} \\
&+\frac{8}{a^4b^2}\sum_{i\neq i'}^{a}\sum_{j=1}^{b}\sum_{|j'-j|<b^h}\sum_{j_1=1}^{b}\sum_{|j_1'-j_1|<b^h}\frac{\left[E(\widehat{\sigma}_{ijj'}\widehat{\sigma}_{ij_1j_1'})E(\widehat{\sigma}_{i'jj'}\widehat{\sigma}_{i'j_1j_1'}) - \sigma_{ijj'}\sigma_{i'jj'}\sigma_{ij_1j_1'}\sigma_{i'j_1j_1'}\right]}{n_i^2n_{i'}^2}.
\end{aligned}
$$

Let $g_2(ijkj'k'j_1k_1j_1'k_1') = (X_{ijk}-\mu_{ij})(X_{ij'k'}-\mu_{ij'})(X_{ij_1k_1}-\mu_{ij_1})(X_{ij_1'k_1'}-\mu_{ij_1'})$. Seperate different cases according to the number of different elements in $\{k_1,k_2,k_3,k_4\}$, we can write

$$
\begin{aligned}
&E(\widehat{\sigma}_{ijj'}\widehat{\sigma}_{ij_1j_1'}) - \sigma_{ijj'}\sigma_{ij_1j_1'} \\
= \;& \frac{1}{(n_i-1)^2}\left[\sum_{k_1,k_2}^{n_i}E(g_2(ijk_1j'k_1j_1k_2j_1'k_2)) - \frac{1}{n_i}\sum_{k_1,k_2,k_3}^{n_i}E(g_2(ijk_1j'k_1j_1k_2j_1'k_3)) - \right. \\
& \left.\frac{1}{n_i}\sum_{k_1,k_2,k_3}^{n_i}E(g_2(ijk_1j'k_2j_1k_3j_1'k_3)) + \frac{1}{n_i^2}\sum_{k_1,k_2,k_3,k_4}^{n_i}E(g_2(ijk_1j'k_2j_1k_3j_1'k_4))\right] - \sigma_{ijj'}\sigma_{ij_1j_1'} \\
= \;& \frac{1}{n_i}[E(g_2(ijk_1j'k_1j_1k_1j_1'k_1)) - \sigma_{ijj'}\sigma_{ij_1j_1'}] + \frac{\sigma_{ijj_1}\sigma_{ij'j_1'} + \sigma_{ijj_1'}\sigma_{ij'j_1}}{n_i(n_i-1)}. \qquad (4.4.18)
\end{aligned}
$$

By an argument similar to that of (4.4.16) and

$$
\sum_{j=1}^{b}\sum_{j'=1}^{b}\sum_{j_1=1}^{b}\sum_{j_1'=1}^{b}\sigma_{i'jj'}\sigma_{i_1j_1j_1'}\sigma_{ijj_1'}\sigma_{ij'j_1} \le \sum_{j=1}^{b}\sum_{j'=1}^{b}\sum_{j_1=1}^{b}\sum_{j_1'=1}^{b}K\alpha^{1/2}_{|j-j'|}\alpha^{1/2}_{|j_1-j_1'|}\alpha^{1/2}_{|j-j_1'|}\alpha^{1/2}_{|j'-j_1|} = o(b^2),
$$

we can see that the first summation on the right hand side of $\mathrm{Var}(\widehat{\zeta}_2)$ is $o(1)$.

$$
\begin{aligned}
&E(\widehat{\sigma}_{ijj'}\widehat{\sigma}_{ij_1j_1'})E(\widehat{\sigma}_{i'jj'}\widehat{\sigma}_{i'j_1j_1'}) - \sigma_{ijj'}\sigma_{i'jj'}\sigma_{ij_1j_1'}\sigma_{i'j_1j_1'} \\
= \;& \left[E(\widehat{\sigma}_{ijj'}\widehat{\sigma}_{ij_1j_1'}) - \sigma_{ijj'}\sigma_{ij_1j_1'}\right]\left[E(\widehat{\sigma}_{i'jj'}\widehat{\sigma}_{i'j_1j_1'}) - \sigma_{i'jj'}\sigma_{i'j_1j_1'}\right] \\
&+\left[E(\widehat{\sigma}_{ijj'}\widehat{\sigma}_{ij_1j_1'}) - \sigma_{ijj'}\sigma_{ij_1j_1'}\right]\sigma_{i'jj'}\sigma_{i'j_1j_1'} + \left[E(\widehat{\sigma}_{i'jj'}\widehat{\sigma}_{i'j_1j_1'}) - \sigma_{i'jj'}\sigma_{i'j_1j_1'}\right]\sigma_{ijj'}\sigma_{ij_1j_1'}.
\end{aligned}
$$

By (4.4.18) and similar proof as that for 4.4.16 and 4.4.17 the second summation on the right hand side of $\mathrm{Var}(\widehat{\zeta}_2)$ can be shown to be $o(1)$. Then we finish the proof.

Lemma 4.4.5. *Suppose $X_{ijk}, j = 1,2\ldots,$ is α-mixing with $\alpha_m = O(m^{-5})$, and let η_i be defined by (4.4.19). Then a consistent estimator of η_i, $i = 1,\cdots,a$, is $\widehat{\eta}_i$, where $\widehat{\eta}_i$ is given in (4.2.9).*

Proof. By Lemma 4.4.1, the limit exists

$$\eta_i = \lim_{b\to\infty} N\mathrm{Var}(\overline{X}_{i..} - E(\overline{X}_{i..})) = \lim_{b\to\infty} \frac{n}{bn_i} \sum_{j=1}^{b} \sum_{j'=1}^{b} E(u_{ijk}u_{ij'k}).$$

$$\begin{aligned} \widehat{\eta}_i &= \frac{n}{bn_i(n_i-1)} \sum_{j=1}^{b} \sum_{j'=-b^h}^{b^h} \sum_{k=1}^{n_i} \left(u_{ijk} - \overline{u}_{ij.}\right)\left(u_{i(j+j')k} - \overline{u}_{i(j+j').}\right) \\ &= \frac{n}{bn_i(n_i-1)} \sum_{j=1}^{b} \sum_{j'=-b^h}^{b^h} \left(\sum_{k=1}^{n_i} u_{ijk}u_{i(j+j')k} - n_i\overline{u}_{ij.}\overline{u}_{i(j+j').} \right) \end{aligned}$$

$$E(\widehat{\eta}_i) = \frac{n}{bn_i} \sum_{j=1}^{b} \sum_{j'=-b^h}^{b^h} E(u_{ijk}u_{i(j+j')k})$$

We will show that $\eta_i - E(\widehat{\eta}_i) \to 0$ and $\widehat{\eta}_i - E(\widehat{\eta}_i) \xrightarrow{p} 0$ as $b \to \infty$. Take $\eta_i(b) = \frac{n}{bn_i}\sum_{j=1}^{b}\sum_{j'=1}^{b} E(u_{ijk}u_{ij'k})$, then obviously $\eta_i - \eta_i(b) \to 0$ and

$$\begin{aligned} |\eta_i(b) - E(\widehat{\eta}_i)| &\le \frac{n}{bn_i} \sum_{j=1}^{b} \sum_{|j'|>b^h} |E(u_{ijk}u_{i(j+j')k})| \\ &\le \frac{n}{bn_i} \sum_{j=1}^{b} \sum_{|j'|>b^h} 8\left(1 + E(u_{ijk}^4) + E(u_{i(j+j')k}^4)\right)\alpha_{|j'|}^{1/2} \\ &\le \frac{n}{n_i} \sum_{|j'|>b^h} K\alpha_{|j'|}^{1/2} \to 0 \text{ as } b \to \infty\,, \end{aligned}$$

where the last equality is because $\sum_{m=1}^{\infty} \alpha_m^{1/2} < \infty$. So $\eta_i - E(\widehat{\eta}_i) \to 0$.

Write $\widehat{\eta}_i = \widehat{\eta}_{i1+} + \widehat{\eta}_{i1-} - \widehat{\eta}_{i2+} - \widehat{\eta}_{i2+}$ and $E(\widehat{\eta}_i) = \eta_{i1+} + \eta_{i1-} - \eta_{i2+} - \eta_{i2-}$, where

$$\widehat{\eta}_{i1+} = \frac{n}{bn_i(n_i-1)} \sum_{j=1}^{b} \sum_{j'=1}^{b^h} \sum_{k=1}^{n_i} u_{ijk} u_{i(j+j')k}, \quad \widehat{\eta}_{i2+} = \frac{n}{b(n_i-1)} \sum_{j=1}^{b} \sum_{j'=1}^{b^h} \overline{u}_{ij.} \overline{u}_{i(j+j').}$$

$$\widehat{\eta}_{i1-} = \frac{n}{bn_i(n_i-1)} \sum_{j=1}^{b} \sum_{j'=-b^h}^{0} \sum_{k=1}^{n_i} u_{ijk} u_{i(j+j')k}, \quad \widehat{\eta}_{i2-} = \frac{n}{b(n_i-1)} \sum_{j=1}^{b} \sum_{j'=-b^h}^{0} \overline{u}_{ij.} \overline{u}_{i(j+j').}$$

and $\eta_{i1+} = E(\widehat{\eta}_{i1+})$, $\eta_{i1-} = E(\widehat{\eta}_{i1-})$, $\eta_{i2+} = E(\widehat{\eta}_{i2+})$, $\eta_{i2-} = E(\widehat{\eta}_{i2-})$.

We only need to show $\widehat{\eta}_{i1+} - \eta_{i1+} \xrightarrow{p} 0$ and $\widehat{\eta}_{i2+} - \eta_{i2+} \xrightarrow{p} 0$. The rest two, $\widehat{\eta}_{i1-} - \eta_{i1-} \xrightarrow{p} 0$ and $\widehat{\eta}_{i2-} - \eta_{i2-} \xrightarrow{p} 0$, will follow similarly. Denote $f_{ikjj'j_1j_1'} = E(u_{ijk}u_{i(j+j')k}u_{ij_1k}u_{i(j_1+j_1')k}) - E(u_{ijk}u_{i(j+j')k})E(u_{ij_1k}u_{i(j_1+j_1')k})$,

$$E\left(\widehat{\eta}_{i1+} - \eta_{i1+}\right)^2 = \frac{n^2}{b^2 n_i^2 (n_i-1)^2} \sum_{k=1}^{n_i} \sum_{j=1}^{b} \sum_{j'=1}^{b^h} \sum_{j_1=1}^{b} \sum_{j_1'=1}^{b^h} f_{ikjj'j_1j_1'}$$

$$= \frac{2n_i n^2}{b^2 n_i^2 (n_i-1)^2} \sum_{j<j_1}^{b} \sum_{j'=1}^{b^h} \sum_{j_1'=1}^{b^h} f_{ikjj'j_1j_1'} + \frac{2n_i n^2}{b^2 n_i^2 (n_i-1)^2} \sum_{j=1}^{b} \sum_{j'=1}^{b^h} \sum_{j_1'=1}^{b^h} f_{ikjj'jj_1'}$$

It is easily seen that the second term in above formula is $o(1)$ since $E(u_{ijk}^{16}) < \infty$. The first term can be split into two parts:

$$\frac{2n_i n^2}{b^2 n_i^2 (n_i-1)^2} \sum_{j<j_1}^{b} \sum_{j'=1}^{b^h} \sum_{j_1'=1}^{b^h} f_{ikjj'j_1j_1'} = \frac{n^2}{n_i(n_i-1)^2}(D_1 + D_2),$$

where

$$D_1 = \frac{2}{b^2} \sum_{j<j_1}^{b} \sum_{j'=1}^{b^h} \sum_{j_1'=1}^{b^h} I(j+j' \leq j_1) f_{ikjj'j_1j_1'}$$

$$D_2 = \frac{2}{b^2} \sum_{j<j_1}^{b} \sum_{j'=1}^{b^h} \sum_{j_1'=1}^{b^h} I(j+j' > j_1 > j) f_{ikjj'j_1j_1'}$$

By α-mixing condition,

$$\begin{aligned}
& I(j+j' \le j_1)|f_{ikjj'j_1j_1'}| \\
\le\ & 8[1+E(u_{ijk}^4 u_{i(j+j')k}^4)+E(u_{ij_1k}^4 u_{i(j_1+j_1')k}^4)]\alpha_{j_1-(j+j')}^{1/2} \\
\le\ & 8[1+E^{\frac{1}{2}}(u_{ijk}^8)E^{\frac{1}{2}}(u_{i(j+j')k}^8)+E^{\frac{1}{2}}(u_{ij_1k}^8)E^{\frac{1}{2}}(u_{i(j_1+j_1')k}^8)]\alpha_{j_1-(j+j')}^{1/2} \\
\le\ & K_1\alpha_{j_1-(j+j')}^{1/2} \text{ for some } K_1.
\end{aligned}$$

So by the fact that $\sum_{m=1}^{\infty}\alpha_m^{\frac{1}{2}} < \infty$,

$$|D_1| \le \frac{2}{b^{2-h}}\sum_{j<j_1}^{b}\sum_{j'=1}^{b^h} I(j+j' \le j_1)K_1\alpha_{j_1-(j+j')}^{1/2} \le \frac{2}{b^{2-h}}\sum_{j=1}^{b}\sum_{j'=1}^{b^h}\sum_{m=1}^{b}\alpha_m^{\frac{1}{2}} = O(b^{2h-1}) \to 0$$

For D_2, first apply Cauchy-Schwarz Inequality twice and use the fact $E(u_{ijk}^{16}) < \infty$, we obtain that there exists $K < \infty$ such that

$$\begin{aligned}
|f_{ikjj'j_1j_1'}| \le\ & E^{\frac{1}{4}}(u_{ijk}^4)E^{\frac{1}{4}}(u_{i(j+j')k}^4)E^{\frac{1}{4}}(u_{ij_1k}^4)E^{\frac{1}{4}}(u_{i(j_1+j_1')k}^4) \\
& +E^{\frac{1}{2}}(u_{ijk}^2)E^{\frac{1}{2}}(u_{i(j+j')k}^2)E^{\frac{1}{2}}(u_{ij_1k}^2)E^{\frac{1}{2}}(u_{i(j_1+j_1')k}^2) < K
\end{aligned}$$

So

$$|D_2| \le \frac{2}{b^2}\sum_{j=1}^{b}\sum_{j'=1}^{b^h}\sum_{j_1=j}^{j+j'}|f_{ikjj'j_1j_1'}| \le \frac{2}{b^2}\sum_{j=1}^{b}\sum_{j'=1}^{b^h} Kj' = \frac{2}{b}\frac{b^h(b^h+1)}{2} = O(b^{2h-1}) \to 0$$

Therefore, $\widehat{\eta}_{i1+} - \eta_{i1+} = o_p(1)$.

$$\widehat{\eta}_{i2+} = \frac{\widehat{\eta}_{i1+}}{n_i} + \frac{n}{bn_i^2(n_i-1)}\sum_{j=1}^{b}\sum_{j'=1}^{b^h}\sum_{k\neq k'}^{n_i} u_{ijk}u_{i(j+j')k'}$$

The first term on the right hand side of above equation converges in probability. Denote the second term on the right hand side of above equation as $\widehat{\eta}_{i22}$, then it remains to show that $\widehat{\eta}_{i22}$

converges in probability and it suffice to show $\mathrm{Var}(\widehat{\eta}_{i22}) \to 0$.

$$\frac{b^2 n_i^4 (n_i-1)^2}{n^2}\mathrm{Var}(\widehat{\eta}_{i22}) = \sum_{j=1}^{b}\sum_{j_1=1}^{b}\sum_{j'=1}^{b^h}\sum_{j_1'=1}^{b^h}\sum_{k\neq k'}^{n_i}\sum_{k_1\neq k_1'}^{n_i}$$

$$\left[E(u_{ijk}u_{i(j+j')k'}u_{ij_1k_1}u_{i(j_1+j_1')k_1'}) - E(u_{ijk}u_{i(j+j')k'})E(u_{ij_1k_1}u_{i(j_1+j_1')k_1'})\right]$$

$$= \sum_{j=1}^{b}\sum_{j_1=1}^{b}\sum_{j'=1}^{b^h}\sum_{j_1'=1}^{b^h}\sum_{k\neq k'}^{n_i}\Big[E(u_{ijk}u_{ij_1k})E(u_{i(j+j')k'}u_{i(j_1+j_1')k'})$$

$$+E(u_{ijk}u_{i(j_1+j_1')k})E(u_{i(j+j')k'}u_{ij_1k'}) - 2E(u_{ijk})E(u_{ij_1k})E(u_{i(j+j')k'})E(u_{i(j_1+j_1')k'})\Big]$$

$$\begin{aligned}
&|E(u_{ijk}u_{ij_1k})E(u_{i(j+j')k'}u_{i(j_1+j_1')k'}) - E(u_{ijk})E(u_{ij_1k})E(u_{i(j+j')k'})E(u_{i(j_1+j_1')k'})| \\
= \quad & |\left[E(u_{ijk}u_{ij_1k}) - E(u_{ijk})E(u_{ij_1k})\right]\left[E(u_{i(j+j')k'}u_{i(j_1+j_1')k'}) - E(u_{i(j+j')k'})E(u_{i(j_1+j_1')k'})\right] \\
&- \left[E(u_{ijk}u_{ij_1k}) - E(u_{ijk})E(u_{ij_1k})\right]E(u_{i(j+j')k'})E(u_{i(j_1+j_1')k'}) \\
&- \left[E(u_{i(j+j')k'}u_{i(j_1+j_1')k'}) - E(u_{i(j+j')k'})E(u_{i(j_1+j_1')k'})\right]E(u_{ijk})E(u_{ij_1k})| \\
\leq \quad & K_2\alpha^{\frac{1}{2}}_{j_1-j}\alpha^{\frac{1}{2}}_{|j_1+j_1'-j-j'|} + K_3\alpha^{\frac{1}{2}}_{j_1-j} + K_4\alpha^{\frac{1}{2}}_{|j_1+j_1'-j-j'|},
\end{aligned}$$

where the last inequality is due to α-mixing condition. Similarly,

$$|E(u_{ijk}u_{i(j_1+j_1')k})E(u_{i(j+j')k'}u_{ij_1k'}) - E(u_{ijk})E(u_{ij_1k})E(u_{i(j+j')k'})E(u_{i(j_1+j_1')k'})|$$
$$\leq K_5\alpha^{\frac{1}{2}}_{j_1+j_1'-j}\alpha^{\frac{1}{2}}_{|j_1-j-j'|} + K_6\alpha^{\frac{1}{2}}_{j_1+j_1'-j} + K_7\alpha^{\frac{1}{2}}_{|j_1-j-j'|}$$

Therefore, $\frac{b^2 n_i^4 (n_i-1)^2}{n^2}\mathrm{Var}(\widehat{\eta}_{i22}) = O((b+b^h)n_i^2)$ and so $\mathrm{Var}(\widehat{\eta}_{i22}) \to 0$. Thus we finished the proof for $\widehat{\eta}_{i2+} - E(\widehat{\eta}_{i2+}) = o_p(1)$ as $b \to \infty$. Then we finished the proof of the Lemma.

Proof of Theorem 4.2.3 Note that $\sqrt{N}(\overline{X}_{i..} - E(\overline{X}_{i..})) = \frac{\sqrt{n}}{\sqrt{bn_i}}\sum_{k=1}^{n_i}\sum_{j=1}^{b} u_{ijk}$. Apply inequality (4.4.15), there is a finite A such that $E\left(\frac{\sqrt{n}}{n_i}\sum_{k=1}^{n_i} u_{ijk}\right)^{16} \leq A\frac{n^8}{n_i^{16}}n_i^8 E(u_{ijk}^{16}) < \infty$.

By Lemmas 4.4.2 and 4.4.1, $\sqrt{N}[\overline{X}_{i..} - E(\overline{X}_{i..})] \xrightarrow{d} N(0,\eta_i)$, where η_i is defined by

$$\mathrm{Var}(\sqrt{N}[\overline{X}_{i..} - E(\overline{X}_{i..})]) = \frac{n}{bn_i}\sum_{j=1}^{b}\sum_{j'=1}^{b} E(u_{ijk}u_{ij'k}) \to \eta_i. \tag{4.4.19}$$

By independence, $\sqrt{N}(\mathbf{W}-E(\mathbf{W})) \stackrel{d}{\to} N_a(\mathbf{0},\mathbf{V})$, where $\mathbf{V} = diag(\eta_1,\cdots,\eta_a)$. Therefore by Continuous Maping Theorem, $\sqrt{N}\mathbf{C}_a\mathbf{W} \stackrel{d}{\to} N_{a-1}(\mathbf{0},\mathbf{C}_a\mathbf{V}\mathbf{C}_a')$, as $b \to \infty$ regardless of whether $n_i \to \infty$ or not. By Lemma 4.5, $\widehat{\eta}_i$ are consistent estimators of η_i, $i = 1,\cdots,a$. The proof is finished by an application of Slutsky's theorem.

4.4.2 Proofs for rank tests

Before we prove the theorems, we will give a basic result first as a lemma. All other lemmas used in the proof of the theorems are stated and proved after the proof of corresponding theorem.

Define $\mathbf{F} = (F_{11},\cdots,F_{ab})'$ and $\hat{\mathbf{F}} = (\widehat{F}_{11},\cdots,\widehat{F}_{ab})'$, where $\widehat{F}_{ij}(x) = n_i^{-1}\sum_{k=1}^{n_i} c(X_{ijk},x)$. Let $\overline{\widehat{\mathbf{F}}} = (\overline{\widehat{F}}_{1.},\cdots,\overline{\widehat{F}}_{a.})'$ and $\overline{\mathbf{F}} = (\overline{F}_{1.},\cdots,\overline{F}_{a.})'$.

Lemma 4.4.6. *If $b \to \infty$, assume X_{ijk}, $j = 1,2,\ldots,$ is α-mixing with $\alpha_m = O\left(m^{-5}\right)$ for all i,k. Then as $b \to \infty$, regardless of whether the n_i remain fixed or $n(a) \to \infty$,*

$$\sqrt{N}\int(\widehat{H}-H)\,d(\overline{\widehat{\mathbf{F}}}-\overline{\mathbf{F}}) \stackrel{p}{\to} \mathbf{0}.$$

This result is a slightly different version of a result shown in Akritas and Brunner (1997) for $n(a) \to \infty$ and b fixed. For a proof of Lemma 4.4.6 see Wang (2004) or Akritas and Wang (2004).

Proof. The component of the vector $\sqrt{N}\int(\widehat{H}-H)\,d(\hat{\mathbf{F}}-\mathbf{F})$ is

$$
\begin{aligned}
&\sqrt{N}\int(\widehat{H}-H)\,d(\widehat{F}_{ij}-F_{ij})\\
=\;& \frac{1}{\sqrt{N}}\sum_{i_1,j_1} n_{i_1}\int(\widehat{F}_{i_1j_1}-F_{i_1j_1})\,d(\widehat{F}_{ij}-F_{ij})\\
=\;& \frac{1}{\sqrt{N}}\sum_{i_1,j_1}\left\{\frac{n_{i_1}}{n_i}\sum_{k=1}^{n_i}(\widehat{F}_{i_1j_1}(X_{ijk})-F_{i_1j_1}(X_{ijk}))-\sum_{k_1=1}^{n_{i_1}}\left[1-F_{ij}(X_{i_1j_1k_1})-\int F_{i_1j_1}\,dF_{ij}\right]\right\}\\
=\;& \frac{1}{\sqrt{N}}\sum_{i_1,j_1}\frac{1}{n_i}\sum_{k=1}^{n_i}\sum_{k_1=1}^{n_{i_1}}\left\{c(X_{i_1j_1k_1},X_{ijk})-F_{i_1j_1}(X_{ijk})-\left[1-F_{ij}(X_{i_1j_1k_1})-\int F_{i_1j_1}\,dF_{ij}\right]\right\}\\
=\;& \frac{1}{\sqrt{N}}\sum_{j_1=1}^{b} nh(\mathbf{X}_{j_1},\mathbf{X}_{ij}),
\end{aligned}
$$

where

$$
h(\mathbf{X}_{j_1},\mathbf{X}_{ij})=\frac{1}{nn_i}\sum_{i_1,k_1}\sum_{k=1}^{n_i}\left\{c(X_{i_1j_1k_1},X_{ijk})-F_{i_1j_1}(X_{ijk})-\left[1-F_{ij}(X_{i_1j_1k_1})-\int F_{i_1j_1}\,dF_{ij}\right]\right\}, \tag{4.4.20}
$$

$\mathbf{X}_{j_1}=(X_{1j_11},\cdots,X_{aj_1n_a})'$, and $\mathbf{X}_{ij}=(X_{ij1},\cdots,X_{ijn_i})'$. Note that $h(\mathbf{X}_{j_1},\mathbf{X}_{ij})$ is uniformly bounded by 4 and satisfies

$$
Eh(\mathbf{X}_{j_1},\mathbf{X}_{ij})=E[h(\mathbf{X}_{j_1},\mathbf{X}_{ij})|\mathbf{X}_{j_1}]=E[h(\mathbf{X}_{j_1},\mathbf{X}_{ij})|\mathbf{X}_{ij}]=0. \tag{4.4.21}
$$

The component of the vector $\sqrt{N}\int(\widehat{H}-H)\,d(\widehat{\overline{\mathbf{F}}}-\overline{\mathbf{F}})$ is

$$
\frac{\sqrt{N}}{b}\sum_{j=1}^{b}\int(\widehat{H}-H)\,d(\widehat{F}_{ij}-F_{ij})=\frac{\sqrt{N}}{bN}\sum_{j=1}^{b}\sum_{j_1=1}^{b} nh(\mathbf{X}_{j_1},\mathbf{X}_{ij})
$$

and

$$
\begin{aligned}
&E\left(\frac{\sqrt{N}}{b}\sum_{j=1}^{b}\int(\widehat{H}-H)\,d(\widehat{F}_{ij}-F_{ij})\right)^2\\
=\;& \frac{n^2}{b^2N}\sum_{j=1}^{b}\sum_{j_1=1}^{b}\sum_{j_2=1}^{b}\sum_{j_3=1}^{b}E\left[h(\mathbf{X}_{j_1},\mathbf{X}_{ij})h(\mathbf{X}_{j_2},\mathbf{X}_{ij_3})\right]\\
=\;& \frac{n^2}{b^2N}\sum_{j=1}^{b}\sum_{j_3=1}^{b}\left[2\sum_{j_1<j_2}^{b}E\left[h(\mathbf{X}_{j_1},\mathbf{X}_{ij})h(\mathbf{X}_{j_2},\mathbf{X}_{ij_3})\right]+\sum_{j_1=1}^{b}E\left[h(\mathbf{X}_{j_1},\mathbf{X}_{ij})h(\mathbf{X}_{j_1},\mathbf{X}_{ij_3})\right]\right]\\
=\;& u_{D1}+u_{D2}+u_{D3}+u_{D4}+u_{D5},
\end{aligned}
$$

where

$$u_{D1} = \frac{2n^2}{b^2N} \sum_{j_1<j_2}^{b} \sum_{j<j_3}^{b} E\left[h(\mathbf{X}_{j_1}, \mathbf{X}_{ij})h(\mathbf{X}_{j_2}, \mathbf{X}_{ij_3})\right]$$

$$u_{D2} = \frac{2n^2}{b^2N} \sum_{j_1<j_2}^{b} \sum_{j>j_3}^{b} E\left[h(\mathbf{X}_{j_1}, \mathbf{X}_{ij})h(\mathbf{X}_{j_2}, \mathbf{X}_{ij_3})\right]$$

$$u_{D3} = \frac{2n^2}{b^2N} \sum_{j_1<j_2}^{b} \sum_{j=1}^{b} E\left[h(\mathbf{X}_{j_1}, \mathbf{X}_{ij})h(\mathbf{X}_{j_2}, \mathbf{X}_{ij})\right]$$

$$u_{D4} = \frac{2n^2}{b^2N} \sum_{j_1=1}^{b} \sum_{j<j_3}^{b} E\left[h(\mathbf{X}_{j_1}, \mathbf{X}_{ij})h(\mathbf{X}_{j_1}, \mathbf{X}_{ij_3})\right]$$

$$u_{D5} = \frac{n^2}{b^2N} \sum_{j=1}^{b} \sum_{j_1=1}^{b} E[h^2(\mathbf{X}_{j_1}, \mathbf{X}_{ij})].$$

Obviously, $u_{D5} = O(n/(bn(a))) = o(1)$, since $E(h(\cdot,\cdot)^2) = O(n_i^{-1})$ which can be verified by examining the expected value in further detail.

$$u_{D1} = \frac{2n^2}{b^2N} \sum_{j<j_3}^{b} \sum_{j_1<j_2}^{b} E\left[E[h(\mathbf{X}_{j_1}, \mathbf{X}_{ij})h(\mathbf{X}_{j_2}, \mathbf{X}_{ij_3})|(\mathbf{X}_{j_1}, \mathbf{X}_{ij})]\right] = 0$$

Similarly $u_{D2} = 0$ and

$$\begin{aligned}
|u_{D3}| &\leq \frac{2n^2}{b^2N} \sum_{j_1<j_2}^{b} \sum_{j=1}^{b} \left|E\left[h(\mathbf{X}_{j_1}, \mathbf{X}_{ij})h(\mathbf{X}_{j_2}, \mathbf{X}_{ij})\right]\right| \\
&= \frac{2n^2}{b^2N} \sum_{j_1<j_2}^{b} \sum_{j=1}^{b} \left|E(E\left[h(\mathbf{X}_{j_1}, \mathbf{X}_{ij})h(\mathbf{X}_{j_2}, \mathbf{X}_{ij})|\mathbf{X}_{ij}\right])\right| \\
&= \frac{2n^2}{b^2N} \sum_{j_1<j_2}^{b} \sum_{j=1}^{b} I(j = j_1 \text{ or } j = j_2)\left|E(E\left[h(\mathbf{X}_{j_1}, \mathbf{X}_{ij})h(\mathbf{X}_{j_2}, \mathbf{X}_{ij})|\mathbf{X}_{ij}\right])\right| \\
&= 0
\end{aligned}$$

where the last equality is due to (4.4.21). The proof of $|u_{D4}| = o(1)$ is similar to that of $|u_{D3}| = o(1)$. Hence $\frac{\sqrt{N}}{b} \sum_{j=1}^{b} \int (\widehat{H} - H)\, d(\widehat{F}_{ij} - F_{ij}) = o_p(1)$ and we finish the proof.

Lemma 4.4.7. *Assume $X_{ijk}, j = 1,2,\ldots,$ is α-mixing with $\alpha_m = O\left(m^{-5}\right)$ for all i,k, Let $\mathbf{R}^c = (R_{111} - \mu_{R,11}, R_{112} - \mu_{R,11}, \ldots, R_{11n_1} - \mu_{R,11}, R_{121} - \mu_{R,12}, \ldots, R_{abn_a} - \mu_{R,ab})$, where $\mu_{R,ij} = NE(H(X_{ijk})) + 1/2$, and let $P_{MSE}(\cdot)$ be the function defined in relation (4.4.12). Then*

then as $b \to \infty$

1. *if n_i remain fixed,*

   (a) $\frac{\sqrt{b}}{N^2}(MSE_R - P_{MSE}(\mathbf{R}^c)) = o_p(1).$

   (b) $MSE_R \xrightarrow{p} \tilde{\sigma}^2$ *, where*

$$\tilde{\sigma}^2 = \lim_{b\to\infty} \frac{1}{ab} \sum_{i,j,k} \frac{\tilde{\sigma}_{ij}^2}{n_i^2}$$

   provided that the limit exists.

2. *if $n(a) \to \infty$ as b goes to ∞, then*

   (a) $n(a)\sqrt{b}[MSE_R - P_{MSE}(\mathbf{R}^c)] = o_p(1).$

   (b) $n(a)MSE_R \xrightarrow{p} \tilde{\sigma}_*^2$ *, where*

$$\tilde{\sigma}_*^2 = \lim_{b\to\infty} \frac{1}{ab} \sum_{i,j,k} \frac{n(a)}{n_i^2} \tilde{\sigma}_{ij}^2$$

   provided that the limit exists.

Proof. Use the decomposition, $MSE_R = P_{MSE}(\mathbf{R}^c) + D_1(\mathbf{R}^c) + D_2(\mathbf{R}^c)$, given in the proof of Lemma 4.4.3. Define $\mathbf{Y}^c = (Y_{111} - p_{11}, \ldots, Y_{11n_1} - p_{11}, Y_{121} - p_{12}, \ldots, Y_{abn_a} - p_{ab})$, where $p_{ij} = E(H(X_{ijk})) = E(Y_{ijk})$. By Lemma 4.4.3, the proof would be done if we will show that in both cases,

$$n(a)\sqrt{b}(D_1(\mathbf{R}^c)/N^2 - D_1(\mathbf{Y}^c)) = o_p(1) \tag{4.4.22}$$

$$n(a)\sqrt{b}(D_2(\mathbf{R}^c)/N^2 - D_2(\mathbf{Y}^c)) = o_p(1) \tag{4.4.23}$$

$$n(a)(P_{MSE}(\mathbf{R}^c)/N^2 - P_{MSE}(\mathbf{Y}^c)) = o_p(1). \tag{4.4.24}$$

We will first show (4.4.22).

$$\begin{aligned} & D_1(\mathbf{R}^c)/N^2 - D_1(\mathbf{Y}^c) \\ = \; & -\frac{1}{ab(b-1)}\sum_{i=1}^{a}\sum_{k=1}^{n_i}\sum_{j\neq j'}^{b}\left[\frac{(Z_{ijk}-p_{ij})(Z_{ij'k}-p_{ij'})-(Y_{ijk}-p_{ij})(Y_{ij'k}-p_{ij'})}{n_i(n_i-1)}\right] \\ = \; & D_{11}+D_{12}, \end{aligned}$$

where

$$\begin{aligned} D_{11} &= -\frac{1}{ab(b-1)}\sum_{i=1}^{a}\sum_{k=1}^{n_i}\sum_{j\neq j'}^{b}\frac{(Z_{ijk}-Y_{ijk})(Z_{ij'k}-Y_{ij'k})}{n_i(n_i-1)} \\ D_{12} &= -\frac{2}{ab(b-1)}\sum_{i=1}^{a}\sum_{k=1}^{n_i}\sum_{j\neq j'}^{b}\frac{(Z_{ijk}-Y_{ijk})(Y_{ij'k}-p_{ij'})}{n_i(n_i-1)}. \end{aligned}$$

By Lemma 4.4.6, $sup_x|\widehat{H}(x)-H(x)| = O_p(N^{-1/2})$. So $D_{11} = O_p(N^{-1}n(a)^{-1})$ and $n(a)\sqrt{b}D_{11} = o_p(1)$.

$$\begin{aligned} D_{12} &= \frac{-2}{ab(b-1)}\sum_{i=1}^{a}\sum_{k=1}^{n_i}\sum_{j\neq j'}^{b}\frac{(Z_{ijk}-Y_{ijk})(Y_{ij'k}-p_{ij'})}{n_i(n_i-1)} \\ &= \frac{-2}{ab(b-1)N}\sum_{i=1}^{a}\sum_{k=1}^{n_i}\sum_{j\neq j'}^{b}\sum_{i_1=1}^{a}\sum_{j_1=1}^{b}\sum_{k_1=1}^{n_{i_1}}\frac{\left[c(X_{i_1j_1k_1},X_{ijk})-F_{i_1j_1}(X_{ijk})\right]\left(Y_{ij'k}-p_{ij'}\right)}{n_i(n_i-1)}, \end{aligned}$$

$$\begin{aligned} & E(D_{12}^2) \\ = \; & \frac{4}{a^2b^2(b-1)^2N^2}\sum_{i=1}^{a}\sum_{k=1}^{n_i}\sum_{j\neq j'}^{b}\sum_{i_1=1}^{a}\sum_{j_1=1}^{b}\sum_{k_1=1}^{n_{i_1}}\sum_{i_2=1}^{a}\sum_{k_2=1}^{n_{i_2}}\sum_{j_2\neq j_3}^{b}\sum_{i_4=1}^{a}\sum_{j_4=1}^{b}\sum_{k_4=1}^{n_{i_4}} E\left\{\frac{\left(Y_{ij'k}-p_{ij'}\right)}{n_i(n_i-1)}\right. \\ & \left.\frac{\left(Y_{i_2j_3k_2}-p_{i_2j_3}\right)\left[c(X_{i_1j_1k_1},X_{ijk})-F_{i_1j_1}(X_{ijk})\right]\left[c(X_{i_4j_4k_4},X_{i_2j_2k_2})-F_{i_4j_4}(X_{i_2j_2k_2})\right]}{n_{i_2}(n_{i_2}-1)}\right\} \end{aligned}$$

Note that $E(c(X_{i_1j_1k_1},X_{ijk})-F_{i_1j_1}(X_{ijk})|X_{ijk}) = 0$, so the expectation under the summation is zero if the number of different elements in set $\{k,k_1,k_2,k_4\}$ is four. If the number of different elements in set $\{j,j',j_1,j_2,j_3,j_4\}$ is four or less, the summation is of order $O(b^{-2}n(a)^{-1})$.

When all elements in the set $\{j, j', j_1, j_2, j_3, j_4\}$ are different, without loss of generality, we can consider a representative case in which $j < j' < j_1 < j_2 < j_3 < j_4$ and $k_4 = k$ (note: the expectation under the summation is not zero only when the number of different elements in $\{k, k_1, k_2, k_4\}$ is three or less.)

$$\begin{aligned}
& \sum_{j'<j_1<j_3<j_4}^{b} E\left\{\left(Y_{ij'k} - p_{ij'}\right)\left[c(X_{i_1j_1k_1}, X_{ijk}) - F_{i_1j_1}(X_{ijk})\right]\right. \\
& \qquad \left.\left(Y_{i_2j_3k_2} - p_{i_2j_3}\right)\left[c(X_{i_4j_4k_4}, X_{i_2j_2k_2}) - F_{i_4j_4}(X_{i_2j_2k_2})\right]\right\} \\
= & \sum_{j'<j_1<j_3<j_4}^{b} E\left(E\left\{\left(Y_{ij'k} - p_{ij'}\right)\left[c(X_{i_1j_1k_1}, X_{ijk}) - F_{i_1j_1}(X_{ijk})\right]\right.\right. \\
& \qquad \left.\left.\left(Y_{i_2j_3k_2} - p_{i_2j_3}\right)\left[c(X_{i_4j_4k_4}, X_{i_2j_2k_2}) - F_{i_4j_4}(X_{i_2j_2k_2})\right]\right\} \middle| (X_{ijk}, X_{i_2j_2k_2})\right) \\
\leq & \sum_{j'<j_1<j_3<j_4}^{b} 4 \times 2 \times 8\alpha_{j_1-j'}^{1/2},
\end{aligned}$$

where the last inequality is because both $\left(Y_{ij'k} - p_{ij'}\right)$ and $\left[c(X_{i_1j_1k_1}, X_{ijk}) - F_{i_1j_1}(X_{ijk})\right]\left(Y_{i_2j_3k_2} - p_{i_2j_3}\right)\left[c(X_{i_4j_4k_4}, X_{i_2j_2k_2}) - F_{i_4j_4}(X_{i_2j_2k_2})\right]$ have zero mean and bounded by 2 and 4, respective, so we can apply the lemma in Billingley for α-mixing sequences and the condition that the observations from different subjects are independent.. Similarly, $\sum_{j'<j_1<j_3<j_4}^{b} 4 \times 4 \times 4\alpha_{j_3-j_1}^{1/2}$ is also a upper bound. So

$$\begin{aligned}
& \sum_{j'<j_1<j_3<j_4}^{b} E\left\{\left(Y_{ij'k} - p_{ij'}\right)\left[c(X_{i_1j_1k_1}, X_{ijk}) - F_{i_1j_1}(X_{ijk})\right]\right. \\
& \qquad \left.\left(Y_{i_2j_3k_2} - p_{i_2j_3}\right)\left[c(X_{i_4j_4k_4}, X_{i_2j_2k_2}) - F_{i_4j_4}(X_{i_2j_2k_2})\right]\right\} \\
\leq & \sum_{j'<j_1<j_3<j_4}^{b} 64 \min(\alpha_{j_1-j'}^{1/2}, \alpha_{j_3-j_1}^{1/2}) \qquad (4.4.25) \\
& = O(b^2),
\end{aligned}$$

The detailed reasoning for the last step is given in the proof of Lemma 4.4.1.

$$\frac{4}{a^2b^2(b-1)^2N^2}\sum_{i=1}^{a}\sum_{k=1}^{n_i}\sum_{i_1=1}^{a}\sum_{k_1=1}^{n_{i_1}}\sum_{i_2=1}^{a}\sum_{k_2=1}^{n_{i_2}}\sum_{i_4=1}^{a}\sum_{j<j'<j_1<j_2<j_3<j_4}^{b} E\left\{\frac{\left(Y_{ij'k}-p_{ij'}\right)}{n_i(n_i-1)}\right.$$
$$\left.\frac{\left(Y_{i_2j_3k_2}-p_{i_2j_3}\right)\left[c(X_{i_1j_1k_1},X_{ijk})-F_{i_1j_1}(X_{ijk})\right]\left[c(X_{i_4j_4k_4},X_{i_2j_2k_2})-F_{i_4j_4}(X_{i_2j_2k_2})\right]}{n_{i_2}(n_{i_2}-1)}\right\}$$
$$\leq \frac{4}{a^2b^2(b-1)^2N^2}\sum_{i=1}^{a}\sum_{k=1}^{n_i}\sum_{i_1=1}^{a}\sum_{k_1=1}^{n_{i_1}}\sum_{i_2=1}^{a}\sum_{k_2=1}^{n_{i_2}}\sum_{i_4=1}^{a}\sum_{j=1}^{b}\sum_{j_2=1}^{b}\sum_{j'<j_1<j_3<j_4}^{b}\frac{\min(\alpha_{j_1-j'}^{1/2},\alpha_{j_3-j_1}^{1/2})}{n_i(n_i-1)n_{i_2}(n_{i_2}-1)}$$
$$= O\left(b^{-1}N^{-1}n^{-2}(a)\right),$$

The situation when the number of different elements in set $\{j, j', j_1, j_2, j_3, j_4\}$ is five is similar.

Therefore, $E(D_{12}^2) = o(n^{-2}(a)b^{-1})$. That is $n(a)\sqrt{b}D_{12} = o_p(1)$. Hence (4.4.22) holds.

The proof of (4.4.23) follows the same argument as that of (4.4.22).

$$D_2(\mathbf{R}^c)/N^2 - D_2(\mathbf{Y}^c)$$
$$= -\frac{1}{ab(b-1)}\sum_{i=1}^{a}\sum_{k=1}^{n_i}\sum_{k'=1}^{n_i}\sum_{j\neq j'}^{b}\left[\frac{(Z_{ijk}-p_{ij})(Z_{ij'k}-p_{ij'})-(Y_{ijk'}-p_{ij})(Y_{ij'k'}-p_{ij'})}{n_i^2(n_i-1)}\right]$$
$$= D_{21}+D_{22},$$

where

$$D_{21} = -\frac{1}{ab(b-1)}\sum_{i=1}^{a}\sum_{k=1}^{n_i}\sum_{k'=1}^{n_i}\sum_{j\neq j'}^{b}\frac{(Z_{ijk}-Y_{ijk})(Z_{ij'k'}-Y_{ij'k'})}{n_i^2(n_i-1)}$$
$$D_{22} = -\frac{2}{ab(b-1)}\sum_{i=1}^{a}\sum_{k=1}^{n_i}\sum_{k'=1}^{n_i}\sum_{j\neq j'}^{b}\frac{(Z_{ijk}-Y_{ijk})(Y_{ij'k'}-p_{ij'})}{n_i^2(n_i-1)}.$$

By Lemma 4.4.6, $sup_x|\widehat{H}(x)-H(x)| = O_p(N^{-1/2})$. So $D_{21} = O_p(N^{-1}n(a)^{-1})$ and $n(a)\sqrt{b}D_{21} = o_p(1)$.

$$D_{22} = \frac{-2}{ab(b-1)}\sum_{i=1}^{a}\sum_{k=1}^{n_i}\sum_{k'=1}^{n_i}\sum_{j\neq j'}^{b}\frac{(Z_{ijk}-Y_{ijk})(Y_{ij'k'}-p_{ij'})}{n_i^2(n_i-1)}$$
$$= \frac{2}{ab(b-1)N}\sum_{i=1}^{a}\sum_{k=1}^{n_i}\sum_{k'=1}^{n_i}\sum_{j\neq j'}^{b}\sum_{i_1=1}^{a}\sum_{j_1=1}^{b}\sum_{k_1=1}^{n_{i_1}}\frac{\left[c(X_{i_1j_1k_1},X_{ijk})-F_{i_1j_1}(X_{ijk})\right]\left(Y_{ij'k'}-p_{ij'}\right)}{n_i^2(n_i-1)},$$

$$
\begin{aligned}
& E(D_{22}^2) \\
= \; & \frac{4}{a^2b^2(b-1)^2N^2} \sum_{i=1}^{a} \sum_{k=1}^{n_i} \sum_{k'=1}^{n_i} \sum_{j \neq j'}^{b} \sum_{i_1=1}^{a} \sum_{j_1=1}^{b} \sum_{k_1=1}^{n_{i_1}} \sum_{i_2=1}^{a} \sum_{k_2=1}^{n_{i_2}} \sum_{k_2'=1}^{n_{i_2}} \sum_{j_2 \neq j_3}^{b} \sum_{i_4=1}^{a} \sum_{j_4=1}^{b} \sum_{k_4=1}^{n_{i_4}} E \left\{ \frac{\left(Y_{ij'k'} - p_{ij'}\right)}{n_i^2(n_i-1)} \right. \\
& \left. \frac{\left(Y_{i_2j_3k_2'} - p_{i_2j_3}\right) \left[c(X_{i_1j_1k_1}, X_{ijk}) - F_{i_1j_1}(X_{ijk})\right] \left[c(X_{i_4j_4k_4}, X_{i_2j_2k_2}) - F_{i_4j_4}(X_{i_2j_2k_2})\right]}{n_{i_2}^2(n_{i_2}-1)} \right\}
\end{aligned}
$$

Note that $E(c(X_{i_1j_1k_1}, X_{ijk}) - F_{i_1j_1}(X_{ijk})|X_{ijk}) = 0$, so the expectation under the summation is zero if the number of different elements in set $\{k, k', k_1, k_2, k_2', k_4\}$ is five or six. If the number of different elements in set $\{j, j', j_1, j_2, j_3, j_4\}$ is four or less, the summation is of order $O(b^{-2}n(a)^{-2})$. When all elements in the set $\{j, j', j_1, j_2, j_3, j_4\}$ are different, without loss of generality, we can consider a representative case in which $j < j' < j_1 < j_2 < j_3 < j_4$, $k_4 = k$ and $k' = k_2'$ (note: the expectation under the summation is not zero only when the number of different elements in $\{k, k', k_1, k_2, k_2', k_4\}$ is four or less.)

$$
\begin{aligned}
& \sum_{j'<j_1<j_3<j_4}^{b} E\left\{ \left(Y_{ij'k'} - p_{ij'}\right) \left[c(X_{i_1j_1k_1}, X_{ijk}) - F_{i_1j_1}(X_{ijk})\right] \right. \\
& \qquad \left. \left(Y_{i_2j_3k_2'} - p_{i_2j_3}\right) \left[c(X_{i_4j_4k_4}, X_{i_2j_2k_2}) - F_{i_4j_4}(X_{i_2j_2k_2})\right] \right\} \\
= \; & \sum_{j'<j_1<j_3<j_4}^{b} E\left\{ \left[c(X_{i_1j_1k_1}, X_{ijk}) - F_{i_1j_1}(X_{ijk})\right] \left[c(X_{i_4j_4k}, X_{i_2j_2k_2}) - F_{i_4j_4}(X_{i_2j_2k_2})\right] \right. \\
& \qquad \left. \left(Y_{ij'k'} - p_{ij'}\right) \left(Y_{i_2j_3k'} - p_{i_2j_3}\right) \right\} \text{ (note that } k_4 = k \text{ and } k' = k_2' \text{) .} \qquad (4.4.26)
\end{aligned}
$$

If $k' \notin \{k, k_1, k_2\}$,

$$
\begin{aligned}
(4.4.26) \;\; \leq \; & \sum_{j_1<j_4}^{b} E\left| \left[c(X_{i_1j_1k_1}, X_{ijk}) - F_{i_1j_1}(X_{ijk})\right] \left[c(X_{i_4j_4k}, X_{i_2j_2k_2}) - F_{i_4j_4}(X_{i_2j_2k_2})\right] \right| \\
& \sum_{j'<j_3} E\left| \left(Y_{ij'k'} - p_{ij'}\right) \left(Y_{i_2j_3k'} - p_{i_2j_3}\right) \right| \qquad (4.4.27) \\
\leq \; & \sum_{j_1<j_4}^{b} 4 \times 2 \times 2\alpha_{j_4-j_1}^{1/2} \sum_{j'<j_3}^{b} 4 \times 2 \times 2\alpha_{j_3-j'}^{1/2} \\
= \; & O(b^2)
\end{aligned}
$$

If $k' \in \{k, k_1, k_2\}$, the total number of different elements in set $\{k, k', k_1, k_2, k_2', k_4\}$ is three or less, the proof of $(4.4.26) = O(b^2)$ consists of the following 3 cases:

- All four terms under the expetation are correlated, which can be delt with similarly as $E(D_{12}^2)$.
- The four terms under the expectation form two independent groups with two terms in each group correlated. This situation can be dealt with similarly as (4.4.27).
- The four terms under the expectation form two independent groups and one of the groups contains three correlated terms, or the four terms form three or four independent groups. In this situation, the expectation is zero.

Therefore, $E(D_{22}^2) = O(b^{-2}n^{-2}(a))$ and so $\sqrt{b}n(a)D_{22} = o_p(1)$.

To show (4.4.24), write $P_{MSE}(\mathbf{R}^c)/N^2 - P_{MSE}(\mathbf{Y}^c) = P_{RY1} - P_{RY2}$, where

$$
\begin{aligned}
P_{RY1} &= \frac{1}{ab}\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n_i}\frac{(Z_{ijk}-p_{ij})^2-(Y_{ijk}-p_{ij})^2}{n_i^2} \\
&= \frac{1}{ab}\sum_{i,j,k}\frac{(Z_{ijk}-Y_{ijk})^2}{n_i^2}+\frac{2}{ab}\sum_{i,j,k}\frac{(Z_{ijk}-Y_{ijk})(Y_{ijk}-p_{ij})}{n_i^2} \\
P_{RY2} &= \frac{1}{ab}\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k\neq k'}^{n_i}\frac{(Z_{ijk}-p_{ij})(Z_{ijk'}-p_{ij})-(Y_{ijk}-p_{ij})(Y_{ijk'}-p_{ij})}{n_i^2(n_i-1)} \\
&= \frac{1}{ab}\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k\neq k'}^{n_i}\frac{(Z_{ijk}-Y_{ijk})(Z_{ijk'}-Y_{ijk'})}{n_i^2(n_i-1)}+\frac{2}{ab}\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k\neq k'}^{n_i}\frac{(Z_{ijk}-Y_{ijk})(Y_{ijk'}-p_{ij})}{n_i^2(n_i-1)}.
\end{aligned}
$$

The first summations in both P_{RY1} and P_{RY2} are $O_p(N^{-1}n(a)^{-1})$. The second summations in both P_{RY1} and P_{RY2} are $o_p(n(a)^{-1})$ because

$$
\left|\frac{2}{ab}\sum_{i,j,k}\frac{(Z_{ijk}-Y_{ijk})(Y_{ijk}-p_{ij})}{n_i^2}\right| \leq \frac{2}{ab}\sum_{i,j,k}\frac{|Z_{ijk}-Y_{ijk}|}{n_i^2} = O_p(N^{-1/2}n^{-1}(a)) = o_p(n^{-1}(a)),
$$

and the other one is similar.

Then we finished the proof of this lemma.

Lemma 4.4.8. *Assume X_{ijk}, $j = 1,2\dots$, is α-mixing with with $\alpha_m = O\left(m^{-5}\right)$ for all i,k, Let $P_\beta(\cdot)$, $P_\gamma(\cdot)$ be the functions defined in (4.4.13) and $\mathbf{R}^c$ be given in Lemma 4.4.7. Then as $b \to \infty$ while a remains fixed,*

(a) if n_i remain fixed,

$$\text{under } H_0(\widetilde{\beta}),\ \sqrt{b}(MS\beta_R - P_\beta(\mathbf{R}^c))/N^2 \xrightarrow{p} 0; \qquad \text{under } H_0(\widetilde{\gamma}),\ \sqrt{b}(MS\gamma - P_\gamma(\mathbf{R}^c))/N^2 \xrightarrow{p} 0$$

(b) if $n(a) \to \infty$ as b go to ∞,

$$\text{under } H_0(\widetilde{\beta}),\ n(a)\sqrt{b}(MS\beta_R - P_\beta(\mathbf{R}^c))/N^2 \xrightarrow{p} 0$$

$$\text{under } H_0(\widetilde{\gamma}),\ n(a)\sqrt{b}(MS\gamma_R - P_\gamma(\mathbf{R}^c))/N^2 \xrightarrow{p} 0.$$

Proof. Note that $MS\beta_R - P_\beta(\mathbf{R}^c) = -D_3(\mathbf{R}^c)$, $\quad MS\gamma_R - P_\gamma(\mathbf{R}^c) = \dfrac{D_3(\mathbf{R}^c) - D_4(\mathbf{R}^c)}{a-1}$, where $D_3(\mathbf{R}^c)$ and $D_4(\mathbf{R}^c)$ are similarly defined as $D_3(\mathbf{u})$ and $D_4(\mathbf{u})$ in the proof of Lemma 4.4.4 .

Note that the expression of $D_3(\mathbf{R}^c)$ and $D_4(\mathbf{R}^c)$ is very close to $D_2(\mathbf{R}^c)$. When $n(a) \to \infty$ as $b \to \infty$, the proof of $n(a)\sqrt{b}(D_4(\mathbf{R}^c)/N^2 - D_4(\mathbf{Y}^c)) = o_p(1)$ follows that of (4.4.23) (see the proof of Lemma 4.4.7) when we have the further assumption $n(a)^2 = o(b)$. Due to independence of the observations in different groups, the proof of $n(a)\sqrt{b}D_3(\mathbf{R}^c)/N^2 - D_3(\mathbf{Y}^c)) = o_p(1)$ is not much different from that of (4.4.23).

When n_i are fixed, treat $n(a)$ as a fixed number in above argument. Then we finished the proof of the lemma.

Proof of Theorem 4.2.4 Let $P_{MSE}(\mathbf{R}^c)$, $P_\beta(\mathbf{R}^c)$ and $P_\gamma(\mathbf{R}^c)$ be the functions defined by $P_{MSE}(\cdot)$ in (4.4.12), $P_\beta(\cdot)$ and $P_\gamma(\cdot)$ in (4.4.13), respectively, with argument $\mathbf{R}^c$ given in (4.4.7). By Lemma 4.4.7 and 4.4.8, we only need to consider the asymptotic distribution of the projections $n(a)\sqrt{b}(P_\beta(\mathbf{R}^c) - P_{MSE}(\mathbf{R}^c))/N^2$ and $n(a)\sqrt{b}(P_\gamma(\mathbf{R}^c) - P_{MSE}(\mathbf{R}^c)))/N^2$ when n_i go to ∞ as $b \to \infty$, under $H_0(\widetilde{\beta})$ and $H_0(\widetilde{\gamma})$, respectively. Let

$$P_1(\mathbf{R}^c) = \frac{1}{ab}\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k\neq k'}^{n_i}\frac{(R_{ijk} - \mu_{ij})(R_{ijk'} - \mu_{ij})}{n_i(n_i-1)},\ P_2(\mathbf{R}^c) = \frac{1}{ab}\sum_{i\neq i'}^{a}\sum_{j=1}^{b}(\overline{R}_{ij.} - \mu_{ij})(\overline{R}_{i'j.} - \mu_{i'j}),$$

then $P_\beta(\mathbf{R}^c)-P_{MSE}(\mathbf{R}^c)=P_1(\mathbf{R}^c)+P_2(\mathbf{R}^c)$ and $P_\gamma(\mathbf{R}^c)-P_{MSE}(\mathbf{R}^c)=P_1(\mathbf{R}^c)-P_2(\mathbf{R}^c)/(a-1)$. We also have $MS\phi_R - MSE_{\phi,R} = P_1(\mathbf{R}^c)-P_2(\mathbf{R}^c)/(a-1)$. If we can show the following

$$n(a)\sqrt{b}\left(\frac{P_1(\mathbf{R}^c)}{N^2}-P_1(\mathbf{Y}^c)\right)=o_p(1) \text{ and } n(a)\sqrt{b}\left(\frac{P_2(\mathbf{R}^c)}{N^2}-P_2(\mathbf{Y}^c)\right)=o_p(1). \quad (4.4.28)$$

The results based on ranks would follow those based on original observations applying on Y_{ijk} (see proof of Theorem 4.2.1).

$$n(a)\sqrt{b}\left(\frac{P_1(\mathbf{R}^c)}{N^2}-P_1(\mathbf{Y}^c)\right)=D_5+D_6, \qquad n(a)\sqrt{b}\left(\frac{P_2(\mathbf{R}^c)}{N^2}-P_2(\mathbf{Y}^c)\right)=D_7+D_8,$$

where

$$D_5 = \frac{n(a)}{a\sqrt{b}}\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k\neq k'}^{n_i}\frac{(Z_{ijk}-Y_{ijk})(Z_{ijk'}-Y_{ijk'})}{n_i(n_i-1)}$$

$$D_6 = \frac{2n(a)}{a\sqrt{b}}\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k\neq k'}^{n_i}\frac{(Z_{ijk}-Y_{ijk})(Y_{ijk'}-p_{ij})}{n_i(n_i-1)}$$

$$D_7 = \frac{n(a)}{a\sqrt{b}}\sum_{i\neq i'}^{a}\sum_{j=1}^{b}(\overline{Z}_{ij.}-\overline{Y}_{ij.})(\overline{Z}_{i'j.}-\overline{Y}_{i'j.})$$

$$D_8 = \frac{2n(a)}{a\sqrt{b}}\sum_{i\neq i'}^{a}\sum_{j=1}^{b}(\overline{Z}_{ij.}-\overline{Y}_{ij.})(\overline{Y}_{i'j.}-p_{i'j}).$$

It is not hard to show that D_5 and D_7 are both $O_p(b^{-1/2})=o_p(1)$. The argument for $D_6=o_p(1)$ and $D_8=o_p(1)$ are similar. We will show only the first one.

$$\begin{aligned}
&E(D_6^2)\\
= \quad & \frac{4n^2(a)}{a^2b}\sum_{i,j}\sum_{k\neq k'}^{n_i}\sum_{i_2,j_2}\sum_{k_2\neq k_2'}^{n_i}\frac{E[(Z_{ijk}-Y_{ijk})(Y_{ijk'}-p_{ij})(Z_{i_2j_2k_2}-Y_{i_2j_2k_2})(Y_{i_2j_2k_2'}-p_{i_2j_2})]}{n_{i_2}(n_i(n_i-1)n_{i_2}-1)}\\
= \quad & \frac{4n^2(a)}{a^2bN^2}\sum_{i,j}\sum_{k\neq k'}^{n_i}\sum_{i_2,j_2}\sum_{k_2\neq k_2'}^{n_i}\sum_{i_1,j_1,k_1}\sum_{i_3,j_3,k_3}E\left\{\frac{(Y_{ijk'}-p_{ij})(Y_{i_2j_2k_2'}-p_{i_2j_2})}{(n_i-1)(n_{i_2}-1)}\right.\\
&\left.\frac{1}{n_{i_2}n_i}[c(X_{i_1j_1k_1},X_{ijk})-F_{i_1j_1}(X_{ijk})][c(X_{i_3j_3k_3},X_{i_2j_2k_2})-F_{i_3j_3}(X_{i_2j_2k_2})]\right\}.
\end{aligned}$$

Note that the expectation under the summation is zero if the number of different elements in set $\{k,k',k_2,k_2',k_1,k_3\}$ is five or six. When the number of different elements in set $\{j,j_2,j_1,j_3\}$ is at most two, the summation is $O(b^{-1})$.

For $j \neq j_2 \neq j_1 \neq j_3$, the expectation is not zero only for the following two possible situations:

- All four terms under the expectation are correlated. In this situation, we must have $i = i_1 = i_2 = i_3$ and $k' = k_1 = k_2' = k_3$. A representative term in above summation is

$$\sum_{j<j_2<j_1<j_3} E\left\{\frac{(Y_{ijk'}-p_{ij})(Y_{ij_2k'}-p_{ij_2})[c(X_{ij_1k'},X_{ijk})-F_{ij_1}(X_{ijk})]}{(n_i-1)(n_i-1)}\right.$$
$$\left.\frac{1}{n_in_i}[c(X_{ij_3k'},X_{ij_2k'})-F_{ij_3}(X_{ij_2k'})]\right\}$$
$$\leq \sum_{j<j_2<j_1<j_3}^{b} 64\min(\alpha_{j_2-j}^{1/2},\alpha_{j_1-j_2}^{1/2})\frac{1}{(n_i-1)^2n_i^2}$$

  where the inequality follows from same argument as (4.4.25).

- The four terms under the expectation form two independent groups with two correlated terms in each group. The proof for this situation is similar like (4.4.27).

When the number of different elements in set $\{j,j_2,j_1,j_3\}$ is three, the summation is $O(b^{-1})$ and the proof is similar to that when $j \neq j_2 \neq j_1 \neq j_3$. Thus $D_6 = o_p(1)$.

Then the proof is done.

Proof of Proposition 4.2.5

We will show only $\widehat{\tilde{\zeta}}_1/N^4 - \tilde{\zeta}_1 \xrightarrow{p} 0$. The other one is similar and is omitted. To make symbols easier, we will write $\widehat{\zeta}_1(\mathbf{Z})$ and $\widehat{\zeta}_1(\mathbf{Y})$ as the statistics when the X_{ijk} in $\widehat{\zeta}_1$ is replaced by Z_{ijk} and Y_{ijk} respectively. Similar notations for $\widehat{\sigma}_{ijj'}(\mathbf{Z})$ and $\widehat{\sigma}_{ijj'}(\mathbf{Y})$ will be used. Note that $\widehat{\tilde{\zeta}}_1 = N^4\widehat{\zeta}_1(\mathbf{Z})$. Apply Proposition 4.2.2 on Y_{ijk}, we have $\widehat{\zeta}_1(\mathbf{Y}) - \tilde{\zeta}_1 \xrightarrow{p} 0$. So it remains to show that $\widehat{\zeta}_1(\mathbf{Z}) - \widehat{\zeta}_1(\mathbf{Y}) = o_p(1)$. Write each of the difference $Z_{ijk_1} - Z_{ijk_2}$ as $Z_{ijk_1} - Z_{ijk_2} - (Y_{ijk_1} - Y_{ijk_2}) + (Y_{ijk_1} - Y_{ijk_2})$, and note that Z_{ijk}, Y_{ijk} are uniformly bounded by 1, we have

$$\begin{aligned}
& \widehat{\sigma}_{ijj'}(\mathbf{Z}) - \widehat{\sigma}_{ijj'}(\mathbf{Y}) \\
\leq \; & \sum_{k_1 \neq k_2 \neq k_3 \neq k_4}^{n_i} \frac{4[Z_{ijk_1} - Z_{ijk_2} - (Y_{ijk_1} - Y_{ijk_2})][Z_{ij'k_1} - Z_{ij'k_2} - (Y_{ij'k_1} - Y_{ij'k_2})]}{n_i(n_i-1)(n_i-2)q(n_i-3)} \\
& + \sum_{k_1 \neq k_2 \neq k_3 \neq k_4}^{n_i} \frac{6[Z_{ijk_1} - Z_{ijk_2} - (Y_{ijk_1} - Y_{ijk_2})](Y_{ij'k_1} - Y_{ij'k_2})}{n_i(n_i-1)(n_i-2)q(n_i-3)} \\
& + \sum_{k_1 \neq k_2 \neq k_3 \neq k_4}^{n_i} \frac{6[Z_{ij'k_1} - Z_{ij'k_2} - (Y_{ij'k_1} - Y_{ij'k_2})](Y_{ijk_1} - Y_{ijk_2})}{n_i(n_i-1)(n_i-2)q(n_i-3)} \\
= \; & O_p(N^{-1/2})
\end{aligned}$$

So $\widehat{\zeta}_1(\mathbf{Z}) - \widehat{\zeta}_1(\mathbf{Y}) = O_p\left(b^{-1}N^{-1/2}\sum_{j=1}^{b}\sum_{j'=C_l(j,h)}^{C_u(j,h)}\right) = O_p(b^h/\sqrt{N}) = o_p(1)$.

Proof of Theorem 4.2.6

By Slutsky's Theorem, we only need to show

$$\frac{1}{\sqrt{N}}\mathbf{C}_a\mathbf{W}_R \xrightarrow{d} N(, \mathbf{C}_a \text{ diag}(\eta_{Y1}, \cdots, \eta_{Ya})\mathbf{C}_a') \tag{4.4.29}$$

$$\widehat{\eta}_{Ri}/N^2 - \eta_{Yi} = o_p(1), \tag{4.4.30}$$

where $\eta_{Yi} = \lim_{b\to\infty} N\text{Var}\left((\overline{Y}_{i..} - E(\overline{Y}_{i..})\right)$. Let $\widehat{\eta}_{Yi}$ be similarly defined as $\widehat{\eta}_{Ri}$ with R replaced by Y. The consistency of $\widehat{\eta}_{Yi}$ for η_{Yi} is easily seen because $\mathbf{Y}_{ik} = (Y_{i1k}, \cdots, Y_{ibk})'$ are independent vectors. To show (4.4.30), we only need to show $\widehat{\eta}_{Ri}/N^2 - \widehat{\eta}_{Yi} = o_p(1)$.

$$\begin{aligned}
&\frac{bn_i(n_i-1)}{n}\left(\widehat{\eta}_{Ri}/N^2 - \widehat{\eta}_{Yi}\right) \\
= &\sum_{j=1}^{b}\sum_{j'=-b^{\frac{1}{4}}}^{b^{\frac{1}{4}}}\sum_{k=1}^{n_i}\left[\left(Z_{ijk}-\overline{Z}_{ij.}\right)\left(Z_{i(j+j')k}-\overline{Z}_{i(j+j').}\right)-\left(Y_{ijk}-\overline{Y}_{ij.}\right)\left(Y_{i(j+j')k}-\overline{Y}_{i(j+j').}\right)\right] \\
= &\sum_{j=1}^{b}\sum_{j'=-b^{\frac{1}{4}}}^{b^{\frac{1}{4}}}\sum_{k=1}^{n_i}\left[Z_{ijk}-\overline{Z}_{ij.}-Y_{ijk}+\overline{Y}_{ij.}\right]\left[Z_{i(j+j')k}-\overline{Z}_{i(j+j').}-Y_{i(j+j')k}+\overline{Y}_{i(j+j').}\right] \\
&+\sum_{j=1}^{b}\sum_{j'=-b^{\frac{1}{4}}}^{b^{\frac{1}{4}}}\sum_{k=1}^{n_i}\left[Z_{ijk}-\overline{Z}_{ij.}-Y_{ijk}+\overline{Y}_{ij.}\right]\left[Y_{i(j+j')k}+\overline{Y}_{i(j+j').}\right] \\
&+\sum_{j=1}^{b}\sum_{j'=-b^{\frac{1}{4}}}^{b^{\frac{1}{4}}}\sum_{k=1}^{n_i}\left[Z_{i(j+j')k}-\overline{Z}_{i(j+j').}-Y_{i(j+j')k}+\overline{Y}_{i(j+j').}\right]\left[Y_{ijk}+\overline{Y}_{ij.}\right]
\end{aligned}$$

The first term is $O(2n_ib^{\frac{1}{4}}/n)$ and the second and the third terms are both $O(\sqrt{b}n_i/\sqrt{n})$. So $\widehat{\eta}_{Ri}/N^2 - \widehat{\eta}_{Yi} = O_p\left(\frac{\sqrt{n}}{(n_i-1)b^{\frac{1}{4}}}\right) = o_p(1)$.

To show (4.4.29), we introduce the following notation. Let $\widehat{F}_{ij}(x) = n_i^{-1}\sum_{k=1}^{n_i} c(X_{ijk},x)$, $\overline{\widehat{F}}_{i.}(x) = b^{-1}\sum_{j=1}^{b}\widehat{F}_{ij}(x)$, and $\overline{\widehat{\mathbf{F}}} = (\overline{\widehat{F}}_{1.},\cdots,\overline{\widehat{F}}_{a.})'$. Then

$$\int \widehat{H}\,d\overline{\widehat{F}}_{i.} = \frac{1}{b}\sum_{j=1}^{b}\frac{1}{n_i}\sum_{k=1}^{n_i}\widehat{H}(X_{ijk}) = N^{-1}\left(\overline{R}_{i..}-\frac{1}{2}\right).$$

So $\mathbf{W}_R = N\int \widehat{H}\,d\overline{\widehat{\mathbf{F}}} + \frac{1}{2}\cdot\mathbf{1}_a$, where $\mathbf{1}_a$ is a a dimentional vector of one's. Because $\mathbf{C}_a$ is a contrast matrix,

$$\text{under } \widetilde{H}(A), \quad \frac{1}{\sqrt{N}}\mathbf{C}_a\mathbf{W}_R = \sqrt{N}\mathbf{C}_a\int \widehat{H}\,d\overline{\widehat{\mathbf{F}}} = \sqrt{N}\mathbf{C}_a\int \widehat{H}\,d(\overline{\widehat{\mathbf{F}}}-\overline{\mathbf{F}})$$

It suffices to find the asymptotic distribution of $\sqrt{N}\int \widehat{H}\,d(\overline{\widehat{\mathbf{F}}}-\overline{\mathbf{F}})$. Note that

$$\begin{aligned}
\sqrt{N}\int H\,d(\overline{\widehat{\mathbf{F}}}-\overline{\mathbf{F}}) &= \sqrt{N}\left(\overline{Y}_{1..}-E(\overline{Y}_{1..}),\cdots,\overline{Y}_{a..}-E(\overline{Y}_{a..})\right)' \\
&= \sqrt{N}[\mathbf{W}_Y - E(\mathbf{W}_Y)],
\end{aligned}$$

where $\mathbf{W}_Y = (\overline{Y}_{1..}, \cdots, \overline{Y}_{a..})'$

Apply the proof for Theorem 4.2.3 on Y_{ijk}, we get $\sqrt{N}(\overline{Y}_{i..} - E(\overline{Y}_{i..})) \xrightarrow{d} N(0, \eta_{Yi})$. Note that $\eta_{Yi} < \infty$ is guranteed by the α-mixing condition. By independence of the observations from different group, we have

$$\sqrt{N}[\mathbf{W}_Y - E(\mathbf{W}_Y)] \xrightarrow{d} N_a(\mathbf{0}, \mathbf{V}_Y) \text{ as } a \to \infty, c \to \infty \, ,$$

where $\mathbf{V}_Y = diag(\eta_{Y1}, \cdots, \eta_{Ya})$. Apply Continuous Mapping Theorem, we have $\sqrt{N}\mathbf{C}_a\mathbf{W}_Y \xrightarrow{d} N_{a-1}(\mathbf{0}, \mathbf{C}_a\mathbf{V}_Y\mathbf{C}_a')$. Therefore, the proof of the theorem will be done if we can show

$$\sqrt{N}\int(\widehat{H} - H)\, d(\widehat{\overline{\mathbf{F}}} - \overline{\mathbf{F}}) \xrightarrow{p} \mathbf{0}. \tag{4.4.31}$$

This is shown in Lemma 4.4.6.

Chapter 5

Summary and Future Research

5.1 Summary

In conclusion, this thesis is concerned with fully nonparametric inference in designs with large number of factor levels and possibly only very small group sizes. This situation includes the so-called *sparse* testing problem where a faint signal may be present among a large number of sources, and is relevant to such modern problems as the outbreak of infectious diseases and terrorism monitoring. More common applications can be found in agricultural screening trials that often involve many different treatments (cultivars, inbred lines, etc.) with limited replication and mostly non-normal response variable, and in microarray data analysis where the number of genes is very large from a small number of subjects. Moreover, this area serves as link to a number of difficult problems in lack-of-fit testing in nonparametric regression and ANCOVA.

The first part of my thesis considers the use of rank statistics for testing hypotheses in balanced and unbalanced, homoscedastic and heteroscedastic one-way and two-way ANOVA models when the number of levels of at least one factor is large. The second part of my thesis deals with nonparametric modeling and test procedures for multifactor (also unbalanced and heteroscedastic) designs, with possibly very small group sizes, where at least some of the factors have a large number of levels. Asymptotic results based on both original observations and their (mid-)ranks are presented for the same general setting. The third part of my thesis presents a fully nonparametric method for testing hypotheses in longitudinal or functional data, and for evaluating the effect of several crossed factors on the curve and their interactions with time. The asymptotics, which rely on the large number of measurements per curve (subject) and not on large group sizes, hold under the general assumption of α-mixing without specifying the covariance structure, and do not require the measurements to be continuous or homoscedastic. A competing set of (mid-)rank procedures is also developed. The procedures in all three parts can be applied to both continuous and discrete ordinal observations. The rank tests are robust to outliers, invariant under monotone transformations, and do not require any restrictive moment conditions. Simulation studies reveal that the (mid-)rank procedures outperform those based on the original observations in all non-normal situations in terms of both type I and type II error rates and they do not loose much power when normality holds. Applications to several data sets are analyzed.

In summary, this thesis generalizes, both from the methodological and scope of applications points of view, previous work by Boos and Brownie (1995), Akritas and Arnold (2000), Bathke (2002), Akritas and Papadatos (2003), Wang and Akritas (2002, 2003), and is closely related to Portnoy (1984) and Li, Lindsay and Waterman (2003). The analysis involved is quite challenging and the problems treated concern novel statistical methodology which is immediately applicable and addresses immediate needs of current interest in the statistical community.

5.2 Future Research

One direct extension of current research is to generalize the methodology for multifactor designs to dependent data with arbitrary but fixed number of factors when the number of repeated measurements per subject is either large or small. Moreover, though the results for dependent data in this thesis provide powerful tools for functional data with identical sampling points (which often happen when the response is automatically recorded like seismic recordings etc.), it is tempting to consider inference for longitudinal or functional data when the subjects are observed at possibly different time points as often is the case. In such situation, factors other than time (repeated measurements) having large or small number of factor levels can both be included.

My future research will deal also with problems in regression and ANCOVA designs. Thinking of covariates as factors with many levels, it is possible to make a connection between testing in designs with a large number of factor levels and the areas of nonparametric lack-of-fit testing in regression and testing for covariate and covariate-adjusted factor effects in the nonparametric ANCOVA model of Akritas, Arnold and Du (2000). The connection with nonparametric ANCOVA has been made in Akritas, Antoniou and Wang (2003) and in Wang (2003), while the connection with lack-of-fit testing in nonparametric regression has been made in Wang, Akritas and VanKeilegom (2003). The last three mentioned papers considered designs with only one covariate and test procedures based only on the exact observations. Using the general multifactor formulation of this thesis, it is promising to extend these to designs with several covariates and will also develop rank procedures which, as my simulations indicate, are much more powerful for non-normal data.

It should be mentioned that rank-based inference and smoothing methods are two areas that, except for sharing the name ’nonparametrics’, are quite distinct both in terms of techniques and applications. Using their fully nonparametric model for ANCOVA, Akritas, Arnold and Du (2000) merged the two areas but considered only testing for covariate-adjusted factor effects (i.e. not for the other common hypotheses of no covariate or covariate-factor interaction effects). Tsangari and Akritas (2003) extended their results to up to three covariates, but curse of dimensionality effects prevents further extension. Preliminary evidence indicates that test procedures which utilize the aforementioned connection between this thesis and testing in ANCOVA will be much less sensitive to curse of dimensionality effects. Thus, future research in this direction expect to be able to produce easy-to-use procedures, for all common hypotheses, which handle more than three covariates in a completely nonparametric way for a very wide class of designs including both independent and dependent data.

APPENDICES

Appendix A

Splus Code for Two-way Heteroscedastic ANOVA

```
## This program is for two-way Heteroscedastic ANOVA
## store the data, they are stored in "data" like this:
# 1 1 1 x111
# 1 1 2 x112
# 1 2 1 x121
# 1 2 2 x122
# 1 2 3 x123
# x_{ijk}, k=1, ..., n_{ij} are iid from distribution defined as a function distr
# The overall rank is calculated and put at the last colum
# coln=4 or 5. if coln=4, test use org data ;  if coln=5, test use rank, 5 is default.
#  mn is a matrix with (i,j) element n_{ij}

# calculate the u-stat of vector x
sigij4_function(x){
        nij_length(x)
sigmaij4_ 0
for (m1 in 1:nij){
 for (m2 in 1:nij){
   for (m3 in 1:nij){
     for (m4 in 1:nij){
```

```
     flag_ (m1!=m2)&(m1 !=m3)&(m1 !=m4) &(m2 !=m3) &(m2 !=m4) & (m3 !=m4)
     sigmaij4_sigmaij4+ (flag==T)*(x[m1]-x[m2])^2 * (x[m3]-x[m4])^2
      }
    }
   }
  }
 sigmaij4_sigmaij4/(nij*(nij-1)*(nij-2)*(nij-3) )
 sigmaij4
}

Heter.test.2_function(data,r,c, mn,coln=5){
d1_data[, 1]
d2_data[, 2]
        data<-cbind(data, cbind(data, rank(data[, 4]))
ranks_data[, coln]

Rij_as.matrix(tapply(ranks, list(d1, d2), mean) )
   # matrix with \bar{R}_{ij.} as the (i,j) element
Ri_apply(Rij, 1, mean)  # returns a vector (\bar{R}_{1..}, ..., \bar{R}_{r..})
Rj_apply(Rij, 2, mean)
Rim_kronecker(Ri, t(as.vector(rep(1,c))) )
  # a r by c matrix with all elements of the ith row same as \bar{R}_{i..}
Rjm_kronecker(t(Rj), as.vector(rep(1,r)))
  # a r by c matrix with all elements of the ith column same as \bar{R}_{.j.}
## calculate test statistics and  \hat{\sigma}_{ij}^2= S_{ij, R}^2

MSalpha_ c* sum( (Ri-mean(Ri))^2 ) /(r-1)
MSgamma_ sum((Rij-Rim-Rjm+mean(Ri) )^2 )/((r-1)*(c-1) )
```

```
sum2ij_as.matrix(tapply(ranks^2, list(d1, d2), sum) )
  # returns a matrix with the (i, j) element  \sum_{k=1}^n_{ij} R_{ijk}^2
hatsigmaij2_sum2ij/(mn-1) - mn* Rij^2 /(mn-1)
  # returns a matrix with the (i, j) element S_{ij, R}^2, the estimate of
  # \sigma_{ij}^2
sigmaij2_(coln==4)*hatsigmaij2 +(coln==5)*hatsigmaij2/(sum(mn))^2
MSE3_mean(hatsigmaij2/mn )

Falpha_ MSalpha/MSE3
Fgamma_MSgamma/MSE3
TSalpha_sqrt(r)*(Falpha-1)
TSgamma_sqrt(r)*(Fgamma-1)

### estimate \sigma_{ij}^4 using U-statistic

Sigma4_as.matrix(tapply(ranks, list(d1, d2), sigij4) )
Sigmaij4_(coln==4)*Sigma4+(coln==5)*Sigma4/(sum(mn))^4
phi4_mean( Sigmaij4/(mn*(mn-1)) )
eta4_ mean( ( apply(sigmaij2/mn, 1, sum ) )^2 )/c- mean(Sigmaij4/mn^2)
nu42_ (coln==4)*MSE3 +(coln==5)*MSE3/(sum(mn))^2

varTSalpha_2*(phi4+eta4)/(c*nu42^2)
varTSgamma_2*(phi4*(c-1)^2+eta4)/(c*nu42^2)

palpha_ 1-pnorm( TSalpha/sqrt(varTSalpha) )
pgamma_ 1- pnorm(TSgamma/sqrt(varTSgamma) )
list(palpha=palpha, pgamma=pgamma)
}
```

Appendix B

Splus Code for Three-way Heteroscedastic ANOVA

```
#______________________________________________________________________________#
# Heter.test3 is for unbalanced heteroscedastic case 3-way ANOVA
# Assume factors A and C have large number of levels and B have
# small number of levels. Other situations are analygous.
# The output is a list reporting P-values for testing
# main factor A effect, B effect, A,C interaction, A, B interaction,
# and A,B, and C interaction.

# Arguments in the function:
# a, b, and c are the number of levels for factor A, B, and C respectively.
# mn is an array with (i,j, k) element n_{ijk}, the sample size for cell (i, j, k)
# The data are stored in matrix "data" like this:
# 1 1 1 1 x_1111
# 1 1 1 2 x_1112
# 1 1 2 1 x_1121
# 1 1 2 2 x_1122
# 1 1 2 3 x_1123
# 1 2 1 1 x_1211
# 1 2 1 2 x_1212
```

```
# 1 2 2 1 x_1221

# 1 2 2 2 x_1222

# 1 2 2 3 x_1223

# coln=5 or 6. if coln=5, the test uses original data ;

#              if coln=6, the test uses (mid-)rank; 6 is default.

Heter.test3<-function(data,a, b, c, mn,coln=6){

d1<-data[, 1]

d2<-data[, 2]

        d3<-data[, 3]

        if (coln==6) data<- cbind(data, rank(data[, 5]))

ranks<-data[, coln]


Rijk<-as.array(tapply(ranks, list(d1, d2, d3), mean) )

   # array with \bar{R}<-{ijk.} as the (i,j, k) element

        Rij<-as.matrix(apply(Rijk, c(1, 2), mean))

                                # returns a matrix with (i, j) element \wtR<-{ij..}

        Rik<-as.matrix(apply(Rijk, c(1, 3), mean))

                                # returns a matrix with (i, k) element \wtR<-{i.k.}

        Rjk<-as.matrix(apply(Rijk, c(2, 3), mean))

                                # returns a matrix with (j, k) element \wtR<-{.j.k.}

Ri<-apply(Rij, 1, mean)  # returns a vector (\wtR<-{1...}, ..., \wtR<-{a...})

Rj<-apply(Rij, 2, mean)  # returns a vector (\wtR<-{.1..}, ..., \wtR<-{.b..})

        Rk<-apply(Rik, 2, mean) # returns a vector (\wtR<-{..1.}, ..., \wtR<-{..c.})

##_____________________ calculate MSE______________________________#

sum2ijk<-as.array(tapply(ranks^2, list(d1, d2, d3), sum) )

        #returns an array with the (i, j, k) element  \sum_{m=1}^n_{ijk} R_{ijkm}^2

hatsigmaijk2<-sum2ijk/(mn-1) - mn* Rijk^2 /(mn-1)
```

```
# returns an array with the (i, j, k) element S<-{ijk, R}^2,
sigmaijk2<-(coln==5)*hatsigmaijk2 +(coln==6)*hatsigmaijk2/(sum(mn))^2
        # an array with the (i, j, k) element being the estimate of  \sigma_{ijk}^2
MSE3<-mean(hatsigmaijk2/mn )
         ### estimate \sigma<-{ijk}^4 using U-statistic
Sigma4<-as.array(tapply(ranks, list(d1, d2, d3), sigijk4) )
Sigmaijk4<-(coln==5)*Sigma4+(coln==6)*Sigma4/(sum(mn))^4
          # returns an array with [i, j, k] element being the unbiased
          # estimate of \sigma<-{ijk}^4
        # nu42<-(coln==5)*MSE3 +(coln==6)*MSE3/(sum(mn))^2
             # it is an estimate of the sqrt of $\sigma^4$ in the
             #denominator of the \asd varaince
nu44<- (mean(sigmaijk2/mn))^2-mean((sigmaijk2^2-Sigmaijk4)/mn^2)/(a*b*c)
   # unbiased estimate of $\sigma^4$ in the denominator of the \asd varaince


##____________________For test of H<-0(A) and H<-0(AB)______________#
Rim<-kronecker(Ri, t(as.vector(rep(1,b))) )
  # a a by b matrix with all elements of the ith row same as \bar{R}<-{i..}
Rjm<-kronecker(t(Rj), as.vector(rep(1,a)))
  # a a by b matrix with all elements of the ith column same as \bar{R}<-{.j.}
## calculate test statistics and  \hat{\sigma}<-{ij}^2= S<-{ij, R}^2
MSA<- b*c* sum( (Ri-mean(Ri))^2 ) /(a-1)
MSAB<- c*sum((Rij-Rim-Rjm+mean(Ri) )^2 )/((a-1)*(b-1) )
        FA<-MSA/MSE3
FAB<-MSAB/MSE3
TSA<-sqrt(a)*(FA-1)
```

```
TSAB<-sqrt(a)*(FAB-1)


        tmptau1<- 2*mean( ( apply(sigmaijk2/mn, 1, mean ) )ˆ2 )
        tau1<-tmptau1-2*mean(sigmaijk2ˆ2 /mnˆ2)/(b*c) +2*mean(Sigmaijk4/mnˆ2)/(b*c)
        #correct for bias by using the unbiased estimate of \sigma<-{ijk}ˆ4
        tmptau2<- ( apply(sigmaijk2/mn, c(1, 2), mean ) )ˆ2 -
                  mean(sigmaijk2ˆ2/mnˆ2-Sigmaijk4/mnˆ2)/c
        tau2<-2*(b-2)/(b-1)ˆ2 * mean(tmptau2 )
varTSA<-tau1/nu44
varTSAB<-(tau1/(b-1)ˆ2+tau2)/nu44


pA<- 1-pnorm( TSA/sqrt(varTSA) )
pAB<- 1- pnorm(TSAB/sqrt(varTSAB) )


##____________________For test of H<-0(AC) ________________________#
Rimk<-kronecker(Ri, t(as.vector(rep(1,c))) )
  # a a by b matrix with all elements of the ith row same as \bar R_{i..}
Rkm<-kronecker(t(Rk), as.vector(rep(1,a)))
# a a by b matrix with all elements of the ith column same as \bar R_{.j.}
        MSAC<- b*sum((Rik-Rimk-Rkm+mean(Ri) )ˆ2 )/((a-1)*(c-1) )
FAC<-MSAC/MSE3
TSAC<-sqrt(a*c)*(FAC-1)
        mu2<-2*mean(( apply(sigmaijk2/mn, c(1, 3), mean ) )ˆ2 )-
             2*mean(sigmaijk2ˆ2/mnˆ2-Sigmaijk4/mnˆ2)/b
        mu1<-2*mean(Sigmaijk4/(mnˆ2*(mn-1)) )/b
        tau4<-mu1+mu2
        varTSAC<-tau4/nu44
```

```
        pAC<- 1- pnorm(TSAC/sqrt(varTSAC) )

##____________________For test of H<-0(ABC)________________________#
        MSABC<-(sum(Rijk^2)-b*sum(Rik^2)-a*sum(Rjk^2)+
               a*b*sum(Rk^2) )/((a-1)*(b-1)*(c-1) ) -   MSAB/(c-1)
        FABC<-MSABC/MSE3
TSABC<-sqrt(a*c)*(FABC-1)
mu3<-2*(b-2)*mean(Sigmaijk4/(mn^2) )/((b-1)^2)
        tau5<-mu1+mu2/(b-1)^2 + mu3
        varTSABC<-tau5/nu44
        pABC<- 1- pnorm(TSABC/sqrt(varTSABC) )

#________________________For test of H<-0(B) ________________________#
        CB<-cbind(as.vector(rep(1, b-1)), -diag(rep(1, b-1)))
        etaj<-apply(hatsigmaijk2/mn, 2, mean)*sum(mn)/(a*c)
        V<-diag(etaj)
        W<-as.vector(Rj)
         TSB<-sum(mn)*t(W) %*% t(CB) %*% solve(CB%*% V %*% t(CB))%*% CB %*% W
        pB<-1-pchisq(TSB, (b-1))

        list(pA=pA, pB=pB, pAC=pAC, pAB=pAB,  pABC= pABC)
}
# calculate the u-stat of vector x
#
sigijk4<-function(x){
        nijk<-length(x)
sigmaijk4<- 0
```

```
for (m1 in 1:nijk){
 for (m2 in 1:nijk){
   for (m3 in 1:nijk){
     for (m4 in 1:nijk){
     flag<- (m1!=m2)&(m1 !=m3)&(m1 !=m4) &(m2 !=m3) &(m2 !=m4) & (m3 !=m4)
     sigmaijk4<-sigmaijk4+ (flag==T)*(x[m1]-x[m2])^2 * (x[m3]-x[m4])^2
      }
    }
   }
  }
 sigmaijk4<-sigmaijk4/(4*nijk*(nijk-1)*(nijk-2)*(nijk-3) )
 sigmaijk4
}
```

Appendix C

Splus Code for Hypotheses Tests in Functional Data

```
## store the data, they are stored in data like this:
# 1 1 1 x111
# 1 1 2 x112
# 1 2 1 x121
# 1 2 2 x122
# 1 2 3 x123
# x<-{ijk}, k=1, ..., n<-{ij} are iid from distribution defined as a function distr
# The overall rank is calculated and put at the last colum
# coln=4 or 5. if coln=4, test use org data ;  if coln=5, test use rank, 5 is default.
#  mn is a matrix with (i,j) element n<-{ij}

# eu is a function to calaculate residue x<-{ijk}-\bar{x}<-{ij.}
eu<-function(x, coln, Rij, Rik){
        d1<-x[ 1]
  d2<-x[ 2]
        d3<-x[ 3]
        R<-x[ coln]
        # e<- R-Rij[d1,d2]+Rik[d1, d3]-mean(Rij[d1,])
        u<- R-Rij[d1,d2]
```

```
        result<-rbind(u,u)   #used to be rbind(e,u)
        result
        }
# returns a vector with two elements. The first one is e<-{ijk} and the second is u<-{ijk}

# in the following function, u, d1, d2 are all b*sum(mn) dimentional vectors,
#same as the last column of dat

#  Unbiased estimate of \sigma<-{ijj1}ˆ2
# calculate the u-stat of vectors x=(x_1, x_2, ..., x_{ni}) y=(y_1, y_2, ..., y_{ni}),
# where $X_j$ and $Y_j$ are correlated, but $X_j$ and $Y_j1$ are indept if j \ne j1.
# $\sum_{k1 \ne k2 \ne k3 \ne k4} (x_{k1}-x_{k2)})(y_{k1}-y_{k2)})
# (x_{k3}-x_{k4)}) (y_{k3}-y_{k4)})$
# unbiased est. of  4*ni*(ni-1)*(ni-2)*(ni-3) [E(X_{ijk}-\mu_{ij} u_{ij1k}) ]ˆ2
#fun.sigijj12$sigmaijj12 will give the unbiased est of $\sigma_{ijj1}ˆ2$
#fun.sigijj12$ssijj1 will give the unbiased est of $\sigma_{ijj} \sigma_{ij_1j_1}$.

fun.sigijj12<-function(x, y){
        ni<-length(x)
sigmaijj12<- 0
        ssijj1<- 0
for (m1 in 1:ni){
 for (m2 in 1:ni){
   for (m3 in 1:ni){
     for (m4 in 1:ni){
     flag<- (m1!=m2)&(m1 !=m3)&(m1 !=m4) &(m2 !=m3) &(m2 !=m4) & (m3 !=m4)
     sigmaijj12<-
```

```
sigmaijj12+ (flag==T)* ( (x[m1]-x[m2])*(y[m1]-y[m2])* (x[m3]-x[m4])*(y[m3]-y[m4]) )
             ssijj1<-ssijj1+(flag==T)*  (x[m1]-x[m2])^2*(y[m3]-y[m4])^2
      }
    }
   }
  }
 sigmaijj12<-sigmaijj12/(4*ni*(ni-1)*(ni-2)*(ni-3) )
        ssijj1 <- ssijj1/(4*ni*(ni-1)*(ni-2)*(ni-3) )
 result<- list(sigmaijj12=sigmaijj12, ssijj1=ssijj1)
 result
}

###
# tauphi2 have five components corresponds to partial sum
# up to b^c(1/4, 1/3, 2/5, 9/20)

taufun<-function(u, ranks, d1, d2, d3, a, b, mcon, coln){
   R<-ranks
   usigmaijj1<-usigmaijj12<-array(rep(0, a*b*b), c(a, b, b))
   usigma2<-0
   us <-numeric()
   for (i in 1:a){
     us[i]<-0
    for (j in 1:b){
    for (j1 in 1:b){
    usigmaijj1[i, j, j1]<-sum(u[((d1==i)&(d2==j))] * u[((d1==i)&(d2==j1))])  /(mn[i]-1)
# unbiased est. of $\sigma<-{ijj1}$
```

```
# usigmaijj12 is sample covariance:
# $\frac{1}{n_i-1}  \sum_{k=1}^{n_i} (X_{ijk}-\bar X_{ij.})(X_{ij_1k}-\bar X_{ij_1.})$
# it gives unbiased estimate of E[(X_{ijk}-\mu_{ij}) (X_{ij<-1k}-\mu_{ij_1}) ]
# and asymptotically unbiased estimate of E(e_{ijk}e_{ij’k})
    x<-R[((d1==i)&(d2==j))]
    y<-R[((d1==i)&(d2==j1))]
    usigmaijj12[i, j, j1]<-sigijj12jack(x ,  y)
   }  # end of j1
   usigma2<-usigma2+usigmaijj1[i, j, j]/mn[i]
   us[i]<-us[i]+usigmaijj1[i, j, j]/mn[i]
 }
 }

  usigma2<-usigma2/(a*b)
 # consistent est. of \frac{1}{ab} \sum_{ij} \sigma_{ijj} /n<-i  i.e. E(MSE_{\phi})
  EMSEphi<-usigma2

#  ucsi1<-2*sum(apply(usigmaijj12, 1, sum) /(mn*(mn-1))  )/(a^2*b)
       # estimate of $\zeta<-1$ in thm 2.1
#  tmpucsi2<-sum( (apply(usigmaijj1/mn, c(2, 3), sum) )^2 ) - sum(usigmaijj1^2  / mn^2)
#  ucsi2<-2*tmpucsi2/(a^2*b)  #estimate of $\zeta<-2$ in thm 2.1.
     #  note when $i\ne i’$, they are indept. so we can use  usigmaijj1 directly.

pucsi1<-2*apply(usigmaijj12 /(mn*(mn-1)), c(2, 3), sum)   /(a^2*b)
pucsi2<-2*((apply(usigmaijj1/mn, c(2, 3), sum) )^2 -
        apply(usigmaijj1^2  / mn^2, c(2, 3), sum) )/(a^2*b)
ucsi1<-sum(pucsi1)  # estimate of $\zeta<-1$ in thm 2.1 using all sum
```

```
ucsi2<-sum(pucsi2)  #estimate of $\zeta<-2$ in thm 2.1. using all sum


psum<-apply(usigmaijj1/mn, c(2, 3), sum)


zeta1<-zeta2<-partsum<-numeric()
#mc<-c(1/4, 1/3, 2/5, 9/20)
mc<-mcon
ll<-0
for (l3 in mc[-length(mc)]){
  ll<-ll+1
  tu1<-tu2<-parts0<-0
 for (j0 in 1:b){
# for (j2 in seq(round(-b^l3 ), round(b^l3))){
for (j2 in seq(-l3, l3)){
          if ((j0+j2>0)& (j0+j2<=b)  )  {
tu1<-tu1+pucsi1[j0, j0+j2]
tu2<-tu2+pucsi2[j0, j0+j2]
parts0<- parts0+ psum[j0, j0+j2]
           }
  }
  }
zeta1[ll]<-tu1
zeta2[ll]<-tu2
partsum[ll]<-parts0
}

 zeta1[length(mc)]<- ucsi1
```

```
 zeta2[length(mc)]<- ucsi2
 partsum[length(mc)]<- sum(usigmaijj1/mn)
 esigma2<- a*b*usigma2/(a*(b-1))- partsum/(a*b*(b-1))
# estimate of E(MSE) if use obs
  N<-sum(mn)*b
tauphi2<-taubeta2<- taugamma2<-numeric()

 tauphi2<-zeta1 + zeta2/(a-1)^2
 taubeta2<-zeta1 + zeta2
 taugamma2<-zeta1 + zeta2/(a-1)^2
  if (coln==5) {
      tauphi2<-tauphi2/N^4
      taubeta2<-taubeta2/N^4
      taugamma2<-taugamma2/N^4
      esigma2<-esigma2/N^2
      usigma2<-usigma2/N^2
     }
  tauphi2[(tauphi2 <=0)] <- 10^(-15)
  taubeta2[(taubeta2 <=0)] <-10^(-15)
  taugamma2[(taugamma2 <=0)] <-10^(-15)
  list( EMSE=esigma2, tauphi2=tauphi2, taubeta2=taubeta2,
    EMSEphi=usigma2,taugamma2=taugamma2, zeta1=zeta1, zeta2=zeta2)
  }

# calculate the u-stat of vector x=(x<-1, x<-2, \cdots, x<-{nij})  where
# X<-i are iid with variance \sigma^2. This u-stat will give
# unbiased estimate of  \sigma^4
```

```
# : \sum<-{k1 \ne k2 \ne k3 \ne k4} (x<-{k1}-x<-{k2)})^2 (x<-{k3}-x<-{k4)})^2
#
sigij4<-function(x){
        nij<-length(x)
sigmaij4<- 0
for (m1 in 1:nij){
 for (m2 in 1:nij){
   for (m3 in 1:nij){
     for (m4 in 1:nij){
     flag<- (m1!=m2)&(m1 !=m3)&(m1 !=m4) &(m2 !=m3) &(m2 !=m4) & (m3 !=m4)
     sigmaij4<-sigmaij4+ (flag==T)*(x[m1]-x[m2])^2 * (x[m3]-x[m4])^2
      }
    }
   }
  }
 sigmaij4<-sigmaij4/(4*nij*(nij-1)*(nij-2)*(nij-3) )
 sigmaij4
}

Heter.test<-function(data,a,b, mn, mcon, coln=5, Ca=cbind(as.vector(rep(1, a-1)),
     -diag(a-1)) ){
        N<-sum(mn)*b
d1<-data[, 1]
d2<-data[, 2]
d3<-data[, 3]
ranks<-data[, coln]
Rij<-as.matrix(tapply(ranks, list(d1, d2), mean) )
```

```
# matrix with \bar{R}<-{ij.} as the (i,j) element
Rik<-as.matrix(tapply(ranks, list(d1, d3), mean) )
# matrix with \bar{R}<-{i.k} as the (i,j) element  might have NA if unbalanced.
Rik<-replace(Rik, is.na(Rik), 0)    # replace NA's in matrix Rik by 0.
# note: after replace, the number of rows still correct, but the number of columns
# would be all same as max<-i n<-i instead of n<-i columns for the ith row.

Ri<-apply(Rij, 1, mean)  # returns a vector (\wtR<-{1..}, ..., \wtR<-{a..})
Rj<-apply(Rij, 2, mean)

Rim<-kronecker(Ri, t(as.vector(rep(1,b))) )
  # a a by b matrix with all elements of the ith row same as \wtR<-{i..}
Rjm<-kronecker(t(Rj), as.vector(rep(1,a)))
  # a a by b matrix with all elements of the ith column same as \wtR<-{.j.}

## calculate test statistics

MSbeta<- a* sum( (Rj-mean(Rj))^2 ) /(b-1)
MSgamma<- sum((Rij-Rim-Rjm+mean(Ri) )^2 )/((a-1)*(b-1) )
MSphi<- sum((Rij-Rjm)^2 )/((a-1)*b  )

tmp1<-tapply(ranks^2, d1, sum)
# returns a vector  with ith element \sum<-{j=1}^b \sum<-{k=1}^{n<-i} R<-{ijk}^2
tmp2<-sum( tmp1/(mn*(mn-1)) )     # \sum<-{i,j, k}\frac{X<-{ijk}^2}{n<-i(n<-i-1)}
MSEphi<- tmp2/(a*b)   - mean( apply(Rij^2, 1, mean)/(mn-1) )
MSE<- MSEphi *b/(b-1) - mean(apply(Rik^2, 1, sum)/(mn*(mn-1)) ) *b/(b-1) +
              mean(Ri^2/(mn-1)) *b/(b-1)
```

```
## another way to calculate MSE
#   MSE_ tmp2/(a*(b-1)) - mean(apply(Rikˆ2, 1, sum)/(mn*(mn-1)) ) *b/(b-1) -
#        mean(apply((Rij-Rim)ˆ2, 1, sum)/(mn-1) )/(b-1)

Fbeta<-MSbeta/MSE
Fgamma<-MSgamma/MSE
Fphi<-MSphi/MSEphi
Dbeta<-MSbeta-MSE
Dgamma<-MSgamma-MSE
Dphi<-MSphi-MSEphi
## and

        euijk<-apply(data, 1, e<-function(x) eu(x, coln, Rij, Rik))
e<-euijk[1,]  # returns the e<-{ijk} as a vector, same as ranks structure
u<-euijk[2,]  # returns the u<-{ijk} as a vector, same as ranks structure
vars<-taufun(u, ranks, d1, d2, d3, a, b, mcon, coln)

TSbeta<-sqrt(b)*(Fbeta-1)
TSgamma<-sqrt(b)*(Fgamma-1)
TSphi<-sqrt(b)*(Fphi-1)
DSbeta<-sqrt(b)*Dbeta
DSgamma<-sqrt(b)*Dgamma
DSphi<-sqrt(b)*Dphi
        if (coln==5) {
        DSbeta<-DSbeta/Nˆ2
DSgamma<-DSgamma/Nˆ2
        DSphi<-DSphi/Nˆ2
```

```
        }

varTSbeta<-vars$taubeta2/vars$EMSEˆ2
varTSgamma<-vars$taugamma2/vars$EMSEˆ2
varTSphi<-vars$tauphi2/vars$EMSEphiˆ2

pbeta<- 1-pnorm( TSbeta/sqrt(varTSbeta) )
pphi<- 1-pnorm( TSphi/sqrt(varTSphi) )
pgamma<- 1- pnorm(TSgamma/sqrt(varTSgamma) )

Dpbeta<- 1-pnorm( DSbeta/sqrt(vars$taubeta2) )
Dpphi<- 1-pnorm( DSphi/sqrt(vars$tauphi2) )
Dpgamma<- 1- pnorm(DSgamma/sqrt(vars$taugamma2) )

# test for main group effect
        Ca<-cbind(as.vector(rep(1, a-1)), -diag(a-1))
        N<-sum(mn)*b
        etai<-matrix(0, length(mcon), a)
        ll<-0
        for (l in mcon){
        ll<-ll+1
        for (i in 1:a){
         for (j in 1:b){
#          for (j1 in seq(round(-bˆl ), round(bˆl))){
   for (j1 in seq(-l, l) ){
          if ((j+j1>0)& (j+j1<=b)  )  {
         tmpinc<- sum((ranks[((d1==i)&(d2==j))] -Rij[i, j])*
```

```
      (ranks[((d1==i)&(d2==(j+j1)))] -Rij[i, (j+j1)]))*sum(mn)/(b*mn[i]*(mn[i]-1))
         etai[ll, i]<-etai[ll, i]+tmpinc
         }
        } #end of j1
        } #end of j
       }  #end of i
     } # end of l
       Ri<-as.vector(Ri)
      TSalphastat<-function(etaii, Ca, N, Ri) N * t(Ri)%*% t(Ca)%*%
                   solve( Ca%*% diag(etaii)%*% t(Ca) ) %*% Ca %*% Ri
   TSalpha<- apply(etai, 1, TSalpha<-function(etaii) {TSalphastat(etaii, Ca, N, Ri)})
      palpha<-1-pchisq(TSalpha, nrow(Ca))
      palpha1<- palpha
 list(TSalpha=TSalpha, Dpalpha=palpha1, Dpbeta=Dpbeta, Dpgamma=Dpgamma,
Dpphi=Dpphi, palpha=palpha1, pbeta=pbeta, pgamma=pgamma, pphi=pphi)
}

CF<-function(data,a,b, mn, coln=4){
       ranks<-data[, coln]
       d1<-data[, 1]
d2<-data[, 2]
d3<-data[, 3]
Rij<-as.matrix(tapply(ranks, list(d1, d2), mean) )
   # matrix with \bar{R}<-{ij.} as the (i,j) element
Rik<-as.matrix(tapply(ranks, list(d1, d3), mean) )
# matrix with \bar{R}<-{i.k} as the (i,j) element  might have NA if unbalanced.
Rik<-replace(Rik, is.na(Rik), 0)    # replace NA's in matrix Rik by 0.
```

```
# note: after replace, the number of rows still correct, but the number of columns
# would be all same as max<-i n<-i instead of n<-i columns for the ith row.
Rbari<-as.vector(tapply(ranks, d1, mean) ) #vector with ith element \bar{R}<-{i..}
Rbarj<-as.vector(tapply(ranks, d2, mean) ) #vector with ith element \bar{R}<-{.j.}
Rim<-kronecker(Rbari, t(as.vector(rep(1,b))) )
  # a a by b matrix with all elements of the ith row same as \wtR<-{i..}
Rjm<-kronecker(t(Rbarj), as.vector(rep(1,a)))
  # a a by b matrix with all elements of the ith column same as \wtR<-{.j.}

## calculate test statistics
MSalpha<- sum(mn*(Rbari-mean(ranks))^2)*b/(a-1)
MSb<-  (sum(Rik^2)-sum(Rbari^2*mn) )*b/(sum(mn)-a)
Falpha<-MSalpha/MSb

MSphi<- sum(mn*apply((Rij-Rjm)^2 , 1, sum) )/((a-1)*b  )
MSEphi<- (sum(ranks^2)- sum(mn*apply(Rij^2, 1, sum))  )/(b*(sum(mn)-a))
Fphi<-MSphi/MSEphi

MSbeta<- sum(mn)* sum( (Rbarj-mean(ranks))^2 ) /(b-1)
MSgamma<- sum(mn*apply((Rij-Rim-Rjm+mean(ranks) )^2, 1, sum ))/((a-1)*(b-1) )
MSE<- MSEphi*b/(b-1) - MSb/(b-1)
Fbeta<-MSbeta/MSE
Fgamma<-MSgamma/MSE

palpha<-1-pf(Falpha, a-1, sum(mn)-a )
pbeta<-1-pf(Fbeta, b-1, (sum(mn)-a)*(b-1) )
pgamma<-1-pf(Fgamma, (a-1)*(b-1), (sum(mn)-a)*(b-1) )
```

```
pphi<-1-pf(Fphi, (a-1)*b, (sum(mn)-a)*b )

list(palpha=palpha, pbeta=pbeta, pgamma=pgamma, pphi=pphi)
}

sigijk4<-function(x){
        nijk<-length(x)
sigmaijk4<- 0
for (m1 in 1:nijk){
 for (m2 in 1:nijk){
   for (m3 in 1:nijk){
     for (m4 in 1:nijk){
     flag<- (m1!=m2)&(m1 !=m3)&(m1 !=m4) &(m2 !=m3) &(m2 !=m4) & (m3 !=m4)
     sigmaijk4<-sigmaijk4+ (flag==T)*(x[m1]-x[m2])^2 * (x[m3]-x[m4])^2
      }
    }
   }
  }
 sigmaijk4<-sigmaijk4/(4*nijk*(nijk-1)*(nijk-2)*(nijk-3) )
 sigmaijk4
}

thetahat<-function(x) (mean((x-mean(x))^2 ))^2
thetahatijj1<-function(x, y) (mean((x-mean(x))*(y-mean(y))  ))^2

sigijk4jack<-function(x){
n<-length(x)
```

```
s4hat<- thetahat(x)
result <- n *s4hat
        for (i in 1:n) result<-result- (n-1)/n* thetahat(x[-i])
        result
        }

sigijk4boot<-function(x){
n<-length(x)
thetahatstar<-mean(apply(matrix(sample(x, 3*1000, replace = T), 1000, 3), 1, thetahat))
result<-2*thetahat(x)-thetahatstar
result
}

## Jackknife estimate of \sigma_{ijj'}^2
## x=(X_{ij1}, \cdots, X_{ijn_i}), y=(X_{ij'1}, \cdots, X_{ij'n_i}),

sigijj12jack<-function(x, y){
n<-length(x)
s4hat<- thetahatijj1(x, y)
result <- n *s4hat
        for (i in 1:n) result<-result- (n-1)/n* thetahatijj1(x[-i], y[-i])
        result
        }
```

Bibliography

[1] Akritas, M. G (1990). The Rank Transform Method in Some Two-Factor Designs, *J. Amer. Statist. Assoc. 85*, 73–78.

[2] Akritas, M.G., Antoniou E., and Wang L. (2003), Fully Nonparametric ANCOVA with Fixed Window Sizes, submitted.

[3] Akritas, M.G. and Arnold, S. (1994) Fully Nonparametric Hypotheses for Factorial Designs I: Multivariate Repeated Measures Designs, *Journal of the American Statistical Association*, **89**, 336-343.

[4] Akritas, M.G. and Arnold, S. (2000). Asymptotics for ANOVA when the Number of Levels is Large, *Journal of the American Statistical Association*, 95, 212-226.

[5] Akritas, M. G, Arnold, S. F. and Brunner, E. (1997) Nonparametric Hypotheses and Rank Statistics for Unbalanced Factorial Designs. *Journal of the American Statistical Association* **92**, 258-265.

[6] Akritas, M.G. and Brunner, E. (1997). ”A unified approach to rank tests in mixed models,” *Journal of Statistical Planning and Inference*, **61**, 249-277.

[7] Akritas, M. G, Arnold, S. F. and Du, Y. (2000). Nonparametric models and methods for nonlinear analysis of covariance. *Biometrika* 87, 507-526.

[8] Akritas, M.G. and Papadatos, N. (2004). Heteroscedastic one-way ANOVA and lack-of-fit test. to appear in *Journal of the American Statistical Association*.

[9] Altman N. S. (1990), ”kernel smoothing of data with correlated errors”. JASA **85**, 749-59

[10] Arnold, S. (1980). Asymptotic Validity of F Tests for the Ordinary Linear Model and the Multiple Correlation Model. *J. Am. Statist. Assoc.* **75** 890-894.

[11] Arnold, S.F. (1981). *The Theory of Linear Models and Multivariate Analysis*, New York: Wiley.

[12] Bathke, Arne (2002). ANOVA for a Large Number of Treatments. *Mathematical Methods of Statistics.* Volume 11, No 1, 118-132.

[13] Billingsley, P. (1986). *Convergence of Probability Measures*, Second Edition. J. Wiley & Sons, New York.

[14] Boos, D.D. and Brownie, C. (1995). ANOVA and Rank Tests when the Number of Treatments is Large, *Statist. Probab. Lett.*, 23, 183–191.

[15] Box, G. E.P., (1954)a, 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. *Annals of Mathematical Statistics* 25, 2, 290-302.

[16] Box, G.E.P., (1954)b, Some theorems on quadratic forms applied in the study of analysis of variance problems, II. Effects of inequality of variance and of correlation between errors in the two-way classification. *Annals of Mathematical Statistics* 25, 3, 484-498.

[17] Box, G. E. P. and Anderson, S. L. (1955). Permutation Theory in the Derivation of Robust Criteria and the Study of Departures from Assumption. *J. R. Statist. Soc.* **B 17,** 1–26.

[18] Brillinger, D.R. (1973), ”The analysis of time series collected in an experiment design”, in *Multivariate Analysis - III*, (ed. P.R. Krishnaiah), New York, Academic, pp. 241-256

[19] Brillinger, D.R. (1980), ”Some aspects of the analysis of evoked response experiments,” in *Statistics and related topics*, eds. M. Csorgo, D.A. Dawson, J.N.K. Rao, and A.K.Saleh, Amsterdam: North-Holland, pp. 15-168.

[20] Brownie, C. and Boos, D.D. (1994). Type I Error Robustness of ANOVA and ANOVA on Ranks when the Number of Treatments is Large, *Biometrics*, 50, 542-549.

[21] Brumback B.A. and Rice J.A., 1998, ”Smoothing spline models for the analysis of nested and crossed samples of curves”, *Journal of the American Statistical Association*, **93**, 961-976.

[22] Brunner, E., Domhof, S. and Langer, F. (2002). *Nonparametric Analysis of Longitudinal Data in Factorial Experiments.* NY: Wiley.

[23] Chiang, C.-T., Rice, J.A. and Wu, C. O. (2001). Smoothing spline estimation for varying coefficient models with repeatedly measured dependent variables. *Journal of the American Statistical Association*, **96**, 605-619.

[24] Cliff, N. (1993). Dominance Statistics: Ordinal Analyses to Answer Ordinal Questions. *Psychological Bulletin* **114**, 494-509.

[25] Davidian, M. and Gallant, A.R. (1993). ”The nonlinear mixed effects model with a smooth random effects density.” *Biometrika*. *80*, 475-488.

[26] Davidian, M. and Giltinan, D.M. (1995). *Nonlinear models for repeated measurement data.* New York: Chapman and Hall.

[27] Diggle, P.J., Liang, K.-Y. and Zeger, S.L. (1994) *Analysis of Longitudinal Data.* Oxford University Press, Oxford.

[28] Fan J. and Lin (1998) ”Test of significance when data are curves”, *Journal of the American Statistical Association*, **93**, 1007-1021.

[29] Hand D. and Crowder M. (1996) *Practical longitudinal data analysis.* Chapman and Hall.

[30] Hart, J.D. (1991). ”Kernel regression estimation with time series errors. ” *Journal of Royal Statistical Society*, Ser. B, **53**, 173-87.

[31] Hart, J.D. and Wehrly, T.E. (1986) ”Kernel regression estimation using repeated measurements data.” *Journal of the American Statistical Association*, **81**, 1080-8.

[32] Hoover, D. R., Rice, J. A., Wu, C.O., Yang, L. (1998) ”Nonparametric smoothing estimates of time-varing coefficient models with longitudinal data”, *Biometrika*, **85**, 4, 809-822.

[33] Kerr M.K., Afshari C.A. , Bennett L., Bushel P., Martinez J., Walker N., and Churchill G.A. (2002) Statistical Analysis of a Gene Expression Microarray Experiment with Replication. *Statistica Sinica*, 2002, 203-217.

[34] Kerr, M.K., Martin, M., and Churchill, G.A. 2001. Analysis of variance for gene expression microarray data. To appear in *Journal of Computational Biology* , www.jax.org/research/churchill/pubs/index.html.

[35] Krutchkoff, R. G. (1989). Two-way Fixed Effects Analysis of Variance when the Error Variances may be Unequal. *J. Stat. Comp. Simul.* **32,** 177–183.

[36] Li, H., Lindsay. B.G and Waterman, R.P. (2003). Efficiency of projected score methods in rectangular array asymptotics. *J. R. Statist. Soc.* B. **65,** 191-208.

[37] Liang, K.-Y. and Zeger, S.L. (1986). ”Longitudinal data analysis using generalized linear models.” *Biometrika*, **73**, 13-22.

Lin and Ying (2001)

[38] Mallet, A. (1986). ”A maximum likelihood estimation method for random coefficient regression models.” *Biometrika*, **73**, 645-656.

[39] Miller, R. G., Jr. (1986). Beyond ANOVA: Basics of Applied Statistics. John Wiley, New York. (Reprinted 1996, Chapman & Hall, New York.)

[40] Moyeed, R.A. and Diggle, P.J. (1994). ”Rates of convergence in semi-parametric modeling of longitudinal data. ” *Aust. J. Statist.* **36**, 75-93.

[41] Pinheiro, J.C. and Bates, D.M. (1995). ”Approximations to the loglikelihood function in the nonlinear mixed effects model.” *Journal of Computational and Graphical Statistics*, **4**, 12-35.

[42] Portnoy, S. (1984). Asymptotic Behavior of M-estimators of p Regression Parameters when p^2/n is Large. I. Consistency. *Annals of Statistics* **12,** 1298–1309.

[43] Sahai, H and Ageel, M. (2000). The analysis of variance : fixed, random, and mixed models. Springer Verlag, New York.

[44] Scheff*é*, H. (1959). The Analysis of Variance. John Wiley, New York.

[45] Wang, L. and Akritas, M.G. (2002). Two-way Heteroscedastic ANOVA when the Number of Levels is Large. Submitted.

[46] Smith, A.F.M. and Roberts, G.O. (1993). ”Bayesian computation via the Gibbs sampler and related Markov chain Monte Carlo methods.” *Journal of the Royal Statistical Statistical Society,* Ser. B, **55**, 3-23.

[47] Rice, J. and Silverman, B.W. (1991) ”Estimating the mean and covariance structure nonparametrically when the data are curves”, ”*Journal of Royal Statistical Society,* Ser. B, **53**, 233-43.

Ramsay and Li (1998)

[48] Ramsay J.O. and Silverman, B.W. (1997) *Functional data analysis*, Springer.

[49] Tsangari, H. and Akritas, M. G. (2003a). Nonparametric ANCOVA with two and three covariates. *Journal of Multivariate Analysis.* In press.

[50] Tsangari, H. and Akritas, M. G. (2003b). Nonparametric models and methods for ANCOVA with dependent data. *Journal of Nonparametric Statistics* In press.

[51] Vonesh,E. F. Chinchilli, V M. (1997) *Linear and nonlinear models for the analysis of repeated measurements*

[52] Wang, H. and Akritas, M.G. (2003). Rank Tests for ANOVA with Large Number of Factor Levels. In press in *Journal of Nonparametric Statistics*.

[53] Wang, L. and Akritas, M.G. (2002). Two-way Heteroscedastic ANOVA when the Number of Levels is Large. Submitted.

[54] Wang, L., Akritas, M.G., and VanKeilegom I., (2003) Nonparametric Goodness-of-fit Test for Heteroscedastic Regression Models, submitted.

[55] Wang, L. (2003). Testing in Heteroscedastic ANOVA with Large Number of Levels and in Nonparametric ANCOVA. Ph.D. Thesis. The Pennsylvania State University.

[56] Wolfinger, R.D., Gibson, G., Wolfinger, E.D., Bennett, L., Hamadeh, H., Bushel, P., Afshari, C., and Paules, R.S (2001). Assessing Gene Significance from cDNA Microarray Expression Data via Mixed Models. *Journal of Computational Biology* **Volume 8** number 6, 625-637. http://statgen.ncsu.edu/ggibson/Publications/WGetc.pdf

[57] Wang, Y. (1998) "Mixed-effects Smoothing Spline ANOVA, " *Journal of Royal Statistical Society*, Ser. B, 60, 159-174.

[58] Yates, F. (1934). The analysis of multiple classifications with unequal numbers in the different classes. *J. Am. Statist. Assoc.* **29** 51-56.

[59] Zeger, S. L. and Diggle, P.J. (1994). ”Semiparametric models for longitudinal data with application to CD4 cell numbers in HIV seroconverters.” *Biometrics* **50**, 689-99.

[60] Zeger, S.L. and Liang, K.-Y. (1986). ”Longitudinal data analysis for discrete and continuous outcomes.” *Biometrics*, **44**, 175-188.

Vita

Haiyan Wang was born in Hailong, Meihekou, Jilin Province, P.R. China on October 1, 1973. Her mother is Guizhen Jiang and her father is Xuewang Wang. She obtained her Bachelor's and Master's degree in Probability and Statistics at Beijing University, Beijing, China, in 1996 and 1999, respectively. In the Fall of 1999, she enrolled in the Ph.D. program of Statistics at the Pennsylvania State University, University Park.

This thesis was typeset with LaTeX 2ε[‡] by the author.

[‡] LaTeX 2ε document preparation system was developed by Leslie Lamport as a special version of Donald Knuth's TeX program for computer typesetting. TeX is a trademark of the American Mathematical Society. The LaTeX 2ε macro package for The Pennsylvania State University Ph.D. format was written and modified by Chanseok Park.