The Pennsylvania State University
The Graduate School
Department of Statistics

THRESHOLDING METHODS FOR HYPOTHESIS TESTING IN HANOVA

A Dissertation in
Statistics
by
Min Hee Kim

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Philosophy

August 2010

The dissertation of Min Hee Kim was reviewed and approved* by the following:

Michael G. Akritas
Professor of Statistics
Dissertation Advisor
Chair of Committee

Bing Li
Professor of Statistics

Runze (Richard) Li
Professor of Statistics

Leonid Berlyand
Professor of Mathematics

Bruce G. Lindsay
Professor of Statistics
Head of the Department of Statistics

*Signatures are on file in the Graduate School.

Abstract

A new thresholding method, based on L-statistics and called *order thresholding*, is proposed as a technique for improving the power when testing against high-dimensional alternatives. The new method allows great flexibility in the choice of the threshold parameter. This results in improved power over the soft and hard thresholding methods. Moreover, order thresholding is not restricted to the normal distribution. An extension of the basic order threshold statistic to high-dimensional ANOVA is presented. Furthermore, the application of order thresholding to Pearson's chi-square test is developed. The performance of the basic order threshold statistic and its extensions is evaluated with extensive simulations.

Table of Contents

List of Figures

List of Tables

Acknowledgments

I am grateful to Dr. Michael G. Akritas for his advice, encouragement and everything he has done for me. This work was supported by the Korea Science and Engineering Foundation Grant.

Dedication

This work was supported by the Korea Science and Engineering Foundation Grant (No. you Grant #).

Chapter 1

Introduction

It is well known that, when testing against a high-dimensional alternative, omnibus tests designed to detect any departure from the null hypothesis have low power. Neyman's (1937) truncation idea, though motivated by a different type of problem, served as the spring board for the development of modern related approaches. Soft and hard thresholding were introduced in the context of nonparametric function estimation using wavelets by Donoho and Johnstone (1994). Johnstone and Silverman (2004) elaborate on a number of additional applications of thresholding including image processing, model selection, and data mining. Beran (2004) considered applications to the one-way ANOVA design. Spokoiny (1996), Fan (1996), Fan and Lin (1998), and Fan and Huang (2001) consider applications of thresholding methods to testing problems. Fan (1996) found that hard thresholding outperforms both soft thresholding and adaptive Neyman's truncation.

This dissertation proposes a new thresholding method based on L-statistics, which is termed *order thresholding*. Order thresholding allows great flexibility in the choice of the threshold parameter, can be used for distributions other than the normal, and extends naturally to factorial design settings and goodness-of-fit testing problem.

In the simple context where the X_i are independent $N(\theta_i, 1)$, $i = 1, \ldots, n$, and we wish to test $H_0 : \theta_1 = \cdots = \theta_n = 0$ vs. $H_a : \theta_i \neq 0$, for some i, the hard thresholding and order thresholding test statistics are, respectively,

$$T_H(\delta_n) = \sum_{i=1}^{n} Y_i I\{Y_i > \delta_n\}, \quad \text{and} \quad T_L(k_n) = \sum_{i=1}^{n} c_{in} Y_{i,n}, \tag{1.1}$$

where $Y_i = X_i^2$, $Y_{1,n} < \cdots < Y_{n,n}$ are the ordered Y_i's, $c_{in} = I(i > n - k_n)$, and δ_n, k_n are the corresponding threshold parameters. Thus, $T_L(k_n)$ is an L-statistic based on the largest k_n squared observations. Both order thresholding and hard thresholding rely on a statistic that is the sum of the largest order statistics, but they differ because in order thresholding the number of the considered observations is fixed, whereas in hard thresholding δ_n is fixed but the number of considered observations is a random variable. Conceptually, the connection between hard thresholding and order thresholding is similar to that between type I and type II censoring. The main difference being that the threshold parameters in type I and type II censoring (the cut-off point and the proportion of observations included, respectively) remain fixed, while in the present case they change with the sample size. As we will see, this distinction implies very different asymptotic behavior.

The idea behind both statistics in (1.1) is similar to that of Neyman's truncation. Namely, when the "signal" is known to be concentrated in a few locations, the accumulation of stochastic errors has a negative impact on the performance of the procedure based on the chi-square statistic

$$T_L(n) = \sum_{i=1}^{n} Y_i. \tag{1.2}$$

Since the signal locations are not known, the statistics in (1.1) attempt to minimize the accumulation of noise by focusing on the observations with the largest absolute values (or the largest squared values). The asymptotic theory of soft and hard thresholding, however, requires the threshold parameter to tend to infinity at a prescribed rate. Moreover, simulations show a lack of robustness to the choice of the threshold parameter which has to be chosen the same regardless of the signal-to-noise ratio. (See Tables A.1, A.2, and Figure A.1; the detailed description is given in Section 2.2.) The practical implication of this is potentially compromised power. In contrast, the asymptotic theory of order thresholding allows very flexible choice of the threshold parameter. Moreover, extensive simulations suggest that order thresholding outperforms both the Simes (1986) multiple testing procedure and hard thresholding.

The success of any threshold procedure, however, rests on the premise that few signals are hidden in a lot of noise. For example, let $X_i = \theta_i + \epsilon_i$, $i = 1, \ldots, n$, where $\epsilon_i \sim$ i.i.d. $N(0, 1)$, and consider testing $H_0 : \theta_i = 0$, $\forall i$. If all $\theta_i \neq 0$ (so high signal-to-noise ratio), then Wald's chi-square test is preferable to any threshold procedure; on the other hand, if $\theta_i \neq 0$ only for a few cases (low signal-to-noise ratio), then thresholding methods can achieve far superior power than the chi-square test. To apply thresholding methods usefully also in high signal-to-noise ratio situations, one has to apply either the Fourier transformation or the wavelet transformation. The effect of such transformations is presented in Section 5.1.

Since either of these transformations takes the form of multiplying the data vector by an orthonormal matrix $\mathbf{\Gamma}$, it is seen that

$$\mathbf{X}_{\mathbf{\Gamma}} \equiv \mathbf{\Gamma}\mathbf{X} = \mathbf{\Gamma}\boldsymbol{\theta} + \mathbf{\Gamma}\boldsymbol{\epsilon} \ \sim \ N\left(\mathbf{\Gamma}\boldsymbol{\theta}, \mathbf{I}_n\right).$$

Thus, under the null hypothesis $H_0 : \theta_i = 0$, $\forall i$, the transformed data continue to be i.i.d. $N(0, 1)$, while under the alternative, $||\boldsymbol{\theta}|| = ||\mathbf{\Gamma}\boldsymbol{\theta}||$. Thus, Wald's chi-square test will have the same power when applied to the transformed data, since the noncentrality parameter, $\sum_{i=1}^{n} \theta_i^2$, remains the same after the transformation. On the other hand, such transformations should improve the power of threshold methods in high signal-to-noise ratio situations.

The remaining chapters of this dissertation are organized as follows. In Chapter 2, we consider the hard threshold statistic and Simes procedure to improve the power of the chi-square statistic against high-dimensional alternatives. In Chapter 3, we represent a special form of the order statistics using data from an exponential distribution and develop the order threshold procedure for testing normal means in settings where the number of parameters increases with the sample size. Moreover, we present simulation results and give a recommendation for choosing a data-driven value of the order threshold parameter. Chapter 4 extends the order thresholding test procedure to the high-dimensional ANOVA setting (called HANOVA in Fan and Lin, 1998), presents simulation results comparing the power of the classical

F and order threshold statistics, and gives a recommendation for a data-driven choice of the order threshold parameter. Chapter 5 extends the applicability of the order thresholding test procedure to Pearson's chi-square test when the number of cells grows with the sample size. The order threshold statistics based on the cell counts and on the cell averages, with the Fourier and wavelet transformations, are developed and compared with Pearson's χ^2 statistic. A discussion summarizing the developments is given in Chapter 6. Finally, all tables, figures, and proofs are given in the Appendix.

Chapter 2

Common Approaches: Hard Thresholding and Simes Procedure

Hard thresholding is a method for increasing the power of test procedures in settings where the number of parameters being tested increases with the sample size. In the first section, we motivate the need for improved power in the simple case of testing a simple null versus simple alternative. Section 2.2 considers a simple null versus composite alternative, and introduces hard thresholding. This section is adapted from Fan (1996). Section 2.3 includes (a power-enhanced version of) the Simes statistic which helps to relate the present work with the modern literature on multiple testing.

2.1 Simple Null versus Simple Alternative

Let $\mathbf{X} = (X_1, \ldots, X_n)^T$ be an $n-$dimensional random vector. We assume that the X_i, $i = 1, \ldots, n$, are independent $N(\theta_i, 1)$, where the θ_i are known. We wish to test the null hypothesis

$$H_0 : \theta_1 = \cdots = \theta_n = \theta_0$$

against the alternative hypothesis

$$H_a : \theta_i = \theta_{ia} \quad \text{for} \quad i = 1, \ldots, n, \tag{2.1}$$

where θ_0 and θ_{ia}, $i = 1, \ldots, n$, are completely specified. The ratio between probability density functions under H_0 and H_a is given by

$$\frac{f_{H_a}(x_1, \ldots, x_n)}{f_{H_0}(x_1, \ldots, x_n)} = \frac{\prod_{i=1}^{n} \exp\left(-\frac{(x_i - \theta_{ia})^2}{2}\right)}{\prod_{i=1}^{n} \exp\left(-\frac{(x_i - \theta_0)^2}{2}\right)}$$
$$= \exp\left(\sum_{i=1}^{n}(\theta_{ia} - \theta_0)x_i - \sum_{i=1}^{n}\left(\frac{\theta_{ia}^2 - \theta_0^2}{2}\right)\right).$$

We can consider the following test statistic $\sum_{i=1}^{n}(\theta_{ia} - \theta_0)x_i$ by the Neyman-Pearson fundamental theorem, or equivalently,

$$S_n(\boldsymbol{\theta}^0, \boldsymbol{\theta}^a) = \sum_{i=1}^{n}(\theta_{ia} - \theta_0)\,(x_i - E_{H_0}(X_i)) = \sum_{i=1}^{n}(\theta_{ia} - \theta_0)(x_i - \theta_0), \tag{2.2}$$

where $\boldsymbol{\theta}^0 = (\theta_0, \ldots, \theta_0)^T$ and $\boldsymbol{\theta}^a = (\theta_{1a}, \ldots, \theta_{na})^T$. Note that the expected value of the test statistic $S_n(\boldsymbol{\theta}^0, \boldsymbol{\theta}^a)$ under the alternative hypothesis (2.1) is

$$E_{H_a}\left[\sum_{i=1}^{n}(\theta_{ia} - \theta_0)\,(X_i - \theta_0)\right] = \sum_{i=1}^{n}(\theta_{ia} - \theta_0)^2 > 0.$$

This suggests that the rejection region will be of the form $S_n(\boldsymbol{\theta}^0, \boldsymbol{\theta}^a) > C$ for some positive value of C. Theorem 2.1 shows the distributions of $S_n(\boldsymbol{\theta}^0, \boldsymbol{\theta}^a)$ under the null and the alternative hypotheses, and the rejection region.

Theorem 2.1. *Let the X_i, $i = 1, \ldots, n$, be a sequence of normally distributed independent random variables with mean θ_i and variance 1, and let $S_n(\boldsymbol{\theta}^0, \boldsymbol{\theta}^a)$ be given in (2.2).*
(a) Under $H_0 : \theta_1 = \cdots = \theta_n = \theta_0$, where θ_0 is specified,

$$S_n^*(\boldsymbol{\theta}^0, \boldsymbol{\theta}^a) = \frac{S_n(\boldsymbol{\theta}^0, \boldsymbol{\theta}^a)}{\|\boldsymbol{\theta}^a - \boldsymbol{\theta}^0\|} \sim N(0, 1).$$

Thus, the test rejects the null hypothesis if

$$S_n^*(\boldsymbol{\theta}^0, \boldsymbol{\theta}^a) \; > \; \Phi^{-1}(1-\alpha)\,, \tag{2.3}$$

where Φ *is a standard normal distribution function and* α *is a significance level.*
(b) Under $H_a : \theta_i = \theta_{ia}$ *for each* i*, where* θ_{ia}*'s are specified,*

$$S_n^*(\boldsymbol{\theta}^0, \boldsymbol{\theta}^a) - \|\boldsymbol{\theta}^a - \boldsymbol{\theta}^0\| = \frac{S_n(\boldsymbol{\theta}^0, \boldsymbol{\theta}^a) - \|\boldsymbol{\theta}^a - \boldsymbol{\theta}^0\|^2}{\|\boldsymbol{\theta}^a - \boldsymbol{\theta}^0\|} \;\sim\; N(0,1).$$

Thus, the power of the test (2.3), evaluated at the parameter values specified by the simple alternative hypothesis is $1 - \Phi\left(z_\alpha - \|\boldsymbol{\theta}^a - \boldsymbol{\theta}^0\|\right)$*, where* $1 - \Phi(z_\alpha) = \alpha$*.*

Note that the power of $S_n^*(\boldsymbol{\theta}^0, \boldsymbol{\theta}^a)$ at the parameter values specified by the simple alternative hypothesis tends to 1 as $\|\boldsymbol{\theta}^a - \boldsymbol{\theta}^0\| \to \infty$. Theorem 2.2 derives the power of $S_n^*(\boldsymbol{\theta}^0, \boldsymbol{\theta}^a)$ at parameter values different from the values specified by the simple alternative hypothesis.

Theorem 2.2. *Consider the settings of Theorem 2.1. We assume that the true values are* θ_{ic}*,* $i = 1, \ldots, n$*, where* $\theta_{ic} \neq \theta_{ia}$*. Then, the power of* $S_n^*(\boldsymbol{\theta}^0, \boldsymbol{\theta}^a)$ *at the parameter values* θ_{ic}*,* $i = 1, \ldots, n$*, may not tend to 1 even though* $\|\boldsymbol{\theta}^a - \boldsymbol{\theta}^0\| \to \infty$*. However, if* $(\theta_{ic} - \theta_0) = k\,(\theta_{ia} - \theta_0)$ *for any positive number* k *and* $\|\boldsymbol{\theta}^a - \boldsymbol{\theta}^0\| \to \infty$*, then the power of* $S_n^*(\boldsymbol{\theta}^0, \boldsymbol{\theta}^a)$ *always tends to 1.*

2.2 Simple Null versus Composite Alternative, and Hard Thresholding

In Section 2.1, we studied the simple case of testing a simple null versus simple alternative. In this case, the test statistic $S_n^*(\boldsymbol{\theta}^0, \boldsymbol{\theta}^a)$ includes the specific alternatives $\theta_i = \theta_{ia}$, $i = 1, \ldots, n$. Moreover, the power of the test (2.3) also depends on the parameter values and the specific alternatives. From now, we consider testing a simple null versus composite alternative. In particular, the null and the alternative hypotheses are written as

$$H_0 : \theta_1 = \cdots = \theta_n = \theta_0, \;\; \theta_0 \text{ known},$$

versus

$$H_a: \text{ At least one of } \theta_i \neq \theta_0, \quad 1 \leq i \leq n. \tag{2.4}$$

In order to test the null hypothesis against the alternative hypothesis (2.4), we can use the test statistic

$$S_n(\boldsymbol{\theta}^0) = \sum_{i=1}^{n} (x_i - \theta_0)^2. \tag{2.5}$$

Note that $S_n(\boldsymbol{\theta}^0)$ is obtained from $S_n(\boldsymbol{\theta}^0, \boldsymbol{\theta}^a)$ by replacing θ_{ia} by its "estimator" x_i. Since, under the alternative hypothesis (2.4), we expect $S_n(\boldsymbol{\theta}^0)$ to be large, the rejection region will be of the form $S_n(\boldsymbol{\theta}^0) > C$ for some positive value of C. Theorem 2.3 shows the asymptotic theory of $S_n(\boldsymbol{\theta}^0)$ under the null and the alternative hypotheses, and the rejection region.

Theorem 2.3. *Let the X_i, $i = 1, \ldots, n$, be a sequence of normally distributed independent random variables with mean θ_i and variance 1, and let $S_n(\boldsymbol{\theta}^0)$ be given in (2.5).*
(a) Under $H_0 : \theta_1 = \cdots = \theta_n = \theta_0$, θ_0 known,

$$S_n^*(\boldsymbol{\theta}^0) = \frac{S_n(\boldsymbol{\theta}^0) - n}{\sqrt{2n}} \xrightarrow{d} N(0,1), \quad \text{as} \quad n \to \infty.$$

Thus, the test rejects the null hypothesis if

$$S_n^*(\boldsymbol{\theta}^0) \; > \; \Phi^{-1}(1-\alpha).$$

(b) Under $H_a : \theta_i = \theta_i^n$, where $\theta_i^n = \theta_0 + \dfrac{\Delta v_i}{\sqrt{n}}$ with $\sum_{i=1}^n v_i^2 = o\left(n^{3/2}\right)$,

$$\frac{S_n(\boldsymbol{\theta}^0) - E_{H_a}\left(S_n(\boldsymbol{\theta}^0)\right)}{\sqrt{Var_{H_a}\left(S_n(\boldsymbol{\theta}^0)\right)}} \xrightarrow{d} N(0,1), \quad \text{as} \quad n \to \infty.$$

Note that under a contiguous sequence of alternatives $\theta_i = \theta_i^n$, the expectation of $S_n(\boldsymbol{\theta}^0)$ is $E_{H_a}\left(S_n(\boldsymbol{\theta}^0)\right) = n + \sum_{i=1}^n \left(\theta_i^n - \theta_0\right)^2$ and the variance of $S_n(\boldsymbol{\theta}^0)$ is $Var_{H_a}\left(S_n(\boldsymbol{\theta}^0)\right) = 2n + 4\sum_{i=1}^n \left(\theta_i^n - \theta_0\right)^2$. The power of $S_n^*(\boldsymbol{\theta}^0)$ against alternatives θ_i^n is approximately $1 - \Phi\left(z_\alpha - \sum_{i=1}^n \left(\theta_i^n - \theta_0\right)^2 / \sqrt{2n}\right)$.

Note that the asymptotic power of $S_n(\boldsymbol{\theta}^0)$ against alternatives θ_i^n tends to α as n increases. Theorem 2.3 suggests that the test statistic $S_n(\boldsymbol{\theta}^0) = \sum_{i=1}^n (x_i - \theta_0)^2$ cannot detect alternatives of the order $\|\boldsymbol{\theta}^n - \boldsymbol{\theta}^0\|^2 = o\left(\sqrt{n}\right)$.

Now, we will present a procedure, called hard thresholding, which has been used in the literature for improving the power of $S_n(\boldsymbol{\theta}^0)$ by mitigating the effect of noise. The main idea of this method is to focus the statistic on the "large" absolute (or squared) values of $(x_i - \theta_0)$. The idea is similar to that of Neyman's truncation. Namely, when the coefficients θ_i different from θ_0 (called "signal") are known to be concentrated in a few locations, the accumulation of stochastic errors has a negative impact on the performance of the procedure based on the test statistic $S_n(\boldsymbol{\theta}^0)$. Since the signal locations are not known, hard thresholding attempts to minimize the accumulation of noise by focusing on the observations with the "large" absolute (or squared) values of $(x_i - \theta_0)$. The hard threshold statistic will be of the form

$$T_H(\delta_n) = \sum_{i=1}^n Y_i I\left\{Y_i > \delta_n\right\}, \tag{2.6}$$

where $Y_i = (X_i - \theta_0)^2$ and δ_n is the hard threshold parameter. The asymptotic theory of $T_H(\delta_n)$ requires the threshold parameter to tend to infinity at specific rates. In particular, it must be of the form $\delta_n = 2\log(na_n)$, where $a_n = c(\log n)^{-d}$, for $c > 0$ and $d > 0.5$. With this choice of the hard threshold parameter δ_n, the hard threshold test rejects the null hypothesis if

$$T_H^*(\delta_n) = \sigma_H^{-1}(\delta_n)\Big(T_H(\delta_n) - \mu_H(\delta_n)\Big) > \Phi^{-1}(1-\alpha), \tag{2.7}$$

where

$$\mu_H(\delta_n) = \sqrt{2/\pi}\; a_n^{-1}\delta_n^{1/2}(1+\delta_n^{-1}) \quad\text{and}\quad \sigma_H^2(\delta_n) = \sqrt{2/\pi}\; a_n^{-1}\delta_n^{3/2}(1+3\delta_n^{-1}).$$

The rejection region (2.7) is supported by the following theorem.

Theorem 2.4 (Fan, 1996)**.** *Under the null hypothesis,* $T_H^*(\delta_n) \xrightarrow{d} N(0,1)$*, with* $\delta_n = 2\log n + 2\log c - 2d\log\log n$*, for* $c > 0$ *and* $d > 0.5$.

Note that if we let $k_H(\delta_n)$ denote the random number of observations considered in $T_H(\delta_n)$, the asymptotic theory of $T_H(\delta_n)$ requires $E\left[k_H(\delta_n)\right]$ to converge to infinity at the rate of

$$\frac{(\log n)^d}{\sqrt{\log n + d\log(c^{1/d}(\log n)^{-1})}},$$

or, roughly, $(\log n)^{d-0.5}$.

While the asymptotic theory of $T_H(\delta_n)$ allows some flexibility in the choice of δ_n, the convergence of the distribution of $T_H(\delta_n)$ to its limiting distribution is very slow unless $c = 1$ and $d = 2$ (Fan, 1996). Tables A.1 and A.2 show that small departures in the recommended value of d, while keeping $c = 1$, have significant effect on the level of the test. The results are based on 30,000 simulation runs.

To fully appreciate the results reported in Table A.1, we mention that for $n = 500$ the recommended δ_{500} value is 5.1216, while the value $\delta_{500} - 2$ corresponds to $c = 1$ and $d = 2.5474$. We see that even with this small departure from the recommended value, the achieved alpha level is 0.0327 even with $n = 500$. In Table A.1 with $n = 500$, the $E\left[k_H(\delta_n + h)\right]$ ranges from 38.63 for $h = -2.0$ to 11.81 for $h = 0$. Thus, the deterioration of the achieved alpha levels occurs as $E\left[k_H(\delta_n)\right]$ increases over a relatively small range (in each case the variance of $k_H(\delta_n)$ is slightly smaller than its expected value). In Table A.2 with $n = 500$, the $E\left[k_H(\delta_n + h)\right]$ ranges from 9.39 for $h = 0.4$ to 3.80 for $h = 2.0$, and for this range of values the type I error rate does not change much. In both tables with $n = 500$, the variance of the binomial random variable $k_H(\delta_n + h)$ is slightly smaller than its expected value because $P(Y_i \leq \delta_n + h) > 0.92$ for $-2.0 \leq h \leq 2.0$.

Under the null hypothesis $H_0 : \theta_1 = \cdots = \theta_n = \theta_0$, θ_0 known, the asymptotic distribution of $T_H(\delta_n)$ converges to the standard normal distribution function even though the rate is very slow (unless $c = 1$ and $d = 2$). However, under the alternative hypothesis H_a : H_0 is false, the random variables $Y_i = (X_i - \theta_0)^2$, $i = 1, \ldots, n$, are independent, but not identically distributed. Thus, it is difficult to derive the asymptotic alternative distribution of $T_H(\delta_n)$ and calculate the exact power of the hard threshold statistic. However, the lower bound of the power of hard thresholding is shown by Fan (1996), and it is restated in Theorem 2.5. Note that when $\sum_{i=1}^{n} (\theta_i - \theta_0)^2 I\left(|\theta_i - \theta_0| > \sqrt{2\log n}\right) \gg \sigma_H(\delta_n)$, the asymptotic power of the hard threshold statistic $T_H(\delta_n)$ tends to 1.

Theorem 2.5 (Fan, 1996)**.** *The power of the hard thresholding test $T_H^*(\delta_n)$ given in (2.7) is at least*

$$P_{\boldsymbol{\theta}^n}\left\{\sum_{i\in S_0}\left((X_i-\theta_0)^2-1\right)\geq m_0\delta_n+\sigma_H(\delta_n)(z_\alpha-Z_n)\right\},$$

provided that $m_0 = o\left(na_n^{1/2}\right)$, where S_0 is the "oracle" best subset, $m_0 = \sharp(S_0)$, and Z_n is a sequence of random variables converging to the standard normal random variable and independent of X_i's, $i \in S_0$.

Remark 2.1. *We may, without loss of generality, set $\theta_0 = 0$. In the aforementioned context, the adaptive Neyman's and soft thresholding test statistics are, respectively,*

$$T_{AN} = \max_{1\leq m\leq n}\left\{\left(\sqrt{2m}\right)^{-1}\sum_{j=1}^{m}(X_j^2-1)\right\}, \quad \text{and} \quad T_S(\delta_n) = \sum_{j=1}^{n}\psi_{\delta_n}^2(X_j),$$

where $\psi_{\delta_n}(x) = sgn(x)(|x|-\sqrt{\delta_n})_+$ and $\delta_n = 2\log n + 2\log c - 2d\log\log n$, with t_+ denoting the positive part of t, and $c > 0$, $d > 2.5$. The asymptotic null distribution of the adaptive Neyman's statistic is followed from Darling and Erdös (1956), while that of the soft threshold statistic is followed from Fan (1996). The hard threshold statistic, however, outperforms the adaptive Neyman's and soft threshold statistics. Thus these test statistics are not considered in this testing problem.

2.3 The Simes Procedure and Variations

This section proposes a power-enhanced version of the Simes (1986) statistic which helps to relate the present work with the modern literature on multiple testing. First, we present the original Simes multiple testing procedure. Next, we improve the power by applying the Simes statistic at level

$$\alpha/(1-\sharp\,\{\text{nonzero means}\}\,/\,\text{sample size}).$$

In the simple context where the X_i are independent $N(\theta_i, 1)$, $i = 1, \ldots, n$, and we wish to test $H_0 : \theta_1 = \cdots = \theta_n = 0$ versus $H_a : \theta_i \neq 0$, for some i (Assume, without loss of generality, that $\theta_0 = 0$, from now on), the original Simes multiple testing procedure is as follows. First set up n hypothesis testing problems of the form $H_0^{(i)} : \theta_i = 0$, $i = 1, \ldots, n$. For each such hypothesis $H_0^{(i)}$, calculate the p-value, denoted by P_i. Note that the p-value is $P_i = 2(1 - \Phi(|X_i|))$, $i = 1, \ldots, n$, where Φ is a standard normal distribution function. Lastly obtain the vector $P_{(1)}, \ldots, P_{(n)}$ of the ordered p-values. Then the original Simes statistic

$$T_S = \min_{1 \leq i \leq n} \left\{ nP_{(i)}/i \right\}$$

rejects the global null hypothesis, H_0^G, that all $H_0^{(i)}$ are true if $T_S < \alpha$.

A power-enhanced version of the original Simes test procedure uses $\alpha/(1 - k_n^{opt}/n)$ instead of α, where k_n^{opt} is the number of false null hypotheses. When we do not have any information about k_n^{opt}, we can use the estimator of k_n^{opt} as a data-driven choice. A recommendation for choosing k_n^{opt} is given in Section 3.3. The idea is obtained from Storey (2002, 2003), Efron, Tibshirani, Story, and Tusher (2001), and the simulation results.

Chapter 3

Order Thresholding

The asymptotic theory for hard thresholding (Fan, 1996) is specific to the normality assumption and to the choice of the hard threshold parameter δ_n. The centering and scaling of $T_H(\delta_n)$ in (2.6) are specific to the normality assumption, and δ_n is required to tend to infinity at a rate that is specific to the normality assumption. For example, δ_n tending to infinity is clearly not appropriate if the X_i have bounded support. Intuitively, if the signal is present in more locations, it is advantageous to lower the value of the hard threshold parameter. However, Tables A.1 and A.2 show that small departures from the value of δ_n recommended in Fan (1996) ($c = 1$, $d = 2$) have significant effect on the level of the test. Johnstone and Silverman (2004) make a strong case for the need for flexibility in the choice of the threshold parameter. This chapter is organized as follows. In the first section, we represent a special form of the order statistics using exponential data and briefly review the methodology of Chernoff, Gastwirth, and Johns (1967). Section 3.2 develops the order threshold procedure for testing normal means in settings where the number of parameters increases with the sample size, and presents simulation results comparing the hard thresholding, a power-enhanced version of the Simes (1986), and order thresholding test statistics. The simulation results suggest that choosing the order threshold parameter equal to the number of the false null hypotheses maximizes the power. Section 3.3 gives a recommendation for choosing a data-driven value of the order threshold parameter using the idea of Storey (2002, 2003).

3.1 From Order Statistics to Order Thresholding: An Overview

In the late 60s when the asymptotic theory of linear combinations of order statistics (L-statistics) was developed (cf. Bickel, 1967, Chernoff, Gastwirth, and Johns, 1967, Shorack, 1969, Stigler, 1969), the main emphasis was in the estimation of the location parameter. Therefore, the conditions in these papers do not yield automatically the asymptotic distribution of L-statistics that assign positive weight to only the largest order statistics. Such L-statistics were considered by Nagaraja (1982) in his study of the selection differential for applications to outlier detection. Using results from Hall (1978) and Stigler (1973), he obtained the asymptotic distribution in the extreme and quantile cases, respectively. Here we will use the conditions from the paper of Chernoff, Gastwirth, and Johns (1967), CGJ1967 from now on. Their approach is based on a special representation of the order statistics from the exponential distribution, which we now review.

Let $V_1, \ldots, V_n$ be i.i.d. from the standard exponential distribution, let $V_{1,n} < \cdots < V_{n,n}$ be the corresponding order statistics, and consider the order threshold statistic

$$T_{E,L}(k_n) = \sum_{i=1}^{n} c_{in} V_{i,n} = \sum_{i=n-k_n+1}^{n} V_{i,n}. \tag{3.1}$$

The method of CGJ1967 for establishing the asymptotic distribution of the statistic $T_{E,L}(k_n)$ rests on the following well known property (cf. David and Nagaraja, 2003, pages 17-18).

Lemma 3.1. *The vector of order statistics* $(V_{1,n}, \ldots, V_{n,n})$ *may be represented in distribution by*

$$(V_{1,n}, \ldots, V_{n,n}) \stackrel{d}{=} (Y_1, \ldots, Y_n),$$

where

$$Y_i = \frac{V_1}{n} + \frac{V_2}{n-1} + \cdots + \frac{V_i}{n-i+1} = \sum_{j=1}^{i} \frac{V_j}{n-j+1}.$$

Thus, with $T_{E,L}(k_n)$ given by (3.1), it can be represented in distribution as

$$T_{E,L}(k_n) \stackrel{d}{=} \sum_{j=1}^{n} \alpha_{E,jn}(k_n) V_j, \tag{3.2}$$

where $\alpha_{E,jn}(k_n) = k_n/(n-j+1)$ for $j \le n-k_n$ and $\alpha_{E,jn}(k_n) = 1$ for $j > n-k_n$.

Relation (3.2) expresses $T_{E,L}(k_n)$ as a linear combination of the independent random variables $V_1, \ldots, V_n$ which enables the use of standard asymptotic results for establishing conditions for its asymptotic distribution. This is given in the following.

Theorem 3.1. *Let k_n, $n \ge 1$, be any sequence of integers which satisfies $k_n \to \infty$, as $n \to \infty$, and $k_n \le n$, and let $T_{E,L}(k_n)$ be given in (3.1). Then we have*

$$T^*_{E,L}(k_n) = \frac{T_{E,L}(k_n) - \sum_{i=1}^{n} \alpha_{E,in}(k_n)}{\sqrt{\sum_{i=1}^{n} \alpha_{E,in}(k_n)^2}} \stackrel{d}{\to} N(0,1), \quad \textit{as} \quad n \to \infty. \tag{3.3}$$

Note that the (exact) mean of $T_{E,L}(k_n)$ is $E(T_{E,L}(k_n)) = \sum_{i=1}^{n} \alpha_{E,in}(k_n)$ and the (exact) variance of $T_{E,L}(k_n)$ is $Var(T_{E,L}(k_n)) = \sum_{i=1}^{n} \alpha_{E,in}(k_n)^2$. Using Lindeberg-Feller central limit theorem, the asymptotic normality of $T_{E,L}(k_n)$ is proved in Appendix.

In the case where the observations Y_i, $i = 1, \ldots, n$, come from a distribution function F, the CGJ1967 approach for obtaining the asymptotic distribution of the order threshold statistic

$$T_{F,L}(k_n) = \sum_{i=1}^{n} c_{in} Y_{i,n} = \sum_{i=n-k_n+1}^{n} Y_{i,n}, \tag{3.4}$$

where $Y_{1,n} < \cdots < Y_{n,n}$ are the ordered Y_i's, is based on the expression $Y_{i,n} = \tilde{H}_F(V_{i,n})$, where $\tilde{H}_F = F^{-1} \circ G$, and G is the standard exponential distribution function, and the use of Taylor expansion to obtain

Lemma 3.2 (Chernoff, Gastwirth, and Johns, 1967)**.** *Let $T_{F,L}(k_n)$ be given by (3.4). Then,*

$$n^{-1} T_{F,L}(k_n) \stackrel{d}{=} \mu_{F,n}(k_n) + Q_{F,n}(k_n) + R_{F,n}(k_n),$$

where

$$\mu_{F,n}(k_n) = \frac{1}{n}\sum_{i=1}^{n} c_{in}\tilde{H}_F(\tilde{\nu}_{in}),$$

$$Q_{F,n}(k_n) = \frac{1}{n}\sum_{i=1}^{n} \alpha_{F,in}(k_n)(V_i - 1),$$

and

$$R_{F,n}(k_n) = \frac{1}{n}\sum_{i=1}^{n} c_{in}\left\{\left(\tilde{H}_F(V_{i,n}) - \tilde{H}_F(\tilde{\nu}_{in})\right) - (V_{i,n} - \tilde{\nu}_{in})\tilde{H}'_F(\tilde{\nu}_{in})\right\}$$

with $\alpha_{F,in}(k_n) = \dfrac{1}{n-i+1}\sum_{j=i}^{n} c_{jn}\tilde{H}'_F(\tilde{\nu}_{jn})$ *and* $\tilde{\nu}_{in} = \sum_{j=1}^{i}\dfrac{1}{n-j+1}$.

From now we will provide Assumptions A, B*, and C of CGJ1967 under which $Q_{F,n}(k_n)$ is asymptotically normally distributed and the remainder term, $R_{F,n}(k_n)$, tends to zero in probability.

Assumption A: $\tilde{H}_F(v)$ *is continuously differentiable for* $0 < v < \infty$.

*Assumption B**: Let* $u_{jn}(\epsilon)$, $u^{jn}(\epsilon)$, $v_{jn}(\epsilon)$, *and* $v^{jn}(\epsilon)$, *be such that*

$$P\left\{u_{jn}(\epsilon) < U_{j,n} < u^{jn}(\epsilon),\ 1 \le j \le n\right\} \ge 1-\epsilon, \quad n \ge 1, \tag{3.5}$$

$$P\left\{v_{jn}(\epsilon) < V_{j,n} < v^{jn}(\epsilon),\ 1 \le j \le n\right\} \ge 1-\epsilon, \quad n \ge 1, \tag{3.6}$$

where $U_{j,n}$, $j = 1, \ldots, n$, *are the ordered observations from an i.i.d. sequence of* $Uniform(0,1)$ *random variables and*

$$\begin{aligned} v_{jn}(\epsilon) &= -\log(1-u_{jn}(\epsilon)), & v^{jn}(\epsilon) &= -\log(1-u^{jn}(\epsilon)),\\ u_{jn}(\epsilon) &< j/(n+1) < u^{jn}(\epsilon), & \nu_{jn} &= -\log(1-j/(n+1)),\\ \tilde{\nu}_{in} &= \sum_{j=1}^{i}\frac{1}{n-j+1}, & v_{jn}(\epsilon) &< \nu_{jn} < \tilde{\nu}_{jn} < v^{jn}(\epsilon). \end{aligned}$$

Let $\alpha_{F,in}(k_n)$ *be given in Lemma 3.2. For each* $\epsilon > 0$,

$$\sum_{j=1}^{n}\left[|c_{jn}|\left\{\sup_{v_{jn}(\epsilon)<v<v^{jn}(\epsilon)}|G_{F,jn}(v)|\right\}\sqrt{\frac{j}{n-j+1}}\right] = o(n\sigma_{F,n}(k_n)),$$

where

$$G_{F,jn}(v) = \begin{cases} \dfrac{\tilde{H}_F(v) - \tilde{H}_F(\tilde{\nu}_{jn})}{v - \tilde{\nu}_{jn}} - \tilde{H}'_F(\tilde{\nu}_{jn}) & \text{if } v \neq \tilde{\nu}_{jn} \\ 0 & \text{if } v = \tilde{\nu}_{jn} \end{cases}$$

and

$$\sigma^2_{F,n}(k_n) = \frac{1}{n}\sum_{j=1}^{n} \alpha^2_{F,jn}(k_n).$$

Assumption C:

$$\max_{1\le j\le n} |\alpha_{F,jn}(k_n)| = o(n^{1/2}\sigma_{F,n}(k_n)).$$

Lemma 3.3 expresses the simultaneous bounds of the exponential and uniform order statistics, i.e., $u_{jn}(\epsilon)$, $u^{jn}(\epsilon)$, $v_{jn}(\epsilon)$, and $v^{jn}(\epsilon)$. Lemmas 3.3, 3.4, and 3.5 show that these simultaneous bounds satisfy relations (3.5) and (3.6).

Lemma 3.3. *Let $U_{i,n}$, $i = 1,\dots,n$, be order statistics from the uniform distribution in $(0,1)$, and set $V_{i,n} = -\log(1-U_{i,n})$. For any $0 < \epsilon < 1$ and some $1 - \log\left(n - \sqrt{\frac{n}{2}\log\left(\frac{58}{\epsilon}\right)} + 1\right)\Big/ \log n \le \delta(n) < 1 - \log\left(\frac{n}{2}\log\left(\frac{58}{\epsilon}\right)\right)\Big/(2\log n)$, set*

$$\begin{cases} u_{jn}(\epsilon) = \begin{cases} \max\left\{0,\ \frac{j}{n} - \sqrt{\frac{1}{2n}\log\left(\frac{58}{\epsilon}\right)}\right\}, & 1 \le j < n^{1-\delta(n)}, \\ 1 - e^{-\tilde{\nu}_{jn}} e^{\sqrt{2/\epsilon}}, & n^{1-\delta(n)} \le j \le n, \end{cases} \\ u^{jn}(\epsilon) = \begin{cases} \frac{j-1}{n} + \sqrt{\frac{1}{2n}\log\left(\frac{58}{\epsilon}\right)}, & 1 \le j < n^{1-\delta(n)}, \\ 1 - e^{-\tilde{\nu}_{jn}} e^{-\sqrt{2/\epsilon}}, & n^{1-\delta(n)} \le j \le n, \end{cases} \end{cases}$$

where $\tilde{\nu}_{jn} = \sum_{i=1}^{j} 1/(n-i+1)$. Then, the sequences of constants

$$v_{jn}(\epsilon) = -\log\left(1 - u_{jn}(\epsilon)\right),\quad v^{jn}(\epsilon) = -\log\left(1 - u^{jn}(\epsilon)\right)$$

satisfy (3.6).

Lemma 3.4. *Let $u_{jn}(\epsilon)$ and $u^{jn}(\epsilon)$ be given in Lemma 3.3. Then, the sequences of constants $u_{jn}(\epsilon)$ and $u^{jn}(\epsilon)$, $j = 1, \ldots, n$, satisfy the relation*

$$u_{jn}(\epsilon) < \frac{j}{n+1} < u^{jn}(\epsilon).$$

Remark 3.1. *Assume that $1 - n^{-\delta(n)} \to 0$, as $n \to \infty$. Then, the sequences of constants $u_{jn}(\epsilon)$ and $u^{jn}(\epsilon)$, given in Lemma 3.3, satisfy the relation $\sup_{1\le j\le n} (u^{jn}(\epsilon) - u_{jn}(\epsilon)) = o(1)$ (cf. Glivenko-Cantelli theorem).*

Remark 3.2. *If we take all $u_{jn}(\epsilon)$ and $u^{jn}(\epsilon)$ from the Kolmogorov's inequality, then $u_{jn}(\epsilon) = 1 - e^{-\tilde{\nu}_{jn} + \sqrt{2/\epsilon}}$ and $u^{jn}(\epsilon) = 1 - e^{-\tilde{\nu}_{jn} - \sqrt{2/\epsilon}}$, $j = 1, \ldots, n$. Under these settings, $\sup_{1\le j\le n} (u^{jn}(\epsilon) - u_{jn}(\epsilon)) \neq o(1)$. Also, the positive function $R(j)$, defined in Lemma B.1, is not increasing on $1 \le j \le n$.*

Lemma 3.5. *Let $\tilde{\nu}_{jn}$ be given in Lemma 3.3, and let $\nu_{jn} = -\log(1 - j/(n+1))$, $j = 1, \ldots, n$. Assume that $1 - n^{-\delta(n)} \to 0$ and $n^{1/2}(1 - n^{-\delta(n)}) \to \infty$, as $n \to \infty$. Then, the sequences of constants $v_{jn}(\epsilon)$ and $v^{jn}(\epsilon)$, given in Lemma 3.3, satisfy the relations*

$$v_{jn}(\epsilon) < \nu_{jn} < \tilde{\nu}_{jn} < v^{jn}(\epsilon) \quad \text{and} \quad v^{jn}(\epsilon) - v_{jn}(\epsilon) \le K(\epsilon),$$

where $K(\epsilon)$ is independent of n.

In this dissertation, we use the original forms of Assumptions A and C (restate above for convenience), but a slightly stronger version of Assumption B* (say Assumption B; it is stated below). (Note that the simultaneous bounds of the exponential order statistics, $v_{jn}(\epsilon)$ and $v^{jn}(\epsilon)$ used in Assumption B, are given in Lemma 3.3 and are different from those in CGJ1967.)

Assumption B: For each $\epsilon > 0$,

$$\begin{aligned} A_n = &\sum_{j=n-k_n+1}^{n} \left[\left\{ \sup_{v_{jn}(\epsilon) < v < v^{jn}(\epsilon)} \left| \tilde{H}'_F(v) - \tilde{H}'_F(\tilde{\nu}_{jn}) \right| \right\} \sqrt{\frac{j}{n-j+1}} \right] \\ &= o(n\sigma_{F,n}(k_n)), \end{aligned}$$

where $v_{jn}(\epsilon)$, $v^{jn}(\epsilon)$, and $\tilde{\nu}_{jn}$ are given in Lemma 3.3.

The main result of CGJ1967 is stated next.

Theorem 3.2 (Chernoff, Gastwirth, and Johns, 1967). *If Assumptions A, B, and C are satisfied, then we have*

$$T^*_{F,L}(k_n) = \frac{T_{F,L}(k_n) - n\mu_{F,n}(k_n)}{\sqrt{n}\sigma_{F,n}(k_n)} \overset{d}{\to} N(0,1), \quad as \quad n \to \infty.$$

3.2 Single Sequence of $N(0,1)$ Random Variables

In this section we will apply the approach of CGJ1967 to develop order threshold test procedures for testing the simple hypothesis

$$H_0 : \theta_i = 0, \ \forall i \ \text{ versus } \ H_a : H_0 \text{ is false} \tag{3.7}$$

based on a sequence of observations $X_i, i = 1, \ldots, n$, where $X_i \sim N(\theta_i, 1)$. The asymptotic null distribution of the order threshold statistic given by (1.1) is derived in the next subsection, while simulation results comparing the power of the hard threshold statistic, a power-enhanced version of the Simes (1986) statistic, and that of order threshold statistics are presented in Subsection 3.2.2.

3.2.1 The Asymptotic Null Distribution

Let $X_i, i = 1, \ldots, n$, be standard normal random variables, and let

$$T_L(k_n) = \sum_{i=1}^{n} c_{in} Y_{i,n} = \sum_{i=n-k_n+1}^{n} Y_{i,n}, \tag{3.8}$$

where $Y_i = X_i^2$, $Y_{1,n} < \cdots < Y_{n,n}$ are the ordered Y_i's, $c_{in} = I(i > n - k_n)$, and k_n is the order threshold parameter. The approach of CGJ1967 is based on the representation

$$T_L(k_n) \overset{d}{=} \sum_{i=n-k_n+1}^{n} \tilde{H}\left(V_{i,n}\right),$$

where $V_{i,n},\ i = 1, \ldots, n$, are the ordered observations from an i.i.d. sequence of $Exp(1)$ random variables, and

$$\tilde{H}(v) = F^{-1} \circ G(v)$$

with $F(y) = \dfrac{1}{\sqrt{2\pi}} \displaystyle\int_0^y t^{-1/2} e^{-t/2} \mathrm{d}t,\ y > 0$, and $G(v) = 1 - e^{-v}$, $v \geq 0$. Let

$$\mu_n(k_n) = \frac{1}{n} \sum_{i=1}^{n} c_{in} \tilde{H}\left(\tilde{\nu}_{in}\right), \quad \sigma_n^2(k_n) = \frac{1}{n} \sum_{i=1}^{n} \alpha_{in}^2(k_n), \tag{3.9}$$

where

$$\alpha_{in}(k_n) = \frac{1}{n-i+1} \sum_{j=i}^{n} c_{jn} \tilde{H}'\left(\tilde{\nu}_{jn}\right), \quad \tilde{\nu}_{in} = \sum_{j=1}^{i} \frac{1}{n-j+1}. \tag{3.10}$$

The term of $\alpha_{in}(k_n)$ can be reexpressed as $\alpha_{in}(k_n) = \dfrac{1}{n-i+1} \sum_{j=n-k_n+1}^{n} \tilde{H}'\left(\tilde{\nu}_{jn}\right)$ for $i \leq n - k_n$ and $\alpha_{in}(k_n) = \dfrac{1}{n-i+1} \sum_{j=i}^{n} \tilde{H}'\left(\tilde{\nu}_{jn}\right)$ for $i > n - k_n$ with $\tilde{H}'\left(\tilde{\nu}_{jn}\right) = \dfrac{e^{-\tilde{\nu}_{jn}}}{f\left(F^{-1}\left(1 - e^{-\tilde{\nu}_{jn}}\right)\right)}$ and the function f is the derivative of F. With this notation we have the following.

Theorem 3.3. *Let $Y_i,\ i = 1, \ldots, n$, be a sequence of i.i.d. random variables having the central chi-squared distribution with 1 degree of freedom. Let $k_n,\ n \geq 1$, be any sequence of integers which satisfies $k_n \to \infty$, as $n \to \infty$, and $k_n \leq n$. Let $\mu_n(k_n)$ and $\sigma_n^2(k_n)$ be as in (3.9) with $c_{in} = I\left(i > n - k_n\right)$, and let $T_L(k_n)$ be given in (3.8). Then we have*

$$T_L^*(k_n) = \frac{T_L(k_n) - n\mu_n(k_n)}{\sqrt{n}\sigma_n(k_n)} \xrightarrow{d} N(0,1), \quad as \quad n \to \infty. \tag{3.11}$$

Note that the asymptotic mean of $T_L(k_n)$ is $n\mu_n(k_n)$ and the asymptotic variance of $T_L(k_n)$ is $n\sigma_n^2(k_n)$ as k_n tends to infinity with n.

3.2.2 Simulations

Table A.3 presents the achieved alpha levels of the order threshold statistic $T_L(k_n)$ for different values of the order threshold parameter k_n. The results are based on 30,000 simulations runs. For $T_L(k_n)$, note that in Table A.3 with $n = 500$, k_n ranges from 2 to 500, and the achieved alpha levels are close to the true value of 0.05. In contrast, for the hard threshold statistic $T_H(\delta_n)$, note that in Table A.1 the deterioration of the achieved alpha levels happens as the hard threshold parameter δ_n decreases over a relatively small range. Moreover, we note that the slightly liberal α levels of the $T_L(k_n)$ statistic can be corrected by the use of a multiple of a χ^2 distribution to approximate its finite sample distribution. Thus, using the approximation $T_L(k_n) \dot{\sim} b\chi^2_\nu$, where b and ν are chosen to match the mean and variance of $T_L(k_n)$, results in the type I error rates shown in Table A.4. Note that the results of Table A.4 are a little more conservative than those of Table A.3.

The greater flexibility in the choice of the threshold parameter that the order threshold statistic offers does not come at the expense of the rate with which it converges to its asymptotic distribution. To emphasize this aspect, Figure A.1 presents the estimated densities of the hard thresholding (solid lines in the upper panel) and order thresholding test statistics (solid lines in the lower panel), based on 20,000 simulated values of each statistic using $n = 200$. The threshold parameters of the hard thresholding and order thresholding test statistics have been chosen so that the average number of observations included in the two statistics are the same in each column. We see that the estimated densities of the order threshold statistic are closer to the standard normal density (dash-dot line) than those of the hard threshold statistic. In particular, the estimated densities in the upper panel show the rapid deterioration of the quality of the normal approximation to the distribution of $T_H(\delta_{200})$ as δ_{200} shifts away from recommended value of $\delta_{200} = 2\log(200\log^{-2}200) = 3.9271$.

From now, we compare the empirical power of the order threshold statistic using several values of the threshold parameter with those of the hard threshold and a power-enhanced version of the Simes (1986) statistics. The simulations reported here use samples of size $n = 500$ generated from the normal distribution with variance 1. The threshold parameter k_{500} of the order threshold statistics takes values of 15, 40, 70, 100, 200, 500, as well as a data-driven value, denoted by $\widehat{k}_{500}^{opt}$, whose description is given in Section 3.3. The empirical power using the approximation $T_L(\widehat{k}_{500}^{opt}) \dot{\sim} b\chi_\nu^2$ is reported together with that using the normal approximation to $T_L(\widehat{k}_{500}^{opt})$. The hard threshold statistic we consider uses the recommended value of the threshold parameter which is $\delta_{500} = 2\log(500\log^{-2}500) = 5.1216$. All results are based on 3,000 simulation runs. Since the the global hypothesis H_0^G, all $H_0^{(i)} : \theta_i = 0$ are true, is the same for all three simulation settings, the type I error rates reported in the last row of Table A.5 pertain also to Tables A.6 and A.7. Note that all achieved significance levels are below 0.06. The alternatives considered have 30 of the 500 mean values different from zero. In particular, we consider the following sequence of alternatives indexed by r:

$$H_r : \theta_j = \eta_{j+r-1} \quad \text{for} \quad j = 1, \dots, 500, \quad r = 1, \dots, 30,$$

where η_j, $j = 1, 2, \dots$, is a given sequence. The following are examples with different values of $\boldsymbol{\eta}$.

Example 3.1. *We generate the values of η_j, $j = 1, \dots, 30$, from $N(1.5, 1)$. The rest values of η_j are 0. The values different from 0 are as follows:*

$$(1.0674, -0.1656, 1.6253, 1.7877, 0.3535, 2.6909, 2.6892, 1.4624, 1.8273, 1.6746, 1.3133, 2.2258, 0.9117, 3.6832, 1.3636, 1.6139, 2.5668, 1.5593, 1.4044, 0.6677, 1.7944, 0.1638, 2.2143, 3.1236, 0.8082, 2.7540, -0.0937, 0.0590, 2.0711, 2.3579).$$

Note that $\#\big(j : 0 < |\eta_j| \le 1,\ j = 1, 2, \dots\big) = 8$, $\#\big(j : 1 < |\eta_j| \le 2,\ j = 1, 2, \dots\big) = 12$, $\#\big(j : 2 < |\eta_j| \le 3,\ j = 1, 2, \dots\big) = 8$, and $\#\big(j : |\eta_j| > 3,\ j = 1, 2, \dots\big) = 2$.

Example 3.2. *We generate the values of η_j, $j = 1, \dots, 30$, from the standard exponential distribution. The remaining values of η_j are 0. The values different from 0 are as follows:*

$$(0.0512, 1.4647, 0.4995, 0.7216, 0.1151, 0.2716, 0.7842, 3.7876, 0.1967, 0.8103, 0.4854, 0.2332, 0.5814, 0.3035, 1.7357, 0.9021, 0.0667, 0.0867, 0.8909, 0.1124, 2.8491, 1.0416, 0.2068, 2.6191, 1.9740, 1.5957, 1.6158, 0.5045, 1.3012, 1.6153).$$

Note that $\#\big(j : 0 < \eta_j \leq 1,\ j = 1, 2, \dots\big) = 19$, $\#\big(j : 1 < \eta_j \leq 2,\ j = 1, 2, \dots\big) = 8$, $\#\big(j : 2 < \eta_j \leq 3,\ j = 1, 2, \dots\big) = 2$, *and* $\#\big(j : \eta_j > 3,\ j = 1, 2, \dots\big) = 1$.

Example 3.3. *In this example the values of* η_j, $j = 1, \dots, 30$, *are 2.0 and the rest are zero.*

As expected, the power in each column decreases by increasing r because the number of $\boldsymbol{\theta}$ with values different from zero (denoted by k_{500}^{opt}) decreases. When the θ_i with the large value such as 3.6832, 3.1236 (in Example 3.1), and 3.7876 (in Example 3.2) is excluded at the alternative, the large decrement in the power occurs. For each alternative the statistic $T_L(15)$ or $T_L(40)$ achieves better power than the order threshold statistics with the other specified values of the threshold parameter. This is a consequence of the fact that the number of mean values that are different from zero never exceeds 30. Thus, less noise is incorporated in $T_L(k_{500}^{opt})$ than the other order threshold statistics. Note that with the chosen value of $\delta_{500} = 5.1216$, the hard threshold statistic uses, on average, 12 observations. Thus, it is rather surprising that the empirical power of the hard threshold statistic is always smaller than that of $T_L(15)$. In all three tables, the empirical power using the approximation $T_L(\widehat{k}_{500}^{opt}) \dot{\sim} \Phi$ is similar to that of $T_L(k_{500}^{opt})$, and always greater than the empirical powers of the hard threshold and Simes statistics. The empirical power using the approximation $T_L(\widehat{k}_{500}^{opt}) \dot{\sim} b\chi_\nu^2$ is a little bit smaller than that using the normal approximation, however, it is still greater than the empirical powers of the hard threshold and Simes statistics. In Table A.7, for large number of the false null hypotheses the Simes statistic T_S performs much worse than the hard threshold statistic, the order threshold statistic, and even the chi-square statistic $T_L(500)$. In all three tables, the power of $T_H(5.1216)$ is similar (though somewhat smaller) to that of $T_L(100)$. Finally, all order threshold statistics achieved higher power than the chi-square statistic $T_L(500)$.

3.3 Choosing k_n

The simulation results and the discussion in the closing paragraph of Subsection 3.2.2 suggest that the power of $T_L(k_n)$ is largest when k_n equals the number of mean values different from zero (denoted by k_n^{opt}). As a data-driven choice of k_n, we propose to use the estimate of k_n^{opt} suggested by Storey (2002, 2003) and Efron, Tibshirani, Storey, and Tusher (2001), which is

$$\widehat{k}_n^{opt}(\lambda) \quad = \quad \max\left\{\frac{n\widehat{G}_n(\lambda) - n\lambda - 1}{1-\lambda},\ \log^{3/2} n\right\},$$

where $\widehat{G}_n$ is the empirical cdf of $\mathbf{P}^n = (P_1, \ldots, P_n)$, the P_is are the p-values of the individual hypotheses $H_0^{(i)} : \theta_i = 0$, $i = 1, \ldots, n$, and λ is the median of the P_is. The recommended lower bound $\log^{3/2} n$ of $\widehat{k}_n^{opt}(\lambda)$ was found to be preferable in the simulations we performed. Interestingly, $\log^{3/2} n$ equals the expected number of observations in hard thresholding with the recommended threshold parameter of $\delta_n = 2\log(n \log^{-2} n)$.

Chapter 4

One-Way HANOVA

Let the X_{ij}, $i = 1, \dots, a$, $j = 1, \dots, n$, be independent $N(\theta_i, \sigma^2)$, where the θ_i and σ^2 are all unknown. Let $\alpha_i = \theta_i - \overline{\theta}$ denote the "effect" of the ith group, and consider testing $H_0 : \alpha_1 = \cdots = \alpha_a = 0$ vs. $H_a : H_0$ is false. Akritas and Papadatos (2004) show that the asymptotic power of the optimal invariant ANOVA F test equals its level of significance even when $||\boldsymbol{\alpha}|| \to \infty$, as $a \to \infty$, with $||\boldsymbol{\alpha}||^2 = o(\sqrt{a})$. Because the power of the chi-square statistic (1.2) has a similar property, an extension of the order thresholding to the one-way HANOVA setting is expected to result in similar gains in power over the ANOVA F test.

In Section 4.1, we extend the applicability of order thresholding to the one-way HANOVA context, while Section 4.2 illustrates the improved power of order thresholding via simulation. Finally, using the idea of Storey (2002, 2003) and the simulation results, we present a recommendation for a data-driven choice of the order threshold parameter in Section 4.3.

4.1 Order Thresholding in One-Way HANOVA

The classical F statistic is given by

$$F_a = \frac{MST}{MSE}, \tag{4.1}$$

where

$$MST = \frac{1}{a-1}\sum_{i=1}^{a} n\left(\overline{X}_{i.} - \overline{X}_{..}\right)^2, \quad MSE = \frac{1}{N-a}\sum_{i=1}^{a}\sum_{j=1}^{n}\left(X_{ij} - \overline{X}_{i.}\right)^2,$$

with $\overline{X}_{i.} = n^{-1}\sum_{j=1}^{n} X_{ij}$, $\overline{X}_{..} = N^{-1}\sum_{i=1}^{a}\sum_{j=1}^{n} X_{ij}$, and $N = an$. Note that

$$(a-1)F_a = \sum_{i=1}^{a}\left(\frac{\sqrt{n}\left(\overline{X}_{i.} - \overline{X}_{..}\right)}{\sqrt{MSE}}\right)^2 \tag{4.2}$$

differs from the chi-square statistic (1.2) only in that the random variables which are being summed are not independent, and their distribution is not χ_1^2. Set

$$\widetilde{Z}_i = \frac{\sqrt{n}\left(\overline{X}_{i.} - \overline{X}_{..}\right)}{\sqrt{MSE}}, \quad \widehat{Z}_i = \frac{\sqrt{n}\left(\overline{X}_{i.} - \overline{X}_{..}\right)}{\sigma}.$$

Thus,

$$\widetilde{Z}_i = s\widehat{Z}_i, \quad \text{where} \quad s = \frac{\sigma}{\sqrt{MSE}}.$$

Threshold versions of (4.2) are of the form

$$\widehat{T}_L(k_a) = \sum_{i=1}^{a} c_{ia}\widetilde{Z}_{i,a}^2 = s^2\sum_{i=1}^{a} c_{ia}\widehat{Z}_{i,a}^2, \tag{4.3}$$

where $\widehat{Z}_{1,a}^2 < \cdots < \widehat{Z}_{a,a}^2$ are the ordered $\widehat{Z}_i^2$'s, $\widetilde{Z}_{i,a} = s\widehat{Z}_{i,a}$, $c_{ia} = I(i > a - k_a)$, and k_a is the order threshold parameter. For suitable centering and scaling constants, $\widehat{\mu}_a(k_a)$ and $\widehat{\sigma}_a(k_a)$, the asymptotic theory of $\widehat{T}_L(k_a)$ will use the decomposition

$$\begin{aligned}\frac{\widehat{T}_L(k_a) - a\widehat{\mu}_a(k_a)}{\sqrt{a}\widehat{\sigma}_a(k_a)} &= s^2\frac{1}{\sqrt{a}\widehat{\sigma}_a(k_a)}\left(\sum_{i=1}^{a} c_{ia}\widehat{Z}_{i,a}^2 - a\widehat{\mu}_a(k_a)\right) \\ &+ \frac{\sqrt{a}}{\widehat{\sigma}_a(k_a)}\widehat{\mu}_a(k_a)\left(s^2 - 1\right). \end{aligned} \tag{4.4}$$

The two components in (4.4) are independent, so it suffices to show the asymptotic normality of each one separately. To deal with the first component, let θ_0 denote the common value of the θ_i under H_0 and write

$$\widehat{Z}_i = Z_i + \frac{t}{\sqrt{a}}, \quad \text{where} \quad Z_i = \frac{\sqrt{n}\left(\overline{X}_{i\cdot} - \theta_0\right)}{\sigma}, \quad \text{and} \quad t = -\frac{\sqrt{N}\left(\overline{X}_{\cdot\cdot} - \theta_0\right)}{\sigma}. \tag{4.5}$$

Our approach for obtaining the asymptotic distribution of $\sum_{i=1}^{a} c_{ia}\widehat{Z}_{i,a}^2$ is to first derive its asymptotic distribution treating the t in (4.5) as fixed, and then to show that the convergence is uniform over all values of t bounded by any positive constant M. By Slutsky's theorem, the asymptotic distribution of the first component of (4.4) is the same as that of $\sum_{i=1}^{a} c_{ia}\widehat{Z}_{i,a}^2$. The asymptotic distribution of the second component of (4.4) is easily derived since

$$\sqrt{a}(s^2 - 1) \ \xrightarrow{d}\ N\left(0, \frac{2}{n-1}\right), \quad \text{as} \quad a \to \infty,$$

and, as it will be shown in lemmas described in Subsection 4.1.3, $\widehat{\mu}_a(k_a)/\widehat{\sigma}_a(k_a) \to \mu_r/\sigma_r$, provided that $k_a/a \to r$ for some $0 \le r \le 1$ and $k_a \to \infty$, as $a \to \infty$.

4.1.1 Asymptotic Distribution when t is Fixed

When t is fixed, we set

$$Z_{t,i} = Z_i + \frac{t}{\sqrt{a}}, \ i = 1, \ldots, a, \quad T_L^t(k_a) = \sum_{i=1}^{a} c_{ia} Z_{t,(i)}^2, \tag{4.6}$$

where $Z_{t,(1)}^2 < \cdots < Z_{t,(a)}^2$ are the order statistics of $Z_{t,1}^2, \ldots, Z_{t,a}^2$. (Note that for t as defined in (4.5), $Z_{t,i}$ becomes $\widehat{Z}_i$.) It follows that the $Z_{t,i}^2$ are independent $\chi_1^2\left(t^2/a\right)$ so that their density and cumulative distribution functions are given by

$$\begin{aligned} g_{a,t}(y) &= \frac{e^{-\frac{1}{2}\left(y + \frac{t^2}{a}\right)} y^{-1/2}}{2^{1/2}} \sum_{k=0}^{\infty} \frac{(\frac{t^2}{a}y)^k}{2^{2k}k!\Gamma\left(k+1/2\right)}, \quad y > 0, \quad \text{and} \\ G_{a,t}(y) &= \int_0^y g_{a,t}(u)\mathrm{d}u = \sum_{k=0}^{\infty} e^{-\frac{t^2}{2a}} \frac{1}{2^k \cdot k!}\left(\frac{t^2}{a}\right)^k G_{2k+1}(y), \quad y > 0, \end{aligned}$$

respectively, where

$$G_k(y) = \frac{1}{2^{k/2}\Gamma(k/2)} \int_0^y u^{k/2-1} e^{-u/2} \mathrm{d}u, \quad y > 0,$$

is the cumulative distribution function of $\chi_k^2(0)$. Let

$$\mu_a^t(k_a) = \frac{1}{a}\sum_{i=1}^{a} c_{ia} G_{a,t}^{-1}(1-e^{-\tilde{\nu}_{ia}}) \text{ and } (\sigma_a^t(k_a))^2 = \frac{1}{a}\sum_{i=1}^{a} \left(\alpha_{ia}^t(k_a)\right)^2, \tag{4.7}$$

where $\alpha_{ia}^t(k_a) = \dfrac{1}{a-i+1} \sum_{j=i}^a c_{ja} \dfrac{e^{-\tilde{\nu}_{ja}}}{g_{a,t}\left(G_{a,t}^{-1}(1-e^{-\tilde{\nu}_{ja}})\right)}$ and $\tilde{\nu}_{ia} = \sum_{j=1}^i \dfrac{1}{a-j+1}$. We note that the order threshold parameter k_a depends on a, not n. This detail aside, the parameters in (4.7) with $t = 0$ are the same as those in (3.9). With this notation we have the following lemma. This lemma is related to Lemma 3.2.

Lemma 4.1 (Chernoff, Gastwirth, and Johns, 1967)**.** *Let $T_L^t(k_a)$ and $\mu_a^t(k_a)$ be as defined in* (4.6) *and* (4.7)*, respectively. Let $V_1, \ldots, V_a$ be i.i.d. from Exp*(1) *random variables and let $V_{1,a} < \cdots < V_{a,a}$ be the corresponding order statistics. Then $a^{-1}T_L^t(k_a)$ can be decomposed as*

$$a^{-1}T_L^t(k_a) \stackrel{d}{=} \mu_a^t(k_a) + Q_a^t(k_a) + R_a^t(k_a),$$

where

$$Q_a^t(k_a) = \frac{1}{a}\sum_{i=1}^{a} \alpha_{ia}^t(k_a)(V_i - 1) \tag{4.8}$$

and

$$R_a^t(k_a) = \frac{1}{a}\sum_{i=1}^{a} c_{ia} \left\{ \left(G_{a,t}^{-1}\left(1-e^{-V_{i,a}}\right) - G_{a,t}^{-1}\left(1-e^{-\tilde{\nu}_{ia}}\right)\right) - \frac{\left(V_{i,a}-\tilde{\nu}_{ia}\right) e^{-\tilde{\nu}_{ia}}}{g_{a,t}\left(G_{a,t}^{-1}\left(1-e^{-\tilde{\nu}_{ia}}\right)\right)} \right\}$$

with $\alpha_{ia}^t(k_a) = \dfrac{1}{a-i+1} \sum_{j=i}^a c_{ja} \dfrac{e^{-\tilde{\nu}_{ja}}}{g_{a,t}\left(G_{a,t}^{-1}(1-e^{-\tilde{\nu}_{ja}})\right)}$ and $\tilde{\nu}_{ia} = \sum_{j=1}^i \dfrac{1}{a-j+1}$.

Theorem 4.1. *For any fixed value of t, let $Z_{t,i}^2$, $i = 1, \ldots, a$, be a sequence of i.i.d. random variables having the noncentral chi-squared distribution with 1 degree of freedom and non-centrality parameter t^2/a. Let k_a, $a \geq 1$, be any sequence of* *integers which satisfies* $k_a \to \infty$, *as* $a \to \infty$, *and* $k_a \leq a$. *Let* $\mu_a^t(k_a)$ *and* $(\sigma_a^t(k_a))^2$ *be as in (4.7) with* $c_{ia} = I(i > a - k_a)$, *and let* $T_L^t(k_a)$ *be given in* (4.6). *Then we have*

$$T_L^{t\,*}(k_a) = \frac{T_L^t(k_a) - a\mu_a^t(k_a)}{\sqrt{a}\sigma_a^t(k_a)} \overset{d}{\to} N(0,1), \quad as \quad a \to \infty. \tag{4.9}$$

4.1.2 Uniformity of the Convergence in Distribution

This subsection shows that, for any positive constant M, the distribution function of (4.9) converges to the standard normal distribution uniformly on $|t| < M$.

Lemma 4.2. *Consider the setting of Theorem 4.1. Let* $H_{a,t}$ *be the distribution function of* $\sqrt{a}Q_a^t(k_a)/\sigma_a^t(k_a)$, *where* $Q_a^t(k_a)$ *is given in (4.8), and let* Φ *be the standard normal distribution function. Then, for any* $M > 0$,

$$\sup_{\substack{-M<t<M \\ -\infty<x<\infty}} |H_{a,t}(x) - \Phi(x)| \ \to \ 0, \quad as \quad a \to \infty.$$

Lemma 4.3. *Consider the setting of Theorem 4.1, and let* $R_a^t(k_a)$ *be as given in Lemma 4.1. Then, for any* $M > 0$,

$$\sup_{-M<t<M} \left| \frac{\sqrt{a}R_a^t(k_a)}{\sigma_a^t(k_a)} \right| \overset{p}{\to} 0, \quad as \quad a \to \infty.$$

Lemma 4.4. *Consider the setting of Theorem 4.1. Let* $F_{a,t}$ *be the distribution function of* $T_L^{t\,*}(k_a)$ *given in (4.9) and let* Φ *be the standard normal distribution function. Then, for any* $M > 0$,

$$\sup_{\substack{-M<t<M \\ -\infty<x<\infty}} |F_{a,t}(x) - \Phi(x)| \ \to \ 0, \quad as \quad a \to \infty.$$

Theorem 4.2. *Let* k_a, $a \geq 1$, *be any sequence of integers which satisfies* $k_a \to \infty$, *as* $a \to \infty$, *and* $k_a \leq a$. *For* t *as defined in* (4.5), *let* $\widehat{Z}_i$, $\widehat{\mu}_a(k_a)$ *and* $(\widehat{\sigma}_a(k_a))^2$ *be as in* (4.6), (4.7), *respectively. Then we have*

$$\widehat{T}_L^*(k_a) = \frac{\sum_{i=1}^a c_{ia}\widehat{Z}_{i,a}^2 - a\widehat{\mu}_a(k_a)}{\sqrt{a}\widehat{\sigma}_a(k_a)} \overset{d}{\to} N(0,1), \quad as \quad a \to \infty, \tag{4.10}$$

where $\widehat{Z}^2_{1,a} < \cdots < \widehat{Z}^2_{a,a}$ *are the ordered* $\widehat{Z}^2_i$*'s and* $c_{ia} = I(i > a - k_a)$.

4.1.3 Asymptotic Normality of Order Threshold Statistics

In this subsection, it is first shown that $\mu^t_a(k_a)$ and $\sigma^t_a(k_a)$ converge to $\mu^0_a(k_a)$ and $\sigma^0_a(k_a)$, respectively, uniformly on $|t| < M$. This fact is then used in Theorem 4.3 for obtaining the asymptotic normality of the order threshold statistic given in (4.3).

Lemma 4.5. *Let* $k_a,\ a \geq 1$*, be any sequence of integers which satisfies* $k_a \to \infty$*, as* $a \to \infty$*, and* $k_a \leq a$*. Let* $(\sigma^t_a(k_a))^2$ *and* $(\sigma^0_a(k_a))^2$ *be as in (4.7) with any* t *and fixed value of* $t = 0$*, respectively. Then, for any* $M > 0$*,*

$$\sup_{-M<t<M} \left| \frac{\sigma^t_a(k_a)}{\sigma^0_a(k_a)} - 1 \right| \ \to\ 0, \quad as \ \ a \to \infty.$$

Lemma 4.6. *Let* $k_a,\ a \geq 1$*, be any sequence of integers which satisfies* $k_a \to \infty$*, as* $a \to \infty$*, and* $k_a \leq a$*. Let* $\mu^t_a(k_a)$*,* $\mu^0_a(k_a)$ *and* $(\sigma^0_a(k_a))^2$ *be as in (4.7) with any* t*, fixed value of* $t = 0$*, respectively. Then, for any* $M > 0$*,*

$$\sup_{-M<t<M} \left| \frac{\sqrt{a}\,(\mu^t_a(k_a) - \mu^0_a(k_a))}{\sigma^0_a(k_a)} \right| \ \to\ 0, \quad as \ \ a \to \infty.$$

Lemma 4.7. *Let* $\mu^0_a(k_a)$ *and* $(\sigma^0_a(k_a))^2$ *be as in (4.7) with the fixed value of* $t = 0$*. Then, provided that* $k_a/a \to r$ *for some* $0 \leq r \leq 1$ *and* $k_a \to \infty$*, as* $a \to \infty$*, we have*

$$\mu^0_a(k_a) \to \mu_r \quad and \quad (\sigma^0_a(k_a))^2 \to \sigma^2_r, \quad as \ \ a \to \infty,$$

where

$$\mu_r = \int_0^1 I(t > 1 - r) G^{-1}_{a,0}(t)\mathrm{d}t$$

and

$$\sigma^2_r = \int_0^1 \int_0^1 I(t > 1 - r) I(s > 1 - r)(\min(t, s) - ts)\mathrm{d}G^{-1}_{a,0}(t)\mathrm{d}G^{-1}_{a,0}(s).$$

Remark 4.1. *If* $r = 1$, *then*

$$\frac{\mu_a^0(k_a)}{\sigma_a^0(k_a)} \rightarrow \frac{1}{\sqrt{2}}, \quad \text{as} \ \ a \rightarrow \infty.$$

From Theorem 4.2 and lemmas described earlier in this subsection, we can obtain the following theorem.

Theorem 4.3. *Let* $\mu_a^0(k_a)$, $(\sigma_a^0(k_a))^2$, μ_r, *and* σ_r^2 *be as in Lemma 4.7, and let* $\widehat{T}_L(k_a)$ *be given in (4.3). Then, provided that* $k_a/a \rightarrow r$ *for some* $0 \leq r \leq 1$ *and* $k_a \rightarrow \infty$, *as* $a \rightarrow \infty$, *we have*

$$\widetilde{T}_L(k_a) = \frac{\widehat{T}_L(k_a) - a\mu_a^0(k_a)}{\sqrt{a}\sigma_a^0(k_a)} \xrightarrow{d} N\left(0, 1 + \frac{2\mu_r^2}{\sigma_r^2(n-1)}\right), \quad \text{as} \ \ a \rightarrow \infty. \tag{4.11}$$

4.2 Simulations

In this section, we compare the performance of the classical F statistic, given in (4.1), and the order threshold statistics $\widetilde{T}_L(k_a)$, given in (4.11).

We remark that Fan and Lin (1998) applied the thresholding methodology to the problem of comparing I curves with data arising from the model $X_{ij}(t) = f_i(t) + \epsilon_{ij}(t),\ t = 1, \ldots, T,\ j = 1, \ldots, n_i,\ i = 1, \ldots, I$. Their asymptotic theory pertains to the case where the number of curves which are compared, I, remains fixed, while T and the sample sizes n_i tend to infinity. This problem is fundamentally different from that considered here, and their procedure is not a competitor to ours.

Figure A.2 presents the estimated densities of $\widetilde{T}_L(k_{500})$ (solid line) and the density of the limiting normal distribution (dash-dot line). The estimated densities are based on 20,000 simulated values, using $a = 500$ and $n = 3$, when the threshold parameter k_{500} takes the values of $[a^{1/2}] = 22$, $[a^{3/4}] = 105$, and $[a^{7/8}] = 229$. It can be seen that the approximation is quite good especially for $k_{500} = 105$ and 229. Similar figures (e.g. Figure A.3) with different values of n suggest that the rate of convergence of the order threshold statistic to its limiting distribution is mainly driven by a, not n.

The results reported in Table A.8 are based on 20,000 simulation runs. As expected, the distributions of $\widetilde{T}_L(k_a)$ converge to the normal distribution function and the achieved alpha levels are close to the true value of 0.05. Thus, the asymptotic theory of the order threshold statistics provides a good approximation. More exactly, when the number of groups are larger than 200, all order threshold statistics are robust for the 0.05 significance level. In particular, the achieved alpha level of $\widetilde{T}_L(k_{1000})$ is 0.0507 when $a = 1000$, $n = 5$, $k_{1000} = [a^{7/8}] = 421$, and 0.0484 when $a = 1000$, $n = 5$, $k_{1000} = [a^{3/4}] = 177$.

From now, we compare the empirical power of $\widetilde{T}_L(k_{1000})$ using several values of the threshold parameter with that of the classical F statistic. The simulations use samples of size $a = 1000$ and $n = 5$ generated from the normal distribution with variance 1. The threshold parameter k_{1000} is 20, 50, 100, 250, 500, and 1000. All results are based on 20,000 simulation runs. The alternatives here have 20 of the 1000 θ_i values different from zero. In particular, we consider the following sequence of alternatives indexed by r:

$$H_r : \theta_j = \eta_{j+r-1} \ \text{ for } \ j = 1, \ldots, 1000, \quad r = 1, \ldots, 20,$$

where η_j, $j = 1, 2, \ldots$, is a given sequence. The following are examples with different values of $\boldsymbol{\eta}$.

Example 4.1. *We generate the values of η_j, $j = 1, \ldots, 20$, from $Uniform(-2, 2)$. The remaining values of η_j are 0. The values different from 0 are as follows:*

$$\begin{aligned}(&1.8005, -1.0754, 0.4274, -0.0561, 1.5652, 1.0484, -0.1741, -1.9260, 1.2856,\\ &-0.2212, 0.4617, 1.1677, 1.6873, 0.9528, -1.2949, -0.3772, 1.7419, 1.6676,\\ &-0.3589, 1.5746).\end{aligned}$$

Note that $\#\big(j : 0 < |\eta_j| \leq 1,\ j = 1, 2, \ldots\big) = 8$ and $\#\big(j : |\eta_j| > 1,\ j = 1, 2, \ldots\big) = 12$.

Example 4.2. *We generate the values of η_j, $j = 1, \ldots, 20$, from $Exp(0.7)$. The remaining values of η_j are 0. The values different from 0 are as follows:*

$$\begin{aligned}(&1.0949, 0.5511, 1.7587, 0.1128, 0.4033, 0.7991, 0.6868, 0.0993, 0.6919, 1.8255,\\ &1.1272, 2.1041, 0.3975, 1.4730, 0.4549, 1.5015, 0.1830, 0.6865, 0.1360, 2.1458).\end{aligned}$$

Note that $\#\big(j : 0 < \eta_j \leq 1,\ j = 1, 2, \ldots\big) = 12$, $\#\big(j : 1 < \eta_j \leq 2,\ j = 1, 2, \ldots\big) = 6$, *and* $\#\big(j : \eta_j > 2,\ j = 1, 2, \ldots\big) = 2$.

As expected, the power in each column decreases as r increases and $\widetilde{T}_L(20)$ has the highest power. Since the number of θ_is that are different from zero does not exceed 20, $\widetilde{T}_L(20)$ minimizes the accumulation of noise, compared to the other order threshold statistics. For each alternative, the largest power differences between F_{1000} and $\widetilde{T}_L(20)$ are about 0.5 (alternative H_{13} in Table A.9) and 0.54 (alternative H_{12} in Table A.10). In both tables, the power of $\widetilde{T}_L(1000)$ is similar to that of F_{1000} because $\widetilde{T}_L(1000)$ is a standardized version of F_{1000}. Finally, all order threshold statistics achieved higher power than the classical F statistic, F_{1000}.

4.3 Choosing k_a

The simulation results and the discussion in the closing paragraph of Section 4.2 suggest that choosing k_a equal to the number of groups with nonzero effects, k_a^{opt}, maximizes the power. Our recommendation for the choice of the threshold parameter is based again on the idea of Storey (2002, 2003) for enhancing the power of Simes statistic for testing the constructed set of hypothesis testing problems $H_0^{(i)} : \theta_i = \overline{X}_{\cdot\cdot}$, $i = 1, \ldots, a$, where $\overline{X}_{\cdot\cdot}$ is the overall sample mean. The p-value for each hypothesis is approximated by

$$P_i = 2(1 - \Phi(|Z_i|)), \quad i = 1, \ldots, a,$$

with

$$Z_i = \frac{\overline{X}_{i\cdot} - \overline{X}_{\cdot\cdot}}{\sqrt{S_p^2/n}}, \quad i = 1, \ldots, a,$$

where $\overline{X}_{i\cdot}$ is the sample mean from the ith group and S_p^2 is the pooled sample variance. The power-enhanced version of the Simes statistic

$$T_S = \min_{1 \leq i \leq a} \left\{ aP_{(i)}/i \right\}$$

rejects the global null hypothesis if $T_S < \alpha/(1 - \widehat{k}_a^{opt}/a)$, with

$$\widehat{k}_a^{opt}(\lambda) \quad = \quad \max\left\{\frac{a\widehat{G}_a(\lambda) - a\lambda - 1}{1-\lambda},\ \log^{3/2} a\right\}, \tag{4.12}$$

where $\widehat{G}_a$ is the empirical cdf of $\mathbf{P}^a = (P_1, \ldots, P_a)$, $P_{(1)} < \cdots < P_{(a)}$ are the ordered P_i's, and λ is the median of the P_i's.

The simulation results shown in Table A.11 suggest that the power of $\widetilde{T}_L(\widehat{k}_{1000}^{opt})$ is similar to that of $\widetilde{T}_L(k_{1000}^{opt})$. These results are based on 2,000 simulation runs; the type I error rate of $\widetilde{T}_L(\widehat{k}_{1000}^{opt})$ was 0.048.

Chapter 5

Goodness-of-Fit Testing

Let $Y_1, \ldots, Y_n$ be a sample from a continuous population with distribution function G. We will consider testing the hypothesis $H_0 : G = G_0$, where G_0 is a completely specified distribution function. The classical Karl Pearson's 1900 chi-square goodness-of-fit test partitions the sample space in n_{bin} cells and considers the corresponding null hypothesis for the parameters of a multinomial distribution. As such, the chi-square test has been widely applied also in contingency tables. The asymptotic theory of this statistic, based on the vector of observed minus expected frequencies, was originally obtained under the assumption that the number of cells, n_{bin}, remains fixed as the number of observations, n, gets large. It was soon pointed out, however, that it is rather unnatural to keep n_{bin} fixed as the sample size tends to infinity. Mann and Wald (1942) were the first to establish the power advantages of letting n_{bin} tend to infinity with n, and found $n_{bin} = n^{2/5}$ to be the optimal rate. The corresponding development in the area of contingency tables was driven by the need for an asymptotic theory under a sparse multinomials scenario rather than optimality considerations; see Holst (1972), and Morris (1975) for theoretical results and Koehler and Larntz (1980) for empirical comparisons.

In this chapter, we will extend the applicability of order thresholding to Pearson's chi-square test when the number of cells grows with the sample size. A competing statistic, based on the cell averages of a score function will also be developed and compared with that based on the cell counts. The asymptotic theory pertains also to the use of a Fourier or wavelet transformation on the standardized observed minus expected cell frequencies, since such transformations have the potential of enhancing the power in both high and low signal-to-noise ratio settings, i.e. settings where either many or only few of the cell probabilities differ from the null probabilities. The main theoretical development lies in extending the asymptotic theory of the order threshold statistic to the case where the variables X_i (which are either the standardized cell counts or the standardized cell averages of the chosen score function) are only approximately normally distributed, and not quite independent. This double asymptotics scenario is approached by a strong approximation of the empirical process by a Brownian process in order to represent each X_i as a standard normal deviate plus an error term which tends to zero. This representation, and an argument that accounts for the fact that the ordering of the X_i's is not exactly the same as the ordering of the standard normal deviates, are used to establish the asymptotic theory of the order threshold statistic on the X_i's from that on the standard normal deviates.

Using the transformation $U_i = G_0(Y_i)$ we may, without loss of generality, consider testing the hypothesis $H_0 : G(t) = t,\ \forall 0 < t < 1$, where now G stands for the distribution function of the U_i. Moreover, testing the hypothesis of a uniform distribution is relevant in multiple testing.

Section 5.1 shows that the Fourier or wavelet transformation is used to ensure power enhancement in both high and low signal-to-noise ratio alternatives. In Section 5.2, we present the two order threshold statistics, based on the cell counts and the cell averages of the chosen score function, after a transformation. Section 5.3 considers the asymptotic theory of the two order threshold statistics, while Section 5.4 provides the basic steps for proving the asymptotic results. Simulation results comparing the power of the Mann and Wald (1942) version of Pearson's chi-square test, the Kolmogorov-Smirnov test, the Cramér-Von Mises test, and the hard threshold and the two order threshold statistics after each of the two transformations, are presented in Section 5.5.

5.1 Motivation of Transformation

In Chapter 3, we introduced the order thresholding procedure, whose asymptotic theory allows very flexible choice of the threshold parameter. Subsection 3.2.2 suggests that the order threshold statistic outperforms the chi-square, Simes, and hard threshold statistics. However, the success of any threshold procedure rests on the premise that few signals are hidden in a lot of noise. For example, let $X_i = \theta_i + \epsilon_i$, $i = 1, \ldots, n$, where $\epsilon_i \sim$ i.i.d. $N(0,1)$, and consider testing $H_0 : \theta_i = 0,\ \forall i$. If all $\theta_i \neq 0$ (so high signal-to-noise ratio), then Wald's chi-square test is preferable to any threshold procedure including order thresholding; on the other hand, if $\theta_i \neq 0$ only for a few cases (low signal-to-noise ratio), then thresholding methods can achieve far superior power than the chi-square test. Especially, order thresholding achieves the highest power in low signal-to-noise ratio alternatives. To apply thresholding methods usefully also in high signal-to-noise ratio alternatives, we should apply either the discrete Fourier transformation (DFT) or the discrete wavelet transformation (DWT). The effect of such transformations is explained in the following figures (Figures A.4 and A.5). The top panel shows different alternatives to the null hypothesis that all means are zero, while the bottom panel shows the same alternatives after a DFT (in Figure A.4) or DWT (in Figure A.5). Note that the three panels on the left are examples of high signal-to-noise ratio and the right panel is an example of low signal-to-noise ratio. The effect of these transformations is to concentrate most of signal in a few locations. In the three high signal-to-noise ratio alternatives, the DFT does a better job in concentrating the signal in a few locations than the DWT, while the opposite is true in the low signal-to-noise ratio alternative. Since either of these transformations takes the form of multiplying the data vector by an orthonormal matrix $\mathbf{\Gamma}$, it is seen that

$$\mathbf{X}_{\mathbf{\Gamma}} \equiv \mathbf{\Gamma}\mathbf{X} = \mathbf{\Gamma}\boldsymbol{\theta} + \mathbf{\Gamma}\boldsymbol{\epsilon} \ \sim \ N\left(\mathbf{\Gamma}\boldsymbol{\theta}, \mathbf{I}_n\right).$$

Thus, under the null hypothesis $H_0 : \theta_i = 0,\ \forall i$, the transformed data continue to be i.i.d. $N(0,1)$, while under the alternative $||\boldsymbol{\theta}|| = ||\mathbf{\Gamma}\boldsymbol{\theta}||$. Thus, Wald's chi-square test will have the same power when applied to the transformed data, since the noncentrality parameter, $\sum_{i=1}^{n} \theta_i^2$, remains the same after the transformation. On the other hand, such transformations should improve the power of threshold methods in high signal-to-noise ratio situations.

The empirical powers, based on the chi-square and order threshold statistics before and after applying such transformations, are reported in Tables A.12 and A.13. The first four alternatives considered in these tables correspond to those in Figures A.4 and A.5, while the fifth is of the form $H_a^{(5)}: \theta_i = 1.8,\ i = n-10, \ldots, n$, 0, otherwise. The simulations in Table A.12, which refer to the DFT, use samples of size $n = 255$, while those in Table A.13, which refer to the DWT, use samples of size $n = 256$. The order threshold statistic uses a data-driven value of the threshold parameter which is suggested in Section 3.3. The results are based on 2,000 simulation runs. Simulations confirm the expectations that a) compared to the order threshold statistic on the untransformed data, the chi-square statistic achieved higher power under high signal-to-noise ratio alternatives, and lower power under low signal-to-noise ratio alternatives, and b) the chi-square statistic has the same power after taking either of the transformations. As shown in the two tables, both transformations improved the power of the order threshold statistic in the three high signal-to-noise ratio alternatives ($H_a^{(1)}$, $H_a^{(2)}$, $H_a^{(3)}$), with the DFT achieving the larger power gains. On the other hand, the DWT achieved larger power gains in both low signal-to-noise ratio alternatives ($H_a^{(4)}$, $H_a^{(5)}$), with the DFT resulting in loss of power under alternative $H_a^{(5)}$ (though it still achieved a somewhat higher power than the chi-square test). Similar results (not shown here) are obtained when the number of $\boldsymbol{\theta}$ with values different from zero is smaller than 10.

5.2 Transform-Based Order Threshold Statistics for Lack of Fit

Let $U_1, \ldots, U_n$ be i.i.d. random variables having distribution function G with support in $(0, 1)$. To test $H_0 : G(t) = t,\ \forall 0 < t < 1$, we will partition the interval $(0, 1]$ into the n_{bin} subintervals $J_i = ((i-1)/n_{bin}, i/n_{bin}]$, $i = 1, \ldots, n_{bin}$, and define

$$N_j = n \int_{J_j} \mathrm{d}\widehat{G}_n(x), \quad M_j = n \int_{J_j} \psi(x) \mathrm{d}\widehat{G}_n(x), \quad j = 1, \ldots, n_{bin},$$

where $\widehat{G}_n(x) = \dfrac{1}{n} \sum_{i=1}^{n} I(U_i \leq x)$ is the empirical distribution function of $U_1, \ldots, U_n$, and ψ is a known score function. Under H_0, the marginal distribution of each N_j is Binomial with parameters n and $1/n_{bin}$, while the mean and variance of the marginal distribution of M_j are

$$\mu_j = nE_{H_0}\left(I(U \in J_j)\psi(U)\right), \quad \sigma_j^2 = n\left[E_{H_0}\left(I(U \in J_j)\psi(U)^2\right) - (\mu_j/n)^2\right].$$

Assuming that $n_{bin} \to \infty$ and $n/n_{bin} \to \infty$ as $n \to \infty$, we have

$$S_j = \frac{N_j - n/n_{bin}}{\sqrt{n(n_{bin}-1)/n_{bin}^2}} \ \dot{\sim} \ N(0,1), \quad \text{and} \quad \widetilde{S}_j = \frac{M_j - \mu_j}{\sigma_j} \ \dot{\sim} \ N(0,1), \tag{5.1}$$

for $j = 1, \ldots, n_{bin}$, under H_0, and components of the vectors $\mathbf{S} = (S_1, \ldots, S_{n_{bin}})^T$ and $\widetilde{\mathbf{S}} = (\widetilde{S}_1, \ldots, \widetilde{S}_{n_{bin}})^T$, are approximately independent. Let

$$\mathbf{X}_{\mathbf{\Gamma}} = (X_{\mathbf{\Gamma},1}, \ldots, X_{\mathbf{\Gamma},n_{bin}})^T, \quad \widetilde{\mathbf{X}}_{\mathbf{\Gamma}} = (\widetilde{X}_{\mathbf{\Gamma},1}, \ldots, \widetilde{X}_{\mathbf{\Gamma},n_{bin}})^T \tag{5.2}$$

be the discrete Fourier or wavelet transforms of $\mathbf{S}$ and $\widetilde{\mathbf{S}}$, respectively. Since either of these transforms is orthonormal, the vectors $\mathbf{X}_{\mathbf{\Gamma}}$ and $\widetilde{\mathbf{X}}_{\mathbf{\Gamma}}$ are approximately $N(\mathbf{0}, \mathbf{I}_{n_{bin}})$ distributed, under H_0. We may thus consider the order threshold statistics

$$T(k_{n_{bin}}) = \sum_{j=n_{bin}-k_{n_{bin}}+1}^{n_{bin}} X_{\mathbf{\Gamma},(j)}^2, \quad \text{and} \quad \widetilde{T}(k_{n_{bin}}) = \sum_{j=n_{bin}-k_{n_{bin}}+1}^{n_{bin}} \widetilde{X}_{\mathbf{\Gamma},(j)}^2, \tag{5.3}$$

where $X_{\mathbf{\Gamma},(1)}^2 < \cdots < X_{\mathbf{\Gamma},(n_{bin})}^2$ are ordered $X_{\mathbf{\Gamma},j}^2$'s, $\widetilde{X}_{\mathbf{\Gamma},(1)}^2 < \cdots < \widetilde{X}_{\mathbf{\Gamma},(n_{bin})}^2$ are ordered $\widetilde{X}_{\mathbf{\Gamma},j}^2$'s, and $k_{n_{bin}}$ is the order threshold parameter. In the next section, the asymptotic normality of these order threshold statistics is derived under some conditions.

Remark 5.1. *Fan (1996) considered a goodness-of-fit criteria based on the Fourier or the wavelet transformation for the vector* $\mathbf{S}^*$ *with components*

$$S_j^* = 2(\sqrt{N_j} - \sqrt{n/n_{bin}}) \ \dot{\sim} \ N(0,1), \quad j = 1, \ldots, n_{bin}.$$

In our simulations, use of the vector $\mathbf{S}^*$ *resulted in more liberal procedure than that based on* $\mathbf{S}$*. For this reason, the order threshold statistic using* $\mathbf{S}^*$ *is not considered in this dissertation.*

5.3 Asymptotic Theory

The derivation of the asymptotic null distribution of the order threshold statistics given in (5.3) is based on the following stochastic representations of S_j and $\widetilde{S}_j$. In all formulas that follow, we used $\psi(x) = x$ in the definition of $\widetilde{S}_j$.

Proposition 5.1. *Let S_j, $\widetilde{S}_j$, be defined in (5.1), and $X_{\mathbf{\Gamma},j}$, $\widetilde{X}_{\mathbf{\Gamma},j}$ be defined in (5.2). Then, there exists a version of the Brownian motion W_n, such that, under H_0:*

(a) The random variables S_j can be represented as

$$S_j = \epsilon_j + V_j, \quad j = 1, \dots, n_{bin},$$

where $\epsilon_j \sim$ i.i.d. $N(0,1)$ and $V_j = -n_{bin}^{-1/2} W_n(1) + O_r\left\{(n^{-1} n_{bin} \log^2 n)^{1/2}\right\}$.

(b) The random variables $X_{\mathbf{\Gamma},j}$ can be represented as

$$X_{\mathbf{\Gamma},j} = X_j + T_j, \quad j = 1, \dots, n_{bin}, \tag{5.4}$$

where X_j is the jth component of $\mathbf{\Gamma}(\epsilon_1, \dots, \epsilon_{n_{bin}})^T$, and the T_j satisfy

$$\max_j |T_j| = O_r\left\{(n^{-1} n_{bin}^2 \log^2 n)^{1/2}\right\}.$$

(c) The random variables $\widetilde{S}_j$ can be represented as

$$\widetilde{S}_j = \epsilon_j + \widetilde{V}_j, \quad j = 1, \dots, n_{bin},$$

where $\epsilon_j \sim$ i.i.d. $N(0,1)$ and

$$\begin{aligned} \widetilde{V}_j &= -\left(\frac{12j^2 - 12j + 3}{12j^2 - 12j + 4}\right)^{1/2} n_{bin}^{-1/2} W_n(1) \\ &\quad + O_r\left\{\left(\frac{n(3j^2 - 3j + 1)}{3n_{bin}^3}\right)^{-1/2} \log n\right\}. \end{aligned}$$

(d) The random variables $\widetilde{X}_{\mathbf{\Gamma},j}$ *can be represented as*

$$\widetilde{X}_{\mathbf{\Gamma},j} = X_j + \widetilde{T}_j, \quad j = 1, \ldots, n_{bin}, \tag{5.5}$$

where the X_j *are defined in part (b) and the* $\widetilde{T}_j$ *satisfy*

$$\max_j |\widetilde{T}_j| = O_r\left\{\left(n^{-1}n_{bin}^4 \log^2 n\right)^{1/2}\right\}.$$

When we square both sides of (5.4) and (5.5), we have

$$X^2_{\mathbf{\Gamma},j} = X_j^2 + W_j, \quad \widetilde{X}^2_{\mathbf{\Gamma},j} = X_j^2 + \widetilde{W}_j, \quad j = 1, \ldots, n_{bin}, \tag{5.6}$$

where $W_j = T_j^2 + 2X_jT_j$, and $\widetilde{W}_j = \widetilde{T}_j^2 + 2X_j\widetilde{T}_j$. Note that the first of the order threshold statistics in (5.3), $T(k_{n_{bin}})$, is based on the ordered $X^2_{\mathbf{\Gamma},j}$'s and the second, $\widetilde{T}(k_{n_{bin}})$, is based on the ordered $\widetilde{X}^2_{\mathbf{\Gamma},j}$'s. On the other hand, the asymptotic theory of order threshold statistics given in Section 3.2 pertains to the ordered X_j^2's. Theorems 5.2, 5.3 below assert that each of the two order threshold statistics in (5.3) has the same asymptotic null distribution as the order threshold statistic based on the ordered X_j^2's. To set the notation, it is convenient to restate the relevant result from the aforementioned section (Section 3.2).

Theorem 5.1. *Let* F, f *be the cumulative distribution function, density function, respectively, of the central chi-square distribution with one degree of freedom, and set* $\tilde{H}(v) = F^{-1}(1 - e^{-v})$, $v \geq 0$. *Let* $Y_1, \ldots, Y_{n_{bin}}$ *be i.i.d.[F], and let* $k_{n_{bin}}$, $n_{bin} \geq 1$, *be any sequence of integers which satisfies* $k_{n_{bin}} \to \infty$, *as* $n_{bin} \to \infty$, *and* $k_{n_{bin}} \leq n_{bin}$. *Let*

$$\mu_{n_{bin}}(k_{n_{bin}}) = \sum_{j=n_{bin}-k_{n_{bin}}+1}^{n_{bin}} \frac{\tilde{H}\left(\tilde{\nu}_{jn_{bin}}\right)}{n_{bin}}, \quad \sigma^2_{n_{bin}}(k_{n_{bin}}) = \sum_{j=1}^{n_{bin}} \frac{\alpha^2_{jn_{bin}}(k_{n_{bin}})}{n_{bin}}, \tag{5.7}$$

where

$$\alpha_{jn_{bin}}(k_{n_{bin}}) = \begin{cases} \dfrac{1}{n_{bin}-j+1} \displaystyle\sum_{i=n_{bin}-k_{n_{bin}}+1}^{n_{bin}} \tilde{H}'\left(\tilde{\nu}_{in_{bin}}\right), & j \le n_{bin}-k_{n_{bin}}, \\ \dfrac{1}{n_{bin}-j+1} \displaystyle\sum_{i=j}^{n_{bin}} \tilde{H}'\left(\tilde{\nu}_{in_{bin}}\right), & j > n_{bin}-k_{n_{bin}}, \end{cases}$$

with $\tilde{H}'\left(\tilde{\nu}_{in_{bin}}\right) = \dfrac{e^{-\tilde{\nu}_{in_{bin}}}}{f\left(F^{-1}\left(1-e^{-\tilde{\nu}_{in_{bin}}}\right)\right)}$ *and* $\tilde{\nu}_{in_{bin}} = \displaystyle\sum_{j=1}^{i} \frac{1}{n_{bin}-j+1}$. *Then we have*

$$\frac{T_O(k_{n_{bin}}) - n_{bin}\mu_{n_{bin}}(k_{n_{bin}})}{\sqrt{n_{bin}}\sigma_{n_{bin}}(k_{n_{bin}})} \xrightarrow{d} N(0,1), \quad as \;\; n_{bin} \to \infty, \tag{5.8}$$

where $T_O(k_{n_{bin}}) = \sum_{j=n_{bin}-k_{n_{bin}}+1}^{n_{bin}} Y_{j,n_{bin}}$ *with* $Y_{1,n_{bin}} < \cdots < Y_{n_{bin},n_{bin}}$ *are ordered* Y_j*'s.*

The main results are stated next.

Theorem 5.2. *Let* $k_{n_{bin}}$*,* $n_{bin} \ge 1$*, be any sequence of integers which satisfies* $k_{n_{bin}} \to \infty$*, as* $n_{bin} \to \infty$*, and* $k_{n_{bin}} \le n_{bin}$*. Let* $\mu_{n_{bin}}(k_{n_{bin}})$ *and* $\sigma^2_{n_{bin}}(k_{n_{bin}})$ *be as in (5.7), and let* $T(k_{n_{bin}})$ *be given in (5.3). Provided that* $n^{-1/2} n_{bin}^{3/2+\delta} \log^{3/2} n \to 0$*, some* $\delta > 0$*, and* $k_{n_{bin}} \to \infty$*, we have*

$$T^*(k_{n_{bin}}) = \frac{T(k_{n_{bin}}) - n_{bin}\mu_{n_{bin}}(k_{n_{bin}})}{\sqrt{n_{bin}}\sigma_{n_{bin}}(k_{n_{bin}})} \xrightarrow{d} N(0,1), \quad as \;\; n_{bin} \to \infty,$$

under H_0*.*

Theorem 5.3. *Let* $k_{n_{bin}}$*,* $n_{bin} \ge 1$*, be any sequence of integers which satisfies* $k_{n_{bin}} \to \infty$*, as* $n_{bin} \to \infty$*, and* $k_{n_{bin}} \le n_{bin}$*. Let* $\mu_{n_{bin}}(k_{n_{bin}})$ *and* $\sigma^2_{n_{bin}}(k_{n_{bin}})$ *be as in (5.7), and let* $\widetilde{T}(k_{n_{bin}})$ *be given in (5.3). Provided that* $n^{-1/2} n_{bin}^{5/2+\delta} \log^{3/2} n \to 0$*, some* $\delta > 0$*, and* $k_{n_{bin}} \to \infty$*, we have*

$$\widetilde{T}^*(k_{n_{bin}}) = \frac{\widetilde{T}(k_{n_{bin}}) - n_{bin}\mu_{n_{bin}}(k_{n_{bin}})}{\sqrt{n_{bin}}\sigma_{n_{bin}}(k_{n_{bin}})} \xrightarrow{d} N(0,1), \quad as \;\; n_{bin} \to \infty,$$

under H_0*.*

5.4 Proofs of the Main Results

5.4.1 Proof of Theorem 5.2

Let $X^2_{\mathbf{\Gamma},j}$ be defined in (5.6), and set $Y_j = X_j^2$, $j = 1, \ldots, n_{bin}$. Let $Y_{(j)}$ and $W_{(j)}$ correspond to the jth order statistic of $X^2_{\mathbf{\Gamma},j}$, $X^2_{\mathbf{\Gamma},(j)}$. Thus, $X^2_{\mathbf{\Gamma},(j)} = Y_{(j)} + W_{(j)}$, $j = 1, \ldots, n_{bin}$. Then we have

$$\begin{aligned} T^*(k_{n_{bin}}) &= \frac{\sum_{j=n_{bin}-k_{n_{bin}}+1}^{n_{bin}} X^2_{\mathbf{\Gamma},(j)} - n_{bin}\mu_{n_{bin}}(k_{n_{bin}})}{\sqrt{n_{bin}}\sigma_{n_{bin}}(k_{n_{bin}})} \\ &= \frac{\sum_{j=n_{bin}-k_{n_{bin}}+1}^{n_{bin}} Y_{(j)} - n_{bin}\mu_{n_{bin}}(k_{n_{bin}})}{\sqrt{n_{bin}}\sigma_{n_{bin}}(k_{n_{bin}})} + \frac{\sum_{j=n_{bin}-k_{n_{bin}}+1}^{n_{bin}} W_{(j)}}{\sqrt{n_{bin}}\sigma_{n_{bin}}(k_{n_{bin}})}. \end{aligned} \quad (5.9)$$

The theorem will be shown by showing that the second term on the right hand side converges in probability to zero, while the first term converges to the standard normal distribution. The former is shown in Lemma 5.1 below. The proof of the later uses essentially Theorem 5.1 after accounting for the fact that, in this case, $Y_{(j)}$ are not really the ordered statistics of the i.i.d. chi-square random variables $Y_1, \ldots, Y_{n_{bin}}$. (Recall that the order statistics are denoted by $Y_{1,n_{bin}}, \ldots, Y_{n_{bin},n_{bin}}$.) The basic idea for accounting for the lack of ordering of the $Y_{(j)}$'s is to create a "buffer zone" which separates "most" of the largest $k_{n_{bin}}$ order statistics that we want to include in the order threshold statistic (i.e. $Y_{n_{bin}-k_{n_{bin}}+1,n_{bin}}, \ldots, Y_{n_{bin},n_{bin}}$) from "most" of the smallest ones that we want to exclude. To make this idea precise, suppose we can find a sequence of integers $a_{n_{bin}}$ which satisfies

1. $k_{n_{bin}} < a_{n_{bin}} \leq n_{bin}$.

2. The convergence in (5.8) remains true if $T_O(k_{n_{bin}})$ is replaced by $T_O(a_{n_{bin}})$. That is,

$$\frac{\sum_{j=n_{bin}-a_{n_{bin}}+1}^{n_{bin}} Y_{j,n_{bin}} - n_{bin}\mu_{n_{bin}}(k_{n_{bin}})}{\sqrt{n_{bin}}\sigma_{n_{bin}}(k_{n_{bin}})} \xrightarrow{d} N(0,1), \text{ as } n_{bin} \to \infty. \quad (5.10)$$

3. The relation

$$\left\{Y_{(n_{bin}-k_{n_{bin}}+1)}, \ldots, Y_{(n_{bin})}\right\} \subset \left\{Y_{n_{bin}-a_{n_{bin}}+1,n_{bin}}, \ldots, Y_{n_{bin},n_{bin}}\right\}, \quad (5.11)$$

holds with probability tending to 1.

Then, using (5.8), it is easily seen that the first term of (5.9) converges in distribution to the standard normal. In Lemma B.5 of Appendix B, it is shown that if the sequence $a_{n_{bin}}$ satisfies

$$\frac{a_{n_{bin}} - k_{n_{bin}}}{\sqrt{k_{n_{bin}}}} \log\left(\frac{n_{bin}+1}{k_{n_{bin}}+1}\right) \to 0, \quad \text{as} \quad n_{bin} \to \infty, \tag{5.12}$$

then relations (5.10) and (5.11) hold. This completes the proof of the theorem.

Remark 5.2. *It is easily seen that a sequence of integers $a_{n_{bin}}$ of the form*

$$k_{n_{bin}} + o\left(\frac{\sqrt{k_{n_{bin}}}}{\log((n_{bin}+1)/(k_{n_{bin}}+1))}\right), \quad \textit{or} \quad k_{n_{bin}} + k_{n_{bin}}^{1/2-\delta} \textit{ for some } 0 < \delta < 1/2,$$

satisfies condition (5.12).

Lemma 5.1. *Let $X^2_{\mathbf{\Gamma},j}$ and W_j be given in (5.6), and let $W_{(j)}$ correspond to the jth order statistic of $X^2_{\mathbf{\Gamma},j}$. Then,*

$$\frac{1}{\sqrt{n_{bin}}\sigma_{n_{bin}}(k_{n_{bin}})} \sum_{j=n_{bin}-k_{n_{bin}}+1}^{n_{bin}} W_{(j)} \xrightarrow{p} 0, \quad \textit{as} \quad n_{bin} \to \infty, \tag{5.13}$$

provided that $n^{-1/2} n_{bin}^{3/2} \log^{3/2} n \to 0$.

Proof: Set $\gamma_{n_{bin}} = \max_j |W_j|$. Then, for some $C > 0$,

$$\frac{1}{\sqrt{n_{bin}}\sigma_{n_{bin}}(k_{n_{bin}})} \sum_{j=n_{bin}-k_{n_{bin}}+1}^{n_{bin}} W_{(j)} \le \frac{k_{n_{bin}}\gamma_{n_{bin}}}{\sqrt{n_{bin}}\sigma_{n_{bin}}(k_{n_{bin}})} \le C k_{n_{bin}}^{1/2} \gamma_{n_{bin}},$$

where the last inequality follows from the proof of Lemma B.5 of the appendix. Thus, relation (5.13) will follow from

$$k_{n_{bin}}^{1/2} \gamma_{n_{bin}} \xrightarrow{p} 0, \quad \text{as} \quad n_{bin} \to \infty. \tag{5.14}$$

However, we have

$$\begin{aligned}\gamma_{n_{bin}} &= \max_j |T_j^2 + 2X_jT_j| \leq \max_j |T_j|^2 + 2\max_j |X_jT_j| \\ &= O_r\left(n^{-1/2} n_{bin} \log n \log^{1/2} n_{bin}\right),\end{aligned}$$

where the last equality is justified by Proposition 5.1 (b) and the fact that $\max_j X_j = O_r(\log^{1/2} n_{bin})$ (cf. Fan (1996)). This, together with the condition

$$n^{-1/2} n_{bin}^{3/2} \log^{3/2} n \to 0,$$

imply (5.14) which completes the proof of the lemma.

Remark 5.3. *Under the conditions of Theorem 5.2, the relation (5.14) can be strengthened to*

$$k_{n_{bin}}^{1/2+\delta} \gamma_{n_{bin}} \xrightarrow{p} 0, \quad \textit{for some } \delta > 0, \textit{ as } n_{bin} \to \infty.$$

We will use this stronger version in the proof of Lemma B.7.

5.4.2 Proof of Theorem 5.3

Let $\widetilde{X}^2_{\mathbf{\Gamma},j}$ be defined in (5.6), and set $Y_j = X_j^2$, $j = 1, \ldots, n_{bin}$. Let $Y_{(j)}$ and $\widetilde{W}_{(j)}$ correspond to the jth order statistic of $\widetilde{X}^2_{\mathbf{\Gamma},j}$, $\widetilde{X}^2_{\mathbf{\Gamma},(j)}$. Thus, $\widetilde{X}^2_{\mathbf{\Gamma},(j)} = Y_{(j)} + \widetilde{W}_{(j)}$, $j = 1, \ldots, n_{bin}$. Then we have

$$\begin{aligned}&\widetilde{T}^*(k_{n_{bin}}) \\ &= \frac{\sum_{j=n_{bin}-k_{n_{bin}}+1}^{n_{bin}} Y_{(j)} - n_{bin}\mu_{n_{bin}}(k_{n_{bin}})}{\sqrt{n_{bin}}\sigma_{n_{bin}}(k_{n_{bin}})} + \frac{\sum_{j=n_{bin}-k_{n_{bin}}+1}^{n_{bin}} \widetilde{W}_{(j)}}{\sqrt{n_{bin}}\sigma_{n_{bin}}(k_{n_{bin}})}. \end{aligned} \tag{5.15}$$

From Lemma B.8 of Appendix B, it is seen that the first term on the right hand side of (5.15) converges to the standard normal distribution. Moreover, in Lemma 5.2, it is shown that

$$\frac{1}{\sqrt{n_{bin}}\sigma_{n_{bin}}(k_{n_{bin}})} \sum_{j=n_{bin}-k_{n_{bin}}+1}^{n_{bin}} \widetilde{W}_{(j)} \xrightarrow{p} 0, \quad \text{as } n_{bin} \to \infty. \tag{5.16}$$

Thus the proof of the theorem is complete.

Lemma 5.2. *Let* $\widetilde{X}^2_{\mathbf{\Gamma},j}$ *and* $\widetilde{W}_j$ *be given in (5.6), and let* $\widetilde{W}_{(j)}$ *correspond to the* j*th order statistic of* $\widetilde{X}^2_{\mathbf{\Gamma},j}$*. Then, the convergence in (5.16) holds, provided that* $n^{-1/2}n_{bin}^{5/2}\log^{3/2} n \to 0$.

Proof: Set $\widetilde{\gamma}_{n_{bin}} = \max_j |\widetilde{W}_j|$. Then, by the same argument of proof of Lemma 5.1, the relation (5.16) will follow from

$$k_{n_{bin}}^{1/2}\widetilde{\gamma}_{n_{bin}} \xrightarrow{p} 0, \quad \text{as} \quad n_{bin} \to \infty. \tag{5.17}$$

By the fact that $\widetilde{\gamma}_{n_{bin}} = O_r\left(n^{-1/2}n_{bin}^2 \log n \log^{1/2} n_{bin}\right)$, which is justified by Proposition 5.1 (d) and the fact that $\max_j X_j = O_r(\log^{1/2} n_{bin})$ (cf. Fan (1996)), the convergence in (5.16) is true under the condition $n^{-1/2}n_{bin}^{5/2}\log^{3/2} n \to 0$.

Remark 5.4. *Under the conditions of Theorem 5.3, the relation (5.17) can be strengthened to*

$$k_{n_{bin}}^{1/2+\delta}\widetilde{\gamma}_{n_{bin}} \xrightarrow{p} 0, \quad \textit{for some } \delta > 0, \textit{ as } \quad n_{bin} \to \infty.$$

We will use this stronger version in the proof of Lemma B.9.

5.5 Simulations

All simulation results reported in this section are based on 10,000 simulation runs. Table A.14 shows the achieved alpha levels, at nominal α level 0.05 and sample size $n = 200$, for the Pearson's χ^2 (based on 31 bins), Kolmogorov-Smirnov (KS), and Cramér-Von Mises (CVM) test statistics. Table A.15 shows the achieved alpha levels, at nominal α level 0.05 and sample size $n = 200$, for the hard threshold (T_H), adaptive Neyman's, and the two proposed order threshold statistics after applying a DFT (using 31 bins) and a DWT (using 32 bins). From these tables, it is seen that all type I error rates are quite close to the nominal level except for those of hard thresholding and adaptive Neyman's test, which appear to be quite liberal. The threshold parameter for the hard threshold statistic was computed according to the formula $\delta_{n_{bin}} = 2\log(n_{bin}\log^{-2} n_{bin})$, which is suggested in Fan (1996); thus,

$\delta_{n_{bin}} = 1.9331$ or 1.9598 for 31 and 32 bins, respectively. The threshold parameter for the order threshold statistic was $k_{n_{bin}} = 6$ in both cases. Table A.15 clearly shows the proposed order thresholding procedures are more conservative than hard thresholding and adaptive Neyman's procedures. Use of the data-driven threshold parameter $\widehat{k}^{opt}_{n_{bin}}$ for the order threshold statistics (see Subsection 3.3), resulted in type I error rates below 0.06. Due to the additional computational time required for using $\widehat{k}^{opt}_{n_{bin}}$, we used the order threshold statistics with threshold parameter fixed at $k_{n_{bin}} = 6$. It should be pointed out, however, that the order threshold statistics with the data-driven threshold parameter achieve at least as good power as the order threshold statistics with fixed threshold parameter.

Figure A.6 shows the estimated and theoretical (asymptotic) densities of T_H, adaptive Neyman's, and the two order threshold statistics, after the DFT (top panel) and after the DWT (bottom panel). The red dash-dot line represents the theoretical (asymptotic) densities and the blue solid line represents the estimated densities. As expected from the reported type I error rates, the distributions of the two order threshold statistics appear to converge to their limiting distribution faster than those of T_H and adaptive Neyman's statistic. In Figure A.6, the theoretical (asymptotic) densities of the two order threshold statistics that we use are the standard normal. If we use a multiple of a χ^2 distribution, the approximation will perform better than the normal approximation, but the difference is not big.

Remark 5.5. *We also considered transforming the data to follow the standard normal distribution under the null hypothesis and applying the order threshold statistic (3.8) without binning the transformed data. While this statistic is quite powerful for detecting location alternatives, its power was relatively low against the present distributional alternatives, and thus it was not included in the reported simulation results.*

Example 5.1. *Consider the testing problem* $H_0 : G = Uniform(0,1)$ *versus* $H_a : G = Beta(c,c)$, *with c =1.8, 1.6, 1.5, 1.4, 1.3, 1.2, 0.8, and 0.7.*

Example 5.2. *This example is similar to Example 5 of Fan (1996). We want to* *test* $H_0 : G = Uniform(0,1)$ *versus* $H_a : G = G_\mu$, *where*

$$G_\mu(y) = \begin{cases} y, & 0 \le y < \frac{1-\mu}{2}, \\ \frac{1}{\mu^2}\left\{\frac{16y^3}{3} + (4\mu - 8)y^2 + (\mu^2 - 4\mu + 4)y + \frac{-2\mu^3+6\mu-4}{6}\right\}, & \frac{1-\mu}{2} \le y < \frac{1}{2}, \\ \frac{1}{\mu^2}\left\{-\frac{16y^3}{3} + (4\mu + 8)y^2 + (\mu^2 - 4\mu - 4)y + \frac{-2\mu^3+6\mu+4}{6}\right\}, & \frac{1}{2} \le y < \frac{1+\mu}{2}, \\ y, & \frac{1+\mu}{2} \le y \le 1, \end{cases}$$

with $\mu =$ *0.4, 0.367, 0.333, 0.3, 0.267, 0.233, 0.2, and 0.1.*

Example 5.3. *Consider the testing problem*

$$H_0 : G = N(0,1) \text{ versus } H_a : G = 0.8N\left(0, \frac{1}{0.8 + 0.2\sigma^2}\right) + 0.2N\left(0, \frac{\sigma^2}{0.8 + 0.2\sigma^2}\right),$$

with $\sigma =$ *0.125, 0.15, 0.175, 0.2, 0.25, 0.3, and 0.4.*

Example 5.4. *Consider the testing problem*

$$H_0 : G = N(0,1) \text{ versus } H_a : G = 0.7N(\mu/0.7, 1) + 0.3N(-\mu/0.3, 1),$$

with $\mu =$ *0.4, 0.367, 0.333, 0.3, 0.267, 0.233, and 0.2.*

Example 5.5. *Consider the testing problem* $H_0 : G = Uniform(0,1)$ *versus* $H_a : G = G_{a,b}$, *where*

$$G_{a,b}(y) = \begin{cases} y, & 0 \le y < b, \\ b(1-a) + ay, & b \le y < (b+1)/2, \\ -1 + a + (2-a)y, & (b+1)/2 \le y \le 1, \end{cases}$$

with a =0.1, 0.2, 0.3, 0.4, 0.5, 0.6, and b =0.75, 0.85.

Example 5.6. *Consider the testing problem* $H_0 : G = Uniform(0,1)$ *versus* $H_a : G = Beta(1.1, c)$, *with c =1.8, 1.7, 1.6, 1.5, 1.4, 1.3, 1.2, 1.1, 1.0, 0.9, and 0.8.*

The power of the test procedures was considered against alternatives described in Examples 5.1-5.6. The hard threshold statistic achieved mainly higher power than either of the two order threshold statistics (after either the DFT or the DWT) in all examples except for Examples 5.3 and 5.4. This higher power, however, has to be interpreted in the context of the very liberal type I error rates of the hard threshold statistic. Excluding the hard threshold statistic, the results indicate that, in Examples 5.1-5.3, the proposed order threshold procedures based on the cell counts and on the cell averages ($T^*(6)$ and $\widetilde{T}^*(6)$), with the DF transform, performed overall best and second best, respectively. In Example 5.4 the best and second best procedures were the order threshold procedures based on the cell counts with the DF and DW transforms, respectively. In Examples 5.5 and 5.6 the two order threshold procedures with the DW transform were the best performers. Overall, the Mann and Wald (1942) version of Pearson's χ^2 had good power; in particular, it outperformed both the KS and the CVM statistics (often by considerable amounts) in all examples except Example 5.6.; it even outperformed $\widetilde{T}^*(6)$ with the DW transform in Example 5.4. The power of the order threshold statistics with the DW transform is almost always lower than that of the corresponding order threshold statistics with the DF transform, and approximately tied Pearson's χ^2. However, Examples 5.5 and 5.6 have the reverse results.

Chapter 6

Discussion

The asymptotic theory of test statistics based on hard and soft thresholding pertain the normal distribution and require the threshold parameter to tend to infinity at a strictly prescribed rate. This second feature results in potentially compromised power of the hard and soft threshold statistics.

Order thresholding, a new thresholding method based on order statistics, is proposed. The asymptotic theory, developed under the normal distribution in this dissertation, allows great flexibility in the choice of the threshold parameter. A data-driven choice of the order threshold parameter is given. Extensions to a one-way HANOVA setting and a goodness-of-fit testing problem are presented. Simulation studies with normal data suggest that order thresholding can have great power advantage over hard thresholding. Additional simulations with data generated under a one-way HANOVA design suggest even larger power gains over the traditional ANOVA F-test. Finally, simulations with uniform data suggest that order thresholding, coupled with the discrete Fourier or wavelet transformation, achieves accurate type I error rates and improves the power of the chi-square statistic. The Fourier transformation was most effective in improving the power in high signal-to-noise ratio alternatives, while the wavelet transformation was most effective in low-signal-to-noise ratio alternatives. Moreover, the proposed order threshold procedures are often considerably more powerful than the Kolmogorov-Smirnov and Cramér-Von Mises tests providing thus a useful alternative method for establishing statistical evidence against the overall null hypothesis in multiple testing problems.

Future work will consider testing the composite null hypothesis which is relevant for model assessment with applications in uncertainty quantification and Bayesian analysis (cf. Johnson, 2004). We will study the order threshold methodology applied to the chi-square statistic when the MLE is used to estimate the unknown parameters. Furthermore, applications of the order thresholding approach to testing in high-dimensional ANOVA without the normality assumption, and to multiple testing problems will also be pursed in future.

Appendix A

TABLES AND FIGURES

A.1 Tables of Chapter 2

Table A.1. Type I Errors of $T_H(\delta_n)$ for Different Values of the Hard Threshold Parameter

	$\delta_n - 2.0$	$\delta_n - 1.6$	$\delta_n - 1.2$	$\delta_n - 0.8$	$\delta_n - 0.4$	δ_n
$n = 50$	0.0003	0.0099	0.0231	0.0341	0.0431	0.0493
$n = 100$	0.0101	0.0229	0.0324	0.0390	0.0461	0.0504
$n = 200$	0.0231	0.0316	0.0382	0.0439	0.0484	0.0507
$n = 500$	0.0327	0.0388	0.0422	0.0465	0.0502	0.0535

Table A.2. Type I Errors of $T_H(\delta_n)$ for Different Values of the Hard Threshold Parameter

	$\delta_n + 0.4$	$\delta_n + 0.8$	$\delta_n + 1.2$	$\delta_n + 1.6$	$\delta_n + 2.0$
$n = 50$	0.0543	0.0588	0.0614	0.0654	0.0663
$n = 100$	0.0552	0.0559	0.0597	0.0616	0.0631
$n = 200$	0.0539	0.0562	0.0590	0.0601	0.0627
$n = 500$	0.0540	0.0563	0.0583	0.0604	0.0623

A.2 Tables and Figures of Chapter 3

Table A.3. Type I Errors of $T_L(k_n)$ for Different Values of the Order Threshold Parameter

	$[\log^{1/2} n]$	$[\log n]$	$[\log^{3/2} n]$	$[n^{1/2}]$	$[n^{2/3}]$	$[n^{3/4}]$	$[n^{7/8}]$	n
$n = 50$	0.0696	0.0685	0.0646	0.0646	0.0635	0.0630	0.0623	0.0626
$n = 100$	0.0669	0.0640	0.0606	0.0600	0.0591	0.0577	0.0585	0.0589
$n = 200$	0.0667	0.0620	0.0603	0.0589	0.0582	0.0583	0.0555	0.0560
$n = 500$	0.0665	0.0631	0.0577	0.0559	0.0536	0.0536	0.0535	0.0547

Table A.4. Type I Error Rates Using the Approximation $T_L(k_n) \sim b\chi^2_\nu$

	$[\log^{1/2} n]$	$[\log n]$	$[\log^{3/2} n]$	$[n^{1/2}]$	$[n^{2/3}]$	$[n^{3/4}]$	$[n^{7/8}]$	n
$n = 50$	0.0565	0.0545	0.0531	0.0531	0.0532	0.0534	0.0540	0.0546
$n = 100$	0.0566	0.0540	0.0520	0.0522	0.0519	0.0507	0.0518	0.0520
$n = 200$	0.0555	0.0547	0.0536	0.0531	0.0526	0.0523	0.0530	0.0516
$n = 500$	0.0589	0.0556	0.0552	0.0530	0.0520	0.0521	0.0508	0.0505

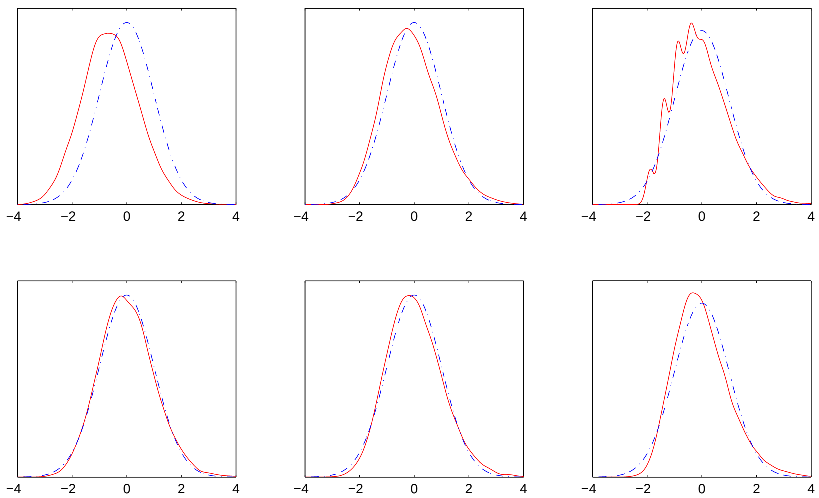

Figure A.1. Top panel: Estimated and theoretical densities of $T_H(\delta_{200})$ for $\delta_{200} = 1.842$, 3.927, and 5.672. Bottom panel: Estimated and theoretical densities of $T_L(k_{200})$ for $k_{200} = 35$, 10, and 3.

Table A.5. Power Calculations in Example 3.1

	k_{500}^{opt}	T_S	$T_H(5.12)$	$T_L(\widehat{k}_{500}^{opt})$	$b\chi^2_\nu$	T_L with 15	40	70	100	200	500
H_1	30	0.843	0.944	0.977	0.975	0.976	0.973	0.968	0.960	0.938	0.913
H_2	29	0.839	0.944	0.975	0.973	0.974	0.971	0.964	0.959	0.938	0.913
H_3	28	0.845	0.942	0.978	0.975	0.976	0.975	0.969	0.961	0.937	0.910
H_4	27	0.839	0.928	0.970	0.969	0.969	0.964	0.956	0.950	0.918	0.893
H_5	26	0.840	0.926	0.972	0.970	0.971	0.966	0.956	0.943	0.911	0.879
H_6	25	0.846	0.927	0.971	0.969	0.971	0.964	0.955	0.946	0.919	0.887
H_7	24	0.796	0.893	0.950	0.948	0.949	0.942	0.929	0.915	0.880	0.851
H_8	23	0.777	0.845	0.933	0.928	0.932	0.915	0.891	0.875	0.818	0.775
H_9	22	0.768	0.836	0.919	0.915	0.918	0.909	0.881	0.861	0.811	0.772
H_{10}	21	0.764	0.817	0.908	0.900	0.907	0.891	0.868	0.841	0.785	0.744
H_{11}	20	0.766	0.792	0.905	0.899	0.906	0.883	0.853	0.832	0.764	0.712
H_{12}	19	0.764	0.783	0.903	0.897	0.903	0.873	0.841	0.812	0.751	0.709
H_{13}	18	0.750	0.752	0.881	0.875	0.880	0.845	0.804	0.776	0.709	0.662
H_{14}	17	0.739	0.734	0.864	0.858	0.869	0.836	0.789	0.760	0.694	0.649
H_{15}	16	0.559	0.574	0.724	0.707	0.723	0.671	0.633	0.608	0.541	0.495
H_{16}	15	0.526	0.564	0.707	0.693	0.707	0.660	0.611	0.574	0.517	0.484
H_{17}	14	0.532	0.529	0.675	0.661	0.677	0.625	0.574	0.542	0.467	0.432
H_{18}	13	0.464	0.435	0.584	0.568	0.590	0.534	0.496	0.458	0.404	0.373
H_{19}	12	0.483	0.402	0.570	0.556	0.574	0.500	0.459	0.427	0.374	0.347
H_{20}	11	0.470	0.380	0.547	0.533	0.551	0.475	0.425	0.395	0.343	0.308
H_{21}	10	0.467	0.390	0.555	0.540	0.559	0.490	0.433	0.402	0.341	0.319
H_{22}	9	0.460	0.364	0.534	0.515	0.535	0.454	0.402	0.368	0.313	0.281
H_{23}	8	0.460	0.362	0.517	0.503	0.522	0.447	0.389	0.351	0.301	0.279
H_{24}	7	0.417	0.290	0.450	0.434	0.455	0.375	0.318	0.288	0.248	0.230
H_{25}	6	0.254	0.174	0.260	0.244	0.262	0.214	0.197	0.182	0.160	0.153
H_{26}	5	0.258	0.174	0.260	0.245	0.262	0.217	0.188	0.175	0.155	0.145
H_{27}	4	0.147	0.110	0.149	0.140	0.150	0.124	0.118	0.110	0.101	0.096
H_{28}	3	0.144	0.104	0.146	0.136	0.145	0.119	0.109	0.110	0.106	0.101
H_{29}	2	0.151	0.119	0.156	0.145	0.153	0.135	0.126	0.115	0.105	0.102
H_{30}	1	0.121	0.086	0.117	0.108	0.117	0.103	0.090	0.084	0.081	0.081
H_0^G	0	0.052	0.050	0.059	0.057	0.057	0.054	0.052	0.051	0.052	0.055

Table A.6. Power Calculations in Example 3.2

	k_{500}^{opt}	T_S	$T_H(5.12)$	$T_L(\widehat{k}_{500}^{opt})$	$b\chi^2_\nu$	T_L with 15	40	70	100	200	500
H_1	30	0.650	0.574	0.759	0.745	0.760	0.699	0.651	0.608	0.548	0.513
H_2	29	0.680	0.584	0.755	0.741	0.761	0.700	0.649	0.612	0.544	0.504
H_3	28	0.652	0.565	0.745	0.729	0.747	0.684	0.640	0.602	0.540	0.498
H_4	27	0.657	0.557	0.739	0.726	0.745	0.682	0.622	0.589	0.528	0.489
H_5	26	0.661	0.546	0.745	0.734	0.748	0.675	0.614	0.579	0.505	0.461
H_6	25	0.666	0.549	0.728	0.717	0.732	0.667	0.625	0.591	0.521	0.479
H_7	24	0.677	0.562	0.743	0.729	0.745	0.686	0.632	0.591	0.522	0.482
H_8	23	0.666	0.536	0.716	0.703	0.724	0.657	0.612	0.569	0.508	0.478
H_9	22	0.351	0.335	0.447	0.428	0.446	0.417	0.380	0.362	0.325	0.310
H_{10}	21	0.340	0.350	0.449	0.434	0.445	0.418	0.394	0.367	0.333	0.317
H_{11}	20	0.352	0.336	0.449	0.432	0.446	0.417	0.384	0.366	0.332	0.318
H_{12}	19	0.351	0.342	0.444	0.426	0.443	0.410	0.383	0.362	0.341	0.305
H_{13}	18	0.342	0.330	0.456	0.442	0.450	0.416	0.388	0.367	0.335	0.316
H_{14}	17	0.350	0.331	0.448	0.432	0.451	0.412	0.377	0.363	0.325	0.300
H_{15}	16	0.337	0.334	0.432	0.416	0.431	0.402	0.375	0.356	0.327	0.307
H_{16}	15	0.330	0.294	0.406	0.393	0.403	0.371	0.338	0.319	0.293	0.274
H_{17}	14	0.357	0.282	0.399	0.387	0.403	0.352	0.323	0.305	0.267	0.252
H_{18}	13	0.325	0.290	0.393	0.378	0.390	0.358	0.329	0.312	0.276	0.261
H_{19}	12	0.337	0.296	0.413	0.396	0.412	0.368	0.337	0.314	0.277	0.255
H_{20}	11	0.343	0.291	0.399	0.383	0.399	0.349	0.314	0.296	0.270	0.250
H_{21}	10	0.346	0.290	0.405	0.391	0.404	0.356	0.321	0.306	0.268	0.248
H_{22}	9	0.224	0.198	0.264	0.251	0.262	0.237	0.220	0.208	0.195	0.189
H_{23}	8	0.196	0.190	0.257	0.242	0.253	0.228	0.216	0.197	0.191	0.182
H_{24}	7	0.207	0.182	0.256	0.245	0.253	0.225	0.212	0.200	0.186	0.179
H_{25}	6	0.117	0.110	0.148	0.137	0.144	0.130	0.131	0.126	0.132	0.129
H_{26}	5	0.094	0.107	0.123	0.113	0.119	0.113	0.106	0.105	0.103	0.098
H_{27}	4	0.078	0.089	0.106	0.100	0.099	0.092	0.088	0.086	0.085	0.083
H_{28}	3	0.058	0.068	0.078	0.072	0.076	0.077	0.074	0.070	0.068	0.070
H_{29}	2	0.065	0.075	0.083	0.078	0.079	0.081	0.081	0.077	0.079	0.078
H_{30}	1	0.061	0.065	0.072	0.068	0.071	0.065	0.061	0.060	0.061	0.058

Table A.7. Power Calculations in Example 3.3

						T_L with					
	k_{500}^{opt}	T_S	$T_H(5.12)$	$T_L(\widehat{k}_{500}^{opt})$	$b\chi^2_\nu$	15	40	70	100	200	500
H_1	30	0.674	0.959	0.973	0.970	0.969	0.976	0.982	0.981	0.970	0.962
H_2	29	0.669	0.957	0.969	0.966	0.962	0.976	0.973	0.971	0.964	0.954
H_3	28	0.643	0.954	0.966	0.960	0.955	0.974	0.974	0.973	0.960	0.947
H_4	27	0.617	0.935	0.957	0.954	0.945	0.965	0.963	0.961	0.950	0.934
H_5	26	0.600	0.917	0.945	0.940	0.934	0.961	0.958	0.953	0.938	0.924
H_6	25	0.598	0.900	0.936	0.931	0.926	0.947	0.943	0.941	0.922	0.903
H_7	24	0.571	0.899	0.925	0.918	0.910	0.934	0.931	0.927	0.907	0.890
H_8	23	0.566	0.872	0.912	0.905	0.902	0.920	0.917	0.911	0.891	0.865
H_9	22	0.547	0.853	0.896	0.890	0.883	0.903	0.899	0.891	0.867	0.837
H_{10}	21	0.529	0.831	0.877	0.869	0.862	0.889	0.886	0.875	0.849	0.817
H_{11}	20	0.508	0.800	0.845	0.839	0.833	0.859	0.854	0.846	0.816	0.790
H_{12}	19	0.509	0.777	0.837	0.828	0.821	0.843	0.833	0.821	0.786	0.753
H_{13}	18	0.481	0.740	0.816	0.803	0.802	0.813	0.803	0.785	0.738	0.703
H_{14}	17	0.472	0.715	0.785	0.773	0.772	0.784	0.773	0.763	0.710	0.671
H_{15}	16	0.448	0.674	0.748	0.732	0.736	0.749	0.735	0.715	0.669	0.633
H_{16}	15	0.418	0.630	0.715	0.700	0.702	0.706	0.686	0.668	0.624	0.585
H_{17}	14	0.393	0.569	0.658	0.645	0.645	0.646	0.629	0.610	0.562	0.523
H_{18}	13	0.368	0.522	0.629	0.616	0.623	0.620	0.597	0.573	0.523	0.489
H_{19}	12	0.341	0.498	0.593	0.577	0.582	0.582	0.552	0.525	0.486	0.451
H_{20}	11	0.328	0.441	0.539	0.527	0.539	0.519	0.491	0.472	0.436	0.407
H_{21}	10	0.306	0.390	0.487	0.470	0.480	0.464	0.436	0.421	0.382	0.353
H_{22}	9	0.285	0.354	0.439	0.423	0.438	0.422	0.393	0.379	0.344	0.317
H_{23}	8	0.260	0.298	0.393	0.374	0.386	0.367	0.342	0.318	0.292	0.276
H_{24}	7	0.245	0.265	0.349	0.333	0.346	0.315	0.296	0.283	0.255	0.236
H_{25}	6	0.221	0.224	0.300	0.286	0.295	0.272	0.257	0.245	0.228	0.213
H_{26}	5	0.190	0.186	0.249	0.236	0.244	0.230	0.212	0.206	0.184	0.172
H_{27}	4	0.157	0.154	0.207	0.196	0.205	0.182	0.167	0.161	0.147	0.143
H_{28}	3	0.130	0.121	0.162	0.153	0.161	0.140	0.131	0.121	0.117	0.115
H_{29}	2	0.122	0.093	0.125	0.119	0.125	0.107	0.106	0.102	0.086	0.082
H_{30}	1	0.079	0.073	0.086	0.079	0.085	0.080	0.078	0.074	0.071	0.070

A.3 Tables and Figures of Chapter 4

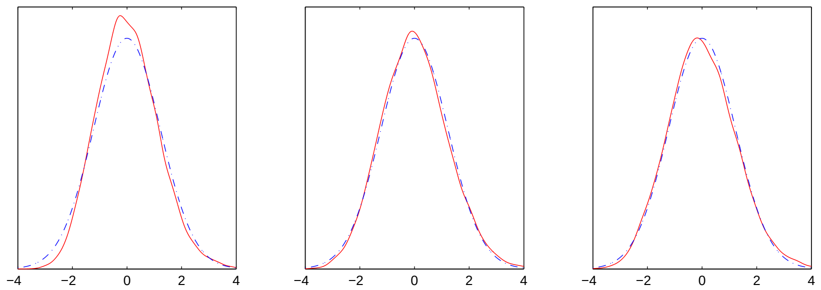

Figure A.2. Estimated and theoretical densities of $\widetilde{T}_L(k_{500})$ for $a = 500$, $n = 3$, and $k_{500} = 22$, 105, and 229.

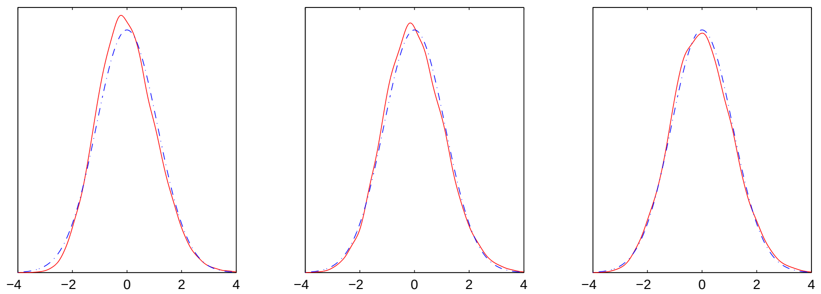

Figure A.3. Estimated and theoretical densities of $\widetilde{T}_L(k_{500})$ for $a = 500$, $n = 5$, and $k_{500} = 22$, 105, and 229.

Table A.8. Type I Errors of Order Threshold Statistics, $\widetilde{T}_L(k_a)$, for Different Values of the Threshold Parameter

	$[\log^{1/2} a]$	$[\log a]$	$[\log^{3/2} a]$	$[a^{1/2}]$	$[a^{2/3}]$	$[a^{3/4}]$	$[a^{7/8}]$	a
$a = 50$ and $n = 3$	0.0522	0.0551	0.0601	0.0601	0.0623	0.0635	0.0637	0.0669
$a = 50$ and $n = 5$	0.0551	0.0583	0.0591	0.0591	0.0588	0.0600	0.0612	0.0619
$a = 100$ and $n = 3$	0.0506	0.0521	0.0561	0.0563	0.0594	0.0607	0.0617	0.0634
$a = 100$ and $n = 5$	0.0539	0.0541	0.0541	0.0549	0.0571	0.0578	0.0596	0.0604
$a = 200$ and $n = 3$	0.0436	0.0440	0.0490	0.0497	0.0552	0.0571	0.0601	0.0597
$a = 200$ and $n = 5$	0.0548	0.0520	0.0505	0.0504	0.0515	0.0529	0.0542	0.0549
$a = 500$ and $n = 3$	0.0436	0.0437	0.0452	0.0466	0.0515	0.0558	0.0593	0.0589
$a = 500$ and $n = 5$	0.0533	0.0492	0.0474	0.0481	0.0510	0.0518	0.0532	0.0534
$a = 1000$ and $n = 3$	0.0427	0.0403	0.0405	0.0411	0.0475	0.0513	0.0548	0.0557
$a = 1000$ and $n = 5$	0.0517	0.0486	0.0459	0.0453	0.0466	0.0484	0.0507	0.0521

Table A.9. Power Calculations in Example 4.1

	k_{1000}^{opt}	F_{1000}	$\widetilde{T}_L(20)$	$\widetilde{T}_L(50)$	$\widetilde{T}_L(100)$	$\widetilde{T}_L(250)$	$\widetilde{T}_L(500)$	$\widetilde{T}_L(1000)$
H_1	20	0.8612	0.9992	0.9975	0.9877	0.9482	0.8923	0.8682
H_2	19	0.7887	0.9963	0.9889	0.9685	0.9000	0.8270	0.7978
H_3	18	0.7561	0.9957	0.9878	0.9623	0.8762	0.7971	0.7658
H_4	17	0.7505	0.9952	0.9848	0.9588	0.8743	0.7924	0.7601
H_5	16	0.7541	0.9949	0.9841	0.9591	0.8801	0.7944	0.7633
H_6	15	0.6785	0.9901	0.9712	0.9275	0.8175	0.7238	0.6891
H_7	14	0.6434	0.9859	0.9634	0.9116	0.7856	0.6887	0.6563
H_8	13	0.6432	0.9855	0.9623	0.9100	0.7876	0.6905	0.6547
H_9	12	0.5091	0.9422	0.8861	0.8008	0.6505	0.5518	0.5193
H_{10}	11	0.4434	0.9191	0.8399	0.7351	0.5794	0.4868	0.4553
H_{11}	10	0.4444	0.9191	0.8399	0.7355	0.5742	0.4855	0.4561
H_{12}	9	0.4448	0.9230	0.8414	0.7333	0.5760	0.4847	0.4562
H_{13}	8	0.3896	0.8894	0.7869	0.6756	0.5132	0.4264	0.4007
H_{14}	7	0.2887	0.7710	0.6364	0.5169	0.3835	0.3185	0.2989
H_{15}	6	0.2615	0.7437	0.6051	0.4866	0.3537	0.2903	0.2724
H_{16}	5	0.2095	0.6603	0.5037	0.3878	0.2803	0.2321	0.2187
H_{17}	4	0.2089	0.6560	0.5002	0.3869	0.2742	0.2319	0.2169
H_{18}	3	0.1356	0.4002	0.2874	0.2250	0.1686	0.1482	0.1421
H_{19}	2	0.0816	0.1736	0.1287	0.1106	0.0943	0.0884	0.0867
H_{20}	1	0.0812	0.1743	0.1277	0.1095	0.0934	0.0880	0.0862

Table A.10. Power Calculations in Example 4.2

	k_{1000}^{opt}	F_{1000}	$\widetilde{T}_L(20)$	$\widetilde{T}_L(50)$	$\widetilde{T}_L(100)$	$\widetilde{T}_L(250)$	$\widetilde{T}_L(500)$	$\widetilde{T}_L(1000)$
H_1	20	0.7680	0.9978	0.9886	0.9657	0.8877	0.8089	0.7769
H_2	19	0.7275	0.9968	0.9861	0.9550	0.8603	0.7732	0.7366
H_3	18	0.7241	0.9960	0.9842	0.9533	0.8563	0.7669	0.7330
H_4	17	0.6278	0.9893	0.9640	0.9048	0.7740	0.6731	0.6394
H_5	16	0.6253	0.9886	0.9624	0.9052	0.7702	0.6730	0.6373
H_6	15	0.6188	0.9892	0.9624	0.9031	0.7681	0.6667	0.6306
H_7	14	0.6011	0.9872	0.9577	0.8891	0.7464	0.6462	0.6119
H_8	13	0.5871	0.9872	0.9519	0.8829	0.7369	0.6337	0.5982
H_9	12	0.5831	0.9870	0.9530	0.8819	0.7406	0.6342	0.5962
H_{10}	11	0.5614	0.9849	0.9467	0.8730	0.7151	0.6097	0.5750
H_{11}	10	0.4476	0.9526	0.8704	0.7600	0.5872	0.4900	0.4598
H_{12}	9	0.4009	0.9411	0.8435	0.7224	0.5399	0.4405	0.4121
H_{13}	8	0.2521	0.7461	0.5879	0.4612	0.3297	0.2770	0.2612
H_{14}	7	0.2495	0.7446	0.5843	0.4573	0.3319	0.2742	0.2597
H_{15}	6	0.1831	0.6204	0.4465	0.3361	0.2419	0.2026	0.1913
H_{16}	5	0.1820	0.6119	0.4411	0.3383	0.2407	0.2014	0.1898
H_{17}	4	0.1283	0.4346	0.2941	0.2197	0.1613	0.1412	0.1356
H_{18}	3	0.1296	0.4389	0.2959	0.2195	0.1654	0.1434	0.1363
H_{19}	2	0.1195	0.4202	0.2793	0.2084	0.1515	0.1308	0.1258
H_{20}	1	0.1176	0.4207	0.2763	0.2041	0.1532	0.1296	0.1238

Table A.11. Power Calculations Using a Data-Driven Choice of the Threshold Parameter in Example 4.1

	H_1	H_2	H_3	H_4	H_5	H_6	H_7	H_8	H_9	H_{10}
k_{1000}^{opt}	20	19	18	17	16	15	14	13	12	11
$\widetilde{T}_L(\widehat{k}_{1000}^{opt})$	1.000	0.996	0.993	0.994	0.995	0.987	0.980	0.981	0.938	0.904
	H_{11}	H_{12}	H_{13}	H_{14}	H_{15}	H_{16}	H_{17}	H_{18}	H_{19}	H_{20}
k_{1000}^{opt}	10	9	8	7	6	5	4	3	2	1
$\widetilde{T}_L(\widehat{k}_{1000}^{opt})$	0.911	0.900	0.883	0.774	0.722	0.674	0.661	0.398	0.182	0.175

A.4 Tables and Figures of Chapter 5

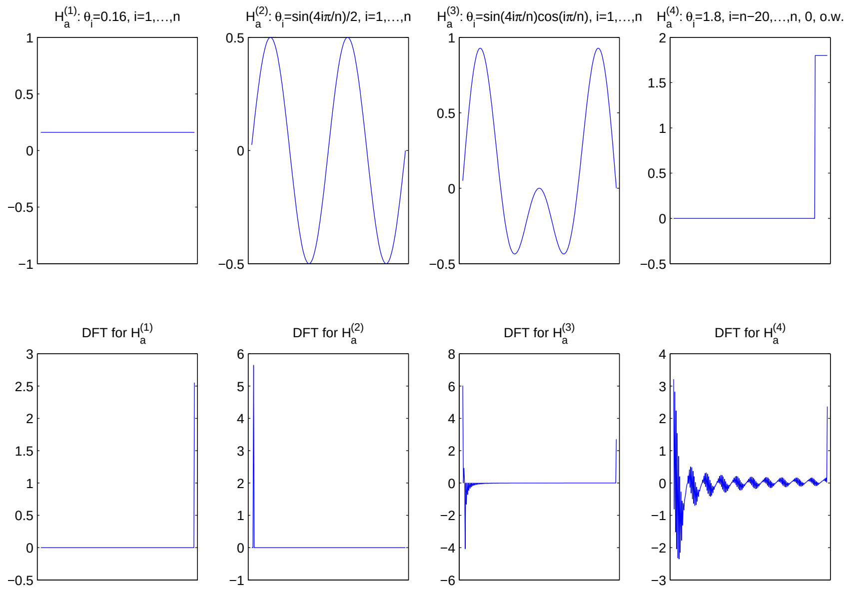

Figure A.4. Top panel: Plot of θ_i versus i. Bottom panel: Plot of $\sum_{j=1}^{n} \gamma_{i,j}\theta_j$ versus i, with an orthonormal matrix $\mathbf{\Gamma} = (\gamma_{i,j})$ corresponding to the discrete Fourier transform (DFT for short).

Table A.12. Power Calculations

Before DFT	$\sum_{i=1}^{255} X_i^2$	$T_L(\widehat{k}_{255}^{opt})$	After DFT	$\sum_{i=1}^{255} X_{\mathbf{\Gamma},i}^2$	$T_L(\widehat{k}_{255}^{opt})$
H_0	0.051	0.056	H_0	0.051	0.075
$H_a^{(1)}$	0.091	0.087	$H_a^{(1)}$	0.091	0.148
$H_a^{(2)}$	0.397	0.294	$H_a^{(2)}$	0.397	0.841
$H_a^{(3)}$	0.816	0.671	$H_a^{(3)}$	0.816	0.992
$H_a^{(4)}$	0.850	0.877	$H_a^{(4)}$	0.850	0.925
$H_a^{(5)}$	0.439	0.535	$H_a^{(5)}$	0.439	0.447

Table A.13. Power Calculations

Before DWT	$\sum_{i=1}^{256} X_i^2$	$T_L(\widehat{k}_{256}^{opt})$	After DWT	$\sum_{i=1}^{256} X_{\mathbf{\Gamma},i}^2$	$T_L(\widehat{k}_{256}^{opt})$
H_0	0.051	0.062	H_0	0.051	0.058
$H_a^{(1)}$	0.093	0.094	$H_a^{(1)}$	0.093	0.096
$H_a^{(2)}$	0.396	0.313	$H_a^{(2)}$	0.396	0.400
$H_a^{(3)}$	0.813	0.661	$H_a^{(3)}$	0.813	0.881
$H_a^{(4)}$	0.861	0.880	$H_a^{(4)}$	0.861	0.998
$H_a^{(5)}$	0.463	0.559	$H_a^{(5)}$	0.463	0.815

Table A.14. Type I Errors of Pearson's Chi-Square, KS, and CVM Test Statistics with $n = 200$

Pearson's χ^2	KS Test	CVM Test
0.0482	0.0484	0.0493

Table A.15. Type I Errors with $n = 200$ after applying a DFT and a DWT

After DFT				After DWT			
$T_H(1.9331)$	Neyman	$T^*(6)$	$\widetilde{T}^*(6)$	$T_H(1.9598)$	Neyman	$T^*(6)$	$\widetilde{T}^*(6)$
0.1012	0.1405	0.0514	0.0503	0.1200	0.1280	0.0473	0.0479

Table A.16. Power Calculations for Example 5.1

	Pearson's			After DFT			After DWT		
	χ^2	KS Test	CVM Test	$T_H(1.93)$	$T^*(6)$	$\widetilde{T}^*(6)$	$T_H(1.96)$	$T^*(6)$	$\widetilde{T}^*(6)$
$c = 1.8$	.9357	.8738	.9545	.9861	.9850	.9767	.9880	.8979	.8738
$c = 1.6$	.7203	.5971	.7095	.8829	.8589	.8214	.8978	.6495	.6139
$c = 1.5$	.5576	.4265	.5101	.7545	.7108	.6606	.7685	.4834	.4485
$c = 1.4$	.3695	.2680	.2911	.5672	.4807	.4366	.5891	.3162	.2945
$c = 1.3$	.2212	.1570	.1532	.3646	.2871	.2608	.3962	.2037	.1920
$c = 1.2$	.1265	.0933	.0842	.2204	.1525	.1427	.2336	.1138	.1076
$c = 0.8$	.2482	.1639	.1603	.2802	.2808	.2198	.2931	.2519	.1987
$c = 0.7$	.6661	.4053	.4567	.5959	.7151	.5631	.6275	.6621	.5079

Table A.17. Power Calculations for Example 5.2

	Pearson's			After DFT			After DWT		
	χ^2	KS Test	CVM Test	$T_H(1.93)$	$T^*(6)$	$\widetilde{T}^*(6)$	$T_H(1.96)$	$T^*(6)$	$\widetilde{T}^*(6)$
$\mu = .400$	.9943	.9360	0.6974	1	.9981	.9981	1	.9948	.9950
$\mu = .367$	.9797	.8752	0.5446	.9997	.9919	.9919	.9995	.9838	.9833
$\mu = .333$	.9605	.7891	0.3960	.9993	.9793	.9796	.9993	.9660	.9657
$\mu = .300$	.9251	.6763	0.2696	.9955	.9457	.9443	.9975	.9436	.9435
$\mu = .267$	.8620	.5564	0.1920	.9906	.8779	.8790	.9897	.8848	.8839
$\mu = .233$	.7741	.4148	0.1226	.9693	.7853	.7861	.9686	.7762	.7767
$\mu = .200$	.6634	.2983	0.0886	.8965	.6637	.6669	.9506	.6376	.6370
$\mu = .100$	.2397	.0879	0.0577	.4412	.2360	.2367	.4359	.2411	.2439

Table A.18. Power Calculations for Example 5.3

	Pearson's			After DFT			After DWT		
	χ^2	KS Test	CVM Test	$T_H(1.93)$	$T^*(6)$	$\widetilde{T}^*(6)$	$T_H(1.96)$	$T^*(6)$	$\widetilde{T}^*(6)$
$\sigma = .125$	.8310	.5747	0.3632	.7882	.8805	.8743	.8012	.8458	.8404
$\sigma = .150$	.7232	.4935	0.3062	.6894	.8009	.7885	.7057	.7336	.7249
$\sigma = .175$	.5922	.4199	0.2545	.5891	.6841	.6726	.6013	.5976	.5896
$\sigma = .200$	.4693	.3530	0.2052	.5033	.5676	.5581	.5195	.4726	.4618
$\sigma = .250$	.3028	.2519	0.1429	.3687	.3792	.3683	.3848	.2975	.2877
$\sigma = .300$	.2055	.1860	0.1083	.2790	.2489	.2435	.2975	.2004	.1940
$\sigma = .400$	.1094	.0972	0.0680	.1795	.1305	.1256	.1992	.1118	.1096

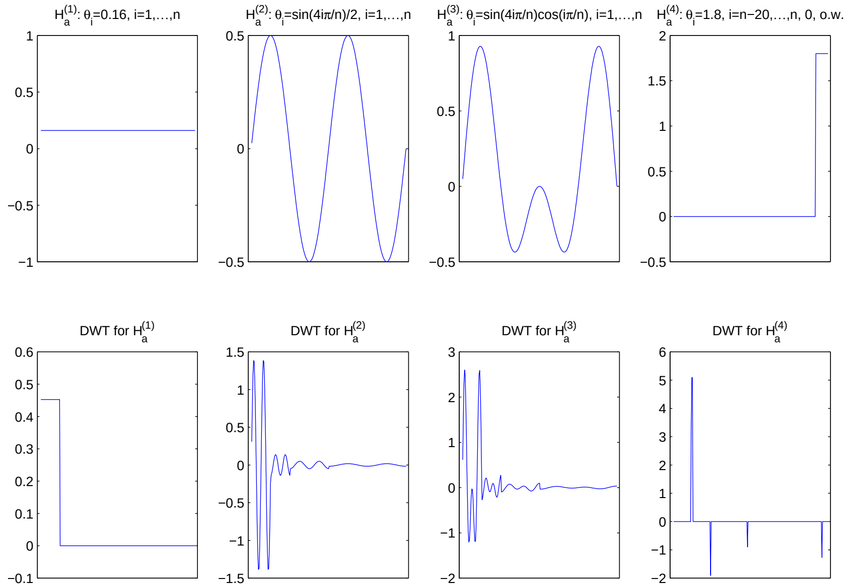

Figure A.5. Top panel: Plot of θ_i versus i. Bottom panel: Plot of $\sum_{j=1}^{n} \gamma_{i,j}\theta_j$ versus i, with an orthonormal matrix $\mathbf{\Gamma} = (\gamma_{i,j})$ corresponding to the discrete wavelet transform (DWT for short).

Table A.19. Power Calculations for Example 5.4

	Pearson's			After DFT			After DWT		
	χ^2	KS Test	CVM Test	$T_H(1.93)$	$T^*(6)$	$\widetilde{T}^*(6)$	$T_H(1.96)$	$T^*(6)$	$\widetilde{T}^*(6)$
$\mu = .400$	.9668	.8344	.8967	.9273	.9804	.8628	.9410	.9637	.8087
$\mu = .367$	.8777	.6692	.7393	.8206	.9124	.7237	.8334	.8777	.6573
$\mu = .333$	.7161	.4881	.5413	.6553	.7724	.5573	.6795	.7161	.4817
$\mu = .300$	.5087	.3244	.3473	.4787	.5675	.3804	.4951	.5110	.3235
$\mu = .267$	.3184	.2132	.2150	.3322	.3711	.2504	.3594	.3276	.2174
$\mu = .233$	.1940	.1480	.1396	.2346	.2194	.1557	.2526	.2025	.1499
$\mu = .200$	.1059	.0935	.0891	.1579	.1215	.1003	.1830	.1180	.0959

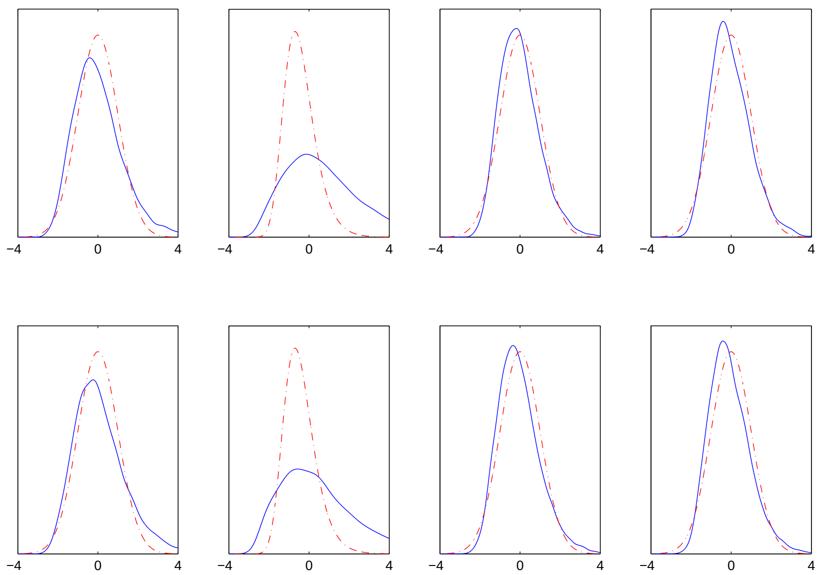

Figure A.6. Top panel: Estimated and theoretical densities of $T_H(1.9331)$, adaptive Neyman, and the two order threshold statistics after the DFT. Bottom panel: Estimated and theoretical densities of $T_H(1.9598)$, adaptive Neyman, and the two order threshold statistics after the DWT.

Table A.20. Power Calculations for Example 5.5

	Pearson's			After DFT			After DWT		
(a, b)	χ^2	KS Test	CVM Test	$T_H(1.93)$	$T^*(6)$	$\widetilde{T}^*(6)$	$T_H(1.96)$	$T^*(6)$	$\widetilde{T}^*(6)$
(.1, .75)	.9770	.7776	.2688	.9988	.9749	.9745	.9998	.9994	.9993
(.1, .85)	.6134	.2016	.0762	.8301	.6016	.6010	.8674	.7062	.7084
(.2, .75)	.9081	.6451	.2056	.9731	.9064	.9068	.9897	.9889	.9887
(.2, .85)	.4880	.1378	.0697	.6584	.4795	.4791	.7038	.5631	.5663
(.3, .75)	.7418	.4732	.1570	.8688	.7502	.7498	.9149	.9293	.9299
(.3, .85)	.3657	.0994	.0670	.4963	.3628	.3638	.5479	.4312	.4353
(.4, .75)	.5620	.3133	.1164	.6914	.5750	.5727	.7516	.7877	.7869
(.4, .85)	.2557	.0702	.0602	.3580	.2612	.2609	.4018	.3078	.3088
(.5, .75)	.3731	.1992	.0940	.4911	.3887	.3862	.5519	.5660	.5649
(.5, .85)	.1806	.0593	.0598	.2685	.1917	.1912	.3027	.2134	.2120
(.6, .75)	.2322	.1184	.0830	.3286	.2414	.2414	.3796	.3595	.3590
(.6, .85)	.1228	.0524	.0526	.1927	.1311	.1337	.2192	.1416	.1421

Table A.21. Power Calculations for Example 5.6

	Pearson's			After DFT			After DWT		
	χ^2	KS Test	CVM Test	$T_H(1.93)$	$T^*(6)$	$\widetilde{T}^*(6)$	$T_H(1.96)$	$T^*(6)$	$\widetilde{T}^*(6)$
$c = 1.8$	.9842	1	1	.9987	.9926	.9922	.9983	.9949	.9944
$c = 1.7$	.9391	.9999	1	.9885	.9595	.9596	.9890	.9703	.9703
$c = 1.6$	.8171	.9972	.9988	.9437	.8550	.8547	.9467	.8789	.8788
$c = 1.5$	.6126	.9720	.9853	.8187	.6631	.6644	.8289	.6809	.6859
$c = 1.4$	.3871	.8505	.8880	.5939	.4285	.4280	.6247	.4284	.4345
$c = 1.3$	.2045	.5330	.5841	.3597	.2324	.2318	.3909	.2183	.2181
$c = 1.2$	.1169	.1964	.2110	.2012	.1329	.1312	.2195	.1158	.1173
$c = 1.1$	.0687	.0552	.0532	.1332	.0837	.0829	.1542	.0771	.0773
$c = 1.0$	.0741	.1697	.1957	.1406	.0850	.0799	.1586	.0806	.0777
$c = 0.9$	.2113	.5684	.6523	.2806	.2215	.2113	.3123	.2214	.2181
$c = 0.8$	.6243	.9303	.9609	.6083	.5956	.5873	.6455	.6732	.6638

Appendix B

PROOFS OF CHAPTERS

B.1 Proofs for Theorems 2.1, 2.2, and 2.3

Proof of Theorem 2.1

The power of the test (2.3), evaluated at the parameter values specified by the simple alternative hypothesis, is given by

$$\begin{aligned}
\text{Power}\left(\boldsymbol{\theta}^{a}\right) &= P_{\boldsymbol{\theta}^{a}}\left[\sum_{i=1}^{n}\left(\theta_{ia}-\theta_{0}\right)\left(X_{i}-\theta_{0}\right) > \sqrt{\sum_{i=1}^{n}\left(\theta_{ia}-\theta_{0}\right)^{2}}\; z_{\alpha}\right] \\
&= P_{\boldsymbol{\theta}^{a}}\left[Z > \frac{\sqrt{\sum_{i=1}^{n}\left(\theta_{ia}-\theta_{0}\right)^{2}}\; z_{\alpha}-\sum_{i=1}^{n}\left(\theta_{ia}-\theta_{0}\right)^{2}}{\sqrt{\sum_{i=1}^{n}\left(\theta_{ia}-\theta_{0}\right)^{2}}}\right] \\
&= 1-\Phi\left(z_{\alpha}-\sqrt{\sum_{i=1}^{n}\left(\theta_{ia}-\theta_{0}\right)^{2}}\right) \rightarrow 1 \quad \text{as} \quad \left\|\boldsymbol{\theta}^{a}-\boldsymbol{\theta}^{0}\right\| \rightarrow \infty.
\end{aligned}$$

Proof of Theorem 2.2

Assume that the true values are θ_{ic}, $i = 1, \ldots, n$, but the statistic has been defined in terms of $\theta_{ia} \neq \theta_{ic}$. The power of the test (2.3) at the parameter values

θ_{ic} is given by

$$\begin{aligned}\text{Power}\left(\boldsymbol{\theta}^c\right) &= P_{\boldsymbol{\theta}^c}\left[\sum_{i=1}^{n}\left(\theta_{ia}-\theta_0\right)\left(X_i-\theta_0\right) > \sqrt{\sum_{i=1}^{n}\left(\theta_{ia}-\theta_0\right)^2}\ z_\alpha\right] \\ &= P_{\boldsymbol{\theta}^c}\left[Z > \frac{\sqrt{\sum_{i=1}^{n}\left(\theta_{ia}-\theta_0\right)^2}\ z_\alpha - \sum_{i=1}^{n}\left(\theta_{ia}-\theta_0\right)\left(\theta_{ic}-\theta_0\right)}{\sqrt{\sum_{i=1}^{n}\left(\theta_{ia}-\theta_0\right)^2}}\right] \\ &= 1-\Phi\left(z_\alpha - \frac{\sum_{i=1}^{n}\left(\theta_{ia}-\theta_0\right)\left(\theta_{ic}-\theta_0\right)}{\sqrt{\sum_{i=1}^{n}\left(\theta_{ia}-\theta_0\right)^2}}\right).\end{aligned}$$

Even though $\|\boldsymbol{\theta}^a-\boldsymbol{\theta}^0\| \to \infty$, power($\boldsymbol{\theta}^c$) may not tend to 1. Power($\boldsymbol{\theta}^c$) is based on the relationship between $(\theta_{ia}-\theta_0)$ and $(\theta_{ic}-\theta_0)$ for each i. However, if $(\theta_{ic}-\theta_0) = k\left(\theta_{ia}-\theta_0\right)$ for any positive value k, then

$$\text{Power}\left(\boldsymbol{\theta}^c\right) = 1-\Phi\left(z_\alpha - k\sqrt{\sum_{i=1}^{n}\left(\theta_{ia}-\theta_0\right)^2}\right) \ \to\ 1 \quad \text{as} \quad \|\boldsymbol{\theta}^a-\boldsymbol{\theta}^0\| \to \infty.$$

Proof of Theorem 2.3 (a)

Since $X_i \sim$ i.i.d. $N(\theta_0, 1)$ under the null hypothesis, $(X_i-\theta_0)^2 \sim \chi^2_1$. By the properties of the chi-square distribution,

$$E\left[(X_i-\theta_0)^2\right] = 1 \quad \text{and} \quad Var\left[(X_i-\theta_0)^2\right] = 2.$$

Thus, by the central limit theorem,

$$S_n^*(\boldsymbol{\theta}^0) = \frac{\sum_{i=1}^{n}(X_i-\theta_0)^2 - n}{\sqrt{2n}} \ \xrightarrow{d}\ N(0,1), \quad \text{as} \quad n \to \infty.$$

Proof of Theorem 2.3 (b)

Under alternatives $\theta_i^n = \theta_0 + \dfrac{\Delta v_i}{\sqrt{n}}$, $(X_i-\theta_0) \sim N\left(\left(\theta_i^n-\theta_0\right), 1\right)$. By the properties of the non-central chi-square distribution, we obtain

$$(X_i - \theta_0)^2 \sim \chi_1^2 \left((\theta_i^n - \theta_0)^2 \right),$$

$$E_{\boldsymbol{\theta}^n} \left[(X_i - \theta_0)^2 \right] = 1 + (\theta_i^n - \theta_0)^2, \; Var_{\boldsymbol{\theta}^n} \left[(X_i - \theta_0)^2 \right] = 2 \left(1 + 2 (\theta_i^n - \theta_0)^2 \right).$$

Let $T_n = \sum_{i=1}^n Y_i$, where $Y_i = (X_i - \theta_0)^2 - E_{\boldsymbol{\theta}^n} [(X_i - \theta_0)^2]$. Then,

$$T_n = S_n(\boldsymbol{\theta}^0) - E_{\boldsymbol{\theta}^n} \left(S_n(\boldsymbol{\theta}^0) \right) = \sum_{i=1}^n (X_i - \theta_0)^2 - n - \sum_{i=1}^n (\theta_i^n - \theta_0)^2.$$

Denote $s_n^2 = Var_{\boldsymbol{\theta}^n}(T_n) = Var_{\boldsymbol{\theta}^n}(S_n(\boldsymbol{\theta}^0)) = 2n + 4 \sum_{i=1}^n (\theta_i^n - \theta_0)^2$. To show the asymptotic normality, we need to check Lyapunov condition:

$$\begin{aligned}
\frac{1}{s_n^4} \sum_{i=1}^n E_{\boldsymbol{\theta}^n}(Y_i^4) &= \frac{\sum_{i=1}^n \left\{ 12 \left(1 + 2 (\theta_i^n - \theta_0)^2 \right)^2 + 48 \left(1 + 4 (\theta_i^n - \theta_0)^2 \right) \right\}}{\left(2n + 4 \sum_{i=1}^n (\theta_i^n - \theta_0)^2 \right)^2} \\
&= \frac{\sum_{i=1}^n \left\{ 60 + 240 (\theta_i^n - \theta_0)^2 + 48 (\theta_i^n - \theta_0)^4 \right\}}{4n^2 + 16 \left(\sum_{i=1}^n (\theta_i^n - \theta_0)^2 \right)^2 + 16n \sum_{i=1}^n (\theta_i^n - \theta_0)^2} \\
&= \frac{60 + \dfrac{240\Delta^2}{n^2} \sum_{i=1}^n v_i^2 + \dfrac{48\Delta^4}{n^3} \sum_{i=1}^n v_i^4}{4n + \dfrac{16\Delta^4}{n^3} (\sum_{i=1}^n v_i^2)^2 + \dfrac{16\Delta^2}{n} \sum_{i=1}^n v_i^2} \to 0 \quad \text{as} \quad n \to \infty,
\end{aligned}$$

with $\sum_{i=1}^n v_i^2 = o\left(n^{3/2}\right)$. Thus, by Lindeberg-Feller central limit theorem, we can conclude that

$$\begin{aligned}
\frac{T_n}{s_n} &= \frac{S_n(\boldsymbol{\theta}^0) - E_{\boldsymbol{\theta}^n} (S_n(\boldsymbol{\theta}^0))}{\sqrt{Var_{\boldsymbol{\theta}^n} (S_n(\boldsymbol{\theta}^0))}} \\
&= \frac{\sum_{i=1}^n (X_i - \theta_0)^2 - n - \sum_{i=1}^n (\theta_i^n - \theta_0)^2}{\sqrt{2n + 4 \sum_{i=1}^n (\theta_i^n - \theta_0)^2}} \xrightarrow{d} N(0,1), \quad \text{as} \quad n \to \infty.
\end{aligned}$$

That is, under the contiguous sequence of alternatives $\theta_i = \theta_i^n$, the proof of the asymptotic normality of $S_n(\boldsymbol{\theta}^0)$ is completed.

The power of $S_n(\boldsymbol{\theta}^0)$ against alternatives θ_i^n is

$$\begin{aligned}
\text{Power}\,(\boldsymbol{\theta}^n) &= P_{\boldsymbol{\theta}^n}\left\{\frac{\sum_{i=1}^n (X_i-\theta_0)^2-n}{\sqrt{2n}} > z_\alpha\right\} \\
&= P_{\boldsymbol{\theta}^n}\left\{Z > \frac{z_\alpha - \sum_{i=1}^n \left(\theta_i^n-\theta_0\right)^2/\sqrt{2n}}{\sqrt{1+2\sum_{i=1}^n \left(\theta_i^n-\theta_0\right)^2/n}}\right\} \\
&\approx P_{\boldsymbol{\theta}^n}\left\{Z > z_\alpha - \sum_{i=1}^n \left(\theta_i^n-\theta_0\right)^2/\sqrt{2n}\right\}.
\end{aligned}$$

We can rewrite the power of $S_n(\boldsymbol{\theta}^0)$ as

$$P_{\boldsymbol{\theta}^n}\left\{Z > z_\alpha - \frac{\Delta^2}{n\sqrt{2n}}\sum_{i=1}^n v_i^2\right\} \;\to\; \alpha \quad \text{as} \;\; n\to\infty,$$

with $\sum_{i=1}^n v_i^2 = o(n^{3/2})$.

B.2 Proofs for Lemmas 3.1, 3.3-3.5, and Theorem 3.1

Proof of Lemma 3.1

We have

$$\begin{aligned}
T_{E,L}(k_n) &= \sum_{i=n-k_n+1}^{n}\sum_{j=1}^{i}\frac{V_j}{n-j+1} \\
&= \sum_{j=1}^{n-k_n+1}\frac{V_j}{n-j+1} + \sum_{j=1}^{n-k_n+2}\frac{V_j}{n-j+1} + \cdots + \sum_{j=1}^{n}\frac{V_j}{n-j+1} \\
&= k_n\sum_{j=1}^{n-k_n+1}\frac{V_j}{n-j+1} + V_{n-k_n+2} + \cdots + V_n \\
&= \sum_{j=1}^{n}\frac{V_j}{n-j+1}\sum_{i=j}^{n} I\left(i > n-k_n\right).
\end{aligned}$$

Thus, $\alpha_{E,jn}(k_n) = \sum_{i=j}^n I\left(i>n-k_n\right)/(n-j+1)$. It completes the proof.

Proof of Theorem 3.1

We want to derive the asymptotic normality by using Lindeberg-Feller central limit theorem. First we need to center $\alpha_{E,in}(k_n)V_i$ by subtracting the mean value. Let

$$Y_i = \alpha_{E,in}(k_n)\left(V_i - 1\right) \quad \text{and} \quad s_n^2 = Var\left(\sum_{i=1}^{n} Y_i\right) = \sum_{i=1}^{n} \alpha_{E,in}(k_n)^2.$$

Next we will check the Lyapunov condition:

$$\frac{1}{s_n^4}\sum_{i=1}^{n} E\left(Y_i^4\right) = \frac{9\sum_{j=1}^{n}\alpha_{E,jn}(k_n)^4}{\left(\sum_{j=1}^{n}\alpha_{E,jn}(k_n)^2\right)^2} \le \frac{9\max_{1\le i\le n}\alpha_{E,in}(k_n)^2}{\sum_{j=1}^{n}\alpha_{E,jn}(k_n)^2}.$$

Since $\alpha_{E,jn}(k_n) = \dfrac{1}{n-j+1}\sum_{i=j}^{n} I\left(i > n-k_n\right)$, $\alpha_{E,jn}(k_n)$ is $\dfrac{k_n}{n-j+1}$ if $1 \le j \le n-k_n+1$, and 1 otherwise. Thus, $\max_{1\le j\le n}\alpha_{E,jn}(k_n)^2 = 1$ and

$$\sum_{j=1}^{n}\alpha_{E,jn}(k_n)^2 = \sum_{j=1}^{n-k_n+1}\frac{k_n^2}{(n-j+1)^2} + (k_n - 1) = \sum_{i=k_n}^{n}\frac{k_n^2}{i^2} + (k_n-1)\,.$$

Since

$$\frac{n-k_n+1}{k_n(n+1)} = \int_{k_n}^{n+1}\frac{1}{x^2}\mathrm{d}x = \sum_{i=k_n}^{n}\int_{i}^{i+1}\frac{1}{x^2}\mathrm{d}x \le \sum_{i=k_n}^{n}\frac{1}{i^2},$$

we have

$$\frac{1}{s_n^4}\sum_{i=1}^{n} E\left(Y_i^4\right) \le 9\left(\frac{k_n(n-k_n+1)}{n+1} + k_n - 1\right)^{-1} \to 0 \quad \text{as} \quad n\to\infty \text{ and } k_n \to \infty.$$

By Lindeberg-Feller central limit theorem, we can conclude that

$$\begin{aligned} T^*_{E,L}(k_n) &= \frac{\sum_{i=1}^{n}\alpha_{E,in}(k_n)\left(V_i-1\right)}{\sqrt{\sum_{i=1}^{n}\alpha_{E,in}(k_n)^2}} \\ &= \frac{T_{E,L}(k_n) - \sum_{i=1}^{n}\alpha_{E,in}(k_n)}{\sqrt{\sum_{i=1}^{n}\alpha_{E,in}(k_n)^2}} \;\overset{d}{\to}\; N(0,1), \quad \text{as} \quad n\to\infty. \end{aligned}$$

Proof of Lemma 3.3

Let $\mathbb{U}_n$ be the uniform empirical process, so

$$\begin{aligned}&P\left(\|\ \mathbb{U}_n\ \|<\lambda\right)\\&=P\left(\max\left\{0,\ \frac{j}{n}-\frac{\lambda}{\sqrt{n}}\right\}<U_{j,n}<\min\left\{\frac{j-1}{n}+\frac{\lambda}{\sqrt{n}},\ 1\right\},\ 1\leq j\leq n\right)\end{aligned}$$

for any $\lambda>0$ (cf. Shorack and Wellner (1986), p. 358). Using

$$P\left(\|\ \mathbb{U}_n\ \|\geq\lambda\right)\leq 58\exp(-2\lambda^2)\ \text{ for all }\ \lambda\geq 0,$$

cf. Shorack and Wellner (1986), p. 354, it follows that

$$\begin{aligned}&P\left(\max\left\{0,\ \frac{j}{n}-\frac{\lambda}{\sqrt{n}}\right\}<U_{j,n}<\min\left\{\frac{j-1}{n}+\frac{\lambda}{\sqrt{n}},\ 1\right\},\ 1\leq j\leq n\right)\\&\geq 1-58\exp(-2\lambda^2).\end{aligned}$$

Take $\epsilon = 58\exp(-2\lambda^2)$. Since, for $1\leq j<n^{1-\delta(n)}\leq n-\sqrt{\frac{n}{2}\log\left(\frac{58}{\epsilon}\right)}+1$, $\min\left\{\frac{j-1}{n}+\sqrt{\frac{1}{2n}\log\left(\frac{58}{\epsilon}\right)},\ 1\right\}=\frac{j-1}{n}+\sqrt{\frac{1}{2n}\log\left(\frac{58}{\epsilon}\right)}$, we have

$$P\left(u_{jn}(\epsilon)<U_{j,n}<u^{jn}(\epsilon),\ \ 1\leq j<n^{1-\delta(n)}\right)\ \ \geq\ \ 1-\epsilon.$$

Expressing this in terms of order statistics, $V_{1,n},\ldots,V_{n,n}$, of the standard exponential distribution, we have

$$P\left(v_{jn}(\epsilon)<V_{j,n}<v^{jn}(\epsilon),\ \ 1\leq j<n^{1-\delta(n)}\right)\ \ \geq\ \ 1-\epsilon. \tag{B.1}$$

Set $V_i=-\log(1-U_i)$ and $X_i=(V_i-1)/(n-i+1)$, $i=1,\ldots,n$. Then $\sum_{i=1}^{j}X_i=V_{j,n}-\tilde{\nu}_{jn}$. Using the Kolmogorov's inequality, for any $\lambda^*>0$,

$$P\left(\max_{1\leq j\leq n}\left|V_{j,n}-\tilde{\nu}_{jn}\right|\geq\lambda^*\right)\leq\frac{1}{\lambda^{*2}}\sum_{j=1}^{n}\frac{1}{(n-j+1)^2}<\frac{2}{\lambda^{*2}}.$$

Take $\epsilon=2/\lambda^{*2}$. Then,

$$P\left(\left|V_{j,n}-\tilde{\nu}_{jn}\right|<\sqrt{\frac{2}{\epsilon}},\quad n^{1-\delta(n)}\leq j\leq n\right)\geq 1-\epsilon. \tag{B.2}$$

Both (B.1) and (B.2) yield (3.6).

Proof of Lemma 3.4

Since

$$\frac{j}{n+1} - u_{jn}(\epsilon) = \begin{cases} \frac{j}{n+1}, & 1 \le j < \sqrt{\frac{n}{2}\log\left(\frac{58}{\epsilon}\right)}, \\ -\frac{j}{n(n+1)} + \sqrt{\frac{1}{2n}\log\left(\frac{58}{\epsilon}\right)}, & \sqrt{\frac{n}{2}\log\left(\frac{58}{\epsilon}\right)} \le j < n^{1-\delta(n)}, \\ -\left(1-\frac{j}{n+1}\right) + e^{-\tilde{\nu}_{jn}} e^{\sqrt{2/\epsilon}}, & n^{1-\delta(n)} \le j \le n, \end{cases}$$

and

$$u^{jn}(\epsilon) - \frac{j}{n+1} = \begin{cases} -\frac{n-j+1}{n(n+1)} + \sqrt{\frac{1}{2n}\log\left(\frac{58}{\epsilon}\right)}, & 1 \le j < n^{1-\delta(n)}, \\ \left(1-\frac{j}{n+1}\right) - e^{-\tilde{\nu}_{jn}} e^{-\sqrt{2/\epsilon}}, & n^{1-\delta(n)} \le j \le n, \end{cases}$$

the relation $u_{jn}(\epsilon) < \frac{j}{n+1} < u^{jn}(\epsilon)$ is clear when $1 \le j < n^{1-\delta(n)}$. As obtained in CGJ1967, $-\log\left(1-\frac{j}{n+1}\right) \le \tilde{\nu}_{jn} < -\log\left(1-\frac{j}{n+1}\right) + \frac{j}{(2n+1)(n-j+0.5)}$, we have

$$\left(1-\frac{j}{n+1}\right) e^{-2n/(2n+1)} < e^{-\tilde{\nu}_{jn}} \le 1-\frac{j}{n+1},$$

so that the relation $u_{jn}(\epsilon) < \frac{j}{n+1} < u^{jn}(\epsilon)$ is also satisfied when $n^{1-\delta(n)} \le j \le n$.

Proof of Lemma 3.5

Using the approximation to the integral,

$$-\log\left(1-\frac{j}{n+1}\right) = \int_{n-j+1}^{n+1} \frac{1}{x}\mathrm{d}x < \tilde{\nu}_{jn} < \int_{n-j}^{n} \frac{1}{x}\mathrm{d}x = -\log\left(1-\frac{j}{n}\right),$$

so that the relation $\nu_{jn} < \tilde{\nu}_{jn}$ is clear. Since also $u_{jn}(\epsilon) < j/(n+1) < u^{jn}(\epsilon)$ (shown in Lemma 3.4), the relation $v_{jn}(\epsilon) < \nu_{jn} < v^{jn}(\epsilon)$ is satisfied. The remaining part is to prove that $\tilde{\nu}_{jn} < v^{jn}(\epsilon)$. When $1 \le j < n^{1-\delta(n)}$,

$$u^{jn}(\epsilon) - \frac{j}{n} = -\frac{1}{n} + \sqrt{\frac{1}{2n}\log\left(\frac{58}{\epsilon}\right)} > 0,$$

so that $\tilde{\nu}_{jn} < v^{jn}(\epsilon)$ holds in this case. When $n^{1-\delta(n)} \le j \le n$,

$$v^{jn}(\epsilon) = \tilde{\nu}_{jn} + \sqrt{\frac{2}{\epsilon}} > \tilde{\nu}_{jn}.$$

Thus, the relation $v_{jn}(\epsilon) < \nu_{jn} < \tilde{\nu}_{jn} < v^{jn}(\epsilon)$ is satisfied in both cases. Next, we study $v^{jn}(\epsilon) - v_{jn}(\epsilon)$, $j = 1, \ldots, n$. When $1 \le j < \sqrt{\frac{n}{2}\log\left(\frac{58}{\epsilon}\right)}$, $v^{jn}(\epsilon) - v_{jn}(\epsilon) = -\log\left(1 - \frac{j-1}{n} - \sqrt{\frac{1}{2n}\log\left(\frac{58}{\epsilon}\right)}\right)$. Thus

$$\sup_{1\le j<\sqrt{\frac{n}{2}\log\left(\frac{58}{\epsilon}\right)}} \left|v^{jn}(\epsilon) - v_{jn}(\epsilon)\right| = -\log\left(1 + \frac{1}{n} - \sqrt{\frac{2}{n}\log\left(\frac{58}{\epsilon}\right)}\right) \to 0,$$

as $n \to \infty$. When $\sqrt{\frac{n}{2}\log\left(\frac{58}{\epsilon}\right)} \le j < n^{1-\delta(n)}$,

$$v^{jn}(\epsilon) - v_{jn}(\epsilon) = \log\left(1 + \frac{-\frac{1}{n} + \sqrt{\frac{2}{n}\log\left(\frac{58}{\epsilon}\right)}}{\frac{n-j+1}{n} - \sqrt{\frac{1}{2n}\log\left(\frac{58}{\epsilon}\right)}}\right).$$

Thus, we have

$$\begin{aligned}&\sup_{\sqrt{\frac{n}{2}\log\left(\frac{58}{\epsilon}\right)}\le j<n^{1-\delta(n)}} \left|v^{jn}(\epsilon) - v_{jn}(\epsilon)\right| \\ &= \log\left(1 + \frac{-\frac{1}{\sqrt{n}} + \sqrt{2\log\left(\frac{58}{\epsilon}\right)}}{\sqrt{n}(1-n^{-\delta(n)}) + \frac{1}{\sqrt{n}} - \sqrt{\frac{1}{2}\log\left(\frac{58}{\epsilon}\right)}}\right) \to 0,\end{aligned}$$

as $n \to \infty$. When $n^{1-\delta(n)} \le j \le n$, $v^{jn}(\epsilon) - v_{jn}(\epsilon) = \sqrt{8/\epsilon}$.

B.3 Lemma B.1, and Proofs for Lemma B.1 and Theorems 3.3

Lemma B.1. *Let $v_{yn}(\epsilon)$ and $v^{yn}(\epsilon)$, $1 \le y \le n$, be given in Lemma 3.3, and let $\tilde{H} = F^{-1} \circ G$, where F is the central chi-squared distribution function with 1 degree of freedom and G is the standard exponential distribution function. Then,*

1. *$\tilde{H}'$ is increasing, positive, concave, and $\tilde{H}'(v) \to 2$, as $v \to \infty$.*

2. *$\tilde{H}''$ is a decreasing positive function, and $\tilde{H}''(v) \to 0$, as $v \to \infty$.*

3. *$\tilde{H}(v)\tilde{H}''(v) \to 0$, as $v \to \infty$.*

4. *$\dfrac{\tilde{H}'''(v)}{\tilde{H}''(v)}(1-e^{-v}) \to 0$, as $v \to 0$, and $\dfrac{\tilde{H}'''(v)}{\tilde{H}''(v)} \to 0$, as $v \to \infty$.*

5. *Assume that $1-n^{-\delta(n)} \to 0$ and $n^{1/2}(1-n^{-\delta(n)}) \to \infty$, as $n \to \infty$. The positive function*

$$R(y) = (v^{yn}(\epsilon) - v_{yn}(\epsilon))\tilde{H}''\left(v_{yn}(\epsilon)\right)\sqrt{\frac{y}{n-y+1}}$$

is increasing on $1 \le y < n^{1-\delta(n)}$. Moreover, for sufficiently large n, $R(y)$ is also increasing on $n^{1-\delta(n)} \le y \le n$.

Proof of Lemma B.1

Parts 1. and 2. The function $\tilde{H}'(v)$ is positive and tends to 2 as $v \to \infty$ since

$$\begin{aligned}
\lim_{v\to\infty} \tilde{H}'(v) &= \lim_{v\to\infty} \frac{e^{-v}}{f\left(F^{-1}(1-e^{-v})\right)} \\
&= \lim_{v\to\infty} \frac{-f\left(F^{-1}(1-e^{-v})\right)}{f'\left(F^{-1}(1-e^{-v})\right)} = \lim_{w\to\infty} \frac{-f(w)}{f'(w)} = 2.
\end{aligned}$$

The function $\tilde{H}''(v)$ is positive and converges to 0 as $v \to \infty$ since

$$\begin{aligned}
\tilde{H}''(v) &= \frac{-e^{-v} f\left(F^{-1}(1-e^{-v})\right) - e^{-v} f'\left(F^{-1}(1-e^{-v})\right)\frac{e^{-v}}{f(F^{-1}(1-e^{-v}))}}{\left[f\left(F^{-1}(1-e^{-v})\right)\right]^2} \\
&= \frac{-e^{-v}}{f\left(F^{-1}(1-e^{-v})\right)} - \left(\frac{e^{-v}}{f\left(F^{-1}(1-e^{-v})\right)}\right)^2 \frac{f'\left(F^{-1}(1-e^{-v})\right)}{f\left(F^{-1}(1-e^{-v})\right)} \\
&= -\tilde{H}'(v) + \left(\tilde{H}'(v)\right)^2\left(\frac{1}{2\tilde{H}(v)} + \frac{1}{2}\right) \to 0 \quad \text{as} \quad v \to \infty.
\end{aligned}$$

Because $\tilde{H}''(v) > 0$, $\tilde{H}'$ is an increasing function. Finally, $\tilde{H}'''(v)$ is an increasing negative function and tends to 0 as $v \to \infty$ by the fact that

$$\begin{aligned}\tilde{H}'''(v) &= -\tilde{H}''(v) + \tilde{H}'(v)\tilde{H}''(v)\left(\frac{1}{\tilde{H}(v)} + 1\right) - \left(\tilde{H}'(v)\right)^2 \frac{\tilde{H}'(v)}{2(\tilde{H}(v))^2} \\ &= \tilde{H}''(v)\left(-1 + \frac{\tilde{H}'(v)}{\tilde{H}(v)}\right) + \tilde{H}'(v)\left(\tilde{H}''(v) - \frac{(\tilde{H}'(v))^2}{2(\tilde{H}(v))^2}\right) \to 0,\end{aligned}$$

as $v \to \infty$. Since $\tilde{H}'''(v) < 0$, $\tilde{H}''$ is decreasing and $\tilde{H}'$ is concave.

Part 3. We have

$$\lim_{v\to\infty} \tilde{H}(v)\tilde{H}''(v) = \lim_{v\to\infty} \tilde{H}(v)\left(\tilde{H}'(v) - 2\right) + 2 = B_{31} + 2 = 0,$$

where

$$\begin{aligned}B_{31} &= \lim_{v\to\infty} \tilde{H}(v)\left(\tilde{H}'(v) - 2\right) = \lim_{w\to\infty} \frac{1 - F(w) - 2f(w)}{f(w)/w} \\ &= \lim_{w\to\infty} \frac{-1 - 2\frac{f'(w)}{f(w)}}{\frac{1}{w^2}\left(w\frac{f'(w)}{f(w)} - 1\right)} = \lim_{w\to\infty} \frac{-2w}{w+3} = -2.\end{aligned}$$

Part 4. We observe that

$$\begin{aligned}&\lim_{v\to 0} \frac{\tilde{H}'''(v)}{\tilde{H}''(v)}\left(1 - e^{-v}\right) \\ &= \lim_{v\to 0}\left(1 - e^{-v}\right)\left(-1 + \frac{\tilde{H}'(v)}{\tilde{H}(v)} + \tilde{H}'(v) - \frac{(\tilde{H}'(v))^3}{2(\tilde{H}(v))^2\tilde{H}''(v)}\right) \\ &= \lim_{v\to 0} \frac{\left(1 - e^{-v}\right)\tilde{H}'(v)}{\tilde{H}(v)} - \frac{1}{2}\lim_{v\to 0} \frac{\left(1 - e^{-v}\right)(\tilde{H}'(v))^3}{(\tilde{H}(v))^2\tilde{H}''(v)} = B_1 - \frac{B_2}{2} = 0,\end{aligned}$$

where

$$\begin{aligned}B_1 &= \lim_{v\to 0} \frac{\left(1 - e^{-v}\right)\tilde{H}'(v)}{\tilde{H}(v)} = \lim_{w\to 0} \frac{F(w)(1 - F(w))}{wf(w)} = \lim_{w\to 0} \frac{F(w)}{wf(w)} \\ &= \lim_{w\to 0} \frac{f(w)}{f(w) + wf'(w)} = \lim_{w\to 0} \frac{1}{1 + wf'(w)/f(w)} = \lim_{w\to 0} \frac{1}{1/2 - w/2} = 2\end{aligned}$$

and

$$\begin{aligned} B_2 &= \lim_{v\to 0} \frac{(1-e^{-v})\,(\tilde{H}'(v))^3}{(\tilde{H}(v))^2\tilde{H}''(v)} = \lim_{w\to 0} \frac{F(w)}{-w^2 f(w)^2 + \frac{wf(w)}{2} + \frac{w^2 f(w)}{2}} \\ &= \lim_{w\to 0} \frac{f(w)}{-2wf(w)^2 - 2w^2 f(w) f'(w) + \frac{f(w)}{2} + \frac{wf'(w)}{2} + wf(w) + \frac{w^2 f'(w)}{2}} \\ &= \lim_{w\to 0} \frac{1}{\frac{1}{4} + \frac{w}{2} - \frac{w^2}{4}} = 4. \end{aligned}$$

Also, we have

$$\lim_{v\to\infty} \frac{\tilde{H}'''(v)}{\tilde{H}''(v)} = 1 - \frac{4}{\lim_{v\to\infty} \tilde{H}''(v)(\tilde{H}(v))^2} = 1 - \frac{4}{B_3} = 0,$$

where

$$\begin{aligned} B_3 &= \lim_{v\to\infty} \tilde{H}''(v)(\tilde{H}(v))^2 \\ &= \lim_{v\to\infty} \tilde{H}'(v)\left\{-(\tilde{H}(v))^2 + \frac{\tilde{H}(v)\tilde{H}'(v)}{2} + \frac{(\tilde{H}(v))^2\tilde{H}'(v)}{2}\right\} \\ &= \lim_{v\to\infty} \tilde{H}'(v)\left\{\frac{(\tilde{H}(v))^2}{2}(\tilde{H}'(v)-2) + \frac{\tilde{H}(v)\tilde{H}'(v)}{2} - \tilde{H}(v) + \tilde{H}(v)\right\} \\ &= \lim_{v\to\infty} \left\{\tilde{H}(v)(\tilde{H}'(v)-2)\tilde{H}(v) + \tilde{H}(v)(\tilde{H}'(v)-2) + 2\tilde{H}(v)\right\} \\ &= \lim_{v\to\infty} \tilde{H}(v)(\tilde{H}'(v)-2) + \lim_{v\to\infty} \tilde{H}(v)\left[\tilde{H}(v)(\tilde{H}'(v)-2)+2\right] \\ &= B_{31} + B_{32} = 4, \end{aligned}$$

with

$$\begin{aligned} B_{32} &= \lim_{v\to\infty} \tilde{H}(v)\left[\tilde{H}(v)(\tilde{H}'(v)-2)+2\right] \\ &= \lim_{w\to\infty} \frac{w(1-F(w)-2f(w)) + 2f(w)}{f(w)/w} \\ &= \lim_{w\to\infty} \frac{(1-F(w)-2f(w)) + w(-f(w)-2f'(w)) + 2f'(w)}{(f'(w)w - f(w))/w^2} \\ &= \lim_{w\to\infty} \frac{\frac{1-F(w)-2f(w)}{f(w)} + w\left(-1-2\frac{f'(w)}{f(w)}\right) + 2\frac{f'(w)}{f(w)}}{\frac{1}{w^2}\left(\frac{f'(w)}{f(w)}w - 1\right)} \end{aligned}$$

$$= \lim_{w\to\infty} \frac{\frac{1-F(w)-2f(w)}{f(w)/w} - 1}{\frac{1}{w}\left(-\frac{w}{2} - \frac{3}{2}\right)} = 6.$$

Part 5. The first derivative of $R(y)$ at y is given by

$$\begin{aligned} R'(y) &= \left\{\frac{\partial}{\partial y}(v^{yn}(\epsilon) - v_{yn}(\epsilon))\right\} \tilde{H}''(v_{yn}(\epsilon)) \sqrt{\frac{y}{n-y+1}} \\ &\quad + (v^{yn}(\epsilon) - v_{yn}(\epsilon)) \left\{\tilde{H}'''(v_{yn}(\epsilon)) \frac{\partial}{\partial y} v_{yn}(\epsilon)\right\} \sqrt{\frac{y}{n-y+1}} \\ &\quad + (v^{yn}(\epsilon) - v_{yn}(\epsilon)) \tilde{H}''(v_{yn}(\epsilon)) \left\{\frac{\partial}{\partial y}\sqrt{\frac{y}{n-y+1}}\right\} \\ &= R(y) \left\{\frac{\frac{\partial}{\partial y}(v^{yn}(\epsilon) - v_{yn}(\epsilon))}{v^{yn}(\epsilon) - v_{yn}(\epsilon)} + \frac{\tilde{H}'''(v_{yn}(\epsilon))}{\tilde{H}''(v_{yn}(\epsilon))} \frac{\partial}{\partial y} v_{yn}(\epsilon) + \frac{n+1}{2y(n-y+1)}\right\}. \end{aligned}$$

Case 1 $\left(1 \le y < \sqrt{\frac{n}{2}\log\left(\frac{58}{\epsilon}\right)}\right)$: Since $\frac{\tilde{H}'''(v)}{\tilde{H}''(v)}(1 - e^{-v}) \to 0$, as $v \to 0$ (shown in Lemma B.1(4)), and $v_{yn}(\epsilon) = 0$, we have $\frac{\tilde{H}'''(v_{yn}(\epsilon))}{\tilde{H}''(v_{yn}(\epsilon))} \frac{\partial}{\partial y} v_{yn}(\epsilon) \approx 0$. Since also $R(y)$, $\frac{\frac{\partial}{\partial y}(v^{yn}(\epsilon) - v_{yn}(\epsilon))}{v^{yn}(\epsilon) - v_{yn}(\epsilon)}$, $\frac{n+1}{2y(n-y+1)} > 0$, the function $R'(y)$ is positive. Thus, $R(y)$ is positive and increasing on $1 \le y < \sqrt{\frac{n}{2}\log\left(\frac{58}{\epsilon}\right)}$.

Case 2 $\left(\sqrt{\frac{n}{2}\log\left(\frac{58}{\epsilon}\right)} \le y < n^{1-\delta(n)}\right)$: Since $v_{yn}(\epsilon) = -\log\left(1 - \frac{y}{n} + \sqrt{\frac{1}{2n}\log\left(\frac{58}{\epsilon}\right)}\right)$ and $v^{yn}(\epsilon) = -\log\left(1 - \frac{y-1}{n} - \sqrt{\frac{1}{2n}\log\left(\frac{58}{\epsilon}\right)}\right)$, we have

$$v^{yn}(\epsilon) - v_{yn}(\epsilon) = \log\left(1 + \frac{\sqrt{2n\log\left(\frac{58}{\epsilon}\right)} - 1}{n - y + 1 - \sqrt{\frac{n}{2}\log\left(\frac{58}{\epsilon}\right)}}\right).$$

Since also $\frac{\partial}{\partial y} v_{yn}(\epsilon) = \frac{1}{n-y+\sqrt{\frac{n}{2}\log\left(\frac{58}{\epsilon}\right)}}$ and $\frac{\partial}{\partial y} v^{yn}(\epsilon) = \frac{1}{n-y+1-\sqrt{\frac{n}{2}\log\left(\frac{58}{\epsilon}\right)}}$, we have

$$\begin{aligned} &\frac{\frac{\partial}{\partial y}(v^{yn}(\epsilon) - v_{yn}(\epsilon))}{v^{yn}(\epsilon) - v_{yn}(\epsilon)} \\ &= \frac{\sqrt{2n\log\left(\frac{58}{\epsilon}\right)} - 1}{\left(n - y + 1 - \sqrt{\frac{n}{2}\log\left(\frac{58}{\epsilon}\right)}\right)\left(n - y + \sqrt{\frac{n}{2}\log\left(\frac{58}{\epsilon}\right)}\right)} \end{aligned}$$

$$\times \frac{1}{\log\left(1+\frac{\sqrt{2n\log\left(\frac{58}{\epsilon}\right)}-1}{n-y+1-\sqrt{\frac{n}{2}\log\left(\frac{58}{\epsilon}\right)}}\right)}.$$

Since $R(y)>0$ and, as shown in computer,

$$\begin{aligned}
&\frac{\frac{\partial}{\partial y}(v^{yn}(\epsilon)-v_{yn}(\epsilon))}{v^{yn}(\epsilon)-v_{yn}(\epsilon)}+\frac{\tilde{H}'''(v_{yn}(\epsilon))}{\tilde{H}''(v_{yn}(\epsilon))}\frac{\partial}{\partial y}v_{yn}(\epsilon)+\frac{n+1}{2y(n-y+1)}\\
=\;&\frac{\sqrt{2n\log\left(\frac{58}{\epsilon}\right)}-1}{\left(n-y+1-\sqrt{\frac{n}{2}\log\left(\frac{58}{\epsilon}\right)}\right)\left(n-y+\sqrt{\frac{n}{2}\log\left(\frac{58}{\epsilon}\right)}\right)\log\left(1+\frac{\sqrt{2n\log\left(\frac{58}{\epsilon}\right)}-1}{n-y+1-\sqrt{\frac{n}{2}\log\left(\frac{58}{\epsilon}\right)}}\right)}\\
&+\frac{\tilde{H}'''(v_{yn}(\epsilon))}{\tilde{H}''(v_{yn}(\epsilon))}\frac{1}{\left(n-y+\sqrt{\frac{n}{2}\log\left(\frac{58}{\epsilon}\right)}\right)}+\frac{n+1}{2y(n-y+1)},\\
>\;&\frac{1}{\left(n-y+\sqrt{\frac{n}{2}\log\left(\frac{58}{\epsilon}\right)}\right)}\left(1+\frac{\tilde{H}'''(v_{yn}(\epsilon))}{\tilde{H}''(v_{yn}(\epsilon))}+\frac{n+1}{2y}\right)>0,
\end{aligned}$$

we have $R'(y)>0$, so that $R(y)$, $\sqrt{\frac{n}{2}\log\left(\frac{58}{\epsilon}\right)}\le y<n^{1-\delta(n)}$, is an increasing positive function.

Case 3 $\left(n^{1-\delta(n)}\le y\le n\right)$: It is easily verified that $R(y)>0$ and $\frac{\partial}{\partial y}(v^{yn}(\epsilon)-v_{yn}(\epsilon))=0$. Also, $\frac{\partial}{\partial y}v_{yn}(\epsilon)=\frac{\partial}{\partial y}\tilde{\nu}_{yn}\approx\tilde{\nu}_{yn}-\tilde{\nu}_{y-1,n}=\frac{1}{n-y+1}$. Thus,

$$R'(y)\approx\frac{R(y)}{n-y+1}\left\{\frac{\tilde{H}'''(v_{yn}(\epsilon))}{\tilde{H}''(v_{yn}(\epsilon))}+\frac{n+1}{2y}\right\}.$$

For sufficiently large n and small $\delta(n)$, as shown in computer,

$$\frac{\tilde{H}'''(v_{yn}(\epsilon))}{\tilde{H}''(v_{yn}(\epsilon))}+\frac{n+1}{2y}>0\quad\left(\text{cf. }\frac{\tilde{H}'''(v_{yn}(\epsilon))}{\tilde{H}''(v_{yn}(\epsilon))}\to 0,\ \text{as } n\to\infty\ \text{(Lemma B.1(4))}\right),$$

so that $R'(y)>0$. Thus, $R(y)$ is positive and increasing on $n^{1-\delta(n)}\le y\le n$.

Proof of Theorem 3.3

We need to check Assumptions A, B, and C of CGJ1967. For convenience, we restate these assumptions below.

Assumption A: $\tilde{H}(v)$ *is continuously differentiable for* $0 < v < \infty$.

Assumption B: For each $\epsilon > 0$,

$$A_n = \sum_{j=n-k_n+1}^{n} \left[\left\{ \sup_{v_{jn}(\epsilon)<v<v^{jn}(\epsilon)} \left| \tilde{H}'(v) - \tilde{H}'(\tilde{\nu}_{jn}) \right| \right\} \sqrt{\frac{j}{n-j+1}} \right] = o(n\sigma_n(k_n)),$$

where $v_{jn}(\epsilon)$, $v^{jn}(\epsilon)$, and $\tilde{\nu}_{jn}$ are given in Lemma 3.3.

Assumption C: $\max_{1\le j\le n} |\alpha_{jn}(k_n)| = o(n^{1/2}\sigma_n(k_n))$.

Assumption A is clearly satisfied. To verify Assumption C, use Lemma B.1(1) to write

$$\frac{\max_{1\le j\le n} |\alpha_{jn}(k_n)|}{\sqrt{n}\sigma_n(k_n)} = \frac{\tilde{H}'(\tilde{\nu}_{nn})}{\sqrt{\sum_{j=1}^{n} \alpha_{jn}^2(k_n)}} \tag{B.3}$$

$$\le \frac{2}{\sqrt{\sum_{j=n-k_n+1}^{n} \left\{ (n-j+1)^{-1} \sum_{i=j}^{n} \tilde{H}'(\tilde{\nu}_{in}) \right\}^2}} \le \frac{2}{\sqrt{\sum_{j=n-k_n+1}^{n} \left\{ \tilde{H}'(\tilde{\nu}_{jn}) \right\}^2}}.$$

Suppose first that $k_n/n \to 0$ as $n \to \infty$. Then

$$\sum_{j=n-k_n+1}^{n} \left\{ \tilde{H}'(\tilde{\nu}_{jn}) \right\}^2 \ge k_n \left\{ \tilde{H}'(\tilde{\nu}_{n-k_n+1,n}) \right\}^2 \to \infty, \quad \text{as} \quad n \to \infty,$$

so that (B.3) tends to zero and Assumption C is satisfied in this case. Next, suppose that $k_n/n \to r$ as $n \to \infty$, for some $0 < r \le 1$. Using the approximation (2.9) of CGJ1967, that is,

$$\tilde{\nu}_{jn} \simeq \nu_{jn} = -\log\left(1 - \frac{j}{n+1}\right),$$

it follows that

$$\frac{1}{n} \sum_{j=n-k_n+1}^{n} \left\{ \tilde{H}'(\tilde{\nu}_{jn}) \right\}^2$$
$$\simeq \frac{1}{n} \sum_{j=1}^{n} I\left(\frac{j}{n+1} > \frac{n-k_n}{n+1}\right) \left\{ \left(1 - \frac{j}{n+1}\right) (F^{-1})'\left(\frac{j}{n+1}\right) \right\}^2$$

$$\to \int_0^1 I(t > 1-r) \left\{ \frac{(1-t)}{f(F^{-1}(t))} \right\}^2 \mathrm{d}t > 0, \quad \text{as} \quad n \to \infty,$$

so that Assumption C is also satisfied. To show Assumption B, we use Lemma B.1(1) and Lemma 3.5 to write $\sup_{v_{jn}(\epsilon)<v<v^{jn}(\epsilon)} \left| \tilde{H}'(v) - \tilde{H}'(\tilde{\nu}_{jn}) \right| \le \tilde{H}'(v^{jn}(\epsilon)) - \tilde{H}'(v_{jn}(\epsilon)) = (v^{jn}(\epsilon) - v_{jn}(\epsilon))\tilde{H}''(\tilde{v}_{jn}(\epsilon))$, $\tilde{v}_{jn}(\epsilon) \in (v_{jn}(\epsilon), v^{jn}(\epsilon))$. Thus,

$$\begin{aligned} \frac{A_n}{n\sigma_n(k_n)} &\le \frac{\sum_{j=n-k_n+1}^n \left[\left\{ (v^{jn}(\epsilon) - v_{jn}(\epsilon))\tilde{H}''(\tilde{v}_{jn}(\epsilon)) \right\} \sqrt{\frac{j}{n-j+1}} \right]}{\sqrt{n \sum_{j=n-k_n+1}^n \left\{ \tilde{H}'(\tilde{\nu}_{jn}) \right\}^2}} \\ &\le \frac{\frac{1}{\sqrt{n}} \sum_{j=n-k_n+1}^n \left[\left\{ (v^{jn}(\epsilon) - v_{jn}(\epsilon))\tilde{H}''(v_{jn}(\epsilon)) \right\} \sqrt{\frac{j}{n-j+1}} \right]}{\sqrt{\sum_{j=n-k_n+1}^n \left\{ \tilde{H}'(\tilde{\nu}_{jn}) \right\}^2}} \qquad \text{(B.4)} \end{aligned}$$

where the last inequality is justified by the fact that $\tilde{H}''$ is a decreasing positive function (Lemma B.1(2)). We need to prove that (B.4) tends to zero as $n \to \infty$. Suppose first that $k_n/n \to 0$, as $n \to \infty$. Divide numerator and denominator of (B.4) by $k_n^{1/2}$ and consider first the numerator. Then,

$$\begin{aligned} &\frac{1}{\sqrt{nk_n}} \sum_{j=n-k_n+1}^n \left[\left\{ (v^{jn}(\epsilon) - v_{jn}(\epsilon))\tilde{H}''(v_{jn}(\epsilon)) \right\} \sqrt{\frac{j}{n-j+1}} \right] \\ &= \begin{cases} \sqrt{\frac{8}{\epsilon}} \frac{1}{\sqrt{nk_n}} \sum_{j=n-k_n+1}^n \left[\tilde{H}''(v_{jn}(\epsilon)) \sqrt{\frac{j}{n-j+1}} \right], & \text{if } k_n < \sqrt{\frac{n}{2} \log\left(\frac{58}{\epsilon}\right)}, \\ \frac{1}{\sqrt{nk_n}} \sum_{j=n-k_n+1}^{n^{1-\delta(n)}-1} \left[\left\{ (v^{jn}(\epsilon) - v_{jn}(\epsilon))\tilde{H}''(v_{jn}(\epsilon)) \right\} \sqrt{\frac{j}{n-j+1}} \right] & \text{o.w.,} \\ +\frac{1}{\sqrt{nk_n}} \sum_{j=n^{1-\delta(n)}}^{n-k_n^{1/4}} \left[\left\{ (v^{jn}(\epsilon) - v_{jn}(\epsilon))\tilde{H}''(v_{jn}(\epsilon)) \right\} \sqrt{\frac{j}{n-j+1}} \right] & \\ +\frac{1}{\sqrt{nk_n}} \sum_{j=n-k_n^{1/4}+1}^{n} \left[\left\{ (v^{jn}(\epsilon) - v_{jn}(\epsilon))\tilde{H}''(v_{jn}(\epsilon)) \right\} \sqrt{\frac{j}{n-j+1}} \right], & \end{cases} \end{aligned}$$

with $1 - \log\left(n - \sqrt{\frac{n}{2}\log\left(\frac{58}{\epsilon}\right)} + 1\right) \Big/ \log n \le \delta(n) < 1 - \log\Big(n - c_\epsilon n^{3/16} k_n^{5/8} +$

$1)\Big/\log n$ $\left(\text{This range is applied only when } k_n \geq \sqrt{\frac{n}{2}\log\left(\frac{58}{\epsilon}\right)}\right)$. Assume that $1-n^{-\delta(n)} \to 0$, $n^{1/2}(1-n^{-\delta(n)}) \to \infty$, and $n^{1/2}(1-n^{-\delta(n)})^{3/2} \to d$ for some $d > 0$, as $n \to \infty$. If $k_n < \sqrt{\frac{n}{2}\log\left(\frac{58}{\epsilon}\right)}$, then $\frac{1}{\sqrt{nk_n}}\sum_{j=n-k_n+1}^{n}\left[\tilde{H}''\left(v_{jn}(\epsilon)\right)\sqrt{\frac{j}{n-j+1}}\right] < \left(\frac{1}{2}\log\left(\frac{58}{\epsilon}\right)\right)^{1/4} n^{1/4}\tilde{H}''\left(v_{nn}(\epsilon)\right) \to 0$, as $n \to \infty$. This inequality is justified by Lemma B.1(5), and the fact that $n^{1/4}\tilde{H}''\left(v_{nn}(\epsilon)\right)$ tends to zero. Suppose that $k_n \geq \sqrt{\frac{n}{2}\log\left(\frac{58}{\epsilon}\right)}$ and $k_n/n \to 0$, as $n \to \infty$. Set $a_n = n^{1-\delta(n)}-1$ and $b_n = n-k_n^{1/4}$. Using Lemma B.1(5) and Lemma 3.5, we have

$$\frac{1}{\sqrt{nk_n}}\sum_{j=n-k_n+1}^{n}\left[\left\{(v^{jn}(\epsilon)-v_{jn}(\epsilon))\tilde{H}''\left(v_{jn}(\epsilon)\right)\right\}\sqrt{\frac{j}{n-j+1}}\right]$$

$$\leq \sqrt{\frac{k_n}{n(1-n^{-\delta(n)})}}(v^{a_n,n}(\epsilon)-v_{a_n,n}(\epsilon))\tilde{H}''\left(v_{a_n,n}(\epsilon)\right) \tag{B.5}$$

$$+\, c_\epsilon\sqrt{\frac{8}{\epsilon}}n^{3/16}\tilde{H}''\left(v_{b_n,n}(\epsilon)\right) + \sqrt{\frac{8}{\epsilon}}k_n^{-1/4}\tilde{H}''(v_{nn}(\epsilon)). \tag{B.6}$$

Since $(n-a_n+1)/n^{1/2} \to \infty$ as $n \to \infty$, using a one-term Taylor expansion we have $v^{a_n,n}(\epsilon) - v_{a_n,n}(\epsilon) \approx \frac{\sqrt{2n\log\left(\frac{58}{\epsilon}\right)-1}}{n-a_n+1-\sqrt{\frac{n}{2}\log\left(\frac{58}{\epsilon}\right)}} = O\left(\frac{1}{n^{1/2}(1-n^{-\delta(n)})}\right)$. Thus, we have

$$(B.5) = O\left(\frac{1}{n^{1/2}(1-n^{-\delta(n)})^{3/2}}\cdot\frac{k_n^{1/2}}{n^{1/2}}\tilde{H}''\left(v_{a_n,n}(\epsilon)\right)\right) \to 0, \quad \text{as} \quad n \to \infty,$$

where it is justified by Lemma B.1(2) and the fact that $v_{a_n,n}(\epsilon)$ tends to infinity. From Lemma B.1(2), the second term of (B.6) tends to 0 as $n \to \infty$. Moreover, as shown in computer, the first term of (B.6) tends to 0 as $n \to \infty$ (even $b_n = n - n^{1/4}$). Since also $\left(\frac{1}{k_n}\sum_{j=n-k_n+1}^{n}\left\{\tilde{H}'\left(\tilde{\nu}_{jn}\right)\right\}^2\right)^{-1/2} \leq \left(\tilde{H}'(\tilde{\nu}_{n-k_n+1,n})\right)^{-1} < \infty$, (B.4) tends to zero and Assumption B is satisfied when $k_n/n \to 0$ as $n \to \infty$. Next, we suppose that for some $0 < r \leq 1$, $k_n/n \to r$ as $n \to \infty$. Divide numerator and denominator of (B.4) by $n^{1/2}$ and consider the numerator and denominator separately. Since

$$\frac{1}{n}\sum_{j=n-k_n+1}^{n}\left[\left\{(v^{jn}(\epsilon)-v_{jn}(\epsilon))\tilde{H}''\left(v_{jn}(\epsilon)\right)\right\}\sqrt{\frac{j}{n-j+1}}\right]$$

$$
\begin{aligned}
&= \frac{1}{n}\sum_{j=n-k_n+1}^{n^{1-\delta(n)}-1}\left[\left\{(v^{jn}(\epsilon)-v_{jn}(\epsilon))\tilde{H}''\left(v_{jn}(\epsilon)\right)\right\}\sqrt{\frac{j}{n-j+1}}\right] \\
&\quad + \frac{1}{n}\sum_{j=n^{1-\delta(n)}}^{n-n^{1/4}}\left[\left\{(v^{jn}(\epsilon)-v_{jn}(\epsilon))\tilde{H}''\left(v_{jn}(\epsilon)\right)\right\}\sqrt{\frac{j}{n-j+1}}\right] \\
&\quad + \frac{1}{n}\sum_{j=n-n^{1/4}+1}^{n}\left[\left\{(v^{jn}(\epsilon)-v_{jn}(\epsilon))\tilde{H}''\left(v_{jn}(\epsilon)\right)\right\}\sqrt{\frac{j}{n-j+1}}\right] \\
&\leq O\left(\frac{\tilde{H}''\left(v_{n^{1-\delta(n)}-1,n}(\epsilon)\right)}{n^{1/2}(1-n^{-\delta(n)})^{3/2}}\right) \\
&\quad + \sqrt{\frac{8}{\epsilon}}n^{3/16}\tilde{H}''\left(v_{n-n^{1/4},n}(\epsilon)\right) + \sqrt{\frac{8}{\epsilon}}n^{-1/4}\tilde{H}''(v_{nn}(\epsilon)) \\
&\to 0 \quad \text{as} \quad n\to\infty,
\end{aligned}
$$

with $1-\log\left(n-\sqrt{\frac{n}{2}\log\left(\frac{58}{\epsilon}\right)}+1\right)\Big/\log n \leq \delta(n) < 1-\log\left(n-n^{13/16}+1\right)\Big/\log n$, and $\left(\frac{1}{n}\sum_{j=n-k_n+1}^{n}\left\{\tilde{H}'\left(\tilde{\nu}_{jn}\right)\right\}^2\right)^{-1/2} < \infty$, the term (B.4) converges to 0 as $n\to\infty$ in this case. Thus, Assumption B holds for both cases. Since Assumptions A, B, and C of CGJ1967 are satisfied, the proof is done.

B.4 Lemmas B.2, B.3, B.4 and Their Proofs

Lemma B.2. *For any* $0<\epsilon<1$ *and some* $\delta(a)$ *which satisfies* $1-\log\Big(a-\sqrt{\frac{a}{2}\log\left(\frac{58}{\epsilon}\right)}+1\Big)\Big/\log a \leq \delta(a) < 1-\log\left(\frac{a}{2}\log\left(\frac{58}{\epsilon}\right)\right)\Big/(2\log a)$, $1-a^{-\delta(a)}\to 0$, $a^{1/2}(1-a^{-\delta(a)})\to\infty$, *and* $a^{1/2}(1-a^{-\delta(a)})^{3/2}\to d$ *for some* $d>0$, *as* $a\to\infty$, *let*

$$
\begin{cases}
u_{ja}(\epsilon) &= \begin{cases}\max\left\{0,\ \frac{j}{a}-\sqrt{\frac{1}{2a}\log\left(\frac{58}{\epsilon}\right)}\right\}, & 1\leq j< a^{1-\delta(a)},\\ 1-e^{-\tilde{\nu}_{ja}}e^{\sqrt{2/\epsilon}}, & a^{1-\delta(a)}\leq j\leq a,\end{cases}\\
u^{ja}(\epsilon) &= \begin{cases}\frac{j-1}{a}+\sqrt{\frac{1}{2a}\log\left(\frac{58}{\epsilon}\right)}, & 1\leq j< a^{1-\delta(a)},\\ 1-e^{-\tilde{\nu}_{ja}}e^{-\sqrt{2/\epsilon}}, & a^{1-\delta(a)}\leq j\leq a,\end{cases}
\end{cases}
$$

where $\tilde{\nu}_{ja} = \sum_{i=1}^{j} 1/(a-i+1)$, *and set*

$$v_{ja}(\epsilon) = -\log\big(1 - u_{ja}(\epsilon)\big), \quad v^{ja}(\epsilon) = -\log\big(1 - u^{ja}(\epsilon)\big).$$

For any $M \geq 0$, *let* $\tilde{H}_{a,M}(v) = G_{a,M}^{-1}(1 - e^{-v})$, *where* $G_{a,M}$ *is the noncentral chi-squared distribution function with 1 degree of freedom and non-centrality parameter* M^2/a. *Then,*

1. $\tilde{H}'_{a,M}$ *is bounded and* $\tilde{H}'_{a,M}(v) \to 2$, *as* $v \to \infty$ *and* $a \to \infty$.

2. $\tilde{H}''_{a,M}(v) = B_{a,M}(v) - (\tilde{H}'_{a,M}(v))^2 J_{a,M}(v)$, *where*

$$\begin{aligned} B_{a,M}(v) &= -\tilde{H}'_{a,M}(v) + \Big(\tilde{H}'_{a,M}(v)\Big)^2 \left(\frac{1}{2} + \frac{1}{2\tilde{H}_{a,M}(v)}\right), \\ J_{a,M}(v) &= \frac{\sum_{k=1}^{\infty} \left\{ \frac{(M^2/a)^k \left(\tilde{H}_{a,M}(v)\right)^{k-1}}{2^{2k}(k-1)!\Gamma(k+1/2)} \right\}}{\sum_{k=0}^{\infty} \left\{ \frac{(M^2/a)^k \left(\tilde{H}_{a,M}(v)\right)^{k}}{2^{2k}k!\Gamma(k+1/2)} \right\}}. \end{aligned}$$

   Note that $B_{a,M}$ *is a decreasing positive function,* $B_{a,M}(v) \to 0$ *as* $v \to \infty$ *and* $a \to \infty$, *and* $J_{a,M}$ *is bounded by* $M^2/(2a)$.

3. *The positive function*

$$R_M(y) = \left(v^{ya}(\epsilon) - v_{ya}(\epsilon)\right) B_{a,M}(v_{ya}(\epsilon)) \sqrt{\frac{y}{a-y+1}}$$

   is increasing on $1 \leq y < a^{1-\delta(a)}$. *Moreover, for sufficiently large* a, $R_M(y)$ *is also increasing on* $a^{1-\delta(a)} \leq y \leq a$.

Proof of Lemma B.2

Part 1. We have

$$\lim_{v\to\infty} \tilde{H}'_{a,M}(v) = \lim_{w\to\infty} \frac{-g_{a,M}(w)}{g'_{a,M}(w)} = \lim_{w\to\infty} \left(\frac{1}{2} + \frac{1}{2w} - \frac{\sum_{k=1}^{\infty} \left\{ \frac{(M^2/a)^k w^{k-1}}{2^{2k}(k-1)!\Gamma(k+1/2)} \right\}}{\sum_{k=0}^{\infty} \left\{ \frac{(M^2/a)^k w^{k}}{2^{2k}k!\Gamma(k+1/2)} \right\}} \right)^{-1}.$$

Since the ratio between the sums is bounded by $M^2/(2a)$, $\tilde{H}'_{a,M}(v) \to 2$, as $v \to \infty$

and $a \to \infty$. Since also $\tilde{H}'_{a,M}$ is continuous and $\tilde{H}'_{a,M}(v) \to 0$ as $v \to 0$, $\tilde{H}'_{a,M}$ is bounded.

Part 2. The second derivative of $\tilde{H}_{a,M}(v)$ at v is given by

$$\begin{aligned}
&\tilde{H}''_{a,M}(v) \\
&= \frac{-e^{-v}}{g_{a,M}\left(G^{-1}_{a,M}(1-e^{-v})\right)} - \left(\frac{e^{-v}}{g_{a,M}\left(G^{-1}_{a,M}(1-e^{-v})\right)}\right)^2 \frac{g'_{a,M}\left(G^{-1}_{a,M}(1-e^{-v})\right)}{g_{a,M}\left(G^{-1}_{a,M}(1-e^{-v})\right)} \\
&= -\tilde{H}'_{a,M}(v) + \left(\tilde{H}'_{a,M}(v)\right)^2 \left(\frac{1}{2} + \frac{1}{2\tilde{H}_{a,M}(v)} - \frac{\sum_{k=1}^{\infty}\left\{\frac{(M^2/a)^k\left(\tilde{H}_{a,M}(v)\right)^{k-1}}{2^{2k}(k-1)!\Gamma(k+1/2)}\right\}}{\sum_{k=0}^{\infty}\left\{\frac{(M^2/a)^k\left(\tilde{H}_{a,M}(v)\right)^{k}}{2^{2k}k!\Gamma(k+1/2)}\right\}}\right) \\
&= B_{a,M}(v) - \left(\tilde{H}'_{a,M}(v)\right)^2 J_{a,M}(v).
\end{aligned}$$

Since $B_{a,M}(v) > 0$ and $B'_{a,M}(v) < 0$, $B_{a,M}$ is a decreasing positive function. Also, from Lemma B.2(1), we observe that $B_{a,M}(v) \to 0$ as $v \to \infty$ and $a \to \infty$. Set $A_{a,k}(v) = \frac{(M^2/a)^k\left(\tilde{H}_{a,M}(v)\right)^k}{2^{2k}k!\Gamma(k+1/2)}$. Then,

$$\sum_{k=1}^{\infty}\left\{\frac{(M^2/a)^k\left(\tilde{H}_{a,M}(v)\right)^{k-1}}{2^{2k}(k-1)!\Gamma\left(k+1/2\right)}\right\} = \frac{M^2}{4a}\sum_{k=0}^{\infty}\left\{\frac{2}{2k+1}A_{a,k}(v)\right\}.$$

Thus, we have

$$J_{a,M}(v) = \frac{\sum_{k=1}^{\infty}\left\{\frac{(M^2/a)^k\left(\tilde{H}_{a,M}(v)\right)^{k-1}}{2^{2k}(k-1)!\Gamma(k+1/2)}\right\}}{\sum_{k=0}^{\infty}\left\{\frac{(M^2/a)^k\left(\tilde{H}_{a,M}(v)\right)^{k}}{2^{2k}k!\Gamma(k+1/2)}\right\}} = \frac{\frac{M^2}{4a}\sum_{k=0}^{\infty}\left\{\frac{2}{2k+1}A_{a,k}(v)\right\}}{\sum_{k=0}^{\infty}\left\{A_{a,k}(v)\right\}} \le \frac{M^2}{2a}.$$

Part 3. It is shown by the similar argument of the proof of Lemma B.1(5) that $R_M(y)$, $1 \le y \le a$, is a positive and increasing function for sufficiently large a and small $\delta(a)$.

Lemma B.3. *Consider the setting of Lemma B.2. Let $g_{a,0}$ and $g_{a,M}$ be the density functions of $\chi^2_1(0)$ and $\chi^2_1(M^2/a)$, respectively. Set $y_{a,0} = \tilde{H}_{a,0}(v_{aa}(\epsilon)) = G^{-1}_{a,0}(1 - e^{-v_{aa}(\epsilon)})$ and $y_{a,M} = \tilde{H}_{a,M}(v_{aa}(\epsilon)) = G^{-1}_{a,M}(1 - e^{-v_{aa}(\epsilon)})$. Then,* *1.* $y_{a,M}$ *is bounded by*

$$\left(2\log(a+1) - 2\log\left(\sqrt{\pi/2}\right) + 2\log\left(e^{M/(2\sqrt{a})} + e^{-M/(2\sqrt{a})}\right)\right)^2.$$

2. $\dfrac{g_{a,0}(y_{a,M})}{g_{a,M}(y_{a,M})} \to 1$*, as* $a \to \infty$*. Also,* $a^{1/4}\left(\dfrac{g_{a,0}(y_{a,M})}{g_{a,M}(y_{a,M})} - 1\right) \to 0$*, as* $a \to \infty$*.*

3. $\dfrac{g_{a,0}(y_{a,0}) - g_{a,0}(y_{a,M})}{g_{a,0}(y_{a,M})} \approx -\left(\dfrac{y_{a,M}^{-1}+1}{2}\right)\dfrac{M^2}{2a}C_{a,M}$*, where* $C_{a,M}$ *is defined in the proof. In particular,* $C_{a,M} = O(y_{a,M})$*.*

4. $a^{1/4}\left(\dfrac{y_{a,0}}{y_{a,M}} - 1\right) \to 0$*, as* $a \to \infty$*. Also,* $a^{1/4}\left(\dfrac{g_{a,0}(y_{a,0})}{g_{a,0}(y_{a,M})} - 1\right) \to 0$*, as* $a \to \infty$*.*

5. $a^{1/4}\left(\dfrac{e^{-v_{aa}(\epsilon)}}{g_{a,0}(y_{a,M})} - 2 + \dfrac{2}{y_{a,0}}\right) \to 0$*, as* $a \to \infty$*.*

6. $a^{1/4}\left(\dfrac{e^{-v_{aa}(\epsilon)}}{g_{a,M}(y_{a,M})} - 2 + \dfrac{2}{y_{a,M}}\right) \to 0$*, as* $a \to \infty$*.*

Proof of Lemma B.3

Part 1. When Φ is a standard normal distribution function, $1-\Phi(v) = P(Z > v) = \int_v^\infty \frac{1}{\sqrt{2\pi}}e^{-t^2/2}\mathrm{d}t \le \int_v^\infty \frac{1}{\sqrt{2\pi}}e^{-t/2}\mathrm{d}t = \sqrt{\frac{2}{\pi}}e^{-v/2}$, $v \ge 1$. The cumulative distribution function of $\chi_1^2(M^2/a)$ is

$$\begin{aligned} G_{a,M}(v) &= P\left(\left(Z + \frac{M}{\sqrt{a}}\right)^2 \le v\right) \\ &= 1 - P\left(Z > \sqrt{v} - \frac{M}{\sqrt{a}}\right) - P\left(Z > \sqrt{v} + \frac{M}{\sqrt{a}}\right) \\ &\ge 1 - \sqrt{\frac{2}{\pi}}e^{-\sqrt{v}/2}\left(e^{M/(2\sqrt{a})} + e^{-M/(2\sqrt{a})}\right), \quad v \ge \left(1 + \frac{M}{\sqrt{a}}\right)^2. \end{aligned}$$

Thus, $1 - e^{-v_{aa}(\epsilon)} = G_{a,M}(y_{a,M}) \ge 1 - \sqrt{2/\pi}e^{-\sqrt{y_{a,M}}/2}\left(e^{M/(2\sqrt{a})} + e^{-M/(2\sqrt{a})}\right)$. Since $v_{aa}(\epsilon) < \nu_{aa} = \log(a+1)$ (from Lemma 3.5),

$$1/(a+1) \le \sqrt{2/\pi}e^{-\sqrt{y_{a,M}}/2}\left(e^{M/(2\sqrt{a})} + e^{-M/(2\sqrt{a})}\right).$$

Thus, $y_{a,M} \le \left(2\log(a+1) - 2\log\left(\sqrt{\pi/2}\right) + 2\log\left(e^{M/(2\sqrt{a})} + e^{-M/(2\sqrt{a})}\right)\right)^2$.

Part 2. Since

$$g_{a,M}(y_{a,M}) = \frac{e^{-y_{a,M}/2-M^2/(2a)}}{2^{1/2}} \sum_{k=0}^{\infty} \left\{ \frac{(M^2/a)^k}{2^{2k}k!\Gamma(k+1/2)} y_{a,M}^{k-1/2} \right\}$$
$$= \frac{1}{2^{1/2}\Gamma(1/2)} y_{a,M}^{-1/2} e^{-y_{a,M}/2} \left[e^{-M^2/2a}\Gamma\left(\frac{1}{2}\right) \sum_{k=0}^{\infty} \left\{ \frac{(M^2/a)^k}{2^{2k}k!\Gamma(k+1/2)} y_{a,M}^{k} \right\} \right],$$

we have

$$\frac{g_{a,M}(y_{a,M})}{g_{a,0}(y_{a,M})} = e^{-M^2/2a}\Gamma\left(\frac{1}{2}\right) \sum_{k=0}^{\infty} \left\{ \frac{(M^2/a)^k}{2^{2k}k!\Gamma(k+1/2)} y_{a,M}^{k} \right\}.$$

Thus,

$$\lim_{a\to\infty} \frac{g_{a,0}(y_{a,M})}{g_{a,M}(y_{a,M})} = \lim_{a\to\infty} \frac{e^{M^2/2a}}{1+\Gamma\left(\frac{1}{2}\right)\sum_{k=1}^{\infty}\left\{\left(\frac{M^2}{4}\cdot\frac{y_{a,M}}{a}\right)^k \frac{1}{k!\Gamma(k+1/2)}\right\}} = 1,$$

since $y_{a,M}/a \to 0$ as $a \to \infty$ (from Lemma B.3(1)). Also, we have

$$\lim_{a\to\infty} a^{1/4}\left(\frac{g_{a,0}(y_{a,M})}{g_{a,M}(y_{a,M})} - 1\right)$$
$$= \lim_{a\to\infty} a^{1/4}\left(\frac{e^{M^2/2a}}{1+\Gamma\left(\frac{1}{2}\right)\sum_{k=1}^{\infty}\left\{\left(\frac{M^2}{4}\cdot\frac{y_{a,M}}{a}\right)^k \frac{1}{k!\Gamma(k+1/2)}\right\}} - 1\right)$$
$$= \lim_{a\to\infty} a^{1/4}\left(e^{M^2/2a} - 1 - \Gamma\left(\frac{1}{2}\right)\sum_{k=1}^{\infty}\left\{\left(\frac{M^2}{4}\cdot\frac{y_{a,M}}{a}\right)^k \frac{1}{k!\Gamma(k+1/2)}\right\}\right)$$
$$= \lim_{a\to\infty} a^{1/4}\left(-\Gamma\left(\frac{1}{2}\right)\sum_{k=1}^{\infty}\left\{\left(\frac{M^2}{4}\cdot\frac{y_{a,M}}{a}\right)^k \frac{1}{k!\Gamma(k+1/2)}\right\}\right)$$
$$= \lim_{a\to\infty} \frac{1}{a^{1/4}}\left(-\Gamma\left(\frac{1}{2}\right)\sum_{k=1}^{\infty}\left\{\left(\frac{M^2}{4}\cdot\frac{y_{a,M}}{\sqrt{a}}\right)^k \frac{1}{k!\Gamma(k+1/2)a^{k/2-1/2}}\right\}\right) = 0,$$

since $a^{1/4}(e^{M^2/2a} - 1) \to 0$ and $y_{a,M}/\sqrt{a} \to 0$ (Lemma B.3(1)) as $a \to \infty$.

Part 3. By Taylor expansion, we write

$$G_{a,M}(v) = G_{a,0}(v) + \frac{M^2}{2a}\left\{\phi'\left(\sqrt{v} - \frac{t^*}{\sqrt{a}}\right) + \phi'\left(\sqrt{v} + \frac{t^*}{\sqrt{a}}\right)\right\},$$

where $t^* \in (0, M)$. From the definition of the quantile function, we have

$$
\begin{aligned}
G_{a,M}^{-1}(u) &= \left\{ v \geq 0: \; G_{a,M}(v) = u \right\} \\
&= \left\{ v \geq 0: \; G_{a,0}(v) = u - \frac{M^2}{2a} \left\{ \phi' \left(\sqrt{v} - \frac{t^*}{\sqrt{a}} \right) + \phi' \left(\sqrt{v} + \frac{t^*}{\sqrt{a}} \right) \right\} \right\} \\
&= G_{a,0}^{-1} \left(u - \frac{M^2}{2a} \left\{ \phi' \left(\sqrt{G_{a,M}^{-1}(u)} - \frac{t^*}{\sqrt{a}} \right) + \phi' \left(\sqrt{G_{a,M}^{-1}(u)} + \frac{t^*}{\sqrt{a}} \right) \right\} \right).
\end{aligned}
$$

Applying Taylor expansion, it follows that

$$
y_{a,M} = y_{a,0} - \frac{M^2}{2a} \frac{\phi' \left(\sqrt{y_{a,M}} - t^*/\sqrt{a} \right) + \phi' \left(\sqrt{y_{a,M}} + t^*/\sqrt{a} \right)}{g_{a,0}(y_{a,\tilde{t}})} = y_{a,0} - \frac{M^2}{2a} C_{a,M},
$$

where $\tilde{t} \in (0, M)$. In particular, we observe that

$$
\begin{aligned}
C_{a,M} &= \frac{\phi' \left(\sqrt{y_{a,M}} - t^*/\sqrt{a} \right) + \phi' \left(\sqrt{y_{a,M}} + t^*/\sqrt{a} \right)}{g_{a,0}(y_{a,\tilde{t}})} \\
&\leq \frac{\left| \phi' \left(\sqrt{y_{a,M}} - t^*/\sqrt{a} \right) \right|}{g_{a,0}(y_{a,M})} + \frac{\left| \phi' \left(\sqrt{y_{a,M}} + t^*/\sqrt{a} \right) \right|}{g_{a,0}(y_{a,M})} = O(y_{a,M}).
\end{aligned}
$$

Since $C_{a,M}/a \to 0$ as $a \to \infty$, using a one-term Taylor expansion we have

$$
g_{a,0}(y_{a,0}) = g_{a,0} \left(y_{a,M} + \frac{M^2}{2a} C_{a,M} \right) \approx g_{a,0}(y_{a,M}) + g'_{a,0}(y_{a,M}) \frac{M^2}{2a} C_{a,M}.
$$

Thus, we obtain

$$
\frac{g_{a,0}(y_{a,0}) - g_{a,0}(y_{a,M})}{g_{a,0}(y_{a,M})} \approx \frac{g'_{a,0}(y_{a,M})}{g_{a,0}(y_{a,M})} \frac{M^2}{2a} C_{a,M} = - \left(\frac{y_{a,M}^{-1} + 1}{2} \right) \frac{M^2}{2a} C_{a,M}.
$$

Part 4. From the proof of Lemma B.3(3), we have

$$
\frac{y_{a,0}}{y_{a,M}} = 1 + \frac{M^2}{2a} \cdot \frac{C_{a,M}}{y_{a,M}},
$$

where $C_{a,M}$ is defined in the proof of Lemma B.3(3). Thus,

$$
a^{1/4} \left(\frac{y_{a,0}}{y_{a,M}} - 1 \right) = \frac{M^2}{2a^{3/4}} \cdot \frac{C_{a,M}}{y_{a,M}} \to 0, \quad \text{as} \quad a \to \infty. \quad \text{(Lemma B.3(3))}
$$

Using Lemma B.3(3) and the facts that $C_{a,M}/a^{3/4},\ y_{a,M}^{-1} \to 0$ as $a \to \infty$, we have

$$a^{1/4}\left(\frac{g_{a,0}(y_{a,0})}{g_{a,0}(y_{a,M})} - 1\right) \approx -\left(\frac{y_{a,M}^{-1}+1}{2}\right)\frac{M^2}{2a^{3/4}}C_{a,M} \to 0, \quad \text{as} \quad a \to \infty.$$

Part 5. We have

$$\begin{aligned}
&a^{1/4}\left(\frac{e^{-v_{aa}(\epsilon)}}{g_{a,0}(y_{a,M})} - 2 + \frac{2}{y_{a,0}}\right)\\
&\quad = a^{1/4}\left(\frac{e^{-v_{aa}(\epsilon)}}{g_{a,0}(y_{a,0})} - 2 + \frac{2}{y_{a,0}}\right) + a^{1/4}\left(\frac{e^{-v_{aa}(\epsilon)}}{g_{a,0}(y_{a,M})} - \frac{e^{-v_{aa}(\epsilon)}}{g_{a,0}(y_{a,0})}\right)\\
&\quad = a^{1/4}\left(\frac{e^{-v_{aa}(\epsilon)}}{g_{a,0}(y_{a,0})} - 2 + \frac{2}{y_{a,0}}\right) + \frac{e^{-v_{aa}(\epsilon)}}{g_{a,0}(y_{a,0})}\, a^{1/4}\left(\frac{g_{a,0}(y_{a,0})}{g_{a,0}(y_{a,M})} - 1\right).
\end{aligned}$$

From the proof of Theorem 3.3, we obtain

$$\begin{aligned}
&\lim_{a\to\infty} a^{1/4}\tilde{H}''_{a,0}(v_{aa}(\epsilon)) = \lim_{a\to\infty} a^{1/4}\left(\tilde{H}'_{a,0}(v_{aa}(\epsilon)) - 2 + \frac{2}{\tilde{H}_{a,0}(v_{aa}(\epsilon))}\right)\\
&\quad = \lim_{a\to\infty} a^{1/4}\left(\frac{e^{-v_{aa}(\epsilon)}}{g_{a,0}(y_{a,0})} - 2 + \frac{2}{y_{a,0}}\right) = 0.
\end{aligned}$$

From Lemma B.1(1) and Lemma B.3(4), the second term also tends to 0 as $a \to \infty$.

Part 6. We have

$$\begin{aligned}
&a^{1/4}\left(\frac{e^{-v_{aa}(\epsilon)}}{g_{a,M}(y_{a,M})} - 2 + \frac{2}{y_{a,M}}\right)\\
&= a^{1/4}\left(\frac{e^{-v_{aa}(\epsilon)}}{g_{a,M}(y_{a,M})} - 2 + \frac{2}{y_{a,M}} - \frac{2g_{a,0}(y_{a,M})}{g_{a,M}(y_{a,M})} + \frac{2g_{a,0}(y_{a,M})}{g_{a,M}(y_{a,M})} + \frac{2}{y_{a,0}} - \frac{2}{y_{a,0}}\right)\\
&= a^{1/4}\left(\frac{e^{-v_{aa}(\epsilon)}}{g_{a,M}(y_{a,M})} - 2 + \frac{2}{y_{a,0}} - \frac{2g_{a,0}(y_{a,M})}{g_{a,M}(y_{a,M})} + \frac{2g_{a,0}(y_{a,M})}{g_{a,M}(y_{a,M})}\right)\\
&\quad + \frac{2a^{1/4}}{y_{a,0}}\left(\frac{y_{a,0}}{y_{a,M}} - 1\right)\\
&= \frac{g_{a,0}(y_{a,M})}{g_{a,M}(y_{a,M})}\, a^{1/4}\left(\frac{e^{-v_{aa}(\epsilon)}}{g_{a,0}(y_{a,M})} - 2 + \frac{2}{y_{a,0}}\right) - \frac{2}{y_{a,0}}a^{1/4}\left(\frac{g_{a,0}(y_{a,M})}{g_{a,M}(y_{a,M})} - 1\right)\\
&\quad + 2a^{1/4}\left(\frac{g_{a,0}(y_{a,M})}{g_{a,M}(y_{a,M})} - 1\right) + \frac{2}{y_{a,0}}a^{1/4}\left(\frac{y_{a,0}}{y_{a,M}} - 1\right).
\end{aligned}$$

From Lemmas B.3(2), B.3(4), B.3(5), and the fact that $y_{a,0}$ tends to infinity, we see that the above term tends to zero.

Remark B.1. *For any $M \geq 0$, we write $\tilde{H}'_{a,M}(v_{aa}(\epsilon)) = \frac{e^{-v_{aa}(\epsilon)}}{g_{a,M}(y_{a,M})}$ and*

$$B_{a,M}(v_{aa}(\epsilon)) = \frac{e^{-v_{aa}(\epsilon)}}{2g_{a,M}(y_{a,M})}\left(\frac{e^{-v_{aa}(\epsilon)}}{g_{a,M}(y_{a,M})} - 2 + \frac{e^{-v_{aa}(\epsilon)}}{g_{a,M}(y_{a,M})} \cdot \frac{1}{y_{a,M}}\right).$$

From Lemmas B.2(1), B.2(2), and B.3(6), we obtain that $a^{1/4}B_{a,M}(v_{aa}(\epsilon)) \to 0$, as $a \to \infty$.

Lemma B.4. *Consider the setting of Lemma B.3. Let $b_a = a - k_a^{1/4}$ with $k_a \geq \sqrt{\frac{a}{2}\log\left(\frac{58}{\epsilon}\right)}$. Set $x_{a,0} = \tilde{H}_{a,0}(v_{b_a,a}(\epsilon)) = G_{a,0}^{-1}(1 - e^{-v_{b_a,a}(\epsilon)})$, $x_{a,M} = \tilde{H}_{a,M}(v_{b_a,a}(\epsilon))$ $= G_{a,M}^{-1}(1 - e^{-v_{b_a,a}(\epsilon)})$. Then,*

1. *$x_{a,M}$ is bounded by*

$$\left(2\log(a+1) - 2\log(k_a^{1/4}+1) - 2\log\left(\sqrt{\pi/2}\right) + 2\log\left(e^{\frac{M}{2\sqrt{a}}} + e^{-\frac{M}{2\sqrt{a}}}\right)\right)^2.$$

2. *$\dfrac{g_{a,0}(x_{a,M})}{g_{a,M}(x_{a,M})} \to 1$, and $a^{3/16}\left(\dfrac{g_{a,0}(x_{a,M})}{g_{a,M}(x_{a,M})} - 1\right) \to 0$, as $a \to \infty$.*

3. *$\dfrac{g_{a,0}(x_{a,0}) - g_{a,0}(x_{a,M})}{g_{a,0}(x_{a,M})} \approx -\left(\dfrac{x_{a,M}^{-1}+1}{2}\right)\dfrac{M^2}{2a}C'_{a,M}$, where*

$$C'_{a,M} = \frac{\phi'\left(\sqrt{x_{a,M}} - t^*/\sqrt{a}\right) + \phi'\left(\sqrt{x_{a,M}} + t^*/\sqrt{a}\right)}{g_{a,0}(x_{a,\tilde{t}})} \quad \text{with} \quad t^*, \tilde{t} \in (0, M).$$

   In particular, $C'_{a,M} = O(x_{a,M})$.

4. *$a^{3/16}\left(\dfrac{x_{a,0}}{x_{a,M}} - 1\right) \to 0$, and $a^{3/16}\left(\dfrac{g_{a,0}(x_{a,0})}{g_{a,0}(x_{a,M})} - 1\right) \to 0$, as $a \to \infty$.*

5. *$a^{3/16}\left(\dfrac{e^{-v_{b_a,a}(\epsilon)}}{g_{a,0}(x_{a,M})} - 2 + \dfrac{2}{x_{a,0}}\right) \to 0$, as $a \to \infty$.*

6. *$a^{3/16}\left(\dfrac{e^{-v_{b_a,a}(\epsilon)}}{g_{a,M}(x_{a,M})} - 2 + \dfrac{2}{x_{a,M}}\right) \to 0$, as $a \to \infty$.*

Proof of Lemma B.4

Part 1. By the same argument of the proof of Lemma B.3(1), $(k_a^{1/4}+1)/(a+1) = e^{-\nu_{b_a,a}} \leq \sqrt{2/\pi}e^{-\sqrt{x_{a,M}}/2}\left(e^{M/(2\sqrt{a})} + e^{-M/(2\sqrt{a})}\right)$. Thus, it is satisfied.

Parts 2.-4. We obtain $x_{a,M}/\sqrt{a} \to 0$ as $a \to \infty$ from Lemma B.4(1). By the same argument of the proof of Lemma B.3(3), we get $C'_{a,M}/a^{3/4} \to 0$ as $a \to \infty$. Also, $x_{a,M} \to \infty$ as $a \to \infty$, since $v_{b_a,a}(\epsilon) \to \infty$ as $a \to \infty$. Thus, using the similar arguments of the proof of Lemmas B.3(2)$-$(4), they are satisfied.

Part 5. From the proof of Theorem 3.3, we obtain

$$\lim_{a\to\infty} a^{3/16}\tilde{H}''_{a,0}(v_{b_a,a}(\epsilon)) = \lim_{a\to\infty} a^{3/16}\left(\frac{e^{-v_{b_a,a}(\epsilon)}}{g_{a,0}(x_{a,0})} - 2 + \frac{2}{x_{a,0}}\right) = 0.$$

Thus, we have

$$\begin{aligned} &a^{3/16}\left(\frac{e^{-v_{b_a,a}(\epsilon)}}{g_{a,0}(x_{a,M})} - 2 + \frac{2}{x_{a,0}}\right) \\ &\quad = a^{3/16}\left(\frac{e^{-v_{b_a,a}(\epsilon)}}{g_{a,0}(x_{a,0})} - 2 + \frac{2}{x_{a,0}}\right) + \frac{e^{-v_{b_a,a}(\epsilon)}}{g_{a,0}(x_{a,0})}\, a^{3/16}\left(\frac{g_{a,0}(x_{a,0})}{g_{a,0}(x_{a,M})} - 1\right) \to 0, \end{aligned}$$

as $a \to \infty$, where the fact that the second term tends to zero is justified by Lemmas B.1(1) and B.4(4).

Part 6. We have

$$\begin{aligned} &a^{3/16}\left(\frac{e^{-v_{b_a,a}(\epsilon)}}{g_{a,M}(x_{a,M})} - 2 + \frac{2}{x_{a,M}}\right) \\ &\quad = \frac{g_{a,0}(x_{a,M})}{g_{a,M}(x_{a,M})}\, a^{3/16}\left(\frac{e^{-v_{b_a,a}(\epsilon)}}{g_{a,0}(x_{a,M})} - 2 + \frac{2}{x_{a,0}}\right) - \frac{2}{x_{a,0}}a^{3/16}\left(\frac{g_{a,0}(x_{a,M})}{g_{a,M}(x_{a,M})} - 1\right) \\ &\qquad + 2a^{3/16}\left(\frac{g_{a,0}(x_{a,M})}{g_{a,M}(x_{a,M})} - 1\right) + \frac{2}{x_{a,0}}a^{3/16}\left(\frac{x_{a,0}}{x_{a,M}} - 1\right) \to 0, \quad \text{as } a \to \infty. \end{aligned}$$

It is justified by Lemmas B.4(2), B.4(4), B.4(5), and the fact that $x_{a,0} \to \infty$ as $a \to \infty$.

Remark B.2. *From Lemmas B.2(1), B.2(2), and B.4(6), we obtain that*

$$a^{3/16}B_{a,M}(v_{a-k_a^{1/4},a}(\epsilon)) \to 0,$$

as $a \to \infty$.

B.5 Proof of Theorem 4.1

For simplicity, let $\tilde{H}_{a,t}(v) = G_{a,t}^{-1}(1 - e^{-v})$. Then,

$$\alpha_{ia}^t(k_a) = \frac{1}{a-i+1} \sum_{j=i}^{a} c_{ja} \frac{e^{-\tilde{\nu}_{ja}}}{g_{a,t}\left(G_{a,t}^{-1}(1 - e^{-\tilde{\nu}_{ja}})\right)} = \frac{1}{a-i+1} \sum_{j=i}^{a} c_{ja} \tilde{H}'_{a,t}\left(\tilde{\nu}_{ja}\right)$$

and

$$\left(\sigma_a^t(k_a)\right)^2 = \frac{1}{a} \sum_{i=1}^{a} (\alpha_{ia}^t(k_a))^2.$$

Let us check Assumptions A, B, and C of CGJ1967, which we restated in the proof of Theorem 3.3. For given any $|t| < M$, Assumption A is clearly satisfied. Next, it is easily verified that for any fixed values of a and v, $\tilde{H}'_{a,t}(v)$ increases as $|t|$ increases. Thus, $\alpha_{ia}^t(k_a)$ and $\sigma_a^t(k_a)$ increase as $|t|$ increases. Let us check Assumption C: For given any $|t| < M$,

$$\frac{\max_{1 \le j \le a} \left|\alpha_{ja}^t(k_a)\right|}{\sqrt{a}\sigma_a^t(k_a)} \le \frac{\max_{1 \le j \le a} \left|\alpha_{ja}^M(k_a)\right|}{\sqrt{a}\sigma_a^0(k_a)} \le \frac{\max_{a-k_a+1 \le j \le a} \tilde{H}'_{a,M}\left(\tilde{\nu}_{ja}\right)}{\sqrt{\sum_{j=a-k_a+1}^{a} \left\{\tilde{H}'_{a,0}(\tilde{\nu}_{ja})\right\}^2}} \to 0,$$

as $a \to \infty$, provided that $k_a \to \infty$, as $a \to \infty$. It is justified by the facts that $\max_{a-k_a+1 \le j \le a} \tilde{H}'_{a,M}\left(\tilde{\nu}_{ja}\right)$ is bounded (Lemma B.2(1)) and $\sum_{j=a-k_a+1}^{a} \left\{\tilde{H}'_{a,0}(\tilde{\nu}_{ja})\right\}^2$ $\to \infty$ as k_a tends to infinity with a. (It was shown in the proof of Theorem 3.3 because it becomes the central chi-square case when $t = 0$.) In order to verify Assumption B, it suffices to show that

$$\frac{\sum_{j=a-k_a+1}^{a} \left[\left\{\sup_{v_{ja}(\epsilon) < v < v^{ja}(\epsilon)} \left|\tilde{H}'_{a,M}(v) - \tilde{H}'_{a,M}\left(\tilde{\nu}_{ja}\right)\right|\right\} \sqrt{\frac{j}{a-j+1}}\right]}{\sqrt{a \sum_{j=a-k_a+1}^{a} \left\{\tilde{H}'_{a,0}\left(\tilde{\nu}_{ja}\right)\right\}^2}} = o(1),$$

where $v_{ja}(\epsilon)$, $v^{ja}(\epsilon)$, and $\tilde{\nu}_{ja}$ are given in Lemma B.2. Using Lemma B.2(2) we write

$$\begin{aligned}
\sup_{v_{ja}(\epsilon)<v<v^{ja}(\epsilon)} \left|\tilde{H}'_{a,M}(v) - \tilde{H}'_{a,M}\left(\tilde{\nu}_{ja}\right)\right| &\leq \left(v^{ja}(\epsilon) - v_{ja}(\epsilon)\right) \cdot \left|\tilde{H}''_{a,M}(v^*_{ja})\right| \\
&\leq \left(v^{ja}(\epsilon) - v_{ja}(\epsilon)\right) B_{a,M}(v^*_{ja}) + \left(v^{ja}(\epsilon) - v_{ja}(\epsilon)\right) \left(\tilde{H}'_{a,M}(v^*_{ja})\right)^2 J_{a,M}(v^*_{ja}) \\
&\leq \left(v^{ja}(\epsilon) - v_{ja}(\epsilon)\right) B_{a,M}(v_{ja}(\epsilon)) + \left(v^{ja}(\epsilon) - v_{ja}(\epsilon)\right) \left(\tilde{H}'_{a,M}(v^*_{ja})\right)^2 \frac{M^2}{2a},
\end{aligned}$$

with some $v^*_{ja} \in (v_{ja}(\epsilon), v^{ja}(\epsilon))$. From the above inequality, we have

$$\begin{aligned}
&\frac{\sum_{j=a-k_a+1}^{a}\left[\left\{\sup_{v_{ja}(\epsilon)<v<v^{ja}(\epsilon)}\left|\tilde{H}'_{a,M}(v) - \tilde{H}'_{a,M}\left(\tilde{\nu}_{ja}\right)\right|\right\}\sqrt{\frac{j}{a-j+1}}\right]}{\sqrt{a\sum_{j=a-k_a+1}^{a}\left\{\tilde{H}'_{a,0}\left(\tilde{\nu}_{ja}\right)\right\}^2}} \\
&\leq \frac{\frac{1}{\sqrt{a}}\sum_{j=a-k_a+1}^{a}\left[\left\{\left(v^{ja}(\epsilon) - v_{ja}(\epsilon)\right) B_{a,M}(v_{ja}(\epsilon))\right\}\sqrt{\frac{j}{a-j+1}}\right]}{\sqrt{\sum_{j=a-k_a+1}^{a}\left\{\tilde{H}'_{a,0}\left(\tilde{\nu}_{ja}\right)\right\}^2}} \qquad \text{(B.7)} \\
&\quad + \frac{\frac{1}{\sqrt{a}}\sum_{j=a-k_a+1}^{a}\left[\left\{\left(v^{ja}(\epsilon) - v_{ja}(\epsilon)\right)\left(\tilde{H}'_{a,M}(v^*_{ja})\right)^2 \frac{M^2}{2a}\right\}\sqrt{\frac{j}{a-j+1}}\right]}{\sqrt{\sum_{j=a-k_a+1}^{a}\left\{\tilde{H}'_{a,0}\left(\tilde{\nu}_{ja}\right)\right\}^2}}. \qquad \text{(B.8)}
\end{aligned}$$

To show that (B.8) tends to zero, we use Lemma B.2(1) and Lemma 3.5 to write

$$\begin{aligned}
\text{(B.8)} \quad &\leq \quad C_\epsilon \frac{1}{a\sqrt{a}} \sum_{j=a-k_a+1}^{a} \sqrt{\frac{j}{a-j+1}} \cdot \left(\sum_{j=a-k_a+1}^{a} \left\{\tilde{H}'_{a,0}\left(\tilde{\nu}_{ja}\right)\right\}^2\right)^{-1/2} \\
&\leq \quad C_\epsilon \cdot \frac{k_a}{a} \cdot \left(\sum_{j=a-k_a+1}^{a} \left\{\tilde{H}'_{a,0}\left(\tilde{\nu}_{ja}\right)\right\}^2\right)^{-1/2} \qquad \text{for some } \; 0 < C_\epsilon < \infty.
\end{aligned}$$

Suppose first that $k_a/a \to 0$ as $a \to \infty$. Then

$$\left(\frac{1}{k_a} \sum_{j=a-k_a+1}^{a} \left\{\tilde{H}'_{a,0}\left(\tilde{\nu}_{ja}\right)\right\}^2\right)^{-1/2} \leq \frac{1}{\tilde{H}'_{a,0}(\tilde{\nu}_{a-k_a+1,a})} < \infty, \qquad \text{(B.9)}$$

so that (B.8) tends to 0 as $a \to \infty$. For some $0 < r \leq 1$, if $k_a/a \to r$ as $a \to \infty$, then

$$\left(\frac{1}{a}\sum_{j=a-k_a+1}^{a}\left\{\tilde{H}'_{a,0}\left(\tilde{\nu}_{ja}\right)\right\}^2\right)^{-1/2} < \infty. \tag{B.10}$$

Thus (B.8) tends to 0 as $a \to \infty$ in both cases. Since (B.8) converges to zero, the remaining part is to prove that (B.7) tends to 0, provided that $k_a \to \infty$ as $a \to \infty$. Suppose first that $k_a < \sqrt{\frac{a}{2}\log\left(\frac{58}{\epsilon}\right)}$. Divide numerator and denominator of (B.7) by $k_a^{1/2}$ and consider first the numerator. From Lemma B.2(3), we have

$$\begin{aligned}&\frac{1}{\sqrt{ak_a}}\sum_{j=a-k_a+1}^{a}\left[\left\{(v^{ja}(\epsilon)-v_{ja}(\epsilon))B_{a,M}(v_{ja}(\epsilon))\right\}\sqrt{\frac{j}{a-j+1}}\right]\\&\leq\sqrt{\frac{8}{\epsilon}}\sqrt{k_a}B_{a,M}(v_{aa}(\epsilon))<\sqrt{\frac{8}{\epsilon}}\left(\frac{1}{2}\log\left(\frac{58}{\epsilon}\right)\right)^{1/4}a^{1/4}B_{a,M}\left(v_{aa}(\epsilon)\right).\end{aligned}$$

Using Lemmas B.2(1), B.2(2), B.3(6), and (B.9), the term (B.7) tends to zero and Assumption B is satisfied in this case. Next, we suppose that $k_a \geq \sqrt{\frac{a}{2}\log\left(\frac{58}{\epsilon}\right)}$ and $k_a/a \to 0$, as $a \to \infty$. Then, from Lemmas B.2(3) and 3.5,

$$\begin{aligned}&\frac{1}{\sqrt{ak_a}}\sum_{j=a-k_a+1}^{a}\left[\left\{(v^{ja}(\epsilon)-v_{ja}(\epsilon))B_{a,M}(v_{ja}(\epsilon))\right\}\sqrt{\frac{j}{a-j+1}}\right]\\&\leq\sqrt{\frac{k_a}{a(1-a^{-\delta(a)})}}(v^{a^{1-\delta(a)}-1,a}(\epsilon)-v_{a^{1-\delta(a)}-1,a}(\epsilon))B_{a,M}\left(v_{a^{1-\delta(a)}-1,a}(\epsilon)\right)\\&+c_\epsilon\sqrt{\frac{8}{\epsilon}}a^{3/16}B_{a,M}\left(v_{a-k_a^{1/4},a}(\epsilon)\right)+\sqrt{\frac{8}{\epsilon}}k_a^{-1/4}B_{a,M}(v_{aa}(\epsilon)),\end{aligned}$$

where $1-\log\left(a-\sqrt{\frac{a}{2}\log\left(\frac{58}{\epsilon}\right)}+1\right)\Big/\log a \leq \delta(a) < 1-\log\left(a-c_\epsilon a^{3/16}k_a^{5/8}+1\right)\Big/\log a$. From Lemmas B.2(1), B.2(2), B.4(6), (B.9), the fact that $v^{a^{1-\delta(a)}-1,a}(\epsilon)-v_{a^{1-\delta(a)}-1,a}(\epsilon)=O\left(\frac{1}{a^{1/2}(1-a^{-\delta(a)})}\right)$, the term (B.7) also tends to zero and Assumption B is satisfied in this case. Lastly, we suppose that for some $0 < r \leq 1$, $k_a/a \to r$ as $a \to \infty$. Divide numerator and denominator of (B.7) by $a^{1/2}$ and consider the numerator and denominator separately. Since (B.10) and

$$
\begin{aligned}
&\frac{1}{a}\sum_{j=a-k_a+1}^{a}\left[\left\{(v^{ja}(\epsilon)-v_{ja}(\epsilon))B_{a,M}\left(v_{ja}(\epsilon)\right)\right\}\sqrt{\frac{j}{a-j+1}}\right]\\
&\le \frac{1}{\sqrt{1-a^{-\delta(a)}}}(v^{a^{1-\delta(a)}-1,a}(\epsilon)-v_{a^{1-\delta(a)}-1,a}(\epsilon))B_{a,M}\left(v_{a^{1-\delta(a)}-1,a}(\epsilon)\right)\\
&\quad+\sqrt{\frac{8}{\epsilon}}a^{3/16}B_{a,M}\left(v_{a-a^{1/4},a}(\epsilon)\right)+\sqrt{\frac{8}{\epsilon}}a^{-1/4}B_{a,M}(v_{aa}(\epsilon))\\
&\to\ 0\quad\text{as}\quad a\to\infty,
\end{aligned}
$$

with $1-\log\left(a-\sqrt{\frac{a}{2}\log\left(\frac{58}{\epsilon}\right)}+1\right)\Big/\log a\le\delta(a)<1-\log\left(a-a^{13/16}+1\right)\Big/\log a$, the term (B.7) converges to 0 as $a\to\infty$ in this case. Thus, Assumption B holds as k_a tends to infinity with a. Since Assumptions A, B, and C of CGJ1967 are satisfied, the proof is done.

B.6 Proofs for Lemmas 4.2-4.4 and Theorem 4.2

Proof of Lemma 4.2

We observe that

$$
\frac{\sqrt{a}Q_a^t(k_a)}{\sigma_a^t(k_a)}=\frac{1}{\sqrt{a}\sigma_a^t(k_a)}\sum_{i=1}^{a}\alpha_{ia}^t(k_a)(V_i-1),
$$

where V_i are i.i.d. random variables with the distribution function $G(v)=1-e^{-v}$, $v\ge0$. Note that $g(v)=e^{-v}$, $v\ge0$, $E(V_i-1)=0$, $Var(V_i-1)=1$, and $E\left(|V_i-1|^3\right)=\int_0^1(1-v)^3e^{-v}\mathrm{d}v+\int_1^\infty(v-1)^3e^{-v}\mathrm{d}v=12/e-2$. Thus the mean, variance, absolute third moment of $\alpha_{ia}^t(k_a)(V_i-1)$ are $E\left(\alpha_{ia}^t(k_a)(V_i-1)\right)=0$, $Var\left(\alpha_{ia}^t(k_a)(V_i-1)\right)=\left(\alpha_{ia}^t(k_a)\right)^2$, $E\left(|\alpha_{ia}^t(k_a)(V_i-1)|^3\right)=|\alpha_{ia}^t(k_a)|^3\times(12/e-2)$.

Let

$$s_a^2 = \sum_{i=1}^{a} Var\left(\alpha_{ia}^t(k_a)(V_i - 1)\right) = \sum_{i=1}^{a} \left(\alpha_{ia}^t(k_a)\right)^2 = a\left(\sigma_a^t(k_a)\right)^2,$$
$$\beta_a^3 = \sum_{i=1}^{a} E\left(|\alpha_{ia}^t(k_a)(V_i - 1)|^3\right) = \left(\frac{12}{e} - 2\right)\sum_{i=1}^{a} |\alpha_{ia}^t(k_a)|^3,$$
$$r_a = \frac{\beta_a^3}{s_a^3} = \left(\frac{12}{e} - 2\right)\frac{\sum_{i=1}^{a} |\alpha_{ia}^t(k_a)|^3}{a^{3/2}\left(\sigma_a^t(k_a)\right)^3}.$$

Using Berry-Esseen theorem of Galambos (1995, page 180), we have

$$\sup_{-\infty<x<\infty} \left|S_a^t\left(xs_a\right) - \Phi(x)\right| \leq 0.8 r_a, \quad \text{as} \quad a \to \infty,$$

where S_a^t is a distribution function of $\sum_{i=1}^{a} \alpha_{ia}^t(k_a)(V_i - 1)$ and Φ is a standard normal distribution function. Thus, we have

$$\sup_{\substack{-M<t<M \\ -\infty<x<\infty}} |H_{a,t}(x) - \Phi(x)| = \sup_{\substack{-M<t<M \\ -\infty<x<\infty}} \left|S_a^t\left(x\sqrt{a}\sigma_a^t(k_a)\right) - \Phi(x)\right|$$
$$\leq 0.8\left(\frac{12}{e} - 2\right) \sup_{-M<t<M} \left\{\frac{\max_{1\leq i\leq a} |\alpha_{ia}^t(k_a)|}{\sqrt{a}\sigma_a^t(k_a)}\right\} \to 0 \quad \text{as} \quad a \to \infty,$$

provided that $k_a \to \infty$, as $a \to \infty$.

Proof of Lemma 4.3

For convenience, we rewrite

$$R_a^t(k_a) = \frac{1}{a} \sum_{j=a-k_a+1}^{a} \Big\{ \left(V_{j,a} - \tilde{\nu}_{ja}\right) G_{ja}^t\left(V_{j,a}\right) \Big\},$$

where

$$G_{ja}^t(v) = \begin{cases} \dfrac{\tilde{H}_{a,t}(v) - \tilde{H}_{a,t}(\tilde{\nu}_{ja})}{v - \tilde{\nu}_{ja}} - \tilde{H}'_{a,t}\left(\tilde{\nu}_{ja}\right) & \text{if} \ \ v \neq \tilde{\nu}_{ja} \\ 0 & \text{if} \ \ v = \tilde{\nu}_{ja}. \end{cases}$$

Let

$$g_{ja}^t(\epsilon) = \sup_{v_{ja}(\epsilon)<v<v^{ja}(\epsilon)} \left|G_{ja}^t(v)\right|,$$

where $v_{ja}(\epsilon)$ and $v^{ja}(\epsilon)$ are given in Lemma B.2. Then, we have

$$P\left\{ \left| R_a^t(k_a) \right| \le \frac{1}{a} \sum_{j=a-k_a+1}^{a} g_{ja}^M(\epsilon) \left| V_{j,a} - \tilde{\nu}_{ja} \right| \text{ for all } |t| < M \right\} \ge 1 - \epsilon. \quad \text{(B.11)}$$

It follows from (B.11) that

$$P\left\{ \sup_{-M<t<M} \left| R_a^t(k_a) \right| \le \frac{1}{a} \sum_{j=a-k_a+1}^{a} g_{ja}^M(\epsilon) \left| V_{j,a} - \tilde{\nu}_{ja} \right| \right\} \ge 1 - \epsilon.$$

From Assumption B and Proposition 2 of CGJ1967, we have

$$\sum_{j=a-k_a+1}^{a} g_{ja}^M(\epsilon) \left| V_{j,a} - \tilde{\nu}_{ja} \right| = o_p\left(\sqrt{a} \sigma_a^M(k_a) \right),$$

so that $\sup_{-M<t<M} |\sqrt{a} R_a^t(k_a)| = o_p\left(\sigma_a^M(k_a) \right)$. Also, it is easily verified that $\sigma_a^M(k_a)/\sigma_a^0(k_a) = O(1)$ (Lemma 4.5), provided that $k_a \to \infty$ as $a \to \infty$. Consequently,

$$\begin{aligned} \sup_{-M<t<M} \left| \frac{\sqrt{a} R_a^t(k_a)}{\sigma_a^t(k_a)} \right| &\le \sup_{-M<t<M} \frac{|\sqrt{a} R_a^t(k_a)|}{\sigma_a^0(k_a)} \\ &= \sup_{-M<t<M} \frac{|\sqrt{a} R_a^t(k_a)|}{\sigma_a^M(k_a)} \cdot \frac{\sigma_a^M(k_a)}{\sigma_a^0(k_a)} = o_p(1). \end{aligned}$$

Proof of Lemma 4.4

We have already proved that for given any $|t| < M$,

$$T_L^{t\,*}(k_a) = \frac{\sqrt{a} Q_a^t(k_a)}{\sigma_a^t(k_a)} + \frac{\sqrt{a} R_a^t(k_a)}{\sigma_a^t(k_a)} \xrightarrow{d} N(0,1), \quad \text{as } a \to \infty, \quad \text{(Theorem 4.1)}$$

provided that $k_a \to \infty$, as $a \to \infty$. From Lemmas 4.2 and 4.3, we have

$$\sup_{\substack{-M<t<M \\ -\infty<x<\infty}} |F_{a,t}(x) - \Phi(x)| \to 0, \quad \text{as } a \to \infty,$$

provided that $k_a \to \infty$, as $a \to \infty$.

Proof of Theorem 4.2

For any given $\delta_1 > 0$, there exists $M > 0$ such that

$$P\left(|t| \geq M\right) < \delta_1.$$

From Lemma 4.4, any given $\delta_2 > 0$, there exists a_0 such that

$$\left|P\left(\widehat{T}_L^*(k_a) \leq x \middle| \, |t| < M\right) - \Phi(x)\right| < \delta_2 \quad \text{for all } a > a_0.$$

Thus, we have

$$\begin{aligned}
&\left|P\left(\widehat{T}_L^*(k_a) \leq x\right) - \Phi(x)\right| \\
&\leq \left|P\left(\widehat{T}_L^*(k_a) \leq x \middle| \, |t| < M\right) - \Phi(x)\right| \cdot P\left(|t| < M\right) \\
&\quad + \left|P\left(\widehat{T}_L^*(k_a) \leq x \middle| \, |t| \geq M\right) - \Phi(x)\right| \cdot P\left(|t| \geq M\right) \\
&\leq \left|P\left(\widehat{T}_L^*(k_a) \leq x \middle| \, |t| < M\right) - \Phi(x)\right| + P\left(|t| \geq M\right) \\
&< \delta_1 + \delta_2 \quad \text{for all } a > a_0.
\end{aligned}$$

Take $\epsilon/2 = \max\{\delta_1, \delta_2\}$. Then,

$$\left|P\left(\widehat{T}_L^*(k_a) \leq x\right) - \Phi(x)\right| < \epsilon \quad \text{for all } a > a_0.$$

Thus, provided that $k_a \to \infty$, as $a \to \infty$, we have

$$\widehat{T}_L^*(k_a) \stackrel{d}{\to} N(0,1), \quad \text{as} \quad a \to \infty.$$

B.7 Proofs for Lemmas 4.5-4.7 and Theorem 4.3

Proof of Lemma 4.5

We need to show that

$$\sup_{-M<t<M} \left|\frac{\sigma_a^t(k_a)}{\sigma_a^0(k_a)} - 1\right| = \frac{\sigma_a^M(k_a) - \sigma_a^0(k_a)}{\sigma_a^0(k_a)} \to 0, \quad \text{as} \quad a \to \infty,$$

provided that $k_a \to \infty$ as $a \to \infty$. Suppose first that $k_a/a \to 0$ as $a \to \infty$. Since $\sqrt{a/k_a}\sigma_a^0(k_a) > 0$, it is enough to show that $\sqrt{a/k_a}\left(\sigma_a^M(k_a) - \sigma_a^0(k_a)\right) \to 0$, as $a \to \infty$. We firstly study $a(\sigma_a^M(k_a))^2$.

$$\begin{aligned}
a(\sigma_a^M(k_a))^2 &\le \left(k_a^2 \sum_{i=1}^{a-k_a}\left(\frac{1}{a-i+1}\right)^2 + k_a\right)\left\{\max_{a-k_a+1\le j\le a} \tilde{H}'_{a,M}(\tilde{\nu}_{ja})\right\}^2 \\
&\le \left(k_a\left(2-\frac{k_a}{a}\right)\right)\left\{\max_{a-k_a+1\le j\le a} \tilde{H}'_{a,M}(\tilde{\nu}_{ja})\right\}^2.
\end{aligned}$$

We also have

$$\begin{aligned}
a(\sigma_a^0(k_a))^2 &\ge \left(k_a^2 \sum_{i=1}^{a-k_a}\left(\frac{1}{a-i+1}\right)^2 + k_a\right)\left\{\tilde{H}'_{a,0}(\tilde{\nu}_{a-k_a+1,a})\right\}^2 \\
&\ge \left(k_a\left(1+\frac{k_a(a-k_a)}{(a+1)(k_a+1)}\right)\right)\left\{\tilde{H}'_{a,0}(\tilde{\nu}_{a-k_a+1,a})\right\}^2.
\end{aligned}$$

Consequently, we obtain

$$\begin{aligned}
&\sqrt{\frac{a}{k_a}}\left(\sigma_a^M(k_a) - \sigma_a^0(k_a)\right) \\
&\le \sqrt{2-\frac{k_a}{a}}\left\{\max_{a-k_a+1\le j\le a} \tilde{H}'_{a,M}(\tilde{\nu}_{ja})\right\} \\
&\quad - \sqrt{1+\frac{k_a(a-k_a)}{(a+1)(k_a+1)}}\left\{\tilde{H}'_{a,0}(\tilde{\nu}_{a-k_a+1,a})\right\} \\
&\to 0, \quad \text{as} \quad a \to \infty.
\end{aligned}$$

Next, we suppose that for some $0 < r \le 1$, $k_a/a \to r$ as $a \to \infty$. Then $\sigma_a^0(k_a) > 0$, so we need to prove that $\sigma_a^M(k_a) - \sigma_a^0(k_a) \to 0$, as $a \to \infty$. We observe that

$$\begin{aligned}
&(\sigma_a^M(k_a))^2 - (\sigma_a^0(k_a))^2 \\
&= \frac{1}{a}\sum_{i=1}^{a-k_a}\left(\frac{1}{a-i+1}\right)^2\left\{\left(\sum_{j=a-k_a+1}^{a}\tilde{H}'_{a,M}(\tilde{\nu}_{ja})\right)^2 - \left(\sum_{j=a-k_a+1}^{a}\tilde{H}'_{a,0}(\tilde{\nu}_{ja})\right)^2\right\} \\
&+\frac{1}{a}\sum_{i=a-k_a+1}^{a}\left\{\left(\frac{1}{a-i+1}\sum_{j=i}^{a}\tilde{H}'_{a,M}(\tilde{\nu}_{ja})\right)^2 - \left(\frac{1}{a-i+1}\sum_{j=i}^{a}\tilde{H}'_{a,0}(\tilde{\nu}_{ja})\right)^2\right\}
\end{aligned}$$

$$
\begin{aligned}
&\le \frac{k_a}{a}\left(2-\frac{k_a}{a}\right)\left(\max_{a-k_a+1\le j\le a}\tilde{H}'_{a,M}(\tilde{\nu}_{ja})+2\right) \\
&\times\left[\max_{a-k_a+1\le j\le a}\left(\tilde{H}'_{a,M}(\tilde{\nu}_{ja})-\tilde{H}'_{a,0}(\tilde{\nu}_{ja})\right)\right] \\
&= K\left[\max_{a-k_a+1\le j\le a}\left(\tilde{H}'_{a,M}(\tilde{\nu}_{ja})-\tilde{H}'_{a,0}(\tilde{\nu}_{ja})\right)\right]\to 0, \quad \text{as} \ \ a\to\infty.
\end{aligned}
$$

Since $(\sigma_a^M(k_a))^2-(\sigma_a^0(k_a))^2=\left(\sigma_a^M(k_a)+\sigma_a^0(k_a)\right)\left(\sigma_a^M(k_a)-\sigma_a^0(k_a)\right)$, we have

$$
\sigma_a^M(k_a)-\sigma_a^0(k_a) \ \to \ 0, \quad \text{as} \ \ a\to\infty.
$$

Proof of Lemma 4.6

We hope to show that

$$
\sup_{-M<t<M}\left|\frac{\sqrt{a}(\mu_a^t(k_a)-\mu_a^0(k_a))}{\sigma_a^0(k_a)}\right| = \frac{\sqrt{a}(\mu_a^M(k_a)-\mu_a^0(k_a))}{\sigma_a^0(k_a)}\to 0, \quad \text{as} \ \ a\to\infty,
$$

provided that $k_a\to\infty$, as $a\to\infty$. From the fact that $G_{a,M}^{-1}(1-e^{-\tilde{\nu}_{ia}})-G_{a,0}^{-1}(1-e^{-\tilde{\nu}_{ia}})$ is increasing in i and Taylor expansion, we have

$$
\begin{aligned}
&\mu_a^M(k_a)-\mu_a^0(k_a)=\frac{1}{a}\sum_{i=a-k_a+1}^{a}\left(G_{a,M}^{-1}(1-e^{-\tilde{\nu}_{ia}})-G_{a,0}^{-1}(1-e^{-\tilde{\nu}_{ia}})\right) \\
&\le \frac{k_a}{a}\left(G_{a,M}^{-1}(1-e^{-\tilde{\nu}_{aa}})-G_{a,0}^{-1}(1-e^{-\tilde{\nu}_{aa}})\right)=\frac{k_a}{a}\cdot O\left(\frac{M^2}{a}G_{a,M}^{-1}(1-e^{-\tilde{\nu}_{aa}})\right).
\end{aligned}
$$

Note that the last equality is justified by the similar argument of the proof of Lemma B.3(3). Applying the same argument of the proof of Lemma B.3(1), it follows that

$$
\begin{aligned}
&G_{a,M}^{-1}(1-e^{-\tilde{\nu}_{aa}}) \\
&\le \left(2\log(a+1)-2\log\left(\sqrt{\pi/2}\right)+2\log\left(e^{M/(2\sqrt{a})}+e^{-M/(2\sqrt{a})}\right)\right)^2.
\end{aligned}
$$

Suppose first that $k_a/a\to 0$ as $a\to\infty$. Since $\sqrt{a/k_a}\sigma_a^0(k_a)>0$, it is enough to show that

$$\frac{a}{\sqrt{k_a}}(\mu_a^M(k_a) - \mu_a^0(k_a)) \to 0, \quad \text{as} \ \ a \to \infty. \tag{B.12}$$

Since $\sqrt{\frac{k_a}{a}} \cdot \frac{G_{a,M}^{-1}(1-e^{-\tilde{\nu}_{aa}})}{\sqrt{a}} \to 0$ as $a \to \infty$, (B.12) is satisfied. Next, we suppose that for some $0 < r \leq 1$, $k_a/a \to r$ as $a \to \infty$. Then, $\sigma_a^0(k_a) > 0$, so we need to prove that

$$\sqrt{a}(\mu_a^M(k_a) - \mu_a^0(k_a)) \to 0, \quad \text{as} \ \ a \to \infty. \tag{B.13}$$

Since $\frac{k_a}{a} \cdot \frac{G_{a,M}^{-1}(1-e^{-\tilde{\nu}_{aa}})}{\sqrt{a}} \to 0$ as $a \to \infty$, (B.13) is also satisfied.

Proof of Lemma 4.7

Suppose that $k_a/a \to r$, $0 \leq r \leq 1$, and $k_a \to \infty$, as $a \to \infty$. Then,

$$\begin{aligned}
\mu_a^0(k_a) &= \frac{1}{a}\sum_{i=a-k_a+1}^{a} \tilde{H}_{a,0}(\tilde{\nu}_{ia}) \simeq \frac{1}{a}\sum_{i=1}^{a} I\left(\frac{i}{a+1} > \frac{a-k_a}{a+1}\right) G_{a,0}^{-1}\left(\frac{i}{a+1}\right) \\
&\to \int_0^1 I(t > 1-r) G_{a,0}^{-1}(t)\mathrm{d}t \ = \ \int_{G_{a,0}^{-1}(1-r)}^{\infty} u g_{a,0}(u)\mathrm{d}u,
\end{aligned}$$

as $a \to \infty$. Note that if $r = 1$, $\mu_a^0(k_a) \to 1$, as $a \to \infty$. Also, we have

$$\begin{aligned}
\left(\sigma_a^0(k_a)\right)^2 &\simeq \frac{1}{a^2}\sum_{j=1}^{a}\sum_{l=1}^{a}\left\{ I\left(\frac{j}{a+1} > \frac{a-k_a}{a+1}\right) I\left(\frac{l}{a+1} > \frac{a-k_a}{a+1}\right)\right. \\
&\quad \times \left(1 - \frac{j}{a+1}\right)\left(1 - \frac{l}{a+1}\right) \frac{\min\{j/(a+1), l/(a+1)\}}{1 - \min\{j/(a+1), l/(a+1)\}} \\
&\quad \left. \times \frac{1}{g_{a,0}\left(G_{a,0}^{-1}\left(j/(a+1)\right)\right)} \frac{1}{g_{a,0}\left(G_{a,0}^{-1}\left(l/(a+1)\right)\right)} \right\} \\
&\to \int_0^1\int_0^1 I(t > 1-r) I(s > 1-r)(\min(t,s) - ts) \\
&\quad \times \frac{1}{g_{a,0}\left(G_{a,0}^{-1}(t)\right)} \frac{1}{g_{a,0}\left(G_{a,0}^{-1}(s)\right)} \mathrm{d}t\mathrm{d}s \\
&= \int_0^1\int_0^1 I(t > 1-r) I(s > 1-r)(\min(t,s) - ts)\mathrm{d}G_{a,0}^{-1}(t)\mathrm{d}G_{a,0}^{-1}(s).
\end{aligned}$$

Note that if $r = 1$, $\sigma_a^0(k_a) \to \sqrt{2}$, as $a \to \infty$.

Proof of Theorem 4.3

From Theorem 4.2, Lemmas 4.5, 4.6, 4.7, and Slutsky's theorem, it follows that

$$\begin{aligned}
\widetilde{T}_L(k_a) &= \frac{\widehat{T}_L(k_a) - a\mu_a^0(k_a)}{\sqrt{a}\sigma_a^0(k_a)} \\
&= s^2 \frac{\widehat{\sigma}_a(k_a)}{\sigma_a^0(k_a)} \widehat{T}_L^*(k_a) + \frac{\sqrt{a}\,(s^2\widehat{\mu}_a(k_a) - \mu_a^0(k_a))}{\sigma_a^0(k_a)} \\
&= s^2 \frac{\widehat{\sigma}_a(k_a)}{\sigma_a^0(k_a)} \widehat{T}_L^*(k_a) + s^2 \frac{\sqrt{a}\,(\widehat{\mu}_a(k_a) - \mu_a^0(k_a))}{\sigma_a^0(k_a)} + \frac{\mu_a^0(k_a)}{\sigma_a^0(k_a)} \sqrt{a}\left(s^2 - 1\right) \\
&\overset{d}{\to} N\left(0, 1 + \frac{2\mu_r^2}{\sigma_r^2(n-1)}\right), \quad \text{as} \quad a \to \infty.
\end{aligned} \tag{B.14}$$

This completes the proof.

B.8 Proof of Proposition 5.1

Part (a): This result is obtained in page 687 of Fan (1996).

Part (b): The $\mathbf{\Gamma} = (\gamma_{j,k})$ is an $n_{bin} \times n_{bin}$ orthonormal matrix, and the transformed vector $\mathbf{X}_{\mathbf{\Gamma}}$ is decomposed as

$$\mathbf{X}_{\mathbf{\Gamma}} = \mathbf{X} + \mathbf{T},$$

with components $X_j = \sum_{k=1}^{n_{bin}} \gamma_{j,k}\epsilon_k$ and $T_j = \sum_{k=1}^{n_{bin}} \gamma_{j,k} V_k$. We note that $X_j \sim$ i.i.d. $N(0, 1)$ and

$$\max_{1 \le j \le n_{bin}} |T_j| \le \max_{1 \le j \le n_{bin}} \left\{ \sum_{k=1}^{n_{bin}} |\gamma_{j,k}| \max_{1 \le i \le n_{bin}} |V_i| \right\} \le n_{bin}^{1/2} \max_{1 \le i \le n_{bin}} |V_i|,$$

where the last inequality is justified by the Cauchy-Schwartz inequality. Since

$$\max_{1 \le i \le n_{bin}} |V_i| = O_r \left\{ (n^{-1} n_{bin} \log^2 n)^{1/2} \right\},$$

we obtain this result.

Part (c): Set $\psi(x) = x$ and $t_j = j/n_{bin}$. Then we have

$$n^{-1}(M_j - \mu_j) = \int_{t_{j-1}}^{t_j} x\mathrm{d}\big(\widehat{G}_n(x) - G(x)\big), \quad j = 1, \ldots, n_{bin}. \tag{B.15}$$

Using integration by parts,

$$\begin{aligned}(B.15) = \quad & \big(\widehat{G}_n(t_j) - G(t_j)\big)t_j - \big(\widehat{G}_n(t_{j-1}) - G(t_{j-1})\big)t_{j-1} \\ & - \int_{t_{j-1}}^{t_j} \big(\widehat{G}_n(x) - G(x)\big)\mathrm{d}x.\end{aligned}$$

When we multiply both sides of (B.15) by n and use the Hungarian embedding of Komlòs, Major, and Tusnády (1975), we get

$$\begin{aligned}M_j - \mu_j = \quad & n^{1/2}B_n(G(t_j))t_j - n^{1/2}B_n(G(t_{j-1}))t_{j-1} \\ & - \int_{t_{j-1}}^{t_j} n^{1/2}B_n(G(x))\mathrm{d}x + O_r(\log n), \quad \text{for any } r > 0,\end{aligned}$$

where B_n is a Brownian bridge. Let $W_n(t)$ be the Brownian motion with $W_n(0) = 0$. Then, possibly by changing the probability space, $B_n(t) = W_n(t) - tW_n(1)$. Thus we have

$$\begin{aligned}& M_j - \mu_j \\ &= n^{1/2}\left\{t_jW_n(G(t_j)) - t_{j-1}W_n(G(t_{j-1})) - \int_{t_{j-1}}^{t_j} W_n(G(x))\mathrm{d}x\right\} \\ &\quad -n^{1/2}W_n(1)\left\{t_jG(t_j) - t_{j-1}G(t_{j-1}) - \int_{t_{j-1}}^{t_j} G(x)\mathrm{d}x\right\} + O_r(\log n) \\ &= n^{1/2}A_{j-1,j} - n^{-1/2}W_n(1)\mu_j + O_r(\log n),\end{aligned}$$

where $A_{j-1,j} = t_jW_n(G(t_j)) - t_{j-1}W_n(G(t_{j-1})) - \int_{t_{j-1}}^{t_j} W_n(G(x))\mathrm{d}x$. Since

$$E(A_{j-1,j}) = 0, \quad Var(A_{j-1,j}) = \frac{3j^2 - 3j + 1}{3n_{bin}^3}, \quad Cov(A_{j-1,j}, A_{k-1,k}) = 0,$$

for $j, k = 1, \ldots, n_{bin}$, $j \neq k$, and $A_{j-1,j}$ is normally distributed and independent, we obtain

$$M_j - \mu_j = n^{1/2} \left(\frac{3j^2 - 3j + 1}{3n_{bin}^3} \right)^{1/2} \epsilon_j - n^{-1/2} W_n(1) \mu_j + O_r(\log n),$$

where $\epsilon_j \sim$ i.i.d. $N(0,1)$, $j = 1, \ldots, n_{bin}$. Using the facts that

$$\mu_j = \frac{n(2j-1)}{2n_{bin}^2} \quad \text{and} \quad \frac{\sigma_j}{n^{1/2}\left[(3j^2 - 3j + 1)/(3n_{bin}^3)\right]^{1/2}} \to 1, \quad \text{as} \quad n, n_{bin} \to \infty,$$

we get the desired result.

Part (d): Applying the same argument of the proof of Proposition 5.1 (b), the transformed data $\widetilde{X}_{\mathbf{\Gamma},j}$ are represented as

$$\widetilde{X}_{\mathbf{\Gamma},j} = X_j + \widetilde{T}_j, \quad j = 1, \ldots, n_{bin},$$

where $X_j = \sum_{k=1}^{n_{bin}} \gamma_{j,k} \epsilon_k$ and $\widetilde{T}_j$ satisfy

$$\max_{1 \le j \le n_{bin}} |\widetilde{T}_j| \le n_{bin}^{1/2} \max_{1 \le j \le n_{bin}} |\widetilde{V}_j|.$$

From the fact that

$$\max_{1 \le j \le n_{bin}} \left(\frac{3j^2 - 3j + 1}{3} \right)^{1/2} |\widetilde{V}_j| = O_r \left\{ (n^{-1} n_{bin}^3 \log^2 n)^{1/2} \right\},$$

we have

$$\max_{1 \le j \le n_{bin}} |\widetilde{V}_j| = O_r \left\{ (n^{-1} n_{bin}^3 \log^2 n)^{1/2} \right\}.$$

This completes the proof.

B.9 Lemmas B.5, B.6, B.7 and Their Proofs

Lemma B.5. *Consider the sequence of integers $a_{n_{bin}}$ which satisfies (5.12).*

1. *Relation (5.10) holds.*

2. *Relation (5.11) holds.*

Proof of Lemma B.5

Part 1. In view of (5.8), the convergence in (5.10) will follow by showing that

$$E\left(\frac{1}{\sqrt{n_{bin}}\sigma_{n_{bin}}(k_{n_{bin}})}\sum_{j=n_{bin}-a_{n_{bin}}+1}^{n_{bin}-k_{n_{bin}}} Y_{j,n_{bin}}\right) \to 0, \quad \text{as } n_{bin}\to\infty, \tag{B.16}$$

and

$$Var\left(\frac{1}{\sqrt{n_{bin}}\sigma_{n_{bin}}(k_{n_{bin}})}\sum_{j=n_{bin}-a_{n_{bin}}+1}^{n_{bin}-k_{n_{bin}}} Y_{j,n_{bin}}\right) \to 0, \quad \text{as } n_{bin}\to\infty. \tag{B.17}$$

To prove (B.16), we firstly study the denominator, the square of which is:

$$\begin{aligned} n_{bin}\sigma^2_{n_{bin}}(k_{n_{bin}}) &= \sum_{j=1}^{n_{bin}} \alpha^2_{jn_{bin}}(k_{n_{bin}}) \\ &\geq \sum_{j=n_{bin}-k_{n_{bin}}+1}^{n_{bin}} \left(\frac{1}{n_{bin}-j+1}\sum_{i=j}^{n_{bin}} \tilde{H}'\left(\tilde{\nu}_{in_{bin}}\right)\right)^2 \\ &\geq \sum_{j=n_{bin}-k_{n_{bin}}+1}^{n_{bin}} \left(\tilde{H}'\left(\tilde{\nu}_{jn_{bin}}\right)\right)^2 && \text{(B.18)} \\ &\geq k_{n_{bin}}\left(\tilde{H}'\left(\tilde{\nu}_{n_{bin}-k_{n_{bin}}+1,n_{bin}}\right)\right)^2, && \text{(B.19)} \end{aligned}$$

since $\tilde{H}'$ is increasing. Next, using the approximation $E\left(Y_{j,n_{bin}}\right) \simeq \tilde{H}\left(\tilde{\nu}_{jn_{bin}}\right)$, and the fact that $\tilde{H}$ is also increasing, the expectation in the numerator is

$$\begin{aligned} E\left(\sum_{j=n_{bin}-a_{n_{bin}}+1}^{n_{bin}-k_{n_{bin}}} Y_{j,n_{bin}}\right) &\simeq \sum_{j=n_{bin}-a_{n_{bin}}+1}^{n_{bin}-k_{n_{bin}}} \tilde{H}\left(\tilde{\nu}_{jn_{bin}}\right) \\ &\leq \left(a_{n_{bin}}-k_{n_{bin}}\right)\tilde{H}\left(\tilde{\nu}_{n_{bin}-k_{n_{bin}},n_{bin}}\right) \\ &\simeq \left(a_{n_{bin}}-k_{n_{bin}}\right)F^{-1}\left(\frac{n_{bin}-k_{n_{bin}}}{n_{bin}+1}\right), && \text{(B.20)} \end{aligned}$$

where the last approximation follows from the approximation in relation (2.9) of Chernoff, Gastwirth, and Johns (1967), that is, $\tilde{\nu}_{jn_{bin}} \simeq \nu_{jn_{bin}} = -\log(1-j/(n_{bin}+1))$. Suppose first that $k_{n_{bin}}/n_{bin} \to 0$, as $n_{bin}\to\infty$. Then, using (B.19), (B.20) and, from $1-\sqrt{2/\pi}e^{-x/2} \leq F(x)$, the easily derived relation

$$F^{-1}\left(\frac{n_{bin}-k_{n_{bin}}}{n_{bin}+1}\right) \leq 2\log\left(\sqrt{\frac{2}{\pi}}\left(\frac{n_{bin}+1}{k_{n_{bin}}+1}\right)\right),$$

we have, for some positive number C,

$$\begin{aligned} &E\left(\frac{1}{\sqrt{n_{bin}}\sigma_{n_{bin}}(k_{n_{bin}})}\sum_{j=n_{bin}-a_{n_{bin}}+1}^{n_{bin}-k_{n_{bin}}} Y_{j,n_{bin}}\right) \\ &\leq C\frac{a_{n_{bin}}-k_{n_{bin}}}{\sqrt{k_{n_{bin}}}}\log\left(\frac{n_{bin}+1}{k_{n_{bin}}+1}\right) \to 0, \end{aligned}$$

as $n_{bin} \to \infty$, by condition (5.12). Next, suppose that $k_{n_{bin}}/n_{bin} \to r$, as $n_{bin} \to \infty$, for some $0 < r \leq 1$. Then, using (B.18) and (B.20),

$$\begin{aligned} &E\left(\frac{1}{\sqrt{n_{bin}}\sigma_{n_{bin}}(k_{n_{bin}})}\sum_{j=n_{bin}-a_{n_{bin}}+1}^{n_{bin}-k_{n_{bin}}} Y_{j,n_{bin}}\right) \\ &\leq \frac{1}{\sqrt{\frac{1}{n_{bin}}\sum_{j=n_{bin}-k_{n_{bin}}+1}^{n_{bin}}\left(\tilde{H}'\left(\tilde{\nu}_{jn_{bin}}\right)\right)^2}}\left(\frac{a_{n_{bin}}-k_{n_{bin}}}{\sqrt{n_{bin}}}\right)F^{-1}\left(\frac{n_{bin}-k_{n_{bin}}}{n_{bin}+1}\right) \\ &\to 0, \end{aligned}$$

as $n_{bin} \to \infty$, by condition (5.12) since, by the proof of Theorem 3.3,

$$\frac{1}{n_{bin}}\sum_{j=n_{bin}-k_{n_{bin}}+1}^{n_{bin}}\left\{\tilde{H}'\left(\tilde{\nu}_{jn_{bin}}\right)\right\}^2 \to \int_{1-r}^{1}\left\{\frac{(1-t)}{f(F^{-1}(t))}\right\}^2 \mathrm{d}t > 0, \quad \text{as} \quad n_{bin} \to \infty.$$

Thus (B.16) holds for both cases. To prove (B.17), write

$$\begin{aligned} &Var\left(\sum_{j=n_{bin}-a_{n_{bin}}+1}^{n_{bin}-k_{n_{bin}}} Y_{j,n_{bin}}\right) \\ &= \sum_{j=1}^{n_{bin}-a_{n_{bin}}}\left\{\frac{1}{n_{bin}-j+1}\sum_{i=n_{bin}-a_{n_{bin}}+1}^{n_{bin}-k_{n_{bin}}}\tilde{H}'\left(\tilde{\nu}_{in_{bin}}\right)\right\}^2 \end{aligned}$$

$$
\begin{aligned}
&+\sum_{j=n_{bin}-a_{n_{bin}}+1}^{n_{bin}-k_{n_{bin}}}\left\{\frac{1}{n_{bin}-j+1}\sum_{i=j}^{n_{bin}-k_{n_{bin}}}\tilde{H}'\left(\tilde{\nu}_{in_{bin}}\right)\right\}^2\\
&\leq\left\{\sum_{j=1}^{n_{bin}-a_{n_{bin}}}\left(\frac{1}{n_{bin}-j+1}\right)^2\right\}(a_{n_{bin}}-k_{n_{bin}})^2\left(\tilde{H}'\left(\tilde{\nu}_{n_{bin}-k_{n_{bin}},n_{bin}}\right)\right)^2\\
&+(a_{n_{bin}}-k_{n_{bin}})\left(\tilde{H}'\left(\tilde{\nu}_{n_{bin}-k_{n_{bin}},n_{bin}}\right)\right)^2.
\end{aligned}
$$

Using the further approximation

$$
\sum_{j=1}^{n_{bin}-a_{n_{bin}}}\left(\frac{1}{n_{bin}-j+1}\right)^2\leq\int_0^{n_{bin}-a_{n_{bin}}}\frac{1}{(n_{bin}-x+1)^2}\mathrm{d}x=\frac{n_{bin}-a_{n_{bin}}}{(a_{n_{bin}}+1)(n_{bin}+1)},
$$

we have

$$
\begin{aligned}
&Var\left(\sum_{j=n_{bin}-a_{n_{bin}}+1}^{n_{bin}-k_{n_{bin}}}Y_{j,n_{bin}}\right)\\
&\leq\left(\tilde{H}'\left(\tilde{\nu}_{n_{bin}-k_{n_{bin}},n_{bin}}\right)\right)^2(a_{n_{bin}}-k_{n_{bin}})\left\{\frac{(n_{bin}-a_{n_{bin}})(a_{n_{bin}}-k_{n_{bin}})}{(a_{n_{bin}}+1)(n_{bin}+1)}+1\right\}.
\end{aligned}
$$

Thus, using (B.19), we have

$$
\begin{aligned}
&Var\left(\frac{1}{\sqrt{n_{bin}}\sigma_{n_{bin}}(k_{n_{bin}})}\sum_{j=n_{bin}-a_{n_{bin}}+1}^{n_{bin}-k_{n_{bin}}}Y_{j,n_{bin}}\right)\\
&\leq\frac{a_{n_{bin}}-k_{n_{bin}}}{k_{n_{bin}}}\left\{\frac{(n_{bin}-a_{n_{bin}})(a_{n_{bin}}-k_{n_{bin}})}{(a_{n_{bin}}+1)(n_{bin}+1)}+1\right\}\to 0,
\end{aligned}
$$

as $n_{bin}\to\infty$, by condition (5.12). This completes the proof of part 1.

Part 2. Since $Y_{n_{bin}-a_{n_{bin}}+1,n_{bin}}<\cdots<Y_{n_{bin},n_{bin}}$ and $n_{bin}-k_{n_{bin}}+1>n_{bin}-a_{n_{bin}}+1$, the relation (5.11) is equivalent to

$$
Y_{(n_{bin}-k_{n_{bin}}+l)}\geq Y_{n_{bin}-a_{n_{bin}}+1,n_{bin}},\quad\forall l=1,2,\ldots,k_{n_{bin}}.
$$

Let $m_{bin}=n_{bin}-k_{n_{bin}}$. Then, by Remark 5.2, $n_{bin}-a_{n_{bin}}=m_{bin}-k_{n_{bin}}^{1/2-\delta}$ for some

$0 < \delta < 1/2$ and relation (5.11) is equivalent to

$$Y_{(m_{bin}+l)} \geq Y_{m_{bin}-k_{n_{bin}}^{1/2-\delta}+1,n_{bin}}, \quad \forall l = 1, 2, \ldots, k_{n_{bin}}. \tag{B.21}$$

Let $\gamma_{n_{bin}} = \max_j |W_j|$, where W_j is defined in (5.6). In Lemma B.6 (2), it will be shown that

$$X^2_{\mathbf{\Gamma},(m_{bin}+l)} \geq Y_{m_{bin}+l,n_{bin}} - \gamma_{n_{bin}}$$

holds for all $l = 1, \ldots, k_{n_{bin}}$. This leads to the following bound of the left hand side of (B.21):

$$Y_{(m_{bin}+l)} \geq X^2_{\mathbf{\Gamma},(m_{bin}+l)} - \gamma_{n_{bin}} \geq Y_{m_{bin}+l,n_{bin}} - 2\gamma_{n_{bin}}.$$

Therefore, we have

$$\begin{aligned} & P\left(Y_{m_{bin}+1,n_{bin}} - 2\gamma_{n_{bin}} \geq Y_{m_{bin}-k_{n_{bin}}^{1/2-\delta}+1,n_{bin}}\right) \\ & < P\left(Y_{m_{bin}+l,n_{bin}} - 2\gamma_{n_{bin}} \geq Y_{m_{bin}-k_{n_{bin}}^{1/2-\delta}+1,n_{bin}},\ \forall l\right) \\ & < P\left(Y_{(m_{bin}+l)} \geq Y_{m_{bin}-k_{n_{bin}}^{1/2-\delta}+1,n_{bin}},\ \forall l\right). \end{aligned}$$

Thus, it suffices to show

$$\begin{aligned} & P\left(Y_{m_{bin}+1,n_{bin}} - 2\gamma_{n_{bin}} \geq Y_{m_{bin}-k_{n_{bin}}^{1/2-\delta}+1,n_{bin}}\right) \\ = \ & P\left(Y_{m_{bin}+1,n_{bin}} - Y_{m_{bin}-k_{n_{bin}}^{1/2-\delta}+1,n_{bin}} \geq 2\gamma_{n_{bin}}\right) \to 1, \quad \text{as} \quad n_{bin} \to \infty, \end{aligned}$$

or

$$P\left(\sum_{j=m_{bin}-k_{n_{bin}}^{1/2-\delta}+2}^{m_{bin}+1} \left(Y_{j,n_{bin}} - Y_{j-1,n_{bin}}\right) \geq 2\gamma_{n_{bin}}\right) \to 1, \quad \text{as} \quad n_{bin} \to \infty. \tag{B.22}$$

The spacings $Y_{j,n_{bin}} - Y_{j-1,n_{bin}}$ may be expressed as

$$Y_{j,n_{bin}} - Y_{j-1,n_{bin}} = F^{-1}\left(1 - \exp(-E_{j,n_{bin}})\right) - F^{-1}\left(1 - \exp(-E_{j-1,n_{bin}})\right),$$

where $E_1, \ldots, E_{n_{bin}}$ are independent standard exponential random variables and $E_{1,n_{bin}} < \cdots < E_{n_{bin},n_{bin}}$ are the corresponding order statistics. The spacings $E_{j,n_{bin}} - E_{j-1,n_{bin}}$ are related to the spacings $Y_{j,n_{bin}} - Y_{j-1,n_{bin}}$ by

$$\begin{aligned} E_{j,n_{bin}} - E_{j-1,n_{bin}} &= -\log\left(1 - F(Y_{j,n_{bin}})\right) + \log\left(1 - F(Y_{j-1,n_{bin}})\right) \\ &= \left(Y_{j,n_{bin}} - Y_{j-1,n_{bin}}\right) \frac{f(\widetilde{Y}_j)}{1 - F(\widetilde{Y}_j)}, \quad \text{with} \quad Y_{j-1,n_{bin}} < \widetilde{Y}_j < Y_{j,n_{bin}}. \end{aligned}$$

Since $f(y)/(1 - F(y))$ is decreasing and the value is less than 1 for large y, relation (B.22) will follow from

$$P\left(\sum_{j=m_{bin}-k_{n_{bin}}^{1/2-\delta}+2}^{m_{bin}+1} \left(E_{j,n_{bin}} - E_{j-1,n_{bin}}\right) \geq 2\gamma_{n_{bin}} \right) \to 1, \quad \text{as} \quad n_{bin} \to \infty, \quad \text{(B.23)}$$

which is shown in Lemma B.7. This completes the proof of relation (5.11).

Lemma B.6. *Let* $X^2_{\mathbf{\Gamma},j} = Y_j + W_j$ *be defined in (5.6), let* $X^2_{\mathbf{\Gamma},(1)} < \cdots < X^2_{\mathbf{\Gamma},(n_{bin})}$ *be the ordered* $X^2_{\mathbf{\Gamma},j}$*'s, and let* $Y_{(j)}$ *and* $W_{(j)}$ *correspond to* $X^2_{\mathbf{\Gamma},(j)}$*. Thus,* $X^2_{\mathbf{\Gamma},(j)} = Y_{(j)} + W_{(j)}$*. Further, let* $Y_{1,n_{bin}} < \cdots < Y_{n_{bin},n_{bin}}$ *be the ordered* Y_j*'s, and let* $X^2_{\mathbf{\Gamma},j,n_{bin}}$ *and* $W_{j,n_{bin}}$ *correspond to* $Y_{j,n_{bin}}$*. Thus,* $X^2_{\mathbf{\Gamma},j,n_{bin}} = Y_{j,n_{bin}} + W_{j,n_{bin}}$*. Then*

1. $X^2_{\mathbf{\Gamma},(l)} \leq Y_{l,n_{bin}} + \gamma_{n_{bin}}$, $l = 1, \ldots, n_{bin}$, *where* $\gamma_{n_{bin}} = \max_j |W_j|$.
2. $X^2_{\mathbf{\Gamma},(l)} \geq Y_{l,n_{bin}} - \gamma_{n_{bin}}$, $l = 1, \ldots, n_{bin}$, *where* $\gamma_{n_{bin}} = \max_j |W_j|$.

Proof of Lemma B.6

Part 1. When $l = 1$, we consider two cases.

Case 1: Suppose that $X^2_{\mathbf{\Gamma},(1)} = X^2_{\mathbf{\Gamma},1,n_{bin}}$. Then, $X^2_{\mathbf{\Gamma},(1)} = Y_{1,n_{bin}} + W_{1,n_{bin}} \leq Y_{1,n_{bin}} + \gamma_{n_{bin}}$.

Case 2: Suppose that $X^2_{\mathbf{\Gamma},(j)} = X^2_{\mathbf{\Gamma},1,n_{bin}}$ for any $j \geq 2$. Since $X^2_{\mathbf{\Gamma},(1)} \leq X^2_{\mathbf{\Gamma},(j)}$, we have $X^2_{\mathbf{\Gamma},(1)} \leq Y_{1,n_{bin}} + W_{1,n_{bin}} \leq Y_{1,n_{bin}} + \gamma_{n_{bin}}$.

When $l \geq 2$, we also consider two cases.

Case 1: Suppose that $X^2_{\mathbf{\Gamma},(l)} = X^2_{\mathbf{\Gamma},m,n_{bin}} = Y_{m,n_{bin}} + W_{m,n_{bin}}$, for some $m \leq l$. Then $X^2_{\mathbf{\Gamma},(l)} \leq Y_{l,n_{bin}} + W_{m,n_{bin}} \leq Y_{l,n_{bin}} + \gamma_{n_{bin}}$.

Case 2: Suppose that $X^2_{\mathbf{\Gamma},(l)} = X^2_{\mathbf{\Gamma},m,n_{bin}} = Y_{m,n_{bin}} + W_{m,n_{bin}}$ for some $m > l$. Then

(a) If $X^2_{\mathbf{\Gamma},(q)} = X^2_{\mathbf{\Gamma},l,n_{bin}} = Y_{l,n_{bin}} + W_{l,n_{bin}}$ with $q < l$, we would have that for some $j > l$ and some $i < l$ the equality $X^2_{\mathbf{\Gamma},(j)} = X^2_{\mathbf{\Gamma},i,n_{bin}} = Y_{i,n_{bin}} + W_{i,n_{bin}}$ would hold. It is justified by the reason that $\sharp\{i < l : X^2_{\mathbf{\Gamma},i,n_{bin}} \leq X^2_{\mathbf{\Gamma},(l)}\}$ is at most $(l-2)$. Thus we have $X^2_{\mathbf{\Gamma},(l)} \leq X^2_{\mathbf{\Gamma},(j)} = Y_{i,n_{bin}} + W_{i,n_{bin}} \leq Y_{l,n_{bin}} + W_{i,n_{bin}} \leq Y_{l,n_{bin}} + \gamma_{n_{bin}}$.

(b) If $X^2_{\mathbf{\Gamma},(q)} = X^2_{\mathbf{\Gamma},l,n_{bin}} = Y_{l,n_{bin}} + W_{l,n_{bin}}$ with $q > l$, we have $X^2_{\mathbf{\Gamma},(l)} \leq X^2_{\mathbf{\Gamma},(q)} = Y_{l,n_{bin}} + W_{l,n_{bin}} \leq Y_{l,n_{bin}} + \gamma_{n_{bin}}$.

Part 2. When $l = 1$, we consider two cases.

Case 1: Suppose that $X^2_{\mathbf{\Gamma},(1)} = X^2_{\mathbf{\Gamma},1,n_{bin}}$. Then, $X^2_{\mathbf{\Gamma},(1)} = Y_{1,n_{bin}} + W_{1,n_{bin}} \geq Y_{1,n_{bin}} - \gamma_{n_{bin}}$.

Case 2: Suppose that $X^2_{\mathbf{\Gamma},(1)} = X^2_{\mathbf{\Gamma},j,n_{bin}}$ for any $j \geq 2$. Then, $X^2_{\mathbf{\Gamma},(1)} = Y_{j,n_{bin}} + W_{j,n_{bin}} \geq Y_{1,n_{bin}} + W_{j,n_{bin}} \geq Y_{1,n_{bin}} - \gamma_{n_{bin}}$.

When $l \geq 2$, we also consider two cases.

Case 1: Suppose that $X^2_{\mathbf{\Gamma},(l)} = X^2_{\mathbf{\Gamma},m,n_{bin}} = Y_{m,n_{bin}} + W_{m,n_{bin}}$, for some $m < l$. Then

(a) If $X^2_{\mathbf{\Gamma},(q)} = X^2_{\mathbf{\Gamma},l,n_{bin}}$ with $q < l$, we have $X^2_{\mathbf{\Gamma},(l)} \geq X^2_{\mathbf{\Gamma},(q)} = Y_{l,n_{bin}} + W_{l,n_{bin}} \geq Y_{l,n_{bin}} - \gamma_{n_{bin}}$.

(b) If $X^2_{\mathbf{\Gamma},(q)} = X^2_{\mathbf{\Gamma},l,n_{bin}}$ with $q > l$, it can be shown by contradiction that $X^2_{\mathbf{\Gamma},(i)} = X^2_{\mathbf{\Gamma},j,n_{bin}}$ holds for some $j > l$ and some $i < l$. Indeed, if this statement does not hold we would have that $\forall\ i < l$ the equality $X^2_{\mathbf{\Gamma},(i)} = X^2_{\mathbf{\Gamma},j,n_{bin}}$ would hold for some $j < l$, which contradicts the assumption of Case 1. It follows that $X^2_{\mathbf{\Gamma},(l)} \geq X^2_{\mathbf{\Gamma},(i)} = Y_{j,n_{bin}} + W_{j,n_{bin}} \geq Y_{l,n_{bin}} + W_{j,n_{bin}} \geq Y_{l,n_{bin}} - \gamma_{n_{bin}}$.

Case 2: Suppose that $X^2_{\mathbf{\Gamma},(l)} = X^2_{\mathbf{\Gamma},m,n_{bin}} = Y_{m,n_{bin}} + W_{m,n_{bin}}$ for some $m \geq l$. Then $X^2_{\mathbf{\Gamma},(l)} \geq Y_{l,n_{bin}} + W_{m,n_{bin}} \geq Y_{l,n_{bin}} - \gamma_{n_{bin}}$.

Lemma B.7. *Let $E_1, \ldots, E_{n_{bin}}$ be independent standard exponential random variables and let $E_{1,n_{bin}} < \cdots < E_{n_{bin},n_{bin}}$ be the corresponding order statistics. Then, under the assumptions of Theorem 5.2, relation (B.23) holds.*

Proof of Lemma B.7

Set $E_{0,n_{bin}} = 0$ and let $D_j = E_{j,n_{bin}} - E_{j-1,n_{bin}}$, $j = 1, \ldots, n_{bin}$. It suffices to show that $\forall \epsilon > 0 \; \exists \; n_0(\epsilon)$ such that

$$P\left(\sum_{j=m_{bin}-k_{n_{bin}}^{1/2-\delta}+2}^{m_{bin}+1} D_j \geq 2\gamma_{n_{bin}} \right) > 1-\epsilon, \quad \forall \; n_{bin} > n_0(\epsilon).$$

From the fact that the sequence $\{(n_{bin} - j + 1)D_j, \; j = 1, \ldots, n_{bin}\}$ is i.i.d. from the standard exponential distribution (Pyke (1965)), it follows that

$$\frac{\sum_{j=m_{bin}-k_{n_{bin}}^{1/2-\delta}+2}^{m_{bin}+1} (n_{bin} - j + 1)D_j - k_{n_{bin}}^{1/2-\delta}}{k_{n_{bin}}^{1/4-\delta/2}} \xrightarrow{d} N(0,1),$$

as $k_{n_{bin}}$ tends to infinity with n_{bin}. Next, according to Remark 5.3, we have that $\forall \epsilon > 0 \; \exists \; n_0(\epsilon)$ such that

$$P\left(2\gamma_{n_{bin}}(k_{n_{bin}}^{1/2+\delta} + 1) < 1\right) > 1-\epsilon, \quad \forall \; n_{bin} > n_0(\epsilon). \tag{B.24}$$

Let now $A_{n_{bin}} = \left[2\gamma_{n_{bin}}(k_{n_{bin}}^{1/2+\delta} + 1) < 1\right]$, and set $P_{A_{n_{bin}}} = P(E \cap A_{n_{bin}})$ for any measurable set E. Thus,

$$\begin{aligned}
&P\left(\sum_{j=m_{bin}-k_{n_{bin}}^{1/2-\delta}+2}^{m_{bin}+1} D_j \geq 2\gamma_{n_{bin}} \right) \\
&\geq P_{A_{n_{bin}}}\left(\sum_{j=m_{bin}-k_{n_{bin}}^{1/2-\delta}+2}^{m_{bin}+1} D_j \geq 2\gamma_{n_{bin}} \right) \\
&= P_{A_{n_{bin}}}\left(\frac{\sum_{j=m_{bin}-k_{n_{bin}}^{1/2-\delta}+2}^{m_{bin}+1} \left(k_{n_{bin}} + k_{n_{bin}}^{1/2-\delta}\right) D_j - k_{n_{bin}}^{1/2-\delta}}{k_{n_{bin}}^{1/4-\delta/2}} \geq \frac{2\gamma_{n_{bin}}\left(k_{n_{bin}} + k_{n_{bin}}^{1/2-\delta}\right) - k_{n_{bin}}^{1/2-\delta}}{k_{n_{bin}}^{1/4-\delta/2}} \right) \\
&\geq P_{A_{n_{bin}}}\left(\frac{\sum_{j=m_{bin}-k_{n_{bin}}^{1/2-\delta}+2}^{m_{bin}+1} (n_{bin} - j + 1) D_j - k_{n_{bin}}^{1/2-\delta}}{k_{n_{bin}}^{1/4-\delta/2}} \geq \frac{2\gamma_{n_{bin}}\left(k_{n_{bin}} + k_{n_{bin}}^{1/2-\delta}\right) - k_{n_{bin}}^{1/2-\delta}}{k_{n_{bin}}^{1/4-\delta/2}} \right) \\
&= P_{A_{n_{bin}}}\left(\frac{\sum_{j=m_{bin}-k_{n_{bin}}^{1/2-\delta}+2}^{m_{bin}+1} (n_{bin} - j + 1) D_j - k_{n_{bin}}^{1/2-\delta}}{k_{n_{bin}}^{1/4-\delta/2}} \geq -k_{n_{bin}}^{1/4-\delta/2} \left\{1 - 2\gamma_{n_{bin}}\left(k_{n_{bin}}^{1/2+\delta} + 1\right)\right\} \right) \\
&\to P(A_{n_{bin}}), \quad \text{as} \quad n_{bin} \to \infty.
\end{aligned}$$

This, together with (B.24), completes the proof.

B.10 Lemmas B.8, B.9 and Their Proofs

Lemma B.8. *Consider the sequence of integers $a_{n_{bin}}$ which satisfies (5.12). Then relations (5.10) and (5.11) are true under the condition $n^{-1/2}n_{bin}^{5/2+\delta}\log^{3/2} n \to 0$, some $\delta > 0$. Thus, the first term on the right hand side of (5.15) converges to the standard normal distribution.*

Proof of Lemma B.8

Using the same argument as in the proof of Lemma B.5 (1), relation (5.10) holds. To prove relation (5.11), it suffices to show that

$$P\left(\sum_{j=m_{bin}-k_{n_{bin}}^{1/2-\delta}+2}^{m_{bin}+1}\left(E_{j,n_{bin}}-E_{j-1,n_{bin}}\right)\geq 2\widetilde{\gamma}_{n_{bin}}\right)\to 1, \quad \text{as} \quad n_{bin}\to\infty, \quad \text{(B.25)}$$

where $m_{bin} = n_{bin} - k_{n_{bin}}$, $E_1, \ldots, E_{n_{bin}}$ are independent standard exponential random variables, $E_{1,n_{bin}} < \cdots < E_{n_{bin},n_{bin}}$ are the corresponding order statistics, and $\widetilde{\gamma}_{n_{bin}} = \max_j |\widetilde{W_j}|$. It is shown in Lemma B.9, so that the proof of relation (5.11) is completed. Thus, using (5.8), it is easily seen that the first term on the right hand side of (5.15) converges in distribution to the standard normal.

Lemma B.9. *Let $E_1, \ldots, E_{n_{bin}}$ be independent standard exponential random variables and let $E_{1,n_{bin}} < \cdots < E_{n_{bin},n_{bin}}$ be the corresponding order statistics. Then, under the assumptions of Theorem 5.3, relation (B.25) holds.*

Proof of Lemma B.9

Set $E_{0,n_{bin}} = 0$ and let $D_j = E_{j,n_{bin}} - E_{j-1,n_{bin}}$, $j = 1, \ldots, n_{bin}$. It suffices to show that $\forall \epsilon > 0 \; \exists \; n_0(\epsilon)$ such that

$$P\left(\sum_{j=m_{bin}-k_{n_{bin}}^{1/2-\delta}+2}^{m_{bin}+1} D_j \geq 2\widetilde{\gamma}_{n_{bin}}\right) > 1-\epsilon, \quad \forall \; n_{bin} > n_0(\epsilon).$$

From the fact that the sequence $\{(n_{bin} - j + 1)D_j, \; j = 1, \ldots, n_{bin}\}$ is i.i.d. from the standard exponential distribution (Pyke (1965)), it follows that

$$\frac{\sum_{j=m_{bin}-k_{n_{bin}}^{1/2-\delta}+2}^{m_{bin}+1}(n_{bin}-j+1)D_j - k_{n_{bin}}^{1/2-\delta}}{k_{n_{bin}}^{1/4-\delta/2}} \xrightarrow{d} N(0,1),$$

as $k_{n_{bin}}$ tends to infinity with n_{bin}. Next, according to Remark 5.4, we have that $\forall\epsilon > 0\ \exists\ n_0(\epsilon)$ such that

$$P\left(2\widetilde{\gamma}_{n_{bin}}(k_{n_{bin}}^{1/2+\delta}+1) < 1\right) > 1-\epsilon, \quad \forall\ n_{bin} > n_0(\epsilon). \tag{B.26}$$

Let now $\widetilde{A}_{n_{bin}} = \left[2\widetilde{\gamma}_{n_{bin}}(k_{n_{bin}}^{1/2+\delta}+1) < 1\right]$, and set $P_{\widetilde{A}_{n_{bin}}} = P(E \cap \widetilde{A}_{n_{bin}})$ for any measurable set E. Thus,

$$\begin{aligned}
&P\left(\sum_{j=m_{bin}-k_{n_{bin}}^{1/2-\delta}+2}^{m_{bin}+1} D_j \geq 2\widetilde{\gamma}_{n_{bin}}\right)\\
&\geq P_{\widetilde{A}_{n_{bin}}}\left(\sum_{j=m_{bin}-k_{n_{bin}}^{1/2-\delta}+2}^{m_{bin}+1} D_j \geq 2\widetilde{\gamma}_{n_{bin}}\right)\\
&= P_{\widetilde{A}_{n_{bin}}}\left(\frac{\sum_{j=m_{bin}-k_{n_{bin}}^{1/2-\delta}+2}^{m_{bin}+1}\left(k_{n_{bin}}+k_{n_{bin}}^{1/2-\delta}\right)D_j - k_{n_{bin}}^{1/2-\delta}}{k_{n_{bin}}^{1/4-\delta/2}} \geq \frac{2\widetilde{\gamma}_{n_{bin}}\left(k_{n_{bin}}+k_{n_{bin}}^{1/2-\delta}\right)-k_{n_{bin}}^{1/2-\delta}}{k_{n_{bin}}^{1/4-\delta/2}}\right)\\
&\geq P_{\widetilde{A}_{n_{bin}}}\left(\frac{\sum_{j=m_{bin}-k_{n_{bin}}^{1/2-\delta}+2}^{m_{bin}+1}(n_{bin}-j+1)D_j - k_{n_{bin}}^{1/2-\delta}}{k_{n_{bin}}^{1/4-\delta/2}} \geq \frac{2\widetilde{\gamma}_{n_{bin}}\left(k_{n_{bin}}+k_{n_{bin}}^{1/2-\delta}\right)-k_{n_{bin}}^{1/2-\delta}}{k_{n_{bin}}^{1/4-\delta/2}}\right)\\
&= P_{\widetilde{A}_{n_{bin}}}\left(\frac{\sum_{j=m_{bin}-k_{n_{bin}}^{1/2-\delta}+2}^{m_{bin}+1}(n_{bin}-j+1)D_j - k_{n_{bin}}^{1/2-\delta}}{k_{n_{bin}}^{1/4-\delta/2}} \geq -k_{n_{bin}}^{1/4-\delta/2}\left\{1-2\widetilde{\gamma}_{n_{bin}}\left(k_{n_{bin}}^{1/2+\delta}+1\right)\right\}\right)\\
&\to P(\widetilde{A}_{n_{bin}}), \quad \text{as} \quad n_{bin} \to \infty.
\end{aligned}$$

This, together with (B.26), completes the proof.

Bibliography

AKRITAS, M. G. and PAPADATOS, N. (2004). Heteroscedastic One-Way ANOVA and Lack-of-Fit Tests. *Journal of the American Statistical Association.* **99** 368–382.

BERAN, R. (2004). Hybrid Shrinkage Estimators Using Penalty Bases For the Ordinal One-Way Layout. *The Annals of Statistics.* **32** 2532–2558.

BICKEL, P. J. (1967). Some Contributions to the Theory of Order Statistics. *Proceedings of the Fifth Berkeley Symposium on Mathematical Statististics and Probability.* Berkeley: University of California Press. **1** 575–591.

CHERNOFF, H., GASTWIRTH, J. L., and JOHNS, M. V. (1967). Asymptotic Distribution of Linear Combinations of Functions of Order Statistics with Applications to Estimation. *The Annals of Mathematical Statistics.* **38** 52–72.

DARLING, D. A. and ERDÖS, P. (1956). A Limit Theorem for the Maximum of Normalized Sums of Independent Random Variables. *Duke Mathematical Journal.* **23** 143–155.

DAVID, H. A. and NAGARAJA, H. N. (2003). *Order Statistics.* Wiley, New York.

DONOHO, D. L. and JOHNSTONE, I. M. (1994). Ideal Spatial Adaptation by Wavelet Shrinkage. *Biometrika.* **81** 425–455.

EFRON, B., TIBSHIRANI, R., STOREY, J. D., and TUSHER, V. (2001). Empirical Bayes Analysis of a Microarray Experiment. *Journal of the American Statistical Association.* **96** 1151–1160.

FAN, J. (1996). Test of Significance Based on Wavelet Thresholding and

Neyman's Truncation. *Journal of the American Statistical Association.* **91** 674–688.

Fan, J. and Huang, L. (2001). Goodness-of-Fit Tests for Parametric Regression Models. *Journal of the American Statistical Association.* **96** 640–652.

Fan, J. and Lin, S. K. (1998). Test of Significance When Data Are Curves. *Journal of the American Statistical Association.* **93** 1007–1021.

Galambos, J. (1995). *Advanced probability theory.* Marcel Dekker Inc., New York.

Hall, P. (1978). Representations and Limit Theorems for Extreme Value Distributions. *Journal of Applied Probability.* **15** 639–644.

Holst, L. (1972). Asymptotic Normality and Efficiency for Certain Goodness-of-Fit Tests. *Biometrika.* **59** 137–145.

Johnson, V. E. (2004). A Bayesian χ^2 Test for Goodness-of-Fit. *The Annals of Statistics.* **32** 2361–2384.

Johnstone, I. M. and Silverman B. W. (2004). Needles and Straw in Haystacks: Empirical Bayes Estimates of Possibly Sparse Sequences. *The Annals of Statistics.* **32** 1594–1649.

Koehler, K. J. and Larntz, K. (1980). An Empirical Investigation of Goodness-of-Fit Statistics for Sparse Multinomials. *Journal of the American Statistical Association.* **75** 336–344.

Komlòs, J., Major, P., and Tusnády, G. (1975). An Approximation of Partial Sums of Independent RV's, and the Sample DF. I. *Wahrscheinlichkeitsheorie verw. Gebiete.* **32** 111–131.

Mann, H. B. and Wald, A. (1942). On the Choice of the Number of Class Intervals in the Application of the Chi Square Test. *The Annals of Mathematical Statistics.* **13** 306–317.

Morris, C. (1975). Central Limit Theorems for Multinomial Sums. *The Annals of Statistics.* **3** 165–188.

Nagaraja, H. N. (1982). Some Nondegenerate Limit Laws for the Selection Differential. *The Annals of Statisitcs.* **10** 1306–1310.

Neyman, J. (1937). Smooth Test for Goodness of Fit. *Skandinavisk Aktuarietidskrift*. **20** 149–199.

Pearson, K. (1900). On the Criterion That and Given System of Deviations From the Probable in the Case of a Correlated System of Variables Is Such That It Can Be Reasonably Supposed to Have Arisen From Random Sampling. *Philosophical Magazine*. **50** 157–175.

Pyke, R. (1965). Spacings. *Journal of the Royal Statistical Society. Series B (Methodological)*. **27** 395–449.

Shorack, G. R. (1969). Asymptotic Normality of Linear Combinations of Functions of Order Statistics. *The Annals of Mathematical Statistics*. **40** 2041–2050.

Shorack, G. R. and Wellner, J. A. (1986). *Empirical Processes with Applications to Statistics*. Wiley, New York.

Simes, R. J. (1986). An improved Bonferroni procedure for multiple tests of significance. *Biometrika*. **73** 751–754.

Spokoiny, V. G. (1996). Adaptive Hypothesis Testing Using Wavelets. *The Annals of Statistics*. **24** 2477–2498.

Stigler, S. M. (1969). Linear Functions of Order Statistics. *The Annals of Mathematical Statistics*. **40** 770–788.

Stigler, S. M. (1973). The Asymptotic Distribution of the Trimmed Mean. *The Annals of Statistics*. **1** 472–477.

Storey, J. D. (2002). A Direct Approach to False Discovery Rates. *Journal of the Royal Statistical Society: Series B*. **64** 479–498.

Storey, J. D. (2003). The Positive False Discovery Rate: A Bayesian Interpretation and the q-value. *The Annals of Statistics*. **31** 2013–2035.

Vita

Min Hee Kim

Min Hee Kim was born in Seoul, South Korea, on July 25, 1979. In 1997 she entered Ewha Womans University and graduated summa cum laude in August 2000. She received an M.S. in Statistics from Ewha Womans University in August 2002, and a Ph.D. in Statistics from Pennsylvania State University in August 2010.