Majorization in Distributions on Hyperplanes and its Applications

垣内, 逸郎

https://doi.org/10.11501/3106930

出版情報：九州大学, 1995, 博士（数理学）, 論文博士
バージョン：

MAJORIZATION IN DISTRIBUTIONS ON HYPERPLANES AND ITS APPLICATIONS

垣　内　逸　郎

①

Majorization in Distributions on Hyperplanes and Its Applications

ITSURO KAKIUCHI

Kobe University

Preface

The concept of majorization concerns the diversity of the components of a vector. Although the basic idea of majorization is simple, it has been used as a useful and powerful tool for deriving inequalities in many areas of mathematics and statistics. Moreover, the derivation of an inequality by methods of majorization is often very helpful both for suggesting a unified theoretical framework and for providing a deeper understanding. Marshall and Olkin (1979) offer a comprehensive introduction on this topic.

In this paper we shall study majorization in multivariate distributions on hyperplanes with a location vector parameter and give their applications to testing problems. Let us consider the probability that a random vector having a multivariate distribution on a hyperplane with a location vector parameter takes the values in a certain set. Then, under some conditions it is shown that the probability can be compared through majorization order of vector parameters. The infimum and the supremum of the probability on a given parameter set can be evaluated by the largest of all parameters majorized by the parameter set and the smallest of all parameters majorizing the parameter set, respectively. Therefore we shall propose a general method to seek such parameters. By applying this method to tests for approximate equality of several location parameters we can construct robust tests and discuss their powers. These are done from both parametric and nonparametric situations.

Chapter 1 presents a basic theorem of majorization inequalities concering a radom vector with exchangeable components whose sum is constant. The result is used to obtain a stochastic ordering result for certain linear functions of order statistics. Then applications to the detection of outliers are discussed, and some unbiasedness properties of certain tests are given.

Chapter 2 presents vector parameters (called least favorable parameter configurations) which attain the infimum and the supremum of a probability depending on the parameter. Unless least favorable parameter configurations are available, vector parameters which give a lower and an upper bounds as close as posible are obtained by using majorization methods on hyperplanes and the majorization result in Chapter 1. To certain robust testing problems of location parameters, these results are used to determine the critical values of certain tests and to evaluate their powers.

Chapter 3 presents the asymptotic testing problems of k-sample approximate equality and gives k-sample robust rank tests with truncated scores for the problems. When underlying distributions vary in gross error neighborhoods by Rieder (1978), lower and upper bounds for limiting values of the probability that k-sample rank statistics take the values in a certain set are quite effectively obtained by using majorization methods on hyperplanes in Chapter 2. These bounds enable us to construct asymptotic level α rank tests and to give lower bounds of their asymptotic minimum powers for the problems. Based on these lower bounds robustness of k-sample rank tests is also studied.

Acknowledgments

I was extremely fortunate to have carried out this research under supervision of Professor Takashi Yanagawa. I am much indebted to him for all the help and direction he has given me.

To Professor Miyoshi Kimura, I wish to express my sincere thanks for his suggestion, discussion, thoughtfulness and especialy for his friendliness.

Contents

Appendix

CHAPTER

1

A Majorization Inequality for Distributions on Hyperplanes and Its Applications to Tests for Outliers

A theorem related to a random vector with exchangeable components whose sum is constant is established. The theorem is applied to obtain a stochastic ordering for certain linear functions of order statistics. It is seen that a number of test statistics for outliers of k location parameters are of these types. The unbiasedness properties of the tests based on such statistics are given. The theorem concering distributions on hyperplanes plays an important role similar to Marshall and Olkin's (1974) theorem.

1.1. Introduction.

The concept of majorization concerns the diversity of the components of a vector and it has been used as a fundamental tool for deriving inequalities in mathematics and statistics. A comprehensive introduction of the theory and applications of majorization is given in Marshall and Olkin (1979), and also Tong (1980).

Let $\mathbf{X}$ be a random vector taking values in the Euclidean k-space $\mathbf{R}^k$ and $\mathbf{D}$ a subset of $\mathbf{R}^k$. For each $\boldsymbol{\theta} \in \mathbf{R}^k$ let $\psi(\boldsymbol{\theta}) = \Pr\{\mathbf{X} + \boldsymbol{\theta} \in \mathbf{D}\}$ denote the probability of $\mathbf{X} + \boldsymbol{\theta}$

taking values in $\mathbf{D}$. Here $\boldsymbol{\theta}$ denotes a location parameter. Based on majorization theory Marshall and Olkin (1974) obtained a fundamental theorem that if the density f of $\mathbf{X}$ is Schur-concave and $\mathbf{D}$ is Schur-convex, then $\psi(\boldsymbol{\theta})$ is Schur-concave in $\boldsymbol{\theta}$. From this theorem, various inequalities can be constructed by evaluating values of functions at points ordered by majorization. The Marshall and Olkin's (1974) theorem weakens the condition of a special case of Mudholkar's (1966) theorem which is a generalization of Anderson's theorem (1955). Mudholkar assumes that f is permutation invariant and convex unimodal, and that $\mathbf{D}$ is permutation invariant and convex. The Marshall and Olkin's weaker condition has significant advantage of being checked more easily than the convexity.

In this chapter we shall consider majorization in distributions of a random vector $\mathbf{Z} = (Z_1, \ldots, Z_k)$ with exchangeable components whose sum is zero, and show that a Marshall-Olkin type theorem also holds for $\mathbf{Z}$ on the hyperplane $\boldsymbol{\Omega} = \{\boldsymbol{\mu} = (\mu_1, \ldots, \mu_k) \mid \sum_{i=1}^k \mu_i = 0\}$, the $k-1$ dimensional hyperplane of $\mathbf{R}^k$, that is, if the density of $(Z_1, \ldots, Z_{k-1})$ is Schur-concave and $\mathbf{D}$ is Schur-convex, then $\hat{\psi}(\boldsymbol{\mu}) = \Pr\{\mathbf{Z} + \boldsymbol{\mu} \in \mathbf{D}\}$ is Schur-concave function of $\boldsymbol{\mu}$ on $\boldsymbol{\Omega}$. An extension of the Marshall-Olkin and Kimura-Kakiuchi theorems (1989) (Theorem 1.2) from the point of view of linear transformations was given by Dean and Verducci (1990).

In Section 1.2, the Marshall-Olkin theorem in addition to definitions used throughout this paper is presented. A theorem related to a random vector with exchangeable components whose sum is constant is established. These two theorems are our starting point. From this theorem we can derive a stochastic ordering for certain linear functions of order statistics. It is seen that a number of test statistics which appear in testing problem of outliers of k location parameters are of these types.

In Section 1.3, the unbiasedness properties of the tests to the detection of outliers based on such statistics are given. The theorem concering distributions on hyperplanes enables us to compare their powers at points ordered by majorization. It plays an important part similar to Marshall and Olkin's (1974) theorem and gives broad applications of majorization inequalities in statistics.

1.2. A majorization inequality for distributions on hyperplanes.

Definitions. (i) A vector $\mathbf{x} \in \mathbf{R}^k$ is said to be majorized by a vector $\mathbf{y} \in \mathbf{R}^k$, written in symbol $\mathbf{x} \prec \mathbf{y}$, if

$$\sum_{i=1}^{k} x_i = \sum_{i=1}^{k} y_i \quad \text{and} \quad \sum_{i=1}^{r} x_{[i]} \le \sum_{i=1}^{r} y_{[i]}, \quad r = 1, \ldots, k-1, \tag{1.1}$$

where $x_{[1]} \ge \cdots \ge x_{[k]}$ and $y_{[1]} \ge \cdots \ge y_{[k]}$ denote the components of $\mathbf{x} = (x_1, \ldots, x_k)$ and $\mathbf{y} = (y_1, \ldots, y_k)$ in decreasing order.

(ii) A real valued function ψ is said to be Schur-concave (Schur-convex), if

$$\mathbf{x} \prec \mathbf{y} \Rightarrow \psi(\mathbf{x}) \ge (\le)\ \psi(\mathbf{y}).$$

A nonnegative function ψ on $\mathbf{R}^k$ is said to be centrally symmetric if $\psi(\mathbf{x}) = \psi(-\mathbf{x})$ for all $\mathbf{x}$. A nonnegative function ψ on $\mathbf{R}^k$ is said to be convex unimodal if for every $c > 0$, the set $\{\mathbf{x} \mid \psi(\mathbf{x}) \geq c\}$ is a centrally symmetric convex set. A nonnegative function ψ on $\mathbf{R}^k$ is said to be logconcave if for every $\mathbf{x}$, $\mathbf{y} \in \mathbf{R}^k$ and for every $\alpha \in (0, 1)$

$$\psi(\alpha\mathbf{x} + (1-\alpha)\mathbf{y}) \geq [\psi(\mathbf{x})]^\alpha [\psi(\mathbf{y})]^{1-\alpha}.$$

A nonnegative function ψ is said to be permutation invariant if $\psi(\mathbf{x}) = \psi(g\mathbf{x})$ for all permutations g and all $\mathbf{x} \in \mathbf{R}^k$.

(iii) A set $\mathbf{D}$ of $\mathbf{R}^k$ is said to be Schur-convex, if

$$\mathbf{y} \in \mathbf{D} \quad \text{and} \quad \mathbf{y} \succ \mathbf{x} \Rightarrow \mathbf{x} \in \mathbf{D}, \tag{1.2}$$

that is, if the indicator function of $\mathbf{D}$ is Schur-concave. A set $\mathbf{D} \subset \mathbf{R}^k$ is said to be centrally symmetric if $\mathbf{x} \in \mathbf{D} \Rightarrow -\mathbf{x} \in \mathbf{D}$. A set $\mathbf{D}$ is said to be shift invariant if $\mathbf{D} = \mathbf{D} + a\mathbf{1}_k$ with $\mathbf{1}_k = (1, \ldots, 1) \in \mathbf{R}^k$ and $a \in \mathbf{R}$. A set $\mathbf{D}$ is said to be permutation invariant if $g\mathbf{D} = \mathbf{D}$ for all permutation g.

Theorem 1.1 (Marshall and Olkin, 1974). *Suppose that $X_1, \ldots, X_k$ are exchangeable random variables with a joint density f that is Schur-concave. If $\mathbf{D} \subset \mathbf{R}^k$ is Schur-convex, then $\Pr\{\mathbf{X} + \boldsymbol{\theta} \in \mathbf{D}\}$ is a Schur-concave function of $\boldsymbol{\theta}$, where $\mathbf{X} = (X_1, \ldots, X_k)$ and $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_k)$.*

Theorem 1.1 was obtained by weakening the conditions for a special case of Mudholkar's (1966) generalization of Anderson's theorem (1955). Mudholkar (1966) assumes that f is permutation invariant and convex unimodal, and that $\mathbf{D}$ is permutation invariant and convex. We note that if f is permutation invariant and convex unimodal, then it is Schur-concave and that if $\mathbf{D}$ is permutation invariant and convex, then it is Schur-convex. As pointed out in Marshall and Olkin (1974), the weaker condition of Theorem 1.1 have the following advantages: It is often much easier to check than convexity, and the convexity condition is strong for certain applications.

The following result, which are useful in deriving various probability inequalities for distributions on hyperplanes, is derived from Theorem 1.1.

Theorem 1.2. *Let $Z_1, \ldots, Z_k$ ($k \geq 3$) be exchangeable random variables satisfying $\sum_{i=1}^k Z_i = 0$ such that $(Z_1, \ldots, Z_{k-1})$ has a Schur-concave density. Let $\mathbf{D} \subset \mathbf{R}^k$ be Schur-convex. Then, $\Pr\{\mathbf{Z} + \boldsymbol{\mu} \in \mathbf{D}\}$ is a Schur-concave function of $\boldsymbol{\mu}$, where $\mathbf{Z} = (Z_1, \ldots, Z_k)$ and $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_k)$.*

Proof. Let $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_k)$ and $\hat{\boldsymbol{\mu}} = (\hat{\mu}_1, \ldots, \hat{\mu}_k)$ be any vectors such that $\boldsymbol{\mu} \prec \hat{\boldsymbol{\mu}}$. Then it is sufficient to show $\Pr\{\mathbf{Z} + \boldsymbol{\mu} \in \mathbf{D}\} \geq \Pr\{\mathbf{Z} + \hat{\boldsymbol{\mu}} \in \mathbf{D}\}$. By Muirhead's result (Marshall and Olkin, 1979, p.21) $\boldsymbol{\mu}$ can be derived from $\hat{\boldsymbol{\mu}}$ by successive applications of a finite number of $T-$transformations. A $T-$transformation is a linear transformation whose matrix has the form $T = \lambda I + (1-\lambda)Q$, where $0 \le \lambda \le 1$, I denotes the identity matrix and Q denotes a permutation matrix that just interchanges two components. Note that the vector derived from $\hat{\boldsymbol{\mu}}$ by an application of a single $T-$transformation is different from $\hat{\boldsymbol{\mu}}$ in at most two components and majorized by $\hat{\boldsymbol{\mu}}$. Thus we can assume without loss of generality that $\boldsymbol{\mu}$ and $\hat{\boldsymbol{\mu}}$ differ in two components only. Since $k \ge 3$, there exists at least one common components of $\boldsymbol{\mu}$ and $\hat{\boldsymbol{\mu}}$. We can take $\mu_k = \hat{\mu}_k$ without loss of generality.

Let $h : \mathbf{R}^k \to \mathbf{R}^{k-1}$ be the projection defined by $h(x_1, \dots, x_k) = (x_1, \dots, x_{k-1})$. Then the restriction of h to $\boldsymbol{\Omega} = \{\mathbf{x} \mid \sum_{i=1}^k x_i = 0\}$ is a bijection from $\boldsymbol{\Omega}$ to $\mathbf{R}^{k-1}$. Letting $\mathbf{D}' = \mathbf{D} \cap \boldsymbol{\Omega}$, then we have

$$\Pr\{\mathbf{Z} + \boldsymbol{\mu} \in \mathbf{D}\} = \Pr\{h(\mathbf{Z}) + h(\boldsymbol{\mu}) \in h(\mathbf{D}')\}. \tag{1.3}$$

Now we shall show that the set $h(\mathbf{D}')$ is Schur-convex. Let $\mathbf{v} \in h(\mathbf{D}')$ and $\mathbf{u} \prec \mathbf{v}$. Then their uniquly exist $\mathbf{x} \in \boldsymbol{\Omega}$ and $\mathbf{y} \in \mathbf{D}'$ such that $\mathbf{u} = h(\mathbf{x})$ and $\mathbf{v} = h(\mathbf{y})$, i.e., $u_i = x_i$ and $v_i = y_i$, $i = 1, \dots, k-1$. Since $\sum_{i=1}^{k-1} u_i = \sum_{i=1}^{k-1} v_i$, we can see $x_k = y_k$. Thus $\mathbf{x} \prec \mathbf{y}$ follows from $\mathbf{u} \prec \mathbf{v}$. As $\mathbf{D}$ satisfies (1.2), we have $\mathbf{x} \in \mathbf{D}$ and hence $\mathbf{x} \in \mathbf{D}'$. This implies $\mathbf{u} \in h(\mathbf{D}')$.

Since the density function of $h(\mathbf{Z})$ is Schur-concave by the assumption, it is seen from Theorem 1.1 that the righthand side of (1.3) is Schur-concave in $h(\boldsymbol{\mu})$. It is also clear that $h(\boldsymbol{\mu}) \prec h(\hat{\boldsymbol{\mu}})$ follows from $\boldsymbol{\mu} \prec \hat{\boldsymbol{\mu}}$ and $\mu_k = \hat{\mu}_k$. Hence we obtain

$$\Pr\{\mathbf{Z} + \boldsymbol{\mu} \in \mathbf{D}\} \ge \Pr\{\mathbf{Z} + \hat{\boldsymbol{\mu}} \in \mathbf{D}\}.$$

This completes the proof of the theorem. □

Corollary 1.1. *Let* $\mathbf{Z}$ *and* $\mathbf{D}$ *be as given in Theorem 1.2 and let* V *be a positive random variable independent of* $\mathbf{Z}$*. Then* $\Pr\{(\mathbf{Z} + \boldsymbol{\mu})/V \in \mathbf{D}\}$ *is a Schur-concave function of* $\boldsymbol{\mu}$*.*

Proof. Since $\mathbf{Z}$ and V are independent, we have

$$\Pr\{(\mathbf{Z} + \boldsymbol{\mu})/V \in \mathbf{D}\} = \int \Pr\{\mathbf{Z} + \boldsymbol{\mu} \in v\mathbf{D}\}\, F(dv),$$

where F is the distribution function of V. It is easy to see that $v\mathbf{D}$ is Schur-convex for every $v \in \mathbf{R}$. Hence, from Theorem 1.2 it follows that $\Pr\{\mathbf{Z} + \boldsymbol{\mu} \in v\mathbf{D}\}$ is Schur-concave in $\boldsymbol{\mu}$ for every $v \in \mathbf{R}$. This implies Corollary 1.1 holds. □

Examples. The following is examples of Schur-convex sets $\mathbf{D}(\lambda)$ for each real λ.

$$\mathbf{D}_{1r}(\lambda) = \{\mathbf{x} \mid \sum_{i=1}^{r} x_{[i]} \le \lambda\}, \quad r = 1, \dots, k-1, \tag{1.4}$$

$$\mathbf{D}_{2r}(\lambda) = \{\mathbf{x} \mid \max(\sum_{i=1}^{r} x_{[i]}, -\sum_{i=k-r+1}^{k} x_{[i]}) \leq \lambda\}, \quad r = 1, \ldots, k-1, \tag{1.5}$$

$$\mathbf{D}_{3r_1r_2}(\lambda) = \{\mathbf{x} \mid \sum_{i=1}^{r_1} x_{[i]} - \sum_{i=k-r_2+1}^{k} x_{[i]} \leq \lambda\}, \tag{1.6}$$

$$r_1, r_2 = 1, \ldots, k-1, r_1 + r_2 \leq k-1,$$

$$\mathbf{D}_4(\lambda) = \{\mathbf{x} \mid \sum_{i=1}^{k} x_i^2 \leq \lambda\}, \tag{1.7}$$

where $\mathbf{x} = (x_1, \ldots, x_k)$ and $x_{[1]} \geq \ldots \geq x_{[k]}$ denotes the components of $\mathbf{x}$ in decreasing order.

Theorem 1.2 is applicable to obtain some results concering a stochastic ordering.

Corollary 1.2. *Let* $\mathbf{Z}$ *and* V *be as given in Corollary 1.1. Let* ϕ *be a measurable Schur-convex function on* $\mathbf{R}^k$. *Then,* $\Pr\{\phi(\mathbf{Z}+\boldsymbol{\mu}) \leq \lambda\}$ *and* $\Pr\{\phi((\mathbf{Z}+\boldsymbol{\mu})/V) \leq \lambda\}$ *are Schur-concave functions in* $\boldsymbol{\mu}$ *for a real number* λ.

Proof. Let $\mathbf{D} = \{\mathbf{u} \mid \phi(\mathbf{u}) \leq \lambda\}$. Then $\mathbf{D}$ is Schur-convex. By Theorem 1.2, $\Pr\{\mathbf{Z}+\boldsymbol{\mu} \in \mathbf{D}\} = \Pr\{\phi(\mathbf{Z}+\boldsymbol{\mu}) \leq \lambda\}$ is Schur-concave in $\boldsymbol{\mu}$. This fact and Corollary 1.1 prove the rest of the theorem. □

Corollary 1.2 says that if $\boldsymbol{\mu} \prec \hat{\boldsymbol{\mu}}$, then $\phi(\mathbf{Z}+\boldsymbol{\mu})$ and $\phi((\mathbf{Z}+\boldsymbol{\mu})/V)$ are stochastically smaller than $\phi(\mathbf{Z}+\hat{\boldsymbol{\mu}})$ and $\phi((\mathbf{X}+\hat{\boldsymbol{\mu}})/V)$, respectively.

The following theorem shows how the random vector $\mathbf{Z}$ on the hyperplane $\boldsymbol{\Omega}$ can be obtained from $\mathbf{X}$ on $\mathbf{R}^k$.

Theorem 1.3. *Suppose that* $X_1, \ldots, X_k (k \geq 3)$ *be exchangeable random variables with a joint density* f *and let* $Z_i = X_i - \bar{X}, i = 1, \ldots, k$. *Then the following statements hold.*

(i) $Z_1, \ldots, Z_k$ $(k \geq 3)$ *be exchangeable random variables with* $\sum_{i=1}^{k} Z_i = 0$.

(ii) *If* f *is Schur-concave, then* $(Z_1, \ldots, Z_{k-1})$ *has a Schur-concave density.*

(iii) *If* f *is logconcave, then* $(Z_1, \ldots, Z_{k-1})$ *has a logconcave density.*

(iv) *If* f *is centrally symmetric, then* $(Z_1, \ldots, Z_{k-1})$ *has a centrally symmetric density.*

Proof. (i) The assertion is obvious.

(ii) Let g be the density of $(Z_1, \ldots, Z_{k-1}, \bar{X})$ and h the density of $(Z_1, \ldots, Z_{k-1})$. Then

$$g(z_1, \ldots, z_{k-1}, \bar{x}) = Kf(l(z_1, \ldots, z_{k-1}, \bar{x})), \tag{1.8}$$

and

$$h(z_1, \dots, z_{k-1}) = \int g(z_1, \dots, z_{k-1}, \bar{x})\, d\bar{x}, \tag{1.9}$$

where $l(z_1, \dots, z_{k-1}, \bar{x}) = (z_1 + \bar{x}, \dots, z_{k-1} + \bar{x}, -\sum_{j=1}^{k-1} z_j + \bar{x})$ and K is a positive constant. Let $(z_1, \dots, z_{k-1}) \prec (z_1^*, \dots, z_{k-1}^*)$. Then for every $\bar{x}$ we have

$$l(z_1, \dots, z_{k-1}, \bar{x}) \prec l(z_1^*, \dots, z_{k-1}^*, \bar{x}).$$

Hence, it follows from the Schur-concavity of f that for every $\bar{x}$

$$g(z_1, \dots, z_{k-1}, \bar{x}) \geq g(z_1^*, \dots, z_{k-1}^*, \bar{x}).$$

This implies

$$h(z_1, \dots, z_{k-1}) \geq h(z_1^*, \dots, z_{k-1}^*).$$

Thus h is Schur-concave.

(iii) The logconcavity of g is easily obtained from the logconcavity of f and (1.8). Hence from Theorem 2.16 of Dharmadhikari and Joag-dev (1988) it follows that h is logconcave.

(iv) The assertion is an immediate consequence from (1.8) and (1.9). □

1.3. Applications to tests for outliers.

As applications of the previous results we are concerned with unbiased properties of testing problems of k location parameters. The following is a list of the problems and relevant test statistics.

Let $X_1, \dots, X_k$ be exchangeable random variables distributed with a joint density $\sigma^{-k} f((\mathbf{x} - \boldsymbol{\theta})/\sigma)$, where f is Schur-concave and $\boldsymbol{\theta} = (\theta_1, \dots, \theta_k)$. We are interested in testing the null hypothesis

$$\mathbf{H}_0: \quad \theta_i = \theta, \quad i = 1, \dots, k, \tag{1.10}$$

where θ is unknown.

(i) σ is unknown but an independent estimator of σ is available.

$$T_{1r} = \sum_{i=1}^{r}(X_{[i]} - \bar{X})/V, \quad r = 1, \dots, k-1, \tag{1.11}$$

$$T_{2r} = \max\left\{\sum_{i=1}^{r}(X_{[i]} - \bar{X})/V, - \sum_{i=k-r+1}^{k}(X_{[i]} - \bar{X})/V\right\}, \tag{1.12}$$
$$r = 1, \dots, k-1,$$

$$T_{3r_1r_2} = \sum_{i=1}^{r_1}(X_{[i]} - \bar{X})/V - \sum_{i=k-r_2+1}^{k}(X_{[i]} - \bar{X})/V, \tag{1.13}$$
$$r_1, r_2 = 1, \dots, k-1, \ r_1 + r_2 \leq k-1,$$

where V is an estimator of σ^2 independent of $X_1, \dots, X_k$, $X_{[i]}$ is the i–the order statistics of $X_1, \dots, X_k$ in decreasing order and $\bar{X} = k^{-1}\sum_{i=1}^{k} X_i$.

(ii) σ is known.

$$T'_{1r} = T_{1r} \quad \text{with } V \text{ replaced by } \sigma, \tag{1.14}$$
$$T'_{2r} = T_{2r} \quad \text{with } V \text{ replaced by } \sigma, \tag{1.15}$$
$$T'_{3r_1r_2} = T_{3r_1r_2} \quad \text{with } V \text{ replaced by } \sigma. \tag{1.16}$$

The statistics T_{1r} and T'_{1r} are used for r upper outliers. On the other hand, the remaining statistics are used for both-sided outliers. T_{2r} and T'_{2r} are statistics for r lower and r upper outliers. $T_{3r_1r_2}$ and $T'_{3r_1r_2}$ are statistics for r_1 lower and r_2 upper outliers (see Barnett and Lewis, 1984, section 6.3). In particular, when $r = r_1 = r_2 = 1$, these statistics reduce to

$$T_{11} = \max_{1\leq i\leq k}(X_i - \bar{X})/V, \quad T_{21} = \max_{1\leq i\leq k}|X_i - \bar{X}|/V \quad \text{and} \quad T_{311} = (X_{[1]} - X_{[k]})/V.$$

Consider the following test:

$$\varphi = 0, \ 1 \text{ according as } T \leq \lambda \text{ or } T > \lambda, \tag{1.17}$$

where T is a statistic and λ is a real constant. We call such φ a test based on T.

Theorem 1.4. *Each of the tests of* $\mathbf{H}_0$ *based on the statistics (1.11)-(1.17) is unbiased for the alternative* $\mathbf{K}_1 : \theta_i = \theta$ *for at least one* i $(1 \leq i \leq k)$.

Proof. Let

$$Z_i = (X_i - \bar{X}) - (\theta_i - \bar{\theta}), \quad i = 1, \dots, k,$$

where $\bar{\theta} = (1/k)\sum_{i=1}^{k}\theta_i$. It is clear from Theorem 1.3 that $Z_1, \ldots, Z_k$ satisfy all the assumptions in Theorem 1.2. For each $\boldsymbol{\theta}$ let $\mu_i(\boldsymbol{\theta}) = \theta_i - \bar{\theta}$, $i = 1, \ldots, k$. Then we have

$$E_{\boldsymbol{\theta}}\varphi_i(\mathbf{X}) = \Pr\{T_i > \lambda\} = 1 - \Pr\{(\mathbf{Z} + \boldsymbol{\mu}(\boldsymbol{\theta}))/V \in \mathbf{D}_i(\lambda)\}, \quad i = 1, 2, 3, \tag{1.18}$$

where φ_i, $i = 1, 2, 3$, denote the tests based on T_i given by (1.11), (1.12) and (1.13), $\mathbf{D}_i$, $i = 1, 2, 3$, are given by (1.4), (1.5) and (1.6), $\mathbf{Z} = (Z_1, \ldots, Z_k)$ and $\boldsymbol{\mu}(\boldsymbol{\theta}) = (\mu_1(\boldsymbol{\theta}), \ldots, \mu_k(\boldsymbol{\theta}))$. By Theorem 1.2, the power functions (1.18) of the tests φ_i based on T_i are Schur-convex in $\boldsymbol{\mu}(\boldsymbol{\theta})$. It is easy to see that $\sum_{i=1}^{k}\mu_i(\boldsymbol{\theta}) = 0$ holds for every $\boldsymbol{\theta} \in \mathbf{R}^k$ and $\boldsymbol{\mu}(\boldsymbol{\theta}) = \mathbf{0}$, which is minimum, is equivalent to that all θ_i are equal. Thus the assertion holds for the test φ_i, $i = 1, 2, 3$. The assertion for the tests based on T_i' given by (1.14), (1.15) and (1.16) are shown in the same way. □

Remark 1.1. Let us consider the tests of $\mathbf{H}_0$ based on the types of the statistics $\phi(X_1 - \bar{X}, \ldots, X_k - \bar{X})$ and $\phi((X_1 - \bar{X})/V, \ldots, (X_k - \bar{X})/V)$ according as σ is known or unknown, where ϕ is a Schur-convex function. Then, in the same way as Theorem 1.4 we can easily get the unbiasedness properties of a variety of tests which are given by changing ϕ.

CHAPTER

2

Majorization Methods on Hyperplanes and Their Applications to Robust Testing

Let $\mathbf{Z}$ be a $k(\geq 3)$ dimensional random vector with exchangeable components whose sum is zero, $\mathbf{D}$ a Schur-convex subset of the Euclidean k-space $\mathbf{R}^k$, and $\mathbf{\Omega}$ the $k-1$ dimensional hyperplane of $\mathbf{R}^k$ consisting of all vector parameters whose components sum up to zero. Let $\hat{\psi}(\boldsymbol{\mu}), \boldsymbol{\mu} \in \mathbf{\Omega}$, denote the probability of $\mathbf{Z} + \boldsymbol{\mu}$ taking values in D. The present chapter derives parameters at which $\hat{\psi}$ attains its infimum and supremum on a given $\mathbf{\Gamma} \subset \mathbf{\Omega}$ or takes their approximate values. This is achieved by using majorization methods. The results are applied to robust testing of several location parameters.

1. Introduction

In this chapter we shall further develope the study of majorization in distributions on hyperplanes which was started by Kimura and Kakiuchi (1989).

Let $\mathbf{Z} = (Z_1, \ldots, Z_k)$ be a random vector with exchangeable components whose sum is zero, $\mathbf{D}$ a subset of the Euclidean k-space $\mathbf{R}^k$ and $\mathbf{\Omega} = \{\boldsymbol{\mu} = (\mu_1, \ldots, \mu_k) \mid \sum_{i=1}^{k} \mu_i = 0\}$ the $k-1$ dimensional hyperplane of $\mathbf{R}^k$. Then we showed in Section 1.2 of Chapter 1 that if the density of $(Z_1, \ldots, Z_{k-1})$ is Schur-concave and $\mathbf{D}$ is Schur-convex, then $\hat{\psi}(\boldsymbol{\mu}) = \Pr\{\mathbf{Z} + \boldsymbol{\mu} \in \mathbf{D}\}$ is a Schur-concave function of $\boldsymbol{\mu}$ on $\boldsymbol{\Omega}$.

For a given subset $\boldsymbol{\Gamma}$ of $\boldsymbol{\Omega}$ we are interested in evaluating the infimum and supremum of $\hat{\psi}$ on $\boldsymbol{\Gamma}$. To this end we need parameters $\boldsymbol{\mu}^{LFC} \in \boldsymbol{\Omega}$ (called least favorable configurations) at which $\hat{\psi}$ attains its infimum and supremum on $\boldsymbol{\Gamma}$. Unless $\boldsymbol{\mu}^{LFC}$ are available, we shall instead want parameters $\boldsymbol{\mu}^M$ and $\boldsymbol{\mu}^s$ in $\boldsymbol{\Omega}$ at which $\hat{\psi}$ takes a lower and an upper bound as close as possible to the infimum and supremum on $\boldsymbol{\Gamma}$, respectively. The purpose of this chapter is to seek these parameters $\boldsymbol{\mu}^{LFC}, \boldsymbol{\mu}^M$ and $\boldsymbol{\mu}^s$ by using majorization methods on hyperplanes and to give their applications to robust testing.

In Section 2.2, for any general $\boldsymbol{\Gamma}$ we shall propose a new majorization method for constructing $\boldsymbol{\mu}^M$ and $\boldsymbol{\mu}^s$. The parameters $\boldsymbol{\mu}^M$ and $\boldsymbol{\mu}^s$ are the best ones in the sense that they can no longer be improved by majorization. This majorization method is very useful and powerful when $\boldsymbol{\mu}^{LFC}$ are not available.

In Section 2.3, we shall treat six special parameter forms of $\boldsymbol{\Gamma}$ to seek $\boldsymbol{\mu}^{LFC}$ or their candidates. In order to obtain $\boldsymbol{\mu}^{LFC}$ we shall make use of Theorem 1.2 in Section 1.2.

Let $\mathbf{X}$ be a random vector taking values in the Euclidean k-space $\mathbf{R}^k$. For each $\boldsymbol{\theta} \in \mathbf{R}^k$ let $\psi(\boldsymbol{\theta}) = \Pr\{\mathbf{X} + \boldsymbol{\theta} \in \mathbf{D}\}$. Marshall and Olkin (1974) obtained a fundamental theorem that if the density f of $\mathbf{X}$ is Schur-concave and $\mathbf{D}$ is Schur-convex, then $\psi(\boldsymbol{\theta})$ is Schur-concave in $\boldsymbol{\theta}$.

In Section 2.4, under the shift invariance of $\mathbf{D}$ we shall consider majorization in non-singular distributions, that is, the majorization problem with $\mathbf{Z}$, $\boldsymbol{\Gamma}$ and $\hat{\psi}$ replaced by $\mathbf{X}$, $\boldsymbol{\Theta}$ and ψ, respectively, where $\boldsymbol{\Theta}$ is a subset of $\mathbf{R}^k$. By the arguments similar to Sections 2.2 we shall get $\boldsymbol{\theta}^M$ and $\boldsymbol{\theta}^s$ for a general $\boldsymbol{\Theta}$. The problem of seeking least favorable configurations $\boldsymbol{\theta}^{LFC}$ for special sets $\boldsymbol{\Theta}$ was treated by Giani and Finner (1991), Chen, Lam and Xiong (1993), Finner and Roters (1993), and Kakiuchi and Kimura (1994). In order to obtain $\boldsymbol{\theta}^{LFC}$ they applied the Mudholkar's theorem or the Krein-Milman's theorem (see Lemma 1 in Chen et al., 1993) which require $\mathbf{D}$ to be convex and/or f to have monotone likelihood ratio. In the same way as in Sections 2.3, under weaker conditions we shall obtain some of their results more easily and more systematically.

In Section 2.5, we shall give some applications of our majorization results on hyperplanes to certain robust testing problems of sevral location parameters. It is seen that the parameters $\boldsymbol{\mu}^{LFC}$, $\boldsymbol{\mu}^M$ and $\boldsymbol{\mu}^s$ are effectively utilized for constructing robust tests and for evaluating their powers on a parameter set $\boldsymbol{\Gamma}$ induced from $\boldsymbol{\Theta}$. These applications reveal that our majorization methods are useful and powerful.

2.2. Majorization method for general parameter sets.

Let $\boldsymbol{\Gamma}$ be a nonempty and bounded subset of $\boldsymbol{\Omega} = \{\boldsymbol{\mu} \mid \sum_{i=1}^k \mu_i = 0\}$. Let α_r and β_r $(r = 1, \ldots, k)$ be defined by

$$\alpha_r = \inf\{\sum_{i=1}^r \mu_{[i]} \mid \boldsymbol{\mu} \in \boldsymbol{\Gamma}\}, \quad \beta_r = \sup\{\sum_{i=1}^r \mu_{[i]} \mid \boldsymbol{\mu} \in \boldsymbol{\Gamma}\}. \tag{2.1}$$

By using these α_r and β_r, we define two particular parameters $\boldsymbol{\mu}^M$ and $\boldsymbol{\mu}^s$ whose r-th components μ_r^M and μ_r^s are respectively given by

$$\mu_r^M = \beta_r - \beta_{r-1}, \qquad r = 1, \ldots, k \tag{2.2}$$

and

$$\mu_r^s = \begin{cases} \alpha_r - \alpha_{r-1}, & r = 1, \ldots, s \\ -\dfrac{\alpha_s}{k-s}, & r = s+1, \ldots, k, \end{cases} \tag{3.3}$$

where $\alpha_0 = \beta_0 = \alpha_k = \beta_k = 0, -\infty < \alpha_r \leq \beta_r < +\infty$ and $s(1 \leq s \leq k)$ is some fixed integer.

Note that the relation $\boldsymbol{\mu}^k \succ \boldsymbol{\mu}^{k-1} \succ \cdots \succ \boldsymbol{\mu}^1 \succ \mathbf{0}_k$ holds, where $\mathbf{0}_k = (0, \ldots, 0) \in \mathbf{R}^k$.

We consider the following conditions:

Condition 1.

$$\beta_r - \beta_{r-1} \geq \beta_{r+1} - \beta_r, \quad r = 1, \ldots, k-1.$$

Condition 2.

$$\alpha_r - \alpha_{r-1} \geq \alpha_{r+1} - \alpha_r, \quad r = 1, \ldots, s-1$$
$$\alpha_s - \alpha_{s-1} \geq \alpha_{r+1} - \alpha_r, \quad r = s, \ldots, k-1.$$

Then the following lemma is basic.

Lemma 2.1. *The following statements hold.*

(i) $\boldsymbol{\mu} \prec \boldsymbol{\mu}^M$ *holds for every* $\boldsymbol{\mu} \in \boldsymbol{\Gamma}$.

(ii) *If Condition 2 is satisfied, then* $\boldsymbol{\mu}^s \prec \boldsymbol{\mu}$ *holds for every* $\boldsymbol{\mu} \in \boldsymbol{\Gamma}$.

Proof. For every $\mu \in \boldsymbol{\Gamma}$,

$$\sum_{i=1}^r \mu_{[i]} \leq \beta_r = (\beta_1 - \beta_0) + \cdots + (\beta_r - \beta_{r-1}) = \sum_{i=1}^r \mu_i^M \leq \sum_{i=1}^r \mu_{[i]}^M, \quad r = 1, \ldots, k-1,$$

and $\sum_{i=1}^k \mu_i = \sum_{i=1}^k \mu_i^M = 0$. These imply the assertion (i).

Assume Condition 2. Then,

$$\sum_{i=1}^r \mu_{[i]} \geq \alpha_r = (\alpha_1 - \alpha_0) + \cdots + (\alpha_r - \alpha_{r-1}) = \sum_{i=1}^r \mu_{[i]}^s, \quad r = 1, \ldots, s,$$

$$\sum_{i=1}^r \mu_{[i]} = (\alpha_1 - \alpha_0) + \cdots + (\alpha_s - \alpha_{s-1}) + \left((\sum_{i=1}^s \mu_{[i]} - \alpha_s) + \mu_{[s+1]} \right) + \sum_{i=s+2}^r \mu_{[i]}$$

$$\geq \sum_{i=1}^r \mu_{[i]}^s, \quad r = s+1, \ldots, k-1,$$

because

$$\left((\sum_{i=1}^{s} \mu_{[s]} - \alpha_s) + \mu_{[s+1]}, \mu_{[s+2]}, \dots, \mu_{[k]} \right) \succ (-\frac{1}{k-s})(\underbrace{\alpha_s, \dots, \alpha_s}_{(k-s)\text{-times}})$$

and

$$\alpha_1 - \alpha_0 \geq \cdots \geq \alpha_s - \alpha_{s-1} \geq -\frac{1}{k-s}\alpha_s.$$

Hence, the assertion (ii) follows from $\sum_{i=1}^{k} \mu_i = \sum_{i=1}^{k} \mu_i^s = 0$. □

The following lemma suggests that in the sense of preordering of majorization $\boldsymbol{\mu}^M$ is the smallest of all the parameters majorizing $\boldsymbol{\Gamma}$ and $\boldsymbol{\mu}^s$ is the largest of all the parameters majorized by $\boldsymbol{\Gamma}$.

Here we mean by $\boldsymbol{\mu} \not\succ \boldsymbol{\nu}$ $(\boldsymbol{\mu}, \boldsymbol{\nu} \in \boldsymbol{\Omega})$ that $\boldsymbol{\mu}$ is not majorized by $\boldsymbol{\nu}$, that is, $\boldsymbol{\mu} \not\succ \boldsymbol{\nu}$ if and only if there exists an integer $r^*(1 \leq r^* \leq k-1)$ such that $\sum_{i=1}^{r^*} \mu_{[i]} > \sum_{i=1}^{r^*} \nu_{[i]}$.

Lemma 2.2. *The following statements hold.*

(i) *Suppose that Condition 1 holds. If $\boldsymbol{\mu}' \in \boldsymbol{\Omega}$ satisfies $\boldsymbol{\mu}^M \not\succ \boldsymbol{\mu}'$, then there exists $\boldsymbol{\mu} \in \boldsymbol{\Gamma}$ such that $\boldsymbol{\mu} \not\succ \boldsymbol{\mu}'$.*

(ii) *Suppose that Condition 2 with $s = k$ holds. If $\boldsymbol{\mu}' \in \boldsymbol{\Omega}$ satisfies $\boldsymbol{\mu}' \not\succ \boldsymbol{\mu}^k$, then there exists $\boldsymbol{\mu} \in \boldsymbol{\Gamma}$ such that $\boldsymbol{\mu}' \not\succ \boldsymbol{\mu}$.*

Proof. To show the assertion (i) let $\boldsymbol{\mu}' \in \boldsymbol{\Omega}$ be such that $\boldsymbol{\mu}^M \not\succ \boldsymbol{\mu}'$. Then, by the definition of $\not\succ$, there exists r^* such that $\sum_{i=1}^{r^*} \mu'_{[i]} < \sum_{i=1}^{r^*} \mu^M_{[i]}$. By Condition 1 we have $\sum_{i=1}^{r^*} \mu^M_{[i]} = \beta_{r^*}$. Hence, by the definition of β_{r^*} there exists $\boldsymbol{\mu} \in \boldsymbol{\Gamma}$ such that $\sum_{i=1}^{r^*} \mu'_{[i]} < \sum_{i=1}^{r^*} \mu_{[i]}$. This implies the assertion (i).

The assertion (ii) is similarly verified. □

The following theorem is an immeadiate consequence of Theorem 1.2 in Section 1.2 of Chapter 1 and Lemma 2.1.

Theorem 2.1. *Let $Z_1, \dots, Z_k$ $(k \geq 3)$ be exchangeable random variables satisfying $\sum_{i=1}^{k} Z_i = 0$ such that $(Z_1, \dots, Z_{k-1})$ has a Schur-concave density and let denote $\mathbf{Z} = (Z_1, \dots, Z_k)$ and $\boldsymbol{\mu} = (\mu_1, \dots, \mu_k)$. Let $\mathbf{D} \subset \mathbf{R}^k$ be Schur-convex. Then the following hold.*

(i) $\Pr\{\mathbf{Z} + \boldsymbol{\mu} \in \mathbf{D}\} \geq \Pr\{\mathbf{Z} + \boldsymbol{\mu}^M \in \mathbf{D}\}$ *for every $\boldsymbol{\mu} \in \boldsymbol{\Gamma}$.*

(ii) *If Condition 2 is satisfied, then*

$\Pr\{\mathbf{Z} + \boldsymbol{\mu} \in \mathbf{D}\} \leq \Pr\{\mathbf{Z} + \boldsymbol{\mu}^s \in \mathbf{D}\}$ *for every $\boldsymbol{\mu} \in \boldsymbol{\Gamma}$.*

Similarly we have the following corollary from Corollary 1.1 in Section 1.2 and Lemma 2.1.

Corollary 2.1. *Let* $\mathbf{Z}$ *and* $\mathbf{D}$ *be as given in Theorem 2.1 and let* V *be a positive random variable independent of* $\mathbf{Z}$*. Then the following statements hold.*

(i) $\Pr\{(\mathbf{Z}+\boldsymbol{\mu})/V \in \mathbf{D}\} \geq \Pr\{(\mathbf{Z}+\boldsymbol{\mu}^M)/V \in \mathbf{D}\}$ *for every* $\boldsymbol{\mu} \in \boldsymbol{\Gamma}$.

(ii) *If Condition 2 is satisfied, then*

$$\Pr\{(\mathbf{Z}+\boldsymbol{\mu})/V \in \mathbf{D}\} \leq \Pr\{(\mathbf{Z}+\boldsymbol{\mu}^s)/V \in \mathbf{D}\} \quad \text{for every } \boldsymbol{\mu} \in \boldsymbol{\Gamma}.$$

We call the inequalities (i) and (ii) of Theorem 2.1 and Corollary 2.1 majorization inequalities.

We now define $\mathbf{M}(\boldsymbol{\Gamma})$ by

$$\mathbf{M}(\boldsymbol{\Gamma}) = \{\, \boldsymbol{\mu} \mid \alpha_r \leq \sum_{i=1}^{r} \mu_{[i]} \leq \beta_r, r = 1, \ldots, k\}.$$

Then Lemmas 2.1, 2.2, Theorem 2.1 and Corollary 2.1 remain true for $\boldsymbol{\Gamma}$ replaced by $\mathbf{M}(\boldsymbol{\Gamma})$. We note that $\mathbf{M}(\boldsymbol{\Gamma}) \supset \boldsymbol{\Gamma}$ is the largest parameter set such that the majorization inequalities are valid. It is not guaranteed that the parameters $\boldsymbol{\mu}^M$ and $\boldsymbol{\mu}^s$ are in $\mathbf{M}(\boldsymbol{\Gamma})$. However, it is easily seen that $\boldsymbol{\mu}^M$ and $\boldsymbol{\mu}^s$ are in $\mathbf{M}(\boldsymbol{\Gamma})$ under Conditions 1 and 2 with $s = k$, respectively.

Remark 2.1.

1. The assertion (i) of Lemma 2.1 is a generalization of the idea in Corollary 4 of Kimura and Kakiuchi (1989).
2. The assertion (ii) of Lemma 2.1 is a new idea. Kimura and Kakiuchi (1989) did not treat a lower bound of a parameter set $\boldsymbol{\Gamma}$ in the sense of partial order of majorization.
3. Although Condition 2 seems to be eccentric and limited, it is reasonable and not limited. Condition 2 requires that every vector of $\boldsymbol{\Gamma}$ has a certain degree of diversity of its components. For example, see (2.29), Problems C and D in Section 2.5.
4. Since our majorization method is valid for any $\boldsymbol{\Gamma}$ and for any Schur-convex set $\mathbf{D}$, Theorem 2.1 and Corollary 2.1 have broad applicability. In particular, it can be effectively applied to obtain probability inequalities for asymptotic joint distributions (singular normal distributions on hyperplanes in $\mathbf{R}^k$) of k-sample rank statistics when underlying distributions vary in some shrinking neighborhoods. In this case we encounter another problem of determining $\boldsymbol{\Gamma}$. See Chapter 3 for details, and also Kakiuchi and Kimura (1995).

2.3. Least favorable configurations for special parameter sets.

First we consider the following two special forms $\mathbf{\Gamma}_{01}$ and $\mathbf{\Gamma}_{11}$ of $\mathbf{\Gamma}$:

$$\mathbf{\Gamma}_{01} = \{\boldsymbol{\mu} \in \mathbf{\Omega} \mid \mu_{[1]} - \mu_{[k]} \leq \delta\}, \qquad \mathbf{\Gamma}_{11} = \{\boldsymbol{\mu} \in \mathbf{\Omega} \mid \mu_{[1]} - \mu_{[k]} \geq \delta\}. \tag{2.4}$$

where δ is a specified positive real number. We define

$$\boldsymbol{\mu}_{01}(r) = \Big(\underbrace{\frac{(k-r)\delta}{k}, \ldots, \frac{(k-r)\delta}{k}}_{r\text{-times}}, \underbrace{-\frac{r\delta}{k}, \ldots, -\frac{r\delta}{k}}_{(k-r)\text{-times}}\Big), \quad r = 1, \ldots, k-1. \tag{2.5}$$

The following theorem gives the candidates of the least favorable configurations for $\mathbf{\Gamma}_{01}$.

Theorem 2.2. *Let $Z_1, \ldots, Z_k (k \geq 3)$ be exchangeable random variables satisfying $\sum_{i=1}^k Z_i = 0$ and let $(Z_1, \ldots, Z_{k-1})$ have a logconcave density. Let $\mathbf{D} \subset \mathbf{R}^k$ be Schur-convex. Then*

$$\inf_{\boldsymbol{\mu} \in \mathbf{\Gamma}_{01}} \Pr\{\mathbf{Z} + \boldsymbol{\mu} \in \mathbf{D}\} = \min_{1 \leq r \leq k-1} \Pr\{\mathbf{Z} + \boldsymbol{\mu}_{01}(r) \in \mathbf{D}\}. \tag{2.6}$$

Furthermore, if the density of $(Z_1, \ldots, Z_{k-1})$ and $\mathbf{D}$ are centrally symmetric, then

$$\inf_{\boldsymbol{\mu} \in \mathbf{\Gamma}_{01}} \Pr\{\mathbf{Z} + \boldsymbol{\mu} \in \mathbf{D}\} = \min_{1 \leq r \leq [k/2]} \Pr\{\mathbf{Z} + \boldsymbol{\mu}_{01}(r) \in \mathbf{D}\}. \tag{2.7}$$

Proof. We note that convex unimodality follows from logconcavity and that convex unimodality together with permutation invariance implies Schur-concavity. Thus the density of $(Z_1, \ldots, Z_{k-1})$ is Schur-concave.

Let $\boldsymbol{\mu}$ be any nonzero parameter of $\mathbf{\Gamma}_{01}$. We can let $\mu_1 \geq \ldots \geq \mu_k$ without loss of generality. Let $\boldsymbol{\nu} = (\nu_1, \ldots, \nu_k) = (\delta/\tau)\boldsymbol{\mu}$, where $\tau = \mu_1 - \mu_k \leq \delta$. Then, we have $\boldsymbol{\mu} \prec \boldsymbol{\nu}$ and $\boldsymbol{\nu} \in \mathbf{\Gamma}_{01}$, where $\nu_1 \geq, \ldots, \geq \nu_k$ and $\nu_1 - \nu_k = \delta$. From $\sum_{i=1}^k \nu_i = 0$, we have $\delta/k \leq \nu_1 \leq (k-1)\delta/k$. Hence $\nu_1 = \delta/k$ or $\nu_1 \in ((k-r-1)\delta/k, (k-r)\delta/k]$ for some $r\,(1 \leq r \leq k-2)$.

First assume $\nu_1 \in ((k-r-1)\delta/k, (k-r)\delta/k]$ and define a parameter $\widehat{\boldsymbol{\nu}}_r(\eta) \in \mathbf{\Gamma}_{01}$ by

$$\widehat{\boldsymbol{\nu}}_r(\eta) = (\underbrace{\nu_1, \ldots, \nu_1}_{r\text{-times}}, \eta, \underbrace{\nu_k, \ldots, \nu_k}_{(k-r-1)\text{-times}}),$$

where $\eta = -\{r\nu_1 + (k-r-1)\nu_k\}, (0 \leq \eta \leq 1)$. Then we have $\boldsymbol{\nu} \prec \widehat{\boldsymbol{\nu}}_r(\eta)$ and hence by Theorem 1.2 in Section 1.2

$$\Pr\{\mathbf{Z} + \boldsymbol{\mu} \in \mathbf{D}\} \geq \Pr\{\mathbf{Z} + \widehat{\boldsymbol{\nu}}_r(\eta) \in \mathbf{D}\}. \tag{2.8}$$

We can easily see that

$$\widehat{\boldsymbol{\nu}}_r(\eta) = (1-\lambda)\boldsymbol{\mu}_{01}(r) + \lambda\boldsymbol{\mu}_{01}(r+1)$$

where $\boldsymbol{\mu}_{01}(r)\,(1 \leq r \leq k-1)$ are given by (2.5) and

$$\lambda = \{(k-r)\delta/k - \nu_1\}/(\delta/k), \quad (0 \leq \lambda \leq 1).$$

Let $h : \mathbf{R}^k \to \mathbf{R}^{k-1}$ be the projection defined by $h(x_1, \dots, x_k) = (x_1, \dots, x_{k-1})$. Then the restriction of h to $\boldsymbol{\Omega} = \{\mathbf{x} \mid \sum_{i=1}^k x_i = 0\}$ is a bijection from $\boldsymbol{\Omega}$ to $\mathbf{R}^{k-1}$. Letting $\mathbf{D}' = \mathbf{D} \cap \boldsymbol{\Omega}$, then we have

$$\begin{aligned}
&\Pr\{\mathbf{Z} + \widehat{\boldsymbol{\nu}}_r(\eta) \in \mathbf{D}\} \\
&= \Pr\{h(\mathbf{Z}) + h(\widehat{\boldsymbol{\nu}}_r(\eta)) \in h(\mathbf{D}')\} \\
&= \Pr\{h(\mathbf{Z}) \in (1-\lambda)\big(h(\mathbf{D}') - h(\boldsymbol{\mu}_{01}(r))\big) + \lambda\big(h(\mathbf{D}') - h(\boldsymbol{\mu}_{01}(r+1))\big)\} \\
&\geq \Big\{\Pr\{h(\mathbf{Z}) \in h(\mathbf{D}') - h(\boldsymbol{\mu}_{01}(r))\}\Big\}^{1-\lambda}\Big\{\Pr\{h(\mathbf{Z}) \in h(\mathbf{D}') - h(\boldsymbol{\mu}_{01}(r+1))\}\Big\}^{\lambda} \\
&\geq \min\Big\{\Pr\{\mathbf{Z} + \boldsymbol{\mu}_{01}(r) \in \mathbf{D}\}, \Pr\{\mathbf{Z} + \boldsymbol{\mu}_{01}(r+1) \in \mathbf{D}\}\Big\}. \qquad (2.9)
\end{aligned}$$

The above first inequality follows from the fact that the distribution of $h(\mathbf{Z})$ is logconcave. We note that by Theorem 2 in Prékopa (1973) the logconcavity of the density implies the logconcavity of the probability measure, where a probability measure P is said to be logconcave if for every nonempty set $\mathbf{A}$, $\mathbf{B} \subset \mathbf{R}^k$ and for every $\alpha \in (0,1)$ $P(\alpha\mathbf{A} + (1-\alpha)\mathbf{B}) \geq [P(\mathbf{A})]^\alpha[P(\mathbf{B})]^{1-\alpha}$.

Next let us consider the case of $\nu_1 = \delta/k$. Then we have $\boldsymbol{\nu} = \boldsymbol{\mu}_{01}(k-1)$. Hence, from (2.8) and (2.9) it follows that

$$\Pr\{\mathbf{Z} + \boldsymbol{\mu} \in \mathbf{D}\} \geq \min_{1 \leq r \leq k-1} \Pr\{\mathbf{Z} + \boldsymbol{\mu}_{01}(r) \in \mathbf{D}\}.$$

Noting that this inequality holds for $\boldsymbol{\mu} = \mathbf{0}$ and that $\boldsymbol{\mu}_{01}(r) \in \boldsymbol{\Gamma}_{01}$, $r = 1, \dots, k-1$, we obtain the first assertion of the theorem.

To show the second assertion of the theorem we first note that $\boldsymbol{\mu}_{01}(r) = -\boldsymbol{\mu}_{01}(k-r)$, $r = 1, \dots, k-1$. Since $\mathbf{D}$ and the density of $(Z_1, \dots, Z_{k-1})$ are centrally symmetric, we have

$$\begin{aligned}
\Pr\{\mathbf{Z} + \boldsymbol{\mu}_{01}(r) \in \mathbf{D}\} &= \Pr\{-\mathbf{Z} - \boldsymbol{\mu}_{01}(r) \in -\mathbf{D}\} \\
&= \Pr\{\mathbf{Z} + \boldsymbol{\mu}_{01}(k-r) \in \mathbf{D}\}, \quad r = 1, \dots, k-1.
\end{aligned}$$

This implies the second assertion of the theorem. □

Remark 2.2. When we do not know which one of $\boldsymbol{\mu}_{01}(1), \dots, \boldsymbol{\mu}_{01}(k-1)$ is the best, we may be interested in the smallest parameter that majorizes all of these $k-1$ parameters. This smallest parameter is seen to be $\boldsymbol{\mu}^M$ defined by (2.2) with $\boldsymbol{\Gamma} = \boldsymbol{\Gamma}_{01}$, where the r-th component of $\boldsymbol{\mu}^M$ is given by

$$\mu_r^M = \frac{(k-2r+1)\delta}{k}, \quad r = 1, \dots, k. \qquad (2.10)$$

By Theorem 2.1 we have

$$\min_{1\le r\le k-1} \Pr\{\mathbf{Z}+\boldsymbol{\mu}_{01}(r)\in\mathbf{D}\} \ge \Pr\{\mathbf{Z}+\boldsymbol{\mu}^M\in\mathbf{D}\}. \tag{2.11}$$

The following theorem gives the least favorable configuration for $\boldsymbol{\Gamma}_{11}$.

Theorem 2.3. *Let $Z_1,\dots,Z_k (k\ge 3)$ be exchangeable random variables satisfying $\sum_{i=1}^k Z_i = 0$ and let $(Z_1,\dots,Z_{k-1})$ have a centrally symmetric and logconcave density. Let $\mathbf{D}\subset\mathbf{R}^k$ be centrally symmetric and Schur-convex. Then*

$$\sup_{\boldsymbol{\mu}\in\boldsymbol{\Gamma}_{11}} \Pr\{\mathbf{Z}+\boldsymbol{\mu}\in\mathbf{D}\} = \Pr\{\mathbf{Z}+\boldsymbol{\mu}_{11}\in\mathbf{D}\}, \tag{2.12}$$

where

$$\boldsymbol{\mu}_{11} = (\frac{\delta}{2},0,\dots,0,-\frac{\delta}{2}). \tag{2.13}$$

Proof. Since the density of $(Z_1,\dots,Z_{k-1})$ is Schur-concave, by Theorem 1.2 in Section 1.2 $\Pr\{\mathbf{Z}+\boldsymbol{\mu}\in\mathbf{D}\}$ is Schur-concave in $\boldsymbol{\mu}$.

Let $\boldsymbol{\mu}$ be any parameter of $\boldsymbol{\Gamma}_{11}$. We can let $\mu_1\ge\dots\ge\mu_k$ without loss of generality. Let $\boldsymbol{\nu}=(\nu_1,\dots,\nu_k)=(\delta/\tau)\boldsymbol{\mu}$, where $\tau=\mu_1-\mu_k\ge\delta$. Then, we have $\boldsymbol{\mu}\succ\boldsymbol{\nu}$ and $\boldsymbol{\nu}\in\boldsymbol{\Gamma}_{11}$, where $\nu_1\ge\dots\ge\nu_k$ and $\nu_1-\nu_k=\delta$. Let

$$\boldsymbol{\xi} = (\nu_1,\nu_k,-\frac{\nu_1+\nu_k}{k-2},\dots,-\frac{\nu_1+\nu_k}{k-2}).$$

Then, clearly we see $\boldsymbol{\xi}\prec\boldsymbol{\nu}$ and hence $\boldsymbol{\xi}\prec\boldsymbol{\mu}$. Therefore, by Theorem 1.2 in Section 1.2

$$\Pr\{\mathbf{Z}+\boldsymbol{\xi}\in\mathbf{D}\} \ge \Pr\{\mathbf{Z}+\boldsymbol{\mu}\in\mathbf{D}\}. \tag{2.14}$$

Let h and $\mathbf{D}'$ be as given in the proof of Theorem 2.2. Then we have

$$\Pr\{\mathbf{Z}+\boldsymbol{\xi}\in\mathbf{D}\} = \Pr\{h(\mathbf{Z})+h(\boldsymbol{\xi})\in h(\mathbf{D}')\}. \tag{2.15}$$

It is easy to see that $h(\mathbf{D}')$ is centrally symmetric and Schur-convex. Let G_1 be a group of all permutations of the components of $\mathbf{x}$ and let $G_2=\{g_{21},g_{22}\}$ be a group of two transformations such that $g_{21}(\mathbf{x})=\mathbf{x}$ and $g_{22}(\mathbf{x})=-\mathbf{x}$. Let G be the direct product group of G_1 and G_2. Then the density function of $(Z_1,\dots,Z_{k-1})$ and $h(\mathbf{D}')$ are G-invariant. Let $g_0\in G$ be a transformation such that $g_0(x_1,x_2,x_3,\dots,x_{k-1}) = (-x_2,-x_1,-x_3,\dots,-x_{k-1})$. Then we can see

$$h(\widehat{\boldsymbol{\mu}}) = \frac{1}{2}\big(h(\boldsymbol{\xi})+g_0(h(\boldsymbol{\xi}))\big),\ \text{where } \widehat{\boldsymbol{\mu}}=(\delta/2,-\delta/2,0,\dots,0).$$

Hence, from the logconcavity of the density of $h(\mathbf{Z})$ it follows that

$$\begin{aligned}&\Pr\{h(\mathbf{Z})+h(\widehat{\boldsymbol{\mu}})\in h(\mathbf{D}')\}\\ &\geq \left[\Pr\{h(\mathbf{Z})+h(\boldsymbol{\xi})\in h(\mathbf{D}')\}\right]^{1/2}\left[\Pr\{h(\mathbf{Z})+g_0(h(\boldsymbol{\xi}))\in h(\mathbf{D}')\}\right]^{1/2} \\ &= \Pr\{h(\mathbf{Z})+h(\boldsymbol{\xi})\in h(\mathbf{D}')\}.\end{aligned} \tag{2.16}$$

Therefore, from (2.14), (2.15) and (2,16) it follows that

$$\Pr\{\mathbf{Z}+\widehat{\boldsymbol{\mu}}\in\mathbf{D}\}=\Pr\{\mathbf{Z}+\boldsymbol{\mu}_{11}\in\mathbf{D}\}\geq\Pr\{\mathbf{Z}+\boldsymbol{\mu}\in\mathbf{D}\}\quad\text{for every }\boldsymbol{\mu}\in\boldsymbol{\Gamma}_{11}.$$

This fact and $\boldsymbol{\mu}_{11}\in\boldsymbol{\Gamma}_{11}$ imply the theorem. □

Next, we consider the least favorable configurations for the following four parameter sets:

$$\boldsymbol{\Gamma}_{02}=\{\boldsymbol{\mu}\in\boldsymbol{\Omega}\mid\max_{1\leq i\leq k}|\mu_i|\leq\delta\},\qquad \boldsymbol{\Gamma}_{12}=\{\boldsymbol{\mu}\in\boldsymbol{\Omega}\mid\max_{1\leq i\leq k}|\mu_i|\geq\delta\},$$

$$\boldsymbol{\Gamma}_{03}=\{\boldsymbol{\mu}\in\boldsymbol{\Omega}\mid\sum_{i=1}^{k}|\mu_i|\leq\delta\},\qquad \boldsymbol{\Gamma}_{13}=\{\boldsymbol{\mu}\in\boldsymbol{\Omega}\mid\sum_{i=1}^{k}|\mu_i|\geq\delta\}.$$

Let $\boldsymbol{\mu}_{ij}\in\boldsymbol{\Gamma}_{ij}$ $(i=0,1,j=2,3)$ be defined by

$$\boldsymbol{\mu}_{02}=\begin{cases}(\underbrace{\delta,\ldots,\delta}_{k/2\text{-times}},\underbrace{-\delta,\ldots,-\delta}_{k/2\text{-times}}), & k\text{ is even}\\ (\underbrace{\delta,\ldots,\delta}_{[k/2]\text{-times}},0,\underbrace{-\delta,\ldots,-\delta}_{[k/2]\text{-times}}), & k\text{ is odd}\end{cases} \tag{2.17}$$

$$\boldsymbol{\mu}_{12}=(\delta,\underbrace{-\frac{\delta}{k-1},\ldots,-\frac{\delta}{k-1}}_{(k-1)\text{-times}}), \tag{2.18}$$

$$\boldsymbol{\mu}_{03}=\boldsymbol{\mu}_{11}=(\frac{\delta}{2},\underbrace{0,\ldots,0}_{(k-2)\text{-times}},-\frac{\delta}{2}), \tag{2.19}$$

$$\boldsymbol{\mu}_{13}(l)=(\underbrace{\frac{\delta}{2(k-l)},\ldots,\frac{\delta}{2(k-l)}}_{(k-l)\text{-times}},\underbrace{-\frac{\delta}{2l},\ldots,-\frac{\delta}{2l}}_{l\text{-times}}),\quad l=1,\ldots,k-1. \tag{2.20}$$

Theorem 2.4. *Let* $\mathbf{Z}$ *and* $\mathbf{D}$ *be as given in Theorem 2.1. Then the following statements hold.*

(i) $\inf_{\boldsymbol{\mu}\in\boldsymbol{\Gamma}_{02}}\Pr\{\mathbf{Z}+\boldsymbol{\mu}\in\mathbf{D}\}=\Pr\{\mathbf{Z}+\boldsymbol{\mu}_{02}\in\mathbf{D}\}$.

(ii) $\sup_{\boldsymbol{\mu}\in\boldsymbol{\Gamma}_{12}}\Pr\{\mathbf{Z}+\boldsymbol{\mu}\in\mathbf{D}\}=\max\big\{\Pr\{\mathbf{Z}+\boldsymbol{\mu}_{12}\in\mathbf{D}\},\Pr\{\mathbf{Z}-\boldsymbol{\mu}_{12}\in\mathbf{D}\}\big\}$.

Furthermore, if the density of $(Z_1, \dots, Z_{k-1})$ *and* $\mathbf{D}$ *are centrally symmetric, then*

$$\sup_{\boldsymbol{\mu} \in \boldsymbol{\Gamma}_{12}} \Pr\{\mathbf{Z} + \boldsymbol{\mu} \in \mathbf{D}\} = \Pr\{\mathbf{Z} + \boldsymbol{\mu}_{12} \in \mathbf{D}\}.$$

(iii) $\inf_{\boldsymbol{\mu} \in \boldsymbol{\Gamma}_{03}} \Pr\{\mathbf{Z} + \boldsymbol{\mu} \in \mathbf{D}\} = \Pr\{\mathbf{Z} + \boldsymbol{\mu}_{03} \in \mathbf{D}\}.$

(iv) $\sup_{\boldsymbol{\mu} \in \boldsymbol{\Gamma}_{13}} \Pr\{\mathbf{Z} + \boldsymbol{\mu} \in \mathbf{D}\} = \max_{1 \le l \le k-1} \Pr\{\mathbf{Z} + \boldsymbol{\mu}_{13}(l) \in \mathbf{D}\}.$

Furthermore, if the density of $(Z_1, \dots, Z_{k-1})$ *and* $\mathbf{D}$ *are centrally symmetric, then*

$$\sup_{\boldsymbol{\mu} \in \boldsymbol{\Gamma}_{13}} \Pr\{\mathbf{Z} + \boldsymbol{\mu} \in \mathbf{D}\} = \max_{1 \le l \le [k/2]} \Pr\{\mathbf{Z} + \boldsymbol{\mu}_{13}(l) \in \mathbf{D}\}.$$

Proof. By Theorem 1.2 in Section 1.2 it is sufficient to show that $\boldsymbol{\mu}_{0j} \succ \boldsymbol{\mu}$ holds for every $\boldsymbol{\mu} \in \boldsymbol{\Gamma}_{0j}$ $(j = 2, 3)$, that at least one of $\boldsymbol{\mu}_{12} \prec \boldsymbol{\mu}$ and $-\boldsymbol{\mu}_{12} \prec \boldsymbol{\mu}$ holds for every $\boldsymbol{\mu} \in \boldsymbol{\Gamma}_{12}$ and that at least one of $\boldsymbol{\mu}_{13}(l) \prec \boldsymbol{\mu}$, $l = 1, \dots k-1$, holds for every $\boldsymbol{\mu} \in \boldsymbol{\Gamma}_{13}$.

(i) For any nonzero $\boldsymbol{\mu} \in \boldsymbol{\Gamma}_{02}$ let $\max_{1 \le i \le k} |\mu_i| = \tau (\le \delta)$ and $\boldsymbol{\nu} = (\delta/\tau)\boldsymbol{\mu}$. Then we see $\boldsymbol{\nu} \succ \boldsymbol{\mu}$ and $\boldsymbol{\nu} \in \boldsymbol{\Gamma}_{02}$, where $\max_{1 \le i \le k} |\nu_i| = \delta$. Clearly, for $l = 1, \dots, [k/2]$ we have

$$\sum_{i=1}^{l} \nu_{[i]} \le l\delta = \sum_{i=1}^{l} \mu_{02[i]} \qquad \text{and} \qquad \sum_{i=1}^{l} \nu_{[k-i+1]} \ge -l\delta.$$

Also, since $l \ge [k/2] + 1$ implies $k - l \le k/2$, we have for $l = [k/2] + 1, \dots, k-1$

$$\sum_{i=1}^{l} \nu_{[i]} = -\sum_{i=1}^{k-l} \nu_{[k-i+1]} \le (k-l)\delta = \sum_{i=1}^{l} \mu_{02[i]}.$$

This implies $\boldsymbol{\nu} \prec \boldsymbol{\mu}_{02}$. Thus $\boldsymbol{\mu} \prec \boldsymbol{\mu}_{02}$ holds for every $\boldsymbol{\mu} \in \boldsymbol{\Gamma}_{02}$ (note $\mathbf{0} \prec \boldsymbol{\mu}_{02}$).

(ii) As in the proof of (i), for every $\boldsymbol{\mu} \in \boldsymbol{\Gamma}_{12}$ there exists $\boldsymbol{\nu} \in \boldsymbol{\Gamma}_{12}$ such that $\boldsymbol{\nu} \prec \boldsymbol{\mu}$ and $\max_{1 \le i \le k} |\nu_i| = \delta$. If $\nu_{[1]} = \delta$, then $\boldsymbol{\mu}_{12} \prec \boldsymbol{\nu}$ follows from $(-\delta/(k-1), \dots, -\delta/(k-1)) \prec (\nu_{[2]}, \dots, \nu_{[k]})$. If $\nu_{[k]} = -\delta$, then $-\boldsymbol{\mu}_{12} \prec \boldsymbol{\nu}$. Hence, for every $\boldsymbol{\mu} \in \boldsymbol{\Gamma}_{12}$ at least one of $\boldsymbol{\mu}_{12} \prec \boldsymbol{\mu}$ and $-\boldsymbol{\mu}_{12} \prec \boldsymbol{\mu}$ hold. The second assertion is easily shown from the symmetry condition.

(iii) For any nonzero $\boldsymbol{\mu} \in \boldsymbol{\Gamma}_{03}$ let $\sum_{i=1}^{k} |\mu_i| = \tau (\le \delta)$ and $\boldsymbol{\nu} = (\delta/\tau)\boldsymbol{\mu}$. Then we see $\boldsymbol{\nu} \in \boldsymbol{\Gamma}_{03}$ such that $\boldsymbol{\nu} \succ \boldsymbol{\mu}$ and $\sum_{i=1}^{k} |\nu_i| = \delta$. Hence we have $\sum_{i \in M} \nu_i = \delta/2$ and $\sum_{i \in M^c} \nu_i = -\delta/2$, where $M = \{j \mid \nu_j \ge 0\}$. This implies $\boldsymbol{\nu} \prec \boldsymbol{\mu}_{03}$. Thus $\boldsymbol{\mu} \prec \boldsymbol{\mu}_{03}$ holds for every $\boldsymbol{\mu} \in \boldsymbol{\Gamma}_{03}$.

(iv) As in the proof of (iii), for every $\boldsymbol{\mu} \in \boldsymbol{\Gamma}_{13}$ there exists $\boldsymbol{\nu} \in \boldsymbol{\Gamma}_{13}$ such that $\boldsymbol{\nu} \prec \boldsymbol{\mu}$ and $\sum_{i=1}^{k} |\nu_i| = \delta$. Let l denote the number of negative components of $\boldsymbol{\nu}$. Then $\boldsymbol{\mu}_{13}(l) \prec \boldsymbol{\mu}$ follows from $\sum_{i=1}^{k-l} \nu_{[i]} = -\sum_{i=k-l+1}^{k} \nu_{[i]} = \delta/2$. Hence, at least one of $\boldsymbol{\mu}_{13}(l) \prec \boldsymbol{\mu}, l = 1, \dots, k-1$, holds for every $\boldsymbol{\mu} \in \boldsymbol{\Gamma}_{13}$. The second assertion is easily shown from $\boldsymbol{\mu}_{13}(l) = -\boldsymbol{\mu}_{13}(k-l), l = 1, \dots, k-1$. □

Remark 2.3. Let V be as given in Corollary 2.1 and let $\boldsymbol{\mu}_{ij} (i = 0, 1, j = 1, 2, 3)$ denote the least favorable configurations for the special parameter sets $\boldsymbol{\Gamma}_{ij}$ in Theorems

2.2, 2.3 and 2.4, respectively. We note that if $\mathbf{Z} + \boldsymbol{\mu}$ and $\mathbf{Z} + \boldsymbol{\mu}_{ij}$ in Theorems 2.2, 2.3 and 2.4 are replaced by $(\mathbf{Z} + \boldsymbol{\mu})/V$ and $(\mathbf{Z} + \boldsymbol{\mu}_{ij})/V$, respectively, then all the theorems remain true.

2.4. Majorization for nonsingular distributions.

Let $X_1, \ldots, X_k$ be exchangeable random variables with a joint density f and let $\mathbf{D}$ and $\boldsymbol{\Theta}$ be subsets of $\mathbf{R}^k$. The problem is to seek least favorable configurations $\boldsymbol{\theta}^{LFC}$ for $\psi(\boldsymbol{\theta}) = \Pr\{\mathbf{X} + \boldsymbol{\theta} \in \mathbf{D}\}$ on $\boldsymbol{\Theta}$ or parameters $\boldsymbol{\theta}^M$ and $\boldsymbol{\theta}^s$ at which ψ takes a lower and an upper bound close to its infimum and supremum on $\boldsymbol{\Theta}$, respectively. In this section we assume that $\mathbf{D}$ is shift invariant. In this case, since $\mathbf{D}$ is shift invariant, we have

$$\Pr\{\mathbf{X} + \boldsymbol{\theta} \in \mathbf{D}\} = \Pr\{\mathbf{X} + \boldsymbol{\mu}(\boldsymbol{\theta}) \in \mathbf{D}\}, \tag{2.21}$$

where $\mathbf{X} = (X_1, \ldots, X_k)$, $\boldsymbol{\mu}(\boldsymbol{\theta}) = \boldsymbol{\theta} - \bar{\theta}\mathbf{1}_k$ and $\bar{\theta} = (1/k)\sum_{i=1}^k \theta_i$. This fact implies that the problem on $\boldsymbol{\Theta}$ is reduced to that on $\boldsymbol{\Gamma}(\boldsymbol{\Theta}) = \{\boldsymbol{\mu}(\boldsymbol{\theta}) \mid \boldsymbol{\theta} \in \boldsymbol{\Theta}\}$, a subset of $\boldsymbol{\Omega}$. Therefore we can use the results in previous sections to achieve our purpose.

As for least favorable configurations let us consider the following six special parameter sets

$$\boldsymbol{\Theta}_{0j} = \{\boldsymbol{\theta} \mid d_j(\boldsymbol{\theta}) \le \delta\} \text{ and } \boldsymbol{\Theta}_{1j} = \{\boldsymbol{\theta} \mid d_j(\boldsymbol{\theta}) \ge \delta\}, \quad j = 1, 2, 3,$$

where $d_1(\boldsymbol{\theta}) = \theta_{[1]} - \theta_{[k]}$, $d_2(\boldsymbol{\theta}) = \max_{1 \le i \le k} |\theta_i - \bar{\theta}|$ and $d_3(\boldsymbol{\theta}) = \sum_{i=1}^k |\theta_i - \bar{\theta}|$. We note that $\boldsymbol{\Gamma}(\boldsymbol{\Theta}_{ij}) = \boldsymbol{\Gamma}_{ij}$ $(i = 0, 1, j = 1, 2, 3)$, where $\boldsymbol{\Gamma}_{ij}$ are given in Section 2.3.

Theorem 2.5. *The following statements hold.*

(i) *If f is Schur-concave and $\mathbf{D}$ is Schur-convex, then (i) and (ii) in Theorem 2.1 with $\mathbf{Z}$, $\boldsymbol{\Gamma}$, $\boldsymbol{\mu}$, $\boldsymbol{\mu}^M$ and $\boldsymbol{\mu}^s$ replaced by $\mathbf{X}$, $\boldsymbol{\Theta}$, $\boldsymbol{\theta}$, $\boldsymbol{\theta}^M$ and $\boldsymbol{\theta}^s$ holds, where $\boldsymbol{\theta}^M$ and $\boldsymbol{\theta}^s$ are defined by (2.1), (2.2) and (2.3) with $\boldsymbol{\Gamma}$ replaced by $\boldsymbol{\Gamma}(\boldsymbol{\Theta})$.*

(ii) *If f is logconcave and $\mathbf{D}$ is Schur-convex, then (2.6) in Theorem 2.2 with $\mathbf{Z}$, $\boldsymbol{\Gamma}_{01}$ and $\boldsymbol{\mu}$ replaced by $\mathbf{X}$, $\boldsymbol{\Theta}_{01}$ and $\boldsymbol{\theta}$ holds. Furthermore, if f and $\mathbf{D}$ is centrally symmetric, then (2.7) in Theorem 2.2 with $\mathbf{Z}$, $\boldsymbol{\Gamma}_{01}$ and $\boldsymbol{\mu}$ replaced by $\mathbf{X}$, $\boldsymbol{\Theta}_{01}$ and $\boldsymbol{\theta}$ holds.*

(iii) *If f is centrally symmetric and logconcave and $\mathbf{D}$ is centrally symmetric and Schur-convex, then (2.12) in Theorem 2.3 with $\mathbf{Z}$, $\boldsymbol{\Gamma}_{11}$ and $\boldsymbol{\mu}$ replaced by $\mathbf{X}$, $\boldsymbol{\Theta}_{11}$ and $\boldsymbol{\theta}$ holds.*

(iv) *If f is Schur-concave and $\mathbf{D}$ is Schur-convex, then (i), (iii) and the first assertions of (ii) and (iv) in Theorem 2.4 with $\mathbf{Z}$, $\boldsymbol{\Gamma}_{ij}$ $(i = 0, 1, j = 2, 3)$ and $\boldsymbol{\mu}$ replaced by $\mathbf{X}$, $\boldsymbol{\Theta}_{ij}$ and $\boldsymbol{\theta}$ hold. Furthermore, if f and $\mathbf{D}$ are centrally symmetric, then the second assertions of (ii) and (iv) in Theorem 2.4 with $\mathbf{Z}$, $\boldsymbol{\Gamma}_{ij}$ and $\boldsymbol{\mu}$ replaced by $\mathbf{X}$, $\boldsymbol{\Theta}_{ij}$ $(i = 0, 1, j = 2, 3)$ and $\boldsymbol{\theta}$ hold.*

Proof. The assertion (i) follows from (2.21), Theorem 1.1 in Section 1.2 and Lemma 2.1.

The assertions (ii) and (iii) are proved by using (5.1) and Theorem 1.1 in a similar way to Theorems 2.2 and 2.3, respectively.

The assertion (iv) follows from (2.21), Theorem 1.1 and the fact that $\boldsymbol{\mu}_{ij}$, $i = 0, 1, j = 2, 3$, defined by (2.17), (2.18), (2.19) and (2.20) are the smallest or the largest in $\boldsymbol{\Gamma}_{ij}$, or their candidates, as shown in Theorem 2.4. □

Remark 2.4.

1. The result (i) of Theorem 2.5 gives majorization inequalities for nonsingular distributions, which are new results.
2. The results (ii) and (iii) of Theorem 2.5 correspond to Theorems 2.1 and 2.2 of Giani and Finner (1991) and weaken their conditions. In particular, the convexity of $\mathbf{D}$ is replaced by the Schur-convexity of $\mathbf{D}$.
3. The results in (iv) of Theorem 2.5 are more general and simpler than those of Chen, Lam and Xiong (1993, Theorems 1, 2 and the results in Section 2.4) and Finner and Roters (1993, Theorems 3.1 and 3.2).

Remark 2.5. The majorization inequalities and the least favorable configurations for $\Pr\{(\mathbf{X}+\boldsymbol{\theta})/V \in \mathbf{D}\}$ on the parameter sets $\boldsymbol{\Theta}$ and $\boldsymbol{\Theta}_{ij}$ $(i = 0, 1, j = 1, 2, 3)$ are given by the same parameter vectors as in Theorem 2.5, respectively, where V is a positive random variable independent of $\mathbf{X}$.

2.5. Applications to robust testing.

We consider robust testing of k location parameters as some applications of our majorization methods on hyperplanes.

Let $X_1, \ldots, X_k$ be exchangeable random variables distributed with a joint density $f(\mathbf{X} - \boldsymbol{\theta})$, where f is Schur-concave. For each parameter $\boldsymbol{\theta}^* = (\theta_1^*, \ldots, \theta_k^*) \in \mathbf{R}^k$ let

$$J(\boldsymbol{\theta}^*) = [-\epsilon + \theta_1^*, \theta_1^* + \epsilon] \times \cdots \times [-\epsilon + \theta_k^*, \theta_k^* + \epsilon]$$

be a k-dimensional interval with the center $\boldsymbol{\theta}^*$, where ϵ is a given positive constant. For each subset $\boldsymbol{\Theta}^*$ of $\mathbf{R}^k$ let

$$J(\boldsymbol{\Theta}^*) = \bigcup_{\boldsymbol{\theta}^* \in \boldsymbol{\Theta}^*} J(\boldsymbol{\theta}^*).$$

We are interested in the following type of testing problems of $\theta_1, \ldots, \theta_k$:

$$\mathbf{H}_0 : \boldsymbol{\theta} \in J(\boldsymbol{\Theta}_0^*) \quad \textbf{vs.} \quad \mathbf{H}_1 : \boldsymbol{\theta} \in J(\boldsymbol{\Theta}_1^*) \tag{2.22}$$

where $\mathbf{\Theta}_0^*$ and $\mathbf{\Theta}_1^*$ are subsets of $\mathbf{R}^k$, and $J(\mathbf{\Theta}_0^*) \cap J(\mathbf{\Theta}_1^*) = \phi$. Since each vector $\boldsymbol{\theta}^*$ is accompanied by the interval $J(\boldsymbol{\theta}^*)$, the problem (2.22) is regarded as a robust version of

$$\mathbf{H}_0 : \boldsymbol{\theta} \in \mathbf{\Theta}_0^* \quad \textbf{vs.} \quad \mathbf{H}_1 : \boldsymbol{\theta} \in \mathbf{\Theta}_1^*. \tag{2.23}$$

Note that when $\epsilon = 0$, the problem (2.22) reduces to the problem (2.23). We treat the following four cases of the problem (2.22):

A. $\mathbf{\Theta}_{0\mathrm{A}}^* = \{\boldsymbol{\theta}^* \mid \boldsymbol{\theta}^* = \boldsymbol{\theta}_0^* + a\mathbf{1}, a \in \mathbf{R}\}$, $\mathbf{\Theta}_{1\mathrm{A}}^* = \{\boldsymbol{\theta}^* \mid \boldsymbol{\theta}^* = \boldsymbol{\theta}_1^* + a\mathbf{1}, a \in \mathbf{R}\}$.

B. $\mathbf{\Theta}_{0\mathrm{B}}^* = \{\boldsymbol{\theta}^* \mid \boldsymbol{\theta}^* = a\mathbf{1}, a \in \mathbf{R}\}$, $\mathbf{\Theta}_{1\mathrm{B}}^* = \{\boldsymbol{\theta}^* \mid \theta_{[1]}^* - \theta_{[k]}^* > \eta\}$, $\eta \geq 4\epsilon$.

C. $\mathbf{\Theta}_{0\mathrm{C}}^* = \{\boldsymbol{\theta}^* \mid \boldsymbol{\theta}^* = a\mathbf{1}, a \in \mathbf{R}\}$, $\mathbf{\Theta}_{1\mathrm{C}}^* = \{\boldsymbol{\theta}^* \mid \theta_{[i]}^* - \theta_{[i+1]}^* > \eta, i = 1, \ldots, k-1\}$, $\eta \geq 2\epsilon$.

D. $\mathbf{\Theta}_{0\mathrm{D}}^* = \{\boldsymbol{\theta}^* \mid \boldsymbol{\theta}^* = a\mathbf{1}, a \in \mathbf{R}\}$, $\mathbf{\Theta}_{1\mathrm{D}}^* = \{\boldsymbol{\theta}^* \mid \theta_{[1]}^* - \theta_{[2]}^* \geq \eta\}$, $\eta \geq 4\epsilon$.

The set $\mathbf{\Theta}_{0\mathrm{B}}^* (= \mathbf{\Theta}_{0\mathrm{C}}^* = \mathbf{\Theta}_{0\mathrm{D}}^*)$ is a special case of $\mathbf{\Theta}_{0\mathrm{A}}^*$ with $\boldsymbol{\theta}_0^* = \mathbf{0}$, and the null hypothesis $\mathbf{H}_0 : \boldsymbol{\theta} \in J(\mathbf{\Theta}_{0\mathrm{B}}^*)$ expresses approximate equality of $\theta_1, \ldots, \theta_k$. The sets $\mathbf{\Theta}_{1\mathrm{C}}^*$ and $\mathbf{\Theta}_{1\mathrm{D}}^*$ are subsets of $\mathbf{\Theta}_{1\mathrm{B}}^*$. Since the problems A, B, C and D are shift invariant, it is natural to consider the following tests based on a maximal invariant statistics $(X_1 - \bar{X}, \ldots, X_k - \bar{X})$ and $\mathbf{D}(\lambda)$:

$$\varphi(\mathbf{X}) = 0, \quad 1 \qquad if \quad (X_1 - \bar{X}, \ldots, X_k - \bar{X}) \in, \quad \notin \mathbf{D}(\lambda), \tag{2.24}$$

where $\mathbf{X} = (X_1, \ldots, X_k), \bar{X} = (1/k)\sum_{i=1}^k X_i$ and $D(\lambda)$ is a Schur-convex set depending on a real number λ. The λ usually means a critical point. Let

$$Z_i = (X_i - \bar{X}) - (\theta_i - \bar{\theta}), \quad i = 1, \ldots, k,$$

where $\bar{\theta} = (1/k)\sum_{i=1}^k \theta_i$. By Theorem 1.3 in Section 1.2 of Chapter 1, $Z_1, \ldots, Z_k$ satisfy all the assumptions in Theorem 1.2 of Section 1.2.

For every $\boldsymbol{\theta}^* \in \mathbf{R}^k$ the maximum size and the minimum power of φ on $J(\boldsymbol{\theta}^*)$ are given by

$$\alpha(\varphi, \boldsymbol{\theta}^*) = \sup_{\boldsymbol{\theta} \in J(\boldsymbol{\theta}^*)} E_{\boldsymbol{\theta}}\varphi(\mathbf{X}) = 1 - \inf_{\boldsymbol{\mu} \in \boldsymbol{\Gamma}(\boldsymbol{\theta}^*)} \Pr\{\mathbf{Z} + \boldsymbol{\mu} \in \mathbf{D}(\lambda)\},$$

$$\beta(\varphi, \boldsymbol{\theta}^*) = \inf_{\boldsymbol{\theta} \in J(\boldsymbol{\theta}^*)} E_{\boldsymbol{\theta}}\varphi(\mathbf{X}) = 1 - \sup_{\boldsymbol{\mu} \in \boldsymbol{\Gamma}(\boldsymbol{\theta}^*)} \Pr\{\mathbf{Z} + \boldsymbol{\mu} \notin \mathbf{D}(\lambda)\},$$

where $\boldsymbol{\Gamma}(\boldsymbol{\theta}^*) = \{\boldsymbol{\mu}(\boldsymbol{\theta}) = (\theta_1 - \bar{\theta}, \ldots, \theta_k - \bar{\theta}) \mid \boldsymbol{\theta} \in J(\boldsymbol{\theta}^*)\}$. The α_r and β_r in (2.1) with $\boldsymbol{\Gamma} = \boldsymbol{\Gamma}(\boldsymbol{\theta}^*)$ become

$$\begin{aligned} \alpha_r(\boldsymbol{\theta}^*) &= \sum_{i=1}^r (\theta_{[i]}^* - \bar{\theta}^*) - \frac{2r(k-r)\epsilon}{k}, \\ \beta_r(\boldsymbol{\theta}^*) &= \sum_{i=1}^r (\theta_{[i]}^* - \bar{\theta}^*) + \frac{2r(k-r)\epsilon}{k}. \end{aligned} \tag{2.25}$$

Then $\boldsymbol{\mu}_r^M$ and $\boldsymbol{\mu}_r^s$ in (2.2) and (2.3) are reduced to

$$\mu_r^M(\boldsymbol{\theta}^*) = (\theta^*_{[r]} - \bar{\theta}^*) + \frac{2(k-2r+1)\epsilon}{k}, \qquad r = 1, \dots, k, \tag{2.26}$$

$$\mu_r^s(\boldsymbol{\theta}^*) = \begin{cases} (\theta^*_{[r]} - \bar{\theta}^*) - \dfrac{2(k-2r+1)\epsilon}{k}, & r = 1, \dots, s \\ -\dfrac{1}{k-s}\{ \sum_{i=1}^s (\theta^*_{[i]} - \bar{\theta}^*) - \dfrac{2s(k-s)\epsilon}{k}\}, & r = s+1, \dots, k. \end{cases} \tag{2.27}$$

Let $\boldsymbol{\mu}^M(\boldsymbol{\theta}^*) = (\mu_1^M(\boldsymbol{\theta}^*), \dots, \mu_k^M(\boldsymbol{\theta}^*))$ and $\boldsymbol{\mu}^s(\boldsymbol{\theta}^*) = (\mu_1^s(\boldsymbol{\theta}^*), \dots, \mu_k^s(\boldsymbol{\theta}^*))$. By (i) of Theorem 2.1 we have

$$\alpha(\varphi, \boldsymbol{\theta}^*) \leq 1 - \Pr\{\mathbf{Z} + \boldsymbol{\mu}^M(\boldsymbol{\theta}^*) \in \mathbf{D}(\lambda)\}. \tag{2.28}$$

If Condition 2 holds, then by (ii) of Theorem 2.1

$$\beta(\varphi, \boldsymbol{\theta}^*) \geq 1 - \Pr\{\mathbf{Z} + \boldsymbol{\mu}^s(\boldsymbol{\theta}^*) \in \mathbf{D}(\lambda)\}. \tag{2.29}$$

We note that $\beta_r(\boldsymbol{\theta}^*), r = 1, \dots, k$ satisfy Condition 1 and that Condition 2 is expressed as

$$\begin{aligned} \theta^*_{[r]} - \theta^*_{[r+1]} &\geq \frac{4\epsilon}{k}, \qquad r = 1, \dots, s-1 \\ \theta^*_{[s]} - \theta^*_{[r+1]} &\geq \frac{4\epsilon}{k}(r - s + 1), \quad r = s, \dots, k-1. \end{aligned} \tag{2.30}$$

Problem A. We first note that $\boldsymbol{\Gamma}(\boldsymbol{\theta}_i^*) = \boldsymbol{\Gamma}(\Theta^*_{i\mathrm{A}})$ holds for $i = 0, 1$. Let λ_α be a real number such that

$$1 - \Pr\{\mathbf{Z} + \boldsymbol{\mu}^M(\boldsymbol{\theta}_0^*) \in \mathbf{D}(\lambda_\alpha)\} = \alpha. \tag{2.31}$$

Then, by (2.28) the test φ based on $\mathbf{D}(\lambda_\alpha)$ is of level α. If Condition 2 holds, then the minimum power $\beta(\varphi, \boldsymbol{\theta}_1^*)$ satisfies

$$\beta(\varphi, \boldsymbol{\theta}_1^*) \geq 1 - \Pr\{\mathbf{Z} + \boldsymbol{\mu}^s(\boldsymbol{\theta}_1^*) \in \mathbf{D}(\lambda_\alpha)\}. \tag{2.32}$$

We see that $\boldsymbol{\mu}^M(\boldsymbol{\theta}_0^*) \prec \boldsymbol{\mu}^s(\boldsymbol{\theta}_1^*)$ is a sufficient condition for the unbiasedness of φ. When Condition 2 with $s = k$ holds, $\boldsymbol{\mu}^M(\boldsymbol{\theta}_0^*) \prec \boldsymbol{\mu}^k(\boldsymbol{\theta}_1^*)$ if and only if

$$\sum_{i=1}^r \{(\theta^*_{1[i]} - \bar{\theta}^*_1) - (\theta^*_{0[i]} - \bar{\theta}^*_0)\} \geq \frac{4r(k-r)\epsilon}{k}, \quad r = 1, \dots, k-1.$$

□

Problem B. We assume the additional conditions that f is centrally symmetric and logconcave, and $\mathbf{D}$ is centrally symmetric.

First we note

$$J(\Theta_{0B}^*) = \{\boldsymbol{\theta} \mid \theta_{[1]} - \theta_{[k]} \leq 2\epsilon\}, \qquad J(\Theta_{1B}^*) = \{\boldsymbol{\theta} \mid \theta_{[1]} - \theta_{[k]} > \frac{\eta}{2}\} \quad (\eta \geq 4\epsilon).$$

It is clear that $\boldsymbol{\Gamma}(\Theta_{0B}^*) = \boldsymbol{\Gamma}_{01}$ with $\delta = 2\epsilon$ and $\boldsymbol{\Gamma}(\Theta_{1B}^*) = \mathring{\boldsymbol{\Gamma}}_{11}$ (the interior of $\boldsymbol{\Gamma}_{11}$) with $\delta = \eta/2$, where $\boldsymbol{\Gamma}_{01}$ and $\boldsymbol{\Gamma}_{11}$ are given by (2.4). Since we treat continuous distributions, we can identify $\mathring{\boldsymbol{\Gamma}}_{11}$ with $\boldsymbol{\Gamma}_{11}$. Let $\alpha(\varphi) = \sup_{\boldsymbol{\theta} \in J(\Theta_{0B}^*)} E_{\boldsymbol{\theta}}\{\varphi(\mathbf{X})\}$ be the maximum size of φ on $J(\Theta_{0B}^*)$. Then

$$\alpha(\varphi) = \sup_{\boldsymbol{\theta}^* \in \Theta_{0B}^*} \alpha(\varphi, \boldsymbol{\theta}^*) = 1 - \inf_{\boldsymbol{\mu} \in \boldsymbol{\Gamma}_{01}} \Pr\{\mathbf{Z} + \boldsymbol{\mu} \in \mathbf{D}(\lambda)\}.$$

Let λ_α be a real number such that

$$1 - \min_{1 \leq r \leq [k/2]} \Pr\{\mathbf{Z} + \boldsymbol{\mu}_{01}(r) \in \mathbf{D}(\lambda_\alpha)\} = \alpha, \tag{2.33}$$

where $\boldsymbol{\mu}_{01}(r)$ is given by (2.5) with $\delta = 2\epsilon$. Then, by Theorem 2.2 φ based on $\mathbf{D}(\lambda_\alpha)$ is of level α. If Condition 2 holds, then a lower bound for the local minimum power $\beta(\varphi, \boldsymbol{\theta}^*)$ of φ on $J(\boldsymbol{\theta}^*)$, $\boldsymbol{\theta}^* \in \Theta_{1B}^*$, is given by the right hand side of (2.32) with $\boldsymbol{\theta}_1^* = \boldsymbol{\theta}^*$. The minimum power $\beta(\varphi)$ of φ on $J(\Theta_{1B}^*)$ is expressed as

$$\beta(\varphi) = \inf_{\boldsymbol{\theta}^* \in \Theta_1^*} \beta(\varphi, \boldsymbol{\theta}^*) = 1 - \sup_{\boldsymbol{\mu} \in \boldsymbol{\Gamma}_{11}} \Pr\{\mathbf{Z} + \boldsymbol{\mu} \in \mathbf{D}(\lambda_\alpha)\}.$$

Hence, from Theorem 2.3 it follows that

$$\beta(\varphi) = 1 - \Pr\{\mathbf{Z} + \boldsymbol{\mu}^* \in \mathbf{D}(\lambda_\alpha)\},$$

where

$$\boldsymbol{\mu}^* = (\eta/4, 0, \ldots, 0, -\eta/4).$$

□

Problem C. We assume that f is logconcave.

First we note that the critical point λ_α of a level α test φ is determined by (2.33) in which minimum ranges over $i = 1, \ldots, k$. It is easy to see that every $\boldsymbol{\theta}^* \in \Theta_{1C}^*$ satisfies Condition 2 with $s = k$. Hence, from (2.29) it follows that for every $\boldsymbol{\theta}^* \in \Theta_{1C}^*$

$$\beta(\varphi, \boldsymbol{\theta}^*) \geq 1 - \Pr\{\mathbf{Z} + \boldsymbol{\mu}^k(\boldsymbol{\theta}^*) \in \mathbf{D}(\lambda_\alpha)\},$$

where $\boldsymbol{\mu}^k(\boldsymbol{\theta}^*)$ is defined by (2.27). Let $\tilde{\boldsymbol{\theta}}^*$ be a parameter whose components satisfy $\tilde{\theta}_{[r]}^* - \tilde{\theta}_{[r+1]}^* = \eta, r = 1, \ldots, k-1$. Then we have $\tilde{\boldsymbol{\theta}}^* \in \bar{\Theta}_{1C}^*$ (the closure of Θ_{1C}^*) and $\boldsymbol{\mu}^k(\tilde{\boldsymbol{\theta}}^*) \prec \boldsymbol{\mu}^k(\boldsymbol{\theta}^*)$ for every $\boldsymbol{\theta}^* \in \Theta_{1C}^*$. It is readily checked that

$$\mu_r^k(\tilde{\boldsymbol{\theta}}^*) = \frac{1}{2}(k - 2r + 1)\eta - \frac{2}{k}(k - 2r + 1)\epsilon, \quad r = 1, \ldots, k.$$

Hence, from Theorem 1.2 in Section 1.2 we obtain

$$\beta(\varphi) \geq 1 - \Pr\{\mathbf{Z} + \boldsymbol{\mu}^k(\tilde{\boldsymbol{\theta}}^*) \in \mathbf{D}(\lambda_\alpha)\}.$$

Letting $\boldsymbol{\mu}^M(\mathbf{0})$ be given by (2.26) with $\boldsymbol{\theta}^* = \mathbf{0}$, we can see $\boldsymbol{\mu}_{01}(r) \prec \boldsymbol{\mu}^M(\mathbf{0}) \prec \boldsymbol{\mu}^k(\tilde{\boldsymbol{\theta}}^*)$ for $r = 1, \ldots, k$ and $\eta \geq 2\epsilon$, where $\boldsymbol{\mu}_{01}(r)$ is given by (2.5) with $\delta = 2\epsilon$. Thus φ is unbiased for every $\mathbf{D}(\lambda_\alpha)$. □

Problem D. We assume that f is logconcave. The critical point λ_α is the same with that in problem C. As easily seen, Condition 2 with $s = 1$ is satisfied for every $\boldsymbol{\theta}^* \in \boldsymbol{\Theta}^*_{1\mathrm{D}}$. Hence, $\boldsymbol{\mu}^1(\boldsymbol{\theta}^*) \prec \boldsymbol{\mu}(\boldsymbol{\theta}^*)$ holds for every $\boldsymbol{\theta}^* \in \boldsymbol{\Theta}^*_{1\mathrm{D}}$. Let $\hat{\boldsymbol{\theta}}^*$ be a parameter whose components satisfy $\hat{\theta}^*_{[1]} - \eta = \hat{\theta}^*_{[2]} = \cdots = \hat{\theta}^*_{[k]}$. Then we have $\hat{\boldsymbol{\theta}}^* \in \bar{\boldsymbol{\Theta}}^*_{1\mathrm{D}}$, and $\boldsymbol{\mu}^1(\hat{\boldsymbol{\theta}}^*) \prec \boldsymbol{\mu}^1(\boldsymbol{\theta}^*)$ holds for every $\boldsymbol{\theta}^* \in \boldsymbol{\Theta}^*_{1\mathrm{D}}$, where

$$\mu^1_1(\hat{\boldsymbol{\theta}}^*) = \frac{k-1}{k}(\eta - 2\epsilon)$$
$$\mu^1_r(\hat{\boldsymbol{\theta}}^*) = -\frac{1}{k}(\eta - 2\epsilon), \quad r = 2, \ldots, k.$$

Therefore, from (2.29) we obtain

$$\beta(\varphi) \geq 1 - \Pr\{\mathbf{Z} + \boldsymbol{\mu}^1(\hat{\boldsymbol{\theta}}^*) \in \mathbf{D}(\lambda_\alpha)\}.$$

□

We can get various tests by changing Schur-convex sets $\mathbf{D}(\lambda)$. Let us consider

$$\mathbf{D}_1(\lambda) = \{\mathbf{x} \mid \sum_{i=1}^k x_i^2 \leq \lambda\}, \qquad \mathbf{D}_2(\lambda) = \{\mathbf{x} \mid x_{[1]} - x_{[k]} \leq \lambda\}.$$

Let φ_i denote the test (2.24) based on $\mathbf{D}_i(\lambda), i = 1, 2$. Then φ_1 denotes a test based on $\sum_{i=1}^k (X_i - \bar{X})^2$, and φ_2 denotes a test based on the range $X_{[1]} - X_{[k]}$. When $f(\mathbf{x} - \boldsymbol{\theta})$ is a normal density with mean $\boldsymbol{\theta}$ and covariance matrix $\sigma^2 I$, where I is the $k \times k$ identity matrix and σ^2 is known, $(1/\sigma^2)\sum_{i=1}^k (X_i - \bar{X})^2$ has the noncentral $\mathcal{X}^2$-distribution with $k-1$ degrees of freedom and noncentrality parameter $(1/\sigma^2)\sum_{i=1}^k (\theta_i - \bar{\theta})^2$. Hence we can easily see that the minimum in (2.33) with $\mathbf{D}_1(\lambda)$ is attained by $\boldsymbol{\mu}_{01}([k/2])$. Giani and Finner (1991) and Chen, Lam and Xiong (1993) showed that under some additional conditions of f the minimum in (2.33) with $\mathbf{D}_2(\lambda)$ is also attained by $\boldsymbol{\mu}_{01}([k/2])$. Also, let us consider the test φ_3 based on $\mathbf{D}_3(\lambda) = \{\mathbf{x} \mid \max_{1 \leq i \leq k} |x_i| \leq \lambda\}$. The φ_3 based on the statistic $\max_{1 \leq i \leq k} |X_i - \bar{X}|$ is a test for a single outlier (Problem D). A list of Schur-convex sets $\mathbf{D}$ which produce various tests for outliers is given in Section 1.2 of Chapter 1.

Finally, let us consider robust testing of location parameters with an unknown scale parameter. Let $X_1, \ldots, X_k$ be exchangeable random variables with a joint density $\sigma^{-k} f((\mathbf{x}-\boldsymbol{\theta})/\sigma)$ where f is Schur-concave and $\sigma > 0$ is unknown. Let V be an estimator of σ independent of $\mathbf{X}$. We wish to test the problems A, B, C and D. To this end it is natural to use the following tests φ based on $((X_1 - \bar{X})/V, \ldots, (X_k - \bar{X})/V)$ and $\mathbf{D}(\lambda)$:

$$\varphi(\mathbf{X}, V) = 0, \quad 1 \quad if \quad (\frac{X_1 - \bar{X}}{V}, \ldots, \frac{X_k - \bar{X}}{V}) \in, \quad \notin \mathbf{D}(\lambda), \tag{2.34}$$

where $\mathbf{D}(\lambda)$ is Schur-convex. It is easy to see that when $\mathbf{Z} + \boldsymbol{\mu}$ is replaced by $(\mathbf{Z} + \boldsymbol{\mu})/V$, all the results for the problems A, B, C and D are valid.

CHAPTER

3

Robustness of Rank Tests for k-Sample Approximate Equality in the Presence of Gross Errors

The probabilities that k-sample rank statistics take the values in an arbitrary specified Schur-convex set in $\mathbf{R}^k$ are considered under a k-sample local asymptotic gross error model. When the underlying distributions vary in gross error neighborhoods shrinking at the rate of $n^{-1/2}$, lower and upper bounds for limiting values of the probabilities are obtained by using majorization methods on hyperplanes. These bounds enable us to construct asymptotic level α rank tests for k-sample approximate equality whose acceptance regions are Schur-convex sets, and to give lower bounds for their asymptotic minimum powers. Robustness of rank tests is studied based on the lower bounds and a measure of asymptotic efficiency is proposed. A number of rank tests are obtained from various choices of the Schur-convex sets.

3.1. Introduction.

Rieder (1981) extended the classical nonparametric hypotheses of symmetry and equality of distributions to hypotheses of approximate symmetry and approximate equality by allowing for gross errors described by ϵ-contamination and total variation neighborhoods. He investigated asymptotic properties of sequences of one- and two-sample rank tests in his local framework (Rieder 1978). His results reveal that the classical optimum rank tests are sensitive to the deviation from the assumptions of rather stringent symmetry and equality, and hence their use are dangerous for practical situations where the deviation may occur.

In this chapter, we shall extend the classical nonparametric hypotheses of k-sample equality to hypotheses of k-sample approximate equality in Rieder's (1981) way, and explore asymptotic properties of sequences of rank tests under a k-sample local asymptotic gross errors model. Because we have to treat k gross error neighborhoods simultaneously, there arise some difficulties unlike one- and two-sample cases such as asymptotic maximum sizes and minimum powers of k sample rank tests are not obtained. To cope with these difficulties, we shall try to apply the majorization methods on hyperplanes in Section 2.2 of Chapter 2. In these applications it is seen that the majorization methods are quite effectively utilized for constructing asymptotic level α rank tests and giving the lower bounds for thier asymptotic minimum powers. Based on these bounds we shall consider robustness of rank tests and propose a measure of asymptotic efficiency (=ARE*). The consideration shows that ARE* is similar to asymptotic relative efficiency in one- and two-sample cases and rank tests with truncated scores are recommendable for practical use. We should point out that new ideas and devices of majorization developed in this paper are essential for overcoming some peculiar difficulties arising in k-sample case.

In section 3.2 we shall present a k-sample asymptotic model. The gross error neighborhoods described in terms of ϵ-contamination and total variation shrink at the rate of $n^{-1/2}$. It is shown that the limiting multivariate normal distribution of k-sample rank statistics is determined by the limit of centering vectors.

In section 3.3 we shall obtain the limiting values of probabilities that k-sample rank statistics take the values in an arbitrarily specified Schur-convex set in $\mathbf{R}^k$ under the asymptotic model. We wish to get a lower and an upper bounds for the limiting probabilities when the underlying distributions range over gross error neighborhoods. It is seen from Theorem 1.2 (Kimura and Kakiuchi, 1989) in Section 1.2 of Chapter 1 that the limiting probability is Schur-concave in the limiting centering vector. Therefore, the problem is to seek a vector (a smaller one is better) majorizing the set of all the limits of centering vectors and a vector (a larger one is better) majorized by the set. To this end, we shall apply the majorization methods on hyperplanes in Section 2.2.

In section 3.4 we shall formulate asymptotic testing problems of k-sample approximate equality and consider rank tests whose acceptance regions are Schur-convex. A number of rank tests are obtained from various choices of the Schur-convex sets. The use of the results in Section 3 enables us to construct asymptotic level α rank tests and to get lower bounds for their asymptotic minimum powers. Based on these lower bounds we shall study robustness of rank tests which includes ARE^*. The consideration shows that the results similar to one- and two-sample cases are derived and rank tests with truncated scores are recommendable.

The Appendix presents the proofs of Lemmas 3.1 and 3.2.

3.2. Framework of asymptotic study.

Let $X_{i1}, \dots , X_{in}$ be independent random variables distributed with continuous distribution functions $G_{i1}(x), \dots , G_{in}(x)$ for $i = 1, \dots , k$, respectively. Let R_{ij} be the rank of X_{ij} among all $N(= kn)$ random variables $X_{11}, X_{12}, \dots , X_{kn}$.

We consider the following k-sample rank statistics

$$T_{Ni} = \frac{1}{n}\sum_{j=1}^{n} a_N(R_{ij}), \qquad i = 1, \dots , k, \tag{3.1}$$

where the scores $a_N(r)$ are generated by a function $a : (0, 1) \to (-\infty, +\infty)$ in either one of the two ways,

$$\begin{aligned} a_N(r) &= a(\frac{r}{N+1}), \qquad & r = 1, \dots , N, \\ a_N(r) &= E\big(a(U_N^{(r)})\big), \qquad & r = 1, \dots , N, \end{aligned} \tag{3.2}$$

where $U_N^{(r)}$ denotes the r-th order statistic in a random sample of size N from the uniform distribution on $(0, 1)$. The scores generating function a is assumed to satisfy the following conditions throughout this chapter:

(i) a is nondecreasing and nonconstant, absolutely continuous inside $(0, 1)$ and

$$\int_0^1 t^{1/2}(1-t)^{1/2}\, da(t) < \infty.$$

(ii) a is Lipschitz bounded of order 1 on $[t_0, 1 - t_0]$, concave on $(0, t_0)$ and convex on $(1 - t_0, 1)$ for some $t_0 \in (0, 1/2\,]$.

Let $\{\, F_\theta \mid \theta \in \boldsymbol{\Theta}\}$ be a parametric family of continuous distribution functions, whose parameter space $\boldsymbol{\Theta}$ is a subset of $\mathbf{R}$ and contains zero in its interior, and which satisfies the following conditions:

(i) F_θ is absolutely continuous with respect to F_0 for every $\theta \in \boldsymbol{\Theta}$.

(ii) There exists a function $\Lambda \in L_2(dF_0)$ such that

$$\frac{f_\theta^{1/2} - 1}{\theta} \to \frac{1}{2}\Lambda \quad \text{in} \quad L_2(dF_0) \quad \text{as} \quad \theta \to 0, \tag{3.3}$$

where f_θ denotes the density of F_θ with respect to F_0.

Let $\mathcal{M}_c$ denote the set of all continuous distribution functions on $\mathbf{R}$, and let

$$\underline{G}(x) = \big((1-\epsilon)F_\theta(x) + \epsilon + \delta\big) \wedge 1, \tag{3.4}$$

$$\bar{G}(x) = \big((1-\epsilon)F_\theta(x) - \delta\big)^+, \tag{3.5}$$

where $a \wedge 1 = \min(a, 1)$, $t^+ = \max(0, t)$ and $\epsilon \geq 0$ and $\delta \geq 0$ are such that $\epsilon + \delta > 0$. The following gross error neighborhood $\mathcal{P}(\theta; \epsilon, \delta)$ of F_θ is a generalization of ϵ-contamination and total variation neighborhoods:

$$\mathcal{P}(\theta; \epsilon, \delta) = \{\, G \in \mathcal{M}_c \mid \bar{G}(x) \leq G(x) \leq \underline{G}(x) \text{ for every } x \in [-\infty, +\infty] \,\}.$$

Let a probability measure G be identified with its distribution function, that is, $G(x) = G((-\infty, x\,])$. The symbol $\otimes_{i=1}^{k} \otimes_{j=1}^{n} G_{ij}$ stands for the stochastic product of $G_{11}, G_{12}, \ldots, G_{kn}$, which denotes the joint distribution of $X_{11}, X_{12}, \ldots, X_{kn}$, and we put

$$\mathbf{W}_N = \otimes_{i=1}^{k} \otimes_{j=1}^{n} G_{ij}.$$

The k-sample local asymptotic gross error neighborhood $\mathcal{P}^{(N)}(\boldsymbol{\theta}_n; \epsilon_n, \delta_n)$ shrinking at order $n^{-1/2}$ is defined by

$$\mathcal{P}^{(N)}(\boldsymbol{\theta}_n; \epsilon_n, \delta_n) = \{\, \mathbf{W}_N \mid G_{ij} \in \mathcal{P}(\theta_{ni}; \epsilon_n, \delta_n),\ i = 1, \ldots, k; j = 1, \ldots, n \},$$

where $\boldsymbol{\theta}_n = (\theta_{n1}, \ldots, \theta_{nk})$, ϵ_n and δ_n are given by

$$\theta_{ni} = n^{-1/2}\theta_i \ (i = 1, \ldots, k), \quad \epsilon_n = n^{-1/2}\epsilon, \quad \delta_n = n^{-1/2}\delta,$$

for all positive integers n and $\theta_i \in \boldsymbol{\Theta}$.

Let us define

$$\mu_{Ni} = \int_{-\infty}^{+\infty} a(G_N(x))\, dG_{ni}(x), \quad i = 1, \ldots, k, \tag{3.6}$$

$$A_a^2 = \int_0^1 \left(a(t) - \bar{a}\right)^2 dt, \tag{3.7}$$

where

$$G_{ni}(x) = \frac{1}{n}\sum_{j=1}^{n} G_{ij}(x), \quad G_N(x) = \frac{1}{k}\sum_{i=1}^{k} G_{ni}(x), \quad \bar{a} = \int_0^1 a(t)\, dt.$$

The following theorem of asymptotic normality of k-sample rank statistics under $\mathbf{W}_N \in \mathcal{P}^{(N)}(\boldsymbol{\theta}_n; \epsilon_n, \delta_n)$ is basic.

Theorem 3.1. *For every $\mathbf{W}_N \in \mathcal{P}^{(N)}(\boldsymbol{\theta}_n; \epsilon_n, \delta_n)$, the statistics*

$$n^{1/2}(T_{N1} - \mu_{N1}, \ldots, T_{Nk} - \mu_{Nk})/A_a$$

has the limiting normal distribution with mean vector zero and covariance matrix $\Sigma = (\sigma_{ij})$, *where*

$$\sigma_{ij} = \begin{cases} \dfrac{k-1}{k}, & i=j=1,\dots,k \\ -\dfrac{1}{k}, & i \neq j; i,j=1,\dots,k, \end{cases} \tag{3.8}$$

Proof. For every $G_{ij} \in \mathcal{P}(\theta_{in};\epsilon,\delta_n)$ and $G_{hl} \in \mathcal{P}(\theta_{hn};\epsilon,\delta_n)$, $i,h=1,\dots,k; j,l=1,\dots,n$, we have

$$\limsup_{n\to\infty} n^{1/2} \sup_x \mid G_{ij}(x) - G_{hl}(x) \mid < \infty.$$

By Theorem 2.4 of Hajéck (1968) and Theorem 1 of Hoeffding (1973) (also see Corollary 2.5.3 and Theorem 2.6.1 in Puri and Sen, 1985) it is easy to see that for every nonzero $\boldsymbol{\lambda} = (\lambda_1,\dots,\lambda_k)'$ and for every $\mathbf{W}_N \in \mathcal{P}(\boldsymbol{\theta};\epsilon_n,\delta_n)$ the distribution of $n^{1/2}\sum_{i=1}^k \lambda_i(T_{Ni} - \mu_{Ni})/A_a$ converges in law to the normal distribution with mean zero and variance $\boldsymbol{\lambda}'\boldsymbol{\Sigma}\boldsymbol{\lambda}$. This implies that the theorem holds. □

We are interested in the class of all asymptotic distributions of the statistics

$$n^{1/2}(T_{N1} - \bar{a},\dots,T_{Nk} - \bar{a})/A_a \tag{3.9}$$

when $\mathbf{W}_N$ ranges over $\mathcal{P}^{(N)}(\boldsymbol{\theta}_n;\epsilon_n,\delta_n)$. By noting that

$$\begin{aligned} & n^{1/2}(T_{N1} - \bar{a},\dots,T_{Nk} - \bar{a})/A_a \\ & = n^{1/2}(T_{N1} - \mu_{N1},\dots,T_{Nk} - \mu_{Nk})/A_a + n^{1/2}(\mu_{N1} - \bar{a},\dots,\mu_{Nk} - \bar{a})/A_a, \end{aligned}$$

we obtain from Theorem 3.1 that the asymptotic distribution of (3.9) entirely depends on the limiting vector of

$$n^{1/2}(\mu_{N1} - \bar{a},\dots,\mu_{Nk} - \bar{a})/A_a$$

under $\mathbf{W}_N \in \mathcal{P}^{(N)}(\boldsymbol{\theta}_n;\epsilon_n,\delta_n)$. Hence, we let

$$\boldsymbol{\Omega}(\boldsymbol{\theta}) = \{ \lim_{n\to\infty} n^{1/2}(\mu_{N1} - \bar{a},\dots,\mu_{Nk} - \bar{a}) \mid W_N \in \mathcal{P}^{(N)}(\boldsymbol{\theta}_n;\epsilon_n,\delta_n),\ n=1,2,\dots \},$$

and we wish to seek the smallest of all vectors majorizing $\boldsymbol{\Omega}(\boldsymbol{\theta})$ and the largest of all vectors majorized by $\boldsymbol{\Omega}(\boldsymbol{\theta})$ in the sense of majorization preorder. Since $\sum_{i=1}^k \mu_{Ni} = k\bar{a}$, we have $\boldsymbol{\Omega}(\boldsymbol{\theta}) \subset \boldsymbol{\Omega}$, where $\boldsymbol{\Omega}$ is the $k-1$ dimensional hyperplane *i.e.*,

$$\boldsymbol{\Omega} = \{\boldsymbol{\mu} = (\mu_1,\dots,\mu_k) \in \mathbf{R}^k \mid \sum_{i=1}^k \mu_i = 0\}.$$

Hence, to obtain such the smallest and the the largest vectors, we can make use of majorization methods on hyperplanes in Section 2.2 of Chapter 2.

3.3. Majorization inequalities for k-sample rank statistics.

Let $\mathcal{J}(m)$ be the family of all subsets of the set $J = \{1, \cdots, k\}$ consisting of m elements and let $A(m) \in \mathcal{J}(m)$. We put

$$K_n(t) = \frac{m}{k}t + \frac{k-m}{k}\underline{G}_{n,J-A(m)}(\bar{G}_{n,A(m)}^{-1}(t)), \quad 0 < t < 1, \tag{3.10}$$

where

$$\bar{G}_{n,A(m)}(x) = \frac{1}{m}\sum_{i\in A(m)} \bar{G}_{ni}(x), \tag{3.11}$$

$$\underline{G}_{n,J-A(m)}(x) = \frac{1}{k-m}\sum_{i\in J-A(m)} \underline{G}_{ni}(x), \tag{3.12}$$

$$\bar{G}_{ni}(x) = \big((1-\epsilon_n)F_{\theta_{ni}}(x) - \delta_n\big)^+ + (\epsilon_n + \delta_n)\lambda_{+\infty}(x), \tag{3.13}$$

$$\underline{G}_{ni}(x) = \big((1-\epsilon_n)F_{\theta_{ni}}(x) + \epsilon_n + \delta_n\big) \wedge 1, \tag{3.14}$$

for any $x \in [-\infty, +\infty]$ and $i = 1, \ldots, k$, and $\lambda_{+\infty}(x)$ has the point mass 1 at $x = +\infty$.

We obtain the following two lemmas. Their proofs are given in Appendix.

Lemma 3.1. *It holds that*

$$\sup\{\sum_{i\in\mathbf{A}(m)} \mu_{Ni} \mid \mathbf{W}_N \in \mathcal{P}^{(N)}(\boldsymbol{\theta}_n; \epsilon_n, \delta_n)\} = m\int_0^1 a\big(K_n(t)\big)\,dt.$$

Lemma 3.2. *It holds that*

(i) $$\lim_{n\to\infty} n^{1/2}\sup\{\sum_{i\in\mathbf{A}(m)} (\mu_{Ni} - \bar{a}) \mid \mathbf{W}_N \in \mathcal{P}^{(N)}(\boldsymbol{\theta}_n; \epsilon_n, \delta_n)\}$$

$$= \sum_{i\in\mathbf{A}(m)} (\theta_i - \bar{\theta})\int_0^1 \Lambda\big(F_0^{-1}(t)\big)a(t)\,dt + \frac{m(k-m)}{k}(\epsilon + 2\delta)\big(a(1) - a(0)\big),$$

(ii) $\lim_{n\to\infty} n^{1/2} \inf \{ \sum_{i \in \mathbf{A}(m)} (\mu_{Ni} - \bar{a}) \mid \mathbf{W}_N \in \mathcal{P}^{(N)}(\boldsymbol{\theta}_n; \epsilon_n, \delta_n)\}$

$$= \sum_{i\in\mathbf{A}(m)} (\theta_i - \bar{\theta}) \int_0^1 \Lambda\big(F_0^{-1}(t)\big)a(t)\,dt - \frac{m(k-m)}{k}(\epsilon + 2\delta)\big(a(1) - a(0)\big),$$

where $a(0) = a(+0)$, $a(1) = a(1-0)$ *and* $\bar{\theta} = \frac{1}{k}\sum_{i=1}^k \theta_i$.

When a is unbounded, $\boldsymbol{\Omega}(\boldsymbol{\theta})$ is not bounded. Therefore, we assume that a is bounded, and use majorization methods on hyperplanes in Section 2.2 of Chapter 2.

Let us define α_i and β_i, $i = 1, \dots, k$, by

$$\alpha_i = \sum_{j=1}^{i}(\theta_{[j]} - \bar{\theta}) \int_0^1 \Lambda\big(F_0^{-1}(t)\big)a(t)\,dt - \frac{i(k-i)}{k}(\epsilon + 2\delta)\big(a(1) - a(0)\big), \tag{3.15}$$

$$\beta_i = \sum_{j=1}^{i}(\theta_{[j]} - \bar{\theta}) \int_0^1 \Lambda\big(F_0^{-1}(t)\big)a(t)\,dt + \frac{i(k-i)}{k}(\epsilon + 2\delta)\big(a(1) - a(0)\big), \tag{3.16}$$

where $\theta_{[1]} \geq \cdots \geq \theta_{[k]}$ denote the components of $\boldsymbol{\theta}$ in decreasing order. Also, let $\boldsymbol{\mu}_a^M(\boldsymbol{\theta})$ and $\boldsymbol{\mu}_a^s(\boldsymbol{\theta})$, $s = 1, \dots, k$, be defined by

$$\boldsymbol{\mu}_a^M(\boldsymbol{\theta}) = (\beta_1 - \beta_0, \beta_2 - \beta_1, \dots, \beta_k - \beta_{k-1}),$$
$$\boldsymbol{\mu}_a^s(\boldsymbol{\theta}) = (\alpha_1 - \alpha_0, \dots, \alpha_s - \alpha_{s-1}, -\frac{\alpha_s}{k-s}, \dots, -\frac{\alpha_s}{k-s}).$$

Let $\mu_{ai}^M(\boldsymbol{\theta})$ and $\mu_{ai}^s(\boldsymbol{\theta})$ denote the i-th components of $\boldsymbol{\mu}_a^M(\boldsymbol{\theta})$ and $\boldsymbol{\mu}_a^s(\boldsymbol{\theta})$, respectively. Then we have

$$\mu_{ai}^M(\boldsymbol{\theta}) = (\theta_{[i]} - \bar{\theta}) \int_0^1 \Lambda\big(F_0^{-1}(t)\big)a(t)\,dt + \frac{1}{k}(k - 2i + 1)(\epsilon + 2\delta)\big(a(1) - a(0)\big), \quad i = 1, \dots, k, \tag{3.17}$$

$$\mu_{ai}^s(\boldsymbol{\theta}) = \begin{cases} (\theta_{[i]} - \bar{\theta}) \int_0^1 \Lambda\big(F_0^{-1}(t)\big)a(t)\,dt \\ \qquad - \frac{1}{k}(k - 2i + 1)(\epsilon + 2\delta)\big(a(1) - a(0)\big), & i = 1, \dots, s \\ -\frac{1}{k-s}\sum_{j=1}^{s}(\theta_{[j]} - \bar{\theta}) \int_0^1 \Lambda\big(F_0^{-1}(t)\big)a(t)\,dt \\ \qquad + \frac{s}{k}(\epsilon + 2\delta)\big(a(1) - a(0)\big), & i = s+1, \dots, k. \end{cases} \tag{3.18}$$

According to Lemmas 2.1 and 2.2 in Section 2.2, $\boldsymbol{\mu}_a^M(\boldsymbol{\theta})$ and $\boldsymbol{\mu}_a^s(\boldsymbol{\theta})$ are the smallest of all parameters majorizing $\boldsymbol{\Omega}(\boldsymbol{\theta})$ and the largest of all parameters majorized by $\boldsymbol{\Omega}(\boldsymbol{\theta})$, respectively, if Conditions 1 and 2 in Section 2.2 are satisfied. The following condition corresponds to Condition 2.

Condition I(s):

$$\begin{cases} (\theta_{[i]} - \theta_{[i+1]}) \int_0^1 \Lambda(F_0^{-1}(t))a(t)\,dt \\ \qquad\qquad -\dfrac{2}{k}(\epsilon + 2\delta)\big(a(1) - a(0)\big) \geq 0, \quad i = 1, \ldots, s-1, \\ (\theta_{[s]} - \theta_{[i+1]}) \int_0^1 \Lambda(F_0^{-1}(t))a(t)\,dt \\ \qquad\qquad -\dfrac{2}{k}(i - s + 1)(\epsilon + 2\delta)\big(a(1) - a(0)\big) \geq 0, \quad i = s, \ldots, k-1 \end{cases}$$

where s $(1 \leq s \leq k)$ is some fixed integer.

This condition I(s) implies that $\int_0^1 \Lambda(F_0^{-1}(t))a(t)\,dt \geq 0$. Hence, Condition 1 is satisfied whenever Condition I(s) holds.

Theorem 3.2. *Let $\mathbf{D} \subset \mathbf{R}^k$ be a Schur-convex set and let $\{\mathbf{D}_n\}$ be a sequence of subsets of $\mathbf{R}^k$ converging to $\mathbf{D}$ as $n \to \infty$. Let a random vector $\mathbf{Z} = (Z_1, \ldots, Z_k)$ satisfying $\sum_{i=1}^k Z_i = 0$ have the normal distribution with mean vector zero and covariance matrix Σ given by (2.8). If $-\infty < a(0) < a(1) < +\infty$, then it holds that*

(i) *for every $\mathbf{W}_N \in \mathcal{P}^{(N)}(\boldsymbol{\theta}_n; \epsilon_n, \delta_n)$,*

$$\lim_{n\to\infty} \Pr\{n^{1/2}(T_{N1} - \bar{a}, \ldots, T_{Nk} - \bar{a})/A_a \in \mathbf{D}_n\} \geq \Pr\{\mathbf{Z} + \boldsymbol{\mu}_a^M(\boldsymbol{\theta})/A_a \in \mathbf{D}\},$$

(ii) *if Condition I(s) are satisfied, then for every $\mathbf{W}_N \in \mathcal{P}^{(N)}(\boldsymbol{\theta}_n; \epsilon_n, \delta_n)$*

$$\lim_{n\to\infty} \Pr\{n^{1/2}(T_{N1} - \bar{a}, \ldots, T_{Nk} - \bar{a})/A_a \in \mathbf{D}_n\} \leq \Pr\{\mathbf{Z} + \boldsymbol{\mu}_a^s(\boldsymbol{\theta})/A_a \in \mathbf{D}\},$$

where $\boldsymbol{\mu}_a^M(\boldsymbol{\theta})$ and $\boldsymbol{\mu}_a^s(\boldsymbol{\theta})$ are given by (3.17) and (3.18), respectively.

Proof. Let $\mathbf{W}_N \in \mathcal{P}^{(N)}(\boldsymbol{\theta}_n; \epsilon_n, \delta_n)$. Then it follows from Theorem 3.1 that

$$\begin{aligned} &\lim_{n\to\infty} \Pr\{n^{1/2}(T_{N1} - \bar{a}, \ldots, T_{Nk} - \bar{a})/A_a \in \mathbf{D}_n\} \\ &= \lim_{n\to\infty} \Pr\{n^{1/2}(T_{N1} - \mu_{N1}, \ldots, T_{Nk} - \mu_{Nk})/A_a \\ &\qquad\qquad + n^{1/2}(\mu_{N1} - \bar{a}, \ldots, \mu_{N1} - \bar{a})/A_a \in \mathbf{D}_n\} \\ &= \Pr\{\mathbf{Z} + (\mu_1, \ldots, \mu_k)/A_a \in \mathbf{D}\}, \end{aligned}$$

where $\mathbf{Z}$ has normal distribution with mean vector $\mathbf{0}$ and covariance matrix Σ, and $\mu_i = \lim_{n\to\infty} n^{1/2}(\mu_{Ni} - \bar{a})$, $i = 1, \ldots, k$. Since $Z_1, \ldots, Z_k$ are exchangeable with

$\sum_{i=1}^{k} Z_i = 0$ and $(Z_1, \ldots, Z_{k-1})$ has a Schur-concave density (see Marshall and Olkin, 1979, page 300), it follows from Theorem 1.2 in Section 1.2 of Chapter 1 that $\Pr\{\mathbf{Z} + (\mu_1, \ldots, \mu_k)/A_a \in \mathbf{D}\}$ is the Schur-concave function of $(\mu_1, \ldots, \mu_k)$. Hence, from Theorem 2.1 in Section 2.2 of Chapter 2 we obtain the theorem. □

3.4. Asymptotic testing problems for k-sample approximate equality.

Let $\tau \in (0, +\infty)$ be a constant satisfying

$$0 < \frac{\epsilon + 2\delta}{\tau} < \int \Lambda^+ \, dF_0 \tag{3.19}$$

for given constants $\epsilon \geq 0$ and $\delta \geq 0$ such that $\epsilon + \delta > 0$. Here we note that $\theta_{[1]} - \theta_{[k]} \geq \tau$ implies that

$$\mathcal{P}(\theta_{ni}; \epsilon_n, \delta_n) \cap \mathcal{P}(\theta_{nj}; \epsilon_n, \delta_n) = \phi$$

holds for at least one pair (i, j) and sufficiently large n, which is an asymptotic disjointness condition (see Rieder, 1978). For each $\boldsymbol{\theta} \in \mathbf{R}^k$, let

$$\mathcal{W}(\boldsymbol{\theta}) = \{(\mathbf{W}_N) \mid \mathbf{W}_N \in \mathcal{P}^{(N)}(\boldsymbol{\theta}_n; \epsilon_n, \delta_n) \quad \text{for } n = 1, 2, \ldots\}.$$

For some specified constant $\xi_0 \geq 0$ let us consider two special parameter sets:

$$\boldsymbol{\Theta}_0 = \{\, \boldsymbol{\theta} \mid \theta_{[1]} - \theta_{[k]} \leq \xi_0 \} \quad \text{and} \quad \boldsymbol{\Theta}_1 \text{ is a subset of } \{\, \boldsymbol{\theta} \mid \theta_{[1]} - \theta_{[k]} \geq \xi_0 + \tau \}.$$

We are interested in the following asymptotic testing problem:

$$\mathbf{H}_0 : (\mathbf{W}_N) \in \bigcup_{\boldsymbol{\theta} \in \boldsymbol{\Theta}_0} \mathcal{W}(\boldsymbol{\theta}) \quad \textbf{vs.} \quad \mathbf{H}_1 : (\mathbf{W}_N) \in \bigcup_{\boldsymbol{\theta} \in \boldsymbol{\Theta}_1} \mathcal{W}(\boldsymbol{\theta}). \tag{3.20}$$

When $\xi_0 = 0$, we have $\boldsymbol{\Theta}_0 = \boldsymbol{\Theta}_0^* = \{\boldsymbol{\theta} \mid \theta_1 = \cdots = \theta_k\}$. Hence the null hypothesis $\mathbf{H}_0$ expresses an asymptotic approximate equality of k samples. Also, when $\xi_0 > 0$, $\mathbf{H}_0$ represents an extended asymptotic approximate equality of k samples.

For the problem (3.20) we consider the following sequences $(\varphi_{N,a})$ of k-sample rank tests based on $T_{N1}, \ldots, T_{Nk}$ and Schur-convex sets $\mathbf{D}(\lambda_{N,a}) \in \mathbf{R}^k$:

$$\varphi_{N,a}(T_{N1}, \ldots, T_{Nk}) = 0,\ 1 \ if \ n^{1/2}(T_{N1} - \bar{a}, \ldots, T_{Nk} - \bar{a})/A_a \in, \ \notin \mathbf{D}(\lambda_{N,a}). \tag{3.21}$$

The $\lambda_{N,a}$ usually means critical values of $\varphi_{N,a}$. The maximum size $\alpha_n(\varphi_{N,a}; \boldsymbol{\theta})$ and the minimum power $\beta_n(\varphi_{N,a}; \boldsymbol{\theta})$ of $\varphi_{N,a}$ at each $\boldsymbol{\theta} \in \boldsymbol{\Theta}_0$ and $\boldsymbol{\theta} \in \boldsymbol{\Theta}_1$ are defined by

$$\alpha_n(\varphi_{N,a}; \boldsymbol{\theta}) = \sup\{E_{\mathbf{W}_N}(\varphi_{N,a}) \mid \mathbf{W}_N \in \mathcal{P}^{(N)}(\boldsymbol{\theta}_n; \epsilon_n, \delta_n)\},$$
$$\beta_n(\varphi_{N,a}; \boldsymbol{\theta}) = \inf\{E_{\mathbf{W}_N}(\varphi_{N,a}) \mid \mathbf{W}_N \in \mathcal{P}^{(N)}(\boldsymbol{\theta}_n; \epsilon_n, \delta_n)\}.$$

By investigating limiting values of $\alpha_n(\varphi_{N,a};\boldsymbol{\theta})$ and $\beta_n(\varphi_{N,a};\boldsymbol{\theta})$ we can construct asymptotic level α rank tests $(\varphi_{N,a})$ and evaluate their asymptotic minimum powers.

3.4.1. Determining critical values.

By the definition of $(\varphi_{N,a})$ and (i) of Theorem 3.2, we can immediately obtain the following lemma.

Lemma 3.3. *If $\lambda_{N,a} \to \lambda_a$ and $\mathbf{D}(\lambda_{N,a}) \to \mathbf{D}(\lambda_a)$ as $n \to \infty$, then for every $\boldsymbol{\theta} \in \Theta_0$ it holds that*

$$\lim_{n\to\infty} \alpha_n(\varphi_{N,a};\boldsymbol{\theta}) \le 1 - \Pr\{\mathbf{Z} + \boldsymbol{\mu}_a^M(\boldsymbol{\theta})/A_a \in \mathbf{D}(\lambda_a)\}, \tag{3.22}$$

where $\mathbf{Z}$ and $\boldsymbol{\mu}_a^M(\boldsymbol{\theta})$ are given in Theorem 3.2.

When $\xi_0 = 0$, we have

$$\lim_{n\to\infty} \alpha_n(\varphi_{N,a};\boldsymbol{\theta}) \le 1 - \Pr\{\mathbf{Z} + \boldsymbol{\mu}_a^*/A_a \in \mathbf{D}(\lambda_a)\},$$

where $\boldsymbol{\mu}_a^* = (\mu_{a1}^*, \dots, \mu_{ak}^*) = \boldsymbol{\mu}_a^M(\mathbf{0})$ is given by

$$\mu_{ai}^* = \frac{1}{k}(k - 2i + 1)(\epsilon + 2\delta)\big(a(1) - a(0)\big), \quad i = 1, \dots, k. \tag{3.23}$$

Also, when $\xi_0 > 0$, by Theorem 2.2 in Section 2.3 of Chapter 2 we have

$$\inf_{\boldsymbol{\theta}\in\Theta_0} \Pr\{\mathbf{Z} + \boldsymbol{\mu}_a^M(\boldsymbol{\theta})/A_a \in \mathbf{D}(\lambda_a)\} = \min_{1\le l\le k-1} \Pr\{\mathbf{Z} + \boldsymbol{\mu}_a^M(\boldsymbol{\theta}_0(l))/A_a \in \mathbf{D}(\lambda_a)\},$$

where

$$\boldsymbol{\theta}_0(l) = \Bigg(\underbrace{\frac{(k-l)\xi_0}{k}, \dots, \frac{(k-l)\xi_0}{k}}_{l\text{-times}}, \underbrace{-\frac{l\xi_0}{k}, \dots, -\frac{l\xi_0}{k}}_{(k-l)\text{-times}}\Bigg), \quad l = 1, \dots, k-1.$$

Hence

$$\lim_{n\to\infty} \sup_{\boldsymbol{\theta}\in\Theta_0} \alpha_n(\varphi_{N,a};\boldsymbol{\theta}) \le 1 - \min_{1\le l\le k-1} \Pr\{\mathbf{Z} + \boldsymbol{\mu}_a^M(\boldsymbol{\theta}_0(l))/A_a \in \mathbf{D}(\lambda_a)\}.$$

For $\alpha \in (0,1)$ let $\lambda_{\alpha,a}$ be determined by

$$1 - \Pr\{\mathbf{Z} + \boldsymbol{\mu}_a^*/A_a \in \mathbf{D}(\lambda_{\alpha,a})\} = \alpha, \text{ when } \xi_0 = 0 \tag{3.24}$$

$$1 - \min_{1 \le l \le k-1} \Pr\left\{\mathbf{Z} + \boldsymbol{\mu}_a^M(\boldsymbol{\theta}_0(l))/A_a \in \mathbf{D}(\lambda_{\alpha,a})\right\} = \alpha, \text{ when } \xi_0 > 0. \tag{3.25}$$

The following theorem shows that an asymptotic level α rank test $(\varphi_{N,a})$ for the problem (3.20) can be constructed by (3.24) and (3.25).

Theorem 3.3. *Let $(\varphi_{N,a})$ be the sequences of (3.21) with the values $\lambda_{N,a}$ satisfying $\lim_{n\to\infty} \lambda_{N,a} = \lambda_{\alpha,a}$ and $\lim_{n\to\infty} \mathbf{D}(\lambda_{N,a}) = \mathbf{D}(\lambda_{\alpha,a})$, where $\lambda_{\alpha,a}$ is given by (3.24) or (3.25) according as $\xi_0 = 0$ or $\xi_0 > 0$. Then it holds that*

$$\lim_{n\to\infty} \sup_{\boldsymbol{\theta}\in\Theta_0} \alpha_n(\varphi_{N,a};\boldsymbol{\theta}) \le \alpha.$$

The asymptotic critical value $\lambda_{\alpha,a}$ in (3.24) is independent of F_0 because $\boldsymbol{\mu}_a^*/A_a$ is independent of F_0. This means the asymptotic distribution freeness of $(\varphi_{N,a})$ with $\lambda_{\alpha,a}$ in (3.24). On the other hand, $\boldsymbol{\mu}_a^M(\boldsymbol{\theta}_0(l))/A_a$ depends on F_0 and hence we can determine $\lambda_{\alpha,a}$ in (3.25) only when F_0 is known.

4.2. Evaluating asymptotic minimum powers.

Let $\boldsymbol{\Theta}^s$ be the set of all $\boldsymbol{\theta}$ satisfying Condition I(s). Clearly, we have $\boldsymbol{\Theta}^1 \supset \boldsymbol{\Theta}^2 \cdots \supset \boldsymbol{\Theta}^k$. The following lemma is an immediate consequence of (ii) of Theorem 3.2, which gives a lower bound for the asymptotic minimum power of $(\varphi_{N,a})$ at every $\boldsymbol{\theta} \in \boldsymbol{\Theta}_1 \cap \boldsymbol{\Theta}^s$.

Lemma 3.4. *Let $(\varphi_{N,a})$ be as given in Theorem 3.3. Then, for every $\boldsymbol{\theta} \in \boldsymbol{\Theta}_1 \cap \boldsymbol{\Theta}^s$ it holds that*

$$\lim_{n\to\infty} \beta_n(\varphi_{N,a};\boldsymbol{\theta}) \ge 1 - \Pr\left\{\mathbf{Z} + \boldsymbol{\mu}_a^s(\boldsymbol{\theta})/A_a \in \mathbf{D}(\lambda_{\alpha,a})\right\}, \tag{3.26}$$

where $\boldsymbol{\mu}_a^s(\boldsymbol{\theta})$ is given by (3.18).

Now we consider the following condition

$$\int_0^1 \Lambda(F_0^{-1}(t))a(t)\,dt \ge \frac{2}{k}\eta(a(1) - a(0)); \qquad \eta = \frac{(\epsilon + 2\delta)}{\tau}. \tag{3.27}$$

For each s $(1 \le s \le k)$ let $\boldsymbol{\Theta}_{1s}^*$ be the set of all $\boldsymbol{\theta}$ satisfying

$$\begin{aligned} &\theta_{[i]} - \theta_{[i+1]} \ge \tau, \quad i = 1, \ldots, s-1 \\ &\theta_{[s]} - \theta_{[i+1]} \ge (i+1-s)\tau, \quad i = s, \ldots, k-1. \end{aligned} \tag{3.28}$$

We note that $\Theta_{11}^* \supset \Theta_{12}^* \supset \cdots \supset \Theta_{1k}^*$. Then the following lemma is obvious.

Lemma 3.5. *If the condition (3.27) is satisfied, then* $\Theta^s = \Theta_{1s}^*$ *holds for each* $s = 1, \ldots, k$.

In what follows, we treat asymptotic level α rank tests $(\varphi_{N,a})$ in case of $\Theta_0 = \Theta_0^*$ and $\Theta_1 = \{\boldsymbol{\theta} \mid \theta_{[1]} - \theta_{[k]} > \tau\}$ and investigate their asymptotic minimum powers on the subset Θ_{1s}^* of Θ_1. Let $\boldsymbol{\mu}_a^{s*} = (\mu_{a1}^{s*}, \ldots, \mu_{ak}^{s*})$ be defined by

$$\mu_{ai}^{s*} = \begin{cases} (k-2i+1)\Big\{\dfrac{\tau}{2}\displaystyle\int_0^1 \Lambda(F_0^{-1}(t))a(t)\,dt \\ \qquad\qquad -\dfrac{(\epsilon+2\delta)}{k}(a(1)-a(0))\Big\}, \quad i = 1, \ldots, s \\ -s\Big\{\dfrac{\tau}{2}\displaystyle\int_0^1 \Lambda(F_0^{-1}(t))a(t)\,dt \\ \qquad\qquad -\dfrac{(\epsilon+2\delta)}{k}(a(1)-a(0))\Big\}, \quad i = s+1, \ldots, k \end{cases} \tag{3.29}$$

Theorem 3.4. *Let* $(\varphi_{N,a})$ *be the sequences of (3.21) with the values* $\lambda_{N,a}$ *satisfying* $\lim_{n\to\infty} \lambda_{N,a} = \lambda_{\alpha,a}$ *and* $\lim_{n\to\infty} \mathbf{D}(\lambda_{N,a}) = \mathbf{D}(\lambda_{\alpha,a})$, *where* $\lambda_{\alpha,a}$ *is given by (3.24). If the condition (3.27) is satisfied, then it holds that*

$$\lim_{n\to\infty} \inf_{\boldsymbol{\theta}\in\Theta_{1s}^*} \beta_n(\varphi_{N,a};\boldsymbol{\theta}) \geq 1 - \Pr\big\{\mathbf{Z} + \boldsymbol{\mu}_a^{s*}/A_a \in \mathbf{D}(\lambda_{\alpha,a})\big\}. \tag{3.30}$$

Proof. Let $\boldsymbol{\theta}_1$ be a parameter such that $\theta_{[i]} - \theta_{[i+1]} = \tau$, $i = 1, \ldots, s-1$, and $\theta_{[s]} - \theta_{[i+1]} = (i+1-s)\tau$, $i = s, \ldots, k-1$. Then it is easy to see that $\boldsymbol{\mu}_a^{s*} = \boldsymbol{\mu}_a^s(\boldsymbol{\theta}_1) \prec \boldsymbol{\mu}_a^s(\boldsymbol{\theta})$ holds for every $\boldsymbol{\theta} \in \Theta_{1s}^*$ where $\boldsymbol{\mu}_a^s(\boldsymbol{\theta})$ is given by (3.18). Hence, the thorem follows from Theorem 1.2 in Section 1.2 and Lemma 3.4. □

Now let us consider the following stronger condition than (3.27).

$$\int_0^1 \Lambda(F_0^{-1}(t))a(t)\,dt \geq \frac{4}{k}\eta(a(1)-a(0)). \tag{3.31}$$

Then we can obtain the following theorem which states the asymptotic unbiasedness of $(\varphi_{N,a})$ for the problem (3.20) with $\Theta_0 = \Theta_0^*$ and $\Theta_1 = \Theta_{1k}^*$.

Theorem 3.5. *Let* $(\varphi_{N,a})$ *be as given in Theorem 3.4. If the condition (3.31) is satisfied, then it holds that*

$$\lim_{n\to\infty} \inf_{\boldsymbol{\theta}\in\Theta_{1k}^*} \beta_n(\varphi_{N,a};\boldsymbol{\theta}) \geq \alpha.$$

Proof. By (3.24) and Theorem 3.4 it is sufficient to show that $\boldsymbol{\mu}_a^{k*} \succ \boldsymbol{\mu}_a^*$ holds, where $\boldsymbol{\mu}_a^{k*}$ is given by (3.29) with $s = k$. It is easy to see that $\boldsymbol{\mu}_a^{k*} \succ \boldsymbol{\mu}_a^*$ is equivalent to the condition (3.31). □

3.4.3. Robustness of k-sample rank tests.

For the problem (3.20) with $\Theta_0 = \Theta_\bullet^*$ and $\Theta_1 = \Theta_{1k}^*$ we consider to find a scores generating function a which makes the righthand side of (3.30) with $s = k$ as large as possible.

Without loss of generality, we assume that $\int_0^1 a(t)\,dt = 0$. Let us define a truncated scores generating function a^* by

$$a^*(t) = d_0 \vee \Lambda(F_0^{-1}(t)) \wedge d_1, \tag{3.32}$$

where truncation points d_0 and d_1 are real numbers determined by

$$\int_0^1 \left(d_0 - \Lambda(F_0^{-1}(t))\right)^+ dt = \frac{4}{k}\eta = \int_0^1 \left(\Lambda(F_0^{-1}(t)) - d_1\right)^+ dt. \tag{3.33}$$

It follows from the disjointness condition (3.19) that d_0 and d_1 uniquely exist for $k \geq 4$. Clearly, d_0 and d_1 are decreasing and increasing in k, respectively. When $k = 3$, we may not have the solutions d_0 and d_1 of the equation (3.33) for given values $\eta = (\epsilon + 2\delta)/\tau$. In this case, to get d_0 and d_1 we have to take a larger value τ such that $0 < 4\eta/3 < \int \Lambda^+ \, dF_0$. Moreover, to guarantee that the conditions (i) and (ii) of scores generating function in Section 3.2 are satisfied by this a^*, we assume the conditions (5.12) and (5.13) in Rieder (1981).

Lemma 3.6. *For any scores generating function a it holds that*

$$\frac{\boldsymbol{\mu}_a^{k*} - \boldsymbol{\mu}_a^*}{A_a} \prec \frac{\boldsymbol{\mu}_{a^*}^{k*} - \boldsymbol{\mu}_{a^*}^*}{A_{a^*}}.$$

Proof. We note that the i-th component of the vector $(\boldsymbol{\mu}_a^{k*} - \boldsymbol{\mu}_a^*)/A_a$ is

$$(k - 2i + 1)\frac{\tau}{2}\Big\{ \int_0^1 \Lambda(F_0^{-1}(t))a(t)\,dt - \frac{4}{k}\eta\big(a(1) - a(0)\big)\Big\}/A_a.$$

It follows from (3.32) that

$$\begin{aligned}
&\int_0^1 \Lambda(F_0^{-1}(t))a(t)\,dt \\
&= \int_0^1 a(t)a^*(t)\,dt + \int_0^1 a(t)\big(\Lambda(F_0^{-1}(t)) - a^*(t)\big)\,dt \\
&\le A_a A_{a^*} + a(1)\int_0^1 \big(\Lambda(F_0^{-1}(t)) - d_1\big)^+\,dt - a(0)\int_0^1 \big(d_0 - \Lambda(F_0^{-1}(t))\big)^+\,dt.
\end{aligned}$$

Hence, from (3.33) we obtain for any scores generating function a

$$\begin{aligned}
&\Big\{\int_0^1 \Lambda(F_0^{-1}(t))a(t)\,dt - \frac{4}{k}\eta\big(a(1) - a(0)\big)\Big\}/A_a \\
&\le A_{a^*} + a(1)\Big\{\int_0^1 \big(\Lambda(F_0^{-1}(t)) - d_1\big)^+\,dt - \frac{4}{k}\eta\Big\}/A_a \\
&\qquad\qquad - a(0)\Big\{\int_0^1 \big(d_0 - \Lambda(F_0^{-1}(t))\big)^+\,dt - \frac{4}{k}\eta\Big\}/A_a \\
&= A_{a^*}
\end{aligned}$$

We can easily see that the i-th component of the vector $(\boldsymbol{\mu}_{a^*}^{k*} - \boldsymbol{\mu}_{a^*}^*)/A_{a^*}$ is $(k - 2i + 1)\frac{\tau}{2}A_{a^*}$, which completes the proof of the lemma. □

The righthand side of (3.30) with $s = k$ is written as

$$1 - \Pr\{\mathbf{Z} + \boldsymbol{\mu}_a^*/A_a + (\boldsymbol{\mu}_a^{k*} - \boldsymbol{\mu}_a^*)/A_a \in \mathbf{D}(\lambda_{\alpha,a})\}. \tag{3.34}$$

The directions of vectors $\boldsymbol{\mu}_a^*$ and $\boldsymbol{\mu}_a^{k*} - \boldsymbol{\mu}_a^*$ are the same if (3.31) is satisfied and the sums of their components are zero. As the size condition (3.24) holds for every scores generating function a, we can regard the magnitude $\|(\boldsymbol{\mu}_a^{k*} - \boldsymbol{\mu}_a^*)/A_a\|$ of the vector $(\boldsymbol{\mu}_a^{k*} - \boldsymbol{\mu}_a^*)/A_a$ as a measure of the magnitude of the lower bound (3.24), where $\|\mathbf{x}\|^2 = \sum_{i=1}^k x_i^2$. As $\|\mathbf{x}\|^2$ is a Schur-convex function of $\mathbf{x}$, Lemma 3.6 implies that a^* is recomendable. Therefore, under the condition (3.31), as a measure of asymptotic efficiency of $(\varphi_{N,a})$ relative to (φ_{N,a^*}), we propose the following quantity

$$\mathrm{ARE}^*\big((\varphi_{N,a}) : (\varphi_{N,a^*})\big) = \frac{\|(\boldsymbol{\mu}_a^{k*} - \boldsymbol{\mu}_a^*)/A_a\|^2}{\|(\boldsymbol{\mu}_{a^*}^{k*} - \boldsymbol{\mu}_{a^*}^*)/A_{a^*}\|^2}.$$

This definition of ARE* is based on the lower bounds of asymptotic minimum powers and it is different from that of the classical Pitman-Noether efficiency.

Note that

$$A_{a^*}^2 = \int_0^1 \Lambda(F_0^{-1}(t))a^*(t)\,dt - \frac{4}{k}\eta\big(a^*(1) - a^*(0)\big)$$

holds by (3.32) and (3.33). Then we have the following theorem.

Theorem 3.6. *If the condition (3.31) is satisfied, then*

$$ARE^*\big((\varphi_{N,a}):(\varphi_{N,a^*})\big) = \frac{\left\{\int_0^1 \Lambda(F_0^{-1}(t))a(t)\,dt - \frac{4}{k}\eta\big(a(1)-a(0)\big)\right\}^2}{A_a^2 A_{a^*}^2}.$$

Examples. Let us consider normal distributions with a location parameter θ such that $F_\theta(x) = \Phi(x-\theta)$ for θ, $x \in [-\infty, +\infty]$. It follows from (3.19) that $0 < \eta < 1/\sqrt{2\pi}$. The truncated normal scores rank tests (φ_{N,a^*}) are given by the scores generating function

$$a^*(t) = (-d_1) \vee \Phi^{-1}(t) \wedge d_1$$

where d_1 is uniquely determined by the equation

$$\varphi(d_1) - d_1\big(1-\Phi(d_1)\big) = \frac{4}{k}\eta, \qquad 0 < \frac{4}{k}\eta \le 1/\sqrt{2\pi}$$

because the righthand side of the above equation is decreasing in d_1. Then the condition (3.31) is written by $2\Phi(d_1) - 1 \ge \frac{8}{k}\eta d_1$, which is satisfied by all nonnegative d_1. We note that $A_{a^*}^2 = 2\Phi(d_1) - 1 - \frac{8}{k}\eta d_1$.

Wilcoxon type scores rank tests $(\varphi_{N,\hat{a}})$ are based on the scores generating function $\hat{a}(t) = 2t - 1$. In this case, it is easily seen that the condition (3.31) is $\frac{4}{k}\eta \le 1/2\sqrt{\pi}$.

From Theorem 3.6 it follows that

$$\mathrm{ARE}^*\big((\varphi_{N,\hat{a}}):(\varphi_{N,a^*})\big) = \frac{\big(\sqrt{3/\pi} - 8\sqrt{3}\eta/k\big)^2}{A_{a^*}^2}, \qquad 0 \le \frac{4}{k}\eta \le 1/2\sqrt{\pi} \tag{3.35}$$

Let $\frac{4}{k}\bar{\eta} = \varphi(\bar{d}) - \bar{d}(1-\Phi(\bar{d}))$, where $\bar{d}$ is the unique solution of $d = \sqrt{\pi}(2\Phi(d)-1)$. As $\frac{4}{k}\eta \to 1/2\sqrt{\pi}$, the ARE* first increases until $\eta = \bar{\eta}$, it attains a unique maximum of value $(3/\bar{d})(1/\sqrt{\pi} - 8\bar{\eta}/k)$ at $\eta = \bar{\eta}$ and then it tends decreasingly to 0. The ARE* in (3.35) corresponds to ARE$_{\mathrm{mx}}$ for one-sample in Rieder (1981, page 261), that is, the former is obtained by replacing η with $\frac{4}{k}\eta$ in the latter. This suggests that ARE* can be regarded as k-sample version of ARE$_{\mathrm{mx}}$. As a numerical example, we have $\bar{d} = 1.56336$, $\frac{4}{k}\bar{\eta} = 0.02533$ and then ARE$^* = 0.985450$.

We note that our ARE* is valid for any Schur-convex set $\mathbf{D}$. By changing Schur-convex sets $\mathbf{D}$ we can get various tests with the scores generating function a^*. The followings are important examples of $\mathbf{D}$ for the problem (3.20) with $\Theta_0 = \Theta_0^*$ and $\Theta_1 = \Theta_{1k}^*$.

$$\mathbf{D}_1(\lambda) = \{\mathbf{x} \mid \sum_{i=1}^{k} x_i^2 \leq \lambda\}, \qquad \mathbf{D}_2(\lambda) = \{\mathbf{x} \mid x_{[1]} - x_{[k]} \leq \lambda\}.$$

The tests (3.21) based on $\mathbf{D}_i(\lambda_{\alpha,a}), i = 1, 2$, are based on the statistics $\sum_{i=1}^{k}(T_{Ni} - \bar{a})^2$ and the range $T_{N[1]} - T_{N[k]}$, where $T_{N[i]}$ denotes the i-th largest statistic among $T_{N1}, \ldots, T_{Nk}$. A list of Schur-convex sets $\mathbf{D}$ which produce various tests for outliers is given in Section 1.3 of Chapter 1.

We have argued that a^* is recommendable in case of $\Theta_0 = \Theta_0^*$ and $\Theta_1 = \Theta_{1k}^*$. The scores generating function a^* seems to be also recommendable for general Θ_0 and Θ_1.

Appendix

1. Proof of Lemma 3.1.

Let

$$G_{n,A(m)}(x) = \frac{1}{m} \sum_{i \in A(m)} G_{ni}(x) \quad \text{and} \quad G_{n,J-A(m)}(x) = \frac{1}{k-m} \sum_{j \in J-A(m)} G_{nj}(x),$$

where $G_{ni}(x) = \frac{1}{n} \sum_{j=1}^{n} G_{ij}(x)$. Then we see that

$$\sum_{i \in \mathbf{A}(m)} \mu_{Ni} = m \int_{-\infty}^{+\infty} a\Big(\frac{m}{k} G_{n,A(m)}(x) + \frac{k-m}{k} G_{n,J-A(m)}(x)\Big)\, dG_{n,A(m)}(x)$$

$$= m \int_{0}^{1} a\Big(\frac{m}{k} t + \frac{k-m}{k} G_{n,J-A(m)}\big(G_{n,A(m)}^{-1}(t)\big)\Big)\, dt.$$

Also we see that

$$\begin{aligned} G_{n,A(m)}^{-1}(t) &\le \bar{G}_{n,A(m)}^{-1}(t) && \text{for all } t \in (0,1), \\ G_{n,J-A(m)}(x) &\le \underline{G}_{n,J-A(m)}(x) && \text{for all } x \in [-\infty, +\infty] \end{aligned}$$

where $\bar{G}_{n,A(m)}(x)$ and $\underline{G}_{n,J-A(m)}(x)$ are given by (3.11) and (3.12), respectively. Since a is increasing function, it follows that

$$\frac{1}{m} \sum_{i \in \mathbf{A}(m)} \mu_{Ni} \le \int_{0}^{1} a\Big(\frac{m}{k} t + \frac{k-m}{k} \underline{G}_{n,J-A(m)}\big(\bar{G}_{n,A(m)}^{-1}(t)\big)\Big)\, dt \tag{A.1}$$

$$= \int_{0}^{1} a\big(K_n(t)\big)\, dt.$$

Let

$$\bar{G}_{n,b,A(m)}(x) = \frac{1}{m} \sum_{i \in A(m)} \bar{G}_{ni,b_i}(x) \quad \text{and} \quad \underline{G}_{n,b,J-A(m)}(x) = \frac{1}{k-m} \sum_{j \in J-A(m)} \underline{G}_{nj,b_j}(x),$$

where

$$\bar{G}_{ni,b_i}(x) = \big((1-\epsilon_n)F_{\theta_{ni}}(x) - \delta_n\big)^+ + (\epsilon_n+\delta_n)\lambda_{(b_i,b_i+1)}(x),$$
$$\underline{G}_{nj,b_j}(x) = \big((1-\epsilon_n)F_{\theta_{nj}}(x) + (\epsilon_n+\delta_n)\lambda_{(-b_j-1,-b_j)}(x)\big) \wedge 1$$

and $\lambda_{(a,b)}$ denotes Lebesgue measure on (a,b). The $\bar{G}_{n,b,A(m)}(x)$ and $\underline{G}_{n,b,J-A(m)}(x)$ denote continuous distribution functions approximating $\bar{G}_{n,A(m)}(x)$ and $\underline{G}_{n,J-A(m)}(x)$. Then we shall show that the righthand side of (A.1) can be approximated arbitrarily close by

$$\int_0^1 a(\frac{m}{k}t + \frac{k-m}{k}\underline{G}_{n,b,J-A(m)}(\bar{G}^{-1}_{n,b,A(m)}(t)))\,dt.$$

Assume that $b_i > 0$ are so large and satisfy

$$(1-\epsilon_n)F_{\theta_{ni}}(-b_i) - \delta_n \le 0 \quad \text{and} \quad (1-\epsilon_n)F_{\theta_{ni}}(b_i) + \epsilon + \delta_n \ge 1.$$

Let

$$\hat{b} = \min_{1\le i\le k} b_i, \quad t_1 = \bar{G}_{n,b,A(m)}(-\hat{b}) \quad \text{and} \quad t_2 = \bar{G}_{n,b,A(m)}(\hat{b}).$$

Then $\hat{b} \to +\infty$ implies $t_1 \to 0$ and $t_2 \to 1-(\epsilon_n+\delta_n)$. We see that

$$\bar{G}_{n,A(m)}(x) = \bar{G}_{n,b,A(m)}(x) \qquad \text{for every } x \in (-\infty, \hat{b}],$$
$$\underline{G}_{n,J-A(m)}(x) = \underline{G}_{n,b,J-A(m)}(x) \qquad \text{for every } x \in [-\hat{b}, +\infty).$$

Hence,

$$\begin{aligned}
&\int_0^1 a(K_n(t))\,dt - \int_0^1 a(\frac{m}{k}t + \frac{k-m}{k}\underline{G}_{n,b,J-A(m)}(\bar{G}^{-1}_{n,b,A(m)}(t)))\,dt \\
&= \int_0^{t_1}\Big\{a\Big(\frac{m}{k}t + \frac{k-m}{k}\underline{G}_{n,J-A(m)}(\bar{G}^{-1}_{n,A(m)}(t))\Big) \\
&\qquad - a\Big(\frac{m}{k}t + \frac{k-m}{k}\underline{G}_{n,b,J-A(m)}(\bar{G}^{-1}_{n,A(m)}(t))\Big)\Big\}\,dt \\
&+ \int_{t_2}^1\Big\{a\Big(\frac{m}{k}t + \frac{k-m}{k}\underline{G}_{n,J-A(m)}(\bar{G}^{-1}_{n,A(m)}(t))\Big) \\
&\qquad - a\Big(\frac{m}{k}t + \frac{k-m}{k}\underline{G}_{n,J-A(m)}(\bar{G}^{-1}_{n,b,A(m)}(t))\Big)\Big\}\,dt.
\end{aligned} \tag{A.2}$$

Since $0 \le \underline{G}_{n,b,J-A(m)}(\bar{G}^{-1}_{n,A(m)}(t)) \le \underline{G}_{n,J-A(m)}(\bar{G}^{-1}_{n,A(m)}(t))$ for every $t \in (0,t_1)$, it follows that the first integral of (A.2) is nonnegative and it is bounded from above by

$$\begin{aligned}
&\int_0^{t_1}\{a(\frac{m}{k}t + \frac{k-m}{k}\underline{G}_{n,J-A(m)}(\bar{G}^{-1}_{n,A(m)}(t))) - a(\frac{m}{k}t)\}\,dt \\
&= \frac{k}{m}\int_{\frac{k-m}{k}\underline{G}_{n,J-A(m)}(-\hat{b})}^{\frac{m}{k}t_1 + \frac{k-m}{k}\underline{G}_{n,J-A(m)}(-\hat{b})} a(t)\,dt - \frac{k}{m}\int_0^{\frac{m}{k}t_1} a(t)\,dt.
\end{aligned} \tag{A.3}$$

Noting $a \in L_2(dt)$ and $t_1 \to 0$ as $\hat{b} \to +\infty$, we obtain from the dominated convergence theorem that the righthand side of (A.3) converges to 0 as $\hat{b} \to +\infty$.

Since $\bar{G}^{-1}_{n,b,A(m)}(t) \leq \bar{G}^{-1}_{n,A(m)}(t)$ for every $t \in (t_2, 1)$, the second integral of (A.2) is nonnegative and it is bounded from above by

$$\begin{aligned}
&\int_{t_2}^{1} \Big\{ a\Big(\frac{m}{k}t + \frac{k-m}{k} \underline{G}_{n,J-A(m)}(\bar{G}^{-1}_{n,A(m)}(t))\Big) - a\Big(\frac{m}{k}t + \frac{k-m}{k} \underline{G}_{n,J-A(m)}(\hat{b})\Big) \Big\}\, dt \\
&\leq \int_{t_2}^{1} \Big\{ a(\frac{m}{k}t + \frac{k-m}{k}) - a\Big(\frac{m}{k}t + \frac{k-m}{k} \underline{G}_{n,J-A(m)}(\hat{b})\Big) \Big\}\, dt \\
&= \frac{k}{m} \int_{\frac{m}{k}t_2 + \frac{k-m}{k}}^{1} a(t)\, dt - \frac{k}{m} \int_{\frac{m}{k}t_2 + \frac{k-m}{k} \underline{G}_{n,J-A(m)}(\hat{b})}^{\frac{m}{k} + \frac{k-m}{k} \underline{G}_{n,J-A(m)}(\hat{b})} a(t)\, dt.
\end{aligned}$$

This upper bound converges to 0 as $\hat{b} \to +\infty$ because $a \in L_2(dt)$, and $t_2 \to 1 - (\epsilon_n + \delta_n)$ and $\underline{G}_{n,J-A(m)}(\hat{b}) \to 1$ as $\hat{b} \to +\infty$.

Therefore, (A.2) converges to 0 as $\hat{b} \to +\infty$, which completes the proof of the lemma. □

2. Proof of Lemma 3.2.

In order to prove the assertion (i), from Lemma 3.1 it is sufficient to show that

$$\begin{aligned}
&\lim_{n \to \infty} n^{1/2} \int_0^1 \{a(K_n(t)) - a(t)\}\, dt \\
&= \frac{1}{m} \sum_{i \in \mathbf{A}(m)} (\theta_i - \bar{\theta}) \int_0^1 \Lambda(F_0^{-1}(t)) a(t)\, dt + \frac{k-m}{k}(\epsilon + 2\delta)(a(1) - a(0)). \qquad \text{(A.4)}
\end{aligned}$$

The assertion (ii) readily follows from the assertion (i) and th fact that

$$\begin{aligned}
&\inf \{ \sum_{i \in \mathbf{A}(m)} (\mu_{Ni} - \bar{a}) \mid W_N \in \mathcal{P}(\boldsymbol{\theta}_n; \epsilon_n, \delta_n) \} \\
&= -\sup \{ \sum_{j \in \mathbf{J} - \mathbf{A}(m)} (\mu_{Nj} - \bar{a}) \mid W_N \in \mathcal{P}(\boldsymbol{\theta}_n; \epsilon_n, \delta_n) \}.
\end{aligned}$$

In order to show (A.4) we verify the following (A.5) and (A.6):

$$\sup_{0<t<1} |K_n(t) - t| = O(n^{-1/2}) \qquad \text{(A.5)}$$

$$\sup_{a_n \leq t \leq 1 - a_n'} |K_n(t) - t - B_n(t)| = o(n^{-1/2}) \qquad \text{(A.6)}$$

where $K_n(t)$ is given by (3.10),

$$B_n(t) = -\frac{1}{m}\sum_{i\in\mathbf{A}(m)}(\theta_{ni}-\bar\theta_n)\int_0^t \Lambda(F_0^{-1}(s))\,ds + \frac{k-m}{k}(\epsilon_n+2\delta_n), \quad \bar\theta_n = n^{-1/2}\bar\theta,$$

$$a_n = \bar G_{n,A(m)}\Big(\max_{1\le i\le k} F_{\theta_{ni}}^{-1}(\frac{\delta_n}{1-\epsilon_n})\Big),$$

$$a_n' = 1-\bar G_{n,A(m)}\Big(\min_{1\le i\le k} F_{\theta_{ni}}^{-1}(\frac{1-\epsilon_n-\delta_n}{1-\epsilon_n})\Big).$$

Since $\bar G^{-1}_{n,A(m)}(t)$ is continuous for $0<t<1-(\epsilon_n+\delta_n)$, it follows that

$$\begin{aligned} K_n(t)-t &= \frac{k-m}{k}\{\underline{G}_{n,J-A(m)}(\bar G^{-1}_{n,A(m)}(t)) - \bar G_{n,A(m)}(\bar G^{-1}_{n,A(m)}(t))\}, \\ &= \frac{1}{km}\sum_{j\in\mathbf{J}-\mathbf{A}(m)}\sum_{i\in\mathbf{A}(m)}\{\underline{G}_{nj}(\bar G^{-1}_{n,A(m)}(t)) - \bar G_{ni}(\bar G^{-1}_{n,A(m)}(t))\}\end{aligned}$$

for $0<t<1-(\epsilon_n+\delta_n)$. Noting that

$$\underline{G}_{nj}(x)-\bar G_{ni}(x)$$
$$=\begin{cases} (1-\epsilon_n)F_{\theta_{nj}}(x)+\epsilon_n+\delta_n, & x<F^{-1}_{\theta_{ni}}(\frac{\delta_n}{1-\epsilon_n}) \\ (1-\epsilon_n)\big(F_{\theta_{nj}}(x)-F_{\theta_{ni}}(x)\big)+\epsilon_n+2\delta_n, & F^{-1}_{\theta_{ni}}(\frac{\delta_n}{1-\epsilon_n})\le x\le F^{-1}_{\theta_{nj}}(\frac{1-\epsilon_n-\delta_n}{1-\epsilon_n}) \\ 1-\big((1-\epsilon_n)F_{\theta_{ni}}(x)-\delta_n\big), & F^{-1}_{\theta_{nj}}(\frac{1-\epsilon_n-\delta_n}{1-\epsilon_n})<x<+\infty \\ 0, & x=+\infty \end{cases}$$

we see that

$$\sup_{0<t<a_n}|K_n(t)-t| \le \frac{1}{km}\sum_{j\in\mathbf{J}-\mathbf{A}(m)}\sum_{i\in\mathbf{A}(m)}\{(1-\epsilon_n)F_{\theta_{nj}}\big((\max_{1\le l\le k}F^{-1}_{\theta_{nl}}(\frac{\delta_n}{1-\epsilon_n})\big)+\epsilon_n+\delta_n\}.$$

We note that

$$F_j^{-1}(t-\parallel F_l-F_j\parallel)\le F_l^{-1}(t)\le F_j^{-1}(t+\parallel F_l-F_j\parallel), \tag{A.7}$$

where $\parallel F_l-F_j\parallel$ denotes supnorm between F_l and F_j, and $F^{-1}(s)=-\infty$ for $s\le 0$, $F^{-1}(s)=+\infty$ for $s\ge 1$. Since (3.3) entails

$$\parallel F_{\theta_{ni}}-F_0\parallel = O(n^{-1/2}) \quad \text{for } i=1,\dots,k,$$

it follows that

$$\begin{aligned}
&\sup_{0<t<a_n} |K_n(t) - t| \\
&\leq \frac{1}{km} \sum_{j\in \mathbf{J}-\mathbf{A}(m)} \sum_{i\in \mathbf{A}(m)} \{(1-\epsilon_n) \max_{1\leq l\leq k} (\frac{\delta_n}{1-\epsilon_n} + \| F_{\theta_{nl}} - F_{\theta_{nj}} \|) + \epsilon_n + \delta_n\} \\
&\leq \frac{k-m}{m} \{2(1-\epsilon_n) \max_{1\leq l\leq k} \| F_{\theta_{nl}} - F_0 \| + \epsilon_n + 2\delta_n\} \\
&= O(n^{-1/2}).
\end{aligned}$$

Similarly,

$$\begin{aligned}
&\sup_{1-a'_n<t<1-\epsilon_n-\delta_n} |K_n(t) - t| \\
&\leq \frac{1}{km} \sum_{j\in \mathbf{J}-\mathbf{A}(m)} \sum_{i\in \mathbf{A}(m)} \{1 - ((1-\epsilon_n) F_{\theta_{ni}} (\min_{1\leq l\leq k} F^{-1}_{\theta_{nl}} (\frac{1-\epsilon_n-\delta_n}{1-\epsilon_n})) - \delta_n)\} \\
&\leq \frac{k}{m} \sum_{j\in \mathbf{J}-\mathbf{A}(m)} \sum_{i\in \mathbf{A}(m)} \{(1-\epsilon_n) \min_{1\leq l\leq k} \| F_{\theta_{nl}} - F_{\theta_{ni}} \| + \epsilon_n + 2\delta_n\} \\
&\leq \frac{k-m}{k} \{2(1-\epsilon_n) \max_{1\leq l\leq k} \| F_{\theta_{nl}} - F_0 \| + \epsilon_n + 2\delta_n\} \\
&= O(n^{-1/2})
\end{aligned}$$

Next, for $a_n \leq t \leq 1 - a'_n$ it holds that

$$\begin{aligned}
&K_n(t) - t \\
&= \frac{1}{km} \sum_{j\in \mathbf{J}-\mathbf{A}(m)} \sum_{i\in \mathbf{A}(m)} \Big((1-\epsilon_n)\{F_{\theta_{nj}}(\bar{G}^{-1}_{n,A(m)}(t)) - F_{\theta_{ni}}(\bar{G}^{-1}_{n,A(m)}(t))\} + \epsilon_n + 2\delta_n\Big).
\end{aligned}$$

It follows from (3.3) that

$$\begin{aligned}
&\sup_{0<t<1} |F_{\theta_{nj}}(\bar{G}^{-1}_{n,A(m)}(t)) - F_{\theta_{ni}}(\bar{G}^{-1}_{n,A(m)}(t)) \\
&\qquad - (\theta_{nj} - \theta_{ni}) \int_{-\infty}^{\bar{G}^{-1}_{n,A(m)}(t)} \Lambda(x)\, dF_0(x)| = o(n^{-1/2})
\end{aligned}$$

and since $\Lambda^2 \in L^2(dF_0)$, it follows that

$$\begin{aligned}
&|\int_{-\infty}^{\bar{G}^{-1}_{n,A(m)}(t)} \Lambda(x)\, dF_0(x) - \int_{-\infty}^{F_0^{-1}(t)} \Lambda(x)\, dF_0(x)|^2 \\
&\qquad \leq \| F_0 - \bar{G}_{n,A(m)} \| \int_{-\infty}^{+\infty} \Lambda^2(x)\, dF_0(x) = O(n^{-1/2}).
\end{aligned}$$

Hence, for $a_n \le t \le 1 - a'_n$

$$K_n(t) - t = \frac{1}{km} \sum_{j \in \mathbf{J}-\mathbf{A}(m)} \sum_{i \in \mathbf{A}(m)} (1-\epsilon_n)(\theta_{nj} - \theta_{ni}) \int_0^t \Lambda\big(F_0^{-1}(s)\big)\, ds$$

$$+ \frac{k-m}{k}(\epsilon_n + 2\delta_n) + o(n^{-1/2})$$

$$= -\frac{1}{m} \sum_{i \in \mathbf{A}(m)} (\theta_{ni} - \bar{\theta}_n) \int_0^t \Lambda\big(F_0^{-1}(s)\big)\, ds + \frac{k-m}{k}(\epsilon_n + 2\delta_n) + o(n^{-1/2}),$$

$$= B_n(t) + o(n^{-1/2}).$$

Finally we consider the case of $1 - (\epsilon_n + \delta_n) \le t < 1$. Since $K_n(t) = \frac{m}{k}t + \frac{k-m}{k}$, it holds that

$$|K_n(t) - t| = \frac{k-m}{k}|1-t| \le \frac{k-m}{k}(\epsilon_n + \delta_n) = O(n^{-1/2}).$$

Therefore we obtain (A.5) and (A.6).

Now we shall show the formula (A.4). We consider the following division of the integral

$$n^{1/2} \int_0^1 \{a\big(K_n(t)\big) - a(t)\}\, dt$$

$$= n^{1/2}\{\int_0^{c_n} + \int_{c_n}^{1-c_n} + \int_{1-c_n}^1\}\{a\big(K_n(t)\big) - a(t)\}\, dt \tag{A.8}$$

and furthermore we split the second integral of (A.8), in view of the absolute continuity of a,

$$n^{1/2} \int_{c_n}^{1-c_n} \{a\big(K_n(t)\big) - a(t)\}\, dt$$

$$= n^{1/2} \int_{c_n}^{1-c_n} (K_n(t) - t)\{\frac{a\big(K_n(t)\big) - a(t)}{K_n(t) - t} - a'(t)\}\, dt \tag{A.9}$$

$$+ n^{1/2} \int_{c_n}^{1-c_n} (K_n(t) - t - B_n(t))a'(t)\, dt$$

$$+ n^{1/2} \int_{c_n}^{1-c_n} B_n(t)a'(t)\, dt$$

where

$$c_n = \sup_{0<t<1} |K_n(t) - t|$$

and we note $c_n = O(n^{-1/2})$ by (A.5). The first integral of (A.8) is bounded from above by

$$n^{1/2} \int_0^{c_n} \{a(t + c_n) - a(t)\}\, dt \le n^{1/2} c_n \{a(2c_n) - a(0)\} \to 0$$

and it is bounded from below by $n^{1/2}c_n\{a(0)-a(c_n)\}\to 0$. Hence, it holds

$$n^{1/2}\int_0^{c_n}\{a(K_n(t))-a(t)\}\,dt\to 0 \quad \text{as} \quad n\to\infty.$$

For the third integral of (A.8) it follows from the same argument that

$$n^{1/2}\int_{1-c_n}^{1}\{a(K_n(t))-a(t)\}\,dt\to 0 \quad \text{as} \quad n\to\infty.$$

As for the second integral of (A.8), the formula (A.9) are employed.

At first, in order to show that the first integral of (A.9) converges to zero, we divide the interval $[\,c_n, 1-c_n\,]$ of this integral into the following five parts according to the assumption (ii) of a,

$$n^{1/2}\int_{c_n}^{1-c_n} = n^{1/2}\Big(\int_{c_n}^{t_0-c_n}+\int_{t_0-c_n}^{t_0+c_n}+\int_{t_0+c_n}^{1-t_0-c_n}+\int_{1-t_0-c_n}^{1-t_0+c_n}+\int_{1-t_0+c_n}^{1-c_n}\Big). \tag{A.10}$$

The absolute value of integral over (c_n, t_0-c_n) is bounded from above by

$$n^{1/2}\sup_{0<t<1}|K_n(t)-t|\cdot\Big|\int_{c_n}^{t_0-c_n}\{a'(t+\xi(K_n(t)-t))-a'(t)\}\,dt\Big|,\quad 0<\xi<1.$$

Because $a'(t)$ is nonincreasing on $(0,t_0)$,

$$\int_{c_n}^{t_0-c_n}\{a'(t+\xi(K_n(t)-t))-a'(t)\}\,dt$$

is bounded from above by

$$a(t_0-2c_n)-a(0)-\{a(t_0-c_n)-a(c_n)\}\to 0$$

and bounded from below by

$$a(t_0)-a(2c_n)-\{a(t_0-c_n)-a(c_n)\}\to 0.$$

Hence, the first integral of (A.10) over (c_n,t_0-c_n) converges to zero. Applying the same argument in view of the nondecreasingness of a on $(1-t_0,1)$, the fifth integral of (A.10) over $(1-t_0+c_n, 1-c_n)$ also converges to zero. The second and fouth integrals of (A.10) over (t_0-c_n, t_0+c_n) and $(1-t_0-c_n, 1-t_0+c_n)$ obviously converge to zero. By virture of Lipschitz boundedness of a on $[\,t_0, 1-t_0\,]$ and the dominated convergence theorem,

$$\int_{t_0+c_n}^{1-t_0-c_n}\Big\{\frac{a(K_n(t))-a(t)}{K_n(t)-t}-a'(t)\Big\}\,dt\to 0$$

and consequently, the third integral of (A.10) over $(t_0+c_n, 1-t_0-c_n)$ converges to zero.

Next, as for the second integral of (A.9), we also divide the interval $(c_n, 1-c_n)$ of the integral into $(c_n, a_n), (a_n, 1-a_n')$ and $(1-a_n', 1-c_n)$. By (A.5) and (A.6), it is bounded from above by

$$O(1)\{a(a_n) - a(c_n) + a(1-c_n) - a(1-a_n')\} + o(1)\{a(1-a_n') - a(a_n)\}.$$

As $a_n = O(n^{-1/2})$ and $a_n' = O(n^{-1/2})$ by a monotony argument using (A.7), this bound tends to zero.

Finally, let us consider the third integral of (A.9). It follows from $\Lambda^2 \in L^2(dF_0)$ and the integration by parts that

$$\begin{aligned}
& n^{1/2} \int_{c_n}^{1-c_n} B_n(t) a'(t)\, dt \\
&= n^{1/2}\{B_n(1-c_n)a(1-c_n) - B_n(c_n)a(c_n)\} \\
&\qquad - n^{1/2}\{-\frac{1}{m}\sum_{i\in\mathbf{A}(m)}(\theta_{ni} - \bar{\theta}_n)\int_{c_n}^{1-c_n}\Lambda(F_0^{-1}(t))a(t)\,dt\} \\
&= \frac{1}{m}\sum_{i\in\mathbf{A}(m)}(\theta_i - \bar{\theta})\int_0^1 \Lambda(F_0^{-1}(t))a(t)\,dt + \frac{k-m}{k}(\epsilon + 2\delta)\big(a(1) - a(0)\big) + o(1).
\end{aligned}$$

Therefore we obtain (A.4), which completes the proof of this lemma. □

References

Anderson, T. W. (1955). The integral of a symmetric unimodal function over a symmetric convex set and some probability inequalities. *Proc. Amer. Math. Soc.* **6**, 170-176.

Barnett, V. and T. Lewis (1984). *Outliers in Statistical Data* (Second Edition). Wily, New York.

Chen, H.J. and M. Xiong (1993). Range tests for the dispersion of sevral location parameters. *J. Statist. Plann. Inference* **36**, 15-25.

Dean, A. M. and J.S. Verducci (1990). Linear transformations that preserve majorization, Schur concavity and exchangeablility. *Linear Algebra and Its Appli.* **127**, 121-138.

Dharmadhikari, S. and K. Joag-Dev (1988). *Unimodality, convexity, and applications.* Academic Press, New York.

Finner, H. and M. Roters (1993). Power considerations and LFC-results with respect to Ψ_t-restriction in k-variate location models. *Statistics & Decisions* **11**, 151-166.

Giani, G. and H. Finner (1991). Some general results on least favorable parameter configurations with special reference to equivalence testing and the range statiatics. *J. Statist. Plann. Inference* **28**, 33-47.

Hajék, J. (1968). Asymptotic normality of simple linear rank statistics under alternatives. *Ann. Math. Statist.* **39**, 325-346.

Hoeffding, W. (1973). On the centering of a simple linear rank statistics. *Ann. Statist.* **1**, 54-66.

Kakiuchi, I. and Kimura, M. (1994). Majorization methods on hyperplanes and their applications. To appear in *J. Statist. Plann. Inference*

Kakiuchi, I and M, Kimura (1995). Robustness of rank tests for k-sample approximate equality in the presence of gross errors. Working Paper No.9413, Center for Management Studies, Nanzan Univ.

Kimura, M. and Kakiuchi, I. (1989). A majorization inequality for distributions on hyperplanes and its applications to test for outliers. *J. Statist. Plann. Inference* **21**, 19-26.

Marshall, A. W. and I. Olkin (1974). Majorization in multivariate distribution. *Ann. Statist.* **2**, 1189-1200.

Marshall, A. W. and I. Olkin (1979). *Inequalities: Theory of Majorization and its Applications.* Academic Press, New York.

Mudholkar, G. S. (1966). The integral of an invariant unimodal function over an invariant convex set—an inequality and applications. *Proc. Amer. Math. Soc.* **17**, 1327-1333.

Prékopa, A. (1973). On logarithmic concave measures and functions. *Acta. Sci. Math.* **34**, 335-343.

Puri, M. L. and Sen, P. K. (1985). *Nonparametric Methods in General Linear Models.* Wiley, New York.

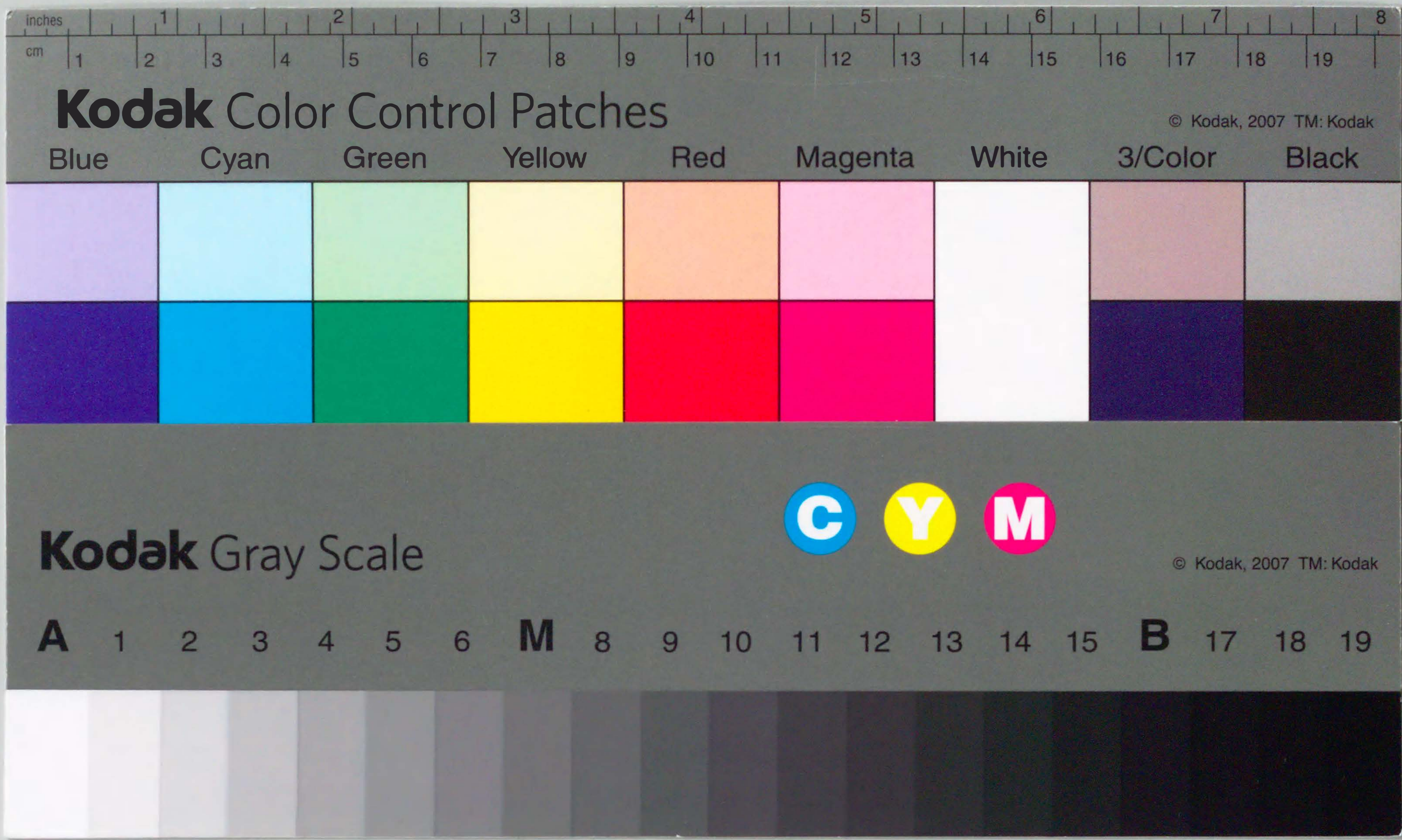
Rieder, H. (1977). Least favorable pairs for special capacities. *Ann. Statist.* **5**, 909-921.

Rieder, H. (1978). A robust asymptotic testing model. *Ann. Statist.* **6**, 1080-1094.

Rieder, H. (1981). Robustness of one- and two-sample rank tests against gross errors. *Ann. Statist.* **9**, 245-265.

Tong, Y. L. (1980). *Probability Inequality in Multivariate Distribution.* Academic Press, New York.

Itsuro KAKIUCHI
Department of Computer and Systems Engineering
Kobe University
Rokkodai 1-1, Nada, Kobe 657, Japan


inches 1 2 3 4 5 6 7 8
cm 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19
Kodak Color Control Patches
© Kodak, 2007 TM: Kodak
Blue Cyan Green Yellow Red Magenta White 3/Color Black
Kodak Gray Scale
C Y M
© Kodak, 2007 TM: Kodak
A 1 2 3 4 5 6 M 8 9 10 11 12 13 14 15 B 17 18 19