Cambridge University Press
978-0-521-09305-7 - Nonparametric Techniques in Statistical Inference
Edited by Madan Lal Puri
Excerpt
More information

PART 1

TESTING AND ESTIMATION 1

MISCELLANEOUS PROBLEMS OF RANK TEST THEORY

JAROSLAV HÁJEK

1 INTRODUCTION AND SUMMARY

The paper deals with diverse problems, partly unpublished, partly published without proofs in Hájek (1969). Section 3 proposes three methods of density estimation which may be used in choosing proper scores for rank tests. The second of these methods is based on a theorem by Jurečková (1969) on asymptotic linearity of simple linear rank statistics, which is presented in a modified form in § 2. The same result by Jurečková is once again used in § 4 to prove that some statistics for testing scale (the symmetric case) may be adapted for the case when both medians are unknown without inflicting their asymptotic distribution. For a less general result of this kind, derived by a different method, *see* Raghavachari (1965). Section 5 provides a theorem to the effect that one partial ordering of permutations is finer than another. This theorem may be utilized in proving Jurečková's theorem, and also in proving unbiasedness of some nonparametric tests (see Lehmann (1966), and Hájek (1969)). The last section provides an explicit form for asymptotic normality of the conditional distribution of averaged rank statistics, when ties are present.

2 THE BEHAVIOUR OF LINEAR RANK STATISTICS UNDER A SHIFT

For each $N \geqslant 1$ let us consider a simple linear rank statistics.

$$S_N = \sum_{i=1}^{N} c_{Ni} a_N(R_{Ni}), \tag{2.1}$$

where the ranks $R_{N1}, \ldots, R_{NN}$ are derived from a sequence of observations $Y_{N1}, \ldots, Y_{NN}$. Introducing

$$u(y) = 1, \quad x \geqslant 0, \tag{2.2}$$
$$= 0, \quad x < 0,$$

and putting $$s_N(y_1, \ldots, y_N) = \sum_{i=1}^{N} c_{Ni} a_N \left(\sum_{j=1}^{N} u(y_i - y_j) \right) \tag{2.3}$$

we may also write $$S_N = s_N(Y_{N1}, \ldots, Y_{NN}). \tag{2.4}$$

We shall study the increment of S_N, if the Y_{Ni}'s are shifted by a random multiple of some numbers d_{Ni}. That is, considering simultaneously the statistic

$$S_N^* = s_N(Y_{N1} + \Delta_N d_{N1}, \ldots, Y_{NN} + \Delta_N d_{NN}) \tag{2.5}$$

we shall try to find an asymptotic expression for the difference $S_N^* - S_N$.

Assumptions A:

(A 1) The regression constants satisfy the Noether condition:

$$\lim_{N\to\infty} \frac{\max\limits_{1\leqslant i\leqslant N} (c_{Ni} - \bar{c}_N)^2}{\sum\limits_{i=1}^{N} (c_{Ni} - \bar{c}_N)^2} = 0 \quad \left(\bar{c}_N = \frac{1}{N}\sum_{i=1}^{N} c_{Ni}\right). \tag{2.6}$$

(A 2) The scores $a_N(i)$ are generated by a function $\phi(t)$, $0 < t < 1$, by either of the following two ways:

$$a_N(i) = \phi\left(\frac{i}{N+1}\right) \quad (1 \leqslant i \leqslant N), \tag{2.7}$$

$$a_N(i) = E\phi(U_N^{(i)}) \quad (1 \leqslant i \leqslant N), \tag{2.8}$$

where $U_N^{(i)}$ denotes the ith order statistic in a sample of size N from the uniform distribution on $(0, 1)$.

(A 3) The score-generating function $\phi(t)$ is nonconstant and expressible as a difference of two *nondecreasing* and *square integrable* functions.

(A 4) The constants d_{Ni} are such that

$$\sup_N \sum_{i=1}^{N} (d_{Ni} - \bar{d}_N)^2 < \infty, \tag{2.9}$$

and

$$\lim_{N\to\infty} \max_{1\leqslant 1\leqslant N} (d_{Ni} - \bar{d}_N)^2 = 0 \quad \left(\bar{d}_N = \frac{1}{N}\sum_{i=1}^{N} d_{Ni}\right). \tag{2.10}$$

(A 5) The constants d_{Ni} and c_{Ni} are concordant in the following sense:

$$(c_{Ni} - c_{Nj})(d_{Ni} - d_{Nj}) \geqslant 0 \quad (1 \leqslant i, j \leqslant N). \tag{2.11}$$

(A 6) The random variables $\Delta_N = \Delta_N(Y_{N1}, \ldots, Y_{NN})$ are *bounded in probability*.

(A 7) $Y_{N1}, \ldots, Y_{NN}$ is a random sample from a distribution with a density f possessing finite Fisher's information.

Put

$$\phi(t, f) = -\frac{f'(F^{-1}(t))}{f(F^{-1}(t))} \quad (0 < t < 1). \tag{2.12}$$

and

$$b_N = \sum_{i=1}^{N} d_{Ni}(c_{Ni} - \bar{c}_N) \int_0^1 \phi(t)\, \phi(t, f)\, dt. \tag{2.13}$$

Theorem 2.1. *Under assumptions* (A 1) *through* (A 7),

$$\frac{S_N^* - S_N - \Delta_N b_N}{\sqrt{(\operatorname{var} S_N)}} \to 0 \tag{2.14}$$

holds in probability.

Proof. The theorem is an easy corollary of a result by J. Jurečková (1969). She showed that for nonrandom Δ_N (2.14) holds uniformly for $|\Delta_N| \leqslant C$. Consequently (2.14) must also hold for Δ_N random and bounded in probability.

Remark 2.1. Instead of (A 4) and (A 5) we could also assume that $d_{Ni} = d'_{Ni} - d''_{Ni}$, where both the constants d'_{Ni} and d''_{Ni} satisfy (A 4) and (A 5).

Remark 2.2. Assumption (A 2) may be replaced by (A 2*): The scores $a_N(i)$ satisfy

$$\frac{1}{N} \sum_{i=1}^{N} [a_N(i) - E\phi(U_N^{(i)})]^2 \to 0. \tag{2.15}$$

3 THREE METHODS OF DENSITY TYPE ESTIMATION

Let $\mathscr{F}_1, \ldots, \mathscr{F}_k$ be k distinct density types generated by some one-dimensional densities $f_1, \ldots, f_k$:

$$\mathscr{F}_j = \{f: f(x) = \lambda f_j(\lambda x - u),\ -\infty < u < \infty,\ \lambda > 0\} \quad (1 \leqslant j \leqslant k). \tag{3.1}$$

Given a sample $(X_1, \ldots, X_N)$ governed by a density of the form

$$p(x_1, \ldots, x_N) = \prod_{i=1}^{N} f(x_i), \tag{3.2}$$

where f belongs to $\bigcup_{j=1}^{k} \mathscr{F}_j$ but otherwise is unknown, let us try to locate the type of f, i.e. find j such that $f \in \mathscr{F}_j$. Thus, a decision procedure for this problem will be a function $\delta(x_1, \ldots, x_N)$ taking its values in the set $\{1, 2, \ldots, k\}$. Assume that the loss incurred by $\delta(x) = d$ under $f \in \mathscr{F}_j$ is given by

$$\begin{aligned} L(j,d) &= 0 \quad \text{if} \quad j = d, \\ &= 1 \quad \text{if} \quad j = d,\ 1 \leqslant j, d \leqslant k. \end{aligned} \tag{3.3}$$

Restricting ourselves to procedures that are invariant relative to the group of positive linear transforms

$$((x_1, \ldots, x_N) \to (\lambda x_1 + u, \ldots, \lambda x_N + u), \lambda > 0)$$

the risk of a procedure given $f \in \mathscr{F}_j$ will equal

$$R(j, \delta) = 1 - P(\delta(X_1, \ldots, X_N) = j \,|\, f \in \mathscr{F}_j). \tag{3.4}$$

Invariant Bayes solutions may be obtained from 'marginal' densities relative to the sub σ-field consisting of events that are invariant with respect to the group of positive linear transforms. These densities can be established for our types as follows:

$$\bar{p}_j(x_1, \ldots, x_N) = \int_0^\infty \int_{-\infty}^\infty \left[\prod_{i=1}^N f_j(\lambda x_i - u)\right] \lambda^{N-2} du\, d\lambda \quad (1 \leqslant j \leqslant k) \quad (3.5)$$

(*see* Hájek and Šidák (1967), § II.2.2). If all types are considered *a priori* equiprobable, then the expected risk $(1/k) \sum_{j=1}^k R(j, \delta)$ is minimized by the procedure δ_0 such that

$$[\delta_0(x_1, \ldots, x_N) = j] \Rightarrow [\bar{p}_j(x_1, \ldots, x_N) = \max_{1 \leqslant h \leqslant k} \bar{p}_h(x_1, \ldots, x_N)]. \quad (3.6)$$

Any such procedure satisfies

$$\sum_{j=1}^k R(j, \delta_0) \leqslant \sum_{j=1}^k R(j, \delta) \quad (\delta \text{ invariant}). \quad (3.7)$$

One should expect that

$$\sum_{j=1}^k R(j, \delta_0) \to 0 \quad \text{as} \quad N \to \infty$$

under fairly general conditions concerning the types $\mathscr{F}_1, \ldots, \mathscr{F}_k$. If the ϕ-functions $\phi(t, f_j)$ given by (2.12) satisfy (A 3), we shall be able to infer

$$\sum_{j=1}^N R(j, \delta_0) \to 0$$

from (3.7) and from the fact that

$$\sum_{j=1}^N R(j, \delta_1) \to 0$$

for some other invariant decision method δ_1 to be described next. We conclude the discussion of the procedure (3.6) by observing that the complexity of the formula for $\bar{p}_j$ makes its applicability doubtful.

Another selection procedure is yielded by Theorem 2.1. First introduce the functions

$$s_{Nj}(y_1, \ldots, y_N) = \sum_{i=1}^{[\frac{1}{2}N]} a_{Nj} \left(\sum_{h=1}^N u(y_i - y_h)\right) \quad (1 \leqslant j \leqslant k), \quad (3.8)$$

where the scores $a_{Nj}(i)$ are generated either by (2.7) or by (2.8) with $\phi(t)$ replaced by $\phi(t, f_j) = -f_j'(F_j^{-1}(t))/f_j(F_j^{-1}(t))$. Further put

$$d_{Ni} = 1/\sqrt{N} \quad (1 \leqslant i \leqslant [\tfrac{1}{2}N]), \quad (3.9)$$
$$= 0 \quad (\tfrac{1}{2}N \leqslant i < N),$$

and denote by $\Delta_N = \Delta_N(y_1, \ldots, y_N)$ any statistic such that

$$\Delta_N(\lambda y_1 + u, \ldots, \lambda y_N + u) = \lambda\Delta_N(y_1, \ldots, y_N) \quad (-\infty < u < \infty,\ \lambda > 0) \tag{3.10}$$

and that $$0 < p \lim_{f=f_j} \Delta_N(X_1, \ldots, X_N) = b_j < \infty \quad (1 \leqslant j \leqslant k). \tag{3.11}$$

For example, if $X_N^{(i)}$ denotes the ith order statistic we may put

$$\Delta_N = M[X_N^{(N+1-j_N)} - X_N^{(j_N)}] \quad \left(\frac{j_N}{N} \to \alpha,\ 0 < \alpha \leqslant \tfrac{1}{2}\right). \tag{3.12}$$

Then obviously,

$$b_j = M[F_j^{-1}(1-\alpha) - F_j^{-1}(\alpha)] \quad \left(F_j(\alpha) = \int_0^\alpha f_j(x)\,dx\right). \tag{3.13}$$

Finally put $$S_{Nj} = s_{Nj}(X_1, \ldots, X_N) \quad (1 \leqslant j \leqslant k) \tag{3.14}$$

and $$S_{Nj}^* = s_{Nj}(X_1 + \Delta_N d_{N1}, \ldots, X_N + \Delta_N d_{NN}) \tag{3.15}$$

and compute the ratios $$l_{Nj} = \frac{S_{Nj}^* - S_{Nj}}{\sqrt{(\operatorname{var} S_{Nj})}}. \tag{3.16}$$

The invariance of l_{Nj} under positive linear transforms and a direct application of Theorem 2.1 provide

$$p \lim_{f \in \mathscr{F}_h} l_{Nj} = p \lim_{f=f_h} l_{Nj} = p \lim_{f=f_h} \frac{\Delta_N \frac{1}{4}\sqrt{N}\int_0^1 \phi(t, f_h)\,\phi(t, f_j)\,dt}{\left[\frac{1}{4}N\int_0^1 \phi^2(t, f_j)\,dt\right]^{\frac{1}{2}}}$$
$$= \tfrac{1}{2} b_h \rho_{jh} \sqrt{I_h}, \tag{3.17}$$

where $$I_h = \int_0^1 \phi^2(t, f_h)\,dt$$

and $$\rho_{jh} = [I_j I_h]^{-\frac{1}{2}} \int_0^1 \phi(t, f_h)\,\phi(t, f_j)\,dt. \tag{3.18}$$

If the types $\mathscr{F}_1, \ldots, \mathscr{F}_k$ are distinct, we have

$$\rho_{jh} < 1 \quad (1 \leqslant j \neq h \leqslant k), \tag{3.19}$$

while $\rho_{jj} = 1$, $1 \leqslant j \leqslant k$. Consequently the decision procedure δ_1 such that

$$[\delta_1(x_1, \ldots, x_N) = j] \Rightarrow [l_{Nj}(x_1, \ldots, x_N) = \max_{1 \leqslant h \leqslant k} l_{Nh}(x_1, \ldots, x_N)] \tag{3.20}$$

is consistent in the sense that $R(j, \delta_1) \to 0$ if $N \to \infty$, $1 \leqslant j \leqslant k$. This entails

$$\sum_{j=1}^k R(j, \delta_1) \to 0,$$

and, in view of (3.7), also $$\sum_{j=1}^k R(j, \delta_0) \to 0.$$

In order to apply the selection procedure δ_1 to the choice of proper scores in a rank statistic of form (2.1), we need δ_1 to be a function of the order statistic $X^{(\cdot)} = (X^{(1)}, \ldots, X^{(N)})$ only.

This goal may be easily achieved by introducing random variables

$$Y_i = X^{(Q_i)} \quad (1 \leqslant i \leqslant N), \tag{3.21}$$

where $Q = (Q_1, \ldots, Q_N)$ is independent of $(X_1, \ldots, X_N)$ and assumes each of $N!$ permutations of $(1, 2, \ldots, N)$ with probability $1/N!$. Then $(Y_1, \ldots, Y_N)$ has the same distribution as $(X_1, \ldots, X_N)$ under all densities of the form (3.2) and the selection procedure may equally well be applied to $Y_1, \ldots, Y_N$ instead of to $X_1, \ldots, X_N$. Thus we have

Theorem 3.1. *Under the above notations, if the functions $\phi(t, f_j)$ are expressible as differences of two nondecreasing and square integrable functions, $1 \leqslant j \leqslant k$, then*

$$\lim_{N\to\infty} P(\delta_1(Y_1, \ldots, Y_N) = j \mid q_{Nj}) = 1 \tag{3.22}$$

holds for any sequence of densities $q_{Nj}(x_1, \ldots, x_N)$ that are contiguous to densities of the form

$$p_{Nj}(x_1, \ldots, x_N) = \lambda_N^N \prod_{i=1}^N f_j(\lambda_N x_i + u_N) \quad (-\infty < u_N < \infty,\ \lambda_N > 0). \tag{3.23}$$

Proof. The above considerations yielded the proof for $q_{Nj} = p_{Nj}$, i.e. if f in (3.2) satisfied $f \in \mathscr{F}_j$ for every $N \geqslant 1$. However, if q_{Nj} are contiguous relative to p_{Nj} in the X-spaces, the induced distributions of $(Y_1, \ldots, Y_N)$ are also contiguous to p_{Nj} in the Y-spaces, as may be easily seen. Then it suffices to note that convergence to 1 under p_{Nj} entails the same under any contiguous alternative. Q.E.D.

Let us resume the discussion of how δ_1 may be applied in rank testing: Let us test the hypothesis of randomness against the alternative of two samples differing in location or against a general location shift alternative. Then the pertinent statistic is of form (2.1), where the scores $a_N(i)$ should correspond to the underlying density f in the well-known way. If the type of f is not known, we put forward a certain number of density types and compute for them the quantities l_{Nj} applied to $Y_1, \ldots, Y_N$ generated as a random permutation of order statistics $X^{(1)}, \ldots, X^{(N)}$. The density type providing the largest l_{Nj} is then chosen to generate the scores and the test is performed as if these scores were decided upon before knowing $X_1, \ldots, X_N$. Since under the hypothesis of randomness the vector of ranks $(R_1, \ldots, R_N)$ is independent of the vector of order statistics, the above selection of scores does not invalidate the significance level. For moderate

N we can be deciding between three types corresponding to the following ϕ-functions, for example:

$$\begin{aligned}
\phi_1(t) &= \Phi^{-1}(t) && \text{(normal type)},\\
\phi_2(t) &= 2t-1 && \text{(logistic type)},\\
\phi_3(t) &= -1 && (0 \leqslant t \leqslant \tfrac{1}{4}),\\
&= 4t-2 && (\tfrac{1}{4} \leqslant t \leqslant \tfrac{3}{4}),\\
&= 1 && (\tfrac{3}{4} \leqslant t \leqslant 1).
\end{aligned} \tag{3.24}$$

(Φ^{-1} denotes the inverse normal distribution function.) It is advisable to use $\phi_3(t)$ instead of $\phi_4(t) = \operatorname{sign}(2t-1)$, which corresponds to the double-exponential type, since the discontinuity of ϕ_4 at $t = \frac{1}{2}$ increases unduly the variability of the corresponding ratio l_{Nj}. Thus the method would consist in testing sensitiveness of the van der Waerden test, the Wilcoxon test, and a test lying 'between' the Wilcoxon test and the median test, to the shift of location of the sample $Y_1, \ldots, Y_{[\frac{1}{2}N]}$. The test which is most sensitive is then applied to the original problem, as far as the scores are concerned. If ϕ_3 is selected, we can apply the median test as well.

The third selection procedure is based on a partial ordering of ϕ-function and of corresponding types. Within the family of skew symmetric ϕ-functions, we shall say, that ϕ increases more rapidly than ψ, if

$$\phi(t) = b(t)\,\psi(t) \quad (\tfrac{1}{2} < t < 1), \tag{3.25}$$

where $b(t)$ is nondecreasing. Inspecting (3.24), it is easy to see that ϕ_1 increases more rapidly than ϕ_2, and ϕ_2 increases more rapidly than ϕ_3, and ϕ_3 increases more rapidly than $\phi_5(t) = \sqrt{2}\sin[\pi(2t-1)]$, which corresponds to the Cauchy distribution. Further, given two types $\mathscr{F}$ and $\mathscr{G}$ of symmetric densities, we shall say that the densities from $\mathscr{F}$ have shorter (lighter) tails than the densities from $\mathscr{G}$ if for any $f \in \mathscr{F}$ and $g \in \mathscr{G}$, such that their medians are zero, the corresponding quantile function functions $F^{-1}(t)$ and $G^{-1}(t)$ satisfy

$$F^{-1}(t) = a(t)\,G^{-1}(t) \quad (\tfrac{1}{2} < t < 1), \tag{3.26}$$

with $a(t)$ nonincreasing.

In Hájek (1969), § 34, there is given a theorem according to which (3.25) with $b(t)$ nondecreasing entail (3.26) with $a(t)$ nonincreasing, provided $\phi(t) = \phi(t,f)$ and $\psi(t) = \phi(t,g)$, using notation (2.12). Consequently, the normal type has shorter tails than the logistic type, the latter type has shorter tails than the type corresponding to ϕ_3, and the last type has shorter tails than the Cauchy type.

Since (3.26) with $a(t)$ nonincreasing entails that

$$F^{-1}(G(x)) \quad (x>0), \quad \text{is concave,}$$

$$G^{-1}(F(x)) \quad (x>0), \quad \text{is convex,}$$

we can test whether the true ϕ-function increases more rapidly than ϕ_0 as follows: We plot $[\frac{1}{2}N]$ points with co-ordinates

$$\left[x^{(N+1-i)}-x^{(i)},\ F_0^{-1}\left(-\frac{N+1-i}{N+1}\right)\right] \quad (1 \leqslant i \leqslant [\tfrac{1}{2}N]),$$

where F_0^{-1} is the quantile function corresponding to ϕ_0.

If the curve suggested by these $[\frac{1}{2}N]$ points is convex, than we may prefer a ϕ-function which is increasing more slowly than ϕ_0; if it is concave, we may prefer a ϕ which increases more rapidly than ϕ_0; if it is straight, we keep ϕ_0. The effectiveness of this method and the possibility of its formalization may be in doubt. On the other hand, it surely is the quickest of all three methods explained in this section.

4 TESTING SCALE WHEN BOTH MEDIANS ARE UNKNOWN

Jurečková's theorem may also be applied to an asymptotic treatment of nuissance medians in testing scale alternatives. First observe that the increment $(S_N^* - S_N)/\sqrt{(\operatorname{var} S_N)}$ will be asymptotically zero if

$$\int_0^1 \phi(t)\,\phi(t,f)\,dt = 0. \tag{4.1}$$

Such a situation occurs if ϕ is symmetric,

$$\phi(t) = \phi(1-t) \quad (0<t<1), \tag{4.2}$$

and $\phi(t,f)$ is skew symmetric,

$$\phi(t,f) = -\phi(1-t,f) \quad (0<t<1). \tag{4.3}$$

Obviously, the skew symmetry of $\phi(t,f)$ is equivalent to the symmetry of f about its median.

Consider two samples $X_1, \ldots, X_m$ and $X_{m+1}, \ldots, X_{m+n}$ and assume that the respective densities are of the form $\sigma f[\sigma(x-\mu)]$ and $\lambda f[\lambda(x-\nu)]$. If $\mu=\nu$, proper rank tests for testing $\sigma=\lambda$ against $\sigma \neq \lambda$ (or $\sigma>\lambda$, or $\sigma<\lambda$) are based on statistics

$$S_N = s_N(X_1, \ldots, X_N), \tag{4.4}$$

where
$$s_N(x_1, \ldots, x_N) = \sum_{i=1}^{m_N} a_N\left(\sum_{j=1}^{N} u(x_i - x_j)\right) \tag{4.5}$$

and the scores $a_N(i)$ are generated by (2.7) or by (2.8). If $\mu \neq \nu$, but we have some estimates $\hat{\mu}_N$ and $\hat{\nu}_N$ for μ and ν_1 respectively, we may try to prove that the statistic

$$S_N^* = s_N(X_1 - \hat{\mu}_N, \ldots, X_m - \hat{\mu}_N, X_{m+1} - \hat{\nu}_N, \ldots, X_N - \hat{\nu}_N) \quad (4.6)$$

possesses the same limiting distribution as S_N.

Assumptions B:

(B 1) The scores $a_N(i)$ are generated by (2.7) or (2.8), or satisfy (2.15), where $\phi(t)$ is square integrable, fulfils (4.2), and is nondecreasing for $t \in (\frac{1}{2}, 1)$.

(B 2) Random vector **X** is governed by the density

$$p_N(x_1, \ldots, x_N) = \prod_{i=1}^{m_N} f(x_i - \mu) \prod_{j=1}^{n_N} f(x_{m+j} - \nu) \quad (m_N + n_N = N), \quad (4.7)$$

where μ and ν are arbitrary and f is a fixed density which is symmetric and possesses finite Fisher's information.

(B 3) The estimates $\hat{\mu}_m$ and $\hat{\nu}_n$ are square-root consistent, that is, random variables $\sqrt{(m)}\,(\hat{\mu}_m - \mu)$ and $\sqrt{(n)}\,(\hat{\nu}_n - \nu)$ are bounded in probability.

(B 4) $$\min(m_N, n_N) \to \infty.$$

Theorem 4.1. *Let assumptions B be satisfied and denote by ES_N and* var S_N *the expectation and variance of S_N under the hypothesis of randomness.*

Then S_N^* is asymptotically normal with parameter (ES_N, var S_N).

Moreover, if densities q_N are contiguous to densities p_N of (4.7), *and if S_N is asymptotically normal with parameters (a_N,* var S_N) *under q_N, then S_N^* is also asymptotically normal (a_N,* var S_N) *under q_N.*

Proof. Since S_N is a rank statistic, we may write equivalently

$$S_N = s_N(X_1 - \mu + \Delta_N d_{N1}, \ldots, X_m - \mu + \Delta_N d_{Nm}, X_{m+1} - \nu + \Delta_N d_{N,m+1}, \ldots, X_N - \nu + \Delta_N d_{NN}), \quad (4.8)$$

where

$$\Delta_N = (\mu - \hat{\mu}_N - \nu + \hat{\nu}_N)\left(\frac{1}{m_N} + \frac{1}{n_N}\right)^{-\frac{1}{2}} \quad (n_N = N - m_N,\ \hat{\mu}_N = \hat{\mu}_{m_N},\ \hat{\nu}_N = \hat{\nu}_{n_N}) \quad (4.9)$$

and

$$d_{Ni} = \left(\frac{1}{m_N} + \frac{1}{n_N}\right)^{\frac{1}{2}} \quad (1 \leqslant i \leqslant m_N), \quad (4.10)$$

$$= 0 \quad (m_N < i \leqslant N).$$