A Unified Theory of Hypothesis Testing Based on Rankings

Jianhong Pan
Department of Mathematics
University of Ottawa, Canada

supervised by

Mayer Alvo
Department of Mathematics
University of Ottawa, Canada

January 31, 1994

National Library of Canada

Acquisitions and Bibliographic Services Branch

395 Wellington Street
Ottawa, Ontario
K1A 0N4

Bibliothèque nationale du Canada

Direction des acquisitions et des services bibliographiques

395, rue Wellington
Ottawa (Ontario)
K1A 0N4

Your file Votre référence

Our file Notre référence

ISBN 0-315-95944-4

Canada

UNIVERSITÉ D'OTTAWA
UNIVERSITY OF OTTAWA

A Ph.D. Thesis

submitted to the School of Graduate Studies
and Research in partial fufilment of the requirement
for the degree of Doctor of Philosophy in Mathematics*

University of Ottawa
Ottawa, Ontario
Canada

(January 31, 1994)

*The Ph.D. Program is a joint program
with Carleton University, administered by the
Ottawa-Carleton Institute of Mathematics and Statistics

To my parents, sisters and brother

Acknowledgement

Most sincere thanks go to Professor Mayer Alvo for his valuable yet workable suggestions and patient guidances. Without his help, this work could hardly have been done.

Abstract

A unified theory of hypothesis testing based on the ranks of the data is proposed. A hypothesis testing problem often gives rise to two separate permutation sets corresponding to the data and to the alternative respectively. By defining the distance between permutation sets as the average of all distances between pairs of permutations, one from each set, it is possible to obtain various test statistics. The limiting distributions of test statistics derived by the unified approach herein are obtained under both the null hypothesis and contiguous alternatives.

The unified approach produces not only some well-known test statistics but also some new yet plausible test statistics. The corresponding results are extensions of the simple linear rank statistics defined by Hajek and Sidak (1967) to the generalized linear rank statistics and of the two-sample case to the the multi-sample case. Furthermore, a combined method was developed for the case of composite alternatives.

Contents

Chapter 1

Test Statistics

1.1 Introduction

The notion of distance functions between permutations has been exploited recently to produce new test statistics for some well-known non-parametric testing problems. For example, Diaconis and Graham (1977) defined a new measure of rank correlation in terms of Spearman's footrule and studied its asymptotic properties. Alvo, Cabilio and Feigin (1982) obtained a new test of concordance based on the Kendall's tau measure. Alvo and Cabilio (1991) also introduced the notion of compatibility and used it to define a notion of distance between incomplete permutations. This led to a new statistic for the problem of concordance in the case of a balanced incomplete block design. Critchlow (1985) also defined a notion of distance and used it to analyze partially ranked data. By introducing a measure of distance between sets of permutations, Critchlow (1986) later proposed a new approach to hypothesis testing based on ranks. The measure Critchlow employed was defined to be the minimum distance between pairs of rankings, one from each permutation sets, that is

$$d(\{\pi\}, E) = \min_{\mu \in \{\pi\},\ \nu \in E} d(\mu, \nu),$$

here the related notation are referred to section 1.2. This measure resulted in a number of new test statistics for the common testing problems. However no distributional results based on this measure were obtained, and consequently, no comparisons among the statistics were possible.

In this thesis, a modification of Critchlow's original approach is used to develop new test statistics. Specifically, our new approach is defined by (1.1) in which the measure of distance between permutation sets is taken to be the average of all distances between pairs of rankings, one from each permutation set. By this approach, it is possible to carry out the necessary steps to solve various hypothesis testing problems, though some difficulties are met. This unified approach not only gives some well-known statistics such as the Wilcoxon statistic and the Mann-Whitney statistic but also produces some new yet plausible statistics in non-parametric analysis. The corresponding results can be viewed as extensions of the simple linear rank statistics defined by Hajek and Sidak (1967) in the multi-sample case.

In Chapter 1, we derive the test statistics for different testing problems corresponding to the distances due to Spearman's rho, Kendall's tau, Spearman's footrule and Hamming. In Chapters 2 and 3, we obtain the limiting distributions of these statistics under the null and contiguous alternative hypothesis. In Chapter 4, we compute the asymptotic efficiencies. Finally, in Chapter 5, we discuss the multi-sample location problem with unordered alternative and develop a method for combining tests.

1.2 The Test Procedure

Let $P_n = \{\mu = [i_1, \ldots, i_n] : [i_1, \ldots, i_n]$ is a permutation of the integers $1, ..., n\}$. For any $\mu, \nu \in P_n$, define $d(\mu, \nu)$ as a distance function between μ and ν, where d is a function from $P_n \times P_n$ to R^1 and is usually chosen in practice to measure a particular characteristic. Some examples of distance functions are as follows:

1) Spearman's Rho $d_S(\mu, \nu) = \sum_{i=1}^{n} [\mu(i) - \nu(i)]^2$

2) Kendall's Tau $d_K(\mu, \nu) = \sum_{1=i<j=n} \{1 - sgn[\mu(i) - \mu(j)] sgn[\nu(i) - \nu(j)]\}$

3) Spearman's Footrule $d_F(\mu, \nu) = \sum_{i=1}^{n} |\mu(i) - \nu(i)|$

4) Hamming $d_H(\mu, \nu) = \sum_{i=1}^{n} I[\mu(i) \neq \nu(i)]$, where I is the indicator function.

It is easy to check that all these distance measures are right invariant in the sense that $d(\mu, \nu) = d(\mu\sigma, \nu\sigma)$ for any $\mu, \nu, \sigma \in P_n$. Hence, a relabelling of the items being ranked does not alter the value of the distance measure.

Consider the following problem of testing

$$H_0 \text{ against } H_1.$$

Let $X_1, ..., X_n$ be a collection of random variables and let $\pi = [\pi(X_1), ..., \pi(X_n)]$ be the corresponding vector of ranks. Assume that the underlying distributions of the X's are all continuous. Hence, $\pi \in P_n$. For simplicity, write $\pi = [\pi(1), ..., \pi(n)]$.

Define $\{\pi\}$ to be the subclass of permutations "equivalent" to π in the sense that ranks occupied by identically distributed random variables are exchangeable. Further, let E be the subclass of permutations most "in agreement" with the alternative H_1. For example, assume that $n = 4$ and let X_1, X_2 come from distribution $F_1(x)$ while X_3, X_4 come from distribution $F_2(x)$ with the observation data (X_1, X_2, X_3, X_4)=(3.5, 4.6, 3.9, 5.6). Then $[\pi(X_1), \pi(X_2), \pi(X_3), \pi(X_4)] = [1,3,2,4]$. Therefore $\{\pi\} = \{[1,3,2,4], [3,1,2,4], [1,3,4,2], [3,1,4,2]\}$. If the alternative hypothesis H_1 is such that $F_2(x) < F_1(x)$, then $E = \{[1,2,3,4], [2,1,3,4], [1,2,4,3], [2,1,4,3]\}$. Thus $\{\pi\}$ is derived from the sample whereas E depends entirely on H_1. The detailed features of $\{\pi\}$ and E will be made clearer when dealing with particular testing problems in which H_0 and H_1 are specified. To test H_0 against H_1, define the distance between $\{\pi\}$ and E:

$$d(\{\pi\}, E) = \sum_{\mu \in \{\pi\}} \sum_{\nu \in E} d(\mu, \nu), \tag{1.1}$$

which is the average distance between $\{\pi\}$ and E up to a constant. The test of H_0 against H_1 consists of rejecting H_0 whenever $d(\{\pi\}, E)$ is small. In this Chapter, we will consider some hypothesis testing situations and obtain the statistics corresponding to the distance functions 1)—4) defined above.

1.3 The Two-sample Case

Let $F_1(x)$ and $F_2(x)$ be two continuous distribution functions, and suppose that we wish to test

$$H_0 : F_1(x) = F_2(x) \text{ against } H_1 : F_1(x) \geq F_2(x) \tag{1.2}$$

with strict inequality for some x, i.e. population 1 is assumed to be stochastically larger. Let $X_1, ..., X_{n_1}$ be a random sample from $F_1(x)$ and $X_{n_1+1}, ..., X_{n_1+n_2}$ be a random sample from $F_2(x)$. By convention, write

$$\pi = [\pi(1), ..., \pi(n_1)|\pi(n_1+1), ..., \pi(n_1+n_2)] \tag{1.3}$$

so that the first n_1 items refer to the ranks from population 1 while the next n_2 items refer to the ranks from population 2. Right invariance of the distance measures justifies the convention. The equivalence subclass $\{\pi\}$ consists of all permutations which assign the same set of ranks to population 1 and population 2 respectively as π does, while E consists of all permutations which assign ranks $1,...,n_1$ to population 1 and ranks $n_1+1,...,n_1+n_2$ to population 2. The cardinality of $\{\pi\}$ = the cardinality of $E = n_1!n_2!$. Formally, we have

$$\begin{aligned}\{\pi\} &= \{[\pi(i_1), ..., \pi(i_{n_1})|\pi(n_1+j_1), ..., \pi(n_1+j_{n_2})] : [i_1, ..., i_{n_1}] \in P_{n_1}, \\ &\qquad [j_1, ..., j_{n_2}] \in P_{n_2}\} \qquad (1.4)\\ E &= \{[i_1, ..., i_{n_1}|n_1+j_1, ..., n_1+j_{n_2}] : [i_1, ..., i_{n_1}] \in P_{n_1}, \\ &\qquad [j_1, ..., j_{n_2}] \in P_{n_2}\} \qquad (1.5)\end{aligned}$$

Let $n = n_1 + n_2$. We are now in a position to derive the test statistics corresponding to the distance functions in section 1.2.

1) Spearman's Rho

$$\begin{aligned}d_{S_2}(\{\pi\}, E) &= \sum_{\mu\in\{\pi\}}\sum_{\nu\in E} d_{S_2}(\mu,\nu) = \sum_{\mu\in\{\pi\}}\sum_{\nu\in E}\sum_{i=1}^{n_1+n_2}[\mu(i)-\nu(i)]^2 \\ &= \sum_{\mu\in\{\pi\}}\sum_{\nu\in E}\sum_{i=1}^{n_1+n_2}[\mu^2(i)+\nu^2(i)] - 2\sum_{\mu\in\{\pi\}}\sum_{\nu\in E}\sum_{i=1}^{n_1+n_2}\mu(i)\nu(i) \\ &= \sum_{\mu\in\{\pi\}}\sum_{\nu\in E}\sum_{i=1}^{n_1+n_2} 2\times i^2 - 2\sum_{i=1}^{n_1+n_2}[\sum_{\mu\in\{\pi\}}\mu(i)][\sum_{\nu\in E}\nu(i)] \\ &= \frac{(n_1!n_2!)^2 n(n+1)(2n+1)}{3} - 2\{\sum_{i=1}^{n_1}[\sum_{j=1}^{n_1} n_2!(n_1-1)!\pi(j)][\sum_{j=1}^{n_1} n_2!(n_1-1)!j] + \\ &\quad \sum_{i=n_1+1}^{n_1+n_2}[\sum_{j=n_1+1}^{n_1+n_2} n_1!(n_2-1)!\pi(j)][\sum_{j=n_1+1}^{n_1+n_2} n_1!(n_2-1)!j]\} \\ &= \frac{(n_1!n_2!)^2 n(n+1)(2n+1)}{3} - (n_1!n_2!)^2[(n_1+1)\sum_{i=1}^{n_1}\pi(i) + (n+n_1+1)\sum_{j=n_1+1}^{n_1+n_2}\pi(j)] \\ &= \frac{(n_1!n_2!)^2 n(n+1)(4n-3n_1-1)}{6} - (n_1!n_2!)^2 n\sum_{i=n_1+1}^{n_1+n_2}\pi(i). \qquad (1.6)\end{aligned}$$

This test statistic is equivalent to the Wilcoxon statistic which rejects H_0 whenever the sum of the ranks of the second population is large.

2) Kendall's Tau

$$d_{K_2}(\{\pi\}, E) = \sum_{\mu\in\{\pi\}}\sum_{\nu\in E} d_{K_2}(\mu,\nu) = \sum_{\mu\in\{\pi\}}\sum_{\nu\in E}\sum_{1\le i<j\le n}\{1 - sgn[\mu(i)-\mu(j)]sgn[\nu(i)-\nu(j)]\}$$

$$= \sum_{\mu\in\{\pi\}} \sum_{\nu\in E} \sum_{1\le i<j\le n} 1 - \sum_{\mu\in\{\pi\}} \sum_{\nu\in E} \sum_{1\le i<j\le n} sgn[\mu(i)-\mu(j)]sgn[\nu(i)-\nu(j)]$$

$$= \frac{(n_1!n_2!)^2 n(n-1)}{2} - \sum_{1\le i<j\le n} \{ \sum_{\mu\in\{\pi\}} sgn[\mu(i)-\mu(j)]\}\{\sum_{\nu\in E} sgn[\nu(i)-\nu(j)]\}$$

$$= \frac{(n_1!n_2!)^2 n(n-1)}{2} - (\sum_{i<j \text{ in the same population}} + \sum_{i<j \text{ not in the same population}}) \cdot$$

$$\{ \sum_{\mu\in\{\pi\}} sgn[\mu(i)-\mu(j)]\}\{\sum_{\nu\in E} sgn[\nu(i)-\nu(j)]\}$$

$$= \frac{(n_1!n_2!)^2 n(n-1)}{2} -$$

$$(0 + \sum_{i<j \text{ not in the same population}} \{ \sum_{\mu\in\{\pi\}} sgn[\mu(i)-\mu(j)]\}\{\sum_{\nu\in E} sgn[\nu(i)-\nu(j)]\})$$

$$= \frac{(n_1!n_2!)^2 n(n-1)}{2} - \sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n_1+n_2} \{ \sum_{\mu\in\{\pi\}} sgn[\mu(i)-\mu(j)]\}\{\sum_{\nu\in E} sgn[\nu(i)-\nu(j)]\}$$

$$= \frac{(n_1!n_2!)^2 n(n-1)}{2} - \sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n_1+n_2} \{\frac{n_1!n_2!}{n_1 n_2} \sum_{i_1=1}^{n_1} \sum_{j_1=1}^{n_1+n_2} sgn[\pi(i_1)-\pi(j_1)]\}\{-n_1!n_2!\}$$

$$= \frac{(n_1!n_2!)^2 n(n-1)}{2} + (n_1!n_2!)^2 \sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n_1+n_2} sgn[\pi(i)-\pi(j)]. \tag{1.7}$$

Hence the use of the Kendall's Tau distance leads to the Mann-Whitney statistic which rejects H_0 whenever there are too many observations from $F_1(x)$ which have small valued rankings. Here $sgn[\pi(i) - \pi(j)] = 1$ if the i^{th} observation from $F_1(x)$ is greater than the j^{th} observation from $F_2(x)$ and -1 otherwise. It is known that the Mann-Whitney and Wilcoxon statistics are equivalent. In fact, it is straightforward to show that

$$\sum_{i=1}^{n_1+n_2} sgn[\pi(i)-\pi(j)] \quad = \quad n_1+n_2+1-2\pi(j).$$

Hence,

$$\sum_{i=1}^{n_1} sgn[\pi(i)-\pi(j)] \quad = \quad n_1+n_2+1-2\pi(j) - \sum_{i=n_1+1}^{n_1+n_2} sgn[\pi(i)-\pi(j)]$$

and

$$\sum_{j=n_1+1}^{n_1+n_2} \sum_{i=1}^{n_1} sgn[\pi(i)-\pi(j)] \quad = \quad n_2(n_1+n_2+1) - 2\sum_{j=n_1+1}^{n_1+n_2} \pi(j).$$

3) Spearman's Footrule

$$d_{F_2}(\{\pi\}, E) = \sum_{\mu\in\{\pi\}} \sum_{\nu\in E} d_{F_2}(\mu,\nu) = \sum_{\mu\in\{\pi\}} \sum_{\nu\in E} \sum_{i=1}^{n} |\mu(i)-\nu(i)|$$

$$= \sum_{i=1}^{n_1} \sum_{\mu \in \{\pi\}} \sum_{\nu \in E} |\mu(i) - \nu(i)| + \sum_{i=n_1+1}^{n_1+n_2} \sum_{\mu \in \{\pi\}} \sum_{\nu \in E} |\mu(i) - \nu(i)|$$

$$= \sum_{i=1}^{n_1} [n_2!(n_1-1)!]^2 \sum_{j_1=1}^{n_1} \sum_{j_2=1}^{n_1} |\pi(j_1) - j_2| + \sum_{i=n_1+1}^{n_1+n_2} [n_1!(n_2-1)!]^2 \sum_{j_1=n_1+1}^{n_1+n_2} \sum_{j_2=n_1+1}^{n_1+n_2} |\pi(j_1) - j_2|$$

$$= n_1[n_2!(n_1-1)!]^2 \sum_{j_1=1}^{n_1} \sum_{j_2=1}^{n_1} \{2[\pi(j_1) - j_2] I_{[\pi(j_1)>j_2]} - [\pi(j_1) - j_2]\} +$$

$$n_2[n_1!(n_2-1)!]^2 \sum_{j_1=n_1+1}^{n_1+n_2} \sum_{j_2=n_1+1}^{n_1+n_2} \{2[\pi(j_1) - j_2] I_{[\pi(j_1)>j_2]} - [\pi(j_1) - j_2]\}$$

$$= 2[n_1!n_2!]^2 \{ \sum_{j_1=1}^{n_1} \sum_{j_2=1}^{n_1} \frac{[\pi(j_1) - j_2] I_{[\pi(j_1)>j_2]}}{n_1} + \sum_{j_1=n_1+1}^{n_1+n_2} \sum_{j_2=n_1+1}^{n_1+n_2} \frac{[\pi(j_1) - j_2] I_{[\pi(j_1)>j_2]}}{n_2} \}$$

$$= 2[n_1!n_2!]^2 \{ - \sum_{j_1=n_1+1}^{n_1+n_2} \{ \sum_{j_2=1}^{n_1} \frac{[\pi(j_1) - j_2] I_{[\pi(j_1)>j_2]}}{n_1} - \sum_{j_2=n_1+1}^{n_1+n_2} \frac{[\pi(j_1) - j_2] I_{[\pi(j_1)>j_2]}}{n_2} \} +$$

$$\sum_{j_1=1}^{n_1+n_2} \sum_{j_2=1}^{n_1} \frac{[\pi(j_1) - j_2] I_{[\pi(j_1)>j_2]}}{n_1} \}$$

$$= 2[n_1!n_2!]^2 \{ - \sum_{j_1=n_1+1}^{n_1+n_2} \{ \sum_{j_2=1}^{n_1} \frac{[\pi(j_1) - j_2] I_{[\pi(j_1)>j_2]}}{n_1} - \sum_{j_2=n_1+1}^{n_1+n_2} \frac{[\pi(j_1) - j_2] I_{[\pi(j_1)>j_2]}}{n_2} \} + \frac{n_1^2 + 3nn_2 - 1}{6} \}. \tag{1.8}$$

The final term in the brackets equals to $\sum_{j_1=1}^{n_1+n_2} \sum_{j_2=1}^{n_1} \{[j_1 - j_2] I_{[j_1>j_2]}\}/n_1$. Here $[\pi(j_1) - j_2] I_{[\pi(j_1)>j_2]}$ for $1 \le j_2 \le n_1 + n_2$ and $n_1 < j_1 \le n_1 + n_2$ measures the excess in the ranks by which an observation from $F_2(x)$ exceeds the positive integer number j_2. Hence the use of the Spearman's Footrule distance leads to a new statistic which rejects H_0 whenever the weighted sum of the excesses is large.

4) Hamming

First introduce some notation as follows:

$$A_1 = \{1, ..., n_1\} \text{ and } A_2 = \{n_1 + 1, ..., n_1 + n_2\}$$
$$B_1 = \{\pi(1), ..., \pi(n_1)\} \text{ and } B_2 = \{\pi(n_1 + 1), ..., \pi(n_1 + n_2)\}$$
$$C_1 = A_1 \bigcap B_1,\ C_2 = A_2 \bigcap B_2, \text{ and } C_0 = (C_1 \bigcup C_2)^c$$

Y_1=the cardinality of C_1, Y_2=the cardinality of C_2, and Y_0=the cardinality of C_0.

Obviously $Y_0 = n - (Y_1 + Y_2)$. Note that under the permutation scheme prescribed by the distance between equivalence sets, a ranking in population 1 will be counted as being in order $(n_1 - 1)!n_2!$ times with respect to H_1 and hence out of order $n_1!n_2! - (n_1 - 1)!n_2! = (n_1 - 1)!n_2!(n_1 - 1)$ times. Similarly a ranking in population 2 will be counted as being out of order $(n_2 - 1)!n_1!(n_2 - 1)$ times. The total number of out of order ranking is $n_1!n_2!$. Therefore it is easy to see that

$$d_{H_2}(\{\pi\}, E) = \sum_{\mu \in \{\pi\}} \sum_{\nu \in E} d_{H_2}(\mu, \nu) = \sum_{\mu \in \{\pi\}} \sum_{\nu \in E} \sum_{i=1}^{n} I[\mu(i) \ne \nu(i)]$$

$$= \sum_{\mu \in \{\pi\}} \{ \sum_{i=1}^{n} \sum_{\nu \in E} I[\mu(i) \ne \nu(i)] \}$$

$$= \sum_{\mu\in\{\pi\}} \{\sum_{i\in C_1}\sum_{\nu\in E} I[\mu(i)\neq\nu(i)] + \sum_{i\in C_2}\sum_{\nu\in E} I[\mu(i)\neq\nu(i)] + \sum_{i\in C_0}\sum_{\nu\in E} I[\mu(i)\neq\nu(i)]\}$$
$$= \sum_{\mu\in\{\pi\}} \{\sum_{i\in C_1}(n_1-1)!n_2!(n_1-1) + \sum_{i\in C_2} n_1!(n_2-1)!(n_2-1) + \sum_{i\in C_0} n_1!n_2!\}$$
$$= \sum_{\mu\in\{\pi\}} \{Y_1(n_1-1)!n_2!(n_1-1) + Y_2 n_1!(n_2-1)!(n_2-1) + Y_0 n_1!n_2!\}$$
$$= n_1!n_2!\{(n_1-1)!(n_2-1)!(nn_1n_2 - n_2Y_1 - n_1Y_2)\}$$
$$= (n_1!n_2!)^2(n - \frac{Y_1}{n_1} - \frac{Y_2}{n_2}). \tag{1.9}$$

The new test statistic obtained from using the Hamming distance rejects H_0 whenever the sum $\frac{Y_1}{n_1}+\frac{Y_2}{n_2}$ is large, i.e. the average number of indices in order as specified by H_1 for $F_1(x)$ and for $F_2(x)$ is large.

1.4 The Multi-sample Case

The generalization of the location problem to r samples, $r > 2$, becomes

$$\begin{aligned} &H_0 : F_1(x) = \cdots = F_r(x) \textit{ against} \\ &H_1 : F_r(x) \leq \cdots \leq F_1(x) \end{aligned} \tag{1.10}$$

with strict inequality for some x; that is, the populations are stochastically ordered. Let $N_0 = 0$ and $N_k = n_1 + \cdots + n_k$ for $1 \leq k \leq r$, where n_k is the k^{th} sample size taken from $F_k(x)$, and assume that sample data $X_{N_{k-1}+1}, \ldots, X_{N_k}$ come from the k^{th} population. As in the two-sample case, write

$$\pi = [\pi(1), \ldots, \pi(N_1)|\pi(N_1+1), \ldots, \pi(N_2)|\cdots|\pi(N_{r-1}+1), \ldots, \pi(N_r)] \tag{1.11}$$

so that the n_k items in the k^{th} block refer to the ranks from the k^{th} population, for $1 \leq k \leq r$. Define

$$\pi(\mu) \stackrel{\text{def}}{=} [\pi(i_1), \ldots, \pi(i_n)], \text{ for } \mu = [i_1, \ldots, i_n] \in P_n$$

and set

$$(u) + \mu \stackrel{\text{def}}{=} [u + i_1, \ldots, u + i_n] \text{ for a positive integer } u \text{ and a permutation } \mu \in P_n$$
$$(u) + P_n \stackrel{\text{def}}{=} \{(u) + \mu : \mu \in P_n\}.$$

Then it can be seen that

$$\begin{aligned} \{\pi\} &= \{[\pi(\mu_1)\cdots\pi(\mu_r)] : \mu_k \in (N_{k-1}) + P_{n_k} \textit{ for } 1 \leq k \leq r\} \\ E &= \{[\mu_1\cdots\mu_r] : \mu_k \in (N_{k-1}) + P_{n_k} \textit{ for } 1 \leq k \leq r\}. \end{aligned} \tag{1.12}$$

We are now in a position to obtain the test statistics for each of the distance functions in section 1.2.

1) Spearman's Rho

$$d_{S_r}(\{\pi\}, E) = \sum_{\mu\in\{\pi\}}\sum_{\nu\in E} d_{S_r}(\mu,\nu) = \sum_{\mu\in\{\pi\}}\sum_{\nu\in E}\sum_{i=1}^{N_r}[\mu(i)-\nu(i)]^2$$
$$= \sum_{\mu\in\{\pi\}}\sum_{\nu\in E}\sum_{i=1}^{N_r} 2i^2 - 2\sum_{k=1}^{r}\sum_{i=N_{k-1}+1}^{N_k}[\sum_{\mu\in\{\pi\}}\mu(i)][\sum_{\nu\in E}\nu(i)]$$

$$= \frac{\prod_{k=1}^{r}(n_k!)^2 N_r(N_r+1)(2N_r+1)}{3} - 2\sum_{k=1}^{r}\sum_{i=N_{k-1}+1}^{N_k}[\sum_{j=N_{k-1}+1}^{N_k}\frac{\prod_{k=1}^{r}n_k!}{n_k}\pi(j)][\sum_{j=N_{k-1}+1}^{N_k}\frac{\prod_{k=1}^{r}n_k!}{n_k}j]$$

$$= \frac{\prod_{k=1}^{r}(n_k!)^2 N_r(N_r+1)(4N_r-1)}{6} - \prod_{k=1}^{r}(n_k!)^2\sum_{k=1}^{r}\sum_{i=N_{k-1}+1}^{N_k}(N_{k-1}+N_k)\pi(i). \tag{1.13}$$

This test statistic which generalizes the Wilcoxon statistic rejects H_0 whenever the sum of the weighted ranks from the r populations is large.

2) Kendall's Tau

$$d_{K_r}(\{\pi\},E) = \sum_{\mu\in\{\pi\}}\sum_{\nu\in E} d_{K_r}(\mu,\nu) = \sum_{\mu\in\{\pi\}}\sum_{\nu\in E}\sum_{1\le i<j\le N_r}\{1 - sgn[\mu(i)-\mu(j)]sgn[\nu(i)-\nu(j)]\}$$

$$= \frac{\prod_{k=1}^{r}(n_k!)^2 N_r(N_r-1)}{2} - \sum_{1\le i<j\le N_k}\{\sum_{\mu\in\{\pi\}} sgn[\mu(i)-\mu(j)]\}\{\sum_{\nu\in E} sgn[\nu(i)-\nu(j)]\}$$

$$= \frac{\prod_{k=1}^{r}(n_k!)^2 N_r(N_r-1)}{2} - (\sum_{i<j \text{ in the same population}} + \sum_{i<j \text{ not in the same population}})\cdot$$

$$\{\sum_{\mu\in\{\pi\}} sgn[\mu(i)-\mu(j)]\}\{\sum_{\nu\in E} sgn[\nu(i)-\nu(j)]\}$$

$$= \frac{\prod_{k=1}^{r}(n_k!)^2 N_r(N_r-1)}{2} - (0 + \sum_{i<j \text{ not in the same population}})\cdot$$

$$\{\sum_{\mu\in\{\pi\}} sgn[\mu(i)-\mu(j)]\}\{\sum_{\nu\in E} sgn[\nu(i)-\nu(j)]\}$$

$$= \frac{\prod_{k=1}^{r}(n_k!)^2 N_r(N_r-1)}{2} - \sum_{k=1}^{r-1}\sum_{t=k}^{r-1}\sum_{i=N_{k-1}+1}^{N_k}\sum_{j=N_t+1}^{N_{t+1}}\{\sum_{\mu\in\{\pi\}} sgn[\mu(i)-\mu(j)]\}\{\sum_{\nu\in E} sgn[\nu(i)-\nu(j)]\}$$

$$= \frac{\prod_{k=1}^{r}(n_k!)^2 N_r(N_r-1)}{2} + \prod_{k=1}^{r}(n_k!)^2\sum_{k=1}^{r-1}\sum_{t=k}^{r-1}\sum_{i=N_{k-1}+1}^{N_k}\sum_{j=N_t+1}^{N_{t+1}}\{sgn[\pi(i)-\pi(j)]\}. \tag{1.14}$$

This generalized Mann-Whitney statistic rejects H_0 whenever the sum of all terms $sgn[\pi(i)-\pi(j)]$ is small where i and j come from the different populations. It is no longer the case that (1.14) and (1.13) are equivalent as in the two sample case. Instead it will be shown in section 2.4 of Chapter 2 that these statistics are asymptotically equivalent in terms of variance.

3) Spearman's Footrule

$$d_{F_r}(\{\pi\},E) = \sum_{\mu\in\{\pi\}}\sum_{\nu\in E} d_{F_r}(\mu,\nu) = \sum_{\mu\in\{\pi\}}\sum_{\nu\in E}\sum_{i=1}^{N_r}|\mu(i)-\nu(i)|$$

$$= \sum_{k=1}^{r}\sum_{N_{k-1}+1}^{N_k}\sum_{\mu\in\{\pi\}}\sum_{\nu\in E}|\mu(i)-\nu(i)|$$

$$= \sum_{k=1}^{r}\sum_{N_{k-1}+1}^{N_k}\frac{\prod_{k=1}^{r}(n_k!)^2}{n_k^2}\sum_{j_1=N_{k-1}+1}^{N_k}\sum_{j_2=N_{k-1}+1}^{N_k}|\pi(j_1)-j_2|$$

$$= \prod_{k=1}^{r}(n_k!)^2 \sum_{k=1}^{r} \sum_{j_1=N_{k-1}+1}^{N_k} \sum_{j_2=N_{k-1}+1}^{N_k} \frac{2[\pi(j_1)-j_2]I_{[\pi(j_1)>j_2]} - [\pi(j_1)-j_2]}{n_k}$$

$$= 2\prod_{k=1}^{r}(n_k!)^2 \sum_{k=1}^{r} \sum_{j_1=N_{k-1}+1}^{N_k} \sum_{j_2=N_{k-1}+1}^{N_k} \frac{[\pi(j_1)-j_2]I_{[\pi(j_1)>j_2]}}{n_k}. \tag{1.15}$$

This new statistic rejects H_0 whenever the weighted sum of all terms $[\pi(j_1)-j_2]I_{[\pi(j_1)>j_2]}$ is small.

4) Hamming

For $1 \le k \le r$, let

$$A_k = \{N_{k-1}+1, \cdots, N_k\}$$
$$B_k = \{\pi(N_{k-1}+1), \cdots, \pi(N_k)\}$$
$$C_k = A_k \bigcap B_k,\ C_0 = (C_1 \bigcup \cdots \bigcup C_r)^c$$
$$Y_k = \text{the cardinality of } C_k,\ Y_0 = \text{the cardinality of } C_0.$$

It is easy to see that $Y_0 = N_r - (Y_1 + \cdots + Y_r)$. As in the two-sample case, we can obtain

$$\sum_{\mu\in\{\pi\}} \sum_{\nu\in E} d_{H_r}(\mu,\nu) = \sum_{\mu\in\{\pi\}} \sum_{\nu\in E} \sum_{i=1}^{N_r} I[\mu(i)\neq\nu(i)]$$
$$= \sum_{\mu\in\{\pi\}} \{\sum_{i=1}^{N_r} \sum_{\nu\in E} I[\mu(i)\neq\nu(i)]\}$$
$$= \sum_{\mu\in\{\pi\}} \{\sum_{k=0}^{r} \sum_{i\in C_k} \sum_{\nu\in E} I[\mu(i)\neq\nu(i)]\}$$
$$= \sum_{\mu\in\{\pi\}} \{\sum_{k=1}^{r} \sum_{i\in C_k} \frac{n_1!\dots n_r!(n_k-1)}{n_k} + \sum_{i\in C_0} n_1!\dots n_r!\}$$
$$= \sum_{\mu\in\{\pi\}} \{\sum_{k=1}^{r} Y_k \frac{n_1!\dots n_r!(n_k-1)}{n_k} + Y_0 n_1!\dots n_r!\}$$
$$= \sum_{\mu\in\{\pi\}} \{-\sum_{k=1}^{r} Y_k \frac{n_1!\dots n_r!}{n_k} + N_r n_1!\dots n_r!\}$$
$$= \prod_{k=1}^{r}(n_k!)^2 (N_r - \sum_{k=1}^{r} \frac{Y_k}{n_k}). \tag{1.16}$$

This test statistic is a new statistic which rejects H_0 whenever the average $\frac{Y_1}{n_1} + \cdots + \frac{Y_r}{n_r}$ is large.

1.5 The Two-way Layout Case: the ordered alternative

Let $X_i^j, j = 1\dots b$ and $i = 1\dots r$ be rb mutually independent random variables such that X_i^j has a continuous distribution function $F_i(x)$. In this section we consider the problem of testing

$$H_0 : F_1(x) = \cdots = F_r(x) \textit{ against}$$
$$H_1 : F_r(x) \le \cdots \le F_1(x) \quad \textit{with strict inequality for at least one } x. \tag{1.17}$$

The null hypothesis is that for fixed i, each $X_i^j, j = 1 \ldots b$, has the same distribution $F_i(x)$. This is referred to as a two-way layout problem with no interaction. The subscript i on X_i^j may be thought of as designating the treatment whereas the superscript j refers to the block. In that case H_0 is the hypothesis that there are no differences among the r treatments.

Let π be a permutation ranking of rb items. For each fixed $j = 1 \ldots b$, denote by $R^j = [R_1^j, \ldots, R_r^j]$ a permutation of the integers $1, \ldots, r$ which preserves the relative order of $X_1^j \ldots X_r^j$ prescribed by π. Under H_0, the random variables $X_1^j \ldots X_r^j$ constitute a sample from a continuous distribution function, say $F(x)$, and consequently R^j is uniformly distributed over P_r, (Randles and Wolfe (1979), Theorem 2.3.6). In view of the independence between blocks, any statistic which is a function of the data only through the within-blocks ranks $R^1, \cdots, R^b$, will be distribution-free over the class of continuous distributions, (Randles and Wolfe (1979), Corollary 2.3.6).

For convenience, write $\mu \in \{\pi\}, \nu \in E$ as

$$\begin{aligned} \mu &= [\mu_1(1), \ldots, \mu_1(r) | \ldots | \mu_b(1), \ldots, \mu_b(r)] \\ \nu &= [\nu_1(1), \ldots, \nu_1(r) | \ldots | \nu_b(1), \ldots, \nu_b(r)]. \end{aligned} \tag{1.18}$$

Hence the two equivalent subclasses are

$$\{\pi\} = \{\mu \in P_{rb} : [\mu_j(1), \ldots, \mu_j(r)] \textit{ has the same relative order as } R^j \textit{ for } j = 1, \ldots, b\} \tag{1.19}$$

$$E = \{\nu \in P_{rb} : [\nu_j(1), \ldots, \nu_j(r)] \textit{ satisfies } \nu_j(1) < \ldots < \nu_j(r) \textit{ for } j = 1, \ldots, b\}. \tag{1.20}$$

Before determining the test statistic, we introduce the concept of compatibility due to Alvo and Cabilio (1991).

Compatibility A complete ranking $\mu \in P_n$ is said to be compatible with an incomplete ranking μ^* if the relative ranking of every pair of objects ranked in μ^* coincides with their relative ranking in μ.

Ordering the complete rankings in P_n in some way, say $P_n = \{\mu_i : i = 1 \ldots n!\}$, we may associate with every incomplete ranking μ^* an $n! \times 1$ compatibility vector, denoted by C_{μ^*}, whose i^{th} component is 1 or 0 according to whether μ_i is compatible with μ^* or not.

The concept of compatibility can be extended to apply to a subclass of rankings G in the way of associating with every subclass of rankings G an $n! \times 1$ compatibility vector, denoted by C_G, whose i^{th} component is 1 or 0 according to whether μ_i is included in G or not.

Thus by the above definition, $C_{\{\pi\}}, C_E$ are $(rb)! \times 1$ vectors compatible with $\{\pi\}, E$ respectively, with $n = rb$. Formally, for $1 \leq i \leq r, 1 \leq j \leq b$, the $(r(j-1)+i)^{th}$ component of $C_{\{\pi\}}$ is

$$C_{\{\pi\}}(r(j-1)+i) = \begin{cases} 1 & \text{if } \mu_{r(j-1)+i} \in \{\pi\} \\ 0 & \text{otherwise.} \end{cases}$$

The expression for C_E is similar. Let $Q = (d(\mu_s, \mu_t))_{(rb)! \times (rb)!}$ be an $(rb)! \times (rb)!$ matrix, where $\mu_s, \mu_t \in P_{rb}$ for $1 \leq s, t \leq n!$. It follows immediately that

$$d(\{\pi\}, E) = \sum_{\mu \in \{\pi\}} \sum_{\nu \in E} d(\mu, \nu) = C_E' Q C_{\{\pi\}}. \tag{1.21}$$

The following theorems due to Feigin and Alvo (1986) characterize the matrices Q corresponding to the distances due to Spearman's Rho, Kendall's Tau and to Spearman's Footrule.

Theorem 1.1 *(Feigin and Alvo, 1986) For $\mu \in P_{rb}$, let*

$$t_S(\mu) = (\mu(1) - (rb+1)/2, \ldots, \mu(rb) - (rb+1)/2)'$$

and define the $(rb) \times (rb)!$ matrix $T_S = (t_S(\mu_1), \ldots, t_S(\mu_{(rb)!}))$. Then

$$Q_S = \frac{rb(rb+1)(rb-1)}{12} J - T_S' T_S.$$

Theorem 1.2 *(Feigin and Alvo, 1986) For $\mu \in P_{rb}$, let*

$$(t_K(\mu))(s) = sgn[\mu(j) - \mu(i)], s = 1 \dots rb(rb-1)/2$$
$$\text{where } s = (i-1)(rb - i/2) + (j-i), 1 \leq i < j \leq rb.$$

Let the $[rb(rb-1)/2] \times (rb)!$ matrix $T_K = (t_K(\mu_1), \dots, t_K(\mu_{(rb)!}))$. Then

$$Q_K = \frac{rb(rb-1)}{2} J - T_K' T_K.$$

Theorem 1.3 *(Feigin and Alvo, 1986) For $\mu \in P_{rb}$, let*

$$(t_F(\mu))((i-1)rb + j) = I[\mu(i) \leq j] - \frac{j}{rb}, 1 \leq i, j \leq rb.$$

Let the $(rb)^2 \times (rb)!$ matrix $T_F = (t_F(\mu_1), \dots, t_F(\mu_{(rb)!}))$. Then

$$Q_F = \frac{(rb+1)(rb-1)}{6} J - T_F' T_F.$$

Set, for $1 \leq j \leq b$,

$$\{\pi_j\} = \{\mu \in P_{rb} : [\mu_j(1) \dots \mu_j(r)] \text{ keeps the same relative order as } R^j \text{ and no order for other blocks}\} \tag{1.22}$$

$$E_j = \{\nu \in P_{rb} : [\nu_j(1) \dots \nu_j(r)] \text{ satisfies } \nu_j(1) < \dots < \nu_j(r) \text{ and no order for other blocks}\}. \tag{1.23}$$

We may now to determine the test statistics for the distance functions in section 1.2.

1) Spearman's Rho

Lemma 1.1 *For the Spearman's Rho distance,*

$$1)\ T_S C_{\{\pi\}} = \frac{1}{(r!)^{b-1}} (T_S C_{\{\pi_1\}} + \dots + T_S C_{\{\pi_b\}})$$
$$2)\ (T_S C_{\{\pi_p\}})'(T_S C_{\{\pi_s\}}) = 0, for\ 1 \leq p, s \leq b, p \neq s.$$

Proof: Let $(T_S C_{\{\pi\}})_{(i,j)}$ be the component representing the i^{th} treatment in the j^{th} block. From now on, similar notation will be used for other distance case. Noting the notation (1.18), for $1 \leq j \leq b, 1 \leq i \leq r$, we have

$$(T_S C_{\{\pi\}})_{(i,j)} = \sum_{s=1}^{(rb)!} [(\mu_s)_j(i) - (rb+1)/2] C_{\{\pi\}}(s)$$
$$= \sum_{\mu \in \{\pi\}} [\mu_j(i) - (rb+1)/2]$$
$$= \sum_{\mu \in \{\pi\}} \mu_j(i) - \frac{rb+1}{2} \binom{br}{r} \cdots \binom{2r}{r}$$
$$= \sum_{t=R_i^j}^{rb-(r-R_i^j)} t \binom{t-1}{R_i^j - 1} \binom{rb-t}{r - R_i^j} \binom{(b-1)r}{r} \cdots \binom{2r}{r} - \frac{rb+1}{2} \binom{br}{r} \cdots \binom{2r}{r}$$

$$= \sum_{t=R_i^j}^{rb-(r-R_i^j)} R_i^j \begin{pmatrix} t \\ R_i^j \end{pmatrix} \begin{pmatrix} rb-t \\ r-R_i^j \end{pmatrix} \begin{pmatrix} (b-1)r \\ r \end{pmatrix} \cdots \begin{pmatrix} 2r \\ r \end{pmatrix} - \frac{rb+1}{2} \begin{pmatrix} br \\ r \end{pmatrix} \cdots \begin{pmatrix} 2r \\ r \end{pmatrix}$$

$$= R_i^j \begin{pmatrix} rb+1 \\ r+1 \end{pmatrix} \begin{pmatrix} (b-1)r \\ r \end{pmatrix} \cdots \begin{pmatrix} 2r \\ r \end{pmatrix} - \frac{rb+1}{2} \begin{pmatrix} br \\ r \end{pmatrix} \cdots \begin{pmatrix} 2r \\ r \end{pmatrix}$$

$$= \frac{(rb+1)!}{(r+1)!(r!)^{(b-1)}}[R_i^j - \frac{(r+1)}{2}]. \tag{1.24}$$

The fourth equality above stems from the fact that $\mu_j(i) = t$, where t is any of the $R_i^j, \ldots, rb-(r-R_i^j)$, with multiplicity $\begin{pmatrix} t-1 \\ R_i^j - 1 \end{pmatrix} \begin{pmatrix} rb-t \\ r-R_i^j \end{pmatrix}$. The remaining components in the populations other than the j^{th} block can be filled in $\begin{pmatrix} (b-1)r \\ r \end{pmatrix} \cdots \begin{pmatrix} 2r \\ r \end{pmatrix}$ ways.

Similarly, for a fixed $t, 1 \le t \le b$, we have

$$(T_S C_{\{\pi_t\}})_{(i,j)} = \sum_{s=1}^{(rb)!}[(\mu_s)_j(i) - (rb+1)/2]C_{\{\pi_t\}}(s)$$

$$= \sum_{\mu \in \{\pi_t\}} \mu_j(i) - \sum_{\mu \in \{\pi_t\}} \frac{rb+1}{2}$$

$$= \sum_{\mu \in \{\pi_t\}} \mu_j(i) - \frac{(rb+1)[(b-1)r]!}{2} \begin{pmatrix} br \\ r \end{pmatrix}$$

$$= \begin{cases} \sum_{t=R_i^j}^{rb-(r-R_i^j)} t \begin{pmatrix} t-1 \\ R_i^j - 1 \end{pmatrix} \begin{pmatrix} rb-t \\ r-R_i^j \end{pmatrix} [(b-1)r]! - \frac{(rb+1)[(b-1)r]!}{2} \begin{pmatrix} br \\ r \end{pmatrix} & \text{if } t = j \\ \sum_{t=1}^{rb} t \begin{pmatrix} rb-1 \\ r \end{pmatrix} (rb-1-r)! - \frac{(rb+1)[(b-1)r]!}{2} \begin{pmatrix} br \\ r \end{pmatrix} & \text{if } t \ne j \end{cases}$$

$$= \begin{cases} R_i^j \begin{pmatrix} br+1 \\ r+1 \end{pmatrix} [(b-1)r]! - \frac{(rb+1)[(b-1)r]!}{2} \begin{pmatrix} br \\ r \end{pmatrix} & \text{if } t = j \\ 0 & \text{if } t \ne j \end{cases}$$

$$= \begin{cases} \frac{(rb+1)!}{(r+1)!}[R_i^j - \frac{r+1}{2}] & \text{if } t = j \\ 0 & \text{if } t \ne j. \end{cases} \tag{1.25}$$

By (1.24) and (1.25) we have

$$(T_S C_{\{\pi\}} - \frac{1}{(r!)^{b-1}}(T_S C_{\{\pi_1\}} + \ldots + T_S C_{\{\pi_b\}}))_{(i,j)}$$

$$= \frac{(rb+1)!}{(r+1)!(r!)^{(b-1)}}[R_i^j - \frac{(r+1)}{2}] - \frac{(rb+1)!}{(r+1)!(r!)^{(b-1)}}[R_i^j - \frac{(r+1)}{2}]$$

$$= 0. \tag{1.26}$$

This completes the proof of 1). Part 2) follows directly from (1.25).

Lemma 1.2 *Let π and π^* be any two permutations in P_{rb}, $\pi \ne \pi^*$. Then*

1) $(T_S C_{\{\pi_p\}})'(T_S C_{\{\pi_s^*\}}) = 0$, *for* $1 \le p,\ s \le b$ *and* $p \ne s$

2) $(T_S C_{\{\pi\}})'(T_S C_{\{\pi^*\}})$

$$= \frac{1}{(r!)^{2(b-1)}}[(T_S C_{\{\pi_1\}})'(T_S C_{\{\pi_1^*\}} + \cdots + (T_S C_{\{\pi_b\}})'(T_S C_{\{\pi_b^*\}})].$$

Proof: The proof follows directly from Lemma 1.1.

The next theorem shows that the Spearman's Rho distance induces Page's statistic, Page (1963).

Theorem 1.4

$$d_S(\{\pi\}, E) = \frac{rb(rb+1)(2rb+1)[(rb)!]^2}{6(r!)^{2b}} - \frac{[(rb+1)!]^2}{(r!)^{2b}(r+1)^2}\sum_{j=1}^{b}\sum_{i=1}^{r} iR_i^j. \quad (1.27)$$

Proof: It follows from (1.21), Theorem 1.1 and Lemma 1.2 that

$$\begin{aligned} d_S(\{\pi\}, E) &= \sum_{\mu\in\{\pi\}}\sum_{\nu\in E} d_S(\mu,\nu) = C_E' Q_S C_{\{\pi\}} \\ &= C_E'[\frac{rb(rb+1)(rb-1)}{12}J - T_S'T_S]C_{\{\pi\}} \\ &= \frac{rb(rb+1)(rb-1)}{12}C_E' J C_{\{\pi\}} - C_E' T_S' T_S C_{\{\pi\}} \\ &= \frac{rb(rb+1)(rb-1)}{12}[\binom{rb}{r}\cdots\binom{2r}{r}]^2 - \frac{(T_S C_{\{\pi_1\}})'(T_S C_{E_1}) + \cdots + (T_S C_{\{\pi_b\}})'(T_S C_{E_b})}{(r!)^{2(b-1)}} \\ &= \frac{rb(rb+1)(rb-1)[(rb)!]^2}{12(r!)^{2b}} - \frac{(T_S C_{\{\pi_1\}})'(T_S C_{E_1}) + \cdots + (T_S C_{\{\pi_b\}})'(T_S C_{E_b})}{(r!)^{2(b-1)}}. \end{aligned}$$

The theorem follows immediately from (1.25).

2) Kendall's Tau

Lemma 1.3 *For the Kendall's Tau distance, there exists a* $\binom{rb}{2} \times 1$ *vector* $C_{\{\pi_t\}}$ *such that*

$$\begin{aligned} &1)\ (C_{\{\pi_0\}})'(C_{\{\pi_0\}}) = \text{a constant depending only on } r \text{ and } b, \\ &2)\ (T_K C_{\{\pi_1\}} + \ldots + T_K C_{\{\pi_b\}})' C_{\{\pi_0\}} = 0, \\ &3)\ (T_K C_{\{\pi_p\}})'(T_K C_{\{\pi_s\}}) = 0, \text{ for } 1 \le p,\ s \le b, p \ne s, \\ \text{and}\quad &4)\ T_K C_{\{\pi\}} = \frac{1}{(r!)^{b-1}}(T_K C_{\{\pi_1\}} + \ldots + T_K C_{\{\pi_b\}}) + C_{\{\pi_0\}}. \end{aligned}$$

Proof: Noting that, *for* $j = 1\ldots b$, $T_K C_{\{\pi_j\}}$ is a $\binom{rb}{2} \times 1$ vector, we write its component as $(T_K C_{\{\pi_j\}})_{(i_1,j_1;i_2,j_2)}$ where either $1 \le j_1 = j_2 \le b, 1 \le i_1 < i_2 \le r$ *or* $1 \le j_1 < j_2 \le b, 1 \le i_1, i_2 \le r$. A similar notation is used for $T_K C_{\{\pi\}}$. Furthermore define the set $\{\pi\}_{j_1,j_2}$ to be the subclass of all permutations in P_{2r} such that $\mu \in \{\pi\}_{j_1,j_2}$ iff $[\mu_{j_t}(1), \cdots, \mu_{j_t}(r)]$ has the same relative order as $R^{j_t} = [R_1^{j_t}, \cdots, R_r^{j_t}]$, for $t = 1, 2$. Then

if $1 \le j_1 = j_2 \le b, 1 \le i_1 < i_2 \le r$,

$$\begin{aligned} (T_K C_{\{\pi_j\}})_{(i_1,j_1;i_2,j_2)} &= \begin{cases} \binom{rb}{r}[(b-1)r]!sgn[R_{i_2}^j - R_{i_1}^j] & \text{if } j = j_1 = j_2 \\ 0 & \text{if } j \ne j_1 = j_2 \end{cases} \\ &= \begin{cases} \frac{(rb)!}{r!}sgn[R_{i_2}^j - R_{i_1}^j] & \text{if } j = j_1 = j_2 \\ 0 & \text{if } j \ne j_1 = j_2 \end{cases} \quad (1.28) \end{aligned}$$

if $1 \le j_1 < j_2 \le b, 1 \le i_1, i_2 \le r$,

$$(T_K C_{\{\pi_j\}})_{(i_1,j_1;i_2,j_2)} = \begin{cases} \binom{rb}{2r} [(b-2)r]!(r-1)! a[R^j_{i_1}, \cdot] & \text{if } j = j_1 \\ \binom{rb}{2r} [(b-2)r]!(r-1)! a[\cdot, R^j_{i_2}] & \text{if } j = j_2 \\ 0 & \text{otherwise} \end{cases}$$

$$= \begin{cases} \frac{(rb)!(r-1)!}{(2r)!} a[R^j_{i_1}, \cdot] & \text{if } j = j_1 \\ \frac{(rb)!(r-1)!}{(2r)!} a[\cdot, R^j_{i_2}] & \text{if } j = j_2 \\ 0 & \text{otherwise} \end{cases} \quad (1.29)$$

and

$$(T_K C_{\{\pi\}})_{(i_1,j_1;i_2,j_2)} = \begin{cases} \binom{rb}{r} \cdots \binom{2r}{r} sgn[R^{j_1}_{i_2} - R^{j_1}_{i_1}] & \text{if } j_1 = j_2 \\ \binom{rb}{r} \cdots \binom{3r}{r} a[R^{j_1}_{i_1}, R^{j_2}_{i_2}] & \text{if } j_1 < j_2 \end{cases}$$

$$= \begin{cases} \frac{(rb)!}{(r!)^b} sgn[R^{j_1}_{i_2} - R^{j_1}_{i_1}] & \text{if } j_1 = j_2 \\ \frac{(rb)!}{(r!)^{b-2}(2r)!} a[R^{j_1}_{i_1}, R^{j_2}_{i_2}] & \text{if } j_1 < j_2 \end{cases} \quad (1.30)$$

where, for $j_1 < j_2$,

$$\begin{cases} a[R^{j_1}_{i_1}, R^{j_2}_{i_2}] = \sum_{\mu \in \{\pi\}_{j_1,j_2}} sgn[\mu_{j_2}(i_2) - \mu_{j_1}(i_1)] \\ a[\cdot, R^{j_2}_{i_2}] = \sum_{i_1=1}^{r} a[R^{j_1}_{i_1}, R^{j_2}_{i_2}] \\ a[R^{j_1}_{i_1}, \cdot] = \sum_{i_2=1}^{r} a[R^{j_1}_{i_1}, R^{j_2}_{i_2}] \\ a[\cdot,\cdot] = \sum_{i_1=1}^{r} \sum_{i_2=1}^{r} a[R^{j_1}_{i_1}, R^{j_2}_{i_2}] = 0 & \text{by symmetry.} \end{cases} \quad (1.31)$$

Note that, if $\mu_{j_2}(i_2) < \mu_{j_1}(i_1)$, then $\mu_{j_1}(i_1)$ can take values ranging from $R^{j_1}_{i_1} + R^{j_2}_{i_2}$ to $r + R^{j_1}_{i_1}$. Hence

$$\begin{aligned} a[R^{j_1}_{i_1}, R^{j_2}_{i_2}] &= \binom{2r}{r} - 2 \sum_{\mu \in \{\pi\}_{j_1,j_2},\ \mu_{j_2}(i_2) < \mu_{j_1}(i_1)} 1 \\ &= \binom{2r}{r} - \sum_{t=R^{j_1}_{i_1}+R^{j_2}_{i_2}}^{r+R^{j_1}_{i_1}} \binom{t-1}{R^{j_1}_{i_1}-1} \binom{2r-t}{r-R^{j_1}_{i_1}} \\ &= \binom{2r}{r} - \sum_{t=0}^{r-R^{j_2}_{i_2}} \binom{t+R^{j_1}_{i_1}+R^{j_2}_{i_2}-1}{R^{j_1}_{i_1}-1} \binom{2r-t-R^{j_1}_{i_1}-R^{j_2}_{i_2}}{r-R^{j_1}_{i_1}}. \end{aligned}$$

By the combinatorial formula in Feller (1950), page 65,

$$\begin{aligned} a[\cdot, R^{j_2}_{i_2}] &= \sum_{i_1=1}^{r} a[R^{j_1}_{i_1}, R^{j_2}_{i_2}] = \sum_{i_1=1}^{r} a[i_1, R^{j_2}_{i_2}] \\ &= \binom{2r}{r} - 2 \sum_{t=0}^{r-R^{j_2}_{i_2}} \sum_{i_1=1}^{r} \binom{t+i_1+R^{j_2}_{i_2}-1}{i_1-1} \binom{2r-t-i_1-R^{j_2}_{i_2}}{r-i_1} \end{aligned}$$

$$= \begin{pmatrix} 2r \\ r \end{pmatrix} - 2 \sum_{t=0}^{r-R_{i_2}^{j_2}} \begin{pmatrix} 2r \\ r-1 \end{pmatrix}$$

$$= \frac{(2r)!}{(r!)^2} \frac{r+1-2r!}{r+1} + 2 \begin{pmatrix} 2r \\ r-1 \end{pmatrix} R_{i_2}^{j_2}. \tag{1.32}$$

Similarly

$$a[R_{i_1}^{j_1}, \cdot] = \sum_{i_2=1}^{r} a[R_{i_1}^{j_1}, R_{i_2}^{j_2}] = \sum_{i_2=1}^{r} a[R_{i_1}^{j_1}, i_2]$$

$$= \frac{(2r)!}{(r!)^2} \frac{2r! - r - 1}{r+1} - 2 \begin{pmatrix} 2r \\ r-1 \end{pmatrix} R_{i_1}^{j_1}. \tag{1.33}$$

By (1.28), (1.29) and (1.30)

$$(T_K C_{\{\pi\}} - \frac{1}{(r!)^{b-1}} (T_K C_{\{\pi_1\}} + \ldots + T_K C_{\{\pi_b\}}))_{(i_1,j_1;i_2,j_2)}$$

$$= \begin{cases} \frac{(rb)!}{(r!)^{b-2}(2r)!} \{a[R_{i_1}^{j_1}, R_{i_2}^{j_2}] - \frac{a[\cdot, R_{i_2}^{j_2}] + a[R_{i_1}^{j_1}, \cdot]}{r}\} & \text{if } j_1 < j_2 \\ 0 & \text{if } j_1 = j_2. \end{cases} \tag{1.34}$$

Set

$$(C_{\{\pi_0\}})_{(i_1,j_1;i_2,j_2)} = (T_K C_{\{\pi\}} - \frac{1}{(r!)^{b-1}} (T_K C_{\{\pi_1\}} + \ldots + T_K C_{\{\pi_b\}}))_{(i_1,j_1;i_2,j_2)}. \tag{1.35}$$

Then from (1.34) and (1.35),

$$(C_{\{\pi_0\}})'(C_{\{\pi_0\}})$$

$$= \sum_{1 \le j_1 < j_2 \le b} \sum_{1 \le i_1, i_2 \le r} \{ \frac{(rb)!}{(r!)^{b-2}(2r)!} \{a[R_{i_1}^{j_1}, R_{i_2}^{j_2}] - \frac{a[\cdot, R_{i_2}^{j_2}] + a[R_{i_1}^{j_1}, \cdot]}{r}\}\}^2$$

$= a$ *constant depending only on* r *and* b.

This completes the proof of 1). For 2), we have, from (1.28), (1.29), (1.31), (1.34) and (1.35),

$$(T_K C_{\{\pi_1\}} + \ldots + T_K C_{\{\pi_b\}})' C_{\{\pi_0\}}$$

$$= \frac{[(rb)!]^2 (r-1)!}{[(2r)!]^2 (r!)^{b-2}} \sum_{1 \le j_1 < j_2 \le b} \sum_{1 \le i_1, i_2 \le r} \{a[R_{i_1}^{j_1}, R_{i_2}^{j_2}]$$

$$- \frac{a[\cdot, R_{i_2}^{j_2}] + a[R_{i_1}^{j_1}, \cdot]}{r}\} \{a[R_{i_1}^{j_1}, \cdot] + a[\cdot, R_{i_2}^{j_2}]\}$$

$$= \frac{[(rb)!]^2 (r-1)!}{[(2r)!]^2 (r!)^{b-2}} \sum_{1 \le j_1 < j_2 \le b} \{ -\frac{2}{r} a^2[\cdot, \cdot] \}$$

$$= 0.$$

For 3), assuming that $p < s$, it follows from (1.28) and (1.29)

$$(T_K C_{\{\pi_p\}})'(T_K C_{\{\pi_s\}})$$

$$= \frac{[(rb)!(r-1)!]^2}{[(2r)!]^2} \sum_{1 \le i_1, i_2 \le r} a[R^p_{i_1}, \cdot] a[\cdot, R^s_{i_2}]$$

$$= \frac{[(rb)!(r-1)!]^2}{[(2r)!]^2} a[\cdot,\cdot] a[\cdot,\cdot]$$

$$= 0.$$

Part 4) is therefore obvious. This proves Lemma 1.3.

Lemma 1.4 *For any permutation $\pi^* \in P_{rb}$, there exist two $\begin{pmatrix} rb \\ 2 \end{pmatrix} \times 1$ vectors $C_{\{\pi_0\}}$ and $C_{\{\pi_0^*\}}$ such that*

1) $(C_{\{\pi_0\}})' C_{\{\pi_0^*\}} = \frac{[(rb)!]^2}{(r!)^{2(b-2)}[(2r)!]^2} \sum_{1 \le j_1 < j_2 \le b} \sum_{1 \le i_1, i_2 \le r} a[R^{j_1}_{i_1}, R^{j_2}_{i_2}] a[i_1, i_2]$,

2) $(T_K C_{\{\pi_1\}} + \cdots + T_K C_{\{\pi_b\}})' C_{\{\pi_0^*\}} = (T_K C_{\{\pi_1^*\}} + \cdots + T_K C_{\{\pi_b^*\}})' C_{\{\pi_0\}} = 0$,

3) $(T_K C_{\{\pi_p\}})'(T_K C_{\{\pi_s^*\}}) = 0$, *for* $1 \le p,\ s \le b$ *and* $p \ne s$,

and 4) $(T_K C_{\{\pi\}})'(T_K C_{\{\pi^*\}})$

$= \frac{1}{(r!)^{2(b-1)}}[(T_K C_{\{\pi_1\}})'(T_K C_{\{\pi_1^*\}}) + \cdots + (T_K C_{\{\pi_b\}})'(T_K C_{\{\pi_b^*\}})] + (C_{\{\pi_0\}})' C_{\{\pi_0^*\}}$.

Proof: Similar to the proof of Lemma 1.3.

We are now in the position to compute the test statistic corresponding to the Kendall's Tau distance. It will be seen that this does not induce the Jonckheere statistic, Jonckheere (1954).

Theorem 1.5

$$d_K(\{\pi\}, E) = \frac{rb(rb-1)[(rb)!]^2}{2(r!)^{2b}} - \frac{[(rb)!]^2}{(r!)^{2(b-2)}[(2r)!]^2} \sum_{1 \le j_1 < j_2 \le b} \sum_{1 \le i_1, i_2 \le r} a[R^{j_1}_{i_1}, R^{j_2}_{i_2}] a[i_1, i_2] - \frac{[(rb)!]^2}{(r!)^{2b}} \sum_{j=1}^{b} \sum_{1 \le i_1 < i_2 \le r} sgn[R^j_{i_2} - R^j_{i_1}]. \quad (1.36)$$

Proof:

$$d_K(\{\pi\}, E) = \sum_{\mu \in \{\pi\}} \sum_{\nu \in E} d_K(\mu, \nu) = C'_E Q_K C_{\{\pi\}}$$

$$= C'_E [\frac{rb(rb-1)}{2} J - T'_K T_K] C_{\{\pi\}}$$

$$= \frac{rb(rb-1)}{2} C'_E J C_{\{\pi\}} - C'_E T'_K T_K C_{\{\pi\}}$$

$$= \frac{rb(rb-1)}{2} [\begin{pmatrix} rb \\ r \end{pmatrix} \cdots \begin{pmatrix} 2r \\ r \end{pmatrix}]^2 - (T_K C_E)'(T_K C_{\{\pi\}})$$

$$= \frac{rb(rb-1)[(rb)!]^2}{2(r!)^{2b}} - (T_K C_E)'(T_K C_{\{\pi\}})$$

$$= \frac{rb(rb-1)[(rb)!]^2}{2(r!)^{2b}} - (C_{\{\pi_0\}})' C_{E_0} - \frac{1}{(r!)^{2(b-1)}} [(T_K C_{\{\pi_1\}})'(T_K C_{E_1}) + \cdots + (T_K C_{\{\pi_b\}})'(T_K C_{E_b})].$$

The theorem follows from (1.28), (1.29), (1.32) and (1.33).

3) Spearman's Footrule

Lemma 1.5 *For the Spearman's Footrule distance,*

$$1)\ T_F C_{\{\pi\}} = \frac{1}{(r!)^{b-1}}(T_F C_{\{\pi_1\}} + \ldots + T_F C_{\{\pi_b\}})$$

$$2)\ (T_F C_{\{\pi_p\}})'(T_F C_{\{\pi_s\}}) = 0,\ \textit{for}\ 1 \le p,\ s \le b, p \ne s.$$

Proof: By definition, for $1 \le j_1, j_2 \le b, 1 \le i_1, i_2 \le r$,

$$\begin{aligned}
&(T_F C_{\{\pi\}})_{(i_1,j_1;i_2,j_2)} \\
&= \sum_{s=1}^{(rb)!}\{I[(\mu_s)_{j_1}(i_1) \le (j_2-1)r + i_2] - \frac{(j_2-1)r+i_2}{rb}\}C_{\{\pi\}}(s) \\
&= \sum_{\mu\in\{\pi\}}\{I[\mu_{j_1}(i_1) \le (j_2-1)r + i_2] - \frac{(j_2-1)r+i_2}{rb}\} \\
&= \sum_{\mu\in\{\pi\}} I[\mu_{j_1}(i_1) \le (j_2-1)r + i_2] - \frac{(j_2-1)r+i_2}{rb}\binom{rb}{r}\cdots\binom{2r}{r} \\
&= \sum_{t=R^{j_1}_{i_1}}^{rb-(r-R^{j_1}_{i_1})}\binom{t-1}{R^{j_1}_{i_1}-1}\binom{rb-t}{r-R^{j_1}_{i_1}}\binom{(b-1)r}{r}\cdots\binom{2r}{r} I[t \le (j_2-1)r+i_2] - \\
&\qquad \frac{(j_2-1)r+i_2}{rb}\binom{rb}{r}\cdots\binom{2r}{r} \\
&= \frac{[r(b-1)]!}{(r!)^{(b-1)}}\sum_{t=R^{j_1}_{i_1}}^{rb-(r-R^{j_1}_{i_1})}\binom{t-1}{R^{j_1}_{i_1}-1}\binom{rb-t}{r-R^{j_1}_{i_1}} I[t \le (j_2-1)r+i_2] - \\
&\qquad [(j_2-1)r+i_2]\frac{(rb-1)!}{(r!)^b}.
\end{aligned} \tag{1.37}$$

The last equality above comes from the fact that $\mu_{j_1}(i_1)$ can take each value t among $R^{j_1}_{i_1}, \ldots, rb-(r-R^{j_1}_{i_1})$ wiht multiplicity $\binom{t-1}{R^{j_1}_{i_1}-1}\binom{rb-t}{r-R^{j_1}_{i_1}}$ while the remaining components in the populations other than the j^{th} block can be chosen in $\binom{(b-1)r}{r}\cdots\binom{2r}{r}$ ways.

Similarly,

$$(T_F C_{\{\pi_j\}})_{(i_1,j_1;i_2,j_2)} = \begin{cases} \sum_{t=R^{j_1}_{i_1}}^{rb-(r-R^{j_1}_{i_1})}\binom{2r}{r}[(b-1)r]!I[t \le (j_2-1)r+i_2] & \\ \qquad - \frac{(j_2-1)r+i_2}{rb}[(b-1)r]!\binom{rb}{r} & \text{if } j = j_1 \\ \sum_{t=1}^{rb}\binom{rb-1}{r}(rb-1-r)!I[t \le (j_2-1)r+i_2] & \\ \qquad - \frac{(j_2-1)r+i_2}{rb}[(b-1)r]!\binom{br}{r} & \text{if } j \ne j_1 \end{cases}$$

$$= \begin{cases} \sum_{t=R_{i_1}^{j_1}}^{rb-(r-R_{i_1}^{j_1})} [(b-1)r]! I[t \le (j_2-1)r+i_2] \\ \qquad -[(j_2-1)r+i_2]\frac{(rb-1)!}{r!} & \text{if } j = j_1 \\ 0 & \text{if } j \ne j_1. \end{cases} \tag{1.38}$$

Therefore by (1.37) and (1.38)

$$\begin{aligned} &(T_F C_{\{\pi\}} - \frac{1}{(r!)^{b-1}}(T_F C_{\{\pi_1\}} + \ldots + T_F C_{\{\pi_b\}}))_{(i_1,j_1;i_2,j_2)} \\ &= -\frac{(j_2-1)r+i_2}{rb}\begin{pmatrix} rb \\ r \end{pmatrix}\cdots\begin{pmatrix} 2r \\ r \end{pmatrix} - \frac{(j_2-1)r+i_2}{rb}\{-\frac{1}{(r)!^{b-1}}[(b-1)r]!\begin{pmatrix} rb \\ r \end{pmatrix}\} \\ &= -\frac{(j_2-1)r+i_2}{rb}\cdot\frac{(rb)!}{(r!)^b} + \frac{(j_2-1)r+i_2}{rb}\cdot\frac{(rb)!}{(r!)^b} \\ &= 0. \end{aligned}$$

Hence 1) is proved and part 2) is obvious from (1.38).

Lemma 1.6 *For any permutation $\pi^* \in P_{rb}$, we have*

1) $(T_F C_{\{\pi_p\}})'(T_F C_{\{\pi_s^*\}}) = 0, \quad$ *for* $1 \le p, s \le b$ *and* $p \ne s$.

2) $(T_F C_{\{\pi\}})'(T_F C_{\{\pi^*\}}) = \frac{1}{(r!)^{2(b-1)}}[(T_F C_{\{\pi_1\}})'(T_F C_{\{\pi_1^*\}}) + \cdots + (T_F C_{\{\pi_b\}})'(T_F C_{\{\pi_b^*\}})]$.

Proof: It follows immediately from Lemma 1.5.

Theorem 1.6

$$\begin{aligned} d_F(\{\pi\}, E) \;=\; & \frac{(rb)^2 - 6rb - 1)[(rb)!]^2}{6(r!)^{2b}} + C(F) + \frac{\{[(b-1)r]!\}^2}{(r!)^{2(b-1)}} \cdot \\ & \sum_{j=1}^{b}\{\sum_{i=1}^{r}\sum_{s=i}^{r(b-1)+i}\sum_{t=R_i^j}^{r(b-1)+R_i^j}\begin{pmatrix} s-1 \\ i-1 \end{pmatrix}\begin{pmatrix} rb-s \\ r-i \end{pmatrix}\begin{pmatrix} t-1 \\ R_i^j-1 \end{pmatrix}\begin{pmatrix} rb-t \\ r-R_i^j \end{pmatrix} max(s,t)\} \end{aligned} \tag{1.39}$$

where $C(F)$ is a constant depending on b, r.

Proof:

$$\begin{aligned} &\sum_{\mu\in\{\pi\}}\sum_{\nu\in E} d_F(\mu,\nu) = C_E' Q_K C_{\{\pi\}} \\ &= C_E'[\frac{(rb+1)(rb-1)}{6}J_F - T_F' T_F]C_{\{\pi\}} \\ &= \frac{(rb+1)(rb-1)}{6}C_E' J_F C_{\{\pi\}} - C_E' T_F' T_F C_{\{\pi\}} \\ &= \frac{(rb+1)(rb-1)}{6}[\begin{pmatrix} rb \\ r \end{pmatrix}\cdots\begin{pmatrix} 2r \\ r \end{pmatrix}]^2 - (T_F C_E)'(T_F C_{\{\pi\}}) \\ &= \frac{(rb+1)(rb-1)[(rb)!]^2}{6(r!)^{2b}} - \\ &\quad \frac{1}{(r!)^{2(b-1)}}[(T_F C_{\{\pi_1\}})'(T_F C_{E_1}) + \cdots + (T_F C_{\{\pi_b\}})'(T_F C_{E_b})]. \end{aligned}$$

The theorem follows from (1.38).

4) Hamming

Theorem 1.7

$$d_H(\{\pi\},E)=\frac{rb[(rb)!]^2}{(r!)^{2b}}-\frac{\{[r(b-1))]!\}^2}{(r!)^{2(b-1)}}$$
$$\sum_{j=1}^{b}\sum_{i=1}^{r}\sum_{t=max(R_i^j,i)}^{(b-1)r+min(R_i^j,i)}\binom{t-1}{R_i^j-1}\binom{rb-t}{r-R_i^j}\binom{t-1}{i-1}\binom{rb-t}{r-i}. \qquad (1.40)$$

Proof:

$$d_H(\{\pi\},E)=\sum_{\mu\in\{\pi\}}\sum_{\nu\in E}d_H(\mu,\nu)$$
$$=\sum_{\mu\in\{\pi\}}\sum_{\nu\in E}\sum_{t=1}^{rb}I[\mu(t)\neq\nu(t)]$$
$$=\sum_{\mu\in\{\pi\}}\sum_{\nu\in E}\sum_{t=1}^{rb}1-\sum_{\mu\in\{\pi\}}\sum_{\nu\in E}\sum_{t=1}^{rb}I[\mu(t)=\nu(t)]$$
$$=rb[\binom{rb}{r}\cdots\binom{2r}{r}]^2-\sum_{\mu\in\{\pi\}}\sum_{\nu\in E}\sum_{t=1}^{rb}I[\mu(t)=\nu(t)]$$
$$=rb[\binom{rb}{r}\cdots\binom{2r}{r}]^2-\sum_{j=1}^{b}\sum_{i=1}^{r}\sum_{\mu\in\{\pi\}}\sum_{\nu\in E}I[\mu_j(i)=\nu_j(i)] \qquad \textit{see notation (1.18)}$$
$$=rb[\binom{rb}{r}\cdots\binom{2r}{r}]^2-\sum_{j=1}^{b}\sum_{i=1}^{r}\sum_{t=R_i^j}^{rb-(r-R_i^j)}\binom{t-1}{R_i^j-1}\binom{rb-t}{r-R_i^j}\binom{(b-1)r}{r}\cdots\binom{2r}{r}\cdot$$
$$\sum_{\nu\in E}I[t=\nu((j-1)r+i)]$$
$$=rb[\binom{rb}{r}\cdots\binom{2r}{r}]^2-\sum_{j=1}^{b}\sum_{i=1}^{r}\sum_{t=R_i^j}^{rb-(r-R_i^j)}\binom{t-1}{R_i^j-1}\binom{rb-t}{r-R_i^j}\binom{(b-1)r}{r}\cdots\binom{2r}{r}\cdot$$
$$\sum_{s=i}^{rb-(r-i)}\binom{s-1}{i-1}\binom{rb-s}{r-i}\binom{(b-1)r}{r}\cdots\binom{2r}{r}I[t=s]$$
$$=rb[\binom{rb}{r}\cdots\binom{2r}{r}]^2-[\binom{(b-1)r}{r}\cdots\binom{2r}{r}]^2\cdot$$
$$\sum_{j=1}^{b}\sum_{i=1}^{r}\sum_{t=R_i^j}^{rb-(r-R_i^j)}\sum_{s=i}^{rb-(r-i)}\binom{t-1}{R_i^j-1}\binom{rb-t}{r-R_i^j}\binom{s-1}{i-1}\binom{rb-s}{r-i}I[t=s]$$
$$=rb[\binom{rb}{r}\cdots\binom{2r}{r}]^2-[\binom{(b-1)r}{r}\cdots\binom{2r}{r}]^2\cdot$$

$$\sum_{j=1}^{b}\sum_{i=1}^{r}\sum_{t=max(R_i^j,i)}^{min[rb-(r-R_i^j),rb-(r-i)]}\binom{t-1}{R_i^j-1}\binom{rb-t}{r-R_i^j}\binom{t-1}{i-1}\binom{rb-t}{r-i}$$

$$=\frac{rb[(rb)!]^2}{(r!)^{2b}}-\frac{\{[r(b-1))]!\}^2}{(r!)^{2(b-1)}}\sum_{j=1}^{b}\sum_{i=1}^{r}\sum_{t=max(R_i^j,i)}^{(b-1)r+min(R_i^j,i)}\binom{t-1}{R_i^j-1}\binom{rb-t}{r-R_i^j}\binom{t-1}{i-1}\binom{rb-t}{r-i}.$$

We conclude based on the above results that the test statistics corresponding to Spearman's Rho, Spearman's Footrule and Hamming are all sums of independent random variables. Hence their asymptotic normalities follow from the central limit theorem for the sum of independent random variables. Furthermore, the statistics are functions only of the relative ranks within each block.

As for the Kendall's Tau statistic, although it is not the sum of independent random variables, it will be shown in section 2.3 of Chapter 2 that it is asymptotically equivalent to a nondegenerated U-statistic as $b \to \infty$. Therefore its asymptotic normality is ensured from the property of U-statistics.

In view of these properties, it is sufficient to compute the means and variances as are done in section 2.3 of Chapter 2.

1.6 The Dispersion Case: the multi-sample case

Consider the following testing problem of

$$\begin{aligned} &H_0: F_1(x) = \cdots = F_r(x) = F[x \cdot exp(-\bar{d})] \ against \\ &H_1: F_k(x) = F[x \cdot exp(-d_k)], 1 \le k \le r \ and \ 0 < d_1 < \cdots < d_r < 1. \end{aligned} \tag{1.41}$$

This is referred to the problem of testing dispersion in the multi-sample. In addition to the notation introduced in section 1.4 of Chapter 1, assume further that $n_k = 2m_k$ and $M_k = m_1 + \cdots m_k$ and $M_0 = 0$, for $1 \le k \le r$. The two equivalent subclasses of P_{N_r} are as follows:

$$\begin{aligned} \{\pi\} = \ &\{[\pi(i_1^1), \cdots, \pi(i_{m_1}^1)| \cdots |\pi(M_{r-2} + i_1^{r-1}), \cdots, \pi(M_{r-2} + i_{m_{r-1}}^{r-1})| \\ &\pi(M_{r-1} + i_1^r), \cdots, \pi(M_{r-1} + i_{n_r}^r)|\pi(N_r - M_{r-1} + i_{m_{r-1}+1}^{r-1}), \cdots, \pi(N_r - M_{r-1} + i_{n_{r-1}}^{r-1})| \\ &\cdots |\pi(N_r - M_1 + i_{m_1+1}^1), \cdots, \pi(N_r - M_1 + i_{n_1}^1)] : \\ &[i_1^k, \cdots, i_{m_k}^k, i_{m_k+1}^k, \cdots, i_{n_k}^k] \in P_{n_k}, \ for \ 1 \le k \le r\} \end{aligned} \tag{1.42}$$

$$\begin{aligned} E = \ &\{[i_1^1, \cdots, i_{m_1}^1| \cdots |M_{r-2} + i_1^{r-1}, \cdots, M_{r-2} + i_{m_{r-1}}^{r-1}| \\ &M_{r-1} + i_1^r, \cdots, M_{r-1} + i_{n_r}^r|N_r - M_{r-1} + i_{m_{r-1}+1}^{r-1}, \cdots, N_r - M_{r-1} + i_{n_{r-1}}^{r-1}| \\ &\cdots |N_r - M_1 + i_{m_1+1}^1, \cdots, N_r - M_1 + i_{n_1}^1] : \\ &[i_1^k, \cdots, i_{m_k}^k, i_{m_k+1}^k, \cdots, i_{n_k}^k] \in P_{n_k}, \ for \ 1 \le k \le r\}. \end{aligned} \tag{1.43}$$

In (1.42) and (1.43) the items labelled $M_{k-1}+1, \cdots, M_k$ and $N_r - M_k + 1, \cdots, N_r - M_{k-1}$ refer to the k^{th} population. With this notation, it can be seen that in addition to the usual properties of $\{\pi\}$ and E sets, there is an invariance with respect to the medians.

We are now in a position to obtain the test statistics for each of the distance functions in section 1.2.

1) Spearman's Rho

$$
\begin{aligned}
d_{S_d}(\{\pi\},E) &= \sum_{\mu\in\{\pi\}}\sum_{\nu\in E} d_{S_d}(\mu,\nu) = \sum_{\mu\in\{\pi\}}\sum_{\nu\in E}\sum_{i=1}^{N_r}[\mu(i)-\nu(i)]^2 \\
&= \sum_{\mu\in\{\pi\}}\sum_{\nu\in E}\sum_{i=1}^{N_r} 2i^2 - 2\sum_{k=1}^{r}\Big(\sum_{i=M_{k-1}+1}^{M_k} + \sum_{i=N_r-M_k+1}^{N_r-M_{k-1}}\Big)\Big[\sum_{\mu\in\{\pi\}}\mu(i)\Big]\Big[\sum_{\nu\in E}\nu(i)\Big] \\
&= \frac{\prod_{k=1}^{r}(n_k!)^2 N_r(N_r+1)(2N_r+1)}{3} - 2\sum_{k=1}^{r}\Big(\sum_{i=M_{k-1}+1}^{M_k} + \sum_{i=N_r-M_k+1}^{N_r-M_{k-1}}\Big) \\
&\qquad \Big[\Big(\sum_{j=M_{k-1}+1}^{M_k} + \sum_{j=N_r-M_k+1}^{N_r-M_{k-1}}\Big)\frac{\prod_{k=1}^{r} n_k!}{n_k}\pi(j)\Big]\Big[\Big(\sum_{j=M_{k-1}+1}^{M_k} + \sum_{j=N_r-M_k+1}^{N_r-M_{k-1}}\Big)\frac{\prod_{k=1}^{r} n_k!}{n_k} j\Big] \\
&= \frac{\prod_{k=1}^{r}(n_k!)^2 N_r(N_r+1)(2N_r+1)}{3} - \prod_{k=1}^{r}(n_k!)^2\frac{N_r(N_r+1)^2}{2} \\
&= \frac{\prod_{k=1}^{r}(n_k!)^2 N_r(N_r+1)(N_r-1)}{6},
\end{aligned}
\qquad (1.44)
$$

which degenerates to a constant. Consequently, the use of the Spearman's Rho distance in this case does not lead to a test statistic.

2) Kendall's Tau

$$
\begin{aligned}
d_{K_d}(\{\pi\},E) &= \sum_{\mu\in\{\pi\}}\sum_{\nu\in E} d_{K_d}(\mu,\nu) = \sum_{\mu\in\{\pi\}}\sum_{\nu\in E}\sum_{1\le i<j\le N_r}\{1 - sgn[\mu(i)-\mu(j)]sgn[\nu(i)-\nu(j)]\} \\
&= \frac{\prod_{k=1}^{r}(n_k!)^2 N_r(N_r-1)}{2} - \sum_{1\le i<j\le N_k}\Big\{\sum_{\mu\in\{\pi\}} sgn[\mu(i)-\mu(j)]\Big\}\Big\{\sum_{\nu\in E} sgn[\nu(i)-\nu(j)]\Big\} \\
&= \frac{\prod_{k=1}^{r}(n_k!)^2 N_r(N_r-1)}{2} - \Big(\sum_{i<j \text{ in the same population}} + \sum_{i<j \text{ not in the same population}}\Big)\cdot \\
&\qquad \Big\{\sum_{\mu\in\{\pi\}} sgn[\mu(i)-\mu(j)]\Big\}\Big\{\sum_{\nu\in E} sgn[\nu(i)-\nu(j)]\Big\} \\
&= \frac{\prod_{k=1}^{r}(n_k!)^2 N_r(N_r-1)}{2} - \Big(0 + \sum_{i<j \text{ not in the same population}}\Big)\cdot \\
&\qquad \Big\{\sum_{\mu\in\{\pi\}} sgn[\mu(i)-\mu(j)]\Big\}\Big\{\sum_{\nu\in E} sgn[\nu(i)-\nu(j)]\Big\} \\
&= \frac{\prod_{k=1}^{r}(n_k!)^2 N_r(N_r-1)}{2} - \\
&\qquad \sum_{k=1}^{r}\Big\{\sum_{i=M_{k-1}+1}^{M_k}\Big(\sum_{j=M_k+1}^{N_r-M_k} + \sum_{i=N_r-M_{k-1}+1}^{N_r}\Big)\Big\{\sum_{\mu\in\{\pi\}} sgn[\mu(i)-\mu(j)]\Big\}\Big\{\sum_{\nu\in E} sgn[\nu(i)-\nu(j)]\Big\} + \\
&\qquad \sum_{i=N_r-M_k+1}^{N_r-M_{k-1}}\ \sum_{j=N_r-M_{k-1}+1}^{N_r}\Big\{\sum_{\mu\in\{\pi\}} sgn[\mu(i)-\mu(j)]\Big\}\Big\{\sum_{\nu\in E} sgn[\nu(i)-\nu(j)]\Big\}\Big\}
\end{aligned}
$$

$$= \frac{\prod_{k=1}^r (n_k!)^2 N_r(N_r-1)}{2} -$$

$$\sum_{k=1}^{r} \sum_{i=M_{k-1}+1}^{M_k} \sum_{j=M_k+1}^{N_r-M_k} \{ \sum_{\mu\in\{\pi\}} sgn[\mu(i)-\mu(j)]\}\{ \sum_{\nu\in\{\pi\}} sgn[\nu(i)-\nu(j)]\} -$$

$$\sum_{k=1}^{r} \sum_{i=M_{k-1}+1}^{M_k} \sum_{i=N_r-M_{k-1}+1}^{N_r} \{ \sum_{\mu\in\{\pi\}} sgn[\mu(i)-\mu(j)]\}\{ \sum_{\mu\in\{\pi\}} sgn[\nu(i)-\nu(j)]\} -$$

$$\sum_{i=N_r-M_k+1}^{N_r-M_{k-1}} \sum_{j=N_r-M_{k-1}+1}^{N_r} \{ \sum_{\mu\in\{\pi\}} sgn[\mu(i)-\mu(j)]\}\{ \sum_{\nu\in E} sgn[\nu(i)-\nu(j)]\}$$

$$= \frac{\prod_{k=1}^r (n_k!)^2 N_r(N_r-1)}{2} - \sum_{k=1}^{r} \sum_{i=M_{k-1}+1}^{M_k} \sum_{j=M_k+1}^{N_r-M_k} \{ \sum_{\mu\in\{\pi\}} sgn[\mu(i)-\mu(j)]\}\{ \sum_{\nu\in\{\pi\}} (-1)\} -$$

$$\sum_{k=1}^{r} \sum_{i=M_{k-1}+1}^{M_k} \sum_{i=N_r-M_{k-1}+1}^{N_r} \{ \sum_{\mu\in\{\pi\}} sgn[\mu(i)-\mu(j)]\}\{0\} -$$

$$\sum_{i=N_r-M_k+1}^{N_r-M_{k-1}} \sum_{j=N_r-M_{k-1}+1}^{N_r} \{ \sum_{\mu\in\{\pi\}} sgn[\mu(i)-\mu(j)]\}\{0\}$$

$$= \frac{\prod_{k=1}^r (n_k!)^2 N_r(N_r-1)}{2} +$$

$$\sum_{k=1}^{r} \sum_{i=M_{k-1}+1}^{M_k} \sum_{t=k}^{r-1} (\sum_{j=M_t+1}^{M_{t+1}} + \sum_{j=N_r-M_{t+1}+1}^{N_r-M_t})\{ \sum_{\mu\in\{\pi\}} sgn[\mu(i)-\mu(j)]\}\{\prod_{k=1}^{r} n_k!\}$$

$$= \frac{\prod_{k=1}^r (n_k!)^2 N_r(N_r-1)}{2} + \sum_{k=1}^{r} \sum_{i=M_{k-1}+1}^{M_k} \sum_{t=k}^{r-1} (\sum_{j=M_t+1}^{M_{t+1}} + \sum_{j=N_r-M_{t+1}+1}^{N_r-M_t}) \cdot$$

$$(\sum_{i_1=M_{k-1}+1}^{M_k} + \sum_{i_1=N_r-M_k+1}^{N_r-M_{k-1}})(\sum_{j_1=M_t+1}^{M_{t+1}} + \sum_{j_1=N_r-M_{t+1}+1}^{N_r-M_t})\{ \frac{\prod_{k=1}^r n_k!}{n_k n_t} sgn[\pi(i_1)-\pi(j_1)]\}\{\prod_{k=1}^{r} n_k!\}$$

$$= \frac{\prod_{k=1}^r (n_k!)^2 N_r(N_r-1)}{2} + \prod_{k=1}^{r} (n_k!)^2/2 \sum_{k=1}^{r} \sum_{t=k}^{r-1} \cdot$$

$$(\sum_{i=M_{k-1}+1}^{M_k} + \sum_{i=N_r-M_k+1}^{N_r-M_{k-1}})(\sum_{j=M_t+1}^{M_{t+1}} + \sum_{j=N_r-M_{t+1}+1}^{N_r-M_t}) sgn[\pi(i)-\pi(j)]. \tag{1.45}$$

This new statistic rejects H_0 whenever the sum of all terms $sgn[\pi(i)-\pi(j)]$ is small.

3) Spearman's Footrule

$$d_{F_d}(\{\pi\}, E) = \sum_{\mu\in\{\pi\}} \sum_{\nu\in E} d_{F_d}(\mu,\nu) = \sum_{\mu\in\{\pi\}} \sum_{\nu\in E} \sum_{i=1}^{N_r} |\mu(i)-\nu(i)|$$

$$= \sum_{k=1}^{r} (\sum_{i=M_{k-1}+1}^{M_k} + \sum_{i=N_r-M_k+1}^{N_r-M_{k-1}}) \sum_{\mu\in\{\pi\}} \sum_{\nu\in E} |\mu(i)-\nu(i)|$$

$$= \sum_{k=1}^{r}\left(\sum_{i=M_{k-1}+1}^{M_k} + \sum_{i=N_r-M_k+1}^{N_r-M_{k-1}}\right)\frac{\prod_{k=1}^{r}(n_k!)^2}{n_k^2}\cdot$$

$$\left(\sum_{j_1=M_{k-1}+1}^{M_k} + \sum_{j_1=N_r-M_k+1}^{N_r-M_{k-1}}\right)\left(\sum_{j_2=M_{k-1}+1}^{M_k} + \sum_{j_2=N_r-M_k+1}^{N_r-M_{k-1}}\right)|\pi(j_1)-j_2|$$

$$= \prod_{k=1}^{r}(n_k!)^2\sum_{k=1}^{r}\left(\sum_{j_1=M_{k-1}+1}^{M_k} + \sum_{j_1=N_r-M_k+1}^{N_r-M_{k-1}}\right)\left(\sum_{j_2=M_{k-1}+1}^{M_k} + \sum_{j_2=N_r-M_k+1}^{N_r-M_{k-1}}\right)\frac{[\pi(j_1)-j_2]I_{[\pi(j_1)>j_2]}}{n_k}. \tag{1.46}$$

This new statistic rejects H_0 whenever the weighted sum of all terms $[\pi(j_1)-j_2]I_{[\pi(j_1)>j_2]}$ is small.

4) Hamming

For $1 \le k \le r$, let

$$A'_k = \{M_{k-1}+1, \cdots, M_k, N_r - M_k + 1, \cdots, N_r - M_{k-1}\}$$
$$B'_k = \{\pi(M_{k-1}+1), \cdots, \pi(M_k), \pi(N_r - M_k + 1), \cdots, \pi(N_r - M_{k-1})\}$$
$$C'_k = A'_k \bigcap B'_k,\ C'_0 = (C'_1 \bigcup \cdots \bigcup C'_r)^c$$
Y'_k=the cardinality of C'_k, Y'_0=the cardinality of C'_0.

It is easy to see that $Y'_0 = N_r - (Y'_1 + \cdots + Y'_r)$. As in the two-sample case, we obtain

$$d_{H_a}(\{\pi\}, E) = \sum_{\mu\in\{\pi\}}\sum_{\nu\in E} d_{H_a}(\mu,\nu) = \sum_{\mu\in\{\pi\}}\sum_{\nu\in E}\sum_{i=1}^{N_r} I[\mu(i)\neq\nu(i)]$$
$$= \sum_{\mu\in\{\pi\}}\{\sum_{i=1}^{N_r}\sum_{\nu\in E} I[\mu(i)\neq\nu(i)]\}$$
$$= \sum_{\mu\in\{\pi\}}\{\sum_{k=0}^{r}\sum_{i\in C'_k}\sum_{\nu\in E} I[\mu(i)\neq\nu(i)]\}$$
$$= \sum_{\mu\in\{\pi\}}\{\sum_{k=1}^{r}\sum_{i\in C'_k}\frac{n_1!\ldots n_r!(n_k-1)}{n_k} + \sum_{i\in C'_0} n_1!\ldots n_r!\}$$
$$= \sum_{\mu\in\{\pi\}}\{\sum_{k=1}^{r} Y'_k\frac{n_1!\ldots n_r!(n_k-1)}{n_k} + Y'_0 n_1!\ldots n_r!\}$$
$$= \sum_{\mu\in\{\pi\}}\{-\sum_{k=1}^{r} Y'_k\frac{n_1!\ldots n_r!}{n_k} + N_r n_1!\ldots n_r!\}$$
$$= \prod_{k=1}^{r}(n_k!)^2(N_r - \sum_{k=1}^{r}\frac{Y'_k}{n_k}). \tag{1.47}$$

This new statistic rejects H_0 whenever the sum $\frac{Y'_1}{n_1} + \cdots + \frac{Y'_r}{n_r}$ is large.

Chapter 2

Asymptotic Distributions Under the Null Hypotheses

In Chapter 1, we have obtained various test statistics for different testing problems by using the general approach (1.1). In this chapter we find the limiting distributions of these test statistics under the null hypotheses. The limiting distribution of any statistic X_n is said to be asymptotically normal with mean μ_n and variance σ_n^2 whenever

$$\frac{X_n - \mu_n}{\sigma_n} \xrightarrow{d} N(0,1), \text{ as } n \to \infty.$$

In this case, we write $X_n \sim N(\mu_n, \sigma_n^2)$, *as* $n \to \infty$. In the following discussions we compute mostly the asymptotic means and variances although in some cases the computations may be exact. The following two theorems will be useful. First let $\varphi(u), 0 < u < 1$ be a real function satisfying

$$\int_0^1 [\varphi(u) - \bar{\varphi}]^2 du > 0, \tag{2.1}$$

with notation $\bar{\varphi} = \int_0^1 \varphi(u)du$. In this case $\varphi(u)$ is called a square integrable funtion.

Theorem 2.1 *(Hoeffding's central limit theorem (1951)) Let $\{c_n(i,j)\}$ be n^2 real numbers and set $d_n(i,j) = c_n(i,j) - \bar{c}_n(i,\cdot) - \bar{c}_n(\cdot,j) + \bar{c}_n(\cdot,\cdot)$. The distribution of*

$$S_n = \sum_{i=1}^n c_n(i, \pi(i))$$

is asymptotically normal with

$$mean = (1/n)\sum_{i,j=1}^n c_n(i,j)$$
$$and \quad variance = [1/(n-1)]\sum_{i,j=1}^n d_n^2(i,j);$$

if

$$\lim_{n\to\infty} \frac{max_{1 \le i,j \le n} d_n^2(i,j)}{\frac{1}{n}\sum_{i,j=1}^n d_n^2(i,j)} = 0. \tag{2.2}$$

This theorem will be used to prove the asymptotic normality of general rank statistics. For the special case of a linear rank statistic we have the following results due to Hajek and Sidak.

Theorem 2.2 *(V 1.6 theorem a, Hajek-Sidak (1967)) Let $c_n(i,\pi(i)) = c_n(i)a_n(\pi(i))$. Define $\bar{c_n} = \sum_{i=1}^{n} c_n(i)/n, \bar{a_n} = \sum_{i=1}^{n} a_n(i)/n$. The distribution of*

$$S_n = \sum_{i=1}^{n} c_n(i)a_n(\pi(i))$$

is asymptotically normal with

$$mean = n\,\bar{c_n}\bar{a_n}$$
$$and\ \ variance = \sum_{i=1}^{n}[c_n(i) - \bar{c_n}]^2 \int_0^1 [\varphi(u) - \bar{\varphi}]^2 du;$$

if

$$\lim_{n\to\infty} \frac{\sum_{i=1}^{n}[c_n(i) - \bar{c_n}]^2}{max_{1\leq i\leq n}[c_n(i) - \bar{c_n}]^2} = \infty \tag{2.3}$$

$$and \quad \lim_{n\to\infty} \int_0^1 \{a_n(1+[un]) - \varphi(u)\}^2 du = 0. \tag{2.4}$$

where $\varphi(u)$ satisfies (2.1). In this case, $a_n(i)$ is called a score and $\varphi(u)$ is called a square integrable function corresponding to $a_n(i)$. The expression $[un]$ in (2.4) means the greatest integer less than or equal to un.

2.1 The Two-sample Case

Recall the null hypothesis H_0 in (1.2).

1) Spearman's Rho

Let

$$S_2 = \sum_{i=n_1+1}^{n_1+n_2} \frac{\pi(i)}{n_1+n_2+1}. \tag{2.5}$$

Then, by (1.6),

$$d_{S_2}(\{\pi\}, E) \equiv S_2 = \sum_{i=n_1+1}^{n_1+n_2} \frac{\pi(i)}{n_1+n_2+1}. \tag{2.6}$$

here " $\equiv$ " means equivalent in the sense that the terms in both sides differ only by some constants depending on sample size. Therefore instead of dealing with $d_{S_2}(\{\pi\}, E)$, we discuss S_2 directly. Hence the critical region for the Spearman's Rho statistic consists of rejecting H_0 whenever $S_2 > c$, where c is a constant determined by the significance level α and the limiting null distribution of S_2.

Let

$$\varphi(u) = \begin{cases} u & \text{if } 0 \leq u \leq 1 \\ 0 & \text{otherwise.} \end{cases} \tag{2.7}$$

Let the score function $a(i) = \frac{i}{n_1+n_2+1}$. Then we can write $S_2 = \sum_{i=n_1+1}^{n_1+n_2} a(\pi(i))$ where $\varphi(u)$ is the square integrable function corresponding to the score.

By Theorem 2.2,

$$S_2 \overset{H_0}{\sim} N(\mu_{S_2}, \sigma^2_{S_2}), \text{ as } min(n_1, n_2) \to \infty$$

$$\text{with} \quad \mu_{S_2} = \frac{n_2}{n_1+n_2} \sum_{i=1}^{n_1+n_2} \frac{i}{n_1+n_2+1} = \frac{n_2}{2}$$

$$\sigma^2_{S_2} = \frac{n_1 n_2}{n_1+n_2} \int_0^1 [\varphi(u) - \bar{\varphi}]^2 du = \frac{n_1 n_2}{12(n_1+n_2)}. \tag{2.8}$$

2) Kendall's Tau

It follows from section 1.3 of Chapter 1 that

$$d_{K_2}(\{\pi\}, E) \equiv d_{S_2}(\{\pi\}, E) \equiv S_2. \tag{2.9}$$

Hence the asymptotic distribution of Kendall's Tau statistic under H_0 is given by (2.8).

3) Spearman's Footrule

Let

$$F_2 = \sum_{i=n_1+1}^{n_1+n_2} a(\pi(i)) \tag{2.10}$$

where

$$a(i) = \begin{cases} \frac{i(i-1)}{(n_1+n_2)^2} & \text{if } 1 \le i \le n_1 \\ -\frac{n_1}{n_2} \cdot \frac{i^2}{(n_1+n_2)^2} + \frac{n_1[2(n_1+n_2)+1]}{n_2(n_1+n_2)} \cdot \frac{i}{n_1+n_2} - \frac{n_1(n_1+1)}{n_2(n_1+n_2)} & \text{if } n_1 < i \le n_1+n_2. \end{cases} \tag{2.11}$$

Then, by (1.8),

$$d_{F_2}(\{\pi\}, E) \equiv F_2 = \sum_{i=n_1+1}^{n_1+n_2} a(\pi(i)) \tag{2.12}$$

Hence the critical region for the Spearman's Footrule statistic is $\{F_2 > c\}$, where c is a constant dependent on the significance level α and the limiting null distribution of F_2. The square integrable function corresponding to the score (2.11) is found to be

$$\varphi(u) = \begin{cases} u^2 & \text{if } 0 < u \le \lambda \\ \lambda - \frac{\lambda}{1-\lambda}(u-1)^2 & \text{if } \lambda < u < 1 \end{cases} \tag{2.13}$$

provided that $\frac{n_1}{n_1+n_2} \to \lambda$ *as* $min\{n_1, n_2\} \to \infty$.

By Theorem 2.2,

$$F_2 \overset{H_0}{\sim} N(\mu_{F_2}, \sigma^2_{F_2}), \text{ as } min(n_1, n_2) \to \infty$$

$$\text{with} \quad \mu_{F_2} = \frac{n_2}{n_1+n_2} \sum_{i=1}^{n_1+n_2} a(i) = \frac{n_1 n_2 (n_1 + 2n_2)}{3(n_1+n_2)^2}$$

$$\sigma^2_{F_2} = \frac{n_1 n_2}{n_1+n_2} \cdot \frac{4\lambda^2(-2\lambda^2 + 2\lambda + 1)}{45}. \tag{2.14}$$

4) Hammimg

Let

$$H_2 = \sum_{i=1}^{n_1+n_2} a_{i\pi(i)} \tag{2.15}$$

where

$$a_{ij} = \begin{cases} \frac{1}{n_1} & \text{if } i,j \in \{1\ldots n_1\} \\ \frac{1}{n_2} & \text{if } i,j \in \{n_1+1\ldots n_1+n_2\} \\ 0 & \text{otherwise.} \end{cases} \tag{2.16}$$

Then, by (1.9),

$$d_{H_2}(\{\pi\}, E) \equiv H_2 = \sum_{i=1}^{n_1+n_2} a_{i\pi(i)}. \tag{2.17}$$

Hence the critical region for the Hamming statistic is $\{H_2 > c\}$, where c is a constant dependent on the significance level α and the limiting null distribution of H_2. Since H_2 is not a linear statistic, Theorem 2.2 can not be employed.

In order to apply Hoeffding's central limit theorem for statistic H_2, note that

$$\begin{cases} \bar{a}_{i.} = \frac{1}{n_1+n_2}\sum_{j=1}^{n_1+n_2} a_{ij} = \frac{1}{n_1+n_2} \\ \bar{a}_{.j} = \frac{1}{n_1+n_2}\sum_{i=1}^{n_1+n_2} a_{ij} = \frac{1}{n_1+n_2} \\ \bar{a}_{..} = \frac{1}{(n_1+n_2)^2}\sum_{i,j=1}^{n_1+n_2} a_{ij} = \frac{1}{n_1+n_2} \end{cases} \tag{2.18}$$

Then

$$\begin{aligned} d_n(i,j) &= a_{ij} - \bar{a}_{i.} - \bar{a}_{.j} + \bar{a}_{..} \\ &= a_{ij} - \frac{1}{n_1+n_2}. \end{aligned} \tag{2.19}$$

Therefore

$$max|d_n(i,j)| = max\{\frac{1}{n_1+n_2}, \frac{n_2}{n_1(n_1+n_2)}, \frac{n_1}{n_2(n_1+n_2)}\} = \frac{1}{n_1+n_2} max\{\frac{n_2}{n_1}, \frac{n_1}{n_2}\} \tag{2.20}$$

$$\frac{1}{n_1+n_2}\sum_{i,j=1}^{n_1+n_2} d_n^2(i,j) = \frac{1}{n_1+n_2}\sum_{i,j=1}^{n_1+n_2}\{a_{ij}^2 - 2a_{ij}\frac{1}{n_1+n_2} + \frac{1}{(n_1+n_2)^2}\} = \frac{1}{n_1+n_2}. \tag{2.21}$$

Hence by (2.20) and (2.21) when $min(n_1, n_2) \to \infty$,

$$\frac{max_{1\le i,j\le n_1+n_2} d_n^2(i,j)}{\frac{1}{n_1+n_2}\sum_{i,j=1}^{n_1+n_2} d_n^2(i,j)} = \frac{1}{n_1+n_2}(max\{1, \frac{n_2}{n_1}, \frac{n_1}{n_2}\})^2 \to 0. \tag{2.22}$$

By Theorem 2.1,

$$\begin{aligned} & H_2 \overset{H_0}{\sim} N(\mu_{H_2}, \sigma_{H_2}^2), \ as\ min(n_1,n_2) \to \infty \\ with \quad & \mu_{H_2} = 1 \\ & \sigma_{H_2}^2 = \frac{1}{n_1+n_2-1}. \end{aligned} \tag{2.23}$$

2.2 The Multi-sample Case

Recall the notation in section 1.4 of Chapter 1 and assume further that

$$\frac{n_k}{N_r} \rightarrow t_k \text{ as } min\{n_1 \cdots n_r\} \rightarrow \infty. \tag{2.24}$$

Write $T_k = t_1 + \cdots + t_k$ *for* $1 \leq k \leq r$ *and* $T_0 = 0$. Then we have

$$T_k = \lim_{min\{n_1 \cdots n_r\} \rightarrow \infty} \frac{N_k}{N_r}. \tag{2.25}$$

1) Spearman's Rho

Let

$$S_r = \sum_{i=1}^{N_r} c(i) \frac{\pi(i)}{N_r + 1} \tag{2.26}$$

where

$$c(i) = N_{k-1} + N_k \quad if\ N_{k-1} < i \leq N_k,\ for\ 1 \leq k \leq r. \tag{2.27}$$

Then, by (1.13),

$$d_{S_r}(\{\pi\}, E) \equiv S_r = \sum_{i=1}^{N_r} c(i) \frac{\pi(i)}{N_r + 1}. \tag{2.28}$$

Hence the critical region for the Spearman statistic has the form $\{S_r > c\}$, where c is a constant determined by the significance level α and the limiting null distribution of S_r. The score function in this case is $a(i) = \frac{i}{N_r+1}$ and it is easy to see that the square integrable function corresponding to the above score is given by

$$\varphi(u) = \begin{cases} u & \text{if } 0 \leq u \leq 1 \\ 0 & \text{otherwise} \end{cases} \tag{2.29}$$

Write $\bar{c} = \frac{1}{N_r} \sum_{i=1}^{N_r} c(i)$. Let the notation "$A(n) \approx B(n)$" mean that $A(n)$ is approxmiately equal to $B(n)$ as $n \rightarrow \infty$, where $A(n), B(n)$ are some real numbers depending on n. Then when $min\{n_1 \cdots n_r\} \rightarrow \infty$,

$$\begin{aligned} \bar{c} &= \frac{1}{N_r} \sum_{i=1}^{N_r} c(i) \\ &= \sum_{k=1}^{r} \frac{n_k}{N_r}(N_{k-1} + N_k) = N_r \end{aligned} \tag{2.30}$$

$$\begin{aligned} max_{1 \leq i \leq N_r}[c(i) - \bar{c}]^2 &= max_{1 \leq k \leq r}(N_{k-1} + N_k - \bar{c})^2 \\ &= N_r^2 max_{1 \leq k \leq r}(\frac{N_{k-1} + N_k}{N_r} - 1)^2 \\ &\approx N_r^2 max_{1 \leq k \leq r}(T_{k-1} + T_k - 1)^2 \end{aligned} \tag{2.31}$$

$$\sum_{i=1}^{N_r}[c(i) - \bar{c}]^2 = \sum_{k=1}^{r} n_k(N_{k-1} + N_k - \bar{c})^2$$

$$\approx \quad N_r^3 \sum_{k=1}^{r} t_k(T_{k-1} + T_k - 1)^2$$
$$= \quad N_r^3 \sum_{k=1}^{r} t_k T_k T_{k-1}. \tag{2.32}$$

Hence when $min\{n_1 \cdots n_r\} \to \infty$

$$\frac{\sum_{i=1}^{N_r}[c(i) - \bar{c}]^2}{max_{1\leq i\leq N_r}[c(i) - \bar{c}]^2} \tag{2.33}$$
$$= \frac{\sum_{k=1}^{r} t_k T_k T_{k-1}}{max_{1\leq k\leq r}(T_{k-1} + T_k - 1)^2} N_r \to \infty. \tag{2.34}$$

By Theorem 2.2,

$$S_r \overset{H_0}{\sim} N(\mu_{S_r}, \sigma^2_{S_r}), \ as\ min(n_1 \cdots n_r) \to \infty$$
$$with \qquad \mu_{S_r} = \bar{c} \sum_{i=1}^{N_r} \frac{\pi(i)}{N_r + 1} = N_r/2$$
$$\sigma^2_{S_r} = \sum_{i=1}^{N_r}[c(i) - \bar{c}]^2 \int_0^1 [\varphi(u) - \bar{\varphi}]^2 du$$
$$\approx \frac{N_r^3}{12} \sum_{k=1}^{r} t_k T_k T_{k-1}. \tag{2.35}$$

2) Kendall's Tau

It will be shown in section 2.4 that when $min(n_1 \cdots n_r) \to \infty$,

$$Var_{H_0}\{\frac{d_{K_r}(\{\pi\}, E) - E_{H_0}\{d_{K_r}(\{\pi\}, E)\}}{[Var_{H_0}\{d_{K_r}(\{\pi\}, E)\}]^{1/2}} - \frac{d_{S_r}(\{\pi\}, E) - E_{H_0}\{d_{S_r}(\{\pi\}, E)\}}{[Var_{H_0}\{d_{S_r}(\{\pi\}, E)\}]1/2}\}$$
$$\to 0. \tag{2.36}$$

Therefore the Kendall's Tau statistic and the Spearman's Rho statistic are asymptotically equivalent in terms of variance. Thus in the multi-sample case, the asymptotic distribution of the Kendall's Tau statistic under H_0 is given by (2.35).

3) Spearman's Footrule

Let

$$F_r = \sum_{i=1}^{N_r} c_{i\pi(i)} \tag{2.37}$$

where

$$c_{ij} = \sum_{s=1}^{N_r} a_{is}(j - s)I[j > s] \tag{2.38}$$

$$and \qquad a_{is} = \begin{cases} \frac{1}{n_k} & \text{if } i, s \in \{N_{k-1}+1, \cdots, N_k\} \\ 0 & \text{otherwise.} \end{cases} \tag{2.39}$$

Then, by (1.15),

$$d_{F_r}(\{\pi\}, E) \equiv F_r = \sum_{i=1}^{N_r} c_{i\pi(i)}. \tag{2.40}$$

Hence the critical region for the Spearman's Footrule statistic is $\{F_r < c\}$, where c is a constant determined by the significance level α and the limiting null distribution of F_r.

For $N_{p-1} < i \leq N_p, 1 \leq p \leq r$,

$$\begin{aligned} c_{ij} &= \sum_{s=1}^{N_r} a_{is}(j-s)I[j>s] \\ &= \sum_{s=N_{p-1}+1}^{N_p} \frac{1}{n_p}(j-s)I[j>s] \\ &= \begin{cases} 0 & \text{if } 1 \leq j \leq N_{p-1} \\ \frac{(j-N_{p-1})(j-N_{p-1}-1)}{2n_p} & \text{if } N_{p-1} < j \leq N_p \\ j - \frac{N_{p-1}+N_p+1}{2} & \text{if } N_p < j \leq N_r. \end{cases} \end{aligned} \tag{2.41}$$

Then we have, for $N_{p-1} < i \leq N_p, 1 \leq p \leq r$,

$$\begin{aligned} \bar{c_{i\cdot}} &\stackrel{\text{def}}{=} \frac{1}{N_r}\sum_{j=1}^{N_r} c_{ij} \\ &= \frac{1}{N_r}[\frac{n_p^2-1}{6} + \frac{(N_r-N_p)(N_r-N_{p-1})}{2}] \end{aligned} \tag{2.42}$$

and, for $N_{q-1} < j \leq N_q, 1 \leq q \leq r$,

$$\begin{aligned} \bar{c_{\cdot j}} &\stackrel{\text{def}}{=} \frac{1}{N_r}\sum_{i=1}^{N_r} c_{ij} \\ &= \frac{1}{N_r}[jN_{q-1} - \frac{N_{q-1}^2+N_{q-1}}{2} + \frac{(j-N_{q-1})(j-N_{q-1}-1)}{2}] \end{aligned} \tag{2.43}$$

$$\begin{aligned} \bar{c_{\cdot\cdot}} &\stackrel{\text{def}}{=} \frac{1}{N_r^2}\sum_{i,j=1}^{N_r} c_{ij} \\ &= \frac{1}{N_r^2}\sum_{p=1}^{r} n_p[\frac{n_p^2-1}{6} + \frac{(N_r-N_p)(N_r-N_{p-1})}{2}] \\ &= \frac{1}{N_r^2}\sum_{p=1}^{r} \frac{n_p(n_p^2+3N_pN_{p-1}-1)}{6} \\ &\approx \frac{N_r}{6}. \end{aligned} \tag{2.44}$$

Let

$$d(i,j) \stackrel{\text{def}}{=} c_{ij} - \bar{c_{i\cdot}} - \bar{c_{\cdot j}} + \bar{c_{\cdot\cdot}} \tag{2.45}$$

From (2.41), (2.42), (2.43) and (2.44), it is easy to see that

$$d^2(i,j) = O(N_r^2). \tag{2.46}$$

Also it will be shown in section 3.6 of Chapter 3 that

$$\begin{aligned}\frac{1}{N_r}\sum_{i,j=1}^{N_r} d^2(i,j) &\approx \frac{N_r^3}{180}\sum_{k=1}^{r}(-5T_k^4T_{k-1} - 5T_k^3T_{k-1}^2 + 5T_k^2T_{k-1}^3 + 5T_kT_{k-1}^4 + \\ &\quad 9T_k^3T_{k-1} + 9T_k^2T_{k-1}^2 - 21T_kT_{k-1}^3 + 3T_{k-1}^4) \\ &= MN_r^3 + o(N_r^3),\end{aligned} \tag{2.47}$$

where M is a computable constant $\neq 0$. Here $O(N_r^2), o(N_r^3)$ mean that when $min\{n_1\cdots n_r\} \to \infty$, $O(N_r^2)/N_r^2 \to$ constant, $o(N_r^3)/N_r^3 \to 0$, respectively. Hence when $min\{n_1\cdots n_r\} \to \infty$

$$\frac{max_{1\le i,j\le N_r} d^2(i,j)}{\frac{1}{N_r}\sum_{i,j=1}^{N_r} d^2(i,j)} \to 0. \tag{2.48}$$

By Theorem 2.1,

$$\begin{aligned} &F_r \overset{H_0}{\sim} N(\mu_{F_r}, \sigma_{F_r}^2),\ as\ min(n_1\cdots n_r) \to \infty \\ with \quad &\mu_{F_r} = \frac{1}{N_r}\sum_{i,j=1}^{N_r} c_{ij} = N_r\,\bar{c}_{..} = \frac{N_r^2}{6} \\ &\sigma_{F_r}^2 = \frac{1}{N_r - 1}\sum_{i,j=1}^{N_r} d^2(i,j),\end{aligned} \tag{2.49}$$

where $\sigma_{F_r}^2$ is approximated by (2.47).

4) Hamming

Let

$$H_r = \sum_{i=1}^{N_r} a_{i\pi(i)} \tag{2.50}$$

where

$$a_{ij} = \begin{cases} \frac{1}{n_k} & \text{if } i,j \in \{N_{k-1}+1,\cdots,N_k\}, 1 \le k \le r \\ 0 & \text{otherwise.}\end{cases} \tag{2.51}$$

Then, by (1.16),

$$d_{H_r}(\{\pi\}, E) \equiv H_r = \sum_{i=1}^{N_r} a_{i\pi(i)}. \tag{2.52}$$

Hence the critical region for the Hamming statistic is $\{H_r > c\}$, where c is a constant determined by the significance level α and the limiting null distribution of H_r.

Let

$$\bar{a}_{i\cdot} \stackrel{\text{def}}{=} \frac{1}{N_r}\sum_{j=1}^{N_r} a_{ij} = \frac{1}{N_r} \tag{2.53}$$

$$\bar{a}_{\cdot j} \stackrel{\text{def}}{=} \frac{1}{N_r}\sum_{i=1}^{N_r} a_{ij} = \frac{1}{N_r} \tag{2.54}$$

$$\bar{a}_{\cdot\cdot} \stackrel{\text{def}}{=} \frac{1}{N_r^2}\sum_{i,j=1}^{N_r} a_{ij} = \frac{1}{N_r}. \tag{2.55}$$

Then

$$d(i,j) \stackrel{\text{def}}{=} a_{ij} - \bar{a}_{i\cdot} - \bar{a}_{\cdot j} + \bar{a}_{\cdot\cdot} = a_{ij} - \frac{1}{N_r}, \tag{2.56}$$

and

$$\begin{aligned} max_{1\le i,j\le N_r} d^2(i,j) &\le max_{1\le i,j\le N_r} a_{ij}^2 \\ &= max_{1\le k\le r}(\frac{1}{n_k})^2 \\ &\approx \frac{1}{N_r^2} max_{1\le k\le r}\frac{1}{l_k^2}. \end{aligned} \tag{2.57}$$

Also

$$\begin{aligned} \frac{1}{N_r}\sum_{i,j=1}^{N_r} d^2(i,j) &= \frac{1}{N_r}\sum_{i,j=1}^{N_r}(a_{ij} - \frac{1}{N_r})^2 \\ &= \frac{1}{N_r}\sum_{i,j=1}^{N_r}(a_{ij}^2 - 2a_{ij}\frac{1}{N_r} + \frac{1}{N_r^2}) \\ &= \frac{r}{N_r} - \frac{2}{N_r} + \frac{1}{N_r} \\ &= \frac{r-1}{N_r}. \end{aligned} \tag{2.58}$$

Hence when $min(n_1 \cdots n_r) \to \infty$, (2.57) and (2.58) ensure that

$$\frac{max_{1\le i,j\le N_r} d^2(i,j)}{\frac{1}{N_r}\sum_{i,j=1}^{N_r} d^2(i,j)} \to 0. \tag{2.59}$$

By Theorem 2.1,

$$\begin{aligned} &H_r \stackrel{H_0}{\sim} N(\mu_{H_r}, \sigma_{H_r}^2), \text{ as } min\{n_1 \cdots n_r\} \to \infty \\ with \quad &\mu_{H_r} = \frac{1}{N_r}\sum_{i,j=1}^{N_r} a_{ij} = 1 \\ &\sigma_{H_r}^2 = \frac{1}{N_r - 1}\sum_{i,j=1}^{N_r} d^2(i,j) = \frac{r-1}{N_r-1}. \end{aligned} \tag{2.60}$$

2.3 The Two-way Layout Case: the ordered alternative

As shown in section 1.5 of Chapter 1, the test statistics $d_S(\{\pi\},E)$, $d_F(\{\pi\},E)$ and $d_H(\{\pi\},E)$ are all sums of independent random variables. Consequently their asymptotic normalities follow from the central limit theorem for independent random variables. The asymptotic normality of $d_K(\{\pi\},E)$ is established in the proof of Theorem 2.4 by showing the fact that it is equivalent to a U-statistic for which the asymptotic normality is given in Serfling (1980), page 192. Therefore, we only need to obtain the corresponding asymptotic means and variances.

Theorem 2.3 *For* $b \to \infty$,

$$d_S(\{\pi\},E) \overset{H_0}{\sim} N(\mu_S, \sigma_S^2),$$

$$\text{with} \qquad \mu_S = \frac{rb(rb+1)(rb-1)[(rb)!]^2}{12(r!)^{2b}}$$
$$\sigma_S^2 = \frac{br^2(r-1)[(rb+1)!]^4}{144(r+1)^2(r!)^{4b}}. \tag{2.61}$$

Proof: Noting the expression (1.22), we have, for $1 \le j \le b$,

$$E_{H_0}(C_{\{\pi_j\}}) = \tfrac{1}{r!}(1,\cdots,1)'_{(rb)!\times 1}.$$

Hence the calculation for μ_S follows from (1.27) and the fact that

$$E_{H_0}(T_S C_{\{\pi_j\}}) = \tfrac{1}{r!}T_S \cdot (1,\cdots,1)'_{(rb)!\times 1} = 0.$$

For σ_S^2, let $A = \frac{[(rb+1)!]^2}{(r!)^{2b}(r+1)^2}$. Then, from (1.27),

$$\begin{aligned}
\sigma_S^2 &= Var_{H_0}\{d_S(\{\pi\},E)\} = A^2\sum_{j=1}^{b} Var_{H_0}\{\sum_{i=1}^{r} iR_i^j\} \\
&= A^2\sum_{j=1}^{b} E_{H_0}\{\sum_{i=1}^{r} iR_i^j - r(r+1)^2/4)\}^2 \\
&= A^2\sum_{j=1}^{b} E_{H_0}\{\sum_{i\neq s=1}^{r} iR_i^j sR_s^j + \sum_{i=1}^{r} i^2(R_i^j)^2 - r(r+1)^2/2\sum_{i=1}^{r} iR_i^j + r^2(r+1)^4/16\} \\
&= A^2\sum_{j=1}^{b}\{\sum_{i\neq s=1}^{r}\sum_{p\neq q=1}^{r}\frac{ipsq}{r(r-1)} + \sum_{i=1}^{r}\sum_{s=1}^{r}\frac{i^2s^2}{r} - (r+1)^2/2\sum_{i=1}^{r}\sum_{s=1}^{r}\frac{is}{r} + r^2(r+1)^4/16\} \\
&= A^2 b\frac{r^2(r+1)^2(r-1)}{144} \\
&= \frac{br^2(r-1)[(rb+1)!]^4}{144(r+1)^2(r!)^{4b}}.
\end{aligned} \tag{2.62}$$

This completes the proof of the theorem.

Theorem 2.4 *When $b \to \infty$,*

$$d_K(\{\pi\}, E) \overset{H_0}{\sim} N(\mu_K, \sigma_K^2),$$

$$\text{with} \quad \mu_K = \frac{rb(rb-1)[(rb)!]^2}{2(r!)^{2b}}$$
$$\sigma_K^2 = \frac{b^3 r(r+1)^2(13r^2-2r+1)[(rb)!]^4}{144(r-1)(r!)^{4(b-2)}[(2r)!(r-1)!(r+1)!]^2}. \tag{2.63}$$

Proof: To establish the asymptotic normality, let $R^j = [R_1^j, \cdots, R_r^j]$ for $1 \le j \le b$ and set $f(R^{j_1}, R^{j_2}) = \sum_{1\le i_1, i_2 \le r} a[R_{i_1}^{j_1}, R_{i_2}^{j_2}]a[i_1, i_2]$ for $j_1 < j_2$. It follows from the definition (1.31) that f is symmetric. Since $R^1, \cdots, R^b$ are independent, the following statistic

$$U_b = \sum_{1\le j_1 < j_2 \le b} f(R^{j_1}, R^{j_2})$$

is a U-Statistic. Serfling (1980), page 183, shows that when $b \to \infty$,

$$Var_{H_0}\{U_b\} \approx \frac{b^3(r-1)(r+1)^2[(2r)!]^4}{9[(r-1)!(r+1)!]^4}.$$

On the other hand,

$$Var_{H_0}\{\sum_{j=1}^{b} \sum_{1\le i_1<i_2\le r} sgn[R_{i_2}^j - R_{i_1}^j]\}$$
$$= \sum_{j=1}^{b} Var_{H_0}\{\sum_{1\le i_1<i_2\le r} sgn[R_{i_2}^j - R_{i_1}^j]\}$$
$$\approx bM, \text{ where } M \text{ is a constant.}$$

Therefore the asymptotic distribution of $d_K(\{\pi\}, E)$ in (1.36) is dominated by U_b when b is getting large. Hence the value of μ_K immediately follows from the same arguments as in Theorem 2.3, and the value of σ_K as $b \to \infty$, is given by

$$\begin{aligned} Var_{H_0}\{d_K(\{\pi\}, E)\} &\approx \frac{[(rb)!]^4}{(r!)^{4(b-2)}[(2r)!]^4} Var_{H_0}\{U_b\} \\ &\approx \frac{[(rb)!]^4}{(r!)^{4(b-2)}[(2r)!]^4} \frac{b^3(r-1)(r+1)^2[(2r)!]^4}{9[(r-1)!(r+1)!]^4} \\ &= \frac{b^3(r-1)(r+1)^2[(rb)!]^4}{9(r!)^{4(b-2)}[(r-1)!(r+1)!]^4}. \end{aligned}$$

This completes the proof.

Theorem 2.5 *When $b \to \infty$,*

$$d_F(\{\pi\}, E) \overset{H_0}{\sim} N(\mu_F, \sigma_F^2),$$

$$\text{with} \quad \mu_F = \frac{(rb+1)(rb-1)[(rb)!]^2}{6(r!)^{2b}}$$
$$\sigma_F^2 = \frac{b\{[(b-1)r]!\}^4}{(r!)^{4(b-1)}} \cdot \frac{1}{r-1} \sum_{i,j=1}^{r} F_{ij}^2,$$

where

$$F_{ij} = f_{ij} - \bar{f}_{\cdot j} - \bar{f}_{i\cdot} + \bar{f}_{\cdot\cdot}$$

$$\text{and} \quad f_{ij} = \sum_{s=i}^{r(b-1)+i} \sum_{t=j}^{r(b-1)+j} \binom{s-1}{i-1}\binom{rb-s}{r-i}\binom{t-1}{j-1}\binom{rb-t}{r-j} max(s,t). \tag{2.64}$$

Proof: This follows directly by applying Theorem 2.1 to (1.39).

Theorem 2.6 *When $b \to \infty$,*

$$d_H(\{\pi\}, E) \overset{H_0}{\sim} N(\mu_H, \sigma_H^2),$$

$$\text{with} \quad \mu_H = \frac{rb[(rb)!]^2}{(r!)^{2b}} - \frac{b\{[r(b-1)]!\}^2}{r(r!)^{2(b-1)}} \cdot \sum_{j=1}^{r}\sum_{i=1}^{r} \sum_{t=max(j,i)}^{(r-1)b+min(j,i)} \binom{t-1}{j-1}\binom{rb-t}{r-j}\binom{t-1}{i-1}\binom{rb-t}{r-i}$$

$$\sigma_H^2 = \frac{b\{[r(b-1)]!\}^4}{(r!)^{4(b-1)}} \cdot \frac{1}{r-1} \sum_{i,j=1}^{r} H_{ij}^2,$$

where

$$H_{ij} = h_{ij} - \bar{h}_{\cdot j} - \bar{h}_{i\cdot} + \bar{h}_{\cdot\cdot}$$

$$\text{and} \quad h_{ij} = \sum_{t=max(j,i)}^{(r-1)b+min(j,i)} \binom{t-1}{j-1}\binom{rb-t}{r-j}\binom{t-1}{i-1}\binom{rb-t}{r-i}. \tag{2.65}$$

Proof: This follows directly by applying Theorem 2.1 to (1.40).

We can see from Theorem 2.3 to 2.6 that, for the two-way layout of the multi-sample case, the involved means and variances for the various test statistics are computable, even though some appear a little complicated. With the help of modern computers, it should be easier to complete the calculations involved.

2.4 The Asymptotic Equivalence of $d_{K_r}(\{\pi\}, E)$ and $d_{S_r}(\{\pi\}, E)$ Under the Null Hypothesis in the Multi-sample Case

Theorem 2.7 *Under the null hypothesis, the Spearman's Rho and Kendall's Tau statistics in the multi-sample case are asymptotically equivalent, i.e. when $min(n_1, \cdots, n_r) \to \infty$*

$$Var_{H_0}\left\{\frac{d_{K_r}(\{\pi\}, E) - E_{H_0}\{d_{K_r}(\{\pi\}, E)\}}{[Var_{H_0}\{d_{K_r}(\{\pi\}, E)\}]^{1/2}} - \frac{d_{S_r}(\{\pi\}, E) - E_{H_0}\{d_{S_r}(\{\pi\}, E)\}}{[Var_{H_0}\{d_{S_r}(\{\pi\}, E)\}]1/2}\right\} \to 0. \tag{2.66}$$

Proof: Let $S_r = \sum_{i=1}^{N_r} c(i)\frac{\pi(i)}{N_r+1}$ be given by (2.26) and define

$$K_r = \sum_{k=1}^{r} \sum_{i=N_{k-1}+1}^{N_k} \sum_{j=N_k+1}^{N_r} sgn[\pi(i) - \pi(j)]. \tag{2.67}$$

Then, by (1.13) and (1.14),

$$\begin{aligned} &\frac{d_{K_r}(\{\pi\},E) - E_{H_0}\{d_{K_r}(\{\pi\},E)\}}{[Var_{H_0}\{d_{K_r}(\{\pi\},E)\}]^{1/2}} - \frac{d_{S_r}(\{\pi\},E) - E_{H_0}\{d_{S_r}(\{\pi\},E)\}}{[Var_{H_0}\{d_{S_r}(\{\pi\},E)\}]1/2} \\ &= \frac{K_r - E_{H_0}K_r}{[Var_{H_0}K_r]^{1/2}} + \frac{S_r - E_{H_0}S_r}{[Var_{H_0}S_r]^{1/2}}. \end{aligned} \tag{2.68}$$

Now it is enough to prove that $K_r \approx -S_r$. By (2.35),

$$Var_{H_0}(S_r) = \sigma^2_{S_r} = \frac{N_r^3}{12}\sum_{k=1}^{r} t_k T_k T_{k-1}. \tag{2.69}$$

By symmetry, $E_{H_0}(K_r) = 0$. Therefore

$$\begin{aligned} Var_{H_0}(K_r) &= E(K_r^2) \\ &= E_{H_0}\{\sum_{k=1}^{r} \sum_{i=N_{k-1}+1}^{N_k} \sum_{j=N_k+1}^{N_r} \sum_{i_1=N_{k-1}+1}^{N_k} \sum_{j_1=N_k+1}^{N_r} sgn[\pi(i)-\pi(j)]sgn[\pi(i_1)-\pi(j_1)] + \\ &\quad 2\sum_{1\le k<k_1\le r} \sum_{i=N_{k-1}+1}^{N_k} \sum_{j=N_k+1}^{N_r} \sum_{i_1=N_{k_1-1}+1}^{N_{k_1}} \sum_{j_1=N_{k_1}+1}^{N_r} sgn[\pi(i)-\pi(j)]sgn[\pi(i_1)-\pi(j_1)]\} \\ &\overset{\text{def}}{=} A_1(first\ term) + A_2(second\ term). \end{aligned} \tag{2.70}$$

Also we have, for mutually different i, j, j_1,

$$\begin{aligned} &E_{H_0}\{sgn[\pi(i)-\pi(j)]sgn[\pi(i)-\pi(j_1)]\} \\ &= \sum_{i=1}^{N_r}\sum_{i\ne j=1}^{N_r}\sum_{s\ne i,j} \frac{sgn[j-i]sgn[s-i]}{N_r(N_r-1)(N_r-2)} \\ &= \sum_{i=1}^{N_r}\sum_{i\ne j=1}^{N_r}\sum_{s\ne j} \frac{sgn[j-i]sgn[s-i]}{N_r(N_r-1)(N_r-2)} \\ &= \sum_{i=1}^{N_r}\sum_{i\ne j=1}^{N_r}\sum_{s=1}^{N_r} \frac{sgn[j-i]sgn[s-i] - sgn[j-i]^2}{N_r(N_r-1)(N_r-2)} \\ &= \frac{\sum_{i=1}^{N_r}(N_r+1-2i)^2 - N_r(N_r-1)}{N_r(N_r-1)(N_r-2)} \\ &= \frac{4N_r(N_r+1)(N_r+2)/6 - 4N_r(N_r+1)(N_r+1)/2 + N_r(N_r+1)(N_r+1) - N_r(N_r-1)}{N_r(N_r-1)(N_r-2)} \\ &= \frac{1}{3} \end{aligned} \tag{2.71}$$

and, for mutually different i, j, i_1, j_1,

$$
\begin{aligned}
& E_{H_0}\{sgn[\pi(i)-\pi(j)]sgn[\pi(i_1)-\pi(j_1)]\} \\
& = E_{H_0}\{E_{H_0}\{sgn[\pi(i)-\pi(j)]sgn[\pi(i_1)-\pi(j_1)]\}|\pi(j_1),\pi(i_1)\} \\
& = 0 \qquad \textit{by symmetry.}
\end{aligned} \tag{2.72}
$$

Therefore

$$
\begin{aligned}
& A_1(\textit{first term}) \\
& = E_{H_0}\{\sum_{k=1}^{r}(\sum_{i=i_1=N_{k-1}+1}^{N_k} +2\sum_{N_{k-1}<i<i_1\leq N_k})\sum_{j=N_k+1}^{N_r}\sum_{j_1=N_k+1}^{N_r} sgn[\pi(i)-\pi(j)]sgn[\pi(i_1)-\pi(j_1)]\} \\
& = E_{H_0}\{\sum_{k=1}^{r}\sum_{i=N_{k-1}+1}^{N_k}\sum_{j=N_k+1}^{N_r}\sum_{j_1=N_k+1}^{N_r} sgn[\pi(i)-\pi(j)]sgn[\pi(i)-\pi(j_1)]\} + \\
& \quad 2E_{H_0}\{\sum_{k=1}^{r}\sum_{N_{k-1}<i<i_1\leq N_k}\sum_{j=N_k+1}^{N_r}\sum_{j_1=N_k+1}^{N_r} sgn[\pi(i)-\pi(j)]sgn[\pi(i_1)-\pi(j_1)]\} \\
& = E_{H_0}\{\sum_{k=1}^{r}\sum_{i=N_{k-1}+1}^{N_k}\sum_{j=N_k+1}^{N_r} sgn[\pi(i)-\pi(j)]sgn[\pi(i)-\pi(j)] + \\
& \quad \sum_{k=1}^{r}\sum_{i=N_{k-1}+1}^{N_k} 2\sum_{N_k<j<j_1\leq N_r} sgn[\pi(i)-\pi(j)]sgn[\pi(i)-\pi(j_1)]\} + \\
& \quad 2E_{H_0}\{\sum_{k=1}^{r}\sum_{N_{k-1}<i<i_1\leq N_k}\sum_{j=N_k+1}^{N_r} sgn[\pi(i)-\pi(j)]sgn[\pi(i_1)-\pi j)]\} \\
& = \{\sum_{k=1}^{r}\sum_{i=N_{k-1}+1}^{N_k}\sum_{j=N_k+1}^{N_r} 1 + \sum_{k=1}^{r}\sum_{i=N_{k-1}+1}^{N_k} 2\sum_{N_k<j<j_1\leq N_r}\frac{1}{3}\} + 2\{\sum_{k=1}^{r}\sum_{N_{k-1}<i<i_1\leq N_k}\sum_{j=N_k+1}^{N_r}\frac{1}{3}\} \\
& = \{\sum_{k=1}^{r} n_k(N_r-N_k) + \sum_{k=1}^{r}\frac{2}{3}n_k\frac{(N_r-N_k)(N_r-N_k+1)}{2}\} + \sum_{k=1}^{r}\frac{2}{3}\frac{n_k(n_k-1)}{2}(N_r-N_k) \\
& \approx N_r^3\frac{2}{3}\{\sum_{k=1}^{N_r}\frac{t_k(1-T_k)^2}{2} + \sum_{k=1}^{r}\frac{t_k^2(1-T_k)}{2}\} \\
& = N_r^3\frac{1}{3}\sum_{k=1}^{N_r} t_k(1-T_k)(1-T_{k-1}) \\
& = N_r^3\frac{1}{3}\sum_{k=1}^{N_r} t_kT_kT_{k-1}.
\end{aligned} \tag{2.73}
$$

$$
\begin{aligned}
& A_2(\textit{second term}) \\
& = E_{H_0}\{2\sum_{1\leq k<k_1\leq r}\sum_{i=N_{k-1}+1}^{N_k}\sum_{j=N_k+1}^{N_r}\sum_{i_1=N_{k_1-1}+1}^{N_{k_1}}\sum_{j_1=N_{k_1}+1}^{N_r} sgn[\pi(i)-\pi(j)]sgn[\pi(i_1)-\pi(j_1)]\}
\end{aligned}
$$

$$= E_{H_0}\{2 \sum_{1\le k<k_1\le r} \sum_{i=N_{k-1}+1}^{N_k} (\sum_{j=N_k+1}^{N_{k_1-1}} + \sum_{j=N_{k_1-1}+1}^{N_{k_1}} + \sum_{j=N_{k_1}+1}^{N_r}) \cdot$$

$$\sum_{i_1=N_{k_1-1}+1}^{N_{k_1}} \sum_{j_1=N_{k_1}+1}^{N_r} sgn[\pi(i)-\pi(j)]sgn[\pi(i_1)-\pi(j_1)]\}$$

$$= E_{H_0}\{2 \sum_{1\le k<k_1\le r} \sum_{i=N_{k-1}+1}^{N_k} \sum_{j=N_k+1}^{N_{k_1-1}} \sum_{i_1=N_{k_1-1}+1}^{N_{k_1}} \sum_{j_1=N_{k_1}+1}^{N_r} sgn[\pi(i)-\pi(j)]sgn[\pi(i_1)-\pi(j_1)]\} +$$

$$E_{H_0(2} \sum_{1\le k<k_1\le r} \sum_{i=N_{k-1}+1}^{N_k} \sum_{j=N_{k_1-1}+1}^{N_{k_1}} \sum_{i_1=N_{k_1-1}+1}^{N_{k_1}} \sum_{j_1=N_{k_1}+1}^{N_r} sgn[\pi(i)-\pi(j)]sgn[\pi(i_1)-\pi(j_1)]\} +$$

$$E_{H_0}\{2 \sum_{1\le k<k_1\le r} \sum_{i=N_{k-1}+1}^{N_k} \sum_{j=N_{k_1}+1}^{N_r} \sum_{i_1=N_{k_1-1}+1}^{N_{k_1}} \sum_{j_1=N_{k_1}+1}^{N_r} sgn[\pi(i)-\pi(j)]sgn[\pi(i_1)-\pi(j_1)]\}$$

$$= 0 + E_{H_0}\{2 \sum_{1\le k<k_1\le r} \sum_{i=N_{k-1}+1}^{N_k} \sum_{j=i_1=N_{k_1-1}+1}^{N_{k_1}} \sum_{j_1=N_{k_1}+1}^{N_r} sgn[\pi(i)-\pi(j)]sgn[\pi(i_1)-\pi(j_1)]\} +$$

$$E_{H_0}\{2 \sum_{1\le k<k_1\le r} \sum_{i=N_{k-1}+1}^{N_k} \sum_{j=j_1=N_{k_1}+1}^{N_r} \sum_{i_1=N_{k_1-1}+1}^{N_{k_1}} sgn[\pi(i)-\pi(j)]sgn[\pi(i_1)-\pi(j_1)]\}$$

$$= 2 \sum_{1\le k<k_1\le r} \sum_{i=N_{k-1}+1}^{N_k} \sum_{j=i_1=N_{k_1-1}+1}^{N_{k_1}} \sum_{j_1=N_{k_1}+1}^{N_r} (-\frac{1}{3}) +$$

$$+2 \sum_{1\le k<k_1\le r} \sum_{i=N_{k-1}+1}^{N_k} \sum_{j=j_1=N_{k_1}+1}^{N_r} \sum_{i_1=N_{k_1-1}+1}^{N_{k_1}} \frac{1}{3}$$

$$= 2 \sum_{1\le k<k_1\le r} n_k n_{k_1-1}(N_r - N_{k_1})(-\frac{1}{3}) + 2 \sum_{1\le k<k_1\le r} n_k n_{k_1-1}(N_r - N_{k_1})\frac{1}{3}$$

$$= 0. \tag{2.74}$$

Therefore

$$Var_{H_0}(K_r) = E_{H_0}(K_r^2) = A_1(first\ term) + A_2(second\ term) \approx \frac{N_r^3}{3} \sum_{k=1}^{r} t_k T_k T_{k-1}. \tag{2.75}$$

Furthermore we have, for $i \neq j$,

$$E_{H_0}\{sgn[\pi(i)-\pi(j)]\pi(j)\}$$

$$= \sum_{i=1}^{N_r} \sum_{j\neq i=1}^{N_r} \frac{sgn[j-i]i}{N_r(N_r-1)}$$

$$= \sum_{i=1}^{N_r} \sum_{j=1}^{N_r} \frac{sgn[j-i]i}{N_r(N_r-1)}$$

$$= \sum_{i=1}^{N_r} \sum_{j=1}^{N_r} \frac{(N_r+1-2i)i}{N_r(N_r-1)}$$

$$= \{\frac{(N_r+1)N_r(N_r+1)}{2} - 2\frac{(N_r+1)N_r(2N_r+1)}{6}\}/\{N_r(N_r-1)\}$$
$$= (\frac{1}{2}-\frac{2}{3})(N_r+1)$$
$$= -\frac{1}{6}(N_r+1) \tag{2.76}$$

and for $s \neq i, s \neq j$,

$$E\{sgn[\pi(i)-\pi(j)]\pi(s)\} = 0. \tag{2.77}$$

Then, by (2.76) and (2.77),

$$E_{H_0}\{(K_r - E_{H_0}K_r)(S_r - E_{H_0}S_r)\} = E_{H_0}\{K_r \cdot S_r\}$$
$$= E_{H_0}\{\sum_{k=1}^{r}\sum_{i=N_{k-1}+1}^{N_k}\sum_{j=N_k+1}^{N_r} sgn[\pi(i)-\pi(j)]\}\{\sum_{k=1}^{r}\sum_{i=N_{k-1}+1}^{N_k}(N_k+N_{k-1})\frac{\pi(i)}{N_r+1}\}$$
$$= E_{H_0}\{\sum_{k=1}^{r}\sum_{i=N_{k-1}+1}^{N_k}\sum_{i_1=k}^{r-1}\sum_{j=N_{i_1}+1}^{N_{i_1+1}} sgn[\pi(i)-\pi(j)]\}\{\sum_{k=1}^{r}\sum_{i=N_{k-1}+1}^{N_k}(N_k+N_{k-1})\frac{\pi(i)}{N_r+1}\}$$
$$= E_{H_0}\{\sum_{k=1}^{r}\sum_{i=N_{k-1}+1}^{N_k}\sum_{i_1=k}^{r-1}\sum_{j=N_{i_1}+1}^{N_{i_1+1}} sgn[\pi(i)-\pi(j)]\}(N_k+N_{k-1})\frac{\pi(i)}{N_r+1}\}$$
$$+E_{H_0}\{\sum_{k=1}^{r}\sum_{i=N_{k-1}+1}^{N_k}\sum_{i_1=k}^{r-1}\sum_{j=N_{i_1}+1}^{N_{i_1+1}} sgn[\pi(i)-\pi(j)]\}(N_{i_1}+N_{i_1+1})\frac{\pi(j)}{N_r+1}\}$$
$$= \{\sum_{k=1}^{r}\sum_{i=N_{k-1}+1}^{N_k}\sum_{i_1=k}^{r-1}\sum_{j=N_{i_1}+1}^{N_{i_1+1}}(\frac{1}{6})(N_k+N_{k-1})\}$$
$$+\{\sum_{k=1}^{r}\sum_{i=N_{k-1}+1}^{N_k}\sum_{i_1=k}^{r-1}\sum_{j=N_{i_1}+1}^{N_{i_1+1}}(-\frac{1}{6})(N_{i_1}+N_{i_1+1})\}$$
$$= \frac{1}{6}\{\sum_{k=1}^{r} n_k(N_r-N_k)(N_k+N_{k-1})\}$$
$$-\frac{1}{6}\{\sum_{k=1}^{r} n_k(N_r^2-N_k^2)\}$$
$$= -\frac{1}{6}\{\sum_{k=1}^{r} n_k(N_r-N_k)(N_r-N_{k-1})\}$$
$$\approx -\frac{N_r^3}{6}\{\sum_{k=1}^{r} t_k(1-T_k)(1-T_{k-1})\}$$
$$= -\frac{N_r^3}{6}\sum_{k=1}^{r} t_k T_k T_{k-1}. \tag{2.78}$$

Thus

$$E_{H_0}\{\frac{K_r - E_{H_0}K_r}{(Var_{H_0}K_r)^{1/2}} + \frac{S_r - E_{H_0}S_r}{(Var_{H_0}S_r)^{1/2}}\}^2$$

$$= 2 + 2\frac{E_{H_0}\{(K_r - E_{H_0}K_r)(S_r - E_{H_0}S_r)\}}{(Var_{H_0}K_r)^{1/2}(Var_{H_0}S_r)^{1/2}}$$
$$\approx 2 - 2\{\frac{N_r^3}{6}\sum_{k=1}^{r} t_k T_k T_{k-1}\}/\{(\frac{1}{3}\sum_{k=1}^{r} t_k T_k T_{k-1})^{1/2}(\frac{1}{12}N_r^3\sum_{k=1}^{r} t_k T_k T_{k-1})^{1/2}\}$$
$$= 2 - 2 = 0. \tag{2.79}$$

This completes the proof of the asymptotic equivalence of $d_{K_r}(\{\pi\}, E)$ and $d_{S_r}(\{\pi\}, E)$.

2.5 The Dispersion Case: the multi-sample case

Consider the test hypothesis (1.41).

1) Spearman's Rho

It follows from (1.44) that the use of the Spearman's Rho distance yields a constant.

2) Kendall's Tau

Noting that (1.45) is essentially the same as (1.14), the asymptotic normality of the Kendall's Tau statistic $d_{K_d}(\{\pi\}, E)$ is given by (2.35) since $d_{S_r}(\{\pi\}, E)$ and $d_{K_r}(\{\pi\}, E)$ are asymptotically equivalent under H_0.

3) Spearman's Footrule

Let

$$F_d = \sum_{i=1}^{N_r} c_{i\pi(i)} \tag{2.80}$$

where

$$c_{ij} = \sum_{s=1}^{N_r} a_{is}(j - s)I[j > s] \tag{2.81}$$

$$and \qquad a_{is} = \begin{cases} \frac{1}{m_k} & \text{if } i, s \in \{M_{k-1}+1, \cdots, M_k, N_r - M_k + 1, \cdots, N_r - M_{k-1}\} \\ 0 & \text{otherwise.} \end{cases} \tag{2.82}$$

Then, by (1.46),

$$d_{F_d}(\{\pi\}, E) \equiv F_d = \sum_{i=1}^{N_r} c_{i\pi(i)}. \tag{2.83}$$

Hence the critical region for the Spearman's Footrule statistic is $\{F_d < c\}$, where c is a constant determined by the significance level α and the limiting null distribution of F_d.

For $M_{p-1} < i \leq M_p$ or $N_r - M_p < i \leq N_r - M_{p-1}$, $1 \leq p \leq r$,

$$c_{ij} = 2\sum_{s=1}^{N_r} a_{is}(j - s)I[j > s]$$

$$= \Big(\sum_{s=M_{p-1}+1}^{M_p} + \sum_{s=N_r-M_p}^{N_r-M_{p-1}} \Big) \frac{1}{m_p}(j-s)I[j>s]$$

$$= \begin{cases} 0 & \text{if } 1 \le j \le M_{p-1} \\ \frac{(j-M_{p-1})(j-M_{p-1}-1)}{2m_p} & \text{if } M_{p-1} < j \le M_p \\ j - \frac{M_{p-1}+M_p+1}{2} & \text{if } M_p < j \le N_r - M_p \\ j - \frac{M_{p-1}+M_p+1}{2} + \frac{(j-N_r+M_{p-1})(j-N_r+M_{p-1}-1)}{2m_p} & \text{if } N_r - M_p < j \le N_r - M_{p-1} \\ 2j-(N_r+1) & \text{if } N_r - M_{p-1} < j \le N_r. \end{cases} \tag{2.84}$$

Then we have, for $M_{p-1} < i \le M_p$ or $N_r - M_p < i \le N_r - M_{p-1}, 1 \le p \le r$,

$$\begin{aligned} \bar{c_{i\cdot}} &\overset{\text{def}}{=} \frac{1}{N_r}\sum_{j=1}^{N_r} c_{ij} \\ &= \frac{1}{N_r}\Big[\frac{n_p^2-4}{12} + \frac{N_r(2N_r - N_p - N_{p-1})}{4} + \frac{N_p N_{p-1}}{4}\Big]. \end{aligned} \tag{2.85}$$

Setting $A_q^1 = \{M_{q-1}+1, \cdots, M_q\}$, $A_q^2 = \{N_r - M_q + 1, \cdots, N_r - M_{q-1}\}$, for $1 \le q \le r$, we have

$$\begin{aligned} \bar{c_{\cdot j}} &\overset{\text{def}}{=} \frac{1}{N_r}\sum_{i=1}^{N_r} c_{ij} \\ &= \begin{cases} \frac{1}{N_r}[jM_{q-1} - \frac{M_{q-1}^2+M_{q-1}}{2} + \frac{(j-M_{q-1})(j-M_{q-1}-1)}{2}] & \text{if } j \in A_q^1 \\ \frac{1}{N_r}[j(N_r-M_q) - \frac{(N_r-M_q)^2+(N_r-M_q)}{2} + \frac{(j-N_r+M_q)(j-N_r+M_q-1)}{2}] & \text{if } j \in A_q^2 \end{cases} \end{aligned} \tag{2.86}$$

$$\begin{aligned} \bar{c_{\cdot\cdot}} &\overset{\text{def}}{=} \frac{1}{N_r^2}\sum_{i,j=1}^{N_r} c_{ij} \\ &\approx \frac{1}{N_r^2}\sum_{p=1}^{r} \frac{n_p[n_p^2 + 3N_r(2N_r - N_p - N_{p-1}) + 3N_pN_{p-1}]}{12} \\ &= \frac{N_r}{3}. \end{aligned} \tag{2.87}$$

Let

$$d(i,j) \overset{\text{def}}{=} c_{ij} - \bar{c_{i\cdot}} - \bar{c_{\cdot j}} + \bar{c_{\cdot\cdot}} \tag{2.88}$$

From (2.84), (2.85), (2.86) and (2.87), it is easy to see that

$$d^2(i,j) = O(N_r^2). \tag{2.89}$$

Also it will be shown in section 3.6 of Chapter 3 that

$$\begin{aligned} \frac{1}{N_r}\sum_{i,j=1}^{N_r} d^2(i,j) &\approx \frac{N_r^3}{720}\sum_{k=1}^{r}(-5T_k^4T_{k-1} - 5T_k^3T_{k-1}^2 + 5T_k^2T_{k-1}^3 + 5T_kT_{k-1}^4 - 57T_k^4 + \\ &\qquad 69T_k^3T_{k-1} + 9T_k^2T_{k-1}^2 - 21T_kT_{k-1}^3 - 10) \\ &= MN_r^3 + o(N_r^3) \;\; \textit{with } M \textit{ is a constant } \ne 0. \end{aligned} \tag{2.90}$$

Hence when $min\{n_1 \cdots n_r\} \to \infty$

$$\frac{max_{1 \le i,j \le N_r} d^2(i,j)}{\frac{1}{N_r}\sum_{i,j=1}^{N_r} d^2(i,j)} \to 0. \tag{2.91}$$

By Theorem 2.1,

$$\begin{aligned} & F_d \overset{H_0}{\sim} N(\mu_{F_d}, \sigma^2_{F_d}),\ as\ min(n_1 \cdots n_r) \to \infty \\ with \quad & \mu_{F_d} = \frac{1}{N_r}\sum_{i,j=1}^{N_r} c_{ij} = N_r\ \bar{c}_{..} = \frac{N_r^2}{3} \\ & \sigma^2_{F_d} = \frac{1}{N_r - 1}\sum_{i,j=1}^{N_r} d^2(i,j), \end{aligned} \tag{2.92}$$

where $\sigma^2_{F_d}$ is approximated by (2.90).

4) Hamming

Let

$$H_d = \sum_{i=1}^{N_r} a_{i\pi(i)} \tag{2.93}$$

where

$$a_{ij} = \begin{cases} \frac{1}{n_k} & \text{if } i,j \in \{M_{k-1}+1, \cdots, M_k, N_r - M_k + 1, \cdots, N_r - M_{k-1}\}, 1 \le k \le r \\ 0 & \text{otherwise.} \end{cases} \tag{2.94}$$

Then, by (1.47),

$$d_{H_d}(\{\pi\}, E) \equiv H_d = \sum_{i=1}^{N_r} a_{i\pi(i)}. \tag{2.95}$$

Since (1.47) is similar to (1.16), the asymptotic null distribution of H_d is given by (2.60).

Chapter 3

Asymptotic Distributions Under the Alternatives

In Chapter 1 we have derived several testing statistics for the hypotheses (1.2), (1.10) and (1.41). Our choice of which is best will be based on the criterion of power. Thus we need to find limiting distributions under the alternatives.

Assume that $F(x)$ is a continuous distribution function and $f(x)$ is the density function of $F(x)$. Let $F^{-1}(u)$ be the inverse of $F(x)$; more precisely, $F^{-1}(u) = inf\{x : F(x) \geq u\}$. Put

$$I(f) = \int_{-\infty}^{\infty} [\frac{f'(x)}{f(x)}]^2 f(x) d(x) \tag{3.1}$$

$$\varphi(u, f) = \frac{f'(F^{-1}(u))}{f(F^{-1}(u))}, 0 < u < 1. \tag{3.2}$$

In the following discussions, we always assume that

$$I(f) < \infty. \tag{3.3}$$

Further assume that each sample distribution $F_k(x), 1 \leq k \leq r$ is a continuous function. We consider the following hypothesis to be tested:

$$H_0 : F_1(x) = \cdots = F_r(x) = F(x - \bar{d}) \text{ against}$$

$$H_1 : F_k(x) = F(x - \Delta_k), 1 \leq k \leq r \tag{3.4}$$

$$\Delta_1 < \cdots < \Delta_r \tag{3.5}$$

where $$d_i = \Delta_k \text{ if } N_{k-1} < i \leq N_k, 1 \leq k \leq r \tag{3.6}$$

and $$\bar{d} = \frac{1}{N_r} \sum_{i=1}^{N_r} d_i \tag{3.7}$$

with conditions

$$N_r^{1/2} \Delta_k \to \delta_k, 1 \leq k \leq r \text{ as } min\{n_1 \cdots n_r\} \to \infty. \tag{3.8}$$

Note that, as $min\{n_1 \cdots n_r\} \to \infty$, (3.3) and (3.7) hold if and only if

$$max_{1 \leq i \leq N_r} (d_i - \bar{d})^2 \to 0 \tag{3.9}$$

$$I(f) \sum_{i=1}^{N_r} (d_i - \bar{d})^2 \to b^2, \text{ a constant.} \tag{3.10}$$

3.1 Contiguity

Consider a sequence $\{p_v, q_v\}$ of simple hypotheses p_v and simple alternatives q_v defined on measure spaces (Ω_v, Ξ_v, u_v), $v > 0$, respectively.

Definition If for any sequence of events $\{A_v\}$, $A_v \in \Xi_v$,

$$P_v(A_v) \to 0 \text{ implies } Q_v(A_v) \to 0, \tag{3.11}$$

we say that the densities q_v are contiguous to the densities p_v, where $dP_v = p_v du_v, dQ_v = q_v du_v, v > 0$. Assume that H_v: the underlying distribution is P_v while K_v: the underlying distribution is Q_v.

If H_v is composite, we say that q_v is contiguous to H_v if for each v the convex hull $\bar{H}_v$ of H_v contains a density p_v such that (3.11) holds.

If both H_v and K_v are composite, we say that K_v is contiguous to H_v if (3.11) holds for some $p_v \in \bar{H}_v$ and $q_v \in \bar{K}_v$.

Contiguity implies that any sequence of random variables converging to zero in P_v-probability converges to zero in Q_v-probability, $v \to \infty$.

In this Chapter, we derive the limiting distributions of test statistics in the multi-sample case under the alternatives of contiguity as in Hajek and Sidak (1967).

3.2 Some Theorems

Theorem 3.1 *(VI.2.4 Theorem, Hajek-Sidak (1967)) Let*

$$S_n = \sum_{i=1}^{n}(c_i - \bar{c})a_n(\pi(i)) \tag{3.12}$$

where the scores $a_n(i)$ satisfy (2.4) and $\sum_{i=1}^{n}(c_i - \bar{c})^2/max_{1\le i\le n}(c_i - \bar{c})^2 \to \infty$. Then, under (3.9) and (3.10), S_n is under H_1 asymptotically normal (μ, σ^2) with

$$\mu = \sum_{i=1}^{n}(c_i - \bar{c})(d_i - \bar{d})\int_0^1 \varphi(u)\varphi(u, f)du \tag{3.13}$$

and $$\sigma^2 = \sum_{i=1}^{n}(c_i - \bar{c})^2\int_0^1 [\varphi(u) - \bar{\varphi}]^2 du. \tag{3.14}$$

Theorem 3.2 *(V.1.4 Theorem a, Hajek-Sidak (1967)) Let $\{U_i\}_1^n$ be i.i.d. uniform random variables on [0,1]. Suppose that $\varphi(u), 0 < u < 1$, is a square integrable function with*

$$\int_0^1 \varphi^2(u)du < \infty \tag{3.15}$$

and put $$a_n^{\varphi}(i) = E\{\varphi(U_1)|\pi(1) = i\}, 1 \le i \le n. \tag{3.16}$$

Then

$$\lim_{n\to\infty} E\{a_n^{\varphi}(\pi(1)) - \varphi(U_1)\}^2 = 0. \tag{3.17}$$

The following theorem shows the asymptotic equivalence of a linear rank statistic to a sum of independent random variables in the multi-sample case.

Theorem 3.3 *Let $\{U_i\}_1^{N_r}$ be i.i.d. uniform random variables on* $[0,1]$ *and set*

$$T_1 = \sum_{i=1}^{N_r}(d_i - \bar{d})\varphi(U_i, f) \tag{3.18}$$

$$T_2 = \sum_{i=1}^{N_r}(d_i - \bar{d})\varphi(\frac{\pi(i)}{N_r+1}, f) \tag{3.19}$$

Then, under (3.10),

$$T_1 \overset{H_0}{\sim} T_2, as\ min\{n_1 \cdots n_r\} \to \infty. \tag{3.20}$$

Proof: Let $U^{(\cdot)} = (U^{(1)} \dots U^{(N_r)})$ be the order statistics and $u^{(\cdot)} = (u^{(1)} \dots u^{(N_r)})$ be a value of $U^{(\cdot)}$. Then we have

$$\begin{aligned}
&E_{H_0}\{(T_1 - T_2)^2 | U^{(\cdot)} = u^{(\cdot)}\} \\
&= E_{H_0}\{\{\sum_{i=1}^{N_r}(d_i - \bar{d})[\varphi(U_i, f) - \varphi(\frac{\pi(i)}{N_r+1}, f)]\}^2 | U^{(\cdot)} = u^{(\cdot)}\} \\
&= E_{H_0}\{\sum_{i=1}^{N_r}(d_i - \bar{d})[\varphi(u^{(\pi(i))}, f) - \varphi(\frac{\pi(i)}{N_r+1}, f)]\}^2 \\
&= E_{H_0}\{\sum_{i=1}^{N_r}(d_i - \bar{d})^2[\varphi(u^{(\pi(i))}, f) - \varphi(\frac{\pi(i)}{N_r+1}, f)]^2 + \\
&\quad 2\sum_{1 \le i < j \le N_r}(d_i - \bar{d})(d_j - \bar{d})[\varphi(u^{(\pi(i))}, f) - \varphi(\frac{\pi(i)}{N_r+1}, f)][\varphi(u^{(\pi(j))}, f) - \varphi(\frac{\pi(j)}{N_r+1}, f)]\} \\
&= \sum_{i=1}^{N_r}(d_i - \bar{d})^2 \sum_{k=1}^{N_r}\frac{[\varphi(u^{(k)}, f) - \varphi(\frac{k}{N_r+1}, f)]^2}{N_r}\} + \\
&\quad 2\sum_{1 \le i < j \le N_r}(d_i - \bar{d})(d_j - \bar{d})\sum_{k=1}^{N_r}\sum_{s \ne k=1}^{N_r}\frac{[\varphi(u^{(k)}, f) - \varphi(\frac{k}{N_r+1}, f)][\varphi(u^{(s)}, f) - \varphi(\frac{s}{N_r+1}, f)]}{N_r(N_r-1)} \\
&\approx \frac{b^2}{I(f)}\sum_{k=1}^{N_r}\frac{[\varphi(u^{(k)}, f) - \varphi(\frac{k}{N_r+1}, f)]^2}{N_r} - \\
&\quad \frac{b^2}{I(f)}\sum_{k=1}^{N_r}\sum_{s \ne k=1}^{N_r}\frac{[\varphi(u^{(k)}, f) - \varphi(\frac{k}{N_r+1}, f)][\varphi(u^{(s)}, f) - \varphi(\frac{s}{N_r+1}, f)]}{N_r(N_r-1)} \\
&= \frac{b^2}{I(f)}E_{H_0}\{[\varphi(U_1, f) - \varphi(\frac{\pi(1)}{N_r+1}, f)]^2 | U^{(\cdot)} = u^{(\cdot)}\} - \\
&\quad \frac{b^2}{I(f)}E_{H_0}\{[\varphi(U_1, f) - \varphi(\frac{\pi(1)}{N_r+1}, f)][\varphi(U_2, f) - \varphi(\frac{\pi(2)}{N_r+1}, f)] U^{(\cdot)} = u^{(\cdot)}\}.
\end{aligned} \tag{3.21}$$

Hence

$$E_{H_0}\{(T_1 - T_2)^2\}$$

$$= E_{H_0}\{E_{H_0}[(T_1 - T_2)^2 | U^{(\cdot)} = u^{(\cdot)}]\}$$
$$\approx \frac{b^2}{I(f)} E_{H_0}\{[\varphi(U_1, f) - \varphi(\frac{\pi(1)}{N_r+1}, f)]^2\} -$$
$$\frac{b^2}{I(f)} E_{H_0}\{[\varphi(U_1, f) - \varphi(\frac{\pi(1)}{N_r+1}, f)][\varphi(U_2, f) - \varphi(\frac{\pi(2)}{N_r+1}, f)]\} \quad (3.22)$$

By Theorem 3.2,

$$E_{H_0}[\varphi(U_1, f) - \varphi(\frac{\pi(1)}{N_r+1}, f)]^2 \to 0$$

and
$$E_{H_0}|[\varphi(U_1, f) - \varphi(\frac{\pi(1)}{N_r+1}, f)][\varphi(U_2, f) - \varphi(\frac{\pi(2)}{N_r+1}, f)]|$$
$$\leq \frac{1}{2}\{E_{H_0}[\varphi(U_1, f) - \varphi(\frac{\pi(1)}{N_r+1}, f)]^2 + E_{H_0}[\varphi(u_2, f) - \varphi(\frac{\pi(2)}{N_r+1}, f)]^2\}$$
$$\to 0. \quad (3.23)$$

Therefore

$$E_{H_0}(T_1 - T_2)^2 \to 0,$$

which implies this theorem.

Theorem 3.4 *(LeCam)(VI.1.4 Lemma, Hajek-Sidak (1967)) Let S_n be any statistic and*

$$log L_n = \sum_{i=1}^{n} log \frac{f(x - d_i)}{f(x - \bar{d})}. \quad (3.24)$$

Assume that the pair $(S_n, log L_n)$ is under H_0 asymptotically jointly normal $N_2\left(\begin{pmatrix}\mu_1\\ \mu_2\end{pmatrix}, \begin{pmatrix}\sigma_1, & \sigma_{12}\\ \sigma_{12}, & \sigma_2\end{pmatrix}\right)$ with $\mu_2 = -\frac{1}{2}\sigma_2^2$. Then S_n is under H_1 asymptotically normal $(\mu_1 + \sigma_{12}, \sigma_1^2)$.

Limiting distributions of test statistics S_n under the alternative are important from the point of view of the power properties of the respective tests. Unfortunately, their derivations are considerably more difficult than the proofs of limiting distributions under the hypothesis. Nonetheless, in the contiguity case, the difficulties are essentially diminished by the above theorem.

Theorem 3.5 *(VI.2.1 Theorem, Hajek-Sidak (1967)) Let T_1 be given by (3.18) and $log L_n$ be given by (3.24). Then under (3.9) and (3.10)*

$$(log L_n - T_1 + \frac{1}{2}b^2) \stackrel{H_0}{\to} 0, \text{ as } min\{n_1 \cdots n_r\} \to \infty. \quad (3.25)$$

Moreover, $Log L_n$ is asymptotically normal $(-\frac{1}{2}b^2, b^2)$.

Theorem 3.6 *Consider a generalized linear rank statistic*

$$T = \sum_{i=1}^{N_r} a_{i\pi(i)}. \quad (3.26)$$

Let $d(i,j) \stackrel{\text{def}}{=} a_{ij} - \bar{a_{i\cdot}} - \bar{a_{\cdot j}} + \bar{a}_{\cdot\cdot}$. Under (3.3) and (3.8), if in addition

$$T \stackrel{H_0}{\sim} N(0,1) \quad (3.27)$$

and $$max_{1 \leq i,j \leq N_r} d(i,j) \approx O(N_r^p), p < 0, \quad (3.28)$$

then we have

$$T \overset{H_1}{\sim} N(\sigma_{12}, 1) \tag{3.29}$$

$$\textit{with} \quad \sigma_{12} = E_{H_0}\{T \cdot T_2\}, \textit{ and } T_2 \textit{ is given by } (3.19). \tag{3.30}$$

Proof: Conditions (3.3), (3.8) imply that (3.10) holds. It follows from Theorem 3.3 that

$$T_1 = \sum_{i=1}^{N_r}(d_i - \bar{d})\varphi(U_i, f) \overset{H_0}{\sim} T_2 = \sum_{i=1}^{N_r}(d_i - \bar{d})\varphi(\frac{\pi(i)}{N_r+1}, f) \tag{3.31}$$

From Theorem 3.5

$$\log L_{N_r} \overset{H_0}{\sim} T_1 - \frac{b^2}{2}. \tag{3.32}$$

Then

$$(T, \log L_{N_r}) \overset{H_0}{\sim} (T, T_1 - \frac{b^2}{2}) \overset{H_0}{\sim} (T, T_2 - \frac{b^2}{2}). \tag{3.33}$$

From (3.27), $T \overset{H_0}{\sim} N(0,1)$. Also it is easy to see that $T_1 \overset{H_0}{\sim} N(0, b^2)$, which implies that $T_2 \overset{H_0}{\sim} N(0, b^2)$. Therefore if

$$(T, T_2) \overset{H_0}{\sim} N_2\left(\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \begin{pmatrix} \sigma_1, \sigma_{12} \\ \sigma_{12}, \sigma_2 \end{pmatrix} \right) \tag{3.34}$$

then, by Theorem 3.4 and (3.34),

$$T \overset{H_1}{\sim} N(\mu_1 + \sigma_{12}, \sigma_1^2) = N(\sigma_{12}, 1) \tag{3.35}$$

Note that if (3.34) is true, then $\mu_1 = \mu_2 = 0, \sigma_1 = 1, \sigma_2 = b^2$, *and* $\sigma_{12} = E_{H_0}\{T \cdot T_2\}$.

It remains to prove (3.34), which is equivalent to showing that, for any $c_1, c_2 \in R^1$,

$$c_1 T + c_2 T_2 \overset{H_0}{\sim} \textit{normal distribution} \tag{3.36}$$

Assume

$$Var_{H_0}\{c_1 T + c_2 T_2\} \to c^2 > 0, \textit{ as } \min\{n_1 \cdots n_r\} \to \infty, \tag{3.37}$$

for otherwise the proof is trivial. Let

$$a^*_{ij} = c_1 a_{ij} + c_2(d_i - \bar{d})\varphi(\frac{j}{N_r+1}, f). \tag{3.38}$$

Then

$$c_1 T + c_2 T_2 = \sum_{i=1}^{N_r} a^*_{i\pi(i)}. \tag{3.39}$$

Put

$$\begin{aligned} d^*(i,j) &\overset{\text{def}}{=} a^*_{ij} - \bar{a}^*_{i\cdot} - \bar{a}^*_{\cdot j} + \bar{a}^*_{\cdot\cdot} \\ &= c_1 d(i,j) + c_2(d_i - \bar{d})[\varphi(\frac{j}{N_r+1}, f) - \varphi(\bar{\cdot}, f)]. \end{aligned} \tag{3.40}$$

Since $\varphi(u,f)$ is an integrable function, there exists a constant $M > 0$ such that $|\varphi(u,f)| \leq M$ a.s.. Also noting that, by (3.6)—(3.8),

$$\begin{aligned} |d_i - \bar{d}\,| &\leq max_{1\leq k\leq r}|\Delta_k - \frac{1}{N_r}\sum_{k=1}^{r} n_k\Delta_k| \\ &\approx \frac{1}{N_r^{1/2}} max_{1\leq k\leq r}|\delta_k - \sum_{k=1}^{r} t_k\delta_k| \\ &\approx O(N_r^{-1/2}). \end{aligned} \tag{3.41}$$

we have, by (3.28),

$$\begin{aligned} |d^*(i,j)| &\leq |c_1 d(i,j)| + |c_2(d_i - \bar{d})[\varphi(\frac{j}{N_r+1}, f) - \bar{\varphi}(\cdot, f)]| \\ &\leq |c_1| max_{1\leq i,j\leq N_r}|d(i,j)| + |c_2 M||d_i - \bar{d}\,| \\ &\approx O(N_r^p) + O(N_r^{-1/2}) \approx o(1), \\ i.e. \quad & max_{1\leq i,j\leq N_r}(d^*(i,j))^2 \approx o(1). \end{aligned} \tag{3.42}$$

On the other hand, from (3.40) and Theorem a, Hajek-Sidak (1967), page 57.

$$\begin{aligned} Var_{H_0}\{c_1 T + c_2 T_2\} &= Var_{H_0}\{\sum_{i=1}^{N_r} a^*_{i\pi(i)}\} \\ &= \frac{1}{N_r - 1}\sum_{i,j=1}^{N_r}(d^*(i,j))^2 \approx \frac{1}{N_r}\sum_{i,j=1}^{N_r}(d^*(i,j))^2. \end{aligned} \tag{3.43}$$

Hence, as $min\{n_1, \cdots n_r\} \to \infty$,

$$\frac{1}{N_r}\sum_{i,j=1}^{N_r}(d^*(i,j))^2 \to c^2 > 0, \tag{3.44}$$

and

$$\frac{max_{1\leq i,j\leq N_r}(d^*(i,j))^2}{\frac{1}{N_r}\sum_{i,j=1}^{N_r}(d^*(i,j))^2} \to 0. \tag{3.45}$$

By Theorem 2.1,

$$c_1 T + c_2 T_2 \overset{H_0}{\sim} normal\ distribution. \tag{3.46}$$

This finishes the proof.

REMARK: If the conditions (3.27) and (3.28) are replaced by

$$T \overset{H_0}{\sim} N(\mu_T, \sigma_T^2) \tag{3.47}$$

$$and \quad \frac{max_{1\leq i,j\leq N_r} d(i,j)}{\sigma_T} \approx O(N_r^p), p < 0. \tag{3.48}$$

Then

$$T \overset{H_1}{\sim} N(\mu_T + \sigma_{12}, \sigma_T^2).$$

3.3 The Two-sample Case

In the two-sample case, we assume

$$H_0 : F_1(x) = F_2(x) = F(x - \bar{d}) \text{ against}$$
$$H_1 : F_1(x) = F(x), F_2(x) = F(x - \Delta) \tag{3.49}$$

$$\text{where} \quad d_i = \begin{cases} 0 & \text{if } 1 \le i \le n_1 \\ \Delta & \text{if } n_1 < i \le n_1 + n_2 \end{cases} \tag{3.50}$$

$$\text{and} \quad \bar{d} = \frac{1}{n_1 + n_2} \sum_{i=1}^{n_1+n_2} d_i = \frac{n_2 \Delta}{n_1 + n_2}. \tag{3.51}$$

Further, let

$$max_{1 \le i \le n_1+n_2}(d_i - \bar{d})^2 = (\frac{\Delta}{n_1 + n_2})^2 max\{n_1^2, n_2^2\} \to 0 \tag{3.52}$$

$$\text{and} \quad I(f) \sum_{i=1}^{n_1+n_2} (d_i - \bar{d})^2 = I(f) \frac{\Delta^2 n_1 n_2}{n_1 + n_2} \to b^2, \text{ a constant.} \tag{3.53}$$

1) Spearman's Rho

Let $a(i) = \frac{i}{n_1+n_2+1}$ and

$$c_i = \begin{cases} 0 & \text{if } 1 \le i \le n_1 \\ 1 & \text{if } n_1 < i \le n_1 + n_2 \end{cases}$$

Then S_2 in (2.5) becomes

$$S_2 = \sum_{i=1}^{n_1+n_2} c_i a(\pi(i)).$$

It is easy to see that S_2 satisfies the conditions in Theorem 3.1. Using the square integrable function given in (2.7), we have

$$S_2 \overset{H_1}{\sim} N(\mu^*_{S_2}, \sigma^2_{S_2}), \text{ as } min(n_1, n_2) \to \infty$$
$$\text{with} \quad \mu^*_{S_2} = \frac{n_2}{2} + \Delta \frac{n_1 n_2}{n_1 + n_2} \int_0^1 u\varphi(u, f) du$$
$$\sigma^2_{S_2} = \frac{n_1 n_2}{12(n_1 + n_2)}. \tag{3.54}$$

2) Kendall's Tau

In view of the asymptotic equivalence, the Kendall's Tau statistic has the same distribution under the alternative as the Spearman's Rho statistic in the two-sample case.

3) Spearman's Footrule

Let $c_i = 0$, *if* $1 \le i \le n_1$ and 1, *if* $n_1 \le i \le n_1 + n_2$. Then F_2 in (2.10) becomes

$$F_2 = \sum_{i=1}^{n_1+n_2} c_i a(\pi(i)),$$

with scores $a(i)$ given by (2.11) and its corresponding square integrable function $\varphi(u)$ by (2.13). Applying Theorem 3.1 for F_2, we have

$$
\begin{aligned}
& F_2 \overset{H_1}{\sim} N(\mu^*_{F_2}, \sigma^2_{F_2}), \ as\ min(n_1, n_2) \to \infty \\
with \quad & \mu^*_{F_2} = \frac{n_1 n_2 (n_1 + 2n_2)}{3(n_1+n_2)^2} + \Delta \frac{n_1 n_2}{n_1+n_2} \int_0^1 \varphi(u)\varphi(u,f)du \\
& \sigma^2_{F_2} \approx \frac{n_1 n_2}{n_1+n_2} \frac{4\lambda^2(-2\lambda^2 + 2\lambda + 1)}{45},
\end{aligned}
\tag{3.55}
$$

where $\frac{n_1}{n_1+n_2} \to \lambda$, $as\ min\{n_1, n_2\} \to \infty$.

4) Hamming

Consider H given in (2.15). In view of (2.20)—(2.23), we have that (3.47) and (3.48) hold with $p = -1/2$. Hence it follows that

$$
\begin{aligned}
& H_2 \overset{H_1}{\sim} N(\mu^*_{H_2}, \sigma^2_{H_2}), \ as\ min(n_1, n_2) \to \infty \\
with \quad & \mu^*_{H_2} = 1 + \frac{bI^{-1/2}(f)}{(n_1+n_2)^{1/2}} \{ (\frac{\lambda}{1-\lambda})^{1/2} \int_\lambda^1 \varphi(u,f)du - (\frac{1-\lambda}{\lambda})^{1/2} \int_0^\lambda \varphi(u,f)du \} \\
& \sigma^2_{H_2} = \frac{1}{n_1+n_2-1},
\end{aligned}
\tag{3.56}
$$

where $\frac{n_1}{n_1+n_2} \to \lambda$, $as\ min\{n_1, n_2\} \to \infty$.

3.4 The Multi-sample Case

1) Spearman's Rho

Consider S_r in (2.26) with $c(i)$ given by (2.27) and the scores $a(i) = \frac{i}{N_r+1}$. It is easy to see that Theorem 3.1 can be used; hence we have

$$
\begin{aligned}
& S_r \overset{H_1}{\sim} N(\mu^*_{S_r}, \sigma^2_{S_r}), \ as\ min(n_1 \cdots n_r) \to \infty \\
with \quad & \mu^*_{S_r} = \frac{N_r^2}{2} + \sum_{i=1}^{N_r} (c_i - \bar{c})(d_i - \bar{d}) \int_0^1 u\varphi(u,f)du \\
& \sigma^2_{S_r} = \frac{N_r^3}{12} \sum_{k=1}^{r} t_k T_k T_{k-1}.
\end{aligned}
\tag{3.57}
$$

2) Kendall's Tau

In view of the asymptotic equivalence, the Kendall's Tau statistic has the same asymptotic distribution under the alternative as the Spearman's Rho statistic in the multi-sample case.

3) Spearman's Footrule

For F_r in (2.37), in view of (2.46)—(2.49), we have that (3.47) and (3.48) hold with $p = -1/2$. It follows that

$$\begin{aligned} & F_r \overset{H_1}{\sim} N(\mu^*_{F_r}, \sigma^2_{F_r}),\ as\ min(n_1 \cdots n_r) \to \infty \\ with\quad & \mu^*_{F_r} = \frac{N_r^2}{6} + \nu_{F_r} \\ & \sigma^2_{F_r} = \frac{1}{N_r - 1}\sum_{i,j=1} d^2(i,j), \end{aligned} \tag{3.58}$$

where $\sigma^2_{F_r}$ is given by (2.47). It will bw shown in section 3.6 that

$$\begin{aligned} \nu_{F_r} &= E_{H_0}\{F_r \cdot T_2\} \\ &\approx N_r^{3/2}\sum_{k=1}^{r}(\delta_k - \sum_{j=1}^{r} t_j\delta_j)\{\int_{T_{k-1}}^{T_k} \frac{(u-T_{k-1})^2}{2}[\varphi(u,f) - \bar{\varphi(f)}]du + \\ & \int_{T_k}^{1} t_k(u - \frac{T_k + T_{k-1}}{2})[\varphi(u,f) - \bar{\varphi(f)}]du\}, \end{aligned} \tag{3.59}$$

where $\bar{\varphi(f)} = \int_0^1 \varphi(u,f)du$.

4) Hamming

For H_r in (2.50), it follows that, with $p = -1/2$ in REMARK of Theorem 3.6,

$$\begin{aligned} & H_r \overset{H_1}{\sim} N(\mu^*_{H_r}, \sigma^2_{H_r}),\ as\ min(n_1 \cdots n_r) \to \infty \\ with\quad & \mu^*_{H_r} = 1 + \nu_{H_r} \\ & \sigma^2_{H_r} = \frac{r-1}{N_r - 1}, \end{aligned} \tag{3.60}$$

It will bw shown in section 3.6 that

$$\begin{aligned} \nu_{H_r} &= E_{H_0}\{H_r \cdot T_2\} \\ &\approx N^{-1/2}\sum_{k=1}^{r}\delta_k \int_{T_{k-1}}^{T_k} [\varphi(u,f) - \bar{\varphi(f)}]du. \end{aligned} \tag{3.61}$$

3.5 The Dispersion Case: the multi-sample case

Consider the test hypothesis (1.41) and assume that (3.3)—(3.10) hold with $I(f)$ in (3.1) replaced by

$$I_1(f) = \int_{-\infty}^{\infty} [-1 - x\frac{f'(x)}{f(x)}]^2 f(x)d(x), \tag{3.62}$$

and $\varphi(u,f)$ in (3.2) replaced by

$$\varphi_1(u,f) = -1 - F^{-1}(u)\frac{f'(F^{-1}(u))}{f(F^{-1}(u))}, 0 < u < 1. \tag{3.63}$$

Then, it is obvious that, under the above assumptions and with the same arguments or proofs, all the theorems in section 3.2 of Chapter 3 still hold in the dispersion case.

1) Spearman's Rho

In the dispersion case, (1.44) shows that the Spearman's Rho distance yields a constant.

2) Kendall's Tau

Comparing (1.45) and (1.14) and applying Theorem 3.1, the asymptotic distribution of the Kendall's Tau statistic under H_1 is given by (3.57) with $\varphi(u,f)$ replaced by $\varphi_1(u,f)$ since $d_{S_r}(\{\pi\},E)$ and $d_{K_r}(\{\pi\},E)$ are asymptotically equivalent under H_1.

3) Spearman's Footrule

For F_d in (2.80), in view of (2.89)—(2.92), it follows from REMARK of Theorem 3.6 that, with $p=-1/2$,

$$\begin{aligned} & F_d \overset{H_1}{\sim} N(\mu^*_{F_d},\sigma^2_{F_d}),\ as\ min(n_1\cdots n_r)\to\infty \\ with\quad & \mu^*_{F_d}=\frac{N_r^2}{3}+\nu_{F_d} \\ & \sigma^2_{F_d}=\frac{1}{N_r-1}\sum_{i,j=1}^{N_r} d^2(i,j), \end{aligned} \tag{3.64}$$

where $\sigma^2_{F_d}$ is given by (2.90) and $\nu_{F_d}=E_{H_0}\{F_d\cdot T_3\}$ with

$$T_3=\sum_{i=1}^{N_r}(d_i-\bar{d})\varphi_1(\frac{\pi(i)}{N_r+1},f), \tag{3.65}$$

It will be shown in section 3.6 that

$$\begin{aligned} \nu_{F_d} &= E_{H_0}\{F_d\cdot T_3\} \\ &\approx N_r^{3/2}\sum_{k=1}^{r}(\delta_k-\sum_{j=1}^{r}t_j\delta_j)\{\int_{T_{k-1}/2}^{T_k/2}(u-T_{k-1}/2)^2[\varphi_1(u,f)-\bar{\varphi_1}(f)]du+ \\ &\quad \int_{T_k/2}^{1-T_k/2} t_k(u-\frac{T_k+T_{k-1}}{4})[\varphi_1(u,f)-\bar{\varphi_1}(f)]du+ \\ &\quad \int_{1-T_k/2}^{1-T_{k-1}/2}[t_k(u-\frac{T_k+T_{k-1}}{4})+(u-1+T_{k-1}/2)^2][\varphi_1(u,f)-\bar{\varphi_1}(f)]du+ \\ &\quad \int_{1-T_{k-1}/2}^{1} t_k(2u-1)[\varphi_1(u,f)-\bar{\varphi_1}(f)]du\}, \end{aligned} \tag{3.66}$$

where $\bar{\varphi_1}(f)=\int_0^1 u\varphi(u,f)du$.

4) Hamming

Comparing (1.47), (1.16) and from (2.60), we have

$$\begin{aligned} H_d &\overset{H_1}{\sim} N(\mu^*_{H_d}, \sigma^2_{H_d}),\ as\ min(n_1 \cdots n_r) \to \infty \\ with \quad \mu^*_{H_d} &= 1 + \nu_{H_d} \\ \sigma^2_{H_d} &= \frac{r-1}{N_r - 1}, \end{aligned} \tag{3.67}$$

It will be shown in section 3.6 that

$$\begin{aligned} \nu_{H_d} &= E_{H_0}\{H_d \cdot T_3\} \\ &\approx N^{-1/2} \sum_{k=1}^{r} \delta_k \left(\int_{T_{k-1}/2}^{T_k/2} + \int_{1-T_k/2}^{1-T_{k-1}/2} \right) [\varphi_1(u,f) - \bar{\varphi_1}(f)] du. \end{aligned} \tag{3.68}$$

As far as all the test statistcis are concerned in this paper, we can see the fact by comparing the results in Chapter 2 that all the variances of these test statistics under the alternative remain the same as that under null hypothesis.

3.6 Computations of $\sigma^2_{F_r}$, ν_{F_r}, ν_{H_r}, $\sigma^2_{F_d}$, ν_{F_d} and ν_{H_d}

1. The computation of $\sigma^2_{F_r}$.

From (3.58), $\sigma^2_{F_r} = \frac{1}{N_r - 1} \sum_{i,j=1} d^2(i,j)$, where $d(i,j)$ is given by (2.45). Hence

$$\begin{aligned} &\frac{1}{N_r} \sum_{i,j=1}^{N_r} d^2(i,j) \overset{\text{def}}{=} \frac{1}{N_r} \sum_{i,j=1}^{N_r} (c_{ij} - \bar{c_{i\cdot}} - \bar{c_{\cdot j}} + \bar{c_{\cdot\cdot}})^2 \\ &= \frac{1}{N_r} \sum_{i,j=1}^{N_r} (c_{ij}^2 + \bar{c_{i\cdot}}^2 + \bar{c_{\cdot j}}^2 + \bar{c_{\cdot\cdot}}^2 - 2c_{ij}\bar{c_{i\cdot}} - 2c_{ij}\bar{c_{\cdot j}} + 2c_{ij}\bar{c_{\cdot\cdot}} + \\ &\quad 2\bar{c_{i\cdot}}\bar{c_{\cdot j}} - 2\bar{c_{i\cdot}}\bar{c_{\cdot\cdot}} - 2\bar{c_{\cdot j}}\bar{c_{\cdot\cdot}}) \\ &= \frac{1}{N_r} \sum_{i,j=1}^{N_r} c_{ij}^2 - \sum_{i=1}^{N_r} \bar{c_{i\cdot}}^2 - \sum_{j=1}^{N_r} \bar{c_{\cdot j}}^2 + N_r \bar{c_{\cdot\cdot}}^2 . \end{aligned} \tag{3.69}$$

Recalling (2.24) and (2.25), we have, from (2.41),

$$\begin{aligned} &\sum_{i,j=1}^{N_r} c_{ij}^2 = \sum_{k=1}^{r} \sum_{i=N_{k-1}+1}^{N_k} (\sum_{j=1}^{N_r} c_{ij}^2) \\ &= \sum_{k=1}^{r} \sum_{i=N_{k-1}+1}^{N_k} \{ \sum_{j=N_{k-1}+1}^{N_k} [\frac{(j-N_{k-1})(j-N_{k-1}-1)}{2n_k}]^2 + \sum_{j=N_k+1}^{N_r} [j - \frac{N_{k-1}+N_k+1}{2}]^2 \} \\ &\approx \sum_{k=1}^{r} \sum_{i=N_{k-1}+1}^{N_k} \{ \sum_{j=N_{k-1}+1}^{N_k} \frac{(j-N_{k-1})^4}{4n_k^2} + \sum_{j=N_k+1}^{N_r} [j^2 + \frac{(N_{k-1}+N_k)^2}{4} - (N_{k-1}+N_k)j] \} \end{aligned}$$

$$\approx \sum_{k=1}^{r} n_k \{ \frac{1}{4n_k^2} \frac{n_k^5}{5} + [\frac{N_r^3 - N_k^3}{3} + \frac{(N_{k-1} + N_k)^2 (N_r - N_k}{4} - \frac{(N_r^2 - N_k^2)(N_{k-1} + N_k)}{2}]\}$$

$$\approx N_r^4 \sum_{k=1}^{r} t_k \{ \frac{t_k^3}{20} + [\frac{1 - T_k^3}{3} + \frac{(T_{k-1} + T_k)^2 (1 - T_k)}{4} - \frac{(1 - T_k^2)(T_{k-1} + T_k)}{2}]\}$$

$$= N_r^4 \{ \sum_{k=1}^{r} \frac{1}{60} [-7T_k^3 T_{k-1} + 3T_k^2 T_{k-1}^2 + 3T_k T_{k-1}^3 + T_{k-1}^4 + 15T_k^2 T_{k-1} - 15T_k T_{k-1}^2] + \frac{3}{60} \}. \tag{3.70}$$

From (2.42), we have

$$\sum_{i=1}^{N_r} \bar{c_{i\cdot}}^2 = \sum_{k=1}^{r} \sum_{i=N_{k-1}+1}^{N_k} \{ \frac{1}{N_r} [\frac{n_k^2 - 1}{6} + \frac{(N_r - N_k)(N_r - N_{k-1})}{2}]\}^2$$

$$= \sum_{k=1}^{r} n_k \{ \frac{1}{N_r} [\frac{n_k^2 - 1}{6} + \frac{(N_r - N_k)(N_r - N_{k-1})}{2}]\}^2$$

$$\approx N_r^3 \sum_{k=1}^{r} t_k \{ \frac{t_k^2}{6} + \frac{(1 - T_k)(1 - T_{k-1})}{2} \}^2$$

$$= N_r^3 \frac{1}{36} \sum_{k=1}^{r} \{ T_k^4 T_{k-1} + T_k^3 T_{k-1}^2 - T_k^2 T_{k-1}^3 - T_k T_{k-1}^4 - 6T_k^3 T_{k-1} + 6T_k T_{k-1}^3 + 9T_k^2 T_{k-1} - 9T_k T_{k-1}^2 + 1 \}. \tag{3.71}$$

From (2.43), we have

$$\sum_{j=1}^{N_r} \bar{c_{\cdot j}}^2 = \sum_{k=1}^{r} \sum_{j=N_{k-1}+1}^{N_k} \{ \frac{1}{N_r} [N_{k-1} \cdot j - \frac{N_{k-1}^2 + N_{k-1}}{2} + \frac{(j - N_{k-1})(j - N_{k-1} - 1)}{2}]\}^2$$

$$\approx \frac{1}{N_r^2} \sum_{k=1}^{r} \sum_{j=N_{k-1}+1}^{N_k} \{ N_{k-1} \cdot j - \frac{N_{k-1}^2}{2} + \frac{(j - N_{k-1})^2}{2} \}^2$$

$$= \frac{1}{N_r^2} \sum_{k=1}^{r} \sum_{j=N_{k-1}+1}^{N_k} \{ N_{k-1}^2 \cdot j^2 + \frac{N_{k-1}^4}{4} + \frac{(j - N_{k-1})^4}{4} - N_{k-1}^3 \cdot j + N_{k-1} \cdot j (j - N_{k-1})^2 - \frac{N_{k-1}^2 (j - N_{k-1})^2}{2} \}$$

$$\approx \frac{1}{N_r^2} \sum_{k=1}^{r} \{ N_{k-1}^2 \frac{N_k^3 - N_{k-1}^3}{3} + \frac{N_{k-1}^4}{4} (N_k - N_{k-1}) + \frac{(N_k - N_{k-1})^5}{5 \cdot 4} - N_{k-1}^3 \frac{N_k^2 - N_{k-1}^2}{2} + N_{k-1} [\frac{(N_k - N_{k-1})^4}{4} + N_{k-1} \frac{(N_k - N_{k-1})^3}{3}] - \frac{N_{k-1}^2}{2} \frac{(N_k - N_{k-1})^3}{3} \}$$

$$\approx N_r^3 \sum_{k=1}^{r} \{ T_{k-1}^2 \frac{T_k^3 - T_{k-1}^3}{3} + \frac{T_{k-1}^4}{4} (T_k - T_{k-1}) + \frac{(T_k - T_{k-1})^5}{5 \cdot 4} - T_{k-1}^3 \frac{T_k^2 - T_{k-1}^2}{2} + T_{k-1} [\frac{(T_k - T_{k-1})^4}{4} + T_{k-1} \frac{(T_k - T_{k-1})^3}{3}] - \frac{T_{k-1}^2}{2} \frac{(T_k - T_{k-1})^3}{3} \}$$

$$= N_r^3 \sum_{k=1}^{r} \frac{1}{60}(3T_k^5 - 3T_{k-1}^5)$$

$$= N_r^3 \frac{1}{20}. \tag{3.72}$$

From (2.44), we have

$$\bar{c_{..}}^2 = N_r^2 \frac{1}{36}. \tag{3.73}$$

Therefore

$$\begin{aligned} \sigma_{F_r}^2 &= \frac{1}{N_r - 1} \sum_{i,j=1}^{N_r} d^2(i,j) \\ &\approx \frac{1}{N_r} \sum_{i,j=1}^{N_r} c_{ij}^2 - \sum_{i=1}^{N_r} \bar{c_{i.}}^2 - \sum_{j=1}^{N_r} \bar{c_{.j}}^2 + N_r \bar{c_{..}} \\ &\approx \frac{N_r^3}{180} \sum_{k=1}^{r} (-5T_k^4 T_{k-1} - 5T_k^3 T_{k-1}^2 + 5T_k^2 T_{k-1}^3 + 5T_k T_{k-1}^4 + \\ &\qquad 9T_k^3 T_{k-1} + 9T_k^2 T_{k-1}^2 - 21 T_k T_{k-1}^3 + 3T_{k-1}^4). \end{aligned} \tag{3.74}$$

2. The computation of ν_{F_r}.

From (2.37), (2.41), (3.19) and the expression in (3.59), we have

$$\nu_{F_r} = E_{H_0}\{F_r \cdot T_2\}$$

$$= E_{H_0}\{\sum_{i=1}^{N_r} c_{i\pi(i)} \sum_{j=1}^{N_r} (d_j - \bar{d})\varphi(\frac{\pi(j)}{N_r+1}, f)\}$$

$$= \sum_{i=1}^{N_r} E_{H_0}\{c_{i\pi(i)}(d_i - \bar{d})\varphi(\frac{\pi(i)}{N_r+1}, f)\} + \sum_{i \neq j=1}^{N_r} E_{H_0}\{c_{i\pi(i)}(d_j - \bar{d})\varphi(\frac{\pi(j)}{N_r+1}, f)\}$$

$$= \sum_{i=1}^{N_r} \frac{1}{N_r} \sum_{j=1}^{N_r} c_{ij}(d_i - \bar{d})\varphi(\frac{j}{N_r+1}, f) + \sum_{i \neq j=1}^{N_r} \frac{1}{N_r(N_r-1)} \sum_{k \neq s=1}^{N_r} c_{is}(d_j - \bar{d})\varphi(\frac{k}{N_r+1}, f)$$

$$= \sum_{i=1}^{N_r} \frac{1}{N_r} \sum_{j=1}^{N_r} c_{ij}(d_i - \bar{d})\varphi(\frac{j}{N_r+1}, f) - \sum_{i=1}^{N_r} \frac{1}{N_r(N_r-1)} \sum_{k \neq s=1}^{N_r} c_{is}(d_i - \bar{d})\varphi(\frac{k}{N_r+1}, f)$$

$$= \frac{1}{N_r - 1} \sum_{i=1}^{N_r} \sum_{j=1}^{N_r} c_{ij}(d_i - \bar{d})\varphi(\frac{j}{N_r+1}, f) -$$

$$\frac{1}{N_r - 1} \sum_{i=1}^{N_r} \sum_{s=1}^{N_r} c_{is}(d_i - \bar{d})\, \bar{\varphi}(\cdot, f), \quad where\ \bar{\varphi}(\cdot, f) = \frac{1}{N_r} \sum_{j=1}^{N_r} \varphi(\frac{j}{N_r+1}, f)$$

$$= \frac{1}{N_r - 1} \sum_{i=1}^{N_r} \sum_{j=1}^{N_r} c_{ij}(d_i - \bar{d})\{\varphi(\frac{j}{N_r+1}, f) - \bar{\varphi}(\cdot, f)\}$$

$$= \frac{1}{N_r - 1} \sum_{k=1}^{r} \sum_{i=N_{k-1}+1}^{N_k} \{ \sum_{j=1}^{N_r} c_{ij}(d_i - \bar{d})[\varphi(\frac{j}{N_r+1}, f) - \bar{\varphi(\cdot, f)}]\}$$

$$= \frac{1}{N_r - 1} \sum_{k=1}^{r} \sum_{i=N_{k-1}+1}^{N_k} \{ \sum_{j=N_{k-1}+1}^{N_k} \frac{(j - N_{k-1})(j - N_{k-1} - 1)}{2n_k} (\Delta_k - \bar{d})[\varphi(\frac{j}{N_r+1}, f) - \bar{\varphi(\cdot, f)}] +$$

$$\sum_{j=N_k+1}^{N_r} (j - \frac{N_k + N_{k-1} + 1}{2})(\Delta_k - \bar{d})[\varphi(\frac{j}{N_r+1}, f) - \bar{\varphi(\cdot, f)}]\}$$

$$= \frac{1}{N_r - 1} \sum_{k=1}^{r} (\Delta_k - \bar{d}) \{ \sum_{j=N_{k-1}+1}^{N_k} \frac{(j - N_{k-1})(j - N_{k-1} - 1)}{2} [\varphi(\frac{j}{N_r+1}, f) - \bar{\varphi(\cdot, f)}] +$$

$$\sum_{j=N_k+1}^{N_r} n_k (j - \frac{N_k + N_{k-1} + 1}{2})[\varphi(\frac{j}{N_r+1}, f) - \bar{\varphi(\cdot, f)}]\}$$

$$\approx N_r^{3/2} \sum_{k=1}^{r} (\delta_k - \sum_{j=1}^{r} t_j \delta_j) \{ \int_{T_{k-1}}^{T_k} \frac{(u - T_{k-1})^2}{2} [\varphi(u, f) - \bar{\varphi(f)}] du +$$

$$\int_{T_k}^{1} t_k (u - \frac{T_k + T_{k-1}}{2})[\varphi(u, f) - \bar{\varphi(f)}] du \}, \quad \textit{where } \bar{\varphi(f)} = \int_0^1 \varphi(u, f) du. \tag{3.75}$$

3. The computation of ν_{H_r}.

From (2.50), (2.51), (3.19) and the expression in (3.61), we have as for ν_{F_r}

$$\nu_{H_r} = E_{H_0}\{H_r \cdot T_2\}$$

$$= E_{H_0}\{ \sum_{i=1}^{N_r} a_{i\pi(i)} \sum_{i=1}^{N_r} (d_i - \bar{d}) \varphi(\frac{\pi(i)}{N_r+1}, f)\}$$

$$= \frac{1}{N_r - 1} \sum_{k=1}^{r} \sum_{i=N_{k-1}+1}^{N_k} \{ \sum_{j=1}^{N_r} a_{ij}(d_i - \bar{d})[\varphi(\frac{j}{N_r+1}, f) - \bar{\varphi(\cdot, f)}]\}, \quad \textit{similar to } \nu_{F_r}$$

$$= \frac{1}{N_r - 1} \sum_{k=1}^{r} n_k \{ \sum_{j=N_{k-1}+1}^{N_k} \frac{1}{n_k} (\Delta_k - \bar{d})[\varphi(\frac{j}{N_r+1}, f) - \bar{\varphi(\cdot, f)}]\}$$

$$= \frac{1}{N_r - 1} \sum_{k=1}^{r} (\Delta_k - \bar{d}) \{ \sum_{j=N_{k-1}+1}^{N_k} [\varphi(\frac{j}{N_r+1}, f) - \bar{\varphi(\cdot, f)}]\}$$

$$\approx N^{-1/2} \sum_{k=1}^{r} (\delta_k - \sum_{j=1}^{r} t_k \delta_k) \int_{T_{k-1}}^{T_k} [\varphi(u, f) - \bar{\varphi(f)}] du$$

$$= N^{-1/2} \sum_{k=1}^{r} \delta_k \int_{T_{k-1}}^{T_k} [\varphi(u, f) - \bar{\varphi(f)}] du. \tag{3.76}$$

4. The computation of $\sigma^2_{F_d}$.

From (2.92), $\sigma^2_{F_d} = \frac{1}{N_r-1}\sum_{i,j=1} d^2(i,j)$, where $d(i,j)$ is given by (2.88). Hence

$$\frac{1}{N_r}\sum_{i,j=1}^{N_r} d^2(i,j) \stackrel{\text{def}}{=} \frac{1}{N_r}\sum_{i,j=1}^{N_r}(c_{ij} - \bar{c_{i\cdot}} - \bar{c_{\cdot j}} + \bar{c_{\cdot\cdot}})^2$$

$$= \frac{1}{N_r}\sum_{i,j=1}^{N_r}(c_{ij}^2 + \bar{c_{i\cdot}}^2 + \bar{c_{\cdot j}}^2 + \bar{c_{\cdot\cdot}}^2 - 2c_{ij}\,\bar{c_{i\cdot}} - 2c_{ij}\,\bar{c_{\cdot j}} + 2c_{ij}\,\bar{c_{\cdot\cdot}} +$$

$$2\,\bar{c_{i\cdot}}\bar{c_{\cdot j}} - 2\,\bar{c_{i\cdot}}\bar{c_{\cdot\cdot}} - 2\,\bar{c_{\cdot j}}\bar{c_{\cdot\cdot}})$$

$$= \frac{1}{N_r}\sum_{i,j=1}^{N_r} c_{ij}^2 - \sum_{i=1}^{N_r}\bar{c_{i\cdot}}^2 - \sum_{j=1}^{N_r}\bar{c_{\cdot j}}^2 + N_r\,\bar{c_{\cdot\cdot}}^2. \tag{3.77}$$

Recalling (2.24) and (2.25) and the definition of $M_k's$ in 1.6, we have, from (2.84),

$$\sum_{i,j=1}^{N_r} c_{ij}^2 = \sum_{k=1}^{r}\Big(\sum_{i=M_{k-1}+1}^{M_k} + \sum_{i=N_r-M_k+1}^{N_r-M_{k-1}}\Big)\Big(\sum_{j=1}^{N_r} 4c_{ij}^2\Big)$$

$$\approx N_r^4\sum_{k=1}^{r}\frac{1}{240}[-19T_k^4 + 13T_k^3T_{k-1} + 3T_k^2T_{k-1}^2 + 3T_kT_{k-1}^3 + 15T_k^2T_{k-1} - 15T_kT_{k-1}^2 + 7]. \tag{3.78}$$

From (2.85), we have

$$\sum_{i=1}^{N_r}\bar{c_{i\cdot}}^2 = \sum_{k=1}^{r}\Big(\sum_{i=M_{k-1}+1}^{M_k} + \sum_{i=N_r-M_k+1}^{N_r-M_{k-1}}\Big)\bar{c_{i\cdot}}^2$$

$$\approx N_r^3\frac{1}{144}\sum_{k=1}^{r}\{T_k^4T_{k-1} + T_k^3T_{k-1}^2 - T_k^2T_{k-1}^3 - T_kT_{k-1}^4 - 6T_k^3T_{k-1} +$$

$$6T_kT_{k-1}^3 + 9T_k^2T_{k-1} - 9T_kT_{k-1}^2 + 15\}. \tag{3.79}$$

From (2.86), we have

$$\sum_{j=1}^{N_r}\bar{c_{\cdot j}}^2 = \sum_{k=1}^{r}\Big(\sum_{j=M_{k-1}+1}^{M_k} + \sum_{j=N_r-M_k+1}^{N_r-M_{k-1}}\Big)\bar{c_{\cdot j}}^2$$

$$\approx N_r^3\sum_{k=1}^{r}\frac{1}{60}\{(3T_k^5 - 3T_{k-1}^5) + (3(N_r - M_{k-1})^5 - 3(N_r - M_k)^5)\}$$

$$= N_r^3\frac{1}{20}. \tag{3.80}$$

From (2.87), we have

$$\bar{c_{\cdot\cdot}}^2 = \frac{1}{9}N_r^2. \tag{3.81}$$

Therefore

$$
\begin{aligned}
\sigma^2_{F_d} &= \frac{1}{N_r-1}\sum_{i,j=1}^{N_r} d^2(i,j) \\
&\approx \frac{1}{N_r}\sum_{i,j=1}^{N_r} c_{ij}^2 - \sum_{j=1}^{N_r} \bar{c}_{\cdot j}^{\;2} - \sum_{i=1}^{N_r} \bar{c}_{i\cdot}^{\;2} + N_r\,\bar{c}_{\cdot\cdot}^{\;2} \\
&\approx \frac{N_r^3}{720}\sum_{k=1}^{r}(-5T_k^4T_{k-1} - 5T_k^3T_{k-1}^2 + 5T_k^2T_{k-1}^3 + 5T_kT_{k-1}^4 - 5T_k^4 + \\
&\qquad 69T_k^3T_{k-1} + 9T_k^2T_{k-1}^2 - 21T_kT_{k-1}^3 - 10). \qquad (3.82)
\end{aligned}
$$

5. The computation of ν_{F_d}.

From (2.80), (2.84), (3.65) and the expression in (3.66), as for ν_{F_r}, we have

$$
\begin{aligned}
&\nu_{F_d} = E_{H_0}\{F_d\cdot T_3\} \\
&= \frac{1}{N_r-1}\sum_{i=1}^{N_r}\{\sum_{j=1}^{N_r} c_{ij}(d_i - \bar{d})[\varphi_1 \frac{j}{N_r+1}, f) - \bar{\varphi_1(\cdot,f)}]\}, \quad \textit{where}\ \ \bar{\varphi_1(\cdot,f)} = \frac{1}{N_r}\sum_{j=1}^{N_r}\varphi_1(\frac{j}{N_r+1},f) \\
&\approx \frac{1}{N_r-1}\sum_{k=1}^{r}(\sum_{i=M_{k-1}+1}^{M_k} + \sum_{i=N_r-M_k+1}^{N_r-M_{k-1}})\{\sum_{j=M_{k-1}+1}^{M_k}\frac{(j-M_{k-1})^2}{2m_k}(\Delta_k - \bar{d})[\varphi_1(\frac{j}{N_r+1},f) - \bar{\varphi_1(\cdot,f)}] + \\
&\qquad \sum_{j=M_k+1}^{N_r-M_k}(j - \frac{M_k+M_{k-1}}{2})(\Delta_k - \bar{d})[\varphi_1(\frac{j}{N_r+1},f) - \bar{\varphi_1(\cdot,f)}] + \\
&\qquad \sum_{j=N_r-M_k}^{N_r-M_{k-1}}[j - \frac{M_k+M_{k-1}}{2} + \frac{(j-N_r+M_{k-1})^2}{2m_k}](\Delta_k - \bar{d})[\varphi_1(\frac{j}{N_r+1},f) - \bar{\varphi_1(\cdot,f)}] + \\
&\qquad \sum_{j=N-r-M_{k-1}+1}^{N_r}(2j - N_r)(\Delta_k - \bar{d})[\varphi_1(\frac{j}{N_r+1},f) - \bar{\varphi_1(\cdot,f)}]\} \\
&= \frac{1}{N_r-1}\sum_{k=1}^{r}(\Delta_k - \bar{d})\{\sum_{j=M_{k-1}+1}^{M_k}(j-M_{k-1})^2[\varphi_1(\frac{j}{N_r+1},f) - \bar{\varphi_1(\cdot,f)}] + \\
&\qquad \sum_{j=M_k+1}^{N_r-M_k} n_k(j - \frac{M_k+M_{k-1}}{2})[\varphi_1(\frac{j}{N_r+1},f) - \bar{\varphi_1(\cdot,f)}] + \\
&\qquad \sum_{j=N_r-M_k}^{N_r-M_{k-1}}[n_k(j - \frac{M_k+M_{k-1}}{2}) + (j-N_r+M_{k-1})^2][\varphi_1(\frac{j}{N_r+1},f) - \bar{\varphi_1(\cdot,f)}] + \\
&\qquad \sum_{j=N-r-M_{k-1}+1}^{N_r} n_k(2j-N_r)[\varphi_1(\frac{j}{N_r+1},f) - \bar{\varphi_1(\cdot,f)}] \\
&\approx N_r^{3/2}\sum_{k=1}^{r}(\delta_k - \sum_{j=1}^{r} t_j\delta_j)\{\int_{T_{k-1}/2}^{T_k/2}(u - T_{k-1}/2)^2[\varphi_1(u,f) - \bar{\varphi_1(f)}]du +
\end{aligned}
$$

$$\int_{T_k/2}^{1-T_k/2} t_k\left(u-\frac{T_k+T_{k-1}}{4}\right)[\varphi_1(u,f)-\bar{\varphi_1}(f)]du+$$
$$\int_{1-T_k/2}^{1-T_{k-1}/2}\left[t_k\left(u-\frac{T_k+T_{k-1}}{4}\right)+(u-1+T_{k-1}/2)^2\right][\varphi_1(u,f)-\bar{\varphi_1}(f)]du+$$
$$\int_{1-T_{k-1}/2}^{1} t_k(2u-1)[\varphi_1(u,f)-\bar{\varphi_1}(f)]du\}$$
$$where\ \ \bar{\varphi_1}(f)=\int_0^1 u\varphi(u,f)du. \tag{3.83}$$

6. The computation of ν_{H_4}.

From (2.93), (2.94), (3.68) and the expression in (3.65), as for ν_{H_r}, we have

$$\begin{aligned}\nu_{H_4} &= E_{H_0}\{H_r\cdot T_3\}\\ &\approx N^{-1/2}\sum_{k=1}^{r}\delta_k\left(\int_{T_{k-1}/2}^{T_k/2}+\int_{1-T_k/2}^{1-T_{k-1}/2}\right)[\varphi_1(u,f)-\bar{\varphi_1}(f)]du.\end{aligned} \tag{3.84}$$

Chapter 4

Asymptotic Powers and Efficiencies

In Chapters 2 and 3, we derived the limiting distributions under both the null and the alternative hypotheses for the distances defined in section 1.2 of Chapter 1. These results make it possible for us to compare the available statistics based on the behavior of the limiting distributions under the alternative hypothesis.

Asymptotic Power Let T_n be any statistic with $\lim_{n\to\infty} P_{H_0}(T_n > x) = G(x)$, where $G(x)$ is some distribution function. Given the significance level α, we find k_α such that

$$\alpha = \lim_{n\to\infty} P_{H_0}\left(\frac{T_n - E_{H_0}\{T_n\}}{(Var_{H_0}\{T_n\})^{1/2}} > k_\alpha\right). \tag{4.1}$$

We define the asymptotic power of T_n denoted by $AP(T_n)$ to be

$$AP(T_n) \stackrel{\text{def}}{=} \lim_{n\to\infty} P_{H_1}\left(\frac{T_n - E_{H_0}\{T_n\}}{(Var_{H_0}\{T_n\})^{1/2}} > k_\alpha\right). \tag{4.2}$$

Asymptotic Power Efficiency Let Ξ be a set of test statistics. $T_n' \in \Xi$ is said to be asymptotically most powerful for H_0 against H_1 in the class Ξ if $AP(T_n') \geq max_{T_n\in\Xi} AP(T_n)$. Assume that $T_n' \in \Xi$ is the asymptotically most powerful test statistic for H_0 against H_1 and that it is asymptotically $N(0,\sigma_0^2)$ under H_0 and asymptotically $N(\mu_0,\sigma_0^2)$ under H_1. Consider another test statistic $T_n \in \Xi$ H_0 against H_1 which is asymptotically $N(0,\sigma^2)$ under H_0 and asymptotically $N(\mu,\sigma^2)$ under H_1. Let $\Phi(x)$ be the normal distribution $N(0,1)$. Then the asymptotic powers of T_n' and T_n are respectively equal to

$$1 - \Phi(k_\alpha - \mu_0\sigma_0^{-1}) \ and \ 1 - \Phi(k_\alpha - \mu\sigma^{-1}). \tag{4.3}$$

In this case, we define the asymptotic power efficiency of T_n denoted by $APE(T_n)$ to be the ratio $(\mu\sigma_0/\mu_0\sigma)^2$, i.e. $APE(T_n) = (\mu\sigma_0/\mu_0\sigma)^2$.

Theorem 4.1 *(VII.1.3 Theorem, Hajek-Sidak (1967)) Let*

$$S_{N_r} = \sum_{i=1}^{N_r} c_i a_{N_r}(\pi(i)) \tag{4.4}$$

Consider testing problem (3.4)—(3.7). Then, for (3.9) and (3.10), the following relation holds:

$$AP(S_{N_r}) = 1 - \Phi(k_\alpha - b), \tag{4.5}$$

$$where \quad \alpha = \lim_{min\{n_1\cdots n_r\}\to\infty} P_{H_1}\left(\frac{S_{N_r} - E_{H_0}\{S_{N_r}\}}{(Var_{H_0}\{S_{N_r}\})^{1/2}} > k_\alpha\right). \tag{4.6}$$

The maximum power is asymptotically reached by the test based on the statistic

$$S_{N_r} = \sum_{i=1}^{N_r} d_i a_{N_r}(\pi(i)), \tag{4.7}$$

where the d_i's are given by (3.6). If $I(f)$ in (3.1) is replaced by $I_1(f)$ in (3.62), $\varphi(u,f)$ in (3.2) by $\varphi_1(u,f)$ in (3.63) and (3.4)—(3.6) are replaced by (1.41), the theorem still holds.

In the next sections, we compute the asymptotic power efficiencies for the various test statistics.

4.1 The Two-sample Case

1) Spearman's Rho

By (2.8)

$$\begin{aligned} \alpha &= \lim_{min\{n_1,n_2\}\to\infty} P_{H_0}(\frac{S_2 - E_{H_0}\{S_2\}}{(Var_{H_0}\{S_2\})^{1/2}} > k_\alpha) \\ &= \lim_{min\{n_1,n_2\}\to\infty} P_{H_0}(\frac{S_2 - \frac{n_2}{2}}{[\frac{n_1 n_2}{12(n_1+n_2)}]^{1/2}} > k_\alpha) \\ &= 1 - \Phi(k_\alpha). \end{aligned} \tag{4.8}$$

From (3.53), (3.54) and (3.10)

$$\begin{aligned} AP(S_2) &= \lim_{min\{n_1,n_2\}\to\infty} P_{H_1}(\frac{S_2 - E_{H_0}\{S_2\}}{Var_{H_0}\{S_2\}} > k_\alpha) \\ &= \lim_{min\{n_1,n_2\}\to\infty} P_{H_1}(\frac{S_2 - \frac{n_2}{2}}{[\frac{n_1 n_2}{12(n_1+n_2)}]^{1/2}} > k_\alpha) \\ &= 1 - \Phi(k_\alpha - b[\frac{12}{I(f)}]^{1/2} \int_0^1 u\varphi(u,f)du). \end{aligned} \tag{4.9}$$

It follows from Theorem 4.1 that if $\int_0^1 u\varphi(u,f)du > 0$,

$$APE(S_2) = \frac{12[\int_0^1 u\varphi(u,f)du]^2}{I(f)}. \tag{4.10}$$

2) Kendall's Tau

In view of the asymptotic equivalence shown in section 2.4 of Chapter 2, $APE(K_2) = APE(S_2)$.

3) Spearman's Footrule

By (2.14), (3.55) and (3.10), we have

$$\begin{aligned} AP(F_2) &= \lim_{min\{n1,n_2\}\to\infty} P_{H_1}(\frac{F_2 - E_{H_0}\{F_2\}}{(Var_{H_0}\{F_2\})^{1/2}} > k_\alpha) \\ &= 1 - \Phi(k_\alpha - b[\frac{45}{4\lambda^2(-2\lambda^2+2\lambda+1)I(f)}]^{1/2} \int_0^1 \varphi(u)\varphi(u,f)du) \end{aligned} \tag{4.11}$$

where $\varphi(u)$ is given by (2.13). By Theorem 4.1, if $\int_0^1 u\varphi(u,f)du > 0$,

$$APE(F_2) = \frac{45}{4\lambda^2(-2\lambda^2+2\lambda+1)I(f)}\left[\int_0^1 \varphi(u)\varphi(u,f)du\right]^2, \tag{4.12}$$

4) Hamming

By (2.23) and (3.57), we have

$$\begin{aligned} AP(H_2) &= \lim_{\min\{n_1,n_2\}\to\infty} P_{H_1}\left(\frac{H_2 - E_{H_0}\{H_2\}}{(Var_{H_0}\{H_2\})^{1/2}} > k_\alpha\right) \\ &= 1 - \Phi(k_\alpha - bI^{-1/2}(f)[(\frac{\lambda}{1-\lambda})^{1/2}\int_\lambda^1 \varphi(u,f)du - (\frac{1-\lambda}{\lambda})^{1/2}\int_0^\lambda \varphi(u,f)du]). \end{aligned} \tag{4.13}$$

By Theorem 4.1, if $(\frac{\lambda}{1-\lambda})^{1/2}\int_\lambda^1 \varphi(u,f)du - (\frac{1-\lambda}{\lambda})^{1/2}\int_0^\lambda \varphi(u,f)du] > 0$, we have

$$APE(H_2) = \frac{[(\frac{\lambda}{1-\lambda})^{1/2}\int_\lambda^1 \varphi(u,f)du - (\frac{1-\lambda}{\lambda})^{1/2}\int_0^\lambda \varphi(u,f)du]^2}{I(f)}. \tag{4.14}$$

4.2 Examples

Assume that $\lambda = 1/2$.

1. Normal case. Let $f(x) = \frac{1}{(2\pi)^{1/2}\sigma}exp(\frac{-x^2}{2\sigma^2}), -\infty < x < \infty$. By (3.1) and (3.2)

$$I(f) = \sigma^{-2} \tag{4.15}$$

$$\int_{1/2}^1 \varphi(u,f)du = -\int_0^{1/2}\varphi(u,f)du = \frac{1}{(2\pi)^{1/2}\sigma} \approx 0.3989 \tag{4.16}$$

$$\int_{1/2}^1 u\varphi(u,f)du = \frac{1}{2(2\pi)^{1/2}} + \frac{1}{4(\pi)^{1/2}} \approx 0.3405 \tag{4.17}$$

$$\int_0^{1/2} u\varphi(u,f)du = -\frac{1}{2(2\pi)^{1/2}} + \frac{1}{4(\pi)^{1/2}} \approx -0.0585 \tag{4.18}$$

$$\int_{1/2}^1 u^2\varphi(u,f)du = \frac{1}{4(2\pi)^{1/2}} + \frac{1}{2(\pi)^{3/2}}(arctg\frac{1}{2^{1/2}} + \frac{\pi}{2}) \approx 0.2961 \tag{4.19}$$

$$\int_0^{1/2} u^2\varphi(u,f)du = -\frac{1}{4(2\pi)^{1/2}} - \frac{1}{2(\pi)^{3/2}}(arctg\frac{1}{2^{1/2}} + \frac{\pi}{2}) + \frac{1}{2(\pi)^{1/2}} \approx -0.0141. \tag{4.20}$$

Therefore, from (4.10), (4.12) and (4.14),

$$APE(S_2) = \frac{3}{\pi} \approx 0.9554 \tag{4.21}$$

$$APE(F_2) \approx 30 \times 0.02936 = 0.8808 \tag{4.22}$$

$$APE(H_2) = \frac{2}{\pi} \approx 0.6369. \tag{4.23}$$

2. Double-exponential case. Let $f(x) = \frac{1}{2}exp(-|x|), -\infty < x < \infty$. By (3.1) and (3.2)

$$I(f) = 1 \tag{4.24}$$

$$\int_{1/2}^{1} \varphi(u,f)du = -\int_{0}^{1/2} \varphi(u,f)du = 1/2 \tag{4.25}$$

$$\int_{1/2}^{1} u\varphi(u,f)du = 3/8, \int_{0}^{1/2} u\varphi(u,f)du = -1/8 \tag{4.26}$$

$$\int_{1/2}^{1} u^2\varphi(u,f)du = 7/24, \int_{0}^{1/2} u^2\varphi(u,f)du = -1/24. \tag{4.27}$$

Therefore, from (4.10), (4.12) and (4.14),

$$APE(S_2) = 3/4 \approx 0.7500 \tag{4.28}$$

$$APE(F_2) = 5/6 \approx 0.8333 \tag{4.29}$$

$$APE(H_2) = 1. \tag{4.30}$$

3. Logistic case. Let $f(x) = e^{-x}(1+e^{-x})^2, -\infty < x < \infty$. By (3.1) and (3.2)

$$I(f) = 1/3 \tag{4.31}$$

$$\int_{1/2}^{1} \varphi(u,f)du = -\int_{0}^{1/2} \varphi(u,f)du = 1/4 \tag{4.32}$$

$$\int_{1/2}^{1} u\varphi(u,f)du = 5/24, \int_{0}^{1/2} u\varphi(u,f)du = -1/24 \tag{4.33}$$

$$\int_{1/2}^{1} u^2\varphi(u,f)du = 17/96, \int_{0}^{1/2} u^2\varphi(u,f)du = -1/96. \tag{4.34}$$

Therefore, from (4.10), (4.12) and (4.14),

$$APE(S_2) = 1 \tag{4.35}$$

$$APE(F_2) = 125/128 \approx 0.9765 \tag{4.36}$$

$$APE(H_2) = 3/4 \approx 0.7500. \tag{4.37}$$

4.3 The Multi-sample Case

1) Spearman's Rho

From (2.27), (2.30), (3.6), (3.7) and (3.8), we have, as $min\{n_1 \cdots n_r\} \to \infty$,

$$\begin{aligned} \sum_{i=1}^{N_r}(c_i - \bar{c})(d_i - \bar{d}) &= \sum_{k=1}^{r} n_k(N_k + N_{k-1} - \bar{c})(\Delta_k - \bar{d}) \\ &\approx N_r^{3/2}\sum_{k=1}^{r} t_k(T_k + T_{k-1} - 1)(\delta_k - \sum_{k=1}^{r} t_k\delta_k) \\ &= N_r^{3/2}\sum_{k=1}^{r} t_k\delta_k(T_k + T_{k-1} - 1) \end{aligned} \tag{4.38}$$

$$\sum_{i=1}^{N_r}(d_i - \bar{d})^2 \approx \sum_{k=1}^{r} t_k(\delta_k - \sum_{k=1}^{r} t_k\delta_k)^2. \tag{4.39}$$

By (2.35), (3.57) and (4.38), we have

$$\begin{aligned} AP(S_r) &= \lim_{min\{n_1 \cdots n_r\} \to \infty} P_{H_1}\left(\frac{S_r - E_{H_0}\{S_r\}}{(Var_{H_0}\{S_r\})^{1/2}} > k_\alpha\right) \\ &= 1 - \Phi\left(k_\alpha - \frac{\sum_{k=1}^r t_k \delta_k (T_{k-1} + T_k - 1) \int_0^1 u\varphi(u,f)du}{\{1/12 \sum_{k=1}^r t_k T_k T_{k-1}\}^{1/2}}\right). \end{aligned} \tag{4.40}$$

Hence if $\sum_{k=1}^r t_k \delta_k (T_{k-1} + T_k - 1) \int_0^1 u\varphi(u,f)du > 0$ and we have by (3.10), (4.39) and Theorem 4.1,

$$APE(S_r) = \frac{12\{\sum_{k=1}^r t_k \delta_k (T_{k-1} + T_k - 1)\{\int_0^1 u\varphi(u,f)du\}^2}{I(f) \sum_{k=1}^r t_k T_{k-1} T_k \sum_{k=1}^r t_k (\delta_k - \sum_{k=1}^r t_k \delta_k)^2}. \tag{4.41}$$

2) Kendall's Tau

In view of section 2.4 of Chapter 2, the Kendall's Tau statistic is asymptotically equivalent to the Spearman's Rho statistic in the multi-sample case. Hence $APE(K_r) = APE(S_r)$.

3) Spearman's Footrule

From (2.49), (3.58), (3.74) and (3.75), we have, from Theorem 4.1,

$$\begin{aligned} APE(F_r) &= \frac{\nu_{F_r}^2}{\sigma_{F_r}^2} \cdot \frac{1}{b^2} \\ &\quad 180\{\sum_{k=1}^r (\delta_k - \sum_{j=1}^r t_j \delta_j)\{\int_{T_{k-1}}^{T_k} \frac{(u - T_{k-1})^2}{2}[\varphi(u,f) - \bar{\varphi(f)}]du + \\ &= \frac{\int_{T_k}^1 t_k (u - \frac{T_k + T_{k-1}}{2})[\varphi(u,f) - \bar{\varphi(f)}]du\}\}^2}{I(f) \sum_{k=1}^r t_k (\delta_k - \sum_{k=1}^r \delta_k t_k)^2 \sum_{k=1}^r (-5T_k^4 T_{k-1} - 5T_k^3 T_{k-1}^2 + 5T_k^2 T_{k-1}^3 + 5T_k T_{k-1}^4 + 9T_k^3 T_{k-1} + 9T_k^2 T_{k-1}^2 - 21 T_k T_{k-1}^3 + 3T_{k-1}^4)}. \end{aligned} \tag{4.42}$$

4) Hamming

From (2.60), (3.60) and (3.76), we have, from Theorem 4.1,

$$\begin{aligned} APE(H_r) &= \nu_{H_r}^2 \cdot \frac{1}{b^2} \\ &= \frac{\{\sum_{k=1}^r \delta_k \int_{T_{k-1}}^{T_k} [\varphi(u,f) - \bar{\varphi(f)}]du\}^2}{(r-1) I(f) \sum_{k=1}^r t_k (\delta_k - \sum_{k=1}^r \delta_k t_k)^2}. \end{aligned} \tag{4.43}$$

4.4 The Dispersion Case: the multi-sample case

1) Spearman's Rho

In the dispersion case, the use of the Spearman's Rho distance yields a constant.

2) Kendall's Tau

Comparing (1.14), (1.45) and noting that (1.13) and (1.14) are equivalent, $AP(K_d), APE(K_d)$ are given by (4.40), (4.41), respectively in which $I(f)$ and $\varphi(u,f)$ are replaced by the corresponding expressions $I_1(f)$ *and* $\varphi_1(u,f)$.

3) Spearman's Footrule

From (2.92), (3.64), (3.82) and (3.83), we have, from Theorem 4.1,

$$APE(F_d) = \frac{\nu_{F_d}^2}{\sigma_{F_d}^2} \cdot \frac{1}{b^2}$$

$$= \frac{\begin{array}{l} 720\{\sum_{k=1}^{r}(\delta_k - \sum_{j=1}^{r} t_j\delta_j)\{\int_{T_{k-1}/2}^{T_k/2}(u - T_{k-1}/2)^2[\varphi_1(u,f) - \bar{\varphi_1}(f)]du + \\ \int_{T_k/2}^{1-T_k/2} t_k(u - \frac{T_k + T_{k-1}}{4})[\varphi_1(u,f) - \bar{\varphi_1}(f)]du + \\ \int_{1-T_k/2}^{1-T_{k-1}/2}[t_k(u - \frac{T_k + T_{k-1}}{4}) + (u - 1 + T_{k-1}/2)^2][\varphi_1(u,f) - \bar{\varphi_1}(f)]du + \\ \int_{1-T_{k-1}/2}^{1} t_k(2u-1)[\varphi_1(u,f) - \bar{\varphi_1}(f)]du\}\}^2 \end{array}}{\begin{array}{l} I_1(f)\sum_{k=1}^{r} t_k(\delta_k - \sum_{k=1}^{r}\delta_k t_k)^2 \sum_{k=1}^{r}(-5T_k^4T_{k-1} - 5T_k^3T_{k-1}^2 + 5T_k^2T_{k-1}^3 + 5T_kT_{k-1}^4 + \\ -57T_k^4 + 69T_k^3T_{k-1} + 9T_k^2T_{k-1}^2 - 21T_kT_{k-1}^3 - 10) \end{array}}. \quad (4.44)$$

4) Hamming

Comparing (1.47) and (1.16), the asymptotic distribution of H_d under H_0 is given by (2.60). Then from (3.67) and (3.84), we have, from Theorem 4.1,

$$\begin{aligned} APE(H_d) &= \nu_{H_d}^2 \cdot \frac{1}{b^2} \\ &= \frac{\{\sum_{k=1}^{r}\delta_k(\int_{T_{k-1}/2}^{T_k/2} + \int_{1-T_k/2}^{1-T_{k-1}/2})[\varphi_1(u,f) - \bar{\varphi_1}(f)]du\}^2}{(r-1)I_1(f)\sum_{k=1}^{r} t_k(\delta_k - \sum_{k=1}^{r}\delta_k t_k)^2}. \end{aligned} \quad (4.45)$$

Chapter 5

The Multi-sample Problem in the Unordered Case

In the previous chapters, we discussed the location problem with ordered alternatives. In this Chapter, we deal with the multi-sample location problem with unordered alternatives. Analogous results can be obtained in the dispersion case.

Consider the following testing problem

$$\begin{aligned} H_0 &: F_1(x) = \cdots = F_r(x) \ against \\ H_1^h &: F_{e_{hr}}(x) \leq \cdots \leq F_{e_{h1}}(x), \end{aligned} \tag{5.1}$$

where, for a fixed h, $1 \leq h \leq r!$, $[e_{hr}, \cdots, e_{h1}] \in P_r$. (5.1) is the exact testing hypothesis problem we considered in the previous chapters. This chapter will discuss a more complicated testing hypothesisi problem. we consider the following composite testing hypothesis

$$\begin{aligned} H_0 &: F_1(x) = \cdots = F_r(x) \ against \\ H_1 &: \bigcup_{h=1}^{r!} H_1^h. \end{aligned} \tag{5.2}$$

Suppose that T_h is a testing statistic for the problem in (5.1) with critical region $\{T_h < c\}$. To test the hypothesis in (5.2), we define the test statistic in the following way:

I. *Linearization* : let $\tilde{\alpha} = (\alpha_1 \cdots \alpha_{r!}) \in R^{r!}$ and put

$$T_L(\tilde{\alpha}) = \sum_{h=1}^{r!} \alpha_h T_h$$

II. *Normalization* :

$$T_N(\tilde{\alpha}) = \frac{T_L(\tilde{\alpha}) - E_{H_0}\{T_L(\tilde{\alpha})\}}{[Var_{H_0}\{T_L(\tilde{\alpha})\}]^{1/2}}$$

III. *Minimization* :

$$T_M = min_{\tilde{\alpha} \in R^{r!}} T_N(\tilde{\alpha}).$$

Hence the critical region for testing the hypothesis in (5.2) is $\{T_M < c\}$. First we require the following lemma.

Lemma 5.1 *Let $\tilde{X} = (X_1 \cdots X_n)'$ be n-dimensional random variables and $\Sigma = Cov(\tilde{X})$ with $Rank(\Sigma) = q$. If $q < n$, then there exists an $n \times 1$ vector $\tilde{\alpha} = (\alpha_1 \cdots \alpha_n) \in R^n$ such that $\tilde{\alpha}'\tilde{X} = 0$ w.p.1..*

Proof: Since Σ is symmetric, there exists an $n \times n$ full rank matrix P such that

$$\Sigma = P \cdot diag\{\lambda_1, \cdots, \lambda_q, 0, \cdots, 0\} \cdot P'$$

with $\lambda_i \neq 0$, for $i = 1, \cdots, q$. Let $\tilde{Y} = P^{-1}\tilde{X}$. Then

$$Cov(\tilde{Y}) = diag\{\lambda_1, \cdots, \lambda_q, 0, \cdots, 0\}.$$

Hence $Var(Y_n) = 0$ which implies that $Y_n = 0$ w.p.1.. Let $P^{-1} = (\tilde{\lambda_1} \cdots \tilde{\lambda_n})'$. Then $Y_n = \tilde{\lambda_n}'\tilde{X}$, i.e. $\tilde{\lambda_n}'\tilde{X} = 0$ w.p.1.. Thus finishes the proof.

5.1 Test Statistic

Recall the notation in section 1.4 of Chapter 1 where

$$A_k = \{N_{k-1} + 1, \cdots, N_k\},\ 1 \leq k \leq r. \tag{5.3}$$

Set

$$\{e_{h1}, \cdots, e_{hr}\} = \{1, \cdots, r\}, 1 \leq h \leq r!. \tag{5.4}$$

1. Spearman's Rho

Let

$$S_h = \sum_{k=1}^{r}[\sum_{i \in A_{e_{hk}}} \frac{i}{n_k}][\sum_{i=N_{k-1}+1}^{N_k} \pi(i)]. \tag{5.5}$$

Under the hypothesis (5.1), similar to (1.13), we find that

$$d_{S_{rh}}(\{\pi\}, E) \equiv S_h = \sum_{k=1}^{r}[\sum_{i \in A_{e_{hk}}} \frac{i}{n_k}][\sum_{i=N_{k-1}+1}^{N_k} \pi(i)]. \tag{5.6}$$

Hence the critical region is $\{S_h < c\}$. Following the above approach, we have the linearized statistic

$$\begin{aligned} S_L(\tilde{\alpha}) &= \sum_{h=1}^{r!} \alpha_h S_h \\ &= \sum_{k=1}^{r}[\sum_{h=1}^{r!} \alpha_h \sum_{i \in A_{e_{hk}}} \frac{i}{n_k}][\sum_{i=N_{k-1}+1}^{N_k} \pi(i)] \\ &= \sum_{k=1}^{r} \alpha_k^* \pi_k^{(S)}, \end{aligned} \tag{5.7}$$

where $\tilde{\alpha^*} = (\alpha_1^*, \cdots, \alpha_r^*)'$ is some $r \times 1$ vector in R^r and

$$\pi_k^{(S)} = \sum_{i=N_{k-1}+1}^{N_k} \pi(i). \tag{5.8}$$

Since $\sum_{k=1}^{r} \pi_k^{(S)} = \sum_{i=1}^{N_r} \pi(i) = \frac{N_r(N_r+1)}{2}$,

$$\begin{aligned} S_L(\tilde{\alpha}) &= \sum_{k=1}^{r} \alpha_k^* \pi_k^{(S)} \\ &\equiv \sum_{k=1}^{r-1} \alpha_k^{**} \pi_k^{(S)}, \end{aligned} \tag{5.9}$$

where $\tilde{\alpha^{**}} = (\alpha_1^{**}, \cdots, \alpha_{r-1}^{**})'$ is some $(r-1) \times 1$ vector in R^{r-1}. Put

$$\tilde{\pi_S} = (\pi_1^{(S)} - E_{H_0}\pi_1^{(S)}, \cdots, \pi_{r-1}^{(S)} - E_{H_0}\pi_{r-1}^{(S)})'. \tag{5.10}$$

Then the normalization process yields

$$\begin{aligned} S_N(\tilde{\alpha}) &= \frac{S_L(\tilde{\alpha}) - E_{H_0}\{S_L(\tilde{\alpha})\}}{[Var_{H_0}\{S_L(\tilde{\alpha})\}]^{1/2}} \\ &\equiv \frac{(\tilde{\alpha^{**}})'\tilde{\pi_S}}{[(\tilde{\alpha^{**}})'\Sigma_S \tilde{\alpha^{**}}]^{1/2}}, \end{aligned} \tag{5.11}$$

where $\Sigma_S = Cov_{H_0}\{\tilde{\pi_S}\} \geq 0$.

Lemma (5.1) shows that $\Sigma_S > 0$. Therefore the minimization step leads to

$$\begin{aligned} S_M &= min_{\tilde{\alpha} \in R^{r!}} S_N(\tilde{\alpha}) \\ &\equiv min_{\tilde{\alpha^{**}} \in R^{r-1}} \frac{(\tilde{\alpha^{**}})'\tilde{\pi_S}}{[(\tilde{\alpha^{**}})'\Sigma_S \tilde{\alpha^{**}}]^{1/2}} \\ &= -\tilde{\pi_S}'\Sigma_S^{-1}\tilde{\pi_S}. \end{aligned} \tag{5.12}$$

The critical region under (5.2) is therefore $\{\tilde{\pi_S}'\Sigma_S^{-1}\tilde{\pi_S} > c\}$, where the $(r-1) \times (r-1)$ matrix $\Sigma_S = Cov_{H_0}\{\tilde{\pi_S}\}$ is given by

$$Cov_{H_0}(\pi_k^{(S)}, \pi_{k'}^{(S)}) = \begin{cases} n_k(N_r - n_k)(N_r+1)/12 & \text{if } k = k' \\ -n_k n_{k'}(N_r+1)/12 & \text{if } k \neq k'. \end{cases} \tag{5.13}$$

2. Kendall's Tau

Let

$$K_h = \sum_{1 \leq k < p \leq r} \{\sum_{i \in A_{e_{hk}}} \sum_{j \in A_{e_{hp}}} sgn(i-j)\}\{\sum_{i=N_{k-1}+1}^{N_k} \sum_{j=N_{p-1}+1}^{N_p} sgn[\pi(i) - \pi(j)]\}. \tag{5.14}$$

Under the hypothesis (5.1), similar to (1.14), we can find that

$$d_{K_{rh}}(\{\pi\}, E) \equiv K_h = \sum_{1 \leq k < p \leq r} \{\sum_{i \in A_{e_{hk}}} \sum_{j \in A_{e_{hp}}} sgn(i-j)\}\{\sum_{i=N_{k-1}+1}^{N_k} \sum_{j=N_{p-1}+1}^{N_p} sgn[\pi(i) - \pi(j)]\}. \tag{5.15}$$

Hence the critical region is $\{K_h < c\}$. Following the three steps, we have

$$\begin{aligned} K_L(\tilde{\alpha}) &= \sum_{1 \leq k < p \leq r} \{\sum_{h=1}^{r!} \alpha_h \sum_{i \in A_{e_{hk}}} \sum_{j \in A_{e_{hp}}} sgn(i-j)\}\{\sum_{i=N_{k-1}+1}^{N_k} \sum_{j=N_{p-1}+1}^{N_p} sgn[\pi(i) - \pi(j)]\} \\ &= \sum_{1 \leq k < p \leq r} \alpha_{kp}^* \pi_{kp}^{(K)}, \end{aligned} \tag{5.16}$$

where $\tilde{\alpha^*} = (\alpha^*_{12}, \cdots, \alpha^*_{1r}, \cdots, \alpha^*_{r-1\ r})'$ is some $\frac{r(r-1)}{2} \times 1$ vector in $R^{r(r-1)/2}$ and

$$\pi^{(K)}_{kp} = \sum_{i=N_{k-1}+1}^{N_k} \sum_{j=N_{p-1}+1}^{N_p} sgn[\pi(i) - \pi(j)]. \tag{5.17}$$

Let

$$\tilde{\pi_K} = (\pi^{(K)}_{12} - E_{H_0}\pi^{(K)}_{12}, \cdots, \pi^{(K)}_{1r} - E_{H_0}\pi^{(K)}_{1r}, \cdots, \pi^{(K)}_{r-1\ r} - E_{H_0}\pi^{(K)}_{r-1\ r})'. \tag{5.18}$$

Then

$$\begin{aligned} K_N(\tilde{\alpha}) &= \frac{K_L(\tilde{\alpha}) - E_{H_0}\{K_L(\tilde{\alpha})\}}{[Var_{H_0}\{K_L(\tilde{\alpha})\}]^{1/2}} \\ &= \frac{(\tilde{\alpha^*})'\tilde{\pi_K}}{[(\tilde{\alpha^*})'\Sigma_K\tilde{\alpha^*}]^{1/2}}, \end{aligned} \tag{5.19}$$

where $\Sigma_K = Cov_{H_0}\{\tilde{\pi_K}\} \geq 0$.

By Lemma (5.1), $\Sigma_K > 0$. Therefore

$$\begin{aligned} K_M &= min_{\tilde{\alpha}\in R^{r!}} K_N(\tilde{\alpha}) \\ &= min_{\tilde{\alpha^*}\in R^{r(r-1)/2}} \frac{(\tilde{\alpha^*})'\tilde{\pi_K}}{[(\tilde{\alpha^*})'\Sigma_K\tilde{\alpha^*}]^{1/2}} \\ &= -\tilde{\pi_K}'\Sigma_K^{-1}\tilde{\pi_K}. \end{aligned} \tag{5.20}$$

Hence the critical region under (5.2) is $\{\tilde{\pi_K}'\Sigma_K^{-1}\tilde{\pi_K} > c\}$, where the $\frac{r(r-1)}{2} \times \frac{r(r-1)}{2}$ matrix $\Sigma_K = Cov_{H_0}\{\tilde{\pi_K}\}$ is given by

$$Cov_{H_0}(\pi^{(K)}_{kp}, \pi^{(K)}_{k'p'}) = \begin{cases} n_k n_p(n_k + n_p + 1)/3 & \text{if } k = k', p = p' \\ n_k n_p n_{p'}/3 & \text{if } k = k', p \neq p' \\ -n_k n_{k'} n_{p'}/3 & \text{if } k < p = k' < p' \\ n_k n_{k'} n_p/3 & \text{if } k \neq k', p = p' \\ 0 & \text{otherwise.} \end{cases} \tag{5.21}$$

3. Spearman's Footrule

Let

$$F_h = \sum_{k=1}^{r} \sum_{i=N_{k-1}+1}^{N_k} \sum_{j\in A_{e_{hk}}} |\pi(i) - j|. \tag{5.22}$$

Under the hypothesis (5.1), similar to (1.15), we can find that

$$d_{F_{rh}}(\{\pi\}, E) \equiv F_h = \sum_{k=1}^{r} \sum_{i=N_{k-1}+1}^{N_k} \sum_{j\in A_{e_{hk}}} |\pi(i) - j|. \tag{5.23}$$

Hence the critical region is $\{F_h < c\}$. Following the above three steps, we have

$$F_L(\tilde{\alpha}) = \sum_{k=1}^{r} \sum_{i=N_{k-1}+1}^{N_k} \{\sum_{h=1}^{r!} \alpha_h \sum_{j\in A_{e_{hk}}} |\pi(i) - j|\}$$

$$
\begin{aligned}
&= \sum_{k=1}^{r} \sum_{i=N_{k-1}+1}^{N_k} \{ \sum_{p=1}^{r} \sum_{j=N_{p-1}+1}^{N_p} \alpha_{kp}^{*} |\pi(i) - j| \} \\
&= \sum_{k=1}^{r} \sum_{p=1}^{r} \alpha_{kp}^{*} \pi_{kp}^{(F)},
\end{aligned}
\tag{5.24}
$$

where $\tilde{\alpha^*} = (\alpha_{11}^*, \cdots, \alpha_{1r}^*, \cdots, \alpha_{r1}^*, \cdots, \alpha_{rr}^*)'$ is some $r^2 \times 1$ vector in R^{r^2} and

$$
\pi_{kp}^{(F)} = \sum_{i=N_{k-1}+1}^{N_k} \sum_{j=N_{p-1}+1}^{N_p} |\pi(i) - j|. \tag{5.25}
$$

Since, for any $p, 1 \leq p \leq r$, $\sum_{k=1}^{r} \pi_{kp}^{(F)} = \sum_{i=1}^{N_r} \sum_{j=N_{p-1}+1}^{N_p} |i - j|$ =a constant depending on sample sizes,

$$
\begin{aligned}
F_L &= \sum_{k=1}^{r} \sum_{p=1}^{r} \alpha_{kp}^{*} \pi_{kp}^{(F)} \\
&\equiv \sum_{k=1}^{r-1} \sum_{p=1}^{r} \alpha_{kp}^{**} \pi_{kp}^{(F)},
\end{aligned}
\tag{5.26}
$$

where $\tilde{\alpha^{**}} = (\alpha_{11}^{**}, \cdots, \alpha_{1r}^{**}, \cdots, \alpha_{r-1\ 1}^{**}, \cdots, \alpha_{r-1\ r}^{**})'$ is some $r(r-1) \times 1$ vector in $R^{r(r-1)}$. Put

$$
\tilde{\pi_F} = (\pi_{11}^{(F)} - E_{H_0}\pi_{11}^{(F)}, \cdots, \pi_{1r}^{(F)} - E_{H_0}\pi_{1r}^{(F)}, \cdots, \pi_{r-1\ 1}^{(F)} - E_{H_0}\pi_{r-1\ 1}^{(F)}, \cdots, \pi_{r-1\ r}^{(F)} - E_{H_0}\pi_{r-1\ r}^{(F)})'. \tag{5.27}
$$

Then

$$
\begin{aligned}
F_N(\tilde{\alpha}) &= \frac{F_L(\tilde{\alpha}) - E_{H_0}\{F_L(\tilde{\alpha})\}}{[Var_{H_0}\{F_L(\tilde{\alpha})\}]^{1/2}} \\
&\equiv \frac{(\tilde{\alpha^{**}})'\tilde{\pi_F}}{[(\tilde{\alpha^{**}})'\Sigma_F \tilde{\alpha^{**}}]^{1/2}},
\end{aligned}
\tag{5.28}
$$

where $\Sigma_F = Cov_{H_0}\{\tilde{\pi_F}\} \geq 0$.

By Lemma (5.1), $\Sigma_F > 0$. Therefore

$$
\begin{aligned}
F_M &= min_{\tilde{\alpha} \in R^{r^2}} F_N(\tilde{\alpha}) \\
&\equiv min_{\tilde{\alpha^{**}} \in R^{r(r-1)}} \frac{(\tilde{\alpha^{**}})'\tilde{\pi_F}}{[(\tilde{\alpha^{**}})'\Sigma_F \tilde{\alpha^{**}}]^{1/2}} \\
&= -\tilde{\pi_F}'\Sigma_F^{-1}\tilde{\pi_F}.
\end{aligned}
\tag{5.29}
$$

Hence the critical region under (5.2) is $\{\tilde{\pi_F}'\Sigma_F^{-1}\tilde{\pi_F} > c\}$, where the $[r(r-1)] \times [r(r-1)]$ matrix $\Sigma_F = Cov_{H_0}\{\tilde{\pi_F}\}$ is given by

$$
Cov_{H_0}(\pi_{kp}^{(F)}, \pi_{k'p'}^{(F)})
$$

$$
= \begin{cases} \frac{n_k(N_r - n_k)}{N_r(N_r-1)} \sum_{j=N_{p-1}+1}^{N_p} \sum_{j_1=N_{p'-1}+1}^{N_{p'}} \sum_{s=1}^{N_r} [|s-j| - \sum_{i=1}^{N_r} \frac{|i-j|}{N_r}][|s-j_1| - \sum_{i=1}^{N_r} \frac{|i-j_1|}{N_r}] & \text{if } k = k' \\ \frac{-n_k n_{k'}}{N_r(N_r-1)} \sum_{j=N_{p-1}+1}^{N_p} \sum_{j_1=N_{p'-1}+1}^{N_{p'}} \sum_{s=1}^{N_r} [|s-j| - \sum_{i=1}^{N_r} \frac{|i-j|}{N_r}][|s-j_1| - \sum_{i=1}^{N_r} \frac{|i-j_1|}{N_r}] & \text{if } k < k'. \end{cases}
\tag{5.30}
$$

4. Hamming

For $1 \le k, p \le r$, let

$$\pi_{kp}^{(H)} = the\ cardinality\ of\ \{\pi(N_{k-1}+1), \cdots, \pi(N_k)\} \bigcap A_p. \tag{5.31}$$

Set

$$H_h = \sum_{k=1}^{r} \frac{-\pi_{ke_{hk}}^{(H)}}{n_k}. \tag{5.32}$$

Under the hypothesis (5.1), similar to (1.16), we can find that

$$d_{H_{rh}}(\{\pi\}, E) \equiv H_h = \sum_{k=1}^{r} \frac{-\pi_{ke_{hk}}^{(H)}}{n_k}. \tag{5.33}$$

Hence the critical region is $\{H_h < c\}$. Following the above three steps, we have

$$\begin{aligned} H_L(\tilde{\alpha}) &= \sum_{k=1}^{r}\{\sum_{h=1}^{r!} \alpha_h \frac{-\pi_{ke_{hk}}^{(H)}}{n_k}\} \\ &= \sum_{k=1}^{r}\sum_{p=1}^{r} \alpha_{kp}^{*} \pi_{kp}^{(H)}, \end{aligned} \tag{5.34}$$

where $\tilde{\alpha^*} = (\alpha_{11}^*, \cdots, \alpha_{1r}^*, \cdots, \alpha_{r1}^*, \cdots, \alpha_{rr}^*)'$ is some $r^2 \times 1$ vector in R^{r^2}. Since $\sum_{k=1}^{r} \pi_{kp}^{(H)} = n_p,\ for\ 1 \le p \le r$ and $\sum_{p=1}^{r} \pi_{kp}^{(H)} = n_k,\ for\ 1 \le k \le r$,

$$\begin{aligned} H_L &= \sum_{k=r}^{r}\sum_{p=1}^{r} \alpha_{kp}^{*} \pi_{kp}^{(H)} \\ &\equiv \sum_{k=r}^{r-1}\sum_{p=1}^{r-1} \alpha_{kp}^{**} \pi_{kp}^{(H)}, \end{aligned} \tag{5.35}$$

where $\tilde{\alpha^{**}} = (\alpha_{11}^{**}, \cdots, \alpha_{1\ r-1}^{**}, \cdots, \alpha_{r-1\ 1}^{**}, \cdots, \alpha_{r-1\ r-1}^{**})'$ is some $(r-1)^2 \times 1$ vector in $R^{(r-1)^2}$. Put

$$\begin{aligned} \tilde{\pi_H} = & (\pi_{11}^{(H)} - E_{H_0}\pi_{11}^{(H)}, \cdots, \pi_{1\ r-1}^{(H)} - E_{H_0}\pi_{1\ r-1}^{(H)}, \cdots, \\ & \pi_{r-1\ 1}^{(H)} - E_{H_0}\pi_{r-1\ 1}^{(H)}, \cdots, \pi_{r-1\ r-1}^{(H)} - E_{H_0}\pi_{r-1\ r-1}^{(H)})'. \end{aligned} \tag{5.36}$$

Then

$$\begin{aligned} H_N(\tilde{\alpha}) &= \frac{H_L(\tilde{\alpha}) - E_{H_0}\{H_L(\tilde{\alpha})\}}{[Var_{H_0}\{H_L(\tilde{\alpha})\}]^{1/2}} \\ &\equiv \frac{(\tilde{\alpha^{**}})'\tilde{\pi_H}}{[(\tilde{\alpha^{**}})'\Sigma_H \tilde{\alpha^{**}}]^{1/2}}, \end{aligned} \tag{5.37}$$

where $\Sigma_H = Cov_{H_0}\{\tilde{\pi_H}\} \ge 0$.

By Lemma (5.1), $\Sigma_H > 0$. Therefore

$$\begin{aligned} H_M &= min_{\tilde{\alpha} \in R^{r^2}} H_N(\tilde{\alpha}) \\ &\equiv min_{\tilde{\alpha^{**}} \in R^{(r-1)^2}} \frac{(\tilde{\alpha^{**}})' \tilde{\pi_H}}{[(\tilde{\alpha^{**}})' \Sigma_H \tilde{\alpha^{**}}]^{1/2}} \\ &= -\tilde{\pi_H}' \Sigma_H^{-1} \tilde{\pi_H}. \end{aligned} \tag{5.38}$$

Hence the critical region under (5.2) is $\{\tilde{\pi_H}' \Sigma_H^{-1} \tilde{\pi_H} > c\}$, where the $(r-1)^2 \times (r-1)^2$ matrix $\Sigma_H = Cov_{H_0}\{\tilde{\pi_H}\}$ is given by

$$Cov_{H_0}(\pi_{kp}^{(H)}, \pi_{k'p'}^{(H)}) = \begin{cases} \frac{n_k n_p (N_r - n_k)(N_r - n_p)}{N_r^2 (N_r - 1)} & \text{if } k = k', p = p' \\ \frac{-n_k n_p (N_r - n_k) n_{p'}}{N_r^2 (N_r - 1)} & \text{if } k = k', p \neq p' \\ \frac{-n_k n_{k'} n_p (N_r - n_p)}{N_r^2 (N_r - 1)} & \text{if } k \neq k', p = p' \\ \frac{n_k n_{k'} n_p n_{p'}}{N_r^2 (N_r - 1)} & \text{if } k \neq k', p \neq p'. \end{cases} \tag{5.39}$$

5.2 Asymptotic Distributions Under the Null Hypothesis

For the asymptotic distibutions of S_M, K_M, F_M and H_M under the null hypothesis, we need to find the corresponding asymptotic null distributions of $\tilde{\pi_S}, \tilde{\pi_K}, \tilde{\pi_F}$ and $\tilde{\pi_H}$.

1. Spearman's Rho

To prove the asymptotic normality of $\tilde{\pi_S}$ given by (5.10), consider, for any $\tilde{c_S} = (c_1, \cdots, c_{r-1})' \in R^{r-1}$, the linear combination

$$\begin{aligned} L_S &\stackrel{\text{def}}{=} \tilde{c_S}' \tilde{\pi_S} \\ &= \sum_{k=1}^{r-1} c_k (\pi_k^{(S)} - E_{H_0}\{\pi_k^{(S)}\}) \\ &= \sum_{k-1}^{r-1} c_k \sum_{i=N_{k-1}+1}^{N_k} [\pi(i) - E_{H_0} \pi_i] \\ &\equiv \sum_{k-1}^{r-1} \sum_{i=N_{k-1}+1}^{N_k} c_k \pi(i). \end{aligned} \tag{5.40}$$

Let

$$a_i = \begin{cases} c_k & \text{if } N_{k-1} < i \leq N_k, 1 \leq k \leq r-1 \\ 0 & \text{if } N_{r-1} < i \leq N_r. \end{cases} \tag{5.41}$$

Then

$$L_S \equiv \sum_{i=1}^{N_r} a_i \pi(i). \tag{5.42}$$

When $min\{n_1 \cdots n_r\} \to \infty$,

$$\bar{a} = \frac{1}{N_r} \sum_{i=1}^{N_r} a_i$$

$$
\begin{aligned}
&= \sum_{i=1}^{r-1} \frac{n_k}{N_r} c_k \\
&\to \sum_{i=1}^{r-1} t_k c_k \qquad (5.43)\\
max_{1\le i\le N_r}(a_i - \bar{a})^2 &= max_{1\le k\le r}(c_k - \bar{a})^2, \ with\ c_r = 0 \\
&\to max_{1\le k\le r}(c_k - \sum_{i=1}^{r-1} t_k c_k)^2 \qquad (5.44)\\
\sum_{i=1}^{N_r}(a_i - \bar{a})^2 &= \sum_{k=1}^{r} n_k (c_k - \bar{a})^2 \\
and &\approx N_r \sum_{k=1}^{r} t_k (c_k - \sum_{i=1}^{r-1} t_k c_k)^2. \qquad (5.45)
\end{aligned}
$$

Hence when $min(n_1 \cdots n_r) \to \infty$,

$$
\frac{\sum_{i=1}^{N_r}(a_i - \bar{a})^2}{max_{1\le i\le N_r}(a_i - \bar{a})^2} = \frac{\sum_{k=1}^{r} t_k (c_k - \sum_{k=1}^{r-1} t_k c_k)^2}{max_{1\le k\le r}(c_k - \sum_{1=1}^{r-1} t_k c_k)^2} N_r \to \infty. \qquad (5.46)
$$

It follows from Theorem 2.2 that, as $min(n_1 \cdots n_r) \to \infty$,

$$
L_S \overset{H_0}{\sim} N(0, \tilde{c_S}' \Sigma_S \tilde{c_S}), \qquad (5.47)
$$

i.e.

$$
\tilde{\pi_S} \overset{H_0}{\sim} N_{(r-1)}(0, \Sigma_S). \qquad (5.48)
$$

Immediately, under H_0,

$$
\lim_{min(n_1 \cdots n_r) \to \infty} \tilde{\pi_S}' \Sigma_S^{-1} \tilde{\pi_S} \overset{H_0}{\sim} \chi^2_{r-1}. \qquad (5.49)
$$

2. Kendall's Tau

For $\tilde{\pi_K}$ in (5.18), consider, for any $\tilde{c_K} = (c_{12}, \cdots, c_{1r}, \cdots, c_{r-1\ r})' \in R^{r(r-1)/2}$, the linear combination

$$
\begin{aligned}
L_K &\overset{\text{def}}{=} \tilde{c_K}' \tilde{\pi_K} \\
&\equiv \sum_{1\le k<p\le r} c_{kp} \pi_{kp}^{(K)}. \qquad (5.50)
\end{aligned}
$$

Set A_k be given in (5.3) for $1 \le k \le r$. For $i < j$, let

$$
c_{ij}^* = \begin{cases} c_{kp} & \text{if } i \in A_k \ and\ j \in A_p, k < p \\ 0 & \text{if } i, j \in A_k\ for\ some\ k,\ 1 \le k \le r, \end{cases}
$$

$$
and \qquad h(\pi(i), \pi(j)) = sgn[\pi(i) - \pi(j)].
$$

Then

$$
L_K = \sum_{1\le i<j\le N_r} c_{ij}^* h(\pi(i), \pi(j)).
$$

For $j > i$, put $c^*_{ij} = c^*_{ji}$ and $h(\pi(i), \pi(j)) = h(\pi(j), \pi(i))$. Thus

$$L_K = \frac{1}{2} \sum_{1 \le i \ne j \le N_r} c^*_{ij} h(\pi(i), \pi(j)).$$

It can be seen that $h(\pi(i), \pi(j))$ and $h(U_i, U_j)$ have the same distribution under H_0, where $U_1, \cdots, U_{N_r}$ are independent uniform distributions on $[0, 1]$. Therefore

$$L_K = \frac{1}{2} \sum_{1 \le i \ne j \le N_r} c^*_{ij} h(\pi(i), \pi(j))$$

$$\text{and} \qquad L^*_K = \frac{1}{2} \sum_{1 \le i \ne j \le N_r} c^*_{ij} h(U_i, U_j)$$

have the same distribution under H_0. On the other hand, direct computations give the following

$$Cov_{H_0}\{h(U_1, U_2),\ h(U_1, U_3)\} = 1/3,$$

$$Var_{H_0}\{h(U_1, U_2)\} = 1,$$

$$\sum_{1 \le i < j \le N_r} (c^*_{ij})^2 = 2 \sum_{1 \le k < p \le r} n_k n_p (c_{kp})^2 = O(N_r^2),$$

$$\text{and} \qquad \sum_{i=1}^{N_r} (\sum_{j=1}^{N_r} c^*_{ij})^2 = 4 \sum_{k=1}^{r-1} n_k (\sum_{p=k+1}^{r} n_p c_{kp})^2 = O(N_r^3).$$

It follows from Lemma 2.1 in Shapiro and Hubert (1979) that as $b \to \infty$,

$$L_K \overset{H_0}{\sim} N(0, \tilde{c_K}' \Sigma_K \tilde{c_K}), \tag{5.51}$$

i.e.

$$\tilde{\pi_K} \overset{H_0}{\sim} N_{(r(r-1)/2)}(0, \Sigma_K). \tag{5.52}$$

Immediately, under H_0,

$$\lim_{min(n_1 \cdots n_r) \to \infty} \tilde{\pi_K}' \Sigma_K^{-1} \tilde{\pi_K} \overset{H_0}{\sim} \chi^2_{r(r-1)/2}. \tag{5.53}$$

3. Spearman's Footrule

For $\tilde{\pi_F}$ in (5.27), consider, for any $\tilde{c_F} = (c_{11}, \cdots, c_{1r}, \cdots, c_{r-1\ 1}, \cdots, c_{r-1\ r})' \in R^{r(r-1)}$, the linear combination

$$\begin{aligned} L_F &\overset{\text{def}}{=} \tilde{c_F}' \tilde{\pi_F} \\ &= \sum_{k=1}^{r-1} \sum_{p=1}^{r} c_{kp} \pi_{kp}^{(F)} \\ &= \sum_{k=1}^{r-1} \sum_{p=1}^{r} c_{kp} \sum_{i=N_{k-1}+1}^{N_k} \sum_{s=N_{p-1}+1}^{N_p} [|\pi(i) - j| - E_{H_0}|\pi(i) - s|] \\ &\equiv \sum_{k=1}^{r-1} \sum_{i=N_{k-1}+1}^{N_k} \{\sum_{p=1}^{r} c_{kp} \sum_{s=N_{p-1}+1}^{N_p} |\pi(i) - s|\} \\ &= \sum_{i=1}^{N_r} a_{i\pi(i)}, \end{aligned} \tag{5.54}$$

where

$$a_{ij} = \begin{cases} \sum_{p=1}^{r} c_{kp} \sum_{s=N_{p-1}+1}^{N_p} |j-s| & \text{if } N_{k-1} < i \le N_k, 1 \le k \le r-1 \\ 0 & \text{if } N_{r-1} < i \le N_r. \end{cases} \tag{5.55}$$

Since

$$\begin{aligned} \sum_{s=N_{p-1}+1}^{N_p} \sum_{j=1}^{N_r} |j-s| &= \sum_{s=N_{p-1}+1}^{N_p} \{[(s-1)+\cdots+1]+[1+\cdots+(N_r-s)]\} \\ &= \sum_{s=N_{p-1}+1}^{N_p} \{s^2-(N_r+1)s+\frac{N_r(N_r+1)}{2}\} \\ &\approx \frac{1}{3}(N_p^3-N_{p-1}^3)-\frac{N_r+1}{2}(N_p^2-N_{p-1}^2)+\frac{n_pN_r(N_r+1)}{2} \\ &\approx N_r^3\{\frac{1}{3}(T_p^3-T_{p-1}^3)-\frac{1}{2}(T_p^2-T_{p-1}^2)+\frac{t_p}{2}\} \\ &= N_r^3\beta_p, \end{aligned} \tag{5.56}$$

where $\beta_p = \frac{1}{3}(T_p^3-T_{p-1}^3)-\frac{1}{2}(T_p^2-T_{p-1}^2)+\frac{t_p}{2}$, we have

$$\begin{aligned} \bar{a_{i\cdot}} &= \frac{1}{N_r}\sum_{j=1}^{N_r} a_{ij} \\ &= \begin{cases} \frac{1}{N_r}\sum_{p=1}^{r} c_{kp} \sum_{s=N_{p-1}+1}^{N_p} \sum_{j=1}^{N_r} |j-s| & \text{if } N_{k-1} < i \le N_k, 1 \le k \le r-1 \\ 0 & \text{if } N_{r-1} < i \le N_r \end{cases} \\ &\approx \begin{cases} N_r^2 \sum_{p=1}^{r} c_{kp}\beta_p & \text{if } N_{k-1} < i \le N_k, 1 \le k \le r-1 \\ 0 & \text{if } N_{r-1} < i \le N_r. \end{cases} \end{aligned} \tag{5.57}$$

$$\begin{aligned} \bar{a_{\cdot j}} &= \frac{1}{N_r}\sum_{i=1}^{N_r} a_{ij} \\ &= \frac{1}{N_r}\sum_{k=1}^{r-1}\sum_{p=1}^{r} c_{kp} n_k \sum_{s=N_{p-1}+1}^{N_p} |j-s| \\ &\approx \sum_{k=1}^{r-1}\sum_{p=1}^{r} c_{kp} t_k \sum_{s=N_{p-1}+1}^{N_p} |j-s|. \end{aligned} \tag{5.58}$$

$$\begin{aligned} \bar{a_{\cdot\cdot}} &= \frac{1}{N_r^2}\sum_{i,j=1}^{N_r} a_{ij} \\ &= \frac{1}{N_r^2}\sum_{k=1}^{r-1}\sum_{p=1}^{r} c_{kp} n_k \sum_{s=N_{p-1}+1}^{N_p} \sum_{j=1}^{N_r} |j-s| \end{aligned} \tag{5.59}$$

$$\approx N_r^2 \sum_{k=1}^{r-1}\sum_{p=1}^{r} c_{kp} t_k \beta_p. \tag{5.60}$$

Then

$$
\begin{aligned}
d_{ij} &\stackrel{\text{def}}{=} a_{ij} - \bar{a_{i\cdot}} - \bar{a_{\cdot j}} + \bar{a_{\cdot\cdot}} \\
&\approx \begin{cases} N_r^2 \sum_{p=1}^{r}(c_{kp} - \sum_{k=1}^{r} t_k c_{kp})(\sum_{s=N_{p-1}+1}^{N_p} \frac{|j-s|}{N_r^2} - \beta_p) & \text{if } N_{k-1} < i \le N_k, 1 \le k \le r-1 \\ 0 & \text{if } N_{r-1} < i \le N_r. \end{cases}
\end{aligned}
\tag{5.61}
$$

Since

$$
\begin{aligned}
\sum_{s=N_{p-1}+1}^{N_p} |j-s| &= \begin{cases} \frac{n_p(N_p+N_{p-1}+1)}{2} - n_p j & \text{if } 0 < j \le N_{p-1} \\ \frac{(j-N_{p-1})(j-N_{p-1}+1)}{2} + \frac{(N_p-j)(N_p-j+1)}{2} & \text{if } N_{p-1} < j \le N_p \\ n_p j - \frac{n_p(N_p+N_{p-1}+1)}{2} & \text{if } N_p < j \le N_r \end{cases} \\
&\approx O(N_r^2),
\end{aligned}
\tag{5.62}
$$

it is easy to see that

$$
max_{1 \le i,s \le N_r} d_{ij}^2 \approx O(N_r^4) \tag{5.63}
$$

$$
\frac{1}{N_r} \sum_{i,j=1}^{N_r} d_{ij}^2 \approx N_r^5 M, \; M \neq 0. \tag{5.64}
$$

Hence

$$
\lim_{min(n_1 \cdots n_r) \to \infty} \frac{\frac{1}{N_r} \sum_{i,j=1}^{N_r} d_{ij}^2}{max_{1 \le i,j \le N_r} d_{ij}^2} = \infty. \tag{5.65}
$$

From Theorem 2.1, as $min(n_1 \cdots n_r) \to \infty$,

$$
L_F \stackrel{H_0}{\sim} N(0, \bar{c_F}' \Sigma_F \bar{c_F}), \tag{5.66}
$$

i.e.

$$
\bar{\pi_F} \stackrel{H_0}{\sim} N_{r(r-1)}(0, \Sigma_F). \tag{5.67}
$$

Therefore, under H_0,

$$
\lim_{min(n_1 \cdots n_r) \to \infty} \bar{\pi_F}' \Sigma_F^{-1} \bar{\pi_F} \stackrel{H_0}{\sim} \chi^2_{r(r-1)}. \tag{5.68}
$$

4. Hamming

For $\bar{\pi_H}$ in (5.36), consider, for any $\bar{c_H} = (c_{11}, \cdots, c_{1\ r-1}, \cdots, c_{r-1\ 1}, \cdots, c_{r-1\ r-1})' \in R^{(r-1)^2}$, the linear combination

$$
L_H \stackrel{\text{def}}{=} \bar{c_H}' \bar{\pi_H} = \sum_{k=1}^{r-1} \sum_{p=1}^{r-1} c_{kp} \pi_{kp}^{(H)}. \tag{5.69}
$$

Let

$$
a_{ij} = \begin{cases} 1 & \text{if } j \in \{N_{i-1}, \cdots, N_i\} \\ 0 & \text{otherwise.} \end{cases} \tag{5.70}
$$

Then

$$\pi_{kp}^{(H)} = \sum_{i=N_{k-1}+1}^{N_k} a_{p\pi(i)}. \tag{5.71}$$

Therefore

$$\begin{aligned} L_H &\stackrel{\text{def}}{=} \sum_{k=1}^{r-1}\sum_{p=1}^{r-1} c_{kp}\pi_{kp}^{(H)} \\ &= \sum_{k=1}^{r-1}\sum_{p=1}^{r-1} c_{kp} \sum_{i=N_{k-1}+1}^{N_k} a_{p\pi(i)} \\ &= \sum_{k=1}^{r-1}\sum_{i=N_{k-1}+1}^{N_k}\sum_{p=1}^{r-1} c_{kp} a_{p\pi(i)} \\ &= \sum_{i=1}^{N_r} a_{i\pi(i)}^{*}, \end{aligned} \tag{5.72}$$

where

$$a_{ij}^{*} = \begin{cases} \sum_{p=1}^{r-1} c_{kp}a_{pj} & \text{if } N_{k-1} < i \le N_k, 1 \le k \le r-1 \\ 0 & \text{if } N_{p-1} < i \le N_r. \end{cases} \tag{5.73}$$

Thus, we have

$$\begin{aligned} \bar{a}_{i\cdot}^{*} &= \frac{1}{N_r}\sum_{j=1}^{N_r} a_{ij}^{*} \\ &= \begin{cases} \frac{1}{N_r}\sum_{p=1}^{r-1} c_{kp}n_p & \text{if } N_{k-1} < i \le N_k, 1 < k \le r-1 \\ 0 & \text{if } k = r \end{cases} \\ &\approx \begin{cases} \sum_{p=1}^{r-1} c_{kp}t_p & \text{if } N_{k-1} < i \le N_k, 1 < k \le r-1 \\ 0 & \text{if } N_{r-1} < i \le N_r, \end{cases} \end{aligned} \tag{5.74}$$

$$\begin{aligned} \bar{a}_{\cdot j}^{*} &= \frac{1}{N_r}\sum_{i=1}^{N_r} a_{ij}^{*} \\ &= \frac{1}{N_r}\sum_{k=1}^{r-1}\sum_{i=N_{k-1}+1}^{N_k}\sum_{p=1}^{r-1} c_{kp}a_{pj} \\ &= \frac{1}{N_r}\sum_{k=1}^{r-1}\sum_{p=1}^{r-1} n_k c_{kp}a_{pj} \\ &\approx \sum_{k=1}^{r-1}\sum_{p=1}^{r-1} t_k c_{kp}a_{pj}, \end{aligned} \tag{5.75}$$

$$\bar{a}_{\cdot\cdot}^{*} = \frac{1}{N_r^2}\sum_{i,j=1}^{N_r} a_{ij}^{*}$$

$$= \frac{1}{N_r^2}\sum_{k=1}^{r-1}\sum_{p=1}^{r-1} c_{kp} n_k n_p$$

$$\approx \sum_{k=1}^{r-1}\sum_{p=1}^{r-1} c_{kp} t_k t_p. \tag{5.76}$$

Therefore it is easy to see that

$$max_{1\leq i,j\leq N_r} d_{ij}^2 \approx O(1) \tag{5.77}$$

$$\frac{1}{N_r}\sum_{i,j=1}^{N_r} d_{ij}^2 \approx N_r M, M \neq 0. \tag{5.78}$$

Hence

$$\lim_{min(n_1\cdots n_r)\to\infty} \frac{\frac{1}{N_r}\sum_{i,j=1}^{N_r} d_{ij}^2}{max_{1\leq i,j\leq N_r} d_{ij}^2} \approx \frac{N_r M}{O(1)} \to \infty. \tag{5.79}$$

By Theorem 2.1, as $min(n_1\cdots n_r)\to\infty$,

$$L_H \overset{H_0}{\sim} N(0, \tilde{c_H}'\Sigma_H\tilde{c_H}), \tag{5.80}$$

i.e.

$$\tilde{\pi_H} \overset{H_0}{\sim} N_{((r-1)^2)}(0, \Sigma_H). \tag{5.81}$$

Therefore, under H_0,

$$\lim_{min(n_1\cdots n_r)\to\infty} \tilde{\pi_H}'\Sigma_H^{-1}\tilde{\pi_H} \overset{H_0}{\sim} \chi^2_{(r-1)^2}. \tag{5.82}$$

5.3 Asymptotic Distributions Under the Alternative

For the asymptotic distributions of S_M, K_M, F_M and H_M under the alternative hypothesis, we need to find the corresponding asymptotic non null distributions of $\tilde{\pi_S}, \tilde{\pi_K}, \tilde{\pi_F}$ and $\tilde{\pi_H}$.

1. Spearman's Rho

To prove the asymptotic normality of $\tilde{\pi_S}$ given by (5.10) under H_1, consider L_S in (5.40). In view of (5.47), it can be shown that L_S satisfies the conditions in Theorem 3.6. Hence, as $min(n_1\cdots n_r)\to\infty$,

$$L_S = \tilde{c_S}'\tilde{\pi_S} \overset{H_1}{\sim} N(\tilde{c_S}'\tilde{\mu_S}, \tilde{c_S}'\Sigma_S\tilde{c_S}), \tag{5.83}$$

where

$$\tilde{\mu_S} = E_{H_0}\{\tilde{\pi_S}T_2\} = (E_{H_0}\{\pi_1^{(S)}T_2\}, \cdots, E_{H_0}\{\pi_{r-1}^{(S)}T_2\})'. \tag{5.84}$$

Since $\tilde{c_S}$ is an arbitary vector in R^{r-1},

$$\tilde{\pi_S} \overset{H_1}{\sim} N_{(r-1)}(\tilde{\mu_S}, \Sigma_S). \tag{5.85}$$

Immediately, under H_1,

$$\lim_{min(n_1 \cdots n_r) \to \infty} \tilde{\pi_S}' \Sigma_S^{-1} \tilde{\pi_S} \overset{H_1}{\sim} \chi^2_{\theta_S, r-1}, \tag{5.86}$$

where $\theta_S = \tilde{\mu_S}' \Sigma_S^{-1} \tilde{\mu_S}$.

2. Kendall's Tau

To prove the asymptotic normality of $\tilde{\pi_K}$ given by (5.18) under H_1, consider L_K in (5.50). For L_K^* given in section 5.2, define its projection denoted by L_K^{**} to be $L_K^{**} = \frac{1}{2} \sum_{1 \le i \ne j \le N_r} c_{ij}^* E_{H_0}\{h(U_i, U_j)|U_i\}$. Noting Theorem 3.3, Theorem 3.6 can apply for L_K^{**}. Therefore it follows from Lemma 2.1 in Shapiro and Hubert (1979), and the same arguments in section 5.2 that the asymptotic normality of L_K under H_1 can be established through L_K^{**}. Hence we obtain that as $min(n_1 \cdots n_r) \to \infty$,

$$L_K = \tilde{c_K}' \tilde{\pi_K} \overset{H_1}{\sim} N(\tilde{c_K}' \tilde{\mu_K}, \tilde{c_K}' \Sigma_K \tilde{c_K}), \tag{5.87}$$

where

$$\tilde{\mu_K} = E_{H_0}\{\tilde{\pi_K} T_2\} = (E_{H_0}\{\pi_{12}^{(K)} T_2\}, \cdots, E_{H_0}\{\pi_{1r}^{(K)} T_2\}, \cdots, E_{H_0}\{\pi_{r-1\ r}^{(K)} T_2\})'. \tag{5.88}$$

Since $\tilde{c_K}$ is an arbitary vector in $R^{r(r-1)/2}$,

$$\tilde{\pi_K} \overset{H_1}{\sim} N_{(r(r-1)/2)}(\tilde{\mu_K}, \Sigma_K). \tag{5.89}$$

Immediately, under H_1,

$$\lim_{min(n_1 \cdots n_r) \to \infty} \tilde{\pi_K}' \Sigma_K^{-1} \tilde{\pi_K} \overset{H_1}{\sim} \chi^2_{\theta_K, r(r-1)/2}, \tag{5.90}$$

where $\theta_K = \tilde{\mu_K}' \Sigma_K^{-1} \tilde{\mu_K}$.

3. Spearman's Footrule

To prove the asymptotic normality of $\tilde{\pi_F}$ given by (5.27) under H_1, consider L_F in (5.54). In view of (5.66), it can be shown using Theorem 3.6 that as $min(n_1 \cdots n_r) \to \infty$,

$$L_F = \tilde{c_F}' \tilde{\pi_F} \overset{H_1}{\sim} N(\tilde{c_F}' \tilde{\mu_F}, \tilde{c_F}' \Sigma_F \tilde{c_F}). \tag{5.91}$$

where

$$\tilde{\mu_F} = E_{H_0}\{\tilde{\pi_F} T_2\} = (E_{H_0}\{\pi_{11}^{(F)} T_2\}, \cdots, E_{H_0}\{\pi_{1r}^{(F)} T_2\}, \cdots, E_{H_0}\{\pi_{r-1\ 1}^{(F)} T_2\}, \cdots, E_{H_0}\{\pi_{r-1\ r}^{(F)} T_2\})'. \tag{5.92}$$

Since $\tilde{c_F}$ is an arbitary vector in $R^{r(r-1)}$,

$$\tilde{\pi_F} \overset{H_1}{\sim} N_{(r(r-1))}(\tilde{\mu_F}, \Sigma_F). \tag{5.93}$$

Immediately, under H_1,

$$\lim_{min(n_1 \cdots n_r) \to \infty} \tilde{\pi_F}' \Sigma_F^{-1} \tilde{\pi_F} \overset{H_1}{\sim} \chi^2_{\theta_F, r(r-1)}, \tag{5.94}$$

where $\theta_F = \tilde{\mu_F}'\Sigma_F^{-1}\tilde{\mu_F}$.

4. Hamming

To prove the asymptotic normality of $\tilde{\pi_H}$ given by (5.36) under H_1, consider L_H in (5.69). In view of (5.80), it can be shown using Theorem 3.6 that as $min(n_1 \cdots n_r) \to \infty$,

$$L_H = \tilde{c_H}'\tilde{\pi_H} \overset{H_1}{\sim} N(\tilde{c_H}'\tilde{\mu_H}, \tilde{c_H}'\Sigma_H\tilde{c_H}), \tag{5.95}$$

where

$$\begin{aligned}\tilde{\mu_H} = E_{H_0}\{\tilde{\pi_H}T_2\} &= (E_{H_0}\{\pi_{11}^{(H)}T_2\}, \cdots, E_{H_0}\{\pi_{1\ r-1}^{(H)}T_2\}, \cdots, \\ &\quad E_{H_0}\{\pi_{r-1\ 1}^{(H)}T_2\}, \cdots, E_{H_0}\{\pi_{r-1\ r-1}^{(H)}T_2\})'.\end{aligned} \tag{5.96}$$

Since $\tilde{c_H}$ is an arbitary vector in $R^{(r-1)^2}$,

$$\tilde{\pi_H} \overset{H_1}{\sim} N_{((r-1)^2)}(\tilde{\mu_H}, \Sigma_H). \tag{5.97}$$

Immediately, under H_1,

$$\lim_{min(n_1 \cdots n_r) \to \infty} \tilde{\pi_H}'\Sigma_H^{-1}\tilde{\pi_H} \overset{H_1}{\sim} \chi^2_{\theta_H,(r-1)^2}, \tag{5.98}$$

where $\theta_H = \tilde{\mu_H}'\Sigma_H^{-1}\tilde{\mu_H}$.

5.4 Asymptotic Powers and Efficiencies

For the multi-sample location problem, the test hypothesis (5.2) becomes

$$\begin{aligned}&H_0 : F_1(x) = \cdots = F_r(x) = F(x - \bar{d}) \ against \\ &H_1 : F_k(x) = F(x - \Delta_k), \ for \ 1 \le k \le r,\end{aligned} \tag{5.99}$$

where $\bar{d}$ is given by (3.6) and (3.7) in which there is no order for Δ_k's. This is referred to the two-sided hypothesis in the multi-sample. Assume that conditions (3.9) and (3.10) hold. Given the significance level α, let $\chi^2_{\alpha,n}$ be the critical value such that $P(\chi^2_n \ge \chi^2_{\alpha,n}) = \alpha$. Then the asymptotic powers of the Spearman's Rho, Kendall's Tau, Spearman's Footrule and Hamming statistics are given by

$$AP(S_M) = P(\chi^2_{r-1}(\theta_S) \ge \chi^2_{\alpha,r-1}) \tag{5.100}$$

$$AP(K_M) = P(\chi^2_{r(r-1)/2}(\theta_K) \ge \chi^2_{\alpha,r(r-1)/2}) \tag{5.101}$$

$$AP(F_M) = P(\chi^2_{r(r-1)}(\theta_F) \ge \chi^2_{\alpha,r(r-1)}) \tag{5.102}$$

$$AP(H_M) = P(\chi^2_{(r-1)^2}(\theta_H) \ge \chi^2_{\alpha,(r-1)^2}). \tag{5.103}$$

However, in terms of asymptotic power efficiency, although the test statistics we are considering are of the asymptotic χ^2-type, it is difficult to obtain the corresponding APE's because of the different degrees of freedom involved. However it follows from Theorem VI.4.6, Hajek and Sidak (1967) that when $\theta \to 0$,

$$\begin{aligned}&P(\chi^2_n(\theta) \ge \chi^2_{\alpha,n}) - \alpha \\ &\approx \theta\ 2^{-(n+2)/2}exp(-\frac{1}{2}\chi^2_{\alpha,n})[\Gamma(\frac{n+2}{2})]^{-1}(\chi^2_{\alpha,n})^{n/2}.\end{aligned} \tag{5.104}$$

Thus the asymptotic powers of the Spearman's Rho, Kendall's Tau, Spearman's Footrule and Hamming statistics can be approximated.

On the other hand, it is resulted from the Theorem VII.1.4, Hajek and Sidak (1967) that the maximum asymptotic power concerning the test hypothesis (5.2) can be reached by the test based on $\chi^2_{r-1}(b^2)$. Hence, for any χ^2-type statistic $\chi^2_n(\theta)$, we can denote its asymptotic power efficiency by employing the approximation (5.104) to be

$$\begin{aligned} APE(\chi^2_n(\theta)) &= \frac{P(\chi^2_n(\theta) \geq \chi^2_{\alpha,n}) - \alpha}{P(\chi^2_{r-1}(b^2) \geq \chi^2_{\alpha,r-1}) - \alpha} \\ &\approx 2^{(r-1-n)/2} exp(\frac{\chi^2_{\alpha,r-1} - \chi^2_{\alpha,n}}{2}) \frac{\Gamma(\frac{r+1}{2})}{\Gamma(\frac{n+2}{2})} \frac{(\chi^2_{\alpha,n})^{n/2}}{(\chi^2_{\alpha,r-1})^{(r-1)/2}} \frac{\theta}{b^2}. \end{aligned} \tag{5.105}$$

In our situations, it can be shown that the asymptotic power efficiencies can be approximated as follows:

$$APE(S_M) \approx \frac{\theta_S}{b^2} \tag{5.106}$$

$$\begin{aligned} APS(K_K) \approx\ & 2^{(-r^2+3r-2)/4} exp(\frac{\chi^2_{\alpha,r-1} - \chi^2_{\alpha,r(r-1)/2}}{2}) \\ & \frac{\Gamma(\frac{r+1}{2})}{\Gamma(\frac{r^2-r+4}{4})} \frac{(\chi^2_{\alpha,r(r-1)/2})^{r(r-1)/4}}{(\chi^2_{\alpha,r-1})^{(r-1)/2}} \frac{\theta_K}{b^2} \end{aligned} \tag{5.107}$$

$$\begin{aligned} APE(F_M) \approx\ & 2^{(-r^2+2r-1)/2} exp(\frac{\chi^2_{\alpha,r-1} - \chi^2_{\alpha,r(r-1)}}{2}) \\ & \frac{\Gamma(\frac{r+1}{2})}{\Gamma(\frac{r^2-r+2}{2})} \frac{(\chi^2_{\alpha,r(r-1)})^{r(r-1)/2}}{(\chi^2_{\alpha,r-1})^{(r-1)/2}} \frac{\theta_F}{b^2} \end{aligned} \tag{5.108}$$

$$\begin{aligned} APE(H_M) \approx\ & 2^{(-r^2+3r-2)/2} exp(\frac{\chi^2_{\alpha,r-1} - \chi^2_{\alpha,(r-1)^2}}{2}) \\ & \frac{\Gamma(\frac{r+1}{2})}{\Gamma(\frac{r^2-2r+3}{2})} \frac{(\chi^2_{\alpha,(r-1)^2})^{(r-1)^2/2}}{(\chi^2_{\alpha,r-1})^{(r-1)/2}} \frac{\theta_H}{b^2}. \end{aligned} \tag{5.109}$$

5.5 Computations of $\theta_S, \theta_K, \theta_F, \theta_H$

1. Computation of θ_S

For Σ_S in (5.13), it is easy to see that

$$\Sigma_S^{-1} = \frac{12}{N_r(N_r+1)}(diag\{1/n_1, \cdots, 1/n_r\} + J/n_r). \tag{5.110}$$

Immediately from (3.10) and (5.84), we have

$$\theta_S = \tilde{\mu_S}' \Sigma_S^{-1} \tilde{\mu_S} \approx 12 b^2 (\int_0^1 u\varphi(u,f)du)^2 / I(f). \tag{5.111}$$

2. Computation of θ_K

Let

$$\mu_{kp}^{(K)} = 2t_k t_p(\delta_k - \delta_p)\int_0^1 u\varphi(u,f)du, 1 \le k < p \le r \quad (5.112)$$

$$\text{and} \quad \nu_{kp,k'p'}^{(K)} = \begin{cases} t_k t_p(t_k + t_p)/3 & \text{if } k = k', p = p' \\ t_k t_p t_{p'}/3 & \text{if } k = k', p \neq p' \\ -t_k t_{k'} t_{p'}/3 & \text{if } k < k' = p < p' \\ t_k t_{k'} t_p/3 & \text{if } k \neq k', p = p' \\ 0 & \text{otherwise} \end{cases} \quad (5.113)$$

Then, by some calculations,

$$E_{H_0}\{\pi_{kp}^{(K)} T_2\} = E_{H_0}\{\sum_{i=N_{k-1}+1}^{N_k} \sum_{j=N_{p-1}+1}^{N_p} sgn[\mu(i) - \nu(j)] \sum_{i=1}^{N_r}(d_i - \bar{d})\varphi(\frac{\pi(i)}{N_r+1}, f)\}$$
$$\approx N_r^{3/2}\mu_{kp}^{(K)} \quad (5.114)$$
$$\text{and} \quad Cov_{H_0}(\pi_{kp}^{(K)}, \pi_{k'p'}^{(K)}) \approx N_r^3 \nu_{kp,k'p'}^{(K)}. \quad (5.115)$$

Put

$$\tilde{\mu_K^*} = (\mu_{12}^{(K)}, \cdots, \mu_{1r}^{(K)}, \cdots, \mu_{r-1\ r}^{(K)})' \quad (5.116)$$
$$\text{and} \quad \Sigma_K^* = (\nu_{kp,k'p'}^{(K)})_{\frac{r(r+1)}{2} \times \frac{r(r+1)}{2}}. \quad (5.117)$$

Then

$$\tilde{\mu_K} \approx N_r^{3/2}\tilde{\mu_K^*} \quad (5.118)$$
$$\text{and} \quad \Sigma_K \approx N_r^3 \Sigma_K^*. \quad (5.119)$$

Therefore

$$\theta_K = \tilde{\mu_K}'\Sigma_K^{-1}\tilde{\mu_K} \approx (\tilde{\mu_K^*})'(\Sigma_K^*)^{-1}\tilde{\mu_K^*}. \quad (5.120)$$

3. Computation of θ_F

For $1 \le k \le r-1, 1 \le p \le r$, let

$$\mu_{kp}^{(F)} = t_k(\delta_k - \sum_{k=1}^{r} t_k\delta_k)\{\int_0^{T_{p-1}}(\frac{T_p^2 - T_{p-1}^2}{2} - t_p u)\varphi(u,f)du +$$
$$\int_{T_{p-1}}^{T_p}[u^2 - (T_p + T_{p-1})u + \frac{T_p^2 + T_{p-1}^2}{2}]\varphi(u,f)du + \int_{T_p}^1 (t_p u - \frac{T_p^2 - T_{p-1}^2}{2})\varphi(u,f)du\}. \quad (5.121)$$

$$\text{and} \quad \nu_{kp,k'p'}^{(F)} = \begin{cases} t_k(1-t_k)(\gamma_p - \beta_p^2) & \text{if } k = k', p = p' \\ t_k(1-t_k)(\gamma_{pp'} - \beta_p\beta_{p'}) & \text{if } k = k', p < p' \\ -t_k t_{k'}(\gamma_p - \beta_p^2) & \text{if } k \neq k', p = p' \\ -t_k t_{k'}(\gamma_{pp'} - \beta_p\beta_{p'}) & \text{if } k \neq k', p < p', \end{cases} \quad (5.122)$$

where

$$\beta_p = \int_{T_{p-1}}^{T_p} (u^2 - u + 1/2)du \tag{5.123}$$

$$\begin{aligned}\gamma_p = & \int_0^{T_{p-1}} [\frac{t_p(T_p + T_{p-1})}{2} - t_p u]^2 du + \\ & \int_{T_{p-1}}^{T_p} [u^2 - (T_p + T_{p-1})u + \frac{T_p^2 + T_{p-1}^2}{2}]^2 du + \int_{T_p}^1 [t_p u - \frac{t_p(T_p + T_{p-1})}{2}]^2 du\end{aligned} \tag{5.124}$$

$$\begin{aligned}\gamma_{pp'} = & \int_0^{T_{p-1}} [\frac{t_p(T_p + T_{p-1})}{2} - t_p u][\frac{t_{p'}(T_{p'} + T_{p'-1})}{2} - t_{p'} u]du + \\ & \int_{T_{p-1}}^{T_p} [u^2 - (T_p + T_{p-1})u + \frac{T_p^2 + T_{p-1}^2}{2}][\frac{t_{p'}(T_{p'} + T_{p'-1})}{2} - t_{p'} u]du + \\ & \int_{T_p}^{T_{p'-1}} [t_p u - \frac{t_p(T_p + T_{p-1})}{2}][\frac{t_{p'}(T_{p'} + T_{p'-1})}{2} - t_{p'} u]du + \\ & \int_{T_{p'-1}}^{T_{p'}} [t_p u - \frac{t_p(T_p + T_{p-1})}{2}][u^2 - (T_{p'} + T_{p'-1})u + \frac{T_{p'}^2 + T_{p'-1}^2}{2}]du + \\ & \int_{T_{p'}}^1 [t_p u - \frac{t_p(T_p + T_{p-1})}{2}][t_{p'} u - \frac{t_{p'}(T_{p'} + T_{p'-1})}{2}]du.\end{aligned} \tag{5.125}$$

Then, by some calculations,

$$\begin{aligned}E_{H_0}\{\pi_{kp}^{(F)} T_2\} &= E_{H_0}\{\sum_{i=N_{k-1}+1}^{N_k} \sum_{j=N_{p-1}+1}^{N_p} |\pi(i) - j| \sum_{i=1}^{N_r} (d_i - \bar{d})\varphi(\frac{\pi(i)}{N_r + 1}, f)\} \\ &\approx N_r^{5/2} \mu_{kp}^{(F)} &(5.126) \\ and \quad Cov_{H_0}(\pi_{kp}^{(F)}, \pi_{k'p'}^{(F)}) &\approx N_r^5 \nu_{kp,k'p'}^{(F)}. &(5.127)\end{aligned}$$

Put

$$\begin{aligned}\tilde{\mu_F^*} &= (\mu_{12}^{(F)}, \cdots, \mu_{1r}^{(F)}, \cdots, \mu_{r-1\ r}^{(K)}, \cdots, \mu_{r-1\ r}^{(F)})' &(5.128) \\ and \quad \Sigma_K^* &= (\nu_{kp,k'p'}^{(F)})_{r(r-1)\times r(r-1)}. &(5.129)\end{aligned}$$

Then

$$\begin{aligned}\tilde{\mu_F} &\approx N_r^{5/2} \tilde{\mu_F^*} &(5.130) \\ and \quad \Sigma_F &\approx N_r^5 \Sigma_F^*. &(5.131)\end{aligned}$$

Therefore

$$\theta_F = \tilde{\mu_F}' \Sigma_F^{-1} \tilde{\mu_F} \approx (\tilde{\mu_F^*})'(\Sigma_F^*)^{-1} \tilde{\mu_F^*}. \tag{5.132}$$

4. Computation of θ_H

Let

$$\mu_{kp}^{(H)} = t_k(\delta_k - \sum_{k=1}^{r} t_k \delta_k) \int_{T_{p-1}}^{T_p} \varphi(u, f)du, 1 \le k, p \le r - 1 \tag{5.133}$$

$$\text{and} \quad \nu^{(H)}_{kp,k'p'} = \begin{cases} t_k t_p(1-t_k)(1-t_p) & \text{if } k=k', p=p' \\ -t_k t_p t_{p'}(1-t_k) & \text{if } k=k', p\neq p' \\ -t_k t_{k'} t_p(1-t_p) & \text{if } k\neq k', p=p' \\ t_k t_{k'} t_p t_{p'} & \text{if } k\neq k', p\neq p'. \end{cases} \quad (5.134)$$

Then, by some calculations,

$$E_{H_0}\{\pi^{(H)}_{kp} T_2\} = E_{H_0}\{\sum_{i=N_{k-1}+1}^{N_k} a_{p\pi(i)} \sum_{i=1}^{N_r}(d_i - \bar{d})\varphi(\frac{\pi(i)}{N_r+1}, f)\}$$

$$\approx N_r^{1/2}\mu^{(H)}_{kp} \quad (5.135)$$

$$\text{and} \quad Cov_{H_0}(\pi^{(H)}_{kp}, \pi^{(H)}_{k'p'}) \approx N_r \nu^{(H)}_{kp,k'p'}. \quad (5.136)$$

Put

$$\tilde{\mu^*_H} = (\mu^{(H)}_{12}, \cdots, \mu^{(H)}_{1\ r-1}, \cdots, \mu^{(H)}_{r-1\ r-1}, \cdots, \mu^{(H)}_{r-1\ r-1})' \quad (5.137)$$

$$\text{and} \quad \Sigma^*_H = (\nu^{(H)}_{kp,k'p'})_{(r-1)^2\times(r-1)^2}. \quad (5.138)$$

Then

$$\tilde{\mu_H} \approx N_r^{1/2}\tilde{\mu^*_H} \quad (5.139)$$

$$\text{and} \quad \Sigma_H \approx N_r \Sigma^*_H. \quad (5.140)$$

Therefore

$$\theta_H = \tilde{\mu_H}'\Sigma_H^{-1}\tilde{\mu_H} \approx (\tilde{\mu^*_H})'(\Sigma^*_H)^{-1}\tilde{\mu^*_H}. \quad (5.141)$$

5.6 Examples

In this section, we compare the test statistics in the two-sample location case with unordered alternative. If $r=2$, (3.10) ensures that

$$b^2 = I(f)t_1t_2(\delta_1-\delta_2)^2. \quad (5.142)$$

By (5.112), (5.113), (5.116), (5.117) and (5.120), it is easy to see that

$$\theta_K \approx 12b^2[\int_0^1 u\varphi(u,f)du]^2/I(f). \quad (5.143)$$

Then (5.106), (5.107), (5.111) and (5.143) show that

$$APE(S_M) \approx APE(K_M) \approx 12b^2[\int_0^1 u\varphi(u,f)du]^2/I(f). \quad (5.144)$$

Therefore the asymptotic power efficiencies of the Spearman's Rho and Kendall's Tau statistics are the same in the two-sample location. As a matter of the fact, the two statistics are equivalent by (5.12) and (5.20). Note that $APE(S_M)$ and $APE(K_M)$ do not depend on the significance level.

For θ_F, it follows from (5.121), (5.122), (5.128), (5.129) and (5.132) that

$$\theta_F \approx \frac{b^2}{I(f)}\tilde{w}'\begin{pmatrix} \gamma_1-\beta_1^2 & \gamma_{12}-\beta_1\beta_2 \\ \gamma_{12}-\beta_1\beta_2 & \gamma_2-\beta_2^2 \end{pmatrix}^{-1}\tilde{w}, \quad (5.145)$$

where

$$\bar{w} = \begin{pmatrix} \int_0^{T_1}(u^2 - t_1 u + \frac{t_1^2}{2})\varphi(u,f)du + \int_{T_1}^1 (t_1 u - \frac{T_1^2}{2})\varphi(u,f)du \\ \int_0^{T_1}(\frac{1-T_1^2}{2} - t_2 u)\varphi(u,f)du + \int_{T_1}^1 [u^2 - (1+T_1)u + \frac{1+T_1^2}{2}]\varphi(u,f)du \end{pmatrix}. \tag{5.146}$$

Let $\alpha = 0.05$. We have from (5.108) and (5.146) that

$$\begin{aligned} APE(F_M) &\approx 2^{(-2^2+2\times 2-1)/2} exp(\frac{\chi^2_{0.05,1} - \chi^2_{0.05,2}}{2}) \frac{\Gamma(\frac{3}{2})}{\Gamma(2)} \frac{\chi^2_{0.05,2}}{(\chi^2_{0.05,1})^{1/2}} \frac{\theta_F}{b^2} \\ &\approx 2^{-1/2} exp(\frac{3.841 - 5.991}{2}) \frac{0.8862}{1} \frac{5.991}{(3.841)^{1/2}} \frac{\theta_F}{b^2} \\ &\approx 0.6538 \times I^{-1}(f) \bar{w}' \begin{pmatrix} \gamma_1 - \beta_1^2 & \gamma_{12} - \beta_1\beta_2 \\ \gamma_{12} - \beta_1\beta_2 & \gamma_2 - \beta_2^2 \end{pmatrix}^{-1} \bar{w}. \end{aligned} \tag{5.147}$$

Similarly, we have

$$APE(H_M) \approx \frac{[\int_0^{T_1} \varphi(u,f)du]^2}{t_1 t_2 I(f)}. \tag{5.148}$$

Note that $APE(H_M)$ does not depend on the significance level in the two-sample case. Assume that $t_1 = t_2 = 0.5$. Then (5.123), (5.124) and (5.125) imply that $\gamma_1 = \gamma_2 = 3/80, \beta_1 = \beta_2 = 1/6, \gamma_{12} = 1/48$. Hence

$$\begin{pmatrix} \gamma_1 - \beta_1^2 & \gamma_{12} - \beta_1\beta_2 \\ \gamma_{12} - \beta_1\beta_2 & \gamma_2 - \beta_2^2 \end{pmatrix} \approx \begin{pmatrix} 0.0097 & -0.0069 \\ -0.0069 & 0.0097 \end{pmatrix}. \tag{5.149}$$

Based on the corresponding values in 4.2 and from (5.144), (5.147) and (5.148), it is easy to see that

if $f(x) = \frac{1}{(2\pi)^{1/2}\sigma} exp(\frac{-x^2}{2\sigma^2}), -\infty < x < \infty$, then $\bar{w} = (0.0857,\ -0.0857)'$ and

$$APE(S_M) \approx APE(K_M) \approx \frac{3}{\pi} \approx 0.9554 \tag{5.150}$$

$$APE(F_M) \approx 0.6538 \times 1 \times 0.8849 \approx 0.5626 \tag{5.151}$$

$$APE(H_M) \approx \frac{2}{\pi} \approx 0.6369. \tag{5.152}$$

If $f(x) = \frac{1}{2} exp(-|x|), -\infty < x < \infty$, then $\bar{w} = (0.0833,\ -0.0833)'$ and

$$APE(S_M) \approx APE(K_M) \approx \frac{3}{4} \approx 0.7500 \tag{5.153}$$

$$APE(F_M) \approx 0.6538 \times 1 \times 0.8360 \approx 0.5313 \tag{5.154}$$

$$APE(H_M) \approx \frac{4}{4} = 1. \tag{5.155}$$

If $f(x) = e^{-x}(1 + e^{-x})^{-2}, -\infty < x < \infty$, then $\bar{w} = (0.0521,\ -0.0521)'$ and

$$APE(S_M) \approx APE(K_M) \approx \frac{36}{36} = 1 \tag{5.156}$$

$$APE(F_M) \approx 0.6538 \times 3 \times 0.3270 \approx 0.6237 \tag{5.157}$$

$$APE(H_M) \approx \frac{3}{4} \approx 0.7500. \tag{5.158}$$

Bibliography

[1] Alvo, M. and Cabilio, P. (1991). On the Balanced Incomplete Block Design for Rankings. *The Annals of Statistics, Vol.* **19**, *No. 3, 1597-1613.*

[2] Alvo, M., Cabilio, P. and Feigin, P. (1982). Asymptotic Theory for Measure of Concordance with Special Reference to Average Kendall Tau. *The Annals of Statis' 'cs, Vol.* **10**, *1269-1276.*

[3] Critchlow, Douglas E. (1985). Metric Method for Analyzing Partially Ranked Data. *Lecture Notes in Statistics.*, Springer-Varleg, Berlin.

[4] Critchlow, Douglas E. (1986). A Unified Approach to Constructing Nonparametric Rank Tests. *Technical Report No.* **86-15**, Purdue University, USA.

[5] Diaconis, P. and Graham, P.L. (1977). Spearman's Footrule As a Measure of Disarray. *Journal of the Royal Statistical Society, B, Vol.* **39**, *262-268.*

[6] Feigin, Paul D. and Alvo, Mayer (1986). Intergroup Diversity and Concordance for Ranking Data: an Approach via Metrics for Permutations. *The Annals of Statistics, Vol.* **14**, *No. 2, 691-707.*

[7] Feller, William (1950). *An Introduction to Probability Theory and Its Application*, Volume I, 3rd edition. John Wiley & Sons Inc.

[8] Hajek, J. and Sidak, Z. (1967). *Theory of Rank Tests.* Academic Press, New York.

[9] Hoeffding, W. (1951). A Combinatorial Central Limit Theorem. *Ann. Math. Statist., Vol.* **22**, *558-566.*

[10] Jonckheere, A.R. (1954). A Distribution-free k Sample Test against Ordered Alternatives. *Biometrika* **41**, *131-145.*

[11] Page, E.B. (1963). Ordered Hypotheses for Multiple Treatments: A Significance Test for Linear Ranks. *J. Amer. Statist. Assoc., Vol.* **58**, *216-230.*

[12] Randles, Ronald H. and Wolfe, Douglas A. (1979). *Introduction to The Theory of Nonparametric Statistics.* John Wiley & Sons, World Publishing Corporation.

[13] Serfling, Robert J. (1980). Approximation Theorems of Mathematical Statistics. *John Wiley & Sons, Inc.*

[14] Shapiro, C. P. and Hubert, Lawrence (1979). Asymptotic Normality of Permutation Statistics Derived from Weighted Sums of Bivariate Functions. *The Annals of Statistics, Vol.* **7**, *No. 4, 788-794.*