ISSN 1683–5603

International Journal of Statistical Sciences Vol. 9(Special Issue), 2009, pp 97-112 © 2009 Dept. of Statistics, Univ. of Rajshahi, Bangladesh

Tests of Concordance between Groups of Incomplete Rankings

Jingyu Bu

Surveillance and Epidemiology Unit Cancer Care Nova Scotia Department of Health, Halifax, NS, Canada E-mail: vickey.bu@ccns.nshealth.ca

Paul Cabilio¹

Department of Mathematics and Statistics Acadia University, Wolfville, NS, Canada, B4P 2R6 E-mail: paul.cabilio@acadiau.ca

Ying Zhang

Department of Mathematics and Statistics Acadia University, Wolfville, NS, Canada E-mail: ying.zhang@acadiau.ca

[Received January 1, 2009; Accepted February 17, 2009]

Abstract

In many situations comparisons of two or more populations (treatments) are based on multivariate data collected from individuals in these populations. Such data vectors (blocks) may consist of measurements or of ranks. Here we address the problem of making comparisons in the presence of incomplete data, and develop test statistics for the case of two different population models. Such tests depend on measures of similarity between incomplete rankings developed in Alvo and Cabilio (1995), and are modelled on tests developed in Feigin and Alvo (1986) for the case of complete rankings. The properties of these tests are investigated, and two examples are presented to illustrate the methods.

Keywords and Phrases: Tests of Homogeneity, Spearman, Kendall and Measures of Similarity.

AMS Classification: 62G10; 62G20.

¹Corresponding Author

1 Introduction

Consider the situation in which t objects are ranked by m independent groups of judges, with each group consisting of n_{ℓ} judges, $\ell = 1, 2, ..., m$. Each ranking within a group of judges (blocks) is a random sample from a distribution characterized by a multinomial probability vector $\pi_{\ell} = (\pi_{\ell 1}, \pi_{\ell 2}, ..., \pi_{\ell t!})'$ whose components $\pi_{\ell j}$ are the probabilities of picking rankings indexed by j = 1, 2, ..., t!. It is wished to test the hypothesis of homogeneity of these distributions,

$$\mathcal{H}_0: \boldsymbol{\pi}_1 = \dots = \boldsymbol{\pi}_m \tag{1.1}$$

The case in which all rankings are complete has been considered by several authors, and in particular Feigin & Alvo (1986), who proposed a statistic that measures concordance between groups based on rank correlations.

In the following we consider the situation where it is wished to test the hypothesis (1.1) in the more general setting in which one or more of the rankings in each group are incomplete. The approach we consider is to generalize the Feigin & Alvo (1986) statistic to the incomplete case making use of the results first obtained in Alvo & Cabilio (1991) for the situation of testing the hypothesis that rankings assigned in a randomized incomplete block design have been randomly selected.

In what follows we will restrict ourselves to the situation where m = 2. In Section 2 we review some of the results in the case where all rankings are complete. Section 3 generalizes these results to the situation where the rankings are incomplete, and the results form the basis of tests for two different sample spaces as developed in section 4. Section 5 describes the computational procedures used for conducting such tests, and section 6 applies such methods to two different examples.

2 Complete Case

Rankings of t objects labeled 1 to t are permutations of (1, 2, ..., t), and for any two such permutations $\mu_j = (\mu_j(1), \mu_j(2), ..., \mu_j(t))'$, j = 1, 2, one may define measures of similarity $\mathcal{A}(\mu_1, \mu_2)$ which can be expressed as an inner product $\mathcal{A}(\mu_1, \mu_2) = t(\mu_1)' t(\mu_2)$, where $\mathbf{t} = t(\mu)$ is a column vector whose components are scores which characterize the ranking μ . For example, in the Spearman and Kendall cases such vectors are respectively

$$t_{S}(\boldsymbol{\mu}) = \left[\mu(i) - \frac{t+1}{2}\right]_{i}^{\prime}, 1 \le i \le t$$

$$t_{K}(\boldsymbol{\mu}) = \left[a(i,j)\right]_{(i,j)}^{\prime} = \left[sgn\left(\mu(j) - \mu(i)\right)\right]_{(i,j)}^{\prime}, 1 \le i < j \le {t \choose 2}$$
(2.1)

where sgn(x) is 1 or -1 depending on whether x > 0 or < 0, so that a(i, j) is the pairwise concordance of the object pair indexed by (i, j) with the natural ordering. The resulting measures of similarity are unstandardized rank correlations. The collection of score vectors

98

Bu, Cabilio and Zhang: Tests of Concordance between Groups

 $t(\boldsymbol{\mu})$, as $\boldsymbol{\mu}$ ranges over all its t! possible values, form the columns of a matrix \mathbf{T} , so that the $(t! \times t!)$ matrix $\mathbf{T'T}$ has components $\mathcal{A}(\boldsymbol{\mu}_1, \boldsymbol{\mu}_2)$, with $\boldsymbol{\mu}_1$ and $\boldsymbol{\mu}_2$ ranging over all permutations of $(1, 2, \ldots, t)$. For each sample of n_ℓ rankings, $\mathbf{f}_\ell = (f_{\ell 1}, f_{\ell 2}, \ldots, f_{\ell t!})'$, whose components $f_{\ell j}$ are the frequencies of observed rankings indexed by $j = 1, 2, \ldots, t!$, has a multinomial distribution with mean $n_\ell \pi_\ell$ and covariance matrix $n_\ell \boldsymbol{\Sigma}_\ell$, where $\boldsymbol{\Sigma}_\ell = \boldsymbol{\Pi}_\ell - \boldsymbol{\pi}_\ell \boldsymbol{\pi}'_\ell$ with $\boldsymbol{\Pi}_\ell = diag(\pi_{\ell 1}, \pi_{\ell 2}, \ldots, \pi_{\ell t!})$. Consequently, as $n_\ell \to \infty$

$$\sqrt{n_{\ell}}\left(\frac{\mathbf{f}_{\ell}}{n_{\ell}}-\boldsymbol{\pi}_{\ell}\right) \rightarrow_{\mathcal{L}} \mathcal{N}\left(0,\boldsymbol{\Sigma}_{\ell}\right),$$

and thus

$$\sqrt{n_{\ell}}\mathbf{T}\left(rac{\mathbf{f}_{\ell}}{n_{\ell}}-\boldsymbol{\pi}_{\ell}
ight)
ightarrow_{\mathcal{L}}\mathcal{N}\left(0,\mathbf{T}\boldsymbol{\Sigma}_{\ell}\mathbf{T}'
ight).$$

Letting $N = n_1 + n_2$, it follows that if $N \to \infty$ in such a way that $n_\ell/N \to \lambda_\ell > 0$,

$$\sqrt{N}\mathbf{T}\left(\frac{\mathbf{f}_{\ell}}{n_{\ell}}-\boldsymbol{\pi}_{\ell}\right) \rightarrow_{\mathcal{L}} \mathcal{N}\left(0,\lambda_{\ell}^{-1}\mathbf{T}\boldsymbol{\Sigma}_{\ell}\mathbf{T}'\right),$$

so that under the null hypothesis $\mathcal{H}_0: \boldsymbol{\pi}_1 = \boldsymbol{\pi}_2$,

$$\sqrt{N}\mathbf{T}\left(\left(\frac{\mathbf{f}_1}{n_1} - \boldsymbol{\pi}_1\right) - \left(\frac{\mathbf{f}_2}{n_2} - \boldsymbol{\pi}_2\right)\right) = \sqrt{N}\mathbf{T}\left(\frac{\mathbf{f}_1}{n_1} - \frac{\mathbf{f}_2}{n_2}\right) \to_{\mathcal{L}} \mathcal{N}\left(0, \mathbf{T}\boldsymbol{\Sigma}\mathbf{T}'\right), \quad (2.2)$$

where $\Sigma = (1/\lambda_1) \Sigma_1 + (1/\lambda_2) \Sigma_2$. Using this result, which also holds for the less restrictive null hypothesis \mathcal{H}'_0 : $\mathbf{T}\pi_1 = \mathbf{T}\pi_2$, Feigin & Alvo (1986) derived the test statistic

$$N\left(\frac{\mathbf{f}_1}{n_1} - \frac{\mathbf{f}_2}{n_2}\right)' \mathbf{T}' \left(\mathbf{T}\hat{\boldsymbol{\Sigma}}\mathbf{T}'\right)^{-} \mathbf{T}\left(\frac{\mathbf{f}_1}{n_1} - \frac{\mathbf{f}_2}{n_2}\right) \to_{\mathcal{L}} \chi_v^2,$$

where $\hat{\Sigma}$ is a suitable consistent estimate of Σ , v is the rank of $\mathbf{T}\Sigma\mathbf{T}'$, and $(\mathbf{T}\hat{\Sigma}\mathbf{T}')^{-}$ is the Moore-Penrose (generalized) inverse of $\mathbf{T}\hat{\Sigma}\mathbf{T}'$. An alternative statistic is given by

$$G = N\left(\mathbf{T}\left(\frac{\mathbf{f}_1}{n_1} - \frac{\mathbf{f}_2}{n_2}\right)\right)' \left(\mathbf{T}\left(\frac{\mathbf{f}_1}{n_1} - \frac{\mathbf{f}_2}{n_2}\right)\right) \to_{\mathcal{L}} \sum_{i=1}^{v} \alpha_i z_i^2$$
(2.3)

where the α_i 's are the eigenvalues of $\mathbf{T}\Sigma\mathbf{T}'$, and z_i are independent standard normal variates. In practice, the distribution of the statistic in (2.3) may be approximated by the distribution of $\sum_{i=1}^{v} \hat{\alpha}_i z_i^2$, where $\hat{\alpha}_i$'s are the eigenvalues of $\mathbf{T}\hat{\Sigma}\mathbf{T}'$. In the Spearman case

$$\left[\mathbf{T}_{S}\left(\frac{\mathbf{f}_{1}}{n_{1}}-\frac{\mathbf{f}_{2}}{n_{2}}\right)\right] = \left(\bar{\mu}^{(1)}(1)-\bar{\mu}^{(2)}(1),...,\bar{\mu}^{(1)}(t)-\bar{\mu}^{(2)}(t)\right)'$$

where $\bar{\mu}^{(\ell)}(i) = \frac{1}{n_{\ell}} \sum_{g=1}^{n_{\ell}} \mu_g^{(\ell)}(i)$, the mean of the ranks assigned to object i = 1, ..., t by the n_{ℓ} judges sampled from population $\ell = 1, 2$. The statistic (2.3) simplifies to

$$N\sum_{i=1}^{t} (\bar{\mu}^{(1)}(i) - \bar{\mu}^{(2)}(i))^2.$$
(2.4)

In the Kendall case,

$$\left[\mathbf{T}_{K}\left(\frac{\mathbf{f}_{1}}{n_{1}}-\frac{\mathbf{f}_{2}}{n_{2}}\right)\right] = \left(\left[\bar{a}^{(1)}(1,2)-\bar{a}^{(2)}(1,2)\right], ..., \left[\bar{a}^{(1)}(t-1,t)-\bar{a}^{(2)}(t-1,t)\right]\right)',$$

where $\bar{a}^{(\ell)}(i,j) = \frac{1}{n_{\ell}} \sum_{g=1}^{n_{\ell}} a_g^{(\ell)}(i,j)$, and $a_g^{(\ell)}(i,j)$, defined in (2.1), is the pairwise concordance score assigned to object pairs indexed by (i,j), $1 \le i < j \le {t \choose 2}$, by the g^{th} judge from population $\ell = 1, 2$. In this case, the statistic (2.3) simplifies to

$$N\sum_{i< j} \left(\bar{a}^{(1)}(i,j) - \bar{a}^{(2)}(i,j)\right)^2.$$
(2.5)

Under the null hypothesis $\Sigma_1 = \Sigma_2$, and a pooled estimate of Σ is

$$\hat{\Sigma}_P = \frac{N^2}{n_1 n_2} \frac{1}{N-2} ((n_1 - 1)\hat{\Sigma}_1 + (n_2 - 1)\hat{\Sigma}_2), \qquad (2.6)$$

where $\hat{\boldsymbol{\Sigma}}_{\ell} = \frac{1}{n_{\ell}-1} \left(\mathbf{F}_{\ell} - \mathbf{f}_{\ell} \mathbf{f}_{\ell}' / n_{\ell} \right)$, and $\mathbf{F}_{\ell} = diag(f_{\ell 1}, f_{\ell 2}, \dots, f_{\ell t!})$. In the Spearman case,

$$\mathbf{T}_{S}\hat{\boldsymbol{\Sigma}}_{\ell}\mathbf{T}_{S}^{\prime} = \frac{1}{n_{\ell}-1}\sum_{g=1}^{n_{\ell}} \left(\boldsymbol{\mu}_{g}^{(\ell)} - \bar{\boldsymbol{\mu}}^{(\ell)}\right) \left(\boldsymbol{\mu}_{g}^{(\ell)} - \bar{\boldsymbol{\mu}}^{(\ell)}\right)^{\prime}$$
(2.7)

where $\boldsymbol{\mu}_{g}^{(\ell)} = \left(\boldsymbol{\mu}_{g}^{(\ell)}(1), \boldsymbol{\mu}_{g}^{(\ell)}(2), ..., \boldsymbol{\mu}_{g}^{(\ell)}(t)\right)'$ is the *t*-vector of rankings assigned by judge *g* from population ℓ , and $\bar{\boldsymbol{\mu}}^{(\ell)} = \left(\bar{\boldsymbol{\mu}}^{(\ell)}(1), \bar{\boldsymbol{\mu}}^{(\ell)}(2), ..., \bar{\boldsymbol{\mu}}^{(\ell)}(t)\right)'$ is the vector of mean rankings. In the Kendall case, we take the elements of $t_{K}(\boldsymbol{\mu})$ to be ordered so that the r^{th} row element corresponds to (i, j) for which $r = (i - 1)(t - i/2) + (j - i), 1 \le i < j \le {t \choose 2}$. Using this ordering, we may write $\mathbf{a}_{g}^{(\ell)} = \left(a_{g}^{(\ell)}(1), a_{g}^{(\ell)}(2), ..., a_{g}^{(\ell)}(q)\right)'$, as the $q = {t \choose 2}$ dimensional vector of pairwise concordances assigned by judge *g* from population ℓ . With this notation, in the Kendall case,

$$\mathbf{T}_{K}\hat{\mathbf{\Sigma}}_{\ell}\mathbf{T}_{K}^{\prime} = \frac{1}{n_{\ell}-1}\sum_{g=1}^{n_{\ell}} \left(\mathbf{a}_{g}^{(\ell)} - \bar{\mathbf{a}}^{(\ell)}\right) \left(\mathbf{a}_{g}^{(\ell)} - \bar{\mathbf{a}}^{(\ell)}\right)^{\prime}.$$
(2.8)

3 Incomplete Case: Preliminaries

In the situation where one or more rankings in a group are incomplete, measures of similarity (or equivalently distance) between incomplete rankings may be defined using the methods developed in Alvo & Cabilio (1995, 1999). A complete ranking of t objects is compatible with an incomplete ranking of k of these objects if every pair of the specified k objects is given the same relative ranking in both rankings. An incomplete ranking of k out of a possible t objects, with a specific pattern of t - k missing objects can be denoted by $\mu^* = (\mu(o_1), \mu(o_2), \dots, \mu(o_k))'$, where $1 \le o_1 < o_2 < \dots < o_k \le t$ are the labels of the actual objects being ranked. For a specific pattern of missing observations, each of the k! possible rankings is associated with a set of t!/k! complete rankings which are compatible to it. For a given incomplete ranking μ_j^* , j = 1, 2, ..., k!, we define a t!-vector of compatibility c_j whose row numbers identify each of the t! possible complete rankings ordered in some way, with the element in each row being 1 or 0 depending on whether that complete ranking is or is not compatible with μ_i^* . Each such vector will contain t!/k! 1's, with the remaining elements 0. For a given pattern of missing observations, each of the possible k! incomplete rankings is associated with a different vector of compatibility, and for this specified pattern we define the $(t! \times k!)$ matrix of compatibility $\mathbf{C} = [\mathbf{c}_1, \mathbf{c}_2, ..., \mathbf{c}_{k!}]$, whose columns are orthogonal and sum up to $\mathbf{1}_{[t!]}$, the t! dimensional vector of 1's.

For a given multinomial distribution $\pi_{\ell} = (\pi_{\ell 1}, \pi_{\ell 2}, \dots, \pi_{\ell t !})'$ on the t! rankings in the complete case, the corresponding distribution over the incomplete rankings with a specified pattern of missing observations, is

$$m{\pi}_{\ell}^{*} = \mathbf{C}' m{\pi}_{\ell} = (\pi_{\ell 1}^{*}, \pi_{\ell 2}^{*}, \dots, \pi_{\ell k !}^{*})', \pi_{\ell s}^{*} = \mathbf{c}_{s}^{'} m{\pi}_{\ell} = \sum_{j \in c(s)} \pi_{\ell j},$$

where c(s) is the index set of complete rankings compatible with μ_s^* , the incomplete ranking indexed by s. Now $\mathbf{c}'_i \mathbf{\Pi}_{\ell} \mathbf{c}_i = \pi_i^*, \mathbf{c}'_i \mathbf{\Pi}_{\ell} \mathbf{c}_j = 0$ for $i \neq j$, so that $\mathbf{C}' \mathbf{\Pi}_{\ell} \mathbf{C} = \mathbf{\Pi}_{\ell}^* = diag(\pi_{\ell 1}^*, \pi_{\ell 2}^*, \dots, \pi_{\ell k!}^*)$ and thus $\mathbf{C}' \boldsymbol{\Sigma}_{\ell} \mathbf{C} = \mathbf{C}' (\mathbf{\Pi}_{\ell} - \pi_{\ell} \pi_{\ell}') \mathbf{C} = \mathbf{\Pi}_{\ell}^* - \pi_{\ell}^* \pi_{\ell}'' = \boldsymbol{\Sigma}_{\ell}^*$. Analogous to the complete case, for each sample of n_{ℓ} incomplete rankings, $\mathbf{f}_{\ell} = (f_{\ell 1}, f_{\ell 2}, \dots, f_{\ell k!})'$, where $f_{\ell s}$ is the frequency of the observed incomplete ranking μ_s^* in population ℓ , and for any matrix \mathbf{T}_{ℓ}^* of suitable dimensions, as $n \to \infty$

$$\sqrt{n_{\ell}} \mathbf{T}_{\ell}^* \left(\frac{\mathbf{f}_{\ell}}{n_{\ell}} - \mathbf{C}' \boldsymbol{\pi}_{\ell} \right) \rightarrow_{\mathcal{L}} \mathcal{N} \left(0, \mathbf{T}_{\ell}^* \boldsymbol{\Sigma}_{\ell}^* \mathbf{T}_{\ell}^{*\prime} \right).$$

In the context of incomplete rankings, the matrix \mathbf{T} is replaced by the matrix

$$\mathbf{T}_{\ell}^{*} = \mathbf{T} \mathbf{\Pi}_{\ell} \mathbf{C} \left(\mathbf{C}' \mathbf{\Pi}_{\ell} \mathbf{C} \right)^{-1}$$
(3.1)

which is the collection of the k! conditional expectations $\mathbf{t}^* = E(\mathbf{t}|C(\boldsymbol{\mu}_s^*))$, where $C(\boldsymbol{\mu}_s^*)$ is the set of complete rankings compatible with $\boldsymbol{\mu}_s^*$. Column s of matrix \mathbf{T}_{ℓ}^* is $(\pi_s^*)^{-1} \mathbf{T} \mathbf{\Pi}_{\ell} \mathbf{c}_s$ which has elements $\left[\sum_{i \in c(s)} t_i(j) \pi_{\ell j} / \pi_{\ell s}^*\right]$, with $j = 1, ..., \dim(\mathbf{t})$. In the comparison of two populations for which the rankings are incomplete with the same pattern, so that both

share the same compatibility matrix C and thus $\mathbf{T}^* = \mathbf{T}_{\ell}^*, \ell = 1, 2$, one can proceed in a manner analogous to the complete case, as follows. Under $\mathcal{H}_0 : \pi_1 = \pi_2$, for which $\mathbf{C}' \pi_1 = \mathbf{C}' \pi_2$, the analogue of (2.2) is

$$\sqrt{N}\mathbf{T}^*\left(\frac{\mathbf{f}_1}{n_1} - \frac{\mathbf{f}_2}{n_2}\right) \to_{\mathcal{L}} \mathcal{N}\left(0, \mathbf{T}^*\boldsymbol{\Sigma}^*\mathbf{T}^{*\prime}\right),\tag{3.2}$$

where $\Sigma^* = (1/\lambda_1) \Sigma_1^* + (1/\lambda_2) \Sigma_2^*$, from which a statistic analogous to (2.3) can be derived.

$$G = N\left(\mathbf{T}^*\left(\frac{\mathbf{f}_1}{n_1} - \frac{\mathbf{f}_2}{n_2}\right)\right)'\left(\mathbf{T}^*\left(\frac{\mathbf{f}_1}{n_1} - \frac{\mathbf{f}_2}{n_2}\right)\right) \to_{\mathcal{L}} \sum_{i=1}^{v} \alpha_i^* z_i^2$$
(3.3)

where the α_i^* 's are the eigenvalues of $\mathbf{T}^* \mathbf{\Sigma}^* \mathbf{T}^{*'}$. The difficulty in using such a statistic is that unlike \mathbf{T} in the complete case, \mathbf{T}^* is a function of both π^* , and, through $\mathbf{\Pi}\mathbf{C}$, of π . The probabilities in the complete case cannot be estimated from the incomplete information. In order to make use of a tractable statistic, in what follows we will replace \mathbf{T}^* in (3.1) by $\mathbf{T}^* = (k!/t!) \mathbf{T}\mathbf{C}$. This version of \mathbf{T}^* is in fact the one in which $\pi_i = 1/t!$, the uniform case, for then $\mathbf{\Pi} = (1/t!)\mathbf{I}_{t!}$, and $(\mathbf{\Pi}^*)^{-1} = k!\mathbf{I}_{k!}$, where \mathbf{I}_m is the $(m \times m)$ identity matrix.

With this change in T^* , Alvo & Cabilio (1991) show that in the Spearman and Kendall cases, the columns of T^* are respectively of the form

$$t_{S}^{*}(\boldsymbol{\mu}^{*}) = \frac{t+1}{k+1} \left[\left(\mu^{*}(i) - \frac{k+1}{2} \right) \delta(i) \right]_{i}^{\prime}, 1 \le i \le t$$
$$t_{K}^{*}(\boldsymbol{\mu}^{*}) = \left[a^{*}(i,j) \right]_{(i,j)}^{\prime}, 1 \le i < j \le \binom{t}{2}$$

where $\delta(i) = 1$ or 0, according to whether or not object *i* is ranked, and

$$a^{*}(i,j) = \begin{cases} sgn(\mu^{*}(j) - \mu^{*}(i)) & \delta(i) = \delta(j) = 1\\ 1 - \frac{2\mu^{*}(i)}{k+1} & \delta(i) = 1, \delta(j) = 0\\ \frac{2\mu^{*}(j)}{k+1} - 1 & \delta(i) = 0, \delta(j) = 1\\ 0 & \delta(i) = \delta(j) = 0 \end{cases}$$
(3.4)

For such cases, the statistic (3.3) simplifies to forms analogous to (2.4) and (2.5). In the Spearman case it becomes

$$N\left(\frac{t+1}{k+1}\right)^2 \sum_{i=1}^{t} (\bar{\mu}^{*(1)}(i) - \bar{\mu}^{*(2)}(i))^2$$
(3.5)

where $\bar{\mu}^{*(\ell)}(i) = \frac{1}{n_{\ell}} \sum_{g=1}^{n_{\ell}} \mu_g^{*(\ell)}(i) \delta_g(i)$, the mean of the ranks assigned to the non-missing object i = 1, ..., t by the n_{ℓ} judges sampled from population $\ell = 1, 2$. If we let

$$\nu(i) = \frac{t+1}{k+1} \mu^*(i) \,\delta(i) + (1-\delta(i)) \,\frac{t+1}{2} \tag{3.6}$$

Bu, Cabilio and Zhang: Tests of Concordance between Groups

we may rewrite the elements of $t_{S}^{*}\left(\boldsymbol{\mu}^{*}
ight)$ as $\left(\nu\left(i
ight) - \left(t + 1
ight) / 2
ight)$, and (3.5) becomes

$$N\sum_{i=1}^{t} \left(\bar{\nu}^{(1)}(i) - \bar{\nu}^{(2)}(i)\right)^2 \tag{3.7}$$

where $\bar{\nu}^{(\ell)}(i) = \frac{1}{n_{\ell}} \sum_{g=1}^{n_{\ell}} \nu_g^{(\ell)}(i)$, and $\nu_g^{(\ell)}(i)$ is the score (3.6) assigned to object *i* by the g^{th} judge from population $\ell = 1, 2$. Note that the *t*-dimensional vector $\boldsymbol{\nu}$ with components (3.6), is a vector with element the adjusted incomplete rank $\frac{t+1}{k+1}\mu^*(i)$ if that object is ranked, and $\frac{t+1}{2}$ if it is not.

In the Kendall case (3.3) simplifies to

$$N\sum_{i< j} \left(\bar{a}^{*(1)}(i,j) - \bar{a}^{*(2)}(i,j)\right)^2$$
(3.8)

where $\bar{a}^{*(\ell)}(i,j) = \frac{1}{n_{\ell}} \sum_{g=1}^{n_{\ell}} a_g^{*(\ell)}(i,j)$, and $a_g^{*(\ell)}(i,j)$, defined in (3.4) above, is the pairwise score assigned to object pairs (i,j), $1 \leq i < j \leq {t \choose 2}$, by the g^{th} judge from population $\ell = 1, 2$. For simplicity, in what follows, we may use a reordering of the pairs, such as the one suggested in Section 2, to write $a^*(r) = a^*(i,j)$.where $1 \leq r \leq q = {t \choose 2}$.

suggested in Section 2, to write $a^*(r) = a^*(i, j)$ where $1 \le r \le q = {t \choose 2}$. Under the null hypothesis $\Sigma_1^* = \Sigma_2^*$, and we use a pooled estimate $\hat{\Sigma}_P^* = \frac{N^2}{n_1 n_2} \frac{1}{N-2}((n_1 - 1)\hat{\Sigma}_1^* + (n_2 - 1)\hat{\Sigma}_2^*)$, which is analogous to the one in the complete case in (2.6). The eigenvalues in (3.3) are estimated by those of $\mathbf{T}^* \hat{\Sigma}_P^* \mathbf{T}^*$. The covariance estimates in (2.7) and (2.8) are replaced by their counterparts in the incomplete case.

Thus, if $\boldsymbol{\nu}_{g}^{(\ell)} = \left(\nu_{g}^{(\ell)}(1), \nu_{g}^{(\ell)}(2), ..., \nu_{g}^{(\ell)}(t)\right)', \bar{\boldsymbol{\nu}}^{(\ell)} = \left(\bar{\nu}^{(\ell)}(1), \bar{\nu}^{(\ell)}(2), ..., \bar{\nu}^{(\ell)}(t)\right)'$, then

$$\mathbf{T}_{S}^{*}\hat{\boldsymbol{\Sigma}}_{\ell}^{*}\mathbf{T}_{S}^{*\prime} = \frac{1}{n_{\ell}-1}\sum_{g=1}^{n_{\ell}} \left(\boldsymbol{\nu}_{g}^{(\ell)} - \bar{\boldsymbol{\nu}}^{(\ell)}\right) \left(\boldsymbol{\nu}_{g}^{(\ell)} - \bar{\boldsymbol{\nu}}^{(\ell)}\right)^{\prime},$$

and in the Kendall case, with $\mathbf{a}_{g}^{*(\ell)} = \left(a_{g}^{*(\ell)}(1), a_{g}^{*(\ell)}(2), ..., a_{g}^{*(\ell)}(q)\right)'$,

$$\mathbf{T}_{K}^{*}\hat{\boldsymbol{\Sigma}}_{\ell}\mathbf{T}_{K}^{*\prime} = \frac{1}{n_{\ell}-1}\sum_{g=1}^{n_{\ell}} \left(\mathbf{a}_{g}^{*(\ell)} - \bar{\mathbf{a}}^{*(\ell)}\right) \left(\mathbf{a}_{g}^{*(\ell)} - \bar{\mathbf{a}}^{*(\ell)}\right)^{\prime}.$$

4 Inferential Models

In this section we consider two situations in which complete and various patterns of incomplete blocks occur in the rankings from the two populations. In one case the sample space includes only the t! complete rankings and the (unknown) probabilities associated with such rankings. Complete rankings, if present, occur in both samples, and observed incomplete rankings with the same patterns also occur in both samples. The null hypothesis is that the distributions over the complete rankings are the same for the two populations. The second situation is one in which the sample space is enlarged to include not only permutations of complete rankings, but, as well, all permutations of incomplete rankings with various possible numbers and patterns of missing observations. In such a case the null hypothesis is that the distributions over this expanded sample space are identical. In what follows, we will denote by C_h the compatibility matrix for the pattern of missing observations indexed by h. There are $\sum_{k=2}^{t-1} {t \choose k} = 2^t - 2 - t$ such possible patterns for which $2 \le k \le t - 1$. We note that there are a total of ${t \choose k} k! = (t!/(t-k)!)$ possible different pattern and permutation outcomes for a fixed number t - k of missing observations.

4.1 Case 1: Matched Patterns of Incomplete Rankings

The approach will be demonstrated in the case where complete rankings and incomplete rankings of one specified pattern are present in both samples. In this situation we observe n_{ℓ} rankings from population $\ell = 1, 2$, of which $n_{\ell 1}$ are complete, and $n_{\ell 2}$ are incomplete with the same pattern of t - k missing observations indexed by h. For each sample of $n_{\ell 1}$ complete rankings, $\mathbf{f}_{\ell 1}$ is the t! dimensional frequency vector, and similarly, for each sample of $n_{\ell 2}$ incomplete rankings, $\mathbf{f}_{\ell 2}$ is the corresponding k! dimensional frequency vector. From (2.2) and (3.2) it follows that under $\mathcal{H}_0 : \pi_1 = \pi_2$, if $n_{\ell j}/N \to \lambda_{\ell j}$ as $N \to \infty, \ell, j = 1, 2$,

$$\begin{split} \sqrt{N}\mathbf{T}\left(\frac{\mathbf{f}_{11}}{n_{11}} - \frac{\mathbf{f}_{21}}{n_{21}}\right) &\to \mathcal{N}\left(\mathbf{0},\mathbf{T}\boldsymbol{\Sigma}\mathbf{T}'\right),\\ \sqrt{N}\mathbf{T}_{h}^{*}\left(\frac{\mathbf{f}_{12}}{n_{12}} - \frac{\mathbf{f}_{22}}{n_{22}}\right) &\to \mathcal{N}\left(\mathbf{0},\mathbf{T}_{h}^{*}\boldsymbol{\Sigma}^{*}\mathbf{T}_{h}^{*}\right). \end{split}$$

Since complete and incomplete cases are independent, the test statistic

$$G = N \begin{pmatrix} \mathbf{T}(\frac{\mathbf{f}_{11}}{n_{11}} - \frac{\mathbf{f}_{21}}{n_{21}}) \\ \mathbf{T}_{h}^{*}(\frac{\mathbf{f}_{12}}{n_{12}} - \frac{\mathbf{f}_{22}}{n_{22}}) \end{pmatrix}' \begin{pmatrix} \mathbf{T}(\frac{\mathbf{f}_{11}}{n_{11}} - \frac{\mathbf{f}_{21}}{n_{21}}) \\ \mathbf{T}_{h}^{*}(\frac{\mathbf{f}_{12}}{n_{12}} - \frac{\mathbf{f}_{22}}{n_{22}}) \end{pmatrix}$$
(4.1)

has, as $N \to \infty$, the same distribution as $\sum_{j} \beta_{j} z_{j}^{2}$ where β_{j} are the eigenvalues of $\Gamma = \begin{pmatrix} \mathbf{T} \Sigma \mathbf{T}' & \mathbf{0} \\ \mathbf{0} & \mathbf{T}_{h}^{*} \Sigma_{h}^{*} \mathbf{T}_{h}^{*'} \end{pmatrix}$. In the Spearman case, the test statistic is found to be

$$N\sum_{i=1}^{t} \left((\bar{\mu}^{(1)}(i) - \bar{\mu}^{(2)}(i))^2 + (\frac{t+1}{k+1})^2 (\bar{\mu}^{*(1)}(i) - \bar{\mu}^{*(2)}(i))^2 \right),$$
(4.2)

and in the Kendall case the test statistic is

$$N\sum_{i=1}^{q} \left((\bar{a}^{(1)}(i) - \bar{a}^{(2)}(i))^2 + (\bar{a}^{*(1)}(i) - \bar{a}^{*(2)}(i))^2 \right).$$
(4.3)

The estimates of the components of Γ are generated from the corresponding estimates in the complete and incomplete cases described in Sections 2 and 3.

The approach detailed here may be generalized to the situation where we have several fixed matching patterns of missing observations in both samples.

4.2 Case 2: Expanded Probability Space

The approach in this case is a direct extension of the complete case with an expansion of possible outcomes, and an adjustment of the matrix of score vectors to take into account the incomplete rankings. The probabilities over this expanded space are denoted by π_{ℓ}^* , $\ell = 1, 2$, and the hypothesis of homogeneity \mathcal{H}_0^* : $\pi_1^* = \pi_2^*$, may be tested either conditionally on the observed patterns of missing observations, or unconditionally over all $2^t - 2 - t$ possible patterns. In either case, define a new matrix

$$\mathbf{T}^* = \mathbf{T}\left(\mathbf{I}|\frac{k_1!}{t!}\mathbf{C}_1|\frac{k_2!}{t!}\mathbf{C}_2|\cdots|\frac{k_d!}{t!}\mathbf{C}_d\right)$$

where each \mathbf{C}_h represents a compatibility matrix for a distinct incomplete pattern, and d is the number of incomplete patterns included in the test for homogeneity. Let $\mathbf{f}'_{\ell} = (\mathbf{f}'_{\ell 0}|\mathbf{f}'_{\ell 1}|\cdots|\mathbf{f}'_{\ell d})$ denote the vector of frequencies, where $\mathbf{f}_{\ell 0}$ and the $\mathbf{f}_{\ell h}$, h = 1, ..., d, are, respectively, the vectors of frequencies for the complete case and for the individual incomplete patterns listed in the order corresponding to that in \mathbf{T}^* . Let $n_{\ell h}$ denote the number observed with pattern indexed by h in population ℓ , so that $\sum_{h=0}^{d} n_{\ell h} = n_{\ell}$. Arguments identical to those in Section 2, yield analogous results with \mathbf{T}^* playing the role of \mathbf{T} . In particular, if n_{ℓ} is the total number of observations from population $\ell = 1, 2, N = n_1 + n_2$ and Σ_{ℓ} is the covariance matrix corresponding to the distribution π^*_{ℓ} , it follows that for $\mathcal{H}^*_0 : \pi^*_1 = \pi^*_2$, the test statistic (2.3) becomes

$$G = N\left(\mathbf{T}^*\left(\frac{\mathbf{f}_1}{n_1} - \frac{\mathbf{f}_2}{n_2}\right)\right)'\left(\mathbf{T}^*\left(\frac{\mathbf{f}_1}{n_1} - \frac{\mathbf{f}_2}{n_2}\right)\right).$$
(4.4)

The asymptotic distribution of this statistic is approximated by the distribution of $\sum_i \hat{\alpha}_i z_i^2$, where z_i are independent standard normal variates, and $\hat{\alpha}_i$'s are the eigenvalues of $\mathbf{T}^* \hat{\Sigma} \mathbf{T}^{*\prime}$, where $\hat{\Sigma}$ is an appropriate estimate of $\boldsymbol{\Sigma} = (1/\lambda_1) \boldsymbol{\Sigma}_1 + (1/\lambda_2) \boldsymbol{\Sigma}_2$. In order to simplify (4.4) it is convenient to relabel the observed values of elements of \mathbf{T}^* . In the Spearman case the collection of t-dimensional observed scores of \mathbf{T}_S^* may be written as

$$(\boldsymbol{\mu}_1, \boldsymbol{\mu}_2, ..., \boldsymbol{\mu}_{n_{\ell_0}} | \boldsymbol{\nu}_{11}, \boldsymbol{\nu}_{12}, ..., \boldsymbol{\nu}_{1n_{\ell_1}} | \cdots | \boldsymbol{\nu}_{d1}, \boldsymbol{\nu}_{d2}, ..., \boldsymbol{\nu}_{dn_{\ell_d}}) - \left(\frac{t+1}{2}\right) \mathbf{J},$$

where **J** is a matrix of 1's, μ_j are complete rankings, and ν_{hj} are weighted incomplete rankings with elements as in (3.6), corresponding to incomplete pattern h = 1, ..., d. We denote these observations by $\nu_g^{*(\ell)}$ for $g = 1, ..., n_\ell, \ell = 1, 2$. The same arguments leading to (2.4)

and (3.7) show that (4.4) may be written as

$$N\sum_{i=1}^{t} \left(\bar{\nu}^{*(1)}(i) - \bar{\nu}^{*(2)}(i)\right)^2,\tag{4.5}$$

where $\bar{\nu}^{*(\ell)}(i) = \frac{1}{n_{\ell}} \sum_{g=1}^{n_{\ell}} \nu_g^{*(\ell)}(i)$. The Kendall case follows in a similar manner, with the the collection of $q = {t \choose 2}$ -dimensional observed scores of \mathbf{T}_K^* , written as

 $(\mathbf{a}_{1}, \mathbf{a}_{2}, ..., \mathbf{a}_{n_{\ell 0}} | \mathbf{a}_{11}^{*}, \mathbf{a}_{12}^{*}, ..., \mathbf{a}_{1n_{\ell 1}}^{*} | \cdots | \mathbf{a}_{d1}^{*}, \mathbf{a}_{d2}^{*}, ..., \mathbf{a}_{dn_{\ell d}}^{*})$

where \mathbf{a}_j are vectors of complete pairwise concordances, and \mathbf{a}_{hj}^* h = 1, ..., d, are vectors with elements as in (3.4). Analogously to the Spearman case we denote these observations by $\alpha_g^{*(\ell)}$ for $g = 1, ..., n_\ell, \ell = 1, 2$, and the same arguments leading to (2.5) and (3.8) show that (4.4) may be written as

$$N\sum_{r=1}^{q} \left(\bar{\alpha}^{*(1)}(r) - \bar{\alpha}^{*(2)}(r)\right)^2,\tag{4.6}$$

where $\bar{\alpha}^{*(\ell)}(r) = \frac{1}{n_{\ell}} \sum_{g=1}^{n_{\ell}} \alpha_g^{*(\ell)}(r)$. The estimate of the covariance matrix under $\mathcal{H}_0^* : \pi_1^* = \pi_2^*$ is of the form (2.6), with $\mathbf{F}_{\ell} = diag \left(\mathbf{f}_{\ell 0}' | \mathbf{f}_{\ell 1}' | \cdots | \mathbf{f}_{\ell d}' \right)$ in this expanded case. In the Spearman case

$$\mathbf{T}_{S}^{*}\hat{\boldsymbol{\Sigma}}_{\ell}\mathbf{T}_{S}^{*\prime} = \frac{1}{n_{\ell}-1}\sum_{g=1}^{n_{\ell}} \left(\boldsymbol{\nu}_{g}^{*(\ell)} - \bar{\boldsymbol{\nu}}^{*(\ell)}\right) \left(\boldsymbol{\nu}_{g}^{*(\ell)} - \bar{\boldsymbol{\nu}}^{*(\ell)}\right)'$$

and in the Kendall case

$$\mathbf{T}_{K}^{*} \hat{\mathbf{\Sigma}}_{\ell} \mathbf{T}_{K}^{*\prime} = rac{1}{n_{\ell} - 1} \sum_{g=1}^{n_{\ell}} \left(\boldsymbol{\alpha}_{g}^{*(\ell)} - ar{\boldsymbol{\alpha}}^{*(\ell)}
ight) \left(\boldsymbol{\alpha}_{g}^{*(\ell)} - ar{\boldsymbol{\alpha}}^{*(\ell)}
ight)^{\prime}.$$

5 **Implementation of the Tests**

Under either case 1 or 2 inferential model, we reject the null hypothesis of homogeneity for large observed values of the statistic G in (4.1) or (4.4). For an observed value G_0 of this test statistic we can approximate the $P - value = P(G \ge G_0)$ of the test in several ways which we describe in the following. One approach is to use the asymptotic distribution to approximate the null. Approximations to the distribution of $Q = \sum_{i} \psi_{i} z_{i}^{2}$, where z_{i} are independent standard normal variates, are given in Jensen & Solomon (1972). Letting $\theta_s = \sum_i \psi_i^s$, this approach, known as the Wilson-Hilferty approximation, is to show that the distribution of $(\hat{Q}/\theta_1)^h$ is approximately normal when $h = 1 - (2\theta_1\theta_3/3\theta_2^2)$, with mean and variance

Bu, Cabilio and Zhang: Tests of Concordance between Groups

 $(1 + (\theta_2 h(h-1))/\theta_1^2)$ and $(2\theta_2 h^2)/\theta_1^2$ respectively. In practice, Q is the appropriate statistic G in (4.1) or (4.4), ψ_i are the eigenvalues of the estimated covariances, $\widehat{\Gamma}$ in case 1, or $\mathbf{T}^* \widehat{\Sigma} \mathbf{T}^{*'}$ in case 2. Another easily implemented approach is to simulate the P - value of G_0 by resampling. In case 1, with one set of incomplete observations, at each step we randomly select n_{11} complete rankings from the combined set of observed $n_{11} + n_{21}$ complete rankings, and n_{12} incomplete rankings from the combined set of observed $n_{12} + n_{22}$ incomplete rankings with the same pattern, and thus calculate the corresponding value of G_0 . In case 2 we randomly select n_1 rankings from the combined set of $N = n_1 + n_2$ rankings of all types, and calculate G_0 .

Instead of using the statistic G, another approach that can be considered is the multiresponse permutation procedure (MRPP) developed by Mielke & Berry (2001). This MRPP statistic is the weighted sum of the internal average distances between all rankings within a group of rankings. The weights chosen are the proportional sizes of each group. In our situation this statistic is of the form $\delta = (n_1/N) \xi_1 + (n_2/N) \xi_2$, where ξ_i is the average distance between all rankings in group \mathcal{G}_i

$$\xi_i = \binom{n_i}{2}^{-1} \sum_{r,s \in \mathcal{G}_i} d\left(\mu_r, \mu_s\right).$$

In this setting, small values of δ are significant. The test is implemented through simulations in which the resampling procedure is the same as that described in the *G*-test above.

The distances between rankings are related to the similarity measures $\mathcal{A}(\mu_1, \mu_2)$ through $d(\mu_1, \mu_2) = c - \mathcal{A}(\mu_1, \mu_2)$. In the Spearman case the distance is

$$d_S(\boldsymbol{\mu}_1, \boldsymbol{\mu}_2) = \frac{1}{2} \sum_{i=1}^t \left(\mu_1(i) - \mu_2(i) \right)^2 = \frac{t(t^2 - 1)}{12} - \sum_{i=1}^t \left(\mu_1(i) - \frac{t+1}{2} \right) \left(\mu_2(i) - \frac{t+1}{2} \right)$$

so that $c_S = t(t^2 - 1)/12$, or equivalently

$$d_S(\boldsymbol{\mu}_1, \boldsymbol{\mu}_2) = \frac{t(t+1)(2t+1)}{6} - \sum_{i=1}^t \mu_1(i)\mu_2(i),$$

while in the Kendall case

$$d_{K}(\boldsymbol{\mu}_{1}, \boldsymbol{\mu}_{2}) = \sum_{i < j} (1 - sgn(\mu_{1}(j) - \mu_{1}(i))sgn(\mu_{2}(j) - \mu_{2}(i)))$$

$$= \frac{t(t-1)}{2} - \sum_{i < j} sgn(\mu_{1}(j) - \mu_{1}(i))sgn(\mu_{2}(j) - \mu_{2}(i))$$

so that $c_K = t(t-1)/2$. Analogously, Alvo & Cabilio (1995) defined distances between two incomplete rankings μ_1^* and μ_2^* , with possibly different patterns and number of missing observations $t - k_1$ and $t - k_2$, as $d(\mu_1^*, \mu_2^*) = c - A(\mu_1^*, \mu_2^*)$, where $A(\mu_1^*, \mu_2^*) =$ $t^* (\boldsymbol{\mu}_1^*)' t^* (\boldsymbol{\mu}_2^*)$. In the Spearman case

$$d_{S}^{*}(\boldsymbol{\mu}_{1}^{*}, \boldsymbol{\mu}_{2}^{*}) = c_{S} - \sum_{i=1}^{t} \left(\frac{t+1}{k_{1}+1}\right) \left(\mu_{1}^{*}(i) - \frac{k_{1}+1}{2}\right) \left(\frac{t+1}{k_{2}+1}\right) \left(\mu_{2}^{*}(i) - \frac{k_{2}+1}{2}\right) \delta(i)$$
$$= \frac{t(t+1)(2t+1)}{6} - \sum_{i=1}^{t} \nu_{1}(i)\nu_{2}(i)$$

while in the Kendall case

$$d_{K}^{*}(\boldsymbol{\mu}_{1}^{*}, \boldsymbol{\mu}_{2}^{*}) = c_{K} - \sum_{i < j} a_{1}^{*}(i, j) a_{2}^{*}(i, j).$$

Distances defined this way have the property that the distance between an incomplete ranking and itself is a minimal distance, but this distance is greater than zero. This is not surprising, given the fact that such measures between two incomplete rankings are in fact averages of the distances between the sets of complete rankings compatible with each of the incomplete rankings. In applying MRPP methods, Mielke & Berry (2001) do not require that such distances be metrics, in that the triangle inequality need not maintain, thus allowing squared Euclidian distances, of which the Spearman distance is a special example. However, they do require that this distance, or measure of separation, have the property that the distance from an element to itself be zero, a property which our distances in the incomplete case do not possess. Thus the use of MRPP in this case may not produce reliable results.

6 Examples

In the following we give two examples in which we apply the tests developed here, as well as the MRPP, for both models. All the calculations were conducted using the R programming language, and the code is available from the third author.

6.1 Medical Compliance

Gwadry-Sridhar, et al. (2005), report a study on the effect of additional education on medical compliance of heart failure patients. Pharmacy refill data was collected on patients who had been randomly allocated to receive control (standard information) or treatment (additional information on diet, lifestyle and compliant medication use). The patients of interest received as many as four medications (ACE-inhibitors, diuretics, beta-blockers and digoxin), but not all patients received all four. For each patient, a value of an index of medical compliance was calculated for each medication taken, yielding four dimensional data vectors with possibly missing values. In this example we consider only blocks which have at least two values, and no ties. Table 1 lists the number of observations for each missing pattern, where 'X' denotes an observation, and '_ ' a missing value.

Pattern	Control	Treatment	Total
XXXX	12	12	24
XX_X	11	8	19
XXX_{-}	10	11	21
X_X	5	8	13
XX	5	9	14
	$n_1 = 43$	$n_2 = 48$	N = 91

Table 1. Sample Sizes in Medical Compliance of Heart Failure Data

Tests were conducted for both Case 1 and Case 2 models using the Spearman measure, with the resulting P - values given in Table 2. We note that even though the sample sizes appear to be moderate, in both cases the P - values using the Wilson-Hilferty approximation of the asymptotic distribution are very close to the corresponding simulated values. However the P - values are much higher for the Case 1 null hypothesis. The P - value obtained by the MRPP method is close to the others for Case 2, but is quite different for Case 1. We note that Wilcoxon tests were conducted for each of the four medications individually, and in each case the results were not significant

	Case 1	Case 2
Wilson-Hilferty	0.720	0.220
Resampling	0.727	0.219
MRPP	0.382	0.230

Table 2. P - values for Tests of Homogeneity for Medical Compliance

6.2 Leisure Time Preferences

Hollander & Sethuraman (1978) present data on leisure time preferences of two groups of females, one black, and the other white. The females were in the age group 70-79, and were asked to rank their preferences of the gender of individual they preferred to spend their leisure time with, the choices being "male", "female", or "both". This complete rank data was analyzed in Feigin & Alvo (1986), and the results were found to be highly significant. In order to see how incomplete blocks would affect the conclusions, we randomly deleted one rank from six responses in each group, three from the "female" and three from the "both" categories. Table 3 lists the number of observations for each missing pattern.

Pattern	Group 1	Group 2	Total
XXX	8	7	15
X_X	3	3	6
XX_{-}	3	3	6
	$n_1 = 14$	$n_2 = 13$	N = 27

Table 3. Sample Sizes in Leisure Time Data

Tests were conducted for both Case 1 and Case 2 models using both the Spearman and Kendall measures, with the resulting P - values given in Table 4. In addition the tests were carried out for the complete data and the results are given in Table 5. It is seen that the results are fairly consistent between complete and incomplete cases, statistics used, models tested and measures used.

	Spearman		Kendall	
	Case 1	Case 2	Case 1	Case 2
Wilson-Hilferty	0.0000	0.0001	0.0001	0.0001
Resampling	0.0001	0.0001	0.0000	0.0004
MRPP	0.0001	0.0000	0.0000	0.0003

Table 4. P - values for Tests of Homogeneity for Leisure Time Incomplete Data

	Spearman	Kendall
Wilson-Hilferty	0.0000	0.0002
Resampling	0.0003	0.0002
MRPP	0.0001	0.0002

Table 5. P - values for Tests of Homogeneity for Complete Leisure Time Data

7 Conclusion

We have developed rank based methods for testing homogeneity of populations of rankings, which may be complete or incomplete, based on functions of measures of similarity between such rankings. Although in this development we have restricted attention to two measures of similarity, the Spearman and the Kendall, such an approach is completely generalizable to other measures such as Hamming (Alvo & Cabilio, 1998). All such tests have forms which are easily calculated, and their asymptotic distributions are linear combinations of independent χ^2

variables with one degree of freedom. These test statistics are not consistent with the null hypothesis of uniformity between the populations, but with the more restrictive assumption of uniformity within each population. Such a restriction is necessary, for without it the statistic derived would be intractable. We consider two different population models from which two different statistics emerge. In the first case all incomplete blocks have patterns and numbers of missing observations which are present in the samples from both populations, and the statistic is appropriate for testing the homogeneity between the two populations of probability distributions over the complete rankings. In the second case, the probability space is expanded to include all permutations of incomplete rankings with various number and patterns of missing values, and the appropriate test is derived to test for the homogeneity of such probabilities between the two populations. Two examples are presented in which P - values of the tests are approximated using the asymptotic distribution as well as by simulation methods. In addition, the MRPP procedure developed in Mielke & Berry (2001) was applied to the same examples. It is noted that the distance measures in the incomplete cases are not proper in the sense that the distance between one incomplete ranking and itself is not zero, and this may affect the applicability of the MRPP procedure. In the examples given, resampling and asymptotic results give very similar results for both models, while MRPP results appear quite different in one case. The P - values may differ between the two models as is the case in Example 1.

Acknowledgement

The work of P. Cabilio and Y. Zhang was supported in part by Natural Science and Engineering Research Council of Canada Discovery Grants.

References

- [1] Alvo, M., and Cabilio, P. (1991). On the balanced incomplete block design for rankings. *Annals of Statistics*, **19**, 1597-1613.
- [2] Alvo, M. and Cabilio, P. (1995). Rank correlation methods for missing data. *Canadian Journal of Statistics*, 23, 345-358.
- [3] Alvo, M. and Cabilio, P. (1998). Applications of Hamming distance to the analysis of block designs, in *Asymptotic Methods in Probability and Statistics*, B. Szyszkowicz (Editor), Elsevier, 787-799.
- [4] Alvo, M. and Cabilio, P. (1999). A general rank based approach to the analysis of block data. *Communications in Statistics: Theory and Methods*, 28, 197-215.
- [5] Feigin, P. and Alvo, M. (1986). Intergroup diversity and concordance for ranking data: an approach via metrics for permutations. *Annals of Statistics*, **14**, 691-707.

- [6] Gwadry-Sridhar, F., Arnold, J. M. O., Zhang, Y., Brown, J. E., Marchiori, G., Guyatt, G. (2005). Pilot study to determine the impact of a multidisciplinary educational intervention in patients hospitalized with heart failure. *American Heart Journal*, 150: 982e1-989e9.
- [7] Hollander, M. and Sethuraman, J. (1978). Testing for agreement between two groups of judges. *Biometrika*, **65**, 403-411.
- [8] Jensen, D. R., and Solomon, H. (1972). A Gaussian approximation to the distribution of a definite quadratic form. *J. of the American Statistical Association*, **76**, 898-902
- [9] Mielke, P. W., and Berry, K. L. (2001). *Permutation Methods: A Distance Function Approach*. Springer, New York.