A Unified Approach to Constructing Nonparametric Rank Tests* **

by

Douglas E. Critchlow
Purdue University

Technical Report #86-15

Department of Statistics
Purdue University

June 1986

* This research was supported by National Science Foundation Grant DMS-86-01645, and by Office of Naval Research Contract N00014-86-K-0156.

** This research is also appearing as Technical Report #376, July 1986, Department of Statistics, Stanford University.

A Unified Approach to Constructing Nonparametric Rank Tests* **

by

Douglas E. Critchlow
Purdue University

Technical Report #86-15

Department of Statistics
Purdue University

June 1986

* This research was supported by National Science Foundation Grant DMS-86-01645, and by Office of Naval Research Contract N00014-86-K-0156.

** This research is also appearing as Technical Report #376, July 1986, Department of Statistics, Stanford University.

ABSTRACT

One shortcoming of the present theory of rank tests is that such tests have usually been constructed on a case-by-case basis, in a quite ad hoc (albeit clever) manner. This paper attempts to provide the basis for a more unified approach to rank tests. It investigates a general, yet simple construction, which simultaneously generates many rank test statistics, for a multitude of hypothesis testing situations. The proposed construction uses metrics on the permutation group in a novel way: the proposed test statistic is the distance between two *sets* of permutations.

This new construction is applied systematically to the two-sample and multi-sample location problems, the two-way layout problem, the one-sample location problem, the two-sample dispersion problem with equal medians, and the problem of testing for trend. It is shown that the construction: (1) works in a variety of testing situations; (2) gives rise to many familiar rank test statistics; (3) produces several other test statistics which are less familiar, yet equally plausible; and (4) enables one to extend rank tests to other hypothesis testing situations. Some connections with the existing nonparametric theory are discussed.

TABLE OF CONTENTS

1. Introduction

Nonparametric tests, based on ranks, have proved useful in a variety of applications. They are often preferred to normal theory procedures, since rank tests require few assumptions about the underlying distribution generating the data. An examination of the book of Hollander and Wolfe (1973), for example, reveals that rank tests have been developed for a multitude of hypothesis testing situations, including the two-sample and multi-sample location problems, the two-way layout problem, the one-sample location problem, the two-sample dispersion problem with equal medians, and problems of testing for trend and for independence.

One shortcoming of the present theory of rank tests is that such tests have usually been constructed on a case-by-case basis, in a quite ad hoc (albeit clever) manner. They depend strongly on the particular testing situation under consideration, and often do not extend readily to other hypothesis testing problems. The work of Hájek and Sidák (1967) addresses this issue in some special cases, and contains numerous statements such as the following (page 95): "from the intuitive point of view, the [Ansari-Bradley] test [of difference in scale] appears to be an analogue of the Wilcoxon test [for the two-sample location problem]." Moreover, Hájek and Sidák (page 98) describe the Siegel-Tukey (1960) method of converting tests of location into tests of dispersion, and also discuss the literature on extending two-sample location tests to the multi-sample problem (pages 106–108). However, there is no truly general procedure for extending an arbitrary rank test from one testing situation to another.

This paper introduces a general, yet simple construction, which produces many test statistics for *all* of the hypothesis testing situations mentioned in the first paragraph. Indeed, the construction gives rise to the following familiar rank test statistics:

(1) the Mann-Whitney statistic, the Wilcoxon statistic, the Mood "median test" statis-

tic, the Kolmogorov-Smirnov statistic, and the Wald-Wolfowitz statistic for the two-sample location problem;

(2) the Jonckheere-Terpstra statistic for the multi-sample location problem with ordered alternatives;

(3) the Page test statistic for the two-way layout with ordered alternatives;

(4) the Wilcoxon signed rank test statistic and the Fisher sign test statistic for the one-sample location problem;

(5) the Ansari-Bradley statistic and the "quartile test" statistic for the two-sample dispersion problem with equal medians.

In addition, the construction gives rise to many less familiar, yet equally plausible test statistics. The author is encouraged by these results, and is therefore hopeful that this novel approach may provide the basis for a new, more unified study of rank tests.

2. The Basic Construction

The proposed construction uses distances between permutations, and between equivalence classes of permutations. Such distances, or metrics, have recently enjoyed increased attention in the statistical literature. Several well known metrics on permutations are presented and discussed briefly in Section 3; they will prove essential to our study of rank tests.

The problem of constructing a nonparametric rank test has the following general formulation. Given two hypotheses H_0 and H_1, one wishes to develop a nonparametric test of H_0 versus H_1, on the basis of an observed ranking (or permutation) of n individuals (or items). In this paper, attention is concentrated only on the problem of finding a suitable test statistic, rather than on the closely related (and important) topic of the distribution of this statistic under H_0.

As a simple example, let the hypothesis H_1 be that women perform better than men on a certain test, and H_0 is that they do not. The available data are the test scores of a group of n_1 men and n_2 women, $n_1 + n_2 = n$. If normal distribution theory is not applicable, one wants to find a nonparametric test statistic for H_0 versus H_1, using only the ranks of the n individual scores.

The proposed general solution to such problems — whether they be the two-sample location problem illustrated above, or some other nonparametric testing situation — is based on the distance between two *sets* of permutations. It is implemented in the following steps:

Step 1: Collect all data relevant to the problem, and rank order *all* of these observations. This produces a single permutation, which we denote by π. (To simplify the presentation, we will assume throughout that all data come from *continuous* probability distributions, so that there are *no ties* in the permutation π.)

Step 2: Identify the *set* of permutations, which (for the testing problem at hand) are equivalent to the *observed* permutation π. This is an equivalence class of permutations, which we denote by $[\pi]$.

Step 3: Identify the set E of *extremal* permutations, consisting of all permutations which are least in agreement with H_0, and most in agreement with H_1.

Step 4: To test H_0 versus H_1, let d be a metric on permutations, and consider the minimum pointwise distance between the sets in Steps 2 and 3:

$$d([\pi], E) = \min_{\substack{\alpha \epsilon [\pi] \\ \beta \epsilon E}} \quad d(\alpha, \beta).$$

This is the proposed test statistic.

Since $d([\pi], E)$ measures the distance from $[\pi]$ to the set of permutations which are most in agreement with H_1, it follows that one wants to reject H_0 for *small* values of $d([\pi], E)$. This contrasts with the form of some parametric tests, where one computes the distance from H_0, and rejects H_0 if this distance is large. However, as illustrated especially by the trivial case of testing for trend (Section 12), measuring the distance from H_1 is really the only natural way to proceed in the nonparametric situation.

Another immediate, and intuitively appealing, property of the proposed test statistic $d([\pi], E)$, is that $d([\pi], E)$ attains its extremal value (zero) if and only if the observed permutation π is equivalent to some extremal permutation $e \epsilon E$. In other words, the strongest evidence for rejecting H_0 occurs if and only if π is equivalent to some $e \epsilon E$. This fact follows trivially from the basic metric property that $d(\pi, \sigma) = 0$ if and only if $\pi = \sigma$.

Mathematically, the sets of permutations in Steps 2 and 3 can often be given group-theoretic descriptions. The group theory provides a convenient notation for these sets, and it simplifies the derivations of the induced test statistics, but it does not appear to be absolutely essential to the theory.

Although the basic construction is rather abstractly formulated, it gives rise to concrete results. Indeed, as discussed below, this procedure: (1) works in a variety of testing situations; (2) gives rise to many familiar nonparametric test statistics; (3) produces several other test statistics which are less familiar, yet equally plausible; and (4) enables us to extend our tests to other hypothesis testing situations. Moreover, the tests produced by this procedure have nice statistical properties, which, in turn, suggest possible future areas of investigation.

3. Some Statistically Meaningful Metrics on the Permutation Group

In Section 2 a method of generating rank tests was described briefly, which uses metrics on permutations. Before giving a detailed illustration of this methodology in Section 4, it will be helpful to list several especially important metrics, and to introduce a suitable notation for these metrics.

To any ranking of n items, there corresponds an element π of the permutation group S_n. Thus π: $\{1,\ldots,n\} \to \{1,\ldots,n\}$ is a function. We make the important convention that $\pi(i)$ *is the rank assigned to item* i.

Diaconis (1982) lists the following metrics on S_n, which are among the most widely used in applied scientific and statistical problems:

$K(\pi,\sigma)$ = number of pairs of items, (i,j), such that $\pi(i) < \pi(j)$ and $\sigma(i) > \sigma(j)$

= minimum number of pairwise adjacent transpositions of items needed to transform π into σ

is Kendall's tau;

$F(\pi,\sigma) = \sum_{i=1}^{n} |\pi(i) - \sigma(i)|$ is Spearman's footrule;

$R(\pi,\sigma) = \left(\sum_{i=1}^{n} (\pi(i) - \sigma(i))^2\right)^{1/2}$ is Spearman's rho;

$H(\pi,\sigma) = \# \ \{i = 1,\ldots,n{:}\, \pi(i) \neq \sigma(i)\}$ is Hamming distance;

$U(\pi,\sigma) = n-$ the length of the longest increasing subsequence in $\sigma\pi^{-1}(1),\ldots,\sigma\pi^{-1}(n)$

$= n-$ the maximal number of items ranked in the same relative order by π and σ

is Ulam's distance.

In the preceding definitions of the metrics, a couple of points require clarification:

(1) For any finite set Ω, $\#\Omega$ denotes the number of elements in Ω. Thus, Hamming distance is simply the number of items which are assigned different ranks by the two permutations.

(2) The more commonly used forms of Kendall's tau, Spearman's footrule, and Spearman's rho are obtained by replacing the metric d by $1 - 2d/M$, where $M = M(d) = \max_{\pi,\sigma \epsilon S_n} d(\pi,\sigma)$. This has the advantage of renormalizing the possible values to lie between -1 and 1, so that they can be interpreted in roughly the same way as a correlation coefficient. However, for the purpose of constructing nonparametric tests, we will work with the original metric versions d, rather than the transformed versions $1 - 2d/M$. Indeed, our whole approach is based on using d as a measure of the *distance* between two sets of permutations. The prevailing practice of renormalizing these metrics may explain why a metric-based approach to rank tests has not been previously explored (at least, not in the context of the present paper).

Kendall's tau and Spearman's rho are, of course, widely used nonparametric "measures of association" for two rankings. Spearman's footrule has recently enjoyed renewed interest (Diaconis and Graham (1977), Feigin and Alvo (1986)). Hamming distance is utilized in coding theory to measure the discrepancy between two binary sequences (Berlekamp (1968), Cameron and Van Lint (1980)). Ulam's metric is used in DNA research to measure the distance between two strings of molecules (Ulam (1972, 1981)), and has also been investigated in the statistical literature as a nonparametric measure of association (Gordon (1979)).

Diaconis (1982) has a beautiful and enlightening discussion of these metrics on the permutation group. This discussion is summarized, and extended, in the recent monograph of the present author. In particular, see Critchlow (1985) for a more complete description of these metrics and their defining properties (pages 5–11), and for a host of additional applications of these metrics to a different collection of statistical problems, involving fully ranked and partially ranked data (pages 97–129). Finally, for further related work, see the paper of Feigin and Alvo (1986) on a metric-based approach to comparing two populations of rankers, using Rao's (1982a, 1982b) apportionment of diversity; and see the papers of

Mallows (1957), Feigin and Cohen (1978), and Fligner and Verducci (1985) on metric-based ranking models.

4. The Two-Sample Location Problem, with Ordered Alternatives

The basic construction will be illustrated now, in detail, for the two-sample location problem with ordered alternatives. In particular, we will proceed methodically through the four steps described in Section 2, and will discuss the implementation of each step for the problem at hand.

Recall the example in which the data are the test scores of n_1 men and n_2 women, $n_1 + n_2 = n$. Let F_1 and F_2 be the (population) distribution functions for men's and women's scores, respectively. For the two-sample location problem, the null hypothesis is H_0: $F_1(x) \equiv F_2(x)$ (i.e. men and women perform equally), and one possible formulation of the alternative is H_1: $F_1(x) \geq F_2(x)$, with strict inequality for some x (i.e. women score higher).

Step 1: The relevant permutation π is the observed rank ordering of all n individuals. We make the following conventions: individuals $1, \ldots, n_1$ and $n_1 + 1, \ldots, n_1 + n_2$ are from the first population (men) and second population (women), respectively. (This arbitrary labeling of the individuals can be performed without any loss of generality, because all of the metrics under consideration are *right-invariant*; that is, they satisfy $d(\pi, \sigma) = d(\pi\tau, \sigma\tau)$ for all $\pi, \sigma, \tau \epsilon S_n$. See Diaconis (1982), Critchlow ((1985), pages 10–11), or the Appendix of the present paper for a discussion.)

Also, recall $\pi(i)$ is the rank given to individual i, with low-numbered ranks (i.e. ranks 1, 2, 3, etc.) corresponding to low scores. (In general, in the future, we will continue to follow the convention that "items are ranked from least to greatest.") This completely specifies π as a member of the permutation group S_n.

Step 2: For the two-sample location problem, a permutation α is "equivalent" to the observed permutation π if and only if it assigns the same set of ranks to population 1, i.e. if and only if $\alpha\{1,\ldots,n_1\} = \pi\{1,\ldots,n_1\}$. This definition is justified formally by the theory of maximal invariants (Lehmann (1959)): for this problem, the maximal invariants are the set of ranks assigned to population 1; α is equivalent to π if and only if it yields the same values of the maximal invariants. The *intuitive* idea is that any such α provides the same information as π for the testing problem at hand.

The equivalence class $[\pi]$ is thus a subset of S_n, consisting of $n_1!n_2!$ permutations. In group-theoretic notation, $[\pi]$ is the left coset $\pi S_{n_1} \times S_{n_2}$, where $S_{n_1} \times S_{n_2}$ is the subgroup $\{\sigma \epsilon S_n\colon\ \sigma(i) \le n_1\ \forall\, i \le n_1 \text{ and } \sigma(i) > n_1\ \forall\, i > n_1\}$.

Step 3: The extremal set E consists of all permutations which are most in agreement with H_1, and least in agreement with H_0. Intuitively, these are the permutations which rank all individuals in the first population before all individuals in the second population. Pictorially, E consists of all rankings of the form

$$\underbrace{X\cdots X}_{n_1} \qquad \underbrace{O\cdots O}_{n_2},$$

where X's and O's denote individuals from populations 1 and 2, respectively. Mathematically, E is the subgroup $S_{n_1} \times S_{n_2}$ described in Step 2.

Step 4: The test statistic for H_0 versus H_1 is given by the minimum interpoint distance between the sets constructed in Steps 2 and 3:

$$d([\pi], E) = \min_{\substack{\alpha \epsilon \pi S_{n_1} \times S_{n_2} \\ \beta \epsilon S_{n_1} \times S_{n_2}}} d(\alpha, \beta).$$

It remains to describe these test statistics explicitly, for various choices of the metric d on S_n.

Recall that two test statistics are *equivalent* if one can be obtained from the other by a monotone transformation. The following interesting result says that the test statistics induced by the metrics of Section 3 are, in fact, equivalent to some well known test statistics.

To state the result, we introduce the

Notation: Let N_1 and N_2 be the sets $N_1 = \{1, \ldots, n_1\}$ and $N_2 = \{n_1+1, \ldots, n_1+n_2\}$. Thus N_1 and N_2 are the sets of labels for the items in the first and second populations, respectively.

Let $a_1 < \ldots < a_{n_1}$ be the ranks assigned by π to the first population; that is, $\{a_1, \ldots, a_{n_1}\}$ is the set $\{\pi(1), \ldots, \pi(n_1)\}$ listed in ascending order. Similarly let $a_{n_1+1} < \ldots < a_{n_1+n_2}$ be the ranks assigned by π to the second population.

For $i = 1, \ldots, n$ let

$$t_{i1} = \#\{j \epsilon N_1\colon \ \pi(j) \leq i\},$$

$$t_{i2} = i - t_{i1} = \#\{j \epsilon N_2\colon \ \pi(j) \leq i\}.$$

Thus t_{i1} is the number of items from the first population, which fall among the top i ranks, while t_{i2} is the corresponding number for the second population.

Finally, let $\hat{F}_1(x)$ and $\hat{F}_2(x)$ be the empirical cumulative distribution functions for the first and second populations, respectively.

Theorem 1: The test statistics for the two-sample location problem with ordered alternatives, induced by the metrics of Section 3, are:

$$K(\pi S_{n_1} \times S_{n_2}, S_{n_1} \times S_{n_2}) = \#\{(i,j) \epsilon N_1 \times N_2\colon \ \pi(i) > \pi(j)\}.$$

This is the Mann-Whitney (1947) test statistic.

$$F(\pi S_{n_1} \times S_{n_2}, S_{n_1} \times S_{n_2}) = 2 \sum_{i=1}^{n_1} \pi(i) - n_1(n_1+1).$$

This is equivalent to the Wilcoxon (1945) test statistic $\sum_{i=1}^{n_1} \pi(i)$.

$$R^2(\pi S_{n_1} \times S_{n_2}, S_{n_1} \times S_{n_2}) = \sum_{i=1}^{n}(a_i - i)^2 = \frac{n(n+1)(2n+1)}{3} - 2 \sum_{i=1}^{b} ia_i.$$

This is equivalent to the test statistic $\sum_{i=1}^{n} ia_i$.

$$H(\pi S_{n_1} \times S_{n_2}, S_{n_1} \times S_{n_2}) = 2\#(\{\pi(1), \ldots, \pi(n_1)\} \cap \{n_1+1, \ldots, n_1+n_2\})$$
$$= 2\#(\pi(N_1) \cap N_2).$$

For equal sample sizes ($n_1 = n_2 = n/2$), this is equivalent to the Mood (1950) "median test" statistic $= \#(\pi(N_1) \cap N_2) = \#(\pi(N_1) \cap \{n/2+1, \ldots, n\})$.

$$U(\pi S_{n_1} \times S_{n_2}, S_{n_1} \times S_{n_2}) = n_1 - \max_i \{t_{i1} - t_{i2}\}.$$

For equal sample sizes, this is equivalent to the Kolmogorov-Smirnov test statistic

$$= \max_i \left\{ \frac{t_{i1}}{n_1} - \frac{t_{i2}}{n_2} \right\} = \max_x \left\{ \hat{F}_1(x) - \hat{F}_2(x) \right\}.$$

Proof: The derivations of these results, and of the corresponding results in succeeding sections, are deferred to the Appendix. We simply remark here that some elementary group-theoretic and combinatorial properties are helpful to the arguments. □

Remarks: It is well known that the Mann-Whitney test statistic and the Wilcoxon test statistic are equivalent. Yet they are induced by different metrics — namely, K and F, respectively. This is not particularly surprising, since it is entirely possible for two different metrics to produce the same minimum interpoint distance between the sets $[\pi]$ and E.

The test statistic induced by Spearman's rho has not appeared previously in the literature, as far as the author is aware. However, it bears a strong resemblance to one form of the Cramér-von Mises test statistic

$$n_1 \sum_{i=1}^{n_1} (a_i - i)^2 + n_2 \sum_{i=n_1+1}^{n} (a_i - (i - n_1))^2$$

(Hájek and Sidák (1967), page 93), and also to Page's (1963) test for the two-way layout with ordered alternatives. This latter resemblance is not a coincidence, as Section 8 demonstrates.

Of course, although the test statistics of Theorem 1 are among the most well known, several other rank test statistics have been proposed for the two-sample location problem. As discussed in Sections 13.2 and 13.3, nearly all such statistics are induced by our basic construction, by suitable metrics on permutations.

*　　　　　*　　　　　*

The basic construction carries over to other hypothesis testing situations. As a result, the construction yields not just a collection of test statistics, but also a collection of *families* of test statistics. For example, since Ulam's metric U induces the Kolmogorov-Smirnov statistic for the two-sample location problem with ordered alternatives, we may regard the statistics induced by U in other situations as natural extensions of the Kolmogorov-Smirnov statistic.

For each testing situation, the equivalence class $[\pi]$ and the extremal set E must be identified, before computing the minimum interpoint distance between these two sets.

5. The Two-Sample Location Problem, with Unordered Alternatives

For the two-sample location problem with unordered alternatives, the equivalence class $[\pi] = \pi S_{n_1} \times S_{n_2}$ is the same as for the ordered alternatives case. However, the extremal set E must be expanded to include not only those permutations which rank all of population 1 before population 2, but also those permutations which rank all of population 2 before population 1. In group-theoretic notation, $E = S_2 \odot (S_{n_1} \times S_{n_2})$, where, as in Section 4,

$$S_{n_1} \times S_{n_2} = \{\sigma \epsilon S_n: \ \sigma(i) < \sigma(j) \ \forall \ i \epsilon N_1 \ \forall \ j \epsilon N_2\} \equiv \{\sigma \epsilon S_n: \ \sigma(N_1) < \sigma(N_2)\},$$

and where the "dot product" $S_2 \odot (S_{n_1} \times S_{n_2})$ is defined as $\{\beta \in S_n\colon \exists\, \sigma \in S_2 \colon \beta(N_{\sigma(1)}) < \beta(N_{\sigma(2)})\}$. (Here and henceforth, the notation $\Omega_1 < \Omega_2$, where Ω_1 and Ω_2 are *sets* of numbers, means that $\omega_1 < \omega_2$ for all $\omega_1 \in \Omega_1$ and $\omega_2 \in \Omega_2$.)

Our basic construction now gives rise to the usual two-sided versions of the test statistics presented in Section 4:

Theorem 2: The test statistics for the two-sample location problem with unordered alternatives, induced by the metrics of Section 3, are:

$$K(\pi S_{n_1} \times S_{n_2}, S_2 \odot S_{n_1} \times S_{n_2}) = \frac{n_1 n_2}{2} - |\#\{(i,j) \epsilon N_1 \times N_2\colon\ \pi(i) > \pi(j)\} - \frac{n_1 n_2}{2}|.$$

This is equivalent to the Mann-Whitney (1947) two-sided test statistic $= |\#\{(i,j)\epsilon N_1 \times N_2\colon\ \pi(i) > \pi(j)\} - n_1 n_2/2|$.

$$F(\pi S_{n_1} \times S_{n_2}, S_2 \odot S_{n_1} \times S_{n_2}) = n_1 n_2 - 2|\sum_{i=1}^{n_1} \pi(i) - \frac{n_1(n+1)}{2}|.$$

This equivalent to the Wilcoxon (1945) two-sided test statistic $= |\sum_{i=1}^{n_1} \pi(i) - n_1(n+1)/2|$.

$$R^2(\pi S_{n_1} \times S_{n_2}, S_2 \odot S_{n_1} \times S_{n_2})$$
$$= \frac{n(n+1)(2n+1)}{3} - 2\sum_{i=1}^{n} i a_i - 2\max\left\{0, n_2 \sum_{i=1}^{n_1} \pi(i) - n_1 \sum_{i=n_1+1}^{n} \pi(i)\right\}.$$

This is the two-sided version of the new test statistic introduced in Section 4.

$$H(\pi S_{n_1} \times S_{n_2}, S_2 \odot S_{n_1} \times S_{n_2}) = 2\min\{\#(\pi(N_1) \cap N_2), \#(\pi(N_1) \cap \{1, \ldots, n_2\})\}.$$

For equal sample sizes ($n_1 = n_2 = n/2$), this is equivalent to the Mood (1950) two-sided "median test" statistic

$$= |\#(\pi(N_1) \cap \left\{\frac{n}{2}+1, \ldots, n\right\}) - \frac{n}{2}|.$$

$$U(\pi S_{n_1} \times S_{n_2}, S_2 \odot S_{n_1} \times S_{n_2}) = \frac{n}{2} - \max_i |\frac{n_2 - n_1}{2} + t_{i1} - t_{i2}|.$$

For equal sample sizes, this is equivalent to the Kolmogorov-Smirnov two-sided test statistic

$$= \max_i \left|\frac{t_{i1}}{n_1} - \frac{t_{i2}}{n_2}\right| = \max_x |\hat{F}_1(x) - \hat{F}_2(x)|.$$

Proof: See the Appendix. □

6. The Multi-Sample Location Problem, with Ordered Alternatives

This is a straightforward generalization of the two-sample situation considered in Section 4. Suppose we have samples from $r > 2$ populations, with sample sizes $n_1, \ldots, n_r$, where $\sum_{j=1}^{r} n_j = n$. Let F_j be the cumulative distribution function for the j-th population. The null hypothesis is H_0: $F_1(x) \equiv \ldots \equiv F_r(x)$, and one possible formulation of the alternative is H_1: $F_1(x) \geq F_2(x) \geq \ldots \geq F_r(x)$, where each inequality is strict for some x.

We proceed to implement Steps 1 through 4 of the basic construction.

Step 1: The permutation π is the observed rank ordering of all n items, with the conventions: items $1, \ldots, n_1$ are from the first population, items $n_1 + 1, \ldots, n_1 + n_2$ are from the second population, and so on. Let $N_1, N_2, \ldots, N_r$ be the corresponding sets of labels:

$$N_1 = \{1, \ldots, n_1\},$$

$$N_2 = \{n_1 + 1, \ldots, n_1 + n_2\}$$

$$\cdot$$

$$\cdot$$

$$\cdot$$

$$N_r = \{n_1 + \ldots + n_{r-1} + 1, \ldots, n_1 + \ldots + n_{r-1} + n_r\}.$$

Step 2: The equivalence class $[\pi] = \{\alpha \epsilon S_n\colon\ \alpha(N_j) = \pi(N_j) \forall\ j = 1, \ldots, r\}$ consists of all permutations α which assign the same set of ranks, that π does, to each of the r

populations. Thus $[\pi]$ is the left coset $\pi S_{n_1} \times \ldots \times S_{n_r}$, where $S_{n_1} \times \ldots \times S_{n_r}$ is the subgroup $\{\sigma \in S_n\colon \sigma(N_j) = N_j \ \forall\, j = 1, \ldots, r\}$, or, equivalently, where $S_{n_1} \times \ldots \times S_{n_r} = \{\sigma \in S_n\colon \sigma(N_1) < \sigma(N_2) < \ldots < \sigma(N_r)\}$.

Step 3: The extremal set E is $S_{n_1} \times \ldots \times S_{n_r}$. In words, E consists of the permutations which rank all items in the first population before all items in the second population, which in turn are ranked ahead of everything in the third population, etc.

Step 4: To present the minimum interpoint distances between the sets $[\pi]$ and E, it is necessary to further generalize our previous notation.

Notation: Let $a_1 < \ldots < a_{n_1}$ be the ranks assigned by π to the first population, $a_{n_1+1} < \ldots < a_{n_1+n_2}$ be the ranks assigned by π to the second population, and so on.

For $i = 1, \ldots, n$ and $j = 1, \ldots, r$, let $t_{ij} = \#\{k \in N_j\colon \pi(k) \leq i\}$ be the number of items from the j-th population, which fall among the top i ranks.

Let $\hat{F}_j(x)$ be the empirical cumulative distribution function for the j-th population.

Finally, $\sum\limits_{j<k}$ denotes summation over all ordered pairs (j, k) such that $j < k$ (and *not* $\sum\limits_{j=1}^{k-1}$).

Theorem 3: The test statistics for the r-sample location problem with ordered alternatives, induced by the metrics of Section 3, are:

$$K(\pi S_{n_1} \times \ldots \times S_{n_r}, S_{n_1} \times \ldots \times S_{n_r}) = \sum_{j<k} \#\{(i,m) \in N_j \times N_k\colon \pi(i) > \pi(m)\}\,.$$

This is the Jonckheere (1954)-Terpstra (1952) test statistic.

$$F(\pi S_{n_1} \times \ldots \times S_{n_r}, S_{n_1} \times \ldots \times S_{n_r}) = \sum_{i=1}^{n} |a_i - i|.$$

$$R^2(\pi S_{n_1} \times \ldots \times S_{n_r}, S_{n_1} \times \ldots \times S_{n_r}) = \sum_{i=1}^{n}(a_i - i)^2 = \frac{n(n+1)(2n+1)}{3} - 2\sum_{i=1}^{n} i a_i.$$

$$H(\pi S_{n_1} \times \ldots \times S_{n_r}, S_{n_1} \times \ldots \times S_{n_r}) = n - \sum_{j=1}^{r} \#(\pi(N_j) \cap N_j).$$

$$U(\pi S_{n_1} \times \ldots \times S_{n_r}, S_{n_1} \times \ldots \times S_{n_r}) = (n - n_r) - \max_{i_1 \leq \ldots \leq i_{r-1}} \sum_{j=1}^{r-1}(t_{i_j j} - t_{i_j j+1}).$$

Proof: See the Appendix. □

Remarks: The Jonckheere-Terpstra statistic is usually viewed as the extension of the Mann-Whitney statistic to the multi-sample situation. Similarly, the statistic induced by F is a natural extension of the Wilcoxon test statistic, and the statistics induced by R^2, H, and U are natural extensions of the corresponding statistics in Section 4. In particular:

(1) The Mann-Whitney and Wilcoxon tests are not equivalent for more than two samples.

(2) For equal sample sizes from $r > 2$ populations, the statistic induced by H is equivalent to the following extension of Mood's median test statistic:

$$\sum_{j=1}^{r} \#(\pi(N_j) \cap N_j).$$

(3) For equal sample sizes from $r > 2$ populations, the statistic induced by U is equivalent to the following extension of the Kolmogorov-Smirnov test statistic:

$$\max_{i_1 \leq \ldots \leq i_{r-1}} \sum_{j=1}^{r-1} \left(\frac{t_{i_j j}}{n_j} - \frac{t_{i_j j+1}}{n_{j+1}} \right) = \max_{x_1 \leq \ldots \leq x_{r-1}} \sum_{j=1}^{r-1} (\hat{F}_j(x_j) - \hat{F}_{j+1}(x_j)).$$

7. The Multi-Sample Location Problem, with Unordered Alternatives

For the unordered alternatives case, the equivalence class $[\pi] = \pi S_{n_1} \times \ldots \times S_{n_r}$ is the same as in the preceding section, but the extremal set E now includes all of the extremal permutations for *all* of the $r!$ different possible ordered alternatives. Thus $E = S_r \odot (S_{n_1} \times \ldots \times S_{n_r})$, where the dot product $S_r \odot (S_{n_1} \times \ldots \times S_{n_r})$ equals $\{\beta \in S_n \colon \exists\, \sigma \in S_r \colon \beta(N_{\sigma(1)}) < \ldots < \beta(N_{\sigma(r)})\}$.

The metrics K, F, R, H, and U all induce very natural competitors to the Kruskal-Wallis (1952) test statistic. Each of these competitors is a test statistic which measures the distance to the *closest* ordered alternative. The test statistic is thus a minimum of test statistics of the type presented in the preceding section, with the minimization being done over all the $r!$ possible ordered alternatives. In the statement of Theorem 4, this amounts to minimizing with respect to a permutation $\sigma \in S_r$.

Notation: For $j = 1, \ldots, r$ and $i = 1, \ldots, n_j$, let $a_{1j} < \ldots < a_{n_j j}$ be the ranks assigned by π to population j.

Theorem 4: The test statistics for the r-sample location problem with unordered alternatives, induced by the metrics of Section 3, are:

$$K(\pi S_{n_1} \times \ldots \times S_{n_r}, S_r \odot S_{n_1} \times \ldots \times S_{n_r}) = \min_{\sigma \in S_r} \sum_{j<k} \#\{(i,m) \in N_{\sigma(j)} \times N_{\sigma(k)} \colon \pi(i) > \pi(m)\}.$$

$$F(\pi S_{n_1} \times \ldots \times S_{n_r}, S_r \odot S_{n_1} \times \ldots \times S_{n_r}) = \min_{\sigma \in S_r} \sum_{j=1}^{r} \sum_{i=1}^{n_{\sigma(j)}} |a_{i\sigma(j)} - (i + \sum_{k=1}^{j-1} n_{\sigma(k)})|.$$

$$R^2(\pi S_{n_1} \times \ldots \times S_{n_r}, S_r \odot S_{n_1} \times \ldots \times S_{n_r}) = \min_{\sigma \in S_r} \sum_{j=1}^{r} \sum_{i=1}^{n_{\sigma(j)}} (a_{i\sigma(j)} - (i + \sum_{k=1}^{j-1} n_{\sigma(k)}))^2$$

$$= \frac{n(n+1)(2n+1)}{3} - 2 \max_{\sigma \in S_r} \sum_{j=1}^{r} \sum_{i=1}^{n_{\sigma(j)}} a_{i\sigma(j)} (i + \sum_{k=1}^{j-1} n_{\sigma(k)}).$$

$$H(\pi S_{n_1} \times \ldots \times S_{n_r}, S_r \odot S_{n_1} \times \ldots \times S_{n_r})$$
$$= n - \max_{\sigma \epsilon S_r} \sum_{j=1}^{r} \#\{i \epsilon N_{\sigma(j)}: \sum_{k=1}^{j-1} n_{\sigma(k)} < \pi(i) \leq \sum_{k=1}^{j} n_{\sigma(k)}\}.$$
$$U(\pi S_{n_1} \times \ldots \times S_{n_r}, S_r \odot S_{n_1} \times \ldots \times S_{n_r})$$
$$= n - \max_{\substack{i_1 \leq \ldots \leq i_{r-1} \\ \sigma \epsilon S_r}} \{n_{\sigma(r)} + \sum_{j=1}^{r-1} (t_{i_j \sigma(j)} - t_{i_j \sigma(j+1)})\}.$$

Proof: See the Appendix. □

Remarks: (1) With the aid of modern computing facilities and suitable minimization algorithms, it should be feasible to calculate any of the preceding test statistics, even for moderately large r.

(2) For equal sample sizes from $r > 2$ populations, the statistic induced by U is equivalent to the following extension of the Kolmogorov-Smirnov statistic:

$$\max_{\substack{x_1 \leq \ldots \leq x_{r-1} \\ \sigma \epsilon S_r}} \sum_{j=1}^{r-1} (\hat{F}_{\sigma(j)}(x_j) - \hat{F}_{\sigma(j+1)}(x_j)).$$

8. The Two-Way Layout, with Ordered Alternatives

Suppose there are b blocks, r populations, and r items in each block (one from each population). Let X_{ij} be the observation from the i-th block and the j-th population. The two-way layout model is

$$X_{ij} = \mu + \gamma_i + \tau_j + \varepsilon_{ij}, \qquad \text{for } i = 1, \ldots, b \text{ and } j = 1, \ldots, r,$$

where μ is the unknown overall mean, γ_i is the i-th block effect, τ_j is the j-th population effect, the ε_{ij} are i.i.d. noise variables, and $\sum_{i=1}^{b} \gamma_i = 0 = \sum_{j=1}^{r} \tau_j$.

For the ordered alternatives case, we are testing H_0: $\tau_1 = \ldots = \tau_r$ versus H_1: $\tau_1 \leq \ldots \leq \tau_r$, where at least one of the inequalities is strict. We proceed to implement Steps 1 through 4 of the basic construction.

Step 1: The observed permutation π is the ranking of *all* the $n = rb$ items. We make the convention that item $(i-1)r+j$ is from block i and population j, for $i = 1, \ldots, b$ and $j = 1, \ldots, r$. As usual, $\pi((i-1)r+j)$ is the rank given to this item.

Step 2: For the two-way layout, two permutations π, $\alpha \in S_n$ are equivalent if and only if, within each block, they induce the same permutation of the r populations. Formally, the permutation $\pi_i \in S_r$ within the i-th block, induced by π, is defined by

$$\pi_i(j) = \#\{k = 1, \ldots, r\colon\ \pi((i-1)r+k) \leq \pi((i-1)r+j)\}$$

for $i = 1, \ldots, b$ and $j = 1, \ldots, r$. Similarly, let $\alpha_i \in S_r$ be the permutation within the i-th block induced by α. Then $\pi, \alpha \in S_n$ are equivalent if and only if $(\pi_1, \ldots, \pi_r) = (\alpha_1, \ldots, \alpha_r)$, and the equivalence class $[\pi]$ is $\{\alpha \in S_n\colon\ (\pi_1, \ldots, \pi_r) = (\alpha_1, \ldots, \alpha_r)\}$.

Step 3: The extremal set E is the equivalence class containing the identity permutation e. Thus $E = [e]$ consists of all permutations which, within each block, rank population 1 first, population 2 second, and so on.

Step 4: For each of the metrics $d = K, F, R^2$, and U, the test statistic induced by d turns out to be the sum of the within-block distances; that is,

$$d([\pi],[e]) = \sum_{i=1}^{b} d(\pi_i, e).$$

(There is some abuse of notation here: d on the left hand side is a metric on S_n, whereas d on the right hand side is the corresponding metric on S_r. Also e on the left hand side is the identity permutation in S_n, whereas e on the right is the identity permutation in S_r. Finally, we are referring to R^2 as a "metric," even though it does not satisfy the triangle inequality.)

The induced test statistics are given by Theorem 5 below.

Notation: We introduce the abbreviation "LLIS" for the phrase "length of the longest increasing subsequence."

Theorem 5: The test statistics for the two-way layout with ordered alternatives, induced by the metrics K, F, R, and U of Section 3, are:

$$K([\pi],[e]) = \sum_{i=1}^{b} K(\pi_i, e)$$

$$= \sum_{i=1}^{b} \#\{(j,k) \in \{1,\dots,r\}^2 \colon j < k \text{ and } \pi_i(j) > \pi_i(k)\}.$$

$$F([\pi],[e]) = \sum_{i=1}^{b} F(\pi_i, e) = \sum_{i=1}^{b}\sum_{j=1}^{r} |\pi_i(j) - j|$$

$$R^2([\pi],[e]) = \sum_{i=1}^{b} R^2(\pi_i, e) = \sum_{i=1}^{b}\sum_{j=1}^{r} (\pi_i(j) - j)^2$$

$$= \frac{br(r+1)(2r+1)}{3} - 2\sum_{j=1}^{r} j \sum_{i=1}^{b} \pi_i(j).$$

This is equivalent to the Page (1963) test statistic $\sum_{j=1}^{r} j \sum_{i=1}^{b} \pi_i(j)$.

$$U([\pi],[e]) = \sum_{i=1}^{b} U(\pi_i, e) = n - \sum_{i=1}^{b} LLIS\{\pi_i(1),\dots,\pi_i(r)\}.$$

Proof: See the Appendix. □

Remarks: For other hypothesis testing situations, the test statistics induced by R — in particular, the new two-sample location test statistic introduced in Section 4 — can be regarded now as extensions of Page's test. Similarly, the new tests in the present section are extensions of previous tests.

The statistic induced by the metric H is not given in Theorem 5, and indeed, has not yet been determined. In future sections of this paper, analogous omissions will have a similar meaning.

9. The Two-Way Layout, with Unordered Alternatives

For the unordered alternatives case, the equivalence class $[\pi]$ is the same, but the extremal set E consists of all permutations which, within each block, rank the populations in the same order. Thus E is the union of $r!$ equivalence classes, one for each possible ranking $\sigma \in S_r$ of the r populations. Mathematically,

$$E = S_r \odot [e] \equiv \{\beta \in S_n\colon \exists\, \sigma \in S_r\colon \beta_i = \sigma \;\forall\, i = 1, \ldots, b\}.$$

The metrics K, F, R, and U induce new test statistics, which are extensions of previous tests, and also are natural competitors to the Friedman (1937) - Kendall-Babington Smith (1939) test. Just as for the multi-sample location problem, these statistics measure the distance to the closest ordered alternative. Since there are $r!$ possible ordered alternatives, the expressions for the induced test statistics entail a minimization with respect to a permutation $\sigma \in S_r$.

Theorem 6: The test statistics for the two-way layout with unordered alternatives, induced by the metrics K, F, R, and U of Section 3, are:

$$\begin{aligned} K([\pi], S_r \odot [e]) &= \min_{\sigma \in S_r} \sum_{i=1}^{b} K(\pi_i, \sigma) \\ &= \min_{\sigma \in S_r} \sum_{i=1}^{b} \#\{(j,k) \in \{1, \ldots, r\}^2\colon\ \sigma(j) < \sigma(k) \text{ and } \pi_i(j) > \pi_i(k)\}. \\ F([\pi], S_r \odot [e]) &= \min_{\sigma \in S_r} \sum_{i=1}^{b} F(\pi_i, \sigma) = \min_{\sigma \in S_r} \sum_{i=1}^{b} \sum_{j=1}^{r} |\pi_i(j) - \sigma(j)|. \end{aligned}$$

$$R^2([\pi], S_r \odot [e]) = \min_{\sigma \epsilon S_r} \sum_{i=1}^{b} R^2(\pi_i, \sigma)$$
$$= \min_{\sigma \epsilon S_r} \sum_{i=1}^{b} \sum_{j=1}^{r} (\pi_i(j) - \sigma(j))^2$$
$$= \frac{br(r+1)(2r+1)}{3} - 2 \max_{\sigma \epsilon S_r} \sum_{j=1}^{r} \sigma(j) \sum_{i=1}^{b} \pi_i(j).$$
$$U([\pi], S_r \odot [e]) = \min_{\sigma \epsilon S_r} \sum_{i=1}^{b} U(\pi_i, \sigma)$$
$$= n - \max_{\sigma \epsilon S_r} \sum_{i=1}^{b} LLIS\{\pi_i\sigma(1), \pi_i\sigma(2), \ldots, \pi_i\sigma(r)\}.$$

Proof: See the Appendix. □

Remarks: In the case of the statistic induced by R, the permutation $\sigma \epsilon S_r$ in the preceding theorem is determined by the rule:

$$\sigma(j) < \sigma(k) \text{ whenever } \sum_{i=1}^{b} \pi_i(j) < \sum_{i=1}^{b} \pi_i(k).$$

(If $\sum_{i=1}^{b} \pi_i(j) = \sum_{i=1}^{b} \pi_i(k)$, then the relative order of $\sigma(j)$ and $\sigma(k)$ may be chosen arbitrarily.)

For the other metrics, the permutation $\sigma \in S_r$ can be found numerically.

10. The One-Sample Location Problem

For the one-sample location problem, there are n_1 observations $X_1, \ldots, X_{n_1}$, which are assumed to follow the model $X_i = \mu + \varepsilon_i$. The "noise terms" ε_i are independent random variables, symmetric about 0. We now implement the four steps of the basic construction, considering initially the case of ordered alternatives. The hypotheses are H_0: $\mu \geq 0$ versus H_1: $\mu < 0$.

Step 1: The relative ordering of the observations $X_1, \ldots, X_{n_1}$ provides no information about the location parameter μ. So we create a superpopulation of $n = 2n_1 + 1$ observations, namely: $X_1, \ldots, X_{n_1}, 0, -X_1, \ldots, -X_{n_1}$. The relative ordering of these n items contains all of the information relevant to a rank test of H_0 versus H_1, and thus defines the desired permutation $\pi \in S_n$. For $i = 1, \ldots, n_1$, we make the conventions: observation X_i corresponds to item i, $-X_i$ corresponds to item $n+1-i$, and 0 corresponds to item $(n+1)/2 = n_1 + 1$. As usual, the items are ranked from the least to the greatest, and $\pi(i)$ is the rank given item i.

Step 2: The equivalence class $[\pi]$ consists of all permutations which give $X_1, \ldots, X_{n_1}$ the same set of ranks in the superpopulation as π (and similarly for $-X_1, \ldots, -X_{n_1}$). To obtain a group-theoretic description of $[\pi]$, let $N_1 = \{1, \ldots, n_1\}$ be the set of labels for the observations $X_1, \ldots, X_{n_1}$, let $N_2 = \{n_1 + 1\}$, and let $N_3 = \{n_1 + 2, \ldots, n\}$ be the set of labels for the observations $-X_{n_1}, \ldots, -X_1$. Then $[\pi]$ is the left coset $\pi \; S_{n_1} \times S_1 \times S_{n_1}$, where $S_{n_1} \times S_1 \times S_{n_1}$ is the subgroup $\{\sigma \in S_n \colon \sigma(N_j) = N_j \; \forall \, j = 1, 2, 3\}$.

Step 3: The extremal set E is the subgroup $S_{n_1} \times S_1 \times S_{n_1}$. Thus E consists of all permutations which assign the first n_1 ranks to $X_1, \ldots, X_{n_1}$ (in some order), the middle rank to 0, and the last n_1 ranks to $-X_1, \ldots, -X_{n_1}$. This means that for an extremal permutation, all $X_i < 0$.

Step 4: To describe the induced test statistics, it is helpful to introduce the following

Notation: For $\pi \in S_n$ and for $i = 1, \ldots, n_1$, define $\pi^*(i) = |\pi(i) - (n+1)/2|$. Then, for π constructed from the observations $X_1, \ldots, X_{n_1}$ as above, it follows that π^* is a permutation of $\{1, \ldots, n_1\}$. In fact, $\pi^*(i)$ is the rank of $|X_i|$ in the set $\{|X_1|, \ldots, |X_{n_1}|\}$.

Let $a_1 < \ldots < a_{n_1}$ be an enumeration of the set $\pi(N_1) = \pi\{1, \ldots, n_1\}$, let $a_{n_1+1} = n_1 + 1 = \pi(n_1 + 1)$, and let $a_{n_1+2} < \ldots < a_n$ be an enumeration of the set $\pi(N_3) = \pi\{n_1 + 2, \ldots, n\}$.

Let $\hat{F}_1(x)$ be the empirical cumulative distribution function for the data set $X_1, \ldots, X_{n_1}$, and let $\hat{F}_2(x)$ be the empirical c.d.f. for the data set $-X_1, \ldots, -X_{n_1}$.

Theorem 7: The test statistics for the one-sample location problem with ordered alternatives, induced by the metrics K, F, R, H, and U of Section 3, are:

$$K(\pi S_{n_1} \times S_1 \times S_{n_1}, S_{n_1} \times S_1 \times S_{n_1}) = 2 \sum_{\{i \epsilon N_1:\ X_i > 0\}} \pi^*(i).$$

This is equivalent to the Wilcoxon (1945) signed rank test statistic $= \sum_{\{i \epsilon N_1:\ X_i > 0\}} \pi^*(i)$.

$$F(\pi S_{n_1} \times S_1 \times S_{n_1}, S_{n_1} \times S_1 \times S_{n_1}) = 4 \sum_{\{i \epsilon N_1:\ X_i > 0\}} \pi^*(i).$$

This is also equivalent to the Wilcoxon signed rank test statistic.

$$\begin{aligned} R^2(\pi S_{n_1} \times S_1 \times S_{n_1}, S_{n_1} \times S_1 \times S_{n_1}) &= \sum_{i=1}^{n} (a_i - i)^2 \\ &= \frac{n(n+1)(2n+1)}{3} - 2 \sum_{i=1}^{n} i a_i. \end{aligned}$$

This is equivalent to the statistic $\sum_{i=1}^{n} i a_i$.

$$\begin{aligned} H(\pi S_{n_1} \times S_1 \times S_{n_1}, S_{n_1} \times S_1 \times S_{n_1}) &= 2\#\{i \epsilon N_1:\ \pi(i) > \frac{n+1}{2}\} \\ &= 2\#\{i \epsilon N_1:\ X_i > 0\}. \end{aligned}$$

This is equivalent to the Fisher (1925) sign test statistic $= \#\{i \epsilon N_1:\ X_i > 0\}$.

$$\begin{aligned} & U(\pi S_{n_1} \times S_1 \times S_{n_1}, S_{n_1} \times S_1 \times S_{n_1}) \\ &\quad = \begin{cases} n_1 + 1 - n_1 \max_x\{\hat{F}_1(x) - \hat{F}_2(x)\} & \text{if } \max_x\{\hat{F}_1(x) - \hat{F}_2(x)\} < 1 \\ 0 & \text{if } \max_x\{\hat{F}_1(x) - \hat{F}_2(x)\} = 1. \end{cases} \end{aligned}$$

This is equivalent to the test statistic $\max_x\{\hat{F}_1(x) + \hat{F}_1(-x)\} - 1$ (assuming no ties). Thus the statistic $\max_x\{\hat{F}_1(x) + \hat{F}_1(-x)\}$ can be regarded as an extension of the Kolmogorov-Smirnov statistic to the one-sample location problem.

Proof: See the Appendix. □

For the case of unordered alternatives, the hypotheses are H_0: $\mu = 0$ versus H_1: $\mu \neq 0$. The equivalence class $[\pi]$ is the same as before, but the extremal set E should include the situation where all $X_i > 0$, as well as the case where all $X_i < 0$. Thus

$$E = S_2 \odot (S_{n_1} \times S_1 \times S_{n_1})$$

$$\equiv \{\beta \in S_n\colon \exists \text{ a bijection } \sigma\colon \{1,3\} \to \{1,3\}\colon \beta(N_{\sigma(j)}) = N_j \ \forall\, j = 1,3\}.$$

We obtain the natural two-sided versions of the previous tests:

Notation: For $i = 1, \ldots, n$, define $\tilde{a}_i$ by

$$\tilde{a}_i = \begin{cases} a_{n_1+1+i} & \text{for } i = 1, \ldots, n_1 \\ a_i = n_1 + 1 & \text{for } i = n_1 + 1 \\ a_{i-n_1-1} & \text{for } i = n_1 + 2, \ldots, n. \end{cases}$$

Theorem 8: The test statistics for the one-sample location problem with unordered alternatives, induced by the metrics K, F, R, H, and U of Section 3, are:

$$K(\pi S_{n_1} \times S_1 \times S_{n_1}, S_2 \odot S_{n_1} \times S_1 \times S_{n_1})$$
$$= 2\left\{\frac{n_1(n_1+1)}{4} - \left|\left(\sum_{X_i>0} \pi^*(i)\right) - \frac{n_1(n_1+1)}{4}\right|\right\}.$$

This is equivalent to the two-sided Wilcoxon signed rank test statistic $= \left|\left(\sum_{X_i>0} \pi^*(i)\right) - n_1(n_1+1)/4\right|$.

$$F(\pi S_{n_1} \times S_1 \times S_{n_1}, S_2 \odot S_{n_1} \times S_1 \times S_{n_1})$$
$$= 4\left\{\frac{n_1(n_1+1)}{4} - \left|\left(\sum_{X_i>0} \pi^*(i)\right) - \frac{n_1(n_1+1)}{4}\right|\right\}.$$

This is also equivalent to the two-sided Wilcoxon signed rank test statistic.

$$R^2(\pi S_{n_1} \times S_1 \times S_{n_1}, S_2 \odot S_{n_1} \times S_1 \times S_{n_1})$$
$$= \min\{\sum_{i=1}^{n} (a_i - i)^2, \sum_{i=1}^{n} (\tilde{a}_i - i)^2\}$$
$$= \frac{n(n+1)(2n+1)}{3} - 2\max\{\sum_{i=1}^{n} i a_i, \sum_{i=1}^{n} i \tilde{a}_i\}.$$

This is equivalent to the test statistic $\max\{\sum_{i=1}^{n} ia_i, \sum_{i=1}^{n} i\tilde{a}_i\}$.

$$H(\pi S_{n_1} \times S_1 \times S_{n_1}, S_2 \odot S_{n_1} \times S_1 \times S_{n_1})$$
$$= \{\frac{n_1}{2} - |\#\{i \in N_1\colon\ X_i > 0\} - \frac{n_1}{2}|\}.$$

This is equivalent to the two-sided Fisher sign test statistic $= |\#\{i \in N_1\colon\ X_i > 0\} - n_1/2|$.

$$U(\pi S_{n_1} \times S_1 \times S_{n_1}, S_2 \odot S_{n_1} \times S_1 \times S_{n_1})$$
$$= \begin{cases} n_1 + 1 - n_1 \max_x\{\hat{F}_1(x) - \hat{F}_2(x)\} & \text{if } \max_x\{\hat{F}_1(x) - \hat{F}_2(x)\} < 1 \\ 0 & \text{if } \max_x\{\hat{F}_1(x) - \hat{F}_2(x)\} = 1. \end{cases}$$

This is equivalent to the test statistic $\max_x |\hat{F}_1(x) - \hat{F}_2(x)| = \max_x |\hat{F}_1(x) + \hat{F}_1(-x) - 1|$.

Proof: See the Appendix. □

11. The Two-Sample Dispersion Problem, with Equal Medians

We consider initially the problem of ordered alternatives. Suppose the data come from two populations, with (population) cumulative distribution functions $F_1(x)$ and $F_2(x)$, respectively, each having the same median μ. Further assume that $F_2(x) \equiv F_1(\mu + (x-\mu)/\gamma)$ for some "dispersion parameter" $\gamma > 0$. The hypotheses to be tested are H_0: $\gamma \geq 1$ versus H_1: $\gamma < 1$. Thus the alternative hypothesis H_1 is that the second population is "less spread out" than the first.

Throughout this section we will assume that the two sample sizes, n_1 and n_2, are both *even numbers*. This assumption is made merely for the sake of notational convenience, and to illustrate the basic concepts. Similar results can be obtained when n_1 or n_2, or both, are odd numbers. These results are extensions of those presented below (and are just as intuitive), but it would be slightly messy to consider all of the different cases at once.

We proceed to implement the four steps of the basic construction.

Step 1: As for the two-sample location problem, let items $1, \ldots, n_1$ be from population 1, and items $n_1 + 1, \ldots, n_1 + n_2$ be from population 2. Let N_1 and N_2 be the sets

$N_1 = \{1,\ldots,n_1\}$ and $N_2 = \{n_1+1,\ldots,n_1+n_2\}$, and let $n = n_1 + n_2$. The observed permutation $\pi \in S_n$ is the rank ordering of all the n items.

Step 2: For this two-sample dispersion problem, a permutation $\alpha \in S_n$ is "equivalent" to the observed permutation π if and only if α can be obtained from π by a combination of the following two types of operations: (1) within each population, permuting the labels assigned to the items in that population; and (2) transposing the items ranked in positions i and $n+1-i$, for any $i = 1,\ldots,n$. Operations of type (1) have the same justification as was given in the two-sample location problem, namely, the equivalence class $[\pi]$ should depend only on the *sets* of ranks $\pi(N_1)$ and $\pi(N_2)$ assigned to the two populations. Operations of type (2) are justified by the assumption that the two populations have the same median.

To give a group-theoretic description to the equivalence class $[\pi]$, let $S_{n_1} \times S_{n_2}$ be the subgroup of S_n corresponding to operations of type (1) above:

$$S_{n_1} \times S_{n_2} = \{\sigma \in S_n\colon\ \sigma(N_j) = N_j \ \forall\, j = 1,2\},$$

and let T be the subgroup of S_n generated by the transpositions described above as operations of type (2):

$$T = \{\tau \in S_n\colon\ \forall\, i = 1,\ldots,n:\ \tau(i) = i \text{ or } \tau(i) = n+1-i\}.$$

The equivalence class $[\pi]$ consists of all permutations of the form $\tau\ \pi\ \sigma$, as τ ranges through T and σ ranges through $S_{n_1} \times S_{n_2}$. Hence $[\pi]$ is a so-called "double coset" (Curtis and Reiner (1966)), and is denoted by $T\ \pi\ S_{n_1} \times S_{n_2}$.

Step 3: The extremal set E consists of all permutations which rank the items from the second population "in the middle," and the items from the first population "at the two ends." Pictorially, E consists of all rankings of the form

$$\underbrace{X\cdots X}_{n_1/2} \qquad \underbrace{O\cdots O}_{n_2} \qquad \underbrace{X\cdots X}_{n_1/2},$$

where X's and O's denote items from populations 1 and 2, respectively, and where the left tail of X's and the right tail of X's have the same length. Thus

$$E = \{\beta \in S_n \colon |\beta(i) - \frac{n+1}{2}| \geq |\beta(j) - \frac{n+1}{2}| \; \forall\, i \in N_1 \; \forall\, j \in N_2\}$$
$$\equiv \{\beta \in S_n \colon |\beta(N_1) - \frac{n+1}{2}| \geq |\beta(N_2) - \frac{n+1}{2}|\}.$$

Clearly, E is the equivalence class $[\psi]$ containing the extremal permutation

$$\psi(i) = \begin{cases} i & \text{for } i = 1, \ldots, \frac{n_1}{2} \\ i + n_2 & \text{for } i = \frac{n_1}{2} + 1, \ldots, n_1 \\ i - \frac{n_1}{2} & \text{for } i = n_1 + 1, \ldots, n_1 + n_2. \end{cases}$$

Thus E is the double coset $T\ \psi\ S_{n_1} \times S_{n_2}$.

Step 4: To state the test statistics, we introduce some further

Notation: Let $b_1, \ldots, b_{n_1}$ be an enumeration of $\pi(N_1)$ such that

$$|b_1 - \frac{n+1}{2}| \geq \ldots \geq |b_{n_1} - \frac{n+1}{2}|,$$

and let $b_{n_1+1}, \ldots, b_n$ be an enumeration of $\pi(N_2)$ such that

$$|b_{n_1+1} - \frac{n+1}{2}| \geq \ldots \geq |b_n - \frac{n+1}{2}|.$$

For any real number r, let $\text{int}(r)$ denote the greatest integer which is less than or equal to r.

Theorem 9: The test statistics for the two-sample dispersion problem with equal medians and ordered alternatives, induced by the metrics K, F, R, and H of Section 3, are:

$$K(T\pi S_{n_1} \times S_{n_2}, T\psi S_{n_1} \times S_{n_2}) = \sum_{i=1}^{n_1} \min\{\pi(i), n+1-\pi(i)\} - \frac{n_1(n_1+2)}{4}.$$

This is equivalent to the Ansari-Bradley (1960) test statistic $= \sum_{i=1}^{n_1} \min\{\pi(i), n+1-\pi(i)\}$.

$$F(T\pi S_{n_1} \times S_{n_2}, T\psi S_{n_1} \times S_{n_2}) = 2K(T\pi S_{n_1} \times S_{n_2}, T\psi S_{n_1} \times S_{n_2}).$$

This is obviously also equivalent to the Ansari-Bradley statistic.

$$R^2(T\pi S_{n_1} \times S_{n_2}, T\psi S_{n_1} \times S_{n_2}) = \sum_{i=1}^{n} \left(\frac{n+1}{2} - \left| b_i - \frac{n+1}{2} \right| - \text{ int} \left(\frac{i+1}{2} \right) \right)^2 .$$

This is equivalent to the (new) test statistic $\sum_{i=1}^{n} \left(\text{int}\left(\frac{i+1}{2}\right) |b_i - \frac{n+1}{2}|\right)$.

$$H(T\pi S_{n_1} \times S_{n_2}, T\psi S_{n_1} \times S_{n_2}) = 2n_1 - 2\# \left\{ i \in N_1\colon \pi(i) \leq \frac{n_1}{2} \text{ or } \pi(i) \geq n+1-\frac{n_1}{2} \right\}.$$

For equal sample sizes ($n_1 = n_2 = n/2$), this is equivalent to the "quartile test statistic" (Westenberg (1948))

$$= \#\{i \in N_1\colon \pi(i) \leq \frac{n}{4} \quad \text{or} \quad \pi(i) \geq \frac{3n}{4} + 1\}.$$

Proof: See the Appendix. □

For the unordered alternatives problem, we still have $[\pi] = T\ \pi\ S_{n_1} \times S_{n_2}$, but now

$$\begin{aligned} E &= T\ S_2 \odot\ \psi\ S_{n_1} \times S_{n_2} \\ &\equiv \{\beta \in S_n\colon \exists\ \sigma \in S_2\colon |\beta(i) - \frac{n+1}{2}| \geq |\beta(j) - \frac{n+1}{2}|\ \forall\ i \in N_{\sigma(1)}\ \forall\ j \in N_{\sigma(2)}\} \\ &\equiv \{\beta \in S_n\colon \exists\ \sigma \in S_2\colon |\beta(N_{\sigma(1)}) - \frac{n+1}{2}| \geq |\beta(N_{\sigma(2)}) - \frac{n+1}{2}|\}. \end{aligned}$$

Notation: Define $\tilde{b}_1, \ldots, \tilde{b}_n$ by

$$\tilde{b}_i = \begin{cases} b_{i+n_1} & \text{for } i = 1, \ldots, n_2 \\ b_{i-n_2} & \text{for } i = n_2+1, \ldots, n. \end{cases}$$

Theorem 10: The test statistics for the two-sample dispersion problem with equal medians and unordered alternatives, induced by the metrics K, F, R, and H of Section 3, are:

$$K(T\pi S_{n_1} \times S_{n_2}, TS_2 \odot \psi S_{n_1} \times S_{n_2}) = \frac{n_1 n_2}{4} - \left| \left(\sum_{i=1}^{n_1} \min\{\pi(i), n+1-\pi(i)\} \right) - n_1 \left(\frac{n+2}{4} \right) \right|.$$

This is equivalent to the two-sided Ansari-Bradley statistic (which is the absolute value term in the above expression).

$$F(T\ \pi\ S_{n_1} \times S_{n_2}, T\ S_2\ \odot\ \psi\ S_{n_1} \times S_{n_2}) = 2K(T\ \pi\ S_{n_1} \times S_{n_2}, T\ S_2\ \odot\ \psi\ S_{n_1} \times S_{n_2}).$$

This is obviously also equivalent to the two-sided Ansari-Bradley statistic.

$$\begin{aligned} &R^2(T\ \pi\ S_{n_1} \times S_{n_2}, T\ S_2\ \odot\ \psi\ S_{n_1} \times S_{n_2}) \\ &\quad = \min\left\{ \sum_{i=1}^{n} \left(\frac{n+1}{2} - \left| b_i - \frac{n+1}{2} \right| - \text{int}\left(\frac{i+1}{2}\right) \right)^2, \right. \\ &\qquad\quad \left. \sum_{i=1}^{n} \left(\frac{n+1}{2} - \left| \tilde{b}_i - \frac{n+1}{2} \right| - \text{int}\left(\frac{i+1}{2}\right) \right)^2 \right\}. \end{aligned}$$

This is equivalent to the statistic

$$\min\{\sum_{i=1}^{n} \left(\text{int}\left(\frac{i+1}{2}\right) \left| b_i - \frac{n+1}{2} \right| \right), \sum_{i=1}^{n} \left(\text{int}\left(\frac{i+1}{2}\right) \left| \tilde{b}_i - \frac{n+1}{2} \right| \right)\},$$

which is the two-sided version of the corresponding statistic introduced in Theorem 9.

$$\begin{aligned} &H(T\ \pi\ S_{n_1} \times S_{n_2}, T\ S_2\ \odot\ \psi\ S_{n_1} \times S_{n_2}) \\ &\quad = 2 \min_{j=1,2} \{n_j\ -\ \#\{i \in N_j\colon\ \pi(i) \le \frac{n_j}{2} \quad \text{or} \quad \pi(i) \ge n+1 - \frac{n_j}{2}\}\}. \end{aligned}$$

For equal sample sizes, this equals $n_1 - 2|\ \#\{i \epsilon\ N_1\colon\ \pi(i) \le n/4 \quad \text{or} \quad \pi(i) \ge 3n/4 + 1\} - n_1/2|$, which is equivalent to the two-sided quartile test statistic

$$= |\ \#\{i \epsilon\ N_1\colon\ \pi(i) \le \frac{n}{4} \text{ or } \pi(i) \ge \frac{3n}{4} + 1\} - \frac{n_1}{2}|$$

Proof: See the Appendix. □

The extension to the r-sample dispersion problem ($r > 2$), with equal medians assumed, and with either ordered or unordered alternatives, is relatively straightforward. These test statistics will appear in a future paper, along with investigations of their statistical properties.

12. Testing for Trend

We conclude with a very trivial application of the basic construction, to the problem of testing for trend. Suppose item i is from population i with cumulative distribution function F_i. We are testing H_0: $F_1(x) \equiv \ldots \equiv F_n(x)$ versus H_1: $F_1(x) \geq \ldots \geq F_n(x)$ (where each inequality is strict for some x). Clearly the equivalence class $[\pi]$ consists only of the observed ranking π itself, the extremal set E is just the identity permutation e, and the test statistic induced by the metric d is $d(\pi, e)$. This simple case has been mentioned in order to show that the procedure "works" here, and also to show that measuring the distance from H_1 produces a very natural test statistic.

13. Some Connections with Previous Work

Many important questions are concerned with relating this novel method to the existing nonparametric theory. We will consider four such potential areas of interconnection: (1) the theory of maximal invariants, (2) the theory of locally most powerful rank tests and linear rank tests, (3) the so-called "inverse problem," and (4) partially ranked data. To keep the discussion manageable, we will consider each of these topics in relation to the two-sample location problem (and will therefore use the notation of Sections 4 and 5).

13.1 Maximal Invariants. The theory of maximal invariants (as presented, for example, in Lehmann (1959)) may, in some cases, yield a more precise formulation of the set $[\pi]$ described in Step 2 of the basic construction. Indeed, for the two-sample location problem, the maximal invariants are the set $\pi\{1, \ldots, n_1\}$ of ranks assigned to the first population. On the other hand, the equivalence class $[\pi]$ consists of all permutations α such that $\alpha\{1, \ldots, n_1\} = \pi\{1, \ldots, n_1\}$. Thus, for the two-sample location problem, the equivalence class $[\pi]$ consists of all permutations which produce the same values of the maximal invariants as π.

13.2 Locally Most Powerful Rank Tests, and Linear Rank Tests. Another natural area of inquiry is the relation with the theory of locally most powerful rank tests, as presented, for example, in Hájek and Sidák (1967). In this regard, consider the two-sample location problem with ordered alternatives, for which the hypotheses are H_0: $g(x) \equiv f(x)$ versus H_1: $g(x) \equiv f(x - \Delta)$ for some $\Delta > 0$. Here $f(x)$ and $g(x)$ are the underlying densities for the first and second populations, respectively; they are assumed to be absolutely continuous and to satisfy $\int_{-\infty}^{\infty} |f'(x)|\, dx < \infty$.

Let $\pi \in S_n$ be the observed ranking of all $n = n_1 + n_2$ items, as discussed in Section 4. Then for $\Delta > 0$ sufficiently small, the locally most powerful rank test (Hájek and Sidák, pages 67–68) rejects H_0 for small values of the test statistic

$$\sum_{i=1}^{n_1} E\left\{-\frac{f'\left(X^{(\pi(i))}\right)}{f\left(X^{(\pi(i))}\right)}\right\},$$

where $X^{(1)} < \ldots < X^{(n)}$ are the order statistics for a random sample of size n from the density f.

It is natural to ask whether this locally most powerful test statistic is, in fact, induced by our basic construction, by some metric on permutations. More generally, the locally most powerful rank test statistics are a subclass of the so-called "linear rank test statistics" of the form $\sum_{i=1}^{n_1} h(\pi(i))$, and we could ask whether any such linear statistic is induced by a metric.

The answer to both questions turns out to be yes.

Notation: Define the function h_f: $\{1, \ldots, n\} \to \mathbb{R}$ by

$$h_f(j) = E\left\{-\frac{f'(X^{(j)})}{f(X^{(j)})}\right\} \qquad \text{for } j = 1, \ldots, n.$$

Hájek and Sidák called these quantities the "scores" generated by the density f; we assume henceforth that the $h_f(j)$ are strictly increasing in j. Define the metric d_f on S_n by

$$d_f(\pi, \sigma) = \sum_{i=1}^{n_1} |h_f(\pi(i)) - h_f(\sigma(i))|.$$

Theorem 11: For the two-sample location problem with ordered alternatives, the test statistic induced by the metric d_f is

$$d_f(\pi S_{n_1} \times S_{n_2}, S_{n_1} \times S_{n_2}) = 2\sum_{i=1}^{n_1} h_f(\pi(i)) - 2\sum_{i=1}^{n_1} h_f(i),$$

and thus is equivalent to the locally most powerful rank test statistic $\sum_{i=1}^{n_1} h_f(\pi(i))$ for testing the hypotheses H_0: $g(x) \equiv f(x)$ versus H_1: $g(x) \equiv f(x - \Delta)$.

Proof: Entirely analogous to the proof of Theorem 1 for the Spearman's footrule metric F. In that proof (see the Appendix), substitute $h_f(\pi(i))$ for $\pi(i)$ everywhere, and $h_f(i)$ for i. □

Remarks: (1) The metrics d_f which yield the locally most powerful rank tests are a subclass of the so-called "fixed vector metrics," studied, in other statistical contexts, by Rukhin (1972), Diaconis (1982), and Critchlow (1985, pages 27–30).

(2) The same argument shows that an arbitrary linear rank statistic $\sum_{i=1}^{n_1} h(\pi(i))$ is induced by the metric $d(\pi, \sigma) = \sum_{i=1}^{n_1} |h(\pi(i)) - h(\sigma(i))|$.

13.3. The "Inverse Problem". Given any locally most powerful rank test statistic for the two-sample location problem, there exists a metric on permutations which induces it, and that metric was presented in Section 13.2. The natural generalization of this line of inquiry is the so-called "inverse problem"; namely, given an arbitrary rank test statistic, we can ask whether it is induced by some metric. Theorem 11 of Section 13.2 solves the inverse problem for the locally most powerful statistics for the two-sample location problem, and more generally for all linear rank test statistics. However, for the two-sample location problem, there are important test statistics which are not locally most powerful against any alternative, and which are not linear, such as the Kolmogorov-Smirnov test, the Wald-Wolfowitz (1940) "runs test", the so-called "tests based on exceeding observations" (Hájek and Sidák, pages 89–90), and Lehmann's (1951) "quadruples test". We now discuss the inverse problem for these tests.

Of course, a principal reason for posing the inverse problem for any test statistic is that if a corresponding metric can be found, then that metric can be used to extend the given test statistic to all other hypothesis testing problems.

By Theorem 1 of Section 4, Ulam's metric U induces the Kolmogorov-Smirnov test statistic, for equal sample sizes. The following result shows that the Wald-Wolfowitz test statistic is also induced by a metric.

Definition: Let $\pi \in S_n$ be the observed ranking of the n_1 items from population 1 and the n_2 items from population 2, where $n_1 + n_2 = n$. Define the number of "runs" in π to be

$$\#\{i = 1,\ldots,n_1\colon\ \pi^{-1}(\pi(i)+1) > n_1\} + \#\{i = n_1+1,\ldots,n_1+n_2\colon\ \pi^{-1}(\pi(i)+1) \le n_1\},$$

with the convention that if $\pi(i) = n$, then i contributes 1 to the number of runs. Thus, if X's and O's denote the items from populations 1 and 2, respectively, then

$$\#\text{ runs } = \#(X\text{'s followed by an } O) + \#(O\text{'s followed by an } X) + 1.$$

The Wald-Wolfowitz "runs test" statistic is precisely the number of runs in π.

Theorem 12: For $\pi, \sigma \in S_n$, define

$$W(\pi,\sigma) \equiv \#\{i = 1,\ldots,n-1\colon\ \pi\sigma^{-1}(i+1) \ne \pi\sigma^{-1}(i)+1\}.$$

Then W is a metric on S_n. Moreover, for the two-sample location problem with unordered alternatives, the test statistic induced by W is

$$W(\pi S_{n_1} \times S_{n_2}, S_2 \odot S_{n_1} \times S_{n_1}) = \#(\text{runs in } \pi) - 2,$$

which is equivalent to the Wald-Wolfowitz "runs test" statistic.

Proof: See the Appendix. □

Remarks: (1) We could call $W(\pi, \sigma)$ the "successor metric" because it counts the number of items, among those given the top $n-1$ ranks by σ, which have a different successor under the ranking π than they do under σ.

(2) The metric $W(\pi, \sigma)$ can now be used, in the context of our basic construction, to find "natural extensions" of the Wald-Wolfowitz runs test to all of the other hypothesis testing situations considered in this paper. These extensions will be derived and discussed in a future paper.

(3) It also turns out that all of the "tests based on exceeding observations" (Hájek and Sidák, pages 89–90) are induced by suitable metrics, as is Lehmann's (1951) one-sided "quadruples test" statistic

$$= \#\{(i,j,k,l) \in N_1 \times N_1 \times N_2 \times N_2\colon\ \pi(i) < \pi(j) < \pi(k) < \pi(l)\}.$$

Therefore, these tests also have natural extensions to other testing situations. A subsequent paper will introduce the corresponding metrics and investigate the extensions.

(4) The author is not aware of any procedure which solves the inverse problem in general. Rather, solutions to the inverse problem have thus far been obtained only on a case by case basis. In the case of the Wald-Wolfowitz test, for example, the author just happened to observe that the metric $W(\pi, \sigma)$ "worked".

(5) In Section 13.2, it was shown that all locally most powerful statistics, and indeed all linear statistics, are induced by metrics. Moreover, the present section demonstrates that the Kolmogorov-Smirnov and Wald-Wolfowitz statistics are also induced by metrics,

even though they are *not* linear. Hence, our basic construction actually gives rise to a larger class of test statistics, for the two-sample location problem.

13.4 Partially Ranked Data, and the Hausdorff Metric.

A *full* ranking of n items is simply an ordering of all these items, of the form: first choice, second choice, ..., n-th choice. If two judges each rank the same n items, statisticians and other scientists have used the various metrics of Section 3 to measure the closeness of the two rankings.

A *partial* ranking arises when each judge specifies only his first k choices, where $k < n$. The author's recent monograph (Critchlow (1985)) is concerned with extending the metrics of Section 3 to metrics between partial rankings, and with exploring the many data-analytic applications of such metrics. More complex types of partially ranked data are also investigated.

What is the relationship of that material with the present paper on rank tests? It is perhaps surprising that any connection exists at all, since the two topics are ostensibly quite distinct. However, as shown in the monograph, a partial ranking of n items can be identified with a *right* coset

$$S\pi = \{\alpha \epsilon\ S_n\colon\ \exists\ \sigma\ \epsilon\ S\colon\ \alpha = \sigma\pi\}$$

of a certain subgroup S of S_n. The idea here is to identify the partial ranking with the *set* of full rankings that are consistent with it, and that set turns out to be a right coset of the above form (Critchlow (1985), pages 12–14).

Distances between partial rankings can therefore be identified with distances between right cosets $S\pi$, $S\sigma$. To obtain such a distance, let d be a metric on S_n, and consider the quantity

$$d(S\pi, S\sigma) = \max\left\{\max_{\beta \epsilon S\sigma} \min_{\alpha \epsilon S\pi} d(\alpha,\beta), \qquad \max_{\alpha \epsilon S\pi} \min_{\beta \epsilon S\sigma} d(\alpha,\beta)\right\}.$$

This is called the "Hausdorff distance" between the cosets $S\pi$ and $S\sigma$, induced by d. It satisfies all of the axioms for a metric on the space of such cosets (e.g. Nadler (1978)). The author's monograph computes the Hausdorff distances between such right cosets, induced by the metrics K, F, R, H, and U.

For nonparametric rank tests, the sets of permutations in Steps 2 and 3 of the basic construction are frequently *left* cosets πS of a subgroup S of S_n. As illustrated by the two-sample location problem with ordered alternatives, many useful test statistics can be obtained by computing the minimum interpoint distance between two such left cosets.

Moreover, it can be shown that the minimum interpoint distance between two left cosets actually coincides with the Hausdorff distance between the left cosets:

Proposition: Let G be any finite group, S any subgroup of G, and d any right-invariant metric on G. Then the Hausdorff distance between any two left cosets πS and σS

$$\max \left\{ \max_{\beta \epsilon \sigma S} \min_{\alpha \epsilon \pi S} d(\alpha, \beta), \qquad \max_{\alpha \epsilon \pi S} \min_{\beta \epsilon \sigma S} d(\alpha, \beta) \right\}$$

is equal to the minimum interpoint distance

$$\min_{\substack{\alpha \epsilon \pi S \\ \beta \epsilon \sigma S}} d(\alpha, \beta).$$

Proof: This is the Lemma on pages 21–24 of Critchlow (1985), with the roles of "left" and "right" reversed. □

In general, Hausdorff distances between left cosets are not the same as Hausdorff distances between right cosets. That is, the rank test statistics of Theorem 1 are distinct from the metrics on partial rankings in the author's monograph. However, they can be derived by analogous arguments.

In sum, two different areas of statistical interest — nonparametric rank tests and partially ranked data — can be approached by quite similar analytical tools.

14. Conclusion

The construction of nonparametric rank tests, and the extension of such tests to other hypothesis testing situations, are important statistical problems. Various techniques have been suggested in the statistical literature, for generalizing rank tests to other testing situations. For example, Hájek and Sidák (1967) discuss the literature on extending two-sample location tests to the multi-sample problem, and also describe the Siegel-Tukey (1960) method of converting tests of location into tests of dispersion.

The methodology presented in this paper simultaneously extends an arbitrary rank test to many other hypothesis testing situations. Since the approach gives rise to so many useful test statistics, both old and new, its properties deserve further investigation.

* * *

APPENDIX: DERIVATIONS OF THE INDUCED TEST STATISTICS

In this appendix we will derive the test statistics induced by the metrics K, F, R, H, and U for the two-sample and multi-sample location problems, the two-way layout, the one-sample location problem, and the two-sample dispersion problem. Each testing problem will be considered for both ordered and unordered alternatives. In other words, we will prove Theorems 1 through 10.

Moreover, we will prove Theorem 12, which states that the metric $W(\pi, \sigma)$, described in Section 13.3, induces the Wald-Wolfowitz "runs test" statistic.

Before proceeding to the details of the individual proofs, it will be helpful to discuss a pair of general properties possessed by many metrics on permutations: the right-invariance property (originally noticed by Diaconis (1982)) and the "transposition property". These properties will be used extensively in the derivations. Two additional general properties, the "first and second partition properties", will be introduced in the proof of Theorems 5 and 6, and will be instrumental for Theorems 9 and 10 as well.

1. The Right-Invariance Property

Definition: The metric d on S_n is *right-invariant* if $d(\pi, \sigma) = d(\pi\tau, \sigma\tau)\ \forall\ \pi, \sigma, \tau \in S_n$.

Proposition A.1: All of the metrics on permutations presented in this paper are right-invariant. In particular, K, F, R, H, and U are right-invariant.

Proof: The proof is very easy and is omitted. □

Remarks: Proposition A.1 is not a coincidence, as it is natural to *require* that all of our metrics be right-invariant. Indeed, in Section 3, to set up the correspondence between rankings and permutations, a preliminary step was to assign *arbitrary* numerical labels to the n items: item $1, \ldots,$ item n. It is natural to insist that our distances between rankings should not depend on this arbitrary labeling of the items, and should be invariant under an arbitrary relabeling. Mathematically, this amounts to requiring that the metrics be right-invariant. See Diaconis (1982) or Critchlow (1985, pages 10–11) for further discussion.

The right-invariance property has the following important consequence. Many of the test statistics in this paper are obtained by computing the minimum interpoint distance between a certain subgroup S of S_n, and a left coset πS of that subgroup. The right-invariance property implies that this distance is the same as the minimum interpoint distance between πS and the identity permutation e:

Proposition A.2: Let d be a right-invariant metric on S_n, let S be an arbitrary subgroup of S_n, and let πS be a left coset of S (where $\pi \in S_n$). Then

$$\min_{\substack{\alpha \in \pi S \\ \beta \in S}} d(\alpha, \beta) = \min_{\alpha \in \pi S} d(\alpha, e).$$

Proof:

$$\min_{\substack{\alpha \in \pi S \\ \beta \in S}} d(\alpha, \beta) = \min_{\sigma_1, \sigma_2 \in S} d(\pi\sigma_1, \sigma_2) = \min_{\sigma_1, \sigma_2 \in S} d(\pi\sigma_1\sigma_2^{-1}, e)$$

$$= \min_{\sigma \in S} d(\pi\sigma, e) = \min_{\alpha \in \pi S} d(\alpha, e). \square$$

2. The Transposition Property

A second property of many metrics on permutations, which will be useful for deriving several rank test statistics, is the so-called "transposition property".

Definition: Let d be a metric on S_n. Let $\alpha, \beta, \gamma \in S_n$ be permutations such that α and β differ by a single transposition; that is, there exist integers $p, q \in \{1, \ldots, n\}$ such that

$$\alpha(p) = \beta(q)$$

$$\alpha(q) = \beta(p)$$

$$\alpha(i) = \beta(i) \qquad \forall\, i \neq p, q.$$

Suppose further that $\beta(p) \leq \beta(q)$ and $\gamma(p) \leq \gamma(q)$. If the preceding conditions imply that $d(\beta, \gamma) \leq d(\alpha, \gamma)$, then the metric d is said to possess the "transposition property".

Proposition A.3: The metrics K, F, and R possess the transposition property.

Proof: This is Lemma 2, pages 50–53 of Critchlow (1985). □

Remarks: Critchlow (1985) also gives counterexamples which show that the metrics H and U do *not* have the transposition property.

In the preceding reference, the transposition property is helpful for a different statistical purpose, that of deriving suitable metrics on partially ranked data. Essentially the same property has also been investigated, in yet another statistical context, by Hollander, Proschan, and Sethuraman (1977).

* * *

We proceed now to the proofs of Theorems 1 through 10, and of Theorem 12.

Proof of Theorems 1 and 3: We will prove Theorem 3 (since Theorem 1 then follows as a special case).

For the multi-sample location problem with ordered alternatives, the test statistics (in Theorem 3) are of the form $d(\pi S_{n_1} \times \ldots \times S_{n_r}, S_{n_1} \times \ldots \times S_{n_r})$, where d is one of the metrics K, F, R, H, U. Since all of these metrics are right-invariant, Proposition A.2 implies that the test statistic induced by d is equal to $\min_{\alpha \epsilon \pi S_{n_1} \times \ldots \times S_{n_r}} d(\alpha, e)$.

Let $\alpha_0 \epsilon \pi S_{n_1} \times \ldots \times S_{n_r}$ be a permutation which actually attains the above minimum; that is, α_0 satisfies $d(\alpha_0, e) = \min_{\alpha \epsilon \pi S_{n_1} \times \ldots \times S_{n_r}} d(\alpha, e)$. Obviously α_0 depends on d. The rest of the proof splits into three cases, according to the metric d under consideration: (i) $d = K$, F, or R, (ii) $d = U$, and (iii) $d = H$.

(i) If $d = K, F$, or R, then we can construct α_0 as follows. Recall that $a_{n_1+\ldots+n_{j-1}+1} < a_{n_1+\ldots+n_{j-1}+2} < \ldots < a_{n_1+\ldots+n_{n-1}+n_j}$ are the ranks assigned by π to population j, for $j = 1, \ldots, r$. Define $\alpha_0 \epsilon S_n$ by $\alpha_0(i) = a_i$, for $i = 1, \ldots, n$. Clearly $\alpha_0 \epsilon \pi S_{n_1} \times \ldots \times S_{n_r}$, since α_0 assigns the same set of ranks to population j as does π. Moreover, it is easy to see that α_0 satisfies $d(\alpha_0, e) = \min_{\alpha \epsilon \pi S_{n_1} \times \ldots \times S_{n_r}} d(\alpha, e)$, because each of the metrics K, F, and R possesses the transposition property. Indeed, given any $\alpha \epsilon \pi S_{n_1} \times \ldots \times S_{n_r}$, we can obtain α_0 from α by a sequence of transpositions, each of which satisfies the condition stated in the definition of the transposition property. (The first transposition transposes item 1 with item $\alpha_0^{-1}(a_1)$, and so on; the i-th transposition transposes item i with the item currently assigned the rank a_i.)

Thus if $d = K, F$, or R, the test statistic induced by d is $d(\alpha_0, e)$, with α_0 constructed as above. The results of Theorem 3 follow immediately, for these three metrics.

(ii) If $d = U$, we claim that the same permutation α_0 as constructed in (i) satisfies $U(\alpha_0, e) = \min_{\alpha \epsilon \pi S_{n_1} \times \ldots \times S_{n_r}} U(\alpha, e)$. Since the metric U does not possess the transposition property, this claim must be proved by a different argument than in (i).

Lemma A.1: Let $\alpha \in \pi S_{n_1} \times \ldots \times S_{n_r}$ be arbitrary, and suppose that the ranks $i_1, i_2 \in \{1, \ldots, n\}$ satisfy $i_1 < i_2$ and $\alpha^{-1}(i_1) < \alpha^{-1}(i_2)$. Then $\alpha_0^{-1}(i_1) < \alpha_0^{-1}(i_2)$.

Proof of Lemma A.1: Recall that $\alpha^{-1}(i)$ is the item assigned the rank i, and that items $n_1 + \ldots \times n_{j-1} + 1, \ldots, n_1 + \ldots + n_j$ are from population j. Now if items $\alpha^{-1}(i_1)$ and $\alpha^{-1}(i_2)$ are from the *same* population j, then so are items $\alpha_0^{-1}(i_1)$ and $\alpha_0^{-1}(i_2)$ (since all permutations in the equivalence class $\pi S_{n_1} \times \ldots \times S_{n_r}$ assign the same set of ranks to population j). It follows that $\alpha_0^{-1}(i_1) < \alpha_0^{-1}(i_2)$, because α_0 has been constructed in such a way that within each population, the items are ranked in ascending order.

On the other hand, if items $\alpha^{-1}(i_1)$ and $\alpha^{-1}(i_2)$ are from *different* populations j_1, j_2, then $\alpha^{-1}(i_1) < \alpha^{-1}(i_2)$ implies $j_1 < j_2$, whence $\alpha_0^{-1}(i_1) < \alpha_0^{-1}(i_2)$. The lemma is proved. □

From the lemma it follows by induction that for any $\alpha \in \pi S_{n_1} \times \ldots \times S_{n_r}$, if $i_1 < \ldots < i_k$ satisfy $\alpha^{-1}(i_1) < \ldots < \alpha^{-1}(i_k)$, then $\alpha_0^{-1}(i_1) < \ldots < \alpha_0^{-1}(i_k)$. Hence LLIS $\{\alpha^{-1}(1), \ldots, \alpha^{-1}(n)\} \leq LLIS\{\alpha_0^{-1}(1), \ldots, \alpha_0^{-1}(n)\}$. (Recall "LLIS" stands for "length of the longest increasing subsequence.") Therefore, $U(\alpha_0, e) = \min_{\alpha \epsilon \pi S_{n_1} \times \ldots \times S_{n_r}} U(\alpha, e)$, as claimed.

It follows that the test statistic induced by U is $U(\alpha_0, e)$, with α_0 constructed as in (i). To evaluate $U(\alpha_0, e)$, we notice that picking an increasing subsequence of $\{\alpha_0^{-1}(1), \ldots, \alpha_0^{-1}(n)\}$ amounts to choosing some integers $i_1 \leq i_2 \leq \ldots \leq i_{r-1}$, and then choosing those items from the first population among the first i_1 ranks, those items from the second population among the next $i_2 - i_1$ ranks, and so on. In this subsequence, there will be $t_{i_1 1}$ items from population 1, $t_{i_2 2} - t_{i_1 2}$ items from population 2, and so on, where t_{ij} is the number of items from the j-th population which fall among the top i ranks. Therefore

$$\begin{aligned} U(\alpha_0, e) &= n - LLIS\{\alpha_0^{-1}(1), \ldots, \alpha_0^{-1}(n)\} \\ &= n - \max_{i_1 \leq \ldots \leq i_{r-1}} \left\{ t_{i_1 1} + \left(\sum_{j=2}^{r-1} (t_{i_j j} - t_{i_{j-1} j}) \right) + (n_r - t_{i_{r-1} r}) \right\} \\ &= (n - n_r) - \max_{i_1 \leq \ldots \leq i_{r-1}} \left\{ \sum_{j=1}^{r-1} (t_{i_j j} - t_{i_j j+1}) \right\}. \end{aligned}$$

This completes the proof of Theorem 3 for Ulam's metric. As observed in Sections 4 and 6, for equal sample sizes from $r = 2$ populations, the statistic induced by U is equivalent to the Kolmogorov-Smirnov statistic. When $r > 2$, the statistic induced by U is a natural extension of the Kolmogorov-Smirnov statistic.

(iii) For the metric $d = H$, we construct the permutation α_0 as follows. For $j = 1, \ldots, r$ let $m_j = \#(\pi(N_j) \cap N_j)$, where we recall that $N_j = \{n_1 + \ldots + n_{j-1} + 1, \ldots, n_1 + \ldots + n_j\}$ is the set of labels for the items in population j. Let:

$$c_{j1}, \ldots, c_{jm_j} \text{ be an enumeration of } \pi(N_j) \cap N_j,$$

$$c_{jm_{j+1}}, \ldots, c_{jn_j} \text{ be an enumeration of } \pi(N_j)^c \cap N_j,$$

$$c'_{jm_{j+1}}, \ldots, c'_{jn_j} \text{ be an enumeration of } \pi(N_j) \cap N_j^c.$$

Define $\alpha_0 \in S_n$ by

$$\alpha_0(c_{jk}) = \begin{cases} c_{jk} & \text{for } k = 1, \ldots, m_j \\ c'_{jk} & \text{for } k = m_j + 1, \ldots, n_j. \end{cases}$$

Then $\alpha_0 \in \pi S_{n_1} \times \ldots S_{n_r}$, and $H(\alpha_0, e) = n - \sum_{j=1}^{r} m_j$. On the other hand, for any $\alpha \in \pi S_{n_1} \times \ldots \times S_{n_r}$, $H(\alpha, e) = n - \sum_{j=1}^{r} \#\{i \in N_j \colon \alpha(i) = i\} \geq n - \sum_{j=1}^{r} m_j$. Thus $H(\alpha_0, e) = \min_{\alpha \in \pi S_{n_1} \times \ldots \times S_{n_r}} H(\alpha, e) = n - \sum_{j=1}^{r} m_j$.

This completes the proof of Theorem 3. □

Proof of Theorems 2 and 4: We will prove Theorem 4, since Theorem 2 then follows as a special case.

The test statistics in Theorem 4 are the minimum interpoint distances between the sets $[\pi] = \pi S_{n_1} \times \ldots \times S_{n_r}$ and $E = S_r \odot S_{n_1} \times \ldots \times S_{n_r} = \{\beta \in S_n \colon \exists\, \sigma \in S_r \colon \beta(N_{\sigma(1)}) < \ldots < \beta(N_{\sigma(r)})\}$. These statistics are therefore equal to $\min_{\sigma \in S_r} \min_{\substack{\alpha \in \pi S_{n_1} \times \ldots \times S_{n_r} \\ \beta \in e_\sigma S_{n_1} \times \ldots \times S_{n_r}}} d(\alpha, \beta)$, where, for each $\sigma \in S_r$, e_σ is some extremal permutation for the ordered alternative $H_{1,\sigma}$: $F_{\sigma(1)}(x) \geq \ldots \geq F_{\sigma(r)}(x)$, and where $E_\sigma \equiv e_\sigma S_{n_1} \times \ldots \times S_{n_r} = \{\tau \in S_n \colon \tau(N_{\sigma(1)}) < \ldots < \tau(N_{\sigma(r)})\}$ is the set of all extremal permutations for $H_{1,\sigma}$.

We tackle now the problem of calculating $\min_{\substack{\alpha \in \pi S_{n_1} \times \ldots \times S_{n_r} \\ \beta \in e_\sigma S_{n_1} \times \ldots \times S_{n_r}}} d(\alpha, \beta)$, for fixed $\sigma \in S_r$. Intuitively, this should not be too much extra work, since we have already calculated the test statistics for the particular ordered alternative $F_1(x) \geq \ldots \geq F_r(x)$ in Theorem 3. In fact, we will demonstrate formally that the calculation of $\min_{\substack{\alpha \in \pi S_{n_1} \times \ldots \times S_{n_r} \\ \beta \in e_\sigma S_{n_1} \times \ldots \times S_{n_r}}} d(\alpha, \beta)$ follows directly from Theorem 3.

Definitions: For fixed $\sigma \in S_r$, define the partition $N_1^\sigma, \ldots, N_r^\sigma$ of $\{1, \ldots, n\}$ by $N_j^\sigma = e_\sigma(N_{\sigma(j)})$. Let $S_\sigma \equiv S_{n_{\sigma(1)}} \times \ldots \times S_{n_{\sigma(r)}}$ be the subgroup of S_n given by $\{\tau \in S_n \colon \tau(N_j^\sigma) = N_j^\sigma \ \forall\, j = 1, \ldots, r\}$. Let S denote the subgroup $S_{n_1} \times \ldots \times S_{n_r}$ considered previously.

Lemma A.2: The left coset $e_\sigma S$ equals the right coset $S_\sigma e_\sigma$.

Proof of Lemma A.2: Both cosets are clearly $\{\tau \in S_n \colon \tau(N_j) = N_j^\sigma \ \forall\, j = 1, \ldots, r\}$.
□

Lemma A.3: The minimum interpoint distance between the sets $[\pi] = \pi S$ and

$E_\sigma = e_\sigma S$ can be calculated from the equation

$$\min_{\substack{\alpha \epsilon \pi S \\ \beta \epsilon e_\sigma S}} d(\alpha, \beta) = \min_{\substack{\alpha \epsilon \pi e_\sigma^{-1} S_\sigma \\ \beta \epsilon S_\sigma}} d(\alpha, \beta).$$

Proof of Lemma A.3: By Lemma A.2 and the right invariance property,

$$\begin{aligned} \min_{\substack{\alpha \epsilon \pi S \\ \beta \epsilon e_\sigma S}} d(\alpha, \beta) &= \min_{s_1, s_2 \epsilon S} d(\pi s_1, e_\sigma s_2) = \min_{s \epsilon S} d(\pi s, e_\sigma) \\ &= \min_{s \epsilon S} d(\pi s e_\sigma^{-1}, e) = \min_{s \epsilon S_\sigma} d(\pi e_\sigma^{-1} s, e) = \min_{\substack{\alpha \epsilon \pi e_\sigma^{-1} S_\sigma \\ \beta \epsilon S_\sigma}} d(\alpha, \beta). \qquad \square \end{aligned}$$

The motivation behind Lemma A.3 is that we already know how to find the minimum interpoint distance between the sets $\pi e_\sigma^{-1} S_\sigma$ and S_σ, since Theorem 3 calculates the distance between two sets of precisely this form.

The test statistics of Theorem 4 follow easily. For example, to find the statistic induced by the metric $d = K$, we compute

$$\begin{aligned} \min_{\substack{\alpha \epsilon \pi e_\sigma^{-1} S_\sigma \\ \beta \epsilon S_\sigma}} K(\alpha, \beta) &= \sum_{j<k} \#\{(i,m) \epsilon N_j^\sigma \times N_k^\sigma \colon \pi e_\sigma^{-1}(i) > \pi e_\sigma^{-1}(m)\} \\ &= \sum_{j<k} \#\{(i,m) \epsilon e_\sigma(N_{\sigma(j)}) \times e_\sigma(N_{\sigma(k)}) \colon \pi e_\sigma^{-1}(i) > \pi e_\sigma^{-1}(m)\} \\ &= \sum_{j<k} \#\{(i,m) \epsilon N_{\sigma(j)} \times N_{\sigma(k)} \colon \pi(i) > \pi(m)\}, \end{aligned}$$

and so the statistic induced by K is

$$\min_{\sigma \epsilon S_r} \min_{\substack{\alpha \epsilon \pi e_\sigma^{-1} S_\sigma \\ \beta \epsilon S_\sigma}} K(\alpha, \beta) = \min_{\sigma \epsilon S_r} \sum_{j<k} \#\{(i,m) \epsilon N_{\sigma(j)} \times N_{\sigma(k)} \colon \pi(i) > \pi(m)\}. \qquad \square$$

Proof of Theorems 5 and 6: Two new properties possessed by some metrics on permutations, the "first and second partition properties," are fundamental to deriving the test statistics for the two-way layout.

Definition: Let d be a metric on S_n for each $n \geq 1$. (More precisely, d is a collection of metrics, defined on S_n for each $n \geq 1$). Consider an arbitrary partitioning of the set $\{1, \ldots, n\}$ into two sets Y and Z. Let $y_1 < \ldots < y_{r_1}$ and $z_1 < \ldots < z_{r_2}$ be enumerations of Y and Z, respectively, where $r_1 = \#Y, r_2 = \#Z, r_1 + r_2 = n$. For each $\pi \in S_n$ define $\pi_Y \in S_{r_1}$ and $\pi_Z \in S_{r_2}$ by

$$\pi_Y(j) = \#\{k \in Y\colon \pi(k) \leq \pi(y_j)\} \qquad \text{for } j = 1, \ldots, r_1,$$

$$\pi_Z(j) = \#\{k \in Z\colon \pi(k) \leq \pi(z_j)\} \qquad \text{for } j = 1, \ldots, r_2.$$

If the preceding conditions imply that

$$d(\pi, \sigma) \geq d(\pi_Y, \sigma_Y) + d(\pi_Z, \sigma_Z) \quad \text{for all } \pi, \sigma \in S_n,$$

then the metric d is said to possess the *first partition property.* Further, if the additional conditions

$$y_i < z_j \quad \text{for all } i, j,$$

$$\pi(y_i) < \pi(z_j) \quad \text{for all } i, j,$$

$$\sigma(y_i) < \sigma(z_j) \quad \text{for all } i, j$$

imply that equality is attained above, i.e. $d(\pi, \sigma) = d(\pi_Y, \sigma_Y) + d(\pi_Z, \sigma_Z)$, then the metric d is said to possess the *second partition property.*

The importance of the partition properties is due to the following observation (the statement and proof of which use the notation of Sections 8 and 9).

Proposition A.4: If the metric d possesses both the first and second partition properties, then the induced test statistic for the two-way layout with ordered alternatives is the sum

$$d([\pi], [e]) = \sum_{i=1}^{b} d(\pi_i, e),$$

and the induced test statistic for the two-way layout with unordered alternatives is

$$d([\pi], S_r \odot [e]) = \min_{\sigma \epsilon S_r} \sum_{i=1}^{b} d(\pi_i, \sigma).$$

Proof: The test statistic for the ordered alternatives case is $d([\pi],[e]) = \min_{\substack{\alpha \epsilon [\pi] \\ \beta \epsilon [e]}} d(\alpha, \beta)$. The first partition property implies that $d(\alpha, \beta) \geq \sum_{i=1}^{b} d(\alpha_i, \beta_i) = \sum_{i=1}^{b} d(\pi_i, e)$ for all $\alpha \epsilon [\pi], \beta \epsilon [e]$. On the other hand, if we consider the particular permutation $\alpha_0 \epsilon [\pi]$ defined by

$$\alpha_0((i-1)r + j) = (i-1)r + \pi_i(j) \ \forall \ i = 1, \ldots, b \ \forall \ j = 1, \ldots, r,$$

and if we take $\beta_0 = e \ \epsilon \ [e]$, then the second partition property implies that the preceding bound is actually attained:

$$d(\alpha_0, \beta_0) = \sum_{i=1}^{b} d(\alpha_{0i}, \beta_{0i}) = \sum_{i=1}^{b} d(\pi_i, e).$$

Hence $\min_{\substack{\alpha \epsilon [\pi] \\ \beta \epsilon [e]}} d(\alpha, \beta) = \sum_{i=1}^{b} d(\pi_i, e)$, as claimed.

For the case of unordered alternatives, the set of extremal permutations is $E = \bigcup_{\sigma \epsilon S_r} E_\sigma$, where $E_\sigma = \{\tau \ \epsilon \ S_n \colon \tau_i = \sigma \ \forall \ i = 1, \ldots, b\}$. For each $\sigma \ \epsilon \ S_r$ let $e_\sigma \ \epsilon \ E_\sigma$ be the permutation defined by $e_\sigma((i-1)r + j) = (i-1)r + \sigma(j)$ for all $i = 1, \ldots, b$ and $j = 1, \ldots, r$. The test statistic for the two-way layout with unordered alternatives is

$$\begin{aligned} d([\pi], E) &= \min_{\sigma \epsilon S_r} d([\pi], E_\sigma) \\ &= \min_{\sigma \epsilon S_r} d([\pi], [e_\sigma]) \\ &= \min_{\sigma \epsilon S_r} d([\pi] e_{\sigma^{-1}}), [e_\sigma] e_{\sigma^{-1}}) \\ &= \min_{\sigma \epsilon S_r} d([\pi e_{\sigma^{-1}}]), [e]). \end{aligned}$$

By the result proved in the first part of this proposition (the ordered alternatives case), this last minimum equals

$$\min_{\sigma\epsilon S_r}\sum_{i=1}^{b} d((\pi e_{\sigma^{-1}})_i, e) = \min_{\sigma\epsilon S_r}\sum_{i=1}^{b} d(\pi_i\sigma^{-1}, e) = \min_{\sigma\epsilon S_r}\sum_{i=1}^{b} d(\pi_i, \sigma). \qquad \square$$

Theorems 5 and 6 will now follow directly from the following fact.

Proposition A.5: The metrics K, F, R^2, and U all possess both the first and second partition properties.

Proof: The proof that these metrics possess the second partition property is very straightforward, and is omitted.

To show that K satisfies the first partition property, we need only observe (using the notation in the definition of the partition property) that

$$\begin{aligned} K(\pi,\sigma) &= \#\{(i,j)\epsilon\{1,\ldots,n\}^2\colon\ \pi(i) < \pi(j) \text{ and } \sigma(i) > \sigma(j)\} \\ &\geq \#\{(i,j)\epsilon Y \times Y\colon\ \pi(i) < \pi(j) \text{ and } \sigma(i) > \sigma(j)\} \\ &\qquad + \#\{(i,j)\epsilon Z \times Z\colon\ \pi(i) < \pi(j) \text{ and } \sigma(i) > \sigma(j)\} \\ &= \#\{(i,j)\epsilon\{1,\ldots,r_1\}^2\colon\ \pi_Y(i) < \pi_Y(j) \text{ and } \sigma_Y(i) > \sigma_Y(j)\} \\ &\qquad + \#\{(i,j)\epsilon\{1,\ldots,r_2\}^2\colon\ \pi_Z(i) < \pi_Z(j) \text{ and } \sigma_Z(i) > \sigma_Z(j)\} \\ &= K(\pi_Y,\sigma_Y) + K(\pi_Z,\sigma_Z). \end{aligned}$$

To prove the first partition property for U, let $i_1,\ldots,i_m$ be a maximal collection of items ranked in the same relative order by $\pi,\sigma \;\epsilon\; S_n$. Let $y_{j_1},\ldots,y_{j_p}$ and $z_{k_1},\ldots,z_{k_q}$ be enumerations of the sets $\{i_1,\ldots,i_m\}\cap Y$ and $\{i_1,\ldots,i_m\}\cap Z$, respectively, where $p+q=m$. Then $y_{j_1},\ldots,y_{j_p}$ are ranked in the same relative order by π and σ, and hence $j_1,\ldots,j_p$ are ranked in the same relative order by $\pi_Y,\sigma_Y \;\epsilon\; S_{r_1}$. Similarly $k_1,\ldots,k_q$ are ranked in the same relative order by $\pi_Z,\sigma_Z \;\epsilon\; S_{r_2}$. It follows that

$$U(\pi,\sigma) = n - m = (r_1 - p) + (r_2 - q) \geq U(\pi_Y,\sigma_Y) + U(\pi_Z,\sigma_Z).$$

We sketch the main ideas of the proof of the first partition property for F, which is relatively difficult. The proof is by induction on n. Assume the property is true for $n-1$, and let $\pi, \sigma \; \epsilon S_n$ be arbitrary. Let $m \epsilon \{1, \ldots, n\}$ minimize $|\pi(\cdot) - \sigma(\cdot)|$. Without loss of generality we assume $m \; \epsilon \; Y$, and so $m = y_{i_m}$ for some index i_m. Let $\pi', \sigma' \; \epsilon S_{n-1}$ be the permutations obtained by deleting m, i.e.

$$\pi'(i) = \begin{cases} \pi(i) & \text{if } i < m \text{ and } \pi(i) < \pi(m) \\ \pi(i) - 1 & \text{if } i < m \text{ and } \pi(i) > \pi(m) \\ \pi(i+1) & \text{if } i \geq m \text{ and } \pi(i+1) < \pi(m) \\ \pi(i+1) - 1 & \text{if } i \geq m \text{ and } \pi(i+1) > \pi(m) \end{cases}$$

and similarly for σ'. Finally, let $\Omega_1 = \{i = 1, \ldots, n\colon \; \pi(i) < \pi(m) \text{ and } \sigma(i) > \sigma(m)\}$, let $\Omega_2 = \{i = 1, \ldots, n\colon \; \pi(i) > \pi(m) \text{ and } \sigma(i) < \sigma(m)\}$, and let $\Omega = \Omega_1 \cup \Omega_2$.

It is straightforward to verify the following four equations:

$$F(\pi_Z, \sigma_Z) = F(\pi'_Z, \sigma'_Z),$$

$$F(\pi, \sigma) = F(\pi', \sigma') + \#\Omega + |\pi(m) - \sigma(m)|,$$

$$F(\pi_Y, \sigma_Y) \leq F(\pi'_Y, \sigma'_Y) + \#(\Omega \cap Y) + |\pi_Y(i_m) - \sigma_Y(i_m)|,$$

$$\begin{aligned} |\pi_Y(i_m) - \sigma_Y(i_m)| &= |\pi(m) - \sigma(m) - (\#(\Omega_1 \cap Z) - \#(\Omega_2 \cap Z))| \\ &\leq |\pi(m) - \sigma(m)| + \#(\Omega \cap Z). \end{aligned}$$

From these equations and the inductive assumption, we obtain

$$\begin{aligned} F(\pi, \sigma) &= F(\pi', \sigma') + \#\Omega + |\pi(m) - \sigma(m)| \\ &\geq F(\pi'_Y, \sigma'_Y) + F(\pi'_Z, \sigma'_Z) + \#(\Omega \cap Y) + \#(\Omega \cap Z) \\ &\qquad + |\pi_Y(i_m) - \sigma_Y(i_m)| - \#(\Omega \cap Z) \\ &\geq F(\pi_Y, \sigma_Y) + F(\pi_Z, \sigma_Z), \end{aligned}$$

as was to be shown.

The proof that R^2 possesses the first partition property follows essentially the same lines as the preceding proof for F, and is omitted. □

Proof of Theorems 7 and 8: These two theorems present the test statistics for the one-sample location problem, with ordered and with unordered alternatives, respectively. We will use the notation of Section 10.

For the case of ordered alternatives (Theorem 7), the test statistic is $d(\pi S_{n_1} \times S_1 \times S_{n_1}, S_{n_1} \times S_1 \times S_{n_1})$. Therefore, this statistic can be inferred directly as a special case of Theorem 3, where the statistic $d(\pi S_{n_1} \times \ldots \times S_{n_r}, S_{n_1} \times \ldots \times S_{n_r})$ was derived. Thus, although Theorem 3 deals with a different testing situation (the multi-sample location problem with ordered alternatives), the two statistics are *algebraically* equivalent.

As an example, Theorem 3 implies that the one-sample location statistic induced by the metric F is $F(\pi S_{n_1} \times S_1 \times S_{n_1}, S_{n_1} \times S_1 \times S_{n_1}) = \sum_{i=1}^{n} |a_i - i|$. It is now simply a matter of adapting to the notation of Section 10, to obtain the form of this statistic given in Theorem 7. Indeed, for the one-sample location problem, one has $a_{n+1-i} = n + 1 - a_i$, and so $\sum_{i=1}^{n} |a_i - i| = 2\sum_{i=1}^{n_1}(a_i - i) = (2\sum_{i=1}^{n_1} \pi(i)) - n_1(n_1 + 1)$. Moreover, from the definition of π^*,

$$\pi(i) = \begin{cases} \pi^*(i) + \frac{n+1}{2} & \text{if } X_i > 0 \\ \frac{n+1}{2} - \pi^*(i) & \text{if } X_i < 0, \end{cases}$$

and so the preceding statistic becomes

$$\begin{aligned} &2 \sum_{\{i \epsilon N_1:\ X_i > 0\}} (\pi^*(i) + \frac{n+1}{2}) \ + \ 2 \sum_{\{i \epsilon N_1:\ X_i < 0)\}} (\frac{n+1}{2} - \pi^*(i)) - n_1(n_1 + 1) \\ &= 2n_1 \frac{n+1}{2} - n_1(n_1 + 1) \ + \ 2 \sum_{\{i \epsilon N_1:\ X_i > 0\}} \pi^*(i) - \ 2 \sum_{\{i \epsilon N_1:\ X_i < 0\}} \pi^*(i) \\ &= n_1(n_1 + 1) \ + \ 2 \sum_{\{i \epsilon N_1:\ X_i > 0\}} \pi^*(i) \ - \ 2 \left[\frac{n_1(n_1 + 1)}{2} - \sum_{\{i \epsilon N_1:\ X_i < 0)\}} \pi^*(i) \right] \\ &= 4 \sum_{\{i \epsilon N_1:\ X_i > 0\}} \pi^*(i), \end{aligned}$$

which is equivalent to the Wilcoxon signed rank statistic, as claimed in Theorem 7.

For the case of unordered alternatives (Theorem 8), the test statistic is $d(\pi S_{n_1} \times S_1 \times S_{n_1}, S_2 \odot S_{n_1} \times S_1 \times S_{n_1})$. This statistic is most easily derived by noting that the set of extremal permutations

$$S_2 \odot S_{n_1} \times S_1 \times S_{n_1} = (S_{n_1} \times S_1 \times S_{n_1}) \cup (\tilde{e}S_{n_1} \times S_1 \times S_{n_1}) = S \cup \tilde{e}S,$$

where $S = S_{n_1} \times S_1 \times S_{n_1}$, and where $\tilde{e} \in S_n$ is defined by $\tilde{e}(i) = n+1-i$, for $i = 1, \ldots, n$. Therefore the test statistic $d(\pi S, S_2 \odot S)$ equals $\min\{d(\pi S, S), d(\pi S, \tilde{e}S)\}$.

The distance $d(\pi S, S)$ was computed in Theorem 7. To evaluate $d(\pi S, \tilde{e}S)$, we first note that $\tilde{e}^{-1} = \tilde{e}$ and hence $S\tilde{e}^{-1} = S\tilde{e} = \tilde{e}S$ (because all of these cosets equal $\{\tau \in S_n\colon \tau(N_1) = N_3, \tau(N_2) = N_2, \tau(N_3) = N_1\}$). Therefore (also using the right-invariance property) $d(\pi S, \tilde{e}S) = d(\pi S, S\tilde{e}) = d(\pi S\tilde{e}^{-1}, S) = d(\pi\tilde{e}S, S)$. This last distance $d(\pi\tilde{e}S, S)$ can be found directly from Theorem 3 (or Theorem 7).

As an illustration, the statistic induced by F (for unordered alternatives) will now be calculated explicitly. This statistic is $\min\{d(\pi S, S), d(\pi\tilde{e}S, S)\}$, which, by Theorem 3, equals $\min\{\sum_{i=1}^{n} |a_i - i|, \sum_{i=1}^{n} |\tilde{a}_i - i|\}$, where $a_1 < \ldots < a_{n_1}$ is an enumeration of the set $\pi(N_1)$, $a_{n_1+2} < \ldots < a_n$ is an enumeration of $\pi(N_3)$, $\tilde{a}_1 < \ldots < \tilde{a}_{n_1}$ is an enumeration of $\pi\tilde{e}(N_1)$, $\tilde{a}_{n_1+2} < \ldots < \tilde{a}_n$ is an enumeration of $\pi\tilde{e}(N_3)$, and where $a_{n_1+1} = \pi(n_1+1)$, $\tilde{a}_{n_1+1} = \pi\tilde{e}(n_1+1)$. For the one-sample location problem, it is easy to verify that

$$\tilde{a}_i = \begin{cases} a_{n_1+1+i} & \text{for } i = 1, \ldots, n_1 \\ a_i = n_1 + 1 & \text{for } i = n_1 + 1 \\ a_{i-n_1-1} & \text{for } i = n_1 + 2, \ldots, n. \end{cases}$$

The test statistic is thus

$$\begin{aligned} \min\{\sum_{i=1}^{n}|a_i - i|, \sum_{i=1}^{n}|\tilde{a}_i - i|\} &= \min\{2\sum_{i=1}^{n_1}(a_i - i), 2\sum_{i=1}^{n_1}(a_{n_1+1+i} - i)\} \\ &= \min\{2\sum_{i=1}^{n_1}(a_i - i),\ 2n_1(n_1+1) - 2\sum_{i=1}^{n_1}(a_i - i)\} \\ &= \min\{4 \sum_{\{i \in N_1\colon X_i > 0} \pi^*(i),\ 2n_1(n_1+1) - 4 \sum_{\{i \in N_1\colon X_i > 0\}} \pi^*(i)\} \\ &= 4\left\{\frac{n_1(n_1+1)}{4} - \left|(\sum_{\{i \in N_1\colon X_i > 0\}} \pi^*(i)) - \frac{n_1(n_1+1)}{4}\right|\right\}. \end{aligned}$$

(The next to last equality uses Theorem 7.) This is equivalent to the two-sided Wilcoxon signed rank statistic, as claimed in Theorem 8. □

Proof of Theorems 9 and 10: These theorems present the test statistics for the two-sample dispersion problem, with ordered and unordered alternatives, respectively. The derivations of these statistics will make use of *all* of the general metric properties, which we have thus far considered: the right-invariance property, the transposition property, and the first and second partition properties.

We will use the notation of Section 11, and let S denote the subgroup $S_{n_1} \times S_{n_2}$ described there. For simplicity, we assume (as in Section 11) that the sample sizes n_1 and n_2 are both even numbers.

For the ordered alternatives problem, the extremal set is $E = T\psi S = \psi S$, and so the test statistic is

$$d(T\pi S, \psi S) = d(T\pi S, \psi) = \min_{\alpha \epsilon T\pi S} d(\alpha, \psi)$$

(by right-invariance). To derive this statistic we must therefore construct a "minimizing" permutation $\alpha_0 \epsilon T\pi S$ such that $d(\alpha_0, \psi) = \min_{\alpha \epsilon T\pi S} d(\alpha, \psi)$, and then we must compute $d(\alpha_0, \psi)$.

It is straightforward to derive the statistic for the metric $d = H$. Indeed, let

$$m_1 = \#\{i \,\epsilon\, N_1\colon\ \pi(i) \leq \frac{n_1}{2} \text{ or } \pi(i) \geq n + 1 - \frac{n_1}{2}\} = \#(\pi(N_1) \cap \psi(N_1));$$

we must show that

$$\min_{\alpha \epsilon T\pi S} H(\alpha, \psi) = 2n_1 - 2m_1.$$

Let $m_2 = \#(\pi(N_2) \cap \psi(N_2)) = m_1 + n_2 - n_1$. For $j = 1, 2$, let $c_{j1}, \ldots, c_{jm_j}$ be an enumeration of the set $\pi(N_j) \cap \psi(N_j)$. Choose any $\sigma_0 \,\epsilon\, S$ such that for all j and k, $\sigma_0(\psi^{-1}(c_{jk})) = \pi^{-1}(c_{jk})$. Let $\alpha_0 = \pi\sigma_0 \,\epsilon\, T\pi S$. Then $\alpha_0(\psi^{-1}(c_{jk})) = c_{jk} = \psi(\psi^{-1}(c_{jk}))$, and hence $H(\alpha_0, \psi) = n - \#\{i = 1, \ldots, n\colon\ \alpha_0(i) = \psi(i)\} = n - (m_1 + m_2) = 2n_1 - 2m_1$. Thus $\min_{\alpha \epsilon T\pi S} H(\alpha, \psi) \leq 2n_1 - 2m_1$. To prove the reverse inequality, we simply observe that for any $\alpha = \tau\pi\sigma \ \epsilon \ T\pi S$,

$$\begin{aligned} H(\tau\pi\sigma, \psi) &= n - \sum_{j=1}^{2} \#\{i \ \epsilon \ N_j\colon\ \tau\pi\sigma(i) = \psi(i)\} \\ &\geq n - \sum_{j=1}^{2} \#(\tau\pi\sigma(N_j) \cap \psi(N_j)) \\ &= n - \sum_{j=1}^{2} \#(\pi(N_j) \cap \psi(N_j)) \\ &= n - (m_1 + m_2) \\ &= 2n_1 - 2m_1. \end{aligned}$$

To derive the statistics induced by K, F, and R^2, we need the following result, the proof of which uses the fact that these three metrics satisfy *both* the transposition property and the first and second partition properties.

Lemma A.4: For $d = K, F$, or R^2, the permutation $\alpha_0 \ \epsilon \ T\pi S$ which minimizes $d(\cdot, \psi)$ may be chosen so as to satisfy the following constraints:

$$\#(\alpha_0(N_1) \cap \{1, \ldots, \frac{n}{2}\}) = \frac{n_1}{2},$$

$$\alpha_0(1) < \ldots < \alpha_0(n_1),$$

$$\alpha_0(n_1 + 1) < \ldots < \alpha_0(n_1 + n_2).$$

Proof of Lemma A.4: Let $\alpha \ \epsilon \ T\pi S$ be arbitrary. Let $\alpha' = \alpha\sigma \ \epsilon \ T\pi S$, where $\sigma \ \epsilon \ S$ is chosen so that

$$\alpha'(1) < \ldots < \alpha'(n_1), \qquad \alpha'(N_1) = \alpha(N_1),$$

$$\alpha'(n_1 + 1) < \ldots < \alpha'(n), \qquad \alpha'(N_2) = \alpha(N_2).$$

Then $d(\alpha, \psi) \geq d(\alpha', \psi)$, by the transposition property (since α' can be obtained from α by a sequence of transpositions which satisfy the defining conditions of the transposition property).

Next we partition $\{1, \ldots, n\}$ into the two sets $Y = \{1, \ldots, n/2\}$ and $Z = \{n/2 + 1, \ldots, n\}$. As defined in the proof of Theorems 5 and 6, let α'_Y and α'_Z be the permutations induced by α' for these two sets, and similarly define ψ_Y, ψ_Z. Then by the first partition property, $d(\alpha', \psi) \geq d(\alpha'_Y, \psi_Y) + d(\alpha'_Z, \psi_Z)$.

Next let $\alpha'' = \tau\alpha' \in T\pi S$, where $\tau \in T$ is the composition of $|n_1/2 - \#(\alpha'(N_1) \cap \{1, \ldots, n/2\})|$ transpositions, chosen so that $\#(\alpha''(N_1) \cap \{1, \ldots, n/2\}) = n_1/2$. Let $\alpha_0 = \alpha''\sigma_0 \in T\pi S$, where $\sigma_0 \in S$ is chosen so that

$$\alpha_0(1) < \ldots < \alpha_0(n_1), \qquad \alpha_0(N_1) = \alpha''(N_1),$$

$$\alpha_0(n_1 + 1) < \ldots < \alpha_0(n), \qquad \alpha_0(n_2) = \alpha''(N_2).$$

Then α_0 also satisfies the constraint $\#(\alpha_0(N_1) \cap \{1, \ldots, n/2\}) = n_1/2$. Moreover, a repeated application of the transposition property shows that $d(\alpha'_Y, \psi_Y) \geq d(\alpha_{0Y}, \psi_Y)$ and $d(\alpha'_Z, \psi_Z) \geq d(\alpha_{0Z}, \psi_Z)$, and hence

$$d(\alpha'_Y, \psi_Y) + d(\alpha'_Z, \psi_Z) \geq d(\alpha_{0Y}, \psi_Y) + d(\alpha_{0Z}, \psi_Z).$$

Finally, by the second partition property,

$$d(\alpha_{0Y}, \psi_Y) + d(\alpha_{0Z}, \psi_Z) = d(\alpha_0, \psi).$$

Putting together all of the pieces, we obtain the desired result:

$$\begin{aligned} d(\alpha, \psi) &\geq d(\alpha', \psi) \text{ (by the transposition property)} \\ &\geq d(\alpha'_Y, \psi_Y) + d(\alpha'_Z, \psi_Z) \text{ (by the first partition property)} \\ &\geq d(\alpha_{0Y}, \psi_Y) + d(\alpha_{0Z}, \psi_Z) \text{ (by the transposition property)} \\ &= d(\alpha_0, \psi) \text{ (by the second partition property).} \quad \square \end{aligned}$$

The preceding lemma is the key step in deriving the statistics induced by K, F, and R^2. Indeed, for $d = K$ or F, it is relatively easy to verify that *all* permutations $\alpha_0 \in T\pi S$ which satisfy the conditions of the lemma will produce the *same* value of $d(\alpha_0, \psi)$, namely, the value given in the statement of Theorem 9. The details are omitted.

For $d = R^2$, the minimizing permutation α_0 may be constructed as follows. Recall (from Section 11) that $b_1, \ldots, b_{n_1}$ is an enumeration of $\pi(N_1)$ such that

$$|b_1 - \frac{n+1}{2}| \geq \ldots \geq |b_{n_1} - \frac{n+1}{2}|.$$

The minimizing α_0 satisfies the conditions of the lemma, plus the additional constraints:

$$\begin{array}{ll} \alpha_0^{-1}(b_i) \in N_1 & \text{whenever either } i \text{ is odd and } b_i \leq \dfrac{n+1}{2}, \\ & \text{or } i \text{ is even and } b_i > \dfrac{n+1}{2}; \\ \alpha_0^{-1}(n+1-b_i) \in N_1 & \text{whenever either } i \text{ is odd and } b_i > \dfrac{n+1}{2}, \\ & \text{or } i \text{ is even and } b_i \leq \dfrac{n+1}{2}. \end{array}$$

(Intuitively, this says that the item from the first population whose rank is furthest from the median is placed to the left of the median, the next furthest is placed to the right of the median, and so on. Of course the items within each population must then be relabeled to ensure that $\alpha_0(1) < \ldots < \alpha_0(n_1)$ and $\alpha_0(n_1+1) < \ldots < \alpha_0(n)$.)

The proof is by contradiction. If α_0 does not satisfy these conditions, one can find a transposition $\tau \in T$ and a "relabeling permutation" $\sigma \in S$ such that $R^2(\tau\alpha_0\sigma, \psi) < R^2(\alpha_0, \psi)$. The details are omitted.

For the case of unordered alternatives (Theorem 10), the test statistic is

$$d(T\pi S, TS_2 \odot \psi S) = \min\{d(T\pi S, T\psi S), d(T\pi S, T\tilde{\psi} S)\},$$

where $\tilde{\psi}$ is the "other type of extremal permutation":

$$\tilde{\psi}(i) = \begin{cases} i + \frac{n_2}{2} & \text{for } i = 1, \ldots, n_1 \\ i - n_1 & \text{for } i = n_1+1, \ldots, n_1 + \frac{n_2}{2} \\ i & \text{for } i = n_1 + \frac{n_2}{2} + 1, \ldots, n_1 + n_2. \end{cases}$$

The distance $d(T\pi S, T\psi S)$ was found in Theorem 9. On the other hand, it can also be shown that computing $d(T\pi S, T\tilde{\psi}S)$ again reduces to Theorem 9. (Intuitively, one need only relabel the two populations.) The details of the derivation are similar to the details for the unordered alternatives cases in Theorems 2, 4, 6, and 8, and are omitted. □

Proof of Theorem 12: This theorem asserts that

$$W(\pi,\sigma) \equiv \#\{i = 1,\ldots,n-1\colon\ \pi\sigma^{-1}(i+1) \neq \pi\sigma^{-1}(i)+1\}$$

is a metric on S_n, and that it induces the Wald-Wolfowitz statistic for the two-sample location problem with unordered alternatives. To prove it, we use the notation of Sections 4, 5, and 13.3.

We first show that W is a metric on S_n. Since W is obviously right-invariant, to check that $W(\pi,\sigma) \equiv W(\sigma,\pi)$, it is enough to show that $W(\pi,e) \equiv W(e,\pi)$. For any $\pi \in S_n$ define $\bar{\pi} \in S_{n+1}$ by

$$\bar{\pi}(i) = \begin{cases} \pi(i) & \text{for } i = 1,\ldots,n \\ n+1 & \text{for } i = n+1. \end{cases}$$

Then

$$\begin{aligned} W(\pi,e) &= \#\{i = 1,\ldots,n-1\colon\ \pi(i+1) \neq \pi(i)+1\} \\ &= \#\{i = 1,\ldots,n\colon\ \bar{\pi}(i+1) \neq \bar{\pi}(i)+1\} - 1_{\{\pi(n)\neq n\}} \\ &= \#\{i = 1,\ldots,n\colon\ i+1 \neq \bar{\pi}^{-1}(\bar{\pi}(i)+1)\} - 1_{\{\pi^{-1}(n)\neq n\}} \\ &= \#\{i = 1,\ldots,n\colon\ \bar{\pi}^{-1}(i)+1 \neq \bar{\pi}^{-1}(i+1)\} - 1_{\{\pi^{-1}(n)\neq n\}} \\ &= \#\{i = 1,\ldots,n-1\colon\ \pi^{-1}(i)+1 \neq \pi^{-1}(i+1)\} \\ &= W(e,\pi), \end{aligned}$$

as claimed. (The fourth equality uses the fact that as i ranges from 1 to n, so does $\bar{\pi}^{-1}(i)$.)

The other metric property which is non-trivial to verify is the triangle inequality. By right-invariance, this reduces to showing $W(\pi,\sigma) \leq W(\pi,e) + W(e,\sigma)$. We observe that for $i = 1, \ldots, n-1$, if both $\sigma^{-1}(i+1) = \sigma^{-1}(i) + 1$ and $\pi(\sigma^{-1}(i)+1) = \pi(\sigma^{-1}(i)) + 1$, then $\pi\sigma^{-1}(i+1) = \pi\sigma^{-1}(i) + 1$. Hence

$$
\begin{aligned}
W(\pi, \sigma) &= \#\{i = 1, \ldots, n-1 \colon\ \pi\sigma^{-1}(i+1) \neq \pi\sigma^{-1}(i) + 1\} \\
&\leq \#\{i = 1, \ldots, n-1 \colon\ \sigma^{-1}(i+1) \neq \sigma^{-1}(i) + 1\} \\
&\qquad + \#\{i = 1, \ldots, n-1 \colon\ \pi(\sigma^{-1}(i)+1) \neq \pi(\sigma^{-1}(i)) + 1 \text{ and } \sigma^{-1}(i+1) = \sigma^{-1}(i) + 1\} \\
&\leq W(e, \sigma) + \#\{i = 1, \ldots, n-1 \colon\ \pi(\sigma^{-1}(i)+1) \neq \pi(\sigma^{-1}(i)) + 1 \text{ and } \sigma^{-1}(i) \leq n-1\} \\
&\leq W(e, \sigma) + \#\{i = 1, \ldots, n-1 \colon\ \pi(i+1) \neq \pi(i) + 1\} \\
&= W(e, \sigma) + W(\pi, e),
\end{aligned}
$$

as desired.

Finally, we show that W induces the Wald-Wolfowitz "runs test" statistic, for the two-sample location problem with unordered alternatives. The statistic induced by W is

$$W(\pi S_{n_1} \times S_{n_2}, S_2 \odot S_{n_1} \times S_{n_2}) = \min\{\min_{\alpha \epsilon \pi S_{n_1} \times S_{n_2}} W(\alpha, e), \min_{\alpha \epsilon \pi S_{n_1} \times S_{n_2}} W(\alpha, \tilde{e})\},$$

where

$$\tilde{e}(i) = \begin{cases} i + n_2 & i = 1, \ldots, n_1 \\ i - n_1 & i = n_1 + 1, \ldots, n_1 + n_2. \end{cases}$$

To find $\min_{\alpha \epsilon \pi S_{n_1} \times S_{n_2}} W(\alpha, e)$, let X's and O's denote items from populations 1 and 2, respectively. In the ranking π, let us say that "an X is followed by an O" whenever an item from population 1 is followed by an item from population 2; and let us say that "an X followed by a space" occurs if the item assigned the rank n belongs to population 1.

We notice first that

$$
\begin{aligned}
\min_{\alpha \epsilon \pi S_{n_1} \times S_{n_2}} W(\alpha, e) \geq\ & \#\ (O\text{'s followed by an } X \text{ or a space}) - 1 \\
&+\ \max\{0, \#(X\text{'s followed by an } O) - 1\} \\
&+ \#(X\text{'s followed by a space}).
\end{aligned}
$$

Indeed, any O followed by an X or a space must contribute 1 to $W(\alpha, e)$, unless it happens to be item $n_1 + n_2$. Similarly, any X followed by an O must contribute 1 unless it is item n_1; and any X followed by a space contributes 1.

On the other hand, we can construct $\alpha_0 \in \pi S_{n_1} \times S_{n_2}$ which actually *attains* the preceding lower bound, as follows. Let $a_1 < \ldots < a_{n_1}$ and $a_{n_1+1} < \ldots < a_{n_1+n_2}$ be the ranks assigned by π to populations 1 and 2, respectively. Choose any pair $(j,k) \in N_1 \times N_2$ such that $a_k = a_j + 1$ (i.e. choose any X followed by an O). Let

$$\alpha_0(i) = \begin{cases} a_{i+j} & i = 1, \ldots, n_1 - j \\ a_{i+j-n_1} & i = n_1 - j + 1, \ldots, n_1 \\ a_{i+k-n_1-1} & i = n_1 + 1, \ldots, n_1 + n - k + 1 \\ a_{i+k-n-1} & i = n_1 + n - k + 2, \ldots, n_1 + n_2. \end{cases}$$

(If no such pair (j,k) exists, just let $\alpha_0(i) \equiv a_i$.) It is easy to check that $\alpha_0 \in \pi S_{n_1} \times S_{n_2}$, and that

$$\begin{aligned} W(\alpha_0, e) = \#\ &(O\text{'s followed by an } X \text{ or a space}) - 1 \\ &+ \max\{0, \#(X\text{'s followed by an } O) - 1\} \\ &+ \#(X\text{'s followed by a space}). \end{aligned}$$

Hence $\min_{\alpha \in \pi S_{n_1} \times S_{n_2}} W(\alpha, e)$ is the right hand side above. By a similar argument,

$$\begin{aligned} \min_{\alpha \in \pi S_{n_1} \times S_{n_2}} W(\alpha, \tilde{e}) = \#&(X\text{'s followed by an } O \text{ or a space}) - 1 \\ &+ \max\{0, \#(O\text{'s followed by an } X) - 1\} \\ &+ \#(O\text{'s followed by a space}). \end{aligned}$$

It follows that

$$\begin{aligned} W(\pi S_{n_1} \times S_{n_2}, S_2 \odot S_{n_1} \times S_{n_2}) &= \min\{\min_{\alpha \in \pi S_{n_1} \times S_{n_2}} W(\alpha, e), \min_{\alpha \in \pi S_{n_1} \times S_{n_2}} W(\alpha, \tilde{e})\} \\ &= \#(X\text{'s followed by an } O) + \#(O\text{'s followed by an } X) - 1 \\ &= \#(\text{runs in } \pi) - 2, \end{aligned}$$

which is equivalent to the Wald-Wolfowitz statistic, as claimed. □

Acknowledgements. The author thanks Persi Diaconis, Prem Puri, and Andrew Rukhin for helpful discussions. To these and to all others who provided encouragement, the author expresses his sincere appreciation. Finally, the author thanks Mohan Delampady for pointing out that the test statistic induced by H in Section 4 is, in fact, equivalent to Mood's median test statistic.

REFERENCES

ANSARI, A. R. and BRADLEY, R. A. (1960). Rank-sum tests for dispersions. *Ann. Math. Statist.* **31** 1174–1189.

BERLEKAMP, E. R. (1968). *Algebraic Coding Theory.* McGraw-Hill, New York.

CAMERON, P. J. and VAN LINT, J. H. (1980). *Graphs, Codes, and Designs.* London Mathematical Society Lecture Notes Series, Volume 43. Cambridge University Press.

CRITCHLOW, D. E. (1985). *Metric Methods for Analyzing Partially Ranked Data.* Lecture Notes in Statistics, Volume 34. Springer-Verlag, New York.

CURTIS, C. W. and REINER, I. (1966). *Representation Theory of Finite Groups and Associative Algebras.* Pure and Applied Mathematics, Volume 11. Interscience, New York.

DIACONIS, P. (1982). *Group Theory in Statistics.* Harvard University Lecture Notes. To appear in the Institute of Mathematical Statistics Lecture Notes-Monograph Series.

DIACONIS, P. and GRAHAM, R. L. (1977). Spearman's footrule as a measure of disarray. *J. Roy. Statist. Soc. B.* **39** 262–268.

FEIGIN, P. D. and ALVO, M. (1986). Intergroup diversity and concordance for ranking data: An approach via metrics for permutations. To appear in *Ann. Statist.* **14**.

FEIGIN, P. D. and COHEN, A. (1978). On a model for concordance between judges. *J. Roy. Statist. Soc. B.* **40** 203–213.

FISHER, R. A. (1925). *Statistical Methods for Research Workers.* First Edition. Oliver & Boyd, Edinburgh.

FLIGNER, M. A. and VERDUCCI, J. S. (1985). Distance based ranking models. Technical Report, Department of Statistics, The Ohio State University.

FRIEDMAN, M. (1937). The use of ranks to avoid the assumption of normality implicit in the analysis of variance. *J. Amer. Statist. Assoc.* **32** 675–701.

GORDON, A. D. (1979). A measure of the agreement between rankings. *Biometrika* **66** 7–15.

HÁJEK, J. and SIDÁK, Z. (1967). *Theory of Rank Tests.* Academic Press, New York.

HETTMANSPERGER, T. (1985). *Statistical Inference Based on Ranks.* Wiley, New York.

HOLLANDER, M. , PROSCHAN, F., and SETHURAMAN, J. (1977). Functions decreasing in transposition and their application in ranking problems. *Ann. Statist.* **5** 722–733.

HOLLANDER, M. and WOLFE, D. (1973). *Nonparametric Statistical Methods.* Wiley, New York.

JONCKHEERE, A. R. (1954). A distribution-free k-sample test against ordered alternatives. *Biometrika* **41** 131–145.

KENDALL, M. G. (1962). *Rank Correlation Methods.* Third Edition. Griffin, New York.

KENDALL, M. G. and BABINGTON SMITH, B. (1939). The problem of m rankings. *Ann. Math. Statist.* **10** 275–287.

KRUSKAL, W. H. and WALLIS, W. A. (1952). Use of ranks in one-criterion variance analysis. *J. Amer. Statist. Assoc.* **47** 583–621.

LEHMANN, E. L. (1951). Consistency and unbiasedness of certain nonparametric tests. *Ann. Math. Statist.* **22** 165–179.

LEHMANN, E. L. (1959). *Testing Statistical Hypotheses.* Wiley, New York.

MALLOWS, C. L. (1957). Non-null ranking models I. *Biometrika* **44** 114–130.

MANN, H. B. and WHITNEY, D. R. (1947). On a test of whether one of two random variables is stochastically larger than the other. *Ann. Math. Statist.* **18** 50–60.

MOOD, A. M. (1950). *Introduction to the Theory of Statistics.* McGraw-Hill, New York.

NADLER, S. B. (1978). *Hyperspaces of Sets.* Marcel Dekker, New York.

PAGE, E. B. (1963). Ordered hypotheses for multiple treatments: A significance test for linear ranks. *J. Amer. Statist. Assoc.* **58** 216–230.

RAO, C. R. (1982a). Diversity and dissimilarity coefficients: A unified approach. *Journal of Theoretical Population Biology* **22**.

RAO, C. R. (1982b). Gini-Simpson index of diversity: A characterization, generalization, and applications. *Utilitas Math.* **21** 273–282.

RUKHIN, A. (1972). Estimation of a rotation parameter on a sphere. *Zap. Nauchn. Sem. Leningrad. Otdel. Mat. Inst. Steklov.* **29** 74–91 (in Russian). English Translation (1975): *J. Soviet Math.* **3** 777-791.

SAVAGE, I. R. (1956). Contributions to the theory of rank order statistics — the two-sample case. *Ann. Math. Statist.* **27** 590–615.

SIEGEL, S. and TUKEY, J. W. (1960). A nonparametric sum of ranks procedure for relative spread in unpaired samples. *J. Amer. Statist. Assoc.* **55** 429–445. Correction (1961): *J. Amer. Statist. Assoc.* **56** 1005.

TERPSTRA, T. J. (1954). A nonparametric test for the problem of k samples. *Nederl. Akad. Wetensch. Indag. Math.* **16** 505–512.

ULAM, S. M. (1972). Some ideas and prospects in biomathematics. *Annual Review of Biophysics and Bioengineering* **1** 277–292.

ULAM, S. M. (1981). Future applications of mathematics in the natural sciences. *American Mathematical Heritage: Algebra and Applied Mathematics.* Texas Tech. University Mathematics Series **13** 101–114.

WALD, A. and WOLFOWITZ, J. (1940). On a test of whether two samples are from the same population. *Ann. Math. Statist.* **11** 147–162.

WESTENBERG, J. (1948). A significance test for median and interquartile range in samples from continuous populations of any form. *Nederl. Akad. Wetensch. Proc. Ser. A* **51** 252–261.

WILCOXON, F. (1945). Individual comparisons by ranking methods. *Biometrics Bulletin* **1** 80–83.