Ann. Inst. Statist. Math.
Vol. 52, No. 4, 680–697 (2000)

NONPARAMETRIC METHODS FOR CHECKING THE VALIDITY OF PRIOR ORDER INFORMATION

GUOHUA PAN

Department of Mathematics and Statistics, Oakland University, Rochester, MI 48309, U.S.A.

(Received August 17, 1998; revised June 7, 1999)

Abstract. A large number of statistical procedures have been proposed in the literature to explicitly utilize available information about the ordering of treatment effects at increasing treatment levels. These procedures are generally more efficient than those ignoring the order information. However, when the assumed order information is incorrect, order restricted procedures are inferior and, strictly speaking, invalid. Just as any statistical model needs to be validated by data, order information to be used in a statistical analysis should also be justified by data first. A common statistical format for checking the validity of order information is to test the null hypothesis of the ordering representing the order information. Parametric tests for ordered null hypotheses have been extensively studied in the literature. These tests are not suitable for data with nonnormal or unknown underlying distributions. The objective of this study is to develop a general distribution-free testing theory for ordered null hypotheses based on rank order statistics and score generating functions. Sufficient and necessary conditions for the consistency of the proposed general tests are rigorously established.

Key words and phrases: Distribution-free test, lack-of-fit, ordered null hypothesis, order restricted inferences, partial order.

1. Introduction

In some scientific investigations, prior subject matter knowledge and past experience may provide valuable information about the ordering of treatment effects at increasing treatment levels, even though the exact magnitudes of the treatment effects are still unknown and need to be studied. This type of information, while still incomplete, can be useful in an investigation aimed to acquire a more specific understanding of the magnitudes of the responses. Many researchers have devised statistical procedures that explicitly utilize available order information. As expected, these procedures are typically more efficient than the corresponding omnibus ones without using the order information. The books by Barlow *et al.* (1972) and Robertson *et al.* (1988) summarize many order restricted procedures. Since then, an additional large number of publications on order restricted inferences have emerged.

However, perceived order information may deviate severely from the true model underlying the current data. When the assumed order information is moderately to severely misspecified, any order restricted statistical procedure performs poorly and, strictly speaking, is invalid. Therefore, just as model checking is indispensable in any statistical analysis, it is necessary to justify the validity of order information based on the current data before using it. In other words, the data should be allowed to speak for themselves about the assumed order information.

As a concrete example, consider the study conducted by Hundal (1969) to "assess the purely motivational effects of knowledge of performance in a repetitive industrial task. The task was to grind a metallic piece to a specified size and shape. Eighteen male workers were divided randomly into three groups. The subjects in the control group A received no information about their output, subjects in group B were given a rough estimate of their output, and subjects in group C were given accurate information about their output and could check it further by referring to a figure that was placed before them." (also see Hollander and Wolfe (1973), p. 121.) The basic data in Table 10 (Section 4) consist of the numbers of pieces processed by each subject in the experimental period. One question of interest to the investigator is whether different degrees of knowledge of performance affect the output. This can be answered by testing the null hypothesis of equal means for the three groups against the alternative hypothesis of unequal means. The Kruskal-Wallis test yields a P-value of 0.113, which indicates insufficient statistical evidence to reject the null hypothesis. Now suppose that past experience indicates that increased degree of knowledge of performance will not decrease productivity in this particular task. Then, the alternative space is restricted to increasing treatment effects, for which the Jonckheere-Terpstra test can be used. The resulting P-value is 0.0231, which indicates strong evidence of motivational effects on productivity. Order information (assuming its validity) not only provides an investigator with more efficient procedures, it can also help him select proper statistical inferences. For instance, in the motivational effects example above, it is probably more relevant to compare *successive* degrees of knowledge of performance using Lee and Spurrier (1995*a*, 1995*b*) than to conduct all pairwise comparisons. With all the benefits of utilizing order information, however, one should remember that the very basis for the use of order restricted procedures is the validity of the assumed order information. One needs to ask : *Do the data agree with the assumed order information?*

A common statistical format for checking the validity of order information is to test the null hypothesis of the ordering representing the order information. The nature of this type of testing is the same as that of a goodness-of-fit test. As D'Agostino and Stephens (1986) point out, "The major focus is on the measurement of agreement of the data with the null hypothesis; in fact, it is usually hoped to accept that H_0 (the null hypothesis) is true".

Several authors studied parametric procedures for testing the null hypotheses of a given ordering. Eeden (1958) considered the situation in which the null hypothesis imposed a simply ordered trend on the normal means. Perlman (1969) considered testing for null hypotheses defined by closed convex cones assuming multivariate normal distribution. Robertson and Wegman (1978) studied the distributions of likelihood ratio tests for ordered null hypotheses under normality and extended the results to exponential families. Mukerjee *et al.* (1986, 1987) studied multiple contrast tests and their power functions. Wollan and Dykstra (1986) considered the use of a conditional test and found that it was less biased than the likelihood ratio tests. Shi (1988) derived likelihood ratio statistics when the order restriction was defined by the positive orthant. Testing for ordered null hypothesis has also been studied in some nonnormal parametric settings. Robertson (1978) and Lee (1987) considered testing for the null hypothesis of an order restriction on multinomial parameters. Dykstra and Robertson (1982) obtained likelihood ratio statistics for testing the null hypothesis of star-shaped multinomial parameters. Singh and Wright (1990) considered testing for an order restriction in fixed effects models. Li and Sinha (1995) considered ordered null hypotheses about gamma scale parameters.

Prior order information is most valuable in applications with small to moderate sample sizes. In these applications, the assumption of parametric underlying distributions is usually untrue and cannot be efficiently verified. Due to insufficient sample sizes, the assumption of normal distributions cannot be justified by central limit theorems. A traditional solution to this kind of problem is to use a nonparametric test, which does not assume specific underlying distributions. Parsons (1979) considered rank analogies to the likelihood ratio tests for ordered null hypotheses. Shiraishi (1982) considered scores other than the Wilcoxon scores.

The major objective of this study is to develop a general distribution-free testing theory for ordered null hypotheses based on rank order statistics and general score functions. Special cases of this theory are several intuitive distribution-free tests analogous to the likelihood ratio and multiple contrast tests in a parametric setting. Sufficient and necessary conditions for the proposed tests to be consistent are rigorously established.

The layout of this paper is as follows. Section 2.1 introduces a general testing theory and proves the distribution-free property. Section 2.2 provides sufficient and necessary conditions for the consistency of the proposed tests. Section 3 presents power characterizations of several special tests based on simulation. Section 4 analyzes the motivational effects example using the proposed method. A summary is given in Section 5.

The rest of this section gives a mathematical description of the research question. Since most previous researches on this topic assume one-way models, this paper considers the corresponding nonparametric location-shift model. Extension of the proposed methodology to other nonparametric models such as two-way layouts is straightforward and will be discussed in Section 5.

Let $X_{11},\ldots,X_{1n_1},\ldots,X_{k1},\ldots,X_{kn_k}$ be k independent random samples from continuous distributions with cumulative distribution functions $F(X-\theta_1),\ldots,F(X-\theta_k)$, respectively, where θ_i denotes the median of the i-th population and $F(x)$ is completely unknown.

Most order information about $\theta_1,\ldots,\theta_k$ can be represented by inequalities among $\theta_1,\ldots,\theta_k$; that is, $\theta_s\le\theta_t$ for some pairs (s,t), where $1\le s,t\le k$. The order information, or equivalently, the corresponding inequalities, defines a subset Ω of R^k as the restricted parameter space for $\boldsymbol{\theta}=(\theta_1,\ldots,\theta_k)$. Some commonly seen orderings about $\theta_1,\ldots,\theta_k$, are as follows:

(i) Simple ordering: $\Omega_S=\{\boldsymbol{\theta}:\theta_1\le\cdots\le\theta_k\}$,

(ii) Simple tree ordering: $\Omega_T=\{\boldsymbol{\theta}:\theta_1\le\theta_i, i=2,\ldots,k\}$

(iii) Umbrella ordering with peak p: $\Omega_{Up}=\{\boldsymbol{\theta}:\theta_1\le\cdots\le\theta_p\ge\cdots\ge\theta_k\}$.

The question of checking the validity of prior order information is to test

$$H_0:\boldsymbol{\theta}\in\Omega \quad \text{versus} \quad H_1:\boldsymbol{\theta}\in R^k-\Omega. \tag{1.1}$$

A pair of populations, (s,t), $1\le s\ne t\le k$, will be called related in Ω if $\theta_s\le\theta_t$ for any $\boldsymbol{\theta}\in\Omega$. Furthermore, if there is no q, $1\le q\ne s$, $t\le k$, such that $\theta_s\le\theta_q\le\theta_t$ for all $\boldsymbol{\theta}\in\Omega$, then related pair (s,t) will be called directly related in Ω. Let $DC(\Omega)=\{(s,t)$: (s,t) is directly related in Ω, $1\le s\ne t\le k\}$. Then, $DC(\Omega)$ is the smallest set of related pairs that contains the complete order information. This is because deleting any pair from $DC(\Omega)$ leads to a less restrictive or informative ordering about $\theta_1,\ldots,\theta_k$. Let $C(\Omega)=\{(s,t)$: (s,t) is related in Ω, $1\le s\ne t\le k\}$. Then, $C(\Omega)$ is the largest set of related pairs that contains the complete order information. This is because adding any new pair to $C(\Omega)$ leads to a more restrictive or informative ordering about $\theta_1,\ldots,\theta_k$. The minimal set $DC(\Omega)$ and maximal set $C(\Omega)$ for the three orderings (i), (ii) and (iii) are, respectively,

(i) Simple ordering: $DC(\Omega_s)=\{(i,i+1):i=1,\ldots,k-1\}$

$$C(\Omega_s)=\{(i,j):1\le i<j\le k\},$$

(ii) Simple tree ordering: $DC(\Omega_T) = C(\Omega_T) = \{(1,i) : i = 2, \ldots, k\}$,
(iii) Umbrella ordering with peak p:

$$DC(\Omega_{U_p}) = \{(i,j) : j = i+1 \text{ for } 1 \le i \le p-1, \text{ and } j = i-1 \text{ for } p+1 \le i \le k\}$$
$$C(\Omega_{U_p}) = \{(i,j) : 1 \le i < j \le p, \text{ and } p \le j < i \le k\}.$$

Obviously, $DC(\Omega_{U_k}) = DC(\Omega_s)$ and $C(\Omega_{U_k}) = C(\Omega_s)$.

It is easy to see that any set $B(\Omega)$ satisfying $DC(\Omega) \subseteq B(\Omega) \subseteq C(\Omega)$ contains the complete order information about $\theta_1, \ldots, \theta_k$. Thus, each such $B(\Omega)$ provides a basis for checking the validity of the order information.

2. A general nonparametric testing theory

Section 2.1 provides a general method for constructing distribution-free tests for testing question (1.1): $H_0 : \boldsymbol{\theta} \in \Omega$ versus $H_1 : \boldsymbol{\theta} \in R^k - \Omega$. Section 2.2 establishes sufficient and necessary conditions for the proposed tests to be consistent.

2.1 *Test Statistics and Distribution-Free Property*

Let $R_j^{(st)}$ be the rank of X_{tj} in the sample of $m_{st} = n_s + n_t$ items obtained by combining the s-th and t-th samples. Obviously, $R_j^{(st)}$ is also a function of the joint rank vector $\boldsymbol{R} = (R_{11}, \ldots, R_{1n_1}, \ldots, R_{k1}, \ldots, R_{kn_k})$, where R_{ij} is the rank of X_{ij} among all the $N = n_1 + \ldots + n_k$ observations from the k samples combined.

Consider general scores $\{a_{m_{st}}(i)\}_{i=1}^{m_{st}}$ for the s-th and t-th samples satisfying

$$a_{m_{st}}(1) \le a_{m_{st}}(2) \le \cdots \le a_{m_{st}}(m_{st}).$$

A two-sample linear rank statistic for samples s and t is of the form

$$M'_{st}(\boldsymbol{R}) = \sum_{j=1}^{n_t} a_{m_{st}}(R_j^{(st)}).$$

As an example, the Mann-Whitney-Wilcoxon rank sum statistic

$$W'_{st}(\boldsymbol{R}) = \sum_{j=1}^{n_t} R_j^{(st)}$$

corresponds to scores $a_{m_{st}}(i) = i$, $i = 1, \ldots, m_{st}$. The expected value and variance of $M'_{st}(\boldsymbol{R})$ under $\theta_s = \theta_t$ are, respectively,

$$E_{\theta_s=\theta_t}(M'_{st}) = n_t \bar{a}_{m_{st}}$$

and

$$\sigma^2_{0:m_{st}} = \text{Var}_{\theta_s=\theta_t}(M'_{st}(\boldsymbol{R})) = \frac{n_s n_t}{m_{st}}(m_{st}-1)\sum_{i=1}^{m_{st}}(a_{m_{st}}(i) - \bar{a}_{m_{st}})^2,$$

where

$$\bar{a}_{m_{st}} = \frac{1}{m_{st}}\sum_{i=1}^{m_{st}} a_{m_{st}}(i).$$

To facilitate the proof of consistency of the proposed tests in the next subsection, the following version of the standardized $M'_{st}(\boldsymbol{R})$ is used:

$$M_{st}(\boldsymbol{R}) = \frac{M'_{st}(\boldsymbol{R}) - n_t a_{m_{st}}}{\sqrt{N}\sigma_{0:m_{st}}}.$$

The standardized Mann-Whitney-Wilcoxon statistic used in this paper is thus

$$W_{st}(\boldsymbol{R}) = \frac{W'_{st}(\boldsymbol{R}) - E_{\theta_s=\theta_t}(W'_{st}(\boldsymbol{R}))}{\sqrt{N}\sqrt{\mathrm{var}_{\theta_s=\theta_t}(W'_{st}(\boldsymbol{R}))}} = \frac{\sum_{j=1}^{n_t} R_j^{(st)} - n_t(m_{st}+1)/2}{\sqrt{N}\sqrt{n_s n_t(m_{st}+1)/12}}.$$

See Chapter 9 of Randles and Wolfe (1991) or Chapter 3 of Hettmansperger (1991) for details.

For any fixed $B(\Omega)$ satisfying $DC(\Omega) \subseteq B(\Omega) \subseteq C(\Omega)$, let $|B|$ denote the number of elements of $B(\Omega)$. For example, $|DC(\Omega_S)| = k-1$ and $|C(\Omega_S)| = k(k+1)/2$. Without loss of generality, the $|B|$ statistics, $M_{st}(\boldsymbol{R})$, $(s,t) \in B(\Omega)$, will be arranged according to their indices (s,t)'s so that the $|DC|$ pairs $(s,t) \in DC(\Omega)$ are in lexicographic order, followed by the remaining $|B|-|DC|$ pairs $(s,t) \in B(\Omega) - DC(\Omega)$, again, in lexicographic order. The arranged $M_{st}(\boldsymbol{R})$, $(s,t) \in B(\Omega)$ will be denoted by $M_1(\boldsymbol{R}), \ldots, M_{|B|}(\boldsymbol{R})$.

Suppose that $l_{\boldsymbol{n}}(x_1, \ldots, x_{|B|})$ is a function on $R^{|B|}$ and $k(\boldsymbol{n}, B(\Omega)) > 0$ is a positive scaling constant. This paper proposes the following test statistic:

$$T(\boldsymbol{R}, \boldsymbol{n}, B(\Omega)) = k(\boldsymbol{n}, B(\Omega)) \cdot l_{\boldsymbol{n}}(M_1(\boldsymbol{R}), \ldots, M_{|B|}(\boldsymbol{R})),$$

where $\boldsymbol{n} = (n_1, \ldots, n_k)$.

The following are three natural special cases of $T(\boldsymbol{R}, \boldsymbol{n}, B(\Omega))$:

(1) *Nonparametric Multiple Pairwise Contrast Test Statistic*:

$$C^*_{B(\Omega)}(\boldsymbol{R}) = \min_{(s,t)\in B(\Omega)} W_{st}(\boldsymbol{R})$$

corresponding to $k(\boldsymbol{n}, B(\Omega)) = \sqrt{N}$ and $l_{\boldsymbol{n}}^{(C^*)}(x_1, \ldots, x_{|B|}) = \min_{1\le i\le |B|} x_i$.

(2) *Test Statistics Based on Distances to the Null Hypothesis Space*:

$$D_{B(\Omega)}(\boldsymbol{R}) = \sum_{(s,t)\in B(\Omega)} g(\max\{-W_{st}(\boldsymbol{R}), 0\})$$

corresponding to $k(\boldsymbol{n}, B(\Omega)) = \sqrt{N}$ and $l_{\boldsymbol{n}}^{(D)}(x_1, \ldots, x_{|B|}) = -\sum_{i=1}^{|B|} g(\max\{-x_i, 0\})$, where $g(x)$ is an increasing function of $x \ge 0$. The most natural choices of $g(x)$ are $g(x) = x$ and $g(x) = x^2$.

(3) *Tryon-Hettmansperger* (1973) *Statistic*:

$$H_{\boldsymbol{h}}(\boldsymbol{R}) = \sum_{(s,t)\in B(\Omega)} h_{st} W_{st}$$

corresponding to $k(\boldsymbol{n}, B(\Omega)) = \sqrt{N}$ and $l_{\boldsymbol{n}}^{(H)}(x_1, \ldots, x_{|B|}) = \sum_{i=1}^{|B|} h_i x_i$, where $\boldsymbol{h} = \{h_{st} > 0, (s,t) \in B(\Omega)\}$ are positive coefficients.

Obviously, the constant $k(\boldsymbol{n}, B(\Omega))$ in the test statistic $T(\boldsymbol{R}, \boldsymbol{n}, B(\Omega))$ does not affect the power performance of the test and can be chosen for convenience. It is usually desirable and relatively straightforward to choose $k(\boldsymbol{n}, B(\Omega))$ so that $T(\boldsymbol{R}, \boldsymbol{n}, B(\Omega))$ has a nondegenerate limiting distribution.

This paper proposes the following test for $H_0 : \boldsymbol{\theta} \in \Omega$ versus $H_1 : \boldsymbol{\theta} \in R^k - \Omega$.

Test $T(\boldsymbol{R}, \boldsymbol{n}, B(\Omega))$: Reject $H_0 : \boldsymbol{\theta} \in \Omega$ if and only if

$$T(\boldsymbol{R}, \boldsymbol{n}, B(\Omega)) < t(\boldsymbol{n}, B(\Omega)),$$

where the cutoff value $t(\boldsymbol{n}, B(\Omega))$ is the largest constant satisfying

$$P\{T(\boldsymbol{R}^*, \boldsymbol{n}, B(\Omega)) \geq t(\boldsymbol{n}, B(\Omega))\} \geq 1 - \alpha,$$

and $\boldsymbol{R}^*$ is uniformly distributed over the set of all permutations of $\{1, \ldots, N\}$.

The next theorem states that test $T(\boldsymbol{R}, \boldsymbol{n}, B(\Omega))$ is of a constant size no more than α under a very mild condition about $l_{\boldsymbol{n}}(x_1, \ldots, x_{|B|})$. The proof is in the Appendix.

THEOREM 2.1. *Suppose that for any* $i = 1, \ldots, |B|$, $l_{\boldsymbol{n}}(x_1, \ldots, x_{|B|})$ *is an increasing function of* x_i *for fixed* $x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_{|B|}$. *Then*

$$\max_{\boldsymbol{\theta} \in H_0, F(x)} P_{\boldsymbol{\theta}, F(x)}\{T(\boldsymbol{R}, \boldsymbol{n}, B(\Omega)) < t(\boldsymbol{n}, B(\Omega))\} = c(l_{\boldsymbol{n}}, B) \leq \alpha,$$

where $c(l_{\boldsymbol{n}}, B)$ *is a constant that does not depend on* $F(x)$.

Theorem 2.1 can be applied easily to show that $C^*_{B(\Omega)}(\boldsymbol{R})$, $D_{B(\Omega)}(\boldsymbol{R})$, and $H_{\boldsymbol{h}}(\boldsymbol{R})$ are all level α tests, because the functions $l_{\boldsymbol{n}}^{(C^*)}(x_1, \ldots, x_{|B|})$, $l_{\boldsymbol{n}}^{(D)}(x_1, \ldots, x_{|B|})$, and $l_{\boldsymbol{n}}^{(H)}(x_1, \ldots, x_{|B|})$ used in these tests are all increasing functions.

Since the cutoff value $t(\boldsymbol{n}, B(\Omega))$ does not depend on $F(x)$, the following corollary is immediate. Thus, tests $C^*_{B(\Omega)}(\boldsymbol{R})$, $D_{B(\Omega)}(\boldsymbol{R})$, and $H_{\boldsymbol{h}}(\boldsymbol{R})$ are all distribution-free.

COROLLARY 2.1. *Under the condition of Theorem* 2.1, *test* $T(\boldsymbol{R}, \boldsymbol{n}, B(\Omega))$ *is distribution-free.*

2.2 *Sufficient and necessary conditions for consistency*

This subsection provides easy-to-check sufficient and necessary conditions for the proposed test $T(\boldsymbol{R}, \boldsymbol{n}, B(\Omega))$ to be consistent. In virtually all the useful rank-based procedures, the scores $\{a_{m_{st}}\}_{i=1}^{m_{st}}$ can be obtained from a strictly increasing square integrable score generating function $\phi : (0, 1) \to R$ as follows:

$$a_{m_{st}}(i) = b_{m_{st}} \phi\left(\frac{i}{m_{st} + 1}\right) + d_{m_{st}}, \qquad i = 1, \ldots, m_{st},$$

where $b_{m_{st}} > 0$ and $d_{m_{st}}$ are constants, and

$$0 < \int_0^1 (\phi(u) - \bar{\phi})^2 \, du < \infty \tag{2.1}$$

with $\bar{\phi} = \int_0^1 \phi(u) du$. It is easy to check that $M_{st}(\boldsymbol{R})$ does not depend on $b_{m_{st}}$ or $d_{m_{st}}$. Therefore, one can always let $b_{m_{st}} = 1$ and $d_{m_{st}} = 0$, which will be the case in the rest of this paper. A mild condition for the limiting normality of $M_{st}(\boldsymbol{R})$ is given by

$$|\phi(u)| \leq k[u(1-u)]^{\delta - 1/2} \quad \text{and} \quad |\phi'(u)| \leq k[u(1-u)]^{\delta - 3/2} \tag{2.2}$$

for some $\delta > 0$ and $k > 0$. See p. 95 of Puri and Sen (1971) for details. Virtually all commonly used score generating functions satisfy this condition.

The following theorem provides a sufficient condition for the consistency of $T(\boldsymbol{R}, \boldsymbol{n}, B(\Omega))$. The proof is in the Appendix.

THEOREM 2.2. *Assume conditions* (2.1) *and* (2.2) *about the score generating function* $\phi(u)$. *Suppose that* $n_i/N \to \lambda_i \in (0,1)$ *as* $N = n_1 + \cdots + n_k \to +\infty$. *Assume that increasing function* $l_{\boldsymbol{n}}(x_1, \ldots, x_{|B|})$ *is continuous and converges to a continuous function* $l(x_1, \ldots, x_{|B|})$ *as* $N \to +\infty$. *If*

$$l(x_1, \ldots, x_{|B|}) < l(0, \ldots, 0) \quad \textit{whenever} \quad \min_{1 \le i \le |B|} x_i < 0, \tag{2.3}$$

then, for any fixed $\boldsymbol{\theta} \in H_1$ and $F(x)$,

$$\lim_{N \to +\infty} P_{\boldsymbol{\theta}, F(x)} \{T(\boldsymbol{R}, \boldsymbol{n}, B(\Omega)) < t(\boldsymbol{n}, B(\Omega))\} = 1.$$

The sufficient condition (2.3) states that if $l_{\boldsymbol{n}}(x_1, \ldots, x_{|B|})$ is less than $l_{\boldsymbol{n}}(0, \ldots, 0)$ whenever one or more x_i's are negative, then the corresponding test $T(\boldsymbol{R}, \boldsymbol{n}, B(\Omega))$ is consistent.

It is easy to check that the functions $l_{\boldsymbol{n}}^{(C^*)}(x_1, \ldots, x_{|B|})$ and $l_{\boldsymbol{n}}^{(D)}(x_1, \ldots, x_{|B|})$ used in tests $C^*_{B(\Omega)}(\boldsymbol{R})$ and $D_{B(\Omega)}(\boldsymbol{R})$ satisfy condition (2.3). Therefore, tests $C^*_{B(\Omega)}(\boldsymbol{R})$ and $D_{B(\Omega)}(\boldsymbol{R})$ are consistent. However, the function $l_{\boldsymbol{n}}^{(H)}(x_1, \ldots, x_{|B|})$ used in test $H_{\boldsymbol{h}}(\boldsymbol{R})$ does not satisfy this condition.

The following slightly different version of the sufficient condition puts more emphasis on $DC(\Omega)$. It may be more convenient to use when $B(\Omega)$ is much larger than $DC(\Omega)$. The proof is essentially the same as that of Theorem 2.2 using the fact that for any fixed $\boldsymbol{\theta} \in H_1$, there must exist some $(s,t) \in DC(\Omega)$ such that $\theta_s > \theta_t$.

COROLLARY 2.2. *A sufficient condition for the consistency of test* $T(\boldsymbol{R}, \boldsymbol{n}, B(\Omega))$ *is to replace condition* (2.3) *in Theorem* 2.2 *by*

$$l(x_1, \ldots, x_{|B|}) < l(0, \ldots, 0) \quad \textit{whenever} \quad \min_{1 \le i \le |DC|} x_i < 0.$$

It can be shown that condition (2.3) may not be necessary for the consistency of test $T(\boldsymbol{R}, \boldsymbol{n}, B(\Omega))$, but a slightly weaker condition is. The following theorem states that a necessary condition for the consistency of test $T(\boldsymbol{R}, \boldsymbol{n}, B(\Omega))$ is that the function $l_{\boldsymbol{n}}(x_1, \ldots, x_{|B|})$ is less than $l_{\boldsymbol{n}}(0, \ldots, 0)$ whenever one or more x_i's are negative in a *neighborhood* of $(0, \ldots, 0)$. Notice that the only difference between the necessary and sufficient conditions in Theorems 2.2 and 2.3 is that the sufficient condition (2.3) requires the above inequality hold for *any* $(x_1, \ldots, x_{|B|})$ with one or more negative x_i's. The proof of the following theorem is in the Appendix.

THEOREM 2.3. *Suppose that* $n_i/N \to \lambda_i \in (0,1)$ *as* $N = n_1 + \cdots + n_k \to +\infty$. *Assume that increasing function* $l_{\boldsymbol{n}}(x_1, \ldots, x_{|B|})$ *is continuous and converges to a continuous function* $l(x_1, \ldots, x_{|B|})$ *as* $N \to +\infty$. *If test* $T(\boldsymbol{R}, \boldsymbol{n}, B(\Omega))$ *is consistent*; *that is*, *for any fixed* $\boldsymbol{\theta} \in H_1$ and fixed $F(x)$,

$$\lim_{N \to +\infty} P_{\boldsymbol{\theta}, F(x)} \{T(\boldsymbol{R}, \boldsymbol{n}, B(\Omega)) < t(\boldsymbol{n}, B(\Omega))\} = 1,$$

Table 1. Powers for simple ordering when $(n_1, n_2, n_3) = (5, 5, 5)$.

$(\theta_1, \theta_2, \theta_3)$	$B(\Omega_S)$	C^*	D_1	D_2	H
(0, 0, 0)	C	.0433	.0448	.0497	.0461
	$t(\boldsymbol{n}, C(\Omega_S))$	−1.984	−3.551	−5.793	−3.238
	DC	.0319	.0461	.0381	.0452
	$t(\boldsymbol{n}, DC(\Omega_S))$	−1.984	−2.089	−3.938	−1.671
(0, 1, 0)	C	.216	.068	.147	.028
	DC	.211	.215	.213	.024
(0, 2, 1)	C	.204	.002	.079	.000
	DC	.204	.204	.204	.000
(1, 0, 0)	C	.331	.403	.420	.378
	DC	.228	.341	.292	.330
(1, 0, 2)	C	.210	.002	.085	.000
	DC	.210	.210	.210	.026
(1, 1, 0)	C	.334	.400	.424	.375
	DC	.219	.334	.284	.328
(1, 2, 0)	C	.695	.525	.686	.217
	DC	.671	.683	.682	.120
(2, 0, 1)	C	.700	.526	.686	.215
	DC	.676	.687	.685	.117
(2, 1, 0)	C	.733	.868	.846	.879
	DC	.399	.757	.642	.836

then there exists a $\delta > 0$ *such that*

(2.4) $\quad l(x_1, \ldots, x_{|B|}) < l(0, \ldots, 0)$ *whenever* $|x_i| < \delta,\ i = 1, \ldots, |B|$, *and* $\min_{1 \le i \le |B|} x_i < 0$.

Condition (2.4) in the above theorem can be used to show the inconsistency of a given test. Consider the function $l_{\boldsymbol{n}}^{(H)}(x_1, \ldots, x_{|B|}) = \sum_{i=1}^{|B|} h_i x_i$ used in test $H_{\boldsymbol{h}}(\boldsymbol{R})$ with $|B| \ge 2$. Obviously, $l_{\boldsymbol{n}}^{(H)}(0, \ldots, 0) = 0$. Without loss of generality, assume that $h_1 \ge h_2 > 0$. Then, $l_{\boldsymbol{n}}^{(H)}(x+1, -x, 0, \ldots, 0) = (h_1 - h_2)x + h_1 > 0$ for any $-x < 0$, which violates condition (2.4). Therefore, test $H_{\boldsymbol{h}}(\boldsymbol{R})$ is inconsistent except for the trivial case when $|B| = 1$.

3. Characterization of Power

This section studies the power characterizations of the four tests $C^*_{B(\Omega)}(\boldsymbol{R})$, $D_{B(\Omega)}(\boldsymbol{R})$ with $g(x) = x$ (denoted as D_1), $D_{B(\Omega)}(\boldsymbol{R})$ with $g(x) = x^2$ (denoted as D_2), and $H_{\boldsymbol{h}}(\boldsymbol{R})$. While test $H_{\boldsymbol{h}}(\boldsymbol{R})$ is not consistent, it is included here as a comparison. Three types of ordered null hypotheses are considered: simple ordering (Ω_S), simple tree ordering (Ω_T), and umbrella ordering (Ω_{U_p}). For each type of null hypothesis, two choices of $B(\Omega)$ are investigated: $B(\Omega) = C(\Omega)$ and $B(\Omega) = DC(\Omega)$. Since $C(\Omega)$ and $DC(\Omega)$ are the largest and smallest $B(\Omega)$, their contrast will shed light on the impact of $B(\Omega)$ on the power performance. The number k of samples considered in the simulation study ranges from 3 to 6. A number of combinations of balanced and unbalanced sample

Table 2. Powers for simple ordering when $(n_1, n_2, n_3) = (5, 10, 15)$.

$(\theta_1, \theta_2, \theta_3)$	$B(\Omega_S)$	C^*	D_1	D_2	H
(0, 0, 0)	C	.0477	.0500	.0499	.0500
	$t(\boldsymbol{n}, C(\Omega_S))$	−2.052	−3.566	−6.045	−3.327
	DC	.0436	.0481	.0481	.0499
	$t(\boldsymbol{n}, DC(\Omega_S))$	−1.960	−2.108	−3.840	−1.764
(0, 1, 0)	C	.573	.18	.433	.064
	DC	.620	.551	.621	.086
(0, 2, 1)	C	.576	.032	.386	.000
	DC	.625	.547	.625	.001
(1, 0, 0)	C	.468	.556	.545	.507
	DC	.356	.452	.414	.426
(1, 0, 2)	C	.343	.004	.137	.013
	DC	.343	.282	.343	.006
(1, 1, 0)	C	.681	.693	.725	.638
	DC	.633	.719	.689	.655
(1, 2, 0)	C	.994	.936	.989	.582
	DC	.995	.994	.996	.535
(2, 0, 1)	C	.881	.749	.834	.204
	DC	.877	.833	.878	.052
(2, 1, 0)	C	.939	.979	.971	.982
	DC	.780	.963	.917	.979

sizes $(n_1, \ldots, n_k)$ are examined. The International Mathematics and Statistics Library (IMSL) of FORTRAN subroutines is used. The critical values $t(\boldsymbol{n}, B(\Omega))$ and sizes $\max_{\boldsymbol{\theta}\in H_0,\, F(x)} P_{\boldsymbol{\theta},F(x)}\{T(\boldsymbol{R},\boldsymbol{n},B(\Omega)) < t(\boldsymbol{n},B(\Omega))\} = c(l_{\boldsymbol{n}}, B)$ of a test are estimated based on 1000000 iterations. The power of a test $P_{\boldsymbol{\theta}\in H_1,F(x)}\{T(\boldsymbol{R},\boldsymbol{n},B(\Omega)) < t(\boldsymbol{n},B(\Omega))\}$ at an alternative $\boldsymbol{\theta} \in H_1$ is estimated based on 10000 iterations. Note that the sizes of the tests are not necessarily equal to α due to the discreteness of the test statistics. One can adjust the sizes to α by using the corresponding randomized tests. However, the nonrandomized tests are the ones usually applied in practice. As can be seen later, the power comparisons are more informative without the adjustments of the sizes.

The general observations in the next paragraph are based on the extensive study described above. But for the sake of space, only the simulated powers for selected cases of k, $\boldsymbol{n} = (n_1, \ldots, n_k)$ and $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_k)$ and normal underlying distribution $F(x)$ are reported here in Tables 1–9. Since the size of a test is the power of the test when $\theta_1 = \cdots = \theta_k = 0$, the rows in Tables 1–9 corresponding to $\theta_1 = \cdots = \theta_k = 0$ provide the sizes and critical values of the tests.

The following observations can be made on the power characterizations. First, no test is uniformly most powerful. This is expected since the alternative space H_1 is typically a very vast region. Second, test C^* has the best overall power performance among the four tests. Test C^* is especially powerful at an alternative where only a small number of pairs of related populations violate the null ordering. Test D_2 outperforms test D_1, but both tests can be very inefficient for some conceivable alternatives in practice when k is moderate, say, $k = 5$. Third, for a given alternative and for the same type of test, the power is higher if $B(\Omega)(\supseteq DC(\Omega))$ includes all and only the pairs of related populations in the alternative that violate the null ordering. Therefore, if the most likely

Table 3. Powers for simple ordering when $(n_1, n_2, n_3, n_4, n_5, n_6) = (5, 5, 5, 5, 5, 5)$.

$(\theta_1, \theta_2, \theta_3, \theta_4, \theta_5, \theta_6)$	$B(\Omega_S)$	C^*	D_1	D_2	H
(0, 0, 0, 0, 0, 0)	C	.0478	.0491	.0500	.0492
	$t(\boldsymbol{n}, C(\Omega_S))$	−2.402	−12.429	−19.669	−9.505
	DC	.0393	.0376	.0500	.0460
	$t(\boldsymbol{n}, DC(\Omega_S))$	−2.193	−3.760	−6.415	−1.984
(0, 0, 0, 0, 1, 0)	C	.098	.022	.040	.006
	DC	.152	.158	.191	.038
(0, 0, 1, 1, 0, 0)	C	.224	.118	.197	.034
	DC	.148	.129	.160	.058
(0, 0, 2, 1, 1, 0)	C	.449	.147	.352	.006
	DC	.250	.427	.403	.124
(0, 1, 0, 1, 1, 0)	C	.236	.094	.193	.017
	DC	.253	.336	.369	.037
(0, 1, 1, 1, 0, 2)	C	.174	.007	.062	.008
	DC	.143	.054	.127	.000
(0, 1, 2, 1, 1, 2)	C	.132	.001	.015	.002
	DC	.131	.014	.091	.003
(0, 2, 0, 1, 1, 2)	C	.412	.005	.086	.003
	DC	.534	.096	.438	.001
(0, 2, 1, 1, 1, 1)	C	.215	.082	.196	.011
	DC	.151	.117	.170	.005
(1, 0, 1, 1, 0, 2)	C	.230	.038	.148	.001
	DC	.250	.331	.370	.002
(1, 0, 2, 1, 0, 1)	C	.487	.310	.480	.046
	DC	.343	.653	.605	.099
(1, 1, 0, 1, 0, 0)	C	.344	.615	.599	.509
	DC	.257	.446	.412	.301
(1, 1, 1, 0, 2, 2)	C	.175	.005	.053	.010
	DC	.147	.142	.171	.016
(1, 1, 2, 1, 0, 2)	C	.455	.157	.371	.011
	DC	.255	.303	.358	.005
(1, 2, 0, 1, 0, 2)	C	.591	.359	.600	.016
	DC	.592	.499	.672	.001
(1, 2, 1, 1, 0, 1)	C	.476	.564	.630	.357
	DC	.244	.307	.345	.039
(2, 0, 1, 0, 2, 2)	C	.559	.068	.339	.000
	DC	.586	.610	.714	.041
(2, 0, 2, 0, 2, 1)	C	.692	.506	.734	.024
	DC	.808	.970	.965	.252
(2, 1, 0, 0, 2, 0)	C	.754	.885	.920	.402
	DC	.644	.948	.907	.759
(2, 1, 1, 0, 1, 2)	C	.462	.314	.485	.042
	DC	.247	.314	.350	.045
(2, 1, 2, 0, 1, 1)	C	.609	.718	.781	.419
	DC	.597	.633	.721	.163
(2, 2, 0, 0, 1, 0)	C	.825	.982	.982	.898
	DC	.603	.743	.784	.570
(2, 2, 1, 0, 0, 2)	C	.745	.892	.92	.404
	DC	.263	.414	.400	.120

Table 4. Powers for simple tree ordering when $(n_1, n_2, n_3) = (5, 5, 5)$.

$(\theta_1, \theta_2, \theta_3)$	$B(\Omega_T)$	C^*	D_1	D_2	H
(0, 0, 0)	$C = DC$	.0295	.0496	.0479	.0496
	$t(\boldsymbol{n}, B(\Omega_T))$	−1.984	−2.716	−4.811	−2.716
(1, 0, 0)	$C = DC$	.319	.509	.483	.509
(1, 0, 1)	$C = DC$	.210	.174	.229	.174
(1, 0, 2)	$C = DC$	.211	.018	.145	.018
(2, 0, 0)	$C = DC$	.828	.951	.942	.951
(2, 0, 1)	$C = DC$	.701	.765	.798	.765
(2, 0, 2)	$C = DC$	.685	.320	.629	.320

Table 5. Powers for simple tree ordering when $(n_1, n_2, n_3) = (5, 10, 15)$.

$(\theta_1, \theta_2, \theta_3)$	$B(\Omega_T)$	C^*	D_1	D_2	H
(0, 0, 0)	$C = DC$	.0453	.0496	.0498	.0495
	$t(\boldsymbol{n}, B(\Omega_T))$	−1.877	−3.014	−4.868	−3.014
(1, 0, 0)	$C = DC$	.539	.604	.596	.604
(1, 0, 1)	$C = DC$	.392	.205	.296	.203
(1, 0, 2)	$C = DC$	.390	.031	.229	.013
(2, 0, 0)	$C = DC$	.969	.985	.984	.985
(2, 0, 1)	$C = DC$	.912	.863	.898	.862
(2, 0, 2)	$C = DC$	.909	.438	.801	.383

Table 6. Powers for simple tree ordering when $(n_1, n_2, n_3, n_4, n_5, n_6) = (5, 5, 5, 5, 5, 5)$.

$(\theta_1, \theta_2, \theta_3, \theta_4, \theta_5, \theta_6)$	$B(\Omega_T)$	C^*	D_1	D_2	H
(0, 0, 0, 0, 0, 0)	$C = DC$	.0325	.0498	.0500	.0499
	$t(\boldsymbol{n}, B(\Omega_T))$	−2.193	−6.267	−9.829	−6.162
(1, 0, 0, 0, 0, 0)	$C = DC$	0.355	0.606	0.582	0.609
(1, 0, 0, 0, 0, 2)	$C = DC$	0.320	0.421	0.469	0.256
(1, 0, 0, 0, 1, 2)	$C = DC$	0.276	0.237	0.335	0.136
(1, 0, 0, 1, 2, 2)	$C = DC$	0.220	0.048	0.156	0.022
(1, 0, 1, 2, 2, 2)	$C = DC$	0.130	0.005	0.019	0.004
(1, 0, 2, 2, 2, 2)	$C = DC$	0.125	0.001	0.003	0.001
(1, 0, 1, 2, 2, 1)	$C = DC$	0.133	0.016	0.046	0.010
(2, 0, 0, 0, 0, 0)	$C = DC$	0.854	0.983	0.977	0.984
(2, 0, 0, 0, 1, 2)	$C = DC$	0.781	0.861	0.903	0.785
(2, 0, 0, 1, 2, 2)	$C = DC$	0.705	0.555	0.751	0.455
(2, 0, 1, 2, 2, 2)	$C = DC$	0.564	0.230	0.388	0.208

alternatives are expected to violate the null ordering for some pairs of related populations not in $DC(\Omega)$, then $DC(\Omega)$ can be augmented with these pairs accordingly to obtain a $B(\Omega)$ that improves the power performance. As an example, if the null hypothesis of nondecreasing treatment effects, i.e., simple ordering Ω_S, is under consideration and the alternatives are most likely to have an up-then-down pattern with peaks at high

Table 7. Powers for umbrella ordering when $(n_1, n_2, n_3) = (5, 5, 5)$ and $p = 2$.

$(\theta_1, \theta_2, \theta_3)$	$B(\Omega_{U_2})$	C^*	D_1	D_2	H
(0, 0, 0)	$C = DC$	.0295	.0497	.0478	.0497
	$t(\boldsymbol{n}, B(\Omega_{U_2}))$	−1.984	−2.716	−4.811	−2.716
(0, 0, 1)	$C = DC$	.205	.170	.222	.170
(0, 0, 2)	$C = DC$	.679	.317	.616	.317
(1, 0, 0)	$C = DC$	.217	.178	.237	.178
(1, 0, 1)	$C = DC$	.327	.517	.490	.517
(1, 0, 2)	$C = DC$	.693	.760	.792	.760
(2, 0, 0)	$C = DC$	.684	.331	.626	.331
(2, 0, 2)	$C = DC$	.823	.948	.939	.948
(0, 1, 2)	$C = DC$	.209	.021	.141	.021

Table 8. Powers for umbrella ordering when $(n_1, n_2, n_3) = (5, 10, 15)$ and $p = 2$.

$(\theta_1, \theta_2, \theta_3)$	$B(\Omega_{U_2})$	C^*	D_1	D_2	H
(0, 0, 0)	$C = DC$	.0474	.0500	.0497	.0499
	$t(\boldsymbol{n}, B(\Omega_{U_2}))$	−1.941	−2.787	−4.584	−2.780
(0, 0, 1)	$C = DC$	.617	.387	.565	.343
(0, 0, 2)	$C = DC$	.994	.930	.989	.721
(1, 0, 0)	$C = DC$	.396	.240	.334	.228
(1, 0, 1)	$C = DC$	.698	.758	.754	.757
(1, 0, 2)	$C = DC$	.995	.981	.994	.972
(2, 0, 0)	$C = DC$	.909	.591	.850	.448
(2, 0, 2)	$C = DC$	.998	1.000	.999	1.000
(0, 1, 2)	$C = DC$	.616	.232	.539	.053

treatment levels, then a reasonable choice of $B(\Omega_S)$ is to augment $DC(\Omega_S) = \{(i, i+1) : i = 1, \ldots, k-1\}$ with $\{(i,j) : (k+1)/2 \le i, 2 \le j - i; i, j = 1, \ldots, k\}$. Thus, for instance, for $k = 5$, use $B(\Omega_S) = \{(1,2), (2,3), (3,4), (4,5), (3,5)\}$; and for $k = 6$, use $B(\Omega_S) = \{(1,2), (2,3), (3,4), (4,5), (5,6), (3,5), (3,6), (4,6)\}$. In most applications, however, no information regarding the alternatives is available. As a practical guide for potential users of the proposed tests, a rule of thumb is to use test C^* with $B(\Omega) = DC(\Omega)$. The reason is that in many applications essential violations of a null ordering at an alternative usually occur for a small number of pairs of directly related populations. For these alternatives, test C^* using $B(\Omega) = DC(\Omega)$ is generally competitive compared with test C^* using $B(\Omega) = C(\Omega)$, but the latter can be inconvenient to implement when k is moderately large.

4. Implementation and an example

To compute the cutoff value $t(\boldsymbol{n}, B(\Omega))$ in test $T(\boldsymbol{R}, \boldsymbol{n}, B(\Omega))$, one can exhaust all $N!/(\prod_{i=1}^k n_i)$ assignments of ranks to the k groups and compute $T(\boldsymbol{R}, \boldsymbol{n}, B(\Omega))$ for each assignment. The frequency table of all the $N!/(\prod_{i=1}^k n_i)$ values of $T(\boldsymbol{R}, \boldsymbol{n}, B(\Omega))$ provides the exact distribution of $T(\boldsymbol{R}^*, \boldsymbol{n}, B(\Omega))$ and thus $t(\boldsymbol{n}, B(\Omega))$. This method is only feasible when N is small or moderate. When N is relatively large, one can use standard Monte

Table 9. Powers for umbrella ordering when $(n_1, n_2, n_3, n_4, n_5, n_6) = (5, 5, 5, 5, 5, 5)$ and $p = 4$.

$(\theta_1, \theta_2, \theta_3, \theta_4, \theta_5, \theta_6)$	$B(\Omega_{U_4})$	C^*	D_1	D_2	H
(0, 0, 0, 0, 0, 0)	C	.0309	.0491	.0500	.0483
	$t(\boldsymbol{n}, C(\Omega_{U_4}))$	−2.402	−8.460	−13.364	−7.207
	DC	.0384	.0450	.0500	.0498
	$t(\boldsymbol{n}, DC(\Omega_{U_4}))$	−2.193	−4.178	−6.949	−2.820
(1, 1, 1, 0, 2, 2)	C	.552	.746	.780	.661
	DC	.562	.691	.721	.668
(2, 0, 1, 0, 2, 2)	C	.747	.951	.960	.849
	DC	.798	.950	.955	.861
(2, 0, 2, 0, 2, 1)	C	.734	.882	.935	.419
	DC	.862	.984	.981	.543
(2, 1, 0, 0, 2, 0)	C	.689	.898	.923	.666
	DC	.646	.865	.863	.440
(2, 1, 1, 0, 1, 2)	C	.600	.940	.922	.946
	DC	.414	.881	.792	.942
(2, 1, 2, 0, 1, 1)	C	.573	.810	.837	.668
	DC	.617	.829	.826	.675
(2, 2, 0, 0, 1, 0)	C	.732	.909	.949	.701
	DC	.598	.551	.689	.341
(2, 2, 1, 0, 0, 2)	C	.759	.983	.980	.967
	DC	.657	.915	.902	.884
(0, 0, 0, 0, 1, 0)	C	.097	.069	.087	.042
	DC	.150	.106	.167	.034
(0, 0, 2, 1, 1, 0)	C	.081	.001	.008	.001
	DC	.137	.072	.126	.001
(0, 1, 0, 1, 1, 0)	C	.090	.005	.019	.001
	DC	.143	.013	.055	.001
(0, 1, 2, 1, 1, 2)	C	.175	.071	.189	.004
	DC	.235	.217	.313	.039
(0, 2, 0, 1, 1, 2)	C	.470	.296	.488	.043
	DC	.595	.410	.669	.032
(0, 2, 1, 1, 1, 1)	C	.137	.045	.106	.005
	DC	.150	.083	.130	.013
(0, 1, 1, 1, 0, 2)	C	.402	.043	.166	.009
	DC	.533	.095	.342	.017
(1, 0, 1, 1, 0, 2)	C	.452	.164	.341	.046
	DC	.599	.437	.660	.089
(1, 0, 2, 1, 0, 1)	C	.205	.063	.146	.005
	DC	.333	.512	.533	.045
(1, 1, 0, 1, 0, 0)	C	.141	.038	.104	.013
	DC	.145	.031	.079	.005
(1, 1, 2, 1, 0, 2)	C	.449	.166	.336	.039
	DC	.601	.441	.658	.076
(1, 2, 0, 1, 0, 2)	C	.643	.592	.775	.200
	DC	.780	.678	.893	.052
(1, 2, 1, 1, 0, 1)	C	.205	.152	.231	.067
	DC	.243	.191	.313	.037

Table 10. Motivational effects of knowledge of performance.

Control no information	Group B rough information	Group C accurate information
40	38	48
35	40	40
38	47	45
43	44	43
44	40	46
41	42	44

Carlo simulation to obtain an estimate of $t(\boldsymbol{n}, B(\Omega))$. In this case, it is more convenient to estimate the P-value instead of $t(\boldsymbol{n}, B(\Omega))$. To carry out the simulation, generate k samples of random numbers from uniform (0, 1) distribution with sample sizes $n_1, \ldots, n_k$, respectively. Compute the joint ranks $\boldsymbol{R}^*$ and the test statistic $T(\boldsymbol{R}^*, \boldsymbol{n}, B(\Omega))$. Repeat this process for M times. The proportion of $T(\boldsymbol{R}^*, \boldsymbol{n}, B(\Omega))$ less than and equal to the computed $T(\boldsymbol{R}, \boldsymbol{n}, B(\Omega))$ based on the data provides an estimate of the exact P-value.

Motivational effects example revisited. In this example, one wants to check the validity of the prior belief that increased motivational effects will not decrease productivity in this particular task. This question can be answered by testing $H_0 : \theta_1 \leq \theta_2 \leq \theta_3$ versus H_1: not H_0. Based on Section 3, the test $C^*_{B(\Omega)}(\boldsymbol{R})$ with $B(\Omega) = DC(\Omega) = \{(1,2),(2,3)\}$ is a reasonable choice. This test rejects the null hypothesis for smaller values of $C^*_{DC(\Omega)}(\boldsymbol{R}) = \sqrt{N} \min\{W_{12}(\boldsymbol{R}), W_{23}(\boldsymbol{R})\}$, where N is the total sample size. The P-value is given by $P\{\sqrt{N} \min\{W_{12}(\boldsymbol{R}^*), W_{23}(\boldsymbol{R}^*)\} \leq C^*_{DC(\Omega)}(\boldsymbol{R})\}$, where $\boldsymbol{R}^*$ is uniformly distributed over the set of permutations of $\{1, \ldots, 18\}$. The basic data in Table 10 consist of the numbers of pieces processed by each subject in the experimental period. The Mann-Whitney-Wilcoxon statistics for samples 1 and 2 and samples 2 and 3, are, respectively, $W'_{12}(\boldsymbol{R}) = 43$ and $W'_{23}(\boldsymbol{R}) = 47.5$. The standardized Mann-Whitney-Wilcoxon statistics are, respectively, $W_{12}(\boldsymbol{R}) = .151$ and $W_{23}(\boldsymbol{R}) = .320$. The test statistic is therefore $C^*_{C(\Omega)}(\boldsymbol{R}) = .641$. An estimate of the exact P-value via Monte Carlo simulation based on $M = 10000$ iterations is 0.98. Hence, at any reasonable level of significance, there is no statistical evidence against the belief that increased knowledge of performance will not negatively affect productivity when performing this industrial task.

5. Conclusion

This study provides a general nonparametric theory for testing ordered null hypotheses. The proposed theory not only has the desirable property of being distribution-free, but also accommodates various types of intuitive nonparametric test statistics proposed in this paper. The sufficient and necessary conditions for the consistency of the proposed tests are simple and easy to use.

While this paper concentrates on one-way location-shift models, the methodology is directly applicable to some other models involving k treatments by replacing the pairwise linear rank statistics for one-way models with their corresponding nonparametric counterparts under other models. For example, in two-way layouts with no interactions, all the theorems and proofs in this paper hold if the pairwise statistics $M_{st}(\boldsymbol{R})$ are replaced by their corresponding Wilcoxon signed rank statistics.

Acknowledgements

The author would like to thank the reviewers for their valuable suggestions that improved the paper.

Appendix

A.1 *Proof of Theorem* 2.1

Let $X_{ij}^* = X_{ij} - \theta_i$, $i = 1, \ldots, k$, $j = 1, \ldots, n_i$. Then, X_{ij}^*, $i = 1, \ldots, k$, $j = 1, \ldots, n_i$, are independently and identically distributed according to continuous distribution function $F(x)$. Let R_{ij}^* be the rank of X_{ij}^* among all X_{ij}^*, $i = 1, \ldots, k$, $j = 1, \ldots, n_i$. It follows from Theorem 2.1 of Randles and Wolfe (1991) that the rank vector $\boldsymbol{R}^* = (R_{11}^*, \ldots, R_{1n_1}^*, \ldots, R_{k1}^*, \ldots, R_{kn_k}^*)$ does not depend on $F(x)$ and is uniformly distributed over the set of all permutations of $\{1, \ldots, N\}$. By the definition of $T(\boldsymbol{R}, \boldsymbol{n}, B(\Omega))$, one can see that

$$\text{(A.1)} \qquad \begin{aligned} &P_{\boldsymbol{\theta}, F(x)}\{T(\boldsymbol{R}^*, \boldsymbol{n}, B(\Omega)) \geq t(\boldsymbol{n}, B(\Omega))\} \\ &= P\{T(\boldsymbol{R}^*, \boldsymbol{n}, B(\Omega)) \geq t(\boldsymbol{n}, B(\Omega))\} \geq 1 - \alpha. \end{aligned}$$

Let $S_j^{(st)}$ be the rank of X_{tj}^* among $X_{s1}^*, \ldots, X_{sn_s}^*$, $X_{t1}^*, \ldots, X_{tn_t}^*$. Then $M_{st}'(\boldsymbol{R}^*) = \sum_{j=1}^{n_t} a_{m_{st}}(S_j^{(st)})$. Notice that for any $(s,t) \in B(\Omega)$, $\theta_s \leq \theta_t$. This implies that, for any $(s,t) \in B(\Omega)$ and any $F(x)$, $S_j^{(st)} \leq R_j^{(st)}$ for $j = 1, \ldots, n_t$, which further implies that $M_{st}'(\boldsymbol{R}^*) \leq M_{st}'(\boldsymbol{R})$ and therefore $M_{st}(\boldsymbol{R}^*) \leq M_{st}(\boldsymbol{R})$. Hence, for any $\boldsymbol{\theta} \in H_0$ and $F(x)$,

$$\text{(A.2)} \qquad \begin{aligned} T(\boldsymbol{R}^*, \boldsymbol{n}, B(\Omega)) &= k(\boldsymbol{n}, B(\Omega)) l_{\boldsymbol{n}}(M_1(\boldsymbol{R}^*), \ldots, M_{|B|}(\boldsymbol{R}^*)) \\ &\leq k(\boldsymbol{n}, B(\Omega)) l_{\boldsymbol{n}}(M_1(\boldsymbol{R}), \ldots, M_{|B|}(\boldsymbol{R})) \\ &= T(\boldsymbol{R}, \boldsymbol{n}, B(\Omega)). \end{aligned}$$

It follows from (A.1) and (A.2) that, for any $\boldsymbol{\theta} \in H_0$ and $F(x)$,

$$\begin{aligned} &P_{\boldsymbol{\theta}, F(x)}\{T(\boldsymbol{R}, \boldsymbol{n}, B(\Omega)) < t(\boldsymbol{n}, B(\Omega))\} \\ &\leq P_{\boldsymbol{\theta}, F(x)}\{T(\boldsymbol{R}^*, \boldsymbol{n}, B(\Omega)) < t(\boldsymbol{n}, B(\Omega))\} \\ &= P\{T(\boldsymbol{R}^*, \boldsymbol{n}, B(\Omega)) < t(\boldsymbol{n}, B(\Omega))\} \triangleq c(l_{\boldsymbol{n}}, B) \leq \alpha. \end{aligned}$$

Theorem 2.1 is therefore established.

A.2 *Proof of Theorem* 2.2

Without loss of generality, assume that $k(\boldsymbol{n}, B(\Omega)) = 1$ since consistency and power performance do not depend on $k(\boldsymbol{n}, B(\Omega))$. For convenience, only the case when $l_{\boldsymbol{n}}(x_1, \ldots, x_{|B|}) = l(x_1, \ldots, x_{|B|})$ will be discussed in detail. The extension of the proof to a general case is straightforward using the continuity of $l_{\boldsymbol{n}}(x_1, \ldots, x_n)$ and $l(x_1, \ldots, x_{|B|})$. As mentioned in Section 2.2, test $T(\boldsymbol{R}, \boldsymbol{n}, B(\Omega))$ does not depend on $b_{m_{st}}$ or $d_{m_{st}}$. Therefore, it suffices to let $b_{m_{st}} = 1$ and $d_{m_{st}} = 0$, which leads to scores $a_{m_{st}}(i) = \phi(i/(m_{st}+1))$, $i = 1, \ldots, m_{st}$. According to (9.2.5) in Randles and Wolfe (1991),

$$\text{(A.3)} \qquad \lim_{N \to +\infty} \frac{\sigma_{0:m_{st}}}{\sqrt{m_{st}}} = \frac{\sqrt{\lambda_s \lambda_t}}{\lambda_s + \lambda_t} \int_0^1 (\phi(u) - \bar{\phi})^2 du$$

and

$$\text{(A.4)} \qquad \lim_{n\to+\infty} \frac{n_t \bar{a}_{m_{st}}}{m_{st}} = \lim_{n\to+\infty} \left[\frac{n_t}{m_{st}} \cdot \frac{1}{m_{st}} \sum_{i=1}^{m_{st}} \phi\left(\frac{i}{m_{st}+1}\right) \right] = \frac{\lambda_t}{\lambda_s+\lambda_t} \bar{\phi}.$$

It follows from (3.5.12) in Hettmansperger (1991) that

$$\text{(A.5)} \qquad \frac{1}{m_{st}} \sum_{j=1}^{n_t} \phi\left(\frac{R_j^{(st)}}{m_{st}+1}\right) \xrightarrow{P} \frac{\lambda_t}{\lambda_s+\lambda_t} \int_{-\infty}^{+\infty} \phi\left[\frac{\lambda_s}{\lambda_s+\lambda_t} F(y-\theta_t-\theta_s) + \frac{\lambda_t}{\lambda_s+\lambda_t} F(y)\right] dF(y) \quad \text{as} \quad N \to +\infty.$$

It follows from (A.3)–(A.5) that

$$\text{(A.6)} \qquad M_{st}(\boldsymbol{R}) = \sqrt{\frac{m_{st}}{N}} \cdot \frac{\frac{1}{m_{st}} \sum_{j=1}^{n_t} \phi\left(\frac{R_j^{(st)}}{m_{st}+1}\right) - \frac{n_t \bar{a}_{m_{st}}}{m_{st}}}{\sigma_{0:m_{st}}/\sqrt{m_{st}}} \xrightarrow{P} a_{st} \cdot b(\theta_s, \theta_t, F),$$

where $a_{st} = \sqrt{\lambda_t(1+\lambda_t/\lambda_s)}[\int_0^1 (\phi(u)-\bar{\phi})^2 du]^{-1}$ and

$$\begin{aligned} \text{(A.7)} \qquad b(\theta_s,\theta_t,F) &= \int_{-\infty}^{\infty} \phi\left[\frac{\lambda_s}{\lambda_s+\lambda_t} F(y+\theta_t-\theta_s) + \frac{\lambda_t}{\lambda_s+\lambda_t} F(y)\right] dF(y) - \bar{\phi} \\ &= \int_{-\infty}^{\infty} \left\{ \phi\left[\frac{\lambda_s}{\lambda_s+\lambda_t} F(y+\theta_t-\theta_s) + \frac{\lambda_t}{\lambda_s+\lambda_t} F(y)\right] \right. \\ &\qquad \left. - \phi\left[\frac{\lambda_s}{\lambda_s+\lambda_t} F(y) + \frac{\lambda_t}{\lambda_s+\lambda_t} F(y)\right] \right\} dF(y). \end{aligned}$$

Note that since $F(x)$ is monotone and $\phi(u)$ strictly increasing,

$$\text{(A.8)} \qquad b(\theta_s,\theta_t,F) \begin{matrix} > 0 \\ = 0 \\ < 0 \end{matrix} \quad \text{when} \quad \begin{matrix} \theta_s < \theta_t \\ \theta_s = \theta_t \\ \theta_s > \theta_t. \end{matrix}$$

By (A.6) and the continuity of $l(x_1, \ldots, x_{|B|})$,

$$\text{(A.9)} \qquad T(\boldsymbol{R}, \boldsymbol{n}, B(\Omega)) \xrightarrow{P} q(\boldsymbol{\theta}, F) = l(a_1 b_1, \ldots, a_{|B|} b_{|B|}), \qquad \text{as} \quad N \to +\infty,$$

where $a_1 b_1, \ldots, a_{|B|} b_{|B|}$ are $a_{st} b(\theta_s, \theta_t, F)$, $(s,t) \in B(\Omega)$, arranged in the same lexicogical order as that of $M_i(\boldsymbol{R})$, $i = 1, \ldots, |B|$. It follows from (A.8) and (A.9) that, when $\theta_1 = \cdots = \theta_k$,

$$\text{(A.10)} \qquad T(\boldsymbol{R}, \boldsymbol{n}, B(\Omega)) \xrightarrow{P} l(0, \ldots, 0), \qquad \text{as} \quad N \to +\infty.$$

Since the cutoff value $t(\boldsymbol{n}, B(\Omega))$ is the α-th percentile of $T(\boldsymbol{R}, \boldsymbol{n}, B(\Omega))$ under $\theta_1 = \cdots = \theta_k$, it is easy to show by contradiction that

$$\text{(A.11)} \qquad \lim_{N\to+\infty} t(\boldsymbol{n}, B(\Omega)) = l(0, \ldots, 0).$$

For any fixed $\boldsymbol{\theta} \in H_1$; that is $\boldsymbol{\theta} \notin \Omega$, there exist some $(s,t) \in C(\Omega)$ such that $\theta_s > \theta_t$. According to (A.8), the $M_{st}(\boldsymbol{R})$'s corresponding to these (s,t)'s converge in probability to negative values $a_{st}b(\theta_s, \theta_t, F)$'s. It follows from condition (2.3) that $q(\boldsymbol{\theta}, F) < l(0,\ldots,0)$ for $\boldsymbol{\theta} \in H_1$ and any fixed $F(x)$.

According to (A.8), there exists an $N_0 > 0$ such that, when $N > N_0$,

$$t(\boldsymbol{n}, B(\Omega)) > \Delta(\boldsymbol{\theta}, F) = \frac{1}{2}(q(\boldsymbol{\theta}, F) + l(0,\ldots,0)).$$

According to (A.9) and the fact that $q(\boldsymbol{\theta}, F) < l(0,\ldots,0)$, for any $\epsilon > 0$ there exists an $N_1 > 0$ such that, when $N > N_1$, $P_{\boldsymbol{\theta},F(X)}\{T(\boldsymbol{R},\boldsymbol{n}, B(\Omega)) < \Delta(\boldsymbol{\theta}, F)\} \geq 1-\epsilon$. Therefore, for any $\epsilon > 0$, when $N > \max\{N_0, N_1\}$, $P_{\boldsymbol{\theta},F(x)}\{T(\boldsymbol{R},\boldsymbol{n}, B(\Omega)) < t(\boldsymbol{n}, B(\Omega))\} \geq P_{\boldsymbol{\theta},F(x)}\{T(\boldsymbol{R},\boldsymbol{n}, B(\Omega)) < \Delta(\boldsymbol{\theta}, F)\} \geq 1-\epsilon$. Sufficiency of condition (2.3) for consistency is hence proved.

A.3 *Proof of Theorem* 2.3

(Proof by contradiction) Let Φ be the cumulative distribution function of standard normal distribution. Define

$$A = \min_{(s,t)\in B(\Omega)} \left\{ \lim_{\theta_t - \theta_s \to -\infty} |b(\theta_s, \theta_t, \Phi)|, \lim_{\theta_t - \theta_s \to +\infty} b(\theta_s, \theta_t, \Phi) \right\}.$$

It is easy to see from (A.7) that $A > 0$.

Suppose that the conclusion in Theorem 2.3 is not true. Then, for any $\delta > 0$, condition (2.4) does not hold. As a special case, condition (2.4) does not hold for $\delta^* = \frac{A}{|B|^2} > 0$. Therefore, without loss of generality, there exist $x_1^0 < 0, \ldots, x_p^0 < 0$, $x_{p+1}^0, \ldots, x_{|B|}^0$, such that $|x_i^0| < \delta^*$, for $i = 1, \ldots, |B|$, but

$$l(x_1^0, \ldots, x_p^0, x_{p+1}^0, \ldots, x_{|B|}^0) > l(0,\ldots,0). \tag{A.12}$$

Since $l(x_1, \ldots, x_{|B|})$ is increasing componentwise, $l(x_1^0, m^0, \ldots, m^0) > l(0,\ldots,0)$ where $m^0 = \max\{x_2^0, \ldots, x_{|B|}^0\} < \delta^*$.

Suppose that in the arrangement convention used throughout this paper the first argument x_1 in the notation $l(x_1, \ldots, x_{|B|})$ corresponds to the pair $(s_1, t_1) \in B(\Omega)$. Now construct the following $\boldsymbol{\theta}^*$ for which test $T(\boldsymbol{R},\boldsymbol{n}, B(\Omega))$ is not consistent. Let $\theta_{s_1}^* = 0$. Choose $\theta_{t_1}^* < 0$ so that $a_{s_1t_1}b(\theta_{s_1}^*, \theta_{t_1}^*, \Phi) = x_1^0$, where Φ is the cumulative distribution function of standard normal distribution. This is always possible since $a_{s_1t_1}b(\theta_{s_1}^*, \theta_{t_1}^*, \Phi)$ is a continuous function of $\theta_{t_1}^*$; and $b(\theta_{s_1}^*, \theta_{s_1}^*, \Phi) = 0$ and $\lim_{\theta_{t_1}^* - \theta_{s_1}^* \to -\infty} |b(\theta_{s_1}^*, \theta_{t_1}^*, \Phi)| > \delta^*$. By using the similar argument, it is possible to choose the remaining $\theta_i^*, i \neq s_1, t_1$, such that $\theta_s^* < \theta_t^*$ and $a_{st}b(\theta_s^*, \theta_t^*, \Phi) \geq m^0$ for all $(s,t) \in B(\Omega) - \{(s_1,t_1)\}$. Notice that $\boldsymbol{\theta}^* \in H_1$ since $\theta_{s_1}^* > \theta_{t_1}^*$ and $(s_1,t_1) \in B(\Omega)$. However, according to (A.9), under $\boldsymbol{\theta}^*$ and $F(x) = \Phi(x)$, $T(\boldsymbol{R},\boldsymbol{n}, B(\Omega)) \xrightarrow{P} q(\boldsymbol{\theta}^*, \Phi) = l(a_1b_1, \ldots, a_{|B|}b_{|B|}) \geq l(x_1^0, m^0, \ldots, m^0) > l(0,\ldots,0)$ as $N \to +\infty$. Therefore, $\lim_{N\to+\infty} P_{\boldsymbol{\theta}^*,\Phi}\{T(\boldsymbol{R},\boldsymbol{n}, B(\Omega)) < t(\boldsymbol{n}, B(\Omega))\} = 0$, which means that test $T(\boldsymbol{R},\boldsymbol{n}, B(\Omega))$ is inconsistent under $\boldsymbol{\theta}^*$ and $F(x) = \Phi(x)$. By contradiction, the necessity of condition (2.4) for the consistency of test $T(\boldsymbol{R},\boldsymbol{n}, B(\Omega))$ is established.

References

Barlow, R. E., Bartholomew, D. J., Bremner, J. M. and Brunk, H. D. (1972). *Statistical Inference Under Order Restrictions*, Wiley, New York.

D'Agostino, R. B. and Stephens, M. A. (1986). *Goodness-of-Fit Techniques*, Marcel Dekker, Inc., New York.

Dykstra, R. L. and Robertson, T. (1982). Order Restricted Statistical Tests on Multinomial and Possion Parameters: The Starshaped Restriction, *Ann. Statist.*, **10**, 1246–1252.

Eeden, C. V. (1958). Testing and estimating ordered parameters of probability distributions, Doctoral Dissertation, University of Amsterdam, Studentendrukkerij Poortpers, Amsterdam.

Hettmansperger, T. P. (1991). *Statistical Inference Based on Ranks*, Krieger Publishing Company, Malabar, Florida.

Hollander, M. and Wolfe, D. A. (1973). *Nonparametric Statistical Methods*, Wiley, New York.

Hundal, P. S. (1969). Knowledge of performance as an incentive in repetitive industrial work, *Journal of Applied Psychology*, **53**, 224–226.

Lee, C. I. C. (1987). Chi-square tests for and against an order restriction on multinomial parameters, *J. Amer. Statist. Assoc.*, **82**, 611–618.

Lee, R. E. and Spurrier, J. D. (1995*a*). Distribution-free multiple comparisons between successive treatments, *J. Nonparametr. Statist.*, **5**, 261–273.

Lee, R. E. and Spurrier, J. D. (1995*b*). successive comparisons between ordered treatments, *J. Statist. Plan. Inference*, **43**, 323–330.

Li, T. and Sinha, B. K. (1995). Tests of ordered hypotheses for gamma scale parameters, *J. Statist. Plan. Inference*, **43**, 311–321.

Mukerjee, H., Robertson, T. and Wright, F. T. (1986). Multiple contrast tests for testing against a simple tree ordering, *Advances in Order Restricted Statistical Inference* (eds. R. L. Dykstra, T. Robertson and F. T. Wright), 203–230, Springer, New York.

Mukerjee, H., Robertson, T. and Wright, F. T. (1987). Comparison of several treatments with a control using multiple contrasts, *J. Amer. Statist. Assoc.*, **82**, 902–910.

Parsons,V. L. (1979). A nonparametric test for trend, Tech. Report, Department of Mathematics, University of Cincinnati.

Perlman, M. D. (1969). One-sided problems in multivariate analysis, *Ann. Math. Statist.*, **40**, 549–567.

Puri, M. L. and Sen, P. K. (1971). *Nonparametric Methods in Multivariate Analysis*, Wiley, New York.

Randles, R. H. and Wolfe, D. A. (1991). *Introduction to the theory of nonparametric statistics*, Krieger Publishing Company, Malabar, Florida.

Robertson, T. (1978). Testing for and against an order restriction on multinomial parameters, *J. Amer. Statist. Assoc.*, **73**, 197–202.

Robertson, T. and Wegman, E. J. (1978). Likelihood ratio tests for ordered restrictions in exponential families, *Ann. Statist.*, **6**, 485–505.

Robertson, T., Wright, F. T. and Dykstra, R. L. (1988). *Order Restricted Statistical Inference*, Wiley, New York.

Shi, N. Z. (1988). Testing the null hypothesis that a normal mean vector lies in the positive orthant, *Memoirs of the Faculty of Science. Kyshu University. series A.*, **42**, 109–122.

Shiraishi, T. (1982). Testing homogeneity against trend based on rank in one-way layout, *Comm. Statist. Theory Methods*, **11**, 1255–1268.

Singh, B. and Wright, F. T. (1990). Testing for and against an order restriction in mixed-effects models, *Statist. Probab. Lett.*, **9**, 195–200.

Tryon, P. V. and Hettmansperger, T. P. (1973). A class of nonparametric tests for homogeneity against ordered alternative, *Ann. Statist.*, **1**, 1061–1070.

Wollan, P. C., and Dykstra, R. L. (1986). Conditional Tests With an Order Restriction As a Null Hypothesis (eds. R. L. Dykstra, T. Robertson and F. T. Wright), 279–295, *Advances in Order Restricted Statistical Inference*, Springer, New York.