A SUFFICIENT CONDITION FOR

ADMISSIBILITY OF THE WILCOXON TEST

IN THE CLASSICAL TWO-SAMPLE PROBLEM

Qiqing Yu

Department of Mathematics Sciences
State University of New York at Binghamton, NY 13902

Current version: 12/6/2001
Short title: Classical Two-Sample Problem.
AMS 1991 subject classification: Primary 62 C15; Secondary 62 G10.
Key words and phrases: continuous distribution functions, nonparametric hypothesis tests, linear rank tests.
Abstract. We consider nonparametric two-sample problems of testing the equality of two continuous distribution functions F and G. Whether or not the Wilcoxon test is admissible within the class of all tests and whether or not the two-sided Wilcoxon test is admissible within the class of all rank tests are two longstanding open questions (Lehmann (1959 and 1986) and Ferguson (1967)). In this paper, we establish a sufficient condition that the Wilcoxon test is admissible in these two-sample problems. As an application, we show that for some special cases, the Wilcoxon test is admissible within the class of all tests and the two-sided Wilcoxon test is admissible within the class of all rank tests. The author believes that the sufficient condition can be used to solve the longstanding open questions in Lehmann (1959) on a case-by-case basis, but is unable to produce a unified proof for all cases.

1. Introduction. One of the basic problems of statistics is the two-sample problem of testing the equality of two distributions. A typical example is the comparison of a treatment with a control, where the hypothesis of no treatment effect is tested against the alternative of a beneficial effect. When normality of samples distributions is in doubt, people may make no specific assumption on distribution functions other than continuity in the two-sample testing problem.

Let X_1, ..., X_m be a random sample from a continuous distribution function F and Y_1, ..., Y_n be another random sample from a continuous distribution function G. Assume X_is and Y_js are independent. One wishes to test H_0: $F = G$ ($F(x) = G(x)$ for all x) against the alternative H_1: $F \leq G$ ($F(x) \leq G(x)$ for all x) yet $F(x) < G(x)$ for some x. Another alternative is H_2: $F \geq G$ yet $F(x) > G(x)$ for some x. The testing problem is invariant under the group of all continuous and strictly monotone transformations. For this problem, the terms "invariant test" and "rank test" are synonymous. It is known that there is no uniformly most powerful rank test of H_0 against H_1. The Wilcoxon rank sum test and the Fisher-Yates test are two most commonly applied procedures for this two-sample problem (see Ferguson (1967)). Both tests are locally best and thus are admissible within the class of all rank tests (see Ferguson (1967)). However, the admissibility of these two tests in the class of all tests has been a difficult and unsolved problem (Lehmann (1959, p. 240, and 1986, p. 322), and Ferguson (1967)).

One two-sided alternative hypothesis is H_3: either H_1 or H_2 is true. The testing problem is still invariant under the group of continuous and strictly monotone transformations. However, "The theory here is still less satisfactory than in the one-sided case" (Lehmann (1959, p. 240)). The Wilcoxon test and the Fisher-Yates test need not even be unbiased (Sugiura (1965)), and "it is not known whether they are admissible within the class of all rank tests" (Lehmann (1959, p. 240) and (1986, p. 322)).

Another two-sided alternative is H_4: $F(x) \neq G(x)$ for some x. Note that the only invariant test in this testing problem is a constant, *i.e.*, $\psi_\alpha(\cdot) \equiv \alpha$. However, ψ_α is inadmissible at every significant level α ($\in (0,1)$) (Lehmann (1986)).

In this note, we establish a sufficient condition that a member from a class of linear rank tests, including the Wilcoxon test, the Fisher-Yates test, the Savage test and the median test, is admissible for testing H_0 against H_i under the continuous set-up within the class of all tests or within the class of all rank tests, where $i = 1, 2, 3, 4$. In particular, we apply the result to show that for some special cases, the Wilcoxon test and the Fisher-Yates test are admissible. The results partially answer the two longstanding open questions in Lehmann (1959) for the first time.

In Section 2, we present notations. In Section 3, we introduce a continuous parametric two-sample problem for which the admissibility problem is much easier to attack. In Section 4, we shall relate this problem to the classical two-sample problem and establish a sufficient condition for admissibility in the latter problem. In Section 5, we consider the application of the result in Section 4. In Section 6, we make some comments.

2. Notations. Let Θ be a class of distribution functions under consideration. Denote by $\mathcal{F}$ the class of all continuous distribution functions. Let $\mathbf{X} = (X_1, ..., X_m)'$ and $\mathbf{Y} = (Y_1, ..., Y_n)'$ be two independent random samples from two populations with distribution functions F and $G \in \Theta$, respectively, where $\mathbf{X}'$ is the transpose of the vector $\mathbf{X}$. Let $X_{(1)} < \cdots < X_{(m)}$ and $Y_{(1)} < \cdots < Y_{(n)}$ be order statistics of $\mathbf{X}$ and $\mathbf{Y}$, respectively, and let $N = m + n$. Denote the pooled sample $\mathbf{Z}' = (Z_1, ..., Z_N) = (\mathbf{X}', \mathbf{Y}')$ and denote $\mathcal{R}(Z_j)$ the rank of Z_j in the pooled sample, and denote $R_i = \mathcal{R}(X_{(i)})$, $R_{m+j} = \mathcal{R}(Y_{(j)})$ and $\mathbf{R} = (R_1, ..., R_N)$. It is obvious that the joint distribution function of $\mathbf{Z}$ is $F_{\mathbf{Z}}(\mathbf{z}) = \prod_{i=1}^{m} F(z_i) \prod_{j=n+1}^{n+m} G(z_j)$.

Recall that a test ψ satisfies $\psi(\cdot) \in [0, 1]$. Let $E_{F,G}(\psi(\mathbf{Z}))$ (or simply $E_{F,G}\psi$) denote the expectation of $\psi(\mathbf{Z})$ with respect to F and G. For testing against H_i, where $i = 1, ..., 4$, a test ψ is said to be admissible if there is no test ϕ_0 such that $E_{F,F}\phi_0 \leq E_{F,F}\psi$ and $E_{F,G}\phi_0 \geq E_{F,G}\psi$ for all F, $G \in \Theta$, where F and G satisfy either H_0 or H_i, and at least one strict inequality holds. A test ϕ_0 is said to be as good as ϕ, if $E_{F,F}\phi_0 \leq E_{F,F}\phi$ and $E_{F,G}\phi_0 \geq E_{F,G}\phi$ for each pair of F and G where F and G satisfy either H_0 or H_i and $F, G \in \Theta$. Given a subset A of the real line, denote A^N the product set $A \times \cdots \times A$ of N factors.

The Wilcoxon test is of the following form:

$$\phi(\mathbf{Z}) = 1_{[L(\mathbf{Z}) \notin [l,r]]} + \gamma(\mathbf{Z}) 1_{[L(\mathbf{Z}) = l \text{ or } r]}, \tag{2.1}$$

where $l < r$, $1_{[\cdot]}$ is an indicator function, $\gamma(\cdot) \in [0, 1)$ on $\{\mathbf{z} : L(\mathbf{z}) = l \text{ or } r\}$,

$$L(\mathbf{Z}) = \sum_{i=1}^{m} c_1 \mathcal{S}(R_i) + \sum_{j=1}^{n} c_2 \mathcal{S}(R_{m+j}),$$

c_1 and c_2 are distinct constants, $c_1 < c_2$, and $\mathcal{S}(\cdot)$ is a real-valued strictly increasing function. A common treatment is to set $\gamma(\cdot) \equiv c$, a constant. By properly defining $\gamma(\cdot)$, we can obtain a test with a desirable size α. When $\mathcal{S}(r) = r$ for all r and $(c_1, c_2) = (0, 1)$, ϕ is the Wilcoxon test. When $\mathcal{S}(r)$ is the expected value of the rth order statistic of a sample of size N from a standard normal distribution and $(c_1, c_2) = (0, 1)$, ϕ is the Fisher-Yates test. (2.1) also includes the Savage test and the median test. The test given in (2.1) actually includes both one-sided and two-sided tests by allowing $l = -\infty$ or $r = \infty$. For example, if $l = -\infty$ and $\gamma = 0$, (2.1) becomes $\phi(\mathbf{Z}) = 1_{[L(\mathbf{Z}) > r]}$.

3. A Continuous Parametric Two-sample Problem. In order to attack the more difficult classical problem, we first formulate an appropriate problem for testing two continuous distribution functions from the regular exponential family.

Let ξ be a $N \times 1$ vector with coordinates $\xi_1 < \cdots < \xi_N$. Let F_q and G_p be two continuous distribution functions with density functions

$$f_q = f_{q,\xi} = 2k \sum_{i=1}^{N} q_i 1_{[x \in (\xi_i - 1/k, \xi_i + 1/k)]},$$

$$g_p = g_{p,\xi} = 2k \sum_{i=1}^{N} p_i 1_{[x \in (\xi_i - 1/k, \xi_i + 1/k)]},$$

where $k > 2/\min\{|\xi_i - \xi_j|, i \neq j\}$ and $\mathbf{q} = (q_1, ..., q_N)$ and $\mathbf{p} = (p_1, ..., p_N)$ are probability vectors. Let $\mathcal{F}_k(\xi)$ be the collection of distribution functions with the above forms and with fixed k and ξ. Note that $\mathcal{F}_k(\xi)$ belongs to the regular exponential family. Let $\mathbf{t}$, $\mathbf{s}$ and $\mathbf{u}$ be $(N-1) \times 1$ vectors, with coordinates

$$t_i = t_i(\mathbf{Z}) = \sum_{j=1}^{n} 1_{[Y_j \in [\xi_i - 1/k, \xi_i + 1/k]]}, \quad s_i = s_i(\mathbf{Z}) = \sum_{j=1}^{m} 1_{[X_j \in [\xi_i - 1/k, \xi_i + 1/k]]}$$

and $u_i = t_i + s_i$, respectively. The t_i's and s_j's are jointly complete and sufficient for F_q and G_p (or $(\mathbf{q}, \mathbf{p})$) and the joint probability density function f of $\mathbf{Z}$ is proportional to $\prod_{i=1}^N p_i^{t_i} \prod_{j=1}^N q_j^{s_j}$, i.e.,

$$f(\mathbf{z}) \propto \binom{N}{\mathbf{t}}\binom{N}{\mathbf{u}-\mathbf{t}} 1_{[\mathbf{z}\in\Omega]} e^{\mathbf{t}(\mathbf{z})\cdot\mathbf{w}+\mathbf{u}(\mathbf{z})\cdot\theta} \quad = h(\mathbf{u},\mathbf{t}) e^{\mathbf{t}(\mathbf{z})\cdot\mathbf{w}+\mathbf{u}(\mathbf{z})\cdot\theta},$$

where $\mathbf{t}\cdot\mathbf{s}$ is the inner product of the vectors $\mathbf{t}$ and $\mathbf{s}$, $\binom{N}{\mathbf{t}} = N!/\prod_{i=1}^N t_i!$, $t_N = n - \sum_{i=1}^{N-1} t_i$, $s_N = m - \sum_{i=1}^{N-1} s_i$, $h(\mathbf{u},\mathbf{t}) = \binom{N}{\mathbf{t}}\binom{N}{\mathbf{u}-\mathbf{t}} 1_{[\mathbf{z}\in\Omega]}$, $\Omega = (\cup_{i=1}^N [\xi_i - 1/k, \xi_i + 1/k])^N$, θ and $\mathbf{w}$ are all $N-1$ dimensional vectors with coordinates

$$\theta_j = \log(q_j/q_N) \text{ and } w_j = \log((p_j/q_j)/(p_N/q_N)),$$

respectively. It is ready to see that H_0 is the same as $\mathbf{w} = \mathbf{0}$ (*i.e.*, $w_i = 0 \; \forall \; i$), H_4 is corresponding to $\mathbf{w} \neq \mathbf{0}$, and $\mathbf{w} \geq \mathbf{0}$ (*i.e.* $w_j \geq 0 \; \forall \; j$) implies H_1.

Given a test ψ which is a function of $(\mathbf{u}, \mathbf{t})$, let $C_{\mathbf{u},\psi} = \{\mathbf{t} : \; \psi(\mathbf{u},\mathbf{t}) < 1\}$. Let $\mu_{F_q,G_p}(\mathbf{u},\cdot)$ denote the measure induced by the conditional distribution function of $\mathbf{t}$ given $\mathbf{u}$. We say $C_{\mathbf{u},\phi_1}$ is convex a.s.$\mu_{F_p,G_q}(u,\cdot)$, in the sense that if $\mathbf{t}_i \in C_{\mathbf{u},\phi_1}$, where $\mu_{F_p,G_q}(\mathbf{u},\mathbf{t}_i) > 0$, $i = 1, 2$, and if $\mathbf{t}_3 = a\mathbf{t}_1 + (1-a)\mathbf{t}_2$, where $a \in (0,1)$ and $\mu_{F_p,G_q}(\mathbf{u},\mathbf{t}_3) > 0$, then $\mathbf{t}_3 \in C_{\mathbf{u},\phi_1}$. It can be shown that ϕ given in (2.1) is not a function of the complete and sufficient statistic $(\mathbf{u},\mathbf{t})$. Thus we define a new test

$$\phi_1 = E_{F_q,G_p}(\phi(\mathbf{Z})|\mathbf{u},\mathbf{t}). \tag{3.1}$$

Verify that ϕ_1 is a function of $(\mathbf{u},\mathbf{t})$ and $\phi_1 \neq \phi$, but

$$E_{F_q,G_p}\phi_1 = E_{F_q,G_p}\phi \quad \forall \; F_q, G_p \in \mathcal{F}_k(\xi). \tag{3.2}$$

By ϕ_1, we convert the continuous nonparametric two-sample problems to a continuous parametric two-sample problem.

Lemma 3.1. *When $\Theta = \mathcal{F}_k(\xi)$, for $i = 1, 2, 3, 4$, the test $\phi_1 = \phi_1(\mathbf{u},\mathbf{t})$ for testing $\mathbf{w} = \mathbf{0}$ against H_i is admissible, if for all $\mathbf{u}$, (1) $C_{\mathbf{u},\phi_1}$ is convex a.s.$\mu_{F_p,G_q}(\mathbf{u},\cdot)$, (2) $\phi_1(\mathbf{u},\mathbf{t}_0) > 0$ implies that $\exists$ a vector $\mathbf{b}$ such that $\mathbf{b}\cdot(\mathbf{t}-\mathbf{t}_0) < 0 \; \forall \; \mathbf{t} \in C_{\mathbf{u},\phi_1}$ and $\mathbf{t} \neq \mathbf{t}_0$, where $\mathbf{b} \geq \mathbf{0}$ for $i = 1$, $\mathbf{b} \leq \mathbf{0}$ for $i = 2$, $\mathbf{b} \geq \mathbf{0}$ or $\mathbf{b} \leq \mathbf{0}$ for $i = 3$, and $\mathbf{b}$ is arbitrary for $i = 4$.*

The proof of the lemma is a minor modification of Theorem 4.1 in Yu (2000). For the convenience of readers, we include the proof in the appendix. The following lemma points out that one only needs to verify condition (2) of Lemma 3.1 for ϕ_1.

Lemma 3.2. *Consider the problem of testing H_0 against H_i with $\Theta = \mathcal{F}_k(\xi)$, $i = 1$ 2, 3, 4. Let ϕ and ϕ_1 be given by (2.1) and (3.1), respectively. Then condition (1) of Lemma 3.1 holds for ϕ_1.*

Proof. We only give the proof for the test against H_1 as the proofs for the rest of the cases are similar. We shall give the proof in 4 steps.

Step 1 (Notations). The coordinates of the random vector $\mathbf{Z}$, Z_i's, are distinct a.s.. Let $Z_{(1)} \leq \cdots \leq Z_{(N)}$ be order statistics of Z_1, ..., Z_N, and let

$$t_i^o = \textstyle\sum_{j=1}^n 1_{[Y_j = Z_{(i)}]} \text{ and } s_i^o = \sum_{j=1}^m 1_{[X_j = Z_{(i)}]}, \; i = 1, \ldots, N. \tag{3.3}$$

Then $t_i^o + s_i^o = 1$ for all i a.s.. Let $\mathbf{t}^o = (t_1^o, \ldots, t_N^o)'$. Verify that $\mathbf{t}$ and $\mathbf{t}^o$ satisfy

$$t_j = \sum_{i>\sigma_{j-1}}^{\sigma_j} t_i^o, \; j = 1, \ldots, N-1, \text{ where } \sigma_j = \textstyle\sum_{k\leq j} u_k, \; u_0 = 0 \text{ and } \sigma_0 = 0. \tag{3.4}$$

Step 2 (Linearity of $\mathbf{t}$ as a function in $\mathbf{t}^o$). Given $\mathbf{u}$, let $T^{\mathbf{u}}$ be the set of all the possible values of $\mathbf{t}$, and for each $\mathbf{t} \in T^{\mathbf{u}}$, let $T_{\mathbf{t}} = \{\mathbf{t}^o : \; \mathbf{t}^o \text{ satisfies (3.4)}\}$. Eq. (3.4) can be viewed as a linear map from $\cup_{\mathbf{t}\in T^{\mathbf{u}}} T_{\mathbf{t}}$ to $T^{\mathbf{u}}$, say

$$\mathbf{t} = A_{\mathbf{u}}\mathbf{t}^o \text{ for each } \mathbf{t}^o \in T_{\mathbf{t}} \text{ and for each } \mathbf{t} \in T^{\mathbf{u}}. \tag{3.5}$$

The entries of the $(N-1) \times N$ matrix $A_{\mathbf{u}}$ can easily be identified by (3.4). Verify that $A_{\mathbf{u}}$ does not depend on $\mathbf{t}$ for each fixed $\mathbf{u}$.

Step 3 (Linearity of L in $\mathbf{t}^o$). Verify that $\sum_{i=1}^N t_i^o = n$, $\sum_{j=1}^N s_j^o = m$ and for each $\mathbf{Z}$ whose coordinates are all distinct,

$$\begin{aligned} L(\mathbf{Z}) =& \sum_{i=1}^m c_1 \mathcal{S}(R_i) \sum_{h=1}^N 1_{[X_{(i)}=Z_{(h)}]} + \sum_{j=1}^n c_2 \mathcal{S}(R_{m+j}) \sum_{h=1}^N 1_{[Y_{(j)}=Z_{(h)}]} \\ =& \sum_{h=1}^N \mathcal{S}(h)(c_2 - c_1) t_h^o + c_1 \sum_{h=1}^N \mathcal{S}(h) \\ =& \mathbf{a} \cdot \mathbf{t}^o + c, \end{aligned} \tag{3.6}$$

where $\mathbf{a} = (a_1, ..., a_N)$ and c are defined in an obvious way. Verify that $\mathbf{a}$ and c are not functions of $\mathbf{u}$ and (3.6) holds w.p.1 as $\mathbf{Z}$ has distinct coordinates w.p.1. Furthermore, $a_1 < \cdots < a_N$ as $\mathcal{S}$ is strictly increasing and $c_1 < c_2$. Note that ϕ is a function of $\mathbf{t}^o$ a.s., thus abusing notations, we write

$$\phi = \phi(\mathbf{Z}) = \phi(\mathbf{u}, \mathbf{t}) = \phi(\mathbf{t}^o).$$

Step 4 (Conclusion). Fix a $\mathbf{Z}$ with distinct coordinates and fix $(\mathbf{u}, \mathbf{t}) = (\mathbf{u}(\mathbf{Z}), \mathbf{t}(\mathbf{Z}))$. Let $|T_\mathbf{t}|$ be the cardinality of the set $T_\mathbf{t}$. Then (3.1) yields

$$\phi_1(\mathbf{u}, \mathbf{t}) = \sum_{\mathbf{t}^o \in T_\mathbf{t}} \frac{1}{|T_\mathbf{t}|} \phi(\mathbf{t}^o). \tag{3.7}$$

In view of (3.6), we can write

$$\phi = 1_{[\mathbf{a}\cdot\mathbf{t}^o + c \notin [l,r]]} + \gamma(\mathbf{t}^o) 1_{[\mathbf{a}\cdot\mathbf{t}^o + c = l \text{ or } r]}. \tag{3.8}$$

We say that $\{\mathbf{t}^o : \phi(\mathbf{t}^o) < 1\}$ is convex in $\mathbf{t}^o$, if $\mathbf{t}_3 = a\mathbf{t}_1 + (1-a)\mathbf{t}_2$, where $a \in (0,1)$, $\mathbf{t}_i \in \{\mathbf{t}^o : \phi(\mathbf{t}^o) < 1\}$, $i = 1$ and 2, and $\mathbf{t}_3 \in \{\mathbf{t}^o : \phi(\mathbf{t}^o) \le 1\}$, then $\mathbf{t}_3 \in \{\mathbf{t}^o : \phi(\mathbf{t}^o) < 1\}$. It is trivially true that $\{\mathbf{t}^o : \phi(\mathbf{t}^o) < 1\}$ is convex in $\mathbf{t}^o$, as there do not exist $\mathbf{t}_1 \neq \mathbf{t}_2$ such that $\mathbf{t}_3 = a\mathbf{t}_1 + (1-a)\mathbf{t}_2$ with $a \in (0,1)$ and $\mathbf{t}_i \in \{\mathbf{t}^o : \phi(\mathbf{t}^o) \le 1\}$. The last statements is due to the fact that the coordinates of $\mathbf{t}_i$ are either 0 or 1. In view of (3.7)

$$C_{\mathbf{u},\phi_1} = \{\mathbf{t} \in T^\mathbf{u} : \ \mathbf{t} = A_\mathbf{u} \mathbf{t}_0^o, \ \mathbf{t}_0^o \in \{\mathbf{t}^o : \ \phi(\mathbf{t}^o) < 1\}\},$$

and thus $C_{\mathbf{u},\phi_1}$ is also convex in $\mathbf{t}$. ▫

4. The main result. We present a sufficient condition for determining whether a linear rank test is admissible in the classical two-sample problem.

Theorem 4.1. *Consider the problem of testing H_0 against H_i (i =1, 2, 3 or 4). Let ϕ be a test of form (2.1). Then ϕ is admissible within the class of all tests when $\Theta = \mathcal{F}$, if for every k and ξ condition (2) of Lemma 3.1 holds for ϕ_1 defined in (3.1) when $\Theta = \mathcal{F}_k(\xi)$.*

Proof. Let i be an arbitrary integer among 1, 2, 3, and 4. Suppose ϕ_1 satisfies condition (2) of Lemma 3.1. Then by Lemmas 3.1 and 3.2, ϕ_1 is admissible for testing H_0 against H_i when $\Theta = \mathcal{F}_k(\xi)$.

Given a measure ν, by ν^N, we mean a product measure. We say that a measurable function $d(\mathbf{z})$ ($\mathbf{z} = (z_1, ..., z_N)$) is approximately continuous at a point $\mathbf{z}_0$ with respect to a measure ν, if $\forall\ \epsilon, \delta > 0$, $\exists$ a neighborhood $O(\mathbf{z}, r)$ of $\mathbf{z}_0$ with radius r such that

$$\frac{\nu^N(\{(\mathbf{z}) \in O(\mathbf{z}, r) : |d(\mathbf{z}) - d(\mathbf{z}_0)| > \epsilon\})}{\nu^N(\{(\mathbf{z}) \in O(\mathbf{z}, r)\})} \le \delta.$$

Denote μ_F the measure induced by a distribution function F. To prove the theorem, it suffices to show the following statement:

Given a measure ν induced by a continuous distribution function, if $E_{F,F}\phi_0 \le E_{F,F}\phi$ and $E_{F,G}\phi_0 \ge E_{F,G}\phi$ for each pair of F, $G \in \mathcal{F}$ where μ_F and μ_G are absolutely continuous with respect to ν and $F \neq G$, then $\phi_0 = \phi$ a.s. ν^N.

Without loss of generality, we only need to consider the case that ν^N is the same as the Lebesgue measure μ^N on the N-dimensional Euclidean space. We shall show that if $\phi = \phi_0$ a.s. is not true, then it leads to a contradiction.

By (3.8), ϕ is constant in a neighborhood of every point $\mathbf{z}$ whose coordinates are all distinct. Since ϕ_0 is measurable, it is approximately continuous almost everywhere (Munroe (1953)). If $\phi_0 = \phi$ a.e. is not true, then there is a point $\eta = (\eta_1, ..., \eta_N)$ such that its coordinates are all distinct,

$$\phi_0 \text{ is approximately continuous at } \eta \text{ and } \phi_0(\eta) \neq \phi(\eta). \tag{4.1}$$

Let $\xi_1 < \cdots < \xi_N$ be the order statistics of $\eta_1, ..., \eta_N$. Since $(\mathbf{u}, \mathbf{t})$ is a sufficient and complete statistic for (F_p, G_q) or $(\mathbf{p}, \mathbf{q})$, we can define a test

$$\phi_2(\mathbf{u}, \mathbf{t}) = E_{F_p, G_q}(\phi_0(\mathbf{Z})|\mathbf{u}, \mathbf{t}). \tag{4.2}$$

By definition and (3.1), ϕ_1 and ϕ_2 are both constant in the neighborhood of η, $O(\eta, 1/k)$. Verify that $\phi(\eta) = \phi_1(\eta)$, as ϕ is constant in a neighborhood of every point $\mathbf{z}$ whose coordinates are all distinct and η is such a point. Then by (4.1) and by taking k large enough

$$\phi_2 \neq \phi_1 \text{ for each } \mathbf{z} \in O(\eta, 1/k). \tag{4.3}$$

Furthermore,

$$E_{F_q, G_p}(\phi_1 - \phi_2) = E_{F_q, G_p}(\phi - \phi_0).$$

Since ϕ_0 is as good as ϕ,

$$E_{F_q, G_p}(\phi_1 - \phi_2) = E_{F_q, G_p}(\phi - \phi_0) = 0 \text{ for all } F_q, G_p \in \mathcal{F}_k(\xi),$$

as ϕ is admissible when $\Theta = \mathcal{F}_k(\xi)$. It follows that $\phi_1 = \phi_2$ on Ω, as ϕ_1 and ϕ_2 are both functions of $(\mathbf{u}, \mathbf{t})$, which is complete and sufficient for (F_q, G_p). Thus $\phi_1 = \phi_2$ a.s on $O(\eta, 1/k)$ $(\subset \Omega)$, which contradicts (4.3).

Verify that ϕ_0 is arbitrary, thus the contradiction shows that ϕ is admissible within the class of all tests. We can also assume ϕ_0 is a rank test, thus the above contradiction shows that ϕ is admissible within the class of all rank tests. This completes the proof of the theorem. ▫

5. Applications. In this section, we apply the theorem to show that the tests of form (2.1) are admissible in some special cases. In particular, in Theorem 5.1, we assume $\max\{m, n\} \leq 2$ but the size of the test is arbitrary, and in Theorem 5.2, we assume that the size of the test is $\leq \frac{4}{\binom{N}{n}}$ but m and n are arbitrary.

Theorem 5.1. *Consider the problem of testing H_0 against H_i with $\Theta = \mathcal{F}$, where i =1, 2, 3, 4. Suppose either (1) min$\{n, m\} = 1$ or (2)* $\max\{n, m\} = 2$ *but $\gamma(\cdot) = 0$. Let ϕ be a test of form (2.1). Then ϕ is admissible within the class of all tests*

Proof. We only give the proof for testing against H_1, as the others are very similar. Replacing $c + l$ by l in (3.6), (2.1) becomes

$$\phi = 1_{[\mathbf{a}\cdot\mathbf{t}^o < l]} + \gamma(\mathbf{t}^o)1_{[\mathbf{a}\cdot\mathbf{t}^o = l]}, \quad \text{where } \gamma(\cdot) \in [0, 1).$$

In view of Theorem 4.1, in order to prove the admissibility of ϕ, it suffices to verify condition (2) of Lemma 3.1. Hereafter, we fix $\mathbf{u}$ and assume that $\mathbf{t}_0$ satisfies $\phi_1(\mathbf{u}, \mathbf{t}_0) > 0$.

Case $n = 1$. If all the coordinates of $\mathbf{t}_0$ are zero, then $\exists\ \mathbf{t}_0^o \in T_{\mathbf{t}_0}$ such that $\mathbf{a} \cdot \mathbf{t}_0^o \leq l$ as $\phi_1(\mathbf{u}, \mathbf{t}_0) > 0$. For each $\mathbf{t}$, every $\mathbf{t}^o \in T_{\mathbf{t}}$ satisfies $\mathbf{a} \cdot \mathbf{t}^o < \mathbf{a} \cdot \mathbf{t}_0^o \leq l$. Consequently $\phi(\mathbf{t}^o) = 1$ and thus $\phi(\mathbf{u}, \mathbf{t}) = 1$ in view of (3.6). Thus, $\mathbf{t} \notin C_{\mathbf{u}, \phi_1}$.

Without loss of generality, we can assume that there is only one coordinate of $\mathbf{t}_0$, say, the i_0th coordinate, that is not zero, and it must be 1. Let $\mathbf{b}$ be a $N - 1$ dimensional vector whose i_0th coordinate is 2 and the rest are 0. Note that $\mathbf{b} \geq 0$. Then $\mathbf{b} \cdot (\mathbf{t} - \mathbf{t}_0) = 2 \cdot (0 - 1) < 0$ for all $\mathbf{t} \neq \mathbf{t}_0$, thus condition (2) holds.

Case $m = n = 2$. There are at most two coordinates of $\mathbf{t}_0$ that are not zero: either (2.a) there are exactly two coordinates of $\mathbf{t}_0$ being 1 in $\mathbf{t}_0$, say, the indexes of the two coordinates are k_0 and j_0, or (2.b) the k_0th coordinate of $\mathbf{t}_0$ is 2, or (2.c) the k_0th coordinate of $\mathbf{t}_0$ is 1 and the rest coordinates are all zero, or (2.d) all the coordinates of $\mathbf{t}_0$ are 0.

Consider case (2.a). Since $\sum_{i=1}^{N-1} u_i \leq N = 4$, either (2.a.1) at least one of the k_0th and j_0th coordinates of $\mathbf{u}$ (say the k_0th) is 1, or (2.a.2) both the k_0th and j_0th coordinates of $\mathbf{u}$ are two.

In case (2.a.1), let $\mathbf{b} = (b_1, b_2, b_3)$ be such that $b_{k_0} = 2$, $b_{j_0} = 1$, and the other $b_h = 0$. Verify that $\mathbf{b}$ satisfies condition (2). In fact in case (2.a.1), every point $\mathbf{t}$ satisfies $\mathbf{b}\cdot(\mathbf{t}-\mathbf{t}_0) < 0$ if $\mathbf{t} \neq \mathbf{t}_0$, as the k_0th coordinate of $\mathbf{t}$ is at most 1 and the j_0th coordinate of $\mathbf{t}$ is at most 2.

In case (2.a.2), let $k_0 < j_0$ and define $\mathbf{b}$ as above, then the only point $\mathbf{t} \neq \mathbf{t}_0$ satisfying $\mathbf{b}\cdot(\mathbf{t}-\mathbf{t}_0) \geq 0$ has the k_0th coordinate being 2. We shall show next that the latter point $\mathbf{t} \notin C_{\mathbf{u},\phi_1}$. Since $\phi_1(\mathbf{u}, \mathbf{t}_0) > 0$, by (3.7) there is a $\mathbf{t}_0^o \in T_{\mathbf{t}_0}$ such that $\phi(\mathbf{t}_0^o) > 0$. Verify that $\mathbf{a}\cdot\mathbf{t}^o < \mathbf{a}\cdot\mathbf{t}_0^o$ $\forall$ $\mathbf{t}^o \in T_{\mathbf{t}}$, as $k_0 < j_0$ and $a_1 < \cdots < a_N$, where $\mathbf{a} = (a_1, ..., a_N)$. As consequences, $\phi(\mathbf{t}^o) = 1$ by (3.8) and $\phi_1(\mathbf{u}, \mathbf{t}) = 1$ by (3.7). It follows that $\mathbf{t} \notin C_{\mathbf{u},\phi_1}$. Thus condition (2) holds.

Consider case (2.b) or (2.c). Let $\mathbf{b}$ be such that its k_0th coordinate is 1 and the rest are zero. Then condition (2) holds.

Consider case (2.d). By assumption in the theorem, ϕ is a non-randomized test. Since $\phi(\mathbf{u}, \mathbf{t}_0) > 0$, there is $\mathbf{t}_0^o \in T_{\mathbf{t}_0}$ such that $\phi(\mathbf{t}_0^o) = 1$. Verify that for each $\mathbf{t} \in T^{\mathbf{u}}$, if $\mathbf{t}^o \in T_{\mathbf{t}}$ then $\mathbf{a}\cdot\mathbf{t}^o \leq \mathbf{a}\cdot\mathbf{t}_0^o$. Consequently, $\phi_1(\mathbf{u}, \mathbf{t}) = 1$. Thus condition (2) holds.

The proof for the case $m = 1$ is similar. This concludes the proof of the theorem. □

Theorem 5.2. *Consider the problem of testing H_0 against H_3 (or H_4). Let $\phi = 1_{[\mathbf{a}\cdot\mathbf{t}^o \notin (l,r)]} + \gamma 1_{[\mathbf{a}\cdot\mathbf{t}^o = l \text{ or } r]}$, where $l = \frac{n(n+1)}{2} + 1$, $r = \frac{n(m+1+m+n)}{2} - 1$ and $\mathbf{a} = (1, 2, ..., N)$. Then ϕ is admissible within the class of all tests.*

Proof. It suffices to show that for each $\mathbf{u}$,

(B) if $\phi(\mathbf{u}, \mathbf{t}_0) > 0$, then there is a vector $\mathbf{b}$ such that (1) $\mathbf{b} \geq \mathbf{0}$ or $\mathbf{b} \leq \mathbf{0}$ and (2) for each $\mathbf{t}$, $\mathbf{b}\cdot(\mathbf{t}-\mathbf{t}_0) \geq 0$ and $\mathbf{t} \neq \mathbf{t}_0$ imply that $\phi(\mathbf{t}^o) = 1$ for all $\mathbf{t}^o \in T_{\mathbf{t}}$.

By (3.7), statement (B) implies that $\phi_1(\mathbf{u}, \mathbf{t}) = 1$ and thus $\mathbf{t} \notin C_{\mathbf{u},\phi_1}$. It follows that condition (2) of Lemma 3.1 holds. As a consequence, ϕ is admissible by Theorem 4.1.

If $\phi(\mathbf{u}, \mathbf{t}_0) > 0$, then by (3.7) there exists a $\mathbf{t}_0^o \in T_{\mathbf{t}_0}$ such that $\phi(\mathbf{t}_0^o) > 0$. By the assumption on ϕ in the theorem, one of the following must be true:

1. The first n coordinates of $\mathbf{t}_0^o$ are 1 and the rest are zero;
2. The first $n-1$ coordinates and the $(n+1)$st coordinate of $\mathbf{t}_0^o$ are 1 and the rest are zero;
3. The last n coordinates of $\mathbf{t}_0^o$ are 1 and the rest are zero;
4. The last $n-1$ coordinates and the mth coordinate of $\mathbf{t}_0^o$ are 1 and the rest are zero.

Let t_{0,i_1} and t_{0,i_1+j} be first and the last nonzero coordinates of $\mathbf{t}_0$, respectively.

In the first two cases, let $\mathbf{b}$ be the vector such that its $(i_1, ..., i_1+j)$th coordinates are $(j+1, j, ..., 1)$ and the rest are zero. It is obvious that $\mathbf{b} \geq \mathbf{0}$.

Verify that if case 1 is true, then there is no $\mathbf{t} \neq \mathbf{t}_0$ such that $\mathbf{b}\cdot(\mathbf{t}-\mathbf{t}_0) \geq 0$. The reason is as follows. By (3.3), (3.4) and (3.5), the (1, ..., i_1+j-1)th coordinates of $\mathbf{u}$ and $\mathbf{t}_0$ are the same and $u_{i_1+j} \geq$ the (i_1+j)th coordinate of $\mathbf{t}_0$, where u_i is the ith coordinate of $\mathbf{u}$. Then $\mathbf{b}\cdot(\mathbf{t}-\mathbf{t}_0) \geq 0$ implies $\mathbf{t} = \mathbf{t}_0$.

Now if case 2 is true and if $\mathbf{b}\cdot(\mathbf{t}-\mathbf{t}_0) \geq 0$ and $\mathbf{t} \neq \mathbf{t}_0$, then (3.3), (3.4) and (3.5) and the structure of $\mathbf{t}_0^o$ in case 2 imply that $T_{\mathbf{t}}$ consists of only one element $\mathbf{t}^o$ and the first n coordinates of $\mathbf{t}^o$ are 1 and the rest are zero. Verify that the point $\mathbf{t}$ satisfies $\mathbf{a}\cdot\mathbf{t}^o < \mathbf{a}\cdot\mathbf{t}_0^o < l$, thus $\phi(\mathbf{t}^o) = 1$ and consequently, $\phi_1(\mathbf{u}, \mathbf{t}) = 1$. So statement (B) holds.

On the other hand, if either case 3 or case 4 is true, let $\mathbf{b} = -(N-1, ..., 2, 1)$. Thus $\mathbf{b} \leq \mathbf{0}$.

Verify that if case 3 is true, then there is no $\mathbf{t} \neq \mathbf{t}_0$ such that $\mathbf{b}\cdot(\mathbf{t}-\mathbf{t}_0) \geq 0$. The argument is somewhat similar to that for case 1. So statement (B) holds.

Now if case 4 is true and if $\mathbf{b}\cdot(\mathbf{t}-\mathbf{t}_0) \geq 0$ and $\mathbf{t} \neq \mathbf{t}_0$, then (3.3), (3.4) and (3.5) imply that $T_{\mathbf{t}}$ consists of only one element $\mathbf{t}^o$ and the last n coordinates of $\mathbf{t}^o$ are 1 and the rest are zero. Verify that the point $\mathbf{t}^o$ satisfies $\mathbf{a}\cdot\mathbf{t}^o > \mathbf{a}\cdot\mathbf{t}_0^o > r$, thus $\phi(\mathbf{t}^o) = 1$ and consequently, $\phi_1(\mathbf{u}, \mathbf{t}) = 1$. So statement (B) holds. □

Remark 1. Unlike Theorem 5.1, the test ϕ stated in Theorem 5.2 is the Wilcoxon test. It is worth mentioning that for the cases considered in Theorem 5.2, the Wilcoxon test is a representative of other tests of form (2.1). In fact, as pointed out in the proof, there are only four cases that $\phi > 0$, and the test $\phi = 1$ if case 1 or 3 is true and $= \gamma$ if case 2 or 4 is true. In the form of (2.1), given $c_1 < c_2$ and $\mathbf{b} \neq (1, ..., N)$ but $b_1 < b_2 < \cdots < b_N$, it is easy to find l_b and r_b such that the test ϕ in Theorem 5.2 is the same as

$1_{[\mathbf{b}\cdot\mathbf{t}^o \notin (l_b, r_b)]} + \gamma 1_{[\mathbf{b}\cdot\mathbf{t}^o = l_b \text{ or } r_b]}$. Consequently, the Wilcoxon test, the Fisher-Yates test and the median test are admissible in the cases mentioned in Theorems 5.1 and 5.2.

Remark 2. It is easy to modify the proof of Theorem 5.2 to show the following results:

1. For testing against H_1, $\phi = 1_{[\mathbf{a}\cdot\mathbf{t}^o < l]} + \gamma 1_{[\mathbf{a}\cdot\mathbf{t}^o = l]}$ is admissible within the class of all test when $\Theta = \mathcal{F}_k$ if the size of ϕ is $\leq \frac{2}{\binom{N}{n}}$.
2. For testing against H_2, $\phi = 1_{[\mathbf{a}\cdot\mathbf{t}^o > r]} + \gamma 1_{[\mathbf{a}\cdot\mathbf{t}^o = r]}$ is admissible within the class of all test when $\Theta = \mathcal{F}_k$ if the size of ϕ is $\leq \frac{2}{\binom{N}{n}}$.

Remark 3. Admissibility within the class of all rank tests is different from admissibility within the class of all tests. A standard t-test is not a rank test. If a test is admissible within the class of all tests, it is admissible within the class of all rank tests, but not vice-verse.

6. Comments. The paper makes progress in attacking two well-known open questions in Lehmann's famous textbook "Testing statistical hypotheses" (1959). To the best of the author's knowledge, the open questions were not settled in any special case before. The main difficulty is that the problems were not tractable.

The significance of the current paper is Theorem 4.1, not Theorems 5.1 and 5.2. Theorem 4.1 provides a sufficient condition for admissibility of the Wilcoxon test and makes the open problems solvable. Theorems 5.1 and 5.2 demonstrate that Theorem 4.1 can indeed be used to solve the open questions for some special sample sizes, or for some special significance levels α. At this moment, the author believes that Theorem 4.1 can be used to show that the non-randomized Wilcoxon test is admissible for each sample size n and each attainable significance level α on a case-by-case basis, but is unable to produce a unified proof.

Note that the sufficient condition in Theorem 4.1 is applicable for a wide class of linear rank tests including the Fisher-Yates test and the median test. In Section 5, we apply the theorem and show that the Wilcoxon test is admissible in some special cases. The cases considered in Section 5 are some special cases in in practice. They were chosen because the proofs are relatively easy to follow and thus it makes the paper more readable. It is possible to establish admissibility results for the Wilcoxon test in additional cases other than those listed in Theorems 5.1 and 5.2 and Remark 2. Thus the results in Section 5 are **not** the only cases that Theorem 4.1 is applicable.

We further point out that Theorem 4.1 is a sufficient condition and may not be a necessary condition. In view of the inadmissibility results on the randomized Wilcoxon tests, it is possible that the tests of form (2.1) maybe inadmissible within the class of all tests if $\Theta = \mathcal{F}_k(\xi)$ for some special cases. However, it is still not clear whether the tests of form (2.1) is admissible within the class of all tests when $\Theta = \mathcal{F}$. Indeed, Theorem 5.2 already presents a different result on the continuous two-sample problem from the one on the discrete set-up. If $n, m \geq 4$ and the assumption of Theorem 5.2 holds, each non-randomized Wilcoxon test with $\gamma = c \in (0,1)$ is inadmissible in the discrete set-up (see Yu (2000)) but is admissible in the continuous set-up.

The test of form (2.1) is a rank test if γ is a constant. However, if γ is not a constant, it may not be a rank test. The test of form (2.1) may not be of practical importance, as probably nobody would use a randomized test in the application. However, it is important in the statistics theory, as is well known that without the concept of randomized tests, the Lehmann-Pearson Lemma would not exist. The latter lemma is a foundation of the theory on testing statistical hypotheses.

Acknowledgement. Helpful discussion on the classical nonparametric two-sample problem with Professor Larry D. Brown is acknowledged. He kindly pointed out an error in an earlier version of the paper. The author also thanks two referees and the editor for invaluable comments.

References.

* Ferguson, T. S. (1967). Mathematical statistics, a decision theoretic approach. *Academic*, New York, 250-255.
* Lehmann, E. L. (1959). Testing statistical hypotheses. 1st ed. *Wiley*, New York,
* Lehmann, E. L. (1986). Testing statistical hypotheses. 2nd ed. *Wiley*, New York.
* Munroe, M. E. (1953). Introduction to measure and integration. *Addison -Wesley*, Massachusetts, 291-292.

* Sugiura, N. (1965). An example of the two-sided Wilcoxon test which is not unbiased. *Ann. Inst. Statist. Math.* 17, 261-263.
* Yu, Q. Q. (2000). Inadmissibility and admissibility of randomized Wilcoxon tests in discrete two-sample problems. *Statistics & Decisions*, 18, 35-48.

Appendix To the Paper

A SUFFICIENT CONDITION FOR ADMISSIBILITY OF THE WILCOXON TEST IN THE CLASSICAL TWO-SAMPLE PROBLEM

For the convenience of readers, in this appendix, we give the proof of Lemma 3.1, which is a minor modification of Theorem 4.1 in Yu (2000). Given an arbitrary test ϕ, we say a $\mathbf{u}$-section of ϕ is admissible for testing $\mathbf{w}=0$ against H_1 if there is no test ψ such that $E_{F_p,G_q}([\phi-\psi]|\mathbf{u})\leq 0$ and $E_{F_p,F_p}([\phi-\psi]|\mathbf{u})\geq 0$ for all F_p, $G_q\in\mathcal{F}_k(\xi)$ with $F_p\leq G_q$, and with at least one strict inequality.

For testing against H_1, Lemma 3.1 follows from Lemmas A.1 and A.2 below. For testing against H_2 or H_3 or H_4, the proof is almost identical to the proof given below.

Lemma A.1. *For testing H_0 against H_1 when $\Theta=\mathcal{F}_k(\xi)$, a test ϕ as (2.1) is admissible if every $\mathbf{u}$-section of ϕ is admissible for testing $\mathbf{w}=\mathbf{0}$ against H_1.*

Lemma A.2. *Consider the problem of testing H_0 against H_1 when $\Theta=\mathcal{F}_k(\xi)$. Let ϕ be a one-sided test satisfying conditions (1) and (2) of Lemma 3.1, then every $\mathbf{u}$-section of ϕ is admissible for testing $\mathbf{w}=\mathbf{0}$ against H_1.*

Proof of Lemma A.1. Assume every $\mathbf{u}$-section of ϕ is admissible for testing $\mathbf{w}=0$. Since $q_i\leq p_i$, $i<N$, imply $F_q\leq G_p$, it suffices to consider testing against H_1^*: $q_i\leq p_i$ for $i=1$, ..., $N-1$, but $\mathbf{p}\neq\mathbf{q}$, or equivalently, $\mathbf{w}\geq\mathbf{0}$ but $\mathbf{w}\neq\mathbf{0}$. Let us denote the set of all possible values of $\mathbf{u}$ by $\mathcal{U}$, and let $\mathbf{u}_0$ be an extreme point of the convex hull of $\mathcal{U}$. Note that $\mathcal{U}$ consists of finitely many elements. If ϕ_2 is as good as ϕ, then

$$\sum_{\mathbf{u}\in\mathcal{U}}e^{\theta\cdot\mathbf{u}}\sum_{\mathbf{t}}h(\mathbf{u},\mathbf{t})e^{\mathbf{w}\cdot\mathbf{t}}[\phi_2(\mathbf{u},\mathbf{t})-\phi(\mathbf{u},\mathbf{t})]\geq 0\ \ \forall\,\theta \text{ and } \forall\,\mathbf{w}\geq\mathbf{0}, \tag{7.1}$$

with equality when $\mathbf{w}=\mathbf{0}$. Consider a hyper-plane $\mathbf{b}\cdot(\mathbf{u}-\mathbf{u}_0)=0$ which supports the convex hull of $\mathcal{U}$ at $\mathbf{u}$ and such that $\mathbf{b}\cdot(\mathbf{u}-\mathbf{u}_0)<0$ for all $\mathbf{u}\in\mathcal{U}$ and $\mathbf{u}\neq\mathbf{u}_0$. It is important to notice that θ in (7.1) is arbitrary, though $\mathbf{w}$ is restricted. In (7.1) let $\theta=v\mathbf{b}$, and multiply (7.1) by $e^{-v\mathbf{b}\cdot\mathbf{u}_0}$. Letting $v\to\infty$ yields

$$\sum_{\mathbf{t}}h(\mathbf{u}_0,\mathbf{t})e^{\mathbf{w}\cdot\mathbf{t}}[\phi_2(\mathbf{u}_0,\mathbf{t})-\phi(\mathbf{u}_0,\mathbf{t})]\geq 0 \text{ for all } \mathbf{w}\geq 0, \tag{7.2}$$

with equality when $\mathbf{w}=\mathbf{0}$. Since each $\mathbf{u}$-section of ϕ is admissible, (7.2) implies

$$\sum_{\mathbf{t}}h(\mathbf{u}_0,\mathbf{t})e^{\mathbf{w}\cdot\mathbf{t}}[\phi_2(\mathbf{u}_0,\mathbf{t})-\phi(\mathbf{u}_0,\mathbf{t})]=0 \text{ for all } \mathbf{w}\geq\mathbf{0}.$$

Since $\{\mathbf{w}:\ \mathbf{w}\geq\mathbf{0}\}$ contains a nonempty $N-1$ dimensional open set, $\mathbf{t}$ is complete and sufficient for the conditional distribution $\mu_{F_q,G_p}(\mathbf{u}_0,\cdot)$. Thus $\phi(\mathbf{u}_0,\mathbf{t})=\phi_2(\mathbf{u}_0,\mathbf{t})$ for all possible $\mathbf{t}$. In the latter case, we can replace $\mathcal{U}$ by $\mathcal{U}\setminus\{\mathbf{u}_0\}$ in (7.1) and repeat the argument for an extreme point of $\mathcal{U}\setminus\{\mathbf{u}_0\}$. After finitely many steps we must either arrive at a contradiction or conclude $\phi_2=\phi$ for all possible $(\mathbf{u},\mathbf{t})$. □

Proof of Lemma A.2. We shall show that if the $\mathbf{u}$-section of ϕ is not admissible, then we can reach a contradiction.

Suppose a $\mathbf{u}$-section of ϕ is dominated by a different test ψ. Then there must exist some point $(\mathbf{u},\mathbf{t}_0)$ such that $\phi(\mathbf{u},\mathbf{t}_0)>\psi(\mathbf{u},\mathbf{t}_0)$. Otherwise, it could not be true that $E_{F_p,F_p}([\phi-\psi]|\mathbf{u})=0$. Because $\phi=0$ for all non-extreme points $\mathbf{t}$ of $C_{\mathbf{u},\phi}$ by condition (2), we see that $\mathbf{t}_0$ must belong to the complement of $C_{\mathbf{u},\phi}$ or be an extreme point of $C_{\mathbf{u},\phi}$. Consequently, $\mathbf{w}=v\mathbf{b}\geq 0\ \forall\ v>0$, where $\mathbf{b}$ is the vector in condition (2). Thus $\mathbf{w}=v\mathbf{b}$ is a proper parameter under H_1. Letting $\mathbf{w}=v\mathbf{b}$, the fact that $\mathbf{u}$-section of ψ dominates the one of ϕ gives

$$\begin{aligned}0 \leq\, & e^{-\mathbf{w}\cdot\mathbf{t}_0}\sum_{\mathbf{t}}[\psi(\mathbf{u},\mathbf{t})-\phi(\mathbf{u},\mathbf{t})]h(\mathbf{u},\mathbf{t})e^{\mathbf{w}\cdot\mathbf{t}} \\ \leq\, & \sum_{\mathbf{t}\in C_{\mathbf{u},\phi}\setminus\{\mathbf{t}_0\}}[\psi(\mathbf{u},\mathbf{t})-\phi(\mathbf{u},\mathbf{t})]h(\mathbf{u},\mathbf{t})e^{v\mathbf{b}\cdot(\mathbf{t}-\mathbf{t}_0)}+[\psi(\mathbf{u},\mathbf{t}_0)-\phi(\mathbf{u},\mathbf{t}_0)]h(\mathbf{u},\mathbf{t}_0)\end{aligned} \tag{7.3}$$

(as $\phi=1$ on $C^c_{\mathbf{u},\phi}$), where $C^c_{\mathbf{u},\phi}$ is the complement of $C_{\mathbf{u},\phi}$. Taking limits on both sides of (7.3) as $v\to\infty$ yields $0\leq\psi(\mathbf{u},\mathbf{t}_0)-\phi(\mathbf{u},\mathbf{t}_0)$ by condition (2), which contradicts the assumption that $\psi<\phi$ at $(\mathbf{u},\mathbf{t}_0)$. This completes the proof of the lemma. □