Nonparametric one-sided testing for the mean and related extremum problems

Norbert Gaffke

University of Magdeburg, Faculty of Mathematics
D-39016 Magdeburg, PF 4120, Germany
E-mail: norbert.gaffke@mathematik.uni-magdeburg.de

Abstract

We consider the nonparametric model of n i. i. d. nonnegative real random variables whose distribution is unknown. An interesting parameter of that distribution is its expectation μ. Wang & Zhao (2003) studied the problem of testing the one-sided hypotheses $H_0 : \mu \leq \mu_0$ vs. $H_1 : \mu > \mu_0$ (with a given $\mu_0 > 0$, where w.l.g. one may take $\mu_0 = 1$). For $n = 1$ there is a UMP nonrandomized level α test. Somewhat surprisingly, for $n = 2$ Wang & Zhao obtained a UMP nonrandomized monotone symmetric level α test. However, they conjectured that the result will not carry over to larger sample size $n \geq 3$. Unfortunately, their conjecture is true as we will show. Also, we present an alternative proof of their (positive) result for $n = 2$. Our derivations are based on a study of related classes of extremum problems on products of probability measures.

Key words: Monotone symmetric test, UMP test, order statistics, probability measures, weak topology, semi-continuity.

2000 Mathematics Subject Classification: Primary 62G10, secondary 28C15.

1 Introduction

Let $X_1, \ldots, X_n$ be *nonnegative* real i.i.d. random variables whose distribution P^{X_i} is unknown. We are interested in the expectation parameter $\mu = \mathrm{E}(X_i)$. Note that $\mu \in [0\,,\,\infty]$. Consider the one-sided testing problem,

$$H_0 \;:\; \mu \leq \mu_0 \quad \text{vs.} \quad H_1 \;:\; \mu > \mu_0 \;,$$

for a given $\mu_0 > 0$. W.l.g. we may assume that $\mu_0 = 1$, since for any given value $\mu_0 > 0$ we can transform to that case via $\widetilde{X}_i = X_i/\mu_0$ $(1 \leq i \leq n)$. So in the following we will restrict to the testing problem

$$H_0 \;:\; \mu \leq 1 \quad \text{vs.} \quad H_1 \;:\; \mu > 1 \;. \tag{1}$$

The recent paper [1] deals with testing the one-sided hypotheses (1). For $n = 1$ the UMP nonrandomized level α test is given by $\phi^* = \mathbb{1}_{[1/\alpha\,,\,\infty)}$. For $n = 2$ there was derived in [1] a UMP test within a class of tests slighly larger than the class of all nonrandomized monotone (nondecreasing) symmetric level α tests (see Section 3). This is is a highly nontrivial and somewhat surprising result. Note that a test ϕ is said to be symmetric iff it is a symmetric function of the observations $x_1, \ldots, x_n$, and ϕ is said to be monotone (nondecreasing) iff $x, z \in [\,0\,,\,\infty)^n$ and $x \leq z$ (w.r.t. the componentwise semi-ordering) imply $\phi(x) \leq \phi(z)$. Unfortunately, that result cannot be extended to larger sample size $n \geq 3$, as we will show in Section 5 and as it was conjectured in [1], p. 76. Also, we will give a proof of the positive result for $n = 2$ which, as we think, is more tranparent than that presented in [1]. Our derivations are based on a study of related classes of extremum problems. E. g., an extremum problem arises when the true level of a given test for (1) has to be determined. The auxiliary results on such extremum problems presented in Section 2 might also be of some independent interest.

2 Related extremum problems

Let S be a nonempty compact subset of $\mathbb{R}^n$, and let $\mathrm{Prob}(S)$ denote the set of all Borel probability measures on S. Endowed with the weak topology (inducing weak convergence) $\mathrm{Prob}(S)$ is a compact space (cf. [2], Satz 31.2 and Bemerkung 1 on p. 237). Clearly, $\mathrm{Prob}(S)$ is a convex set, i.e., if $Q_1, Q_2 \in \mathrm{Prob}(S)$ and $0 \leq \lambda \leq 1$ then $\lambda Q_1 + (1-\lambda) Q_2 \in \mathrm{Prob}(S)$. A real function $g : S \longrightarrow \mathbb{R}$ is said to be upper semi-continuous (u.s.c.) iff for any given $u \in \mathbb{R}$ the set $\{\, z \in S \,:\, g(z) \geq u\}$ is closed. Equivalently, g is u.s.c. iff for any sequence $z_k \in S$ ($k \in \mathbb{N}$) converging to some $z_0 \in S$ one has $\limsup_{k\to\infty} g(z_k) \leq g(z_0)$. Similarly, we have the notion of upper semi-continuity of a real function on $\mathrm{Prob}(S)$. A function $G : \mathrm{Prob}(S) \longrightarrow \mathbb{R}$ is said to be u.s.c. iff for any given $u \in \mathbb{R}$ the set $\{\, Q \in \mathrm{Prob}(S) \,:\, G(Q) \geq u\,\}$ is a closed set w.r.t. the weak topology. Equivalently, the u.s.c. property of G means that whenever $Q_k \in \mathrm{Prob}(S)$ ($k \in \mathbb{N}$) is a sequence converging weakly to some $Q_0 \in \mathrm{Prob}(S)$ then $\limsup_{k\to\infty} G(Q_k) \leq G(Q_0)$.

Lemma 2.1 *Let S be a nonempty compact subset of $\mathbb{R}^n$ and $g : S \longrightarrow \mathbb{R}$ be a nonnegative u.s.c. function. Then, by $G(Q) = \mathrm{E}_Q(g)$ (the expectation of g w.r.t. Q), for all $Q \in \mathrm{Prob}(S)$, we have a nonnegative real u.s.c. function on $\mathrm{Prob}(S)$.*

Proof. Firstly, we consider the special case that g is the indicator function of a closed subset C of S, i.e., $g = \mathbb{1}_C$ on S. Then $G(Q) = Q(C)$ for all $Q \in \mathrm{Prob}(S)$. If Q_k ($k \in \mathbb{N}$) is a sequence in $\mathrm{Prob}(S)$ which converges weakly to some $Q_0 \in \mathrm{Prob}(S)$, then by [3], Theorem 2.1, we have $\limsup_{k\to\infty} Q_k(C) \leq Q_0(C)$, i.e., the function G is u.s.c. Now let g be an arbirary nonnegative u.s.c. function on S. As an u.s.c. function on the compact set S the function g is bounded above, and hence $0 \leq g(z) \leq v$ for all $z \in S$ with some finite $v > 0$. In particular, $G(Q) = \mathrm{E}_Q(g)$ is nonnegative and finite for all $Q \in \mathrm{Prob}(S)$. Also, together with a well-known formula for expectations of nonnegative functions, we have for all $Q \in \mathrm{Prob}(S)$, abbreviating $C_g(u) = \{\, z \in S \,:\, g(z) \geq u\,\}$,

$$\mathrm{E}_Q(g) \;=\; \int_0^\infty Q(\,C_g(u)\,)\,\mathrm{d}u \;=\; \int_0^v Q(\,C_g(u)\,)\,\mathrm{d}u\;.$$

Let Q_k be a sequence in $\mathrm{Prob}(S)$ which converges weakly to $Q_0 \in \mathrm{Prob}(S)$. Since for any $0 \leq u \leq v$ the set $C_g(u)$ is closed, we have by the above, $\limsup_{k\to\infty} Q_k(\,C_g(u)\,) \leq Q_0(\,C_g(u)\,)$ for all $0 \leq u \leq v$, and hence (by a result from integration theory, cf. [2],

p. 100, Ex. 1),

$$\limsup_{k\to\infty} \int_0^v Q_k(\, C_g(u)\,)\,\mathrm{d}u \;\le\; \int_0^v \limsup_{k\to\infty} Q_k(\, C_g(u)\,)\,\mathrm{d}u \;\le\; \int_0^v Q_0(\, C_g(u)\,)\,\mathrm{d}u \;,$$

i.e., $\limsup_{k\to\infty} G(Q_k) \;\le\; G(Q_o)$, and so G is u.s.c. □

The *support* of a probability measure $Q \in \mathrm{Prob}(S)$, denoted by $\mathrm{supp}(Q)$, is the smallest closed subset of S which has Q-probability equal to one. That is, $\mathrm{supp}(Q)$ is a closed subset of S with $Q(\mathrm{supp}(Q)) \;=\; 1$, and if A is any closed subset of S with $Q(A) \;=\; 1$ then $\mathrm{supp}(Q) \subset A$. More explicitly, $\mathrm{supp}(Q)$ consists of all points $z \in S$ such that for each open neighborhood U_z of z in $\mathbb{R}^n$ one has $Q(\, U_z \cap S\,) \;>\; 0$. Clearly, if Q is an r-point distribution (where $r \in \mathbb{N}$) giving probabilities $q_1, \ldots, q_r > 0$ to distinct points $z_1, \ldots, z_r \in S$, resp., (where $\sum_{i=1}^r q_i = 1$), for short

$$Q \;=\; \begin{pmatrix} z_1 & \ldots & z_r \\ q_1 & \ldots & q_r \end{pmatrix} \;,$$

then $\mathrm{supp}(Q) \;=\; \{\, z_1, \ldots, z_r \,\}$.

Theorem 2.2 *For a given real number $c > 1$ consider the compact interval $I \;=\; [\,0\,,\,c\,]$ of the real line, and let $g \;:\; I \longrightarrow \mathbb{R}$ be a nonnegative u.s.c. function. Consider the extremum problem*

$$\text{(EP1)} \qquad \begin{array}{l} \text{maximize} \quad G(Q) \;=\; \mathrm{E}_Q(g) \\ \text{s.t.} \quad Q \in \mathrm{Prob}(I)\,, \quad \mathrm{E}[Q] \;=\; 1\;, \end{array}$$

where $\mathrm{E}[Q] \;=\; \int_I z\,\mathrm{d}Q(z)$ denotes the expectation of Q. Then:

(i) *There exists an optimal solution Q_0 to problem* (EP1)*.*

(ii) *Let $Q_0 \in \mathrm{Prob}(I)$ with $\mathrm{E}[Q_0] = 1$ be given. Q_0 is an optimal solution to* (EP1) *if and only if there exist $\rho \in [\,0\,,\,\infty)$ and $\tau \in \mathbb{R}$ such that*

$$g(z) \le \rho + \tau z \quad \forall\; z \in I \qquad \textit{and} \quad g(z) = \rho + \tau z \quad \forall\; z \in \mathrm{supp}(Q_0)\;.$$

(iii) *There exists a one- or two-point distribution Q_0 which is an optimal solution to* (EP1)*.*

Proof. The feasible region of (EP1) is compact w.r.t. the weak topology and, by Lemma 2.1, the objective function G is u.s.c. w.r.t. the weak topology. This proves (i). To prove the 'only if' part of (ii), let Q_0 be an optimal solution to (EP1). Consider the subset of $\mathbb{R}^2$,

$$C \;=\; \{\,(\,\mathrm{E}[Q]\,,\,G(Q)\,)\;:\; Q \in \mathrm{Prob}(I)\,\}\;.$$

Because of $\mathrm{E}[\lambda Q_1 + (1-\lambda)Q_2] \;=\; \lambda\mathrm{E}[Q_1] + (1-\lambda)\mathrm{E}[Q_2]$ and $G(\lambda Q_1 + (1-\lambda)Q_2)) = \lambda G(Q_1) + (1-\lambda)G(Q_2)$ for all $Q_1, Q_2 \in \mathrm{Prob}(I)$ and all $0 \le \lambda \le 1$, the set C is convex. Denote by γ the optimum value of (EP1), i.e., $\gamma = G(Q_0)$. So the point $(1, \gamma)$ belongs to C and, clearly, it must be a boundary point of C. So there is a supporting straight line for C at $(1, \gamma)$, i.e., there exist $a, b \in \mathbb{R}$ not both equal to zero and such that

$$a\mathrm{E}[Q] + bG(Q) \;\le\; a + b\gamma \quad \forall\; Q \in \mathrm{Prob}(I)\,. \tag{2}$$

Specializing to one-point distributions, (2) yields

$$az + bg(z) \;\le\; a + b\gamma \quad \forall\, z \in I\,. \tag{3}$$

The expectation w.r.t. Q_0 of the l.h.s. of (3) equals $a + b\gamma$, and hence

$$az + bg(z) \;=\; a + b\gamma \quad Q_0\text{-a.s.} \tag{4}$$

Case 1: $b < 0$.
Then (3) rewrites as

$$g(z) \;\ge\; \frac{a}{b} + \gamma - \frac{a}{b}z \quad \forall\, z \in I\,.$$

Suppose that there is some $z_0 \in I$ such that the strict inequality holds, i.e., $g(z_0) > \frac{a}{b} + \gamma - \frac{a}{b}z_0$. Then we can construct a one- or a two-point distribution $Q_1 \in \mathrm{Prob}(I)$ which has z_0 as a support point and with $\mathrm{E}[Q_1] = 1$. So Q_1 is feasible for (EP1) and $G(Q_1) = \mathrm{E}_{Q_1}(g) > \gamma$, which is a contradiction. Hence it follows that

$$g(z) \;=\; \frac{a}{b} + \gamma - \frac{a}{b}z \quad \forall\, z \in I\,,$$

and thus, taking $\rho = \frac{a}{b} + \gamma$ and $\tau = -\frac{a}{b}$, we have $g(z) = \rho + \tau z$ for all $z \in I$. In particular, $\rho = g(0) \ge 0$ and Q_0 satisfies the necessary condition of (ii).

Case 2: $b = 0$.
Then $a \neq 0$ and (3) rewrites as $az \le a$ for all $z \in I$. This yiels either $z \le 1$ for all $z \in I$ or $z \ge 1$ for all $z \in I$, which is a contradiction ($I = [\,0\,,\,c\,]$ and $c > 1$). So actually, case 2 cannot occur.

Case 3: $b > 0$.
Then, with $\rho = \frac{a}{b} + \gamma$ and $\tau = -\frac{a}{b}$, (3) and (4) rewrite as

$$g(z) \;\le\; \rho + \tau z \quad \forall\, z \in I\,, \quad \text{and} \quad g(z) \;=\; \rho + \tau z \quad Q_0\text{-a.s.}$$

The latter can be stated as

$$Q_0\Big(\,\{z \in I : \, g(z) \ge \rho + \tau z\}\,\Big) \;=\; 1\,,$$

and since g is u.s.c. the set of all $z \in I$ with $g(z) \ge \rho + \tau z$ is closed. So the support of Q_0 must be a subset of that set, and we have

$$g(z) \;=\; \rho + \tau z \quad \forall\, z \in \mathrm{supp}(Q_0)\,.$$

Clearly, $\rho \ge 0$ follows from $0 \le g(0) \le \rho$.
To prove the 'if' part of (ii), let $Q_0 \in \mathrm{Prob}(I)$ with $\mathrm{E}[Q_0] = 1$, and $\rho \ge 0$, $\tau \in \mathbb{R}$ such that $g(z) \le \rho + \tau z$ for all $z \in I$ and $g(z) = \rho + \tau z$ for all $z \in \mathrm{supp}(Q_0)$. Then, for any $Q \in \mathrm{Prob}(I)$ with $\mathrm{E}[Q] = 1$, we obtain

$$\mathrm{E}_Q(g) \;\le\; \rho + \tau \quad \text{and} \quad \mathrm{E}_{Q_0}(g) \;=\; \rho + \tau\,,$$

hence $G(Q) \le G(Q_0)$, and so Q_0 is an optimal solution to (EP1).

It remains to prove (iii). Choose any optimal solution Q_0 of (EP1) according to (i), and choose ρ and τ according to (ii). If $1 \in \mathrm{supp}(Q_0)$ then the one-point distribution at 1, δ_1

say, is trivially feasible to (EP1), and $G(\delta_1) = g(1) = \rho + \tau = \mathrm{E}_{Q_0}(g) = G(Q_0)$, i.e.. δ_1 is also an optimal solution to (EP1). Now let $1 \notin \mathrm{supp}(Q_0)$. Because of $\mathrm{E}[Q_0] = 1$, the support of Q_0 is neither a subset of the subinterval $[\,0\,,\,1\,]$ nor a subset of the subinterval $[\,1\,,\,c\,]$. Hence there exist $z_1, z_2 \in \mathrm{supp}(Q_0)$ with $z_1 < 1 < z_2$. Choose $0 < \lambda < 1$ with $\lambda z_1 + (1-\lambda) z_2 = 1$. So the two-point distribution

$$Q_1 \;=\; \begin{pmatrix} z_1 & z_2 \\ \lambda & 1-\lambda \end{pmatrix}$$

is feasible for (EP1) and $g(z) = \rho + \tau z$ on the support $\{z_1, z_2\}$ of Q_1. By (ii), Q_1 is also an optimal solution to (EP1). □

Now, for $I = [\,0\,,\,c\,]$ with a given finite $c > 1$ as above, we consider the n-dimensional cube I^n, where $n \in \mathbb{N}$, $n \geq 2$. For $Q_1, \ldots, Q_n \in \mathrm{Prob}(I)$ we denote by $\bigotimes_{i=1}^n Q_i$ the product of $Q_1, \ldots, Q_n$, i.e., the joint distribution of n stochastically independent random variables $X_1, \ldots, X_n$ with distributions $Q_1, \ldots, Q_n$, resp., which is a member of $\mathrm{Prob}(I^n)$. If $Q_1 = \ldots = Q_n = Q$, say, then we write Q^n instead of $\bigotimes_{i=1}^n Q$. Let $g : I^n \longrightarrow \mathbb{R}$ be a given nonnegative u.s.c. function on the cube I^n. Firstly, we will restrict to the case that g is *symmetric*, i.e., permutationally invariant,

$$g(z_{\pi(1)}, \ldots, z_{\pi(n)}) \;=\; g(z_1, \ldots, z_n)$$

for all $(z_1, \ldots, z_n) \in I^n$ and all permutations π of $1, \ldots, n$. We consider the following extremum problem.

$$\text{(EP2)} \qquad \begin{array}{l} \text{maximize} \quad G_n(Q) \;=\; \mathrm{E}_{Q^n}(g) \\ \text{s.t.} \quad Q \in \mathrm{Prob}(I)\,, \quad \mathrm{E}[Q] \;=\; 1\,. \end{array}$$

A *necessary* condition that a given feasible (for (EP2)) Q_0 is an optimal solution to (EP2) is that the directional derivatives of the objective function at Q_0 for all feasible directions are nonpositive, i.e.,

$$\lim_{\varepsilon \to 0+} \frac{1}{\varepsilon} \Big(G_n(\,(1-\varepsilon)Q_0 + \varepsilon Q\,) \;-\; G_n(Q_0) \Big) \;\leq\; 0 \quad \forall\, Q \in \mathrm{Prob}(I) \text{ with } \mathrm{E}[Q] = 1\,. \tag{5}$$

In fact, the directional derivative (the limit on the l.h.s. of (5)) exists and is given by

$$n \Big(\mathrm{E}_{Q_0^{n-1} \otimes Q}(g) \;-\; G_n(Q_0) \Big)\,, \tag{6}$$

which can be seen as follows. By the symmetry of g we get for any $0 < \varepsilon < 1$,

$$\begin{aligned} G_n(\,(1-\varepsilon)Q_0 + \varepsilon Q\,) \;&=\; \mathrm{E}_{(\,(1-\varepsilon)Q_0 + \varepsilon Q\,)^n}(g) \\ &=\; (1-\varepsilon)^n \mathrm{E}_{Q_0^n}(g) \;+\; \sum_{j=1}^{n-1} (1-\varepsilon)^j \varepsilon^{n-j} \tbinom{n}{j}\, \mathrm{E}_{Q_0^j \otimes Q^{n-j}}(g) \;+\; \varepsilon^n\, \mathrm{E}_{Q^n}(g)\,. \end{aligned}$$

Subtracting by $G_n(Q_0) = \mathrm{E}_{Q_0^n}(g)$, dividing by ε, and letting $\varepsilon \to 0$ the limit becomes

$$\Big(\lim_{\varepsilon \to 0+} \tfrac{(1-\varepsilon)^n - 1}{\varepsilon} \Big)\, G_n(Q_0) \;+\; \tbinom{n}{n-1}\, \mathrm{E}_{Q_0^{n-1} \otimes Q}(g)\,,$$

and formula (6) follows from $\lim_{\varepsilon\to 0+} \frac{(1-\varepsilon)^n-1}{\varepsilon} = -n$ and $\binom{n}{n-1} = n$. We note that the expectation in (6) can be written as

$$\mathrm{E}_{Q_0^{n-1}\otimes Q}(g) \;=\; \mathrm{E}_Q(g_{Q_0})\;, \text{ where} \tag{7}$$

$$g_{Q_0}(t) \;=\; \int_{I^{n-1}} g(z_1,\ldots,z_{n-1},t)\,\mathrm{d}Q_0^{n-1}(z_1,\ldots,z_{n-1}) \quad \forall\, t\in I\,. \tag{8}$$

Theorem 2.3 *Let $I = [0\,,\,c]$ with a real $c > 1$, let $n \in \mathbb{N}$, $n \geq 2$, and $g : I^n \longrightarrow \mathbb{R}$ be nonnegative, symmetric, and u.s.c. Then, for the extremum problem* (EP2) *above one has:*

(i) *There exists an optimal solution to* (EP2)*.*

(ii) *If Q_0 is an optimal solution to* (EP2) *then there exist $\rho \in [0\,,\,\infty)$ and $\tau \in \mathbb{R}$ such that the function g_{Q_0} from* (8) *satisfies*

$$g_{Q_0}(t) \leq \rho + \tau t \quad \forall\, t \in I \quad \textit{and} \quad g_{Q_0}(t) = \rho + \tau t \quad \forall\, t \in \operatorname{supp}(Q_0)\;,$$

and for the optimum value of (EP2) *we have $G_n(Q_0) = \rho + \tau$.*

Proof. By Lemma 2.1 the function $G : \mathrm{Prob}(I^n) \longrightarrow \mathbb{R}$, $G(R) = \mathrm{E}_R(g)$ is u.s.c. w.r.t. the weak topology on $\mathrm{Prob}(I^n)$. Since $Q \longrightarrow Q^n$ is a continuous mapping from $\mathrm{Prob}(I)$ to $\mathrm{Prob}(I^n)$ (w.r.t. the weak topologies), we see from $G_n(Q) = G(Q^n)$, $Q \in \mathrm{Prob}(I)$, that the objective function G_n of (EP2) is u.s.c. By the compactness of the feasible region of (EP2) statement (i) follows. To prove (ii), let Q_0 be an optimal solution to (EP2). Then, for any $Q \in \mathrm{Prob}(I)$ with $\mathrm{E}[Q] = 1$ and any $0 < \varepsilon < 1$ the convex combination $(1-\varepsilon)Q_0+\varepsilon Q$ is again feasible for (EP2) and hence $G_n(\,(1-\varepsilon)Q_0+\varepsilon Q\,) \leq G_n(Q_0)$. So the directional derivatives from (5) must be nonpositive and together with (6), (7), (8) we get $\mathrm{E}_Q(g_{Q_0}) \leq \mathrm{E}_{Q_0}(g_{Q_0})$. In other words, Q_0 is also an optimal solution to the extremum problem

$$\begin{aligned} &\text{maximize} \quad G_{Q_0}(Q) \;=\; \mathrm{E}_Q(g_{Q_0}) \\ &\text{s.t.} \quad Q \in \mathrm{Prob}(I)\,, \quad \mathrm{E}[Q] \;=\; 1\;, \end{aligned}$$

which is of type (EP1) considered in Theorem 2.2. To apply Theorem 2.2 we have to verify the conditions that g_{Q_0} is nonnegative real and u.s.c. As a nonnegative u.s.c. function g satisfies $0 \leq g(z) \leq v$ for all $z \in I^n$ for some finite $v > 0$, from which we immediately get $0 \leq g_{Q_0}(t) \leq v$ for all $t \in I$. Let $t_k \in I$ $(k \in \mathbb{N})$ be a sequence converging to some $t_0 \in I$. The u.s.c. property of g implies $\limsup_{k\to\infty} g(z_1,\ldots,z_{n-1},t_k) \leq g(z_1,\ldots,z_{n-1},t_0)$ for all $(z_1,\ldots,z_{n-1}) \in I^{n-1}$, and thus

$$\begin{aligned} &\limsup_{k\to\infty} g_{Q_0}(t_k) \;=\; \limsup_{k\to\infty} \int_{I^{n-1}} g(z_1,\ldots,z_{n-1},t_k)\,\mathrm{d}Q_0^{n-1}(z_1,\ldots,z_{n-1}) \\ &\leq\; \int_{I^{n-1}} \limsup_{k\to\infty} g(z_1,\ldots,z_{n-1},t_k)\,\mathrm{d}{Q_0}^{n-1}(z_1,\ldots,z_{n-1}) \\ &\leq\; \int_{I^{n-1}} g(z_1,\ldots,z_{n-1},t_0)\,\mathrm{d}{Q_0}^{n-1}(z_1,\ldots,z_{n-1}) \;=\; g_{Q_0}(t_0)\;. \end{aligned}$$

This proves the u.s.c. property of g_{Q_0}. Now we can apply Theorem 2.2, part (ii), which shows that Q_0 satisfies the condition in (ii) of Theorem 2.3, where $G_n(Q_0) = \rho+\tau$ follows from

$$G_n(Q_0) \;=\; \mathrm{E}_{Q_0}(g_{Q_0}) \;=\; \rho + \tau \mathrm{E}[Q_0] \;=\; \rho + \tau \;.$$

□

We consider still another extremum problem,

$$\text{(EP3)} \qquad \begin{array}{ll} \text{maximize} & \widetilde{G}_n(Q_1,\ldots,Q_n) \;=\; \mathrm{E}_{\bigotimes_{i=1}^n Q_i}(g) \\ \text{s.t.} & Q_1,\ldots,Q_n \in \mathrm{Prob}(I)\,, \quad \mathrm{E}[Q_i] \;=\; 1 \;,\; 1 \le i \le n\,. \end{array}$$

Again, the given function $g \;:\; I^n \longrightarrow \mathbb{R}$ is assumed to be nonnegative and u.s.c., but we do not impose here a symmetry condition on g.

Theorem 2.4 *Let $I = [0\,,\,c]$ with a real $c > 1$, let $n \in \mathbb{N}$, $n \ge 2$, and $g \;:\; I^n \longrightarrow \mathbb{R}$ be nonnegative and u.s.c. Then for the extremum problem* (EP3) *from above there exists an optimal solution $(Q_1^*,\ldots,Q_n^*)$ such that each Q_i^* $(1 \le i \le n)$ is a one- or two-point distribution.*

Proof. By Lemma 2.1 the function $G \;:\; \mathrm{Prob}(I^n) \longrightarrow \mathbb{R}$, $G(R) \;=\; \mathrm{E}_R(g)$ is u.s.c. w.r.t. the weak topology on $\mathrm{Prob}(I^n)$. Since $(Q_1,\ldots,Q_n) \longrightarrow \bigotimes_{i=1}^n Q_i$ is a continuous mapping from $(\mathrm{Prob}(I))^n$ to $\mathrm{Prob}(I^n)$ (w.r.t. the weak topologies), we see from $\widetilde{G}_n(Q_1,\ldots,Q_n) \;=\; G\Big(\bigotimes_{i=1}^n Q_i\Big)$ that the objective function $\widetilde{G}_n$ of (EP3) is u.s.c. By the compactness of the feasible region of (EP3) it follows that there exists an optimal solution $(Q_1^*,\ldots,Q_n^*)$ to (EP3). It remains to show that each Q_i^* can be chosen as a one- or two-point distribution. To this end let $(Q_1^*,\ldots,Q_n^*)$ be any optimal solution to (EP3). For a given $i \in \{1,\ldots,n\}$ we can write

$$\widetilde{G}_n(Q_1^*,\ldots,Q_n^*) \;=\; \int_I g_i^*\,\mathrm{d}Q_i^* \;, \quad \text{where}$$

$$g_i^*(t) \;=\; \int_{I^{n-1}} g(z_1,\ldots,z_{i-1},t,z_{i+1},\ldots,z_n)\,\mathrm{d}\Big(\bigotimes_{\substack{j=1\\ j\ne i}}^{n} Q_j^*\Big)(z_1,\ldots,z_{i-1},z_{i+1},\ldots,z_n)$$

for all $t \in I$. Along similar lines as in the proof of Theorem 2.3 one can see that g_i^* is a nonnegative real u.s.c. function on I. Considering the (EP1),

$$\begin{array}{ll} \text{maximize} & G_i^*(Q) \;=\; \mathrm{E}_Q(g_i^*) \\ \text{s.t.} & Q \in \mathrm{Prob}(I)\;, \quad \mathrm{E}[Q] = 1\;, \end{array}$$

we get from Theorem 2.2, part (iii), the existence of a one- or two-point distribution Q_i^{**} which is an optimal solution to that (EP1). Observing that

$$\begin{aligned} \mathrm{E}_{Q_i^{**}}(g_i^*) &= \widetilde{G}_n(Q_1^*,\ldots,Q_{i-1}^*,Q_i^{**},Q_{i+1}^*,\ldots,Q_n^*)\;, \\ \mathrm{E}_{Q_i^{*}}(g_i^*) &= \widetilde{G}_n(Q_1^*,\ldots,Q_n^*)\;, \end{aligned}$$

we have

$$\widetilde{G}_n(Q_1^*,\ldots,Q_{i-1}^*,Q_i^{**},Q_{i+1}^*,\ldots,Q_n^*) \;\ge\; \widetilde{G}_n(Q_1^*,\ldots,Q_n^*)\;,$$

and, since $(Q_1,^*, \dots, Q_n^*)$ is an optimal solution to our original (EP3), the inequality is actually an equality. So we have shown that if $(Q_1^*, \dots, Q_n^*)$ is an optimal solution to (EP3) then for any given $i \in \{1, \dots, n\}$ there is a one- or two-point distribution Q_i^{**} such that $(Q_1^*, \dots, Q_{i-1}^*, Q_i^{**}, Q_{i+1}^*, \dots, Q_n^*)$ is also an optimal solution to (EP3). Applying this successively for $i = 1, \dots, n$ the result follows. $\square$

We close this section by adding a technical lemma, which will enable us for certain extremum problems below to cut down probability distributions on the real half line $[0, \infty)$ to a compact interval $I = [0, c]$ (with a $c > 1$) as considered in the above extremum problems. By $\mathrm{Prob}([0, \infty))$ we denote the set of all Borel probability measures on $[0, \infty)$. A function $g : [0, \infty)^n \longrightarrow \mathbb{R}$ is said to be *nondecreasing* iff $x, y \in [0, \infty)^n$ and $x \le y$ (componentwise) imply $g(x) \le g(y)$.

Lemma 2.5 *Let $g : [0, \infty)^n \longrightarrow \mathbb{R}$ be nonnegative and nondecreasing (and Borel-measurable). Assume that there is a real constant $c > 1$ such that*

$$g(z_1, \dots, z_n) = g(h_c(z_1), \dots, h_c(z_n)) \quad \forall\, z_1, \dots, z_n \in [0, \infty) ,$$
$$\textit{where we denote} \quad h_c : [0, \infty) \longrightarrow [0, c] , \quad h_c(t) = \min\{t, c\} .$$

Let $Q_1, \dots, Q_n \in \mathrm{Prob}([0, \infty))$ with $\mathrm{E}[Q_i] \le 1$ $(1 \le i \le n)$ be given. Denote by $Q_i^{h_c} \in \mathrm{Prob}([0, c])$ the distribution of h_c under Q_i $(1 \le i \le n)$, by $\delta_c \in \mathrm{Prob}([0, c])$ the one-point distribution at c, and define $Q_{i,c} \in \mathrm{Prob}([0, c])$ $(1 \le i \le n)$ by

$$Q_{i,c} = (1 - \lambda_i)Q_i^{h_c} + \lambda_i \delta_c , \textit{ where } \lambda_i = (1 - \mathrm{E}[Q_i^{h_c}])/(c - \mathrm{E}[Q_i^{h_c}]) , \quad (1 \le i \le n).$$

Then: $\mathrm{E}[Q_{i,c}] = 1$ $(1 \le i \le n)$ *and*

$$\int_{[0, \infty)^n} g \,\mathrm{d}\bigotimes_{i=1}^n Q_i \le \int_{[0, c]^n} g \,\mathrm{d}\bigotimes_{i=1}^n Q_{i,c} .$$

Proof. Obviously, $\mathrm{E}[Q_i^{h_c}] \le \mathrm{E}[Q_i] \le 1$, hence $0 \le \lambda_i < 1$, and the $Q_{i,c}$ are in fact probability distributions on $[0, c]$. Also, by the definition of λ_i,

$$\mathrm{E}[Q_{i,c}] = (1 - \lambda_i)\mathrm{E}[Q_i^{h_c}] + \lambda_i c = 1 .$$

Clearly, for each $i = 1, \dots, n$, the distribution $Q_i^{h_c}$ is stochastically smaller than or equal to $Q_{i,c}$, and since g is nondecreasing we have (cf. [4], Theorem 4.B.10, part (b)),

$$\int_{[0, c]^n} g \,\mathrm{d}\bigotimes_{i=1}^n Q_i^{h_c} \le \int_{[0, c]^n} g \,\mathrm{d}\bigotimes_{i=1}^n Q_{i,c} .$$

The integral on the l.h.s. equals

$$\int_{[0, \infty)^n} g(h(z_1), \dots, h(z_n)) \,\mathrm{d}\bigotimes_{i=1}^n Q_i ,$$

which is the same as $\int_{[0,\infty)^n} g \,\mathrm{d} \bigotimes_{i=1}^n Q_i$ because of $g(h_c(z_1), \dots, h_c(z_n)) = g(z_1, \dots, z_n)$ by the assumed condition on g. $\square$

Remark. If in the lemma the Q_i are all equal, $Q_1 = \ldots = Q_n = Q$, say, then the $Q_{i,c}$ are all equal as well, $Q_{1,c} = \ldots = Q_{n,c} = Q_c$, say.

3 An important statistic

We return to the nonparametric statistical model of Section 1, i.e., we suppose that the values $x_1, \ldots, x_n$ of n i.i.d. nonnegative real random variables $X_1, \ldots, X_n$ are observed whose distribution is unknown, and we consider the one-sided hypotheses (1) about the expectation $\mu = \mathrm{E}(X_i)$. Let $0 < \alpha < 1$ be a given level. For $n = 1$, it was shown by Wang & Zhao in [1], pp. 82-83, that $\phi^* = \mathbb{1}_{[1/\alpha,\infty)}$ is UMP level α among all nonrandomized level α tests. Also they showed that this is no longer true if one admits *randomized* tests. For $n = 2$, they constructed a UMP test within the class of all *strongly monotone* and symmetric level α tests (p. 84 of [1]), which is a highly nontrivial result. We employ the notion of a *strongly monotone* test from [1] for arbitrary n. A test $\phi : [0, \infty)^n \longrightarrow [0, 1]$ is said to be strongly monotone (nondecreasing) iff, firstly, $\phi(x) > 0$, $x \le z$, $x \ne z$ imply $\phi(z) = 1$ and, secondly, $\phi(x) < 1$, $x \ge z$, $x \ne z$ imply $\phi(z) = 0$. Clearly, strong monotonicity of ϕ implies monotonicity of ϕ, i.e., ϕ is nondecreasing ($x \le z$ implies $\phi(x) \le \phi(z)$). For a nonrandomized test ϕ strong monotonicity of ϕ is the same as monotonicity of ϕ (ϕ is nondecreasing). So, the class of all strongly monotone symmetric level α tests for H_0 is slightly larger than the class of all nonrandomized monotone symmetric level α tests for H_0.
The following statistic introduced by Wang & Zhao in [1] for $n = 2$ is meaningful for the testing problem (1). We introduce that statistic for arbitrary n,

$$W(x) = \sup_{H_0} P\Big(X_{(i)} \ge x_{(i)} \ \ \forall\, i = 1, \ldots, n \Big) , \quad x = (x_1, \ldots, x_n) \in [0, \infty)^n , \tag{9}$$

where $X_{(1)} \le X_{(2)} \le \ldots \le X_{(n)}$ denote the order statistics of the random variables $X_1, \ldots, X_n$ and $x_{(1)} \le x_{(2)} \le \ldots \le x_{(n)}$ denote the nondecreasingly rearranged components of $x = (x_1, \ldots, x_n)$. Our notation in (9) is somewhat loose; the sup on the r.h.s. of (9) is to mean that one takes into account all possible distributions $Q = P^{X_i}$ of n i.i.d. nonnegative real random variables $X_1, \ldots, X_n$ with $\mathrm{E}(X_i) \le 1$ and takes the supremum over of all resulting probabilities $P(X_{(i)} \ge x_{(i)} \ \forall\, i = 1, \ldots, n)$, for fixed $x = (x_1, \ldots, x_n) \in [0, \infty)^n$. That is, the value $W(x)$ is the optimum value of the extremum problem,

$$\begin{aligned} &\text{maximize} \quad G_{x,n}(Q) = Q^n(C_x) \\ &\text{s.t.} \quad Q \in \mathrm{Prob}([0,\infty)) , \quad \mathrm{E}[Q] \le 1 , \end{aligned} \tag{10}$$

where $x = (x_1, \ldots, x_n) \in [0, \infty)^n$ is given and C_x denotes the set

$$C_x = \{ z = (z_1, \ldots, z_n) \in [0, \infty)^n : z_{(i)} \ge x_{(i)} \ \ \forall\, i = 1, \ldots, n \} . \tag{11}$$

Obviously, the function W on $[0, \infty)^n$ is symmetric and nonincreasing. Below we will see that W is continuous (see Lemma 3.5). So the following result is obtained by analogous arguments as those in [1], Corollary A.1 and its proof on pp. 91-92.

Lemma 3.1 *Consider the (nonrandomized) test $\psi^*_\alpha = \mathbb{1}_{\{W \le \alpha\}}$. Then, for any strongly monotone symmetric level α test ϕ for H_0 we have $\phi \le \psi^*_\alpha$. In particular, if ψ^*_α keeps level α on H_0 then ψ^*_α is UMP within the class of all strongly monotone symmetric level α tests for H_0 vs. H_1.* □

As it will turn out, the test ψ^*_α keeps level α on H_0 only for $n \le 2$; moreover, for $n \ge 3$ no UMP strongly monotone symmetric level α test for H_0 vs. H_1 exists. To see this we need a more explicit description of the statistic W than that from the definition (9) or (10), resp. For $n = 1$ we have $C_x = [\,x\,,\,\infty)$, i.e.,

$$W(x) \;=\; \sup\Big\{\, Q(\,[\,x\,,\,\infty)\,) \,:\, Q \in \mathrm{Prob}(\,[\,0\,,\,\infty)\,)\;,\; \mathrm{E}[Q] \le 1 \Big\}\;,$$

from which one easily obtains (using the Tchebychev-Markov inequality in case $x > 1$),

$$W(x) \;=\; \begin{cases} 1 & \text{, if } \; x \le 1 \\ 1/x & \phantom{\text{, if }}\; x > 1 \end{cases}\;.$$

So for $n = 1$ we have $\psi^*_\alpha \;=\; \mathbb{1}_{[1/\alpha\,,\infty)}$, which keeps level α on H_0, again by the Tchebychev-Markov inequality. Now let $n \ge 2$. From the definition (9) or (10) we have trivially

$$W(x) \;=\; 1 \quad \text{for all } x = (x_1, \ldots, x_n) \in [\,0\,,\,\infty)^n \text{ with } x_{(n)} \le 1\;.$$

In the nontrivial case that $x_{(n)} > 1$ the extremum problem (10) can be brought into an (EP2) problem from Section 2 by using Lemma 2.5. The function $g \;=\; \mathbb{1}_{C_x}$ satisfies the assumptions of Lemma 2.5 with $c \;=\; x_{(n)}$. By that lemma the optimum value $W(x)$ of the extremum problem (10) remains the same when we restrict in (10) to probability distributions Q on $[\,0\,,\,x_{(n)}]$ with $\mathrm{E}[Q] = 1$. So we can apply Theorem 2.3 on an (EP2), which yields the following (cp. [1], Lemma A1 for the case $n = 2$).

Lemma 3.2 *Let $n \ge 2$. Then, for a given $x = (x_1, \ldots, x_n) \in [\,0\,,\,\infty)^n$ with $x_{(n)} > 1$, the maximum value $W(x)$ of (10) is attained by some Q_x for which $\mathrm{E}[Q_x] = 1$ and*

$$\mathrm{supp}(Q_x) \;=\; \{\,0, x_{(1)}, \ldots, x_{(n)}\} \quad \textit{or} \quad \mathrm{supp}(Q_x) \;=\; \{\,x_{(1)}, \ldots, x_{(n)}\}\;.$$

Proof. Theorem 2.3 with $c = x_{(n)}$, $g \;=\; \mathbb{1}_{C_x}$, and C_x from (11), yields the existence of an optimal solution $Q_x \in \mathrm{Prob}(\,[\,0\,,\,x_{(n)}]\,)$ to (10), and the existence of $\rho \ge 0$, $\tau \in \mathbb{R}$ such that

$$g_{Q_x}(t) \le \rho + \tau t \;\forall\; t \in [\,0\,,\,x_{(n)}] \quad \text{and} \quad g_{Q_x}(t) = \rho + \tau t \;\forall\; t \in \mathrm{supp}(Q_x), \tag{12}$$

$$\text{where} \quad g_{Q_x}(t) = Q_x^{n-1}(\,C_x(t)\,)\;,\; C_x(t) = \{s \in [\,0\,,\,x_{(n)}]^{n-1} \,:\, (s,t) \in C_x\,\}\,. \tag{13}$$

The distribution Q_x must assign positive probability to the value $x_{(n)}$, since otherwise we would have $Q_x^n(C_x) \;=\; 0$ which cannot be true (*any* probability distribution $Q \in \mathrm{Prob}([\,0\,,\,x_{(n)}])$ with $\mathrm{E}[Q] = 1$ which assigns positive probability to $x_{(n)}$ has $Q^n(C_x) > 0$). Clearly, the sets $C_x(t)$ are nondecreasing in $t \in [\,0\,,\,x_{(n)}]$, and hence the function g_{Q_x} is nondecreasing. Also, g_{Q_x} is constant on the subinterval $0 \le t < x_{(1)}$ (if this is nonempty), and on the subinterval $x_{(k)} \le t < x_{(k+1)}$ (if this is nonempty) for each $k = 1, \ldots, n-1$. More explicitly, it is easily seen that

$$C_x(t) \;=\; \emptyset\;, \quad \text{if } 0 \le t < x_{(1)}\;; \tag{14}$$

$$C_x(t) \;=\; \Big\{s \in [\,0\,,\,x_{(n)}]^{n-1} \,:\, s_{(i)} \ge x_{(i)}\;(1 \le i \le k-1)\,,\; s_{(i)} \ge x_{(i+1)}\;(k \le i \le n-1)\Big\}\;, \tag{15}$$

$$\text{if } x_{(k)} \le t < x_{(k+1)} \text{ and } 1 \le k \le n-1\;;$$

$$C_x(x_{(n)}) \;=\; \Big\{s \in [\,0\,,\,x_{(n)}]^{n-1} \,:\, s_{(i)} \ge x_{(i)}\;(1 \le i \le n-1)\Big\}\;. \tag{16}$$

Since g_{Q_x} is nondecreasing we must have $\tau \geq 0$ in (12) (otherwise $x_{(n)}$ would be the only support point of Q_x contradicting $\mathrm{E}[Q_x] = 1$ and $x_{(n)} > 1$).
Suppose $\tau = 0$. Then, $g_{Q_x}(t) \leq \rho$ for all $t \in [\,0\,,\,x_{(n)}]$ and the equality $g_{Q_x}(t) = \rho$ holds for all $t \in \mathrm{supp}(Q_x)$. The point s in $\mathbb{R}^{n-1}$ having all components equal to $x_{(n)}$ belongs to $C_x(x_{(n)})$ by (16) and hence $\rho = g_{Q_x}(x_{(n)}) \geq \Big(Q_x(\{x_{(n)}\}) \Big)^{n-1} > 0$. So we have $\rho > 0$ and, because of $g_{Q_x}(t) = 0$ if $0 \leq t < x_{(1)}$, we have $\mathrm{supp}(Q_x) \subset [\,x_{(1)},\,x_{(n)}]$, in particular $x_{(1)} < x_{(n)}$. Let t_0 be the smallest support point of Q_x. There is a $k \in \{1, \ldots, n-1\}$ with $x_{(k)} \leq t_0 < x_{(k+1)}$. The interval $I_k = [\,x_{(k)}, x_{(k+1)})$ has positive Q_x-probability, since it contains the smallest support point t_0. Consider the set

$$D \;=\; \{\, s = (s_1, \ldots, s_{n-1}) \,:\, s_i \in I_k \;(1 \leq i \leq k)\,,\; s_i = x_{(n)} \;(k+1 \leq i \leq n-1)\,\}\;.$$

Then $Q_x^{n-1}(D) = (Q_x(I_k))^k (Q_x(\{x_{(n)}\}))^{n-1-k} > 0$. For each $s \in D$ we see from (15) that $s \in C_x(x_{(k+1)})$, but $s \notin C_x(x_{(k)})$. Hence $D \subset C_x(x_{(k+1)}) \setminus C_x(x_{(k)})$ and $Q_x^{n-1}(D) \leq g_{Q_x}(x_{(k+1)}) - g_{Q_x}(x_{(k)})$, in particular $g_{Q_x}(x_{(k)}) < g_{Q_x}(x_{(k+1)})$. On the other hand we have $\rho = g_{Q_x}(t_0) = g_{Q_x}(x_{(k)})$ and $\rho = g_{Q_x}(x_{(n)})$, hence $g_{Q_x}(x_{(k)}) = g_{Q_x}(x_{(k+1)})$ and thus a contradiction, i.e., $\tau = 0$ cannot hold.
So we have $\tau > 0$. By (12) the increasing straight line $\rho + \tau t$ can touch the graph of the nondecreasing step-function $g_{Q_x}(t)$ only at the very left points of its plateaus, the abscissa of which are $0, x_{(1)}, \ldots, x_{(n)}$. Hence

$$\mathrm{supp}(Q_x) \;\subset\; \{\,0\,,\,x_{(1)}\,,\,\ldots\,,\,x_{(n)}\,\}\;. \tag{17}$$

It remains to show that each $x_{(k)}$ $(k = 1, \ldots, n)$ is in fact a support point of Q_x. We know that this is true for $k = n$. Suppose that for some $k \in \{2, \ldots, n\}$ we have $x_{(k)} \in \mathrm{supp}(Q_x)$, but $x_{(k-1)} \notin \mathrm{supp}(Q_x)$, and thus, in particular, $x_{(k-1)} < x_{(k)}$. Clearly, the Q_x^{n-1}-probability of a set $C_x(t)$ remains the same when restrict to the subset of those points s in $C_x(t)$ with components in $\mathrm{supp}(Q_x)$. By (15) we have

$$\begin{aligned} C_x(x_{(k)}) \;&=\; \{\, s \,:\, s_{(i)} \geq x_{(i)} \;(1 \leq i \leq k-1)\,,\; s_{(i)} \geq x_{(i+1)} \;(k \leq i \leq n-1)\,\}\;, \\ C_x(x_{(k-1)}) \;&=\; \{\, s \,:\, s_{(i)} \geq x_{(i)} \;(1 \leq i \leq k-2)\,,\; s_{(i)} \geq x_{(i+1)} \;(k-1 \leq i \leq n-1)\,\}\;. \end{aligned}$$

But for s with components in $\mathrm{supp}(Q_x)$ we get from (17) and $x_{(k-1)} \notin \mathrm{supp}(Q_x)$ that the inequality $s_{(k-1)} \geq x_{(k-1)}$ is equivalent to $s_{(k-1)} \geq x_{(k)}$, which shows that the sets $C_x(x_{(k)})$ and $C_x(x_{(k-1)})$ contain the same set of points s with components in $\mathrm{supp}(Q_x)$. Hence it follows that

$$g_{Q_x}(x_{(k)}) \;=\; Q_x^{n-1}(C_x(x_{(k)})) \;=\; Q_x^{n-1}(C_x(x_{(k-1)})) \;=\; g_{Q_x}(x_{(k-1)})\;,$$

and by (12),

$$\rho + \tau x_{(k)} \;=\; g_{Q_x}(x_{(k)}) \;=\; g_{Q_x}(x_{(k-1)}) \;\leq\; \rho + \tau x_{(k-1)}\;,$$

which is a contradiction to $\tau > 0$ and $x_{(k)} > x_{(k-1)}$. Hence we must have $x_{(k-1)} \in \mathrm{supp}(Q_x)$ as well. This completes the proof of our lemma. □

From Lemma 3.2 it follows that the value $W(x)$ (in the nontrivial case $x_{(n)} > 1$) equals the maximum of $Q^n(C_x)$ taken over all discrete probability distributions Q with

$$\{x_{(1)}, \ldots, x_{(n)}\} \subset \mathrm{supp}(Q) \subset \{0, x_{(1)}, \ldots, x_{(n)}\} \quad \text{and} \quad \mathrm{E}[Q] = 1\;.$$

For $n = 2$ that maximum was explicitly determined in [1], Lemma A.1, giving the following.

Lemma 3.3 (Wang & Zhao)
Let $n = 2$. Then, for all $x = (x_1, x_2) \in [\,0\,,\,\infty)^2$, we have

$$W(x) \;=\; \begin{cases} 1 & ,\ \textit{if}\ x_{(2)} \le 1 \\ \dfrac{1}{x_{(1)}(2x_{(2)} - x_{(1)})} & ,\ \textit{if}\ x_{(2)} > 1\,,\ x_{(1)} > x_{(2)} - \sqrt{x_{(2)}^2 - x_{(2)}} \\ 1 - \dfrac{(x_{(2)}-1)^2}{(x_{(2)} - x_{(1)})^2} & ,\ \textit{if}\ x_{(2)} > 1\,,\ x_{(1)} \le x_{(2)} - \sqrt{x_{(2)}^2 - x_{(2)}} \end{cases}$$

□

For $n \ge 3$ no explicit description of W seems to be possible. However, an alternative (implicit) representation of the statistic will be useful; in particular, it will help to prove continuity of W (for arbitrary n) below.

Lemma 3.4 *Let $U_{(1)} \le U_{(2)} \le \ldots \le U_{(n)}$ be the order statistics of n i.i.d. standard uniformly distributed (on the interval $(\,0\,,\,1\,)$) random variables $U_1, \ldots, U_n$. Then, for any $x = (x_1, \ldots, x_n) \in [\,0\,,\,\infty)^n$ with $x_{(n)} > 1$, we have*

$$W(x) \;=\; \max\Big\{ P(\,U_{(1)} \ge u_1\,,\,\ldots\,,\,U_{(n)} \ge u_n)\ :\ 0 \le u_1 \le \ldots \le u_n \le 1\,,\ \sum_{i=1}^{n} u_i(x_{(i)} - x_{(i-1)}) = x_{(n)} - 1 \Big\}\,,$$

where $x_{(0)} = 0$. Also, as we already know, $W(x) = 1$ whenever $x_{(n)} \le 1$.

Proof. Firstly, we note that if $X_1, \ldots, X_n$ are i.i.d. real random variables, F denoting the (right continuous) c.d.f. of X_i, and if $a_1, \ldots, a_n \in \mathbb{R}$, then

$$P\Big(\,X_{(i)} \ge a_i\ \forall\ i = 1, \ldots, n\Big) \;=\; P\Big(\,U_{(i)} \ge F_\ell(a_i)\ \forall\ i = 1, \ldots, n\Big)\,, \tag{18}$$

where $U_1, \ldots, U_n$ are i.i.d. standard uniform random variables and $F_\ell(t) = \lim_{v \to t,\, v<t} F(v)$ for $t \in \mathbb{R}$, (the *left* continuous c.d.f. of X_i). Eq. (18) can be seen as follows. Consider the pseudo-inverse of F,

$$F^-(u) \;=\; \min\{\,t \in \mathbb{R}\ :\ F(t) \ge u\,\} \quad \forall\ u \in (\,0\,,\,1\,)\,.$$

Then, as it is well-known, we may write $X_i = F^-(U_i)$ $(1 \le i \le n)$, with some i.i.d. standard uniform random variables $U_1, \ldots, U_n$. It is easily seen that for any $a \in \mathbb{R}$ and any $u \in (\,0\,,\,1\,)$ the following implications are true,

$$u > F_\ell(a) \implies F^-(u) \ge a \implies u \ge F_\ell(a)\,.$$

This gives, observing that $X_{(i)} = F^-(U_{(i)})$ $(1 \le i \le n)$,

$$\Big\{U_{(i)} > F_\ell(a_i)\ \forall\ i\Big\} \subset \Big\{X_{(i)} \ge a_i\ \ \forall\ i\Big\} \subset \Big\{U_{(i)} \ge F_\ell(a_i)\ \forall\ i\Big\},$$

and since the event $U_{(i)} = F_\ell(a_i)$ has probability zero for all i, we conclude (18).

Now let $x = (x_1, \ldots, x_n) \in [\,0\,,\,\infty)^n$ with $x_{(n)} > 1$ be given. Denote by $m(x)$ the maximum on the r.h.s. of the asserted formula of the lemma. Note that the term 'maximum' (instead of 'supremum') is correct, since the probability $P(U_{(i)} \geq u_i \;\forall\; i = 1, \ldots, n)$ is a continuous function of $u = (u_1, \ldots, u_n)$ and the set of u's over which the supremum is taken is a compact (and nonempty) set. We have to show that $m(x) = W(x)$.
By definition of $W(x)$ through (9) or (10) and by Lemma 3.2, $W(x)$ is the maximum of all probabilities

$$P\Big(\, X_{(i)} \geq x_{(i)} \;\forall\; i = 1, \ldots, n\,\Big)\;,$$

taken over all i.i.d. random variables $X_1, \ldots, X_n$ with values in $\{0, x_{(1)}, \ldots, x_{(n)}\}$ and with $\mathrm{E}(X_i) = 1$. For any such random variables $X_1, \ldots, X_n$ denote by F the (right continuous) c.d.f. of X_i and by F_ℓ its left continuous version. A well-known formula for the expectation of a nonnegative random variable gives $1 = \mathrm{E}(X_1) \;=\; \int_0^\infty (1 - F_\ell(t))\mathrm{d}t\,$.
Since X_1 has its values in $\{0, x_{(1)}, \ldots, x_{(n)}\}$ we can write

$$1 - F_\ell(t) \;=\; \sum_{i=1}^{n} (1 - F_\ell(x_{(i)}))\, \mathbb{1}_{(\,x_{(i-1)}, x_{(i)}\,]}(t) \quad \forall\, t > 0\,,$$

and hence

$$1 = \sum_{i=1}^{n} (1 - F_\ell(x_{(i)})\,(x_{(i)} - x_{(i-1)})\,,\ \text{i.e.,}\ \sum_{i=1}^{n} F_\ell(x_{(i)})\,(x_{(i)} - x_{(i-1)}) \;=\; x_{(n)} - 1\;.$$

So, choosing $u_i = F_\ell(x_{(i)})$ $(1 \leq i \leq n)$, we have $0 \leq u_1 \leq \ldots \leq u_n \leq 1$, $\sum_{i=1}^n u_i(x_{(i)} - x_{(i-1)}) = x_{(n)} - 1$, and, together with (18),

$$P(\, X_{(i)} \geq x_{(i)} \;\forall\; i = 1, \ldots, n) \;\;=\;\; P(\, U_{(i)} \geq u_i \;\forall\; i = 1, \ldots, n)\;,$$

which proves $W(x) \leq m(x)$.
To prove the reverse inequality let $0 \leq u_1 \leq \ldots \leq u_n \leq 1$ with $\sum_{i=1}^n u_i(x_{(i)} - x_{(i-1)}) = x_{(n)} - 1$ be given. Then, obviously, the function

$$F(t) \;=\; \sum_{i=1}^{n} u_i \mathbb{1}_{[\,x_{(i-1)}, x_{(i)}\,)}(t) \;+\; \mathbb{1}_{[\,x_{(n)}, \infty)}(t)\;, \quad t \in \mathbb{R}\,,$$

is the (right continuous) c.d.f. of some probability distribution Q with support in the set $\{0, x_{(1)}, \ldots, x_{(n)}\}$. The left continuous version F_ℓ satisfies $F_\ell(x_{(i)}) \;\leq\; u_i$ $(1 \leq i \leq n)$, and for the expectation of Q we have

$$\mathrm{E}[Q] \;=\; \int_0^\infty (1 - F(t))\mathrm{d}t \;\;=\; \sum_{i=1}^{n} (1 - u_i)(x_{(i)} - x_{(i-1)}) \;=\; 1\;.$$

Let $X_1, \ldots, X_n$ be i.i.d. Q-distributed random variables. From (18) together with $F_\ell(x_{(i)}) \leq u_i$ $(1 \leq i \leq n)$ we get

$$P(\, X_{(i)} \geq x_{(i)} \;\forall\; i = 1, \ldots, n) \;\;\geq\;\; P(\, U_{(i)} \geq u_i \;\forall\; i = 1, \ldots, n)\;,$$

which shows $W(x) \geq m(x)$. □

Lemma 3.5 *The function W is continuous on $[0, \infty)^n$.*

Proof. Since W is constantly equal to 1 on the cube $C = [0, 1]^n$, it suffices to prove
(i) continuity of W on the set $C^c = [0, \infty)^n \setminus C$, and
(ii) $W(x) \to 1$ whenever $x \in C^c$ approaches the boundary of C.
In view of Lemma 3.4, the following abbreviations will be convenient.

$$f(u) \;=\; P(\, U_{(i)} \geq u_i \;\forall\; i = 1, \ldots, n \,)\;, \quad \text{for } u = (u_1, \ldots, u_n) \in \mathbb{R}^n\,,$$

which is a continuous function on $\mathbb{R}^n$,

$$D \;=\; \{\, u \in \mathbb{R}^n \,:\, 0 \leq u_1 \leq \ldots \leq u_n \leq 1 \,\}\;,$$
$$D(x) \;=\; \Big\{\, u \in D \,:\, \sum_{i=1}^{n} u_i (x_{(i)} - x_{(i-1)}) = x_{(n)} - 1 \Big\} \quad \text{for } x \in C^c\,.$$

Clearly, D and $D(x)$ (for $x \in C^c$) are nonempty compact and convex sets. It is not hard to see that the vertices of $D(x)$ are given by

$$(\underbrace{\xi_i, \ldots, \xi_i}_{i \text{ times}}, \underbrace{1, \ldots, 1}_{n-i \text{ times}}) \quad \text{for } 1 \leq i \leq n \text{ with } x_{(i)} \geq 1\,, \text{ where } \xi_i = \tfrac{x_{(i)}-1}{x_{(i)}}\;;$$

$$(\underbrace{0, \ldots, 0}_{i \text{ times}}, \underbrace{\xi_{ij}, \ldots, \xi_{ij}}_{j-i \text{ times}}, \underbrace{1, \ldots, 1}_{n-j \text{ times}}) \quad \text{for } 1 \leq i < j \leq n \text{ with } x_{(i)} < 1 < x_{(j)},$$
$$\text{where } \xi_{ij} = (x_{(j)} - 1)/(x_{(j)} - x_{(i)})\,.$$

By Lemma 3.5 we have

$$W(x) \;=\; \max_{u \in D(x)} f(u) \quad \text{for } x \in C^c. \tag{19}$$

<u>Ad (i):</u> Let $x^{(0)} \in C^c$ and $x^{(k)} \in C^c$ ($k \in \mathbb{N}$) be a sequence with $\lim_{k\to\infty} x^{(k)} = x^{(0)}$. For each $k \in \mathbb{N}$ let $u^{(k)} \in D(x^{(k)})$ with $W(x^{(k)}) = f(u^{(k)})$. Denote $a = \limsup_{k\to\infty} f(u^{(k)})$. Choose an infinite subsequence $u^{(k)}$, $k \in K \subset \mathbb{N}$, such that $f(u^{(k)})$ ($k \to \infty$, $k \in K$) converges to a, and due to the compactness of D choose a further convergent subsequence $u^{(k)}$, $k \in K' \subset K$, with some limit $u^{(0)} \in D$. From $\sum_{i=1}^n u_i^{(k)}(x_{(i)}^{(k)} - x_{(i-1)}^{(k)}) = x_{(n)}^{(k)} - 1$ for all $k \in K'$ we get by taking the limits, $\sum_{i=1}^n u_i^{(0)}(x_{(i)}^{(0)} - x_{(i-1)}^{(0)}) = x_{(n)}^{(0)} - 1$, i.e., $u^{(0)} \in D(x^{(0)})$. Hence by continuity of f and by (19),

$$a \;=\; \lim_{k\to\infty,\, k \in K'} f(u^{(k)}) \;=\; f(u^{(0)}) \;\leq\; W(x^{(0)})\;,$$

which shows that $\limsup_{k\to\infty} W(x^{(k)}) \leq W(x^{(0)})$. It remains to prove the inequality $\liminf_{k\to\infty} W(x^{(k)}) \geq W(x^{(0)})$. To this end it suffices to show that for any given point $u^{(0)} \in D(x^{(0)})$ one can find a sequence $u^{(k)} \in D(x^{(k)})$ ($k \in \mathbb{N}$) converging to $u^{(0)}$. Then, taking $u^{(0)} \in D(x^{(0)})$ with $f(u^{(0)}) = W(x^{(0)})$, we get $\lim_{k\to\infty} f(u^{(k)}) = f(u^{(0)}) = W(x^{(0)})$ and $f(u^{(k)}) \leq W(x^{(k)})$ for all k, from which $W(x^{(0)}) \leq \liminf_{k\to\infty} W(x^{(k)})$ follows. Now, since $D(x^{(k)})$ is a convex set for each k, the set of all limits of convergent sequences $u^{(k)} \in D(x^{(k)})$ ($k \in \mathbb{N}$) is again a convex set, and it is a subset of $D(x^{(0)})$ (see our arguments above). Thus, that set of limits coincides with $D(x^{(0)})$ iff each vertex of $D(x^{(0)})$ is a limit of some sequence $u^{(k)} \in D(x^{(k)})$. But this is readily verified in view of our description of the vertices of a set $D(x)$ from above. In fact, the vertices of

$D(x^{(0)})$ are the limits of the corresponding vertices of $D(x^{(k)})$ with one exception: If $x^{(0)}_{(i)} = 1$ for some $i \in \{1, \dots, n-1\}$ and $x^{(k)}_{(i)} < 1$ for infinitely many k, then the vertex $(0, \dots, 0, 1, \dots, 1)$ (where 0 appears i times) of $D(x^{(0)})$ is the limit of $u^{(k)} \in D(x^{(k)})$ for $k \to \infty$, where

$$u^{(k)} \;=\; (\underbrace{\xi_i^{(k)}, \dots, \xi_i^{(k)}}_{i \text{ times}}, 1, \dots, 1) \quad \text{with } \xi_i^{(k)} = \tfrac{x^{(k)}_{(i)} - 1}{x^{(k)}_{(i)}}\,, \text{ if } x^{(k)}_{(i)} \geq 1\,,$$

$$u^{(k)} \;=\; (\underbrace{0, \dots, 0}_{i \text{ times}}, \xi_{in}^{(k)}, \dots, \xi_{in}^{(k)}) \quad \text{with } \xi_{in}^{(k)} \;=\; \tfrac{x^{(k)}_{(n)} - 1}{x^{(k)}_{(n)} - x^{(k)}_{(i)}}\,, \text{ if } x^{(k)}_{(i)} < 1\,.$$

<u>Ad (ii):</u> Let $x^{(k)} \in C^c$ ($k \in \mathbb{N}$) converge to some boundary point $x^{(0)}$ of C, hence in particular $\lim_{k\to\infty} x^{(k)}_{(n)} = 1$. For each k consider the point $u^{(k)}$ with components all equal to $(x^{(k)}_{(n)} - 1)/x^{(k)}_{(n)}$, which belongs to $D(x^{(k)})$. Obviously, $\lim_{k\to\infty} u^{(k)} = (0, \dots, 0)$. Again by (19), $W(x^{(k)}) \geq f(u^{(k)})$ for all k and thus, by continuity of f,

$$\liminf_{k\to\infty} W(x^{(k)}) \;\geq\; f(0, \dots, 0) \;=\; 1\;,$$

i.e., $\lim_{k\to\infty} W(x^{(k)}) = 1$. □

4 The Wang-Zhao result for $n = 2$

Here we present an alternative proof of [1], Theorem 5.2.

Theorem 4.1 (Wang & Zhao)
Let $n = 2$. Then, for any given $\alpha \in (\,0\,,\,1\,)$, we have

$$\sup_{Q \in H_0} Q^2(\,W \leq \alpha\,) \;=\; \alpha\;,$$

where $Q \in H_0$ means that $Q \in$ Prob$(\,[\,0\,,\,\infty)\,)$ with E$[Q] \leq 1$. *So, by* Lemma 3.1*, the test $\psi^*_\alpha = \mathbb{1}_{\{W \leq \alpha\}}$ is UMP among all strongly monotone symmetric level α tests for H_0 vs. H_1.*

Proof. In what follows we denote

$$c_\alpha \;=\; \tfrac{1}{1-\sqrt{1-\alpha}} \quad \text{and} \quad I_\alpha = [\,0\,,\,c_\alpha]\,.$$

By the formula of Lemma 3.3, $W(0, c_\alpha) \;=\; \alpha$. Since W is symmetric and nonincreasing, we see that the indicator function $g = \mathbb{1}_{\{W \leq \alpha\}}$ satisfies the conditions of Lemma 2.5 (with $c = c_\alpha$). By that lemma the sup in the assertion equals the maximum value of the (EP2),

$$\begin{aligned} &\text{maximize} \quad Q^2(\,W \leq \alpha\,)\,, \\ &\text{s.t.} \quad Q \in \text{Prob}(I_\alpha)\,, \quad \text{E}[Q] \;=\; 1\,. \end{aligned} \tag{20}$$

By Lemma 3.2, resolving the inequality $W(x_1,x_2)\le\alpha$ for $x_1,x_2\in I_\alpha$, yields (after some lengthy but elementary calculations, cp. [1], p. 84),

$$W(x_1,x_2)\;\le\;\alpha\quad\Longleftrightarrow\quad \tfrac{1}{\sqrt{\alpha}}\le x_{(2)}\le c_\alpha\;,\;x_{(1)}\ge w_\alpha(x_{(2)})\;,\quad (x_1,x_2\in I_\alpha)\;, \tag{21}$$

where we have introduced the function

$$w_\alpha\,:\,[\,\tfrac{1}{\sqrt{\alpha}}\,,\,c_\alpha]\longrightarrow[\,0\,,\,\tfrac{1}{\sqrt{\alpha}}\,]\;,\quad w_\alpha(t)\;=\;\begin{cases} t-\sqrt{t^2-\frac{1}{\alpha}} & ,\text{ if } \frac{1}{\sqrt{\alpha}}\le t\le\frac{1}{\alpha}\\ t-\frac{t-1}{\sqrt{1-\alpha}} & ,\text{ if } \frac{1}{\alpha}<t\le c_\alpha\end{cases}\;. \tag{22}$$

Also we get

$$W(x_1,x_2)\;=\;\alpha\quad\Longleftrightarrow\quad \tfrac{1}{\sqrt{\alpha}}\le x_{(2)}\le c_\alpha\;,\;x_{(1)}=w_\alpha(x_{(2)})\;,\quad (x_1,x_2\in I_\alpha)\;. \tag{23}$$

It is easily seen that the function w_α given by (22) is a continuous and decreasing one-to-one function from $[\,\frac{1}{\sqrt{\alpha}}\,,\,c_\alpha\,]$ onto $[\,0\,,\,\frac{1}{\sqrt{\alpha}}\,]$, and so there is the inverse function w_α^{-1}.
We firstly show that for the (EP2) from (20) feasible probability distributions Q with $Q^2(W\le\alpha)\;=\alpha$ are the following. Let $x^*=(x_1^*,x_2^*)\in I_\alpha^2$ be such that $W(x^*)=\alpha$ and choose Q_{x^*} according to Lemma 3.2, i.e.,

$$\{\,x^*_{(1)},x^*_{(2)}\}\subset\operatorname{supp}(Q_{x^*})\subset\{\,0\,,\,x^*_{(1)},x^*_{(2)}\}\;,\quad \mathrm{E}[Q_{x^*}]\;=\;1\;, \tag{24}$$

$$\text{and}\quad W(x^*)\;=\;Q^2_{x^*}(\{z\in I^2_\alpha\;:\;z_{(1)}\ge x^*_{(1)}\;\;z_{(2)}\ge x^*_{(2)}\})\;. \tag{25}$$

From (21), (23), (24) we see that an $x\;=\;(x_1,x_2)\;\in\;\operatorname{supp}(Q^2_{x^*})$ satisfies $W(x)\;\le\;\alpha\,(=W(x^*))$ if and only if $x_{(1)}\;=\;x^*_{(1)}$, $x_{(2)}\;=\;x^*_{(2)}$ or $x_{(1)}\;=\;x_{(2)}\;=\;x^*_{(2)}$, i.e., if and only if $x_{(1)}\ge x^*_{(1)}$, $x_{(2)}\ge x^*_{(2)}$. Hence, together with (25), we obtain

$$\begin{aligned}&Q^2_{x^*}(\,W\le\alpha\,)\;=\;Q^2_{x^*}(\,\{x\in\operatorname{supp}(Q^2_{x^*})\;:\;x_{(1)}\ge x^*_{(1)}\,,\;x_{(2)}\ge x^*_{(2)}\}\,)\\ &\;=\;W(x^*)\;=\;\alpha\;.\end{aligned}$$

So we have proved that for any $x^*\in I^2_\alpha$ with $W(x^*)\;=\;\alpha$ the probability distribution Q_{x^*} from Lemma 3.2 satisfies $Q^2_{x^*}(\,W\;\le\;\alpha\,)\;=\;\alpha$. The rest of our proof will show, in particular, that those distributions Q_{x^*} are the optimal solutions to the (EP2) from (20).
Let Q_0 be any optimal solution to (20) (which exists by Theorem 2.3, part (i)). By part (ii) of Theorem 2.3 there exist real numbers $\rho\ge0$ and τ such that

$$g_{Q_0}(t)\le\rho+\tau t\quad\forall\,t\in I_\alpha\quad\text{and}\quad g_{Q_0}(t)=\rho+\tau t\quad\forall\,t\in\operatorname{supp}(Q_0)\;, \tag{26}$$

$$\text{where}\quad g_{Q_0}(t)\;=\;Q_0(\,\{s\in I_\alpha\;:\;W(s,t)\le\alpha\}\,)\quad\forall\,t\in I_\alpha\,, \tag{27}$$

and then $Q_0^2(\,W\le\alpha\,)\;=\;\rho+\tau$. So we have to conclude that $\rho+\tau\le\alpha$.
From (27) and (21) we obtain

$$g_{Q_0}(t)\;=\;\begin{cases} Q_0(\,[w_\alpha(t)\,,\,c_\alpha]\,) & ,\text{ if } \frac{1}{\sqrt{\alpha}}\le t\le c_\alpha\\ Q_0(\,[w_\alpha^{-1}(t)\,,\,c_\alpha]\,) & ,\text{ if } 0\le t\le\frac{1}{\sqrt{\alpha}}\end{cases}\;. \tag{28}$$

The function g_{Q_0} is nondecreasing and thus (26) forces $\tau\;\ge\;0$. Suppose $\tau\;=\;0$. Since $0\;\le\;g_{Q_0}(t)\;\le\;1$ for all $t\;\in\;[\,0\,,\,c_\alpha]$ and $g_{Q_0}(c_\alpha)\;=\;1$, (26) forces $\rho\;\ge\;1$ and hence $\rho=1$. The minimum support point s_0 of Q_0 must be less than or equal to 1 (because of

$\mathrm{E}[Q_0]=1$), in particular $s_0 < \frac{1}{\sqrt{\alpha}}$, hence by (26) and (28) $Q_0(\,[\,w_\alpha^{-1}(s_0)\,,\,c_\alpha]\,) \,=1$. But $w_\alpha^{-1}(s_0) > \frac{1}{\sqrt{\alpha}}$ and hence $s_0 \notin \mathrm{supp}(Q_0)$ which is a contradiction. So we have $\tau > 0$.
For the support of Q_0 we conclude further:

$$\text{If } t \in \mathrm{supp}(Q_0) \text{ and } \tfrac{1}{\sqrt{\alpha}} \le t \le c_\alpha, \text{ then } w_\alpha(t) \in \mathrm{supp}(Q_0)\,; \tag{29}$$
$$\text{if } s \in \mathrm{supp}(Q_0) \text{ and } 0 < s < \tfrac{1}{\sqrt{\alpha}}, \text{ then } w_\alpha^{-1}(s) \in \mathrm{supp}(Q_0)\,. \tag{30}$$

This can be seen as follows. Let $t \in \mathrm{supp}(Q_0)$ and $\frac{1}{\sqrt{\alpha}} \le t \le c_\alpha$, but suppose that $w_\alpha(t) \notin \mathrm{supp}(Q_0)$ (hence $t > \frac{1}{\sqrt{\alpha}}$ since $w_\alpha(\frac{1}{\sqrt{\alpha}}) = \frac{1}{\sqrt{\alpha}}$). So there is an $\varepsilon > 0$ such that

$$w_\alpha(t) + \varepsilon \le \tfrac{1}{\sqrt{\alpha}} \quad \text{and} \quad Q_0(\,[\,w_\alpha(t)\,,\,w_\alpha(t)+\varepsilon\,]\,) \,=\, 0\,.$$

Since w_α is continuous and decreasing, there is a $\delta > 0$ such that $t-\delta > \frac{1}{\sqrt{\alpha}}$ and $w_\alpha(t) < w_\alpha(t-\delta) \le w_\alpha(t)+\varepsilon$, hence $Q_0(\,[\,w_\alpha(t)\,,\,w_\alpha(t-\delta)\,]\,) \,=0$, and we get from (26), (28),

$$\rho+\tau t \,=\, g_{Q_0}(t) \,=\, Q_0(\,[\,w_\alpha(t)\,,\,c_\alpha]\,) \,=\, Q_0(\,[\,w_\alpha(t-\delta)\,,\,c_\alpha]\,) \,=\, g_{Q_0}(t-\delta) \,\le\, \rho+\tau(t-\delta)\,,$$

which is a contradiction in view of $\tau > 0$.
The arguments are similar for the case that $s \in \mathrm{supp}(Q_0)$ and $0 < s < \frac{1}{\sqrt{\alpha}}$; suppose that $w_\alpha^{-1}(s) \notin \mathrm{supp}(Q_0)$. Then there is an $\varepsilon > 0$ such that

$$w_\alpha^{-1}(s) + \varepsilon \le c_\alpha \quad \text{and} \quad Q_0(\,[\,w_\alpha^{-1}(s)\,,\,w_\alpha^{-1}(s)+\varepsilon\,]\,) \,=\, 0\,.$$

Since w_α^{-1} is continuous and decreasing, there is a $\delta > 0$ such that $s-\delta \ge 0$ and $w_\alpha^{-1}(s) < w_\alpha^{-1}(s-\delta) \le w_\alpha^{-1}(s)+\varepsilon$, hence $Q_0(\,[\,w_\alpha^{-1}(s)\,,\,w_\alpha^{-1}(s-\delta)\,]\,) \,=0$, and we get from (26), (28),

$$\rho+\tau s \,=\, g_{Q_0}(s) \,=\, Q_0(\,[\,w_\alpha^{-1}(s)\,,\,c_\alpha]\,) \,=\, Q_0(\,[\,w_\alpha^{-1}(s-\delta)\,,\,c_\alpha]\,) \,=\, g_{Q_0}(s-\delta) \,\le\, \rho+\tau(s-\delta)\,,$$

which is a contradiction in view of $\tau > 0$.
Let t_0 denote the maximum support point of Q_0; by (29), (30), $\frac{1}{\sqrt{\alpha}} \le t_0 \le c_\alpha$. Denote $s_0 = w_\alpha(t_0)$. By (26), (28), and (29),

$$\rho+\tau s_0 \,=\, Q_0(t_0) \quad \text{and} \quad \rho+\tau t_0 \,=\, Q_0(\,[\,s_0,\,t_0]\,)\,, \tag{31}$$

where for short we write $Q_0(t_0)$ instead of $Q_0(\{t_0\})$. We have to distinguish three cases to prove $\rho+\tau \le \alpha$.
Case 1: $t_0 = c_\alpha$.
Case 2: $t_0 < c_\alpha$ and $0 \in \mathrm{supp}(Q_0)$.
Case 3: $t_0 < c_\alpha$ and $0 \notin \mathrm{supp}(Q_0)$.
Ad case 1: Then $s_0 = 0$ and (31) rewrites as

$$\rho \,=\, Q_0(c_\alpha) \quad \text{and} \quad \rho+\tau c_\alpha \,=\, 1\,,$$

hence $c_\alpha\rho \,=\, c_\alpha Q_0(c_\alpha) \,\le\, \mathrm{E}[Q_0]=1$ and

$$\rho+\tau \,=\, \rho+\tfrac{1-\rho}{c_\alpha} \,=\, (1-\tfrac{1}{c_\alpha})\rho+\tfrac{1}{c_\alpha} \,\le\, (1-\tfrac{1}{c_\alpha})\tfrac{1}{c_\alpha}+\tfrac{1}{c_\alpha} \,=\, \alpha\,.$$

Ad case 2: Then, by (26) and (28), $\rho = Q_0(c_\alpha) = 0$, and together with (31), $\tau t_0 = Q_0(\,[s_0, t_0]\,) < 1$. If $t_0 \geq \frac{1}{\alpha}$ this yields $\rho + \tau = \tau < 1/t_0 \leq \alpha$. Now let $t_0 < \frac{1}{\alpha}$. From

$$1 \;=\; \mathrm{E}[Q_0] \;\geq\; s_0 Q_0(\,[s_0, t_0)\,) + t_0 Q_0(t_0) \;=\; s_0 Q_0(\,[s_0, t_0]\,) + (t_0 - s_0) Q_0(t_0) \qquad (32)$$

we obtain, observing (31) and $\rho = 0$,

$$1 \;\geq\; s_0 \tau t_0 + (t_0 - s_0)\tau s_0 \;=\; \tau s_0\,(2t_0 - s_0)\ .$$

Because of $\frac{1}{\sqrt{\alpha}} \leq t_0 < \frac{1}{\alpha}$ and $s_0 = w_\alpha(t_0)$ we see from (22) that $s_0(2t_0 - s_0) = \frac{1}{\alpha}$, and we conclude $\rho + \tau = \tau \leq \alpha$.

Ad case 3: Then, by (29), (30), s_0 is the minimum support point of Q_0 and hence $Q_0(\,[s_0, t_0]\,) = 1$. For short let us write $\lambda_0 = Q_0(t_0)$. So (31) rewrites as

$$\rho + \tau s_0 \;=\; \lambda_0 \quad \text{and} \quad \rho + \tau t_0 \;=\; 1\ .$$

Resolving for ρ and τ yields

$$\rho \;=\; \frac{t_0 \lambda_0 - s_0}{t_0 - s_0}\ , \quad \tau \;=\; \frac{1 - \lambda_0}{t_0 - s_0}\ .$$

Since $\rho \geq 0$, we have $\lambda_0 \geq s_0/t_0$; since (32) is generally true, we have $s_0 + (t_0 - s_0)\lambda_0 \leq 1$, hence $\lambda_0 \leq (1-s_0)/(t_0-s_0)$. Thus, $s_0/t_0 \leq (1-s_0)/(t_0-s_0)$, which after straightforward calculations gives $s_0 \leq t_0 - \sqrt{t_0^2 - t_0}$. Since $s_0 = w_\alpha(t_0)$ we see from (22) that $t_0 \geq \frac{1}{\alpha}$, and hence

$$s_0 \;=\; w_\alpha(t_0) \;=\; t_0 - \tfrac{t_0 - 1}{\sqrt{1-\alpha}}\ .$$

So we obtain,

$$\begin{aligned} \rho + \tau \;&=\; \frac{t_0 \lambda_0 - s_0}{t_0 - s_0} + \frac{1 - \lambda_0}{t_0 - s_0} \;=\; \frac{(t_0 - 1)\lambda_0 + 1 - s_0}{t_0 - s_0} \\ &\leq\; \frac{(t_0 - 1)\frac{1 - s_0}{t_0 - s_0} + 1 - s_0}{t_0 - s_0} \;=\; \frac{1 - s_0}{t_0 - s_0}\,\Big(\frac{t_0 - 1}{t_0 - s_0} + 1\Big)\ . \end{aligned}$$

Now, $\frac{1 - s_0}{t_0 - s_0} = 1 - \sqrt{1 - \alpha}$ and $\frac{t_0 - 1}{t_0 - s_0} = \sqrt{1 - \alpha}$, and thus

$$\rho + \tau \;\leq\; (1 - \sqrt{1 - \alpha})\,(1 + \sqrt{1 - \alpha}) \;=\; \alpha\ .$$

□

5 The negative result for $n \geq 3$

As announced earlier, we will prove that if $n \geq 3$ then the test ψ_α^* from Lemma 3.1 does not keep the level α on H_0, i.e.,

$$\sup_{Q \in H_0} Q^n(\,W \leq \alpha\,) \;>\; \alpha\ ,$$

and, moreover, there does not exist a UMP strongly monotone symmetric level α test for H_0 vs. H_1. To this end we need the following lemma.

Lemma 5.1 *Let $n \geq 3$ and $\alpha \in (\,0\,,\,1\,)$ be given. There exist real numbers $b > a > 1$ such that*

$$W(\underbrace{0,\ldots,0}_{n-3\ \textit{times}},a,a,a) \;=\; W(\underbrace{0,\ldots,0}_{n-2\ \textit{times}},a,b) \;=\; \alpha\ .$$

Proof. For any $\beta > 1$ and $1 \leq j \leq n$ we get from Lemma 3.2,

$$W(\underbrace{0,\ldots,0}_{n-j\ \text{times}},\underbrace{\beta,\ldots,\beta}_{j\ \text{times}}) \;=\; \sum_{i=j}^{n}\binom{n}{i}(\tfrac{1}{\beta})^i(1-\tfrac{1}{\beta})^{n-i}\ . \tag{33}$$

This shows, in particular, that for any sequence $x^{(k)} \in [\,0\,,\infty)^n$ ($k \in \mathbb{N}$) with $x^{(k)}_{n)} \to \infty$ for $k \to \infty$ we have $\lim_{k\to\infty} W(x^{(k)}) = 0$, since with $\beta_k = x^{(k)}_{(n)}$ and $j = 1$ in (33)

$$W(x^{(k)}) \;\leq\; W(\underbrace{0,\ldots,0}_{n-1\ \text{times}},\beta_k) \;=\; 1-\Big(1-\frac{1}{\beta_k}\Big)^n \overset{(k\to\infty)}{\longrightarrow} 0\,.$$

Consider $W(0,\ldots,0,\beta,\beta,\beta)$ (zero appearing $n-3$ times) for $1 \leq \beta < \infty$. For $\beta = 1$, $W(0,\ldots,0,1,1,1) = 1$ and for $\beta \to \infty$, $W(0,\ldots,0,\beta,\beta,\beta) \to 0$. By the continuity of W there exists an $a > 1$ such that $W(0,\ldots,0,a,a,a) \;=\; \alpha$. The r.h.s. of (33) decreases when j increases, and in particular

$$\gamma \;=\; W(\underbrace{0,\ldots,0}_{n-2\ \text{times}},a,a) \;>\; W(\underbrace{0,\ldots,0}_{n-3\ \text{times}},a,a,a) \;=\; \alpha\ .$$

Consider $W(0,\ldots,0,a,\beta)$ (zero appearing $n-2$ times) for $a \leq \beta < \infty$. For $\beta = a$, $W(0,\ldots,0,a,a) = \gamma > \alpha$, and for $\beta \to \infty$, $W(0,\ldots,0,a,\beta) \to 0$. By the continuity of W there exists a $b > a$ such that $W(0,\ldots,0,a,b) \;=\; \alpha$. □

Theorem 5.2 *Let $n \geq 3$ and $\alpha \in (\,0\,,\,1\,)$. Choose $b > a > 1$ as in* Lemma 5.1. *Then:*

(i) *For $x^* \;=\; (0,\ldots,0,a,b)$ (with $n-2$ zeros) and Q_{x^*} from* Lemma 3.2 *we have*

$$Q^n_{x^*}(\,W \leq \alpha\,) \;>\; \alpha\,.$$

In particular, the test ψ^*_α does not keep the level α on H_0.

(ii) *The tests ϕ_1 and ϕ_2, defined by*

$$\phi_1(x) \;=\; \begin{cases} 1 & ,\ \textit{if}\ x_{(n-2)} \geq a \\ 0 & ,\ \textit{else} \end{cases} \quad,\quad \phi_2(x) \;=\; \begin{cases} 1 & ,\ \textit{if}\ x_{(n-1)} \geq a\ ,\ x_{(n)} \geq b \\ 0 & ,\ \textit{else} \end{cases} \quad,$$

are nonrandomized monotone symmetric level α tests for H_0, and no strongly monotone symmetric level α test for H_0 can be at least as powerful (uniformly on H_1) as both ϕ_1 and ϕ_2. In particular, there does not exist a UMP strongly monotone symmetric level α test for H_0 vs. H_1.

Proof.

(i) Because of $\mathrm{E}[Q_{x^*}] = 1$ and $b > a > 1$ we have $\mathrm{supp}(Q_{x^*}) \;=\; \{\,0,a,b\,\}$. The set $\{W \leq \alpha\}$ has as a subset

$$C_{x^*} \;=\; \Big\{\,x \in [\,0\,,\,\infty)^n \;:\; x_{(i)} \geq x^*_{(i)}\ \forall\ i = 1,\ldots,n\,\Big\}\ ,$$

since W is symmetric and nonincreasing and $W(x^*) = \alpha$. Another subset of $\{W \le \alpha\}$ is given by

$$D^* = \Big\{ x \in [0, \infty)^n : x_{(i)} = 0, \ 1 \le i \le n-3, \text{ and } x_{(n-2)} = x_{(n-1)} = x_{(n)} = a \Big\},$$

again by the symmetry of W and by $W(0, \ldots, 0, a, a, a) = \alpha$. Clearly, C_{x^*} and D^* are disjoint, and hence

$$Q^n_{x^*}(W \le \alpha) \ge Q^n_{x^*}(C_{x^*}) + Q^n_{x^*}(D^*),$$

and the assertion follows from

$$Q^n_{x^*}(C_{x^*}) = W(x^*) = \alpha \quad \text{and} \quad Q_{x^*}(D^*) = \tbinom{n}{3}(Q_{x^*}(0))^{n-3}(Q_{x^*}(a))^3 > 0.$$

(ii) The tests ϕ_1 and ϕ_2 are clearly nonrandomized, monotone, and symmetric. By definition of W we have, with $x^* = (0, \ldots, 0, a, b)$ (zero appearing $n-2$ times) as above, $y^* = (0, \ldots, 0, a, a, a)$ (zero appearing $n-3$ times), and notation C_{x^*}, C_{y^*} from (11),

$$\sup_{Q \in H_0} \mathrm{E}_{Q^n}(\phi_1) = \sup_{Q \in H_0} Q^n(C_{x^*}) = W(x^*) = \alpha,$$
$$\sup_{Q \in H_0} \mathrm{E}_{Q^n}(\phi_2) = \sup_{Q \in H_0} Q^n(C_{y^*}) = W(y^*) = \alpha.$$

So ϕ_1 and ϕ_2 are nonrandomized monotone symmetric level α tests for H_0. Now let ϕ be any strongly monotone symmetric level α test for H_0 which is uniformly at least as powerful on H_1 as ϕ_1, i.e.,

$$\mathrm{E}_{Q^n}(\phi) \ge \mathrm{E}_{Q^n}(\phi_1) \quad \forall\, Q \in \mathrm{Prob}([0, \infty)) \text{ with } \mathrm{E}[Q] > 1. \tag{34}$$

We have to show that ϕ is <u>not</u> uniformly at least as powerful on H_1 as ϕ_2. To this end we specialize (34) to the two-point distributions $Q = \begin{pmatrix} 0 & a \\ 1-\lambda & \lambda \end{pmatrix}$ with $1/a < \lambda < 1$, which yields

$$\mathrm{E}_{Q^n}(\phi) = \sum_{i=0}^{n} \varphi_i \binom{n}{i} \lambda^i (1-\lambda)^{n-i} \ge \mathrm{E}_{Q^n}(\phi_1) = \sum_{i=3}^{n} \binom{n}{i} \lambda^i (1-\lambda)^{n-i}, \tag{35}$$

where we have denoted

$$\varphi_i = \phi(\underbrace{0, \ldots, 0}_{n-i \text{ times}}, \underbrace{a, \ldots, a}_{i \text{ times}}), \quad 0 \le i \le n.$$

Hence it follows that $\varphi_3 = 1$, i.e.,

$$\phi(0, \ldots, 0, a, a, a) = 1. \tag{36}$$

For, suppose that $\varphi_3 < 1$. Then, by strong monotonicity of ϕ, $\varphi_i = 0$ for $i = 0, 1, 2$ which contradicts (35). As a next step we show that

$$\theta = \phi(0, \ldots, 0, a, b) < 1. \tag{37}$$

Suppose that $\theta = 1$. Then $\phi \ge \phi_2$ by the symmetry and monotonicity of ϕ. By Lemma 3.2 there is a probability distribution Q_0 with support equal to $\{0, a, b\}$ and

with $\mathrm{E}[Q_0] = 1$ such that $\alpha = W(0,\ldots,0,0,a,b) = \mathrm{E}_{Q_0^n}(\phi_2)$. But $\mathrm{E}_{Q_0^n}(\phi) \le \alpha$, and thus we must have $\phi(x) = \phi_2(x)$ for all x from the support of Q_0^n, i.e., for all x with components in $\{0,a,b\}$, and in particular $\phi(0,\ldots,0,a,a,a) = \phi_2(0,\ldots,0,a,a,a) = 0$, contradicting (36). Hence (37) follows.
Consider the sets

$$\begin{aligned} C &= \{x = (x_1,\ldots,x_n) : x_i \in \{0,a,b\} \ (1 \le i \le n),\ x_{(n-2)} = 0,\ \ x_{(n-1)} = a,\ x_{(n)} = b\}\,, \\ D &= \{x = (x_1,\ldots,x_n) : x_i \in \{0,a,b\} \ (1 \le i \le n),\ x_{(n)} \le a\}\,. \end{aligned}$$

Then we have

$$\begin{aligned} &\phi(x) + (1-\theta)\,\mathbb{1}_C(x) \ \le\ \phi_2(x) + \mathbb{1}_D(x) \\ &\forall\ x = (x_1,\ldots,x_n) \text{ with } x_i \in \{0,a,b\} \ (1 \le i \le n)\,, \end{aligned} \tag{38}$$

which can be seen as follows. The l.h.s. of (38) does not exceed 1 because of $\phi(x) = \theta$ on C, while the r.h.s. of (38) yields either 0, 1, or 2. So it remains to verify inequality (38) for all x with components in $\{0,a,b\}$ and $\phi_2(x) = \mathbb{1}_D(x) = 0$. The latter means that $x_{(n-1)} = 0$ and $x_{(n)} = b$, hence $\phi(x) = \phi(0,\ldots,0,b)$ which equals zero by (37) and the strong monotonicity of ϕ, and clearly $\mathbb{1}_C(x) = 0$. This shows (38). Hence it follows that for any probability distribution Q with $\mathrm{supp}(Q) = \{0,a,b\}$ we have

$$\begin{aligned} &\mathrm{E}_{Q^n}(\phi_2) - \mathrm{E}_{Q^n}(\phi) \ \ge\ (1-\theta)\,Q^n(C) - Q^n(D) \\ &= n(n-1)(1-\theta)\,Q(b)Q(a)(Q(0))^{n-2} - (1-Q(b))^n\,. \end{aligned}$$

Taking $Q(0) = Q(a) = \varepsilon$, $Q(b) = 1 - 2\varepsilon$ with $0 < \varepsilon < 1/2$ and $\mathrm{E}[Q] > 1$, i.e., $\varepsilon < (b-1)/(2b-a)$, one gets

$$\begin{aligned} &\mathrm{E}_{Q^n}(\phi_2) - \mathrm{E}_{Q^n}(\phi) \ \ge\ n(n-1)(1-\theta)(1-2\varepsilon)\varepsilon^{n-1} - (2\varepsilon)^n \\ &= \varepsilon^{n-1}\Big(n(n-1)(1-\theta) - [2n(n-1)(1-\theta) + 2^n]\varepsilon\Big)\,. \end{aligned}$$

The latter expression is positive for some positive small ε. So we have shown that there exists a probability distribution Q on $[\,0\,,\,\infty)$ with $\mathrm{E}[Q] > 1$ and $\mathrm{E}_{Q^n}(\phi) < \mathrm{E}_{Q^n}(\phi_2)$, i.e., ϕ is not uniformly at least as powerful on H_1 as ϕ_2. □

References

[1] Wang, W., Zhao, L.H., *Nonparametric tests for the mean of a non-negative population.* J. Statist. Plann. Inference 110 (2003), pp. 75-96.

[2] Bauer, H., *Maß und Integrationstheorie.* [In German]. De Gruyter, Berlin, 1992.

[3] Billingsley, P., *Weak Convergence of Measures: Applications in Probability.* Soc. for Industrial and Applied Mathematics, Philadelphia, 1971.

[4] Shaked, M., Shanthikumar, J.G., *Stochastic Orders and Their Applications.* Academic Press, San Diego, 1994.