Three test statistics for a nonparametric one-sided hypothesis on the mean of a nonnegative variable

Norbert Gaffke*

University of Magdeburg, Faculty of Mathematics
D-39016 Magdeburg, PF 4120, Germany

Abstract

Assume the nonparametric model of n i. i. d. nonnegative real random variables whose distribution is unknown. Consider the one sided hypotheses on the expectation, $H_0 : \mu \leq 1$ vs. $H_1 : \mu > 1$. Wang & Zhao (2003) studied several statistics for significance testing. Here we focus on three statistics. One was introduced in Wang & Zhao (2003), W say, another is the nonparametric likelihood ratio statistic (R) also studied in that paper, and last but not least we propose a new statistic (K). Either of these statistics has its values between zero and one, and it seems reasonable to reject the null hypothesis iff the value is smaller than or equal to α (the nominal significance level). However, when doing so, the question is whether the desired level α is really kept. For $n \leq 2$ the answer is positive as shown by Wang & Zhao (2003) for W and R, and hence positive for K as well, since we will show that $W \leq K \leq R$ (for arbitray n). For $n \geq 3$ the answer is negative for W as shown by Gaffke (2004), but the definite answers for R and K are unknown. We will report some numerical evidence and an asymptotic result on the statistic K which let us conjecture that the answer for K (hence for R as well) is positive for arbitrary sample size. Somewhat surprisingly, the numerics indicate that this should be true even when we suspend the assumption of *identically* distributed observations. For $n = 2$ this is proved.

Key words: Level of a test, UMP test, order statistics, stochastic ordering, asymptotic distribution, finite sample distribution.

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,

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

There is a close relation to the problem of constructing lower confidence bounds for μ. In particular, if ϕ is a nonrandomized level α test for (1) then a lower $(1-\alpha)$-confidence bound for μ is obtained by

$$\widehat{\mu}_\ell(x) = \inf\{ \beta > 0 : \phi(x/\beta) = 0 \} , \quad (x \in [0, \infty)^n).$$

* *E-mail:* norbert.gaffke@mathematik.uni-magdeburg.de

Lower $(1-\alpha)$-confidence bounds were derived in a few papers, Breth (1976), Breth, Maritz, Williams (1978), and Kaplan (1987). The recent paper of Wang & Zhao (2003) deals with testing the one-sided hypotheses (1). For $n = 1$ the UMP non-randomized level α test is given by $\phi^*_\alpha = 1\!\!1_{[1/\alpha\,,\infty)}$. Somewhat surprisingly, for $n = 2$ Wang & Zhao (2003) derived a UMP level α test within the class of all *strongly monotone (nondecreasing) symmetric* level α tests. This is a highly nontrivial result though, unfortunately, not extendable to larger sample size $n \geq 3$ (cf. Gaffke (2004)). The Wang-Zhao test (for $n = 2$) is $\phi^*_\alpha = 1\!\!1_{\{W\leq\alpha\}}$, where the statistic W is given by

$$W(x) \;=\; \sup_{H_0} P\Big(\, X_{(1)} \geq x_{(1)}\,,\, X_{(2)} \geq x_{(2)}\Big)\;, \quad x = (x_1, x_2) \in [\,0\,,\,\infty)^2\;. \tag{2}$$

Note that (2) means that for a fixed sample point $x = (x_1, x_2)$ the sup is taken over all possible distributions from H_0 of the i.i.d. random variables X_1, X_2, and $X_{(1)} \leq X_{(2)}$ denote their order statistics, $x_{(1)} \leq x_{(2)}$ denote the ordered values of x_1, x_2. An explicit formula for $W(x)$ was derived in Wang & Zhao (2003), p. 90. The straightforward extension of (2) to sample size $n \geq 3$,

$$W(x) \;=\; \sup_{H_0} P\Big(\, X_{(i)} \geq x_{(i)} \;\forall\, i = 1, \ldots, n\Big)\;, \quad x = (x_1, \ldots, x_n) \in [\,0\,,\,\infty)^n\;, \tag{3}$$

fails to yield a level α test via $\phi = 1\!\!1_{\{W\leq\alpha\}}$, i.e., we have

$$\sup_{H_0} P\Big(\, W \leq \alpha\Big) \;>\; \alpha \quad \text{whenever } n \geq 3,$$

as it was shown in Gaffke (2004). Also, Wang & Zhao (2003) considered the nonparametric likelihood ratio statistic R (see Section 2), and the critical region $\{R \leq \alpha\}$ for (1). We will introduce a third statistic K, and the critical region $\{K \leq \alpha\}$. It will be shown in Section 2 that the three statistics are related by $W \leq K \leq R$. An interesting question is whether the critical regions for (1) given by K or R are of level α (when $n \geq 3$). Posing that question simultaneously for all $\alpha \in (\,0\,,\,1\,)$, we are faced with the question whether the distributions of K or R under H_0 are stochastically larger than the standard uniform distribution (on the interval $(\,0\,,\,1\,)$). There is some evidence that the answer will be affirmative for K (and hence also for R), based on numerical experiments and an asymptotic result (see Sections 3 and 4). Somewhat surprisingly, the numerics indicate that this may be true even if we dispense the assumption of *identically* distributed random variables $X_1, \ldots, X_n$, assuming only $X_1, \ldots, X_n$ to be independent with expectations less than or equal to one.

2 Three statistics

An alternative representation of the statistic W from (3) is the following, as it was proved in Gaffke (2004), Lemma 3.4.

$$\begin{aligned} W(x) \;&=\; \max\Big\{\, P\big(U_{(1)} \geq u_1\,,\,\ldots\,,\,U_{(n)} \geq u_n\big) \;:\; 0 \leq u_1 \leq \ldots \leq u_n \leq 1\,, \\ &\qquad\qquad \sum_{i=1}^n u_i\big(x_{(i)} - x_{(i-1)}\big) = x_{(n)} - 1 \,\Big\}\;, \quad \text{if } x_{(n)} > 1, \qquad (4) \\ W(x) \;&=\; 1\;, \qquad \text{if } x_{n)} \leq 1, \end{aligned}$$

for any $x = (x_1, \ldots, x_n) \in [0, \infty)^n$, where $U_{(1)} \leq \ldots \leq U_{(n)}$ are the order statistics of n i.i.d. standard uniformly distributed (on the interval $(0, 1)$) random variables $U_1, \ldots, U_n$, and $x_{(0)} = 0$.
The nonparametric likelihood ratio statistic for problem (1) is defined by

$$R(x) = \sup_{Q \in H_0} L(Q, x) \Big/ \sup_{Q \in H_0 \cup H_1} L(Q, x) , \quad x = (x_1, \ldots, x_n) \in [0, \infty)^n , \qquad (5)$$

where $L(Q, x)$ is the nonparametric likelihood,

$$L(Q, x) = \prod_{i=1}^{n} Q\big(\{x_i\}\big) , \qquad (6)$$

(cf. Wang & Zhao (2003), p. 81). Here, H_0 and H_1 stand for the sets of all probability distributions on $[0, \infty)$ with expectations $\mu(Q) \leq 1$ and $\mu(Q) > 1$, resp. Of course, for $L(Q, x)$ and $R(x)$ only those Q from H_0 and H_1 need to be considered which have point masses at the observations $x_1, \ldots, x_n$ (for any other Q one has $L(Q, x) = 0$). A nearly explicit formula for $R(x)$ is the following.

$$R(x) = \min_{0 \leq t \leq 1} \prod_{i=1}^{n} \big(1 - t + t\, x_i\big)^{-1} , \qquad (7)$$

which was derived in Wang & Zhao (2003), Theorem 4.1, for the case that all components of x are distinct; we will give a proof of (7) for arbitrary x the appendix.
A third statistic we will introduce here, based on an idea in Gaffke & Zöllner (2003), is

$$K(x) = P\Big(\sum_{i=1}^{n} x_{(i)} \big(U_{(i+1)} - U_{(i)}\big) \leq 1 \Big) , \quad x = (x_1, \ldots, x_n) \in [0, \infty)^n , \qquad (8)$$

where $U_{(1)} \leq \ldots \leq U_{(n)}$ are the order statistics of n i.i.d. standard uniformly distributed random variables $U_1, \ldots, U_n$, and $U_{(n+1)} = 1$. A motivation for using $\{K \leq \alpha\}$ as a rejection region for the null hypothesis H_0 of (1) (where α is a given nominal level) is as follows. Let Q be the true (but unknown) underlying probability distribution on $[0, \infty)$, and denote by F its (right continuous) c.d.f. and by F^- the pseudo-inverse of F. Then we obtain i.i.d. Q-distributed random variables $X_1, \ldots, X_n$ by $X_i = F^-(U_i)$ $(1 \leq i \leq n)$. So the observed values $x_1, \ldots, x_n$ emerge from values $u_1, \ldots, u_n$ of $U_1, \ldots, U_n$, resp., via $x_i = F^-(u_i)$ $(1 \leq i \leq n)$. Now the expectation $\mu = \mu(Q)$ can be written as $\mu = \int_0^1 F^-(v)\, \mathrm{d}v$, and since F^- is nondecreasing we have

$$\mu \geq \sum_{i=1}^{n} F^-(u_{(i)}) \big(u_{(i+1)} - u_{(i)}\big) = \sum_{i=1}^{n} x_{(i)} \big(u_{(i+1)} - u_{(i)}\big) ,$$

where $u_{(n+1)} = 1$. So, if the lower bound $\sum_{i=1}^{n} x_{(i)} \big(u_{(i+1)} - u_{(i)}\big)$ exceeds 1 then the alternative Hypothesis $H_1 : \mu > 1$ is true. However, that lower bound cannot be computed since the values $u_1, \ldots, u_n$ are not observable. So the idea is to suppose a resampling of all possible values of $u_1, \ldots, u_n$, i.e., to replace them by the random variables $U_1, \ldots, U_n$, and to look at the probability

$$P\Big(\sum_{i=1}^{n} x_{(i)} \big(U_{(i+1)} - U_{(i)}\big) > 1 \Big) .$$

If this is large (greater than or equal to $1-\alpha$) or, equivalently, if $K(x) \le \alpha$, then we decide for H_1, i.e., reject the null hypothesis H_0. However, this is only a heuristic view which is mathematically not conclusive. So, in the next sections of the paper, we will focus the question whether the rejection region $\{K \le \alpha\}$ really keeps the nominal level α on H_0.
Two alternative formulas for the statistic K are the following. Firstly,

$$K(x) \;=\; P\Big(\sum_{i=1}^{n} x_i D_i \;\le\; 1\Big)\;, \tag{9}$$

where $D = (D_1, \ldots, D_n)$ denotes a random variable which is uniformly distributed over the unit simplex in $\mathbb{R}^n$,

$$\Big\{\, d = (d_1, \ldots, d_n) \in \mathbb{R}^n \;:\; d_i \ge 0 \;(1 \le i \le n)\,,\; \sum_{i=1}^{n} d_i \le 1 \Big\}\;;$$

and secondly we have

$$K(x) \;=\; P\Big(\sum_{i=1}^{n} (x_i - 1) Z_i \;\le\; Z_0\Big)\;, \tag{10}$$

where $Z_0, Z_1, \ldots, Z_n$ are i.i.d. standard exponentially distributed random variables.

Formula (9) is obtained by observing that for i.i.d. standard uniformly distributed random variables the vector of spacings $U_{(i+1)} - U_{(i)}$ $(1 \le i \le n)$, where $U_{(n+1)} = 1$, is uniformly distributed over the unit simplex, and because of permutational symmetry of that distribution the ordered values $x_{(1)} \le \ldots \le x_{(n)}$ in (8) can be replaced by the original values $x_1, \ldots, x_n$. Formula (10) results from the fact (cf. e.g. David (1981), p. 103) that if $Z_0, Z_1, \ldots, Z_n$ are i.i.d. standard exponentially distributed random variables then the joint distribution of $D_i = Z_i \big/ \sum_{j=0}^{n} Z_j$ $(1 \le i \le n)$ is the uniform distribution on the unit simplex.

Theorem 2.1 *For arbitrary sample size $n \ge 1$ we have*

$$W(x) \;\le\; K(x) \;\le\; R(x) \quad \textit{for all } x \in [\,0\,,\,\infty)^n\,.$$

Proof. To prove $W \le K$ we use the representation (4) of $W(x)$. Firstly, if x is such that $x_{(n)} \le 1$ then $W(x) = 1$ and also, by (8), $K(x) = 1$. Now let $x_{(n)} > 1$. According (4) let $u_1, \ldots, u_n$ be given with $0 \le u_1 \le \ldots \le u_n \le 1$ and $\sum_{i=1}^{n} u_i (x_{(i)} - x_{(i-1)}) = x_{(n)} - 1$, and let $U_1, \ldots, U_n$ be i.i.d. standard uniform random variables. Then, clearly,

$$\Big\{\, U_{(i)} \ge u_i \;\forall\; i = 1, \ldots, n \Big\} \;\subset\; \Big\{\, \sum_{i=1}^{n} U_{(i)} (x_{(i)} - x_{(i-1)}) \ge x_{(n)} - 1 \Big\}\;, \quad \text{hence}$$

$$P\Big(\, U_{(i)} \ge u_i \;\forall\; i = 1, \ldots, n \Big) \;\le\; P\Big(\sum_{i=1}^{n} U_{(i)} (x_{(i)} - x_{(i-1)}) \ge x_{(n)} - 1 \Big)\;. \tag{11}$$

The r.h.s. of (11) is equal to $K(x)$ since

$$x_{(n)} - \sum_{i=1}^{n} U_{(i)} (x_{(i)} - x_{(i-1)}) \;=\; \sum_{i=1}^{n} x_{(i)} \big(U_{(i+1)} - U_{(i)}\big)$$

(note that $x_{(0)} = 0$ and $U_{(n+1)} = 1$). By (4), this proves $W(x) \le K(x)$.

To prove $K(x) \le R(x)$ we use representations (7) and (9). Let $x \in [\,0\,,\,\infty)^n$ be given. We have to show, for any $t \in [\,0\,,\,1\,]$, that

$$K(x) \;\le\; \prod_{i=1}^{n} (1 - t \,+\, t\,x_i)^{-1} \;, \tag{12}$$

and we may assume that t is such that $1 - t + tx_i \; > 0$ for all $i = 1, \ldots, n$ (otherwise the r.h.s. of (12) is defined to be infinity). According to (9) denote by $D_1, \ldots, D_n$ random variables whose joint distribution is the uniform distribution on the unit simplex, and we have

$$K(x) \;=\; P\Big(\sum_{i=1}^{n} x_i D_i \;\le 1 \Big) \;=\; n!\,\mathrm{vol}_n\big(B\big) \;\text{, where}$$

$$B \;=\; \big\{ d \in \mathbb{R}^n \;:\; d_i \ge 0 \;\forall\, i \;,\; \sum_{i=1}^{n} d_i \le 1 \;,\; \sum_{i=1}^{n} x_i d_i \le 1 \,\big\} \;,$$

and vol_n stands for the n-dimensional volume. We get

$$\begin{aligned} B \;&=\; \Big\{ d \in \mathbb{R}^n \;:\; d_i \ge 0 \;\forall\, i \;,\; \max\big\{ \sum_{i=1}^{n} d_i \,,\, \sum_{i=1}^{n} x_i d_i \big\} \;\le 1 \Big\} \\ &\subset\; \Big\{ d \in \mathbb{R}^n \;:\; d_i \ge 0 \;\forall\, i \;,\; (1-t) \sum_{i=1}^{n} d_i \;+\; t \sum_{i=1}^{n} x_i d_i \;\le 1 \Big\} \\ &=\; \Big\{ d \in \mathbb{R}^n \;:\; d_i \ge 0 \;\forall\, i \;,\; \sum_{i=1}^{n} (\,1 - t + tx_i) d_i \;\le 1 \Big\} \;. \end{aligned}$$

The volume of the latter set equals $\frac{1}{n!} \prod_{i=1}^{n} (1-t+tx_i)^{-1}$, which is thus an upper bound for $\mathrm{vol}_n(B)$. From this (12) follows. □

As pointed out in the introduction, a crucial question is whether the statistics W, K, or R are, under the null-hypothesis H_0, stochastically larger than a standard uniform random variable. That is, if $X_1, \ldots, X_n$ are any i.i.d. nonnegative real random variables with $\mathrm{E}(X_i) \le 1$, is it true that, (for $T = W$, K, or R),

$$T(X_1, \ldots, X_n) \;\overset{\mathcal{D}}{\ge}\; U \quad \text{(a standard uniform r.v.)} \tag{13}$$

$$\text{i.e.,} \quad P\Big(T(X_1, \ldots, X_n) \;\le\; \alpha \Big) \;\le\; \alpha \quad \forall\, \alpha \in (\,0\,,\,1\,) \;? \tag{14}$$

For $n \le 2$, (13) is true for W, K, and R by the Wang-Zhao results and by Theorem 2.1. For $n \ge 3$, (13) is generally *not* true for W (cf. Gaffke (2004)), but is still an open question for K and R. If we restrict to a two-point distribution of the i.i.d. random variables X_i, then we get the following minor result.

Lemma 2.2 *If* $X_1, \ldots, X_n$ *are i.i.d. two-point distributed nonnegative random variables with* $\mathrm{E}(X_i) \le 1$*, then*

$$W(X_1, \ldots, X_n) \;\overset{\mathcal{D}}{\ge}\; U \;,$$

(where U *is a standard uniform random variable).*

Proof. Since $W(x)$, $x \in [\,0\,,\,\infty)^n$, is a nonincreasing function (w.r.t. the componentwise semi-ordering on $[\,0\,,\,\infty)^n$), it suffices to consider the case of i.i.d. two-point distributed nonnegative X_i with $\mathrm{E}(X_i) = 1$. Moreover, we may assume that one support point of the two-point distribution is zero, which can be seen as follows. Denote the support points by a and b, where $0 \le a < 1 < b$. Then, $\widetilde{X}_i = (X_i - a)/(1-a)$, $i = 1, \ldots, n$, again are i.i.d. two-point distributed nonnegative random variables with $\mathrm{E}(\widetilde{X}_i) = 1$, and the support points of their distribution are 0 and $\widetilde{b} = (b-a)/(1-a)$. We have by definition (2) of $W(x)$,

$$W(\widetilde{X}_1, \ldots, \widetilde{X}_n) \;=\; \sup_{H_0} P^*\big(\, X^*_{(i)} \ge \widetilde{X}_{(i)} \ \forall\ i = 1, \ldots, n \,\big) \ ,$$

the sup being taken over all i.i.d. nonnegative random variables $X^*_1, \ldots, X^*_n$ with distribution in H_0; our notation P^* is to indicate that the probability refers only to the X^*_i random variables, while the values of the $\widetilde{X}_i$ are considered fixed. Now, the inequalities $X^*_{(i)} \ge \widetilde{X}_{(i)}$ rewrite as $(1-a)X^*_{(i)} + a \ge X_{(i)}$; observing that $Y^*_i = (1-a)X^*_i + a$, $i = 1, \ldots, n$, again are i.i.d. nonnegative random variables with distribution in H_0, we see that each probability

$$P^*\big(\, X^*_{(i)} \ge \widetilde{X}_{(i)} \ \forall\ i = 1, \ldots, n \,\big)$$

obtained from all i.i.d. nonnegative $X^*_1, \ldots, X^*_n$ with distribution in H_0 also appears among the probabilities

$$P^*\big(\, X^*_{(i)} \ge X_{(i)} \ \forall\ i = 1, \ldots, n \,\big)$$

obtained from all i.i.d. nonnegative $X^*_1, \ldots, X^*_n$ with distribution in H_0.
Since $W(X_1, \ldots, X_n)$ is the sup over the latter probabilities we see that

$$W(\widetilde{X}_1, \ldots, \widetilde{X}_n) \;\le\; W(X_1, \ldots, X_n) \ .$$

Therefore it suffices to prove the assertion for any i.i.d. random variables $X_1, \ldots, X_n$ having a two point distribution $\begin{pmatrix} 0 & b \\ 1 - \frac{1}{b} & \frac{1}{b} \end{pmatrix}$, for some $b > 1$.
For an $x = (0, \ldots, 0, b, \ldots, b)$, where b appears $k \ge 1$ times, it is readily seen from (4) that

$$W(x) \;=\; P\Big(\, U_{(n-k+1)} \ge 1 - \tfrac{1}{b} \Big) \;=\; 1 - F_{n,1/b}(k-1) \ ,$$

where $F_{n,1/b}$ denotes the (right continuous) c.d.f. of the binomial-$(n, 1/b)$-distribution, and also for $k = 0$ trivially $W(0, \ldots, 0) = 1 = 1 - F_{n,1/b}(-1)$. We have thus, denoting $B = \sum_{i=1}^n \mathbb{1}_{\{b\}}(X_i)$ (which is a binomially-$(n, 1/b)$-distributed random variable),

$$W(X_1, \ldots, X_n) \;=\; 1 - F_{n,1/b}(B-1) \ .$$

Hence, for any $\alpha \in (\,0\,,\,1\,)$,

$$W(X_1, \ldots, X_n) \le \alpha \iff F_{n,1/b}(B-1) \ge 1 - \alpha \iff B > F^-_{n,1/b}(1-\alpha)\,,$$

where

$$F^-_{n,1/b}(1-\alpha) \;=\; \min\{k \,:\, F_{n,1/b}(k) \ge 1 - \alpha\}\,.$$

We conclude that

$$P\Big(\, W(X_1, \ldots, X_n\,) \,\le \alpha \Big) \;=\; 1 -\, F_{n,1/b}\big(\, F^-_{n,1/b}(1-\alpha)\,\big) \;\le\; \alpha\ . \tag{15}$$

Remarks.

1. For a vector x with all components from $\{0, b\}$, where $b > 1$, we have by the proof and by (8) that $W(x) = K(x) = 1 - F_{n,1/b}(k-1)$ (k denoting the number of components of x equal to b), whereas by (7) the value $R(x)$ turns out to be larger, unless $x = (b, \ldots, b)$ in which case $W(x) = K(x) = R(x) = 1/b^n$.

2. For a given $\alpha \in (0, 1)$ there are two-point distributions $\begin{pmatrix} 0 & b \\ 1 - \frac{1}{b} & \frac{1}{b} \end{pmatrix}$ (with $b > 1$) such that for n i.i.d. random variables $X_1, \ldots, X_n$ with that distribution,

$$P\Big(W(X_1, \ldots, X_n) \leq \alpha \Big) = \alpha ,$$

From (15) we see that those b have to be such that $F_{n,1/b}\big(F^{-}_{n,1/b}(1-\alpha) \big) = 1 - \alpha$, i.e., b has to be such that

$$F_{n,1/b}(k) = 1 - \alpha \quad \text{for some } k \in \{0, 1, \ldots, n-1\}.$$

In fact, there are n distinct such b's, $b_0(\alpha) > b_1(\alpha) > \ldots > b_{n-1}(\alpha) > 1$, where $b_k(\alpha)$ is the solution (for $b > 1$) of $F_{n,1/b}(k) = 1 - \alpha$, ($k = 0, 1, \ldots, n-1$). Or equivalently, by $F_{n,1/b}(k) = P(U_{(k+1)} > 1/b)$, the value $1/b_k(\alpha)$ equals the α-quantile of the order statistic $U_{(k+1)}$. In particular, we obtain $b_0(\alpha) = 1/\big(1-(1-\alpha)^{1/n}\big)$ and $b_{n-1}(\alpha) = \alpha^{-1/n}$.

3 Asymptotics of K

Theorem 3.1 *Let X_i ($i \in \mathbb{N}$) be an infinite sequence of nonnegative real i.i.d. random variables and $\mu = \mathrm{E}(X_i)$. Denote $K_n = K\big(X_1, \ldots, X_n\big)$ for all $n \in \mathbb{N}$. We have the following results on the asymptotic behaviour of K_n.*

If $\mu < 1$ then $\lim_{n\to\infty} K_n = 1$ P-a.s.;
if $\mu > 1$ then $\lim_{n\to\infty} K_n = 0$ P-a.s.;
if $\mu = 1$ and X_i has a positive finite variance, then K_n converges in distribution to a standard uniform random variable.

Proof. By the SLLN, for any i.i.d. sequence Y_i ($i \in \mathbb{N}$) of nonnegative real random variables the following convergence results hold.

$$\frac{1}{n} \sum_{i=1}^{n} Y_i \overset{(n\to\infty)}{\longrightarrow} \mathrm{E}(Y_1) \quad P\text{-a.s.}\,; \tag{16}$$

if $\mathrm{E}(Y_1) < \infty$ then

$$\frac{1}{n} \max_{1\leq i\leq n} Y_i \overset{(n\to\infty)}{\longrightarrow} 0 \quad P\text{-a.s.}\,, \quad \text{and} \quad \frac{1}{n^2} \sum_{i=1}^{n} Y_i^2 \overset{(n\to\infty)}{\longrightarrow} 0 \quad P\text{-a.s.}\,; \tag{17}$$

if $0 < \mathrm{E}(Y_1) < \infty$ then

$$\max_{1\leq i\leq n} Y_i \Big/ \Big(\sum_{j=1}^{n} Y_j \Big) \overset{(n\to\infty)}{\longrightarrow} 0 \quad P\text{-a.s.}\,. \tag{18}$$

In fact, (16) is the SLLN (which also holds true in the case $\mathrm{E}(Y_1) = \infty$). In the case that $\mathrm{E}(Y_1) < \infty$, (17) is obtained from the SLLN as follows. Consider any path y_i ($i \in \mathbb{N}$ of the sequence Y_i having the SLLN property $\frac{1}{n}\sum_{i=1}^n y_i \to \mathrm{E}(Y_1)$; if that path is bounded from above then trivially $\max_{1\le i\le n} y_i/n \to 0$; otherwise, consider the record times n_k ($k \in \mathbb{N}$) defined by $n_1 = 1$ and $n_k = \min\{i > n_{k-1} : y_i > y_{n_{k-1}}\}$ (for $k \ge 2$); it follows that

$$y_{n_k}/n_k = \frac{1}{n_k}\sum_{i=1}^{n_k} y_i - \frac{n_k - 1}{n_k}\frac{1}{n_k - 1}\sum_{i=1}^{n_k-1} y_i \overset{(k\to\infty)}{\longrightarrow} \mathrm{E}(Y_1) - \mathrm{E}(Y_1) = 0\,,$$

which clearly implies $\max_{1\le i\le n} y_i/n \to 0$ for $n \to \infty$. Now,

$$\frac{1}{n^2}\sum_{i=1}^n y_i^2 \le \Big(\frac{1}{n}\max_{1\le i\le n} y_i\Big)\Big(\frac{1}{n}\sum_{i=1}^n y_i\Big) \overset{(n\to\infty)}{\longrightarrow} 0\,.$$

(18) follows from (17) and (16), since

$$\max_{1\le i\le n} Y_i \Big/ \Big(\sum_{j=1}^n Y_j\Big) = \frac{1}{n}\max_{1\le i\le n} Y_i \Big/ \Big(\frac{1}{n}\sum_{j=1}^n Y_j\Big)\,.$$

Now let x_i ($i \in \mathbb{N}$) be any path of the sequence X_i. By the SLLN we may assume that $\overline{x}_n = \frac{1}{n}\sum_{i=1}^n x_i \to \mu$ (for $n \to \infty$). Let $\mu < \infty$. By (17) applied to $Y_i = |X_i - 1|$ we may assume that $\frac{1}{n^2}\sum_{i=1}^n (x_i - 1)^2 \to 0$ for $n \to \infty$. We will use the representation (10) of the function K. So we introduce an infinite sequence $Z_0, Z_1, \ldots, Z_n, \ldots$ of i.i.d. standard exponentially distributed random variables. By (10), we have for all n,

$$K(x_1, \ldots, x_n) = P\Big(\tfrac{1}{n}\sum_{i=1}^n (x_i - 1)Z_i - \tfrac{1}{n}Z_0 \le 0\Big)\,. \tag{19}$$

The sequence of random variables $(x_i - 1)Z_i$ ($i \in \mathbb{N}$) satisfies the WLLN, i.e.,

$$\frac{1}{n}\sum_{i=1}^n \Big((x_i - 1)Z_i - (x_i - 1)\Big) \overset{(n\to\infty)}{\longrightarrow} 0 \quad \text{in prob.}, \tag{20}$$

since (cf. e.g. Bauer (1996), Theorem 10.2)

$$\frac{1}{n^2}\sum_{i=1}^n \mathrm{Var}\Big((x_i - 1)Z_i\Big) = \frac{1}{n^2}\sum_{i=1}^n (x_i - 1)^2 \overset{(n\to\infty)}{\longrightarrow} 0\,.$$

Because of $\frac{1}{n}\sum_{i=1}^n (x_i - 1) \longrightarrow \mu - 1$ for $n \to \infty$, (20) is the same as

$$\frac{1}{n}\sum_{i=1}^n (x_i - 1)Z_i \overset{(n\to\infty)}{\longrightarrow} \mu - 1 \quad \text{in prob.}$$

Together with (19) we see, that if $\mu < 1$ then $K(x_1, \ldots, x_n) \to 1$ for $n \to \infty$ and if $\mu > 1$ then $K(x_1, \ldots, x_n) \to 0$ for $n \to \infty$. The latter also holds true if $\mu = \infty$, since in that case one can easily construct another i.i.d. sequence $\widetilde{X}_i$ ($i \in \mathbb{N}$) of nonnegative real random variables such that $\widetilde{X}_i \le X_i$ for all i and $1 < \widetilde{\mu} = \mathrm{E}(\widetilde{X}_i) < \infty$. So almost every path $\widetilde{x}_i$ ($i \in \mathbb{N}$) of the sequence $\widetilde{X}_i$ satisfies $K(\widetilde{x}_1, \ldots, \widetilde{x}_n) \to 0$ for $n \to \infty$ by the above, and hence also $K(x_1, \ldots, x_n) \to 0$ for $n \to \infty$ because of $\widetilde{x}_i \le x_i$ for all i and $K(x_1, \ldots, x_n) \le K(\widetilde{x}_1, \ldots, \widetilde{x}_n)$ for all n.

Now let $\mu = 1$ and $0 < \sigma^2 = \mathrm{Var}(X_1) < \infty$. Applying (18) on $Y_i = (X_i - 1)^2$, we see that almost every path x_i ($i \in \mathbb{N}$) of the sequence X_i satisfies

$$\max_{1\le i\le n} (x_i - 1)^2 \Big/ \Big(\sum_{j=1}^{n} (x_j - 1)^2 \Big) \stackrel{(n\to\infty)}{\longrightarrow} 0 \; .$$

Hence the sequence of random variables $(x_i - 1) Z_i$ ($i \in \mathbb{N}$) satisfies the CLT (cf. e.g. Billingsley (1995), Problem 27.6), i.e.,

$$\Big(\sum_{i=1}^{n} (x_i - 1) Z_i \; - \; \sum_{i=1}^{n} (x_i - 1) \Big) \Big/ \Big(\sum_{i=1}^{n} (x_i - 1)^2 \Big)^{1/2} \stackrel{\mathcal{D}}{\longrightarrow} N^* \quad (\text{for } n \to \infty)$$

(convergence in distribution), where N^* is a standard normal random variable. Rewriting (19) as

$$\begin{aligned} & K(x_1, \ldots, x_n) \; = \\ & P\Big(\frac{\sum_{i=1}^n (x_i-1) Z_i - \sum_{i=1}^n (x_i - 1)}{\big(\sum_{i=1}^n (x_i-1)^2\big)^{1/2}} \; - \; \frac{Z_0}{\big(\sum_{i=1}^n (x_i-1)^2\big)^{1/2}} \;\; \le \; - \frac{\sum_{i=1}^n (x_i - 1)}{\big(\sum_{i=1}^n (x_i-1)^2\big)^{1/2}} \Big), \end{aligned}$$

and observing that $Z_0 \Big/ \big(\sum_{i=1}^n (x_i - 1)^2 \big)^{1/2} \to 0$ for $n \to \infty$, we have thus obtained

$$\Big| \, K(x_1, \ldots, x_n) \; - \; \Phi\Big(- \frac{\sum_{i=1}^n (x_i - 1)}{\big(\sum_{i=1}^n (x_i - 1)^2 \big)^{1/2}} \Big) \Big| \stackrel{(n\to\infty)}{\longrightarrow} 0 \; ,$$

where Φ denotes the c.d.f. of the standard normal distribution. In other words, we have for the random variables $K_n = K(X_1, \ldots, X_n)$ that

$$\Big| \, K_n \; - \; \Phi\big(-T_n \big) \Big| \stackrel{(n\to\infty)}{\longrightarrow} 0 \quad P\text{-a.s.} \tag{21}$$

$$\text{where} \quad T_n \; = \; n^{1/2} \frac{\frac{1}{n} \sum_{i=1}^n (X_i - 1)}{\big(\frac{1}{n} \sum_{i=1}^n (X_i - 1)^2 \big)^{1/2}} \; .$$

Since $\frac{1}{n} \sum_{i=1}^n (X_i - 1)^2 \to \sigma^2 = \mathrm{Var}(X_1)$ P-a.s. for $n \to \infty$, it follows again by the CLT that T_n, and hence also $-T_n$, converges in distribution to a standard normal random variable N, say. Thus, by (21),

$$K_n \stackrel{\mathcal{D}}{\longrightarrow} \Phi(N) \quad (\text{for } n \to \infty),$$

and $\Phi(N)$ is standard uniformly distributed.

4 Numerics on the finite sample distribution of K

Here we will numerically examine a stronger question on K than that posed in (13) of Section 2, since we dispense now the assumption of *identical* distributions of the random variables $X_1, \ldots, X_n$. So, let $X_1, \ldots, X_n$ be independent nonnegative random variables with $\mathrm{E}(X_i) \le 1$ for all $i = 1, \ldots, n$; does this entail

$$K(X_1, \ldots, X_n) \overset{\mathcal{D}}{\ge} U \quad \text{(a standard uniform r.v.)} \tag{22}$$

$$\text{i.e.,} \quad P\Big(K(X_1, \ldots, X_n) \le \alpha \Big) \le \alpha \quad \forall\, \alpha \in (\,0\,,\,1\,)\ ? \tag{23}$$

Clearly, since $K(x)$, $x \in [\,0\,,\,\infty)^n$ is a nonincreasing function, it suffices to examine (22) for the case that $\mathrm{E}(X_i) = 1$ for all $i = 1, \ldots, n$. Mathematically, an advantage of not assuming identical distributions of the X_i is that we can further restrict to *two-point* distributed independent nonnegative random variables $X_1, \ldots, X_n$ (with expectations equal to 1). This can be seen from Gaffke (2004), Theorem 2.4 and Lemma 2.5 of that paper, applied to the extremum problem of maximizing (for a fixed α) the probability on the l.h.s. of (23) over all independent nonnegative random variables $X_1, \ldots, X_n$ with expectations equal to 1. So, let $X_1, \ldots, X_n$ be independent random variables with

$$P^{X_i} = \begin{pmatrix} a_i & b_i \\ 1-\lambda_i & \lambda_i \end{pmatrix}, \tag{24}$$

$$\text{where} \quad 0 \le a_i \le 1 < b_i\,, \quad \lambda_i = \frac{1-a_i}{b_i - a_i} \quad (1 \le i \le n)\,. \tag{25}$$

Then, using representation (10) of K, the random variable $K(X_1, \ldots, X_n)$ has a discrete distribution supported by 2^n values

$$\kappa(I) = P\Big(\sum_{i \in I} (b_i - 1) Z_i \le Z_0 + \sum_{j \in I^c} (1 - a_j) Z_j \Big) \quad \forall\, I \subset \{1, \ldots, n\}\,, \tag{26}$$

where according to (10) $Z_0, Z_1, \ldots, Z_n$ are i.i.d. standard exponentially distributed random variables, $I^c = \{1, \ldots, n\} \setminus I$, and the probability given to $\kappa(I)$ is

$$\pi(I) = \Big(\prod_{i \in I} \lambda_i \Big) \Big(\prod_{j \in I^c} (1 - \lambda_j) \Big)\,. \tag{27}$$

The numerical computation of the values $\kappa(I)$ can be done very accurately by using a recursive formula due to Kaminsky, Luks, and Nelson (cf.. Johnson, Kotz, Balakrishnan (1994), p. 553).
To compute the c.d.f. of $K(X_1, \ldots, X_n)$ one has to arrange the values $\kappa(I)$ ($I \subset \{1, \ldots, n\}$) in nondecreasing order, $\kappa_1 \le \ldots \le \kappa_{N-1} \le \kappa_N = 1$, say, where $N = 2^n$, and permute the weights accordingly, $\pi_1, \ldots, \pi_{N-1}, \pi_N$, say. Then (22) to be checked is equivalent to the system of inequalities

$$\sum_{i=1}^{j} \pi_i \le \kappa_j \quad \forall\, j = 1, \ldots, N-1\,. \tag{28}$$

Because of $N = 2^n$ our numerical experiments found their limits when $n = 15$; but within that range in all the (thousands of) instances of input values $a_1, \ldots, a_n \in [\,0\,,\,1\,]$ and $b_1, \ldots, b_n > 1$ the inequalities turned out to be true. A proof (or disproof) is outstanding, unless $n = 2$ (see the lemma below). What we can do further for chosing the input values in a systematic way is to perform an optimization procedure to

$$\text{maximize} \quad P\Big(K(X_1, \ldots, X_n) \le \alpha \Big)$$

over all two-point distributions from (24), (25), (where $X_1, \ldots, X_n$ are stochastically independent), for a given α. We used a heuristics which at each step maximizes that probability w.r.t. one two-point distribution $P^{X_{i_0}}$ while keeping the other $n-1$ two-point distributions P^{X_i}, $i \neq i_0$, fixed; the index i_0 cycles over $1, \ldots, n$. In fact, the maximization w.r.t. one single distribution can be done by using a result of Gaffke (2004), (Theorem 2.2 of that paper). In all the cases ($n \leq 15$, various α, various starting points) the procedure ended up with distributions satisfying (22). Some c.d.f. 's of $K(X_1, \ldots, X_n)$ obtained by the procedure are graphed in Figures 1a/b and 2a/b.

Lemma 4.1 *For $n = 2$,* (28) *and hence* (22) *hold true.*

Proof. For $I = \emptyset, \{1\}, \{2\}, \{1,2\}$ we get from (26) and (27), using the recursion formula of Kaminsky, Luks, and Nelson (cf.. Johnson, Kotz, Balakrishnan (1994), p. 553),

$$\begin{aligned}
&\kappa(\emptyset) = 1 \ , \quad \pi(\emptyset) = (1-\lambda_1)(1-\lambda_2) = \tfrac{(b_1-1)(b_2-1)}{(b_1-a_1)(b_2-a_2)} \ , \\
&\kappa(\{1\}) = \tfrac{(b_1-1)(1-a_2)}{b_1(b_1-a_2)} + \tfrac{1}{b_1} \ , \quad \pi(\{1\}) = \lambda_1(1-\lambda_2) = \tfrac{(1-a_1)(b_2-1)}{(b_1-a_1)(b_2-a_2)} \ , \\
&\kappa(\{2\}) = \tfrac{(b_2-1)(1-a_1)}{b_2(b_2-a_1)} + \tfrac{1}{b_2} \ , \quad \pi(\{2\}) = (1-\lambda_1)\lambda_2 = \tfrac{(b_1-1)(1-a_2)}{(b_1-a_1)(b_2-a_2)} \ , \\
&\kappa(\{1,2\}) = \tfrac{1}{b_1 b_2} \ , \quad \pi(\{1,2\}) = \lambda_1\lambda_2 = \tfrac{(1-a_1)(1-a_2)}{(b_1-a_1)(b_2-a_2)} \ .
\end{aligned}$$

From (26) we see also that $\kappa(\{1,2\}) \leq \kappa(\{1\}), \kappa(\{2\})$; w.l.g. we may assume that $\kappa(\{1\}) \leq \kappa(\{2\})$ (otherwise interchange the random variables X_1 and X_2). So the nondecreasing arrangement of the κ-values, $\kappa_1 \leq \kappa_2 \leq \kappa_3 \leq \kappa_4$, is $\kappa_1 = \kappa(\{1,2\})$, $\kappa_2 = \kappa(\{1\})$, $\kappa_3 = \kappa(\{2\})$, and $\kappa_4 = \kappa(\emptyset) = 1$, and the corresponding weights $\pi_1, \pi_2, \pi_3, \pi_4$ are defined accordingly. We have to show that the following three inequalities hold true,

$$\pi_1 \leq \kappa_1 \ , \quad \pi_1 + \pi_2 \leq \kappa_2 \ , \quad \pi_1 + \pi_2 + \pi_3 \leq \kappa_3 \ . \tag{29}$$

The first inequality holds true because of

$$\pi_1 \ = \ \tfrac{(1-a_1)(1-a_2)}{(b_1-a_1)(b_2-a_2)} \leq \tfrac{1}{b_1 b_2} = \kappa_1 \ ,$$

where we have used that for any given $b > 1$ the ratio $(1-a)/(b-a)$ is decreasing in $a \in [\,0\,,\,1\,]$. The second inequality of (29) follows from

$$\pi_1 + \pi_2 \ = \ \tfrac{(1-a_1)(1-a_2)+(1-a_1)(b_2-1)}{(b_1-a_1)(b_2-a_2)} \ = \ \tfrac{1-a_1}{b_1-a_1} \ \leq \ \tfrac{1}{b_1} \ \leq \ \kappa_2 \ .$$

For a proof of the third inequality in (29) we note that, firstly, $\pi_1 + \pi_2 + \pi_3 = 1 - \pi_4 = 1 - \frac{(b_1-1)(b_2-1)}{(b_1-a_1)(b_2-a_2)}$ and, secondly for any given $a \in [\,0\,,\,1\,]$ the ratio $(b-1)/(b-a)$ is nondecreasing in $b > 1$. We distinguish two cases.

Case 1: $b_1 \leq b_2$. Then,

$$1 - \pi_4 \leq 1 - \tfrac{(b_1-1)(b_2-1)}{b_1(b_2-a_2)} \leq 1 - \tfrac{(b_1-1)^2}{b_1(b_1-a_2)} = \tfrac{(b_1-1)(1-a_2)}{b_1(b_1-a_2)} + \tfrac{1}{b_1} = \kappa_2 \leq \kappa_3 \, .$$

Case 2: $b_1 > b_2$. Then,

$$1 - \pi_4 \leq 1 - \tfrac{(b_1-1)(b_2-1)}{(b_1-a_1)b_2} \leq 1 - \tfrac{(b_2-1)^2}{(b_2-a_1)b_2} = \tfrac{(b_2-1)(1-a_1)}{(b_2-a_1)b_2} + \tfrac{1}{b_2} = \kappa_3 \, .$$

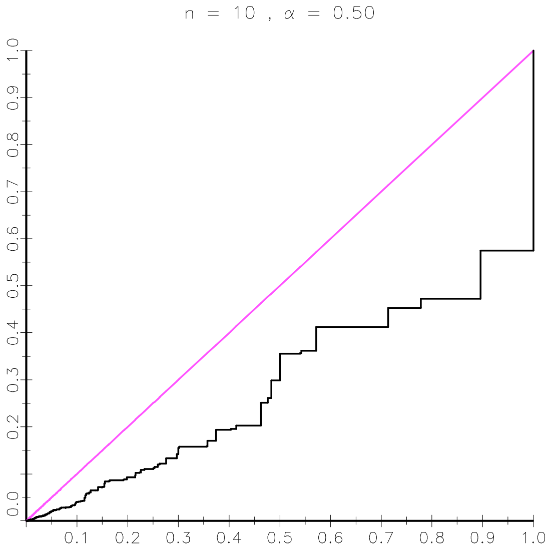

Fig. 1a c.d.f. of K after 1 iteration ($n = 10$, $\alpha = 0.50$)

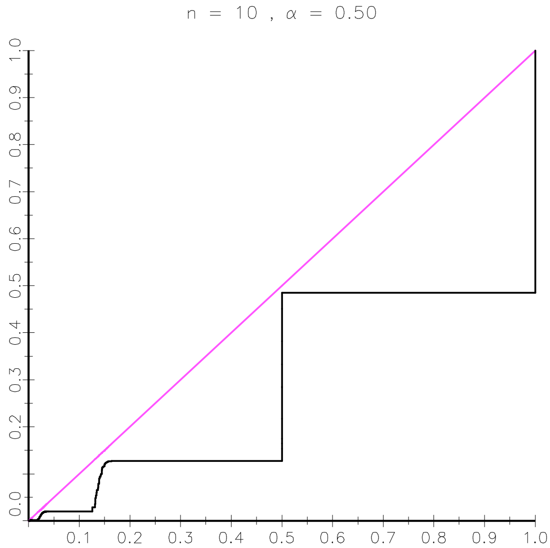

Fig. 1b c.d.f. of K after 20 iterations ($n = 10$, $\alpha = 0.50$)

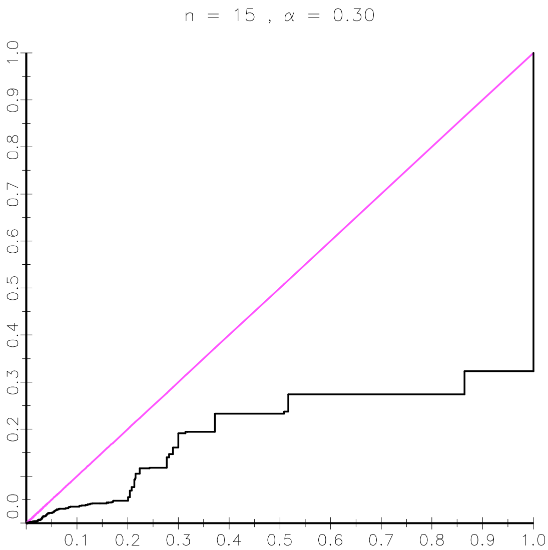

Fig. 2a c.d.f. of K after 1 iteration ($n = 15$, $\alpha = 0.30$)

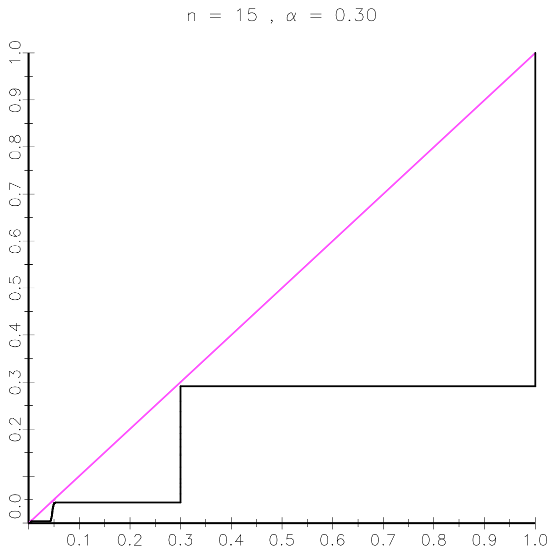

Fig. 2b c.d.f. of K after 30 iterations ($n = 15$, $\alpha = 0.30$)

A Appendix: Proof of eq. (7)

In the following we represent the given observations $x_1, \ldots, x_n \in [\,0\,,\,\infty)$ by

$$\begin{pmatrix} z_1 & \ldots & z_r \\ n_1 & \ldots & n_r \end{pmatrix},$$

where r is the number of distinct values among the x_i, $z_1 < \ldots < z_r$ are the increasingly ordered distinct values of the x_i, and n_j is the multiplicity of the value z_j among the x_i, for $j = 1, \ldots, r$. Clearly, $\sum_{j=1}^r n_j = n$. So the nonparametric likelihood from (6), for a discrete probability distribution Q on $[\,0\,,\,\infty)$, rewrites as

$$L(x, Q) \;=\; \prod_{j=1}^r Q(z_j)^{n_j} \;.$$

It is easily seen that the denominator of the ratio $R(x)$ from (5) equals

$$\sup\Big\{ \prod_{j=1}^r \lambda_j^{n_j} \;:\; \lambda_j > 0 \;(1 \le j \le r)\,,\; \sum_{j=1}^r \lambda_j \le 1 \Big\}\,, \tag{30}$$

and the numerator in (5) equals

$$\sup\Big\{ \prod_{j=1}^r \lambda_j^{n_j} \;:\; \lambda_j > 0 \;(1 \le j \le r)\,,\; \sum_{j=1}^r \lambda_j \le 1\,,\; \sum_{j=1}^r \lambda_j z_j \le 1 \Big\}\,. \tag{31}$$

Taking logarithms,

$$\ln\Big(\prod_{j=1}^r \lambda_j^{n_j} \Big) \;=\; \sum_{j=1}^r n_j \ln(\lambda_j)\;,$$

it is easy to see that the sup in (30) is attained for $\lambda_j \;=\; n_j/n$ $(1 \le j \le r)$ which gives $\prod_{j=1}^r \big(n_j/n\big)^{n_j}$ for the denominator (30) of $R(x)$. To calculate the numerator of $R(x)$, i.e., the sup in (31), we have to solve the extremum problem,

$$\text{maximize} \quad f(\lambda) \;=\; \sum_{j=1}^r n_j \ln(\lambda_j) \quad \text{s.t.} \quad \lambda = (\lambda_1, \ldots, \lambda_r) \in C\,, \tag{32}$$

$$\text{where} \quad C \;=\; \Big\{ \lambda \in (\,0\,,\,\infty)^r \;:\; \sum_{j=1}^r \lambda_j \le 1\,,\; \sum_{j=1}^r \lambda_j z_j \le 1 \Big\}\,.$$

The objective function f is differentiable and concave on $(\,0\,,\,\infty)^r$ and C is a convex subset of that domain. So, a feasible point $\lambda^* \in C$ is an optimal solution to (32) if and only if all the directional derivatives of f at λ^* into feasible directions are nonpositive, i.e.,

$$\sum_{j=1}^r \frac{n_j}{\lambda_j^*} \big(\lambda_j - \lambda_j^*\big) \;\le\; 0 \quad \forall\, \lambda \in C\,,$$

(see Rockafellar (1972), Theorem 27.4), or equivalently,

$$\sum_{j=1}^n \frac{n_j}{n} \frac{1}{\lambda_j^*} \lambda_j \;\le\; 1 \quad \forall\, \lambda \in C\;. \tag{33}$$

We will verify that the point λ^* given next belongs to C and satisfies (33) (and is thus an optimal solution to (32)), where we have to distinguish three cases.

(1) If $\sum_{j=1}^{r} \frac{n_j}{n} z_j \le 1$, then $\lambda_j^* = \frac{n_j}{n}$, $(1 \le j \le r)$;

(2) if $\sum_{j=1}^{r} \frac{n_j}{n} \frac{1}{z_j} \le 1$, then $\lambda_j^* = \frac{n_j}{n} \frac{1}{z_j}$, $(1 \le j \le r)$;

(3) if $\sum_{j=1}^{r} \frac{n_j}{n} z_j > 1$ and $\sum_{j=1}^{r} \frac{n_j}{n} \frac{1}{z_j} > 1$, then

$$\lambda_j^* = \frac{n_j}{n} (1 - t^* + t^* z_j)^{-1}, \quad (1 \le j \le r),$$

where $0 < t^* < 1$ is the unique solution to the equation

$$\sum_{j=1}^{r} n_j \frac{z_j - 1}{1 - t + t z_j} = 0, \quad t \in [0, 1).$$

Note that if $z_1 = 0$ we define $\sum_{j=1}^{r} \frac{n_j}{n} \frac{1}{z_j} = \infty$; in case (3) we have

$$\sum_{j=1}^{r} n_j \frac{z_j - 1}{1 - t + t z_j} = \frac{\mathrm{d}}{\mathrm{d}t} \sum_{j=1}^{r} n_j \ln\bigl(1 - t + t z_j\bigr),$$

which is a dereasing continuous function of $t \in [0, 1)$, and for $t = 0$ and $t \to 1$, resp., it gives the values $\sum_{j=1}^{r} n_j z_j - n > 0$ and $n - \sum_{j=1}^{r} n_j z_j^{-1} < 0$; so, in fact, that function has a unique root $t^* \in (0, 1)$.

<u>Case (1).</u> We have $\sum_{j=1}^{r} \frac{n_j}{n} = 1$ and $\sum_{j=1}^{r} \frac{n_j}{n} z_j \le 1$, hence $\lambda^* \in C$. For any $\lambda \in C$ we obtain

$$\sum_{j=1}^{r} \frac{n_j}{n} \frac{1}{\lambda_j^*} \lambda_j = \sum_{j=1}^{r} \lambda_j \le 1,$$

and so λ^* satisfies (33).

<u>Case (2).</u> We have $\sum_{j=1}^{r} \frac{n_j}{n} z_j^{-1} \le 1$ and $\sum_{j=1}^{r} \frac{n_j}{n} z_j^{-1} z_j = 1$, hence $\lambda^* \in C$, and for any $\lambda \in C$ we obtain

$$\sum_{j=1}^{r} \frac{n_j}{n} \frac{1}{\lambda_j^*} \lambda_j = \sum_{j=1}^{r} z_j \lambda_j \le 1,$$

and so λ^* satisfies (33).

<u>Case (3).</u> By definition of t^* and the λ_j^* we have

$$\sum_{j=1}^{r} z_j \lambda_j^* = \sum_{j=1}^{r} \frac{n_j}{n} \frac{z_j}{1 - t^* + t^* z_j} = \sum_{j=1}^{r} \frac{n_j}{n} \frac{1}{1 - t^* + t^* z_j} = \sum_{j=1}^{r} \lambda_j^*,$$

and from $t^* z_j \lambda_j^* + (1 - t^*) \lambda_j^* = n_j / n$ $(1 \le j \le r)$,

$$t^* \sum_{j=1}^{r} z_j \lambda_j^* + (1 - t^*) \sum_{j=1}^{r} \lambda_j^* = \sum_{j=1}^{r} \frac{n_j}{n} = 1.$$

Hence it follows that $\sum_{j=1}^{r} z_j\lambda_j^* = \sum_{j=1}^{r} \lambda_j^* = 1$, and thus $\lambda^* \in C$. For any $\lambda \in C$ we obtain

$$\sum_{j=1}^{r} \frac{n_j}{n}\frac{1}{\lambda_j^*}\lambda_j = (1-t^*)\sum_{j=1}^{r}\lambda_j + t^*\sum_{j=1}^{r} z_j\lambda_j \le 1 ,$$

and so λ^* satisfies (33).

We can summarize the results of the above three cases by defining $t^* = 0$ in case (1) and $t^* = 1$ in case (2). Then in either cases we can write the optimal solution to (32) as $\lambda_j^* = (n_j/n)\,(1-t^*+t^*z_j)^{-1}$ $(1 \le j \le r)$. Now, the function of $t \in [\,0\,,\,1\,]$ given by

$$h(t) = -\sum_{j=1}^{r} n_j \ln\bigl(1-t+tz_j\bigr)$$

is convex (where in case $z_1 = 0$ its value at $t = 1$ is defined to be ∞). It is easily verified by considering its derivative on $[\,0\,,1\,)$,

$$h'(t) = -\sum_{j=1}^{r} n_j\,\frac{z_j-1}{1-t+tz_j} ,$$

that in either of the three cases t^* is the minimizer of h over $t \in [\,0\,,\,1\,]$. We have therefore

$$\prod_{j=1}^{r}(1-t^*+t^*z_j)^{-n_j} = \min_{0\le t\le 1}\prod_{j=1}^{r}(1-t+tz_j)^{-n_j} ,$$

and for the likelihood ratio $R(x)$ we have,

$$R(x) = \Bigl(\prod_{j=1}^{r}\lambda_j^{*n_j}\Bigr)\Big/\Bigl(\prod_{j=1}^{r}(n_j/n)^{n_j}\Bigr) = \prod_{j=1}^{r}\Bigl(\frac{n\,\lambda_j^*}{n_j}\Bigr)^{n_j}$$

$$= \prod_{j=1}^{r}(1-t^*+t^*z_j)^{-n_j} = \min_{0\le t\le 1}\prod_{j=1}^{r}(1-t+tz_j)^{-n_j} .$$

Clearly, $\prod_{j=1}^{r}(1-t+tz_j)^{-n_j} = \prod_{i=1}^{n}(1-t+tx_i)^{-1}$, which completes the proof of (7).

References

Bauer, H., 1996. Probability Theory. De Gruyter, Berlin.

Billingsley, P., 1995. Probability and Measure. Wiley, New York.

Breth, M., 1976. Non-parametric confidence intervals for a mean using censored data. J. Roy. Statist. Soc. B 38, 251-254.

Breth, M., Maritz, J.S., Williams, E.J., 1978. On distribution-free lower confidence limits for the mean of a nonnegative random variable. Biometrika 65, 529-534.

David, H.A., 1981. Order Statistics. Wiley, New York.

Gaffke, N., 2004. Nonparametric one-sided testing for the mean and related extremum problems. Preprint No. 6 (2004), Faculty of Mathematics, University of Magdeburg.

Gaffke, N., Zöllner, A., 2003. A resampling approach for under-estimating a finite population total from a censored sample. Comm. Statist.–Theory Meth. 32, 2305-2320.

Kaplan, H.M., 1987. A method of one-sided nonparametric inference for the mean of a nonnegative population. Amer. Statistician 41, 157-158.

Johnson, N.L., Kotz, S., Balakrishnan, N., 1994. Continuous Univariate Distributions. Volume 1. Wiley, New York.

Rockafellar, R.T., 1972. Convex Analysis. Princeton Univ. Press.

Wang, W., Zhao, L.H., 2003. Nonparametric tests for the mean of a non-negative population. J. Statist. Plann. Inference 110, 75-96.