SA

stichting
mathematisch
centrum

MC

AFDELING MATHEMATISCHE STATISTIEK SW 33/74 AUGUST

P. GROENEBOOM, Y. LEPAGE and F.H. RUYMGAART

RANK TESTS FOR INDEPENDENCE WITH BEST STRONG EXACT BAHADUR SLOPE

PREPUBLICATION

2e boerhaavestraat 49 amsterdam

Printed at the Mathematical Centre, 49, 2e Boerhaavestraat, Amsterdam.

The Mathematical Centre, founded the 11-th of February 1946, is a non-profit institution aiming at the promotion of pure mathematics and its applications. It is sponsored by the Netherlands Government through the Netherlands Organization for the Advancement of Pure Research (Z.W.O), by the Municipality of Amsterdam, by the University of Amsterdam, by the Free University at Amsterdam, and by industries.

AMS (MOS) Classification scheme (1970): 62G20

RANK TESTS FOR INDEPENDENCE WITH BEST STRONG EXACT BAHADUR SLOPE

by

P. Groeneboom , Y. Lepage(*) and F.H. Ruymgaart
Mathematisch Centrum, Amsterdam

ABSTRACT

Hájek [2] has shown that there are rank statistics for the two-sample problem which have the best possible strong exact slope. Here we prove the same kind of result in the case of rank statistics for testing independence against a general fixed dependence alternative. Our proof is also based on a strong law of large numbers for the rank statistics involved. Moreover, in the appendix we prove that the crucial property A of Woodworth [4] is satisfied for both exact and approximate score functions, derived from a suitable function on the open unit square. This function has to satisfy rather mild integrability conditions, may have certain discontinuities and need not be of product type.

(*) This work was done while the author was on leave from the Université de Montréal and visiting the Mathematisch Centrum, Amsterdam, with a post-doctoral fellowship from the National Research Council of Canada.

1. INTRODUCTION

Let $(\Omega,\mathcal{A},P)$ be a probability space on which a pair (X,Y) of random variables (rvs) is defined, having joint distribution function (df) $H(x,y) = P(\{X \le x, Y \le y\})$ and marginal dfs $F(x) = P(\{X \le x\})$ and $G(y) = P(\{Y \le y\})$ for all $x,y \in (-\infty,\infty)$. Let be given a sequence of mutually independent and identically distributed (iid) random vectors (X_1,Y_1), $(X_2,Y_2),\ldots$, all defined on the probability space mentioned above and all possessing the bivariate df H. To display the underlying df H, the probability measure will occasionally be denoted by P_H rather than P.

Given a positive interger N, the joint empirical df based on the first N random vectors in the sequence is defined by $NH_N(x,y) = \#\{(X_n,Y_n) : X_n \le x,\ Y_n \le y,\ n = 1,\ldots,N\}$ and its marginal empirical dfs $F_N(x)$ and $G_N(y)$ by $NF_N(x) = \#\{X_n : X_n \le x,\ n = 1,\ldots,N\}$ respectively $NG_N(y) = \#\{Y_n : Y_n \le y,\ n = 1,\ldots,N\}$. (For any finite set S we denote the number of its elements by # S.) The rank R_{nN} of X_n will be defined as $\#\{X_m : X_m \le X_n,\ m = 1,\ldots,N\}$ and the rank Q_{nN} of Y_n as $\#\{Y_m : Y_m \le Y_n,\ m = 1,\ldots,N\}$. The set of ordered first and second coordinates will be denoted by $X_{1:N} \le \ldots \le X_{N:N}$ and $Y_{1:N} \le \ldots \le Y_{N:N}$ respectively.

For any non-decreasing right-continuous function Ψ on $(-\infty,\infty)$, satisfying $\lim_{z\to-\infty}\Psi(z) = 0$ and $\lim_{z\to\infty}\Psi(z) = 1$ let us define an inverse Ψ^{-1} on $(0,1]$ of this function by

$$\Psi^{-1}(u) = \inf_{z\in(-\infty,\infty)} \{z : \Psi(z) \ge u\} \text{ for } u \in (0,1].$$

When $\{z : \Psi(z) \ge 1\} = \emptyset$ by convention we define $\Psi^{-1}(1) = \infty$. By this definition Ψ^{-1} is left-continuous on $(0,1]$. We obviously have the useful relations

$$(1.1)\qquad F_N(X_n) = R_{nN}/N,\quad G_N(Y_n) = Q_{nN}/N,$$

$$(1.2)\qquad F_N^{-1}(n/N) = X_{n:N},\ G_N^{-1}(n/N) = Y_{n:N},$$

for $n = 1,\ldots,N$.

The rank statistics that will be considered here are suitable for testing the hypothesis of independence and have the form

(1.3) $$T_N = N^{-1} \sum_{n=1}^{N} C_N(R_{nN}, Q_{nN}),$$

where the numbers $C_N(m,n)$, called scores, are defined and finite for $m,n=1,\ldots,N$ and $N = 1,2,\ldots$ Let us define a score function J_N on $(0,1] \times (0,1]$ by

(1.4) $$J_N(s,t) = C_N(m,n) \text{ for } (s,t) \in ((m-1)/N, m/N] \times ((n-1)/N, n/N],$$

$m,n = 1,\ldots,N$. It follows from (1.1) that we may write T_N alternatively as

(1.5) $$T_N = \iint_{x,y\in(-\infty,\infty)} J_N(F_N(x), G_N(y))dH_N(x,y) = \iint J_N(F_N, G_N)dH_N.$$

Still another representation of T_N may be obtained in terms of the modified bivariate empirical df $\bar{H}_N$, defined on $(0,1] \times (0,1]$ by

(1.6) $$\bar{H}_N(s,t) = H_N(F_N^{-1}(s), G_N^{-1}(t)) \text{ for } (s,t) \in (0,1] \times (0,1].$$

We may think of $\bar{H}_N$ as a scalefree version of H_N which assigns mass 1 to $(0,1] \times (0,1]$ and has the property that $N\bar{H}_N(s,t) = \#\{(R_{nN}, Q_{nN}) : R_{nN}/N \le s, Q_{nN}/N \le t, n=1,\ldots,N\}$. Note that (with P_H-probability 1) $\bar{H}_N(n/N,0) = \bar{H}_N(0,n/N) = 0$ and $\bar{H}_N(n/N,1) = \bar{H}_N(1,n/N) = n/N$ for $n = 1,\ldots,N$. Combining (1.2), (1.5) and (1.6) it follows that

(1.7) $$T_N = \iint_{s,t\in(0,1]} J_N(s,t)d\bar{H}_N(s,t).$$

Similarly let us introduce

(1.8) $$\bar{H}(s,t) = H(F^{-1}(s), G^{-1}(t)) \text{ for } (s,t) \in (0,1] \times (0,1],$$

and observe that $\bar{H}(s,t) = P(\{F(X) \le s, G(Y) \le t\})$ so that it assigns mass 1 to the unit square and has uniform (0,1) marginal dfs.

Hájek [2] investigated the almost sure convergence of linear rank statistics for the two-sample problem and exhibited the existence of linear rank statistics with best strong exact slope for testing against a fixed simple alternative. It is our purpose to prove similar results for statistics of type (1.7). In section 2 we prove the almost sure convergence of the random variables (1.7) and section 3 is devoted to the construction of linear rank statistics - i.e. statistics of type (1.3) or, equivalently, (1.7) - with best strong exact slope. Some results concerning the best strong exact slope of rank-likelihood ratio statistics are presented in section 4. Finally, in the appendix we prove that the conditions for theorem 2.1 and theorem 3.1, one of which is the crucial property A of Woodworth [4], are satisfied in the case where the J_N are either the exact or the approximate score functions, derived from a sufficiently smooth function J on the open unit interval.

2. ALMOST SURE CONVERGENCE

It will be convenient to introduce the class $\mathcal{H}$ of bivariate dfs, defined by

$$\mathcal{H} = \{H{:}H \text{ is a bivariate df, continuous on } (-\infty,\infty) \times (-\infty,\infty)\}.$$

For any $H \in \mathcal{H}$ it follows that

$$(2.1) \qquad P_H\left(\left\{\lim_{N\to\infty} \sup_{s,t\in(0,1]} |\bar{H}_N(s,t)-\bar{H}(s,t)| = 0\right\}\right) = 1.$$

To see this let us observe that, with probability 1 (under P_H),

$$\begin{aligned}
&\sup_{s,t\in(0,1]} |\bar{H}_N(s,t)-\bar{H}(s,t)| \le \\
&\le \sup_{s,t\in(0,1]} \Big| H_N\Big(F^{-1}\big(F(F_N^{-1}(s))\big),G^{-1}\big(G(G_N^{-1}(t))\big)\Big) \\
&\qquad - H\Big(F^{-1}\big(F(F_N^{-1}(s))\big),G^{-1}\big(G(G_N^{-1}(t))\big)\Big)\Big| \\
&+ \sup_{s,t\in(0,1]} |\bar{H}\big(F(F_N^{-1}(s)),G(G_N^{-1}(t))\big)-\bar{H}(s,t)|.
\end{aligned}$$

The first term in this bound converges to 0 with probability 1 by the Glivenko-Cantelli theorem in two dimensions. Because $F(F_N^{-1})$ and $G(G_N^{-1})$ behave on $(0,1]$ as inverse empirical dfs based on random samples from the uniform $(0,1)$ distribution and since $\bar{H}$ is uniformly continuous on $[0,1] \times [0,1]$ the second term in this bound also converges to 0 with probability 1. This proves (2.1).

In the sequel $\phi \geq 0$ and $\psi \geq 0$ will be continuous functions on $(0,1)$, satisfying

$$(2.2) \qquad \int_0^1 [\phi(s)]^\xi ds < \infty, \int_0^1 [\psi(t)]^\eta dt < \infty \text{ for some } \xi,\eta \geq 1 \text{ with } \xi^{-1}+\eta^{-1} = 1.$$

We shall also need the functions ϕ_N and ψ_N on $(0,1]$, related to ϕ and ψ according to the relations

$$(2.3) \qquad \begin{aligned} &\phi_N = \phi \text{ on } (0,1-N^{-1}],\ \phi_N = \phi(1-N^{-1}) \text{ on } (1-N^{-1},1], \\ &\psi_N = \psi \text{ on } (0,1-N^{-1}],\ \psi_N = \psi(1-N^{-1}) \text{ on } (1-N^{-1},1]. \end{aligned}$$

For any real function f, by $f^{(\tau)}$ we shall understand the truncated function

$$(2.4) \qquad f^{(\tau)} = f \text{ on } \{|f| \leq \tau\},\ f^{(\tau)} = 0 \text{ on } \{|f| > \tau\}.$$

If f is a function of two variables defined on a bounded rectangle $[a,A] \times [b,B]$ in the plane, for any two partitions $a = a_0 \leq a_1 \leq \dots \leq a_N = A$ and $b = b_0 \leq b_1 \leq \dots \leq b_N = B$, let us put

$$(2.5) \qquad \Delta_{m,n} f = f(a_m,b_n) - f(a_{m-1},b_n) + f(a_{m-1},b_{n-1}) - f(a_m,b_{n-1}).$$

The total variation of f on $[a,A] \times [b,B]$ is defined by

$$(2.6) \qquad V_{[a,A]\times[b,B]}(f) = \sup \textstyle\sum_{m=1}^N \sum_{n=1}^N |\Delta_{m,n} f|,$$

the supremum being taken over all pairs of partitions.

If J_N is defined as in (1.4) the condition $|J_N(s,t)| \leq \phi(s)\psi(t)$ for all

$s,t \in (0,1)$ is equivalent to the condition that $|J_N(s,t)| \le \phi_N(s)\psi_N(t)$ for all $s,t \in (0,1]$, see (2.3). Each of these conditions on the J_N entails that (for $H \in \mathcal{H}$)

$$(2.7) \qquad P_H(\{\lim_{\tau\to\infty} \sup_{N=1,2,\ldots} \iint |J_N - J_N^{(\tau)}| d\bar{H}_N = 0\}) = 1,$$

$$(2.8) \qquad \lim_{\tau\to\infty} \sup_{H\in\mathcal{H}, N=1,2,\ldots} \iint |J_N - J_N^{(\tau)}| d\bar{H} = 0.$$

Both (2.7) and (2.8) follow from Hölder's inequality. By way of an example let us prove (2.7) and note that

$$\iint_{\{|J_N|>\tau\}} |J_N| d\bar{H}_N \le \iint_{\{\phi_N\psi_N>\tau\}} \phi_N\psi_N d\bar{H}_N \le$$

$$\le \iint_{\{\phi_N>\tau^{\frac{1}{2}}\}\times(0,1]} \phi_N\psi_N d\bar{H}_N + \iint_{(0,1]\times\{\psi_N>\tau^{\frac{1}{2}}\}} \phi_N\psi_N d\bar{H}_N,$$

for each $\tau \ge 0$. By symmetry we need only consider the first term in the latter bound. Applying Hölder's inequality we see that the supremum over all $N = 1,2,\ldots$ of the first term is bounded by

$$\sup_{N=1,2,\ldots} \left\{\int_{\{\phi_N>\tau^{\frac{1}{2}}\}} [\phi_N(s)]^{\xi} d\bar{H}_N(s,1)\right\}^{1/\xi}$$

$$\times \left\{\int_{(0,1]} [\psi_N(t)]^{\eta} d\bar{H}_N(1,t)\right\}^{1/\eta} \to 0 \text{ with probability 1, as } \tau \to \infty,$$

because $\bar{H}_N(\cdot,1)$ and $\bar{H}_N(1,\cdot)$ are discrete probability measures, no longer depending on $\omega \in \Omega$ except for a set with P_H-measure 0, restricted to the points n/N for $n = 1,\ldots,N$ to each of which they assign mass $1/N$.

We also have to introduce a measurable function J on $(0,1) \times (0,1)$ which is, in a sense to be made precise, the limit of the functions J_N and therefore referred to as limiting score function. The condition that $|J(s,t)| \le \phi(s)\psi(t)$ for all $s,t \in (0,1)$ entails

$$(2.9) \qquad \lim_{\tau\to\infty} \sup_{H\in\mathcal{H}} \iint |J - J^{(\tau)}| d\bar{H} = 0 \text{ and } \sup_{H\in\mathcal{H}} \iint |J| d\bar{H} < \infty,$$

in a completely similar way.

THEOREM 2.1. *Suppose the following conditions are satisfied:*

(a) $|J_N(s,t)|$, $|J(s,t)| \le \phi(s)\psi(t)$ *for* $s,t \in (0,1)$ *and* $N = 1,2,\ldots$ *(see (2.2))*,

(b) $\lim_{N\to\infty} |\iint (J_N - J)\,d\bar{H}| = 0$,

(c) $\sup_{N=1,2,\ldots} V_{[1/N,1]\times[1/N,1]}(J_N^{(\tau)}) \le V(\tau) < \infty$ *for each* $\tau \ge 0$.

Then we have, for each fixed $H \in \mathcal{H}$,

$$P_H(\{\lim_{N\to\infty} T_N = \iint J d\bar{H}\}) = 1. \tag{2.10}$$

PROOF. Let us first observe that

$$|T_N - \iint J d\bar{H}| \le |\iint (J_N - J)d\bar{H}| + |\iint (J_N - J_N^{(\tau)})d\bar{H}_N|$$
$$+ |\iint (J_N^{(\tau)} - J_N)d\bar{H}| + |\iint J_N^{(\tau)} d(\bar{H}_N - \bar{H})|, \tag{2.11}$$

for each $\tau \ge 0$. Condition (a) implies (2.7) and (2.8) so that, on account of condition (b) it suffices to prove that for each $\tau \ge 0$ the last term on the right in (2.11) converges to 0 with probability 1 (under P_H), as $N \to \infty$.

By definition $J_N^{(\tau)}$ is constant on each square $((m-1)/N,m/N] \times ((n-1)/N,n/N]$ for $m,n = 1,\ldots,N$. Because $\bar{H}_N(n/N,0) - \bar{H}(n/N,0) = \bar{H}_N(n/N,1) - \bar{H}(n/N,1) = \bar{H}_N(0,n/N) - \bar{H}(0,n/N) = \bar{H}_N(1,n/N) - \bar{H}(1,n/N) = 0$ with probability 1 for $n = 1,\ldots,N$, it follows that

$$|\iint J_N^{(\tau)} d(\bar{H}_N - \bar{H})| = \tag{2.12}$$
$$= |\textstyle\sum_{m=1}^{N-1} \sum_{n=1}^{N-1} [\bar{H}_N(m/N,n/N) - \bar{H}(m/N,n/N)]\Delta_{m+1,n+1}\, J_N^{(\tau)}| \le$$
$$\le \sup_{s,t\in(0,1]} |\bar{H}_N(s,t) - \bar{H}(s,t)|V(\tau) \to 0, \text{ with probability 1, as } N \to \infty.$$

Here we use (2.1), which holds by continuity of H, and condition (c). The difference operator $\Delta_{m+1,n+1}$, defined in (2.5), refers to the partitions $\{m/N, m=1,\ldots,N\}$ and $\{n/N, n=1,\ldots,N\}$ of $[1/N,1]$. □

The theorem, properly modified, remains true in some situations where the sample elements do not all have the same underlying bivariate df. More specifically this is the case if, for some fixed positive integer k, $N = N_1 + \ldots + N_k$ and N_i sample elements have df $H_{(i)} \in \mathcal{H}$, provided $N_i/N \to \nu_i \in [0,1]$ as $N \to \infty$ and $\bar{H}$ in (2.10) is replaced by $\bar{H}_{(\nu)}$, where $H_{(\nu)} = \sum_{i=1}^{k} \nu_i H_{(i)}$. The bivariate empirical df H_N and, consequently, its marginals F_N and G_N are based on the combined sample in this case.

Let us specialize this generalization to the case where $k = 2$ and $H_{(i)}(x,y) = F_{(i)}(x)U_{(i)}(y)$, where $F_{(i)}$ is a continuous univariate df and $U_{(i)}$ is the uniform df on $(i-1,i)$ for $i = 1,2$. Choosing, in addition $J_N(s,t) = K_N(s)\chi_N(t)$, where $\chi_N(t) = 1$ for $t \in (0,N_1/N]$ and $\chi_N(t) = 0$ for $t \in (N_1/N,1]$, (1.7) reduces to the (1-dim.) two-sample statistic, because

$$(2.13) \qquad P\left(\left\{T_N = N^{-1} \sum_{n:X_n \text{ has df } F_{(1)}} K_N(R_{nN}/N)\right\}\right) = 1.$$

For the limiting score function we take $J(s,t) = K(s)\chi_\nu(t)$, where $\chi_\nu(t) = 1$ for $t \in (0,\nu_1]$ and $\chi_\nu(t) = 0$ for $t \in (\nu_1,1)$. In order that the J_N and J satisfy condition (a) of the theorem it suffices that the $|K_N|$ and $|K|$ are bounded by a continuous function ϕ satisfying (2.2) with $\xi = 1$ because the χ_N and χ_ν are bounded by 1, so that we may take $\psi = 1$ on $(0,1)$ and hence $\eta = \infty$. Conditions (b) and (c) can accordingly be modified. In view of (2.13) we obtain

$$(2.14) \qquad P\left(\left\{\lim_{N\to\infty} N^{-1} \sum_{n:X_n \text{ has df } F_{(1)}} K_N(R_{nN}/N) = \nu_1 \int K dF_{(1)}(F_{(\nu)}^{-1})\right\}\right) = 1,$$

where $F_{(1)}$ and $F_{(\nu)}$ are the first marginals of $H_{(1)}$ and $H_{(\nu)}$ respectively. This is essentially theorem 1 of Hájek [2].

It should be noted that if $H \in \mathcal{H}$ has a density with respect to Lebesgue measure, theorem 2.1 still holds if condition (b) is replaced by

$$(b') \quad \lim_{N\to\infty} \int_0^1\int_0^1 |J_N(s,t)-J(s,t)|dsdt = 0.$$

For special choices of J_N and J satisfying conditions (a) and (b) of the theorem we refer to the appendix.

3. BEST STRONG EXACT SLOPE

In this section we restrict our attention to a subclass $\mathcal{H}^*$ of smooth dfs. If we say that a df has a density, we shall tacitly assume that it is a density with respect to Lebesgue measure. Similarly, if we say that a relation holds a.e. this is also with respect to Lebesgue measure. The subclass is given by

$$(3.1) \qquad \mathcal{H}^* = \{H \in \mathcal{H} : H \text{ has a density } h \text{ on } (-\infty,\infty) \times (-\infty,\infty)\}.$$

We shall also consider the special families

$$(3.2) \qquad \mathcal{H}_0^* = \{H \in \mathcal{H}^* : H = F{\times}G\},\ \mathcal{H}_1^* = \mathcal{H}^* - \mathcal{H}_0^* \ .$$

For any $H \in \mathcal{H}^*$ the univariate marginal dfs F and G possess densities f and g respectively, and the transformed df $\bar{H}$ has a density $\bar{h}$ with support contained in $(0,1) \times (0,1)$. Each $F \times G \in \mathcal{H}_0^*$ has the transformed df $\bar{H}_{(0)}(s,t) = st$ for all $s,t \in (0,1)$ with density $\bar{h}_{(0)} = 1$, a.e. on $(0,1) \times (0,1)$.

From now on let us fix $H_{(1)} \in \mathcal{H}_1^*$ with density $h_{(1)}$. For each N we wish to test, on the basis of the sample $(X_1,Y_1),\ldots,(X_N,Y_N)$, the composite hypothesis that the underlying df H satisfies $H \in \mathcal{H}_0^*$ against the simple alternative $H = H_{(1)}$. The probability measure on (Ω,A) corresponding to $F \times G \in \mathcal{H}_0^*$ will be denoted by $P_{F\times G}$, or simply by P_0 if the probability of an event involving ranks only is considered. The probability measure corresponding to $H_{(1)}$ is denoted by P_1. For this testing problem we shall be concerned with the properties of tests based on the statistics T_N as defined in (1.7), where the score functions J_N are related to the special fixed limiting score function

$$(3.3) \qquad J_{(1)} = \log(\bar{h}_{(1)}) \ .$$

Again we suppose that $|J_{(1)}(s,t)| \le \phi(s)\psi(t)$ for all $s,t \in (0,1)$. This entails that we only consider alternatives $H_{(1)}$ which satisfy the condition $\bar{h}_1(s,t) > 0$ for all $s,t \in (0,1)$.

To formulate theorem 3.1 we shall have to introduce some more notation and to give a short review of some results obtained by Raghavachari [3] and

Woodworth [4]. For any two densities p and q of probability measures with respect to a σ-finite measure μ on a measurable space, denote the Kullback-Leibler information number by

$$(3.4) \qquad K(q,p) = \int_{\{q>0\}} q \log (q/p)d\mu .$$

The property that $0 \leq K(q,p) \leq \infty$, where $K = 0$ if and only if $\mu(\{p \neq q\}) = 0$, plays an essential role in the sequel without explicit reference. A change of variables entails $K(h,f\times g) = K(\bar{h},\bar{h}_{(0)}) = \iint \log (\bar{h})d\bar{H}$, for each $H \in H^*$. In particular for the fixed $H_{(1)} \in H_1^*$ we have

$$(3.5) \qquad \inf_{F\times G\in H_0^*} K(h_{(1)},f\times g) = K(h_{(1)},f_{(1)}\times g_{(1)}) = K(\bar{h}_{(1)},\bar{h}_{(0)}).$$

For $J_{(1)}$ as in (3.3) let us introduce

$$\pi(J_{(1)},\bar{h}) = \iint J_{(1)}d\bar{H}, \quad \underline{\pi}(J_{(1)}) = \iint J_{(1)}d\bar{H}_{(0)},$$

$$\bar{\pi}(J_{(1)}) = \sup_{H\in H^*} \iint J_{(1)}d\bar{H}.$$

It can be shown that

$$(3.6) \qquad -\infty < \underline{\pi}(J_{(1)}) < K(\bar{h}_{(1)},\bar{h}_{(0)}) < \bar{\pi}(J_{(1)}) < \infty.$$

We shall also employ the notation

$$(3.7) \qquad \rho(t) = \inf_{H\in H^*} \{K(\bar{h},\bar{h}_{(0)}):\pi(J_{(1)},\bar{h}) \geq t\} \text{ for } t \in (\underline{\pi}(J_{(1)}),\bar{\pi}(J_{(1)})).$$

In view of (3.6) this function ρ is defined at $K(\bar{h}_{(1)},\bar{h}_{(0)})$, where it assumes the value

$$(3.8) \qquad \rho(K(\bar{h}_{(1)},\bar{h}_{(0)})) = K(\bar{h}_{(1)},\bar{h}_{(0)}).$$

By way of an example let us prove (3.8). It follows from Jensen's inequality that $\pi(J_{(1)},\bar{h}) - K(\bar{h},\bar{h}_{(0)}) = \iint \log (\bar{h}_{(1)}/\bar{h})d\bar{H} \leq \log (\iint d\bar{H}_{(1)}) = 0$, so that

(3.9) $$\mathcal{r}(J_{(1)},\bar{h}) \leq K(\bar{h},\bar{h}_{(0)}),$$

for all $H \in H^*$. We have to compute the infimum in (3.7) for $t = K(\bar{h}_{(1)},\bar{h}_{(0)})$. Because $\mathcal{r}(J_{(1)},\bar{h}_{(1)}) = K(\bar{h}_{(1)},\bar{h}_{(0)})$ it follows that this infimum is not greater than $K(\bar{h}_{(1)},\bar{h}_{(0)})$. On the other hand, by (3.9) we have, for any density $\bar{h}'$ satisfying $\mathcal{r}(J_{(1)},\bar{h}') \geq K(\bar{h}_{(1)},\bar{h}_{(0)})$, the relations $K(\bar{h}',\bar{h}_{(0)}) \geq \mathcal{r}(J_{(1)},\bar{h}') \geq K(\bar{h}_{(1)},\bar{h}_{(0)})$. Combination of these results yields (3.8).

Let $S_N = S_N(X_1,Y_1,\ldots,X_N,Y_N)$ be an arbitrary extended real-valued measurable function and define

$$L_N(S_N;s) = \sup_{F\times G\in H_0^*} P_{F\times G}\ (\{S_N \geq s\}) \text{ for } s \in [-\infty,\infty].$$

According to Bahadur [1] the level attained by S_N is defined as the rv $L_N(S_N;S_N)$, for brevity denoted by $\ell(S_N)$. It follows from theorem 1 of Raghavachari [3] and (3.5) that

(3.10) $$P_1\left(\left\{\liminf_{N\to\infty} N^{-1} \log\ (\ell(S_N)) \geq -K(\bar{h}_{(1)},\bar{h}_{(0)})\right\}\right) = 1.$$

Because T_N is a rank statistic we have $L_N(T_N;t) = P_0(\{T_N \geq t\})$. Under a suitable condition (actually condition (b) of theorem 3.1 below), theorem 1 of Woodworth [4] yields that

(3.11) $$\lim_{N\to\infty} N^{-1} \log\ (P_0(\{T_N \geq t\})) = -\rho(t) \text{ for } t \in (\underline{\mathcal{r}}(J_{(1)}),\bar{\mathcal{r}}(J_{(1)}))\ ,$$

where ρ is defined in (3.7).

<u>THEOREM 3.1</u>. *Let* $J_{(1)}$ *be the function defined in (3.3) and suppose that the following conditions hold:*

(a) $|J_N(s,t)|,\ |J_{(1)}(s,t)| \leq \phi(s)\psi(t)$ *for* $s,t \in (0,1)$ *and* $N = 1,2,\ldots$ *(see (2.2)),*

(b) $\lim_{N\to\infty} \sup_{H\in H^*} |\int\int(J_N-J_{(1)})d\bar{H}| = 0,$

(c) $\sup_{N=1,2,\ldots} V_{[1/N,1]\times[1/N,1]}(J_N^{(\tau)}) \leq V(\tau) < \infty$ *for all* $\tau \geq 0$.

Then the level $\ell(T_N)$ *attained by the linear rank statistics* T_N *satisfies*

$$(3.12) \qquad P_1\left(\left\{\lim_{N\to\infty} N^{-1} \log(\ell(T_N)) \geq -K(\bar{h}_{(1)},\bar{h}_{(0)})\right\}\right) = 1,$$

so that $\{T_N\}$ *has strong exact slope* $2K(\bar{h}_{(1)},\bar{h}_{(0)})$ *for testing* H_0^* *against* $H_{(1)}$. *Moreover, this is the best strong exact slope for this testing problem.*

PROOF. Under the present conditions theorem 2.1 applies, which yields that $P_1\left(\left\{\lim_{N\to\infty} T_N = \iint J_{(1)}d\bar{H}_{(1)} = K(\bar{h}_{(1)},\bar{h}_{(0)})\right\}\right) = 1$. Relation (3.11) and the continuity of ρ imply that $(1/N)\log(\ell(T_N)) \to -\rho(K(\bar{h}_{(1)},\bar{h}_{(0)})$, as $N \to \infty$, on a set with probability 1 under P_1. The assertion (3.12) follows at once by applying (3.8).

The result in (3.10) implies that $2K(\bar{h}_{(1)},\bar{h}_{(0)})$ is the best possible strong exact slope. □

Condition (b) of the present theorem is stronger than condition (b) of theorem 2.1 (for $H \in H^*$). In the appendix it will be shown, however, that the present condition (b) is still satisfied in some interesting special cases.

4. RANK-LIKELIHOOD RATIO STATISTICS

For a fixed positive integer k let us introduce the composite alternative

$$(4.1) \qquad \{H_{(i)} \in H_1^*,\ i = 1,\ldots,k\},$$

and denote the probability measure on (Ω,A) corresponding to $H_{(i)}$ by P_i. For each $N = 1,2,\ldots$ and $i = 1,\ldots,k$ we introduce the function

$$(4.2) \qquad \lambda_{iN}(r_1,q_1,\ldots,r_N,q_N) = \log\left((N!)^2 P_i\left(\bigcap_{n=1}^{N} \{R_{nN}=r_n, Q_{nN}=q_n\}\right)\right),$$

where $(r_1,\ldots,r_N)$ and $(q_1,\ldots,q_N)$ are two arbitrary permutations of the numbers $1,\ldots,N$.

Let us consider the random variables

(4.3) $$\Lambda_{iN} = \lambda_{iN}(R_{1N},Q_{1N},\ldots,R_{NN},Q_{NN}) \text{ for } i = 1,\ldots,k,$$

(4.4) $$M_N = \max_{1\le i\le k} \Lambda_{iN},$$

defined on Ω, where λ_{iN} is given by (4.2). The Λ_{iN} are called rank-likelihood ratio statistics. It will be convenient to compare the Λ_{iN} with special linear rank statistics (see Hájek [2]), T_{iN} say, given by (1.7) with score functions satisfying

(4.5) $$J_{iN}(m/N,n/N) = \int_0^1\int_0^1 J_{(i)}(s,t)b_{m,N-m+1}(s)b_{n,N-n+1}(t)dsdt.$$

Here $J_{(i)}$ is defined according to (3.3) with $\bar{h}_{(1)}$ replaced by $\bar{h}_{(i)}$ and $b_{\mu,\nu}$ is the beta-density with parameters μ and ν.

<u>THEOREM 4.1</u>. *Suppose that* $J_{(i)}$ *and the* J_{iN} *satisfy conditions (a)-(c) of theorem 3.1 for* $i = 1,\ldots,k$. *Then the sequence* $\{M_N\}$ *has strong exact slope* $2K(\bar{h}_{(i)},\bar{h}_{(0)})$ *for testing* H_0^* *against* $H_{(i)}$ *in the alternative (4.1),* $i = 1,\ldots,k$. *These exact slopes are the best attainable.*

<u>PROOF</u>. The proof is a straightforward modification of the approach in sections 3 and 4 of Hájek [2]. To avoid a needless repetition of arguments let us just observe that the proof centers around the remark that

(4.6) $$N^{-1} \sum_{n=1}^{N} J_{iN}(r_n/N,q_n/N) \le N^{-1} \lambda_{iN}(r_1,q_1,\ldots,r_N,q_N).$$

To indicate the proof of this relation let us define the random variables

(4.7) $$\tilde{\Lambda}_{iN} = \sum_{n=1}^N J_{(i)}(F(X_n), G(Y_n)),$$

where F and G are marginal dfs of $H_{(i)}$.

By $E_0(\cdot\,|R_{1N}=r_1,Q_{1N}=q_1,\ldots,R_{NN}=r_N,Q_{NN}=q_N)$ we understand a conditional expectation given $R_{nN} = r_n$, $Q_{nN} = q_n$ for $n = 1,\ldots,N$, computed under the null hypothesis that X_n and Y_n are independent, so that $H = F \times G$ and each vector $(F(X_n),G(Y_n))$ has the uniform distribution on the unit square. We have the identity

$$(4.8) \qquad \lambda_{iN}(r_1,q_1,\ldots,r_N,q_N) = \\ = \log\left(E_0(e^{\tilde{\Lambda}_{iN}} \mid R_{1N}=r_1,Q_{1N}=q_1,\ldots,R_{NN}=r_N,Q_{NN}=q_N)\right),$$

and the proof proceeds like in the paper by Hâjek [2]. □

In the proof of theorem 4.1 we use theorem 2.1 and relation (3.10) which follows from a result of Raghavachari [3]. However, it should be noted that, still in accordance with the results of Hâjek [2], we do not need here the large deviation result (3.11) of Woodworth [4]. Conditions (a) and (b) of theorem 3.1 are considered in the appendix.

5. APPENDIX

In this appendix we shall investigate conditions (a) and (b) of theorems 2.1 and 3.1 in the two special cases where the J_N are either the exact or the approximate score functions derived from a fixed function J.

Let us start with condition (a) and introduce the notation

$$(A.1) \qquad R(u) = [u(1-u)]^{-1} \quad \text{for } u \in (0,1).$$

In the sequel we shall exclusively deal with measurable functions J on $(0,1) \times (0,1)$ that satisfy

$$(A.2) \qquad |J(s,t)| \le c[R(s)]^{\alpha}[R(t)]^{\beta} \text{ for } \alpha,\beta \in (0,1) \text{ with } \alpha + \beta < 1,$$

and for $s,t \in (0,1)$. Here c is an arbitrary positive constant. Let us observe that, for $k_1,k_2 \in (0,\infty)$, the functions

$$(A.3) \qquad \phi(s) = k_1[R(s)]^{\alpha},\ \psi(t) = k_2[R(t)]^{\beta} \text{ for } s,t \in (0,1),$$

satisfy (2.2) with $\xi = (\alpha+\beta)/\alpha$ and $\eta = (\alpha+\beta)/\beta$.

As before, $b_{\mu,\nu}$ will denote the beta-density with parameters μ and ν.

For any function J on $(0,1) \times (0,1)$ as described above let us define the exact score functions

$$J_{e,N}(s,t) = \int_0^1\int_0^1 J(u,v) b_{m,N-m+1}(u) b_{n,N-n+1}(v)\,du\,dv, \tag{A.4}$$

(derived from J for the sample size N), and the approximate score functions

$$J_{a,N}(s,t) = J(m/(N+1), n/(N+1)), \tag{A.5}$$

(derived from J for the sample size N), for $(s,t) \in ((m-1)/N, m/N] \times ((n-1)/N, n/N]$ and $m,n = 1,\ldots,N$. In this way the functions $J_{e,N}$ and $J_{a,N}$ are defined throughout $(0,1]$.

THEOREM A.1. *Suppose that* J *satisfies (A.2). Then there is a constant* $\tilde{c} \in (0,\infty)$ *such that*

$$|J_{e,N}(s,t)|,\ |J_{a,N}(s,t)|,\ |J(s,t)| \le \tilde{c}[R(s)]^{\alpha}[R(t)]^{\beta}, \tag{A.6}$$

for $s,t \in (0,1)$ *and* $N = 1,2,\ldots$ *Here* $J_{e,N}$ *and* $J_{a,N}$ *are given in (A.4) and (A.5).*

PROOF. It suffices to prove the theorem in the special case where $J(s,t) = [R(s)]^{\alpha}[R(t)]^{\beta}$. This function trivially satisfies the condition of the theorem, so that we need only consider the corresponding $J_{e,N}$ and $J_{a,N}$ which are, in this special case, also of product type. By symmetry, we may restrict attention to the first factors, $K_{e,N}$ and $K_{a,N}$ say, given by

$$K_{e,N}(s) = \int_0^1 [R(u)]^{\alpha} b_{m,N-m+1}(u)\,du, \tag{A.7}$$

$$K_{a,N}(s) = [R(m/(N+1))]^{\alpha}, \tag{A.8}$$

for $s \in ((m-1)/N, m/N]$ and $m = 1,\ldots,N$.

The properties of the function R^{α} and the fact that $K_{e,N}$ and $K_{a,N}$ are simple step functions entail that we only have to prove that (for some $c_1 \in (0,\infty)$)

$$|K_{e,N}(m/N)|,\ |K_{a,N}(m/N)| \le c_1[R(m/N)]^\alpha \text{ for } m = 1,\dots,N-1,$$

$$|K_{e,N}(1)|,\ |K_{a,N}(1)| = O(N^\alpha) \text{ as } N \to \infty.$$

It is immediate from (A.1) and (A.7) that

$$|K_{e,N}(m/N)| \le \Gamma(N+1)[\Gamma(m)\Gamma(N-m+1)]^{-1} \int_0^1 s^{m-\alpha-1}(1-s)^{N-m-\alpha}ds =$$

$$= \Gamma(m-\alpha)\Gamma(N-m-\alpha+1)[\Gamma(m)\Gamma(N-m+1)\Gamma(N-2\alpha+1)]^{-1} \le$$

$$\le c_2 m^{-\alpha}(N-m+1)^{-\alpha}(N+1)^{2\alpha} = c_2[R(m/(N+1))]^\alpha$$

for some constant $c_2 \in (0,\infty)$ and $m = 1,\dots,N$. Hence, according to (A.8) it suffices to show that, for some constant $c_3 \in (0,\infty)$,

$$R(m/(N+1)) \le c_3 R(m/N) \text{ for } m = 1,\dots,N-1,$$

$$R(N/(N+1)) = O(N^\alpha) \text{ as } N \to \infty.$$

These relations follow immediately from the properties of the function R^α. □

Next we shall consider condition (b) of theorem 3.1, which is the crucial property A of Woodworth [4], in the even stronger form where the supremum is taken over all $H \in \mathcal{H}$. It is clear that in this form the condition is also stronger than condition (b) of theorem 2.1. We shall say that the function J is piecewise continuous on $(0,1) \times (0,1)$ if there exist partitions $0 = s_0 < s_1 < \dots < s_p = 1$ and $0 = t_0 < t_1 < \dots < t_q = 1$ such that J is continuous on

$$\bigcup_{i=1}^{p} \bigcup_{j=1}^{q} (s_{i-1},s_i) \times (t_{j-1},t_j).$$

<u>THEOREM A.2</u>. *Suppose that* J *is piecewise continuous on* $(0,1) \times (0,1)$ *and satisfies (A.2). Then we have*

$$\text{(A.9)} \qquad \lim_{N\to\infty} \sup_{H\in\mathcal{H}} \left|\iint (J_{e,N}-J)d\bar{H}\right| = 0,$$

(A.10) $$\lim_{N\to\infty} \sup_{H\in\mathcal{H}} |\iint (J_{a,N}-J)d\bar{H}| = 0.$$

PROOF. Let us choose an arbitrary $\varepsilon > 0$. For sufficiently small $\gamma > 0$ let us consider the sets

$$D_{1\gamma} = \bigcup_{i=0}^{p} (s_i-\gamma, s_i+\gamma) \cap (0,1),\ D_{2\gamma} = \bigcup_{j=0}^{q} (t_j-\gamma, t_j+\gamma) \cap (0,1).$$

For ϕ and ψ as in (A.3) with $\xi = (\alpha+\beta)/\alpha$ and $\eta = (\alpha+\beta)/\beta$ we have, because each $\bar{H}$ has uniform (0,1) marginals,

$$\sup_{H\in\mathcal{H}} \iint_{D_{1\gamma}\times(0,1)} \phi(s)\psi(t)d\bar{H}(s,t) \leq$$
$$\leq \{\int_{D_{1\gamma}} [\phi(s)]^{\xi}ds\}^{1/\xi}\{\int_0^1 [\psi(t)]^{\eta}dt\}^{1/\eta} \to 0 \text{ as } \gamma \downarrow 0.$$

A similar result holds for integration over the set $(0,1) \times D_{2\gamma}$. Consequently, in view of theorem A.1, there exists a sufficiently small fixed $\gamma = \gamma(\varepsilon) > 0$ independent of $H \in \mathcal{H}$ such that the contribution to the integrals in (A.9), when integration is restricted to the set $\{D_{1\gamma}\times(0,1)\} \cup \{(0,1)\times D_{2\gamma}\}$, is bounded by ε for all $H \in \mathcal{H}$.

To prove the theorem it suffices to show that J is uniformly approximated by both $J_{e,N}$ and $J_{a,N}$ on the closed subset $\{(0,1)-D_{1\gamma}\} \times \{(0,1)-D_{2\gamma}\}$. Since J is uniformly continuous on this closed subset, this is trivially true for the $J_{a,N}$. As far as the exact scores are concerned it suffices to prove that the $J_{e,N}$ approximate J uniformly on the set

(A.11) $$S_{\gamma} = [\gamma, s_1-\gamma] \times [\gamma, t_1-\gamma],$$

which is one of the rectangles constituting the set $\{(0,1)-D_{1\gamma}\} \times \{(0,1)-D_{2\gamma}\}$.

The uniform continuity of J on the set (A.11) implies the existence of a number $0 < \zeta = \zeta(\varepsilon) < \gamma/2$ such that for any two points in an arbitrary square with sides of length ζ and centre in $[\gamma,1-\gamma] \times [\gamma,1-\gamma]$, the difference of the corresponding values of J does not differ by more than ε in absolute value. Let us, for brevity, introduce the notation

$$O_{m,N} = (m/N-\zeta/2, m/N+\zeta/2),\ O_{m,n,N} = O_{m,N} \times O_{n,N},$$

$$\max_{u \notin O_{m,N}} b_{m,N-m+1}(u) = \pi_{m,N}.$$

We shall use the property

$$\text{(A.12)} \qquad \lim_{N\to\infty} \max_{m:m/N\in[\gamma,1-\gamma]} \pi_{m,N} = 0 .$$

By uniform continuity of J it follows that $|J(u,v)-J(m/N,n/N)| \le \varepsilon$ for all $(u,v) \in O_{m,n,N}$ and all m,n such that $(m/N,n/N) \in S_\gamma$. Hence we obtain

$$\max_{m,n:(m/N,n/N)\in S_\gamma} |J_{e,N}(m/N,n/N)-J(m/N,n/N)| \le$$

$$\le \max_{m,n:(m/N,n/N)\in S_\gamma} \iint_{(u,v)\in O_{m,n,N}} |J(u,v)-J(m/N,n/N)|$$
$$\times\, b_{m,N-m+1}(u) b_{n,N-n+1}(v)dudv$$

$$+ \max_{m,n:(m/N,n/N)\in S_\gamma} \iint_{(u,v)\notin O_{m,n,N}} |J(u,v)-J(m/N,n/N)|$$

$$\times\, b_{m,N-m+1}(u) b_{n,N-n+1}(v)dudv \le 2\varepsilon ,$$

for N sufficiently large, because the second term in this bound is less than $K \times \max_{m:m/N\in[\gamma,1-\gamma]} \pi_{m,N}$ which tends to 0 as $N \to \infty$ by (A.12). □

REFERENCES

[1] BAHADUR, R.R.: Rates of convergence of estimates and test statistics. *Ann. Math. Statist. 38*, 303-324 (1967).

[2] HÁJEK, J.: Asymptotic sufficiency of the vector of ranks in the Bahadur sense. *Ann. Statist. 2*, 75-83 (1974).

[3] RAGHAVACHARI, M.: On a theorem of Bahadur on the rate of convergence of test statistics. *Ann. Math. Statist. 41*, 1695-1699 (1970).

[4] WOODWORTH, G.: Large deviations and Bahadur efficiency in linear rank statistics. *Ann. Math. Statist. 41*, 251-283 (1970).

List of Symbols

Latin						Greek	Mathematics		
Normal		Italics		Script	IBM symbol 10				
A	a	A	a	$\mathcal{A}$			α	$\mathbb{R}$:	real line
B	b	B	b				β	0 :	zero
C						Γ	γ	1 :	one
D							Δ	∞ :	infinity
E							ε	$\int$:	integral
F							η	$\sum$:	sum
G							κ	$\times$:	multiplication
H				$\mathcal{H}$		Λ	λ	+ :	summation
I							μ	− :	subtraction
J							ν	inf :	infimum
K				$\mathcal{K}$			π	sup :	supremum
L					ℓ		ρ	lim :	limit
M							σ	$\in$:	element of
N							τ	# :	number of elements in a set
O							ϕ		
P						Ψ	ψ	$\emptyset$:	void set
Q							χ	$\cup$:	union
R				$\mathcal{r}$		Ω	ω	$\cap$:	intersection
S									
T									
U									
V									
W									
X									
Y									
Z									