The Annals of Mathematical Statistics
1970, Vol. 41, No. 1, 251–283

LARGE DEVIATIONS AND BAHADUR EFFICIENCY OF LINEAR RANK STATISTICS[1]

By George G. Woodworth

Stanford University

1. Introduction. $R_1, R_2, \cdots, R_N$ are the ranks of N random variables $Z_1, \cdots, Z_N$. A *linear rank statistic* is one of the form

$$T_N = \sum_{j=1}^N a_N(R_j/N+1, j/N+1), \tag{1.1}$$

where $a_N(u, v)$ is a function on the unit square called the weight function. For example, let $X_1, \cdots, X_m$ and $Y_1, \cdots, Y_n$ be two samples and define $(Z_1, \cdots, Z_N) = (X_1, \cdots, X_m, Y_1, \cdots, Y_n)$, $N = m+n$. A two-sample "scores" statistic (Chernoff-Savage statistic) can be written in the form

$$T_N = \sum_{j=1}^N J_N(R_j/N+1)L_N(j/N+1) \tag{1.2}$$

where $L_N(u) = 0$ or 1 according as $u \leqq$ or $> m/N+1$. Some well known choices of the *score function*, J_N, in (1.2) and the names of the corresponding two-sample test statistics are

Name	Score Function
Wilcoxon-Mann-Whitney	$J_N(u) = J(u) = u - \frac{1}{2}$
Fisher-Yates (normal scores)	$J_N(u) = \mu_{j \mid N}, j-1 \leqq Nu < j, j = 1, \cdots, N$
median	$J_N(u) = J(u) = \operatorname{sgn}(u - \frac{1}{2})$,

where $\mu_{j \mid N}$ is the mean of the jth smallest of N independent standard normal random variables.

This paper is concerned with large deviations of linear rank statistics under the *null hypothesis* that $(R_1, \cdots, R_N)$ is equally likely to be any of the $N!$ permutations of $(1, \cdots, N)$. The main result (Theorem 1) extends the work of M. Stone ([8], [9]) and can in fact be derived from Hoadley's Theorem 1 [5]; it is not an extension of Hoadley's theorem since its only concern is linear rank statistics under the null hypothesis. However, for such statistics the results in this paper are more general than Hoadley's and the proof is simpler; to give two examples, the results of this paper apply to tests of independence or trend (such as Spearman's *rho*), while Hoadley considers k-sample tests only, and they apply to two-sample scores statistics with unbounded scores (the normal-scores statistic for example) while Hoadley, page 362 line 22, requires bounded scores.

2. Asymptotic properties of the probability of a large deviation. Consider the following special case. T_N is defined by (1.1) but the weight function is a step function over a rectangular grid; i.e.,

$$a_N(u, v) = a(u, v) = a_{ij}, \qquad (u, v) \in C_{ij}, \qquad 1 \leqq i \leqq l, \qquad 1 \leqq j \leqq k, \tag{2.1}$$

Received October 10, 1968; revised August 6, 1969.
[1] Research supported by the Office of Naval Research under Contract Nonr 225 (53).

where C_{ij} is the rectangle: $C_{ij} = \{(u,v) \mid u_{i-1} \leqq u < u_i, v_{j-1} \leqq v < v_j\}$, and $0 = u_0 < u_1 < \cdots < u_l = 1, 0 = v_0 < v_1 < \cdots < v_k = 1$ are constants.

Define the random matrix $\mathbf{X}^{(N)} = \{X_{ij}^{(N)}; 1 \leqq i \leqq l, 1 \leqq j \leqq k\}$ as follows: $X_{ij}^{(N)} = \#\{\alpha \mid (R_\alpha/N+1, \alpha/N+1) \in C_{ij}\}$, where "#" stands for "the number of integers in".

It follows from this definition and (2.1) that

$$T_N = \sum_i \sum_j a_{ij} X_{ij}^{(N)} \tag{2.2}$$

Let $\mathbf{x}$ denote a realization of $\mathbf{X}^{(N)}$, thus $\mathbf{x} = \{x_{ij};\ i = 1, \cdots, l, j = 1, \cdots, k\}$ is a matrix of nonnegative real numbers with fixed marginal totals[2]:

$$x_{i\cdot} = \sum_j x_{ij} = m_i, \qquad i = 1, \cdots, l, \quad \text{and} \tag{2.3}$$

$$x_{\cdot j} = \sum_i x_{ij} = n_j, \qquad j = 1, \cdots, k, \quad \text{where}$$

$m_i = \#[(N+1)u_{i-1},\ (N+1)u_i)$, $i = 1, \cdots, l$, and $n_j = \#[(N+1)v_{j-1},\ (N+1)v_j)$, $j = 1, \cdots, k$.

The distribution of $\mathbf{X}^{(N)}$ is multi-hypergeometric, to wit:

$$P[\mathbf{X}^{(N)} = \mathbf{x}] = \prod_i m_i! \prod_j n_j! / N! \prod_{ij} x_{ij}!,$$

provided of course $\mathbf{x}$ satisfies (2.3).

The key to the main result of this paper is the simple fact that $\mathbf{X}^{(N)}$ has a conditioned multinomial distribution; in fact, suppose that $\mathbf{Y}^{(N)} = \{Y_{ij}^{(N)} \mid 1 \leqq i \leqq l, 1 \leqq j \leqq k\}$ has a multinomial distribution with sample size N and cell probabilities

$$p_{ij} = (u_i - u_{i-1})(v_j - v_{j-1}) = \mu_i \nu_j, \quad \text{say.} \tag{2.4}$$

Then, for any $\mathbf{x}$ satisfying (2.3)

$$\begin{aligned} P[\mathbf{X}^{(N)} = \mathbf{x}] &= P[\mathbf{Y}^{(N)} = \mathbf{x} \mid Y_{i\cdot}^{(N)} = m_i, Y_{\cdot j}^{(N)} = n_j, \quad \text{for all} \quad i,j] \\ &= P[\mathbf{Y}^{(N)} = \mathbf{x} \text{ and } Y_{i\cdot}^{(N)} = m_i, Y_{\cdot j}^{(N)} = n_j, \quad \text{for all} \quad i,j] \\ &\qquad \cdot [\textstyle\prod_i (m_i!\, \mu_i^{-m_i})/N!][\prod_j (n_j!\, \nu_j^{-n_j})/N!] \\ &= P[\mathbf{Y}^{(N)} = \mathbf{x} \quad \text{and} \quad Y_{i\cdot}^{(N)} = m_i, Y_{\cdot j}^{(N)} = n_j, \quad \text{for all} \quad i,j] \exp[N\varepsilon_N] \end{aligned} \tag{2.5}$$

where, after application of Stirlings approximation,

$$\begin{aligned} \varepsilon_N = \textstyle\sum_j (n_j/N) \log(n_j/N\nu_j) + O(\log(\min_j(n_j))/N) \\ + \textstyle\sum_i (m_i/N) \log(m_i/N\mu_i) + O(\log(\min_i(m_i))/N). \end{aligned}$$

Clearly

$$m_i/N \to \mu_i \quad \text{and} \quad n_j/N \to \nu_j \quad \text{as} \quad N \to \infty, \tag{2.6}$$

thus $\varepsilon_N \to 0$ as $N \to \infty$.

[2] The dot notation used throughout this paper indicates summation (not averaging) over the dotted subscript.

It follows from (2.2) and (2.5) that, for any sequence of constants $\{r_N\}$,

$$-N^{-1}\log P[T_N \geqq Nr_N] = -N^{-1}\log P[\sum_i\sum_j a_{ij}Y_{ij}^{(N)} \geqq Nr_N \quad\text{and}\quad Y_{i\cdot}^{(N)} = m_i, \quad Y_{\cdot j}^{(N)} = n_j, \quad \text{for all} \quad i,j] - \varepsilon_N.$$

It follows from this and Hoeffding[3] [6, Theorem 2.1] that

$$-N^{-1}\log P[T_N \geqq Nr_N] = \min_x\{\sum_i\sum_j (x_{ij}/N)\log(x_{ij}/Np_{ij}) \,|\, \sum_i\sum_j a_{ij}x_{ij} \geqq Nr_N, x_{i\cdot} = m_i, x_{\cdot j} = n_j, \quad x_{ij} \text{ are integers } \geqq 0\} + o(1) \tag{2.7}$$

where $p_{ij} = \mu_i\nu_j$, defined in (2.4).

That part of (2.7) (and similar expressions below) lying between "|" and "}" will be called the *constraint*.

LEMMA 1. *Let* $\mathbf{q} = \{q_{ij} \,|\, 1 \leqq i \leqq l, 1 \leqq j \leqq k\}$ *be nonnegative real numbers and define*

$$\underline{r}(a) = \sum_i\sum_j a_{ij}p_{ij}, \tag{2.8}$$

$$\bar{r}(a) = \sup_{\mathbf{q}}\{\sum_i\sum_j a_{ij}q_{ij} \,|\, q_{i\cdot} = \mu_i, q_{\cdot j} = \nu_j, \quad \textit{for all} \quad i,j\}, \tag{2.9}$$

with p_{ij}, μ_i *and* ν_j *as in* (2.4), *and define*

$$I(r;a) = \inf_{\mathbf{q}}\{\sum_i\sum_j q_{ij}\log(q_{ij}/p_{ij}) \,|\, \sum_i\sum_j a_{ij}q_{ij} \geqq r, \quad q_{i\cdot} = \mu_i, q_{\cdot j} = \nu_j, \quad \textit{for all} \quad i,j\}. \tag{2.10}$$

If T_N *is defined by* (1.1) *and satisfies* (2.1) *and* $\{r_N\}$ *is a sequence of constants approaching a constant* $r < \bar{r}(a)$ *then* $\lim_{N\to\infty}\{-N^{-1}\log P[T_N \geqq Nr_N]\} = I(r;a)$. *If* $r \leqq \underline{r}(a)$, *the above limit is zero.*

PROOF. The case $r \leqq \underline{r}(a)$ is trivial since T_N/N converges in probability to $\underline{r}(a)$.

Let $\mathbf{x}^{(N)}$ be a value of $\mathbf{x}$ for which the minimum[4] is attained on the right side of (2.7). Since $0 \leqq x_{ij}^{(N)}/N \leqq 1$ there exists a subsequence of N's (which for simplicity of notation will be taken to be the original sequence) such that for each i and j, $x_{ij}^{(N)}/N$ converges to some number q_{ij}^0. Clearly $\mathbf{q}^0 = \{q_{ij}^0; i = 1,\cdots,l, j = 1,\cdots,k\}$ will satisfy the constraint in (2.10). Thus from (2.7) and (2.10)

$$\lim_{N\to\infty}\{-N^{-1}\log P[T_N \geqq Nr_N]\} = \lim_{N\to\infty}\sum_i\sum_j (x_{ij}^{(N)}/N)\log(x_{ij}^{(N)}/Np_{ij}) = \sum_i\sum_j q_{ij}^0\log(q_{ij}^0/p_{ij}) \geqq I(r;a). \tag{2.11}$$

Suppose that the above inequality is strict. Then it is possible to select $\delta > 0$ small enough that

$$\lim_{N\to\infty}\{-N^{-1}\log P[T_N \geqq Nr_N]\} \geqq I(r;a) + 3\delta. \tag{2.12}$$

It follows from the definition (2.10) that $I(r;a)$ is a nonnegative, nondecreasing convex function of r, hence it is continuous where it is finite. Since the inequality in

[3] Warning: Hoeffding and I use the notation $I(:;\cdot)$ in different ways.

[4] $\mathbf{x}^{(N)}$ is unique but I neither need nor prove the fact here.

(2.11) was assumed to be strict $I(r;a)$ must be finite and thus it is possible to select $\varepsilon > 0$ small enough that

$$r+2\varepsilon < \bar{r}(a) \quad \text{and} \tag{2.13}$$

$$I(r+2\varepsilon;a) \leqq I(r;a)+\delta. \tag{2.14}$$

It follows from (2.9), the definition (2.10) of $I(r+2\varepsilon;a)$, (2.13) and (2.14) that there exists $\mathbf{q}^1$ such that

$$\textstyle\sum_i \sum_j q_{ij}^1 a_{ij} \geqq r+2\varepsilon, \qquad q_{i\cdot}^1 = \mu_i, \qquad q_{\cdot j}^1 = \nu_j, \quad \text{for all} \quad i,j, \quad \text{and} \tag{2.15}$$

$$\textstyle\sum_i \sum_j q_{ij}^1 \log(q_{ij}^1/p_{ij}) \leqq I(r+2\varepsilon;a)+\delta \leqq I(r;a)+2\delta. \tag{2.16}$$

Define

$$z_{ij}^{(N)} = [q_{ij}^1 m_i' n_j'/Np_{ij}], \qquad 1 \leqq i \leqq l, \quad 1 \leqq j \leqq k, \tag{2.17}$$

where $[\cdot]$ denotes the greatest integer function and

$$m_i' = \min(m_i, [N\mu_i]), \qquad n_j' = \min(n_j, [N\nu_j]). \tag{2.18}$$

Clearly from (2.15), (2.17) and (2.18) one has

$$z_{i\cdot}^{(N)} \leqq m_i \quad \text{and} \quad z_{\cdot j}^{(N)} \leqq n_j \quad \text{for all} \quad i,j, \tag{2.19}$$

and from this and (2.17) it follows that

$$\lim_{N\to\infty} z_{i\cdot}^{(N)}/N = \mu_i, \lim_{N\to\infty} z_{\cdot j}^{(N)}/N = \nu_j, \quad \text{for all} \quad i,j. \tag{2.20}$$

Define $c_i^{(N)} = m_i - z_{i\cdot}^{(N)}, d_j^{(N)} = n_j - z_{\cdot j}^{(N)}$ and for each N select[5] nonnegative integers $\{y_{ij}^{(N)}; 1 \leqq i \leqq l, 1 \leqq j \leqq k\}$ such that $y_{i\cdot}^{(N)} = c_i^{(N)}$ and $y_{\cdot j}^{(N)} = d_j^{(N)}$. Now set $x_{ij}^{(N)} = z_{ij}^{(N)} + y_{ij}^{(N)}$, thus $x_{i\cdot}^{(N)} = m_i$ and $x_{\cdot j}^{(N)} = n_j$ and, by (2.4), (2.6), (2.17), (2.18) and (2.20),

$$\lim_{N\to\infty} x_{ij}^{(N)}/N = q_{ij}^1. \tag{2.21}$$

It follows from (2.15) and (2.21) that, for large enough $N, \sum_i \sum_j x_{ij}^{(N)} a_{ij}/N \geqq r+\varepsilon$. Thus $\mathbf{x}^{(N)} = \{x_{ij}^{(N)}\}$ satisfies the constraint in (2.7) with r_N set equal to $r+\varepsilon$ and from this fact and (2.7), (2.16) and (2.21) it follows that

$$\begin{aligned} \lim_{N\to\infty} \{-N^{-1}\log P[T_N \geqq Nr_N]\} &\leqq \lim_{N\to\infty} \{-N^{-1}\log P[T_N \geqq N(r+\varepsilon)]\} \\ &\leqq \lim_{N\to\infty} \textstyle\sum_i \sum_j (x_{ij}^{(N)}/N)\log(x_{ij}^{(N)}/Np_{ij}) \\ &= \textstyle\sum_i \sum_j q_{ij}^1 \log(q_{ij}^1/p_{ij}) \leqq I(r;a)+2\delta. \end{aligned}$$

Since the latter contradicts (2.12), the inequality in (2.11) cannot be strict and the lemma is proved.

Requiring, as in (2.1), that $a_N(\cdot,\cdot)$ be a step function makes Lemma 1 too

[5] Such a selection is always possible: for example, assume that $c_l \leqq d_k$ (otherwise reverse the roles of the c's and d's) and set $y_{lk} = c_l$ and $y_{lj} = 0$ for $j = 1, \cdots, k-1$. Now define $k' = k$, $l' = l-1$, $c_i' = c_i, i = 1, \cdots, l', d_j' = d_j, j = 1, \cdots, k'-1$ $d_{k'}' = d_k - c_l$ and repeat the previous sentence.

restricted to cover many interesting statistics so it is desirable to widen the class of weight functions to which this lemma applies. To achieve this end let us define a pseudometric d as follows:[6]

$$d(a,b) = \sup_{h\in\mathscr{H}} \left| \int\int (a(u,v)-b(u,v))h(u,v)\,du\,dv \right|, \tag{2.22}$$

where a and b are real functions over the unit square and

$$\mathscr{H} = \{h(\cdot,\cdot) \mid h \geqq 0, \quad \int h(u,v)\,du = 1 = \int h(u,v)\,dv\} \tag{2.23}$$

is the set of all bivariate densities with uniform marginals.

It will be assumed hereafter that the sequence $\{a_N(\cdot,\cdot)\}$ satisfies

Property A.

(i) For each N, a_N is constant over the rectangles $\{i-l \leqq Nu < i, j-1 \leqq Nv < j\}$, $1 \leqq i,j \leqq N$.

(ii) There exists a function $a(\cdot,\cdot)$ over the unit square such that

$$d(a_N, a) = \sup_{h\in\mathscr{H}} \left|\int\int (a_N - a)h\right| \to 0, \quad \text{as} \quad N \to \infty. \tag{2.24}$$

Despite its formidable appearance, Property A is satisfied by all the standard linear rank statistics, in particular by any statistic satisfying the Chernoff-Savage [4] conditions ((1)–(3) page 974) or Bhuchongkul's [3] conditions ((1)–(5) page 139) or more generally, the sufficient condition proved below (Theorem 2).

DEFINITION 2.1. A sequence of statistics $\{T_N\}$, each of form (1.1) with weight functions $a_N(\cdot,\cdot)$ satisfying property A will be called a *type A sequence of linear rank statistics.*

For a function $a(\cdot,\cdot)$ over the unit square satisfying $d(a,0) < \infty$, let

$$\underline{r}(a) = \int\int a, \tag{2.25}$$

$$\bar{r}(a) = \sup\{\int\int ah \mid h \in \mathscr{H}\} \tag{2.26}$$

and, for $r < \bar{r}(a)$, define

$$I(r;a) = \inf\{\int\int h\log(h) \mid \int\int ah \geqq r, h \in \mathscr{H}\}. \tag{2.27}$$

With a little effort one can see that the above reduce to (2.8), (2.9) and (2.10) when $a(\cdot,\cdot)$ satisfies (2.1), and that $I(r;a)$ is nonnegative, nondecreasing and convex in r (hence continuous).

Moreover if a_ε is a function such that

$$d(a,a_\varepsilon) \leqq \varepsilon, \tag{2.28}$$

then

$$I(r-\varepsilon;a) \leqq I(r;a_\varepsilon) \leqq I(r+\varepsilon;a). \tag{2.29}$$

Another property of $I(r;a)$ which will prove useful is the following: for

[6] When the range of integration is unspecified it is understood to be (0, 1).

any integrable functions $a_1(u)$ and $a_2(v)$ $0 < u < 1$, $0 < v < 1$, and any positive constant c

$$(2.30) \qquad I(r; a) = I(cr + \bar{a}_1 + \bar{a}_2; \qquad ca + a_1 + a_2),$$

where $\bar{a}_1 = \int a_1$ and $\bar{a}_2 = \int a_2$.

THEOREM 1. *If* $\{T_N\}$ *is a type* A *sequence of linear rank statistics* (*Definition* 2.1) *and* $\{r_N\}$ *is a sequence of constants approaching a constant r then*

$$\begin{aligned} \lim_{N\to\infty}\{-N^{-1}\log P[T_N \geqq Nr_N]\} &= I(r; a) \qquad r < \bar{r}(a), \\ &= 0 \qquad r \leqq \underline{r}(a); \end{aligned}$$

where $\underline{r}(a)$, $\bar{r}(a)$ *and* $I(r; a)$ *are defined by* (2.25), (2.26) *and* (2.27).

PROOF. Define the bivariate density $\tilde{h}_N$ as follows

$$\begin{aligned} (2.31) \qquad \tilde{h}_N(u, v) &= N \qquad (u, v) \in D_j, \quad j = 1, \cdots, N, \\ &= 0 \qquad \text{elsewhere}; \end{aligned}$$

where $D_j = \{(u, v) \mid (R_j - 1)/N \leqq u < R_j/N, j - 1/N \leqq v < j/N\}$ and $(R_1, \cdots, R_N)$ are defined in the first line of this paper. Clearly $\tilde{h}_N$ has uniform marginals and, by (1.1) and part (i) of Property A,

$$(2.32) \qquad T_N = N \int\int a_N \tilde{h}_N.$$

It follows from (2.24) that for each $\varepsilon > 0$ there is a function $a_\varepsilon(\cdot, \cdot)$ over the unit square such that a_ε satisfies (2.1) for some $k = k(\varepsilon)$ and $l = l(\varepsilon)$ and such that $d(a_\varepsilon, a_N) < \varepsilon$ for large enough N and $d(a_\varepsilon, a) < \varepsilon$. Let $T_{N\varepsilon} = \sum_{j=1}^N a_\varepsilon(R_j/N+1, j/N+1)$. Since $a_\varepsilon(R_j/N+1, j/N+1) \neq N^2 \int_{Dj}\int a_\varepsilon$ only if there is some α or some β such that $R_j - 1 \leqq Nu_\alpha < R_j$ or $j - 1 \leqq Nv_\beta < j$, where u_α and v_β are defined after (2.1), and since this can happen at most once for each $\alpha \leqq k(\varepsilon)$ or $\beta \leqq l(\varepsilon)$ it follows that

$$(2.33) \qquad |T_{N\varepsilon}/N - \int\int a_\varepsilon \tilde{h}_N| \leqq (\max a_\varepsilon - \min a_\varepsilon)(k(\varepsilon) + l(\varepsilon))/N = \delta_N.$$

The latter is non-random and approaches zero as $N \to \infty$, because $a_\varepsilon(\cdot, \cdot)$ takes on at most $k(\varepsilon) \cdot l(\varepsilon)$ distinct values.

Since $\tilde{h}_N \in \mathscr{H}$ it follows from (2.23), (2.32) and (2.33) that

$$|T_N - T_{N\varepsilon}|/N \leqq \int\int |a_\varepsilon - a_N| \tilde{h}_N + |T_{N\varepsilon}/N - \int\int a_\varepsilon \tilde{h}_N| \leqq \varepsilon + \delta_N$$

for sufficiently large N. Consequently, with $r_N' = r_N - \delta_N$, $r_N'' = r_N + \delta_N$, for large N,

$$P[T_{N\varepsilon} \geqq N(r_N'' + \varepsilon)] \leqq P[T_N \geqq Nr_N] \leqq P[T_{N\varepsilon} \geqq N(r_N' - \varepsilon)].$$

It follows from the above, (2.29) and Lemma 1 applied to $T_{N\varepsilon}$ that

$$\begin{aligned} (2.34) \qquad I(r - 2\varepsilon, a) &\leqq I(r - \varepsilon; a_\varepsilon) \leqq \lim_{N\to\infty}\{-N^{-1}\log P[T_N \geqq Nr_N]\} \\ &\leqq I(r + \varepsilon; a_\varepsilon) \leqq I(r + 2\varepsilon, a). \end{aligned}$$

Since ε is arbitrary and $I(r; a)$ continuous, the theorem is proved.

It was remarked earlier that Property A is usually satisfied; the following theorem makes this statement more specific.

THEOREM 2. *If* $a_N(u, v)$ *is of the form*

$$a_N(u,v) = \sum_{l=1}^{p} J_{Nl}(u)L_{Nl}(v), \tag{2.35}$$

where p *is fixed and finite, and the* $2p$ *functions,* J_{Nl}, L_{Nl}, $l = 1, \cdots, p$, *are constant over intervals like* $[(i-1)/N, i/N)$, $i = 1, \cdots, N$ *and converge in quadratic mean* (qm) *to square integrable functions* J_l, L_l, $l = 1, \cdots, p$, *then Property* A *is satisfied.*

PROOF. Clearly part (i) of Property A holds.
Define

$$a(u,v) = \sum_{l=1}^{p} J_l(u)L_l(v). \tag{2.36}$$

For convenience in establishing part (ii) consider $p = 1$ and drop the l-subscripts; the generalization to any fixed finite p is routine. Since J and L are square integrable they can be approximated arbitrarily well in qm by functions constant on the intervals $[i-1/k, i/k), i = 1, \cdots, k$ for k sufficiently large. The result then follows in an obvious way from the fact that for any square integrable functions J^*, L^* and any density h with uniform marginals one has

$$\int\int |J^*(u)L^*(v) - J(u)L(v)|\, h(u,v)\, du\, dv$$
$$\leqq \{\int [J^*(u)-J(u)]^2\, du \int [L^*(v)]^2\, dv\}^{\frac{1}{2}} + \{\int [L^*(v)-L(v)]^2 \int [J(u)]^2\}^{\frac{1}{2}}.$$

COROLLARY 1. *If* $\{a_N\}$ *has property* A, *then for every* $\varepsilon > 0$ *there exists a bounded, continuous function* a_ε^* *such that* $\lim_{N\to\infty} d(a_N, a_\varepsilon^*) < \varepsilon$ *and* $d(a, a_\varepsilon^*) < \varepsilon$.

PROOF. Let N_0 be large enough that $d(a_N, a) < \varepsilon$ for $N \geqq N_0$. Let $a_\varepsilon = a_{N_0}$. Clearly $a_\varepsilon(u,v) = \sum_{i=1}^{p} J_i(u) \cdot L_i(v)$ where $p = N_0, J_i$ and L_i are elementary functions. Now proceed as in the proof of Theorem 2 except approximate J_i and L_i by bounded, continuous functions.

3. Evaluating $I(r; a)$. Suppose $\lambda > 0$ is an arbitrary constant and $s(v), 0 \leqq v \leqq 1$ is an arbitrary function such that for almost every v, $0 \leqq v \leqq 1$,

$$t(u) = \log \{\int \exp [\lambda(a(u,v) - s(v))]\, dv\}/\lambda < \infty.$$

If one defines $g(u,v) = \exp [\lambda(a(u,v) - s(v) - t(u))]$, then for each u, $g(u,v)$ is a density on $0 < v < 1$. If $f(u,v)$ is a density with uniform marginals, then

$$\int \log (g(u,v)/f(u,v))f(u,v)\, dv \leqq \log \{\int_{[f>0]} g(u,v)\, dv\} \leqq 0.$$

Consequently $\int\int f \log f \geqq \lambda[\int\int a(u,v)f(u,v) - \int s(v) - \int t(u)]$, and if $\int\int af \geqq r$ then

$$\int\int f \log f \geqq \lambda[r - \int s(v) - \int t(u)] \tag{3.1}$$
$$= \lambda[r - \int s(v)] - \int \log \{\int \exp [\lambda(a(u,v) - s(v))]\, dv\}\, du.$$

If λ and $s(v)$ can be chosen so that $g(u,v)$ is a density with uniform marginals and $\int\int ag = r$ then equality can be attained in (3.1) by setting $f = g$. The above remarks imply

THEOREM 3. *If there exists a constant* $\lambda > 0$ *and a function* $s(u)$ *such that*

$$(3.2) \qquad 1 = \int \frac{\exp[\lambda(a(u,v)-s(v))]}{\int \exp[\lambda(a(u,v')-s(v'))]\,dv'}\,du \qquad 0<v<1 \quad \textit{and}$$

$$(3.3) \qquad r = \int \frac{\int a(u,v)\exp[\lambda(a(u,v)-s(v))]\,dv}{\int \exp[\lambda(a(u,v)-s(v))]\,dv}\,du, \qquad \textit{then}$$

$$(3.4) \qquad I(r;a) = \lambda(r-\textstyle\int s)-\int \log\{\int \exp[\lambda(a(u,v)-s(v))]\,dv\}\,du.$$

REMARK. The roles of u and v can be reversed in the above discussion.

EXAMPLE 1. The Fisher-Yates (normal-scores) correlation coefficient is a statistic of form (1.1) with $a_N(u,v) = J_N(u)\cdot J_N(v)$, where $J_N(u) = EZ_{j\mid N}$, $(j-1)/N \leqq u < j/N$, and $Z_{j\mid N}$ is the jth smallest of N independent standard normal random variables. It is well known that $J_N(u)$ converges in quadratic mean to $\Phi^{-1}(u)$, the inverse of the standard normal cdf. Thus the sufficient condition (Theorem 2) for Property A is satisfied with $a(u,v) = \Phi^{-1}(u)\Phi^{-1}(v) = a_{FYC}(u,v)$, say.

The function $s(v) = b[\Phi^{-1}(v)]^2$, where $b = [-1+(1+4\lambda^2)^{\frac{1}{2}}]/4\lambda$ satisfies (3.2) with $a = a_{FYC}$. By solving for $\lambda \geqq 0$ in (3.3), which can be done for $0 < r < 1$, and substituting the result in (3.4) one obtains

$$(3.5) \qquad I(r;a_{FYC}) = -\tfrac{1}{2}\log(1-r^2), \qquad 0<r<1.$$

k-Sample scores statistics. In the k-sample problem let $n_1,\cdots,n_k$ denote the sample sizes, $n_1+\cdots+n_k = N$, and suppose that $n_j/N \to \rho_j(\neq 0,1)$ as $N\to\infty$. Let S_{ij} be the rank in the combined sample of the jth (unordered) observation from the ith sample. Under the null hypothesis that the k samples were drawn from identical continuous populations, the ranks $(R_1,\cdots,R_N) = (S_{11},\cdots,S_{1n_1};\ S_{21},\cdots,S_{2n_2};\cdots;\ S_{k1},\cdots,S_{kn_k})$ are equally likely to be any permutation of $(1,\cdots,N)$.

A k-sample scores statistic[7] is one of the form:

$$(3.6) \qquad \textstyle\sum_{i=1}^{k}\sum_{j=1}^{n_i} J_{Ni}(S_{ij}/N+1),$$

where $J_{Ni}(u)$, $i = 1,\cdots,k$ are functions which are constant over the intervals $i-1 \leqq Nu < i$, $i = 1,\cdots,N$. Let $v_{Ni} = (n_1+\cdots+n_i)/N$, $i = 1,\cdots,k$, $v_{N0} = 0$, and define $L_{Ni}(v)$ to be 1 for $v_{N,i-1} \leqq v < v_{Ni}$ and 0 otherwise. Then the above becomes $\sum_{j=1}^{N} a_N(R_j/N+1, j/N+1)$, where $a_N(u,v) = \sum_{i=1}^{k} J_{Ni}(u)\cdot L_{Ni}(v)$.

Since $v_{Nj}\to v_j = \rho_1+\cdots+\rho_j$ it follows that if $J_{Ni}(u)\to J_i(u)$ in quadratic mean, then, by Theorem 2, a_N has Property A with

$$(3.7) \qquad \begin{aligned} a(u,v) &= J_j(u), \qquad v_{j-1}\leqq v < v_j, \qquad j = 1,\cdots,k,\\ &= a(u,v;\mathbf{J},\rho), \quad \text{say} \end{aligned}$$

where $\mathbf{J} = (J_1,\cdots,J_k)$ and $\rho = (\rho_1\ \cdots,\rho_k)$.

Writing $I(r; a(\cdot,\cdot;\mathbf{J},\rho))$ more simply as $I(r;\mathbf{J},\rho)$ one has:

---

[7] Such statistics were considered by Andrews and Truax [1] for $k \geqq 3$ and, of course, for $k = 2$ they have been extensively studied.

THEOREM 4. *If* $a(u,v)$ *is of form* (3.7), $J_j(u)$, $j = 1,\cdots,k$ *are integrable, and* r *satisfies*:

$$\sum_i \rho_i \int J_i = \underline{r} \leqq r < \bar{r} = \sup\{\sum_i \rho_i \int J_i f_i \mid \sum_i \rho_i f_i = 1\}, \tag{3.8}$$

($f_1,\cdots,f_k$ *being densities over* $(0, 1)$) *then there exist constants* $\lambda, s_1,\cdots,s_k$ *such that*

$$I(r;\mathbf{J},\rho) = \lambda r - \lambda \sum \rho_j s_j - \int \log\{\sum \rho_j [\exp(\lambda J_j(u) - s_j)]\}\, du, \tag{3.9}$$

where $(\lambda, s_1,\cdots,s_k)$ *is the essentially unique solution with* $\lambda \geqq 0$ *of*

$$1 = \int \exp[\lambda(J_j(u) - s_j)]/\sum \rho_i \exp[\lambda(J_i(u) - s_i)]\, du, \quad j = 1,\cdots,k, \quad \textit{and} \tag{3.10}$$

$$r = \int \sum \rho_j J_j(u) \exp[\lambda(J_j(u) - s_j)]/\sum \rho_j \exp[\lambda(J_j(u) - s_j)]\, du. \tag{3.11}$$

"Essentially unique" means that if $(\lambda', s_1',\cdots,s_k')$ is another solution with $\lambda' \geqq 0$, then $\lambda = \lambda'$ and $s_i' = s_i + c$ for $i = 1,\cdots,k$ and some constant c. If $r < \underline{r}$, then $I(r;\mathbf{J},\rho) = 0$.

PROOF. If a solution $(\lambda, s_1,\cdots,s_k)$ does exist, and one defines $s(v) = s_j$, $v_{j-1} \leqq v < v_j$, then clearly λ and $s(v)$ satisfy (3.2) and (3.3) so that (3.9), which is simply (3.4) specialized to the present example, holds.

By Corollary 1 of the appendix, for any $\lambda \geqq 0$ there is a solution $s_1(\lambda),\cdots,s_k(\lambda)$,to (3.10) having the uniqueness property described above. With s_j replaced by $s_j(\lambda), j = 1,\cdots,k$, the right side of (3.11), call it $m(\lambda)$, becomes:

$$m(\lambda) = \int \sum \rho_j J_j(u) \exp[\lambda(J_j(u) - s_j(\lambda))]/\sum \rho_j \exp[\lambda(J_j(u) - s_j(\lambda))]\, du. \tag{3.12}$$

To complete the proof it suffices to show that the equation $m(\lambda) = r$ has a root for every $r \in [\underline{r}, \bar{r})$. Since $m(0) = \underline{r}$, it is enough to show that $m(\lambda)$ is strictly increasing, continuous and $m(\lambda) \to \bar{r}$ as $\lambda \to \infty$. This is proved in Lemmas 4, 5 and 6 of the appendix.

Application to two-sample scores statistics. These statistics, which were defined by (1.2) are also of form (3.6) with $n_1 = m, n_2 = n$, $J_{N1}(u) = 0$ and $J_{N2}(u) = J_N(u) = J_{iN}, i - 1/N \leqq u < i/N$, where $J_{N1},\cdots,J_{NN}$ are the scores on which the test is based. If $J_N(u)$ converges in quadratic mean to a function $J(u)$, then, with $\rho = \rho_2 = \lim(n/N)$, $\bar{\rho} = \rho_1 = 1 - \rho$, $s = s_2 - s_1$ and $I(r;J,\rho) = I(r;(0,J),(\rho,\bar{\rho}))$, it follows from Theorem 4 that

$$I(r;J,\rho) = \lambda(r - \rho s) - \int \log\{\bar{\rho} + \rho \exp[\lambda(J(u) - s)]\}\, du \tag{3.13}$$

where (λ, s) is the unique solution of

$$1 = \int \exp[\lambda(J(u) - s)]/(\bar{\rho} + \rho \exp[\lambda(J(u) - s)]), \quad \text{and} \tag{3.14}$$

$$r = \int \rho J(u) \exp[\lambda(J(u) - s)]/(\bar{\rho} + \rho \exp[\lambda(J(u) - s)]), \tag{3.15}$$

for any r such that

$$\rho \int J \leqq r < \sup_f \{\rho \int J f \mid 0 \leqq f \leqq \rho^{-1}, \int f = 1\}. \tag{3.16}$$

This result was first reported by M. Stone [9] who required slightly stronger conditions; in addition to qm convergence of J_N to J he required that $J_N{}^+$ converge to J^+ in $2+\delta$th moment for some $\delta > 0$.

If $J(u)$ is nondecreasing then (3.16) becomes

$$\rho \int J(u)\,du \leqq r < \int_{\bar{\rho}}^{1} J(u)\,du. \tag{3.17}$$

For the special case $\rho = \frac{1}{2}$ and $J(u) = -J(1-u)$, for any λ the solution of (3.14) is $s = 0$. Thus

$$I(r; J, \tfrac{1}{2}) = \lambda r - \int \log \cosh(\tfrac{1}{2}\lambda J(v))\,dv, \tag{3.18}$$

where λ is the solution of

$$2r = \int J(v) \tanh(\tfrac{1}{2}\lambda J(v))\,dv. \tag{3.19}$$

EXAMPLE 2. The two-sample median test is based on the number of observations from the second sample greater than the median of the combined sample or equivalently upon the difference between the numbers of observations above and below the median. Thus the median test is based on the two-sample scores statistic with $J_N(u) = J(u) = \operatorname{sgn}(u-\frac{1}{2})$. Starting from (3.13)–(3.17) a routine calculation yields $I(r;\ \text{Median},\ \rho) = K(\rho) - \frac{1}{2}[K(\bar{\rho}+r)+K(\bar{\rho}-r)]$, $0 \leqq r < \min(\rho, \bar{\rho})$, where $K(x) = -(x\log(x)+(1-x)\log(1-x))$. Values of $I(r;\ \text{Median}, \rho)$ for $\rho = \frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{16}$ and various r-values are found in Table 2b.

EXAMPLE 3. The Wilcoxon test is based on the two sample scores statistic with $J_N(u) = J(u) = u - \frac{1}{2}$. From (3.13)–(3.17) one obtains after a little manipulation

$$\begin{aligned} I(r; \text{Wilcoxon}) = {} & 2\lambda r + \rho \log(\exp(\rho\lambda)-1) \\ & + \bar{\rho}\log(\exp(\bar{\rho}\lambda)-1) - \log(\exp(\lambda)-1) + K(\rho) + \lambda\rho\bar{\rho} \end{aligned}$$

where λ is the unique solution of

$$r = \int_0^1 \frac{u \exp[\lambda(u-\rho)]}{\exp[\lambda(u-\rho)] + (1-\exp(\bar{\rho}\lambda))(1-\exp(\rho\lambda))^{-1}}\,du - \tfrac{1}{2}\rho, \qquad 0 \leqq r < \tfrac{1}{2}\rho\bar{\rho}.$$

Both Hoadley and Stone ([5], [8]) report this result[8]; the correspondence between the present notation and Hoadley's is as follows: $r = \rho\bar{\rho}\varepsilon$, $\rho_1 = \bar{\rho}$, $\rho_2 = \rho$, and his $\lambda/\rho_1\rho_2$ corresponds to λ. Values of $I(r; \text{Wilcoxon}, \rho)$ for $\rho = \frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{16}$ and selected r-values are found in Table 2c.

EXAMPLE 4. The Fisher-Yates (normal scores, *ns*) test is based on the two-sample scores statistic with $J_N(u) \to J(u) = \Phi^{-1}(u)$, the inverse of the standard normal distribution function. Again from (3.13) through (3.17) one obtains:

$$I(r; ns, \rho) = \lambda r - \lambda\rho s - \int_{-\infty}^{\infty} \log[\bar{\rho} + \rho\exp(\lambda(x-s))]\varphi(x)\,dx,$$

[8] Gerald L. Sievers also reported this result in his 1967 thesis.

where $\varphi(x) = (d/dx)\Phi(x)$ and (λ, s) is the unique solution (with $\lambda \geqq 0$) of

$$1 = \int_{-\infty}^{\infty} \frac{\exp[\lambda(x-s)]}{\bar{\rho}+\rho\exp[\lambda(x-s)]}\varphi(x)\,dx \qquad \text{and}$$

$$r = \int_{-\infty}^{\infty} \frac{\rho x\exp[\lambda(x-s)]}{\bar{\rho}+\rho\exp[\lambda(x-s)]}\varphi(x)\,dx, \qquad 0 \leqq r < \varphi(\Phi^{-1}(\bar{\rho})).$$

This was first obtained by Stone [9].

Values of $I(r; ns, \rho)$ for $\rho = \frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{16}$ and various r-values are found in Table 2a.

Type A *linear rank statistics.* If T_N is type A (Definition 2.1), then for any $\varepsilon > 0$ there exists a function a_ε satisfying (3.7) for some $k = k(\varepsilon)$ such that $d(a, a_\varepsilon) \leqq \varepsilon$. From this and (2.29) it follows that $I(r; a)$ can be approximated arbitrarily closely by $I(r; a_\varepsilon)$. But one can use (3.9)–(3.11) to evaluate $I(r; a_\varepsilon)$, so $I(r; a)$ can be calculated as accurately as one wishes.

In fact, from (2.29) and the convexity of $I(r; a_\varepsilon)$ one can derive the bound

$$(3.20) \quad |I(r,a)-I(r;a_\varepsilon)| \leqq \varepsilon(r''-r')^{-1}[I(r'';a_\varepsilon)-I(r';a_\varepsilon)], \text{ for } r \leqq r', \quad r+\varepsilon \leqq r''.$$

This method was used to calculate values of $I(r; a)$ for Spearman's rank correlation coefficient rho (Table 1); i.e. for $a(u, v) = 12(u-\frac{1}{2})(v-\frac{1}{2})$. The approximating function was $a_\varepsilon(u, v) = 12(u-\frac{1}{2})(k^{-1}(j-\frac{1}{2})-\frac{1}{2}), (j-1)/k \leqq v < j/k, 1 \leqq j \leqq k$, with $k = 320$. In order to use (3.20) one needs to know an upper bound for

$$(3.21) \qquad \varepsilon = \sup\{\textstyle\int\int |a-a_\varepsilon|\,h; h \in \mathscr{H}\}.$$

Since $a(u, v) - a_\varepsilon(u, v) = a(u, v+j/k) - a_\varepsilon(u, v+j/k)$ it is clear that one can, without loss of generality, assume $h(u, v) = h(u, v+j/k)$ in (3.21). Thus $h^*(u, v) = h(u, v/k)$ is in $\mathscr{H}$, and

$$\textstyle\int\int (a-a_\varepsilon)h = 12k^{-1}\int_0^1\int_0^1 (u-\frac{1}{2})(v-\frac{1}{2})h^*(u,v)\,du\,dv$$

$$\leqq 12k^{-1}[\textstyle\int_0^1 (u-\frac{1}{2})^2]^{\frac{1}{2}}[\int_0^1 (v-\frac{1}{2})^2]^{\frac{1}{2}} = 1/k.$$

Consequently $\varepsilon \leqq k^{-1} = .003125$. Error estimates using (3.20) are included in Table 1. Notice that for small and large r the error estimates are quite large; for small r this appears to be due to the crudeness of the error bound as the following argument suggests.

An expansion for $I(r; a)$. If one assumes that (3.2) has a solution $s(v; \lambda)$ and that $s(v; \lambda) = -\log[g(v; \lambda)]/\lambda$ can for $\lambda \approx 0$ be expressed as $g(v; \lambda) = 1 + \lambda g_1(v) + \lambda^2 g_2(v) + \cdots$; then by solving for $g_1(v), g_2(v), \cdots$ in (3.2) and substituting the result into (3.3) and (3.4) one obtains for $\lambda \approx 0$

$$(3.22) \qquad r = \lambda c_1 + \tfrac{1}{2}\lambda^2 c_2 + \tfrac{1}{6}\lambda^3 c_3 + \cdots$$

$$I(r; a) = \tfrac{1}{2}\lambda^2 c_1 + \tfrac{1}{3}\lambda^3 c_2 + \tfrac{1}{8}\lambda^4 c^4 + \cdots,$$

TABLE 1
Index of large deviations of Spearman's rho*

r	$I(r; rho)$†	Error‡ Bound	r	$I(r; rho)$	Error Bound
.03	$.0_3 4505$**	312	.57	.1886	14
.06	$.0_2 1803$	131	.60	.2133	13
.09	$.0_2 4063$	81	.63	.2406	13
.12	$.0_2 7240$	59	.66	.2708	13
.15	.01135	46	.69	.3046	13
.18	.01640	38	.72	.3426	13
.21	.02243	33	.75	.3857	13
.24	.02946	28	.78	.4354	14
.27	.03751	25	.81	.4935	15
.30	.04664	23	.84	.5634	16
.33	.05689	21	.87	.6496	18
.36	.06831	19	.90	.7613	21
.39	.08097	18	.93	.9172	29
.42	.09495	17	.96	1.169	—***
.45	.1103	16	.99	1.820	—***
.48	.1273	15	.999	2.944	—
.51	.1458	15	.9999	4.086	—
.54	.1662	15			

* Spearman's *rho* is $T_N = 12(N+1)^{-2}\Sigma R_j(j-\frac{1}{2}(N+1))$.
† $I(r; rho) = -\lim N^{-1} \log P[T_N \geqq Nr_N]$, when $r_N \to r$ as $N \to \infty$.
‡ The error bound is in parts per thousand and is obtained from (3.20).
** The notation 0_3 means 000.
*** From this point on $I(r; rho)$ is calculated from (3.26); no error estimate is available.

where

$$c_1 = \int\int a^2, \qquad c_2 = \int\int a^3,$$

$$c_3 = \int\int a^4 - 3\int\int\int a^2(u,v)a^2(u,w)\,du\,dv\,dw - 3\int\int\int a^2(u,v)a^2(w,v)\,du\,dv\,dw + 3\left(\int\int a^2\right)^2 \quad \text{and}$$

$$c_4 = \int\int a^4 - 2\int\int\int a^2(u,v)a^2(u,w)\,du\,dv\,dw - 2\int\int\int a^2(u,v)a^2(w,v)\,du\,dv\,dw - \int\left[\int a^2(u,v)\,du\right]dv - \int\left[\int a^2(u,v)\,dv\right]du + 3\left(\int\int a^2\right)^2, \quad \text{provided}$$

$$\int a(u,v)\,du = \int a(u,v)\,dv = 0. \tag{3.23}$$

If $a(u,v) = J(u)L(v)$, then the coefficients reduce to $c_1 = J_2L_2$, $c_2 = J_3L_3$, and $c_3 = c_4 = J_4L_4 - 3J_2{}^2L_4 - 3J_4L_2{}^2 + 3J_2{}^2L_2{}^2$, where $J_r = \int J^r, L_r = \int L^r$, provided $J_1 = L_1 = 0$.

By inverting (3.22) to obtain λ as a power series in r one obtains:

$$I(r;a) = r^2\left(\frac{1}{2c_1}\right) - r^3\left(\frac{c_2}{6c_1{}^3}\right) + r^4\left(\frac{c_4}{8c_1{}^4} - \frac{c_2{}^2}{8c_1{}^5} - \frac{c_3}{6c_1{}^4}\right) + \cdots. \tag{3.24}$$

If (3.23) is not true then one should replace $a(u, v)$ by $a^*(u, v) = a(u, v) - a_1(u) - a_2(v) + a_{12}$ where $a_1(u) = \int a(u, v)\,dv$, $a_2(v) = \int a(u, v)\,du$, $a_{12} = \iint a = \underline{r}(a)$. Also on the right side of (3.24) r should be replaced by $r - \underline{r}(a) = r - a_{12}$. These remarks follow from (2.30).

In particular for Spearman's rho (3.24) becomes

$$I(r; \text{rho}) = .5r^2 + .19r^4 + \cdots \tag{3.25}$$

This agrees with Table 1 within one part in a thousand in the range $.03 \leqq r \leqq .24$ and suggests considerably more accuracy there than does (3.20).

For large $r (r \approx 1)$ another intuitive argument[9] suggests

$$\begin{aligned} I(r; \text{rho}) &\doteq -\tfrac{1}{2}\log(1-r) - \tfrac{1}{2}(\log(\tfrac{1}{3}\pi) + 1) + .40604(1-r)^{\frac{1}{2}} + \cdots \\ &\doteq -\tfrac{1}{2}\log(1-r) - .52305 + .40604(1-r)^{\frac{1}{2}} + \cdots \end{aligned} \tag{3.26}$$

4. Bahadur efficiency of type A linear rank statistics. For a definition of Bahadur efficiency see Part II, Sections 4 and 5, of [2]. One method of evaluating Bahadur efficiency (given in [2]) is as follows:

Let $(\mathscr{X}_N, \mathscr{A}_N)$; $N = 1, 2, \cdots$, be measurable spaces and let $P = \{P_N\}$ and $Q = \{Q_N\}$ be sequences of measures on these spaces. $S_N^{(1)}$ and $S_N^{(2)}$ are two statistics defined on $(\mathscr{X}_N, \mathscr{A}_N)$. Suppose one uses rejection regions of the form $S_N^{(i)} \geqq k_N^{(i)}$, $i = 1$ or 2, to test the null hypothesis[10] P_N versus the alternative Q_N and the test statistics converge in probability under Q to some constants, call them $r_i(Q) i = 1, 2$; i.e.,

$$S_N^{(i)}/N \to_Q r_i(Q), \qquad i = 1, 2. \tag{4.1}$$

If there are continuous functions I_i, $i = 1, 2$, such that for any sequence of constants x_N converging to a constant x

$$\lim_{N\to\infty} \{-N^{-1} \log P[S_N^{(i)} \geqq Nx_N]\} = I_i(x), \qquad i = 1, 2. \tag{4.2}$$

then the Bahadur efficiency of $S_N^{(1)}$ compared to $S_N^{(2)}$ for rejecting P in favor of Q is

$$I_1(r_1(Q))/I_2(r_2(Q)), \tag{4.3}$$

provided numerator and denominator are neither zero nor infinity.

The quantity $2I_i(r_i(Q))$ is called the *exact slope* of $S_N^{(i)}$ at the alternative Q; thus (4.3) states that the Bahadur relative efficiency of one test statistic compared to another is the ratio of their exact slopes.

Notice that if T_N is a type A linear rank statistic (Definition 2.1) and if T_N/N converges in probability under a sequence Q of simple alternatives to, say, $r(Q)$ then, by Theorem 1 the exact slope of T_N against Q is

$$2I(r(Q); a). \tag{4.4}$$

---

[9] See [11] page 14 for details; note that r of this paper is 12 times the r of [11]. The constant and logarithmetic terms are thought to be exact.

[10] More generally we may take P_N (or Q_N) to be classes of distributions provided a test statistic has the same distribution throughout a class.

Thus to calculate exact slopes one needs to be able to evaluate the probability limit of T_N/N for alternative hypotheses of interest.

Two examples of probability limits.

EXAMPLE 1. T_N is a type A linear rank statistic and $R_1, \cdots, R_N$ are the ranks of N independent random variables $Z_1, \cdots, Z_N$. Under the alternative Q_N the Z's have distribution functions $F_{N1}, \cdots, F_{NN}$, which are assumed to have no discrete probability points in common.

Let $H_N^*(x, v) = N^{-1} \sum_{j=1}^N F_{Nj}(x) I[j \leqq Nv]$, where $I[\cdot]$ is the indicator function.

If $H_N^*(x, v)$ converges to a bivariate cdf $H^*(x, v)$ at continuity points of $H^*(x, v)$, then T_N/N converges in probability and the probability limit is given by:

$$T_N/N \to_Q \int_{-\infty}^{\infty} \int_0^1 a(F(x), v)\, dH^*(x, v), \tag{4.5}$$

where $F(x) = H^*(x, \infty)$.

PROOF. Define $N\hat{H}_N^*(x, v)$ to be the number of observations among $Z_1, \cdots, Z_j$, $j = [Nv]$, which are less than or equal to x. Then $\hat{F}_N(x) = \hat{H}_N^*(x, \infty)$ is the empirical cdf of $Z_1, \cdots, Z_N$ and, since $R_j = N\hat{F}_N(X_{Nj})$, it follows from (1.1) that

$$\begin{aligned} T_N/N &= \int_{-\infty}^{\infty} \int_0^1 a_N\left(\frac{N}{N+1} \hat{F}_N(x), \frac{N}{N+1} v\right) d\hat{H}_N^*(x, v) \\ &= \int_0^1 \int_0^1 a_N\left(\frac{N}{N+1} u, \frac{N}{N+1} v\right) d\hat{H}_N(u, v), \end{aligned}$$

where $\hat{H}_N(u, v)$ is the cdf which puts probability N^{-1} at the points $(R_j/N+1, j/N+1)$.

Now let a_ε^* be the uniformly continuous function guaranteed by Corollary 1, and let $\tilde{H}_N(u, v)$ be the cdf corresponding to the density $\tilde{h}_N$ of (2.31). If δ_N^* is the modulus of continuity of a_ε^* over a square of size N^{-1}, then clearly

$$\left|\iint a_\varepsilon^*\, d(\hat{H}_N - \tilde{H}_N)\right| \leqq 2\delta_N^*.$$

Thus $|T_N/N - \iint a_\varepsilon^*\, d\hat{H}_N| = |\iint a_N\, d\tilde{H}_N - \iint a_\varepsilon^*\, d\hat{H}_N| \leqq d(a_N, a_\varepsilon^*) + 2\delta_N^* \leqq \varepsilon + 2\delta_N^*$ for large N. Since

$$\iint a_\varepsilon^*\, d\hat{H}_N = \iint a_\varepsilon^*\left(\frac{N}{N+1} F_N(x), \frac{N}{N+1} v\right) d\hat{H}_N^*(x, v)$$

the result follows at once from the uniform continuity of a_ε^* and the easily verified fact $\sup_{x,v} |\hat{H}_N^*(x,v) - H^*(x, v)| \to_Q 0$.

EXAMPLE 2. $(X_1, Y_1), \cdots, (X_N, Y_N)$ are independent and identically distributed random vectors with cdf $H(x, y)$ having continuous marginals $F(x)$ and $G(y)$. The null hypothesis P is that $H = FG$ and the alternative Q is that H is some fixed cdf $\neq FG$.

Let $(X_{(j)}, Y_{[j]}), j = 1, \cdots, N$, be the sample arranged so that $X_{(1)} \leqq \cdots \leqq X_{(N)}$ and let $R_1, \cdots, R_N$ denote the ranks of $Y_{[1]}, \cdots, Y_{[N]}$. Under the null hypothesis,

$(R_1, \cdots, R_N)$ is equally likely to be any permutation of $(1, \cdots, N)$. If T_N is a type A linear rank statistic, then under the alternative Q

$$T_N/N \to_Q \iint a(F, G)\, dH. \tag{4.6}$$

Proof. Let $\hat{H}_N(x, y)$ be the bivariate empirical cdf and let $\hat{F}_N(x), \hat{G}_N(y)$ be its marginals, then clearly

$$T_N/N = \iint a_N\left(\frac{N}{N+1}\hat{F}_N, \frac{N}{N+1}\hat{G}_N\right) d\hat{H}_N.$$

An argument similar to the above yields the desired result.

To conclude this section the results of this paper are applied to several testing problems.

The two-sample case. In the two-sample case described in the introduction, let F and G denote the distributions and m and n the sample sizes of the X and Y samples, respectively. The null hypothesis P is that $F = G$, continuous, and the alternative Q is that F and G are some fixed cdf's, $F \neq G$. If $n/N \to \rho \neq 0, 1$ as $N \to \infty$ and $H_N{}^*$ is defined as in Example 1, then

$$\begin{aligned} H_N{}^*(x, v) \to H^*(x, v) &= F(x), \qquad 0 < v < \bar{\rho}, \\ &= G(x), \qquad \bar{\rho} \leqq v \leqq 1; \end{aligned}$$

where $\bar{\rho} = 1 - \rho$. If T_N is a two-sample scores statistic (see (1.2)) with score function J_N converging in quadratic mean to J, then as in the sentence containing (3.7) it follows that T_N is type A.

Thus by Example 1,

$$T_N/N \to_Q \rho \textstyle\int_{-\infty}^{\infty} J(\bar{\rho}F(x) + \rho G(x))\, dG(x), \tag{4.7}$$

If $G(x) = F(x - \theta)$, then

$$T_N/N \to_Q \rho \textstyle\int_{-\infty}^{\infty} J(\bar{\rho}F(x+\theta) + \rho F(x))\, dF(x); \tag{4.8}$$

an interesting special case is $\rho = \frac{1}{2}$ (equal samples), $F(x) = 1 - F(-x)$ and $J(u) = -J(1-u)$, in which case, after some manipulation, (4.7) becomes

(4.9 $\qquad T_N/N \to_Q \int_0^\infty J(\frac{1}{2}(F(x+\theta') + F(x-\theta')))\, dF(x+\theta') - \int_{\frac{1}{2}}^1 J(u)\, du,$

where $\theta' = \frac{1}{2}\theta$.

Exact slopes and Bahadur efficiencies of two-sample scores tests. It was remarked above (see (4.3)) that the Bahadur relative efficiency of two statistics is the ratio of their exact slopes. It follows from (4.4) and (4.7) that the exact slope at the alternative F, G of a two-sample scores statistic with score function J is $2I(r(F, G, \rho); J, \rho)$, where $r(F, G, \rho)$ is the right side of (4.7) and $I(r; J, \rho)$ is given by (3.12); this exact slope (without the 2) is tabulated in Tables 2 through Table 4 for $G(x) = F(x - \theta)$ normal, double exponential, logistic, $\rho = \frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{16}$, various θ values, and J corresponding to normal scores, Wilcoxon and median tests.

For the special case $\rho = \frac{1}{2}$, $J(u) = -J(1-u)$, and $F(x) = 1-F(-x)$ the exact slope at the alternative $G(x) = F(x-\theta)$ is twice the right side of (3.18) with r set equal to the right side of (4.9). An interesting numerical agreement between his work and that of Klotz [7] was pointed out by Hoadley [5] page 381; we are now in a position to explain this agreement. Notice that Klotz's "exponent" $e_s(\rho)$ ((1.4) page 1760) is[11] $I(\frac{1}{2}\rho; J, \frac{1}{2})$, the right side of (3.18) with our $J(u)$ equal to Klotz's $G^{-1}(2u-1)$ for $\frac{1}{2} \leqq u < 1$ and $J(u) = -J(1-u)$ for $0 < u < \frac{1}{2}$ and Klotz's ρ and h correspond to $2r$ and $\frac{1}{2}\lambda$ of this paper.

Let S_N denote the one-sample signed rank statistic given by Klotz's (1.2). Under the alternative Q that the observations in the sample are drawn from $F_\mu(x) = F(x-\mu)$ it is easy to see that $S_N/N \to_Q \rho(\mu) = 2r(\mu)$. where $r(\mu)$ is the right side of (4.9) with $\theta' = \mu, \rho = \frac{1}{2}$.

To obtain relative efficiencies, Klotz evaluates $e_s(\rho)$ with ρ equal to $\rho(\mu)$. Consequently Klotz's "exponent" $e_s(\rho(\mu))$ equals $I(r(\mu); J, \frac{1}{2})$ and

the entry in the "I" column under $\rho = \frac{1}{2}$ opposite θ of	equals	the entry opposite $\mu = \frac{1}{2}\theta$ of Klotz's
Table 2a		Table I, col. 3
Table 2b		Table I, col. 4
Table 2c		Table I, col. 2
Table 3, "median"		Table II, col. 4
Table 3, "Wilcoxon"		Table II, col. 2
Table 3, "normal scores"		Table II, col. 3
Table 4, "median"		Table II, col. 7
Table 4, "Wilcoxon"		Table II, col. 5
Table 4, "normal scores"		Table II, col. 6

By an argument which need not detail us here the "I" column of Table 2d under $\rho = \frac{1}{2}$ corresponds in a similar way to Klotz's Table I, column 6.

Exact slopes and Bahadur efficiencies of some tests of bivariate dependence. The Fisher-Yates normal scores correlation coefficient was defined in Example 1 of Section 3. In this case the right side of (4.6), call it $r(H; FYC)$, is $r(H; FYC) = \iint \Phi^{-1}(F(x))\Phi^{-1}(G(y))\,dH(x,y)$. In particular, if H has normal marginals, then $r(H; FYC)$ equals ρ, the product moment correlation between X and Y. From (3.5) and (4.4), the exact slope for testing independence of X and Y versus the alternative H is $-\log(1-r^2(H; FYC))$ or $-\log(1-\rho^2)$ in case H has normal marginals and $\rho > 0$. Let T_N' denote the sample product-moment correlation coefficient. From Klotz' [7] formula (3.3) and the fact that (under the null hypothesis that X and Y are independent and normally distributed) $(N-2)^{\frac{1}{2}}T_N'/(1-T_N'^2)^{\frac{1}{2}}$ is distributed as Student's t with $N-2$ degrees of freedom, it follows easily that the exact slope of T_N' for testing a normal null against any alternative H with correlation $\rho > 0$ is $-\log(1-\rho^2)$. Thus the Fisher-Yates correlation coefficient has Bahadur efficiency *one* relative to the product-moment correlation coefficient for testing the null

[11] This is a new use of the symbol "ρ".

TABLE 2a

Probability limit and exact slope (times $\frac{1}{2}$) of the two-sample normal scores test against the normal shift alternative: $F(x) = \Phi(x)$, $G(y) = F(y-\theta)$.

ρ	$\frac{1}{2}$		$\frac{1}{4}$	
θ	r	I	r	I
.25	.062017	$.0_27752$	.046601	$.0_25825$
.50	.12126	.03031	.091587	.02289
.75	.17551	.06577	.13349	.05003
1.00	.22333	.1114	.17116	.08544
1.25	.26407	.1642	.20388	.1269
1.50	.29773	.2211	.23140	.1721
1.75	.32474	.2793	.25382	.2189
2.00	.35807	.3365	.27156	.2652
2.25	.36180	.3908	.28520	.3095
2.50	.37361	.4408	.29538	.3506
2.75	.38209	.4860	.30277	.3878
3.00	.38802	.5255	.30798	.4206
3.25	.39204	.5594	.31155	.4490
3.50	.39470	.5881	.31393	.4730
3.75	.39640	.6117	.31546	.4929
4.00	.39746	.6310	.31642	.5092
4.25	.39811	.6464	.31701	.5223
4.50	.39848	.6584	.31735	.5325
4.75	.39870	.6678	.31755	.5405
5.00	.39881	.6749	.31766	.5466
5.25	.39888	.6802	.31772	.5511
5.50	.39891	.6841	.31775	.5549
5.75	.39893	.6869	.31776	.5569
6.00	.39894	.6890	.31777	.5587
∞	.39894	.6932	.31778	.5623

ρ	$\frac{1}{8}$		$\frac{1}{16}$	
θ	r	I	r	I
.25	.027250	$.0_23406$	.014621	$.0_21828$
.50	.053922	.01348	.029071	$.0_27268$
.75	.079396	.02977	.043140	.01618
1.00	.10303	.05147	.056558	.02827
1.25	.12426	.07746	.069019	.04307
1.50	.14271	.1064	.080229	.05995
1.75	.15820	.1369	.089979	.07815
2.00	.17080	.1677	.098172	.09687
2.25	.18072	.1975	.10482	.1154
2.50	.18830	.2256	.11005	.1331
2.75	.19392	.2513	.11402	.1496
3.00	.19795	.2743	.11695	.1644
3.25	.20077	.2943	.11903	.1775
3.50	.20267	.3114	.12047	.1888
3.75	.20392	.3257	.12143	.1984
4.00	.20471	.3375	.12205	.2064
4.25	.20519	.3470	.12244	.2129
4.50	.20549	.3545	.12268	.2181
4.75	.20566	.3604	.12282	.2222
5.00	.20575	.3649	.12290	.2253
5.25	.20580	.3683	.12294	.2277
5.50	.20582	.3707	.12296	.2295
5.75	.20584	.3724	.12297	.2308
6.00	.20585	.3735	.12298	.2316
∞	.20585	.3768	.12299	.2338

hypothesis that (x, y) are independent and normal versus the alternative hypothesis that (x, y) are dependent with normal marginals and positive correlation.

No comparison between the product-moment and normal-scores correlation coefficients has been made for null distributions with nonnormal marginals since the exact slope of the product-moment correlation coefficient in such cases is not known (at least, not by the author).

A widely used test statistic for bivariate dependence is *Spearman's rank correlation coefficient rho.*[12] Let $H(x,y)$ be a bivariate density with continuous marginals $F(x)$ and $G(y)$ such that $H(x,y) \neq F(x)\,G(y)$. It is well known, and follows from (4.6) that the probability limit of rho/N, call it $r(H;\text{rho})$, is

$$r(H;\text{rho}) = 12\int\int F(x)G(y)\,dH(x,y) - 3 \tag{4.10}$$

[12] $a(u,v) = 12(u-\frac{1}{2})(v-\frac{1}{2})$.

By (4.4), the exact slope of rho for testing independence against the simple alternative H is $2I(r(H;\text{ rho}); \text{rho})$. For example, when $H(x,y)$ is bivariate normal with correlation ρ the right side of (4.10) becomes $6\pi^{-1}\arctan(\rho(4-\rho^2)^{-\frac{1}{2}})$; for $\rho = .1256$ the above equals .12 so for this particular ρ-value the exact slope of *rho* is $2I(.12;\text{ rho}) \doteq .01448$ (from Table 1). In comparison, the exact slope of either the product-moment or normal-scores correlation coefficient is $-\log(1-\rho^2) \doteq .01590$ so the Bahadur efficiency against this alternative of *rho* compared to either of these other correlation coefficients is .01448/.01590 = .911.

Another competitor of Spearman's *rho* is Kendall's coefficient *tau*. It is shown in [11] that the exact slope of tau against the alternative H is $2e(r(H;\text{ tau}); \text{tau})$ where

$$r(H;\text{tau}) = 4\int\int H\,dH - 1 \tag{4.11}$$

$$(= 2\pi^{-1}\arctan(\rho(1-\rho^2)^{-\frac{1}{2}})\,,$$

when H is bivariate normal with correlation ρ)

TABLE 2b

Probability limit and exact slope (times $\frac{1}{2}$) of the two-sample median test against the normal shift alternative: $F(x) = \Phi(x)$, $G(y) = F(y-\theta)$.

ρ	$\frac{1}{2}$		$\frac{1}{4}$	
θ	r	I	r	I
.25	.049738	$.0_24956$	.037231	$.0_23707$
.50	.098706	.01961	.073448	.01455
.75	.14617	.04336	.10768	.03172
1.00	.19146	.07522	.13902	.05390
1.25	.23401	.1139	.16670	.07936
1.50	.27337	.1580	.19017	.1061
1.75	.30921	.2058	.20911	.1321
2.00	.34134	.2557	.22355	.1554
2.25	.36971	.3062	.23387	.1747
2.50	.39435	.3558	.24076	.1896
2.75	.41543	.4034	.24503	.2001
3.00	.43319	.4479	.24749	.2070
3.25	.44792	.4886	.24880	.2111
3.50	.45994	.5250	.24947	.2135
3.75	.46960	.5570	.24977	.2147
4.00	.47725	.5846	.24991	.2153
4.25	.48321	.6079	.24997	.2156
4.50	.48778	.6272	.24999	.2157
4.75	.49123	.6429		
5.00	.49379	.6554		
5.25	.49567	.6653		
5.50	.49702	.6728		
5.75	.49798	.6786		
6.00	.49865	.6829		
∞	$\frac{1}{2}$	.69315	$\frac{1}{4}$	.21576

ρ	$\frac{1}{8}$		$\frac{1}{16}$	
θ	r	I	r	I
.25	.021665	$.0_22155$	.011588	$.0_21152$
.50	.042434	$.0_28377$	.022593	$.0_24450$
.75	.061489	.01796	.032509	$.0_29448$
1.00	.078169	.02985	.040965	.01550
1.25	.092042	.04276	.047776	.02188
1.50	.10295	.05542	.052940	.02794
1.75	.11103	.06677	.056620	.03318
2.00	.11663	.07612	.059079	.03734
2.25	.12027	.08322	.060619	.04040
2.50	.12248	.08820	.061524	.04249
2.75	.12374	.09145	.062022	.04381
3.00	.12440	.09341	.062279	.04458
3.25	.12473	.09452	.062404	.04501
3.50	.12489	.09510	.062460	.04522
3.75	.12496	.09538	.062485	.04533
4.00	.12498	.09551	.062494	.04538
4.25	.12499	.09557	.062498	.04540
∞	$\frac{1}{8}$	.095602	$\frac{1}{16}$	.045406

and $e(r;\text{tau}) = \frac{1}{4}\lambda r + \frac{1}{2}\lambda + \log(\lambda) - \log(e^{\lambda}-1)$, λ being the solution of

$$r = 1+4\left[\int_0^{\lambda} \frac{x\,dx}{e^x-1} - \lambda\right]/\lambda^2.$$

An extensive table of $e(r;\text{tau})$ (Table 5), computed since [11] was written, is included here to facilitate comparisons between tau and rho; we emphasize that tau is not a linear rank statistic and cannot be handled by the methods of this paper.

Efficiencies against bivariate normal alternatives. From the above and (4.3) it follows that the Bahadur efficiency of *tau* with respect to *rho* against the bivariate normal alternative with correlation $\rho > 0$ is

$$e(2\pi^{-1}\arctan(\rho(1-\rho^2)^{-\frac{1}{2}});\text{tau})/I(6\pi^{-1}\arctan(\rho(4-\rho^2)^{-\frac{1}{2}});\text{rho}).$$

This ratio was calculated, using Table 1 and Table 5, for various values of ρ, it equals one for $\rho = 0$ and $\rho = 1$ and appears to be always greater than one but no

TABLE 2c

Probability limit and exact slope (times $\frac{1}{2}$) of the two-sample Wilcoxon test against the normal shift alternative: $F(x) = \Phi(x)$, $G(y) = F(y-\theta)$.

p	$\frac{1}{2}$		$\frac{1}{4}$		$\frac{1}{8}$		$\frac{1}{16}$	
θ	r	I	r	I	r	I	r	I
.25	.01754	$.0_27416$	.01315	$.0_25564$	$.0_27674$	$.0_23247$	$.0_24111$	$.0_22421$
.50	.03454	.02914	.02591	.02189	.01511	.01279	$.0_28095$	$.0_26860$
.75	.05051	.06368	.03789	.04793	.02210	.02808	.01284	.01509
1.00	.06506	.1087	.04880	.08208	.02847	.04827	.01525	.02600
1.25	.07791	.1615	.05843	.1224	.03408	.07234	.01826	.03910
1.50	.08889	.2189	.06667	.1667	.03889	.09921	.02083	.05390
1.75	.09801	.2779	.07351	.2129	.04288	.1279	.02297	.06995
2.00	.10534	.3360	.07900	.2592	.04609	.1574	.02469	.08688
2.25	.11105	.3910	.08329	.3040	.04858	.1869	.02603	.1043
2.50	.11536	.4416	.08652	.3459	.05047	.2156	.02704	.1219
2.75	.11852	.4867	.08889	.3841	.05185	.2426	.02778	.1391
3.00	.12076	.5262	.09057	.4179	.05283	.2672	.02830	.1554
3.25	.12231	.5600	.09173	.4471	.05351	.2890	.02867	.1703
3.50	.12333	.5885	.09250	.4717	.05396	.3076	.02891	.1834
3.75	.12400	.6120	.09300	.4920	.05425	.3231	.02906	.1946
4.00	.12442	.6311	.09331	.5086	.05443	.3357	.02916	.2037
4.25	.12467	.6464	.09350	.5219	.05454	.3459	.02922	.2112
4.50	.12482	.6585	.09361	.5323	.05461	.3538	.02925	.2170
4.75	.12490	.6678	.09368	.5404	.05464	.3600	.02927	.2215
5.00	.12495	.6749	.09371	.5468	.05466	.3647	.029285	.2249
5.25	.12497	.6801	.09373	.5511	.05468	.3682	.029291	.2275
5.50	.12499	.6841	.09374	.5544	.05468	.3708	.029293	.2294
5.75	.12499	.6869	.09375	.5569	.05468	.3726	.029295	.2308
6.00	.12500	.6889	.09375	.5586	.05469	.3740	.029296	.2317
∞	.125	.6932	.09375	.5623	7/128 $\doteq$.05469	.3768	15/512 $\doteq$.029297	.2338

TABLE 2d

Exact slope (times $\frac{1}{2}$) of the two-sample t test against the normal shift alternative: $F(x) = \Phi(x)$, $G(y) = F(y-\theta)$.

ρ	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{16}$
θ	I	I	I	I
.25	$.0_2 77521$	$.0_2 58253$	$.0_2 34063$	$.0_2 18277$
.50	.030312	.022905	.013488	$.0_2 72711$
.75	.065788	.050135	.029853	.016214
1.00	.11157	.085425	.051898	.028471
1.25	.16488	.12847	.078886	.043801
1.50	.22314	.17599	.11001	.061921
1.75	.28425	.22688	.14445	.082521
2.00	.34657	.27981	.18145	.10528
2.25	.40893	.33371	.22032	.12988
2.50	.47049	.38780	.26047	.15602
2.75	.53074	.44146	.30138	.18340
3.00	.58933	.49431	.34265	.21177
3.25	.64608	.54604	.38396	.24087
3.50	.70090	.59649	.42504	.27051
3.75	.75377	.64554	.46570	.30051
4.00	.80472	.69315	.50580	.33070
4.25	.85379	.73929	.54522	.36095
4.50	.90106	.78398	.58389	.39116
4.75	.94660	.82725	.62176	.42122
5.00	.99050	.86914	.65879	.45106
5.25	1.0328	.90968	.69497	.48063
5.50	1.0737	.94895	.73030	.50987
5.75	1.1132	.98698	.76479	.53874
6.00	1.1513	1.0238	.79843	.56721

greater than about 1.05 (this value occurs near $\rho = .85$). Because of the large error bounds in Table 1 the above findings are tentative; nevertheless, the author conjectures that *tau* is more efficient than *rho* against the normal alternative for all positive ρ-values. This means that for large N if tau and rho are adjusted to have equal power, then tau will have the smaller type I error; this must be contrasted with van der Waerden's finding [10] that for small and moderate N, the reverse is true.

Let Φ_ρ denote the bivariate normal cdf with zero means, unit variances and correlation ρ; the best test of $H = \Phi_0$ versus $H = \Phi_\rho$ is of course the "simple vs simple" likelihood ratio test (LRT). In [11] it was found that this test has the same exact slope, $-\log(1-\rho^2)$, as the product moment correlation test (PMCT) which was shown above to have the same exact slope as the normal scores correlation test (NSCT). Calculations reported in Figure 4 of [11] show that the Bahadur efficiency of *tau* with respect to the LRT is always greater than the Pitman efficiency $(3/\pi)^2$ and increases to one as $\rho \to 1$.

Summary of efficiency relations of tests of independence against the bivariate normal alternative

LRT = PMCT = NSCT > tau
tau[13] $\geqq$ LRT $\cdot (3/\pi)^2$
1.05 $\cdot$ rho[13] $\geqq$ tau $\geqq$ rho

If one balances efficiency against ease of calculation of the test statistic and availability of tables of its critical values, tau would seem to emerge as the best choice of the nonparametric tests. Of course this statement so far applies only to the normal alternative for large (perhaps *very* large) N. For nonnormal alternatives *tau* need not be more efficient than rho, for example, if the alternative is $H(x, y) = xy(1+\theta(1-x)(1-y))$, $0 \leqq x, y \leqq 1$, $0 < \theta \leqq 1$, then (4.10) and (4.11) become $r(H; \text{rho}) = \theta/3$ and $r(H; \text{tau}) = 2\theta/9$. Thus when $\theta = .36$ $I(r(H; \text{rho}); \text{rho}) = I(.12; \text{rho}) = .007240$ while $e(r(H; \text{tau}); \text{tau}) = e(.08; \text{tau}) = .007219$.

APPENDIX

A fixed point lemma. Let $K_j(u)$, $j = 1, \cdots, k$, be almost everywhere (a.e.) positive functions on $0 < u < 1$ and define $K.(u) = \sum_{j=1}^k \rho_j K_j(u)$, where $\rho_1, \cdots, \rho_k$ are positive and $\sum \rho_j = 1$.

LEMMA 1. *If $K_j(u)/K.(u)$ is bounded away from zero, say $K_j(u)/K.(u) > a > 0$, $0 < u < 1, j = 1, \cdots, k$, then there exist constants $g_1, \cdots, g_k$ such that*

$$\sum \rho_j g_j = 1, \qquad a \leqq g_j \leqq b = 1/\min(\rho_j), \quad \textit{and} \tag{A.1}$$

$$g_j = \int_0^1 \frac{K_j(u)}{\sum \rho_i K_i(u)/g_i}\, du, \qquad j = 1, \cdots, k. \tag{A.2}$$

PROOF. Define the function $\mathbf{T}(\mathbf{x}) = (T_1(\mathbf{x}), \cdots, T_k(\mathbf{x}))$, mapping the positive part of k-dimensional Euclidean space into itself, as follows

$$T_j(\mathbf{x}) = \int_0^1 \frac{K_j(u)}{K.(u)} \frac{K.(u)}{\sum \rho_i K_i(u)/x_i}\, du \div \int_0^1 \frac{K.(u)}{\sum \rho_i K_i(u)/x_i}\, du$$

$$= \int_0^1 \frac{K_j(u)}{K.(u)} m(u; \mathbf{x})\, du, \quad \text{say.}$$

Since $m(u; \mathbf{x})$ is a probability density it is clear that $\sum \rho_j T_j(\mathbf{x}) = 1$ and since $a \leqq K_j(u)/K.(u) \leqq b$ it follows that also $a \leqq T_j(\mathbf{x}) \leqq b$. Thus $\mathbf{T}$ maps the compact convex set $A = \{\mathbf{x}: \sum \rho_j x_j = 1, a \leqq x_j \leqq b\}$ into a subset of itself. Since $\mathbf{T}$ is

[13] Conjectures based on numerical calculations.

continuous on this set there must, by the Brouwer fixed point theorem, exist a point $\boldsymbol{g} \in \mathrm{A}$ such that $\mathbf{T}(\boldsymbol{g}) = \boldsymbol{g}$. Thus,

$$g_j = c \int_0^1 \frac{K_j(u)}{\sum \rho_i K_i(u)/g_i} du, \qquad j = 1, \cdots, k, \quad \text{where} \tag{A.3}$$

$$c^{-1} = \int_0^1 \frac{K.(u)}{\sum \rho_i K_i(u)/g_i} du.$$

Multiplying each side of (A.3) by ρ_j/g_j and summing, one sees that $c = 1$. □

TABLE 3

Exact slopes (times $\frac{1}{2}$) of nonparametric tests against the logistic shift alternative: $F(x) = (1+e^{-x})^{-1}$, $G(y) = F(y-\theta)$.

$\rho = \frac{1}{2}$				$\rho = \frac{1}{4}$			
θ	Median	Wilcoxon	Normal scores	θ	Median	Wilcoxon	Normal scores
.5	.007752	.01031	.009851	.5	.005780	.007736	.007398
1.0	.03030	.03999	.03828	1.0	.02220	.03006	.02886
1.5	.06566	.08561	.08214	1.5	.04668	.06452	.06234
2.0	.1109	.1424	.1370	2.0	.07564	.1077	.1048
2.5	.1628	.2050	.1980	2.5	.1053	.1559	.1527
3.0	.2181	.2689	.2606	3.0	.1325	.2058	.2027
3.5	.2738	.3001	.3215	3.5	.1554	.2547	.2519
4.0	.3278	.3869	.3781	4.0	.1733	.3006	.2982
4.5	.3784	.4373	.4290	4.5	.1866	.3423	.3403
5.0	.4246	.4810	.4739	5.0	.1961	.3793	.3774
5.5	.4659	.5184	.5124	5.5	.2027	.4112	.4095
6.0	.5023	.5500	.5451	6.0	.2072	.4384	.4369
6.5	.5338	.5763	.5725	6.5	.2102	.4612	.4599
7.0	.5608	.5982	.5952	7.0	.2121	.4801	.4791
7.5	.5837	.6162	.6140	7.5	.2134	.4957	.4949
8.0	.6030	.6309	.6294	8.0	.2143	.5085	.5079
8.5	.6192	.6430	.6420	8.5	.2148	.5189	.5186
9.0	.6327	.6528	.6522	9.0	.2151	.5274	.5273
9.5	.6438	.6607	.6605	9.5	.2154	.5342	.5343
10.0	.6530	.6671	.6672	10.0	.2155	.5398	.5400
10.5	.6605	.6724	.6725	10.5		.5444	.5446
11.0	.6667	.6766	.6766	11.0		.5480	.5482
11.5	.6717	.6799	.6798	11.5		.5509	.5511
12.0	.6758	.6826	.6822	12.0		.5532	.5533
12.5	.6792	.6847	.6840	12.5		.5550	.5550
13.0	.6819	.6864	.6852	13.0		.5565	.5562
13.5	.6840	.6878	.6861	13.5		.5577	.5571
14.0	.6859	.6889	.6867	14.0		.5586	.5578
14.5	.6873	.6897	.6871	14.5		.5594	.5582
15.0	.6885	.6904	.6874	15.0		.5599	.5585
∞	.6932	.6932	.6932	∞	.2158	.5623	.5623

LEMMA 2. *Lemma* 1 *remains true with* $a = 0$; *moreover* $g_j > 0, j = 1, \cdots, k$.

PROOF. For arbitrary $\delta > 0$ replace $K_j(u)$ by $\delta + \min(K_j(u), \delta^{-1})$; then the conditions of Lemma 1 are satisfied. Let $g_1(\delta), \cdots, g_k(\delta)$ denote a solution of (A.2) satisfying (A.1) for this modification of K_j. Since $0 \leqq g_j(\delta) \leqq b$, there exists a sequence $\delta_n \to 0$ as $n \to \infty$ such that $g_j(\delta_n) \to g_j$, say. Let $K_{jn}(u) = \delta_n + \min(K_j(u), \delta_n^{-1})$, and $g_{jn} = g_j(\delta_n)$. Since

$$\frac{K_{jn}(u)}{\sum_i \rho_i K_{in}(u)/g_{in}} \leqq \frac{g_{jn}}{\rho_j} \leqq b^2$$

TABLE 3—*continued*

$\rho = \frac{1}{8}$				$\rho = \frac{1}{16}$			
θ	Median	Wilcoxon	Normal scores	θ	Median	Wilcoxon	Normal scores
.5	.003347	.004516	.004323	.5	.001785	.002420	.002318
1.0	.01259	.01758	.01695	1.0	.006628	.009434	.009120
1.5	.02566	.03787	.03692	1.5	.01327	.02073	.01998
2.0	.04004	.06355	.06279	2.0	.02032	.03431	.03424
2.5	.05361	.09267	.09274	2.5	.02674	.05028	.05114
3.0	.06512	.1023	.1248	3.0	.03201	.06745	.06972
3.5	.07414	.1545	.1572	3.5	.03606	.08519	.08902
4.0	.08085	.1847	.1883	4.0	.03902	.1030	.1081
4.5	.08563	.2131	.2172	4.5	.04111	.1203	.1264
5.0	.08895	.2391	.2430	5.0	.04254	.1368	.1430
5.5	.09121	.2623	.2657	5.5	.04352	.1521	.1578
6.0	.09273	.2825	.2852	6.0	.04417	.1658	.1707
6.5	.09373	.2997	.3017	6.5	.04460	.1177	.1818
7.0	.09439	.3140	.3156	7.0	.04489	.1880	.1912
7.5	.09482	.3259	.3270	7.5	.04507	.1966	.1991
8.0	.09510	.3356	.3365	8.0	.04519	.2037	.2056
8.5	.09528	.3436	.3443	8.5	.04527	.2095	.2110
9.0	.09540	.3501	.3507	9.0	.04532	.2143	.2154
9.5	.09547	.3554	.3559	9.5	.04535	.2181	.2190
10.0	.09552	.3596	.3601	10.0	.04537	.2212	.2219
10.5		.3630	.3635	10.5		.2237	.2243
11.0		.3658	.3662	11.0		.2258	.2262
11.5		.3680	.3684	11.5		.2274	.2278
12.0		.3698	.3701	12.0		.2287	.2290
12.5		.3712	.3714	12.5		.2297	.2300
13.0		.3723	.3724	13.0		.2306	.2307
13.5		.3733	.3732	13.5		.2312	.2313
14.0		.3740	.3737	14.0		.2317	.2318
14.5		.3745	.3741	14.5		.2321	.2321
15.0		.3750	.3744	15.0		.2325	.2323
∞	.09560	.3768	.3768	∞	.04541	.2338	.2338

TABLE 4

Exact slopes (times $\frac{1}{2}$) of nonparametric tests against the double exponential shift alternative: $F(x) = \frac{1}{2}[1+\text{sgn}(x)(1-\exp(-|x|))]$, $G(y) = F(y-\theta)$.

	$\rho = \frac{1}{2}$				$\rho = \frac{1}{4}$		
θ	Median	Wilcoxon	Normal scores	θ	Median	Wilcoxon	Normal scores
.5	.02467	.02223	.01908	.5	.01739	.01669	.01433
1.0	.07954	.07910	.06939	1.0	.05259	.05959	.05243
1.5	.1465	.1536	.1380	1.5	.09043	.1163	.1052
2.0	.2158	.2325	.2133	2.0	.1240	.1772	.1643
2.5	.2823	.3074	.2872	2.5	.1508	.2363	.2235
3.0	.3434	.3744	.3551	3.0	.1710	.2904	.2788
3.5	.3979	.4321	.4149	3.5	.1854	.3380	.3281
4.0	.4456	.4806	.4661	4.0	.1955	.3789	.3705
4.5	.4868	.5208	.5089	4.5	.2024	.4133	.4064
5.0	.5219	.5539	.5444	5.0	.2070	.4418	.4362
5.5	.5516	.5809	.5734	5.5	.2101	.4651	.4606
6.0	.5766	.6028	.5971	6.0	.2121	.4841	.4806
6.5	.5975	.6207	.6162	6.5	.2134	.4996	.4968
7.0	.6149	.6351	.6316	7.0	.2142	.5120	.5099
7.5	.6292	.6467	.6441	7.5	.2148	.5221	.5205
8.0	.6411	.6560	.6541	8.0	.2151	.5300	.5290
8.5	.6508	.6635	.6621	8.5	.2154	.5366	.5358
9.0	.6588	.6695	.6685	9.0	.2155	.5418	.5413
9.5	.6653	.6721	.6737	9.5	.2156	.5460	.5457
10.0	.6706	.6782	.6779	10.0	.2157	.5494	.5492
10.5	.6750	.6813	.6812	10.5		.5520	.5520
11.0	.6785	.6837	.6838	11.0		.5541	.5542
11.5	.6813	.6856	.6858	11.5		.5558	.5560
12.0	.6836	.6872	.6874	12.0		.5572	.5574
12.5	.6855	.6884	.6887	12.5		.5582	.5585
13.0	.6870	.6894	.6897	13.0		.5591	.5593
13.5	.6882	.6901	.6904	13.5		.5598	.5600
14.0	.6892	.6908	.6910	14.0		.5603	.5605
14.5	.6900	.6913	.6915	14.5		.5607	.5609
15.0	.6906	.6916	.6919	15.0		.5609	.5612
∞	.6932	.6932	.6932	∞	.2158	.5623	.5623

it follows from the dominated convergence theorem that

$$\text{(A.4)} \quad g_j = \lim_{n\to\infty} g_{jn} = \lim_{n\to\infty} \int_0^1 \frac{K_{jn}(u)}{\sum_i \rho_i K_{in}(u)/g_{in}}\, du = \int \frac{K_j(u)}{\sum_i \rho_i K_i(u)/g_i}\, du.$$

Also

$$\text{(A.5)} \qquad \sum \rho_j g_j = \lim_{n\to\infty} \sum \rho_j g_{jn} = 1.$$

Thus it is not possible that, for example, $g_1 = 0$ since then $K_j(u)/\sum \rho_i K_i(u)/g_i) = 0$, which would by (A.4) imply that $g_j = 0$ for all j contrary to (A.5). □

TABLE 4—*continued*

$\rho = \frac{1}{8}$				$\rho = \frac{1}{16}$			
θ	Median	Wilcoxon	Normal scores	θ	Median	Wilcoxon	Normal scores
.5	.009471	.009750	.008379	.5	.004870	.005228	.004495
1.0	.02710	.03496	.03086	1.0	.01356	.01880	.01662
1.5	.04459	.06872	.06264	1.5	.02191	.03713	.03398
2.0	.05914	.1057	.09935	2.0	.02871	.05752	.05449
2.5	.07024	.1427	.1374	2.5	.03384	.07839	.07646
3.0	.07830	.1778	.1741	3.0	.03753	.09889	.09847
3.5	.08396	.2101	.2078	3.5	.04011	.1185	.1195
4.0	.08786	.2389	.2375	4.0	.04188	.1367	.1385
4.5	.09049	.2638	.2630	4.5	.04308	.1531	.1553
5.0	.09225	.2850	.2843	5.0	.04388	.1675	.1696
5.5	.09342	.3026	.3020	5.5	.04441	.1799	.1816
6.0	.09419	.3171	.3165	6.0	.04476	.1902	.1916
6.5	.09469	.3288	.3284	6.5	.04499	.1987	.1998
7.0	.09502	.3383	.3380	7.0	.04513	.2057	.2065
7.5	.05923	.3460	.3457	7.5	.04524	.2113	.2119
8.0	.09536	.3522	3520	8.0	.04530	.2158	.2163
8.5	.09545	.3572	.3570	8.5	.04534	.2194	.2198
9.0	.09551	.3611	.3611	9.0	.04536	.2224	.2226
9.5	.09554	.3643	.3643	9.5	.04538	.2247	.2249
10.0	.09556	.3669	.3669	10.0	.04539	.2265	.2267
10.5		.3689	.3690	10.5		.2280	.2282
11.0		.3705	.3707	11.0		.2292	.2294
11.5		.3718	.3720	11.5		.2302	.2303
12.0		.3728	.3730	12.0		.2309	.2311
12.5		.3737	.3739	12.5		.2315	.2317
13.0		.3743	.3745	13.0		.2320	.2321
13.5		.3748	.3750	13.5		.2324	.2325
14.0		.3752	.3754	14.0		.2327	.2328
14.5		.3755	.3757	14.5		.2329	.2330
15.0		.3758	.3759	15.0		.2331	.2332
∞	.09560	.3768	.3768	∞	.04541	.2338	.2338

LEMMA 3. *There is only one solution to* (A.2) *satisfying* $\sum \rho_j g_j = 1$.

PROOF. Let g_j' be any other non-zero solution to (A.2) (with $a = 0$ as in Lemma 2). By dividing each side of (A.2) by a constant one can assume $\sum \rho_j g_j' = 1$.

Let $g_j^* = \frac{1}{2}(g_j + g_j')$. It follows from the concavity of $(1/x+1/y)^{-1}$ that for each u

$$(\textstyle\sum_j \rho_j K_j(u)/g_j^*)^{-1} \geqq \frac{1}{2}(\sum_j \rho_j K_j(u)/g_j')^{-1} + \frac{1}{2}(\sum_j \rho_j K_j(u)/g_j)^{-1}$$

TABLE 5
Index of Large Deviations of Kendall's tau*

r	*e*(*r*)†	*r*	*e*	*r*	*e*
.020	$.0_{3}44998$	.184	.038624	.348	.14355
.024	$.0_{3}64813$	.188	.040346	.352	.14706
.028	$.0_{3}88230$	.192	.042108	.356	.15062
.032	$.0_{2}11525$	.196	.043910	.360	.15423
.036	$.0_{2}14588$	.200	.045751	.364	.15789
.040	$.0_{2}18012$	.204	.047632	.368	.16160
.044	$.0_{2}21797$	.208	.049552	.372	.16536
.048	$.0_{2}25945$	.212	.051513	.376	.16918
.052	$.0_{2}30454$	.216	.053514	.380	.17304
.056	$.0_{2}35325$	.220	.055556	.384	.17696
.060	$.0_{2}40560$	.224	.057638	.388	.18093
.064	$.0_{2}46157$	.228	.059761	.392	.18495
.068	$.0_{2}52119$	.232	.061925	.396	.18903
.072	$.0_{2}58444$	.236	.064130	.400	.19317
.076	$.0_{2}65134$	.240	.066377	.404	.19736
.080	$.0_{2}72189$	.244	.068665	.408	.20160
.084	$.0_{2}79609$	.248	.070995	.412	.20590
.088	$.0_{2}87396$	.252	.073367	.416	.21026
.092	$.0_{2}95550$	.256	.075781	.420	.21467
.096	.010407	.260	.078238	.424	.21915
.100	.011296	.264	.080737	.428	.22368
.104	.012222	.268	.083280	.432	.22827
.108	.013185	.272	.085865	.436	.23292
.112	.014185	.276	.088494	.440	.23763
.116	.015222	.280	.091166	.444	.24240
.120	.016296	.284	.093883	.448	.24723
.124	.017407	.288	.096644	.452	.25212
.128	.18556	.292	.099449	.456	.25708
.132	.19742	.296	.10230	.460	.26210
.136	.20966	.300	.10519	.464	.26719
.140	.22228	.304	.10813	.468	.27234
.144	.23527	.308	.11112	.472	.27755
.148	.24864	.312	.11415	.476	.28283
.152	.26240	.316	.11723	.480	.28818
.156	.27653	.320	.12035	.484	.29360
.160	.29105	.324	.12352	.488	.29909
.164	.30594	.328	.12674	.492	.30465
.168	.32123	.332	.13009	.496	.31028
.172	.33690	.336	.13332	.500	.31598
.176	.35295	.340	.13669	.504	.32175
.180	.36940	.344	.14010	.508	.32760

* Kendall's *tau* is $S_N = [N(N-1)]^{-1}\Sigma\Sigma\,\mathrm{sgn}\,(j-i)\,\mathrm{sgn}\,(R_j - R_i)$.
† $e(r) = e(r;\mathrm{tau}) = -\lim N^{-1}\log P[S_N \geqq Nr_N]$, when $r_N \to r$ as $N \to \infty$

TABLE 5—*continued*

r	$e(r)$	r	e	r	e
.512	.33352	.676	.65758	.840	1.2871
.516	.33951	.680	.66805	.844	1.3106
.520	.34559	.684	.67869	.848	1.3348
.524	.35174	.688	.68948	.852	1.3597
.528	.35797	.692	.70045	.856	1.3854
.532	.36428	.696	.71159	.860	1.4118
.536	.37067	.700	.72291	.864	1.4390
.540	.37715	.704	.73442	.868	1.4671
.544	.38370	.708	.74611	.872	1.4961
.548	.39034	.712	.75799	.876	1.5261
.552	.39707	.716	.77007	.880	1.5572
.556	.40389	.720	.78236	.884	1.5894
.560	.41079	.724	.79485	.888	1.6227
.564	.41778	.728	.80756	.892	1.6574
.568	.42487	.732	.82049	.896	1.6934
.572	.43205	.736	.83365	.900	1.7309
.576	.43932	.740	.84704	.904	1.7700
.580	.44669	.744	.86067	.908	1.8108
.584	.45416	.748	.87456	.912	1.8536
.588	.46172	.752	.88869	.916	1.8984
.592	.46940	.756	.90310	.920	1.9455
.596	.47717	.760	.91777	.924	1.9951
.600	.48505	.764	.93273	.928	2.0474
.604	.49303	.768	.94798	.932	2.1029
.608	.50112	.772	.96356	.936	2.1618
.612	.50932	.776	.97942	.940	2.2247
.616	.51764	.780	.99561	.944	2.2920
.620	.52607	.784	1.0121	.948	2.3644
.624	.53462	.788	1.0290	.952	2.4428
.628	.54329	.792	1.0462	.956	2.5281
.632	.55207	.796	1.0638	.960	2.6217
.636	.56099	.800	1.0818	.964	2.7254
.640	.57003	.804	1.1002	.968	2.8416
.644	.57920	.808	1.1190	.972	2.9734
.648	.58850	.812	1.1383	.976	3.1259
.652	.59794	.816	1.1580	.980	3.3066
.656	.60751	.820	1.1782	.984	3.5280
.660	.61723	.824	1.1989	.988	3.8141
.664	.62710	.828	1.2201	.992	4.2179
.668	.63711	.832	1.2418	.996	4.9093
.672	.64727	.836	1.2641		

with equality only if $K_j(u)\cdot(g_j-g_j')=0$ for $j=1,\cdots,k$. But $K_j(u)>0$ for almost all u, thus if $g_j\neq g_j'$ for some j, then

$$\int_0^1 \frac{K_j(u)}{\sum_i \rho_i K_i(u)/g_i^*}\,du > (g_j+g_j')/2 = g_j^*$$

which is impossible. □

COROLLARY 1. *For each* $\lambda \geqq 0$ *there exists a solution* $s_1(\lambda),\cdots,s_k(\lambda)$ *to* (3.10) *and if* $s_1'(\lambda),\cdots,s_k'(\lambda)$ *is any other solution then* $s_j(\lambda)-s_j'(\lambda)$ *is constant in* j *for* $j=1,\cdots,k$.

PROOF. Set $K_j(u)=\exp[\lambda J_j(u)]$. □

Continuity of $m(\lambda)$. Define $m(\lambda)$ as in (3.12) and let $g_j(\lambda)=\exp[\lambda s_j(\lambda)]$. By Corollary 1 one can assume without loss of generality that

(A.6) $$\sum \rho_j g_j(\lambda) = 1.$$

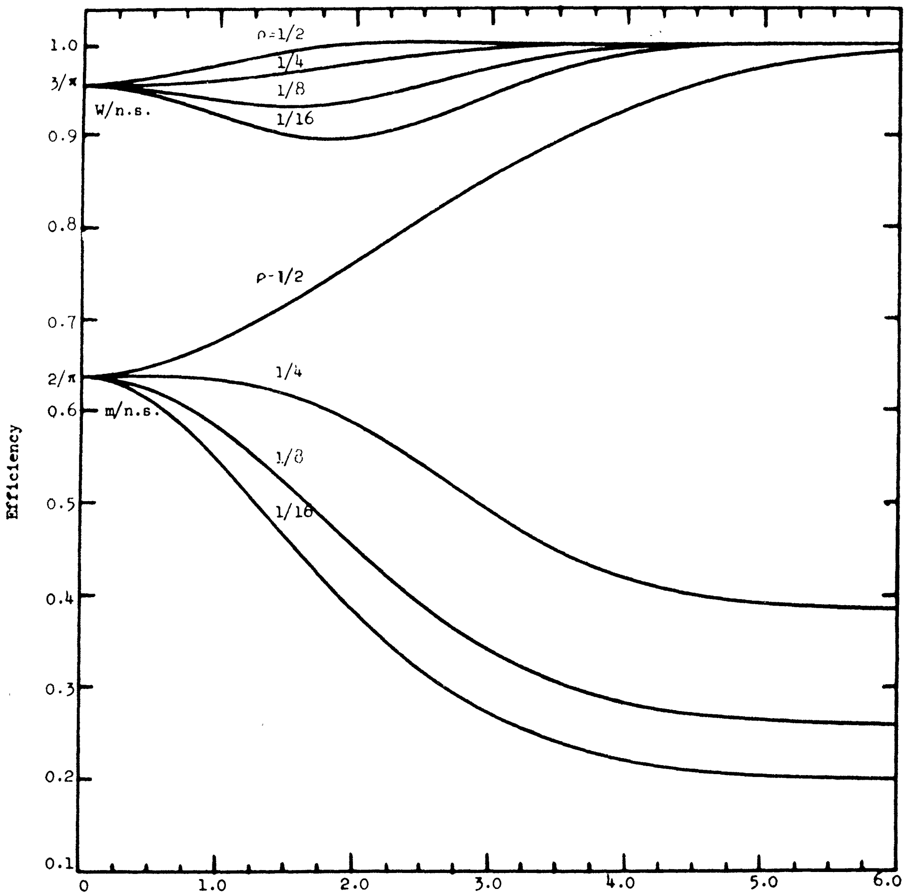

FIG. 1. Bahadur efficiencies for normal shift alternatives.

LEMMA 4. *$m(\lambda)$ is continuous in λ.*

PROOF. Rewrite (3.12) as

$$m(\lambda) = \int\int a(u, v)h_\lambda(u, v)\, du\, dv, \tag{A.7}$$

where $a(u, v)$ is defined by (3.7) and

$$h_\lambda(u, v) = \exp[\lambda(J_j(u) - s_j(\lambda))]/\sum_i \rho_i \exp[\lambda(J_i(u) - s_i(\lambda))], \quad v_{j-1} \leqq v < v_j, \tag{A.8}$$

is, by (3.10), a density with uniform marginals. Thus, by Corollary 1 of Theorem 2, for any $\varepsilon > 0$ there is a bounded, continuous $a_\varepsilon{}^*$ such that $|m(\lambda) - m(\lambda')| \leqq |\int\int a_\varepsilon{}^*(h_\lambda - h_{\lambda'})] + 2\varepsilon$; since $a_\varepsilon{}^*$ is bounded and $h_\lambda \leqq b = 1/\min(\rho_j)$ it is, by the dominated convergence theorem, sufficient to show that the $g_j(\lambda)$, subject to (A.6), are continuous in λ.

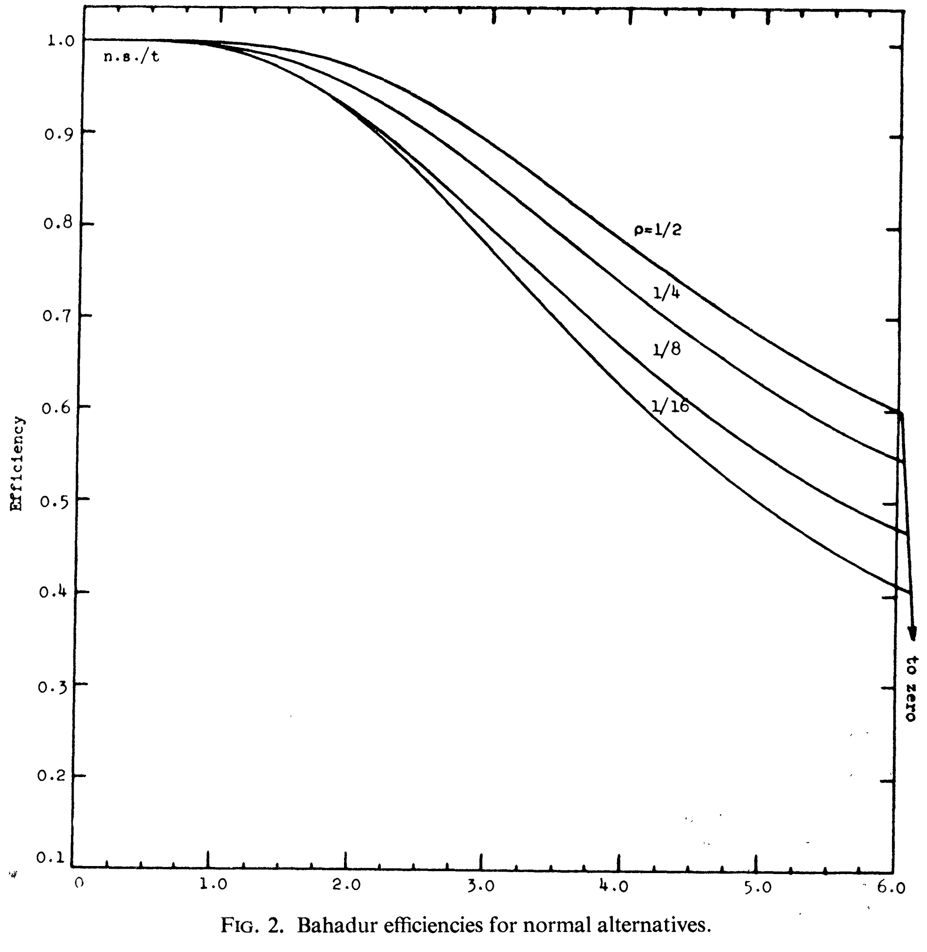

FIG. 2. Bahadur efficiencies for normal alternatives.

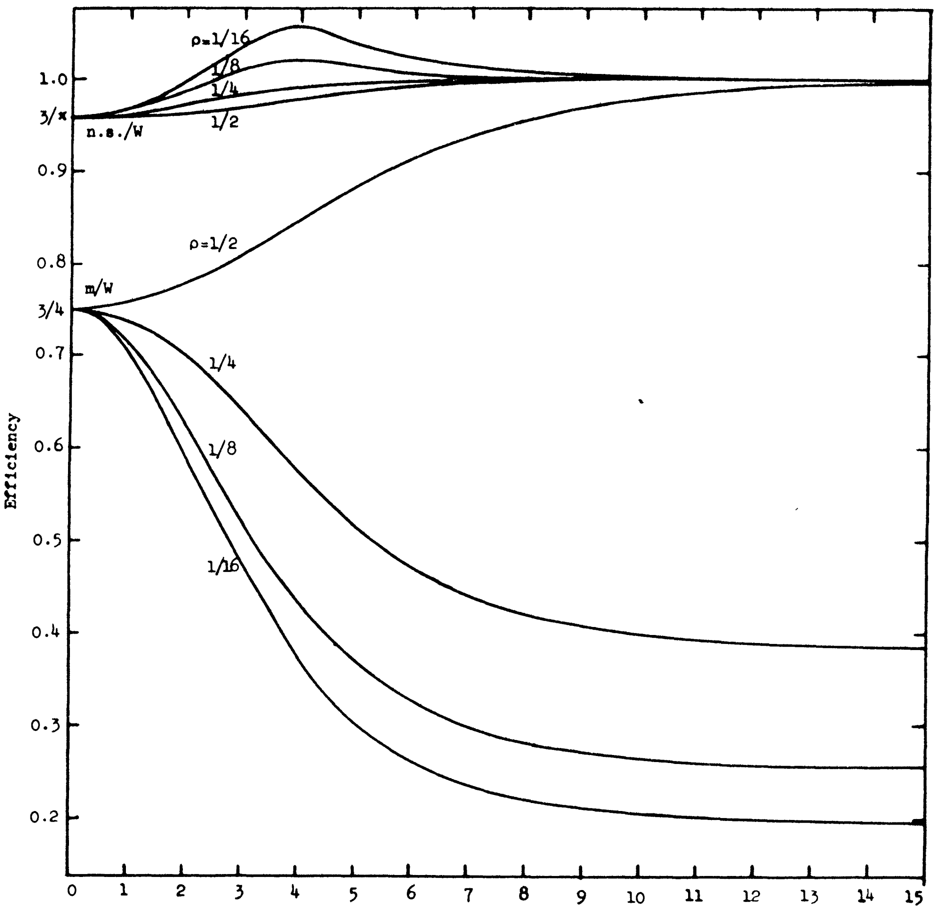

FIG. 3. Bahadur efficiencies for logistic shift alternatives.

Suppose not; then there exists a sequence $\lambda_n \to \lambda_0$ such that, say, $|g_1(\lambda_n) - g_1(\lambda_0)| \geqq \delta > 0$ for all n. By (A.1) and Lemma 2, $0 \leqq g_j(\lambda_n) \leqq b$. Thus one can extract a subsequence, say λ_n for convenience, such that $g_j(\lambda_n) \to g_j^*$, say, as $n \to \infty$ for $j = 1, \cdots, k$. It follows as in the proof of Lemma 2 that g_j^* is a solution to (A.2) satisfying (A.6), thus by Lemma 3 $g_j^* = g_j(\lambda_0)$. □

Strict monotonicity of $m(\lambda)$.

LEMMA 5. *If* $J_j(u)$, $j = 1, \cdots, k$, *are integrable and at least one is not almost everywhere constant then* $m(\lambda)$ *is strictly increasing in* $\lambda \geqq 0$.

PROOF. In the notation of (A.7) what needs to be shown is that $\iint a h_\lambda$ is strictly increasing or, since h_λ has uniform marginals it is enough to show that $\iint [c\, a(u, v) + a_1(u) + a_2(v)] h_\lambda(u, v)\, du\, dv$ is strictly increasing for some $c > 0$ and some integrable functions a_1 and a_2.

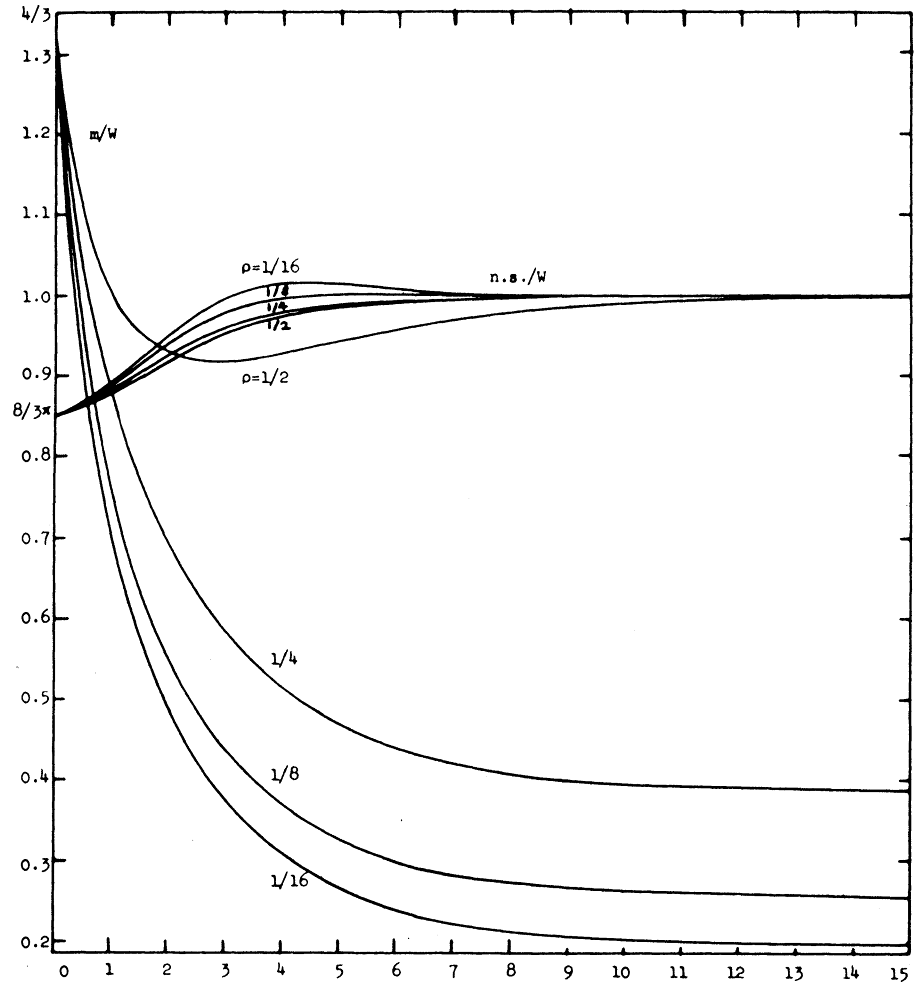

FIG. 4. Bahadur efficiencies for double exponential shift alternatives.

Take any $\lambda_0 < \lambda_1$ and for each fixed v define the density

$$l(u;v,\theta) = h_{\lambda_0}\left[\frac{h_{\lambda_1}}{h_{\lambda_0}}\right]^\theta \div \int h_{\lambda_0}\left[\frac{h_{\lambda_1}}{h_{\lambda_0}}\right]^\theta du$$

where $0 \leqq \theta \leqq 1$. Notice that $l(u;v,1) = h_{\lambda_1}(u,v)$. The above can be rewritten as $l(u;v,\theta) = c(\theta)Q(u)\exp[\theta T(u)]$, where the dependence of c, Q and T on v is suppressed and, by (A.8) and (2.16),

$$T(u) = \log[h_{\lambda_1}/h_{\lambda_0}] = (\lambda_1-\lambda_0)a(u,v)+a_1(u)+a_2(v)$$

with

$$a_1(u) = \log\{\textstyle\sum_i\rho_i\exp[\lambda_0(J_i(u)-s_i(\lambda_0))]/\sum_i\rho_i\exp[\lambda_1(J_i(u)-s_i(\lambda_1))]\},$$

and

$$a_2(v) = \lambda_0 s_j(\lambda_0) - \lambda_1 s_j(\lambda_1), \quad v_{j-1} \leqq v < v_j, \quad j = 1, \cdots, k.$$

Since $T(u)$ is clearly not almost everywhere constant, it follows from well-known properties of exponential families of densities that, for each v,

$$\begin{aligned} &\int[(\lambda_1 - \lambda_0)a + a_1 + a_2]h_{\lambda_0}\, du \\ &\quad = \int T(u)l(u; v, 0)\, du < \int T(u)l(u; v, 1)\, du \\ &\quad = \int[(\lambda_1 - \lambda_0)a + a_1 + a_2]h_{\lambda_1}\, du. \end{aligned}$$

Since a_1 and a_2 are integrable and $\lambda_1 - \lambda_0$ is positive, the lemma is proved. □

LEMMA 6. *$m(\lambda) \to \bar{r}$ as $\lambda \to \infty$.*

PROOF. Recall the definition of $\bar{r}$ as the supremum of $\int \sum \rho_j f_j(u) J_j(u)\, du$ subject to $\sum \rho_j f_j(u) = 1, f_1, \cdots, f_k$ being densities on $0 < u < 1$.

For arbitrary constants $(s_1, \cdots, s_k) = \mathbf{s}$ the modified expression

$$\int \sum \rho_j f_j(u)[J_j(u) - s_j]$$

subject to the weaker constraints $\sum \rho_j f_j(u) = 1$ and $f_j(u) \geqq 0$ is clearly maximized by

$$\text{(A.9)} \qquad \begin{aligned} f_j(u; \mathbf{s}) &= 1, & J_j(u) - s_j &> \max_{i \neq j}(J_i(u) - s_i) \\ &= \lambda_j(u), & &= \\ &= 0, & &< \end{aligned}$$

where $\sum \rho_j \lambda_j(u) = 1$ but otherwise $\lambda_1, \cdots, \lambda_k$ are arbitrary.[14] Clearly one can replace s_j by $s_j - \sum \rho_j s_j$ without changing (A.9) thus one can assume

$$\text{(A.10)} \qquad \sum \rho_j s_j = 0.$$

If it is possible to select $\mathbf{s}$ so that $f_j(u; \mathbf{s})$, $j = 1, \cdots, k$ are densities then $f_j(u; \mathbf{s})$, $j = 1, \cdots, k$, is a solution to the original maximization problem and

$$\text{(A.11)} \qquad \bar{r} = \int_0^1 \sum_{j=1}^k \rho_j f_j(u; \mathbf{s})[J_j(u) - s_j]\, du.$$

Consider $s_1(\lambda), \cdots, s_k(\lambda)$, the solution to (3.10) satisfying (A.10). If, say, $|s_1(\lambda)|$ is unbounded as $\lambda \to \infty$, then there exists a sequence $\lambda_n \to \infty$ such that $s_{j'}(\lambda_n) - s_j(\lambda_n) \to \infty$ for some pair (j, j'). But then for $v_{j-1} \leqq v < v_j$,

$$\begin{aligned} h_{\lambda_n}(u, v) &= \{\sum_{i=1}^k \rho_i \exp[\lambda_n(J_i(u) - J_j(u) - s_i(\lambda_n) + s_j(\lambda_n))]\}^{-1} \\ &\to 0, \qquad \text{as } n \to \infty \end{aligned}$$

which is impossible since $h_\lambda(u, v) \leqq 1/\min(\rho_j)$ and $\int h_\lambda(u, v)\, du = 1$, for every v. Since $|s_j(\lambda)|$ remains bounded there exists a sequence $\lambda_n \to \infty$ such that $s_j(\lambda_n) \to s_j$, say. Clearly for $v_{j-1} \leqq v < v_j$, $h_{\lambda_n}(u, v)$ converges to an expression like the right side

[14] This is a new use of the symbol λ.

of (A.9) and of course $f_j(u; s) = \lim h_{\lambda_n}(u, v)$, $v_{j-1} \leqq v < v_j$, is a density. Since one can write

$$m(\lambda) = \int_0^1 \sum_{j=1}^k \rho_j J_j(u) h_\lambda(u, v_j^*)\, du,$$

where $v_{j-1} < v_j^* < v_j$ and the integrand is dominated by $\sum \rho_j J_j(u)/\min(\rho_j)$, the lemma follows at once. □

REFERENCES

[1] Andrews, F. C. and Truax, D. R. (1964). Locally most powerful rank tests for several sample problems. *Metrika* **8** 16–24.

[2] Bahadur, R. R. (1967). Rates of convergence of estimates and test statistics. *Ann. Math. Statist.* **38** 303–324.

[3] Bhuchongkul, S. (1964). A class of nonparametric tests for independence in bivariate populations. *Ann. Math. Statist.* **35** 138–149.

[4] Chernoff, H. and Savage, I. R. (1958). Asymptotic normality and efficiency of certain nonparametric procedures. *Ann. Math. Statist.* **29** 972–994.

[5] Hoadley, A. B. (1967). On the probability of large deviations of functions of several empirical cdf's. *Ann. Math. Statist.* **38** 360–381.

[6] Hoeffding, W. (1965). Asymptotically optimal tests for multinomial distributions. *Ann. Math. Statist.* **36** 369–401.

[7] Klotz, J. (1965). Alternative efficiencies for signed rank tests. *Ann. Math. Statist.* **36** 1759–1766.

[8] Stone, M. (1967). Extreme tail probabilities of the two-sample Wilcoxon statistic. *Biometrika* **54** 629–640.

[9] Stone, M. (1968). Extreme tail probabilities for sampling without replacement and exact Bahadur efficiency of the two-sample normal scores test. *Biometrika* **55** 371–375.

[10] Waerden, B. L. van der (1957). *Mathematische Statistik*. Page 331. Springer-Verlag, Berlin.

[11] Woodworth, G. (1966). On the asymptotic theory of tests of independence based on bivariate layer ranks. Technical Report No. 75, Department of Statistics, Univ. of Minnesota.

[12] Woodworth, G. (1967). On large deviations of linear rank statistics. Technical Report No. 98, Department of Statistics, Stanford Univ.