The Annals of Statistics
1983, Vol. 11, No. 3, 933-946

A MODIFIED KOLMOGOROV-SMIRNOV TEST SENSITIVE TO TAIL ALTERNATIVES

BY DAVID M. MASON AND JOHN H. SCHUENEMEYER

University of Delaware

It is well known that the Kolmogorov-Smirnov (K-S) test exhibits poor sensitivity to deviations from the hypothesized distribution that occur in the tails. A modified version of the K-S test is introduced that is more sensitive than the K-S test to deviations in the tails. The finite and infinite sample distribution along with the consistency properties of the proposed test are studied. Tables of critical values are provided for two versions of the test (one sensitive to heavy tail alternatives and one sensitive to light tail alternatives) and the finite sample properties of these two versions of the test are investigated.

1. Introduction. Let $X_1, \cdots, X_n$ be independent random variables with common continuous distribution function F and let $X_{1,n} \le \cdots \le X_{n,n}$ denote their order statistics. F_n will denote the right continuous empirical distribution function based on $X_1, \cdots, X_n$. Also let F_0 be any fixed continuous distribution function.

We will write the Kolmogorov-Smirnov (K-S) statistic as

$$K_n = \sup\{n^{1/2} |F_n(x) - F_0(x)| : -\infty < x < \infty\};$$

and the Rényi-type statistics (refer to Rényi, 1953, 1968) as

$$L_{n,1} = \sup\{F_0(x)/F_n(x) : x > X_{1,n}\},$$

$$L_{n,2} = \sup\{F_n(x)/F_0(x) : -\infty < x < \infty\},$$

$$U_{n,1} = \sup\{(1 - F_0(x))/(1 - F_n(x)) : x < X_{n,n}\}, \quad \text{and}$$

$$U_{n,2} = \sup\{(1 - F_n(x))/(1 - F_0(x)) : -\infty < x < \infty\}.$$

(Though these statistics are usually called Rényi-type statistics, $L_{n,1}$ and $U_{n,1}$ were first studied by Chang (1955) and later by Tang (1962); likewise, $L_{n,2}$ and $U_{n,2}$ were first studied by Daniels (1945) and later by Robbins (1954).)

Consider the following hypothesis testing procedure based on the statistics $L_{n,1}$, $L_{n,2}$, $U_{n,1}$, $U_{n,2}$ and K_n for testing

$$H_0 : F = F_0 \quad \text{versus} \quad H_a : F \in \mathscr{F} \quad \text{at level} \quad \alpha,$$

where $\mathscr{F}$ is a specified class of continuous distributions not containing F_0: Reject H_0 if

$$\max\{w_1 L_{n,1}, w_2 L_{n,2}, K_n, w_3 U_{n,1}, w_4 U_{n,2}\} > c,$$

where $w_1, \cdots, w_4$ are predetermined nonnegative weights and $0 < c < \infty$ is a constant (depending on n) chosen so that the probability of rejection is α. (Observe that when $w_1 = w_2 = w_3 = w_4 = 0$ this procedure reduces to the usual K-S test.)

Any version of the above procedure will be called a modified K-S test or M test for short. The object of this paper will be to study the finite and infinite sample properties of various versions of the M test. Most importantly, it will be shown that particular versions of the M test are much more sensitive than the K-S test to deviations from the hypothesized

Received June 1982; revised November 1982.

AMS 1980 *subject classifications.* Primary 62G10; secondary 62E20.

Key words and phrases. Tail alternatives, Kolmogorov-Smirnov test, Rényi tests, weighted empirical distributions.

distribution F_0 that occur in the tails. This will be shown both asymptotically by comparing the behavior of the K-S test to the M test with respect to appropriately chosen "local tail alternatives;" and for finite samples by comparing the power function of two versions of the M test (one sensitive to "heavy tail alternatives" and one sensitive to "light tail alternatives") to the power function of the K-S test calculated for selected families of alternative distributions. Tables of finite and infinite sample critical values are provided for these two versions of the M test.

Versions of the M-test may also be inverted, as is typically done with the usual K-S test, to form confidence bands for F_0. This is discussed in Section 5.

2. The finite sample and asymptotic distribution of the M test under the null hypothesis. In this section we assume that $F = F_0$. Since in this case the M test is distribution free, we will replace $X_1, \cdots, X_n$ by n independent uniform (0, 1) random variables $U_1, \cdots, U_n$; $X_{1,n} \le \cdots \le X_{n,n}$ by the order statistics $U_{1,n} \le \cdots \le U_{n,n}$ of $U_1, \cdots, U_n$; F_n by G_n the empirical distribution based on $U_1, \cdots, U_n$; and F_0 by the uniform (0, 1) distribution. Hence, we will write $\langle L_{n,1}, L_{n,2}, K_n, U_{n,1}, U_{n,2}\rangle$ as

$$\langle \sup\{u/G_n(u) : U_{1,n} < u < 1\}, \quad \sup\{G_n(u)/u : 0 < u < 1\},$$
$$\sup\{n^{1/2} | G_n(u) - u | : 0 < u < 1\}, \quad \sup\{(1-u)/(1-G_n(u)) : 0 < u < U_{n,n}\},$$
$$\sup\{(1-G_n(u))/(1-u) : 0 < u < 1\}\rangle.$$

For any choice of weights $0 \le w_1, \cdots, w_4 \le 1$ and $0 < c < \infty$ the probability of accepting H_0 using the M-test becomes:

$$P(w_1 L_{n,1} \le c,\ w_2 L_{n,2} \le c,\ K_n \le c,\ w_3 U_{n,1} \le c,\ w_4 U_{n,2} \le c). \tag{1}$$

It is routine to show that the probability in (1) can be written equivalently as

$$P(u_i \le U_{i,n} \le v_i : 1 \le i \le n), \tag{2}$$

where u_i and v_i with $u_i \le v_i$ for $i = 1, \cdots, n$ are nondecreasing positive constants determined by $w_1, \cdots, w_4$ and c. This last probability can be calculated by Steck's formula (1971) or Noé's algorithm (1972).

We now obtain the asymptotic distribution of the M test under the null hypothesis. First, it is well known that for any $0 < c < \infty$

$$P(K_n \le c) \to P(\| B \| \le c) = 1 + 2 \textstyle\sum_{k=1}^{\infty} (-1)^k e^{-2k^2c^2}, \tag{3}$$

where B is a Brownian bridge defined on (0, 1) and $\| B \| = \sup\{| B(u)| : 0 < u < 1\}$. (Refer, for instance, to Billingsley, 1968.) Also it can be shown by a simple application of Theorem 7 of Wellner (1977) that for each choice of $1 < a, b < \infty$ that

$$P(L_{n,1} \le a, L_{n,2} \le b) = P(U_{n,1} \le a, U_{n,2} \le b) \to \tag{4}$$

$$P(1/a \le N(t)/t \le b \quad \text{for} \quad E \le t < \infty) \tag{5}$$

where N is a Poisson process with parameter 1 and first jump at E.

We will write

$$G(c) = P(\| B \| \le c)$$

for any choice of $0 < c < \infty$; and

$$H(a, b) = P(1/a \le N(t)/t \le b \quad \text{for} \quad E \le t < \infty)$$

for any choice of $1 < a, b \le \infty$. In particular, when $1 < a < \infty$ and $b = \infty$

$$H(a, \infty) = 1 - \exp(-a) - \textstyle\sum_{k=1}^{\infty} \dfrac{(k-1)^{k-1}}{k!} a^k \exp(-ka)$$

and when $a = \infty$ and $1 < b < \infty$

$$H(\infty, b) = 1 - b^{-1}.$$

(Refer to Pyke, 1959, or Wellner, 1977.) For general values of $1 \le a, b \le \infty$ a recursion formula for $H(a, b)$ is outlined in the Appendix.

In the following theorem and elsewhere in this paper, repeated use will be made of the following fact:

Let $\{k_n\}$ be any sequence of positive integers such that

(K) $$k_n \to \infty \quad \text{and} \quad k_n/n \to 0, \qquad \text{then both}$$

(C1) $$nU_{k_n,n}/k_n \to_P 1, \qquad \text{and}$$

(C2) $$n(1 - U_{n-k_n,n})/k_n \to_P 1.$$

The proof of this fact is elementary. (See page 18 of Balkema and de Haan, 1974.)

THEOREM 1. (*The Asymptotic Distribution of the* M *Test under the Null Hypothesis.*) *Under the null hypothesis, for any choice of* $1 \le a_1, a_2, b_1, b_2 \le \infty$ *and* $0 < c < \infty$

(6) $$P(L_{n,1} \le a_1, L_{n,2} \le b_1, K_n \le c, U_{n,1} \le a_2, U_{n,2} \le b_2) \to H(a_1, b_1)G(c)H(a_2, b_2).$$

PROOF. Let $k_n = [2\ln n] + 1$ and $\ell_n = [(\ln n)^2/2] + 1$. ($[x]$ = greatest integer $\le x$.) Let

$$L_{n,1}(k_n) = \sup\{u/G_n(u) : U_{1,n} < u \le U_{k_n,n}\},$$

$$L_{n,2}(k_n) = \sup\{G_n(u)/u : 0 < u \le U_{k_n,n}\},$$

$$K_n(\ell_n) = \sup\{n^{1/2}|G_n(u) - u| : U_{\ell_n,n} \le u \le U_{n-\ell_n,n}\},$$

$$U_{n,1}(k_n) = \sup\{(1-u)/(1-G_n(u)) : U_{n-k_n,n} \le u < U_{n,n}\}, \qquad \text{and}$$

$$U_{n,2}(k_n) = \sup\{(1-G_n(u))/(1-u) : U_{n-k_n,n} \le u < 1\}.$$

We will show that

(7) $$\langle L_{n,1}, L_{n,2}, K_n, U_{n,1}, U_{n,2}\rangle - \langle L_{n,1}(k_n), L_{n,2}(k_n), K_n(\ell_n), U_{n,1}(k_n), U_{n,2}(k_n)\rangle$$

converges in probability to $\vec{0}$, from which we can conclude from Lemma 1 of Rossberg (1965) (using the fact that $k_n/\ell_n \to 0$) that the random variables $\langle L_{n,1}, L_{n,2}\rangle$, K_n, and $\langle U_{n,1}, U_{n,2}\rangle$ are asymptotically independent and hence from (3), (4) and (5) that the conclusion of the theorem holds.

Observe that $|L_{n,1}(k_n) - L_{n,1}|$ is less than or equal to

(8) $$\sup\{|G_n(u) - u|/G_n(u) : U_{k_n,n} \le u < 1\},$$

and $|L_{n,2}(k_n) - L_{n,2}|$ is less than or equal to

(9) $$\sup\{|G_n(u) - u|/u : U_{k_n,n} \le u < 1\}.$$

Theorem 0 of Wellner (1978) implies that expression (8) converges in probability to zero. The same theorem in combination with (C1) above implies that expression (9) converges in probability to zero. Hence $\langle L_{n,1}, L_{n,2}\rangle - \langle L_{n,1}(k_n), L_{n,2}(k_n)\rangle$ converges in probability to $\langle 0, 0\rangle$. Similarly $\langle U_{n,1}, U_{n,2}\rangle - \langle U_{n,1}(k_n), U_{n,2}(k_n)\rangle$ converges in probability to $\langle 0, 0\rangle$. Now

(10) $$|K_n(\ell_n) - K_n| \le \sup\{n^{1/2}|G_n(u) - u| : u \notin (U_{\ell_n,n}, U_{n-\ell_n,n})\}.$$

It is known for a more general situation (see Shorack, 1979) that

(11) $$\sup\{n^{1/2}|G_n(u) - u| : u \notin ((\ln n)^2/n, 1 - (\ln n)^2/n)\} \to_P 0.$$

Thus by (11) we can conclude as above using (C1) and (C2) that $K_n(\ell_n) - K_n$ converges in probability to zero. □

3. A convenient class of local tail alternatives. Here we introduce a class of "local tail alternatives" that will simplify our study of the consistency properties of the M test with respect to local deviations from the null distribution F_0 that occur in the tails.

For any $\beta > 0$ and $0 \le p \le 1/2$, set

$$Q_{\beta,p}(u) = \begin{cases} 1 - p^{1-\beta}(1-u)^{\beta} & \text{for } 1-p < u < 1 \\ u & \text{for } p \le u \le 1-p \\ p^{1-\beta}u^{\beta} & \text{for } 0 < u < p. \end{cases}$$

Let $F_{\beta,p}$ denote the distribution which has inverse equal to $Q_{\beta,p}$, i.e.,

$$F_{\beta,p}(x) = \begin{cases} 1 - p^{(\beta-1)/\beta}(1-x)^{1/\beta} & 1-p < x < 1 \\ x & p \le x \le 1-p \\ p^{(\beta-1)/\beta}x^{1/\beta} & 0 < x < p. \end{cases}$$

Notice that when $\beta = 1$ each $F_{1,p}$ is equal to the uniform (0, 1) distribution ($U(0, 1)$).

Observe that whenever $\beta > 1$

$$\lim_{u\downarrow 0}(F_{\beta,p}(u) \wedge (1 - F_{\beta,p}(1-u)))/u = \infty, \tag{12}$$

that is, $F_{\beta,p}$ has heavier lower and upper tails than $U(0, 1)$; and whenever $0 < \beta < 1$ the limit in (12) is equal to zero, that is, $F_{\beta,p}$ has lighter upper and lower tails than $U(0, 1)$.

For future reference, we note that the Kolmogorov distance between $F_{\beta,p}$ and the $U(0, 1)$ distribution is

$$\|\Delta\| \equiv \sup_{0\le u\le 1}|F_{\beta,p}(u) - u| = p\beta^{\beta/(\beta-1)}|\beta - 1|.$$

Observe that $\|\Delta\|$ strictly increases from zero to p as a function of β on the interval $[1, \infty)$ and strictly decreases from p to zero on the interval $[0, 1]$.

For any choice of $\beta > 0$ and sequence $\{k_n\}$ satisfying (K), let $\mathscr{F}_{\beta,k_n} = \{F_{\beta,p_n} : n \ge 1\}$, where $p_n = k_n/n$ for each $n \ge 1$. Whenever $\beta > 1$ [respectively $0 < \beta < 1$], $\mathscr{F}_{\beta,k_n}$ will be called a class of local heavy tail alternatives [respectively local light tail alternatives].

It will assist our discussion later on to introduce the following definitions.

Let $\{X_n^{(i)} : i = 1, \cdots, n, n \ge 1\}$ denote a triangular array of random variables. For each $n \ge 1$ let T_n denote a statistic based on the empirical distribution F_n of $X_n^{(1)}, \cdots, X_n^{(n)}$. For any fixed value of $0 < \beta < \infty$ and sequence $\{k_n\}$ satisfying (K)

$$T_n \to_{P_{\beta,n}} \quad \text{and} \quad T_n \to_{d_{\beta,n}}$$

will denote convergence in probability and convergence in distribution respectively assuming that for each $n \ge 1$, $X_n^{(1)}, \cdots, X_n^{(n)}$ are i.i.d. F_{β,p_n}. (When $\beta = 1$, $T_n \to_P$ and $T_n \to_d$ will denote convergence in probability and convergence in distribution respectively.)

DEFINITION 1. A sequence of statistics $\{T_n\}$ will be said to be consistent against local tail alternatives; respectively local heavy tail alternatives; respectively local light tail alternatives if

(D) $T_n \to_d$ to a nondegenerate random variable (say with distribution H) when $\beta = 1$, and

(E) $T_n \to_{P_{\beta,n}} x_\infty$, where $x_\infty = \inf\{x \le \infty : H(x) = 1\}$, for every sequence $\{k_n\}$ satisfying (K) and $0 < \beta < \infty$ such that $\beta \ne 1$; respectively every $1 < \beta < \infty$; respectively every $0 < \beta < 1$.

DEFINITION 2. A sequence of statistics $\{T_n\}$ will be said to be inconsistent against local tail alternatives; respectively local heavy tail alternatives; respectively local light tail alternatives if for every $0 < \beta < \infty$ such that $\beta \ne 1$; respectively $\beta > 1$; respectively every $0 < \beta < 1$ there exists a sequence $\{k_n\}$ satisfying (K) such that T_n does not converge in probability (under $P_{\beta,n}$) to any number greater than or equal to x_∞.

4. Consistency and inconsistency properties of $L_{n,1}$, $L_{n,2}$, $U_{n,1}$, $U_{n,2}$, and K_n against local tail alternatives. Equipped with the definitions given in the previous section, we now establish some important consistency and inconsistency properties of $L_{n,1}$, $L_{n,2}$, $U_{n,1}$, $U_{n,2}$ and K_n against local tail alternatives. These are the following:

PROPERTY 1. $L_{n,2}$ and $U_{n,2}$ are consistent against local heavy tail alternatives, but inconsistent against local light tail alternatives.

PROPERTY 2. $L_{n,1}$ and $U_{n,1}$ are consistent against local light tail alternatives, but inconsistent against local heavy tail alternatives.

PROPERTY 3. K_n is inconsistent against local tail alternatives (both heavy and light).

The following propositions establish these properties. First observe that when $F = F_{\beta,p}$ for some choice of $0 < p \leq ½$ and $\beta > 0$.

$$L_{n,1} =_d \sup\{Q_{\beta,p}(u)/G_n(u) : U_{1,n} < u < 1\}, \tag{13}$$

$$L_{n,2} =_d \sup\{G_n(u)/Q_{\beta,p}(u) : 0 < u < 1\}, \tag{14}$$

$$U_{n,1} =_d \sup\{(1 - Q_{\beta,p}(u))/(1 - G_n(u)) : 0 < u < U_{n,n}\}, \tag{15}$$

$$U_{n,2} =_d \sup\{(1 - G_n(u))/(1 - Q_{\beta,p}(u)) : 0 < u < 1\}, \quad \text{and} \tag{16}$$

$$K_n =_d \sup\{n^{1/2}\,|\,G_n(u) - Q_{\beta,p}(u)| : 0 < u < 1\}. \tag{17}$$

To avoid additional notation, whenever $F = F_{\beta,p}$, $L_{n,1}$, $L_{n,2}$, $U_{n,1}$, $U_{n,2}$ and K_n will denote the right side of (13)–(17).

PROPOSITION 1.
(a) *For every $\beta > 1$ and sequence $\{k_n\}$ satisfying* (K), $L_{n,2} \to_{P_{\beta,n}} \infty$; *whereas*
(b) *for every $0 < \beta < 1$, $L_{n,2} \to_{P_{\beta,n}} 1$ with $k_n = (\ln n)^2$. The same statements hold for $U_{n,2}$.*

PROOF. First consider (a). Choose any $\beta > 1$ and sequence $\{k_n\}$ satisfying (K). Let $\{k'_n\}$ be any sequence of positive integers such that $k'_n \to \infty$ and $k'_n/k_n \to 0$. Observe that

$$(k'_n/k_n)^{\beta-1} L_{n,2} \geq (k'_n/k_n)^{\beta-1}[k'_n/(nQ_{\beta,p_n}(U_{k'_n,n}))] \quad \text{a.s.} \tag{18}$$

(Recall $p_n = k_n/n$.)
Since $nU_{k'_n,n}/k'_n \to 1$, we see that with arbitrarily high probability and all n sufficiently large that the right side of inequality (18) is greater than or equal to

$$(k'_n/(nU_{k'_n,n}))^{\beta}, \tag{19}$$

which converges in probability to 1. Since $\beta > 1$, we see from (18) that (a) is true.
Now to prove part (b). Choose any $0 < \beta < 1$. Notice that

$$L_{n,2} = \max\{I_{1n}(\beta), I_{2n}(\beta)\}$$

where

$$I_{1n}(\beta) = \sup\{G_n(u)/Q_{\beta,p_n}(u) : 0 < u \leq p_n\}, \quad \text{and}$$

$$I_{2n}(\beta) = \sup\{G_n(u)/Q_{\beta,p_n}(u) : p_n \leq u < 1\}.$$

First we claim that

$$I_{1n}(\beta) \to_P 1. \tag{20}$$

Observe that

$$|I_{1n}(\beta) - 1| \leq \sup\{n^{1-\beta}\,|\,G_n(u) - u|/((k_n)^{1-\beta}u^{\beta}) : 0 < u \leq p_n\} \equiv S_{n,\beta}.$$

LEMMA 1. *For every* $0 \le \beta < 1$ *there exists a sequence of positive constants* $\{k_n\}$ *satisfying condition* (K) *such that*

$$S_{n,\beta} \to_P 0.$$

(*The choice* $k_n = (\ln n)^2$ *works for all* $0 \le \beta < 1$.)

PROOF. First assume that $0 \le \beta < \frac{1}{2}$ and $\{k_n\}$ is any sequence of positive constants satisfying (K). There exists a universal constant $K > 0$ such that for every $\varepsilon > 0$

$$P(S_{n,\beta} > \varepsilon) \le Kn^{1-2\beta}k_n^{-2+2\beta}\varepsilon^{-2}\int_0^{p_n} u^{-2\beta}\,du = \varepsilon^{-2}K(1-2\beta)^{-1}k_n^{-1}.$$

(See, for instance, Pyke and Shorack, 1968.) Now assume that $\frac{1}{2} < \beta < 1$. Observe that

$$S_{n,\beta} \le k_n^{-1+\beta}\sup_{0<u<1} n^{1-\beta}\,|\,G_n(u) - u\,|/u^{\beta}, \tag{21}$$

but by Theorem 1 of Mason (1983) the right side of inequality (21) converges in probability to zero for any sequence $\{k_n\}$ satisfying (K).

Finally assume that $\beta = \frac{1}{2}$. It is immediate from Corollary 2 of Jaeschke (1979) that for any sequence of positive constants $\{k_n\}$ such that

$$\ln k_n = o(\ln n) \tag{L}$$

that

$$\sup\{n^{1/2}\,|\,G_n(u) - u\,|\,u^{-1/2}: 0 < u \le p_n\} = 0_p((\ln\ln n)^{1/2}). \tag{22}$$

Hence by (22) if, in addition to (L), k_n satisfies

$$k_n^{-1} = o(1/(\ln\ln n)), \tag{M}$$

$S_{n,1/2}$ converges in probability to zero. Observe that the choice $k_n = (\ln n)^2$ works for all values of $0 \le \beta < 1$. □

Hence we have shown (20). Notice that

$$I_{2n}(\beta) \le (1 - p_n)^{-1}\sup\{G_n(u)/u : p_n < u < 1\}, \tag{23}$$

but Theorem 0 of Wellner (1978) implies that the right side of expression (23) converges in probability to 1. This completes the proof of Proposition 1. □

Since $L_{n,2} \to_d H(\infty, \cdot)$, when $\beta = 1$, we see that Proposition 1 establishes Property 1. (Recall (D) and (E) above and Definition 2.)

PROPOSITION 2.
(a) *For every* $0 < \beta < 1$ *and sequence* $\{k_n\}$ *satisfying* (K), $L_{n,1} \to_{P_{\beta,n}} \infty$; *whereas*

(b) *for every* $1 < \beta < \infty$, $L_{n,1} \to_{P_{\beta,n}} 1$ *with* $k_n = (\ln n)^2$. *The same statements hold for* $U_{n,1}$.

PROOF. First consider (a). Choose any $0 < \beta < 1$ and sequence $\{k_n\}$ satisfying (K). Let $\{k'_n\}$ be any sequence of positive integers such that $k'_n \to \infty$ and $k'_n/k_n \to 0$. We see as in the proof of Proposition 1 that with arbitrarily high probability and all n sufficiently large that

$$(k'_n/k_n)^{1-\beta}L_{n,1} \ge (nU_{k'_n,n}/k'_n)^{\beta}. \tag{24}$$

Since $0 < \beta < 1$ and the right side of (24) converges in probability to 1, the proof of part (a) is complete.

Now consider part (b). Choose any $\beta > 1$. Let

$$J_n(\beta) = \sup\{n^{\beta-1}u^{\beta}k_n^{1-\beta}G_n(u)\colon U_{1,n} < u \le p_n\}.$$

($J_n(\beta) \equiv 0$ if $U_{1,n} > p_n$.)

We claim that

$$J_n(\beta) \to_P 1 \tag{25}$$

for the sequence $k_n = (\ln n)^2$. We will require the following lemma.

LEMMA 2. *For every* $\beta > 1$

$$T_{n,\beta} \equiv \sup\{n^{1-1/\beta}k_n^{1/\beta-1}|G_n(u) - u|/(G_n(u))^{1/\beta}\colon U_{1,n} < u \le p_n\} \to_P 0$$

with $k_n = (\ln n)^2$.

PROOF. Choose any $\beta > 1$.

$$T_{n,\beta} \le S_{n,1/\beta}(\sup\{u/G_n(u)\colon U_{1,n} < u < 1\})^{1/\beta}. \tag{26}$$

Hence, by Lemma 1 and (4) and (5) above, the right side of (26) converges in probability to zero for the sequence $k_n = (\ln n)^2$. □

By Lemma 2 for every $\beta > 1$

$$\sup_{U_{1,n}<u\le p_n} n^{\beta-1}k_n^{1-\beta}(G_n(u))^{\beta-1} - J_n(\beta) \to_P 0 \tag{27}$$

for the sequence $k_n = (\ln n)^2$. Using (C1) above it is easy to show that the first term in (27) converges in probability to 1. Hence we have shown (25).

The rest of the proof of Proposition 2 proceeds very much like the proof of Proposition 1, therefore we omit the details. □

Since $L_{n,1} \to_d H(\cdot, \infty)$, when $\beta = 1$, we see that Proposition 2 establishes Property 2.

Let Q denote the class of nonnegative continuous functions q defined on (0, 1) such that

$$q \uparrow \text{ on } (0, \tfrac{1}{2}) \text{ and is symmetric about } \tfrac{1}{2}, \quad \text{and} \tag{q1}$$

$$\lim_{u\downarrow 0}[u \ln \ln(1/u)]^{-1/2}q(u) = \infty. \tag{q2}$$

It is well known that whenever $q \in Q$

$$K_n(q) \equiv n^{1/2}\sup\{|G_n(u) - u|/q(u)\colon 0 < u < 1\} \to_d \|B/q\|.$$

(See O'Reilly, 1974, and Shorack, 1979.)

To establish Property 3, we will actually prove the more general property:

PROPERTY 4. For every $q \in Q$, $K_n(q)$ is inconsistent against local tail alternatives.

PROPOSITION 3. *For every* $0 < \beta < \infty$ *and* $q \in Q$, $K_n(q) \to_{d_{\beta,n}} \|B/q\|$, *where* $k_n = \ln \ln \ln n$ *for large* n.

PROOF. Choose any $0 < \beta < \infty$. Since by Shorack (1979)

$$\sup\{n^{1/2}|G_n(u) - Q_{\beta,p_n}(u)|/q(Q_{\beta,p_n}(u))\colon p_n < u < 1 - p_n\}$$
$$= \sup\{n^{1/2}|G_n(u) - u|/q(u)\colon p_n < u < 1 - p_n\} \to_d \|B/q\|,$$

it is sufficient to show that for each $0 < \beta < \infty$ both

$$\Delta_n(\beta) = \sup\{n^{1/2}|G_n(u) - Q_{\beta,p_n}(u)|/q(Q_{\beta,p_n}(u))\colon 0 < u \le p_n\}$$

and the corresponding upper tail term converge in probability to zero. Now $q \in Q$ implies that for all u sufficiently small

$$q(u) \geq (u \ln \ln (1/u))^{1/2}. \tag{28}$$

Hence by (28) for all n sufficiently large

$$\begin{aligned} q(Q_{\beta,p_n}(u)) &= q((k_n/n)^{1-\beta}u^{\beta}) \geq (k_n/n)^{(1-\beta)/2}u^{\beta/2}(\ln \ln (n/k_n))^{1/2} \\ &\equiv (k_n/n)^{(1-\beta)/2}u^{\beta/2}b_n \end{aligned}$$

for every $0 < u \leq p_n$. Therefore it is enough to show that

$$\begin{aligned} &\sup_{0<u\leq p_n} n^{1/2}| G_n(u) - p_n^{1-\beta}u^{\beta}|/(u^{\beta/2}b_n(k_n/n)^{(1-\beta)/2}) \\ &\leq \sup_{0<u\leq p_n} n^{1-\beta/2}G_n(u)/(u^{\beta/2}b_n k_n^{(1-\beta)/2}) + k_n^{1/2}/b_n \\ &\equiv I_n(\beta) + k_n^{1/2}/b_n \to_p 0. \end{aligned}$$

The choice of $\{k_n\}$ implies that $k_n^{1/2}/b_n \to 0$. To show that

$$I_n(\beta) \to_P 0, \tag{29}$$

we require the following elementary lemma.

LEMMA 3. *Let $\beta > 0$ and $\{k_n\}$ be any sequence satisfying* (K). *Then for every $x > 0$*

$$P_n(x) \equiv P(\sup\{n^{1-\beta}G_n(u)/u^{\beta}: 0 < u \leq p_n\} > x) \leq \left(\frac{2[k_n] + 1}{x}\right)^{1/\beta} + o(1),$$

where the $o(1)$ term is dependent only on the sequence $\{k_n\}$.

PROOF. Choose $x > 0$.

$$\begin{aligned} P_n(x) &\leq P(\sup\{n^{1-\beta}G_n(u)/u^{\beta}: 0 < u \leq U_{2[k_n]+1,n}\} > x) \\ &\quad + P(U_{2[k_n]+1,n} < p_n) \equiv P_{1,n}(x) + P_{2,n}(x). \end{aligned}$$

But

$$P_{1,n}(x) \leq P((2[k_n] + 1)n^{-\beta}U_{1,n}^{-\beta} > x) \leq \left(\frac{2[k_n] + 1}{x}\right)^{1/\beta}.$$

(C1) above completes the proof.

(A more refined version of this inequality for ½ < $\beta \leq 1$ is given in Mason, 1981.)

Choose any $\varepsilon > 0$. By Lemma 3

$$P(I_n(\beta) > \varepsilon) \leq \left(\frac{2[k_n] + 1}{\varepsilon b_n k_n^{1/2-\beta/2}}\right)^{2/\beta} + o(1).$$

We see by the choice of $\{k_n\}$ that the proof of (29) is complete. □

We should remark here that the K-S test will be consistent against subclasses of local tail alternatives for which (i) $n^{1/2}p_n = k_n n^{-1/2} \to \infty$ and will have non-degenerate asymptotic power when (ii) $p_n = dn^{-1/2}$ for some $d > 0$. To see this, observe that for any positive β not equal to 1 the Kolmogorov distance between F_{β,p_n} and the $U(0, 1)$ distribution has the property that $n^{1/2}\|\Delta\| \to \infty$ if (i) holds and $n^{1/2}\|\Delta\| \to d\beta^{\beta/(1-\beta)}|\beta - 1|$ if (ii) holds.

This same format can be carried out to study the behavior of these statistics against local deviations that occur in the middle of the distribution. Let $\{\ell_n\}$ be any sequence of finite constants such that $|\ell_n| \to \infty$ but $\ell_n/n^{1/2} \to 0$ and choose $0 < \alpha <$ ½.

Let

$$Q^*_{\alpha,\ell_n}(u) = \begin{cases} u \quad \text{for} \quad 1-\alpha < u < 1 \\ \text{piecewise linear from} \quad \langle \alpha, \alpha \rangle \quad \text{to} \quad \langle 1/2, 1/2 + \ell_n/\sqrt{n} \rangle \\ \text{and} \quad \langle 1/2, 1/2 + \ell_n/\sqrt{n} \rangle \quad \text{to} \quad \langle 1-\alpha, 1-\alpha \rangle \quad \text{for} \quad \alpha \le u \le 1-\alpha \\ u \quad \text{for} \quad 0 < u < \alpha. \end{cases}$$

Let G_{α,ℓ_n} denote the distribution with inverse Q^*_{α,ℓ_n}. Any class $G_{\alpha,\ell_n} = \{G_{\alpha,\ell_n}: n \ge 1\}$ will be called a class of local middle alternatives. It can be shown that $L_{n,1}$, $L_{n,2}$, $U_{n,1}$, and $U_{n,2}$ are all inconsistent against local middle alternatives, whereas K_n is consistent against local middle alternatives. (The details are elementary and are left to the reader.) Hence the conclusion of this section is that any M test for which $w_i > 0$ for $i = 1, \cdots, 4$ is consistent against both local tail and middle alternatives.

5. Two versions of the M Test. In this section we examine the finite sample performance of two versions of the M test, one sensitive to light tail alternatives and one sensitive to heavy tail alternatives. These are statistics of the following form:

$$L_n = \max\{wL_{n,1}, K_n, wU_{n,1}\}, \text{ and } H_n = \max\{wL_{n,2}, K_n, wU_{n,2}\},$$

where $w > 0$ is a positive weight to be specified later.

The asymptotic results of Section 4 indicate that L_n should be sensitive to light tail and middle alternatives, but fairly insensitive to heavy tail alternatives; and conversely, H_n should be sensitive to heavy tail and middle alternatives, but fairly insensitive to light tail alternatives. This will be confirmed by the numerical evidence that follows.

One problem that arises in the practical implementation of these tests is in the determination of criteria for the selection of the weight w. For our two examples we have chosen the following:

For a given significance level α, we select that weight w for the L_n test such that

$$P(wL_{n,1} \le x_{\alpha,n}) = P(wU_{n,1} \le x_{\alpha,n}) \to (1-\alpha)^{1/3}, \quad \text{and}$$
$$P(K_n \le x_{\alpha,n}) \to (1-\alpha)^{1/3},$$

where $x_{\alpha,n}$ is the sample size n α-critical value for the L_n test. Similarly, we select that weight w for the H_n test such that

$$P(wL_{n,2} \le y_{\alpha,n}) = P(wU_{n,2} \le y_{\alpha,n}) \to (1-\alpha)^{1/3}, \quad \text{and}$$
$$P(K_n \le y_{\alpha,n}) \to (1-\alpha)^{1/3},$$

where $y_{\alpha,n}$ is the sample size n α-critical value for the H_n test. These criteria say that asymptotically these two versions of the M test assign equal probability of rejecting H_0 when it is true due to deviations that occur in the lower, middle, and upper sample quantiles.

Tables of weights and finite sample critical values for these two versions of the M test are provided. The computations were performed using Noé's algorithm on a DEC-10 in double precision. The tabled numbers are accurate to the significant digits displayed.

The power functions of the K-S test, the L_n test, and the H_n test for sample sizes $n = 25$ and $n = 50$ using the $\alpha = .10$ significance level are plotted in Figures I, II, III and IV with respect to the alternatives $F_{\beta,p}$ for $0 < \beta < \infty$ and the choices of $p = 1/4$ and $p = 1/2$; Figures I and II are plots of the power functions for $p = 1/4$ and Figures III and IV are plots of the power functions for the $p = 1/2$. Values of $0 < \beta < 1$ correspond to light tail alternatives and values of $\beta > 1$ to heavy tail alternatives. (The power functions were computed directly using Noé's algorithm.)

We see that the numerical evidence indicates that the L_n test is much more sensitive to light tail alternatives than the K-S test alone, but less sensitive to heavy tail alternatives;

TABLE I
Weights and critical values for the L_n test

Sample size	α .10 w .2738	α .05 w .2559	α .01 w .2267
5	1.139	1.259	1.495
10	1.185	1.294	1.546
15	1.253	1.353	1.567
20	1.293	1.396	1.608
25	1.317	1.423	1.638
30	1.333	1.441	1.659
35	1.345	1.455	1.675
40	1.354	1.465	1.688
45	1.361	1.473	1.697
50	1.367	1.479	1.705
∞	1.425	1.544	1.788

TABLE II
Weights and critical values for the H_n test

Sample size	α .10 w .0491	α .05 w .0261	α .01 w .006
5	1.336	1.442	1.642
10	1.368	1.480	1.704
15	1.380	1.495	1.726
20	1.387	1.503	1.737
25	1.392	1.509	1.744
30	1.395	1.512	1.749
35	1.398	1.515	1.753
40	1.400	1.517	1.756
45	1.402	1.519	1.758
50	1.403	1.521	1.760
∞	1.425	1.544	1.788

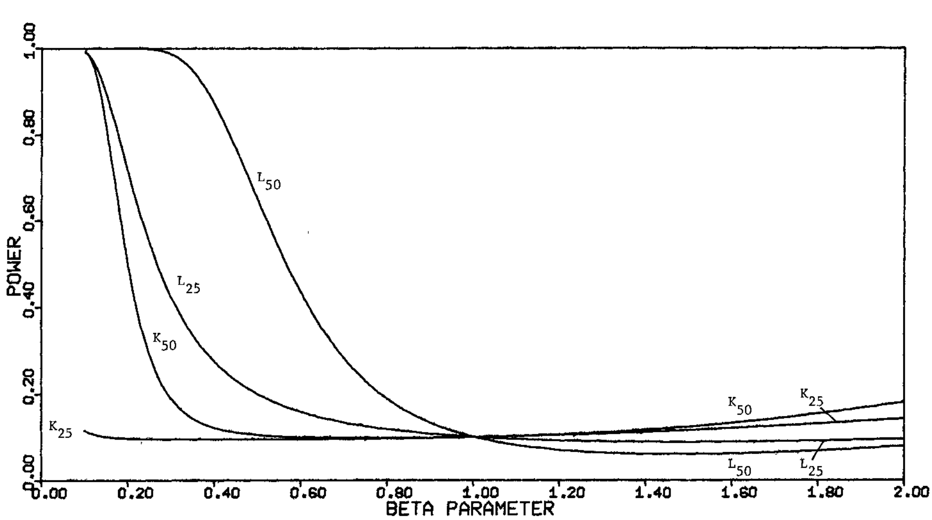

FIG. I. *Power functions of L_n and K_n for $p = 1/4$.*

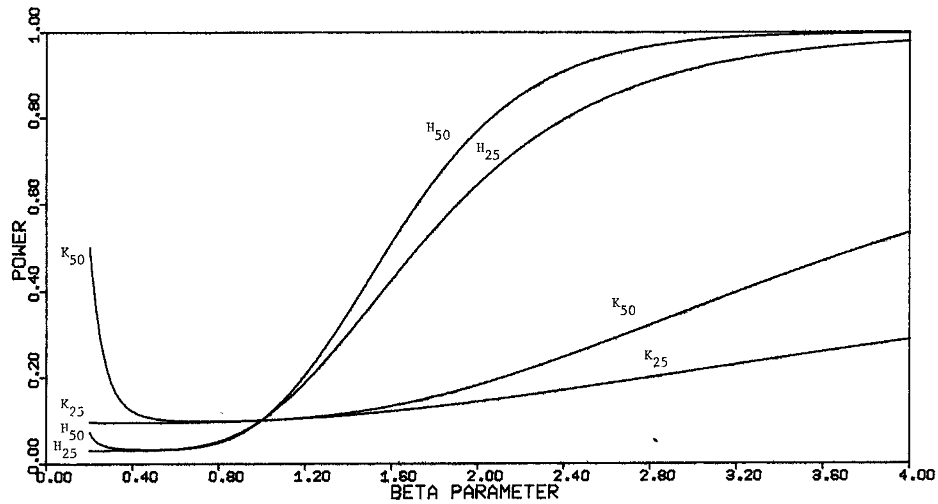

FIG. II. *Power functions of* H_n *and* K_n *for* $p = 1/4$.

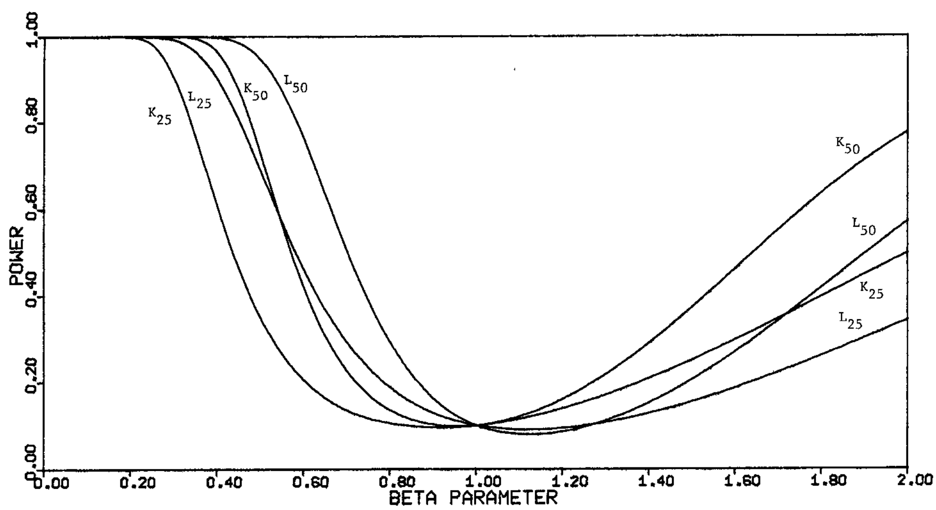

FIG. III. *Power functions of* L_n *and* K_n *for* $p = 1/2$.

whereas the opposite conclusions are true for H_n. This is in agreement with the asymptotic power studies in Section 4. In practical data analysis we would advise using all three tests.

A much more thorough finite sample study of the M test, where positive weights are applied to each component of the statistic will appear elsewhere. We chose these two examples to demonstrate the power of the M test in two extreme cases.

The statistics L_n and H_n may be inverted to form confidence bands for F_0. These are constructed as follows:

For a given critical value $x_{\alpha,n}$ of L_n, $L_n \le x_{\alpha,n}$ if and only if $A_n(x) \le F_0(x) \le B_n(x)$ for all x, where

$$A_n(x) = \max\{F_n(x) - n^{-1/2}x_{\alpha,n}, 1 - w^{-1}(1 - F_n(x))x_{\alpha,n}, 0\}, \qquad \text{and}$$

$$B_n(x) = \min\{F_n(x) + n^{-1/2}x_{\alpha,n}, w^{-1}F_n(x)x_{\alpha,n}, 1\};$$

and for a given critical value $y_{\alpha,n}$ of H_n, $H_n \le y_{\alpha,n}$ if and only if $C_n(x) \le F_0(x) \le D_n(x)$ for

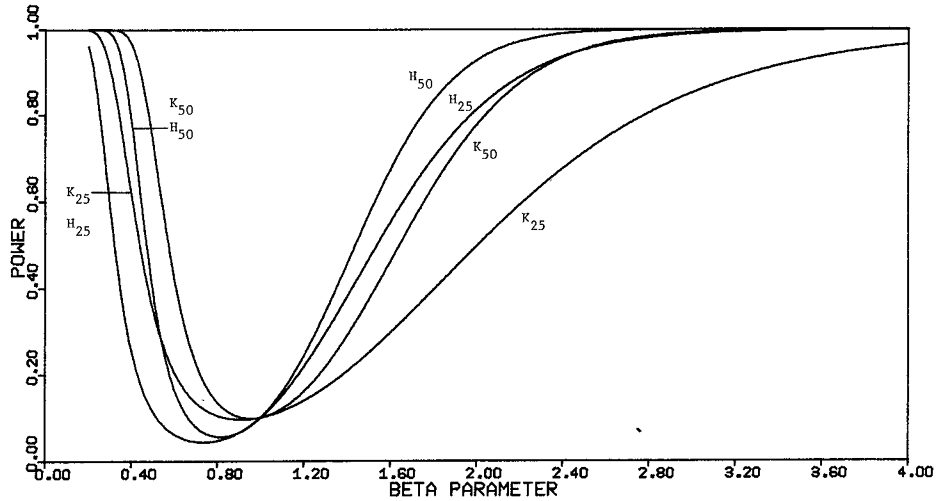

FIG. IV. *Power functions of H_n and K_n for $p = ½$.*

all x, where

$$C_n(x) = \max\{F_n(x) - n^{-1/2}y_{\alpha,n}, wF_n(x)/y_{\alpha,n}, 0\}, \quad \text{and}$$

$$D_n(x) = \min\{F_n(x) + n^{-1/2}y_{\alpha,n}, 1 - w(1 - F_n(x))y_{\alpha,n}^{-1}, 1\}.$$

The confidence bands based on H_n are closely related to the one-sided confidence bands based on $L_{n,2}$ considered by Robbins (1954). His bands are constructed in the obvious way utilizing the fact that $P(L_{n,2} \leq \lambda) = 1 - \lambda^{-1}$ for $\lambda > 1$.

It is easy to show that the confidence bands just described give, for large enough n, narrower bands for the upper and lower tails of F_0 than those that can be obtained by the K-S test alone; but with a sacrifice of slightly wider bands for the middle of F_0. A quick way to see this is to note that the critical values of L_n and H_n are always greater than or equal to the corresponding critical values of the K-S test.

6. Some concluding remarks. Any version of the M test can be written as

$$\sup\{n^{1/2}\Psi_n(F_0(x), F_n(x))\,|F_n(x) - F_0(x)|: -\infty < x < \infty\}$$

where Ψ_n is a positive weight function dependent on n and the choice of w_1, w_2, w_3, and w_4. The asymptotic consistency and inconsistency properties established for the individual components of the M test in Section 4 are in agreement with the remark on page 118 of Eicker (1979) regarding the possible futility of finding a uniformly good *fixed* weight function that will make a weighted K-S test sensitive to local deviations that may occur at any place in the support of the hypothesized distribution. We have not found a uniformly good weight function, but we have shown theoretically that there do exist weight functions *dependent on n* that make particular weighted K-S tests consistent with respect to local deviations that may occur in the tails or the middle of the distribution.

Révész (1982) has recently considered an approach to a modified K-S test sensitive to deviations in the tails that is very similar to ours. His statistic can be written as

$$R_n = \max\{wJ_n, K_n\},$$

where J_n is the Eicker-Jaeschke statistic. (See Eicker, 1979, and Jaeschke, 1979, for details.) He proves that J_n is consistent with respect to a particular class of local tail alternatives, whereas the K-S test is not. The M test is easily shown to be consistent against his class of local tail alternatives, but a slight extension of Proposition 3 shows that J_n is inconsistent against local tail alternatives in our sense. The Révész test suffers from the same problem that the M test does, that is, the determination of reasonable criteria for the choice of the weight w. For the R_n test this problem is compounded by the fact that the finite sample distribution of J_n converges very slowly to its asymptotic distribution. Refer to page 108 of Jaeschke (1979) or compare the asymptotic distribution with the tabled values given in Noé (1972).

7. Appendix. Here we outline a recursion formula for $H(a, b)$. Choose $1 < a, b < \infty$, and set $s_i = ai$ and $t_i = b^{-1}i$ for $i \geq 1$. Let $r = \max\{i: t_i \leq s_1\}$.

First notice that if $N(s_1) = 0, 1$ or $r+1, r+2, \cdots$, then with probability 1 $N(t)$ crosses the line $a^{-1}t$ or bt for some $t \geq E$. Also notice, $H(a, b) = 0$ if $r = 1$. Thus assuming that $r \geq 2$

$$H(a, b) = \textstyle\sum_{i=2}^{r} P(N(s_1) = i, \quad \text{and} \quad a^{-1}t \leq N(t) \leq bt \quad \text{for} \quad t \geq E),$$

but by independence the right side of the above expression is equal to

$$\textstyle\sum_{i=2}^{r} A_i B_i,$$

where

$$A_i = P(N(s_1) = i, \quad \text{and} \quad a^{-1}t \leq N(t) \leq bt \quad \text{for} \quad E \leq t \leq s_1), \quad \text{and}$$

$$\begin{aligned} B_i &= P(N(s_1) = i, \quad \text{and} \quad a^{-1}t - i \leq N(t) - i \leq bt - i \quad \text{for} \quad t \geq s_1) \\ &= P(a^{-1}u + (a^{-1}s_1 - i) \leq N(u) \leq bu + (bs_1 - i) \quad \text{for} \quad u \geq 0). \end{aligned}$$

Since $a^{-1}s_1 - i \leq 0$ and $bs_1 - i \geq 0$ for $2 \leq i \leq r$ a recursion formula for B_i is available on pages 445–447 of Durbin (1971), which for the sake of brevity is not repeated here.

In order to derive an expression for the A_i we must consider two cases.

CASE I. $t_r < s_1$ and $r \geq 2$. In this case for $2 \leq i \leq r$

$$A_i = \textstyle\sum_{(r)}^{(i)} P(N(t_1) = 0, N(t_2) - N(t_1) = k_1, \cdots, N(t_r) - N(t_{r-1}) = k_{r-1}, \\ N(s_1) - N(t_r) = k_r),$$

where the notation $\sum_{(r)}^{(i)}$ denotes summation over all r-tuples of nonnegative integers $(k_1, \cdots, k_r)$ such that

$$\textstyle\sum_{j=1}^{r} k_j = i \quad \text{and} \quad \sum_{j=1}^{\ell} k_j \leq \ell \quad \text{for each} \quad \ell = 1, \cdots, r.$$

Hence

$$A_i = \sum_{(r)}^{(i)} \frac{b^{-i}(ab - r)^{k_r} e^{-a}}{k_1! \cdots k_r!}.$$

CASE II. $t_r = s_1$ and $r \geq 2$. First notice that if $r = 2$, $H(a, b) = 0$, so we assume that $r \geq 3$. In this case

$$A_i = \sum_{(r-1)}^{(i)} \frac{b^{-i} e^{-a}}{k_1! \cdots k_{r-1}!}$$

for $i = 1, \cdots, r-1$ and $A_r = 0$ since $t_r = s_1$.

Acknowledgment. The authors thank J. Eastham for his computational assistance in the early stages of this investigation, and the referee for many helpful comments. In particular, the remark immediately following the proof of Proposition 3 and the suggestion that the M test can be inverted to form confidence bands for F_0 are due to the referee.

REFERENCES

BALKEMA, A. and DE HAAN, L. (1974). Limit laws for order statistics. Colloquia Mathematica Societatis János Bolyai. *Limit Theorems of Probability* (P. Révész, Ed.) Keszthely, Hungary; North-Holland, Amsterdam, London, 17–22.

BILLINGSLEY, P. (1968). *Convergence of Probability Measures.* Wiley, New York.

CHANG, LI-CHIEN (1964). On the ratio of the empirical distribution function to the theoretical distribution function. *Selected Trans. Math. Statist. and Probab. Amer. Math. Soc.* 4 17–38 (Translated from *Acta. Math. Sinica* 5 (1955) 341–368.)

DANIELS, H. E. (1945). The statistical theory of the strength of bundles of threads. *I. Proc. Roy. Soc. London Ser. A.* **183** 405–435.

DURBIN, J. (1971). Boundary-crossing probabilities for Brownian motion and Poisson processes and techniques for computing the power of the Kolmogorov-Smirnov test. *J. Appl. Probab.* **8** 431–453.

EICKER, F. (1979). The asymptotic distribution of standardized empirical processes. *Ann. Statist.* **7** 116–138.

JAESCHKE, D. (1979). The asymptotic distribution of the supremum of the standardized empirical function on subintervals. *Ann. Statist.* **7** 108–115.

MASON, D. M. (1981). Bounds for weighted empirical distribution functions. *Ann. Probab.* **6** 881–884.

MASON, D. M. (1983). The asymptotic distribution of weighted empirical distribution functions. *Stochastic Process. Appl.* **15** 99–109.

NOÉ, M. (1972). The calculation of distributions of two-sided Kolmogorov-Smirnov type statistics. *Ann. Math. Statist.* **43** 58–64.

O'REILLY, N. (1974). On the weak convergence of empirical processes in sup-norm metrics. *Ann. Probab.* **2** 642–651.

PYKE, R. (1959). The supremum and infinum of the Poisson process. *Ann. Math. Statist.* **30** 568–579.

PYKE, R. and SHORACK, G. R. (1968). Weak convergence of a two sample empirical process and a new approach to Chernoff-Savage theorems. *Ann. Math. Statist.* **39** 755–771.

RÉNYI, A. (1953). On the theory of order statistics. *Acta Math. Acad. Sci. Hung.* **4** 191–231.

RÉNYI, A. (1973). On a group of problems in the theory of order statistics. *Selected Transl. in Math. Statist. and Probab. Amer. Math. Soc.* **13** 289–298. (Translated from *Magyar Tud. Akad. Mat. Fiz. Oszt. Közl.* **18** (1968) 23–30.)

RÉVÉSZ, P. (1982). A joint study of the Kolmogorov-Smirnov and the Eicker-Jaeschke statistics. *Statist. and Decisions,* **1** 57–65.

ROBBINS, H. (1954). A one-sided confidence interval for an unknown distribution function. *Ann. Math. Statist.* **25** 409 No. 2.

ROSSBERG, H. J. (1965). Die asymptotische Unabhängigkeit der kleinsten and grössten Werte einer Stichprobe von Stichprobenmittel. *Math. Nachr.* **28** 305–318.

SHORACK, G. R. (1979). Weak convergence of empirical and quantile processes in sup-norm metrics via KMT-construction. *Stochastic Process. Appl.* **9** 95–98.

STECK, G. (1971). Rectangular probabilities for uniform order statistics and the probability that the empirical distribution function lies between two distribution functions. *Ann. Math. Statist.* **42** 1–11.

TANG, S. C. (1962). Some theorems on the ratio of empirical distribution to the theoretical distribution. *Pacific J. Math.* **12** 1107–1114.

WELLNER, J. A. (1977). Distributions related to linear bounds for the empirical distribution function. *Ann. Statist.* **5** 1003–1016.

WELLNER, J. A. (1978). Limit theorems for the ratio of the empirical distribution function to the true distribution function. *Z. Wahrsch. verw. Gebiete* **45** 73–88.

DEPARTMENT OF MATHEMATICAL SCIENCES
COLLEGE OF ARTS AND SCIENCES
501 EWING HALL
UNIVERSITY OF DELAWARE
NEWARK, DELAWARE 19711