The Annals of Statistics
1976, Vol. 4, No. 5, 924-935

POWER BOUNDS FOR A SMIRNOV STATISTIC IN TESTING THE HYPOTHESIS OF SYMMETRY

BY HIRA LAL KOUL AND R. G. STAUDTE, JR.

Michigan State University and La Trobe University

Lower and upper bounds on the power of a Smirnov test for symmetry $H_0: \bar{F} = F$ versus $H_1: \bar{F} \geqq F$, $\sup_x [\bar{F}(x) - F(x)] = \Delta > 0$ are obtained exactly or estimated for selected values of sample size N, level α, and asymmetry Δ. Furthermore the asymptotic power of the test as $N^{\frac{1}{2}}\Delta_N \to c$ is shown to be bounded by $\Phi(c - k_\alpha)$ and 1 if $c \geqq k_\alpha$ and by α and $2\Phi(c - k_\alpha)$ if $c < k_\alpha$, where k_α is the critical point. These bounds compare favorably in some respects with those of the Wilcoxon and other monotone rank tests studied in "Power bounds and asymptotic minimax results for one-sample rank tests," *Ann. Math. Statist.* **42** 12–35.

1. Introduction. Let Ω be the class of continuous cdf's and for each F in Ω let $\bar{F}(x) \equiv 1 - F(-x)$. On the basis of random sample $X_1, \cdots, X_N$ from F in Ω, test the hypothesis of symmetry $H_0: F = \bar{F}$ against alternatives $H_1: \bar{F}(x) \geqq F(x)$ for all x, $\bar{F} \neq F$. Since the Kolmogorov distance of F from the symmetric class $\{F = \bar{F}\}$ is $\frac{1}{2}\sup_x |\bar{F}(x) - F(x)|$, it is natural to measure the asymmetry of F in H_1 by $\sup_x [\bar{F}(x) - F(x)]$ and just as natural to test H_0 versus H_1 with

$$(1.1)\qquad \begin{aligned} \phi_N^+ &= 1 \quad \text{if} \quad A_N^+ \equiv N \sup_x [\bar{F}_N(x) - F_N(x)] \geqq k_\alpha \\ &= 0 \quad \text{if} \quad A_N^+ < k_\alpha , \end{aligned}$$

where F_N is the empirical distribution function. The test (1.1) was proposed in 1969 [2] by C. Butler, who observed that the null distribution of A_N^+ was that of the maximum abscissae of an unrestricted symmetric random walk. A similar test based on the statistic

$$(1.2)\qquad N \sup_x [F_N(x) - \bar{F}_N(x) + 1 - 2F_N(0)]$$

(which has the same null distribution as (1.1)) was proposed by N. V. Smirnov [9] in 1947.

In 1971 Doksum and Thompson [3] defined classes of alternatives $\Omega(\Delta) = \{F; F \in \Omega, \bar{F} \geqq F, \sup_x [\bar{F}(x) - F(x)] \geqq \Delta\}$ and $\bar{\Omega}(\Delta) = \{F; F \in \Omega, \sup_x |\bar{F}(x) - F(x)| \leqq \Delta\}$. They then showed that for monotone rank tests ϕ, $\inf \{E_F \phi; F \in \Omega(\Delta)\} = \inf \{E_{\Delta,a} \phi; 0 \leqq a \leqq 1 - \Delta\}$ and $\sup \{E_F \phi; F \in \bar{\Omega}(\Delta)\} = E_\Delta \phi$ where $E_{\Delta,a}$ and E_Δ are expectations taken with respect to specific distributions $F_{\Delta,a}$ and G_Δ on $[-1, 1]$, respectively. We refer the reader to [3] for further details.

In this work we study the power bounds of the test ϕ_N^+ for the above described symmetry problem. In Section 2 we point out a simple criterion for determining whether or not a rank test is monotone; the monotonicity of ϕ_N^+

Received December 1972; revised January 1976.
AMS 1970 *subject classification.* Primary 62G10.
Key words and phrases. Smirnov statistics, monotone rank tests, power bounds.

follows trivially. Then we obtain an easily computable expression for the maximum power $\sup_{F \in \bar{\Omega}(\Delta)} E_F \phi_N^+$. In Section 3 Monte Carlo estimates of the lower bounds $E_{\Delta,a}\phi_N^+$ are presented for selected values of N, Δ, a and α.

The asymptotic power bounds of ϕ_N^+ are studied in Section 4. We show that for sequences $\{\Delta_N\}$ with $N^{\frac{1}{2}}\Delta_N \to c$, $0 \leqq c < \infty$, $\lim_{N\to\infty} \sup_{F \in \bar{\Omega}(\Delta_N)} E_F \phi_N^+ = \min\{2\Phi(c - k_\alpha), 1\}$ where Φ is the standard normal distribution function and $\Phi(-k_\alpha) = \alpha/2$. Moreover $\lim_{N\to\infty} \inf_{F \in \Omega(\Delta_N)} E_F \phi_N^+$ is approximately $\Phi(c - k_\alpha)$ if $c \geqq k_\alpha$ and equal to α if $c < k_\alpha$. These bounds are compared with those of the linear rank tests studied in [3] (see Remark 4.7). Finally we observe (Corollary 4.8) that in the smaller problem where it is known that $\inf_{F \in \Omega(\Delta)} E_F \Phi = E_{F_{\Delta,0}} \phi$ the asymptotic minimum power of ϕ_N^+ is approximately $\Phi(c - k_\alpha)$ for all sequences $N^{\frac{1}{2}}\Delta_N \to c$, $0 \leqq c < \infty$. This smaller problem contains both the location shift of a symmetric unimodal distribution and Lehmann's alternatives as special cases.

2. Theoretical results for finite samples. First we show that ϕ_N^+ is a monotone rank test, and then we derive an explicit expression for its maximum power against members of $\bar{\Omega}(\Delta)$.

A real-valued function f will be called *monotone* (*with respect to a partial ordering* $\prec$ *of its domain*), if for all a, b the relation $a \prec b$ implies $f(a) \leqq f(b)$. A test function ϕ (any measurable map of the sample space $\mathscr{X} = E^N$ into $[0, 1]$) is called *monotone* if it is monotone with respect to the ordering on $\mathscr{X}$ defined by $\mathbf{x} \prec \mathbf{y}$ if $x_i \leqq y_i$ for $i = 1, \cdots, N$. ϕ is a one-sample *rank test* if ϕ may be written as a composition $\psi \circ V$; the map V: $\mathscr{X} \to 2^N$ is defined by $V(\mathbf{x}) = \mathbf{v} = (v_1, \cdots, v_N)$, where $v_i = 1$ or 0 depending on whether the ith smallest in absolute value observation is positive or negative. Monotone rank tests have the desirable property of unbiasedness, (Corollary 2.1, [3]). The following proposition provides a criterion for determining the monotonicity of a rank test ϕ.

PROPOSITION 2.1. *Order* 2^N *by* $\mathbf{v} \prec \mathbf{w}$ *if and only if* $\sum_{i=j}^N v_i \leqq \sum_{i=j}^N w_i$ *for* $j = 1, \cdots, N$. *A rank test* $\phi = \psi \circ V$ *is monotone on* $\mathscr{X}$ *if and only if* ψ *is monotone on* 2^N.

PROOF. Let r_i^+ and s_i^+ denote the absolute ranks, respectively, of x_i within $\mathbf{x}$ and y_i within $\mathbf{y}$. Then

$$\begin{aligned}\sum_{i=j}^N v_i &= \sum_{k=1}^N I\{r_k^+ \geqq j\} I\{x_k \geqq 0\} \\ &\leqq \sum_{k=1}^N I\{s_k^+ \geqq j\} I\{y_k \geqq 0\} \\ &= \sum_{i=j}^N w_i \qquad \text{for } j = 1, \cdots, N.\end{aligned}$$

Thus the monotonicity of ψ implies that of ϕ.

To prove the converse, let $\mathbf{v} \prec \mathbf{w}$. Define

$$x_i = i(-1)^{1-v_i}, \qquad y_i = i(-1)^{1-w_i}, \qquad i = 1, \cdots, N.$$

Then $\mathbf{v} = V(\mathbf{x}_{(\cdot)})$, $\mathbf{w} = V(\mathbf{y}_{(\cdot)})$ where $\mathbf{x}_{(\cdot)}$ is the order statistic of $\mathbf{x}$. Moreover, $\mathbf{v} \prec \mathbf{w}$ implies

$$\sum_{i=1}^N I\{x_{(i)} \geqq j\} \leqq \sum_{i=1}^N I\{y_{(i)} \geqq j\}, \qquad j = 1, \cdots, N \tag{2.1}$$

which in turn implies that for each positive $x_{(k)}$, $x_{(k)} \leqq y_{(k)}$. Also, $\mathbf{v} \prec \mathbf{w}$ implies

$$\sum_{i=1}^{N} I\{x_{(i)} \leqq -j\} \geqq \sum_{i=i}^{N} \{y_{(i)} \leqq -j\}, \qquad j = 1, \cdots, N \tag{2.2}$$

which can be used to show $x_{(i)} \leqq y_{(i)}$ for each negative $y_{(i)}$. It follows that $x_{(\cdot)} \prec y_{(\cdot)}$ and the monotonicity of ϕ implies that of ψ.

REMARK 2.2. In 1959 I. R. Savage [8] showed that if a parametric family $F(x; \theta)$ has densities $f(x; \theta)$ statisfying certain regularity conditions then under alternatives ($\theta > 0$) to the hypothesis of symmetry ($\theta = 0$) the vectors $\mathbf{v} \prec \mathbf{w}$, $\mathbf{v} \neq \mathbf{w}$, satisfy $P_\theta(V(\mathbf{X}) = \mathbf{v}) < P_\theta(V(\mathbf{X}) = \mathbf{w})$.

COROLLARY 2.3. *The rank test* ϕ^+ *defined by* (1.1) *is monotone.*

PROOF. Observe that $\phi^+ = \psi^+ \circ V$, where

$$\begin{aligned} \psi^+ &= 1 \quad \text{if} \quad \max_{1\leqq j\leqq N} [\textstyle\sum_{i=j}^{N} (2v_i - 1)] \geqq k_\alpha \\ &= 0 \quad \text{if} \quad \max_{1\leqq j\leqq N} [\textstyle\sum_{i=j}^{N} (2v_i - 1)] < k_\alpha . \end{aligned}$$

and apply Proposition 2.1.

REMARK 2.4. The corresponding test based on (1.2) is *not* monotone since (1.2) may be written $\psi \circ V$ where $\psi(\mathbf{v}) = \max_{1\leqq j\leqq N} [\sum_{i=1}^{j} (2v_i - 1)]$. If $N = 5$, $\mathbf{v} = (1\ 1\ 1\ 1\ 0)$ and $\mathbf{w} = (0\ 1\ 1\ 1\ 1)$, then $\mathbf{v} \prec \mathbf{w}$ but $\psi(\mathbf{v}) > \psi(\mathbf{w})$.

We now determine the upper bound on $E_F\phi^+$ for F in $\bar{\Omega}(\Delta)$ defined in [3].

PROPOSITION 2.5 *Assume* ϕ^+ *is defined by* (1.1), *where* k_α *is of the form* $N - 2M - 1$, M *integral. Then*

(i) $\alpha = 2^{1-N} \sum_{j=0}^{M} \binom{N}{j}$ *and*

(ii) $\sup_{F\in\bar{\Omega}(\Delta)} E_F\varphi^+ = \sum_{k=0}^{k_\alpha-1} \binom{N}{k}\Delta^k(1-\Delta)^{N-k}2^{k-N+1} \sum_{j=0}^{M} \binom{N-k}{j} + \sum_{k=k_\alpha}^{N} \binom{N}{k}\Delta^k(1-\Delta)^{N-k}$.

(iii) *If* k_α *is of the form* $N - 2M$, M *integral, then* (i) *and* (ii) *remain true when* $\sum_{j=0}^{M} \binom{N-k}{j}$ *is replaced by* $\sum_{j=0}^{M-1} \binom{N-k}{j} + \frac{1}{2}\binom{N-k}{M}$.

PROOF. (i) Under the null hypothesis of symmetry $P\{A_N^+ = x\} = 2^{-N}[\binom{N}{(N+x)/2} + \binom{N}{(N+x+1)/2}$ where $\binom{N}{k}$ is zero unless k is integral in $[0, N]$, (page 2210, [2]). The result follows immediately by simple computation.

(ii) According to Corollary 2.2 of [3] and Corollary 2.3 above

$$\sup_{F\in\bar{\Omega}(\Delta)} E_F\phi^+ = \sum_{k=0}^{N} \binom{N}{k}\Delta^k(1 - \Delta)^{N-k}A(\phi^+, N, k)2^{k-N} \tag{2.3}$$

where $A(\phi^+, N, k)$ = number of vectors $\mathbf{v}$ in 2^N which have $v_i = 1$, $i = N, N - 1, \cdots, N - k + 1$ and which lead to rejection. If $k \geqq k_\alpha$, $A(\phi^+, N, k) = 2^{N-k}$ since $A_N^+ \geqq k$.

If $0 \leqq k < k_\alpha$, $A(\phi^+, N, k)$ = number of vectors in the critical region of a ϕ^+ test based on $N - k$ observations and having critical point $k_\alpha - k$; to see this observe that conditional on $v_i = 1$, $i = N - k + 1, \cdots N$,

$$\begin{aligned} A_N^+ &= \max\{k, k + \max_{1\leqq j\leqq N-k} \textstyle\sum_{i=j}^{N-k} (2v_i - 1)\} \\ &= \max\{k, k + A_{N-k}^+\}, \end{aligned}$$

where A_{N-k}^+ is based on the remaining $N - k$ observations. Thus when it is known that $v_i = 1$, $i = N - k + 1, \cdots, N$, $A_N^+ \geqq k_\alpha$ if and only if $A_{N-k}^+ \geqq k_\alpha - k$. We conclude by part (i) that $A(\phi^+, N, k) = 2 \sum_{j=0}^{M} \binom{N-k}{j}$ and the proof of (ii) is complete upon substitution of these values for $A(\phi^+, N, k)$ into (2.3).

(iii) The proof of (iii) follows the pattern of (i) and (ii) and will be omitted.

3. Computational results. In this section we first present numerical bounds on the power of ϕ_N^+ which can be compared with the bounds for the Wilcoxon and other tests studied in [3]. Then we examine the power of ϕ_N^+ against the least favorable (see page 16, [3]) distributions $F_{\Delta_N,a}$, $0 \leqq a \leqq 1 - \Delta_N$, for sequences $\{\Delta_N\}$ suggested by the asymptotic results in Section 4 below.

The expressson for the maximum power $E_{G_\Delta}\phi_N^+$ obtained analytically in the previous section (Proposition 2.5) is easily computed for given values of N, α and Δ; see Tables 3.1 and 3.2 for examples. The minimum power $\inf_{F \in \Omega(\Delta)} E_F \phi_N^+ = \inf_{0 \leqq a \leqq 1-\Delta} E_{\Delta,a}\phi_N^+$ (by Theorem 2.1 of [3] and Corollary 2.3 above) where $E_{\Delta,a}$ denotes integration with respect to $F_{\Delta,a}$, $0 \leqq a \leqq 1 - \Delta$. The integrals $E_{\Delta,a}\phi_N^+$ were analytically intractable, so we obtained Monte Carlo estimates of them using the techniques described at the end of this section. The choices $a = 0$ and $a = 1 - \Delta$ are suggested by the asymptotic results. A comparison of the bounds in Tables 3.1 and 3.2 with the corresponding bounds on the Wilcoxon and other tests considered in Tables 2.2–2.6, [3], reveals that while ϕ_N^+ may have smaller maximum power it also has greater minimum power than the latter, markedly so for larger values of Δ (see also Remark 4.7). In Table 3.1 and 3.2, $\phi_N^+ = 1, \gamma$, or 0 depending on whether A_N^+ is greater than, equal to, or less than k_α.

TABLE 3.1
Performance of ϕ_N^+ for $N = 10$

		Δ						
		0	.1	.2	.3	.4	.5	.75
$\alpha = .01$								
$k_\alpha = 8$	$\sup_{F \in \bar{\Omega}(\Delta)} E_F\phi_N^+$		.0223	.0453	.0866	.1555	.2623	.7022
$\gamma = .824$	$E_{\Delta,a}\phi_N^+$ for $a = 0$	.009	.018	.038	.062	.134	.222	.58
	$= \Delta$	.010	.011	.016	.019	.039	.042	
	$= \frac{1}{2}$	.010	.006	.011	.013	.027	.043	
	$= 1 - \Delta$	.007	.011	.013	.007	.016	.043(.047)*	
$\alpha = .05$								
$k_\alpha = 6$	$\sup_{F \in \bar{\Omega}(\Delta)} E_F\phi_N^+$		:0961	.1711	.2828	.4312	.6027	.9431
$\gamma = .649$	$E_{\Delta,a}\phi_N^+$ for $a = 0$	.051	.081	.130	.193	.311	.431	.780
	$= \Delta$	.052	.063	.091	.127	.197	.300	
	$= \frac{1}{2}$	.051	.053	.063	.010	.175	.300	
	$= 1 - \Delta$	.047	.045	.067	.075	.135	.300(.305)	
$\alpha = .10$								
$k_\alpha = 5$	$\sup_{F \in \bar{\Omega}(\Delta)} E_F\phi_N^+$		.1830	.3069	.4679	.6432	.8002	.9881
$\gamma = .787$	$E_{\Delta,a}\phi_N^+$ for $a = 0$	.098	.116	.184	.261	.392	.526	.875
	$= \Delta$	.098	.125	.162	.254	.373	.582	
	$= \frac{1}{2}$	.100	.109	.124	.215	.353	.559	
	$= 1 - \Delta$	.095	.098	.110	.179	.323	.559(.571)	

* Values in parentheses are exact values (to three places) for $a = \Delta = .5$.

TABLE 3.2
Performance of ϕ_N^+ for $N = 20$

		Δ						
		0	.1	.2	.3	.4	.5	.75
$\alpha = .01$								
$k_\alpha = 11$	$\sup_{F \in \bar{\Omega}(\Delta)} E_F \phi_N^+$		.0323	.0883	.2063	.4070	.622	.9919
$\gamma = .607$	$E_{\Delta,a}\phi_N^+$ for $a = 0$	.012	.019	.062	.118	.237	.429	.897
	$= \Delta$	.011	.022	.037	.081	.166	.349	
	$= \frac{1}{2}$	.010	.016	.017	.049	.143	.349	
	$= 1 - \Delta$	.012	.013	.009	.018	.107	.349(.349)	
$\alpha = .05$								
$k_\alpha = 8$	$\sup_{F \in \bar{\Omega}(\Delta)} E_F \phi_N^+$		.1280	.2789	.5078	.7515	.9177	.9998
$\gamma = .233$	$a = 0$	.057	.080	.147	.293	.471	.659	.976
	$= \Delta$	.057	.077	.147	.261	.470	.781	
	$= \frac{1}{2}$	.057	.057	.098	.220	.455	.781	
	$= 1 - \Delta$	.049	.049	.061	.160	.439	.771(.776)	
$\alpha = .1$								
$k_\alpha = 7$	$\sup_{F \in \bar{\Omega}(\Delta)} E_F \phi_N^+$		.2177	.4413	.6925	.8836	.9723	.9999
$\gamma = .586$	$a = 0$	.099	.136	.253	.418	.616	.787	.990
	$= \Delta$	.099	.132	.221	.413	.647	.911	
	$= \frac{1}{2}$	.101	.112	.182	.377	.655	.915	
	$= 1 - \Delta$	.096	.104	.153	.343	.685	.915(.911)	

Table 3.3 below contains Monte Carlo estimates of the power $\beta_N(a) = E_{\Delta_N,a}\phi_N^+$ of the level .05 test ϕ_N^+ for $0 \leqq a < 1 - \Delta_N$ and $\Delta_N = (1.96)N^{-\frac{1}{2}}$. These estimates are to be compared with $\beta(a) = \lim_{N^{\frac{1}{2}}\Delta_N \to c} \beta_N(a)$ obtained in Proposition 4.3, where c equals the asymptotic critical point 1.96.

In Tables 3.4 and 3.5 the level is still .05 but the alternatives $\{\Delta_N\}$ approach 0 through "small" and "large" sequences, respectively. In the first case $N^{\frac{1}{2}}\Delta_N = 1.46$; in the second $N^{\frac{1}{2}}\Delta_N = 2.46$. Again, the limiting power $\beta(a)$ is displayed for comparison. Note that in all three tables for $N \geqq 20$ most values of $\beta_N(a)$ exceed the asymptotic value and appear to converge down to it. This suggests that the asymptotic power $\beta(a)$, $0 \leqq a < 1$, is a conservative lower bound on the power of ϕ_N^+ for all F in $\Omega(\Delta_N)$, $N \geqq 20$.

TABLE 3.3
$\beta_N(a)$ for $\alpha = .05$, $N^{\frac{1}{2}}\Delta_N = 1.96$

N	a															
	0	.05	.1	.15	.2	.25	.3	.35	.4	.45	.5	.55	.6	.65	.7	.75
10	.597	.586	.593	.595	.603	.589	.605	.586								
20	.538	.557	.562	.569	.562	.578	.576	.585	.577	.580	.594	.585				
30	.530	.539	.546	.529	.562	.557	.556	.555	.571	.563	.571	.571	.574			
40	.547	.528	.523	.535	.540	.539	.527	.539	.548	.556	.559	.551	.539	.553		
100	.529					.533					.532					.561
∞*	.500	.500	.500	.500	.500	.500	.500	.500	.500	.500	.500	.501	.501	.502	.503	.504

* The asymptotic power $\beta(a)$ is computed by the method described in Remark 4.5 below.

TABLE 3.4
$\beta_N(a)$ *for* $\alpha = .05$, $N^{\frac{1}{2}}\Delta_N = 1.46$

N	a															
	0	.05	.1	.15	.2	.25	.3	.35	.4	.45	.5	.55	.6	.65	.7	.75
10	.376	.360	.352	.334	.324	.321	.309	.285	.273	.260	.248					
20	.325	.333	.343	.333	.336	.319	.319	.308	.311	.299	.281	.255	.243	.224		
30	.326	.332	.321	.312	.322	.324	.317	.309	.273	.291	.262	.261	.242	.232	.212	
40	.332	.312	.320	.314	.313	.313	.299	.289	.294	.295	.270	.259	.236	.214	.212	.177
100	.329					.299					.273					.188
∞	.309	.304	.298	.295	.289	.283	.281	.270	.262	.257	.246	.236	.229	.217	.198	.179

TABLE 3.5
$\beta_N(a)$ *for* $\alpha = .05$, $N^{\frac{1}{2}}\Delta_N = 2.46$

N	a															
	0	.05	.1	.15	.2	.25	.3	.35	.4	.45	.5	.55	.6	.65	.7	.75
10	.827	.845	.865	.889	.894											
20	.760	.777	.789	.792	.824	.827	.838	.358	.852							
30	.734	.756	.762	.763	.783	.785	.800	.810	.817	.833	.855	.885				
40	.743	.739	.754	.759	.765	.764	.790	.799	.804	.812	.828	.848	.877			
100	.718					.746					.813					.898
∞	.692	.696	.701	.706	.712	.718	.725	.732	.741	.750	.760	.772	.785	.801	.819	.841

In order to obtain the Monte Carlo estimates presented above, we used the antithetic method suggested by Hammersley and Hanscomb [4]. The details of our application follow.

If $\mathbf{U} = (U_1, \cdots, U_N)$ is a vector of i.i.d. uniform (0, 1) variables and $\mathbf{X} = (X_1, \cdots, X_N)$, where $X_i = F_{\Delta,a}^{-1}(U_i)$, we may write $\phi(\mathbf{U}) = \phi(\mathbf{X})$. Let $\mathbf{U}_1, \mathbf{U}_2, \cdots, \mathbf{U}_n$ be independent vectors, each distributed as $\mathbf{U}$. Then an unbiased estimator of $p = E_{F_{\Delta,a}}\phi$ is

$$\hat{p} = \frac{1}{2n} \sum_{j=1}^{n} [\phi(\mathbf{U}_j) + \phi(\mathbf{1} - \mathbf{U}_j)] \quad \text{where} \quad \mathbf{1} = (1, \cdots, 1) .$$

This estimator has variance

$$\operatorname{Var} \hat{p} = \frac{pq}{2n} + \frac{r - p^2}{2n}, \quad \text{where} \quad r = E[\phi(\mathbf{U}_j)(\mathbf{1} - \mathbf{U}_j)] ,$$

so that if $r < p^2$ the estimator $\hat{p}$ (which requires n samples and $2n$ computations) has smaller variance than the crude Monte Carlo estimate $(1/2n) \sum_{j=1}^{2n} \phi(\mathbf{U}_j)$ based on $2n$ samples and $2n$ computations. Since it is necessary that $r \geqq \max\{2p - 1, 0\}$, the optimal estimate obtains when $r = \max\{2p - 1, 0\}$. It is easily checked that under this condition, $\operatorname{Var} \hat{p} \leqq 1/16n$, $0 < p < 1$. In our application of this technique we generated $n = 1000$ samples so that optimally $\sigma_{\hat{p}} \leqq .008$. In fact the sample standard deviation $\hat{\sigma}_{\hat{p}}$ was less than .01 for all estimates obtained above. The uniform (0, 1) observations were generated (on the Michigan State University CDC 6500) by a linear congruential generator: $U_{n+1} = cU_n \bmod (2^{48})$ where $c = 553645$ and $U_0 = 1{,}274{,}321{,}477{,}413{,}155$ (base 8).

4. Asymptotic power bounds for ϕ^+. Throughout this section ϕ_N^+ will denote the test of asymptotic level α which rejects H_0 when $N^{\frac{1}{2}} \sup_x [\bar{F}_N(x) - F_N(x)] \geqq k_\alpha$. Also, $E_{\Delta,a}$ and E_Δ will denote expectations with respect to $F_{\Delta,a}$ and G_Δ defined on page 16, [3]. We determine the asymptotic power of ϕ_N^+ against appropriate sequences $\{F_{\Delta_N,a}\}$ and $\{G_{\Delta_N}\}$ in terms of standard Brownian motion W on [0, 1] and then in terms of the standard normal distribution function Φ. Then we find the asymptotic minimum power $\lim_{N\to\infty} \inf_{F\in\Omega(\Delta_N)} E_F \phi_N^+$.

Assume $\{\Delta_N\}$ satisfies

$$0 \leqq N^{\frac{1}{2}}\Delta_N \to c\,, \qquad \text{for some} \quad 0 \leqq c < \infty\,. \tag{4.1}$$

Define $k = k_\alpha = -\Phi^{-1}(\alpha/2)$,

$$\begin{aligned} \beta_N(a) &= E_{\Delta_N,a}\phi_N^+\,, & 0 \leqq a \leqq 1 - \Delta_N\,, \\ \beta(a) &= P\{\sup_x [W(x) + cI_{\{1-a\}}(x)] \geqq k\}\,, & 0 \leqq a < 1\,, \end{aligned} \tag{4.2}$$

and extend β by continuity:

$$\begin{aligned} \beta(1) = \lim_{a\uparrow 1}\beta(a) &= 1 & c > k \\ &= \frac{\alpha+1}{2} & c = k \\ &= \alpha & c < k\,. \end{aligned}$$

Unless otherwise specified, all supremum are taken over $0 \leqq x \leqq 1$ and all limits are taken as $N \to \infty$.

PROPOSITION 4.1. *Let $\{\Delta_N\}$ satisfy* (4.1). *Then*

(i) *the limiting upper bound on the power is*

$$\lim E_{\Delta_N}\phi_N^+ = P\{\sup W(x) \geqq k - c\}\,;$$

(ii) *if $\{a_N\}$ is any sequence satisfying $0 \leqq a_N \leqq 1 - \Delta_N$ and $a_N \to a$, then for $0 \leqq a < 1$ and for $a = 1$ when $c \neq k$, the limiting power against the least favorable $\{F_{\Delta_N,a_N}\}$ is*

$$\lim \beta_N(a_N) = \beta(a)\,.$$

PROOF. Let F represent F_{Δ_N,a_N} or G_{Δ_N} and define $\mu_N = N^{\frac{1}{2}}[\bar{F} - F]$ and $Z_N = N^{\frac{1}{2}}(\bar{F}_N - F_N) - \mu_N$ on [0, 1]. It can be shown, using e.g., Lemma 2.3 of [6] and the simple structure of $\{Z_N\}$, that $Z_N(\cdot) \Rightarrow W(1 - \cdot)$ where W is a standard Brownian motion on [0, 1] with continuous sample paths.

To prove (ii), note that $\beta_N(a_N) = P\{\sup [Z_N(x) + \mu_N(x)] \geqq k\}$, where

$$\begin{aligned} \mu_N(x) &= N^{\frac{1}{2}}(\Delta_N - |x - a_N|)\,, & |x - a_N| \leqq \Delta_N \\ &= 0 & |x - a_N| > \Delta_N\,. \end{aligned}$$

For any $\eta > 0$ and N sufficiently large

$$(c - \eta)I_{\{a_N\}}(x) \leqq \mu_N(x) \leqq (c + \eta)I_{[a-\eta,a+\eta]}(x)\,. \tag{4.3}$$

The right-hand inequality of (4.3) implies that eventually

$$\beta_N(a_N) \leqq P\{\sup [Z_N(x) + (c + \eta)I_{[a-\eta, a+\eta]}(x)] \geqq k\} .$$

But this bound converges to

$$P\{\sup [W(1 - x) + (c + \eta)I_{[a-\eta, a+\eta]}(x)] \geqq k\} \tag{4.4}$$

since the map $Z(\cdot) \rightarrow \sup [Z(x) + (c + \eta)I(x)]$ is continuous in the sup metric and $Z_N \Rightarrow W$ (see Chapter 13, [1]). Similarly the lower bound on μ_N from (4.3) implies that $\beta_N(a_N)$ is bounded below by a sequence converging to

$$P\{\sup [W(1 - x) + (c - \eta)I_{\{a\}}(x)] \geqq k\} . \tag{4.5}$$

The continuity of W implies that the limiting bounds (4.4) and (4.5) are arbitrarily close for η arbitrarily small (except when both $a = 1$ and $c = k$), and (ii) follows.

The proof of (i) proceeds in the same way, where now $F = G_{\Delta_N}$ and

$$\begin{aligned} \mu_N(x) &= N^{\frac{1}{2}}\Delta_N , \quad 0 \leqq x \leqq 1 - \Delta_N \\ &= 1 - x , \quad 1 - \Delta_N \leqq x \leqq 1 . \end{aligned}$$

REMARK 4.2. While the above proof of (ii) fails for the case $c = k$, $a = 1$, one can still obtain a useful lower bound on the limiting minimum power when $c = k$; see Proposition 4.6 (iii).

PROPOSITION 4.3.

(i) $P\{\sup_x W(x) \geqq k - c\} = \min \{2\Phi(c - k), 1\}$.

(ii) *For* $0 \leqq a < 1$,

$$1 - \beta(a) = \frac{[\Phi((c + k)(1 - a)^{-\frac{1}{2}}) - \Phi((c - k)(1 - a)^{-\frac{1}{2}})]}{\Phi((k - c)(1 - a)^{-\frac{1}{2}})]} \times \int_{-\infty}^{t} \left[2\Phi\left(\frac{k - x(1 - a)^{\frac{1}{2}}}{a^{\frac{1}{2}}}\right) - 1\right] d\Phi(x)$$

where $t = (k - c)(1 - a)^{-\frac{1}{2}}$.

PROOF OF (ii). For any real k,

$$\begin{aligned} 1 - \beta(a) &= P\{\sup_{0 \leqq x \leqq 1} [W(x) + cI_{\{1-a\}}(x)] \leqq k\} \\ &= P\{U \leqq k, X \leqq k - c, V \leqq k\} , \end{aligned}$$

where

$$U = \sup_{0 \leqq x \leqq 1-a} W(x) , \qquad X = W(1 - a) , \qquad \text{and} \qquad V = \sup_{1-a \leqq x \leqq 1} W(x) .$$

The Markov property of the process implies that

$$1 - \beta(a) = \frac{P\{U \leqq k, X \leqq k - c\}P\{V \leqq k, X \leqq k - c\}}{P\{X \leqq k - c\}} . \tag{4.6}$$

In problem (2), page 181, [5], one derives the joint density of U and X to be

$$f_{U,X}(u, x) = \frac{2(2u - x)}{2\pi^{\frac{1}{2}}(1 - a)^{\frac{3}{2}}} \exp[-(2u - x)^2/2(1 - a)] \qquad x \leqq u \geqq 0 . \tag{4.7}$$

For $k \geqq c$

$$P\{U \leqq k, X \leqq k - c\} = \int_0^{k-c} \int_{-\infty}^{u} f_{U,X}(u, x)\, du\, dx + \int_{k-c}^{k} \int_{-\infty}^{k-c} f_{U,X}(u, x)\, du\, dx$$

which upon substitution of $t = (2u - x)^2/(1 - a)$ yields

$$P\{U \leqq k, X \leqq k - c\} = \Phi\left(\frac{k + c}{(1 - a)^{\frac{1}{2}}}\right) - \Phi\left(\frac{c - k}{(1 - a)^{\frac{1}{2}}}\right). \tag{4.8}$$

The case $k \leqq c$ is simpler and also leads to (4.8). The joint density of V and X is derived from (4.7) using the spatial homogeneity of the process; it is

$$f_{V,X}(v, x) = \frac{1}{(a(1 - a))^{\frac{1}{2}}\pi} \exp -[((v - x)^2/2a + x^2/2(1 - a)] , \qquad x \leqq v . \tag{4.9}$$

Thus

$$P\{V \leqq k, X \leqq k - c\} = \int_{-\infty}^{(k-c)/(1-a)^{\frac{1}{2}}} \left[2\Phi\left(\frac{k - x(1 - a)^{\frac{1}{2}}}{a^{\frac{1}{2}}}\right) - 1\right] d\Phi(x) . \tag{4.10}$$

Combining (4.6), (4.8), and (4.10) we obtain the exact expression for $\beta(a)$. The proof of (i) is on page 276, [5].

REMARK 4.4. We have the useful bounds (obtained by taking obvious lower and upper bounds on the integrand in (4.10)).

$$\begin{aligned} &\left[\Phi\left(\frac{c - k}{(1 - a)^{\frac{1}{2}}}\right) + \Phi\left(\frac{-c - k}{(1 - a)^{\frac{1}{2}}}\right)\right] \\ &\qquad \leqq \beta(a) \leqq 1 - \left[2\Phi\left(\frac{c}{a^{\frac{1}{2}}}\right) - 1\right]\left[\Phi\left(\frac{c + k}{(1 - a)^{\frac{1}{2}}}\right) - \Phi\left(\frac{c - k}{(1 - a)^{\frac{1}{2}}}\right)\right] \\ &\qquad\qquad \text{for} \quad 0 \leqq a < 1 . \end{aligned} \tag{4.11}$$

A quick estimate of $\beta(a)$ is implied by these bounds, namely $\Phi((c - k)/(1 - a)^{\frac{1}{2}}) + \Phi((-c - k)/(1 - a)^{\frac{1}{2}})$ or even simply $\Phi((c - k)/(1 - a)^{\frac{1}{2}})$.

REMARK 4.5. The exact value of $\beta(a)$, $0 < a < 1$, can be obtained with the aid of a table of the bivariate normal distribution (8.5, page 184 [7], say). For we may write the integral

$$\int_{-\infty}^{(k-c)/(1-a)^{\frac{1}{2}}} \Phi\left(\frac{k - x(1 - a)^{\frac{1}{2}}}{a^{\frac{1}{2}}}\right) d\Phi(x)$$

in the form

$$P\left\{a^{\frac{1}{2}}Z_1 + (1 - a)^{\frac{1}{2}}Z_2 \leqq \frac{k - c}{(1 - a)^{\frac{1}{2}}}\right\} = P\left\{Z_3 \leqq k, Z_2 \leqq \frac{k - c}{(1 - a)^{\frac{1}{2}}}\right\} ,$$

where Z_1, Z_2, Z_3 each have marginal distribution Φ, Z_1, Z_2 are independent; and $\mathrm{Cov}\,(Z_2, Z_3) = (1 - a)^{\frac{1}{2}}$.

In many applications the point "a" is not given (see Corollary 4.8 for an exception) so that we need to consider the global behavior of $\beta(a)$, $0 \leqq a \leqq 1$. Insofar as $\beta(a)$ is approximately $\Phi((c - k)/(1 - a)^{\frac{1}{2}})$ (see Remark 4.4), we may say that $\beta(a)$ is monotone increasing, constant, or decreasing depending on whether c (and hence $N^{\frac{1}{2}}\Delta_N$) is large, equal or small relative to the critical point k.

The asymptotic minimum power is found in the next proposition.

Let $\underline{\beta}$, $\bar{\beta}$, and β^* denote respectively the liminf, limsup and limit of $\inf\{E_F\phi_N^+ : F \in \Omega(\Delta_N)\}$.

PROPOSITION 4.6. *Let* $\{\Delta_N\}$ *satisfy* (4.1). *Then for* $c \neq k$, $\beta^* = \inf_{0\leqq a\leqq 1}\beta(a)$, *and furthermore*

(i) *if* $c < k$, $\beta^* = \alpha$;

(ii) *if* $c > k$, $\Phi(c - k) \leqq \beta^* \leqq \Phi(c - k) + \Phi(-c - k)$;

(iii) *if* $c = k$, $\frac{1}{2} \leqq \underline{\beta} \leqq \bar{\beta} \leqq \frac{1}{2} + \Phi(-2k)$.

PROOF. By Theorem 2.1, [3], and Corollary 2.3 above, it suffices to consider the limiting behavior of $\inf_{0\leqq a\leqq 1-\Delta_N} \beta_N(a)$. This quantity equals $\inf_{0\leqq a\leqq 1} \beta(a)$ in the limit if

$$\sup_{0\leqq a\leqq 1-\Delta_N} |\beta_N(a) - \beta(a)| \to 0 . \tag{4.12}$$

But if (4.12) fails, the continuity of $\beta_N - \beta$ in $a \in [0, 1 - \Delta_N]$ implies there exists $\eta > 0$ and a convergent subsequence $a_{N'} \to a$, $0 \leqq a \leqq 1$, such that $|\beta_{N'}(a_{N'}) - \beta(a_{N'})| \geqq \eta$ for all N'. But in view of Proposition 4.1 (ii) and the continuity of β this is not possible. Thus for $c \neq k$, $\beta^* = \inf_{0\leqq a\leqq 1} \beta(a)$.

To see (i), note that for $c < k$, $\beta(1) = \alpha$. Moreover, it follows from Corollary 2.1 of [3] and Corollary 2.3 above that ϕ_N^+ is unbiased, so $\beta^* = \alpha$.

When $c > k$ (case (ii)), $\beta(a)$ is bounded below (see (4.11)) by the monotone increasing function $\Phi((c - k)/(1 - a)^{\frac{1}{2}})$ of a, $0 \leqq a \leqq 1$, so that $\beta^* \geqq \Phi(c - k)$. Since $\beta^* = \inf_{0\leqq a\leqq 1} \beta(a) \leqq \beta(0) = \Phi(c - k) + \Phi(-c - k)$, the result follows.

In case (iii), $c = k$, $\bar{\beta} = \limsup \inf_{0\leqq a\leqq 1-\Delta_N} \beta_N(a) \leqq \limsup \beta_N(0) = \beta(0)$ by Proposition 4.1 (ii). On the other hand, for any sequence $0 \leqq a_N \leqq 1 - \Delta_N$,

$$\begin{aligned}\beta_N(a_N) &= P\{\sup [Z_N(x) + cI_{\{a_n\}}(x)] \geqq c\}\\ &\geqq P\{Z_N(a_N) \geqq 0\} \to \tfrac{1}{2}\end{aligned}$$

since $Z_N(a_N)/(\text{Var } Z_N(a_N))^{\frac{1}{2}}$ is a normalized sum of i.i.d. random variables which satisfies the Lindeberg–Feller condition. Hence

$$\underline{\beta} = \liminf_{0\leqq a_N\leqq 1-\Delta_N} \beta_N(a_N) \geqq \tfrac{1}{2} .$$

REMARK 4.7. From Proposition 4.6 we conclude that for sufficiently large N, $\inf\{E_F\phi_N^+ ; F \in \Omega(\Delta_N)\}$ behaves like α when $N^{\frac{1}{2}}\Delta_N < k_\alpha = -z_{\alpha/2}$ and has a lower bound $\Phi(N^{\frac{1}{2}}\Delta_N + z_{\alpha/2})$, when $N^{\frac{1}{2}}\Delta_N > -z_{\alpha/2}$.

By way of comparison, Lemma 4.1 and Theorem 4.1 of [3] show that for sufficiently large N the asymptotic minimax test $\phi(\Delta)$ defined by (4.8) of [3] has $\inf\{E_F\phi(\Delta) ; F \in \Omega(\Delta_N)\} \approx \Phi((3N)^{\frac{1}{2}}\Delta_N^2 + z_\alpha)$. For $N^{\frac{1}{2}}\Delta_N < -z_{\alpha/2}$ and N large, both tests have power near α; for $N^{\frac{1}{2}}\Delta_N > -z_{\alpha/2}$ and N large, the lower bound of the minimum power of ϕ_N^+ is greater than the minimum power of $\phi(\Delta)$ when

$N^{\frac{1}{2}}\Delta_N + z_{\alpha/2} \geqq (3N)^{\frac{1}{2}}\Delta_N^2 + z_\alpha$. Solving this quadratic in Δ_N we find ϕ_N^+ superior to $\phi(\Delta_N)$ on the interval $(2\cdot 3^{\frac{1}{2}})^{-1}\,[1 \pm \{1 - 4(3/N)^{\frac{1}{2}}(z_\alpha - z_{\alpha/2})\}^{\frac{1}{2}}]$ which approaches the interval $[0, (\frac{1}{3})^{\frac{1}{2}}]$ as $N \to \infty$.

If we compare these same two tests in terms of Pitman's efficiency, we find because of the differences in rates that as $\Delta_N \to 0$, ϕ_N^+ is infinitely more efficient than $\phi(\Delta_N)$.

In the previous discussion it has been assumed that the point "a" is unknown, but there are important statistical problems for which it is known that $a = 0$ and then the following corollary may be useful.

COROLLARY 4.8. *Let $\Omega_0(\Delta)$ denote the subclass of $\Omega(\Delta)$ for which*

$$\sup_x [\bar{F}(x) - F(x)] = \bar{F}(0) - F(0),\ F \in \Omega(\Delta)\,. \tag{4.13}$$

Then for any sequence $\{\Delta_N\}$ satisfying (4.1) *we have*

$$\lim_{N\to\infty} \inf_{F\in\Omega_0(\Delta_N)} E_F\phi_N^+ = \Phi(c-k) + \Phi(-c-k)\,, \quad k = -\Phi^{-1}(\alpha/2)\,. \tag{4.14}$$

PROOF. If X has distribution F_X satisfying (4.13) then $Y = F_X(X) - F_X(-X)$ has distribution function F_Y also satisfying (4.13). Thus by Lemma 2.2 and the proof of Theorem 2.1 ([3]), for any monotone rank test ϕ, $\inf_{F\in\Omega_0(\Delta)} E_F\phi = E_{\Delta,0}\phi$. In particular (4.14) follows from Proposition 4.1 above.

This corollary provides a lower bound on the asymptotic minimum power of ϕ_N^+ in any problem for which the alternatives to symmetry form a subclass of $\Omega_0(\Delta)$, $\Delta > 0$.

EXAMPLE 4.9. Let $\Omega_U = \{F_\theta \in \Omega : F_\theta(x) = F(x-\theta)$ where $\theta \geqq 0$ and F has a symmeric unimodal density$\}$. The testing problem $H_0: \theta = 0$ versus $H_1: \theta > 0$ falls within the symmetry problem described in Section 1. Ona may easily check that $\Delta = \sup_x [\bar{F}_\theta(x) - F_\theta(x)] = \bar{F}_\theta(0) - \bar{F}_\theta(0) = 2F(\theta) - 1$ so $\Omega_U(\Delta) \subset \Omega_0(\Delta)$; by Corollary 4.8, $\inf_{F\in\Omega_U(\Delta)} \phi_N^+$ is bounded below by $\Phi(c - k_\alpha)$ for $N^{\frac{1}{2}}\Delta_N$ approaching c.

EXAMPLE 4.10. Let $\Omega_L = \{F_\theta \in \Omega : F_\theta(x) = F^\theta(x)$ where $\theta \geqq 1$ and F is symmetric$\}$. Then for testing $H_0: \theta = 1$ versus $H_1: \theta > 1$ we find $\Omega_L(\Delta) \subset \Omega_0(\Delta)$, $\Delta = \sup [\bar{F}_\theta(x) - F_\theta(x)] = 1 - 2F^\theta(0)$ and so again $\Phi(c - k_\alpha)$ provides an asymptotic lower bound on the power.

Acknowledgment. The authors wish to thank Professor Peter J. Bickel and the referee for helpful suggestions.

REFERENCES

[1] BREIMAN, LEO (1968). *Probability*. Addison-Wesley, Reading, Mass.

[2] BUTLER, CALVIN C. (1969). A test for symmetry using the sample distribution function. *Ann. Math. Statist.* **40** 2209–2210.

[3] DOKSUM, KJELL and THOMPSON, RORY (1971). Power bounds and asymptotic minimax results for one sample rank tests. *Ann. Math. Statist.* **42** 12–34.

[4] HAMMERSLEY, J. M. and HANSCOMB, D. C. (1964). *Monte Carlo Methods*. Methuen, London.
[5] KARLIN, S. (1966). *A First Course in Stochastic Processes*. Academic Press, New York.
[6] KOUL, HIRA LAL (1970). Some convergence theorems for ranks and weighted empirical cumulatives. *Ann. Math. Statist.* **41** 1768-1773.
[7] OWEN, D. B. (1962). *Handbook of Statistical Tables*. Addison-Wesley, Reading, Mass.
[8] SAVAGE, I. R. (1959). Contributions to the theory of rank order tests—the one sample case. *Ann. Math. Statist.* **30** 1018-1023.
[9] SMIRNOV, N. V. (1947). Sur un critère de symétrie de la loi de distribution d'une variable aléatoire. *C. R.* (*Doklady*) *Acad. Sci. URSS* **56** 11-14.

DEPARTMENT OF STATISTICS
MICHIGAN STATE UNIVERSITY
EAST LANSING, MICHIGAN 48824

MATHEMATICS DEPARTMENT
LA TROBE UNIVERSITY
BUNDOORA, VICTORIA
AUSTRALIA 3083