The Annals of Mathematical Statistics
1971, Vol. 42, No. 5, 1595–1605

PITMAN EFFICIENCIES OF KOLMOGOROV–SMIRNOV TESTS[1]

BY C. S. YU

State University of New York at Albany

A comparison, by means of Pitman asymptotic efficiency, is made between the Kolmogorov–Smirnov test and the locally most powerful rank and the locally asymptotically most powerful (Neyman) test for testing two-sided shifts in the two-sample problem under the assumption that the true distribution is different from the one assumed. It is shown that the behavior of the bounds for the Pitman asymptotic efficiencies are the same as those for testing the one-sided shift using the Smirnov test in place of the Kolmogorov–Smirnov test.

1. Introduction. Let $X_1, X_2, \cdots, X_m$ and $Y_1, Y_2, \cdots, Y_n$ be ordered independent random samples from continuous cumulative distributions $F(x)$ and $G(x)$ respectively with $0 < \tau = m/n$. We are interested in testing the following statistical hypothesis:

$$H_0 : F(x) = G(x)$$
$$H_1 : G(x) = F(x-\theta), \qquad \theta \neq 0.$$

A number of optimal (most powerful) tests can be constructed based on $F(x)$; among them are locally most powerful rank and locally asymptotically most powerful ("Neyman") tests. It has been shown that these tests possess certain optimality properties, see Chernoff and Savage (1958) and Neyman (1959). However, in many instances, the distribution of X's is actually $\Psi(x)$ which is different from $F(x)$.

In this paper, we will be concerned with comparing the performance of the above optimal tests with Kolmogorov–Smirnov test when the underlying distribution of the X's is $\Psi(x)$.

The above-mentioned tests are based correspondingly on the following statistics:

(i) The Kolmogorov–Smirnov statistic:

$$K_{m,n} = \sup_x \left(\frac{mn}{m+n}\right)^{\frac{1}{2}} |F_m(x) - G_n(x)|$$

where $F_m(x)$ and $G_n(x)$ are empirical cumulative distributions of the X's and Y's respectively.

(ii) The locally most powerful rank (LMPR) statistic:

$$T_N{}^* = \sum_{j=1}^N E[g(V^{(j)})]Z_j$$

Received June 29, 1971.

[1] Research supported in part by a Faculty Research Fellowship awarded by the Research Foundation, State University of New York and NSF Grant GU 3171.

Institute of Mathematical Statistics is collaborating with JSTOR to digitize, preserve, and extend access to *The Annals of Mathematical Statistics.* www.jstor.org

where $V^{(j)}$ is the jth order statistic in the joint sample of X's and Y's; Z_j equals zero or one according to whether $V^{(j)}$ is an X or Y, and $g(\cdot) = -f'(\cdot)/f(\cdot)$ where f is the density of F.

(iii) The locally asymptotically most powerful ("Neyman") statistic:

$$T_N = \frac{\tau}{1+\tau}\sum_{j=1}^{n} g(Y_j) - \frac{1}{1+\tau}\sum_{i=1}^{m} g(X_i), \qquad \tau = m/n.$$

This test is asymptotically equivalent to the likelihood ratio test. For a detailed discussion of this test, see Neyman (1959).

In order to properly construct the statistics and ensure the asymptotic normality of the rank statistic, and the derivation of certain functionals, see Chernoff and Savage (1958), and Kalish and Mikulski (1971), we shall impose the following conditions on the cumulative distribution function $F(x)$:

CONDITION 1. $F(x)$ is twice continuously differentiable for real x.

CONDITION 2. $f(x) = F'(x)$ for all x and for $c > 0$, we have

$$\lim_{\zeta\to 0}\int_{-\infty}^{\infty}\left[\frac{f'(x)}{f(x)}\right]^2 f(cx+\zeta)\,dx = \int_{-\infty}^{\infty}\left[\frac{f'(x)}{f(x)}\right]^2 f(cx)\,dx < \infty.$$

CONDITION 3. The function $g(x) = -f'(x)/f(x)$ is strictly increasing and $g(0) = 0$.

CONDITION 4. Let $F^*(x) = F[h(x)]$ where $h(x) = g^{-1}(x)$. Let F^{*-1} be the inverse function of F^*. Then

$$\left|\frac{d^k}{du^k}F^{*-1}(u)\right| \leqq C[u(1-u)]^{-k-\frac{1}{2}+\delta}$$

for $k = 0, 1, 2$; for all $u \in (0, 1)$; for some C and some $\delta > 0$.

CONDITION 5. $F(x)$ has finite variance, i.e. $\mathrm{Var}_F X < \infty$.

CONDITION 6. $g(x)$ is twice differentiable; $g''(x)$ is F-integrable and uniformly continuous.

CONDITION 7. $\int_{-\infty}^{\infty} g'(x)f(x)\,dx < \infty$.

We shall take the definition of Pitman efficiency as given by Fraser (1957): Let $\{\phi_N\}$ and $\{\phi_N{}^*\}$ be sequences of tests of size α for testing

$$H_0\colon \theta = \theta_0$$

$$H_{1_i}\colon \theta = \theta_i \qquad \text{such that } \theta_i \to \theta_0 \text{ as } i \to \infty,$$

where i is the subscript of $\{N_i\} \subset \{N\}$ and $\{N_i{}^*\} \subset \{N\}$ such that

$$\lim_{i\to\infty} E_{\theta_i}[\phi_{N_i}] = \lim_{i\to\infty} E_{\theta_i}[\phi^*_{N_i^*}] = \beta, \qquad 0 < \beta < 1.$$

The Pitman efficiency or the relative asymptotic efficiency of $\{\phi_N\}$ with respect to $\{\phi_N{}^*\}$ is defined to be:

$$\underline{e}_{\phi_N\phi_N*} = \lim_{i\to\infty} \frac{N_i{}^*}{N_i} = \frac{\lim_{i\to\infty} N_i{}^*\theta_i{}^2}{\lim_{i\to\infty} N_i\theta_i{}^2},$$

provided that this limit exists and is independent of the choices for $\{N_i\}$ and $\{N_i{}^*\}$; if this limit does not exist, we define the generalized upper and lower Pitman asymptotic efficiencies correspondingly to be:

$$\bar{e}_{\phi_N\phi_N*} = \limsup_{i\to\infty} \frac{N_i{}^*}{N_i} \leqq \frac{\limsup_{i\to\infty} N_i{}^*\theta_i{}^2}{\liminf_{i\to\infty} N_i\theta_i{}^2}$$

$$\underline{e}_{\phi_N\phi_N*} = \liminf_{i\to\infty} \frac{N_i{}^*}{N_i} \geqq \frac{\liminf_{i\to\infty} N_i{}^*\theta_i{}^2}{\limsup_{i\to\infty} N_i\theta_i{}^2}.$$

We shall adopt the following notations:

$\bar{e}_{KT}(F;\Psi)$ = the upper Pitman efficiency computed under $\Psi(x)$ of Kolmogorov–Smirnov test with respect to the Neyman test derived for $F(x)$;

$\bar{e}_{KT^*}(F;\Psi)$ = the upper Pitman efficiency computed under $\Psi(x)$ of Kolmogorov–Smirnov test with respect to LMPR test derived for $F(x)$;

$\underline{e}_{KT}(F;\Psi)$, $\underline{e}_{KT^*}(F;\Psi)$ = the respective lower Pitman efficiencies;

$e_{TT^*}(F;\Psi)$ = Pitman efficiency computed under $\Psi(x)$ of Neyman test with respect to LMPR test derived for $F(x)$.

where $\Psi(x)$ is taken from a class $\mathscr{C}$ of absolutely continuous cumulative distribution functions with $\Psi'(x) = \psi(x)$ a.e. such that for some x_0, $\psi(x)$ is non-decreasing for $x \leqq x_0$ and non-increasing for $x \geqq x_0$, without loss of generality, we can assume $x_0 = 0$.

It will be shown that the above-mentioned upper and lower bounds have the same behavior as those for the Smirnov test described in Kalish and Mikulski (1971) and Mikulski and Yu (1971).

2. Asymptotic bounds for the power of Kolmogorov–Smirnov test. Consider the following sequence of two-sample problems: Let $X_1, X_2, \cdots, X_m$ and $Y_1, Y_2, \cdots, Y_n$ be ordered independent random samples from continuous cumulative distributions $F(x)$ and $G(x)$ respectively; we shall assume that $m/n = \tau > 0$, τ a constant, and we wish to consider the statistical hypotheses

$$H_0: F(x) = G(x)$$

$$H_{1_n}: F(x) = F^{(1)}(x), G(x) = G^{(1)}_{(n)}(x)$$

such that $\sup_x |F^{(1)}(x) - G^{(1)}_{(n)}(x)| = \delta_n = \delta_n{}^{-\frac{1}{2}} + o(n^{-\frac{1}{2}}), \delta > 0$ and the supremum is achieved at some point η_n, i.e. $|F^{(1)}(\eta_n) - G^{(1)}_{(n)}(\eta_n)| = \delta_n$.

The Kolmogorov–Smirnov test as defined in Section 1 can be used to test the above hypothesis; the null hypothesis is rejected if $K_{m,n} \geqq d_n(\alpha)$ where $d_n(\alpha)$ is determined by the relation

$$P\{K_{m,n} \geqq d_n(\alpha) \mid F(x)\} = \alpha.$$

It has been found in Massey (1950) that for large n, the lower bound for the power of this test is given by $1-(2\pi)^{-\frac{1}{2}}\int_{\lambda_1}^{\lambda_2} e^{-\frac{1}{2}t^2}\,dt$ with $\lambda_1 = (A-B)/C$, $\lambda_2 = (A+B)/C$ where

$$A = \delta_n, \quad B = \left(\frac{mn}{m+n}\right)^{-\frac{1}{2}} d_n(\alpha),$$

$$C = \left\{\frac{F^{(1)}(\eta_n)[1-F^{(1)}(\eta_n)]}{m} + \frac{G_{(n)}^{(1)}(\eta_n)[1-G_{(n)}^{(1)}(\eta_n)]}{n}\right\}^{\frac{1}{2}}.$$

It was observed in Capon (1965) that since

$$\lambda_1 \geqq \lambda_1' = 2\left[\delta_n\left(\frac{mn}{m+n}\right)^{\frac{1}{2}} - d_n(\alpha)\right], \quad \text{for} \quad \delta_n\left(\frac{mn}{m+n}\right)^{\frac{1}{2}} - d_n(\alpha) > 0,$$

we have the power of the Kolmogorov–Smirnov test for large n is bounded from below by

$$1-(2\pi)^{-\frac{1}{2}}\int_{\lambda_1'}^{\infty} e^{-\frac{1}{2}t^2}\,dt. \tag{2.1}$$

In order to obtain an upper bound for the power, we first consider the following one-sample problem:

Let $X_1, X_2, \cdots, X_n$ be an ordered random sample from a random variable X with continuous cumulative distribution $F(x)$; we are testing the statistical hypothesis

$$H_0: F(x) = H(x)$$
$$H_1: F(x) = G^{(n)}(x)$$

such that $\sup_x|H(x)-G^{(n)}(x)| = \delta_n = \delta n^{-\frac{1}{2}}+o(n^{-\frac{1}{2}})$, $\delta > 0$ and the supremum is achieved at some point ζ_n, i.e.

$$\delta_n = |H(\zeta_n)-G^{(n)}(\zeta_n)|.$$

It has been found in Quade (1965) that if the Kolmogorov–Smirnov statistic $K_n = \sup_x n^{\frac{1}{2}}|F_n(x)-H(x)|$ is used to test this hypothesis, then, for large n, the power of the Kolmogorov–Smirnov test is bounded from above by the expression

$$\begin{cases} 2\sum_{k=1}^{\infty}(-1)^{k-1}\exp[-2k^2(d_n(\alpha)-n^{\frac{1}{2}}\delta_n)^2] & \text{for } n^{\frac{1}{2}}\delta_n < d_n(\alpha) \\ 1 & \text{for } n^{\frac{1}{2}}\delta_n \geqq d_n(\alpha) \end{cases} \tag{2.2}$$

where $d_n(\alpha)$ is determined by the relation $P\{K_n \geqq d_n(\alpha) \mid H(x)\} = \alpha$.

Now, we shall derive from (2.2) an asymptotic upper bound for the power of the $K_{m,n}$ test for testing the two-sample problem stated at the beginning of this section as follows:

$$\sup_x |F_m(x)-G_n(x)| \leqq \sup_x |F^{(1)}(x)-G_n(x)|+\sup_x |F_m(x)-F^{(1)}(x)|.$$

For $0 < \gamma < 1$, let

$$\left(\frac{mn}{m+n}\right)^{\frac{1}{2}} \sup_x |F^{(1)}(x)-G_n(x)| < (1-\gamma)d_n(\alpha)$$

and

$$\left(\frac{mn}{m+n}\right)^{\frac{1}{2}} \sup_x |F_m(x)-F^{(1)}(x)| < \gamma d_n(\alpha),$$

the two inequalities above imply

$$\left(\frac{mn}{m+n}\right)^{\frac{1}{2}} \sup_x |F_m(x)-G_n(x)| < d_n(\alpha).$$

Hence, by independence of the random samples, we have

$$P\left\{\left(\frac{mn}{m+n}\right)^{\frac{1}{2}} \sup_x |F_m(x)-G_n(x)| < d_n(\alpha) \mid F^{(1)}(x), G^{(1)}(x)\right\}$$
$$\geqq P\left\{n^{\frac{1}{2}} \sup_x |F^{(1)}(x)-G_n(x)| < (1-\gamma)\left(\frac{m+n}{m}\right)^{\frac{1}{2}} d_n(\alpha) \mid G^{(1)}(x)\right\}$$
$$\cdot P\left\{m^{\frac{1}{2}} \sup_x |F_m(x)-F^{(1)}(x)| < \gamma\left(\frac{m+n}{n}\right)^{\frac{1}{2}} d_n(\alpha) \mid F^{(1)}(x)\right\}.$$

Set $\gamma = (n/(m+n))^{\frac{1}{2}}$, we obtain

$$P\{m^{\frac{1}{2}} \sup_x |F_m(x)-F^{(1)}(x)| < d_n(\alpha) \mid F^{(1)}(x)\} = 1-\alpha$$

and

$$P\{n^{\frac{1}{2}} \sup_x |F^{(1)}(x)-G_n(x)| < \frac{(\tau+1)^{\frac{1}{2}}-1}{\tau^{\frac{1}{2}}} d_n(\alpha) \mid G^{(1)}(x)\}$$
$$\geqq 1-2\sum_{k=1}^{\infty} (-1)^{k-1} e^{-2k^2\left(\frac{(\tau+1)^{\frac{1}{2}}-1}{\tau^{\frac{1}{2}}} d_n(\alpha)-n^{\frac{1}{2}}\delta_n\right)^2}.$$

For large n, the quantity $d_n(\alpha)$ can be approximated by $[\frac{1}{2}\log(2/\alpha)]^{\frac{1}{2}}$, for a discussion on this approximation, see Hodges (1957). Incorporate above, we have

$$P\left\{\left(\frac{mn}{m+n}\right)^{\frac{1}{2}} \sup_x |F_m(x)-G_n(x)| < d_n(\alpha) \mid F^{(1)}(x), G^{(1)}(x)\right\}$$
$$\geqq (1-\alpha)\left(1-2\sum_{k=1}^{\infty} (-1)^{k-1} e^{-2k^2\left(\frac{(\tau+1)^{\frac{1}{2}}-1}{\tau^{\frac{1}{2}}}[\frac{1}{2}\log(2/\alpha)]^{\frac{1}{2}}-n^{\frac{1}{2}}\delta_n\right)^2}\right).$$

Finally, we note that the left-hand side of the inequality is one minus the power for the $K_{m,n}$ test; hence the power of the $K_{m,n}$ test is bounded from above by

$$(2.3)\quad \alpha+2(1-\alpha)\sum_{k=1}^{\infty}(-1)^{k-1}e^{-2k^2\left(\frac{(\tau+1)^{\frac{1}{2}}-1}{\tau^{\frac{1}{2}}}[\frac{1}{2}\log(2/\alpha)]^{\frac{1}{2}}-n^{\frac{1}{2}}\delta_n\right)^2}$$

$$< \alpha+2(1-\alpha)e^{-2\left(\frac{(\tau+1)^{\frac{1}{2}}-1}{\tau^{\frac{1}{2}}}[\frac{1}{2}\log(2/\alpha)]^{\frac{1}{2}}-n^{\frac{1}{2}}\delta_n\right)^2}.$$

In our subsequent computations, we shall use the last expression as the upper bound of the power for the Kolmogorov–Smirnov test.

3. Computation of bounds for Pitman efficiencies. In this section, we will compute the upper and lower bounds of the Pitman asymptotic efficiency of the Kolmogorov–Smirnov test with respect to the Neyman test and LMPR test for the following sequence of two-sample problems: Let $X_1, X_2, \cdots, X_{m_i}$ and $Y_1, Y_2, \cdots, Y_{n_i}$ be ordered independent random samples from continuous cumulative distributions $F(x)$ and $G(x)$ respectively with $m_i/n_i = \tau > 0$, τ a constant. We are testing the following sequence of hypotheses:

$$H_0: F(x) = G(x)$$

$$H_{1_i}: G(x) = F(x-\theta_i), \qquad \theta_i \neq 0$$

with θ_i so chosen that $\sup_x |F(x)-F(x-\theta_i)| = \delta_i = \delta n_i^{-\frac{1}{2}}, \quad \delta > 0.$

We note that this problem is a special case of the problem considered at the beginning of Section 2.

LEMMA 1. *Let $\{T_i\}$ be a sequence of α-level Neyman tests for testing the above statistical hypotheses and $\{N_i(T)\}$, $N_i(T) = m_i+n_i$, be a sequence of sample sizes for which the sequence $\{T_i\}$ of Neyman tests have limiting power $\beta > \alpha$. Then*

$$(3.1)\qquad \lim_{i\to\infty} N_i(T)\theta_i^2 \leqq \frac{(K_{\alpha/2}-K_\beta)^2(1+\tau)^2}{\tau(E_\Psi[g'(X)])^2}\operatorname{Var}_\Psi g$$

and

$$(3.2)\qquad \lim_{i\to\infty} N_i(T)\theta_i^2 \leqq \frac{(K_{\alpha/2}-K_{(\beta-\alpha/2)})^2(1+\tau)^2}{\tau(E_\Psi[g'(X)])^2}\operatorname{Var}_\Psi g,$$

where K_p is determined by the relation $(2\pi)^{-\frac{1}{2}}\int_{K_p}^{\infty}e^{-t^2/2}\,dt$.

PROOF. Under the null hypotheses, T_i has an asymptotic distribution

$$N\left(0, N_i(T)\frac{\tau}{(1+\tau)^2}\operatorname{Var}_\Psi g\right)$$

for each sample size $N_i(T)$ there is a critical point C_i determined by $P[|T(N_i(T), 0)| > C_i] = \alpha$. Since T_i converges to a continuous symmetric distribution, we have

$$\lim_{i\to\infty} \frac{C_i}{\left[N_i(T)\frac{\tau}{(1+\tau)^2}\operatorname{Var}_\Psi g\right]^{\frac{1}{2}}} = K_{\alpha/2}. \tag{3.3}$$

Upper bound for $\lim_{i\to\infty} N_i(T)\theta_i^2$. By definition,

$$\beta = \lim_{i\to\infty} P[-C_i > T(N_i(T), \theta_i), T(N_i(T), \theta_i) > C_i],$$

hence

$$\beta \geqq \lim_{i\to\infty} P[T(N_i(T), \theta_i) > C_i]$$

$$= \lim_{i\to\infty} P\left[\frac{T(N_i(T), \theta_i) + \frac{m_i n_i}{m_i+n_i} E_\Psi(g(X)-g(Y))}{\left[N_i(T)\frac{\tau}{(1+\tau)^2}\operatorname{Var}_\Psi g\right]^{\frac{1}{2}}} > \frac{C_i + \frac{m_i n_i}{m_i+n_i} E_\Psi(g(X)-g(Y))}{\left[N_i(T)\frac{\tau}{(1+\tau)^2}\operatorname{Var}_\Psi g\right]^{\frac{1}{2}}}\right]$$

or

$$K_\beta \leqq \frac{C_i + \frac{m_i n_i}{m_i+n_i} E_\Psi(g(X)-g(Y))}{\left[N_i(T)\frac{\tau}{(1+\tau)^2}\operatorname{Var}_\Psi g\right]^{\frac{1}{2}}}$$

and by (3.3), we have

$$\lim_{i\to\infty} \frac{\frac{m_i n_i}{N_i^2(T)}(N_i(T))^{\frac{1}{2}} E_\Psi(g(X)-g(Y))}{\left[\frac{\tau}{(1+\tau)^2}\operatorname{Var}_\Psi g\right]^{\frac{1}{2}}} \geqq K_\beta - K_{\alpha/2}.$$

By the relation $m_i n_i / N_i^2(T) = \tau/(1+\tau)^2$ we have

$$\lim_{i\to\infty} [N_i(T)]^{\frac{1}{2}} E_\Psi(g(X)-g(Y)) > \frac{(K_\beta - K_{\alpha/2})(1+\tau)}{\tau^{\frac{1}{2}}} (\operatorname{Var}_\Psi g)^{\frac{1}{2}}.$$

For $Y = X+\theta_i$, $\theta_i > 0$, we have

$$\begin{aligned}\lim_{i\to\infty} [N_i(T)]^{\frac{1}{2}} E_\Psi[g(X)-g(Y)] &= \lim_{i\to\infty} [N_i(T)]^{\frac{1}{2}} E_\Psi[g(X)-g(X+\theta_i)] \\ &= -\lim_{i\to\infty} [N_i(T)]^{\frac{1}{2}} \theta_i E_\Psi[g'(X) + \tfrac{1}{2}\theta_i g''(\xi)] \qquad \xi \in [x, x+\theta_i] \\ &= -\lim_{i\to\infty} [N_i(T)]^{\frac{1}{2}} \theta_i E_\Psi[g'(X)].\end{aligned}$$

Therefore

$$\lim_{i\to\infty}[N_i(T)]^{\frac{1}{2}}\theta_i < \frac{(K_{\alpha/2}-K_\beta)(\mathrm{Var}_\Psi g)^{\frac{1}{2}}(1+\tau)}{\tau^{\frac{1}{2}}E_\Psi[g'(X)]}. \tag{3.4}$$

Now, suppose $\theta_i < 0$, we have

$$\beta \geqq \lim_{i\to\infty} P[-C_i > T(N_i(T),\theta_i)]$$

$$= \lim_{i\to\infty} P\left[\frac{T(N_i(T),\theta_i)+\frac{m_i n_i}{m_i+n_i}E_\Psi[g(X)-g(Y)]}{\left[N_i(T)\frac{\tau}{(1+\tau)^2}\mathrm{Var}_\Psi g\right]^{\frac{1}{2}}} < \frac{-C_i+\frac{m_i n_i}{m_i+n_i}E_\Psi[g(X)-g(Y)]}{\left[N_i(T)\frac{\tau}{(1+\tau)^2}\mathrm{Var}_\Psi g\right]^{\frac{1}{2}}}\right].$$

Therefore

$$K_\beta \geqq \frac{-C_i+\frac{m_i n_i}{m_i+n_i}E_\Psi[g(X)-g(Y)]}{\left[N_i(T)\frac{\tau}{(1+\tau)^2}\mathrm{Var}_\Psi g\right]^{\frac{1}{2}}},$$

but

$$-\frac{C_i}{\left[N_i(T)\frac{\tau}{(1+\tau)^2}\mathrm{Var}_\Psi g\right]^{\frac{1}{2}}} \to K_{\alpha/2} \quad \text{as} \quad i\to\infty.$$

Hence

$$-K_\beta + K_{\alpha/2} \geqq \frac{\frac{m_i n_i}{m_i+n_i}E_\Psi[g(X)-g(Y)]}{\left[N_i(T)\frac{\tau}{(1+\tau)^2}\mathrm{Var}_\Psi g\right]^{\frac{1}{2}}}$$

or

$$\lim_{i\to\infty}[N_i(T)]^{\frac{1}{2}}E_\Psi[g(X)-g(Y)] \leqq \frac{(K_{\alpha/2}-K_\beta)(1+\tau)}{\tau^{\frac{1}{2}}}(\mathrm{Var}_\Psi g)^{\frac{1}{2}},$$

since for small θ

$$E_\Psi[g(X)-g(Y)] \simeq \zeta E_\Psi[g'(X)] \quad \text{for} \quad \zeta = -\theta,$$

we have

$$\text{(3.5)}\qquad \lim_{i\to\infty}(N_i(T))^{\frac{1}{2}}E_\Psi[g(X)-g(Y)] = \lim_{i\to\infty}(N_i(T))^{\frac{1}{2}}\zeta E_\Psi[g'(X)] \leqq \frac{(K_{\alpha/2}-K_\beta)(1+\tau}{\tau^{\frac{1}{2}}}(\mathrm{Var}_\Psi g)^{\frac{1}{2}}.$$

Combining (3.4) and (3.5), we obtain the inequality (3.1).

Lower bound for $\lim_{i\to\infty} N_i(T)\theta_i^2$. Suppose $\theta_i > 0$,

$$\beta \leqq \lim_{i\to\infty} P[-(C_i-\xi_i) > T(N_i(T),\theta_i), T(N_i(T),\theta_i) > C_i] = \alpha/2+\lim_{i\to\infty} P[T(N_i(T),\theta_i) > C_i]$$

where $\xi_i = (\tau/1+\tau)E_\Psi[g(X+\theta_i)],$.

or

$$\beta-\alpha/2 \leqq \lim_{i\to\infty} P[T(N_i(T),\theta_i) > C_i] = \lim_{i\to\infty} P\left[\frac{T(N_i(T),\theta_i)+\frac{m_in_i}{m_i+n_i}E_\Psi[g(X)-g(Y)]}{\left[N_i(T)\frac{\tau}{(1+\tau)^2}\mathrm{Var}_\Psi g\right]^{\frac{1}{2}}} > \frac{C_i+\frac{m_in_i}{m_i+n_i}E_\Psi[g(X)-g(Y)]}{\left[N_i(T)\frac{\tau}{(1+\tau)^2}\mathrm{Var}_\Psi g\right]^{\frac{1}{2}}}\right],$$

after simple computation, we obtain

$$\text{(3.6)}\qquad \lim_{i\to\infty}(N_i(T))^{\frac{1}{2}}\theta_i > \frac{(K_{\alpha/2}-K_{(\beta-\alpha/2)})(1+\tau)}{[\tau E_\Psi[g'(X)]]^{\frac{1}{2}}}(\mathrm{Var}_\Psi g)^{\frac{1}{2}}.$$

Similarly, for $\theta_i < 0$, by using the relation

$$\beta \leqq \lim_{i\to\infty} P[-C_i > T(N_i(T),\theta_i), T(N_i(T),\theta_i) > (C_i+\xi_i)],$$

we obtain

$$\text{(3.7)}\qquad \lim_{i\to\infty}(N_i(T))^{\frac{1}{2}}\theta_i \geqq \frac{(K_{(\beta-\alpha/2)}-K_{\alpha/2})(1+\tau)}{\tau^{\frac{1}{2}}E_\Psi[g'(X)]}(\mathrm{Var}_\Psi g)^{\frac{1}{2}}.$$

Hence, combining (3.6) and (3.7), we have the inequality (3.2).

LEMMA 2. *Let* $\{K_i\}$ *be a sequence of* α*-level Kolmogorov–Smirnov tests for testing the problem stated at the beginning of this section, and* $\{N_i(K)\}$, $N_i(K) = m_i+n_i$, *be a sequence of sample sizes for which the sequence of tests* $\{K_i\}$ *have limiting power* $\beta > \alpha$. *Then*

$$\limsup_{i\to\infty} N_i(K)\theta_i^2 \leqq \frac{(1+\tau)^2}{\tau\psi^2(0)}\left[\tfrac{1}{2}K_{(1-\beta)}+\left(\tfrac{1}{2}\log\frac{2}{\alpha}\right)^{\frac{1}{2}}\right]^2$$

and

$$\liminf_{i\to\infty} N_i(K)\theta_i^2 \geqq \frac{(\tau+1)}{\psi^2(0)}\left[\frac{(\tau+1)^{\frac{1}{2}}-1}{\tau^{\frac{1}{2}}}\left(\tfrac{1}{2}\log\left(\frac{2}{\alpha}\right)\right)^{\frac{1}{2}} - \left(\log 2\left(\frac{1-\alpha}{\beta-\alpha}\right)\right)^{\frac{1}{2}}\right]^2.$$

PROOF. From (2.1) we have

$$\beta \geqq \limsup_{i\to\infty}\,[1-(2\pi)^{-\frac{1}{2}}\textstyle\int_{\lambda_1'}^{\infty} e^{-t^2/2}\,dt]$$

where

$$\lambda_1' = 2\left[\delta_i\left(\frac{m_i n_i}{m_i+n_i}\right)^{\frac{1}{2}} - d_{n_i}(\alpha)\right];$$

or

$$1-\beta \leqq \limsup_{i\to\infty}(2\pi)^{-\frac{1}{2}}\textstyle\int_{\lambda_1'}^{\infty} e^{-t^2/2}\,dt.$$

Hence

$$\tfrac{1}{2}K_{(1-\beta)} \geqq \limsup_{i\to\infty}\left[\delta_i\left(\frac{m_i n_i}{m_i+n_i}\right) - d_{n_i}(\alpha)\right].$$

Since $d_i(\alpha) \to (\frac{1}{2}\log(2/\alpha))^{\frac{1}{2}}$, we have

$$\tfrac{1}{2}K_{(1-\beta)} + (\tfrac{1}{2}\log(2/\alpha))^{\frac{1}{2}} \geqq \limsup_{i\to\infty}\delta_i\left(\frac{m_i n_i}{m_i+n_i}\right)^{\frac{1}{2}};$$

since

$$\left(\frac{m_i n_i}{m_i+n_i}\right)^{\frac{1}{2}} = (N_i(K))^{\frac{1}{2}}\frac{\tau^{\frac{1}{2}}}{\tau+1}$$

and by Kalish and Mikulski (1971)

$$\limsup_{i\to\infty}\delta_i^2 N_i(K) = \psi^2(0)\limsup_{i\to\infty} N_i(K)\theta_i^2,$$

we have the first part of the lemma.

By virtue of (2.3), we have

$$\beta \leqq \alpha + 2(1-\alpha)\liminf_{i\to\infty}\exp\left[-2\left(\frac{(\tau+1)^{\frac{1}{2}}-1}{\tau^{\frac{1}{2}}}(\tfrac{1}{2}\log(2/\alpha))^{\frac{1}{2}} - n_i^{\frac{1}{2}}\delta_i\right)^2\right]$$

and the second assertion of the lemma can be obtained after a simple calculation.

Now, combining the results of Kalish and Mikulski (1971), Mikulski and Yu (1971), and Lemmas 1 and 2, we have the following:

THEOREM 1. *If the assumptions of Lemmas* 1 *and* 2 *are satisfied then*

(a) *for any distribution* F *and positive constant* C, *there is a distribution* $\Psi \in \mathscr{C}$ *such that* $e_{KT}(F;\Psi) \geqq C$;

(b) *if* $g(x)$ *is bounded, then* $\inf_{\Psi\in\mathscr{C}}\bar{e}_{KT}(F;\Psi) = 0$;

(c) *if* $g'(x)$ *satisfies the condition* $0 < \varepsilon \leqq g'(x) \leqq L$ *for all* x. *Then* $\inf_{\Psi\in\mathscr{C}}\bar{e}_{KT}(F;\Psi) > 0$.

An important special case, when $\Psi = F$, of the above theorem has been given in Capon (1965).

It has been shown in Mikulski (1963) that

$$e_{T^*T}(F;\Psi) = \left[\frac{\int_{-\infty}^{\infty} J'[\psi(x)]\Psi^2(x)\,dx}{E_\Psi[g'(X)]}\right]^2 \frac{\mathrm{Var}_\Psi g(x)}{\mathrm{Var}_F g(x)},$$

where the function $J(u)$ is as defined in Chernoff and Savage (1958), and $J'(u) = dJ(u)/du$. Incorporate this and results obtained in Kalish and Mikulski (1971) and Mikulski and Yu (1971), and we have the following theorem:

THEOREM 2. *If the assumption of Lemmas* 1 *and* 2 *are satisfied, then*

(a) *for any distribution F and any constant C there is a distribution $\Psi \in \mathscr{C}$ such that $\underline{e}_{KT^*}(F;\Psi) \geqq C$,*

(b) *if $g(x)$ is bounded, then* $\inf_{\Psi\in\mathscr{C}} \underline{e}_{KT^*}(F;\Psi) > 0$;

(c) *if $g(x)$ is unbounded, then* $\inf_{\Psi\in\mathscr{C}} \bar{e}_{KT^*}(F;\Psi) = 0$.

Acknowledgment. The author wishes to thank Professor P. W. Mikulski for his careful reading of the manuscript, and valuable comments.

REFERENCES

CAPON, J. (1965). On the asymptotic efficiency of the Kolmogorov–Smirnov test. *J. Amer. Statist. Assoc.* **60** 843–853.

CHERNOFF, H. and SAVAGE, I. R. (1958). Asymptotic normality and efficiency of certain nonparametric test statistics. *Ann. Math. Statist.* **29** 972–994.

FRASER, D. A. S. (1957). *Nonparametric Methods in Statistics.* Wiley, New York.

HODGES, J. L. Jr. (1957). The significance probability of the Smirnov two-sample test. *Ark. Mat.* **43** 469–486.

KALISH, G. and MIKULSKI, P. W. (1971). The asymptotic behavior of the Smirnov test compared to standard optimal procedures. *Ann. Math. Statist.* **42**

MASSEY, F. J. Jr. (1950). A note on the power of a non-parametric procedures in the two sample case. *Ann. Math. Statist.* **20** 440–443.

MIKULSKI, P. W. (1963). On the efficiency of optimal non-parametric procedures in the two sample case. *Ann. Math. Statist.* **34** 22–32.

MIKULSKI, P. W. and YU, C. S. (1971). On upper bounds of asymptotic power and Pitman efficiencies of Kolmogorov and Smirnov tests. Submitted to *Ann. Math. Statist.*

NEYMAN, J. (1959). Optimal asymptotic tests of composite statistical hypothesis. *Probability and Statistics, The Harald Cramer Volume.* Almqvist and Wiksell, Stockholm, 213–234.

QUADE, D. (1965). On the asymptotic power of the one-sample Kolmogorov–Smirnov tests. *Ann. Math. Statist.* **36** 1000–1018.