The Annals of Mathematical Statistics
1970, Vol. 41, No. 5, 1646–1654

BOUNDS FOR THE POWER OF LIKELIHOOD RATIO TESTS AND THEIR ASYMPTOTIC PROPERTIES[1]

BY OLAF KRAFFT AND DETLEF PLACHKY

University of Münster

0. Introduction and summary. Let P_0 and P_1 be two probability distributions on a measurable space $(\mathscr{X}, \mathscr{B})$. We consider the problem of testing the simple hypothesis $H: P = P_0$ against the simple alternative $K: P = P_1$ on the basis of n independent random variables $X_1, X_2, \cdots, X_n$ with common distribution P. The method that is widely used in the literature for investigating large sample properties of optimal tests is the following: according to the Neyman–Pearson lemma an optimal test can be described by means of sums of independent random variables. Theorems from the theory of probabilities of large deviations or ergodic theorems then allow one to obtain results on the asymptotic behavior of the error probabilities of optimal tests. In this paper we use a representation of the power of an optimal test which has its origin in the duality theory of infinite linear programming, in order to derive upper and lower bounds for the power. The bounds hold for any sample size n. This is done in Section 1 for two different types of tests, namely for most powerful tests at level α_n and for tests which minimize a weighted sum of the error probabilities. It turns out that those bounds allow one to derive the well-known asymptotic properties under slightly more general conditions, i.e. α_n is permitted to tend to zero faster than any negative power of n (but not exponentially fast) and the weight λ_n to tend exponentially fast to zero. This and another application are discussed in Section 2.

1. Bounds for the error probabilities. Let P_0 and P_1 be two probability distributions on a measurable space $(\mathscr{X}, \mathscr{B})$ and f_0 and f_1 their densities with respect to a σ-finite measure μ. We consider the problem of testing the simple hypothesis $H: P = P_0$ against the simple alternative $K: P = P_1$ on the basis of n independent random variables $X_1, X_2, \cdots, X_n$ with common distribution P. By $\varphi_{\alpha,n}$ we denote a most powerful test for testing H against K at level α, $0 < \alpha < 1$. Since the customary procedure of maximizing the power with respect to all tests of size α does not seem to be very appropriate when the sample size approaches infinity, Chernoff [2], [3], proposed to consider an unconstrained version of that extremal problem, namely to choose under all tests φ one that minimizes $E_1(1-\varphi)+\lambda E_0\varphi$. Here λ is a positive constant (independent of P_0 and P_1, but in our approach not necessarily independent of n). Such a test is denoted by $\psi_{\lambda,n}$ and called a W-test at weight λ. This kind of test exhibits some nice properties, two of them are recorded below, cf. Krafft [4]: ψ_λ is a W-test if and only if

$$(1) \qquad \psi_\lambda(x) = 1, \quad f_1(x) > \lambda f_0(x), \qquad = 0, \quad f_1(x) < \lambda f_0(x); \qquad \mu\text{-a.s.};$$

Received July 29, 1969; revised April 2, 1970.

[1] Research supported by the Deutsche Forschungsgemeinschaft.

hence ψ_λ is a likelihood ratio test for which one does not have to calculate a quantile in order to determine an optimal test as in using the optimality criterion of Neyman and Pearson. The test ψ_λ can be chosen arbitrarily on the set $\{x: f_1(x) = \lambda f_0(x)\}$; hence there always exists a non-randomized W-test.

The problem of determining a most powerful test at level α can be regarded as a problem of infinite programming, cf. Krafft, Witting [5], to which corresponds a dual problem, namely to determine a number $k \geqq 0$ such that

$$\alpha k + \int (f_1(x) - k f_0(x))^+ \, d\mu = \inf. \tag{2}$$

Similarly, cf. Krafft [4] page 551, to the problem of determining a W-test ψ_λ there corresponds the problem of determining a number u, $0 \leqq u \leqq 1$, such that

$$u + \lambda - \int \max(u f_1(x), \lambda f_0(x)) \, d\mu = \sup. \tag{3}$$

(2) and (3) can obviously be extended to the n observations version by replacing $f_1(x)$ and $f_0(x)$ by $f_1(\mathbf{x}) = \prod_{i=1}^n f_1(x_i)$ and $f_0(\mathbf{x}) = \prod_{i=1}^n f_0(x_i)$, respectively, and μ by the product measure $\mu^{(n)}$. The relations between the primal problems and the dual problems are given in the following lemma, the proof of which is in this form due to the referee.

LEMMA 1. *Let $\varphi_{\alpha,n}$ be a most powerful test at level α, $0 < \alpha < 1$, and $\psi_{\lambda,n}$ be a W-test at weight λ, $\lambda > 0$. Then*

$$E_1 \varphi_{\alpha,n} = \inf_{k_n \geqq 0} \{\alpha k_n + \int (f_1(\mathbf{x}) - k_n f_0(\mathbf{x}))^+ \, d\mu^{(n)}\}, \tag{4}$$

$$E_1(1 - \psi_{\lambda,n}) + \lambda E_0 \psi_{\lambda,n} = \sup_{0 \leqq u_n \leqq 1} \{u_n + \lambda - \int \max(u_n f_1(\mathbf{x}), \lambda f_0(\mathbf{x})) \, d\mu^{(n)}\}. \tag{5}$$

PROOF. Let $\psi_{k_n,n}$ be a W-test at weight $k_n > 0$. It follows from (1) and from the definition of $\psi_{k_n,n}$ that

$$\begin{aligned} 1 - \int (f_1(\mathbf{x}) - k_n f_0(\mathbf{x}))^+ \, d\mu^{(n)} &= E_1(1 - \psi_{k_n,n}) + k_n E_0 \psi_{k_n,n} \\ &\leqq E_1(1 - \varphi_{\alpha,n}) + k_n E_0 \varphi_{\alpha,n} \leqq E_1(1 - \varphi_{\alpha,n}) + k_n \alpha. \end{aligned}$$

Hence $E_1 \varphi_{\alpha,n} \leqq \alpha k_n + \int (f_1(\mathbf{x}) - k_n f_0(\mathbf{x}))^+ \, d\mu^{(n)}$. This inequality holds for $k_n = 0$, too.

Equation (4) follows from the fact that, according to the Neyman–Pearson lemma, there exists a number $k_n^* \geqq 0$ such that $\varphi_{\alpha,n}$ is a W-test at weight k_n^* if $k_n^* > 0$ and $E_1 \varphi_{\alpha,n} = 1$ if $k_n^* = 0$. In order to prove (5) let $\lambda_n = \lambda / u_n$ and $\psi_{\lambda_n,n}$ be a W-test at weight λ_n, $0 < u_n \leqq 1$. Then

$$\begin{aligned} u_n + \lambda - \int \max(u_n f_1(\mathbf{x}), \lambda f_0(\mathbf{x})) \, d\mu^{(n)} &= u_n E_1(1 - \psi_{\lambda_n,n}) + \lambda E_0 \psi_{\lambda_n,n} \\ &\leqq u_n E_1(1 - \psi_{\lambda,n}) + \lambda E_0 \psi_{\lambda,n} \leqq E_1(1 - \psi_{\lambda,n}) + \lambda E_0 \psi_{\lambda,n} \end{aligned}$$

with equality in both inequalities for $u_n = 1$. The case $u_n = 0$ is trivial.

For any two probability densities (w.r.t. μ) $g(x)$ and $h(x)$ let the functions $\rho_{g,h}(t)$ and $I(g:h)$ be defined as

$$\rho_{g,h}(t) = \int g^t(x) h^{1-t}(x) \, d\mu \quad \text{and}$$

$$I(g:h) = \int g(x) \ln(g(x)/h(x)) \, d\mu.$$

Here for $q = 0$ the expressions $p^t q^{1-t} (t > 1)$ and $p \ln (p/q)$ are, respectively, defined to be 0 or ∞ as p is equal to zero or greater than zero. Note that $0 \leqq \rho_{g,h}(t) \leqq 1$ for $0 \leqq t \leqq 1$, that $\int g^t(\mathbf{x}) h^{1-t}(\mathbf{x})\, d\mu^{(n)} = \rho^n_{g,h}(t)$ and that $\rho_{g,h}(t) < \infty$ for some $t > 1$ implies $\int_{\{x:h(x)=0\}} g(x)\, d\mu = 0$. For $\rho_{f_0,f_1}(t)$ we shortly write $\rho(t)$. The quantity $(t-1)^{-1} \ln \rho(t)$, $t \neq 1$, is usually called information measure of order t and denoted by $I_t(f_0 : f_1)$. For $t = 1$ $I_t(f_0 : f_1)$ is defined to be equal to $I(f_0 : f_1)$. For further properties of these notions which are related to the subject of this paper cf. Vajda [9]. Lemma 1 enables us to derive upper and lower bounds for $E_1 \varphi_{\alpha,n}$ and $E_1(1-\psi_{\lambda,n}) + \lambda E_0 \psi_{\lambda,n}$ in terms of $I_t(f_0 : f_1)$.

THEOREM 1. *For the power $E_1 \varphi_{\alpha,n}$ of the most powerful test at level α, $0 < \alpha < 1$, and any fixed sample size n it holds that*

$$(6) \qquad E_1 \varphi_{\alpha,n} \leqq 1 - (1-\alpha)^{t/(t-1)} \exp[-n I_t(f_0 : f_1)], \qquad t > 1,$$

$$(7) \qquad E_1 \varphi_{\alpha,n} \geqq 1 - (1-t) t^{t/(1-t)} \alpha^{t/(t-1)} \exp[-n I_t(f_0 : f_1)], \qquad 0 < t < 1.$$

PROOF. If $\rho(t) = \infty$ for $t > 1$, inequality (6) is trivially satisfied. Assume $\rho(t) < \infty$ for some $t > 1$. In that case we have for $k_n \geqq 0$

$$(8) \qquad \int (f_1(\mathbf{x}) - k_n f_0(\mathbf{x}))^+ \, d\mu^{(n)} = \int f_1(\mathbf{x}) \max [k_n f_0(\mathbf{x})/f_1(\mathbf{x}), 1]\, d\mu^{(n)} - k_n.$$

The relation

$$(9) \qquad \max(z, 1) \leqq a z^t + 1, \qquad t > 1,\ z \geqq 0,\ a = (t-1)^{t-1} t^{-t},$$

(8) and Lemma 1 then imply

$$(10) \qquad E_1 \varphi_{\alpha,n} \leqq \inf_{k_n \geqq 0} \{1 + (\alpha - 1) k_n + a \rho^n(t) k_n^{\,t}\}.$$

The right-hand side of (10) attains its minimum at $k_n^* = [(1-\alpha)/a t \rho^n(t)]^{1/(t-1)}$. Inserting k_n^* into (10) we obtain (6). If $\rho(t) = 0$ for some t, $0 < t < 1$, P_0 and P_1 are orthogonal. Hence $E_1 \varphi_{\alpha,n} = 1$. Assume $\rho(t) > 0$ for $0 < t < 1$. Since for $k_n \geqq 0$

$$(11) \qquad \int (f_1(\mathbf{x}) - k_n f_0(\mathbf{x}))^+ \, d\mu^{(n)} = 1 - \int f_1(\mathbf{x}) \min [k_n f_0(\mathbf{x})/f_1(\mathbf{x}), 1]\, d\mu^{(n)},$$

$$(12) \qquad \min(z, 1) \leqq z^t, \qquad 0 < t < 1,\ z \geqq 0,$$

we get from Lemma 1 that

$$(13) \qquad E_1 \varphi_{\alpha,n} \geqq \inf_{k_n \geqq 0} \{\alpha k_n + 1 - k_n^{\,t} \rho^n(t)\}.$$

The right-hand side of (13) attains its minimum at $k_n^* = [\alpha / t \rho^n(t)]^{1/(1-t)}$. Inserting k_n^* into (13) we obtain (7).

REMARK. For all (arbitrary) tests (6) can also be proved by means of Hölder's inequality. The following example shows that inequalities (6) and (7) are sharp. Let P_0 be the uniform distribution over [0, 1], P_1 be the uniform distribution over $[0, \vartheta]$, $\vartheta > 1$, and μ be the Lebesgue measure. In this case we have $E_1(1-\varphi_{\alpha,n}) = (1-\alpha)/\vartheta^n$, which can be proved e.g. with the help of Lemma 1, and $I_t(f_0 : f_1) = \ln \vartheta$. Then for $t \to \infty$ and $t = \alpha$ equality holds in (6) and (7), respectively.

THEOREM 2. *Let $\psi_{\lambda,n}$ be a W-test at weight $\lambda > 0$. Then for any fixed sample size $n \geqq 1$ and $0 < t < 1$ it holds that*

$$E_1(1-\psi_{\lambda,n})+\lambda E_0\psi_{\lambda,n} \leqq \lambda^t \exp[n(t-1)I_t(f_0:f_1)], \tag{14}$$

and for $n \geqq I_t^{-1}(f_1:f_0)\ln t\lambda/(t-1)$, $t > 1$, it holds that

$$E_1(1-\psi_{\lambda,n})+\lambda E_0\psi_{\lambda,n} \geqq \lambda \exp[-nI_t(f_1:f_0)]. \tag{15}$$

PROOF. Since $E_1(1-\psi_{\lambda,n})+\lambda E_0\psi_{\lambda,n} = \int f_1(\mathbf{x})\min(\lambda f_0(\mathbf{x})/f_1(\mathbf{x}), 1)\,d\mu^{(n)}$, (14) follows from (11) and (12). Essentially the same argument as used in the proof of Theorem 1 applied to equation (5) yields inequality (15). The condition $u_n \leqq 1$ in (5) has as consequence the restriction on n in (15).

COROLLARY 1. *Let φ_n be a likelihood ratio test for testing $H: P = P_0$ against $K: P = P_1$ on the basis of n independent random variables $X_1, X_2, \cdots, X_n$ with common distribution P. Then for $\gamma_n(t) = t^t(1-t)^{1-t}\rho^n(t)$ it holds that*

$$E_0\varphi_n \leqq \inf_{0<t<1}\gamma_n(t) \quad or \quad E_1(1-\varphi_n) \leqq \inf_{0<t<1}\gamma_n(t).$$

Moreover, there always exists a likelihood ratio test φ_n^* such that

$$E_0\varphi_n^* \leqq \inf_{0<t<1}\gamma_n(t) \quad and \quad E_1(1-\varphi_n^*) \leqq \inf_{0<t<1}\gamma_n(t).$$

PROOF. Assume that $E_0\varphi_n > \inf_{0<t<1}\gamma_n(t)$ and $E_1(1-\varphi_n) > \inf_{0<t<1}\gamma_n(t)$. Then for some τ, $0 < \tau < 1$, we have $E_0\varphi_n > \gamma_n(\tau)$ and $E_1(1-\varphi_n) > \gamma_n(\tau)$. Let $\tilde{\varphi}_{\alpha_n,n}$ be a most powerful test at level $\alpha_n = \gamma_n(\tau)$. Since $E_0\tilde{\varphi}_{\alpha_n,n} < E_0\varphi_n$ and φ_n is most powerful at level $E_0\varphi_n$, we get a contradiction from $E_1(1-\varphi_n) \leqq E_1(1-\tilde{\varphi}_{\alpha_n,n}) \leqq \gamma_n(\tau)$. Here the last inequality is a consequence of (7). In order to prove the second assertion, let $\{t_m\}$ be a sequence such that $\lim_{m\to\infty}\gamma_n(t_m) = \inf_{0<t<1}\gamma_n(t)$ and $\{\varphi_{n,m}^*\}_m$ a sequence of most powerful tests at level $\gamma_n(t_m)$. According to the weak compactness theorem for test functions (a proof for the case that $\mathscr{B}$ is not necessarily separable is given in [6]) and to (7) there exists a test φ_n^* such that $E_0\varphi_n^* \leqq \inf_{0<t<1}\gamma_n(t)$ and $E_1(1-\varphi_n^*) \leqq \inf_{0<t<1}\gamma_n(t)$. The most powerful test at level $\inf_{0<t<1}\gamma_n(t)$ then fulfills the second assertion.

REMARK. The uniform distribution example in the remark of Theorem 1 shows that the bounds of Corollary 1 cannot be improved.

2. Applications. In this section two modes of applications, namely large sample properties of tests and a large deviation theorem, shall illustrate the use of the bounds derived in Section 1.

COROLLARY 2. *If $I_t(f_0:f_1) < \infty$ for some $t > 1$ and $\{\alpha_n\}$ is such that $\lim_{n\to\infty}\alpha_n^{1/n} = \lim_{n\to\infty}(1-\alpha_n)^{1/n} = 1$, then*

$$\lim_{n\to\infty}[E_1(1-\varphi_{\alpha_n,n})]^{1/n} = \exp[-I(f_0:f_1)] = \lim_{n\to\infty} c_n^{1/n},$$

where c_n is the critical value of $\varphi_{\alpha_n,n}$ as it is given by the Neyman–Pearson lemma.

PROOF. The left-hand equation is an immediate consequence of (6) and (7), since $\lim_{t\to 1-} I_t(f_0:f_1) = \lim_{t\to 1+} I_t(f_0:f_1) = I(f_0:f_1)$. According to (14) it holds that

$$E_1(1-\varphi_{\alpha_n,n}) \leqq c_n{}^t\rho^n(t) \quad \text{and} \quad E_0\,\varphi_{\alpha_n,n} \leqq c_n{}^{t-1}\rho^n(t), \qquad 0<t<1.$$

This implies

$$c_n{}^{1/n} \geqq [E_1(1-\varphi_{\alpha_n,n})]^{1/n} \quad \text{and} \quad c_n{}^{1/n} \leqq [\alpha_n{}^{1/n}]^{1/(t-1)}[\rho(t)]^{1/(t-1)}$$

which has as consequence the right-hand equation.

REMARK. Again the uniform distribution example can be used to show that the assertion of Corollary 2 is not true if $\lim_{n\to\infty}(1-\alpha_n)^{1/n} < 1$.

The next theorem shows that the assumption on $I_t(f_0:f_1)$ in Corollary 2 can be discarded, if we impose a slightly stronger condition on the sequence $\{\alpha_n\}$. This will be proved by using the conjugate distributions of $Y = \ln(f_0(X)/f_1(X))$ under P_1—a technique which is by now well known in large deviations theory and which was employed by Wald [10], Section 3.4, in sequential analysis. To this end let $g_t(x)$ for all t with $0<\rho(t)<\infty$ be defined as $g_t(x) = f_0{}^t(x)f_1{}^{1-t}(x)/\rho(t)$. By P_t we denote the distribution corresponding to $g_t(x)$ and by E_{g_t} the expectation operator with respect to g_t.

THEOREM 3. *If* $\{\alpha_n\}$ *is such that* $\lim_{n\to\infty}\alpha_n{}^{1/n} = 1$, *then*

$$\limsup_{n\to\infty}[E_1(1-\varphi_{\alpha_n,n})]^{1/n} \leqq \exp[-I(f_0:f_1)]. \tag{16}$$

If, moreover, $\{\alpha_n\}$ *is bounded away from unity, then*

$$\lim_{n\to\infty}[E_1(1-\varphi_{\alpha_n,n})]^{1/n} = \exp[-I(f_0:f_1)], \quad \textit{and} \tag{17}$$

$$\lim_{n\to\infty} c_n{}^{1/n} = \exp[-I(f_0:f_1)], \tag{18}$$

where c_n *is the critical value of* $\varphi_{\alpha_n,n}$ *as it is given by the Neyman–Pearson lemma.*

PROOF. Since $\lim_{t\to 1-} I_t(f_0:f_1) = I(f_0:f_1)$, (16) follows from (7). We now prove (18). If $I(f_0:f_1) = \infty$ and if there is an n_0 such that $c_n = 0$ for $n \geqq n_0$, then (18) is trivially true. Let $\{c_m\}$ be a subsequence of $\{c_n\}$ such that $c_m > 0$ for all m. By the Neyman–Pearson lemma $\varphi_{\alpha_m,m}$ is a W-test at weight $\lambda_m = c_m$. Neglecting the term $E_1(1-\varphi_{\alpha_m,m})$ we get from (14) for $0<t<1$ $\limsup_{m\to\infty} c_m{}^{1/m} \leqq (\lim \alpha_m{}^{1/m})^{1/(t-1)}\exp[-I_t(f_0:f_1)]$, so that

$$\limsup_{m\to\infty} c_m{}^{1/m} \leqq \exp[-I(f_0:f_1)]. \tag{19}$$

If $I(f_0:f_1) = \infty$, then (18) follows from (19), since $c_n \geqq 0$ for all n. Assume that $I(f_0:f_1) < \infty$. In that case we have $\int_{\{x:f_1(x)=0\}} f_0(x)\,d\mu = 0$, so that $0<\eta \leqq 1-\alpha_n$ implies $c_n > 0$ and, moreover, $0<\eta\leqq 1-\alpha_n \leqq P_0(n^{-1}\sum_{i=1}^n Y_i - I(f_0:f_1) \geqq -n^{-1}\ln c_n - I(f_0:f_1))$, where $Y_i = \ln(f_0(X_i)/f_1(X_i))$. Note that $E_0 Y_i = I(f_0:f_1)$. Applying the weak law of large numbers, from the last inequality it is seen that to $\gamma>0$ there does not exist a subsequence $\{c_m\}\subset\{c_n\}$ such that $-m^{-1}\ln c_m - I(f_0:f_1) \geqq \gamma$ for all m. Hence we get $\liminf_{n\to\infty} c_n{}^{1/n} \geqq \exp[-I(f_0:f_1)]$. This together with (19) implies (18).

If $I(f_0:f_1) = \infty$, then (17) is satisfied because of (16) and $E_1(1-\varphi_{\alpha_n,n}) \geqq 0$. Assume $I(f_0:f_1) < \infty$. In that case $\int_{\{x:f_1(x)=0\}} f_0(x)\,d\mu = 0$, so that $\rho(t)$ is positive. Hence $g_t(x)$ is well defined for $0 < t < 1$. Since $g_{t'}(x)/g_t(x) = (f_1(x)/f_0(x))^{t-t'}\rho(t)/\rho(t')$, the class $\{g_\vartheta(x)\}$, $\vartheta = 1-t$, has monotone likelihood ratio in $T(x) = f_1(x)/f_0(x)$. Therefore $\varphi_{\alpha_n,n}$ is a most powerful test for testing $H: P = P_t$, $0 < t < 1$, against $K: P = P_1$ at level $E_{g_t}\varphi_{\alpha_n,n}$, so that from (6) we get for $\tau > 1$

$$E_1(1-\varphi_{\alpha_n,n}) \geqq [E_{g_t}(1-\varphi_{\alpha_n,n})]^{\tau/(\tau-1)} \exp[-nI_\tau(g_t:f_1)].$$

Observing that $I_{1/t}(g_t:f_1) = I_t(f_0:f_1)$ and putting $\tau = 1/t$ we therefore have

$$E_1(1-\varphi_{\alpha_n,n}) \geqq [E_{g_t}(1-\varphi_{\alpha_n,n})]^{1/(1-t)} \exp[-nI_t(f_0:f_1)]. \tag{20}$$

For $b_n = e^{-n\varepsilon}$, $\varepsilon > 0$, and $B_n = \{\mathbf{x}: b_n c_n f_0(\mathbf{x}) \leqq f_1(\mathbf{x})\}$ we have

$$\begin{aligned} E_{g_t}(1-\varphi_{\alpha_n,n}) &\geqq \rho^{-n}(t)\int_{B_n}(1-\varphi_{\alpha_n,n}) f_0{}^t(\mathbf{x}) f_1{}^{1-t}(\mathbf{x})\,d\mu^{(n)} \\ &\geqq (b_n c_n)^{1-t}\rho^{-n}(t)\int_{B_n}(1-\varphi_{\alpha_n,n}) f_0(\mathbf{x})\,d\mu^{(n)} \\ &\geqq (b_n c_n)^{1-t}\rho^{-n}(t)[P_0(B_n)-\alpha_n] \\ &\geqq (b_n c_n)^{1-t}\rho^{-n}(t)[P_0(B_n)+\eta-1]. \end{aligned}$$

Since $P_0(B_n) = P_0(n^{-1}\sum_{i=1}^n Y_i - I(f_0:f_1) \leqq \varepsilon - n^{-1}\ln c_n - I(f_0:f_1))$, according to (18) and the weak law of large numbers $P_0(B_n)$ will be greater than $1-\eta/2$ for sufficiently large n. From (20) we then obtain that

$$\liminf_{n\to\infty}[E_1(1-\varphi_{\alpha_n,n})]^{1/n} \geqq \rho^{1/(t-1)}(t)\exp[-I_t(f_0:f_1)]\lim_{n\to\infty}(b_n c_n)^{1/n},$$

so that, finally, since ε can be chosen arbitrarily small, (18) implies

$$\liminf_{n\to\infty}[E_1(1-\varphi_{\alpha_n,n})]^{1/n} \geqq \exp[-I(f_0:f_1)]. \tag{21}$$

From (21) and (16) we get (17).

REMARK. (16) was obtained by Rao [8] page 381, under the condition that $\lim_{n\to\infty}\alpha_n = \alpha$, $0 < \alpha < 1$, and (17) obtained by Chernoff [3] (with reference to Ch. Stein) under the condition $\alpha_n \equiv \alpha$. Rao [7] then proved that (17) holds if $\lim_{n\to\infty}\alpha_n = \alpha$, $0 < \alpha < 1$. Relation (18) is contained also in that last paper (with reference to Basu).

The next two results deal with W-tests. Although inequalities (14) and (15) are a certain analogue to inequalities (6) and (7), they do not yield the direct proof of an analogue to Theorem 3, since (15) is somewhat weaker than (6). Nevertheless, a technique similar to that used in the proof of Theorem 3 will enable us to transform the problem in such a way that we can apply (6) instead of (15).

THEOREM 4. *Let* $\{\lambda_n\}$ *be a sequence of positive numbers such that* $\lim_{n\to\infty}\lambda_n{}^{1/n} = \lambda > 0$ *and* t^* *with* $\lambda^{t^*}\rho(t^*) = \inf \lambda^t\rho(t)$ *be an interior point of the interval* $\{t: \rho(t) < \infty\}$. *If the set* $B = \{x: f_0(x)f_1(x) > 0\}$ *has positive* μ-*measure and if* $f_1(x)/f_0(x)$ *is not* *equal to* λ μ-a.s. *on* B, *then for* $\psi_{\lambda_n,n}$ *it holds that*

(22) $$\lim_{n\to\infty}[E_{g_t}(1-\psi_{\lambda_n,n})]^{1/n} = 1, \qquad t \geqq t^*,$$
$$= \lambda^{t^*}\rho(t^*)/\lambda^t\rho(t), \qquad t < t^*.$$

(23) $$\lim_{n\to\infty}[E_{g_t}\psi_{\lambda_n,n}]^{1/n} = 1, \qquad t \leqq t^*,$$
$$= \lambda^{t^*}\rho(t^*)/\lambda^t\rho(t), \qquad t > t^*.$$

PROOF. g_t is well defined on $\{t:\rho(t)<\infty\}$, since $\mu(B)>0$ implies $\rho(t)>0$. In the interior of $\{t:\rho(t)<\infty\}$ the second derivative $H''(t)$ of $H(t)=\lambda^t\rho(t)$ can be written as

$$H''(t) = \lambda^t \int f_0^{\,t}(x) f_1^{\,1-t}(x)\left[\ln\frac{\lambda f_0(x)}{f_1(x)}\right]^2 d\mu,$$

so that it follows from the assumption that t^* is the unique minimum of $H(t)$. Let $A(s,t)$ for $t<s$ be defined as $A(s,t)=\lambda_n^{\,s-t}[\rho(s)/\rho(t)]^n$. Then $\{\mathbf{x}: f_1(\mathbf{x})>\lambda_n f_0(\mathbf{x})\} = \{\mathbf{x}: g_t(\mathbf{x})>A(s,t)g_s(\mathbf{x})\}$, so that according to (14) for all τ, $0<\tau<1$,

$$E_{g_t}(1-\psi_{\lambda_n,n}) + A(s,t)E_{g_s}\psi_{\lambda_n,n} \leqq A^\tau(s,t)\rho^n_{g_s,g_t}(\tau).$$

The last inequality implies that $E_{g_s}(1-\psi_{\lambda_n,n}) \geqq 1-A^{\tau-1}(s,t)\rho^n_{g_s,g_t}(\tau)$ and $E_{g_t}(1-\psi_{\lambda_n,n}) \leqq A^\tau(s,t)\rho^n_{g_s,g_t}(\tau)$. Since $H(t^*)<H(t)$ for $t>t^*$ and $\lim_{n\to\infty}\lambda_n^{\,1/n}=\lambda$, we therefore have

(24) $$\liminf_{n\to\infty}[E_{g_s}(1-\psi_{\lambda_n,n})]^{1/n} \geqq 1, \qquad s>t^*,$$

(25) $$\limsup_{n\to\infty}[E_{g_t}(1-\psi_{\lambda_n,n})]^{1/n} \leqq H(t^*)/H(t), \qquad t<t^*.$$

Let $\varphi_{\alpha_n,n}$ be a most powerful test for testing $H:P=P_s$ against $K:P=P_t$ at level $\alpha_n = E_{g_s}\psi_{\lambda_n,n}$. According to (6) and the definition of $\varphi_{\alpha_n,n}$ it holds that for $\tau>1$

$$E_{g_t}(1-\psi_{\lambda_n,n}) = E_{g_t}(1-\varphi_{\alpha_n,n}) \geqq (1-E_{g_s}\psi_{\lambda_n,n})^{\tau/(\tau-1)}[\rho_{g_s,g_t}(\tau)]^{n/(1-\tau)}.$$

From (24) and letting τ tend to one we therefore get for all $s>t^*$ $\liminf_{n\to\infty}[E_{g_t}(1-\psi_{\lambda_n,n})]^{1/n} \geqq \exp[-I(g_s:g_t)]$. Since

(26) $$\exp[-I(g_s:g_t)] = \frac{\rho(s)}{\rho(t)}\exp\left[(t-s)\rho^{-1}(s)\int\left(\ln\frac{f_0(x)}{f_1(x)}\right)f_0^{\,s}(x)f_1^{\,1-s}(x)\,d\mu\right]$$

and t^* is an interior point of the interval $\{t:\rho(t)<\infty\}$ the exponential term on the right-hand side of (26) tends to λ^{t^*}/λ^t as s tends to t^*, so that for all t

$$\liminf_{n\to\infty}[E_{g_t}(1-\psi_{\lambda_n,n})]^{1/n} \geqq H(t^*)/H(t).$$

This result together with (24) and (25) implies (22). (23) follows from (22), since $\tilde{\psi}_{\lambda_n,n} = 1-\psi_{\lambda_n,n}$ is a W-test for testing $H:P=P_1$ against $K:P=P_0$ at weight λ_n^{-1}.

COROLLARY 3. *Under the assumptions of Theorem* 4 *it holds that*

$$\lim_{n\to\infty}[E_1(1-\psi_{\lambda_n,n})+\lambda_n E_0\psi_{\lambda_n,n}]^{1/n} = \inf_{0\leqq t\leqq 1}\lambda^t\rho(t).$$

PROOF. According to (14) we have

$$\limsup_{n\to\infty}[E_1(1-\psi_{\lambda_n,n})+\lambda_n E_0\psi_{\lambda_n,n}]^{1/n} \leqq \inf_{0\leqq t\leqq 1}\lambda^t\rho(t).$$

Since $E_1(1-\psi_{\lambda_n,n})+\lambda_n E_0\psi_{\lambda_n,n} \geqq \max[E_1(1-\psi_{\lambda_n,n}), \lambda_n E_0\psi_{\lambda_n,n}]$, the assertion follows from (22) and (23).

REMARK. Putting $\lambda_n \equiv \lambda_0 > 0$ we obtain Chernoff's result [3] page 16, as a special case of Corollary 3. A further consequence of Theorem 4 is

COROLLARY 4. *If* $\varphi_{\alpha_n,n}$ *is most powerful at level* α_n,

$$\lim_{n\to\infty}\alpha_n^{1/n} = \lim_{n\to\infty}(1-\alpha_n)^{1/n} = 1 \quad \textit{and} \quad I_t(f_0:f_1) < \infty$$

for some $t > 1$, *then*

$$\begin{aligned}\lim_{n\to\infty}[E_{g_t}(1-\varphi_{\alpha_n,n})]^{1/n} &= 1, & t \geqq 1\\ &= \exp[-I(f_0:g_t)], & t < 1;\\ \lim_{n\to\infty}[E_{g_t}\varphi_{\alpha_n,n}]^{1/n} &= 1, & t \leqq 1\\ &= \exp[-I(f_0:g_t)], & t > 1.\end{aligned}$$

PROOF. Since $I_t(f_0:f_1) < \infty$ for some $t > 1$ implies $\int_{\{x:f_1(x)=0\}}f_0(x)\,d\mu = 0$, $\rho(t)$ is positive and, consequently, g_t is well defined. $\varphi_{\alpha_n,n}$ is a W-test at weight c_n, where c_n is the critical value of $\varphi_{\alpha_n,n}$. According to Corollary 2 we have $\lim_{n\to\infty}c_n^{1/n} = \exp[-I(f_0:f_1)] = \lambda > 0$. Differentiating $\lambda^t\rho(t)$ with respect to t yields

$$\ln\lambda = \rho^{-1}(t^*)\int\ln\left(\frac{f_1(x)}{f_0(x)}\right)f_0^{t^*}(x)f_1^{1-t^*}(x)\,d\mu$$

as condition on t^*. Hence $t^* = 1$, so that t^* is an interior point of $\{t:\rho(t) < \infty\}$. Therefore we can apply Theorem 4, if $f_1(x)/f_0(x) \neq \lambda$ μ-a.s. on B. In that case the assertion then follows from the fact that according to (26) for $s = t^* = 1$ we get $\exp[-I(f_0:g_t)] = H(t^*)/H(t)$. If $f_1(x)/f_0(x) = \lambda$ μ-a.s. on B, then the assertion follows from $\lim_{n\to\infty}\alpha_n^{1/n} = \lim_{n\to\infty}(1-\alpha_n)^{1/n} = 1$, Corollary 2 and from the fact that $g_t(x) = f_0(x)$ μ-a.s. for $t > 0$ and $g_t(x) = f_1(x)$ μ-a.s. for $t \leqq 0$.

Some of the results recorded above could have been derived or have been derived from results of the theory of probabilities of large deviations. Our last result shows that, conversely, the method used in this paper leads to large deviation results.

COROLLARY 5. *Let* $\{X_i\}$ *be a sequence of independent random variables with common* (*non-degenerate*) *distribution* P_0 *and let* $S_n = \sum_{i=1}^n X_i$. *If the moment generating function* $m(\tau)$ *of* P_0 *exists for some* $\tau > 0$ *and if* $e^{-a\tau}m(\tau)$ *attains its minimum at* $\tau = \tau^* > 0$ *in the interior of* $\{\tau: m(\tau) < \infty\}$, *then for any sequence* $\{a_n\}$ *with* $\lim_{n\to\infty}(a_n/n) = a$ *it holds that* $\lim_{n\to\infty}[P_0(S_n > na_n)]^{1/n} = e^{-a\tau^*}m(\tau^*$

PROOF. Let μ be a measure dominating P_0 and denote by f_0 a μ-density of P_0. Let P_1 be a distribution given by the μ-density $f_1(x) = f_0(x)m^{-1}(\tau^*) e^{\tau^* x}$. Let $\psi_{\lambda_n, n}$ be a W-test for testing $H: P = P_0$ against $K: P = P_1$ at weight $\lambda_n = m^{-n}(\tau^*) e^{a_n \tau^*}$. It follows that $f_1(\mathbf{x}) > \lambda_n f_0(\mathbf{x})$ if and only if $S_n > na_n$, i.e. that $P_0(S_n > na_n) = E_0 \psi_{\lambda_n, n}$. Theorem 4 with $t^* = 0$ then implies the assertion.

REMARK. The result of Corollary 5 is due to Bahadur [1], Lemma 2.2, who proved it by a different method.

Acknowledgment. We should like to thank the referee for his careful reading of this paper and for a series of helpful comments.

REFERENCES

[1] BAHADUR, R. R. (1960). On the asymptotic efficiency of tests and estimates. *Sankhyā* **22** 229–252.

[2] CHERNOFF, H. (1952). A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations. *Ann. Math. Statist.* **23** 493–507.

[3] CHERNOFF, H. (1956). Large sample theory: Parametric case. *Ann. Math. Statist.* **27** 1–22.

[4] KRAFFT, O. (1967). Eine symmetrische Behandlung des Testproblems. *Arch. Math.* **28** 545–560.

[5] KRAFFT, O. and WITTING, H. (1967). Optimale Tests und ungünstigste Verteilungen. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **7** 289–302.

[6] NÖLLE, G. and PLACHKY, D. (1967). Zur schwachen Folgenkompaktheit von Testfunktionen. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **8** 182–184.

[7] RAO, C. R. (1962). Efficient estimates and optimum inference procedures in large samples. *J. Roy. Statist. Soc. Ser. B* **24** 46–72.

[8] RAO, C. R. (1967). *Linear statistical inference and its applications.* Wiley, New York.

[9] VAJDA, I. (1970). On the amount of information contained in a sequence of observations. To appear in *Kybernetika* (Prague).

[10] WALD, A. (1959). *Sequential Analysis.* Wiley, New York.