ON A MEASURE OF TEST EFFICIENCY PROPOSED BY R. R. BAHADUR[1]

By Leon J. Gleser

Columbia University

1. Introduction. In [4] R. R. Bahadur has shown "that the study (as random variables) of the levels attained when two alternative tests of the same hypothesis are applied to given data affords a method of comparing the performances of the tests in large samples." This method of comparison produces as an end result a measure of asymptotic relative performance which in this paper is called the "Bahadur efficiency."

Both Bahadur [4] and this author [9] have pointed out that the Bahadur efficiency is in general only an approximate measure of asymptotic relative performance. In Section 2 we introduce the Bahadur efficiency by giving sufficient conditions for the measure to be exact. These conditions are generalizations of the conditions given by Bahadur ([4], p. 282).

In Section 3 we show that the conditions required to compute the Bahadur efficiency are much less restrictive than the discussion in [4] might indicate. A very general set of sufficient conditions is given, and the heuristic justification given in [4] by Bahadur for his method is sketched to show the modifications in argument required by this generalization.

In Section 4 a set of sufficient conditions easily applied in practice (but more restrictive than the conditions of Section 3) are given. These conditions are similar to those defining a "standard sequence" in [4], but are relaxed to allow general rates of convergence $L(n)$ rather than the more restrictive rate $L(n) = n$ assumed in [4]. An example of test comparison when $L(n) = (\log n)^{\frac{1}{2}}$ is discussed. Finally, we illustrate the "approximate" character of Bahadur efficiency by finding the Bahadur efficiency of two equivalent tests.

2. A special case of test comparison. For the duration of this paper it is understood that by the general problem of asymptotic test comparison we mean the following: We are given a set of probability measures $\{P_\theta\}$, $\theta \varepsilon \Omega$, defined over an arbitrary space S of points s. For $\Omega_0 \subset \Omega$, H is defined to be the hypothesis that $\theta \varepsilon \Omega_0$. To test H we have two sequences of test statistics $\{T_n^{(1)}\}$ and $\{T_n^{(2)}\}$, $n = 1, 2, \cdots$; we wish to compare the performances of the (asymptotically) optimal tests based on $\{T_n^{(i)}\}$, $i = 1, 2$, as $n \to \infty$ in the hope that some idea of the relative performance of these tests for any value of n may be gained.

Bearing this general context in mind, consider now a special case of our problem.

Received December 26, 1963; revised March 17, 1964.

[1] This paper is partly based upon a Stanford University doctoral thesis [9], sponsored by the National Science Foundation under Foundation Grant NSF-214. Research supported in part by National Science Foundation Grant NSF-GP-2074 at Columbia University.

Assumption 2.1. For $i = 1, 2$, there exist continuous cumulative distribution functions $F_n^{(i)}(x)$ such that for each $\theta_0 \varepsilon \Omega_0$,

$$P_\theta\{T_n^{(i)} \leqq x\} = F_n^{(i)}(x), \qquad \text{all } x. \tag{2.1}$$

Assumption 2.2. For $i = 1, 2$, there exist continuous, strictly increasing functions $L^{(i)}(x)$ mapping $(0, \infty)$ into $(0, \infty)$, $L^{(i)}(x) \to \infty$ as $x \to \infty$, and there exist functions $h_i(\theta)$, $0 < h_i(\theta) < \infty$, defined on $\Omega - \Omega_0$, such that for each $\theta \varepsilon \Omega - \Omega_0$,

$$\text{plim}_{n\to\infty} M_n^{(i)}/L^{(i)}(n) = h_i(\theta) \tag{2.2}$$

where

$$M_n^{(i)} = -2\log[1 - F_n^{(i)}(T_n^{(i)})]. \tag{2.3}$$

Since $M_n^{(i)}$ is a continuous monotone transformation of $T_n^{(i)}$, then for any test of H based on $T_n^{(i)}$ there exists an equivalent (possibly randomized) test of H based on $M_n^{(i)}$, $i = 1, 2$, all n. Thus, comparison of $\{T_n^{(i)}\}$, $i = 1, 2$, is the same as comparison of $\{M_n^{(i)}\}$, $i = 1, 2$.

By Assumption 2.1 and the probability integral transformation,

$$P_\theta\{M_n^{(i)} > x\} = \exp(-\tfrac{1}{2}x) \tag{2.4}$$

for all i, $\theta \varepsilon \Omega_0$, and n. On the other hand, a direct consequence of Assumption 2.2 is that plim $M_n^{(i)} = \infty$ all i, all $\theta \varepsilon \Omega - \Omega_0$. It follows that for large n, large values of $M_n^{(i)}$ are significant for rejecting H. From (2.2) we see that optimal rejection regions for H should be of the form:

$$M_n^{(i)} \geqq ch_i(\theta)L^{(i)}(n), \qquad 0 < c < 1, \tag{2.5}$$

when n is large. The reason for this last remark is that regions of the form (2.5) provide for the fastest convergence to 0 of the type-one probability of error while still allowing the power of the test to converge to 1 at the alternative $\theta \varepsilon \Omega - \Omega_0$, $i = 1, 2$. Since the rejection regions (2.5) hold the power near 1 for n large, the tests based on these regions have asymptotically equivalent performance if the type-one probabilities of error are equal; that is if (from (2.4)):

$$\exp - \tfrac{1}{2}\{ch_1(\theta)L^{(1)}(n)\} = \exp - \tfrac{1}{2}\{ch_2(\theta)L^{(2)}(n)\}$$

or equivalently,

$$h_1(\theta)L^{(1)}(n) = h_2(\theta)L^{(2)}(n). \tag{2.6}$$

If $L^{(1)}(n)/L^{(2)}(n)$ converges to 0 (∞) as $n \to \infty$, then it is apparent that almost any measure of efficiency would indicate that the test based on $\{T_n^{(1)}\}$ is less (more) efficient than the test based on $\{T_n^{(2)}\}$. If $L^{(1)}(n)/L^{(2)}(n)$ converges to b, $0 < b < \infty$, then from (2.2) we can assume $L^{(1)}(x) = L^{(2)}(x) = L(x)$ without any loss of generality. Since a sequence of tests having power converging to 1 has power converging to 1 for any subsequence, our goal should be to find a subsequence of the tests based on $\{T_n^{(2)}\}$, say $\{T_{n'(n)}^{(2)}\}$, such that

$$L(n)/L(n') = [h_1(\theta)]/[h_2(\theta)] \equiv \Psi_{1,2}(\theta).$$

Such a subsequence $\{n'(n)\}$ is called an "asymptotically equivalent subsequence of sample sizes" for $\{T_n^{(2)}\}$ as compared to the sequence $\{n\}$ of sample sizes for $\{T_n^{(1)}\}$.

Since $L(x)$ is continuous and strictly monotone, there exists a function $L^I(x)$ from $(0, \infty)$ to $(0, \infty)$ such that $L^I(L(x)) \equiv x$, all x. Thus the asymptotically equivalent subsequence of sample sizes $n'(n)$ is:

$$n'(n) = L^I(\Psi_{1,2}^{-1}(\theta)L(n)). \tag{2.7}$$

(If $n'(n)$ is not an integer, it is understood that $n'(n)$ expresses a randomized sample size, where randomization is between the greatest integer less than $n'(n)$ and the least integer greater than $n'(n)$. Since $\Psi_{1,2}(\theta) > 0$ and $L(n) \to \infty$ as $n \to \infty$, such (randomized) subsequences also have power converging to 1.) From (2.7) and the preceding discussion we conclude that $\Psi_{1,2}(\theta)$ is a legitimate measure of asymptotic relative efficiency for the tests based on $\{T_n^{(i)}\}$, $i = 1, 2$, at the alternatives $\theta \varepsilon \Omega - \Omega_0$.

Note that whereas the measure of efficiency defined by Hodges and Lehmann [10] is found by holding the type-one probabilities of error fixed and equal and adjusting the type-two probabilities of error, this measure of efficiency holds the type-two errors fixed and equal (approximately) and adjusts the type-one errors.

Assumptions 2.1 and 2.2 are generalizations of the assumptions defining a "standard sequence in the strict sense" in [4], p. 282. As Bahadur remarks, determination of such exact efficiencies is in general as difficult as the determination of the exact efficiencies defined in [5], [10], or [12]. We are thus led to look for a short-cut, easily computed approximation to the efficiency $\Psi_{1,2}(\theta)$. As Bahadur shows, such an approximation can be defined and computed even when Assumptions 2.1 and 2.2 do not hold. This approximation is what we have called the Bahadur efficiency.

3. The Bahadur efficiency. To apply Bahadur's method we only require that:

ASSUMPTION 3.1. There exist continuous cumulative distribution functions $F^{(i)}(x)$, $i = 1, 2$, such that for each $\theta \varepsilon \Omega_0$,

$$\lim_{n\to\infty} P_\theta\{T_n^{(i)} \leqq x\} = F^{(i)}(x), \qquad \text{all } x. \tag{3.1}$$

ASSUMPTION 3.2. There exists a continuous, strictly increasing function $L(x)$ from $(0, \infty)$ to $(0, \infty)$, $L(x) \to \infty$ as $x \to \infty$, and functions $h_i(\theta)$, $0 < h_i(\theta) < \infty$, defined on $\Omega - \Omega_0$, $i = 1, 2$, such that

$$\text{plim } K_n^{(i)}/L(n) = h_i(\theta), \qquad \text{all } \theta \varepsilon \Omega - \Omega_0, \tag{3.2}$$

where

$$K_n^{(i)} = -2 \log[1 - F^{(i)}(T_n^{(i)})]. \tag{3.3}$$

NOTE. As in Section 2, we can restrict attention to the case where $K_n^{(1)}$ and $K_n^{(2)}$ converge at the same rate $L(n)$. Also as in Section 2, comparison of $\{K_n^{(i)}\}$, $i = 1, 2$, is equivalent to comparison of $\{T_n^{(i)}\}$, $i = 1, 2$.

Repeating the argument in Section 2 and estimating $P_\theta\{K_n^{(i)} \geqq x\}$ by its limit (i.e., by $P\{\chi_2^2 \geqq x\} = \exp(-\frac{1}{2}x))$ as $n \to \infty$, we can again conclude that

$$\varphi_{1,2}(\theta) \equiv [h_1(\theta)]/[h_2(\theta)], \qquad \theta \varepsilon \Omega - \Omega_0, \tag{3.4}$$

is an (approximate) measure of asymptotic relative efficiency at $\theta \varepsilon \Omega - \Omega_0$ for $\{K_n^{(1)}\}$ and $\{K_n^{(2)}\}$, and that $n'(n) = L^I(L(n)\varphi_{1,2}^{-1}(\theta))$ is the (approximate) asymptotically equivalent sample size for test sequence 2 as compared to a sample size of n for test sequence 1 when n is large.

The error in the approximation comes from the fact that in general

$$P_{\theta_0}\{K_n^{(1)} \geqq x\} \neq P_{\theta_0}\{K_{n'}^{(2)} \geqq x\} \tag{3.5}$$

for every $\theta_0 \varepsilon \Omega_0$. Thus the type-one probabilities of error for the tests based on $K_n^{(1)}$ and $K_{n'}^{(2)}$ may be of different orders of magnitude for large n even though

$$\lim_{n\to\infty} P_{\theta_0}\{K_n^{(1)} \geqq x\} = \lim_{n\to\infty} P_{\theta_0}\{K_{n'}^{(2)} \geqq x\} = \exp(-\tfrac{1}{2}x).$$

If, however,

$$\lim_{n\to\infty} \frac{P_{\theta_0}\{K_n^{(1)} \geqq \gamma L(n)\}}{P_{\theta_0}\{K_{n'}^{(2)} \geqq \gamma L(n')\}} = 1 \tag{3.6}$$

for all $\theta_0 \varepsilon \Omega_0$, all $0 < \gamma < \infty$, all $\theta \varepsilon \Omega - \Omega_0$, then by the arguments of Section 2 it should be apparent that $\varphi_{1,2}(\theta)$ is an exact measure of efficiency. Of course, if $P_{\theta_0}\{T_n^{(i)} \leqq x\} = F^{(i)}(x)$ for all n, all $\theta_0 \varepsilon \Omega_0$, $i = 1, 2$, then Assumptions 2.1 and 2.2 are satisfied, (3.6) holds trivially, and in fact we have $\varphi_{1,2}(\theta) = \Psi_{1,2}(\theta)$ for all $\theta \varepsilon \Omega - \Omega_0$.

Even in the cases where $\varphi_{1,2}(\theta)$ is not exact, it is still possible to find relationships between $\varphi_{1,2}$ and the characteristics of the tests based on $\{K_n^{(i)}\}$, $i = 1, 2$. First, as we have noted before, the choice of sample size $n'(n)$ for a given $\theta \varepsilon \Omega - \Omega_0$ makes the power of the (asymptotically) optimal tests based on $K_n^{(1)}$ and $K_{n'}^{(2)}$ approximately equal at that given θ. If

$$\lim_{n\to\infty} K_n^{(i)}/L(n) = h_i(\theta) \quad \text{a.s.} \tag{3.7}$$

for all $\theta \varepsilon \Omega - \Omega_0$, $i = 1, 2$, then we can establish an even stronger connection between $\varphi_{1,2}$ and the sample sizes needed to give equivalent power for the elements of sequences 1 and 2.

THEOREM 3.1. *For $i = 1, 2$, and any $v > 0$, let $N_i^+(v, s)$ be the smallest m such that $K_n^{(i)}(s) > v$ all $n \geqq m$ ($N_i^+ = \infty$ if no such m exists) and let $N_i^-(v, s)$ be the smallest n such that $K_{n+1}^{(i)}(s) \geqq v$ ($N_i^- = \infty$ if no such n exists). Under the convention that $\infty/\infty = 1$, Condition (3.7) implies that*

$$\lim_{v\to\infty} L(N_2^-)/L(N_1^+) = \lim_{v\to\infty} L(N_2^+)/L(N_1^-) = \varphi_{1,2}(\theta) \quad \text{a.s.} \tag{3.8}$$

at every $\theta \varepsilon \Omega - \Omega_0$.

PROOF. The proof is a trivial extension of the proof given for $L(n) = n$ in [4], pp. 279–280. Simply replace n by $L(n)$ in (15), and replace N_i^- and N_i^+ by $L(N_i^-)$ and $L(N_i^+)$ in (16). Q.E.D.

NOTE. Bahadur [4] has conjectured that Theorem 3.1 might hold under condition (3.2) provided that we only demanded convergence in probability in (3.8). However, if $\Omega = \{0, 1\}$, $\Omega_0 = \{0\}$, $S = [0, 1]$, C is the Cantor set, $P_0(s)$ is the uniform distribution on S, and $P_1(s)$ places mass 1 on C, then let

$$K_n^{(1)}(s) = n \quad , \quad s \varepsilon C$$
$$= -2 \log s, \quad s \not\varepsilon C,$$

and

$$K_n^{(2)}(s) = 0 \quad , \quad s \varepsilon C \cap I_n$$
$$= n \quad , \quad s \varepsilon C - I_n$$
$$= -2 \log s, \quad s \not\varepsilon C,$$

where $I_1 = [0, \frac{1}{2}]$, $I_2 = [0, 1]$, $I_3 = [0, \frac{1}{3}]$, $I_4 = [\frac{1}{3}, \frac{2}{3}]$, $I_5 = [\frac{2}{3}, 1]$ etc. Under Ω_0 the distributions of $K_n^{(1)}$ and $K_n^{(2)}$ are χ_2^2, while under $\Omega - \Omega_0$ we have that $K_n^{(1)}/n \to 1$ a.s., plim $K_n^{(2)}/n = 1$, but $L(N_2^+(v, s)) = N_2^+(v, s) = \infty$ all v, all $s \varepsilon C$, and thus $\text{plim}_{v\to\infty} L(N_2^+)/L(N_1^-) = \text{plim}_{v\to\infty} N_2^+/N_1^- = \infty \neq 1 = \varphi_{1,2}(1)$. (I am indebted to Dr. P. Bickel, whose suggestions led to the above example.) The necessary and sufficient conditions under which Bahadur's conjecture holds are somewhat involved and will not be given here.

As a final explanation of the Bahadur efficiency, Bahadur shows that $\varphi_{1,2}$ has a relation to the power functions of the tests based on $\{T_n^{(1)}\}$ and $\{T_n^{(2)}\}$ as follows. Consider any number γ, $0 < \gamma < \infty$. Define $\beta_n(\gamma \mid \theta)$ to be $P_\theta\{K_n^{(i)} > \gamma\}$. Although for $\theta \varepsilon \Omega_0$ it is not necessarily true that $\gamma = -2 \log \alpha$ implies $\beta_n(\gamma \mid \theta) = \alpha$ (i.e. the test is of size α) we can still compare the power functions of tests defined by the rejection regions $\{K_n^{(i)} > \gamma\}$ as if they were the power functions of approximate size $\exp(-\frac{1}{2}\gamma)$ (Bahadur argues that such tests are often used in practice).

Taking note of this approximation, let for each n

$$\delta_n(1, 2 \mid \theta) \equiv \sup_{\gamma \varepsilon (0,\infty)} [\beta_n^{(2)}(\gamma \mid \theta) - \beta_n^{(1)}(\gamma \mid \theta)]. \tag{3.9}$$

It is easily seen from the definition of β_n that $0 \leqq \delta_n \leqq 1$. Let us say that $\{T_n^{(2)}\}$ dominates $\{T_n^{(1)}\}$ (written $T_n^{(2)} \succ T_n^{(1)}$) at $\theta \varepsilon \Omega - \Omega_0$ if:

$$\lim_{n\to\infty} \delta_n(1, 2 \mid \theta) = 0. \tag{3.10}$$

REMARK. In particular, $T_n^{(2)} \succ T_n^{(1)}$ if $1 - \beta_n^{(2)}(\gamma \mid \theta) \geqq 1 - \beta_n^{(1)}(\gamma \mid \theta)$ for all n and all $\gamma \varepsilon (0, \infty)$.

THEOREM 3.2. [4, Appendix 1].
(i) *If* $T_n^{(2)} \succ T_n^{(1)}$ *at* $\theta \varepsilon \Omega - \Omega_0$, *then* $\varphi_{1,2}(\theta) \leqq 1$.
(ii) *If* $\varphi_{1,2}(\theta) < 1$, *then* $T_n^{(2)} \succ T_n^{(1)}$ *at* θ.

PROOF. The proof follows as in [4] replacing n by $L(n)$ where appropriate. Q.E.D.

As a consequence of Theorem 3.2, $\varphi_{1,2} < 1$ if and only if $T_n^{(2)} \succ T_n^{(1)}$ and

$T_n^{(1)} \not\succ T_n^{(2)}$. It also follows that $\varphi_{1,2} = 1$ if each sequence dominates the other or if neither dominates. The latter eventuality is possible, and thus domination is only a partial ordering of sequences $\{T_n^{(i)}\}$.

4. Applications and examples. Under the following conditions we can simplify the computation of the Bahadur efficiency:

ASSUMPTION 4.1. For the $F^{(i)}(x)$ defined in Assumption 3.1, $i = 1, 2$, there exist a $t_i > 0$ and an $a_i > 0$ such that

$$-2\log(1 - F^{(i)}(x)) = a_i x^{t_i}(1 + o(1)), \qquad x \to \infty. \tag{4.1}$$

ASSUMPTION 4.2. There exist continuous, strictly increasing functions $b^{(i)}(x)$ mapping $(0, \infty)$ into $(0, \infty)$, $b^{(i)}(x) \to \infty$ as $x \to \infty$, and functions $c_i(\theta)$, $0 < c_i(\theta) < \infty$, defined on $\Omega - \Omega_0$, $i = 1, 2$, such that

$$\text{plim}\; T_n^{(i)}/b_{(n)}^{(i)} = c_i(\theta), \qquad \text{all } \theta \in \Omega - \Omega_0 .$$

ASSUMPTION 4.3. $(b^{(1)}(x))^{t_1} = (b^{(2)}(x))^{t_2}$.

From these assumptions it is easily shown (as in [4], p. 272) that Assumption 3.2 of Section 2 is met with

$$h_i(\theta) = a_i c_i^{t_i}(\theta), \tag{4.2}$$

and

$$L(x) = [b^{(1)}(x)]^{t_1}. \tag{4.3}$$

That these conditions relax those given in [4] for a "standard sequence" can be seen by noting that in [4], $t_1 = t_2 = 2$ and $b^{(1)} = b^{(2)} = n^{\frac{1}{2}}$.

To apply Assumptions 4.1, 4.2, and 4.3, the most difficult problem will usually be to verify Assumption 4.1. It is shown in [4] that the standard normal distribution $\mathfrak{N}(0, 1)$ and the chi distribution χ_k for any k satisfy (4.1) with $t = 2$ and $a = 1$. Some theorems useful in verifying that other cumulative distributions satisfy (4.1) are given in [9].

In general, most tests of hypothesis conform to the requirements in [4]. If a sequence of test statistics does not meet these requirements, quite often it is possible to find a monotone transformation for the statistics such that the transformed statistics meet the requirements. If, for example, T_n satisfies Assumptions 3.1, 4.1, 4.2, and 4.3 with $L(x) = x$, but satisfies (4.1) with $t = 4$, then T_n^2 satisfies Bahadur's requirements. Another example (where the original T_n may not even satisfy Assumptions 3.1 and 3.2) is the likelihood ratio statistic. A discussion of this last example may be found in [9].

However, there are many cases (e.g., in discussion of random walks, in tests based on order statistics) where Bahadur's conditions cannot be met, and yet where the generalization of these conditions given in this section apply. One such case is the following example of the comparison of tests based on "systematic statistics" [11] drawn from a normal distribution with unknown mean and known variance.

Consider a random sample $x_1, \cdots x_n$, from a normal distribution with known variance σ^2 (without loss of generality we let $\sigma^2 = 1$), but of unknown mean θ. It is desired to test the hypothesis $H: \theta = \theta_0$ against alternatives $\theta > \theta_0$. However, instead of receiving the variables $x_1, \cdots x_n$, we are allowed only to know the m largest and the m smallest values from the sample. We can select m, but at a cost which is monotone increasing in m. This situation might occur, for example, in studies of river levels where only records of extreme floods and droughts have been kept and where the effort to find the mth largest value $x^{(m)}$ and the mth smallest value $x_{(m)}$ increases with m. Since $\frac{1}{2}(x^{(m)} + x_{(m)})$ is an unbiased and consistent estimator of θ, we might be interested in tests of the form: Reject H if $T_n^{(m)} > K$, where $T_n^{(m)} = (2 \log n)^{\frac{1}{2}}(x^{(m)} + x_{(m)} - 2\theta_0)$ and K is a fixed constant. We can then ask what is the relative performance of $T_n^{(i)}$ and $T_n^{(j)}$ for n large.

As is shown in [8], p. 375,

$$T_n^{(m)} = 2(2 \log n)^{\frac{1}{2}}(\theta - \theta_0) + V(n, m) \tag{4.4}$$

where $V(n, m)$ is a random variable whose limiting distribution as $n \to \infty$ is the difference of two independent random variables each with distribution $p(v; m) = \exp\{-mv - e^{-v}\}/\Gamma(m), -\infty \leqq v \leqq \infty$. From this fact it immediately follows that plim $T_n^{(m)}/(\log n)^{\frac{1}{2}} = 2(2)^{\frac{1}{2}}(\theta - \theta_0)$, and that for $\theta = \theta_0$,

$$1 - F^{(m)}(x) = \lim P_{\theta_0}\{T_n^{(m)} > x\} = \frac{\Gamma(2m)}{\Gamma^2(m)} \int_0^{e^{-x}} \frac{p^{m-1}}{(1+p)^{2m}}\, dp.$$

It can easily be shown that there exist constants $C_1^{(m)}$ and $C_2^{(m)}$ such that $C_2^{(m)} e^{-mx} \leqq 1 - F^{(m)}(x) \leqq C_1^{(m)} e^{-mx}$. Thus the test sequences $\{T_n^{(m)}\}$ satisfy Assumptions 3.1, 4.1, 4.2, and 4.3 for all finite m with $b^{(m)}(x) = (\log x)^{\frac{1}{2}}$, $a_m = m$, $t_m = 1$, and $c_m(\theta) = 2(2)^{\frac{1}{2}}(\theta - \theta_0)$. Hence $L(x) = (\log x)^{\frac{1}{2}}$ and $h_m(\theta) = 2(2)^{\frac{1}{2}} m(\theta - \theta_0)$. The efficiency of $\{T_n^{(i)}\}$ with respect to $\{T_n^{(j)}\}$ is thus $\varphi_{i,j} = i/j$.

It would be of interest to know how the Hodges and Lehmann [10] or Pitman [12] efficiencies for this problem compare with the above result. Unfortunately the computation of the Hodges-Lehmann efficiency is a major problem in analysis, and it is not certain whether the Pitman efficiency can be extended to this case. One efficiency that can be computed is the relative efficiency of the statistics $\frac{1}{2}(x^{(m)} + x_{(m)})$ as estimates of the true θ. The ratio of the variance of $\frac{1}{2}(x^{(i)} + x_{(i)})$ and the variance of $\frac{1}{2}(x^{(j)} + x_{(j)})$ is, in the limit as $n \to \infty$, equal to

$$\sum_{l=i}^{\infty} \frac{1}{l^2} \Big/ \sum_{l=j}^{\infty} \frac{1}{l^2}$$

(see [8], p. 377).

We can both compare the Bahadur and Hodges-Lehmann efficiencies and illustrate the nonexactness of the Bahadur efficiency by the following example: Consider a random sample $x_1, \cdots x_n$ from a normal distribution with unknown mean μ and unknown variance σ^2. We wish to test the hypothesis $H: \theta = 0$, where $\theta \equiv \mu/\sigma$. Let $\bar{x}_n = \sum_{i=1}^n x_i/n$, $S_n^2 = \sum_{i=1}^n (x_i - \bar{x}_n)^2/n - 1$, then it is well known that the likelihood ratio statistic is $\lambda = [1 + (n/(n-1))(\bar{x}_n^2/S_n^2)]^{-\frac{n}{2}}$ and a monotone function of λ is the statistic $T_n^{(1)} = n\bar{x}_n^2/S_n^2$. Let $T_n^{(2)} = -2 \log \lambda$. Then $T_n^{(1)}$ and $T_n^{(2)}$ are equivalent test statistics. On the other hand, under H

the limiting distributions of both $T_n^{(1)}$ and $T_n^{(2)}$ are those of a chi-squared random variable with two degrees of freedom. Such a limiting distribution satisfies Assumption 4.1 with $t_i = 1$, $a_i = 1$, $i = 1, 2$. Further plim $T_n^{(1)}/n = \theta^2$ and plim $T_n^{(2)}/n = \log(1 + \theta^2)$. Since Assumptions 4.2 and 4.3 are satisfied, then $L(x) = x$, $h_1(\theta) = \theta^2$, $h_2(\theta) = \log(1 + \theta^2)$, and $\varphi_{1,2}(\theta) = \theta^2/\log(1 + \theta^2)$. As $|\theta| \to \infty$, $\varphi_{1,2}(\theta) \to \infty$. But since $T_n^{(1)}$ and $T_n^{(2)}$ are equivalent test statistics, any exact efficiency for these tests must equal one for all $\theta \varepsilon \Omega - \Omega_0$.

Bahadur has conjectured that in general as $\theta \varepsilon \Omega - \Omega_0$ tends to some $\theta_0 \varepsilon \Omega_0$, the Bahadur efficiency becomes exact. The above example supports that conjecture (since $\lim \varphi_{1,2} = 1$ as $|\theta| \to 0$). Under certain regularity conditions and for $L(x) = x$, he shows in [4] that $\lim \varphi_{1,2}(\theta)$ as $\theta \to \theta_0$ is Pitman's efficiency evaluated at θ_0. This theorem should be easily extendible to the case where $L(x) = x^t$, some $t > 0$. For general $L(x)$ it is not obvious how this extension is to be accomplished.

The relatively simple application of the Bahadur method, its intuitive appeal as an approximate extension of the exact efficiency defined in Section 2, and its formal relation to the power functions of the tests compared (viz. Section 3) justify the use of the Bahadur efficiency in practice as a reasonable first step in the comparison of tests of a given hypothesis. The last example in Section 4 indicates that care must be exercised in using this method, but this warning in no way mitigates the usefulness of the procedure.

Acknowledgments. The author is indebted to Doctors Ingram Olkin, R. R. Bahadur, H. Chernoff, and S. Das Gupta, and to the referee for their comments and assistance in the preparation of this paper.

REFERENCES

[1] Anderson, T. W. (1960). *Introduction to Multivariate Analysis*, **1** (2nd ed.). Wiley, New York.

[2] Bahadur, R. R. (1960). On the asymptotic efficiency of tests and estimates. *Sankhyā* **22** 229–253.

[3] Bahadur, R. R. (1960). Simultaneous comparison of the sign and optimum tests of a normal mean. *Contributions to Probability and Statistics: Essays in Honor of Harold Hotelling*. Stanford Univ. Press.

[4] Bahadur, R. R. (1960). Stochastic comparison of tests. *Ann. Math. Statist.* **31** 276–295.

[5] 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.

[6] Chernoff, H. (1954). On the distribution of the likelihood ratio. *Ann. Math. Statist.* **25** 573–578.

[7] Chernoff, H. (1956). Large sample theory: parametric case. *Ann. Math. Statist.* **27** 1–22.

[8] Cramér, Harald, (1958). *Mathematical Methods of Statistics*, **1** (1st ed.). Princeton Univ. Press.

[9] Gleser, L. J. (1963). The comparison of multivariate tests of hypothesis by means of Bahadur efficiency. *Technical Report No. 8* (N.S.F. Grant 214), Stanford Univ.

[10] Hodges, J. L., Jr. and Lehmann, E. L. (1956). The efficiency of some nonparametric competitors of the *t* test. *Ann. Math. Statist.* **27** 324–335.

[11] Mosteller, F. (1946). On some useful "inefficient" statistics. *Ann. Math. Statist.* **17** 377–408.

[12] Noether, G. E. (1955). A theorem of Pitman. *Ann. Math. Statist.* **26** 64–68.