The Annals of Mathematical Statistics
1968, Vol. 39, No. 1, 179–182

AN EXAMPLE OF LARGE DISCREPANCY BETWEEN MEASURES OF ASYMPTOTIC EFFICIENCY OF TESTS[1]

BY R. K. TSUTAKAWA

University of Chicago

1. Summary. Two statistics for testing that the mean of a normal distribution is zero are compared to show that the measures of asymptotic relative efficiency proposed by Pitman (cf., e.g., [5]), Hodges and Lehmann [4], and Bahadur [1], [2] do not always agree, even locally for alternative means near zero.

2. Example. Let $X_1, X_2, \cdots$ be a sequence of independent random variables with common normal distribution having mean θ and unit variance. Consider testing the hypothesis $\theta = 0$ against the alternatives $\theta \neq 0$. For the first procedure consider the sequence of test statistics

$$T_n^{(1)} = n^{\frac{1}{2}} |\bar{X}_n|, \tag{1}$$

$n = 1, 2, \cdots$, where $\bar{X}_n = \sum_{j=1}^n X_j/n$. Now let p be a constant, $0 < p < 1$, and let k_n equal the integral part of np. In the following, we assume n so large that $1 < k_n < n$. For the second procedure consider the sequence of test statistics

$$\begin{aligned} T_n^{(2)} &= (n - k_n)^{\frac{1}{2}} \bar{Z}_n \quad \text{if} \quad \bar{Y}_n \geqq 0 \\ &= -(n - k_n)^{\frac{1}{2}} \bar{Z}_n \quad \text{if} \quad \bar{Y}_n < 0, \end{aligned} \tag{2}$$

where $\bar{Y}_n = \sum_{j=1}^{k_n} X_j/k_n$ and $\bar{Z}_n = \sum_{j=k_n+1}^{n} X_j/(n - k_n)$. The second procedure is to be interpreted as splitting n observations into two samples and performing a one-sided test with the second sample, using the first sample to determine the direction. (This is a special case of a more general problem of data splitting suggested to me by W. H. Kruskal.) Large values of $T_n^{(i)}$ will be considered significant, $i = 1, 2$. "Efficiency" will henceforth refer to the efficiency of $T_n^{(2)}$ relative to $T_n^{(1)}$; this convention is appropriate since $T_n^{(1)}$ is an optimal statistic under most circumstances.

3. Asymptotic efficiencies. We shall first define a basic non-asymptotic efficiency, say e, and express the Pitman, Hodges-Lehmann, and (exact) Bahadur efficiencies, say e_1, e_2, and e_3, in terms of e.

Choose and fix α and β, $0 < \alpha < 1 - \beta < 1$. We will construct tests from each statistic which have size α and power $1 - \beta$ against a given $\theta \neq 0$, and compare the requisite sample sizes. For each n, let $t_n^{(i)}$, $i = 1, 2$, be constants such that $P(T_n^{(i)} \geqq t_n^{(i)} \mid \theta = 0) = \alpha$. Then $t_n^{(1)} = K_{\alpha/2}$ and $t_n^{(2)} = K_\alpha$, where $\Phi(K_\epsilon) = 1 - \epsilon$ for $0 < \epsilon < 1$, Φ being the standard normal distribution func-

Received 15 June 1967.

[1] This research was sponsored by the Army Research Office, Office of Naval Research, and Air Force Office of Scientific Research by Contract No. Nonr-2121(23), NR 342-043.

tion. Choose and fix $\theta \neq 0$. By symmetry it is enough to take $\theta > 0$. For each i, let $M_i = M_i(\alpha, \beta, \theta)$ be the smallest n such that $1 < k_n < n$ and $P(T_n^{(i)} \geqq t_n^{(i)} \mid \theta) \geqq 1 - \beta$. We define $e(\alpha, \beta, \theta) = M_1(\alpha, \beta, \theta)/M_2(\alpha, \beta, \theta)$.

In the present case e_1 cannot be expressed by "Pitman's formula"; however, as in typical cases, it may be expressed as $e_1 = \lim_{\theta\to 0} e(\alpha, \beta, \theta)$. An alternative method of obtaining e_1 is to apply Pitman's arguments to the present case; we use the latter approach in the following paragraph.

For given α, n, and $\theta \neq 0$, the power of the critical region $\{T_n^{(1)} \geqq t_n^{(1)}\}$ is

$$(3) \qquad 1 - \beta_n^{(1)} = \Phi(-K_{\alpha/2} - n^{\frac{1}{2}}\theta) + \Phi(-K_{\alpha/2} + n^{\frac{1}{2}}\theta)$$

and the power of the region $\{T_n^{(2)} \geqq t_n^{(2)}\}$ is

$$(4) \quad 1 - \beta_n^{(2)} = \Phi(-k_n^{\frac{1}{2}}\theta)\Phi(-K_\alpha - j_n^{\frac{1}{2}}\theta) + \Phi(k_n^{\frac{1}{2}}\theta)\Phi(-K_\alpha + j_n^{\frac{1}{2}}\theta),$$

where $j_n = n - k_n$. Now consider the sequence of alternatives, $\theta_n = \nu_1/n^{\frac{1}{2}}$, $n = 1, 2, \cdots$, where ν_1 is such that, for each n, $\beta_n^{(1)} = \beta$ at the alternative $\theta = \theta_n$. For each n, let $m = m(n)$ be the smallest sample size such that $\beta_m^{(2)} \leqq \beta$ at the alternative $\theta = \theta_n$. By definition, $e_1 = \lim_{n\to\infty} n/m(n)$, provided the limit exists. In the present case, it is easily seen that

$$(5) \qquad e_1 = \lim_{\theta\to 0} e(\alpha, \beta, \theta) = (\nu_1/\nu_2)^2,$$

where ν_2 satisfies

$$(6) \quad 1 - \beta = \Phi(-p^{\frac{1}{2}}\nu_2)\Phi(-K_\alpha - q^{\frac{1}{2}}\nu_2) + \Phi(p^{\frac{1}{2}}\nu_2)\Phi(-K_\alpha + q^{\frac{1}{2}}\nu_2),$$

TABLE 1

Pitman efficiency of the split sample test

Size	1−power				
	10^{-1}	10^{-2}	10^{-4}	10^{-6}	$10^{-\infty}$
		$p = .3$			
10^{-1}	.779	.730	.620	.543	.300
10^{-2}	.781	.763	.737	.696	.300
10^{-4}	.748	.740	.732	.728	.300
10^{-6}	.732	.728	.723	.720	.300
$10^{-\infty}$	.700	.700	.700	.700	
		$p = .5$			
10^{-1}	.639	.604	.575	.562	.500
10^{-2}	.571	.555	.542	.536	.500
10^{-4}	.535	.529	.523	.520	.500
10^{-6}	.523	.520	.516	.515	.500
$10^{-\infty}$	.500	.500	.500	.500	
		$p = .7$			
10^{-1}	.391	.363	.345	.337	.300
10^{-2}	.343	.333	.325	.322	.300
10^{-4}	.321	.317	.314	.312	.300
10^{-6}	.314	.312	.310	.309	.300
$10^{-\infty}$	.300	.300	.300	.300	

where $q = 1 - p$. It is curious that e_1 is not independent of α and β (cf. Table 1). This is because some of the regularity conditions of Pitman's theory (in particular conditions C and D, as well as D', in [5]) are not satisfied.

In the present case, e_2 and e_3 may be expressed as $e_2 = \lim_{\beta\to 0} e(\alpha, \beta, \theta)$ and $e_3 = \lim_{\alpha\to 0} e(\alpha, \beta, \theta)$. In the next two paragraphs we will show directly that

$$e_2 = \min\{p, q\} \tag{7}$$

and

$$e_3 = q \tag{8}$$

for every $\theta \neq 0$. According to e_2 the best choice of p is $\frac{1}{2}$ but according to e_3 it is 0.

Suppose $\theta > 0$. (The case where $\theta < 0$ follows by symmetry.) Then, as $n \to \infty$,

$$\beta_n^{(1)} = \Phi(K_{\alpha/2} - n^{\frac{1}{2}}\theta)\,[1 + o(1)] \tag{9}$$

by (3). Also from (4), as $n \to \infty$,

$$\begin{aligned}\beta_n^{(2)} &= \Phi(K_\alpha - (nq)^{\frac{1}{2}}\theta)[1 + o(1)]\\ &\quad + \Phi(-(np)^{\frac{1}{2}}\theta)[1 + o(1)]\\ &= u_n^{\,n} + v_n^{\,n} \qquad \text{say.}\end{aligned} \tag{10}$$

It is easily seen (cf. [3], p. 166) that $\log u_n \to -q\theta^2/2$ and $\log v_n \to -p\theta^2/2$, as $n \to \infty$. Thus the exact dual slopes of $\{T_n^{(1)}\}$ and $\{T_n^{(2)}\}$ (defined as the limits of $-2n^{-1}\log \beta_n^{(i)}$, $i = 1, 2$, as $n \to \infty$, for fixed α and $\theta \neq 0$) are θ^2 and $\theta^2 \min\{p, q\}$, respectively. Hence the Hodges-Lehmann efficiency, the ratio of the exact dual slopes, is given by (7).

It follows from (1) and (2) that, when $\theta \neq 0$ obtains,

$$\begin{aligned} n^{-\frac{1}{2}}T_n^{(i)} &\to |\theta| \quad \text{if } i = 1\\ &\to q^{\frac{1}{2}}|\theta| \quad \text{if } i = 2\end{aligned} \tag{11}$$

with probability one as $n \to \infty$. It is easily seen that for $0 < t < \infty$, $i = 1, 2$,

$$-2n^{-1}\log P(T_n^{(i)} \geqq n^{\frac{1}{2}}t \mid \theta = 0) \to t^2, \tag{12}$$

as $n \to \infty$. Thus by [2] the exact slopes of $\{T_n^{(1)}\}$ and $\{T_n^{(2)}\}$ (defined, e.g., as the limits of $-2n^{-1}\log \alpha_n^{(i)}$, $i = 1, 2$, as $n \to \infty$, where $\alpha_n^{(i)}$ is the size of the ith test for given n, β, and $\theta \neq 0$) are θ^2 and $q\theta^2$, respectively. Hence the Bahadur efficiency, the ratio of the two exact slopes, is given by (8).

4. Discussion. 1. The present example is in contrast to the comparison of the sign test to the optimal test based on $\bar{X}_n$ for one-sided alternatives, where it is known (cf. [1] and [4]) that both the Hodges-Lehmann and Bahadur efficiencies approach the Pitman efficiency as $\theta \to 0$. Bahadur [1] gives regularity conditions under which the "approximate" Bahadur efficiency, which in the present case coincides with the exact one, approaches the Pitman value, as the alternative approaches the hypothesis being tested. However, some of these conditions are not satisfied by $\{T_n^{(2)}\}$. (In particular the limits of (xii) in [1] cannot be interchanged and condition (xiv) in [1] is not satisfied.) As noted earlier, the Pitman value does not exist (as a single number independent of α and β) in the present case.

2. It may be worthwhile to try to pinpoint the cause of the discrepancy between e_2 and e_3 by considering certain cases related to the present one in which there is no discrepancy. If it is known that $\theta \geqq 0$, and $\bar{Z}_n$ is compared with $\bar{X}_n$, then

$$e_1(\alpha, \beta) = e_2(\alpha, \theta) = e_3(\beta, \theta) = q \tag{13}$$

for all α, β, and $\theta \neq 0$. If the sign of θ is unknown, and $|\bar{Z}_n|$ is compared with $|\bar{X}_n|$, then again (13) holds. Suppose now that the sign of θ is unknown, and we are to compare a reasonable statistic based on $|\bar{Z}_n|$ and W_n = the signs of $\bar{Y}_n$ and $\bar{Z}_n$, with the optimal statistic $|\bar{X}_n|$. Since the present case is an intermediate one as regards the information available to the user of $(|\bar{Z}_n|, W_n)$, it would seem that efficient use of the latter statistic would have efficiency q. According to (8), $T_n^{(2)}$ is an efficient statistic based on $(|\bar{Z}_n|, W_n)$, but according to (7) it is not and one would do better by ignoring W_n and using $|\bar{Z}_n|$.

It is easy to see the analytical reasons for the discrepancy under discussion. If $\theta > 0$ then with probability one the sign of $\bar{Y}_n$ is correct, and hence the level attained by $T_n^{(2)}$ equals the level attained by $\bar{Z}_n$ for all sufficiently large n. (Cf. [2] for the relation between levels attained and exact slopes.) On the other hand, if $\theta > 0$ and if we look at powers with α held fixed, then in using $T_n^{(2)}$ rather than $|\bar{Z}_n|$ one has to take into account the probability of the sign of $\bar{Y}_n$ being wrong; this probability $\rightarrow 0$ but is asymptotically just as important as the probability of $\bar{Z}_n$ being too small.

3. It may be noted that $U_n^{(1)} = T_n^{(1)}/n^{\frac{1}{2}}$ and $U_n^{(2)} = T_n^{(2)}/(nq)^{\frac{1}{2}}$ are consistent estimates of $|\theta|$; the asymptotic efficiency of $U_n^{(2)}$ relative to $U_n^{(1)}$ can be computed e.g. by the methods described in Sections 1 and 2 of [2]. It is interesting that both the methods just cited lead to $q = e_3$ as the estimation efficiency at each θ.

5. Acknowledgments. I am indebted to W. H. Kruskal for introducing me to problems in data splitting. I wish to thank R. R. Bahadur for his encouragement and suggestions.

REFERENCES

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

[2] Bahadur, R. R. (1967). Rates of convergence of estimates and test statistics. *Ann. Math. Statist.* **38** 303–324.

[3] Feller, William (1957). *An Introduction to Probability Theory and its Applications*, **1** (2nd ed.). Wiley, New York.

[4] 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.

[5] Noether, G. E. (1955). On a theorem of Pitman. *Ann. Math. Statist.* **29** 64–68.