ON A THEOREM OF PITMAN

By Gottfried E. Noether

Boston University

Summary. A theorem by Pitman on the asymptotic relative efficiency of two tests is extended and some of its properties are discussed.

1. Introduction. The idea of the relative efficiency of one estimate with respect to another estimate of the same parameter is well established. This cannot be said, however, of the corresponding concept for two tests of the same statistical hypothesis. This paper is concerned with a definition of the relative efficiency of two tests which seems to be due to Pitman (see, e.g., [1] p. 241) and has been used in several recent papers.

Definition. Given two tests of the same size of the same statistical hypothesis, the *relative efficiency* of the second test with respect to the first is given by the ratio n_1/n_2, where n_2 is the sample size of the second test required to achieve the same power for a given alternative as is achieved by the first test with respect to the same alternative when using a sample of size n_1.

In general the ratio n_1/n_2 will depend on the particular alternative chosen (as well as on n_1). However, in the asymptotic case, this somewhat undesirable fact can be avoided. It might be argued that restriction to the asymptotic case is even more undesirable in itself, but the unfortunate fact remains that for many test procedures in current use the asymptotic power function is the only one available.

Now, at least for consistent tests, the power with respect to a fixed alternative is practically 1 if the number of observations is sufficiently large. Therefore, the power no longer provides a worthwhile criterion for preferring one test over another. On the other hand, it is possible to define sequences of alternatives changing with n in such a way that as $n \to \infty$ the power of the corresponding sequence of tests converges to some number less than 1. It seems then reasonable to define the asymptotic relative efficiency of the second test procedure with respect to the first test procedure as the limit of the corresponding ratios n_1/n_2.

A theorem due to Pitman allows us to compute this limit if certain general conditions are satisfied [1]. The purpose of this paper is to give an extension of Pitman's theorem and to discuss some of its properties. The derivation of the present version of the theorem follows Pitman's original method of proof. Since this proof has not appeared in print, full details are given.

2. Asymptotic power. Assume that we want to test the null hypothesis H_0: $\theta = \theta_0$ against alternatives H: $\theta > \theta_0$. As mentioned in the Introduction, we shall assume actually that a particular alternative $\theta = \theta_n$ changes with the sample size n in such a way that $\lim_{n\to\infty} \theta_n = \theta_0$.

Received April 7, 1954.

To be more definite, let the test be based on the static $T_n = T(x_1, \cdots, x_n)$. Let[1] $ET_n = \psi_n(\theta)$ and var $T_n = \sigma_n^2(\theta)$. Assume that

A. $$\psi_n'(\theta_0) = \cdots = \psi_n^{(m-1)}(\theta_0) = 0, \qquad \psi_n^{(m)}(\theta_0) > 0,$$

B. $$\lim_{n\to\infty} n^{-m\delta}\psi_n^{(m)}(\theta_0) \,/\, \sigma_n(\theta_0) = c > 0 \quad \text{for some } \delta > 0.$$

The indicated derivatives are assumed to exist. We shall consider the power of the test based on T_n with respect to the alternative H_1: $\theta_n = \theta_0 + k/n^\delta$, where k is an arbitrary positive constant. In addition to A and B we shall assume

C. $$\lim_{n\to\infty} \psi_n^{(m)}(\theta_n) \,/\, \psi_n^{(m)}(\theta_0) = 1, \qquad \lim_{n\to\infty} \sigma_n(\theta_n) \,/\, \sigma_n(\theta_0) = 1,$$

D. the distribution of $[T_n - \psi_n(\theta)] \,/\, \sigma_n(\theta)$ tends to the normal distribution with mean 0 and variance 1, uniformly in θ, with $\theta_0 \leqq \theta \leqq \theta_0 + d$ for some $d > 0$.

Let $\phi(\lambda) = \int_\lambda^\infty \exp(-\frac{1}{2}x^2)\, dx \,/\, \sqrt{2\pi}$ and find λ_α such that $\phi(\lambda_\alpha) = \alpha$. For sufficiently large n, a critical region of approximate size α is given by

$$T_n \geqq T_n(\alpha), \qquad [T_n(\alpha) - \psi_n(\theta_0)] \,/\, \sigma_n(\theta_0) = \lambda_\alpha .$$

The power of this test with respect to the alternative H_1 is given by

$$L_n(\theta_n) = P\{T_n \geqq T_n(\alpha) \mid \theta_n\} \sim \phi(t_n),$$

where $t_n = [\sigma_n(\theta_0)\lambda_\alpha + \psi_n(\theta_0) - \psi_n(\theta_n)] \,/\, \sigma_n(\theta_n)$. Now

$$\psi_n(\theta_n) = \psi_n(\theta_0 + k/n^\delta) = \psi_n(\theta_0) + (1/m!)(k/n^\delta)^m\psi_n^{(m)}(\hat{\theta}), \qquad \theta_0 < \hat{\theta} < \theta_n$$

and

$$t_n = \frac{\sigma_n(\theta_0)\lambda_\alpha - (1/m!)(k/n^\delta)^m[\psi_n^{(m)}(\hat{\theta})/\psi_n^{(m)}(\theta_0)]\psi_n^{(m)}(\theta_0)}{[\sigma_n(\theta_n)/\sigma_n(\theta_0)]\sigma_n(\theta_0)} \xrightarrow[n\to\infty]{} \lambda_\alpha - \frac{k^m c}{m!}.$$

Thus asymptotically, $L_n(\theta_n) \sim \phi(\lambda_\alpha - k^m c/m!)$.

It follows from the proof that condition (D) can be replaced by the somewhat weaker condition

D′. the distribution of $[T_n - \psi_n(\theta_n)] \,/\, \sigma(\theta_n)$ tends to the normal distribution with mean 0 and variance 1, both under the alternative hypothesis H_1 and under the null hypothesis $\theta_n = \theta_0$.

It is also clear that alternatives of the type $\theta_n < \theta_0$ or $|\theta_n| \neq \theta_0$, or the case when $\psi_n^{(m)}(\theta_0) < 0$, can be handled correspondingly.

Let $\gamma > \delta$ and consider the alternative H_2 : $\theta_n = \theta_0 + k/n^\gamma$ an alternative which converges to θ_0 faster than H_1. Now

$$t_n \sim \lambda_\alpha - \lim_{n\to\infty} \frac{k^m}{m!\,n^{\gamma m}} \frac{\psi_n^{(m)}(\theta_0)}{\sigma_n(\theta_0)} = \lambda_\alpha - \lim_{n\to\infty} \frac{k^m}{m!\,n^{(\gamma-\delta)m}} \frac{\psi_n^{(m)}(\theta_0)}{n^{\delta m}\sigma_n(\theta_0)} = \lambda_\alpha,$$

[1] J. Putter [4] has pointed out that the functions $\psi_n(\theta)$ and $\sigma_n^2(\theta)$ need not necessarily be the mean and variance of T_n, respectively, as long as conditions A, B, C, and D or D′ are satisfied.

and the power of our test with respect to this alternative H_2 is equal to the size of the critical region. The test cannot distinguish between H_0 and H_2. Similarly, if $\gamma < \delta$, the power of our test converges to 1.

3. Asymptotic relative efficiency of two tests. Assume now that we have two tests based on the statistics T_{1n} and T_{2n}. Assume further that $\delta_1 > \delta_2$ and consider the alternative $H_{1'}: \quad \theta_n = \theta_0 + k/n^{\delta_1}$. It follows from our previous results that the second test is ineffective with respect to this alternative, while the power of the first test can be made as large as we please by choosing k sufficiently large. Therefore, the asymptotic relative efficiency of the second test with respect to the first test is zero.

Thus, from now on, we assume that $\delta_1 = \delta_2 = \delta$.

According to our definition of relative efficiency, the two tests must have identical power with respect to identical alternatives. The two tests have identical power if

$$k_1^{m_1} c_1/m_1! = k_2^{m_2} c_2/m_2! \tag{1}$$

The alternatives are identical if

$$k_1/n_1^{\delta} = k_2/n_2^{\delta}. \tag{2}$$

If now $m_1 = m_2 = m$, as it must be if $m_i \delta = \frac{1}{2}$ for $i = 1$ and 2, which is true in most cases, we can proceed as follows. From (1) and (2) we have

$$\frac{n_1}{n_2} = \left(\frac{k_1}{k_2}\right)^{1/\delta} = \left(\frac{c_2}{c_1}\right)^{1/m\delta} = \lim_{n\to\infty} \frac{(1/n)[\psi_{2n}^{(m)}(\theta_0) \,/\, \sigma_{2n}(\theta_0)]^{1/m\delta}}{(1/n)[\psi_{1n}^{(m)}(\theta_0) \,/\, \sigma_{1n}(\theta_0)]^{1/m\delta}} = \lim_{n\to\infty} \frac{R_{2n}^{1/m\delta}(\theta_0)}{R_{1n}^{1/m\delta}(\theta_0)} = E_{21}$$

where

$$R_{in}(\theta) = \psi_{in}^{(m)}(\theta) \,/\, \sigma_{in}(\theta), \qquad i = 1, 2. \tag{3}$$

Pitman has called the quantity $R_{in}^{1/m\delta}(\theta_0)$ the *efficacy* of the ith test in testing the hypothesis $H_0 : \theta = \theta_0$. Thus we get

PITMAN'S THEOREM. *The asymptotic relative efficiency of two tests satisfying* A, B, C, *and* D *or* D′, *with* $\delta_1 = \delta_2$ *and* $m_1 = m_2$, *is given by the limit of the ratio of the efficacies of the two tests.*

For $m = 1$ and $\delta = \frac{1}{2}$, this theorem reduces to the one quoted in [1]. *If* $m\delta = \frac{1}{2}$,

$$E_{21} = \lim_{n\to\infty} R_{2n}^2(\theta_0) \,/\, R_{1n}^2(\theta_0). \tag{4}$$

If, in addition,

$$\lim_{n\to\infty} \psi_{2n}^{(m)}(\theta_0) \,/\, \psi_{1n}^{(m)}(\theta_0) = 1, \tag{5}$$

then (4) reduces to $E_{21} = \lim_{n\to\infty} \sigma_{1n}^2/\sigma_{2n}^2$.

This is the usual expression for measuring the asymptotic relative efficiency of two estimates of the same parameter. Thus, only if (5) is satisfied can we use the ratio of the variances of the two test statistics as a measure of the asymptotic relative efficiency of the two tests. In particular, if T_{1n} and T_{2n} are unbiased estimates of the parameter θ, (5) is satisfied with $m = 1$, and E_{21} is the same as the asymptotic relative efficiency of T_2 and T_1 used as estimators of θ.

4. Comparison with another definition of relative efficiency. Another definition of the relative efficiency of two tests in current use (see, e.g., [2] p. 597) is based on the ratio of the respective sample sizes under the assumption that the power functions of the two tests have equal slope at $\theta = \theta_0$.

We shall show that if $m = 1$ and $\delta = \frac{1}{2}$, the two definitions give the same value for the asymptotic relative efficiency, provided a very general condition is satisfied.

Under the conditions of Pitman's theorem we have, as before, $L_n(\theta) \sim \phi(t_n)$ where

$$t_n = [\lambda_\alpha \sigma_n(\theta_0) + \psi_n(\theta_0) - \psi_n(\theta)] \,/\, \sigma_n(\theta).$$

If $L_n'(\theta)$ converges uniformly to some limit, this limit must be $d\phi(t_n)/d\theta$. Actually, the exact form of $L_n'(\theta)$ will rarely be known, so that the uniform convergence cannot be investigated. However, even in this case it is customary to replace $L_n'(\theta)$ by $d\phi(t_n)/d\theta$ in computing the ratio of the slopes of the two power functions. Thus it seems reasonable to compare the asymptotic relative efficiency based on the slopes $d\phi(t_n)/d\theta$ with E_{21}. Now

$$\left.\frac{d\phi(t_n)}{d\theta}\right|_{\theta=\theta_0} = \frac{-1}{\sqrt{2\pi}} \exp\left(-\tfrac{1}{2}t_n^2\right) \left.\frac{\partial t_n}{\partial \theta}\right|_{\theta=\theta_0}$$

$$= \frac{1}{\sqrt{2\pi}} \exp\left(-\tfrac{1}{2}\lambda_\alpha^2\right) \left[\frac{\psi_n'(\theta_0)}{\sigma_n(\theta_0)} + \lambda_\alpha \frac{\sigma_n'(\theta_n)}{\sigma_n(\theta_0)}\right] \sim \frac{1}{\sqrt{2\pi}} \exp\left(-\tfrac{1}{2}\lambda_\alpha^2\right) c n^{1/2},$$

provided $\sigma_n'(\theta_0)/\sigma_n(\theta_0) = o(\sqrt{n})$, which is very generally true. The requirement that the two power functions have equal slope at the point $\theta = \theta_0$ becomes $c_1\sqrt{n_1} = c_2\sqrt{n_2}$, so that the asymptotic relative efficiency according to this definition is again given by

$$(n_1/n_2) = (c_2/c_1)^2 = E_{21}.$$

5. Efficacy of a test. Still assuming that $m = 1$ and $\delta = \frac{1}{2}$, it is interesting to investigate more closely the efficacy R_n^2, where R_n is given by (3). Consider the function $t = \psi_n(u)$ determined by $E\,T_n = \psi_n(\theta)$. Unless $\psi_n(\theta) = \theta$, T_n is not an unbiased estimate of θ, but may be considered an unbiased estimate of the ficticious parameter $\tau = \psi_n(\theta)$.

Let $u = \varphi_n(t)$ be the inverse of $t = \psi_n(u)$ and define the statistic $U_n = \varphi_n(T_n)$. Then we may write

$$U_n - \theta = \varphi_n(T_n) - \varphi_n(\tau) = (T_n - \tau)\,\varphi_n'(\tau) + \cdots.$$

If it is permissible to neglect terms of higher order in $T_n - \tau$, we find $E\,U_n \sim \theta$ and

$$\operatorname{var} U_n \sim [\varphi_n'(\tau)]^2 \sigma_n^2(\theta) = \sigma_n^2(\theta) \,/\, [\psi_n'(\theta)]^2 = R_n^{-2}(\theta).$$

Thus, if the above conditions are satisfied, asymptotically the efficacy of T_n is the reciprocal of the variance of an asymptotically unbiased estimate[2] of θ based on T_n.

Now, under the regularity conditions for the Cramér-Rao inequality, var $U_n \geqq 1/nE(\partial \log f/\partial\theta)^2$. Therefore

$$R_n^2(\theta) \leqq nE(\partial \log f/\partial\theta)^2.$$

Thus we may use the quantity $R_n^2(\theta) \,/\, nE(\partial \log f/\partial\theta)^2$ as a measure of the asymptotic efficiency of the test based on the statistic T_n.

REFERENCES

[1] G. E. NOETHER, "Asymptotic properties of the Wald-Wolfowitz test of randomness," *Ann. Math. Stat.*, Vol. 21 (1950), pp. 231–246.

[2] N. BLOMQVIST, "On a measure of dependence between two random variables," *Ann. Math. Stat.*, Vol. 21 (1950), pp. 593–600.

[3] A. STUART, "Asymptotic relative efficiencies of distribution-free tests of randomness against normal alternatives," *J. Amer. Stat. Assn.*, Vol. 49 (1954), pp. 147–157.

[4] J. PUTTER, "The treatment of ties in some nonparametric tests." Ph.D. thesis.

[2] Essentially this same result has also been obtained by Stuart [3]. However Stuart uses it even in some cases for which $m = 1$ but $\delta \neq \frac{1}{2}$. That it is then no longer correct can be seen easily from the generalized form of Pitman's Theorem given in Section 3 of this paper.