# THE EFFICIENCY OF TESTS

by

Wassily Hoeffding
University of North Carolina

Institute of Statistics
Mimeograph Series No. 99
Limited Distribution
April 1954

---

[1]This research was supported by the United States Air Force, through the Office of Scientific Research of the Air Research and Development Command.

UNCLASSIFIED
Security Information

Bibliographical Control Sheet

1. O.A.: Institute of Statistics, North Carolina State College of the University of North Carolina

   M.A.: Office of Scientific Research of the Air Research and Development Command

2. O.A.: CIT Report No. 6
   M.A.: OSR Technical Note -

3. THE EFFICIENCY OF TESTS (UNCLASSIFIED)

4. Hoeffding, Wassily

5. April, 1954

6. 21

7. None

8. AF 18(600)-458

9. RDO No. R-354-20-8

10. UNCLASSIFIED

11. None

12. The efficiency of a family of tests is defined. Methods for evaluating the efficiency are discussed. The asymptotic efficiency is obtained for certain families of tests under assumptions which imply that the sample size is large.

# THE EFFICIENCY OF TESTS[1]

By Wassily Hoeffding, University of North Carolina

Summary. The efficiency of a family of tests is defined. Methods for evaluating the efficiency are discussed. The asymptotic efficiency is obtained for certain families of tests under assumptions which imply that the sample size is large.

1. Introduction. Let $w_n^{(1)}$ and $w_n^{(2)}$ be two tests each of which has the same fixed significance level for testing a hypothesis $\theta = \theta_1$, and which are defined for every sample size n. The relative efficiency of test $w_n^{(2)}$ (or rather, of the sequence $\{w_n^{(2)}\}$) with respect to test $w_n^{(1)}$ has been defined as the ratio $n_1/n_2$, where $n_2$ is the least sample size required by a test of the sequence $\{w_n^{(2)}\}$ in order to achieve the same power for a given alternative $\theta = \theta_2$ as is achieved by the test in $\{w_n^{(1)}\}$ which uses a sample of size $n_1$. This is essentially the definition used by Pitman (see Noether [1], p. 241).

In section 2 we extend this definition by replacing sequences of tests by arbitrary families of (nonsequential) tests and the parameters $\theta_1$, $\theta_2$ by two arbitrary classes of distributions. The tests are regarded as general two-decision rules. If $N(\underline{\underline{T}})$ is the least sample size used by a test in family $\underline{\underline{T}}$ whose probabilities of the two kinds of error (corresponding to the two classes of distributions) do not exceed two given numbers, the

1. This research was supported by the United States Air Force, through the office of Scientific Research of the Air Research and Development Command.

ratio $N(\underline{\underline{T}}_1)/N(\underline{\underline{T}}_2)$ is defined as the relative efficiency of family $\underline{\underline{T}}_2$ with respect to family $\underline{\underline{T}}_1$.

In section 3 it is pointed out that the determination of $N(\underline{\underline{T}})$ is closely related to finding a test which maximizes the minimum power.

In section 4 the problem of asymptotic efficiency is considered. In studies of the asymptotic efficiency of tests it is customary to consider a simple hypothesis $\theta = \theta_1$ (say) and a simple alternative, $\theta = \theta_2$, and to assume that as $n \longrightarrow \infty$, $\theta_1$ remains fixed and $\theta_2$ approaches $\theta_1$ in a certain way, for instance by setting $\theta_2 = \theta_1 + kn^{-1/2}$. Neither the restriction to simple hypotheses nor the assumption that $\theta_2$ depends on n seems to be entirely adequate from the point of view of most applications.

In this paper a different approach is used. In section 4 we assume that one or the other decision is undesirable according as $\theta \leq \theta_1$ or $\theta \geq \theta_2$, where $\theta_1$ and $\theta_2 > \theta_1$ are two fixed numbers. Since $\theta_1$ and $\theta_2$ will usually be so chosen that neither decision is strongly preferred when $\theta_1 < \theta < \theta_2$, small values of $\theta_2 - \theta_1$ are of particular interest. We consider families of tests which depend on statistics whose distributions are determined by the parameter $\theta$. Let $N(\underline{\underline{T}})$ denote the least sample size used by a test in family $\underline{\underline{T}}$ for which the probability of a wrong decision does not exceed $\alpha_1$ if $\theta \leq \theta_1$ and does not exceed $\alpha_2$ if $\theta \geq \theta_2$. Thus $N(\underline{\underline{T}})$ is a function of $\theta_1$ and $\theta_2$. Under suitable assumptions we derive asymptotic expressions for $N(\underline{\underline{T}})$ when $\delta = \theta_2 - \theta_1$ tends to zero, while $\alpha_1$ and $\alpha_2$ remain fixed.

2. <u>The efficiency of a family of tests</u>. Let X be a random variable or a random vector with cdf (cumulative distribution function) F, which is assumed to belong to a class $\underline{\underline{C}}$ of cdfs. It is desired to decide, on the

basis of n independent observations $\underline{x}_n = (x_1, \ldots, x_n)$ on X, between two alternative courses of action, $A_1$ and $A_2$. Let $d_i$ denote the decision in favor of $A_i$ $(i = 1, 2)$. Suppose there are given two disjoint subclasses $\underline{\underline{C}}_1$ and $\underline{\underline{C}}_2$ of $\underline{\underline{C}}$ such that decision $d_i$ is preferred if F is in $\underline{\underline{C}}_i$ $(i = 1, 2)$. (In most applications one will choose the classes $\underline{\underline{C}}_1$ and $\underline{\underline{C}}_2$ in such a way that neither decision is strongly preferred if F is neither in $\underline{\underline{C}}_1$ nor in $\underline{\underline{C}}_2$.)

Consider a decision rule, briefly referred to as a test, of the following type. Let $E_n$ denote the space of points $\underline{x}_n$. (Thus if $x_j$ is a vector with k components, $E_n$ may be taken as the nk-dimensional Euclidean space.) Let $w_{1n}$ be a (Borel-) subset of $E_n$, and denote its complement by $w_{2n}$. (All sets of points $\underline{x}_n$ considered in this paper are assumed to be Borel sets, and all functions of $\underline{x}_n$ are understood to be Borel-measurable. This will usually not be explicitly stated.) A test determined by the pair of sets $w_n = (w_{n1}, w_{n2})$ consists in taking n observations on X and taking decision $d_i$ if the observed point $\underline{x}_n$ is in $w_{in}$, $i = 1, 2$. This test will be referred to as the test $w_n$. A test $w_n$ with $w_{in} \subset E_n$ will be called a test based on n observations.

Let $\alpha_1$, $\alpha_2$ be two positive numbers. We shall say that the test $w_n$ solves the problem $(\underline{\underline{C}}_1, \underline{\underline{C}}_2, \alpha_1, \alpha_2)$ if

$$(2.1) \qquad P(\underline{X}_n \in w_{in} \mid F) \geq 1 - \alpha_i \text{ for all } F \text{ in } \underline{\underline{C}}_i,\ i = 1, 2,$$

where $\underline{X}_n = (X_1, \ldots, X_n)$ is a random vector with values in $E_n$, and $P(R \mid F)$ denotes the probability of relation R when $X_1, \ldots, X_n$ are independent with common cdf F.

Let $\underline{\underline{T}}$ be a family of tests. We shall mainly be concerned with families $\underline{\underline{T}}$ which, for every positive integer n, contain at least one test based on n observations. For example, $\underline{\underline{T}}$ may be the family of all tests $w_n$ with $w_{1n} = \{ \underline{x}_n : t_n(\underline{x}_n) < c \}$, $-\infty < c < \infty$, $n = 1, 2, \ldots$, where $\{ t_n \}$ is a given sequence of functions. (Here $\{ \underline{x}_n : R \}$ denotes the set of all points $\underline{x}_n$ such that relation R is satisfied.) Let $N(\underline{\underline{T}}) = N(\underline{\underline{T}}, \underline{\underline{C}}_1, \underline{\underline{C}}_2, \alpha_1, \alpha_2)$ be the least integer n such that the inequalities (2.1) are satisfied for some test in $\underline{\underline{T}}$. Thus $N(\underline{\underline{T}})$ is the least number of observations with which problem $(\underline{\underline{C}}_1, \underline{\underline{C}}_2, \alpha_1, \alpha_2)$ can be solved when we restrict ourselves to tests belonging to family $\underline{\underline{T}}$. $N(\underline{\underline{T}})$ will be termed the efficiency index, or simply the index, of family $\underline{\underline{T}}$ for problem $(\underline{\underline{C}}_1, \underline{\underline{C}}_2, \alpha_1, \alpha_2)$. Evidently $\underline{\underline{T}}$ may contain more than one test based on $N(\underline{\underline{T}})$ observations.

If $\underline{\underline{T}}_1$ and $\underline{\underline{T}}_2$ are two families of tests and at least one of the indices $\underline{\underline{N}}(\underline{\underline{T}}_i) = N(\underline{\underline{T}}_i, \underline{\underline{C}}_1, \underline{\underline{C}}_2, \alpha_1, \alpha_2)$, $i = 1, 2$, is finite, the ratio

$$\text{eff } (\underline{\underline{T}}_1/\underline{\underline{T}}_2) = N(\underline{\underline{T}}_2)/N(\underline{\underline{T}}_1)$$

will be called the efficiency of family $\underline{\underline{T}}_1$ relative to family $\underline{\underline{T}}_2$ for problem $(\underline{\underline{C}}_1, \underline{\underline{C}}_2, \alpha_1, \alpha_2)$. The relative efficiency can thus equal any nonnegative rational number or infinity.

This definition of relative efficiency of two tests is an extension of Pitman's definition of the same term; see Noether [1], p. 241.

Let $\underline{\underline{T}}^*$ be the family of all tests $w_n = (w_{1n}, w_{2n})$, for all $n = 1, 2, \ldots$ . If $N(\underline{\underline{T}}^*) = \infty$, problem $(\underline{\underline{C}}_1, \underline{\underline{C}}_2, \alpha_1, \alpha_2)$ has no solution (within the family $\underline{\underline{T}}$*). If $N(\underline{\underline{T}}*) < \infty$, and $\underline{\underline{T}}$ is any subfamily of $\underline{\underline{T}}*$, then eff(T/T*) may be called the (absolute) efficiency of family $\underline{\underline{T}}$ for problem $(\underline{\underline{C}}_1, \underline{\underline{C}}_2, \alpha_1, \alpha_2)$ (provided we confine ourselves to tests in $\underline{\underline{T}}*$). Clearly

$$\text{eff}\ (\underline{\underline{T}}/\underline{\underline{T}}*) \leq 1,$$

the sign of equality holding if and only if $\underline{\underline{T}}$ contains a test which solves problem $(\underline{\underline{C}}_1, \underline{\underline{C}}_2, \alpha_1, \alpha_2)$ and is based on the least possible number of observations with which the problem can be solved by any test in $\underline{\underline{T}}*$.

All that has been said can immediately be extended to families of randomized tests. A randomized test is determined by a function $\phi_{1n}(\underline{x}_n)$, $0 \leq \phi_{1n}(\underline{x}_n) \leq 1$. Let $\phi_{2n} = 1 - \phi_{1n}$. The test consists in taking n observations $\underline{x}_n$ and performing a random experiment whose two possible outcomes, $e_1$ and $e_2$, have probabilities $\phi_{1n}(\underline{x}_n)$ and $\phi_{2n}(\underline{x}_n)$, respectively. If event $e_i$ occurs, decision $d_i$ is taken. A test determined by the pair of functions $\phi_n = (\phi_{1n}, \phi_{2n})$ will be referred to as the test $\phi_n$. If $\underline{\underline{T}}$ is a family of randomized tests, the index $N(\underline{\underline{T}}, \underline{\underline{C}}_1, \underline{\underline{C}}_2, \alpha_1, \alpha_2)$ is defined as the least n such that the relations

$$E(\phi_{in} \mid F) \geq 1 - \alpha_i \quad \text{for all F in } \underline{\underline{C}}_i, \qquad i=1,2,$$

are satisfied for some $\phi_n$ in $\underline{\underline{T}}$. Here $E(\phi_{in} \mid F)$ denotes the expected value of $\phi_{in}(\underline{X}_n)$ when $X_1, \ldots, X_n$ are independent with the common cdf. F.

If randomized tests are admitted which do not use any observations, we could have $N(\underline{T}) = 0$. This trivial case can be excluded by assuming that $\alpha_1 + \alpha_2 < 1$.

One could extend the notion of efficiency to families of tests such that the choice of the number of observations also depends on a random experiment, the probabilities of whose outcomes may or may not depend on the observations. (In the former case we are dealing with sequential tests.) In this paper we confine ourselves to families of nonsequential tests based on a nonrandom number of observations.

So far we have assumed, to simplify the exposition, that $X_1, X_2, \ldots,$ is a sequence of independent, identically distributed random variables. Suppose, more generally, that for every n the random vector $\underline{X}_n = (X_1, \ldots, X_n)$ has a cdf $G_n$ which belongs to a class $\underline{\underline{C}}_n$. For every n there are given two disjoint subclasses, $\underline{\underline{C}}_{1n}$ and $\underline{\underline{C}}_{2n}$, of $\underline{\underline{C}}_n$. We say that a test $w_n = (w_{1n}, w_{2n})$ solves problem $(\{\underline{\underline{C}}_{1n}\}, \{\underline{\underline{C}}_{2n}\}, \alpha_1, \alpha_2)$ if

$$P(\underline{X}_n \varepsilon\ w_{in} \mid G_n) \geq 1 - \alpha_i \text{ for all } G_n \text{ in } \underline{\underline{C}}_{in}, \qquad i=1,2,$$

where $P(R \mid G_n)$ denotes the probability of relation R when $\underline{X}_n$ has the cdf $G_n$. The definitions of $N(\underline{\underline{T}}) = N(\underline{\underline{T}}, \{\underline{\underline{C}}_{1n}\}, \{\underline{\underline{C}}_{2n}\}, \alpha_1, \alpha_2)$ and $\mathrm{eff}(\underline{\underline{T}}_1/\underline{\underline{T}}_2)$ are obvious extensions of the corresponding definitions in the special case.

3. The determination of $N(\underline{\underline{T}})$. Let $\underline{\underline{T}}$ be a given family of tests, randomized or not. Since a randomized test $\phi_n$ such that $\phi_n(\underline{x}_n) = 0$ or 1

for all $\underline{x}_n$ is equivalent to a nonrandomized test $w_n$, we may denote the tests in $\underline{\underline{T}}$ by $\phi_n$. Let $\underline{\underline{T}}_n$ be the family of all tests in $\underline{\underline{T}}$ which are based on n observations. Denote by $\underline{\underline{T}}_n^*$ the family of all tests $\phi_n$ in $\underline{\underline{T}}_n$ for which

$$(3.1) \qquad E(\phi_{1n} \mid G_n) \geq 1 - \alpha_1 \qquad \text{for all } G_n \text{ in } \underline{\underline{C}}_{1n}.$$

Let

$$M(\phi_n) = \sup_{G_n \varepsilon \underline{\underline{C}}_{2n}} E(\phi_{1n} \mid G_n),$$

$$M_n = \inf_{\phi_n \varepsilon \underline{\underline{T}}_n^*} M(\phi_n) .$$

The following theorem is self-evident.

<u>Theorem 3.1.</u> $N(\underline{\underline{T}}, \{\underline{\underline{C}}_{1n}\}, \{\underline{\underline{C}}_{2n}\}, \alpha_1, \alpha_2)$ <u>is the least integer n for which</u> $M_n \leq \alpha_2$.

Adapting a familiar terminology, we may say that a test which satisfies (3.1) has level $\alpha_1$ with respect to $\underline{\underline{C}}_{1n}$, and we may call $1-M(\phi_n)$ the minimum power of test $\phi_n$ with respect to $\underline{\underline{C}}_{2n}$. If there exists a test $\phi_n$ in $\underline{\underline{T}}_n^*$ such that $M(\phi_n) = M_n$, the test is said to maximize the minimum power with respect to $\underline{\underline{C}}_{2n}$. Tests which maximize the minimum power can sometimes be obtained by a method due to Wald and explicitly applied to this problem by Lehmann [2], Theorem 8.3. For a proof of a special case of the theorem and for illustrations of its use see Lehmann and Stein [3], [4].

4. Asymptotic efficiency. We shall now consider the asymptotic behavior of the efficiency index $N(\underline{\underline{T}})$ for certain families of tests in cases where the "distance" between the classes $\underline{\underline{C}}_{1n}$ and $\underline{\underline{C}}_{2n}$ is small. Let $\theta(G_n)$ be a real-valued function defined for all $G_n$ in $\underline{\underline{C}}_n$ and for every $n=1,2,\ldots$ (or for every sufficiently large n). We assume that the set $\omega$ of values $\theta(G_n)$ when $G_n \varepsilon \underline{\underline{C}}_n$, $n=1,2,\ldots$, is an interval, finite or infinite. For example, $\underline{\underline{C}}_n$ may be the class of all cdfs of the form $G_n(\underline{x}_n)=F(x_1)F(x_2)\ldots F(x_n)$, where F belongs to some class $\underline{\underline{C}}$; in this case $\theta(G_n)=\theta_1(F)$ is a function of F only, such as the mean or $\theta_1(F) = F(0)$. As another example, let $\underline{\underline{C}}_n$ be the class of cdfs $G_n(\underline{x}_n)=F_1(x_1)\ldots F_1(x_m)F_2(x_{m+1})\ldots F_2(x_n)$, where $F_1$ and $F_2$ are normal cdfs with common variance $\sigma^2$ and respective means $\mu_1$ and $\mu_2$, and $m = m(n)$ is a given function of n, and let $\theta(G_n) = (\mu_1 - \mu_2)/\sigma$.

Let $\theta_1$ and $\theta_2$ be two numbers in $\omega$, $\theta_1 < \theta_2$, and let $\underline{\underline{C}}_{1n}$ and $\underline{\underline{C}}_{2n}$ be the classes of all $G_n$ in $\underline{\underline{C}}_n$ such that $\theta(G_n) \leq \theta_1$ and $\theta(G_n) \geq \theta_2$, respectively.

Let $\{t_n(\underline{x}_n)\}$ be a given sequence of functions, defined for all $n=1,2,\ldots$, and suppose that the distribution of $t_n(\underline{X}_n)$ depends on $G_n$ only through $\theta(G_n)$. We shall write $P(t_n \varepsilon S \mid \theta)$ for $P(t_n(\underline{X}_n) \varepsilon S \mid G_n)$ when $\theta(G_n) = \theta$.

Let $\underline{\underline{T}}$ be the family of all tests $w_n$ with

$$(4.1) \qquad w_{1n} = \{ \underline{x}_n : \ t_n(\underline{x}_n) \leq c \}, \quad -\infty < c < \infty \quad n=1,2,\ldots$$

We shall consider the asymptotic behavior of $N(\underline{T})=N(\underline{\underline{T}}, \{\underline{\underline{C}}_{1n}\}, \{\underline{\underline{C}}_{2n}\}, \alpha_1, \alpha_2)$ when $\delta = \theta_2 - \theta_1$ is small. Apart from the assumptions made so far we shall make four special assumptions A, B, C, D.

Assumption A. For every $n$ and every $c$, $P(t_n \leq c \mid \theta)$ is a continuous and nonincreasing function of $\theta$ for $\theta \varepsilon \omega$.

The nonincreasing character of $P(t_n \leq c \mid \theta)$ implies that the inequalities

$$P(\underline{X}_n \varepsilon w_{in} \mid G_n) \geq 1 - \alpha_i \text{ for all } G_n \text{ in } \underline{\underline{C}}_{in}, \quad i=1,2,$$

will be satisfied by test (4.1) if and only if

$$(4.2) \qquad P(t_n \leq c \mid \theta_1) \geq 1 - \alpha_1 , \qquad P(t_n \leq c \mid \theta_2) \leq \alpha_2.$$

Let $c_n$ be the smallest number which satisfies the inequality

$$P(t_n \leq c_n \mid \theta_1) \geq 1 - \alpha_1$$

This number exists since $P(t_n \leq c \mid \theta_1)$ is a nondecreasing function of $c$, continuous on the right.

As an application of Theorem 3.1 we obtain

Lemma 4.1. $N(\underline{\underline{T}})$ is the least integer $n$ for which

$$(4.3) \qquad P(t_n \leq c_n \mid \theta_2) \leq \alpha_2.$$

Assumption B. $\alpha_1 + \alpha_2 < 1$.

From Assumption B and the continuity assumption in A we obtain

Lemma 4.2. If $\theta_1$ is fixed and $\delta = \theta_2 - \theta_1 \longrightarrow 0$, then $N(\underline{T}) \longrightarrow \infty$.

Proof. Given an integer $n'$, we can choose $\delta_1 > 0$ so small that

$$P(t_n \leq c_n \mid \theta_1) - P(t_n \leq c_n \mid \theta_1 + \delta) < 1 - \alpha_1 - \alpha_2$$

if $\delta < \delta_1$, for $n \leq n'$. Then $N(\underline{T}) > n'$.

Assumption C. There exist a positive number $r$, a sequence of everywhere increasing functions $f_n(t_n)$ of $t_n$, $n=1,2,\ldots$, and a family of cdfs $H_d(x)$, defined for every $d \geq 0$, such that for every real $x$ and every $d \geq 0$

$$\lim_{n \to \infty} P\left\{ f_n(t_n) \leq x \mid \theta_1 + d\, n^{-r} \right\} = H_d(x). \tag{4.4}$$

It follows from Assumption A that $H_d(x)$ is a nonincreasing function of $d$. We have to assume more.

Assumption D. $H_d(x)$ is continuous in $x$ and increasing at every $x$ except where $H_d(x) = 0$ or $H_d(x) = 1$. Furthermore, for every $x$ such that $0 < H_0(x) < 1$, $H_d(x)$ is a continuous, everywhere decreasing function of $d$, and

$$\lim_{d \to \infty} H_d(x) = 0.$$

The first part of Assumption D implies that there exists an inverse function $H_d^{-1}(u)$, uniquely defined for $0 < u < 1$, so that $H_d^{-1}(u) = x$ if and only if $H_d(x) = u$. Assumption D further implies that if we set

$$a = H_0^{-1}(1-\alpha_1), \tag{4.5}$$

the equation in D, $H_D(a) = \alpha_2$ has a unique nonnegative root, which will be denoted by $D(\alpha_1,\alpha_2)$. By Assumption B, $D(\alpha_1,\alpha_2) > 0$. Thus $D(\alpha_1,\alpha_2)$ is the unique positive root of

$$H_D^{-1}(\alpha_2) = H_0^{-1}(1-\alpha_1). \tag{4.6}$$

<u>Theorem 4.1.</u> <u>Let</u> $\underline{T}$ <u>be the family of tests</u> (4.1). <u>If Assumptions</u> A, B, C, D <u>are satisfied, then for</u> $\theta_1$ <u>fixed and</u> $\delta \longrightarrow 0$ <u>we have asymptotically</u>

$$N(\underline{\underline{T}}) \sim \left\{ \delta^{-1} D(\alpha_1,\alpha_2) \right\}^{1/r}. \tag{4.7}$$

<u>Proof</u>. Since $t_n' = f_n(t_n)$ is an increasing function of $t_n$, the inequality $t_n \leq c_n$ is equivalent to $t_n' \leq c_n'$, where $c_n' = f_n(c_n)$. To simplify the notation we substitute $t_n'$, $c_n'$ for $t_n$, $c_n$ and omit the primes. Thus equation (4.4) is replaced by

$$\lim_{n \longrightarrow \infty} P(t_n \leq x \mid \theta_1 + d\, n^{-r}) = H_d(x). \tag{4.8}$$

Let $\varepsilon$ be a given small positive number,

$$\varepsilon < D, \tag{4.9}$$

where $D = D(\alpha_1,\alpha_2)$.

Let a be defined by (4.5). The quantity

$$2\gamma = \min \left\{ H_D(a) - H_{D+\varepsilon}(a),\ H_{D-\varepsilon}(a) - H_D(a) \right\} \tag{4.10}$$

is positive by Assumption D.

Choose $\varepsilon_1 > 0$ so small that

$$| H_{D\pm\varepsilon}(x) - H_{D\pm\varepsilon}(a) | < \gamma \qquad \text{if} \quad |x-a| \leq \varepsilon_1 \tag{4.11}$$

It follows from Assumption D that $H_d^{-1}(u)$ is continuous in u for $0 < u < 1$. Choose $\eta > 0$ so small that

$$\eta < \gamma , \tag{4.12}$$

$$|H_0^{-1}(u) - H_0^{-1}(1-\alpha_1) | < \varepsilon_1 \qquad \text{if} \quad |u-(1-\alpha_1) | < \eta. \tag{4.13}$$

Let

$$\eta_n(d) = P(t_n \leq c_n \mid \theta_1 + d\, n^{-r}) - H_d(c_n). \tag{4.14}$$

Since the cdf $H_d(x)$ is continuous in x, the convergence in (4.8) is uniform in x. Hence we can choose $n_1$ so large that

$$(4.15) \qquad |\eta_n(0)| < \eta \quad \text{and} \quad |\eta_n(D \pm \varepsilon)| < \eta \qquad \text{if } n > n_1.$$

In addition we may assume that

$$(4.16) \qquad \left(\frac{n}{n-1}\right)^r < 1 + \varepsilon \qquad \text{if } n > n_1.$$

Finally, by Lemma 4.2, we can choose $\delta_1 > 0$ so small that, with $N=N(\underline{\underline{T}})$,

$$(4.17) \qquad N > n_1 + 1 \qquad \text{if } \delta < \delta_1.$$

Recalling the definition of $c_n$ and making use of (4.14) and (4.15), we have for $n > n_1$

$$1 - \alpha_1 \leq P(t_n \leq c_n \mid \theta_1) < H_0(c_n) + \eta,$$

$$1 - \alpha_1 \geq P(t_n < c_n \mid \theta_1) > H_0(c_n) - \eta.$$

Thus $|H_0(c_n) - (1-\alpha_1)| < \eta$, and hence, by (4.13),

$$(4.18) \qquad |c_n - a| < \varepsilon_1 \quad \text{if } n > n_1.$$

The theorem will be proved if we show that for $\delta$ sufficiently small

$$D - \varepsilon \leq \delta N^r \leq (D+\varepsilon)(1+\varepsilon) \ . \tag{4.19}$$

By (4.16) and (4.17) this will hold if $\delta < \delta_1$ and

$$(D-\varepsilon) N^{-r} \leq \delta \leq (D+\varepsilon)(N-1)^{-r} \ . \tag{4.20}$$

We shall establish the theorem by showing that the assumption that (4.20) is false leads to a contradiction when $\delta < \delta_1$.

By Lemma 4.1 we have

$$P(t_N \leq c_N \mid \theta_1 + \delta) \leq \alpha_2 < P(t_{N-1} \leq c_{N-1} \mid \theta_1 + \delta) \ . \tag{4.21}$$

First assume that $\delta < \delta_1$ and $\delta < (D-\varepsilon) N^{-r}$, so that the first inequality (4.20) is violated. Then, applying successively Assumption A, (4.17), (4.15), (4.18), (4.12), (4.11), (4.10) and (4.6),

$$P(t_N \leq c_N \mid \theta_1 + \delta) \geq P(t_N \leq c_N \mid \theta_1 + (D-\varepsilon) N^{-r})$$

$$> H_{D-\varepsilon}(c_N) - \eta > H_{D-\varepsilon}(a-\varepsilon_1) - \gamma$$

$$> H_{D-\varepsilon}(a) - 2\gamma \geq H_D(a) = \alpha_2.$$

But this contradicts (4.21). Hence the first inequality (4.20) is true for $\delta < \delta_1$.

In a similar way we can show that if $\delta < \delta_1$ and $(D+\varepsilon)(N-1)^{-r} < \delta$, then $P(t_{N-1} \leq c_{N-1} \mid \theta_1 + \delta) < \alpha_2$, which contradicts (4.21). This completes the proof.

Let $\{t_{1n}\}$ and $\{t_{2n}\}$ be two sequences of statistics, and denote by $\underline{\underline{T}}_1$ and $\underline{\underline{T}}_2$ the corresponding families of tests of the form (4.1). Suppose that the distributions of $t_{1n}$ and $t_{2n}$ depend only on $\theta = \theta(G_n)$ and that Assumptions A, B, C, D are satisfied in either case. If $r_i$ and $D_i(\alpha_1, \alpha_2)$, $i=1,2$, denote the values of $r$ and $D(\alpha_1, \alpha_2)$ for the two families, an application of Theorem 4.1 gives immediately

Theorem 4.2. Let $\underline{\underline{T}}_1$ and $\underline{\underline{T}}_2$ be two families of tests of the form (4.1) which both satisfy Assumptions A, B, C, D. Then as $\delta \longrightarrow 0$,

$$\text{eff}(\underline{\underline{T}}_1/\underline{\underline{T}}_2) = \frac{N(\underline{\underline{T}}_2)}{N(\underline{\underline{T}}_1)} \sim \delta^{\frac{r_2 - r_1}{r_1 r_2}} \; \frac{D_2(\alpha_1, \alpha_2)^{1/r_2}}{D_1(\alpha_1, \alpha_2)^{1/r_1}} \quad . \tag{4.22}$$

Thus if $r_1 < r_2$, the efficiency of family $\underline{\underline{T}}_1$ relative to family $\underline{\underline{T}}_2$ tends to zero. If $r_1 = r_2 = r$,

$$\lim_{\delta \longrightarrow 0} \text{eff}(\underline{\underline{T}}_1/\underline{\underline{T}}_2) = \left(\frac{D_2(\alpha_1, \alpha_2)}{D_1(\alpha_1, \alpha_2)}\right)^{1/r} \quad . \tag{4.23}$$

In many applications the statistic $t_n$ or a function of $t_n$ will be asymptotically normally distributed. Let

$$\Phi(x) = (2\pi)^{-\frac{1}{2}} \int_{-\infty}^{x} e^{-y^2/2}\, dy$$

Assumption $C_1$. There exist two functions $\mu(\theta)$ and $\sigma(\theta) > 0$, defined for $\theta \in \omega$, and a sequence of increasing functions $g_n(t_n)$ of $t_n$, $n=1,2,\ldots,$ such that for every real $x$ and every $d \geq 0$

$$(4.24) \qquad \lim_{n \to \infty} P\left\{ n^{1/2} \frac{g_n(t_n)-\mu(\theta_1+ dn^{-1/2})}{\sigma(\theta_1+ dn^{-1/2})} \leq x \;\middle|\; \theta_1+ dn^{-1/2}\right\} = \Phi(x).$$

Assumption $D_1$. The function $\mu(\theta)$ has a continuous and positive derivative at $\theta = \theta_1$. The function $\sigma(\theta)$ is continuous at $\theta = \theta_1$.

By Assumption $D_1$ we can write

$$\mu(\theta_1+ dn^{-1/2}) = \mu(\theta_1) + dn^{-1/2}\, \mu'(\theta_1)(1+\epsilon_n),$$

$$\sigma(\theta_1+ dn^{-1/2}) = \sigma(\theta_1)(1+\epsilon_n'),$$

where

$$\varepsilon_n \longrightarrow 0 \text{ and } \varepsilon_n' \longrightarrow 0 \qquad \text{as } n \longrightarrow \infty.$$

Hence

$$n^{1/2} \frac{g_n(t_n)-\mu(\theta_1+ dn^{-1/2})}{\sigma(\theta_1+ dn^{-1/2})} = n^{1/2} \frac{g_n(t_n) - \mu(\theta_1)}{\sigma(\theta_1)(1+\varepsilon_n')} - d \frac{\mu'(\theta_1)(1+\varepsilon_n)}{\sigma(\theta_1)(1+\varepsilon_n')}$$

has the same limiting distribution as

$$n^{1/2} \frac{g_n(t_n) - \mu(\theta_1)}{\sigma(\theta_1)} - d \frac{\mu'(\theta_1)}{\sigma(\theta_1)} .$$

Thus if we let

$$f_n(t_n) = n^{1/2} \frac{g_n(t_n) - \mu(\theta_1)}{\sigma(\theta_1)}$$

and replace in (4.24) $x$ by $x-d\,\mu'(\theta_1)/\sigma(\theta_1)$, we obtain

$$\lim_{n \longrightarrow \infty} P\left\{ f_n(t_n) \leq x \mid \theta_1+ dn^{-1/2} \right\} = \Phi\left(x-d\frac{\mu'(\theta_1)}{\sigma(\theta_1)}\right) .$$

Assumption C is now satisfied with $r = \frac{1}{2}$ and $H_d(x)= \Phi(x-d\mu'(\theta_1)/\sigma(\theta_1))$.

Assumption D is also satisfied.

Let $\lambda(u)$ be defined by

$$u = 1 - \Phi(\lambda(u)) = \Phi(-\lambda(u)). \tag{4.25}$$

Then $\lambda(1-u) = -\lambda(u)$,

$$H_d^{-1}(u) = -\lambda(u) + d\mu'(\theta_1)/\sigma(\theta_1),$$

and from (4.6) we obtain

$$D(\alpha_1,\alpha_2) = \frac{\sigma(\theta_1)}{\mu'(\theta_1)} \left\{ \lambda(\alpha_1) + \lambda(\alpha_2) \right\} .$$

Hence we can state

Theorem 4.3. Let $\underline{T}$ be the family of tests (4.1). If Assumptions A, B, $C_1$, $D_1$ are satisfied and $\theta_1$ is fixed, we have asymptotically as $\delta \longrightarrow 0$,

$$N(\underline{T}) \sim \left(\frac{\sigma(\theta_1)}{\mu'(\theta_1)}\right)^2 \left(\frac{\lambda(\alpha_1) + \lambda(\alpha_2)}{\delta}\right)^2 . \tag{4.26}$$

The functions $\mu(\theta)$ and $\sigma(\theta)$ depend on the family $\underline{T}$. Let

$$J(\underline{\underline{T}}) = \left(\frac{\mu'(\theta_1)}{\sigma(\theta_1)}\right)^2 .$$

From Theorem 4.3 we obtain

<u>Theorem 4.4. Let $\underline{\underline{T}}_1$ and $\underline{\underline{T}}_2$ be two families of tests of the form</u> (4.1) <u>which both satisfy assumptions</u> A, B, $C_1$, $D_1$. <u>Then</u>

(4.27) $$\lim_{\delta \to 0} \text{eff}(\underline{\underline{T}}_1/\underline{\underline{T}}_2) = \frac{J(\underline{\underline{T}}_1)}{J(\underline{\underline{T}}_2)} .$$

Thus in this case the asymptotic relative efficiency is independent of $\alpha_1$ and $\alpha_2$.

Theorem 4.4 is essentially due to Pitman (see Noether [1], p. 241) who obtained an analogous result under somewhat different assumptions. Pitman's result has been extended by Noether [5].

So far we have considered tests with $w_{1n}$ of the particular form $t_n \leq c$. Since, by Assumptions C and D, the probability of $t_n = c$ tends to zero as n tends to infinity, Theorems 4.1 to 4.4 apply as well to families of tests with $w_{1n}$ of the form $t_n < c$. They also apply to families of tests which prescribe to take decision $d_1$ or $d_2$ according as $t_n < c$ or $t_n > c$, and a random decision when $t_n = c$.

One is frequently interested in the performance of a sequence of tests of the form here considered, with the constant $c = c_n'$ prescribed for every n. From the proof of Theorem 4.1, in particular from inequality (4.18), we obtain the conditions which the sequence $\{c_n'\}$ has to satisfy in order that the theorems be applicable to sequences of tests of this type. We can summarize these results in the following theorem.

Theorem 4.5. Theorems 4.1 and 4.3 remain true if $\underline{T}$ denotes one of the following families.

(a) A family of tests $\{\phi_{n,c}\}$, $-\infty < c < \infty$, $n=1,2,\ldots$, where $\phi_{n,c}$ prescribes to take decision $d_1$ if $t_n(\underline{x}_n) < c$ and decision $d_2$ if $t_n(\underline{x}_n) > c$.

(b) A sequence of tests $\{\phi_{nc'_n}\}$, $n=1,2,\ldots$, defined as in (a), where

$$\lim_n f_n(c'_n) = H_0^{-1}(1-\alpha_1) \qquad \text{(for Theorem 4.1)}$$

or

$$\lim n^{1/2} \frac{g_n(c'_n) - \mu(\theta_1)}{\sigma(\theta_1)} = \lambda(\alpha_1) \text{(for Theorem 4.3).}$$

Theorems 4.2 and 4.4 remain true for any two families of tests of types (a) or (b).

REFERENCES

[1] G. E. Noether, Asymptotic properties of the Wald-Wolfowitz test of randomness. Annals of Math. Stat., Vol. 21 (1950), pp. 231-246.

[2] E. L. Lehmann, Some principles of the theory of testing hypotheses. Annals of Math. Stat., Vol. 21 (1950), pp. 1-26.

[3] E. L. Lehmann and C. Stein, Most powerful tests of composite hypotheses. I. Normal distributions. Annals of Math. Stat., Vol. 19 (1948), pp. 495-516.

[4] E. L. Lehmann and C. Stein, On the theory of some non-parametric hypotheses. Annals of Math. Stat., Vol. 20 (1949), pp. 28-45.

[5] G. E. Noether, Asymptotic relative efficiency of tests of hypotheses. Unpublished.