A LOWER BOUND FOR THE AVERAGE SAMPLE NUMBER OF A SEQUENTIAL TEST[1]

By Wassily Hoeffding

University of North Carolina

Summary. A lower bound is derived for the expected number of observations required by an arbitrary sequential test which satisfies conventional conditions regarding the probabilities of erroneous decisions.

1. Statement of results. Let $X_1, X_2, \cdots$ be a sequence of independent random variables with a common frequency function $f(x, \theta)$ (either a probability density or the elementary probability law of a discrete distribution), where the parameter θ is confined to a set Ω. Let S be an arbitrary (possibly randomized) sequential test for deciding between two alternatives H_0 and H_1 which fulfills the following requirement. Given two disjoint subsets ω_0 and ω_1 of Ω and two numbers α, β between 0 and 1, S satisfies the inequalities

$$(1) \qquad \begin{aligned} &P_\theta(S \text{ accepts } H_1) \leqq \alpha && \text{if } \theta \in \omega_0, \\ &P_\theta(S \text{ accepts } H_0) \leqq \beta && \text{if } \theta \in \omega_1, \end{aligned}$$

where $P_\theta(E)$ denotes the probability of the event E when the common frequency function of the X_i is $f(x, \theta)$.

It will also be assumed that

$$(2) \qquad P_\theta(S \text{ accepts } H_0) + P_\theta(S \text{ accepts } H_1) = 1 \qquad \text{for all } \theta \in \Omega.$$

Let n be the number of observations required to terminate the test S (by accepting H_0 or H_1). It will be shown that if conditions (1) and (2) are satisfied and

$$(3) \qquad \alpha + \beta \leqq 1,$$

we have

$$(4) \qquad E_\theta(n) \geqq \frac{-\log [\alpha^c(1-\beta)^{1-c} + (1-\alpha)^c\beta^{1-c}]}{cE_\theta\left(\log \dfrac{f(X, \theta)}{f(X, \theta_0)}\right) + (1-c)E_\theta\left(\log \dfrac{f(X, \theta)}{f(X, \theta_1)}\right)}$$

for every c, θ_0 and θ_1 such that

$$(5) \qquad 0 < c < 1, \qquad \theta_0 \in \omega_0, \qquad \theta_1 \in \omega_1.$$

Here X denotes a random variable with the same distribution as the X_i, and $E_\theta(U)$ is the expected value of U (a function of $X, X_1, X_2, \cdots$) when the common frequency function is $f(x, \theta)$.

The expected values in the denominator in (4) always exist and are nonnegative, possibly $+\infty$ (as can be seen by applying the inequality $\log x \geqq 1 - x^{-1}$

Received 7/14/52.

[1] Work done with the support of the Office of Naval Research.

if $x > 0$). The numerator in (4) is nonnegative and vanishes if and only if $\alpha + \beta = 1$.

The best inequality that can be obtained from (4) is evidently

$$E_\theta(n) \geqq \sup_{0<c<1} \frac{-\log [\alpha^c(1-\beta)^{1-c} + (1-\alpha)^c\beta^{1-c}]}{ce_0(\theta) + (1-c)e_1(\theta)}, \tag{6}$$

where

$$e_i(\theta) = \inf_{\theta' \epsilon \omega_i} E_\theta\left(\log \frac{f(X,\theta)}{f(X,\theta')}\right), \qquad i = 0, 1. \tag{7}$$

If $\theta \epsilon \omega_1$, then $e_1(\theta) = 0$, and the ratio in (6) can be written as

$$-e_0(\theta)^{-1} \log \left[(1-\beta)\left(\frac{\alpha}{1-\beta}\right)^c + \beta\left(\frac{1-\alpha}{\beta}\right)^c \right]^{1/c}.$$

The expression following log is an increasing function of c. (For a proof see, e.g., Uspensky [1], particularly p. 267.) Letting $c \to 0$, we obtain

$$E_\theta(n) \geqq \frac{\beta \log \dfrac{\beta}{1-\alpha} + (1-\beta) \log \dfrac{1-\beta}{\alpha}}{e_0(\theta)} \qquad \text{if } \theta \epsilon \omega_1. \tag{8}$$

If $\theta \epsilon \omega_0$, inequality (6) reduces in a similar way to

$$E_\theta(n) \geqq \frac{\alpha \log \dfrac{\alpha}{1-\beta} + (1-\alpha) \log \dfrac{1-\alpha}{\beta}}{e_1(\theta)} \qquad \text{if } \theta \epsilon \omega_0. \tag{9}$$

Inequalities (8) and (9) were obtained by Wald ([2], Section 4.7) under the assumption that the sets ω_0 and ω_1 consist of one point each and the signs of equality hold in (1).

The sign of equality can hold in (6) only in certain special cases. If $\alpha + \beta = 1$, conditions (1) can be satisfied without taking any observations, and hence $E_\theta(n)$ can attain the lower bound 0. In (8) and (9) equality can be attained by a sequential probability ratio test for certain special distributions $f(x, \theta)$ and suitable values θ (cf. Wald [2]). In general the greatest lower bound for $E_\theta(n)$ is likely to be a complicated expression. The bounds derived here, although in general not the best ones, have the advantage of being simple.

The greatest lower bound of $E_\theta(n)$ can not exceed the least sample size $N = N(\alpha, \beta)$ of the best nonsequential test which satisfies (1). The following example may serve to compare the bound in (6) with $N(\alpha, \beta)$. Let

$$f(x, \theta) = \frac{1}{\sqrt{2\pi}} e^{-((x-\theta)^2)/2}, \qquad \omega_0 = \{\theta \leqq -\delta\}, \qquad \omega_1 = \{\theta \geqq \delta\}.$$

Then

$$e_0(\theta) = 0 \quad \text{if } \theta \leqq -\delta, \qquad e_0(\theta) = (\theta + \delta)^2/2 \quad \text{if } \theta > -\delta,$$

$$e_1(\theta) = (\theta - \delta)^2/2 \text{ if } \theta < \delta, \qquad e_1(\theta) = 0 \quad \text{if } \theta \geqq \delta.$$

Suppose that $\alpha = \beta$, and let $\theta = 0$. Then the supremum in (6) is attained for $c = \frac{1}{2}$, and we obtain $E_0(n) \geqq -\log[4\alpha(1-\alpha)]/\delta^2 = M$, say. (It can be shown that M is the maximum with respect to θ of the bound in (6).)

The best nonsequential test which satisfies (1) with $\alpha = \beta$ accepts H_0 or H_1 according as $\sum_1^N X_i$ is negative or positive, where

$$N = \frac{\lambda^2}{\delta^2}, \qquad \frac{1}{\sqrt{2\pi}} \int_\lambda^\infty e^{-(x^2)/2}\, dx = \alpha.$$

(We assume for simplicity that λ^2/δ^2 is an integer.) Hence

$$M/N = -\log[4\alpha(1-\alpha)]/\lambda^2,$$

a function of α only which varies between $\frac{1}{2}$ (for $\alpha \to 0$) and $2/\pi$ (for $\alpha \to \frac{1}{2}$).

2. Proof of inequality (4). The proof of (4) will be based on the inequality

$$E_\theta(n) E_\theta\left(\log \frac{f(X,\theta)}{f(X,\theta')}\right) \geqq L(\theta) \log \frac{L(\theta)}{L(\theta')} + [1 - L(\theta)] \log \frac{1 - L(\theta)}{1 - L(\theta')}, \qquad \theta, \theta' \in \Omega, \tag{10}$$

where

$$L(\theta) = P_\theta(S \text{ accepts } H_0). \tag{11}$$

Inequality (10) is due to Wald [2] and is true for every test S which satisfies (2). In Wald's proof the test S is assumed to be nonrandomized (in particular, a first observation is always taken, so that $n \geqq 1$), but it is easy to extend the proof to randomized tests.

To prove (4), put in (10) $\theta' = \theta_0$ and multiply both sides with c; then put $\theta' = \theta_1$ and multiply both sides with $1 - c$. Addition of the corresponding sides of the two resulting inequalities gives

$$\begin{aligned} &E_\theta(n)\left[cE_\theta\left(\log \frac{f(X,\theta)}{f(X,\theta_0)}\right) + (1-c)E_\theta\left(\log \frac{f(X,\theta)}{f(X,\theta_1)}\right)\right] \\ &\quad \geqq L(\theta)\log L(\theta) + [1 - L(\theta)]\log[1 - L(\theta)] - rL(\theta) - s[1 - L(\theta)] \\ &\quad = H(L(\theta)), \end{aligned} \tag{12}$$

say, where

$$r = c \log L(\theta_0) + (1-c)\log L(\theta_1),$$
$$s = c \log[1 - L(\theta_0)] + (1-c)\log[1 - L(\theta_1)].$$

The minimum of the function $H(u)$ is attained at $u = u_0$, where $u_0 = e^r/(e^r + e^s)$, and we find

$$H(u) \geqq H(u_0) = -\log K(1 - L(\theta_0), L(\theta_1)), \tag{13}$$

where

$$K(x,y) = x^c(1-y)^{1-c} + (1-x)^c y^{1-c}. \tag{14}$$

The function $K(x, y)$ is an increasing function of x and an increasing function of y, provided $x + y < 1$. Conditions (1) and (2) imply that $1 - L(\theta_0) \leqq \alpha$, $L(\theta_1) \leqq \beta$. Hence if $\alpha + \beta < 1$, we have

$$K(1 - L(\theta_0), L(\theta_1)) \leqq K(\alpha, \beta). \tag{15}$$

Inequality (4) now follows from relations (12) to (15).

Concerning the conditions for equality, it suffices to observe that in (10) the sign of equality holds if and only if there exist constants C_0 and C_1 such that

$$P_\theta\left\{\prod_{j=1}^{n} \frac{f(X_j, \theta)}{f(X_j, \theta')} = C_i \mid S \text{ accepts } H_i\right\} = 1, \qquad i = 0, 1,$$

where the usual notation for conditional probabilities is used. This can be verified from Wald's proof. The conditions for equality in (12), (13), (15) are obvious.

REFERENCES

[1] J. V. Uspensky, *Introduction to Mathematical Probability*, McGraw-Hill Book Co., New York and London, 1937.

[2] A. Wald, "Sequential tests of statistical hypotheses," *Ann. Math. Stat.*, Vol. 16 (1945), pp. 117–186.

SOME INEQUALITIES ON MILL'S RATIO AND RELATED FUNCTIONS

By M. R. Sampford

University of Oxford

1. Introduction. Mill's ratio is defined as

$$R_x = e^{\frac{1}{2}x^2} \int_x^\infty e^{-\frac{1}{2}u^2}\, du. \tag{1}$$

Gordon [1] and Birnbaum [2] have given, respectively, upper and lower limits for R_x as

$$\tfrac{1}{2}\{\sqrt{4 + x^2} - x\} < R_x < 1/x, \qquad x > 0. \tag{2}$$

Murty [3] has shown how limits to R_x of any required degree of accuracy can be derived for $x > 0$ by the use of successive convergents of Laplace's expression for the normal integral as a continued fraction. No limits have, as yet, been published for $x < 0$.

If the functions $\nu(x)$ and $\lambda(x)$ are defined by $\nu(x) = 1/R_x$, $\lambda(x) = \nu'(x) = \nu(\nu - x)$, the inequalities

$$0 < \lambda < 1, \tag{3}$$

$$\lambda' = \nu\{(\nu - x)(2\nu - x) - 1\} > 0 \tag{4}$$

Received 5/15/52, revised 9/16/52.