The Annals of Statistics
1976, Vol. 4, No. 6, 1240–1243

AN IMPROVED STATEMENT OF OPTIMALITY FOR SEQUENTIAL PROBABILITY RATIO TESTS

BY GORDON SIMONS[1]

University of North Carolina at Chapel Hill

An improved version of the optimality property of sequential probability ratio tests is described. Alternatives to error probabilities, as measurements of precision, are suggested.

Wald and Wolfowitz (1948) first proved that sequential probability ratio tests (SPRT's) possess a strong optimality property. Improvements on their statement of optimality have been given by Burkholder and Wijsman (1960 and 1963) and by J. K. Ghosh (1961). The intent of this paper is to discuss another improvement.

We shall assume the usual i.i.d. model (see Ferguson (1967), page 361) with respect to two probability measures P_0 and P_1. In particular, every observation has the density f_0 under P_0, or f_1 under P_1. Let T be an SPRT $S(A, B)$, $0 < A < 1 < B < \infty$, with error probabilities α_0 and α_1, and with expected sample sizes $E_0 N$ and $E_1 N$. Further, let T' be a competing test with error probabilities α_0' and α_1', and with expected sample sizes $E_0 N'$ and $E_1 N'$. The optimality property stated by Wald and Wolfowitz says that if

$$(1) \qquad \alpha_0' \leqq \alpha_0 \quad \text{and} \quad \alpha_1' \leqq \alpha_1 ,$$

and if

$$(2) \qquad E_0 N' < \infty \quad \text{and} \quad E_1 N' < \infty ,$$

then

$$(3) \qquad E_0 N \leqq E_0 N' \quad \text{and} \quad E_1 N \leqq E_1 N' .$$

We shall show that (1) can be replaced by the weaker assumption

$$(4) \qquad \beta_0' \leqq \beta_0 \quad \text{and} \quad \beta_1' \leqq \beta_1 ,$$

where $\beta_0 = \alpha_0/(1 - \alpha_1)$, $\beta_1 = \alpha_1/(1 - \alpha_0)$ and the quantities β_0' and β_1' are defined analogously. The conditions in (4) imply that $\alpha_0' + \alpha_1' \leqq \alpha_0 + \alpha_1$ but permit one of the inequalities in (1) to be violated.

As an aside, it appears to us that these ratios (the β's) may be more useful measures of error than are error probabilities. For instance, they appear in the fundamental results

$$(5) \qquad \beta_1 \leqq A \quad \text{and} \quad \beta_0 \leqq B^{-1} .$$

Received November 1975; revised April 1976.

[1] The work of the author was supported in part under Grant No. AFOSR-75-2796 and in part by Grant NSF MPS 75-07556.

AMS 1970 *subject classification.* Primary 62L10.

Key words and phrases. Optimality, sequential tests, sequential probability ratio tests, Kullback–Leibler information number, error probabilities.

(See Lehmann (1959), page 99.) Moreover, they appear in various places in the literature of sequential tests. We shall refer to them as *normalized* error probabilities. (Observe that $\beta_0 = P_0(\text{rejecting } H_0)/P_1(\text{rejecting } H_0)$. Thus, the error probability in the numerator is being normalized by a probability of nonerror in the denominator.) Besides the result we have mentioned, we shall discuss another well-known result which can be strengthened by replacing assumptions about error probabilities with assumptions about normalized error probabilities. We suspect that many other examples can be found.

PROOF THAT (4) AND (2) IMPLY (3). For the sake of brevity, we shall draw rather freely from the notation and results appearing on pages 361–368 of Ferguson (1967).

The conditions in (1) are used by Ferguson only to conclude that

$$(1-\pi)w_{01}\alpha_0' + \pi w_{10}\alpha_1' \leqq (1-\pi)w_{01}\alpha_0 + \pi w_{10}\alpha_1\,, \tag{6}$$

where π, $0 < \pi < 1$, is the prior probability that f_1 is the correct density, and where $w_{01} > 0$ and $w_{10} > 0$ are losses due to type I and type II errors, respectively. Thus, it suffices to show that (4) implies (6).

According to (7.65) of Ferguson,

$$\pi_L \leqq \frac{w_{01}}{w_{01} + w_{10}} \leqq \pi_U\,,$$

where π_L and π_U are values in the open interval (0, 1) which satisfy

$$A = \frac{1-\pi}{\pi}\cdot\frac{\pi_L}{1-\pi_L} \qquad \text{and} \qquad B = \frac{1-\pi}{\pi}\cdot\frac{\pi_U}{1-\pi_U}\,.$$

Thus,

$$A \leqq \frac{(1-\pi)w_{01}}{\pi w_{10}} \leqq B\,,$$

which, in view of (5), implies

$$\beta_1 \leqq \frac{(1-\pi)w_{01}}{\pi w_{10}} \leqq \beta_0^{-1}. \tag{7}$$

Finally, (6) follows from (4) and (7). This is most easily seen by first showing that the conditions in (4) are algebraically equivalent to

$$\beta_0(\alpha_1' - \alpha_1) \leqq \alpha_0 - \alpha_0' \qquad \text{and} \qquad \alpha_1' - \alpha_1 \leqq \beta_1(\alpha_0 - \alpha_0')\,;$$

and (6) is algebraically equivalent to

$$(\alpha_1' - \alpha_1) \leqq \frac{(1-\pi)w_{01}}{\pi w_{10}}(\alpha_0 - \alpha_0')\,. \qquad \square$$

REMARKS.

1. Ferguson's statement of the optimality property (Theorem 2, page 365) legitimately ignores the conditions in (2), but his proof implicitly uses them; Burkholder and Wijsman (1963) have shown that the conditions are unnecessary.

Since the proof of our improved statement of optimality, given above, draws heavily on Ferguson's argument, we have included the conditions of (2) in our statement as a matter of convenience. They can be deleted.

2. Ferguson permits A and B to equal 1 by allowing his sequential probability ratio test to stop with no observations. In such a case $E_0 N$ and $E_1 N$ are both zero and, of course, the optimality property (i.e., (3)) holds whether one assumes (4) or not.

We shall now discuss another result which can be strengthened by the introduction of normalized error probabilities. Let $I(f_0, f_1) = E_0 \log (f_0(X)/f_1(X))$ (a Kullback–Leibler information number), where X represents an observation appearing in the random sample. Suppose α_0^* and α_1^* are fixed positive numbers whose sum $\alpha_0^* + \alpha_1^* < 1$. Then, for any test whose stopping variable is N and whose error probabilities are $\alpha_0 \leqq \alpha_0^*$ and $\alpha_1 \leqq \alpha_1^*$,

$$E_0 N \geqq \frac{\alpha_0^* \log \dfrac{\alpha_0^*}{1-\alpha_1^*} + (1-\alpha_0^*) \log \dfrac{1-\alpha_0^*}{\alpha_1^*}}{I(f_0, f_1)} . \tag{8}$$

This follows from the well-known result that

$$E_0 N \geqq \frac{\alpha_0 \log \dfrac{\alpha_0}{1-\alpha_1} + (1-\alpha_0) \log \dfrac{1-\alpha_0}{\alpha_1}}{I(f_0, f_1)} ,$$

and from the fact that the function

$$g(x, y) = x \log \frac{x}{1-y} + (1-x) \log \frac{1-x}{y} , \qquad 0 < x, y < 1 ,$$

is convex in each of its arguments. It is not difficult to show (8) holds when the conditions $\alpha_0 \leqq \alpha_0^*$ and $\alpha_1 \leqq \alpha_1^*$ are replaced by the weaker conditions

$$\beta_0 \leqq \beta_0^* \qquad \text{and} \qquad \beta_1 \leqq \beta_1^* , \tag{9}$$

where $\beta_0^* = \alpha_0^*/(1-\alpha_1^*)$ and $\beta_1^* = \alpha_1^*/(1-\alpha_0^*)$. One must show that (9) implies $g(\alpha_0, \alpha_1) \geqq g(\alpha_0^*, \alpha_1^*)$.

REMARK. After submitting this paper to the *Annals of Statistics*, the author discovered that an improved version of the Neyman–Pearson lemma can be obtained as well in terms of normalized error probabilities. This will be reported on elsewhere.

REFERENCES

[1] BURKHOLDER, D. L. and WIJSMAN, R. A. (1960). Optimum properties and admissibility of sequential tests. *Ann. Math. Statist.* **31** 232.

[2] BURKHOLDER, D. L. and WIJSMAN, R. A. (1963). Optimum properties and admissibility of sequential tests. *Ann. Math. Statist.* **34** 1–17.

[3] FERGUSON, T. S. (1967). *Mathematical Statistics, A Decision Theoretic Approach.* Academic Press, New York.

[4] GHOSH, J. K. (1961). On the optimality of probability ratio tests in sequential and multiple sampling. *Calcutta Statist. Assoc. Bull.* **10** 73–92.

[5] LEHMANN, E. L. (1959). *Testing Statistical Hypotheses.* Wiley, New York.

[6] WALD, A. and WOLFOWITZ, J. (1948). Optimum character of the sequential probability ratio test. *Ann. Math. Statist.* **19** 326–339.

DEPARTMENT OF STATISTICS
UNIVERSITY OF NORTH CAROLINA
CHAPEL HILL, NORTH CAROLINA 27514