REMARK ON THE OPTIMUM CHARACTER OF THE SEQUENTIAL PROBABILITY RATIO TEST

BY J. WOLFOWITZ

Cornell University

There is a small lacuna in the proof ([1]) of the property stated in the title of this note. In some recent papers many pages are devoted to correcting it. Since all “other” proofs of the optimum character of the sequential probability ratio test follow all the principal ideas of [1] and differ from the latter only in very minor details, it seems appropriate to show, as we will, that the lacuna can be filled in a very simple and obvious way. The present note assumes familiarity only with Lemmas 1, 2, and 3 of [1]. The gap in [1] is in Lemma 1, where it is claimed that the test \( S^* \) there constructed minimizes the average risk.

We shall replace Lemmas 1, 2, and 3 of [1] by Lemma A whose statement is that of Lemma 1 plus that of Lemma 2 plus that of Lemma 3. The proof of Lemma A will be that given for Lemma 1, followed by that given for Lemma 2, followed by that given for Lemma 3, followed by the remarks which we now make.

At the end of the proof of Lemma 2 we already have that \( S^* \) is the sequential probability ratio test.

We now prove that \( S^* \) minimizes the average risk. We adopt the notation and terminology of [1]. Suppose there were a test \( S \) such that

\[
R(S) = R(S^*) - \delta, \quad \delta > 0.
\]

We shall construct a sequence of tests \( S_0 (=S), S_1, S_2, \ldots \), such that, for \( i = 0, 1, 2, \ldots \),

(2) \[ R(S_{i+1}) \leq R(S_i) + \delta(2^{i-2}) \]

and

(3) \[ \lim_{i \to \infty} R(S_i) = R(S^*). \]

From this it follows that

(4) \[ R(S^*) \leq R(S) + \delta/2. \]

The contradiction between (1) and (4) proves the desired result.

If \( t \) is any sequential test let \( n(t) \) be its associated stopping variable; the value of \( n(t) \) at the point \( \omega = x_1, x_2, \ldots \) will be denoted by \( n(\omega, t) \). Let \( r_j(\omega) = r_j(x_1, \cdots, x_j) = p_{ij}/p_{0j} \) as in [1]. Let \( T \) be the totality of all sequential tests \( t \) such that \( E[n(t)] < \infty, i = 0, 1, 2 \).

Define

\[
T_0 = \{ t \in T \mid r_j(\omega) \geq A \quad \text{or} \quad \leq B \Rightarrow n(\omega, t) \leq j, j \geq 1 \}
\]

Received 4 November 1965.

\(^1\) Research under contract with the Office of Naval Research.
and

\[ T_j = \{ t \in T_0 \mid P_i(\omega \mid n(\omega, t) = j, B < r_j(\omega) < A) = 0, \ i = 0, 1 \}, \]

where \( P_i \) indicates probability under the hypothesis \( H_i \). We note that, if \( S \) is not in \( T_0 \), we may, by Lemma 3, (c), replace it by an obvious modification \( S' \in T_0 \) such that \( R(S') \leq R(S) \). We therefore assume that \( S \) is in \( T_0 \). Let \( \alpha \) be the integer such that \( S \in \mathcal{T}_\alpha \cap \bigcap_{i=1}^{\alpha-1} T_i \).

Let \( \epsilon > 0 \) be arbitrary. It follows from Lemma 3 and the argument of Lemma 2 that there exists a sequential test \( Z(\epsilon) \) (resp. \( Z'(\epsilon) \)) which requires at least one additional observation, and is such that

\[ l(x_1, \ldots, x_\alpha) - E(L_n \mid (x_1, \ldots, x_\alpha), Z(\epsilon)) > -\epsilon \]

for any \( (x_1, \ldots, x_\alpha) \) such that \( B \leq r_\alpha(x_1, \ldots, x_\alpha) \leq W_{\theta_0}/W_{\theta_1} \) (resp., is such that

\[ l(x_1, \ldots, x_\alpha) - E(L_n \mid (x_1, \ldots, x_\alpha), Z'(\epsilon)) > -\epsilon \]

for any \( (x_1, \ldots, x_\alpha) \) such that \( W_{\theta_0}/W_{\theta_1} < r_\alpha(x_1, \ldots, x_\alpha) \leq A \). Now modify \( S \) as follows: when \( B < r_\alpha(x_1, \ldots, x_\alpha) \leq W_{\theta_0}/W_{\theta_1} \) (resp., when \( W_{\theta_0}/W_{\theta_1} < r_\alpha(x_1, \ldots, x_\alpha) < A \)) and \( S \) calls for stopping taking observations, replace \( S \) by the test \( Z(\delta/4) \) (resp., \( Z'(\delta/4) \)). Call the resulting test \( S_1 \). Since \( S, Z, \) and \( Z' \) fulfill the required measurability and integrability conditions, so does \( S_1 \), which also satisfies (2) for \( i = 0 \).

We repeat the above procedure in an obvious way on \( S_1 \), replacing \( \delta/4 \) by \( \delta/8 \), and obtaining \( S_2 \). We repeat the same procedure on \( S_2, S_3, \ldots \), using \( \delta/16, \delta/32, \ldots \). Since the index \( \alpha \) increases by at least one each time, (3) follows. This completes the proof.

REFERENCE