THEMIS SIGNAL ANALYSIS STATISTICS RESEARCH PROGRAM

OPTIMUM PROPERTY OF THE PARTIAL SEQUENTIAL PROBABILITY RATIO TEST

by

Campbell B. Read

Technical Report No. 53
Department of Statistics THEMIS Contract

January 15, 1970

Research sponsored by the Office of Naval Research
Contract N00014-68-A-0515
Project NR 042-260

DEPARTMENT OF STATISTICS
Southern Methodist University

Optimum Property of the Partial Sequential Probability Ratio Test

by

Campbell B. Read

Southern Methodist University and Southwestern Medical School of the University of Texas.

---

1. Introduction. Let θ be the parameter of interest in the distribution of an i.i.d. sequence $\{X_n\}$ of r.v.'s having common p.d.f. $p_\theta(\cdot)$. The Partial Sequential Probability Ratio Test (PSPRT) procedure is defined on the class $\mathcal{C}_n$ of test procedures requiring at least n observations to be made. Beyond stage n, observations are continued one at a time according as

$$B < p_{1n'}/p_{on'} < A \qquad (n' \geq n)$$

or not, where $0 < B < 1 < A$ is assumed, and we are testing the simple hypothesis H_o: $\theta = \theta_o$ against the simple alternative H_1: $\theta = \theta_1$. $p_{in'}$ is the joint p.d.f. of the first n' observations $\underline{x}_{n'} = (x_1, \ldots, x_{n'})$ when $\theta = \theta_i$ $(i = 0, 1)$. When we stop observing, H_o is accepted if $p_{1n'}/p_{on'} \leq B$, and H_1 is accepted if $p_{1n'}/p_{on'} \geq A$.

Essentially the PSPRT is a two-stage procedure, in which the second stage operates exactly like a Wald SPRT [4]. The latter has an optimum property, established by Wald and Wolfowitz ([5], [6]), namely, that among all tests for which

$$P(\text{reject } H_o|\theta_o) \leq \alpha, \; P(\text{reject } H_1|\theta_1) \leq \beta,$$

and for which $E(N|\theta_i) < \infty$ $(i = 0,1)$, the SPRT with error probabilities α and β minimizes $E(N|\theta_o)$ and $E(N|\theta_1)$.

THEMIS SIGNAL ANALYSIS STATISTICS RESEARCH PROGRAM

OPTIMUM PROPERTY OF THE PARTIAL SEQUENTIAL PROBABILITY RATIO TEST

by

Campbell B. Read

Technical Report No. 53
Department of Statistics THEMIS Contract

January 15, 1970

Research sponsored by the Office of Naval Research
Contract N00014-68-A-0515
Project NR 042-260

DEPARTMENT OF STATISTICS
Southern Methodist University

Optimum Property of the Partial Sequential Probability Ratio Test

by

Campbell B. Read

Southern Methodist University and Southwestern Medical School of the University of Texas.

---

1. Introduction. Let θ be the parameter of interest in the distribution of an i.i.d. sequence $\{X_n\}$ of r.v.'s having common p.d.f. $p_\theta(\cdot)$. The Partial Sequential Probability Ratio Test (PSPRT) procedure is defined on the class $\mathcal{C}_n$ of test procedures requiring at least n observations to be made. Beyond stage n, observations are continued one at a time according as

$$B < p_{1n'}/p_{on'} < A \qquad (n' \geq n)$$

or not, where $0 < B < 1 < A$ is assumed, and we are testing the simple hypothesis H_o: $\theta = \theta_o$ against the simple alternative H_1: $\theta = \theta_1$. $p_{in'}$ is the joint p.d.f. of the first n' observations $\underline{x}_{n'} = (x_1, \ldots, x_{n'})$ when $\theta = \theta_i$ $(i = 0, 1)$. When we stop observing, H_o is accepted if $p_{1n'}/p_{on'} \leq B$, and H_1 is accepted if $p_{1n'}/p_{on'} \geq A$.

Essentially the PSPRT is a two-stage procedure, in which the second stage operates exactly like a Wald SPRT [4]. The latter has an optimum property, established by Wald and Wolfowitz ([5], [6]), namely, that among all tests for which

$$P(\text{reject } H_o|\theta_o) \leq \alpha, \; P(\text{reject } H_1|\theta_1) \leq \beta,$$

and for which $E(N|\theta_i) < \infty$ $(i = 0,1)$, the SPRT with error probabilities α and β minimizes $E(N|\theta_o)$ and $E(N|\theta_1)$.

If θ has a prior distribution $\lambda = (w, 1-w)$ on $\Omega = (\theta_o, \theta_1)$, and if the losses for wrong rejection of H_i are L_i $(i = 0,1)$, with cost per observation c, then the Bayes risk of any test procedure δ is

$$r(w,\delta) = w\left[\alpha L_o + cE(N|\theta_o)\right] + (1-w)\left[\beta L_1 + cE(N|\theta_1)\right] \tag{1}$$

where α and β depend on δ. Wald and Wolfowitz proved the auxiliary result that the Bayes risk (1) is minimized by a SPRT.

In this paper we show that the same results hold for a PSPRT procedure if attention is restricted to tests in the class $\mathcal{C}_n$. The proofs directly extend those of Lehmann and LeCam [1].

In a separate paper [2], Wald's approximations to the OC function and ASN of a SPRT are extended to a PSPRT, and several numerical results are given for the problem of testing a normal mean with known variance. There is a clear indication that the PSPRT may result in a lower ASN than that resulting from a SPRT with the same error probabilities, for some values of θ satisfying $\theta_o < \theta < \theta_1$. This last feature may make the PSPRT an attractive procedure to follow, if the experimenter wishes to allow for the possibility that θ may take some value other than θ_o or θ_1.

2. Main results.

<u>Theorem 1</u>. There exist numbers w' and w", determined by L_o, L_1, and c, but independent of n and w such that the Bayes risk (1) in the class of procedures $\mathcal{C}_n$ terminating with probability one, is minimized by a PSPRT with boundaries

$$B = \frac{w}{1-w} \cdot \frac{1-w''}{w''}, \quad A = \frac{w}{1-w} \cdot \frac{1-w'}{w'} \quad .$$

<u>Corollary 1</u>. If A and B are the boundaries of the SPRT which minimizes

the Bayes risk (1) in the class $\mathcal{C}_o$ of decision procedures, then they are also the boundaries of the PSPRT which minimizes (1) in the class $\mathcal{C}_n$, for any given n.

Theorem 2. Given any two numbers w_o' and w_o'' such that $0 < w_o' < w_o'' < 1$, there exist numbers ℓ $(0 < \ell < 1)$ and $c > 0$ such that the Bayes risk (1) in the class $\mathcal{C}_n$ with $L_o = 1 - \ell$, $L_1 = \ell$ and a prior probability w for θ_o satisfying $w_o' < w < w_o''$ is minimized by a PSPRT with boundaries

$$B = \frac{w}{1-w} \cdot \frac{1-w_o''}{w_o''} \quad , \qquad A = \frac{w}{1-w} \cdot \frac{1-w_o'}{w_o'} \quad .$$

Theorem 3. Among all procedures in the class $\mathcal{C}_n$ for which

$P(\text{reject } H_o|\theta_o) \leq \alpha$,

$P(\text{reject } H_1|\theta_1) \leq \beta$,

and for which $E(N|\theta_o)$ and $E(N|\theta_1)$ are finite, the PSPRT with error probabilities α and β minimizes both $E(N|\theta_o)$ and $E(N|\theta_1)$.

The proofs which follow in Section 3 are closely related to Lehmann's Lemmas 5 and 6, and his Theorem 8 [1], and it is of interest first to examine an intuitive argument. If the PSPRT is denoted by T_n, with boundaries A and B, and the conditional procedure, given $\underline{x}_n$ has been observed at stage n, by $T(\underline{x}_n)$, then $T(\underline{x}_n)$ is a SPRT (possibly with no further observations) having the continuation region

$$B_n' = B \cdot \frac{p_{on}}{p_{1n}} < \frac{p_1(x_{n+1},\ldots, x_n')}{p_o(x_{n+1},\ldots, x_n')} < A \cdot \frac{p_{on}}{p_{1n}} = A_n'$$

If $\alpha(\underline{x}_n)$, $\beta(\underline{x}_n)$ and $E(N|\underline{x}_n,\theta_i)$ are the error probabilities of this conditional procedure, and

$W_n = wp_{on}/(wp_{on}+(1-w)p_{1n})$

is the posterior probability of θ_o at stage n, then it follows from Lehmann's Lemma 5 that $T(\underline{x}_n)$ minimizes the conditional (or posterior) Bayes risk

$$r(w_n, \delta(\underline{x}_n)) = w_n\left[\alpha(\underline{x}_n)L_o + cE(N|\underline{x}_n, \theta_o)\right] + (1-w_n)\left[\beta(\underline{x}_n)L_1 + cE(N|\underline{x}_n, \theta_1)\right]$$

if $B_n' = \frac{w_n}{1-w_n} \cdot \frac{1-w''}{w''}$, $A_n' = \frac{w_n}{1-w_n} \cdot \frac{1-w'}{w'}$ for certain numbers $0 < w' < w'' < 1$.

But it is easy to show that the continuation region

$$\frac{w_n}{1-w_n} \cdot \frac{1-w''}{w''} < \frac{p_1(x_{n+1},\ldots,x_{n'})}{p_o(x_{n+1},\ldots,x_{n'})} < \frac{w_n}{1-w_n} \cdot \frac{1-w'}{w'}$$

is equivalent to

$$B = \frac{w}{1-w} \cdot \frac{1-w''}{w''} < \frac{p_{1n'}}{p_{on'}} < \frac{w}{1-w} \cdot \frac{1-w'}{w'} = A \quad , n' \geq n. \tag{2}$$

Hence, whatever $\underline{x}_n$ may be observed, the SPRT $T(\underline{x}_n)$ which minimizes the conditional Bayes risk, is derived from the same PSPRT T_n. Hence, intuitively, it would appear that T_n minimizes the (unconditional) Bayes risk (1). While the above equivalence relations demonstrate the likelihood principle, the uniqueness of T_n arises also from the fact that w' and w" are independent of w_n, and hence independent of $\underline{x}_n$. Finally, Lehmann's proof shows that the SPRT minimizes the conditional Bayes risk at every stage of the experiment. In the class $\mathcal{C}_n$, then, it seems natural that the PSPRT should do the same, from stage n onwards.

3. Proofs of theorems. In what follows, some familiarity with Lehmann's treatment ([1], 3.10 and 3.12) is assumed.

Proof of Theorem 1. Suppose that n observations have been taken, with outcome $\underline{x}_n$. We first wish to decide whether or not to stop. The loss nc already incurred does not affect the problem, since no further action can retrieve it. An unlimited sequence $X_{n+1}, \ldots$ is available.

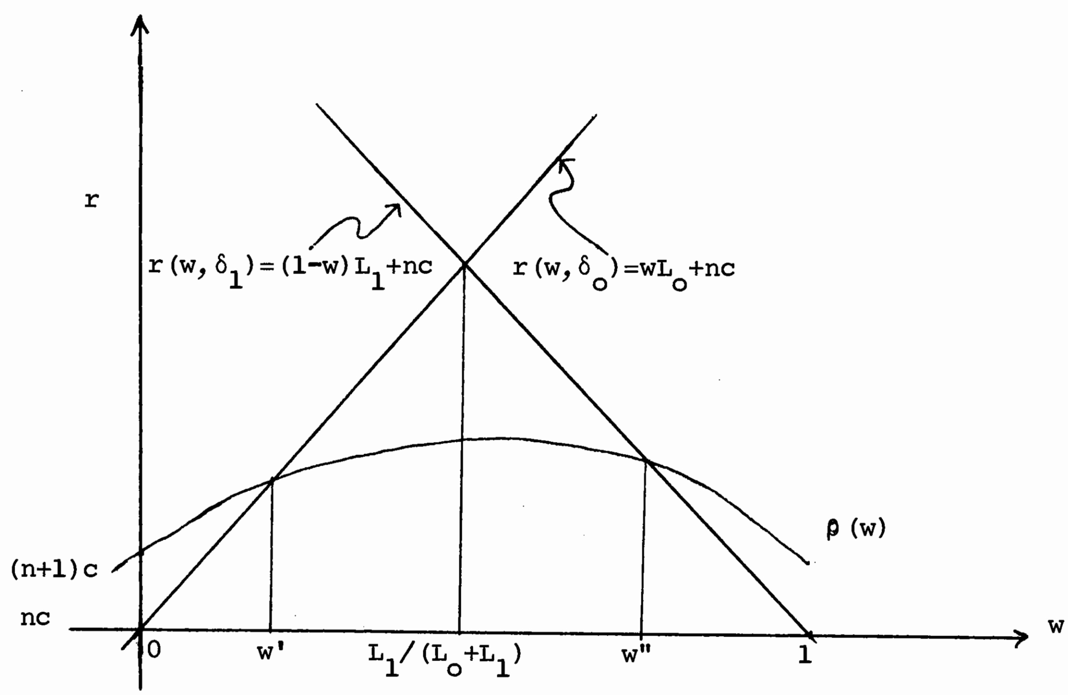
Let δ_o be the procedure which rejects H_o without taking any further observations beyond stage n, and δ_1 the procedure that similarly rejects H_1.

Thus $r(w,\delta_o) = wL_o + cn$

and $r(w,\delta_1) = (1-w)L_1 + cn$.

Let $\rho(w) = \inf_{\delta \in \mathcal{C}_{n+1}} r(w,\delta)$.

Then ρ is concave and continuous in (0,1) (see [1], Lemma 5).



If $\rho(L_1/(L_o+L_1)) < L_oL_1/(L_o+L_1) + nc$, define w' and w" by $r(w', \delta_o) = \rho(w')$, and $r(w", \delta_1) = \rho(w")$.

Otherwise let

$$w' = w" = L_1/(L_o + L_1) .$$

Figure 1 shows that we restrict attention to cases in which $0 < w' < w" < 1$. δ_o minimizes (1) if and only if $w \le w'$, and δ_1 minimizes (1) if and only if $w \ge w"$. So the optimum step at stage n is uniquely established as: if $Pr(H_o$ true) $< w'$ or $> w"$, take no further observations and reject or accept H_o respectively. If $w' < Pr(H_o$ true) $< w"$, observe X_{n+1}.

The proof is completed by induction as in Lehmann's Lemma 5. A similar situation arises at stage n' if $\underline{x}_{n'}$ has been observed $(n' \ge n)$, and the optimum rule is the same, except that

$$Pr(H_o \text{ true}) = w_{n'} = wp_{on'}/\left[wp_{on'}+(1-w)p_{1n'}\right].$$

The rule for continuing to observe, i.e. $w' < Pr(H_o$ true) $< w"$, is equivalent to (2), which defines the required PSPRT.

This establishes Theorem 1.

Corollary 1 follows from the analysis in the preceding proof. w' and w" are uniquely determined, and the argument for deriving them is not affected by n. Lehmann's Lemma 5 leads to the same continuation region (2) in the class C_o (where for the sake of completeness, we may take $p_{1n'}/p_{on'} = 1$ when $n' = 0$).

Proof of Theorem 2. Consider first the class C_o of decision procedures. Then Theorem 2 is Lehmann's Lemma 6.

But the corollary to Theorem 1 implies that a PSPRT with the same boundaries A and B minimizes the Bayes risk (1) in the class C_n, for any given finite positive integer n, and for losses $L_o = \ell$, $L_1 = 1-\ell$, and cost c.

Theorem 2 follows.

Proof of Theorem 3. Lehmann's proof applies with very little change.

For the sake of clarity, it is repeated here.

Consider any PSPRT δ with $0 < B < 1 < A$, and any constant w such that $0 < w < 1$. If

$$w' = w/\left[A(1-w)+w\right], \qquad w'' = w/\left[B(1-w)+w\right],$$

then $0 < w' < w < w'' < 1$ and (2) is satisfied. By Theorem 2, there exist ℓ and c such that $0 < \ell < 1$ and $c > 0$; and δ is Bayes in the class C_n for the situation in which w is the prior probability of H_o, $L_o = 1-\ell$, $L_1=\ell$, and cost c. Let this PSPRT have error probabilities α and β, and ASN $E(N|\theta)$; and consider any other procedure δ^* with error probabilities $\alpha^* \le \alpha$, $\beta^* \le \beta$, and ASN $E^*(N|\theta_i) < \infty$, $i = 0,1$. Since δ minimizes the Bayes risk,

$$w[(1-\ell)\alpha+cE(N|\theta_o)] + (1-w)[\ell\beta + cE(N|\theta_1)]$$
$$\le w[(1-\ell)\alpha^*+cE^*(N|\theta_o)] + (1-w)[\ell\beta^*+cE^*(N|\theta_1)]$$

and so

$$wE(N|\theta_o) + (1-w)E(N|\theta_1) \le wE^*(N|\theta_o) + (1-w)E^*(N|\theta_1) .$$

Since this holds uniformly for all w in $(0,1)$,

$$E(N|\theta_o) \le E^*(N|\theta_o)$$

and $E(N|\theta_1) \le E^*(N|\theta_1)$.

UNCLASSIFIED
Security Classification

DOCUMENT CONTROL DATA - R & D

Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY *(Corporate author)* SOUTHERN METHODIST UNIVERSITY	2a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED 2b. GROUP UNCLASSIFIED

3. REPORT TITLE

Optimum Property of the Partial Sequential Probability Ratio Test

4. DESCRIPTIVE NOTES *(Type of report and inclusive dates)*
Technical Report

5. AUTHOR(S) *(First name, middle initial, last name)*

Campbell B. Read

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
January 15, 1970	8	6

8a. CONTRACT OR GRANT NO. N00014-68-A-0515	9a. ORIGINATOR'S REPORT NUMBER(S)
b. PROJECT NO. NR 042-260	53
c.	9b. OTHER REPORT NO(S) *(Any other numbers that may be assigned this report)*
d.	

10. DISTRIBUTION STATEMENT
This document has been approved for public release and sale; its distribution is unlimited. Reproduction in whole or in part is permitted for any purpose of the United States Government.

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY
	Office of Naval Research

13. ABSTRACT

$\{X_n\}$ is a sequence of i.i.d. r.v.'s with common p.d.f. $p_\theta(\cdot)$, and tests of $H_0:\theta = \theta_0$ against $H_1:\theta = \theta_1$ are discussed. A Partial Sequential Probability Ratio Test operates exactly like a Wald SPRT, except that n observations are made before the stopping rule is applied. It is shown that in the class C_n of test procedures requiring at least n observations, (1) the Bayes risk with fixed losses for wrong decision and fixed cost is minimized by a PSPRT, (2) the boundaries of this PSPRT are those of the Wald SPRT which minimizes the Bayes risk in the class C_0, and (3) for all tests having upper bounds α and β on the probabilities of wrong dicision, and having finite ASN's, the PSPRT with error probabilities α and β minimizes the ASN at $\theta = \theta_0$ and $\theta = \theta_1$. These are the optimum properties which Wald's SPRT has in the class C_0 .

DD FORM 1 NOV 65 1473 (PAGE 1)
S/N 0101-807-6811

UNCLASSIFIED
Security Classification

A-31408