NASA CONTRACTOR REPORT

NASA CR-436

AN OPTIMAL PROPERTY OF SEQUENTIAL DECISION PROCEDURES RELATED TO BAYES SOLUTIONS

by Kenzo Seo

Prepared under Grant No. NSG-569 *by*
COLORADO STATE UNIVERSITY
Fort Collins, Colo.
for

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION • WASHINGTON, D. C. • APRIL 1966

NASA CR-436

AN OPTIMAL PROPERTY OF SEQUENTIAL DECISION PROCEDURES RELATED TO BAYES SOLUTIONS

By Kenzo Seo

Prepared under Grant No. NSG-569 by
COLORADO STATE UNIVERSITY
Fort Collins, Colo.

for

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION

For sale by the Clearinghouse for Federal Scientific and Technical Information
Springfield, Virginia 22151 – Price $0.10

1. Summary

It is well known that for testing a simple statistical hypothesis against a simple alternative hypothesis a sequential probability ratio test is best in the sense that it requires on the average, under both hypotheses, the minimum number of observations. This is usually called the optimal property of a sequential probability ratio test. A similar optimal property is defined for sequential decision procedures for testing k statistical hypotheses and its connection with Bayes solutions is investigated.

2. Definitions

We shall consider the problem of testing k simple statistical hypotheses $H_1, H_2, \ldots, H_k$. Each hypothesis specifies a probability measure on a Borel field of subsets of a space X. The k measures are considered as k possible distributions over X. To each sample point χ in X a decision procedure assigns a pair of integers (m, i). This means that when χ is the sample sequence the procedure observes the first m coordinates of χ and decides that H_i is the true hypothesis. To each decision procedure we shall assign a vector v of its expected sample sizes and error probabilities. Since the properties of decision procedures we shall consider are defined in terms of their expected sample sizes and error probabilities, we need not distinguish two procedures having the same vector of these quantities. Therefore, a decision procedure will be identified with the corresponding vector. We shall denote by $n(v)$ the vector whose i-th component $n_i(v)$ is the expected sample size of the procedure v under the hypothesis H_i and denote by $\alpha(v)$ the matrix whose ij-th element is the probability that v chooses H_j as true when really H_i should be chosen. Let g be a k-dimensional vector with non-negative components g_i such that $g_1 + g_2 + \ldots + g_k = 1$, and let ℓ be a $k \times k$ matrix (ℓ_{ij}) such that $\ell_{ij} \geqq o$ and $\ell_{ii} = o$. If g_i is the a priori probability of H_i and if ℓ_{ij} is the loss for choosing h_j when H_i is true, the risk of a procedure v is defined as $R(g, \ell, v) = \Sigma_i g_i (n_i(v) + \Sigma j \; \ell_{ij} \; \alpha_{ij} \; (v))$.

Let $\mathcal{V}$ be a class of decision procedures. For example $\mathcal{V}$ may be the class of all sequential procedures with finite expected sample sizes, or the class of all procedures which take at most m observations, etc. The analogue of the optimal property of a sequential probability ratio test is now stated as follows:

(P_1): A procedure v^* in $\mathcal{V}$ has property (P_1) if for any v in $\mathcal{V}$ $\alpha_{ij}(v) \leqq \alpha_{ij}(v^*)$ for all $i \neq j$ imply $n_i(v) \geqq n_i(v^*)$. for all i.

Let (P_2) be defined as follows:

(P_2): A procedure v^* in $\mathcal{V}$ has property (P_2) if for any vector g of a priori probability there is a matrix ℓ with $\ell_{ij} \geqq o$ and $\ell_{ii} = o$ such that $R(g, \ell, v^*) \leqq R(g, \ell, v)$ for all v in $\mathcal{V}$. Note that (P_2) is stronger than the property of a Bayes solution in which both g and ℓ are fixed.

3. The Main Result

From the statements of (P_1) and (P_2) it is easily seen that (P_2) implies (P_1). We give a sufficient condition under which a vector satisfying (P_1) also satisfies (P_2).

Theorem 1. In order for a vector v^* in $\mathcal{V}$ with property (P_1) to satisfy (P_2) it is sufficient that

1) $\mathcal{V}$ is a convex subset of the corresponding vector space, and

2) there exists a vector v' in $\mathcal{V}$ such that $\alpha_{ij}(v') < \alpha_{ij}(v^*)$ for all i and j, $i \neq j$.

Note that condition 1) is satisfied in most of the statistical decision problems if we include randomized decision procedures. Condition 2) is satisfied for example if $\mathcal{V}$ contains all sequential procedures with finite expected sample sizes. The proof is based on the following theorem on concave programmings [1].

Let $f(\chi) = \langle f_1(\chi), \ldots, f_m(\chi)\rangle$ and $g(\chi) = \langle g_1(\chi), \ldots, g_n(\chi)\rangle$ be two vector valued functions defined on a convex subset X of Euclidean space. We shall say that a vector χ^* in X is a solution of the "uniform maximum problem" if $g(\chi^*) \geqq o$ and for any χ in X, $g(\chi) \geqq o$ implies $f(\chi) \leqq f(\chi^*)$.

Theorem 2. Assume that functions $f(\chi)$ and $g(\chi)$ are concave on X, and $g(\chi) > o$ for some χ in X. Then a vector χ^* is a solution to the uniform maximum problem if and only if, the following conditions hold: for all $z > o$, there exists a vector $y^* = y^*(z) \geqq o$ such that for Lagrangian expression

$$L(\chi,y,z) = z \cdot f(\chi) + y \cdot g(\chi)$$

the saddle-point inequalities

$$L(\chi,y^*,z) \leqq L(\chi^*,y^*,z) \leqq L(\chi^*,y,z)$$

hold for all χ in X and $y \geqq o$.

Proof of theorem 1. Note that a vector v^* with property (P_1) is a solution to the uniform maximum problem with X replaced by $\mathcal{V}$, $f(\chi)$ by $-n(v)$, and $g(\chi)$ by $\alpha(v^*)-\alpha(v)$. (Here we consider $\alpha(v)$ and ℓ as $k(k-1)$ dimensional vectors of their off diagonal elements.) Therefore, by theorem 2 the conditions 1) and 2) imply the existance of $\ell^* \geqq o$ for each $g > o$ such that the pair (v^*, ℓ^*) is a saddle-point of the corresponding Lagrangian form. It then follows that

$$\Sigma_i g_i n_i(v^*) + \Sigma_{\substack{i,j\\ i\neq j}} \ell^*_{ij}\alpha_{ij}(v^*) \leqq \Sigma_i g_i n_i(v) + \Sigma_{\substack{i,j\\ i\neq j}} \ell^*_{ij}\alpha_{ij}(v)$$

for all v in $\mathcal{V}$. Setting $\ell = (\ell_{ij})$ where $\ell_{ij} = \ell^*_{ij}/g_i$ we see that v^* satisfies (P_2).

<u>Remark</u>. The property (P_2) seems to be easier to check than (P_1) in most of the decision problems.

NASA-Langley, 1966