June 1963

NOTES ON THE OPTIMUM CHARACTER OF THE SEQUENTIAL PROBABILITY RATIO TEST

Prepared by Morris Skibinsky

Special Report No. 8

University of Minnesota

Minneapolis, Minnesota

PREFACE

The set of notes which follows is a revision and extension of lecture notes which I prepared during the summer of 1960 at Purdue University, with partial support from the Purdue Research Foundation. The present version derives from lectures given to a staff seminar at the University of Minnesota during the winter quarter 1962. It contains a proof which follows the essential outline of the original as given by Wald and Wolfowitz (first three sections) and which was contained in somewhat cruder form in the original notes. Section 4 is new and is an expansion of the proof due to Le Cam that the sequential probability ratio test (viewed as a Bayes rule) has the invariance properties which lead to its optimality. (See E. L. Lehman's book "Testing Statistical Hypotheses" pp. 107-109.) Section 4 may be substituted for section 3 with no loss in continuity for the overall proof. The alternative proof of lemma 1.6, p. 23, is also taken from the optimality proof as it appears in Lehman's book.

It was not my original object, nor is it my object in the present notes, to find an essentially new or shorter way of proving the optimality of the sequential probability ratio test, but rather to make clear in a rigorous way the essential mechanisms (somewhat modified in this treatment) of the original authors. The introduction of section 4 and the alternative proof of lemma 1.6 provides the reader with a comparison of techniques. A discussion of the recent papers by Burkholder & Wijsman, and by Mathes, both in the Annals of Mathematical Statistics Volume 34, March 1963, is not included in these notes since their appearance followed the completion of this work. However, with reference to techniques employed in the latter paper, the reader who is interested should also see lemma 3.4 and theorem 3.1, p. 343 of a paper by the undersigned in the Annals of Mathematical Statistics Volume 31, June 1960.

It is apropos here to discuss certain omissions in the original paper with which others besides myself may have had difficulty. In lemma 1 of the original, a Bayes solution to the two-decision problem is exhibited. Lemma 2 (essentially) shows that the rule advanced in lemma 1 is a sequential probability ratio test. The proof as it is given rests upon its Bayes property. Now a rigorous proof that the rule advanced in lemma 1 is a Bayes solution requires that it be shown to terminate with probability one under each hypothesis. The fact that this test is a sequential probability ratio test (which may be shown by considerations not immediately involved to have finite expected sample size) may <u>not</u> be invoked to prove this, since as presented, that fact rests upon its being a Bayes solution. In the notes which follow, this difficulty is circumvented. Moreover, the class of rules in which the optimality of the sequential probability ratio test holds is shown to be unrestricted. We remark that the proof which we employ to show the existence of and exhibit a Bayes solution (Theorem B, p. 36 of these notes) is based upon the corresponding proof of a more general result that is sketched in "Bayes Solutions to Sequential Decision Problems", by Wald & Wolfowitz, Annals of Mathematical Statistics, Volume 21 (1950) pp. 82-99. Lemma 2.5 of these notes derives from the same source.

Lemma 8 of the original paper makes no use of the second limit derived in lemma 7. Such an omission requires that a sequential probability ratio test of one density against another have positive probability of choosing the first density when the second density is true. This in turn requires that the second density be less than the first density on a set of positive probability according to the second density. No assumptions concerning the two densities were made in the original paper except that they were distinct. In these notes, the second limit of lemma 7 (corollary 3.5 in the notes) is employed in the proof of lemma 8 (lemma 3.6 in the notes) and the above mentioned restriction is not required. A proposition, (2.13), p. 31, is proved concerning the error probabilities of sequential probability ratio tests which is required for the proof of lemma 8, but which does not appear in the original paper.

The notation here used is considerably changed from that of the original paper, although some of the forms in the original are maintained. In particular, it should be noted that the terminal decision function of these notes is one minus that of the original. New notation as it is introduced is marked by a number in square brackets at the extreme left of the page on the line in which it occurs. Round brackets containing formula numbers are used in the usual way. As an aid in keeping track of notation, an index of notation is given on the last page of these notes.

Morris Skibinsky
Statistics Department
University of Minnesota
June 17, 1963

NOTES ON THE OPTIMUM CHARACTER OF THE SEQUENTIAL PROBABILITY RATIO TEST BY A. WALD AND J. WOLFOWITZ, A. M. S. VOL. 19 (1948) pp. 326-339

Prepared by Morris Skibinsky

1. Introduction

We are given a sequence of independent random variables

[1] $$X = (X_1, X_2, \ldots)$$

with a common distribution function known a priori to be one of two specified distribution functions. We shall suppose that the probability measures which correspond to these distribution functions are both absolutely continuous with respect to Lebesgue measure or else that both are absolutely continuous with respect to counting measure (our notation will conform to the former case) and denote by

[2] $$f_0,\ f_1$$

corresponding probability densities. In addition, to avoid trivialities, we assume that these probability measures assign positive probability to the set on which the two densities are positive and unequal. Let

[3] $$(\mathfrak{X}, \mathcal{F})$$

be the measurable space determined by the space $\mathfrak{X}$ of infinite sequences (we suppose for convenience that this is the range space of X), and the smallest σ-field, $\mathcal{F}$, of its subsets which contains the cylinder sets with bases which are finite dimensional Borel sets. (See pp. 59-62, M. Loeve "Probability Theory".)

Let

[4] $$P_0,\ P_1$$

denote, respectively, the unique probability measures on $(\mathfrak{X}, \mathcal{F})$ induced by the probability measures that determine f_0, f_1, and consistent with these measures in the precise manner indicated in Theorem B, p. 157, Paul Halmos' "Measure Theory" (which proves their existence). See also pp. 90-94 M Loeve "Probability Theory". Let

[5] $$E_P$$

denote the expectation operator relative to a probability measure P on $(\mathfrak{X}, \mathfrak{F})$. When there is no possibility of confusion we shall write

[6] $$E_i$$

for E_{P_i}.

[7] A (non-randomized) sample size function (s.s.f.) is any measureable function n on $\mathfrak{X}$ to the non-negative integers and ∞ such that

(i) n is the zero function or n is identically positive

(ii) $\{x: n(x) \gtreqless j\}$ are cylinder sets whose respective bases are Borel sets in j dim sp. and form a partition thereof, j=1,2,... .

[8] Define $C_0(x) \overset{x\in\mathfrak{X}}{\equiv} \mathfrak{X}$ and for j=1,2,... and each $x \in \mathfrak{X}$ define

$$C_j(x) = \{z \in \mathfrak{X} : z_1 = x_1, \ldots, z_j = x_j\}$$

Then observe that for any sample size function n, we have

$$n(x) = j \implies n(y) = j, \quad \text{all } y \in C_j(x).$$

[9] A (non-randomized) terminal decision function (t.d.f.) for a sample size function n, is any measureable indicator function φ_n on $\mathfrak{X}$ whose value at any point x of $\mathfrak{X}$ depends only on the first n(x) components of x. i.e.,

$$\varphi_n(x) = i \implies \varphi_n(y) = i, \quad \text{all } y \in C_{n(x)}(x), \quad i=0,1$$

[10] A (non-randomized) rule (for deciding between f_0 and f_1) is a pair (n, φ_n) consisting of a sample size function n and a terminal decision function φ_n for n.

Associated with each rule (n, φ_n) we have the four constants, $E_i n$, i=0,1 and

[11] $$Q_i(\varphi_n) = P_i\{\varphi_n = 1-i\}$$

Let

$$G = \{g=(g_0, g_1): g_0+g_1=1, g_0, g_1 \geq 0\}$$

$$G^0 = \{g \in G: g_0, g_1 > 0\}$$

$$\Lambda = \{\lambda=\begin{pmatrix} 0 & \lambda_0 \\ \lambda_1 & 0 \end{pmatrix}: \lambda_0, \lambda_1 > 0\}$$

A criterion of the goodness of any rule (n, φ_n) relative to $g \in G$ and $\lambda \in \Lambda$ is

[12] $$R(g, \lambda|n, \varphi_n) = \Sigma\, g_i[E_i n + \lambda_i Q_i(\varphi_n)]$$

Let $\mathcal{Q}$ be a class of rules; $g \in G$, $\lambda \in \Lambda$.

[13] A Bayes rule in $\mathcal{Q}$ relative to the pair g, λ (a Bayes g, λ rule in $\mathcal{Q}$) is a rule $(n^*, \varphi^*_{n^*}) \in \mathcal{Q}$ such that

$$R(g, \lambda|n^*, \varphi^*_{n^*}) \leq R(g, \lambda|n, \varphi_n), \text{ all } (n, \varphi_n) \in \mathcal{Q}.$$

Let

[14] $$\rho(g, \lambda|\mathcal{Q}) = \inf_{(n, \varphi_n) \in \mathcal{Q}} R(g, \lambda|n, \varphi_n)$$

<u>Remark</u>

Clearly, any rule in $\mathcal{Q}$ whose average (g, λ) risk [12] can attain to this infimum will be Bayes g, λ in $\mathcal{Q}$ and if no such rule exists, there is no Bayes (g, λ) rule in $\mathcal{Q}$.

We will as convenience dictates sometimes denote a rule (n, φ_n) by a single letter, say S. e.g.,

$$\rho(g, \lambda|\mathcal{Q}) = \inf_{S \in \mathcal{Q}} R(g, \lambda|S).$$

There are three main classes of rules with which we shall be concerned

$$\mathcal{S} = \{\text{rules } (n, \varphi_n): P_i\{n < \infty\} = 1\}$$

$$\mathcal{S}_0 = \{(n, \varphi_n) \in \mathcal{S}: n \text{ is identically positive}\}$$

$$\mathcal{S}_0' = \{(n, \varphi_n): n(x) \equiv 0\}$$

We shall write

[15] $$\rho(g, \lambda) = \rho(g, \lambda|\mathcal{S}), \quad \rho^*(g, \lambda) = \rho(g, \lambda|\mathcal{S}_0), \quad \rho_0(g, \lambda) = \rho(g, \lambda|\mathcal{S}_0')$$

The easiest of the above infima to evaluate is of course $\rho_0(g, \lambda)$ for $\mathcal{S}_0'$ contains precisely two rules S_0 and S_1, say with t.d.f.'s identically 0 and identically 1, respectively.

$$R(g, \lambda|S_i) = g_{1-i}\lambda_{1-i}$$

so that

(1.1) $$\rho_0(g, \lambda) = \min_{i=0,1} R(g, \lambda|S_i) = \min[g_0\lambda_0, g_1\lambda_1]$$

In addition observe that trivially

(1.2) $$\rho(g, \lambda) = \min[\rho_0(g, \lambda), \rho^*(g, \lambda)]$$

We define the operation o between two vectors

$$a = (a_0, a_1), \quad b = (b_0, b_1)$$

of non-negative components by

[16] $$a o b = (a_0 b_0, a_1 b_1)/(a \cdot b), \qquad a \cdot b = a_0 b_0 + a_1 b_1$$

so that $a o b$ is defined so long as $a \cdot b \neq 0$. Note that always $a o b \in G$.

We shall take

$$(a o b)_i = a_i b_i / a \cdot b$$

Let

[17] $$f(x_j) = (f_0(x_j), f_1(x_j)), \quad j=1,2,\ldots.$$

[18] $$f_{ij}(x) = \begin{cases} \prod_{k=1}^{j} f_i(x_k), & j=1,2,\ldots \\ 1, & j=0 \end{cases}, \quad i=0,1$$

[19] $$f^{(j)}(x) = (f_{0j}(x), f_{1j}(x)) \qquad f^{(n)}: \quad (f^{(n)})(x) = f^{(n(x))}(x)$$

<u>Observe that</u>

$$g o f^{(j)}(x) = g o f(x_1) o f(x_2) o \ldots\ldots o f(x_j), \quad j=0,1,2,\ldots.$$

$$= (g_0 f_{0j}(x), g_1 f_{1j}(x))/[g_0 f_{0j}(x) + g_1 f_{1j}(x)]$$

and that this is simply the vector of "a posteriori probabilities" which for $j=0,1,2,\ldots$ is of course always an element of G.

Let

$$r(x|j) = (1,\ 1)\circ f^{(j)}(x).$$

This is the normalized likelihood vector at the $j^{\underline{th}}$ step. Observe that

(1.3) $$g\circ f^{(j)}(x) \equiv g\circ r(x|j),$$

where the equivilence holds for all $g \in G$, for $j=0,1,2,\ldots$ and for all $x \in \mathfrak{X}$ such that both sides are defined.

Let

[20] $$P_g = g_0P_0 + g_1P_1, \quad g \in G$$

Note that for each $g \in G$, P_g is a probability measure on $(\mathfrak{X}, \mathfrak{F})$.

<u>Lemma 1.1</u>

Let $g \in G$, $\lambda \in \Lambda$ be arbitrary, fixed. Then for any t.d.f. φ_n for a s.s.f. such that $P_i\{n < \infty\} = 1$, $i=0,1$, we have that

$$\sum_{i=0}^{1} g_i\lambda_iQ_i(\varphi_n) = E_{P_g} R(g\circ f^{(n)},\ \lambda|S_{\varphi_n})$$

$$= E_{P_g}[(g\circ f^{(n)})_{1-\varphi_n}\lambda_{1-\varphi_n}]$$

An immediate consequence of this lemma is the following

<u>Corollary 1.11</u>

Let $g \in G$, $\lambda \in \Lambda$ be arbitrary, fixed. Then for any rule $(n,\ \varphi_n) \in \mathcal{S}$, we have that

$$R(g,\ \lambda|n,\ \varphi_n) = E_{P_g}[n + (g\circ f^{(n)})_{1-\varphi_n}\lambda_{1-\varphi_n}]$$

Proof of Lemma 1.1

If $n(x) \equiv 0$, the lemma is trivially true. For in this case, either $\varphi_n(x) \equiv 0$ so that $Q_i(\varphi_n) = i$ or $\varphi_n(x) \equiv 1$ so that $Q_i(\varphi_n) = 1-i$.

Now let us suppose that n is a positive s.s.f. .

By hypothesis, $P_i\{n < \infty\} = 1$, $i=0,1$. Hence

$$\mathfrak{X} = \sum_{j=1}^{\infty} \sum_{k=0}^{1} \{n=j,\ \varphi_n=1-k\} + N,$$

where N is null according to both P_0 and P_1, and hence according to P_g, for all $g \in G$.

Thus

$$E_{P_g}[(g\circ f^{(n)})_{1-\varphi_n} \lambda_{1-\varphi_n}] = \sum_{j=1}^{\infty} \sum_{k=0}^{1} \lambda_k \int_{\{n=j,\ \varphi_n=1-k\}} (g\circ f^{(j)})_k \, dP_g$$

Let T_j denote for each positive integer j the projection map on $\mathfrak{X}$ defined by

[21] $$T_j(x) = (x_1, x_2, \ldots, x_j) = \underline{x}_j, \text{ say.}$$

Then, by definition of P_i on $(\mathfrak{X}, \mathfrak{F})$, the fact that $(g\circ f^{(j)})_k$ depends only upon the first j coordinates of its argument, and that $\{n=j,\ \varphi_n=1-k\}$ is a cylinder subset of $\mathfrak{X}$ whose base is a j dimensional Borel set

$$\int_{\{n=j,\ \varphi_n=1-k\}} (g\circ f^{(j)})_k \, dP_g = \int_{T_j[n=j,\ \varphi_n=1-k]} (g\circ f^{(j)})_k \left(\sum_{i=0}^{1} g_i f_{ij} \right) d\mu_j,$$

where the integrand on the right hand side is to be regarded as a function on j-dimensional Euclidean space, rather than on $\mathfrak{X}$, and μ_j is j dimensional Lebesgue (or counting) measure. But then, by [16] and [19], this right hand side may be written

$$g_k \int_{T_j[n=j,\ \varphi_n=1-k]} f_{kj} \, d\mu_j = g_k \int_{\{n=j,\ \varphi_n=1-k\}} dP_k$$

Hence

$$E_{P_g}[(g\circ f^{(n)})_{1-\varphi_n} \lambda_{1-\varphi_n}] = \sum_{k=0}^{1} g_k \lambda_k \sum_{j=1}^{\infty} \int_{\{n=j,\ \varphi_n=1-k\}} dP_k$$

$$= \sum_{k=0}^{1} g_k \lambda_k Q_k(\varphi_n) .$$

Corollary 1.12

Let φ_n be an arbitrary t.d.f. for any s.s.f. n such that $P_i\{n < \infty\}$, i=0,1. Let $\mathcal{D}_n$ denote the class of all rules which have the particular s.s.f. n. Then for arbitrary $g \in G$, $\lambda \in \Lambda$, the following statements are equivalent outside of a P_g null set.

(i) (n,φ_n^*) is a Bayes g, λ rule in $\mathcal{D}_n$.

(ii) $(g\circ f^{(n)})_{1-\varphi_n^*} \lambda_{1-\varphi_n^*} = \rho_0(g\circ f^{(n)}, \lambda)$

(iii) $\varphi_n^*(x) = \begin{cases} 1, & g_0 \lambda_0 f_{0n(x)}(x) < g_1 \lambda_1 f_{1n(x)}(x) \\ 0, & \quad > \end{cases}$.

(iv) $\varphi_n^*(x) = \begin{cases} 1, & r_1(x|n(x)) > \dfrac{g_0 \lambda_0}{g_0 \lambda_0 + g_1 \lambda_1} \\ 0, & \quad < \end{cases}$.

For each $g \in G$, $\lambda \in \Lambda$ and each s.s.f. n such that $P_i\{n < \infty\} = 1$, i=0,1, we define

[22] $$R(g, \lambda|n) = E_{P_g}[n + \rho_0(g\circ f^{(n)}, \lambda)].$$

By the two preceding corollaries, we then have

Corollary 1.13

Let (n, φ_n) be an arbitrary rule in $\mathcal{S}$, and let φ_n^* denote any t.d.f. for n which satisfies either (iii) or (iv) of corollary 1.12, then for arbitrary $g \in G$, $\lambda \in \Lambda$, we have that

$$R(g, \lambda|n, \varphi_n) \geq R(g, \lambda|n,\varphi_n^*) = R(g, \lambda|n).$$

Corollary 1.14

For each $g \in G$, $\lambda \in \Lambda$

$$\rho(g, \lambda) = \inf_n R(g, \lambda | n), \qquad \rho^*(g, \lambda) = \inf_{n > 0} R(g, \lambda | n),$$

where $\inf_n$ is taken to denote the infimum over all s.s.f.'s such that $P_i\{n < \infty\} = 1$, $i=0,1$, and $\inf_{n>0}$, to denote the infimum over all positive s.s.f.'s such that $P_i\{n < \infty\} = 1$, $i=0,1$.

We shall adopt the following notation. For each $v \in \mathfrak{X}$, we shall take

[23] $$\underline{x}_j v = \begin{cases} (x_1, x_2, \ldots, x_j, v_1, v_2, \ldots), & j=1,2,\ldots \\ v, & j=0 \end{cases}$$

For each $g \in G$, $\lambda \in \Lambda$, each non-negative integer j, and each s.s.f. n such that $P_i\{n < \infty\} = 1$, $i=0,1$, we define the function $R_j(g, \lambda | n, \cdot)$ at each point in $\mathfrak{X}$ outside of a P_g null set by

[24] $$R_j(g, \lambda | n, x) = E_{P_{g \circ f^{(j)}(x)}} \left[n(\underline{x}_j \cdot) + \rho_0(g \circ f^{n(\underline{x}_j \cdot)}(\underline{x}_j \cdot), \lambda) \right].$$

We shall assume some arbitrary but fixed constant as the definition for this function on the P_g null set for which the right hand side above is undefined.

Observe that

(1.4) $$R_0(g, \lambda | n, x) \equiv R(g, \lambda | n).$$

When j is a positive integer we have

Lemma 1.2

For each $g \in G$, $\lambda \in \Lambda$, for each positive integer j, and for each s.s.f. n such that $P_i\{n < \infty\} = 1$, i=0,1, the function $R_j(g,\lambda|n, \cdot)$ is a version of the conditional probability

$$E_{P_g}[n + \rho_0(g\circ f^{(n)}, \lambda)|\mathcal{F}_j] .$$

Proof

$R_j(g, \lambda|n, \cdot)$ is clearly measureable with respect to the sub-σ-field of $\mathcal{F}$ given by

$$\{T_j^{-1}[A]: \text{ A is a j-dimensional Borel set}\} = \{A\times\mathfrak{X} : \text{ A is a j-dimensional Borel set}\}.$$

Thus, we need only show that if A is an arbitrarily given j-dimensional Borel set, then

$$(1.5)\qquad \int_{A\times\mathfrak{X}} R_j(g, \lambda|n, \cdot)dP_g = \int_{A\times\mathfrak{X}} [n + \rho_0(g\circ f^{(n)}, \lambda)]dP_g .$$

Observe first that by the definition of s.s.f. [7],

$$(1.6)\qquad n(x) \leq j \implies R_j(g, \lambda|n, x) = n(x) + \rho_0(g\circ f^{n(x)}(x), \lambda) .$$

Hence

$$(1.7)\qquad \int_{A\times\mathfrak{X}\cap\{n\leq j\}} R_j(g, \lambda|n, \cdot)dP_g = \int_{A\times\mathfrak{X}\cap\{n\leq j\}} [n + \rho_0(g\circ f^{(n)}, \lambda)]dP_g$$

On the other hand, by the definition of P_i on $(\mathfrak{X}, \mathcal{F})$, the fact that $R_j(g, \lambda|n, x)$ depends only on the first j coordinates of x, and that $A\times\mathfrak{X}\cap\{n > j\}$ is a cylinder subset of $\mathfrak{X}$ whose base is a j-dimensional Borel set, we have that

$$(1.8) \qquad \int_{A\times\mathfrak{X}\cap\{n>j\}} R_j(g, \lambda|n, \cdot)dP_g = \int_{A\cap T_j[n>j]} R_j(g, \lambda|n, \cdot) \sum_{i=0}^{1} g_i f_{ij}\, d\mu_j$$

where the integrand on the right hand side is to be regarded as a function on j-dimensional Euclidean space, rather than on $\mathfrak{X}$, and μ_j is, again, j-dimensional Lebesgue (or counting) measure.

Now

$$P_{g\circ f^{(j)}(x)} = \sum_{k=0}^{1} (g\circ f^{(j)}(x))_k P_k = \sum_{k=0}^{1} g_k f_{kj}(x) P_k \Big/ \sum_{i=0}^{1} g_i f_{ij}(x)$$

Thus, by [24]

$$(1.9) \qquad R_j(g, \lambda|n, x) \sum_{i=0}^{1} g_i f_{ij}(x)$$

$$= \sum_{k=0}^{1} g_k f_{kj}(x) \int_{\mathfrak{X}} [n(\underline{x}_j\ \cdot) + \rho_0(g\circ f^{n(\underline{x}_j\ \cdot)}(\underline{x}_j\ \cdot), \lambda)]dP_k$$

Now observe that to each $x \in \mathfrak{X}$ such that $n(x) > j$, there corresponds a partition of $\mathfrak{X}$, namely

$$\mathfrak{X} = \sum_{m=j+1}^{\infty} \{v \in \mathfrak{X} : n(\underline{x}_j\ v) = m\} + N_x$$

where N_x is null according to both P_0 and P_1. Hence, for each $x \in \mathfrak{X}$ such that $n(x) > j$, the right hand side of (1.9) may be written

$$\sum_{k=0}^{1} g_k f_{kj}(x) \sum_{m=j+1}^{\infty} \int_{\{v:\ n(\underline{x}_j\ v) = m\}} [m + \rho_0(g\circ f^{(m)}(\underline{x}_j\ \cdot), \lambda)]dP_k$$

By arguments strictly analogous to those used for (1.8), the integrals in the above expression may be respectively replaced by

$$\int_{T_{m-j}[\{v:\ n(\underline{x}_j\ v) = m\}]} [m + \rho_0(g\circ f^{(m)}(\underline{x}_j\ \bullet), \lambda)] f_{k,m-j}\, d\mu_{m-j}$$

where the integrand is to be regarded as a function on m-j dimensional Euclidean space rather than on $\mathfrak{X}$. Thus, the right hand side of (1.8) may be written

$$\sum_{m=j+1}^{\infty} \int_{A\times \mathfrak{X}_{m-j} \cap T_m[n=m]} [m + \rho_0(g\circ f^{(m)}, \lambda)] \sum_{k=0}^{1} g_k f_{km}\, d\mu_m$$

where the integrand for index m is to be regarded as a function on m-dimensional Euclidean space, and $\mathfrak{X}_{m-j}$ is taken to denote m-j dimensional Euclidean space. It now follows by a reversal in application of the arguments used for (1.8) that the right hand side of (1.8) may be written

$$\sum_{m=j+1}^{\infty} \int_{A\times \mathfrak{X} \cap \{n=m\}} [m + \rho_0(g\circ f^{(m)}, \lambda)] dP_g\ .$$

and this in turn is just

$$\int_{A\times \mathfrak{X} \cap \{n > j\}} [n + \rho_0(g\circ f^{(n)}, \lambda)] dP_g\ .$$

Together with (1.7) this now implies the desired result, (1.5).

<u>Remark</u>

As an extension to (1.6) in the above proof, we observe that by the definition of s.s.f. [7] and by [24], (1.10) $$n(x) = j \implies R_j(g, \lambda | n, x) = j + \rho_0(g \circ f^{(j)}(x), \lambda) .$$

For each non-negative integer j and for each $x \in \mathfrak{X}$, we define the class of s.s.f.'s

[25] $$\mathcal{S}_j(x) = \{n: n(x) > j, \ P_i\{n < \infty\} = 1, \ i=0,1\} ,$$

and for each $g \in G$, $\lambda \in \Lambda$, take

[26] $$\nu_j(x|g, \lambda) = \sup_{n \in \mathcal{S}_j(x)} [j + \rho_0(g \circ f^{(j)}(x), \lambda) - R_j(g, \lambda | n, x)] .$$

It is intuitively evident that the sign of this quantity indicates the existence or non-existence of a s.s.f. relative to which it is "worth while" with respect to the particular pair g, λ to continue observing components of X past X_j, when the realization of X_k has been x_k for $k=1,\ldots,j$.

Observe that by [18], [19]

$$g \circ f^{(0)}(x) \equiv g, \quad \text{all } x \in \mathfrak{X}, \ g \in G ,$$

so that by (1.4) and corollary 1.14, we have for each $g \in G$. $\lambda \in \Lambda$, that

(1.11) $$\nu_0(x|g, \lambda) \equiv \rho_0(g, \lambda) - \rho^*(g, \lambda) .$$

<u>Lemma 1.3</u>

Let $g \in G$, $\lambda \in \Lambda$ be arbitrary, fixed. Then $\nu_j(x|g, \lambda)$ depends upon j and x only through

[27] $$r(x|j) = \frac{(f_{0j}(x), f_{1j}(x))}{f_{0j}(x) + f_{1j}(x)} .$$

i.e. To each $h \in G$ such that $g \cdot h \neq 0$ and such that the set

$$(1.12) \qquad \{(k, y): r(y|k) = h\}$$

is non-empty, there corresponds a number, call it

[28] $$\gamma(h|g, \lambda),$$

such that

$$\nu_j(x|g, \lambda) = \gamma(h|g, \lambda)$$

for each pair (j, x) in the set (1.12).

<u>Proof</u>

Let (j, x), (k, y) be two arbitrarily given pairs such that

$$(1.13) \qquad r(x|j) = r(y|k) = h.$$

To prove the lemma, it will be sufficient to show that

$$(1.14) \qquad \nu_j(x|g, \lambda) = \nu_k(y|g, \lambda).$$

Suppose that

$$(1.15) \qquad \nu_j(x|g, \lambda) > \nu_k(y|g, \lambda) .$$

By [26] and the definition of supremum, there exists a s.s.f., n', say, in $\mathcal{S}_j(x)$ with the property that

$$(1.16) \qquad j + \rho_0(g \circ f^{(j)}(x), \lambda) - R_j(g, \lambda|n', x) > \nu_k(y|g, \lambda) ,$$

for otherwise, (1.15) could not hold.

Below, we shall produce a s.s.f.

$$n'' \in \mathcal{S}_k(y)$$

with the property that

$$(1.17) \quad k + \rho_0(g \circ f^{(k)}(y), \lambda) - R_k(g, \lambda|n'', y) = j + \rho_0(g \circ f^{(j)}(x), \lambda) - R_j(g, \lambda|n', y).$$

But this, in view of (1.16) will contradict the definition of $\nu_k(y|g, \lambda)$ and hence imply that

$$\nu_j(x|g, \lambda) \leq \nu_k(y|g, \lambda) .$$

Considerations of symmetry then show that the opposite inequality must also hold so that (1.14) will hold and the lemma be proved.

We make the following remarks which are essentially notational and easily verified. We have for any given non-negative integers s and t and arbitrary $z \in \mathfrak{X}$,

$$(1.18) \qquad g \circ f^{(s)}(z) = g \circ r(z|s)$$

$$(1.19) \qquad r(\underline{z}_s\ v|s+t) \equiv r(z|s) \circ r(v|t) .$$

Hence by (1.13)

$$g \circ f^{(j)}(x) = g \circ f^{(k)}(y) = g \circ h .$$

It follows that to satisfy (1.17) we need only find a s.s.f. $n'' \in \mathcal{S}_k(y)$ such that

$$R_k(g, \lambda|n'', y) - k = R_j(g, \lambda|n', x) - j .$$

Let us choose n" to be any s.s.f. whose definition on $C_k(y)$ is given by

$$(1.20) \qquad n''(\underline{y}_k\ v) = n'(\underline{x}_j\ v) + k - j, \quad v \in \mathfrak{X} .$$

But then since $n' \in \mathcal{S}_j(x)$, it follows that $n'' \in \mathcal{S}_k(y)$. In addition, by (1.18), (1.19), (1.20), we have for all $v \in \mathfrak{X}$,

$$g \circ f^{n''(\underline{y}_k\ v)}(\underline{y}_k\ v) = g \circ f^{n'(\underline{x}_j\ v)}(\underline{x}_j\ v)$$

so that

$$R_k(g, \lambda|n'', y) - k = E_{P_{g \circ h}}[n'(\underline{x}_j\ \cdot) - j + \rho_0(g \circ f^{n'(\underline{x}_j\ \cdot)}(\underline{x}_j\ \cdot), \lambda)]$$

$$= R_j(g, \lambda|n', x) - j .$$

This completes the proof.

[29] Let H denote the set of all points $h \in G$ such that

$$\{(k, y): r(y|k) = h\} \neq \Phi$$

By [28] (with the possible exception, when $g \notin G$, of a point h such that $g \cdot h = 0$), H is for each $g \in G$, $\lambda \in \Lambda$, the domain of definition for $\gamma(\cdot\,|g, \lambda)$. Observe that since $r(x|0) \equiv (\frac{1}{2}, \frac{1}{2})$, it follows that

(1.21) $$(\tfrac{1}{2}, \tfrac{1}{2}) \in H,$$

and hence by (1.11), for each $g \in G$, $\lambda \in \Lambda$,

(1.22) $$\gamma((\tfrac{1}{2}, \tfrac{1}{2})|g, \lambda) = \rho_0(g, \lambda) - \rho^*(g, \lambda) .$$

Lemma 1.4

Let $g \in G$, $\lambda \in \Lambda$ be arbitrary, fixed. Then for all $h \in H$ such that $g \cdot h \neq 0$, we have that

$$\gamma(h|g, \lambda) = \gamma((\tfrac{1}{2}, \tfrac{1}{2})|g \circ h, \lambda)$$

Proof

We note at the outset that the right hand side above is defined for all $h \in G$ such that $g \cdot h \neq 0$ and hence that it is defined for all h such that the left hand side is defined, namely all $h \in H$ such that $g \cdot h \neq 0$.

Let h be an arbitrary, fixed point in H such that $g \cdot h \neq 0$. By [29], there exists a pair (j, x) such that

$$r(x|j) = h$$

and by the previous lemma

$$\gamma(h|g, \lambda) = \nu_j(x|g, \lambda) .$$

In addition

$$g\circ f^{(j)}(x) \;=\; g\circ r(x|j) \;=\; g\circ h$$

and

$$g\circ f^{n(\underline{x}_j v)}(\underline{x}_j v) \;=\; g\circ h\circ f^{n(\underline{x}_j v)-j}(v), \qquad \text{all } n \in \mathcal{S}_j(x) \;.$$

Hence by [26] and then [24]

$$(1.23) \qquad \gamma(h|g, \lambda) \;=\; \rho_0(g\circ h, \lambda) - \inf_{n \in \mathcal{S}_j(x)} [R_j(g, \lambda|n, x) - j]$$

$$= \rho_0(g\circ h, \lambda) - \inf_{n \in \mathcal{S}_j(x)} E_{P_{g\circ h}}[n(\underline{x}_j \,\cdot) - j + \rho_0(g\circ h\circ f^{n(\underline{x}_j \,\cdot)-j}, \lambda)] \;.$$

Now let M denote the mapping

$$M: \quad \mathcal{S}_j(x) \longrightarrow \{n:\, n > 0,\; P_i\{n < \infty\} = 1,\; i=0,1\}$$

defined by

$$M(n) \;=\; \bar{n}(\,\cdot\,|n), \qquad\qquad n \in \mathcal{S}_j(x)$$

where

$$(1.24) \qquad \bar{n}(v|n) \;=\; n(\underline{x}_j v) - j, \qquad v \in \mathfrak{X} \;.$$

M is an onto map. i.e.

$$(1.25) \qquad \{\bar{n}(\cdot|n):\, n \in \mathcal{S}_j(x)\} \;=\; \{n:\, n > 0,\; P_i\{n < \infty\} = 1,\; i=0,1\} \;.$$

By (1.24), the second term on the right hand side of (1.23) may be written

$$\inf_{n \in \mathcal{S}_j(x)} E_{P_{g\circ h}} [\bar{n}(\,\cdot\,|n) + \rho_0(g\circ h\circ f^{n(\,\cdot\,|n)}, \lambda)]$$

By (1.25), [22] and corollary 1.14, this is just $\rho^*(g\circ h, \lambda)$. But by (1.22), this proves the lemma.

As a notational convenience we now define the function $\hat{\gamma}$ on $G \times \Lambda$ by

[30] $$\hat{\gamma}(g, \lambda) = \rho_0(g, \lambda) - \rho^*(g, \lambda), \quad g \in G, \lambda \in \Lambda .$$

We may use this to summarize the results of lemmas 1.3 and 1.4 as follows. For each $g \in G$, $\lambda \in \Lambda$, for each non-negative integer j, and for each $x \in \mathfrak{X}$, we have that

(1.26) $$\nu_j(x|g, \lambda) = \gamma(r(x|j)|g, \lambda) = \hat{\gamma}(g \circ f^{(j)}(x), \lambda).$$

For convenient reference in the proofs which follow, observe that we may write

(1.27) $$\rho_0(g, \lambda) = \begin{cases} g_0\lambda_0, & g_1 \geq \dfrac{\lambda_0}{\lambda_0+\lambda_1} \\ g_1\lambda_1, & \leq \end{cases}, \quad \lambda \in \Lambda,$$

and hence for any rule (n, φ_n), we have by [12] that for each $\lambda \in \Lambda$,

(1.28) $$\rho_0(g, \lambda) - R(g, \lambda|n, \varphi_n)$$

$$= \begin{cases} g_0[-E_0n+\lambda_0(1-Q_0(\varphi_n))] + g_1[-E_1n-\lambda_1Q_1(\varphi_n)], & g_1 \geq \dfrac{\lambda_0}{\lambda_0+\lambda_1} \\ g_0[-E_0n-\lambda_0Q_0(\varphi_n)] + g_1[-E_1n+\lambda_1(1-Q_1(\varphi_n))], & \leq \end{cases}.$$

<u>Lemma 1.5</u>

Let $\lambda \in \Lambda$ be arbitrary, fixed. If

(1.29) $$\{g \in G: \hat{\gamma}(g, \lambda) > 0\}$$

is non-empty, then it is an interval subset of G^0 which contains the point

(1.30) $$\frac{(\lambda_1, \lambda_0)}{\lambda_0+\lambda_1} .$$

If it is not the degenerate interval consisting of this point alone, it is a non-degenerate interval. $\hat{\gamma}(\cdot\ ,\lambda)$ is monotonic on the interval to each side of this point and is maximum there.

Proof

It is first of all easily verified that

$$\hat{\gamma}((0,1),\lambda) = \hat{\gamma}((1,0),\lambda) = -1 , \tag{1.31}$$

from which it follows that (1.29) is of necessity a subset of G^0. Conceivably, (1.29) might consist of the single point (1.30). In this case, the lemma would be trivially true. Now suppose that (1.29) does not consist of the single point (1.30). Since by hypothesis (1.29) is non-empty there exists a point

$$g^* = (g_0^*, g_1^*) \neq \frac{(\lambda_1, \lambda_0)}{\lambda_0+\lambda_1}$$

and such that

$$\hat{\gamma}(g^*, \lambda) > 0. \tag{1.32}$$

Now either

$$g_1^* > \frac{\lambda_0}{\lambda_0+\lambda_1} \quad \text{or} \quad g_1^* < \frac{\lambda_0}{\lambda_0+\lambda_1} \quad . \tag{1.33}$$

Suppose the first of these inequalities to hold. We will prove below that if g is any point in G^0 such that

$$\frac{\lambda_0}{\lambda_0+\lambda_1} \leq g_1 < g_1^* \tag{1.34}$$

then

$$\hat{\gamma}(g, \lambda) \geq \hat{\gamma}(g^*, \lambda) .$$

Supposing the second inequality of (1.33) to hold, a strictly analogous argument (not repeated) leads to the result that

$$g_1^* < g_1 \leq \frac{\lambda_0}{\lambda_0+\lambda_1} \implies \hat{\gamma}(g, \lambda) \geq \hat{\gamma}(g^*, \lambda) .$$

But in view of (1.32), the above results prove the lemma.

Thus, suppose the first inequality of (1.33) to hold. Let

$$0 < \epsilon < \hat{\gamma}(g^*, \lambda)$$

By [30] and corollary 1.14, we have that

$$\hat{\gamma}(g^*, \lambda) = \sup_{n > 0} [\rho_0(g^*, \lambda) - R(g^*, \lambda | n)] ,$$

where $\sup_{n>0}$ is taken to denote the supremum over all positive s.s.f.'s such that $P_i\{n < \infty\} = 1$, $i=0,1$. Hence there exists a positive s.s.f. n', say, such that

$$0 < \hat{\gamma}(g^*, \lambda) - \epsilon < \rho_0(g^*, \lambda) - R(g^*, \lambda | n') . \tag{1.35}$$

By corollary 1.13 and (1.28) and our assumption that the first inequality of (1.33) holds, the right hand side of the above inequality may be written

$$(1-g_1^*)[-E_0 n' + \lambda_0(1-Q_0(\varphi^*_{n',g^*,\lambda}))] + g_1^*[-E_1 n' - \lambda_1 Q_1(\varphi^*_{n',g^*,\lambda})] ,$$

where $\varphi^*_{n',g,\lambda}$ is any t.d.f. for n' which satisfies (iii) or (iv) of corollary 1.12. By (1.35), the above expression is positive. But

$$-E_1 n' - \lambda_1 Q_1(\varphi^*_{n',g^*,\lambda}) < 0 ,$$

since n' is a positive s.s.f. Moreover $g^* \epsilon G^0$. It follows that

$$-E_0 n' + \lambda_0(1-Q_0(\varphi^*_{n',g^*,\lambda})) > 0 .$$

It now becomes clear that if g is any point of G^0 which satisfies (1.34), we must have that

$$\rho_0(g^*, \lambda) - R(g^*, \lambda | n') < \rho_0(g, \lambda) - R(g, \lambda | n', \varphi^*_{n',g^*,\lambda}) .$$

By corollary 1.13 the right hand side of the above inequality is bounded above by

$$\rho_0(g, \lambda) - R(g, \lambda|n'),$$

and this in turn is bounded above by $\hat{\gamma}(g, \lambda)$. Thus, by (1.35) we have that

$$\hat{\gamma}(g^*, \lambda)-\epsilon < \hat{\gamma}(g, \lambda) .$$

Since $\epsilon > 0$ may be arbitrarily small, the desired result follows.

<u>Lemma 1.6</u>

Let $\lambda \in \Lambda$ be arbitrary, fixed. Then $\hat{\gamma}(\cdot, \lambda)$ is continuous on G (one sidedly at the endpoints of G).

<u>Proof</u>

For $i=0,1$, let S_{1i} denote the rule (n, φ_n) defined by

$$n(x) \equiv 1, \qquad \varphi_n(x) \equiv i, \qquad \text{all } x \in \mathfrak{X} .$$

Then for all $g \in G$,

$$R(g, \lambda|S_{1i}) = 1+g_{1-i}\lambda_{1-i} ,$$

so that for all $g \in G$,

$$(1.36) \qquad 1 \leq \rho^*(g, \lambda) \leq \min_{i=0,1} R(g, \lambda|S_{1i}) = 1+\rho_0(g, \lambda).$$

It then follows that

$$(1.37) \qquad -1 \leq \hat{\gamma}(g, \lambda) \leq -1+\rho_0(g, \lambda), \qquad \text{all } g \in G.$$

Since $\rho_0(g, \lambda)$ tends to zero as g tends to either endpoint of G, we have by the above inequality and (1.31) that $\hat{\gamma}(\cdot, \lambda)$ is one sidedly continuous at the endpoints of G.

Now recall again that by definition [30] and corollary 1.14,

$$(1.38) \qquad \hat{\gamma}(g, \lambda) = \sup_{n > 0} [\rho_0(g, \lambda) - R(g, \lambda|n)], \qquad \text{all } g \in G,$$

where $\sup_{n>0}$ denotes supremum over all positive s.s.f.'s such that $P_i\{n < \infty\} = 1$, i=0,1. In addition, we observe for later reference, using (1.28) and corollary 1.13, that if n is any given s.s.f. in the above class and if g', g" are any two points in G such that either

$$g_1', g_1'' \leq \lambda_0/(\lambda_0+\lambda_1) \quad \text{or} \quad g_1', g_1'' \geq \lambda_0/(\lambda_0+\lambda_1) , \tag{1.39}$$

then the following inequality holds.

$$|\rho_0(g', \lambda)-R(g', \lambda|n)-[\rho_0(g'', \lambda)-R(g'', \lambda|n)]| \leq (\lambda_0+\lambda_1+E_0n+E_1n)|g_1'-g_1''| . \tag{1.40}$$

Let g* be an arbitrary, fixed point in G^0 such that

$$g^* \neq (\lambda_1, \lambda_0)/(\lambda_0+\lambda_1) .$$

We will show that $\hat{\gamma}(\cdot, \lambda)$ is continuous at g*. It will then remain only to show that $\hat{\gamma}(\cdot, \lambda)$ is continuous at $(\lambda_1, \lambda_0)/(\lambda_0+\lambda_1)$, for the lemma to be proved.

Let ϵ be an arbitrarily given positive number. To show that $\hat{\gamma}(\cdot, \lambda)$ is continuous at g*, we need only show that there exists a number $\delta_\epsilon > 0$, with the property that if g', g" are arbitrary points in the neighborhood

$$\{g \in G^0: |g_1-g_1^*| < \delta_\epsilon\} = N(g^*), \text{ say,} \tag{1.41}$$

then

$$|\hat{\gamma}(g', \lambda) - \hat{\gamma}(g'', \lambda)| < \epsilon. \tag{1.42}$$

By corollary 1.14, we have that corresponding to each $g \in G^0$, there exists a positive s.s.f. n_g, say, such that

$$R(g, \lambda|n_g) < \rho^*(g, \lambda) + \epsilon/2, \quad \text{all } g \in G^0, \tag{1.43}$$

i.e., by [30], such that

$$\hat{\gamma}(g, \lambda) - \epsilon/2 < \rho_0(g, \lambda) - R(g, \lambda|n_g); \quad \text{all } g \in G^0 . \tag{1.44}$$

By [12] and corollary 1.13

$$g_0E_0n_g+g_1E_1n_g \leq R(g, \lambda|n_g), \qquad g \in G^0,$$

and by (1.36)

$$\rho^*(g, \lambda) \leq 1+\rho_0(g, \lambda) < 1+\lambda_1, \qquad g \in G^0,$$

so that by (1.43)

$$g_0E_0n_g+g_1E_1n_g < 1+\lambda_1+\epsilon/2, \qquad g \in G^0.$$

Hence

$$E_0n_g+E_1n_g < (1+\lambda_1+\epsilon/2)/g_0g_1, \qquad g \in G^0.$$

If we now restrict our consideration to just those points $g \in G^0$, for which

(1.45) $$g_1^*/2 \leq g_1 \leq (1+g_1^*)/2,$$

we then have for all such g, that

(1.46) $$\lambda_0+\lambda_1+E_0n_g+E_1n_g < \lambda_0+\lambda_1+4(1+\lambda_1+\epsilon/2)/g_0^*g_1^* = K_\epsilon, \text{ say.}$$

Let

$$\delta_\epsilon = \tfrac{1}{2}\min[\epsilon/2K_\epsilon, g_0^*, g_1^*, |g_1^*-\lambda_0/(\lambda_0+\lambda_1)|] .$$

Note that with this definition of δ_ϵ, the neighborhood $N(g^*)$ given by (1.41) is an interval subset of G^0 which does <u>not</u> contain the point $(\lambda_1, \lambda_0)/(\lambda_0+\lambda_1)$. In addition, each of its points satisfies (1.45) and hence (1.46). Finally, if g', g" are arbitrary points in $N(g^*)$

(1.47) $$|g_1'-g_1''| \leq 2\delta_\epsilon \leq \epsilon/2K_\epsilon .$$

Since (1.39) must hold for g', g", we have by (1.40), taking $n=n_{g'}$, that

$$\rho_0(g', \lambda)-R(g', \lambda|n_{g'})-[\rho_0(g'', \lambda)-R(g'', \lambda|n_{g'})] \leq (\lambda_0+\lambda_1+E_0n_{g'}+E_1n_{g'})|g_1'-g_1''| .$$

By (1.46) and (1.47), the right hand side of the above inequality is $< \epsilon/2$.

Hence

$$\rho_0(g', \lambda) - R(g', \lambda | n_{g'}) < \rho_0(g'', \lambda) - R(g'', \lambda | n_{g'}) + \epsilon/2 .$$

By (1.44) and (1.38), we then have that

$$\hat{\gamma}(g', \lambda) < \hat{\gamma}(g'', \lambda) + \epsilon .$$

But g', g" were arbitrarily chosen from $N(g^*)$. Hence the above inequality with g' and g" interchanged must also hold. But then the two inequalities taken together yield (1.42) which is the desired result.

Continuity at $(\lambda_1, \lambda_0)/(\lambda_0+\lambda_1)$ may be proved by showing one sided continuity for each side separately using devices strictly analogous to those used above. This completes the proof.

An alternative proof of lemma 1.6 is to be found in E. L. Lehman's book, "Testing Statistical Hypotheses", page 105, as part of a proof for the Optimality Theorem which is given there. This proof makes use of the proposition that <u>a function defined, concave, and bounded below on an open interval is continuous there</u>. A more general statement and proof of the above proposition, put in terms of convex functions bounded above, is to be found in Hardy, Littlewood, and Polya's book, "Inequalities", proposition III, section 3.18, page 91.

<u>Lemma 1.6</u> (alternative proof).

Let $\lambda \in \Lambda$ be arbitrary, fixed. Then $\hat{\gamma}(\cdot, \lambda)$ is continuous on G (one sidedly at the endpoints of G).

<u>Proof</u>

Proof of one sided continuity at the endpoints of G is trivial. (See previous proof.) We show that $\rho^*(\cdot, \lambda)$ is continuous on G^0. Since $\rho_0(\cdot, \lambda)$ is continuous there, this will suffice for the result. Clearly, by corollary 1.14 and [22]

(1.48) $\rho^*(g, \lambda) \geqq 0, \quad \text{all } g \in G^0.$

To show that $\rho^*(\cdot, \lambda)$ is concave on G^0, let h, g', g" be arbitrary, fixed points in G^0 and write

$$h_i g = (h_i g_0, h_i g_1), \quad i=0,1,$$

for any fixed g. We need only show that

(1.49) $$\rho^*(h_0 g' + h_1 g'', \lambda) \geqq h_0 \rho^*(g', \lambda) + h_1 \rho^*(g'', \lambda).$$

The left hand side of this inequality is by [15] and the definitions on the bottom of page 3

$$\inf_{(n, \varphi_n) \in \mathscr{S}_0} R(h_0 g' + h_1 g'', \lambda | n, \varphi_n) .$$

By [12] and (1.48), this may be written

$$\inf_{(n, \varphi_n) \in \mathscr{S}_0} [h_0 R(g', \lambda | n, \varphi_n) + h_1 R(g'', \lambda | n, \varphi_n)] .$$

But this infimim is bounded below by the right hand side of (1.49). In view of the proposition to which reference has already been made, this proves the lemma.

Let

[31] $$b: \Lambda \rightarrow G, \qquad a: \Lambda \rightarrow G$$

denote mappings defined as follows. Take

$$b_1(\lambda) = \sup\{g_1 : 0 \leqq g_1 \leqq \lambda_0/(\lambda_0+\lambda_1), \hat{\gamma}(g, \lambda) \leqq 0\} ,$$

$$a_1(\lambda) = \inf\{g_1 : \lambda_0/(\lambda_0+\lambda_1) \leqq g_1 \leqq 1, \hat{\gamma}(g, \lambda) \leqq 0\} ,$$

$$b_0(\lambda) = 1 - b_1(\lambda), \qquad a_0(\lambda) = 1 - a_1(\lambda) ,$$

$$b(\lambda) = (b_0(\lambda), b_1(\lambda)), \qquad a(\lambda) = (a_0(\lambda), a_1(\lambda)) .$$

Concerning these mappings, we are now in a position to state the following theorem.

Theorem A

Let $\lambda \in \Lambda$ be arbitrarily given. Then

$$0 < b_1(\lambda) \leq \lambda_0/(\lambda_0+\lambda_1) \leq a_1(\lambda) < 1 \ , \tag{1.50}$$

$$\hat{\gamma}\left(\frac{(\lambda_1, \lambda_0)}{\lambda_0+\lambda_1}, \lambda\right) \begin{cases} \leq 0 \implies \begin{cases} \hat{\gamma}(g, \lambda) \leq 0, & \text{all } g \in G \\ b(\lambda) = (\lambda_1, \lambda_0)/(\lambda_0+\lambda_1) = a(\lambda) \end{cases} \\ \geq 0 \implies \hat{\gamma}(b(\lambda), \lambda) = 0 = \hat{\gamma}(a(\lambda), \lambda) \\ > 0 \implies 0 < b_1(\lambda) < \lambda_0/(\lambda_0+\lambda_1) < a_1(\lambda) < 1 \end{cases} \tag{1.51}$$

Proof

The proof follows immediately from the lemmas which precede.

For each $g \in G^0$ and each $\lambda \in \Lambda$, let

[32] $$B(g, \lambda) = \frac{g_0 b_1(\lambda)}{g_1 b_0(\lambda)} \ , \qquad A(g, \lambda) = \frac{g_0 a_1(\lambda)}{g_1 a_0(\lambda)} \ .$$

Corollary A1

The following three statements are equivalent

(i) $\hat{\gamma}(g \circ f^{(j)}(x), \lambda) > 0.$

(ii) $b_1(\lambda) < (g \circ f^{(j)}(x))_1 < a_1(\lambda).$

(iii) $B(g, \lambda) < f_{1j}(x)/f_{0j}(x) < A(g, \lambda).$

If we take $j=0$ in the above corollary, we get

Corollary A2

The following three statements are equivalent.

(i) $\hat{\gamma}(g, \lambda) > 0$.

(ii) $b_1(\lambda) < g_1 < a_1(\lambda)$.

(iii) $B(g, \lambda) < 1 < A(g, \lambda)$.

Corollary A3

$$\lambda \in \Lambda \ , \ \frac{\lambda_0 \lambda_1}{\lambda_0 + \lambda_1} \leq 1 \implies \hat{\gamma}(g, \lambda) \leq 0, \quad \text{all } g \in G.$$

Proof

This follows from (1.51) and (1.37). Otherwise, we may observe that

$$\rho_0(g, \lambda) \leq \rho_0\left(\frac{(\lambda_1, \lambda_0)}{\lambda_0 + \lambda_1}, \lambda\right) = \frac{\lambda_0 \lambda_1}{\lambda_0 + \lambda_1}, \quad g \in G, \lambda \in \Lambda \ .$$

The desired result now follows from (1.37).

2. Bayes g, λ Rules

We define below a family of s.s.f.'s

[33] $$n^*(\cdot | g, \lambda)$$

indexed by points in $G \times \Lambda$. First, we define

(2.1) $$n^*(x | g, \lambda) \equiv 0, \quad \text{whenever } \hat{\gamma}(g, \lambda) \leq 0.$$

For g, λ such that

(2.2) $$\hat{\gamma}(g, \lambda) > 0,$$

we take

$$(2.3)\quad n^*(x|g,\lambda) = \begin{cases} j, & \hat{\gamma}(g\circ f^{(k)}(x),\lambda) > 0,\ k=1,2,\ldots,j-1, \\ & \hat{\gamma}(g\circ f^{(j)}(x),\lambda) \leq 0,\ j=1,2,\ldots. \\ \infty, & \hat{\gamma}(g\circ f^{(k)}(x),\lambda) > 0,\ k=1,2,\ldots. \end{cases}$$

Observe that for arbitrary g, λ which satisfy (2.2), the complement of the set on which the above definition has meaning is a subset of

$$(2.4)\qquad \{x \in \mathfrak{X} : f_{0j}(x)+f_{1j}(x) = 0, \text{ for some positive integer } j\},$$

which is null under both P_0 and P_1. We shall assume some arbitrary definition of $n^*(\cdot|g,\lambda)$ on this complement consistent with the fact that it must be a s.s.f. . We remark that each member of the above family does in fact satisfy the conditions of definition [7].

By corollary A3 and (2,1), we observe that

$$(2.5)\qquad \lambda \in \Lambda,\ \frac{\lambda_0\lambda_1}{\lambda_0+\lambda_1} \leq 1 \implies n^*(x|g,\lambda) \equiv 0, \quad \text{all } g \in G.$$

In addition, we note for later reference the fact that

$$(2.6)\qquad n^*(x|g,\lambda) < \infty \implies \hat{\gamma}(g\circ f^{n^*(x|g,\lambda)},\lambda) \leq 0.$$

Now define

[34] $\varphi^*(\cdot|g,\lambda)$

to be for each $g \in G$, $\lambda \in \Lambda$, a t.d.f. for $n^*(\cdot|g,\lambda)$ which satisfies

$$(2.7)\qquad \varphi^*(x|g,\lambda) = \begin{cases} 1, & r_1(x|n^*(x|g,\lambda)) > \dfrac{g_0\lambda_0}{g_0\lambda_0+g_1\lambda_1} \\ 0, & \qquad\qquad < \end{cases}.$$

By (2.1), [27], and [18], we then have that

$$\hat{\gamma}(g,\lambda) \leq 0 \implies \begin{cases} \varphi^*(x|g,\lambda) \equiv 1, & g_1 > \dfrac{\lambda_0}{\lambda_0+\lambda_1} \\ \varphi^*(x|g,\lambda) \equiv 0, & \qquad < \end{cases}$$

To complete the definition for those g, λ such that $n^*(x|g, \lambda) \equiv 0$, we define (as a matter of later convenience)

$$\varphi^*(x|\frac{(\lambda_1, \lambda_0)}{\lambda_0+\lambda_1}, \lambda) \equiv 0, \quad \text{whenever} \quad \hat{\gamma}(\frac{(\lambda_1, \lambda_0)}{\lambda_0+\lambda_1}, \lambda) \leq 0. \tag{2.8}$$

For g, λ which satisfy (2.2), the formula (2.7) leaves $\varphi^*(x|g, \lambda)$ undefined for some x in the set (2.4) and for all x such that $n^*(x|g, \lambda) = \infty$. We shall assume some arbitrary definition of $\varphi^*(x|g, \lambda)$ for these x's consistent with the fact that $\varphi^*(\cdot|g, \lambda)$ must be a t.d.f. for $n^*(\cdot|g, \lambda)$ and will show later that the set of all such x's is null under both P_0 and P_1 for each pair g, λ which satisfies (2.2).

We show now that

$$\left.\begin{array}{l}\hat{\gamma}(g, \lambda) > 0 \\ n^*(x|g, \lambda) < \infty\end{array}\right\} \Longrightarrow r_1(x|n^*(x|g, \lambda)) \ne \frac{g_0\lambda_0}{g_0\lambda_0+g_1\lambda_1}, \tag{2.9}$$

and hence that (2.7) defines $\varphi^*(\cdot|g, \lambda)$ uniquely on the complement of the above mentioned set, for each pair g, λ which satisfies (2.2). But the left hand side of (2.9) implies, by (1.51), (2.6), and corollary A1 that

$$(g \circ f^{n^*(x|g, \lambda)}(x))_1 \ne \frac{\lambda_0}{\lambda_0+\lambda_1},$$

and some slight manipulation shows this to be equivalent to the inequality on the right hand side of (2.9).

Finally, we take as a short hand notation

[35] $$S^*(g, \lambda) = (n^*(\cdot|g, \lambda), \varphi^*(\cdot|g, \lambda)).$$

The family of Sequential Probability Ratio Tests.

We define below a family of s.s.f.'s

[36] $$\hat{n}(\cdot|u, v)$$

indexed by the set of all number pairs (u, v) which satisfy the inequality

$$0 < u < v.$$

For j=1,2,...., take

$$\hat{n}(x|u, v) = j ,$$

whenever

$$u < f_{1k}(x)/f_{0k}(x) < v, \quad k=1,2,...,j-1,$$

(treat this condition as vacuous for j=1), and either

$$f_{0j}(x) = 0 ,$$

or

$$f_{0j}(x) > 0 \text{ and the inequality } u < f_{1j}(x)/f_{0j}(x) < v \text{ is violated.}$$

Take

$$\hat{n}(x|u, v) = \infty ,$$

whenever

$$u < f_{1k}(x)/f_{0k}(x) < v, \quad k=1,2,.... .$$

We remark that each member of the family [36] is now uniquely defined on $\mathfrak{X}$ and satisfies the definition [7].

Now define

[37] $$\hat{\phi}(\cdot|u, v)$$

for $0 < u < v$, by taking

$$\hat{\phi}(x|u, v) = 0 ,$$

whenever

$$f_{1\hat{n}(x|u, v)}(x)/f_{0\hat{n}(x|u, v)}(x) \leq u ,$$

and otherwise, take

$$\hat{\varphi}(x|u, v) = 1 .$$

It is easy to verify, according to [9], that $\hat{\varphi}(\cdot|u, v)$ is a t.d.f. for $\hat{n}(\cdot|u, v)$.

Now take

[38] $$\hat{S}(u, v) = (\hat{n}(\cdot|u, v), \hat{\varphi}(\cdot|u, v)) .$$

These rules will be called sequential probability ratio tests.

For notational convenience in use below, we take

[39] $$\hat{Q}_i(u, v) = Q_i(\hat{\varphi}(\cdot|u, v))$$

Lemma 2.1

For each number pair u, v such that $0 < u < v$, we have that

$$P_i\{x: \hat{n}(x|u, v) < \infty\} = 1 , \quad i=0,1 .$$

In fact, the stronger result holds that

$$E_i n(\cdot|u, v) < \infty , \quad i=0,1 .$$

Remarks on proof

Proof of the above lemma rests upon the assumption that the densities f_0 and f_1 are positive and unequal on a set of positive probability under both hypotheses, and on the assumption that $X_1, X_2, \ldots$ are independent. A readable proof is given in "A Note on Cumulative Sums" by Charles Stein, A. M. S. volume 17 (1946) pp. 498-499.

Lemma 2.2

Let $g \in G^0$, $\lambda \in \Lambda$ be arbitrary, fixed such that

$$\hat{\gamma}(g, \lambda) > 0 .$$

Then outside of a subset of $\mathfrak{X}$ which is null according to both P_0 and P_1 we have that

$$S^*(g, \lambda) = \hat{S}(B(g, \lambda), A(g, \lambda)) ,$$

where $B(g, \lambda)$, $A(g, \lambda)$ are defined by [32] and

(2.10) $$0 < B(g, \lambda) < 1 < A(g, \lambda).$$

Proof

By corollary A1, (2.3) and [36]

$$n^*(x|g, \lambda) = \hat{n}(x|B(g, \lambda))$$

for all x outside a set which is null according to both P_0 and P_1. Thus, by lemma 2.1,

$$P_i\{x: n^*(x|g, \lambda) < \infty\} = 1, \quad i=0,1.$$

But this means that the right hand side of (2.6) must hold for all x outside of a set which is null under both P_0 and P_1. Hence by corollary A1, (2.7), [37],

$$\varphi^*(x|g, \lambda) = \hat{\varphi}(x|B(g, \lambda), A(g, \lambda))$$

for all x outside such a set. The inequality (2.10) follows from corollary A2.

Lemma 2.3

Let u, v be an arbitrarily given pair of numbers such that

$$0 < u < 1 < v.$$

Then

(2.11) $$\hat{Q}_0(u, v) \leq 1/v, \qquad \hat{Q}_1(u, v) \leq u.$$

(2.12) $$0 < u_1 < u \implies \hat{Q}_1(u_1, v) \leq \hat{Q}_1(u, v), \qquad v_1 \geq v \implies \hat{Q}_0(u, v_1) \leq \hat{Q}_0(u, v).$$

(2.13) $$\hat{Q}_1(u, v) = 0 \implies \begin{cases} \hat{Q}_1(u', v') = 0, & 0 < u' < 1 < v' \\ \hat{Q}_0(u', v) = \hat{Q}_0(u, v) > 0, & 0 < u' < 1. \end{cases}$$

$$\text{(2.14)} \qquad \hat{Q}_0(u, v) = 0 \implies \begin{cases} \hat{Q}_0(u', v') = 0, & 0 < u' < 1 < v' \\ \hat{Q}_1(u, v') = \hat{Q}_1(u, v) > 0, & v' > 1 \end{cases}$$

Proof

The proofs for (2.11), (2.12) are standard and straightforward and will not be given here. We shall prove (2.13). The proof of (2.14) is strictly analogous.

We first show that

$$\text{(2.15)} \qquad \hat{Q}_1(u, v) = 0 \iff P_1\{x: f_0(x_1) > f_1(x_1) > 0\} = 0.$$

For suppose that the right hand side equality holds. Then

$$P_1\{x: 0 \leq \frac{f_1(x_1)}{f_0(x_1)} < 1\} = 0.$$

But for arbitrary u', v' such that $0 < u' < 1 < v'$ and for each positive integer j, we have that

$$P_1\{x: \hat{n}(x|u', v') = j, \ \hat{\phi}(x|u', v') = 0\}$$

$$\leq P_1\{x: 0 \leq \frac{f_{1j}(x)}{f_{0j}(x)} < u'\} \leq j \ P_1\{x: 0 \leq \frac{f_1(x_1)}{f_0(x_1)} < 1\},$$

and hence that

$$\hat{Q}_1(u', v') = \sum_{j=1}^{\infty} P_1\{x: \hat{n}(x|u', v') = j, \hat{\phi}(x|u', v') = 0\} = 0.$$

On the other hand, suppose that the right hand side equality in (2.15) does not hold, i.e., that

$$P_1\{x: f_0(x_1) > f_1(x_1) > 0\} > 0.$$

Now

$$\{x\colon 0 < f_1(x_1) < f_0(x_1)\} = \{x\colon 0 < \frac{f_1(x_1)}{f_0(x_1)} < 1\} = \sum_{j=1}^{\infty} \{x;\ u^{\frac{1}{j-1}} < \frac{f_1(x_1)}{f_0(x_1)} \leq u^{\frac{1}{j}}\},$$

where we interpret

$$u^{\frac{1}{0}} = 0.$$

Since the P_1 probability of this disjoint union is positive, P_1 must assign positive probability to at least one set in the union. That is, there must exist a positive integer N and a positive number δ such that

$$P_1\{x\colon u^{\frac{1}{N-1}} < \frac{f_1(x_1)}{f_0(x_1)} \leq u^{\frac{1}{N}}\} = \delta.$$

But then, because the sequence $X_1, X_2, \ldots$. was taken to be independent,

$$P_1\{x\colon u^{\frac{1}{N-1}} < \frac{f_1(x_k)}{f_0(x_k)} \leq u^{\frac{1}{N}},\quad k=1,2,\ldots,N\} = \delta^N$$

But

$$u^{\frac{1}{N-1}} < \frac{f_1(x_k)}{f_0(x_k)} \leq u^{\frac{1}{N}},\quad k=1,2,\ldots,N \implies u^{\frac{k}{N-1}} < \frac{f_{1k}(x)}{f_{0k}(x)} \leq u^{\frac{k}{N}},\quad k=1,2,\ldots,N$$

$$\implies u < \frac{f_{1k}(x)}{f_{0k}(x)} < v,\quad k=1,2,\ldots,N-1, \quad \text{and} \quad \frac{f_{1N}(x)}{f_{0N}(x)} \leq u.$$

Hence

$$\hat{Q}_1(u, v) \geq P_1\{x\colon \hat{n}(x|u, v) = N,\ \hat{\varphi}(x|u, v) = 0\} \geq \delta^N > 0.$$

This proves (2.15) and with it the first half of our result.

Next observe that

The proof of this is analogous to that of (2.15). Thus by (2.15), (2.16), (2.18), (2.19) and by (2.17), the proof is complete.

We shall have occasion (lemma 3.5) to use the following particularization of a more general inequality.

Lemma 2.4

Let u, v be an arbitrarily given pair of numbers such that $0 < u < 1 \leq v$. Then

$$E_0\hat{n}(\cdot|u, v) \geq \frac{(1-\hat{Q}_0(u, v)) \log u + \hat{Q}_0(u, v)(\log v + \eta)}{E_0 \log(f_{11}/f_{01})},$$

where η is a non-negative constant which is independent of u and v.

Remarks on proof

The above inequality is a special case of A:78, page 172 of A. Wald's "Sequential Analysis". η (denoted η_{θ_0} in that text) is a special case of A:73. The proof of the inequality is given in the above reference and will not be repeated here.

We remark that

(2.20) $$E_0 \log(f_{11}/f_{01}) < 0.$$

This follows, since clearly,

$$E_0(f_{11}/f_{01}) \leq 1,$$

and because the expected value of a non-negative random variable (which is not constant with probability 1) is less than the log of its expected value. f_{11}/f_{01} can obviously not be equal to a constant larger than 1 with P_0 probability 1. In addition it cannot be equal to 0 or to 1 with P_0 probability 1 since we have assumed that our two densities are positive and unequal on a set of positive probability under both hypotheses. Finally f_{11}/f_{01} may conceivably be equal, with P_0 probability 1, to a positive constant less than 1, but in this case (2.20) holds trivially.

For use in the proof of theorem B given below, we remark now that by (2.3), (1.26), and [26], when

$$\hat{\gamma}(g, \lambda) > 0$$

we have for any positive integer j that

$$n^*(x|g, \lambda) = j \iff \tag{2.21}$$

$$\begin{cases} k+\rho_0(g\circ f^{(k)}(x), \lambda) > \inf\limits_{n \in \mathcal{S}_k(x)} R_k(g, \lambda|n, x) \\ j+\rho_0(g\circ f^{(j)}(x), \lambda) \leq \inf\limits_{n \in \mathcal{S}_j(x)} R_j(g, \lambda|n, x) \end{cases}$$

<u>Theorem B</u>

Let $g \in G$, $\lambda \in \Lambda$ be arbitrary, fixed. Then $S^*(g, \lambda)$ is a Bayes g, λ rule in

$$\mathcal{S} = \{(n, \varphi_n): P_i\{n < \infty\} = 1,\ i=0,1\}$$

and hence it is a Bayes g, λ rule in the class of all rules.

<u>Proof</u>

Observe first that by lemmas 2.2 and 2.1,

$$S^*(g, \lambda) \in \mathcal{S}. \tag{2.22}$$

Hence by (2.7), corollary 1.12 and 1.13,

$$R(g, \lambda|S^*(g, \lambda)) = R(g, \lambda|n^*(g, \lambda)).$$

Thus, to show that $S^*(g, \lambda)$ is a Bayes rule in $\mathcal{S}$, we must by corollary 1.14, show that

$$R(g, \lambda | n^*(g, \lambda)) = \rho(g, \lambda).$$

Suppose that

$$\rho_0(g, \lambda) \leq \rho^*(g, \lambda),$$

i.e., that $\hat{\gamma}(g, \lambda) \leq 0$. By (1.2)

$$\rho(g, \lambda) = \rho_0(g, \lambda).$$

But then also by (2.1),

$$n^*(x|g, \lambda) \equiv 0$$

so that by [22]

$$R(g, \lambda | n^*(g, \lambda)) = \rho_0(g, \lambda).$$

It follows from the above argument that the theorem is established for all g, λ such that $\hat{\gamma}(g, \lambda) \leq 0$.

Now suppose that

$$\rho_0(g, \lambda) > \rho^*(g, \lambda),$$

i.e., that $\hat{\gamma}(g, \lambda) > 0$. Then by (1.2)

(2.23) $$\rho(g, \lambda) = \rho^*(g, \lambda)$$

By (2.22),

$$P_g\{x: n^*(x|g, \lambda) = \infty\} = 0.$$

Thus, using

$$n^* = n^*(\cdot|g, \lambda)$$

as a notational convenience which we shall continue for the remainder of this proof, we may write

(2.24) $$R(g, \lambda|n^*) = \lim_{j \to \infty} \sum_{k=1}^{j} \int_{\{n^*=k\}} [k+\rho_0(g\circ f^{(k)}, \lambda)]dP_g .$$

We shall now suppose that $S^*(g, \lambda)$ is not a Bayes g, λ rule in $\mathcal{S}$ and show that this leads to a contradiction. If $S^*(g, \lambda)$ is not a Bayes g, λ rule in $\mathcal{S}$, there must in view of (2.23), exist a positive s.s.f. n_1, say, in $\mathcal{S}$, such that

(2.25) $$\rho^*(g, \lambda) \leq R(g, \lambda|n_1) < R(g, \lambda|n^*).$$

By lemma 1.2, we have

(2.26) $$R(g, \lambda|n_1) = E_{P_g} R_1(g, \lambda|n_1, \cdot) \geq \int_{\{n^*=1\}} R_1(g, \lambda|n_1, \cdot)dP_g .$$

Now

$$\{n^*=1\} = \{n^*=1, n_1=1\} + \{n^*=1, n_1 > 1\}.$$

But by [24],

(2.27) $$n_1(x) = 1 \implies R_1(g, \lambda|n_1, x) = 1+\rho_0(g\circ f^{(1)}(x), \lambda).$$

By (2.21)

$$n^*(x) = 1 \implies \inf_{n \in \mathcal{S}_1(x)} R_1(g, \lambda|n, x) \geq 1+\rho_0(g\circ f^{(1)}(x), \lambda)$$

In addition, it is clear that

$$n_1(x) > 1 \implies R_1(g, \lambda|n_1, x) \geq \inf_{n \in \mathcal{S}_1(x)} R_1(g, \lambda|n, x).$$

Hence

(2.28) $$n^*(x) = 1, n_1(x) > 1 \implies R_1(g, \lambda|n_1, x) \geq 1+\rho_0(g\circ f^{(1)}(x), \lambda).$$

By (2.25), (2.26), (2.27), (2.28), it follows that

(2.29) $$\int_{\{n^*=1\}} [1+\rho_0(g\circ f^{(1)}, \lambda)]dP_g \leq R(g, \lambda|n_1) < R(g, \lambda|n^*)$$

Let

$$K = \{x: n^*(x) > 1, n_1(x) = 1\}$$

and write

$$(2.30) \qquad R(g, \lambda|n_1) = \left(\int_K + \int_{\mathfrak{X} - K} \right) R_1(g, \lambda|n_1, \cdot) dP_g.$$

By (2.21)

$$n^*(x) > 1 \implies \inf_{n \in \mathcal{S}_1(x)} R_1(g, \lambda|n, x) < 1+\rho_0(g \circ f^{(1)}(x), \lambda)$$

But then by (2.27),

$$(2.31) \qquad x \in K \implies \inf_{n \in \mathcal{S}_1(x)} R_1(g, \lambda|n, x) < R_1(g, \lambda|n_1, x)$$

(2.31) implies that there exists a s.s.f. n_2, say, such that

$$n_2 \in \bigcap_{x \in K} \mathcal{S}_1(x)$$

and such that

$$(2.32) \qquad R_1(g, \lambda|n_2, x) < R(g, \lambda|n_1, x), \quad \text{all } x \in K$$

and which in addition may be arbitrarily defined (consistent with its being a s.s.f.) on $\mathfrak{X}$ -K. In particular we shall take

$$n_2(x) = n_1(x), \quad \text{all } x \in \mathfrak{X} - K,$$

so that we have

$$(2.33) \qquad R_1(g, \lambda|n_2, x) = R_1(g, \lambda|n_1, x), \quad \text{all } x \in \mathfrak{X}-K.$$

Thus by (2.29), (2.30), (2.31), (2.32), (2.33), and lemma 1.2,

$$(2.34) \qquad R(g, \lambda|n_2) \leq R(g, \lambda|n_1) < R(g, \lambda|n^*).$$

To obtain a lower bound for $R(g, \lambda|n_2)$ analogous to that obtained for $R(g, \lambda|n_1)$ in (2.29), we proceed as follows. By lemma 1.2

$$R(g, \lambda|n_2) = E_{P_g} R_2(g, \lambda|n_2, \cdot) \geqq \sum_{k=1}^{2} \int_{\{n^*=k\}} R_2(g, \lambda|n_2, \cdot)dP_g. \tag{2.35}$$

Now we may interpret $\{n^*=1\}$ as a cylinder subset of $\mathfrak{X}$ having either a one-dimensional Borel set as base or a two-dimensional Borel set as base. Hence by lemma 1.2,

$$\int_{\{n^*=1\}} R_2(g, \lambda|n_2, \cdot)dP_g = \int_{\{n^*=1\}} R_1(g, \lambda|n_2, \cdot)dP_g$$

Now $\{n^*=1\} \subset \mathfrak{X}-K$ and hence by (2.33), (2.28), (2.27)

$$n^*(x) = 1 \implies R_1(g, \lambda|n_2, x) \geqq 1+\rho_0(g\circ f^{(1)}(x), \lambda). \tag{2.36}$$

In addition, since

$$n_2(x) \geqq 2, \quad \text{all } x \in \mathfrak{X},$$

we may write

$$\{n^*=2\} = \{n^*=2, n_2=2\} + \{n^*=2, n_2 > 2\}.$$

By [24]

$$n_2(x) = 2 \implies R_2(g, \lambda|n_2, x) = 2+\rho_0(g\circ f^{(2)}(x), \lambda). \tag{2.37}$$

Also by (2.21),

$$n^*(x) = 2 \implies \inf_{n \in \mathcal{S}_2(x)} R_2(g, \lambda|n, x) \geqq 2+\rho_0(g\circ f^{(2)}(x), \lambda),$$

so that

$$n^*(x) = 2, n_2(x) > 2 \implies R_2(g, \lambda|n_2, x) \geqq 2+\rho_0(g\circ f^{(2)}(x), \lambda). \tag{2.38}$$

Thus, by (2.34), (2.35), (2.36), (2.37), and (2.38), we have

$$\sum_{k=1}^{2} \int_{\{n^*=k\}} [k+\rho_0(g\circ f^{(k)}, \lambda)]dP_g \leq R(g, \lambda|n_2) \leq R(g, \lambda|n_1) < R(g, \lambda|n^*)$$

We may now, proceeding in the same manner, establish by induction the existence of a sequence $n_1, n_2, n_3, \ldots$ of s.s.f.'s (the first two members as given above) such that for each positive integer j

$$\sum_{k=1}^{j} \int_{\{n^*=k\}} [k+\rho_0(g\circ f^{(k)}, \lambda)]dP_g \leq R(g, \lambda|n_j) \leq R(g, \lambda|n_1) < R(g, \lambda|n^*).$$

But then, taking the limit of the sum at the left as $j \to \infty$ we arrive, by (2.24) at a contradiction.

Lemma 2.5

$$\rho^*(g, \lambda) \equiv 1 + E_{P_g} \rho(g\circ f^{(1)}, \lambda), \qquad g \in G, \lambda \in \Lambda .$$

Proof

Consider first, for each $x \in \mathfrak{X}$, the mapping

$$M_x: \{n: n > 0\} \to \{n: n \geq 0\}$$

defined by

$$M_x(n) = \bar{n}(\cdot|n, x), \qquad n > 0, x \in \mathfrak{X},$$

where (see definition [23])

(2.39) $$\bar{n}(v|n, x) = n(x_1 v) - 1, \qquad \text{all } v \in \mathfrak{X},$$

and observe that for each $x \in \mathfrak{X}$, M_x is an onto map. i.e.,

(2.40) $$\{\bar{n}(\cdot|n, x): n > 0\} = \{n: n \geq 0\}, \qquad \text{for each } x \in \mathfrak{X}.$$

Let $g \in G$, $\lambda \in \Lambda$ be arbitrary, fixed. By definition [24] and (2.39), we have for each $x \in \mathfrak{X}$ and each positive s.s.f. n such that $P_i\{n < \infty\} = 1$, i=0,1, that

$$(2.41) \qquad R_1(g, \lambda|n, x) = 1 + R(g \circ f^{(1)}(x), \lambda|\bar{n}(\cdot|n, x))$$

By (2.40), this means that

$$(2.42) \qquad \inf_{n>0} R_1(g, \lambda|n, x) = 1 + \rho(g \circ f^{(1)}(x), \lambda), \qquad \text{all } x \in \mathfrak{X},$$

where $\inf_{n>0}$ is taken to denote infimum overall positive s.s.f.'s n such that $P_i\{n < \infty\} = 1$, i=0,1. It then follows that

$$(2.43) \qquad E_{P_g} \inf_{n>0} R_1(g, \lambda|n, x) = 1 + E_{P_g} \rho(g \circ f^{(1)}, \lambda).$$

Let n_1 be an <u>arbitrarily</u> given positive s.s.f. . By lemma 1.2,

$$R(g, \lambda|n_1) = E_{P_g} R_1(g, \lambda|n_1, \cdot) \geq E_{P_g} \inf_{n>0} R_1(g, \lambda|n, \cdot).$$

Hence

$$(2.44) \qquad \rho^*(g, \lambda) \geq E_{P_g} \inf_{n>0} R_1(g, \lambda|n, \cdot).$$

On the other hand, let $\epsilon > 0$ be an arbitrarily given positive number. We can find corresponding to each real number x_1, a s.s.f. $n_{x_1,\epsilon}$, say, such that

$$(2.45) \qquad R(g \circ f^{(1)}(x), \lambda|n_{x_1,\epsilon}) < \rho(g \circ f^{(1)}(x), \lambda) + \epsilon, \qquad \text{all } x \in \mathfrak{X}.$$

Now define the <u>positive</u> s.s.f. $n^{(\epsilon)}$, by

$$n^{(\epsilon)}(x_1 v) = 1 + n_{x_1,\epsilon}(v), \qquad \text{all real } x_1, \; v \in \mathfrak{X}.$$

By (2.39)

$$n_{x_1,\epsilon} = \bar{n}(\cdot|n^{(\epsilon)}, x).$$

Hence, using (2.41), (2.45), and (2.42), in that order, we have that

$$R_1(g, \lambda|n^{(\epsilon)}, x) = 1 + R(g\circ f^{(1)}(x), \lambda|n_{x_1,\epsilon})$$

$$< 1 + \rho(g\circ f^{(1)}(x), \lambda) + \epsilon = \inf_{n > 0} R_1(g, \lambda|n, x) + \epsilon.$$

But then integrating the extreme left and right hand sides above over $\mathfrak{X}$ with respect to P_g, we have by lemma 1.2,

$$R(g, \lambda|n^{(\epsilon)}) < E_{P_g} \inf_{n > 0} R_1(g, \lambda|n, \cdot) + \epsilon$$

Since

$$\rho^*(g, \lambda) \leq R(g, \lambda|n^{(\epsilon)})$$

and $\epsilon > 0$ is arbitrary we have

$$\rho^*(g, \lambda) \leq E_{P_g} \inf_{n > 0} R_1(g, \lambda|n, \cdot).$$

But this together with (2.44) and (2.43) proves the lemma.

We now define a family of positive s.s.f.'s

[40] $$\tilde{n}(\cdot|g, \lambda)$$

indexed by points in $G \times \Lambda$, by taking

$$\tilde{n}(x_1 v|g, \lambda) = 1 + n^*(v|g\circ f^{(1)}(x), \lambda) \tag{2.46}$$

for all real x_1 and each $v \in \mathfrak{X}$. Note that the right hand side above depends on x only through its first coordinate x_1 so that this is in fact a valid definition.

If we use the definition [33], it is immediately seen that $\tilde{n}(x|g, \lambda)$ is defined by the right hand side of (2.3) for all $x \in \mathfrak{X}$ and for all $g \in G$, $\lambda \in \Lambda$ without restriction. Thus, in particular, we have that

(2.47) $$\tilde{n}(x|g, \lambda) \equiv n^*(x|g, \lambda), \quad \text{all } x \in \mathfrak{X}, \text{ whenever } \hat{\gamma}(g, \lambda) > 0,$$

and by (1.51)

$$\hat{\gamma}\left(\frac{(\lambda_1, \lambda_0)}{\lambda_0+\lambda_1}, \lambda\right) \leq 0 \implies \tilde{n}(x|g, \lambda) \equiv 1, \quad \text{all } x \in \mathfrak{X}, g \in G.$$

Define

[41] $$\tilde{\varphi}(\cdot|g, \lambda)$$

for each $g \in G$, $\lambda \in \Lambda$ by

$$\tilde{\varphi}(x_1 v|g, \lambda) = \varphi^*(v|g \circ f^{(1)}(x), \lambda)$$

for all real x_1 and each $v \in \mathfrak{X}$. It is easily shown, using the definition [34] that $\tilde{\varphi}(\cdot|g, \lambda)$ is for each $g \in G$, $\lambda \in \Lambda$, a t.d.f. for $\tilde{n}(\cdot|g, \lambda)$ which is optimal with respect to g, λ in the sense of corollary 1.12. Moreover

(2.48) $$\tilde{\varphi}(x|g, \lambda) \equiv \varphi^*(x|g, \lambda), \quad \text{all } x \in \mathfrak{X}, \text{ whenever } \hat{\gamma}(g, \lambda) > 0.$$

Let

[42] $$\tilde{S}(g, \lambda) = (\tilde{n}(\cdot|g, \lambda), \tilde{\varphi}(\cdot|g, \lambda)), \quad g \in G, \ \lambda \in \Lambda.$$

Lemma 2.6

Let $g \in G^0$, $\lambda \in \Lambda$ be arbitrary, fixed, the latter such that

(2.49) $$\hat{\gamma}\left(\frac{(\lambda_1, \lambda_0)}{\lambda_0+\lambda_1}, \lambda\right) > 0.$$

Then outside of a subset of $\mathfrak{X}$ which is null according to both P_0 and P_1

$$\tilde{S}(g, \lambda) = \hat{S}(B(g, \lambda), A(g, \lambda)),$$

where $B(g, \lambda)$, $A(g, \lambda)$ are defined by [32] and

$$0 < B(g, \lambda) < A(g, \lambda). \quad (2.50)$$

<u>Proof</u>

By (2.47), (2.48),

$$\tilde{S}(g, \lambda) = S^*(g, \lambda), \text{ whenever } \hat{\gamma}(g, \lambda) > 0.$$

Hence for such g, λ the proof is immediate, following from lemma 2.2. On the other hand, in view of the remarks which follow definitions [40] and [41], the proof for <u>any</u> $g \in G^0$ and $\lambda \in \Lambda$ satisfying (2.49) is strictly analogous to that for lemma 2.2. The inequality (2.50) follows from the last implication of (1.51), the definitions [32] and the fact that $g \in G^0$.

<u>Lemma 2.7</u>

$$R(g, \lambda|\tilde{n}(\cdot|g, \lambda)) = \rho^*(g, \lambda), \quad \text{for all } g \in G, \lambda \in \Lambda .$$

Hence we may write

$$\hat{\gamma}(g, \lambda) = \rho_0(g, \lambda) - R(g, \lambda|\tilde{n}(\cdot|g, \lambda)), \quad \text{for all } g \in G, \lambda \in \Lambda .$$

<u>Proof</u>

Let $g \in G$, $\lambda \in \Lambda$ be arbitrary, fixed. If we use the notation (2.39) of lemma 2.5, we have by (2.46) that for each $x \in \mathfrak{X}$ and each $v \in \mathfrak{X}$,

$$\bar{n}(v|\tilde{n}(\cdot|g, \lambda), x) = n^*(v|g\circ f^{(1)}(x), \lambda)$$

and hence by (2.41) that

$$R_1(g, \lambda|\tilde{n}(\cdot|g, \lambda), x) = 1 + R(g\circ f^{(1)}(x), \lambda|n^*(\cdot|g\circ f^{(1)}(x), \lambda)).$$

By theorem B, we then have that

$$R_1(g, \lambda|\tilde{n}(\cdot|g, \lambda), x) = 1 + \rho(g\circ f^{(1)}(x), \lambda).$$

If we now integrate both sides above over $\mathfrak{X}$ with respect to P_g and use lemmas 1.2 and 2.5, we get the desired result.

Corollary 2.7

$\hat{\gamma}(g, \cdot)$ is uniformly continuous on Λ, uniformly for all $g \in G$.

Proof

Let ϵ be an arbitrarily given positive number. Let λ', λ'' be arbitrary points in Λ such that

(2.51) $$|\lambda_0'-\lambda_0''| < \frac{1}{3}\epsilon , \quad |\lambda_1'-\lambda_1''| < \frac{1}{3}\epsilon .$$

Then for an arbitrary $g \in G$ we have by (1.28), corollary 1.13, and the above lemma, that

$$\hat{\gamma}(g, \lambda') < \rho_0(g, \lambda'') - R(g, \lambda''|\tilde{n}(\cdot|g, \lambda')) + \epsilon.$$

But by (1.38)

$$\rho_0(g, \lambda'') - R(g, \lambda''|\tilde{n}(\cdot|g, \lambda')) \leq \hat{\gamma}(g, \lambda'').$$

Hence

$$\hat{\gamma}(g, \lambda') < \hat{\gamma}(g, \lambda'') + \epsilon.$$

But λ', λ'' were arbitrarily chosen to satisfy (2.51). Hence the above inequality holds also with λ', λ'' interchanged. This yields the desired result.

Theorem C

Let $\lambda \in \Lambda$ be arbitrary, fixed.

(i) If

(2.52) $$0 \leq g_1 < b_1(\lambda) \quad \text{or} \quad a_1(\lambda) < g_1 \leq 1,$$

then

$$\hat{\gamma}(g, \lambda) < 0.$$

(ii) If

(2.53) $$\hat{\gamma}\left(\frac{(\lambda_1, \lambda_0)}{\lambda_0+\lambda_1}, \lambda\right) > 0,$$

then $\hat{\gamma}(\cdot, \lambda)$ is non-negative and strictly monotone on the interval

(2.54) $$\{g \in G: b_1(\lambda) \leqq g_1 \leqq a_1(\lambda)\}$$

to either side of a maximum at the interior point $(\lambda_1, \lambda_0)/(\lambda_0+\lambda_1)$.

Proof

Note that to prove part (ii), we need, in view of theorem A and lemma 1.5, prove only the strict monotonicity.

We will prove below that if g', g" are any two distinct points in G such that

(2.55) $$\hat{\gamma}(g', \lambda) \geqq 0$$

and either

(2.56) $$\lambda_0/(\lambda_0+\lambda_1) \leqq g_1'' < g_1' \quad \text{or} \quad g_1' < g_1'' \leqq \lambda_0/(\lambda_0+\lambda_1),$$

then

(2.57) $$\hat{\gamma}(g'', \lambda) > \hat{\gamma}(g', \lambda).$$

But then both parts of the theorem will follow. For by corollary A2, if (2.52) holds, then

$$\hat{\gamma}(g, \lambda) \leqq 0.$$

Now suppose that (2.52) holds and $\hat{\gamma}(g, \lambda) = 0$. Choose g'=g so that (2.55) is satisfied and take $g''=b(\lambda)$ or $a(\lambda)$, according as $g_1 < b_1(\lambda)$ or $g_1 > a_1(\lambda)$, so that (2.56) is satisfied. But then (2.57) holds and hence either

$$\hat{\gamma}(b(\lambda), \lambda) > 0 \quad \text{or} \quad \hat{\gamma}(a(\lambda), \lambda) > 0$$

which contradicts theorem A. On the other hand, if (2.53) holds and g', g'' are distinct points of (2.54) which satisfy (2.56), then (2.55) must hold by theorem A and corollary A2. But then (2.57) holds which is the desired result.

We now prove that (2.55) together with (2.56) implies (2.57). Suppose the first condition of (2.56) to hold. By (1.28), Corollary 1.13 and lemma 2.7, we have, taking

$$\tilde{Q}_i(g, \lambda) = \ddot{Q}_i(\tilde{\varphi}(\cdot|g, \lambda), \quad i=0,1,$$

that

$$\hat{\gamma}(g', \lambda) = g_0'[-E_0\tilde{n}(\cdot|g', \lambda) + \lambda_0(1-\tilde{Q}_0(g', \lambda))] + g_1'[-E_1\tilde{n}(\cdot|g', \lambda) - \lambda_1\tilde{Q}_1(g', \lambda)].$$

Since $\tilde{n}(\cdot|g', \lambda)$ is a positive s.s.f.

$$-E_1\tilde{n}(\cdot|g', \lambda) - \lambda_1\tilde{Q}_1(g', \lambda) < 0.$$

But then by (2.55) and the fact that this implies that $g' \in G^0$,

$$-E_0\tilde{n}(\cdot|g', \lambda) + \lambda_0(1-\tilde{Q}_0(g', \lambda)) > 0.$$

Hence again using (1.28) and since the first condition of (2.56) is being taken to hold

$$\hat{\gamma}(g', \lambda) < \rho_0(g'', \lambda) - R(g'', \lambda|\tilde{n}(\cdot|g', \lambda)).$$

But by (1.38) the right hand side above is bounded above by $\hat{\gamma}(g'', \lambda)$. Hence (2.57) holds. A strictly analogous argument yields the identical result under the second condition of (2.56). This completes the proof.

3. Invariance and Optimality Property of the Sequential Probability Ratio Test.

Lemma 3.1

The functions a and b (defined by [31]) are continuous on Λ.

Proof

We shall first suppose that $\hat{\lambda}$ is an arbitrary fixed point in Λ such that

$$\hat{\gamma}((\hat{\lambda}_1, \hat{\lambda}_0)/(\hat{\lambda}_0+\hat{\lambda}_1), \hat{\lambda}) \leq 0$$

and show that a and b are continuous at $\hat{\lambda}$. By theorem A, the above inequality implies that

$$0 < b_1(\hat{\lambda}) = \hat{\lambda}_0/(\hat{\lambda}_0+\hat{\lambda}_1) = a_1(\hat{\lambda}) < 1.$$

Let ϵ be an arbitrarily given positive number and choose g', g" to be any two fixed points in G^0 such that

$$g_1' < \hat{\lambda}_0/(\hat{\lambda}_o+\hat{\lambda}_1) < g_1'' \quad \text{and} \quad g_1''-g_1' < \epsilon.$$

By theorem C, it then follows that

$$\hat{\gamma}(g', \hat{\lambda}) < 0, \quad \hat{\gamma}(g'', \hat{\lambda}) < 0.$$

By corollary 2.7 and the continuity in λ of the ratio $\lambda_0/(\lambda_0+\lambda_1)$, there exists a positive number δ_ϵ such that whenever

$$|\lambda_0-\hat{\lambda}_0| < \delta_\epsilon , \quad |\lambda_1-\hat{\lambda}_1| < \delta_\epsilon ,$$

then

$$\hat{\gamma}(g', \lambda) < 0, \quad \hat{\gamma}(g'', \lambda) < 0 \quad \text{and} \quad g_1' < \lambda_0/(\lambda_0+\lambda_1) < g_1'' .$$

But then by theorem A we must have that

$$g_1' \leq b_1(\lambda) \leq a_1(\lambda) \leq g_1'' .$$

Thus

$$|\lambda_i-\hat{\lambda}_i| < \delta_\epsilon,\ i=0,1 \implies |a_1(\lambda)-a_1(\hat{\lambda})|,\ |b_1(\lambda)-b_1(\hat{\lambda})| < \epsilon.$$

Now suppose that $\hat{\lambda} \in \Lambda$ is arbitrary, fixed such that

$$\hat{\gamma}((\hat{\lambda}_1,\ \hat{\lambda}_0)/(\hat{\lambda}_0+\hat{\lambda}_1),\ \hat{\lambda}) > 0.$$

We shall prove that a_1 and hence a is continuous at $\hat{\lambda}$. A strictly analogous argument, not repeated, holds for continuity of b at $\hat{\lambda}$. By theorem A, the above inequality implies that

$$0 < b_1(\hat{\lambda}) < a_1(\hat{\lambda}) < 1.$$

Let ϵ be an arbitrarily given positive number and choose g', g" to be any two fixed points in G^0 such that

$$b_1(\hat{\lambda}) < g_1' < a_1(\hat{\lambda}) < g_1'' \quad \text{and} \quad g_1''-g_1' < \epsilon.$$

By corollary A2 and theorem C, it then follows that

$$\hat{\gamma}(g',\ \hat{\lambda}) > 0,\ \hat{\gamma}(g'',\ \hat{\lambda}) < 0.$$

By corollary 2.7, there then exists a positive number δ_ϵ such that whenever

$$|\lambda_i-\hat{\lambda}_i| < \delta_\epsilon,\ \ i=0,1,$$

then

$$\hat{\gamma}(g',\ \lambda) > 0,\ \hat{\gamma}(g'',\ \lambda) < 0,$$

so that by theorem A, corollary A2, and the fact that $g_1' < g_1''$,

$$b_1(\lambda) < g_1' < a_1(\lambda) < g_1''.$$

Thus

$$|\lambda_i-\hat{\lambda}_i| < \delta_\epsilon,\ \ i=0,1 \implies |a_1(\lambda)-a_1(\hat{\lambda})| < \epsilon.$$

Hereafter, as notational convenience dictates, we shall regard each 2×2 matrix $\lambda \in \Lambda$, simply as the two-vector (λ_0, λ_1) of its positive components.

Lemma 3.2

Let λ_1 be an arbitrary, fixed positive number. Then

(i) $$\lim_{\lambda_0 \to 0} a_1(\lambda_0, \lambda_1) = 0, \qquad \lim_{\lambda_0 \to \infty} a_1(\lambda_0, \lambda_1) = 1$$

(ii) $a_1(\cdot, \lambda_1)$ is strictly increasing on $(0, \infty)$.

Proof

By corollary A3 and theorem A,

$$\lambda_0 \leqq 1 \implies \hat{r}((\lambda_1, \lambda_0)/(\lambda_0+\lambda_1), \lambda) \leqq 0 \implies a_1(\lambda) = \lambda_0/(\lambda_0+\lambda_1)$$

and this yields the first limit in (i). The second limit of (i) is also immediate since by theorem A

$$\lambda_0/(\lambda_0+\lambda_1) \leqq a_1(\lambda) < 1, \quad \text{all } \lambda_0 > 0.$$

To prove (ii) we first observe that by (1.28), corollary 1.13 and lemma 2.7, we have, taking

$$\mathcal{E}_i(\lambda) = E_i\tilde{n}(\cdot|a(\lambda), \lambda), \qquad \overline{Q}_i(\lambda) = Q_i(\tilde{\varphi}(\cdot|a(\lambda), \lambda)), \quad i=0,1,$$

that

(3.1) $$\hat{r}(a(\lambda), \lambda) = a_0(\lambda)[-\mathcal{E}_0(\lambda)+\lambda_0(1-\overline{Q}_0(\lambda))] - a_1(\lambda)[\mathcal{E}_1(\lambda)+\lambda_1\overline{Q}_1(\lambda)]$$

Let λ_0', λ_0'' be arbitrary, fixed numbers such that

(3.2) $$0 < \lambda_0' < \lambda_0''$$

and let

$$\lambda' = (\lambda_0', \lambda_1), \qquad \lambda'' = (\lambda_0'', \lambda_1) .$$

First suppose that

$$(3.3)\qquad \hat{\gamma}((\lambda_1,\ \lambda_0')/(\lambda_0'+\lambda_1),\ \lambda') \leqq 0\ .$$

Now either

$$\hat{\gamma}(a(\lambda'),\ \lambda'') \leqq 0$$

or

$$(3.4)\qquad \hat{\gamma}(a(\lambda'),\ \lambda'') > 0\ .$$

If the former inequality holds, then by (3.2), (3.3) and theorem A,

$$a_1(\lambda') = \lambda_0'/(\lambda_0'+\lambda_1) \ < \ \lambda_0''/(\lambda_0''+\lambda_1) \ \leqq \ a_1(\lambda'')\ .$$

On the other hand, if (3.4) holds, then by corollary A2, we have again that

$$(3.5)\qquad a_1(\lambda') < a_1(\lambda'')\ .$$

Now suppose that

$$\hat{\gamma}((\lambda_1,\ \lambda_0')/(\lambda_0'+\lambda_1),\ \lambda') > 0\ .$$

By theorem A, this means that

$$\hat{\gamma}(a(\lambda'),\ \lambda') = 0$$

By (3.1) and the fact that

$$\mathcal{E}_i(\lambda') \geqq 1,\quad i=0,1\ ,$$

this in turn implies that

$$1-\bar{Q}_0(\lambda') \ \geqq \ 1/\lambda_0' \ > \ 0\ .$$

But by (3.2) and an already familiar argument this means that

$$0 = \hat{\gamma}(a(\lambda'),\ \lambda') < \rho_0(a(\lambda'),\ \lambda'')-R(a(\lambda'),\ \lambda''|\tilde{n}(\cdot|a(\lambda'),\ \lambda'')) \leqq \hat{\gamma}(a(\lambda'),\ \lambda'')$$

But then (3.4) holds which means that (3.5) again follows. This completes the proof.

Corollary 3.2

There exists a function

[43] $$\lambda_0^*: (0, \infty) \times (0, 1) \longrightarrow (0, \infty)$$

such that

(i) $$a_1(\lambda_0^*(\lambda_1, \delta), \lambda_1) \equiv \delta, \quad 0 < \lambda_1 < \infty, \quad 0 < \delta < 1 .$$

(ii) For arbitrary, fixed $\lambda_1 > 0$, $\lambda_0^*(\lambda_1, \cdot)$ is a strictly increasing, continuous, unbounded function on $(0, 1)$ and

$$\lim_{\delta \to 0} \lambda_0^*(\lambda_1, \delta) = 0 .$$

Lemma 3.3

Let δ be arbitrary, fixed, $0 < \delta < 1$. Then $\lambda_0^*(\cdot, \delta)$ is continuous on $(0, \infty)$.

Proof

Let $\lambda_1 > 0$ be arbitrary, fixed. Take

(3.6) $$\xi_1, \; \xi_2, \ldots.$$

to be any sequence of positive numbers such that

(3.7) $$\lim_{j \to \infty} \xi_j = \lambda_1$$

To prove continuity of $\lambda_0^*(\cdot, \delta)$ at λ_1 it suffices to show that

(3.8) $$\lim_{j \to \infty} \lambda_0^*(\xi_j, \delta) = \lambda_0^*(\lambda_1, \delta) .$$

By corollary 3.2 and theorem A, we have for each positive intejer j

(3.9) $$\frac{\lambda_0^*(\xi_j, \delta)}{\lambda_0^*(\xi_j, \delta) + \xi_j} \leq a_1(\lambda_0^*(\xi_j, \delta), \xi_j) = \delta < 1 .$$

It follows that

(3.10) $$\lambda_0^*(\xi_1, \delta), \; \lambda_0^*(\xi_2, \delta), \ldots..$$

is a bounded sequence of positive numbers. For if it were not, in view of (3.7), the left hand side of (3.9) would tend to 1 as $j \longrightarrow \infty$, which is impossible. Hence the sequence (3.10) must have at least one limit point. Let ℓ_1, ℓ_2 be limit points of (3.10). There must then exist subsequences

$$\{\xi_{s_j}\}, \{\xi_{t_j}\}, \text{ say,}$$

of (3.6) such that

$$(3.11) \qquad \lim_{j \longrightarrow \infty} \lambda_0^*(\xi_{s_j}, \delta) = \ell_1, \qquad \lim_{j \longrightarrow \infty} \lambda_0^*(\xi_{t_j}, \delta) = \ell_2 .$$

By corollary 3.2, we have for each positive integer j

$$(3.12) \qquad a_1(\lambda_0^*(\xi_{s_j}, \delta), \xi_{s_j}) = \delta = a_1(\lambda_0^*(\xi_{t_j}, \delta), \xi_{t_j}) .$$

Now both ℓ_1 and ℓ_2 must be positive numbers. For suppose, for example that $\ell_1 = 0$. Then by (3.11), corollary A3, and theorem A, there exists N such that

$$j > N \Longrightarrow 0 < \lambda_0^*(\xi_{s_j}, \delta) < 1 \Longrightarrow a_1(\lambda_0^*(\xi_{s_j}, \delta), \xi_{s_j}) = \frac{\lambda_0^*(\xi_{s_j}, \delta)}{\lambda_0^*(\xi_{s_j}, \delta) + \xi_{s_j}} .$$

But then in view of (3.7)

$$\lim_{j \longrightarrow \infty} a_1(\lambda_0^*(\xi_{s_j}, \delta), \xi_{s_j}) = 0 ,$$

and this is impossible by (3.12). Since ℓ_1 and ℓ_2 are positive, it follows from lemma 3.1, that a_1 is continuous at each of the points

$$(\ell_1, \lambda_1), \quad (\ell_2, \lambda_2) .$$

But then by (3.12) and by (3.7) and (3.11),

$$a_1(\ell_1, \lambda_1) = \delta = a_1(\ell_2, \lambda_1)$$

By lemma 3.2 and its corollary, we then have

$$\ell_1 = \ell_2 = \lambda_0^*(\lambda_1, \delta) .$$

This shows that (3.8) must hold and hence completes the proof.

Lemma 3.4

For all $\lambda \in \Lambda$ which satisfy the inequality

$$\lambda_0\lambda_1/(\lambda_0+\lambda_1) \geqq 1 ,$$

we have that

$$0 < \frac{1}{\lambda_1} \leqq b_1(\lambda) \leqq \frac{\lambda_0}{\lambda_0+\lambda_1} \leqq a_1(\lambda) \leqq \frac{\lambda_0-1}{\lambda_0} < 1 .$$

Proof

Observe first that

$$\lambda_0\lambda_1/(\lambda_0+\lambda_1) \geqq 1 \iff 1/\lambda_1 \leqq \lambda_0/(\lambda_0+\lambda_1) \leqq (\lambda_0-1)/\lambda_0 .$$

By (1.37), (1.27), we have for $g \in G$, $\lambda \in \Lambda$, that

$$\hat{\gamma}(g, \lambda) \leqq \begin{cases} \lambda_1 g_1 - 1 , & g_1 \leqq \lambda_0/(\lambda_0+\lambda_1) \\ -\lambda_0 g_1 + \lambda_0 - 1 , & \geqq \end{cases} .$$

The desired result now follows from theorem A.

Lemma 3.5

Let $\lambda_0 > 0$ be arbitrary, fixed. Then

$$\lim_{\lambda_1 \to 0} a_1(\lambda) = 1, \qquad \lim_{\lambda_1 \to \infty} a_1(\lambda) = 0.$$

Proof

If $\lambda_0 \leqq 1$, the lemma is immediate. For by corollary A3 and theorem A, we have that

(3.13) $$a_1(\lambda) = \lambda_0/(\lambda_0+\lambda_1) ,$$

for all $\lambda_1 > 0$.

Thus, suppose that $\lambda_0 > 1$. In this case, corollary A3 and theorem A imply that (3.13) holds whenever $0 < \lambda_1 \leqq 1$. But this means that again the first limit to be proved holds true.

It now remains only to prove the second limit for arbitrarily given $\lambda_0 > 1$. Let

(3.14) $$\xi_1, \xi_2, \ldots .$$

be any sequence of numbers such that

(3.15) $$\xi_j \geqq \lambda_0/(\lambda_0+\lambda_1), \quad j=1,2,\ldots, \text{ and } \lim_{j \longrightarrow \infty} \xi_j = \infty .$$

To prove that the second limit of the lemma holds, it suffices to show that

(3.16) $$\lim_{j \longrightarrow \infty} a_1(\lambda_0, \xi_j) = 0 .$$

Because $\lambda_0 > 1$, (3.15) implies that for each positive integer j,

$$\lambda_0 \xi_j/(\lambda_0+\xi_j) \geqq 1$$

and hence, by lemma 3.4, that for each positive integer j

(3.17) $$0 < b_1(\lambda_0, \xi_j) \leqq \lambda_0/(\lambda_0+\xi_j) \leqq a_1(\lambda_0, \xi_j) \leqq (\lambda_0-1)/\lambda_0 < 1.$$

It follows that the sequence

(3.18) $$a_1(\lambda_0, \xi_1), a_1(\lambda_0, \xi_2), \ldots .$$

has at least one limit point in $[0, (\lambda_0-1)/\lambda_0]$ and that every limit point of this sequence must lie in this interval.

Suppose that the above sequence has a limit point ℓ such that

$$0 < \ell \leqq (\lambda_0-1)/\lambda_0 .$$

We will show in what follows, that this supposition leads to a contradiction. But then as a consequence it will follow that (3.16) must hold and hence the lemma will be proved.

Let

$$h_1 = \tfrac{1}{2}\ell, \quad h_0 = 1-h_1 .$$

Since ℓ is a limit point of (3.18), there must exist a subsequence of (3.14), $\{\xi_{t_j}\}$, say, such that

$$(3.19) \qquad \lim_{j \to \infty} a_1(\lambda_0, \xi_{t_j}) = 2h_1.$$

Now let v be an arbitrarily given number such that

$$v > 1.$$

Define

$$q_1 = h_1/(h_0 v+h_1), \quad q_0 = 1-q_1, \quad q = (q_0, q_1) ,$$

then

$$(3.20) \qquad 0 < q_1 < h_1 \quad \text{and} \quad q_0 h_1/q_1 h_0 = v.$$

Thus, by (3.15), (3.17), (3.19), and by theorem A, there exists a positive integer N such that if j is any integer $\geq$ N,

$$(3.21) \qquad 0 < b_1(\lambda_0, \xi_{t_j}) < \lambda_0/(\lambda_0+\xi_{t_j}) < q_1 < h_1 < a_1(\lambda_0, \xi_{t_j}) \leq (\lambda_0-1)/\lambda_0 < 1.$$

If we now take

$$B_j = q_0 b_1(\lambda_0, \xi_{t_j})/q_1 b_0(\lambda_0, \xi_{t_j}), \quad A_j = q_0 a_1(\lambda_0, \xi_{t_j})/q_1 a_0(\lambda_0, \xi_{t_j}) ,$$

then by (3.20), (3.21), whenever j is an integer $\geq$ N,

$$(3.22) \qquad 0 < B_j < q_0\lambda_0/q_1\xi_{t_j} < 1 < v < A_j < q_0(\lambda_0-1)/q_1 .$$

Let

$$\lambda^{(j)} = (\lambda_0, \xi_{t_j})$$

By (3.21) and corollary A2, it is clear that for each integer $j \geq N$,

$$\hat{\gamma}(q, \lambda^{(j)}) > 0.$$

Hence by lemma 2.2 and theorem B, we have that for each integer $j \geq N$,

$$\hat{S}(B_j, A_j) = S^*(q, \lambda^{(j)})$$

is a Bayes $q, \lambda^{(j)}$ rule in the class of all rules.

Now

$$R(q, \lambda^{(j)}|\hat{S}(B_j, A_j)) \geq q_0 E_0 \hat{n}(\cdot|B_j, A_j), \quad \text{all } j \geq N. \tag{3.23}$$

In addition, for each integer $j \geq N$, we have by lemma 2.4, taking

$$\hat{Q}_{0j} = \hat{Q}_0(B_j, A_j),$$

that

$$E_0 \hat{n}(\cdot|B_j, A_j) \geq \frac{(1-\hat{Q}_{0j}) \log B_j + \hat{Q}_{0j}(\log A_j + \eta_0)}{E_0 \log(f_{11}/f_{01})}, \tag{3.24}$$

where η_0 is a non-negative constant independent of j and the denominator on the right hand side is negative. By (2.11) and (3.22), for each $j \geq N$,

$$\hat{Q}_{0j} \leq 1/A_j < 1/v, \qquad \log B_j < 0,$$

and hence

$$(1-\hat{Q}_{0j}) \log B_j < \frac{v-1}{v} \log B_j.$$

Thus, again using (3.22), we find that the numerator on the right hand side of (3.24) is bounded above by

$$[(v-1) \log(q_0\lambda_0/q_1) + \log(q_0(\lambda_0-1)/q_1) + \eta_0 - (v-1) \log \xi_{t_j}]/v.$$

By (3.23), it follows that

$$R(q, \lambda^{(j)}|\hat{S}(B_j, A_j)) \geq K - \frac{q_0(v-1)}{vE_0 \log(f_{11}/f_{01})} \log \xi_{t_j}, \qquad j \geq N,$$

where K is a constant independent of j. But then by (3.15), (2.20),

$$\lim_{j \to \infty} R(q, \lambda^{(j)} | \hat{S}(B_j, A_j)) = \infty . \tag{3.25}$$

On the other hand, if S_1 denotes the rule whose s.s.f. and t.d.f. are respectively identically 0 and identically 1, then for each positive integer j,

$$R(q, \lambda^{(j)} | S_1) = q_0 \lambda_0$$

In view of (3.25), it follows that there exists an integer $N_1 \geqq N$ such that for each integer $j \geqq N_1$,

$$R(q, \lambda^{(j)} | \hat{S}(B_j, A_j) > R(q, \lambda^{(j)} | S_1).$$

But this contradicts the conclusion reached above that for each $j \geqq N_1$, $\hat{S}(B_j, A_j)$ is a Bayes q, $\lambda^{(j)}$ rule. This completes the proof.

Corollary 3.5

Let δ be arbitrary, fixed, $0 < \delta < 1$. Then

$$\lim_{\lambda_1 \to 0} \lambda_0^*(\lambda_1, \delta) = 0, \qquad \lim_{\lambda_1 \to \infty} \lambda_0^*(\lambda_1, \delta) = \infty .$$

Lemma 3.6

Let δ be arbitrary, fixed, $0 < \delta < 1$. Then

$$\lim_{\lambda_1 \to 0} b_1(\lambda_0^*(\lambda_1, \delta), \lambda_1) = \delta, \qquad \lim_{\lambda_1 \to \infty} b_1(\lambda_0^*(\lambda_1, \delta), \lambda_1) = 0.$$

Proof

By corollary A3, theorem A, and corollary 3.2

$$0 < \lambda_1 < 1 \Longrightarrow b_1(\lambda_0^*(\lambda_1, \delta), \lambda_1) = a_1(\lambda_0^*(\lambda_1, \delta), \lambda_1) = \delta,$$

so that the first limit holds.

Let

(3.26) $\xi_1, \xi_2, \ldots\ldots$

be <u>any</u> sequence of positive numbers such that

(3.27) $\lim_{j \to \infty} \xi_j = \infty.$

To prove that the second limit holds, it suffices to show that

(3.28) $\lim_{j \to \infty} b_1(\lambda_0^*(\xi_j, \delta), \xi_j) = 0.$

By theorem A and corollary 3.2, we have for each positive integer j that

(3.29) $$0 < b_1(\lambda_0^*(\xi_j, \delta), \xi_j) \leq \frac{\lambda_0^*(\xi_j', \delta)}{\lambda_0^*(\xi_j, \delta) + \xi_j} \leq \delta < 1 .$$

It follows that the sequence

(3.30) $\{b_1(\lambda_0^*(\xi_j, \delta), \xi_j)\}$

has at least one limit point in $[0, \delta]$ and that every limit point of the sequence must lie in this interval. Suppose that it has a limit point ℓ, say such that

(3.31) $0 < \ell \leq \delta.$

We will show in what follows that this supposition leads to a contradiction. But then as a consequence, it will follow that (3.28) must hold and hence the lemma will be proved.

Let

$$\hat{b}_1 = \tfrac{1}{2}\ell , \qquad \hat{b}_0 = 1-\hat{b}_1 .$$

Since ℓ is a limit point of (3.30), there must exist a subsequence $\{\xi_{t_j}\}$, say, of (3.26) such that, taking

(3.32) $b_{1j} = b_1(\lambda_0^*(\xi_{t_j}, \delta), \xi_{t_j}), \quad j=1,2,\ldots.$, we have

$$\lim_{j \to \infty} b_{1j} = 2\hat{b}_1 . \tag{3.33}$$

Now let v be an arbitrarily given number such that

$$1 < v < 1 + \frac{1}{1-\delta} . \tag{3.34}$$

Define

$$q_1 = \frac{\delta}{\delta+(1-\delta)v} , \qquad q_0 = 1-q_1 , \qquad q = (q_0, q_1) . \tag{3.35}$$

By (3.29), (3.31), and (3.34),

$$0 < \hat{b}_1 \leq \tfrac{1}{2}\delta < q_1 < \delta, \qquad \frac{q_0\delta}{q_1(1-\delta)} = v . \tag{3.36}$$

Take

$$\lambda^{(j)} = (\lambda_0^*(\xi_{t_j}, \delta), \xi_{t_j}), \quad j=1,2,\ldots . \tag{3.37}$$

We shall have repeated occasion, below, to make use of the following remarks. Let u be an arbitrarily given number such that

$$0 < u < 1 .$$

Consider the sequential probability ratio test, $\hat{S}(u, 1/u)$. By lemma 2.1,

$$E_i\hat{n}(\cdot|u, 1/u) < \infty , \quad i=0,1 . \tag{3.38}$$

In addition, by (2.11),

$$\hat{Q}_i(u, 1/u) \leq u , \quad i=0,1 .$$

Hence for each positive integer j, we have

$$q_0\lambda_0^*(\xi_{t_j}, \delta)\, \hat{Q}_0(u, 1/u) + q_1\xi_{t_j}\, \hat{Q}_1(u, 1/u) \leq (q_0\lambda_0^*(\xi_{t_j}, \delta) + q_1\xi_{t_j})u. \tag{3.39}$$

By (3.29) and (3.36)

$$q_0\lambda_0^*(\xi_{t_j}, \delta) \leq vq_1\xi_{t_j} , \qquad j=1,2,\ldots.$$

and hence it follows that for each positive integer j

(3.40) $$q_0\lambda_0^*(\xi_{t_j}, \delta)\, \hat{Q}_0(u, 1/u) + q_1\xi_{t_j}\, \hat{Q}_1(u, 1/u) \leq (1+v)uq_1\xi_{t_j} .$$

We shall now show that there exists a positive integer N_1, say, such that

(3.41) $$b_{1j} < q_1 , \qquad j=N_1, N_1+1, \ldots .$$

For suppose this were not true. Then

$$q_1 \leq b_{1j}$$

for infinitely many positive integers j. But recalling the definitions (3.32) and (3.37), we have by [35], corollary A2, (2.1), and the remark which follows (2.7), that this fact implies that

$$S^*(q, \lambda^{(j)}) = S_0$$

for infinitely many positive integers j, where S_0 is the rule with s.s.f. and t.d.f. both identically zero. But this further implies that

(3.42) $$R(q, \lambda^{(j)} | S^*(q, \lambda^{(j)})) = q_1\xi_{t_j}$$

for infinitely many positive integers, j.

On the other hand, if we choose u to be any fixed number such that

$$0 < u < \frac{1}{2(1+v)}$$

then by (3.40), for each positive integer j,

$$q_0\lambda_0^*(\xi_{t_j}, \delta)\, \hat{Q}_0(u, 1/u) + q_1\xi_{t_j}\, \hat{Q}_1(u, 1/u) < \tfrac{1}{2}q_1\xi_{t_j} .$$

In addition, by (3.38) and (3.27), there exists an integer N_2, say, $\geq N_1$ such that for each integer $j \geq N_2$,

$$\sum_{i=0}^{1} q_i E_i \hat{n}(\cdot | u, 1/u) < \tfrac{1}{2}q_1\xi_{t_j} .$$

By (3.42) it now follows that for infinitely many positive integers j,

$$R(q, \lambda^{(j)}|\hat{S}(u, 1/u)) < R(q, \lambda^{(j)}|S*(q, \lambda^{(j)})).$$

But this contradicts theorem B and hence (3.41) must hold.

Thus by (3.41), (3.33), and (3.36), there exists an integer N_3, say, $\geq N_1$ such that for each integer $j \geq N_3$

$$0 < \hat{b}_1 < b_{1j} < q_1 < \delta < 1 \; . \tag{3.43}$$

Let

$$B = \frac{q_0 \hat{b}_1}{q_1 \hat{b}_0} , \qquad B_j = \frac{q_0 b_{1j}}{q_1 b_{0j}} , \qquad j=1,2,\ldots.$$

By corollary A2 and (3.43)

$$\hat{\gamma}(q, \lambda^{(j)}) > 0 , \qquad j \geq N_3.$$

Hence by lemma 2.3 and (3.36)

$$S*(q, \lambda^{(j)}) = \hat{S}(B_j, v) , \quad j \geq N_3. \tag{3.44}$$

By (3.43) and (3.36)

$$0 < B < B_j < 1 < v, \qquad j \geq N_3. \tag{3.45}$$

Now consider the sequential probability ratio test, $\hat{S}(B, v)$. It is clear that either

$$\hat{Q}_1(B, v) > 0 \quad \text{or} \quad \hat{Q}_1(B, v) = 0.$$

We shall develop our contradiction by demonstrating both alternatives to be impossible. It will follow that (3.31) cannot hold and the lemma will be proved.

By (3.44),

$$R(q, \lambda^{(j)}|S*(q, \lambda^{(j)})) = R(q, \lambda^{(j)}|\hat{S}(B_j, v)), \quad j \geq N_3. \tag{3.46}$$

But

$$R(q, \lambda^{(j)}|\hat{S}(B_j, v)) \geq q_1\xi_{t_j} \hat{Q}_1(B_j, v), \qquad j \geq N_3.$$

In addition, by (2.12) and (3.45)

$$\hat{Q}_1(B_j, v) \geq \hat{Q}_1(B, v), \qquad j \geq N_3.$$

Hence

$$R(q, \lambda^{(j)}|S^*(q, \lambda^{(j)})) \geq q_1\xi_{t_j} \hat{Q}_1(B, v), \qquad j \geq N_3. \tag{3.47}$$

Now suppose that

$$\hat{Q}_1(B, v) > 0, \tag{3.48}$$

and choose u to be any fixed number such that

$$0 < u < \hat{Q}_1(B, v)/2(1+v).$$

By (3.40), we then have for each positive integer j that

$$q_0\lambda_0^*(\xi_{t_j}, \delta) \hat{Q}_0(u, 1/u) + q_1\xi_{t_j} \hat{Q}_1(u, 1/u) < \tfrac{1}{2}q_1\xi_{t_j} \hat{Q}_1(B, v).$$

In addition, by (3.38), (3.27), there exists an integer $N_4 \geq N_3$ such that for each integer $j \geq N_4$

$$\sum_{i=0}^{1} q_i E_i \hat{n}(\cdot|u, 1/u) < \tfrac{1}{2}q_1\xi_{t_j} \hat{Q}_1(B, v).$$

By (3.47), it follows that

$$R(q, \lambda^{(j)}|\hat{S}(B, v)) < R(q, \lambda^{(j)}|S^*(q, \lambda^{(j)})), \qquad j \geq N_4.$$

This contradicts theorem B and hence (3.48) cannot hold.

Now suppose that

$$\hat{Q}_1(B, v) = 0. \tag{3.49}$$

By (2.13) and (3.45) this implies that

$$\hat{Q}_0(B_j, v) = \hat{Q}_0(b, v) > 0, \qquad j \geq N_3.$$

By (3.46) and the above we may write

$$R(q, \lambda^{(j)}|S^*(q, \lambda^{(j)})) \geq q_0\lambda_0^*(\xi_{t_j}, \delta)\, \hat{Q}_0(B, v), \qquad j \geq N_3. \tag{3.50}$$

By (3.32), (3.29), (3.43),

$$\xi_{t_j}/\lambda_0^*(\xi_{t_j}, \delta) \leq b_{0j}/b_{1j} < \hat{b}_0/\hat{b}_1, \qquad j \geq N_3.$$

Hence

$$1 + \frac{q_1\xi_{t_j}}{q_0\lambda_0^*(\xi_{t_j}, \delta)} < \frac{B+1}{B}, \qquad j \geq N_3. \tag{3.51}$$

Now choose u to be any fixed number such that

$$0 < u < \frac{B}{2(B+1)}\hat{Q}_0(B, v).$$

By (3.39), (3.51), we then have that

$$q_0\lambda_0^*(\xi_{t_j}, \delta)\, \hat{Q}_0(u, 1/u) + q_1\xi_{t_j}\, \hat{Q}_1(u, 1/u) < \tfrac{1}{2}q_0\lambda_0^*(\xi_{t_j}, \delta)\, \hat{Q}_0(B, v), \qquad j \geq N_3.$$

By corollary 3.5 and (3.27), there exists an integer $N_5 \geq N_3$ such that for each integer $j \geq N_5$,

$$\sum_{i=0}^{1} q_i E_i \hat{n}(\cdot|u, 1/u) < \tfrac{1}{2}q_0\lambda_0^*(\xi_{t_j}, \delta)\, \hat{Q}_0(B, v).$$

By (3.50) it now follows that

$$R(q, \lambda^{(j)}|\hat{S}(u, 1/u)) < R(q, \lambda^{(j)}|S^*(q, \lambda^{(j)})), \qquad j \geq N_5.$$

But this again contradicts theorem B, so that (3.49) cannot hold. This establishes a contradiction and the lemma is proved.

Theorem D

There exists a mapping

$$\bar{\lambda}: \{(g, u, v): g \in G^0, 0 < u < 1 < v\} \longrightarrow \Lambda$$

such that identically on its domain

$$B(g, \bar{\lambda}(g, u, v)) = u, \qquad A(g, \bar{\lambda}(g, u, v)) = v .$$

Proof

Let (g, u, v) be an arbitrary fixed point in the hypothesised domain of $\bar{\lambda}$. Take

$$\bar{\delta}(g, v) = g_1 v/(g_0+g_1 v).$$

By corollary 3.2 we then have, identically for $\lambda_1 > 0$, that

$$A(g, \lambda_0^*(\lambda_1, \bar{\delta}(g, v)), \lambda_1) = g_0\bar{\delta}(g, v)/g_1(1-\bar{\delta}(g, v)) = v . \tag{3.52}$$

Now

$$0 < \bar{\delta}(g, u) < \bar{\delta}(g, v) .$$

By lemma 3.1, b is continuous on Λ. By lemma 3.3, $\lambda_0^*(\cdot, \delta(g, v))$ is continuous on $(0, \infty)$. Hence by lemma 3.6 there exists a positive value of λ_1, call it $\bar{\lambda}_1(g, u, v)$ such that taking

$$\bar{\lambda}(g, u, v) = (\lambda_0^*(\bar{\lambda}_1(g, u, v), \bar{\delta}(g, v)), \bar{\lambda}_1(g, u, v)),$$

we get

$$b_1(\bar{\lambda}(g, u, v)) = \bar{\delta}(g, u) .$$

But then it follows that

$$B(g, \bar{\lambda}(g, u, v)) = g_0\bar{\delta}(g, u)/g_1(1-\bar{\delta}(g, u)) = u .$$

Since (3.52) is an identity for $\lambda_1 > 0$, we have in addition that

$$A(g, \bar{\lambda}(g, u, v)) = v .$$

This proves the theorem.

Optimality Theorem

Let u, v be arbitrary, fixed numbers such that

$$0 < u < 1 < v$$

and let (n, φ_n) be any rule such that

$$Q_i(\varphi_n) \leq \hat{Q}_i(u, v) , \quad i=0,1 . \tag{3.53}$$

Then

$$E_i n \geq E_i\hat{n}(\cdot|u, v), \quad i=0,1 .$$

Proof

By corollary A2, lemma 2.2, and theorem D

$$S^*(g, \bar{\lambda}(g, u, v)) \equiv \hat{S}(u, v) ,$$

where the identity holds for all $g \in G^0$ and for all u, v such that $0 < u < 1 < v$. Hence by theorem B,

$$R(g, \bar{\lambda}(g, u, v)|\hat{S}(u, v)) \leq R(g, \bar{\lambda}(g, u, v)|n, \varphi_n)$$

where this inequality holds identically over the same domain. Rewriting this inequality we have, again identically,

$$\sum_{i=0}^{1} g_i[E_i\hat{n}(\cdot|u, v) - E_i n] \leq \sum_{i=0}^{1} g_i\bar{\lambda}_i(g, u, v)[Q_i(\varphi_n)-\hat{Q}_i(u, v)] .$$

By (3.53), the right hand side is identically non-positive. It follows that for each u, v, $0 < u < 1 < v$ and for each g_1 such that $0 < g_1 < 1$,

$$E_0\hat{n}(\cdot|u, v) - E_0 n + [(E_1\hat{n}(\cdot|u, v) - E_1 n) - (E_0\hat{n}(\cdot|u, v) - E_0 n)] g_1 \leq 0 .$$

But now taking limits as $g_1 \rightarrow 0$ and $g_1 \rightarrow 1$, we achieve the desired result.

4. Alternative Proof of Invariance Property

The following is an alternative proof of theorem D in section 3 which is due to Le Cam and appears in E. L. Lehman's book, "Testing Statistical Hypotheses". (See the discussion which precedes the alternative proof of lemma 1.6 which also applies here.) The present section may be substituted for section 3 with no loss in continuity for the overall proof.

We now relate Le Cam's proof notationally to that which precedes. We first set up a one to one correspondence between points in Λ and those in the cross product $G^0 \times (0, \infty)$ as follows. Let

[44] $$c = 1/(\lambda_0+\lambda_1), \quad W_0 = \lambda_0/(\lambda_0+\lambda_1), \quad W_1 = 1-W_0, \quad W = (W_0, W_1),$$

then

$$\lambda = W/c$$

Thus to each $\lambda \in \Lambda$ there corresponds a unique $(W, c) \in G^0 \times (0, \infty)$ and precisely one such λ gives rise to this point. In the following, we shall refer interchangeably, as convenience dictates to points (W, c) and their correspondents, λ.

We define a new average risk for a rule (n, φ_n) relative to $g \in G$ and $(W, c) \in G^0 \times (0, \infty)$ by

[45] $$\overline{R}(g, W, c|n, \varphi_n) = \sum_{i=0}^{1} g_i[cE_in + W_iQ_i(\varphi_n)] .$$

By [12] this is just $R(g, \lambda|n, \varphi_n)/(\lambda_0+\lambda_1)$. Let

[46] $$\overline{\rho}(g, W, c) = \inf_{n > 0} \overline{R}(g, W, c|n, \varphi_n), \quad \overline{\rho}_0(g, W) = \inf_{n=0} \overline{R}(g, W, c|n, \varphi_n)$$

It is immediate that

$$\overline{\rho}_0(g, W) = \min(g_0W_0, g_1W_1)$$

(Note that this is independent of c), and using lemma 2.7, that

(4.1) $$\bar{\rho}(g, W, c) = \rho^*(g, \lambda)/(\lambda_0+\lambda_1) = \bar{R}(g, W, c|\tilde{S}(g, W/c)).$$

Let

[47] $$\bar{\gamma}(g, W, c) = \bar{\rho}_0(g, W, c) - \bar{\rho}(g, W, c)$$

and observe that

(4.2) $$\bar{\gamma}(g, W, c) \lesseqgtr 0 \iff \hat{\gamma}(g, \lambda) \lesseqgtr 0.$$

<u>Lemma 4.1</u>

Let $g \in G^0$, $W \in G^0$ be arbitrary, fixed. Then $\bar{\rho}(g, W, \cdot)$ is

(i) concave on $(0, \infty)$ and hence continuous there.

(ii) strictly increasing on $(0, \infty)$.

(iii) $$\lim_{c \to 0} \bar{\rho}(g, W, c) = 0.$$

<u>Proof</u>

Let $h \in G^0$, $0 < c_0 < c_1$ be arbitrary, fixed. We have

$$\bar{\rho}(g, W, h_0c_0+h_1c_1) = \inf_{n > 0} R(g, W, h_0c_0+h_1c_1|n, \varphi_n)$$

$$= \inf_{n > 0} \sum_{i=0}^{1} h_iR(g, W, c_i|n, \varphi_n)$$

$$\geq \sum_{i=0}^{1} h_i\bar{\rho}(g, W, c_i).$$

But this proves (i). Again, take c_0, c_1 to be arbitrary fixed numbers such that $0 < c_0 < c_1$. By (4.1) and definitions [45], [46], we have that

$$\bar{\rho}(g, W, c_1) > \bar{R}(g, W, c_0|\tilde{S}(g, W/c)) \geq \bar{\rho}(g, W, c_0) ,$$

and this proves (ii). Now let ϵ be an arbitrarily given positive number. Let (n, φ_n) be a fixed sample size rule with

$$n(x) \equiv N_\epsilon$$

and N_ϵ so large that

$$Q_i(\varphi_n) < \epsilon/2 , \quad i=0,1 .$$

Such a rule can always be found. Then for arbitrary $c > 0$, we have that

$$\bar{\rho}(g, W, c) \leqq \bar{R}(g, W, c|n, \varphi_n) < cN_\epsilon + \epsilon/2 .$$

It follows that

$$0 < c < \epsilon/2N_\epsilon \implies \bar{\rho}(g, W, c) < \epsilon .$$

This completes the proof of the lemma.

Lemma 4.2

To each $W = (W_0, W_1) \in G^0$ there corresponds a positive number

[48] $$\bar{c}(W)$$

such that

$$0 < c < \bar{c}(W) \iff 0 < b_1(W/c) < W_0 < a_1(W/c) < 1 .$$

$$c \geqq \bar{c}(W) \iff b_1(W/c) = W_0 = a_1(W/c) .$$

Proof

By lemma 4.1, there corresponds to each pair of points $g, W \in G^0$ a positive number $\hat{c}(g, W)$, say, such that

$$c \lesseqgtr \hat{c}(g, W) \iff \bar{\rho}(g, W, c) \lesseqgtr \bar{\rho}(g, W)$$

The result follows from (4.2), the definitions [44] – [47] and theorem A, if we take $g = (W_1, W_0)$ in the above equivalence and define

$$\bar{c}(W) = \hat{c}((W_1, W_0), W) .$$

Lemma 4.3

Let $W = (W_0, W_1) \in G^0$ be arbitrary, fixed. Then

(i) $b_1(W/\cdot)$ is strictly increasing and continuous on $(0, \bar{c}(W))$ and

$$\lim_{c \to 0} b_1(W/c) = 0, \qquad \lim_{c \to \bar{c}(W)} b_1(W/c) = W_0 .$$

(ii) $a_1(W/\cdot)$ is strictly decreasing and continuous on $(0, \bar{c}(W))$ and

$$\lim_{c \to 0} a_1(W/c) = 1, \qquad \lim_{c \to \bar{c}(W)} a_1(W/c) = W_0 .$$

Proof

The proof follows immediately from lemmas 4.1 and 4.2.

Define functions ξ and η on Λ by

[49] $$\xi(\lambda) = \frac{a_0(\lambda)\ b_1(\lambda)}{a_1(\lambda)\ b_0(\lambda)} , \qquad \eta(\lambda) = \frac{b_0(\lambda)}{b_1(\lambda)} ,$$

then

$$b_1(\lambda) = \frac{1}{1+\eta(\lambda)} , \qquad a_1(\lambda) = \frac{1}{1+\xi(\lambda)\ \eta(\lambda)} .$$

Lemma 4.4

Let $W = (W_0, W_1) \in G^0$ be arbitrary, fixed. Then

(i) $\xi(W/\cdot)$ is strictly increasing and continuous on $(0, \bar{c}(W))$ and

$$\lim_{c \to 0} \xi(W/c) = 0, \qquad \lim_{c \to \bar{c}(W)} \xi(W/c) = 1 .$$

(ii) $\eta(W/\cdot)$ is strictly decreasing and continuous on $(0, \bar{c}(W))$ and

$$\lim_{c \to 0} \eta(W/c) = \infty , \qquad \lim_{c \to \bar{c}(W)} \eta(W/c) = W_1/W_0 .$$

Proof

The proof follows immediately from definition [49] and lemma 4.3.

Lemma 4.5

There exists a mapping

[50] $$c^*: \; G^0 \times (0, 1) \longrightarrow (0, \infty)$$

such that

$$0 < c^*(W, u) < \bar{c}(W), \qquad W \in G^0, \quad 0 < u < 1 ,$$

and such that

$$\xi(W/c^*(W, u)) \equiv u , \qquad W \in G^0, \quad 0 < u < 1 .$$

Proof

The proof follows immediately from lemma 4.4.

Corollary 4.5

$$\xi(W/c^*(W, u)) \, \eta(W/c^*(W, u)) \equiv \frac{a_0(W/c^*(W, u))}{a_1(W/c^*(W, u))} , \qquad W \in G^0, \quad 0 < u < 1.$$

Lemma 4.6

There exists a unique mapping W*

$$W^*: \; (0, 1) \times (0, \infty) \longrightarrow G^0$$

such that

$$\eta(W^*(u, z)/c^*(W^*(u, z), u)) \equiv z , \quad 0 < u < 1, \quad 0 < z < \infty .$$

Proof

Observe first of all that by lemma 4.2, definitions [44] – [47], (4.2) and corollary A2, we have for each $W \in G^0$

$$(4.3) \qquad 0 < c < \bar{c}(W) \iff \bar{\gamma}((W_1, W_0), W, c) > 0 .$$

Hence by lemma 2.6 and (4.2) we have for each $W \in G^0$ that

$$(4.4) \qquad 0 < c < \bar{c}(W) \implies \tilde{S}(g, W/c) = \hat{S}(B(g, W/c), A(g, W/c)), \quad \underline{\text{all}}\ g \in G^0,$$

where the equality on the right hand side holds in each case outside a subset of $\mathfrak{X}$ which is null under both P_0 and P_1. Since the above implication is for all $g \in G^0$, it holds in particular for $g = b(W/c)$ and $g = a(W/c)$. Note Further that by definitions [32] and [49],

$$B(a(W/c), W/c) = \xi(W/c) , \qquad A(a(W/c), W/c) = 1$$

$$B(b(W/c), W/c) = 1 \quad , \qquad A(b(W/c), W/c) = 1/\xi(W/c) .$$

By (4.4) and lemma 4.5, we now have for each $W \in G^0$ and each number u, $0 < u < 1$, that outside a subset of $\mathfrak{X}$ which is null under both P_0 and P_1,

$$\tilde{S}(b(W/c^*(W, u)), W/c^*(W, u)) = \hat{S}(1, 1/u)$$

$$\tilde{S}(a(W/c^*(W, u)), W/c^*(W, u)) = \hat{S}(u, 1) .$$

By (4.1) this means that for each $W \in G^0$ and each $u \in (0, 1)$,

$$(4.5) \qquad \begin{aligned} \bar{\rho}(b(W/c^*(W, u)), W, c^*(W, u)) &= \bar{R}(b(W/c^*(W, u)), W, c^*(W, u) | \hat{S}(1, 1/u)) \\ \bar{\rho}(a(W/c^*(W, u)), W, c^*(W, u)) &= \bar{R}(a(W/c^*(W, u)), W, c^*(W, u) | \hat{S}(u, 1)) . \end{aligned}$$

On the other hand, by (4.3), (4.2) and theorem A,

$$0 < c < \bar{c}(W) \implies \bar{\gamma}(b(W/c), W, c) = 0 = \bar{\gamma}(a(W/c), W, c) .$$

In particular, by lemma 4.5, we have for each $W \in G^0$ and each $u \in (0, 1)$, that

$$(4.6) \qquad \bar{\gamma}(b(W/c^*(W, u)), W, c^*(W, u)) = 0 = \bar{\gamma}(a(W/c^*(W, u)), W, c^*(W, u)) .$$

By definition [47], (4.6), (4.5), and by lemma 4.2, we now have for each $W \in G^0$ and each $u \in (0, 1)$ that

$$\bar{R}(b(W/c^*(W, u)), W, c^*(W, u)|\hat{S}(1, 1/u)) - W_1 b_1(W/c^*(W, u)) = 0$$

$$\bar{R}(a(W/c^*(W, u)), W, c^*(W, u)|\hat{S}(u, 1)) - W_0 a_0(W/c^*(W, u)) = 0 .$$

If we divide the first equation by $b_1(W/c^*(W, u))$, the second by $a_1(W/c^*(W, u))$, we obtain, using definition [45] and lemma 4.5 and its corollary, after some minor rearrangement, the equations

$$[\mathcal{E}_0^{(0)}\eta^* + \mathcal{E}_1^{(0)}]c^* + W_0 Q_0^{(0)}\eta^* - W_1(1-Q_1^{(0)}) = 0 ,$$

$$[u\,\mathcal{E}_0^{(1)}\eta^* + \mathcal{E}_1^{(1)}]c^* + W_1 Q_1^{(1)} - W_0 u(1-Q_0^{(1)})\eta^* = 0 ,$$

where we have adopted the following abreviated notation:

$$\mathcal{E}_i^{(0)} = E_i\hat{n}(\cdot|1, 1/u), \qquad Q_i^{(0)} = \hat{Q}_i(1, 1/u), \qquad i=0,1$$

$$\mathcal{E}_i^{(1)} = E_i\hat{n}(\cdot|u, 1), \qquad Q_i^{(1)} = \hat{Q}_i(u, 1), \qquad i=0,1$$

$$c^* = c^*(W, u), \qquad \eta^* = \eta(W/c^*(W, u)) .$$

If we eliminate c^* between these two equations, we get

(4.7)
$$[u\,\mathcal{E}_0^{(1)}\eta^* + \mathcal{E}_1^{(1)}][W_0 Q_0^{(0)}\eta^* - W_1(1-Q_1^{(0)})] + [\mathcal{E}_0^{(0)}\eta^* + \mathcal{E}_1^{(0)}][W_0 u(1-Q_0^{(1)})\eta^* - W_1 Q_1^{(1)}] = 0 .$$

Now let z be an arbitrary, fixed positive number and set

$$\eta^* = \eta(W/c^*(W, u)) = z .$$

If we substitute this into (4.7), we get an equation which is linear in W_1 (recall that $W_0 = 1-W_1$) and which may be easily solved. Denote the solution by $W_1^*(u, z)$. Let $W_0^*(u, z) = 1-W_1^*(u, z)$ and take

$$W^*(u, z) = (W_0^*(u, z), W_1^*(u, z)) .$$

Thus

$$\eta(W/c^*(W, u)) = z \implies W = W^*(u, z) .$$

On the other hand, if we set

$$W_1 = W_1^*(u, z)$$

and regard this as an equation in z, we get a quadratic with coefficient of z^2 positive and constant term negative. But this implies the existence of a unique positive root. In view of (4.7), this root must be the number $\eta(W/c^*(W, u))$. Thus

$$W = W^*(u, z) \implies \eta(W/c^*(W, u)) = z.$$

This completes the proof.

Theorem D (alternative proof)

There exists a mapping

$$\bar{\lambda}: \{(g, u, v): g \in G^0, \; 0 < u < 1 < v\} \longrightarrow \Lambda$$

such that identically on its domain

$$B(g, \bar{\lambda}(g, u, v)) = u , \qquad A(g, \bar{\lambda}(g, u, v)) = v .$$

Proof

Let (g, u, v) be an arbitrary, fixed point in the hypothesized domain of $\bar{\lambda}$. By lemma 4.5 we have identically for all $W \in G^0$ that

$$\xi(W/c^*(W, u/v)) = u/v .$$

By lemma 4.6

$$\eta\left(W^*(u/v,\ g_0/g_1u)\Big/\ c^*(W^*(u/v,\ g_0/g_1u),\ u/v)\right) = g_0/g_1u\ .$$

The conclusion now follows from definitions [32] and [49] if we take

$$\overline{\lambda}(g,\ u,\ v) = \frac{W^*(u/v,\ g_0/g_1u)}{c^*(W^*(u/v,\ g_0/g_1u),\ u/v)}\ .$$

INDEX OF NOTATION