LOCKHEED MARTIN ENERGY RESEARCH LIBRARIES
3 4456 0515741 5

CENTRAL RESEARCH LIBRARY
DOCUMENT COLLECTION

2

ORNL-4208
UC-32 – Mathematics and Computers

INFORMATION THEORY IN SEQUENTIAL ANALYSIS

(Thesis)

Roger Lee Tittle

OAK RIDGE NATIONAL LABORATORY
CENTRAL RESEARCH LIBRARY
DOCUMENT COLLECTION
LIBRARY LOAN COPY
DO NOT TRANSFER TO ANOTHER PERSON
If you wish someone else to see this document, send in name with document and the library will arrange a loan.
UCN-7969
(3 3-67)

OAK RIDGE NATIONAL LABORATORY
operated by
UNION CARBIDE CORPORATION
for the
U.S. ATOMIC ENERGY COMMISSION

Printed in the United States of America. Available from Clearinghouse for Federal
Scientific and Technical Information, National Bureau of Standards,
U.S. Department of Commerce, Springfield, Virginia 22151
Price: Printed Copy $3.00; Microfiche $0.65

LEGAL NOTICE

This report was prepared as an account of Government sponsored work. Neither the United States, nor the Commission, nor any person acting on behalf of the Commission:

A. Makes any warranty or representation, expressed or implied, with respect to the accuracy, completeness, or usefulness of the information contained in this report, or that the use of any information, apparatus, method, or process disclosed in this report may not infringe privately owned rights; or

B. Assumes any liabilities with respect to the use of, or for damages resulting from the use of any information, apparatus, method, or process disclosed in this report.

As used in the above, "person acting on behalf of the Commission" includes any employee or contractor of the Commission, or employee of such contractor, to the extent that such employee or contractor of the Commission, or employee of such contractor prepares, disseminates, or provides access to, any information pursuant to his employment or contract with the Commission, or his employment with such contractor.

ORNL-4208

Contract No. W-7405-eng-26

MATHEMATICS DIVISION

INFORMATION THEORY IN SEQUENTIAL ANALYSIS

Roger Lee Tittle

Submitted as a dissertation to the Graduate Council of The University of Tennessee in partial fulfillment of the requirements for the degree of Doctor of Philosophy

JANUARY 1968

OAK RIDGE NATIONAL LABORATORY
Oak Ridge, Tennessee
operated by
UNION CARBIDE CORPORATION
for the
U.S. ATOMIC ENERGY COMMISSION

LOCKHEED MARTIN ENERGY RESEARCH LIBRARIES
3 4456 0515741 5

ACKNOWLEDGEMENTS

The author wishes to express his sincere appreciation to his major professor, Dr. G. E. Albert, for his many valuable suggestions, time, and encouragement throughout the research and writing of this dissertation.

In addition, grateful acknowledgement is made to the Oak Ridge Associated Universities for financial assistance throughout the duration of this project, and to the Mathematics Division of the Oak Ridge National Laboratory, operated by Union Carbide Corporation for the U. S. Atomic Energy Commission, for the use of office space.

The author also wishes to express his gratitude to Mrs. Charlotte Montgomery for an exceptionally fine typing of a difficult manuscript.

TABLE OF CONTENTS

CHAPTER		PAGE
I.	INTRODUCTION	1
	Review of the Literature	1
	Summary of Contents	3
II.	DEFINITION OF PROBABILITY SPACES	6
III.	INFORMATION IN THREE BOREL FIELDS	14
	Introduction	14
	Background Material	15
	Integral Representations	19
IV.	RELATIONSHIPS BETWEEN THE INFORMATION IN THREE BOREL FIELDS	35
V.	CONCLUDING REMARKS	47
BIBLIOGRAPHY		51

CHAPTER I

INTRODUCTION

Review of the Literature

Let $(\Omega,\mathcal{B})$ be an arbitrary measurable space and let P_1 and P_2 be two probability measures defined on $\mathcal{B}$. For any subfield $\mathcal{J}$ of $\mathcal{B}$ let $P_{i\mathcal{J}}$ denote the restrictions of P_i to $\mathcal{J}$, $i = 1,2$. If $\{F_k\}$ is a finite $\mathcal{B}$-measurable partition of Ω and $\mathcal{J}$ is the finite subfield of $\mathcal{B}$ consisting of all unions of the sets F_k, then the quantity

$$(1.1) \qquad I(\mathcal{J}) = \sum_k P_{1\mathcal{J}}(F_k) \log \left[P_{1\mathcal{J}}(F_k) \,/\, P_{2\mathcal{J}}(F_k) \right]$$

will be called the <u>information in the field $\mathcal{J}$ for discrimination between P_1 and P_2</u>. In (1.1) and its extensions to be made in this paper the natural logarithm will always be used and terms $P_{1\mathcal{J}}(F_k)\log[P_{1\mathcal{J}}(F_k)/P_{2\mathcal{J}}(F_k)]$ will be assigned the value zero whenever $P_{1\mathcal{J}}(F_k) = 0$ regardless of the value of $P_{2\mathcal{J}}(F_k)$. The information $I(\mathcal{J})$ is always non-negative [6]; each of the terms in its defining sum is greater than $-\infty$; and $I(\mathcal{J}) = +\infty$ if and only if $P_{1\mathcal{J}}(F_k) > P_{2\mathcal{J}}(F_k) = 0$ for some k.

The information measure $I(\mathcal{J})$ above has been discussed by several authors [6], [10], [11], and in 1959 the definition was extended by G. Kallianpur [4] to cover an arbitrary Borel subfield $\mathcal{A}$ of $\mathcal{B}$. The basic property of $I(\mathcal{J})$ that underlies Kallianpur's extension of the definition (1.1) is the following. Let $\mathcal{J}_1$ be a finite subfield of $\mathcal{B}$

generated by a refinement of the partition $\{F_k\}$ of Ω, then $I(\mathcal{F}_1) \geq I(\mathcal{F})$. This is easily shown to be true by use of Jensen's inequality for continuous convex functions. It is natural then to extend the definition (1.1) in the following manner. For an arbitrary Borel subfield $\mathcal{A}$ of $\mathcal{B}$ let

$$I(\mathcal{A}) = \sup_{\mathcal{F} \subseteq \mathcal{A}} I(\mathcal{F}) ,$$

where the supremum is taken over all finite subfields of $\mathcal{A}$.

By use of properties of martingales and semi-martingales, Kallianpur derives necessary and sufficient conditions under which $I(\mathcal{A})$ is finite as well as an integral representation for $I(\mathcal{A})$. The main result of his paper is

Theorem 1.1. Let $\mathcal{A}$ be any Borel field of subsets of Ω and let $I(\mathcal{A})$ be defined by (1.2).

(i) If $P_{1\mathcal{A}} \not\ll P_{2\mathcal{A}}$, ($P_{1\mathcal{A}}$ is not absolutely continuous with respect to $P_{2\mathcal{A}}$), then $I(\mathcal{A}) = \infty$.

(ii) If $P_{1\mathcal{A}} \ll P_{2\mathcal{A}}$, then

$$I(\mathcal{A}) = \int_\Omega X \log X dP_{2\mathcal{A}} = \int_\Omega \log X dP_{1\mathcal{A}},$$

where X is the Radon-Nikodym (R-N) derivative of $P_{1\mathcal{A}}$ with respect to $P_{2\mathcal{A}}$, whether $I(\mathcal{A})$ is finite or infinite.

(iii) $I(\mathcal{A})$ is finite if and only if

(a) $P_{1\mathcal{A}} \ll P_{2\mathcal{A}}$

and

(b) $E_{2\mathcal{A}}|X \log X| < \infty$,

where $E_{i\mathcal{A}}$ denotes expectation on the probability space $(\Omega, \mathcal{A}, P_{i\mathcal{A}})$, $i = 1,2$.

Theorem 1.1 is important because it relates the basic definition (1.1) to an integral representation which was assumed to be the information in the general case by Kullback and Leibler in [6]; and which has been applied extensively since that time to the theory of statistical hypotheses testing in the fixed sample size case.

Summary of Contents

In this paper we investigate the information as defined by (1.2) in a sequential test S of simple hypotheses H_i, $i = 1,2$, which is determined by cumulative sums of independent and identically distributed random variables, $X_1, X_2, \ldots X_N, \ldots$, and which terminates with probability 1. For the problem of testing one simple hypothesis against another, of all tests whose probabilities of incorrectly accepting the second hypothesis are bounded above by given bounds, the Wald sequential probability ratio test gives the smallest expectation of the sample size under either hypothesis. For this reason, particular emphasis is given to this special case of the general test S.

By defining probability spaces $(\Omega, \mathcal{B}, P_\ell)$, $\ell = 1,2$, appropriate to any such test S, we derive integral representations for the information $I(\mathcal{A})$, $I(\mathcal{C})$, and $I(\mathcal{D})$ in three particular Borel sub-fields $\mathcal{A}$, $\mathcal{C}$, and $\mathcal{D}$ of $\mathcal{B}$ in terms of the conditional probability densities of the cumulative sum variables. The fields $\mathcal{A}$, $\mathcal{C}$, and $\mathcal{D}$ are chosen so that the quantities $I(\mathcal{A})$, $I(\mathcal{C})$, and $I(\mathcal{D})$ may be appropriately referred to repsectively as the total information in the decision process, the information in the sequential decision variables, and the information in the stopping rule. In the special case of a sequential probability ratio test of hypotheses $H_i : g(u) = g_i(u)$, $i = 1,2$, on the density function $g(u)$ of an observable random variable U, we show that

$$I(\mathcal{A}) = E_{1\mathcal{A}}(n) \int_{-\infty}^{\infty} \log \left[\frac{g_1(u)}{g_2(u)}\right] g_1(u)du \ .$$

This generalizes the familiar additivity property of the information in fixed sample size random sampling, Kullback [7].

In general the Borel fields $\mathcal{A}$, $\mathcal{C}$ and $\mathcal{D}$ have the properties that $I(\mathcal{A}) \geq I(\mathcal{C}) \geq I(\mathcal{D})$, but necessary and sufficient conditions on the conditional density functions of the cumulative sums will be stated under which $I(\mathcal{A}) = I(\mathcal{C})$ and $I(\mathcal{A}) = I(\mathcal{D})$ when $I(\mathcal{A}) < \infty$, and $I(\mathcal{C}) = I(\mathcal{D})$ when $I(\mathcal{C}) < \infty$. Again, in the case of a sequential probability ratio test, it is demonstrated that $I(\mathcal{A}) = I(\mathcal{C})$ whenever $g_i(u)$, $i = 1,2$, belong to an important class of density functions, namely the exponential family.

All of the primary results stated previously are obtained for the case in which all of the random variables under consideration admit a Borel measurable density function; however, by interpreting integrals as sums, and replacing Lebesque measure by the counting measure, the results remain valid for the discrete case as well.

Care has been used to impose only conditions which are accepted as basic in sequential analysis. In this connection see [12].

CHAPTER II

DEFINITION OF PROBABILITY SPACES

In this chapter we discuss the probability spaces (Ω, B, P_i), $i = 1,2$, appropriate for a study of the general sequential tests S mentioned in Chapter I. Such tests S will include those based upon a random walk with independent increments and terminating barriers, -b and a, defined as follows. For any sequence $X_1, X_2, \ldots X_N, \ldots$ of independent random variables with a common density function $f(x)$ define the sequence $Z_1 = X_1$, $Z_2 = \sum_{j=1}^{2} X_j, \ldots$ $Z_N = \sum_{j=1}^{N} X_j, \ldots$. Let the cumulative sums Z_N of the X_j's begin at $Z_0 = 0$ inside $(-b,a)$ and change (walk) by independent increments $Z_N - Z_{N-1} = X_N$ between $(-b,a)$ until an increment X_N takes Z_N outside $(-b,a)$. The process then terminates at Z_N with a type 1 termination if $Z_N \geq a$ and with a type 2 termination if $Z_N \leq -b$.

Historically, Wald suggested using random walks of this special type in his sequential probability ratio tests for deciding between hypothesized density functions $g_1(u)$, $g_2(u)$. We extend Wald's idea to the case of a sequential test which is determined by cumulative sums of independent random variables and a more general type of random walk.

Specifically, for each $k = 1,2,\ldots$, let $\psi_i^k(z)$, $i = 0,1,2$, be three Borel measurable functions defined and non-negative on $-\infty < z < \infty$ and such that for each z,

$$\sum_{i=0}^{2} \psi_i^k(z) = 1. \tag{2.1}$$

For each $j = 1,2,\ldots$, $N = 1,2,\ldots$, let X_j and Z_N be defined as in the random walk with terminating barriers. The walk with independent random increments X_N is terminated at Z_N in a type 1 termination with conditional probability $\psi_1^N(z_N)$, or in a type 2 termination with conditional probability $\psi_2^N(z_N)$, or is continued to Z_{N+1} with conditional probability $\psi_0^N(z_N)$. Clearly the walk with terminating barriers, -b and a, is a simple case of this more general process. We refer to this type of walk as a random walk with randomized terminations.

Now let $g(u)$ be the probability density function of a random variable u and let H_1: $g(u) = g_1(u)$, H_2: $g(u) = g_2(u)$ be simple hypothesis on $g(u)$. Let S be a sequential test of H_1 against H_2 which is determined by cumulative sums of independent random variables, $X_1, X_2, \ldots$, and which specifies the probability decision functions ψ_0^k, ψ_1^k, ψ_2^k with the properties defined above. The decision rule for S is then the random walk with randomized terminations as follows. If the termination of the walk is of type 1, accept H_1. If the termination of the walk is of type 2, accept H_2.

For example, suppose S is a generalized sequential probability ratio test as defined in [13]. Then $X_j = \log \frac{g_1(u_j)}{g_2(u_j)}$, $Z_N = \sum_{j=1}^{N} X_j$ for each $j = 1,2,\ldots$, $N = 1,2,\ldots$, and ψ_0^k, ψ_1^k, ψ_2^k are the indicator functions of the respective intervals $-b_k < z < a_k$, $z_k \leq a$, $z_k \leq -b_k$ for each $k = 1,2,\ldots$.

Before defining the probability spaces appropriate to the general decision rule S, we note that the random variables $X_1, X_2, \ldots X_N, \ldots$ have a common density function $f_i(x)$ according to H_i, and since $Z_j - Z_{j-1} = X_j$ for $j \geq 1$, the conditional density of Z_j given Z_{j-1} must be $f_i(z_j - z_{j-1})$, where Z_0 and z_0 are equal to zero.

To define the probability spaces let Ω be the set of points ω which are infinite sequences

$$\omega = (\alpha_1, \alpha_2, \ldots, \alpha_k, \ldots)$$

in which each component α_k is a pair (z_k, y_k), where z_k is any real number and y_k takes on one of the values of the subscript i of $\psi_i^k(z_k)$. Each α_k will then denote a possible cumulative sum value and a possible decision at that sum value. The following subsets are of particular interest in connection with the decision rule S:

$$(2.2) \quad \left\{ \begin{array}{ll} A_{ky_k} & = \{\omega | \alpha_j = (z_j, 0),\ j \leq k-1 \text{ and } \alpha_k = (z_k, y_k),\ y_k = 0,1,2\} \\ B & = \{\omega | \alpha_j = (z_j, 0),\ j = 1,2,\ldots\} \ . \end{array} \right\}$$

Note that $B = \lim_{k \to \infty} A_{k0}$, and $\{A_{k0}, A_{j1}, A_{j2}\}_{j=1}^{k}$ is a partition of Ω for each $k = 1,2,\ldots$.

The sets A_{ky_k} are special cylinder subsets of Ω. In general a cylinder subset of Ω is defined by any restriction on a finite number

of the α_j's. In particular, Borel cylinder sets are defined by restricting $z_1,\ldots,z_m$, for some finite m, to a Borel set in Euclidean m space and by specifying a set of values of y_i, one for each index $j = 1,2,\ldots,m$. The special sets A_{ky_k} described above are Borel cylinder sets and all Borel cylinder sets are subsets of finite unions of the A_{ky_k}.

It is known that there exists a minimal Borel field $\mathcal{B}$ of subsets of Ω which contains all Borel cylinder subsets of Ω. Moreover it is known that if a function P(A) can be defined for all Borel cylinder subsets A of Ω in such a way that the axioms for a probability measure are satisfied on the Borel cylinder sets, then P(A) can be extended uniquely to a probability measure on the minimal Borel field $\mathcal{B}$. Hence we shall be content with the definition of P(A) for Borel cylinder subsets of Ω.

Let Λ_c be a Borel cylinder subset of Ω. Then Λ_c is described by Borel set type conditions on $\alpha_1,\alpha_2,\ldots,\alpha_m$ for some m. Since $\{A_{m0}, A_{j1}, A_{j2}\}_{j=1}^{m}$ is a partition of Ω, the sets

$$\left\{\begin{array}{l} \Lambda_{ki} = \Lambda_c \cap A_{ki},\ i = 1,2;\ k = 1,2,\ldots,m \\ \Lambda_{m0} = \Lambda_c \cap A_{m0} \end{array}\right\}$$

partition Λ_c.

Now $\omega \in \Lambda_{ki}$, $i = 1,2,$ implies that

$$\left\{\begin{array}{l} y_1 = i \quad \text{if } k = 1 \\ y_1 = \ldots = y_{k-1} = 0,\ y_k = i \text{ if } k = 2,\ldots,m \end{array}\right\}$$

and $\omega \in \Lambda_{m0}$ implies that $\{y_1 = \ldots = y_m = 0\}$. If $k < m$, let $(y^*_{k+1},\ldots,y^*_m)$ be any possible set of the y values of the indices $k+1,\ldots,m$. Let $\Lambda^{(y^*_{k+1},\ldots,y^*_m)}$ be the subset of Λ_{ki} with the given $(y^*_{k+1},\ldots,y^*_m)$ values. Then

$$\left\{\Lambda_{ki}^{(y^*_{k+1},\ldots,y^*_m)}\right\}_{(y^*_{k+1},\ldots,y^*_m)} \text{ partitions } \Lambda_{ki}.$$

For any fixed index sequence $(k,i,y^*_{k+1},\ldots,y^*_m)$, the set $\Lambda^{(y^*_{k+1},\ldots,y^*_m)}$ is specified by $(z_1,\ldots,z_m) \in B_m$, a Borel set in R_m. For $\ell = 1,2$, define

$$P_\ell(\Lambda_{ki}^{(y^*_{k+1},\ldots y^*_m)}) = \left\{\begin{array}{l} \Big(\int_{B_m}\ldots\int\Big[\prod_{j=1}^{k-1}\psi_0^j(z_j)\Big]\Big[\psi_i^k(z_k)\Big]\psi^{k+1}_{y^*_{k+1}}(z_k)\ldots\psi^m_{y^*_m}(z_m) \\ \quad .f_\ell(z_m-z_{m-1})f_\ell(z_{m-1}-z_{m-2})\ldots f_\ell(z_1)dz_1\ldots dz_m\Big), \\ \quad \text{if } 1 < k < m. \\ \\ \int_{B_m}\ldots\int\psi_i^1(z_1)\psi^2_{y^*_2}(z_2)\ldots\psi^m_{y^*_m}(z_m)f_\ell(z_m-z_{m-1})\ldots f_1(z_1) \\ \quad .dz_1\ldots dz_m\ , \\ \quad \text{if } k = 1. \end{array}\right\}$$

and, by additivity,

$$P_\ell(\Lambda_{ki}) = \sum_{(y^*_{k+1},\ldots,y^*_m)} P_\ell(\Lambda_{ki}^{y^*_{k+1},\ldots,y^*_m})$$

for $1 \leq k < m$. If we now define

$$P_\ell(\Lambda_{mi}) = \int_{B_m}\cdots\int\Big[\prod_{k=1}^{m-1}\psi_0^k(z_k)\Big]\psi_i^m(z_m)f_\ell(z_m-z_{m-1})\ldots f_1(z_1)dz_1\ldots dz_m$$

and

$$P_\ell(\Lambda_{m0}) = \int_{B_m}\cdots\int\Big[\prod_{k=1}^{m}\psi_0^k(z_k)\Big]f_\ell(z_m-z_{m-1})\ldots f_\ell(z_1)dz_1\ldots dz_m,$$

then

$$P_\ell(\Lambda_c) = \sum_{k=1}^{m}\Big(\sum_{i=1}^{2}P_\ell(\Lambda_{ki})\Big)+P_\ell(\Lambda_{m0}),$$

and the definition of P_ℓ for an arbitrary Borel cylinder set is complete. Clearly, P_ℓ satisfies the axioms for a probability measure on all Borel cylinder subsets of Ω since $\Omega = A_{10} \cup A_{11} \cup A_{12}$ implies that

$$P_\ell(\Omega) = P_\ell(A_{10}) + P_\ell(A_{11}) + P_\ell(A_{12})$$

$$= \int_R \psi_0^1(z_1)f_\ell(z_1)dz_1 + \int_R \psi_1^1(z_1)f_\psi(z_1)dz_1 + \int_R \psi_2^1(z_1)f_\ell(z_1)dz_1$$

$$= \int_R [\psi_0^1(z_1)+\psi_1^1(z_1)+\psi_2^1(z_1)]\ f_1(z_1)dz_1 = \int_R f_1(z_1)dz_1 = 1$$

from (2.1).

In order to guarantee that the sequential tests under consideration terminate with probability 1, and for other reasons which will become apparent later, we make the following assumptions:

Assumption 1: There is a constant ρ, $0 \leq \rho < 1$, and an integer M such that for all $m \geq M$ the inequality

$$\int_R \cdots \int_R \left[\prod_{k=1}^{m} \psi_0^k(z_k) \right] f_\ell(z_m - z_{m-1}) \cdots f_1(z_1) dz_1 \cdots dz_m \leq \rho$$

is satisfied for $\ell = 1,2$.

Assumption 2:

$$\int_R \left| \log \frac{f_1(x)}{f_2(x)} \right| f_1(x) dx < \infty .$$

For future use we define three random variables, $n(\omega)$, $Z_{n(\omega)}(\omega)$, $Y(\omega)$, which are closely associated with the decision processes.

Definition 2.1. For each $\omega \in \Omega - B$, let $n(\omega)$ be the least positive integer N such that the component $\alpha_N = (z_N, y_N)$ of ω has y_N equal to one or two; that is, $\alpha_{n(\omega)}$ is the component of ω at which the test terminates. Let $Z_{n(\omega)}(\omega)$ be the random variable whose value is the value of z_N when $n(\omega) = N$. Let $Y(\omega)$ be the random variable defined by

$$Y(\omega) = \left\{ \begin{array}{ll} 1 & \text{if} \quad \omega \in \bigcup_{k=1}^{\infty} A_{k1} \\ 2 & \text{if} \quad \omega \in \bigcup_{k=1}^{\infty} A_{k2} \end{array} \right\}$$

where the sets A_{ki}, $i = 1,2$, $k = 1,2,\ldots$, are defined in (2.2).

It should be noted here that the random variables $n(\omega)$, $Z_{n(\omega)}(\omega)$, and $Y(\omega)$ are defined only up to a set of P_ℓ measure 0, (Assumption 1 quarantees that $P_\ell(B) = 0$); however, this will suffice for our purposes. Also note that

$$\{\omega | n(\omega) = k\} = A_{k1} \cup A_{k2} , \tag{2.3}$$

for each $k = 1,2,\ldots.$

CHAPTER III

INFORMATION IN THREE BOREL SUBFIELDS

Introduction

As mentioned in Chapter I, three Borel subfields of $\mathcal{B}$ are of primary importance in connection with the information measure (1.2) applied to the probability spaces $(\Omega, \mathcal{B}, P_\ell)$, $\ell = 1,2$, defined in Chapter II. We are now in a position to define these fields.

Definition 3.1.

(i) Let $\mathcal{A}$ be the Borel field generated by the collection consisting of A_{ki} and all Borel cylinder subsets of A_{ki} for each $k = 1,2,\ldots$, $i = 1,2$, (here A_{ki} is defined by (2.2), and we take a Borel cylinder subset Λ of A_{ki} to mean that $\Lambda \subseteq A_{ki}$ and Λ is specified by $(z_1,\ldots,z_k) \in B_k$, a Borel set in R_k).

(ii) Let $\mathcal{C}$ be the Borel field induced by the random vector (n, Y, Z_n), defined in Definition 2.1.

(iii) Let $\mathcal{D}$ be the Borel field induced by (n,Y), where n and Y are defined in Definition 2.1.

In this chapter we obtain analytical representations for the information measures $I(\mathcal{A})$, $I(\mathcal{C})$, and $I(\mathcal{D})$ in the Borel subfields $\mathcal{A}$, $\mathcal{C}$, and $\mathcal{D}$ described in Definition 3.1.

Background Material

Preliminary to a discussion of $I(\mathcal{A})$, $I(\mathcal{C})$, and $I(\mathcal{D})$, we require the following definitions and lemmas.

Definition 3.2. A Borel field $\mathcal{J}$ is separable if it is generated by (is minimal over) a countable collection of sets.

Definition 3.3. Let $(\mathcal{X}, \mathcal{J})$ be a measurable space. A sequence $\{\mathcal{M}_s\} = \{\{M_s^j\}_j\}_{s=1}^{\infty}$ of finite $\mathcal{J}$-measurable partitions of $\mathcal{X}$ is called regular if for every s, $M_{s+1}^j \in \mathcal{M}_{s+1}$ is a subset of some $M_s^k \in \mathcal{M}_s$.

Relating Definitions 3.2 and 3.3 is

Lemma 3.1. Let $\mathcal{J}$ be a separable Borel field. There exists a regular sequence $\{\mathcal{M}_s\}_{s=1}^{\infty}$ of finite $\mathcal{J}$-measurable partitions of $\mathcal{X}$ such that $\mathcal{J}$ is generated by $\bigcup_{s=1} \mathcal{J}_s$, where $\mathcal{J}_s$ is the Borel field consisting of all unions of sets of $\mathcal{M}_s$, ([8], p. 355).

Of fundamental importance in the proofs of the theorems to follow is Doob's example on the theory of derivatives, ([1]), Example 1, p. 344). We modify and state his basic results as

Lemma 3.2. Let $(\mathcal{X}, \mathcal{J}, Q_i)$, $i = 1,2$, be probability spaces with $Q_1 \ll Q_2$ and let $\{\mathcal{M}_s\}_{s=1}^{\infty} = \{\{M_s^j\}_{j=1}^{k_s}\}_{s=1}^{\infty}$ be a regular sequence of finite $\mathcal{J}$-measurable partitions of $\mathcal{X}$. For each s, let $\mathcal{J}_s$ be the field consisting of all finite unions of the sets of $\mathcal{M}_s$, and let $\mathcal{J}_\infty$ be the Borel field generated by $\bigcup_{s=1}^{\infty} \mathcal{J}_s$.

If we define for each $s = 1,2,\ldots,$

$$X_s(\omega) = \sum_{j=1}^{k_s} \left[Q_1(M_s^j) \;/\; Q_2(M_s^j)\right] \cdot I_{M_s^j}(\omega),$$

where $Q_1(M_s^j)/Q_2(M_s^j) = 0$ if $Q_2(M_s^j) = 0$, and let X_∞ be the R-N derivative of $Q_{1\mathcal{J}_\infty}$ with respect to $Q_{2\mathcal{J}_\infty}$, then

$\{X_s,\ s \geq 1\}$ is a uniformly integrable martingale with $\lim_{s \to \infty} X_s(\omega) = X_\infty(\omega)$ a.s.$(Q_{2\mathcal{J}_\infty})$

Kallianpur proves the following specialization of Theorem 4.1's (ii)(b) of Doob [1].

Lemma 3.3. If (i) $\{X_s,\ s \geq 1\}$ is a uniformly integrable martingale, so that $\lim_{s \to \infty} X_s = X_\infty$ exists with probability 1, and (ii) $\varphi(t)$ is a real, continuous, convex function of the real variable t such that $E|\varphi(X_s)| < \infty$ for every finite s, then $\lim_{s \to \infty} E[\varphi(X_s)] = E[\varphi(X_\infty)]$.

We shall make extensive use of this lemma throughout the remainder of this paper.

Although the next lemma is only used in connection with the Wald sequential probability ratio test, we include it here for completeness.

Lemma 3.4. Let U be a random variable with density function $g(u)$, and let $H_i: g(u) = g_i(u)$, $i = 1,2$, be simple hypotheses on g. Suppose $g_2(u) = 0$ implies $g_1(u) = 0$ except on a set of Lebesque measure zero.

If the random variable

$$X(u) = \begin{cases} g_1(u)/g_2(u) & \text{if } g_2(u) \neq 0 \\ 1 & \text{if } g_2(u) = 0 \end{cases}$$

has probability density functions $f_i(x)$ under H_i, $i = 1,2$, then

$$\int_R \left[\log \frac{g_1(u)}{g_2(u)}\right] g_1(u)dm_1 = \int_R \left[\log \frac{f_1(x)}{f_2(x)}\right] f_1(x)dm_1,$$

where 0 log 0 is assigned the value zero, and m_1 denotes Lebesque measure on the Borel sets $\mathcal{B}_1$ of the real line R.

Proof: Let Q_1 and Q_2 be the probability measures defined on $\mathcal{B}_1$ by g_1 and g_2 respectively; that is,

$$Q_i(E) = \int_E g_i(u)dm_1 \text{ for each } E \in \mathcal{B}_1,\ i = 1,2.$$

Since $g_2(u) = 0$ implies $g_1(u)$ a.e.(m), we have $Q_1 \ll Q_2 \ll m_1$, and it follows from the definition of $X(u)$ and an application of the property $\frac{dQ_1}{dm_1} = \frac{dQ_1}{dQ_2}\frac{dQ_2}{dm_1}$ a.e.(m_1) of the R-N derivates that

$$X(u) = \frac{dQ_1}{dQ_2} \text{ a.s. } (P_2) .$$

Let $Q_X^{(i)}$, $i = 1,2$, be the probability measures induced by $X(u)$ on the Borel sets; that is,

$$Q_X^{(i)}(E') = Q_i(X^{-1}(E')) \text{ for each } E' \in \mathcal{B}_1. \tag{3.1}$$

Since $Q_1 \ll Q_2$, then $Q_X^{(1)} \ll Q_X^{(2)}$, and so it follows from the R-N theorem that exists a non-negative $\mathcal{B}_1$-measurable function h such that

$$(3.2) \qquad Q_X^{(1)}(E') = \int_{E'} h dQ_X^{(2)} \text{ for each } E' \in \mathcal{B}_1 \quad .$$

Now

$$(3.3) \qquad \int_{E'} h dQ_X^{(2)} = \int_E h(X) dQ_2 = \int_E X dQ_2 = Q_1(E), \; E = X^{-1}(E')$$

from (3.1), (3.2) and ([8], Theorem a., p. 342).

Let $\mathcal{B}_0$ be the Borel field induced by X on R, ($\mathcal{B}_0 = \{E \in \mathcal{B}_1 | E = X^{-1}(F)$ for some $F \in \mathcal{B}_1$). We now have

$$Q_{2\mathcal{B}_0}(E) = \int_E X dQ_{2\mathcal{B}_0} = \int_E h(X) dQ_{2\mathcal{B}_0}$$

for each $E \in \mathcal{B}_0$ from (3.3), and hence

$$h(X) = X \quad \text{a.s.}(Q_{1\mathcal{B}_0}) \; .$$

Thus, obviously,

$$\int_R \log [h(X)] \, dQ_1 = \int_R \log X dQ_1$$

and therefore

$$(3.4) \qquad \int_R \log h dQ_X^{(1)} = \int_R \log X dQ_1$$

from [8], Theorem a., p. 342.

By assumption, X has probability density functions f_1 and f_2 under H_1 and H_2 respectively, so

$$\int_R \log h dQ_X^{(1)} = \int_R \log \left[f_1(x) \,/\, f_2(x)\right] f_1(x) dm_1,$$

and finally, from (3.4) and the definition of h,

$$\int_R \log \left[f_1(x) \,/\, f_2(x)\right] f_1(x) dm_1 = \int_R \log \left[g_1(u) \,/\, g_2(u)\right] g_1(u) dm_1.$$

Integral Representations

We are now in a position to establish the

Theorem 3.1. If $(\Omega, \mathcal{B}, P_\ell)$ $\ell = 1,2$, are the probability spaces in Chapter II, and if $P_{1\mathcal{A}} \ll P_{2\mathcal{A}}$, then the information in the Borel field $\mathcal{A}$ of Definition 3.1 has the representation

$$I(\mathcal{A}) = \sum_{k=1}^{\infty} \Big(\sum_{i=1}^{2} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \log \left\{ \prod_{j=1}^{k} \frac{f_1(z_j - z_{j-1})}{f_2(z_j - z_{j-1})} \right\} \left[\prod_{j=1}^{k-1} \psi_0^j(z_j) \right]$$

$$\cdot\ \psi_i^k(z_k) f_1(z_k - z_{k-1}) \cdots f_1(z_1) dz_1 \cdots dz_k,$$

where $z_0 = 0$, and $\prod_{j=1}^{k-1} \psi_0^j(z_j) \equiv 1$ if $k = 1$.

Proof: Throughout this proof, whenever we say that a sequence $\{Y_s\}_{s=1}^{\infty}$ is a uniformly integrable martingale, or converges a.s., these statements will be understood as referring to the probability space $(\Omega, \mathcal{A}, P_{2\mathcal{A}})$, and we shall also let $\prod_{j=1}^{k-1} \psi_0^j(z_j) = 1$ if $k = 1$.

For each integer $k = 1,2,\ldots$, let $\mathcal{B}_k$ denote the Borel sets of $R_k:(z_1,\ldots,z_k)$. Since $\mathcal{B}_k$ is a separable field, by Lemma 3.1 there exists a regular sequence $\{\mathcal{E}_{sk}\}_{s=1}^{\infty} = \{\{E_{sk}^j\}_{j=1}^{\ell_{sk}}\}_{s=1}^{\infty}$ of finite $\mathcal{B}_k$-measurable partitions of R_k such that $\mathcal{B}_k$ is generated by $\bigcup_{s=1}^{\infty} \mathcal{C}_{sk}$, where $\mathcal{C}_{sk}$ is the field consisting of all finite unions of sets of $\mathcal{E}_{sk}$ for each $k = 1,2,\ldots$. For each $E_{sk}^j \in \mathcal{E}_{sk}$, let

$$M_{sk}^{ij} = \{\omega \in A_{ki} \mid (z_1,\ldots,z_k) \in E_{sk}^j\}, \quad i = 1,2;$$

then, for each $s = 1,2,\ldots$, $k = 1,2,\ldots$, $\{M_{sk}^{ij}\}_{j=1}^{\ell_{sk}}$ is a finite partition of A_{ki}.

Define the collections $\mathcal{M}_s$, $s = 1,2,\ldots$, of sets:

$$\mathcal{M}_1 = \left\{A_{10}, \{M_{11}^{1j}\}_{j=1}^{\ell_{11}}, \{M_{11}^{2j}\}_{j=1}^{\ell_{11}}\right\}$$

$$\mathcal{M}_2 = \left\{A_{20}, \{M_{21}^{1j}\}_{j=1}^{\ell_{21}}, \{M_{21}^{2j}\}_{j=1}^{\ell_{21}}, \{M_{22}^{1j}\}_{j=1}^{\ell_{22}}, \{M_{22}^{2j}\}_{j=1}^{\ell_{22}}\right\}$$

$$\vdots$$

$$\mathcal{M}_s = \left\{A_{s0}, \{M_{s1}^{1j}\}_{j=1}^{\ell_{s1}}, \{M_{s1}^{2j}\}_{j=1}^{\ell_{s1}} \ldots \{M_{ss}^{1j}\}_{j=1}^{\ell_{ss}}, \{M_{ss}^{2j}\}_{j=1}^{\ell_{ss}}\right\}$$

$$\vdots$$

Since for each $s = 1,2,\ldots$ the sets comprising $\mathcal{M}_s$ are a finite $\mathcal{A}$-measurable partition of Ω, $A_{s+10} \subseteq A_{s0}$, and $M_{s+1k}^{ij} \in \mathcal{M}_{s+1}$ is a subset of some $M_{sk}^{ip} \in \mathcal{M}_s$, $k = 1,2,\ldots$, $i = 1,2$, then $\{\mathcal{M}_s\}_{s=1}^{\infty}$ is a regular sequence.

Let $\mathcal{J}_s$ be the field consisting of all finite unions of sets of $\mathcal{M}_s$ and let $\mathcal{J}_\infty$ be the Borel field generated by the field $\bigcup_{s=1}^{\infty} \mathcal{J}_s$, then clearly $\mathcal{J}_\infty = \mathcal{A}$.

Define the sequence of functions $V_s(\omega)$, $s = 1,2,\ldots$, by

$$(3.5) \qquad V_s(\omega) = \sum_{k=1}^{s} \sum_{j=1}^{\ell_{sk}} \Big(\sum_{i=1}^{2} \Big[P_{1\mathcal{A}}(M_{sk}^{ij}) \,/\, P_{2\mathcal{A}}(M_{sk}^{ij}) \Big] \cdot I_{M_{sk}^{ij}}(\omega)$$

$$+ \Big[P_{1\mathcal{A}}(A_{s0}) \,/\, P_{2\mathcal{A}}(A_{s0}) \Big] \cdot I_{\mathcal{A}_{s0}}(\omega),$$

where $P_{1\mathcal{A}}(M_{sk}^{ij}) \,/\, P_{2\mathcal{A}}(M_{sk}^{ij}) = 0$ if $P_{2\mathcal{A}}(M_{sk}^{ij}) = 0$.

It follows from Lemma 3.2 that $\{V_s(\omega),\ s \geq 1\}$ is a uniformly integrable martingale and $\lim_{s \to \infty} V_s(\omega) = V_\infty$ a.s., where V_∞ is the R-N derivative of $P_{1\mathcal{A}}$ with respect to $P_{2\mathcal{A}}$ Now let

$$\varphi(t) = \begin{cases} t \log t & \text{if } t > 0 \\ 0 & \text{if } t = 0 \end{cases} .$$

Since $P_{1\mathcal{A}} \ll P_{2\mathcal{A}}$, we have $E_{2\mathcal{A}}[|\varphi(V_s)|] < \infty$ for each $s = 1,2,\ldots$, and it follows from Theorem 1.1 and Lemma 3.3 that

$$\lim_{s \to \infty} E_{2\mathcal{A}}[\varphi(V_s)] = E_{2\mathcal{A}}[\varphi(V_\infty)] = E_{1\varphi}[\log V_\infty] = I(\mathcal{A}).$$

Fix k, i arbitrarily and apply Lemma 3.3 to the uniformly integrable

martingale $\{V_s \cdot I_{A_{ki}},\ s \geq 1\}$ to obtain

$$(3.6) \qquad \lim_{s \to \infty} \int_{A_{ki}} \varphi(V_s) dP_{2\mathcal{A}} = \int_{A_{ki}} \varphi(V_\infty) dP_{2\mathcal{A}} = \int_{A_{ki}} \log V_\infty dP_{1\mathcal{A}} .$$

From the definition of V_s we have

$$(3.7) \qquad \int_{A_{ki}} \varphi(V_s) dP_{2\mathcal{A}} = \sum_{j=1}^{\ell_{sk}} P_{1\mathcal{A}}(M_{sk}^{ij}) \log \left[P_{1\mathcal{A}}(M_{sk}^{ij}) \ / \ P_{2\mathcal{A}}(M_{sk}^{ij}) \right] .$$

Equations (3.6) and (3.7) combine to

$$(3.8) \qquad \lim_{s \to \infty} \sum_{j=1}^{\ell_{sk}} P_{1\mathcal{A}}(M_{sk}^{ij}) \log \left[P_{1\mathcal{A}}(M_{sk}^{ij}) \ / \ P_{2\mathcal{A}}(M_{sk}^{ij}) \right]$$

$$= \int_{A_{ki}} \varphi(V_\infty) dP_{2\mathcal{A}} = \int_{A_{ki}} \log (V_\infty) dP_{1\mathcal{A}} .$$

Thus if we show that the left-hand side of equation (3.8) is equal to

$$(3.9) \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \log \left\{ \prod_{j=1}^{k} \frac{f_1(z_j - z_{j-1})}{f_2(z_j - z_{j-1})} \right\} \left[\prod_{j=1}^{k-1} \psi_0^j(z_j) \right] \psi_i^k(z_k) f_1(z_k - z_{k-1}) \ldots f_1(z_1) dz_1 \ldots dz_k,$$

where $z_0 = 0$,

then the representation stated in the theorem will follow from

$$I(\mathcal{A}) = \int_{\Omega} \varphi(V_\infty) dP_{2\mathcal{A}} = \sum_{k=1}^{\infty} \left(\sum_{i=1}^{2} \int_{A_{ki}} \varphi(V_\infty) dP_{2\mathcal{A}} \right) .$$

To obtain (3.9) for an arbitrary but fixed k and i, define

$$\mu_{\ell ki} = \int_E \cdots \int \left[\prod_{j=1}^{k-1} \psi_0(z_j) \right] \psi_i^k(z_k) f_\ell(z_k - z_{k-1}) \cdots f_\ell(z_1) dz_1 \cdots dz_k$$

for $\ell = 1,2$, and for each $E \in \mathcal{B}_k$.

By the definition of $P_{\ell\mathcal{A}}$, $\mu_{\ell ki}$ is a finite measure on $\mathcal{B}_k$, (We assume $\mu_{\ell ki}$ is not identically zero for otherwise the result is trivial). Let

$$C_{\ell ki} = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \left[\prod_{j=1}^{k-1} \psi_0^j(z_j) \right] \psi_i^k(z_k) f_\ell(z_k - z_{k-1}) \cdots f_\ell(z_1) dz_1 \cdots dz_k$$

and define

$$P_{\ell ki}(E) = \frac{\mu_{\ell ki}(E)}{C_{\ell ki}}, \quad \ell = 1,2, \text{ for each } E \in \mathcal{B}_k.$$

Now $P_{1\mathcal{A}} \ll P_{2\mathcal{A}}$ by assumption, so, from the definitions of $P_{\ell ki}$, $\ell = 1,2$, we have that P_{1ki} and P_{2ki} are probability measures on $\mathcal{B}_k$ and $P_{1ki} \ll P_{2ki} \ll m_k$, where m_k denotes Lebesque measure on $\mathcal{B}_k$. Since $\{\mathcal{E}_{sk}\}_{s=1}^{\infty} = \{\{E_{sk}^j\}_{j=1}^{\ell_{sk}}\}_{s=1}^{\infty}$ is a regular sequence of finite $\mathcal{B}_k$-measurable partitions of R_k such that $\mathcal{B}_k$ is generated by $\bigcup_{s=1}^{\infty} \mathcal{C}_{sk}$, it follows similarly to the first part of the proof that

$$\text{(3.10)} \qquad \lim_{s \to \infty} \sum_{j=1}^{\ell_{sk}} P_{1ki}(E_{sk}^j) \log \left[P_{1ki}(E_{sk}^j) \,/\, P_{2ki}(E_{sk}^j) \right]$$

$$= \int_{R_k} X_{ki} \log X_{ki} \, dP_{2ki} = \int_{R_k} \log X_{ki} \, dP_{1ki},$$

where X_{ki} denotes the R-N derivative of P_{1ki} with respect to P_{2ki}. Consider the Borel measurable functions

$$t_{1ki}(z_1,\ldots,z_k) = \frac{1}{C_{1ki}}\left\{\left[\prod_{j=1}^{k-1}\psi_0^j(z_j)\right]\psi_i^k(z_k)f_1(z_k-z_{k-1})\ldots f_1(z_1)\right\}$$

$$t_{2ki}(z_1,\ldots,z_k) = \frac{1}{C_{2ki}}\left\{\left[\prod_{j=1}^{k-1}\psi_0^j(z_j)\right]\psi_i^k(z_k)f_2(z_k-z_{k-1})\ldots f_2(z_1)\right\}.$$

They have the properties that

$$P_{1ki}(E) = \int_E t_{1ki}dm_k,\ P_{2ki}(E) = \int_E t_{2ki}dm_k \text{ for each } E \in \mathcal{B}_k,$$

and so, by the R-N theorem, are equal a.e. (m_k) to the R-N derivatives $\frac{dP_{2ki}}{dm_k}$ and $\frac{dP_{2ki}}{dm_k}$ respectively. From the definitions of $\mu_{\ell ki}$ and $P_{\ell ki}$, and an application of (3.10) and the property $\frac{dP_{1ki}}{dm_k} = \frac{dP_{1ki}}{dP_{2ik}}\frac{dP_{2ki}}{dm_k}$ of the R-N derivatives, we obtain

$$\lim_{s\to\infty}\sum_{j=1}^{\ell_{sk}} P_{1ki}(E_{sk}^j)\log\left[P_{1ki}(E_{sk}^j)\ /\ P_{2ki}(E_{sk}^j)\right]$$

$$= \log\frac{C_{2ki}}{C_{1ki}} + \frac{1}{C_{1ki}}\lim_{s\to\infty}\sum_{j=1}^{\ell_{sk}}\mu_{1ki}(E_{sk}^j)\log\left[\mu_{1ki}(E_{sk}^j)\ /\ \mu_{2ki}(E_{sk}^j)\right]$$

$$= \int_{R_k} X_{ki}\log X_{ki}\ dP_{2ki}$$

$$= \int_{R_k} \log X_{ki} \, dP_{1ki}$$

$$= \log \frac{C_{1ki}}{C_{2ki}} + \frac{1}{C_{1ki}} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \log \left\{ \prod_{j=1}^{k} \frac{f_1(z_j - z_{j-1})}{f_2(z_j - z_{j-1})} \right\}$$

$$\cdot \left[\prod_{j=1}^{k-1} \psi_0^j(z_j) \right] \psi_i^k(z_k) f_1(z_k - z_{k-1}) \cdots f_1(z_1) dz_1 \cdots dz_k,$$

where $z_0 = 0$.

It now follows that the integral in (3.9) is equal to

$$\lim_{s \to \infty} \sum_{j=1}^{\ell_{sk}} \mu_{1ki}(E_{sk}^j) \log \left[\frac{\mu_{1ki}(E_{sk}^j)}{\mu_{2ki}(E_{sk}^j)} \right]$$

and thus to

$$\lim_{s \to \infty} \sum_{j=1}^{\ell_{sk}} P_{1\mathcal{A}}(M_{sk}^{ij}) \log \left[\frac{P_{1\mathcal{A}}(M_{sk}^{ij})}{P_{2\mathcal{A}}(M_{sk}^{ij})} \right],$$

since

$$\mu_{1ki}(E_{sk}^j) = P_{1\mathcal{A}}(M_{sk}^{ij}) \text{ and } \mu_{2ki}(E_{sk}^j) = P_{2\mathcal{A}}(M_{sk}^{ij}),$$

for each $j = 1,2,\ldots, \ell_{sk}$, $s = 1,2,\ldots$.

If the test S under investigation is a Wald sequential probability ratio test of $H_1: g(u) = g_1(u)$ against $H_2: g(u) = g_2(u)$, then for each $k = 1,2,\ldots$, $N = 1,2,\ldots$,

$$X_k = \log \frac{g_1(u_k)}{g_2(u_k)}, \quad Z_N = \sum_{k=1}^{N} X_k, \text{ and } \psi_0^k, \psi_1^k, \psi_2^k$$

are the indicator functions of the intervals $(-b,a)$, $z \leq -b$, $z \leq a$ respectively, where a and b are positive real numbers. In this case the infinite series for $I(\mathcal{A})$ reduces to

$$(3.11)\quad \sum_{k=1}^{\infty}\Big[\int_{-b}^{a}\int_{-b}^{a}\cdots\int_{-\infty}^{-b}\log\Big\{\prod_{j=1}^{k}\frac{f_1(z_j-z_{j-1})}{f_2(z_j-z_{j-1})}\Big\}f_1(z_k-z_{k-1})\cdots f_1(z_1)dz_1\cdots dz_k$$

$$+\int_{-b}^{a}\int_{-b}^{a}\cdots\int_{a}^{\infty}\log\Big\{\prod_{j=1}^{k}\frac{f_1(z_j-z_{j-1})}{f_2(z_j-z_{j-1})}\Big\}f_1(z_k-z_{k-1})\cdots f_1(z_1)dz_1\cdots dz_k\Big],$$

$z_0 = 0$.

Now let $(R_\infty, \mathcal{S}_\infty, P_\infty)$ be the probability space induced by the above stochastic process when P_∞ is generated by $f_1(x)$, (the components of sequences $X = (x_1,x_2,\ldots)$ in R_∞ all have the same density $f_1(x)$ and they are totally independent random variables). Define $n^*(X)$ on R_∞ to be the least positive integer N such that one of the inequalities $z_N = \sum_{j=1}^{N} x_j \geq a$ or $z_N = \sum_{j=1}^{N} x_j \leq -b$ holds. It follows at once from the definitions of P_∞ and $P_{1\mathcal{A}}$ that the sum (3.11) is equal to

$$(3.12)\qquad E_\infty\Big[\sum_{j=1}^{n^*(X)}\log\frac{f_1(x_j)}{f_2(x_j)}\Big],$$

and

$$(3.13)\qquad E_{1\mathcal{A}}(n) = 1 + \sum_{k=1}^{\infty} P_{1\mathcal{A}}(n > k) = 1 + \sum_{k=1}^{\infty} P_\infty(n^* > k) = E_\infty(n^*)\quad .$$

Thus if it is shown that the quantity (3.12) is equal to

$$E_\infty(n^*) \int_{-\infty}^{\infty} \left[\log \frac{f_1(x)}{f_2(x)}\right] f_1(x)dx ,$$

then by Lemma 3.4 and from (3.13) we shall have

Theorem 3.2. In the case of a Wald sequential probability ratio test,

$$I(\mathcal{A}) = E_{1\mathcal{A}}(n) \int_{-\infty}^{\infty} \log \left[\frac{g_1(u)}{g_2(u)}\right] g_1(u)du$$

Proof: As remarked above, it suffices to show that

$$E_\infty\left[\sum_{j=1}^{n^*(X)} \log \frac{f_1(x_j)}{f_2(x_j)}\right] = E_\infty(n^*) \int_{-\infty}^{\infty} \left[\log \frac{f_1(x)}{f_2(x)}\right] f_1(x)dx.$$

For this purpose, let

$$W_j = \log \frac{f_1(x_j)}{f_2(x_j)} , \quad C = \int_{-\infty}^{\infty} \left[\log \frac{f_1(x_j)}{f_2(x_j)}\right] f_1(x_j)\, dx_j ,$$

and $S_k = \sum_{j=1}^{k} [W_j - C]$ for each $j = 1,2,\ldots,$ $k = 1,2,\ldots.$

Consider $\{(S_k, \mathcal{L}_k)\}_{k=1}^{\infty}$, where $\mathcal{L}_k$ is the Borel subfield of $\mathcal{S}_\infty$ induced by $X_1,\ldots,X_k$. Since W_{k+1} is a Borel measurable function of X_{k+1}, it follows that $X_1,\ldots,X_k$, W_{k+1} are independent and S_k is $\mathcal{L}_k$-measurable.

Thus

$$E_\infty[S_{k+1}|\mathcal{L}_k] = E_\infty[S_k + \{W_{k+1} - C\}|\mathcal{L}_k]$$

$$= E_\infty[S_k|\mathcal{L}_k] + E_\infty[\{W_{k+1} - C\}|\mathcal{L}_k]$$

$$= S_k + E_\infty[W_{k+1}] - C = S_k \quad \text{a.s.},$$

and hence $\{(S_k,\mathcal{L}_k)\}_{k=1}^{\infty}$ is a martingale. Since

$$E_\infty[|S_{k+1} - S_k| \;|\mathcal{L}_k] = E_\infty[|W_{k+1} - C|] \leq E_\infty[|S_1|]$$

for $k \leq n^*(X)$ a.s.(P_∞), then, according to Doob [1], (Theorem 2.2, p. 302), the sequence $\{S_1, S_{n^*(X)}\}$ obtained by optional sampling is a martingale with

$$E_\infty[S_1] = E_\infty[S_{n^*(X)}] \; .$$

It follows that

$$0 = E_\infty[S_{n^*(X)}] = E_\infty\left[\sum_{j=1}^{n^*(X)} W_j\right] - E_\infty[W_1]\; E_\infty(n^*),$$

or

$$E_\infty\left[\sum_{j=1}^{n^*(X)} \log \frac{f_1(x_j)}{f_2(x_j)}\right] = E_\infty(n^*) \int_{-\infty}^{\infty} \log\left[\frac{f_1(x)}{f_2(x)}\right] f_1(x)\; dx,$$

Remark: Kullback [7] has shown that if $H_i : h(v) = h_i(v)$, $i = 1,2$, are simple hypotheses on the density function $h(v)$ of a random variable V, and we define

$$I(1,2;0.) = \int_{-\infty}^{\infty} \log \left[\frac{h_1(v)}{h_2(v)}\right] h_1(v)dv$$

to be the information in a sample of size 1 for discrimination between H_1 and H_2, then the information $I(1,2;0_N)$ in a random sample of size N is $NI(1,2;0_1)$. Theorem 3.2 is thus an extension of Kullbacks result to Wald's sequential probability ratio test. It is easy to obtain Kullback's result from the infinite series of Theorem 3.1 in the following manner.

Let ψ_0^k, ψ_1^k, ψ_2^k be the indicator functions of (a_k, b_k), $z \leq -b_k$, $z \leq a_k$ respectively, where $a_k = b_k = +\infty$ for $1 \leq k < N$ and $a_N = -b_N$ for $k \geq N$.

In this case, from Theorem 3.1,

$$\begin{aligned}
I(\mathcal{A}) &= \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} \log \left\{\prod_{j=1}^{N} \frac{f_1(z_j-z_{j-1})}{f_2(z_j-z_{j-1})}\right\} f_1(z_N-z_{N-1})\cdots f_1(z_1)dz_1\cdots dz_N \\
&= \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} \log \left\{\prod_{j=1}^{N} \frac{f_1(x_j)}{f_2(x_j)}\right\} \prod_{j=1}^{N} f_1(x_j)dx_j \\
&= \sum_{j=1}^{N} \int_{-\infty}^{\infty} \log \left\{\frac{f_1(x_1)}{f_2(x_1)}\right\} f_1(x_1)dx_1 \\
&= N \int_{-\infty}^{\infty} \log \left\{\frac{f_1(x_1)}{f_2(x_1)}\right\} f_1(x_1)dx_1,
\end{aligned}$$

where $z_0 = 0$.

We next obtain a representation for $I(\mathcal{C})$ that is similar to the representation in Theorem 3.1.

Theorem 3.3. Let $\mathcal{C}$ be the Borel field given in Definition 3.1. If $P_{1\mathcal{C}} \ll P_{2\mathcal{C}}$, then

$$I(\mathcal{C}) = \sum_{k=1}^{\infty} \left(\sum_{i=1}^{2} \int_{-\infty}^{\infty} \log \left[\frac{h_{1k}(z_k)}{h_{2k}(z_k)} \right] \psi_i^k(z_k)\, h_{1k}(z_k) dz_k \right),$$

where

$$(3.14) \quad h_{\ell k}(z_k) = \begin{cases} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \left[\prod_{j=1}^{k-1} \psi_0^j(z_j) \right] f_\ell(z_k - z_{k-1}) \cdots f_\ell(z_1) dz_1 \cdots dz_{k-1}, \\ \text{if } k > 1 \\ f_\ell(z_1), \quad \text{if } k = 1 \end{cases}$$

$\ell = 1,2.$

Proof: Since the proof of this theorem resembles very closely that of Theorem 3.1, most of the details will be omitted. Let $\mathcal{B}_1$ denote the Borel sets of the line R. As in the proof of Theorem 3.1, there exists a regular sequence $\{\mathcal{E}_s\}_{s=1}^{\infty} = \{\{E_s^j\}_{j=1}^{\ell_s}\}_{s=1}^{\infty}$ of finite $\mathcal{B}_1$-measurable partitions of R such that $\mathcal{B}_1$ is generated by $\bigcup_{s=1}^{\infty} \mathcal{L}_s$, where $\mathcal{L}_s$ is the field consisting of all finite unions of sets of $\mathcal{E}_s$.

For each $E_s^j \in \mathcal{E}_s$, let

$$G_{sk}^{ij} = \{\omega \in A_{ki} | z_k \in E_s^j\}, \; i = 1,2.$$

For each $s = 1,2,\ldots$, $k = 1,2,\ldots$, $\{G_{sk}^{ij}\}_{j=1}^{\ell_s}$ is a finite partition of A_{ki};

$$\mathcal{G}_s = \{A_{s0}, \{G_{s1}^{1j}\}_{j=1}^{\ell_s}, \{G_{s1}^{2j}\}_{j=1}^{\ell_s}, \ldots, \{G_{ss}^{1j}\}_{j=1}^{\ell_s}, \{G_{ss}^{2j}\}_{j=1}^{\ell_s}\}$$

is a finite $\mathcal{C}$-measurable partition of Ω; $\{\mathcal{G}_s\}_{s=1}^{\infty}$ is a regular sequence; and if $\mathcal{K}_s$ is the field generated by $\mathcal{G}_s$, then the Borel field generated by $\bigcup_{s=1}^{\infty} \mathcal{K}_s$ is $\mathcal{C}$. Proceeding in the same manner as in the proof of Theorem 3.1, we need only show

$$(3.15) \qquad \lim_{s \to \infty} \sum_{j=1}^{\ell_s} P_{1\mathcal{C}}(G_{sk}^{ij}) \log \left[P_{1\mathcal{C}}(G_{sk}^{ij}) \,/\, P_{2\mathcal{C}}(G_{sk}^{ij}) \right]$$

$$= \int_{-\infty}^{\infty} \log \left[\frac{h_{1k}(z_k)}{h_{2k}(z_k)} \right] \psi_i^k(z_k) h_1(z_k) dz_k$$

for arbitrary k, i, where $h_{1k}(z_k)$ and $h_{2k}(z_k)$ are given by (3.14). Fix k, i, and for each $F \in \mathcal{B}_1$, let

$$(3.16) \qquad \lambda_{\ell ki}(F) = \int_F \psi_i^k(z_k) h_{\ell k}(z_k) dz_k, \quad \ell = 1,2 \ .$$

From Fubini's Theorem and the definition of $P_{\ell\lambda}$, the set functions $\lambda_{\ell ki}$, $\ell = 1,2$, are finite measures on $\mathcal{B}_1$, and

$$(3.17) \qquad \lambda_{\ell ki}(E_s^j) = P_{\ell\mathcal{C}}(G_{sk}^{ij}) \text{ for each } j = 1,2,\ldots,\ell_s;\ s = 1,2,\ldots.$$

The absolute continuity of $P_{1\mathcal{C}}$ with respect to $P_{2\mathcal{C}}$ and the definition of $\lambda_{\ell ki}$, $\ell = 1,2$, give $\lambda_{1ki} \ll \lambda_{2ki} \ll m_1$ where m_1 denotes Lebesque measure on $\mathcal{B}_1$. Since the functions $\psi_i^k(z_k)h_{\ell k}(z_k)$ in (3.16) are Borel measurable, it follows precisely as in the proof of Theorem 3.1 that

$$\lim_{s \to \infty} \sum_{j=1}^{\ell_s} \lambda_{1ki}(E_s^j) \log \left[\frac{\lambda_{1ki}(E_s^j)}{\lambda_{2ki}(E_s^j)}\right]$$

$$= \int_{-\infty}^{\infty} \log \left[\frac{h_{1k}(z_k)}{h_{2k}(z_k)}\right] \psi_i^k(z_k)h_{1k}(z_k)dz_k.$$

The result now follows from (3.17)

We conclude this chapter with

Theorem 3.4. Let $\mathcal{D}$ be the Borel field given in Definition 3.1. If $P_{1\mathcal{D}} \ll P_{2\mathcal{D}}$, then

$$I(\mathcal{D}) = \sum_{k=1}^{\infty} \left(\sum_{i=1}^{2} P_{1\mathcal{D}}(A_{ki}) \log \left[P_{1\mathcal{D}}(A_{ki}) \,/\, P_{2\mathcal{D}}(A_{ki})\right]\right).$$

Proof: In Lemma 3.1 let $(\mathcal{X}, \mathcal{J}, Q_\ell)$ be $(\Omega, \mathcal{D}, P_{\ell\mathcal{D}})$, $\ell = 1,2$, and take

$$\{\mathcal{M}_s\}_{s=1}^{\infty} = \{\{A_{s0}, A_{1j}, A_{2j}\}_{j=1}^{s}\}_{s=1}^{\infty} .$$

From the choice of $\{\mathcal{M}_s\}_{s=1}^{\infty}$, it follows directly that

$$\mathcal{J}_\infty = \mathcal{D}$$

and

$$X_s(\omega) = \sum_{j=1}^{s} \left(\sum_{i=1}^{2} P_{1\mathcal{D}}(A_{ij}) \,/\, P_{2\mathcal{D}}(A_{ij}) \right) + P_{1\mathcal{D}}(A_{s0}) \,/\, P_2(A_{s0})$$

for each $s = 1,2,\ldots.$

Also,

$$\lim_{s \to \infty} X(\omega) = X_{\mathcal{D}} \quad \text{a.s.}(P_{2\mathcal{D}}) ,$$

where $X_{\mathcal{D}}$ is the R-N derivative of $P_{1\mathcal{D}}$ with respect to $P_{2\mathcal{D}}$.

Clearly, $E_{2\mathcal{D}}(|X_s \log X_s|) < \infty$ for each $s = 1,2,\ldots,$ and so Lemma 3.2 gives

$$\lim_{s \to \infty} \int_\Omega X_s \log X_s \, dP_{2\mathcal{D}} = \int_\Omega X_{\mathcal{D}} \log X_{\mathcal{D}} \, dP_{2\mathcal{D}} .$$

Now since

$$\lim_{s \to \infty} \int_{\Omega} X_s \log X_s \, dP_{2\vartheta} = \lim_{s \to \infty} \left\{ \sum_{j=1}^{s} \left(\sum_{i=1}^{2} P_{1\vartheta}(A_{ij}) \log \left[\frac{P_{1\vartheta}(A_{sO})}{P_{2\vartheta}(A_{sO})} \right] \right) \right.$$

$$\left. + P_{1\vartheta}(A_{sO}) \log \left[\frac{P_{1\vartheta}(A_{sO})}{P_{2\vartheta}(A_{sO})} \right] \right\} ,$$

and

$$\lim_{s \to \infty} P_{1\vartheta}(A_{sO}) \log \left[\frac{P_{1\vartheta}(A_{sO})}{P_{2\vartheta}(A_{sO})} \right] = 0$$

from Assumption I in Chapter II, the theorem is established.

CHAPTER IV

RELATIONSHIPS BETWEEN THE INFORMATION IN THREE BOREL FIELDS

It follows readily from the definitions of the Borel fields $\mathcal{A}$, $\mathcal{C}$, and $\mathcal{B}$ described in Definition 3.1 and from (1.2) that $I(\mathcal{A}) \geq I(\mathcal{C}) \geq I(\mathcal{B})$ with equality between any pair $\{I(\mathcal{G}), I(\mathcal{L})\}$ of $I(\mathcal{A})$, $I(\mathcal{C})$, and $I(\mathcal{B})$ if and only if

$$\sup_{\mathcal{S} \subset \mathcal{G}} I(\mathcal{S}) = \sup_{K \subset \mathcal{L}} I(K) , \tag{4.1}$$

where the supremum is taken over all finite subfields $\mathcal{S}$ contained in $\mathcal{G}$ and K contained in $\mathcal{L}$. However, it is not clear from the abstract definition (1.2) what restrictions the equality (4.1) imposes on the conditional densities of the cumulative sums and the probability decision functions used to define the probability measures P_ℓ, $\ell = 1,2$, in Chapter II. We now use the rperesentations for $I(\mathcal{A})$, $I(\mathcal{C})$, and $I(\mathcal{B})$ given in Theorems 3.1, 3.3, and 3.4 respectively to obtain necessary and sufficient conditions in terms of the conditional densities of the cumulative sums and the decision functions for $I(\mathcal{A}) = I(\mathcal{B})$, $I(\mathcal{A}) = I(\mathcal{C})$ when $I(\mathcal{A}) < \infty$, and $I(\mathcal{C}) = I(\mathcal{B})$ when $I(\mathcal{C}) < \infty$.

The primary result which will be used to derive the conditions above is the following lemma due to Kullback [7], and which is in essence Jensen's inequality.

Lemma 4.1. Let $(\mathcal{L},\mathcal{J})$ be a measurable space and let $\mu_1 \not\equiv 0$, μ_2 be two finite measures on $\mathcal{J}$ with $\mu_1 \ll \mu_2$. If λ is any measure on $\mathcal{J}$ with $\mu_2 \ll \lambda$ and q_1 and q_2 are the R-N derivatives $\frac{d\mu_1}{d\lambda}$ and $\frac{d\mu_2}{d\lambda}$ respectively, then

$$\int_{\mathcal{L}} \log\left[\frac{q_1}{q_2}\right] q_1 d\lambda \geq \mu_1(\mathcal{L}) \log \frac{\mu_1(\mathcal{L})}{\mu_2(\mathcal{L})} ,$$

with equality if and only if

$$\frac{q_1}{\mu_1(\mathcal{L})} = \frac{q_2}{\mu_2(\mathcal{L})} \quad \text{a.e.}(\lambda) .$$

Theorem 4.1. If $I(\mathcal{A}) < \infty$, then $I(\mathcal{A}) = I(\mathcal{B})$ if and only if

$$\frac{\left[\prod_{j=1}^{k-1} \psi_0^j(z_j)\right]\psi_i^k(z_k)\prod_{j=1}^{k} f_1(z_j-z_{j-1})}{P_1(A_{ki})} = \frac{\left[\prod_{j=1}^{k-1} \psi_0^j(z_j)\right]\psi_i^k(z_k)\prod_{j=1}^{k} f_2(z_j-z_{j-1})}{P_2(A_{ki})} \quad \text{a.e.}(m_k)$$

for any pair of indices (k,i) such that $P_1(Z_{ki}) > 0$, $i = 1,2$, $k = 1,2,\ldots$, where $z_0 = 0$, and $\prod_{j=1}^{k-1} \psi_0^j(z_j) = 1$ if $k = 1$.

Proof: For each $i = 1,2$, $k = 1,2,\ldots$ let

$$C_{ki} = \Big(\int_R \cdots \int_R \log\left\{\prod_{j=1}^{k} \frac{f_1(z_j-z_{j-1})}{f_2(z_j-z_{j-1})}\right\}\left[\prod_{j=1}^{k-1} \psi_0^j(z_j)\right]\psi_i^k(z_k)$$

$$\cdot\, f_1(z_k-z_{k-1})\ldots f_1(z_1)dz_1\ldots dz_k) - P_1(A_{ki})\log \frac{P_1(A_{ki})}{P_2(A_{ki})}\Big) ,$$

where $z_0 = 0$ and $\prod_{j=1}^{k-1} \psi_0^j(z_j) = 1$ if $k = 1$.

It follows at once from Theorems 3.1, 3.4, and the assumed finiteness of $I(\mathcal{A})$ that

$$(4.2) \qquad I(\mathcal{A}) - I(\mathcal{B}) = \sum_{k=1}^{\infty}\left(\sum_{i=1}^{2} C_{ki}\right) .$$

Fix k,i arbitrarily and assume $P_1(A_{ki}) > 0$. Since, (as was indicated in the proof of Theorem 3.1), the set functions defined by

$$\mu_{\ell ki}(E) = \int_E \cdots \int_E \left[\prod_{j=1}^{k-1} \psi_0^j(z_j)\right] \psi_i^k(z_k) f_\ell(z_k - z_{k-1}) \cdots f_\ell(z_1) dz_1 \cdots dz_k ,$$

for each $E \in \mathcal{B}_k$, are finite measures on $\mathcal{B}_k$ with $\mu_{\ell ki}(R_k) = P_{\ell ki}(A_{ki})$, $\ell = 1,2$, and $\mu_{1ki} \ll \mu_{2ki} \ll m_k$, it is easily seen from Lemma 4.1 that

$$(4.3) \qquad C_{ki} \geq 0 ,$$

with equality if and only if

$$\frac{\left[\prod_{j=1}^{k-1} \psi_0^j(z_j)\right] \psi_i^k(z_k) \prod_{j=1}^{k} f_1(z_j - z_{j-1})}{P_1(A_{ki})} = \frac{\left[\prod_{j=1}^{k} \psi_0^j(z_j)\right] \psi_i^k(z_k) \prod_{j=1}^{k} f_2(z_j - z_{j-1})}{P_2(A_{ki})} \text{ a.e.}(m_k)$$

If $P_1(A_{ki}) = 0$ the equality in (4.3) is trivially satisfied. It now follows from the inequality (4.3) and equation (4.2) that $I(\mathcal{A}) = I(\mathcal{C})$ if and only if $C_{ki} = 0$ for each $i = 1,2$, $k = 1,2,\ldots$. The conditions for $C_{ki} = 0$ now establish the theorem.

Theorem 4.2. For each $k = 1,2,\ldots$, let h_{1k}, h_{2k} be the functions defined in (3.14). If $I(\mathcal{C}) < \infty$, then $I(\mathcal{C}) = I(\mathcal{D})$ if and only if

$$\frac{\psi_i^k(z_k)h_{1k}(z_k)}{P_1(A_{ki})} = \frac{\psi_i^k(z_k)h_{2k}(z_k)}{P_2(A_{ki})} \quad \text{a.e.}(m_1)$$

for each pair of indices (k,i) such that $P_1(A_{ki}) > 0$, $i = 1,2$, $k = 1,2,\ldots$.

Proof: For each $\ell = 1,2$, $i = 1,2$, $k = 1,2,\ldots$, let

$$\lambda_{\ell ki}(E) = \int_E \psi_i^k(z_k)h_\ell(z_k)dz_k \ .$$

for each $E \in \mathcal{B}_1$.

From the proof of Theorem 3.1, λ_{1ki}, λ_{2ki} are finite measures on $\mathcal{B}_1$ with the properties

$$\lambda_{1ki} \ll \lambda_{2ki} \ll m_1$$

and

$$\lambda_{1ki}(R) = P_1(A_{ki}), \ \lambda_{2ki}(R) = P_2(A_{ki}).$$

The proof now follows along the same lines as the proof of Theorem 4.1 from Theorems 3.1, 3.4, and the finiteness of $I(\mathcal{C})$.

Theorem 4.3. For each $k = 2,3,\ldots$, let h_{1k}, h_{2k} be the functions defined in (3.14) and let $G_k = \{\eta \in R_1 | h_{1k}(\eta) \neq 0\}$. If $I(\mathcal{A}) < \infty$,

then $I(\mathcal{A}) = I(\mathcal{C})$ if and only if

$$\left[\prod_{j=1}^{k-1} \psi_o^j(z_j)\right]\psi_i^k(\xi) f_1(\xi - z_{k-1}) f_1(z_{k-1} - z_{k-2}) \ldots f_1(z_1) \;/\; h_{1k}(\xi)$$

$$= \left[\prod_{j=1}^{k-1} \psi_0^j(z_j)\right]\psi_i^k(\xi) f_2(\xi - z_{k-1}) f_2(z_{k-1} - z_{k-2}) \ldots f_1(z_1) \;/\; h_{2k}(\xi) \text{ a.e.}(m_{k-1})$$

for each $\xi \in G_k$, except perhaps on a subset of G_k of Lebesque (m_1) measure zero, $i = 1,2$, $k = 2,3,\ldots$.

Proof: For each $i = 1,2$, $k = 2,3,\ldots$, let

$$q_{ik}(\xi) = \psi_i^k(\xi) \int_R \cdots \int_R \log \left\{\frac{f_1(\xi - z_{k-1}) \prod_{j=1}^{k-1} f_1(z_j - z_{j-1})}{f_2(\xi - z_{k-1}) \prod_{j=1}^{k-1} f_2(z_j - z_{j-1})}\right\}\left[\prod_{j=1}^{k-1} \psi_0^j(z_j)\right]$$

$$\cdot\; f_1(\xi - z_{k-1}) f_1(z_{k-1} - z_{k-2}) \ldots f_1(z_1) dz_1 \ldots dz_k$$

and

$$r_{ik}(\xi) = \psi_i^k(\xi)\; h_{1k}(\xi) \log\left[h_{1k}(\xi) \;/\; h_{2k}(\xi)\right]$$

for each $\xi \in R$, where $z_0 = 0$.

It follows from the finiteness of $I(\mathcal{A})$ and from Fubini's theorem that the functions $q_{ik}(\xi)$ and $r_{ik}(\xi)$ are integrable and hence finite except on a set E_{ik} of Lebesque (m_1) measure zero for each $i = 1,2$, $k = 2,3,\ldots$. We thus obtain from Lemma 4.1 the inequality

$$(4.4)\quad \left\{\begin{array}{l} q_{ik}(\xi) \geq r_{ik}(\xi) \text{ for each } \xi \in R - E_{ik}, \text{ with} \\ \text{equality for any point } \xi \in R - E_{ik} \text{ such that'} \\ h_{1k}(\xi) \neq 0 \text{ if and only if} \\ \psi_i^k(\xi)\left[\prod_{j=1}^{k-1} \psi_0^j(z_j)\right] f_1(\psi - z_{k-1}) f_1(z_{k-1} - z_{k-2}) \ldots f_1(z_1) \;/\; h_{1k}(\xi) \\ = \psi_i^k(\xi)\left[\prod_{j=1}^{k-1} \psi_0^j(z_j)\right] f_2(\xi - z_{k-1}) f_2(z_{k-1} - z_{k-2}) \ldots f_2(z_1) \;/\; h_{2k}(\xi) \\ \text{a.e.}(m_{k-1}) \end{array}\right\}$$

for each $i = 1,2$, $k = 2,3,\ldots$.

Now $h_{1k}(\xi) = 0$ implies $q_{ik}(\xi) = r_{ik}(\xi) = 0$ and so from (4.4) we have

$$\int_R q_{ik}(\xi)d\xi - \int r_{ik}(\xi)d\xi \geq 0,$$

with equality if and only if equality holds in (4.4) a.e.(m_1) on G_k. Hence it follows from Theorems 3.1, 3.3, and the finiteness of $I(\mathcal{A})$ that

$$I(\mathcal{A}) - I(\mathcal{C})$$

$$= \sum_{k=2}^{\infty}\left(\sum_{i=1}^{2} \int_R \cdots \int_R \left[\log\left\{\prod_{j=1}^{k} \frac{f_1(z_j - z_{j-1})}{f_2(z_j - z_{j-1})}\right\}\left[\prod_{j=1}^{k-1} \psi_i^k(z_j)\right]\psi_i^k(z_k) f_1(z_k - z_{k-1}) \ldots f_1(z_1)\right.\right.$$

$$\left.\left. \cdot\, dz_1 \ldots dz_k\right] - \int_R \log \frac{h_{1k}(z_k)}{h_{2k}(z_k)}\, \psi_i^k(z_k) h_1(z_k) dz_k\right)$$

$$= \sum_{k=1}^{\infty}\left(\sum_{i=1}^{2}\left[\int_R q_{ik}(\xi)d\xi - \int_R r_{ik}(\xi)d\xi\right]\right) \geq 0,$$

with equality if and only if

$$\int_R q_{ik}(\xi)d\xi - \int_R r_{ik}(\xi)d\xi = 0, \tag{4.5}$$

for each $i = 1,2$, $k = 2,3,\ldots$.

The conditions for equation (4.5) to hold now establish the theorem.

Corollary 4.1. If the common density functions $f_i(x)$ of the random variables $X_1, X_2, \ldots, X_N, \ldots$ under the hypotheses H_i, $i = 1,2$, have the property that $f_1(x) = e^x f_2(x)$ for each $x \in (-\infty,\infty)$, then $I(\mathcal{A}) = I(\mathcal{C})$, whether $I(\mathcal{A})$ is finite or infinite.

Proof: Since $f_1(x) = e^x f_2(x)$, then

$$\prod_{j=1}^{k} f_1(z_j - z_{j-1}) = e^{z_k} \prod_{j=1}^{k} f_2(z_k - z_{k-1})$$

and

$$h_{1k}(z_k) = e^{z_k} \int_R \cdots \int_R \left[\prod_{j=1}^{k-1} \psi_0^j(z_j)\right] f_2(z_k - z_{k-1}) \cdots f_1(z_1) dz_1 \cdots dz_{k-1}$$

$$= e^{z_k} h_{2k}(z_k)$$

for each $k = 1,2,\ldots$, where $z_0 = 0$, and $\prod_{j=1}^{k-1} \psi_0^j(z_j) = 1$ if $k = 1$.

Thus

$$I(\mathcal{A}) = \sum_{k=1}^{\infty}\Big(\sum_{i=1}^{\infty}\int_R\cdots\int_R \log\Big\{\prod_{j=1}^{k}\frac{f_1(z_j-z_{j-1})}{f_2(z_j-z_{j-1})}\Big\}\prod_{j=1}^{k-1}\psi_o(z_j)\psi_i^k(z_k)f_1(z_k-z_{k-1})\cdots f_1(z_1)$$

$$\cdot\, dz_1\cdots dz_k\Big)$$

$$= \sum_{k=1}^{\infty}\Big(\sum_{i=1}^{2}\int_R z_k\psi_i^k(z_k)dz_k\int_R\cdots\int_R\prod_{j=1}^{k-1}\psi_o(z_j)f_1(z_k-z_{k-1})\cdots f_1(z_1)dz_1\cdots dz_{k-1}$$

$$= \sum_{k=1}^{\infty}\Big(\sum_{i=1}^{2}\int_R \log\Big[\frac{h_{1k}(z_k)}{h_{2k}(z_k)}\Big]\psi_i^k(z_k)h_1(z_k)dz_k$$

$$= I(\mathcal{C})$$

from Theorems 3.1, 3.3 and Fubini's Theorem.

Remark: Let the probability density function $g(u;\theta)$ of a random variable U be given by

$$g(u:\theta) = e^{\theta u+r(u)+q(\theta)} \tag{4.6}$$

for u in some subset of the real line, independent of θ. Suppose that S_w is a Wald sequential probability ratio test of the simple hypotheses $H_i:g(u;\theta) = g(u;\theta_i)$, $i = 1,2$, on the parameter θ of g. It is easy to show that $X = \log\frac{g(u;\theta_1)}{g(u;\theta_2)}$ has density functions $f_i(x)$ under H_i, $i = 1,2$, with the property $f_1(x) = e^x f_2(x)$; and thus $I(\mathcal{A}) = I(\mathcal{C})$ for the test S_w from Corollary 4.1. Density functions of the form (4.6) include an important class used in applied statistics. Particular examples are

the normal, Poisson and binomial densities. In the binomial case, for example, the density (probability function with respect to counting measure) is

$$\binom{m}{x}p^x(1-p)^{m-x} = e^{x\log(\frac{p}{1-p})+\log\binom{m}{x}+m\log(1-p)}, \quad 0 < p < 1.$$

It might appear that the hypothesis of Corollary 4.1 not only guarantees $I(\mathcal{A}) = I(\mathcal{C})$ but also $I(\mathcal{A}) = I(\mathcal{C}) = I(\mathcal{D})$. The following example shows that this is not the case.

Example 4.1. Let $g(u)$ be the density function of a random variable U, and let S be a sequential probability ratio test of $H_1: g(u) = g_1(u)$ against $H_2: g(u) = g_2(u)$, where

$$g_1(u) = \begin{cases} 2u, & 0 \le u \le 1 \\ 0, & \text{elsewhere} \end{cases}, \quad g_2(u) = \begin{cases} 1, & 0 \le u \le 1 \\ 0, & \text{elsewhere} \end{cases}.$$

Now $X = \log\frac{g_1(u)}{g_2(u)} = \log 2u$ has density functions

$$f_1(x) = \begin{cases} \frac{1}{2}e^{2x}, & -\infty \le x \le \log 2 \\ 0, & \text{elsewhere} \end{cases}, \quad f_2(x) = \begin{cases} \frac{1}{2}e^{x}, & -\infty \le x \le \log 2 \\ 0, & \text{elsewhere} \end{cases}$$

under H_1, H_2 respectively, so, clearly, $f_1(x) = e^x f_2(x)$ for each $x \in (-\infty,\infty)$, and the hypothesis of Corollary 4.1 is satisfied. For ease of computation, we shall assume in the following that the terminating

barriers, -b and a, associated with S_w satisfy the inequality $a + b \leq \log 2$. From Theorem 3.2, we have

$$I(\mathcal{A}) = E_1(n) \int_{-\infty}^{\infty} \log \left[\frac{f_1(x)}{f_2(x)}\right] f_1(x)dx.$$

Here
$$E_1(n) = 1 + \sum_{N=1}^{\infty} P_1(n > N)$$

$$= 1 + \sum_{N=1}^{\infty} \int_{-b}^{a} \cdots \int_{-b}^{a} \prod_{j=1}^{N} f_1(z_j - z_{j-1})dz_j$$

$$= 1 + \sum_{N=1}^{\infty} \int_{-b}^{a} \cdots \int_{-b}^{a} \frac{1}{2^N} e^{2z_N} dz_1 \cdots dz_N$$

$$= 1 + \left\{\left[\sum_{N=1}^{\infty} \left(\frac{a+b}{2}\right)^{N-1}\right]\left[\frac{e^{2a} - e^{-2b}}{4}\right]\right\},$$

and
$$\int_{-\infty}^{\infty} \log\left[\frac{f_1(x)}{f_2(x)}\right] f_1(x)dx = \frac{1}{2}\int_{-\infty}^{\log 2} xe^x dx = \frac{\log 4 - 1}{2} .$$

So,
$$I(\mathcal{A}) = \left[1 + \frac{e^{2a} - e^{-2b}}{4(1 - \frac{a+b}{2})}\right]\left[\frac{\log 4 - 1}{2}\right] .$$

From Theorem 3.4, $I(\mathcal{D}) = \sum_{N=1}^{\infty}\left(\sum_{i=1}^{2} P_1(A_{Ni}) \log\left[\frac{P_1(A_{Ni})}{P_2(A_{Ni})}\right]\right)$, and it is easy

to show that here we have

$$P_1(A_{Ni}) = \left\{ \begin{array}{l} e^{-2b}/4 \text{ if } N = 1,\ i = 1. \\ (4-e^{2a})/4 \text{ if } N = 1,\ i = 2. \\ \left(\frac{a+b}{2}\right)^{N-2} \left(\frac{a+b}{2}\right) \frac{e^{-2b}}{4} \text{ if } N > 1,\ i = 1. \\ \left(\frac{a+b}{2}\right)^{N-2} \left[\frac{2(e^{2a}-e^{-2b})-e^{2a}(a+b)}{8}\right] \text{ if } N > 1,\ i = 2. \end{array} \right\}$$

Also,

$$P_2(A_{Ni}) = \left\{ \begin{array}{l} e^{-b}/2 \text{ if } N = 1,\ i = 1. \\ (2-e^{a})/2 \text{ if } N = 1,\ i = 2. \\ \left(\frac{a+b}{2}\right)^{N-2} \left(\frac{a+b}{2}\right)(e^{-b}/2) \text{ if } N > 1,\ i = 1. \\ \left(\frac{a+b}{2}\right)^{N-2} \frac{2(e^{a}-e^{-b})-e^{a}(a+b)}{4} \text{ if } N > 1,\ i = 1. \end{array} \right\}$$

Thus, in this case,

$$I(\mathcal{D}) = \frac{e^{-2b}}{4} \log \frac{e^{-b}}{2} + \frac{4-e^{2a}}{4} \log \frac{4-e^{2a}}{2(2-e^{a})}$$

$$+ \left(\sum_{N=2}^{\infty} \left(\frac{a+b}{2}\right)^{N-2}\right) \left\{ \left(\frac{a+b}{2}\right) \frac{e^{-2b}}{4} \log \frac{e^{-b}}{2} \right.$$

$$\left. + \left(\frac{2(e^{2a}-e^{-2b})-e^{2a}(a+b)}{8}\right) \log \left(\frac{2(e^{2a}-e^{-2b})-e^{2a}(a+b)}{2[2(e^{a}-e^{-b})-e^{a}(a+b)]}\right) \right\} .$$

Evaluating $I(\mathcal{A})$ and $I(\mathcal{B})$ for $a = b = \frac{1}{2} \log 2$, we find that

$$I(\mathcal{A}) = .3039 \text{ and } I(\mathcal{B}) = .2529$$

to four significant digits, and hence

$$I(\mathcal{A}) > I(\mathcal{B}) \ .$$

CHAPTER V

CONCLUDING REMARKS

In this paper we considered the information in a sequential test S of simple hypotheses which is determined by cumulative sums of independent random variables; and we derived analytical representations for the information $I(\mathcal{A})$ in the decision process, the information $I(\mathcal{C})$ in the sequential decision variables, and the information $I(\mathcal{B})$ in the stopping rule. These results are not limited to this particular type of sequential test. The probability spaces defined in Chapter II can easily be modified so that representations similar to those given in Theorems 3.1, 3.3, and 3.4 can be found for more general types of sequential tests of simple hypotheses.

In the particular case of a Wald sequential probability ratio test we showed that the information $I(\mathcal{A})$ is equal to $E_1(n)\, E_1[\log \frac{g_1(u)}{g_2(u)}]$, where n is the termination sample size and $g_1(u)$ and $g_2(u)$ are hypothesized density functions of an observable random variable U. It should be clear that with minor modifications of the proof of Theorem 3.2 the same result is also valid for the generalized sequential probability ratio test. The writer conjectures that $I(\mathcal{A}) = E_1(n)E_1[\log \frac{f_1(x)}{f_2(x)}]$ even for the randomized test and this is presently being investigated.

It should also be pointed out that the representations given by Theorems 3.1, 3.3, and 3.4 can be obtained by determining the explicit

form of the R-N derivatives $\frac{dP_{1\mathcal{A}}}{dP_{2\mathcal{A}}}$, $\frac{dP_{1\mathcal{C}}}{dP_{2\mathcal{C}}}$, and $\frac{dP_{1\mathcal{B}}}{dP_{2\mathcal{B}}}$ respectively. However, this approach becomes quite difficult in generalization.

Equality between any pair of $I(\mathcal{A})$, $I(\mathcal{C})$, and $I(\mathcal{B})$ is closely related to the concept of sufficiency of a Borel field, (see [3] for the definition of sufficiency). It is known [6] that a Borel field $\mathcal{J}_1 \subset \mathcal{J}$ is sufficient for the set of probability measures $\{\mu_{1\mathcal{J}}, \mu_{2\mathcal{J}}\}$ if and only if $I(\mathcal{J}_1) = I(\mathcal{J})$. From the viewpoint, Theorems 4.1, 4.2, and 4.3 then give necessary and sufficient conditions for $\mathcal{B}$ to be sufficient conditions for $\mathcal{B}$ to be sufficient for $\{P_{1\mathcal{A}}, P_{2\mathcal{A}}\}$, $\mathcal{B}$ to be sufficient for $\{P_{1\mathcal{C}}, P_{2\mathcal{C}}\}$, and $\mathcal{C}$ to be sufficient for $\{P_{1\mathcal{A}}, P_{2\mathcal{A}}\}$ respectively. It is hoped then that the results of Chapter IV can be used as a basis for the study of sufficiency in sequential tests in terms of the technical sense of information.

By way of further research, in addition to that above, we mention the following topics:

(1) investigate the information in sequential tests of simple hypotheses when the basic random variables are not independent;

(2) investigate the information in multiple decision problems;

(3) use the information measures as a means of comparing sequential tests.

BIBLIOGRAPHY

1. J. L. Doob, Stochastic Processes, John Wiley and Sons, Inc., New York, 1953.

2. P. R. Halmos, Measure Theory, D. Van Nostrand Co., New York, 1950.

3. P. R. Halmos and L. J. Savage, "Application of the Radon-Nikodym Theorem to the Theory of Sufficient Statistics," Ann. of Math. Stat., 20 (1949), 225-241.

4. G. Kallianpur, "Information Contained in a σ-Field," Contributions to Probability and Statistics, Stanford University Press, 1960, 265-271.

5. J. Keifer and L. Weiss, "Some Properties of Generalized Sequential Probability Ratio Tests," Ann. of Math. Stat., 28 (1957), 57-74.

6. S. Kullback and R. A. Leibler, "On Information and Sufficiency," Ann. of Math. Stat., 22 (1951), 79-86.

7. S. Kullback, Information Theory and Statistics, John Wiley and Sons, Inc., New York, 1959.

8. M. Loeve, Probability Theory, D. Van Nostrand Co., New York, 1955.

9. H. L. Royden, Real Analysis, The MacMillan Co., New York, 1963.

10. L. J. Savage, The Foundations of Statistics, John Wiley and Sons, New York, 1954.

11. M. P. Schutzenberger, "On Some Measures of Information Used in Statistics," In Colin Cherry (ed.), Symposium on Information Theory, Academic Press, 1956.

12. A. Wald, Sequential Analysis, J. Wiley and Sons, Inc., New York, 1947.

13. L. Weiss, "Testing One Simple Hypothesis Against Another," Ann. of Math. Stat., 24, (1953), 273-281.

ORNL-4208
UC-32 — Mathematics and Computers

INTERNAL DISTRIBUTION

1. Biology Library
2-4. Central Research Library
5-6. ORNL — Y-12 Technical Library
Document Reference Section
7-26. Laboratory Records Department
27. Laboratory Records, ORNL R.C.
28. R. T. Boughner
29. R. R. Coveyou
30. J. W. Curlin
31. M. H. Davis
32. Manuel Feliciano
33. D. A. Gardiner
34. D. G. Gosslee
35. T. L. Hebble
36. A. S. Householder
37. M. C. Hull
38. W. H. Jordan
39. M. A. Kastenbaum
40. C. E. Larson
41. E. Leach
42. H. G. MacPherson
43. J. S. Olson
44. M. J. Skinner
45. W. J. Stelzman
46. D. A. Sundberg
47-96. R. L. Tittle (Y-12)
97. V. R. Rao Uppuluri
98. D. R. Vondy
99. A. M. Weinberg

EXTERNAL DISTRIBUTION

100. J. H. Barrett, Mathematics Department, University of Tennessee
101. J. A. Swartout, Union Carbide Corporation, New York
102. Laboratory and University Division, AEC, ORO
103-390. Given distribution as shown in TID-4500 under Mathematics and Computers category (25 copies — CFSTI)