THE SELECTION OF AN OPTIMUM SAMPLING PARTITION

By

ARNOLD F. GOODMAN

TECHNICAL REPORT NO. 72
August 3, 1961

PREPARED FOR ARMY, NAVY, AND AIR FORCE
UNDER CONTRACT Nonr-225(52) (NR 342 022)
WITH THE OFFICE OF NAVAL RESEARCH
THE SELECTION OF AN OPTIMUM SAMPLING PARTITION

by

Arnold F. Goodman

TECHNICAL REPORT NO. 72
August 3, 1961

PREPARED FOR ARMY, NAVY AND AIR FORCE UNDER
CONTRACT Nonr-225(52) (NR-342-022)
WITH THE OFFICE OF NAVAL RESEARCH

This work was sponsored by the Army, Navy, and Air Force through the Joint Services Advisory Group for Research Groups in Applied Mathematics and Statistics by Contract Nonr-225(52) (NR-342-022)

Reproduction in Whole or in Part is Permitted for any Purpose of the United States Government

APPLIED MATHEMATICS AND STATISTICS LABORATORIES

STANFORD UNIVERSITY

STANFORD, CALIFORNIA
THE SELECTION OF AN OPTIMUM SAMPLING PARTITION

by

Arnold F. Goodman

1. **Introduction and Summary**

It is desired to test a simple hypothesis concerning the probability distribution of a random variable $X$ against a simple alternative. Rather than sampling $X$ directly, it is sometimes more convenient to partition the range of $X$ into numbered sections and to sample the number $Y$ of that section which contains $X$.

This indirect sampling procedure has prompted the investigation of two types of statistical problems. One type is concerned with the selection of an optimum sampling partition to use with a specified statistical test, and the other type is concerned with the selection of an optimum statistical test to use with a specified sampling partition. The former is a design problem and the latter, a testing problem.

In this paper we consider the design problem. The statistical test which we use is a Wald sequential probability ratio test. An asymptotic property of this test suggests a measure for the quality of a sampling partition upon which we base our definition of an optimum one.

Our treatment makes extensive use of the likelihood ratio $\lambda$ of $X$ and the Kullback-Leibler information numbers $I_0$ and $I_1$ of $Y$. The major results of this paper are theorems 4.1 and 5.1. Theorem 4.1 states that for each sampling partition, there exists one based upon $\lambda$ for which neither $I_0$ nor $I_1$ is decreased. A sufficient condition for the existence of an optimum sampling partition is presented in theorem 5.1.
We introduce notation and assumptions and review related literature in section 2. Section 3 characterizes a useful random variable and its distribution. In section 4, sampling partitions based upon $\lambda$ are investigated. The existence of an optimum sampling partition is considered in section 5. In section 6, we discuss some consequences of sections 4 and 5. Finally, section 7 contains a graphical method for obtaining an optimum sampling partition and three illustrative examples.
2. Preliminary Discussion

2.1 Notation and Assumptions

Our treatment of the problem requires a considerable amount of notation and utilizes some simplifying assumptions. For ease of reference, most of these are introduced and discussed in this subsection.

The notation and assumptions involving the random variable $X$ are presented first. Let $\mathcal{X}$ be the range of $X$ and $P_o$ and $P_1$ be two specified probability measures on a $\sigma$ algebra $\mathcal{A}$ of measurable subsets of $\mathcal{X}$. The simple hypothesis $H_o$ is that the probability distribution of $X$ is given by $P_o$ and the simple alternative hypothesis $H_1$ is that it is given by $P_1$.

We assume that $P_o$ and $P_1$ are non-atomic, unless otherwise stated. In sections 4 and 6, we also consider the special discrete (atomic) case for which $\mathcal{X} = \{x_1, \ldots, x_p\}$ and the elements of $\mathcal{X}$ are equally likely under either $H_o$ or $H_1$.

We assume that $P_o$ and $P_1$ differ for some measurable set. Otherwise, they represent the same probability measure on $\mathcal{A}$ and the problem is degenerate.

Let $P$ denote the probability measure $\frac{P_o + P_1}{2}$. Since $P_o$ is absolutely continuous with respect to $P$, it possesses a finite p.d.f. (probability density function) $f_o$ with respect to $P$ by the Radon-Nikodym theorem [9, p. 128, theorem B]. Similarly, $P_1$ possesses a finite p.d.f. $f_1$ with respect to $P$. It follows immediately that $0 \leq f_o \leq 2$, $0 \leq f_1 \leq 2$ and $f_o + f_1 = 2$ a.e. $(P)$.

We assume that $P_o$ and $P_1$ are absolutely continuous with respect to each other. Otherwise, by the proof of the Lebesgue decomposition
theorem [9, p. 134, theorem C], there exist three disjoint measurable
sets \( \mathcal{X}_0 = \{x | f_0(x) = 2\} \), \( \mathcal{X}_1 = \{x | f_1(x) = 2\} \) and \( \mathcal{X}' = \mathcal{X} - (\mathcal{X}_0 \cup \mathcal{X}_1) \)
such that
\[
P_0(\mathcal{X}_0) \geq P_1(\mathcal{X}_0) = 0, \\
P_1(\mathcal{X}_1) \geq P_0(\mathcal{X}_1) = 0,
\]
and \( P_0 \) and \( P_1 \) are absolutely continuous with respect to each other
on the \( \sigma \) algebra of measurable subsets of \( \mathcal{X}' \). We should then au-
tomatically accept \( H_0 \) whenever any observation is contained in \( \mathcal{X}_0 \),
amatically reject \( H_0 \) whenever any observation is contained in \( \mathcal{X}_1 \),
and test the conditional probability distribution of \( X \) given that
\( X \in \mathcal{X}' \) under \( H_0 \) against that under \( H_1 \) whenever all observations
are contained in \( \mathcal{X}' \). Consequently, the non-trivial part of the prob-
lem would have the two probability measures absolutely continuous with
respect to each other.

Now \( P_0, P_1 \) and \( P \) are all absolutely continuous with respect to
each other. Also \( f_0 \) and \( f_1 \) may be so selected that \( 0 < f_0 < 2 \),
\( 0 < f_1 < 2 \) and \( f_0 + f_1 = 2 \).

The likelihood ratio \( \lambda \) of \( X \) is defined to be \( \frac{f_1}{f_0} \). With \( f_0 \)
and \( f_1 \) selected as above, \( \lambda \) is a positive and finite random vari-
able. In addition, \( \lambda \) is a p.d.f. of \( P_1 \) with respect to \( P_0 \) by
the properties of Radon-Nikodym derivatives [9, p. 133, theorem A and
p. 136, problem 2].

The notation and assumptions involving the indirect sampling of \( X \)
are now presented. If \( m \) is a given integer greater than one, a sampling
partition \( E = (E_1, \ldots, E_m) \) is a partition of \( \mathcal{X} \) into \( m \) numbered dis-
joint measurable sections \( E_1, \ldots, E_m \). For each \( E_i \), the random variable
\( Y = Y_{E_i}(X) \) is defined by
(2.1) \( Y_E(x) = k \) when \( x \in E_k \) for \( k = 1, \ldots, m \).

Let \( P_0 = P_0(E_k) \) and \( P_1 = P_1(E_k) \) for \( k = 1, \ldots, m \). Since \( P_0 \) and \( P_1 \) are absolutely continuous with respect to each other, \( P_0 \) is zero if and only if \( P_1 \) is zero for \( k = 1, \ldots, m \). We note that the distribution of \( Y \) is completely determined by \( P_{01}, \ldots, P_{0m} \) under \( H_0 \) and by \( P_{11}, \ldots, P_{1m} \) under \( H_1 \). The set of all probability vectors \( p = (p_{01}, \ldots, p_{0m}, p_{11}, \ldots, p_{1m}) \) which are determined by some sampling partition is denoted by \( \hat{P} \).

We assume that the cost per observation of \( Y \) is a constant \( c \) which is independent of the sampling partition used and the value of \( Y \) obtained.

Consider a statistical test \( T \) which is based upon \( Y \). Let \( r_0 \) be the loss for \( T \) falsely rejecting \( H_0 \), \( r_1 \) be the loss for it falsely accepting \( H_0 \), \( \alpha \) be the probability of \( T \) falsely rejecting \( H_0 \) and \( \beta \) be the probability of it falsely accepting \( H_0 \). We use \( n \) to denote either the fixed sample size of \( T \) or a particular value of the random sample size of \( T \) and \( n_k \) to denote the number of \( Y \)'s out of \( n \) which assume the value \( k \) for \( k = 1, \ldots, m \). If the expected sample size of \( T \) is \( \mu_0 \) under \( H_0 \) and \( \mu_1 \) under \( H_1 \), then the risks of performing \( T \) under \( H_0 \) and \( H_1 \) are given by

\[
R_0 = r_0 \alpha + c\mu_0 \quad \text{and} \\
R_1 = r_1 \beta + c\mu_1 ,
\]

respectively.

The expected risk \( R \) for a given a priori probability \( w \) of \( H_0 \) is

\[
w R_0 + (1-w)R_1 .
\]
The statistical test which we use is a Wald sequential probability ratio test. This test

(2.2) accepts \( H_0 \) if \[ \sum_{k=1}^{m} n_k \log \frac{p_{1k}}{p_{0k}} \leq B, \]

(2.3) continues sampling if \( B < \sum_{k=1}^{m} n_k \log \frac{p_{1k}}{p_{0k}} < A \) and

(2.4) rejects \( H_0 \) if \[ \sum_{k=1}^{m} n_k \log \frac{p_{1k}}{p_{0k}} \geq A, \]

where \( A \) and \( B \) are selected to minimize \( R \). The class of Wald tests has the following optimum properties:

(i) for a given \( w \), there is a Wald test which minimizes \( R \) among all sequential tests [3, p. 267]; and

(ii) for a given \( \alpha \) and \( \beta \), there is a Wald test which simultaneously minimizes \( \mu_0 \) and \( \mu_1 \) among all sequential tests [3, p. 292, theorem 10.9.1].

Chernoff [6, p. 757] has shown that the expected risk of a sequential probability ratio test is asymptotically, as \( c \) approaches zero, approximated by \( (-c \log c)(\frac{w}{I_0} + \frac{1-w}{I_1}) \). (We use log for logarithm to the base e.) Then a good measure for the quality of a sampling partition in our design problem is

(2.5) \[ J_w = \frac{w}{I_0} + \frac{1-w}{I_1}. \]

We define an optimum sampling partition to be one which minimizes \( J_w \).
The Kullback-Leibler information numbers \( J_0 \) and \( J_1 \) of \( X \) and \( I_0 \) and \( I_1 \) of \( Y \), extensively treated in [10], are defined by

\begin{align*}
J_0 &= -\int \log \lambda(x) dP_0(x) , \\
J_1 &= \int \log \lambda(x) dP_1(x) , \\
I_0 &= \sum_{k=1}^{m} p_{ok} \log \frac{p_{ok}}{p_{lk}} \quad \text{and} \\
I_1 &= \sum_{k=1}^{m} p_{lk} \log \frac{p_{lk}}{p_{ok}}
\end{align*}

We define \( q \log \frac{a}{r} \) to be zero when \( q \) is zero. It follows immediately that \( I_0 \) and \( I_1 \) are finite, that \( J_\omega \) is positive and that all three are unaffected by changes in the numbering of the \( m \) sections of \( E \).

Since \( I_0 \leq J_0 \) and \( I_1 \leq J_1 \) [10, p. 16, corollary 3.2], the efficiency of a sampling partition may be measured by \( \frac{I_0}{J_0} \) under \( H_0 \) and \( \frac{I_1}{J_1} \) under \( H_1 \).

A sampling partition \( E \) is ordered by a random variable \( V = V(X) \) if \( V(x) \leq V(x') \) when \( x \in E_i \) and \( x' \in E_j \) for \( 1 \leq i < j \leq m \) and is strictly ordered by \( V \) if these inequalities are strict for \( 1 \leq i < j \leq m \). Then \( E \) is ordered by \( V \) if and only if there exist

\[ \inf_{x \in \mathcal{X}} V(x) = v_0 \leq v_1 \leq \cdots \leq v_{m-1} \leq v_m = \sup_{x \in \mathcal{X}} V(x) \]

such that \( v_{k-1} \leq V(x) \leq v_k \) when \( x \in E_k \) for \( k = 1, \ldots, m \); and \( E \) is strictly ordered by \( V \) if and only if, in addition, \( (x|V(x) = v_{k-1}) \) is contained in either \( E_{k-1} \) or \( E_k \) for \( k = 1, \ldots, m \).
2.2 Related Literature

Resumés of the relevant portions of three papers treating design problems are presented here. We also mention two papers which treat testing problems.

In these resumés, $\gamma$ is a given positive number and $M$ is a specified positive integer. The assumption that $P_0$ and $P_1$ are absolutely continuous with respect to each other does not apply in this subsection.

Chernoff [5] treats a fixed sample size design problem. He uses a statistical test which rejects $H_0$ if $\sum_{j=1}^{n} Y_j > M$. An optimum sampling partition is one which minimizes $\beta + \gamma \alpha$. Theorem 2 [p. 8] states that there exists an optimum sampling partition which is a.e. (P) ordered by $\lambda$. This assertion is also valid for the special discrete case [p. 10, theorem 3].

A very specialized fixed sample size design problem is solved by Kullback in [7]. $P_0$ and $P_1$ are normal probability measures with unknown means $\theta_0$ and $\theta_1$, respectively, and known variance $\sigma^2$. The statistical test which is used rejects $H_0$ if $\sum_{j=1}^{n} Y_j > M$. He confines his attention to sampling partitions of the form $E = \{(-\infty,a),(a,\infty)\}$ and defines one of these to be optimum if it maximizes $I_0 + I_1$.

Kullback then shows that $((-\infty, -\frac{\theta_0 + \theta_1}{2}), (-\frac{\theta_0 + \theta_1}{2}, \infty))$ is an optimum sampling partition [p. 14].

Tsao assumes that $P_0$ and $P_1$ have continuous p.d.f.'s with respect to Lebesgue measure on $s$-dimensional Euclidean space and that $P(\lambda = \lambda')$ is zero for $\lambda' \geq 0$ in [11], [12] and [13].
He treats a sequential design problem with \( m = 3 \) in [11]. The statistical test which he uses accepts \( H_0 \) if \( n_1 = M_1 \), continues to sample if \( n_1 < M_1 \) and \( n_3 < M_2 \), and rejects \( H_0 \) if \( n_3 = M_2 \). The requirement for an optimum sampling partition is the simultaneous minimization of \( \mu_0 \) and \( \mu_1 \) for a given \( \alpha \) and \( \beta \). An optimum sampling partition does not necessarily exist for each \( \alpha \) and \( \beta \). However, any sampling partition ordered by \( \lambda \) which yields the given \( \alpha \) and \( \beta \) is an optimum one [theorem 1, p. 692].

In [12] Tsao considers a fixed sample size testing problem and in [13] he considers a sequential testing problem. Both of these testing problems involve multinomial distributions and apparently were partially inspired by his work on sampling partitions.
3. A Useful Random Variable, $Z$

A random variable $Z$ is introduced and its distribution is discussed in this section.

Certain technical difficulties in the use of $\lambda$ arise from the fact that it may assume some values with positive probability. A supplementary random variable $\Lambda$ is introduced to enable us to deal with these atoms in the distribution of $\lambda$.

Since $P_0$ is non-atomic, there exists a random variable $\Lambda = \Lambda(x)$ such that $\Lambda$ is zero when $P_0[\lambda = \lambda(x)]$ is zero and

\[
(3.1) \quad P_0[\lambda = \lambda(x), \Lambda < \Lambda(x)] = P_0[\lambda = \lambda(x), \Lambda \leq \Lambda(x)] = \Lambda(x)P_0[\lambda = \lambda(x)]
\]

when $P_0[\lambda = \lambda(x)]$ is positive [8, p. 420, lemma 7]. Then $P_0[\lambda = \lambda(x), \Lambda = \Lambda(x)]$ is zero for $x \in \mathcal{X}$ and when $P_0[\lambda = \lambda']$ is positive, the conditional distribution of $\Lambda$ given that $\lambda = \lambda'$ is the uniform distribution on $[0,1]$.

The c.d.f. (cumulative distribution function) of a continuous random variable on the real line provides a non-decreasing transformation to a random variable which is uniformly distributed on $[0,1]$. An analogous transformation, non-decreasing in $\lambda(x)$ and non-decreasing in $\Lambda(x)$ for a given value of $\lambda(x)$, is provided by the random variable $Z = Z(X)$ which is defined by

\[
(3.2) \quad Z(x) = P_0[\lambda < \lambda(x)] + P_0[\lambda = \lambda(x), \Lambda \leq \Lambda(x)] \text{ for } x \in \mathcal{X}.
\]
In Lemma 3.1, we see that $Z$ is not only uniformly distributed on $[0,1]$ under $H_0$ but also possesses a non-decreasing p.d.f. under $H_1$ which is "essentially equal to $\lambda$.”

It follows immediately from the definitions of $\lambda$ and $Z$ that a sampling partition is ordered by $\lambda$ when it is strictly ordered by $Z$ and that it is ordered by $Z$ when it is strictly ordered by $\lambda$. Let $\mathcal{P}$ denote the set of all probability vectors $p^* = (p_{01}^*, \ldots, p_{om}^*, p_{11}^*, \ldots, p_{lm}^*)$ which are determined by some sampling partition ordered by $\lambda$. We note that $\mathcal{P}$ is also the set of all $p^*$ which are determined by some sampling partition strictly ordered by $Z$.

**Lemma 3.1.** $Z$ is distributed on $[0,1]$, uniformly under $H_0$ and with a non-decreasing p.d.f. under $H_1$ which may be selected to be

$$g(z) = \sup \{ \lambda' \mid P_0(\lambda < \lambda') \leq z \} \text{ for } z \in [0,1].$$

**Proof.** That $Z$ is distributed on $[0,1]$ and $g$ is non-decreasing on $[0,1]$ follows immediately from their definitions. Let $h$ be defined on $[0,1]$ by

$$h(z) = 0 \quad \text{when } P_0(\lambda = g(z)) = 0$$

and

$$h(z) = \frac{z - P_0(\lambda < g(z))}{P_0(\lambda = g(z))} \quad \text{when } P_0(\lambda = g(z)) > 0.$$

It is easily shown that for each $z \in [0,1]$,

$$P_0(\lambda < g(z)) + P_0(\lambda = g(z), \lambda \leq h(z)) = z.$$
and \( Z(x) \leq z \) if and only if either \( \lambda(x) < g(z) \) or \( \lambda(x) = g(z) \)
with \( \lambda(x) \leq h(z) \). Hence

\[
(3.4) \quad P_o(z \leq z) = z \quad \text{for} \quad z \in [0,1].
\]

We now demonstrate that

\[
(3.5) \quad \lambda(x) = g(Z(x)) \quad \text{a.e.} \quad (P_o).
\]

Clearly, \( \lambda(x) \leq g(Z(x)) \) for \( x \in \mathcal{X} \). When \( \lambda(x') < g(Z(x')) \),

\[
W[Z(x')] = [\lambda'] \inf_{\{x \mid Z(x) = Z(x')\}} \lambda(x) < \lambda' \leq g(Z(x'))
\]

is a non-degenerate interval of \( \lambda' \) on which

\[
S_o(\lambda') = P_o(\lambda < \lambda') = Z(x').
\]

The \( W[Z(x')] \) are disjoint for different values of \( Z(x') \) and
\( P_o(\lambda \in W[Z(x')]) \) is zero for each such \( Z(x') \). Thus there are at most
a countable number of the \( W[Z(x')] \), which may be denoted by \( W_1, W_2, \ldots \),
and

\[
\{x' \mid \lambda(x') < g(Z(x'))\} = \bigcup_{i=1}^{\infty} \{x' \mid \lambda(x') \in W_i\}.
\]

Hence, \( P_o(\lambda < g(Z)) \) is zero and equation \((3.5)\) is valid.
Therefore,

\[(3.6) \quad P(Z \leq z) = \int_{\{x|Z(x) \leq z\}} \lambda(x) \, dP(x) = \int_0^z g(t) \, dt \quad \text{for } z \in [0,1]\]

and the proof is completed.

Lemma 3.1 justifies the canonical form proposed by Chernoff [5, p. 11] for the problem. In this canonical form, \(X\) is distributed on \([0,1]\), uniformly under \(H_0\) and with a non-decreasing p.d.f. under \(H_1\).

The following properties of \(g\) and the c.d.f. \(G\) of \(Z\) under \(H_1\) are easily verified:

(i) \(g\) and \(\lambda\) are further related by

\[(3.7) \quad g(z) = \inf \{\lambda' | P_0(\lambda < \lambda') > z\}\]

and

\[(3.8) \quad g(Z(x)) = \lambda(x) \text{ a.e. (P)};\]

(ii) \(g\) is continuous and differentiable a.e. and is right continuous at points of discontinuity with

\[(3.9) \quad \lim_{z \to 0^+} g(z) = \inf \{\lambda' | P_0(\lambda < \lambda') > 0\} = g(0)\]

and

\[(3.10) \quad \lim_{z \to 1^-} g(z) = \sup \{\lambda' | P_0(\lambda < \lambda') < 1\} \leq g(1) = \infty;\]

(iii) \(G\) is the definite integral of \(g\);
(iv) \( G \) is convex;
(v) \( G(z) \) is less than or equal to \( z \); and
(vi) the ratio \( \frac{G(z_2') - G(z_1)}{z_2 - z_1} \) is non-decreasing in both \( z_1 \) and \( z_2 \) with

\[
(3.11) \quad g(z_1) \leq \frac{G(z_2') - G(z_1)}{z_2 - z_1} \leq g(z_2) \quad \text{for } 0 \leq z_1 < z_2 \leq 1.
\]

We now note some useful relationships between the coordinates of \( p^* \in \mathcal{C}^* \) and values of \( Z \). For each \( p \in \mathcal{C} \), there exist

\[
0 \leq z_1 \leq \cdots \leq z_{m-1} \leq 1
\]

such that

\[
(3.12) \quad z_k = \sum_{j=1}^{k} \rho_{o,j} \quad \text{for } k = 1, \ldots, m.
\]

If, in addition, \( p \in \mathcal{C}^* \), then

\[
(3.13) \quad G(z_k) = \sum_{j=1}^{k} \rho_{l,j} \quad \text{for } k = 1, \ldots, m.
\]

Conversely, for each

\[
0 = z_0 \leq z_1 \leq \cdots \leq z_{m-1} \leq z_m = 1
\]

there exists \( p^* \in \mathcal{C}^* \) such that
\begin{align}
(3.14) \quad p_{ok}^* &= z_k - z_{k-1} \\
\text{and} \\
(3.15) \quad p_{lk}^* &= G(z_k) - G(z_{k-1}) \quad \text{for } k = 1, \ldots, m.
\end{align}

The related pair of $p^*$ and $(z_1, \ldots, z_{m-1})$ satisfy the following conditions: either $p_{ol}^* + p_{1l}^*$ is zero or

\begin{equation}
(3.16) \quad \lim_{z \to 0^+} g(z) \leq \frac{p_{1l}^*}{p_{ol}^*} \leq g(z_1);
\end{equation}

either $p_{ok}^* + p_{lk}^*$ is zero or

\begin{equation}
(3.17) \quad g(z_{k-1}) \leq \frac{p_{lk}^*}{p_{ok}^*} \leq g(z_k) \quad \text{for } k = 2, \ldots, m-1;
\end{equation}

and either $p_{om}^* + p_{lm}^*$ is zero or

\begin{equation}
(3.18) \quad g(z_{m-1}) \leq \frac{p_{lm}^*}{p_{om}^*} \leq \lim_{z \to 1^-} g(z).
\end{equation}
4. Sampling Partitions Ordered by $\lambda$

In this section, we prove that the search for an optimum sampling partition may be limited to those ordered by $\lambda$. That this is true for the special discrete case is also demonstrated. These results are consequences of theorems 4.1 and 4.2, respectively.

**Theorem 4.1.** For each sampling partition $E$, there exists a sampling partition $E^*$ ordered by $\lambda$ such that $I^*_0 \geq I_o$ and $I^*_1 \geq I_1$.

**Proof.** Let $E_1, \ldots, E_m$ be renumbered so that $\frac{p_{1k}}{p_{ok}}$ is a non-decreasing function of $k$ on $(k|p_{ok} + p_{1k} > 0)$ and $E_i$ and $E_j$ be two arbitrary sections with $i < j$. We define $\tilde{Z}$ to be such that

\begin{equation}
(4.1) \quad p_o((x|x \in E_i \cup E_j, Z(x) \leq \tilde{Z})) = p_{0i},
\end{equation}

$\tilde{E}_i$ to be the subset of $E_i$ on which $Z > \tilde{Z}$ and $\tilde{E}_j$ to be the subset of $E_j$ on which $Z < \tilde{Z}$. Let $\tilde{E}_i$ be constructed from $E_i$ by replacing $\tilde{E}_i$ with $\tilde{E}_j$ and $\tilde{E}_j$ be constructed from $E_j$ by replacing $\tilde{E}_j$ with $\tilde{E}_i$. It follows immediately that $Z \leq \tilde{Z}$ on $\tilde{E}_i$, $Z > \tilde{Z}$ on $\tilde{E}_j$, $p_{0i} = p_{0i}$, $p_{0j} = p_{0j}$, $p_o(\tilde{E}_i) = p_o(\tilde{E}_j)$ and

\begin{equation}
(4.2) \quad p_{1i} + p_{1j} = p_{1i} + p_{1j} = s.
\end{equation}

Since $\lambda \leq g(\tilde{Z})$ when $Z \leq \tilde{Z}$ and $\lambda \geq g(\tilde{Z})$ when $Z > \tilde{Z}$,

\[ p_i(\tilde{E}_i) \geq g(\tilde{Z}) p_o(\tilde{E}_i) \]

16
and

\[ P_1(\tilde{E}_j) \leq g(\tilde{z}) \ P_0(\tilde{E}_j) = g(\tilde{z}) \ P_0(\tilde{E}_1) . \]

Hence \( p_{1 \tilde{z}}^* \leq p_{1 \tilde{z}} \) and

\[ (4.3) \quad \frac{p_{1 \tilde{z}}}{p_{o i}} \leq \frac{p_{1 \tilde{z}}}{p_{o i}} \leq \frac{p_{1 \tilde{z}}}{p_{o j}} \leq \frac{p_{1 \tilde{z}}}{p_{o j}} = \frac{s-p_{1 \tilde{z}}}{p_{o j}} . \]

The relevant portions of \( I_o \) and \( I_1 \) are defined by

\[ (4.4) \quad Q_o(p_{1 \tilde{z}}) = p_{o i} \log \frac{p_{o i}}{p_{1 \tilde{z}}} + p_{o j} \log \frac{p_{o j}}{s-p_{1 \tilde{z}}} \]

and

\[ (4.5) \quad Q_1(p_{1 \tilde{z}}) = p_{1 \tilde{z}} \log \frac{p_{1 \tilde{z}}}{p_{o i}} + (s-p_{1 \tilde{z}}) \log \frac{s-p_{1 \tilde{z}}}{p_{o j}} , \text{ respectively.} \]

If either \( p_{o i} + p_{1 \tilde{z}} \) or \( p_{o j} + p_{1 \tilde{z}} \) is zero, then \( Q_o(p_{1 \tilde{z}}) = Q_o(p_{1 \tilde{z}}) \) and \( Q_1(p_{1 \tilde{z}}) = Q_1(p_{1 \tilde{z}}) \). Otherwise, both \( Q_o \) and \( Q_1 \) are non-increasing functions of \( p_{1 \tilde{z}} \) when \( p_{1 \tilde{z}} \leq p_{1 \tilde{z}} \); for then

\[ (4.6) \quad \frac{d Q_o}{dp_{1 \tilde{z}}} = \frac{p_{o j}}{s-p_{1 \tilde{z}}} - \frac{p_{o i}}{p_{1 \tilde{z}}} \leq 0 \]

and

\[ (4.7) \quad \frac{d Q_1}{dp_{1 \tilde{z}}} = \log \frac{p_{1 \tilde{z}}}{p_{o i}(s-p_{1 \tilde{z}})} \leq 0 . \]
Hence \( Q_0(p_{11}) \geq Q_0(p_{11}) \) and \( Q_1(p_{11}) \geq Q_1(p_{11}). \)

Now apply this construction procedure sequentially to \( E_1 \) and \( E_2 \), to the once altered \( E_1 \) and \( E_3 \), to the twice altered \( E_1 \) and \( E_4 \), \ldots, to the \( m-2 \) times altered \( E_1 \) and \( E_m \). We note that the maximum value of \( Z \) on the current \( E_1 \) does not increase from step to step. If \( E_1^* \) is the final form of \( E_1 \) (i.e., the \( m-1 \) times altered \( E_1 \)), then the value of \( Z \) on \( \mathcal{E} - E_1^* \) is greater than the maximum value of \( Z \) on \( E_1^* \). Let \( E^* \) be constructed by repeating this process for \( E_2 \), \ldots, \( E_{m-1} \). It is easily shown that \( E^* \) is strictly ordered by \( Z \), \( I_0^* \geq I_0 \) and \( I_1^* \geq I_1 \).

The following corollary is an immediate consequence of theorem 4.1.

**Corollary 4.1.** An optimum sampling partition exists if and only if one ordered by \( \lambda \) exists. Hence the search for one may be limited to those ordered by \( \lambda \).

Chernoff's treatment of the special discrete case suggests that a theorem analogous to theorem 4.1 may be true for this case. We now state and prove that theorem and its corollary.

**Theorem 4.2 (The Special Discrete Case).** Let \( \mathcal{X} \) be \( \{x_1, \ldots, x_D\} \) and the elements of \( \mathcal{X} \) be equally likely under either \( H_0 \) or \( H_1 \). Then for each sampling partition \( E \), there exists a sampling partition \( E^* \) ordered by \( \lambda \) such that \( I_0^* \geq I_0 \) and \( I_1^* \geq I_1 \).

**Proof.** We perform the proof for the case where the elements of \( \mathcal{X} \) are equally likely under \( H_0 \). The proof for the case where they are equally likely under \( H_1 \) involves only minor modifications.
Let $E_1, \ldots, E_m$ be renumbered so that $\frac{p_{1k}}{p_{0k}}$ is a non-decreasing function of $k$ on $\{k | p_{0k} + p_{1k} > 0\}$ and $E_i$ and $E_j$ be two arbitrary sections with $i < j$. If there are $D_i$ elements in $E_i$ and $D_j$ elements in $E_j$, and these elements are renumbered from 1 to $D_i + D_j$ so that

\[(4.8) \quad \lambda(x_1) \leq \cdots \leq \lambda(x_{D_i + D_j}),\]

then $\oplus_i$ and $\oplus_j$ are defined to be \{\{x_1, \ldots, x_{D_i}\} and \{x_{D_i+1}, \ldots, x_{D_i+D_j}\}, respectively. It follows immediately that $\lambda \leq \lambda(x_{D_i})$ on $\oplus_i$, $\lambda \geq \lambda(x_{D_i})$ on $\oplus_j$, $p_{oi} = p_{oi}$, $p_{oj} = p_{oj}$, $p_{ii} \leq p_{ij}$ and

\[(4.9) \quad p_{ii} + p_{ij} = p_{ii} + p_{ij} = s.\]

The proof may be completed by using the same argument as that in the last two paragraphs of the proof of theorem 4.1.

**Corollary 4.2 (The Special Discrete Case).** If $\mathcal{X}$ is \{\{x_1, \ldots, x_D\} and the elements of $\mathcal{X}$ are equally likely under either $H_0$ or $H_1$, then an optimum sampling partition ordered by $\lambda$ exists. Hence the search for one may be limited to those ordered by $\lambda$.

**Proof.** Since there are only a finite number of sampling partitions, an optimum one must exist. Theorem 4.2 then implies that an optimum one ordered by $\lambda$ exists.
It is natural to wonder, as did Chernoff in [5], if the same result is valid for the general discrete case. The construction procedure used in the proofs of theorems 4.1 and 4.2 breaks down for the general discrete case because $p_{01}^*$ cannot necessarily be made to equal $p_{01}$. However, no counter-example to the validity of this result for the general discrete case has been found.
5. Existence of an Optimum Sampling Partition

This section is concerned with the existence of an optimum sampling partition. One exists, by corollary 4.1, if and only if \( J_w \) attains its minimum value on \( \mathcal{P}^* \). The lemma below states a useful property of \( \mathcal{P}^* \).

Lemma 5.1. \( \mathcal{P}^* \) is a compact subset of 2m-dimensional Euclidean space.

Proof. \( \mathcal{P}^* \) is bounded because the coordinates of \( p^* \) are probabilities. By the Heine-Borel theorem [1, p. 53, theorem 3-38], it suffices to prove that \( \mathcal{P}^* \) is closed. To show this, let

\[
p^*_r = (p^*_r, \ldots, p^*_r, p^*_{rl}, \ldots, p^*_{r_2}) \quad \text{for} \quad r = 1, 2, \ldots,
\]

be a sequence of points from \( \mathcal{P}^* \) which converges to a point \( p = (p_{r_1}, \ldots, p_{rl}, \ldots, p_{r_2}) \). Since \( \mathcal{P} \) is closed by an extension of the Liapounoff theorem on the range of a vector measure [4, p. 725, theorem 1], \( p \in \mathcal{P} \).

For \( k = 1, \ldots, m-1 \) and \( r = 1, 2, \ldots \), we define \( z_{rk} \) and \( z_k \) by

\[
(5.1) \quad z_{rk} = \sum_{j=1}^{k} p^*_{rj}
\]

and

\[
(5.2) \quad z_k = \sum_{j=1}^{k} p^*_{rj}.
\]

Then

\[
(5.3) \quad G(z_{rk}) = \sum_{j=1}^{k} p^*_{rj} \quad \text{for} \quad k = 1, \ldots, m-1 \quad \text{and} \quad r = 1, 2, \ldots.
\]

21
The convergence of \( p_r^k \) to \( p \) and equations (5.1), (5.2) and (5.3) imply that

\[
z_{rk} \rightarrow \sum_{j=1}^{k} p_{0j} = z_k
\]

and

\[
G(z_{rk}) \rightarrow \sum_{j=1}^{k} p_{ij} \text{ for } k = 1, \ldots, m-1.
\]

Since \( G \) is continuous [1, p. 214, theorem 9-31],

\[
(5.4) \quad G(z_k) = \sum_{j=1}^{k} p_{ij} \text{ for } k = 1, \ldots, m-1
\]

and \( p \in \mathcal{P}^* \). Hence \( \mathcal{P}^* \) is closed and the proof is complete.

We investigate the behavior of \( J_w \) on \( \mathcal{P}^* \) via that of \( I_o \) and \( I_1 \) on \( \mathcal{P}^* \). A sufficient condition for \( J_w \) to attain its minimum value on \( \mathcal{P}^* \) is that both \( I_o \) and \( I_1 \) be continuous on \( \mathcal{P}^* \), because a continuous function on a compact set attains its minimum value [1, p. 73, theorem 4-20]. Hence the continuity properties of \( I_o \) and \( I_1 \) on \( \mathcal{P}^* \) are of particular interest. Example 5.1 shows that \( I_o \) may not only fail to be continuous but also fail to be bounded.

**Example 5.1.** Suppose that \( X \) is uniformly distributed on \([0,1]\) under \( H_0 \) and that

\[
\lambda(x) = \begin{cases} 
2x^3e^{-x^2} & \text{for } x \in [0, \frac{1}{2}] \\
2(1-e^{-2}) & \text{for } x \in (\frac{1}{2}, 1]
\end{cases}
\]
We consider the case \( m = 2 \). Then \( I_0 \) may be regarded as a function of \( p^*_{01} \) alone. It is easily verified that

\[
p^*_{11} = \int_0^\infty 2x^3 e^{-x} dx = e \left( \frac{1}{p^*_{01}} \right)^2 \quad \text{for} \quad p^*_{01} \in [0, \frac{1}{2}]
\]

and

\[
\lim_{p^*_{01} \to 0^+} I_0(p^*_{01}) = \lim_{p^*_{01} \to 0^+} p^*_{01} \log \frac{p^*_{01}}{p^*_{11}} = \lim_{p^*_{01} \to 0^+} \frac{1}{p^*_{01}} = \infty.
\]

Hence \( I_0 \) is not continuous, is not bounded and does not attain its maximum value.

This divergence of \( I_0 \) to infinity as \( p^*_{01} \) approaches zero is caused by the very rapid decrease of \( \lambda \) to zero as \( x \) approaches zero.

The following theorem, corollary and conjecture treat the behavior of \( I_0 \), \( I_1 \) and \( J_\psi \) on \( \Phi^* \).

**Theorem 5.1.** \( I_0 \) attains its maximum value on \( \Phi^* \) if

\[
(5.5) \quad \lim_{p^*_{01} \to 0} \frac{p^*_{01} \log \frac{p^*_{01}}{p^*_{11}}}{p^*_{11}} = 0 \quad \text{for} \quad p^* \in \Phi^*
\]

and only if there exists \( \tilde{p}^* \in \Phi^* \) such that

\[
(5.6) \quad \limsup_{p^*_{01} \to 0} \frac{p^*_{01} \log \frac{p^*_{01}}{p^*_{11}}}{p^*_{11}} \leq \tilde{p}^* \log \frac{p^*_{01}}{p^*_{11}} \quad \text{for} \quad p^* \in \Phi^*
\]
$I_1$ attains its maximum value on $\Theta^*$ if

\begin{equation}
\lim_{p^*_\text{om} \to 0, \ p^* \in \Theta^*} \frac{p^*_\text{lm}}{p^*_\text{om}} \log \frac{p^*_\text{lm}}{p^*_\text{om}} = 0
\end{equation}

and only if there exists $\hat{p}^* \in \Theta^*$ such that

\begin{equation}
\limsup_{p^*_\text{om} \to 0, \ p^* \in \Theta^*} \frac{p^*_\text{lm}}{p^*_\text{om}} \log \frac{p^*_\text{lm}}{p^*_\text{om}} < \hat{p}^*_\text{lm} \log \frac{\hat{p}^*_\text{lm}}{p^*_\text{om}}.
\end{equation}

For each $w \in (0,1)$, $J_w$ attains its minimum value on $\Theta^*$ and an optimum sampling partition exists if both equations (5.5) and (5.7) are valid.

**Proof.** We first prove the assertion concerning $I_0$. A similar argument proves the assertion concerning $I_1$.

Since $\Theta^*$ is compact, the sufficiency of equation (5.5) is established if it implies that $I_0$ is continuous on $\Theta^*$ [1, p. 73, theorem 4-20]. We demonstrate below that it does.

The statements in the following paragraph are true for $k=1,\ldots,m$. Except possibly at points for which $p^*_\text{ok} + p^*_\text{lk}$ is zero, $p^*_\text{ok} \log \frac{p^*_\text{lk}}{p^*_\text{ok}}$ is a continuous function on $\Theta^*$. Since $g$ is non-decreasing,

\[
p^*_\text{lk} = \int_{z_k}^{z_k-z_{k-1}} g(z)dz \geq \int_{z_{k-1}}^{0} g(z)dz = g(p^*_\text{ok})
\]

and so

\[
p^*_\text{ok} \log p^*_\text{ok} \leq p^*_\text{ok} \log \frac{p^*_\text{lk}}{p^*_\text{ok}} \leq p^*_\text{ok} \log \frac{p^*_\text{ok}}{g(p^*_\text{ok})}.
\]
Equation (5.5) then implies that

\[
(5.9) \quad \lim_{p^* \to 0} p^* \log \frac{p^*_{ok}}{p^*_{lk}} = 0. \\
\text{for } p^* \in \varphi^*.
\]

Hence \( p^* \log \frac{p^*_{ok}}{p^*_{lk}} \) is continuous on \( \varphi^* \).

It follows immediately that \( I_o^* \) is continuous on \( \varphi^* \).

To prove the necessity of equation (5.6), we fix \( p^*_{ok} \) and \( p^*_{lk} \)
for \( k = 3, \ldots, m \) and let

\[
t = 1 - \sum_{k=3}^{m} p^*_{ok}.
\]

Now \( I_o^* \) may be regarded as a function of \( p^*_{ol} \) alone. Since \( t \) is a
constant,

\[
p^*_{o2} \log \frac{p^*_{o2}}{p^*_{l2}} = (t-p^*_{ol}) \log \frac{t-p^*_{ol}}{G(t)-G(p^*_{ol})}
\]

increases as \( p^*_{ol} \) decreases to zero. When \( \varphi^* \) does not exist,

\[
p^*_{ol} \log \frac{p^*_{ol}}{G(p^*_{ol})} < \limsup z \log \frac{z}{G(z)} \text{ for } p^*_{ol} \in [0,t]
\]

and so \( I_o^* \) does not attain its maximum value for each fixed
\( p^*_{o3}, \ldots, p^*_{om}, p^*_{l3}, \ldots, p^*_{lm} \). Hence it does not attain its maximum
value on \( \varphi^* \).
The assertion concerning $J_w$ follows from the subsequent continuity of $I_o$ and $I_1$ on $\mathcal{O}^*$. 

It would have been preferable to phrase theorem 5.1 in terms of $\lambda$ rather than $p^*$, but technical difficulties prevented this. However, equations analogous to (5.5) and (5.7) can be stated in terms of $\lambda$ when some additional conditions are satisfied. This is done in corollary 5.1. Conjecture 5.1 indicates a variation of theorem 5.1 which we believe to be true.

**Corollary 5.1.** $I_o$ attains its maximum value on $\mathcal{O}^*$ if either $\inf(\lambda' | P_o(\lambda < \lambda') > 0)$ is positive or $S_o(\lambda') = P_o(\lambda < \lambda')$ is continuous and increasing in some neighborhood $U_o$ of $\lambda' = 0$ with

$$
(5.10) \quad \lim_{\lambda' \to 0^+} \frac{P_o(\lambda < \lambda')}{P_1(\lambda < \lambda')} \log \frac{P_o(\lambda < \lambda')}{P_1(\lambda < \lambda')} = 0.
$$

$I_1$ attains its maximum value on $\mathcal{O}^*$ if either $\sup(\lambda' | P_o(\lambda < \lambda') < 1)$ is finite or $S_o(\lambda')$ is continuous and increasing in some neighborhood $U_\infty$ of $\lambda' = \infty$ with

$$
(5.11) \quad \lim_{\lambda' \to \infty} \frac{P_1(\lambda > \lambda')}{P_o(\lambda > \lambda')} \log \frac{P_1(\lambda > \lambda')}{P_o(\lambda > \lambda')} = 0.
$$

**Proof.** We prove the assertion concerning $I_o$. That concerning $I_1$ is proved similarly.
Since either $p_{ol}^* + p_{ll}^*$ is zero or $\frac{p_{ol}^*}{p_{ll}^*}$ is bounded above by
\[
\frac{1}{\inf(\lambda' | P_o(\lambda < \lambda') > 0)}
\]

\[
p_{ol}^* \log p_{ol}^* \leq p_{ol}^* \log \frac{p_{ol}^*}{p_{ll}^*} \leq p_{ol}^* \log \frac{1}{\inf(\lambda' | P_o(\lambda < \lambda') > 0)}
\]

Equation (5.5) is then valid when $\inf(\lambda' | P_o(\lambda < \lambda') > 0)$ is positive.

There is a one-to-one correspondence between $U_o$ and the neighborhood $S_o(U_o)$ of $p_{ol}^* = 0$. Equation (5.5) is then valid when equation (5.10) is valid.

In either case, $I_o$ attains its maximum value on $\Theta^*$ by theorem 5.1.

**Conjecture 5.1.** $I_o$ attains its maximum value on $\Theta^*$ if

\[
(5.12) \lim_{\lambda' \to 0^+} P_o(\lambda < \lambda') \log \frac{P_o(\lambda < \lambda')}{P_l(\lambda < \lambda')} = 0
\]

and only if there exists $\lambda_o$ such that

\[
(5.13) \limsup_{\lambda' \to 0^+} P_o(\lambda < \lambda') \log \frac{P_o(\lambda < \lambda')}{P_l(\lambda < \lambda')} \leq P_o(\lambda < \lambda_o) \log \frac{P_o(\lambda < \lambda_o)}{P_l(\lambda < \lambda_o)}
\]

$I_l$ attains its maximum value on $\Theta^*$ if

27
\[(5.14) \quad \lim_{\lambda' \to \infty} P_1(\lambda > \lambda') \log \frac{P_1(\lambda > \lambda')}{P_0(\lambda > \lambda')} = 0\]

and only if there exists \( \lambda_\infty \) such that

\[(5.15) \quad \limsup_{\lambda' \to \infty} P_1(\lambda > \lambda') \log \frac{P_1(\lambda > \lambda')}{P_0(\lambda > \lambda')} \leq P_1(\lambda > \lambda_\infty) \log \frac{P_1(\lambda > \lambda_\infty)}{P_0(\lambda > \lambda_\infty)}.\]

For each \( w \in (0,1) \), \( J_w \) attains its minimum value on \( \mathcal{D}^* \) and an optimum sampling partition exists if both equations (5.12) and (5.14) are valid.
6. Consequences

The results of sections 4 and 5 apply to \( I_0 \) and \( I_1 \), in general, as well as to \( J_w \), in particular. We note some consequences of these results here.

Whenever the appropriate measure for the quality of a sampling partition is a monotonic function of \( I_0 \) and \( I_1 \),

(i) the search for an optimum sampling partition may be limited to those ordered by \( \lambda \) in both the non-atomic and special discrete cases;

(ii) an optimum sampling partition ordered by \( \lambda \) exists for the special discrete case; and

(iii) the existence of an optimum sampling partition for the non-atomic case can be investigated by means of theorem 5.1 and corollary 5.1.

We now consider a generalization of our design problem in which a sampling partition may be selected prior to each sequential observation. In addition, this selection may be based upon all previous observations.

Chernoff in [6] and Bessler in [2] treat the sequential design of experiments, of which this problem is a special case. Chernoff's results apply to the special discrete case and Bessler's to the non-atomic case.

We introduce the following notation: \( \theta \) is a parameter which assumes the value \( \theta_0 \) under \( H_0 \) and \( \theta_1 \) under \( H_1 \), \( I(\theta_0) = I_0 \), \( I(\theta_1) = I_1 \) and \( \hat{\theta}_n \) is the maximum likelihood estimate of \( \theta \) based upon the first \( n \) observations. In these papers, an optimum sampling partition (experiment) after \( n \) observations is one which maximizes \( I(\hat{\theta}_n) \).
Since $I(\hat{n})$ is a monotonic function of $I_0$ and $I_1$, our previous discussion is applicable to this problem. In particular, an optimum sampling partition exists at each stage for the non-atomic case if both equations (5.5) and (5.7) are valid and only if both $\tilde{p}^*$ and $\hat{p}^*$ exist.
7. A Graphical Method and Examples

In this section, we present a graphical method for obtaining an optimum sampling partition and three illustrative examples. These examples treat a real random variable $X$ with location parameter $\theta_0$ under $H_0$ and $\theta_1$ under $H_1$.

The problem of finding a sampling partition which minimizes $J_w$ is reminiscent of the problem of finding a Bayes strategy in a statistical game for which Nature has two actions. Let $L$ be the set of all points $(\frac{1}{I_0}, \frac{1}{I_1})$ which are determined by some sampling partition and $L^*$ be the lower left-hand boundary of $L$. By the results of section 4, $L^*$ is also the lower left-hand boundary of the set of all points $(\frac{1}{I_0}, \frac{1}{I_1})$ which are determined by some sampling partition ordered by $\lambda$.

Suppose that the line

$$\frac{w}{I_0} + \frac{1-w}{I_1} = J$$

is moved from the origin to the right until it first intersects $L^*$. It follows immediately that the minimum value of $J_w$ on $L$ is attained at the points of this intersection, if it is attained at all. Hence any sampling partition which determines a point of this intersection is an optimum one.

The sampling partitions which determine $L^*$ are, therefore, of particular interest to us.

For convenience, we define the random variable $X'$ and the parameter $\Delta$ by
(7.1) \[ X' = X - \frac{\theta_0 + \theta_1}{2} \]

and

(7.2) \[ \Delta = \frac{\theta_1 - \theta_0}{2} . \]

Then \( X' \) has the same type of distribution as \( X \), with location parameter \(-\Delta\) under \( H_0 \) and \( \Delta \) under \( H_1 \).

**Example 7.1.** Let \( X \) be normally distributed with unit variance and mean \( \theta_0 \) under \( H_0 \) and \( \theta_1 \) under \( H_1 \). It is easily verified that

\[ \lambda(x') = e^{2\Delta x'} \text{ for } x' \in (-\infty, \infty) \]

and

\[ \Delta_0 = \Delta_1 = 2\Delta^2 . \]

In this example, we first treat the cases for which \( m = 2 \) and \( \Delta = .1, .5, 2.5 \) and then treat the case for which \( m = 3 \) and \( \Delta = .5 \).

For each sampling partition \( E^* = (E_1^*, E_2^*) \) ordered by \( \lambda \), there exists \(-\infty \leq a \leq \infty \) such that

(7.3) \[ E^* = ((-\infty, a], (a, \infty)) \]

or

(7.4) \[ E^* = ((-\infty, a), [a, \infty)) . \]
If $\phi$ is the c.d.f. of the normal distribution with zero mean and unit variance and $E^*$ is defined by equation (7.3) or (7.4), then

$$I_o = \phi(a+\Delta) \log \frac{\phi(a+\Delta)}{\phi(a-\Delta)} + [1-\phi(a+\Delta)] \log \frac{1-\phi(a+\Delta)}{1-\phi(a-\Delta)}$$

and

$$I_1 = \phi(a-\Delta) \log \frac{\phi(a-\Delta)}{\phi(a+\Delta)} + [1-\phi(a-\Delta)] \log \frac{1-\phi(a-\Delta)}{1-\phi(a+\Delta)}.$$

We note that $I_o(-a) = I_1(a)$.

Figures 1a, 1b and 1c depict $L^*$ for $m = 2$ and $\Delta = .1, .5, 2.5$, respectively. We label selected points of $L^*$ with corresponding values of $a$.

For each sampling partition $E^* = (E^*_1, E^*_2, E^*_3)$ ordered by $\lambda$, there exist $-\infty \leq a_1 \leq a_2 \leq \infty$ such that

$$E^* = ((-\infty,a_1],(a_1,a_2),[a_2,\infty)),$$

$$E^* = ((-\infty,a_1],[a_1,a_2],[a_2,\infty)),$$

$$E^* = ((-\infty,a_1],[a_1,a_2),[a_2,\infty)) \text{ or }$$

$$E^* = ((-\infty,a_1],[a_1,a_2],[a_2,\infty)).$$

When $E^*$ is defined by equation (7.5), (7.6), (7.7) or (7.8),

$$I_o = \phi(a_1+\Delta) \log \frac{\phi(a_1+\Delta)}{\phi(a_1-\Delta)} + [\phi(a_2+\Delta)-\phi(a_1+\Delta)] \log \frac{\phi(a_2+\Delta)-\phi(a_1+\Delta)}{\phi(a_2-\Delta)-\phi(a_1-\Delta)}$$

$$+ [1-\phi(a_2+\Delta)] \log \frac{1-\phi(a_2+\Delta)}{1-\phi(a_2-\Delta)}$$

and

33
\[ I_1 = \phi(a_1-\Delta) \log \frac{\phi(a_1-\Delta)}{\phi(a_1+\Delta)} + [\phi(a_2-\Delta)-\phi(a_1-\Delta)] \log \frac{\phi(a_2-\Delta)-\phi(a_1-\Delta)}{\phi(a_2+\Delta)-\phi(a_1+\Delta)} \]

\[ + [1-\phi(a_2-\Delta)] \log \frac{1-\phi(a_2-\Delta)}{1-\phi(a_2+\Delta)}. \]

We see that \( I_o(-a_2,-a_1) = I_1(a_1,a_2). \)

The graph of \( L^* \) for \( m = 3 \) and \( \Delta = .5 \) appears in figure 1d.

We note that \( L^* \) is the envelope of a family of arcs which are generated by fixing \( a_1 \) and varying \( a_2 \). Selected points of \( L^* \) are labeled with corresponding values of \( (a_1,a_2) \). These values have been found by interpolation. For reference, two other points on selected arcs are labeled with corresponding values of \( (a_1,a_2) \).

Example 7.2. Let the distribution of \( X \) be the Cauchy distribution with unit dispersion parameter and location parameter \( \theta_o \) under \( H_c \) and \( \theta_1 \) under \( H_1 \). Then

\[ \lambda(x') = \frac{1+(x'+\Delta)^2}{1+(x'-\Delta)^2} \text{ for } x' \in (-\infty, \infty), \]

\[ \mathcal{A}_o = \int_{-\infty}^{\infty} \frac{1}{\pi[1+(x'+\Delta)^2]} \log \frac{1+(x'+\Delta)^2}{1+(x'-\Delta)^2} dx', \]

\[ \mathcal{A}_1 = \int_{-\infty}^{\infty} \frac{1}{\pi[1+(x'-\Delta)^2]} \log \frac{1+(x'+\Delta)^2}{1+(x'-\Delta)^2} dx'. \]

It is easily verified that

\[ \mathcal{A}_o = \mathcal{A}_1 < 2\Delta^2. \]
In addition, the equation \( \lambda(x') = \tilde{x} \) has the two solutions \( a_1 \) and \( a_2 \), with \( a_1 < a_2 \), which are given by

\[
\frac{\Delta(\tilde{x}+1) + \sqrt{4\Delta^2 \tilde{x} - (\tilde{x}-1)^2}}{\tilde{x}-1} \quad \text{for } \tilde{x} \neq 1
\]

and

\[
2\Delta^2 + 1 - \sqrt{(2\Delta^2 + 1)^2 - 1} \leq \tilde{x} \leq 2\Delta^2 + 1 + \sqrt{(2\Delta^2 + 1)^2 - 1}.
\]

We note that \( \lambda(0) = 1 \). The cases for which \( m = 2 \) and \( \Delta = 0.1, 0.5, 2.5 \) are considered.

It follows immediately that for each sampling partition \( E^* \) ordered by \( \lambda \), there exist either \( a_1 < a_2 \leq 0 \) such that

\[
E^* = ([a_1, a_2], (-\infty, a_1) \cup (a_2, \infty))
\]

or

\[
E^* = ((a_1, a_2), (-\infty, a_1) \cup [a_2, \infty))
\]

or \( 0 \leq a_1 < a_2 \) such that

\[
E^* = ((-\infty, a_1) \cup (a_2, \infty), [a_1, a_2])
\]

or

\[
E^* = ((-\infty, a_1] \cup [a_2, \infty), (a_1, a_2)).
\]
When $E^*$ is defined by equation (7.9), (7.10), (7.11) or (7.12) and

$$F(u) = \frac{1}{2} + \frac{1}{\pi} \arctan u \text{ for } u \in (-\infty, \infty),$$

$$I_o = [F(a_2 + \Delta) - F(a_1 + \Delta)] \log \frac{F(a_2 + \Delta) - F(a_1 + \Delta)}{F(a_2 - \Delta) - F(a_1 - \Delta)}$$

$$+ [1 + F(a_1 + \Delta) - F(a_2 + \Delta)] \log \frac{1 + F(a_1 + \Delta) - F(a_2 + \Delta)}{1 + F(a_1 - \Delta) - F(a_2 - \Delta)}$$

and

$$I_1 = [F(a_2 - \Delta) - F(a_1 - \Delta)] \log \frac{F(a_2 - \Delta) - F(a_1 - \Delta)}{F(a_2 + \Delta) - F(a_1 + \Delta)}$$

$$+ [1 + F(a_1 - \Delta) - F(a_2 - \Delta)] \log \frac{1 + F(a_1 - \Delta) - F(a_2 - \Delta)}{1 + F(a_1 + \Delta) - F(a_2 + \Delta)}.$$

We see that $I_o(-a_2, -a_1) = I_1(a_1, a_2)$.

Figures 2a, 2b and 2c depict $L^*$ for $\Delta = .1, .5, 2.5$, respectively. We label selected points of $L^*$ with corresponding values of $\lambda$ and $(a_1, a_2)$.

**Example 7.3.** Suppose that $X$ is distributed according to the Cauchy distribution with unit dispersion parameter and location parameter $\theta_0$ under $H_0$ and $\theta_1$ under $H_1$, and that we are going to select the sampling partition which minimizes $J_w$ among those of the form

$$E = ((-\infty, a], (a, \infty)) \text{ for some } -\infty < a < \infty.$$
The case for which \( m = 2 \) and \( \Delta = .5 \) is treated. It follows immediately that when \( E \) is defined by equation (7.13),

\[
I_0 = F(a+\Delta) \log \frac{F(a+\Delta)}{F(a-\Delta)} + [1-F(a+\Delta)] \log \frac{1-F(a+\Delta)}{1-F(a-\Delta)}
\]

and

\[
I_1 = F(a-\Delta) \log \frac{F(a-\Delta)}{F(a+\Delta)} + [1-F(a-\Delta)] \log \frac{1-F(a-\Delta)}{1-F(a+\Delta)}.
\]

We note that \( I_0(-a) = I_1(a) \).

For purposes of comparison, figure 2b also contains the graph of the lower left-hand boundary of the set of all points \( (\frac{1}{I_0}, \frac{1}{I_1}) \) which are determined by some sampling partition defined by equation (7.13). Selected points are labeled with corresponding values of \( a \).

We conclude this section with brief comments concerning the figures. They enable us to observe the effect of a change in \( \Delta \) for a constant \( m \), the effect of a change in \( m \) for a constant \( \Delta \) and the effect of using a sampling partition which minimizes \( J_w \) only among those defined by equation (7.13).

As \( \Delta \) increases in examples 7.1 and 7.2 for \( m = 2 \),

1. it becomes less difficult to discriminate between the distributions under \( \mathcal{H}_0 \) and \( \mathcal{H}_1 \);
2. \( I_0 \) and \( I_1 \) increase;
3. the efficiency of a sampling partition, as measured by \( \frac{I_0}{S_0} \) or \( \frac{I_1}{S_1} \), increases;
4. \( L^* \) moves toward the origin; and
5. the shape of \( L^* \) remains about the same.
For $\Delta = .5$ in example 7.1, an increase in $m$ causes

(i) an increase in $I_o$ and $I_l$;

(ii) an increase in the efficiency of a sampling partition, as measured by $\frac{I_o}{A_o}$ and $\frac{I_l}{A_l}$;

(iii) $L^*$ to move toward the origin; and

(iv) the shape of $L^*$ to change appreciably.

Finally, using a sampling partition which minimizes $J_w$ among those defined by equation (7.13) is almost as good as using an optimum sampling partition for $\Delta = .5$ in example 7.2.
\[ \Delta_0 = \Delta_1 = 0.02 \]
\[ E^* = ((-\infty, a], (a, \infty)) \text{ or } \]
\[ E^* = ((-\infty, a), [a, \infty)) \]

Figure 1a

NORMAL DISTRIBUTION, \( \Delta = 0.1 \)
$\phi_0 = \phi_1 = .50$

$E^* = ((-\infty, a], (a, \infty))$ or

$E^* = ((-\infty, a), [a, \infty))$

Figure 1b: NORMAL DISTRIBUTION, $\Delta = .5$
$t_0 = t_1 = 12.50$

$E^* = ((-\infty, a], (a, \infty))$ or $E^* = ((-\infty, a], [a, \infty))$

Figure 1c

NORMAL DISTRIBUTION, $\Delta = 2.5$
$$\lambda_0 = \lambda_1 = .5$$

$$E^* = ((-\infty, a_1], (a_1, a_2], (a_2, \infty))$$,

$$E^* = ((-\infty, a_1), [a_1, a_2], (a_2, \infty))$$,

$$E^* = ((-\infty, a_1], (a_1, a_2), [a_2, \infty))$$ or

$$E^* = ((-\infty, a_1), [a_1, a_2), [a_2, \infty))$$

\[A = (-1.0, .10) \quad K = (-.6, .90)\]
\[B = (-.9, .30) \quad L = (-.4, .75)\]
\[C = (-1.0, .50) \quad M = (-.5, 1.00)\]
\[D = (-.8, .40) \quad N = (-.3, .80)\]
\[E = (-.9, .60) \quad O = (-.4, 1.10)\]
\[F = (-.7, .50) \quad P = (-.2, .85)\]
\[G = (-.8, .40) \quad Q = (-.3, 1.20)\]
\[H = (-.6, .60) \quad R = (-.2, 1.30)\]
\[I = (-.7, .80) \quad S = (-.1, .90)\]
\[J = (-.5, .70) \quad T = (-.1, 1.40)\]

$$(a_1, a_2)$$

**Figure 1a**

NORMAL DISTRIBUTION, \( m = 3, \Delta = .5 \)

42
Figure 2a

CAUCHY DISTRIBUTION, \( \Delta = .1 \)
\( e_0 = e_1 = .22 \)

\[ E^* = ([a_1, a_2], (-\infty, a_1) \cup (a_2, \infty)) \]

\[ E^* = ((-\infty, a_1) \cup (a_2, \infty), [a_1, a_2]) \]

\[ E^* = ((-\infty, a_1) \cup (a_2, \infty), (a_1, a_2)) \]

\[ E = ((-\infty, a), (a, \infty)) \]

---

**Corresponding Values of \( \lambda, a_1 \) and \( a_2 \)**

<table>
<thead>
<tr>
<th>( \lambda )</th>
<th>( a_1 )</th>
<th>( a_2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>.82</td>
<td>-9.99</td>
<td>-.12518</td>
</tr>
<tr>
<td>.84</td>
<td>-11.39</td>
<td>-.10974</td>
</tr>
<tr>
<td>.86</td>
<td>-13.19</td>
<td>-.09476</td>
</tr>
<tr>
<td>.88</td>
<td>-15.59</td>
<td>-.08020</td>
</tr>
<tr>
<td>.90</td>
<td>-18.93</td>
<td>-.06602</td>
</tr>
<tr>
<td>.92</td>
<td>-23.95</td>
<td>-.05220</td>
</tr>
<tr>
<td>.94</td>
<td>-32.29</td>
<td>-.03871</td>
</tr>
<tr>
<td>.96</td>
<td>-48.97</td>
<td>-.02552</td>
</tr>
<tr>
<td>.98</td>
<td>-98.98</td>
<td>-.01263</td>
</tr>
<tr>
<td>1.00</td>
<td>-( \infty ), 0</td>
<td></td>
</tr>
</tbody>
</table>

---

**Figure 2b**

CAUCHY DISTRIBUTION, \( \Delta = .5 \)

44
\[ \lambda \quad (a_1, a_2) \]

\[ \frac{\lambda}{I_0} \]

\[ 1.2 \]

\[ 1.1 \]

\[ .15 \]

\[ (-5.43, -1.3353) \]

\[ .20 \]

\[ (-6.86, -1.1399) \]

\[ .25 \]

\[ (-7.35, -.9869) \]

\[ .30 \]

\[ (-8.43, -.8605) \]

\[ .35 \]

\[ (-9.63, -.7527) \]

\[ .40 \]

\[ (-11.01, -.6586) \]

\[ .45 \]

\[ (-12.61, -.5751) \]

\[ .50 \]

\[ (-14.50, -.5000) \]

\[ .55 \]

\[ (-19.63, -.3693) \]

\[ .60 \]

\[ (-28.08, -.2582) \]

\[ .65 \]

\[ (-44.84, -.1617) \]

\[ .70 \]

\[ (-94.92, -.0764) \]

\[ 1 \]

\[ (-\infty, 0) \]

\[ \Delta_0 = \Delta_1 = 1.9 \]

\[ E^* = ([a_1, a_2], (-\infty, a) U (a_2, \infty)) \]

\[ E^* = ((a_1, a_2), (-\infty, a) U [a_2, a]) \]

\[ E^* = ((-\infty, a_1) U (a_2, \infty), [a_1, a_2]) \]

or \[ E^* = ((-\infty, a_1) U [a_2, \infty), (a_1, a_2)) \]

Figure 2c

CAUCHY DISTRIBUTION, \[ \Delta = 2.5 \]
BIBLIOGRAPHY


[5] H. Chernoff, "Classification Procedures of the Form $\Sigma \varphi(X_i) \leq k$ Where $\varphi(X)$ Assumes Only $m$ Successive Integral Values," Technical Report No. 5, prepared under Contract N6onr-25140 for Office of Naval Research, October 19, 1951.


<table>
<thead>
<tr>
<th>Position</th>
<th>Name</th>
<th>Address 1</th>
<th>Address 2</th>
<th>Address 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Commanding Officer</td>
<td>Armed Services Technical Information Agency</td>
<td>Arlington Hall Station</td>
<td>Arlington, Virginia</td>
<td>1</td>
</tr>
<tr>
<td>Commanding Officer</td>
<td>Office of Naval Research</td>
<td>Navy Br. 100, Fleet P.O.</td>
<td>New York, N.Y.</td>
<td>3</td>
</tr>
<tr>
<td>Commanding Officer</td>
<td>Office of Naval Research</td>
<td>1000 G Street</td>
<td>San Francisco 9, California</td>
<td>1</td>
</tr>
<tr>
<td>Commanding Officer</td>
<td>Office of Naval Research</td>
<td>2nd Floor, The John Creech Library Bldg.</td>
<td>Chicago 10, Illinois</td>
<td>1</td>
</tr>
<tr>
<td>Commanding Officer</td>
<td>Office of Naval Research</td>
<td>345 Broadway</td>
<td>New York 13, N.Y.</td>
<td>1</td>
</tr>
<tr>
<td>Commanding Officer</td>
<td>Diamond Ordnance Fuze Labs.</td>
<td>Washington 25, D.C.</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Commanding Officer</td>
<td>Picatinny Arsenal</td>
<td>Dover, New Jersey</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Commanding Officer</td>
<td>Watertown Arsenal 2nd</td>
<td>Watertown, Massachusetts</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Commanding Officer</td>
<td>Watertown Arsenal</td>
<td>Watertown, Massachusetts</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Commanding Officer</td>
<td>Watervilet Arsenal</td>
<td>New York</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Commanding Officer</td>
<td>Attn: Inspection Division</td>
<td>Springfield Armory</td>
<td>Springfield, Massachusetts</td>
<td>1</td>
</tr>
<tr>
<td>Commanding Officer</td>
<td>Signal Corps Electronics Research Unit</td>
<td>9500 Technical Service Unit</td>
<td>Mountain View, California</td>
<td>1</td>
</tr>
<tr>
<td>Commanding Officer</td>
<td>Army Liaison Group, Project Mich</td>
<td>2500 Technical Service Unit</td>
<td>Fort Huachuca, Arizona</td>
<td>1</td>
</tr>
<tr>
<td>Commanding Officer</td>
<td>Engineering Research &amp; Development Labs.</td>
<td>Fort Belvoir, Virginia</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Commanding Officer</td>
<td>Franklin Mint</td>
<td>Library Branch, GOFO, Bldg. 40</td>
<td>Philadelphia 37, Pennsylvania</td>
<td>10</td>
</tr>
<tr>
<td>Commanding General</td>
<td>Rock Island Arsenal</td>
<td>Rock Island, Illinois</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Commanding General</td>
<td>Redstone Arsenal (GREDW-GS)</td>
<td>Huntsville, Alabama</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Commanding General</td>
<td>White Sands Proving Ground</td>
<td>Las Cruces, New Mexico</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Commanding General</td>
<td>Signal Corps Engineering Laboratory</td>
<td>Fort Monmouth, New Jersey</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Commanding General</td>
<td>Research Branch</td>
<td>Rock Island, Illinois</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Commander</td>
<td>Wright Air Development Center</td>
<td>Wright-Patterson Air Force Base, Ohio</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Commander</td>
<td>Western Development Division, WDET</td>
<td>Ingleside, California</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Chief, Research Division</td>
<td>Office of Research &amp; Development</td>
<td>U.S. Army</td>
<td>Washington 25, D.C.</td>
<td>1</td>
</tr>
<tr>
<td>Chief, Computing Laboratory</td>
<td>Ballistic Research Laboratory</td>
<td>Aberdeen Proving Ground, Maryland</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Director</td>
<td>National Security Agency</td>
<td>Fort George G. Meade, Maryland</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Director</td>
<td>Operations Analysis Div., AFOSR</td>
<td>Wright-Patterson Air Force Base</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Director</td>
<td>Snow, Ice &amp; Permafrost Research Establishment</td>
<td>Wright-Patterson Air Force Base</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Director</td>
<td>South Dakota State University</td>
<td>Brookings, South Dakota</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Director</td>
<td>Lincoln Laboratory</td>
<td>Lexington, Massachusetts</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Department of Mathematics</td>
<td>Michigan State University</td>
<td>East Lansing, Michigan</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Headquarters</td>
<td>Oklahoma City Air Materiel Area</td>
<td>United States Air Force</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Institute of Statistics</td>
<td>North Carolina State College</td>
<td>Raleigh, North Carolina</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Jet Propulsion Laboratory</td>
<td>California Institute of Technology</td>
<td>Pasadena 7, California</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Librarian</td>
<td>The RAND Corporation</td>
<td>1700 Main Street</td>
<td>Santa Monica, California</td>
<td>1</td>
</tr>
<tr>
<td>Library Division</td>
<td>Naval Missile Center Command</td>
<td>U.S. Naval Missile Center</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Naval Inspector of Ordnance</td>
<td>U.S. Naval Gun Factory</td>
<td>Washington 25, D.C.</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Office, Attn: Chief of Staff, G-1</td>
<td>Department of the Army</td>
<td>Washington 25, D.C.</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Office of Technical Services</td>
<td>Department of Commerce</td>
<td>Washington 25, D.C.</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Technical Information Officer</td>
<td>Naval Research Laboratory</td>
<td>Washington 25, D.C.</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Technical Director</td>
<td>Naval Research Laboratory</td>
<td>Washington 25, D.C.</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>Technical Information Service</td>
<td>Reference Branch</td>
<td>Fort Huachuca, Arizona</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Technical Library Branch</td>
<td>U.S. Naval Ordnance Laboratory</td>
<td>Attn: Clayborn Graves</td>
<td>Corona, California</td>
<td>1</td>
</tr>
</tbody>
</table>

March 31, 1961
JOINT SERVICES ADVISORY GROUP

Dr. Merle M. Andrew, Chief Mathematics
Division Air Force Office of
Scientific Research
Washington 25, D. C. 1

Mr. James J. Fleming, Head
Operational Research Branch
U. S. Naval Research Laboratory
Washington 25, D. C. 1

Mr. Fred Frishman, Chairman
Army Research Office
Arlington Hall Station
Arlington, Virginia 1

Mrs. Dorothy M. Gilford
Logistics and Mathematical
Statistics Branch
Office of Naval Research
Washington 25, D. C. 3

Dr. Robert Lundegard
Logistics and Mathematical
Statistics Branch
Office of Naval Research
Washington 25, D. C. 1

Dr. Clifford Maloney
Applied Sciences Division
Chemical Corps, U. S. Army
Fort Detrick, Maryland 1

Mr. R. H. Noyes
Office of Technical Plans, USASRDLC
Fort Monmouth, New Jersey 1

Major Oliver A. Shaw, Jr.
Office of Scientific Research
Air Force - Room 2718, Temp. X
Washington 25, D. C. 2

Dr. Horace M. Trent, Head
Applied Mathematics Branch
U. S. Naval Research Laboratory
Washington 25, D. C. 1

Mr. J. Weinstein
Institute for Exploratory Research
USASRDLC
Fort Monmouth, New Jersey 1