

AD-A052 811

FLORIDA STATE UNIV TALLAHASSEE DEPT OF STATISTICS F/G 12/1  
A COMPARISON OF TWO SAMPLING SCHEMES WHEN TESTING A SIMPLE HYPO--ETC(U)  
AUG 77 J LYNCH, J SETHURAMAN DAAG29-76-G-0238

UNCLASSIFIED

FSU-TR-424

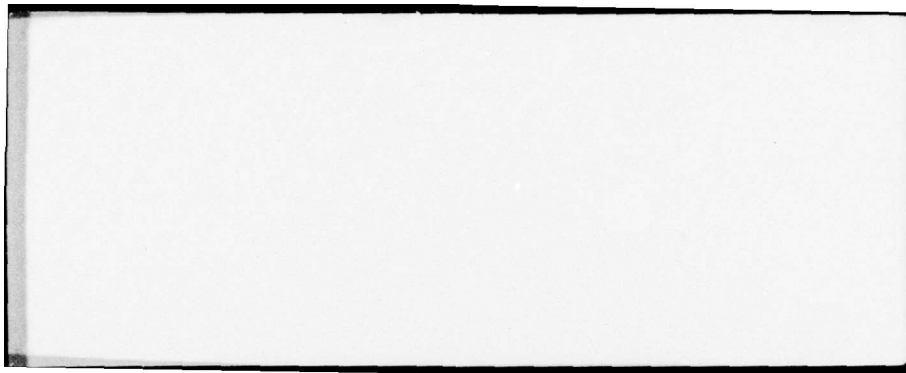
ARO-13888.9-MX

NL

1 OF 1  
ADA  
052811



END  
DATE  
FILMED  
5 - 78  
DDC



5

ABSTRACT for	
NTIS	White Section <input checked="" type="checkbox"/>
ODG	Buff Section <input type="checkbox"/>
UNANNOUNCED	<input type="checkbox"/>
JUSTIFICATION	
BY	
DISTRIBUTION/AVAILABILITY CODE	
Dist.	AVAIL. and/or SPECIAL
A	

6 A COMPARISON OF TWO SAMPLING SCHEMES WHEN TESTING A SIMPLE HYPOTHESIS VERSUS A SIMPLE ALTERNATIVE .

by

10 James Lynch and Jayaram Sethuraman

The Pennsylvania State University

9 Technical rept.,

and 11 Aug 77

12 2pp.

The Florida State University

14 FSU-TR-424, TR-23

FSU Technical Report No. 424  
USARO-D Technical Report No. 23

15 DAA29-76-G-0238, PHS-5-T01-GM0-0913

Key Words: Large deviations, Bayes risk, asymptotic Bayes cost ratio, tests of hypotheses, likelihood ratio, random number of observations.

AMS 1970 subject classification numbers

60F10, 62F20, 62F05, 62F15.

18 ARO

19 13889.9-MX

<sup>1</sup> Research sponsored by the United States Army Research Office, Durham, under Grant No. DAA29-76-G-0238

<sup>2</sup> Research sponsored by a grant from the National Institutes of Health No. 5-T01-GM00913.

400277

DISTRIBUTION STATEMENT A  
Approved for public release;  
Distribution Unlimited

DDC  
RECEIVED  
APR 18 1978  
RECEIVED

D  
MT/2mc

### Abstract

A large deviation result is established for sequences of random variables with random indices. This result is used to compare two different sampling schemes when testing for a simple hypothesis versus a simple alternative. It turns out that the sampling scheme yielding more expected number of observations for the same sampling cost is not necessarily the more profitable procedure. Properties other than the mean of the distribution of the occurrence of the observations play a role in determining the more profitable sampling rule.

#### 1. Introduction and summary.

Consider a situation where observations occur randomly over time and the statistician has to choose between two schemes by which he can pay for the observations. In scheme I he must pay  $c_1$  per observation while in scheme II he must pay  $c_2$  per unit time and collect the random number of observations that fall in that period of time. In this paper we compare the profitability of these schemes, when the purpose of the statistician is to test a hypothesis. Since  $c_1$  and  $c_2$  are costs for doing different things one cannot come up with a simple number which can be called the asymptotic efficiency. Instead, we propose, in Section 5, a cost ratio based on Bayesian considerations. This Asymptotic Bayes Cost Ratio (ABCR(I,II)) is the limiting ratio of the two costs of sampling which achieve equal Bayes risks when the best available testing procedure and best sample size or stopping time are used in each case. Thus, when the actual cost ratio  $c_1/c_2$  is less than ABCR(I,II), scheme I is more profitable and when  $c_1/c_2$  exceeds ABCR(I,II), scheme II is more profitable.

It turns out that the scheme that produces more expected number of observations for the same cost is not necessarily the more profitable scheme. Properties other than the mean of the occurrence of the observations play a role in determining the more profitable scheme.

The calculation of efficiency of tests leads quite naturally to deviation theory. For example, the computation of Bayes risk efficiency as formulated by Chernoff (1952) or the Bahadur efficiency (Bahadur (1960, (1967), and (1971))) requires results from large deviation theory, while the computation of Bayes risk efficiency as formulated by Rubin and Sethuraman (1965a) and (1965b) requires moderate deviation theory. This is also characteristic of the ABCR(I,II). In Section 4, a large deviation result for sequences of random variables with random indices, required in the computation of ABCR(I,II), is established.

We conclude the paper with an example. The ABCR(I,II) is computed when the observations occur according to a Poisson process, and have a common normal or Laplace distribution.

## 2. Preliminaries.

Let  $X_1, X_2, \dots$  be i.i.d. with distribution  $P_\theta$  determined by the parameter  $\theta$ . For  $n = 1, 2, \dots$ , let  $(X_1, \dots, X_n) = X_{(n)}$  and let  $X_{(0)}$  denote the event that no observations are taken. We assume that the observations occur randomly over time, i.e., the number which occur up until time  $t$ , say  $N_t$ , is a random variable. For simplicity and convenience we also assume  $N_t$  is independent of the observations.

Consider the problem of testing the simple hypothesis  $H_0: \theta = \theta_0$  against the simple alternative  $H_1: \theta = \theta_1$ . To avoid trivialities, we assume that  $P_{\theta_0} \neq P_{\theta_1}$  and that  $P_{\theta_0}$  and  $P_{\theta_1}$  are not mutually singular, i.e., there does not

exist a set  $A$  such that  $P_{\theta_0}(A) = 1 = P_{\theta_1}(A^c)$ . A statistical test of  $H_0$  against  $H_1$  is defined as follows.

Definition. Given a statistic  $T$ , (possibly multidimensional), a function  $\phi$  is said to be a test of  $H_0$  against  $H_1$  based on  $T$  if

- (i)  $\phi$  is a measurable function of  $T$ ,
- (ii)  $0 \leq \phi \leq 1$ , and
- (iii) the test consists of rejecting  $H_0$  with probability  $\phi(T)$  when  $T$  is observed.

The type I and type II error probabilities are

$$\alpha(\phi) = E_{\theta_0}(\phi(T)) \text{ and } \beta(\phi) = E_{\theta_1}(1-\phi(T)),$$

respectively.

We assume that a loss  $k_1(k_2)$  ( $0 < k_i < \infty$ ,  $i = 1, 2$ ) is incurred for a type I (II) error. We place a prior on  $\{\theta_0, \theta_1\}$  by assigning masses  $\pi$  and  $1 - \pi$  ( $0 < \pi < 1$ ) to  $\theta_0$  and  $\theta_1$ , respectively. Then the Bayes loss of the test  $\phi$  is given by

$$(2.1) \quad B(\phi) = \pi k_1 \alpha(\phi) + (1-\pi) k_2 \beta(\phi),$$

and the Bayes risk, which is the sum of loss and observation cost, is given by

$$(2.2) \quad M(\phi, c_T) = c_T + B(\phi),$$

where  $c_T$  denotes the cost of observing the statistic  $T$ .

For testing  $H_0$  against  $H_1$ , we assume the following two sampling schemes are both available.

Scheme I. Take  $n$  observations at a cost  $c_1$  ( $0 < c_1 < \infty$ ) per observation.

Scheme II. Take  $N_t$  (a random variable) observations by observing till some fixed time  $t$  at a cost  $c_2$  ( $0 < c_2 < \infty$ ) per unit time.

The quantities  $c_1$  and  $c_2$  are referred to as the unit costs. It is easily seen that the Bayes risks of a test  $\phi_I$  based on  $X_{(n)}$  (i.e. Scheme I) and a test  $\phi_{II}$  based on  $X_{(N_t)}$  (i.e. Scheme II) are

$$c_1 n + B(\phi_I) \text{ and } c_2 t + B(\phi_{II}),$$

respectively.

Let  $f_1$  and  $f_2$  denote the density functions of  $P_{\theta_1}$  and  $P_{\theta_2}$ , respectively, with respect to some common  $\sigma$ -finite measure  $\nu$  (which is easy to construct.)

For  $n > 0$ , let

$$(2.3) \quad \phi_n(X_{(n)}) = \begin{cases} 1 & \text{if } \sum_{i=1}^n \log[f_1(X_i)/f_2(X_i)] \geq \log[k_1\pi/k_2(1-\pi)] \\ 0 & \text{otherwise} \end{cases}$$

and for  $n = 0$ , let

$$\phi_0(X_{(0)}) = \begin{cases} 1 & \text{if } k_1\pi \geq k_2(1-\pi) \\ 0 & \text{otherwise.} \end{cases}$$

It is well known that for a fixed number of observations  $n$  or for a random number,  $N_t$ , of observations obtained by observing till time  $t$ , the Bayes loss is minimized by basing the test procedure on the likelihood ratio, for instance see Ferguson ((1967), p. 292). Thus

$$(2.4) \quad \inf\{B(\phi) : \phi \text{ a test based on } X_{(n)}\} = B(\phi_n)$$

and

$$(2.5) \quad \inf\{B(\phi) : \phi \text{ a test based on } X_{(N_t)}\} = B(\phi_{N_t}).$$

In the next section we compare tests based on the competing sampling schemes I and II, on the basis of their Bayes risks. In each case we choose the best available test and the optimal sample size or the optimal stopping time. In view of (2.4) and (2.5), we can restrict our attention exclusively to likelihood ratio tests.

### 3. Asymptotic Bayes Cost Ratio.

Let  $\phi_n$  ( $n=0,1,\dots$ ) be as defined in (2.3). Let

$$B_I(c_1) = \inf_n \{c_1 n + B(\phi_n)\}$$

and

$$B_{II}(c_2) = \inf_t \{c_2 t + B(\phi_{N_t})\}.$$

The quantities  $B_I(c_1)$  and  $B_{II}(c_2)$  correspond, respectively, to the risks involved when the sample size  $n$  for Scheme I and the stopping time  $t$  for Scheme II are chosen optimally. In typical cases,  $B_I(c)$  and  $B_{II}(c)$  both tend to 0 as  $c$  tends to 0. When this is the case, we define the Bayes risk efficiency as follows.

Definition. Let  $c_1(\cdot)$  be a function of  $c_2$  such that  $B_I(c) \leq B_{II}(c_2)$  for  $0 < c < c_1(c_2)$  and  $B_I(c_1(c_2)) \geq B_{II}(c_2)$ . Then the asymptotic Bayes cost ratio (ABCR) of Scheme II relative to Scheme I is

$$ABCR(I,II) = \lim_{c_2 \rightarrow 0} c_1(c_2)/c_2$$

when this limit exists.

To put the above definition into more applicable terms, the Bayes risk criterion suggests that it is more profitable for the statistician to use Scheme I when the ratio of the unit costs  $c_1/c_2 \leq ABCR(I,II)$  and to use



Scheme II otherwise. Of course, we are assuming that the unit costs are both relatively small.

When both  $c_1$  and  $c_2$  are small, the statistician is likely to observe a large number of observations since other costs remain fixed. One might, as a first guess, say that the sampling scheme which yields more expected number of observations for the same sampling cost is more profitable. For instance, suppose under Scheme II that  $E(N_t/t) \rightarrow \mu$  as  $t \rightarrow \infty$ , where  $\mu$  is some fixed constant. If the statistician observes until time  $t$  ( $t$  large), he will observe approximately  $\mu t$  observations at a sampling cost of  $tc_2$ . Under Scheme I, he can observe  $tc_2/c_1$  observations for the same sampling cost. Thus, one might guess that Scheme II would be more profitable than Scheme I if  $\mu t > tc_2/c_1$ , i.e. if  $c_1/c_2 > 1/\mu$ . This may not be always correct. We show in Theorem 5.1 that distribution properties of  $N_t$  other than the behavior of  $E(N_t)/t$  as  $t \rightarrow \infty$  are involved. In fact, Theorem 5.1 states that under certain conditions,

$$ABCR(I,II) = -\rho/J(-\rho),$$

where  $\rho$  depends on the distribution of  $X_1$  through its m.g.f. and  $J$  depends on the distribution of  $N_t$  through its m.g.f.. However, from Lemma 5.1, wherein it is shown that  $ABCR(I,II) \geq 1/\mu$ , it follows that Scheme I is more profitable than Scheme II if  $c_1/c_2 \leq 1/\mu$ .

#### 4. Large deviations.

Let  $\{T_n\}$  be a sequence of random variables, defined on some probability space  $(X, \mathcal{B}, P)$ , for which  $T_n \rightarrow \mu$  in probability as  $n \rightarrow \infty$ . Let  $A$  be a Borel measurable set such that the closure of  $A$  does not contain  $\mu$ . Then  $P(T_n \in A) \rightarrow 0$

Scheme II otherwise. Of course, we are assuming that the unit costs are both relatively small.

When both  $c_1$  and  $c_2$  are small, the statistician is likely to observe a large number of observations since other costs remain fixed. One might, as a first guess, say that the sampling scheme which yields more expected number of observations for the same sampling cost is more profitable. For instance, suppose under Scheme II that  $E(N_t/t) \rightarrow \mu$  as  $t \rightarrow \infty$ , where  $\mu$  is some fixed constant. If the statistician observes until time  $t$  ( $t$  large), he will observe approximately  $\mu t$  observations at a sampling cost of  $tc_2$ . Under Scheme I, he can observe  $tc_2/c_1$  observations for the same sampling cost. Thus, one might guess that Scheme II would be more profitable than Scheme I if  $\mu t > tc_2/c_1$ , i.e. if  $c_1/c_2 > 1/\mu$ . This may not be always correct. We show in Theorem 5.1 that distribution properties of  $N_t$  other than the behavior of  $E(N_t)/t$  as  $t \rightarrow \infty$  are involved. In fact, Theorem 5.1 states that under certain conditions,

$$ABCR(I,II) = -\rho/J(-\rho),$$

where  $\rho$  depends on the distribution of  $X_1$  through its m.g.f. and  $J$  depends on the distribution of  $N_t$  through its m.g.f.. However, from Lemma 5.1, wherein it is shown that  $ABCR(I,II) \geq 1/\mu$ , it follows that Scheme I is more profitable than Scheme II if  $c_1/c_2 \leq 1/\mu$ .

#### 4. Large deviations.

Let  $\{T_n\}$  be a sequence of random variables, defined on some probability space  $(X, \mathcal{B}, P)$ , for which  $T_n \rightarrow \mu$  in probability as  $n \rightarrow \infty$ . Let  $A$  be a Borel measurable set such that the closure of  $A$  does not contain  $\mu$ . Then  $P(T_n \in A) \rightarrow 0$

as  $n \rightarrow \infty$ . In typical cases  $P(T_n \in A) \rightarrow 0$  at an exponential rate, i.e., there exists a positive number  $I_A$  such that

$$(4.1) \quad \lim_{n \rightarrow \infty} \frac{1}{n} \log P(T_n \in A) = -I_A.$$

When (4.1) holds, we say  $A$  is a large deviation event and call  $I_A$  the index of large deviation of  $\{T_n \in A\}$  or of  $T_n$  if the set  $A$  is otherwise understood.

The earliest research in the theory of large deviations dealt primarily with sums of independent random variables. Most notable is the following theorem due to Chernoff.

Theorem 4.1. (Chernoff (1952), p. 494). Let  $Y_1, Y_2, \dots$  be a sequence of i.i.d. random variables with mean  $\mu < \infty$ . ( $\mu$  may equal  $-\infty$ ). Then for  $T_n = \sum_{i=1}^n Y_i/n$ ,

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log P(T_n \geq a) = -I_{[a, \infty)},$$

where

$$-I_{[a, \infty)} = \inf\{-\lambda a + \log E(e^{\lambda Y_1}) : \lambda \geq 0\}.$$

As a consequence of the above theorem, Chernoff was able to determine the exponential rate at which  $B(\phi_n)$  goes to zero. This is given in the following corollary.

Let

$$(4.2) \quad -\rho = \log \inf\{g(\lambda) : 0 < \lambda < 1\},$$

where

$$g(\lambda) = \int [f_1(x)]^\lambda [f_0(x)]^{1-\lambda} d\nu(x).$$

Then,

Corollary 4.2. Let  $B(\phi_n)$  be as defined in (2.4). Then,

$$(4.3) \quad \lim_{n \rightarrow \infty} \frac{1}{n} \log B(\phi_n) = -\rho.$$

Remark. Under our assumption that  $P_0$  and  $P_1$  are not mutually singular, it should be noted that  $B(\phi_n) \rightarrow 0$  as  $n \rightarrow \infty$  neither too fast nor too slow; i.e.,  $0 < \rho < \infty$ .

To suitably approximate the Bayes risk for Scheme II, we need a result analogous to Corollary 4.2 for  $B(\phi_{N_t})$ . This will follow from Corollary 4.4 below which is a consequence of Theorem 4.3.

Theorem 4.3. Let  $\{a_n\}$  be a sequence of non-negative integers such that

$$(4.4) \quad \lim_{n \rightarrow \infty} (1/n) \log a_n = -I,$$

where  $0 < I < \infty$ . Let  $\psi_t(\lambda) = E(e^{\lambda N_t})$  and let

$$(4.5) \quad \lim_{t \rightarrow \infty} (1/t) \log \psi_t(\lambda) = J(\lambda)$$

exist and be continuous and strictly increasing in a neighborhood of  $-I$ . Let  $h(t) = \sum_{n=0}^{\infty} a_n P(N_t = n)$ . Then

$$(4.6) \quad \lim_{t \rightarrow \infty} (1/t) \log h(t) = J(-I).$$

Proof. Let  $\epsilon > 0$ . From (4.4),

$$e^{(-I-\epsilon)k} \leq a_k \leq e^{(-I+\epsilon)k}$$

for all  $k \geq k(\epsilon)$ . Thus

$$(4.7) \quad \psi_t(-I-\epsilon) - \sum_1^{k(\epsilon)} e^{m(-I-\epsilon)} P(N_t = m) \leq h(t) \leq \psi_t(-I+\epsilon) + a(\epsilon) P(N_t \leq k(\epsilon)),$$

where  $a(\epsilon) = \max\{a_k: 0 \leq k \leq k(\epsilon)\}$ . Assume that  $I > 0$  and choose  $\epsilon$  such that  $I - \epsilon > 0$ . From a Chebycheff inequality for  $P\{N_t \leq k(\epsilon)\}$ ,

$$h(t) \leq \psi_t(-I+\epsilon) [1 + a(\epsilon) e^{k(\epsilon)(I-\epsilon)}]$$

and

$$\overline{\lim}_t (1/t) \log h(t) \leq J(-I+\epsilon).$$

Since  $J$  is continuous at  $-I$ , we can allow  $\epsilon$  to decrease to 0 and obtain

$$(4.8) \quad \overline{\lim}_t (1/t) \log h(t) \leq J(-I).$$

If  $I = 0$ , (4.8) is immediate since  $J(0) = 0$ . Now, choose  $\epsilon_1$  to satisfy  $0 < \epsilon_1 < \epsilon$  and  $J(-I-\epsilon_1) - J(-I-\epsilon) = 3\alpha > 0$ . From (4.5), there is a  $t_0$  such that for  $t \geq t_0$

$$e^{t[J(-I-\epsilon_1)-\alpha]} \leq \psi_t(-I-\epsilon_1) \leq \psi_t(-I-\epsilon) \leq e^{t[J(-I-\epsilon)+\alpha]}.$$

Thus, for  $t \geq t_0$ , it follows that

$$\begin{aligned} h(t) &\geq \psi_t(-I-\epsilon_1) - \sum_{m=0}^{k(\epsilon_1)} e^{m(-I-\epsilon_1)} P(N_t = m) \\ &\geq \psi_t(-I-\epsilon_1) - e^{k(\epsilon_1)(\epsilon-\epsilon_1)} \sum_{m=0}^{k(\epsilon)} e^{m(-I-\epsilon)} P(N_t = m) \\ &\geq \psi_t(-I-\epsilon_1) - e^{k(\epsilon_1)(\epsilon-\epsilon_1)} \psi_t(I-\epsilon) \\ &\geq e^{t[J(-I-\epsilon_1)-\alpha]} - e^{k(\epsilon_1)(\epsilon-\epsilon_1) + t[J(-I-\epsilon)+\alpha]} \\ &\geq e^{t[J(-I-\epsilon_1)-\alpha]} [1 + o(1)] \end{aligned}$$

as  $t \rightarrow \infty$ , thus establishing

$$\lim_{t \rightarrow \infty} (1/t) \log h(t) \geq J(-I - \epsilon_1) - \alpha.$$

Since  $J$  is continuous at  $-I$ , we may allow  $\epsilon \rightarrow 0$  which entails  $\epsilon_1 \rightarrow 0$  and  $\alpha \rightarrow 0$  and we therefore obtain

$$(4.9) \quad \lim_{t \rightarrow \infty} (1/t) \log h(t) \geq J(-I).$$

Relations (4.8) and (4.9) established Theorem 4.2.  $\square$

Corollary 4.4. Let  $\rho$  be as defined in (4.2). Let  $J(\lambda)$  as defined in (4.5) exist, be continuous and strictly increasing in a neighborhood of  $-\rho$ . Then

$$(4.10) \quad \lim_{t \rightarrow \infty} (1/t) \log B(\phi_{N_t}) = J(-\rho).$$

Proof. Since it was assumed that  $N_t$  is independent of  $X_1, X_2, \dots$ , we have

$$B(\phi(N_t)) = \sum_{m=0}^{\infty} B(\phi_m) P(N_t = m).$$

Corollary 4.4 follows immediately from Theorem 4.2.  $\square$

Remark. Let  $\{N_t, t \geq 0\}$  be a non homogeneous Poisson process with intensity function  $\mu(t)$  satisfying

$$(4.11) \quad M(t)/t \rightarrow \mu, \text{ where } M(t) = \int_0^t \mu(s) ds.$$

Then,

$$\psi_t(\lambda) = E(e^{\lambda N_t}) = e^{M(t)(e^\lambda - 1)},$$

and condition (4.6) of Theorem 4.2 is satisfied, with

$$(4.12) \quad J(\lambda) = (e^\lambda - 1)\mu.$$

## 5. General results on ABCR(I,II).

In the remainder of this paper, the following assumption is made concerning the process  $\{N_t: t > 0\}$ .

Assumption A. Let  $\rho$  be as defined in (4.2). There is a function  $J$  which is continuous, strictly increasing and satisfies (4.5) in some neighborhood of  $-\rho$ .

Theorem 5.1. Let assumption A hold. Then

$$(5.1) \quad \text{ABCR(I,II)} = -\rho/J(-\rho).$$

Proof. The theorem follows immediately from the asymptotic rates of  $B_I(c)$  and  $B_{II}(c)$  in (5.2) and (5.3) below, which we proceed to establish:

$$(5.2) \quad B_I(c) \sim -c \log c/\rho, \text{ and}$$

$$(5.3) \quad B_{II}(c) \sim c \log c/J(-\rho),$$

where  $f(c) \sim g(c)$  denotes  $f(c)/g(c) \rightarrow 1$  as  $c \rightarrow 0$ . We will prove only (5.3). The proof of (5.2) is analogous and will be omitted.

Since,  $0 < \rho < \infty$  by the Remark following Corollary 4.2, it follows that  $0 < -J(-\rho) < \infty$ . Choose  $\epsilon$  such that  $0 < \epsilon < -J(-\rho)$ . Recall that the Bayes risk for Scheme II with unit cost  $c$  and observing till time  $t$  is

$$M(\phi_{N_t}, ct) = ct + B(\phi_{N_t}).$$

Thus, it follows from (4.10) of Corollary 4.4 that there exists a  $t_0$  independent of  $c$  such that for  $t \geq t_0$

$$\begin{aligned}
M_{-\epsilon,c}(t) &= e^{t[J(-\rho)-\epsilon]} + ct \\
&\leq M(\phi_{N_t}, ct) \\
&\leq e^{t[J(-\rho)+\epsilon]} + ct = \bar{M}_{\epsilon,c}(t).
\end{aligned}$$

By differentiating  $M_{-\epsilon,c}(\cdot)$  and  $\bar{M}_{\epsilon,c}(\cdot)$ , it is seen that  $M_{-\epsilon,c}(\cdot)$  and  $\bar{M}_{\epsilon,c}(\cdot)$  attain their minimums at

$$t_c = -[J(-\rho)-\epsilon]^{-1} \log[-J(-\rho)-\epsilon]c^{-1}$$

and

$$\bar{t}_c = -[J(-\rho)+\epsilon]^{-1} \log[-J(-\rho)+\epsilon]c^{-1},$$

respectively. Since  $t_c$  and  $\bar{t}_c$  tend to infinity as  $c \rightarrow 0$ , it follows that

$$(5.4) \quad M_{-\epsilon,c}(t_c) \leq \inf\{M(\phi_{N_t}, ct) : t \geq t_0\} \leq \bar{M}_{\epsilon,c}(\bar{t}_c),$$

for all sufficiently small  $c$ . As  $c \rightarrow 0$ , both the extreme terms of (5.4) tend to zero. Thus

$$(5.5) \quad \inf\{M(\phi_{N_t}, ct) : t \geq t_0\} \rightarrow 0 \text{ as } c \rightarrow 0.$$

Since  $0 < J(-\rho) < \infty$ , it follows from (4.10) that  $B(\phi_{N_t}) > 0$  for all large  $t$ .

Since  $B(\phi_{N_t})$  is monotonic decreasing in  $t$  we can conclude that

$$\inf\{M(\phi_{N_t}, ct) : t < t_0\} \geq B(\phi_{N_{t_0}}) > 0$$

and, from (5.5), also that



$$(5.6) \quad \inf\{M(\phi_{N_t}, ct) : t \geq t_0\} = B_{II}(c)$$

if  $c$  is small. Since

$$\lim_{\epsilon} \frac{\lim_{c} M_{\epsilon, c}(t_c)/c \log c = [J(-\rho)]^{-1}}$$

and

$$\lim_{\epsilon} \frac{\lim_{c} M_{\epsilon, c}(\bar{t}_c)/c \log c = [J(-\rho)]^{-1}},$$

it follows from (5.4) and (5.6) that  $B_{II}(c) \sim c \log c / J(-\rho)$ . This completes the proof of Theorem 5.1.  $\square$

Remark. Let  $t_c \sim (\log c) / J(-\rho)$ . Then it will follow from the proof above and (4.10) that

$$M(\phi_{N_{t_c}}, ct_c) \sim B_{II}(c).$$

Thus for Scheme II with unit cost  $c$ ,  $t_c$  is an asymptotically optimal stopping time. It follows similarly that  $n_c \sim \log c / -\rho$  is an asymptotically optimal sample size for Scheme I with unit cost  $c$ .

The following two lemmas show the relationship of the ABCR(I,II) to the  $E(N_t/t)$  and  $N_t/t$  as  $t \rightarrow \infty$ .

Lemma 5.2. The

$$\text{ABCR}(I, II) \geq 1 / \lim_{t} E(N_t/t).$$

In fact, the above inequality is strict when  $J$  is strictly convex.

Proof. Let  $\epsilon > 0$ . Since  $\log \psi_t(\lambda)$  is the cumulant generating function of  $N_t$ , it is convex. Hence, for sufficiently small  $\epsilon$ ,

$$\begin{aligned}
\frac{1}{-\rho t} \log \psi_t(-\rho) &= \frac{1}{-\rho t} [\log \psi_t(-\rho) - \log \psi_t(0)] \\
&\leq \frac{1}{(-\rho+\epsilon)t} [\log \psi_t(-\rho+\epsilon) - \log \psi_t(0)] \\
&\leq \frac{1}{t} \frac{d}{d\lambda} \log \psi_t(\lambda) \Big|_{\lambda=0} = E(N_t/t).
\end{aligned}$$

The lemma follows by taking  $\lim \inf$ 's in the above and by observing that, when  $J(\lambda)$  is strictly convex,  $J(-\rho)/-\rho < J(-\rho+\epsilon)/(-\rho+\epsilon)$ .  $\square$

Lemma 5.3. If  $N_t/t \rightarrow \mu$  in probability as  $t \rightarrow \infty$ , then

$$\text{ABCR(I,II)} \geq 1/\mu \geq 1/\underline{\lim} E(N_t/t).$$

Proof. The right hand inequality follows readily from Fatou's lemma. By Jensen's inequality, for  $t \geq 1$ ,

$$\frac{1}{t} \log \psi_t(\lambda) \geq \log E(e^{\lambda N_t/t}).$$

By another application of Fatou's lemma

$$\lim_t t^{-1} \log \psi_t(\lambda) = J(\lambda) \geq \lambda \mu.$$

Thus  $-\rho/J(-\rho) \geq 1/\mu$ . This completes the proof of Lemma 5.3.  $\square$

Example. Let  $X_1, X_2, \dots$  be i.i.d. with mean  $\theta$  and variance  $\sigma^2$ . We now compute the  $\text{ABCR(I,II)}$  for testing  $H_0: \theta = 0$  against  $H_1: \theta = \theta_1$  when the observations occur according to a Poisson process  $\{N_t: t > 0\}$  with mean rate  $\mu$  and have a common normal or Laplace distribution. Then the  $\text{ABCR(I,II)}$  is

$$\text{ABCR(I,II)} = -\rho/J(-\rho),$$

where, from (4.2),

$$\rho = \theta_1^2 / (8\sigma^2) \text{ if the } X\text{'s are normal, and}$$

$$\rho = -\log(1 + \theta_1 / (\sigma\sqrt{2})) + \theta_1 / (\sigma\sqrt{2})$$

if the  $X$ 's are Laplace, and

$$J(\lambda) = \mu(e^\lambda - 1).$$

The table below lists  $\mu$  ABCR(I,II) for various values of  $\theta_1/\sigma$ . One interprets the table as follows: Scheme II is more or less profitable than Scheme I according to whether or not  $\mu c_1/c_2$  is greater than the appropriate entry in the table. One sees immediately from the table that as the alternative hypothesis recedes from the null hypothesis the unit cost  $c_2$  necessary for Scheme II to be as profitable as Scheme I with a fixed unit cost  $c_1$  must decrease. Also, as is expected from Lemma 5.2,  $\mu$  ABCR  $\geq 1$ .

Values of  $\mu$  ABCR(I,II), when under Scheme II observations according to a Poisson process with parameter  $\mu$  and when testing for

$\theta/\sigma$	$N(0,\sigma)$ versus $N(\theta,\sigma)^*$	$L(0,\sigma)$ versus $L(\theta,\sigma)^*$
.125	1.001	1.002
.250	1.004	1.007
.375	1.009	1.015
.5	1.016	1.026
.625	1.025	1.038
.75	1.036	1.053
.875	1.049	1.070
1	1.064	1.089
1.5	1.147	1.178
2	1.271	1.290
2.5	1.440	1.421
3	1.666	1.571
3.5	1.954	1.738
4	2.313	1.921
4.5	2.750	2.119
5	3.269	2.332

\* $N(\theta,\sigma)$  = Normal distribution with mean  $\theta$  and variance  $\sigma^2$ .  
 $L(\theta,\sigma)$  = Laplace distribution with mean  $\theta$  and variance  $\sigma^2$ .

#### REFERENCES

- Bahadur, R. R. (1960). Stochastic comparison of tests. Ann. Math. Statist. 31 276-295.
- Bahadur, R. R. (1967). Rates of convergence of estimates and test statistics. Ann. Math. Statist. 38 303-324.
- Bahadur, R. R. (1971). Some Limit Theorems in Statistics. SIAM, Philadelphia, Penn.
- Chernoff, H. (1952). A measure of asymptotic efficiency for tests of hypotheses based on the sum of observations. Ann. Math. Statist. 23 493-507.
- Ferguson, T. (1967) Mathematical Statistics: A Decision Theoretic Approach. Academic Press, New York.
- Rubin, H. and Sethuraman, J. (1965a) Probabilities of moderate deviations. Sankhyā. Series A 27 325-346.
- Rubin, H. and Sethuraman, J. (1965b). Bayes risk efficiency. Sankhyā Series A 27 347-356.

## UNCLASSIFIED

## SECURITY CLASSIFICATION OF THIS PAGE

## REPORT DOCUMENTATION PAGE

1. REPORT NUMBER USARO-D No. 23 ✓	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and subtitle)  A Comparison of Two Sampling Schemes When Testing a Simple Hypothesis versus a Simple Alternative.	5. TYPE OF REPORT & PERIOD COVERED  Technical Report	
	6. PERFORMING ORG. REPORT NO.  FSU Technical Report No. 424 ✓	
7. AUTHOR(s)  James Lynch and Jayaram Sethuraman	8. CONTRACT OR GRANT NUMBER(s)  USARO-DAA29-76- <sup>6</sup> 0238	
9. PERFORMING ORGANIZATION NAME AND ADDRESS  The Florida State University Department of Statistics ✓ Tallahassee, Florida 32306	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS	
11. CONTROLLING OFFICE NAME AND ADDRESS  U. S. Army Research Office Post Office Box 12211 Research Triangle Park, NC 27709	12. REPORT DATE  August, 1977	
	13. NUMBER OF PAGES  16	
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)	15. SECURITY CLASS (of this report)  Unclassified	
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE  NA	
16. DISTRIBUTION STATEMENT (of this Report)  Approved for public release; distribution unlimited.		
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)  NA		
18. SUPPLEMENTARY NOTES  The findings in this report are not to be construed as an official Department of the Army position, unless so designated by other authorized documents.		

---

REPORT DOCUMENTATION PAGE

---

19. KEY WORDS

Large deviations  
asymptotic Bayes cost ratio  
likelihood ratio

Bayes risk  
tests of hypotheses  
random number of observations

---

20. ABSTRACT

A large deviation result is established for sequences of random variables with random indices. This result is used to compare two different sampling schemes when testing for a simple hypothesis versus a simple alternative. It turns out that the sampling scheme yielding more expected number of observations for the same sampling cost is not necessarily the more profitable procedure. Properties other than the mean of the distribution of the occurrence of the observations play a role in determining the more profitable sampling rule.

---

