EFFICIENCY AND MINIMAXITY OF BAYES SEQUENTIAL PROCEDURES IN SIMPLE VERSUS SIMPLE HYPOTHESIS TESTING FOR A GENERAL NONREGULAR MODELS

by

Hyun Sook Oh and Anirban DasGupta
Purdue University

Technical Report #93-25

Department of Statistics
Purdue University

May 1993

◇

Efficiency and Minimaxity of Bayes Sequential Procedures in Simple versus Simple Hypothesis Testing for a General Nonregular Models

by
HYUN SOOK OH AND ANIRBAN DASGUPTA
Purdue University

May 17, 1993

1 Introduction

A sequential problem is considered in which independent observations are taken on a random variable X which is distributed as $Uniform(0, \theta)$, where the parameter $\theta > 0$ is an unknown constant. Suppose that we want to test

$$H_0 : \theta = \theta_0 \quad against \quad H_1 : \theta = \theta_1 \tag{1}$$

where $\theta_1 = 1$ and $\theta_0 > 1$. Let π_0 denote the prior probability for H_0 and let c be the constant cost of sampling. It is assumed that the decision loss is $0 - l_i$ loss, i.e., $L(\theta_0, a_0) = L(\theta_1, a_1) = 0, L(\theta_0, a_1) = l_1$ and $L(\theta_1, a_0) = l_0$. Let n denote the number of observations ultimately taken. It is assumed that the overall loss is

$$L(\theta_i, a_j, n) = L(\theta_i, a_j) + nc, \quad i = 0, 1, \;\; j = 0, 1.$$

The Bayes stopping time is, in general, of the form: Stop either at time 0 or at the first time n such that

$$\rho_0(\pi^n) \leq \rho^*(\pi^n)$$

where $\rho_0(\pi^n)$ is the posterior Bayes decision risk in the fixed sample size problem with a sample of size n and data $\mathbf{X}^n$ and $\rho^*(\pi^n)$ is the minimum Bayes risk that can be attained if at least $n+1$ observations are taken. (The "decision" risk does not include the cost of sampling c.) Notice that $\rho_0(\pi^n)$ does not involve the cost c while $\rho^*(\pi^n)$ does. It turns out that the Bayes procedure is just the immediate Bayes decision with no observation or a SPRT(sequential probability ratio test) which is of the following form:(Berger(1985))
At stage $n (n \geq 1)$,

if $L_n \leq A$, stop sampling and decide a_0;
if $L_n \geq B$, stop sampling and decide a_1;
if $A < L_n < B$, take another observation;

here, $A < 1$ and $B > 1$ are appropriate stopping boundaries and L_n is the likelihood ratio of θ_1 to θ_0 at stage n,

$$L_n = \frac{\prod_{i=1}^{n} f(x_i|\theta_1)}{\prod_{i=1}^{n} f(x_i|\theta_0)}.$$

(For the SPRT, as a nature of Bayes test, see Wald(1947), Wald and Wolfowitz (1948), Ferguson(1967) and Berger(1985).) Observe that

$$\begin{aligned} L_n &= \frac{I_{(x_i \leq 1, \forall i=1,\ldots,n)}(\theta)}{I_{(x_i \leq \theta_0, \forall i=1,\ldots,n)}(\theta)} \\ &= \begin{cases} 0 & \text{if } 1 \leq x_{(n)} < \theta_0 \\ \theta_0^n & \text{if } x_{(n)} < 1 \end{cases} \end{aligned} \tag{2}$$

where $x_{(n)}$ is the nth order statistic. So only two things can happen:
a either we get $x_n > 1$ at some stage; then, clearly, we should stop and accept H_0 or
b we keep getting $x_i's \leq 1$ and therefore since $\theta_0 > 1$, by (2) the ratio sooner or later goes above the fixed bound B and we will reject H_0. Thus Bayes rules must be one of the following rules:
d_0 : stop and take the optimal action without taking any observations; $d_J (J \geq 1)$: necessarily stop before $n \leq J$; if $\exists n < J \ni x_n \geq 1$, then stop at stage n and accept H_0, while if $\forall n < J, x_n < 1$, then continue sampling untill the Jth obsevation and accept H_0 if $x_J \geq 1$ and decide H_1 otherwise.

In section 2, efficiency of the Bayes sequential procedure with respect to the optimal fixed sample size Bayes procedure is considered. Let $r_{\pi_0}(c)$ be

Bayes sequential risk and let $r^F_{\pi_0}(c)$ be the optimal fixed sample size Bayes risk. Then it is proved that

$$\lim_{c\to 0}\frac{r^F_{\pi_0}(c)}{r_{\pi_0}(c)} = \frac{1}{1-\pi_0}.$$

On the otherhand, for a fixed value of π_0, for the Bayes sequential procedure and the optimal fixed sample size Bayes procedure to have the same Bayes risk, i.e., $r_{\pi_0}(c) = r^F_{\pi_0}(c^F)$, we prove that the ratio of the sampling costs satisfies

$$\lim_{c\to 0}\frac{c^F}{c} = 1-\pi_0.$$

Notice that the two definitions of efficiency are not equivalent. In section 3, the minimax sequential procedure is considered. The minimax sequential rule is determined among the set of Bayes rules, $d_J, J \geq 0$. This is justified, as established in Brown, Cohen and Strawderman(1980). It will be shown that the minimax sequential risk increases as the sampling cost increases while for minimax sequential strategy d_{J^m}, J^m decreases as c increases.

The results of this paper hold also for the general nonregular case in which independent observations have a common density of the form $b(\eta)h(x)I_{(x<\eta)}$ and (1) is replaced by

$$H_0 : \eta = \eta_0 \quad vs \quad H_1 : \eta = \eta_1$$

for some $\eta_0 > \eta_1$. This is easily seen on majing a transformation of the form $Y = 1/(\eta_1 b(X))$.

2 Asymptotic Efficiency of Bayes Sequential Procedure with respect to the Optimal Fixed Sample Size Bayes Procedure

Let $r(\pi_0, d_J)$ denote the Bayes risk for the procedure d_J is defined in section 1. Since a Bayes sequential rule is d_J, for some $J \geq 0$, for a given π_0, the Bayes sequential procedure is determined by minimizing $r(\pi_0, d_J)$ over $J \geq 0$.

For the procedure $d_J, J \geq 1$, (Berger(1985))

$$\begin{aligned}
\alpha_0 &= P_{\theta=\theta_0}(reject\ \ H_0) \\
&= P_{\theta=\theta_0}(x_i < 1, \forall i = 1, \ldots, J) \\
&= (\frac{1}{\theta_0})^J, \\
\alpha_1 &= P_{\theta=1}(accept\ \ H_0) \\
&= 1 - P_{\theta=1}(reject\ \ H_0) \\
&= 0.
\end{aligned}$$

Let N_J be a stopping time for the procedure $d_J, J \geq 1$. Then

$$\begin{aligned}
E(N_J|H_0) &= \sum_{n=0}^{J-1} P_{H_0}(N > n) \\
&= \sum_{n=0}^{J-1} P_{\theta=\theta_0}(x_i < 1, \forall i = 1, \ldots, n) \\
&= \sum_{n=0}^{J-1} \theta_0^{-n} \\
&= \frac{1-\theta_0^{-J}}{1-\theta_0^{-1}}, \\
E(N_J|H_1) &= J.
\end{aligned}$$

Thus, the Bayes sequntial risk for the procedure $d_J, J \geq 1$, is given by

$$r(\pi_0, d_J) = \pi_0(\theta_0^{-J} l_1 + c\frac{1-\theta_0^{-J}}{1-\theta_0^{-1}}) + (1-\pi_0)cJ.$$

Also, $r(\pi_0, d_0) = \min\{\pi_0 l_1, (1-\pi_0)l_0\}$. Let $f(J) = r(\pi_0, d_J), J \geq 1$. Pretending that J is a continuous variable and differentiating with respect to J gives

$$f'(J) = \pi_0(\frac{\theta_0}{\theta_0 - 1}c - l_1)\theta_0^{-J} \log \theta_0 + (1-\pi_0)c.$$

Thus $f'(J) > 0$ if $l_1 < \theta_0/(\theta_0 - 1)c$. So $J = 0$ is the optimal value for $l_1 < \theta_0/(\theta_0 - 1)c$. If $l_1 \geq \theta_0/(\theta_0 - 1)c$, the second derivative of $f(J)$ is positive, so $f(J)$ is strictly convex function in J. Setting $f'(J) = 0$ and solving gives the approximate optimal value of J which is

$$\frac{\log(\pi_0(l_1 - \frac{c}{1-1/\theta_0})\log\theta_0) - \log((1-\pi_0)c)}{\log\theta_0}.$$

Let $r_{\pi_0}(c)$ denote the Bayes sequential risk. Thus if $l_1 \geq \frac{\theta_0}{\theta_0 - 1}c$, since J^* is not an integer value,

$$\begin{aligned} r_{\pi_0}(c) &\approx f(J^*) \\ &= \pi_0(\theta_0^{-J^*} l_1 + c\frac{1-\theta_0^{-J^*}}{1-\frac{1}{\theta_0}}) + (1-\pi_0)cJ^*, \end{aligned} \tag{3}$$

where

$$J^* = \max\{0, \frac{\log(\pi_0(l_1 - \frac{c}{1-1/\theta_0})\log\theta_0) - \log((1-\pi_0)c)}{\log\theta_0}\}.$$

And if $l_1 < \frac{\theta_0}{\theta_0-1}c$,

$$r_{\pi_0}(c) = \min\{\pi_0 l_1, (1-\pi_0)l_0\}.$$

Now, let us consider the optimal fixed sample size procedure.
If $\mathbf{X}^n = (X_1, \ldots, X_n)$ is observed, the Bayes decision rule is to select a_0 if

$$\begin{aligned} & E^\pi E_\theta^x L(\theta, a_0) \leq E^\pi E_\theta^x L(\theta, a_1) \\ \Leftrightarrow\ & (1-\pi_0) l_0 I_{(x_{(n)}<1)} \leq \pi_0 l_1 (\frac{1}{\theta_0})^n I_{(x_{(n)}<\theta_0)}. \end{aligned}$$

Thus the Bayes decision rule is

$$\delta_\pi^n = \begin{cases} a_0 & \text{if } x_{(n)} > 1 \text{ or } (1-\pi_0)l_0 \leq \pi_0 l_1 (1/\theta_0)^n \\ a_1 & \text{otherwise.} \end{cases}$$

Let $r^n(\pi)$ denote the Bayes decision risk for δ_π^n. Then

$$\begin{aligned} r^n(\pi) &= E^\pi E_\theta^x L(\theta, \delta_\pi^n) \\ &= \pi_0 l_1 P_{\theta_0=\theta_0}(x_{(n)} < 1 \ \ and \ \ (1-\pi_0)l_0 \leq \pi_0 l_1 (1/\theta_0)^n) \\ &\quad + (1-\pi_0) l_0 (P_{\theta_1=1}(x_{(n)} \geq 1) \\ &\quad + P_{\theta_1=1}(x_{(n)} < 1 \ \ and \ \ (1-\pi_0)l_0 > \pi_0 l_1 (1/\theta_0)^n) \\ &= \begin{cases} (\frac{1}{\theta_0})^n \pi_0 l_1 & \text{if } n > \frac{\log(\pi_0 l_1) - \log((1-\pi_0)l_0)}{\log\theta_0} \\ (1-\pi_0) l_0 & \text{otherwise.} \end{cases} \end{aligned}$$

If we let $r^F_{\pi_0}(c)$ denote the optimal fixed sample size Bayes risk, then

$$r^F_{\pi_0}(c) = \min_{n\geq 0}(r^n(\pi) + nc).$$

($n = 0$ corresponds to making a decision without taking observations so that $r^0(\pi) = \min\{\pi_0 l_1, (1-\pi_0)l_0\}$.) Let $g(n) = (1/\theta_0)^n \pi_0 l_1 + nc$. Pretending that n is a continuous variable and differentiating with repect to n gives

$$g'(n) = -\theta_0^{-n}\pi_0 l_1 \log\theta_0 + c.$$

Setting equal to zero and solving gives

$$n^* = \frac{\log(\pi_0 l_1 \log\theta_0) - \log c}{\log\theta_0}.$$

Since the second derivative of $g(n)$ is positive, $g(n)$ is strictly convex in n. Define $q = \frac{\log(\pi_0 l_1) - \log((1-\pi_0)l_0)}{\log\theta_0}$ for notational convenience.
(i) Suppose that $q \leq 0 (\Leftrightarrow \pi_0 l_1 \leq (1-\pi_0)l_0)$.
Then $r^n(\pi) = (1/\theta_0)^n \pi_0 l_1$, so

$$r^F_{\pi_0}(c) \approx (1/\theta_0)^{n^*}\pi_0 l_1 + n^* c \tag{4}$$

unless $n^* < 0$, in which case $r^F_{\pi_0}(c) = \pi_0 l_1$.
(ii) Suppose that $q > 0 (\Leftrightarrow \pi_0 l_1 > (1-\pi_0)l_0)$.
If $n^* > q$,

$$r^F_{\pi_0}(c) \approx \min\{(1/\theta_0)^{n^*}\pi_0 l_1 + n^* c, (1-\pi_0)l_0\}. \tag{5}$$

And if $n^* \leq q$,

$$r^F_{\pi_0}(c) = (1-\pi_0)l_0.$$

The asymptotic efficiency of the Bayes sequential procedure with respect to the optimal fixed sample size Bayes procedure in terms of their risks is given in the next theorem.

Theorem 1. *For a fixed prior probability π_0, the asymptotic ratio of the Bayes sequential risk to the optimal fixed sample size Bayes risk has the property*

$$\lim_{c\to 0}\frac{r^F_{\pi_0}(c)}{r_{\pi_0}(c)} = \frac{1}{1-\pi_0}.$$

Proof:

Let π_0 be fixed and let $c > 0$ be sufficiently small.

(a) For the Bayes sequential procedure, from (3),

$$\begin{aligned}
r_{\pi_0}(c) &\approx \pi_0\left(\frac{(1-\pi_0)l_1c}{\pi_0(l_1 - \frac{c}{1-\theta_0^{-1}})\log\theta_0} + \frac{c}{1-\theta_0^{-1}}\left(1 - \frac{(1-\pi_0)c}{\pi_0(l_1 - \frac{c}{1-\theta_0^{-1}})}\right)\right) \\
&\quad + (1-\pi_0)c\frac{\log(\pi_0(l_1 - \frac{c}{1-\theta_0^{-1}})\log\theta_0) - \log((1-\pi_0)c)}{\log\theta_0} \\
&\approx \pi_0\left(\frac{(1-\pi_0)c}{\pi_0\log\theta_0} + \frac{c}{1-\theta_0^{-1}}\left(1 - \frac{(1-\pi_0)c}{\pi_0 l_1\log\theta_0}\right)\right. \\
&\quad + (1-\pi_0)c\frac{log(\pi_0 l_1\log\theta_0) - \log((1-\pi_0)c)}{\log\theta_0} \\
&\approx \frac{c}{\log\theta_0}\left(1 - \pi_0 + \frac{1}{1-\theta_0^{-1}}\left(\pi_0\log\theta_0 - \frac{(1-\pi_0)c}{l_1}\right)\right. \\
&\quad + (1-\pi_0)(\log\pi_0 l_1\log\theta_0 1 - \pi_0 - \log c)) \\
&= O(c) - \frac{1-\pi_0}{\log\theta_0}c\log c \\
&= O\left(-\frac{1-\pi_0}{\log\theta_0}c\log c\right)
\end{aligned}$$

as $c \to 0$.

(b) For the optimal fixed sample size Bayes procedure, since

$$\begin{gathered}
n^* \le 0 \Leftrightarrow \pi_0 l_1\log\theta_0 \le c \ \ and \\
n^* \le q \Leftrightarrow (1-\pi_0)l_0 \le c,
\end{gathered}$$

$n^* > 0$ & $q \le 0$ or $n^* > q$ & $q > 0$ for small c. Also, $n^*c \approx \frac{-c\log c}{\log\theta_0} = o(1)$ as $c \to 0$, Thus by (4) and (5),

$$\begin{aligned}
r_{\pi_0}^F(c) &\approx (1/\theta_0)^{n^*}\pi_0 l_1 + n^*c \\
&= \frac{c}{\pi_0 l_1\log\theta_0}\pi_0 l_1 + \frac{\log(\pi_0 l_1\log\theta_0) - \log c}{\log\theta_0}c \\
&= \frac{c}{\log\theta_0} + c\frac{\log(\pi_0 l_1\log\theta_0) - \log c}{\log\theta_0} \\
&= O(c) - \frac{c\log c}{\log\theta_0}
\end{aligned}$$

$$= O(\frac{-c \log c}{\log \theta_0})$$

as $c \longrightarrow 0$.

This proves the stated assertion. □

Next, an efficiency will be considered in terms of sampling cost of each procedure.

Theorem 2. *For given $c > 0$, let c^F be such that $r_{\pi_0}(c) = r^F_{\pi_0}(c^F)$ for a given π_0. Then*

$$\lim_{c \to 0} \frac{c^F}{c} = 1 - \pi_0.$$

Proof :

Assume that c is sufficiently small. From the above theorem

$$r_{\pi_0}(c) \approx -\frac{1 - \pi_0}{\log \theta_0} c \log c. \tag{6}$$

Again from the above theorem, since

$$\frac{r^F_{\pi_0}(c)}{r_{\pi_0}(c)} = \frac{1}{1 - \pi_0} > 1,$$

$r_{\pi_0}(c) < r^F_{\pi_0}(c)$ for small c. Thus to have them equal, c^F must be less than c, implying $\lim_{c \to 0} c^F = 0$. Now, for c^F small,

$$r^F_{\pi_0}(c^F) \approx -\frac{c^F \log c^F}{\log \theta_0}. \tag{7}$$

The Theorem now follows from (6) and (7). □

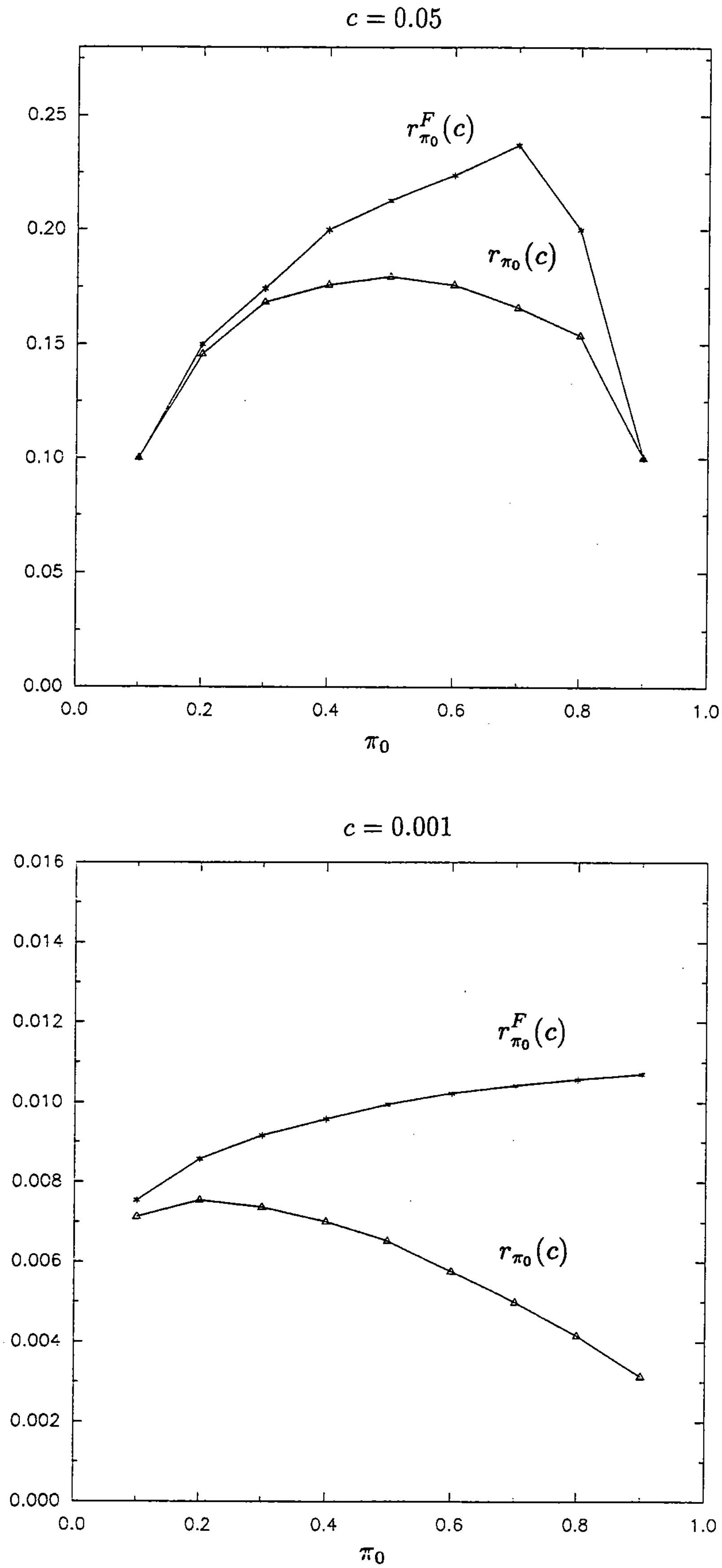

Figure 1: Plots of sequential Bayes risk and the optimal sample size Bayes risk for $0-1$ decision loss.

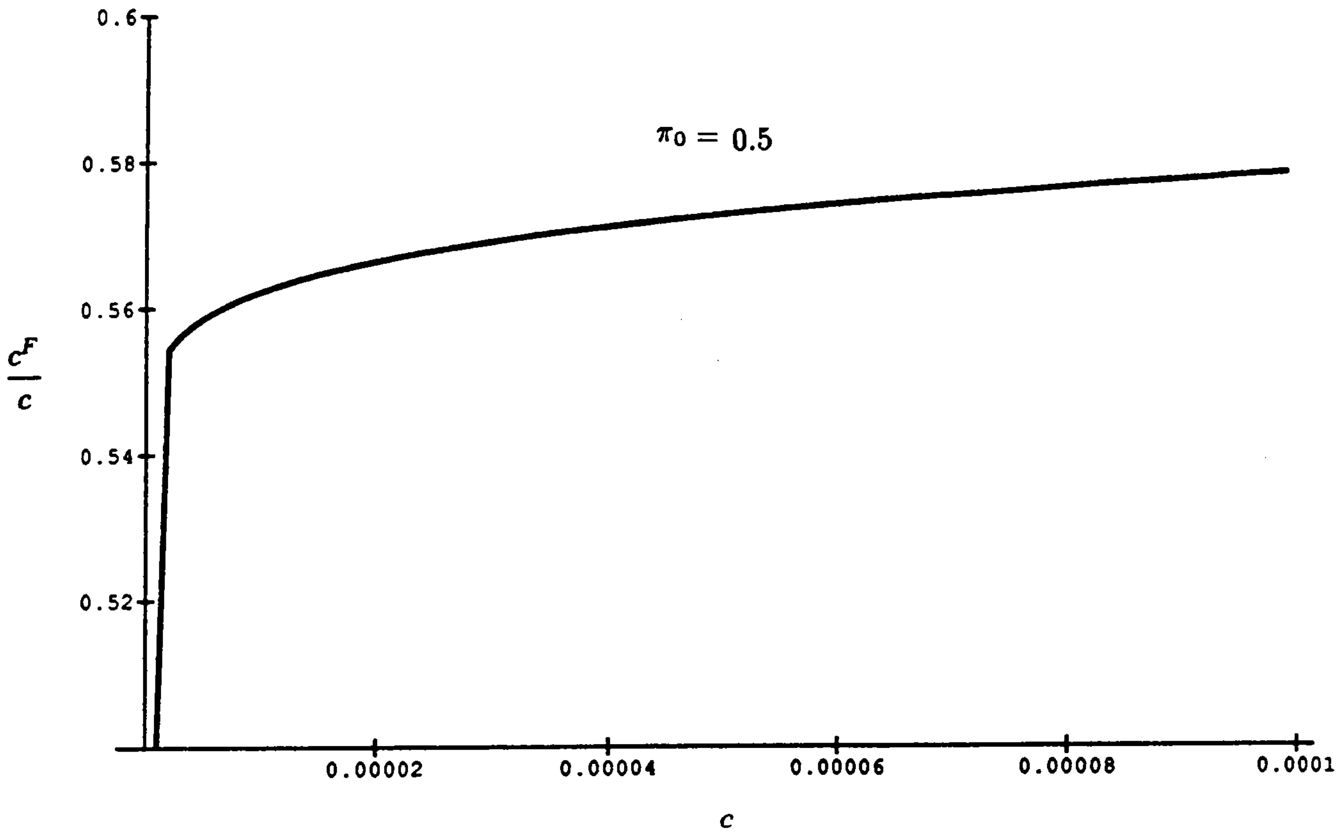

Figure 2: Plot of $\frac{c^F}{c}$ for $0-1$ decision loss.

3 Minimax Strategy and the Minimax Risk

We now consider the minimax sequential procedure for the problem. A minimax sequential procedure is a procedure which minimizes $\sup_\theta R(\theta, \mathsf{d})$ among all proper sequential procedures. We begin with definitions of the needed concepts.

Definition 1. A sequential rule δ_1 is R-better than a sequential rule δ_2 if $R(\theta, \delta_1) \leq R(\theta, \delta_2)$ for all $\theta \in \Theta$, with strict inequality for some θ.

Definition 2. A class $\mathcal{C}$ of sequential rules is said to be complete if, for any sequential rule δ not in $\mathcal{C}$, there is a sequential rule $\delta' \in \mathcal{C}$ which is R-better than δ.

It is shown in Brown, Cohen, and Strawderman(1980). that for simple versus simple testing problems the Bayes sequential rules form a complete class. Since a Bayes rule is represented by d_J for some $J \geq 0$, the minimax sequential procedure can be considered only among the procedures $d_J, J \geq 0$. Let $r_m(c)$ denote the minimax risk. Then

$$r_m(c) = \inf_{\mathsf{d}} \sup_{\theta} R(\theta, \mathsf{d}) = \inf_{d_J} \sup_{\theta} R(\theta, d_J).$$

Let us recall from the previous section that the Bayes risk for the procedure $d_J, J \geq 1$, is

$$r(\pi_0, d_J) = \pi_0(\theta_0^{-J} l_1 + c\frac{1-\theta_0^{-J}}{1-\theta_0^{-1}}) + (1-\pi_0)cJ.$$

Since for $J \geq 1$

$$\begin{aligned} \sup_{\theta} R(\theta, d_J) &= \max\{\theta_0^{-J} l_1 + c\frac{1-\theta_0^{-J}}{1-\theta_0^{-1}}, cJ\} \\ &= \max_{\pi_0} r(\pi_0) \end{aligned}$$

and

$$\sup_{\theta} R(\theta, d_0) = max\{l_0, l_1\},$$

the minimax sequential risk is

$$r_m(c) = \min\{\min_{J\geq 1}(\max\{\{\theta_0^{-J} l_1 + c\frac{1-\theta_0^{-J}}{1-\theta_0^{-1}}, cJ\}), \max\{l_0, l_1\}\}.$$

The following lemma gives that for all small c we will need more than 1 observation under the alternative as the minimax sequential strategy.

Lemma 1. *Let $d_{J^m(c)}$ denote the minimax sequential strategy. Then $J^m(c)$ is at least 1 for* $0 < c \leq l_1(\theta_0 - 1)/\theta_0$.

Proof :

Let $g_1(x) = \theta_0^{-x} l_1 + c\frac{1-\theta_0^{-x}}{1-\theta_0^{-1}}$ and let $g_2(x) = cx$ for $x \geq 0$. Since the first derivative of $g_1(x)$ is positive if $c \leq l_1(\theta_0 - 1)/\theta_0$, g_1 is decreasing function with $g_1(0) = l_1$. Thus $g_1(x)$ meets with $g_2(x)$ exactly once and the crossing point will minimize $\max\{g_1(x), g_2(x)\}$. Let $x^m(c)$ be the crossing point for a given c. Solving the equation $g_1(x) = g_2(x)$ gives that

$$c = \frac{l_1(\theta_0 - 1)}{\theta_0 - \theta_0^{x^m(c)}(\theta_0 + x^m(c)(1 - \theta_0))}. \tag{8}$$

Let

$$h_1(x) = \theta_0 - \theta_0^x(\theta_0 + x(1 - \theta_0)). \tag{9}$$

Then

$$h_1'(x) > 0 \Leftrightarrow x > 1 + \frac{1}{\theta_0 - 1} - \frac{1}{\log\theta_0}. \tag{10}$$

Since

$$0 < 1 + \frac{1}{\theta_0 - 1} - \frac{1}{\log\theta_0} < 1$$

and $h_1(0) = h_1(1) = 0$, $h_1(x) \leq 0$ for $0 \leq x \leq 1$. It follows that this contradicts (8), since c is necessarially positive. Hence $x^m(c) > 1$ and therefore $J^m(c) \geq 1$ for $0 < c \leq l_1(\theta_0 - 1)/\theta_0$. □

Theorem 3. *For the minimax sequential strategy $d_{J^m(c)}$, $J^m(c)$ is monotonically decreasing in c.*

PROOF :

Let $0 < c_1 < c_2 < l_1(\theta_0 - 1)/\theta_0$ and let $x_1 = x^m(c_1)$ and $x_2 = x^m(c_2)$. $c_1 - c_2 < 0$ implies by virtue of (8),

$$\frac{(\theta_0 - 1)(\theta_0^{x_1}(\theta_0 + x_1(1-\theta_0)) - \theta_0^{x_2}(\theta_0 + x_2(1-\theta_0)))}{(\theta_0 - \theta_0^{x_1}(\theta_0 + x_1(1-\theta_0)))(\theta_0 - \theta_0^{x_2}(\theta_0 + x_2(1-\theta_0)))} < 0. \quad (11)$$

Again, let $h_1(x) = \theta_0 - \theta_0^x(\theta_0 + x(1-\theta_0))$. Then, from (10), $h_1(x)$ is strictly increasing for $x \geq 1$. But $h_1(1) = 0$. Thus $h_1(x) > 0$ for all $x > 1$. Hence

$$(11) \Leftrightarrow$$
$$\theta_0 x_1(\theta_0 + x_1(1-\theta_0)) < \theta_0 x_2(\theta_0 + x_2(1-\theta_0)). \quad (12)$$

Let $h_2(x) = \theta_0^x(\theta_0 + x(1-\theta_0))$. Then

$$\begin{aligned} h_2'(x) &= \theta_0^x((1-\theta_0) + (\theta_0 + x(1-\theta_0))\log\theta_0) \\ &< \theta_0^x(1 - \theta_0 + \log\theta_0) < 0. \end{aligned}$$

i.e., $h_2(x)$ is decreasing in x. Thus

$$(12) \Rightarrow x_1 > x_2.$$

Also, by Lemma 1, $x_2 > 1$.
Suppose $n < x_m(c_2) < x_m(c_1) < n+1$ for some integer $n \geq 1$. Then

$$\begin{aligned} J^m(c_1) &= \begin{cases} n & \text{if } g_1(n) < g_2(n+1) \\ n+1 & \text{otherwise} \end{cases} \\ &= \begin{cases} n & \text{if } \frac{1}{c_1} < \frac{1}{l_1}(\frac{\theta_0}{\theta_0-1} - (\frac{1}{\theta_0-1} - n)\theta_0^n) \\ n+1 & \text{otherwise} \end{cases} \end{aligned}$$

and

$$\begin{aligned} J^m(c_2) &= \begin{cases} n & \text{if } g_(n) < g_2(n+1) \\ n+1 & \text{otherwise} \end{cases} \\ &= \begin{cases} n & \text{if } \frac{1}{c_2} < \frac{1}{l_1}(\frac{\theta_0}{\theta_0-1} - (\frac{1}{\theta_0-1} - n)\theta_0^n) \\ n+1 & \text{otherwise.} \end{cases} \end{aligned}$$

Since

$$1/c_2 < 1/c_1 < \frac{1}{l_1}(\frac{\theta_0}{\theta_0 - 1} - (\frac{1}{\theta_0 - 1} - n)\theta_0^n),$$

$$J^m(c_1) = n \Longrightarrow J^m(c_2) = n.$$

Thus $J^m(c_1) \geq J^m(c_2)$. Suppose that $n < x_1$ and $m < x_2$ where $n > m > 0$. Then obviously $J^m(c_1) > J^m(c_2)$. Hence $J^m(c)$ is monotone decreasing of c if $0 < c \leq l_1(\theta_0 - 1)/\theta_0$.
Now, if $c \geq l_1(\theta_0 - 1)/\theta_0$,

$$g_1(x) = \theta_0^{-x} l_1 + c\frac{1-\theta_0^{-x}}{1-\theta_0^{-1}}$$

is increasing in x. Thus $\max\{g_1(x), cx\}$ is minimized at $x = 0$. Hence $J^m(c) = 0$ if $c \geq l_1(\theta_0 - 1)/\theta_0$. This proves $J^m(c)$ is monotone decreasing for all c. □
Next, it will be shown that the minimax sequential risk $r_m(c)$ is monotonely decreasing in c.

Theorem 4. *The minimax risk*

$$r_m(c) = \min_J \max_{\pi_0} r(\pi_0, d_J)$$

is monotone increasing in c.

Proof :
(a) Assume that $0 < c < l_1(\theta_0 - 1)/\theta_0$. Let $0 < c_1 < c_2 < l_1(\theta_0 - 1)/\theta_0$ be given. Let $J_1 = J^m(c_1), J_2 = J^m(c_2)$. Then by the above theorem, $J_1 \geq J_2 > 1$. And

$$r_m(c_i) = \max\{\theta_0^{-J_i} l_1 + c_i\frac{1-\theta_0^{-J_i}}{1-1/\theta_0}, c_i J_i\}, \quad i = 1, 2.$$

(i) Suppose that $J_1 = J_2$.; Then obviously $r_m(c_1) \leq r_m(c_2)$.
(ii) Suppose that $J_1 = J_2 + n$ for some $n \geq 1$. ; If $r_m(c_1) = c_1 J_1$,

$$\begin{aligned} c_1 J_1 &< l_1\theta_0^{J_1-1} + c\frac{1-\theta_0^{J_1-1}}{1-1/\theta_0} (\textit{by the definition of } J_1) \\ &\leq l_1\theta_0^{J_2} + c_1\frac{1-\theta_0^{J_2}}{1-1/\theta_0} \quad (\textit{since } g_1(x) \textit{ is decreasing in } x) \\ &< l_1\theta_0^{J_2} + c_2\frac{1-\theta_0^{J_2}}{1-1/\theta_0} \quad (c_1 < c_2) \\ &\leq r_m(c_2). \end{aligned}$$

If $r_m(c_1) = l_1\theta_0^{-J_1} + c_1\frac{1-\theta_0^{-J_1}}{1-1/\theta_0}$, then since $g_1(x)$ is decresing of x and $c_1 < c_2$,

$$\begin{aligned} r_m(c_1) &= l_1\theta_0^{-J_1} + c_1\frac{1-\theta_0^{-J_1}}{1-1/\theta_0} \\ &< l_1\theta_0^{-J_2} + c_2\frac{1-\theta_0^{-J_2}}{1-1/\theta_0} \\ &\leq r_m(c_2). \end{aligned}$$

(b) Assume that $c > l_1(\theta_0 - 1)/\theta_0$. Then $J^m(c) = 0$ from the above theorem. Thus

$$r_m(c) = \max\{l_0, l_1\}.$$

Since $r_m(c) < l_1$ for $0 < c < l_1(\theta_0 - 1)/\theta_0$, $r_m(c)$ is monotone increasing in c.
□

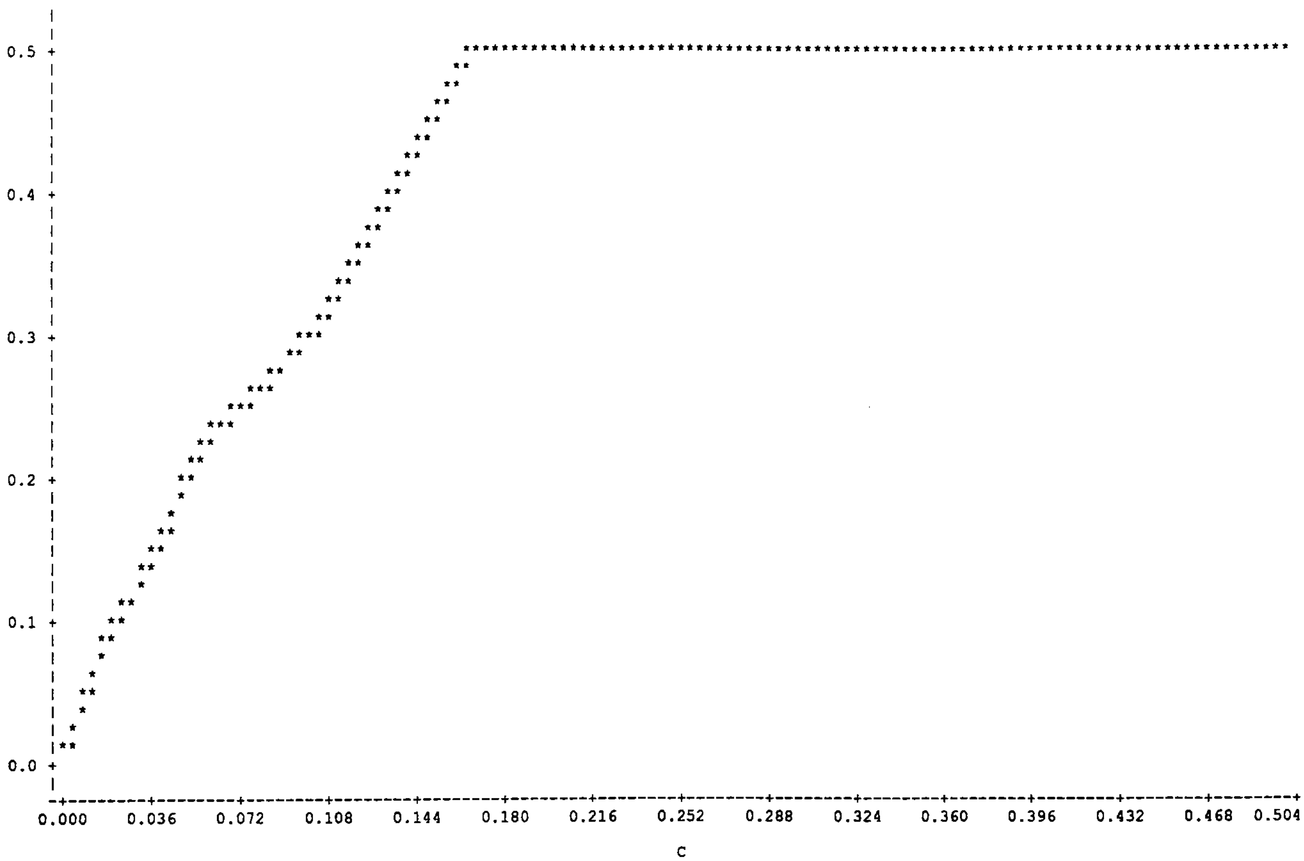

Figure 3: Plot of Minimax risk and the sampling cost when $\theta_0 = 2$ for $0-1$ decision loss.

4 Acknowledgement

We are thankful to Herman Rubin for a number of helpful diserssions.

References

[1] ARROW, K. J., BLACKWELL, D. and GIRSHICK, M. A. (1949). Bayes and minimax solutions of sequential decision problems. *Econometrica* **17** *213-244*.

[2] BERGER, J. (1985). *Statistical Decision Theory and Bayesian Analysis.* Springer-Verlag, New York. (2nd Edition)

[3] BROWN, L. D., COHEN, A. and STRAWDERMAN, W. E. (1980). Complete classes for sequential tests of Hypotheses. *Ann. Statist.* **8** *377-398.*

[4] BROWN, L. D., COHEN, A. and STRAWDERMAN, W. E. (1979). Monotonicity of Bayes Sequential Tests. *Ann. Statist.* **7** *1222-1230.*

[5] CHERNOFF, H.(1972) *Sequential Analysis and Optimal Design.* SIAM, Philadelphia.

[6] FERGUSON, T. S. (1967) *Mathematical Statistics: A Decision-Theoretic Approach.* Academic Press, New York.

[7] SOBEL, M.(1953) An Essentially Complete Class of Decision Functions for Certain Standard Sequential Problems. *Ann. Math. Statist.* **24** *319-337.*

[8] WALD, A. (1947). *Sequential Analysis.* New York, Wiley.

[9] WALD, A. (1950). *Statistical Decision Functions.* John Wiley and Sons.

[10] WALD, A. and WOLFOWITZ, J. (1948) Optimum charater of the sequential probability ratio test. *Ann. Math. Statist.* **19** *326-339.*