Sequential Batch Sampling

by

Eitan Greenshtein

Technical Report # 90–12

Department of Statistics
Purdue University

March, 1990

SEQUENTIAL BATCH SAMPLING

by

Eitan Greenshtein
Purdue University
West Lafayette, IN 47907

ABSTRACT

Let $\{X_i\}$ be i.i.d. $\{F_\theta\}$. We consider the testing problem $H_1 : \theta < \theta_0$ vs $H_2 : \theta \geq \theta_0$, where sampling is done sequentially in batches with variable batch sizes.

For some parametric families $\{F_\theta\}$, and under natural cost structure, we characterize the Bayes procedures. This characterization of Bayes procedures lead to a similar characterization of essentially complete class of procedures.

Introduction

Let $\{X_i\}_{i=1}^N, N \leq \infty$ be a sequence of independent random variable. Assume $X_i \sim F_\theta$ where F_θ is a one dimensional exponential family.

Assume for each $m = 1, 2, \ldots$, after observing $X_1, \ldots, X_m$ the statistician may decide one of the terminal decision $H_1 : \theta < \theta_0$ or $H_2 : \theta \geq \theta_0$, or may decide to sample one more observation at a cost c_1, or to sample a batch of two more observations at a cost c_2. This formulation will be generalized later (section 3) to treat problems in which more than two batch sizes are allowed. We will assume that an initial set of observations $X_1, \ldots, X_I$ $I > 0$ is always given. It will be called the initial observation, and the first decision is made after observing it.

A little research was done about sequential decision problems in this setting, for some reference see [6] [10] for example. The purpose of this work is to characterize Bayes procedures, the characterization is by monotonicity (see definition 1). The definition of monotone procedures is a natural extension to the batch sample problem of the definition given by Sobel [9], and Brown, Cohen, Strawderman [1], for monotone sequential procedures.

Section 1: Preliminaries

We will now describe the problem in a decision theoretic framework. Let $A_m = \{d_1^m, d_2^m, b_1^m, b_2^m, \eta^m\}$ be the actions available at stage m. d_i^m should be interpreted as the terminal decision H_i, and b_i^m as the decision to sample a further batch of size i. The meaning of the action η^m will be clear later.

Let the terminal loss $L(\theta, d_i), i = 1, 2$, be a real valued functions. Assume:

(i) $L(\theta, d_1)$ is non decreasing

(ii) $L(\theta, d_2)$ is non increasing

We will consider decision rules based on sufficient statistics $S_m = X_1 + \ldots + X_m$, $S_m \sim F_\theta^m, m = I, I+1, \ldots$. For reasons to be clear we add a point $\tilde{\eta}$ to the real line, denote $\overline{R} = R \cup (\tilde{\eta})$. A decision rule is a collection of randomizations $\delta_m(\cdot|s_m)$ on A_{m-I+1}, $s_m \in \overline{R}_m$ $m = I, I+1, I+2, \ldots$. The randomizations satisfy the following: $\sum_{a \in A_{m-I+1}} \delta_m(a|s_m) = 1, \delta_m(\eta|\tilde{\eta}) = 1, \delta_m(\eta|s_m) = 0$ if $s_m \neq \eta$. A set $D = \{\delta_m\}$ and a parameter θ, determine a stochastic process on $(A_1 \times \overline{R}_I) \times (A_2 \times \overline{R}_{I+1}) \times \ldots$. Its description is the following: Observe $s_I \in \overline{R}_I$ according to the law $F_\theta^I(ds)$, choose an action

$a_1 \in A_1$ according to the law $\delta_I(\cdot|s_I)$. If $a_1 = b_1^1$, then observe a point $s_{I+1} \in \overline{R}_{I+1}$ according to the distribution law of S_{I+1} conditioning on $S_I = s_I$, choose an action $a_2 \in A_2$ according to the law $\delta_{I+1}(\cdot|S_{I+1})$ and so on. If $a_1 = b_2^1$ then observe s_{I+2} according to the distribution law of S_{I+2}, conditioning on $S_I = s_I$, set $s_{I+1} = \tilde{\eta}$ and accordingly choose the action η^{I+1}, then choose an action $a_3 \in A_3$ according to the law $\delta_{I+2}(\cdot|s_{I+2})$, and so on. Thus the value $\tilde{\eta}^m$ indicates that at stage $m-1$ a batch of size two was taken and the "action" denoted η^m indicates that no "significant" new decision can be made at that stage because it is only the first observation in a batch of size two. Finally if $a_k = d_i^k$ $i = 1, 2$, set $a_m = \eta$ and $s_m = \tilde{\eta}$ for $m > k$.

In the case $N < \infty$ we will assume $A_{N-1} = \{d_1^{N-1}, d_2^{N-1}, b_1^{N-1}, \eta^{N-1}\}, A_N = \{d_1^N, d_2^N, \eta^N\}$. Denote the measure induced by θ and D as $H_{\theta D}$; denote its marginal on $A_1 \times A_2 \times \ldots \times A_N$ as $\mu_{\theta,D}$. Let $\pi(\theta)$ be a prior distribution, and denote $H_{\pi,D}(da) = \int H_{\theta,D}(da)d\pi(\theta)$.

For a sequence $a = (a_1, a_2, \ldots) \in A_1 \times A_2 \times \ldots$, let k be the first index such that $a_k = d_1^k$ or $a_k = d_2^k$. Let n_1 be the number of indices i such that $a_i = b_1^i$ $i < k$, let n_2 be the number of indices i such that $a_i = b_2^i, i < k$. Define the loss function:

$$\tilde{L}(\theta, (a_1, a_2, \ldots)) = c_1 \cdot n_1 + c_2 \cdot n_2 + L(\theta, a_k).$$

Define the risk function:

$$R(\theta, D) = \int \tilde{L}(\theta, a) d\mu_{\theta,D}(a).$$

A Bayes procedures D° satisfy for some prior:

$$\operatorname*{Min}_D \int R(\theta, D) d\pi(\theta) = \int R(\theta, D^\circ) d\pi(\theta).$$

Total-Positivity

The concept of total-positivity will be used in the sequel, and we will review some facts about it now. Some references on this subject are [2] and [7].

Definition 1.1: The function $\varphi(x) : R \to R$ changes signs at most n times if and only if there exist $-\infty = a_0 < a_1 < \ldots < a_{n+1} = \infty$ such that $\varphi(x)$ preserves its sign on $(a_i, a_{i+1}), i = 0, \ldots, n$, i.e., it is either non negative or non positive.

Let $\{G_\theta\}$ $\theta \in \Theta \subseteq \mathcal{R}$ be a family of distributions on the real line.

Definition 1.2: $\{G_\theta\}$ is TP_n if for any function $\varphi(x)$, that changes signs at most $n-1$ times, $L(\theta) = E_\theta\varphi(x)$ changes signs at most $n-1$ times, and if it does change sign $n-1$ times, then it does so in the same order as φ. $\{G_\theta\}$ is STP_n if in addition for any φ as above which is not identically zero, the function $L(\theta)$ changes sign at most $n-1$ times in the stronger sense that there are $-\infty < a_1 \leq \ldots \leq a_{n-1} < \infty$ as in Definition 1 such that $L(\theta)$ can be zero only at $a_i, i = 1, \ldots, n-1$.

Suppose $X_i \sim F_\theta$ $\theta \in \Theta$ are i.i.d., $\{F_\theta\}$ is an exponential family and $\pi(\theta)$ is a prior distribution on Θ. Denote:

$$d\nu_{s_n}^{(n+i,n)} = d\nu^{(n+i,n)}(s_{n+i}|S_n = s_n)$$

the conditional distribution of S_{n+i} given $S_n = s_n$ and a prior $\pi(\theta)$.

In the sequel we will require $\{\nu_{s_n}^{(n+i,n)}\}$ to be STP_3 with respect to the parameter s_n. Denote $F_\theta^{(n+i,n)}(ds_{n+i}|S_n = s_n)$, the conditional distribution of S_{n+i} given $S_n = s_n$.

Proposition 1.1: Suppose for some $\theta_0 \in \Theta, \{F_{\theta_0}^{(n+i,n)}(ds_{n+i}|S_n = s_n)\}$ is $(S)TP_n$, with respect to the parameter s_n, then for every $\pi(\theta)$ $\{\nu_{s_n}^{(n+i,n)}\}$ is $(S)TP_n$.

Proof: [5].

Proposition 1.2: For the cases where $\{F_\theta\}$ are: Binomial $\theta = P$, Poisson with $\theta = \lambda$ its expectation, Geometric $\theta = P$, Exponential $\theta = \lambda, \lambda^{-1}$ its expectation, Normal $\theta = \mu$ its expectation, $dF_{\theta_0}^{(n_i,n)}(ds_{n_i}|S_n = s_n)$ is STP_∞ (i.e., STP_n for every n).

Section 2: Main Theorem

Definition 2.1: A batch sampling sequential decision procedure is monotone if for every $\delta_m \in D$ there exist numbers $-\infty \le C_1^m \le C_2^m \le C_3^m \le C_4^m \le \infty$ such that $\delta_m(d_1|s_m) = 1$ if $S_m < C_1^m, \delta_m(b_1|s_m) = 1$ if $s_m \in (C_1^m, C_2^m) \cup (C_3^m, C_4^m)$, $\delta_m(b_2|s_m) = 1$ if $s_m \in (C_2^m, C_3^m)$ and $\delta_m(d_2|s_m) = 1$ if $s_m > C_4^m$, for almost every real number s_m under $H_{\pi,0}$. Randomization are allowed when $S_m = C_J^m$ between the actions taken for $s_m \in (C_{J-1}^m, C_j)$ and $s_m \in (C_j^m, C_{J+1}^m)$. (Obvious modification in the randomization description is needed when $C_J^m = C_{J+1}^m$.)

Theorem 1: Assume $\{\nu_{s_m}^{(m+i,m)}\}$ in STP_3 for every m and for $i = 1, 2$. Then every Bayes procedure in a batch sampling problem is monotone.

We will first consider the case of finite horizon, i.e. when the number of available observations N, is finite. First some notations and lemmas. Denote:

$$\rho_i^m(s) = \int L(\theta, d_i) d\pi(\theta|S_m = s).$$

In words $\rho_i^m(s)$ is the conditional on $S_m = s$ additional expected loss for deciding the terminal decision d_i; here $\pi(\theta|s_m = s)$ is the posterior distribution of θ. ("Additional" to the cost of sampling so far.) Let $\beta_n^{m+n}(s)$ be the conditional on $S_m = s$ expected additional loss for a Bayes procedure among those that sample at least one more batch after observing $S_m = s$, where the number of available observations is $N = m + n, n \neq 0\ n < \infty$. Define $\beta_m^N(s) = \underset{i=1,2}{\text{Min}}(\rho_i^m(s))$ if $N \geq m$. Then:

$$\beta_m^{m+n}(s_m) = \underset{i=1,2}{\text{Min}} \left[\int \text{Min} \ \left[c_i + \beta_{m+i}^{m+n}(s), c_i + \rho_1^{m+i}(s), c_i + \rho_2^{m+i}(s)\right] d\nu^{(m+i,m)}(s|s_m) \right]$$

Lemma 1: $\rho_1^m(s)$ is non decreasing and ρ_2^m is non increasing.

Proof: Sobel [9].

Lemma 2: $\rho_J^m(s_m) = \int \rho_J^{m+i}(s) d\nu^{(m+i,m)}(s|s_m)$.

Proof: Sobel [9].

Lemma 3:

(*i*) $\beta_m^{m+n}(s_m) - \rho_1^m(s_m)$ is monotone decreasing.

(*ii*) $\beta_m^{m+n}(s_m) - \rho_2^m(s_m)$ is monotone increasing.

Proof: The proof is similar to B.C.S. [1]. It is by induction on n. The general induction step is a s follows:

(*iii*)

$$\begin{aligned} &\beta_m^{m+n}(s_m) - \rho_2^m(s_m) = \\ &= \beta_m^{m+n}(s_m) - \int \rho_2^{m+i}(s) d\nu^{(m+i,m)}(s|s_m) \\ &= \underset{i=1,2}{\text{Min}} \left\{ \int \text{Min} \ \left[c_i + \beta_{m+1}^{m+n}(s) - \rho_2^m(s), c_i + \rho_1^{m+i}(s) - \rho_2^{m+i}(s), c_i\right] \right. \\ &\qquad \left. d\nu^{(m+i,m)}(s|s_m) \right\} \end{aligned}$$

The first function in the internal brackets is monotone increasing by the induction hypothesis, the second is monotone decreasing by Lemma 1, and the third is a constant.

Hence the function Min $[\cdot,\cdot,\cdot]$ is non decreasing. The family $\{d\nu_{s_m}^{(m_i,m)}\}$ is STP_3 and thus stochastically increasing. Therefore

$$g_i(s_m) = \int \text{Min } [\cdot,\cdot,\cdot] d\nu^{(m+i,m)}(s|s_m)$$

is non decreasing. The fact that $\{d\nu_{s_m}^{(m+i,m)}\}$ is STP_3 implies easily that for every real number w, $g_i(s) - w$ is zero at most at one point and hence $g_i(s)$ is monotone increasing. This implies $\beta_m^{m+n}(s_m) = \underset{i=1,2}{\text{Min}}(g_i(s_m))$ is monotone increasing. Similarly we can prove the other part.

Remarks on Lemma 3.

(i) Lemma 3 is still valid in the more general case when sampling a batch of size i at stage m costs $c_i^m \geq 0$.

(ii) Lemma 3 is still valid in the more general case in which at some stages, only one size of a batch is allowed.

Notice the interesting case is when $c_1 < c_2 < 2c_1$; otherwise the best policy is to always sample a size one batch, or always to sample a size two batch. Thus we assume w.l.o.g. $c_1 < c_2 < 2c_1$.

The following notations are needed for the proof of the next lemma. Let $r_m^{m+n}(D/s)$ be the conditional on $S_n = s$ additional expected Bayes loss, when there are n remaining observations $S_{m+1}, \ldots, S_{m+n}$ using the procedure $D, n \leq \infty$. Let $\Delta_m^{m+n}(s) = r_m^{m+n}(D^2|s_m) - r_m^{m+n}(D^1|s_m)$, where D^i is the procedure that takes a batch of size i after observing S_m and then proceeds optimally. Denote by $\tilde{r}_m^{m+n}(D^i|s_{m+1})$ the conditional expected additional loss using D^i conditional upon $S_{m+1} = s_{m+1}$ when there are n remaining observations $S_{m+1}, \ldots, S_{m+n}$. Denote:

$$\tilde{\Delta}_m^{m+n}(s_{m+1}) = \tilde{r}_m^{m+n}(D^2|s_{m+1}) - \tilde{r}_m^{m+n}(D^1|s_{m+1}), n \leq \infty.$$

Lemma 2.4: $\Delta_m^{m+n}(s_m)$ changes signs at most twice, and if it does, then there exists (a_1^m, a_2^m) such that the function is negative if and only if $s_m \in (a_1^m, a_2^m)$. Moreover $\Delta_n^{m+n}(\cdot)$ can be zero only at its crossing points.

Proof: Notice that:

$$\Delta_m^{m+n}(s_m) = \int \tilde{\Delta}_m^{m+n}(s_{m+1}) d\nu^{(m+1,n)}(s_{m+1}|s_m).$$

By STP_3 of $\{d\nu_{s_m}^{(m+1,m)}\}$ it is enough to show $\tilde{\Delta}_m^{m+n}(\cdot)$ changes signs at most twice and in the right order.

Examine the following auxiliary problem, which is a variant of a batch-sampling problem. Prior $\pi(\theta)$ as in the original problem, the initial observation is $X_1, \ldots, X_{m+1}$; the available actions are $A_1 = \{d_1^1, d_2^1, b_1^1, b_2^1, r\}$ and $A_J = \{d_1^J, d_2^J, b_1^J, b_2^J, \eta^J\}$ for $1 < J$. d_i^J, b_i^J and η^J have the usual meaning. The action $r \in A_1$ has the meaning of sampling a batch of size one at a cost $c_2 - c_1$. The cost of a batch of size one using b_1^J is c_1, of a batch of size two using b_2^J is c_2 as in the original problem. The terminal loss is $\hat{L}(\theta, d_i) = c_1 + L(\theta, d_i)$.

Notice that: At the first decision stage in the auxiliary problem, the conditional loss conditioning on $S_{m+1} = s_{m+1}$ and given that the Bayes action only among $\{d_1^1, d_2^1, b_1^1, b_2^1\}$ is taken (respectively, given that the action r is taken) is $\tilde{r}_m^{m+n}(D^1|s)$ (respectively, is $\tilde{r}_m^{m+n}(D^2|s)$). Paying attention to these definitions we conclude:

$$\tilde{r}_m^{m+n}(D^1|s) \geq \tilde{r}_m^{m+n}(D^2|s)$$

if and only if the Bayes action in the auxiliary problem given $S_{m+1} = s_{m+1}$ is r.

The Bayes action in the auxiliary problem cannot be b_1^1, the action m is better since $c_2 - c_1 < c_1$; the Bayes action in the auxiliary problem can neither be b_2^1, because it is better to take the action r and then either to stop or to take b_1^2. Hence in the auxiliary problem we need only to consider Bayes action among $A_1' = \{d_1^1, d_2^1, r\}$. From the remarks following Lemma 3 we conclude: The Bayes action in the auxiliary problem is r if and only if $s_{m+1} \in (E_1^{m+1}, E_2^{m+1})$ for certain $E_1^{m+1} \leq E_2^{m+1}$. Thus $\tilde{\Delta}_m^{m+n}(s_{m+1})$ changes signs at most twice, and the proof follows. □

Proof of Theorem 1: For the case $n < \infty$, the proof follows from Lemmas 1, 3, 4. In the case $n = \infty$, we consider the M truncated sequential decision problem. As in Chow Robbins and Siegmund [3] $\beta_m^\infty(s) = \lim_{M\to\infty} \beta_m^M(s)$. Since $\beta_m^M(s) - \rho_2^m(s)$ is monotone increasing for every $M < \infty$, we get: $\beta_m^\infty(s) - \rho_2^m(s)$ is nondecreasing. Now the equation (iii) in Lemma 3 is true also for $n = \infty$. Hence we can conclude as in Lemma 3 that $\beta_m^\infty(s) - \rho_2^m(s)$ is monotone increasing. Similarly for $\beta_m^\infty(s) - \rho_1^m(s)$. As in Lemma 4, we conclude $\tilde{\Delta}_m^\infty(s)$ changes sign at most twice, and by STP_3 of $d\nu_{s_m}^{(m+1,m)}$, $\Delta_m^\infty(s)$ changes sign at most twice and it is zero only at its crossing points. The proof now follows.

Proposition 3.1: The class of monotone procedures is an essentially complete class.

Proof: This can be shown as in [1] and [9]. The reason is that every admissible procedure is a limit of Bayes rules, and limit of monotone procedures is a monotone procedure. □

Section 4: Generalization

Suppose at each stage the size of the next sample can be $1, 2, \ldots, M$. Define $b_1, \ldots, b_M$,

$D^1, \ldots, D^M$ and $r_m^{m+n}(D^i|s)$ in the obvious analogy to the definition in Section 2. Suppose the cost of a batch of size i is $p + c \cdot i$.

Theorem 2: Suppose $\{d\nu_{s_m}^{(m+i,m)}\}$ is STP_3 for every m and i. Then:

i) For $\ell > k, \tilde{r}_m^{m+n}(D^\ell|s) - \tilde{r}_m^{m+n}(D^k|s)$ changes signs at most twice, and if it does it is first positive. Moreover the function is zero only at its crossing points.

ii) If the available batches at stage m are of size $1, 2, \ldots, \min(n, M)$, then the Bayes action is $b_{\min(n,M)}^m$ if and only if $s_m \in (E_1^m, E_2^m)$ for certain $E_1^m \leq E_2^m$.

We find it convenient to deal with a slightly wider class of batch sampling problems to be described. The available batches at each stage m are $1, 2, \ldots, \min(M, n)$ but at the first stage the available batches are of size $1, \ldots, M_1$. $M_1 \leq \min(M, n)$. Also the cost of a batch size i is $p + c \cdot i$, but at the first stage the cost of the batch size M_1 is $c_{M_1} \leq p + c \cdot M_1$. We will denote the cost of a batch of size i, c_i.

Proof of the Theorem: (ii) follows from (i), Lemma 1 and a modification of Lemma 3 to the multi batch case.

(i) The proof is very similar to that of Lemma 4. We will briefly describe it. We use a double induction argument. Suppose we have proved (i) for every pair ℓ, k $\ell > k$ such that $\ell < L$, we will prove it for $\ell = L$.

The proof for $\ell = L$ will be by induction on the number of remaining observations n, assuming the total number is $N = m + n$. Define:

$$\Delta_n^{m+n}(s) = r_m^{m+n}(D^\ell|s) - r_m^{m+n}(D^k|s)$$

Define $\tilde{r}_n^{m+n}(D^k|s)$ the risk using D^k conditional on $S_{m+k} = s$. Similarly define $\tilde{r}_m^{m+n}\ (D^\ell|s)$. Denote:

$$\tilde{\Delta}_m^{m+n}(s) = \tilde{r}_m^{m+n}(D^\ell|s) - \tilde{r}_m^{m+n}(D^k|s).$$

Then:

$$\Delta_m^{m+n}(s_m) = \int \tilde{\Delta}_m^{m+n}(s_{m+k})d\nu^{(m+i,m)}(s_{m+k}|s_m).$$

By STP_3 it is enough to show $\tilde{\Delta}_m^{m+n}(s_{m+k})$ changes signs twice. Consider the auxiliary problem with initial observations $X_1, \ldots, X_{m+k}$, actions $A_1 = \{d_1^1, d_2^1, b_1^1, \ldots, b_{M_1}^1, r\}$, $A_J = \{d_1^j, d_2^j, b_1^j, \ldots, b_{M_j}^j\} J > 1$, $M_j = \{M, N-J\}$, terminal loss function $\hat{L}(\theta, d_i) = c_k + L(\theta, d_i)$. The action r has the meaning of taking a batch of size $\ell - k$ at a cost $c_{r-k} = c_\ell - P - c \cdot k$. Notice that $c_\ell \leq p + c \cdot \ell$ and hence $c_{r-k} \leq p + c \cdot (\ell - k)$. It is easy to see from the cost structure that the actions $b_{\ell-k}, b_{\ell-k+1}, \ldots$ are inadmissible and in the auxiliary problem we may consider only $A_1' = \{d_1^1, d_2^1, b_1^1, \ldots, b_{\ell-k-1}, r\}$. As in Lemma 4 $\tilde{\Delta}_m^{m+n}(s)$ is negative if and only if the Bayes action conditional upon $S_{m+k} = s$ is r. By induction hypothesis there exist an interval (E_1^{m+k}, E_2^{m+k}) such that the Bayes action is r if and only if $s_{m+k} \in (E_1^{m+k}, E_2^{m+k})$. The proof now follows by STP_3 of $\{\nu_{s_m}^{(m+k,m)}\}$. For the case $n = \infty$ we proceed as in Theorem 1. □

We will now extend the definition of monotonicity to the multi choice sequential batch sampling problem.

Definition 4.1: A multi choice sequential batch sampling problem, with largest available batch of size M, is monotone if there exist numbers $-\infty < E_1^m \leq \ldots \leq E_n^m \leq B_n^m \leq \ldots \leq B_1^m < \infty$ such that $\delta_m(b_i^m|s) = 1$ if $s \in (E_i^m, E_{i+1}^m) \cup (B_{i+1}^m, B_i^m)$, $\delta_m(d_1^m|s) = 1$ if

$s \in (-\infty, E_1^m), \delta_m(d_2^m|s) = 1$ if $s \in (B_1^m, \infty)$. Obvious randomizations are allowed when $s_m = E_i^m$ or $s_m = B_i^m$.

Theorem 2 together with Lemmas 1 and 3 modified for the multi choice problem, do not yield a monotonicity theorem. The following structure does not contradict them: Let $\infty < E_1^m < E_2^m < E_3^m < E_4^m < \infty; \delta_m(b_3^m|s) = 1$ if $s \in (E_1^m, E_2^m), \delta_m(b_1^m|s) = 1$ if $s \in (E_2^m, E_3^m), \delta_m(b_2^m|s) = 1$ if $s \in (E_3^m, E_4^m)$, $\delta(d_1^m|s) = 1$ if $s \in (-\infty, E_1^m)$, $\delta(d_2^m|s) = 1$ if $s \in (E_4^m, \infty)$.

When the problem has enough symmetry, a monotonicity theorem can be proved for the multi batch problem as in the following two examples.

<u>Example 1:</u> Consider a multi-batch sampling problem where $Y_i \sim N(\theta, 1) H_1 : \theta < 0$ v.s. $H_2 : \theta > 0$, the prior $\pi(\theta)$ is $N(\mu, \sigma^2)$ and the loss is $0 - 1$.

<u>Example 2:</u> Consider the problem in Example 1 only with the following different loss function

$$L(\theta, d_i) = \begin{cases} |\theta| & \text{if } \theta \notin \Theta_i \\ 0 & \text{otherwise} \end{cases}$$

Here $\Theta_1 = \{\theta | \theta \leq 0\}$ $\Theta_2 = \{\theta | \theta > 0\}$. This problem was considered by Chernoff [4] and other authors in the ordinary sequential framework.

The following consideration applies for both problems. There exists a value s° such that if the first observation is s° then $\pi(\theta|s^\circ)$ is symmetric around 0. A Bayes procedure should obviously be symmetric around s°. This fact together with Lemmas 1 and 3 and Theorem 2 yield the desired monotonicity.

Acknowledgement

The author would like to thank Professor L.D. Brown for helpful discussions concerning this problem, and for comments that lead to a better presentation.

References

[1] Brown, L.D., Cohen, A., and Strawderman, W.E. (1979). Monotonicity of Bayes Sequential Tests.

[2] Brown, L.D., Johnstone, I.M., MacGibon, K.B. (1981). Variation diminishing transformations. *J. Am. Stat. Assoc.*, **76**, 824–832.

[3] Chow, Y.S., Robbins, H., and Siegmund, D. (1971). *Great expectations*, Houghton Mifflin Company.

[4] Chernoff, H. (1972). Sequential Analysis and Optimal Design. Regional Conf. in App. Math. Siam.

[5] Greenshtein, E., Ph.D. thesis (1990), Cornell University, Ithaca, NY.

[6] Hayre, L.S. Group Sequential Sampling with Variable Group Sizes. *J. R. Stat. Soc. B*, (1985), 90–97.

[7] Karlin, S. (1968). *Total-Positivity*, Stanford University Press, Stanford, Ca.

[8] Karlin, S. (1955). Decision Theory for Polya Type Distributions. Third Berkeley Symp.

[9] Sobel, M. (1952). An Essentially Complete Class of Decision Functions for Certain Standard Sequential Problems. *Ann. Math. Stat.*, **23**, 319–332.

[10] Spahn, M., and Ehrenfeld, S. (1974). Optimal and Suboptimal Procedures in Group Sequential Sampling. *Nav. Res. Log. Q.*, **21**, 53–68.