Sequential Analysis: Design Methods and Applications
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/lsqa20

Construction of an Optimal Sequential Plan for Testing a Treatment for an Adverse Effect

Sumihiro Suzuki [a] & Michael I. Baron [b]

[a] Department of Biostatistics, University of North Texas, Health Science Center, Fort Worth, Texas, USA

[b] Department of Mathematical Sciences, University of Texas at Dallas, Richardson, Texas, USA

Available online: 04 Aug 2011

To cite this article: Sumihiro Suzuki & Michael I. Baron (2011): Construction of an Optimal Sequential Plan for Testing a Treatment for an Adverse Effect, Sequential Analysis: Design Methods and Applications, 30:3, 261-279

To link to this article: http://dx.doi.org/10.1080/07474946.2011.593917

Sequential Analysis, 30: 261–279, 2011

ISSN: 0747-4946 print/1532-4176 online
DOI: 10.1080/07474946.2011.593917

Taylor & Francis
Taylor & Francis Group

Construction of an Optimal Sequential Plan for Testing a Treatment for an Adverse Effect

Sumihiro Suzuki[1] and Michael I. Baron[2]
[1]Department of Biostatistics, University of North Texas, Health Science Center, Fort Worth, Texas, USA
[2]Department of Mathematical Sciences, University of Texas at Dallas, Richardson, Texas, USA

Abstract: Classical sequential procedures that collect a single observation at a time are often found impractical, expensive, and time consuming. *Sequentially planned procedures*, or simply sequential plans, extend and generalize the concepts of sequential analysis by allowing observations to be collected in groups of variable sizes. After every group, all of the previously collected data are used to determine the next course of action. An optimal (Bayes) sequential plan minimizes the (Bayes) risk function that combines the decision loss, observation (variable) cost, and group (fixed) cost. In general, determining the optimal sequential plan remains an open and challenging problem mainly because it requires risk optimization over a huge and rather unstructured set of all sequential plans. This article demonstrates how to obtain the optimal solution for a particular class of problems that may arise in testing a treatment for a rare but severe adverse effect. This solution is obtained by studying a number of properties of the Bayes sequential plan such as *transitivity* and *monotonicity*. This allows one to reduce the search to a small, manageable set of sequential plans within which the optimal plan can be calculated.

Keywords: Adverse effect; Group sequential methods; Monotonicity; Optimal sequential plan; Risk function; Sequentially planned design; Transitivity.

Subject Classifications: 62L05; 62L10; 62F03; 62P05.

1. INTRODUCTION

In classical (pure) sequential analysis, the data are collected one observation at a time, and the decision to terminate sampling is taken based on all of the data already

Received May 13, 2010, Revised December 13, 2010, and February 14, 2011, Accepted February 20, 2011
Recommended by Wolfgang Schmid
Address correspondence to Sumihiro Suzuki, Department of Biostatistics, University of North Texas, Health Science Center, 3500 Camp Bowie Blvd., Fort Worth, TX 76107, USA; E-mail: Sumihiro.Suzuki@unthsc.edu

collected. In practice, sampling one observation at a time is often expensive and time consuming. As long as there is a cost per group in addition to a cost per observation, for example, the time factor, it is cheaper to sample groups rather than single units. Group sequential methods are very popular in clinical trials; see Pocock (1977, 1983), Lewis and Berry (1994), Jennison and Turnbull (2000), and others.

Further optimization of sampling costs is possible if sampled groups are allowed to have variable sizes calculated in an optimal way at each interim point based on all of the already available data. Such sequential designs are called *sequentially planned procedures* or simply *sequential plans*. A formal definition follows.

Definition 1.1. A sequential plan $\mathcal{N}$ is a family of integer-valued functions

$$\mathcal{N} = \left\{ N^{(t)} : \mathbf{X}^{(t)} \longrightarrow \{0, 1, 2, \dots\} \right\}, \tag{1.1}$$

where $N^{(t)}$ is a random variable that returns the sample size of the next group given a sample of size t, $\mathbf{X}^{(t)} = (X_1, X_2, \dots, X_t)$. If $N^{(t)}(\mathbf{X}^{(t)}) = 0$, sampling terminates after observing $\mathbf{X}^{(t)}$, and a terminal decision is taken based on $\mathbf{X}^{(t)}$.

The ultimate goal is to find an *optimal sequential plan* from a set of all possible plans. Optimality is understood in the decision theoretic sense of minimizing the overall risk function $R(\theta, \mathcal{N}, \delta)$ that includes the expected decision loss and sampling cost,

$$R(\theta, \mathcal{N}, \delta) = \mathbf{E}_\theta \left\{ L(\theta, \delta(\mathbf{X})) + \sum_j T_j C_j \right\},$$

where $\mathcal{N}$ is a sequential plan, $\theta \in \Theta$ is the unknown parameter or state of nature, $\delta(\mathbf{X}) \in \mathcal{A}$ is a terminal decision that is based on the collected data $\mathbf{X} = (X_1, X_2, \dots)$ after sampling is terminated, $L(\theta, \delta)$ is the loss function corresponding to this decision, C_j is the cost of sampling a group of size j units, and T_j is the total number of sampled groups of size j. Sequential plan $\mathcal{N}_1$ is *R-better* than sequential plan $\mathcal{N}_2$ if and only if its risk is uniformly not higher for all values θ and is strictly lower than the risk of $\mathcal{N}_2$ for at least one θ. Plan $\mathcal{N}_1$ is *inadmissible* if and only if there exists a plan $\mathcal{N}_2$ that is R-better than $\mathcal{N}_1$.

Given a prior distribution $\pi(\theta)$ on Θ, the *Bayes risk* of a sequential plan is

$$r(\pi, \mathcal{N}, \delta) = \mathbf{E}_\pi R(\theta, \mathcal{N}, \delta).$$

Naturally, sequential plan $\mathcal{N}$ is *Bayes* if and only if it minimizes the Bayes risk over all possible sets of sequential plans, and plan $\mathcal{N}_1$ is *preferred* to plan $\mathcal{N}_2$ if and only if it has a lower Bayes risk.

Theoretical foundation of optimal sequential planning is described in Schmitz (1993). In particular, existence of Bayes sequential plans is shown. However, how should one determine the best sequential plan given a particular situation with specified losses and costs?

Basic principles of sufficiency and non randomization of optimal sequential plans are established in Schmegner and Baron (2004). Methods of risk evaluation are proposed in Schmegner and Baron (2007), allowing to compare the given sequential plans and choosing the optimal one. The asymptotic behavior of

minimum Bayes risks of sequential plans, as the sampling cost tends to zero, is obtained in Bartroff (2007). However, in general, determining the optimal sequential plan remains an open problem. Solution of this problem requires risk-optimization over a very large and rather unstructured set of all possible sequential plans of the form (1.1), and no algorithm is obtained so far that delivers the optimal solution, either analytically or numerically.

For certain relatively simple classes of problems, the optimal sequential plan can be calculated. In this article, we focus on the construction of an optimal sequential plan for a class of problems which may arise in a clinical trial.

Consider a clinical trial where a treatment is being tested for the presence of a rare but severe and adverse side effect such as a stroke or a heart attack. Depending on the treatment, even one case of this adverse side effect will be deemed unacceptable. Thus, we consider the following test:

$$H_0 : p = p_0 = 1 \quad \text{vs.} \quad H_1 : p = p_1 = u, \tag{1.2}$$

given $X_1, X_2, \ldots,$ independent and identically distributed (i.i.d.) observations from *Bernoulli*(p) where $u \in (0, 1)$. Here, a failure $(X_i = 0)$ is equivalent to an adverse effect, and thus under H_0 we assume that no such adverse effect ever occurs. The goal is to compute the Bayes sequential plan for this test under the prior distribution

$$\pi(p) = \begin{cases} \pi_0 & \text{if } p = 1, \\ 1 - \pi_0 & \text{if } p = u. \end{cases}$$

where $\pi_0 \in (0, 1)$ and the $0 - K$ loss function

$$L(\theta = p_i, \delta = p_j) = \begin{cases} 0 & \text{if } i = j, \\ 1 & \text{if } i \neq j, \end{cases} \quad i, j = 0, 1, \tag{1.3}$$

where $K = 1$ without loss of generality.

Obviously, as soon as a single failure is observed, sampling is terminated and the null hypothesis H_0 is rejected because there is no chance that it can be true if $X_i = 0$ is observed. Otherwise, sampling should continue until sufficient evidence is collected to accept H_0. In this case, after t sampled units, none of which registered an adverse effect,

$$\sum_{i=1}^{t} X_i = t;$$

hence, the decision on whether to continue sampling or not and what the optimal size of the next group is will only depend on the number of units already collected; that is, $N^{(t)}\left(\mathbf{X}^{(t)}\right) = N(t)$. Thus, all sequential plans considered will be of the form

$$\mathcal{N} = \left\{N^{(t)} : \{1, 2, \ldots\} \longrightarrow \{0, 1, 2, \ldots\}\right\}.$$

This also follows from the *sufficiency principle* of optimal sequential planning; see Schmegner and Baron (2004). Given the first t observations, we either stop and accept H_0 if $N^{(t)}(t) = 0$ or collect another batch of $N^{(t)}(t)$ observations $X_{t+1}, \ldots, X_{t+N(t)}$, if $N^{(t)}(t) > 0$.

In this article, we develop a methodology of calculating the optimal sampling function $N(t)$. By eliminating classes of inadmissible sequential plans, we reduce the search to a small manageable set of plans within which the optimal plan can be calculated.

2. ADMISSIBLE AND BAYES SEQUENTIAL PLANS

The following methodology of optimal sequential planning is proposed. First, we study admissible sequential plans and determine several necessary conditions for a sequential plan to be admissible.

Any plan that violates any of these conditions is inadmissible. After eliminating the large classes of inadmissible plans, the search reduces to a relatively small set of plans where the Bayes plan can be calculated directly.

One necessary condition of admissibility is that a sequential plan never rejects H_0 unless at least one failure is observed. Any plan to the contrary is inadmissible. Another necessary condition of admissibility is *monotonicity*.

Definition 2.1. Sequential plan $\mathcal{N}$ is *monotone* if and only if the function $N(t)$ is nondecreasing; that is, during the sampling process, the size of every next sampled group is never smaller than that of the previous group.

Intuitively, while testing (1.2), we become more confident that H_0 is true as we obtain more observations without a failure. Consequently, during later stages of this clinical trial, it should be safer to sample larger groups. We prove monotonicity of admissible plans by introducing the concept of *priorities*.

2.1. Terminal Decisions

For any sequential plan $\mathcal{N}$ and any prior distribution π, rejection of H_0 after observing data $\mathbf{X}$ is the Bayes terminal decision under the loss function (1.3) if and only if $P(H_1 \mid \mathbf{X}) \geq P(H_0 \mid \mathbf{X})$.

In particular, it means that any sequential plan $\mathcal{N}$ that takes at least one observation and then rejects H_0 without observing a failure $X_i = 0$ is *inadmissible* for testing (1.2).

Indeed, the posterior terminal decision loss for rejecting H_0 for $\mathcal{N}$ is $P\left(H_0 \mid \mathbf{X}^{(t)}\right)$. Sampling more observations without a failure provides a stronger support for the null hypothesis. Mathematically, for the event $A_t = \{X_1 = \cdots = X_t = 1\}$, we have

$$P\left(H_0 \mid \mathbf{X}^{(t)}\right) = \frac{\pi_0 P(A_t \mid H_0)}{\pi_0 P(A_t \mid H_0) + (1-\pi_0)P(A_t \mid H_1)} = \frac{\pi_0}{\pi_0 + (1-\pi_0)u^t},$$

which is an increasing function of the number of observed successes t.

Thus, if plan $\mathcal{N}$ is Bayes, and it rejects H_0 after observing t successes, then it is necessarily true that

$$\pi_0 < P\left(H_0 \mid \mathbf{X}^{(t)}\right) \leq P\left(H_1 \mid \mathbf{X}^{(t)}\right) < 1 - \pi_0. \tag{2.1}$$

Compare $\mathcal{N}$ with a sequential plan $\mathcal{N}_0$ that does not sample a single observation and makes a decision based solely on the prior distribution. In view of (2.1), the terminal decision of $\mathcal{N}_0$ is to reject H_0, and the terminal loss for $\mathcal{N}_0$ is π_0, which is strictly less than the terminal loss of $\mathcal{N}$. Moreover, the overall risk of $\mathcal{N}$ will have, in addition, some cost due to sampling t observations, whereas $\mathcal{N}_0$ will not. Hence, $\mathcal{N}_0$ is preferred over $\mathcal{N}$, in turn making any plan that rejects H_0 without a failure inadmissible.

This is the first step in our search for the optimal sequential plan. Next, we eliminate the mentioned inadmissible plans and focus on sequential plans that reject H_0 after observing a single failure and accept H_0 after observing a sufficient number of successes.

2.2. Priority Matrix for the Comparison of Sequential Plans

In nontrivial cases, analysis of admissibility requires comparison of risks of different plans. First, we develop a tool for comparing two sequential plans that are only different in the order of two stages. Then, two sequential plans can be compared by building a connecting sequence of "intermediate" sequential plans where any two adjacent plans coincide except for two stages.

Consider a sequential plan *plan* $\mathcal{N}_1$ that takes t observations in some predetermined manner, then one group of size k followed by one group of size m, and finally at most s more observations sampled according to a known plan S (Figure 1(a)). Also consider another sequential plan *plan* $\mathcal{N}_2$ that starts and ends the same way as plan $\mathcal{N}_1$ except that the k-group is interchanged with the m-group (Figure 1(b)). Without loss of generality, we can assume that $k < m$ (see below).

To compute the risks of $\mathcal{N}_1$ and $\mathcal{N}_2$, consider the moment when t units without a failure are already observed. Then the posterior risk r_i of each plan $\mathcal{N}_i$ $(i = 1, 2)$ can be written out explicitly. Conceptually, the risk is written as

$$\text{risk} = P(H_0)\mathbf{E}\,(\text{cost} + \text{decision loss}\,|\,H_0) + P(H_1)\mathbf{E}\,(\text{cost} + \text{decision loss}\,|\,H_1)\,.$$

There is no difference between the two plans until t observations are collected, and they are all successes. Thus, for the comparison of $\mathcal{N}_1$ and $\mathcal{N}_2$, it is equivalent to replace the prior probability π_i by $P\,(H_i\,|\,A_t)$ for $i = 1, 2$ and consider only the costs that occur after t observations are collected. Also, assume temporarily that $(t + k + m)$ observed successes are sufficient for both plans to stop and accept H_0 and, thus, sampling never proceeds beyond the k- and m-groups in Figure 1.

Then, consider the risks in connection to the k- and m-groups. Under H_0, failures are impossible; thus, both k- and m-groups are sampled almost surely. Furthermore, as shown above, an optimal plan never rejects H_0 unless a failure is observed and, therefore, its expected decision loss is 0; that is, it cannot reject incorrectly. Hence,

$$\mathbf{E}\,(\text{cost}\,|\,H_0) = C_k + C_m$$
$$\mathbf{E}\,(\text{loss}\,|\,H_0) = 0.$$

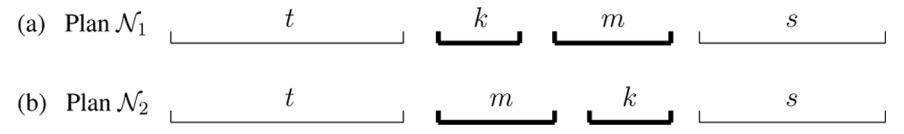

Figure 1. Sequential plans $\mathcal{N}_1$ and $\mathcal{N}_2$: changing the order of a k-group and an m-group.

Under H_1, a decision error can only occur when there are no failures. Thus, the expected loss is proportional to the probability (u) that we do not see a failure in the k- or m-groups. Under the 0-1 loss,

$$\mathbf{E}\,(\text{loss} \mid H_1) = P\,(A_{k+m} \mid H_1) = P\left(\bigcap_{i=1}^{k+m} \{X_i = 1\} \mid H_1\right) = u^{k+m}.$$

The costs of the two plans are different depending on the plan and depending on whether or not a failure was observed in the preceding group. For $\mathcal{N}_1$, the k-group is sampled almost surely, but the m-group is sampled only if no failure occurred in the k-group. This failure occurs with probability u^k, and thus

$$\mathbf{E}\,(\text{cost of } \mathcal{N}_1 \mid H_1) = C_k + u^k C_m.$$

Similarly, for plan $\mathcal{N}_2$,

$$\mathbf{E}\,(\text{cost of } \mathcal{N}_2 \mid H_1) = C_m + u^m C_k.$$

Combining the above expressions, we obtain the risks of $\mathcal{N}_1$ and $\mathcal{N}_2$, respectively, as

$$r_1 = \pi(C_k + C_m) + (1 - \pi)\{u^{k+m} + C_k + u^k C_m\},$$
$$r_2 = \pi(C_k + C_m) + (1 - \pi)\{u^{k+m} + C_m + u^m C_k\}.$$

Suppose now that plans $\mathcal{N}_1$ and $\mathcal{N}_2$ continue sampling beyond $(t + k + m)$ successes. Note that the probability of this (event A_{t+k+m}) is the same for the two plans. Thus, the risk (loss + cost) for sampling the remaining s observations is also the same. We denote this risk by

$$\mathbf{E}L_S = \mathbf{E}\left(\text{loss} + \text{cost of } S \mid \mathbf{X}^{(t+k+m)}\right),$$

a term that should be added to r_1 and r_2 in this case.

The difference Δ_{km} between the two risks, not affected by $\mathbf{E}L_S$, determines which plan is preferred,

$$\begin{aligned}\Delta_{km} &= r_1 - r_2 \\ &= (1 - \pi)\{C_k + u^k C_m - (C_m + u^m C_k)\} \\ &= (1 - \pi)\{C_k(1 - u^m) - C_m(1 - u^k)\}. \end{aligned} \tag{2.2}$$

If $\Delta_{km} < 0$, then *plan* $\mathcal{N}_1$ has a lower risk; hence, it is preferred over *plan* $\mathcal{N}_2$ and vice versa if $\Delta_{km} > 0$. We can compute Δ_{km} for any positive integers k and m given u, C_k, and C_m. Note that Δ_{km} is antisymmetric; that is, $\Delta_{km} = -\Delta_{mk}$. Thus, there is no loss of generality in assuming $k < m$.

Notice that the sign of Δ_{km} is independent of t and S. Hence, if the Bayes plan includes a k-group and an m-group, these groups will be sampled in a predetermined order, an order that is the same regardless of where these two groups appear in the course of the whole sequential plan.

Also, Δ_{km} is proportional to the difference of frequentist risks of plans $\mathcal{N}_1$ and $\mathcal{N}_2$ under H_1 whose sign determines which plan is R-better. Therefore, any admissible sequential plan should have these two groups in the same order as the Bayes plan.

Finally, we mention a few words for the case when $\Delta_{km} = 0$. The trivial case where this occurs is when $k = m$, but suppose $k < m$ and $\Delta_{km} = 0$. This means that sampling a k-group produces the same risk as sampling an m-group. In such cases, by convention, we sample the k-group first.

Based on Δ_{km}, we now define the concept of *priorities* p_{km}.

Definition 2.2. A group of size k has priority over a group of size m if and only if k-groups are always sampled prior to m-groups, in which case we let priority $p_{km} = 0$ and $p_{km} = 1$ otherwise.

Priorities can be defined from Δ_{km}, the difference of risks,

$$p_{km} = \begin{cases} 0 & \text{if } \Delta_{km} \leq 0, \\ 1 & \text{if } \Delta_{km} > 0. \end{cases}$$

If $p_{km} = 0$, plan $\mathcal{N}_1$ that samples a k-group and an m-group consecutively is R-better than (and preferred to) plan $\mathcal{N}_2$ that samples the same groups in reverse order. Thus, any admissible (and any Bayes) plan that includes these consecutive groups samples them in the same order. We summarize these results in the following lemma.

Lemma 2.1 (Order of Two Consecutively Sampled Groups)**.** *Any Bayes or admissible sequential plan for testing* (1.2) *samples any of its two consecutive groups in the same order determined only by the group sizes and costs, independent of the prior distribution and any groups collected earlier or later in the course of this plan. For any k and m, priority p_{km} is invariant for all admissible and Bayes sequential plans.*

To summarize the comparisons of risks between any given k and m, we arrange the priorities p_{km} into an $n \times n$ *priority matrix*.

Definition 2.3. A priority matrix $\mathbf{P}$ is the matrix of priorities $\{p_{km} \mid k = 1, \ldots, n; m = 1, \ldots, n\}$.

The priority matrix can be used to quickly determine whether or not the Bayes sequential plan for testing (1.2) is monotone. Notice that all diagonal elements are 0. This is the case when $k = m$, and hence the two plans coincide, $\mathcal{N}_1 = \mathcal{N}_2$. If $\mathbf{P}$ is lower triangular, it implies that $p_{km} = 0$ for all $k \leq m$. In such cases, the next sampled group can only be as large or larger than the preceding group. Consequently, such a sequential plan is monotone; for example,

$$\mathbf{P_1} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \end{bmatrix}, \quad \mathbf{P_2} = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}.$$

The dimension n of each priority matrix is the maximum size of a group that can be sampled under the given sequential plan.

2.3. Transitivity

A natural question that arises is whether or not Δ_{km} is *transitive*.

Definition 2.4. We say that Δ_{km} is transitive if and only if $\Delta_{km} < 0$ and $\Delta_{ml} < 0$, then $\Delta_{kl} < 0$ for any given k, m, and l.

If transitivity holds, then knowing that a k-group should always precede an m-group, and an m-group should always precede an l-group, we can conclude that a k-group should always precede an l-group.

However, if this does not hold true, then priorities are not consistent, and such a priority matrix serves no purpose. For example, suppose that $\Delta_{12} < 0$, $\Delta_{23} < 0$, and $\Delta_{13} > 0$. According to such a priority matrix, 1-groups should precede 2-groups, which in turn should precede 3-groups; however, 3-groups should precede 1-groups. Clearly, these three constraints are in contradiction to each other and, thus, there is no Bayes sequential plan satisfying these priorities.

Lemma 2.2 (Transitivity)**.** *For any given group sizes k, l, and m, with the corresponding costs C_k, C_m, and C_l, respectively, if $\Delta_{km} < 0$ and $\Delta_{ml} < 0$, then $\Delta_{kl} < 0$.*

Proof. From (2.2), the inequality $\Delta_{km} < 0$ holds if and only if

$$\frac{C_k}{1-u^k} < \frac{C_m}{1-u^m}.$$

Similarly, $\Delta_{ml} < 0$ is equivalent to

$$\frac{C_m}{1-u^m} < \frac{C_l}{1-u^l}.$$

Combining these two inequalities, we get

$$\frac{C_k}{1-u^k} < \frac{C_m}{1-u^m} < \frac{C_l}{1-u^l},$$

which is equivalent to $\Delta_{kl} < 0$. □

By extending transitivity to more than three groups, we can draw the following corollaries.

Corollary 2.1. *For a Bayes sequential plan $\mathcal{N}$:*

(i) *If $\Delta_{k_i,k_{i+1}} < 0$ for group sizes $k_1, \ldots, k_j$, then $\Delta_{k_1,k_j} < 0$, where $i = 1, 2, \ldots, j-1$.*
(ii) *If $\Delta_{k,m} < 0$ for any group sizes k and m, then the Bayes sequential plan for testing* (1.2) *samples any k-group before any m-group, even if these groups are not sampled consecutively.*
(iii) *There is a Bayes sequential plan that samples groups of the same size consecutively.*

Proof. Assertion (i) is a direct extension of Lemma 2.2 that can be proven by induction.

To prove (ii), assume that after sampling an m-group, the Bayes plan samples groups of sizes $k_1, \ldots, k_j$, followed by a k-group. According to the priority matrix, this can only happen if $\Delta_{m,k_1} < 0$, $\Delta_{k_i,k_{i+1}} < 0$ for $i = 1, \ldots, j-1$, and $\Delta_{k_j,k} < 0$. By the generalized transitivity (i), this yields $\Delta_{k,m} \geq 0$. Thus, if $\Delta_{k,m} < 0$, the Bayes plan cannot sample an m-group before a k-group anywhere during the sampling design.

To show (iii), assume that k-groups are not sampled consecutively. That is, there is a group of a different size, m, between the k-groups. According to (ii), if the plan is Bayes, it implies that $\Delta_{km} \geq 0$ and $\Delta_{mk} \geq 0$. Therefore, $\Delta_{km} = 0$ for each group that is planned between two k-groups. On the other hand, if $\Delta_{km} = 0$, then k-groups and m-groups can be sampled in *any order* without changing the risk. Thus, a permutation of k-groups and all of the groups between them that samples groups of the same sizes consecutively has the same Bayes risk and thus it is also a Bayes plan. □

Due to transitivity, the priority matrix can be used to determine the order in which groups of different sizes should be sampled by the optimal plan. Furthermore, we conclude that groups of the same size should be sampled consecutively.

2.4. Monotonicity

In this section, we show that any Bayes sequential plan is monotone. We prove that if the priority matrix is not lower triangular, then the optimal sequential plan does not sample any groups of size k whenever $\Delta_{km} > 0$ for some $k < m$. Note that because the priority matrix is problem specific and not plan specific, the structure of the matrix is predetermined, and hence we may have a situation where $\Delta_{km} > 0$ for some $k < m$. However, we show that in such cases the Bayes plan will never sample a k-group. That is, a plan that samples an m-group prior to a k-group is preferred over a plan that does the opposite, but there is a plan that is preferred over both where no k-groups are sampled.

Here and later, let t_k be the maximum possible number of k-groups sampled (consecutively, as we know) by a sequential plan $\mathcal{N}$.

Theorem 2.1 (Monotonicity)**.** *Let $\mathcal{N}$ be a Bayes sequential plan. If $\Delta_{km} > 0$ for some $k < m$, then $t_k = 0$. That is, the Bayes plan samples groups in nondecreasing order. If the priority matrix indicates otherwise, then the number of times the smaller group is sampled is* 0 *almost surely.*

Proof. Suppose that plan $\mathcal{N}$ is Bayes, although it takes at least one k-group while $\Delta_{km} > 0$ for some $k < m$. Without loss of generality, we can assume that the k-group is the first group sampled in the course of this plan. If $t > 0$, we replace the prior distribution with the posterior given t observations and consider the remaining part of the plan that starts at t. The reason we can do this is because we assume $\mathcal{N}$ is Bayes, and thus at any time t, it chooses the Bayesian (optimal) continuation. That is, for every t observations collected, the next group size equals the first group size for a Bayes sequential plan under the prior distribution $\pi\left(\mathbf{X}^{(t)}\right)$, which is equivalent to the posterior probability of H_0 given all t observations. Note that this is invariant with respect to how the t observations were collected; that is, number of previous groups and sizes.

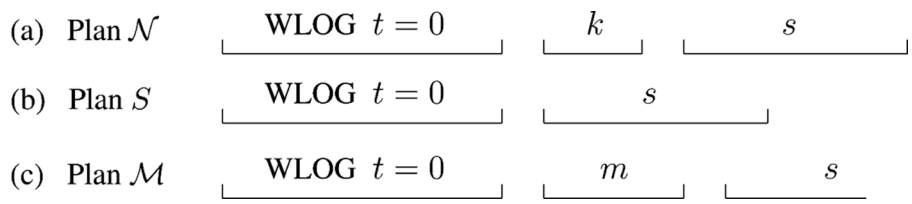

Figure 2. Three plans for the proof of monotonicity.

Consider three sequential plans. The given Bayes plan $\mathcal{N}$ starts with a k-group and then proceeds according to some arbitrary plan S (Figure 2(a)). The second plan is the plan S itself, which is the original plan $\mathcal{N}$ without sampling its first planned group (Figure 2(b)). The third plan $\mathcal{M}$ samples an m-group instead of a k-group and then proceeds according to plan S (Figure 2(c)).

Bayes risks of these plans can be written out explicitly as

$$\begin{aligned} r_S &= r(S) = \pi C_S + (1-\pi)R_1(S), \\ r_{\mathcal{N}} &= r(k, S) = C_k + \pi C_S + (1-\pi)u^k R_1(S), \\ r_{\mathcal{M}} &= r(m, S) = C_m + \pi C_S + (1-\pi)u^m R_1(S), \end{aligned}$$

where C_S and $R_1(S)$ are, respectively, the cost and expected loss plus cost functions of plan S.

Because plan $\mathcal{N}$ is Bayes, its Bayes risk is the lowest of the three, and thus

$$r_S - r_{\mathcal{N}} = (1-u^k)(1-\pi)R_1(S) - C_k \geq 0, \tag{2.3}$$

and

$$r_{\mathcal{M}} - r_{\mathcal{N}} = C_m - C_k + (u^m - u^k)(1-\pi)R_1(S) \geq 0. \tag{2.4}$$

From (2.3) we get

$$R_1(S) \geq \frac{C_k}{(1-u^k)(1-\pi)}.$$

Combining this with (2.4) and noticing that $u^m < u^k$ for $m > k$, we obtain

$$\begin{aligned} 0 &\leq C_m - C_k - (u^k - u^m)(1-\pi)R_1(S) \\ &\leq C_m - C_k - \frac{(u^k - u^m)(1-\pi)C_k}{(1-u^k)(1-\pi)} \\ &= \frac{C_m(1-u^k) - C_k(1-u^m)}{1-u^k}. \end{aligned}$$

This is a contradiction because $\Delta_{km} > 0$ implies that

$$C_k(1-u^m) - C_m(1-u^k) > 0.$$

Therefore, if $\Delta_{km} > 0$, then $t_k = 0$; that is, no k-groups should be sampled. □

Thus, the group sizes should monotonically increase as sampling continues. All nonmonotone plans can be eliminated from our search for the Bayes sequential plan. The next result yields further reduction of the scope of our search.

2.5. Maximum Number of Groups

Consider a *k-group sequential plan*, that is, a family of integer-valued functions $\mathcal{N}$ such that

$$\mathcal{N} = \left\{ N^{(t)} : \mathbf{X}^{(t)} \longrightarrow \{0, k\}, k = 1, 2, \dots \right\}.$$

Such a plan can only sample groups of size k. Intuitively, the Bayes sequential plan would not need as many k-groups as the k-group sequential plan because it has the option of taking groups of other sizes and, therefore, it can attain the same loss at a lower cost. In this section, we show that the Bayes sequential plan takes no more k-groups than the optimal number of groups in the k-group sequential plan for any k.

We first show that the *Bayes k-group sequential plan* $\mathcal{K}$ (which minimizes the Bayes risk among all k-group sequential plans) has a closed form. Let t_k be the maximum number of k-groups under plan $\mathcal{K}$ and T_k be the actual (random) number of sampled k-groups. The Bayes risk of this plan is

$$\begin{aligned}
r(t_k) &= \mathbf{E}^{\pi}\mathbf{E}_{\theta}\left\{L(\theta, \delta) + C_k T_k\right\} \\
&= \pi_0 \mathbf{E}_{H_0}\left\{L(\theta, \delta) + C_k T_k\right\} + (1 - \pi_0)\mathbf{E}_{H_1}\left\{L(\theta, \delta) + C_k T_k\right\} \\
&= \pi_0(0 + C_k t_k) + (1 - \pi_0)\left\{ P\left(\sum_{i=1}^{kt_k} X_i = kt_k\right)(1 + C_k t_k) \right. \\
&\quad \left. + C_k \sum_{j=1}^{t_k} jP\left(\sum_{i=1}^{k(j-1)} X_i = k(j-1)\right) P\left(\sum_{k(j-1)+1}^{kj} X_i < k\right)\right\} \\
&= \pi_0 C_k t_k + (1 - \pi_0)\left\{ u^{kt_k}(1 + C_k t_k) + C_k \sum_{j=1}^{t_k} j u^{k(j-1)}(1 - u^k)\right\} \\
&= \pi_0 C_k t_k + (1 - \pi_0)\left\{ u^{kt_k} + C_k \frac{1 - u^{kt_k}}{1 - u^k}\right\}.
\end{aligned}$$

Differentiating with respect to t_k, it is easy to see that the maximum number of k-groups taken by $\mathcal{K}$ is either the floor or ceiling (whichever one has the smaller risk) of

$$\tilde{t}_{k_g} = (k \ln u)^{-1} \ln\left\{\left(\frac{\pi_0}{1 - \pi_0}\right)\frac{C_k(1 - u^k)}{k|\ln u|(1 - u^k - C_k)}\right\}. \tag{2.5}$$

Now we show that this quantity for any k serves as the upper bound for the number of k-groups in the overall Bayes sequential plan.

Theorem 2.2. *Let t_{k_g} be the maximum number of k-groups for the Bayes k-group sequential plan under a given prior (or posterior) distribution π. Let t_k be the maximum number of k-groups for the unrestricted Bayes plan given the same distribution π. Then,*

$$t_k \leq t_{k_g} \quad \text{for all } k = 1, 2, \dots.$$

Proof. Assume $t_k > t_{k_g}$ for some k. Let t be the number of observations collected before any of the k-groups, and suppose that according to the priority matrix, k-groups should be sampled next.

Consider the following two sequential plans. Plan $\mathcal{N}$ takes a maximum of t_k k-groups and stops. Plan $\mathcal{K}$ takes a maximum of t_{k_g} k-groups and then stops. Denote the Bayes risks of these plans by $r(t_k)$ and $r(t_{k_g})$, respectively, which can be written out explicitly as

$$r_{\mathcal{N}} = r(t_k) = \pi C_k t_k + (1-\pi)\left\{u^{kt_k} + C_k \mathbf{E}_{H_1} T_k\right\},$$

and

$$r_{\mathcal{K}} = r(t_{k_g}) = \pi C_k t_{k_g} + (1-\pi)\left\{u^{kt_{k_g}} + C_k \mathbf{E}_{H_1} T_{k_g}\right\},$$

where T_k and T_{k_g} again denote the actual (random) number of k-groups taken. Notice that by stopping after the k-groups, plan $\mathcal{N}$ essentially proceeds according to a k-group sequential scheme. Then, because t_{k_g} is Bayes among the k-group sequential plans,

$$\begin{aligned} 0 &\le r(t_k) - r(t_{k_g}) \\ &= \pi C_k (t_k - t_{k_g}) + (1-\pi)\left\{u^{kt_k} - u^{kt_{k_g}} + C_k \mathbf{E}_{H_1}(T_k - T_{k_g})\right\} \\ &= C_k \mathbf{E}_\pi (T_k - T_{k_g}) + (1-\pi)(u^{kt_k} - u^{kt_{k_g}}). \end{aligned}$$

Hence,

$$C_k \mathbf{E}_\pi (T_k - T_{k_g}) \ge (1-\pi)(u^{kt_{k_g}} - u^{kt_k}). \tag{2.6}$$

Now consider two more sequential plans. Suppose that plan $\mathcal{N}_S$ is the unrestricted Bayes plan that takes at most t_k k-groups followed by some scheme S. Let $\mathcal{K}_S$ be a plan that takes at most t_{k_g} k-groups and then proceeds according to the same S. The Bayes risks of these plans are

$$r_{\mathcal{N}_S} = r(t_k, S) = \pi(t_k C_k + C_s) + (1-\pi)\left\{u^{kt_k+s} + C_k \mathbf{E}_{H_1} T_k + u^{kt_k} \mathbf{E}_{H_1} C_S\right\},$$

and

$$r_{\mathcal{K}_S} = r(t_{k_g}, S) = \pi(t_{k_g} C_k + C_s) + (1-\pi)\left\{u^{kt_{k_g}+s} + C_k \mathbf{E}_{H_1} T_{k_g} + u^{kt_{k_g}} \mathbf{E}_{H_1} C_S\right\},$$

where s is the maximum number of observations obtained by sampling S, and C_s and C_S are the maximum and actual (random) costs of sampling S, respectively. Because the first plan is the unrestricted Bayes plan,

$$\begin{aligned} 0 &\le r(t_{k_g}, S) - r(t_k, S) \\ &= \pi\{C_k(t_{k_g} - t_k)\} + (1-\pi)\{u^s(u^{kt_{k_g}} - u^{kt_k}) + C_k \mathbf{E}_{H_1}(T_{k_g} - T_k) + \mathbf{E}_{H_1} C_S (u^{kt_{k_g}} - u^{kt_k})\} \\ &= C_k \mathbf{E}_\pi (T_{k_g} - T_k) + (1-\pi)(u^{kt_{k_g}} - u^{kt_k})(u^s + \mathbf{E}_{H_1} C_S) \\ &= -C_k \mathbf{E}_\pi (T_k - T_{k_g}) + (1-\pi)(u^{kt_{k_g}} - u^{kt_k}) R_1(S). \end{aligned}$$

Hence,

$$C_k \mathbf{E}_\pi(T_k - T_{k_g}) \leq (1-\pi)(u^{kt_{k_g}} - u^{kt_k}) R_1(S). \tag{2.7}$$

Now subtracting the corresponding sides of (2.7) from (2.6) yields

$$0 \geq (1-\pi)(u^{kt_{k_g}} - u^{kt_k}) \left[1 - R_1(S)\right].$$

Of the three quantities on the right-hand side, $1-\pi > 0$ because it is the posterior probability of H_1, and $u^{kt_{k_g}} - u^{kt_k} > 0$ because of our initial assumption, $t_{k_g} < t_k$. Let us interpret the third term, $1 - R_1(S)$. Under the $0-1$ loss function that we use, the maximum overall decision loss is 1. If $1 - R_1(S) \leq 0$, then our expected loss plus cost of sampling S (under H_1) would be at least 1. Such a plan S would not be Bayes or even admissible because the trivial plan that takes no more observations and randomly accepts or rejects H_0 is preferred. The cost of said plan is 0 and its loss is less than 1. Hence, $1 - R_1(S) > 0$, which leads to a contradiction. Therefore,

$$t_k \leq t_{k_g} \quad \text{for all } k = 1, 2, \ldots. \qquad \square$$

This theorem provides an important tool for the reduction of our search for the Bayes sequential plan. Indeed, it bounds the maximum number of k-groups by the t_{k_g}, the number of k-groups sampled by the Bayes group sequential plan, which is given explicitly by (2.5).

Moreover, comparing with the mentioned trivial no-data plan, the maximum group size of the Bayes plan is bounded by $C^{-1}(1/2)$, the smallest group size whose cost is at least $1/2$.

Concluding the results of Sections 2.1–2.5, the search for the Bayes sequential plan is now reduced to a finite set of plans that take groups of equal sizes consecutively, are monotone with a nondecreasing sampling function $N(t)$, with at most $C^{-1}(1/2)$ sampling units in each sampled group and at most t_{k_g} groups of size k for each $k = 1, \ldots, C^{-1}(1/2)$.

3. EXAMPLES

Applying all of the obtained criteria, the class of sequential plans we need to consider is reduced significantly. Therefore, in most situations, the unrestricted Bayes plan can be calculated by manual comparison of risks. Below are a few examples, where N_j represents the size of the jth group. We let $C_k = a + ck$, where a is the cost associated with the group and c is the cost of one observation (fixed cost plus variable cost, as in Schmitz, 1993; Schmegner and Baron, 2004; Schmegner, 2009).

1. $\pi_0 = 0.5$, $u = 0.7$, $a = 0.01$, $c = 0.01$
   $N_1 = 3$, $N_2 = 6$,
   Bayes risk $= 0.1072$
2. $\pi_0 = 0.5$, $u = 0.9$, $a = 0.01$, $c = 0.01$
   $N_1 = 6$, $N_2 = 6$, $N_3 = 8$
   Bayes risk $= 0.2421$

3. $\pi_0 = 0.5$, $u = 0.7$, $a = 0.01$, $c = 0.001$
   $N_1 = 15$
   Bayes risk $= 0.0274$
4. $\pi_0 = 0.5$, $u = 0.9$, $a = 0.0001$, $c = 0.01$
   $N_1 = N_2 = \cdots = N_{21} = 1$
   Bayes risk $= 0.2057$

Notice that every plan follows the aforementioned properties and makes intuitive sense. The second plan samples bigger groups and more observations overall than the first plan because the value of u is larger and closer to H_0, making it a more difficult test. Under the alternative, it is harder to see a failure, and hence we need more observations. In the third example, the group cost is high relative to the observation cost; thus, we should sample bigger group(s). On the other hand, in the fourth example, the group cost is low relative to the observation cost; thus, we sample more groups of smaller sizes. We compare the performance of the above examples to the optimal pure sequential scheme that always takes one observation at a time. The differences show the expected savings due to sampling in groups.

1. $\pi_0 = 0.5$, $u = 0.7$, $a = 0.01$, $c = 0.01$
   Minimum Bayes risk $= 0.1072$
   Minimum pure sequential risk $= 0.1402$
   Difference $= 0.033$
2. $\pi_0 = 0.5$, $u = 0.9$, $a = 0.01$, $c = 0.01$
   Minimum Bayes risk $= 0.2421$
   Minimum pure sequential risk $= 0.3315$
   Difference $= 0.0894$
3. $\pi_0 = 0.5$, $u = 0.7$, $a = 0.01$, $c = 0.001$
   Minimum Bayes risk $= 0.0274$
   Minimum pure sequential risk $= 0.0869$
   Difference $= 0.0595$
4. $\pi_0 = 0.5$, $u = 0.9$, $a = 0.0001$, $c = 0.01$
   Minimum Bayes risk $= 0.2057$
   Minimum pure sequential risk $= 0.2057$
   Difference $= 0$ (the Bayes plan is pure sequential)

The difference in risks should be judged in terms of the given costs. Hence, in the first example, the expected saving equals about the cost of three observations.

4. EXTENSION: A RARE ADVERSE EFFECT

A natural extension to the above problem is to consider the test

$$H_0 : p = p_0 = 1 - \varepsilon \quad \text{vs.} \quad H_1 : p = p_1 = u, \tag{4.1}$$

where $0 < \varepsilon \ll 1$. Essentially, the treatment will be accepted if the probability of the severe adverse effect is sufficiently low (ε) although not necessarily zero as in testing (1.2) of the previous sections. For sufficiently small ε, a single failure ($X_i = 0$) still leads to sampling being terminated as soon as possible. We shed some light on

the value of ε by considering the type I error for the extension. Note that the type I error rate for the previous case was 0 because we cannot reject H_0 incorrectly. Suppose the maximum sample size collected for the plan is n; that is, when there are no failures. Then the type I error rate is

$$P(\text{at least one } X_i = 0 \mid H_0) = 1 - (1 - \varepsilon)^n.$$

Suppose the test is to be conducted at some level α and thus we must have that

$$1 - (1 - \varepsilon)^n \leq \alpha.$$

Now solving for ε yields

$$\varepsilon \leq 1 - \exp\left[\frac{\ln(1 - \alpha)}{n}\right].$$

In addition, we note that because there are costs associated to sampling groups as well as observations, there is almost always an upper bound on the maximum sample size n. Thus, given this n and some α, we can determine whether or not ε is sufficiently small to implement our proposed sequential plan.

In this section, we establish results for testing (4.1) analogous to testing (1.2), starting with the priority matrix.

4.1. Priority Matrix (Extension)

Suppose we are at the point where we have collected t observations without a failure, and we are choosing between two plans, $\mathcal{N}_1$, which takes a k-group next, and $\mathcal{N}_2$, which takes an m-group next, before proceeding with the same continuation S. The corresponding Bayes risks are

$$r_1 = \pi\{1 - (1 - \varepsilon)^{k+m} + C_k + (1 - \varepsilon)^k C_m\} + (1 - \pi)\{u^{k+m} + C_k + u^k C_m\} + r(S),$$
$$r_2 = \pi\{1 - (1 - \varepsilon)^{k+m} + C_m + (1 - \varepsilon)^m C_k\} + (1 - \pi)\{u^{k+m} + C_m + u^m C_k\} + r(S),$$

where $\pi = \pi(p \mid \mathbf{X}^{(t)})$ is the posterior probability of H_0 after t observations without a failure and $r(S)$ is the expected loss plus cost of sampling S. Now subtracting r_2 from r_1 and denoting this quantity by Δ^*_{km}, we have

$$\begin{aligned}\Delta^*_{km} &= r_1 - r_2 \\ &= \pi\{C_k[1 - (1 - \varepsilon)^m] - C_m[1 - (1 - \varepsilon)^k]\} + (1 - \pi)\{C_k(1 - u^m) - C_m(1 - u^k)\} \\ &= \pi\{C_k[1 - (1 - \varepsilon)^m] - C_m[1 - (1 - \varepsilon)^k]\} + \Delta_{km},\end{aligned}$$

where Δ_{km} is the difference in risks for the testing (1.2) when $\varepsilon = 0$. Denoting by

$$P_k = \pi(1 - \varepsilon)^k + (1 - \pi)u^k$$

the marginal probability of observing no failure among the first k sampled units, we obtain that

$$\begin{aligned}\Delta^*_{km} &= C_k - C_m - C_k P_k + C_m P_m \\ &= C_k(1 - P_m) - C_m(1 - P_k).\end{aligned}$$

Again, the sign of Δ^*_{km} indicates the better plan. If $\Delta^*_{km} < 0$, then *plan* $\mathcal{N}_1$ has a lower risk; hence, it is preferred over *plan* $\mathcal{N}_2$, and vice versa if $\Delta^*_{km} > 0$. Then for any positive integer k and m, given C_k, C_m, ε, and u, we can determine which group should precede the other in the course of the Bayes plan. As before, arrange priorities p^*_{km} in an $n \times n$ *priority matrix* $\mathbf{P}^*$, where

$$p^*_{km} = \begin{cases} 0 & \text{if } \Delta^*_{km} \le 0, \\ 1 & \text{if } \Delta^*_{km} > 0. \end{cases}$$

Similar to the previous case of $\epsilon = 0$, we use Δ^*_{km} to show transitivity and monotonicity of the Bayes sequential plan for testing (4.1).

4.2. Transitivity (Extension)

As in the previous section, transitivity holds for the Bayes sequential plan. That is, for any given costs C_k, C_m, and C_l, $k < l < m$, having $\Delta^*_{km} < 0$ and $\Delta^*_{ml} < 0$ yields $\Delta^*_{kl} < 0$. Indeed, if $\Delta^*_{km} < 0$,

$$C_k(1 - P_m) - C_m(1 - P_k) < 0$$

and therefore,

$$\frac{C_k}{1 - P_k} < \frac{C_m}{1 - P_m}.$$

Similarly, $\Delta^*_{ml} < 0$ is equivalent to

$$\frac{C_m}{1 - P_m} < \frac{C_l}{1 - P_l}.$$

Combining these two inequalities we get

$$\frac{C_k}{1 - P_k} < \frac{C_m}{1 - P_m} < \frac{C_l}{1 - P_l},$$

which implies $\Delta^*_{kl} < 0$.

Due to transitivity, the priority matrix can be used to determine the order in which various group sizes should be sampled. Using transitivity arguments repeatedly, we can conclude that all groups of the same size should be sampled one after another.

4.3. Monotonicity (Extension)

We now show that under certain conditions, the Bayes sequential plan for testing (4.1) is monotone.

Theorem 4.1 (Monotonicity). *Let $\mathcal{N}$ be the Bayes sequential plan. If $\Delta^*_{km} > 0$ for some $k < m$, then $t_k = 0$ for $\mathcal{N}$, where t_k is the (maximum) number of k-groups sampled, provided that*

$$\left| \frac{C_m(1 - P_k) - C_k(1 - P_m)}{1 - P_k} \right| > |Q|, \tag{4.2}$$

where

$$Q = \pi[(1-\varepsilon)^k - (1-\varepsilon)^m] - (P_k - P_m)\frac{\pi[1-(1-\varepsilon)^k]}{1-P_k}.$$

Proof. Assume that $\Delta^*_{km} > 0$ for some $k < m$, and there is no sequential plan that is preferred to a plan that takes at least one k-group. Without loss of generality, assume that no observations are sampled before the k-groups; otherwise, replace the prior distribution with the posterior, given these sampled observations.

Consider three sequential plans. One plan proceeds according to some arbitrary plan S. The second plan takes one k-group before proceeding to S. The third plan takes one m-group before S. Denote the Bayes risks of these plans as $R(S)$, $R(k, S)$, and $R(m, S)$, respectively. The risks can be written out explicitly as

$$\begin{aligned} R(S) &= r(S) \\ R(k,S) &= C_k + \pi[1-(1-\varepsilon)^k] + P_k r(S), \\ R(m,S) &= C_m + \pi[1-(1-\varepsilon)^m] + P_m r(S), \end{aligned}$$

where $r(S)$ is the expected loss plus cost of executing plan S. Because of our assumption of the Bayes plan, we have

$$R(S) - R(k,S) = (1-P_k)r(S) - C_k - \pi[1-(1-\varepsilon)^k] \geq 0, \tag{4.3}$$

$$R(m,S) - R(k,S) = C_m - C_k - (P_k - P_m)r(S) + \pi[(1-\varepsilon)^k - (1-\varepsilon)^m] \geq 0, \tag{4.4}$$

From (4.3) we get

$$r(S) \geq \frac{C_k + \pi[1-(1-\varepsilon)^k]}{1-P_k}.$$

Combining this with (4.4),

$$\begin{aligned} 0 &\leq C_m - C_k - (P_k - P_m)r(S) + \pi[(1-\varepsilon)^k - (1-\varepsilon)^m] \\ &\leq C_m - C_k - (P_k - P_m)\frac{C_k + \pi[1-(1-\varepsilon)^k]}{1-P_k} + \pi[(1-\varepsilon)^k - (1-\varepsilon)^m] \\ &= \frac{C_m(1-P_k) - C_k(1-P_k) - C_k(P_k - P_m)}{1-P_k} \\ &\quad + \pi[(1-\varepsilon)^k - (1-\varepsilon)^m] - (P_k - P_m)\frac{\pi[1-(1-\varepsilon)^k]}{1-P_k} \\ &= \frac{C_m(1-P_k) - C_k(1-P_k)}{1-P_k} + Q. \end{aligned}$$

However, $\Delta^*_{km} > 0$ implies that,

$$C_k(1-P_m) - C_m(1-P_k) > 0,$$

hence,

$$\frac{C_m(1-P_k) - C_k(1-P_m)}{1-P_k} < 0.$$

This is a contradiction, because

$$\left|\frac{C_m(1-P_k)-C_k(1-P_m)}{1-P_k}\right| > |Q|,$$

and hence

$$\frac{C_m(1-P_k)-C_k(1-P_m)}{1-P_k} + Q < 0.$$

Therefore, if $\Delta^*_{km} > 0$, then no k-groups are sampled and, therefore, the Bayes plan is monotone. □

Hence, as long as the additional condition of (4.2) is satisfied, the Bayes sequential plan for testing (4.1) would still be monotone.

5. CONCLUSION

In general, determining the optimal sequential plan remains an open problem mainly because it requires risk optimization over a rather unstructured set of all plans. For a simple class of problems that may arise in testing a treatment for a rare but severe adverse effect, a number of properties of the Bayes sequential plan such as transitivity and monotonicity were proven. This allows one to reduce the overall scope of the search to a small manageable set of plans.

ACKNOWLEDGMENTS

We thank both anonymous reviewers and the Associate Editor for their thoughtful and constructive comments that led to a seriously revised and improved version of this article. Research support of S. Suzuki by the National Heart, Lung, and Blood Institute, National Institute on Alcohol Abuse and Alcoholism, and National Cancer Institute is gratefully acknowledged. Research support of M. Baron by the National Science Foundation and the National Security Agency is gratefully acknowledged.

REFERENCES

Bartroff, J. (2007). Asymptotically Optimal Multistage Tests of Simple Hypotheses, *Annals of Statistics* 35: 2075–2105.

Jennison, C. and Turnbull, B. W. (2000). *Group Sequential Methods with Applications to Clinical Trials*, Boca Raton: Chapman & Hall.

Lewis, R. J. and Berry, D. A. (1994). Group Sequential Clinical Trials: A Classical Evaluation of Bayesian Decision-Theoretic Designs, *Journal of American Statistical Association* 89: 1528–1534.

Pocock, S. J. (1977). Group Sequential Methods in the Design and Analysis of Clinical Trials, *Biometrika* 64: 191–199.

Pocock, S. J. (1983). *Clinical Trials: A Practical Approach*, New York: Wiley.

Schmegner, C. (2009). Truncated Sequential Procedures, *Sequential Analysis* 28: 406–422.

Schmegner, C. and Baron, M. (2004). Principles of Optimal Sequential Planning, *Sequential Analysis* 23: 11–32.

Schmegner, C. and Baron, M. (2007). Sequential Plans and Risk Evaluation, *Sequential Analysis* 26: 335–354.

Schmitz, N. (1993). *Optimal Sequentially Planned Decision Procedures*, New York: Springer-Verlag.