The Annals of Statistics
1986, Vol. 14, No. 1, 283–297

BAYESIAN MULTISTAGE DECISION PROBLEMS

BY JEFFREY A. WITMER

University of Florida

Two treatments that yield Bernoulli outcomes are available in a clinical trial. One success probability is known. A probability distribution reflects opinion about the other success rate. N patients are to be treated, with N possibly unknown, in a multistage trial. The goal is to maximize the total number of successes on the N patients.

Optimal lengths for each stage and optimal treatment allocations are found for two-stage trials with N known.

When N is unknown the problem is shown to be equivalent to that of discounting future observations. Optimal stage lengths and treatment allocations are characterized for distributions on N that yield *regular* discount sequences. This class of distributions includes the geometric family, which is given special consideration.

It is shown that if the number of stages in the trial is fixed and if the distribution on N yields a regular discount sequence, then it is optimal to use the known treatment in the last stage only. This extends the work of Berry and Fristedt (1979).

1. Introduction. Experimenters are often interested in comparing two treatments that yield dichotomous response. For convenience, we consider the setting of a clinical trial in which patients receive one of two medical treatments. The results obtained herein apply to other types of trials as well.

Within the clinical trial, application of a treatment results in either recovery (success) or no recovery (failure). The goal of the trial is to maximize the total number of successes in N patients, where N may be unknown. This is equivalent to maximizing the total utility of the experiment when a success has utility 1 and a failure has utility 0.

For each of the N patients in the trial the experimenter must choose one of two treatments to be applied. Treatment 1 has success probability θ_1 and treatment 2 has unknown success rate θ_2. We assume that treatment 1 is a well-known standard already in use. We may wish to model θ_1 as being unknown, as the "well-known" standard may produce surprises in the current trial. However, for simplicity we assume throughout that θ_1 is known.

Opinion about treatment 2 available separate from the trial is expressed by a prior distribution on θ_2. The outcomes are assumed to be Bernoulli random variables. When treatment 1 is applied the random variables are indexed by θ_1 and are independent. When treatment 2 is applied the random variables are indexed by θ_2, are exchangeable, and are independent of the variables indexed by θ_1. The patients are viewed as being exchangable for the purpose of the trial.

The outcome for a patient is frequently assumed to be known before the next patient is treated. Petkau (1978), Upton and Lee (1976), and Berry (1972, 1978),

Received May 1984; revised July 1985.
AMS 1980 *subject classifications.* Primary 62C10.
Key words and phrases. Multistage trial, regular discounting.

among others, have considered sequential allocation of Bernoulli processes. Of particular interest is Berry and Fristedt (1979).

While mathematically appealing, these procedures may be too cumbersome to be used in practice. They require that the patients respond almost immediately to treatment and, often, that the experimenter make a continuing series of calculations and decisions as the trial progresses. Of course, the availability of microcomputers lessens the second obstacle.

A more realistic assumption is that the data are collected at intervals throughout the trial. Calculations can be made at these times and the future course of the trial can be altered. Such a trial consists of several stages. While it is possible that only partial information is available from one stage to the next, the setting considered here is that in which all of the patients from the previous stage respond before the experimenter decides on an allocation scheme for the next stage. If the number of patients in each stage is 1 then we are in the classical sequential setting.

The problem of maximizing the total number of successes when an experiment takes place in two or more stages has received little attention. Cornfield, Halperin, and Greenhouse (1969), Colton (1963), and Donner (1979) discussed this problem when the outcomes are normal random variables. Canner (1970) considered Bernoulli processes in a multistage setting with both treatments unknown, as did Pearson (1980). Pearson (1980) also considered the case in which one treatment is known. However, he required N to be fixed.

We now discuss the notion of a strategy. The number of stages in the trial is chosen before the trial begins and is not part of a strategy. A strategy, often denoted by τ, specifies $K(\tau)$, the length of the first stage. Denote by $K_i(\tau)$ the number of planned first-stage observations on treatment i, $i = 1$ or 2. Often we abbreviate $K_i(\tau)$ to Ki and $K(\tau)$ to K. Then $K = K1 + K2$. As well, τ specifies $t_1, t_2, \ldots, t_K$, where t_i is 1 or 2 according as treatment 1 or 2 is assigned to the ith patient. Then it specifies the length and treatment allocation scheme for the second stage, in the same manner as for the first stage, for each possible outcome of the first stage, and so on.

The length and treatment allocation scheme for a particular stage are functions of the success rate θ_1, the number of stages remaining, and the current distributions on θ_2 and N. The results from previous stages are reflected in the current distribution of θ_2. The quantities N and θ_2 are assumed to be independent.

The utility of a strategy is the average utility of the histories possible when following that strategy; the average is with respect to θ_1 and the initial distributions of θ_2 and N. An optimal strategy is one that yields maximal expected utility.

An example of a two-stage strategy can be described as follows. Take exactly one first-stage observation on treatment 2. Thus $K = 1$ and $t_1 = 2$. If a success is obtained use only treatment 2 for the second stage. That is, set $t_2 = t_3 = \cdots = 2$. If a failure is observed use only treatment 1 in the second stage ($t_2 = t_3 \cdots = 1$).

For each ordered pair $(K1, K2)$ there are various possible allocation vectors $(t_1, \ldots, t_K)$. Also, there are many continuations for each possible outcome of the first stage. Thus, many strategies share a common $(K1, K2)$.

For fixed θ_1, Π, and Q, let $U((K1, K2), \theta_1, \Pi, Q)$ denote the maximal expected utility over all strategies with given $(K1, K2)$. Let

$$U^*(\theta_1, \Pi, Q) = \sup_{(K1, K2)} U((K1, K2), \theta_1, \Pi, Q).$$

Optimal strategies have utility $U^*(\theta_1, \Pi, Q)$.

It may be that more than one ordered pair $(K1, K2)$ is optimal. Let $K1^*$ denote the smallest value of $K1$ that is optimal. That is, $K1^*$ is the least value of $K1$ among all $(K1, K2)$ that satisfy

$$U((K1, K2), \theta_1, \Pi, Q) = U^*(\theta_1, \Pi, Q). \tag{1}$$

Similarly, denote by K^* the smallest value of $K1^* + K2$ among all $(K1^*, K2)$ that satisfy (1). Of course, K^* may sometimes be zero, in which case the first stage contains no observations. Let $K2^* = K^* - K1^*$.

Throughout the remainder of the discussion we denote $E(\theta_2|\Pi)$ by μ and the number of successes form the K2 first-stage observations on treatment 2 by S. In Section 2 we consider two-stage trials in which the total number of patients, N, is known. In Section 3 we discuss distributions on N and develop their relation to the discounting of future observations. Special attention is given to a class of discount sequences that are called regular. The main results, which extend the work of Berry and Fristedt (1979), are in Section 3, where we consider two-stage and higher-stage trials in which N is unknown.

2. Known trial length with two stages. In this section we consider the case in which N, the total number of patients to be treated, is known and the trial has two stages. We modify our notation by writing $U((K1, K2), \theta_1, \Pi, N)$ in place of $U((K1, K2), \theta_1, \Pi, Q)$.

Our goal is to maximize the total number of successes on the N patients. Given the results of the first stage, the optimal conditional strategy for the second stage clearly is to use the treatment with larger posterior mean on all remaining patients. Thus, our problem is to choose K and to assign treatments to each of the first K patients.

When N is known we may, without loss, take all observations on treatment 1 before any observations on treatment 2, i.e., $(t_1, \ldots, t_K)$ is of the form $(1, 1, \ldots, 1, 2, 2, \ldots, 2)$. Thus, the ordered pair (K1, K2) uniquely identifies a strategy. It is understood that an optimal continuation will be used for the second stage.

The advantage of a two-stage procedure over a one-stage procedure is that with a two-stage procedure we can learn about the treatments during the first stage and apply our knowledge during the second. Since θ_1 is known it seems that no first-stage observations on treatment 1 should be necessary. Theorem 1, which is similar to a result by Pearson (1980), confirms this.

THEOREM 1. *Let N be known. For all θ_1 and for any prior, Π, on θ_2, $K1^* = 0$.*

COMMENT. $K1 = 0$ may not be uniquely optimal. For example, if $P(\theta_2 < \theta_1) = 1$ then treatment 2 should never be used, and any value of $K1$ such that $0 \leq K1 \leq N$ is optimal.

PROOF. Note that

$$P(S=j) = EP(S=j|\theta_2) = \binom{K2}{j}\int \theta^j(1-\theta)^{K2-j}\,d\Pi(\theta)$$

and that the support of S is $0,1,2,\ldots,K2$. Let $K1 + K2 \le N$. Then

$$U((K1, K2), \theta_1, \Pi, N) = K1\cdot\theta_1 + K2\cdot\mu + (N - K1 - K2) \times\left(\sum_{j=0}^{K2}\max\{\theta_1, E(\theta_2|S=j)\}P(S=j)\right),$$

$$\begin{aligned}\sum_{j=0}^{K2}\max\{\theta_1, E(\theta_2|S=j)\}P(S=j) &= U((0, K2), \theta_1, \Pi, N) + K1\cdot\theta_1 \\ &\quad - K1\left(\sum_{j=0}^{K2}\max\{\theta_1, E(\theta_2|S=j)\}P(S=j)\right) \\ &\le U((0, K2), \theta_1, \Pi, N) + K1\cdot\theta_1 \\ &\quad - K1\left(\sum_{j=0}^{K2}\theta_1 P(S=j)\right) \\ &= U((0, K2), \theta_1, \Pi, N).\ \square\end{aligned}$$

Consider the case in which Π is a uniform distribution on the interval $(0,1)$. Here $P(S=j|\Pi) = 1/(K+1)$ for $j = 0,1,\ldots,K$. Also,

$$E(\theta_2|\Pi) = \frac{1}{2}; \qquad E(\theta_2|S=j,\Pi) = \frac{j+1}{K+2}.$$

Thus

$$U(K, \theta_1, \Pi, N) = \frac{K}{2} + \frac{N-K}{K+1}\left(\sum_{j=0}^{K}\max\left\{\theta_1, \frac{j+1}{K+2}\right\}\right). \tag{2}$$

Pearson (1980) derived the approximation

$$K^* \doteq \left[(N+1)\left(\theta_1^{-1} - 1\right)\right]^{1/2} - 1$$

as follows:

Let $[|x|]$ denote the greatest integer less than or equal to x.

$$\begin{aligned}U(K, \theta_1, \Pi, N) &\doteq \frac{K}{2} + \frac{N-K}{K+1}\left(\sum_{j=0}^{[|\theta_1(K+2)-1|]}\theta_1 + \sum_{j=[|\theta_1(K+2)|]}^{K}\frac{j+1}{K+2}\right) \\ &= \frac{K}{2} + \frac{N-K}{K+1}\left(\theta_1^2\frac{K+2}{2} - \frac{\theta_1}{2} + \frac{K+1}{2}\right) \\ &= \tfrac{1}{2}\left\{N - \theta_1(N-K)(K+1)^{-1}\right. \\ &\qquad \left. + \theta_1^2(K+2)(N-K)(K+1)^{-1}\right\} \\ &= \tfrac{1}{2}f(K, \theta_1, N), \quad \text{say.}\end{aligned}$$

Setting $d/dK\, f(K, \theta_1, N) = 0$ gives the desired result.

3. Distributions on N and discounted future. In section 4 we consider trials in which N is unknown. The purpose of this section is to explore the relationship between distributions on N and the commonly considered approach of discounting future observations.

It is often for computational convenience that N is assumed to be known. In applications it is unlikely that N is known precisely: patients may drop out of the trial, new treatments may be discovered, etc. We model the case in which N is unknown by assuming that it is a random variable with known distribution Q on the positive integers.

Let τ_m denote the observation on patient m when following strategy τ. When N is known, the utility, $u(\tau|N)$, of strategy τ is the expected sum of the observations:

$$u(\tau|N) = E\sum_{m=1}^{N} \tau_m. \tag{3}$$

Suppose that N has distribution Q, under which $P(N = i) = p_i$. Let $\alpha_n = \sum_{i=n}^{\infty} p_i$, the probability that N is greater than or equal to n. We assume that all α_i are independent of all τ_i. That is, the probability that the trial ends at a given time is independent of the results obtained on the patients. A more realistic, although more difficult to model, assumption is that the results obtained during the trial affect the decision of when to terminate the trial by, for example, affecting the rate of development of other treatments.

THEOREM 2. *The sequence $A = (\alpha_1, \alpha_2, \ldots)$, which is nonincreasing, is a discount sequence for the trial.*

PROOF. For any strategy τ, with N distributed as Q,

$$\begin{aligned} u(\tau|Q) &= \sum_n u(\tau|N = n, Q)P(N = n|Q) \\ &= \sum_n P(N = n|Q)\left[\sum_{m=1}^{n} E(\tau_m|N = n, Q)\right]. \end{aligned}$$

The strategy τ depends on Q. Thus $E(\tau_m|N = n, Q)$ depends on Q but not on $N = n$. Hence

$$\begin{aligned} &u(\tau|Q) + p_1E(\tau_1|Q) + \cdots + p_nE(\tau_1 + \cdots + \tau_n|Q) + \cdots \\ &\quad = E(\tau_1|Q)(p_1 + p_2 + \cdots) + E(\tau_2|Q)(p_2 + p_3 + \cdots) + \cdots \\ &\qquad + E(\tau_m|Q)(p_m + p_{m+1} + \cdots) + \cdots \\ &\quad = E(\tau_1|Q)\alpha_1 + \cdots + E(\tau_m|Q)\alpha_m + \cdots. \end{aligned}$$

Thus

$$u(\tau|Q) = E\sum_{m=1}^{\infty} \alpha_m\tau_m = \sum_{m=1}^{\infty} \alpha_m E\tau_m, \tag{4}$$

suppressing the dependence of τ on Q. □

The discount sequence A assigns to each potential observation in the trial weight equal to the chance that it is indeed observed.

EXAMPLE 1. Suppose Q is a one-point distribution at N_0, i.e., N is known to be N_0. Then $\alpha_1 = \cdots = \alpha_{N_0} = 1$ and $\alpha_{N_0+1} = \cdots = 0$. So

$$A = (1,1,\ldots,1,0,0,\ldots) \tag{5}$$

and for any strategy τ,

$$u(\tau) = E \sum_{m=1}^{\infty} \alpha_m \tau_m = E \sum_{m=1}^{N} \tau_m.$$

All discount sequences generated by distributions on N can be expressed as mixtures of discount sequences of form (5).

Only the first N patients in a sequence of possible patients are treated in the trial. Note that observations are not actually discounted as they are observed. N, however, is unknown (before the trial begins) and, hence, in finding an optimal strategy we discount future observations by fixed factors $\alpha_1, \alpha_2, \ldots$.

Although we learn about N as the trial progresses, this learning does not change the nature of the problem. The conditional distribution of N at any given time is proportional to Q, the original distribution. Given that $N > m$, our goal at time m is to maximize

$$E \sum_{n=m+1}^{\infty} \beta_n \tau_n,$$

where

$$\beta_n = P(N \geq n | N > m, Q) = \frac{P(N \geq n|Q)}{P(N > m|Q)} = \frac{\alpha_n}{\alpha_{m+1}}.$$

But this is equivalent to maximizing

$$E \sum_{n=m+1}^{\infty} \alpha_n \tau_n.$$

The problem of finding optimal strategies can be greatly simplified for a large class of distributions on N. This is the class of distributions that are regular. We define a distribution Q to be regular if it yields a regular discount sequence, as defined by Berry and Fristedt (1979).

DEFINITION. A discount sequence $A = (\alpha_1, \alpha_2, \ldots)$ is regular if, for each m,

$$\gamma_m \gamma_{m+2} \leq \gamma_{m+1}^2,$$

where $\gamma_n = \sum_{i=n}^{\infty} \alpha_i$.

Note that all geometric discount sequences are regular, but just barely: $\gamma_{m+1}^2 = \gamma_m \gamma_{m+2}$ for all m.

4. Trials with unknown N. In this section we consider the case in which N is unknown. We consider the special case of a geometric distribution on N separately before developing several theorems that lead to the main result on regular distributions. This main result (Theorem 7) extends the sequential work of Berry and Fristedt (1979).

If $\theta_1 > E\theta_2$ we are faced with two conflicting goals. We wish to maximize the probability of success on the current patient. This requires that we use treatment 1. On the other hand, we want to explore the possibility that $\theta_2 > \theta_1$. To gather information about θ_2 we can assign treatment 2 to the current patient.

Although it may benefit future patients, use of treatment 2 when $\theta_1 > E\theta_2$ may be inconsistent with medical ethics. We shall avoid the issue of medical ethics and proceed with the understanding that some combination of "the good of the current patient" and "the good of the whole" is appropriate. Such a combination is consistent with, and can be reflected in, the discount sequence A discussed in the previous section.

Each distribution Q yields a unique discount sequence $A(Q)$. If $A(Q)$ is of the form $(\alpha_1, \ldots, \alpha_n, 0, 0, \ldots)$, then only a finite number of strategies are possible; hence there exists an optimal strategy. Theorem 3 shows that there exists an optimal strategy for any discount sequence.

THEOREM 3. *For any distributions* Π *and* Q, *for any* θ_1, *and for any* (*fixed*) *number of stages there exists an optimal strategy.*

NOTE. This result is very similar to Lemma 1.1 in Berry and Fristedt (1979). We adapt their proof.

PROOF. Let D denote the number of stages in the trial. As always, D is fixed before the trial begins.

Suppose that Q yields discount sequence $(\alpha_1, \alpha_2, \ldots)$. Further, suppose that, for each n, τ^n is an optimal D-stage strategy for θ_1, Π, and the discount sequence $(\alpha_1, \ldots, \alpha_n, 0, \ldots)$. There is a strategy τ that through any term m agrees with at least one τ^n. We construct τ recursively as follows: At term m define τ so that it agrees with infinitely many of the τ^n through term m. The utility of τ is

$$u(\tau|Q) = \sum_{i=1}^{\infty} \alpha_i E\tau_i = \lim_{m\to\infty} \sum_{i=1}^{m} \alpha_i E\tau_i$$

$$= \lim_{m\to\infty} \sum_{i=1}^{m} \alpha_i E\tau_i^m.$$

Let $\hat{\tau}$ be any other strategy. Then

$$u(\tau|Q) = \sum_{i=1}^{\infty} \alpha_i E\hat{\tau}_i = \lim_{m\to\infty} \sum_{i=1}^{m} \alpha_i E\hat{\tau}_i$$

$$\leq \lim_{m\to\infty} \sum_{i=1}^{m} \alpha_i E\tau_i^m = u(\tau|Q)$$

since

$$\sum_{i=1}^{m} \alpha_i \hat{\tau}_i \leq \sum_{i=1}^{m} \alpha_i \tau_i^m$$

by the definition of τ^m. Thus τ is an optimal strategy for θ_1, Π, and Q. □

In Section 2 we considered N to be known and concluded that $K1^* = 0$. This is not the case when N is unknown, as the following example shows.

EXAMPLE 2. Suppose $Q(1) = 0.9$, $Q(10) = 0.1$, $\theta_1 = 0.6$, and $\theta_2 \sim$ uniform(0, 1). Further suppose that when $K1 > 0$ we take all $K1$ observations on treatment 1 before any of the $K2$ observations on treatment 2; i.e., $(t_1, \ldots, t_K)$ is of the form $(1, \ldots, 1, 2, \ldots, 2)$. This will be shown in Theorem 5 to characterize optimal strategies. Then $U(K1, 0) = (0.9)[(0.6)(1)] + (0.1)[(0.6)(10)] = 1.14$ for $K1 = 0, 1, \ldots, 10$, and $U(1, 1) = 1.156$. Similar calculations show that $U(K1, K2) < U(1, 1)$ for all $(K1, K2) \neq (1, 1)$. Thus, the ordered pair $(1, 1)$ is optimal and $K1^* > 0$.

Nevertheless, we can often set $K1$ equal to zero.

THEOREM 4. *For any prior distributions* Π *on* θ_2 *and* Q *on* N, *if* $\theta_1 \leq \mu$ *then* $K1^* = 0$.

PROOF. Suppose $K1(\tau) > 0$ for some strategy τ. Consider an alternative strategy τ' derived from τ as follows: Set $K(\tau') = K(\tau)$, $K1(\tau') = 0$, and $K2(\tau') = K(\tau')$. Then under strategy τ', $t_i = 2$ for $i \leq K(\tau')$.

The expected utility from the first stage when following τ' is no less than the expected utility from the first stage when following τ, since $\theta_1 \leq \mu$. The same relationship holds for the second stage, as may be seen by noting that

$$E_{X_{j+1}}\left[\max\left\{\theta_1, E\left(\theta_2 | \mathbf{X}_{j+1}\right)\right\} | \mathbf{X}_j\right] \geq \max\left\{\theta_1, E\left(\theta_2 | \mathbf{X}_j\right)\right\},$$

where $\mathbf{X}_j$ is the random vector of responses for the first j observations on treatment 2. Thus $u(\tau) \leq u(\tau')$ and the result follows. □

Clearly, when $K1 = 0$ the order of treatment assignments in the first stage is not an issue. When N is unknown and both $K1 > 0$ and $K2 > 0$, however, the order of the first-stage observations can be important.

When $\theta_1 \leq \mu$, $K1^* = 0$ by Theorem 4. When $\theta_1 > \mu$ we want all first-stage observations on treatment 1 to precede any first-stage observations on treatment 2. Theorem 5, which was used in Example 1, makes this idea precise. This theorem states that for any prior distributions Π on θ_2 and Q on N, if $\theta_1 > \mu$ then it is optimal to take any first-stage observations on treatment 1 before any and all first-stage observations on treatment 2.

THEOREM 5. *For any prior distributions* Π *on* θ_2 *and* Q *on* N, *if* $\theta_1 > \mu$ *then there is an optimal strategy of the form:* $t_1 = \cdots = t_{K1} = 1$, $t_{K1+1} = \cdots = t_K = 2$.

PROOF. Suppose strategy τ is such that $t_i = 2$ and $t_{i+j} = 1$ for some i, j with $i + j \leq K(\tau)$. Modify τ to τ' only as follows: Set $t_i = 1$ and $t_{i+j} = 2$ under τ'. Then $u(\tau') - u(\tau) = (\alpha_i - \alpha_{i+j})(\theta_1 - \mu) \geq 0$. Proceed to exchange all such deviant pairs. □

Suppose that N has a geometric distribution with mean $1/p$ (p is known). As noted earlier, geometric distributions are regular. We will see that the case $N \sim G(p)$ behaves much like that in which N is known and equal to $[1/p]$, the greatest integer in $1/p$, provided $1/p$ is large ($1/p > 20$, say).

THEOREM 6. *Let* $N \sim G(p)$ *with* $0 < p < 1$. *Then for any prior* Π *on* θ_2 *and for any* θ_1, $K1^*(\theta_1, \Pi, G) = 0$.

PROOF. If $\theta_1 \leq \mu$ then $K1^* = 0$ by Theorem 4. Assume that $\theta_1 > \mu$. Let $q = 1 - p$. We now derive a closed form expression for the utility function. We have

$$U((K1, K_2), \theta_1, \Pi, G) = U(K1, K2) = \sum_{i=1}^{\infty} q^{i-1} p U(K1, K2|N = i).$$

By Theorem 5 we can, without loss, consider only those strategies for which all $K1$ observations on treatment 1 precede any of the $K2$ observations on treatment 2. Thus,

$$U(K1, K_2) = \sum_{i=1}^{K1} q^{i-1} p \theta_1 i + \sum_{i=K1+1}^{K1+K2} q^{i-1} p \left[\theta_1 K1 + \mu(i - K1)\right]$$

$$+ \sum_{i=K1+K2+1}^{\infty} q^{i-1} p \left[\theta_1 K1 + \mu K2 + (i - K1 - K2) H(K2, \theta_1, \Pi)\right],$$

where $H = H(K2, \theta_1, \Pi) = \Sigma_{j=0}^{K2} \max\{\theta_1, E(\theta_2|S = j)\} P(S = j)$. This reduces to

$$U((K1, K2), \theta_1, \Pi, G) = \theta_1(1 + qp^{-1}) + q^{K1}\mu p^{-1}(1 - \theta_1\mu^{-1}) \tag{6}$$

$$+ q^{K1+K2}\mu p^{-1}(H\mu^{-1} - 1).$$

Thus, $U(K1, K2)$ is of the form

$$U(K1, K2) = q^{K1} f(K2) + C.$$

So

$$\max_{(K1, K2)} U(K1, K2) = \max_{K1} \max_{K2} \{q^{K1} f(K2) + C\}.$$

If $\max_{K2} f(K2)$ is nonnegative then $U(K1, K2)$ is maximized by setting $K1 = 0$, which maximizes q^{K1}. If $\max_{K2} f(K2)$ is negative then $U(K1, K2)$ is

maximized by setting $K1 = \infty$, which minimizes q^{K1}. But this yields utility C, as does $(K1, K2) = (0,0)$. Thus $K1^*$ is always zero. □

Since $K1^* = 0$, we can set $K2 = K$ and simplify (6) to

$$U(K, \theta_1, \Pi, G) = \mu p^{-1} + q^K \mu p^{-1}(H\mu^{-1} - 1). \tag{7}$$

Expanding $q^K = (1-p)^K$ and deleting terms involving $p^2, p^3, \ldots, p^K$ yields

$$\begin{aligned} U(K, \theta_1, \Pi, G) &\doteq \mu p^{-1} + \mu(p^{-1} - K)(H\mu^{-1} - 1) \\ &= \mu K + (p^{-1} - K)H. \end{aligned} \tag{8}$$

In particular, if $\theta_2 \sim$ uniform$(0,1)$ then

$$U(K, \theta_1, G) \doteq \tfrac{1}{2}K + (p^{-1} - K)(K+1)^{-1}[M(K, \theta_1)],$$

where $M(K, \theta_1) = \sum_{j=0}^{K} \max\{\theta_1, E(\theta_2 | S = j)\}$.

When compared with the results from Section 2, we see that this is the same as (2) with N replaced by p^{-1}, the mean of the geometric distribution. Thus

$$K^*(G(p)) \doteq \left[(p^{-1} + 1)(\theta_1^{-1} - 1)\right]^{1/2} - 1.$$

We now present a generalization of Theorem 6: If Q yields a discount sequence A that is regular then $K1^*(Q) = 0$. The presence of this condition greatly simplifies the design problem.

To prove this result we need several lemmas. This first lemma follows immediately from the definition of regularity and is presented without proof.

LEMMA 1. *If the discount sequence A is regular then $\alpha_i \gamma_{m+1} \leq \alpha_m \gamma_{i+1}$ for $i \leq m$.*

LEMMA 2. *Suppose that Q is regular and that $\mu < \theta_1$. Then*

$$g_1(K2) \equiv [H(K2) - \mu]\gamma_{K1+K2+1} + [\mu - \theta_1]\gamma_{K1+1} > 0 \tag{9}$$

implies

$$g(K2) \equiv [H(K2) - \mu]\alpha_{K1+K2} + [\mu - \theta_1]\alpha_{K1} > 0. \tag{10}$$

PROOF. Inequality (9) implies $\gamma_{K1+K2+1} > 0$, which implies

$$\alpha_{K1+K2} > 0, \qquad \gamma_{K1+1} > 0$$

and, hence, $\alpha_{K1} > 0$. Thus

$$[H(K2) - \mu] + [\mu - \theta_1]\frac{\gamma_{K1+1}}{\gamma_{K1+K2+1}} > 0,$$

which implies

$$[H(K2) - \mu]\alpha_{K1+K2} + [\mu - \theta_1]\alpha_{K1} > 0,$$

since, by Lemma 1,

$$\frac{\gamma_{K1+1}\alpha_{K1+K2}}{\gamma_{K1+K2+1}\alpha_{K1}} \geq 1. \,\square$$

The next lemma makes use of the following fact. Suppose $\mu < \theta_1$. Let Q_{K1} be the conditional distribution of N given that $N > K1$. Let $K2' = K2'(Q_{K1}, K1)$ maximize $U(K1, K2)$ over $K2$ with $K1$ fixed. That is, $K2'$ is the smallest optimal value of $K2$, for distribution Q, given $K1$. Then, in view of Theorem 5, $K2'$ is also the value of $K2$ that maximizes $U(0, K2), \theta_1, \Pi, Q_{K1})$. This holds because the $K1$ observations on treatment 1 do not change Π, they only modify Q to Q_{K1}. Thus

$$\max_{K2} U((0, K2), \theta_1, \Pi, Q_{K1}) = U((0, K2'), \theta_1, \Pi, Q_{K1}). \tag{11}$$

Lemma 3 shows that when Q is regular the maximal utility, $U^*(K1)$, for $K1$ fixed is no greater than the utility of using the ordered pair $(K1 - 1, K2'(Q_{K1}, K1))$. This fact is crucial to the proof of Theorem 7 and will be used in an obvious way.

LEMMA 3. *Let Q be regular, let $\theta_1 > \mu$, and let $K1 > 0$ be fixed. Then*

$$U^*(K1) = U((K1, K2'(Q_{K1}, K_1)), \theta_1, \Pi, Q) \leq U((K1 - 1, K2'(Q_{K1}, K1)), \theta_1, \Pi, Q).$$

PROOF.

$$U^*(K1) = U(K1, K2') = \theta_1 \sum_{i=1}^{K1} \alpha_i + \mu \sum_{i=K1+1}^{K1+K2'} \alpha_i + H(K2') \sum_{i=K1+K2'+1}^{\infty} \alpha_i.$$

Thus

$$U(K1 - 1, K2'(Q_{K1}, K1)) - U(K1, K2'(Q_{K1}, K1)) = [\mu - \theta_1]\alpha_{K1} + [H(K2') - \mu]\alpha_{K1+K2'} = g(K2'),$$

where $g(K2)$ is defined by (10).

Define

$$U^\dagger(K2) = \sum_{i=K1+1}^{\infty} U(K2|N = i)P(N = i|Q) = \mu \sum_{i=K1+1}^{K1+K2} \alpha_i + H(K2) \sum_{i=K1+K2+1}^{\infty} \alpha_i.$$

For convenience, consider

$$U^\dagger(K2) - U^\dagger(0) = [\mu - \theta_1] \sum_{i=K1+1}^{K1+K2} \alpha_i + [H(K2) - \theta_1] \sum_{i=K1+K2+1}^{\infty} \alpha_i$$

$$= [H(K2) - \mu]' \sum_{i=K1+K2+1}^{\infty} \alpha_i + [\mu - \theta_1] \sum_{i=K1+1}^{\infty} \alpha_i$$

$$= g_1(K2) \quad \text{as defined by (9).}$$

Since, in view of (11), $K2'$ maximizes $U^\dagger(K2)$ and since $g_1(0) = 0$,

$$K2'(Q_{K1}, K1) > 0 \quad \text{implies} \quad g_1(K2'(Q_{K1}, K1)) > 0.$$

Now, if $K2' > 0$ then $g_1(K2') > 0$, which implies $g(K2') > 0$ by Lemma 2. If $K2' = 0$ then $g(K2') = g(0) = 0$, since $H(0) = \theta_1$. Thus $g(K2') \geq 0$ for all $K2'$. □

We are now ready to prove the main two-stage result. Theorem 7 is similar to Theorem 2.1 of Berry and Fristedt (1979), who consider sequential allocation of treatments.

THEOREM 7. *If Q is regular then $K1^*(Q) = 0$ for all θ_1 and Π.*

PROOF. If $\mu \geq \theta_1$ then $K1^* = 0$ by Theorem 4. Suppose $\mu < \theta_1$.

$$U(K1, K2) = \theta_1 \sum_{i=1}^{K1} \alpha_i + \mu \sum_{i=K1+1}^{K1+K2} \alpha_i + H(K2) \sum_{i=K1+K2+1}^{\infty} \alpha_i.$$

$K2'(Q_{K1}, K1)$ maximizes

$$\mu \sum_{i=K1+1}^{K1+K2} \alpha_i + H(K2) \sum_{i=K1+K2+1}^{\infty} \alpha_i$$

in view of (11). That is, $U^*(K1) = U(K1, K2'(Q_{K1}, K1))$.

Suppose $K1 > 0$.

$$U(K1, K2'(Q_{K1}, K_1)) \leq U(K1 - 1, K2'(Q_{K1}, K1)) \tag{12}$$

by Lemma 3. But

$$\begin{aligned} U(K1 - 1, K2'(Q_{K1}, K1)) &\leq U(K1 - 1, K2'(Q_{K1-1}, K1 - 1)) \\ &= U^*(K1 - 1), \end{aligned} \tag{13}$$

where the equality holds by (11).

Combining (12) and (13), we see that

$$U^*(K1) \leq U^*(K1 - 1) \quad \text{for all } K1 > 0.$$

Since $K1$ must be at least zero, $K1^* = 0$. □

There are settings in which $K1 = 0$ is not uniquely optimal. For some Q, θ_1, and Π it is optimal to never use treatment 2. In such a case any value of $K1$ is optimal.

The converse of the Berry and Fristedt result is true, while the converse of Theorem 7 is false. The following example shows that regularity is not a necessary condition for $K1^*$ to be zero.

EXAMPLE 3. Suppose Q is of the form $(0, a, 0, b, 0, 0, 0, \dots)$ where $a + b = 1$ and $0 < b < \frac{1}{2}$. Since $A = (1, 1, b, b, 0, 0, 0, \dots)$, no distribution Q in this class is regular, yet $K1^*(Q) = 0$ for all θ_1 and for all Π. This may be shown by process of elimination, noting that $K1 + K2 \leq 4$ since $N \leq 4$ with probability 1.

We have now established that for two-stage trials with regular discount sequences it is optimal to take no observations from the known treatment in the first stage ($K1^* = 0$). This result holds for any (fixed) number of stages.

THEOREM 8. *Let a trial consist of D stages, where $D \geq 2$ is fixed before the trial begins. If the discount sequence generated by Q is regular then $K1^*(Q) = 0$ for all θ_1 and Π.*

PROOF. We establish this by considering two cases. For each case we use an induction argument on the number of stages. We shall implicitly use the fact that a regular discount sequence modified by the fact that $N > c$ (a constant) is in turn regular. We will use a subscript to keep track of the number of remaining stages. For example, $K1_2^*$ is the smallest optimal number of first-stage observations on treatment 1 when there are two stages remaining.

CASE 1. $\theta_1 \leq \mu$. By Theorem 4, $K1_2^* = 0$. Assume that $K1_D^* = 0$. We now show that $K1_{D+1}^* = 0$. This requires the examination of three subcases. Note that

$$U_{D+1}(K1, K2) = \sum_i U_{D+1}(K1, K2|N = i)P(N = i|Q).$$

Consider $U_{D+1}(K1, K2|N = i)$ and let L denote the number of observations on treatment 1 in the first i observations.

SUBCASE 1.1. $i < K2$. Here $0 \leq L \leq K1$ and $i - L < K2$. Thus

$$\begin{aligned} U_{D+1}(K1, K2|N = i) &= L\theta_1 + (i - L)\mu \leq i\mu \\ &= U_{D+1}(0, K2|N = i). \end{aligned}$$

SUBCASE 1.2. $K2 \leq i < K1 + K2$. In this subcase

$$\begin{aligned} U_{D+1}(K1, K2|N = i) &= L\theta_1 + (i - L)\mu \\ &\leq (i - K2)\theta_1 + K2(\mu) \\ &= U_{D+1}(0, K2|N = i). \end{aligned}$$

SUBCASE 1.3. $K1 + K2 \leq i$. Here

$$U_{D+1}(K1, K2|N = i) = K1(\theta_1) + K2(\mu) + E_{\Pi'}[U_D^*(\Pi', Q_{K1+K2})|N = i],$$

where Π' is the posterior distribution of θ_2 after the first $K1 + K2$ observations and Q_{K1+K2} is the conditional distribution of N given that $N > K1 + K2$. When $K1 = 0$ we have

$$U_{D+1}(0, K2|N = i) = K2(\mu)' + E_{\Pi'}[U_D^*(\Pi', Q_{K2})|N = i].$$

The induction hypothesis implies that

$$E_{\Pi'}[U_D^*(\Pi', Q_{K2})|N = i] \geq K1(\theta_1) + E_{\Pi'}[U_D^*(\Pi', Q_{K1+K2})|N = i].$$

That is, with D stages remaining, it is optimal to take no observations on

treatment 1 in the first of the remaining stages. Hence

$$U_{D+1}(K_1, K2|N = i) \leq U_{D+1}(0, K2|N = i).$$

For all three subcases we have shown that

$$U_{D+1}(K1, K2|N = i) \leq U_{D+1}(0, K2|N = i).$$

Thus $U_{D+1}(K1, K2) \leq U_{D+1}(0, K2)$, which implies that $K1^*_{D+1} = 0$. This completes the induction argument for Case 1.

CASE 2. $\theta_1 > \mu$. The proof of Theorem 5 generalizes easily. Thus, it is clear in this case that, for any number of stages, it is optimal to take all first-stage observations on treatment 1 before any first-stage observations on treatment 2. We adopt this convention.

By Theorem 7, $K1^*_2 = 0$. Assume that $K1^*_D = 0$. Let τ denote the best strategy among those that dictate that $K1$ first-stage observations be taken on treatment 1. Then $U^*_{D+1}(K1) = u_{D+1}(\tau)$. We shall show that $U^*_{D+1}(K1) \leq U^*_{D+1}(K1 - 1)$. To do this, we shall employ the same ideas used in the two-stage setting of Theorem 7. As in the two-stage setting, let $K2' = K2'(Q_{K1}, K1)$ be the smallest optimal value of $K2$, for distribution Q, given $K1$. Then $U^*_{D+1}(K1) = U_{D+1}(K1, K2')$.

Consider a strategy τ' that differs from τ in only one respect: Let τ' begin by taking $K1 - 1$ first-stage observations on treatment 1. From observation $K1$ on, let τ' assign to patient i that treatment that τ assigns to patient $i + 1$. Further, let τ' dictate the same stage lengths as those of τ for all stages subsequent to the first.

The induction hypothesis guarantees that all observations taken in stages other than the first and the last will be on treatment 2. The total number of observations on treatment 2 during stages one through D will, in general, be a random variable depending on $K2$; call it J.

Let $R(\mathbf{X}_J) = E[\max\{\theta_1, E(\theta_2|\mathbf{X}_J)\}|J]$. Generalizing the ideas in the proof of Lemma 3, we have

$$\begin{aligned} U_{D+1}(\tau') - U_{D+1}(\tau) &= [\mu - \theta_1]\alpha_{K1} + E[R(\mathbf{X}_J) - \mu]\alpha_{K1+J} \\ &\equiv h(K2'), \quad \text{say.} \end{aligned}$$

We wish to show that $h(K2') \geq 0$ for all $K2'$. Now, $K2'$ maximizes $U^\dagger_{D+1}(K2)$ where

$$U^\dagger_{D+1}(K2) = \sum_{i=K1+1}^{\infty} U_{D+1}(K2|N = i)P(N = i).$$

Thus

$$\begin{aligned} U^\dagger_{D+1}(K2) - U^\dagger_{D+1}(0) &= E\sum_{i=K1+1}^{K1+J} \alpha_i[\mu - \theta_1] + E\sum_{i=K1+J+1}^{\infty} \alpha_i[R(\mathbf{X}_J) - \theta_1] \\ &= E[R(\mathbf{X}_J) - \mu]\gamma_{K1+J+1} + [\mu - \theta_1]\gamma_{K1+1} \\ &\equiv h_1(K2'), \quad \text{say.} \end{aligned}$$

When necessary we have taken expectation with respect to the random variable J.

Now, $h_1(0) = 0$. Hence $K2' > 0$ implies $h_1(K2') > 0$. If $K2' = 0$, then $h(K2') = h_1(K2') = 0$. If $K2' > 0$, then $h_1(K2') > 0$, which implies $h(K2') > 0$. This follows from Lemma 2 with J in place of $K2$.

Thus $h(K2') \geq 0$ for all $K2'$, which implies $U_{D+1}(\tau') \geq U_{D+1}(\tau)$. Clearly $U^*_{D+1}(K1 - 1) \geq U(\tau')$. So we have

$$U^*_{D+1}(K1) = U_{D+1}(\tau) \leq U_{D+1}(\tau') \leq U^*_{D+1}(K1 - 1).$$

This holds for all $K1 > 0$. Thus $K1^*_{D+1} = 0$, which was to be shown. This completes the induction argument for Case 2.

We have now shown that for either possible case $K1^*_{D+1} = 0$. □

COROLLARY. *For any (fixed) number of stages, for any θ_1, and for any Π, if the discount sequence generated by Q is regular then it is optimal to use treatment* 1 *in the last stage only.*

Acknowledgments. I thank Professor D. A. Berry for his invaluable help and guidance during the preparation of my thesis, from which this paper is adapted. I also thank the referees for their useful suggestions.

REFERENCES

BERRY, D. A. (1972). A Bernoulli two-armed bandit. *Ann. Math. Statist.* **43** 871–897.

BERRY, D. A. (1978). Modified two-armed bandit strategies for certain clinical trials. *J. Amer. Statist. Assoc.* **73** 339–345.

BERRY, D. A. and FRISTEDT, B. (1979). Bernoulli one-armed bandits—arbitrary discount sequences. *Ann. Statist.* **7** 1086–1105.

CANNER, P. L. (1970). Selecting one of two treatments when the responses are dichotomous. *J. Amer. Statist. Assoc.* **65** 293–306.

COLTON, T. (1963). A model for selecting one of two medical treatments. *J. Amer. Statist. Assoc.* **58** 388–400.

CORNFIELD, J., HALPERIN, M., and GREENHOUSE, S. W. (1969). An adaptive procedure for sequential clinical trials. *J. Amer. Statist. Assoc.* **64** 759–770.

DONNER, A. (1977). The use of auxiliary information in the design of a clinical trial. *Biometrics* **33** 305–314.

PEARSON, L. M. (1980). Treatment allocation for clinical trials in stages. Ph.D. thesis, School of Statistics, University of Minnesota, unpublished.

PETKAU, A. J. (1978). Sequential medical trials for comparing an experimental with a standard treatment. *J. Amer. Statist. Assoc.* **73** 328–338.

UPTON, G. J. G. and LEE, R. D. (1976). The importance of the patient horizon in the sequential analysis of binomial clinical trials. *Biometrika* **63** 335–342.

DEPARTMENT OF STATISTICS
UNIVERSITY OF FLORIDA
GAINESVILLE, FLORIDA 32611