In *Computing Science and Statistics* **24** (1992), pp. 597-601.

Optimal Adaptive Equal Allocation Rules

Janis P. Hardwick[1] Quentin F. Stout[2]
Statistics Department EECS Department
University of Michigan, Ann Arbor, MI 48109

Abstract

Suppose we wish to decide which of two treatments is better, where the outcomes are Bernoulli random variables, the success probabilities of which, themselves, are modeled as independent beta random variables. Assume that the maximal population size for the experiment is fixed, but that the length of the study and the number and order of patients assigned to each treatment may be random. Our goal is to maximize the likelihood of making the correct decision by utilizing a curtailed equal allocation rule, but we wish to do so with a minimal average study length.

We show that this experimental design problem reduces to a problem of optimal adaptive allocation which can be solved exactly using dynamic programming. We compare the optimal allocation procedure to the commonly-used approach of curtailed alternating allocation and show that the optimal allocation procedure is noticeably superior. The evaluations of allocation procedures are all exact, calculated via backward induction. Since the optimal adaptive allocation procedure can be easily determined and evaluated on workstations, and stored on personal computers for ready access during experiments, it is a practical improvement over alternating allocation.

1 Introduction

We are interested in the simple and common decision problem of trying to pick the better of two Bernoulli populations. For concreteness the problem will be described in terms of a clinical trial, but it is equally applicable to other areas such as product testing.

Suppose two treatments are available to treat a certain disease and we need to design an experiment to decide which is better. A maximum of N exchangeable patients can be used in the trial, but the trial may be stopped earlier. (To simplify exposition we assume N is even.) We assume that the patients enter the trial sequentially, and that the outcomes of patients $1, \ldots, i$ are known before patient $i+1$ is assigned. Responses to treatment will be either *successes* or *failures*, where, for each patient, the probability of success on Treatment i is p_i, $i = 1, 2$. Our study design is set up to allow the incorporation of prior information on the success rates p_1 and p_2. This information is modeled in the form of a joint distribution function ξ on (p_1, p_2) and is taken to be the product of two independent beta random variables:

$$(1) \qquad p_1 \sim Be(a_0, b_0) \quad \textit{and} \quad p_2 \sim Be(c_0, d_0)$$

for $(p_1, p_2) \in \Omega = (0,1) \times (0,1)$.

As was mentioned, our main goal is to select the better treatment, but we have a secondary goal as well. We wish to make the decision based on as few patients as possible, and we take advantage of the sequential nature of the data to do this. A *design* is comprised of two parts - an allocation procedure, $\gamma(\xi)$, and a decision rule. An allocation procedure is a rule for deciding what to do at each stage of the trial. Let $\Gamma(\xi)$ represent the class of all allocation procedures with the following features: At any stage i, the procedure may indicate one of three options: assign patient i to Treatment 1, to Treatment 2 or stop the trial. At each stage, the decision of how to proceed may depend only on the prior distribution and the information available from preceding patients.

Regardless of what procedure is used, the form of the optimal decision rule remains the same. The rule states that we select the population with the higher observed mean at the end of the trial.

2 Optimal Decision Making

Unfortunately, our two goals of making good terminal decisions and keeping study size to a minimum are somewhat contradictory, so our concern is to study tradeoffs between these criteria. First, however, it must be noted that, even for fixed sample size experiments, there are no allocation procedures that make optimal decisions for all $(p_1, p_2) \in \Omega$. A useful method of examining optimality in a more restricted sense is to consider procedures that offer optimality along lines of constant difference: $|p_2 - p_1| = \Delta$.

Let $\mathcal{P}^\gamma[(p_1, p_2); N]$ represent the probability of making a correct decision after N observations using procedure γ, when (p_1, p_2) are the true treatment success

[1]Research supported in part by National Science Foundation under grants DMS-8914328 and DMS-9157715.

[2]Research supported in part by National Science Foundation/DARPA under grant CCR-9004727.

rates. Next, define

$$\mathcal{P}^{\gamma}_{\Delta} = \min_{|p_2 - p_1| \geq \Delta} \mathcal{P}^{\gamma}\left[(p_1, p_2); N\right];$$

then we say a procedure, γ^*, is Δ-*optimal* if

$$P^{\gamma^*}_{\Delta} = \max_{\gamma \in \Gamma} \mathcal{P}^{\gamma}_{\Delta}.$$

3 Curtailed Allocation Procedures

Let $\mathcal{C}$ be the class containing all procedures that are Δ-optimal for all $\Delta \in (0,1)$. A popular sub-class of $\mathcal{C}$ is the set of *fixed horizon equal allocation rules*, where the term 'horizon' refers to the number of patients actually observed in the study. This class of procedures, $\mathcal{C}_{EA}$, contains all allocation rules that assign $\frac{N}{2}$ patients to each treatment alternative. Since we are concerned both with making good decisions and keeping study length to a minimum, we work here with a class of procedures that retain the optimal properties of those in $\mathcal{C}_{EA}$, but which also may be stopped prior to N.

Before characterizing the procedures in this new class, we must first define what is meant by a *state* of the general process that represents our clinical trial. Suppose that we are at Stage M, $0 \leq M \leq N$, of the study and have observed

$i = \#$ patients succeeding on Treatment 1,

$j = \#$ patients failing on Treatment 1,

$k = \#$ patients succeeding on Treatment 2, and

$l = \#$ patients failing on Treatment 2.

Note that $M = i+j+k+l$. The vector (i,j,k,l) which is sufficient for (p_1, p_2), coupled with the prior density function ξ, is referred to as the state $(i,j,k,l;\xi)$. Thus, at any time M, the state provides all information incorporated in the *posterior density function,* $\xi(p_1, p_2 \mid i,j,k,l)$, which is simply the product of the individual posterior densities

$$(p_1 \mid i,j) \sim Be(a,b), \qquad (p_2 \mid k,l) \sim Be(c,d),$$

where

$$a = i + a_0 \qquad b = j + b_0 \qquad c = k + c_0 \qquad d = l + d_0.$$

Often, the prior density function will be understood, and only the stage of the trial and the sufficient statistics will represent the state.

If, for some $M \geq 1$, the trial is terminated at a state $(i,j,k,l;\xi)$, then our *decision rule* is to declare as better the treatment with the higher observed mean – selecting

$$\left.\begin{array}{c} 1 \\ 2 \\ tie \end{array}\right\} \; if \; \frac{i}{i+j} \left\{\begin{array}{c} > \\ < \\ = \end{array}\right\} \frac{k}{k+l} \quad ,$$

and where we randomize to select a winner if a tie is declared. Throughout, we will assume this decision rule is used whenever the trial is terminated. Note that the state also contains enough information to determine the study length (namely M), the number of failures ($j+l$), and the number of patients assigned to the inferior treatment (either $k+l$ or $i+j$, depending on whether Treatment 1 or 2 is declared the better).

Now, suppose we are using some allocation procedure γ, and that during an experiment a state α is reached. Further, suppose that γ will allocate future patients so that all terminal states that can be reached from α will have the same decision D. In such a case, one can go ahead and make decision D at state α without affecting the Δ-optimal status of γ. This simple concept is called *pruning* or *curtailing,* and we denote the class of all curtailed equal allocation procedures as $\mathcal{C}_{CEA}$.

Consider, for example, the equal allocation procedure referred to as *alternating allocation,* which is a commonly used oblivious allocation procedure which ignores all previous outcomes and simply alternates back and forth between treatment assignments. If alternating allocation is being used with $N = 10$, and after the 7th patient the state is (3,1,0,3), then no matter what happens on the 8th, 9th and 10th patients (assigned Treatments 2, 1, and 2, respectively), the decision will be that Treatment 1 is the better. Clearly, one may as well prune the decision tree (or curtail the experiment) and declare Treatment 1 the better.

4 Minimum Average Study Length

When considering our second criteria for procedure optimality, study length, we must again resort to a restricted notion of optimality. Since no procedure offers the minimum expected study length for all $(p_1, p_2) \in \Omega$, we turn to a Bayesian concept which we refer to as the *average study length* of a procedure γ : $\mathcal{L}^{\gamma}$. Let $L^{\gamma}(p_1, p_2)$ denote the expected number of patients in a trial when the procedure γ is used for the fixed parameter configuration (p_1, p_2). Then

$$\mathcal{L}^{\gamma} = \int_{\Omega} L^{\gamma}(p_1, p_2)\, d\xi(p_1, p_2),$$

and it is our goal to find procedures that simultaneously minimize $\mathcal{L}$ and are contained in $\mathcal{C}_{CEA}$. In other words, we seek procedures $\gamma \in \Gamma^*$, where

$$\Gamma^* = \{\gamma' : \mathcal{L}^{\gamma'} = \min_{\gamma \in \mathcal{C}_{CEA}} \mathcal{L}^{\gamma}\}.$$

Our goal is minimize the average study length among the curtailed equal allocation rules, but one potential difficulty is that there are a very large number of equal allocation procedures, totaling $2^{2^{\Theta(N)}}$. Fortunately the use

Comment: This algorithm determines the minimal average study length, $ML(0,0,0,0)$, among all curtailed equal allocation rules, for a given prior distribution.

For all states α with N patients, $ML(\alpha) := N$.

For $M := N-1$ down to 0 do
    For all states α with M patients
        If α *prunable* then $ML(\alpha) := M$
        else
            If *Treatment 1 permissible*
            then $ML_1 :=$ expected minimal study length allocating Treatment 1
            else $ML_1 := \infty$.

            If *Treatment 2 permissible*
            then $ML_2 :=$ expected minimal study length allocating Treatment 2
            else $ML_2 := \infty$.

            $ML(\alpha) := \min(ML_1, ML_2)$.

Figure 1: **Algorithm for Minimal Study Length among Equal Allocation Rules**

of sufficient statistics reduces the number of nonequivalent procedures to $2^{\Theta(N^4)}$ since there are only $\Theta(N^4)$ states for which allocation decisions are need.

The algorithm for determining the minimal average study length appears in Figure 1. For the given fixed prior distribution (1), it determines the minimal possible average study length among the class of curtailed equal allocation rules. It works in the following way: For each state α, the algorithm determines $ML(\alpha)$, the minimum average study length given that the study reaches α and then proceeds optimally to the end of the study. The value $ML(\alpha)$ is computed via dynamic programming, which works from the last stage towards the first. The variable M in Figure 1 denotes the number of patients treated so far, and the final value calculated, $ML(0,0,0,0)$, is the answer.

While this is similar to Bellman's classic dynamic programming algorithm, [1], to minimize expected failures, there is a critical difference. With equal allocation procedures, there are states for which allocating one or the other of the treatments is prohibited. To decide if *Treatment 1 permissible*, note that in the state (i,j,k,l), an equal allocation rule can assign another patient to Treatment 1 if and only if the number of patients that have previously been assigned to Treatment 1 is less than $N/2$; i.e., if and only if

$$i + j < N/2 \ .$$

Similarly, the condition *Treatment 2 permissible* is

$$k + l < N/2 \ .$$

Note that, at any stage $M < N$, at least one treatment must be permissible.

To decide if state (i,j,k,l) is *prunable*, note that one can stop and declare Treatment 1 the better if and only if the number of successes on Treatment 1 exceeds the number of successes Treatment 2 would have if all future assignments to Treatment 2 resulted in successes. That is, Treatment 1 will be the winner in all possible final states that are reachable from this state by equal allocation rules if and only if $i > N/2-l$. Similarly, Treatment 2 will be the winner if and only if $k > N/2-j$. While ties can occur, in this problem no state with $M < N$ is prunable with a tie outcome; although such states can have tie outcomes among the reachable states, they must also have at least one reachable state with a nontie outcome. Thus the condition α *prunable* is

$$(i > N/2 - l) \ or \ (k > N/2 - j) \ .$$

Finally, ML_1 and ML_2 are the minimal possible average study lengths if the study reaches (i,j,k,l) and the next patient is assigned to Treatment 1 or Treatment 2, respectively. To determine ML_1, note that assigning the next patient (patient $M+1$) to Treatment 1 will either result in the state $(i+1,j,k,l)$ or the state $(i,j+1,k,l)$, depending on whether the treatment is a success or failure. Given ξ and M, the posterior probability of success is

$$\frac{a_0 + i}{a_0 + i + b_0 + j} \ ,$$

$N \to$	20	50	100	200	400
Curt. Alt.	16.2	39.4	78.1	155.3	309.8
Optimal	15.2	36.1	70.8	140.2	278.8

Table 1: $p_1 \sim Be(1,1), \quad p_2 \sim Be(1,1)$

$N \to$	20	50	100	200	400
Curt. Alt.	16.7	41.0	81.4	162.3	324.0
Optimal	15.9	38.4	75.9	150.8	300.5

Table 2: $p_1 \sim Be(1,1), \quad p_2 \sim Be(25,25)$

and thus ML_1 is

$$ML_1(i,j,k,l) = \frac{a_0+i}{a_0+i+b_0+j} \cdot ML(i+1,j,k,l) + \frac{b_0+j}{a_0+i+b_0+j} \cdot ML(i,j+1,k,l) .$$

Similarly, ML_2 is given by

$$ML_2(i,j,k,l) = \frac{c_0+k}{c_0+k+d_0+l} \cdot ML(i,j,k+1,l) + \frac{d_0+l}{c_0+k+d_0+l} \cdot ML(i,j,k,l+1) .$$

The assignment $ML(\alpha) := \min(ML_1, ML_2)$ is the essence of the *principle of optimality* which is being exploited by dynamic programming. The principle indicates that the minimal average study length of any study passing through the given state is the smallest of the minimal average study lengths for each permissible treatment.

The algorithm does not explicitly show the allocation procedure that achieves the minimal average study length, but it can be determined by noting for each state whether the algorithm decides to terminate or, if not, which treatment gives the smaller expected study length. This information can be stored and used to conduct the study.

5 Results

In Tables 1-4 we have given results comparing curtailed alternating allocation to the optimal curtailed equal allocation. Each table corresponds to a different configuration of the prior parameters a_0, b_0, c_0, d_0, and gives the average study length for several different values of N. All the entries are exact to within roundoff errors.

The basic heuristics to be learned from the tables are as follows:

1. For similar means, smaller variances correspond to greater average study lengths;

$N \to$	20	50	100	200	400
Curt. Alt.	16.1	39.2	77.7	154.6	308.3
Optimal	15.1	36.1	71.0	140.7	280.0

Table 3: $p_1 \sim Be(1,1), \quad p_2 \sim Be(40,10)$

$N \to$	20	50	100	200	400
Curt. Alt.	18.0	44.6	88.9	177.5	354.7
Optimal	16.1	38.7	76.3	151.6	302.2

Table 4: $p_1 \sim Be(4,1), \quad p_2 \sim Be(40,10)$

2. As the difference between the means becomes smaller, the average study length becomes greater.

The first point is illustrated by comparing Table 1, where both treatments have a uniform prior, with Table 2, where Treatment 2 has a prior whose mean is still 1/2 but whose variance is much smaller. The second point is illustrated by comparing Tables 2 and 3, where in Table 3, Treatment 1 continues to have a uniform prior, but now Treatment 2 has a prior with a mean of 4/5.

Finally, in Table 4, we are trying to choose the better treatment when one has a prior of Be(4,1), and the other has a prior of Be(40,10). Comparing to Table 3, we see that the equal means in Table 4 increases the expected study length for both procedures, and does so much more dramatically for alternating allocation than for the optimal allocation.

6 Further Remarks

In this short conference paper, we have illustrated only the ability to minimize the average study length within the class of curtailed equal allocation rules, which were selected because they are Δ-optimal. However, the same algorithmic approach can be applied to other situations. Within the class of curtailed equal allocation rules, one can apply the constrained dynamic programming method described in Section 4 to such problems as

1. Minimizing the expected number of failures;
2. Minimizing the expected number of patients assigned to the inferior treatment; and
3. Minimizing the expected cost of the experiment when different costs may be incurred for each different (treatment, result) possibility,

where expectation is taken to mean average with respect to the prior distribution on (p_1, p_2).

For these alternative optimizations, the only changes needed are in the assignment of ML at states that are terminal. For example, to minimize the number of failures, the new value of ML at terminal states is just $j + l$.

Another extension that is easy to incorporate is that in which a a third decision is allowed at the end of the study. For example, rather than randomly picking one of the treatments as better if the observed means are similar, one may prefer to declare a study outcome of "no difference". In this case, a new decision rule would be specified and the test for prunability of a state would become slightly more complicated. Otherwise the algorithm would be just as in Figure 1.

There are many other easy variations. For example, the present program allows us to optimize on one criteria, such as study length, and then to evaluate the resulting procedure on other criteria, such as number of failures. Another goal may be to determine how robust a design is. Here, one can create a procedure using one set of priors but evaluate it using a different set. All such evaluations would be exact, performed using backward induction.

Historically, the use of backward induction to help solve sequential allocation problems has been limited due to the method's computational intensity. Lately, however, it has become quite practical given the current speeds and memory capacity of computers. For example, on a large workstation one can carry out designs and evaluations for $N > 400$, and on a parallel computer one can handle $N > 1000$. Further, having used a workstation or departmental computer to design and evaluate an allocation procedure for, say $N = 400$, the procedure can be compressed and stored on a personal computer, available for ready access in many testing situations. These computational points will be explored more fully in an expanded version of this paper and in [3].

The binomial selection problem has been examined in settings far too numerous to note here. One paper that is particularly worthwhile to reference here, however, is an unpublished work of Berry and Eick [2] that was called to our attention during the presentation of the present paper. The special relevance of Berry and Eick [2] is that it is one of the few papers of which we are aware that utilizes backward induction to perform exact evaluations of non-recursive procedures. In their paper, the goal was to attain a tradeoff between failures and the probability of selecting the better treatment, without curtailment. Instead of maintaining optimal probability of selecting the better treatment by searching through equal allocation rules, dynamic programming was used to minimize the sum of the number of failures observed during a study of N patients, plus the expected number of failures in a future population of size P, assuming that the treatment declared better from the study was used for this population. The expected failure rate in the second population was the posterior estimate of the failure rate for the better treatment, based on the prior and the observations during the study.

Intuitively, in such an approach, for fixed N, increasing P increases the importance of correctly selecting the better treatment, while decreasing P increases the importance of minimizing failures during the initial study. While this idea is certainly not unique to Berry and Eick [2], the approach that they take with respect to "selecting the better treatment" is a Bayesian notion. In this paper we use a frequentist version of the properties of the decision rule and a Bayesian version of average study length. This latter approach is part of a general program of blending Bayesian and frequentist ideas in both the design and analysis phases of an experiment. Such a blending, while often controversial, is highly flexible. Not only can it be useful in a wide range of applications, but it can also yield designs satisfying a variety of statistical criteria.

References

[1] Bellman, R. (1957) *Dynamic Programming.* Princeton University Press.

[2] Berry, D.A. and Eick, S.G. (1987), Decision analysis of randomized clinical trials: comparison with adaptive procedures. *Unpublished manuscript.*

[3] Hardwick, J. and Stout, Q. F. (1992), Computational aspects of sequential allocation for testing with multiple criteria. In preparation.