In *Computing Science and Statistics* **30** (1998), pp. 219–223.

Adaptive Allocation in the Presence of Missing Outcomes

Janis Hardwick[1] Robert Oehmke Quentin F. Stout[1]

University of Michigan, Ann Arbor, Michigan 48109 USA

Abstract

We consider the ways in which adaptive allocation is altered when some of the observations become unavailable. Such "missing outcomes" are a strong form of censoring. The problems analyzed involve adaptive sampling from two Bernoulli populations. Both fully sequential designs and few-stage designs are examined. For each type of design, we developed an algorithm to determine the optimal design. Prior approaches were ad hoc and did not fully optimize under the censoring assumptions. Perhaps one reason for this is that censoring turns the populations into 3-outcome populations, which increases the complexity of the problem. We also convert designs for uncensored problems into natural designs for the censored problem, and compare to the optimal design.

Keywords and phrases: censoring, dynamic programming, sequential sampling, few-stage, design of experiments, algorithms

1 Introduction

An *adaptive allocation* problem is one in which an investigator has the option to determine how to sample while the experiment is being carried out, using the results that have been observed so far. This is in contrast with the standard technique of *fixed allocation*, in which all sampling decisions are made prior to beginning the experiment. Adaptive allocation is useful because in certain situations it can be dramatically superior to standard techniques. For example, it can often allow significant reductions in costs or fatalities over fixed allocation without sacrificing statistical objectives such as maximizing the probability of determining the best treatment. Unfortunately adaptive allocation is more complex to analyze, which has often inhibited its utilization.

To help make adaptive allocation more practical, we have been developing a collection of algorithms to help in their design, optimization, and analysis. As part of this process, we are adding the ability to deal with various real-world factors. One such factor is the censoring of observations. There are many forms of censoring that can take place, and here we consider only the strongest form, "missing outcomes", in which no information is obtained about the population being sampled. For example, one may assign a patient to a drug therapy, but then an accident unrelated to the therapy occurs and the effects of the drug become unknown. This is in contrast with other forms of censoring, such as right-censoring in survival data, in which at least some information is obtained from each censored observation.

In this paper we begin an investigation of how censoring affects adaptive allocation. For a given setting, there are several natural questions which arise, a few of which are:

Q1: What is the optimal design for the censored problem?

Q2: How well can one predict the expected outcome of the optimal design for the censored problem, by using the outcome for the optimal design for the uncensored problem?

Q3: How well do interesting suboptimal designs perform relative to the optimal design?

Q4: How can one utilize a design for the uncensored problem in situations where censoring occurs?

We provide answers, or mechanisms for obtaining the answer, for these questions for several classes of adaptive allocation problems. To the best of our knowledge, this is the first investigation of the design of adaptive algorithms that are optimal for censoring. The closest work we have been able to find is that of Eick [1] and Hayre and Turnbull [4], but these papers consider models that are quite different than ours.

In Section 2, we define our censoring model and the types of adaptive allocation problems we consider. In Section 3, we describe the approaches we use for answering Q2 and Q4, and in Section 4 we describe the algorithms which generate the optimal design for Q1 and which provide the analyses for arbitrary designs. In Section 5 we describe the results obtained when these procedures were applied to two sample problems, and mention further research.

2 Models and Problems Considered

Our initial work is for adaptive designs in which there are two Bernoulli populations, denoted Pop 1 and Pop 2, and a fixed

[1]Research supported in part by National Science Foundation under grants DMS-9157715 and DMS-9504980.

sample size n. The fixed sample size is merely for convenience of explanation, and is not a requirement of the procedures. Sampling Pop i at the j^{th} observation has outcome

$$X_{ij} = \begin{cases} 0 & \text{"failure" wp } 1 - P_i \\ 1 & \text{"success" wp } P_i \end{cases}$$

for $i = 1, 2$ and $j = 1, \ldots$, where the population success parameters P_i are unknown.

The censoring is modeled as another Bernoulli independent process $\mathcal{C}$ with outcomes:

$$\mathcal{C}_j = \begin{cases} 0 & \text{"not censored" wp } 1 - P_c \\ 1 & \text{"censored" wp } P_c \end{cases}$$

for $j = 1, \ldots$.

To combine censoring and the experiment we overlay the two population sampling problem with the censoring process to get outcomes as follows:

1. For trial j select a target population, Pop 1 or Pop 2.

2. Next, sample from the $\mathcal{C}$ process.

3. If $\mathcal{C}_j = 1$ then the j^{th} outcome is **censored**.

4. If $\mathcal{C}_j = 0$ then observe a sample from the population targeted in (1): Then the j^{th} outcome is either a **success** or a **failure** from the population sampled.

Thus the outcome of each trial is either a success or a failure on the targeted population, or it is censored.

Note that the censoring is assumed to be independent of the population being sampled, which is appropriate for many, although not all, situations. Another model would be to have two independent censoring processes, one for each of the populations in the experiment. Here, due to time and space constraints, we investigate only the former model, and we plan to investigate the latter model later.

At stage $j = 0, \ldots, n$, let (s_i, f_i) represent the number of successes and failures observed on Pop i for $i = 1, 2$, and let c represent the number of censored observations. Then $\langle s_1, s_2, f_1, f_2, c \rangle$ forms a vector of sufficient statistics for this model, which we call a *state*. Also, we utilize a Bayesian approach, where on the parameters (P_1, P_2, P_c) we assume a prior distribution that is taken to be the product of the three independent prior distributions.

2.1 Objective Functions

We assume that there is some objective function $\mathcal{O}$ which depends on the final state of the experiment, and that the goal is to optimize the expected value of $\mathcal{O}$ at the end of the experiment. This is a quite general framework, and all of our design and analysis algorithms will work for arbitrary objective functions. However, to illustrate results it is necessary to pick specific objective functions, so we have selected the following two as examples.

2-arm bandit In this problem, the objective is to maximize the number of successes. This is a well-known problem with an extensive literature and applications to a great many areas.

product of means This is a non-linear estimation problem, where the goal is to minimize the mean squared error of the posterior estimate of $P_1 \cdot P_2$. Several people have studied fixed and adaptive designs for this problem, see the references in [3].

2.2 Types of Adaptive Designs

There are many different ways that adaptive experiments can be structured, and in general each different structure requires a different set of algorithms for design and analysis. In this work, we restricted ourselves to considering three types of designs: fully sequential, 1-stage, and 2-stage.

In *fully sequential allocation* the trials occur one at a time, and each sampling decision can be based on the outcomes of all previous trials. This includes the most powerful designs, as it permits designs which can make increasingly better decisions about which populations to sample from next, and thus quickly adapt to achieve the goals of the problem.

In contrast to sequential allocation, in *staged allocation*, trials occur in groups. The trials in each group can overlap, but different groups don't overlap. This allows the results of all previous stages to be available before making the decision for the next stage. Here we only examine two types of staged allocation: 1-stage and 2-stage. 1-stage allocation, also known as "fixed" allocation, is non-adaptive, as the only decision that is allowed is how many trials to assign to each population. The 2-stage sampling design is more complex. First, a decision is made about how many trials to use in the first or "pilot" stage, and to which populations to assign them. The second stage is then just like the 1-stage procedure, except we now have the information gathered from the first stage to use in making the allocation decision.

3 Methodology

Once one has selected an objective function, sample size, prior distributions, and type of adaptive experiment desired, the optimal design can be determined by the algorithms described in Section 4. However, as noted in the Introduction, there are other natural questions that arise. For example, because most approaches assume no censoring, uncensored designs may already have been evaluated before the censoring was considered, and hence one may wish to predict an expected value from a censoring design for the problem based on an expected value from the design assuming no censoring.

To describe some options, let $\mathcal{D}$ represent a family of designs for a given problem and type of adaptive experiment. For example, it might be a family of 2-stage designs for bandits which uses equal allocation on a first stage of size $\sqrt{n}$, followed by allocating all observations on the final stage to the population with the highest posterior mean at the end of the first stage.

3.1 Predicting Censored Performance

If $\mathcal{D}$ includes designs for the censored problem, then let $D(n, F(P_c))$ represent the design for sample size n with and prior distribution F on P_C. Note that $D(n, F(P_c))$ also depends on the prior distributions for P_1 and P_2, but these will be omitted from the notation throughout the rest of the paper. Let $\mathbf{E}[\mathcal{O}(D(n, F(P_c)))]$ represent the expected value of the objective function obtained by using $D(n, F(P_c))$.

One naive answer to Q3 is to estimate this by ignoring censoring, equivalently, by assuming F is a point mass at 0, in which case the estimate is merely $\mathbf{E}[\mathcal{O}(D(n, 0))]$. In general, a far better answer can be obtained by adjusting the sample size. Since the expected number of uncensored observations is $n(1-\mu(F))$, this provides a natural "equivalent" sample size, and thus, a natural approximation for Q3 is $\mathbf{E}[\mathcal{O}(D(n(1-\mu(F)), 0))]$.

3.2 Extending Uncensored Designs

Extending designs not intended for censoring into a form that can be applied when censoring occurs has varying degrees of difficulty. For example, a common approach to fully sequential designs is to use a myopic approach (also known as a greedy approach), in which the next population to be sampled is the one that will minimize the expected value of the objective function if the experiment were immediately terminated. Such a design can be naturally used in the presence of censoring, and whenever a given population is sampled the sampling on it continues until an uncensored result is obtained.

However, many other adaptive designs are more complicated, and base their decisions, either explicitly or implicitly, both on the observed data and on the number of observations remaining. To describe how an arbitrary family can be extended, suppose $\mathcal{D}_u$ denotes a design family for the uncensored problem. Let $D_u(m \mid \langle s_1, f_1, s_2, f_2 \rangle)$ denote the decision made by the member of $\mathcal{D}_u$ which has observed state $\langle s_1, f_1, s_2, f_2 \rangle$ and which has m observations remaining. For example, in a two-stage design family, $\langle s_1, f_1, s_2, f_2 \rangle$ might represent the results of the first stage, and m would be the length of the second stage. Note that there might not be a natural member of $\mathcal{D}_u$ which has observed state $\langle s_1, f_1, s_2, f_2 \rangle$. For example, if $\mathcal{D}_u$ is a 2-stage design which always allocates equally in the first stage, when $s_1 + f_1 \neq s_2 + f_2$ the state $\langle s_1, f_1, s_2, f_2 \rangle$ could not be observed at the end of a first stage. Thus, in such a situation, one would first need to expand the family to decide what to do in the second stage.

There are two natural ways to extend $\mathcal{D}_u$ to handle censoring. The simplest, which we call the *oblivious* extension, is to make each decision using the uncensored observations obtained, assuming that no further censoring will occur. Thus, for 2-stage design, the first stage allocation would be that of $D_u(n)$, and if the resulting state was $\langle s_1, f_1, s_2, f_2, c \rangle$ then the second stage allocation would be the second stage allocation of $D_u(n - s_1 - f_1 - s_2 - f_2 - c \mid \langle s_1, f_1, s_2, f_2 \rangle)$. For 1-stage design there is no change, and for fully sequential designs the changes are as for the 2-stage design.

Another extension, which we call *cognizant*, is to make the decision using the observations obtained so far and assuming that there is an equivalent uncensored sample size remaining. If $\hat{P}_c$ represents the posterior estimate of P_c, and $m = s_1 + f_1 + s_2 + f_2 + c$ is the number of trials so far, then the equivalent sample size remaining is $(n-m)(1-\hat{P}_c)$, and for a fully sequential problem the decision is that of $D_u((n-m)(1-\hat{P}_c) \mid \langle s_1, f_1, s_2, f_2 \rangle)$. For staged allocation one must also scale the allocation. For example, for 1-stage allocation, the equivalent sample size is $n(1-\hat{P}_c)$. If $D_u(n(1-\hat{P}_c) \mid \emptyset)$ would allocate o_i observations to Pop i, then the cognizant extension would allocate $o_i/(1-\hat{P}_c)$ observations to Pop i.

4 Algorithms Developed

For each type of design considered (fully sequential, 1-stage, or 2-stage), we developed algorithms to

- determine the fully optimal censoring design;
- determine the expected value of the objective for an arbitrary censoring design;
- determine the expected value of the objective for the oblivious and cognizant extensions of an arbitrary uncensored design.

In this section we briefly describe the algorithms, and their computational complexity. Due to space limitations, the details of the algorithms are omitted. In general we only refer to the changes needed, compared to the algorithms for the uncensored situation. See [2, 3] for a more detailed explanation and pointers to the algorithms for the uncensored situation.

4.1 Optimal Designs

Determining the optimal design for a censored environment is very similar to determining the optimal design for an uncensored one. One uses dynamic programming. However, the censored programs are more challenging because of the increase in the number of outcomes possible. Note that a different dynamic programming algorithm is needed for each type of experiment, but that the algorithm (as opposed to its output) does not depend on the objective function or priors.

For fully sequential designs, the time to find the optimal uncensored design is $\Theta(n^4)$, which is the size of the state space. For censored outcomes, the number of states is $\binom{n+5}{5} \approx n^5/5!$, and there are 3 outcomes for each of the two possible options, rather than 2 outcomes as in the uncensored case. The total time is $\Theta(n^5)$.

For 1-stage allocation, there are $n+1$ possible different options (i.e., one need only determine how many observations to sample from Pop 1 and place those remaining on Pop 2). In the uncensored problem there are $O(n^2)$ outcomes per option, so a simple summation over them determines the optimal option in $\Theta(n^3)$ time, which is linear in the number of terminal states. For many specific objective functions it is possible to algebraically evaluate an option in constant time, which would reduce the time to $\Theta(n)$, but here our analyses will be for the worst-case general situation where no such reductions are assumed.

However, in the censored case, each option may result in $O(n^4)$ outcomes, so a simple summation would take $O(n^4)$ time per option, and $\Theta(n^5)$ time overall, which is superlinear in the number of terminal states ($\binom{n+4}{4} = \Theta(n^4)$). To achieve time linear in the number of terminal states, we create terminal meta-states $\langle o_1, o_2, c \rangle$, where o_i denotes the number of uncensored observations on Pop i. Since each terminal state contributes to exactly one meta-state, the expected value of the objective function for all of the meta-states can be easily determined in time linear in the number of true terminal states. Then one can evaluate each option by a weighted sum over all of the meta-states it could produce. There are only $O(n^2)$ meta-outcomes per option, and the total time is reduced to $\Theta(n^4)$.

For 2-stage allocation, the fastest algorithm known for the uncensored situation first determines, for each intermediate state, the optimal final stage from that state, and then determines the optimal initial state. There are $\Theta(n^4)$ intermediate states, with $O(n)$ options per state and $O(n^2)$ outcomes per option. Through some algebraic manipulation, the total time for determining the optimal decision at each intermediate state can be reduced to the product of the number of states and the number of options, giving $\Theta(n^5)$ time. The time for determining the optimal first stage is $\Theta(n^4)$, so the total time is $\Theta(n^5)$. For the censored situation, the number of intermediate states is $\Theta(n^5)$, with $O(n)$ options per state and $O(n^4)$ outcomes per option. Similar algebraic manipulations keep the total time down to a constant per state-option, or $\Theta(n^6)$. The optimal first stage can then be found in $\Theta(n^5)$ time, so the total is $\Theta(n^6)$.

4.2 Evaluating Designs

To determine the expected value of an objective function for a given fully sequential design, one typically uses backward induction. This is a close relative of dynamic programming, but there is no optimization involved and only one option to consider per state (the option the design chooses). Only designs with deterministic decisions at each state will be considered here — random sampling decisions can be similarly evaluated, but the time increases depending on the number of possibilities. In general, for fully sequential allocation one needs to evaluate all states, so the time for the censoring model would again be $\Theta(n^5)$.

For 1-stage allocation there are $\Theta(n^4)$ possible outcomes, so the time would be $\Theta(n^4)$. For 2-stage allocation there are $\Theta(n^4)$ possible outcomes at the end of the first stage. Evaluating each of these separately would take $\Theta(n^8)$ time, so instead one would use an approach similar to the dynamic programming approach for optimization, and merely record the value corresponding to the design's decision at each intermediate state. This reduces the time to $\Theta(n^6)$.

The above analyses assumed that the design's decisions could be determined in time no worse than the evaluation time. For the oblivious and cognizant designs there may be additional computational challenges. For example, we evaluated the oblivious and cognizant versions of the optimal uncensored designs. For the sequential case this required determining the solution for all sample sizes less than or equal to n, which takes $\Theta(n^5)$ time. While this is not worse than the analysis time, it is a nontrivial increase over the $\Theta(n^4)$ time needed to determine only the optimal uncensored solution for n. The space required increased similarly.

5 Conclusions

To answer questions Q2 and Q3, the algorithms described above were run for various sample sizes n and for each combination of experiment type and the two sample objective functions. For each combination we also compared the optimal censored design to the oblivious and cognizant versions of the optimal uncensored design.

For Q2 and the 2-armed bandit problem, we found that the expected value of the uncensored design applied to an equivalent sample size, as discussed in Section 3.1, was a poor estimator of the actual value obtained from the optimal censoring version. In fact, the prediction by the non-censoring version varied by as much as 10% from the censoring value.

For Q3 and the 2-arm bandit, the oblivious extension of the uncensored design, as discussed in Section 3.2, was fairly good as a method for handling censoring. Among all the cases we tried it was at most 1% worse than the optimal. The cognizant extension was even better, differing by at most 0.1% over all of the cases we tried. It is interesting that this was true for all of the types of experiments considered.

The results for the product of means problem were very similar to those for the 2-armed bandit problem. Again, we found that the uncensored design was a poor estimator of the performance obtained from the optimal censoring design. This time the uncensored estimate varied by as much as 30% from the true censoring value. The oblivious and cognizant versions behaved basically the same as for the 2-arm bandit, differing by at most 1% and 0.1% respectively.

Our data point to the unexpected result that simple modifications to adaptive allocation procedures not originally designed for censoring can yield procedures that perform extremely well when confronted with it. Unfortunately, an analysis of the expected outcome obtained by these methods required the same amount of computational resources as the optimal method. In general, then, our experimental results can be summarized as

> *For adaptive allocation with a single censoring mechanism, censoring is more of an analysis nuisance than an optimization problem.*

In the future, we plan to pursue various extensions of this work. One is to investigate how having different censoring rates for the different populations would affect the performance of the oblivious and cognizant design extensions (with refinements to the cognizant extension to cope with the differing rates). We believe that for the staged allocation, the oblivious extension will fare poorly in such a setting if the censoring rates are significantly different, although the oblivious fully sequential extension may still show acceptable performance. We believe that the oblivious staged allocation performed well for the single censoring mechanism in the problems considered because the optimal allocation quickly converges to a fixed ratio allocation between the populations within the initial stage, with the second stage converging to a fixed ratio dependent on the outcome of the first stage. Since the oblivious allocation converged to the same ratio, it resulted in very similar allocation, although the initial stage was too small. If the censoring rates differ on the different populations, then the oblivious allocation will no longer be in the correct ratio and we can expect performance to suffer.

Another area to explore is the manner in which censoring affects a more varied selection of problems and types of experimental designs. Because the design and analysis is more computationally challenging than the uncensored problem, evaluating useful sample sizes may require porting some of the programs to parallel computers.

References

[1] Eick, S.G. (1988), "The two-armed bandit with delayed responses", *Annals of Statistics* **16**, pp. 254–264.

[2] Hardwick, J. and Stout, Q.F. (1993), "Exact computational analyses for adaptive designs", *Adaptive Designs* (B. Rosenberger & N. Flournoy, ed.'s), Institute Math. Stat. Lec. Notes **25**, pp. 223–237.

[3] Hardwick, J. and Stout, Q.F. (1998), "Optimal few-stage allocation procedures", submitted.

[4] Hayre, L.S. and Turnbull, B.W. (1981), "A class of simple approximate sequential tests for adaptive comparison of two treatments", *Comm. in Statistics A — Theory and Methods* **10**, pp. 2339–2360.