

Simulation-Based Sequential Bayesian Design¹

Peter Müller*, Don A. Berry

*Department of Biostatistics, University of Texas M.D. Anderson Cancer Center,
Houston, TX, U.S.A.*

Andy P. Grieve,

Pfizer Global Research & Development, Sandwich, UK

Michael Smith

Pfizer Clinical Statistics, Sandwich, U.K.

Michael Krams

Wyeth Research, Collegeville, PA²

Abstract

We consider simulation-based methods for exploration and maximization of expected utility in sequential decision problems. We consider problems which require backward induction with analytically intractable expected utility integrals at each stage. We propose to use forward simulation to approximate the integral expressions, and a reduction of the allowable action space to avoid problems related to an increasing number of possible trajectories in the backward induction. The artificially reduced action space allows strategies to depend on the full history of earlier observations and decisions only indirectly through a low dimensional summary statistic. The proposed rule provides a finite-dimensional approximation to the unrestricted infinite-dimensional optimal decision rule. We illustrate the proposed approach with an application to an optimal stopping problem in a clinical trial.

Key words: Backward induction; Forward simulation; Monte Carlo simulation; Optimal design; Sequential decision.

1991 MSC: 62C10, 62K05, 62L05, 62L15

1 Introduction

We consider simulation-based methods for exploration and maximization of expected utility in sequential decision problems. Formally, decision making under uncertainty is choosing an action d to maximize expected utility $U(d) = \int u(d, y, \theta) p_d(\theta, y)$. Here, $u(d, y, \theta)$ is the utility function modeling preferences over consequences and $p_d(\theta, y)$ is a probability distribution of parameter θ and observation y , possibly influenced by the chosen action d . See Chaloner and Verdinelli (1995) and Verdinelli (1992) for reviews of Bayesian approaches to decision problems traditionally known as optimal design. Spiegelhalter et al. (1994), Berry (1993) and Berry and Stangl (1996) discuss general issues related to the use of Bayesian optimal design methods in medical decision problems.

In many applications decisions are made sequentially. Let $d = (d_1, \dots, d_T)$, and $y = (y_1, \dots, y_T)$. Assume that y_t depends on d_1, \dots, d_t only and that decision d_t is made after decision d_{t-1} and after observing y_{t-1} . Then $d_t = d_t(d_1, \dots, d_{t-1}, y_1, \dots, y_{t-1})$ is a policy depending on the observed outcome from the first $t - 1$ decision periods. Such sequential decision problems where later decisions depend on earlier outcomes are notoriously difficult. A complete solution requires, in general, backward induction involving an exponentially increasing number of possible scenarios. See, for example, Berger (1985, chapter 7). In problems with continuous outcomes the set of possible decision rules is infinite dimensional. The proposed simulation methods overcome two important practical problems which hinder a routine application of backward induction. These are related to the large number of scenarios that need to be considered, and the evaluation of many possibly analytically intractable expected utility integrals.

Many sequential decision problems involve stopping: y_t are independent observations from a common density $p(y_t|\theta)$, and one decides after each observation whether to stop sampling and make an immediate final decision or take another observation. Let $y^t = (y_1, \dots, y_t)$. It is convenient to write $d = (\tau, \delta)$, where $\tau = (\tau_1, \tau_2, \dots)$ is a stopping rule with $\tau_t(y^t)$ defining the probability of stopping at time t , and $\delta(y^t)$ specifies the final decision to be made if we stop at time t . The final decision could be, for example, a hypothesis test, or the

* Corresponding author. Tel.: +1-713-563-4296, Fax: +1-713-4242

Department of Biostatistics, M. D. Anderson Cancer Center,
1515 Holcombe Boulevard, Box 447, Houston, TX 77030-4009

Email addresses: pmueller@mdanderson.org (Peter Müller),
dberry@mdanderson.org (Don A. Berry), andy.p.grieve@pfizer.com (Andy P. Grieve),
mike.k.smith@pfizer.com (Michael Smith), kramsm@wyeth.com (Michael Krams).

¹ *brief running title:* "Sequential Bayesian Design"

² research carried out while at Pfizer, Sandwich, UK

decision whether to abandon a drug’s development or seek regulatory permission to market it. Although sequential sampling is a convenient example to bear in mind for the following development, the proposed methods apply more generally. Our approach is appropriate for any decisions that do not influence the probability model beyond identifying a subset of possible responses. Beyond this, there are no constraints on the probability model and the structure of the utility function. The allowable decisions include those that terminate an experiment at a particular time, like sequential sampling, and also those that select from alternative probability models, e.g., probability models associated with alternative treatments, as is the case, for example, in bandit problems (Berry and Fristedt, 1985). Also, we consider only problems with a finite horizon T . The underlying probability model is (almost) entirely unconstrained. In particular, the model can involve random functions. This is useful, for example, in the context of a dose finding trial with an unknown dose-response curve.

In some problems the optimal policy d is characterized by a low dimensional summary vector of control parameters. For example, Carlin et al. (1998) address stopping a clinical trial at each of K interim analyses. In their setup they show that the optimal Bayes sequential procedure can be characterized by a set of $2K$ critical cutoffs for the posterior mean on a parameter which quantifies the advantage of a treatment over placebo. The problem is reduced to choosing these cutoffs. Similar simplifications apply whenever the loss function and the posterior distribution depend on the current history only indirectly through a low dimensional statistic. Lewis and Berry (1994) discuss optimal sequential design under binomial sampling and a 0-1 type inference loss related to a hypothesis test about the unknown success probabilities. In their setup, at any given time the state of information is given by the number of patients treated and the parameters of the beta posterior distribution on the binomial success probabilities. Christen and Nakamura (2003) discuss sequential stopping rules for species accumulation. They show that the optimal rule depends on the current history only indirectly through a bivariate summary, namely the number of species accumulated and the area under the accumulation curve. Hsiao and Clayton (2001) show that truncated repeated significance boundaries (Lan and DeMets, 1983) can be optimal from a Bayesian perspective. They assume a decision problem related to testing a hypothesis about the drift parameter in a Brownian motion with drift. Hsiao and Clayton demonstrate how under an appropriate choice of loss function and sampling cost repeated significance test boundaries can be optimal from a Bayesian perspective.

Alternative Bayesian approaches to optimal sequential design in medical decision problems are discussed, among other references, in Thall et al. (1995) who define stopping criteria based on posterior probabilities of clinically meaningful events. Similarly, Thall and Russell (1998) define a sequential procedure based on monitoring posterior probabilities of certain events. Using reasonable ad-

hoc rules based on these probabilities they define designs and evaluate their frequentist performance. Vlachos and Gelfand (1998) follow a similar strategy. Whitehead and Brunier (1995) and Whitehead and Williamson (1998) use what is essentially a Bayesian m -step look-ahead procedure to find the optimal dose to assign to the next m patients in a dose-finding study. Spiegelhalter et al. (1994) propose a Bayesian approach to monitoring clinical trials based on posterior intervals of the unknown treatment success parameters. Depending on the location of the posterior interval relative to some critical thresholds they propose to consider appropriate decisions.

For inference problems with a low dimensional sufficient statistic to summarize the posterior distribution Brockwell and Kadane (2003) propose a grid-based scheme for numerical optimal sequential design that is similar to the solution proposed in this paper. They use one-step ahead forward simulation to evaluate expected utilities. They focus on problems with inference loss related to parameter estimation, although the algorithm could be easily used for more general loss functions.

The methods described in this paper could in theory also be applied for non-sequential problems. However, we do not recommend to do so. Without the complications arising from the sequential nature of the problem other methods are more efficient and preferable. A discussion of simulation-based methods for non-sequential decision problems appears in Müller et al. (2004) and Amzal et al. (2003).

In Section 2 we introduce the motivating case study. In Section 3 we review simulation based optimal design in non-sequential problems. In Section 4 we introduce an approach for simulation based sequential design. Section 5 illustrates the approach in two examples. Section 6 is the concluding discussion.

2 Optimal Stopping in a Clinical Trial

The motivation for the proposed methods is a dose-finding clinical trial. Patients are recruited into the trial at participating centers. Based on the information available at each time period t , say once a week, we have to decide (d_t) whether to terminate the trial and abandon the drug's development ($d_t = D0$), continue with dose-finding ($d_t = D1$), or terminate the dose-finding trial and switch to a pivotal trial ($d_t = D2$). The decision is allowed to depend on all data observed up to time t , and trivially depends on earlier decisions d_1, \dots, d_{t-1} by force of the fact that we only reach time t if all earlier decisions were to continue.

Let T denote the maximum number of periods that the trial is allowed to

last. Let $p(y_t, t = 1, 2, \dots, T \mid \theta)$ denote a probability model for the data, parametrized by $\theta = (\theta_1, \dots, \theta_p)$, but not dependent upon d except for the fact that we only get to observe y_t if we decided to continue the trial at least until time t . The probability model includes a dose-response curve $f_\theta(z)$ which gives the mean response of a patient treated at dose z . The model is completed with a prior probability distribution $p(\theta)$. For a given parameter vector θ , the difference $\Delta = f_\theta(z^*) - f_\theta(0)$ corresponds to the advantage of the proposed new treatment (at the eventually recommended dose z^*) over placebo ($z = 0$). Let $y^t = (y_1, y_2, \dots, y_t)$ denote the observations up to time t . The posterior moments $m_t = E(\Delta \mid y^t)$ and $s_t^2 = Var(\Delta \mid y^t)$ feature prominently in the proposed decision rule. All we require from the probability model is that we can generate (approximate) Monte Carlo samples $\theta \sim p(\theta \mid y^t)$ by appropriate Markov chain Monte Carlo simulation. A detailed description of the probability model and the posterior simulation scheme is discussed in Berry et al. (2001). The ability to accommodate a complex hierarchical probability model and to use a utility function which involves non-linear summaries of the posterior distribution is key to the successful implementation in that application.

We use a utility function to quantify the worth of consequences of possible decisions. We assume a fixed sampling cost c_1 per patient in the trial. The payoff is $c_2 \cdot \bar{\Delta}$ if we decide to initiate a pivotal trial *and* the pivotal trial concludes that the drug at the recommended dose is in fact an effective treatment. There is no payoff if the drug's development is stopped or if the pivotal trial turns out to be negative. Here $\bar{\Delta}$ is the advantage of the new treatment over placebo, as estimated in the pivotal trial at the finally recommended dose. For example, c_1 could be \$10,000, and c_2 could be \$10,000,000. The decision to continue the trial will depend on the tradeoff c_1/c_2 between sampling cost and benefit.

Implementing optimal stopping using the simulation based method proposed in this paper has substantial advantages. First, the number of patients in a sequential trial will usually be substantially smaller than when using standard designs. This has important economical and ethical implications. Second, the proposed approach allows a seamless transition between the dose-finding and confirmatory stages. This eliminates the time required to set up a second trial. Details are described in Berry et al. (2001). A related report in an industry journal (Farr-Jones, 2001) and an article by Malakoff (1999) highlight the practical impact of the proposed design.

In this paper we discuss only the decision of terminating versus continuation. Another important decision problem in a dose-finding clinical trial is the assignment of doses to patients assuming that the trial is continued. See Berry et al. (2001) for a (non-sequential) approach to that decision problem.

3 Optimal Design by Simulation

Consider a non-sequential decision problem in which

$$U(d^*) = \max_{d \in D} U(d) = \int u(d, y, \theta) dp_d(\theta, y). \quad (1)$$

Decision d^* is said to be optimal. The probability model $p_d(\theta, y) = p(\theta)p_d(y|\theta)$ is typically given as a prior distribution $p(\theta)$ on the parameters and a sampling distribution $p_d(y|\theta)$ for the observations given the parameters. In inferential problems, i.e., when the decision is related to inference about the unknown parameter θ , utility is typically a function of (d, θ) only. Negative expected utility $-U(d)$ is known as Bayes risk of the decision rule d , and $-U(d^*)$ is called the Bayes risk.

Even if the expected utility integration (1) is analytically intractable, it easily can be approximated by Monte Carlo simulation if the prior and likelihood are both available for efficient random variate (r.v.) generation, and if the utility function $u(d, y, \theta)$ can be evaluated for any given realization of the experiment (θ, y) . Efficient r.v. generation from the probability model is typically feasible. However, evaluating the utility function can be difficult. For example, if we wish to choose covariates in a non-linear regression to minimize expected posterior variances, then evaluation of $u(d, y, \theta)$ requires the posterior variance integral and this may well be analytically intractable and may require numerical quadrature.

Assuming that r.v. generation is feasible, and that the utility function can be evaluated pointwise, we can solve (1) by simulation as follows. Simulate experiments $(\theta_i, y_i) \sim p_d(\theta, y)$, $i = 1, \dots, M$, and evaluate for each simulated experiment the observed utility $u_i = u(d, y_i, \theta_i)$. Use $\hat{U}(d) = \frac{1}{M} \sum u_i$ to approximate $U(d)$. Using the approximate evaluations $\hat{U}(d)$ we could proceed with a suitable maximization method to find the optimal design $d^* = \arg \max \hat{U}(d)$. Carlin et al. (1998) use such Monte Carlo evaluation of expected utilities to find optimal thresholds to define stopping times in a sequential sampling problem. An attractive feature of Monte Carlo integration is that the probability model $p(\theta)$ is not restricted to any particular form; it needs only to be accessible for sampling. In many problems decisions have to be made conditionally on available data x , in which case $p(\theta)$ is replaced by $p(\theta|x)$. Wakefield (1994) considers choosing an optimal dose in a pharmacokinetic study, using Markov chain Monte Carlo simulation to generate from $p(\theta|x)$ where x is data available from earlier patients.

4 Constrained Backward Induction and Forward Simulation

4.1 Backward Induction

Solving a sequential design problem is complicated by the fact that later decisions can depend on earlier outcomes. To simplify notation we write y^t for (y_1, \dots, y_t) , and d^t for (d_1, \dots, d_t) . Assuming a finite horizon T , let

$$U_T(d_T, d^{T-1}, y^{T-1}) = \int u(d^T, y^T, \theta) p_d(y_T|\theta) dy_T p(\theta|y^{T-1}) d\theta \quad (2)$$

denote the posterior expected utility of the decision d_T at the end of the final period, conditioning on $I_T = \{d^{T-1}, y^{T-1}\}$, and marginalizing over the relevant posterior and posterior predictive distribution on the unknown parameter vector θ and the final observation y_T . Let $d_T^* = d_T^*(d^{T-1}, y^{T-1})$ denote the posterior Bayes decision which maximizes this expected utility, and let $U_T^*(d^{T-1}, y^{T-1}) = U_T(d_T^*, d^{T-1}, y^{T-1})$. Similarly

$$U_{T-1}(d_{T-1}, d^{T-2}, y^{T-2}) = \int U_T^*(d^{T-1}, y^{T-1}) dp_{d^{T-1}}(y_{T-1}|y^{T-2}) \quad (3)$$

is the expected utility at time $T-1$, assuming decision d_T^* in the final period. The optimal decision d_{T-1}^* is the one that maximizes U_{T-1} . In the special case of sequential sampling we interpret (3) with the understanding that if d_{T-1} specifies stopping at time $T-1$, then the data y_{T-1} is the empty set, and the right hand side of (2) reduces to $\int u(d^{T-1}, y^{T-1}, \theta) p(\theta | y^{T-1}) d\theta$.

Extending analogous definitions to $t = T-2, \dots, 1$, we arrive at the definition of the sequential decision problem as an alternating sequence of expectations (to find U_t) and maximizations (to find d_t^*). We write E_x for an expected value with respect to x . The relevant distributions are clear from definitions (2) and (3).

$$\begin{aligned} U(d_1^*) &= \max_{d_1} E_{y_1} \max_{d_2} E_{y_2} \dots \max_{d_T} E_{y_T, \theta} u(d^T, y^T, \theta) = \\ &= \max_{d_1} E_{y_1} \max_{d_2} E_{y_2} \dots \max_{d_T} U_T(d^T, y^{T-1}) = \\ &= \max_{d_1} E_{y_1} \max_{d_2} E_{y_2} \dots U_T^*(d^{T-1}, y^{T-1}) = \\ &\dots\dots\dots \\ &= \max_{d_1} E_{y_1} \max_{d_2} U_2(d_2, d_1, y_1) = \\ &= \max_{d_1} E_{y_1} U_2^*(d_1, y_1) = \\ &= \max_{d_1} U_1(d_1). \end{aligned} \quad (4)$$

A traditional solution of (4) starts by solving the maximization problem for d_T^* for all possible scenarios $I_T = \{d^{T-1}, y^{T-1}\}$. Having a table of solutions

for $d_T^*(I_T)$ and expected utilities $U_T^*(I_T)$ we can proceed to solve the maximization problem for d_{T-1}^* , substituting the appropriate values for d_T^* and U_T^* . Considering $t = T - 2, \dots, 1$, in sequence, we eventually find the optimal initial decision d_1^* .

There are at least two difficulties in implementing this backward induction scheme. First, it can require a great many maximizations. Even if the outcomes y_t are discrete, or can be appropriately discretized, the problem requires keeping track of solutions $d_t^*(d^{t-1}, y^{t-1})$ for an exponentially increasing number of scenarios. Second, the solution involves calculating expected utility integrals of the form (2) and (3), and these are typically analytically intractable. The proposed approach resolves both difficulties. We use a strategy of constrained decision spaces to reduce the number of possible scenarios to something manageable, and we use forward simulation and Monte Carlo integration to evaluate the required integrals.

4.2 Constrained Backward Induction

Although each decision d_t could depend on all earlier data and decisions $I_t = \{d^{t-1}, y^{t-1}\}$, typically only some critical summary of I_t is important. For example, Carlin et al. (1998) show that in the specific setup they consider the optimal decision depends on I_t only indirectly through the current posterior mean $E(\theta \mid y^{t-1})$. Although such a simplification may not always be possible, it motivates an approximate solution strategy. We replace $d_t(I_t)$, which is allowed to depend on the full history at time t , by a reduced decision space which allows the decision d_t to depend on I_t only indirectly through some low-dimensional summary $S_t(I_t)$. Additionally, we consider a finite grid over possible values of S_t . Effectively, this means considering a finite discrete S_t . To simplify notation we will write $S_t = j$ to indicate that the value of S_t falls in the j th grid cell, with the understanding that the grid would typically be two- or three-dimensional. Also, we shall write $U_t(d_t, j)$ for the approximate evaluation of $U_t(d_t, d^{t-1}, y^{t-1})$ if $S_t(d^{t-1}, y^{t-1}) = j$. This notation is meaningful since the numerical integration scheme, details of which are described below in Section 4.3, depends on (d^{t-1}, y^{t-1}) only indirectly through S_t . For each cell (t, j) on the grid, starting with $t = T$, we report the expected utility $U_t(d_t, j)$ under alternative decisions d_t , the optimal strategy $d_t^*(S_t = j)$ and its value $U_t^*(j) = U_t(d_t^*, j)$. To compute $U_t(\cdot)$ we use the already tabulated values for $U_{t+1}^*(\cdot)$ to evaluate (3). The number of grid cells (t, j) remains constant over t , thus avoiding the progressively increasing number of possible scenarios which we would have to consider in an unconstrained backward induction. The remaining problem is to evaluate the integral expressions (2) and (3) required for $U_t(\cdot)$.

A critical step in the proposed algorithm is the choice of the summary statistic S_t . Desirable characteristics of a good statistic are as follows. The statistic should be a good approximation of a sufficient statistic for the posterior predictive distribution. A good candidate for S_t are estimators for parameters and functions of parameters that feature prominently in the utility function. Finally, to be practicable the statistic should be low dimensional, say at most three-dimensional.

4.3 Forward Simulation

To evaluate expected utility integrals we use forward simulation. Recall that we only consider decisions related to stopping or choice among finitely many probability models, and with a finite horizon T . Thus we can always generate *all* possibly observed data. Let $p(y|\theta)$ denote the appropriate probability model. We generate M possible experiments $\omega_i = (\theta_i, y_i^T)$, $i = 1, \dots, M$, using the prior probability model to generate $\theta_i \sim p(\theta)$, and $p(y | \theta)$ to generate $y_i^T \sim p(y^T | \theta_i)$. To simplify exposition, in the following description of forward simulation we focus on the special case of sequential sampling only. In this case $p(y | \theta)$ is the probability model which generates responses for all periods, $t = 1, \dots, T$, although for a specific decision we will only get to observe a subset of these. Extension beyond sequential sampling to decisions which select from alternative probability models is straightforward.

For each simulated experiment ω_i we record $S_{ti} = S_t(d^{t-1}, y_i^{t-1})$ at $t = 1, \dots, T$, assuming continuation decisions d . Each experiment corresponds to a trajectory on a (t, S_t) grid. This is illustrated in Figure 1. Starting with the last period, T , we can now approximately evaluate integrals $U_t(\cdot)$ as a sample average. For the j -th cell on the (T, S_T) grid, denote with A_{Tj} the subset of indices $i \in \{1, \dots, M\}$ corresponding to trajectories which terminate in that cell. Let $M_{Tj} = |A_{Tj}|$ denote the number of indices in A_j . Replacing the integral in (2) by a sample average we propose using

$$\hat{U}_T(d_T, S_T = j) = \frac{1}{M_{Tj}} \sum_{i \in A_{Tj}} u(d_T, y_i^T, \theta_i) \quad (5)$$

as an approximate evaluation of $U_T(d_T, d^{T-1}, y^{T-1})$ for all $I_T = \{d^{T-1}, y^{T-1}\}$ with $S_T(I_T)$ falling within the j th cell. Note that at time T continuation is not possible because of the finite horizon. Having recorded $\hat{U}_T(d_T, j)$ we can find the optimal decision $d_T^*(S_T = j)$ for each grid cell, and the corresponding value $\hat{U}_T^*(j) = \hat{U}_T(d_T^*, j)$. From here we proceed similarly for periods $t = T-1, \dots, 1$. At each step t we approximate the expected utility of continuation as

$$\hat{U}_t(d_t, S_t = j) = \frac{1}{M_{tj}} \sum_{i \in A_{tj}} \hat{U}_{t+1}^*(S_{t+1, i}), \quad (6)$$

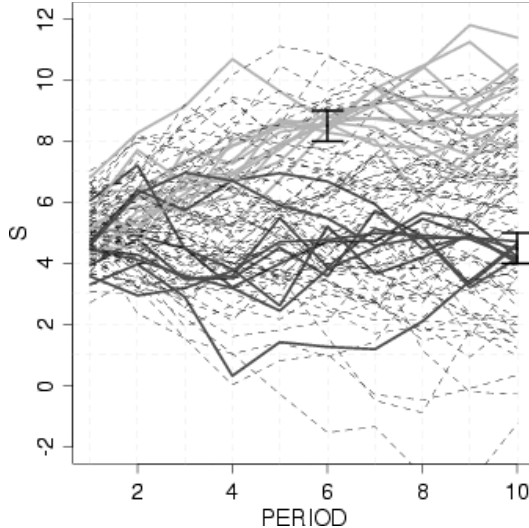


Fig. 1. Trajectories of simulated experiments on a grid over (t, S_t) . The trajectories passing through grid cells $(t = 10, 4 \leq S < 5)$ are shown as solid lines. Trajectories passing through $(t = 6, 8 \leq S < 9)$ are marked as unbroken grey lines. Other simulations are shown as dashed grey lines. The simulations are a stylized example for illustration.

with A_{tj} defined as the subset of size M_{tj} of all simulated experiments whose trajectories at time t pass through cell j , and d_t being the action corresponding to continuation. For any decision d_t which involves stopping at time t we use

$$\hat{U}_t(d_t, S_t = j) = \frac{1}{M_{tj}} \sum_{i \in A_{tj}} u(d_t, y_i^t, \theta_i). \quad (7)$$

We can now find the optimal decision $d_t^*(S_t = j)$ for each grid cell, and the corresponding expected utilities $\hat{U}_t^*(j) = \hat{U}_t(d_t^*, j)$. At the end of the recursion, at time $t = 1$, we are left with the optimal decision d_1^* for the first period. Note that (6) defines the numerical evaluation of (3). Equation (7) provides an alternative expression for the special case of sequential sampling and decisions involving stopping at time t , i.e., decisions which do not require backward induction. The selection of the subset A_{tj} is the simulation equivalent of conditioning on the history S_{ti} . In other words, we implement conditioning on S_{ti} by prior simulation (forward simulation) followed by subset selection. A practical problem occurs when A_{tj} is empty, i.e., when we wish to condition on a subset $S_{ti} = j$, but do not find any (or not many) forward simulations that pass through the selected cell. In that case we propose to re-launch the forward simulation, using the posterior predictive distribution conditional on the partially observed data instead of the prior predictive distribution. Consider, for example, a trial that ends up with $S_3 = 0$ in Figure 1. There are no saved forward simulations i that match $S_{3i} = 0$. We could now generate a new set of forward simulations starting in $S_t = 0$.

A minor variation of the algorithm described here allows using a grid of S_t

values only instead of on (t, S_t) . The problem is that without t , there is no natural start for the backward induction. This problem can be circumvented by using an iterative scheme. Without loss of generality assume $d = 0$ corresponds to continuation, and all other decisions involve stopping. Start out with an initial guess $\hat{U}^o(d = 0, j)$, $j = 1, \dots, J$, for the expected utilities $U_t(d_t = 0, S_t = j)$ under continuation. We will now record expected utilities on a grid over S_t only. Therefore we drop the index t on $\hat{U}^o(d, j)$. To evaluate $\hat{U}^o(d, j)$, $d \neq 0$, use an appropriate modification of (7):

$$\hat{U}^o(d, j) = \frac{1}{M_j} \sum_{i \in A_j} u(d, y_i^t, \theta_i), \quad d \neq 0,$$

where A_j is the set of all indices i with $S_{ti} = j$ for some t . Analogous to the above discussion, let $d^{*o}(j) = \arg \max_d \hat{U}^o(d, j)$ and $\hat{U}^{*o}(j) = \hat{U}^o[d^{*o}(j), j]$ denote the optimal decision and expected utility under d^{*o} . We use an iterative scheme to update \hat{U} . Scan over all grid cells $j = 1, \dots, J$, and replace $\hat{U}^o(d = 0, j)$ by

$$\hat{U}^1(d = 0, j) \equiv \frac{1}{M_j} \sum_{i \in A_j} \hat{U}^{*o}(S_{t+1, i}).$$

For $d \neq 0$ the estimates remain unchanged, $\hat{U}^1(d, j) = \hat{U}^o(d, j)$. Again, set $d^{*1}(j) = \arg \max_d \hat{U}^1(d, j)$ and define $\hat{U}^{*1}(j) = \hat{U}^1[d^{*1}(j), j]$. Repeat the process until updating leaves all decisions unchanged, i.e., $d^{*k}(j) = d^{*k-1}(j)$, $\forall j$.

The approach is particularly attractive if S_t already includes some summary that is closely related to t , for example posterior variance of some (function) of parameter of interest. In this case, there is a natural sequence to update the table. Proceeding from large to smaller posterior variances is almost equivalent to the recursion over time. In practical implementations we expect around 10 iterations to suffice.

5 Examples

Example 1 (Berger 1985, chapter 7).

For illustration we consider an example with an analytically known optimal decision. Assume $y_t \sim \text{Bern}(\theta)$, $t = 1, \dots, T$, is a sequential sample from a Bernoulli distribution, with a prior distribution $p(\theta = 0.4) = p(\theta = 0.6) = 0.5$. Consider the decision problem of choosing between $H_1: \theta = 0.4$ versus $H_2: \theta = 0.6$. After each observation, possible decisions d_t are to terminate and decide for H_1 ($d_t = 1$); terminate and decide for H_2 ($d_t = 2$); or to continue sampling ($d_t = 0$). Let N be the observed stopping time, i.e., $N = \min\{t : d_t \neq 0\}$. Let $d = (d_1, \dots, d_N)$ and assume a “0-K” decision loss and a linear sampling cost

of $c = 1$ per observation.

$$-u(d, Y, \theta) = N + \begin{cases} 0 & \text{if } (\theta = 0.4, d_N = 1) \text{ or } (\theta = 0.6, d_N = 2) \\ K & \text{if } (\theta = 0.4, d_N = 2) \text{ or } (\theta = 0.6, d_N = 1) \end{cases}$$

Let $x_t = \sum_{i=1}^t y_i$. It can be shown (Berger 1985, chapter 7) that the Bayes sequential decision rule d_B stops sampling for the first t for which $|2x_t - t| = k$, where k is some integer depending on K . For example, for $K = 100$ the cutoff is $k = 4$. Conditional on stopping at time τ , the optimal decision is the Bayes decision rule given y^τ , i.e., $d_\tau = 2$ if $x_\tau/\tau > 1/2$.

Implementing the proposed simulation-based algorithm we represent the pattern of information as the pair $(t, p_t = x_t/t)$, and consider a horizon of $T = 50$ periods. Since the known optimal rule d_B can be written in terms of (t, p_t) we expect the numerical solution to approximately reproduce d_B . We simulated $M = 1000$ experiments and proceeded as described in Section 4. Figure 2 plots the estimated expected utilities under alternative decisions, on a (t, p_t) grid. We discretized p_t on a grid of size 50, resulting in a 50×50 grid for (t, p_t) . Evaluation of the expected utilities $U_t(d_t = 1, p_t)$ and $U_t(d_t = 2, p_t)$ requires no backward induction. We use appropriate summaries of the forward simulation as in (7) to evaluate them. In this case, since the expected utilities depend on I_t only through the chosen summary statistic we could analytically evaluate the expected utilities exactly. Evaluation of $U_t(d_t = 0, p_t)$ is done by backward induction, starting at $t = T$.

Figure 3 shows the corresponding estimated optimal decisions $d_t^*(t, p_t)$. For comparison we show the exact Bayes sequential decision rule. The differences between the estimated and the analytical solution around the boundary are due to numerical errors in evaluating the Monte Carlo averages for the expected utility evaluations. The differences close to $N = 50$ are due to the upper bound on N in the simulation. Enforcing monotonicity of the decision rule can remove almost all discrepancies. For a given value N , it is reasonable to enforce that for increasing values of p_t the decision rule change from $d = 1$ to $d = 0$ to $d = 2$. The clever use of such ad-hoc rules to adjust the estimated optimal decision rule is important, but naturally highly problem-specific.

Example 2 (continued from Section 2.)

Let $n_1(d)$ denote the number of patients recruited in the dose-finding trial under decision d . Let $A = A(d)$ denote the event that $d = (d_1, \dots, d_T)$ calls for a pivotal trial, i.e., $d_t = D2$, for some $t \leq T$. Let $y = y^t$ denote the data observed in the dose-finding phase. Since the decision rule is sequential, A depends on y implicitly through d . If a pivotal trial is initiated, let $n_2(d, y)$ denote the number of patients included in the pivotal trial. The sample size of

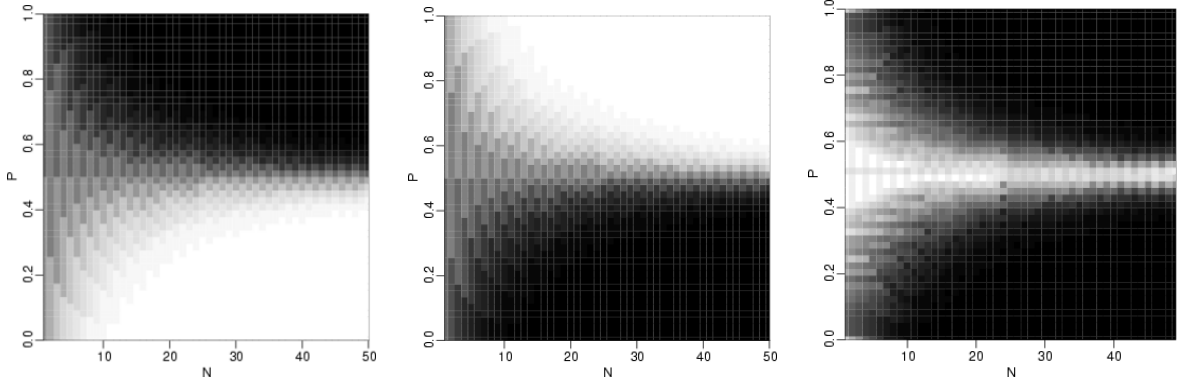


Fig. 2. Example 1. Estimated loss $-\hat{U}_t(d_t = 1, p_t)$ (left panel), $-\hat{U}_t(d_t = 2, p_t)$ (center panel), $-\hat{U}_t(d_t = 0, p_t)$ (right panel).

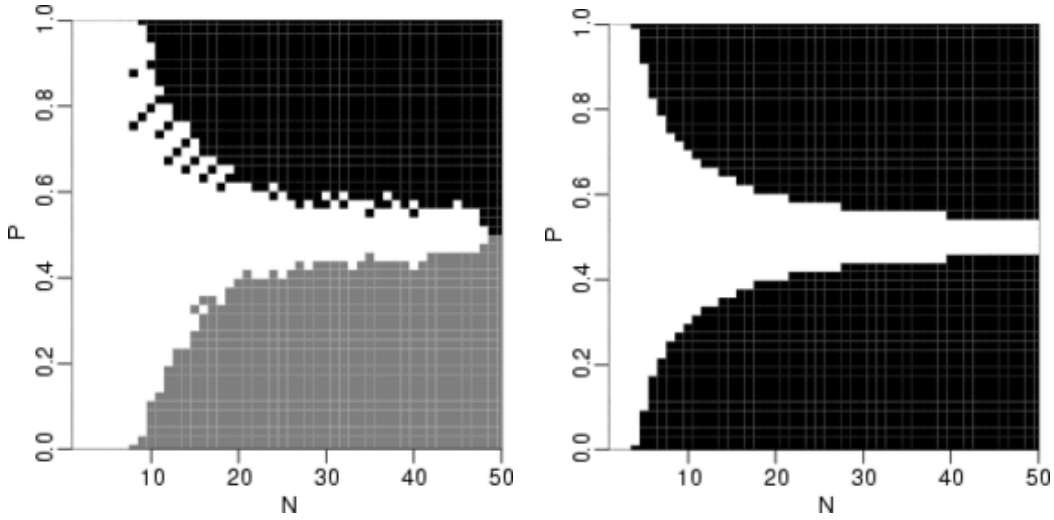


Fig. 3. Example 1. Recommended actions $d_t^*(t, p_t)$ (left panel), and analytic solution $d_t^*(t, I_t)$ (right panel). The grey shades in the left panel indicate $d_t^* = 1, 2$ and 0 , where $d_t^* = 1$ is grey, $d_t^* = 2$ is black, and $d_t^* = 0$ is white. The plot in the right panel indicates $d_t^* \in \{1, 2\}$ (black) versus $d_t^* = 0$ (white).

the pivotal trial can depend on the data collected in the first phase, thus the dependence on y . Let y^p denote the data collected in the pivotal trial, and let $B = B(y^p)$ denote the event that the pivotal trial concludes that the drug is effective. The utility function discussed in Section 2 is formally defined as

$$u(d, y, y^p, \theta) = \begin{cases} -c_1 n_1(d) & \text{if } A^c \\ -c_1 \{n_1(d) + n_2(d, y)\} & \text{if } A \cap B^c \\ -c_1 \{n_1(d) + n_2(d, y)\} + c_2 \bar{\Delta}(y, y^p) & \text{if } A \cap B, \end{cases}$$

and $u(d, y, \theta)$ replaces $\bar{\Delta}(y, y^p)$ by $p(B | y, \theta, A) E(\bar{\Delta}(y, y^p) | y, \theta, A, B)$ and combines the last two cases. Here $\bar{\Delta}(y, y^p)$ denotes the posterior mean on the advantage over placebo, Δ , conditional on the data at the completion of the confirmatory phase. Besides the constant sampling cost, the utility function is

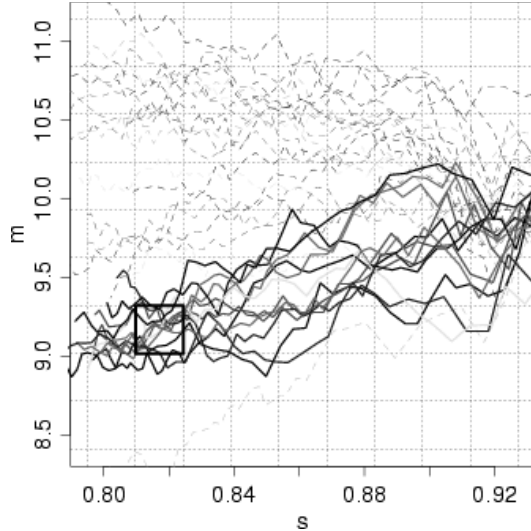


Fig. 4. Example 2. Trajectories on a grid over the bivariate summary statistic $S_t = (m_t, s_t)$. Consider, for example, the grid cell j highlighted with a bold outline, around $s = 0.82$ and $m = 9.0$. To compute expected utilities \hat{U}_j we use an average over all simulations which pass through this grid cell. The corresponding trajectories are plotted in bold.

determined only by the advantage Δ of treatment over placebo. This motivates considering S_t to be a summary of the current inference on Δ . Let $m_t = E(\Delta|y^t)$, $s_t = SD(\Delta|y^t)$ denote posterior mean and standard deviation of Δ . We use $S_t = (m_t, s_t)$ as summary statistic in the constrained backward induction. We simulated $M = 1000$ experiments and record $S_t = (m_t, s_t)$ for each week t . Figure 4 shows some of the simulated trajectories. Based on this forward simulation we computed $\hat{U}(d, S_t = j)$ for j on a 20×20 grid over S_t . Evaluating the expected utilities we included an additional step to reduce numerical uncertainties due to finite simulations. Namely, after computing estimates \hat{U}_j as described above, we fit a smooth surface \tilde{U}_j through the pairs (j, \hat{U}_j) . The smooth fit \tilde{U}_j formalizes “borrowing strength” across simulations for neighboring grid cells, and allows for interpolating for grid cells with few or no simulations (with the usual caveat about extrapolation beyond the range of the data). This is shown in Figure 5.

6 Conclusion

We proposed a simulation-based method for solving sequential design problems. The method is broadly applicable in that only minimal constraints are assumed for the probability model and the utility function. Essentially, the method applies to any model which allows posterior Markov chain Monte Carlo simulation. The only constraint on the utility function is that it must be possible to evaluate utility $u(d, y, \theta)$ for a given simulated experiment (θ, y)

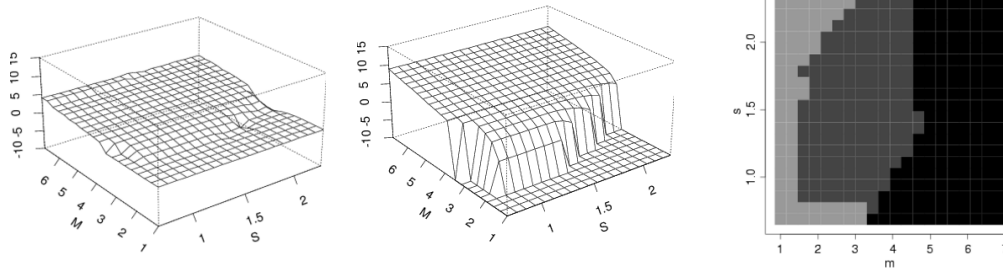


Fig. 5. Example 2. Expected utilities on the grid over $S_t = (m_t, s_t)$. The left panel shows $\hat{U}(d = D1, S_t = j)$. The center panel shows $\hat{U}(d = D2, S_t = j)$. The right panel shows $d^*(j)$ with light grey indicating $d^*(j) = D0$, grey for $d^*(j) = D1$ and black for $d^*(j) = D2$.

for any particular decision d . The space of decisions is limited to those that do not influence the probability model beyond identifying a subset of possible responses. The allowable decisions include those that terminate an experiment at a particular time, and also those that select from alternative probability models, e.g., probability models associated with alternative treatments.

The solution we propose is not exact; it is an approximation with the quality of the approximation depending on the summary statistic S_t , the number M of the forward simulations, and the extent of discretization when defining the grid on S_t .

We developed the algorithm in the context of a sequential decision problem at the conclusion of a dose-finding clinical trial, but the methods apply for any problem which fits into the described framework.

Acknowledgments

Research was supported by NIH/NCI grants R33 CA97534-01 and R01 CA075981.

References

- Amzal, B., Bois, F., Parent, E., Robert, C., 2003. Bayesian optimal design via interacting mcm. Tech. rep., Université Paris-Dauphine/Ceremade.
- Berger, J. O., 1985. Statistical Decision Theory and Bayesian Analysis, 2nd Edition. Springer-Verlag, New York.
- Berry, D., 1993. A case for Bayesianism in clinical trials (with discussion). *Statistics in Medicine* 12, 1377–1404.
- Berry, D., Fristedt, B., 1985. Bandit Problems: Sequential Allocation of Experiments. Chapman and Hall, London.

- Berry, D. A., Mueller, P., Grieve, A. P., Smith, M., Parke, T., Blazek, R., Mitchard, N., Krams, M., 2001. Adaptive bayesian designs for dose-ranging drug trials. In: Gatsonis, C., Kass, R. E., Carlin, B., Carriquiry, A., Gelman, A., Verdinelli, I., West, M. (Eds.), *Case Studies in Bayesian Statistics*, Volume V. *Lecture Notes in Statistics*. Springer-Verlag, New York.
- Berry, D. A., Stangl, D. K. (Eds.), 1996. *Bayesian Biostatistics*. Vol. 151 of *Statistics: Textbooks and Monographs*. Marcel Dekker, New York.
- Brockwell, A. E., Kadane, J. B., 2003. A gridding method for bayesian sequential decision problems. *Journal of Computational and Graphical Statistics* 12, 566–584.
- Carlin, B., Kadane, J., Gelfand, A., 1998. Approaches for optimal sequential decision analysis in clinical trials. *Biometrics* 54, 964–975.
- Chaloner, K., Verdinelli, I., 1995. Bayesian experimental design: a review. *Statistical Science* 10, 273–304.
- Christen, J. A., Nakamura, M., 2003. Sequential stopping rules for species accumulation. *Journal of Agricultural, Biological and Environmental Statistics* 8, 184–195.
- Farr-Jones, S., 2001. Better statistics. *BioCentury* 9 (12), 1–6.
- Hsiao, C.-F., Clayton, M. K., 2001. Bayes discrete sequential boundaries for clinical trials. *Communications in Statistics: Theory and Methods* 30, 1381–1395.
- Lan, K. K. G., DeMets, D. L., 1983. Bayesian approaches to randomized trials. *Biometrika* 157, 357–416.
- Lewis, R., Berry, D., 1994. Group sequential clinical trials: A classical evaluation of bayesian decision-theoretic designs. *Journal of the American Statistical Association* 89, 1528–1534.
- Malakoff, D., 1999. Bayes offers a ‘new’ way to make sense of numbers. *Science* 286, 1460–1464.
- Müller, P., Sansó, B., DeIorio, M., 2004. Optimal bayesian design by inhomogeneous markov chain simulation. *Journal of the American Statistical Association* 99, 788–798.
- Spiegelhalter, D. J., Freedman, L. S., Parmar, M. K. B., 1994. Bayesian approaches to randomized trials. *Journal of the Royal Statistical Society, Series A, General* 157, 357–387.
- Thall, P., Russell, K., 1998. A strategy for dose-finding and safety monitoring based on efficacy and adverse outcomes in phase I/II clinical trials. *Biometrics* 54, 251–64.
- Thall, P. F., Simon, R. M., Estey, E. H., 1995. Bayesian sequential monitoring designs for single-arm clinical trials with multiple outcomes. *Statistics in Medicine* 14, 357–379.
- Verdinelli, I., 1992. Advances in Bayesian experimental design. In: *Bayesian Statistics 4. Proceedings of the Fourth Valencia International Meeting*. pp. 467–475.
- Vlachos, P., Gelfand, A., 1998. Nonparametric bayesian group sequential design. In: Dey, D., Müller, P., Sinha, D. (Eds.), *Practical Nonparametric and*

- Semiparametric Bayesian Statistics. Springer-Verlag, New York, pp. 115–132.
- Wakefield, J., 1994. An expected loss approach to the design of dosage regimens via sampling-based methods. *The Statistician* 43, 13–29.
- Whitehead, J., Brunier, H., 1995. Bayesian decision procedures for dose determining experiments. *Statistics in Medicine* 14, 885–893.
- Whitehead, J., Williamson, D., 1998. Bayesian decision procedures based on logistic regression models for dose-finding studies. *Journal of Biopharmaceutical Statistics* 8, 445–467.