A SAMPLE-SIZE OPTIMAL BAYESIAN PROCEDURE FOR SEQUENTIAL PHARMACEUTICAL TRIALS

BY

NOEL CRESSIE
JONATHAN BIELE

TECHNICAL REPORT NO. 451
MARCH 5, 1992

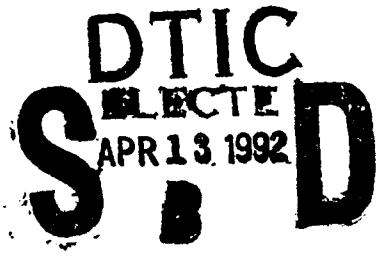
PREPARED UNDER CONTRACT
N00014-92-J-1264 (NR-042-267)
FOR THE OFFICE OF NAVAL RESEARCH

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA

A SAMPLE-SIZE OPTIMAL BAYESIAN PROCEDURE FOR SEQUENTIAL PHARMACEUTICAL TRIALS

BY

NOEL CRESSIE
JONATHAN BIELE



TECHNICAL REPORT NO. 451
MARCH 5, 1992

Prepared Under Contract
N00014-92-J-1264 (NR-042-267)
For the Office of Naval Research

Herbert Solomon, Project Director

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA

A Sample-Size Optimal Bayesian Procedure for Sequential Pharmaceutical Trials

by

Noel Cressie and *Jonathan Biele*

SUMMARY

Consider a pharmaceutical trial where the consequences of different decisions are expressed on a financial scale. The efficacy of the new drug under consideration has a prior distribution obtained from the underlying biological process, animal experiments, clinical experience, and so forth. In an important paper, Berry and Ho (1988) show how these components are used to establish an optimal (Bayes) sequential procedure, assuming a known constant sample size at each decision point. We show in this article how it is also possible to optimize with respect to the sample-size rule. This last component of the design, which is missing from most sequential procedures, has the potential to yield considerably larger expected net gains.

Keywords: Backward induction; clinical trials; group-sequential procedure; optimal stopping; predictive distribution; VPRT.

1. Introduction.

Berry and Ho (1988) argue cogently for a Bayesian decision-theoretic approach to the design and analysis of pharmaceutical trials. They (and we) address the question of when a drug development program should be stopped in the face of evidence that suggests insufficient drug efficacy. The purpose of this article is not to enter into a debate about Bayesian versus frequentist methods for the design and analysis of sequential clinical trials. but rather to implement *fully* the Bayesian technology. Breslow (1990), and the discussion following his article, contrasts the two methods.

In the pharmaceutical setting, it is often reasonable to assume that the consequences related to the objectives of the trial can be converted to a financial scale. In decision-theoretic terms (e.g., Raiffa and Schlaifer. 1961: DeGroot, 1970), we assume that the drug company is able to construct a realistic gain (or loss) function. Such a function might come from a cost-benefit analysis. What we are proposing in this paper is an optimal way to conduct the clinical-trials part of the drug development program, but the same ideas could be used in later parts of the program (where one must choose between several possibilities, one of which is to put off the decision and obtain more data to help make a decision at the next time point). In both circumstances, the Bayesian approach provides the decision-maker with not only an optimal stopping and terminal-decision rule, but also with an optimal sample-size rule.

Consider sequential decision procedures that allow sampling in other than a one-at-a-time manner. Although group-sequential sampling was suggested by Wald (1947, pp. 101-104), it was not until much later that interest in these techhniques (with constant group size) was initiated; see, for example, Ghosh (1970, p. 224), Pocock (1977), and Gupta and Miescke (1984). In all these studies, at each decision point the sample size is either assumed to be known in some way or is chosen in an *ad hoc* (perhaps stochastic) manner. Here we focus attention on the *optimal* choice of sample sizes as a component of a sequential decision procedure for pharmaceutical trials. In the related problem of sequential hypothesis testing, Whittle (1965), Ehrenfeld (1972), and Spahn and Ehrenfeld (1974) considered special cases; Cressie and Morgan (1988) gave a general formulation for two simple hypotheses.

Section 2 defines the pharmaceutical trial under consideration in this article and establishes the necessary notation. In Section 3, the optimal (Bayes) sequential procedure of Berry and Ho (1988) is re-established and then improved by further optimizing over sample

sizes. Two illustrative examples are given in Section 4; in one, the benefits of sample-size optimization are seen to be substantial. Section 5 contains concluding remarks.

2. Sequential Pharmaceutical Trial.

We now give a brief description of the problem considered by Berry and Ho (1988) and show how our proposed sample-size optimization maximizes the expected net gain (ENG) over all sequential procedures. The purpose of the sequential trial is to test an experimental treatment compared to a control; at each time the same number of data is collected from the treatment population as from the control population.

The data from the control population is assumed to be independent and identically distributed (i.i.d.) $\text{Gau}(\mu_{\text{ctl}}, \sigma^2/2)$ and that from the treatment population is assumed to be i.i.d. $\text{Gau}(\mu_{\text{trt}}, \sigma^2/2)$, distribution with mean μ and variance σ^2. (Here, $\text{Gau}(a, b^2)$ denotes a Gaussian, or normal, distribution with mean a and variance b^2.) Define

$$\delta \equiv \mu_{\text{trt}} - \mu_{\text{ctl}} \,. \tag{2.1}$$

We wish to test sequentially the composite hypotheses

$$H_0 : \delta \leq \delta_0 \quad \text{versus} \quad H_1 : \delta > \delta_0 \,, \tag{2.2}$$

where without loss of generality we take $\delta_0 = 0$. Suppose that a sample of size n_t is requested at time t and obtained at time $(t+1)$, for testing (2.2). Let $\bar{x}_{\text{trt},t}$ and $\bar{x}_{\text{ctl},t}$ denote the corresponding sample means. Then

$$w_t \equiv \bar{x}_{\text{trt},t} - \bar{x}_{\text{ctl},t} \sim \text{Gau}(\delta, \sigma^2/n_t); \quad t = 0, 1, \cdots, T-1 \,, \tag{2.3}$$

where T denotes the *truncation time* by which a decision about H_0 or H_1 *must* be made. Clearly, the statistics $w_1, w_2, \cdots$ are sufficient for inference on δ.

Define

$$\bar{w}_{t+1} \equiv \left(\sum_{i=0}^{t} n_i w_i \right) \Big/ \left(\sum_{i=0}^{t} n_i \right); \quad t = 0, 1, \cdots, T-1 \,, \tag{2.4}$$

which is the sample mean of all differences between treatment and control up to time $(t+1)$. It will be seen that (2.4) is the basis of inference on δ made at time $(t+1)$. Group-sequential testing usually assumes $n_0 = n_1 = \cdots = n_{T-1} = k$, where k is a predetermined sample size. Berry and Ho (1988) make this assumption, but in this article we demonstrate

that considerable gains are possible by optimizing with respect to sample-size choice at *each* time point.

We take a Bayesian approach to the sequential analysis of pharmaceutical trials. That is, the consequences of right or wrong decisions about δ can be distilled into a *gain function.* sampling has a *cost* associated with it, and we have a *prior* opinion about δ that is here expressed as,

$$\delta \sim \text{Gau}(\nu_0, \tau_0^2) \,. \tag{2.5}$$

In line with (2.4), define

$$\overline{\omega}_0 \equiv \nu_0 \,. \tag{2.6}$$

A Bayes procedure maximizes the expected gain net of sampling costs or expected net gain (ENG). It is usually defined as a decision procedure that minimizes risk or expected net loss; however, an ENG-maximizing procedure is easily seen to be Bayes by defining the loss function to be the negative of the specified gain function.

The gain function g for a sequential pharmaceutical trial is a function of the value of δ and which of H_0 (i.e., $\Delta_t = 0$) or H_1 (i.e., $\Delta_t = 1$) is specified at time t. Potentially, the function itself could also vary with t. In what is to follow, we use Berry and Ho's (1988) specification:

$$g(\delta, \Delta_t) = \begin{cases} 0, & \text{if} \quad \Delta_t = 0 \\ -L, & \text{if} \quad \Delta_t = 1 \quad \text{and} \quad \delta \leq 0 \\ K\delta, & \text{if} \quad \Delta_t = 1 \quad \text{and} \quad \delta > 0, \end{cases} \tag{2.7}$$

where L and K are given positive constants. Berry and Ho (1988, p. 222) provide a justification for this gain function (they actually specify a loss function) in the context of pharmaceutical trials.

Sampling is costly; it is assumed that each observation in the sample costs one unit so that at time t the cost of the next treatment sample and control sample is $2n_t$. At each time point t, we are faced with a decision whether to stop the trial or not. A well-informed decision will weight up the gains of stopping versus the net gains to be expected from continuing to sample (after debiting sampling costs). In the event that the trial is continued to the time point $(t+1)$, how many observations should be sampled? We shall show that there is an optimal number n_t^*, of treatment and control samples, that depends only on the posterior distribution of δ and the exogeneous constants K and L of the gain structure.

In general, a sequential procedure can be completely specified through a stopping rule $S \equiv \{S_t : t = 0.1, \cdots, T\}$, a terminal-decision rule $\Delta \equiv \{\Delta_t : t = 0.1, \cdots, T\}$ and a

sample-size rule $n \equiv \{n_t : t = 0, 1, \cdots, T\}$, all of which are functions of the data. (For the problem we consider here, S_t, Δ_t, and n_t will be functions of $\bar{w}_t$ given by (2.4).) When $S_t = 1$, the procedure stops at time t; otherwise $S_t = 0$, in which case sampling continues. Suppose the procedure stops at time t: When $\Delta_t = 0$, the hypothesis H_0 is chosen; otherwise $\Delta_t = 1$, in which case H_1 is chosen. The function n_t is nonnegative-integer valued; should sampling continue at time t, it specifies the sample size to be taken between time t and time $(t + 1)$. In much of the sequential-analysis literature, the sample-size rule n is not mentioned, since it is usually assumed that one sample (or a specified constant number of samples) is taken at each time point.

Now, since a decision must be made by (truncation) time T, $S_T \equiv 1$ and $n_T \equiv 0$. Berry and Ho (1988) make further restrictions on their sequential procedure. They argue that one should stop investigating an ineffective treatment, but obtain as much information as possible on a good one. That is, their procedure is one-sided and the combination $S_t = 1$, $\Delta_t = 1$ never happens for $t = 0, 1, \cdots, T - 1$. (In Section 5, we show that the one-sided restriction can be relaxed.) Also, they assume $n_t = k$, a prespecified integer, for $t = 0, 1, \cdots, T - 1$. For most of what is to follow, we retain all aspects of Berry and Ho's procedure *except* for the restriction on the sample-size rule. We shall demonstrate that there can be considerable benefits to be had from obtaining optimal sample sizes.

3. Bayes sequential procedures.

To simplify notation, we shall call the Bayes sequential procedure proposed by Berry and Ho (1988) the B-procedure. Upon relaxing the sample-size restriction in the B-procedure, the ENG-maximal (Bayes) sequential procedure can be obtained, which we call the B^*-procedure. By definition, since the B^*-procedure is optimal over a less restrictive set of procedures than the B-procedure, the former's ENG is larger than or equal to the latter's. We shall demonstrate in Section 4 that the increase in the B^*-procedure's ENG can sometimes be substantial.

First, we shall describe the B-procedure, generalized slightly here to handle any prespecified sample-size rule n. Berry and Ho (1988) show it to be Bayes using backward induction (e.g., DeGroot, 1970, Section 12.4). Assuming the ENG of the sequential procedure at time $(t+1)$ is known, an optimal procedure at time t can be obtained by comparing the ENGs of stopping and choosing various hypotheses with the ENG of continuing to sample (which involves an average over the known ENG at time $t + 1$). The optimal stopping, terminal-decision, and sample-size rules at time t are obtained by taking whatever action

achieves the maximum of all the competing ENGs.

From (2.5), the prior probability that H_0 is true is

$$p_0 = \Phi(-\nu_0/\tau_0)\,, \tag{3.1}$$

where

$$\Phi(x) \equiv \int_{-\infty}^{x} (2\pi)^{-1/2} \exp(-s^2/2) ds\,. \tag{3.2}$$

Let ν_t and τ_t^2 be the posterior mean and variance respectively of δ at time t. Then, it is straightforward to show that

$$\nu_t = \{\nu_0\sigma^2 + \tilde{\omega}_t(\sum_{i=0}^{t-1} n_i)\tau_0^2\}/\{\sigma^2 + (\sum_{i=0}^{t-1} n_i)\tau_0^2\} \tag{3.3}$$

$$\tau_t^2 = \sigma^2\tau_0^2/\{\sigma^2 + (\sum_{i=0}^{t-1} n_i)\tau_0^2\}\,. \tag{3.4}$$

Thus, the posterior probability that H_0 is true is

$$p_t = \Phi(s_t)\,, \tag{3.5}$$

where

$$s_t \equiv -\nu_t/\tau_t\,; \quad t = 0, 1, \cdots, T\,. \tag{3.6}$$

Since Φ given by (3.2) is a one-to-one mapping, there is no loss of information from considering the pair of posterior parameters p_t, τ_t (or s_t, τ_t), rather than ν_t, τ_t.

One further result that is needed, to obtain the Bayes procedure. is the preditive (or marginal) distribution of w_{t+1} given p_t, τ_t. Upon integrating out the conditional density of δ given p_t, τ_t, we obtain the *predictive density* q_{t+1} of w_{t+1}; that is.

$$w_{t+1}|p_t, \tau_t \sim \text{Gau}(-\Phi^{-1}(p_t)\tau_t, \tau_t^2 + (\sigma^2/n_t))\,; \quad t = 0, \cdots. T-1\,. \tag{3.7}$$

Following Berry and Ho (1988), the ENG for stopping at time t and choosing H_1 is (from (2.7)):

$$\begin{aligned} h(p_t, \tau_t) &= -Lp_t + E(K\delta|\delta > 0, p_t, \tau_t)(1-p_t) \\ &= -Lp_t + K\tau_t\{\phi(\Phi^{-1}(p_t)) + \Phi^{-1}(p_t)(1-p_t)\}\,; \quad t = 0.1. \cdots. T\,. \end{aligned} \tag{3.8}$$

where

$$\phi(x) \equiv \Phi'(x) = (2\pi)^{-1/2} \exp(-x^2/2)\,. \tag{3.9}$$

Notice that the data and prior appear in (3.8) via the posterior parameters given by (3.3) through (3.6). Similarly, from (2.7), the ENG of stopping at time t and choosing H_0 is:

$$z(p_t, \tau_t) \equiv 0 \; ; \quad t = 0, 1, \cdots, T \; . \tag{3.10}$$

Let V_t^T denote the expected net gain (ENG), at time t, of the Bayes procedure whose truncation time is T. At $t = T$, H_0 or H_1 must be chosen. The optimal terminal-decision rule is, from (3.8) and (3.10),

$$\Delta_T = \begin{cases} 0, & \text{if} \quad h(p_T, \tau_T) \leq z(p_T, \tau_T) \\ 1, & \text{otherwise,} \end{cases} \tag{3.11}$$

with ENG

$$V_T^T(p_T, \tau_T) = \max\{z(p_T, \tau_T), h(p_T, \tau_T)\} \; . \tag{3.12}$$

At $t = T - 1$, the options are either stopping or continuing the trial by sampling according to the sample-size rule n_{T-1}. Conditional on the posterior parameters p_{T-1}, τ_{T-1} (and the sample size rule), the ENG of continuing is

$$g_{T-1}(p_{T-1}, \tau_{T-1}, n_{T-1}) \equiv E_{F_{T-1}}\{V_T^T(p_T, \tau_T) | p_{T-1}, \tau_{T-1}, n_{T-1}\} - 2n_{T-1} \; , \tag{3.13}$$

where $E_{F_{T-1}}$ denotes expectation with respect to the predictive distribution of w_{T-1} given by (3.7). Notice that the sampling cost $2n_{T-1}$ is subtracted from the expected gain of continuing to sample. Since V_T^T is given by (3.12), the quantity (3.13) can be computed and compared with the ENG of stopping at time $t = T - 1$.

More generally, at time t, the ENG of continuing the trial to time $(t+1)$ by sampling according to n_t is

$$g_t(p_t, \tau_t, n_t) \equiv E_{F_t}\{V_{t+1}^T(p_{t+1}, \tau_{t+1}) | p_t, \tau_t, n_t\} - 2n_t \; ; \quad t = 0, \cdots, T-1 \; , \tag{3.14}$$

where E_{F_t} denotes expectation with respect to the predictive distribution of w_{t+1} given by (3.7). Thus, in order to complete the backward induction for the B-procedure, we need formulas for $\{V_{T-\ell}^T(p_{T-\ell}, \tau_{T-\ell}) : \ell = 1, 2, \cdots, T\}$. (The formula for V_T^T is given by (3.12).)

Recall that, for the purposes of comparison to Berry and Ho's results, we have a one-sided procedure for $t = 0, \cdots, T-1$. Thus, from (3.10) and (3.14), we obtain

$$V_t^T(p_t, \tau_t) = \max\{z(p_t, \tau_t), g_t(p_t, \tau_t, n_t)\} : \quad t = 0, \cdots, T-1 \; . \tag{3.15}$$

so that V_t^T depends on (the expected value of) V_{t+1}^T. Since V_T^T is given by (3.12), the backward induction can proceed to V_{T-1}^T to V_{T-2}^T, and so forth, to V_0^T. In terms of $\{(S_t, \Delta_t, n_t) : t = 0, \cdots, T-1\}$, the B-procedure is given by

$$S_t = \begin{cases} 1, & \text{if } \ g_t(p_t, \tau_t, n_t) \leq z(p_t, \tau_t) \\ 0, & \text{otherwise :} \end{cases} \tag{3.16}$$

$$\Delta_t \equiv 0 \, ; \tag{3.17}$$

and n_t is prespecified ($t = 0, \cdots, T-1$). Recall that, for $t = T$, $S_T \equiv 1$ and Δ_T is given by (3.11).

Having given an optimal stopping rule and an optimal terminal-decision rule, let us now turn our attention to the B^*-procedure, which in addition incorporates an *optimal sample-size rule*. Define the optimal sample size at time t by the value of n_t that maximizes (3.14); specifically,

$$n_t^*(p_t, \tau_t) \equiv \text{argmax}_{n_t \geq 0}\{g_t(p_t, \tau_t, n_t)\}; \quad t = 0, \cdots, T-1 \, , \tag{3.18}$$

where $\text{argmax}_{n \geq 0}\{\cdots\}$ denotes a value of the argument n that achieves the maximum for the expression in braces. If there are several such values, choose the smallest. Upon substitution of $\{n_t^* : t = 0, \cdots, T-1\}$ in place of $\{n_t : t = 0, \cdots, T-1\}$ into $\{S_t : t = 0, \cdots, T\}$ and $\{\Delta_t : t = 0, \cdots, T\}$, we obtain the B^*-procedure. Likewise, from (3.15), its ENG is

$$V_t^{T*}(p_t, \tau_t) \equiv \max\{z(p_t, \tau_t), g_t(p_t, \tau_t, n_t^*)\} \, ; \quad t = 0, \cdots, T-1 \, . \tag{3.19}$$

By definition, $V_t^{T*} \geq V_t^T$. In the examples given in the next section, we show that the B^*-procedure can yield a substantially larger ENG. In order to compute the B-procedure and B^*-procedure from given gain constants K and L, prior parameters ν_0 and τ_0^2, and model parameter σ^2, various analytical results about h and g_t are needed. These can be found in Berry and Ho (1988) and guarantee the existence of break-even values $b_0, \cdots, b_{T-1}$ such that (3.16) can be written as

$$S_t = \begin{cases} 1, & \text{if } \ p_t < b_t \\ 0, & \text{otherwise : } \ t = 0, \cdots, T-1. \end{cases} \tag{3.20}$$

Further, there exists a b_T such that (3.11) can be written as

$$\Delta_T = \begin{cases} 0, & \text{if } \ p_T < b_T \\ 1, & \text{otherwise .} \end{cases} \tag{3.21}$$

Finally, since g_t given by (3.13) is bounded above as a function of n_t, (3.18) is well defined and hence the B^*-procedure can also be given in terms of break-even values $b_0^*, \cdots, b_T^*$.

4. Examples.

This section presents numerical results for two different choices of K, L, τ_0^2, and σ^2. (Notice that specification of the prior mean ν_0 of δ is not needed for performance evaluations, since the optimal sequential procedures can be given in terms of $\{p_t\}$ rather than $\{\bar{w}_t\}$. To implement the procedures on data, ν_0 has to be specified.) Using backward induction to do the computations, means that the time taken to implement the procedure is *linear* in T.

We deliberately chose one example to be that considered by Berry and Ho (1988), in order to validate our results. In this example, their choice of $n_t = 30$ $(t = 0, 1, 2 = T)$ turns out to be close to optimal. The second example was chosen to show that such an *ad hoc* choice of sample size can be far from optimal and that the B^*-procedure can offer considerable improvement.

4.1. Example 1.

Consider the constants specified by Berry and Ho (1988), namely, $K = 5000$, $L = 2000$, $\tau_0^2 = 1$, and $\sigma^2 = 2$. It is instructive to consider initially the case $T = 1$. Figure 1 shows a contour plot of the ENG $g_0(p_0, 1, n_0)$, given by (3.14), as a function of p_0 on the horizontal axis and n_0 on the vertical axis. The optimal n_0^*, as a function of p_0, will follow a ridge, but it is clear that the contour lines are essentially parallel and that Berry and Ho's choice of $n_0 = 30$ is as optimal as any sample size in the range $[20, 50]$.

Figure 1 here

Another way to make the comparison is to superimpose g_0 as a function of p_0, for several chosen values of n_0. Figure 2 shows $g_0(\bullet, 1, n_0)$; $n_0 = 0, 1, 30$, and 100, on the same graph. For most of p_0, $n_0 = 30$ has the largest ENG.

Figure 2 here

Figure 3 shows the optimal sample size n_0^* as a function of p_0 for the cases $T = 1, 2$, and 3. As T increases, initial optimal sample sizes tend to decrease, which is sensible since there are more times at which further samples could be taken. The break-even values are $b_1^* = 0.979, 0.984$, and 0.984, for $T = 1, 2$, and 3, respectively.

Figure 3 here

The ENG $g_0(p_0, 1, n_0^*(p_0))$ is the largest possible ENG associated with the decision to continue sampling; selected values for $T = 1, 2,$ and 3 are given in Table 1. For fixed p_0, these values are increasing in T. Interestingly, the table shows that there is little to gain by specifying a truncation time beyond $T = 2$.

Table 1 here

4.2. Example 2.

Consider the constants $K = 100$, $L = 5000$, $\tau_0^2 = 1$, and $\sigma^2 = 0.5$. these values were chosen from among several combinations because they illustrate clearly the advantage of the B^*-procedure over the B-procedure. To compare with Example 1, consider the case $T = 1$. Figure 4 shows a contour plot of the ENG $g_0(p_0, 1, n_0)$ as a function of p_0 and n_0 (cf. Figure 1). The ridge that traces out the optimal value $n_0^*(p_0)$ as a function of p_0 is rather pronounced and gives values of n_0^* around 5. An *ad hoc* choice of $n_0 = 30$ is far from optimal. To illustrate this point, Figure 5 shows the percentage increase of the ENG $g_0(p_0, 1, n_0^*(p_0))$ compared to the ENG $g_0(p_0, 1, 30)$. Substantial improvements in expected net gain are possible using an optimal choice of sample size.

Figure 5 here

Figure 6 shows the optimal sample size n_0^* as a function of p_0 for the cases $T = 1, 2,$ and 3. The break-even values are $b_1^* = 0.737, 0.754, 0.754$, for $T = 1, 2,$ and 3, respectively.

Figure 6 here

5. Discussion.

The Bayesian sequential approach relies heavily on the specification of a gain structure. In medical trials where human lives are at stake, this will be a difficult (if not impossible) task. However, the classical sequential trial (e.g., Whitehead, 1983) is once removed from this type of specification.

In this article, we have demonstrated that the Bayesian approach to sequential sampling can be extended to include an optimal sample-size rule, which is a function of the posterior distribution. The rule can be calculated *before* any sampling takes place. It can either be adhered to strictly throughout the sequential trial or can be used to obtain an approximately optimal sample size for a group-sequential procedure.

For example, consider the case $T = 1$ in Section 4.1; $n_0 = 20$ yields an ENG that is close to $g_0(p_0, 1, n_0^*(p_0))$ over much of $0 \leq p_0 \leq 1$. Values of n_0^* for larger T tend to be less

than for $T = 1$ and, for fixed $T > 1$, this is also true for $n_1^*, \cdots, n_{T-1}^*$. Thus, a conservative specification of the total sample size needed for the B^*-procedure is $20T$. Alternatively, a group-sequential B-procedure with $n_0 = \cdots = n_{T-1} = 20$ would achieve ENGs close to those of the B^*-procedure. (Berry and Ho's choice of $n_0 = \cdots = n_{T-1} = 30$ would yield a B-procedure with similar ENG properties, however it would need more samples.)

Similar reasoning applied to Section 4.2 yields a conservative specification of the B^*-procedure's total sample size to be $5T$. Or, a group-sequential B-procedure with $n_0 = \cdots = n_{T-1} = 5$ would achieve ENGs close to those of the B^*-procedure. However, a B-procedure with $n_0 = \cdots = n_{T-1} = 30$ would not be appropriate at all.

Although Berry and Ho (1988) chose to carry out a one-sided sequential pharmaceutical trial, it is easy to adapt their approach to handle the two-sided case. In that case, (3.16) becomes

$$S_t = \begin{cases} 1, & \text{if} \quad g_t(p_t, \tau_t, n_t) \leq \max\{z(p_t, \tau_t), h(p_t, \tau_t)\} \\ 0, & \text{otherwise ;} \quad t = 0, \cdots, T-1, \end{cases} \tag{6.1}$$

(3.17) becomes

$$\Delta_t = \begin{cases} 0, & \text{if} \quad h(p_t, \tau_t) \leq z(p_t, \tau_t) \\ 1, & \textit{otherwise:} \quad t = 0, \cdots, T-1. \end{cases} \tag{6.2}$$

(3.15) becomes

$$V_t^T(p_t, \tau_t) = \max\{z(p_t, \tau_t), h(p_t, \tau_t), g_t(p_t, \tau_t, n_t)\} \; ; \quad t = 0, \cdots, T-1 \, , \tag{6.3}$$

and S_T, Δ_T, and V_T^T remain unchanged. To optimize on sample size, again use (3.18) to choose n_t^* and substitute $\{n_t^*\}$ in place of $\{n_t\}$.

Although the probability of type 1 error(α), power(π), and expected total sample size are not considered by Bayes sequential procedures, there is evidence that a superior performance can also be expected in terms of these criteria. In the case of a simple null hypothesis versus a simple alternative hypothesis for testing Gaussian means, Cressie, Biele, and Morgan (1991) demonstrate the B^*-procedure's small α and large π over much of $0 < p_0 < 1$. Moreover, Cressie and Morgan (1992) show that of all sequential procedures with a given α and π, the B^*-procedure minimizes the expected total sample size.

Acknowledgement.

Part of this article was written while the first author was visiting the Department of Statistics, Stanford University.

References

Berry, D. A., and Ho, C. H. (1988). One-sided sequential stopping boundaries for clinical trials: A decision-theoretic approach. *Biometrika*. 44, 219-227.

Breslow, N. (1990). Biostatistics and Bayes (with discussion). *Statistical Science*. 5, 269-298.

Cressie, N., Biele, J., and Morgan, P. B. (1991). Sample-size-optimal sequential testing. Unpublished manuscript. Iowa State University.

Cressie, N., and Morgan, P. B. (1992). The VPRT: A sequential testing procedure dominating the SPRT. Tentatively accepted for publication in *Econometric Theory*.

Cressie, N., and Morgan, P. B. (1988). The VPRT: Optimal sequential and nonsequential testing, in *Proceedings of the Fourth Purdue Symposium on Decision Theory and Related Topics, Vol. 2*, eds. S. S. Gupta and J. O. Berger. Springer, New York. 107-118.

DeGroot, M. H. (1970). *Optimal Statistical Decisions*. McGraw-Hill, New York.

Ehrenfeld, S. (1972). On group sequential sampling. *Technometrics*, **14**, 167-174.

Ghosh, B. K. (1970). *Sequential Tests of Statistical Hypotheses*. Addison-Wesley, Reading, Massachusetts.

Pocock, S. J. (1987). Group sequential methods in the design and analysis of clinical trials. *Biometrika*, **64**, 191-199.

Raiffa, H., and Schlaifer, R. (1961). *Applied Statistical Decision Theory*. MIT Press. Cambridge, Massachusetts.

Spahn, M., and Ehrenfeld, S. (1974). Optimal and suboptimal procedures in group sequential sampling. *Naval Research Logistics Quarterly*, **21**, 53-68.

Wald, A. (1947). *Sequential Analysis*. Wiley, New York.

Whitehead, J. (1983). *The Design and Analysis of Sequential Clinical Trials*. Wiley, New York.

Whittle, P. (1965). Some general results in sequential design (with discussion). *Journal of the Royal Statistical Society, Series B*, **27**, 371-394.

FIGURE CAPTIONS

Figure 1: Contour plot of ENG $g_0(p_0, 1, n_0)$ as a function of p_0 and n_0. On the horizontal axis, $0.1 \leq p_0 \leq 0.6$, and on the vertical axis, n_0 ranges from 5 to 50 in steps of size 5. Constants are given in Section 4.1: $T = 1$.

Figure 2: Plot of ENG $g_0(p_0, 1, n_0)$ as a function of p_0, for $n_0 = 0, 1, 30$, and 100. On the horizontal axis, $0.10 \leq p_0 \leq 1.00$. Constants are given in Section 4.1; $T = 1$.

Figure 3: Optimal sample sizes n_0^* and associated continue sampling intervals $[0, b_1^*)$ for $T = 1, 2$, and 3. Constants are given in Section 4.1.

Figure 4: Same as for Figure 1 except n_0 ranges from 1 to 21 in steps of size 2 and constants are given in Section 4.2.

Figure 5: Percentage increase of ENG $g_0(p_0, 1, n^*(p_0))$ over ENG $g_0(p_0, 1, 30)$. Constants are given in Section 4.2: $T = 1$.

Figure 6: Same as for Figure 3 except constants are given in Section 4.2.

Figure 1

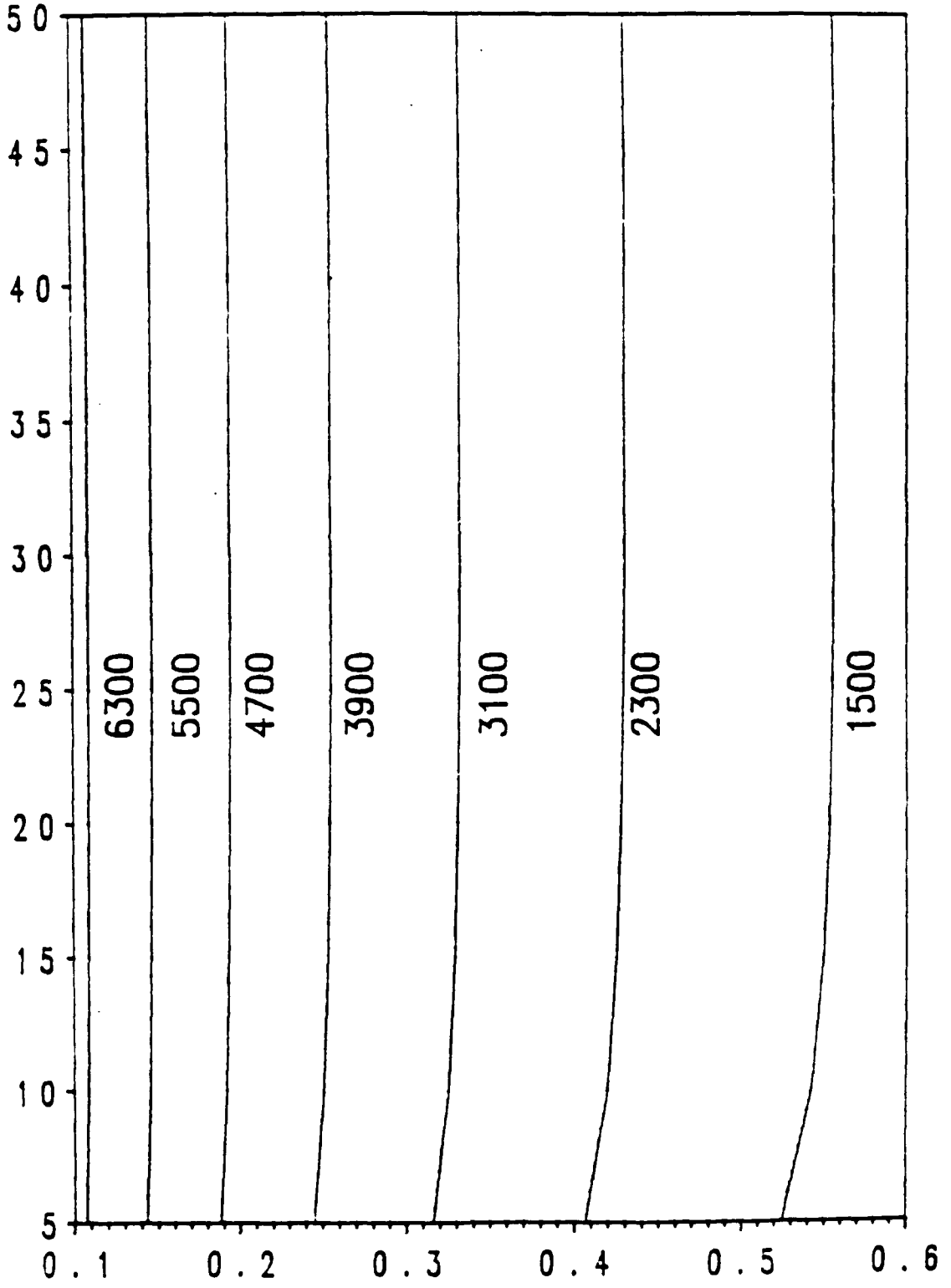

50
45
40
35
30
25
20
15
10
5
0.1
0.2
0.3
0.4
0.5
0.6
6300
5500
4700
3900
3100
2300
1500

Figure 2

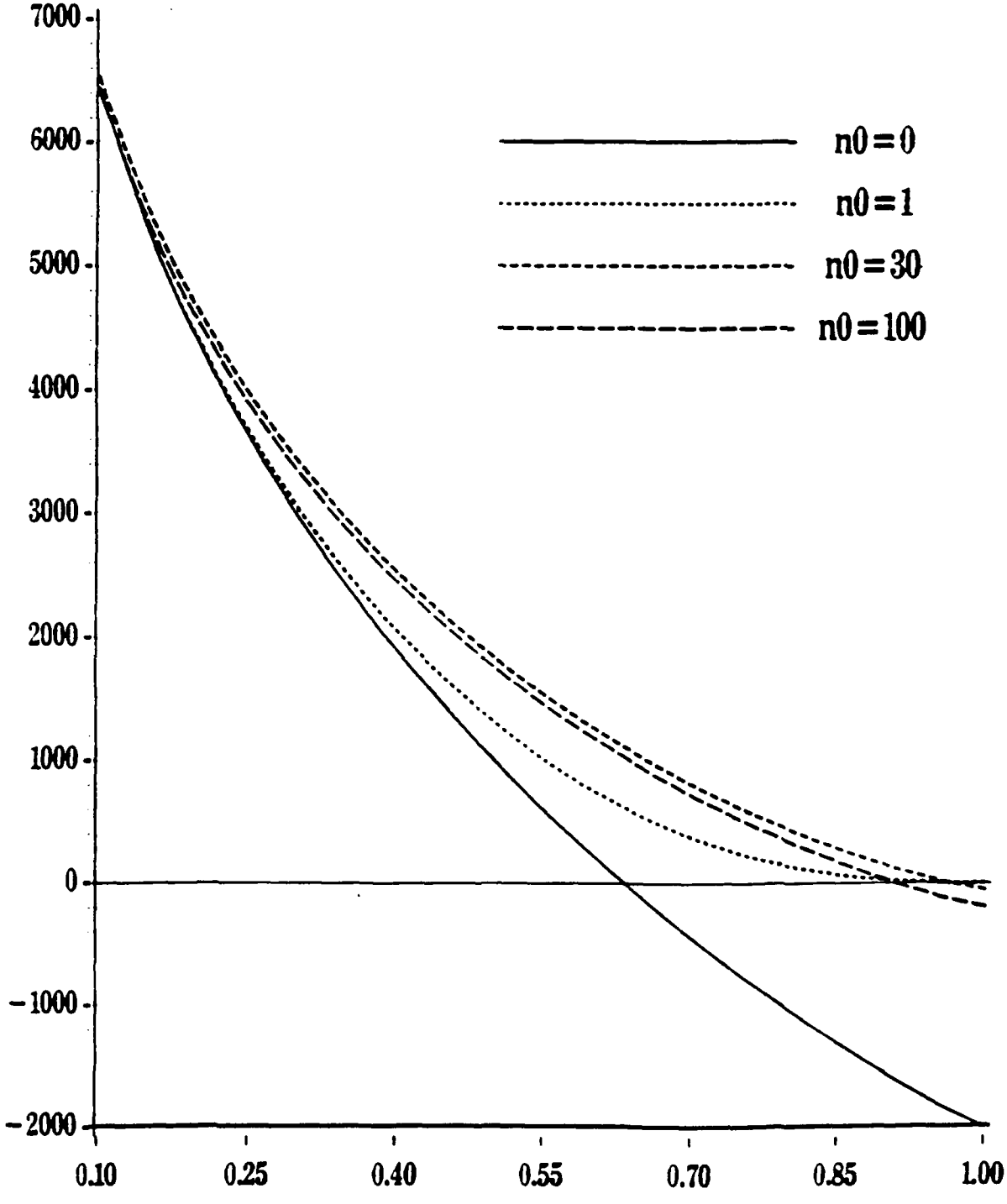

7000
6000
5000
4000
3000
2000
1000
0
-1000
-2000
0.10
0.25
0.40
0.55
0.70
0.85
1.00
n0=0
n0=1
n0=30
n0=100

Figure 3

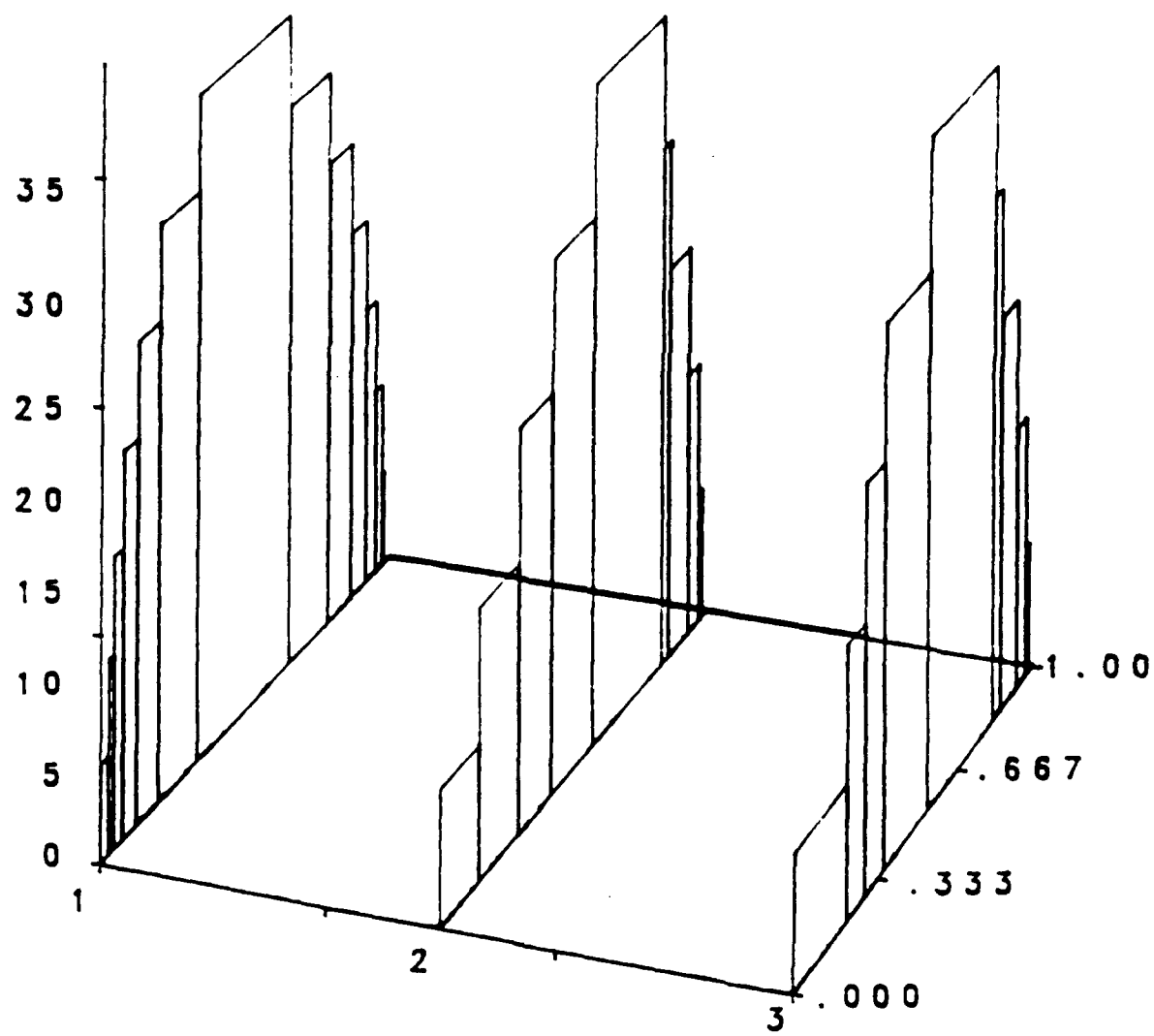


Figure 4

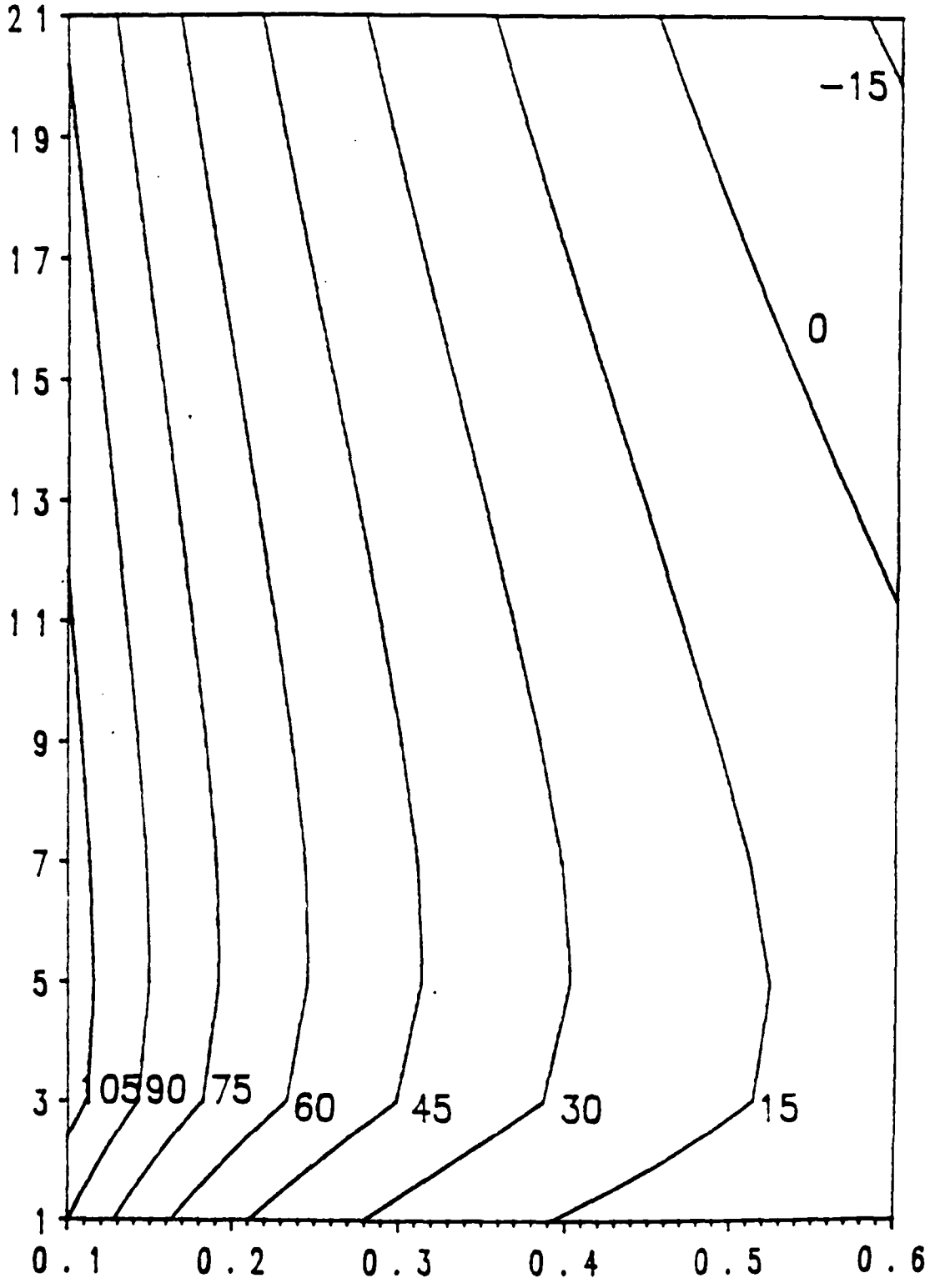


Figure 5

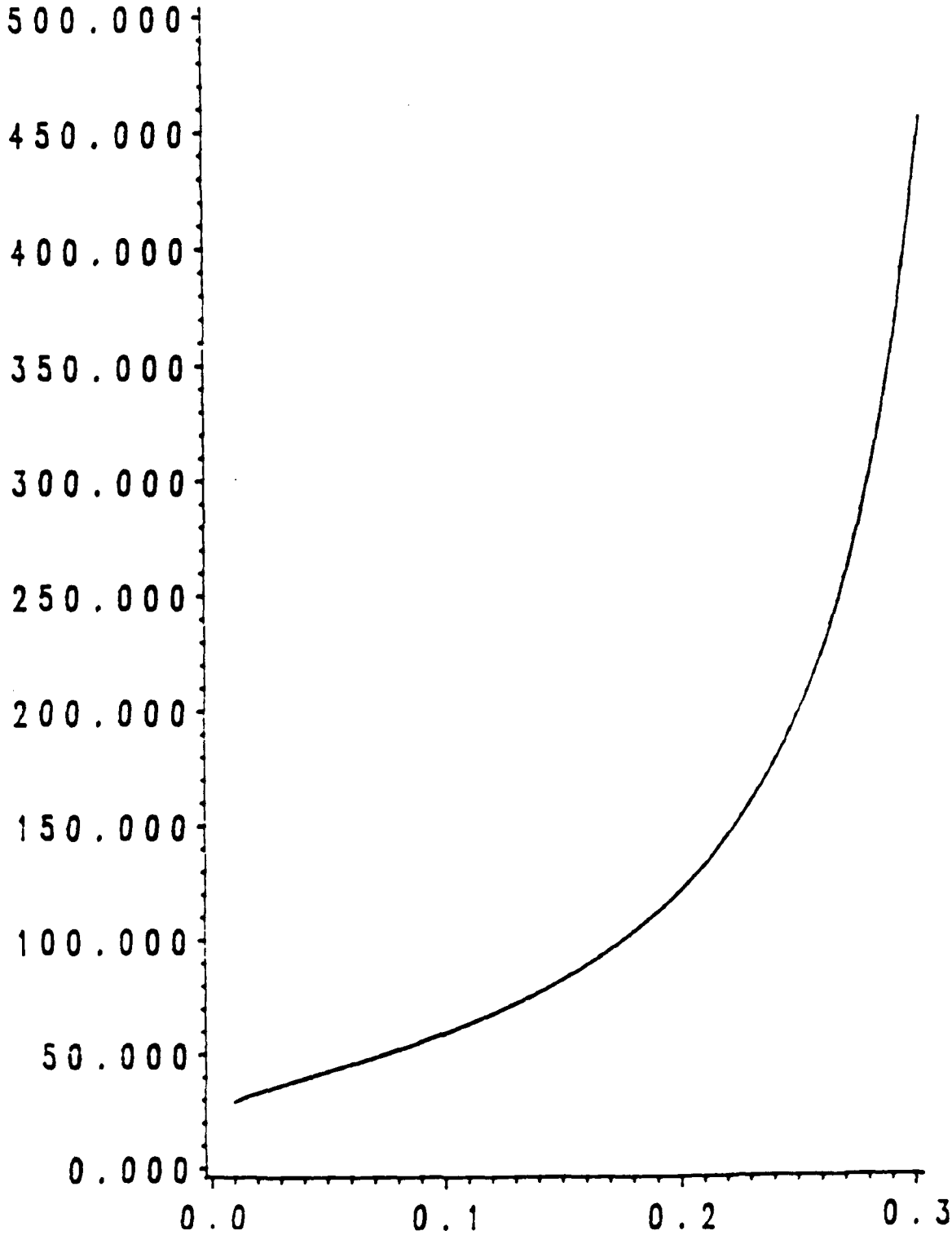
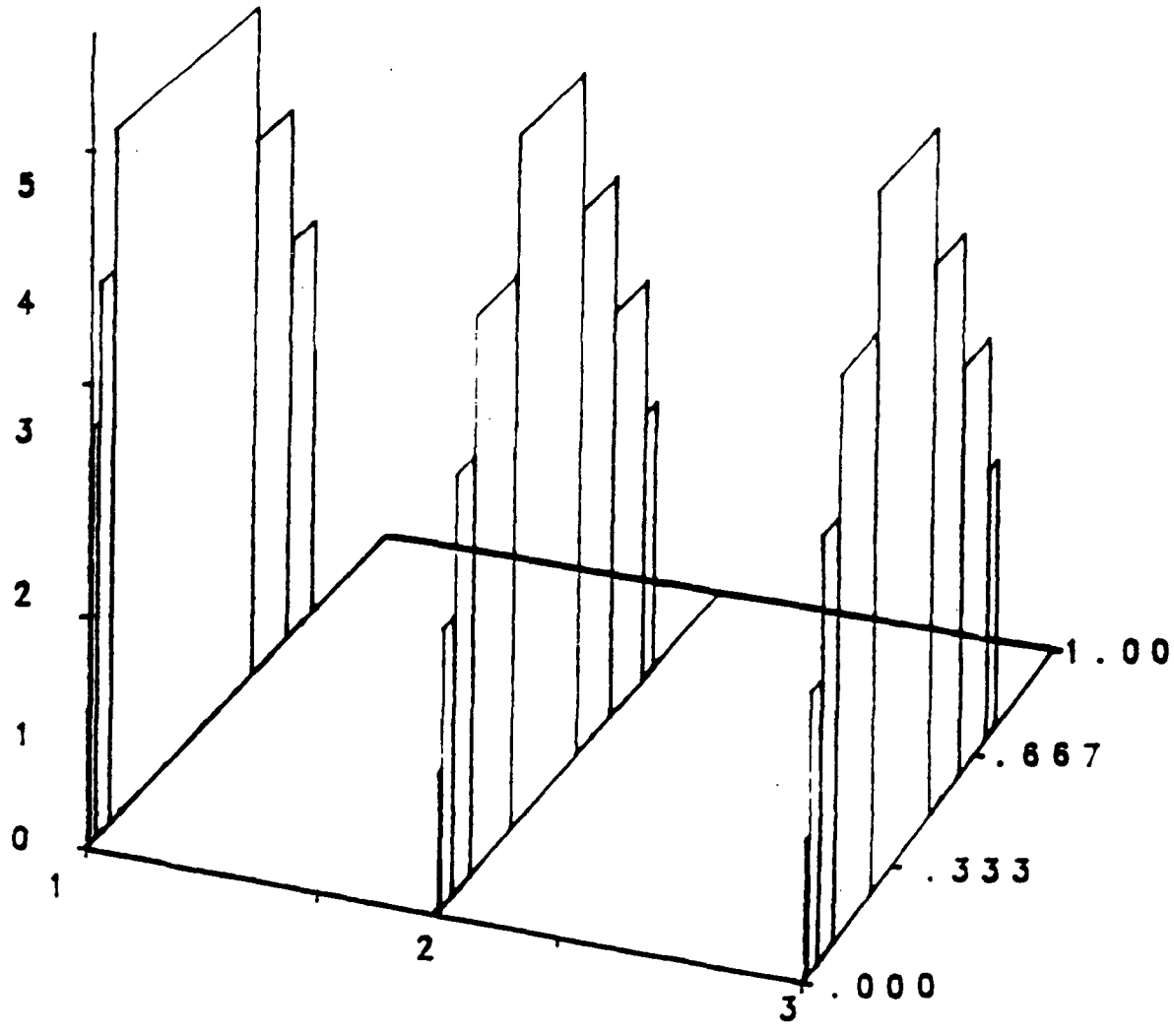

500.000
450.000
400.000
350.000
300.000
250.000
200.000
150.000
100.000
50.000
0.000
0.0
0.1
0.2
0.3

Fig. 6



UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: 451	2. GOVT ACCESSION NO.
3. RECIPIENT'S CATALOG NUMBER	
4. TITLE (and Subtitle): A Sample-Size Optimal Bayesian Procedure for Sequential Pharmaceutical Trials	5. TYPE OF REPORT & PERIOD COVERED: Technical Report
6. PERFORMING ORG. REPORT NUMBER	
7. AUTHOR(s): Noel Cressie and Jonathan Biele	8. CONTRACT OR GRANT NUMBER(s): N0025-92-J-1264
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Department of Statistics, Stanford University, Stanford, CA 94305	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: NR-042-267
11. CONTROLLING OFFICE NAME AND ADDRESS: Office of Naval Research, Statistics & Probability Program Code 1111	12. REPORT DATE: March 5, 1992
	13. NUMBER OF PAGES: 20
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)	15. SECURITY CLASS. (of this report): Unclassified
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

Also issued as a technical report, Departmentof Statistics, Iowa State University

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Backward induction; clinical trials; group-sequential procedure; optimal stopping; predictive distributipn; VPRT

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

see attached

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE S/N 0102-014-6601

UNCLASSIFIED

Consider a pharmaceutical trial where the consequences of different decisions are expressed on a financial scale. The efficacy of the new drug under consideration has a prior distribution obtained from the underlying biological process, animal experiments, clinical experience, and so forth. In an important paper, Berry and Ho (1988) show how these components are used to establish an optimal (Bayes) sequential procedure, assuming a known constant sample size at each decision point. We show in this article how it is also possible to optimize with respect to the sample size rule. This last component of the design, which is missing from most sequential procedures, has the potential to yield considerably larger expected net gains.