THE APPLICATION OF TWO-ARMED BANDIT STRATEGIES TO CLINICAL TRIALS

by

Donald A. Berry

University of Minnesota

Technical Report No. 265

May 1976

THE APPLICATION OF TWO-ARMED BANDIT STRATEGIES TO CLINICAL TRIALS

by

Donald A. Berry*

University of Minnesota

Technical Report No. 265

ABSTRACT

A procedure which maximizes the expected number of successes in a clinical trial involving two treatments can usually be found only by backward induction. Not only is it difficult to find an optimal procedure but, once found, it is difficult to describe and cumbersome to communicate. A procedure is suggested which depends on the information present concerning the treatments, is easy to calculate, and approximates an optimal procedure quite well. The procedure is applicable to trials for which the number of patients is unknown as well as those of fixed duration. It can be modified to apply as well in trials in which there is a delay between treatment application and response.

*This work was supported by Public Health Service Research Grant No. GM 22234-01 from the National Institutes of Health. Computations were performed by Larry M. Pearson while supported by a Research Assistantship from the Graduate School of the University of Minnesota.

THE APPLICATION OF TWO-ARMED BANDIT STRATEGIES TO CLINICAL TRIALS

by

Donald A. Berry

1. Introduction.

Suppose that two treatments are available for use in a clinical trial. Further suppose that the response to treatment is either positive, a success, or negative and that the patients can be treated one at a time, with each patient's response available before the next patient is to be treated. Since this assumption is seldom realistic, modifications will be considered in Section 6.

The number of patients in the trial is N, which may be considered fixed or random; for convenience the consideration of the latter possibility will be delayed until Section 7. The main purpose of the trial is to treat these N patients as effectively as possible. An inevitable and obviously beneficial result of any trial is the acquisition of information concerning the treatments. An effect of this accumulating information is that it allows for better treatment of patients appearing late in the trial than those appearing early. The treatment allocation problem for early phases of the trial is then governed by two possibly conflicting desirata: (1) A patient should be given the treatment which is apparently superior. (2) A patient should be given the treatment about which more is likely to be learned.

Suppose that the probability of success using Treatment 1 is p_1 and using Treatment 2 is p_2. If p_1 and p_2 are known precisely then there is no information to be collected during the course of the trial, and the best plan for the trial is to allocate all patients to the treatment which has the

larger p_i, and for that allocation the expected number of successes is $N \max\{p_1, p_2\}$. If the p_i are unknown then there are at least two approaches to the problem. One, the classical approach, suggests a class of allocation schemes or sampling procedures and finds the expected number of successes, for example, using each procedure as a function of (p_1, p_2). Of course, no one procedure is better than all other procedures uniformly in (p_1, p_2), so, in a sense, the procedure which one should use depends on knowing p_1 and p_2. But if p_1 and p_2 are known the best procedure is the trivial one described above! Some of these procedures will be discussed in Section 2. Another, the Bayesian approach, asks that current information concerning (p_1, p_2) be quantified in the form of a probability distribution. It is this approach which will be used in most of this paper. A distinct advantage of the approach is that accumulating information can be handled in a unified way: Bayes' theorem is used to modify the probability distribution on (p_1, p_2). The effectiveness of a procedure can then be averaged over (p_1, p_2) and, hopefully, an optimal procedure found.

If one of the p_i is known, say p_1, and p_2 has probability measure $\nu(p_2)$, then the procedure selection problem is called a "one-armed bandit;" cf. Bradt, Karlin, and Johnson (1956). There are two frequently occurring clinical trials for which this problem is appropriate: (1) Treatment 1 is a standard treatment and has been used many times, and Treatment 2 is experimental; (2) Treatment 1 is a placebo and Treatment 2 experimental. If $Ep_2 > p_1$, where expectation E is with respect to ν, then using Treatment 2 not only has more information value but also has a greater probability of success (this being the expected value of p_2) associated with it; it seems clear that Treatment 2 should be used, at least until such time as $Ep_2 < p_1$. It may not be as clear that Treatment 2 may be optimal even though $Ep_2 < p_1$.

In fact, Treatment 2 may be used optimally even though it yields only failures for an arbitrarily long time, depending on N, and provided $\nu(p_2 > p_1) > 0$. A variant of this problem which avoids the latter difficulty discounts successful treatment of future patients, so that the present patient is worth 1 while the next is worth γ and the following γ^2, etc., for $0 \leq \gamma < 1$; cf. Section 7, Bellman (1956), and Berry (1976).

An intuitively appealing characteristic of optimal procedures for the one-armed bandit is that the initial patients constitute an information gathering stage, which may be empty or may exhaust the trial, during which the experimental Treatment 2 is used, and then the standard Treatment 1 is used for the remainder of the trial. Whether information gathering is initiated depends on ν; for example, a necessary condition is $\nu(p_2 > p_1) > 0$. The length of this stage depends on the information accumulated concerning p_2 and, to a decreasing extent, the initial distribution ν of p_2.

If p_1 and p_2 are both unknown and subject to a probability measure $\mu(p_1, p_2)$, then the problem is a "two-armed bandit." Unlike the case in which p_1 is known, sampling can switch from one treatment to the other any number of times using an optimal procedure. With one important exception to be mentioned immediately, an optimal procedure can only be found using a backward induction requiring on the order of N^3 storage locations and is, in general, extremely difficult to implement since it must specify a treatment to use in each of 4^N different situations. If however, $\mu(\alpha, \beta) + \mu(\beta, \alpha) = 1$ for any fixed pair (α, β), then an optimal procedure is especially easy: Treatment 1 is used whenever the current probability that $p_i = \max\{\alpha, \beta\}$ is at least $1/2$. (This result is due to Feldman (1962), and was generalized by Fabius and van Zwet (1970), Berry (1972), and Kelley (1974).) The only calculation that is necessary in implementing this procedure is that of the posterior

probability that $(p_1, p_2) = (\alpha, \beta)$. Such procedures are discussed further in Section 3 of this paper.

A method is suggested in Section 4 for approximating μ with such a two-point measure. The loss in effectiveness in using the procedure that is optimal for this approximating measure is examined in Section 5.

2. The classical approach.

The two-armed bandit problem was considered by Robbins (1952), who suggested the so-called "play-the-winner rule:" randomly select one of the treatments to be used on the first patient and henceforth use the same treatment on the next patient whenever the response is positive and the other treatment whenever the response is negative. Though this procedure is not optimal in general for any μ, it behaves well despite its simplicity. If μ is such that p_1 and p_2 are independent, then Berry (1972) shows that an optimal procedure always "stays on a winner," but sometimes switches and sometimes stays on a loser. Bradt, Karlin, and Johnson (1956) give an example in which p_1 and p_2 are dependent and for which the unique optimal procedure switches on a winner and stays on a loser!

Robbins' approach has been generalized to allow the performance of the treatments on the previous k patients to dictate the treatment used on the next patient; this is the so-called "finite memory" approach (for references see Smith and Pyke (1965)).

There is a very large "selection and ranking" literature which examines different sampling procedures. Many of the better procedures use Robbins' play-the-winner rule in the course of the trial, but effectively treating the patients in the trial is regarded as being of secondary importance - of primary importance is determining which treatment has the larger p at the end of the trial. There are many papers which use this approach referred to

in Simon et al. (1975).

A procedure which is designed to treat successfully as many patients as possible in the course of the trial was suggested by Zelen (1969). Zelen uses Robbins' play-the-winner rule for the first n patients in the trial and then uses the apparently superior treatment for the remaining $N-n$ patients. Zelen finds that $n \doteq N/3$ yields nearly the maximal expected proportion of successes for the entire trial. In one example he supposes that $N = 100$, $p_1 = .75$, and $p_2 = .25$. The maximal expected proportion of successes using his procedure is .726, which corresponds to $n = 14$, while the proportion is .709 for $n = 33$.

To compare Zelen's procedure with the optimal procedure, which is Feldman's, fix $\mu(.75, .25) = \mu(.25, .75) = 1/2$. Feldman's procedure is especially simple when $\mu(\alpha, 1-\alpha) = \mu(1-\alpha, \alpha) = 1/2$; for the current patient it requires using the treatment for which the current difference between number of successes and number of failures is greater (cf. (3.2)), and uses either when these differences are the same. The total expected proportion of successes when $N = 100$ using Feldman's procedure is .740. Now this number is, as usual, an average with respect to μ, which is a distinctly Bayesian concept. However, in view of the special nature of μ, this number is constant over its entire support so this number has a non-Bayesian interpretation as well and is directly comparable with .709 and .726 of Zelen's procedures - of course, no procedure can do better than Feldman's. Incidentally, .740 compares favorably with .75, the expected proportion attainable if the better treatment were known a priori. Put in these terms, it is clear that Feldman's procedure uses the better treatment ($p_i = 3/4$) an average of 98 of the 100 trials!

At least two other papers of importance should be mentioned here. Anscombe (1963) and Colton (1963) independently proposed a two-stage procedure similar to Zelen's except that both treatments are used the same number of times

in the initial, information gathering stage. This was in turn generalized by Cornfeld et al. (1969) to allow for an arbitrary proportion of patients on Treatment 1 in the initial stage.

3. Myopic procedures.

There is a slightly different way of viewing Feldman's procedure that can help put it in perspective. If $N = 1$ then an optimal procedure is to use Treatment 1 if $Ep_1 \geq Ep_2$ and Treatment 2 otherwise, where expectation is with respect to μ. If $\mu(\alpha, \beta) = r = 1-\mu(\beta, \alpha)$ and $\alpha > \beta$ then $Ep_1 - Ep_2 = r\alpha + (1-r)\beta - [(1-r)\alpha + r\beta] = 2(\alpha-\beta)(r-\frac{1}{2})$ has the same sign as $r - \frac{1}{2}$. Therefore, Feldman's procedure behaves as though each patient is the only patient in the trial; that is, the procedure is "myopic."

At any stage of the trial let s_1, f_1, s_2, f_2 be, respectively, the numbers of successes and failures on Treatments 1 and 2. The probability (posterior probability) that $p_1 = \alpha$ is then

$$\mu(\alpha, \beta | s_1, f_1; s_2, f_2) = \left[1 + \frac{1-r}{r}\left(\frac{\beta}{\alpha}\right)^{s_1-s_2}\left(\frac{1-\beta}{1-\alpha}\right)^{f_1-f_2}\right]^{-1}$$

which is $> \frac{1}{2}$ if

$$(3.1) \qquad (s_1-s_2)\log\frac{\alpha}{\beta} > (f_1-f_2)\log\frac{1-\beta}{1-\alpha} + \log\frac{1-r}{r}.$$

In case $\beta = 1-\alpha$ this is equivalent to

$$[(s_1-f_1)-(s_2-f_2)]\log\frac{\alpha}{1-\alpha} > \log\frac{1-r}{r},$$

and further, if $r = \frac{1}{2}$ it is equivalent to

$$(3.2) \qquad s_1-f_1 > s_2-f_2,$$

as was stated in Section 2.

While the optimal procedure for general μ is not known, it has been partially characterized in Berry (1972) when p_1 and p_2 are independent. The myopic procedure for (α, β) arbitrary has at least one characteristic in common with the optimal procedure: it "stays on a winner." A success on Treatment 1 increases the left-hand side of (3.1) while a success on Treatment 2 decreases it. Also, a failure may leave inequality (3.1) unchanged or it may reverse it; the difference between a myopic procedure and an optimal procedure is that one may switch on a failure when the other does not.

4. The class of procedures $B(\mu)$.

The procedure $B(\alpha(\mu), \beta(\mu))$, abbreviated $B(\mu)$, is defined to be Feldman's procedure with $\alpha = \alpha(\mu)$, $\beta = \beta(\mu)$, and $r = r(\mu)$; that is, Treatment 1 is used whenever inequality (3.1) holds and Treatment 2 is used whenever the inequality in (3.1) is reversed. When the inequality in (3.1) is replaced with equality then either treatment can be used following Feldman's procedure; $B(\mu)$ may strictly prefer one of the treatments depending on μ, as will be described shortly.

There is no obviously best way to define $\alpha(\mu)$, $\beta(\mu)$, and $r(\mu)$. While there are many candidates, the definition proposed here is natural and, as will be seen, gives excellent results.

Make the simplifying assumption that μ is such that $\mu(p_1=p_2) = 0$. This is without loss of generality since if $\mu(p_1=p_2) > 0$ and μ is changed to

$$\mu'(p_1, p_2) = \begin{cases} \mu(p_1, p_2)/\mu(p_1=p_2), & \text{for } p_1 \neq p_2 \\ 0, & \text{for } p_1 = p_2 \end{cases}$$

then the optimal procedure is unchanged. The point is that when $p_1 = p_2$ the treatments are the same and on the line $p_1 = p_2$ <u>every</u> procedure has the same effectiveness.

For arbitrary μ define

$$r(\mu) = \mu(p_1 > p_2) = \int_{p_1 > p_2} d\mu(p_1, p_2).$$

Assume $r(\mu) \geq \frac{1}{2}$; if it is not then reverse the names Treatment 1 and Treatment 2. This means that Treatment 1 is used initially when following $B(\mu)$, except possibly when $r(\mu) = \frac{1}{2}$. Consider the measure

$$\mu^*(p_1, p_2) = \begin{cases} \mu(p_1, p_2)/r(\mu), & \text{for } p_1 > p_2 \\ 0, & \text{for } p_1 \leq p_2 \end{cases}$$

and define

(4.1) $$\alpha(\mu) = E^* p_1, \qquad \beta(\mu) = E^* p_2,$$

where E^* denotes expectation with respect to μ^*. So defined, $\alpha(\mu)$ and $\beta(\mu)$ are easy to calculate and are simply the conditional expectations of p_1 and p_2 given $p_1 > p_2$. An alternative definition that comes to mind immediately corresponds to conditioning on $p_1 < p_2$ instead; however, comparisons suggest that conditioning on the triangle with larger probability is usually preferable.

Procedure $B(\mu)$ is myopic and is not optimal in general, no matter how $\alpha(\mu)$, $\beta(\mu)$, and $r(\mu)$ are defined. Kelley (1974) explores conditions under which every optimal procedure is myopic. $B(\mu)$ is compared with an optimal procedure in Section 5 for various μ which have densities in the plane:

(4.2) $\mu(p_1\ p_2) = \nu_1(p_1)\nu_2(p_2); \quad 0 < p_i < 1, \quad i = 1,2,$

and ν_1 and ν_2 are beta densities:

(4.3) $\nu_1(x) = k_i x^{a_i-1}(1-x)^{b_i-1}, \quad 0 < x < 1.$

The ν_i are densities for $a_i > 0$ and $b_i > 0$. Under this assumption p_1 and p_2 and independent and $Ep_i = a_i/(a_i + b_i)$, $i = 1,2$. Also, where s_i and f_i are defined in Section 3, for $i = 1, 2$,

$$E(p_i|s_1, f_1; s_2, f_2) = \frac{a_i + s_i}{a_i + b_i + s_i + f_i}.$$

Table 1 gives the values of $\alpha(\mu)$, $\beta(\mu)$, and $r(\mu)$ for various ν_1 and ν_2 as given by the pairs (a_1, b_1) and (a_2, b_2).

The initial treatment to be used following $B(\mu)$ is as yet not specified when $r(\mu) = \frac{1}{2}$, or at any later stage when equality replaces the inequality in (3.1). In such a case $B(\mu)$ is defined to use Treatment i if "less is known" about p_i than p_j, $j \neq i$. When μ is of the form (4.2), this is taken to mean using Treatment i if $a_i + b_i < a_j + b_j$; when they are equal either treatment is used. When p_1 and p_2 are not independent beta variables then the definition of "less is known" is left open. While there are several reasonable definitions of the information contained in a probability distribution, these are avoided here because the matter is usually an insignificant one. Indeed, μ may be such that it is impossible for the probability of $(p_1 > p_2)$ ever to become exactly $\frac{1}{2}$. It can be important if a great deal is known about Treatment 1, say, and little about Treatment 2. In such a circumstance $B(\mu)$ could be effectively modified by requiring the use of the Treatment 2 whenever $r(\mu)$ is somewhat less than $\frac{1}{2}$ - this is the character of the optimal procedure in the one-armed bandit, in which one treatment is precisely known. A limitation of $B(\mu)$ is that there is no way of letting a two-point

measure reflect the possibility that "more is known" about one treatment than the other.

5. Comparison of $B(\mu)$ and optimal procedure.

Consider first the special case in which μ is the product of uniform measures on $(0, 1)$, so that p_1 and p_2 are independent and, in terms of the notation of (4.2) and (4.3), $a_1 = b_1 = a_2 = b_2 = 1$. Then, by symmetry, $r(\mu) = \frac{1}{2}$, and

$$\alpha(\mu) = 2\int_0^1\int_0^{p_1} p_1 dp_2 dp_1 = 2/3,$$

$$\beta(\mu) = 2\int_0^1\int_0^{p_1} p_2 dp_2 dp_1 = 1/3.$$

Since $r(\mu) = \frac{1}{2}$ and $\beta(\mu) = 1-\alpha(\mu)$, procedure $B(\mu)$ uses Treatment 1 whenever (3.2) holds and Treatment 2 whenever the inequality is reversed in (3.2); whenever (3.2) holds with equality then Treatment 1 is used if $s_1 + f_1 \leq s_2 + f_2$ and Treatment 2 otherwise. $B(\mu)$ is an optimal procedure for values of $N < 5$ and imitates an optimal procedure much of the time for larger values of N. This is seen in Table 2 which gives the expected proportion of success for $B(\mu)$ and for optimal procedures when $N = 1(1)10(5)40(20)100$. It is to be stressed that these expectations for $B(\mu)$ are calculated with respect to μ and <u>not</u> with respect to the indicated two-point measure; the clinician is acting as though μ were a two-point measure but he knows the actual μ. As $N \to \infty$ the limiting proportion for $B(\mu)$ and optimal procedures is

$$E \max\{p_1, p_2\} = 2/3,$$

and the convergence is monotonic in both.

Table 3 gives the expected proportion of success using $B(\mu)$ and using an optimal procedure when $N = 25$ for each of the pairs of beta densities considered in Table 4.1. Table 4 compares these procedures when $N = 50$. These numbers were computed by backward induction. In addition, the values of $\max\{Ep_1, Ep_2\}$ and of $E \max\{p_1, p_2\}$ are given for each pair. These two numbers give effective lower and upper bounds for all reasonable procedures and all N. It is readily seen that in quite a broad range of measures μ the procedure $B(\mu)$ compares favorably with an optimal procedure.

As indicated before, the manner in which $r(\mu)$, $\alpha(\mu)$, $\beta(\mu)$ were defined in Section 4 is not necessarily "optimal" - the procedure $B(\mu)$ may not be the best myopic procedure. Calculations not reported here were made for various triples (r, α, β) and various μ and N; they make it clear that when $B(\mu)$ is not optimal among myopic procedures it is nearly so.

6. Delays in treatment response.

In most clinical trials there is a nontrivial time delay between the administration of a treatment (as a drug) and a response to the treatment. During this delay other patients may have to be treated - possibly for considerations of time as well as of patient health. If this delay is so great that none of the patients will respond before the Nth patient is treated then the procedure is easy: simply assign all patients to the treatment with the larger probability of success, given by EP_1 or Ep_2. As is evident here, and is true more generally, there is a decrease in the maximal expected proportion of success when there is response delay.

Suppose that the delay is in terms of a fixed number, n, or patients, so that $n + 1$ patients are treated before the first patient's response becomes known. When μ is such that $\mu(\alpha, \beta) + \mu(\beta, \alpha) = 1$ then an induction on n shows that the myopic procedure is again optimal. In this case, of course, the current probability that $p_1 > p_2$ is used to determine which treatment to use, and since this probability does not change until the $(n+1)$st

patient is treated, at the least the first $n + 1$ patients are given the same treatment. Also, except for the first n patients one response becomes known immediately after each treatment and the information contained in a response is independent of the treatment used, so that the last N-n patients are treated as though they constitute a trial with no response delay. If, as in the example in Section 3, $\mu(.75, .25) = \mu(.25, .75) = \frac{1}{2}$, assume $N = 200$ and $n = 100$; then the expected proportion of success over the trial is

$$\frac{100}{200}(\tfrac{1}{2}(.75) + \tfrac{1}{2}(.25)) + \frac{100}{200}(.740) = .250 + .370 = .620,$$

which compares with .5 if $n \geq 199$ and .745 if $n = 0$. If n is considered to be random, and varying from one patient to another, then this result extends to show that the myopic procedure is again optimal.

The determination of an optimal procedure for general μ does not follow as easily from the case of no response delay. It is not in general true, for example, that the first $n + 1$ patients should be given the same treatment. The obvious generalization of $B(\mu)$ to this case, approximating μ with a two-point measure and allocating accordingly, is not as efficient as when $n = 0$. It seems reasonable to modify $B(\mu)$ so that the first n patients are allocated between the treatments so that the expected amount of information present after n responses become known will be approximately the same for both treatments. Another possibility, which could be combined with the first, is to assign $r(\mu)$ of the first n patients to Treatment 1 and $1-r(\mu)$ of them to Treatment 2. The remaining N-n patients can then be allocated following $B(\mu)$. The effectiveness of these procedures has not been considered here. The optimal procedures for the delayed response problem have never been systematically studied. If n is allowed to be random rather than fixed then it is not obvious how best to modify $B(\mu)$ and comparisons become increasingly complicated. There are a variety of kinds of "randomness" to consider. The delay in response may be less for one treatment (which may not be known) or it may depend on whether or not the response is positive. In such situations there is still a third consideration in using a treatment: the desire to use a treatment which will likely yield

information more quickly so that information can be considered in the treatment of more patients. This desirata is inseparable from the other two that have been discussed previously - using a treatment with a large expected probability of success and one which yields a large amount of information.

7. Unknown trial length.

It is frequently the case that N is not precisely known, but is subject to a probability distribution. Assume that N is independent of (p_1, p_2) and of the way in which the trial is conducted. If $\mu(\alpha, \beta) + \mu(\beta, \alpha) = 1$ then the optimal procedure, being myopic, is not affected. The procedure $B(\mu)$ does not depend on N though, of course, the optimal procedure does depend on N. Roughly speaking, an optimal procedure is more inclined to sacrifice immediate payoff for information that can be used later. $B(\mu)$ is offered here as a reasonable approximation to the optimal procedure for all values of N and, therefore, it is reasonable for any distribution of N. In fact, since the one-point distributions are the extremes, the distribution of N for which $B(\mu)$ suffers most in comparison with the optimal procedure is one in which N is known. In this sense, not knowning N actually improves the appropriateness of $B(\mu)$.

These remarks hold as well if N is effectively infinite and a success on the jth patient is discounted by the factor γ^{j-1}, where $0 \leq \gamma < 1$. Then the expected discounted number of successes is bounded above by $(1-\gamma)^{-1}$ which plays the role of N(though not in every sense). Posed thusly, this problem is similar to the one in which N is random, with a geometric distribution; there is a constant probability, $1-\gamma$, for each patient that the trial will terminate with that patient.

8. Summary and conclusions.

A procedure which maximizes the expected number of successes in a clinical trial has been shown to be reasonably approximable with a procedure that is relatively easy to calculate. This procedure, $B(\mu)$, is determined from μ, the

available information concerning the treatments, using elementary calculations, and without the need for backward induction. $B(\mu)$ is actually optimal for certain μ and compares favorably with an optimal procedure in situations for which it is not optimal. The procedure is appropriate, possibly with modification, when there is a delay between treatment application and response and when the length of the trial is not known.

REFERENCES

Anscombe, F. J.(1963). Sequential medical trials, Journal of the American Statistical Association, 58, 365-383.

Bellman, R. E.(1956). A problem in the sequential design of experiments, Sankhya, 16, 221-229.

Berry, D. A.(1972). A Bernoulli two-armed bandit, Annals of Mathematical Statistics, 43, 871-897.

Berry, D. A.(1976). One-armed bandits and related problems. Technical Report No. 267, School of Statistics, University of Minnesota.

Bradt, R. N., Johnson, S. M., and Karlin, S.(1956). On sequential designs for maximizing the sum of n observations, Annals of Mathematical Statistics, 27, 1060-1070.

Colton, T.(1963). A model for selecting one of two medical treatments, Journal of the American Statistical Association, 58, 388-400.

Cornfield, J., Halperin, M., and Greenhouse, S. W.(1969). An adaptive procedure for sequential clinical trials, Journal of the American Statistical Association, 64, 759-770.

Fabius, J. and van Zwet, W. R.(1970). Some remarks on the two-armed bandit, Annals of Mathematical Statistics, 41, 1906-1916.

Feldman, D.(1962). Contributions to the "two-armed bandit" problem, Annals of Mathematical Statistics, 33, 847-856.

Kelley, T. A.(1974). A note on the Bernoulli two-armed bandit problem, Annals of Statistics, 2, 1056-1062.

Robbins, H. E.(1952). Some aspects of the sequential design of experiments, Bulletin of the American Mathematical Society, 55, 527-535.

Simon, R., Weiss, G. H., and Hoel, D. G.(1975). Sequential analysis of binomial clinical trials, Biometrika, 62, 195-200.

Smith, C. V. and Pyke, R.(1965). The Robbins-Isbell two-armed bandit problem with finite memory, Annals of Mathematical Statistics, 36, 1375-1386.

Zelen, M.(1969). Play-the-winner rule and controlled clinical trials, Journal of the American Statistical Association, 64, 131-140.

TABLE 1. Values of $r(\mu)$, $\sigma(\mu)$, $\beta(\mu)$ when μ is the Product of Densities of the Form (4.3).

(a_2,b_2) / (a_1,b_1)	(4, 1)	(6, 2)	(1, ½)	(2, 1)	(½, ½)	(1, 1)	(2, 2)	(4, 4)	(6, 6)	(1, 2)	(½, 1)	(2, 6)	(1, 4)
(4, 1)	.50 .89 .71	.62 .88 .70	.60 .84 .50	.67 .86 .57	.73 .83 .35	.80 .83 .42	.86 .83 .46	.89 .83 .48	.91 .83 .49	.93 .82 .31	.89 .82 .27	.99 .81 .25	.99 .81 .20
(6, 2)		.50 .83 .67	.53 .79 .45	.58 .80 .53	.68 .77 .32	.75 .78 .39	.82 .78 .44	.87 .78 .47	.89 .78 .48	.92 .76 .30	.86 .76 .25	.99 .75 .25	.99 .75 .19
(1, ½)			.50 .83 .51	.53 .86 .57	.64 .77 .34	.66 .80 .40	.68 .81 .45	.69 .82 .47	.70 .83 .48	.80 .76 .29	.79 .75 .25	.86 .74 .24	.89 .73 .18
(2, 1)				.50 .80 .53	.63 .73 .30	.67 .75 .38	.70 .76 .43	.72 .77 .46	.73 .77 .47	.83 .72 .28	.80 .71 .24	.92 .70 .24	.93 .69 .18
(½, ½)					.50 .70 .30	.50 .75 .38	.50 .78 .43	.50 .80 .47	.50 .80 .48	.62 .70 .27	.63 .67 .22	.67 .69 .23	.72 .65 .17
(1, 1)						.50 .67 .33	.50 .70 .40	.50 .72 .44	.50 .73 .46	.67 .63 .25	.67 .60 .20	.75 .61 .22	.80 .58 .17
(2, 2)							.50 .63 .37	.50 .65 .42	.50 .66 .45	.70 .57 .24	.69 .56 .19	.82 .56 .22	.86 .54 .17
(4, 4)								.50 .60 .41	.50 .61 .43	.72 .54 .23	.70 .53 .18	.87 .53 .22	.89 .52 .17
(6, 6)									.50 .58 .42	.73 .53 .23	.70 .52 .17	.89 .52 .22	.91 .51 .17
(1, 2)										.50 .47 .20	.53 .43 .14	.58 .47 .20	.67 .43 .14
(½, 1)											.50 .50 .17	.52 .55 .21	.59 .50 .15
(2, 6)												.50 .33 .17	.62 .30 .12
(1, 4)													.50 .29 .11

TABLE 2. Expected Proportion of Success when $a_1 = b_1 = a_2 = b_2 = 1$.

N	Optimal Procedure	$B(\mu)$
1	.50000	.50000
2	.54167	.54167
3	.55556	.55556
4	.56944	.56944
5	.57778	.57611
6	.58472	.58403
7	.59028	.58812
8	.59494	.59346
9	.59866	.59625
10	.60218	.60017
15	.61410	.61046
20	.62156	.61746
25	.62679	.62162
30	.63066	.62515
35	.63371	.62743
40	.63617	.63410
60	.64271	.63470
80	.64657	.63757
100	.64918	.63943

TABLE 3. Values of $\max(Ep_1, Ep_{\prime\prime})$, the Expected Proportion of Success Using $B(\mu)$ and an Optimal Procedure for $N = 25$, and $E \max(p_1, p_{\prime\prime})$ for Distributions μ as in Table 1.

(a_2,b_2)	(4, 1)	(6, 2)	(1, ⅓)	(2, 1)	(⅔, ⅔)	(1, 1)	(2, 2)	(4, 4)	(6, 6)	(1, 2)	(½, 1)	(2, 6)	(1, 4)
(a_1,b_1)													
	.80000	.80000	.80000	.80000	.80000	.80000	.80000	.80000	.80000	.80000	.80000	.80000	.80000
	.85800	.83549	.83610	.82651	.81489	.80746	.80320	.80030	.79993	.79681	.79953	.79702	.79728
(4,1)	.85941	.83659	.84435	.82895	.82216	.81065	.80520	.80311	.80259	.80087	.80291	.80003	.80001
	.88889	.86364	.87388	.85714	.84922	.83333	.82143	.81414	.81154	.80952	.81818	.80152	.80159
		.75000	.75000	.75000	.75000	.75000	.75000	.75000	.75000	.75000	.75000	.75000	.75000
		.79115	.79559	.78308	.76741	.75760	.75227	.74986	.74905	.74729	.74906	.74821	.74792
(6, 2)		.79796	.81520	.78939	.78229	.76275	.75461	.75208	.75153	.75044	.75259	.75000	.75000
		.83077	.84361	.82222	.81219	.79167	.77576	.76573	.76212	.76111	.77255	.75128	.75152
			.66667	.66667	.66667	.66667	.66667	.66667	.66667	.66667	.66667	.66667	.66667
			.79484	.78199	.74350	.73202	.72133	.71339	.71045	.68892	.69591	.67067	.66821
(1, ½)			.80052	.78742	.75190	.73651	.72607	.72105	.71966	.69245	.70091	.67655	.67320
			.83333	.81905	.78294	.76667	.75397	.74572	.74260	.71429	.72603	.68922	.68485
				.66667	.66667	.66667	.66667	.66667	.66667	.66667	.66667	.66667	.66667
				.76029	.72345	.71111	.70009	.69323	.69037	.67510	.68192	.66451	.66377
(2, 1)				.76351	.73694	.71595	.70239	.69603	.69415	.67831	.68706	.66890	.66799
				.80000	.77083	.75000	.73333	.72222	.71795	.70000	.71429	.67778	.67619
					.50000	.50000	.50000	.50000	.50000	.50000	.50000	.50000	.50000
					.66001	.63763	.61956	.60656	.60171	.56794	.58028	.52956	.52334
(⅔, ⅔)					.66772	.65290	.64270	.63726	.63555	.57419	.58530	.54006	.52837
					.70264	.68750	.67578	.66823	.66539	.60417	.61628	.56219	.54922
						.50000	.50000	.50000	.50000	.50000	.50000	.50000	.50000
						.62162	.60065	.58563	.57961	.54643	.55616	.51549	.51024
(1, 1)						.62679	.61148	.60310	.60063	.54957	.56464	.51977	.51311
						.66667	.65000	.63889	.63462	.58333	.60000	.54167	.53333
							.50000	.50000	.50000	.50000	.50000	.50000	.50000
							.58407	.56732	.56021	.52752	.53686	.50418	.50136
(2, 2)							.58667	.57430	.57035	.53107	.55013	.50672	.50408
							.62857	.61304	.60769	.56667	.58730	.52576	.52143
								.50000	.50000	.50000	.50000	.50000	.50000
								.55338	.54578	.51417	.52321	.49859	.49700
(4, 4)								.55443	.54859	.51980	.54427	.50121	.50063
								.59518	.58752	.57905	.57905	.51573	.51414
									.50000	.50000	.50000	.50000	.50000
									.53953	.50817	.51801	.49750	.49703
(6, 6)									.54010	.51650	.54257	.50025	.50011
									.57893	.55128	.57593	.51212	.51154
										.33333	.33333	.33333	.33333
										.42292	.43630	.36880	.35864
(1, 2)										.42613	.44893	.37297	.35993
										.46667	.48571	.40556	.39048
											.33333	.33333	.33333
											.45734	.38781	.37327
(½, 1)											.46281	.39673	.37845
											.50000	.42694	.40722
												.25000	.25000
												.29328	.27438
(2, 6)												.29397	.27869
												.33077	.31364
													.20000
													.25280
(1, 4)													.25387
													.28889

TABLE 4. Values of $\max\{Ep_1, Ep_2\}$, the Expected Proportion of Success Using $B(\mu)$ and an Optimal Procedure for N = 50, and E $\max\{p_1, p_2\}$ for Distributions μ as in Table 1.

(a_2,b_2)	(4, 1)	(6, 2)	(1, ½)	(2, 1)	(½, ½)	(1, 1)	(2, 2)	(4, 4)	(6, 6)	(1, 2)	(½, 1)	(2, 6)	(1, 4)
(a_1,b_1)													
	.80000	.80000	.80000	.80000	.80000	.80000	.80000	.80000	.80000	.80000	.80000	.80000	.80000
	.86548	.84174	.84371	.83995	.83063	.81191	.80505	.80226	.80115	.79871	.80233	.79694	.79718
(4, 1)	.86810	.84376	.85428	.83688	.83061	.81653	.80862	.80516	.80422	.80215	.80614	.80012	.80009
	.88889	.86364	.87388	.85714	.84922	.83333	.82143	.81414	.81154	.80952	.81818	.80152	.80159
		.75000	.75000	.75000	.75000	.75000	.75000	.75000	.75000	.75000	.75000	.75000	.75000
		.80496	.80454	.79122	.77437	.76317	.75605	.75191	.75046	.74903	.75248	.74792	.74807
(6, 2)		.80661	.82398	.79905	.79253	.77062	.75880	.75418	.75306	.75164	.75681	.75003	.75002
		.83077	.84361	.82022	.81219	.79167	.77576	.76573	.76212	.76111	.77255	.75128	.75152
			.66667	.66667	.66667	.66667	.66667	.66667	.66667	.66667	.66667	.66667	.66667
			.80375	.79035	.75137	.73927	.72793	.71935	.71574	.69353	.70163	.67343	.67071
(1, ½)			.81195	.79764	.76246	.74607	.73448	.72828	.72632	.69872	.70867	.67982	.67609
			.83333	.81905	.78294	.76667	.75397	.74572	.74260	.71429	.72603	.68922	.68485
				.66667	.66667	.66667	.66667	.66667	.66667	.66667	.66667	.66667	.66667
				.76987	.73197	.71910	.70726	.69894	.69557	.67977	.68758	.66611	.66532
(2, 1)				.77500	.74867	.72636	.71109	.70292	.70035	.68375	.69498	.67059	.66912
				.80000	.77083	.75000	.73333	.72222	.71795	.70000	.71429	.67778	.67619
					.50000	.50000	.50000	.50000	.50000	.50000	.50000	.50000	.50000
					.67014	.64857	.63154	.61926	.61462	.57547	.58801	.53525	.52786
(½, ½)					.68029	.66456	.65335	.64697	.64520	.58372	.59551	.54665	.53454
					.70264	.68750	.67578	.66823	.66539	.60417	.61628	.56219	.54932
						.50000	.50000	.50000	.50000	.50000	.50000	.50000	.50000
						.63257	.61188	.59752	.59161	.55460	.56539	.52034	.51446
(1, 1)						.63993	.62334	.61363	.61070	.55974	.57601	.52560	.51836
						.66667	.65000	.63889	.63462	.58333	.60000	.54167	.53333
							.50000	.50000	.50000	.50000	.50000	.50000	.50000
							.59493	.57799	.57082	.53544	.54615	.50785	.50453
(2, 2)							.59909	.58509	.58044	.54112	.56209	.51080	.50760
							.62857	.61364	.60769	.56667	.58730	.52576	.52143
								.50000	.50000	.50000	.50000	.50000	.50000
								.56306	.55505	.52154	.53175	.50071	.49933
(4, 4)								.56496	.55811	.52943	.55478	.50318	.50226
								.59518	.58752	.57905	.57905	.51573	.51414
									.50000	.50000	.50000	.50000	.50000
									.44810	.51475	.52625	.49876	.49838
(6, 6)									.54918	.52599	.55235	.50121	.50088
									.57893	.55128	.57593	.51212	.51154
										.33333	.33333	.33333	.33333
										.43323	.44720	.37695	.36573
(1, 2)										.43865	.46064	.38194	.38832
										.46667	.48571	.40556	.39048
											.33333	.33333	.33333
											.76726	.39561	.38082
(½, 1)											.47540	.40573	.38702
											.50000	.42694	.40722
												.25000	.25000
												.30174	.28269
(2, 6)												.30341	.28803
												.33077	.31364
													.20000
													.26154
(1, 4)													.26377
													.28889