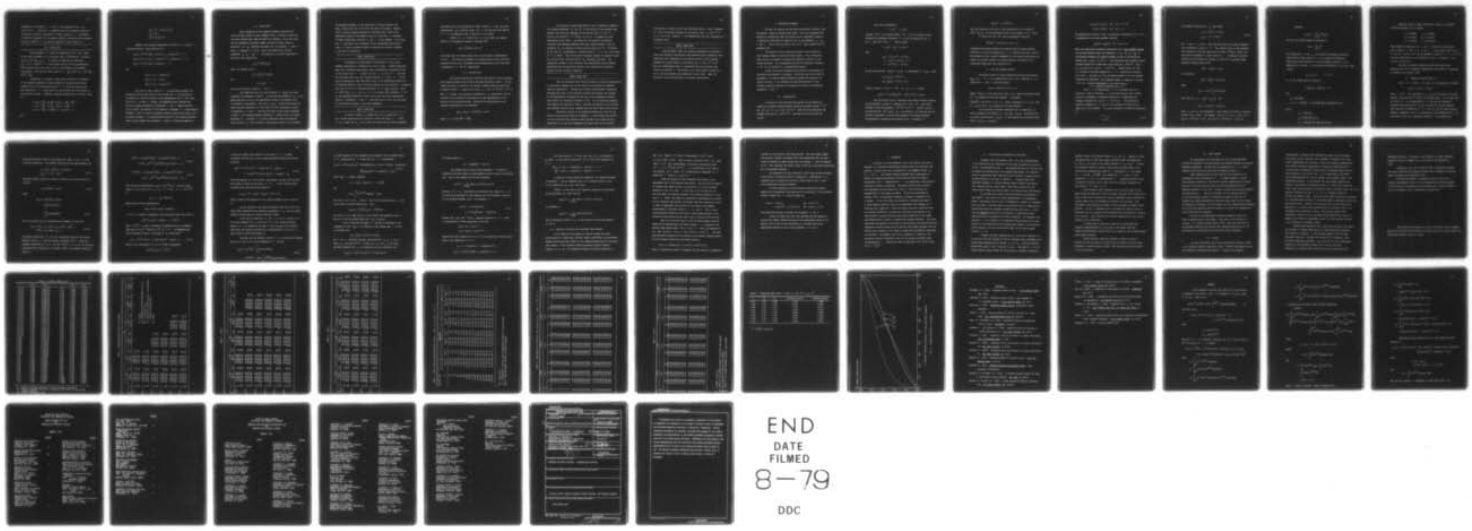
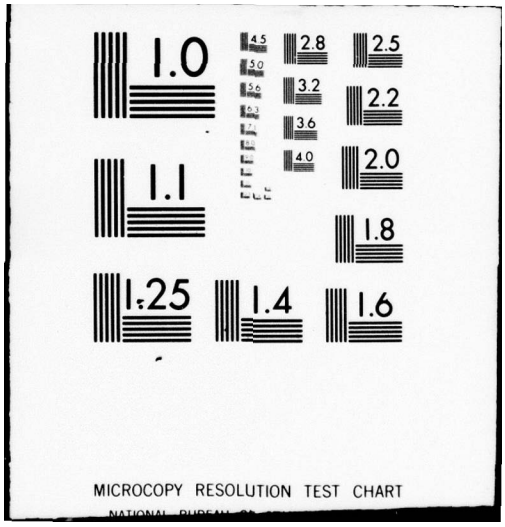
AD-A070 825 MASSACHUSETTS INST OF TECH CAMBRIDGE DEPT OF MATHEMATICS F/G 12/1
SEQUENTIAL MEDICAL TRIALS INVOLVING PAIRED DATA.(U)
MAY 79 H CHERNOFF, A J PETKAU N00014-75-C-0555
UNCLASSIFIED TR-14 NL

| OF |
AD
A070825



END
DATE
FILMED
8-79
DDC


1.0
1.1
1.25
1.4
1.6
1.8
2.0
2.2
2.5
2.8
3.2
3.6
4.0
4.5
5.0
5.6
6.3
7.1
8.0
9.0
10.0
MICROCOPY RESOLUTION TEST CHART

LEVEL II

(1) B.S.

AD A070825

SEQUENTIAL MEDICAL TRIALS INVOLVING PAIRED DATA

BY

HERMAN CHERNOFF
DEPARTMENT OF MATHEMATICS
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

AND

A. JOHN PETKAU
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF BRITISH COLUMBIA

TECHNICAL REPORT NO. 14
MAY 22, 1979

PREPARED UNDER CONTRACT
N00014-75-C-0555 (NR-042-331)
FOR THE OFFICE OF NAVAL RESEARCH

DEPARTMENT OF MATHEMATICS
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
CAMBRIDGE, MASSACHUSETTS

Sequential Medical Trials Involving Paired Data

by

Herman Chernoff
Department of Mathematics
Massachusetts Institute of Technology

and

A. John Petkau
Department of Mathematics
University of British Columbia

-A-

Sequential Medical Trials Involving Paired Data

Herman Chernoff
Department of Mathematics
Massachusetts Institute of Technology

A. John Petkau
Department of Mathematics
University of British Columbia

SUMMARY

A continuous time version of Anscombe's formulation of the problem of comparing two treatments in the context of medical trials is considered and the Bayes sequential procedure is explicitly determined. Various suboptimal procedures are proposed, evaluated and compared to the optimal procedure; the approximation to the optimal procedure proposed by Anscombe turns out to be surprisingly efficient. Comparison with the discrete time version demonstrates that the continuous time version provides accurate approximations for clinical trials involving horizon sizes as small as 100. The optimal procedure determined here provides a design which is relevant for clinical trials involving either normal or Bernoulli responses.

Some key words: clinical trials; optimal stopping; Wiener process; free boundary problem.

AMS 1970 subject classifications: Primary 62L10; secondary 62C10, 62L15.

1. INTRODUCTION

A natural formulation for many statistical problems is that combining Bayesian, sequential and decision theoretic aspects. There exists a body of literature developing an approach to this formulation where sums of successive observations are replaced by a continuous time Wiener process and the heat equation plays a prominent role in the resulting analysis. In particular, optimal procedures may be characterized in terms of solutions of free boundary problems related to the heat equation.

Some of the resulting analysis is quite complicated and more work remains to be done on many aspects of this analysis. It would seem that the formidable appearance of some of the technicalities of this literature has distracted potential users from taking advantage of other aspects which are easy or routine and which contribute clarity, asymptotic results with important theoretical implications, and numerical descriptions of optimal sequential procedures and the resulting risks.

We have become aware of recent work by Siegmund together with some of his colleagues (1978) and by Begg and Mehta (1978) on a sequential medical trials model first proposed by Anscombe (1963). This same model, except for the Bayesian aspect, had been considered earlier by Maurice (1959) in the context of industrial production. We suggest that there is much to be said about this problem in terms of the approach and literature described above. In this paper we develop the model and present these results with two objectives. First, we wish to present the optimal procedure and its properties and show how it compares with several suboptimal procedures. Second, we wish to use the Anscombe model as an illustration to show how such results may be obtained for other similar sequential decision theoretic

problems.

The formulation that we refer to as the Anscombe model for sequential medical trials involving paired data is the following: There is a horizon of N patients to be treated by one of two available treatments. In the initial (experimental) phase, n pairs of patients are treated sequentially, with different treatments randomly assigned to the patients in each pair. The differences, X_i in the values of the outcomes for the i-th $(i=1,2,\cdots,n)$ pair are assumed to be independently and normally distributed with unknown mean μ and known variance σ^2 (and are assumed to be instantaneously available after treatment). After n is selected by some sequential decision rule, the remaining $N-2n$ patients are all assigned to the treatment which is inferred to be superior. The expected loss involved has two components. The first is $E(n|\mu|)$ which represents the expected cost in patient benefit incurred during the experimental period where n of the $2n$ patients treated were assigned the inferior treatment. The second is the expected cost due to the possibility of selecting the wrong treatment for the final stage and thus losing $(N-2n)|\mu|$. An optimal solution exists for this sequential decision problem if it is posed in a Bayesian framework with μ being given a prior normal distribution with mean μ_0 and variance σ_0^2. For any particular specification of the parameters, this Bayes sequential design may be determined, at least in principle, by the backward induction method of dynamic programming. Although such computations have been carried out by Day (1969) and Siegmund (1978) for a few specific cases, for larger values of N this rapidly becomes unfeasible. Further, the method fails to characterize the optimal solution in terms of the parameters involved in an explicit fashion.

Without some theory, the results are less illuminating than they could be.

In Sections 2 and 3 we describe the continuous time model and some of the results that may be derived and computed from it. One result is that the optimal procedure can be described by a single curve in the (t,β) plane where, at any time, t represents that proportion of the potential information that is currently available and β is the current nominal significance level for testing the hypothesis $\mu = 0$. Thus the optimal procedure may be described as a sequence of repeated significance tests with the appropriate significance level varying with the amount of information available. The experimental period terminates when an appropriate significance level is achieved. The fact that the solution can be so described implies that one curve applies for an entire family of problems independent of the parameters μ_0, σ_0^2, σ^2 and N. Moreover, it suggests that this same curve is meaningful for variations of this problem where the data and the priors are not necessarily normal.

Two other results are that the suboptimal procedure proposed by Anscombe (1963) is remarkably close to optimal (this result was first discovered by Siegmund (1978)) and that the expected cost due to ignorance of μ is of the order of magnitude of $(\log N)^2$ which may seen surprisingly small.

2. MODEL AND PROCEDURES

2.1. Posterior probabilities and risks

Upon observing the differences $X_1, X_2, \cdots, X_n$, the posterior distribution of μ becomes $N(Y_n^*, s_n^*)$ where

$$Y_n^* = (\sigma_0^{-2}\mu_0 + \sigma^{-2}\sum_{i=1}^{n} X_i)/(\sigma_0^{-2} + n\sigma^{-2}) ,$$
$$s_n^* = (\sigma_0^{-2} + n\sigma^{-2})^{-1} . \qquad (2.1.1)$$

Further, it can be shown that for $n > m$, the marginal distribution of $Y_n^* - Y_m^*$ (treating μ as random) is $N(0, s_m^* - s_n^*)$ and $Y_n^* - Y_m^*$ is independent of Y_m^*. Thus as sampling continues, Y_n^*, the posterior mean of μ, behaves like a Gaussian process of independent increments starting from $Y_0^* = \mu_0$. Once the experimental phase is concluded, the preferred choice of treatment for the remaining $N - 2n$ patients is clearly decided by the sign of Y_n^*. Also, the expected loss or posterior risk associated with stopping after treating exactly n pairs of patients is $nE(|\mu|) + (N-2n)E[\max\{0, -\operatorname{sgn}(Y_n^*)\mu\}]$, where E represents expectation with respect to the posterior distribution of μ given Y_n^*. It is easy to verify that $E(|\mu|) = 2s_n^{*1/2}\psi(Y_n^* s_n^{*-1/2})$ and $E[\max\{0, -\operatorname{sgn}(Y_n^*)\mu\}] = s_n^{*1/2}\psi(Y_n^* s_n^{*-1/2}) - |Y_n^*|/2$ which leads to the expression $Ns_n^{*1/2}\psi(Y_n^* s_n^{*-1/2}) - (N/2-n)|Y_n^*|$ for the posterior risk where

$$\psi(u) = \phi(u) + u\{\Phi(u) - 1/2\} \qquad (2.1.2)$$

and ϕ and Φ are the standard normal density and cumulative respectively. Using (2.1.1) to substitute for n in terms of s_n^*, the posterior risk can be written as

$$d_1(Y_n^*, s_n^*) = N s_n^{*\frac{1}{2}} \psi(Y_n^* s_n^{*-\frac{1}{2}}) - \sigma^2(s_*^{-1} - s_n^{*-1})|Y_n^*| \tag{2.1.3}$$

where

$$s_*^{-1} = \sigma_0^{-2} + N\sigma^{-2}/2 \tag{2.1.4}$$

may be regarded as the <u>total potential information</u> available for estimating μ. The problem of selecting the best sequential procedure for stopping is equivalent to the optimal stopping problem where the Gaussian process Y_n^* is observed and one selects the stopping time n to minimize the expected risk $Ed_1(Y_n^*, s_n^*)$.

2.2. Continuous time version

A natural approximation to the above problem results if the discrete sequence of partial sums, $\sum_{i=1}^{n} X_i$, is replaced by the continuous time Wiener process, $X(t^*)$, with drift μ and variance σ^2 per unit time. We may write $E\{dX(t^*)\} = \mu dt^*$ and $Var\{dX(t^*)\} = \sigma^2 dt^*$ for $0 \le t^* \le N/2$. The posterior distribution of μ given $X(t')$ for $0 \le t' \le t^*$ is $N(Y^*(s^*), s^*)$ where

$$Y^*(s^*) = \{\sigma_0^{-2}\mu_0 + \sigma^{-2}X(t^*)\}/(\sigma_0^{-2} + t^*\sigma^{-2}),$$

$$s^* = (\sigma_0^{-2} + t^*\sigma^{-2})^{-1}.$$

In parallel with Section 2.1, $Y^*(s^*)$ is a Wiener process with drift 0 and variance 1 per unit in the $-s^*$ scale, and originates at the initial point (y_0^*, s_0^*) where $s_0^* = \sigma_0^2$ and $y_0^* = Y^*(s_0^*) = \mu_0$. Note that as t^* increases from 0 to N/2, s^* <u>decreases</u> from s_0^* to s_*, as defined in (2.1.4).

Once more the posterior risk corresponding to stopping at $(Y^*(s^*), s^*)$ is given by $d_1(Y^*(s^*), s^*)$, as specified in (2.1.3), and our current problem is also an optimal stopping problem which differs from the original problem only in that it involves the continuous time process Y^*. Indeed, the original problem may be regarded as a particular version of the continuous time problem where the possible stopping times in the s^* scale are restricted to only those values of s^* which are of the form $(\sigma_0^{-2} + n\sigma^{-2})^{-1}$. From this point of view it is clear that the expected risk for the discrete time stopping problem is larger than for the continuous version.

For each initial point (y_0^*, s_0^*), there is a corresponding optimal stopping problem. Let $\rho_1(y_0^*, s_0^*)$ be the optimal risk as a function of the initial point. Then $\rho_1(y_0^*, s_0^*) \leq d_1(y_0^*, s_0^*)$ and it pays to continue taking observations if and only if $\rho_1(y_0^*, s_0^*) < d_1(y_0^*, s_0^*)$. But, since $Y^*(s^*)$ is a process of independent increments, $\rho_1(y^*, s^*)$ is the optimal conditional risk given $Y(s^*) = y^*$. Hence one may characterize the optimal procedure as the rule to stop sampling as soon as $\rho_1(Y^*(s^*), s^*) = d_1(Y^*(s^*), s^*)$. Thus the optimal rule is determined by an <u>optimal stopping set</u> S_1^* in the (y^*, s^*) space where $\rho_1 = d_1$ or by its complement the <u>optimal continuation set</u> C_1^* where $\rho_1 < d_1$. This

characterization suggests that in searching for the optimal procedure we restrict attention to procedures defined by stopping or continuation sets in (y^*, s^*) space. In what follows we will refer to the optimal procedure as procedure O .

2.3. Suboptimal procedures

Working in the context of the original discrete time problem, Anscombe (1963, p. 376) proposed a procedure which he argued should be close to optimal. In our notation his procedure, which we shall label A, consists of stopping as soon as

$$1 - \Phi(|Z|) \leq t/2 , \tag{2.3.1}$$

where

$$Z = Y^*(s^*)s^{*-\frac{1}{2}} \tag{2.3.2}$$

and

$$t = (\sigma_0^{-2} + t^*\sigma^{-2})/(\sigma_0^{-2} + N\sigma^{-2}/2) , \qquad 0 \leq t^* \leq N/2 , \tag{2.3.3}$$

is that proportion of the potential information that is currently available.

Also working in the context of the original discrete time problem, Begg and Mehta (1978) proposed a procedure which consists of stopping as soon as there is no fixed additional time of observation which does as well as stopping. We shall call the analogue of this procedure for our continuous time version of the problem procedure F . This procedure has been considered previously in the context of other continuous time optimal stopping problems by Chernoff (1965a) and Petkau (1978).

In the remainder of this paper we shall be particularly concerned with comparisons of procedures O, A and F . However, for discussion purposes it will occasionally be useful to refer to two other procedures. The first of these, which we shall label FS , is the best fixed sampling time procedure given the initial prior. The second, which we shall label ND , is the no decision procedure where the entire set of N patients is evenly divided between the two treatments. The Bayes risk for this procedure can easily be evaluated as $N\sigma_0\psi(\mu_0/\sigma_0)$ which is of the order of magnitude of N .

There are several additional procedures of interest which we shall not discuss in this paper. One of these would be a Wald type procedure which would consist of a stopping region determined by two horizontal lines in the original $(X(t^*),t^*)$ scale. Such procedures have been considered previously in the context of the discrete time problem by Maurice (1959), Colton (1963) and Siegmund (1978).

3. RESULTS

3.1. Stopping rules

The optimal procedure for the continuous time problem may be expressed by the curve (stopping boundary), $\tilde{\beta}_0(t) = 1 - \Phi(\tilde{z}_0(t))$, presented in detail in Table 1. It is interpreted as follows: Let Z and t be as defined in (2.3.2) and (2.3.3). If at any time $\hat{p} = 1 - \Phi(|Z|) \leq \tilde{\beta}_0(t)$, stop taking observations and for the remaining $N-2t^*$ units of time use the treatment in accord with the sign of Y^* . Since the posterior distribution of μ is $N(Y^*(s^*),s^*)$, Z is simply the number of standard deviations that the current Bayes estimate of μ is away from zero and β is the observed P value for a one-sided test of $\mu = 0$ based on the data and the prior. At time t , the curve $\tilde{z}_0(t)$ specifies the number of standard derivations required for stopping and $\tilde{\beta}_0(t)$ is the corresponding nominal significance level. Thus the optimal procedure is a type of repeated significance test with the nominal significance level varying with the amount of information available. Note that as the proportion of information available increases from 0 to 1 , the nominal significance level becomes less stringent, increasing from 0 to 1/2.

Table 1 about here

The definition of procedure F makes it clear that this procedure always prescribes early stopping relative to the optimal procedure. We do not present a tabulation of procedure F since, as we shall demonstrate, this is quite an inefficient procedure. On the other hand, there is no need to tabulate Anscombe's procedure A since the nominal significance level for that procedure is simply $\tilde{\beta}_A(t) = 1 - \Phi(\tilde{z}_A(t)) = t/2$. The stopping

boundaries for procedures 0 , A and F are presented in the (β,t) scale in Fig. 1. Note that F prescribes very early stopping relative to the optimal procedure, particularly for small values of t . Procedure A is much more comparable to the optimal procedure, prescribing later stopping for small values of t and earlier stopping for large values of t .

Fig. 1 about here

While procedure FS cannot be represented in the (β,t) scale as the other procedures, it is relatively easy to characterize. In particular, for the case $\mu_0 = 0$, it consists of sampling the fixed time $t^*_{FS} = N/\{(9 + 4N\sigma_0^2/\sigma^2)^{\frac{1}{2}} + 3\}$ which agrees exactly with the corresponding result obtained by Colton (1963, p.393) in the context of the discrete time problem. Note that for large values of N , $t^*_{FS} \approx N^{\frac{1}{2}}\sigma/\,2\sigma_0 - 3\sigma^2/\,4\sigma_0^2 + 9\sigma^3 N^{-\frac{1}{2}}/\,16\sigma_0^3$.

Although Fig. 1 provides a clear overall picture of the behavior of the stopping boundaries of these procedures, the exact form of these boundaries near the distinguished points $t = 0$ where few patients have been treated and $t = 1$ where nearly all the patients have been treated is of particular interest. Asymptotic expansions indicate that for small values of t

$$- 2 \log t \approx \tilde{z}_0^2 + \log \tilde{z}_0^2 + \log(2\pi) + 2\tilde{z}_0^{-2} + \tilde{z}_0^{-4} + \cdots ,$$

$$- 2 \log t \approx \tilde{z}_A^2 + \log \tilde{z}_A^2 + \log(\pi/2) + 2\tilde{z}_A^{-2} + \cdots ,$$

$$- 2 \log t \approx \tilde{z}_F^2 + 5 \log \tilde{z}_F^2 + 2 + \log(\pi/8) + \cdots ,$$

and

$$\tilde{\beta}_0 \approx t[1 + 3(\log t)^{-2}/4] ,$$

$$\tilde{\beta}_A = t/2 ,$$

$$\tilde{\beta}_F \approx et(-\log t)^2 .$$

Somewhat less important expansions for values of t close to 1 are easily derived. These indicate that

$$\tilde{z}_0(t) \approx (1-t)^{\frac{1}{2}}\{0.7642 + 0.2737(1-t) + 0.1659(1-t)^2 + \cdots\} ,$$

$$\tilde{z}_A(t) \approx (1-t)\{1.2533 + 0.3281(1-t)^2 + 0.1804(1-t)^4 + \cdots\} ,$$

$$\tilde{z}_F(t) \approx (1-t)^{\frac{1}{2}}\{0.3854 + 0.1528(1-t) + \cdots\} ,$$

and

$$\tilde{\beta}_0(t) \approx 0.5 - 0.3049(1-t)^{\frac{1}{2}} ,$$

$$\tilde{\beta}_A(t) = 0.5 - 0.5(1-t) ,$$

$$\tilde{\beta}_F(t) \approx 0.5 - 0.1537(1-t)^{\frac{1}{2}} .$$

The region of small values of t is particularly relevant for problems involving large values of the horizon size N and therefore it is important to note the accuracy of the approximation $\tilde{\beta}_0 \approx t$ for small values of t in Table 1. Further, the expansions above indicate that while the stopping boundaries of procedures 0 and A behave similarly in this region, the behavior of the boundary of procedure F is qualitatively different. Thus, at least for problems involving large horizon sizes, while we expect procedure F to perform poorly relative to the optimal procedure, there is some evidence that procedure A may be a reasonable competitor.

3.2. Bayes risks

While comparison of the stopping boundaries indicates how these procedures differ in their stopping rules, of greater interest are the risks incurred when these procedures are employed. How do the risks for the alternative procedures compare and what are their orders of magnitude? Let R_{1P} represent the Bayes risk for procedure P where P may be 0 (optimal), A, F or FS. These risks depend upon the four parameters μ_0, σ_0, σ and N. For simplicity in tabular presentation we may use the normalization

$$R_{1P} = \sigma^2 \sigma_0^{-1} \phi(z_0) R'_P$$

where R'_P depends only on

$$t_0 = \sigma_0^{-2}/(\sigma_0^{-2} + N\sigma^{-2}/2)$$

and

$$z_0 = \mu_0/\sigma_0$$

which are the initial values of t and Z.

The normalized risks R'_P and the ratios $E = R'_0/R'_P$ for these procedures are presented in Table 2. In each case the proportion r_P of the Bayes risk which is due to the experimental period of the medical trial is tabulated in parenthesis. Each point (t_0, z_0) at which the normalized risks have been tabulated is in the interior of the continuation regions of procedures 0, A and F. Note that for every initial point considered in Table 2, the Anscombe procedure dominates F which in turn, of course, dominates FS. Procedure F is quite inefficient even for relatively large values of t_0 and its behavior deteriorates rapidly as t_0 decreases.

The Anscombe procedure, on the other hand, is highly efficient for R_0'/R_A' varies from 0.94 to 1.00 over the range of the table. Further, while a slightly greater proportion of the Bayes risk is due to the experimental period of the medical trial for A than for 0 , procedure F differs dramatically from both 0 and A in this respect. It is clear that F prescribes early termination of the experimental period which leads to a greatly increased chance of selecting the inferior treatment at the time of termination.

Table 2 about here

The apparent low efficiencies of procedures F and FS are somewhat exaggerated by considering the ratios R_0'/R_P' . Another perspective comes from considering some special cases corresponding to different values of the horizon size N . In Table 3 we let $\mu_0 = 0$, $\sigma_0^2 = \sigma^2 = 1$, and then list the Bayes risks R_{1P} for various values of $N = 2(t_0^{-1} - 1)$. In each case the proportion of the risk which is due to the experimental period and the expected number of pairs of patients sampled are also tabulated. Considering that the risk for the no decision procedure ND is $R_{1ND} = N\sigma_0\psi(z_0)$ which is $N/\sqrt{2\pi}$ when $\mu_0 = 0$ and $\sigma_0 = 1$, it is evident that procedures F and FS are reasonably effective in spite of the fact that they behave poorly relative to procedures 0 and A . From this point of view, the efficiency of the Anscombe procedure is even more remarkable.

Table 3 about here

To return to Table 2, we remark that if t_0 is small and μ is not, not many observations are required to reveal the sign of μ . Hence if t_0 is small and z_0 is not large, the Bayes risks for these procedures

are mainly due to the contribution of small values of μ and the rough approximation $R_{1P} \approx \sigma^2 g(0) R_P'$ where $g(0)$ is the non-zero prior density at $\mu = 0$ is meaningful even if the prior is not normal.

Finally, if t_0 is small and z_0 is not large, the leading behavior of an asymptotic expansion indicates that

$$R_{1O} \approx \sigma^2 \sigma_0^{-1} \phi(z_0)(\log t_0)^2$$

which means that the optimal Bayes risk is of the order of magnitude of $(\log N)^2$. This should be compared to the Bayes risks of the no decision procedure and the best fixed sample size procedure which are of the order of magnitude of N and $N^{\frac{1}{2}}$ respectively.

3.3. Discrete time

The original problem was a discrete time version of the continuous time stopping problem whose solution has been tabulated in Table 1. We may regard the former as one where the possible stopping times are restricted to those for which t takes on the values $t_n = (\sigma_0^{-2} + n\sigma^{-2})/(\sigma_0^{-2} + N\sigma^{-2}/2)$. When N is large the successive differences in t_n become small and the solution of the continuous time problem is a good approximation to the solution of the original problem. Moreover this approximation can be improved considerably by the adjustment

$$\tilde{z}_0'(t) = \tilde{z}_0(t) - kt^{-\frac{1}{2}}(\sigma^2\sigma_0^{-2} + N/2)^{-\frac{1}{2}}$$

where $k = -\zeta(1/2)/\sqrt{2\pi} = 0.5826$.

The accuracy of these approximations can be examined by comparing the stopping boundaries $\tilde{z}_0$ and $\tilde{z}_0'$ to the solution of the discrete time problem. The latter was computed for the specific case $N = 100$, $\sigma^2 = 1$ and for a few values of $a = \sigma^2/N\sigma_0^2 = t_0/2(1-t_0)$ to compare with some tabulations by Day (1969). Grid sizes for the numerical integrations involved in the backward induction were taken refined enough so that the solutions $\hat{z}_0$ are accurate to within relative errors of 0.3%. We present $\tilde{z}_0$, $\tilde{z}_0'$ and $\hat{z}_0$ corresponding to integer values of $t^* = n$ for a few values of a in Table 4. Note the excellent agreement between the adjusted values $\tilde{z}_0'$ and the "exact" discrete results $\hat{z}_0$ throughout the table. The exceptional accuracy of the continuous time approximation may be somewhat surprising since the particular discrete time problem being considered involves at most 50 pairs of patients.

Table 4 about here

That the continuous time solution provides an excellent approximation to the optimal stopping rule for the discrete time problem has important practical implications. Although this solution would provide a suboptimal procedure when used in any particular discrete time problem, the accuracy of the approximation indicates that the use of this suboptimal procedure would result in a negligible increase in risk. For all practical purposes, the stopping rule tabulated in Table 1 provides the solution to all discrete time problems involving large, moderate and even fairly small horizon sizes.

Theoretical results on the difference between the risks for the discrete and continuous cases are incomplete. As has already been pointed out, the continuous time problem is more favorable since sampling may be terminated at any time and consequently the Bayes risk for the discrete

time problem is larger than for the continuous version. Table 5 compares R'_0 with its discrete analogue for the specific case $N = 100$, $\sigma^2 = 1$, $\mu_0 = 0$ for a few values of a and demonstrates that this discrepancy is reasonably small.

Table 5 about here

Begg and Mehta (1978) have recently considered the procedure F within the context of the discrete time problem and have provided a table presenting their computation of the discrete version of the stopping boundary for various values of the horizon size N from 10^2 to 10^6. When the corresponding table is constructed from the continuous time version of procedure F, the entries in all cases agree to within one unit in the last decimal place displayed in their table. Again the continuous time problem provides an excellent approximation.

4. DERIVATION OF RESULTS

We shall now indicate how the results presented in Section 3 may be obtained. There are several major steps. First the continuous time problem is normalized so that its solution can be expressed in terms of a stopping set which does not depend on any of the parameters μ_0, σ_0, N and σ. This is why one table in the (β,t) space suffices for all parameter sets.

The solution of this normalized problem is then related to a free boundary problem involving the heat equation. This relationship leads directly to methods of deriving the asymptotic behavior of the optimal procedure. The corresponding results for suboptimal procedures F and FS are obtained in a straightforward manner.

Finally the relation between the solutions of the discrete and continuous time problems is indicated. This gives rise to an effective adjustment in a simple backward induction algorithm for computing the optimal solution. It is indicated how minor modifications of this backward induction algorithm can be applied to evaluate the suboptimal procedures A and F.

4.1. Normalization

In Section 2, the continuous time version of our problem was stated as an optimal stopping problem involving the Wiener process $Y^*(s^*)$ for $\sigma_0^2 = s_0^* \le s^* \le s_* = (\sigma_0^{-2} + N\sigma^{-2}/2)^{-1}$ where the loss associated with stopping was given by $d_1(Y^*(s^*),s^*)$ and where the initial point was (μ_0,σ_0^2).

Since the transformation

$$Y = aY^* , \quad s = a^2 s^*$$

replaces $Y^*(s^*)$ by a Wiener Process $Y(s)$ (in the -s scale), we may normalize our problem by selecting a so that the terminal value s_* of s^* goes into $a^2 s_* = 1$. Thus we select

$$a = s_*^{-\frac{1}{2}} = (\sigma_0^{-2} + N\sigma^{-2}/2)^{\frac{1}{2}} .$$

Then

$$Y(s) = Y^*(s^*)(\sigma^{-2} + N\sigma^{-2}/2)^{\frac{1}{2}} ,$$

$$s = s^*(\sigma^{-2} + N\sigma^{-2}/2) = t^{-1} ,$$

$$Z = s^{*-\frac{1}{2}} Y^*(s^*) = s^{-\frac{1}{2}} Y(s) ,$$

and the initial point $(y_0^*, s_0^*) = (\mu_0, \sigma_0^2)$ is transformed to (y_0, s_0) where

$$y_0 = \mu_0(\sigma_0^{-2} + N\sigma^{-2}/2)^{\frac{1}{2}} ,$$

$$s_0 = \sigma_0^2(\sigma_0^{-2} + N\sigma^{-2}/2) = t_0^{-1} .$$

Finally, setting $y = ay^*$, $s = a^2 s^*$, for $s_0 \geq s \geq 1$ we have

$$d_1(y^*, s^*) = Na^{-1} s^{\frac{1}{2}} \psi(ys^{-\frac{1}{2}}) - a\sigma^2(1-s^{-1})|y| \equiv d_2(y,s) .$$

Thus the solution of our continuous time optimal stopping problem can be described in terms of a stopping set in the (y,s) space which involves at most one parameter, $Na^{-1}/a\sigma^2 = Na^{-2}\sigma^{-2}$. However we shall now show that the $s^{\frac{1}{2}}\psi$ term in d_2 is irrelevant to the solution and our problem is equivalent to solving the parameter free stopping problem corresponding to minimizing the expected value, on stopping, of

$$d_3(y,s) = -(1-s^{-1})|y| \quad .$$

Then the parameters enter only in the determination of the starting point (y_0,s_0) and the translation back to the original (X,t^*) scale.

To verify that the $s^{\frac{1}{2}}\psi$ term is irrelevant, note that

$$s^{\frac{1}{2}}\psi(ys^{-\frac{1}{2}}) = E\{|Y(0)|/2 \mid Y(s) = y\}$$

represents the expected payoff that would be made if another observer continued to follow the process until $s = 0$ and was then paid $|Y(0)|/2$. The presence of such a payoff independent of our stopping time should not affect the optimal strategy, nor should replacing the payoff by its conditional expectation upon stopping do so.

4.2 The free boundary problem

The Wiener process is closely associated with the heat equation $\frac{1}{2}u_{yy} = u_s$. If S is a stopping set in the (y,s) space and d is a function of y and s, then the function

$$u(y,s) = E\{d(Y(S),S) \mid Y(s) = y\} ,$$

where $(Y(S),S)$ is the first point after (y,s) where the process enters S, is a solution of the heat equation. For example $s^{\frac{1}{2}}\psi(ys^{-\frac{1}{2}}) = E\{|Y(0)/2| \mid Y(s) = y\}$, which corresponds to $d = |y|/2$ and $S = \{(y,s): s=0\}$, is a solution of the heat equation.

For an arbitrary procedure associated with a stopping set S and our stopping cost function d_2, the risk $u(y_0,s_0)$ associated with an initial starting point (y_0,s_0) or, equivalently the conditional risk given $Y(s_0) = y_0$, satisfies

$$\tfrac{1}{2}u_{yy}(y,s) = u_s(y,s) \quad \text{for} \quad (y,s) \in C = S^c ,$$

$$u(y,s) = d_2(y,s) \quad \text{for} \quad (y,s) \in S .$$

The stopping set for which $u(y,s)$ is minimized (uniformly for all (y,s)) is determined by the extra boundary condition

$$u_y(y,s) = d_{2y}(y,s) \quad \text{for} \quad (y,s) \in \partial S .$$

Hence the optimization problem is associated with a <u>free boundary problem</u> for the heat equation where the conditions $u = d_2$ and $u_y = d_{2y}$ on the boundary determine the optimal boundary as well as the associated optimal risk $\rho_2(y,s) = \rho_{\frac{1}{2}}(y^*,s^*)$. This associated free boundary problem provides another way to show that we may ignore the $s^{\frac{1}{2}}\psi(ys^{-\frac{1}{2}})$ term in d_2 and deal with $d_3(y,s) = -(1-s^{-1})|y|$. Observe that since $s^{\frac{1}{2}}\psi$ is a solution of the heat equation for $s > 0$, subtracting a multiple of it from u and from d_2 will not affect whether or not the boundary conditions are satisfied. The interested reader is referred to Chernoff (1972) for a detailed developement of the above results.

4.3. Asymptotic behavior near t = 1.

Near $t = 1$ or equivalently, $s = 1$, our method will be to construct a class of solutions ρ_3 of the heat equation and to modify them and the boundary in successive steps so as to approximate the boundary conditions for $d_3(y,s) = -(1-s^{-1})|y|$. This particular problem is symmetric in y and so it is convenient to deal with the upper half of the boundary. It is also convenient to transform coordinates to

$$\begin{aligned} r &= s - 1, \\ v &= yr^{-\frac{1}{2}} . \end{aligned} \tag{4.3.1}$$

The boundary conditions for ρ_3 then become

$$\rho_3 = - vr^{\frac{1}{2}}(1 - 1/(1+r)) ,$$
$$\rho_{3y} = - (1 - 1/(1+r)) . \qquad (4.3.2)$$

For r near 0 (s near 1), the right hand sides of the above equations may be expanded in powers of $r^{\frac{1}{2}}$. This suggests that we seek separable, even solutions of the heat equation of the form $r^{n/2}H_n(v)$ where $H_n(v)$ may itself be expanded in a power series in v. Solutions of the heat equation of this form may be obtained in terms of the confluent hypergeometric functions which solve

$$H_n''(v) + vH_n'(v) = nH_n(v)$$

or by defining

$$H_n(v) = [G_n(v) + G_n(-v)]/2$$

where

$$G_n(v) = (1/n!) \int_{-v}^{\infty} (v+\varepsilon)^n \phi(\varepsilon) d\varepsilon$$

Note that for $n \geq 1$, $G_n'(v) = G_{n-1}(v)$ and

$$G_n(v) = P_n(v)\Phi(v) + Q_n(v)\phi(v)$$

where P_n and Q_n are polynomials, a few of which are listed in Chernoff and Ray (1965, p.1394). For example, $G_0(v) = \Phi(v)$, $G_1(v) = v\Phi(v) + \phi(v)$, $2G_2(v) = (v^2+1)\Phi(v) + v\phi(v)$ and $6G_3(v) = (v^3+3v)\Phi(v) + (v^2+2)\phi(v)$.

Assuming

$$\rho_3(y,s) = \sum_{n=1}^{\infty} a_n r^{n/2} H_n(v) ,$$

and that the optimal boundary corresponds to

$$\bar{v}(s) = \sum_{n=0}^{\infty} c_n r^n ,$$

the coefficients a_n and c_n are obtained alternately by matching coefficients of equal powers of r in the equations obtained by substituting these expressions into the boundary conditions (4.3.2). Expanding the H_n about c_0 and matching coefficients, we are led to the results:

$$a_1 = a_{2n} = 0 , \qquad n = 1,2,\cdots ,$$

c_0 is the unique positive solution of

$$(1-c_0^2)\phi(c_0) = c_0^3[\Phi(c_0) - \tfrac{1}{2}] ,$$

and

$$c_1 = 2c_0/(c_0^2+5) ,$$

$$c_2 = [1 + (9+c_0^2)c_1^2 - 7(1+c_1^2)(c_0^4+10c_0^2+5)/(c_0^4+16c_0^2+35)]/2c_0 ,$$

while

$$a_3 = -2c_0^3/\phi(c_0) ,$$

$$a_5 = 40c_0^3/(c_0^2+5)\phi(c_0) ,$$

$$a_7 = -1680c_0^3(1+c_1^2)/(c_0^4+16c_0^2+35)\phi(c_0) .$$

Numerical values of these coefficients, correct to the number of digits displayed, are as follows:

$$c_0 = 0.764226, \qquad a_3 = -2.996457,$$
$$c_1 = 0.273718, \qquad a_5 = 10.732217,$$
$$c_2 = -0.107795, \qquad a_7 = -60.547318.$$

These asymptotic expansions for ρ_3 and $\tilde{v}$ translate into equivalent ones in terms of y and s or z and t. In particular, from (4.3.1), $\tilde{v} = \tilde{y}/(s-1)^{\frac{1}{2}} = \tilde{z}_0(t)/(1-t)^{\frac{1}{2}}$ can be expanded in powers of $r = s - 1 = t^{-1} - 1$ and this leads directly to the expansion of $\tilde{z}_0(t)$ in powers of $(1-t)$ as given in Section 3.1.

The type of argument required to prove that these formal expansions are in fact asymptotic expansions for the solution is presented in Breakwell and Chernoff (1964).

4.4. Asymptotic behavior near t = 0

Near $t = 0$ or where s is large, it is convenient to introduce

$$d_4(y,s) = d_3(y,s) + 2s^{\frac{1}{2}}\psi(z)$$

where $z = ys^{-\frac{1}{2}}$. The term added to d_3 is an even solution of the heat equation, and hence does not affect our optimal policy. On the other hand for large $|z|$, it is approximately $|y|$ and has the attribute of cancelling the major part of d_3 with a solution of the heat equation. (The term that we dropped out of d_2 to get d_3 had the same attribute. However it depended, in part, on a parameter of the problem and using d_2 would then involve this unnecessary parameter). This modification

facilitates matching terms on the boundary for large s and z in the following exapansions. The boundary conditions on the upper boundary are

$$\rho_4 = s^{-\frac{1}{2}}z + 2s^{\frac{1}{2}}\{\phi(z) - z[1-\Phi(z)]\} ,$$
$$\rho_{4y} = s^{-1} - 2[1 - \Phi(z)]. \tag{4.4.1}$$

Following Chernoff (1965a) we consider solutions of the heat equation of the form

$$\rho_4 = Ks^{-\frac{1}{2}}\phi(z) + g(z,s) \tag{4.4.2}$$

where

$$g(z,s) = E\{f[Y(0)] \mid Y(s)=y\}$$
$$= \int_{-\infty}^{\infty} f(y+\epsilon s^{\frac{1}{2}})\phi(\epsilon)d\epsilon ,$$
$$= \int_{-\infty}^{\infty} \phi(b-z)f(s^{\frac{1}{2}}b)db , \tag{4.4.3}$$

while the boundary will be represented by an expansion of the form

$$\log s = \tilde{z}^2/2 + a_{-1}\log\tilde{z} + a_0 + a_1\tilde{z}^{-2} + a_2\tilde{z}^{-4} + \cdots ,$$
$$= \tilde{z}^2/2 + a_{-1}\log\tilde{z} + a_0 + \eta . \tag{4.4.4}$$

The unknown coefficients a_i are to be determined along with the unknown symmetric function f from the boundary conditions (4.4.1). Along the boundary represented by (4.4.4), the term $Ks^{-\frac{1}{2}}\phi(z)$ in (4.4.2) is relatively negligible and, proceeding as in Chernoff (1965a), the boundary conditions reduce to

$$f(s^{\frac{1}{2}}z) + (1/2!)sf^{(2)}(s^{\frac{1}{2}}z) + (1\cdot 3/4!)s^2 f^{(4)}(s^{\frac{1}{2}}z) + \cdots$$

$$= s^{-\frac{1}{2}}\{z + 2z^{a_{-1}-2}(2\pi)^{-\frac{1}{2}}e^{a_0}e^{\eta}[1-3z^{-2}+15z^{-4}- \cdots]\} , \tag{4.4.5}$$

$$f^{(1)}(s^{\frac{1}{2}}z) + (1\cdot 3/3!)sf^{(3)}(s^{\frac{1}{2}}z) + (1\cdot 3\cdot 5/5!)s^2 f^{(5)}(s^{\frac{1}{2}}z) + \cdots$$

$$= s^{-1}\{1 - 2z^{a_{-1}-1}(2\pi)^{-\frac{1}{2}}e^{a_0}e^{\eta}[1-z^{-2}+3z^{-4}- \cdots]\} . \tag{4.4.6}$$

With the initial approximation $f_0(x) = 2|x|^{-1}\log x^2$, the main terms match in (4.4.5) and (4.4.6) if $a_{-1} - 1 = 0$ and $-1 = 1 - 2(2\pi)^{-\frac{1}{2}}e^{a_0}$ or

$$a_{-1} = 1 , \quad a_0 = \log(2\pi)/2 .$$

Applying the resulting approximation

$$\log s = \tilde{z}^2/2 + \log \tilde{z} + \log(2\pi)/2$$

to (4.4.5), we obtain a discrepancy (left side minus right side) which is

$$(s^{\frac{1}{2}}z)^{-1}[3 \log z^2 + \log(2\pi) - 1 + 0^*(z^{-2})]$$

where $0^*(z^{-2k})$ is used to represent an expression which is bounded by some power of $\log z^2$ divided by z^{2k} as $z \to \infty$. To compensate for this discrepancy we apply a correction to f_0 which gives

$$f_1(x) = |x|^{-1}\{2 \log x^2 - 3 \log(2 \log x^2) - \log(2\pi) + 1\} . \tag{4.4.7}$$

Applying this approximation to (4.4.6) yields a discrepancy

$$(s^{\frac{1}{2}}z)^{-2}[2(a_1-1) + 0^*(z^{-2})]$$

in which the leading term vanishes if the choice $a_1 = 1$ is made. Proceeding one more step in this iterative process we obtain the further correction

$$f_2(x) = |x|^{-1}[2 \log x^2 - 3 \log(2 \log x^2) - \log(2\pi) + 1 + (2 \log x^2)^{-1}\{9 \log(2 \log x^2) + 3 \log(2\pi) - 4\}] \tag{4.4.8}$$

which when applied to (4.4.6) yields a discrepency, the main part of which can be made to vanish by the choice $a_2 = 1/2$. To this limited number of terms we then have the formal expansion

$$2 \log s \sim \tilde{z}^2 + \log \tilde{z}^2 + \log(2\pi) + 2\tilde{z}^{-2} + \tilde{z}^{-4}$$

which is exactly the expansion for the optimal boundary given in Section 3.1.

One may continue in the fashion described above and a proof that these formal expansions yield asymptotic expansions to ρ_4 and the optimal boundary follows along the lines of Chernoff (1965a).

If one were to substitute f_2 for f in (4.4.3) instead of in the formal expansion for ρ_4 given by (4.4.5), the integral diverges. However, if f is bounded or set equal to zero for some finite interval about zero, the integral converges and is approximated by (4.4.5) for z large, which is where the boundary is.

The Bayes risk for moderate values of z is of particular interest. From (4.4.2) and (4.4.3) and the symmetry of f, we have

$$\rho_4(y,s) = Ks^{-\frac{1}{2}}\phi(z) + \int_{-\infty}^{\infty} \phi(b-z) f(s^{\frac{1}{2}}b)\,db\,, \tag{4.4.9}$$

$$= Ks^{-\frac{1}{2}}\phi(z) + 2\bar{\phi}(z) \int_0^{\infty} e^{-\alpha^2/2} \cosh(bz) f(s^{\frac{1}{2}}b)\,db\,.$$

A careful analysis of this expression (see appendix) with a bounded version of f_2 substituted for f yields that for $s \to \infty$ and bounded z

$$\rho_4(y,s) \approx s^{-\frac{1}{2}}\phi(z)[(\log s)^2 - 3(\log s)\log(\log s) + \{2h(z) + c\}\{2\log s - 3\log(\log s)\} + \frac{9}{4}\{\log(\log s)\}^2 + 4\log(\log s) + O(1)] \tag{4.4.10}$$

where (here γ = Euler's constant)

$$c = 2 - \log 2 - \frac{1}{2}\log \pi - \gamma = 0.15727$$

and

$$h(z) = \int_0^\infty b^{-1}e^{-b^2/2}[\cosh(bz) - 1]db .$$

Note that $h'(z) = [\Phi(z) - 1/2]/\phi(z)$ and, for any fixed value of z, $h(z)$ can be easily evaluated numerically. Since

$$\rho_2(y,s) = \sigma^2\sigma_0^{-1}s^{\frac{1}{2}}\rho_4(y,s) - 2\sigma^2\sigma_0^{-1}\psi(z) ,$$

the result (4.4.9) leads directly to the leading term expansion given in Section 3.2 for the optimal Bayes risk $R_{10} = \rho_2(y_0,s_0)$.

Given an arbitrary procedure P, one may be interested in the (frequentist) risk $R_{1P}(\mu)$ as a function of the unknown mean μ as well as the Bayes risk

$$R_{1P} = \int R_{1P}(\mu)\ \sigma_0^{-1}\phi((\mu-\mu_0)/\sigma_0)d\mu .$$

Let P be O, the Bayes procedure associated with μ_0, σ_0, σ and N where $t_0 = \sigma_0^{-2}/(\sigma_0^{-2} + N\sigma^{-2}/2)$ is small and $z_0 = \mu_0/\sigma_0$ is not large. The above expansion for ρ_4 is consistent with the conjecture that

$$R_{10}(\mu) \approx \sigma^2|\mu|^{-1}\{2\log \mu^2 - 3\log(2\log \mu^2)\}$$

for large values of μ .

4.5. Procedures F and FS

The procedure which we have called procedure F consists of stopping at any point where no fixed sampling time will reduce the expected risk. Thus on the stopping set for this procedure

$$d_3(y,s) = \inf_{1\le s_1\le s} \int d_3(y + (s-s_1)^{\frac{1}{2}}\varepsilon, s_1)\phi(\varepsilon)d\varepsilon \ .$$

Defining $\Delta^2 = s - s_1$ and setting the derivative with respect to s_1 of the right hand side equal to zero leads to the two determining relations for the stopping boundary $\tilde{y}_F(s)$ for procedure F ,

$$\begin{aligned} d_3(y,s) &= -\ 2\Delta(1-s_1^{-1})\psi(y/\Delta) \\ 0 &= \Delta^{-1}(1-s_1^{-1})\phi(y/\Delta) - 2\Delta s_1^{-2}\psi(y/\Delta) \end{aligned} \qquad (4.5.1)$$

Assuming that $\tilde{y}_F/\Delta$ and $\Delta^2/(s-1)$ approach constants as $s \to 1$ leads to consideration of formal expansions of the form

$$\begin{aligned} \tilde{y}_F/\Delta &= a_0 + a_1(s-1) + \cdots \ , \\ s_1 - 1 &= b_0(s-1) + b_1(s-1)^2 + \cdots \ . \end{aligned}$$

Substituting these expansions into (4.5.1) and successively matching terms leads to the expansions as $s \to 1$,

$$s_1 \approx 1 + 0.793165(s-1) - 0.099480(s-1)^2 + \cdots \ ,$$

$$\tilde{y}_F(s) \approx (s-1)^{\frac{1}{2}}[0.385387 + 0.152838(s-1) + \cdots] \ .$$

For large values of s, both $\tilde{y}_F/\Delta$ and s/s_1 are expected to be large. In this case the relations (4.5.1) lead to the expansions as $s \to \infty$,

$$2s/s_1 \approx 2 \log s - 5 \log(2\log s) - \log(\pi/8) + 3 + \cdots ,$$
$$\tilde{z}_F^2 \approx 2 \log s - 5 \log(2\log s) - \log(\pi/8) - 2 + \cdots .$$

In addition to these asymptotic expansions, the stopping boundary of procedure F can be tabulated since it is relatively easy to solve (4.5.1) numerically for $s_1(s)$ and $\tilde{y}_F(s)$.

Finally, we note that we have implicity carried out the analysis for the procedure FS where the risk

$$\rho_{3FS}(y,s) = \inf_{1 \le s_1 \le s} \int d_3(y + (s-s_1)^{\frac{1}{2}}\epsilon, s_1)\phi(\epsilon)d\epsilon$$

is evaluated as

$$\rho_{3FS}(y,s) = \inf_{1 \le s_1 \le s} \{-2\Delta(1-s_1^{-1})\psi(y/\Delta)\}$$

and the appropriate value of s_1 is the solution of the second equation in (4.5.1).

4.6. Relation of discrete and continuous time problems

In the context of the sequential analysis problem of testing for the sign of a normal mean, Chernoff (1965b) established the relationship between the discrete time version of the stopping problem and the continuous time version. If the intervals between successive possible stopping values of s is δs, the difference between the optimal boundaries in

the (y,s) space is $\delta\tilde{y}$ which is approximately $k(\delta s)^{\frac{1}{2}}$ where $k = -\zeta(1/2)/\sqrt{2\pi} = 0.5826$. This is easily translated to the (z,t) space where $t = s^{-1}$ and, corresponding to successive observations where $\delta t^* = 1$, $\delta t = \sigma^{-2}/(\sigma_0^{-2} + N\sigma^{-2}/2)$. It is interesting to note that in the original (x,t^*) space, $\delta\tilde{x}$ is approximately independent of t^* except for t^* close to 0 and $N/2$.

Given the solution to the continuous time problem, we may approximate that of the discrete time version. How do we get the former? In a manner that seems circular, one may obtain it by using backward induction in the (y,s) scale. In fact this is not circular for one may first apply the backward induction with stopping permitted at a fine grid of s values, then apply the correction to approximate the solution of the continuous time problem, and finally apply the correction to that in order to approximate the solution of the discrete time problem corresponding to arbitrary values of σ, σ_0 and N. Thus a single refined backward induction can be used to obtain approximations to the continuous time solution and the whole class of discrete time solutions.

What is of even more value is that the same backward induction method can be employed where the Wiener process Y is replaced by the discrete time process where $Y(s-\eta) = Y(s) \pm \eta^{\frac{1}{2}}$, each with probability 1/2. Here $E\{Y(s-\eta) - Y(s)\} = 0$ and $E\{Y(s-\eta) - Y(s)\}^2 = \eta$. The backward induction evaluation of the solution of the optimal stopping problem for this process involves the very simple equation

$$\rho(y,s+\eta) = \min[d(y,s+\eta),\ \{\rho(y+\eta^{\frac{1}{2}},s) + \rho(y-\eta^{\frac{1}{2}},s)\}/2]$$

which is considerably simpler to implement than the numerical integration

required for the discrete time normal process. With this simple random walk process, Chernoff and Petkau (1976) have established that the same correction applies as before except that the constant k must be replaced by 0.5. This technique was used by Petkau (1978) and it and some elaborations will be discussed elsewhere.

The application of this technique to derive very refined estimates would require an exorbitant amount of computation. Nevertheless, it is extremely easy to program and relatively coarse grids on the s axis yield surprisingly accurate estimates.

A slight modification of this backward induction calculation permits one to approximate the Bayes risk for an arbitrary (not necessarily optimal) stopping set S. Here we simply use

$$\rho(y,s+\eta) = d(y,s+\eta) \qquad \text{for} \quad (y,s+\eta) \in S ,$$
$$= \{\rho(y+\eta^{\frac{1}{2}},s) + \rho(y-\eta^{\frac{1}{2}},s)\}/2 \qquad \text{for} \quad (y,s+\eta) \in C = S^{c} .$$

This method was applied to evaluate the procedures A and F.

Finally we remark that this last technique was also applied to evaluate the (Bayes) expected duration of the experimental period of the medical trial as well as the contribution to the Bayes risk of the experimental period for each of the procedures O, A and F.

5. DISCUSSION

In Section 3 we have presented a few of the results that can be obtained in a relatively straightforward manner using the continuous time approach. We have shown that, irrespective of the parameter values μ_0, σ_0, σ and N, the optimal stopping procedure may be expressed as a sequence of repeated significance tests where the (one-tail) nominal significance level β depends only on t, the proportion of the potential information currently accumulated. The simple function which displays this dependence has been tabulated in Table 1; this facilitates the implementation of the optimal procedure. As was first discovered by Siegmund (1978), the procedure proposed by Anscombe is extremely close to optimal. On the other hand, the procedure of stopping when there is no fixed sampling time that would be an improvement over stopping is a relatively poor competitor. There are simple continuity corrections that are extremely effective in relating the continuous time problem to the (more relevant) original discrete time normal problem and to the random walk problem which is useful in computing the optimal procedure. When these corrections are applied, the solution of the continuous time problem provides remarkably accurate results even for discrete time normal problems involving horizon sizes as small as 100. Orders of magnitude and asymptotic expansions are available. For example, when little information is available, the appropriate nominal significance level for the optimal procedure should be approximately t. Further, the order of magnitude of the optimal Bayes risk is $(\log N)^2$.

5.1. Limitations and variations of the model

Anscombe (1963) and Armitage (1963, 1975) have discussed many of the restrictions and limitations of the model as applied to clinical trials. Two problems are that σ, the standard deviation of the data, may be unknown and the data may not be normally distributed. We conjecture that our significance level table would still provide an (asymptotically) optimal procedure in the case of unknown σ. Arguments have been presented elsewhere in slightly different models (Moriguti and Robbins, 1962; Chernoff and Ray, 1965; Petkau, 1978) which indicate that the Wiener process theory and results apply if we have Bernoulli observations.

In our model the situation involving Bernoulli observations is more complex. With success or failure observations, if circumstances differ considerably from one pair of treatments to another or if matched pairs are used, it may make sense to pay attention only to those pairs of observations where one treatment succeeds and the other fails. In that case the effective horizon size is reduced by a random factor whose expectation must be estimated as the data accumulate. We conjecture that replacing N by an estimate of the effective horizon would lead to good results if our significance level procedure were applied. The correction factor required to go from the Bernoulli model to the Wiener process involves $k = 0.5$.

A somewhat different approach may be a little more effective if it were felt that the outcomes for the two treatments were independent with fixed unknown probabilities of success p_1 and p_2. Here the relevant significance level is that in testing $p_1 = p_2$. The value of t is not clearly defined mainly because the total amount of potential information

depends in part on the unknown values of p_1 and p_2. However it would seem appropriate to make some current estimates based on assuming that p_1 and p_2 are equal to the current estimate of $(p_1 + p_2)/2$. Finally the transition from this discrete time, discrete variable model to the Wiener process resembles one where as s decreases by δ, the change in Y is either $\pm \delta^{\frac{1}{2}}/\sqrt{2p^*}$ with probability p^* or else 0 with probability $1 - 2p^*$ for an appropriate p^*. Reasoning similar to that in Chernoff and Petkau (1976) shows that the continuity correction for this model is $k\delta^{\frac{1}{2}}$ with $k = 0.5$.

While Anscombe (1963) suggested that the general model of this paper treats the physician's ethical problem, Armitage (1963) argued that it fails to deal with the physician's ethical requirement that he provide his current patient with the treatment he believes to be the best. That requirement often frustrates attempts to gain knowledge to benefit future patients. One way to compromise here is to modify the model so that an additional cost is attached to each treatment which the physician believes to be inferior. This "ethical" cost could be proportional to the current estimate of the inferiority.

It should also be noted that unless one is prohibited from reopening experimentation as is the case in this model, some more complex two-armed bandit issues arise as the results of the non-experimental phase become available.

Another problem, indirectly referred to before, is that of dealing with the case where the horizon size is random or not well known. Estimates in related problems (Petkau, 1978) indicated that this problem is not very serious.

5.2. Open problems

The techniques we have discussed are not yet fully developed. It would be desirable to have more general and powerful methods of generating asymptotic expansions. In particular the expansions for $t \to 0$ seem to converge slowly and are not very good for moderate values of t.

These expansions have not been applied to get results for the (frequentist) risks as functions of μ. It seems plausible to expect that suitable inversions of the bilateral Laplace transform may be helpful here but no such work has been done for either the expansions or the numerical estimates of the risks. Presently one must solve the heat equation for each value of μ. While our procedure for the numerical solution of the free boundary problem is convenient and seems to work well for computing the optimal procedure and evaluating Bayes risks for the optimal and suboptimal procedures, one would expect that more sophisticated techniques for solving heat equations would be required to evaluate the (frequentist) risks through the heat equation.

We have not considered Wald sequential probability ratio test type of stopping rules in this paper. Such rules are not difficult to evaluate. In a related problem (Petkau, 1978), such rules were found to be reasonably efficient when applied with thresholds appropriate for the particular prior distribution and horizon size being considered.

5.3. History

The first substantial use of this model was due to Maurice (1959) in an industrial sampling inspection context. Minimax fixed sample and Wald type sequential procedures were applied. In papers which appeared

simultaneously, Anscombe (1963) proposed this model and his procedure for clincial trials, Armitage (1963) responded to Anscombe, and Colton (1963) supplemented some of Maurice's ideas with Bayesian considerations. Both Maurice and Colton approximated the operating characteristics of the Wald type of procedure by assuming that the expected sample size was small compared to the horizon size. All of these authors used Wiener process approximations but did not attempt a free boundary attack on the optimization problem. In the meantime the free boundary approach was applied to the problem of deciding the sign of a normal mean by Chernoff (1961, 1965a, 1965b), Breakwell and Chernoff (1964), and Bather (1962) and to the problem of deciding the sign of $p - 1/2$ in a binomial problem by Moriguti and Robbins (1962). Chernoff and Ray (1965) applied this technique to a one-armed bandit problem in a rectified sampling inspection context with fixed horizon size. Chernoff (1967) pointed out how those results were applicable to a clinical trials problem where the efficacy of one of two treatments being considered is known. More recently Petkau (1978) considered a variation of this one-armed bandit problem where there is a sampling cost during the experimental phase (an issue raised by Armitage in his discussion of Anscombe's paper).

Recently Begg and Mehta (1978) rediscovered the model and studied the procedure of stopping when there is no fixed sample size procedure that improves on stopping. In the meantime, Siegmund (1978) has made substantial advances in bounding and estimating the frequentist risks of several suboptimal procedures. He pointed out to us the approximate optimality of the Anscombe procedure and it was his work that stimulated us to explore what could be done by the techniques we had

developed previously. At present a major weakness of these techniques compared to those of Siegmund are in the attack on the frequentist risks.

Incidentally, for lack of space we have not said anything about the elegant and effective techniques Bather (1962) developed to bound the optimal procedure and its Bayesian risks.

The interested reader is referred to Chernoff (1972) for a detailed developement of many of the techniques which have been employed in this paper and illustrations of their application to other problems as well as an extensive list of references to related work.

This research was supported in part by the office of Naval Research under Contract N0001-75-C-0555 (NR-042-331) and in part by a research grant from the National Research Council of Canada.

Table 1. Optimal stopping rule*

t	$\tilde{z}_0(t)$	$\tilde{\beta}_0(t)$	t	$\tilde{z}_0(t)$	$\tilde{\beta}_0(t)$
1.0(-6)	4.747	1.033(-6)	0.16	1.234	0.1086
2.0(-6)	4.606	2.054(-6)	0.17	1.208	0.1135
3.0(-6)	4.520	3.095(-6)	0.18	1.183	0.1185
4.0(-6)	4.460	4.101(-6)	0.19	1.158	0.1234
5.0(-6)	4.412	5.115(-6)	0.20	1.136	0.1281
6.0(-6)	4.372	6.155(-6)	0.22	1.092	0.1374
7.0(-6)	4.339	7.168(-6)	0.24	1.052	0.1464
8.0(-6)	4.310	8.185(-6)	0.26	1.015	0.1550
9.0(-6)	4.283	9.219(-6)	0.28	0.980	0.1636
1.0(-5)	4.261	1.020(-5)	0.30	0.947	0.1718
2.0(-5)	4.102	2.045(-5)	0.32	0.916	0.1799
3.0(-5)	4.006	3.083(-5)	0.34	0.886	0.1877
4.0(-5)	3.939	4.094(-5)	0.36	0.858	0.1955
5.0(-5)	3.884	5.135(-5)	0.38	0.830	0.2032
6.0(-5)	3.838	6.198(-5)	0.40	0.804	0.2107
7.0(-5)	3.801	7.213(-5)	0.42	0.779	0.2180
8.0(-5)	3.768	8.221(-5)	0.44	0.754	0.2253
9.0(-5)	3.738	9.262(-5)	0.46	0.731	0.2325
1.0(-4)	3.711	1.031(-4)	0.48	0.707	0.2397
2.0(-4)	3.530	2.077(-4)	0.50	0.684	0.2469
3.0(-4)	3.422	3.112(-4)	0.52	0.662	0.2540
4.0(-4)	3.342	4.166(-4)	0.54	0.640	0.2611
5.0(-4)	3.279	5.201(-4)	0.56	0.619	0.2680
6.0(-4)	3.227	6.263(-4)	0.58	0.598	0.2751
7.0(-4)	3.183	7.297(-4)	0.60	0.577	0.2821
8.0(-4)	3.143	8.364(-4)	0.62	0.556	0.2891
9.0(-4)	3.107	9.437(-4)	0.64	0.536	0.2961
0.001	3.077	0.001047	0.66	0.515	0.3033
0.002	2.865	0.002088	0.68	0.495	0.3104
0.003	2.735	0.003122	0.70	0.474	0.3176
0.004	2.641	0.004131	0.72	0.454	0.3249
0.005	2.566	0.005138	0.74	0.433	0.3324
0.006	2.505	0.006127	0.76	0.413	0.3400
0.007	2.452	0.007109	0.78	0.391	0.3478
0.008	2.405	0.008079	0.80	0.370	0.3557
0.009	2.364	0.009041	0.82	0.348	0.3639
0.01	2.326	0.009997	0.84	0.325	0.3725
0.02	2.074	0.01905	0.86	0.302	0.3814
0.03	1.920	0.02740	0.88	0.277	0.3908
0.04	1.808	0.03529	0.90	0.251	0.4009
0.05	1.720	0.04268	0.92	0.223	0.4119
0.06	1.646	0.04984	0.94	0.191	0.4241
0.07	1.584	0.05659	0.95	0.174	0.4309
0.08	1.529	0.06307	0.96	0.155	0.4383
0.09	1.480	0.06939	0.97	0.134	0.4467
0.10	1.437	0.07542	0.98	0.109	0.4566
0.11	1.396	0.08135	0.99	0.077	0.4694
0.12	1.359	0.08703	0.995	0.054	0.4784
0.13	1.325	0.09267	0.999	0.024	0.4904
0.14	1.293	0.09801	0.9995	0.017	0.4932
0.15	1.263	0.1033	1.0000	0.000	0.5000

*t = currently available proportion of total potential information
$\tilde{z}_0$ = number of standard deviations of posterior mean from 0 at the optimal stopping boundary
$\tilde{\beta}_0$ = nominal significance level

Table 2. Normalized Bayes risks of procedures

		z_0 = 0.0		0.5		1.0		1.5		2.0		3.0	
t_0	P	$R'_P(r_P)$	E_P	$R'_P(r_P)$	E_P	$R'_P(r_P)$	E_P	$R'_P(r_P)$	E_P	$R'_P(r_P)$	E_P	$R'_P(r_P)$	E_P
1(-1)	O	4.47(.61)		4.60(.57)									
	A	4.53(.69)	.99	4.68(.67)	.98								
	F	5.97(.18)	.75	6.51(.11)	.71								
	FS	6.38(.47)	.70	6.96(.43)	.66								
5(-2)	O	6.39(.63)		6.66(.61)									
	A	6.53(.73)	.98	6.85(.73)	.97								
	F	9.17(.18)	.70	10.06(.14)	.66								
	FS	10.11(.48)	.63	11.15(.45)	.60								
2(-2)	O	9.53(.67)		9.98(.65)		11.44(.62)							
	A	9.82(.78)	.97	10.32(.78)	.97	11.96(.77)	.96						
	F	15.20(.17)	.63	16.76(.14)	.60	22.50(.06)	.51						
	FS	17.47(.49)	.55	19.46(.47)	.51	25.06(.42)	.46						
1(-2)	O	12.42(.68)		13.07(.68)		15.17(.66)							
	A	12.86(.80)	.97	13.58(.80)	.96	15.91(.80)	.95						
	F	21.38(.17)	.58	23.57(.14)	.55	31.70(.08)	.48						
	FS	25.77(.49)	.48	28.81(.48)	.45	37.99(.45)	.40						
5(-3)	O	15.82(.70)		16.65(.70)		19.51(.69)		25.46(.65)					
	A	16.44(.82)	.96	17.33(.82)	.96	20.45(.82)	.95	27.03(.82)	.94				
	F	29.62(.16)	.53	32.66(.14)	.51	44.06(.09)	.44	73.87(.02)	.34				
	FS	37.49(.49)	.42	42.03(.49)	.40	56.21(.46)	.35	80.01(.38)	.32				
2(-3)	O	21.17(.72)		22.31(.72)		26.34(.72)		34.80(.70)					
	A	22.05(.84)	.96	23.27(.84)	.96	27.60(.85)	.95	36.78(.85)	.95				
	F	44.16(.15)	.48	48.59(.13)	.46	65.27(.10)	.40	108.46(.04)	.32				
	FS	60.74(.50)	.35	68.25(.49)	.33	92.33(.48)	.29	137.65(.44)	.25				

$z_0 = \mu_0/\sigma_0$

t_0 = proportion of total information in prior

$R'_P = \sigma_0 \sigma^{-2} \phi(z_0)^{-1} R_{1P}$

r_P = proportion of risk R_{1P} due to experimental period of clinical trial

$E_P = R'_0/R'_P$

Table 2. Cont'd

		z_0											
		0.0		0.5		1.0		1.5		2.0		3.0	
t_0	P	$R'_P(r_P)$	E_P	$R'_P(r_P)$	E_P	$R'_P(r_P)$	E_P	$R'_P(r_P)$	E_P	$R'_P(r_P)$	E_P	$R'_P(r_P)$	E_P
1(-3)	O	25.92(.74)		27.28(.74)		32.18(.74)		42.52(.73)					
	A	27.02(.85)	.96	28.46(.86)	.96	33.68(.86)	.96	44.77(.86)	.95				
	F	58.93(.14)	.44	64.67(.13)	.42	86.20(.10)	.37	141.34(.06)	.30				
	FS	86.94(.50)	.30	97.80(.49)	.28	133.02(.49)	.24	202.15(.46)	.21				
5(-4)	O	31.33(.75)		32.96(.75)		38.81(.76)		51.56(.75)		80.48(.74)			
	A	32.66(.87)	.96	34.38(.87)	.96	40.56(.87)	.96	54.11(.88)	.95	85.11(.87)	.95		
	F	77.51(.13)	.40	84.91(.12)	.39	112.84(.10)	.34	185.10(.06)	.28	387.56(.02)	.21		
	FS	123.99(.50)	.25	139.58(.50)	.24	190.55(.49)	.20	293.17(.47)	.18	475.22(.42)	.17		
2(-4)	O	39.52(.77)		41.57(.77)		48.54(.78)		64.44(.78)		100.37(.77)			
	A	41.19(.88)	.96	43.33(.88)	.96	50.64(.88)	.96	67.38(.89)	.96	105.45(.89)	.95		
	F	109.68(.12)	.36	119.86(.11)	.35	157.42(.10)	.31	254.72(.07)	.25	523.79(.03)	.19		
	FS	197.50(.50)	.20	222.49(.50)	.19	304.69(.49)	.16	473.61(.48)	.14	795.16(.46)	.13		
1(-4)	O	46.56(.78)		48.82(.78)		56.59(.79)		75.27(.79)		117.10(.79)			
	A	48.48(.89)	.96	50.84(.89)	.96	58.96(.89)	.96	78.52(.90)	.96	122.51(.90)	.96		
	F	140.72(,11)	.33	153.33(.11)	.32	200.07(.09)	.28	322.89(.07)	.23	660.74(.04)	.18		
	FS	280.34(.50)	.16	315.92(.50)	.15	433.32(.50)	.13	676.87(.49)	.11	1154.13(.47)	.10		
5(-5)	O	54.33(.79)		56.97(.80)		65.94(.80)		86.67(.81)		134.85(.81)			
	A	56.52(.89)	.96	59.26(.90)	.96	68.60(.90)	.96	90.20(.91)	.96	140.60(.91)	.96		
	F	179.03(.11)	.30	194.69(.10)	.29	252.20(.09)	.26	400.12(.07)	.22	806.18(.04)	.17		
	FS	397.50(.50)	.14	448.05(.50)	.13	615.23(.50)	.11	964.29(.49)	.09	1661.08(.48)	.08		
2(-5)	O	65.78(.81)		68.79(.81)		79.16(.81)		103.29(.82)		160.10(.83)			
	A	68.33(.90)	.96	71.45(.90)	.96	82.19(.91)	.96	107.23(.91)	.96	166.20(.92)	.96		
	F	242.27(.10)	.27	262.52(.10)	.26	337.27(.08)	.23	530.52(.07)	.19	1062.54(.04)	.15		
	FS	629.95(.50)	.10	710.21(.50)	.10	976.15(.50)	.08	1534.53(.50)	.07	2666.24(.49)	.06		

Table 2. Cont'd

		z_0 = 0.0		0.5		1.0		1.5		2.0		3.0	
t_0	P	$R'_P(r_P)$	E_P	$R'_P(r_P)$	E_P	$R'_P(r_P)$	E_P	$R'_P(r_P)$	E_P	$R'_P(r_P)$	E_P	$R'_P(r_P)$	E_P
1(-5)	O	75.34(.82)		78.54(.82)		90.34(.82)		116.50(.83)		181.02(.84)		854.89(.81)	
	A	78.17(.91)	.96	81.48(.91)	.96	93.67(.91)	.96	120.72(.92)	.97	187.61(.92)	.96	894.00(.92)	.96
	F	301.71(.09)	.25	325.73(.09)	.24	415.87(.08)	.22	643.39(.07)	.18	1272.34(.05)	.14	11362.51(.01)	.08
	FS	891.93(.50)	.08	1005.66(.50)	.08	1382.90(.50)	.07	2177.15(.50)	.05	3798.70(.49)	.05	14214.98(.41)	.06
5(-6)	O	85.70(.83)		89.34(.83)		101.86(.83)		131.13(.84)		200.88(.85)		959.19(.83)	
	A	88.82(.91)	.96	92.57(.91)	.97	105.47(.92)	.97	135.67(.92)	.97	207.76(.93)	.97	997.86(.93)	.96
	F	372.36(.09)	.23	401.36(.09)	.22	507.90(.08)	.20	781.62(.06)	.17	1530.36(.05)	.13	13495.52(.01)	.07
	FS	1262.41(.50)	.07	1423.49(.50)	.06	1958.13(.50)	.05	3085.94(.50)	.04	5400.07(.49)	.04	21108.91(.44)	.05
2(-6)	O	100.63(.84)		104.43(.84)		118.54(.84)		151.93(.85)		228.83(.86)		1114.96(.85)	
	A	104.13(.92)	.97	108.03(.92)	.97	122.54(.92)	.97	156.90(.93)	.97	236.15(.94)	.97	1155.96(.94)	.96
	F	589.10(.08)	.17	630.69(.08)	.17	786.12(.07)	.15	1178.37(.06)	.13	2254.35(.05)	.10	19233.94(.02)	.06
	FS	1997.50(.50)	.05	2252.52(.50)	.05	3099.46(.50)	.04	4889.09(.50)	.03	8577.20(.50)	.03	34689.64(.47)	.03
1(-6)	O	112.88(.84)		117.08(.84)		132.46(.85)		167.30(.86)		253.58(.87)		1228.28(.86)	
	A	116.67(.92)	.97	120.98(.92)	.97	136.77(.93)	.97	172.56(.93)	.97	261.30(.94)	.97	1268.96(.94)	.97
	F	589.10(.08)	.19	630.69(.08)	.19	786.12(.07)	.17	1178.37(.06)	.14	2254.35(.05)	.11	19233.94(.02)	.06
	FS	2825.93(.50)	.04	3186.81(.50)	.04	4385.71(.50)	.03	6921.18(.50)	.02	12157.66(.50)	.02	49957.03(.48)	.02

Table 3. Bayes risks and expected sampling times --- Case $\mu_0 = 0$, $\sigma_0^2 = 1$, $\sigma^2 = 1$*.

Horizon Size	Procedure							
	O		A		F		FS	
N	$R_{10}(r_0)$	$E_0(n)$	$R_{1A}(r_A)$	$E_A(n)$	$R_{1F}(r_F)$	$E_F(n)$	$R_{1FS}(r_{FS})$	$E_{FS}(n)$
18	1.78(.61)	1.76	1.81(.69)	2.02	2.38(.18)	0.63	2.55(.47)	1.50
38	2.55(.63)	2.91	2.61(.73)	3.46	3.66(.18)	1.02	4.03(.48)	2.42
98	3.80(.66)	5.31	3.92(.78)	6.49	6.06(.17)	1.77	6.97(.49)	4.26
198	4.95(.68)	8.11	5.13(.80)	10.05	8.53(.17)	2.63	10.28(.49)	6.33
398	6.31(.70)	12.19	6.56(.82)	15.24	11.82(.16)	3.83	14.96(.49)	9.25
998	8.45(.72)	20.53	8.80(.84)	25.85	17.62(.15)	6.22	24.23(.50)	15.06
1,998	10.34(.74)	30.15	10.78(.85)	38.08	23.51(.14)	8.91	34.68(.50)	21.61
3,998	12.50(.75)	44.00	13.03(.87)	55.65	30.92(.13)	12.73	49.46(.50)	30.87
9,998	15.77(.77)	71.90	16.43(.88)	91.00	43.78(.12)	20.32	78.79(.50)	49.25
19,998	18.57(.78)	103.73	19.34(.89)	131.27	56.14(.11)	28.94	111.84(.50)	69.96
39,998	21.67(.79)	149.08	22.55(.89)	188.58	71.42(.11)	41.16	158.58(.50)	99.25
99,998	26.24(.81)	239.73	27.26(.90)	303.00	96.65(.10)	65.56	251.32(.50)	157.36
199,998	30.06(.82)	342.35	31.19(.91)	432.42	120.37(.09)	93.18	355.83(.50)	222.86
399,998	34.19(.83)	487.99	35.43(.91)	615.99	148.55(.09)	132.40	503.63(.50)	315.48
999,998	40.15(.84)	777.63	41.54(.92)	980.88	193.91(.08)	210.48	796.89(.50)	499.25
1,999,998	45.03(.84)	1104.72	46.54(.92)	1392.81	235.02(.08)	298.77	1127.38(.50)	706.36

*R_{1P} = Bayes risk

r_P = proportion of risk R_{1P} due to experimental period of clinical trial

$E_P(n)$ = expected number of pairs of patients sampled

Table 4. Approximations to discrete time stopping rules --- Case $N = 100$, $\sigma^2 = 1$*.

a	0.020			0.050			0.125			0.500		
σ_0^2	0.50			0.20			0.08			0.02		
n	$\tilde{z}_0$	$\tilde{z}'_0$	$\hat{z}_0$	$\tilde{z}_0$	$\tilde{z}'_0$	$\hat{z}_0$	$\tilde{z}_0$	$\tilde{z}'_0$	$\hat{z}_0$	$\tilde{z}_0$	$\tilde{z}'_0$	$\hat{z}_0$
0	1.823	1.412	1.426	1.477	1.216	1.218	1.135	0.971	0.969	0.684	0.602	0.600
1	1.663	1.327	1.333	1.399	1.162	1.163	1.101	0.942	0.941	0.673	0.592	0.590
2	1.545	1.254	1.258	1.334	1.113	1.114	1.068	0.915	0.913	0.662	0.581	0.580
3	1.452	1.192	1.194	1.276	1.070	1.070	1.037	0.889	0.888	0.651	0.571	0.569
4	1.376	1.138	1.139	1.225	1.031	1.030	1.008	0.865	0.863	0.640	0.561	0.559
5	1.310	1.090	1.090	1.178	0.994	0.993	0.980	0.841	0.839	0.630	0.551	0.549
6	1.252	1.046	1.045	1.135	0.960	0.959	0.954	0.818	0.816	0.619	0.541	0.539
7	1.200	1.006	1.005	1.096	0.928	0.926	0.928	0.796	0.794	0.608	0.531	0.529
8	1.153	0.969	0.968	1.059	0.897	0.896	0.904	0.775	0.773	0.598	0.521	0.519
9	1.110	0.935	0.933	1.025	0.869	0.868	0.880	0.755	0.753	0.587	0.511	0.510
10	1.070	0.902	0.901	0.993	0.842	0.840	0.858	0.735	0.733	0.577	0.501	0.500
11	1.034	0.872	0.870	0.962	0.816	0.814	0.836	0.716	0.714	0.566	0.492	0.490
12	0.999	0.843	0.841	0.933	0.791	0.790	0.815	0.697	0.695	0.556	0.482	0.480
13	0.966	0.815	0.814	0.905	0.768	0.766	0.794	0.678	0.676	0.546	0.472	0.471
14	0.935	0.789	0.787	0.878	0.745	0.743	0.774	0.660	0.658	0.535	0.463	0.461
15	0.905	0.764	0.762	0.853	0.722	0.720	0.754	0.643	0.641	0.525	0.453	0.451
16	0.877	0.740	0.738	0.828	0.701	0.699	0.735	0.626	0.624	0.515	0.443	0.442
17	0.850	0.716	0.715	0.804	0.680	0.678	0.716	0.609	0.607	0.505	0.434	0.432
18	0.824	0.694	0.692	0.781	0.659	0.657	0.698	0.592	0.590	0.495	0.424	0.422
19	0.799	0.672	0.670	0.759	0.640	0.637	0.680	0.576	0.574	0.484	0.414	0.412
20	0.775	0.651	0.648	0.736	0.620	0.618	0.662	0.560	0.557	0.474	0.405	0.403
21	0.751	0.630	0.628	0.715	0.601	0.599	0.645	0.544	0.541	0.464	0.395	0.393
22	0.728	0.609	0.607	0.694	0.582	0.580	0.628	0.528	0.526	0.454	0.385	0.383
23	0.706	0.590	0.587	0.674	0.564	0.561	0.610	0.513	0.510	0.444	0.375	0.373
24	0.684	0.570	0.568	0.654	0.546	0.543	0.593	0.497	0.494	0.433	0.365	0.363
25	0.663	0.551	0.548	0.635	0.528	0.525	0.577	0.482	0.479	0.423	0.356	0.354
26	0.642	0.532	0.529	0.615	0.510	0.508	0.560	0.466	0.464	0.412	0.346	0.344
27	0.621	0.513	0.510	0.596	0.493	0.490	0.544	0.451	0.448	0.402	0.336	0.333
28	0.601	0.494	0.492	0.577	0.475	0.472	0.527	0.436	0.433	0.391	0.325	0.323
29	0.581	0.476	0.473	0.558	0.458	0.455	0.511	0.420	0.418	0.381	0.315	0.313

Table 4. cont'd

a	0.020			0.050			0.125			0.500		
σ_0^2	0.50			0.20			0.08			0.02		
n	$\tilde{z}_0$	$\tilde{z}_0'$	$\hat{z}_0$	$\tilde{z}_0$	$\tilde{z}_0'$	$\hat{z}_0$	$\tilde{z}_0$	$\tilde{z}_0'$	$\hat{z}_0$	$\tilde{z}_0$	$\tilde{z}_0'$	$\hat{z}_0$
30	0.561	0.458	0.455	0.539	0.441	0.438	0.495	0.405	0.402	0.370	0.305	0.303
31	0.541	0.439	0.437	0.521	0.423	0.421	0.478	0.390	0.387	0.359	0.294	0.292
32	0.521	0.421	0.418	0.502	0.406	0.403	0.462	0.375	0.372	0.348	0.284	0.281
33	0.502	0.403	0.400	0.483	0.389	0.386	0.445	0.359	0.356	0.337	0.273	0.270
34	0.482	0.385	0.382	0.465	0.372	0.368	0.429	0.344	0.341	0.325	0.262	0.259
35	0.463	0.367	0.363	0.446	0.354	0.351	0.412	0.328	0.325	0.314	0.250	0.248
36	0.443	0.348	0.345	0.428	0.337	0.333	0.396	0.312	0.309	0.302	0.239	0.236
37	0.423	0.330	0.326	0.409	0.319	0.315	0.379	0.296	0.292	0.290	0.227	0.224
38	0.403	0.311	0.307	0.389	0.301	0.297	0.361	0.279	0.276	0.277	0.215	0.212
39	0.382	0.291	0.288	0.370	0.282	0.278	0.344	0.262	0.259	0.264	0.203	0.200
40	0.362	0.272	0.268	0.350	0.263	0.259	0.325	0.245	0.241	0.251	0.190	0.187
41	0.340	0.251	0.247	0.329	0.244	0.239	0.307	0.227	0.223	0.237	0.176	0.173
42	0.318	0.230	0.226	0.308	0.223	0.219	0.287	0.208	0.204	0.223	0.162	0.159
43	0.295	0.208	0.204	0.286	0.202	0.197	0.267	0.189	0.184	0.208	0.147	0.144
44	0.271	0.185	0.180	0.263	0.180	0.175	0.246	0.168	0.163	0.192	0.131	0.127
45	0.246	0.161	0.155	0.238	0.156	0.150	0.223	0.146	0.141	0.174	0.114	0.110
46	0.218	0.134	0.128	0.212	0.130	0.124	0.198	0.122	0.116	0.155	0.096	0.091
47	0.188	0.104	0.097	0.182	0.101	0.094	0.171	0.095	0.088	0.134	0.075	0.069
48	0.152	0.070	0.061	0.148	0.068	0.059	0.138	0.063	0.056	0.109	0.050	0.044

*$a = \sigma^2/N\sigma_0^2 = t_0/2(1-t_0)$

$\tilde{z}_0$ = value of z at optimal boundary; continuous time problem

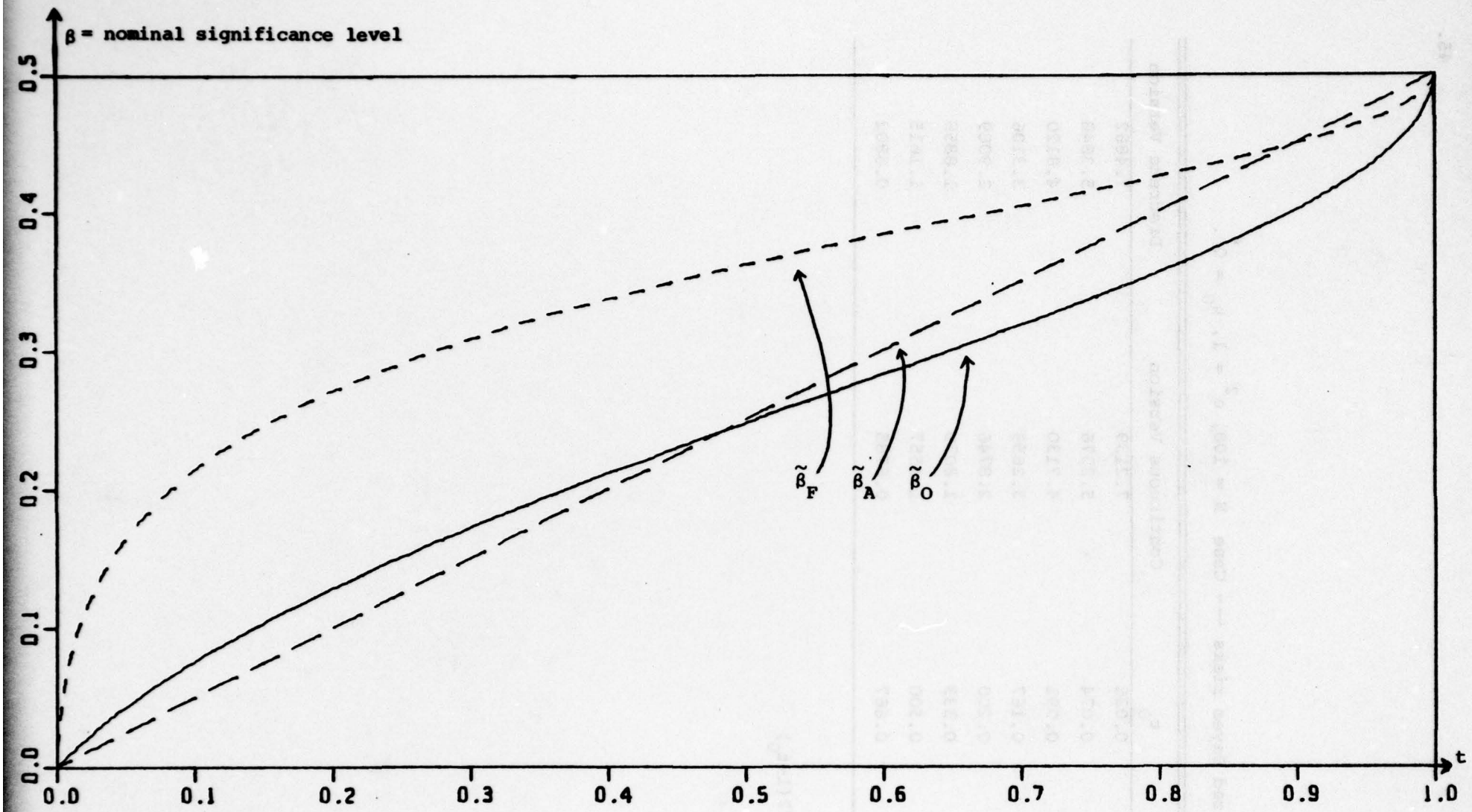
$\tilde{z}_0'$ = $\tilde{z}_0$ corrected for discrete time problem

$\hat{z}_0$ = value of z at optimal boundary; discrete time problem

Table 5. Normalized Bayes risks --- Case $N = 100$, $\sigma^2 = 1$, $\mu_0 = 0$*.

a	σ_0^2	t_0	Continuous Version	Discrete Version
0.020	0.50	0.038	7.2139	7.4862
0.040	0.25	0.074	5.2576	5.3848
0.050	0.20	0.091	4.7130	4.8120
0.100	0.10	0.167	3.2659	3.3106
0.125	0.08	0.200	2.8746	2.9089
0.250	0.04	0.333	1.8079	1.8858
0.500	0.02	0.500	1.1557	1.1615
1.000	0.01	0.667	0.6785	0.6802

*$a = \sigma^2/N\sigma_0^2 = t_0/2(1-t_0)$



t = currently available proportion of total potential information

Fig. 1. Stopping boundaries of procedures

REFERENCES

Anscombe, F.J. (1963). Sequential medical trials. J. Am. Statist. Assoc. 58, 365-83.

Armitage, P. (1963). Sequential medical trials. Some comments on F.J. Anscombe's paper. J. Am. Statist. Assoc. 58, 384-7.

Armitage, P. (1975). Sequential Medical Trials, 2nd edition. Oxford: Blackwell.

Bather, J. (1962). Bayes procedures for deciding the sign of a normal mean. Proc. Cambridge Philos. Soc. 58, 599-620.

Begg, C.B. and Mehta, C.R. (1978). Sequential analysis of comparative clinical trials. Biometrika, to appear.

Breakwell, J. and Chernoff, H. (1964). Sequential tests for the mean of a normal distribution II. Ann. Math. Statist. 35, 162-73.

Chernoff, H. (1961). Sequential tests for the mean of a normal distribution. Proc. 4th Berkeley Symp. 1, 79-91.

Chernoff, H. (1965a). Sequential tests for the mean of a normal distribution III. Ann. Math. Statist. 36, 28-54.

Chernoff, H. (1965b). Sequential tests for the mean of a normal distribution IV. Ann. Math. Statist. 36, 55-68.

Chernoff, H. (1967). Sequential models for clinical trials. Proc. 5th Berkeley Symp. 4, 805-12.

Chernoff, H. (1972). Sequential Analysis and Optimal Design. SIAM monograph. Philadelphia.

Chernoff, H. and Petkau, A.J. (1976). An optimal stopping problem for sums of dichotomous random variables. Ann. Prob. 4, 875-89.

Chernoff, H. and Ray, S.N. (1965). A Bayes sequential sampling inspection plan. Ann. Math. Statist. 36, 1387-407.

Colton, T. (1963). A model for selecting one of two medical treatments. J. Am. Statist. Assoc. 58, 388-400.

Day, N.E. (1969). A comparison of some sequential designs. Biometrika 56, 301-11.

Maurice, R.J. (1959). A different loss function for the choice between two populations. J.R. Statist. Soc. B 21, 203-13.

Moriguti, S. and Robbins, H. (1962). A Bayes test of $p \leq 1/2$ versus $p > 1/2$. Rep. Statist. Appl. Res., Un. Japan Sci. Engrs. 9, 39-60.

Petkau, A.J. (1978). Sequential medical trials for comparing and experimental with a standard treatment. J. Am. Statist. Assoc. 73, 328-38.

Siegmund, D. O. (1978). private communication.

APPENDIX

In this appendix we provide some details for the derivation of asymptotic (for fixed z and $s \to \infty$) expansion of $\rho_4(y,s)$ given in (4.4.10). From (4.4.9),

$$\rho_4(y,s) = Ks^{-\frac{1}{2}}\phi(z) + 2\phi(z) \int_0^\infty e^{-b^2/2} \cosh(bz) f(s^{\frac{1}{2}}b)\,db , \qquad (1)$$

and from (4.4.8),

$$f_2(x) = |x|^{-1}[2 \log x^2 - 3 \log(\log x^2) + C_1 + (2 \log x^2)^{-1}\{9 \log(\log x^2) + C_2\}] ,$$

where

$$C_1 = \log(2^{-4}\pi^{-1}e) ,$$

$$C_2 = \log(2^{12}\pi^3 e^{-4}) .$$

Requiring $\alpha > 1$, but otherwise arbitrary, set $f(x) = f_2(x)\, I(|x| \geq \alpha)$. With this choice of f, (1) becomes

$$\rho_4(y,s) = Ks^{-\frac{1}{2}}\phi(z) + 2s^{-\frac{1}{2}}\phi(z)[\{2\log s - 3\log(\log s) + C_1\} I_1 + 4I_2 - 3I_3 + (2\log s)^{-1}\{9\log(\log s) + C_2\}I_4 + 9(2\log s)^{-1}I_5] , \qquad (2)$$

where

$$I_1 = \int_{\alpha s^{-\frac{1}{2}}}^\infty b^{-1} e^{-b^2/2} \cosh(bz)\,db ,$$

$$I_2 = \int_{\alpha s^{-\frac{1}{2}}}^\infty b^{-1}(\log b) e^{-b^2/2} \cosh(bz)\,db ,$$

$$I_3 = \int_{\alpha s^{-\frac{1}{2}}}^\infty b^{-1} \log(1 + 2\log b/\log s) e^{-b^2/2} \cosh(bz)\,db ,$$

$$I_4 = \int_{\alpha s^{-\frac{1}{2}}}^{\infty} b^{-1}(1 + 2\log b/\log s)^{-1} e^{-b^2/2} \cosh(bz)\,db ,$$

$$I_5 = \int_{\alpha s^{-\frac{1}{2}}}^{\infty} b^{-1}(1 + 2\log b/\log s)^{-1} \log(1 + 2\log b/\log s)\, e^{-b^2/2} \cosh(bz)\,db .$$

It remains only to evaluate these integral expressions.

Rewrite

$$I_1 = \int_{\alpha s^{-\frac{1}{2}}}^{\infty} b^{-1} e^{-b^2/2}\{\cosh(bz)-1\}db + \int_{\alpha s^{-\frac{1}{2}}}^{1} b^{-1}(e^{-b^2/2}-1)\,db + \int_{\alpha s^{-\frac{1}{2}}}^{1} b^{-1}db + \int_{1}^{\infty} b^{-1} e^{-b^2/2}db,$$

$$= \int_{0}^{\infty} b^{-1} e^{-b^2/2}\{\cosh(bz)-1\}db + \int_{0}^{1} b^{-1}(e^{-b^2/2}-1)\,db + \int_{1}^{\infty} b^{-1} e^{-b^2/2}db + \int_{\alpha s^{-\frac{1}{2}}}^{1} b^{-1}db$$

$$- \int_{0}^{\alpha s^{-\frac{1}{2}}} b^{-1} e^{-b^2/2}\{\cosh(bz)-1\}db - \int_{0}^{\alpha s^{-\frac{1}{2}}} b^{-1}(e^{-b^2/2}-1)\,db .$$

Thus,

$$I_1 = h(z) + K - \log \alpha + \frac{1}{2}\log s + O(s^{-1}) ,$$

where

$$h(z) = \int_{0}^{\infty} b^{-1} e^{-b^2/2}\{\cosh(bz)-1\}db$$

and

$$K = \int_{0}^{1} b^{-1}(e^{-b^2/2}-1)\,db + \int_{1}^{\infty} b^{-1} e^{-b^2/2}db ,$$

$$= \int_{0}^{\infty} b(\log b)\, e^{-b^2/2}db ,$$

$$= (\log 2 - \gamma)/2 ,$$

where γ = Euler's constant. Similar analyses yield

$$I_2 = \int_{\alpha s^{-\frac{1}{2}}}^{1} b^{-1}(\log b)\,db + O(1) ,$$

$$= -\frac{1}{8}(\log s)^2 + \frac{1}{2}\log\alpha \log s + O(1) ,$$

$$I_3 = \int_{\alpha s^{-\frac{1}{2}}}^{1} b^{-1}\log(1 + 2\log b/\log s)\,db + O(1) ,$$

$$= -\frac{1}{2}\log s + \log\alpha \log(\log s) + O(1) ,$$

$$I_4 = \int_{\alpha s^{-\frac{1}{2}}}^{1} b^{-1}(1+2\log b/\log s)^{-1}db + O(1) ,$$

$$= \frac{1}{2}\log s\{\log(\log s) - \log(\log \alpha^2)\} + O(1) ,$$

$$I_5 = \int_{\alpha s^{-\frac{1}{2}}}^{1} b^{-1}(1 + 2\log b/\log s)^{-1}\log(1 + 2\log b/\log s)\,db + O(1) ,$$

$$= -\frac{1}{4}\log s\{\log(\log s) - \log(\log \alpha^2)\}^2 + O(1) .$$

Substituting these expressions into (2) and simplifying leads directly to

$$\rho_4(y,s) = s^{-\frac{1}{2}}\phi(z)[(\log s)^2 - 3\log s \log(\log s) + \{2h(z)+c\}\{2\log s - 3\log(\log s)\} + \frac{9}{4}\{\log(\log s)\}^2 + 4\log(\log s) + O(1)]$$

where

$$c = 2K + (C_1+3)/2 ,$$

$$= 2 - \log 2 - \frac{1}{2}\log\pi - \gamma = 0.157272 ,$$

and

$$h(z) = \int_0^{\infty} b^{-1}e^{-b^2/2}\{\cosh(bz)-1\}db .$$

Note that the constant α disappears in terms larger than $O(1)$.

OFFICE OF NAVAL RESEARCH
STATISTICS AND PROBABILITY PROGRAM

BASIC DISTRIBUTION LIST
FOR
UNCLASSIFIED TECHNICAL REPORTS

JANUARY 1979

	Copies
Statistics and Probability Program (Code 436) Office of Naval Research Arlington, VA 22217	3
Defense Documentation Center Cameron Station Alexandria, VA 22314	12
Office of Naval Research New York Area Office 715 Broadway - 5th Floor New York, New York 10003	1
Commanding Officer Office of Naval Research Branch Office Attn: D. A.L. Powell Building 114, Section D 666 Summer Street Boston, MA 02210	1
Commanding Officer Office of Naval Research Branch Office Attn: Director for Science 536 South Clark Street Chicago, Illinois 60605	1
Commanding Officer Office of Naval Research Branch Office Attn: Dr. Richard Lau 1030 East Green Street Pasadena, CA 91101	1
Office of Naval Research San Francisco Area Office One Hallidie Plaza - Suite 601 San Francisco, CA 94102	1
Office of Naval Research Scientific Liaison Group Attn: Scientific Director American Embassy - Tokyo APO San Francisco 96503	1
Applied Mathematics Laboratory David Taylor Naval Ship Research and Development Center Attn: Mr. G. H. Gleissner Bethesda, Maryland 20084	1
Commandant of the Marine Corps (Code AX) Attn: Dr. A.L. Slafkosky Scientific Advisor Washington, DC 20380	1
Director National Security Agency Attn: Mr. Stahly and Dr. Maar (R51) Fort Meade, MD 20755	2
Navy Library National Space Technology Laboratory Attn: Navy Librarian Bay St. Louis, MS 39522	1

	Copies
U.S. Army Research Office P.O. Box 12211 Attn: Dr. J. Chandra Research Triangle Park, NC 27706	1
Naval Sea Systems Command (NSEA 03F) Attn: Miss B. S. Orleans Crystal Plaza #6 Arlington, VA 20360	1
Office of the Director Bureau of The Census Attn: Mr. H. Nisselson Federal Building 3 Washington, DC 20233	1
OASD (I&L), Pentagon Attn: Mr. Charles S. Smith Washington, DC 20301	1
ARI Field Unit-USAREUR Attn: Library c/o ODCSPER HQ USAREUR & 7th Army APO New York 09403	1
Naval Underwater Systems Center Attn: Dr. Derrill J. Bordelon Code 21 Newport, Rhode Island 02840	1
Library, Code 1424 Naval Postgraduate School Monterey, California 93940	1
Technical Information Division Naval Research Laboratory Washington, DC 20375	1

OFFICE OF NAVAL RESEARCH
STATISTICS AND PROBABILITY PROGRAM

MODELING AND ESTIMATION DISTRIBUTION LIST
FOR
UNCLASSIFIED TECHNICAL REPORTS

JANUARY 1979

	Copies
Technical Library Naval Ordnance Station Indian Head, Maryland 20640	1
Bureau of Naval Personnel Department of the Navy Technical Library Washington, DC 20370	1
Library Naval Ocean Systems Center San Diego, CA 92152	1
Professor Robert Serfling Department of Statistics Florida State University Tallahassee, FL 32306	1
Professor Ralph A. Bradley Department of Statistics Florida State University Tallahassee, FL 32306	1
Professor G. S. Watson Department of Statistics Princeton University Princeton, NJ 08540	1
Professor P. J. Bickel Department of Statistics University of California Berkeley, CA 94720	1
Professor H. Robbins Department of Mathematics Columbia University New York, New York 10027	1
Professor W. M. Hirsch Courant Institute of Mathematical Sciences New York University New York, New York 10453	1
Professor F. J. Anscombe Department of Statistics Yale University New Haven, CT 06520	1
Professor S. S. Gupta Department of Statistics Purdue University Lafayette, Indiana 47907	1
Professor R. E. Bechhofer Department of Operations Research Cornell University Ithaca, New York 14850	1
Professor D. B. Owen Department of Statistics Southern Methodist University Dallas, Texas 75275	1
Professor A.F. Veinott Department of Operations Research Stanford University Stanford, CA 94305	1

	Copies
Professor D. L. Iglehart Department of Operations Research Stanford University Stanford, CA 94305	1
Professor Herbert Solomon Department of Statistics Stanford University Stanford, CA 94305	1
Professor P.A.W. Lewis Department of Operations Research Naval Postgraduate School Monterey, CA 93940	1
Professor R. L. Disney Dept. of Industrial Engineering and Operations Research Virginia Polytechnic Institute and State University Blacksburg, VA 24061	1
Professor H. Chernoff Department of Mathematics Massachusetts Institute of Technology Cambridge, MA 02139	1
Dr. D. E. Smith Desmatics, Inc. P.O. Box 618 State College, PA 16801	1
Professor J.A. Muckstadt Department of Operations Research Cornell University Ithaca, New YOrk 19850	1
Professor R. W. Madsen Department of Statistics University of Missouri Columbia, Missouri 65201	1
Professor F. A. Tillman Department of Industrial Engineering Kansas State University Manhattan, Kansas 66506	1

	Copies
Professor D. P. Gaver Department of Operations Research Naval Postgraduate School Monterey, CA 93940	1
Dr. M. J. Fischer Defense Communications Agency Defense Communications Engineering Center 1860 Wiehle Avenue Reston, Virginia 22090	1
Defense Logistics Studies Information Exchange Army Logistics Management Center Attn: Mr. J. Dowling Fort Lee, Virginia 23801	1
Professor D. O. Siegmund Department of Statistics Stanford University Stanford, CA 94305	1
Professor M. L. Puri Department of Mathematics Indiana University Foundation P.O. Box F Bloomington, Indiana 47401	1
Professor S. M. Ross College of Engineering University of California Berkeley, CA 94720	1
Professor J. E. Boyer, Jr. Department of Statistics Southern Methodist University Dallas, Texas 75275	1
Professor Grace Wahba Department of Statistics University of Wisconsin Madison, Wisconsin 53706	1
Mr. David S. Siegel Code 210T Office of Naval Research Arlington, VA 22217	1

	Copies
Reliability Analysis Center (RAC) RADC/RBRAC Attn: I. L. Krulac Data Coordinator/ Government Programs Griffiss AFB, New York 13441	1
Mr. Jim Gates Code 9211 Fleet Material Support Office U.S. Navy Supply Center Mechanicsburg, PA 17055	1
Mr. Ted Tupper Code 92 Fleet Material Support Office U.S. Navy Supply Center Mechanicsburg, PA 17055	1
Mr. Barnard H. Bissinger Mathematical Sciences Capitol Campus Pennsylvania State University Middletown, PA 17057	1
Professor Walter L. Smith Department of Statistics University of North Carolina Chapel Hill, NC 27514	1
Professor S. E. Fienberg Department of Applied Statistics University of Minnesota St. Paul, Minnesota 55108	1
Professor Gerald L. Sievers Department of Mathematics Western Michigan University Kalamazoo, Michigan 49008	1
Professor Richard L. Dykstra Department of Statistics University of Missouri Columbia, Missouri 65201	1

	Copies
Professor Franklin A. Graybill Department of Statistics Colorado State University Fort Collins, Colorado 80523	1
Professor J. S. Rustagi Department of Statistics Ohio State University Research Foundation Columbus, Ohio 43212	1
Mr. F. R. Del Priori Code 224 Operational Test and Evaluation Force (OPTEVFOR) Norfolk, Virginia 23511	1

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE	READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER (14) TR-14	2. GOVT ACCESSION NO.
3. RECIPIENT'S CATALOG NUMBER	
4. TITLE (and Subtitle) (6) SEQUENTIAL MEDICAL TRIALS INVOLVING PAIRED DATA	5. TYPE OF REPORT & PERIOD COVERED (9) Technical Report
6. PERFORMING ORG. REPORT NUMBER	
7. AUTHOR(s) (10) Herman Chernoff and A. John Petkau	8. CONTRACT OR GRANT NUMBER(s) (15) N00014-75-C-0555
9. PERFORMING ORGANIZATION NAME AND ADDRESS Department of Mathematics Massachusetts Institute of Technology Cambridge, Massachusetts 02139	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS (NR-042-331)
11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research Statistics & Probability Program Code 436 Arlington, Virginia 22217	12. REPORT DATE (11) May 1979 13. NUMBER OF PAGES 51
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) (12) 61 p.	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)

18. SUPPLEMENTARY NOTES

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

clinical trials: optimal stopping: Wiener process: free boundary problem

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

See reverse side

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

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

220021 /B

A continuous time version of Anscombe's formulation of the problem of comparing two treatments in the context of medical trials is considered and the Bayes sequential procedure is explicitly determined. Various suboptimal procedures are proposed, evaluated and compared to the optimal procedure; the approximation to the optimal procedure proposed by Anscombe turns out to be surprisingly efficient. Comparison with the discrete time version demonstrates that the continuous time version provides accurate approximations for clinical trials involving horizon sizes as small as 100. The optimal procedure determined here provides a design which is relevant for clinical trials involving either normal or Bernoulli responses.