SEQUENTIAL MEDICAL TRIALS FOR COMPARING
AN EXPERIMENTAL WITH A STANDARD TREATMENT

BY

ALBERT JOHN PETKAU:

TECHNICAL REPORT NO. 221
JULY 30, 1975

THIS RESEARCH WAS SPONSORED BY THE ARMY RESEARCH OFFICE
OFFICE OF NAVAL RESEARCH, AND AIR FORCE OFFICE OF
SCIENTIFIC RESEARCH BY CONTRACT NO.
NO0014-67-A-0112-0085 (NR-042-267)

Reproduction in Whole or in Part is Permitted
for the purpose of the United States Government
Approved for public release; distribution unlimited.

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA
SEQUENTIAL MEDICAL TRIALS FOR COMPARING
AN EXPERIMENTAL WITH A STANDARD TREATMENT

by

Albert John Petkau

Technical Report No. 221
July 30, 1975

Prepared under Contract N00014-67-A-0112-0085
(NR-042-267)
Office of Naval Research
Herbert Solomon, Project Director

Reproduction in Whole or in Part is Permitted for
any Purpose of the United States Government

Approved for public release; distribution unlimited

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA
<table>
<thead>
<tr>
<th>Chapter</th>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>PROBLEM ATTACKED</td>
<td>5</td>
</tr>
<tr>
<td>2.1</td>
<td>The Model</td>
<td>5</td>
</tr>
<tr>
<td>2.2</td>
<td>Objectives</td>
<td>6</td>
</tr>
<tr>
<td>2.3</td>
<td>The Optimal Solution</td>
<td>7</td>
</tr>
<tr>
<td>2.4</td>
<td>A Related Free Boundary Problem</td>
<td>11</td>
</tr>
<tr>
<td>3</td>
<td>A RELATED WIENER PROCESS PROBLEM</td>
<td>19</td>
</tr>
<tr>
<td>3.1</td>
<td>Formulation of the Problem</td>
<td>19</td>
</tr>
<tr>
<td>3.2</td>
<td>An Associated Free Boundary Problem</td>
<td>24</td>
</tr>
<tr>
<td>3.3</td>
<td>Asymptotic Behavior Near $s = 1$</td>
<td>27</td>
</tr>
<tr>
<td>3.4</td>
<td>Asymptotic Behavior for Large $s$</td>
<td>32</td>
</tr>
<tr>
<td>3.5</td>
<td>Suboptimal Procedures</td>
<td>41</td>
</tr>
<tr>
<td>3.5.A</td>
<td>Procedure W</td>
<td>42</td>
</tr>
<tr>
<td>3.5.B</td>
<td>Procedure F</td>
<td>43</td>
</tr>
<tr>
<td>4</td>
<td>COMPUTATIONAL RESULTS</td>
<td>47</td>
</tr>
<tr>
<td>4.1</td>
<td>Solution of the Wiener Process Problem</td>
<td>47</td>
</tr>
<tr>
<td>4.2</td>
<td>Solution of the Binomomial Problem</td>
<td>64</td>
</tr>
<tr>
<td>4.3</td>
<td>Evaluation of Suboptimal Procedures</td>
<td>78</td>
</tr>
<tr>
<td>4.3.A</td>
<td>Procedure W</td>
<td>78</td>
</tr>
<tr>
<td>4.3.B</td>
<td>Procedure F</td>
<td>81</td>
</tr>
<tr>
<td>Chapter</td>
<td>Page</td>
<td></td>
</tr>
<tr>
<td>---------</td>
<td>------</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4.4</td>
<td>84</td>
<td></td>
</tr>
<tr>
<td>4.4.A</td>
<td>84</td>
<td></td>
</tr>
<tr>
<td>4.4.B</td>
<td>92</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>101</td>
<td></td>
</tr>
</tbody>
</table>

COMPUTATIONAL RESULTS (cont.)

4.4 Comparisons of Optimal and Suboptimal Procedures

4.4.A Boundaries

4.4.B Risks

DISCUSSION AND CONCLUSIONS

REFERENCES

110
### TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Boundaries and Risks of Procedure W for Different Starting Points and Different Values of the Parameter $\gamma$</td>
<td>80</td>
</tr>
<tr>
<td>2</td>
<td>Efficiencies of Procedures W and F for Different Starting Points and Different Values of the Parameter $\gamma$</td>
<td>96</td>
</tr>
<tr>
<td>3</td>
<td>Risks Incurred in Two Different Clinical Trials Problems Corresponding to the Wiener Process Problem With $\gamma = 0.1$</td>
<td>98</td>
</tr>
<tr>
<td>4</td>
<td>Risks Incurred in Two Different Clinical Trials Problems Corresponding to the Wiener Process Problem With $\gamma = 2.0$</td>
<td>99</td>
</tr>
<tr>
<td>Figure</td>
<td>Illustration</td>
<td>Page</td>
</tr>
<tr>
<td>--------</td>
<td>-----------------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>1</td>
<td>Solution of Wiener Process Problem -- ((z,t)) Plane</td>
<td>51</td>
</tr>
<tr>
<td>2</td>
<td>Solution of Wiener Process Problem -- ((\alpha,t)) Plane</td>
<td>52</td>
</tr>
<tr>
<td>3</td>
<td>Solution of Wiener Process Problem -- ((\alpha,t)) Plane -- Small Values of (t)</td>
<td>53</td>
</tr>
<tr>
<td>4</td>
<td>Solution of Wiener Process Problem -- ((\Phi(\alpha),t)) Plane</td>
<td>55</td>
</tr>
<tr>
<td>5</td>
<td>Solution of Wiener Process Problem -- ((\Phi(\alpha),t)) Plane -- Small Values of (t)</td>
<td>56</td>
</tr>
<tr>
<td>6</td>
<td>Asymptotic Expansions -- (t) Large -- (\gamma = 2.0)</td>
<td>57</td>
</tr>
<tr>
<td>7</td>
<td>Asymptotic Expansions -- (t) Large -- (\gamma = 0.5)</td>
<td>58</td>
</tr>
<tr>
<td>8</td>
<td>Asymptotic Expansions -- (t) Large -- (\gamma = 0.1)</td>
<td>59</td>
</tr>
<tr>
<td>9</td>
<td>Asymptotic Expansions -- (t) Small -- (\gamma = 2.0)</td>
<td>61</td>
</tr>
<tr>
<td>10</td>
<td>Asymptotic Expansions -- (t) Small -- (\gamma = 0.5)</td>
<td>62</td>
</tr>
<tr>
<td>11</td>
<td>Asymptotic Expansions -- (t) Small -- (\gamma = 0.1)</td>
<td>63</td>
</tr>
<tr>
<td>12</td>
<td>Comparison of Wiener Process Solution and Binomial Solution -- (\gamma = 0.1), (p_0 = 0.5)</td>
<td>69</td>
</tr>
<tr>
<td>13</td>
<td>Comparison of Wiener Process Solution and Binomial Solution -- (\gamma = 0.5), (p_0 = 0.5)</td>
<td>70</td>
</tr>
<tr>
<td>14</td>
<td>Comparison of Wiener Process Solution and Binomial Solution -- (\gamma = 2.0), (p_0 = 0.5)</td>
<td>71</td>
</tr>
<tr>
<td>15</td>
<td>Comparison of Wiener Process Solution and Binomial Solution -- (\gamma = 0.1), (p_0 = 0.1)</td>
<td>72</td>
</tr>
<tr>
<td>16</td>
<td>Comparison of Wiener Process Solution and Binomial Solution -- (\gamma = 0.5), (p_0 = 0.1)</td>
<td>73</td>
</tr>
<tr>
<td>17</td>
<td>Comparison of Wiener Process Solution and Binomial Solution -- (\gamma = 2.0), (p_0 = 0.1)</td>
<td>74</td>
</tr>
<tr>
<td>18</td>
<td>Comparison of Wiener Process Solution and Binomial Solution -- (\gamma = 0.1), (p_0 = 0.9)</td>
<td>75</td>
</tr>
</tbody>
</table>
ILLUSTRATIONS (cont.)

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>19</td>
<td>Comparison of Wiener Process Solution and Binomial Solution -- $\gamma = 0.5$, $p_0 = 0.9$</td>
<td>76</td>
</tr>
<tr>
<td>20</td>
<td>Comparison of Wiener Process Solution and Binomial Solution -- $\gamma = 2.0$, $p_0 = 0.9$</td>
<td>77</td>
</tr>
<tr>
<td>21</td>
<td>Boundaries of Procedure $W$ -- Case $s_0 = 1001$, $y_0 = 0$ -- $(a,t)$ Plane</td>
<td>82</td>
</tr>
<tr>
<td>22</td>
<td>Boundaries of Procedure $W$ -- Case $s_0 = 1001$, $y_0 = 0$ -- $(\phi(a),t)$ Plane</td>
<td>83</td>
</tr>
<tr>
<td>23</td>
<td>Boundaries of Procedure $F$ -- $(z,t)$ Plane</td>
<td>85</td>
</tr>
<tr>
<td>24</td>
<td>Boundaries of Procedure $F$ -- $(a,t)$ Plane</td>
<td>86</td>
</tr>
<tr>
<td>25</td>
<td>Boundaries of Procedure $F$ -- $(\phi(a),t)$ Plane</td>
<td>87</td>
</tr>
<tr>
<td>26</td>
<td>Comparison of Optimal Procedure and Procedure $F$ -- $\gamma = 2.0$</td>
<td>89</td>
</tr>
<tr>
<td>27</td>
<td>Comparison of Optimal Procedure and Procedure $F$ -- $\gamma = 0.5$</td>
<td>90</td>
</tr>
<tr>
<td>28</td>
<td>Comparison of Optimal Procedure and Procedure $F$ -- $\gamma = 0.1$</td>
<td>91</td>
</tr>
</tbody>
</table>
Chapter 1

Introduction

In recent years there has been an emphasis in the literature on a decision theoretic approach to the design of sequential medical trials for comparing two treatments. Such an approach has been strongly advocated by Anscombe [2], who does not believe that the traditional Neyman-Pearson theory can be appropriately applied in this context. He considers a model in which a total of $N$ patients are to receive one of the two available treatments. Patients are paired and one of each pair is randomly selected for one treatment while the other is given the second treatment. The number of pairs to be treated is determined sequentially, after which the remainder of the patients are given the treatment estimated to be superior. The loss function considered was based on only one cost, the consequence of treating a patient with the inferior of the two treatments. It was assumed that a quantitative measure of response was obtained for each patient and that these responses were normally distributed with unknown means and known variances. Anscombe considered the problem of finding the optimum Bayes sequential plan for his formulation of the problem and was able to obtain approximate solutions. In research that was carried out concurrently but independently, Colton [12] considered exactly this same formulation. He determined the optimal fixed sample size procedure for both the minimax and Bayes criteria. For both of these criteria, he also determined approximations to the best among a class of sequential procedures which we shall call
Wald-type because of their resemblance to the Wald sequential probability ratio test. Subsequently, Cornfield, Halperin, and Greenhouse [15] have investigated an adaptive procedure for this same formulation. A similar formulation in which the response variables are assumed to be dichotomous has also received some attention. Zelen [40] has investigated the behavior of the "play the winner" allocation rule in this context. Subsequently this allocation rule has received a great deal of attention [21, 22, 26, 27, 29, 34, 35, 37]. Canner [5] also investigates this formulation of the problem and considers the effect of including in the loss function the cost of experimentation. Various authors have contributed work that is related to one or the other of the above formulations [13, 14, 16, 17, 28].

The case in which more than two treatments are to be compared has also received some attention. Generally speaking, however, this problem has not been considered in the particular context of clinical trials, but has been investigated as a problem in selecting the population with the largest mean [25, 30, 31, 32, 33, 36].

A closely related problem is that of comparing two treatments in the context of clinical trials when one of the available treatments is a standard treatment. This problem appears to have received very limited attention. Donner [18] has considered one- and two-stage procedures for a particular formulation of this problem and Chernoff [9] has considered this problem from the sequential point of view. The generalization of this problem to the case of more than two available treatments does not appear to have been investigated from the sequential point of view.
Over the period of years in which this research was appearing in the literature, a number of papers have appeared which attempt to address the underlying statistical and medical questions that arise in the planning and conduct of these clinical trials. These papers are particularly useful in clarifying the issues involved for those of us that are not that well acquainted with this particular area of application of statistics. A few of these papers are [2], [3], and [39].

The particular problem considered in this investigation is the problem of comparing an experimental treatment with a standard treatment in the context of clinical trials. The model considered is similar to that of Canner [5]. A finite number of patients are to be treated with one of the two available treatments. The responses are assumed to be dichotomous. One of the costs incurred is assumed to be proportional to the number of failures. The remaining cost is that of experimentation in the portion of the plan prior to the time at which a decision is made as to which of the two treatments is superior and should be applied to the remaining patients. In Chapter 2, the model is formulated in detail. The Bayes solution is determined in Section 2.3, and in Section 2.4 we indicate that asymptotically this Bayes solution can be represented as the solution of a free boundary problem.

In Chapter 3 we study a related Wiener process problem that arises from the consideration of the normal version of the clinical trials problem. This related problem, whose solution is also represented by the solution of the free boundary problem mentioned above,
is then investigated with the help of techniques developed by Chernoff [4], [6], [7], [8] in connection with sequentially testing for the sign of the mean of a normal distribution. In Section 3.5 two suboptimal procedures are proposed for the Wiener process problem.

Chapter 4 contains the bulk of the computational work carried out in this research. In Section 4.1 a computational method of solving the Wiener process problem is presented and carried out. In Section 4.2 the approximate solution provided by the solution of the Wiener process problem is compared to the solution of the clinical trials problem. In Section 4.3 the suboptimal procedures proposed in Section 3.5 are evaluated and in the next section the performance of these procedures is compared to that of the optimal procedure within the context of the Wiener process problem.

In Chapter 5 some open problems are indicated, several questions concerning the validity of the model are discussed and the robustness of the Bayes procedure against misspecification of the total number of available patients is examined. The insights gained from this investigation are indicated at the conclusion of Chapter 5.
Chapter 2

Problem Attacked (Formulation of the Problem)

2.1 The Model

The model proposed is a variation of the model considered by Canner [5]. Whereas Canner was concerned with the problem of comparing two experimental treatments, we restrict ourselves to consideration of the problem of comparing two treatments, one of which is an experimental treatment while the other is a standard treatment, that is, one whose characteristics are known.

The following assumptions are made:

(1) All patients must be treated with one or the other of the two available treatments. The total number of patients, that is, the horizon size, is denoted by \( N \). \( N \) is considered to be a large specified quantity.

(2) The response variables are dichotomous, that is, the result of the use of either of the treatments can be classified simply as either a success or a failure. Further, the response is assumed to be instantaneous.

(3) The standard treatment is characterized by a known probability of success denoted by \( p_0 \), while the experimental treatment is characterized by an unknown probability of success denoted by \( p \).

(4) The trials are carried out in the following manner: Sampling is initiated with the experimental treatment and this sampling is continued until a decision is made in favor of one of the two
available treatments. The remaining patients are then treated with the favored treatment. The portion of the plan during which sampling continues will be called the experimental period.

(5) Two costs are considered. One is the cost incurred each time the use of either of the treatments results in a failure. This cost is assigned a value of 1. The other cost is the cost of experimentation. This cost is assigned a value of \( c \) per patient treated in the experimental period.

Several objections to the use of this model can immediately be raised. The assumption of instantaneous response may be unrealistic in many situations. Another serious difficulty is the specification of the horizon size. Objections could be raised to the particular cost structure assumed. Finally, and perhaps most important in the present context, is the lack of control groups. Previous authors have discussed these objections. We shall proceed, keeping these difficulties in mind, but optimistic that this formulation of the problem is relevant enough so that its solution will provide some insight for the more realistic situation.

2.2 Objectives

This research was undertaken with several specific objectives in mind. The first objective was to determine the optimal solution and to characterize this solution as a function of the parameters of the problem in as simple a manner as possible. A second objective was to propose various suboptimal procedures in the hope of obtaining a
procedure that is simple to implement and whose performance is comparable to that of the optimal procedure. The final objective is to gain insight for the more realistic situation as mentioned in the previous section.

2.3 The Optimal Solution

Since the standard treatment is characterized by a known probability of success, the only reasonable procedures are those which begin sampling with the experimental treatment and continue to do so until sufficient evidence accumulates to enable the experimenter to decide in favor of one or the other of the two available treatments. It should be kept in mind that the word reasonable is applied in the context of the model as we have formulated it. A small cost of sampling or a large horizon size encourages the experimenter to continue sampling.

Let $s$ be the number of successes observed in the first $n$ patients sampled. Since $s$ is a sufficient statistic for $p$ based on the "past history", the decision to stop or continue sampling may be made to depend on $s$. Thus we may restrict our attention to procedures defined by two mutually exclusive and exhaustive sets in the $(n,s)$ plane called the continuation and stopping sets. If at any stage the current position $(n,s)$ belongs to the continuation set we sample another patient. If it belongs to the stopping set we terminate sampling and treat the remaining patients with the treatment which seems better. Obviously all points with $n > N$ or $s > N$ are stopping points. Now define the boundary set as the set of stopping
points accessible from some continuation point. Further define the
upper boundary set (lower boundary set) as that subset of the boundary
set at which the appropriate decision is a decision in favor of the
experimental (standard) treatment.

Due to the nature of the costs associated with a procedure, one
would expect that for an admissible plan, the upper boundary set con-
sists of points \( \{(n, s_n^-)\} \) with \( s_n^- \) monotonically nondecreasing in \( n \)
while the lower boundary set consists of points \( \{(n, s_n^+)\} \) with \( s_n^+ \)
monotonically nondecreasing in \( n \) and the points \((n, s)\) are continua-
tion points or stopping points according as \( s_n^- < s < s_n^+ \) or alter-
nately either \( s < s_n^- \) or \( s > s_n^+ \). Such procedures are thus defined
by a set of boundary points \( \{(n, s_n^-)\} \) and \( \{(n, s_n^+)\} \). The problem of
determining the optimal solution thus becomes the problem of deter-
mining the set of boundary points.

The technique used to define an optimal solution is the introduc-
tion of a beta prior distribution for the unknown probability of
success of the experimental treatment and the subsequent determination
of the Bayes sequential procedure. It is well known that the backward
induction method of dynamic programming enables one to determine the
Bayes sequential procedure for any truncated sequential decision prob-
lem (see [20]).

Suppose we assume a beta prior distribution with parameters
\((a_o, b_o)\) for \( p \). After sampling \( n \) observations, the posterior dis-
tribution of \( p \) is a beta distribution with parameters \((a_o + s, b_o + f)\)
where \( s \) and \( f \) are respectively the number of successes and number
of failures obtained with the \( n \) observations. Thus the state of the
experiment at any stage can be completely specified by a three tuple, 
(a,b,k) say, where (a,b) denotes the parameters of the posterior 
distribution of p and k denotes the number of patients that remain 
to be treated. Note that prior to sampling the state of the experiment 
is specified by \((a_o, b_o, N)\) and further note that at any stage 
\(a + b + k = a_o + b_o + N\).

After sampling \(n\) observations, a choice of one of the following 
three decisions must be made:

(1) Terminate sampling and treat the remaining patients with the 
    standard treatment.

(2) Terminate sampling and treat the remaining patients with the 
    experimental treatment.

(3) Sample the next patient and proceed optimally thereafter.

The choice among these decisions is made on the basis of consideration 
of the expected future costs which result.

Suppose the experiment is in state \((a,b,k)\). The present esti-
mate of \(p\) is given by the expectation of the posterior distribution, 
namely \(a/(a+b)\). Choosing the first decision results in an expected 
future cost given by the expected number of failures that will be 
incurred if all the remaining patients are treated with the standard 
treatment, namely \(k \cdot (1-p_o)\). Similarly choosing the second decision 
results in an expected future cost given by \(k \cdot (1 - a/(a+b))\). If, on 
the other hand, the third decision is chosen, a cost of \(c\) is incurred 
due to the sampling of the next patient. Moreover the state of the 
experiment becomes \((a+1,b,k-1)\) if the observation sampled results in
a success and \((a, b+1, k-l)\) if the observation sampled results in a failure. In the later case a cost of 1 is incurred due to the failure.

From these considerations we conclude that the Bayes risk with respect to the beta prior with parameters \((a_o, b_o)\) is given by \(r_N(a_o, b_o)\) where \(r_N(a, b)\) is computed recursively from the equations:

\[
r_0(a, b) = 0
\]

and

\[
r_i(a, b) = \min(r_i^*(a, b), c + [a/(a+b)]r_{i-1}(a+1, b) + [b/(a+b)]r_{i-1}(a, b+1)+1)
\]

for \(i \geq 1\), where \(r_i^*(a, b) = i \cdot \min\{1-p_o, 1-a/(a+b)\}\). The Bayes stopping rule for the above prior distribution calls for stopping after \(n\) observations consisting of \(s\) successes and \(f\) failures if and only if

\[
r_{N-n}(a_o+s, b_o+f) = r_{N-n}^*(a_o+s, b_o+f).
\]

Thus the Bayes procedure is characterized by a set of stopping points in the \((a, b)\) plane where \(a\) and \(b\) represent \(a_o+s\) and \(b_o+f\) respectively.

The dependence of the Bayes risk and hence the Bayes stopping rule upon \(a_o, b_o, s, f\) and \(N\) can be expressed in terms \(a = a_o+s\), \(b = b_o+f\) and \(M = a_o+b_o+N\) (since \(N-n = M-a-b\)). Thus the single
set of stopping points in the \((a,b)\) plane calculated for fixed \(M\) suffices to give the stopping rules for various combinations of beta prior distributions with parameters \((a'_o, b'_o)\) and horizon sizes \(N'\) such that \(a'_o + b'_o + N' = M\). In other words, the Bayes risk with respect to a beta prior distribution with parameters \((a,b)\) for the problem with horizon size \(M - a - b\) is given by \(r(a,b)\) where \(r(a,b)\) satisfies the recursion relation

\[
  r(a,b) = 0 \quad \text{for} \quad a + b = M,
\]

\[ (2.3.1) \]

\[
  r(a,b) = \min(r^*(a,b), c + [a/(a+b)]r(a+1,b) + [b/(a+b)][r(a,b+1)+1]) \quad \text{for} \quad a + b \leq M - 1,
\]

and where \(r^*(a,b) = [M - a - b] \cdot \min(1-p_o, 1-a/(a+b))\) corresponds to stopping. Furthermore the point \((a,b)\) is in the stopping set if and only if \(r(a,b) = r^*(a,b)\).

Thus for a given set of parameters, \(a_o, b_o, p_o, c\) and \(N\), the Bayes procedure for the problem can be determined simply by performing the calculations involved in carrying out this backward induction algorithm with \(M = a_o + b_o + N\). This does not, however, explicitly characterize the solution in terms of the parameters involved.

2.4 A Related Free Boundary Problem

In this section we indicate heuristically that for large values of \(M\), the problem of obtaining the optimal stopping set and the Bayes risk reduces to a free boundary problem (FBP) involving a diffusion equation.
Let
\[ r'(a,b) = r(a,b) - [M-a-b][1-a/(a+b)] , \quad 0 < a+b \leq M . \]

It follows from (2.3.1) that \( r'(a,b) \) satisfies the recursion relation \( r'(a,b) = 0 \) for \( a+b = M \), and

\[
(2.4.1) \quad r'(a,b) = \min \left\{ \left[ \frac{1}{2} r'(a+1,b) + r'(a,b+1) \right], 0, c 
+ \frac{1}{2} \left[ r'(a+1,b) + r'(a,b+1) \right] 
+ \left[ \frac{a}{(a+b)-1}/2 \right] [r'(a+1,b) - r'(a,b+1)] \right\}
\]

for \( a+b \leq M-1 \). Now normalize with the following transformation:

\[
z = \frac{a-(a+b)p_o}{[M_p(1-p_o)]^{1/2}}
\]

(2.4.2)

\[
t = (a+b)/M
\]

\[
E(z,t) = r'(a,b)/[M_p(1-p_o)]^{1/2}
\]

and let

\[
y_1 = (1-p_o)/[p_o(1-p_o)]^{1/2}
\]

\[
y_2 = -p_o/[p_o(1-p_o)]^{1/2} \cdot
\]

Note that the variable \( t \) represents the fraction of the total (including the prior) available information that has already been collected. The equations (2.4.1) then become
\[ B(z,t) = 0, \quad t = 1 \]

\[ B(z,t) = \min \left\{ \frac{z(1-t)}{t}, 0, c/[Mp_0(1-p_0)]^{1/2} \right\} + \frac{1}{2} \{ B(z+\gamma_1M^{-1/2}, t+M^{-1}) + B(z+\gamma_2M^{-1/2}, t+M^{-1}) \}
\]

\[ + (z[p_0(1-p_0)]^{1/2}/tM^{1/2}+p_0-1/2)(B(z+\gamma_1M^{-1/2}, t+M^{-1}) - B(z+\gamma_2M^{-1/2}, t+M^{-1})) \}, \quad t \leq 1-M^{-1}. \]

Now assuming \( B(z,t) \) is smooth enough to be differentiated with respect to both of its arguments the first few times, the above equations become:

\[ B(z,t) = 0, \quad t = 1 \]

\[ B(z,t) = \min \left\{ z(1-t)/t, 0, c/[Mp_0(1-p_0)]^{1/2} + B(z,t) \right\} + \left\{ \frac{1}{2}B_{zz}(z,t) + (z/t)B_z(z,t) + B_t(z,t)M^{-1} + O(M^{-3/2}) \right\} , \quad t \leq 1-M^{-1}. \]

Setting

(2.4.3) \[ \gamma = cM^{1/2}/[p_0(1-p_0)]^{1/2} \]

these equations become:
\[ B(z, t) = 0, \quad t = 1 \]

\[ (2.4.4) \]
\[
B(z, t) = \min \left\{ \frac{z(1-t)}{t}, 0, B(z, t) + (1/2)B_{zz} + (z/t)B_z + B_t + \gamma M^{-1} + O(M^{-3/2}) \right\}, \quad t \leq 1 - M^{-1}.
\]

Thus in the limit as \( M \to \infty \), \( B(z, t) \) satisfies the following partial differential equation:

\[ (2.4.5) \quad (1/2)B_{zz} + (z/t)B_z + B_t + \gamma = 0 \]

in the continuation region in the \((z, t)\) plane,

\[ z_\ell(t) \leq z \leq z_u(t), \quad 0 \leq t \leq 1 \]

where \( z_\ell(t) \) and \( z_u(t) \) are the lower and upper boundaries respectively to be determined along with \( B(z, t) \).

One set of boundary conditions is given by the requirement of continuity as:

\[ (2.4.6) \quad B(z, t)\bigg|_{z=z_u(t)} = 0, \]

\[ (2.4.7) \quad B(z, t)\bigg|_{z=z_\ell(t)} = z(1-t)/t. \]

Another set of boundary conditions is obtained by the following argument: Suppose \((a, b)\) is a point (in the \((a, b)\) plane) that
corresponds to a point \((z,t)\) (in the \((z,t)\) plane) on the upper boundary \(z_u(t)\), while \((a,b+1)\) corresponds to a point in the continuation region and \((a+1,b)\) corresponds to a point in the stopping region. Then \(r'(a,b) = 0 = r'(a+1,b)\) and (2.4.1) suggests that

\[
0 = c + [1 - a/(a+b)] \cdot r'(a,b+1) .
\]

Using the transformation (2.4.2), expanding \(B\) about \((z,t)\) and remembering that \(B(z,t) = 0\) for points on the upper boundary, the above equation becomes

\[
0 = (1-p_o) \cdot \gamma^2 \cdot B_z + O(M^{-1/2}) ,
\]

which suggests that in the limit as \(M \to \infty\), we expect

\[(2.4.8) \quad B_z(z,t) \bigg|_{z = z_u(t)} = 0 .\]

A similar argument, slightly more complicated algebraically, applies at the lower boundary: Suppose \((a,b)\) corresponds to a point \((z,t)\) on the lower boundary \(z_b(t)\), while \((a,b+1)\) corresponds to a point in the stopping region and \((a+1,b)\) corresponds to a point in the continuation region. Then (2.4.1) suggests that

\[
[M-a-b][a/(a+b)-p_o] = c + [1 - a/(a+b)][M-a-b-1][a/(a+b+1)-p_o] + [a/(a+b)] r'(a+1,b) ,
\]

which leads to
\[ 0 = p_0 \cdot \gamma_1 \cdot B_z - (1-t)[p_0 (1-p_0)]^{1/2}/t + O(M^{-1/2}). \]

This suggests that in the limit as \( M \to \infty \), we expect

\[ (2.4.9) \quad B_z(z,t) \big|_{z=z_k(t)} = (1-t)/t. \]

Thus the problem of obtaining the optimum boundary and the Bayes risk corresponds, in the limit as \( M \to \infty \), to the following FEP:

Solve for \( z_k(t) \), \( z_u(t) \) and \( B(z,t) \) where \( B(z,t) \) satisfies the partial differential equation (2.4.5) in the continuation region subject to the boundary conditions (2.4.6), (2.4.7), (2.4.8) and (2.4.9).

For reasons which will become clear in the following chapter, consider the transformation:

\[ s = 1/t \]

\[ (2.4.10) \quad y = z/t \]

\[ u(y,s) = B(z,t) - z/t + \gamma \cdot t. \]

In the \((y,s)\) plane the FEP becomes:

Solve for \( y_k(s) \), \( y_u(s) \) and \( u(y,s) \) where \( u(y,s) \) satisfies the partial differential equation

\[ (2.4.11) \quad (1/2)u_{yy}(y,s) = u_s(y,s) \]
in the continuation region in the \((y,s)\) plane,

\[ y_L(s) \leq y \leq y_u(s), \quad \infty > s \geq 1 \]

subject to the boundary conditions

\[ u(y,s) \big|_{y=y_u(s)} = \gamma/s - y , \quad (2.4.12) \]

\[ u(y,s) \big|_{y=y_L(s)} = \gamma/s - y/s , \quad (2.4.13) \]

\[ u_y(y,s) \big|_{y=y_u(s)} = -1 , \quad (2.4.14) \]

\[ u_y(y,s) \big|_{y=y_L(s)} = -1/s , \quad (2.4.15) \]

Thus the reduction of the original problem to the FBP indicates that the solution can be characterized, after normalization, in terms of the single parameter \(\gamma\) (see (2.4.3)). This represents a reduction from the three parameters \(c, p,\) and \(M\) of the original problem. The parameters of the prior distribution determine an initial point in the appropriate plane. In particular, the initial point \((a_o, b_o)\) in the \((a,b)\) plane corresponding to the parameters of the beta prior distribution corresponds to the initial point \((z_o, t_o)\) in the \((z,t)\) plane where.
\[ z_o = [a_o - (a_o + b_o)p_o]/[M_p(1-p_o)]^{1/2} \]

(2.4.16)

\[ t_o = (a_o + b_o)/M \]

and corresponds to the initial point \((y_o, s_o)\) in the \((y, s)\) plane where

\[ y_o = M^{1/2}[a_o/(a_o + b_o) - p_o]/[p_o(1-p_o)]^{1/2} \]

(2.4.17)

\[ s_o = M/(a_o + b_o) \]
Chapter 3

A Related Wiener Process Problem

In this chapter we consider a sequential decision problem for a Wiener process. The problem is formulated in Section 3.1 and in this same section we indicate that this problem is closely related to the problem considered in the previous chapter. Further, the problem is reduced to a stopping problem for a Wiener process. In Section 3.2 we reduce the solution of this stopping problem to a free boundary problem. This FBP is exactly the same as that of Section 2.4. We then consider the asymptotic behavior of the solution of the FBP in two different regions in Sections 3.3 and 3.4. Finally in Section 3.5 we propose two suboptimal procedures for the sequential decision problem formulated in Section 3.1 and briefly consider their behavior.

3.1 Formulation of the Problem

In the preceding chapter we have formulated our problem as a sequential decision problem involving binomial random variables. Consider now the following closely related sequential decision problem obtained by replacing the dichotomous 0, 1 observations counting successes in the clinical trials by normal random variables: Two normal processes are available for observation. One process produces independent and identically distributed random variables \( X_i \) whose common distribution is normal with mean \( \mu \) (unknown) and variance \( \sigma^2 \) (known). This process will be termed the unknown process. The other process produces independent and identically distributed random
variables whose common distribution is normal with mean \( \mu_s \), which we will take to be 0, and variance \( \sigma_s^2 \) (known). This process will be termed the standard process. The experimenter is required to observe a total of \( N^* \) random variables \( (N^* \text{ large}) \). Sampling is initiated with the unknown process and is continued with this process during an experimental period until a decision is made in favor of one of the two available processes. The remainder of the \( N^* \) random variables are then observed from the favored process. The experimenter's gains are given by the sum of the random variables observed. However, a cost of \( c^* \) is incurred for each random variable observed during the experimental period, that is, prior to the time of decision. The unknown parameter \( \mu \) is assumed to have a normal prior distribution with mean \( \mu_o \) and variance \( \sigma_o^2 \). What is the optimal procedure?

It is well-known and easily verified that upon observing \( n \) random variables from the unknown process, the posterior distribution of \( \mu \) is given by a normal distribution with mean \( Y_n^* \) and variance \( s_n^* \)

\[
y_n^* = \left[ \sigma^{-2} \sum_{i=1}^{n} x_i + \mu_o \sigma_o^{-2} \right] / \left[ n \sigma^{-2} + \sigma_o^{-2} \right]
\]

\[
s_n^{-1} = n \sigma^{-2} + \sigma_o^{-2}.
\]

Furthermore, it can be shown that for \( n > m \) the distribution of \( Y_n^* - Y_m^* \) is normal with mean 0 and variance \( s_m^* - s_n^* \) and is independent of \( Y_m^* \). Thus as sampling with the unknown process continues,
the posterior mean of \( \mu \) behaves like a Gaussian process with independent increments, starting from \( Y^*_0 = \mu^*_0 \).

The experimenter's gains upon observing \( n \) random variables from the unknown process are given by \( \sum_{1}^{n} X_1 - c^* n \). The choice between the two normal processes will, of course, be made on the basis of the current estimate of \( \mu \), namely \( Y^*_n \). If the decision is in favor of the unknown process, the expected future gains are given by \( (N^*-n)Y^*_n \). If, on the other hand, the decision is in favor of the standard process, the expected future gains are given by \( (N^*-n) \cdot 0 = 0 \).

Thus the choice is made on the basis of the sign of \( Y^*_n \) and the expected future gains are given by \( (N^*-n) \max\{Y^*_n, 0\} \). Using the relations (3.1.1) we obtain the result that the posterior expected gain associated with deciding at the \( n^{th} \) observation is \( g^*(Y^*_n, s^*_n) \) where

\[
(3.1.2) \quad g^*(y^*, s^*) = c^2(y^*/s^*)I_{\{y^*<0\}} + (N^* \sigma^2 + s^2) \sigma^2 y^* I_{\{y^*>0\}}
\]

\[- c^* \sigma^2 / s^* + c^* \sigma^2 / s^* - \mu^*_0 \sigma^2 / s^* .
\]

With this result the problem of determining the optimal procedure for this problem has been reduced to a stopping problem.

Note that the decision to stop or continue after the \( n^{th} \) observation depends on \( \sum_{1}^{n} X_1 \). Thus we are interested in the process \( \sum_{1}^{n} X_1 \), which is for fixed \( \mu \) a Gaussian process of independent increments. Since we are concerned with \( N^* \) large, we expect to observe this sum for many values of \( n \). A limiting version of this process is the Wiener process where the sum \( \sum_{1}^{n} X_1 \) is replaced by the
continuous Gaussian process $X(t)$ with independent increments described by

$$E[dX(t)] = \mu dt$$

(3.1.3)

$$\text{Var}[dX(t)] = \sigma^2 dt.$$ 

When $\mu$ is regarded as a random variable, the limiting form of the $(Y_n^*, s_n^*)$ process is the Wiener process (in the $-s^*$ scale), $Y^*(s^*)$, where

$$E[dY^*(s^*)] = 0$$

(3.1.4)

$$\text{Var}[dY^*(s^*)] = -ds^*$$

with $Y^*(s_o^*) = \mu_o$ at $s_o^* = \sigma_o^2$ and

(3.1.5) 

$$s^{*-1} = t_o^{-2} + \sigma_o^{-2}.$$ 

Note that as $t$ increases from 0 to $N^*$, $s^*$ decreases from $\sigma_o^2$ to $(\sigma_o^{-2} + N^* \sigma_o^{-2})^{-1}$. Hence a limiting form of this problem is a special case of the following general stopping problem: Given a Gaussian process $(Y^*(s^*), s^* \in G^*)$ of independent increments in the $-s^*$ scale, with $E[dY^*(s^*)] = 0$, $\text{Var}[dY^*(s^*)] = -ds^*$, starting at $Y^*(s_o^*) = y_o^*$, find a stopping time $S$ to maximize

$$E[g(Y^*(S), S)].$$
Certain simplifying transformations can be made. First, the constant term appearing in (3.1.2) has no effect on the optimal procedure and may be ignored. Second, the transformation

\[ Y(s) = ay^*(s^*) \]

(3.1.6)

\[ s = a^2 s^* \]

converts the process \( y^*(s^*) \) to the process \( Y(s) \) which is also a Wiener process with \( \text{E}[dY(s)] = 0 \), \( \text{Var}[dY(s)] = -ds \). Then taking

(3.1.7)

\[ a = (\sigma_o^{-2} + N \sigma_s^{-2})^{1/2} \]

(3.1.2) becomes

(3.1.8)

\[ g^*(y^*,s^*) = (\sigma_o^{-2} + N \sigma_s^{-2})^{1/2} \sigma^2 \cdot \{g(y,s)\} + \frac{c^*}{\sigma_o} \sigma^2 - \nu_0 \sigma_0^{-2} \sigma^2 \]

where

(3.1.9)

\[ g(y,s) = \frac{y}{s} I_{\{y \leq 0\}} + y I_{\{y > 0\}} - c^* a / s \]

Thus the problem may be reduced by this transformation to that of dealing with the particular stopping gain \( g(y,s) \) defined in (3.1.9). Note that \( s \) is restricted to \( a^2 \sigma_o^2 = s_o \geq s \geq 1 \). If we define
d(y,s) = -g(y,s), then d(y,s) may be considered as the corresponding stopping cost function. If we further set \( \gamma = c^*a \), we are concerned with the particular stopping problem defined by the stopping cost function

\[
d(y,s) = \begin{cases} 
\gamma/s - y, & y > 0 \\
\gamma/s - y/s, & y \leq 0 
\end{cases}
\]  

(3.1.10)

where \( s \) is restricted to \( s \geq 1 \).

3.2 **An Associated Free Boundary Problem**

Given the general stopping problem for a Wiener process \( \{Y(s), s \in \mathbb{G}\} \) in the \(-s\) scale originating from \((y_0, s_0)\), as described in the previous section, Chernoff [10] has shown that a characterization of an optimal procedure (under regularity conditions sufficient to imply the existence of an optimal procedure) is given by the following:

Let \( \rho(y_0, s_0) = \inf_{S} b(y_0, s_0) \) where \( b(y_0, s_0) \) is the risk associated with a particular stopping time \( S \) and the infimum is taken over all such stopping times. Note that \( \rho(y_0, s_0) \leq d(y_0, s_0) \). But \( \rho(y, s) \) can be defined for all \((y, s)\) with \( s \in \mathbb{G} \). Since \( Y(s) \) is a process with independent increments (in the \(-s\) scale), it follows that \( \rho(y, s) \) also represents the best that can be expected once \( Y(s) = y \) has been reached, irrespective of how it was reached. Therefore a characterization of an optimal procedure is described by the
stopping time \( S_0 \) defined as:

\[(3.2.1) \quad S_0: \text{Stop as soon as } \rho(y(s),s) = d(y(s),s).\]

Since the optimal procedure \( S_0 \) is characterized by the continuation set \( \mathcal{C}_0 = \{(y,s) : \rho(y,s) < d(y,s)\} \) and the stopping set \( \mathcal{S}_0 = \{(y,s) : \rho(y,s) = d(y,s)\} \), we shall restrict attention to procedures that can be represented by a continuation set \( \mathcal{C}_0 \) and its complement, the stopping set \( \mathcal{S}_0 \).

Note that this characterization of \( \mathcal{C}_0 \) does not depend upon the initial point \((y_0, s_0)\) and thus it yields the solution for all initial points simultaneously.

Chernoff [10] has shown that associated with any procedure described by a continuation set \( \mathcal{C}_0 \) we have a risk function \( b(y,s) \) which satisfies the heat equation within \( \mathcal{C}_0 \) subject to the boundary condition \( b(y,s) = d(y,s) \). The optimal continuation region \( \mathcal{C}_0 \) yields \( b(y,s) = \rho(y,s) \) which uniformly minimizes \( b(y,s) \) everywhere.

Chernoff [10] has demonstrated that one should expect the solution \((\rho, \mathcal{C}_0)\) of the optimization problem to be a solution of the following FBP:

\[(3.2.2)\]

\[
(1/2)u_{yy}(y,s) = u_y(y,s), \quad (y,s) \in \mathcal{C}_0
\]

\[
u(y,s) = d(y,s), \quad (y,s) \in \mathcal{S}_0
\]

\[
u_y(y,s) = d_y(y,s), \quad (y,s) \in \partial \mathcal{C}_0.
\]
Furthermore, Chernoff [10] gives conditions on the solution of the FBP (3.2.2) that are sufficient to insure that the solution of the FBP is a solution of the optimization problem.

Thus our optimization problem of Section 3.1 has been reduced to the following FBP:

Determine the function \( \rho(y,s) \) and the continuation set \( \mathcal{C}_o \) such that:

\[
(1/2)\rho_{yy}(y,s) = \rho_s(y,s) \quad , \quad (y,s) \in \mathcal{C}_o
\]

\[
(3.2.3) \quad \rho(y,s) = \begin{cases} 
\gamma/s - y & , \quad (y,s) \in \mathcal{C}_o \cap \{y > 0\} \\
\gamma/s - y/s & , \quad (y,s) \in \mathcal{C}_o \cap \{y < 0\}
\end{cases}
\]

\[
\rho_y(y,s) = \begin{cases} 
-1 & , \quad (y,s) \in \mathcal{C}_o \cap \{y > 0\} \\
-1/s & , \quad (y,s) \in \mathcal{C}_o \cap \{y < 0\}
\end{cases}
\]

where \( s \geq 1 \).

Comparing the FBP (3.2.3) with the FBP (2.4.11) - (2.4.15) developed in Section 2.4, we see that the two problems are identical. The FBP (2.4.11) - (2.4.15) was developed as an approximate representation of the optimal procedure for the clinical trials problem. The solution of the FBP (3.2.3), on the other hand, is expected to correspond exactly to the optimal procedure for the Wiener process problem. In the next two sections we shall concern ourselves with the description of the asymptotic behavior of the solution of the FBP, near \( s = 1 \).
and for $s$ large, within the context of the Wiener process problem. This behavior can then be translated to obtain approximate solutions to our clinical trials problem.

3.3 Asymptotic Behavior Near $s = 1$

In this section we shall present a formal expansion for the optimal boundary and the optimal risk in the region $s$ close to 1. We represent the optimal continuation region $\mathcal{C}_0$ as the set

\[ \{(y,s); y_L(s) \leq y \leq y_U(s), s \geq 1\}. \]

From the description of the problem in Section 3.1, we expect that $y_L(s) \leq 0$ and $y_U(s) \geq 0$. We then proceed to obtain formal expansions for both $y_L(s)$, $y_U(s)$ and $\rho(y,s)$ in the region $s$ close to 1. The techniques employed are similar to the techniques developed by Breakwell and Chernoff [4] and Chernoff and Ray [11].

We transform the FBP of (3.2.3) as follows:

\[ r = s - 1 \]  
\[ (3.3.1) \]
\[ v(y,r) = \rho(y,s) - \gamma + y. \]

Then $v(y,r)$ satisfies the heat equation

\[ (1/2)v_{yy}(y,r) = v_r(y,r) \]  
\[ (3.3.2) \]

in the continuation region $y_L(r) \leq y \leq y_U(r), \ r \geq 0$ with the boundary conditions:
\begin{align}
(3.3.3) \quad v(y,r) &= \gamma [1/(1+r) - 1], \quad y = y_u(r) \\
(3.3.4) \quad v(y,r) &= \gamma [1/(1+r) - 1] - y[1/(1+r) - 1], \quad y = y_g(r) \\
(3.3.5) \quad v_y(y,r) &= 0, \quad y = y_u(r) \\
(3.3.6) \quad v_y(y,r) &= r/(1+r), \quad y = y_g(r).
\end{align}

For \( s \) close to 1, i.e., for \( r \) close to 0, the righthand sides of (3.3.3) - (3.3.6) may be expanded in a power series of \( r \). We are then led to seek separable solutions for \( v(y,r) \) of the form

\begin{equation}
(3.3.7) \quad r^{n/2} g_n(y/r^{1/2}), \quad n = 1, 2, 3, \ldots,
\end{equation}

where \( g_n(y/r^{1/2}) \) may itself be expanded in a power series of \( y/r^{1/2} \).

Solutions of the heat equation (3.3.2) of the form (3.3.7) may be obtained by defining:

\begin{equation}
(3.3.8) \quad g_n(\alpha) = \int_{-\infty}^{\infty} (\alpha + \epsilon)^n \phi(\epsilon) \, d\epsilon / n!
\end{equation}

where

\begin{equation}
(3.3.9) \quad \alpha = y/r^{1/2}
\end{equation}

and

\begin{equation}
(3.3.10) \quad \phi(\epsilon) = (2\pi)^{-1/2} \exp(-\epsilon^2/2).
\end{equation}
These solutions are special cases of the separable solutions of the heat equation of the form

\[ r^{n/2} H_n(\alpha) \]

where \( H_n(\alpha) \) satisfies the ordinary differential equation

\[ H_n''(\alpha) + \alpha H_n'(\alpha) = nH_n(\alpha) \]

and hence, for each \( n \), is a linear combination of two hypergeometric functions.

Note that \( g_n'(\alpha) = g_{n-1}(\alpha) \) for \( n > 1 \). We write down the first few of these polynomials for future reference:

\[ g_1(\alpha) = \alpha \]
\[ g_2(\alpha) = (1+\alpha^2)/2 \]
\[ g_3(\alpha) = (3\alpha+\alpha^3)/3! \]
\[ g_4(\alpha) = (3+6\alpha^2+\alpha^4)/4! . \]

After a certain amount of trial and error, we are led to consider expansions of the form:

\[ v(y,r) = \sum_{n=2}^{\infty} a_n r^{n/2} g_n(\alpha) \]

(3.3.12)
(3.3.13) \[ a_u(r) = \sum_{n=1}^{\infty} b_n r^{n-1/2} \]

(3.3.14) \[ a_x(r) = \sum_{n=1}^{\infty} c_n r^{n-1/2} \]

where \( y_u(r) = r^{1/2} \cdot a_u(r) \) and \( y_x(r) = r^{1/2} \cdot a_x(r) \).

The coefficients \( a_n, b_n \) and \( c_n \) are then determined by matching coefficients of equal powers of \( r \) in the equations obtained by substituting the representations (3.3.12) - (3.3.14) in the boundary conditions (3.3.3) - (3.3.6). The values of the first few of these coefficients are given by

\[ b_1 = b = (4\gamma)^{-1} \]

(3.3.15) \[ b_2 = b(1-4b^2/3) \]

\[ b_3 = -4b^3(1-4b^2/9) \]

and

\[ c_1 = -b \]

(3.3.16) \[ c_2 = -b(1-8b^2/3) \]

\[ c_3 = 4b^3(5/3 - 3b^2/9) \]

while

30
\[ a_2 = -1/2b \]
\[ a_3 = 1 \]
\[ a_4 = 2(1-b^2)/b \]
\[ a_5 = -4(1+b^2)/3 \]
\[ a_6 = -12(1-2b^4)/b \]
\[ a_7 = 24(1+14b^4)/9 \]
\[ a_8 = 96(1+11b^4/3 - 91b^6/6) \]

In terms of the original FEP (3.2.3), these results yield the following formal asymptotic expansions in the region \( s \) close to 1:

\[ (3.3.18) \quad \rho(y,s) = \gamma - y + \sum_{n=2}^{\infty} a_n (s-1)^n / \varepsilon_n (y/(s-1)^{1/2}) \]

\[ (3.3.19) \quad y_u(s) = \sum_{n=1}^{\infty} b_n (s-1)^n \]

\[ (3.3.20) \quad y_z(s) = \sum_{n=1}^{\infty} c_n (s-1)^n \]

The behavior of these formal expansions will be examined in Section 4.1.

By use of the appropriate transformations, these formal expansions can be transformed to yield expansions in any of the planes of interest. The transformation (3.1.6) - (3.1.7) would transform the results to the \((y^*, s^*)\) plane of Section 4.1. To relate the expansions to the
solution of the clinical trials problem, the results could be transformed to either the \((z,t)\) plane of Section 2.4 or the \((a,b)\) plane of Section 2.3. This can be accomplished by the use of the transformations (2.4.10) and (2.4.2).

3.4 Asymptotic Behavior for Large \(s\)

In this section we shall present a formal expansion for the optimal boundary and the optimal risk in the region close to the boundary for large values of \(s\). The same representation of \(Q_0\) is used as was used in the previous section. The techniques employed are similar to the techniques developed by Chernoff [7] and Chernoff and Ray [11].

We introduce the variable \(\alpha\) defined by

\[
(3.4.1) \quad \alpha = y/s^{1/2}
\]

and note from the distribution theory developed in Section 3.1 that \(\alpha\) is simply the number of standard deviations that our current estimate of \(\mu\) is away from zero. Thus for large values of \(s\), we expect \(\alpha_u(s) = y_u(s)/s^{1/2}\) and \(\alpha_\lambda(s) = y_\lambda(s)/s^{1/2}\) to approach \(+\infty\) and \(-\infty\) respectively. We also introduce the functions \(\psi^+(x)\) and \(\psi^-(x)\) defined by

\[
(3.4.2) \quad \psi^+(x) = \phi(x) - x[1 - \phi(x)]
\]

\[
(3.4.3) \quad \psi^-(x) = \phi(x) + x\phi(x)
\]
where \( \phi(x) \) has been defined by (3.3.10) and

\[
(3.4.4) \quad \phi(x) = \int_{-\infty}^{x} \phi(u) \, du
\]

is the standard normal cumulative. Notice that \( \psi^+(x) = \psi^-(x) - x \), \( \psi^+(x) = \psi^-(x) \), and in addition both \( s^{1/2} \psi^+(\alpha) \) and \( s^{1/2} \psi^-(\alpha) \) are solutions of the heat equation.

We transform the FBP of (3.2.3) by considering the \((\alpha, s)\) plane and by adding the special solution of the heat equation given by \( s^{1/2} \psi^-(\alpha) \) to the function \( \rho(y, s) \). Then \( v(\alpha, s) \) defined by

\[
(3.4.5) \quad v(\alpha, s) = \rho(y, s) + s^{1/2} \psi^-(\alpha)
\]

satisfies the transformed heat equation

\[
(3.4.6) \quad \frac{\partial v(\alpha, s)}{\partial \alpha} + \frac{\partial v(\alpha, s)}{\partial s} = 2s \psi_s(\alpha, s)
\]

in the continuation region \( \alpha_{\ell}(s) \leq \alpha \leq \alpha_u(s) \), \( s \geq 1 \), with the boundary conditions

\[
(3.4.7) \quad v(\alpha, s) = g_s^{-1} + s^{1/2} \psi^+(\alpha), \quad \alpha = \alpha_u(s)
\]

\[
(3.4.8) \quad v(\alpha, s) = g_s^{-1} - \alpha s^{-1/2} + s^{1/2} \psi^-(\alpha), \quad \alpha = \alpha_{\ell}(s)
\]

\[
(3.4.9) \quad v(\alpha, s) = -s^{1/2}[1 - \phi(\alpha)], \quad \alpha = \alpha_u(s)
\]

\[
(3.4.10) \quad v(\alpha, s) = -s^{1/2} + s^{1/2} \phi(\alpha), \quad \alpha = \alpha_{\ell}(s).
\]
We now consider

\[(3.4.11) \quad \int s^{-1/2} \phi[s^{-1/2}(y-u)]f(u)du ,\]

which is a heat potential with sources distributed along the \(y\) axis according to the density \(f(u)\). This is a solution of the heat equation provided that the integral converges. Transforming to the \((a,s)\) plane, \(\int s^{-1/2} \phi(a-s^{-1/2}u)f(u)du\) is a solution of the transformed heat equation (3.4.6). Thus we are led to consider a representation of the form:

\[(3.4.12) \quad v(a,s) = \int_{-\infty}^{\infty} \phi(a-b)f(s^{1/2}b)db .\]

Further, after a certain amount of trial and error, we are led to consider representations for the boundaries of the following form:

\[(3.4.13) \quad \ln s = a_u^{2/3} + a_1 + a_2/a_u^2 + a_3/a_u^4 + \cdots\]

\[(3.4.14) \quad \ln s = a_l^{2/2} + a_1^{2/2} + b_1 + b_2/a_l^2 + b_3/a_l^4 + \cdots .\]

From (3.4.12) we obtain

\[(3.4.15) \quad v_\alpha(a,s) = \int_{-\infty}^{\infty} -(a-b)\phi(a-b)f(s^{1/2}b)db .\]

Setting \(b = a + \varepsilon\) in both (3.4.12) and (3.4.15) leads to
\[ (3.4.16) \quad v(a,s) = \int_{-\infty}^{\infty} f(s^{1/2}(a+\varepsilon))\phi(\varepsilon)d\varepsilon = \mathbb{E}[f(s^{1/2}(a+\varepsilon))] \]

\[ (3.4.17) \quad v_a(a,s) = \int_{-\infty}^{\infty} e^s f(s^{1/2}(a+\varepsilon))\phi(\varepsilon)d\varepsilon = \mathbb{E}[e^s f(s^{1/2}(a+\varepsilon))] \]

where \( \mathbb{E} \) denotes expectation with respect to the normal random variable \( \varepsilon \) which has mean zero and variance one. Recall that for \( s \) large we expect \( a_u(s) \) and \( a_l(s) \) to approach \( +\infty \) and \( -\infty \) respectively. Since we shall be concerned with expansions where \( f(u) \to 0 \) as \( |u| \to \infty \), the asymptotic behavior of \( v(a,s) \) and \( v_a(a,s) \) is largely determined by the integrals (3.4.16) and (3.4.17) in the region where \( \varepsilon \) is close to zero.

We have reduced our boundary conditions to

\[ (3.4.18) \quad \mathbb{E}[f(s^{1/2}(a_u+\varepsilon))] = \gamma s^{-1} + s^{-1/2} \psi^+(a_u) \]

\[ (3.4.19) \quad \mathbb{E}[f(s^{1/2}(a_l+\varepsilon))] = \gamma s^{-1} - a_l s^{-1/2} + s^{1/2} \psi^-(a_l) \]

\[ (3.4.20) \quad \mathbb{E}[e^s f(s^{1/2}(a_u+\varepsilon))] = -s^{1/2}[1-\Phi(a_u)] \]

\[ (3.4.21) \quad \mathbb{E}[e^s f(s^{1/2}(a_l+\varepsilon))] = -s^{-1/2} + s^{1/2} \phi(a_l) \].

The determination of the expansion has now been reduced to two separate problems since equations (3.4.18) and (3.4.20) involve only \( a_u \) and the behavior of \( f \) at \( +\infty \) while equations (3.4.19) and (3.4.21) involve only \( a_l \) and the behavior of \( f \) at \( -\infty \).
Now transform equations (3.4.18) and (3.4.20) using

\[ y' = \gamma^{-1/3} y \]

(3.4.22) \[ s' = \gamma^{-2/3} s \]

\[ f_\perp(x) = \gamma^{-1/3} f(\gamma^{1/3} x) \]

to obtain

(3.4.23) \[ E(f_\perp[s^{1/2}(\alpha_u' + \varepsilon)]) = s'^{-1} + s'^{1/2} \psi(\alpha_u') \]

(3.4.24) \[ E(\varepsilon f_\perp[s^{1/2}(\alpha_u' + \varepsilon)]) = -s'^{1/2}[1 - \delta(\alpha_u')] \].

This is exactly the set of equations considered by Chernoff [7] (page 31, equations 35a and 35b). The formal expansions are given there as:

(3.4.25) \[ \ln s' = \alpha_u'^2/3 + \ln(8\pi)/s + 2\alpha_u'^{-2} + \cdots \]

(3.4.26) \[ f_\perp(x) = x^{-2}[3\ln x^2 - 3\ln[3\ln x^2] - [\ln 8\pi + 1] - \cdots] \].

Chernoff [7] also considers an alternate method of obtaining an expansion which leads to the formal expansion for the boundary:

\[ \alpha_u'^2 = 3w(1 - 2w^{-2}/3 - 5w^{-3}/9 - 22w^{-4}/9 - \cdots) \]

where
\[ w = \ln s' - \frac{\ln(8\pi)}{3} \, . \]

Upon using (3.4.22) to transform back to our \((y,s)\) plane, these formal expansions become:

(3.4.27) \[ \ln s = a_u^2/3 + \ln(8\pi y^2)/3 + 2a_u^{-2} + \cdots \]

(3.4.28) \[ f(x) \approx \gamma x^{-2}(3\ln x^2 - 3\ln[3\ln x^2] - [\ln(8\pi y^2) + 1] - \cdots) \]

and the alternative formal expansion becomes:

(3.4.29) \[ a_u^2 = 3w[1 - 2w^{-2}/3 - 5w^{-3}/9 - 22w^{-4}/9 - \cdots] \]

where

(3.4.30) \[ w = \ln s - \frac{\ln(8\pi y^2)}{3} \, . \]

For the lower boundary, transform equations (3.4.19) and (3.4.21) using

\[ \beta = -a_u \]

(3.4.31) \[ \epsilon^* = -\epsilon \]

\[ f_2(x) = f(-x) \]

to obtain

(3.4.32) \[ E[f_2[s^{1/2}(\beta+\epsilon^*)}] = \gamma s^{-1} + \beta s^{-1/2} + s^{1/2} f^+ (\beta) \]
\begin{align}
(3.4.33) \quad E^* \{ \epsilon^* f_2 [s^{1/2} (\beta + \epsilon^*)] \} &= s^{-1/2} + s^{1/2} [1 - \phi(\beta)]
\end{align}

where $E^*$ denotes expectation with respect to the normal random variable $\epsilon^*$ which has mean zero and variance one. Along the boundary $\alpha_s$, the term $\gamma s^{-1}$ in equation (3.4.32) is relatively negligible.

To be more specific, along the boundary represented by (3.4.14),

$\beta = [2n \ln s - \ln (2n \ln s) + O(1)]^{1/2}$, which implies that

$\beta s^{-1/2} - s^{1/2} (\ln s)^{1/2}$. Further, upon using the expansion

\begin{align}
(3.4.34) \quad 1 - \phi(\beta) &= \phi(\beta) (\beta^{-1} - \beta^{-3} + (1.3) \beta^{-5} - (1.3.5) \beta^{-7} - \ldots)
\end{align}

which leads to

\begin{align}
(3.4.35) \quad \Psi^+(\beta) &= \phi(\beta) (\beta^{-2} - (1.3) \beta^{-4} + (1.3.5) \beta^{-6} - \ldots)
\end{align}

we conclude $s^{1/2} \Psi^+(\beta) - s^{1/2} (\ln s)^{-1/2}$. Since the last two terms on the right hand side of (3.4.32) can be expanded as power series in $s^{-1/2} (\ln s)^{-k/2}$ for $k = -1, 0, 1, 2, \ldots$, we may ignore the $\gamma s^{-1}$ term.

Thus we are led to consideration of the system

\begin{align}
(3.4.36) \quad E^* \{ f_2 [s^{1/2} (\beta + \epsilon^*)] \} &= \beta s^{-1/2} + s^{1/2} \Psi^+(\beta)
\end{align}

\begin{align}
(3.4.37) \quad E^* \{ \epsilon^* f_2 [s^{1/2} (\beta + \epsilon^*)] \} &= s^{-1/2} + s^{1/2} [1 - \phi(\beta)]
\end{align}

together with the assumed form of the boundary.
\begin{align*}
(3.4.38) \quad \ln s &= \beta^2/2 + \ln \beta^2/2 + b_1 + b_2 \beta^{-2} + b_3 \beta^{-4} + \cdots .
\end{align*}

This is exactly the set of equations considered by Chernoff and Ray [11] (pages 1397-1398, equations 6.3 - 6.10). The formal expansions are given there as

\begin{align*}
(3.4.39) \quad \ln s &= \beta^2/2 + \ln \beta^2/2 + \ln(\pi)/2 + \beta^{-2} + \beta^{-4}/2 + \cdots \\
(3.4.40) \quad f_2^-(x) &= x^{-1} \{2\ln x^2 - 3\ln[2\ln x^2] - [\ln(\pi) - 1] \\
&\quad \quad + (2\ln x^2)^{-1}[9\ln(2\ln x^2) + 3\ln 8\pi - 4] + \cdots \}.
\end{align*}

This formal expansion obtained for the boundary does not in any way depend upon the parameter $\gamma$ since we were able to ignore the $\gamma s^{-1}$ term in (3.4.32). It would be of interest to estimate what the effect of this term is. That is, we would like to determine a higher order correction to the formal expansion obtained for the boundary through consideration of the effect of the $\gamma s^{-1}$ term.

Let $\beta_0$ and $f_{20}$ denote the solution (3.4.39) and (3.4.40) obtained to the equations (3.4.36) and (3.4.37). Perturb this solution by considering

\begin{align*}
\beta &= \beta_0 + \delta \\
(3.4.41) \quad f_2^-(x) &= f_{20}(x) + f_{21}(x).
\end{align*}
Substituting these expressions in equations (3.4.32) and (3.4.33) and proceeding in the manner indicated by Chernoff and Ray [11], we obtain the following formal expansions

(3.4.42) \[ \delta = -(\gamma/2) s^{-1/2}(\ln s)^{-1}\{1+(2\ln s)^{-1}[\ln(2\ln s)+\ln(8\pi)-4]+\cdots\} \]

(3.4.43) \[ f_{21}(x) = \gamma x^{-2}(2\ln x^2 - 3\ln(2\ln x^2) - [\ln(8\pi) + 3] + \cdots) \].

It is of interest to note that, at least for the first two terms, we have the relation \( f_{21}(x) = \gamma x^{-1} f_{20}(x) \). This would seem to indicate that an expansion of the form \( f_2(x) = f_{20}(x) \cdot (1 + c_1 x^{-1} + c_2 x^{-2} + \cdots) \) might be appropriate. This has not been investigated further.

To summarize, we transform these formal expansions for the boundaries back to the \((y, s)\) plane. Equation (3.4.27) yields the formal expansion for the upper boundary (denoted \( y_{u1} \))

(3.4.44) \[ y_{u1}^2 = s \cdot \{3\ln s - \ln(8\pi^2) - 2(\ln s)^{-1} - \cdots\} \).

Equation (3.4.29) yields the alternate expansion for the upper boundary (denoted \( y_{u2} \))

(3.4.45) \[ y_{u2}^2 = 3sw(1 - 2w^{-2}/3 - 5w^{-3}/9 - 22w^{-4}/9 - \cdots) \]

where \( w \) is given by equation (3.4.30). Equation (3.4.39) yields the formal expansion for the lower boundary obtained by neglecting the
\( y_{\lambda_0}^{-\frac{1}{2}} \) term (denoted by \( y_{\lambda_0} \))

\[ (3.4.46) \quad y_{\lambda_0}^2 = 2\lambda \ln s - \ln(2\lambda \ln s) - \ln(8\pi) - (\ln s)^{-1} - \cdots \]

Equation (3.4.42) yields the higher order correction which leads to the expansion (denoted by \( y_{\lambda_c} \))

\[ (3.4.47) \quad y_{\lambda_c} = y_{\lambda_0} + (\gamma / 2)(\ln s)^{-1} \left[ 1 + (2\ln s)^{-1} \ln(2\ln s) \right. \]

\[ + \left. (\ln(8\pi) - 4) \right] + \cdots \]

The behavior of the formal expansions of this section will also be examined further in Section 4.1.

As indicated at the end of Section 3.3, use of the appropriate transformations would transform the results to any of the planes of interest.

3.5 Suboptimal Procedures

For purposes of comparison with the optimal procedure, we now define two different suboptimal procedures. The first of these is a Wald-type procedure which we will call procedure \( W \). The second is the procedure which consists of stopping at any point at which there is no fixed sample size procedure which does as well as stopping. This procedure will be called procedure \( F \). Procedure \( F \) regarded as a one-stage sequential stopping procedure, call it \( F_1 \), could be generalized to procedure \( F_m \) which is defined inductively. Let \( F_0 \)
denote the procedure which consists simply of stopping. Procedure \( F_m \) consists of continuing for a fixed time \( t_m \geq 0 \) after which time procedure \( F_{m-1} \) is used, where the value of \( t_m \) is determined so as to minimize the risk incurred. (See [23], [24].) This generalization has not been investigated here. We will be concerned with the comparison of procedures \( F \) and \( W \) with the optimal procedure in terms of both the nature of the region of continuation and in terms of the risks incurred.

3.5.A Procedure \( W \)

For the continuous-time stopping problem developed in Section 3.1, a Wald-type procedure consists of continuation as long as \( B^* < X(t) < A^* \) where \( B^* < 0 \) and \( A^* > 0 \). Transforming to the \( Y(s) \) process, this procedure becomes:

"Continue as long as \( Bs < Y(s) < As \), where \( B < 0 \) and \( A > 0 \)."

Thus the Wald-type procedure corresponds to straight line boundaries for the \( Y(s) \) process. Both of the straight line boundaries pass through the point \((0,0)\) in the \((y,s)\) plane although the range \( s < 1 \) is, of course, meaningless for our problem. For a particular initial point \((y_0,s_0)\) of the \( Y(s) \) process, the best Wald-type procedure is that particular procedure that yields the minimum risk at \((y_0,s_0)\).

Thus the \( A \) and \( B \) which specify the procedure depend upon the initial point \((y_0,s_0)\). The values of \( A \) and \( B \) that define the best Wald-type procedure for a particular initial point \((y_0,s_0)\) can be determined computationally, as will be indicated in Section 4.3.
3.5. B Procedure $F$

For the continuous-time stopping problem developed for the $Y(s)$ process in Section 3.1, consider the procedure which consists of stopping at any point at which there is no fixed sample size procedure which does as well as stopping. For a given value of $s$ the stopping boundary, denoted by $y_F(s)$, is defined by

$$(3.5.B.1) \quad d(y_F(s), s) = \inf_{1 \leq s_{\perp} \leq s} \int_{-\infty}^{\infty} d[y_F(s)+s_{\perp}, \frac{1}{2}, \epsilon, s_{\perp}][\phi(\epsilon)] d\epsilon$$

where $d(y, s)$ is defined by $(3.1.10)$. Defining

$$(3.5.B.2) \quad \Delta^2 = s - s_{\perp}$$

and setting the derivative of the right-hand side of $(3.5.B.1)$ with respect to $s_{\perp}$ equal to zero leads to the two determining relations

$$(3.5.B.3) \quad d(y_F(s), s) = \gamma/s_{\perp} - y_F - \Delta(1-s_{\perp}^{-1})y^+(y_F/\Delta)$$

$$(3.5.B.4) \quad 0 = -\gamma/s_{\perp}^2 + \frac{1}{2} \Delta^{-1}(1-s_{\perp}^{-1})\phi(y_F/\Delta) - \Delta s_{\perp}^{-2}y^+(y_F/\Delta)$$

where $\phi$ and $y^+$ are defined by $(3.3.10)$ and $(3.4.2)$ respectively. These two determining relations yield two sets of two equations in two unknowns $y_F$ and $s_{\perp}$ for each value of $s$; one set for $y_F \geq 0$ and the other for $y_F \leq 0$. 

43
We now consider the asymptotic behavior of the solutions of these equations for large values of $s$. At the upper boundary, $\gamma_F = \gamma_{Fu} > 0$ and we assume $s/s_1$ to be large. After a certain amount of algebra, we are led to the formal expansion

\begin{equation}
(3.5.B.5) \quad \gamma_{Fu}^2 = s \{3\ln s - 4\ln[3\ln s] - \ln(\pi^2/2) - \cdots \}.
\end{equation}

At the lower boundary, $\gamma_F = \gamma_{Fl} < 0$ and we assume $s/s_1$ to be large. In a similar fashion we are led to the formal expansion

\begin{equation}
(3.5.B.6) \quad \gamma_{Fl}^2 = s \{2\ln s - 5\ln[2\ln s] - \ln(\pi/2) - \cdots \}.
\end{equation}

Notice how these expansions compare to the expansions (3.4.44) and (3.4.46) obtained for the boundaries of the optimal procedure.

Now consider the asymptotic behavior of the solutions of these equations in the region $s$ close to 1. At the upper boundary we assume that as $s$ approaches 1, $\gamma_{Fu}/\Delta$ approaches a constant denoted by $c_u$. If we further assume that $\Delta^2 = o(s-1)$, after a certain amount of algebra we are led to the relations $\phi(c_u) = 2c_u[1-\phi(c_u)]$ and $\Delta = \frac{1}{2} (s-1) \phi(c_u)/\gamma$. These relations lead to

\begin{equation}
(3.5.B.7) \quad s_1 = s - \left[\frac{1}{2}(s-1) \phi(c_u)/\gamma\right]^2
\end{equation}

\begin{equation}
(3.5.B.8) \quad \gamma_{Fu} = \frac{1}{2} (s-1) c_u \phi(c_u)/\gamma
\end{equation}
where $c_u = .612$. An identical analysis at the lower boundary leads
to the same $s_1$ and

\[(3.5.B.9) \quad y_{F_0} \approx -y_{F_u} \cdot\]

Notice how these expansions compare to the expansions (3.3.19) and
(3.3.20) obtained for the boundaries of the optimal procedure.

The determining relations (3.5.B.3) and (3.5.B.4) could also be
used to obtain, for each fixed value of $s$, the value of both $y_{F_u}(s)$
and $y_{F_0}(s)$ by numerical methods. This was in fact carried out and
will be discussed in Section 4.3.

Finally from (3.5.B.1) we obtain the result that the function
$h(y,s)$ defined by

\[(3.5.B.10) \quad h(y,s) = \inf_{1 < s_1 < s} \int_{-\infty}^{\infty} d[y+(s-s_1)^{1/2}, s_1]q(\varepsilon) d\varepsilon\]

represents the risk of the best fixed sample size procedure at the
point $(y,s)$. In particular, for $y = 0$,

$h(0,s) = \gamma/s_1 - (s-s_1)^{1/2}(1-s_1^{-1})(2\pi)^{-1/2}$

where $s_1$ is the solution of

$0 = -\gamma/s_1^2 + \frac{1}{2} (s-s_1)^{-1/2}(1-s_1^{-1})(2\pi)^{-1/2} - (s-s_1)^{1/2}(2\pi)^{-1/2}/s_1^2$.
A simple computation yields

\[(3.5.B.11) \quad h(0,s) = \gamma/s - (s-1)^2/(8\pi\gamma) + \cdots\]

for \(s\) close to 1 and

\[(3.5.B.12) \quad h(0,s) = -(s/2\pi)^{1/2} + \pi^{-1/2} + (\pi^{-1/2}\gamma^{-1/4})/(2\pi\gamma)^{1/2} + \cdots\]

for \(s\) large.

These expansions could be used for the purposes of comparing the best fixed sample size procedure with the various other procedures that have been proposed. However, since procedure \(F\) is an improvement on the best fixed sample size procedure, this procedure will not be considered further.
Chapter 4

Computational Results

4.1 Solution of the Wiener Process Problem

In Section 3.1 it has been demonstrated that the problem of interest is a continuous-time stopping problem for the Wiener process $Y(s)$ of independent increments (in the $-s$ scale) with the properties $E[dY(s)] = 0$ and $\text{Var}[dY(s)] = -ds$ and defined on the region $s \geq 1$. Further, this stopping problem is defined by the stopping cost function $d(y,s)$ as given in (3.1.10).

Consider now the discrete-time discrete-step process $Y'(s')$ defined on the set of points $S_\Delta = \{s' : s' = 1 + k \cdot \Delta, \ k = 0, 1, 2, \ldots\}$ subject to $Y'(s' - \Delta) = Y'(s') \pm \Delta^{1/2}$ each with probability $1/2$ independent of the past, that is, independent of $\{Y'(s' + k \Delta), k = 0, 1, 2, \ldots\}$. Then $Y'(s')$ is a process with independent increments and with mean zero and variance one per unit change in $-s'$. Thus given the stopping problem for the Wiener process $Y(s)$ defined by $d(y,s)$, we can imitate it by the use of a small $\Delta$ in the $Y'(s')$ process. For this process, however, the backward induction algorithm reduces to:

\[ (4.1.1) \quad \rho'(y',s') = \min\{d(y',s'), \frac{1}{\Delta} [\rho'(y'+\Delta^{1/2},s'-\Delta) \]
\[ + \rho'(y'-\Delta^{1/2},s'-\Delta)] \]

which is a very simple algorithm to implement. As $\Delta$ approaches zero, the result of this algorithm will converge to the solution of the Wiener process stopping problem.
Considerations identical to those of Chernoff [8] indicate that the following stopping problem is relevant for the consideration of the difference between the boundaries obtained from the $Y'(s')$ process and the optimal boundaries for the Wiener process stopping problem:

Let $Y_n$ be a process starting at $y_0$ for $n = n_0$, $n_0$ a positive integer. As $n$ decreases, let $Y_{n-1} = Y_n + u_n$ where $\{u_n\}$ are independent random variables with common distribution given by $u_n = \pm 1$ each with probability $1/2$. The statistician may stop at any integer time $n$ ($0 < n \leq n_0$) and receive a payoff of zero or at a cost of one decrease $n$ by one. If the player continues until $n = 0$, the game ends and he receives a payoff of zero if $Y_0 > 0$ and $Y_0^2$ if $Y_0 \leq 0$.

This game is subject to analysis and yields the following solution: There is a stopping boundary $\bar{y}_n$ such that at time $n$ the player stops if $Y_n \geq \bar{y}_n$ and continues if $Y_n < \bar{y}_n$. The values of $\bar{y}_n$ can be computed by means of backward induction, and the first few values of $\bar{y}_n$ are given by:

\[
\begin{align*}
\bar{y}_1 &= \bar{y}_2 = 1 - 2^{1/2} \approx -0.414, \\
\bar{y}_3 &= \bar{y}_4 = (5 - 2(10)^{1/2})/3 \approx -0.442, \\
\bar{y}_5 &= \bar{y}_6 = (11 - 4(11)^{1/2})/5 \approx -0.453, \\
\bar{y}_7 &= \bar{y}_8 = (93 - 8(186)^{1/2})/35 \approx -0.460, \\
\bar{y}_9 &= \bar{y}_{10} = -(386)^{1/2}/((386)^{1/2} + (512)^{1/2}) \approx -0.465, \\
\bar{y}_{11} &= \bar{y}_{12} = -(793)^{1/2}/((793)^{1/2} + (1024)^{1/2}) \approx -0.468, \ldots.
\end{align*}
\]
Moreover, it can be shown that as $n \to +\infty$, $y_n \to -1/2$ monotonically. This result leads to what has been called a "continuity correction" by Chernoff [6], relating the boundaries obtained from the $Y'(s')$ process and the optimal boundaries for the Wiener process problem. If we denote the upper and lower boundaries for the Wiener process ($Y'(s')$ process) by $y_u$ and $y_l$ ($y_u'(\Delta)$ and $y'(\Delta)$) respectively, then the use of this "continuity correction" leads to the following relations:

\begin{align*}
(4.1.2) & \quad y_u = y_u'(\Delta) + (1/2)\Delta^{1/2} + o(\Delta^{1/2}) \\
(4.1.3) & \quad y_l = y_l'(\Delta) - (1/2)\Delta^{1/2} + o(\Delta^{1/2}) .
\end{align*}

Thus by computing the boundaries for the $Y'(s')$ process (for a given value of $\Delta$) and by subsequently applying this "continuity correction" we are able to obtain close approximations to the boundaries for the Wiener process problem.

The computation of the boundaries has been carried out for the cases $\gamma = 0.1, 0.5$ and 2.0. For the purpose of saving computation time, the computation of these boundaries was carried out in a slightly different fashion than indicated above. Beginning with a small value of $\Delta$, successively larger values of $\Delta$ were employed to determine the boundaries in successively larger intervals of $s$. These boundaries, determined in overlapping intervals, were then superimposed to obtain the optimal boundary. For each value of $\gamma$, the optimal boundary was determined in the region $1 < s < 10,000$. 

49
Recall that in terms of \( y \) and \( s \), the variables \( \alpha \), \( z \), and \( t \) are defined by \( \alpha = y/s^{1/2} \), \( z = y/s \), and \( t = 1/s \). The optimal boundaries for different values of \( \gamma \) were graphically compared in the \((z,t)\) plane. This graph appears as Figure 1. In Figure 2 these optimal boundaries are compared in the \((\alpha,t)\) plane. Recalling the work of Section 3.4, we see that in the \((\alpha,t)\) scale, the optimal boundaries indicate the number of standard deviations at which it is desirable to stop and make a decision. Due to the nature of these boundaries, it is of interest to examine the region \( t \) close to zero in more detail. This has been accomplished in Figure 3 in which the region \( 0.00 \leq t \leq 0.05 \) of Figure 2 has been redrawn.

An enlightening interpretation of these boundaries is the following: Suppose that at time \( t \) (note from (2.4.2) that \( t \) is the fraction of the available information that has already been collected) the player stopped and decided to test the hypothesis \( \mu = 0 \). Then \( \alpha_u \) is the number of standard deviations required for stopping and deciding in favor of the unknown treatment, while \( \alpha_k \) is the number of standard deviations required for stopping and deciding in favor of the known treatment. Then \( 1 - \Phi(\alpha_u) \) and \( \Phi(\alpha_k) \) are corresponding nominal significance levels. From a classical point of view, one can regard the player as continuously testing the hypothesis \( \mu = 0 \) against each of two alternatives. As the proportion of the total available information already collected varies from 0 to 1, both the nominal significance levels required to stop and make a decision become less stringent, increasing from 0 to 1/2. Although these nominal significance levels serve as a convenient description of the procedure,
FIGURE 2
Solution of Wiener Process Problem -- ($\alpha$,t) Plane
their use should not be confused with that of the standard significance levels which are not applicable here. In Figure 4 we have graphed the boundaries in the \((\phi(s), t)\) plane. Again due to the nature of these boundaries, it is of interest to examine the region \(t\) close to zero in more detail. This is accomplished in Figure 5 in which the region \(0.00 \leq t \leq 0.05\) of Figure 4 has been redrawn.

In Section 3.3 we have determined asymptotic expansions for the optimal boundaries in the region \(s\) close to 1, which corresponds to the region \(t\) close to 1. In Figures 6, 7, and 8 the behavior of these expansions is considered for the cases \(\gamma = 2.0, 0.5\) and 0.1 respectively. In particular, in each of these graphs, the optimal boundary is compared to the first, second and third order expansions. As can be seen from these figures, the behavior of the expansion deteriorates rapidly as the value of \(\gamma\) decreases. For the case \(\gamma = 2.0\), the asymptotic expansion provides a remarkably close fit in the region \(0.5 \leq t \leq 1.0\). However, for the case \(\gamma = 0.5\), the fit is adequate only in the region \(0.75 \leq t \leq 1.00\) and for the case \(\gamma = 0.1\) the fit is adequate only for \(0.98 \leq t \leq 1.00\). This behavior is not particularly surprising, since the expansions were assumed to be a power series in \((s-1)\) and the coefficients then were determined to be polynomials in inverse powers of \(\gamma\). (See (3.3.12) - (3.3.16).) Perhaps if some other form of expansion were considered, this poor behavior could be avoided. This has not been investigated further here.

In Section 3.4 we have determined asymptotic expansions for the optimal boundaries in the region \(s\) large which corresponds to the
region \( t \) close to 0. In Figures 9, 10, and 11 the behavior of the expansion (3.4.45) (up to four terms) is compared to the optimal upper boundary. In addition the expansion (3.4.46) (up to four terms) is compared to the optimal lower boundary. The comparisons are considered in the range \( 0.0001 < t < 0.01 \). These comparisons clearly indicate that for the values of the parameter \( \gamma \) considered here, the formal expansions provide an adequate fit only in the region \( t \) very close to zero, that is, in the region of very large values of \( s \).

As a final note, we indicate what the asymptotic expansions obtained in Sections 3.3 and 3.4 imply as regards the behavior of the optimal boundaries in the \((\phi(\alpha), t)\) scale. This is relevant for our interpretation of the problem as a problem of continuously testing the hypothesis \( \mu = 0 \) against the two alternatives. The asymptotic expansions imply that as \( t \to 0 \),

\[
1 - \phi(\alpha_u) = 2\gamma t^{3/2} (-3\ln t)^{-1/2}
\]

and

\[
\phi(\alpha_2) = 2t
\]

while as \( t \to 1 \),

\[
1 - \phi(\alpha_u) = \frac{1}{2} - t^{-1/2} (1-t) / 4\gamma (2\pi)^{1/2}
\]

and

\[
\phi(\alpha_2) = \frac{1}{2} - t^{-1/2} (1-t) / 4\gamma (2\pi)^{1/2}
\]
These results indicate the rate at which the nominal significance
levels approach 0 and 1/2 as t approaches 0 and 1 respectively.

4.2 Solution of the Binomial Problem

In Section 2.3 we have obtained the optimal solution for the
binomial problem in terms of the five parameters $p_0$, $c$, $N$ and
$(a_0,b_0)$. We have seen that in fact the optimal solution can be speci-
fied in terms of the three parameters $p_0$, $c$ and $M$ with $(a_0,b_0)$
representing the starting point in the stopping problem. If we define
a grid point in the $(z,t)$ plane to be any possible $(z,t)$ point
arising from the transformation (2.4.2), the optimal solution classi-
fies each $(z,t)$ grid point as either a stopping point or a continua-
tion point. For the purposes of this section, we have restricted both
$a_0$ and $b_0$ to be integer-valued. Note that in this case the grid
is simply determined by the parameters $M$ and $p_0$. (See (2.4.2).)

The solution of the FBP (3.2.3) which is a solution of the Wiener
process problem provides an approximation to the solution of the
binomial problem. The only parameter appearing in the Wiener process
problem is the parameter $\gamma$ which corresponds to $c[M/p_0(1-p_0)]^{1/2}$
in the binomial problem. The solution to the Wiener process problem
provides a pair of curves $z_u(t) = y_u(s)/s$ and $z_\gamma(t) = y_\gamma(s)/s$
(where $t = 1/s$) which define a region of continuation in the $(z,t)$
plane. The evaluation of the performance of the approximation can then
be considered in terms of the question: Do the curves $z_u(t)$ and
$z_\gamma(t)$ derived from the Wiener process problem correctly classify the $(z,t)$ grid points of the binomial problem as either stopping points or continuation points?

The question was considered in the following manner: For a given value of the parameter $\gamma$, we have seen in Section 4.1 how to determine the solution to the Wiener process problem. For each of a set of values of $(p_0, M)$, determine that value of $c$ that yields this given value of $\gamma$. Then for each of these sets of parameters $p_0$, $M$ and $c$, determine the optimal solution to the binomial problem. For each particular set of parameters it can then easily be determined which of the $(z,t)$ grid points are classified incorrectly by the curves $z_u(t)$ and $z_\gamma(t)$. Different values of $\gamma$ can then be considered.

It is important to remember that the curves $z_u(t)$ and $z_\gamma(t)$ are appropriate only for the Wiener process problem. Thus, for large values of the parameter $M$, these curves would be expected to provide an excellent approximation in the region in which the usual normal approximation to the binomial is adequate, namely, for $t$ removed from zero. It would be of interest, however, to determine some way of adjusting these curves in order to improve the accuracy of the approximation for moderate values of $M$ where the discrete nature of $t$ becomes important in the binomial problem. Recall that we have demonstrated in Section 4.1 that the curves that are appropriate for the Wiener process problem can be obtained from the solution of the discrete-time discrete-step process problem by means of the "continuity correction"
summarized in equations (4.1.2) and (4.1.3). These results suggest that the accuracy of the approximation can be improved by the use of a similar "continuity correction" applied in reverse. More specifically, if we denote by \( z_u(t, M) \) and \( z_\lambda(t, M) \) the boundaries to be used in the \((z, t)\) plane for a particular value of the parameter \( M \), we are led to the corrections given by:

\[
\begin{align*}
(4.2.1) \quad z_u(t, M) &= z_u(t) - (1/2)M^{-1/2} \\
(4.2.2) \quad z_\lambda(t, M) &= z_\lambda(t) + (1/2)M^{-1/2}.
\end{align*}
\]

From the nature of the discrete-time discrete-step process \( Y'(s') \) from which these corrections were obtained, we recognize that these corrections can be expected to perform well only for those particular binomial problems in which \( p_0 \) is very nearly equal to 0.5.

Presumably, by considering more general discrete-time discrete-step processes, the appropriate corrections for other values of \( p_0 \) could be obtained. These corrections should be of the form:

\[
\begin{align*}
(4.2.3) \quad z_u(t, M, p_0) &= z_u(t) - a_u(p_0) M^{-1/2} \\
(4.2.4) \quad z_\lambda(t, M, p_0) &= z_\lambda(t) + a_\lambda(p_0) M^{-1/2}
\end{align*}
\]

where the constants \( a_u(p_0) \) and \( a_\lambda(p_0) \) have yet to be determined except for the case \( p_0 = 0.5 \), in which case we have already determined \( a_u(0.5) = a_\lambda(0.5) = 1/2 \). Although these corrections may be
fairly substantial for small values of $M$, they become negligible as $M$ becomes large. We emphasize that we have not provided a derivation to justify this use of this correction. However, in view of Chernoff [8], we anticipate no difficulty in establishing the validity of this correction.

The accuracy of the approximation provided by the solution of the Wiener process problem has been examined for the parameter values $\gamma = 0.1, 0.5$ and 2.0; $p_0 = 0.1, 0.5$ and 0.9; $M = 100$ and 400. These results appear in Figures 12-20. Figure 12 is an examination of the accuracy of approximation provided by the Wiener process solution to two separate binomial problems: the binomial problem with $p_0 = 0.5$, $M = 100$ and that value of $c$ ($= 0.0050$) that yields a value of $\gamma = 0.1$ and the binomial problem with $p_0 = 0.5$, $M = 400$ and that value of $c$ ($= 0.0025$) that yields a value of $\gamma = 0.1$. In this figure, the curves $z_u(t)$ and $z_l(t)$ that form the solution to the Wiener process problem with $\gamma = 0.1$ have been graphed, the $(z,t)$ grid has been indicated for both the cases $M = 100$ and 400 and the grid points that are incorrectly classified by the curves $z_l(t)$ and $z_u(t)$ are indicated for both the cases $M = 100$ and 400. In addition, the curves $z_u(t,100)$ and $z_l(t,100)$ as given by equations (4.2.1) and (4.2.2) are graphed. Since for any fixed value of $t$, $z_u(t,400)$ lies midway between $z_u(t,100)$ and $z_l(t)$ and $z_l(t,400)$ lies midway between $z_l(t,100)$ and $z_u(t)$, these curves were not graphed. Figures 13 and 14 then consider the cases $\gamma = 0.5$ and 2.0 respectively with the value of $p_0 = 0.5$. In an identical fashion, Figures 15, 16 and 17 consider the case $p_0 = 0.1$ and Figures 18,
19 and 20 consider the case $p_0 = 0.9$. As noted above, the $(z,t)$ grid is determined simply by the parameters $M$ and $p_0$. For this reason the grid has been indicated only on Figures 12, 15 and 18. Further, Figure 12 is the only figure that is completely labeled.

Examination of the Figures 12-14, which deal with $p_0 = 0.5$, reveals that the solution of the Wiener process problem together with the corrections (4.2.1) and (4.2.2) leads to an exceptionally accurate approximation to the optimal solution for each of the binomial problems considered in these particular figures. In every case, the curves $z_u(t,M)$ and $z_\gamma(t,M)$ classify virtually all of the $(z,t)$ grid points correctly as either stopping or continuation points. The only region where there appears to be any difficulty in these cases is the region $t$ close to zero. As mentioned above, this difficulty was anticipated a priori due to the inadequacy of the usual normal approximation to the binomial in this region. It is remarkable that such an accurate approximation is obtained in these cases in which the horizon sizes that we are considering are moderately small, namely $M = 100$ and 400.

In the same fashion as Figures 12-14 deal with the case $p_0 = 0.5$, Figures 15-17 deal with the case $p_0 = 0.1$ and Figures 18-20 deal with the case $p_0 = 0.9$. Examination of these figures leads us to the immediate observation that the approximation provided by the solution of the Wiener process problem behaves much more poorly in these cases. The curves $z_u(t)$ and $z_\gamma(t)$ misclassify a much larger proportion of the total number of grid points in each of the cases considered in
Comparison of Wiener Process Solution and Binomial Solution -- $\gamma = 0.5$, $P_0 = 0.1$
Figures 15-20 than they do in any of the cases considered in Figures 12-14. Further, in these cases the corrections (4.2.1) and (4.2.2), designed for \( p_0 = 0.5 \), appear to be inadequate. This inadequacy was foreseen since \( p_0 \neq 0.5 \) for these cases. Presumably, the appropriate corrections as given by (4.2.3) and (4.2.4) would substantially improve the adequacy of the approximation. Indeed Figures 15-20 indicate that appropriate corrections of this form will perform well. This good performance led us to carry out the remainder of the work appearing here within the context of the Wiener process problem.

4.3 Evaluation of Suboptimal Procedures

4.3.A Procedure W

In Section 3.5.A we have defined the suboptimal procedure \( W \). This procedure can be specified by two constants \( A \) and \( B \) (both positive). However, these constants must be determined in such a way as to yield the minimum Bayes risk at the starting point \( (y_0, s_0) \).

This determination was carried out in the \( Y'(s') \) process in the following way: For a fixed value of \( \gamma \) and a particular starting point \( (y_0, s_0) \), choose a pair \( (A, B) \) and obtain the risk at the point \( (y_0, s_0) \) for this pair \( (A, B) \) by means of a backward induction algorithm that incorporates the region \( \{(y, s): -B \leq y \leq A, \ s \geq 1\} \) as the continuation region. Then scan on both the values of \( A \) and the values of \( B \). In this manner, a table of risks at the point \( (y_0, s_0) \) is obtained as a function of the values of \( A \) and \( B \). From this table, the optimal \( (A, B) \) pair and the optimal risk at the point
\((y_0, s_0)\) were approximately determined by fitting the risk as a quadratic surface in a neighborhood of the apparent minimum. It is important to note that the risk at any given starting point \((y_0, s_0)\) does not vary greatly with the particular choice of \((A, B)\), as long as the choice of \((A, B)\) is in a neighborhood of the optimal \((A, B)\) pair. Thus even though the optimal \((A, B)\) pair may not be located with great accuracy, the minimum value of the risk at \((y_0, s_0)\) is determined fairly accurately.

These computations were carried out for the parameter values \(\gamma = 0.1, 0.5\) and \(2.0\) and for twelve different starting points, namely:

\[
\begin{align*}
s_0 &= 101, \quad y_0 = 0, 10 \\
s_0 &= 501, \quad y_0 = 0, \pm 10, 20 \\
s_0 &= 1001, \quad y_0 = 0, \pm 10, \pm 20, 30.
\end{align*}
\]

The results of these computations are summarized in Table 1.

Due to the fact that the boundaries obtained depend upon the particular starting point considered, it was not feasible to graphically compare the boundaries obtained for each of the various different starting points considered. Instead the starting point \(s_0 = 1001, \quad y_0 = 0\) was considered as a particular example. From Table 1 we see that the boundaries are given as follows:
Table 1
Boundaries and Risks of Procedure W for Different Starting Points and Different Values of the Parameter $\gamma$

<table>
<thead>
<tr>
<th>$s_0$</th>
<th>$y_0$</th>
<th>$\gamma = 0.1$</th>
<th></th>
<th>$\gamma = 0.5$</th>
<th></th>
<th>$\gamma = 2.0$</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>A</td>
<td>B</td>
<td>Risk</td>
<td></td>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>101</td>
<td>0</td>
<td>.342</td>
<td>.207</td>
<td>-3.694</td>
<td></td>
<td>.289</td>
<td>.201</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>.396</td>
<td>.261</td>
<td>-10.727</td>
<td></td>
<td>.337</td>
<td>.262</td>
</tr>
<tr>
<td>501</td>
<td>-10</td>
<td>.379</td>
<td>.204</td>
<td>-4.571</td>
<td></td>
<td>.334</td>
<td>.201</td>
</tr>
<tr>
<td>1001</td>
<td>-20</td>
<td>.384</td>
<td>.182</td>
<td>-4.853</td>
<td></td>
<td>.334</td>
<td>.181</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>.386</td>
<td>.196</td>
<td>-12.450</td>
<td></td>
<td>.335</td>
<td>.193</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>.388</td>
<td>.201</td>
<td>-18.097</td>
<td></td>
<td>.300</td>
<td>.196</td>
</tr>
<tr>
<td></td>
<td>30</td>
<td>.395</td>
<td>.203</td>
<td>-32.823</td>
<td></td>
<td>.300</td>
<td>.198</td>
</tr>
</tbody>
</table>
\[ \gamma = 0.1, \quad A = 0.386, \quad B = 0.196 \]

\[ \gamma = 0.5, \quad A = 0.335, \quad B = 0.193 \]

\[ \gamma = 2.0, \quad A = 0.296, \quad B = 0.186 \]

In the \((z,t)\) plane, the boundaries for the procedure \(W\) specified by \((A,B)\) are given simply by \(z_u(t) = A, \quad z_v(t) = -B\). Due to the simple form of these boundaries, they were not represented graphically. The boundaries for the particular starting point \(s_0 = 1001, \quad y_0 = 0\) and for the different values of the parameter \(\gamma\) were compared in the \((a,t)\) and \((\Phi(a),t)\) planes. These comparisons appear in Figures 21 and 22 respectively. In order to compare procedure \(W\) for this particular starting point to the optimal procedure, these figures should be compared to Figures 2 and 4 respectively.

Comparisons in terms of both boundaries and risks will be discussed further in Section 4.4.

4.3.B Procedure \(F\)

In Section 3.5.B we have defined the suboptimal procedure \(F\). The boundaries for this procedure were determined numerically by means of a Newton-Raphson iterative scheme designed to solve the two equations \((3.5.B.3)\) and \((3.5.B.4)\). The initial guess of the value of the solution was provided by the asymptotic expansions of Section 3.5.B in the region \(s\) close to 1. For larger values of \(s\), the initial guess was determined by a quadratic extrapolation making use of the value of the solution for smaller values of \(s\). These boundaries were evaluated
FIGURE 22
Boundaries of Procedure W -- Case \( s_0 = 1001, \ y_0 = 0 \) -- \((\phi(x),t)\) Plane
for the parameter values $\gamma = 0.1, 0.5$ and 2.0. Note that these boundaries are exact for the Wiener process problem (up to the accuracy of the Newton-Raphson scheme). The $Y'(s')$ process was not used in obtaining these boundaries. The boundaries for the different values of $\gamma$ were graphically compared in the $(z,t)$, $(a,t)$ and $(\Phi(a),t)$ planes in Figures 23, 24 and 25 respectively. In order to compare procedure $F$ to the optimal procedure, these figures should be compared to Figures 1, 2 and 4 respectively.

Once the boundaries have been obtained, the risks incurred by the use of this procedure can easily be evaluated using the $Y'(s')$ process with a backward induction algorithm that incorporates these boundaries as defining the continuation region.

Comparisons in terms of both boundaries and risks will be discussed further in Section 4.4.

4.4 Comparisons of Optimal and Suboptimal Procedures

4.4.A Boundaries

We have determined the boundaries for the optimal procedure and for the suboptimal procedures $F$ and $W$ for the parameter values $\gamma = 0.1, 0.5$ and 2.0, as indicated in Section 4.3. Comparison of these procedures in terms of their boundaries indicates how these procedures differ in terms of their stopping rules.

For each value of $\gamma$, the boundaries of procedure $F$ have been compared to those of the optimal procedure in the $(z,t)$ plane. These comparisons appear in Figures 26, 27 and 28. The boundaries for the
procedure $W$ were not compared to those of the other procedures since the boundaries of procedure $W$ depend upon the particular point considered as the starting point. However, due to the simple nature of these boundaries in the $(z,t)$ plane ($z_u(t) = A, \quad z_v(t) = -B$), it is easy to see how the boundaries for a particular procedure $W$ compare to those of procedure $F$ and the optimal procedure simply from the values of the optimal $(A,B)$ pair as evaluated in Table 1. More precisely, for the value $\gamma = 2.0$ the $(A,B)$ pair for the different starting points considered in all cases falls in the region $0.20 \leq A \leq 0.31, \quad 0.18 \leq B \leq 0.23$. Similarly, for $\gamma = 0.5$, the region is $0.29 \leq A \leq 0.34, \quad 0.18 \leq B \leq 0.26$ and for $\gamma = 0.1$, the region is $0.34 \leq A \leq 0.41, \quad 0.18 \leq B \leq 0.26$. These regions can be located on Figures 26, 27 and 28 respectively in order to compare the boundaries of procedure $W$ with those of procedure $F$ and the optimal procedure.

Comparison (in the $(z,t)$ plane) of the boundaries corresponding to procedure $F$ and those corresponding to the optimal procedure demonstrates the fact that procedure $F$ is a conservative procedure in the sense that it always calls for stopping prior to the time the optimal procedure calls for stopping. Of greater interest, however, is the fact that procedure $F$ mimics the form of the optimal boundary. That is, even though procedure $F$ always prescribes early stopping, the shape of the boundary corresponding to procedure $F$ bears a striking resemblance to the shape of the boundary corresponding to the optimal procedure. This observation lends support to the conjecture.
Comparison of Optimal Procedure and Procedure $F = y = 0.5$
that the extension of procedure \( F \) to an \( m \)-stage sequential procedure would yield a very efficient suboptimal procedure. Unfortunately, since procedure \( F \) is defined in an implicit manner, the \( m \)-stage version is difficult to determine.

Comparison of the boundaries corresponding to procedure \( W \) with those of procedure \( F \) and the optimal procedure leads to the observation that whereas the boundaries of the optimal procedure and procedure \( F \) are of the same general shape, the shape of the boundaries corresponding to procedure \( W \) are radically different. This, in fact, is one reason why procedure \( W \) is important. The simple shape of the boundaries makes this a very easy procedure to implement. Thus procedure \( W \) is an interesting alternative to the optimal procedure. Of course, it remains to study the effectiveness of procedure \( W \) as compared to the optimal procedure. This study is contained within the next section.

4.4.B Risks

Of greater interest for the purpose of comparison of suboptimal procedures with the optimal procedure are the risks incurred when these procedures are employed. In order to enable us to make meaningful comparisons of the risks, we should develop some measure of efficiency for our situation. To this end, consider once again the Wiener process problem defined in Section 3.1. Consider the hypothetical experimenter with a crystal ball who is informed of the value of \( \mu \) after it is selected according to the \( N(\mu_0, \sigma_0^2) \) prior probability distribution. Since his gains are given by the output of the particular Wiener process
observed, his action would be to decide to use the unknown process for the entire time $N^*$ if $\mu > 0$ and to use the standard process for the entire time $N^*$ if $\mu \leq 0$. Thus his gain $G$ is given by $N^*\mu$ if $\mu > 0$ and 0 otherwise. Therefore his expected gain is

$$EG = N^* \int_0^\infty \mu (2\pi \sigma_0^2)^{-1/2} \exp\{-(\mu-\mu_0)^2/2\sigma_0^2\} d\mu$$

$$= N^* \sigma_0 \{\phi(\mu_0/\sigma_0) + (\mu_0/\sigma_0) \phi(\mu_0/\sigma_0)\}.$$  

Transforming to the $Y(s)$ process (see (3.1.6) and (3.1.7)) and removing the constant multiplier of (3.1.8), the normalized expected gain is given by the quantity

$$(4.4.3.1) \quad G_{CB}(y_0,s_0) = (1-s_0^{-1})s_0^{1/2} \{\phi(y_0/s_0)^{1/2}$$

$$+ (y_0/s_0)^{1/2} \phi(y_0/s_0)^{1/2}\}.$$  

Note that this quantity is independent of the parameter $\gamma$.

Let us denote the risk of any arbitrary procedure $P$ by $R_P$. Then we have computed the risks $R_{CB} = -G_{CB}$, $R_{opt}$, $R_W$, and $R_F$. Since $R_{opt}$ is the best that can be achieved by any statistician while $R_{CB}$ is the risk under perfect information, we shall define the efficiency of a procedure $P$ in the problem where $(y_0,s_0)$ is the starting point of the Wiener process as:

$$(4.4.3.2) \quad E(P;y_0,s_0) = \frac{R_{opt}(y_0,s_0) - R_{CB}(y_0,s_0)}{R_{F}(y_0,s_0) - R_{CB}(y_0,s_0)}.$$  

93
There is a slight problem with the use of this measure of efficiency. This is due to the fact that whereas $R_{opt}$, $R_W$ and $R_F$ are approximated by the use of the $Y'(s')$ process, $R_{GB}$ is derived using the $Y(s)$ process. Thus it would be of interest to determine what effect the choice of the particular value of $\Delta$ used to perform the computations in the $Y'(s')$ process (see Section 4.1) has upon the evaluation of these risks. We have already seen in Section 4.1 that by using a simple "continuity correction", we are able to transform the boundaries for the $Y'(s')$ process, obtained by the use of a particular value of $\Delta$, to the optimal boundaries for the Wiener process problem. It would clearly be of interest to obtain a similar "continuity correction" for the risks of the various procedures, both optimal and suboptimal. The measure of efficiency (4.4.B.2) proposed above could then be computed using these transformed risks and the resulting values could then be meaningfully interpreted. This has not been accomplished here; instead we have attempted to bypass this problem in an empirical manner. The manner in which this was accomplished is outlined as follows: The risks for the optimal procedure were approximated by the use of the $Y'(s')$ process for the choices of $\Delta = 100, 25, 4,$ and $1$. For the different values of $\gamma$ and for the starting points of interest, the risks obtained were then examined to determine the effect of the changing value of the step size $\Delta$. Generally, the change in the risk obtained resulting from a change in $\Delta$ from $25$ to $4$ appeared in (at worst) the third significant digit, while the change in the risk obtained resulting from a change in $\Delta$ from $4$ to $1$ appeared in (at
worst) the fourth significant digit. From these considerations, it was concluded that the use of the step size \( \Delta = 4 \) was adequate for our purposes. The subsequent computation of the risks \( R_{\text{opt}} \), \( R_F \) and \( R_W \) was carried out using the step size \( \Delta = 4 \) in all cases.

Comparisons of the various procedures in terms of the risks incurred were carried out for the values of the parameter \( \gamma = 0.1 \), \( 0.5 \) and \( 2.0 \) and for twelve different starting points, namely:

\[
\begin{align*}
  s_0 &= 101, \quad y_0 = 0, 10 \\
  s_0 &= 501, \quad y_0 = 0, \pm 10, 20 \\
  s_0 &= 1001, \quad y_0 = 0, \pm 10, \pm 20, 30 .
\end{align*}
\]

These comparisons appear in Table 2.

Examination of Table 2 leads to the conclusion that whereas procedure \( F \) is relatively inefficient (efficiencies for the various different cases range from 27 percent to 81 percent with most cases yielding efficiencies quite close to 50 percent), procedure \( W \) is surprisingly efficient (efficiencies for the various cases range from 73 percent to 100 percent with most cases yielding efficiencies quite close to 92 percent). These results must be treated with a healthy suspicion due to the problems associated with the use of this particular measure of efficiency as mentioned above. The other point to be kept in mind is the fact that only these twelve particular starting points have been considered. This constitutes a fairly small data base.
Table 2

Efficiencies of Procedures W and F for Different Starting Points and Different Values of the Parameter γ

<table>
<thead>
<tr>
<th>Procedure</th>
<th>s₀ = 101 y₀ = 0</th>
<th>R - Rₑ</th>
<th>E</th>
<th>s₀ = 101 y₀ = 10</th>
<th>R - Rₑ</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>Opt</td>
<td>.266</td>
<td>.288</td>
<td>.370</td>
<td>.098</td>
<td>.032</td>
<td>.115</td>
</tr>
<tr>
<td>W</td>
<td>.276</td>
<td>.96</td>
<td>.314</td>
<td>.92</td>
<td>.395</td>
<td>.94</td>
</tr>
<tr>
<td>F</td>
<td>.344</td>
<td>.77</td>
<td>.380</td>
<td>.75</td>
<td>.456</td>
<td>.81</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Procedure</th>
<th>s₀ = 501 y₀ = -10</th>
<th>R - Rₑ</th>
<th>E</th>
<th>s₀ = 501 y₀ = 0</th>
<th>R - Rₑ</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>Opt</td>
<td>.270</td>
<td>.224</td>
<td>.272</td>
<td>.186</td>
<td>.198</td>
<td>.242</td>
</tr>
<tr>
<td>W</td>
<td>.232</td>
<td>.91</td>
<td>.236</td>
<td>.95</td>
<td>.281</td>
<td>.97</td>
</tr>
<tr>
<td>F</td>
<td>.396</td>
<td>.53</td>
<td>.406</td>
<td>.55</td>
<td>.463</td>
<td>.59</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Procedure</th>
<th>s₀ = 501 y₀ = 10</th>
<th>R - Rₑ</th>
<th>E</th>
<th>s₀ = 501 y₀ = 20</th>
<th>R - Rₑ</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>Opt</td>
<td>.109</td>
<td>.125</td>
<td>.178</td>
<td>.058</td>
<td>.068</td>
<td>.103</td>
</tr>
<tr>
<td>W</td>
<td>.116</td>
<td>.94</td>
<td>.129</td>
<td>.97</td>
<td>.181</td>
<td>.98</td>
</tr>
<tr>
<td>F</td>
<td>.187</td>
<td>.58</td>
<td>.205</td>
<td>.61</td>
<td>.278</td>
<td>.64</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Procedure</th>
<th>s₀ = 1001 y₀ = -10</th>
<th>R - Rₑ</th>
<th>E</th>
<th>s₀ = 1001 y₀ = 0</th>
<th>R - Rₑ</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>Opt</td>
<td>.185</td>
<td>.195</td>
<td>.227</td>
<td>.182</td>
<td>.191</td>
<td>.221</td>
</tr>
<tr>
<td>W</td>
<td>.205</td>
<td>.90</td>
<td>.213</td>
<td>.92</td>
<td>.242</td>
<td>.94</td>
</tr>
<tr>
<td>F</td>
<td>.400</td>
<td>.46</td>
<td>.491</td>
<td>.46</td>
<td>.448</td>
<td>.51</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Procedure</th>
<th>s₀ = 1001 y₀ = 10</th>
<th>R - Rₑ</th>
<th>E</th>
<th>s₀ = 1001 y₀ = 10</th>
<th>R - Rₑ</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>Opt</td>
<td>.146</td>
<td>.158</td>
<td>.198</td>
<td>.124</td>
<td>.133</td>
<td>.166</td>
</tr>
<tr>
<td>W</td>
<td>.160</td>
<td>.91</td>
<td>.169</td>
<td>.93</td>
<td>.206</td>
<td>.96</td>
</tr>
<tr>
<td>F</td>
<td>.284</td>
<td>.51</td>
<td>.296</td>
<td>.53</td>
<td>.351</td>
<td>.56</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Procedure</th>
<th>s₀ = 1001 y₀ = 20</th>
<th>R - Rₑ</th>
<th>E</th>
<th>s₀ = 1001 y₀ = 30</th>
<th>R - Rₑ</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>Opt</td>
<td>.075</td>
<td>.086</td>
<td>.122</td>
<td>.047</td>
<td>.054</td>
<td>.080</td>
</tr>
<tr>
<td>W</td>
<td>.082</td>
<td>.91</td>
<td>.091</td>
<td>.95</td>
<td>.127</td>
<td>.96</td>
</tr>
<tr>
<td>F</td>
<td>.144</td>
<td>.52</td>
<td>.158</td>
<td>.54</td>
<td>.213</td>
<td>.57</td>
</tr>
</tbody>
</table>
from which to make general conclusions. However, the claim that procedure $W$ is a fairly efficient procedure seems to be strongly supported, at least in the region of the parameter space examined here. This is of importance for possible applications since this indicates that the use of the suboptimal procedure $W$, which is easier to implement than is the optimal procedure, will result in very little increase in the risk incurred.

In order to emphasize this point even more dramatically, the results obtained for the Wiener process problem have been translated into the context of two different binomial problems in Table 3 and 4. (Procedures $M(\cdot)$ appearing in these tables will be defined in the next chapter.) Table 3 translates the results obtained for the Wiener process problem with cost parameter $\gamma = 0.1$ and starting point given by $s_0 = 100$, $y_0 = 0$ while Table 4 considers the Wiener process problem with the same starting point but with cost parameter $\gamma = 2.0$. In both tables the results are translated into the context of two different corresponding binomial problems. The first binomial problem has parameters $p_0 = 0.5$, $a_0 = 1$, $b_0 = 1$, $N = 2000$ and $c = \gamma/(8,008)^{1/2}$, while the second binomial problem has parameters $p_0 = 0.5$, $a_0 = 4$, $b_0 = 4$, $N = 8000$ and $c = \gamma/(32,032)^{1/2}$. The risks incurred using the various procedures are displayed in the tables.

From these tables two points are dramatically demonstrated within the context of the clinical trials problem. The first point is that these statistical procedures are effective in avoiding failures; the second point is that the suboptimal procedures perform very well. For

97
Table 3

Risks Incurred in Two Different Clinical Trials Problems
Corresponding to the Wiener Process Problem with $\gamma = 0.1$

<table>
<thead>
<tr>
<th>Procedure</th>
<th>Wiener Process $u(0,1001)$</th>
<th>First Binomial $r(1,1)$</th>
<th>Second Binomial $r(4,4)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>CB</td>
<td>-12.609</td>
<td>717.903</td>
<td>3435.805</td>
</tr>
<tr>
<td>Opt</td>
<td>-12.463</td>
<td>721.167</td>
<td>3442.335</td>
</tr>
<tr>
<td>W</td>
<td>-12.450</td>
<td>721.474</td>
<td>3442.948</td>
</tr>
<tr>
<td>F</td>
<td>-12.326</td>
<td>724.250</td>
<td>3448.500</td>
</tr>
<tr>
<td>M(5)</td>
<td>-12.266</td>
<td>725.586</td>
<td>3451.172</td>
</tr>
<tr>
<td>M(20)</td>
<td>-11.905</td>
<td>733.669</td>
<td>3467.339</td>
</tr>
</tbody>
</table>
Table 4
Risks Incurred in Two Different Clinical Trials Problems Corresponding to the Wiener Process Problem with \( Y = 2.0 \)

<table>
<thead>
<tr>
<th>Procedure</th>
<th>First Binomial</th>
<th>Second Binomial</th>
</tr>
</thead>
<tbody>
<tr>
<td>CB</td>
<td>-12.609</td>
<td>3435.737</td>
</tr>
<tr>
<td>Opt</td>
<td>-12.411</td>
<td>722.298</td>
</tr>
<tr>
<td>W</td>
<td>-12.403</td>
<td>722.469</td>
</tr>
<tr>
<td>( V )</td>
<td>-12.259</td>
<td>725.709</td>
</tr>
<tr>
<td>( M(\frac{1}{2}) )</td>
<td>-12.116</td>
<td>728.231</td>
</tr>
<tr>
<td>( M(\frac{1}{20}) )</td>
<td>-8.075</td>
<td>812.297</td>
</tr>
</tbody>
</table>

Wiener Process Problem: \( Y = 2.0 \), \( s_0 = 1001 \), \( y_0 = 0 \)

First Binomial Problem: \( p_0 = 0.5 \), \( a_0 = 1 \), \( b_0 = 1 \), \( M = 2000 \), \( C = 0.0223 \)

Second Binomial Problem: \( p_0 = 0.5 \), \( a_0 = 4 \), \( b_0 = 4 \), \( M = 8000 \), \( C = 0.0112 \)
instance, examination of Table 3 reveals that in a clinical trial situation with \( p_0 = 0.5 \) and horizon size 2000, whereas a priori we would expect to incur 1000 failures (or to phrase it more colorfully, we would expect 1000 of our 2000 patients to die), use of the optimal procedure results in a reduction of this number to 722. Further, use of procedure W instead of the optimal procedure results in less than 1 more expected failure while use of procedure F results in less than 4 more expected failures. Similar interpretations follow for the second binomial problem of Table 3 and also for Table 4. These tables will be examined further in the next chapter.
Chapter 5

Discussion and Conclusions

The objectives of the present research were presented in Section 2.2. These objectives have been met. The introduction of the beta prior distribution for the unknown probability of success of the experimental treatment led to the determination of the optimal solution. Passing to the limit \((M \to \infty)\) led to a characterization of the optimal solution, at least in this asymptotic sense, as the solution of a FBP involving a single parameter. This same FBP was shown to be the solution of the related Wiener process problem formulated in Section 3.1.

The formal asymptotic expansions for the boundaries of the continuation region, as developed within the context of the related Wiener process problem, served to make the characterization of the optimal solution more explicit. A numerical method of solving the related Wiener process problem was presented and applied in Section 4.1. These computations yield both the boundaries of the continuation region and the Bayes risk within the continuation region. The proposal of the suboptimal procedures \(W\) and \(F\) was carried out within the context of the related Wiener process problem. Numerical methods for evaluating these suboptimal procedures were presented and applied in Section 4.3. Subsequently the performance of these suboptimal procedures was compared to that of the optimal procedure. In Section 4.4.B, procedure \(W\) was determined to be of relatively high efficiency in the range of parameters considered and in retrospect this appears to be the most significant result of this body of research.
There are a number of open problems remaining in connection with this work. The various expansions that we have obtained are simply formal expansions; we have made no attempt to establish that these formal expansions do in fact represent asymptotic expansions. Presumably this could be accomplished by the use of techniques similar to those of Breakwell and Chernoff [4] and Chernoff [7]. In a similar fashion, presumably techniques similar to those of Chernoff [8] could be used to prove that the "continuity correction" introduced in Section 4.1 is in fact appropriate.

There is also some further work which would be relevant to this body of research. For the purpose of relating the optimal solution of the Wiener process problem to the clinical trials situation with small values of $M$, it would be of interest to determine the constants $a_u(p_0)$ and $a_x(p_0)$ to be employed in equations (4.2.3) and (4.2.4) for those cases with $p_0 \neq 0.5$. Further, as indicated in Section 4.4, it would be of interest to obtain a "continuity correction" which would transform the risk of any of the various procedures (within the region of continuation) as evaluated from the $Y'(s')$ process into the risk for the same procedure within the context of the Wiener process. Also, it would be of interest to determine expressions for the expected sample sizes and probabilities of error of the optimal and suboptimal procedures developed in this body of research. Consideration of other suboptimal procedures would also be of interest. In particular, the examination of the extension of procedure $F$ to a $k$-stage sequential stopping procedure could be of great interest. From our comparisons of
procedure \( F \) and the optimal procedure we have seen that even though procedure \( F \) appears to mimic the optimal procedure in terms of its stopping rule, it is nonetheless a relatively inefficient procedure according to the measure of efficiency we have used. A conjecture is that the extension of procedure \( F \) to even a two-stage sequential procedure would greatly improve its efficiency.

As indicated in the introductory chapter, whereas the problem of comparing an experimental treatment with a standard has received limited attention in the literature, the problem of comparing two experimental treatments in the context of clinical trials has been the object of a great deal of interest. It would be of interest to examine this more general problem in a formulation similar to that considered here for our more specific problem. The formulation of this more general problem presents no difficulties and the introduction of beta prior distributions for the unknown probabilities of success of the two treatments would enable us to determine the optimal solution for a specified set of parameters by carrying out a slightly more complicated backward induction algorithm. The extension to more than two treatments would also be of interest.

The model discussed in this body of research is clearly oversimplified. Several possible objections to the model were indicated in Section 2.1.

The assumption of instantaneous response appears to be fundamental in our formulation of the problem. Although our model requires that the response of any patient to treatment be observed prior to the time
that the next patient enters the experiment, this problem becomes significant only if the response time is very long or erratic compared to the rate of influx of patients. For formal discussions of the response time problem in a more general situation see Anderson [1], Ehrenfeld [19], and Suzuki [38].

The lack of control groups is a serious difficulty in the situation where these two treatments are being compared at a variety of locations, at each of which patients and procedures differ substantially. In such a situation, the model proposed by Donner [18] may be more appropriate.

The cost structure of our particular formulation could also be criticized. The cost $c$, which represents the cost of including a patient in the experimental period relative to the cost of a failure, has been assumed to remain constant throughout the experimental period. Ethical considerations would seem to indicate that an additional cost assessed as a function of the present estimated effectiveness of the experimental treatment relative to the standard treatment should be included in the formulation.

Another serious difficulty is the specification of the horizon size. This difficulty has been discussed in some detail by Anscombe [2]. The effect of the misspecification of the horizon size could be examined in detail for our particular problem. Simply stated, over specification of the horizon size would lead to continuing for a longer (than optimal) period of time while underspecification would lead to continuing for a shorter (than optimal) period of time. The
misspecification of the horizon size translates into a misspecification of the boundaries of the region of continuation and consequently results in the use of a suboptimal procedure.

Consider an experimenter in the clinical trials situation with parameters $a_0$, $b_0$, $p_0$, $c$ and $N$. Suppose he incorrectly specifies the horizon size to be $N^*_N$. Upon transformation to the $(y,s)$ plane, this results in a misspecification of the parameter $\gamma$. We define the misspecification factor $b$ by:

$$b^2 = \frac{(a_0 + b_0 + N^*_N)/(a_0 + b_0 + N)}{(5.1)}$$

and denote the suboptimal procedure corresponding to a misspecification factor $b$ by $M(b)$. The boundaries of the continuation region corresponding to the procedure $M(b)$ and the risk incurred can be determined by our previous methods. These computations have been carried out for a few cases and the results for two seemingly typical cases are displayed in Tables 3 and 4. Although the computations are performed in the context of the Wiener process problem, the results are translated to two corresponding binomial situations for ease of interpretation. Note that Table 3 considers the case of overspecification ($b > 1$) while Table 4 considers the case of underspecification ($b < 1$).

Although the empirical evidence presented in Tables 3 and 4 is somewhat limited, these particular cases do seem to be typical. We are thus led to the following two observations: First, the effect of underspecification appears to be considerably more severe than the effect of overspecification. Second, neither of these effects appears to be as severe
as might have been expected. For example, for the first binomial problem considered in Table 3, a misspecification factor of 5 (that is, overestimating the horizon size by a factor of 25) results in an increase in the risk which is equivalent to less than 5 more expected failures (out of a total of 2000 patients) while a misspecification factor of 20 results in an increase in the risk which is equivalent to less than 13 more expected failures (out of a total of 2000 patients). These observations are somewhat reassuring since as Anscombe [2] has pointed out, the natural tendency of the experimenter would be to overestimate the horizon size.

The discussion of the misspecification of the horizon size has to this point, used the criterion of risk. However, we could also consider the following question: If each trial of the experimental treatment leads to a failure, how many successive failures would be required before the experimenter would stop the use of the experimental treatment? Since underspecification leads to stopping early, we will only consider overspecification. For any particular binomial problem this question can be answered directly from the graphs of the boundaries corresponding to the various procedures in the (z,t) plane. For the first binomial problem considered in Table 3, we determine that the procedures Opt, M(5), and M(20) stop after 7, 12, and 17 successive failures respectively while for the second binomial problem considered in Table 3 the same three procedures stop after 12, 18, and 23 successive failures respectively. Thus we see that this criterion is considerably more sensitive to overspecification than the
criterion of risk. However, it must be kept in mind that we are considering quite serious overspecifications here and when this is taken into account these effects also appear to be relatively insignificant.

The above discussion indicates that while it may indeed be a difficult problem for the experimenter to accurately specify his horizon size, he may still proceed confidently since the effects of misspecification (particularly overspecification) appear to be relatively minor.

It should be noted in passing that procedure \( F \) always prescribes early stopping and consequently performs well in comparison to the optimal procedure with respect to this criterion. From the form of the boundaries of procedure \( W \), we expect that procedure \( W \) will perform poorly in comparison to the optimal procedure with respect to this criterion. Indeed, asymptotic considerations suggest that whereas the number of successive failures required for stopping varies as \( \log N \) for the optimal procedure, this number varies as \( N^{1/2} \) for procedure \( W \). Thus, particularly for very large values of the horizon size, procedure \( W \) will perform poorly with respect to this criterion. Due to the simple nature of the boundaries, it is easy to calculate exactly how many successive failures must be incurred in any given binomial problem before procedure \( W \) calls for stopping. In particular, for the Wiener process problem of Table 3, we obtain the lower boundary \( z_{\xi}(t) = -0.196 \) (from Table 1) for procedure \( W \). From this boundary, it is easy to determine that for the first binomial problem of Table 3 procedure \( W \) calls for stopping after 9 successive failures while
for the second binomial problem of Table 3 procedure \( W \) calls for stopping after 18 successive failures. These numbers are to be compared with the values 7 and 12 respectively as determined above for the optimal procedure.

The only question that remains to be addressed is: What insights has this body of research provided? The first insight is the result that the Wiener process provides an excellent approximation for our clinical trials situation. This was demonstrated in Section 4.2 in which the optimal solution to the related Wiener process problem was compared to the optimal solution of the clinical trials problem for a few specific cases. This insight is not very surprising and would almost surely have been conjectured a priori. What is perhaps surprising concerning this insight is that the approximation appears to work very well even for fairly moderate values of the parameter \( M \). Further, the determination of the appropriate "continuity corrections" (4.2.3) and (4.2.4) should further enhance the adequacy of the approximations.

The second insight which is important as far as potential applications of the present work is concerned is the fact that all of the procedures that have been proposed appear to be very effective in reducing total expected failures. Although this fact is somewhat hidden in Table 2, in which efficiencies are examined within the context of the Wiener process problem, it is dramatically demonstrated in Tables 3 and 4, in which the results obtained within the context of the Wiener process problem are translated to the clinical trials situation for a few particular cases. These tables also indicate that the cost of ignorance
of the probability of success of the experimental treatment is equivalent to the cost of a few additional failures since the risk associated with the use of the optimal procedure is only slightly greater than the crystal ball risk. The final insight is the result that, at least for the cases considered here, the Wald-type procedure performs exceptionally well, except possibly if there are many successive failures at the beginning of sampling. This is clearly demonstrated in Table 2 and even more dramatically in Tables 3 and 4. This result is particularly surprising in light of the great difference in the shape of the boundaries of procedure W and the optimal procedure. This conclusion must be taken with a grain of salt, however, since we have really considered only moderate values of the horizon size in this research. If we recall the fact that the boundaries of the procedure W are horizontal lines in the (z,t) plane, we are led to suspect that at least for very large values of the horizon size the procedure may behave poorly. This point has been previously made by Chernoff [9].
REFERENCES


<table>
<thead>
<tr>
<th>1. REPORT NUMBER</th>
<th>2. GOVT ACCESSION NO.</th>
<th>3. RECIPIENT'S CATALOG NUMBER</th>
</tr>
</thead>
<tbody>
<tr>
<td>221</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>4. TITLE (and Subtitle)</th>
<th>5. TYPE OF REPORT &amp; PERIOD COVERED</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sequential Medical Trials for Comparing an Experimental with a Standard Treatment</td>
<td>Technical Report</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>7. AUTHOR(s)</th>
<th>8. CONTRACT OR GRANT NUMBER(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Albert John Petkau</td>
<td>N00014-67-A-0112-0085</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>9. PERFORMING ORGANIZATION NAME AND ADDRESS</th>
<th>10. PROGRAM ELEMENT, PROJECT, TASK AREA &amp; WORK UNIT NUMBERS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Department of Statistics</td>
<td>(NR-042-267)</td>
</tr>
<tr>
<td>Stanford University</td>
<td></td>
</tr>
<tr>
<td>Stanford, Calif. 94305</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>11. CONTROLLING OFFICE NAME AND ADDRESS</th>
<th>12. REPORT DATE</th>
<th>13. NUMBER OF PAGES</th>
</tr>
</thead>
<tbody>
<tr>
<td>Office of Naval Research</td>
<td>July 30, 1975</td>
<td>115</td>
</tr>
<tr>
<td>Statistics &amp; Probability Program Code 436</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arlington, Virginia 22217</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>14. MONITORING AGENCY NAME &amp; ADDRESS (if different from Controlling Office)</th>
<th>15. SECURITY CLASS. (of this report)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Unclassified</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>16. DISTRIBUTION STATEMENT (of this Report)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Approved for Public Release: Distribution Unlimited.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)</th>
<th>18. SUPPLEMENTARY NOTES</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Also issued as Technical Report No. 1, Department of Mathematics, Massachusetts Institute of Technology, under Contract Number N00014-75-C-0555 (NR-042-331).</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>19. KEY WORDS (Continue on reverse side if necessary and identify by block number)</th>
</tr>
</thead>
<tbody>
<tr>
<td>clinical trials, backward induction, dynamic programming, free boundary problem, optimal stopping, Wiener process</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>20. ABSTRACT (Continue on reverse side if necessary and identify by block number)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(see reverse side)</td>
</tr>
</tbody>
</table>
20. This research is concerned with a cost function approach to the design of sequential medical trials for comparing an experimental with a standard treatment. It is assumed that all available patients must be treated with one of the two treatments. The standard and experimental treatments are characterized by known and unknown probabilities of success respectively. Two costs are considered; the cost incurred every time either of the treatments fails and a special cost per patient treated in the experimental portion of the plan.

The Bayes sequential procedure is determined and it is indicated that as the number of available patients increases indefinitely and with the appropriate normalization of the cost parameters, the problem reduces to a free boundary problem. This same free boundary problem is shown to arise as the solution of a continuous time optimal stopping problem for the Wiener process. The asymptotic behavior of the solution of the Wiener process problem is examined. The results for the Wiener process problem then furnish approximations for our problem. These approximations are compared with the optimal procedure. Several suboptimal procedures are proposed and their behavior is evaluated and compared with that of the optimal procedure.