The Korean Communications
in Statistics Vol. 6, No. 1, 1999
pp. 207-220

Accuracy of Brownian Motion Approximation in Group Sequential Methods[1)]

Euy Hoon Suh[2)]

Abstract

In this paper, some of the issues about a group sequential method are considered in the Bayesian context. The continuous time optimal stopping boundary can be used to approximate the optimal stopping boundary for group sequential designs. The exact stopping boundary for group sequential design is obtained by using the backward induction method and is compared with the continuous optimal stopping boundary and the corrected continuous stopping boundary.

1. Introduction

In the long-term clinical trials, where the patients are entering sequentially, the strict application of fixed sample size designs is unjustified on ethical grounds. On the other hand fully sequential designs may be impractical due to need for continuous assessment of accumulating data. The planned use of the group sequential designs has been advocated as the convenient approach to the monitoring of clinical trials.

In the literature there are many ad-hock group sequential designs, to name a few, Pocock (1977), O'Brien-Fleming (1979), and Lan-DeMets (1983). For a good review one can see in Whitehead(1992). But relative merits of the different types of group sequential designs are seldom seriously investigated. In most of these procedures either one uses Brownian Motion approximation to evaluate error probabilities or uses some numerical methods. The numerical methods are usually very time consuming. In this presentation we would like to resolve some of the issues in the Bayesian context. In this manuscript we will focus on the following issues:

(a) How to select the number of groups in a group sequential method, without loosing much "information".

(b) In the decision theoretic framework, how a continuous time stopping problem with

1) 이 논문은 1996년도 경상대학교허영중연구장학재단학술연구조성비에 의하여 연구되었음.

2) Professor, Department of Statistics, Gyeongsang National University, Chinju, 660-701, Korea.

Wiener process can approximate a discrete time group sequential procedure.

(c) How good is continuous time "optimal" stopping boundary as an approximation to "optimal" group sequential stopping boundary.

2. Statement of the Problem

In Anscombe (1963) introduced a decision theoretic approach to clinical trial in the fully sequential context. He assumed that the patients are treated in pairs during the experimental phase of the study, where the difference in response (treatment 1-treatment 2) is distributed as $N(\mu, \sigma^2)$ r.v. with μ unknown. He also assumed that a patient horizon N, the total number of patients ever to receive either of the treatments, is known. The loss function is defined as $|\mu|$ times the number of patients receiving the poorer treatment. After experimenting on n pairs of patients, let S_n denote the sum of the response differences. Anscombe assumed that the remaining $(N-2n)$ patients would be treated according to the sign of S_n, that is, if $S_n>0$ then the remaining patients will receive treatment 1. If we assume that μ has a prior distribution $\lambda(\mu)$, then the posterior expected loss can be written

$$E(L) = nE[|\mu|] + (N-2n)E[\max(0, -\mu\, sgn(S_n))], \qquad (2.1)$$

where the expectation is taken with respect to the posterior distribution of μ after observing S_n.

In this manuscript, we will use Anscombe's decision theoretic framework, in the group sequential context.

3. Fully Sequential Discrete and Continuous Time Problem

Chernoff-Petkau (1981) considered the Bayes sequential problem for the anscombe loss function (2.1) and the prior distribution $\lambda(\mu)$ to be a normal distribution with mean μ_0 and variance σ_0^2. After observing the differences $X_1, \cdots, X_n$ in effect of treatment 1 minus treatment 2, the posterior distribution of μ becomes $N(Y_n^*, s_n^*)$, where

$$Y_n^* = \frac{\left[\sigma_0^{-2}\mu_0 + \sigma^{-2}\sum_{i=1}^{n} X_i\right]}{\sigma_0^{-2} + n\sigma^{-2}}, \quad s_n^* = \left(\sigma_0^{-2} + n\sigma^{-2}\right)^{-1}. \tag{3.1}$$

From Chernoff (1972), we have the following lemma.

Lemma: The distribution of $(Y_n^* - Y_m^* | Y_m^*)$ is a $N(0, s_m^* - s_n^*)$, and $Y_n^* - Y_m^*$ is independent of Y_m^*, for $n \geq m \geq 0$.

Therefore, Y_n^* behave like a Gaussian process of independent increment starting from $Y_0^* = \mu_0$ and $s_0^* = \sigma_0^2$. Since the preferred choice of treatment for the remaining $N-2n$ patients is indicated by the sign of Y_n^*, the expected loss or the posterior risk associated with stopping after treating n pairs of patients is

$$nE(|\mu|) + (N-2n)\, E[\max\{0, -sgn(Y_n^*)\mu\}],$$

where E represents the expectation with respect to the posterior distribution of μ given Y_n^*.

By simple calculations, the posterior risk can be expressed as

$$N(s_n^*)^{\frac{1}{2}} \psi\left[(Y_n^*)(s_n^*)^{-\frac{1}{2}}\right] - \frac{1}{2}(N-2n)\,|Y_n^*|,$$

where

$$\psi(u) = \phi(u) + u\left\{\Phi(u) - \frac{1}{2}\right\},$$

and $\phi(u)$, $\Phi(u)$ are the standard normal density and cumulative respectively. Using (3.1), the posterior risk can be written as $d_1(Y_n^*, s_n^*)$, where

$$d_1(y^* \cdot s^*) = N s^{*\frac{1}{2}} \psi\left[y^* s^{*-\frac{1}{2}}\right] - \sigma^2\left[s_0^{-1} - s^{*-1}\right]|y^*|, \quad s_0^{-1} = \sigma_0^{-2} + \frac{1}{2}N\sigma^{-2}. \tag{3.2}$$

The problem of selecting the best sequential procedure for terminating the experimental phase is equivalent to the optimal stopping problem where the Gaussian process Y_n^* is

observed and one selects the stopping time τ to minimize the expected risk $E\{d_1(Y_\tau^*, s_\tau^*)\}$., where the expectation is taken over the distribution of the stopping time τ.

A natural approximation to the above problem is to replace the discrete sequence of partial sums $\sum_{i=1}^{n} X_i$ by the continuous time Wiener process $X(t^*)$ with drift μ and variance σ^2 per unit in the t^* scale $(0 < t^* < N/2)$. The posterior distribution of μ, given $X(t')$ for $0 \le t' \le t^*$, is $N(Y^*, s^*)$, where

$$Y^* = Y^*(s^*) = \frac{\{\sigma_0^{-2}\mu_0 + \sigma^{-2}X(t^*)\}}{\sigma_0^{-2} + t^*\sigma^{-2}}, \quad s^* = \{\sigma_0^{-2} + t^*\sigma^{-2}\}^{-1}. \tag{3.3}$$

From Chernoff (1972), we know that $Y^*(s^*)$ is a Wiener process with drift 0 and variance 1 per unit time in the $-s^*$ scale, and originates at the initial point (y_0^*, s_0^*), where $s_0^* = \sigma_0^2$, $y_0^* = Y^*(s_0^*) = \mu_0$. As t^* increases from 0 to $N/2$, s^* decreases from s_0^* to $\tilde{s}_0 = \left[\sigma_0^{-2} + \frac{N}{2}\sigma^{-2}\right]^{-1}$. When $\sigma_0^2 \to \infty$, that corresponds to negligible prior information.

By using the transformation, $Y(s) = aY^*(s^*)$ and $s = a^2 s^*$, one can convert the $Y^*(s)$ to the $Y(s)$ process. This is also a Gaussian process of independent increments with $E[dY(s)] = 0$ and $Var[dY(s)] = -a^2 ds^* = -ds$.

By using (3.2), we get

$$d(y, s) = Na^{-1}s^{\frac{1}{2}}\psi\left[ys^{-\frac{1}{2}}\right] - \sigma^2[\tilde{s}_0^{-1} - a^2 s^{-1}]a^{-1}|y|.$$

If we choose $a = \left(\sigma_0^{-2} + \frac{1}{2}N\sigma^{-2}\right)^{\frac{1}{2}}$, then

$$d(y, s) = Na^{-1}s^{\frac{1}{2}}\psi\left[ys^{-\frac{1}{2}}\right] - \sigma^2 a(1 - s^{-1})|y|,$$

where $1 < s < s_0 = \sigma_0^2\left(\sigma_0^{-2} + \frac{1}{2}N\sigma^{-2}\right)^{-1} = t_0^{-1}$. Let

$$d_1(y, s) = -(1 - s^{-1})|y|.$$

Since the first term of $d(y, s)$ is a martingale, and $(\sigma^2 a)$ is a constant, the optimal

stopping boundary for the cost functions d and d_1 are same. Hence $d_1(y,s)$ is the normalized version of the original problem. For the normalized version of the Anscombe's problem Chernoff-Petkau (1981) computed the continuous time optimal stopping boundary. These are tabulated in Table 1 and plotted in Figure 1. We will investigate how the continuous time optimal stopping boundary can be used to approximate the optimal stopping boundary for group sequential designs.

4. Group Sequential Designs

The group sequential design can be described as follows: Suppose the total number of pairs of patients $0.5N$ is split into K groups of m pairs of patients, $Km = 0.5N$. Then the stopping is allowed only at the values $n = im$, for $i = 1, \cdots, K$, and the stopping is enforced after the Kth group when $n = Km = 0.5N$. So the discrete time group sequential designs can be interpreted as the continuous time problem where the stopping is enforced only at the points $n = im$, $i = 1, \cdots, K$. In the Bayesian framework the group sequential problem can be described as the following problem: We observe

$$\overline{X_{im}} = \frac{1}{m}\sum_{j=1}^{m} X_{ij}, \quad i = 1, 2, \cdots, K$$

which are independently and identically distributed normal random variables with mean μ and variance $\tilde{\sigma}^2 = \frac{\sigma^2}{m}$. Where μ has a prior distribution that is normal with mean μ_0 and variance σ_0^2.

The posterior risk associated with stopping at ith group is given by

$$\begin{aligned} & imE(|\mu|) + (2Km - 2im)E[\max\{0, -sgn(Y_n^*)\mu\}] \\ = \; & m[iE(|\mu|) + 2(K-i)E[\max\{0, sgn(Y_n^*)\mu\}]], \end{aligned}$$

where the expectation E is with respect to the posterior distribution of μ. Thus the group sequential problem is same as the fully sequential problem, if we replace n by i, $\frac{N}{2}$ by K and σ^2 by $\tilde{\sigma}^2 = \frac{\sigma^2}{m}$. Hence we can approximate optimal Bayes boundary and Bayes risk for the discrete time problem from the corresponding continuous time problem. For the continuous

time problem one can write the posterior cost of stopping at (y^*, s^*) as

$$d^*(y^*, s^*) = m\left[2Ks^{*\frac{1}{2}}\phi\left(y^* s^{*\frac{1}{2}}\right) - (K-i)|y^*|\right] .$$

By using the transformation $Y(s) = aY^*(s^*)$ and $s = a^2 s^*$ and by choosing $a^2 = (\sigma_0^{-2} + \tilde{\sigma}^{-2}K)$,

$$d^*(y, s) = m\left[2Ka^{-1}s^{\frac{1}{2}}\phi\left(ys^{-\frac{1}{2}}\right) - \tilde{\sigma}^2 a(1-s^{-1})|y|\right], 1 \le s \le s_0 = \sigma_0^2(\sigma_0^{-2} + K\tilde{\sigma}^{-2}). \quad (4.1)$$

Let $d_1^*(y, s) = m\tilde{\sigma}^2 a(1-s^{-1})|y|$. Since difference between $d^*(y, s)$ and $d_1^*(y, s)$ is a martingale, the optimal stopping boundary for both cost functions are same. Comparing $d_1(y, s)$ and $d_1^*(y, s)$ one can see the optimal group sequential stopping boundary can be related with optimal continuous time boundary.

5. Corrections to the Continuous Time Boundary to Get the Group Sequential Boundary.

From (3.3), if we let $\sigma_0^2 \to \infty$, we get

$$Y^* = Y^*(s^*) = \frac{\{X(t^*)\}}{t^*}, \quad s^* = (t^*\sigma^{-2})^{-1}. \quad (5.1)$$

Let $$W = \frac{\sum_{i=1}^{i}\sum_{j=1}^{m} X_{ij}}{\sqrt{im}} \ge b_i ,$$

where b_i is the standardized continuous time optimal stopping boundary evaluated at $t = i/K$. Using (4.1) and (5.1), one can relate the continuous time boundary with the group sequential boundary in the following way:

$$Y(s) = \frac{X(t)}{t} = \frac{1}{i}\sum_{i=1}^{i} x_i = \frac{1}{\sqrt{im}} W, \; i = 1, \cdots, K, \; s = \frac{1}{im} \text{ or } \frac{Y(s)}{\sqrt{s}} \equiv W.$$

If we use Chernoff (1965) boundary correction to the above problem, when $t = i/K$, $\tilde{Y}(s)$

the boundary, and $s=(2/N)(1/t)$, we get

$$\tilde{Y}_{dis}(s)\equiv \tilde{Y}_{dis}\left(\frac{2}{N}\cdot\frac{1}{t}\right)=\tilde{Y}_{cont}\left(\frac{2}{N}\cdot\frac{1}{t}\right)-0.5826\sqrt{\Delta_t}\,,$$

where $\Delta_t = s_i - s_{i+1} = \dfrac{1}{mi(i+1)}$. Hence

$$\tilde{Y}_{dis}(s) = \tilde{Y}_{cont}(s)-0.5826\sqrt{\frac{1}{mi(i+1)}} \quad \text{or}$$

$$\frac{\tilde{Y}_{dis}(s)}{\sqrt{s}} = b_i - 0.5826\sqrt{\frac{1}{i+1}}\,,\quad i=1,2,\cdots,K\,.$$

The corrected continuous time stopping boundary is tabulated in Table 2 and plotted in Figure 2.

6. Optimal Group Sequential Boundary by Using Backward Induction.

In the normalized version of the problem, we observe a standard Wiener process $Y(s)$ in s scale, $1\leq s\leq\infty$. If we stop at $Y(s)=y$, then our stopping cost is

$$d_2(y,s) = -(1-s^{-1})|y|.$$

We can stop only at $s_i=(1/im)$, $i=1,2,\cdots,K$.

Let $\rho(y,s)$ be the risk corresponding to the optimal stopping rule starting at (y,s). Then

$$\rho(y,1) = d_2(y,1).$$

We can start the backward induction as follows:

$$\rho(y,s_i) = \min\left[d_2(y,s_i), E(\rho(y+z\sqrt{\Delta_t}, s_{i+1}))\right], \tag{6.1}$$

where z is standard normal random variable and $\Delta_t = s_i - s_{i+1}$. If

$$\rho(y,s_i) \geq d_2(y,s_i),$$

then (y, s_i) is a stopping point, otherwise it is a continuation point. By linear extrapolation one can get the exact stopping boundary. For $\sigma=1$, we use (6.1) to compute $\rho(y, s_i)$. By using (4.1), we computed the Bayes risk for our original problem with the stopping cost $d^*(y, s)$. For $N=1000$, the boundaries using the backward induction method for $K=5$, 10, 20, 50 are in Table 2 and are plotted in Figure 3. The difference between the continuous time optimal stopping boundary and the boundary using the backward induction method is plotted in Figure 4. The difference between the corrected continuous time stopping boundary and the boundary using the backward induction method is plotted in Figure 5. These results show that the corrected continuous time stopping boundary can be a good approximation to the optimal group sequential stopping boundary. The Bayes risks $\rho_K(0, t)$ are tabulated in Table 3 and are plotted in Figure 6, for $K=5$, 10, 20, 50, 100.

7. A Measure of Efficiency in the Bayesian Context.

For two procedure P_1 and P_2, we define the Bayes Efficiency of P_1 with respect to P_2 is defined as follows:

$$BE(P_1, P_2|(y, s)) = \frac{\rho_{P_2}(y, s)}{\rho_{P_1}(y, s)} ,$$

where $\rho_P(y, s)$ is the posterior risk for procedure P at (y, s). We define the percentage loss of information denoted by PLI as

$$PLI = (1 - BE) \times 100\% .$$

For Bayes group sequential procedures, we will use PLI to measure the loss of information due to grouping. For computation of BE, we compare $\rho_K(0, t)$, $K=5$, 10, 20, 50 with $\rho_{100}(0, t)$ as the base. The Bayes Efficiency and the percent loss of information, for $K=5$, 10, 20, 50 are plotted in Figures 7 and 8, which suggest that there is a significant loss of information due to grouping from $K=5$ to $K=10$.

Table 1. Continuous Time Optimal Bayes Sequential Boundary

t	$b(t)$	t	$b(t)$
0.000001	4.747	0.16	1.234
0.000002	4.606	0.18	1.183
0.000005	4.412	0.20	1.136
0.00001	4.261	0.25	1.033
0.00002	4.102	0.3	0.947
0.00005	3.884	0.35	0.872
0.0001	3.711	0.4	0.804
0.0002	3.530	0.5	0.684
0.0005	3.279	0.6	0.577
0.001	3.077	0.7	0.474
0.002	2.865	0.8	0.370
0.005	2.566	0.85	0.314
0.01	2.326	0.9	0.251
0.02	2.074	0.95	0.174
0.04	1.808	0.97	0.134
0.06	1.646	0.99	0.077
0.08	1.529	0.995	0.054
0.1	1.437	0.999	0.024
0.12	1.359	0.9995	0.017
0.14	1.293	1.000	0.0

Table 2. The Backward Induction and Corrected Continuous Time Boundary

	t	B.I BND	Corrected	Cont-B.I	Corr-B.I
$K=5$	1.000	0.00000	0.00000	0.00000	0.00000
	0.800	0.00400	0.10945	0.36600	0.10545
	0.600	0.22170	0.28570	0.35530	0.06400
	0.400	0.38467	0.46764	0.41933	0.08297
	0.200	0.57400	0.72404	0.56200	0.15004
$K=10$	1.000	0.00000	0.00000	0.00000	0.00000
	0.900	0.00600	0.06677	0.24500	0.06077
	0.800	0.14708	0.17580	0.22292	0.02872
	0.700	0.24341	0.26802	0.23059	0.02461
	0.600	0.32823	0.35680	0.24877	0.02857
	0.500	0.41591	0.44615	0.26809	0.03025
	0.400	0.50800	0.54345	0.29600	0.03545
	0.300	0.60968	0.65570	0.33732	0.04602
	0.200	0.72973	0.79964	0.40627	0.06990
	0.100	0.88600	1.02504	0.55100	0.13904
$K=20$	1.000	0.00000	0.00000	0.00000	0.00000
	0.900	0.10182	0.11734	0.14918	0.01552
	0.800	0.21600	0.22870	0.15400	0.01270
	0.700	0.31430	0.32357	0.15970	0.00927
	0.600	0.40876	0.41542	0.16824	0.00665
	0.500	0.49964	0.50834	0.18436	0.00870
	0.400	0.59397	0.60980	0.21003	0.01583
	0.300	0.71035	0.72680	0.23665	0.01645
	0.200	0.84800	0.87545	0.28800	0.02745
	0.100	1.03520	1.10064	0.40180	0.06543
$K=50$	1.000	0.00000	0.00000	0.00000	0.00000
	0.900	0.16100	0.16510	0.09000	0.00410
	0.800	0.27828	0.27901	0.09172	0.00073
	0.700	0.37863	0.37690	0.09537	−0.00173
	0.600	0.47104	0.47236	0.10596	0.00132
	0.500	0.57000	0.56974	0.11400	−0.00026
	0.400	0.67976	0.67687	0.12424	−0.00290
	0.300	0.79783	0.80135	0.14917	0.00352
	0.200	0.95501	0.96034	0.18099	0.00533
	0.100	1.18064	1.19915	0.25636	0.01851
	0.080	1.24400	1.26845	0.28500	0.02445

Table 3. Bayes Risk ($N=1000$ and $\sigma=1$)

	t	$\rho_K(0,t)$	$\rho_{100}(0,t)$
$K=5$	1.000	17.84124	17.84124
	0.800	19.94413	19.00924
	0.600	20.72656	19.71729
	0.400	21.41810	20.00779
	0.200	22.00463	19.50591
$K=10$	1.000	17.84124	17.84124
	0.900	18.80417	18.50855
	0.800	19.28009	19.00924
	0.700	19.71605	19.40882
	0.600	20.07546	19.71729
	0.500	20.35382	19.92460
	0.400	20.53382	20.00779
	0.300	20.57943	19.91149
	0.200	20.41521	19.50591
	0.100	19.89269	13.37302
$K=20$	1.000	17.84124	17.84124
	0.900	18.58906	18.50855
	0.800	19.10467	19.00924
	0.700	19.52879	19.40882
	0.600	19.86624	19.71729
	0.500	20.10978	19.92460
	0.400	20.23802	20.00779
	0.300	20.20402	19.91149
	0.200	19.89630	19.50591
	0.100	18.96817	13.37302
$K=50$	1.000	17.84124	17.84124
	0.900	18.52453	18.50855
	0.800	19.03427	19.00924
	0.700	19.44190	19.40882
	0.600	19.75769	19.71729
	0.500	19.97500	19.92460
	0.400	20.07008	20.00779
	0.300	19.99012	19.91149
	0.200	19.60855	19.50591
	0.100	18.51805	18.36337

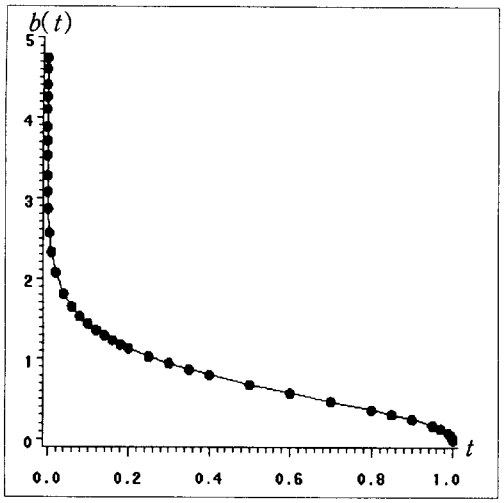

Figure 1. Continuous Time Standardized Continuous Boundary

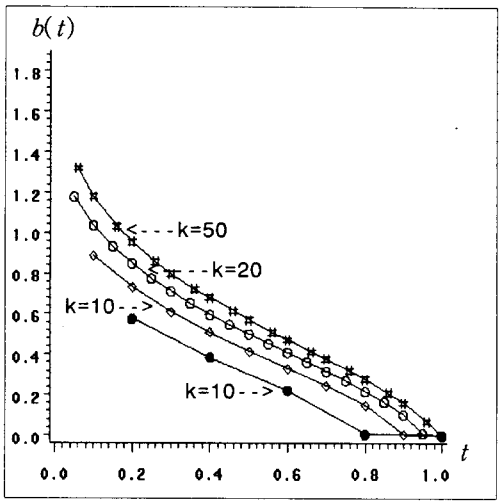

Figure 2. Standardized Corrected Time Boundary

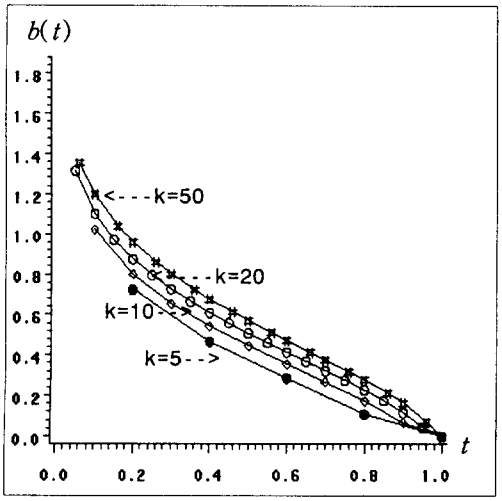

Figure 3. Standard Stopping Boundary Using Backward Induction

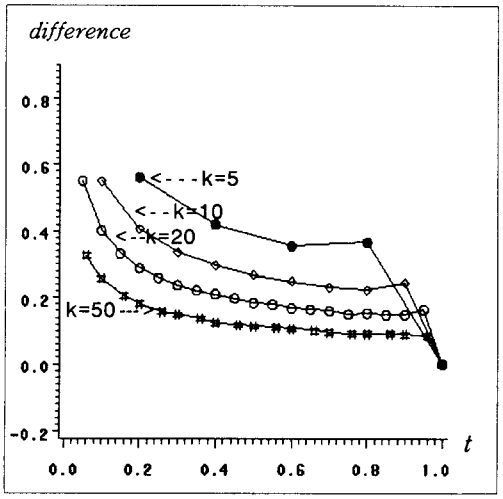

Figure 4. Difference Between Continuous Time Boundary and Backward Induction Boundary

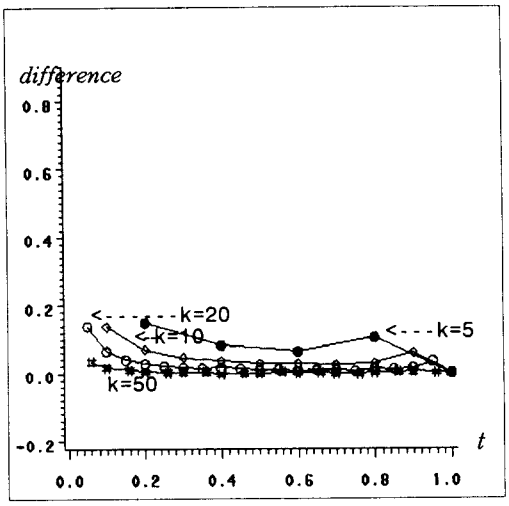

Figure 5. Difference Between Corrected Boundary and Backward Induction Boundary

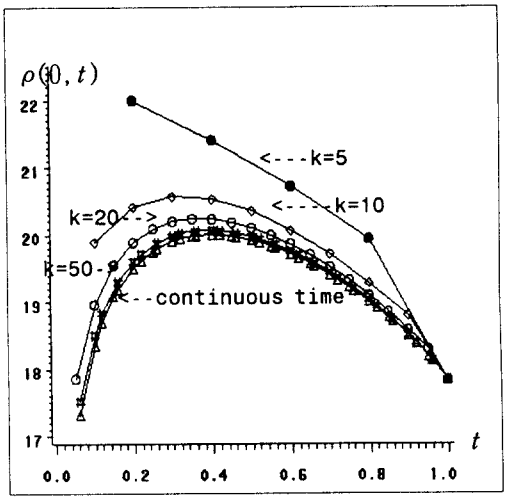

Figure 6. Bayes Risk $\rho(0, t)$ ($N=1000$ and $\sigma=1$)

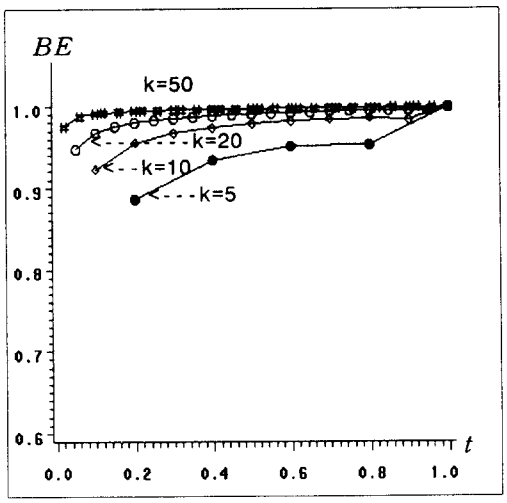

Figure 7. Bayes Efficiency

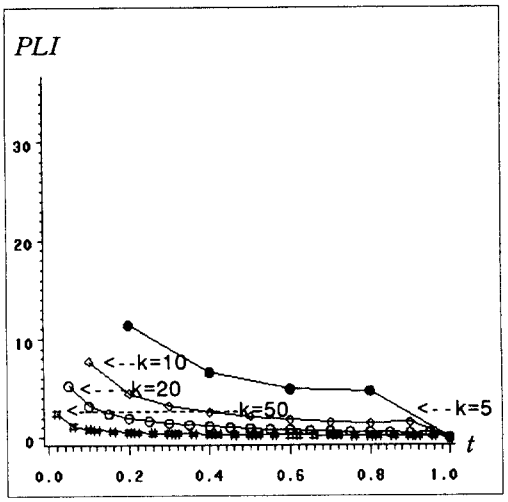

Figure 8. Percent Loss of Information

REFERENCES

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

[2] Chernoff, H. (1965). Sequential tests for the mean of a normal distribution IV (discrete case), *Ann. Math.Statist.*, 36, 55-68.

[3] Chernoff, H. (1972). *Sequential analysis and optimal design*. SIAM, J.W.Arrowsmith, Bristol.

[4] Chernoff, H. and Petkau, A. J. (1981). Sequential medical trials involving Paired data, *Biometrika*, 68, 119-32.

[5] Chernoff, H. and Petkau, A. J. (1986). Numerical methods for bayes sequential decision problems, *SIAM Journal on Scientific and Statical Computing*, 7, 46-59.

[6] Eales, J. D. and Jennison, C. (1995). Optimal two-sided group sequential tests, *Sequential Analysis*, 14(4), 273-286.

[7] Lan, K. K. G. and DeMets, D. (1983). Discrete sequential boundaries for clinical trials. *Biometrics*, 70, 659-663.

[8] O'Brien, P.C. and Fleming, T. R. (1979). A multiple testing procedure for clinical trials. *Biometrics*, 35, 549-556.

[9] Pocock, S. J. (1977). Group sequential methods in the design and analysis of clinical trials. *Biometrika*, 64, 191-199.

[10] Pocock, S. J. (1979). Allocation of Patients to treatment in clinical trials. *Biometrics*, 35, 183-197.

[11] Whitehead, J. (1992). *The Design and Analysis of Sequential Clinical Trials*, Ellis Harwood, Chichester.