Biometrika (1992), **79**, 2, *pp.* 347–53
Printed in Great Britain

Estimation after sequential testing: A simple approach for a truncated sequential probability ratio test

BY MICHAEL WOODROOFE
Statistics Department, The University of Michigan, Ann Arbor, Michigan 48109, U.S.A.

SUMMARY

An approximate pivot is constructed for the problem of estimating a normal mean θ following a truncated sequential probability ratio test and shown to provide a useful method for constructing confidence bounds and intervals. Letting t denote the sample size and S_t the sum of observations, the approximate pivot is constructed by standardizing $S_t^* = t^{-\frac{1}{2}}(S_t - t\theta)$ the mean and variance of which are no longer 0 and 1, due to the optional stopping. The truncation of the sequential probability ratio test is done in a nonstandard way in order to smooth the boundary.

Some key words: Asymptotic normality; Confidence levels; Error probabilities; Expected sample size; Means and variances; Simulation.

1. INTRODUCTION

Sequential tests have long been advocated as a means of reducing the ethical problems inherent in randomized clinical trials on human patients. There are now several good sequential tests which are finding increasingly widespread applications. See, for example, the books by Armitage (1975), Siegmund (1985) and Whitehead (1983) for complementary accounts of these developments.

Of course, the use of a sequential test complicates the problem of estimating parameters, after the test has been concluded. In response to this problem, Armitage (1957) proposed a confidence procedure which explicitly incorporates the stopping rule into the calculation of coverage probabilities for binary data; and the proposal was developed in some detail by Siegmund (1978) for normal data. In effect, Armitage and Siegmund order the points on the stopping boundary in a natural way; they then use this ordering to construct a family of tests of hypotheses of the form $\theta \leq \theta_0$ for arbitrary θ_0; and then they invert this family of tests to form confidence bounds. See also Siegmund (1985, §§ 3.4, 4.5). This approach has attracted substantial interest recently (Bather, 1988; Facey & Whitehead, 1990; Kim, 1987). On the other hand, it is moderately complicated.

The purpose of the present paper is to present an alternative confidence procedure, which starts with an approximately pivotal quantity and then proceeds in natural ways. The alternative seems both conceptually and technically simpler.

The alternative is developed in the context of a nonstandard truncated sequential probabability ratio test about a normal mean, although the method is fairly general. Truncated sequential probability ratio tests are described in § 2, and the alternative approach in § 3. In § 4 the results of a simulation study of the alternative procedure are reported. Section 5 contains some brief remarks on possible modifications of the alternative procedure and its domain of applicability; § 6 contains a rough derivation of an approximation to the error probabilities.

2. TRUNCATED SEQUENTIAL PROBABILITY RATIO TESTS

Let $X_1, X_2, \ldots$ denote independent random variables which are normally distributed with an unknown mean θ and unit variance, under a probability measure pr_θ, where $-\infty<\theta<\infty$ is unknown; and consider testing H_0: $\theta\leqslant 0$ versus H_1: $\theta>0$. Let $S_n = X_1+\ldots+X_n$ for $n=1,2,\ldots$.

Then the simplest formulation of a symmetric, truncated sequential probability ratio test of H_0 versus H_1 depends on the two design parameters $0<c<\infty$ and an integer $N\geqslant 1$. Sampling is continued until either $|S_n|>c$ or $n\geqslant N$, and H_0 is rejected if and only if $S_n>0$ when the test is terminated.

The rationale for the method described here is clearer for tests with smooth boundaries than for those with corners. So it is convenient to modify the truncation. Let $m=N/2$, $\delta=2c/N$ and

$$t=\inf[n\geqslant 1\colon n\leqslant m \text{ and } |S_n|>c \text{ or } n>m \text{ and } |S_n|>\delta\{n(N-n)\}^{\frac{1}{2}}].$$

The test requires a total of t observations, and H_0 is rejected if and only if $S_t>0$. Thus, the continuation region has horizontal boundaries up to $m=N/2$ and then parabolic boundaries meeting at N. Observe that

$$t=\inf\{n\geqslant 1\colon ng(S_n/n)>c\}, \tag{1}$$

where

$$g(x)=\begin{cases}\frac{1}{2}(\delta^2+x^2)/\delta & \text{if } |x|\leqslant\delta,\\ |x| & \text{if } |x|\geqslant\delta.\end{cases}$$

Observe further that g is symmetric, bounded away from zero, and continuously differentiable with derivative $g'(x)=x/\delta$ for $0\leqslant x\leqslant\delta$, and $g'(x)=1$ for $\delta<x<\infty$. In fact, g' is absolutely continuous and $g''(x)=1/\delta$ and 0 for $|x|<\delta$ and $|x|>\delta$.

The expected sample size of the test may be approximated from nonlinear renewal theory following Hagwood & Woodroofe (1982): as $c\to\infty$ and $N\to\infty$ in such a manner that $\delta=2c/N$ remains fixed,

$$E_\theta(t)=\frac{1}{g(\theta)}\{c+\rho(\theta)-\tfrac{1}{2}g''(\theta)\}+o(1) \tag{2}$$

for all $\theta\notin\{-\delta,0,\delta\}$, where $\rho(\theta)$ denotes the limiting expected excess over the boundary for the perturbed random walk $Z_n=ng(S_n/n)$, $n=1,2,\ldots$. Of course, (2) fails at $\theta=\pm\delta$, since g'' is discontinuous there.

Let $\nu=g(\theta)/|g'(\theta)|$, $(\theta\neq 0)$. Then, since $Z_n=ng(\theta)+g'(\theta)(S_n-n\theta)$ plus terms which are slowly changing, it follows from the Nonlinear Renewal Theorem of Lai & Siegmund (1977) that $\rho(\theta)=|g'(\theta)|\rho_\nu$, where ρ_ν denotes the limiting expected excess over the boundary for a normal random walk with drift ν and unit variance. Following Siegmund (1985, § 3.5), ρ_ν is approximated by ρ_0 and ρ_0 by 0·583, below.

Let $c'=c+\rho_0$ and $\eta_\theta=N\theta^2/(\delta^2+\theta^2)$. Then the following approximation to the error probabilities may be derived: for $0<\theta<\infty$,

$$\mathrm{pr}_{-\theta}\{S_t>0\}\simeq 1-\Phi(N^{\frac{1}{2}}\theta)+1-\Phi(\eta_\theta^{-\frac{1}{2}}c'+\eta_\theta^{\frac{1}{2}}\theta)+e^{-2\theta c'}\{1-\Phi(\eta_\theta^{-\frac{1}{2}}c'-\eta_\theta^{\frac{1}{2}}\theta)\}, \tag{3}$$

where Φ denotes the standard normal distribution function. See § 6 for a plausibility argument leading to (3).

3. The estimation problem

When a truncated sequential probability ratio test is used, the estimation problem is to estimate θ from the sufficient statistic (t, S_t). Then $\hat{\theta}_t = S_t/t$ is still the maximum likelihood estimator, since the likelihood function is unaffected by optional stopping; but it may be badly biased. Let

$$S_t^* = t^{-\frac{1}{2}}(S_t - t\theta).$$

Then S_t^* is asymptotically standard normal as $c \to \infty$, under pr_θ, by a simple application of Anscombe's (1952) Theorem; but this approximation may not be very accurate for moderate values of c. In fact, the simulations presented in § 4 indicate that the standard normal approximation to the distribution of S_t^* may be quite poor.

Let μ denote the expected value of S_t^*; that is,

$$\mu(\theta) = E_\theta(S_t^*) \quad (-\infty < \theta < \infty).$$

Then the results of Woodroofe (1986, 1989) suggest that $S_t^* - \mu(\hat{\theta}_t)$ should be more nearly standard normal than S_t^*. Formally, a very weak asymptotic expansion of its distribution agrees with the standard normal up to terms of order $o(N^{-\frac{1}{2}})$. It is easily seen that

$$\mu(\theta) \simeq c^{-\frac{1}{2}}(g^{\frac{1}{2}})'(\theta) + o(c^{-\frac{1}{2}}) \tag{4}$$

for all $-\infty < \theta < \infty$, as $c, N \to \infty$ with $0 < \delta = 2c/N < \infty$ fixed. In fact, neglecting the excess over the boundary suggests that $t \simeq c/g(S_t/t)$; and Wald's Lemma then leads to

$$\mu(\theta) = E_\theta\{t^{-\frac{1}{2}}(S_t - t\theta)\} \simeq c^{-\frac{1}{2}}E[\{g^{\frac{1}{2}}(S_t/t) - g^{\frac{1}{2}}(\theta)\}(S_t - t\theta)]$$
$$\simeq c^{-\frac{1}{2}}E\left\{(g^{\frac{1}{2}})'(\theta)\frac{(S_t - t\theta)^2}{t}\right\} \simeq c^{-\frac{1}{2}}(g^{\frac{1}{2}})'(\theta).$$

This approximation may be justified along the lines of Siegmund (1978) or Woodroofe (1986). Accordingly, μ is estimated by $\hat{\mu}_t = c^{-\frac{1}{2}}(g^{\frac{1}{2}})'(\hat{\theta}_t)$. Continuing, let

$$\sigma^2(\theta) = E_\theta\{(S_t^* - \hat{\mu}_t)^2\} \quad (-\infty < \theta < \infty). \tag{5}$$

Then, again using Taylor series expansions, as in the derivation of (4), it may be shown that

$$\sigma^2(\theta) = 1 + c^{-1}(g^{\frac{1}{2}})'(\theta)^2 + o(c^{-1})$$

as $c \to \infty$ for all $\theta \neq \pm\delta$. Let

$$\hat{\sigma}_t^2 = 1 + c^{-1}(g^{\frac{1}{2}})'(\hat{\theta}_t)^2$$
$$S_t^\# = (S_t^* - \hat{\mu}_t)/\hat{\sigma}_t, \tag{6}$$

and $F_\theta^\#(z) = \text{pr}_\theta\,(S_t^\# \leq z)$ for all $-\infty < \theta,\ z < \infty$. Then it may be shown that $F_\theta^\#$ differs from the standard normal distribution by $o(N^{-1})$ in the very weak sense of Woodroofe (1986). The approach investigated here is to treat $S_t^\#$ as an approximate pivot and to approximate $F_\theta^\#$ by the standard normal distribution.

To the extent that this approximation is accurate,

$$\mathscr{I}_t = \left(\hat{\theta}_t - \frac{1}{\sqrt{t}}\hat{\mu}_t\right) \pm \frac{\hat{\sigma}_t}{\sqrt{t}}\Phi^{-1}\left(\frac{1+\gamma}{2}\right) \tag{7}$$

is an approximate $100\gamma\%$ confidence interval for any $0 < \gamma < 1$. The simulations presented in § 4 suggest that the approximation is quite accurate. Of course, the midpoint of the interval (7) may be regarded as an estimator of θ, which is median unbiased to a high order in the very weak sense used here. It is amusing to note that this estimator effectively subtracts half of the estimated conventional bias of $\hat{\theta}_t$, since $E_\theta(\hat{\theta}_t - \theta) \simeq g'(\theta)/c$ (Siegmund, 1978), and $\hat{\mu}_t/\sqrt{t} \simeq \frac{1}{2}g'(\hat{\theta}_t)/c$ with high probability.

4. Simulations

To study the accuracy of the approximations described above, simulation studies were conducted for selected values of c and N. Those for the case $c = 9{\cdot}0$ and $N = 72$, corresponding to a test with nominal power $\beta = 0{\cdot}95$ at $\theta = 0{\cdot}2$ from (3), are reported in detail. Some others are described briefly.

Monte Carlo estimates of the error probabilities and expected sample size are reported in Table 1, along with the approximations (2), (3) and (4). The approximation (3) appears to underestimate the error probabilities slightly for small θ, that is $\theta \leqslant 0{\cdot}1$. The differences are small, but significant. The approximation (2) to $E_\theta(t)$ is quite accurate, except near $\theta = 0$ and $\theta = \delta$ where (2) fails.

The most striking aspect of Table 1 is the magnitude of $\mu(\theta)$. The approximation (4) overestimates $\mu(\theta)$ everywhere and does so seriously for θ near to δ, where μ is largest. To some extent this is an advantage, since $(g^{\frac{1}{2}})'(\hat{\theta}_t)$ can be at most $\max_\theta (g^{\frac{1}{2}})'(\theta)$. Still it is surprising that the Monte Carlo estimates of $E_\theta(S_t^*)$ and $\{E_\theta(S_t^{*2})\}^{\frac{1}{2}}$ agree with the nominal values, 0 and 1, so well in the second and third columns of Table 2.

Monte Carlo estimates of $F_\theta^*(z)$ for $z = -1{\cdot}96$, $-1{\cdot}65$, $1{\cdot}65$ and $1{\cdot}96$, and of $F_\theta^*(z) - F_\theta^*(-z)$ for $z = 1{\cdot}65$ and $z = 1{\cdot}96$ are reported in Table 2. These show quite good agreement with the normal values. Of the 21 values of $F_\theta^*(1{\cdot}96) - F_\theta^*(-1{\cdot}96)$, two are significantly less than the nominal value 0·95 at level 0·05. Three values of $F_\theta^*(1{\cdot}645) - F_\theta^*(-1{\cdot}645)$

Table 1. *Error probabilities and expected sample sizes*

θ	$\mathrm{pr}_\theta(S_t<0)$ MC	$\mathrm{pr}_\theta(S_t<0)$ Approx. (3)	$E_\theta(t)$ MC	$E_\theta(t)$ Approx. (2)	$\mu(\theta)$ MC	$\mu(\theta)$ Approx. (4)
0·00	0·502	0·500	52·78	56·00	−0·004	0·000
0·05	0·346	0·336	51·75	54·77	0·074	0·092
0·10	0·208	0·199	48·75	49·93	0·144	0·175
0·15	0·111	0·107	44·69	43·29	0·192	0·243
0·20	0·049	0·052	39·78	36·49	0·228	0·294
0·25	0·020	0·022	34·93	38·33	0·244	0·333
0·30	0·007	0·008	36·10	31·94	0·245	0·304
0·35	0·003	0·003	27·03	27·38	0·232	0·282
0·40	0·001	0·001	23·94	23·96	0·226	0·264
0·45	0	0	21·48	21·30	0·215	0·248
0·50	0	0	19·40	19·17	0·208	0·236
0·55	0	0	17·69	17·42	0·200	0·225
0·60	0	0	16·26	15·97	0·189	0·215
0·65	0	0	15·01	14·74	0·186	0·207
0·70	0	0	13·97	13·69	0·178	0·198
0·75	0	0	13·07	12·78	0·175	0·192
0·80	0	0	12·26	11·98	0·169	0·186
0·85	0	0	11·56	11·27	0·164	0·181
0·90	0	0	10·92	10·65	0·158	0·176
0·95	0	0	10·36	10·09	0·155	0·171
1·00	0	0	9·87	9·58	0·151	0·167

MC, Monte Carlo estimates. Based on 10 000 replications.

Table 2. *Monte Carlo estimates of moments and coverage probabilities*

			$F^*(z)$				$F^*(z)-F^*(-z)$	
			$z=$	$z=$	$z=$	$z=$	$z=$	$z=$
θ	$E_\theta(S_t^*)$	$\surd E_\theta(S_t^{*2})$	$-1{\cdot}960$	$-1{\cdot}645$	$1{\cdot}645$	$1{\cdot}986$	$1{\cdot}960$	$1{\cdot}645$
0·00	−0·004	1·02	0·025	0·051	0·952	0·976	0·950	0·900
0·05	−0·004	1·02	0·025	0·051	0·952	0·976	0·951	0·902
0·10	−0·003	1·02	0·026	0·060	0·952	0·976	0·950	0·892
0·15	−0·006	1·00	0·032	0·066	0·953	0·978	0·946	0·887
0·20	−0·002	0·99	0·034	0·053	0·953	0·978	0·944	0·900
0·25	−0·001	0·99	0·027	0·045	0·952	0·977	0·951	0·908
0·30	−0·002	0·99	0·023	0·044	0·952	0·978	0·954	0·908
0·35	−0·012	0·99	0·020	0·044	0·954	0·979	0·959	0·909
0·40	−0·010	0·99	0·020	0·045	0·953	0·979	0·959	0·908
0·45	−0·012	0·99	0·021	0·048	0·954	0·978	0·957	0·906
0·50	−0·010	0·99	0·022	0·051	0·953	0·976	0·955	0·901
0·55	−0·009	1·00	0·023	0·050	0·953	0·978	0·955	0·904
0·60	−0·013	1·00	0·025	0·053	0·950	0·977	0·952	0·896
0·65	−0·009	1·00	0·025	0·052	0·953	0·976	0·951	0·901
0·70	−0·011	1·00	0·025	0·052	0·951	0·976	0·951	0·900
0·75	−0·008	1·00	0·026	0·053	0·950	0·977	0·951	0·896
0·80	−0·009	1·00	0·027	0·051	0·950	0·976	0·949	0·899
0·85	−0·009	1·00	0·025	0·051	0·952	0·975	0·950	0·901
0·90	−0·011	1·00	0·025	0·053	0·953	0·976	0·951	0·901
0·95	−0·009	1·00	0·026	0·054	0·951	0·978	0·952	0·897
1·00	−0·008	1·00	0·026	0·054	0·947	0·979	0·953	0·893

Based on 10 000 replications.

are significantly less than the nominal value. For purposes of setting confidence intervals, the normal approximation appears to be either adequate or conservative, except in the left tail for $\theta = 0{\cdot}10$, $0{\cdot}15$ and $0{\cdot}20$.

Similar results were found for two other combinations of design parameters, $c = 6{\cdot}5$ and $N = 52$, and $c = 11$ and $N = 128$, corresponding to tests with power at least 0·95 for $|\theta| \geqslant 0{\cdot}25$ and at least 0·94 for $|\theta| \geqslant 0{\cdot}15$, respectively. In these cases too, normal approximation appears to be adequate or conservative except in the left tail for $\delta/2 \leqslant \theta \leqslant \delta$.

5. Remarks

One possible modification to the procedure of § 3 is to replace c by $c' = c + \rho_0$ in (4) and the definition of $\hat{\mu}_t$, in an attempt to recover the neglected excess over the boundary (Siegmund, 1985, § 3.5). This reduces the amount by which (4) overestimates $\mu(\theta)$ slightly. A similar change in (5) and the definition of $\hat{\sigma}_t^2$ has a negligible effect.

A conservative modification is to replace $\hat{\sigma}_t$ by $1 + 1/(4\delta c)$, the maximum of the asymptotic expression in (5). For $c = 9{\cdot}0$ and $N = 72$, this leads to minimum Monte Carlo estimates of 0·898 and 0·952 for $F_\theta^*(1{\cdot}645) - F_\theta^*(-1{\cdot}645)$ and $F_\theta^*(1{\cdot}96) - F_\theta^*(-1{\cdot}96)$, while increasing the length of $\mathcal{I}_t$ in (7) by at most 5·5%.

The derivation described in § 3 is applicable to other stopping regions for which the stopping time may be written in the form (1), with a different g. Since the approach requires estimating $\mu(\theta) \simeq c^{-\frac{1}{2}}(g^{\frac{1}{2}})'(\theta)$, difficulty may be expected if the stopping region has corners, which reflect themselves in discontinuities in the derivative. For example, if the approach is used on Schwarz's (1962) test, for which $g(x) = (\delta + |x|)^2$ $(-\infty < x < \infty)$, then resulting approximations are not very accurate. Surprisingly, the approach worked quite well for the triangular test, for which $g(x)=\delta+|x|$ $(-\infty<x<\infty)$. For $\delta=0\cdot25$ and $c=12\cdot5$, the minimum and maximum Monte Carlo estimates of $F_\theta^\#(1\cdot96)-F_\theta^\#(-1\cdot96)$ over $\theta=0$ (0·05) 1 were 0·949 and 0·954. Such agreement may be expected for large values of θ, but not for small ones, in view of the discontinuity of g' at 0.

6. On the derivation of equation (3)

The derivation of (3) is discussed in this section. All limits are taken as c, $N\to\infty$ with fixed $0<\delta=2c/N<\infty$. For $\theta>0$, write

$$\mathrm{pr}_{-\theta}(S_t>0)=\mathrm{pr}_{-\theta}(S_t>0,\ S_N\leqslant 0)+\mathrm{pr}_{-\theta}(S_t>0,\ S_N>0). \tag{8}$$

An asymptotic approximation to the second term on the right is derived, along with an asymptotic upper bound for the first. Since these are of different orders of magnitude, it is difficult to assign a conventional mathematical meaning to the final approximation.

For the second term on the right-hand side of (8), write

$$\mathrm{pr}_{-\theta}(S_t>0,\ S_N>0)=1-\Phi(\theta N^{\frac{1}{2}})-\mathrm{pr}_{-\theta}(S_t<0,\ S_N>0).$$

The last probability on the right requires a very large increase from S_t to S_N and should, therefore, have small probability. To see that this is the case, let φ denote the standard normal density, let $\mathrm{pr}^x(.)=\mathrm{pr}(.|S_N=x)$ for $-\infty<x<\infty$, observe that pr^x does not depend on θ, and write

$$\begin{aligned}\mathrm{pr}_{-\theta}(S_t<0,\ S_N>0)&=\int_0^\infty \mathrm{pr}^x(S_t<0)N^{-\frac{1}{2}}\varphi(N^{-\frac{1}{2}}x+N^{\frac{1}{2}}\theta)\,dx\\&\leqslant\frac{1}{\theta}N^{-\frac{1}{2}}\varphi(N^{\frac{1}{2}}\theta)\int_0^\infty \mathrm{pr}^x(S_t<0)\theta\,e^{-\theta x}\,dx\\&=O[N\{1-\Phi(N^{\frac{1}{2}}\theta)\}\{1-\Phi(N^{\frac{1}{2}}\delta)\}],\end{aligned}$$

since

$$\begin{aligned}\mathrm{pr}^x(S_t<0)&\leqslant \mathrm{pr}_0\left[S_k-\frac{k}{N}S_N\leqslant-\delta\{k(N-k)\}^{\frac{1}{2}},\quad 1\leqslant k<N\right]\\&\leqslant\sum_{k=1}^{N-1}\mathrm{pr}_0\left[S_k-\frac{k}{N}S_N\leqslant-\delta\{k(N-k)\}^{\frac{1}{2}}\right]\leqslant N\{1-\Phi(N^{\frac{1}{2}}\delta)\}\end{aligned}$$

for all $x\geqslant 0$. So, $\mathrm{pr}_{-\theta}(S_t>0,\ S_N>0)\simeq 1-\Phi(N^{\frac{1}{2}}\theta)$.

For the first term on the right-hand side of (8), conditioning on (t, S_t) leads to

$$\mathrm{pr}_{-\theta}(S_N\leqslant 0\mid t, S_t)=\Phi\left\{\frac{(N-t)\theta-S_t}{(N-t)^{\frac{1}{2}}}\right\}\leqslant 1-\Phi\{\delta t^{\frac{1}{2}}-\theta(N-t)^{\frac{1}{2}}\}.$$

Observe that $\theta(N-t)^{\frac{1}{2}}-\delta t^{\frac{1}{2}}\geqslant 0$ if and only if $t\leqslant\eta_\theta=N\theta^2/(\delta^2+\theta^2)$. So,

$$\begin{aligned}&\mathrm{pr}_{-\theta}(\eta_\theta<t<N,\ S_t<0,\ S_N\leqslant 0)\\&\quad=\int_{\{\eta_\theta<t<N,\ S_t>0\}}[1-\Phi\{\delta t^{\frac{1}{2}}-\theta(N-t)^{\frac{1}{2}}\}]d\,\mathrm{pr}_{-\theta}\\&\quad\leqslant\sum_{\eta_\theta<k<N}[1-\Phi\{\delta k^{\frac{1}{2}}-\theta(N-k)^{\frac{1}{2}}\}]\,\mathrm{pr}_{-\theta}[S_k>\delta\{k(n-k)\}^{\frac{1}{2}}]\\&\quad\leqslant\sum_{\eta_\theta<k<N}[1-\Phi\{\delta k^{\frac{1}{2}}-\theta(N-k)^{\frac{1}{2}}\}][1-\Phi\{\delta(N-k)^{\frac{1}{2}}+\theta k^{\frac{1}{2}}\}]\\&\quad\leqslant\sum_{\eta_\theta<k<N}\exp\{-\tfrac{1}{2}(\delta^2+\theta^2)N\}\leqslant N\exp\{-\tfrac{1}{2}(\delta^2+\theta^2)N\}\end{aligned}$$

since $1-\Phi(x) \leqslant \exp(-x^2/2)$ for all $0 \leqslant x < \infty$. Thus, $\text{pr}_{-\theta}(t > \eta_\theta, S_t > 0, S_N \leqslant 0)$ is small. For the remaining term,

$$\text{pr}_{-\theta}(t \leqslant \eta_\theta, S_t > 0, S_N \leqslant 0) \leqslant \text{pr}_{-\theta}(t \leqslant \eta_\theta, S_t > 0)$$
$$\leqslant 1 - \Phi(\eta_\theta^{-\frac{1}{2}} c + \eta_\theta^{\frac{1}{2}} \theta) + e^{-2\theta c}\{1 - \Phi(\eta_\theta^{-\frac{1}{2}} c - \eta_\theta^{\frac{1}{2}} \theta)\},$$

by Bartlett's formula (Siegmund, 1985, equation (3.13)). Following Siegmund (1985, § 3.5), it is expected that a better approximation may be obtained by replacing c with $c' = c + \rho_0$.

Acknowledgements

The work reported here is an outgrowth of the Workshop on Sequential Methods in Statistics, held in Edinburgh July 4-July 15, 1988. In particular, David Siegmund and John Whitehead made valuable comments. The research for this paper was supported by the U.S. Army Research Office and the U.S. National Science Foundation.

References

Anscombe, F. (1952). Large sample theory of sequential estimation. *Proc. Camb. Phil. Soc.* **48**, 600–7.

Armitage, P. (1957). Numerical studies in the estimation of a binomial parameter. *Biometrika* **44**, 1–15.

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

Bather, J. (1988). Stopping rules and ordered families of distributions. *Seq. Anal.* **7**, 111–26.

Facey, K. & Whitehead, J. (1990). An improved approximation for calculation of a confidence intervals after a sequential clinical trial. *Statist. Med.* **9**, 1227–35.

Hagwood, C. & Woodroofe, M. (1982). On the expansion for the expected sample size in nonlinear renewal theory. *Ann. Prob.* **10**, 844–8.

Kim, K. (1987). An improved approximation for estimation following closed sequential tests. *Biometrika* **75**, 121–8.

Lai, T. L. & Siegmund, D. (1977). A nonlinear renewal theory I with applications to sequential analysis. *Ann. Statist.* **5**, 946–54.

Schwarz, G. (1962). Asymptotic shapes for Bayes sequential testing regions. *Ann. Math. Statist.* **33**, 224–36.

Siegmund, D. (1978). Estimation following sequential testing. *Biometrika* **65**, 341–9.

Siegmund, D. (1985). *Sequential Analysis.* New York: Springer.

Whitehead, J. (1983). *The Design and Analysis of Sequential Clinical Trials.* Chichester: Ellis Horwood.

Woodroofe, M. (1986). Very weak expansions for sequential confidence levels. *Ann. Statist.* **14**, 1049–67.

Woodroofe, M. (1989). Very weak expansions for sequentially designed experiments: linear models. *Ann. Statist.* **17**, 1087–102.

[*Received January* 1991. *Revised June* 1991]