The Annals of Statistics
1988, Vol. 16, No. 4, 1643–1666

REPEATED SIGNIFICANCE TESTS FOR EXPONENTIAL FAMILIES

BY INCHI HU

University of Pennsylvania

The problem of approximating the power and significance levels of repeated significance tests (RST) and modified repeated significance tests (MRST) is considered. The method due to Siegmund in the special case of normal observations with known variance is generalized. The main advantages that are claimed for this method are two-fold. First, it can be used to approximate the power of RSTs. Second, it can also be used to approximate the power and significance levels of MRSTs. Numerical and Monte Carlo results are also given for the repeated t-test.

1. Introduction.

1.1. In this article we study the significance levels and power of repeated significance tests (RST). An RST is the sequential version of a generalized likelihood ratio test. To be more precise, let us consider the following testing problem. Let $X_1, X_2, \dots$ be i.i.d. according to the distribution function F_θ, where $\{F_\theta\colon \theta \in \Theta\}$ forms a multiparameter exponential family. By that we mean F_θ has the form $F_\theta(dx) = e^{\theta' x - \psi(\theta)} F_0(dx)$ for some smooth function $\psi(\cdot)$ from the parameter space Θ into R^1 and some distribution function F_0 over R^d. Throughout the sequel we shall assume that F_0 has density with respect to Lebesgue measure on R^d and there exists an integer n_0 such that the n_0th convolution of this density is bounded. It is well known that $E_\theta X_1 = \mu(\theta) = \nabla\psi(\theta)$. Moreover, there is no loss of generality in assuming that $\mu(0) = 0$. Sometimes it is convenient to index this family by μ and write F_μ. We also use $\Sigma(\mu)$ to denote the covariance matrix of X_1 under F_μ. Let Θ_0 be a proper subset of $\Theta \subset R^d$. We are interested in testing $H_0\colon \theta \in \Theta_0$ against $H_1\colon \theta \notin \Theta_0$. The generalized log-likelihood ratio statistic after observing $X_1, X_2, \dots, X_n$ for this testing problem is

$$n\Lambda(S_n/n) = \sup_{\theta\in\Theta} l_n(\theta) - \sup_{\theta\in\Theta_0} l_n(\theta) = n\phi(S_n/n) - n\phi_0(S_n/n),$$

where $l_n(\theta) = \theta' S_n - n\psi(\theta)$ is the log-likelihood after observing $X_1, \dots, X_n$, $S_n = \sum_{i=1}^n X_i$ and

$$\phi(x) = \sup_{\theta\in\Theta} [\theta' x - \psi(\theta)], \qquad \phi_0(x) = \sup_{\theta\in\Theta_0} [\theta' x - \psi(\theta)].$$

An RST is defined in terms of the stopping rule

$$T = \inf\{n\colon n \ge m_0,\ n\Lambda(S_n/n) > a\}.$$

Received October 1985; revised September 1987.
AMS 1980 *subject classification*. Primary 62L10.
Key words and phrases. Exponential family, nonlinear renewal theory, sequential test.

Sampling stops at $\min(T, m)$ and H_0 is rejected when $T \le m$. Significance levels and power of an RST are given by

$$\sup_{\theta \in \Theta_0} P_\theta\{T \le m\} \text{ and } P_\theta\{T \le m\}, \qquad \theta \notin \Theta_0,$$

respectively, where P_θ denotes a probability law under which $X_1, X_2, \ldots$ are i.i.d. according to the distribution function F_θ. In some cases when one expects a small deviation from the null hypothesis and wants to increase the power, one may use a modified version of an RST. An MRST rejects H_0 when either $T \le m$ or $T > m$ and $m\Lambda(S_m/m) > c$ for some $c < a$.

Observe that if we fix c and let a tend to ∞, then it is unlikely that the log-likelihood ratio process $n\Lambda(S_n/n)$ will cross the level a before time m. In this case, the rejection region of the corresponding MRST reduces to $\{m\Lambda(S_m/m) > c\}$, which is exactly the rejection region of a fixed sample test. On the other hand, if we set $a = c$, then the corresponding MRST is just an RST. So an MRST can be thought of as a family of tests interpolating the fixed sample size test and RST.

Underlying this interpolation, there is a trade-off between the expected sample size and power. That is, as a moves from c to ∞ the power of an MRST increases to that of a fixed sample test at the cost of increasing the expected sample size. So with an MRST at hand the designer of an experiment has one more degree of freedom to choose from in fulfilling his need. If he thinks the power is more important he may choose an MRST with a substantially larger than c. If smaller expected sample size is desired he may choose a close to c.

The power of an MRST is given by

$$\begin{aligned} &P_\theta\{T \le m\} + P_\theta\{T > m, m\Lambda(S_m/m) > c\} \\ &\quad = P_\theta\{m\Lambda(S_m/m) > c\} + P_\theta\{T < m, m\Lambda(S_m/m) \le c\}. \end{aligned} \tag{1.1}$$

The quantity (1.1) also appears on other occasions. Siegmund (1985) suggests defining attained significance levels of an RST in the following way:

(i) If $T = m_0$ and $m_0\Lambda(S_{m_0}/m_0) = z > a$, then the attained level is $\sup_{\theta \in \Theta_0} P_\theta\{m_0\Lambda(S_{m_0}/m_0) > z\}$.

(ii) If $T = n \in (m_0, m]$, the attained level is $\sup_{\theta \in \Theta_0} P_\theta\{T \le n\}$.

(iii) If $T > m$ and $m\Lambda(S_m/m) = c$, then the attained level is

$$\begin{aligned} &\sup_{\theta \in \Theta_0} \left[P_\theta\{T \le m\} + P_\theta\{T > m, m\Lambda(S_m/m) > c\}\right] \\ &\quad = \sup_{\theta \in \Theta_0} \left[P_\theta\{m\Lambda(S_m/m) > c\} + P_\theta\{T < m, m\Lambda(S_m/m) \le c\}\right]. \end{aligned}$$

In case (iii) the attained significance level is of the same form as (1.1).

In this article we only consider a special kind of Θ_0,

$$\Theta_0 = \{\theta\colon \theta_1 = \cdots = \theta_{d_1} = 0\}, \qquad d_1 \le d.$$

By reparametrization Θ_0 can be generalized to $\Theta_0' = \{A\theta\colon \theta \in \Theta\}$, where A is a $d \times d$ matrix.

Typically the significance levels and powers of RSTs and MRSTs cannot be computed exactly and some approximation is required. Approximations for significance levels of RSTs in exponential families have been provided by Woodroofe (1978) and Lalley (1983). Their setting is more general than that given previously, but their methods are not as successful in approximating the power of RSTs and the power and significance levels of MRSTs. In what follows, we shall exhibit with a simple example three methods which have been developed by previous authors.

Let $X_1, X_2, \dots$ be i.i.d. according to $N(\theta, 1)$. We want to test $\theta = 0$ against $\theta \neq 0$. The RST in this case is defined by the stopping rule

$$T = \inf\{n\colon n \geq m_0,\ S_n^2/(2n) > a\}.$$

1.2. *The forward method.* The essential ingredients of this method are the likelihood ratio of a mixture measure Q and the probability measure P_0 under the null hypothesis and Wald's likelihood ratio identity. Let $Q(A) = \int_{-\infty}^{\infty} P_\theta(A)\, d\theta$. Then

$$dQ/dP_0(S_1, \dots, S_n) = \int_{-\infty}^{\infty} \exp(\theta S_n - n\theta^2/2)\, d\theta = (2\pi/n)^{1/2} \exp(S_n^2/2n).$$

Here the notation $d\mu/d\nu(Y)$ means that μ and ν are considered to be measures on a σ-field that contains $\sigma(Y)$, and $d\mu/d\nu(Y)$ is the Radon–Nikodym derivative of the restricted measures.

By Wald's likelihood ratio identity,

$$\begin{aligned} P_0\{T \leq m\} &= E_Q\{(T/2\pi)^{1/2} e^{-S_T^2/2T};\ T \leq m\} \\ &= \int_{-\infty}^{\infty} E_\theta\{(T/2\pi)^{1/2} \exp[-(a + R_m(T))];\ T \leq m\}\, d\theta \\ &= (a/\pi)^{1/2} e^{-a} \int_{-\infty}^{\infty} E_\theta\{(T/2a)^{1/2} e^{-R_m(T)};\ T \leq m\}\, d\theta, \end{aligned}$$

where $R_m(T) = (S_T^2/2T - a)$ is the corresponding excess over the boundary for this problem.

Before going any further, we introduce some notation. Throughout this work, let $R(T)$ denote the excess over the boundary corresponding to the stopping time T. Usually the stopping time depends on a parameter m. To emphasize the dependence on m, sometimes we write $R_m(T)$ or R_m and use $R_\infty(T)$ or R_∞ to denote the corresponding limit in distribution as $m \to \infty$.

If $a, m, m_0 \to \infty$ in such a way that $(2a/m)^{1/2} = \theta_1 < \theta_0 = (2a/m_0)^{1/2}$, then an argument using the strong law of large numbers shows that with P_θ-probability 1,

$$(T/2a)^{1/2} 1_{\{m_0 < T \leq m\}} \to \theta^{-1} 1_{[\theta_1, \theta_0]}$$

and

$$P_0\{T \leq m\} \sim (a/\pi)^{1/2} e^{-a} \int_{\theta_1}^{\theta_0} \theta^{-1} E_\theta\{e^{-R_\infty(T)}\}\, d\theta.$$

$E_\theta\{e^{-R_\infty(T)}\}$ can be approximated using nonlinear renewal theory developed by Lai and Siegmund (1977) or Woodroofe (1982) and the approximation is completed. This method has been generalized to RSTs for curved exponential families by Lalley (1983).

1.3. *The backward method.* The backward method due to Siegmund (1985, 1986) sets as its primary goal the approximation of the conditional probability $P_\xi^{(m)}(A) = P\{A|S_m = \xi\}$, which by the sufficiency of S_m is independent of θ. Then powers and significance levels may be obtained by unconditioning with respect to the distribution of S_m. In this article, we shall generalize this method to multiparameter exponential families.

The essence of this method involves randomizing the starting point of a conditional process, then regarding it as a process running backward from the point of conditioning. Let $P_{\lambda,\xi}^{(m)}(A) = P\{A|S_0 = \lambda,\ S_m = \xi\}$ and $T^* = \sup\{n\colon n \le m,\ S_n^2/2n > a\}$. Observe that $P_{0,\xi}^{(m)}\{T \le m\} = P_{0,\xi}^{(m)}\{T^* \ge m_0\}$. Let

$$\tilde{P}_\xi^{(m)}(A) = \int_{-\infty}^{\infty} P_{\lambda,\xi}^{(m)}(A)(2\pi m)^{-1/2}\exp\{-[(\lambda-\xi)^2/2m]\}\,d\lambda.$$

Then

$$\frac{d\tilde{P}_\xi^{(m)}}{dP_{0,\xi}^{(m)}}(S_n,\dots,S_m) = \left(\frac{n}{m}\right)^{1/2}\exp\left(\frac{S_n^2}{2n} - \frac{\xi^2}{2m}\right).$$

Since under the reversed time scale T^* is a stopping time, Wald's likelihood ratio identity gives

$$P_\xi^{(m)}\{T^* \ge m_0\} = \tilde{E}_\xi^{(m)}\left\{\left(\frac{m}{T^*}\right)^{1/2}\exp\left(\frac{\xi^2}{2m} - \frac{S_{T^*}^2}{2T^*}\right);\ T^* \ge m_0\right\}.$$

The $\tilde{P}_\xi^{(m)}$ distributions of S_n, $n = m, m-1,\dots,$ running backward from $S_m = \xi$ is the same as the P_0 distributions of $\xi - S_n$, $n = 0, 1, \dots,$ running forward.

Hence the previous expectation equals

$$E_0\left\{\left(\frac{m}{m-\tau}\right)^{1/2}\exp\left[\frac{\xi^2}{2m} - \frac{(S_\tau+\xi)^2}{2(m-\tau)}\right];\ \tau \le m - m_0\right\},$$

where $\tau = \inf\{n\colon\ n \ge 1,\ (\xi + S_n^2)/[2(m-n)] > a\}$. Assume that $\theta_1 = (2a/m_1)^{1/2}$ and $\xi_0 = \xi/m$. A law of large numbers argument shows that $\tau/m \to 1 - (\xi_0/\theta_1)^2$ with probability 1 as $m \to \infty$. The preceding quantity is approximated by

$$(m\theta_1/\xi)\exp(-a + \xi^2/2m)E_0\{e^{-R_m(\tau)}\},$$

where

$$R_m(\tau) = \{(S_\tau - \xi)^2/[2(m-\tau)]\} - a$$

is the excess over the boundary at the stopping time τ. Again nonlinear renewal theory can be used to obtain the asymptotic distribution of $R_\infty(\tau)$ and the approximation to $P_\xi^{(m)}\{T \le m\}$ is completed. Unconditioning ξ using the marginal distribution of S_m under P_0 yields the approximation to $P_0\{T \le m\}$.

We may uncondition ξ using P_θ with $\theta \neq 0$ and obtain approximations to the power. Unfortunately, the result is not a bona fide asymptotic expression, although numerical results show that it is a very good approximation. See Siegmund (1985), Section 9.3, for details.

1.4. *Woodroofe's method.* This method, which was developed by Woodroofe, is quite different from the two methods described previously. It does not use Wald's likelihood ratio identity. The method first approximates $P_0\{T = m\}$ then estimates $P_0\{T \le m\}$ by summation. Observe that $P_0\{T = n\} \sim 2P\{T_+ = n\}$, where

$$T_+ = \inf\{n: n \ge m_0, S_n > \sqrt{2na}\}$$

and

$$P_0\{T_+ = n\} = \int_{\sqrt{2an}}^{\infty} P_\xi^{(n)}\{T_+ > n-1\}(2\pi n)^{-1/2}\exp(-\xi^2/2n)\,d\xi.$$

It is easy to see that the only values of ξ which are of first-order importance are $\sqrt{2an} + O(1)$. In this range we can approximate the curve $\sqrt{2na}$ by its tangent and the conditional random walk by an unconditional one (with drift ξ_0). That is, let $\xi = \sqrt{2an} + y$, where y is arbitrary but fixed,

$$\begin{aligned} P_\xi^{(n)}\{T_+ > n-1\} &= P_0\{S_k < \sqrt{2ak} \text{ for all } m_0 \le k \le n-1 | S_n = \sqrt{2an} + y\} \\ &= P_0\{S_n - S_k > y + \sqrt{2a}(n^{1/2} - k^{1/2}) \\ &\qquad \text{for all } m_0 \le k < n | S_n = \sqrt{2an} + y\} \\ &= P_0\{S_i > y + \sqrt{2a}[n^{1/2} - (n-i)^{1/2}] \\ &\qquad \text{for all } 1 \le i \le n - m_0 | S_n/n = \mu^* + O(1)\}. \end{aligned}$$

Observe that $\sqrt{2n}[n^{1/2} - (n-i)^{1/2}] = \frac{1}{2}\sqrt{2a/n}\,i + O(\sqrt{2a/n}\,n^{-3/2}i^2) \to \frac{1}{2}\mu^* i$ if $n \to \infty$, $a \to \infty$ in such a way that $(2a/n)^{1/2} \to \mu^*$. The preceding conditional probability is asymptotically equivalent to

$$P_{\mu^*}\{S_i > y + \tfrac{1}{2}\mu^* i \text{ for all } i \ge 1\} = P_{\mu^*/2}\{S_i > y \text{ for all } i \ge 1\}.$$

To continue, we need

LEMMA [Woodroofe (1982)]. *Assume* $\mu = EX_1 > 0$. *Let* $M = \min(S_1, S_2 \dots)$. *Then for* $x > 0$,

$$[E(S_{\tau_+})]^{-1}P\{S_{\tau_+} > x\} = \mu^{-1}P\{M > x\}, \quad \textit{where } \tau_+ = \inf\{n: n \ge 1, S_n > 0\}.$$

By the lemma and the argument given previously,

$$\begin{aligned} P_0\{T = n\} &\sim 2P_0\{T_+ = n\} \\ &\sim \int_0^\infty \mu^*[E_{\mu^*/2}(S_{\tau_+})]^{-1}P_{\mu^*/2}\{S_{\tau_+} > y\}(2\pi n)^{-1/2} \\ &\qquad \times \exp\{-[(\sqrt{2an} + y)^2]/2n\}\,dy \\ &\sim n^{-1}(a/\pi)^{1/2}e^{-a}\int_0^\infty [E_{\mu^*/2}(S_{\tau_+})]^{-1}P_{\mu^*/2}\{S_{\tau_+} > y\}e^{-\mu^* y}\,dy. \end{aligned}$$

The preceding integral equals $\lim_{a\to\infty} E_{\mu^*}\{\exp[-R_a(T)]\}$ by renewal theorem, where $R_a(T) = S_T^2/2T - a$. Summing over n and approximating the sum by an integral yields the desired result. Now we are in a position to make brief comments on the three methods described previously.

If one were only concerned with the significance levels of RSTs, then the forward method is the most general of the three. If one wants second-order approximations to significance levels, then Woodroofe's method appears to be the appropriate method to use (cf. Woodroofe, 1977, and Woodroofe and Takahashi, 1982). But if we restrict our attention to the linear hypothesis, then the backward method produces the most fruitful results. It can be used to approximate significance levels, power and p-values of RSTs and MRSTs. One of the major contributions of this article is to generalize the backward method to multiparameter exponential families.

The rest of the article is organized as follows. The main results are given in Section 2. In Section 3, we prove the main theorem which allows us to generalize the backward method. Section 4 contains the proofs of two theorems on repeated t-tests. In Section 5, we discuss the duality of the forward and backward method, highlighting Theorem 7, which relates the excess over the boundary of these two methods. The numerical results on repeated t-tests are given in Section 6.

2. Main results.

2.1. *The simple null hypothesis.* In the case of simple null hypothesis (i.e., Θ_0 contains only one point), the backward method can be generalized to multiparameter exponential families in a straightforward fashion. Without loss of generality we may assume $\Theta_0 = \{0\}$. In this case the stopping rule is given by

$$T = \{n\colon n \ge m_0,\ n\phi(S_n/n) > a\}.$$

It will be convenient to use the notation

$$P_\xi^{(m)}(A) = P_0(A \mid S_m = \xi),$$

where A belongs to the σ-field generated by $X_1, \dots, X_m$.

Throughout the sequel we shall let $\xi = m\xi_0$, $a = ma_0$. Let $H_1(t, x) = (1-t)\phi((\xi_0 - x)/(1-t))$ and define

$$t_1 = \inf\{t\colon t > 0,\ a_0/(1-t) = \phi[\xi_0/(1-t)]\}. \tag{2.1}$$

Let Z_i, $i = 1, 2, 3, \dots$, be a sequence of i.i.d. random variables and each Z_i has the same distribution as $\nabla H_1(t_1, 0)\cdot(1, X_1)$, where X_1 is distributed according to F_0. Put $V_n = \sum_{i=1}^n Z_i$ and define

$$\tau_+ = \inf\{n\colon n > 0,\ V_n > 0\}.$$

THEOREM 1. *Suppose $a \to \infty$, $m \to \infty$, $m_0 \to \infty$ in such a way that $a/m = a_0 > 0$, $1 > m_0/m = a_1 > 0$. Then for each ξ_0 which satisfies $\phi(\xi_0) < a_0$ and $t_1 < 1 - a_1$, we have*

$$\begin{aligned} P_\xi^{(m)}\{T \le m\} \sim{}& (1-t_1)^{-d/2} |\Sigma(\xi_0)|^{1/2} |\Sigma[\xi_0/(1-t_1)]|^{-1/2} \\ &\times \exp\{-m[a_0 - \phi(\xi_0)]\}\nu_1(\xi_0), \end{aligned} \tag{2.2}$$

where

$$\nu_1(\xi_0) = \left[E(V_{\tau_+})\right]^{-1} \int_0^{\infty} e^{-x} P\{V_{\tau_+} > x\}\, dx.$$

COROLLARY 1.1. *Suppose in addition to the asymptotic relations of Theorem 1 that* $c/m = c_0 < a_0$, *and the set* $A = \{\xi_0\colon \phi(\xi_0) \le c_0,\ t_1 < 1 - m_0/m\}$ *contains an open set of* R^d. *Then*

$$P_0\{T < m,\ m\phi(S_m/m) \le c\}$$
$$\sim (m/2\pi)^{d/2} e^{-a} \int_A \left|\Sigma[\xi_0/(1-t_1)]\right|^{-1/2} (1-t_1)^{-d/2} \nu_1(\xi_0)\, d\xi_0 \qquad (2.3)$$

and

$$P_\theta\{T < m,\ m\phi(S_m/m) \le c\}$$
$$\sim (m/2\pi)^{d/2} \exp[-m(a_0 + \psi(\theta) - \theta\cdot\beta)] \qquad (2.4)$$
$$\times \left|\Sigma[\beta/(1-t_\beta)]\right|^{-1/2} (1-t_\beta)^{-d/2} \nu_1(\beta),$$

where β *maximizes* $\theta\cdot\xi_0$ *over the set* A, *and* t_β *is defined in the same way as* t_1 *in* (2.1) *with* ξ_0 *replaced by* β.

The proof of Theorem 1 is omitted here since it is similar to the arguments given in Section 1.2 [see Hu (1985) for details]. Corollary 1.1 follows from unconditioning (2.2) with respect to the distribution of S_m.

2.2. *The composite null hypothesis.* For any vector V in R^d, let $V^{(1)}, V^{(2)}$ be two vectors such that $V = (V^{(1)}, V^{(2)})$, where $V^{(1)} \in R^{d_1}$, $V^{(2)} \in R^{d_2}$, $d_1 + d_2 = d$.

The null hypothesis we consider here takes the form

$$H_0\colon \theta = (\theta^{(1)}, \theta^{(2)}) \in \Theta_0 = \{\theta\colon \theta^{(1)} = 0^{(1)}\},$$

where $0^{(1)}$ denotes the zero vector in R^{d_1}. $\theta^{(2)}$ plays the role of nuisance parameters. The stopping rule is given by

$$T = \inf\{n\colon n \ge m_0,\ n\Lambda(S_n/n) > a\}.$$

All the following notation corresponds in the obvious way to that of the simple hypothesis case. Let $H(t, x) = (1-t)\Lambda[(\xi_0 - x)/(1-t)]$ and define

$$t_0 = \inf\{t\colon t > 0,\ (1-t)\Lambda[(\xi_0 - t\mu_0)/(1-t)] = a_0\},$$

where $\mu_0 = \mu(\theta_0)$, with $\theta_0 = (0^{(1)}, \theta^{(2)}(\xi_0^{(2)}))$. Let W_i, $i = 1, 2, 3, \ldots$, be a sequence of i.i.d. random variables and each W_i has the same distribution as $\nabla H(t_0, t_0\mu_0)\cdot(1, X)$, where X is distributed according to distribution function F_{θ_0}. Put $V_n = \sum_{i=1}^n W_i$ and define $\tau_+ = \inf\{n\colon n > 0,\ V_n > 0\}$.

THEOREM 2. *Suppose* $a \to \infty$, $m \to \infty$, $m_0 \to \infty$ *in such a way that* $a/m = a_0 > 0$, $1 - m_0/m = a_1 > 0$. *Then for each* ξ_0 *such that* $t_0 < 1 - m_0/m$ *and*

$\Lambda(\xi_0) < a_0$, *we have*

$$P_\xi^{(m)}\{T \le m\} \sim (1 - t_0)^{-d_1/2} |\ddot{\Sigma}(\xi_0)|^{1/2} \times |\ddot{\Sigma}[(\xi_0 - t_0\mu_0)/(1 - t_0)]|^{-1/2} \nu(\xi_0) \times \exp\{-m[a_0 - \Lambda(\xi_0)]\}, \tag{2.5}$$

where

$$\nu(\xi_0) = \int_0^\infty e^{-x} P(V_{\tau_+} > x)\, dx / E(V_{\tau_+}). \tag{2.6}$$

2.3. *Repeated t-tests.* Theorem 2 can be applied to obtain results on repeated t-tests. Assume $X_1, X_2, \dots$ are independent and normally distributed with unknown mean μ and variance σ^2 and that we are interested in testing H_0: $\mu = 0$ against H_1: $\mu \ne 0$. Let $l_{\mu,\sigma}(\cdot)$ denote the log-likelihood of X_1. Simple algebra shows that the (generalized) log-likelihood ratio statistic is

$$n\Lambda\left(\frac{S_n}{n}\right) = n\left[\phi\left(\frac{S_n}{n}\right) - \phi_0\left(\frac{S_n^{(2)}}{n}\right)\right] = \left(\frac{n}{2}\right)\log\left\{\frac{S_n^{(2)}/n}{S_n^{(2)}/n - \left(S_n^{(1)}/n\right)^2}\right\},$$

where

$$S_n = \left(S_n^{(1)}, S_n^{(2)}\right) = \left(\sum_{i=1}^n X_i, \sum_{i=1}^n X_i^2\right),$$

$$\phi(x_1, x_2) = \sup_{\sigma,\mu} l_{\mu,\sigma}(x_1, x_2) = \left[(x_2 - 1) - \log\left(x_2 - x_1^2\right)\right]/2,$$

$$\phi_0(x_2) = \sup_{\sigma} l_{0,\sigma}(x_2) = \left[(x_2 - 1) - \log x_2\right]/2.$$

The repeated t-test is defined in terms of the stopping rule

$$T = \inf\{n : n \ge m_0, n\Lambda(S_n/n) > a\}.$$

The repeated t-test rejects H_0 when $T \le m$, while the modified test rejects H_0 if the sample path belongs to

$$\{T \le m\} \cup \{T > m, m\Lambda(S_m/m) > c\}.$$

Observe that the test statistic $\Lambda(S_n/n)$ is scale invariant. So under the null hypothesis, the probability of any event which is measurable with respect to the σ-field generated by $\Lambda(S_i/i)$, $i = 1, 2, 3, \dots$, is independent of the variance σ^2, and we may write

$$P_0\{T \le m\}$$

and

$$P_0\{T \le m\} + P_0\{T > m, m\Lambda(S_m/m) > c\} = P_0\{m\Lambda(S_m/m) > c\} + P_0\{T < m, m\Lambda(S_m/m) \le c\},$$

for significance levels of the repeated t-test and the modified test, respectively. Let $Y_1, Y_2, \dots$ be an i.i.d. sequence of random variables, each having the same distribution as

$$Y = -\left[\theta^2/2(1+\theta^2)\right]Z^2 + \left[\theta/(1+\theta^2)\right]Z + \left[\theta^2/(1+\theta^2) + \log(1+\theta^2)\right]/2, \tag{2.7}$$

where Z is standard normally distributed. Put $U_n = \sum_{i=1}^n Y_i$ and let

$$\theta_0 = \left(e^{2a/m_0} - 1\right)^{1/2}, \tag{2.8}$$

$$\theta_1 = \left(e^{2a/m} - 1\right)^{1/2}, \tag{2.9}$$

$$\theta_2 = \left(e^{2c/m} - 1\right)^{1/2}. \tag{2.10}$$

Define $\nu_+(\theta) = \int_0^\infty e^{-x} P\{U_{\tau_+} > x\}\, dx / E(U_{\tau_+})$.

THEOREM 3. *Assume $a \to \infty$, $m \to \infty$, $m_0 \to \infty$, $c \to \infty$ such that $0 < \theta_2 < \theta_1 < \theta_0 < \infty$. Then*

$$P_0\{T \le m,\, m\Lambda(S_m/m) < c\} \sim 2(a/\pi)^{1/2} e^{-a} \int_{\tilde\theta}^{\theta_0} \left[\log(1+\theta^2)\right]^{-1/2} \nu_+(\theta)\, d\theta, \tag{2.11}$$

where $\tilde\theta$ is the positive solution of the equation

$$\frac{\tilde\theta^2}{(1+\tilde\theta^2)\log(1+\tilde\theta^2)} = \frac{\theta_2^2}{(1+\theta_2^2)\log(1+\theta_1^2)}. \tag{2.12}$$

Replacing c by a in Theorem 3 yields

COROLLARY 3.1.

$$P_0\{T \le m\} \sim 2(a/\pi)^{1/2} e^{-a} \int_{\theta_1}^{\theta_0} \left[\log(1+\theta^2)\right]^{-1/2} \nu_+(\theta)\, d\theta. \tag{2.13}$$

REMARK. Theorem 3 of Woodroofe (1978) contains a mistake (an incorrect Jacobian), so the results on the repeated t-test in Woodroofe (1979) are also incorrect.

It is easy to see that the power of the modified test depends on the parameters μ, σ^2 only through the ratio $\eta = \mu/\sigma$, so without loss of generality we may take $\sigma = 1$ and $\mu = \eta$. By symmetry the power at η equals that at $-\eta$, so we may assume $\eta > 0$. The power of the modified t-test is given by

$$P_\eta\{T \le m\} + P_\eta\{T > m,\, m\Lambda(S_m/m) > c\} = P_\eta\{T < m,\, m\Lambda(S_m/m) \le c\} + P_\eta\{m\Lambda(S_m/m) > c\}. \tag{2.14}$$

The second term on the r.h.s. of (2.14) can be obtained by approximating the tail probabilities of noncentral t-distribution. It is therefore sufficient to construct an approximation for the first term. Let $\alpha = (1 - e^{-2c/m})^{1/2}$ and define

$g(x) = (x - 1 - \log x)/2 - \eta \alpha x^{1/2}$. Put

$$\gamma_2 = \left[\left(\alpha^2\eta^2 + 4\right)^{1/2} + \alpha\eta\right]^2/4, \tag{2.15}$$

$$\gamma_1 = \alpha\gamma_2^{1/2} = (\alpha/2)\left[\left(\alpha^2\eta^2 + 4\right)^{1/2} + \alpha\eta\right]. \tag{2.16}$$

Let t_γ be the positive solution of the equation

$$(1 - t_\gamma)\log\left[1 - \gamma_1^{-1}\gamma_2^2/(1 - t_\gamma)^2\right] = -2a_0. \tag{2.17}$$

THEOREM 4. *Under the same asymptotic relations as in Theorem* 3, *we have*

$$P_\eta\{T < m,\ m\Lambda(S_m/m) \le c\}$$

$$\sim m^{-1/2}\exp\left[-m\left(a_0 + g(\gamma_2) + \eta^2/2\right)\right]\nu_+\left\{\left[\exp\left(2a_0/(1 - t_\gamma)\right) - 1\right]^{1/2}\right\}$$

$$\times\exp\left[3a_0/(1 - t_\gamma)\right]\eta^{-1}a_0\left\{\exp\left[2a_0/(1 - t_\gamma)\right] - 1 - a_0/(1 - t_\gamma)\right\}^{-1}$$

$$\times\left[2\pi\left(\gamma_2 + \alpha\eta\gamma_2^{3/2}/2\right)(1 - t_\gamma)^3\right]^{-1/2}.$$

3. Proof of Theorem 2. The most tricky part of the proof of Theorem 2 is the construction of a measure Q whose likelihood ratio relative to $P_\xi^{(m)}$ is a simple function of the stopping rule (asymptotically). We will construct Q by randomizing the starting point of the sufficient process S_n according to a conditional distribution. Define $T^* = \sup\{n\colon n\Lambda(S_n/n) > a\}$ and $P_{\lambda,\xi}^{(m)}(A) = p(A \mid S_0 = \lambda,\ S_m = \xi)$. It is easy to see that

$$\frac{dP_{\lambda,\xi}^{(m)}}{dP_{0,\xi}^{(m)}}(S_n, \dots, S_m) = \frac{f_n(S_n - \lambda)f_m(\xi)}{f_n(S_n)f_m(\xi - \lambda)},$$

where $f_n(\cdot)$ is the density function of S_n under P_0.

Let $f_m^{(2)}(\cdot)$ denote the density function of $S_n^{(2)}$ under P_0 and let $0^{(2)}$ be the zero vector in R^{d_2}. Define

$$Q(A) = \int_{R^{d_1}} P_{\lambda_0,\xi}^{(m)}(A)\frac{f_m(\xi - \lambda_0)}{f_m^{(2)}(\xi^{(2)})}\,d\lambda^{(1)},$$

where $\lambda_0 = (\lambda^{(1)}, 0^{(2)})$. Observe that $f_m(\xi - \lambda_0)/f_m^{(2)}(\xi^{(2)})$ is the conditional density of S_m given $S_m^{(2)} = \xi^{(2)}$. The likelihood ratio of Q relative to $P_\xi^{(m)}$ is found to be

$$\frac{dQ}{dP_\xi^{(m)}}(S_n, \dots, S_m) = \int_{R^{d_1}} \frac{dP_{\lambda_0,\xi}^{(m)}}{dP_{0,\xi}^{(m)}}(S_n, \dots, S_m)\frac{f_m(\xi - \lambda_0)}{f_m^{(2)}(\xi^{(2)})}\,d\lambda^{(1)}$$

$$= \frac{f_m(\xi)f_n^{(2)}(S_n^{(2)})}{f_m^{(2)}(\xi^{(2)})f_n(S_n)}.$$

The distributions of S_n under Q, $n = m, m - 1, \dots,$ running backward from $S_m = m\xi_0$ are the same as the conditional P_0 distributions of $m\xi_0 - S_n$, $n =$

$0, 1, \ldots,$ running forward and tied down at $S_m^{(2)} = m\xi_0^{(2)}$. Under the reverse time scale, T^* is a stopping time, so Wald's likelihood ratio identity gives

$$
\begin{aligned}
P_\xi^{(m)}\{T < m\} &= P_\xi^{(m)}\{T^* \geq m_0\} = E_Q\left\{\frac{dP_\xi^{(m)}}{dQ}(S_{T^*}, \ldots, S_m);\ T^* \geq m_0\right\} \\
&= E_Q\left\{\frac{f_{T^*}(S_{T^*}) f_m^{(2)}(\xi^{(2)})}{f_{T^*}^{(2)}(S_{T^*}^{(2)}) f_m(\xi)};\ T^* \geq m_0\right\}.
\end{aligned} \tag{3.1}
$$

Define

$$
\tau = \inf\left\{n\colon (m-n)\Lambda\left(\frac{\xi - S_n}{m-n}\right) > a,\ n \leq m - m_0\right\}.
$$

It is easy to see that the distribution of T^* under Q is the same as the distribution of $m - \tau$ under $P_{\xi^{(2)}}^{(m)}(A) = P_0(A \mid S_m^{(2)} = \xi^{(2)})$, so the last term of (3.1) can be replaced by

$$
E_{\xi^{(2)}}^{(m)}\left\{\frac{f_{m-\tau}(\xi - S_\tau) f_m^{(2)}(\xi^{(2)})}{f_{m-\tau}^{(2)}(\xi^{(2)} - S_\tau^{(2)}) f_m(\xi)};\ \tau \leq m - m_0\right\}, \tag{3.2}
$$

where $E_{\xi^{(2)}}^{(m)}$ denotes the expectation corresponding to the conditional probability $P_{\xi^{(2)}}^{(m)}$. To proceed further, we need

PROPOSITION 5 [Borovkov and Rogozin (1965)]. *If for some integer n_0, S_{n_0} has a bounded continuous density f_{n_0} with respect to the Lebesgue measure on R^d, then as $n \to \infty$,*

$$
f_n(nx) \sim (2\pi n)^{-d/2} |\Sigma(x)|^{-1/2} \exp[-n\phi(x)]. \tag{3.3}
$$

Moreover the preceding limit is attained uniformly over compact subsets.

By Proposition 5 we have

$$
\begin{aligned}
\frac{f_m^{(2)}(\xi^{(2)})}{f_m(\xi)} &\sim (2\pi m)^{d_1/2} |\Sigma(\xi_0)|^{1/2} |\Sigma^{(2)}(\xi_0^{(2)})|^{-1/2} \\
&\quad \times \exp\{m[\phi(\xi_0) - \phi_0(\xi_0^{(2)})]\} \\
&= (2\pi m)^{d_1/2} |\Sigma(\xi_0)|^{1/2} |\Sigma^{(2)}(\xi_0^{(2)})|^{-1/2} \exp[m\Lambda(\xi_0)],
\end{aligned} \tag{3.4}
$$

where $\Sigma^{(2)}(\mu)$ is the covariance matrix of $S_1^{(2)}$ under P_μ. To approximate the remaining part of (3.2), we need the following lemma whose proof will be given in the Appendix.

LEMMA 6. *Under the same assumptions as in Theorem 2, there exists a compact set K of R^d such that*

$$
\begin{aligned}
&E_{\xi^{(2)}}^{(m)}\left\{\frac{f_{m-\tau}(\xi - S_\tau)}{f_{m-\tau}^{(2)}(\xi^{(2)} - S_\tau^{(2)})};\ \tau \leq m - m_0,\ S_\tau \notin mK\right\} \\
&\quad = o\left(E_{\xi^{(2)}}^{(m)}\left\{\frac{f_{m-\tau}(\xi - S_\tau)}{f_{m-\tau}^{(2)}(\xi^{(2)} - S_\tau^{(2)})};\ \tau \leq m - m_0,\ S_\tau \in mK\right\}\right),
\end{aligned}
$$

where $mK = \{mx: x \in K\}$.

By Lemma 6 we have

$$E^{(m)}_{\xi^{(2)}}\left\{\frac{f_{m-\tau}(\xi - S_\tau)}{f^{(2)}_{m-\tau}\left(\xi^{(2)} - S^{(2)}_\tau\right)}; \tau \le m - m_0\right\} \tag{3.5}$$
$$\sim E^{(m)}_{\xi^{(2)}}\left\{\frac{f_{m-\tau}(\xi - S_\tau)}{f_{m-\tau}\left(\xi^{(2)} - S^{(2)}_\tau\right)}; \tau \le m - m_0, S_\tau \in mK\right\}.$$

Since (3.3) holds uniformly over compact sets, the r.h.s. of (3.5) is asymptotically equivalent to

$$E^{(m)}_{\xi^{(2)}}\left\{2\pi(m-\tau)^{-d_1/2}\left|\Sigma^{(2)}\left(\frac{\xi^{(2)} - S^{(2)}_\tau}{m-\tau}\right)\right|^{1/2}\left|\Sigma\left(\frac{\xi - S_\tau}{m-\tau}\right)\right|^{-1/2}\right.$$
$$\left.\times\exp\left\{-(m-\tau)\left[\phi\left(\frac{\xi - S_\tau}{m-\tau}\right) - \phi_0\left(\frac{\xi^{(2)} - S^{(2)}_\tau}{m-\tau}\right)\right]\right\}; \tau \le m - m_0, S_\tau \in mK\right\}$$

$$= E^{(m)}_{\xi^{(2)}}\left\{2\pi(m-\tau)^{-d_1/2}\left|\Sigma^{(2)}\left(\frac{\xi^{(2)} - S^{(2)}_\tau}{m-\tau}\right)\right|^{1/2}\left|\Sigma\left(\frac{\xi - S_\tau}{m-\tau}\right)\right|^{-1/2}\right. \tag{3.6}$$
$$\times\exp\left\{-\left[(m-\tau)\Lambda\left(\frac{\xi - S_\tau}{m-\tau}\right) - a\right]\right\};$$
$$\left.\tau \le m - m_0, S_\tau \in mK\right\}e^{-a}.$$

Observe that the exponent inside the expectation on the r.h.s. of (3.6) is exactly the excess over the boundary at stopping time τ. The nonlinear renewal theory for conditional random walks developed in Chapter 4 of Hu (1985) will be employed to finish the proof.

The following limit relations are valid [see Hu (1985), Chapter 4]. For all $\varepsilon > 0$,

$$\lim_{m\to\infty} P^{(m)}_{\xi^{(2)}}\{|S_\tau/\tau - \mu_0| > \varepsilon\} = 0, \tag{3.7}$$

$$\lim_{m\to\infty} P^{(m)}_{\xi^{(2)}}\{|\tau/m - t_0| > \varepsilon\} = 0. \tag{3.8}$$

By (3.7) and (3.8) we have

$$P^{(m)}_{\xi^{(2)}}\left\{\left|\det\left(\Sigma^{(2)}\left[\left(\xi^{(2)} - S^{(2)}_\tau\right)/(m-\tau)\right]\right) - \det\left(\Sigma^{(2)}\left(\xi^{(2)}\right)\right)\right| > \varepsilon\right\} \to 0, \tag{3.9}$$

$$P^{(m)}_{\xi^{(2)}}\left\{\left|\det\left(\Sigma\left[(\xi - S_\tau)/(m-\tau)\right]\right)\right.\right. \tag{3.10}$$
$$\left.\left. -\det\left(\Sigma\left[(\xi_0 - t_0\mu_0)/(1-t_0)\right]\right)\right| > \varepsilon\right\} \to 0,$$

as $m \to \infty$. In view of (3.8)–(3.10), the r.h.s. of (3.6) is asymptotically equivalent to

$$(3.11)\quad \begin{aligned} &m^{-d_1/2}e^{-a}(1-t_0)^{-d_1/2}\left|\Sigma^{(2)}\left(\xi_0^{(2)}\right)\right|^{1/2}\left|\Sigma\left[(\xi_0 - t_0\mu_0)/(1-t_0)\right]\right|^{-1/2} \\ &\quad \times E_{\xi^{(2)}}^{(m)}\left\{e^{-R_m(\tau)};\ \tau \le m - m_0,\ S_\tau \in mK\right\}, \end{aligned}$$

where $R_m(\tau) = (m-\tau)\Lambda((\xi - S_\tau)/(m-\tau)) - a$. Under the assumption $t_0 < 1 - m_0/m$, the last factor in (3.11) is asymptotically equivalent to

$$E_{\xi^{(2)}}^{(m)}\{e^{-R_m(\tau)}\}.$$

By Theorem 6 of Chapter 4 of Hu (1985), we have

$$(3.12)\qquad E_{\xi^{(2)}}^{(m)}\{e^{-R_m(\tau)}\} \to \nu(\xi_0).$$

Substituting (3.4) and (3.11) into (3.2) and using (3.12), we have

$$\begin{aligned} P_{\xi^{(2)}}^{(m)}\{T \le m\} &\sim (1-t_0)^{-d_1/2}\left|\Sigma(\xi_0)\right|^{1/2}\left|\Sigma\left[(\xi_0 - t_0\mu_0)/(1-t_0)\right]\right|^{-1/2}\nu(\xi_0) \\ &\quad \times \exp\{-m[a_0 - \Lambda(\xi_0)]\}, \end{aligned}$$

which is the desired result.

4. Proofs of Theorems 3 and 4.

4.1. *Proof of Theorem* 3. The starting point of the proof is Theorem 2. We will identify those quantities on the r.h.s. of (2.5) in the special case of the repeated t-test given in the following discussion. Let $\xi/m = \xi_0 = (\xi_1, \xi_2)$, $a_0 = a/m$. Define

$$(4.1)\qquad t_0 = \inf\{t\colon (1-t)\Lambda[\xi_1/(1-t), \xi_2] = a_0\},$$

where

$$\Lambda(x_1, x_2) = (1/2)\log\left[x_2/\left(x_2 - x_1^2\right)\right].$$

Denote by $\Sigma(\xi_1, \xi_2)$ the covariance matrix of (X, X^2), where X is normally distributed with first and second moments given by $E(X) = \xi_1$, $E(X^2) = \xi_2$. A simple calculation shows that the determinant of Σ is

$$(4.2)\qquad |\Sigma(\xi_0)| = 2\left(\xi_2 - \xi_1^2\right)^3.$$

Define

$$H(t, x_1, x_2) = (1-t)\Lambda[(\xi_0 - x)/(1-t)].$$

Let W_i, $i = 1, 2, \dots$, be i.i.d. random variables having the same distributions as $\nabla H(t_0, 0, t_0\xi_2)\cdot(1, \xi_0^{1/2}Z, \xi_2 Z^2)$, where Z is distributed according to the standard normal. An easy calculation shows that W_i has the same distribution as

$$(4.3)\quad \begin{aligned} W = {}&\frac{Z^2}{2}\frac{\xi_1^2}{2\left[\xi_2(1-t_0)^2 - \xi_1^2\right]} - Z\frac{\xi_1(1-t_0)}{\xi_2^{3/2}(1-t_0)^2 - \xi_1^2\xi_2^{1/2}} \\ &+ \frac{1}{2}\left\{\frac{\xi_1^2}{\xi_2(1-t_0)^2 - \xi_1^2} - \log\left[\frac{\xi_2(1-t_0)^2}{\xi_2(1-t_0)^2 - \xi_1^2}\right]\right\}. \end{aligned}$$

Let $V_n = \Sigma_{i=1}^n W_i$ and

$$\nu_-(\xi_0) = \int_0^\infty e^{-x} P\{V_{\tau_+} > x\}\, dx / EV_{\tau_+}.$$

By Theorem 2 we have

$$(4.4)\quad \begin{aligned} P_\xi^{(m)}\{T \le m\} \sim{}& (1 - t_0)^{-1/2} |\Sigma(\xi_0)|^{1/2} |\Sigma[\xi_1/(1 - t_0), \xi_2]|^{-1/2} \nu_-(\xi_0) \\ &\times \exp\{-m[a_0 - \Lambda(\xi_0)]\}. \end{aligned}$$

In order to obtain significance levels from (4.4) we need to uncondition it with respect to $P_0\{S_m \in m\, d\xi_0\}$. Observe that each term on the r.h.s. of (4.4) is a function of $z = \xi_1/\xi_2^{1/2}$, hence a function of $y = (m-1)^{1/2} z (1 - z^2)^{-1/2}$. This reduces the conditional probability $P_\xi^{(m)}\{T \le m\}$, which in general is a function of two variables, to a function of one variable. This is because the likelihood ratio statistic is scale invariant. It is easy to see that the random variable corresponding to y has a t-distribution with $m - 1$ degrees of freedom. Multiplying (4.4) by the t-density and integrating over the appropriate range yields

$$(4.5)\quad \begin{aligned} &P_0\{T \le m;\ m\Lambda(S_m/m) < c\} \\ &\quad \sim e^{-a} \int_D (\xi_2 - \xi_1^2)^{3/2} [\xi_2 - \xi_1^2 (1 - t_0)^{-2}]^{-3/2} (1 - t_0)^{-1/2} \\ &\qquad \times \nu_-(\xi_0) e^{m\Lambda(\xi_0)} g_{m-1}(y)\, dy, \end{aligned}$$

where

$$(4.6)\quad \begin{aligned} g_{m-1}(y) &= \frac{\Gamma(m/2)}{[(m-1)\pi]^{1/2} \Gamma[(m-1)/2]} \left(\frac{y^2}{m-1} + 1 \right)^{-m/2}, \\ D &= \{y: 0 < t_0 < 1 - m_0/m,\ \Lambda(\xi_0) < c_0\}. \end{aligned}$$

The last factor on the r.h.s. of (4.6) cancels with

$$e^{m\Lambda(\xi_0)} = \exp\{(m/2)\log[\xi_2/(\xi_2 - \xi_1^2)]\} = [y^2/(m-1) + 1]^{m/2}.$$

After calculation, we have

$$(4.7)\quad \frac{dy}{dz} = (m-1)^{1/2} (1 - z^2)^{-3/2} = (m-1)^{1/2} [\xi_2/(\xi_2 - \xi_1^2)]^{3/2}$$

and

$$(4.8)\quad \xi_2^{3/2} [\xi_2 - \xi_1^2 (1 - t_0)^{-2}]^{-3/2} = \exp[3a_0/(1 - t_0)].$$

Applying Stirling's formula to approximate the first factor on the r.h.s. of (4.5) and using (4.7) and (4.8), (4.5) becomes

$$(4.9)\quad \begin{aligned} &P_0\{T < m;\ m\Lambda(S_m/m) \le c\} \\ &\quad \sim (m/2\pi)^{1/2} e^{-a} \int_C \exp[3a_0/(1 - t_0)] (1 - t_0)^{-1/2} \nu_-[\xi_0(z)]\, dz, \end{aligned}$$

where

$$C = \{z: 0 < t_0 < 1 - m_0/m,\ \Lambda(\xi_0) < c_0\}.$$

We need to make another change of variables. Let

$$\theta = \frac{z/(1-t_0)}{\left[1 - z^2/(1-t_0)^2\right]^{1/2}}. \tag{4.10}$$

By (4.1)

$$\log(1+\theta^2) = -\log\left[1 - z^2/(1-t_0)^2\right] = 2a_0/(1-t_0). \tag{4.11}$$

Let $\theta_0, \theta_1, \theta_2$ be the same as in (2.8)–(2.10). Observe that

$$0 < t_0 < 1 - m_0/m \Leftrightarrow \theta_1 < \theta < \theta_0 \text{ or } -\theta_0 < \theta < -\theta_1. \tag{4.12}$$

Let $\tilde{\theta}$ be the solution of (4.10) corresponding to $z = \theta_2^2/(1+\theta_2^2)$. Then by (4.10) $\tilde{\theta}$ satisfies (2.12). Notice that $\theta_1 < \tilde{\theta}$ if $c < a$ and that $c = a$ implies $\theta_1 = \tilde{\theta}$. So

$$\Lambda(\xi_0) < c_0 \Leftrightarrow |z| < \theta_2^2/(1+\theta_2^2) \Leftrightarrow \theta > \tilde{\theta} \text{ or } \theta < -\tilde{\theta}. \tag{4.13}$$

(4.12) and (4.13) determine the range of integration

$$\{0 < t_0 < 1 - m_0/m,\ \Lambda(\xi_0) < c_0\} \Leftrightarrow \{\tilde{\theta} < \theta < \theta_0\}\{-\theta_0 < \theta < -\tilde{\theta}\}. \tag{4.14}$$

An easy calculation shows that

$$\left|\frac{dz}{d\theta}\right| = (1-t_0)(1+\theta^2)^{-3/2}\left[2\theta^2 - \log(1+\theta^2)\right]\left[\log(1+\theta^2)\right]^{-1}. \tag{4.15}$$

In view of (4.11), (4.14) and (4.15), (4.9) becomes

$$\begin{aligned} &P_0\{T < m,\ m\Lambda(S_m/m) \le c\} \\ &\quad \sim 2(a/\pi)^{1/2} e^{-a} \int_{\tilde{\theta}}^{\theta_0} \left[\log(1+\theta^2)\right]^{-3/2} \\ &\qquad \times \left[2\theta^2 - \log(1+\theta^2)\right] \nu_-[\xi_0(\theta)]\, d\theta. \end{aligned} \tag{4.16}$$

Clearly (4.3) can be rewritten in terms of θ. That is,

$$\begin{aligned} W &= (\theta^2/2)Z^2 - \theta(1+\theta^2)^{1/2} Z + (1/2)\left[\theta^2 - \log(1+\theta^2)\right] \\ &= (\theta^2/2)\left[Z - (1+\theta^2)^{1/2}/\theta\right]^2 - (1/2)\left[1 + \log(1+\theta^2)\right]. \end{aligned}$$

Comparing $-W$ with Y defined by (2.7), we find that $-W$ and Y have the same support, $(-\infty, (1/2)[1 + \log(1+\theta^2)])$. In fact, $-W$ and Y are distributed as $-(1/2)[\theta^2/(1+\theta^2)]\chi^2(\theta^{-2}) + (1/2)[1 + \log(1+\theta^2)]$ and $-(1/2)\theta^2\chi^2(\theta^{-2}+1) + (1/2)[1 + \log(1+\theta^2)]$, respectively, where $\chi^2(\rho)$ denotes the noncentral χ^2-distribution with one degree of freedom and noncentrality parameter ρ. So the likelihood ratio of Y relative to $-W$ exists.

$\chi^2(\rho)$ has density in the form [see, e.g., Ferguson (1967), page 103]

$$\sum_{j=0}^{\infty} e^{-\rho/2}(\rho/2)^j (j!)^{-1}\left[\Gamma(j+1/2)\right]^{-1} 2^{-(j+1/2)} e^{-x/2} x^{j-1/2}, \qquad 0 < x < \infty.$$

So Y has density

$$f_Y(y) = 2(1+\theta^2)\theta^{-2}\sum_{j=0}^{\infty} e^{-1/2\theta^2}(2\theta^2)^{-j}(j!)^{-1}[\Gamma(j+1/2)]^{-1}2^{-(j+1/2)}$$
$$\times\exp\{(1+\theta^2)[y-1/2-(1/2)\log(1+\theta^2)]/\theta^2\}$$
$$\times\left\{\theta^2(1+\theta^2)^{-1}[1+\log(1+\theta^2)-2y]^{j-1/2}\right\}$$
$$=\theta^{-1}\exp\left\{(1+\theta^2)y-[2+\theta^2+\log(1+\theta^2)](2\theta^2)^{-1}\right\}$$
$$\times\sum_{j=0}^{\infty}(\theta^{-4}+\theta^{-2})^j[\Gamma(j+1/2)]^{-1}2^{-2j+1/2}$$
$$\times(j!)^{-1}[1+\log(1+\theta^2)-2y]^{j-1/2},$$
$$-\infty < y < (1/2)[1+\log(1+\theta^2)].$$

$-W$ has density

$$f_{-W}(y) = 2\theta^{-2}\sum_{j=0}^{\infty}\exp\left[-(2\theta^2)^{-1}(1+\theta^2)\right]\left[(2\theta^2)^{-1}(1+\theta^2)\right]^j(j!)^{-1}$$
$$\times[\Gamma(j+1/2)]^{-1}2^{-(j+1/2)}$$
$$\times\exp\{\theta^{-2}[y-1/2-(1/2)\log(1+\theta^2)]\}$$
$$\times\{\theta^{-2}[1+\log(1+\theta^2)-2y]\}^{j-1/2}$$
$$=\theta^{-1}\exp\left\{\theta^{-2}y-[2+\theta^2+\log(1+\theta^2)](2\theta^2)^{-1}\right\}$$
$$\times\sum_{j=0}^{\infty}(\theta^{-4}+\theta^{-2})^j[\Gamma(j+1/2)]^{-1}2^{-2j+1/2}$$
$$\times(j!)^{-1}[1+\log(1+\theta^2)-2y]^{j-1/2},$$
$$-\infty < y < (1/2)[1+\log(1+\theta^2)].$$

The ratio of these two densities is surprisingly simple:

$$f_Y(y)/f_{-W}(y) = e^y.$$

Appealing to Theorem 7, we have

$$\nu_-(\theta) = \nu_+(\theta)EY/EW. \tag{4.17}$$

The expectations of Y and W are easy to evaluate:

$$EW = \theta^2 - (1/2)\log(1+\theta^2) \tag{4.18}$$

and

$$EY = (1/2)\log(1+\theta^2). \tag{4.19}$$

Substituting (4.17)–(4.19) into (4.16) yields the desired result.

4.2. *Proof of Theorem* 4. By Proposition 5, S_m has P_η asymptotic density in the form

$$(4.20)\quad f_{m,\eta}(m\xi_0) \sim (2\pi m)^{-1}|\Sigma(\xi_0)|^{-1/2}\exp\{-m[\phi(\xi_0) - \eta\xi_1 - \eta^2/2]\}.$$

Unconditioning (2.5) with respect to the r.h.s. of (4.20) gives

$$P_\eta\{T < m,\ m\Lambda(S_m/m) \le c\}$$

$$(4.21)\quad \sim (m/2\pi)e^{-m(a_0+\eta^2/2)}\int\int_B (1-t_0)^{-1/2}\left|\Sigma\left[\xi_1(1-t_0)^{-1}, \xi_2\right]\right|^{-1/2}$$

$$\times \nu_-(\xi_0)\exp\{-m[\phi_0(\xi_2) - \eta\xi_1]\}\, d\xi_2\, d\xi_1,$$

where

$$B = \{(\xi_1, \xi_2)\colon 0 < t_0 \le 1 - m_0/m,\ \Lambda(\xi_0) \le c_0/m\}.$$

After calculation, we have

$$B = \{(\xi_1, \xi_2)\colon \alpha_1 \le \xi_1/\xi_2^{1/2} \le \alpha\},$$

where

$$\alpha = (1 - e^{-2c/m})^{1/2},$$

$$\alpha_1 = \frac{m_0}{m}(1 - e^{-2a/m_0})^{1/2}.$$

Although the double integral on the r.h.s. of (4.21) is taken over the set B, the only part which contributes to the first-order approximation is the integral over a small neighborhood N of $\gamma = (\gamma_1, \gamma_2)$, where (γ_1, γ_2) minimizes $\phi_0(\xi_2) - \eta\xi_1$ over B. It is not hard to see that the minimum of $\phi_0(\xi_2) - \eta\xi_1$ over B occurs on the curve $\xi_1\xi_2^{-1/2} = \alpha$. On this curve, $\phi_0(\xi_2) - \eta\xi_1$ equals

$$g(\xi_2) = -\eta\alpha\xi_2^{1/2} - (\xi_2 - 1 - \log\xi_2)/2.$$

A simple calculation shows that γ_1 and γ_2 are given by (2.15) and (2.16), respectively. The nonexponential factor inside the integral on the rh.s. of (4.21) is approximately constant over N. Namely,

$$(4.22)\quad \begin{aligned}&(1-t_0)^{-1/2}\left|\Sigma\left[\xi_1(1-t_0)^{-1}, \xi_2\right]\right|^{-1/2}\nu_-(\xi_0)\\ &\quad\approx (1-t_\gamma)^{-1/2}\left|\Sigma\left[\gamma_1(1-t_\gamma)^{-1}, \gamma_2\right]\right|^{-1/2}\nu_-(\gamma),\end{aligned}$$

where t_γ is defined by (2.17).

It remains to approximate $\int\int_B e^{-m[\phi_0(\xi_2)-\eta\xi_1]}\, d\xi_2\, d\xi_1$. The following argument uses a change of variables to convert the double integral into an integral of a single variable, which can then be approximated by the Laplace method. Let

$y_1 = \xi_1 \xi_2^{-1/2},\ y_2 = \xi_2,$

$$\begin{aligned}
&\iint_B \exp\{-m[\phi_0(\xi_2) - \eta\xi_1]\}\, d\xi_2\, d\xi_1 \\
&\quad = \int_0^\infty \int_{\alpha_1}^{\alpha} \exp\{-m[\phi_0(y_2) - \eta y_1 y_2^{1/2}]\} y_2^{1/2}\, dy_2\, dy_1 \\
&\quad = \int_0^\infty \exp[-m\phi_0(y_2)]\, y_2^{1/2} \left[\int_{\alpha_1}^{\alpha} \exp(m\eta y_1 y_2^{1/2})\, dy_1\right] dy_2 \\
&\quad = \int_0^\infty \exp[-m\phi_0(y_2)]\, y_2^{1/2} \{(m\eta y_2^{1/2})^{-1} \\
&\qquad\qquad \times [\exp(m\eta_2 y_2^{1/2}\alpha) - \exp(m\eta y_2^{1/2}\alpha_1)]\}\, dy_2 \\
&\quad \sim \int_0^\infty \exp[-m\phi_0(y_2)](m\eta)^{-1} \exp(m\eta y_2^{1/2}\alpha)\, dy_2 \\
&\quad = (m\eta)^{-1} \int_0^\infty e^{-mg(y_2)}\, dy_2.
\end{aligned}$$

The Laplace method gives

$$(m\eta)^{-1} \int_0^\infty e^{-mg(y_2)}\, dy_2 \sim (m\eta)^{-1} \{2\pi[mg''(\gamma_2)]^{-1}\}^{1/2} e^{-mg(\gamma_2)}. \tag{4.23}$$

Substituting (4.22) and (4.23) into (4.21), we have

$$\begin{aligned}
&P_\eta\{T < m,\ m\Lambda(S_m/m) \le c\} \\
&\sim \exp\{-m[a_0 + \eta^2/2 + g(\gamma_2)]\} \qquad (4.24)\\
&\quad \times \nu_-(\gamma)\eta^{-1}\{2\pi m g''(\gamma_2)(1 - t_\gamma)|\Sigma[\gamma_1/(1 - t_\gamma), \gamma_2]|\}^{-1/2} \\
&= m^{-1/2} \exp\{-m[a_0 + \eta^2/2 + g(\gamma_2)]\}\nu_+\{[\exp[2a_0/(1 - t_\gamma)] - 1]^{1/2}\} \\
&\quad \times \exp[3a_0/(1 - t_\gamma)]\, a_0 \eta^{-1}\{\exp[2a_0/(1 - t_\gamma)] - 1 - a_0/(1 - t_\gamma)\}^{-1} \qquad (4.25)\\
&\quad \times \{2\pi[\gamma_2 + (1/2)\alpha\eta\gamma_2^{3/2}](1 - t_\gamma)^3\}^{-1/2}.
\end{aligned}$$

In the preceding equality we have used

$$\begin{aligned}
\nu_-(\gamma) = \nu_+\{[\exp[2a_0/(1 - t_\gamma)] - 1]^{1/2}\}[a_0/(1 - t_\gamma)] \\
\times \{\exp[2a_0/(1 - t_\gamma)] - 1 - a_0/(1 - t_\gamma)\}^{-1},
\end{aligned}$$

which is the immediate consequence of (4.11) and (4.17)–(4.19), and

$$|\Sigma[\gamma_1/(1 - t_\gamma), \gamma_2]|^{-1/2} = 2^{-1/2}\gamma_2^{-3/2} \exp[3a_0/(1 - t_\gamma)],$$

which follows from (4.2) and (4.8). This completes the proof of Theorem 4.

5. Duality of the forward and backward methods. A variant of the forward method [see Theorem 5.29 of Siegmund (1985)], which involves taking the likelihood ratio of the maximum invariant process $X_1^{-1}X_2, X_1^{-1}X_2, \dots, X_1^{-1}X_n, \dots$ then mixing it by Lebesgue measure over the invariant parameter space, can be employed to derive Corollary 3.1. This method cannot be easily modified to prove Theorem 3 because of the existence of a subtle measurability problem. In short, the two-dimensional sufficient process $(\Sigma_{i=1}^n X_i, \Sigma_{i=1}^n X_i^2)$ is not measurable with respect to the σ-field generated by the maximum invariant process.

Although the method mentioned previously cannot be employed to prove Theorem 3, it suggests the succinct form of Theorem 3. The following theorem is crucial in proving Theorem 3.

THEOREM 7. *Let $Y_1, Y_2, \dots$ and $Z_1, Z_2, \dots$ be two sequences of independent identically distributed random variables and let $U_n = \Sigma_{i=1}^n Y_i$, $V_n = \Sigma_{i=1}^n Z_i$. Let $f_Y(\cdot)$ and $f_{-Z}(\cdot)$ denote the probability density functions of Y_i and $-Z_i$, respectively. Assume*

$$f_Y(y)/f_{-Z}(y) = e^y \tag{5.1}$$

and $EZ_1 > 0$, $EY_1 > 0$. Then

$$\frac{\int_0^\infty e^{-x}P\{U_{\tau_+} > x\}\,dx/EU_{\tau_+}}{\int_0^\infty e^{-x}P\{V_{\tau_+} > x\}\,dx/EV_{\tau_+}} = \frac{EZ_1}{EY_1},$$

where τ_+ denotes the first time the corresponding random walk is positive.

When there is a possibility of confusion, we use a superscript to denote the random walk we refer to, e.g., $\tau_+^U = \inf\{n > 0: U_n > 0\}$.

PROOF OF THEOREM 7. Integration by parts gives

$$\int_0^\infty e^{-x}P\{U_{\tau_+} > x\}\,dx/EU_{\tau_+} = (EU_{\tau_+})^{-1}\{1 - E[\exp(-U_{\tau_+})]\}. \tag{5.2}$$

By Wald's lemma we have

$$EU_{\tau_+} = EY_1 \cdot E\tau_+^U. \tag{5.3}$$

By the duality lemma

$$(E\tau_+^U)^{-1} = P\{\tau_-^U = \infty\}, \tag{5.4}$$

where τ_-^U is the first time the corresponding random walk is nonpositive. Appealing to Wald's likelihood ratio identity, we have

$$\begin{aligned} P\{\tau_+^{-V} < \infty\} &= E\left\{\prod_{i=1}^{\tau_+^U} \frac{f_{-Z}(Z_i)}{f_Y(Z_i)};\ \tau_+^U < \infty\right\} \\ &= E[\exp(-U_{\tau_+})], \end{aligned} \tag{5.5}$$

where in the second equality given previously we have used (5.1). Substituting (5.3)–(5.5) into (5.2), we have

$$(5.6)\qquad \int_0^\infty e^{-x}P\{U_{\tau_+} > x\}\,dx/EU_{\tau_+} = P\{\tau_-^U = \infty\}P\{\tau_+^{-V} = \infty\}/EY_1.$$

A similar argument shows that

$$(5.7)\qquad \int_0^\infty e^{-x}p\{V_{\tau_+} > x\}\,dx/EV_{\tau_+} = P\{\tau_-^V = \infty\}P\{\tau_+^{-U} = \infty\}/EZ_1.$$

By the obvious scale property

$$(5.8)\qquad P\{\tau_+^{-U} = \infty\} = P\{\tau_-^U = \infty\}$$

and

$$(5.9)\qquad P\{\tau_+^{-V} = \infty\} = P\{\tau_-^V = \infty\}.$$

Dividing (5.6) by (5.7), using (5.8) and (5.9), the desired result follows. □

REMARK. The relation (5.1) holds not only for the repeated t-test but also for other cases; see Hu (1985), Chapter 2, for details.

Although Theorem 7 is not difficult to prove, its importance cannot be overemphasized. It provides the crucial link to proving the equivalence of the forward and backward approximations in cases in which both methods work. Also, for cases in which only the backward method is applicable, Theorem 7 allows one to convert backward representation into forward representation efficiently. This reduces the programming work in numerical computation. To be more specific, in evaluating the excess over the boundary numerically, Theorem 7 guarantees that only one subroutine is enough.

6. Numerical results. In the numerical computation of the approximations given by Theorems 3 and 4, the main programming task is to evaluate numerically $\nu_+(\cdot)$, the Laplace transform of the excess over the boundary. It turns out that $\nu_+(\cdot)$ relates to the characteristic function of Z_1, the increment of the random walk V_n which generates the excess over the boundary, through an integral formula. See the theorem in Woodroofe (1979).

Sometimes it is inconvenient to take the data continuously, so it is helpful to consider group sequential tests on these occasions. The stopping rule we are interested in is

$$T_k = \inf\left\{n\colon n = m_0 + ik,\ i = 0,\dots,\left[\frac{-m - m_0}{k}\right],\ n\Lambda(S_n/n) > a\right\},$$

where k is the size of a group. It is easy to see that T_k is a stopping time. After a moment's reflection, we find that approximation to the corresponding significance levels and power are the same as $k = 1$, except for the excess over the boundary part. To find the excess over the boundary, it is sufficient to identify the increment of the random walk which generates the excess over the boundary. In the repeated t-test it is given by $U_k = \Sigma_{i=1}^{k} Y_i$, where the distribution of Y_i is given by (2.7).

Tables 1–6 give some examples of the approximations to power and significance levels of RSTs and MRSTs. For comparison, the results of Monte Carlo experiments are also included.

TABLE 1
Significance levels of repeated t-tests

a	m_0	m	Analytic approximation	Monte Carlo (2000 replications)[a]
3.8	7	30	0.050	0.053 ± 0.001
4.0	8	50	0.047	0.048 ± 0.001
4.5	10	75	0.032	0.033 ± 0.0006
5.0	10	110	0.024	0.023 ± 0.0004

[a] Importance sampling [cf. Siegmund (1976)] is used in the preceding Monte Carlo experiments.

TABLE 2
The power of repeated t-tests

a	m_0	m	η	Analytic approximation	Monte Carlo (2000 replications)
3.8	7	30	0.8	0.946	0.951 ± 0.005
			0.6	0.734	0.742 ± 0.010
4.0	8	30	0.6	0.934	0.933 ± 0.006
			0.4	0.584	0.596 ± 0.008
4.0	10	75	0.5	0.950	0.948 ± 0.005
			0.3	0.518	0.522 ± 0.011
5.0	10	110	0.4	0.882	0.889 ± 0.007
			0.3	0.581	0.581 ± 0.011

TABLE 3
Significance levels of modified repeated t-tests

a	c	m_0	m	Analytic approximation	Monte Carlo (6000 replications)
3.8	3.6	7	30	0.050	0.053 ± 0.001
3.95	3.6	7	40	0.050	0.052 ± 0.001
4.0	3.6	8	50	0.048	0.049 ± 0.0009
4.7	4.2	10	80	0.028	0.027 ± 0.0007
5.0	4.5	10	100	0.023	0.023 ± 0.0066

TABLE 4
The power of modified repeated t-tests

a	c	m_0	m	η	Analytic approximation	Monte Carlo (2000 replications)
3.8	3.6	7	30	0.8	0.952	0.956 ± 0.005
3.95	3.6	7	40	0.7	0.960	0.959 ± 0.004
				0.5	0.717	0.727 ± 0.010
4.0	3.6	8	30	0.6	0.946	0.943 ± 0.005
				0.4	0.613	0.626 ± 0.011
4.7	4.2	10	80	0.5	0.947	0.937 ± 0.006
				0.4	0.779	0.770 ± 0.010
5.0	4.5	10	100	0.45	0.940	0.938 ± 0.005
				0.3	0.553	0.550 ± 0.001

TABLE 5
Significance levels of group repeated t-tests

Number of observations in a group	a	m_0	m	Analytic approximation	Monte Carlo (2000 replications)[a]
2	3.65	8	40	0.050	0.052 ± 0.001
3	3.6	10	55	0.049	0.049 ± 0.001
4	3.6	10	70	0.051	0.052 ± 0.001
5	3.6	10	80	0.050	0.052 ± 0.001
7	3.6	15	120	0.047	0.047 ± 0.001

[a] Importance sampling is used in the preceding Monte Carlo experiments.

TABLE 6
The power of group repeated t-tests

Number of observations in a group	a	m_0	m	η	Analytic approximation	Monte Carlo (2000 replications)
2	3.65	8	40	0.7	0.962	0.961 ± 0.004
				0.6	0.880	0.888 ± 0.007
				0.5	0.726	0.741 ± 0.010
3	3.6	10	55	0.6	0.969	0.966 ± 0.004
				0.4	0.681	0.685 ± 0.010
4	3.6	10	70	0.5	0.949	0.940 ± 0.005
				0.3	0.527	0.518 ± 0.011
5	3.6	10	80	0.5	0.973	0.961 ± 0.004
				0.4	0.855	0.843 ± 0.008
				0.3	0.590	0.574 ± 0.011
7	3.6	15	120	0.4	0.970	0.966 ± 0.004
				0.3	0.790	0.773 ± 0.009
				0.2	0.420	0.414 ± 0.011

APPENDIX

PROOF OF LEMMA 6.

$$E_{\xi^{(2)}}^{(m)}\left\{\frac{f_{m-\tau}(\xi-S_\tau)}{f_{m-\tau}^{(2)}\left(\xi^{(2)}-S_\tau^{(2)}\right)};\ \tau\le m-m_0,\ S_\tau\notin mK\right\}$$

$$\text{(A.1)}\qquad =\sum_{i=1}^{m-m_0}E_{\xi^{(2)}}^{(m)}\left\{\frac{f_{m-i}(\xi-S_i)}{f_{m-i}^{(2)}\left(\xi^{(2)}-S_i^{(2)}\right)};\ \tau=i,\ S_i\notin mK\right\}$$

$$=\sum_{i=1}^{m-m_0}E_{\theta_0}\left\{\frac{dP_{\xi^{(2)}}^{(m)}}{dP_{\theta_0}}(S_1,\dots,S_i)\frac{f_{m-i}(\xi-S_i)}{f_{m-i}^{(2)}\left(\xi^{(2)}-S_i^{(2)}\right)};\ \tau=i,\ S_i\notin mK\right\},$$

where in the last equality we have used Wald's likelihood ratio identity. A simple calculation shows that

$$\text{(A.2)}\qquad \frac{dP_{\xi^{(2)}}^{(m)}}{dP_{\theta_0}}(S_1,\dots,S_i)=\frac{f_{m-i}^{(2)}\left(\xi^{(2)}-S_i^{(2)}\right)}{\left\{f_m^{(2)}\left(\xi^{(2)}\right)\exp\left[S_i^{(2)}\theta_0^{(2)}-i\psi^{(2)}\left(\theta_0^{(2)}\right)\right]\right\}}.$$

Replacing the likelihood ratio in (A.1) by the r.h.s. of (A.2), we find that the last term of (A.1) equals

$$\sum_{i=1}^{m-m_0}E_{\theta_0}\left\{\frac{f_{m-i}(\xi-S_i)}{f_m^{(2)}\left(\xi^{(2)}\right)}\exp\left[-S_i^{(2)}\theta_0^{(2)}+i\psi^{(2)}\left(\theta_0^{(2)}\right)\right];\ \tau=i,\ S_i\notin mK\right\},$$

which is less than

$$\text{(A.3)}\qquad \begin{aligned}&\sum_{i=1}^{m-m_0}E_{\theta_0}\left\{\frac{f_{m-i}(\xi-S_i)}{f_m^{(2)}\left(\xi^{(2)}\right)}\exp\left[-S_i^{(2)}\theta_0^{(2)}+i\psi^{(2)}\left(\theta_0^{(2)}\right)\right];\ S_i\notin mK\right\}\\ &\quad=\sum_{i=0}^{m-m_0}\int_{\{x\notin mK\}}\frac{f_{m-i}(\xi-x)}{f_m^{(2)}\left(\xi^{(2)}\right)}f_i(x)\,dx.\end{aligned}$$

By the assumption we made in Section 1, $|f_{m-i}(\xi-x)|\le b$ for some constant b, if $m-i\ge n_0$. By Proposition 5

$$f_m^{(2)}\left(\xi^{(2)}\right)\ge(1/2)(m/2\pi)^{d_2/2}\left|\Sigma^{(2)}\left(\xi_0^{(2)}\right)\right|^{1/2}\exp\left[m\phi_0\left(\xi_0^{(2)}\right)\right],$$

for m sufficiently large.

So the r.h.s. of (A.3) is bounded by

$$(b/2)(m/2\pi)^{d_2/2}\left|\Sigma^{(2)}\left(\xi_0^{(2)}\right)\right|^{1/2}\exp\left[m\phi^{(2)}\left(\xi_0^{(2)}\right)\right]\sum_{i=1}^{m-m_0}P_0\{S_i\notin mK\}.$$

Standard exponential Chebyshev inequalities show that $P_0\{S_i\notin mK\}$ is exponentially small with exponent depending on the size of K. Choosing K large enough yields the desired result. □

Acknowledgments. This work is based on a portion of the author's doctoral dissertation, which was written at Stanford University under the supervision of Professor David Siegmund. His guidance and encouragement are gratefully acknowledged. The author would also like to thank the Associate Editor and two referees for their helpful suggestions on exposition.

REFERENCES

BOROVKOV, A. A. and ROGOZIN, B. A. (1965). On the multidimensional central limit theorem. *Theory Probab. Appl.* **10** 55–62.

FERGUSON, T. (1967). *Mathematical Statistics: A Decision Theoretic Approach.* Academic, New York.

HU, I. (1985). Repeated significance tests for exponential families. Ph.D. dissertation, Stanford Univ.

LAI, T. L. and SIEGMUND, D. (1977). A nonlinear renewal theory with application to sequential analysis. I. *Ann. Statist.* **5** 946–954.

LALLEY, S. (1983). Repeated likelihood ratio tests for curved exponential families. *Z. Wahrsch. verw. Gebiete* **62** 293–321.

SIEGMUND, D. (1976). Importance sampling in the Monte Carlo study of sequential tests. *Ann. Statist.* **4** 673–684.

SIEGMUND, D. (1985). *Sequential Analysis. Tests and Confidence Intervals.* Springer, New York.

SIEGMUND, D. (1986). Boundary crossing probabilities and statistical applications. *Ann. Statist.* **14** 361–404.

WOODROOFE, M. (1977). Second order approximation for sequential point and interval estimation. *Ann. Statist.* **5** 984–995.

WOODROOFE, M. (1978). Large deviations of likelihood ratio statistics with application to sequential testing. *Ann. Statist.* **6** 72–84.

WOODROOFE, M. (1979). Repeated likelihood ratio tests. *Biometrika* **66** 453–463.

WOODROOFE, M. (1982). *Nonlinear Renewal Theory in Sequential Analysis.* SIAM, Philadelphia.

WOODROOFE, M. and TAKAHASHI, H. (1982). Asymptotic expansions for the error probabilities of some repeated significance tests. *Ann. Statist.* **10** 895–908.

DEPARTMENT OF STATISTICS
UNIVERSITY OF PENNSYLVANIA
PHILADELPHIA, PENNSYLVANIA 19104-6302