Qiwei Yao

Conditional boundary crossing probabilities and two-stage tests for a change-point

Article (Accepted version) (Refereed)

CONDITIONAL BOUNDARY CROSSING PROBABILITIES AND TWO-STAGE TESTS FOR A CHANGE-POINT

Qiwei Yao

University of Kent at Canterbury, UK

Running headline: Boundary Crossing Probabilities

Abstract

For normal random walks S_1, S_2, ..., a large deviation approximation is obtained for the probability of the event that the stochastic process $\{S_n\}$ crosses over the boundary $\{n(1-n/m)\}^{1/2}$ in the time interval $m_0 \le n \le m_1(< m)$ conditionally on $S_{m_0} = S_{m_1}$ being fixed. The result is applied to a change-point problem to approximate the significance level of the two-stage test, which is defined as a stochastic convex combination of the modified likelihood ratio test and Pettitt's test.

Key words: change-point, boundary crossing probability, large deviation, likelihood ratio test, Pettitt's test.

1 Introduction

One way to develop approximations for a boundary crossing probability is to write the probability as an expectation of a conditional boundary crossing probability given an appropriate random variable, and then to develop a large deviation approximation for the conditional probability. Such a method has been used with some degree of success, as measured by the accuracy in the cases that the sample sizes are moderate or even small, by Siegmund (1985, 1986) and James, James and Siegmund (1987, 1992) etc. It has also been adapted to produce some satisfactory results for boundary crossing problems of random fields (cf. Siegmund 1986, Yao 1993a,b, etc).

Let $Y_1,\ Y_2,\ \ldots$ be independent identically distributed $N(0,\sigma^2)$ random variables, and $S_n = Y_1 + \cdots + Y_n$. Using a likelihood ratio argument, Siegmund (1985, 1986) studied the asymptotic behavior of the conditional probabilities

$$P\{\, S_n \geq b_1\sqrt{n(1-n/m)} \text{ for some } m_0 \leq n \leq m_1 \,|\, S_{m_1} = my\,\},$$

where $b_1 > 0$ is a constant. The results were applied to approximate tail probabilities in some sequential tests and change-point problems. In this paper, we develop large deviation approximations for the conditional probabilities

$$P\{\, S_n \ \geq\ b_1\sqrt{n(1\ -\ n/m)} \text{ for some } m_0 \leq n \leq m_1 \mid S_{m_0} \ =\ S_{m_1} \ = my\,\}, \tag{1.1}$$

and apply the result to some change-point problems.

Our main result is stated and proved in §2. The proof uses the method of Woodroofe (1982) which splits the conditional probability into a sum in terms of the first crossing time, and then evaluates the approximation for each summand. The method of likelihood ratios or mixtures of likelihood ratios (cf. Siegmund 1985), which seems particularly simple for certain problems, appears to be difficult to be adapted to the present situation.

Our studying of the conditional probabilities (1.1) is motivated by the following observation on change-point problems. Since Page (1954) proposed the problem of detecting a parameter change in the context of quality control, there has been considerable studies on the test of the hypothesis that $Y_1, \ldots, Y_m$ are independent identically distributed random variables against the alternative that they are independent, but for some value ρ $(1 \leq \rho < m)$, $Y_1, \ldots, Y_\rho$ are identically distributed and $Y_{\rho+1}, \ldots, Y_m$ are also identically distributed but with a distribution different from that of Y_1. Various statistics, such as the Bayesian statistic (Chernoff and Zacks 1964), the recursive residual statistic (Brown, Durbin and Evans 1975), Pettitt's statistic (Pettitt 1980), and the modified likelihood ratio statistic (Siegmund 1986), were developed. James, James and Siegmund (1987) compared the different tests, and have shown that no method is overwhelmingly superior to the others. More specifically, when the change-point ρ is near 1 or m, the modified likelihood ratio test performs better than the others; when ρ is around $m/2$, Pettitt's test has the largest power. They remarked: "one possible conclusion is that one should choose a test statistic on a subjective basis, depending on where one 'expect' a change to take place, should there be one" (James, James and Siegmund 1987, p.82). Therefore, one possible alternative is to carry out a preliminary inference to identify the location of the possible change-point before performing a formal test with a properly chosen statistic. This strategy is particularly justified if there is *no prior information* available. In §3, some two-stage tests are defined as stochastic convex combinations of the (modified) likelihood ratio test and Pettitt's test. Comparisons of the power among the different tests are made by simulation. The approximation for the significance level of a new test is obtained by using Theorem 1 in §2 below. Some technical derivations are relegated to an appendix.

To simplify the presentation, we use $a_m \sim b_m$ to indicate that $a_m/b_m \to 1$ as $m \to \infty$, and

$$P\{X \in y + dx\} \equiv P\{X \in (y + x, y + x + dx]\} = f(y + x)dx,$$

where $f(.)$ is the density function of the random variable X.

2 Main Results

Throughout this section, it is always assumed that $Y_1, \ldots, Y_m$ are independent standard normal random variables, and $S_n = Y_1 + \cdots + Y_n$, $n = 1, \ldots, m$. Theorem 1 presents the large deviation approximations for the conditional boundary crossing probability (1.1), which is not only useful in estimating the significance level of a two-stage test in §3 below, but also of independent interest.

Theorem 1. Suppose that $m \to \infty$, $m_0 \to \infty$, and $b_1 \to \infty$ in such a way that $m_0/m = t_0 \in (0, 1/2)$, $b_1/\sqrt{m} = c > 0$ fixed. Let $m_1 = m - m_0$. Then as $m \to \infty$,

(i) uniformly for y in closed subintervals of $(\, ct_0\sqrt{t_0/(1-t_0)},\ 2ct_0(1-t_0)\,)$,

$$P\{\, S_n \ \geq\ b_1\sqrt{n(1\ -\ n/m)}\ \text{ for some } m_0 \leq n \leq m_1 \mid S_{m_0}\ =\ S_{m_1}\ = my\,\}$$

$$\sim \frac{1}{\sqrt{2}}\left[\frac{c-2y}{2ct_0(1-t_0)-y}\right]^{1/2}\nu\left(2\frac{c-2y}{1-2t_0}\right)\exp\{-\frac{m}{2(1-2t_0)}(c-2y)^2\};$$

(ii) uniformly for y in closed subintervals of $(\,2ct_0(1-t_0),\ c\sqrt{t_0(1-t_0)}\,)$,

$$P\{\,S_n \ \geq\ b_1\sqrt{n(1-n/m)} \text{ for some } m_0 \leq n \leq m_1 \mid S_{m_0} = S_{m_1} = my\,\}$$

$$\sim \frac{1-2t_0}{\sqrt{t_0(1-t_0)}}\,\nu\left(\frac{1-2t_0}{t_0(1-t_0)}y\right)\exp\{-\frac{m}{2}(1-2t_0)(c^2 - \frac{y^2}{t_0(1-t_0)})\}, \tag{2.1}$$

where

$$\nu(x) = 2x^{-2}\exp\{-2\sum_{k=0}^{\infty}k^{-1}\,\Phi(-\frac{1}{2}xk^{1/2})\} \qquad (x>0), \tag{2.2}$$

and $\Phi(\cdot)$ denotes the standard normal distribution function.

For numerical purposes, it is often sufficient to use the following approximation for the function $\nu(\cdot)$ given in (2.2) (see (4.38) of Siegmund 1985),

$$\nu(x) = \exp(-\,0.583x) + o(x^2).$$

To prove Theorem 1, the Woodroofe's method (cf. Chapter 8 of Woodroofe 1982) is adapted. To simplify the statement, we introduce some notations. For $t_0 = m_0/m \in (0,\ 1/2)$, $c = b_1/\sqrt{m} > 0$, and $y > 2ct_0(1-t_0)$, let

$$t_1^* = \frac{1}{2}(1-\sqrt{1-4c^2t_0^2(1-t_0)^2/y^2}\,), \qquad t_2^* = \frac{1}{2}(1+\sqrt{1-4c^2t_0^2(1-t_0)^2/y^2}\,),$$

$$\psi(t) = (c\sqrt{t(1-t)}-y)^2/[t(1-t)-t_0(1-t_0)], \quad t \in (0,1).$$

In fact, the principal contribution to the conditional probability (1.1) comes from the process S_n exceeding $b_1\sqrt{n(1-n/m)}$ for some n in a neighbourhood of $n = [m/2]$ when $y < 2ct_0(1-t_0)$ (see (i) — (iii) of Lemma 1 below), and in two symmetric neighbourhoods centred at t_1^* and t_2^* respectively when $y > 2ct_0(1-t_0)$ (see (iv) — (vi) of Lemma 1 below).

Lemma 1. Let A_1, and A_2 be closed subintervals of $(\,ct_0\sqrt{t_0/(1-t_0)},\ 2ct_0(1-t_0)\,)$, and $(\,2ct_0(1-t_0),\ c\sqrt{t_0(1-t_0)}\,)$ respectively. As $m \to \infty$,

(i) uniformly for $|n-m/2| \leq m^{7/12}$, $x \in [0,\ \log m]$, and $y \in A_1$,

$$P\{\,S_n \in b_1\sqrt{n(1-n/m)} + dx \mid S_{m_0} = S_{m_1} = my\,\}$$

$$\sim \sqrt{\frac{2\,m^{-1}}{\pi(1-2t_0)}}\exp\{-\frac{m}{2(1-2t_0)}(c-2y)^2\}\exp\{-\frac{2}{1-2t_0}(c-2y)x\}$$

$$\times\ \exp[-\frac{m}{2}(1-2t_0)\{\psi(n/m)-\psi(1/2)\}]\,dx;$$

(ii) uniformly for $|n - m/2| \leq m^{7/12}$, and $y \in A_1$,

$$P\{ S_n \ \geq \ b_1 \sqrt{n(1-n/m)} \ + \ \log m \,|\, S_{m_0} = S_{m_1} = my \}$$

$$= \ o\Big(m^{-1} \exp\{ -\frac{m}{2(1-2t_0)} \, (c-2y)^2 \} \Big) \ ;$$

(iii) uniformly for $|n - m/2| > m^{7/12}$, and $y \in A_1$,

$$P\{ S_n \ \geq \ b_1 \sqrt{n(1-n/m)} \,|\, S_{m_0} = S_{m_1} = my \} \ = \ o\Big(m^{-1} \exp\{ -\frac{m}{2(1-2t_0)} \, (c-2y)^2 \} \Big) \ ;$$

(iv) uniformly for $|n - mt_i^*| \leq m^{7/12}$ $(i = 1,\ 2)$, $x \in [0,\ \log m]$, and $y \in A_2$,

$$P\{ S_n \ \in \ b_1 \sqrt{n(1-n/m)} \ + \ dx \,|\, S_{m_0} = S_{m_1} = my \}$$

$$\sim \ \sqrt{\frac{1-2t_0}{2\pi m}} \, \frac{y t_0^{-1} (1-t_0)^{-1}}{\sqrt{c^2 - y^2/[t_0(1-t_0)]}} \ \exp\{ -\frac{m}{2} (1-2t_0)(c^2 - \frac{y^2}{t_0(1-t_0)}) \}$$

$$\times \ \exp\{ -\frac{1-2t_0}{t_0(1-t_0)} \, yx \} \ \exp[-\frac{m}{2} \, (1-2t_0) \{ \psi(n/m) - \psi(t_1^*) \}] \, dx;$$

(v) uniformly for $|n - mt_i^*| \leq m^{7/12}$ $(i = 1,\ 2)$, and $y \in A_2$,

$$P\{ S_n \ \geq \ b_1 \sqrt{n(1-n/m)} \ + \ \log m \,|\, S_{m_0} = S_{m_1} = my \}$$

$$= \ o\Big(m^{-1} \exp\{ -\frac{m}{2} (1-2t_0)(c^2 - \frac{y^2}{t_0(1-t_0)}) \} \Big) \ ;$$

(vi) uniformly for $|n - mt_i^*| > m^{7/12}$ $(i = 1,\ 2)$, and $y \in A_2$,

$$P\{ S_n \ \geq \ b_1 \sqrt{n(1-n/m)} \, , | \, S_{m_0} = S_{m_1} = my \}$$

$$= \ o\Big(m^{-1} \exp\{ -\frac{m}{2} (1-2t_0)(c^2 - \frac{y^2}{t_0(1-t_0)}) \} \Big) \ .$$

Lemma 1 follows from the fact that given $S_{m_0} = S_{m_1} = my$, the conditional distribution of S_n $(m_0 < n < m_1)$ is normal with mean my and variance $(n - m_0)[1 - (n - m_0)/(m - 2m_0)]$, and some standard estimates used, for example, in the proof of Lemma 1 of James, James and Siegmund (1988).

Proof of Theorem 1. We prove (ii) only. (i) can be shown in a similar but simpler way.

The following decomposition is obvious

$$P\{ \ S_n \ \geq \ b_1 \sqrt{n(1 \ - \ n/m)} \ \text{ for some } m_0 \leq n \leq m_1 \ | \ S_{m_0} \ = \ S_{m_1} \ = my \ \}$$

$$= \sum_{n=m_0}^{m_1} \int_0^\infty P\{ S_n \in b_1 \sqrt{n(1-n/m)} + dx \,|\, S_{m_0} = S_{m_1} = my \}$$

$$\times \;\; P\{ S_j < b_1\sqrt{j(1-j/m)} \text{ for all } m_0 \le j < n \,|\, S_{m_0} = my, \; S_n = b_1\sqrt{n(1-n/m)} + x \}$$

$$= \sum_{n\in\mathcal{A}_1} \int_0^{\log m} + \sum_{n\in\mathcal{A}_1} \int_{\log m}^\infty + \sum_{n\in\mathcal{A}_2} \int_0^\infty \;\equiv\; p_1(m) + p_2(m) + p_3(m), \quad \text{say}, \tag{2.3}$$

where $\mathcal{A}_1 = \{m_0 \le n \le m_1 : |n - mt_i^*| \le m^{7/12} \text{ for } i = 1 \text{ or } 2\}$, $\mathcal{A}_2 = \{m_0 \le n \le m_1 : n \notin \mathcal{A}_1\}$. By the symmetry in t_1^* and t_2^*,

$$p_1(m) = 2 \sum_{|n-mt_1^*|\le m^{7/12}} \int_0^{\log m} P\{S_n \in b_1\sqrt{n(1-n/m)} + dx | S_{m_0} = S_{m_1} = my\}$$

$$\times \;\; P\{S_j < b_1\sqrt{j(1-j/m)} \text{ for all } m_0 \le j < n | S_{m_0} = my, S_n = b_1\sqrt{n(1-n/m)} + x\}. \tag{2.4}$$

Along the same lines in the proof of Lemma 2 in Yao (1993b), it can be proved that given $S_{m_0} = my$ and $S_n = b_1\sqrt{n(1-n/m)} + x$, if $S_j \ge b_1\sqrt{j(1-j/m)}$ for some $m_0 \le j < n$, this event with overwhelming probability occurs for some j close to n, say $n - (\log m)^2 \le j \le n$. For such a j, and $|n - mt_1^*| \le m^{7/12}$,

$$b_1 \sqrt{n(1-n/m)} \; - \; b_1 \sqrt{j(1-j/m)} \; = \; (n-j) \, \frac{c(1-2t_1^*)}{2\sqrt{t_1^*(1-t_1^*)}} \; + \; o(1).$$

Hence,

$$P\{ S_j < b_1\sqrt{j(1-j/m)} \text{ for all } m_0 \le j < n | \, S_{m_0} = my, \; S_n = b_1\sqrt{n(1-n/m)} + x \}$$

$$= \;\; P\{ S_n - S_j - (n-j)\, c(1-2t_1^*) \,/\, [2\sqrt{t_1^*(1-t_1^*)}\,] \; > \; x \; \text{ for all } n - (\log m)^2 \le j \le n$$

$$| \, S_{m_0} = my, \; S_n = b_1\sqrt{n(1-n/m)} + x \} \; + \; o(1).$$

It can also be proved that for $n - (\log m)^2 \le j \le n$ and $|n - mt_1^*| \le m^{7/12}$, given $S_{m_0} = my, \; S_n = b_1\sqrt{n(1-n/m)} + x$, the process $S_n - S_j - (n-j)\, c(1-2t_1^*)/[2\sqrt{t_1^*(1-t_1^*)}\,]$ $(j = 1, 2, \ldots)$ behaves asymptotically like a normal random walk with the mean value $\mu^* \equiv y(1-2t_0)/[2t_0(1-t_0)]$. (See the brief proof of (9.87) in Siegmund (1985), and also the proof of Lemma 3 of Yao (1993b) for a more complex result.) Consequently, by Theorem 2.7 of Woodroofe (1982),

$$P\{ S_j < b_1\sqrt{j(1-j/m)} \text{ for all } m_0 \le j < n | \, S_{m_0} = my, \; S_n = b_1\sqrt{n(1-n/m)} + x \}$$

$$\to \; P\{\min_{n\ge1} V_k > x\} \; = \; \mu^* P\{V_{\tau_+} > x\}/E(V_{\tau_+}), \tag{2.5}$$

where $V_n = \sum_{i=1}^{n} Z_i$, $\{Z_i\}$ is a sequence of independent $N(\mu^*, 1)$ random variables, and $\tau_+ = \min\{n \geq 1 : V_n > 0\}$. By (8.46) of Siegmund (1985) and (2.10) of Siegmund (1986), it holds that

$$\frac{1}{E(V_{\tau_+})} \int_0^\infty e^{-2\mu^* x} P\{V_{\tau_+} > x\} dx = \frac{1}{2\mu^* E(V_{\tau_+})} \{1 - E \exp(-2\mu^* V_{\tau_+})\} = \nu(2\mu^*). \qquad (2.6)$$

By Lemma 1 (iv), relations (2.4), (2.5) and (2.6) imply that

$$\begin{aligned} p_1(m) &= 2\sqrt{\frac{1-2t_0}{2\pi}} \frac{y\, t_0^{-1}(1-t_0)^{-1}}{\sqrt{c^2 - y^2/[t_0(1-t_0)]}} \mu^* \nu(2\mu^*) \exp\{-\frac{m}{2}(1-2t_0)(c^2 - \frac{y^2}{t_0(1-t_0)})\} \\ &\times \sum_{|n-mt_1^*| \leq m^{7/12}} m^{-1/2} \exp[-\frac{m}{2}(1-2t_0)\{\psi(n/m) - \psi(t_1^*)\}]. \qquad (2.7) \end{aligned}$$

Let $\dot{\psi}(x) = d\psi(x)/dx$, and $\ddot{\psi}(x) = d^2\psi(x)/dx^2$. Note that $\dot{\psi}(t_1^*) = 0$. Thus, the sum in the above expression is equal to

$$\begin{aligned} & \sum_{|\frac{n}{\sqrt{m}} - \sqrt{m}t_1^*| \leq m^{1/12}} \frac{1}{\sqrt{m}} \exp\{-\frac{1}{4}(1-2t_0)\ddot{\psi}(t_1^*)(\frac{n}{\sqrt{m}} - \sqrt{m}t_1^*)^2 + O(m^{-1/4})\} \\ &= \int_{-m^{1/12}}^{m^{1/12}} \exp\{-\frac{1}{4}(1-2t_0)\ddot{\psi}(t_1^*)x^2 + O(m^{-1/4})\}dx + o(1) \\ &\to \int_{-\infty}^{\infty} \exp\{-\frac{1-2t_0}{2} \frac{y^4}{t_0^2(1-t_0)^2} \frac{x^2}{c^2 t_0(1-t_0) - y^2}\}dx \\ &= \sqrt{2\pi}\, t_0(1-t_0)\sqrt{c^2 t_0(1-t_0) - y^2} \,/\, [y^2\sqrt{1-2t_0}\,]. \end{aligned}$$

By substituting the RHS of the above expression into (2.7), we have proved that $p_1(m)$ is asymptotically equivalent to the RHS of (2.1). The proof is completed by establishing the relations

$$p_i(m) = o\left(\exp\{-\frac{m}{2}(1-2t_0)(c^2 - \frac{y^2}{t_0(1-t_0)})\}\right) \quad \text{for } i = 2, 3,$$

which follow from Lemma 1 (v) and (vi) immediately (see (2.3)).

3 Two-Stage Tests for a Change-Point

3.1 Two-stage tests

Let $Y_1, \ldots, Y_m$ be independent random variables with $Y_n \sim \mathrm{N}(\mu_n, 1)$, and $S_n = Y_1 + \cdots + Y_n, 1 \leq n \leq m$. A simple change-point problem is to test the following hypotheses

$$\begin{aligned} H_0 &: \mu_1 = \ldots = \mu_m = \mu; \\ H_1 &: \text{for some } 1 \leq \rho < m,\ \mu_1 = \ldots = \mu_\rho = \mu,\ \mu_{\rho+1} = \ldots = \mu_m = \mu + \delta, \qquad (3.1) \end{aligned}$$

where μ, $\delta > 0$ are the nuisance parameters. In what follows, P_0 denotes the probability measure under hypothesis H_0. For the hypotheses (3.1), the square root of the generalized log likelihood ratio statistic is calculated to be

$$U_1 = \max_{m_o \le n \le m_1} (n\, S_m/m - S_n)/\sqrt{n(1-n/m)} \tag{3.2}$$

for some $1 \le m_0 < m_1 < m$. Here we have generalized the statistic slightly to take the maximum over $m_0 \le n \le m_1$ instead of over $1 \le n < m$. This kind of generalization was suggested by Siegmund (1986) for the following reasons: it is intrinsically difficult to detect a change that occurs near 1 or m, and the likelihood ratio statistic (before the generalization) pays for its efforts to do so by giving up power near $\rho = [m/2]$. The introduction of m_0 and m_1 in the statistic gives the statistician the flexibility to give up a little power to detect changes occurring near the two endpoints in return for an increase in power near $m/2$. To simplify our discussion, we let $m_1 = m - m_0$.

Under H_0, the process $n\, S_m/m - S_n, n = 1, \ldots, m$, is independent of S_m. Hence the significance level of the likelihood ratio test with the statistic U_1 can be expressed as the conditional boundary crossing probability

$$\begin{aligned} P_0\{\, U_1 \ge b_1 \,\} &= P_0\{-S_n \ge b_1\sqrt{n(1-n/m)} \text{ for some } m_0 \le n \le m_1 \mid S_m = 0\} \\ &= P_0\{S_n \ge b_1\sqrt{n(1-n/m)} \text{ for some } m_0 \le n \le m_1 \mid S_m = 0\}, \end{aligned} \tag{3.3}$$

where $b_1 > 0$ is a constant. The second equality in (3.3) follows from the symmetry of $\{S_n;\ 1 \le n \le m\}$ under the measure $P_0(\,\cdot\,|S_m = 0)$. By Theorem 11.30 of Siegmund (1985), we have the following asymptotic approximation

$$P_0\{\, U_1 \ge b_1 \,\} \approx \frac{1}{2} b_1\, \varphi(b_1) \int_{m_0/m}^{m_1/m} \frac{1}{t(1-t)} \nu(b_1/\sqrt{mt(1-t)}\,)\, dt, \tag{3.4}$$

where φ denotes the standard normal density function, and $\nu(\cdot)$ is as given in (2.2).

On the other hand, the testing problem (3.1) is invariant under common shift in location of all the observations. This suggests that one should restrict consideration to invariant procedures, i.e. those which depend on Y's only through $X_n \equiv Y_n - Y_1$, $n = 2, \ldots, m$. For given ρ and δ, the log likelihood ratio of X's under H_1 relative to H_0 is easily calculated to be

$$\delta\,(\,\rho\, S_m/m \;-\; S_\rho\,) \;-\; \rho\,(1 \;-\; \rho/m)\,\delta^2\,/2.$$

By differentiating this statistic with respect to δ, setting $\delta = 0$ and then maximizing over ρ, we obtain the score-like statistic

$$U_2 \;=\; \max_{1 \le n < m} (n\, S_m/m \;-\; S_n) \tag{3.5}$$

for testing the hypotheses (3.1). This statistic was proposed originally by Pettitt (1980) for testing a change-point in zero-one observations, and used as a motivation in developing tests of the hypotheses (3.1) by James, James and Siegmund (1987).

Similar to (3.3), the significance level of U_2 can be expressed as

$$P_0(\, U_2 \ge b_2 \,) \;=\; P_0\{S_n \ge b_2 \text{ for some } 1 \le n \le m \;\mid\; S_m = 0\}, \tag{3.6}$$

where $b_2 > 0$ is a constant. By equation (10.43) of Siegmund (1985), we have the approximation

$$P_0(\, U_2 \ge b_2 \,) \;\approx\; \nu(4b_2/m)\, \exp\{-2b_2^2/m\}. \tag{3.7}$$

With (3.3) and (3.6), it is easy to see that in order to keep the two tests at the same significance level, b_2 must be less than $b_1\sqrt{n(1-n/m)}$ in a neighborhood of $n = [m/2]$ for m_0 near 1 and m_1 near m (cf. Fig. 1 of Siegmund 1986). Note that under H_1,

$$E\{nS_m/m - S_n\} = \begin{cases} n(1-\rho/m)\delta & 1 \le n \le \rho, \\ \rho(1-n/m)\delta & \rho < n \le m, \end{cases} \tag{3.8}$$

which attains the maximum at $n = \rho$. Hence it seems intuitively clear that the primary contribution to the power of the test comes from the probability that the process $nS_m/m - S_n$ exceeds the boundary for some n in a neighborhood of $n = \rho$. Consequently, we can expect that U_2 has greater power than U_1 when ρ is about $m/2$, whereas the converse is true for ρ near the endpoints 1 and m. Some numerical results have lent support to the above heuristic argument (cf. Table 1 of James, James and Siegmund 1987).

From the above discussion, we would desire a test which performs as U_1 when ρ is near 1 or m, and as U_2 when ρ is about $m/2$. Therefore, it is pertinent to consider the strategy of a two-stage test. In the first stage, we carry out a preliminary inference to identify the location of a possible change-point. Then we conduct the test for hypotheses (3.1) using either U_1 or U_2 accordingly. The resulted statistic will be a stochastic convex combination of statistics U_1 and U_2.

There are several ways to carry out the preliminary inference in the first stage. For example, we may carry out a preliminary test of the hypothesis that ρ is near either 1 or m against the alternative that ρ is around $m/2$. We define

$$Z_1 \;=\; S_m/m - S_{m_0}/m_0, \qquad Z_2 \;=\; (S_m - S_{m_1})/m_0 - S_m/m.$$

It is easy to see that under H_1

$$\begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix} \sim N\left(\begin{pmatrix} \delta(1-\frac{\rho}{m}) \\ \delta\frac{\rho}{m} \end{pmatrix}, \frac{m_1}{m_0 m} \begin{pmatrix} 1 & \frac{m_0}{m_1} \\ \frac{m_0}{m_1} & 1 \end{pmatrix} \right). \tag{3.9}$$

Hence, Z_1, or Z_2 can be approximately considered as a normal random variable with mean 0 in the case of ρ near m, or ρ near 1 respectively. This suggests the test statistic

$$U = I_{\{Z_1<h \text{ or } Z_2<h\}}(U_1 - b_1) + I_{\{Z_1 \geq h \text{ and } Z_2 \geq h\}}(U_2 - b_2),$$

and H_0 will be rejected if and only if $U \geq 0$. In the above expression, $h > 0$ is a constant. There is a drawback to formulate the preliminary inference as a test due to the asymmetry of statistical tests for hypotheses. As a compensation, we would choose h such that the two kinds of errors in the preliminary test are about the same. For example, we choose h such that

$$P\{Z_1 \geq h \,\text{and}\, Z_2 \geq h \,|\, \rho = m_0 \,\text{or}\, m_1\} = P\{Z_1 < h \text{ or } Z_2 < h \,|\, \rho = [m/2]\},$$

if the above equation has a positive solution. By (3.9), this implies that $h > 0$ should satisfy the equation

$$\begin{aligned} &\int_{\mathcal{B}_1(h)} \exp\left\{-\frac{m m_0 m_1}{2(m_1^2 - m_0^2)}(x_1^2 - 2\frac{m_0}{m_1}x_1x_2 + x_2^2)\right\} dx_1dx_2 \\ =& \frac{2\pi\sqrt{m_1^2 - m_0^2}}{m_0 m} - \int_{\mathcal{B}_2(h)} \exp\left\{-\frac{m m_0 m_1}{2(m_1^2 - m_0^2)}(x_1^2 - 2\frac{m_0}{m_1}x_1x_2 + x_2^2)\right\} dx_1dx_2, \end{aligned} \tag{3.10}$$

where

$$\begin{aligned} \mathcal{B}_1(h) &= \{x_1 \geq h - \delta(1 - m_0/m),\ x_2 \geq h - \delta m_0/m\}, \\ \mathcal{B}_2(h) &= \{x_1 \geq h - \delta(1 - [0.5m]/m),\ x_2 \geq h - \delta[0.5m]/m\}. \end{aligned}$$

In practice, we use

$$\hat{\delta} = (S_m - S_{\hat{\rho}})/(m - \hat{\rho}) \tag{3.11}$$

instead of unknown δ in the above expressions, where $\hat{\rho}$ is the maximum likelihood estimate of ρ under H_1, i.e.

$$\hat{\rho} = \underset{1 \leq n < m}{\text{argmax}} \,(nS_m/m - S_n)/\sqrt{n(1 - n/m)}. \tag{3.12}$$

Remark 1. For a given test level, there is one degree of freedom in the choice of (b_1, b_2) in the two-stage tests. In practice, we can use the values of b_1 and b_2 determined by (3.4) and (3.7) respectively as the starting values, then adjust them further using (3.14) below.

Remark 2. It is arguable whether the two-stage test statistic U is a good choice. Intuitively, the preliminary test defined above is not very effective in identifying the location of the possible change-point. Two alternative statistics will be given in §3.4 below. We focus on the statistic U partially because with U it is feasible to pursue the asymptotic approximations, in the form of large deviations, for the significance levels of the test.

3.2 An approximation for the test level

Based on the Theorem 1 in §2, it is possible to derive an analytic approximation for the significance level of the test U. Obviously the significance level can be written as

$$P_0(U \geq 0) = P_0(U_1 \geq b_1,\ Z_1 < h \text{ or } Z_2 < h) + P_0(U_2 \geq b_2,\ Z_1 \geq h,\ Z_2 \geq h). \tag{3.13}$$

From Propositions 1, and 2 below, we have the following approximation

$$\begin{aligned} P_0(U \geq 0) &\approx 2b_1\,\varphi(b_1)\int_{\xi_1}^{\xi_0} x^{-1}\nu[x/t_0 + b_1^2 t_0/(xm)]\,dx \\ &+ \frac{m}{t_0\sqrt{1-2t_0}}\varphi(\frac{2b_2}{\sqrt{m-2m_0}})\int_{\xi_0}^{\zeta}\int_{\xi_0}^{\zeta}\nu[\frac{2(2\zeta-y-z)}{1-2t_0}] \\ &\times \exp[-\frac{m}{1-2t_0}\{\frac{1-t_0}{2t_0}(y^2+z^2)+yz-2\zeta(y+z)\}]\,dy\,dz, \end{aligned} \tag{3.14}$$

where $t_0 = m_0/m,\ \zeta = b_2/m,\ \xi_0 = t_0 h,\ \xi_1 = b_1 m^{-1/2} t_0\sqrt{t_0/(1-t_0)}$.

To assess the accuracy of (3.14), we conduct some Monte Carlo experiments with 20000 replications for m=30, 40 and 50 and h=0.15 and 0.25. The results are recorded in Table 1. These figures along with similar ones not reported here show that when h is small, the approximation (3.14) offers a reasonably good approximation to the significance level.

(Table 1 is about here)

To justify the approximation (3.14), the following two propositions are established based on Theorem 1 and some other large deviation approximations for boundary crossing probabilities in Siegmund (1985, 1986). Their proofs are given in the appendix.

Similar to (3.3), the two terms on the RHS of (3.13) can be expressed as follows

$$\begin{aligned} \alpha_1(m) &\equiv P_0(U_1 \geq b_1,\ Z_1 < h \text{ or } Z_2 < h) \\ &= P_0^{(m)}\{\,S_{m_0} < m\xi_0 \text{ or } S_{m_1} < m\xi_0,\ \max_{m_0\leq n\leq m_1} S_n/\sqrt{n(1-n/m)} \geq b_1\,\} \end{aligned}$$

$$
\begin{aligned}
&= 2\,P_0^{(m)}\{\, S_{m_1} < m\xi_0,\ \max_{m_0\le n\le m_1} S_n/\sqrt{n(1-n/m)} \ge b_1 \,\} \\
&- P_0^{(m)}\{\, S_{m_0} < m\xi_0,\ S_{m_1} < m\xi_0,\ \max_{m_0\le n\le m_1} S_n/\sqrt{n(1-n/m)} \ge b_1 \,\}, \qquad (3.15)
\end{aligned}
$$

$$
\begin{aligned}
\alpha_2(m) &\equiv P_0(U_2 \ge b_2,\ Z_1 \ge h,\ Z_2 \ge h) \\
&= P_0^{(m)}\{\, S_{m_0} \ge m\xi_0,\ S_{m_1} \ge m\xi_0,\ \max_{1\le n\le m} S_n \ge b_2 \,\}, \qquad (3.16)
\end{aligned}
$$

where $P_0^{(m)}(\cdot)$ denotes the conditional probability measure $P_0(\,\cdot\,|S_m = 0)$.

Proposition 1. Suppose that $m \to \infty$, $m_0 \to \infty$, $b_1 \to \infty$ in such a way that $m_0/m = t_0 \in (0, 1/2)$, $b_1/\sqrt{m} = c > 0$ fixed. Let $m_1 = m - m_0$, $\xi_1 = ct_0\sqrt{t_0/(1-t_0)}$. Then for

$$
\xi_0 \in \left(\xi_1,\ \frac{2ct_0(1-t_0)[1-(1-2t_0)/\sqrt{2(1-2t_0)}\,]}{(1+2t_0-4t_0^2)}\right), \qquad (3.17)
$$

$$
\alpha_1(m) \sim 2b_1\varphi(b_1)\int_{\xi_1}^{\xi_0} y^{-1}\nu(y/t_0 + c^2t_0/y)\,dy,
$$

where $\varphi(\cdot)$ is the standard normal density function, and $\nu(\cdot)$ is given in (2.2).

Proposition 2. Suppose that $m \to \infty$, $m_0 \to \infty$, and $b_2 \to \infty$ in such a way that $m_0/m = t_0 \in (0, 1/4)$, and $b_2/m = \zeta > 0$ fixed. Let $m_1 = m - m_0$. Then for $\xi_0 \in (0,\ \zeta)$,

$$
\begin{aligned}
\alpha_2(m) \sim\ & \frac{m}{t_0\sqrt{1-2t_0}}\,\varphi\!\left(\frac{2\zeta\sqrt{m}}{\sqrt{1-2t_0}}\right)\int_{\xi_0}^{\zeta}\!\int_{\xi_0}^{\zeta} \nu\left(\frac{2(2\zeta-y-x)}{1-2t_0}\right) \\
& \times\ \exp[-\frac{m}{1-2t_0}\{\frac{1-t_0}{2t_0}(x^2+y^2) + xy - 2\zeta(x+y)\}]\,dx\,dy,
\end{aligned}
$$

where $\varphi(\cdot)$ is the standard normal density function, $\nu(\cdot)$ is given in (2.2).

3.3 Numerical comparisons

A Monte Carlo experiment based on 20000 replications is conducted to compare the power of the two-stage test U, the modified likelihood ratio test U_1 (LRT), and Pettitt's test U_2. In each case the significance level is 0.025, m=40, m_0=5, and δ is set at three different values: (a) δ=0.4, (b) δ=0.8, (c) δ=1.2. For the two-stage test, we solve equation (3.10) to obtain $h = 0.122$, and 0.264 for $\delta = 0.8$, and 1.2 respectively. For $\delta = 0.4$, (3.10) does not have a positive solution, and we arbitrarily choose $h = 0.1$. The critical values were determined by (3.14), (3.4) and (3.7) first. They have been further adjusted such that the differences among the levels of the three tests in the simulation are not greater than 0.0005. The results are plotted in Figure 1. Note that due to the symmetry in the hypotheses (3.1) and the statistics U_1 and U_2, the power functions are symmetric in ρ and $m-\rho$ when $m_1 = m - m_0$. Thus, only the results with $1 \le \rho \le 20$ are reported. The values corresponding to $\rho = 0$ are the levels of the tests. These curves illustrate that the power of the two-stage test is almost always between the powers of the two other tests. As we pointed out earlier, the two-stage test is appealing only when there is no prior information available on the location of the possible change-point. Other simulation results, not reported here, show that the essential pattern is unchanged over a range of significance levels and sample sizes, although the magnitude of the difference can be more or less.

(Figure 1 is about here.)

3.4 Two alternatives

In order to increase the power of the two-stage test, we may consider some other forms of the preliminary inference which are able to identify the location of the possible change point more efficiently. Further, we may use the genuine likelihood ratio statistic

$$U_1^* = \max_{1 \le n < m} (nS_m/m - S_n)/\sqrt{n(1-n/m)}$$

instead of the modified likelihood ratio statistic U_1.

We consider a test as follows: First, we evaluate the maximum likelihood estimator of ρ. When the estimator is near either 1 or m, we use the likelihood ratio test U_1^*. When the estimator is near $m/2$, we use Pettitt's statistic U_2. This leads to the test statistic

$$T_1 = I_{\{\hat\rho < l \text{ or } \hat\rho > m-l\}}(U_1^* - b_1) + I_{\{l \le \hat\rho \le m-l\}}(U_2 - b_2),$$

where $1 < l < m/2$ is an integer and $\hat\rho$ is given as in (3.12). We reject H_0 if and only if $T_1 \ge 0$. Ideally we would choose l in such a way that U_2 has greater power than U_1^* if and only if $l \le \rho \le m-l$. To evaluate the value of l, we may use the analytic approximations of the power functions with given δ for both tests U_1^* and U_2, which were developed by James, James and Siegmund (1987). A Monte Carlo method can also be used to solve l. In practice, we may use $\hat\delta$ given as in (3.11) instead of unknown δ.

Another alternative is motivated by the following observation. From (3.8), we can see that under hypothesis H_1, the expectation of $(nS_m/m - S_n)$ obtains its maximum at $n = \rho$, and is monotonically decrease when n spreads away from ρ to both sides. This suggests that we can simply compare the values of $(nS_m/m - S_n)$ at typical values of n to determine where ρ is about. For example, we may define the test statistic

$$\begin{aligned} T_2 &= I_{\{\frac{1}{2}S_m - S_{[m/2]} + d < (\frac{m_0}{m}S_m - S_{m_0}) \vee (\frac{m_1}{m}S_m - S_{m_1})\}}(U_1^* - b_1) \\ &+ I_{\{\frac{1}{2}S_m - S_{[m/2]} + d \geq (\frac{m_0}{m}S_m - S_{m_0}) \vee (\frac{m_1}{m}S_m - S_{m_1})\}}(U_2 - b_2), \end{aligned}$$

where $d \in R$ is a constant, $m_0 \geq 1$ is a small integer, and $m_1 = m - m_0$. We reject H_0 if and only if $T_2 \geq 0$. The idea behind the above statistic is that when ρ is near $m/2$, it is likely to occur that $\frac{1}{2}S_m - S_{[m/2]} + d \geq (\frac{m_0}{m}S_m - S_{m_0}) \vee (\frac{m_1}{m}S_m - S_{m_1})$. We would choose d in such a way that under H_1 $E(S_m - S_{[m/2]}) + d \geq E(\frac{m_0}{m}S_m - S_{m_0}) \vee E(\frac{m_1}{m}S_m - S_{m_1})$ when $\rho \in [l, m-l]$ in which U_2 has greater power than U_1. Therefore, d should satisfy the condition

$$\left\{m_0\left(1 - \frac{l}{m}\right) - \frac{[0.5m]}{m}\right\}\delta \leq d < \left\{m_0\left(1 - \frac{l-1}{m}\right) - \frac{[0.5m]}{m}\right\}\delta.$$

In practice, we use $\hat{\delta}$ given in (3.11) instead of δ in the above expression.

Intuitively, the tests T_1 and T_2 seem more powerful in identifying the location of the possible change-point in the first stage than the test U. We do not pursue the discussion on T_1 and T_2 further in this paper since, within our knowledge, it seems formidable to develop the large deviation approximations for the levels of the tests.

Appendix: Proofs of Propositions 1, and 2

We use the same notation as in Section 3. Further, $P_x^{(m)}(\cdot)$ denotes the conditional probability measure $P_0(\cdot\,|S_m = x)$.

Proof of Proposition 1. By Theorem 3.11 of Siegmund (1986), it holds that for any $y \in (\xi_1,\ c\sqrt{t_0(1-t_0)})$

$$\begin{aligned} &P_{my}^{(m_1)}\{\ S_n \geq b_1\sqrt{n(1-n/m)} \text{ for some } m_0 \leq n \leq m_1\ \} \\ \sim\ &\sqrt{t_0(1-t_0)}\,\frac{c}{y}\,\nu\left(\frac{t_0c^2}{y} + \frac{y}{t_0}\right)\exp\{-\frac{m}{2}(c^2 - \frac{y^2}{t_0(1-t_0)})\}; \end{aligned} \tag{A.1}$$

In fact, the above convergence is uniform for y in any closed subintervals of $(\xi_1,\ c\sqrt{t_0(1-t_0)})$ (cf. Appendix 1 of Siegmund 1986, and also the proof of Theorem 1 in §2). In the similar but much simplier way, it can be proved that uniformly for y in closed subintervals of $(-\infty, \xi_1)$

$$P_{my}^{(m_1)}\{\max_{m_0\leq n\leq m_1} S_n/\sqrt{n(1-n/m)} \geq b_1\} = o\left(\exp\{-\frac{m}{2}(c^2 - \frac{y^2}{t_0(1-t_0)}\}\right). \tag{A.2}$$

Note that given S_{m_1} fixed, the process $\{S_n, 1 \le n \le m_1\}$ is independent of S_m. Therefore, we have the following decomposition

$$
\begin{aligned}
& P_0^{(m)}\{\, S_{m_1} < m\xi_0,\ S_n \ge b_1\sqrt{n(1-n/m)} \text{ for some } m_0 \le n \le m_1 \,\} \\
= & \int_{-\infty}^{\xi_1} P_0^{(m)}(\, S_{m_1} \in m\, dy\,)\, P_{my}^{(m_1)}\{\, S_n \ge b_1\sqrt{n(1-n/m)} \text{ for some } m_0 \le n \le m_1 \,\} \\
+ & \int_{\xi_1}^{\xi_0} P_0^{(m)}(\, S_{m_1} \in m\, dy\,)\, P_{my}^{(m_1)}\{\, S_n \ge b_1\sqrt{n(1-n/m)} \text{ for some } m_0 \le n \le m_1 \,\} \\
\equiv & \ \alpha_{11}(m) + \alpha_{12}(m), \quad \text{say.}
\end{aligned} \tag{A.3}
$$

It follows from (A.1) that for any $\varepsilon > 0$

$$
\begin{aligned}
& \int_{\xi_1(1+\varepsilon)}^{\xi_0} P_0^{(m)}(\, S_{m_1} \in m\, dy\,)\, P_{my}^{(m_1)}\{\, S_n \ge b_1\sqrt{n(1-n/m)} \text{ for some } m_0 \le n \le m_1 \,\} \\
\sim & \int_{\xi_1(1+\varepsilon)}^{\xi_0} \sqrt{m}\varphi\left(\sqrt{\frac{m}{t_0(1-t_0)}}y\right) \frac{c}{y}\nu\left(\frac{t_0c^2}{y} + \frac{y}{t_0}\right) \exp\{-\frac{m}{2}(c^2 - \frac{y^2}{t_0(1-t_0)})\}dy \\
= & \ b_1\varphi(b_1)\int_{\xi_1(1+\varepsilon)}^{\xi_0} y^{-1}\nu\left(\frac{t_0c^2}{y} + \frac{y}{t_0}\right) dy \equiv b_1\varphi(b_1)M(\varepsilon), \text{ say.}
\end{aligned} \tag{A.4}
$$

Write the integration on the LHS of the above expression as $\alpha_{12}(m,\varepsilon)$. We have proved that $b_1^{-1}\alpha_{12}(m,\varepsilon)/\varphi(b_1) \to M(\varepsilon)$ for any $\varepsilon > 0$. This implies that for any $\varepsilon' > 0$, there exists $n = n(\varepsilon, \varepsilon')$ for which

$$
b_1^{-1}\alpha_{12}(m,\varepsilon)/\varphi(b_1) \ge M(\varepsilon) - \varepsilon', \quad \text{for all } m \ge n.
$$

Consequently,

$$
b_1^{-1}\alpha_{12}(m)/\varphi(b_1) \ge b_1^{-1}\alpha_{12}(m,\varepsilon)/\varphi(b_1) \ge M(\varepsilon) - \varepsilon', \quad \text{for all } \varepsilon > 0 \text{ and } \ m \ge n(\varepsilon, \varepsilon').
$$

Let $\varepsilon \to 0$ in the above expression, we have $\text{limmin}_{m\to\infty}\, b_1^{-1}\alpha_{12}(m)/\varphi(b_1) \ge M(0) - \varepsilon'$ for any $\varepsilon' > 0$. Thus, $\text{limmin}_{m\to\infty}\, b_1^{-1}\alpha_{12}(m)/\varphi(b_1) \ge M(0)$.

Note that the large values of S_{m_1} are in favour to the event that $S_n \ge b_1\{n(1-n/m)\}^{1/2}$ for some $m_0 \le n \le m_1$. Therefore, it holds that for any $\varepsilon > 0$

$$
\alpha_{12}(m) \le \int_{\xi_1}^{\xi_0} P_0^{(m)}(\, S_{m_1} \in m\, dy\,)\, P_{m(y+\varepsilon)}^{(m_1)}\{\, S_n \ge b_1\sqrt{n(1-n/m)} \text{ for some } m_0 \le n \le m_1 \,\}.
$$

Applying (A.1) to the RHS of the above expression, we can prove that $\text{limmax}_{m\to\infty}\, b_1^{-1}\alpha_{12}(m)/\varphi(b_1)$ $\le M(0)$. Therefore, $b_1^{-1}\alpha_{12}(m)/\varphi(b_1) \sim M(0)$.

By repeating the above argument with (A.2) instead of (A.1), it can be shown that $\alpha_{11}(m) = o(b_1\varphi(b_1))$. Consequently from (A.3), we have that

$$P_0^{(m)}\{S_{m_1} < m\xi_0,\ S_n \ge b_1\sqrt{n(1-n/m)} \text{ for some } m_0 \le n \le m_1\}$$

$$\sim \quad \alpha_{12}(m) \ \sim \ b_1\varphi(b_1)\int_{\xi_1}^{\xi_0} y^{-1}\nu(y/t_0 + c^2 t_0/y)dy. \qquad \text{(A.5)}$$

On the other hand, we have that

$$P_0^{(m)}(\,S_{m_0} < m\xi_0,\ S_{m_1} < m\xi_0,\ S_n \ge b_1\sqrt{n(1-n/m)} \text{ for some } m_0 \le n \le m_1\,\}$$

$$= \quad \int_{\xi_1}^{\xi_0}\int_{\xi_1}^{\xi_0} P\left(\,S_n \ge b_1\sqrt{n(1-n/m)} \text{ for some } m_0 \le n \le m_1 \,|\, S_{m_0} = mx,\ S_{m_1} = my\right)$$

$$\times \quad P_0^{(m)}(\,S_{m_0} \in m\,dx,\ S_{m_1} \in m\,dy\,) \ + \ o\,(b_1\varphi(b_1))\,. \qquad \text{(A.6)}$$

Being symmetrical in x and y, the integral in the above equation is equal to twice the integral on the triangle $\{(x,\ y) :\ \xi_1 \le x \le y,\ \xi_1 \le y \le \xi_0\}$. Note that the large values of S_{m_0} are in favour to the event that the process $\{S_n\}$ crosses over the boundary. Therefore, for $x < y$, the conditional probability in the integrand of (A.6) is less than

$$P\{\,S_n \ge b_1\sqrt{n(1-n/m)} \text{ for some } m_0 \le n \le m_1 \,|\, S_{m_0} = S_{m_1} = my\,\},$$

Theorem 1 (i) implies that for all sufficiently large m, the RHS of (A.6) is less than

$$\frac{m}{2\pi t_0\sqrt{2(1-2t_0)}}\int_{\xi_1}^{\xi_0} dy\int_{\xi_1}^{y}\sqrt{\frac{c-2y}{2ct_0(1-t_0)-y}}\,\nu\left(2\frac{c-2y}{1-2t_0}\right)$$

$$\times\ \exp\{-\frac{m(1-t_0)}{2t_0(1-2t_0)}\,(x-\frac{t_0}{1-t_0}y)^2 - \frac{my^2}{2t_0(1-t_0)} - \frac{m(c-2y)^2}{2(1-2t_0)}\}\,dx. \qquad \text{(A.7)}$$

Some algebraic calculation entails that

$$\frac{my^2}{2t_0(1-t_0)} \ + \ \frac{m}{2(1-2t_0)}\,(c-2y)^2 \ = \ \frac{m}{2}c^2 \ + \ \frac{m}{1-2t_0}\,w(y),$$

where $w(y)$ is a second order polynomial, which is positive for all $y \le \xi_0$ under the condition (3.17). Hence it is easy to see that (A.7) tends to 0 faster than $m^{1/2}\exp\{-mc^2/2\}$ as $m \to \infty$. Therefore the probability on the LHS of (A.6) is $o(b_1\varphi(b_1))$. The conclusion then follows from (3.15) and (A.5).

Proof of Proposition 2. Note that given $S_m = 0$, S_{m_0} and S_{m_1} are jointly normal. It follows from some elementary calculation that for any $\varepsilon > 0$,

$$P_0^{(m)}\{\, S_{m_0} \geq m(\zeta - \varepsilon),\ S_{m_1} \geq m(\zeta - \varepsilon),\, \} = O(m^{-1} \exp\{-m t_0^{-1} (\zeta - \varepsilon)^2\}). \qquad \text{(A.8)}$$

On the other hand,

$$P_0^{(m)}\{\, m\xi_0 < S_{m_0} \leq m(\zeta - \varepsilon),\ m\xi_0 < S_{m_1} \leq m(\zeta - \varepsilon),\ \max_{1 \leq n \leq m} S_n \geq b_2 \,\}$$
$$= \int_{\xi_0}^{\zeta-\varepsilon} \int_{\xi_0}^{\zeta-\varepsilon} P_0^{(m)}\{S_{m_0} \in m\,dx,\ S_{m_1} \in m\,dy\}$$
$$\times\ P_0^{(m)}\{\max_{1 \leq n \leq m} S_n \geq b_2 \,|\, S_{m_0} = mx,\ S_{m_1} = my\}. \qquad \text{(A.9)}$$

It is easy to see that the principal contribution to the conditional probability in the above integrand comes from S_n exceeding b_2 with n between m_0 and m_1 since $t_0 < 1/4$. Hence by (8.78) of Siegmund (1985), the RHS of (A.9) is asymptotically equivalent to

$$\frac{m}{t_0\sqrt{1 - 2t_0}}\, \varphi\!\left(\frac{2\zeta\sqrt{m}}{\sqrt{1 - 2t_0}}\right) \int_{\xi_0}^{\zeta-\varepsilon} \int_{\xi_0}^{\zeta-\varepsilon} \nu(2(2\zeta - y - x)/(1 - 2t_0))$$
$$\times\ \exp[-\frac{m}{1 - 2t_0}\{\frac{1 - t_0}{2t_0}(x^2 + y^2) + xy - 2\zeta(x + y)\}]\, dx\, dy. \qquad \text{(A.10)}$$

It is easy to prove that on $\xi_0 < x,\, y \leq d$ with $d > 2t_0\zeta$, the function $f(x,\, y) \equiv (1 - t_0)(x^2 + y^2)/(2t_0) + xy - 2\zeta(x + y)$ attains the maximum $f(d,\, d)$ at $x = y = d$. This fact implies that the primary contribution to the integral in (A.10) comes from the values of the both x and y near ξ_0. Therefore this integral can be taken over $\xi_0 < x,\, y < \zeta$, and the probability (A.8) can be neglected. Consequently, the proposition follows from (3.16) and (A.8) – (A.10).

Acknowledgements

The author wishes to thank the Editor and the three anonymous referees for many helpful comments and suggestions which led to a considerable improvement of the paper from its earlier version. Especially, the test statistics T_1 and T_2 discussed in §3 were suggested by two referees.

References

Brown, R. L., Durbin, J. & Evans, J. M. (1975). Techniques for testing the constancy of regression relationships over time. *J. R. Statist. Soc.* B 37, 149-192.

Chernoff, H. & Zacks, S. (1964). Estimating the current mean of a normal distribution which is subject to change in time. *Ann. Math. Statist.* 35, 999-1028.

James, B., James, K. L. & Siegmund, D. (1987). Tests for a change-point. *Biometrika* 74, 71-83.

James, B., James, K. L. & Siegmund, D. (1988). Conditional boundary crossing probabilities, with applications to change-point problems. *Ann. Probab.* 16, 825-839.

James, B., James, K. L. & Siegmund, D. (1992). Asymptotic approximations for likelihood ratio tests and confidence regions for a change-point in the mean of a multivariate normal distribution. *Statistica Sinica* **2**, 69-90.

Page, E. S. (1954). Continuous inspection schemes. *Biometrika* 41, 100-115.

Pettitt, A. N. (1980). A simple cumulative sum type statistic for the change-point problem with zero-one observations. *Biometrika* 67, 79-84.

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

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

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

Yao, Q. (1993a) Tests for change-points with epidemic alternatives. *Biometrika*, 80, 179-191.

Yao, Q. (1993b) Boundary crossing probabilities of some random fields related to likelihood ratio tests for epidemic alternatives. *J. Appl. Probab.*, 30, 52-65.

Qiwei Yao, Institute of Mathematics and Statistics, University of Kent at Canterbury, Canterbury, Kent CT2 7NF, UK.

Table 1. Accuracy of the approximation (3.14)

($m_0 = m/10$, $m_1 = m - m_0$)

m	h	(b_1, b_2)	(3.14)	Monte Carlo
		(2.07, 7.01)	0.050	0.071
30	0.15	(1.82, 6.50)	0.086	0.114
		(1.71, 6.34)	0.175	0.201
		(2.07, 7.01)	0.107	0.115
30	0.25	(1.82, 6.50)	0.182	0.186
		(1.71, 6.34)	0.270	0.228
		(2.67, 7.88)	0.023	0.028
40	0.15	(2.21, 7.22)	0.056	0.081
		(1.94, 6.82)	0.146	0.152
		(2.67, 7.88)	0.036	0.033
40	0.25	(2.21, 7.22)	0.103	0.102
		(1.94, 6.82)	0.204	0.185
		(2.33, 7.91)	0.078	0.071
50	0.15	(2.13, 7.41)	0.121	0.115
		(1.92, 7.13)	0.180	0.176
		(2.33, 7.91)	0.108	0.093
50	0.25	(2.13, 7.41)	0.159	0.143
		(1.92, 7.13)	0.225	0.241

Figure Legend

Figure 1. The plots of the estimated power functions against the change-point form a Monte Carlo experiment with 20000 replications with $m = 40$ and $m_0 = 5$. Diamond solid curve — test U; Plus dashed curve — test U_1 ($b_1 = 2.832$); Square shorter-dashed curve — test U_2 ($b_2 = 8.022$). (a) $\delta = 0.4$. ($h = 0.1$ and $(b_1, b_2) = (2.825, 8.015)$ for the test U); (b) $\delta = 0.8$. ($h = 0.122$ and $(b_1, b_2) = (2.825, 8.017)$ the test U); (c) $\delta = 1.2$. ($h = 0.264$ and $(b_1, b_2) = (2.766, 8.017)$ for the test U).