SEQUENTIAL ANALYSIS, 12(3&4), 201-210 (1993)

ASYMPTOTICALLY OPTIMAL DETECTION OF A CHANGE IN A LINEAR MODEL

Qiwei Yao*

Department of Mathematics, Southeast University, Nanjing 210018, China.

Key words and Phrases: stopping time; detection of a change; Cusum procedure; one-sided sequential probability ratio test.

ABSTRACT

For a change in the mean value parameters of a normal linear model, a class of detecting methods is proposed, which are asymptotically optimal in an appropriate sense. If there exists no nuisance parameters, the Cusum procedure is included in this class.

1. INTRODUCTION AND MAIN RESULT

Suppose a process produces a potentially infinite sequence of independent observations $y_1, y_2, \ldots$. Initially the process is 'under control' in the sense that the effects are the same. At some unknown time m the process may change and the effects become 'out of control'. The observer would like to infer from the y's that this change has taken place as soon as possible. Of course the rate of false alarm should be kept low.

Most of the literature on this problem considers the case that the observations before the change and after the change are identically distributed. This means that $y_1, \ldots, y_{m-1}$ have density function f_0 while $y_m, y_{m+1}, \ldots$ have density $f_\theta \neq f_0$, and f_0 is known. For this setup, Page (1954) proposed the *Cusum* procedure. Lorden (1971) proved an asymptotic optimality property of the Cusum procedure. He formulated the problem as follows: Find the

*Present address: Inst. of Math. & Statist., Univ. of Kent, Canterbury, Kent CT2 7NF, UK.

stopping time N which minimizes $E\{(N-m+1)^+ | y_1, \ldots, y_{m-1};$ the change occurs at $m\}$ under the restriction that $E\{N \,|\, \text{no change occurs}\} \geq \gamma$. Lorden considered the asymptotic case as $\gamma \to \infty$. When f_θ is known, the class of asymptotically optimal procedures includes the Cusum procedure. When f_θ is unknown, however f_0 and f_θ can be embedded into an exponential family of distributions, Lorden (1971) gives a feasible asymptotically optimal procedure. Recently, Moustakides (1986) showed that the Cusum procedure is also optimal for a finite γ. Other research on the problem has been done by Shiryayev (1963) and Roberts (1966). Pollak (1985) showed that the Shiryayev-Roberts procedure is asymptotically optimal as $\gamma \to \infty$ under the modified criteria that $\sup_m E\{N-m \,|\, N > m;$ the change occurs at $m\} = \min$. Pollak and Siegmund (1985) showed that the difference in the speed of detection between the Cusum and the Shiryayev-Roberts procedures is small.

In this article, a non-i.i.d. situation is considered: the observations y's are from a normal linear model. Similar results to Lorden (1971) will be shown for this case. Since covariants arise very often in some practical situations, this model has wide applications in econometrics (Quandt 1960), industrial reliability (Worsley 1983), and general regression prediction (Hinkley 1971) etc..

For a given sequence of $p \times 1$ explanatory vectors $x_1, x_2, \ldots$, we assume that the observations y's satisfy the equations

$$\begin{cases} y_i = x_i'\beta + e_i, & \text{for } i = 1, \ldots, m-1; \\ y_i = x_i'(\beta + \theta) + e_i, & \text{for } i = m, m+1, \ldots, \end{cases} \tag{1}$$

where $e_1, e_2, \ldots$ are i.i.d. N(0, 1) variables; β, θ are p-dimensional column vectors; β is known, $\theta \neq 0$ and m are unknown; x_i' denotes the transpose of x_i. Let $P_{m,\theta}$ denote the corresponding probability measure and $E_{m,\theta}$ the expectation under this measure. We allow m to take the value infinite to indicate the case of no change. We simply write P_∞ and E_∞. We also write $Y_n = (y_1, \ldots, y_n)'$, and $X_n = (x_1, \ldots, x_n)'$.

As a possible detecting procedure, we consider a stopping time N with respect to the observed sequence $\{y_n\}$. Thus the event $\{N = n\}$ is determined by $y_1, \ldots, y_n$ (i.e. belongs to the sigma-field generated by $y_1, \ldots, y_n$). N may take the value infinity. If N is finite almost everywhere, we call it a stopping rule. The optimality problem we put as Lorden (1971) did. That is, if N is a stopping time, define

$$D_{m,\theta}(N) = \operatorname{ess\,sup} E_{m,\theta}\{\, (N-m+1)^+ \,|\, Y_{m-1} \,\},$$

$$D_\theta(N) = \sup_{m \geq 1} D_{m,\theta}(N).$$

We want to minimize $D_\theta(N)$ over all stopping times which satisfy the constraint:

$$E_\infty N \;\geq\; \gamma \;>\; 0.$$

We show that the following theorem, similar to Lorden's results (1971), holds under an assumption on the explanatory vectors.

Assumption. As $n \to \infty$,

$$\frac{1}{n} \sum_{i=k+1}^{k+n} x_i x_i' \;\to\; M \qquad \text{uniformly for all } k \geq 0, \tag{2}$$

where M is a $p \times p$ positive definite matrix. We write $\mu_\theta = \theta' M \theta / 2$.

The uniformity of the convergence in (2) also entails the fact that for any $\theta \neq 0$, there exists a common upper bound for $(x_1'\theta)^2,\ (x_2'\theta)^2,\ \ldots$. This fact will be used several times in our arguments.

Theorem. Let assumption (2) hold. For every sequence of stopping time $\{T(\gamma),\ \gamma > 1\}$ with

$$E_\infty T(\gamma) \;\geq\; \gamma \qquad \text{for all } \gamma > 1, \tag{3}$$

the following relation holds

$$D_\theta\{T(\gamma)\} \;\geq\; \frac{\log \gamma}{\mu_\theta}(1 + o(1)) \qquad \text{for all } \theta \neq 0, \text{ as } \gamma \to \infty. \tag{4}$$

Furthermore, there exists such a sequence $\{T(\gamma), \gamma > 1\}$ for which (4) holds with equality.

In fact, the $T(\gamma)$ for which (4) holds with equality can be constructed in terms of some one-sided sequential tests for the null hypothesis $\theta = 0$ against the alternative $\theta \neq 0$ in the following way. Assume $\{N(\alpha),\ \alpha \in (0,1)\}$ is a sequence of such tests with the properties that $P_\infty\{N(\alpha) < \infty\} \leq \alpha$ for all $\alpha < 1$, and

$$E_{1,\theta} N(\alpha) \;\sim\; \frac{|\log \alpha|}{\mu_\theta} \quad \text{as } \alpha \to 0, \qquad \text{for all } \theta \neq 0.$$

Let $N_k(\alpha)$ be the stopping time obtained by applying $N(\alpha)$ to $y_k,\ y_{k+1},\ \ldots,\ k = 1, 2, \ldots$. Obviously, it holds that

$$E_{k,\theta} N_k(\alpha) \;\sim\; \frac{|\log \alpha|}{\mu_\theta} \quad \text{as } \alpha \to 0, \qquad \text{for all } \theta \neq 0, \text{ and } k \geq 1. \tag{5}$$

If this convergence is uniform in k, it can be shown that the stopping time defined by

$$T(\gamma) \;=\; \min_{k \geq 1}\{N_k(\alpha) + k - 1\} \tag{6}$$

with $\gamma = 1/\alpha$, satisfies the inequality (3), and furthermore the relation (4) holds with equality. Lemma 6 in Section 6 presents a feasible example of such a sequence $\{N(\alpha),\ 0 < \alpha < 1\}$.

The proof of the theorem is postponed to Section 4. Section 3 offers the asymptotic theory for the simple case where θ is fixed. In this case one can see that the Cusum procedure is asymptotically optimal. The arguments need some elementary lemmas on stopping times of non-i.i.d. random variables. Since they are of some independent interests, we state them separately in Section 2.

2. SOME LEMMAS ON STOPPING TIMES

In this section, we assume that $\xi_1, \ \xi_2, \ \ldots$ are independent random variables. Let $S_n = \sum_{i=1}^n \xi_i$. Some lemmas are in order.

LEMMA 1. Assume that τ is a topping rule with respect to $\{\xi_n\}$, and $E\tau < \infty$. Let $\sup_{n\geq 1} E|\xi_n| < \infty$. If $\mu = E\xi_n$ for all $n \geq 1$, then $ES_\tau = \mu E\tau$.

LEMMA 2. Suppose that $\{\tau(\alpha), \ \alpha \in (0,1)\}$ is a sequence of stopping rules with respect to $\{\xi_n\}$. Suppose further $E\tau(\alpha) \to \infty$ as $\alpha \to 0$. Let $\sup_{n\geq 1} E|\xi_n| < \infty$, and $\frac{1}{n}\sum_{i=1}^n E\xi_n \to \mu$ as $n \to \infty$ with μ finite. Then $ES_{\tau(\alpha)}/E\tau(\alpha) \to \mu$ as $\alpha \to 0$.

LEMMA 3. Assume that the second moments of random variables $\{\xi_n\}$ have a common bound, and $\frac{1}{n}\sum_{i=1}^n E\xi_n \to \mu$ as $n \to \infty$. For $b > 0$, define $\tau_b = \inf\{n \geq 1 \,|\, S_n \geq b\}$.

(i) If $\mu < 0$ and $P\{\xi_n \leq 0\} > 0$ for every $n \geq 1$, then $P\{\tau_b = \infty\} > 0$.

(ii) If $\mu > 0$, then $\tau_b < \infty$ a.s., and $\tau_b/b \to \mu^{-1}$ a.s. as $b \to \infty$. Also it holds that $\sup_{b>0} E(S_{\tau_b} - b) < \infty$. Moreover if $E\tau_b < \infty$ for all $b > 0$, then $E\tau_b/b \to \mu^{-1}$ as $b \to \infty$.

(iii) If $\min_{n\geq 1} E(\xi_n \wedge c_n) > 0$ for some positive bounded constants $c_1, \ c_2, \ \ldots$, then $E\tau_b \leq K < \infty$ for all $b > 0$ with $K = (b + \max_{n\geq 1} c_n)/\min_{n\geq 1} E(\xi_n \wedge c_n) > 0$.

Lemma 1 presents a modification of Wald's equation, which weakens the assumption that the ξ's are identically distributed. For our application, the ξ's have different expectations. If their average values tend to a finite limit, Lemma 2 shows that Wald's equation holds asymptotically. Lemma 3 (ii) is a renewal theorem for non-identically distributed random variables. Although the present form is rather crude, it is sufficient for our application. More refined results can be found in Chow & Robbins (1963). The proofs of these three lemmas contain the standard techniques on stopping times only, which are omitted here.

The following Lemmas 4 and 5 are related to a sequential testing problem. Now we assume that $\eta_1, \ \eta_2, \ \ldots$ are independent random variables. Consider the testing problem

$$H_0: \ \eta_n \ \sim \ f_{0n} \qquad \text{against} \qquad H_1: \ \eta_n \ \sim \ f_{1n},$$

where f_{0n} and f_{1n}, $n = 1, 2, \ldots$, are density functions. Let $\xi_n = \log\{f_{1n}(\eta_n)/f_{0n}(\eta_n)\}$. The one-sided sequential probability ratio test is defined as $N(\alpha) = \inf\{n \geq 1 \,|\, \sum_{i=1}^n \xi_i \geq$

$|\log\alpha|\}$. Stopping means rejection of H_0. By the Wald's likelihood ratio identity (cf. Proposition 2.24 of Siegmund 1985), $P_0\{N(\alpha) < \infty\} \leq \alpha$. If the sequence $\{\xi_n,\ n \geq 1\}$ satisfies the conditions in Lemma 3, then $E_1 N(\alpha) \sim |\log\alpha|/\mu$ as $\alpha \to 0$. Lemma 4 below shows that $|\log\alpha|/\mu$ is asymptotically the lower bound for any stopping time N which satisfies the inequality $P_0(N < \infty) \leq \alpha$. This means that the one-sided sequential probability ratio test $N(\alpha)$ is asymptotically optimal.

LEMMA 4. Suppose that $\{\tau(\alpha),\ 0 < \alpha \leq 1\}$ is a family of one-sided sequential tests with $P_0\{\tau(\alpha) < \infty\} \leq \alpha$ for all $0 < \alpha \leq 1$. Let $\xi_n = \log\{f_{1n}/f_{0n}\}$, $n \geq 1$, satisfy the conditions of Lemma 2 with $\mu > 0$ under the hypothesis H_1. Then $\liminf_{\alpha\to 0} E_1\tau(\alpha)/|\log\alpha| \geq \mu^{-1}$.

PROOF. We need only to consider the case that $E_1\tau(\alpha) < \infty$. Hence $\tau(\alpha) < \infty$ a.s. under H_1. With this we can easily show that $P_0\{\tau(\alpha) < \infty\} = E_1\exp\{-S_{\tau(\alpha)}\}$. By Jensen's inequality, $-\log P_0\{\tau(\alpha) \leq E_1 S_{\tau(\alpha)}$. Hence $|\log\alpha| \leq E_1 S_{\tau(\alpha)}$. By Lemma 2, $E_1 S_{\tau(\alpha)} \sim \mu E_1\tau(\alpha)$. □

LEMMA 5. Suppose that $\frac{1}{n}\sum_{i=1}^{n} E_1\{\log(f_{1i}/f_{0i})\} \to \mu$ with μ finite. Then for any two sided sequential test with a stopping time N and error probabilities α_1, α_2,

$$(\mu+\varepsilon)E_1 N \geq (1-\alpha_2)|\log\alpha_1| - C(\varepsilon),$$

where $\varepsilon > 0$ is arbitrary, and $C(\varepsilon)$ is a finite constant.

PROOF. Let $E_1 N < \infty$. By Wald's inequality (cf. Proposition 2.39 of Siegmund 1985),

$$E_1 S_N \ \geq\ (1-\alpha_2)\log\frac{1-\alpha_2}{\alpha_1} + \alpha_2\log\frac{\alpha_2}{1-\alpha_1},$$

where $S_n = \sum_{i=1}^{n}\log\{f_{1i}/f_{0i}\}$. Notice that $\alpha_2\log(1-\alpha_1)^{-1}$ is non-negative, and $(1-\alpha_2)\log(1-\alpha_2) + \alpha_2\log\alpha_2$ attains its minimum value $-\log 2$ at $\alpha_2 = 1/2$. Thus $E_1 S_N \geq (1-\alpha_2)\log|\alpha_1| - \log 2$. In terms of Lemma 1, we can prove that for any $\varepsilon > 0$, there exists $C(\varepsilon) > 0$ such that

$$-C(\varepsilon) + (\mu-\varepsilon)E_1 N \ \leq\ E_1 S_N \ \leq\ (\mu+\varepsilon)E_1 N + C(\varepsilon).$$

The conclusion follows immediately. □

3. ASYMPTOTIC THEORY

This section discusses a simple case that θ in model (1) is fixed. Since β is known, there is no loss of generality in assuming $\beta = 0$. Hence under P_∞, the $y_1,\ y_2,\ \ldots$ are identically distributed. To simplify the notation a little, we omit the θ as subscript and write P_m, E_m, and $D(.)$ instead of $P_{m,\theta}$, $E_{m,\theta}$, and $D_\theta(.)$.

PROPOSITION 1. Let N be a stopping time with respect to $y_1,\ y_2,\ \ldots$ such that $P_\infty(N < \infty) \le \alpha$ for some $\alpha \in (0,1)$. For $k = 1,2,\ldots$, let N_k denote the stopping time obtained by applying N to $y_k,\ y_{k+1},\ \ldots$, and define

$$N^* \ = \ \min\{\, N_k + k - 1 \,|\, k = 1,2,\ldots\}.$$

Then N^* is a stopping time with $D(N^*) \le \sup_{m\ge 1} E_m N_m$, and $E_\infty N^* \ge \alpha^{-1}$.

PROOF. Notice that $\{N^* \le n\}$ is the union of $\{N_1 \le n\}$, $\{N_2 \le n-1\}$, $\ldots$, $\{N_n \le 1\}$, which are determined by $y_1,\ \ldots\ y_n$. Hence N^* is a stopping time and $\{N^* \ge m - 1 + k\}$ is a subset of $\{N_m \ge k\}$, which implies that

$$D(N^*) \ = \ \sup_m \operatorname{ess\,sup} E_m\{\,(N^* - m + 1)^+ \,|\, Y_{m-1}\,\} \ \le \ \sup_m E_m(N_m).$$

Since $y_1,\ y_2,\ \ldots$ are i.i.d. under P_∞, the inequality $E_\infty N^* \ge \alpha^{-1}$ follows along the same lines as those in Theorem 2 of Lorden (1971). □

Consider the hypotheses H_0 that $y_1,\ y_2,\ \ldots$ are i.i.d. N(0, 1) variables against H_1 that $y_1,\ y_2,\ \ldots$ are independent and $y_i \sim$N$(\theta' x_i, 1)$. A one-sided sequential probability ratio test can be defined as

$$N(\alpha) \ = \ \inf\{\, n \ge 1 \,|\, S_n \ge |\log\alpha|\,\}, \tag{7}$$

where $S_n = \sum_{i=1}^n \xi_i$ and $\xi_i = \theta' x_i(y_i - \theta' x_i/2)$. By the Wald's likelihood ratio identity, $P_\infty\{N(\alpha) < \infty\} \le \alpha$. Since under assumption (2), the second moments of ξ's have a common bound, it is easy to see that $P_1\{N(\alpha) < \infty\} = 1$ from Lemma 3. Denote by $N_k(\alpha)$ the stopping time obtained by applying $N(\alpha)$ to $y_k,\ y_{k+1},\ \ldots$. For $\gamma = \alpha^{-1}$, define

$$N^*(\gamma) \ = \ \inf\{\, N_k(\alpha) + k - 1 \,|\, k = 1,2,\ldots\}. \tag{8}$$

Then $N^*(\gamma)$ can be expressed by

$$N^*(\gamma) \ = \ \inf\{\, n \ge 1 \,|\, \max_{1\le k\le n}(S_n - S_k) \ge \log\gamma\,\},$$

which is a generalized Cusum procedure (cf. Lorden 1971, Siegmund 1985 §2.6).

PROPOSITION 2. Under assumption (2), $\limsup_{\gamma\to\infty} D\{N^*(\gamma)\}/\log\gamma \le \mu^{-1}$, where $N^*(\gamma)$ is defined as in (8), and $\mu = \theta' M\theta/2$.

PROOF. By Proposition 1, we need only to show that

$$E_m\, N_m(\alpha) \ \sim \ |\log\alpha|/\mu \quad \text{uniformly for } m \ge 1, \qquad \text{as } \alpha \to 0. \tag{9}$$

Let $\mu_i = E_1\xi_i = (\theta' x_i)^2/2$. By assumption (2), $\frac{1}{n}\sum_{i=1}^n \mu_i \to \mu$. Furthermore by the uniformity of the convergence in (2), $\sup_{n\geq 1} E_1|\xi_n| < \infty$, and there exists a positive integer d such that

$$\frac{1}{2}\mu \;\leq\; \frac{1}{d}\sum_{i=k+1}^{k+d} \mu_i \;\leq\; \frac{3}{2}\mu, \quad \text{for all } k \geq 0.$$

Let $S_0 \equiv 0$, $Z_n = S_{nd} - S_{(n-1)d}$, $n \geq 1$. Then under P_1, Z_n is normally distributed with mean ν_n and variance $2\nu_n$, where $\nu_n = \sum_{i=(n-1)d+1}^{nd} \mu_i \;\geq\; \mu d/2 > 0$. Let $\tau_\alpha = \inf\{\, n \geq 1 \,|\, \sum_{i=1}^n Z_i \geq |\log\alpha|\,\}$. Obviously, $N(\alpha) \leq d\cdot\tau_\alpha$ a.s.. Let $K_1 > 0$ such that

$$\nu \;\equiv\; (\mu d)^{1/2}[(\mu d)^{1/2}/2 + K_1\{1-\Phi(K_1)\} - \varphi(K_1)] \;>\; 0,$$

where $\varphi(.)$ denotes the standard normal density and $\Phi(.)$ its distribution function. Let $Z_n' = Z_n \wedge (\nu_n + K_1\sqrt{2\nu_n})$, some integration operation entails that $0 < \nu \leq E_1 Z_n' \leq K_2 < \infty$, where $K_2 = 3\mu d/2 + K_1(3\mu d)^{1/2}$. By Lemma 3 (iii), $E_1\tau_\alpha \leq (|\log\alpha| + K_2)/\nu < \infty$. Consequently,

$$E_1 N(\alpha) \;\leq\; d\cdot E_1\tau_\alpha \;\leq\; K_3 \;<\; \infty, \tag{10}$$

where $K_3 = d(|\log\alpha| + K_2)/\nu$. Therefore from Lemma 3 (ii), (9) holds for $N(\alpha)$. By the definition of $N_m(\alpha)$ and (7), (9) also holds for all $N_m(\alpha)$, for $m \geq 1$. Its uniformity of the convergence follows from the uniformity assumption in (2). □

By Proposition 2,

$$\limsup_{\gamma\to\infty} \inf D(N)/\log\gamma \;\leq\; \mu^{-1}, \tag{11}$$

where the infimum is taken over N with $E_\infty N > \gamma$. Proposition 3 shows that the limit of the inf $D(N)/\log\gamma$ equals μ^{-1}. Hence the Cusum procedure $N^*(\gamma)$ defined in (8) is asymptotically optimal.

PROPOSITION 3. Under assumption (2), it holds that $\inf D(N) \sim \log\gamma/\mu$ as $\gamma \to \infty$, where $\mu = \theta' M\theta/2$, and the infimum is taken over N with $E_\infty N > \gamma$.

PROOF. By the virtue of (11), we need only to show that for any $\varepsilon \in (0,1)$, there is a constant $C(\varepsilon) > 0$ such that for every stopping time N,

$$(\mu+\varepsilon)D(N) \;\geq\; (1-\varepsilon)\log\{E_\infty N\} - C(\varepsilon). \tag{12}$$

For fixed ε, define $T_0 \equiv 0 < T_1 < T_2 < \ldots$ as follows:

$$T_k \;=\; \inf\{\, k \geq T_{k-1} \,|\, (S_k - S_{T_{k-1}}) \geq |\log\varepsilon|\,\} \quad \text{for } k \geq 1,$$

where $S_n = \sum_{i=1}^n \xi_i$, $\xi_i = -\theta' x_i(y_i - \theta' x_i/2)$. Let $R = \inf\{k \geq 1 \,|\, T_k \geq N\}$. Following the proof of Theorem 3 of Lorden (1971) with Lemma 5 instead of the Wald's theorem, we have that

$$(\mu+\varepsilon)D(N) \;\geq\; (1-\varepsilon)\log\{E_\infty R\} - C(\varepsilon), \tag{13}$$

where $C(\varepsilon) > 0$ depends only on ε, which is guaranteed by the uniformity in (2).

It is easy to show that $T_n - T_{n-1}$, $n = 1, 2, \ldots$, are independent, and

$$P_\infty\{T_2 - T_1 = n\} \;=\; \sum_{k=1}^{\infty} P_\infty\{T_1 = k\}\, P_\infty\{T_1^{(k)} = n\} \quad \text{for } n \geq 1,$$

where $T_1^{(k)}$ is T_1 applied to $y_{k+1}, y_{k+2}, \ldots$. By Lemma 3 (ii), $T_1^{(k)} < \infty$ a.s. under P_∞ for all $k \geq 1$. Consequently, $T_2 - T_1 < \infty$ a.s. under P_∞ also. The similar arguments to relation (10) imply that

$$E_\infty T_1 \;\leq\; K \;<\; \infty, \tag{14}$$

where K is equal to K_3 in (10) with ε instead of α. From the uniformity of the convergence in (2), the inequality (14) holds for all $T_1^{(k)}$, $k = 1, 2, \ldots$. Hence

$$E_\infty(T_2 - T_1) \;=\; \sum_{n=1}^{\infty} n\, P_\infty(T_2 - T_1 = n) \;=\; \sum_{k=1}^{\infty} P_\infty(T_1 = k)\, E_\infty T_1^{(k)} \;\leq\; K.$$

Similarly, we can show that $E_\infty(T_n - T_{n-1}) \leq K$ for all $k \geq 1$. Notice that $\{R \geq n\}$ is determined by $\{T_k - T_{k-1},\ k = 1, \ldots, n-1\}$, which is independent to $T_n - T_{n-1}$. Hence

$$E_\infty T_R = E_\infty \sum_{n=1}^{\infty} (T_n - T_{n-1})\, I_{\{R \geq n\}} = \sum_{n=1}^{\infty} E_\infty(T_n - T_{n-1})\, P_\infty\{R \geq n\} \leq K \cdot E_\infty R.$$

From the definition of R, $\log\{E_\infty N\} \;\leq\; \log\{E_\infty T_R\} \;\leq\; \log\{E_\infty R\} + \log K$. The relation (12) follows from this inequality and (13) immediately. □

4. PROOF OF THEOREM

The relation (4) follows directly from Proposition 3. From the following Lemma 6, there exists a sequence of stopping time $\{N(\alpha),\ 0 < \alpha < 1\}$ for which $P_\infty\{N(\alpha) < \infty\} \leq \alpha$, and (5) holds uniformly for $k \geq 1$. Define $T(\gamma)$ as in (6). Proposition 1 implies that $T(\gamma)$ is a stopping time, and for which the inequality (3) holds, and further $D\{T(\gamma)\} \leq \sup_{m \geq 1} E_{m,\theta}\{N_m(\alpha)\}$ with $\alpha = \gamma^{-1}$. Since the convergence in (5) is uniform, the inequality (4) holds for $T(\gamma)$ with equality.

LEMMA 6. Let $\delta_\alpha = |\log \alpha|^{-1}$, and $h_\alpha = |\log \alpha| + \log(|\log \alpha|^{4p+6})$, where p is the dimension of the vector θ. Define

$$N(\alpha) \;=\; \inf\{\, n \geq 1 \mid \sup_{||\theta_n|| \geq \delta_\alpha} (\theta' X_n' Y_n - \frac{1}{2}\theta' X_n' X_n \theta) \geq h_\alpha \,\},$$

where $\theta_n = n^{-1/2}(X_n' X_n)^{1/2}\theta$. Then under assumption (2), $P_\infty\{N(\alpha) < \infty\} \leq \alpha$ for $0 < \alpha < \alpha_0$, where $\alpha_0 \in (0, 1)$ is a constant. Furthermore, the relation (5) holds uniformly for $k \geq 1$.

PROOF. Let $Z_n = n^{1/2}(X_n'X_n)^{-1/2}X_n'Y_n$. For $n \le n_0 \equiv [2h_\alpha/\delta_\alpha^2]$, $\sup_{||\theta_n|| \ge \delta_\alpha}(\theta' X_n' Y_n - \frac{1}{2}\theta' X_n' X_n \theta) \ge h_\alpha$ if and only if

$$||Z_n|| \;\ge\; \inf_{||\theta_n|| \ge \delta_\alpha} \{ \frac{n}{2}||\theta_n|| + h_\alpha ||\theta_n||^{-1} \} \;=\; (2\, h_\alpha\, n)^{1/2}. \tag{15}$$

Under P_∞ the random variable $||Z_n||^2/n$ has the distribution $\chi^2(p)$. Hence

$$P_\infty\{N(\alpha) = n\} \;\le\; P_\infty\{\frac{1}{n}||Z_n||^2 \ge 2h_\alpha\} \;\le\; 3h_\alpha^p e^{-h_\alpha}.$$

Consequently,

$$P_\infty\{N(\alpha) \le n_0\} \;\le\; 3n_0 h_\alpha^p e^{-h_\alpha} \;=\; o(\alpha/|\log \alpha|). \tag{16}$$

For $n > n_0$, the infimum in (15) is attained at a point with $||\theta_n|| = \delta_\alpha$. Since $n_0\delta_\alpha^2/4 + (h_\alpha/\delta_\alpha)^2 \ge h_\alpha - \delta_\alpha^2/4$, some elementary algebra estimates entail that

$$\begin{aligned} P_\infty\{n_0 < N(\alpha) < \infty\} \;&\le\; \sum_{n=n_0+1}^{\infty} P_\infty\{ \frac{1}{n}||Z_n||^2 \ge (\frac{n^{1/2}}{2}\delta_\alpha + n^{-1/2}\frac{h_\alpha}{\delta_\alpha})^2 \} \\ &\le\; \frac{144\, p!}{\delta_\alpha^2} h_\alpha^p (1 + \frac{2h_\alpha}{\delta_\alpha^2})^{p+1} \exp\{-h_\alpha + \frac{\delta_\alpha^2}{8}\} \;=\; O(\alpha/|\log \alpha|). \end{aligned}$$

Based on this and (16), we can choose a sufficiently small $\alpha_0 \in (0, 1)$ such that $P_\infty\{N(\alpha) < \infty\} \le \alpha$ for all $\alpha \in (0,\, \alpha_0)$.

To show (5), we define

$$N_\theta(\alpha) \;=\; \inf\{\, n \ge 1 \,|\, \theta' X_n' Y_n - \frac{1}{2}\theta' X_n' X_n \theta \ge h_\alpha \,\} \quad \text{for } \theta \ne 0.$$

Since $\delta_\alpha \downarrow 0$ as $\alpha \to 0$, $\theta_n \ge \delta_\alpha$ for all sufficiently small α, large n and fixed $\theta \ne 0$. Hence by Lemma 3,

$$h_\alpha^{-1} E_{1,\theta} N(\alpha) \;\le\; h_\alpha^{-1} E_{1,\theta} N_\theta(\alpha) \;\to\; \mu_\theta^{-1} \qquad \text{as } \alpha \to 0.$$

Lemma 4 and (3) imply that $|\log \alpha|^{-1} E_{1,\theta} N(\alpha) \to \mu_\theta^{-1}$. It is easy to see that

$$h_\alpha^{-1} E_{k,\theta} N_k(\alpha) \;\to\; \mu_\theta^{-1}, \quad \text{for } k \ge 1,$$

where $N_k(\alpha)$ is the stopping time obtained by applying $N(\alpha)$ to $y_k,\ y_{k+1},\ \ldots$. The uniformity in convergence of (2) guatantees that the convergence in the above limit is also uniform for all $k \ge 1$. □

ACKNOWLEDGMENTS

The research was carried out during the anthor's stay at the University of Freiburg with partial support by the Alexander von Humboldt Foundation and the National Science Foundation of China. The author also thanks Professor H. R. Lerche for many helpful discussions and suggestions.

BIBLIOGRAPHY

Chow, Y.S. & Robbins, H. (1963). A renewal theorem for random variables which are dependent or nonidentically distributed. *Ann. Math. Statist.* **34**, 390-395.

Hinkley, D.V. (1971). Inference in two phase regression. *J. Amer. Statist. Assoc.* **66**, 736-743.

Lorden, G. (1971). Procedures for reacting to a change in distribution. *Ann. Math. Statist.* **42**, 1897-1908.

Moustakides, G.V. (1986). Optimal stopping times for detection changes in distributions. *Ann. Statist.* **14**, 1379-1387.

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

Pollak, M. (1985). Optimal detection of a change in distribution. *Ann. Statist.* **13**, 206-227.

Pollak, M. & Siegmund, D. (1985). A diffusion process and its applications to detecting a change in the drift of Brownian motion. *Biometrika* **72**, 267-280.

Quandt, R.E. (1960). Tests of the hypothesis that a linear regression system obeys two separate regimes. *J. Amer. Statist. Assoc.* **55**, 324-330.

Roberts, S.W. (1966). A comparison of some control chart procedures. *Technometrics* **8**, 411-430.

Shiryayev, A.N. (1963). On optimal methods in quickest detection problems. *Theor. Probab. Appl.* **8**, 22-46.

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

Worsley, K.J. (1983). Testing for a two phase multiple regression. *Technometrics* **25**, 35-42.

Received July 1991; Revised February 1993

Recommended by P.K. Sen