The Annals of Statistics
1978, Vol. 6, No. 1, 85–91

EXPONENTIALLY BOUNDED STOPPING TIMES OF INVARIANT SPRT's IN GENERAL LINEAR MODELS: FINITE mgf CASE

BY S.-S. PERNG

Howard University

A general theorem which is useful in proving the exponential boundedness of the stopping time of sequential tests for parameters in general linear models is formulated; this theorem is formulated under the assumptions that the squared error has a finite moment-generating function and the sequence of the running averages of the concomitant variables converges. Applications are given.

1. Introduction. Let $y_1, y_2, \cdots$ be independent random variables (vectors) with common distribution P, and for each n let L_n be a function of $y_1, \cdots, y_n$ and n. For $l > 0$, let the stopping time N be defined as

$$N = \min\{n \geqq 1 : L_n \notin (-l, l)\} . \tag{1.1}$$

This research will be concerned with the exponential boundedness of N, i.e.,

$$P[N > n] \leqq c\rho^n , \qquad n = 1, 2, \cdots \tag{1.2}$$

for some $c > 0$ and $0 < \rho < 1$. If (1.2) cannot be satisfied, then P is called obstructive. The stopping time N is said to be finite a.s. (P) if $P[N = \infty] = 0$.

When the y_i's are i.i.d., the exponential boundedness of N has been extensively investigated by Wijsman [16–21] and Lai [9]. See also Savage and Sethuraman [12], Sethuraman [13] and Stein [15]. For the non-i.i.d. case, the field is relatively unexplored. Berk [2] considered the stopping time of SPRT based on exchangeable models. In this paper another situation of the non-i.i.d. case where the y_i's are the observed values of linear models is studied. Some examples of this type were found in Perng [11]. It is noted that the test of hypotheses about the parameters in linear models is widely studied. References can be found in [5] or [6].

To keep the paper from being too long, some generality in the "true" distribution P of the random error and in the concomitant variables is sacrificed. It is assumed throughout that under P the error has 0 mean and its square has a finite moment-generating function (mgf) and that the sequences of the running averages of the concomitant variables and the running averages of the squares converge.

The exponential boundedness of N is proved for the sequential T^2-test of parameters in general linear models under the above assumptions, unless the random error e satisfies

$$P[f(e) = 0] = 1 , \tag{1.3}$$

Received November 1974; revised August 1976.
AMS 1970 *subject classifications.* Primary 62L10; Secondary 62J10, 62H15.
Key words and phrases. Invariant SPRT, exponentially bounded stopping time, linear models.

for a particular function f. Similar results for other sequential tests are also noted.

General theorems about the exponential boundedness of N are given in Section 2.

2. Theorems in exponential boundedness. In this section general theorems are proved that present sufficient conditions for the validity of (1.2) with N defined by (1.1) and L_n being a sequence of random variables satisfying certain conditions. The theorems are generalizations of Theorem 2.1 in [17] suited for the application to the case where the observations come from linear models.

Let $u, u_1, u_2, \cdots$ be i.i.d. random vectors with common distribution P. Write $E(\cdot)$ for $E_P(\cdot)$.

ASSUMPTION A. Let $\{\gamma_n\}$ be a sequence of numbers such that $\gamma_n \to \gamma$ as $n \to \infty$ and $\gamma_{n+1} - \gamma_n = O(n^{-1})$.

ASSUMPTION B. Assume that (i) $E(u) = \xi$ and (ii) $E(\exp t||u||^2) < \infty$ for t in some neighborhood of 0.

ASSUMPTION C. Let $\{d_n\}$ be a sequence of bounded vectors and let $\{D_n\}$ be a sequence of bounded matrices such that as $n \to \infty$, $\bar{d}_n \to d$ and $\bar{D}_n \to D$. (As usual, $\bar{x}_n = (1/n) \sum_1^n x_i$.)

Let $z_n = D_n u_n$. The following theorem is an extension of a theorem of Chernoff [3] and the proof is similar.

THEOREM 2.1. *Under Assumption* B *with* $\xi = 0$, $\bar{z}_n$ *converges to* 0 *exponentially*, i.e., *for any* $\varepsilon > 0$, $P[||\bar{z}_n|| > \varepsilon] \leqq c\rho^n$, $n = 1, 2, \cdots$ *for some* $c > 0$ *and* $0 < \rho < 1$, *provided that* $\{D_n\}$ *is bounded.*

The first corollary is an immediate consequence of the theorem by noting that

$$(1/n) \textstyle\sum_1^n D_k u_k - D\xi = (1/n) \sum_1^n (D_k - D)(u_k - \xi) + (1/n) \sum_1^n D(u_k - \xi) + (1/n) \sum_1^n (D_k - D)\xi \, .$$

COROLLARY 2.1.1. *Under Assumptions* B *and* C, $\bar{z}_n$ *converges to* $z = D\xi$ *exponentially.*

COROLLARY 2.1.2. *In Theorem* 2.1 *or Corollary* 2.1.1 $\bar{z}_n \to z = D\xi$ a.s. (P).

This corollary is an immediate consequence of Theorem 2.1 or Corollary 2.1.1. It also holds without Assumption B(ii) (see, e.g., [4], page 122).

Write $v_n' = (z_n', d_n')$ and $w_n' = (\bar{v}_n, \gamma_n)$. Also, write $v' = (z', d')$ and $w' = (v', \gamma) = (z', d', \gamma)$. In our application L_n may be uniformly approximated by a random variable $n\Phi(w_n)$. To prove the exponential boundedness (and finiteness) of N, we may write $L_n = n\Phi(w_n)$ (cf. e.g., [21]). Let $(\partial/\partial w)\Phi(w)$ denote a column vector of partial derivatives.

ASSUMPTION D. The function Φ has continuous first partial derivatives on a neighborhood V of w. Let $P = (\partial/\partial z)\Phi$ evaluated at $w' = (z', d', \gamma)$. Let $a_n = P'D_n$ so that $P'z_n = a_n'u_n$. Assume that $\bar{a}_n$ converges to a and

$$(2.1) \qquad P[a'(u - \xi) = 0] < 1 .$$

THEOREM 2.2. *Under Assumptions* A, B(i) *and* C, *if* $\Phi(w) \neq 0$, *then* N *is finite* a.s. (P). *If Assumption* B(ii) *also holds*, N *is exponentially bounded.*

PROOF. The proof follows the same lines as those in case 1 of Theorem 2.1 in [17] (cf. also Theorem 2.3 in [21]).

THEOREM 2.3. *Under Assumptions* A, B(i), C *and* D, *if* $\Phi(w) = 0$, *then* N *is finite* a.s. (P). *If Assumption* B(ii) *also holds, then* N *is exponentially bounded.*

PROOF. Without loss of generality, suppose that $w = 0$. Following an argument similar to the one found in the proof of Theorem 2.1, case 2, in [17], we can show (with w_{jr}, L_{jr}, $w_{(j+1)r}$ and $L_{(j+1)r}$ playing the role of $\bar{x}_n$, Φ_n, $\bar{x}_{n+r}$ and Φ_{n+r} respectively, and using Assumption A in deriving the counterpart of (2.12) in [17]) that

$$(2.2) \qquad [w_{jr} \in V; w_{(j+1)r} \in V; |L_{(j+1)r} - L_{jr}| < 2l] \subset [||\omega_{j+1}|| > B_1 \text{ or } |\Delta'\omega_{j+1}| < 2l + 2\delta] = E_{j+1}, \quad \text{say},$$

where V is a small convex neighborhood of $w = 0$, $\Delta - (\partial/\partial w)\Phi(0)$,

$$\omega'_{j+1} = (j+1)rw'_{(j+1)r} - jrw'_{jr} = (\textstyle\sum_{i=1}^{r} v'_{jr+i}, (j+1)r\gamma_{(j+1)r} - jr\gamma_{jr}) ,$$

$\delta > 0$ and $r(B_1)$ is a positive integer (large real number) to be chosen later. Note that the E_i's are independent.

Since $\bar{z}_n \to z(= 0)$ a.s. (P), by Assumptions A and C and Corollary 2.1.2 $w_n \to w(= 0)$ a.s. (P), so that for each $\varepsilon > 0$ there is an integer j_0 such that $P[F] \leqq \varepsilon$, where F is the complement of $[w_{jr} \in V, j \geqq j_0]$. By the following lemma for the proper choice of r and B_1, $P[\bigcap_{n=j_0}^{\infty} E_j] = 0$. Thus following the same argument as in the paragraph containing (2.16) in [17], it can be shown that $P[N = \infty] = 0$.

Next, note that if $w_{jr} \in V$ and $|L_{jr}| \geqq l$, then $N \leqq jr$. Thus

$$(2.3) \qquad P[N > j(r+1)r] \leqq \textstyle\sum_{i=1}^{jr} P[w_{(j+i)r} \notin V] + P[w_{(j+1)r} \in V; |L_{(j+i)r}| < l, i = 0, 1, \cdots, jr] .$$

By Assumptions A and C, when j is sufficiently large, $j \geqq j_1$ say,

$$[w_{(j+i)r} \notin V] = [\bar{z}_{(j+i)r} \notin V_z] , \qquad i = 0, 1, 2, \cdots ,$$

where V_z is the cross-section of V in the space of $\bar{z}_n$. Hence, by Theorem 2.1 or Corollary 2.1.1, for $j \geqq j_1$

$$(2.4) \qquad P[w_{(j+i)r} \notin V] \leqq c_2\rho_2^{(j+i)r} , \qquad i = 0, 1, 2, \cdots ,$$

for some $c_2 > 0$ and $0 < \rho_2 < 1$. By (2.2) and the lemma below, the second term on the right-hand side of (2.3) does not exceed

$$\prod_{i=1}^{jr} P[E_{j+i}] < \rho_1^{j((r+1)c_1/2-1)} \tag{2.5}$$

for $j \geqq j_2$, say. Thus, by (2.4) and (2.5), $P[N > j(r+1)r] < c_3\rho_3^j$ for some $c_3 < 0$ and $0 < \rho_3 < 1$. The exponential boundedness of N follows (cf. [15]).

LEMMA. *For proper choice of r and B_1, there is a set J of positive integers such that for $j \in J$, $P[E_j] < \rho_1 < 1$ and that* $\liminf k'/k = c_1 > 0$, *where k' is the number of integers in J not exceeding k.*

PROOF. Write $\Delta'\omega_j = s_j + d_j^*$, where $s_j = \sum_{i=1}^r a'_{(j-1)r+i} u_{(j-1)r+i}$, the a_i's are defined in Assumption D and $d_j^* = \Delta'\omega_j - s_j$. It can be shown (see the proof of (5.8) in [11]) that for a proper choice of r, there is a set J of positive integers and $\varepsilon > 0$ such that for $j \in J$,

$$P[|\Delta'\omega_j| \geqq 2l + 2\delta] > \varepsilon^r \tag{2.6}$$

and $\liminf k'/k = c_1 > 0$, where k' is the number of integers in J which are less than or equal to k. Next, by Assumptions A and C, $||\omega_j|| \leqq B_3 \sum_{i=1}^r ||u_{(j-1)r+i}||^2 + rB_4$ for some B_3 and B_4. Hence $P[||\omega_j||^2 \leqq B_5] \geqq P[B_3 \sum_{i=1}^r ||u_i||^2 + rB_4 \leqq B_5] \to 1$ as $B_5 \to \infty$. Thus for $j \in J$, by (2.6), B_1 may be chosen so large that

$$P[\tilde{E}_j] = P[||\omega_j|| \leqq B_1; |\Delta'\omega_j| \geqq 2l + 2\delta] \geqq \varepsilon^r/2\,, \tag{2.7}$$

where $\tilde{E}_j$ is the complement of E_j. The lemma follows with $\rho_1 = 1 - \varepsilon^r/2$.

3. Applications. Consider the linear model

$$y_i = \beta_1 x_{1i} + \beta_2 x_{2i} + e_i\,, \qquad i = 1, 2, \cdots, \tag{3.1}$$

where the e_i's are i.i.d. distributed random p-vectors, β_1 and β_2 are $p \times 1$ and $px(q-1)$ parameters, and $\{x_{1i}\}$ and $\{x_{2i}\}$ are sequences of real numbers and $(q-1)$-vectors respectively. Write $x_i' = (x_{1i}, x'_{2i})$. Let $Y_n = (y_1, \cdots, y_n)$. Define X_n, E_n, X_{1n} and X_{2n} similarly. Let

$$K_n = (1/n)(X_nX_n') = \begin{pmatrix} K_{11n} & K_{12n} \\ K_{21n} & K_{22n} \end{pmatrix} = (1/n)\begin{pmatrix} X_{1n}X'_{1n}\,, & X_{1n}X'_{2n} \\ X_{2n}X'_{1n}\,, & X_{2n}X'_{2n} \end{pmatrix}, \tag{3.2}$$

$$F_n = (1/n)E_nX_n' = (F_{1n}, F_{2n}) = (1/n)E_n(X'_{1n}, X'_{2n}) \tag{3.3}$$

$$k_n = K_{11n} - K_{12n}K_{22n}^{-1}K_{21n}\,, \tag{3.4}$$

$$\begin{aligned} U_n &= (nk_n)^{-\frac{1}{2}}Y_n(I_n - X'_{2n}(X_{2n}x'_{2n})^{-1}X_{2n})X'_{1n} \\ &= (n^{-1}k_n)^{-\frac{1}{2}}(k_n\beta_1 + F_{1n} - F_{2n}K_{22n}^{-1}K_{21n})\,, \end{aligned} \tag{3.5}$$

$$W_n = Y_n(I_n - X_n'(X_nX_n')^{-1}X_n)Y_n' = n(M_n - F_nK_n^{-1}F_n')\,, \tag{3.6}$$

where $M_n = E_nE_n'$ and I_n is the n^2 identity matrix.

Throughout this section, it is assumed that:

ASSUMPTION B′. Under the true distribution P, $E(e) = 0$, $E(ee') = \Sigma$, where Σ is positive definite and $E(e^{t||e||^2}) < \infty$ for t in some neighborhood of 0.

ASSUMPTION C′. The sequence $\{x_n\}$ is bounded and $\bar{x}_n' = (\bar{x}_{1n}, \bar{x}_{2n}') \to x_0' = (x_{10}, x_{20}')$. The matrix K_n is positive definite for $n \geqq q$ and $K_n \to K$ as $n \to \infty$, where $K = \begin{pmatrix} K_{11} & K_{12} \\ K_{21} & K_{22} \end{pmatrix}$ is positive definite.

To apply Theorems 2.2 and 2.3, we identify $u_n' = (e_n, e_n e_n')$, $\xi' = (0, \Sigma)$, $D_n = \begin{pmatrix} x_n & 0 \\ 0 & I_p \end{pmatrix}$, $z_n = D_n u_n = \begin{pmatrix} x_n & e_n' \\ e_n & e_n' \end{pmatrix}$ and $d_n = x_n x_n'$. Hence $\bar{z}_n' = (F_n, M_n)$ and $\bar{d}_n = K_n$. Note that as $n \to \infty$

$$(3.7) \qquad \bar{D}_n \to D = \begin{pmatrix} x_0 & 0 \\ 0 & I_p \end{pmatrix} \qquad \text{and} \qquad k_n \to k = K_{11} - K_{12} K_{22}^{-1} K_{21} > 0$$

and that by Corollary 2.1.2

$$(3.8) \qquad \bar{z}_n' = (F_n, M_n) \to (F, \Sigma) = (0, \Sigma) \quad \text{a.s.} \quad (P)\,.$$

T^2-test. Consider the test of $H_1\colon \lambda = \lambda_1$ vs. $H_2\colon \lambda = \lambda_2$, where $\lambda = \beta_1'\Sigma^{-1}\beta_1$ and $0 \leqq \lambda_1 < \lambda_2$. To generate the test, assume that the e_i's are i.i.d. $N(0, \Sigma)$. Then it is shown (cf. [6] or [10]) that an invariant SPRT is based at stage n on the probability ratio R_n of $T_n = (n - q)U_n''W^{-1}U_n$ which is noncentral F-distributed with p and $(n - p - q + 1)$ degrees of freedom and noncentrality $nk_n\lambda_j$ under H_j, $j = 1, 2$. It can be shown by using a result of Skovgaard [14] (a similar result was obtained in [8]) that $L_n = \log R_n$ can be uniformly approximated by n times

$$\Phi(F_n, M_n, K_n, \gamma_n) = \tfrac{1}{2}k_n(\lambda_1 - \lambda_2) - H(\lambda_1, \gamma_n, k_n, \eta_n) + H(\lambda_2, \gamma_n, k_n, \eta_n)\,,$$

where $H(\lambda, \gamma, k, \eta) = \frac{1}{4}(\lambda\gamma k\eta + \xi(\lambda\gamma k\eta))$, $\xi(x) = (x(1+x))^{\frac{1}{2}} + \log(x^{\frac{1}{2}} + (1+x)^{\frac{1}{2}})$, $\gamma_n = n/(4n - 4q - 2p + 4)$ and $\eta_n = T_n/(1 + T_n)$, provided that T_n is bounded away from 0. Note that by (3.5), (3.6), (3.7) and (3.8), $T_n \to k\lambda$ a.s. (P). Hence, if $\beta_1 = 0$ (which implies $\lambda = 0$), then $L_n/n \to \frac{1}{2}k(\lambda_1 - \lambda_2) \neq 0$, a.s. (P); therefore by Theorem 2.2, N is exponentially bounded. From here on assume that $\beta_1 \neq 0$ and replace L_n/n by $\Phi(F_n, M_n, K_n, \gamma_n)$. Again by Theorem 2.2, if $\Phi_0 = \Phi(0, \Sigma, K, \frac{1}{4}) \neq 0$, N is exponentially bounded. Next consider the case $\Phi_0 = 0$. (It can be shown that such a case exists.) It is shown that P in Assumption D is found to be (except for a nonzero factor) $P' = ((\lambda_{ij}), (\pi_{ij}))$ with

$$\lambda_{i1} = 2\beta_1'\sigma^{\cdot i} \quad \text{and for} \quad 2 \leqq j \leqq q\,, \quad \lambda_{ij} = 2\beta_1'\sigma^{\cdot i}k^{j1}k_{11}\,,$$
$$\pi_{ii} = -k(\sigma^{i\cdot}\beta_1)^2 \quad \text{and for} \quad i > j\,, \quad \pi_{ij} = -2k\beta_1'\sigma^{\cdot i}\sigma^{j\cdot}\beta_1\,,$$

with $\Sigma^{-1} = (\sigma^{ij})$, $K^{-1} = (k^{ij})$ and $\sigma^{\cdot i}(\sigma^{i\cdot})$ is the ith column (row) of Σ^{-1}. Thus in Assumption D, $a' = P'D$ and

$$(3.9) \qquad [a'(u - \xi) = 0] = [\text{trace } P'D(e', ee' - \Sigma)' = 0]$$
$$= [\beta_1'\Sigma^{-1}e = b \pm (\lambda + b^2)^{\frac{1}{2}}]\,,$$

where $b = x'k^{\cdot 1}$ and $k^{\cdot 1}$ is the first column of K^{-1}. Since $E(\beta_1'\Sigma_e^{-1}e) = 0$, (3.9) with probability 1 is equivalent to

$$(3.10) \qquad \beta_1'\Sigma^{-1}e = b \pm (\lambda + b^2)^{\frac{1}{2}} \quad \text{with probability} \quad \rho \quad \text{and} \quad 1 - \rho\,,$$

respectively, where $\rho = ((\lambda + b^2)^{\frac{1}{2}} - b)/(2(\lambda + b^2)^{\frac{1}{2}}$. By Theorem 2.3, unless (3.10) holds, N is exponentially bounded.

It is noted that if $\Sigma = \sigma^2 I$, then the T^2-test takes the form of the general F-test. A similar result for the F-test is obtained.

Other tests. For the case $p = 1$, similar results may be obtained for other tests, such as the tests for different values of σ^2, δ and β_1 (when $\sigma^2 = 1$), where $\delta = \beta_1/\sigma$ and $\sigma^2 = \Sigma$. For these tests, L_n can be written as n times a function of W_n, $U_n/W_n^{\frac{1}{2}}$, and U_n, respectively.

REMARK 1. Exponential boundedness of N implies the finiteness a.s. of N. For the latter to be true the finite mgf assumption of $||e||$ in Assumption B′ is not needed (see Theorems 2.2 and 2.3).

REMARK 2. Assumptions B′ and C′ in Section 3 are unnecessarily restrictive. There are cases where N is exponentially bounded with neither condition in the Assumption B′. In Assumption C′, it suffices that $\bar{x}_n$ and K_n have special convergent subsequences (cf. Example 6.1 in [11]). The results for the case without the finite moment assumption will be reported separately.

REMARK 3. The distribution of (3.10) may be termed suspect (see [18], page 1710), i.e., we suspect that this distribution spoils the exponential boundedness of N when $\Phi_0 = 0$. This certainly would be the case if the y_i's were i.i.d. (see [18], Theorem 2.1, page 1710). However, this is not true for the non-i.i.d. case since a counterexample can be constructed.

Acknowledgment. The author wishes to thank Professor Robert A. Wijsman for his guidance during the preparation of the author's thesis. He also thanks the referee and the Associate Editor for comments.

REFERENCES

[1] ANDERSON, T. W. (1958). *An Introduction to Multivariate Statistical Analysis.* Wiley, New York.

[2] BERK, R. H. (1970). Stopping time of SPRT based on exchangeable models. *Ann. Math. Statist.* **41** 979–990.

[3] CHERNOFF, H. (1952). A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations. *Ann. Math. Statist.* **23** 439–507.

[4] CHUNG, K. L. (1968). *A Course in Probability Theory.* Harcourt, Brace and World, New York.

[5] GHOSH, B. K. (1970). *Sequential Tests of Statistical Hypotheses.* Addison-Wesley, Reading, Mass.

[6] HALL, W. J., WIJSMAN, R. A. and GHOSH, J. K. (1965). The relationship between sufficiency and invariance with applications in sequential analysis. *Ann. Math. Statist.* **36** 575–614.

[7] IFRAM, A. F. (1965). On the asymptotic behavior of densities with application to sequential analysis. *Ann. Math. Statist.* **36** 615–637.

[8] IFRAM, A. F. (1965). Hypergeometric functions in sequential analysis. *Ann. Math. Statist.* **36** 1870–1872.

[9] LAI, T. L. (1975). Termination, moments and exponential boundedness of the stopping rule for certain invariant sequential probability ratio tests. *Ann. Statist.* **3** 581–598.

10] Perng, S.-S. (1976). A note on invariant sequential probability ratio tests. Unpublished manuscript.

11] Perng, S.-S. (1977). Some extensions of a theorem of Stein on cumulative sums. *Ann. Statist.* **5** 98–109.

12] Savage, I. R. and Sethuraman, J. (1966). Stopping time of a rank-ordered sequential test based on Lehmann alternatives. *Ann. Math. Statist.* **37** 1154–1160. (Correction note, *Ann. Math. Statist.* **38** 1309.)

13] Sethuraman, J. (1970). Stopping time of a rank order sequential probability ratio test based on Lehmann alternatives—II. *Ann. Math. Statist.* **41** 1322–1333.

14] Skovgaard, H. (1966). *Uniform Asymptotic Expressions of Confluent Hypergeometric Functions and Wittaker Functions*. Jul. Gjellerups Forlag, Copenhagen.

15] Stein, C. (1946). A note on cumulative sums. *Ann. Math. Statist.* **17** 498–499.

16] Wijsman, R. A. (1970). Examples of exponentially bounded stopping time of invariant SPRT's when the model may be false. *Proc. Sixth Berkeley Symp. Math. Statist. Prob.* **1** 109–128. Univ. of California Press.

17] Wijsman, R. A. (1971). Exponentially bounded stopping time of SPRT's for composite hypotheses. *Ann. Math. Statist.* **42** 1859–1869.

18] Wijsman, R. A. (1972). A theorem on obstructive distributions. *Ann. Math. Statist.* **43** 1709–1715.

19] Wijsman, R. A. (1974). Stopping time on sequential sample from multivariate exponential families. *Multivariate Analysis III* (P. R. Krishnaiah, ed.). Academic Press, New York.

20] Wijsman, R. A. (1975). Exponentially bounded stopping time of the sequential *t*-test. *Ann. Statist.* **3** 1006–1010.

21] Wijsman, R. A. (1977). A general theorem with applications on exponentially bounded stopping time, without moment conditions. *Ann. Statist.* **5** 292–315.

Department of Mathematics
Howard University
Washington, D.C. 20001