The Annals of Mathematical Statistics
1968, Vol. 39, No. 3, 1048–1056

BOUNDS ON THE SAMPLE SIZE DISTRIBUTION FOR A CLASS OF INVARIANT SEQUENTIAL PROBABILITY RATIO TESTS[1]

BY R. A. WIJSMAN

University of Illinois

1. Introduction and summary. In a previous paper [8] the asymptotic behavior of a class of invariant sequential probability ratio tests was studied to the extent of establishing termination with probability one. In this paper, under somewhat stronger conditions, certain bounds on the distribution of sample size will be obtained.

The statistical framework is as follows: $Z_1, Z_2, \cdots$ is a sequence of independent and identically distributed (iid) random vectors with values in Euclidean k-space E^k. The common distribution is at first assumed to belong to the family $\mathfrak{N}$ of all nondegenerate k-variate normal distributions. On E^k there acts a group G^* of affine transformations (precise assumptions on G^* are given in Section 2). Let $V_1, V_2, \cdots$ be a maximal invariant sequence obtained from $Z_1, Z_2, \cdots$ under the application of G^*. Since every transformation in G^* sends a member of $\mathfrak{N}$ into a member of $\mathfrak{N}$, G^* also acts on $\mathfrak{N}$. Let γ be a maximal invariant function on $\mathfrak{N}$, then the joint distribution of $V_1, V_2, \cdots$ depends only on γ.

Denote the distribution of $(V_1, \cdots, V_n)$ by $P_{n\gamma}$. Let γ_1, γ_2 be two distinct values of γ, let $r_n = dP_{n\gamma_2}/dP_{n\gamma_1}$ (set equal to ∞ wherever $P_{n\gamma_2}$ is not absolutely continuous with respect to $P_{n\gamma_1}$) and put $R_n = r_n(V_1, \cdots, V_n)$. Then an invariant sequential probability ratio test, based on the sequence $\{R_n\}$, is defined by choosing stopping bounds $B < A$, letting the stopping variable N be the smallest n such that $R_n \leqq B$ or $\geqq A$, and accepting γ_1 or γ_2 according as $R_N \leqq B$ or $\geqq A$.

Once the sequence $\{R_n\}$ has been defined we are at liberty to study its behavior when the actual common distribution of the Z_n is not necessarily on the orbit of γ_1 or of γ_2 or, for that matter, is not even a member of $\mathfrak{N}$. Still assuming $Z_1, Z_2, \cdots$ to be iid, the common distribution P will be assumed to be a member of $\mathcal{P}$, to be defined later, where $\mathcal{P} \supset \mathfrak{N}$. The joint distribution of $Z_1, Z_2, \cdots$ will also be denoted P. The object is to establish a bound on $P(N > n)$ as a function of n, for each $P \in \mathcal{P}$.

Results on sample size distribution of invariant sequential probability ratio tests are very scarce. Ifram [5] considered a certain class of problems and obtained an exponential bound on $P(N > n)$, where P is a member of the original model and should not belong to a certain exceptional set of distributions for which no results could be obtained. Sacks [6] also obtained an exponential bound in the case of the sequential t-test (as a by-product of other results), again

Received 29 June 1967.
[1] This work was supported in part by the National Science Foundation under Grant GP-3814.

excluding certain exceptional P's. Savage and Sethuraman [7] obtained an exponential bound in a nonparametric problem, a sequential rank test. They allowed P to be outside the original model, but again for a certain set of exceptional P's no results could be obtained.

In the present paper we shall establish, under Assumptions A and B (Section 2), an exponential bound of the form $P(N > n) < c\rho^n$ for some $\rho < 1$, except, again, for P's in a certain set. For these exceptional P's, however, we also obtain results, even though weaker, of the form $P(N > n) < c\rho^{n^{1/3}}$ (under an additional assumption on the function Φ of Section 2). This result is still strong enough to assert the existence of all moments of N, although not the existence of a moment generating function as is the case if $P(N > n)$ has an exponential bound. Thus, our results are more general and stronger in some respects than those of Ifram [5], but Ifram does give a precise value of the smallest possible ρ in the exponential bound, whereas we have nothing comparable to offer.

2. Assumptions and a basic lemma. We shall now give the precise conditions on G^* and $\mathcal{P}$. Those on G^* are the same as in [8], whereas the conditions on $\mathcal{P}$ are more restrictive than those on $\mathcal{F}$ in [8].

ASSUMPTION A. $G^* = GH$ where (i) G is a Lie subgroup of the real general linear group $GL(k, R)$; (ii) G is closed in $GL(k, R)$ and $\dim G \geqq 1$; (iii) H is a group of translations of E^k with k-vectors b, the totality of vectors b constituting a subspace invariant under G; (iv) each transformation $g^* = (C, b)$, $C \,\varepsilon\, G$, $b \,\varepsilon\, H$, transforms $(Z_1, Z_2, \cdots)$ according to $Z_n \to CZ_n + b$, $n = 1, 2, \cdots$.

ASSUMPTION B. $\mathcal{P}$ is the family of all distributions P such that (i) $Z_1, Z_2, \cdots$ are iid, (ii) the components Z_{1j}, $1 \leqq j \leqq k$, of Z_1 satisfy $E_P \exp tZ_{1j}^2 < \infty$ for t in an interval about 0, (iii) if M is a k^2 symmetric matrix and b a k-vector then $P(Z_1'MZ_1 + b'Z_1 = \text{constant}) = 1$ implies $M = 0$, $b = 0$.

For $P \,\varepsilon\, \mathfrak{N}$ a sufficient statistic based on $(Z_1, \cdots, Z_n)$ is $(\bar{Z}_n, S_n)$, where $\bar{Z}_n = (1/n) \sum_1^n Z_i$ and $S_n = (1/n) \sum_1^n (Z_i - \bar{Z}_n)(Z_i - \bar{Z}_n)'$. The sample covariance matrix S_n takes, with probability one, values in a space $\mathcal{S}$ of k^2 positive definite matrices. We may regard $\mathcal{S}$ as a subset of a Euclidean space of $k(k + 1)/2$ dimensions. We may write $E^k = E_1 \times E_2$, and choose E_2 to consist of the vectors $b \,\varepsilon\, H$. The coordinate system may be chosen so that E_1 is spanned by the first l $(0 \leqq l \leqq k)$ coordinate axes. For any vector $v \,\varepsilon\, E^k$ let $v^{(1)}$ be its projection on E_1, so that the components of $v^{(1)}$ are the first l components of v. Thus we have e.g. $Z_n^{(1)}$, $\bar{Z}_n^{(1)}$.

After applying an invariance reduction on $Z_1, Z_2, \cdots$ we may further reduce by sufficiency. The same result is obtained by reversing the order in which these two reductions are applied (more detail on this in [8]). Thus we may apply G^* to the sequence of sufficient statistics $\{(\bar{Z}_n, S_n)\}$ on which it acts according to $\bar{Z}_n \to C\bar{Z}_n + b$, $S_n \to CS_nC'$. We may apply H first, obtaining the sequence of maximal invariants $X_n = (\bar{Z}_n^{(1)}, S_n)$, taking values in a space $\mathfrak{X} = E_1 \times \mathcal{S}$ of points $x = (z, s)$. $\mathfrak{X}$ is a subspace of E^q, where $q = l + (k(k + 1)/2)$. For any $P \,\varepsilon\, \mathcal{P}$ let μ and Σ be the mean and covariance matrix of Z_1 (so μ and Σ are functions of P). Put $\theta = (\mu^{(1)}, \Sigma)$, then θ also takes its values in $\mathfrak{X}$.

There still remains the difficult step of obtaining a maximal invariant in $\mathfrak{X}$ under the action of the group G, and, from this, an expression for R_n. The details will not be given here since we shall not need them. Instead, we shall from [8] (essentially Lemma 3 and (3.11) in [8]) take the main result that is needed and shall state this as the following:

LEMMA 2.1. *There exists on* $\mathfrak{X}$ *a continuous function* Φ *such that for any open set* V *on which* $\|z\|$, $\operatorname{tr} s$ *and* $\operatorname{tr} s^{-1}$ *are bounded there is a positive constant* K *(depending on* V *but not on* n*) such that* $X_n \varepsilon V$ *implies*

$$(2.1) \qquad |\ln R_n - n\Phi(X_n)| < d \ln n + K,$$

in which $d = \dim G$.

Results on the behavior of R_n will follow from the behavior of $\Phi(X_n)$ and the approximation given by (2.1).

3. Exponential bounds. It will be convenient in this section to make a few definitions that will be used repeatedly. We shall say that a sequence $\{p_n\}$ of probabilities is *exponentially bounded* if there exists $c > 0$ and $\rho < 1$ such that $p_n < c\rho^n$ for $n = 1, 2, \cdots$. If $\{X_n\}$ is any sequence of random vectors (not necessarily the sequence of Section 2), we shall say that X_n *converges exponentially to* x, written $X_n \to_{\exp} x$, if for every neighborhood V of x the sequence $\{P(X_n \not\varepsilon V)\}$ is exponentially bounded. It will also be convenient to develop a small amount of calculus of exponential convergence.

LEMMA 3.1. *Let* X_{nj} *be the components of* X_n, x_j *the components of* x, *then* $X_n \to_{\exp} x$ *if and only if* $X_{nj} \to_{\exp} x_j$ *for every* j.

LEMMA 3.2. *If* f *is a continuous function from* E^p *to* E^q *and* $X_n \varepsilon E^p$, *then* $X_n \to_{\exp} x$ *implies* $f(X_n) \to_{\exp} f(x)$.

These lemmas are very reminiscent of similar statements for convergence in probability. The proofs are easy and will be omitted.

We shall rely heavily on the following result, which is part (the easy part) of a stronger result of Chernoff ([1], Theorem 1). (Under an additional assumption on the common distribution of the random variables the result also follows from Theorem 6 in [2].)

LEMMA 3.3. *If* $U_1, U_2, \cdots$ *is a sequence of* iid *real valued random variables such that* $E \exp tU_1 < \infty$ *for* t *in some interval about* 0, *then* $(1/n)\sum_1^n U_i \to_{\exp} EU_1$.

We shall show now, under Assumption B (Section 2), that for the sequence $\{X_n\}$ of Section 2 and any $P \varepsilon \mathcal{P}$ we have

$$(3.1) \qquad X_n \to_{\exp} \theta$$

in which $\theta = \theta(P)$. Assumption B implies that $E_p \exp tZ_{1j} < \infty$ (for t in an interval about 0) so that, by Lemma 3.3, $\bar{Z}_{nj} \to_{\exp} \mu_j$ for every j. From Lemma 3.1. it follows then that $\bar{Z}_n \to_{\exp} \mu$ and $\bar{Z}_n^{(1)} \to_{\exp} \mu^{(1)}$. It can be checked easily that Assumption B also implies $E_p \exp tZ_{1j}Z_{1j'} < \infty$ for all j, j'. Therefore, all $k(k+1)/2$ components of the matrix $(Z_1 - \mu)(Z_1 - \mu)'$, considered as a vector, have a moment generating function. We write now

$$(3.2) \qquad S_n = (1/n)\sum_1^n (Z_i - \mu)(Z_i - \mu)' - (\bar{Z}_n - \mu)(\bar{Z}_n - \mu)'.$$

Applying Lemmas 3.3 and 3.1, we find that the first term converges exponentially to Σ. Recalling $\bar{Z}_n - \mu \rightarrow_{\text{exp}} 0$ and applying Lemmas 3.1 and 3.2, we have $(\bar{Z}_n - \mu)(\bar{Z}_n - \mu)' \rightarrow_{\text{exp}} 0$. Several more applications of Lemma 3.1 and 3.2 yield (3.1).

THEOREM 3.1. *If $P \varepsilon \mathcal{P}$ and $\theta = \theta(P)$ is such that $\Phi(\theta) \neq 0$, then there exists $c > 0$ and $\rho < 1$ such that $P(N > n) < c\rho^n$, i.e. $\{P(N > n)\}$ is exponentially bounded.*

PROOF. Suppose $\Phi(\theta) > 0$, the case < 0 being analogous. Since Φ is continuous, there is a neighborhood V of θ, satisfying the conditions of Lemma 2.1, and a constant $\delta > 0$ such that $\Phi(x) > \delta$ for all $x \varepsilon V$. By (2.1) we have that $X_n \varepsilon V$ implies that $\ln R_n > n\Phi(X_n) - d \ln n - K$ and since $\Phi(X_n) > \delta$ we have $\ln R_n > n\delta - d \ln n - K$. The right hand side of the last inequality exceeds $\ln A$ if $n > n_0$, for some n_0. In other words, if $n > n_0$ and $X_n \varepsilon V$ then stopping must have occurred by stage n. Therefore, if $n > n_0$, $P(N > n) \leqq P(X_n \not\varepsilon V)$. Since $X_n \rightarrow_{\text{exp}} \theta$, the sequence $\{P(X_n \not\varepsilon V)\}$ is exponentially bounded, and then so is $\{P(N > n)\}$.

4. A bound in the exceptional case.

The proof for the exponential bound given in Section 3 breaks down for distributions P with $\theta = \theta(P)$ satisfying $\Phi(\theta) = 0$. It is much harder in this case to get any bound at all. In fact, we shall have to make another assumption, this time on the function Φ, in order to get anywhere. In general, Φ need not possess continuous first order partial derivatives in every point. We know only (see [8]) that a derivative exists in each direction in each point, and that in any given point the directional derivative is not identically equal to 0 (Lemma 4 in [8]). On the other hand, in the various known examples Φ is always a very nice, i.e. analytic, function (see e.g. [4] in which the function $h(\theta_2, \cdot) - h(\theta_1, \cdot)$ takes the place of our function Φ). It is therefore not unreasonable to impose a certain smoothness on Φ.

THEOREM 4.1. *Let $P \varepsilon \mathcal{P}$ be such that $\theta = \theta(P)$ satisfies $\Phi(\theta) = 0$ and assume that Φ at θ has continuous first and second partial derivatives. Then there exists $c > 0$ and $\rho < 1$ such that*

$$P(N > n) < c\rho^{n^{1/3}}. \tag{4.1}$$

Before presenting the proof, we shall introduce some useful notation, give a sketch of the idea of the proof, and get some of the essential steps in the proof out of the way in the form of lemmas. In the following, μ and Σ are held fixed and $\Phi(\theta) = 0$. Denote

$$\psi = \text{grad}\, \Phi \text{ at } \theta, \tag{4.2}$$

$$S_n^* = (1/n) \textstyle\sum_1^n (Z_i - \mu)(Z_i - \mu)', \tag{4.3}$$

$$Y_n = (\bar{Z}_n - \mu, S_n^* - \Sigma), \tag{4.4}$$

$$T_n = (Z_n - \mu, (Z_n - \mu)(Z_n - \mu)' - \Sigma), \tag{4.5}$$

then $T_1, T_2, \cdots$ are iid vectors, $ET_1 = 0$, and

$$nY_n = \textstyle\sum_1^n T_i. \tag{4.6}$$

Analogously to (4.4) and (4.5) define

$$(4.7) \qquad Y_n^{(1)} = (\bar{Z}_n^{(1)} - \mu^{(1)}, S_n^* - \Sigma),$$

$$(4.8) \qquad T_n^{(1)} = (Z_n^{(1)} - \mu^{(1)}, (Z_n - \mu)(Z_n - \mu)' - \Sigma),$$

so that

$$(4.9) \qquad nY_n^{(1)} = \sum_1^n T_i^{(1)}.$$

Now define U_n by

$$(4.10) \qquad U_n = \psi' T_n^{(1)}$$

then $U_1, U_2, \cdots$ is a sequence of iid real valued random variables with $EU_1 = 0$. Furthermore, since $\psi \neq 0$ (from [8], Lemma 4), and using Assumption B (iii), we have $U_1 \not\equiv 0$. Comparing (4.9) and (4.10) we see

$$(4.11) \qquad \psi' n Y_n^{(1)} = \sum_1^n U_i.$$

Further, define

$$(4.12) \qquad W_n = \bar{Z}_n - \mu$$

so that by (3.2) and (4.3),

$$(4.13) \qquad S_n = S_n^* - W_n W_n'.$$

Lastly, define

$$(4.14) \qquad X_n^* = (\bar{Z}_n^{(1)}, S_n^*)$$

so that from (4.7) we have

$$(4.15) \qquad X_n^* - \theta = Y_n^{(1)}.$$

We are interested in the process $\{R_n\}$ that is stopped as soon as it moves outside the fixed bounds A and B. In view of (2.1) we shall replace the process $\{\ln R_n\}$ by $\{n\Phi(X_n)\}$, at the same time letting the stopping bounds widen essentially at the rate $\ln n$. If we could replace $\Phi(X_n)$ by $\psi'(X_n - \theta)$ and if we could replace X_n by X_n^*, defined in (4.14), then in view of (4.11) we would have replaced $\{n\Phi(X_n)\}$ by the random walk $\{\sum_1^n U_i\}$. The necessary modification is nontrivial and yields another term $n\|Y_n\|^2$. This follows from Lemma 4.2, and in Lemma 4.3 it is shown how to cope with this extra term. This involves, among other things, replacing the stopping bounds by others that are widening somewhat faster than at the rate $\ln n$, but slower than $n^{\frac{1}{2}}$. Still, the main interest is in the study of the random walk $\{\sum_1^n U_i\}$ between those widening bounds. Since this forms the most essential part of the proof, its contribution is presented in the first lemma below.

LEMMA 4.1. *Let* $x_1, x_2, \cdots$ *be a sequence of* iid *real valued random variables with* $Ex_1 = 0$, $Ex_1^2 = 1$, *and let* $s_n = x_1 + \cdots + x_n$ *be their partial sum. Let* $\{a_n\}$ *be a sequence of numbers such that* $a_n \to \infty$ *and* $a_n n^{-\frac{1}{2}} \to 0$. *Then there exists*

$c > 0$ and $\rho < 1$ such that

$$P\{\max_{i \leqq n} |s_i| \leqq a_n\} < c\rho^{n/a_n^2}, \qquad n = 1, 2, \cdots. \tag{4.16}$$

PROOF. From Erdös and Kac [3] we take the following result: Let $\{x(t), t \geqq 0\}$ be standard Brownian motion and let $a > 0$. Then

$$\lim_{n \to \infty} P\{\max_{i \leqq n} |s_i| \leqq an^{\frac{1}{2}}\} = P\{\max_{t \leqq 1} |x(t)| \leqq a\}. \tag{4.17}$$

Denote

$$\rho(a) = P\{\max_{t \leqq 1} |x(t)| \leqq a\} < 1. \tag{4.18}$$

Let $\alpha_n^2 = [a_n^2] + 1$, then α_n^2 is an integer and $\alpha_n > a_n$ so that

$$P\{\max_{i \leqq \alpha_n^2} |s_i| < 2a_n\} \leqq P\{\max_{i \leqq \alpha_n^2} |s_i| < 2\alpha_n\}.$$

Since $\alpha_n \to \infty$, the right hand side of the above inequality has, by (4.17), the limit $\rho(2)$ given by (4.18) with $a = 2$. Choosing any ρ_1 such that $\rho(2) < \rho_1 < 1$, it follows that there is n_1 such that $n > n_1$ implies

$$P\{\max_{i \leqq \alpha_n^2} |s_i| < 2a_n\} < \rho_1. \tag{4.19}$$

In the following it will be assumed that $n > n_1$. If (4.16) is true with this restriction, it is also true without it.

Denote $b_n = [n/\alpha_n^2]$. From the definition of α_n and the assumption that $a_n n^{-\frac{1}{2}} \to 0$ we deduce that $b_n \to \infty$. We now consider the time interval from 0 to n decomposed in b_n blocks of length α_n^2 each, plus whatever is left over. (We shall ignore what happens in the left-over piece.) For $j = 1, \cdots, b_n$ denote

$$A_j = \{|s_i| < a_n,\ (j-1)\alpha_n^2 \leqq i \leqq j\alpha_n^2\} \quad (\text{define } s_0 = 0),$$

$$B_j = \{|s_i| - |s_{(j-1)\alpha_n^2}| < 2a_n,\ (j-1)\alpha_n^2 \leqq i \leqq j\alpha_n^2\},$$

then $\{|s_{(j-1)\alpha_n^2}| < a_n\} \cap A_j \subset B_j$ so that *a fortiori* $A_{j-1} \cap A_j \subset B_j$ and therefore $\cap A_j \subset \cap B_j$. We have $P\{\max_{i \leqq n} |s_i| < a_n\} \leqq P\{\max_{i \leqq b_n \alpha_n^2} |s_i| < a_n\} = P \cap A_j \leqq P \cap B_j$. Since $x_1, x_2, \cdots$ are iid, the B_j are independent and have equal probabilities, their common value being equal to the left hand side of (4.19). Therefore, $P \cap B_j < \rho_1^{b_n}$. From the definitions of b_n and α_n, by taking ρ slightly larger than ρ_1, (4.16) follows.

LEMMA 4.2. *We can choose a neighborhood V of θ and $b > 0$ such that $X_n \varepsilon V$ implies*

$$|\Phi(X_n) - \psi' Y_n^{(1)}| < b\|Y_n\|^2, \tag{4.20}$$

where ψ, Y_n and $Y_n^{(1)}$ are given by (4.2), (4.4) *and* (4.7), *respectively.*

PROOF. First we write

$$|\Phi(X_n) - \psi' Y_n^{(1)}| \leqq |\Phi(X_n) - \psi'(X_n - \theta)| + |\psi'(X_n - \theta) - \psi' Y_n^{(1)}|. \tag{4.21}$$

The second term on the right hand side can be bounded as follows. Let the last $k(k+1)/2$ components of ψ be denoted $\psi_{jj'}$, $1 \leqq j \leqq j' \leqq k(k+1)/2$, and let W_{nj}, $1 \leqq j \leqq k$, be the components of W_n (defined in (4.12)). From (4.15) we see that $X_n - \theta - Y_n^{(1)} = X_n - X_n^*$. The latter equals $(0, S_n - S_n^*)$ by (4.14), and $S_n - S_n^* = -W_n W_n'$ by (4.13). Hence

$$|\psi'(X_n - \theta) - \psi' Y_n^{(1)}| = \sum_{j \leqq j'} \psi_{jj'} W_{nj} W_{nj'} < c_1 \|W_n\|^2$$

and since $\|W_n\|^2 < \|Y_n\|^2$ we have

$$|\psi'(X_n - \theta) - \psi' Y_n^{(1)}| < c_1 \|Y_n\|^2. \tag{4.22}$$

For the first term on the right hand side in (4.21) we have a bound

$$|\Phi(X_n) - \psi'(X_n - \theta)| < c_2 \|X_n - \theta\|^2 \tag{4.23}$$

provided V is taken sufficiently small. Furthermore, using (4.13),

$$\|X_n - \theta\|^2 = \|\bar{Z}_n^{(1)} - \mu^{(1)}\|^2 + \|S_n^* - \Sigma - W_n W_n'\|^2$$
$$\leqq \|W_n\|^2 + 2 \|S_n^* - \Sigma\|^2 + 2 \|W_n W_n'\|^2$$

in which the symmetric matrix $W_n W_n'$ is regarded as a $(k(k+1)/2)$-dimensional vector. One can easily check that $\|W_n W_n'\|^2 \leqq \|W_n\|^4$. By further restricting V, if necessary, we can make sure that $\|W_n\| \leqq 1$ for $X_n \varepsilon V$, so that $\|W_n\|^4 \leqq \|W_n\|^2$. We have then $\|X_n - \theta\|^2 \leqq 3 \|W_n\|^2 + 2 \|S_n^* - \Sigma\|^2 < 3 \|Y_n\|^2$. Substitution of this into (4.23), and (4.22) and (4.23) into (4.21) yields (4.20).

LEMMA 4.3. *Let $\{t_n\}$ be a sequence of numbers such that $t_n \to \infty$ and $t_n = O(n^{\frac{1}{3}})$. Then there exists $c > 0$ and $\rho < 1$ such that*

$$P\{n \|Y_n\|^2 > t_n\} < c t_n^{-\frac{1}{2}} \rho^{t_n}, \tag{4.24}$$

in which Y_n is defined in (4.4).

PROOF. Letting T_{nj}, $1 \leqq j \leqq q$, be the components of T_n defined in (4.5), we can write $n \|Y_n\|^2$, in view of (4.6), as $\sum_1^q n\bar{T}_{nj}^2$. If this quantity is to be $> t_n$ then for some j we must have $n\bar{T}_{nj}^2 > t_n/q$, so that

$$P\{n \|Y_n\|^2 > t_n\} \leqq \sum_{j=1}^q P\{n\bar{T}_{nj}^2 > t_n/q\}. \tag{4.25}$$

Writing $a_n = (t_n/q)^{\frac{1}{2}}$ we have $a_n = O(n^{\frac{1}{6}})$. The jth term on the right hand side in (4.25) we write as

$$P\{n^{\frac{1}{2}}\bar{T}_{nj} > a_n\} + P\{n^{\frac{1}{2}}\bar{T}_{nj} < -a_n\}.$$

The bounds on these two terms are similar, so we shall only deal with the first term. Let $T_{1j}^*, T_{2j}^*, \cdots$ be a sequence of iid normal random variables with mean 0 and variance σ^2 equal to the variance of T_{1j}. From the results of Cramér [2] we know that $P\{n^{\frac{1}{2}}\bar{T}_{nj} > a_n\}/P\{n^{\frac{1}{2}}\bar{T}_{nj}^* > a_n\}$ has a finite limit as $n \to \infty$. The rest now follows easily from the fact that $P\{n^{\frac{1}{2}}\bar{T}_{nj}^* > a_n\} < c_3 a_n^{-1} \exp[-a_n^2/2\sigma^2]$ and remembering that $a_n^2 = t_n/q$.

The proof of Theorem 4.1 will now be given. A last word of explanation seems in order. In the course of the proof there appear certain terms depending on n, e.g. $P(X_n \not\varepsilon V)$, whose values become small only when n becomes large. Yet, in a bound for $P(N > m)$ all terms with $n \leqq m$ contribute. To get around this difficulty we shall take an integer m (thought of as large) and study the process for n between m and $2m$, ignoring what the process does before m.

PROOF OF THEOREM 4.1. Let V be a neighborhood of θ chosen according to Lemma 4.2. Let a positive integer m be fixed. Define the events

$$A_m = \{X_n \not\in V \quad \text{for some} \quad m \leqq n \leqq 2m\},$$

$$B_m = \{X_n \in V \quad \text{and} \quad B < R_n < A, m \leqq n \leqq 2m\},$$

$$C_m = \{X_n \in V \quad \text{and} \quad |n\Phi(X_n)| < d \ln n + K_1 , m \leqq n \leqq 2m\}$$

in which $K_1 = K + \max(\ln A, -\ln B)$, and K is defined in Lemma 2.1. From (2.1) it follows that $B_m \subset C_m$. Furthermore, the event $(N > 2m) = \{B < R_n < A, n \leqq 2m\}$ implies $A_m \cup B_m$, so that

$$P(N > 2m) \leqq PA_m + PC_m . \tag{4.26}$$

Define the event

$$D_m = \{\textstyle\sum_1^n U_i < bn \|Y_n\|^2 + d \ln n + K_1 , \qquad m \leqq n \leqq 2m\}$$

with U_i defined in (4.10), Y_n in (4.4) and b in Lemma 4.2. By (4.11) and Lemma 4.2 we have $C_m \subset D_m$ so it suffices to find a bound for PD_m. Let $\{t_n\}$ be an increasing sequence of numbers, to be determined later, such that $t_n/\ln n \to \infty$. Then there exists n_0 such that $n > n_0$ implies $t_n > d \ln n + K_1$. We shall assume in the following that $m > n_0$. Define the events

$$E_m = \{bn \|Y_n\|^2 > t_n \quad \text{for some} \quad m \leqq n \leqq 2m\},$$

$$F_m = \{|\textstyle\sum_1^n U_i| < 2t_n , m \leqq n \leqq 2m\},$$

then $D_m \subset E_m \cup F_m$, and we have

$$P(N > 2m) \leqq PA_m + PE_m + PF_m . \tag{4.27}$$

From Section 3 we know that $P(X_n \not\in V) < c\rho_1^n$ for some $c > 0$, $\rho_1 < 1$. Summing this over n from m to $2m$ we have

$$PA_m < c_1\rho_1^m. \tag{4.28}$$

We now impose on t_n the restriction $t_n = O(n^{\frac{1}{3}})$. By Lemma 4.3, with t_n replaced by t_n/b (changing the value of ρ) we have $PE_m \leqq \sum_m^{2m} c_4 t_n^{-\frac{1}{2}} \rho^{t_n} \leqq (m + 1)c_4 t_m^{-\frac{1}{2}} \rho^{t_m}$ in which use has been made of the monotonictiy of $\{t_n\}$. Since t_m goes to ∞ faster than $\ln m$, by taking ρ_2 slightly larger than ρ we can write

$$PE_m < c_2\rho_2^{t_m}. \tag{4.29}$$

If in the expression for F_m we replace t_n by its maximum value t_{2m}, we increase, if anything, the probability. Furthermore, $P\{|\sum_1^n U_1| < 2t_{2m} , m \leqq n \leqq 2m\} \leqq P\{|\sum_{m+1}^n U_i| < 4t_{2m} , m + 1 \leqq n \leqq 2m\}$ and the latter equals $P\{|\sum_1^n U_i| < 4t_{2m} , 1 \leqq n \leqq m\}$ since the U_i are iid. Thus we have

$$PF_m \leqq P\{|\textstyle\sum_1^n U_i| < 4t_{2m} , 1 \leqq n \leqq m\}. \tag{4.30}$$

Let the variance of U_1 be σ^2 (>0 since $U_1 \not\equiv 0$), then the right hand side of (4.30) corresponds to the left hand side of (4.16) if we set $4t_{2m}/\sigma = a_m$. Thus a bound on PF_m is given by the right hand side of (4.16) in which the exponent of ρ is $m\sigma^2/16t_{2m}^2$. This can be written as

$$PF_m < c_3\rho^{m/t_{2m}}. \tag{4.31}$$

After substituting (4.28), (4.29) and (4.31) into (4.27) we see that the exponential bound on the first term on the right hand side in (4.27) goes to zero faster than the bounds on the second and third term. If we increase the rate with which $t_n \to \infty$, the bound on PE_m becomes better, on PF_m worse. A balance is struck by choosing $t_n = cn^{\frac{1}{3}}$ in which case $P(N > 2m) < c_4\rho^{m^{1/3}}$. By suitable redefinition of the constants we obtain (4.1).

REFERENCES

[1] Chernoff, Herman (1952). A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations. *Ann. Math. Statist.* **23** 493–507.

[2] Cramér, Harald (1938). Sur un nouveau théorème-limite de la théorie des probabilités. *Actualités Sci. Indust.* no. 736, 5–23.

[3] Erdös, P. and Kac, M. (1946). On certain limit theorems of the theory of probability. *Bull. Amer. Math. Soc.* **52** 292–302.

[4] Ifram, Adnan F. (1965). On the asymptotic behavior of densities with applications to sequential analysis. *Ann. Math. Statist.* **36** 615–637.

[5] Ifram, Adnan F. (1966). On the sample size and simplification of a class of sequential probability ratio tests. *Ann. Math. Statist.* **37** 425–434.

[6] Sacks, J. (1965). A note on the sequential t-test. *Ann. Math. Statist.* **36** 1867–1869.

[7] Savage, I. Richard and Sethuraman, J. (1966). Stopping time of a rank-order sequential probability ratio test based on Lehmann alternatives. *Ann. Math. Statist.* **37** 1154–1160.

[8] Wijsman, R. A. (1967). General proof of termination with probability one of invariant sequential probability ratio tests based on multivariate normal observations. *Ann. Math. Statist.* **38** 8–24.