The Annals of Mathematical Statistics
1972, Vol. 43, No. 5, 1709-1715

A THEOREM ON OBSTRUCTIVE DISTRIBUTIONS[1]

BY R. A. WIJSMAN

University of Illinois at Urbana-Champaign

Let N be the stopping time of a sequential probability ratio test of composite hypotheses, based on the i.i.d. sequence $Z_1, Z_2, \cdots$ with common distribution P. If for every choice of stopping bounds there exist constants $c>0$, $0<\rho<1$ such that $P\{N>n\}<c\rho^n$ $n=1,2,\cdots$, we say that N is exponentially bounded under P; otherwise P is called obstructive. A theorem is proved giving sufficient conditions for P to be obstructive. By virtue of this theorem it is possible to exhibit families of obstructive distributions in several examples, including the sequential t-test.

1. Introduction. This paper is a continuation of [4] and [5] and supplies partial answers to unresolved questions in those papers.

Let $Z_1, Z_2, \cdots$ be i.i.d. random vectors with common distribution P, and let $\mathscr{P}=\{P_\theta : \theta\in\Theta\}$ be a family of distributions, not necessarily containing P. We shall call $\mathscr{P}$ the *model*. Suppose Θ_1 and Θ_2 are two disjoint subsets of Θ and it is desired to test sequentially against each other the two (usually composite) hypotheses $H_j : \theta\in\Theta_j$, $j=1,2$. Suppose, furthermore, that there is a group of invariance transformations which is transitive over each Θ_j so that the restriction to invariant tests reduces the composite hypotheses to simple ones. In terms of these simple hypotheses let R_n be the probability ratio at the nth stage of sampling, and $L_n=\log R_n$. With an *invariant sequential probability ratio test* is meant a sequential procedure that chooses two stopping bounds l_1 and l_2 and stops sampling at the smallest positive integer n for which

$$l_1 < L_n < l_2 \tag{1.1}$$

is violated (and then accepts H_1 or H_2 according as $L_n \leqq l_1$ or $\geqq l_2$). Let N be the stopping variable thus defined.

The distribution of N depends on the stopping bounds l_1, l_2, and on P, the true distribution of the Z_i (it is emphasized once more that P is not necessarily a member of $\mathscr{P}$). We are interested in particular in the question whether the distribution of N has the following desirable property: for every l_1, l_2 there exist constants $c>0$ and $0<\rho<1$ such that

$$P\{N>n\} < c\rho^n, \qquad n=1,2,\cdots. \tag{1.2}$$

In this case we shall say that N is *exponentially bounded* (under P). If, for some

Received October 1971; revised January 3, 1972.

[1] Research supported by the National Science Foundation under Grant GP 28154.

AMS 1970 *subject classifications.* Primary 62L10; Secondary 62F05.

Key words and phrases. Sequential probability ratio test, sequential t-test, stopping time, exponentially bounded, obstructive.

P, N is not exponentially bounded we shall say that P is *obstructive*. In [4] Section 4 and [5] Section 3 an example was treated where it has been possible to give a complete characterization of the obstructive P's. On the other hand Examples 2 and 3 in [4] provided only incomplete results insofar as the existence of obstructive P's was not demostrated. Shown in each of those two examples was that if P is indeed obstructive it must belong to a certain well-defined family of degenerate distributions. For the sake of discussion it is convenient to call those distributions *suspect*. Theorem 2.1 in Section 2 in the present paper permits now to conclude that in Examples 2 and 3 of [4] certain of these suspect distributions are indeed guilty of being obstructive. This narrows the gap but does not close it since we still do not know how to deal with all the suspect distributions.

Theorem 2.1 is also applied to the sequential t-test. In [5] Section 4 a family of suspect P's was exhibited and with the help of Theorem 2.1 it is possible now to brand all the P's in this family as obstructive. It should be kept in mind, however, that in the case of the sequential t-test we only know how to deal with P's under which Z_1^2 has finite moment generating function (m.g.f.). It is not known whether there are obstructive P's among those under which Z_1^2 has infinite m.g.f.

It deserves mention perhaps that although the applications in this paper are exclusively to invariant sequential probability ratio tests, Theorem 2.1 is not inherently so restricted. Conceivably there may be applications to sequential tests of composite hypotheses that do not arise from an invariance reduction.

2. The theorem. If the model $\mathscr{P}$ is an exponential family, L_n is a function only of n and the sufficient statistic $\bar{X}_n = (1/n) \sum_1^n X_i$, where $X_1, X_2, \cdots$ are i.i.d. random vectors and the components of X_i are certain functions of Z_i. Although Theorem 2.1 is not stated in terms of an exponential family, it is formulated with that kind of an application in mind.

A central role in the theorem is played by the sequence of random variables $\Phi_n = n\Phi(\bar{X}_n)$, and the real-valued function Φ satisfies certain regularity conditions (in the applications Φ arises in a natural way but in the theorem it does not matter where Φ comes from). Φ_n is used as an approximation to L_n and is usually much easier to handle. What Theorem 2.1 proves is that under certain conditions on P and Φ, and appropriately chosen symmetric stopping bounds $-l_1 = l_2 = B$, say, the expected sample size is infinite, which implies the violation of (1.2). The method of proof reduces the problem in successive stages to a generalization of a problem first solved by Blackwell and Freedman [1] Theorem 1 (a one-dimensional generalization is given in [2] Corollary 1 to Theorem 2).

THEOREM 2.1. *Let* $X_1, X_2, \cdots$ *be* i.i.d., *their common distribution* P *being supported on* k*-dimensional Euclidean space* R^k, *such that* $EX_1'X_1 < \infty$; *put* $EX_1 = \xi$ *and* $\bar{X}_n = (1/n) \sum_1^n X_i$. *Let* $\{L_n, n = 1, 2, \cdots\}$ *be a sequence of finite and real-valued random variables such that* L_n *depends only on* $X_1, \cdots, X_n$. *For any* $B > 0$

let $N(B)$ be the smallest integer $n \geqq 1$ such that $|L_n| \geqq B$. Suppose there exists a function $\Phi: R^k \to R$ with the following properties: (1) $\Phi(\xi) = 0$; (ii) *there is a neighborhood of ξ on which Φ has continuous second partial derivatives*; (iii) *let Δ be* grad Φ *evaluated at ξ, then $P\{\Delta'(X_1 - \xi) = 0\} = 1$*; (iv) *using the notation $\Phi_n = n\Phi(\bar{X}_n)$, there exist positive numbers r and B_1 such that for every $n = 1, 2, \cdots$, $||\bar{X}_n - \xi|| < r$ implies $|L_n - \Phi_n| < B_1$. Then for some choice of B we have $EN(B) = \infty$. Hence P is obstructive.*

PROOF. By making a translation in R^k we may suppose $\xi = 0$. Condition (iii) of the hypothesis of the theorem says that with P-probability one the random walk takes place in the hyperplane $\{x \in R^k : \Delta' x = 0\}$ so that in the following we shall restrict attention to this plane. Taking into account (i): $\Phi(0) = 0$, (ii) and (iii), we can assert the existence of $r_1 > 0$ and $B_2 > 0$ such that $|\Phi(x)| < B_2||x||^2$ whenever $||x|| < r_1$. Without loss of generality we may identify r_1 with r of (iv) by choosing the latter sufficiently small. Thus, $||\bar{X}_n|| < r$ implies both $|L_n - \Phi_n| < B_1$ (from (iv)) and $|\Phi_n| < nB_2||\bar{X}_n||^2$. Putting $S_n = \sum_1^n X_i$, these implications can be written as

$$[||S_n|| < nr] \Rightarrow [|L_n - \Phi_n| < B_1, \Phi_n < (1/n)B_2||S_n||^2] . \tag{2.1}$$

Put

$$\sigma^2 = EX_1'X_1 \tag{2.2}$$

which is finite by assumption. Take any constant B_3 such that

$$B_3 > 4B_2\sigma^2 , \tag{2.3}$$

and take any integer

$$n_1 > B_3/(B_2 r^2) . \tag{2.4}$$

Since the variances of the components of X_1 are finite, and $EX_1 = 0$, as $n \to \infty$ the distribution of $S_n/n^{\frac{1}{2}}$ converges to a multivariate normal law (with mean 0). Therefore, given the positive constant

$$c = \tfrac{1}{2}(B_3/B_2)^{\frac{1}{2}} , \tag{2.5}$$

we have, for some $n_2 > n_1$,

$$P\{||S_{n_2}||/n_2^{\frac{1}{2}} < c\} > 0 . \tag{2.6}$$

Recalling the definition of $N(B)$ from the hypothesis of the theorem, and considering for any fixed $n \geqq 1$ the event $\{N(B)\} > n\}$, we have $\{N(B) > n\} \uparrow \Omega$ (= sample space) as $B \uparrow \infty$. Thus, if A is any event, we have $\{A, N(B) > n_2\} \uparrow A$ and therefore $P\{A, N(B) > n_2\} \uparrow PA$, as $B \to \infty$. Apply this to A = event on the left-hand side of (2.6); that is, define the event C (depending on B) as

$$C = \{||S_{n_2}|| < cn_2^{\frac{1}{2}}, N(B) > n_2\} , \tag{2.7}$$

then, taking (2.6) into account, we see that $PC > 0$ for sufficiently large B. Choose any $B > B_1 + B_3$ such that $PC > 0$.

Define N_1 as the smallest integer $n > n_2$ such that $|L_n| \geqq B_1 + B_3$, then on C (where $N(B) > n_2$) we have $N(B) \geqq N_1$ since $B > B_1 + B_3$. This accounts for the second inequality in

$$EN(B) \geqq \int_C N(B)\, dP \geqq \int_C N_1\, dP . \tag{2.8}$$

Now define N_2 as the smallest integer $n > n_2$ such that $||S_n - S_{n_2}|| \geqq cn^{\frac{1}{2}}$. For any $n > n_2$ we have on C (where $||S_{n_2}|| < cn_2^{\frac{1}{2}}$) the implications

$$[||S_n - S_{n_2}|| < cn^{\frac{1}{2}}] \Rightarrow [||S_n|| < c(n^{\frac{1}{2}} + n_2^{\frac{1}{2}})] \Rightarrow [||S_n|| < 2cn^{\frac{1}{2}}] . \tag{2.9}$$

Furthermore, from $n > n_2 > n_1$, (2.4) and (2.5) it follows that

$$[||S_n|| < 2cn^{\frac{1}{2}}] \Rightarrow [||S_n|| < nr] . \tag{2.10}$$

Then from (2.1), (2.5) and (2.10) we obtain

$$\begin{aligned} [||S_n|| < 2cn^{\frac{1}{2}}] \Rightarrow [|L_n| &< B_1 + (1/n)B_2||S_n||^2 \\ &< B_1 + (1/n)B_2(nB_3/B_2) = B_1 + B_3] . \end{aligned} \tag{2.11}$$

Recalling the definitions of N_1 and N_2, from (2.9) and (2.11) we deduce that on C, $[N_2 > n] \Rightarrow [N_1 > n]$, for every n, so that $N_2 \leqq N_1$. Therefore,

$$\int_C N_1\, dP \geqq \int_C N_2\, dP = PC\, EN_2 , \tag{2.12}$$

the equality in (2.12) following from the fact that by (2.7) C is defined only in terms of $X_1, \cdots, X_{n_2}$ whereas N_2 depends only on the X_i for $i > n_2$. Since $PC > 0$, in order to show $EN(B) = \infty$ it is sufficient, by virtue of (2.8) and (2.12), to show $EN_2 = \infty$.

Define N_3 as the smallest integer $n \geqq 1$ such that $||S_n|| \geqq c(n_2 + n)^{\frac{1}{2}}$ then it is easily seen from the i.i.d. character of the X_i that $n_2 + N_3$ has the same distribution as N_2. Hence it suffices to show $EN_3 = \infty$. Finally, define N_4 as the smallest integer $n \geqq 1$ such that $||S_n|| \geqq cn^{\frac{1}{2}}$, then clearly $N_4 \leqq N_3$. We shall prove now $EN_4 = \infty$.

For convenience of notation we shall write N instead of N_4. Let the components of X_i be $X_{ij}, j = 1, \cdots, k$, then the components of S_n are $S_{nj} = \sum_{i=1}^n X_{ij}$ so that $E||S_N||^2 = E \sum_{j=1}^k S_{Nj}^2 = \sum_{j=1}^k E(\sum_{i=1}^N X_{ij})^2$. Now suppose EN were $< \infty$. By the second of Wald's equations [3] (3.1), proved in great generality by Chow, Robbins, and Teicher [2] (17), we have for each $j = 1, \cdots, k$ (remembering that $EX_{1j} = 0$) $E(\sum_{i=1}^N X_{ij})^2 = EN\, EX_{1j}^2$ so that $E||S_N||^2 = EN \sum_{j=1}^k EX_{1j}^2$ which can be written, using (2.2), as

$$E||S_N||^2 = \sigma^2 EN . \tag{2.13}$$

On the other hand, $||S_N||^2 \geqq c^2 N$ by the definition of N, so that

$$E||S_N||^2 \geqq c^2 EN . \tag{2.14}$$

Comparing (2.13) and (2.14) leads to $\sigma^2 \geqq c^2 = B_3/(4B_2)$, using (2.5). But this contradicts (2.3). □

3. Applications. (a) *Sequential t-test* (see [5] Section 4). Under the model

$\mathscr{P}$ the Z_i are $N(\zeta, \sigma^2)$ and the hypotheses H_j are $\zeta/\sigma = \gamma_j$, $j = 1, 2$, with the γ_j given and distinct. The X_i may be taken as $(Z_i^2, Z_i)'$ so that Φ may be restricted to the half plane $\{x = (x_1, x_2)' : x_1 > 0\}$. It was shown in [5] Section 4 that

$$\Phi(x) = \beta(\gamma_2 x_2 x_1^{-\frac{1}{2}}) - \beta(\gamma_1 x_2 x_1^{-\frac{1}{2}}) - \tfrac{1}{2}\gamma_2^2 + \tfrac{1}{2}\gamma_1^2 \tag{3.1}$$

with $\beta(u) = \frac{1}{2}u\alpha(u) + \log \alpha(u)$ and $\alpha(u) = \frac{1}{2}[u + (u^2 + 4)^{\frac{1}{2}}]$, $-\infty < u < \infty$. It is easily seen that Φ has partials of all orders, so that (ii) of Theorem 2.1 is satisfied. Condition (iv) was proved to hold in [5] Section 4. It was shown there also that (iii) implies either $P\{Z_1 = 0\} = 1$ which case we exclude for reasons explained in [5] Section 4, or P is one of the following two-point distributions:

$$P\{Z_1 = (\sigma^2 + \zeta^2)^{\frac{1}{2}}\zeta^{-1}[(\sigma^2 + \zeta^2)^{\frac{1}{2}} \pm \sigma]\} = \tfrac{1}{2}[1 \mp \sigma(\sigma^2 + \zeta^2)^{-\frac{1}{2}}] \tag{3.2}$$

in which $\zeta \neq 0$ and $\sigma > 0$ are the mean and standard deviation of Z_1 under the true distribution P. In order that (i) of Theorem 2.1 be satisfied we deduce from (3.1) that $\eta = \zeta/(\sigma^2 + \zeta^2)^{\frac{1}{2}}$ has to solve the equation

$$\beta(\gamma_2\eta) - \beta(\gamma_1\eta) - \tfrac{1}{2}\gamma_2^2 + \tfrac{1}{2}\gamma_1^2 = 0\,. \tag{3.3}$$

With the unique solution for η from (3.3) the family (3.2) is a one-parameter family of two-point distributions, obtained from a single one by scale transformations (it is assumed here that $\gamma_1^2 \neq \gamma_2^2$ for otherwise the solution of (3.3) is $\eta = 0$ which leads to the excluded case $P\{Z_1 = 0\} = 1$). From Theorem 2.1 in [5] we know that among all P under which Z_1^2 has finite m.g.f. only the family (3.2) with η satisfying (3.3) is suspect, and from Theorem 2.1 in the present paper we see that these distributions are indeed obstructive. Thus, in the case of the sequential t-test the distributions P under which Z_1^2 has finite m.g.f. are now completely classified into those that are obstructive and those under which N is exponentially bounded.

(b). *Tests about the mean vector or about the characteristic roots of the covariance matrix.* In [4] Examples 2 and 3 (Sections 5, 6) the model $\mathscr{P}$ prescribes the Z_i to be d-variate $N(\zeta, \Sigma)$. In Example 2 $d = 2$, $\zeta = 0$ and, for $j = 1, 2$, H_j specifies the values of the characteristic roots of Σ^{-1}. In Example 3 d is arbitrary, $\Sigma =$ identity matrix, and H_j specifies the value of $||\zeta||$. (In the latter example the existence of obstructive distributions was demonstrated by R. H. Berk (private communication).) As shown in [4] both problems—as far as investigation of exponentially bounded N is concerned—are special cases of the following problem: $X_1, X_2, \cdots$ are i.i.d. k-vectors, $|L_n - \Phi_n| < B_1$ for some constant B_1 uniformly in n, and Φ_n is given by

$$\Phi_n = ||\textstyle\sum_1^n X_i|| + a \sum_1^n ||X_i|| + bn \tag{3.4}$$

with the constants a and b depending on the specific problem. It was shown that N is exponentially bounded unless

$$P\{u'X_1 + a||X_1|| + b = 0\} = 1 \tag{3.5}$$

for some unit vector u (i.e. $||u|| = 1$). Thus, in the terminology of Section 1, distributions P of the form (3.5) are suspect. Which of those P can actually be asserted to be obstructive in the light of Theorem 2.1?

For P of the form (3.5) we may replace Φ_n in (3.4) by

$$\Phi_n = ||\textstyle\sum_1^n X_i|| - u' \sum_1^n X_i \tag{3.6}$$

so that

$$\Phi(x) = ||x|| - u'x\,, \qquad x \in R^k\,, \tag{3.7}$$

for some unit vector u. Obviously, Φ possesses continuous partials of all order at every x except at $x = 0$. For the application of Theorem 2.1 it is therefore necessary to exclude those P for which $\xi = 0$. (Theorem 2.1 in [5] is not applicable either to this Φ when $\xi = 0$.) From (3.7) we see that (i) of Theorem 2.1 is satisfied if and only if $\xi/||\xi|| = u$. Furthermore, we compute grad $\Phi(x) = x/||x|| - u$ so that in (iii) $\Delta = \xi/||\xi|| - u$, which $= 0$ if (i) is satisfied. From Theorem 2.1 we conclude then that if P is such that the components of X_1 have finite variances and $\xi \neq 0$ is in the direction of u, then P is obstructive. On the other hand, if $\xi \neq 0$ is not in the direction of u and the components of X_1 have finite m.g.f., then Theorem 2.1 of [5] tells us that N is exponentially bounded (one could say that these P's have been "exonerated"). Thus, among the suspect P's (given by (3.5)) we have delineated two subfamilies, one obstructive, the other not. However, not all suspect P's have thus been classified. For instance, we do not know whether suspect P's under which some of the components of X_1 have infinite variance are obstructive. Also, the special case of suspect P's with $\xi = 0$ still has to be dealt with (in Example 3 of [4] it cannot happen that P is suspect and $\xi = 0$, but in Example 2 it can happen).

After this paper was written R. H. Berk noticed that suspect P's with $\xi = 0$ are indeed obstructive if the X's with P-probability one lie in a one-dimensional linear space. This can be seen as follows: In (3.5) we may suppose that u lies in the same one-dimensional linear space to which the X's are confined under P. Then (3.6) implies that N is the first passage time of a one-dimensional random walk through a single barrier. It is well known that with $\xi = 0$ we have $EN = \infty$ (this follows most easily from the first of Wald's equations: if EN were finite, then $E \sum_1^N u'X_i$ would equal $u'\xi EN = 0$). Therefore, P is obstructive. Note also that (3.5) implies that X_1 has, under P, either a one-point or a two-point distribution. Thus, in the one-dimensional case, X_1 has a finite m.g.f. under any suspect P. Therefore, in the one-dimensional case the suspect P's can be classified according to the value of $u'\xi$, with $u'\xi \geqq 0$ implying that P is obstructive, $u'\xi < 0$ that N is exponentially bounded under P. However, in dimension 2 or higher the question is still open.

Acknowledgment. I would like to thank R. H. Berk for his contribution and for valuable discussion.

REFERENCES

[1] Blackwell, David and Freedman, David (1964). A remark on the coin tossing game. *Ann. Math. Statist.* **35** 1345–1347.

[2] Chow, Y. S., Robbins, Herbert and Teicher, Henry (1965). Moments of randomly stopped sums. *Ann. Math. Statist.* **36** 789–799.

[3] Wald, Abraham (1946). Differentiation under the expectation sign in the fundamental identity of sequential analysis. *Ann. Math. Statist.* **17** 493–497.

[4] Wijsman, R. A. (1972). Examples of exponentially bounded stopping time of invariant sequential probability ratio tests when the model may be false. *Proc. Sixth Berkeley Symp. Math. Statist. Prob.* **1** 109–128.

[5] Wijsman, R. A. (1971). Exponentially bounded stopping time of sequential probability ratio tests for composite hypotheses. *Ann. Math. Statist.* **42** 1859–1869.

Department of Mathematics
University of Illinois
Urbana, Illinois 61801