NOTES

ON A SEQUENTIAL TEST FOR THE GENERAL LINEAR HYPOTHESIS[1]

By Paul G. Hoel

University of California, Los Angeles

1. Introduction. A few years ago I reported [1] on a sequential method for testing the general linear hypothesis, but held up publication until some of the properties of the method had been investigated further. Johnson [2] recently published a paper in which he obtained the same sequential test, but from an entirely different point of view. His method of derivation is based on showing, by means of a theorem of Cox [3], that the likelihood ratio approach to the problem can be used successfully. My method is a direct generalization of Wald's sequential t-test. This note outlines the nature of this generalization, and also points out some additional properties of the test.

2. Method. The general linear hypothesis assumes that the variables x_1, $x_2, \cdots, x_l$ with means $\mu_1, \mu_2, \cdots, \mu_l$ and common variance σ^2 possess the frequency function

$$f(x_1, \cdots, x_l) = (2\pi\sigma)^{-l} \exp\left\{\frac{-1}{2\sigma^2}\left[\sum_{i=1}^{k}(x_i - \mu_i)^2 + \sum_{i=k+1}^{l} x_i^2\right]\right\}. \tag{1}$$

It tests the hypothesis H_0: $\mu_1 = \cdots = \mu_p = 0$, for $p \leqq k$. The means $\mu_{k+1}, \cdots, \mu_l$ have the value zero. The parameters $\mu_{p+1}, \cdots, \mu_k$, and σ are nuisance parameters, and therefore make H_0 a composite hypothesis.

Following Wald's [4] procedure and notation for the sequential t-test, let the parameter space Ω be divided into the three regions ω_a, ω_r, and $\Omega - \omega_a - \omega_r$. The region ω_a will be chosen as that part of Ω where H_0 holds. The region ω_r will be chosen as that part of Ω where

$$\sum_{i=1}^{p} \frac{\mu_i^2}{\sigma^2} \geqq \lambda_0$$

where λ_0 is a selected constant. The boundary of ω_r will be denoted by S_r. As normalized weight functions choose

$$v_a(\theta) = \begin{cases} a_1\, \sigma^{b_1}\, (2c)^{p-k}, & 0 \leqq \sigma \leqq c,\ |\mu_i| \leqq c,\ i = p+1, \cdots, k, \\ 0, & \text{elsewhere}; \end{cases}$$

$$v_r(\theta) = \begin{cases} a_2\, \sigma^{b_2}\, (2c)^{p-k}, & 0 \leqq \sigma \leqq c,\ |\mu_i| \leqq c,\ i = p+1, \cdots, k, \\ 0, & \text{elsewhere}. \end{cases}$$

Received February 19, 1954.

[1] This research was sponsored by the O.N.R.

Here a_1, a_2, b_1, and b_2 are certain related constants to be specified later. With these weight functions and with $\prod_{\alpha=1}^{n} f(x_{1\alpha}, \cdots, x_{l\alpha})$ denoting the likelihood function for a sample of size n from (1), form the quantities

$$p_{1nc} = \int_{S_{rc}} v_r(\theta) \prod_{\alpha=1}^{n} f(x_{1\alpha}, \cdots, x_{l\alpha})\, dS_{rc},$$

$$p_{0nc} = \int_{\omega_{ac}} v_a(\theta) \prod_{\alpha=1}^{n} f(x_{1\alpha}, \cdots, x_{l\alpha})\, d\theta.$$

Here S_{rc} is that part of the surface S_r and ω_{ac} is that part of the region ω_a where $0 \leqq \sigma \leqq c$ and $|\mu_i| \leqq c$ for $i = p + 1, \cdots, k$. These are truncations of S_r and ω_a to permit the existence of the necessary integrals.

Now choose $b_2 = b_1 + 1 - p$, form the ratio p_{1nc} / p_{0nc}, and allow c to become infinite. The resulting limit, which will be denoted by p_{1n} / p_{0n}, is then treated as the likehood ratio in the standard sequential probability ratio test. By choosing $b_1 = -1$ and $l = k$ it can be shown that this test reduces to the test proposed by Johnson.

The procedure, in outline, for obtaining the reduced form of this test is as follows. The numerator integral in p_{1n} / p_{0n}, which is a surface integral over S_r, is expressed as an integral over the $\mu_1, \cdots, \mu_p$ plane, and the denominator integral is expressed as an integral with respect to σ. Then, letting

$$z_{i\alpha} = \begin{cases} x_{i\alpha}, & i = 1, \cdots, p, k + 1, \cdots, l, \\ x_{i\alpha} - \bar{x}_i, & i = p + 1, \cdots, k, \end{cases}$$

it is shown that p_{1n} / p_{0n} is a homogeneous function of degree zero in the $z_{i\alpha}$. This permits the replacing of $z_{i\alpha}$ by $z_{i\alpha} / \sqrt{\sum\sum z_{i\alpha}^2}$, which in turn enables one to evaluate the denominator integral. The numerator integral is evaluated by introducing spherical coordinates, expanding one of the exponential functions in the integral, and integrating term by term. The resulting series will be found to reduce to a confluent hypergeometric function. The result of these manipulations is the expression

$$\text{(2)} \qquad \frac{p_{1n}}{p_{0n}} = e^{-n\lambda_0/2} F(\tfrac{1}{2}[nl + p - k - b_1 - 1], \tfrac{1}{2}p, \delta^2/2\lambda_0)$$

where $\delta^2 = n\lambda_0^2\, n \sum_{i=1}^{p} \bar{x}_i^2 / [\sum_{\alpha=1}^{n} \sum_{i=1}^{l} x_{i\alpha}^2 - n \sum_{i=p+1}^{k} \bar{x}_i^2]$.

This derivation requires that $b_1 \leqq 0$. If b_1 is chosen equal to -1, the preceding test reduces to the Johnson test except for one minor feature. Johnson's formulation of the general linear hypothesis does not include the variables $x_{k+1}, \cdots, x_l$, and as a result our first parameter values in the confluent hypergeometric function are not in agreement unless $l = k$. His derivation applies equally well to the formulation given in (1), and then his test will be identical with (2) for $b_1 = -1$. The choice of $b_1 = -1$ in (2) is probably as good a choice as any other.

3. Properties. If the principle of invariance is applied, the sequential test (2) will possess an interesting property.

From (1) the frequency function for a sample of size n is

$$(3)\qquad L = (2\pi\sigma)^{-nl} \exp\left\{\frac{-1}{2\sigma^2}\sum_{\alpha=1}^{n}\left[\sum_{i=1}^{k}(x_{i\alpha}-\mu_i)^2 + \sum_{i=k+1}^{l} x_{i\alpha}^2\right]\right\}.$$

By an orthogonal transformation for which $x'_{i1} = (x_{i1} + \cdots + x_{in}) / \sqrt{n}$, for $i = 1, \cdots, k$, (3) can be reduced to the canonical form

$$(4)\quad L' = (2\pi\sigma)^{-nl} \exp\left\{\frac{-1}{2\sigma^2}\left[\sum_{\alpha=2}^{n}\sum_{i=1}^{k} x'^2_{i\alpha} + \sum_{\alpha=1}^{n}\sum_{i=k+1}^{l} x'^2_{i\alpha} + \sum_{i=1}^{k}(x'_{i1} - \sqrt{n}\,\mu_i)^2\right]\right\}.$$

From this it follows that the problem specification will remain invariant under the following groups of transformations:

(i) $\tilde{x}_{i1} = x'_{i1} + c, \quad \text{for } i = p + 1, \cdots, k;$

(ii) any orthogonal transformation on the variables $x'_{i\alpha}$ when the variables $x'_{11}, \cdots, x'_{k1}$ are deleted, and any orthogonal transformation on the variables $x'_{11}, \cdots, x'_{p1}$;

(iii) $\tilde{x}_{i\alpha} = cx_{i\alpha}, \quad \text{for } i = 1, \cdots, l, \text{ and } \alpha = 1, \cdots, n.$

Standard methods will show that the maximal invariant under these groups of transformations is

$$\phi_n = \sum_{i=1}^{p} x'^2_{i1} \Big/ \left[\sum_{\alpha=2}^{n}\sum_{i=1}^{k} x'^2_{i\alpha} + \sum_{\alpha=1}^{n}\sum_{i=k+1}^{l} x'^2_{i\alpha}\right].$$

In terms of the original variables $x_{i\alpha}$, the expression for ϕ_n reduces to

$$\phi_n = n\sum_{i=1}^{p} \bar{x}_i^2 \Big/ \left[\sum_{\alpha=1}^{n}\sum_{i=1}^{l} x_{i\alpha}^2 - n\sum_{i=p+1}^{k} \bar{x}_i^2\right].$$

Thus, if the test of H_0 is to be an invariant sequential test, it must be based on the sequence of variables $\phi_1, \phi_2, \phi_3, \cdots$. Since (2) is based on this sequence, it is an invariant test. The distribution of ϕ_n depends on the parameters through the single parameter

$$\lambda = \sum_{i=1}^{p} \frac{\mu_i^2}{\sigma^2}.$$

As a consequence, for any invariant test, testing the hypotheses

$$H_0 : \mu_i = 0 \qquad \text{vs.} \qquad H_1 : \mu_i = \mu'_i, \qquad i = 1, \cdots, p,$$

is equivalent to testing the hypothesis $\lambda = 0$ against the alternative $\lambda = \lambda_0$ where $\lambda_0 = \sum_1^p \mu'^2_i / \sigma^2$.

Johnson has shown that test (2) is the probability ratio test for testing $\lambda = 0$ against $\lambda = \lambda_0$ based on the variables $\phi_1, \phi_2, \phi_3, \cdots$. If the latter variables were independent, it would follow from the optimum property of the probability ratio test that test (2) with $b_1 = -1$ is the optimum invariant test of H_0. However, since these variables are not independent, this conclusion does not follow. It may be that the use of decision limits which depend upon the value of n will produce a better test than (2), which uses the customary constant limits for the sequential probability ratio test.

REFERENCES

[1] P. G. Hoel, "A sequential test for linear hypotheses," *Ann. Math. Stat.*, Vol. 22 (1951), p. 483, (Abstract).

[2] N. L. Johnson, "Some notes on the application of sequential methods in the analysis of variance," *Ann. Math. Stat.*, Vol. 24 (1953), pp. 614–623.

[3] D. R. Cox, "Sequential tests of composite hypotheses," *Proc. Camb. Philos. Soc.*, Vol. 48 (1952), pp. 290–299.

[4] A. Wald, *Sequential Analysis*, John Wiley and Sons, 1947, pp. 81–84, 203–207.

TWO COMMENTS ON "SUFFICIENCY AND STATISTICAL DECISION FUNCTIONS"

By R. R. Bahadur and E. L. Lehmann

Columbia University and *University of California, Berkeley*

In the following comments we employ the notation and definitions of [1]. The first comment answers a question raised in [1] by giving an example of a necessary and sufficient subfield which cannot be induced by a statistic. The second remark clarifies this example somewhat by discussing the connection between statistics and subfields in general. It was hoped that this connection would be so close as to provide the answer to another question raised in [1]: whether the existence of a necessary and sufficient subfield implies that of a necessary and sufficient statistic. However, an example given at the end of the second comment shows that such a result cannot be proved without making deeper use of sufficiency.

1. A counter example. The following result was communicated to us by David Blackwell.

Lemma 1. (*Blackwell*). *Let* $\mathbf{S}_0$ *be a proper subfield of* $\mathbf{S}$ *and suppose that for each* x *the set* $\{x\}$ *consisting of the single point* x *is in* $\mathbf{S}_0$. *Then* $\mathbf{S}_0$ *cannot be induced by a statistic.*

Proof. Suppose there exists such a statistic, say T, and let $\mathbf{T}$ be the field of sets B in the range of T such that $T^{-1}(B) \varepsilon \mathbf{S}$. Since $\{x\} \varepsilon \mathbf{S}_0$, there exists $B \varepsilon \mathbf{T}$ such that $T^{-1}(B) = \{x\}$, and, by definition of $\mathbf{T}$, a set $A \varepsilon \mathbf{S}$ such that $T(A) = B$. We therefore have $T^{-1}[T(A)] = \{x\}$, and since always $T^{-1}[T(A)] \supseteq A$, we have that $T^{-1}[T(x)] = x$ for all x. Therefore, if A is any set in $\mathbf{S}$, we see that $T^{-1}[T(A)] = A$ so that $A \varepsilon \mathbf{S}_0$ and hence our assumption that $\mathbf{S}_0$ is induced by T implies that $\mathbf{S}_0 = \mathbf{S}$.

We now give an example of a necessary and sufficient subfield that cannot be

Received July 24, 1953, revised May 24, 1954.