SEQUENTIAL TESTS FOR VARIANCE RATIOS AND COMPONENTS OF VARIANCE[1]

By Allan Birnbaum

Columbia University, New York, and Imperial College, London

1. Introduction and summary. A general sequential sampling method is given for problems of comparing the variances of two or more normal populations in terms of ratios of variances. Sequential tests are given for a hypothesis specifying the ratio of two variances, including tests for variance components in the analysis of variance (Model II). Such tests provide savings in average required numbers of observations, relative to standard F tests, comparable to those typical of sequential probability ratio tests.

2. Basic sequential sampling rule. Let $x' = (x_1', x_2', \cdots)$, $y' = (y_1', y_2', \cdots)$ be sequences of independent observations from two normal populations with unknown means and unknown respective variances σ_x^2, σ_y^2. Let

$$x_1 = \frac{1}{\sqrt{2}} x_1' - \frac{1}{\sqrt{2}} x_2',$$

$$x_2 = \frac{1}{\sqrt{2 \cdot 3}} x_1' + \frac{1}{\sqrt{2 \cdot 3}} x_2' - \frac{2}{\sqrt{2 \cdot 3}} x_3',$$

$$x_n = \frac{1}{\sqrt{n(n+1)}} x_1' + \cdots + \frac{1}{\sqrt{n(n+1)}} x_n' - \frac{n}{\sqrt{n(n+1)}} x_{n+1}',$$

for $n = 3, 4, \cdots$. Let $y_1, y_2, \cdots$ be similarly defined as functions of $y_1', y_2', \cdots$. Then $x = (x_1, x_2, \cdots)$, $y = (y_1, y_2, \cdots)$ are sequences of independent observations, normally distributed with zero means and respective variances σ_x^2, σ_y^2. (In case the original observations are from populations with known means equal, say, to zero, x and y will denote the sequences of original observations.)

Let g be a given positive number, and let

$$u = (u_1, u_2, \cdots) = (gx_1^2 + gx_2^2, gx_3^2 + gx_4^2, \cdots) \quad \text{and}$$

$$v = (v_1, v_2, \cdots) = (y_1^2 + y_2^2, y_3^2 + y_4^2, \cdots).$$

Let

$$R_i = \sum_{j=1}^{i} u_j \quad \text{and} \quad S_i = \sum_{j=1}^{i} v_j, \qquad \text{for } i = 1, 2, \cdots.$$

Let $T = (T_1, T_2, \cdots)$ be a nondecreasing sequence whose elements are those of $\{R_i\} \cup \{S_i\}$. Since events of the form $R_i = R_j$, $S_i = S_j$, $i \neq j$, and $R_i = S_j$

Received December 14, 1954; revised September, 1956.

[1] Work sponsored under the Office of Naval Research, Contract N6onr-271, T. O. XI, Project 042–034. Reproduction in whole or in part is permitted for any purpose of the United States Government.

have total probability zero, we have with probability one that T is uniquely determined,

$$T_1 = \min(R_1, S_1),$$

$$T_2 = \begin{cases} \min(R_2, S_1) & \text{if } T_1 = R_1, \\ \min(R_1, S_2) & \text{if } T_1 = S_1, \text{ etc.} \end{cases}$$

Given T, let $B = (b_1, b_2, \cdots)$ be defined by

$$b_i = \begin{cases} 1 & \text{if } T_i = \text{some } R_j, \\ 0 & \text{if } T_i = \text{some } S_j, \end{cases} \qquad \text{for } i = 1, 2, \cdots.$$

The statistical decision procedures to be described below are based on observed values b_i only.

A simple rule for sequential sampling of x_i's and y_i's so as to obtain the values b_i is the following

Sampling rule 1:

1. Observe u_1 and v_1 (that is, observe x_1, x_2, y_1 and y_2, and compute $u_1 = g(x_1^2 + x_2^2)$ and $v_1 = (y_1^2 + y_2^2)$.
2. If an additional observation b_2 is required, then if $u_1 < v_1$, observe u_2; if $u_1 \geqq v_1$, observe v_2.
3. Similarly at every stage, if an additional observation b_i is required, then if $\sum u_i < \sum v_i$, observe an additional u_i, while if $\sum u_i \geqq \sum v_i$, observe an additional v_i.
4. Discontinue sampling when the observations $b_1, \cdots, b_m$ thus far obtained suffice to determine a decision according to the particular decision procedure being used.

It is clear that to obtain m observations $b_1, \cdots, b_m$, a total of $m + 1$ observations u_i and v_i are required; thus the above rule is most efficient for sampling u_i's and v_i's to obtain b_i's.

A minor gain in efficiency here becomes possible if the sampling rule is described in terms of x_i's and y_i's. This follows from the following observation: If only x_1, x_2, and y_1 have been observed, and are such that

$$u_1 \equiv g(x_1^2 + x_2^2) < y_1^2 \leqq v_1,$$

then $u_1 < v_1$ and $b_1 = 1$ are known without need to observe y_2. Similarly if $gx_1^2 > y_1^2 + y_2^2$ is observed, $b_1 = 0$ is known without need to observe x_2. It is clear that the determination of b_1 in this way (that is, by observing whether $u_1 \leqq v_1$, without observing the exact numerical values of both u_1 and v_1) does not alter any mathematical or statistical properties of b_1, since b_1 is defined in terms of an inequality in u_1 and v_1. Analogous observations hold at each stage of sampling, and are the basis for the following

Basic sequential sampling rule:

1. Observe x_1 and y_1.
2. If $gx_1^2 < y_1^2$, observe x_2; if $gx_1^2 \geqq y_1^2$, observe y_2.

TABLE 1

Example of sampling rule (with $g = 1$) and computation of b_i's.

Observed values, in order of sampling		$\Sigma gx_i^2 - \Sigma y_i^2$	b_i values
x_i^2	y_i^2		
1. $x_1^2 = 4$	$y_1^2 = 2$	2	
2.	$y_2^2 = 1$	1	$b_1 = 0$
3.	$y_3^2 = 3$	-2	
4. $x_2^2 = 1$		-1	$b_2 = 1$
5. $x_3^2 = 3$		2	
6.	$y_4^2 = 5$	-3	
7. $x_4^2 = 1$		-2	$b_3 = 1$
8. $x_5^2 = 3$		1	$b_4 = 0$

3. Similarly at every stage, letting $\sum gx_i^2$ and $\sum y_i^2$ denote summations over all observations thus far obtained, if $\sum gx_i^2 < \sum y_i^2$, observe an additional x_i; if $\sum gx_i^2 \geqq \sum y_i^2$, observe an additional y_i.
4. Discontinue sampling when the observations $b_1, \cdots b_m$, thus far obtained suffice to determine a decision according to the particular decision procedure being used.

A convenient tabular method for carrying out the sampling and computing the required b_i's is illustrated in Table 1, for the case $g = 1$ of the sampling rule.

The computation of b_i's may be described thus: As soon as $2r$ or more x_i's have been observed *and* $\sum gx_i^2 - \sum y_i^2 < 0$ is observed, the r^{th} unity value of b_i is observed. As soon as $2r$ or more y_i's have been observed *and* $\sum gx_i^2 - \sum y_i^2 \geqq 0$ is observed, the r^{th} zero value of b_i is observed. In applications where population means are unknown, the relation

$$\sum_1^n x_i^2 = \sum_1^{n+1} \left(x_i' - \frac{1}{n+1} \sum_1^{n+1} x_i'\right)^2 = \sum_1^{n+1} (x_i')^2 - \left(\sum_1^{n+1} x_i'\right)^2 \Big/ (n+1)$$

allows simple application of the method directly to the original observations. It is readily verified that this rule minimizes the number of x_i's and y_i's which must be observed to determine the required b_i's. Since this rule requires at least $(2m + 1)$ and at most $(2m + 2)$ observations x_i and y_i, it affords a saving of at most one such observation (and in terms of expected number of observations, a saving of a fraction of one observation) as compared with the preceding rule.

Hence rule 1 may often be preferred because of its greater simplicity. However only the basic sampling rule will be considered in the following sections.

Clearly the sampling rule depends on the given value of g. Criteria for the choice of g are discussed below.

3. Distribution theory. $\frac{1}{2} u_1, \frac{1}{2} u_2, \cdots$ are independently distributed with common density function

$$f(w) = \frac{1}{g\sigma_x^2} e^{-w/g\sigma_x^2}, \qquad w \geqq 0.$$

Similarly $\frac{1}{2}v_1$, $\frac{1}{2}v_2$, $\cdots$ are independently distributed with common density function

$$h(w) = \frac{1}{\sigma_y^2} e^{-w/\sigma_y^2}, \qquad w \geqq 0.$$

Hence (cf. [1] or [2])

(a) The sequences u and v are distributed as the "waiting times" between successive events in two independent Poisson processes with respective parameters

$$\lambda_1 = \frac{1}{2g\sigma_x^2} \quad \text{and} \quad \lambda_2 = \frac{1}{2\sigma_y^2}.$$

(b) If two such processes are observed simultaneously as one process, the new process is Poisson with mean $\lambda = \lambda_1 + \lambda_2$ and waiting times T_1, $T_2 - T_1$, $T_3 - T_2$, $\cdots$.

(c) $b_i = 1$ (i.e. $T_i =$ some R_j) denotes that the i^{th} event observed in the new process occurred in the first of the two original processes.

(d) The b_i's are independent, with

$$p = \text{Prob}\ \{b_1 = 1\} = \left(1 + \frac{\lambda_2}{\lambda_1}\right)^{-1} = \left(1 + g\frac{\sigma_x^2}{\sigma_y^2}\right)^{-1}$$

for all i.

Thus we may apply to B any sequential or nonsequential methods for statistical inferences concerning a binomial parameter p. By use of the relation $\sigma_x^2/\sigma_y^2 = (1/g)\ (1/p - 1)$, any interval estimate of p provides an interval estimate of σ_x^2/σ_y^2, and any procedure for testing a hypothesis on p provides a procedure for testing a hypothesis on σ_x^2/σ_y^2. An unbiased estimate of σ_x^2/σ_y^2 is given by $(1/g)\ (1/\tilde{p} - 1)$, where $1/\tilde{p}$ is an unbiased estimate of $1/p$ based on inverse binomial sampling of b_i's.

The generalization of the basic sampling rule and distribution theory to the case of comparisons of three or more variances is immediate. In this generalization, the distribution of b_i's would be multinomial instead of binomial.

4. Efficiency. A number of questions concerning the efficiencies of tests based on the above sampling rule are discussed in the following sections.

4a. Comparisons with standard F-tests. The standard method of testing a hypothesis $H_0 : \sigma_x^2/\sigma_y^2 = \rho_0$, for any given $\rho_0 > 0$, is to take fixed numbers n_x, n_y of observations x_i, y_i respectively, and to use the statistic

$$F = \frac{n_y \sum_{i=1}^{n_x} x_i^2}{n_x \rho_0 \sum_{i=1}^{n_y} y_i^2}$$

which under H_0 has the F-distribution with n_x, n_y degrees of freedom. Tables 8.3 and 8.4 of [3] give the operating characteristics of such tests. For example, to test $H_0 : \sigma_x^2/\sigma_y^2 = .3404$ against $H_1 : \sigma_x^2/\sigma_y^2 = (.3404)^{-1} = 2.938$, with $\alpha =$

Type I error $= .01$ and $\beta =$ Type II error $= .01$, a total of $n_x + n_y = 40$ observations suffices provided $n_x = n_y = 20$. If the above sampling rule is used, taking $g = 1$, then H_0 is equivalent to

$$H_0': p = \text{Prob } \{b_i = 1\} = p_0 = (1 + .3404)^{-1} = .746$$

and H_1 is equivalent to $H_1': p = p_1 = (1 + 2.938)^{-1} = .254$. By use of binomial probability tables [4] we find Prob $\{\sum_{i=1}^{21} b_i \geqq 11 \mid p = .26\} =$ Prob $\{\sum_{i=1}^{21} b_i \leqq 10 \mid p = .74\} = .0088$; thus a nonsequential test of H_0' against H_1' based on 21 observations b_i has

$$\alpha = \beta < .0088 < .01$$

The sampling rule requires either 43 or 44 observations x_i, y_i to generate 21 observations b_i. Thus the efficiency of the standard F-test is approximately matched by the nonsequential test based on b_i's in this case. (More exact comparisons of efficiency can be made, for example by computing the required exact untabled binomial probabilities and using randomized binomial sample sizes, but this does not seem necessary for present purposes.) Comparisons of this kind are given for a number of other cases in Table 2 below. The properties of the four F-tests are taken from [3], with $n_x = n_y$ in each case. The approximately matching non-sequential binomial tests are each based on the case $g = 1$ of the sampling rule; m is the required binomial sample size in each case. Since for m

TABLE 2

	Standard F-tests					Approximately equivalent non-sequential binomial tests						Value of σ_x^2/σ_y^2 corresponding exactly to H_1'
	Value of σ_x^2/σ_y^2		Strength		$n_x + n_y$	Value of p		Strength		m	$n_x + n_y \doteq 2m + \frac{3}{2}$	
	H_0	H_1	α	β		H_0'	H_1'	α	β			
Test 1	.3404	2.938	.01	.01	40	.73	.27	.0119	.0119	21	43.5	2.704
	.4707	2.124	.05	.05	40	.67	.33	.0520	.0520	21	43.5	2.030
Test 2	1	4.512	.05	.05	40	.5	.15	.0318	.0537	19	39.5	5.667
						.5	.19	.0835	.0532	19	39.5	4.263
						.5	.18	.0577	.0537	20	41.5	4.555
Test 3	1	2.866	.05	.05	80	.5	.26	.0541	.0597	39	79.5	2.846
	1	2.866	.05	.05	80	.5	.25	.0403	.0544	40	81.5	3.000
	1	1.693	.05	.50	80	.5	.37	.0541	.4850	39	79.5	1.703
	1	1.693	.05	.50	80	.5	.36	.0403	.4807	40	81.5	1.778
Test 4	1	4.470	.01	.01	80	.5	.17	.0119	.0099	39	79.5	4.882
		4.470	.01	.01	80	.5	.17	.0083	.0124	40	81.5	4.882
		3.579	.01	.05	80	.5	.21	.0119	.0505	39	79.5	3.762
		3.579	.01	.05	80	.5	.20	.0083	.0432	40	81.5	4.000
		2.114	.01	.50	80	.5	.32	.0019	.4889	39	79.5	2.125
		2.114	.01	.50	80	.5	.31	.0083	.4777	40	81.5	2.226

observations b_i the sampling rule requires $2m + 1$ or $2m + 2$ observations x_i, y_i, $2m + 3/2$ ($\doteq n_x + n_y$) is given for each binomial test for comparison with the $n_x + n_y$ required by the corresponding F-test. In each case investigated, the F-test is approximately matched in efficiency by a nonsequential binomial test based on b_i's.

The F-tests are no doubt preferable to the nonsequential binomial tests, but evidently simplicity of application is the only important basis for this preference. Sequential probability ratio tests [5] are directly applicable to the b_i's. Such tests of $H_0': p = p_0$ against $H_1': p = p_1 > p_0$ require average sample sizes of approximately $m/2$ or less when $p \leqq p_0$ and when $p \geqq p_1$. Thus by use of the above sampling rule, with application of a sequential test to the b_i's, gains in efficiency over the standard F-test are obtained. These gains can be calculated, to close approximation, by use of the average-sample-size function for a sequential binomial test on b_i's corresponding to any given F-test.

Two-sided sequential tests on σ_x^2/σ_y^2 based on the b_i's would require two-sided sequential probability ratio tests on a binomial parameter. Such tests are not yet available, but can be constructed by application of the method of sections 4.1.2 and 4.1.3 of [5].

4b. Choice of the scale-factor g. Each of the binomial tests considered in the preceding sections was based on the particular case $g = 1$ of the basic sequential sampling rule. The efficiencies of such binomial tests will in general be still further increased by suitable choices of values of g. When a problem of testing a variance ratio $\rho = \sigma_x^2/\sigma_y^2$ is specified by given values of ρ_0, ρ_1, α, and β, it is natural to define a best value of g as follows: Consider the problem of testing $H_0': p = (1 + g\rho_0)^{-1}$ against $H_1': p = (1 + g\,\rho_1)^{-1}$, at strength α, β. Let $n(g)$ be the binomial sample size required for a nonsequential test of strength (at least) α, β. Then a best value of g may be defined as one which minimizes $n(g)$. The calculation of an optimal value of g, by use of binomial probability tables, is elementary.

In the case $\alpha = \beta$, symmetry considerations suggest that $g = (\rho_0\rho_1)^{-1/2}$ is a best value; the same conclusion can be reached more formally by use of the normal approximation to the binomial probability $\alpha = \beta$. This case occurs in some of the examples above: For example, to test $H_0: \rho = .4707$ against $H_1: \rho = (.4707)^{-1} = 2.124$ at strength $\alpha = \beta = .05$, the binomial test based on $g = 1$ requires the (non-sequential) sample size $m = 21$. Taking the non-optimal value $g = 21.24$, the corresponding binomial problem is one of testing $H_0': p = .09$ against $H_1': p = .02$ at strength $\alpha = \beta = .05$, for which a (non-sequential) binomial sample size $m = 50$ is required.

If only ρ_0 and ρ_1 are given, it is interesting to consider whether a value of g exists which is best in the above sense simultaneously for all possible values of (α, β). This is a problem in the "comparison of experiments" [6, pp. 334–6]: For any ρ_0, ρ_1 ($\rho_1 \neq \rho_0$) *and* any two values g_1, g_2 of g ($g_1 \neq g_2$), the "binomial dichotomy" experiment E_1, testing $H_0^{(1)}: p = (1 + g_1\rho_0)^{-1}$ against $H_1^{(1)}: p =$

$(1 + g_1\rho_1)^{-1}$, can be shown to be "not comparable with" the experiment E_2, testing $H_0^{(2)}: p = (1 + g_2\rho_0)^{-1}$ against $H_1^{(2)}: p = (1 + g_2\rho_1)^{-1}$. It follows that a best value of g depends on the desired strength (α, β) as well as on ρ_0, ρ_1 in any particular testing problem.

While the value $g = 1$ may not be optimal in those testing problems of Table 1 in which $\rho_1 \neq \rho_0^{-1}$ or $\alpha \neq \beta$, it is evident that this value suffices to provide the savings in expected numbers of observations pointed out in Section 4a above.

4c. Comparison with other sampling rules. It is of interest to compare the above sampling rule with the one considered by Girshick in [7, pp. 134–6]. Girshick's rule is: observe x_1 and y_1; if more observations are needed, observe x_2 and y_2; continue taking such pairs of observations (x_i, y_i) as long as required to terminate a particular inference procedure. (For non-sequential F tests of the hypothesis $H_0: \sigma_x^2/\sigma_y^2 = \rho_0 < 1$ against $H_1: \sigma_x^2/\sigma_y^2 = \rho_0^{-1}$, with a total of $2n$ observations and equal Type I and Type II error rates, $n_x = n_y = n$ are optimal sample sizes. This case is formally the instance of Girshick's sampling rule in which the observation of exactly n pairs (x_i, y_i) is prescribed.) Girshick gave an optimal sequential probability ratio test based on this sampling rule for deciding which of the two variances σ_x^2, σ_y^2 is the larger, and showed that the power functions of such tests depend just on $(1/\sigma_x^2 - 1/\sigma_y^2)$, a parameter which is not of interest in most applications. This suggests that in order to obtain a test whose power function depends just on the variance ratio (which is generally the parameter of interest), we must

(a) use Girshick's sampling rule and apply a test other than Girshick's, which must then lack certain efficiency properties of the sequential probability ratio test, or

(b) use some other sampling rule as a basis for a test.

Rushton [8] and Johnson [9] have given procedures which represent alternative (a). Johnson's Procedure I is based on Girshick's sampling rule and the sequence of statistics

$$\frac{x_1^2}{y_1^2}, \quad \frac{x_1^2 + x_2^2}{y_1^2 + y_2^2}, \cdots \frac{\sum_1^r x_i^2}{\sum_1^r y_i^2}, \cdots,$$

and consists of a sequential probability ratio test applied to this sequence of non-independent statistics. The average sample sizes required by this procedure are not known. Johnson also gives some alternative procedures based on Girshick's sampling rule which are evidently less efficient but have approximately-known average sample sizes.

Alternative (b) is represented by the sampling rule and tests of the preceding sections.

This comparison of sampling rules illustrates a general problem of designing sequential sampling rules which are appropriate and optimal for various problems of testing composite hypotheses. Other illustrations are provided by various procedures for comparing Poisson processes [2]. Some general methods for the design of appropriate sampling rules will be given in another paper.

The above sampling rule and tests illustrate a remark in [2, p. 257]: "$\cdots$ problems dealing with variances of normal populations have direct analogues in problems dealing with parameters of Poisson processes. $\cdots$ " The above methods are analogues of Methods 1–3, pp. 261–2, in [2].

4d. Near-optimal properties. Consider the problem of testing $H_0: \sigma_x^2/\sigma_y^2 = \rho_0$ against $H_1: \sigma_x^2/\sigma_y^2 = \rho_1$, $\rho_1 > \rho_0 > 0$, at a given significance level α. Suppose that sequential sampling is conducted according to the above basic sampling rule, with any given value of $g > 0$, and any given termination rule which with probability one leads to termination. Then, relative to the *given* sampling and termination procedure (i.e. relative to the corresponding given sample space), the present problem may be viewed as one of nonsequential testing between two composite hypotheses, on the basis of a single (vector) observation. The parameter space consists of points (ρ, τ), where $\rho = \sigma_x^2/\sigma_y^2$, and $\tau = \sigma_y^2$.

Consider the simple hypothesis $H_0^*: (\rho, \tau) = (\rho_0, \tau_0)$, and the simple alternative $H_1^*: (\rho, \tau) = (\rho_1, \tau_1)$, where $\tau_1 = \tau_0 (\rho_0(\rho_1 + 1))/(\rho_1(\rho_0 + 1))$, and τ_0 has any fixed positive value. By the Neyman-Pearson lemma, every best test of H_0^* against H_1^* has a critical region of the form

$$W_k = \{(x, y) \mid \lambda(x, y) \geqq k\},$$

for some given $k \geqq 0$, where

$$\lambda(x, y) = \frac{f(x_1, \cdots x_{n_x}, y_1, \cdots y_{n_y} \mid H_1^*)}{f(x_1, \cdots x_{n_x}, y_1, \cdots y_{n_y} \mid H_0^*)}$$

is the ratio of likelihood functions. Such a test has significance level α^* which, when k is suitably chosen, equals the prescribed α. This test then has power $1 - \beta^*$ under H_1^* which is the maximum attainable under the stated restrictions.

On the other hand, a best test of H_0^* against H_1^* based only on the observed values of b_i's provided by the given sampling procedure has a rejection region of the form

$$W_k' = \{(x, y) \mid \lambda'(x, y) \geqq k\},$$

where

$$\lambda'(x, y) = \left(\frac{p_1}{p_0}\right)^{\Sigma_1^m b_i} \left(\frac{1 - p_1}{1 - p_0}\right)^{m - \Sigma_1^m b_i}$$

is the likelihood ratio of the observed b_i's, with

$$p_j = \text{Prob } \{b_i = 1 \mid H_j^*\} = (1 + g\rho_j)^{-1}, \qquad j = 0, 1.$$

The purpose of the present section is to show, in a qualitative and heuristic manner, that under appropriate restrictions

$$\text{Prob } \{W_k \mid H_j^*\} \doteq \text{Prob } \{W_k' \mid H_j^*\} \qquad \text{for } j = 0, 1$$

and any τ_0, which implies the approximate optimality of a test based on b_i's only, for the given sequential sampling and termination rule. The following discussion could be formulated more quantitatively, but this seems unnecessary since

(a) the tests based on b_i's have simple and known properties,
(b) the operating characteristics of tests based on $\lambda(x, y)$ would apparently be difficult to determine exactly and more difficult to apply than the test on b_i's,
(c) a qualitative indication of the approximate equivalence of the tests serves practically as an additional recommendation for use of the tests based on b_i's.

Now

$$\lambda(x, y) = \frac{(2\pi\rho_1\tau_1)^{-n_x/2}(2\pi\tau_1)^{-n_y/2}\exp\left[-\frac{1}{2\rho_1\tau_1}\sum_1^{n_x} x_i^2 - \frac{1}{2\tau_1}\sum_1^{n_y} y_i^2\right]}{(2\pi\rho_0\tau_0)^{-n_x/2}(2\pi\tau_0)^{-n_y/2}\exp\left[-\frac{1}{2\rho_0\tau_0}\sum_1^{n_x} x_i^2 - \frac{1}{2\gamma_0}\sum_i^{n_y} y_i^2\right]}$$

$$= \left(\frac{\rho_0 + 1}{\rho_1 + 1}\right)^{n_x/2}\left(\frac{\rho_1(\rho_0 + 1)}{\rho_0(\rho_1 + 1)}\right)^{n_y/2}\exp\left[-\frac{1}{2}\left(\sum_1^{n_x} x_i^2 - \sum_1^{n_y} y_i^2\right)\left(\frac{\rho_0 - \rho_1}{\tau_0\rho_0(\rho_1 + 1)}\right)\right].$$

Now the case $g = 1$ of the basic sampling rule is such that at every stage of sampling the quantity $(\sum x_i^2 - \sum y_i^2)$ will be increased if it is negative and will be decreased if it is positive; thus this quantity will under all hypotheses have a distribution concentrated near the value zero, and the exponential factor of $\lambda(x, y)$ will tend to have a value near unity. For cases $g \neq 1$ of the sampling rule, a change of scale in x_i values, $x_i \rightarrow \sqrt{g}x_i$, before computation of $\lambda(x, y)$ gives the same result. We continue the discussion just for the case $g = 1$.

Thus if the rule for termination of sampling is such that $n_x + n_y$ is not small,

$$\text{Prob}\,\{\lambda(x, y) \geqq k \mid H_j^*\} \doteq \text{Prob}\left\{\left(\frac{\rho_0 + 1}{\rho_1 + 1}\right)^{n_x/2}\left(\frac{\rho_1(\rho_0 + 1)}{\rho_0(\rho_1 + 1)}\right)^{n_y/2} \geqq k \mid H_j^*\right\}$$

and the last quantity is independent of the "nuisance parameter" τ_0, for $j = 0, 1$. From the definition of the b_i's we readily obtain

$$\frac{n_y}{2} - 1 \leqq m - \sum_1^m b_i \leqq \frac{n_y}{2} \quad \text{and} \quad \frac{n_x}{2} - 1 \leqq \sum_1^m b_i \leqq \frac{n_x}{2}.$$

Since

$$p_i = \frac{1}{\rho_i + 1}, \qquad 1 - p_i = \frac{\rho_i}{\rho_i + 1},$$

for $i = 0, 1$, we have

$$\lambda'(x, y) = \left(\frac{p_1}{p_0}\right)^{\sum_1^m b_i}\left(\frac{1 - p_1}{1 - p_0}\right)^{m - \sum_1^m b_i}$$

$$\doteq \left(\frac{\rho_0 + 1}{\rho_1 + 1}\right)^{n_x/2}\left(\frac{\rho_1(\rho_0 + 1)}{\rho_0(\rho_1 + 1)}\right)^{n_y/2}.$$

Hence

$$\text{Prob } \{\lambda(x, y) > k \mid H_j^*\} \doteq \text{Prob } \{\lambda'(x, y) > k \mid H_j^*\},$$

for $j = 0, 1$, and for each value of the "nuisance parameter" τ_0.

5. Sequential Tests on Components of Variance. The testing problems arising in Model II of the analysis of variance, and their usual non-sequential solution based on F tests, are described, for example, by Mood in [10, Chapter 14). The method of the preceding sections can be adapted to provide sequential tests for such problems, as indicated below. Such sequential tests provide savings like those described above in the required numbers of observations.

Consider the "one-way layout" problem in which

$$y_{ij} = \mu + a_i + e_{ij}, \qquad \text{for } i = 1, 2, \cdots, j = 1, 2, \cdots.$$

Here μ is an unknown constant, a_i and e_{ij} are normally distributed random variables with zero means and unknown respective variances σ_a^2 and σ_e^2, and all a_i's and e_{ij}'s are mutually independent. The statistical problem is to test a hypothesis specifying the value of $\rho = \sigma_a^2 / \sigma_e^2$. For present purposes we consider that a doubly infinite array of the random variables y_{ij} is available, and that we are free to take observations y_{ij} throughout this array in any manner. Let

$$T = (t_{ij}) = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 & 0 & \cdots \\ \frac{1}{\sqrt{2\cdot 3}} & \frac{1}{\sqrt{2\cdot 3}} & -\frac{2}{\sqrt{2\cdot 3}} & 0 & \cdots \\ \vdots & \cdot & \cdot & \cdot & \\ \cdot & \cdot & \cdot & \cdot & \\ \cdot & \cdot & \cdot & \cdot & \end{pmatrix}$$

denote the unique "doubly infinite orthogonal matrix" characterized as indicated, i.e., by the requirements that $t_{ij} = 0$ if $j \geqq i + 2$, that $t_{i1} > 0$, and that $\sum_{j=1}^{\infty} t_{ij} = 0$, for all i, j. Let the random variables $(r_1, r_2, \cdots)$ be defined by

$$(r_1, r_2, \cdots)' = T(y_{11}, y_{12}, \cdots)':$$

i.e., $r_\alpha = \sum_{j=1}^{\alpha+1} t_{\alpha j} y_{1j}$, for $\alpha = 1, 2, \cdots$. Let the random variables $(s_1, s_2, \cdots)$ be defined by

$$(s_1, s_2, \cdots)' = T(y_{21}, y_{31}, \cdots)';$$

i.e., $s_\alpha = \sum_{j=1}^{\alpha+1} t_{\alpha j} y_{j+1,1}$ for $\alpha = 1, 2, \cdots$. Then all r_α's and s_α's have independent normal distributions with zero means; and $\text{Var}(r_\alpha) = \sigma_r^2 = \sigma_e^2$, $\text{Var}(s_\alpha) = \sigma_s^2 = \sigma_e^2 + \sigma_a^2$.

The sequential test procedures given above can be applied directly to the sequences of r_α's and s_α's to test any hypothesis on $\rho = \sigma_s^2 / \sigma_r^2$, say $H_0: \rho = \rho_0$

against $H_1 : \rho = \rho_1 > \rho_0$, at specified size and power. But since

$$\rho = \frac{\sigma_s^2}{\sigma_r^2} = \frac{\sigma_e^2 + \sigma_a^2}{\sigma_e^2} = 1 + \frac{\sigma_a^2}{\sigma_e^2},$$

H_0 is equivalent to $H_0^* : \sigma_a^2/\sigma_e^2 = \rho_0 - 1$, and H_1 is equivalent to $H_1^* : \sigma_a^2/\sigma_e^2 = \rho_1 - 1$. (Here $\rho_0 \geqq 1$ is required if H_0^* is to be meaningful.) Thus the usual hypotheses of interest in terms of variance components, namely those specifying a positive value for $\sigma_a^2/\sigma_e^2 = \rho_0 - 1$, and those specifying $\sigma_a^2 = 0$ (and hence $\rho_0 = 1$), can be tested sequentially with gains in efficiency as described above.

REFERENCES

[1] A. Birnbaum, "Some procedures for comparing Poisson processes or populations", *Biometrika*, Vol. 40 (1953), pp. 447–449.

[2] A. Birnbaum, "Statistical methods for Poisson processes and exponential populations", *J. Amer. Stat. Assn.*, Vol. 49 (1954), pp. 254–266.

[3] Statistical Research Group, Columbia University, *Selected techniques of statistical analysis*, McGraw-Hill Book Co., Inc., New York, 1947.

[4] National Bureau of Standards, *Tables of the Binomial Probability distribution*, U. S. Government Printing Office, Washington, 1949.

[5] A. Wald, *Sequential Analysis*. John Wiley and Sons, Inc., New York, 1947.

[6] D. Blackwell and M. A. Girshick, *Theory of games and statistical decisions*, John Wiley and Sons, Inc., New York, 1954.

[7] M. A. Girshick, "Contributions to the theory of sequential analysis. I", *Ann. Math. Stat.*, Vol. 17 (1946), pp. 123–143.

[8] S. Rushton, "On sequential tests of the equality of variances of two normal populations with known means," *Sankhyā*, Vol. 12 (1952), pp. 63–78.

[9] N. L. Johnson, "Sequential procedures in certain component of variance problems," *Ann. Math. Stat.*, Vol. 25 (1954), pp. 357–366.

[10] A. Mood, *Intro. to the Theory of Statistics*. McGraw-Hill Book Co., Inc., New York, 1950.