A SINGLE-SAMPLE MULTIPLE DECISION PROCEDURE FOR RANKING VARIANCES OF NORMAL POPULATIONS[1]

By Robert E. Bechhofer and Milton Sobel

Cornell University

Summary. A single-sample multiple decision procedure for ranking variances of normal populations is described. Exact small-sample methods and a large-sample method are given for computing the sample sizes necessary to guarantee a preassigned probability of a correct ranking under specified conditions on certain variance ratios. Some tables computed by these methods are provided.

1. Introduction. In an earlier paper [3], one of the present authors proposed a single-sample multiple decision procedure for ranking means of normal populations with known variances. Although the procedure described in that paper can be used for ranking variances if the sample sizes are sufficiently large, the question as to which type of large-sample approximation would give satisfactory results required further study. In addition, since much applied statistics involves small sample sizes, it was felt that it would be desirable to develop an exact small-sample theory for ranking variances of normal populations. The formulation of the ranking problem as given in this paper is the same as the one given in the earlier paper. However, the earlier paper treats the problem somewhat more generally, and the reader is referred to it for additional background and motivation.

Neyman and Pearson's [10] L_1 test as modified by Bartlett [2] (with the new tables of Thompson and Merrington [13]) is the best known and most widely used test[2] for the homogeneity of variances. However, even in situations where the test is appropriate,[3] it has a very important deficiency—namely, that its power against various types of alternatives has not been determined.

In many situations the test is used inappropriately, particularly when the experimenter has strong a priori reasons for believing that the population variances can not, in fact, be exactly equal. Many times in such situations the experimenter would like to know, for example, which population has the smallest variance. What he requires is a decision procedure which will tell him which population to choose, and an operating characteristic curve which will tell him

Received 6/18/53, revised 2/2/54.

[1] This research was supported by the United States Air Force under Contract No. AF 18(600)331 monitored by the Office of Scientific Research.

[2] Recently Hartley [7] introduced an easier (from the computional viewpoint) but less powerful test of the same hypothesis based on the maximum F-ratio. See also Cochran's test [5], [6] for the significance of the largest of a set of sample estimates of variance.

[3] For example, before attempting an analysis of variance test, an experimenter might want to do some preliminary sampling in order to obtain information concerning the validity of the assumption of homogeneous variances.

the probability of his making a correct choice if he follows the given decision procedure. The ranking procedure described in the next section is designed to handle this latter situation.

2. The ranking (multiple decision) approach.

2.1. *The mathematical model and related definitions.* Let X_{ij} be normally and independently distributed chance variables $N(X_{ij} \mid \mu_i, \sigma_i^2)$,

$$(i = 1, 2, \cdots, k; j = 1, 2, \cdots, N_i).$$

We assume that the μ_i are known, and that the σ_i^2 are unknown. (If the μ_i are known linear combinations of parameters which themselves are unknown and which must be estimated from the data (a typical situation in regression problems), the only effect will be to change the degrees of freedom associated with the estimate of each σ_i^2.) Let

$$\sigma_{[1]}^2 \leqq \sigma_{[2]}^2 \leqq \cdots \leqq \sigma_{[k]}^2 \tag{1}$$

be the ranked σ_i^2, and let

$$\theta_{ij} = \sigma_{[i]}^2 / \sigma_{[j]}^2 \qquad i, j = 1, 2, \cdots, k \tag{2}$$

be the variance ratios; we assume that it is not known which population is associated with $\sigma_{[i]}^2$. We further assume that for each population, the only parameter of interest is the population variance, the "best" population being the one having the smallest variance, the "second best" being the one having the second smallest variance, etc. Alternatively, we might have defined the "best" population as being the one having the largest variance, etc.; the mathematical theory is similar for both cases.

The k populations might be k different lots of ammunition, and σ_i^2 might be the (population) target dispersion of the ith lot, or the k populations might be k different measuring instruments, and σ_i^2 might be the (population) variance of measurement of the ith instrument. This variance, which characterizes the reproducibility of repeated measurements of the same quantity, can be used as an index of the precision of the measuring instrument. We would like on the basis of a sample of $\sum_{i=1}^{k} N_i$ independent observations to make some inferences about the "bestness" of the populations.

Our inferences will be based on the sample variances, by which we mean the *best unbiased estimates* of the corresponding population variances. The sample variance from the ith population, and the number of degrees of freedom (d.f.) associated with this estimate, will be denoted by s_i^2 and n_i, respectively. (For simplicity of notation no attempt will be made in this paper to distinguish between chance variables and their observed values.) The sample variance associated with the population having population variance $\sigma_{[i]}^2$ and the number of d.f. associated with this estimate will be denoted by $s_{(i)}^2$ and $n_{(i)}$, respectively.

Thus

(3) $$n_{(i)}s^2_{(i)}/\sigma^2_{[i]} = \chi^2_{n(i)} \qquad i = 1, 2, \cdots, k.$$

The ranked s_i^2 will be denoted by

(4) $$s^2_{[1]} < s^2_{[2]} < \cdots < s^2_{[k]}.$$

(If two or more s_i^2 are equal, they should be "ranked" by using a randomized procedure which assigns equal probability to each ordering.)

2.2. *The goals, requirements, and procedures.* Different goals are appropriate for different practical situations. We shall assume that in each situation it is the experimenter's responsibility to decide, *before taking any observations*, precisely what his goal is. Two representative goals will be considered here. All problems of dividing the k populations into groups will be special cases of these two, or will require a similar development.

GOAL I. To divide the k populations into two groups, the t "best" and the $k - t$ "worst," the t best being *unordered* and the $k - t$ worst being unordered $(1 \leqq t \leqq k - 1)$.

GOAL II. To divide the k populations into $t + 1$ groups, the t "best" and the $k - t$ "worst," the t best being *ordered* and the $k - t$ worst being unordered $(1 \leqq t \leqq k - 1)$.

It should be clear that, for Goal I, the problem of choosing the t "best" is logically equivalent to choosing the $k - t$ "worst." It should also be noted that, for Goal II when $t = k - 1$, the problem is that of requiring a *complete ranking*. The goals coincide for $t = 1$.

For Goal I it is assumed that the experimenter can specify a smallest value of $\theta_{t+1,t}$, say $\theta^*_{t+1,t}$, that he desires to detect. The experimenter also must specify the smallest acceptable probability of achieving Goal I when $\theta_{t+1,t} \geqq \theta^*_{t+1,t}$.

For Goal II it is assumed that the experimenter can specify a smallest value of each $\theta_{i+1,i}$, say $\theta^*_{i+1,i}$ $(i = 1, 2, \cdots, t)$ that he desires to detect. He also must specify the smallest acceptable probability of achieving Goal II when $\theta_{i+1,i} \geqq \theta^*_{i+1,i}$ $(i = 1, 2, \cdots, t)$.

The statistical procedure for achieving these goals is essentially the same for the two cases. The experimenter takes a predetermined number N_i (depending on the goal and the problem) of independent observations from the ith population. He computes the k sample variances s_i^2 and makes the ranking (4). He then makes the obvious decision. For Goal I he states that the populations that gave rise to the t smallest sample variances are the "best" populations, and the $k - t$ remaining populations are the "worst" populations. For Goal II he states that the populations that gave rise to the smallest, second smallest, ... , tth smallest sample variances are the "best," "second best," ... , "tth best" populations, respectively, and the $k - t$ remaining populations are the "worst" populations. The probability of achieving the goal, that is, the probability of a correct ranking, depends for each goal on the $\theta_{i+1,i}$ $(i = 1, 2, \cdots, k - 1)$ and the

n_i (d.f.) associated with the s_i^2 ($i = 1, 2, \cdots, k$). We shall show, for Goals I and II, how to determine the n_i so that the experimenter's requirements will be satisfied.

2.3. *Confidence statements associated with the procedure.* It is important to point out that if one adopts the procedure described above, it is possible for him to make useful confidence statements. These are given below without proof.

For Goal I if the n_i are chosen so that the probability of a correct ranking is P when $\theta_{t+1,t} = \theta^*_{t+1,t}$ and $\theta_{t,1} = \theta_{k,t+1} = 1$ (see Section 3.2), then after having taken the required number of observations, the experimenter can assert with confidence coefficient at least P that

$$(5) \qquad 1 \leqq \frac{\max\{\sigma^2_{(1)}, \sigma^2_{(2)}, \cdots, \sigma^2_{(t)}\}}{\sigma^2_{[t]}} \leqq \theta^*_{t+1,t}$$

where $\sigma^2_{(i)}$ denotes the variance of the population which yielded $s^2_{[i]}$ ($i = 1, 2, \cdots, t$). Similarly, for Goal II if the n_i are chosen so that the probability of a correct ranking is P when $\theta_{i+1,i} = \theta^*_{i+1,i}$ ($i = 1, 2, \cdots, t$) and $\theta_{k,t+1} = 1$ (see Section 3.2), then after having taken the required number of observations, the experimenter can assert with confidence coefficient at least P that

$$(6) \qquad 1/\theta^*_{i,i-1} \leqq \sigma^2_{(i)}/\sigma^2_{[i]} \leqq \theta^*_{i+1,i}$$

simultaneously for $i = 1, 2, \cdots, t$. (Here $\theta^*_{1,0} = 1$.)

For example, if $t = 1$ the goals coincide and the confidence statement becomes

$$(7) \qquad 1 \leqq \sigma^2_{(1)}/\sigma^2_{[1]} \leqq \theta^*_{2,1}.$$

This statement holds with confidence coefficient at least P regardless of the true configuration of the population variances. Thus, without knowing whether $\theta_{2,1} < \theta^*_{2,1}$ or $\theta_{2,1} \geqq \theta^*_{2,1}$, the experimenter still can assert with confidence coefficient at least P that the variance $\sigma^2_{(1)}$ of the population which he chose as having the smallest variance is not greater than $\theta^*_{2,1}\sigma^2_{[1]}$.

It follows from the above that the problem for $t = 1$ (say) could have been formulated in the following equivalent way: "How many observations must I take from each population in order that I will be able to assert with confidence coefficient at least P that the variance of the population that I choose as having the smallest variance is either the smallest one or at least not greater than $\theta^*_{2,1}$ times the smallest one?"

3. The probability of a correct ranking expressed as an integral.

3.1. *Arbitrary configuration of the population variances.* The probability of a correct ranking can be represented as

$$(8) \qquad \Pr[\max\{s^2_{(1)}, \cdots, s^2_{(t)}\} < \min\{s^2_{(t+1)}, \cdots, s^2_{(k)}\}],$$

$$(9) \qquad \Pr[s^2_{(1)} < s^2_{(2)} < \cdots < s^2_{(t)} < \min\{s^2_{(t+1)}, \cdots, s^2_{(k)}\}]$$

for Goals I and II, respectively. We shall give integral expressions for each of these probabilities.

We note that (8) can be written as

$$\sum_{j=1}^{t} \Pr[\max\{s^2_{(1)}, \cdots, s^2_{(j-1)}, s^2_{(j+1)}, \cdots, s^2_{(t)}\} < s^2_{(j)} < \min\{s^2_{(t+1)}, \cdots, s^2_{(k)}\}]$$

$$(10) \quad = \sum_{j=1}^{t} \Pr\left[s^2_{(\alpha)} < s^2_{(j)} < s^2_{(\beta)}; \quad \begin{matrix}(\alpha = 1, 2, \cdots, j-1, j+1, \cdots, t) \\ (\beta = t+1, t+2, \cdots, k)\end{matrix}\right]$$

$$= \sum_{j=1}^{t} \Pr\begin{bmatrix}\chi^2_{n(\alpha)} < (n_{(\alpha)}/n_{(j)})\,\theta_{j\alpha}\chi^2_{n(j)}; & (\alpha = 1, 2, \cdots, j-1, j+1, \cdots, t) \\ \chi^2_{n(\beta)} > (n_{(\beta)}/n_{(j)})\,\theta_{j\beta}\chi^2_{n(j)}; & (\beta = t+1, t+2, \cdots, k).\end{bmatrix}$$

If for each j the above probability is evaluated for $\chi^2_{n(j)}$ fixed (say at y), and the expectation is taken over y, then (10) can be written as

$$(11) \quad \sum_{j=1}^{t} \int_0^{\infty} \left[\prod_{\substack{\alpha=1 \\ \alpha \neq j}}^{t} F_{n(\alpha)}\left(\frac{n_{(\alpha)}}{n_{(j)}}\,\theta_{j\alpha}\,y\right)\right]\left[\prod_{\beta=t+1}^{k}\left\{1 - F_{n(\beta)}\left(\frac{n_{(\beta)}}{n_{(j)}}\,\theta_{j\beta}\,y\right)\right\}\right] f_{n(j)}(y)\,dy$$

where $f_{n(i)}(y)$ and $F_{n(i)}(y)$ are the probability density function (p.d.f.) and cumulative distribution function (c.d.f.), respectively, of the Gamma variable $\frac{1}{2}\chi^2_{n(i)}$.

The probability (8) can be evaluated for arbitrary n_i and $\theta_{i,j}$ $(i, j = 1, 2, \cdots, k)$ using (11). However, we shall be concerned with the case

$$(12) \qquad n_1 = n_2 = \cdots = n_k = n \text{ (say)},$$

and future probability calculations will be made for this special case.

The probability (8) also can be evaluated using an alternative expression which we give for the special case (12). The expression is

$$(13) \quad \sum_{j=1}^{t} \int \cdots \int \Gamma(kn/2) \prod_{i=1}^{k-1} u_i^{(n-2)/2}\,du_i \Big/ [\Gamma(n/2)]^k \left(1 + \sum_{i=1}^{k-1} u_i\right)^{nk/2}$$

where the limits of integration for $u_1, u_2, \cdots, u_{k-1}$ are $(0, \theta_{j,1})$, $(0, \theta_{j,2})$, $\cdots$, $(0, \theta_{j,j-1})$, $(0, \theta_{j,j+1})$, $(0, \theta_{j,j+2})$, $\cdots$, $(0, \theta_{j,t})$, $(\theta_{j,t+1}, \infty)$, $(\theta_{j,t+2}, \infty)$, $\cdots$, $(\theta_{j,k}, \infty)$, respectively. The above expression is derived by considering the joint distribution of the k independent $s^2_{(i)}$, making the transformation $u_i = \theta_{j,i}s^2_{(i)}/s^2_{(j)}$ $(i = 1, 2, \cdots, j-1, j+1, j+2, \cdots, k)$, $u_j = n\sum_{i=1}^{k} s^2_{(i)}/2\sigma^2_{[i]}$, and integrating out u_j as a Gamma function. Then renaming the u_i to make the subscripts consecutive, we obtain (13).

The probability (9) can be evaluated using different expressions. The expression corresponding to (11) will be omitted; the expression corresponding to (13) is derived in a similar way as (13) and is given by

$$(14) \qquad \int \cdots \int \frac{\Gamma(kn/2) \prod_{i=t}^{k-1} u_i^{(n-2)/2} \prod_{j=1}^{t-1} u_j^{[(k-j)n-2]/2} \prod_{i=1}^{k-1} du_i}{[\Gamma(n/2)]^k \left[1 + \sum_{i=1}^{k-1}\left(\prod_{j=1}^{i} u_j \Big/ \prod_{\alpha=t}^{i-1} u_\alpha\right)\right]^{nk/2}}$$

where we understand that $\prod_{\alpha=m}^{n} u_\alpha = 1$ if $m > n$, and the limits of integration for $u_1, \cdots, u_{t-1}; u_t, \cdots, u_{k-1}$ are $(\theta_{1,2}, \infty), (\theta_{2,3}, \infty), \cdots, (\theta_{t-1,t}, \infty)$; $(\theta_{t,t+1}, \infty), (\theta_{t,t+2}, \infty), \cdots, (\theta_{t,k}, \infty)$, respectively. The density functions contained in both (13) and (14) are multivariate generalizations of the F-distribution. The expressions (13) and (14) can be regarded as the *operating characteristic curves* with respect to a correct ranking for the procedures of Goals I and II, respectively.

3.2. *Least favorable configuration of the population variances.* For both Goal I and Goal II we are interested in finding the smallest value of n which will guarantee the requirements specified in Section 2.2. In order to do this it will be convenient to define a *least favorable configuration of the population variances.* For Goal I this configuration is defined by

$$\theta_{t,1} = \theta_{k,t+1} = 1; \qquad \theta_{t+1,t} \geqq 1; \tag{15}$$

and for Goal II it is defined by

$$\theta_{k,t+1} = 1; \qquad \theta_{i+1,i} \geqq 1 \qquad i = 1, 2, \cdots, t. \tag{16}$$

Since the probabilities (8) and (9) obviously are increasing functions of the $\theta_{i+1,i}$ $(i = 1, 2, \cdots, k-1)$, we see that in order to guarantee our requirements it is sufficient to evaluate these probabilities at

$$\theta_{t,1} = \theta_{k,t+1} = 1 \quad \text{and} \quad \theta_{t+1,t} = \theta^*_{t+1,t} = \theta^* \text{ (say)} \tag{17}$$

for Goal I, and at

$$\theta_{k,t+1} = 1 \quad \text{and} \quad \theta_{i+1,i} = \theta^*_{i+1,i} \qquad i = 1, 2, \cdots, t \tag{18}$$

for Goal II. The desired value of n then is the smallest integer which will make the probabilities, evaluated at these points, equal to or greater than the preassigned probability specified by the experimenter.

When (12) and (17) hold, the expressions (11) and (13) simplify considerably and we obtain

$$t \int_0^\infty [F_n(y)]^{t-1}[1 - F_n(y/\theta^*)]^{k-t} f_n(y)\, dy \tag{19}$$

$$t \int_{1/\theta^*}^\infty \cdots \int_{1/\theta^*}^\infty \int_0^1 \cdots \int_0^1 g_n(u_1, \cdots, u_{k-1})\, du_1 \cdots du_{t-1}\, du_t \cdots du_{k-1}, \tag{20}$$

respectively, where $g_n(u_1, \cdots, u_{k-1})$ is the same density function as is displayed in (13). When (12) and (18) hold, there is no corresponding simplification in (14), but $\theta_{t,i}$ is replaced by $1/\theta^*_{t+1,t}$ $(i = t+1, \cdots, k)$ and $\theta_{i,i+1}$ is replaced by $1/\theta^*_{i+1,i}$ $(i = 1, 2, \cdots, t-1)$.

4. Evaluation of the probability integrals. When n is even, the integrals (14), (19), and (20) can be evaluated in a straightforward manner, and the results can be expressed as rational functions of the $\theta_{i,j}$. However, when $k \geqq 3$ this method of evaluation becomes increasingly tedious as n increases and is inefficient even for small values of n. In some cases the probabilities also can be expressed as a finite sum of incomplete Beta functions, and using [12] the computations can be simplified in some cases.

When n is odd the integrations are more involved. For $k = 2$ and 3 the results can be expressed in terms of rational functions and inverse trigonometric functions of the $\theta_{i,j}$; for $k \geqq 4$ and n odd, no results were obtained.

For $k = 2$, the probabilities (13) and (14) coincide and are given by the incomplete Beta function $I_{\theta_{2,1}/(1+\theta_{2,1})}(n/2, n/2)$.

For $k = 3$ and n even, the probabilities (13) and (14) can be expressed in terms of finite sums of incomplete Beta functions. We give here three such sums. For Goal I when (15) holds and $\theta_{t+1,t} = \theta$, we have for $t = 1$ and 2, respectively

$$2I_{(\theta+1)/(\theta+2)}(n, n/2)I_{\theta/(1+\theta)}(n/2, n/2) - \sum_{j=1}^{n/2} b(n/2 - 1; n - 1 - j, 1/2)I_{\theta/(\theta+2)}(n/2 + j, n - j) \tag{21}$$

$$2I_{\theta/(1+2\theta)}(n/2, n)I_{\theta/(1+\theta)}(n/2, n/2) - \sum_{j=1}^{n/2} b(n/2 - 1; n + j - 2, 1/2)I_{2\theta/(2\theta+1)}(n + j - 1, n/2 - j + 1) \tag{22}$$

For Goal II when (16) holds and $\theta_{3,2} = \theta_{2,1} = \theta$, we have

$$I_{\theta^2/(1+\theta+\theta^2)}(n/2, n)I_{\theta/(1+\theta)}(n/2, n/2) - \theta(1 + \theta)^{-1} \sum_{j=1}^{n/2} b(n/2 - 1; n + j - 2, \theta/(1 + \theta)) \cdot I_{(\theta+\theta^2)/(1+\theta+\theta^2)}(n + j - 1, n/2 - j + 1). \tag{23}$$

In the above, the symbol $b(x; n, p)$ is the binomial probability and is equal to $C_x^n p^x (1 - p)^{n-x}$.

For $k = 3$ and n odd, general formulas were obtained for the probabilities (13) and (14), but for simplicity we shall give the results only for $n = 1$ and 3. For Goal I when (15) holds and $\theta_{t+1,t} = \theta$, we have

$$\frac{2}{\pi} \arctan\left\{\frac{\sqrt{\theta}\,[2(\theta - 1) + \sqrt{\theta + 2}]}{4\theta - 1}\right\} \qquad t = 1,\ n = 1 \tag{24}$$

$$\frac{2}{\pi} \arctan\left\{\frac{\sqrt{\theta}\,[2(\theta - 1) + \sqrt{\theta + 2}]}{4\theta - 1}\right\} + \frac{4\sqrt{\theta}\,(\theta - 1)}{\pi(\theta + 1)^2} - \frac{2\sqrt{\theta}\,(\theta - 1)(\theta^2 + 7\theta + 8)}{\pi(\theta + 1)^2(2 + \theta)^{5/2}} \qquad t = 1,\ n = 3 \tag{25}$$

$$\frac{2}{\pi} \arctan \frac{\theta}{\sqrt{1 + 2\theta}} \qquad t = 2,\ n = 1 \tag{26}$$

$$\frac{2}{\pi} \arctan \frac{\theta}{\sqrt{1 + 2\theta}} + \frac{2\theta(\theta - 1)(8\theta^2 + 7\theta + 1)}{\pi(\theta + 1)^2(2\theta + 1)^{5/2}}. \qquad t = 2,\ n = 3 \tag{27}$$

For Goal II when (16) holds and $\theta_{3,2} = \theta_{2,1} = \theta$, we have

$$(28)\quad \frac{2}{\pi}\arctan\{\sqrt{\theta}\,[\theta + 1 - \sqrt{\theta^2 + \theta + 1}]\} \qquad t = 2,\ n = 1$$

$$(29)\quad \frac{2}{\pi}\arctan\{\sqrt{\theta}\,[\theta + 1 - \sqrt{\theta^2 + \theta + 1}]\} + \frac{2\sqrt{\theta}\,(\theta - 1)}{\pi(\theta + 1)^2}\cdot\left[1 + \frac{(\theta - 1)(\theta^4 + 3\theta^3 + 5\theta^2 + 3\theta + 1)}{(\theta^2 + \theta + 1)^{5/2}}\right] \qquad t = 2,\ n = 3$$

Because of their simplicity we give in addition two general results for $n = 2$. For Goal I when (15) holds we have

$$(30)\quad t!\left[\prod_{i=1}^{t}\left(\frac{k - t}{\theta_{t+1,t}} + i\right)\right]^{-1},$$

while for Goal II when (16) holds we have

$$(31)\quad \left[\prod_{i=1}^{t}\sum_{j=i}^{k}\theta_{i,j}\right]^{-1}.$$

5. Large sample approximation to the probability. In Section 4 we pointed out that it is extremely tedious to compute exact probabilities when n is even and large, and that when n is odd these difficulties multiply considerably even for small n. In this section we shall show how large-sample theory can be used to find very good approximations to the required probabilities even for relatively small n.

We shall illustrate the method using a particular problem. The extension of the method to the general problem will be straightforward. Our principal tools will be the use of the transformation $y = \log_e s^2$ (see [2]), and the approach of certain multivariate distributions to multivariate normal distributions.

As our particular problem we shall consider Goal I for $k = 3$, $t = 1$ when (12) holds. Letting

$$(32)\quad X_i = \log_e (s^2_{(i)}/\sigma^2_{[i]}) \qquad i = 1, 2, 3$$

we see that we can write the probability of achieving our goal as

$$(33)\quad \Pr[s^2_{(1)} < s^2_{(2)},\ s^2_{(1)} < s^2_{(3)}] = \Pr[X_2 - X_1 > -\log_e \theta_{2,1},\ X_3 - X_1 > -\log_e \theta_{3,1}].$$

Now it can be shown (see [2]) that the expectation and variance are

$$(34)\quad E\{X_i\} = -\left(\frac{1}{n} + \frac{1}{3n^2}\right) + O(n^{-3}) \qquad i = 1, 2, 3$$

$$(35)\quad \mathrm{Var}\{X_i\} = \frac{d^2}{dx^2}[\log_e \Gamma(x)]\Big|_{x=n/2} = \frac{2}{n} + \frac{2}{n^2} + \frac{4}{3n^3} + O(n^{-5}) \sim \frac{2}{n - 1} \qquad i = 1, 2, 3.$$

Thus

$$\text{E}\,\{X_i - X_1\} = 0$$

$$\text{(36)}\qquad \text{Var}\,\{X_i - X_1\} = 2\left(\frac{2}{n} + \frac{2}{n^2} + \frac{4}{3n^3}\right) + O(n^{-5}) \sim \frac{4}{n-1} \qquad i = 2, 3$$

$$\text{Correlation}\,\{X_2 - X_1, X_3 - X_1\} = \tfrac{1}{2}.$$

Using the method of characteristic functions, it can be shown that the joint distribution of the chance variables

$$\text{(37)}\qquad Y_i = \sqrt{(n-1)/4}\,(X_{i+1} - X_1) \qquad i = 1, 2$$

approaches the bivariate normal distribution with means zero, variances unity, and correlation coefficient plus one-half. Thus the probability is given approximately by

$$\text{(38)}\qquad \int_{-\frac{1}{2}\sqrt{n-1}\,\log_e\theta_{2,1}}^{+\infty} \int_{-\frac{1}{2}\sqrt{n-1}\,\log_e\theta_{3,1}}^{+\infty} \frac{1}{\pi\sqrt{3}} \exp\left\{-\tfrac{2}{3}(y_1^2 - y_1y_2 + y_2^2)\right\} dy_1\,dy_2$$

This integral is tabulated [11]. When (17) holds the common value of the two lower limits is

$$\text{(39)}\qquad -\tfrac{1}{2}\sqrt{n-1}\,\log_e \theta^*.$$

More generally, for Goal I when (12) and (15) hold, the probability (8) can be expressed as

$$\text{(40)}\qquad t\int_{-\infty}^{+\infty} [G_n(y)]^{t-1}[1 - G_n(y - \sqrt{\tfrac{1}{2}(n-1)}\,\log_e \theta_{t+1,t})]^{k-t} g_n(y)\,dy$$

where $g_n(y)$ and $G_n(y)$ are the common p.d.f. and c.d.f., respectively, of the chance variables $\sqrt{\frac{1}{2}(n-1)}X_i$ $(i = 1, 2, 3)$. Since $g_n(y)$ and $G_n(y)$ approach the p.d.f. $f(y)$ and c.d.f. $F(y)$ of the standardized normal chance variable, it follows that the expression (40) approaches

$$\text{(41)}\qquad t\int_{-\infty}^{+\infty} [F(y)]^{t-1}[1 - F(y - d)]^{k-t} f(y)\,dy$$

where $d = \sqrt{\frac{1}{2}(n-1)}\,\log_e \theta_{t+1,t}$. A tabulation[4] [9] of (41) has been made as a function of d for certain pairs (t, k). These tables therefore can be used to find an approximation to the probability (40). The reader should note that (41) can also be written as

$$\text{(42)}\qquad (k - t)\int_{-\infty}^{+\infty} [F(y + d)]^{t}[1 - F(y)]^{k-t-1} f(y)\,dy$$

and it is this expression which appears in [9].

[4] These tables were computed by the National Bureau of Standards at the Institute for Numerical Analysis, Los Angeles. They are the basic tables from which Table I in [3] was derived.

6. Tables. Tables of the probability of a correct ranking have been prepared to assist the experimenter in designing experiments for ranking variances. All of the tables are computed for the case when n is the same for each population, and the least favorable configuration of the population variances holds, that is, (15) holds for Goal I and (16) holds for Goal II. The following key describes the tables.

Table Number	Goal Number	Value of		Conditions on the $\theta_{i,j}$
		k	t	
I	I or II	2	1	$\theta_{2,1} = \theta$
II	I or II	3	1	$\theta_{3,1} = \theta_{2,1} = \theta$
III	I	3	2	$\theta_{3,1} = \theta_{3,2} = \theta$
IV	II	3	2	$\theta_{3,2} = \theta_{2,1} = \theta$
V	I	4	1	$\theta_{4,1} = \theta_{3,1} = \theta_{2,1} = \theta$

All of the tables give the probabilities for $n = 1(1)20$ and $\theta = 1.0(0.2)2.2$.

Two probabilities are given in each cell of the tables (except for some of the cells in Table V). The correct probability, $P_c(\theta, n)$, is given to five decimal places; the normal approximation, $P_a(\theta, n)$, (see Section 5) to the correct probability, is given to four decimal places. The purpose of giving $P_a(\theta, n)$ is to indicate the magnitude of the error of the approximation, and to show for various goals, k, and t how this error varies as a function of θ and n. The magnitude of the error cannot be judged for most of the $P_a(\theta, n)$ in Table V since the $P_c(\theta, n)$ are given only for $n = 2(2)12$. Formulae had been developed for the computation of these $P_c(\theta, n)$ for n even, but such computations were found to be too laborious for $n > 12$; no similar formulae had been developed for n odd.

If we let $D(\theta, n) = P_c(\theta, n) - P_a(\theta, n)$, then the following properties would appear to hold for *all* of the tables: 1) For any fixed θ, $\lim_{n=\infty} D(\theta, n) = 0$; 2) For all n, $D(1, n) = 0$; 3) For any fixed n, $D(\theta, n)$ is continuous in θ; and 4) For any fixed $n > 1$, $\lim_{\theta=\infty} D(\theta, n) = 0$.

Based on the behavior of $P_c(\theta, n)$ and $P_a(\theta, n)$ in the range computed, the following additional properties would appear to hold:

(a) For Tables I, III, and IV: 1) For all n and $\theta > 1$, $D(\theta, n) > 0$; 2) For any fixed $n > 1$ there exists a value θ_n^0 of θ such that $D(\theta, n)$ is strictly increasing for $1 < \theta < \theta_n^0$ and strictly decreasing for $\theta > \theta_n^0$, while for $n = 1$, $\theta_1^0 = \infty$; 3) θ_n^0 is strictly decreasing with n; 4) For any fixed $\theta > 1$, $D(\theta, n)$ is strictly decreasing with n; and 5) $\text{Max}_{\theta>1} D(\theta, n)$ is strictly decreasing with n.

(b) For Tables II and V: 1) For $\theta > 1$, $D(\theta, 1) > 0$; 2) For any fixed $\theta > 1$ there exists a value n_θ^0 of n such that $D(\theta, n)$ is strictly decreasing for $n < n_\theta^0$ ($D(\theta, n_\theta^0) < 0$) and strictly increasing for $n > n_\theta^0$; and 3) n_θ^0 is strictly decreasing with θ.

The normal approximation is the same for Tables II and III. In general, for fixed k the *normal approximation* to the probability of achieving Goal I will be

TABLE I

Probability of a correct ranking as a function of the true variance ratio θ and the number of degrees of freedom (n) from each population:*

$P[s^2_{(1)} < s^2_{(2)}]$. *True variance ratio:* $\sigma^2_{[2]}/\sigma^2_{[1]} = \theta$

Degrees of freedom (n)	θ						
	1.0	1.2	1.4	1.6	1.8	2.0	2.2
1	0.50000 0.5000	0.52898 0.5000	0.55330 0.5000	0.57412 0.5000	0.59223 0.5000	0.60817 0.5000	0.62236 0.5000
2	0.50000 0.5000	0.54545 0.5363	0.58333 0.5668	0.61538 0.5929	0.64286 0.6156	0.66667 0.6355	0.68750 0.6533
3	0.50000 0.5000	0.55779 0.5513	0.60561 0.5940	0.64560 0.6302	0.67938 0.6612	0.70821 0.6880	0.73301 0.7114
4	0.50000 0.5000	0.56799 0.5627	0.62384 0.6146	0.67000 0.6580	0.70845 0.6946	0.74074 0.7258	0.76807 0.7526
5	0.50000 0.5000	0.57685 0.5723	0.63951 0.6317	0.69071 0.6808	0.73274 0.7217	0.76749 0.7559	0.79641 0.7848
6	0.50000 0.5000	0.58476 0.5808	0.65338 0.6466	0.70879 0.7004	0.75364 0.7445	0.79012 0.7808	0.81999 0.8110
7	0.50000 0.5000	0.59196 0.5883	0.66588 0.6599	0.72488 0.7176	0.77195 0.7642	0.80964 0.8020	0.83999 0.8329
8	0.50000 0.5000	0.59861 0.5953	0.67731 0.6719	0.73939 0.7329	0.78822 0.7816	0.82670 0.8204	0.85718 0.8515
9	0.50000 0.5000	0.60481 0.6017	0.68786 0.6829	0.75260 0.7469	0.80281 0.7971	0.84176 0.8365	0.87210 0.8676
10	0.50000 0.5000	0.61064 0.6078	0.69767 0.6931	0.76473 0.7596	0.81600 0.8110	0.85515 0.8508	0.88515 0.8815
11	0.50000 0.5000	0.61614 0.6134	0.70685 0.7026	0.77594 0.7713	0.82800 0.8237	0.86714 0.8635	0.89662 0.8937
12	0.50000 0.5000	0.62137 0.6188	0.71548 0.7116	0.78633 0.7821	0.83897 0.8352	0.87791 0.8748	0.90677 0.9045
13	0.50000 0.5000	0.62635 0.6239	0.72364 0.7200	0.79602 0.7922	0.84903 0.8457	0.88765 0.8850	0.91578 0.9140
14	0.50000 0.5000	0.63112 0.6288	0.73136 0.7279	0.80508 0.8016	0.85830 0.8553	0.89646 0.8943	0.92381 0.9224
15	0.50000 0.5000	0.63570 0.6335	0.73870 0.7355	0.81357 0.8104	0.86686 0.8643	0.90447 0.9026	0.93098 0.9299
16	0.50000 0.5000	0.64011 0.6380	0.74569 0.7427	0.82155 0.8186	0.87479 0.8725	0.91177 0.9102	0.93741 0.9366
17	0.50000 0.5000	0.64436 0.6423	0.75237 0.7495	0.82908 0.8264	0.88214 0.8801	0.91843 0.9172	0.94317 0.9426
18	0.50000 0.5000	0.64846 0.6465	0.75875 0.7561	0.83618 0.8337	0.88898 0.8872	0.92452 0.9235	0.94836 0.9480
19	0.50000 0.5000	0.65243 0.6505	0.76487 0.7623	0.84290 0.8406	0.89535 0.8938	0.93011 0.9293	0.95303 0.9528
20	0.50000 0.5000	0.65629 0.6544	0.77075 0.7683	0.84926 0.8472	0.90129 0.8999	0.93523 0.9346	0.95725 0.9571

* The five- and four-decimal place numbers in the body of the table are the correct probabilities and normal approximations to the correct probabilities, respectively.

TABLE II

Probability of a correct ranking as a function of the true variance ratio θ and the number of degrees of freedom (n) from each population:*

$P[s^2_{(1)} < \min(s^2_{(2)}, s^2_{(3)})]$. *True variance ratio:* $\sigma^2_{[2]}/\sigma^2_{[1]} = \sigma^2_{[3]}/\sigma^2_{[1]} = \theta$

Degrees of freedom (n)	θ						
	1.0	1.2	1.4	1.6	1.8	2.0	2.2
1	0.33333 0.3333	0.35835 0.3333	0.38020 0.3333	0.39958 0.3333	0.41696 0.3333	0.43269 0.3333	0.44705 0.3333
2	0.33333 0.3333	0.37500 0.3704	0.41176 0.4027	0.44444 0.4312	0.47368 0.4567	0.50000 0.4795	0.52381 0.5003
3	0.33333 0.3333	0.38792 0.3861	0.43633 0.4325	0.47928 0.4734	0.51746 0.5096	0.55149 0.5420	0.58192 0.5710
4	0.33333 0.3333	0.39880 0.3983	0.45702 0.4556	0.50846 0.5059	0.55379 0.5501	0.59375 0.5892	0.62901 0.6237
5	0.33333 0.3333	0.40838 0.4087	0.47519 0.4752	0.53390 0.5333	0.58515 0.5839	0.62976 0.6280	0.66858 0.6664
6	0.33333 0.3333	0.41703 0.4179	0.49156 0.4925	0.55663 0.5572	0.61284 0.6131	0.66113 0.6611	0.70254 0.7023
7	0.33333 0.3333	0.42498 0.4262	0.50654 0.5081	0.57725 0.5787	0.63766 0.6389	0.68886 0.6900	0.73209 0.7331
8	0.33333 0.3333	0.43238 0.4339	0.52042 0.5225	0.59618 0.5982	0.66016 0.6621	0.71362 0.7155	0.75807 0.7599
9	0.33333 0.3333	0.43933 0.4411	0.53339 0.5358	0.61369 0.6162	0.68070 0.6832	0.73590 0.7383	0.78106 0.7835
10	0.33333 0.3333	0.44590 0.4478	0.54559 0.5483	0.62998 0.6328	0.69956 0.7024	0.75605 0.7588	0.80153 0.8043
11	0.33333 0.3333	0.45209 0.4542	0.55713 0.5600	0.64523 0.6483	0.71697 0.7200	0.77436 0.7774	0.81982 0.8228
12	0.33333 0.3333	0.45813 0.4603	0.56808 0.5712	0.65955 0.6628	0.73310 0.7363	0.79106 0.7942	0.83622 0.8393
13	0.33333 0.3333	0.46386 0.4662	0.57852 0.5818	0.67305 0.6765	0.74809 0.7514	0.80633 0.8096	0.85097 0.8541
14	0.33333 0.3333	0.46937 0.4718	0.58849 0.5918	0.68580 0.6893	0.76205 0.7655	0.82033 0.8236	0.86427 0.8674
15	0.33333 0.3333	0.47469 0.4772	0.59805 0.6015	0.69788 0.7015	0.77508 0.7785	0.83319 0.8365	0.87628 0.8794
16	0.33333 0.3333	0.47983 0.4824	0.60723 0.6107	0.70934 0.7130	0.78727 0.7907	0.84503 0.8483	0.88715 0.8902
17	0.33333 0.3333	0.48481 0.4874	0.61606 0.6196	0.72024 0.7239	0.79869 0.8022	0.85594 0.8592	0.89699 0.9000
18	0.33333 0.3333	0.48965 0.4923	0.62456 0.6282	0.73061 0.7343	0.80940 0.8128	0.86601 0.8692	0.90592 0.9088
19	0.33333 0.3333	0.49434 0.4971	0.63277 0.6364	0.74050 0.7442	0.81947 0.8229	0.87532 0.8784	0.91403 0.9168
20	0.33333 0.3333	0.49892 0.5017	0.64069 0.6444	0.74994 0.7536	0.82893 0.8323	0.88393 0.8869	0.92140 0.9241

* The five- and four-decimal place numbers in the body of the table are the correct probabilities and normal approximations to the correct probabilities, respectively.

TABLE III

Probability of a correct ranking as a function of the true variance ratio θ and the number of degrees of freedom (n) from each population:*

$P[\max(s^2_{(1)}, s^2_{(2)}) < s^2_{(3)})]$. *True variance ratio:* $\sigma^2_{[3]}/\sigma^2_{[1]} = \sigma^2_{[3]}/\sigma^2_{[2]} = \theta$

Degrees of freedom (n)	θ						
	1.0	1.2	1.4	1.6	1.8	2.0	2.2
1	0.33333 0.3333	0.36729 0.3333	0.39650 0.3333	0.42200 0.3333	0.44450 0.3333	0.46456 0.3333	0.48258 0.3333
2	0.33333 0.3333	0.38503 0.3704	0.42982 0.4027	0.46886 0.4312	0.50311 0.4567	0.53333 0.4795	0.56019 0.5003
3	0.33333 0.3333	0.39827 0.3861	0.45473 0.4325	0.50372 0.4734	0.54633 0.5096	0.58350 0.5420	0.61609 0.5710
4	0.33333 0.3333	0.40927 0.3983	0.47539 0.4556	0.53244 0.5059	0.58156 0.5501	0.62388 0.5892	0.66046 0.6237
5	0.33333 0.3333	0.41889 0.4087	0.49340 0.4752	0.55726 0.5333	0.61165 0.5839	0.65790 0.6280	0.69730 0.6664
6	0.33333 0.3333	0.42755 0.4179	0.50954 0.4925	0.57929 0.5572	0.63803 0.6131	0.68732 0.6611	0.72867 0.7023
7	0.33333 0.3333	0.43549 0.4262	0.52426 0.5081	0.59920 0.5787	0.66157 0.6389	0.71318 0.6900	0.75583 0.7331
8	0.33333 0.3333	0.44286 0.4339	0.53787 0.5225	0.61741 0.5982	0.68282 0.6621	0.73618 0.7155	0.77960 0.7599
9	0.33333 0.3333	0.44978 0.4411	0.55056 0.5358	0.63421 0.6162	0.70216 0.6832	0.75681 0.7383	0.80058 0.7835
10	0.33333 0.3333	0.45632 0.4478	0.56247 0.5483	0.64981 0.6328	0.71988 0.7024	0.77542 0.7588	0.81920 0.8043
11	0.33333 0.3333	0.46253 0.4542	0.57372 0.5600	0.66438 0.6483	0.73620 0.7200	0.79229 0.7774	0.83582 0.8228
12	0.33333 0.3333	0.46846 0.4603	0.58438 0.5712	0.67804 0.6628	0.75130 0.7363	0.80766 0.7942	0.85071 0.8393
13	0.33333 0.3333	0.47414 0.4662	0.59453 0.5818	0.69089 0.6765	0.76530 0.7514	0.82170 0.8096	0.86409 0.8541
14	0.33333 0.3333	0.47960 0.4718	0.60423 0.5918	0.70302 0.6893	0.77832 0.7655	0.83455 0.8236	0.87614 0.8674
15	0.33333 0.3333	0.48487 0.4772	0.61350 0.6015	0.71449 0.7015	0.79047 0.7785	0.84636 0.8365	0.88702 0.8794
16	0.33333 0.3333	0.48996 0.4824	0.62241 0.6107	0.72537 0.7130	0.80182 0.7907	0.85721 0.8483	0.89686 0.8902
17	0.33333 0.3333	0.49490 0.4874	0.63096 0.6196	0.73570 0.7239	0.81245 0.8022	0.86721 0.8592	0.90578 0.9000
18	0.33333 0.3333	0.49968 0.4923	0.63919 0.6282	0.74553 0.7343	0.82241 0.8128	0.87644 0.8692	0.91388 0.9088
19	0.33333 0.3333	0.50433 0.4971	0.64713 0.6364	0.75489 0.7442	0.83176 0.8229	0.88496 0.8784	0.92123 0.9168
20	0.33333 0.3333	0.50885 0.5017	0.65480 0.6444	0.76382 0.7536	0.84055 0.8323	0.89285 0.8869	0.92791 0.9241

* The five- and four-decimal place numbers in the body of the table are the correct probabilities and normal approximations to the correct probabilities, respectively.

TABLE IV

Probability of a correct ranking as a function of the true variance ratio θ and the number of degrees of freedom (n) from each population:*

$P[s^2_{(1)} < s^2_{(2)} < s^2_{(3)}]$. *True variance ratio:* $\sigma^2_{[2]}/\sigma^2_{[1]} = \sigma^2_{[3]}/\sigma^2_{[2]} = \theta$

Degrees of freedom (n)	θ						
	1.0	1.2	1.4	1.6	1.8	2.0	2.2
1	0.16667	0.19716	0.22510	0.25067	0.27412	0.29567	0.31554
	0.1667	0.1667	0.1667	0.1667	0.1667	0.1667	0.1667
2	0.16667	0.21578	0.26223	0.30531	0.34484	0.38095	0.41387
	0.1667	0.2053	0.2412	0.2744	0.3051	0.3334	0.3597
3	0.16667	0.23033	0.29168	0.34869	0.40061	0.44740	0.48934
	0.1667	0.2225	0.2759	0.3257	0.3716	0.4136	0.4519
4	0.16667	0.24273	0.31696	0.38572	0.44763	0.50251	0.55078
	0.1667	0.2362	0.3037	0.3668	0.4244	0.4762	0.5227
5	0.16667	0.25379	0.33951	0.41846	0.48857	0.54962	0.60229
	0.1667	0.2480	0.3279	0.4022	0.4691	0.5284	0.5804
6	0.16667	0.26388	0.36009	0.44799	0.52488	0.59063	0.64624
	0.1667	0.2587	0.3497	0.4337	0.5083	0.5732	0.6289
7	0.16667	0.27325	0.37912	0.47495	0.55749	0.62675	0.68416
	0.1667	0.2684	0.3696	0.4622	0.5432	0.6122	0.6703
8	0.16667	0.28206	0.39693	0.49984	0.58703	0.65883	0.71719
	0.1667	0.2775	0.3882	0.4884	0.5746	0.6466	0.7061
9	0.16667	0.29040	0.41368	0.52292	0.61394	0.68750	0.74613
	0.1667	0.2861	0.4055	0.5125	0.6030	0.6772	0.7372
10	0.16667	0.29834	0.42953	0.54443	0.63857	0.71323	0.77162
	0.1667	0.2943	0.4219	0.5350	0.6290	0.7046	0.7644
11	0.16667	0.30594	0.44459	0.56456	0.66121	0.73645	0.79417
	0.1667	0.3021	0.4374	0.5559	0.6528	0.7291	0.7884
12	0.16667	0.31326	0.45896	0.58347	0.68209	0.75745	0.81419
	0.1667	0.3096	0.4522	0.5756	0.6746	0.7513	0.8096
13	0.16667	0.32031	0.47269	0.60127	0.70139	0.77650	0.83201
	0.1667	0.3168	0.4663	0.5940	0.6948	0.7713	0.8284
14	0.16667	0.32714	0.48585	0.61806	0.71927	0.79383	0.84794
	0.1667	0.3237	0.4798	0.6114	0.7134	0.7895	0.8451
15	0.16667	0.33375	0.49849	0.63393	0.73587	0.80963	0.86219
	0.1667	0.3304	0.4928	0.6278	0.7307	0.8060	0.8600
16	0.16667	0.34018	0.51065	0.64895	0.75131	0.82405	0.87497
	0.1667	0.3370	0.5052	0.6433	0.7467	0.8210	0.8733
17	0.16667	0.34643	0.52236	0.66319	0.76569	0.83725	0.88646
	0.1667	0.3433	0.5172	0.6580	0.7616	0.8347	0.8853
18	0.16667	0.35252	0.53365	0.67671	0.77911	0.84934	0.89680
	0.1667	0.3495	0.5287	0.6719	0.7755	0.8473	0.8960
19	0.16667	0.35847	0.54454	0.68955	0.79163	0.86044	0.90612
	0.1667	0.3555	0.5398	0.6851	0.7885	0.8588	0.9056
20	0.16667	0.36428	0.55508	0.70177	0.80334	0.87064	0.91454
	0.1667	0.3614	0.5506	0.6976	0.8006	0.8693	0.9143

* The five- and four-decimal place numbers in the body of the table are the correct probabilities and normal approximations to the correct probabilities, respectively.

TABLE V

Probability of a correct ranking as a function of the true variance ratio θ and the number of degrees of freedom (n) from each population:*

$P[s^2_{(1)} < \min(s^2_{(2)}, s^2_{(3)}, s^2_{(4)})]$. *True variance ratio:* $\sigma^2_{[2]}/\sigma^2_{[1]} = \sigma^2_{[3]}/\sigma^2_{[1]} = \sigma^2_{[4]}/\sigma^2_{[1]} = \theta$

Degrees of freedom (n)	θ						
	1.0	1.2	1.4	1.6	1.8	2.0	2.2
1	0.25000						
	0.2500	0.2500	0.2500	0.2500	0.2500	0.2500	0.2500
2	0.25000	0.28571	0.31818	0.34783	0.37500	0.40000	0.42308
	0.2500	0.2843	0.3149	0.3424	0.3674	0.3903	0.4112
3	0.25000						
	0.2500	0.2991	0.3437	0.3841	0.4207	0.4541	0.4845
4	0.25000	0.30799	0.36195	0.41153	0.45679	0.49792	0.53523
	0.2500	0.3107	0.3664	0.4169	0.4626	0.5039	0.5411
5	0.25000						
	0.2500	0.3206	0.3859	0.4450	0.4982	0.5458	0.5882
6	0.25000	0.32541	0.39643	0.46150	0.52010	0.57229	0.61848
	0.2500	0.3295	0.4033	0.4700	0.5296	0.5822	0.6285
7	0.25000						
	0.2500	0.3375	0.4192	0.4927	0.5577	0.6145	0.6638
8	0.25000	0.34027	0.42588	0.50370	0.57256	0.63244	0.68395
	0.2500	0.3450	0.4339	0.5136	0.5834	0.6435	0.6949
9	0.25000						
	0.2500	0.3521	0.4477	0.5330	0.6069	0.6698	0.7227
10	0.25000	0.35349	0.45201	0.54060	0.61741	0.68247	0.73677
	0.2500	0.3587	0.4607	0.5511	0.6286	0.6936	0.7476
11	0.25000						
	0.2500	0.3650	0.4730	0.5681	0.6487	0.7155	0.7699
12	0.25000	0.36556	0.47571	0.57349	0.65643	0.72475	0.77999
	0.2500	0.3711	0.4847	0.5842	0.6675	0.7355	0.7901
13	0.25000						
	0.2500	0.3769	0.4960	0.5994	0.6850	0.7539	0.8083
14	0.25000						
	0.2500	0.3825	0.5067	0.6138	0.7015	0.7709	0.8249
15	0.25000						
	0.2500	0.3879	0.5171	0.6276	0.7169	0.7866	0.8399
16	0.25000						
	0.2500	0.3931	0.5271	0.6407	0.7314	0.8011	0.8535
17	0.25000						
	0.2500	0.3982	0.5367	0.6532	0.7450	0.8146	0.8660
18	0.25000						
	0.2500	0.4031	0.5460	0.6651	0.7579	0.8270	0.8773
19	0.25000						
	0.2500	0.4079	0.5550	0.6766	0.7700	0.8386	0.8876
20	0.25000						
	0.2500	0.4126	0.5637	0.6875	0.7815	0.8494	0.8970

* The five- and four-decimal place numbers in the body of the table are the correct probabilities and normal approximations to the correct probabilities, respectively.

the same for choosing the t smallest as for choosing the t largest (i.e., the $k - t$ smallest). The following relationships hold for the entries in Tables II and III: 1) For all n and $\theta > 1$, $P_c(\theta, n \mid II) < P_c(\theta, n \mid III)$ and 2) For sufficiently large n and $\theta > 1$, $P_c(\theta, n \mid \text{II}) < P_a(\theta, n) < P_c(\theta, n \mid \text{III})$.

It should be noted that $P_a(\theta, n)$ is very close to $P_c(\theta, n)$ throughout the range of the various tables, and that therefore the normal approximation could be used with very good results to fill out Tables I to V for $n > 20$. The approximation also could be used (together with [9] or tables in [3]) for the construction of additional tables ($k \geqq 4$) for which exact formulae are unavailable or available but unwieldy.

All of the $P_c(\theta, n)$ should be correct to the five decimal places which are given. For Table I exact formulae were used in preference to interpolating in the incomplete Beta function tables. For Tables II, III and IV exact formulae were obtained for $n = 1$ (1) 8 (2) 20, and probabilities were computed to 8 decimal places; for $n = 9$ (2) 19 the probabilities were computed by interpolation on the values for even n using Everett's interpolation formula. For Table V exact formulae were obtained for $n = 2$ (2) 12.

All of the $P_a(\theta, n)$ were computed by setting Var $\{X_i\}$ equal to $2/(n - 1)$, (see equation (35)). The univariate normal, bivariate normal, and trivariate normal probabilities were found by interpolating in [8], [4], and [9], respectively. An empirically noted fact which was not only interesting, but also extremely useful from the viewpoint of checking the tables, was that for given θ the first differences of the probability as a function of n were strictly decreasing, and all of the higher differences were strictly increasing.

7. Example. The following is an example to show how the tables are to be used. The model of Section 2.1 is assumed throughout.

Given three populations. Suppose that it is desired to find which population has the smallest variance, and to guarantee that the probability of correctly choosing that population will be at least a) 0.60, b) 0.90, when $\sigma_{[2]}^2/\sigma_{[1]}^2 \geqq 1.8$. How many observations must be taken from each population? (The information in the tables is given in terms of d.f. The conversion of number of d.f. to number of observations will depend on the nature of the problem at hand.)

a) Refer to Table II. We see that 6 d.f. from each population will meet the requirements.

b) Refer to Table II. We see that 20 d.f. from each population is too small to meet the requirements. To estimate the correct number we proceed as follows. We compute $\sqrt{\frac{1}{2}(n - 1)} \log_e \theta_{2,1}^* = \sqrt{\frac{1}{2}(n - 1)} \log_e 1.8$ and set it equal to 2.2302. (The number 2.2302 is obtained from [3], Table I, column headed $k = 3$, $t = 1$ opposite $P = 0.90$). Solving for n we find that 30 d.f. from each population will meet the requirements.

In terms of the problems considered in this paper, the quantities given in the body of Tables I and II of [3] are $\sqrt{\frac{1}{2}(n - 1)} \log_e \theta_{t+1,t}$ and $\sqrt{\frac{1}{2}(n - 1)} \log_e \theta_{i+1,i}$ ($i = 1, 2$), respectively. These tables can be used for Goal I ($10 \geqq k \geqq$

2) and Goal II ($k = 3$), respectively. No corresponding tables exist for Goal II ($k \geqq 4$).

8. Acknowledgement. The writers would like to express their appreciation here for the efforts of Mr. Seiji Sugihara and Mrs. Shirley Hockett who spent long hours on the tedious and exacting job of computing the $P_c(\theta, n)$.

REFERENCES

[1] M. S. BARTLETT, "Properties of sufficiency and statistical tests," *Proc. Roy. Soc. London*, Vol. 160A (1937), pp. 268–282.

[2] M. S. BARTLETT AND D. G. KENDALL. "The statistical analysis of variance-heterogeneity and the logarithmic transformation," *J. Roy. Stat. Soc. Suppl.*, Vol. 8 (1946), pp. 128–138.

[3] R. E. BECHHOFER, "A single-sample multiple decision procedure for ranking means of normal populations with known variances," *Ann. Math. Stat.*, Vol. 25 (1954), pp. 16–39.

[4] G. BLANCH, *NAML Report* 52–53, National Bureau of Standards, Los Angeles, January 5, 1952.

[5] W. G. COCHRAN, "The distribution of the largest of a set of estimated variances as a fraction of their total," *Ann. Eugenics*, Vol. 11 (1941), pp. 47–52.

[6] C. EISENHART AND H. SOLOMON, "Significance of the largest of a set of sample estimates of variance," *Selected Techniques of Statistical Analysis*, McGraw-Hill Book Co., 1947, pp. 385–394.

[7] H. O. HARTLEY, "The maximum F-ratio as a shortcut test for heterogeneity of variance," *Biometrika*, Vol. 37 (1950), pp. 308–312.

[8] *Tables of Probability Functions*, Vol. 2 (MT14), National Bureau of Standards, 1942.

[9] National Bureau of Standards. Personal communication on unpublished tables.

[10] J. NEYMAN AND E. S. PEARSON, "On the problem of k samples," *Bull. Int. Acad. Polon. Sci. Lett. Cl. Sci. Math. Nat. Ser. A. Sci. Math.* (1931), pp. 460–481.

[11] K. PEARSON, *Tables for Statisticians and Biometricians*, Part II, 1st ed., Cambridge University Press, 1931.

[12] K. PEARSON, *Tables of the Incomplete Beta Function*, Cambridge University Press, 1934.

[13] C. THOMPSON AND M. MERRINGTON, "Tables for testing the homogeneity of a set of estimated variances," *Biometrika*, Vol. 33 (1946), pp. 296–304.