ON THE SIMULTANEOUS ANALYSIS OF VARIANCE TEST[1, 2]

BY K. V. RAMACHANDRAN[3]

University of North Carolina

1. Summary. In this paper we have solved certain distribution problems connected with the simultaneous analysis of variance test [1] and have proved that the power function of this test has the monotonicity property.

2. Introduction. It is well known that in situations involving the testing of the significance of k mean squares, the usual method of analysis of variance gives tests which are not independent. In these situations Ghosh [1] has recommended a test for the k mean squares which can be derived by the union-intersection principle [9].

The theory of simultaneous analysis of variance test and its use in certain problems in Public Health is given by Ghosh [1]. We shall be dealing only with certain distribution problems connected with the test and shall prove that the power function of the test has the monotonicity property. For further references, consult [6], [8], and [10].

3. Statement of the problem. Suppose we have k F-statistics

$$F_i = (S_i/S)\frac{m}{t_i}, \tag{3.1}$$

where $S_1, S_2, \cdots, S_k$ and S are mutually independent, S_i/σ^2 having a χ^2 distribution under the null hypothesis with t_i d.f., and S/σ^2 having a χ^2 distribution with m d.f. For example, S_1 might be the sum of squares for row effects; S_2, for column effects; and S, for error in a two-way layout. In the simultaneous analysis of variance situation [1] we are interested in evaluating

$$P[F_i \leqq a_i; i = 1, 2, \cdots, k] \tag{3.2}$$

for given a_i or for a given α to find the a_i's such that

$$P[F_i \leqq a_i; i = 1, 2, \cdots, k] = 1 - \alpha. \tag{3.3}$$

The optimum choice of a_i is not known. Ghosh [1] has intuitively suggested choosing a_i as proportional to t_i. A method of evaluating the probability on the left-hand side of (3.3) will be presented in Sections 4–6. The special case $t_i = 1$ $(i = 1, 2, \cdots, k)$ has been solved by Nair [3]. We shall also prove in Section 7 that the power of the simultaneous analysis of variance test has the monotonicity property.

Received Feb. 1, 1955.

[1] Work done under the sponsorship of the Office of Naval Research.

[2] Presented at the meetings of the Institute of Mathematical Statistics at Pasadena, California, June, 1954.

[3] Present address: Department of Statistics, University of Baroda, India.

4. Evaluation of the probability statement given on the left-hand side of (3.3). From (3.3) it is evident that

$$P\left[\frac{S_i}{S} \leqq a_i \frac{t_i}{m}; i = 1, 2, \cdots, k\right] = 1 - \alpha. \tag{4.1}$$

Hence the simultaneous analysis of variance test depends on the evaluation of expressions of the form

$$c \int_0^{b_1} \cdots \int_0^{b_k} \left\{\prod_{i=1}^k G_i^{(t_i-2)/2}\, dG_i \Big/ \left[1 + \sum_1^k G_i\right]^{(\Sigma t_i+m)/2}\right\}, \tag{4.2}$$

where c is a function of $t_1, t_2, \cdots, t_k$, and m and $b_i = (a_i t_i/m)$,

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

Usually we will be interested in obtaining $b_1, b_2, \cdots, b_k$ such that

$$c \int_0^{b_1} \cdots \int_0^{b_k} \left\{\prod_{i=1}^k G_i^{(t_i-2)/2}\, dG_i \Big/ \left[1 + \sum_1^k G_i\right]^{(\Sigma t_i+m)/2}\right\} = 1 - \alpha. \tag{4.3}$$

This can be evaluated as follows: Denoting the left-hand side of (4.3) by

$$I(b_1, b_2, \cdots, b_k; t_1, t_2, \cdots, t_k; m),$$

we get, by integration by parts,

$$\begin{aligned}
&I(b_1, b_2, \cdots, b_k; t_1, t_2, \cdots, t_k; m)\\
&\quad = \frac{-2c(t_1, t_2, \cdots, t_k; m)}{\left(\sum_1^k t_i + m - 2\right)}\\
&\quad \cdot \int_0^{b_1} \cdots \int_0^{b_{k-1}} \left[\left\{\sum_{i=1}^{k-1} G_i^{(t_i-2)/2}\, dG_i G_k^{(t_k-2)/2} \Big/ \left[1 + \sum_1^k G_i\right]^{(\Sigma t_i+m-2)/2}\right\}\right]_0^{b_k}\\
&\quad + \frac{t_k - 2}{2} \frac{c(t_1, t_2, \cdots, t_k; m)}{\dfrac{\sum t_i + m - 2}{2}}\\
&\quad \cdot \int_0^{b_1} \cdots \int_0^{b_k} \left\{\prod_{i=1}^{k-1} G_i^{(t_i-2)/2} G_k^{(t_k-4)/2} \prod_1^k dG_i \Big/ \left[1 + \sum_1^k G_i\right]^{(\Sigma t_i+m-2)/2}\right\},\\
&\quad = \frac{-b_k^{t_k-2/2} c(t_1, t_2, \cdots, t_k; m)}{(1 + b_k)^{(t_k+m-2)/2} \left(\dfrac{\sum t_i + m - 2}{2}\right)}\\
&\quad \cdot \int_0^{b_1/(1+b_k)} \cdots \int_0^{b_{k-1}/(1+b_k)} \left\{\prod_1^{k-1} G_i^{(t_i-2)/2}\, dG_i \Big/ \left[1 + \sum_1^{k-1} G_i\right]^{(\Sigma t_i+m-2)/2}\right.\\
&\quad + \frac{t_k - 2}{\sum t_i + m - 2} \frac{c(t_1, t_2, \cdots, t_k; m)}{c(t_1, t_2, \cdots, t_{k-1}, t_k - 2; m)}\\
&\quad \cdot I(b_1, b_2, \cdots, b_k; t_1, t_2, \cdots, t_{k-1}, t_k - 2; m).
\end{aligned} \tag{4.4}$$

Successive reduction will leave us with the evaluation of integrals of the form

$$\int_0^{l_1} \int_0^{l_2} \cdots \int_0^{l_j} \left\{ \prod_1^j G_i^{-\frac{1}{2}} \, dG_i \Big/ \left[1 + \sum_1^j G_i \right]^{(p+j)/2} \right\}. \tag{4.5}$$

Now it is easy to see that (4.5) is equivalent to

$$\int_0^{\infty} dV \int_0^{l_1 V} \cdots \int_0^{l_j V} \frac{e^{-V} V^{(p-2)/2}}{\Gamma\left(\frac{p+j}{2}\right)} \prod_1^j u_i^{-\frac{1}{2}} e^{-u_i} \, du_i. \tag{4.6}$$

Also from [6] we get

$$\int_0^x u^{-\frac{1}{2}} e^{-u} \, du = 2x^{1/2} e^{-x/3} \sum_0^{\infty} k_i^{(1)} x^i. \tag{4.7}$$

The convergence of the series on the right-hand side of (4.6) has been proved [7].

Thus the evaluation of (4.3) for given values of $b_1, b_2, \cdots, b_k$ can be carried out successively for different values of $t_1, t_2, \cdots, t_k$ and m by using the reduction formula (4.4). When m is large, (4.3) can be evaluated using tables of the incomplete gamma function [5].

It is easy to notice that the tabulation of (4.3) is rather tedious because of the large number of parameters involved. In the next section we shall consider the special but important case where $t_i = t$ $(i = 1, 2, \cdots, k)$. The general reduction formula given by (4.4) can be used also when $t_i = t$. But the special method which we shall use for the special case seems to be easier to handle than the use of the general formula. Also, it can easily be noticed that even though the method used for the special case generalises, the use of this method is rather tedious. Hence for the general case the general reduction formula is to be used; and for the special case, the special method to be discussed in the next section is to be used.

5. Special case when $t_i = t$ $(i = 1, 2, \cdots, k)$. In this case we have to obtain a "b" such that

$$1 - \alpha = c(k, t; m) \int_0^b \cdots \int_0^b \left\{ \prod_1^k G_i^{(t-2)/2} \, dG_i \Big/ \left[1 + \sum_1^k G_i \right]^{(kt+m)/2} \right\}, \tag{5.1}$$

where

$$c(k, t; m) = \frac{\Gamma\left(\frac{kt+m}{2}\right)}{\Gamma^k\left(\frac{t}{2}\right) \Gamma\left(\frac{m}{2}\right)}.$$

It is evident that (5.1) can be rewritten as

$$P\left[\frac{S_i}{S} \leqq b, i = 1, 2, \cdots, k\right] = P\left[\frac{S_{\max}}{S} \leqq b\right] = 1 - \alpha. \tag{5.2}$$

Let us call the statistics $u_k = S_{\max}/S$, the studentized largest chi-square. In order to obtain a "b" such that (5.2) is satisfied, we shall study the distribution of the studentized largest chi-square.

We shall derive in the next section certain mathematical results which we shall use to derive the distribution of u_k.

6. Power series expansion for the incomplete gamma-type integrals. Let

$$I(n, k; x) = \left[\int_0^x \frac{u^n e^{-u/2}}{2^{n+1}\Gamma(n+1)}\,du\right]^k. \tag{6.1}$$

Using methods similar to those given in [7], we find that an appropriate expansion for $I(n, k; x)$ is given by

$$I(n, k; x) = \frac{x^{k(n+1)}}{\Gamma^k(n+2)} \frac{\exp\left[-\frac{1}{2}(n+1)kx/(n+2)\right]}{2^{k(n+1)}} \sum_0^\infty A_i^{(k)} x^i, \tag{6.2}$$

where the A's satisfy the recurrence relation

$$\begin{aligned} A_i^{(k)}&\left[1 + \frac{i}{k(n+1)}\right] \\ &= \left[A_i^{(k-1)} - \frac{1}{2(n+2)} A_{i-1}^{(k-1)} + \cdots + \frac{(-1)^i}{i!} \frac{1}{2^i(n+2)^i} A_0^{(k-1)}\right] \\ &\quad + \frac{1}{2(n+2)} A_{i-1}^{(k)} \qquad (i = 0, 1, 2, \cdots). \end{aligned} \tag{6.3}$$

Notice that $A_0^{(k)} = 1$ and $A_1^{(k)} = 0$ for all k.

The convergence of the series $\sum_0^\infty A_i^{(k)} x^i$ is proved in the Appendix.

7. Distribution of the studentized largest chi-square. The p.d.f. of $S_{\max} = v$ is

$$\begin{aligned} p(v) &= \frac{k v^{(t-2)/2} e^{-v/2}}{2^{t/2}\Gamma\left(\frac{t}{2}\right)} \left[\int_0^v \frac{x^{(t-2)/2} e^{-x/2}}{2^{t/2}\Gamma\left(\frac{t}{2}\right)}\,dx\right]^{k-1} \\ &= \frac{k t v^{(kt-2)/2}}{2^{(kt+2)/2}\Gamma^k\left(\frac{t+2}{2}\right)} \exp\left[-\tfrac{1}{2}(kt+2)v/t+2)\right] \sum_0^\infty A_i^{(k-1)} v^i, \end{aligned} \tag{7.1}$$

using (6.2). And the p.d.f. of $S = y$ is

$$p(y) = \frac{y^{m/2-1} e^{-y/2}}{2^{m/2}\Gamma\left(\frac{m}{2}\right)}. \tag{7.2}$$

Multiplying (7.1) and (7.2), using the transformation $u = (v/y)$ and integrating with respect to y in the interval 0 to ∞, we get

$$p(u) = \frac{\frac{kt}{2}}{\Gamma^k\left(\frac{t+2}{2}\right)\Gamma\left(\frac{m}{2}\right)} \sum_0^\infty \frac{A_i^{(k-1)}\Gamma\left\{\frac{kt+m}{2} + i\right\} 2^i u^{(kt/2)+i-1}}{\left[1 + \frac{kt+2}{t+2}u\right]^{i+(kt+m)/2}}. \tag{7.3}$$

TABLE 1

Upper 5 per cent points of $u = S_{max}/S$ for different values of m and t (see formula 7.3) when $k = 2$

m	t = 1	2	3	4	6	8	10	12	16	20
5	9.55	7.88	7.15	6.70	6.20	5.92	5.72	5.59	5.42	5.28
6	8.50	6.90	6.21	5.79	5.32	5.04	4.87	4.75	4.59	4.46
7	7.84	6.28	5.62	5.21	4.76	4.50	4.35	4.22	4.09	3.96
8	7.39	5.86	5.23	4.81	4.38	4.14	3.98	3.86	3.74	3.61
10	6.81	5.32	4.70	4.32	3.90	3.66	3.51	3.40	3.28	3.16
12	6.43	4.99	4.38	4.01	3.61	3.38	3.23	3.12	3.00	2.88
16	6.02	4.62	4.02	3.66	3.27	3.05	2.90	2.79	2.66	2.56
20	5.81	4.41	3.81	3.46	3.06	2.85	2.71	2.60	2.47	2.37
24	5.66	4.29	3.69	3.34	2.95	2.74	2.59	2.49	2.36	2.25
∞	5.02	3.69	3.12	2.79	2.41	2.19	2.05	1.94	1.80	1.71

From (7.3) it is evident that the distribution of u can be tabulated using tables of the incomplete beta function [4]. Upper 5 per cent points of u are given in Table 1 for $k = 2$ and for different values of t and m. The methods presented in Sections 4–7 will enable us to evaluate integrals of the form

$$\int_0^\infty dy \int_0^{b_1 y} \cdots \int_0^{b_k y} y^m e^{-y/2}\, dy \prod_1^k x_i^{n_i} e^{-x_i/2} \cdot dx_i. \tag{7.4}$$

These integrals are found to be useful in obtaining lower bounds to the power of the Hartley test for equality of several variances from univariate normal populations [2].

8. Power function of the simultaneous analysis of variance test. In this situation, S_i/σ^2 has a noncentral χ^2 distribution with t_i d.f. and noncentrality parameter λ_i $(i = 1, 2, \cdots, k)$ and S/σ^2 has a χ^2 distribution with m d.f. We shall now prove the following theorem.

THEOREM. *The power function of the simultaneous analysis of variance test is a monotonic increasing function of the absolute value of the square root of each of the deviation parameters $\lambda_1, \lambda_2, \cdots, \lambda_k$ separately.*

PROOF. The second kind of error of the simultaneous analysis of variance test can easily be shown to be equal to

$$\beta = c \int_{\mathfrak{D}} \exp\left[-\tfrac{1}{2}\left(\sum_1^k \sum_1^{t_i} x_{i_j}^2 + \sum_1^m y_i^2\right)\right] \prod_1^k \prod_1^{t_i} dx_{i_j} \prod_1^m dy_i, \tag{8.1}$$

where the domain of integration $\mathfrak{D}$ is given by

$$\mathfrak{D}: \left\{\begin{array}{l} 0 \leqq \left(x_{11} + \sqrt{\lambda_1}\right)^2 + \sum_2^{t_1} x_{1i}^2 \leqq b_1 \sum_1^m y_i^2 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ 0 \leqq (x_{k_1} + \sqrt{\lambda_k})^2 + \sum_2^{t_k} x_{k_i}^2 \leqq b_k \sum_1^m y_i^2 \end{array}\right\},$$

and $c > 0$ is a pure constant independent of the λ's.

It is easy to see that each λ enters β in the same way. Hence we shall prove the theorem only for λ_1. For any other λ_i, the theorem is immediate. Notice that λ_1 occurs only with x_{11}. From (8.1) we get the limits of x_{11} to be

$$(8.2)\quad -\left(b_1 \sum_1^m y_i^2 - \sum_2^{t1} x_{1i}^2\right)^{1/2} - \sqrt{\lambda_1} \leqq x_{11} \leqq \left(b_1 \sum_1^m y_i^2 - \sum_2^{t1} x_{1i}^2\right)^{1/2} - \sqrt{\lambda_1}.$$

Now in (8.1) let us first perform the integration over x_{11}. The contribution to the total p.d.f. made by x_{11} is const. exp $(-\frac{1}{2}x_{11}^2)$. The upper and lower limits of the x_{11} integration are l_1 and l_2, given by

$$(8.3)\qquad l_1 = \left(b_1 \sum_1^m y_i^2 - \sum_2^{t1} x_{1i}^2\right)^{1/2} - \sqrt{\lambda_1},$$

and

$$l_2 = -\left(b_1 \sum_1^m y_i^2 - \sum_2^{t1} x_{1i}^2\right)^{1/2} - \sqrt{\lambda_1}.$$

If we now differentiate with respect to $\sqrt{\lambda_1}$, we get, through the x_{11} integral, an integrand which is

$$(8.4)\qquad \exp\left(-\frac{l_2^2}{2}\right) - \exp\left(-\frac{l_1^2}{2}\right).$$

For all positive values of $\sqrt{\lambda_1}$, the expression in (8.4) will be negative; and for all negative values of $\sqrt{\lambda_1}$, it will be positive. Thus

$$(8.5)\qquad \frac{\partial \beta}{\partial \sqrt{\lambda_1}} < 0 \quad \text{if} \quad \sqrt{\lambda_1} > 0, \qquad \text{and} \qquad > 0 \quad \text{if} \quad \sqrt{\lambda_1} < 0.$$

Similarly for any other λ_i. Therefore the second kind of error of the simultaneous analysis of variance test is a decreasing functon of each $|\sqrt{\lambda_i}|$ separately, and consequently the power of the test $(=1 - \beta)$ is an increasing function of $|\sqrt{\lambda_i}|$ separately. Hence the theorem.

9. Concluding remarks. The distribution of the studentized largest chi-square given in (7.3) has been noticed to be useful in obtaining useful simultaneous confidence bounds on certain parameters connected with the main effects of factorial experiments having m factors at s levels each. Extensive tables of the distribution of the studentized largest chi-square are being prepared and will be offered elsewhere. Using techniques similar to that given in Section 8, it is possible to prove that the power function of the test for the hypothesis that the m main effects of the s^m factorial experiments are simultaneously zero has the monotonicity property.

10. Acknowledgment. The author wishes to express his indebtedness to Prof. S. N. Roy for suggesting this problem, and for his help and guidance in the preparation of this paper.

APPENDIX

Convergence of the series on the right-hand side of (6.2). Consider

$$I(n; x) = \int_0^x \frac{u^n e^{-u}}{\Gamma(n+1)}\, du = \frac{x^{n+1}}{\Gamma(n+2)} \cdot \exp\left[-(n+1)x/(n+2)\right] \sum_0^\infty A_i^{(1)} x^i, \tag{A.1}$$

where

$$A_i^{(1)}\left[1 + \frac{i}{n+1}\right] = \frac{(-1)^i}{i!} \frac{1}{(n+2)^i} + \frac{1}{n+2} A_{i-1}^{(1)}. \tag{A.2}$$

Since we will be interested in cases where n is of the form $(r/2)$

$$(r = -1, 0, 1, \cdots),$$

we shall prove the convergence of the series on the right-hand side of (A.1) for the case $n = (r/2)$ $(r = -1, 0, 1, \cdots)$. The case when $r = -1$ has been already considered [7].

CASE 1. $n = 0$, i.e., $r = 0$.

In this case,

$$\begin{aligned} A_{2i+1}^{(1)} &= 0 && (i = 0, 1, 2, \cdots) \\ A_{2i}^{(1)} &= \frac{1}{2^{2i} 2 \cdot 3 \cdots (2i+1)} && (i = 1, 2, \cdots) \end{aligned} \tag{A.3}$$

and

$$A_0^{(1)} = 1.$$

Hence

$$\frac{A_{2i}^{(1)}}{A_{2i-2}^{(1)}} = \frac{1}{4 \cdot 2i(2i+1)} < \frac{1}{16i^2}. \tag{A.4}$$

Consequently $\sum_0^\infty A_i^{(1)}$ is convergent, and the value of the ratio of the ith to the $(i-1)$th term of the power series in (A.1) is less than $x^2/16i^2$. Therefore in this situation the series $\sum_0^\infty A_i^{(1)} x^i$ is absolutely convergent, and so the powers of the series are also convergent. It may be noticed that the series (6.2) is rather rapidly convergent, so that for a relatively small x only a few terms of the series will suffice for any degree of accuracy desired in practice.

CASE 2. $n > 0$, i.e., $r > 0$.

Now from (A.2), after a little simplification, we get

$$0 \leqq A_i^{(1)} = \frac{(n+1)^i}{(n+2)^i} \frac{(n+1)!}{(n+i+1)!} \cdot \left[\frac{n+2}{n+1} \frac{1}{2!} - \frac{(n+2)(n+3)}{(n+1)^2\, 3!} + \cdots + \frac{(-1)^i}{i!} \frac{(n+i)!}{(n+1)!(n+1)^{i-1}}\right]. \tag{A.5}$$

Hence

$$\frac{A_i^{(1)}}{A_{i-1}^{(1)}} = \frac{n+1}{(n+2)(n+i+1)}$$

(A.6)

$$\cdot \frac{\text{sum of first } (i+1) \text{ terms in } \left(1 + \dfrac{1}{n+1}\right)^{-(n+1)}}{\text{sum of first } i \text{ terms in } \left(1 + \dfrac{1}{n+1}\right)^{-(n+1)}}.$$

Therefore if i is large,

$$\frac{A_i^{(1)}}{A_{i-1}^{(1)}} < \frac{1}{i}. \tag{A.7}$$

Consequently $\sum_0^\infty A_i$ is convergent, and the value of the ratio of the ith to the $(i-1)$th term of the power series in (A.1) is less than x/i. Hence in this situation the series $\sum_0^\infty A_i^{(1)} x^i$ is absolutely convergent, and therefore the powers of the series are also convergent.

REFERENCES

[1] M. N. Ghosh, "Simultaneous test of linear hypotheses by the method of analysis of variance" (unpublished manuscript), 1953.

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

[3] K. R. Nair, "The studentized form of the extreme mean square test in the analysis of variance", *Biometrika*, Vol. 35 (1948), pp. 16–31.

[4] K. Pearson, *Tables of the Incomplete Beta Function*, The Biometrika Office, London, 1934.

[5] K. Pearson, *Tables of the Incomplete Gamma Function*, Cambridge University Press, 1946.

[6] K. C. S. Pillai, "On the distributions of midrange and semirange in samples from a normal population", *Ann. Math. Stat.*, Vol. 21 (1950), pp. 100–105.

[7] K. C. S. Pillai and K. V. Ramachandran, "On the distribution of the ratio of the ith observation in an ordered sample from a normal population to an independent estimate of the standard deviation", *Ann. Math. Stat.*, Vol. 25 (1954), pp. 565–572.

[8] K. V. Ramachandran, "Contributions to simultaneous confidence interval estimation" (to be submitted for publication).

[9] S. N. Roy, "On a heuristic method of test construction and its use in multivariate analysis", *Ann. Math. Stat.*, Vol. 24 (1953), pp. 220–238.

[10] S. N. Roy, "A report on some aspects of multivariate analysis", Institute of Statistics Mimeograph Series, No. 121 (1954).