ON DEPENDENT TESTS OF SIGNIFICANCE IN THE ANALYSIS OF VARIANCE[1]

By A. W. Kimball
Oak Ridge National Laboratory

1. Introduction. Some statisticians and other practitioners of the analysis of variance have expressed concern over the fact that many experimental designs lead to multiple tests of significance which are not independent in the probability sense. Factorials, latin squares, lattices, etc. have the advantage of enabling a research worker to test several hypotheses in one experiment, but all tests ordinarily depend on the same estimate of population variance. It is argued that whatever error is present in this estimate for a particular experiment will affect all tests of hypothesis in the same manner, and one tends either to accept or reject a large proportion of the hypotheses when the population variance is respectively overestimated or underestimated. The difficulty can be avoided by performing a separate experiment for each hypothesis to be tested, but this would contradict the whole philosophy of experimental design.

This paper deals with an attempt to evaluate the effect of dependency among the tests of significance when each experiment is treated as a unit regardless of the number of hypotheses tested per experiment. From this point of view if all null hypotheses are true, an error is committed if one or more of the hypotheses are rejected. It is shown that the probability of making no errors of the first kind in one experiment is greater when the tests are dependent than when they are independent. For those who prefer this way of looking at the problem, the doubts expressed in the first paragraph should be dispelled. The situation in which risks are calculated using the hypothesis rather than the experiment as a unit is not considered.

In the following sections it is assumed that samples are taken independently from normal populations having the same variance and having means additively related in a manner defined by the design of the experiment. These are the usual assumptions associated with analysis of variance models in which the parameters are population means (as distinguished from components of variance models).

2. Case of two dependent tests of hypothesis. We shall consider first the case of an analysis of variance in which two hypotheses are tested using the same error variance for each test. A well known example of this case occurs in the analysis of variance with two criteria of classification where the effects of both rows and columns are to be tested. In the usual cases, formulation as a general linear hypothesis leads to three quadratic forms, q_1, q_2, and q_3, which are independently distributed as χ^2 with n_1, n_2, and n_3 degrees of freedom, respectively.[2] The likelihood ratio statistics for testing the two hypotheses are then

$$F_1 = \frac{q_1/n_1}{q_3/n_3} \quad \text{and} \quad F_2 = \frac{q_2/n_2}{q_3/n_3}.$$

[1] This work was begun while the author was at the USAF School of Aviation Medicine, Randolph Field, Texas.

[2] For a more complete statement, see [1], p. 177.

If the critical region for the rejection of each null hypothesis is of size α, the probability of making no errors of the first kind is given by

$$P\{F_1 \leq F_{1\alpha}, F_2 \leq F_{2\alpha}\},$$

where $F_{1\alpha}$ and $F_{2\alpha}$ are the 100α per cent points of the distributions of F_1 and F_2, respectively. We shall prove[3] that

$$(1) \qquad P\{F_1 \leq F_{1\alpha}, F_2 \leq F_{2\alpha}\} > P\{F_1 \leq F_{1\alpha}\} \cdot P\{F_2 \leq F_{2\alpha}\}.$$

Since q_1, q_2, and q_3 are independent, their joint density is the product of three χ^2 densities. Clearly (1) may be written

$$(2) \qquad P\{q_1 \leq k_1 q_3, q_2 \leq k_2 q_3\} > P\{q_1 \leq k_1 q_3\} \cdot P\{q_2 \leq k_2 q_3\},$$

where $k_1 = n_1 F_{1\alpha}/n_3$, $k_2 = n_2 F_{2\alpha}/n_3$. Expressed in integral form, (2) become

$$(3) \qquad \int_0^\infty f_1(q_3) f_2(q_3) f_3(q_3)\, dq_3 > \int_0^\infty f_1(q_3) f_3(q_3)\, dq_3 \int_0^\infty f_2(q_3) f_3(q_3)\, dq_3,$$

where for $i = 1$ or 2, $f_i(q_3)$ is the integral from zero to $k_i q_3$ of a χ^2 density with n_i degrees of freedom, while $f_3(q_3)$ is the χ^2 density function with n_3 degrees of freedom. Since $f_1(q_3)$ and $f_2(q_3)$ are positive strictly monotonically increasing functions of q_3, and $f_3(q_3)$ is a density function, (3) may be written

$$(4) \qquad E[f_1(q_3) f_2(q_3)] > E[f_1(q_3)] \cdot E[f_2(q_3)],$$

where the expected values are taken over the probability distribution of χ^2.

The inequality expressed in (4) may be proved as a special case of the following theorem.[4]

THEOREM. *If $f(x) \geq 0$ and $g(x) \geq 0$ are both strictly monotonically increasing functions of a random variable x having the probability density $h(x)$ $(0 \leq x \leq \infty)$, and if both $f(x)$ and $g(x)$ have finite expectations, then*

$$E[f(x)g(x)] - E[f(x)] \cdot E[g(x)] > 0.$$

PROOF. We may write

$$E[f(x)\, g(x)] - E[f(x)] \cdot E[g(x)] = \int_0^\infty f(x) \{g(x) - E[g(x)]\} h(x)\, dx$$
$$= I,$$

say. Because of the monotonicity of $g(x)$, there must exist a quantity $x_0 > 0$ such that $g(x_0) = E[g(x)]$. It follows that

$$I = - \int^{x_0} f(x) \{E[g(x)] - g(x)\} h(x)\, dx$$
$$+ \int_{x_0}^\infty f(x) \{g(x) - E[g(x)]\} h(x)\, dx$$
$$= - I_1 + I_2,$$

[3] The trivial cases in which either $F_{1\alpha}$ or $F_{2\alpha}$ or both are either zero or infinite are excluded.

[4] The author is indebted to Dr. Max Halperin for the proof of this theorem.

say. Since

$$\int_0^\infty \{g(x) - E[g(x)]\}h(x)\, dx = 0,$$

we must have

$$-\int_0^{x_0} \{g(x) - E[g(x)]\}h(x)\, dx = \int_{x_0}^\infty \{g(x) - E[g(x)]\}h(x)\, dx$$
$$= J,$$

say. Furthermore, since $f(x)$ is a strictly monotonically increasing function of x, it follows that

$$I_1 < f(x_0)J, \qquad I_2 > f(x_0)J.$$

Therefore, $I_2 - I_1 = I > 0$, and the theorem is proved.

It is obvious that the foregoing theorem may be applied directly to prove the validity of (4). This in turn verifies (1).

Although the proof in this section was introduced by reference to a specific model in the analysis of variance, it is clearly valid for any two F-tests of significance which satisfy the relationships with respect to q_1, q_2, and q_3, and in general for any n_1, n_2, and n_3.

3. Extension to several dependent tests of hypothesis. The extension of (1) to more than two tests of significance is straightforward. If there are three F-tests, we must show that

$$(5) \quad \int_0^\infty f_0(q_3) f_1(q_3) f_2(q_3) f_3(q_3)\, dq_3$$
$$> \int_0^\infty f_0(q_3) f_3(q_3)\, dq_3 \int_0^\infty f_1(q_3) f_3(q_3)\, dq_3 \int_0^\infty f_2(q_3) f_3(q_3)\, dq_3,$$

where $f_0(q_3)$ is a function similar to $f_1(q_3)$ and $f_2(q_3)$ resulting from the third test of significance. From Section 2 we know that

$$(6) \quad \int_0^\infty f_0(q_3) f_1(q_3) f_2(q_3) f_3(q_3)\, dq_3 > \int_0^\infty f_0(q_3) f_3(q_3)\, dq_3 \int_0^\infty f_1(q_3) f_2(q_3) f_3(q_3)\, dq_3,$$

since $f_0(q_3)$ and $f_1(q_3)f_2(q_3)$ satisfy the requirements of the theorem. But from (3) we may make an obvious substitution in the right-hand side of (6) which reduces it to (5). Clearly this simple procedure may be repeated as often as necessary to prove the extension of (1) to any number of F-tests of significance in which the numerators of the test statistics are mutually independent, and each is independent of the denominator which is the same for all statistics.

The author wishes to thank Professors J. W. Tukey and H. Levene for helpful suggestions in the preparation of this manuscript.

REFERENCE

[1] S. S. Wilks, *Mathematical Statistics*, Princeton University Press, 1946.