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.