EXTENDED VALIDITY OF RESULTS FOR VARIANCE COMPONENTS MODELS.

John E. Walsh
Southern Methodist University

ABSTRACT

Models used for investigation of variance components in analysis of variance are moderately restrictive for point estimation and quite restrictive, assuming normality, when tests and confidence regions are developed. However, the same results as occur for these models can be obtained for generalizations that are much less restrictive. In particular, results based on normality for the random variables of the model, and on their independence, remain valid for situations where each observation can have a different nonnormal distribution and dependence between each two observations can be different. The level of generalization for a model depends on what is being investigated and decreases as the number of variance components being investigated increases. Considered are the N-way classification models of various kinds and some specific examples. Similar extensions can be made for virtually all types of variance components models. The verification consists in showing that the statistics involved have the same joint distribution for the extended model as they do for the original model. Some ways of investigating outliers of nested designs, which also apply to the generalizations, are discussed.

INTRODUCTION

Consideration of variance components models for analysis of

variance can be useful in statistical investigations (for example, see the discussion in Chapter 16 of ref. 1). However, the assumptions made in such linear models are moderately restrictive for cases of point estimation and quite restrictive for cases involving significance tests and confidence regions. More specifically, for point estimation the random variables of the model are required to be mutually uncorrelated and every random variable of a given type is required to have the same variance. For tests and confidence regions, these random variables are also required to have a joint normal distribution (implies mutual independence of the random variables and the same distribution for variables of a given type).

The statistical results that occur for the usual variance components models can also be obtained for generalizations of these models that use substantially less restrictive assumptions. In particular, this is the case for results based on the assumptions of normality, mutual independence of the random variables, and a common distribution for variables of the same type. These results can be valid for situations where each observation has a different nonnormal distribution and each two observations are dependent in a different way.

To illustrate the generalization approach, consider the usual model for a balanced one-way classification. An observation is denoted by \( Y_{ij} \), where \( i = 1, \ldots, I \) and \( j = 1, \ldots, J \). The model assumes that

\[
Y_{ij} = \mu + a_i + b_{ij},
\]

where \( \mu \) is a constant, each \( a_i \) is a random variable with zero mean and variance \( \sigma_a^2 \), and each \( b_{ij} \) is a random variable with zero mean and variance \( \sigma_b^2 \). The statistics used for investigating \( \sigma_a^2 \) are (for example, see ref.1)

\[
\sum_{i=1}^{I} \sum_{j=1}^{J} (Y_{ij} - Y_{..})^2 \quad \text{and} \quad \sum_{i=1}^{I} (Y_{i.} - Y_{..})^2
\]

where

\[
Y_{i.} = \sum_{j=1}^{J} Y_{ij}/J, \quad Y_{..} = \sum_{i=1}^{I} \sum_{j=1}^{J} Y_{ij}/IJ.
\]

The extension of model (1) permits the presence of an additional