VARIANCE COMPONENTS IN THE UNBALANCED 2-WAY NESTED CLASSIFICATION

By S. R. Searle

New Zealand Dairy Board, Wellington, New Zealand

Introduction. Sampling variances of estimates of components of variance obtained from data that are balanced (having the same number of observations in all subclasses) are easily derived because the mean squares in the analysis of variance are independent and distributed as χ^2. The variance component estimates are linear functions of the mean squares and their variances can be derived accordingly, although their distributions are, in general, unknown. When the data are not balanced, however, and there are unequal numbers of observations in the subclasses the mean squares are no longer independent and they do not have χ^2-distributions. Methods of deriving expressions for the sampling variances of the variance component estimates are developed for these situations in an earlier paper [3] and applied to the 1-way classification. A second paper [4] gives these expressions for the 2-way factorial classification, and extension to the 2-way hierarchical (nested) classification is presented here.

Model and analysis of variance. The earlier work discussed sampling variances of variance component estimates obtained by Henderson's Method 1 [2] from data having unequal subclass numbers, based on the completely random model, namely Eisenhart's Model II, [1]. The same situation is considered here for the 2-way nested classification.

The linear model for an observation x_{ijk} is taken as

$$x_{ijk} = \mu + \alpha_i + \beta_{ij} + e_{ijk}$$

where μ is the general mean, α_i is the effect due to the ith main classification, β_{ij} is the effect due to the jth sub-class within the ith main classification, and e_{ijk} is the residual error term peculiar to x_{ijk}. We suppose the number of classes in the main classification is a, so that $i = 1, \cdots, a$; and that there are c_i sub-classes within each of these so that $j = 1, \cdots, c_i$. The total number of such sub-classes will be represented by b, giving $b = \sum_{i=1}^{a} c_i$. The number of observations in the jth subclass of the ith class is taken as n_{ij}. All terms of the model (except μ) are assumed to be normally distributed random variables with zero means and variances σ_α^2, σ_β^2 and σ_e^2. These are the variance components to be estimated, along with the sampling variances of their estimates.

Received December 4, 1960; revised June 30, 1961.