Fisher, R.A. 1935. *Statistical Methods for Research Workers*. London: Oliver and Boyd.

Greenhouse, S. W., and Geisser, S. 1959. On methods in the analysis of profile data. *Psychometrika*, 32: 95–112.

Huynh, H., and Feldt, L. S. 1976. Estimation of the box correction for degrees of freedom from sample data in the randomized block and split plot designs, *Journal of Educational Statistics*, 1: 69–82.

Laird, N. 1988. Missing data in longitudinal studies. *Statistics in Medicine*, 7: 305–315.

Kenward, M.G., and Roger, J.H. 1997. Small sample inference for fixed effects from restricted maximum likelihood, *Biometrics*, 53: 983–997.

McCulloch, C. E., and Searle, S. R. 2000. *Generalized, Linear, and Mixed Models*. New York: Wiley.

Scheffe, H. 1959. *The Analysis of Variance*. New York: Wiley.

Searle, S. R. (1970). Linear Models, New York: Wiley.

Searle, S. R., Casella, G., and McCulloch, C. E. 1992. *Variance Components*. New York: Wiley.

Tietjen, G.L. 1974. Exact and approximate tests for unbalanced random effects designs. *Biometrics*, 30: 573–581.

Comment: Anova as a Tool for Structuring and Understanding Hierarchical Models

Andrew Gelman

I agree with McCulloch that hierarchical models (which consider the persons in the experiment as a random sample from a hypothetical population) are a good idea for repeated measures data. As McCulloch points out, this assumption is typically well motivated by the goal of extrapolating the experimental findings to the population. He also explains why classical ANOVA is not the best tool for exploring such data.

However, if we think about ANOVA more broadly—as a way of structuring statisical analyses rather than as one particular set of computations—I believe a unification is possible that will give us the benefits of hierarchical modeling (efficient estimation, even under imbalance, missing data, nonnormality, and other realistic data conditions, as discussed by McCulloch), while also preserving the benefits of ANOVA (the summary of a complicated model in terms of batches of coefficients and variance parameters). In my own areas of application, I have not found much use for F-tests and p-values, but I have found concepts, such as decomposition of degrees of freedom, and estimation of the importance of different components of variation, to be helpful. In the embrace of maximum likelihood (or, more generally, Bayesian) estimation, I do not want to lose these helpful summaries.

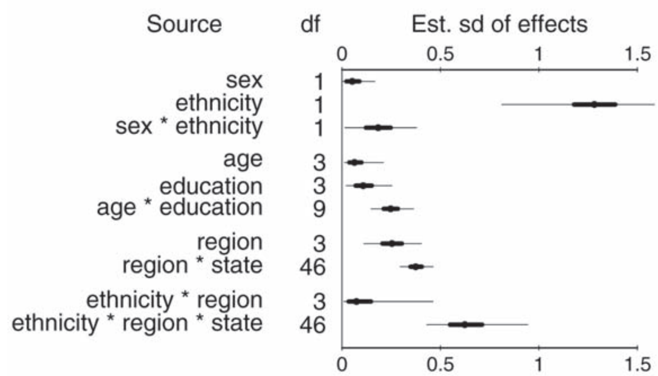

Figure 1. ANOVA display for a multilevel logistic regression model fit to survey data on voter preferences. The estimate, 50% interval, and 95% interval are shown for the finite-sample standard deviation of each batch of effects in the model. From Gelman (2005).

So, how can we get the most out of ANOVA in a likelihood/Bayesian modeling context? Each row of the ANOVA table corresponds to, and labels, a different batch of coefficients in the linear model. For example, McCulloch's Table 1 has five rows, and his equation (1) has five subsetted coefficients. I would like to see, for each row of the ANOVA table, an estimate of the standard deviation of its batch of coefficients. This idea is discussed fully in Gelman (2005); for an example, see Figure 1.

Thus, I agree with McCulloch's point that hierarchical models and likelihood-based estimation can work better than classical ANOVA, especially in complex settings; but I would like to reserve a role for the concepts of ANOVA to help us understand fitted models. The goal here is not to test hypotheses of zero effects but rather to summarize the importance of each batch of coefficients.

One of the most important advantages of model-summary tools is that they can facilitate the fitting of multiple models. With regard to McCulloch's particular application discussed here, I would be interested in seeing interactions between person and alcohol and between person and pregnenolone. These batches of interactions would represent random samples from the interactions in the entire population and thus would have additional variance components, each estimated from n-1 degrees of freedom. An ANOVA display along the lines of Figure 1. could help us understand such a model and could lead to further investigation of treatment effects of interest. Hierarchical estimation tools of the sort discussed by McCulloch are crucial in allowing us to fit these models.

Reference

Gelman, A. 2005. Analysis of variance: Why it is more important than ever (with discussion). *Annals of Statistics*, 33: 1-53.