A pragmatic bootstrap for inference about variance components in multilevel models

Daniel Farewell[1] and Vern Farewell[2]

[1] Department of Primary Care and Public Health, Cardiff University, UK
[2] MRC Biostatistics Unit, Cambridge, UK

In a large Welsh study of neighbourhood social cohesion, the analysis of multilevel questionniare data focussed on variance components. In general, inference about the sampling distribution of estimates of variance components in multilevel models is difficult. These estimates tend to arise from numerical optimisation routines, rather than as closed-form functions of the data, so their properties are not immediately apparent. Use of bootstrap methodology provides a possible approach to inference, but was computationally demanding for the Welsh study. We present a novel method of inference that mimics certain features of a bootstrap, while avoiding the computational expense of refitting the multilevel model for each bootstrap realisation.
In brief, we explicitly express the updating step of an EM algorithm for the variance components as a function of the posterior second moments of the random effects. Provided the data may be partitioned into independent groups, we can imitate a non-parametric bootstrap by sampling (with replacement) from this set of posterior second moments. This replicate then updates the bootstrap maximum likelihood estimates, a procedure that is computationally trivial and so may be repeated as many times as desired for bootstrap inference. The approximation is improved by adjusting for the rate of convergence of the EM alogrithm, and is over 300 times faster than full bootstrap inference.