

Resampling Multilevel Models

Rien van der Leeden¹, Erik Meijer², and Frank M. T. A. Busing¹

¹ Leiden University, Department of Psychology

² University of Groningen, Faculty of Economics, and RAND Corporation

11.1 Introduction

Estimation in (linear) multilevel models usually relies on maximum likelihood (ML) methods. The various computer programs for multilevel analysis employ versions of full information (FIML) and restricted maximum likelihood (REML) methods. Two vital assumptions underlying ML theory are that (a) the residuals are i.i.d. with a distribution from a specified class, usually the multivariate normal, and (b) the sample size is (sufficiently) large. More specifically, the attractive properties of FIML estimators—consistency, (asymptotic) efficiency and (asymptotic) normality—are derived from the supposition that the sample size goes to infinity. In practice, however, these assumptions will frequently be met only approximately, which may lead to severely biased estimators and incorrect standard errors [7].

Resampling methods can be used to obtain consistent estimators of bias and standard errors, and to obtain confidence intervals and bias-corrected estimates of model parameters. A number of general resampling approaches are found throughout the literature, of which we mention the bootstrap and the jackknife, permutation, and cross-validation. Bayesian Markov chain Monte Carlo methods [e.g., 16, 23] and simulation-based estimators for mixed non-linear models [e.g., 27, 61, 63] are also closely related to these resampling methods. Particularly, bootstrap and jackknife procedures have proven to be methods that yield satisfactory results in small sample situations under minimal assumptions. In this chapter, we discuss resampling of multilevel data by means of bootstrap and jackknife procedures. In cases where the assumptions underlying ML methods for estimating multilevel models are violated, bootstrap and jackknife estimation may provide useful alternatives.

The application of the bootstrap and the jackknife to multilevel models is not straightforward. For the bootstrap, there are several possibilities, depending on the nature of the data and the assumptions one is willing to make. Each of them, however, has its own associated problems. In this chapter, we discuss three different approaches, which are derived from general principles of bootstrap theory and apply concepts adapted from bootstrapping regression models.

The application of the jackknife to multilevel analysis is based on a version of the delete- m jackknife approach [57, Section 2.3]. In this procedure, subsamples are obtained from the original sample by successively removing mutually exclusive groups of size m . For the application to multilevel analysis, the delete- m jackknife has been adapted for groups of unequal size.

Bootstrap and jackknife estimation in the context of multilevel analysis have been studied by several authors and for various models and situations. Laird and Louis [33, 34] discuss empirical Bayes confidence intervals based on bootstrap samples and Moulton and Zeger [45] study bootstrapping a model for repeated measurements. Bellmann et al. [3] use a parametric bootstrap for a panel data model that is essentially a multilevel model and Booth [4] similarly uses a parametric bootstrap for generalized linear mixed models. Goldstein [24] presents an iterated bootstrap based on the results of Kuk [32]. A theoretical analysis of nonparametric bootstrapping of balanced two-level models without covariates is given in Davison and Hinkley [14, pp. 100–102]. Our discussion largely follows the lines of the systematic development of resampling methods for multilevel models in Busing et al. [8, 9, 10, 12], Van der Leeden et al. [68], and Meijer et al. [42, 43].

In this chapter, we focus on FIML estimation for multilevel linear models with two levels. The ideas, however, are directly applicable to REML estimators and generalize straightforwardly to models with three or more levels.

In Section 11.2, we define the model upon which we center our discussion and we elaborate on the consequences of violating the assumptions of ML estimation in multilevel models. In Section 11.3, we briefly discuss the general ideas of the bootstrap and the jackknife, and the specific issues involved in the application of the bootstrap to regression models. Section 11.4 provides an extensive discussion of the three methods for bootstrap implementation, as well as a number of approaches to construct confidence intervals. In Section 11.5, the application of the jackknife to multilevel models is discussed. Section 11.6 briefly discusses the availability of resampling options in existing software for multilevel analysis. In Section 11.7, we discuss some results of evaluation studies of the various resampling approaches and in Section 11.8, we briefly discuss application of the presented approaches to other types of multilevel models, and we mention various possible extensions and alternatives to the (bootstrap) resampling methods presented.