Multilevel Item Response Theory Models

The item response data structure is hierarchical since item responses are nested within respondents. Often respondents are also grouped into larger units and variables are available that characterize the respondents and the higher-level units. An item response modeling framework is discussed that includes a multilevel population model for the respondents and takes such a hierarchical data structure into account. An important application area is in education, where response observations are grouped in students and students grouped in schools. Several school effectiveness research studies are discussed. The hierarchical item response model is extended in several directions to handle latent explanatory variables, model latent individual growth, and identify clusters of respondents.

6.1 Introduction: School Effectiveness Research

School effectiveness research is concerned with exploring differences within and between schools. The objective is to investigate the relationship between explanatory and outcome factors. This involves choosing an outcome variable, such as student’s ability, and studying differences among schools after adjusting for relevant background variables. Interest is focused on the relative size of school differences and the factors that explain these differences and influence student learning.

Typically, in school effectiveness research, students are nested in classrooms, classrooms in schools, schools within school systems, and so on. A generally acceptable statistical model in the assessment of school effectiveness therefore requires the deployment of multilevel analysis techniques. A multilevel model takes the hierarchical structure into account, and variance components are modeled at each sampling level. As a result, homogeneity of results of individual pupils in the same school is accounted for since pupils in the same school share common experiences. Specifically, a multilevel model
describes relationships between one or more “outcome” variables (examination results, attitudes), school and teacher characteristics (teacher’s attitude, financial resources, class size), and student characteristics (achievements, social background). In the study of school effectiveness research, multilevel modeling has become so conspicuous since it allows for the analysis of individual and group-level effects and cross-level interactions.

The appropriateness of multilevel models in the assessment of school effectiveness was shown by Aitkin and Longford (1986). Since that time, most of the research has focused on multilevel modeling of hierarchically structured educational data and the assessment of relevant input and output indicators (e.g., Goldstein, 2003; Longford, 1993; Raudenbush and Bryk, 2002).

6.2 Nonlinear Mixed Effects Models

Two approaches for analyzing variables from different levels at one single level have been criticized. The first disaggregates all higher-order variables to the individual level. That is, data from higher levels are assigned to a much larger number of units at level 1. All disaggregated values are assumed to be independent of each other, which is a misspecification that threatens the validity of the inferences. In the second approach, observations at level 1 are aggregated to the higher level. As a result, all within-group information is lost. Relations between aggregated variables can be much stronger and different from the relations between nonaggregated variables. Snijders and Bosker (1999) give a complete overview of potential (statistical) errors when the clustered structure of the data is ignored.

Statistical models under a broad variety of names have been developed that can handle the different levels of the data. The models capture the between- and within-subject variances by modeling the data in two stages. At the first stage, a regression function is specified for the observations for each subject. Each subject has its own regression function. The same covariates are used across subjects, but the regression coefficients are allowed to vary. At the second stage, the regression coefficients that are allowed to vary are considered to be random outcome variables. These random outcome variables are referred to as random (regression) effects or random coefficients. This explains the term random coefficient models that is often used in sociological research and econometrics (De Leeuw and Kreft, 1986; Longford, 1993). However, there are a variety of names in the literature to describe versions of the same model. In educational and sociological research, the name multilevel model is often used (Goldstein, 2003; Snijders and Bosker, 1999). In biometric research, mixed effects model or random effects model are common terms (Laird and Ware, 1982; Hedeker and Gibbons, 2006; Longford, 1987). Other common names are variance component models (Dempster, Rubin and Tsutakawa, 1981) and hierarchical linear models (Raudenbush and Bryk, 2002).