Module 12: Cross-Classified Multilevel Models

George Leckie
Centre for Multilevel Modelling

Pre-requisites
• Modules 1-5,11

Contents

If you find this module helpful and wish to cite it in your research, please use the following citation:

Leckie, G. (2013). Cross-Classified Multilevel Models - Concepts. LEMMA VLE Module 12, 1-54. http://www.bristol.ac.uk/cmm/learning/course.html

Address for correspondence:

George Leckie
Centre for Multilevel Modelling
University of Bristol
35 Berkeley Square
Bristol, BS8 1JA
UK

g.leckie@bristol.ac.uk

What are Cross-Classified Multilevel Models?

In the previous modules we illustrated two- and three-level models for analysing hierarchical data structures whereby lower level units, such as students, are nested within higher level units, such as schools, and where these higher level units may in turn be nested within further clusters (or groupings) such as school districts, regions or countries. With hierarchical data structures each lower level unit belongs to one and only one higher level unit. For example, each student attends one school, each school is located within one school district and so on. However, social reality is more complicated than this and so social and behavioural data often do not follow strict hierarchies. Two types of non-hierarchical data structures which often appear in practice are cross-classified and multiple membership structures. In this module, we describe cross-classified data structures and cross-classified multilevel models which can be used to analyse them. We leave discussion of multiple membership data structures and multiple membership multilevel models until Module 13.

In cross-classified data, lower level units do not belong to one and only one higher level unit. Rather, lower level units belong to pairs or combinations of higher level units formed by crossing two or more higher level classifications with one another. An example in educational research arises in studies of student attainment where students are nested within schools and are separately nested within neighbourhoods. Schools and neighbourhoods are not typically nested within one another as not all students from the same school live in the same neighbourhood, nor do all students from the same neighbourhood attend the same school. Rather, schools and neighbourhoods are crossed with one another, with each student potentially belonging to any combination of school and neighbourhood. Students are described as nested within the cells of the two-way cross-classification of schools by neighbourhoods. An example in health services research arises in studies of hospital patient outcomes. Hospitals and general practitioners (GPs, i.e. family doctors) are cross-classified as GPs tend to refer their patients to different hospitals depending on patient need while hospitals typically treat patients who have been referred by many different GPs. There is nothing to stop data structures being even more complex and having three or more higher classifications and we shall consider examples of such data in this module. Many further examples of cross-classified structures are described in C4.4 of Module 4.

It is important to incorporate cross-classified structures in to our models when they arise in the data and lead the higher level clusters to differ substantially from one another on the response variable under study. Naively fitting the nearest equivalent hierarchical model to cross-classified data will lead us to misattribute response variation to the included levels (van Landeghem et al., 2005; Moerbeek, 2004; van den Noortgate et al., 2005; Tranmer and Steele, 2001). This in turn may lead us to draw misleading conclusions about the relative importance of different sources of influence on the response. For example, fitting a students-within-schools two-level model of student attainment and ignoring the fact that students are simultaneously, but separately, nested within neighbourhoods will likely lead us to overstate the importance of schools as a source of variation in student attainment. Some of the variation that we attribute to schools may be better characterised as neighbourhood-to- neighbourhood differences in attainment. Our naïve analysis would therefore overstate the importance of schools on student attainment and would ignore the role of neighbourhoods (i.e. neighbourhood policies, practices, context and compositional effects). Furthermore, by incorrectly modelling the dependency in the data we will likely obtain biased standard errors for the predictor variables, particularly those measured at higher levels. We therefore run the risk of making incorrect inferences and drawing misleading conclusions about the relationships being studied. For example, including neighbourhood-level predictor variables in our students-within-schools two-level model, but ignoring neighbourhood as a level in the model will typically lead us to severely underestimate the standard errors on these neighbourhood-level variables. When we then go on to test the significance of these variables, we will run the risk of making type 1 errors of inference.

Introduction to the Example Dataset

We will illustrate multiple membership models in the context of the same school effectiveness application which was analysed in Module 11. Readers familiar with this application may wish to skip the next two paragraphs.

In educational research, there is considerable interest in measuring the effects that schools have on students' educational achievements. Measuring the effects that schools have on their students is after all a necessary first step to learning how schools' policies and practices combine to generate differences between schools. Governments are also often interested in measuring school effects, typically for school accountability purposes, but often to also provide parents with information to help guide school choice. However, in nearly all education systems, there are substantial differences between schools in their students' attainments at intake (i.e. when students first arrive at their schools). For the purposes of researching the effects of schools' policies and practices, holding schools accountable, or informing school choice, schools should not be compared simply in terms of their average exam results as these differences will, at least in part, be driven by these initial differences.

Traditional studies of school effects attempt to measure the 'true' effects that schools have on their students by fitting two-level students-within-schools multilevel models to students' exam scores where covariate adjustments are made for students' initial scores, and typically for a range of other student background characteristics. The school-level residuals from these models are then argued to measure the effects that schools have on their students having adjusted for the non random selection of students into schools. These effects are interpreted as measuring the influences schools have on their students' academic progress (improvement or change in attainment) while they attend their schools. In school effectiveness research these influences are referred to as 'value-added' effects.

In Module 11, we used three-level models to explore the stability of school effects (across cohorts) on students' academic progress during secondary schooling in the English education system. However, it is easy to think of further sources of clustering or influence on child learning which might also be important to consider and explore. One interesting example is the role that schools attended in an earlier phase of education may continue to exert on students after they have left these schools. For example, we might ask: Does primary school attended (ages 4 to 11) predict student academic progress during secondary schooling (ages 11 to 16)? Such long lasting or continued effects of schools attended in earlier phases of schooling are sometimes referred to as carry over school effects. However, as not all students from the same primary school typically go on to attend the same secondary school, the data can no longer be described as hierarchical. Rather, the data are described as cross-classified with students (level 1) nested within the cells of a cross-classification of secondary schools and primary schools (both conceptually at level 2). In this module, we shall introduce cross-classified multilevel models to analyse the potential role of primary schools on students' progress during secondary schooling in the English education system.

We shall then go on to consider the further nesting of students within residential neighbourhoods and additionally the nesting of schools within administrative educational regions referred to as local authorities (LAs).[1] Accounting for neighbourhoods and LAs leads to an even more complex cross-classified data structure.[2]

Another potential influence on child learning is the residential neighbourhood within which a child lives. For example, we might ask: Do communities where adults have few educational qualifications entrench low academic aspirations in children growing up there? Alternatively, we might ask: Do gangs and children playing truant in deprived neighbourhoods disrupt not only their own education but also that of other children in the street? An important issue when considering neighbourhood effects is the spatial scale at which they are purported to operate. There are a multitude of spatial scales in UK geography. We focus on lower super output areas (LSOAs), which were designed using the 2001 UK Census and are defined to be fairly consistent in size, having a mean population of approximately 1,500, and to reflect as far as possible social homogeneity.

In England, secondary schools are organised into 150 LAs. Traditionally, LAs controlled the distribution of government funds across schools, co-ordinated school admissions, and were the direct employers of all teachers and staff in many schools. While over the last few years there has been a reduction of LAs' powers, we might still expect to identify LA effects in the data. If nothing else, we would expect LA effects to pick up geographic variation in student attainment that exists across England.

We shall use data from England's National Pupil Database (NPD), a census of all students in state (i.e. government funded) schools in England. The data are provided by the Department for Education (http://www.education.gov.uk). The NPD records students' academic attainments and a limited number of background characteristics. We focus on the academic cohort of students who sat their General Certificate of Secondary Education (GCSE) examinations (age 16 years) in London schools in 2010 and their Key Stage 2 (KS2) examinations (age 11 years) five years earlier in 2005.[3] [4] This cohort is the third of the three cohorts analysed in Module 11 and the data analysed here are identical to the data analysed for that cohort.

[1] LAs correspond to school districts in the U.S.

[2] Note that pupils may live in one LA, but be schooled in another LA. One could therefore envisage LA of residence effects on pupil attainment as well as the LA of schooling effects which we consider here. See Fielding et al (2006) for a multilevel cross-classified analysis which simultaneously considers both sources of clustering.

[3] GCSE examinations are taken in the last year of secondary schooling. Successful GCSE results are often a requirement for taking A-level examinations (age 18 years) which in turn are a common type of university entrance determinant. For those who leave school at 16 years of age, GCSE results are their main job market qualification.

[4] KS2 examinations are taken in the last year of primary schooling.