AQMeN
Applied Quantitative Methods Network

Multilevel Modelling

Focus on Methods
Helen Brown, AQMeN Senior Statistician, University of Edinburgh

Multilevel Models

Multilevel models (MLMs) are becoming an increasingly popular method of data analysis. While 'conventional' methods such as ANOVA and multiple regression have been used for many years, there are many situations where an improved analysis can be obtained by using a MLM. Here a brief introduction to MLMs is provided and the key differences from the more 'conventional' methods are explained.

How do multilevel models differ from conventional models?

Multilevel models (MLMs) provide a flexible framework for modelling the variances and correlations in a dataset. Often this is achieved by a subtle change in assumptions about groupings in the data. For example, in a study of educational attainment, schools may be assumed to come from a 'population' of schools. School effects are assumed to have a distribution rather than being treated as 'fixed' (as they are in conventional models) and are then described as 'random' effects. Alternatively regression lines or curves may be assumed to vary randomly (e.g. growth curves), or patterns of correlation in the data can be modelled directly, e.g. between repeated responses in a longitudinal study or between a set of multivariate outcomes.

What are the benefits?

Wider inference: In some situations a wider inference can be made. For example, when schools are modelled as random, the mean difference in attainment between males and females will have a wider confidence interval than it would in a conventional model. Inference can then be made to the 'population' of schools rather than just the schools assessed.

Shrunken estimates: Sometimes it is of interest to compare between groups, e.g. attainment level between schools, or health outcome between hospitals. In a MLM the group means will be 'shrunken' compared to the means obtained from a conventional analysis, with the means tending to be closer to the population average. In general shrinkage is greater for smaller schools or hospitals and the chance of obtaining spurious outlying results is reduced.

Gain in efficiency: In some situations the use of a MLM will lead to smaller standard errors particularly when there are few observations within some groups. For example, if only one patient has surgery at certain hospitals, all information from these hospitals is lost in a conventional analysis fitting hospitals as fixed. However, data from patients at these hospitals will still contribute to the analysis in a MLM fitting hospitals as random.

Overcoming problems caused by missing data: An MLM will often overcome the problems caused by missing data. This is an advantage particularly in longitudinal studies (e.g. panel studies) where there are often dropouts. However, it is important that the data can be assumed missing at random.

More appropriate variances and correlations can be modelled using a MLM. For example: in a longitudinal analysis the correlation between observations on the same person may become less for measurements that are further apart in time; measurements on females may be more variable than those on males; there may be less correlation between districts further apart.

Are there any disadvantages?

MLMs are a more complex approach than conventional models and involve more assumptions and practical considerations. This can make the results more difficult to explain to those without a quantitative background. While these disadvantages do not usually outweigh the benefits, it is important to understand the underlying assumptions and to consider how to explain results from MLMs to others.

To learn more…

AQMeN will be holding a training event on MLMs later in 2010 and will include an introductory session in the 'Methods Tester Session' to be held on 13 May 2010 in Edinburgh. A more detailed introduction including links to textbooks and publications is provided by LEMMA, a node of the ESRC's National Centre for Research Methods. LEMMA have also developed an excellent set of eLearning modules which may be accessed free of charge.