15

Software for Multilevel Modelling, Resources and Further Developments

15.1 Software packages and resources

There are now many software packages that will carry out multilevel modelling. Most of the major statistical packages have features for the basic models, and can fit more complex or specialized models using a macro language. We shall not give a detailed review here, but merely list the main packages together with web addresses for further information. Recently published reviews of some of the packages are those of De Leeuw and Kreft (2001), Zhou et al. (1999) and Fein and Lissitz (2000). The Centre for Multilevel Modelling (www.multilevel.ioe.ac.uk) maintains a series of reviews.

Table 15.1 lists the packages, together with the internet address and a brief note.

The methodological literature on multilevel modelling is growing rapidly as is the literature on applications. The Centre for Multilevel Modelling endeavours to maintain a selection of the methodological literature and links to other resources such as websites and training materials. The Centre also produces a twice-yearly electronic newsletter with free subscription: this contains articles about current developments, reviews etc. A collection of datasets together with training materials and a version of the MLwiN package that will work with these datasets, is freely available at http://www.tramss.data-archive.ac.uk. Another useful resource for multilevel modelling is http://www.ats.ucla.edu/stat/mlm/default.htm .

There is a very active email discussion group that can be accessed and joined at http://www.jiscmail.ac.uk/lists/multilevel.html. The group serves as a means of exchanging information and suggestions about data analysis.

15.2 Further developments

While the models and estimation procedures described in preceding chapters are extensive there remain many important areas for future development. Some important ones are, briefly, as follows.

Table 15.1   Available software packages

Name	Web site	Note
aML	http://www.applied-ml.com	Concentrates on event history and multiprocess models. Maximum likelihood estimation
ASREML	http://www.vsn-intl.com/asreml/	Same features as GENSTAT
BAYESX	http://www.stat.uni-meunchen.de/~lang/bayesx/bayesx.html	General-purpose MCMC estimation. Continuous and discrete responses with nested and cross-classified structures. Concentrates on semiparametric regression
BMDP	http://www.statsol.ie/bmdp/bmdp.htm	Variance components model and serial correlations for nested structures. Maximum likelihood and GEE
EGRET	http://www.cytel.com/products/egret	Discrete responses for nested structures up to two levels. Maximum likelihood estimation
GENSTAT	http://www.nag.co.uk/stats/tt_soft.asp	Continuous and discrete responses. Nested and cross-classified structures. Maximum likelihood estimation
HLM	http://www.ssicentral.com/hlm	Continuous and discrete responses up to three levels. Serial correlation structures; measurement errors. Maximum and quasilikelihood estimation
LIMDEP	www.limdep.com	General-purpose econometric software. Continuous and discrete responses. Nested structures. Maximum likelihood estimation
LISREL	http://www.ssicentral.com/lisrel.htm	Multilevel structural equations. Nested data structures. Maximum likelihood estimation
MIXOR, MIXREG	http://tigger.uic.edu/~hedeker/mix.html	A suite of programs for continuous and discrete response multicategory models up to three levels. Maximum likelihood estimation
MLWiN	www.multilevel.ioe.ac.uk	General-purpose package. Continuous and discrete responses for nested, cross-classified and multiple membership structures for any number of levels. Serial correlation structures; event history models; factor analysis; measurement errors. Maximum and quasi likelihood estimation; MCMC estimation
MPLUS	http://www.statmodel.com/mplus	Continuous and discrete responses. Nested structures. Multilevel structural equations. Maximum likelihood estimation
OSWALD	http://www.maths.lancs.ac.uk/Software/Oswald/	Works with s plus for analysis of serial correlation and event history data. Maximum likelihood and GEE estimation

(*Continued*)

Table 15.1 (*Continued*)

Name	Web site	Note
SAS (version 8)	http://www.sas.com/products/ sassystem/release82/	General-purpose package. Continuous and discrete responses for nested and cross-classified structures up to two levels. Serial correlation structures. Maximum and quasilikelihood estimation
S-PLUS 2000	http://www.insightful.com	General-purpose software. Continuous and discrete responses for nested structures
SPSS (version 10)	http://www.spss.com/spss10	General-purpose package. Handles basic continuous and discrete response models for nested structures up to two levels. Maximum likelihood estimation
STATA	http://www.stata.com	General-purpose package. Continuous and discrete responses. Nested and cross-classified structures up to two levels. Structural equation models (GLLAMM program), serial correlations. Maximum likelihood estimation
SYSTAT	http://www.spssscience.com/systat	General-purpose package. Continuous responses. Nested structures up to two levels. Serial correlations. Maximum likelihood estimation.
WINBUGS	http://www.mrc-bsu.cam.ac.uk	General-purpose package. Uses MCMC to fit a very wide range of models via a statistical control language. Continuous and discrete responses, measurement errors, factor analysis, serial correlations and more

Further work on missing data, especially where it is informatively missing, is needed, and sensitivity analysis to the various assumptions that have to be made would be useful. There is a considerable amount of work to be done on measurement error models, including errors at different levels of a data hierarchy and satisfactory ways of estimating measurement error distributions. Further work on the specification and fitting of structural equation models would be useful. More work on optimal design is needed as is work on diagnostics; in both cases a useful start has been made.

On the computational side, with an increasing use of very large datasets, ways of improving the efficiency of existing methods, especially for MCMC modelling, would be very useful, although the increasing power and memory capacity of computers will be of some help here.

While many of the procedures described in these chapters provide powerful tools for the exploration of complex data structures, in many areas there are few datasets that are collected in such a way that allows these models to be applied. Thus, information that identifies units may be absent, or research questions may have been phrased in ways that assume only single-level analysis is to be carried out. One of the most pressing needs, therefore, is for researchers to become familiar with multilevel modelling techniques and their possibilities. Although multilevel modelling is becoming more

widely used and understood, there is still a need for good introductory materials and training generally and the introduction of courses into undergraduate and postgraduate teaching is welcome. Given the ubiquity of multilevel data, in time it should become a standard technique for data analysis in the same way that ordinary regression has been during the latter part of the twentieth century.