Discussiones Mathematicae
Probability and Statistics 31 (2011) 87–89
doi:10.7151/dmps.1131

APPLICATION OF HLM TO DATA WITH MULTILEVEL STRUCTURE

Vítor Valente

Open University and CEAUL
e-mail: vmfvalente@netcabo.pt

AND

Teresa A. Oliveira

Departamento de Ciencias e Tecnologia
Open University and CEAUL
Rua da Escola Politécnica, 141–147, 1269–001 Lisboa, Portugal
e-mail: toliveir@univ-ab.pt

Abstract

Many data sets analyzed in human and social sciences have a multilevel or hierarchical structure. By hierarchy we mean that units of a certain level (also referred micro units) are grouped into, or nested within, higher level (or macro) units. In these cases, the units within a cluster tend to be more different than units from other clusters, i.e., they are correlated. Thus, unlike in the classical setting where there exists a single source of variation between observational units, the heterogeneity between clusters introduces an additional source of variation and complicates the analysis.

Collecting data on Educational Research often does not follow the principles of simple random sample, suspected by classical regression, but rather a sample by nested clusters. Selected to students and also the contextual units to which they belong such as classes, courses, schools, neighborhoods or regions, and so forth.

Using classical regression bias is produced in the typical error of measurement and an increased likelihood of committing errors of inference.

The hierarchical linear or multilevel models are most suitable because they consider the hierarchical relationships and also provide estimates on the contextual variability of regression coefficients.

In practice, often the data structures are not hierarchical, are more complex structures such as cross-classification (level 2 or macro). For example, students (level 1 or micro) to attend different courses at a school while in other schools there are students who attend the same courses.

Two examples of application to academic achievement of students are presented. First, a model of cross-classification of level 2 is used. Second, a hierarchical model of two levels (students and schools) is presented, taking into account the different areas of science - scientific-humanistic courses and technology courses.

Keywords: hierarchical linear model, multilevel model, cross-classification models, academic achievement.

2010 Mathematics Subject Classification: 62–07, 62P25.

References

[1] W. Browne, MCMC Estimation in MLwiN, Centre of Multilevel Modelling (Institute of Education, University of London, 2004).

[2] A. Bryk and S. Raudenbush, Hierarchical Linear Models: applications and data analysis methods (Sage Publications, London, 1992).

[3] H. Goldstein, Multilevel Statistical Models (2nd Edition, Edward Arnold, London, 1995).

[4] J. Hox, Multilevel Analysis: Techniques and Applications (Lawrence Erlbaum Associates, Mahwah, NJ 2002).

[5] I. Kreft and J. de Leeuw, Introducing Multilevel Modeling (Sage Publications, London, 1998).

[6] MLwiN, version 2.23, Centre for Multilevel Modelling (University of Bristol, 2011). Software authors: J. Rasbash, W. Browne, M. Healy, B. Cameron and C. Charlton.

[7] J. Rasbash and W. Browne, Non-Hierarchical Multilevel Models, in: J. de Leeuw and E. Meijer (eds.), Handbook of Multilevel Analysis (Springer, 2008). doi:10.1007/978-0-387-73186-5_8

[8] J. Rasbash, F. Steele, W. Browne and B. Prosser, A User's Guide to MLwiN: version 2.0. Centre for Multilevel Modelling (Institute of Education, University of London, London, 2004).

[9] S. Raudenbush and A. Bryk, Hierarchical Linear Models: applications and data analysis methods (2nd Edition, Sage Publications, Thousand Oaks, California, 2002).

[10] T. Richter, *What Is Wrong With ANOVA and Multiple Regression? Analyzing Sentence Reading Times With Hierarchical Linear Models*, Discourse Processes **41** (3) (2006) 221–250. doi:10.1207/s15326950dp4103_1

[11] T. Snijders and R. Bosker, Multilevel Analysis: An introduction to basic and advanced multilevel modeling (Sage Publications, London, 1999).

[12] L.M. Sullivan, K.A. Dukes and E. Losina, *Tutorial in Biostatistics: an introduction to hierarchical linear modelling*, Statistics in Medicine **18** (1999) 855–888.

[13] V. Valente and T.A. Oliveira, *Hierarchical Linear Models in Education Sciences: an Application*. Biometrical Letters **46** (1) (2009) 71–86.

[14] V. Valente and T.A. Oliveira, *Modelos Lineares Hierárquicos na Educação: Uma aplicação*. In: *Estatística Ciencia Interdisciplinar*, Actas do XIV Congresso Anual da SPE – Covilhã, Edições SPE (2007) 827–837.

[15] V. Valente and T.A. Oliveira, "Diferentes áreas científicas, diferentes notas? Uma aplicação do MLH a alunos do 10º ano", IIWEMC – Workshop de Estatística, Matemática e Computação, Universidade Aberta (Lisboa, 2006) (private communication).

Received 28 May 2011