- Faculté des sciences économiques
- www.unine.ch/seco

Mixed Models and Panel Data

Characteristics

- 3 ECTS credits
- Compulsory course for master in statistics
- Spring Semester
- Course : 2 hours
- Evaluation : written exam 2 hours
- Prerequisite : linear models

Teaching Team

- *Prof. Geert Molenberghs*
  Center for Statistics
  Hasselt University
  Agoralaan - building D
  3590 Diepenbeek
  Belgium
  ☎ +32-11-268238 - ✉ geert.molenberghs@uhasselt.be
  Web: http://www.uhasselt.be/censtat.
- *Prof. Geert Verbeke*
  Biostatistical Centre
  K.U.Leuven
  Kapucijnenvoer 35
  B-3000 Leuven
  Belgium
  ☎ *+32 16 33 68 91* ✉ geert.verbeke[at]med.kuleuven.be
  Web: http://www.kuleuven.be/biostat.

Objectives

As a result of the course, participants should be able to perform a basic analysis for a particular longitudinal data set at hand, using linear, generalized linear and non-linear tools for longitudinal data. Based on a selection of exploratory tools, the nature of the data, and the research questions to be answered in the analyses, they should be able to construct an appropriate statistical model, to fit the model within the SAS framework, and to interpret the obtained results. Further, participants should be aware not only of the possibilities and strengths of a particular selected approach, but also of its drawbacks in comparison to other methods. The course will be explanatory rather than mathematically rigorous. Emphasis is on giving sufficient detail in order for participants to have a general overview of frequently used approaches, with their advantages and disadvantages, while giving reference to other sources where more detailed information is available. Also, it will be explained in detail how the different approaches can be implemented in the SAS package, and how the resulting outputs should be interpreted.

Contents

We first present linear mixed models for continuous hierarchical data. The focus lies on the modeler's perspective and on applications. Emphasis will be on model formulation, parameter estimation, and hypothesis testing, as well as on the distinction between the random-effects (hierarchical) model and the implied marginal model. Apart from classical model building strategies, many of which have been implemented in standard statistical software, a number of flexible extensions and additional tools for model diagnosis will be indicated. Sec-

ond, models for non-Gaussian data will be discussed, with a strong emphasis on generalized estimating equations (GEE) and the generalized linear mixed model (GLMM). To usefully introduce this theme, a brief review of the classical generalized linear modeling framework will be presented.

Similarities and differences with the continuous case will be discussed. The differences between marginal models, such as GEE, and random-effects models, such as the GLMM, will be explained in detail. Third, it is oftentimes necessary to consider fully non-linear models for longitudinal data. We will discuss such situations, and place some emphasis on the non-linear mixed-effects model. Fourth, when analysing hierarchical and longitudinal data, one is often confronted with missing observations, i.e., scheduled measurements have not been made, due to a variety of (known or unknown) reasons. It will be shown that, if no appropriate measures are taken, missing data can cause seriously jeopardize results, and interpretational difficulties are bound to occur. Course Neuchâtel, Verbeke G. & Molenberghs G. 2 Methods to properly analyze incomplete data, under flexible assumptions, are presented. Key concepts of sensitivity analysis are introduced. Throughout the course, it will be assumed that the participants are familiar with basic statistical modelling, including linear models (regression and analysis of variance), as well as generalized linear models (logistic and Poisson regression). Moreover, pre-requisite knowledge should also include general estimation and testing theory (maximum likelihood, likelihood ratio). All developments will be illustrated with worked examples using the SAS System. However, the course is conceived such that it will be of benefit to both SAS users and users of other platforms.

Textbooks

Verbeke, G. and Molenberghs, G. (2000) Linear Mixed Models for Longitudinal Data. New York: Springer.
Molenberghs, G. and Verbeke, G. (2005) Models for Repeated Discrete Data. New York: Springer.