

Clustering for Multivariate Continuous and Discrete Longitudinal Data

ARNOŠT KOMÁREK AND LENKA KOMÁRKOVÁ

SUMMARY: Multiple outcomes, both continuous and discrete are routinely gathered on subjects in longitudinal studies. We propose a model-based statistical method for clustering (classification) of subjects into a prespecified number of groups with a priori unknown characteristics on basis of repeated measurements of all longitudinal outcomes. We start by modeling the evolution of each outcome using the classical generalized linear mixed model. Possible dependence between the values of different outcomes is captured by specifying a joint distribution of all random effects involved in the generalized linear mixed models for each outcome. The basis for subsequent clustering is provided by assuming a heteroscedastic mixture of multivariate normal distributions in the random effect distribution. Mainly for computational reasons we base the inference on a Bayesian specification of the model and simulation based Markov chain Monte Carlo methodology for which we prepared a ready to use software available in the form of R package. We discuss evaluation of uncertainty in the classification and also adapt recently proposed methodology for model comparison based on the posterior distribution of deviances to explore optimal number of clusters. The methodology is exemplified on the analysis of two real datasets (i) from a clinical trial on patients with primary biliary cirrhosis, (ii) from a semilongitudinal study on growth of contemporary Czech children.

KEY WORDS: Cluster analysis; Functional data; Generalized linear mixed model; Multivariate longitudinal data; Repeated observations.

A. KOMÁREK
FACULTY OF MATHEMATICS AND PHYSICS
CHARLES UNIVERSITY IN PRAGUE
SOKOLOVSKÁ 83
CZ-186 75, PRAHA 8
CZECH REPUBLIC
E-MAIL: arnost.komarek@mff.cuni.cz
URL: <http://www.karlin.mff.cuni.cz/~komarek>

L. KOMÁRKOVÁ
FACULTY OF MANAGEMENT
UNIVERSITY OF ECONOMICS IN PRAGUE
JAROŠOVSKÁ 1117
CZ-377 01, JINDŘICHŮV HRADEC
CZECH REPUBLIC
E-MAIL: komarkol@fm.vse.cz