Functional regression analysis using R

Christian Ritz

Department of Natural Sciences, Faculty of Life Sciences, University of Copenhagen, Thorvaldsensvej 40, DK-1871 Frederiksberg C, Denmark
`ritz@life.ku.dk`

Functional data consist of observations which can be treated as functions rather than just numeric vectors. One example is fluorescence curves commonly used in photosynthesis research: The curve reflects the biological processes taking place in a plant in the first second of exposure to sunlight. As several (mostly unexplained) processes are involved the resulting curves can have several local minima and maxima and is not easily described using parametric models. Another example is repeated measurements over time on the same subject, frequently encountered in dietary and growth studies. The resulting curves may fluctuate as a consequence of daily patterns or seasonal trends.

Any variety of models for describing functional data exists. Functional regression models are particularly appealing as they strike a balance between flexible non-parametric modelling of the unknown average curve and semi-parametric modelling of effects due to explanatory variables in much the same way as for ordinary ANOVA models.

This presentation shows how to use **R** for estimation, hypothesis testing and graphical model checking of functional regression models of the form:

$$y(t) = \phi(z)\mu(t) + \epsilon(t)$$

with $y(t), \mu(t)$ and $\epsilon(t)$ denoting the functional observation, the average curve and the error process, respectively. The term $\phi(z)$ is a multiplicative effect modifying the average curve according to the explanatory variable z. The functions μ and ϕ are estimated nonparametrically in a two-step procedure. A quasi-likelihood or glm approach can be used to estimate the effects of the explanatory variables.