Functional linear regression models with distributional response

R. Talská[1], A. Menafoglio[2], K. Hron[1], J. Machalová[1], and E. Fišerová[1]
[1]Department of Mathematical Analysis and Applications of Mathematics, Faculty of Science, Palacký University, Olomouc, Czech Republic; *talskarenata@seznam.cz*
[2]MOX - Department of Mathematics, Politecnico di Milano, Milano, Italy

Abstract

Distributional data result frequently from statistical surveys. When additional explanatory variables are available, they can also occur as the response of a functional regression model. In functional data analysis (FDA) methods to deal with regression models with functional response and real regressors have been widely discussed, but only for the somehow limited case of the space L^2 (Ramsay and Silverman, 2005). Problems arise when the response variable is formed by probability density functions (PDF) because the L^2 space, which is typically employed in FDA, does not account for their geometrical properties. In fact, PDFs represent a special case of functional data carrying relative information; respecting their relative nature – resulting in scale invariance and relative scale properties – is crucial for processing them statistically. Accordingly, PDFs should be represented in the Bayes space with the Hilbert space structure (Boogaart and others, 2014; Egozcue and others, 2006; Hron and others, 2016, Menafoglio and others, 2014). The aim of this contribution is to introduce functional linear regression models with distributional response, by grounding our developments precisely on the theory of Bayes spaces, that properly accounts for the properties of distributional data. In order to extend the widely-used approaches of FDA, especially those based on B-spline representation, we map PDFs from the Bayes space to the L^2 space, using the so called centred logratio transformation. We illustrate methodological developments with real data consisting of measurements of metabolite concentrations. The data set results from standard newborn screening done in the Laboratory of Inherited Metabolic Disorders, Department of Clinical Biochemistry, Faculty Hospital in Olomouc.

References

van den Boogaart, K.G., Egozcue, J.J., Pawlowsky-Glahn, V. (2014). Bayes Hilbert spaces. *Australian & New Zealand Journal of Statistics 56*(2), pp. 171–194.

Egozcue, J.J., Díaz-Barrero, J.L., Pawlowsky-Glahn, V. (2006). Hilbert space of probability density functions based on Aitchison geometry. *Acta Mathematica Sinica (English Series) 22*(4), pp. 1175—1182.

Hron, K., Menafoglio, A., Templ, M., Hrùzová, K., Filzmoser, P. (2016). Simplicial principal component analysis for density functions in Bayes spaces. *Computational Statistics and Data Analysis 94*, pp. 330–350.

Menafoglio, A., Guadagnini, A., Secchi, P. 2014. A kriging approach based on Aitchison geometry for the characterization of particle-size curves in heterogeneous aquifers. *Stoch. Environ. Res. Risk Assess 28*(7), pp. 1835–1851.