**On dimensionality and functional data analysis**

**Philippe Vieu**
Institut de Mathématiques de Toulouse, philippe.vieu@math.univ-toulouse.fr

The talk will be on statistical modelling for problems involving functional data. Functional data are elements of infinite dimensional spaces, and their statistical analysis has been popularized in the last twenty years (mainly with Ramsay-Silverman's books, [9]), leading to a new field of Statistics called Functional Data Analysis (FDA) which received high attention in the statistical community in the last twenty years (see eg [1], [3], [8] for a sample of recent bibliographically oriented works in FDA).

The first aim of the talk is to discuss how the infinite dimensionality of the problems is source of various difficulties relying to various areas of Mathematics:

i) Analytic issues: lack for reference measure as could be Lebesgue one in finite dimensional spaces (see eg [2]);

ii) Probabilistic issues: the lack of reference measure leads to the lack of density function (see eg [4], [5], ... );

iii) Statistical issues: the sparsity of data in high/infinite dimensional space may cause troubles when constructing flexible/non-parametric models ([6]).

The second attempt of the talk is to show, on the simple statistical situation of regression problems, how some dimensionality reduction models based on functional semi-parametric ideas can be succcesfully constructed to overpass the three points raised before. More precisely the talk will be oriented towards the so-called Single Functional Index Modelling (see [7] and references therein.)

The talk will be structured in two parts:

i) The first part, linked with the statements of the main issues of the problem, can be accessible to any mathematical audience not necessarily familiar with Statistics;

ii) The second part, linked with the construction of semiparametric models, will be more in direction of specialists in Statistics. This part will be mainly methodological, and the interest of the functional semiparametric approach will be highlighted by means of a few basic asymptotic results and by means of the treatment of some benchmark functional dataset coming from chemiometrics sciences.

# References

[1] Aneiros, G., Bongiorno, E., Cao, R. and Vieu, P. (2017). An introduction to the 4th edition of the International Workshop on Functional and Operatorial Statistics. In *Functional Statistics and Related Fields*, 1-6, Contributions to Statistics, Springer.

[2] Bogachev, V. (1998). *Gaussian measures.*Mathematical Surveys and Monographs, **62**. American Mathematical Society, Providence.

[3] Cuevas, A. (2014). A partial overview of the theory of statistics with functional data. *J. Statist. Plann. Inference,* **147**, 1-23.

[4] Delaigle, A. and Hall, P. (2010). Defining probability density for a distribution of random functions. *Ann. Statist.*, **38**(2), 1171-1193.

[5] Ferraty, F., Kudraszow, N. and Vieu, P. (2012). Nonparametric estimation of a surrogate density function in infinite-dimensional spaces. *J. Nonparametr. Stat.*, **24**(2), 447-464.

[6] Ferraty, F. and Vieu, P. (2006). *Nonparametric Functional Data Analysis. Theory and Practice.* Springer-Verlag, New York.

[7] Goia, A. and Vieu, P. (2015). A partitioned single functional index model. *Comput. Statist.*, **30**(3), 673-692.

[8] Goia, A. and Vieu, P. (2016). An introduction to recent advances in high/infinite dimensional statistics. *J. Multivariate Anal.*, **146**, 1-6.

[9] Ramsay, J. and Silverman, B. (2005). *Functional Data Analysis* (2nd edition). Springer Series in Statistics. Springer, New York.