PhD course
An Introduction to Functional Data Analysis: theory and practice

Valeria Vitelli, University of Oslo, Norway

March 25th-28th, 2024

Course Description

Functional Data Analysis (FDA) is a branch of statistics focused on the analysis and interpretation of data on a continuum, meaning that they can be described naturally as smooth curves or functions. Recent technological advances allow collection of data at such high-frequency that classical multivariate statistical techniques are not able to fully capture and describe their complexity, resulting in the flourishing of functional data methodology. FDA can easily be thought of as the collection of all statistical methodologies needed to handle samples of curves. These include mostly classical multivariate statistical methods adapted to the infinite dimensional case, such as regression models, inference, dimensional reduction, and clustering. However, methods that are specifically needed in this context have also been developed, such as data smoothing and registration. FDA methods have been applied quite broadly, from medicine to engineering, from business, to demography and social sciences, affecting nearly all fields of science. This course offers an introduction to FDA from a methodological viewpoint and gives practical guidance to R packages used in such context.

Course Objective

The overall objective of the course is that the PhD students in statistics at the DSEAS of the University of Palermo (Italy) develop skills in FDA, both at the theoretical (understanding the methodologies) and practical level (actual implementation in R). At the end of the course the students should be able to identify situations in which they can treat their data as functional, to apply standard FDA techniques (data description, dimensionality reduction, clustering, regression, and inference) and to visualize the results.
The course will last 4 days, comprising lectures connected to practical hands-on sessions.

Exam

Evaluation of the course learning outcome will be carried out via a take-home exam. The students will receive a dataset on Day 3 at lunch time, and they will be requested to analyse the data via a selected FDA technique among those introduced during the course. Evaluation will be carried out in a seminar, where the students will be requested to motivate the choice of analysis, describe the chosen methodology, visualize the data and results, and properly discuss them.

Course Program: schedule

Monday, March 25th	10:00 – 11:00	Lecture 0
	11:00 – 11:15	Break
	11:15 – 13:00	Lecture 1
	13:00 – 14:00	Lunch
	14:00 – 15:00	Practical Session 1
Tuesday, March 26th	10:00 – 11:30	Lecture 2 – part I
	11:30 – 11:45	Break
	11:45 – 13:00	Lecture 2 – part II
	13:00 – 14:00	Lunch
	14:00 – 15:00	Practical Session 2
Wednesday, March 27th	10:00 – 12:00 (with 10 min break)	Lecture 3
	12:00 – 13:00	Practical Session 3
	13:00 – 14:00	Lunch
	14:00 – ….	Start to work on the take-home exam + Question time
Thursday, March 28th	10:00 – 12:00 (with 10 min break)	Lecture 4
	12:00 – 13:00	Practical Session 4

Course Program: Lecture and Practical sessions plan

- Lecture 0: Overview of function spaces and operator theory
- Lecture 1: FDA fundamentals (descriptive statistics and outliers, smoothing functional data, functional PCA – fPCA)
- Practical session 1: Functional data smoothing and fPCA with R
- Lecture 2 – Part I: Unsupervised learning, functional clustering (distance-based clustering methods, hierarchical clustering)
  Lecture 2 – Part II: Unsupervised learning, data registration (landmark and continuous registration, relation to clustering)
- Practical session 2: Functional clustering and registration with R
- Lecture 3: Functional linear regression models (Scalar-on-Function – SoF, Function-on-Scalar – FoS, and Function-on-Function – FoF regression models)
- Practical session 3: Functional linear regression models in R
- Lecture 4: Inference for Functional data (functional ANOVA, testing and confidence intervals in SOF, FOS, and FOF regression models)
- Practical session 4: Inference for functional data in R, recap of the course