KEYNOTE - Maria Francesca Marino

On the flexibility of finite mixtures: from clustering to regression

Finite mixtures represent a flexible and effective tool of analysis when dealing with data characterized by complex dependence structures. This dependence is typically ascribed to sources of unobserved heterogeneity that, within a finite mixture framework, are fruitfully captured (and thus explained) by means of discrete latent variables. Conditional on them, a simple model for the outcomes of interest can be specified. This clearly offers several advantages, such as (i) ensuring interpretability, (ii) discovering hidden patterns in the data, (iii) providing a dimensionality reduction of potentially large datasets, (iv) accounting for the presence of heterogeneous sub-groups in the sample. Starting from their more basic definition, some recent developments of finite mixtures are presented and discussed. In particular, the focus is on finite mixtures for model-based clustering and finite mixtures of regression models for dependent observations, such as those recorded in longitudinal or multi-level designs. Estimation of model parameters, inference, and model selection are discussed. Some real data examples are shown with the aim of highlighting pros and cons of the methods at hand.