Séminaire de Biostatistique

Multilevel Functional Clustering Analysis

Nicoleta Serban
Georgia Institute of Technology, Atlanta, USA

Mardi 12 Juin 2012 à 16h00
Salle de séminaires, Tour 15-16, 2ème étage, 4 place Jussieu

Résumé

In this presentation, I will introduce clustering methods for multilevel functional data, which consist of repeated random functions observed for a large number of units (e.g. genes) at multiple sub-units (e.g. bacteria types). To describe the within- and between-variability induced by the hierarchical structure in the data, we take a multilevel functional principal components (MFPCA) approach. We develop and compare a hard clustering method applied to the scores derived from the MFPCA and a soft clustering method using an MFPCA decomposition. In a simulation study, we assess the estimation accuracy of the clustering membership and the cluster patterns under a series of settings: small vs. moderate number of time points; various noise levels and varying number of subunits per unit. We demonstrate the applicability of the clustering analysis to a real data set consisting of expression profiles from genes activated by immunity system cells. Prevalent response patterns are identified by clustering the expression profiles using our multilevel clustering analysis.