Clustering multivariate data using finite mixture models

Panagiotis Papastamoulis
Department of Statistics, AUEB

Abstract: Clustering data is a fundamental task in various scientific disciplines. Under a model-based clustering point of view, finite mixture models serve as an all-purpose workhorse for attacking this problem. Despite their flexibility, the estimation of such models is not trivial and computational methods (frequentist or Bayesian) should be incorporated. However, certain computational and inferential difficulties arise, especially in the case of multivariate datasets. This talk will review recent advances on these issues using finite mixtures for clustering various types (e.g.: count or continuous or categorical) of multivariate datasets with an unknown number of latent groups.