

PhD program in Statistics

DSS Statistics Seminar

May 30, 2025, 12:00

In person Room 34 (CU002)

Webinar <https://uniroma1.zoom.us/j/83625004899?pwd=bXCtz0mp759PUh2lkqT0BUoVa0Uegg.1>

Passcode: 123456

*Bi-clustering multivariate
categorical data via extended
mixtures of latent trait analyzers*

Maria Francesca Marino

UNIVERSITÀ DEGLI STUDI DI FIRENZE

Dipartimento di Statistica, Informatica, Applicazioni "G.Parenti" (DISIA)

Multivariate categorical outcomes are common in real-world applications but, frequently, their high dimensionality makes both the analysis and the interpretation particularly challenging. In this context, model-based clustering offers a powerful approach to data reduction and structure discovery. We build upon the Mixture of Latent Trait Analyzers (MLTA) framework to propose a model that enables simultaneous clustering of both rows and columns of the data matrix. Specifically, rows are grouped into components using a finite mixture specification. Within each component, variables are segmented based on a flexible yet parsimonious specification of the linear predictor. To capture residual dependence among observations, we retain the use of a multidimensional latent trait, consistent with the original MLTA formulation. Additionally, the model accommodates the influence of individual-specific covariates on the clustering process via a concomitant variable framework.

Parameter estimation is carried out using maximum likelihood, implemented through an extended Expectation-Maximization (EM) algorithm. Since the likelihood involves integrals without closed-form solutions, we apply a Gauss-Hermite quadrature for numerical approximation. A comprehensive simulation study evaluates the model's ability to accurately recover both clustering structure and parameter values, demonstrating strong performance. Finally, we apply the proposed method to an original dataset on pediatric patients with suspected appendicitis, aiming to identify patient subgroups characterized by distinct patterns of clinical conditions.



SAPIENZA
UNIVERSITÀ DI ROMA