

Statistics and Actuarial Science Seminar

Title: Multivariate random partition learning and prediction

Speaker: Beatrice Franzolini (King's College London)

Date: Thu 26th November 2026 at 3:00PM

Location: N0.20 - Science North

Abstract: A number of models (including mixture models, stochastic block models, and mixtures of experts) capture data complexity by positing a latent partition of the observed data points. However, standard formulations can be restrictive for modelling repeated measurements on the same subjects. Examples include multivariate panels, dynamic networks, and multi-view data where cluster membership may differ across views and evolve over time. Patients who group together on one set of biomarkers may split apart on another, and nodes in a network may move between communities through time. In this talk, I will address the following question: how can a generative model capture these differences, learn what different views or time points have in common, and use evolving clustering processes to predict future observations?

Treating the partition as random within a Bayesian framework helps control overfitting while preserving flexibility. The symmetries of its law govern how learning is structured and which information is used for inference and prediction. Well-established exchangeable random partitions, characterised through Kingman's representation theorem, induce a single univariate random partition (even when exchangeability is relaxed to partial exchangeability), typically resulting in loss of information when they are applied to repeated measurements and multi-view data.

The core of the talk focuses on how to move from univariate to multivariate random partitions and on their probabilistic symmetries and statistical properties.