DEVIL: Efficient and statistically principled differential expression analysis at scale

A seminar by Giovanni Santacatterina

Università di Trieste

Monday 16 Mar 2026 | 14:30-15:30

Room SC30

Department of Statistical Sciences

Standard differential expression analyses assume independent samples, an assumption violated in multi-patient single-cell RNA sequencing, where thousands of correlated cells originate from the same individual. Existing strategies either collapse data to the patient level, losing cellular resolution, or introduce subject-specific parameters that compromise scalability and stability in complex designs.

We introduce DEVIL, a Gamma–Poisson regression framework that restores valid population-level inference by combining efficient estimations with clustered sandwich covariance correction. This approach accounts for within-patient dependence without increasing model dimensionality. By correctly estimating variance and leveraging GPU acceleration, DEVIL scales to tens of millions of cells while maintaining calibrated false discovery control in different experimental settings. Across simulations and real datasets, we show that DEVIL is able to reconcile single-cell resolution with statistically robust population inference at modern data scale.

Università
degli Studi
di Padova

Dipartimento
di Scienze
Statistiche