HERITABILITY ESTIMATION IN HIGH-DIMENSIONAL SPARSE LINEAR MIXED MODELS

Anna Bonnet
Joint work with Elisabeth Gassiat and Céline Lévy-Leduc

The heritability of a biological quantitative feature is defined as the proportion of its variation that can be explained by genetic factors. We propose an estimator for heritability in high dimensional sparse linear mixed models and we study its theoretical properties. We highlight the fact that in the case where the size $N$ of the random effects is too large compared to the number $n$ of observations, a precise estimation for heritability cannot be provided. Since in practice almost all datasets verify the condition $N \gg n$, we perform a variable selection method to reduce the size of the random effects and to improve the accuracy of heritability estimations. However, as shown by our simulations, this kind of approach only works when the number of non-zero components in the random effects (i.e. the genetic variants which have an impact on the phenotypic variations) is small enough. In face of this limitation, we proceeded to define an empirical criterion to determine whether it is possible or not to apply a variable selection approach. As an example of its use, we applied our method to estimate the heritability of the volume of several regions of the human brain.