



Seminar/Talk

Bayesian hierarchical models for large-scale discovery, estimation and prediction analysis of common complex disease

Matthew Robinson

University of Lausanne

Host: Nick Barton

A major challenge across the sciences is to conduct statistical testing for associations within large datasets, where the number of covariates strongly exceeds the number of observations. My group derives novel statistical models and develops computational algorithms, to conduct variable selection and appropriate effect estimation within large-scale data, all with the aim of addressing long-standing questions in the life sciences. This is particularly important for understanding common complex human diseases, such as type-2 diabetes, obesity, and cardiovascular disease, where typically there are many thousands of genetic and environmental risk factors, each of small effect. Currently our ability to characterise these risk factors is limiting our ability to respond optimally, treat and ultimately prevent common disease. In this talk, I will present a flexible hierarchical Bayesian model we have developed that utilises structured graph parallel programming and residual updating, enabling all effects to be estimated conditionally in large-scale data of any type, with low compute resource requirements. Three general extensions of this hierarchical Bayesian model are then presented, each of which is applied to address a long-standing biological problem. First, we allow multiple levels of hierarchy, which enables joint estimation of associations between obesity risk and genetic and epigenetic variation. We show that this facilitates accurate characterization of disease progression and identification of individuals who are on a path to disease, as it leads to the identification of biomarkers of large effect. Second, we further develop this model to provide a full quantification of the different genomic components and molecular mechanisms that underlie common disease risk, which also provides improved disease prediction. Third, we present an approach to model time-to-event data with a Weibull distribution that handles sparsity with spike and slab variable selection, considers left truncation and right censoring, and utilizes adaptive rejection sampling. This enables more accurate discovery and estimation of survival related genomic marker effects and grants novel insight into the genetic architecture of time-to-disease diagnosis. Finally, I will then outline my future interdisciplinary research goals to facilitate a range of large-scale analyses across the sciences, as well as better understand early-life genetic effects.

Thursday, January 31, 2019 10:00am - 11:00am

Mondi Seminar Room 3, Central Building



This invitation is valid as a ticket for the ISTA Shuttle from and to Heiligenstadt Station.
Please find a schedule of the ISTA Shuttle on our webpage:
<https://ista.ac.at/en/campus/how-to-get-here/> The ISTA Shuttle bus is marked ISTA Shuttle (#142) and has the Institute Logo printed on the side.

www.ista.ac.at | Institute of Science and Technology Austria | Am Campus 1 | 3400 Klosterneuburg