Approximate Bayesian Computing (ABC) for model choice: from statistical sufficiency to machine learning

Christian P. Robert

Université Paris-Dauphine & University of Warwick

Abstract

Since its introduction in the late 1990's, the performances of the ABC method have been analysed from several perspectives, starting with the pure practical motivations of the population geneticists who created it to an approximate Bayesian method, to a non-parametric one. We cover in this talk a new vision the specific case of model selection, showing how we originally developed convergent methods for Gibbs random fields, before moving to a pessimistic view of the consistency of the method and producing necessary and sufficient conditions for this consistency to hold, then to the realisation that generic machine learning tools like KNNs and random forests should be put to use to run model selection in the complex models covered by ABC techniques. Our perspective radically alters the way model selection is operated as we ban approximations of posterior probabilities for the models under comparison, since they cannot be reliably estimated, and propose instead to compute the performances of the selection method. As an aside, we argue that both knn and random forest methods can be adapted to the settings of interest, with a recommendation on the automated selection on the tolerance level and sparse implementations of the random forest tree construction, using subsampling and reduced reference tables. This talk is based on joint works with Jean-Marie Cornuet, Arnaud Estoup, Jean-Michel Marin, Natesh Pillai, Pierre Pudlo and Judith Rousseau.