Introduction to "Adaptive ABC model choice and geometric summary statistics for hidden Gibbs random fields" by J. Stoehr, P. Pudlo, L. Cucala

Christian P. Robert

Approximate Bayesian computation techniques are 2000's successors of MCMC methods as handling new models where MCMC algorithms are at a loss, in the same way the latter were able in the 1990's to cover models that regular Monte Carlo approaches could not reach. While they first sounded like "quick-and-dirty" solutions, only to be considered until more elaborate solutions could (not) be found, they have been progressively incorporated within the statistician's toolbox as a novel form of non-parametric inference handling partly defined models. A statistically relevant feature of those ACB methods is that they require replacing the data with smaller dimension summaries or statistics, because of the complexity of the former. In almost every case when calling ABC is the unique solution, those summaries are not sufficient and the method thus implies a loss of statistical information, at least at a formal level since relying on the raw data is out of question. This forced reduction of statistical information raises many relevant questions, from the choice of summary statistics to the consistency of the ensuing inference.

In this paper, Stoehr et al. attack the recurrent problem of selecting summary statistics for ABC in a hidden Markov random field, since there is no fixed dimension sufficient statistics in that case. The paper provides a very broad overview of the issues and difficulties related with ABC model choice, which has been the focus of some advanced research only for a few years. Most interestingly, the authors define a novel, local, and somewhat Bayesian misclassification rate, an error that is conditional on the observed value and derived from the ABC reference table. It is the conditional expected misclassification error

$$\mathrm{P}^{\mathrm{ABC}}\left(\widehat{m}\left(y^{\mathrm{obs}}\right) \neq m \,|\, \mathrm{S}(y^{\mathrm{obs}})\right)$$

integrating in both the model index m and the corresponding random variable Y (and the hidden intermediary parameter) given the observation. Or rather given the transform of the observation by the summary statistic S. The authors even go further to define the error rate of a classification rule based on a first (collection of) statistic, conditional on a second (collection of) statistic (see Definition 1). A notion rather delicate to validate on a fully Bayesian basis. And they advocate the substitution of the unreliable (estimates of the) posterior probabilities by this local error rate, estimated by traditional non-parametric kernel methods. Methods that are calibrated by cross-validation. Given a reference summary statistic, this perspective leads (at least in theory) to select the optimal summary statistic as the one leading to the minimal local error rate. Besides its application to hidden Markov random fields, which is of interest per se, this paper thus opens a new vista on calibrating ABC methods and evaluating their true performances conditional on the actual data. (The advocated abandonment of the posterior probabilities could almost justify the classification of this paper as constituting a *paradigm shift*).

C. P. Robert (✉)
Université Paris-Dauphine, Paris, France
e-mail: xian@ceremade.dauphine.fr

C. P. Robert
University of Warwick, Coventry, UK