

# Approximate Bayesian computation methods

Gilles Celeux

Published online: 5 September 2012  
© Springer Science+Business Media, LLC 2012

Occasionally, *Statistics and Computing* is publishing Special Issues on topics of potential interests. The most recent published Special Issues were concerned with “*Adaptive Methods in Bayesian Computation*”, Guest Editor Paul Fearnhead, Volume 18 Issue 4 (2008), “*Regularisation Methods in Classification and Regression*”, Guest Editor Gerhard Tutz, Volume 20 Issue 2 (2009), and “*Modeling of Computer Experiments for Uncertainty Propagation and Sensitivity Analysis*”, Guest Editors Anestis Antoniadis and Alberto Pisanisi, Volume 22 Issue 3 (2012).

Usually, those special issues are started with a “*Call for papers*” giving their purpose and specifying their desired subjects. This issue is proposing a Special Issue on Approximate Bayesian Computation (ABC) methods, which has not been conceived according to this scheme. Actually, ABC methodology is an emerging domain of computational statistics and there was no need to prepare a formal “*Call for papers*” for this special issue. The project of this special issue simply rises because, since early 2010, *Statistics and Computing* has received a lot of submissions on this new topic of computational statistics. And ABC methodology is typically a good material for *Statistics and Computing*.

Roughly speaking each ten years, a new methodology of computational statistics appears and dominates the scene for a while, it was the bootstrap, Efron (1979), and the algorithms related to the EM algorithm of Dempster et al. (1977) in the eighties, Monte Carlo Markov Chains for Bayesian analysis in the nineties (Gelfand and Smith 1990), regularization methods derived from the Lasso (Tibshirani 1996)

in the noughties, and now it is ABC methodology which is dominating.

ABC methodology is remarkable by the fact that it has been introduced by researchers involved in population genetics problems rather than by statisticians (Pritchard et al. 1999; Beaumont et al. 2002). Obviously as illustrated in the present issue, ABC received an increasing interest in the statistical literature and became an important Monte Carlo method. Typically, ABC is aiming to solve a computational challenge and can be regarded as a general methodology to derive Bayesian statistical estimators for high dimensional models where the likelihood is not practically available.

This special issue on ABC methods proposes eight articles. The first article has been demanded to my friends Jean-Michel Marin and Christian P. Robert who have rapidly become distinguished authors on ABC methodology. The introductory article of Marin, Pudlo, Ryden and Robert can be regarded as a general tutorial on ABC methods with historical notes. This survey gives a clear exposition of the problems encountered when implementing this methodology and a detailed account of the more recent improvements and extensions the basic ABC algorithm has received. In particular, this article includes an interesting and smart section on ABC and model selection.

In their article, C. Barnes, S. Filippi, M. Stumpf and T. Thorpe attacks the important problem of choosing a pseudo-sufficient summary statistics to replace the actual data. Considering summary statistics as data-compression mechanisms, these statistics are constructed in an information-theoretical framework by combining different summary statistics until the loss of information is minimized using one of the algorithms proposed by the authors. Moreover, an automated selection of summary statistics is proposed for model selection.

The next two articles give methodological contributions.

---

G. Celeux (✉)  
Département de Mathématiques, Inria Saclay-Île-de-France,  
Université Paris-Sud, 91405 Orsay, France  
e-mail: [Gilles.Celeux@inria.fr](mailto:Gilles.Celeux@inria.fr)

The article of R. McVinish is proposing a specific ABC algorithm when the ABC methodology is needed because the statistical model involves quantile distributions. This new algorithm includes an efficient Metropolis-Hasting step in this particular context.

G. Peters, Y. Fan and S. Sisson propose a sequential Monte Carlo sampler motivated by applications to the free-likelihood ABC context. A Bayesian stochastic model in insurance for analyzing claims illustrates the efficiency of the sequential ABC algorithm.

A. Jasra, S. Singh, J. Martin and E. McCoy propose an ABC approximation to perform biased filtering for a hidden Markov model when the likelihood function is intractable. This article has an intermediate position between a methodological contribution and a more application-oriented contribution. The new target distribution they propose is removing the need to evaluate the likelihood. Notice that as Peters et al. they use a sequential Monte Carlo algorithm to sample from their ABC distribution. This article includes an illustration for online portfolio optimization.

The next articles are essentially of applied nature and are primarily concerned with real applications of ABC methods.

The article of P. Neal considers ABC methodology to deal with epidemic data. It is proposed to generate a set of values rather than a single value at each simulation by coupling different sets of parameters. An application to analyze final size data for epidemics amongst communities partitioned into households is presented. Moreover, a discussion section analyzed the pro and the cons of different ABC algorithms from the practical point of view.

The article of A. Rau, F. Jaffrezic, J.-L. Foulley and R. Doerge is devoted to an application of ABC methodology to the reverse analysis of a gene network from time-course gene expression data. A non-standard extension of the standard ABC-MCMC algorithm is proposed to enable inference of gene regulatory networks. This is a typical applica-

tion of the ABC methodology since the concerned networks are complex and high-dimensional networks for which the likelihood is prohibitive to calculate.

The article of D. Nott, L. Marshall and M.N. Tran is of more theoretical nature and its title is speaking by itself: “*The ensemble Kalman filter is an ABC algorithm*”.

As it is, this special issue on ABC methods is dealing with much of the aspects of this promising free-likelihood methodology for complex and high dimensional data and I hope it will interest the *Statistics and Computing* readers.

**Acknowledgements** I gratefully thank all the Associate Editors and reviewers who participated to this special issue for their valuable work with a special thank to Ajay Jasra (Associate Editor) who much contributed to this Special Issue. I also gratefully thank the authors for their much interesting contributions with a special thank to Jean-Michel Marin and Christian P. Robert who kindly accepted to prepare a first class survey paper on ABC methodology which leads to a great introduction to this Special Issue. I am also glad to thank Melissa Fearon, Jacqueline Lalic and Malgosia Slowinska for their efficient help in preparing this special issue.

## References

- Beaumont, M., Zhang, W., Balding, D.: Approximate Bayesian computation in population genetics. *Genetics* **162**(4), 2025–2035 (2002)
- Dempster, A.P., Laird, N.M., Rubin, D.B.: Maximum likelihood from incomplete data via the EM algorithm. *J. R. Stat. Soc. B* **39**(1), 1–38 (1977)
- Efron, B.: Bootstrap methods: another look at the jackknife. *Ann. Stat.* **7**(1), 1–26 (1979)
- Gelfand, A., Smith, A.F.M.: Sampling-based approaches to calculating marginal densities. *J. Am. Stat. Assoc.* **85**(410), 398–409 (1990)
- Pritchard, J., Seielstad, M., Perez-Lezaun, A., Feldman, M.: Population growth of human Y chromosomes: a study of Y chromosome microsatellites. *Mol. Biol. Evol.* **16**(12), 1791–1798 (1999)
- Tibshirani, R.: Regression shrinkage and selection via the lasso. *J. R. Stat. Soc. B* **58**(1), 267–288 (1996)