Journal de la Société Française de Statistique
Vol. 160 No. 1 (2019)

Editorial of the Special Issue on Mixture Analysis

Titre: Éditorial du numéro spécial sur l'analyse de mélanges

Gilles Celeux[1]

Since many years, mixture models have taken a great importance in modern statistical inference. Mixture models are hidden structure models allowing to analyse heterogeneity of data often occurring in a lot of domains. Thus mixture models can be viewed as a semi-parametric tool to estimate probability distributions and are also the reference tool in model-based clustering. They apply in many contexts such as clustering of independent data, hidden Markov models, clustering of networks, mixtures of experts models, co-clustering... They involve important, difficult and interesting questions to be addressed. Choosing a relevant number of mixture components is not the least one.

As a mater of fact, mixture models lead to authoritative monographs along the years. The book of [Titterington et al., 1985] can be recommended in particular for its presentation of the basic mathematical characteristics of mixture models. The book of [McLachlan and Peel, 2000] can be recommended in particular for its presentation of model-based clustering. The book of [Frühwirth-Schnatter, 2006] can be recommended in particular for its focus on Bayesian inference. The editorial activity about mixture models is still important. We can cite the Handbook of Mixture Analysis by [Frühwirth-Schnatter et al., 2018] which presents a large overview of the methods and applications of mixture models. We can also mention the forthcoming monography on model-based clustering [Bouveyron et al., 2019].

The present special issue on mixture models deals with multiple aspects of this important field of research: clustering of dynamic networks, mixture of experts models, testing the number of componets of univariate Gaussian mixtures, bootstrap procedures to assess the number of mixture components, and an application in genomics. It is worthwhile to remark that all the articles of this special issue are concerned with model selection and the crucial problem to choose a sensible and proper number of mixture components.

Ricardo Rastelli considers the direct maximisation of the exact integrated completed likelihood in a non informative setting for the Stochastic Block Model with dynamic networks.

Faicel Chamroukhi and Bao-Tuyen Hyung propose regularised estimation and variable selection procedures for mixture of experts models to get parsimonious solutions.

Didier Chauveau, Bernard Garel and Sophie Mercier go back to the basic problem of testing the number of components in univariate Gaussian mixtures from a practical point of view.

Zhivko Taushanov and André Berchtold propose different bootstrap procedures to select a mixture model. They concentrate their illustrations on the Hidden Mixture Transition Distribution model for the clustering of longitudinal data.

[1] Inria Saclay-Île-de-France, Orsay, F-91405, France

 ISSN: 2102-6238

Christine Keribin, Yi Lia, Tatiana Popova and Yves Rozenholc present an application of mixture models on genomic data. They pay special attention to the selection of a relevant model and illustrate the possible interest of the slope heuristics in the context they consider.

I thank all the authors for their contributions and I hope that the readers of the Journal of the SFdS will be interested by this special issue.

References

[Bouveyron et al., 2019] Bouveyron, C., Celeux, G., Murphy, T. B., and Raftery, A. E. (2019). *Model-based Clustering and Classification for Data Science*. Cambridge University Press. (Forthcoming).

[Frühwirth-Schnatter, 2006] Frühwirth-Schnatter, S. (2006). *Finite Mixture and Markov Switching Models*. Springer-Verlag.

[Frühwirth-Schnatter et al., 2018] Frühwirth-Schnatter, S., Celeux, G., and Robert, C. P. (2018). *Handbook of Mixture Analysis*. CRC Press.

[McLachlan and Peel, 2000] McLachlan, G. J. and Peel, D. (2000). *Finite Mixture Models*. John Wiley.

[Titterington et al., 1985] Titterington, D. M., Smith, A. F. M., and Makov, U. E. (1985). *Statistical Analysis of Finite Mixture distributions*. John Wiley.