



XXIX COMCA

Congreso de Matemática Capricornio

July 28th, 29th and 30th, 2021, Copiapó, Chile

The semi-hierarchical Dirichlet Process and its application to clustering homogeneous distributions

Fernando Andrés Quintana*

Departamento de Estadística, Facultad de Matemáticas
Pontificia Universidad Católica de Chile
Santiago, Chile

Abstract

Assessing homogeneity of distributions is an old problem that has received considerable attention, especially in the nonparametric Bayesian literature. To this effect, we propose the semi-hierarchical Dirichlet process, a novel hierarchical prior that extends the hierarchical Dirichlet process of Teh et al. (2006) and that avoids the degeneracy issues of nested processes recently described by Camerlenghi et al. (2019). We go beyond the simple yes/no answer to the homogeneity question and embed the proposed prior in a random partition model; this procedure allows us to give a more comprehensive response to the above question and in fact find groups of populations that are internally homogeneous when $I \geq 2$ such populations are considered. We study theoretical properties of the semi-hierarchical Dirichlet process and of the Bayes factor for the homogeneity test when $I = 2$. Extensive simulation studies and applications to educational data are also discussed.

Joint work with:

Mario Beraha¹, Department of Mathematics, Politecnico di Milano, Milan, Italy.

Alessandra Guglielmi², Department of Mathematics, Politecnico di Milano, Milan, Italy.

References

- [1] Camerlenghi, F., Dunson, D. B., Lijoi, A., Prünster, I., and Rodríguez, A. (2019). Latent Nested Nonparametric Priors (with Discussion). *Bayesian Analysis*, 14(4): 1303–1356.
- [2] Teh, Y. W., Jordan, M. I., Beal, M. J., and Blei, D. M. (2006). Hierarchical Dirichlet Processes. *Journal of the American Statistical Association*, 101(476): 1566–1581.

*Partially supported by gran Fondecyt 1180034, and ANID - Millennium Science Initiative Program - NCN17_059, e-mail: quintana@mat.uc.cl

¹e-mail: mario.beraha@polimi.it

²e-mail: alessandra.guglielmi@polimi.it