The Pyrenees International Workshop on Statistics, Probability and Operations Research, SPO 2011 Jaca, September 13–16th 2011

CID sequences and Bayesian non parametrics

Fabrizio Leisen[1],

SUMMARY

Conditional identity in distribution ([1]) is a new type of dependence for random variables, which generalizes the well-known notion of exchangeability. In Bassetti, Crimaldi, Leisen ([2]) a class of random sequences, called Generalized Species Sampling Sequences, is defined and a condition to have conditional identity in distribution is given. In particular, a class of generalized species sampling sequences that are conditionally identically distributed is introduced and studied: the Generalized Ottawa sequences (GOS). Species sampling sequences is an important tool in Bayesian non parametrics. Indeed, many popular Bayesian Nonparametric priors can be characterized in terms of exchangeable species sampling sequences. One example is the Dirichlet Process prior, that has been increasingly used for modeling purposes in mixture of DP hierarchical models. However, in some applications, the implied exchangeability assumption may not be considered appropriate. In the GOS family it is available a species sampling sequence, characterized by a tractable predictive probability function, and with weights driven by a sequence of independent Beta random variables. We discuss some of the properties that can be useful in applications, and we compare our findings with wellknown properties of the DP and the two parameters Poisson-Dirichlet process. We detail on Markov Chain Monte Carlo posterior sampling, and illustrate the behavior of such priors.

Keywords: Conditionally Identically Distributed Sequences, Species Sampling, Bayesian Non Parametrics, Generalized Ottawa Sequences

AMS Classification: Primary 62C10; Secondary 62G57

References

[1] Berti P., Pratelli L. and Rigo P. (2004). Limit Theorems for a Class of Identically Distributed Random Variables. *Annals of Probability* **32**, 2029–2052.

[2] Bassetti F., Crimaldi I. and Leisen F. (2010). Conditionally identically distributed species sampling sequences. *Advances in applied probability* **42**, 433–459.

[1]Universidad Carlos III de Madrid
fabrizio.leisenl@uc3m.es