Spatial neutral-to-the-left stochastic processes

Vittorio Romaniello

March 28, 2020

1   Summary

A sequence of random variables $X = (X_1, \ldots, X_k) \in \mathbb{R}_+^k$ is said to be neutral-to-the-left (NTL) if the increments

$$R_j = \frac{X_j}{\sum_{i=1}^{j} X_i}$$

form a sequence of mutually independent random variables in $[0,1]$ [Blo+18]. Such processes are common in statistics and can be used to describe, for instance, the generating process of left censored data or the asymptotic degree distribution in preferential attachment random graphs [BO17]. In the statistics and probability literature and, especially, in Bayesian nonparametrics, NTL processes have not received as much attention as their analogous counterpart, neutral-to-the-right (NTR) processes [1]. The disparity in the development of theory for NTL and NTR processes can be attributed to two factors:

1. NTR processes are more common than NTL processes, therefore more focus has been put on the development of theory for the former.

2. Since NTL and NTR are symmetric opposites [Blo+18], an NTL process can be transformed into an NTR for the analysis and back to an NTL after the analysis.

However, the increased attention received by the study of random graphs as well as the increased has made NTL processes more popular and highlighted the need to develop a theory for NTL processes. In fact, while the approach used so far in the study of NTL processes is valid for situations where

Neutral-to-the-left (NTL) stochastic processes are a class of stochastic processes used to

Neutral-to-the-right (NTR) stochastic processes [Dok74] were introduced as priors for Bayesian nonparametrics models of right censored data (e.g survival analysis). In Bayesian statistics, using priors specifically designed for the problem at hand is renowned to be beneficial for effective inference. Hence, having the possibility to improve the descriptiveness of models for right censored data using NTR priors should be seen as a great advantage. However, [Jam06] noticed that, despite the numerous applications suitable for NTR priors, these processes were not being used as extensively as they could have been. The Dirichlet process constituted the only exception and it is still ubiquitously used in Bayesian nonparametrics applications. [Jam06] identified several reasons for the absence of NTR priors from the Bayesian nonparametrics literature and addressed them in his work.

Firstly, he noticed that the NTR process of [Dok74] was defined only on the real line, $\mathbb{R}$, therefore restricting its application to processes in the real space. Hence, NTR processes could not be applied to situations where inference on more complex spaces was required. To address this limitation, [Jam06] proposed the spatial neutral-to-the-right process (SPNTR), an extension of NTR processes to arbitrary Polish spaces.

Secondly, and probably the most important problem addressed in [Jam06], is that NTR processes are, in general, not tractable and there is no simple way to sample from them. On the other hand, Dirichlet processes can easily be sampled through the Chinese restaurant process. Therefore, [Jam06] derives a similar sampling scheme for NTR processes, making their implementation more approachable.

Lastly, to address the issues above, [Jam06] used results from Poisson partition calculus [Jam05]. Although the results from Poisson calculus come from a previous work [Jam05], the approach greatly simplified proofs and allows for a relatively simple generalization of NTR processes. Approaches used in literature prior to this work derived sampling distributions for Dirichlet processes using combinatorial arguments [Ant74; Pit+02]. Adopting a similar strategy in generalizing NTR processes would result challenging and would likely not be viable.

---

[1] Reversing an NTL process gives an NTR process.

2   Mini-proposals

2.1   Proposal 1: MY PROPOSAL TITLE

2.2   Proposal 2: MY OTHER PROPOSAL TITLE

3   Project report

References

[Ant74]   C. E. Antoniak. "Mixtures of Dirichlet processes with applications to Bayesian nonparametric problems". In: *The annals of statistics* (1974), pp. 1152–1174.

[BO17]   B. Bloem-Reddy and P. Orbanz. "Preferential attachment and vertex arrival times". In: *arXiv preprint arXiv:1710.02159* (2017).

[Blo+18]   B. Bloem-Reddy et al. "Sampling and inference for Beta Neutral-to-the-Left models of sparse networks". In: *arXiv preprint arXiv:1807.03113* (2018).

[Dok74]   K. Doksum. "Tailfree and neutral random probabilities and their posterior distributions". In: *The Annals of Probability* (1974), pp. 183–201.

[Jam05]   L. F. James. "Poisson process partition calculus with an application to Bayesian Lévy moving averages". In: (2005).

[Jam06]   L. F. James. "Poisson calculus for spatial neutral to the right processes". In: *The annals of Statistics* 34.1 (2006), pp. 416–440.

[Pit+02]   J. Pitman et al. *Combinatorial stochastic processes*. Tech. rep. Technical Report 621, Dept. Statistics, UC Berkeley, 2002. Lecture notes for . . ., 2002.