Ramsés Mena: Exchangeable random measures and stick-breaking priors

Discussant: Sinead Williamson

Department of Statistics and Data Science
The University of Texas at Austin

O'Bayes 2022
September 9, 2022

Exchangeable random measures and stick-breaking priors

Thank you Ramsés for a great talk!!

Exchangeable random measures and stick-breaking priors

Stick breaking priors

$$w_i = v_i \prod_{j<i} (1 - v_j) \qquad\qquad v_i \sim \nu_i$$

- Dirichlet process: $\nu_i := \nu = \text{Beta}(1, \theta)$
- Pitman-Yor process: $\nu_i = \text{Beta}(1 - d, \theta + id)$
- N-IG process: ν_i depends on current stick length
- NRMI: ν_i depends on $\nu_{<i}$
- σ-stable Poisson-Kingman process: ν_i has parametric form depending on all $v_{<i}$
- Geometric process: $\nu_i = \lambda$

Exchangeable random measures and stick-breaking priors

Stick breaking priors *with iid length variables*

$$w_i = v_i \prod_{j<i}(1 - v_j) \qquad\qquad v_i \sim \nu$$

- Dirichlet process: $\nu_i := \nu = \mathrm{Beta}(1, \theta)$
- ~~Pitman-Yor process: $\nu_i = \mathrm{Beta}(1 - d, \theta + id)$~~
- ~~N-IG process: ν_i depends on current stick length~~
- ~~NRMI: ν_i depends on $\nu_{<i}$~~
- Geometric process: $\nu_i = \lambda$

Exchangeable random measures and stick-breaking priors

Stick breaking priors *with exchangeable length variables*

- Dirichlet and geometric trivially exchangeable...
- Non-trivially exchangeable if we put a prior on θ/λ

Exchangeable random measures and stick-breaking priors

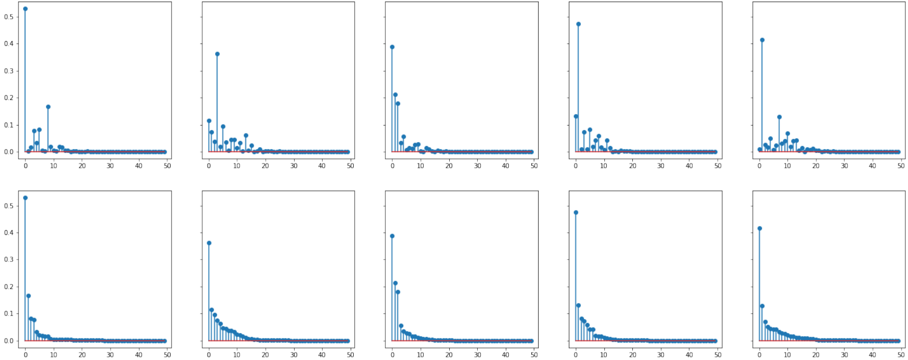
Stick breaking priors *with exchangeable length variables*

- Dirichlet and geometric trivially exchangeable...
- Non-trivially exchangeable if we put a prior on θ/λ
- Alternative: $\nu = \sum_{i=1}^{\infty} \pi_i \delta_{\phi_i}$, where all $\phi_i \in (0, 1)$
  - Can't really specify the iid version... but *can* specify the exchangeable version using species-sampling priors

$$w_i = v_i \prod_{j<i} (1 - v_j) \qquad\qquad v_1, \nu_2, \cdots \sim \mathrm{SSS}$$

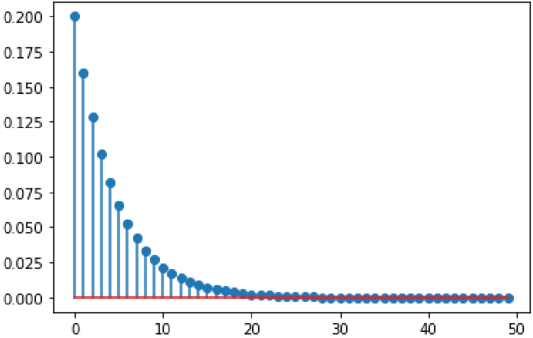
Some examples (all with fixed ν, $E_\nu[v] = 0.2$)

- GEM(4) (ie $\nu = \beta(1, 4)$)



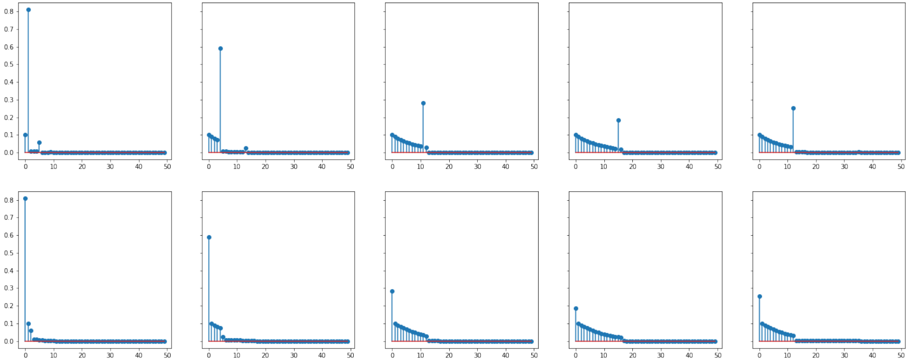
Some examples (all with fixed ν, $E_\nu[v] = 0.2$)

- Geometric(0.2) (ie $\nu = \delta_{0.2}$)



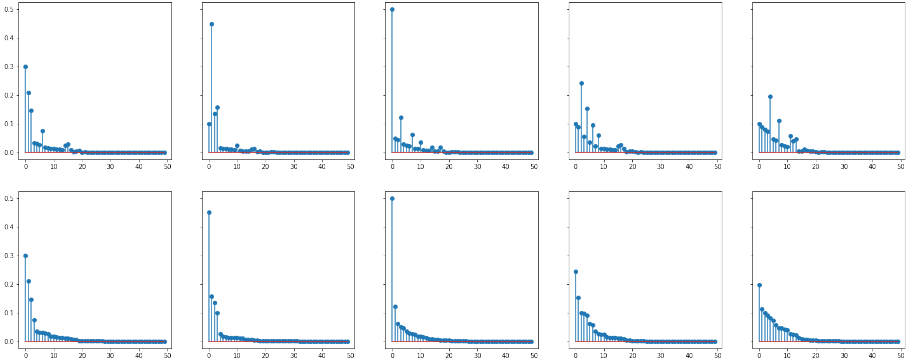
Some examples (all with fixed ν, $E_\nu[v] = 0.2$)

- $\nu = 0.875\delta_{0.1} + 0.125\delta_{0.9}$



Some examples (all with fixed ν, $E_\nu[v] = 0.2$)

- $\nu = 0.6\delta_{0.1} + 0.3\delta_{0.3} + 0.1\delta_{0.5}$



Beyond exchangeability

Stick breaking priors *with Markovian length variables*

- Many NMRIs have dependency in thier stick-breaking process: v_i depends on $v_{<i}$

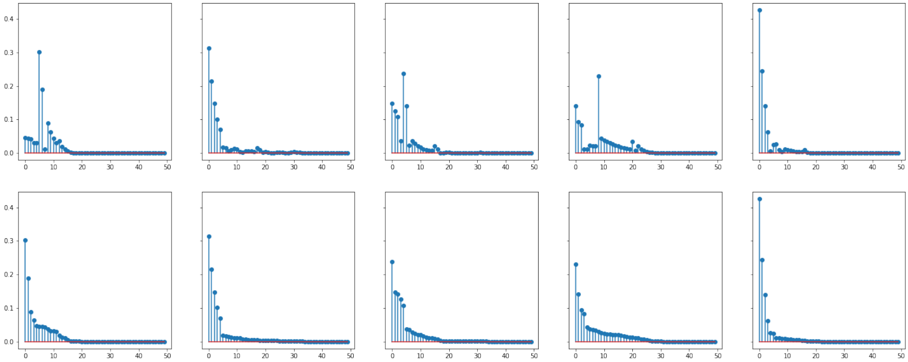
Beyond exchangeability

Stick breaking priors *with Markovian length variables*

- Many NMRIs have dependency in thier stick-breaking process: v_i depends on $v_{<i}$
- Can construct a *stationary Markov sequence* of sticks
  - Example: $x_i \sim \text{Bin}(\kappa, v_i), \qquad v_{i+1} \sim \text{Beta}(1 + x_i, \theta + \kappa - x_i)$
  - Example: $v_i \sim \rho\delta_{v_{i-1}} + (1-\rho)\nu_0$

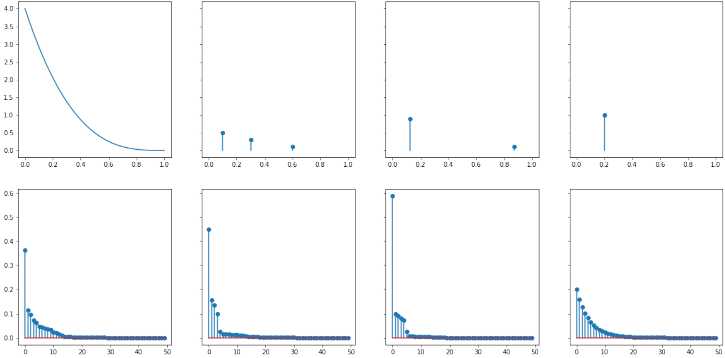
Markov stick-breaking ($E_\nu[v] = 0.2$)

- Stick-breaking $\nu_i = 0.5\delta_{v_{i-1}} + 0.5\text{Beta}(1, 4)$



Behavior of ESB models

- Talk: # clusters is “in between” DP and Geometric
- Talk: Ordering is also ”in between” size-biased and size-ordered
- Empirically/intuitively, support can have more “structured” behavior



- Can we exploit this to learn better predictive performance? How well do we predict future species proportions?

Behavior of ESB models

- Stick breaking means we can construct a slice sampler
- Potential issue: hard to change ordering of variables
- Paper/Talk: We know about the order statistics!
- Can we use this in our samplers?

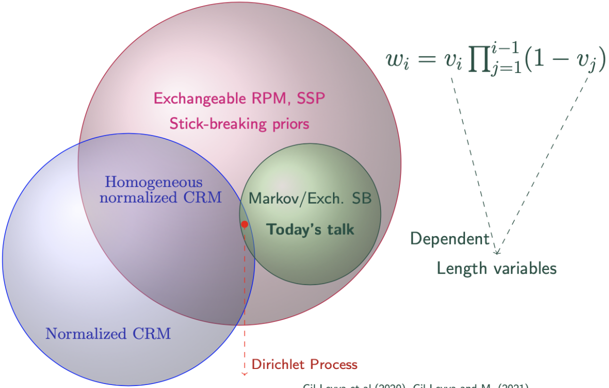
Behavior of ESB models

- Stick breaking means we can construct a slice sampler
- Potential issue: hard to change ordering of variables
- Paper/Talk: We know about the order statistics!
- Can we use this in our samplers?

- Ordering matters... different orders $\rightarrow$ different distributions
- Different ESBs $\leftrightarrow$ different orderings
- Identifiability issues?

More types of stick-breaking priors??

- These works massively expand the scope of stick-breaking priors



Gil-Leyva et al.(2020), Gil-Leyva and M. (2021)

More types of stick-breaking priors??

- Like the DP and GP, ESB processes assume $\nu_i := \nu$
- Markov SBPs introduce dependency in the sticks.
- NRMIs (can) depend on the whole history...
- What do we lose (or gain?) by being Markovian?
- What other options are there? Non-stationarity? Other forms of dependence?