

Title

Subtitle

Bruno Guindani
Elena Zazzetti



POLITECNICO
MILANO 1863

November 22, 2019

<https://github.com/poliprojects/BNPLib>

Title

Non-Parametric statistics

- Goal: density approximation
- **Infinite-dimensional** parameters, e.g. functions

Non-Parametric statistics (the Bayesian way)

- Goal: density approximation
- **Infinite-dimensional** parameters, e.g. functions
- Model:

$$y_i|G \stackrel{\text{iid}}{\sim} G, \quad i = 1, \dots, n$$
$$G \sim \mathcal{P}$$

Non-Parametric statistics (the Bayesian way)

- Goal: density approximation
- **Infinite-dimensional** parameters, e.g. functions
- Model:

$$y_i | G \stackrel{\text{iid}}{\sim} G, \quad i = 1, \dots, n$$
$$G \sim \mathcal{P}$$

$$\mathcal{P} : \Omega \rightarrow M(S) \text{ fixed}$$
$$[\omega \mapsto G(\cdot)]$$

- Model name: **BNP model**

Dirichlet Process prior

$$y_i | G \stackrel{\text{iid}}{\sim} G$$
$$G \sim DP(MG_0)$$

- Parameters: $M > 0$, $G_0 \in M(S)$
- Defining property: $\forall \{B_{1:k}\}$ partition of S ,

$$[G(B_1), \dots, G(B_k)] \sim \text{Dir}(MG_0(B_1), \dots, MG_0(B_k))$$

Dirichlet Process prior

$$y_i | G \stackrel{\text{iid}}{\sim} G$$
$$G \sim DP(MG_0)$$

- Parameters: $M > 0$, $G_0 \in M(S)$
- Defining property: $\forall \{B_{1:k}\}$ partition of S ,

$$[G(B_1), \dots, G(B_k)] \sim \text{Dir}(MG_0(B_1), \dots, MG_0(B_k))$$

- Discreteness** (stick-breaking): $G(\cdot) = \sum_{k=1}^{+\infty} w_k \delta_{m_k}(\cdot)$
- Conjugacy**: $G|\mathbf{y} \sim DP(MG_0 + \sum_i \delta_{y_i})$
- Polya urn representation:

$$\mathcal{L}(y_i | y_1, \dots, y_{i-1}) \propto \sum_{h=1}^{i-1} \delta_{y_h}(y_i) + MG_0(y_i)$$

- $M \log n \ll n$ values

Continuous density estimation

- **Mixtures:**

$$y_i|G \sim f_G(y) = \int f_{\vartheta}(y) \, \mathrm{d}G(\vartheta)$$

Continuous density estimation

- **Mixtures:**

$$y_i|G \sim f_G(y) = \int f_{\vartheta}(y) \mathrm{d}G(\vartheta)$$
$$G \sim DP(MG_0)$$

- Model name: **DPM model**

Continuous density estimation

- **Mixtures:**

$$y_i|G \sim f_G(y) = \int f_{\vartheta}(y) \mathrm{d}G(\vartheta)$$
$$G \sim DP(MG_0)$$

- Model name: **DPM model**

- Equivalent to:

$$y_i|\vartheta_i \stackrel{\text{iid}}{\sim} f_{\vartheta_i}$$
$$\vartheta_i|G \stackrel{\text{iid}}{\sim} G$$
$$G \sim DP(MG_0)$$

- ϑ_i “latent variables” $\forall i = 1, \dots, n$

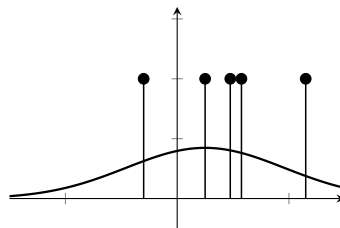
Clustering in the DPM

- Discreteness: the ϑ_i have one of the k **unique values** ϕ_j ($j = 1, \dots, k$)
- $k \simeq M \log(n) \ll n$
- All i s.t. $\vartheta_i = \phi_j$ belong to cluster S_j ($j = 1, \dots, k$), and $n_j = |S_j|$

Clustering in the DPM

- Discreteness: the ϑ_i have one of the k **unique values** ϕ_j ($j = 1, \dots, k$)
- $k \simeq M \log(n) \ll n$
- All i s.t. $\vartheta_i = \phi_j$ belong to cluster S_j ($j = 1, \dots, k$), and $n_j = |S_j|$
- Conditional prior for ϑ_i :

$$\mathcal{L}(\vartheta_i | \boldsymbol{\vartheta}_{-i}) \propto \sum_{j=1}^{k^-} \underset{\uparrow}{n_j^-} \underset{\uparrow}{\delta_{\phi_j^-}}(\vartheta_i) + MG_0(\vartheta_i)$$

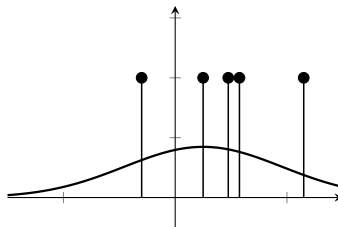


Clustering in the DPM

- Discreteness: the ϑ_i have one of the k **unique values** ϕ_j ($j = 1, \dots, k$)
- $k \simeq M \log(n) \ll n$
- All i s.t. $\vartheta_i = \phi_j$ belong to cluster S_j ($j = 1, \dots, k$), and $n_j = |S_j|$
- Conditional prior for ϑ_i :

$$\mathcal{L}(\vartheta_i | \boldsymbol{\vartheta}_{-i}) \propto \sum_{j=1}^{k-} n_j^- \delta_{\phi_j^-}(\vartheta_i) + MG_0(\vartheta_i)$$

$\uparrow \qquad \qquad \uparrow$





- Via **conjugacy** of the DP:

Title

Title

Stuff

Bibliography

-  Muller, Quintana, *Bayesian Nonparametric Data Analysis*
-  Neal (2000), *Markov Chain Sampling Methods for Dirichlet Process Mixture Models*
-  Ishwaran, James (2001), *Gibbs Sampling Methods for Stick-Breaking Priors*