

BNPlib for density estimation:

A nonparametric C++ library
(part 3)

Bruno Guindani
Elena Zazzetti

February 19th, 2020



POLITECNICO
MILANO 1863

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

Title

Cluster estimation

```
unsigned int cluster_estimate();
```

$$\hat{k} = \arg \min_k \left\| D^{(k)} - \bar{D} \right\|_F^2 = \arg \min_k \sum_{i,j} \left(D_{ij}^{(k)} - \bar{D}_{ij} \right)^2$$

Density estimation

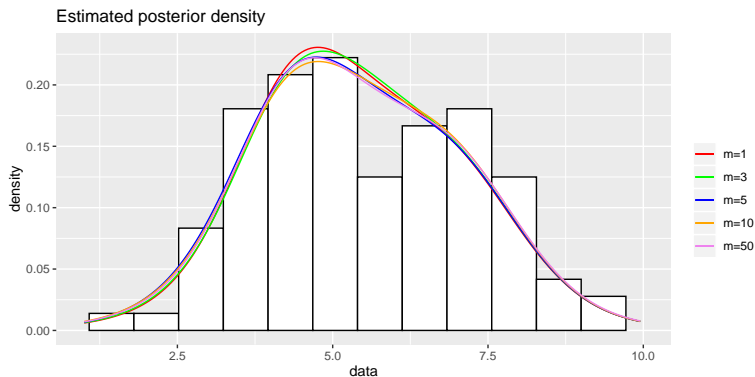
```
void eval_density(const std::vector<double> grid);
```

$$\hat{f}^{(k)}(x) = \sum_j \frac{n_j^{(k)}}{M+n} f(x|\phi_j^{(k)}) + \frac{M}{M+n} m(x)$$

$$\hat{m}(x) = \frac{1}{m} \sum_{h=0}^{m-1} f(x|\phi_h)$$

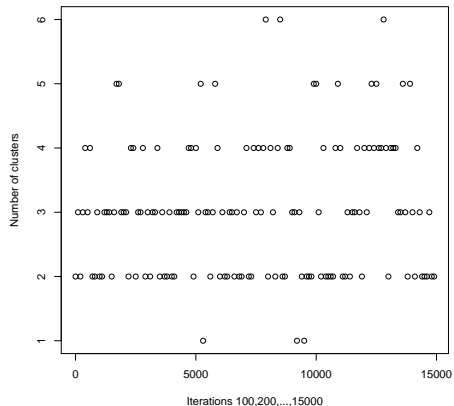
$$\implies \hat{f}(x) = \frac{1}{K} \sum_k \hat{f}^{(k)}(x)$$

Auxiliary parameters

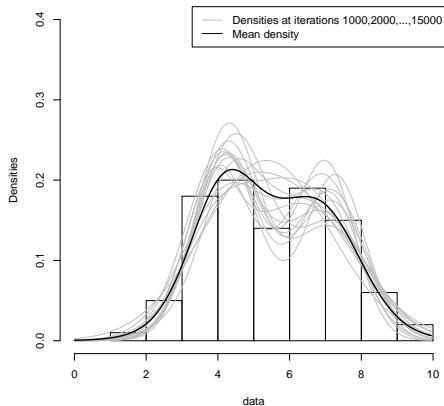


Oscillations

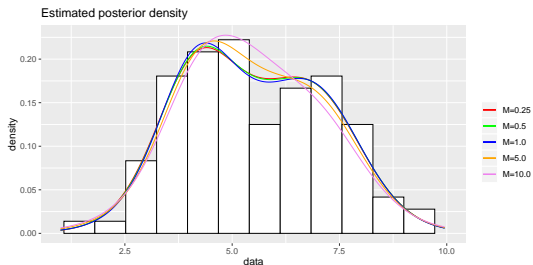
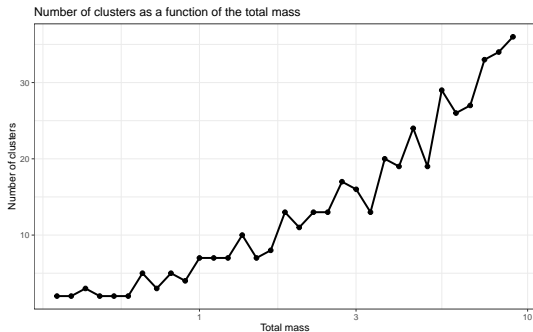
Number of clusters at single iterations



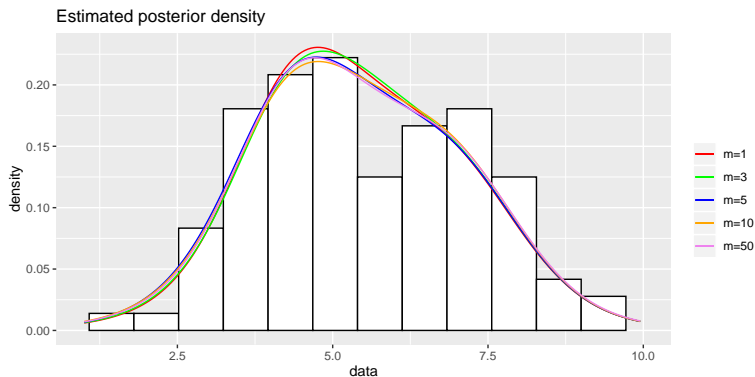
Local density estimates at single iterations



Total mass

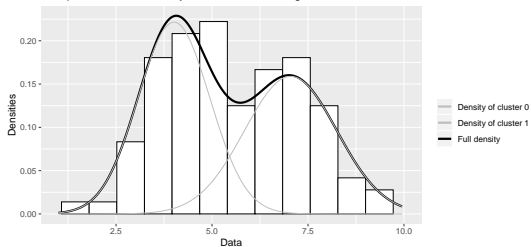


Auxiliary parameters

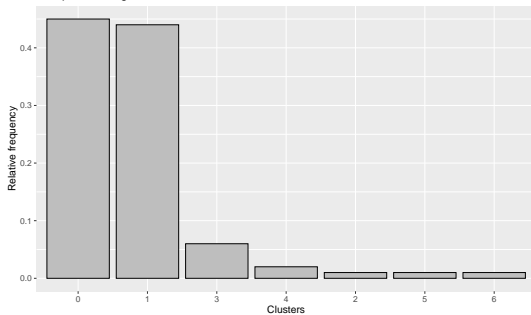


Density components

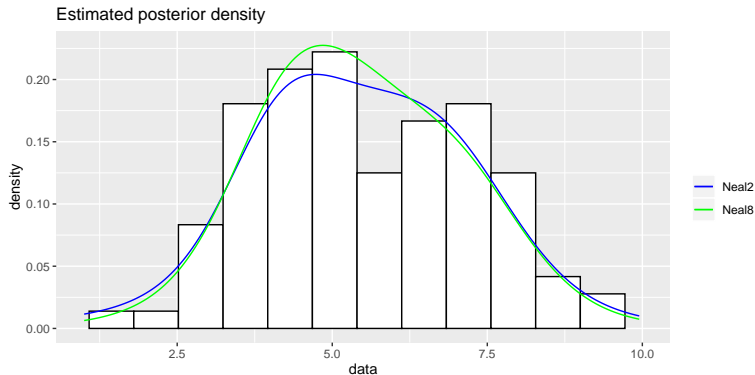
Components of the density at the best clustering










Component weights



Neal2 vs Neal8



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*
-  Murphy (2007), *Conjugate Bayesian analysis of the Gaussian distribution*
-  Protocol Buffers: <https://developers.google.com/protocol-buffers/docs/cpp/tutorial>
-  Stan: <http://mc-stan.org/math>
-  Eigen: <https://eigen.tuxfamily.org/dox>
-  GitHub codes of Mario Beraha and Riccardo Corradin for similar projects