23: Dirichlet Process Models

Taylor

University of Virginia

Introduction

TODO

Definitions

A probability model for the density analagous to the histogram

$$f(y \mid \pi_1, \ldots, \pi_k) = \sum_{h=1}^{k} 1_{\xi_{h-1} < y \leq \xi_h} \frac{\pi_h}{(\xi_h - \xi_{h-1})}$$

where $\xi_0 < \xi_1 < \cdots < \xi_k$ are your **knot points**, and $(\pi_1, \ldots, \pi_k)$ is an unknown probability vector.

Definitions

A probability model for the density analagous to the histogram

$$f(y \mid \pi_1, \ldots, \pi_k) = \sum_{h=1}^{k} 1_{\xi_{h-1} < y \leq \xi_h} \frac{\pi_h}{(\xi_h - \xi_{h-1})}$$

where $\xi_0 < \xi_1 < \cdots < \xi_k$ are your **knot points**, and $(\pi_1, \ldots, \pi_k)$ is an unknown probability vector.

Note

$$\int f(y)\mathrm{d}y = \sum_{h=1}^{k} \mathrm{base}_h \times \mathrm{height}_h = \sum_{h=1}^{k} (\xi_h - \xi_{h-1}) \frac{\pi_h}{(\xi_h - \xi_{h-1})} = 1.$$

Definitions

$$f(y \mid \pi) = \sum_{h=1}^{k} 1_{\xi_{h-1} < y \le \xi_h} \frac{\pi_h}{(\xi_h - \xi_{h-1})}$$

We can put a Dirichlet$(\alpha_1, \ldots, \alpha_k)$ prior on the parameters $\pi = (\pi_1, \ldots, \pi_k)$:

$$p(\pi) = \frac{\Gamma\left(\sum_{h=1}^{k} a_h\right)}{\prod_{h=1}^{k} \Gamma(\alpha_h)} \prod_{h=1}^{k} \pi_h^{a_h - 1}$$

where $a = (a_1, \ldots, a_k)$ are the chosen parameters of the prior.

Definitions

Note that $f(y \mid \pi)$ was for one data point y. Let $\sigma(i) = \{k : \xi_{k-1} < y_i \leq \xi_k\}$. Notice that this function is many-to-one. Then

$$
\begin{aligned}
p(y \mid \pi) &= \prod_{i=1}^{n} f(y_i \mid \pi) \\
&= \prod_{i=1}^{n} \left[\sum_{h=1}^{k} 1_{\xi_{h-1} < y_i \leq \xi_h} \frac{\pi_h}{(\xi_h - \xi_{h-1})} \right] \\
&= \prod_{i=1}^{n} \left[\frac{\pi_{\sigma(i)}}{(\xi_{\sigma(i)} - \xi_{\sigma(i)-1})} \right] \\
&= \prod_{h=1}^{k} \left[\frac{\pi_h}{(\xi_h - \xi_{h-1})} \right]^{n_h}
\end{aligned}
$$

where $n_h = \sum_{i=1}^{n} 1_{\xi_{h-1} < y_i \leq \xi_h}$.

Definitions

Bayes' rule:

$$
\begin{aligned}
p(\pi \mid y) &\propto p(y \mid \pi)p(\pi) \\
&= \left[\prod_{h=1}^{k} \frac{\pi_h}{(\xi_h - \xi_{h-1})}\right]^{n_h} \prod_{h=1}^{k} \pi_h^{a_h-1} \\
&\propto \prod_{h=1}^{k} \pi_h^{a_h+n_h-1}
\end{aligned}
$$

So $p(\pi \mid y) = \text{Dirichlet}(a_1 + n_1, \ldots, a_k + n_k)$.

Definitions

Bayes' rule:

$$
\begin{aligned}
p(\pi \mid y) &\propto p(y \mid \pi) p(\pi) \\
&= \left[\prod_{h=1}^{k} \frac{\pi_h}{(\xi_h - \xi_{h-1})} \right]^{n_h} \prod_{h=1}^{k} \pi_h^{a_h - 1} \\
&\propto \prod_{h=1}^{k} \pi_h^{a_h + n_h - 1}
\end{aligned}
$$

So $p(\pi \mid y) = \text{Dirichlet}(a_1 + n_1, \ldots, a_k + n_k)$.

But bin specification is annoying!