Resume: bigram language model

Define the model:

$$p(\mathbf{w}) = \prod_{i=1}^{k+1} p(w_i|w_{i-1})$$

Estimate the probabilities:

$$p(w_i|w_{i-1}) = \frac{c(w_{i-1}w_i)}{\sum_{w_i} c(w_{i-1}w_i)} = \frac{c(w_{i-1}w_i)}{c(w_{i-1})}$$

It's all about counting!