The Dirichlet distribution

- The Dirichlet distribution is often referred to the distribution of distributions. It is called so because it is the conjugate prior of the multinomial distribution, which is a common distribution used to describe multivariate random variables.

$$p(\boldsymbol{\phi}, \boldsymbol{\alpha}) = \frac{1}{B(\boldsymbol{\alpha})} \prod_{i=1}^{K} \phi_i^{\alpha_i - 1} \quad \text{where} \quad B(\boldsymbol{\alpha}) = \frac{\prod_{i=1}^{K} \Gamma \alpha_i}{\Gamma \left(\sum_{j=1}^{K} \alpha_j \right)}, \quad \Gamma(x) = (x - 1)!$$

- The generative process in LDA is:

  (a) $n \sim Poisson(\xi)$ sample number of words from Poisson distribution (or another suitable distribution)

  (b) $\phi \sim \text{Dir}(\alpha)$ sample multinomial topic distribution from a Dirichlet distribution parameterized by α

  (c) $\forall i \in \{1,...,n\}$ do: loop through each word position in document and do:

       (d) $z_i \sim \text{Mult}(\phi)$ sample topic from multinomial distribution parameterized by ϕ

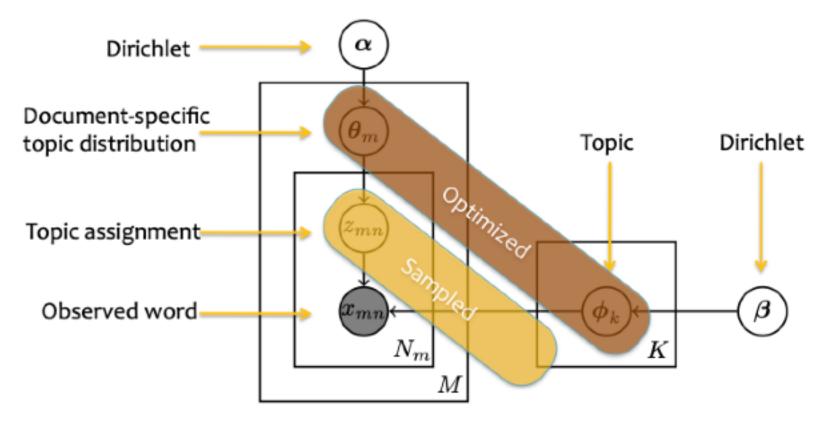   (e) $w_i \sim p(w_i | z_i, \beta)$ sample word from multinomial distribution parameterized by ϕ and conditioned on the topic
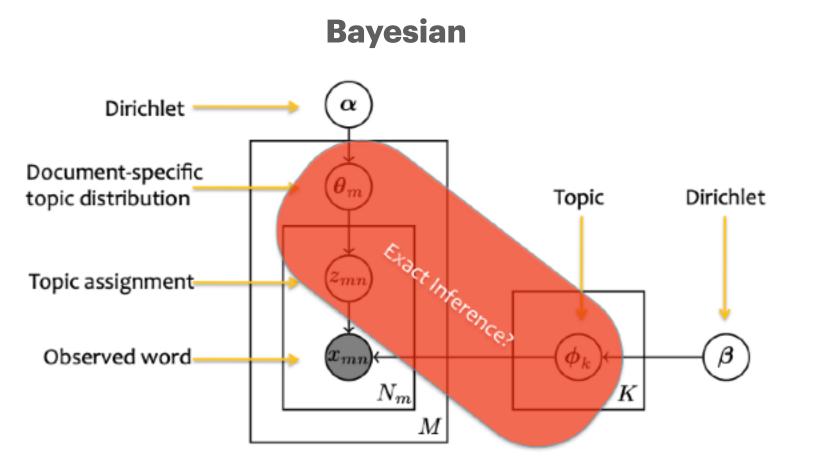
LDA parameter estimation

MLE

MAP

Bayesian

Monte Carlo sampling

Taken from Gormley, 2016