Posteriors and Predictions

in Beta-Bernoulli and Dirichlet-Categorical distributions

David Mareček

October 14, 2021

EUROPEAN UNION
European Structural and Investment Fund
Operational Programme Research,
Development and Education

Charles University
Faculty of Mathematics and Physics
Institute of Formal and Applied Linguistics

Definition and Posterior probability

Beta distribution:

$$Beta(\pi|\alpha,\beta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\pi^{\alpha-1}(1-\pi)^{\beta-1}$$

Dirichlet distribution (generalization of Beta distribution to m outcomes):

$$Dir(\vec{\pi}|\alpha_1,\ldots,\alpha_m) = \frac{\Gamma(\sum_{i=1}^{m}\alpha_i)}{\prod_{i=1}^{m}\Gamma(\alpha_i)}\prod_{i=1}^{m}\pi_i^{\alpha_i-1}$$

Posterior probability for Beta-Bernoulli distribution:

$$p(\pi|D) = \frac{p(\pi|\alpha,\beta)p(D|\pi)}{p(D)} \propto \pi^{\alpha-1}(1-\pi)^{\beta-1}\pi^{k}(1-\pi)^{n-k} = \pi^{\alpha+k-1}(1-\pi)^{\beta+n-k-1}$$

$$p(\pi|D) = Beta(\pi|\alpha+k,\beta+n-k)$$

Posterior probability for Dirichlet-Categorical distribution:

$$p(\vec{\pi}|D) = \frac{p(\vec{\pi}|\vec{\alpha})p(D|\vec{\pi})}{p(D)} \propto \prod_{i=1}^{m}\pi_i^{\alpha_i-1}\pi_i^{k_i} = \prod_{i=1}^{m}\pi_i^{\alpha_i+k_i-1} \propto Dir(\vec{\pi}|\vec{\alpha}+\vec{k})$$

Predictive distributions

Beta-Bernoulli predictions:

$$p(x_{next} = 1|D) = \int_0^1 p(x_{next} = 1|\pi)p(\pi|D)d\pi = \int_0^1 \pi Beta(\pi|\alpha{+}k, \beta{+}n{-}k)d\pi = \frac{\alpha + k}{\alpha + \beta + n}$$

Dirichlet-Categorical predictions:

$$p(x_{next} = j|D) = \int_\triangle p(x_{next} = j|\vec{\pi})p(\vec{\pi}|D)d\vec{\pi} = \int_\triangle \pi_j Dir(\vec{\pi}|\vec{\alpha} + \vec{k})d\vec{\pi} = \frac{\alpha_j + k_j}{\sum_{i=1}^m (\alpha_i + k_i)}$$

The sign $\triangle$ indicates a simplex: the integral is taken across all vectors $\vec{\pi}$ that are valid probability distributions, i.e. $\sum_{i=1}^m \pi_i = 1$

The integrals are equal to expected values of given Beta/Dirichlet posterior distributions.