Dylan DeNicola, Natalie Cardoso
COGS 4210, RPI
Spring 2024

# Homework 3

Problem 1:

Mark all statements which are FALSE.

1. Direct K-fold cross-validation requires K model re-fits, which may be computationally demanding, especially when inverse inference is costly. ***TRUE***
2. Bayes factors (BFs) are relative measures, that is, they cannot differentiate between "equally good" and "equally bad" models. ***TRUE***
3. Marginal likelihoods and, by extension, Bayes factors (BFs) cannot be used to compare models with different likelihoods. ***FALSE***
4. Both the Binomial and the Dirichlet distribution can be formulated as special cases of the Multinomial distribution. ***TRUE***
5. Bayesian leave-one-out cross-validation (LOO-CV) relies on the posterior predictive distribution of left-out data points. ***TRUE***
6. The Akaike Information Criterion (AIC) penalizes model complexity indirectly through the variance of a model's marginal likelihood. ***FALSE***
7. The log-predictive density (LPD) is a relative metric of model complexity. ***FALSE***
8. The LPD can be approximated by evaluating the likelihood of each posterior draw (e.g., as provided by an MCMC sampler) and taking the average of all resulting likelihood values. ***TRUE***
9. Bayes factors do not depend on the prior odds, that is, the ratio of prior model probabilities p(M1)/p(M2). ***TRUE***
10. You should always prefer information criteria to cross-validation in terms of estimation predictive performance. ***FALSE***

Problem 2: Solution is in the corresponding Jupyter notebook on the GitHub.

Problem 3: Solution is in the corresponding Jupyter notebook on the GitHub.

Problem 4: Derive the analytic posterior for the conjugate Dirichlet-Multinomial model.

$$\theta \sim Dirichlet(\alpha)$$

$$y \sim Multinomial(y \mid \theta; N)$$

Assume we have $K$ categories. Thus, we have that $\alpha \in \mathbb{R}_+^K$ and $y \in \mathbb{R}_+^K$ such that $\sum_{k=1}^{K} y_k = 1$.

From $\theta \sim Dirichlet(\alpha)$, we know the prior:

$$p(\theta \mid \alpha) = \frac{1}{\mathrm{B}(\alpha)} \prod_{k=1}^{K} \theta_k^{\alpha_k - 1}$$

$\mathrm{B}(\alpha)$ is the multivariate Beta function.

From $y \sim Multinomial(y \mid \theta; N)$, we know the likelihood:

Dylan DeNicola, Natalie Cardoso
COGS 4210, RPI
Spring 2024

# Homework 3

$$p(y \mid \theta; N) = \frac{N!}{y_1! \ldots y_k!} \prod_{k=1}^{K} \theta_k^{y_k}$$

From Bayes' rule proportionality, we can get the posterior:

$$p(\theta \mid y) \propto p(y \mid \theta; N) * p(\theta \mid \alpha)$$

$$p(\theta \mid y; \alpha) \propto \frac{N!}{y_1! \ldots y_k!} \prod_{k=1}^{K} \theta_k^{y_k} * \frac{1}{\mathrm{B}(\alpha)} \prod_{k=1}^{K} \theta_k^{\alpha_k - 1}$$

Drop the constants:

$$p(\theta \mid y; \alpha) \propto \prod_{k=1}^{K} \theta_k^{y_k} * \prod_{k=1}^{K} \theta_k^{\alpha_k - 1}$$

$$p(\theta \mid y; \alpha) \propto \prod_{k=1}^{K} \theta_k^{y_k + a_k - 1}$$

We see that this is just another Dirichlet distribution. Thus, we have the following posterior:

$$p(\theta \mid y; \alpha) = Dirichlet(\alpha + y)$$

We see that $\alpha \in \mathbb{R}_+^K$ and $y \in \mathbb{R}_+^K$ such that $\sum_{k=1}^{K} y_k = 1$.

Problem 5: Solution is in the corresponding Jupyter notebook on the GitHub.

Problem 6: Solution is in the corresponding Jupyter notebook on the GitHub.