Bayesian Learning - Lecture 9

Mattias Villani

Division of Statistics
Department of Computer and Information Science
Linköping University

Lecture overview

- Hierarchical models
- RStan

An hierarchical binomial model

- Example:

$$y_j|\theta_j \sim Bin(n_j, \theta_j),\ j = 1, ..., J.$$

- We could do inference on each θ_j separately. Problem: n_j may be small for some j. Not much info then about θ_j.
- If you knew θ_j, would that give information about θ_i, $i \neq j$? If so, then inference about the parameters θ_j, $j = 1, ..., J$, may 'borrow strength' from each other.
- Extreme case: assume $\theta_j = \theta$ for all j. Define $y = \sum_{j=1}^{J} y_j$ and $n = \sum_{j=1}^{J} n_j$. Straightforward to analyze θ with the usual Beta-Binomial approach.
- Intermediate case: tie the θ's together by assuming a superpopulation/prior

$$\theta_j \overset{iid}{\sim} Beta(\alpha, \beta).$$

An hierarchical binomial model, cont.

- Model summary

$$y_j|\theta_j \sim Bin(n_j, \theta_j),\ j = 1, ..., J.$$
$$\theta_j \overset{iid}{\sim} Beta(\alpha, \beta).$$
$$\alpha \sim Gamma(a_1, a_2).$$
$$\beta \sim Gamma(b_1, b_2).$$

- Sample from the joint posterior of $p(\theta, \alpha, \beta|y) = p(\theta|\alpha, \beta, y)p(\alpha, \beta|y)$ by sampling from:
  - $\theta_j|\alpha, \beta, y,\ j = 1, ..., J$, which are independent *Beta* distributions.
  - $p(\alpha, \beta|y)$ can be derived in closed form (similar to eq. 5.8), but cannot be sampled directly. Evaluate on grid and sample.

The one-way normal random effects model

- Consider the data model:

$$y_j|\theta_j \sim N(\theta_j, \sigma_j^2), \;\; \sigma_j^2 \text{ known}$$

- At one extreme we may: estimate each θ_j using the mean $\bar{y}_j$ of observations in the jth group.
- At the other extreme we may: assume $\theta_j = \theta$, for all j. Estimate θ with a pooling of group means $\bar{y}_j$.

The one-way normal random effects model, cont.

- Intermediate: a hierarchical model

$$y_j|\theta_j \sim N(\theta_j, \sigma_j^2), \;\; \sigma_j^2 \text{ known}$$

$$\theta_j|\mu, \tau \sim N(\mu, \tau^2)$$

$$p(\mu, \tau) = p(\mu|\tau)p(\tau) \propto p(\tau)$$

- Here we do not assume equal group mean, yet the estimates of each θ_j borrow strength from each other.