StanKorea 2019W Study1

Bayesian Data Analysis Ch.5
Hierarchical Model

발표자: 장헌우
발표일: 2019.01.14 (Mon)
발표장: NextOpt/StanKorea Office

1. Rat Tumor Example(5.1)

1. Directly Estimating Prior Distribution(Non-Bayesian Method)

- Use conjugate family ->Assuming $\theta \sim Beta(\alpha,\beta)$

- Occurring problem : Data would be used twice (Overestimate)

  Could Ignore posterior uncertainty

1. Rat Tumor Example(5.3)

2. Fully Bayesian Analysis - Constructing Posterior Distribution

- Step: Join posterior density $p(\theta, \Phi \mid y)$

  Conditional posterior density $p(\theta \mid \Phi, y)$

  Marginal posterior distribution $p(\Phi \mid y)$

- Hyperparameter $\Phi=(\alpha,\beta)$ from $Beta(\alpha,\beta)$

1. Rat Tumor Example(5.3)

2. Fully Bayesian Analysis - Constructing Posterior Distribution

$$\begin{aligned} p(\theta, \alpha, \beta | y) &\propto p(\alpha, \beta) p(\theta|\alpha, \beta) p(y|\theta, \alpha, \beta) \\ &\propto p(\alpha, \beta) \prod_{j=1}^{J} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \theta_j^{\alpha-1} (1-\theta_j)^{\beta-1} \prod_{j=1}^{J} \theta_j^{y_j} (1-\theta_j)^{n_j - y_j}. \end{aligned} \quad (5.6)$$

$$p(\theta|\alpha, \beta, y) = \prod_{j=1}^{J} \frac{\Gamma(\alpha+\beta+n_j)}{\Gamma(\alpha+y_j)\Gamma(\beta+n_j-y_j)} \theta_j^{\alpha+y_j-1} (1-\theta_j)^{\beta+n_j-y_j-1}. \quad (5.7)$$

$$p(\alpha, \beta | y) \propto p(\alpha, \beta) \prod_{j=1}^{J} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \frac{\Gamma(\alpha+y_j)\Gamma(\beta+n_j-y_j)}{\Gamma(\alpha+\beta+n_j)}. \quad (5.8)$$

1. Rat Tumor Example(5.3)

2. Fully Bayesian Analysis - Constructing Posterior Distribution

- Discussion about p(α,β) (see Exercise 5.9)

- To make marginal posterior distribution proper integral(5.8)

$$p\left(\log\left(\frac{\alpha}{\beta}\right), \log(\alpha+\beta)\right) \propto \alpha\beta(\alpha+\beta)^{-5/2}. \qquad (5.10)$$

1. Rat Tumor Example(5.3)

3. Fully Bayesian Analysis – Drawing simulations

- Step: Draw hyperparameter Φ from $p(\Phi \mid y)$

  Draw parameter θ from $p(\theta \mid \Phi, y)$, given drawn value of Φ

4. Result

- Posterior variability is higher in Bayesian method, reflecting posterior uncertainty in the hyperparameters

2. Exchangeability(5.2)

1. Divorce Rate Example

If we know extra information, further discussion of exchangeability of data must be needed. See case of the eighth sample is Nevada.

2. Rat Tumor Example

If extra information is given, such as experiment condition, partial exchangeability can be assumed as well.

Use multi-level hierarchical model for these cases.

3. Normal Distribution Model(5.4)

1. Constructing Posterior Distribution(hyperparameter: μ,τ from N(μ,τ))

$$
\begin{aligned}
p(\theta,\mu,\tau|y) &\propto p(\mu,\tau)p(\theta|\mu,\tau)p(y|\theta) \\
&\propto p(\mu,\tau)\prod_{j=1}^{J}\mathrm{N}(\theta_j|\mu,\tau^2)\prod_{j=1}^{J}\mathrm{N}(\overline{y}_{.j}|\theta_j,\sigma_j^2), \qquad (5.16)
\end{aligned}
$$

$$\theta_j|\mu,\tau,y \sim \mathrm{N}(\hat{\theta}_j, V_j),$$

$$\hat{\theta}_j = \frac{\frac{1}{\sigma_j^2}\overline{y}_{.j}+\frac{1}{\tau^2}\mu}{\frac{1}{\sigma_j^2}+\frac{1}{\tau^2}} \quad \text{and} \quad V_j = \frac{1}{\frac{1}{\sigma_j^2}+\frac{1}{\tau^2}}. \qquad (5.17)$$

$$p(\mu,\tau|y) \propto p(\mu,\tau)\prod_{j=1}^{J}\mathrm{N}(\overline{y}_{.j}|\mu,\sigma_j^2+\tau^2). \qquad (5.18)$$

3. Normal Distribution Model(5.4)

2. Integrating Over μ

$$p(\mu, \tau | y) = p(\mu | \tau, y) p(\tau | y). \tag{5.19}$$

$$\begin{aligned} p(\tau|y) &\propto \frac{p(\tau) \prod_{j=1}^{J} \mathrm{N}(\overline{y}_{.j} | \hat{\mu}, \sigma_j^2 + \tau^2)}{\mathrm{N}(\hat{\mu} | \hat{\mu}, V_\mu)} \\ &\propto p(\tau) V_\mu^{1/2} \prod_{j=1}^{J} (\sigma_j^2 + \tau^2)^{-1/2} \exp\left(-\frac{(\overline{y}_{.j} - \hat{\mu})^2}{2(\sigma_j^2 + \tau^2)} \right), \end{aligned} \tag{5.21}$$

3. Normal Distribution Model(5.4)

3. Discussion About $p(\tau)$

- Use inverse χ square distribution(see Exercise 5.10)

4. Difficulties of non Bayesian method

- As seen in the rat tumor example, main difficulty is how to consider the uncertainty of posterior distribution

4. Meta-Analysis and Exchangeability(5.6)

1. In Bayesian view, we regard the studies as exchangeable but not necessarily either identical or completely unrelated.

2. Complete pooling makes the assumption that τ is 0.

3. In contrast, no pooling makes the assumption that τ is infinite.

   Select proper assumption using weighted combination;

$$\hat{\theta}_j = \lambda_j \overline{y}_{.j} + (1 - \lambda_j)\overline{y}_{..},$$

5. Another priors(5.7)

1. Non informative priors

   - Uniform prior(0,A) where A -> infinite

   - Inverse gamma prior(ε,ε) where ε->0

2. Weakly informative priors

   -Half Cauchy prior

Improper prior can leads to proper posterior.