Practical: Combining Gibbs and Metropolis-Hastings Kernels

Instructors: Kari Auranen, Elizabeth Halloran and Vladimir Minin

July 17 – July 17, 2019

1 Beta-binomial hierarchical model

Let $\mathbf{x} = (x_1, \dots, x_n)$, where $x_i \,|\, \theta_i \sim \text{Bin}(n_i, \theta_i)$ and x_is are independent given θ_is. We further assume that $\theta_i \overset{\text{iid}}{\sim} \text{Beta}(\alpha, \beta)$. We group all success probabilities into a vector $\boldsymbol{\theta} = (\theta_1, \dots, \theta_n)$ and put a prior distribution on hyper-parameters α and β, $\Pr(\alpha, \beta)$. Under our assumptions, the posterior distribution becomes

$$\Pr(\boldsymbol{\theta}, \alpha, \beta \,|\, \mathbf{x}) \propto \Pr(\mathbf{x} \,|\, \boldsymbol{\theta}, \alpha, \beta) \Pr(\boldsymbol{\theta}, \alpha, \beta) \propto \Pr(\alpha, \beta) \prod_{i=1}^{n} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \theta_i^{\alpha-1} (1-\theta_i)^{\beta-1} \prod_{i=1}^{n} \theta_i^{x_i} (1-\theta_i)^{n_i - x_i}.$$

We can compute the posterior up to a proportionality constant, but this does not mean that we can compute expectations with respect to the posterior. We will tackle this problem with Markov chain Monte Carlo.

The full condition distribution of θ_i is

$$\Pr(\theta_i \,|\, \mathbf{x}, \alpha, \beta, \boldsymbol{\theta}_{-i}) \propto \theta_i^{x_i+\alpha-1} (1-\theta_i)^{n_i - x_i + \beta - 1}.$$

Therefore,

$$\theta_i \mid \mathbf{x}, \alpha, \beta, \boldsymbol{\theta}_{-i} \sim \text{Beta}(x_i + \alpha, n_i - x_i + \beta).$$

Sampling from $\Pr(\alpha, \beta \,|\, \mathbf{x}, \boldsymbol{\theta})$ directly is difficult, so we will use two Metropolis-Hastings steps to update α and β. To propose new values of α and β, we will multiply their current values by $e^{\lambda(U-0.5)}$, where $U \sim U[0,1]$ and λ is a tuning constant. The proposal density is

$$q(y_{\text{new}} \,|\, y_{\text{cur}}) = \frac{1}{\lambda y_{\text{new}}}.$$

This proposal is not symmetric, so we will have to include it into the M-H acceptance ratio.

Your task

Download the file "beta_bin_reduced.R" from the module web site. We will go through this R script together at first. After you become familiar with data structures used in the script, you will fill in two gaps, marked by "TO DO" comments in the script. Your first task is to replace the line "cur.theta = rep(0.5, data.sample.size)" in the script with code that implements the Gibbs update. Your second task is to implement the M-H steps to sample α and β. The file "beta_bin_reduced.R" contains functions that implement the described proposal mechanism and all the pieces necessary for the acceptance probability. The full MCMC algorithm is outlined on the next page.

Algorithm 1 MCMC for the beta-binomial hierarchical model

Start with some initial values $(\boldsymbol{\theta}^{(0)}, \alpha^{(0)}, \beta^{(0)})$.

for $t = 0$ to N **do**

    for $i = 0$ to n **do**

        Sample $\theta_i^{(t+1)} \sim \text{Beta}(x_i + \alpha^{(t)}, n_i - x_i + \beta^{(t)})$

    end for

    Generate $U_1 \sim U[0,1]$ and set $\alpha^* = \alpha^{(t)} e^{\lambda_\alpha (U_1 - 0.5)}$. Generate $U_2 \sim U[0,1]$ and set

$$\alpha^{(t+1)} = \begin{cases} \alpha^* & \text{if } U_2 \leq \min\left\{ \frac{\Pr\left(\boldsymbol{\theta}^{(t+1)}, \alpha^*, \beta^{(t)} \mid \mathbf{x}\right) q(\alpha^{(t)} \mid \alpha^*)}{\Pr\left(\boldsymbol{\theta}^{(t+1)}, \alpha^{(t)}, \beta^{(t)} \mid \mathbf{x}\right) q(\alpha^* \mid \alpha^{(t)})}, 1 \right\}, \\ \alpha^{(t)} & \text{otherwise.} \end{cases}$$

    Generate $U_3 \sim U[0,1]$ and set $\beta^* = \beta^{(t)} e^{\lambda_\beta (U_3 - 0.5)}$. Generate $U_4 \sim U[0,1]$ and set

$$\beta^{(t+1)} = \begin{cases} \beta^* & \text{if } U_4 \leq \min\left\{ \frac{\Pr\left(\boldsymbol{\theta}^{(t+1)}, \alpha^{(t+1)}, \beta^* \mid \mathbf{x}\right) q(\beta^{(t)} \mid \beta^*)}{\Pr\left(\boldsymbol{\theta}^{(t+1)}, \alpha^{(t+1)}, \beta^{(t)} \mid \mathbf{x}\right) q(\beta^* \mid \beta^{(t)})}, 1 \right\}, \\ \beta^{(t)} & \text{otherwise.} \end{cases}$$

end for

return $(\boldsymbol{\theta}^{(t)}, \alpha^{(t)}, \beta^{(t)})$, for $t = 1, \ldots, N$.