Bayesian Statistics (STA306)
Homework III

08/05/2024–18/05/2024

1. Nonconjugate hierarchical models: suppose that in the rat tumor example, we wish to use a normal population distribution on the log-odds scale: $\operatorname{logit}(\theta_j) \sim N(\mu, \tau^2)$, for $j = 1, \ldots, J$. As in Section 5.3, you will assign a noninformative prior distribution to the hyperparameters and perform a full Bayesian analysis.

   (a) Write the joint posterior density, $p(\theta, \mu, \tau | y)$.

   (b) Show that the integral (5.4) has no closed-form expression.

   (c) Why is expression (5.5) no help for this problem?

2. Rejection sampling and importance sampling: Consider the model, $y_j \sim \text{Binomial}\left(n_j, \theta_j\right)$, where $\theta_j = \operatorname{logit}^{-1}\left(\alpha + \beta x_j\right)$, for $j = 1, \ldots, J$, and with independent prior distributions, $\alpha \sim t_4\left(0, 2^2\right)$ and $\beta \sim t_4(0,1)$. Suppose $J = 10$, the x_j values are randomly drawn from a $\mathrm{U}(0,1)$ distribution, and $n_j \sim \text{Poisson}^{+}(5)$, where Poisson $^{+}$is the Poisson distribution restricted to positive values.

   (a) Sample a dataset at random from the model.

   (b) Use rejection sampling to get 1000 independent posterior draws from (α, β).

   (c) Approximate the posterior density for (α, β) by a normal centered at the posterior mode with covariance matrix fit to the curvature at the mode.

   (d) Take 1000 draws from the two-dimensional t_4 distribution with that center and scale matrix and use importance sampling to estimate $\mathrm{E}(\alpha \mid y)$ and $\mathrm{E}(\beta \mid y)$

   (e) Compute an estimate of effective sample size for importance sampling using (10.4) on page 266.

3. Hierarchical binomial model: Exercise 3.8 described a survey of bicycle traffic in Berkeley, California, with data displayed in Table 3.3. For this problem, restrict your attention to the first two rows of the table: residential streets labeled as 'bike routes,' which we will use to illustrate this computational exercise.

(a) Set up a model for the data in Table 3.3 so that, for $j = 1, \ldots, 10$, the observed number of bicycles at location j is binomial with unknown probability θ_j and sample size equal to the total number of vehicles (bicycles included) in that block. The parameter θ_j can be interpreted as the underlying or 'true' proportion of traffic at location j that is bicycles. (See Exercise 3.8.) Assign a beta population distribution for the parameters θ_j and a noninformative hyperprior distribution as in the rat tumor example of Section 5.3. Write down the joint posterior distribution.

(b) Compute the marginal posterior density of the hyperparameters and draw simulations from the joint posterior distribution of the parameters and hyperparameters, as in Section 5.3.

(c) Compare the posterior distributions of the parameters θ_j to the raw proportions, (number of bicycles / total number of vehicles) in location j. How do the inferences from the posterior distribution differ from the raw proportions?

(d) Give a 95% posterior interval for the average underlying proportion of traffic that is bicycles.

4. Consider the following Exponential model for an observation x :

$$p(x \mid a, b) = ab \exp(-abx) \mathbb{1}(x > 0)$$

and suppose the prior is

$$p(a, b) = \exp(-a - b) \mathbb{1}(a, b > 0).$$

You want to sample from the posterior $p(a, b \mid x)$. Find the conditional distributions needed for implementing a Gibbs sampler and write down the specific algorithm.