ST4234: Bayesian Statistics

Tutorial 8, AY 19/20

Instructions: For each group who were scheduled to present on Tuesday, 31 March, please upload **one** copy of your solutions to the folder "Tutorial submission" by **6pm Tuesday, 31 March**.

1. (Continued from Tutorial 6 Question 3) Fifteen reciprocating pumps were tested for a prespecified time and one assumes that the failure times follow the two-parameter exponential distribution

$$p(y|\beta, \mu) = \frac{1}{\beta} \exp\left\{-\frac{y-\mu}{\beta}\right\}, \quad y \geq \mu.$$

Suppose one places a uniform prior on (μ, β). The posterior density is then given by

$$p(\beta, \mu|\text{data}) \propto \frac{1}{\beta^s} \exp\left\{-\frac{t-n\mu}{\beta}\right\}, \quad \mu \leq t_1,$$

where n is the number of items placed on test, t is the total time on test, t_1 is the smallest failure time, and s is the observed number of failures in a sample of size n. When $n = 15$ pumps were tested for a total time of $t = 15962989$, eight failures ($s = 8$) were observed and the smallest failure time was $t_1 = 237217$. Suppose one transforms the parameters to the real line using $\theta_1 = \log \beta$ and $\theta_2 = \log(t_1 - \mu)$.

(a) Use a multivariate t_4 proposal density and the SIR algorithm to simulate a sample of 10^4 draws from the posterior distribution of (θ_1, θ_2).

(Hint: Your t_4 proposal can be centered at the posterior mode with an appropriate scale matrix.)

(b) Suppose one is interested in estimating the reliability at time t_0, defined by

$$R(t_0) = \exp\left(-\frac{t_0-\mu}{\beta}\right).$$

Using your simulated values from the posterior, find the posterior mean and posterior standard deviation of $R(t_0)$ when $t_0 = 10^6$ cycles.

2. (Continued from Tutorial 7 Question 1) Consider the genetic linkage model in Chapter 5. 197 animals are distributed into four categories as follows (Rao 1997):

$$\boldsymbol{y} = (y_1, y_2, y_3, y_4) = (125, 18, 20, 34)$$

with cell probabilities

$$\left(\frac{1}{2} + \frac{\theta}{4}, \ \frac{1}{4}(1-\theta), \ \frac{1}{4}(1-\theta), \frac{\theta}{4}\right),$$

where $\theta \in [0, 1]$ is the parameter. Suppose that θ is assigned a uniform prior on $[0, 1]$. We consider the posterior $p(\eta|\boldsymbol{y})$ where $\eta = \log \dfrac{\theta}{1-\theta}$ is the logit transformation of θ.

(a) Use a Metropolis-Hastings random walk algorithm (`rwmetrop()`) to simulate from the posterior density of η for the length $T = 10^4$ and discard the first 5000 draws as burn-in. Choose the scale parameter to be twice the approximate posterior standard deviation of η found in a normal approximation. (i) Report the acceptance rate. (ii) Compare the histogram of the simulated output of η with the normal approximation. (iii) From the simulation output, find a 95% interval estimate for the parameter of interest θ.

(b) Repeat Part (a) but with a Metropolis-Hastings independence algorithm (`indepmetrop()`) to simulate from the posterior density of η.

3. Haberman (1978) considers an experiment involving subjects reporting one stressful event. The collected data (shown in Table 1) are $\boldsymbol{y} = (y_1, ..., y_{18})$, where y_i is the number of events recalled i months before the interview. Suppose y_i is independently Poisson distributed with mean λ_i, where the λ_i satisfies the loglinear regression model $\log \lambda_i = \beta_0 + \beta_1 i$.

Months	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
y_i	15	11	14	17	5	11	10	4	8	10	7	9	11	3	6	1	1	4

Table 1: Numbers of subjects recalling one stressful event.

One is interested in the posterior density of the regression coefficients (β_0, β_1). Suppose (β_0, β_1) is assigned an improper uniform prior $p(\beta_0, \beta_1) \propto 1$.

(a) Show that the log posterior density is given by

$$\log p(\beta_0, \beta_1|\boldsymbol{y}) = \sum_{i=1}^{18} [y_i(\beta_0 + \beta_1 i) - \exp(\beta_0 + \beta_1 i)] + C,$$

where C is an additive constant not depending on (β_0, β_1).

(b) Find a normal approximation of the posterior density.

(c) Simulate 10^4 iterates from the posterior and compute the posterior mean and standard deviation of β_1 by using a

  (i) Metropolis random walk algorithm via `rwmetrop`.

  (ii) Metropolis independence algorithm via `indepmetrop`.

(d) Construct a table to compare (i) the acceptance rate of the two Metropolis algorithms and (ii) the 5th, 50th and 95th percentiles of β_0 and β_1 obtained using all three computational methods.