Bayesian Econometrics, ECO545A, Semester I, 2020-21. Homework I (100 points)

Instructor: M.A. Rahman

Deadline: 17:00 hours, October 7, 2020.

Please read the instructions carefully and follow them while writing answers.

- *Solutions to homework should be written in A4 size loose sheets. If you are not comfortable writing on white sheets, please ask for biology paper in Tarun Book Store.*
- *Questions should be answered in order as they appear in the homework. Every new question should begin in a new page. Please number all the pages of your homework solution.*
- *Please leave a margin of one inch from all sides. Staple the sheets on the top-left.*
- *Computational assignments (Matlab):* ***Please answer the questions and report the results. Codes should be attached as an appendix.***

---

1. **(10+10=20 points)** This exercise is based on Marin and Robert (2007), which deals with capture-recapture surveys. Eli and Ezra decide to estimate the number N of frogs in their backyard. They set a (harmless) trap and capture n frogs in an hour. Let p equal the probability of capturing a frog in an hour, and assume that p is independent across individuals frogs and independent of N. Then the number of captured frogs have the binomial distribution. Having captured n, you know that $N \geq n$, so that the likelihood function is,

$$p(n|N,p) \propto \binom{N}{n} p^n (1-p)^{N-n}, \qquad N = n, n+1, \cdots .$$

For priors, assume $p \sim Beta(\alpha, \beta)$, and $p(N) \propto 1/N^2$, where $N = 1, 2, \cdots$. Write the joint distribution for (N, p) and then integrate out p to verify that the marginal posterior distribution for N is,

$$p(N|n) \propto \frac{(N-1)!}{N(N-n)!} \frac{\Gamma(N-n+\beta)}{\Gamma(N+\alpha+\beta)}, \qquad N = n, n+1, \cdots .$$

Write a program to evaluate this expression as a function of N for values of n and the hyperparameters α and β. Explore how the results depend on n and the hyperparameters.

2. **(5+5+5 =15 points)** Consider the following two sets of data obtained after tossing a die 100 and 1,000 times, respectively:

n	1	2	3	4	5	6
100	19	12	17	18	20	14
1000	190	120	170	180	200	140

Suppose you are interested in θ_1, the probability of obtaining a one spot. Assume your prior for all the probabilities is a Dirichlet distribution, where each $\alpha_i = 2$. Compute the posterior distribution for θ_1 for each of the sample sizes in the table. Plot the resulting distribution and compare the results. Comment on the effect of having a larger sample.

3. **(9+6 = 15 points)** Consider the multiple linear regression model $y_i = x_i'\beta + \epsilon_i$ where $\epsilon_i | x_i \sim N(0, \sigma^2)$ for all $i = 1, \cdots, n$. Further, assume that the priors on (β, σ^2) are independent and $\pi(\beta, \sigma^2) = \pi(\beta)\pi(\sigma^2) = N_k(\beta_0, B_0)\, IG(\alpha_0/2, \delta_0/2)$. Based on the given setting answer the following without skipping any steps.

   (a) Derive the conditional posterior distribution of β and show that $\pi(\beta|\sigma^2, y) \sim N(\bar{\beta}, B_1)$, where $B_1 = \left[\sigma^{-2}X'X + B_0^{-1}\right]^{-1}$ and $\bar{\beta} = B_1\left[\sigma^{-2}X'y + B_0^{-1}\beta_0\right]$.

   (b) Derive the conditional posterior distribution of σ^2 and show that $\pi(\sigma^2|\beta, y) \sim IG(\alpha_1/2, \delta_1/2)$, where $\alpha_1 = \alpha_0 + n$ and $\delta_1 = \delta_0 + (y - X\beta)'(y - X\beta)$.

4. **(10+10 = 20 points)**. Estimate the mean of a Beta$(3.7, 4.8)$ distribution with (1) an AR algorithm and a Beta$(4, 5)$ proposal density (you will need to determine the value of c needed in Algorithm 5.2); (2) an MH algorithm with a Beta$(4, 5)$ proposal density. After the burn-in sample, graph the values of the mean against the iteration number to monitor convergence. Compare your answers to the true value.

5. **(5+5+5+5+5+5 = 30 points)**. Consider the model,

$$y_i = \beta x_i + u_i, \qquad u_i \sim N(0, 1), \quad i = 1, \ldots, n,$$

with the gamma prior distribution $\beta \sim \text{Ga}(2, 1)$, $\beta > 0$. Verify that the posterior distribution is,

$$\pi(\beta|y) \propto \beta \exp[-\beta] \exp\left[-\frac{1}{2}\sum_{i=1}^{n}(y_i - \beta x_i)^2\right] 1(\beta > 0).$$

Note that this distribution does not have a standard form. Construct an MH algorithm to sample from this distribution with an independence kernel, where the kernel is a Student-t distribution truncated to the region $(0, \infty)$, with five degrees of freedom, mean equal to the value of β that maximizes the posterior distribution ($\hat{\beta}$), and scale factor equal to the negative inverse of the second derivative of the posterior distribution evaluated at $\hat{\beta}$. Verify that

$$\hat{\beta} = \frac{\left(\sum_i x_i y_i - 1\right) + \sqrt{\left(\sum_i x_i y_i - 1\right)^2 + 4\sum_i x_i^2}}{2\sum_i x_i^2},$$

and that the scale factor is $\left(1/\hat{\beta}^2 + \sum_i x_i^2\right)^{-1}$. Generate a dataset by choosing $n = 50$, x_i from $N(0,1)$, and a value of β from its prior distribution. Write a program to implement your algorithm and see how well β is determined. You may try larger values of n (say 300) to explore the effect of sample size, and, depending on the acceptance, you may wish to change the scale factor.

(Note the breakdown of points: 5 points each for verifying the posterior distribution, constructing the MH algorithm, verifying the value of $\hat{\beta}$, verifying the scale factor, program to implement the algorithm, and using a higher value of n.)

References

Marin, J.-M. and Robert, C. P. (2007), *Bayesian Core: A Practical Approach to Computational Bayesian Statistics*, Springer, New York.