Fall 2019 Jingchen (Monika) Hu

MATH 347 Homework 4 (Total 45 points)

Due: Thursday 10/31, at the beginning of class

Name: ______________________________

- **Print out this cover page and staple with your homework.**
- **Show all work. Incomplete solutions will be given no credit.**
- **You may prepare either hand-written or typed solutions, but make sure that they are legible. Answers that cannot be read will be given no credit.**
- **R graphical outputs must be printed instead of hand-drawn.**

1. (*12 points; 3 points each part*)
   Suppose two people, Chrystal and Danny, have different prior beliefs about the average number of ER visits during the 10pm - 11pm time period. Chrystal's prior information is matched to a Gamma distribution with parameters $\alpha = 70$ and $\beta = 10$, and Danny's beliefs are matched to a Gamma distribution with $\alpha = 33.3$ and $\beta = 3.3$.

   (a) Using the `rgamma` or the `dgamma` functions to plot these two gamma priors on the same graph. (Hint: if you use `dgamma`, you can let the range of x be (0, 20].)

   (b) Compare the priors of Chrystal and Danny with respect to average value and spread. Which person believes that there will be more ER visits, on average? Which person is more confident of his/her/their "best guess" at the average number of ER visits?

   (c) Using the `qgamma` function, construct a 90% credible interval for λ using Chrystal's prior and Danny's prior.

   (d) After some thought, Chrystal believes that her best prior guess at λ is correct, but she is less confident in this guess. Explain how Chrystal can adjust the parameters of her gamma prior to reflect this new prior belief. (Hint: the mean and variance of Gamma(α, β) is $\frac{\alpha}{\beta}$ and $\frac{\alpha}{\beta^2}$ respectively.)

2. (*9 points; 3 points each part*)
   Continuing from Question 1. A hospital collects the number of patients in the emergency room (ER) admitted between 10 pm and 11 pm for each day of a week. Table 1 records the day and the number of ER visits for the given day.

Table 1: Data for ER visits in a given week.

Day	Number of ER visits
Sunday	8
Monday	6
Tuesday	6
Wednesday	9
Thursday	8
Friday	9
Saturday	7

Suppose one assumes Poisson sampling for the counts, and a conjugate Gamma prior with parameters $\alpha = 70$ and $\beta = 10$ for the Poisson rate parameter λ (i.e. Chrystal's first prior).

(a) Given the sample shown in Table 1, obtain the posterior distribution for λ through the Gamma-Poisson conjugacy. Obtain a 95% posterior credible interval for λ.

(b) Suppose a hospital administrator states that the average number of ER visits during any evening hour does not exceed 6. By computing a posterior probability, evaluate the validity of the administrator's statement.

(c) The hospital is interested in predicting the number of ER visits between 10 pm and 11 pm for another week. Use simulations to generate posterior predictions of the number of ER visits for another week (seven days).

3. (*12 points; 3 points each part*)
Suppose that we want inference about an unknown number of animals N in a fixed-size population. On five separate days, we take photographs of some areas where they reside, and count the number of animals in the photos $(y_1, \ldots, y_5)$. Suppose further than each animal has a constant probability θ of appearing in a photograph, and that appearances are independent across animals and days. A reasonable model for such data is a binomial distribution, $y_i \sim \text{Binomial}(N, \theta)$. In our setting, neither the number of trials N nor the probability θ are known.

To get a posterior distribution for N and θ, we propose the following system of models (Raftery 1988)

$$\begin{aligned} y_i \mid N, \theta &\sim \text{Binomial}(N, \theta) \\ N \mid \theta, \lambda &\sim \text{Poisson}(\lambda/\theta) \\ f(\lambda, \theta) &\propto 1/\lambda \end{aligned}$$

where $\lambda > 0$ is a continuous random variable introduced to help with computations.

(a) Write the expression of the joint posterior distribution, $p(N, \theta, \lambda \mid y_1, \ldots, y_5)$, up to a multiplicative constant.

(b) Find an expression for the conditional distribution, $f(\lambda \mid y_1, \ldots, y_5, N, \theta)$. Write the name of the distribution, and specify expressions for its parameter values.

(c) Find the kernel of the posterior distribution $f(N, \theta \mid y_1, \ldots, y_5)$ by integrating $f(N, \theta, \lambda \mid y_1, \ldots, y_5)$ with respect to λ. You don't need to name it; just write its mathematical form.

(d) Find the conditional distribution, $f(\theta \mid y_1, \ldots, y_5, N)$. Write the name of the distribution, and specify expressions for its parameter values.

4. (*12 points; 3 points each part*)
Consider a three-component mixture model, where the posterior for parameter μ is:

$$\mu \mid y \sim 0.45 \times \text{Normal}(-3, (1/3)^2) + 0.1 \times \text{Normal}(0, (1/3)^2) + 0.45 \times \text{Normal}(3, (1/3)^2). \qquad (1)$$

How can we draw samples from this posterior?

<u>Approach 1: Monte Carlo approximation</u>
If we introduce a "mixture component indicator", δ, an unobserved latent variable, the sampling can be greatly simplified, as in the following:

- $\delta = 1$, then $\mu \mid \delta, y \sim \text{Normal}(-3, (1/3)^2)$ and $p(\delta = 1 \mid y) = 0.45$.
- $\delta = 2$, then $\mu \mid \delta, y \sim \text{Normal}(0, (1/3)^2)$ and $p(\delta = 2 \mid y) = 0.1$.

- $\delta = 3$, then $\mu \mid \delta, y \sim \text{Normal}(3, (1/3)^2)$ and $p(\delta = 3 \mid y) = 0.45$.

Therefore, we can draw δ and then draw μ given δ through Monte Carlo approximation. Figure 1 is the histogram of μ from 1000 Monte Carlo draws with posterior density as a solid line.

Approach 2: Markov chain Monte Carlo - Gibbs

We can also derive the full conditional posterior distributions for $\pi(\mu \mid \delta, y)$ and $\pi(\delta \mid \mu, y)$, and use a Gibbs sampler to sample the posterior of μ. The full conditional posterior distributions are eliminated for brevity, but be convinced that a Gibbs sampler can be developed for this case. Figure 2 is the histogram of μ from 1000 MCMC/Gibbs draws with posterior density as a solid line. Note that the initial values are $\mu = -6$ and $\delta = 1$.

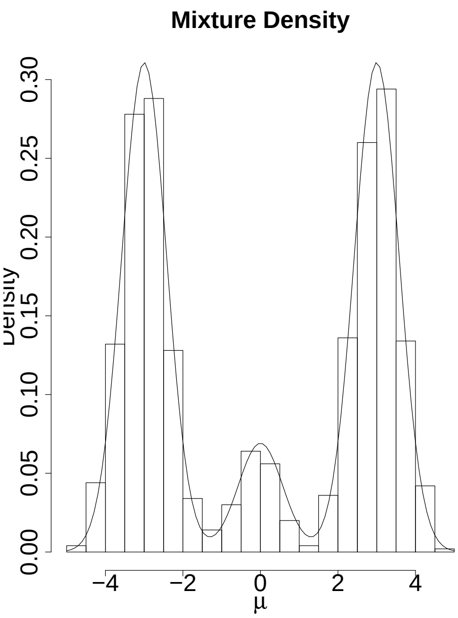


Figure 1: Histogram of 1000 samples of μ from Monte Carlo approximation.

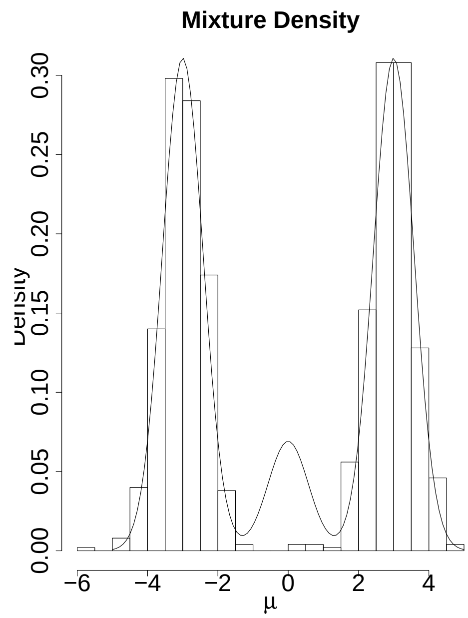
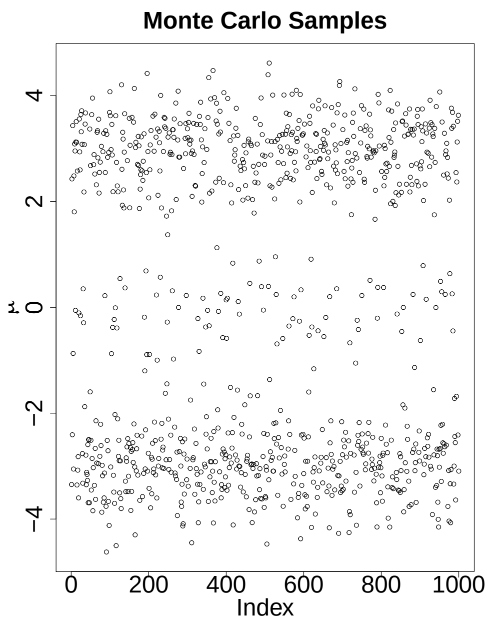


Figure 2: Histogram of 1000 samples of μ from MCMC/Gibbs.

Sometimes, a regular scatterplot could be more descriptive of the problem. See Figure 3 for a scatterplot of 1000 samples of μ from Monte Carlo approximation, and Figure 4 for a scatterplot of 1000 samples of μ from MCMC/Gibbs. Think carefully with respect to the posterior density in Equation (1).



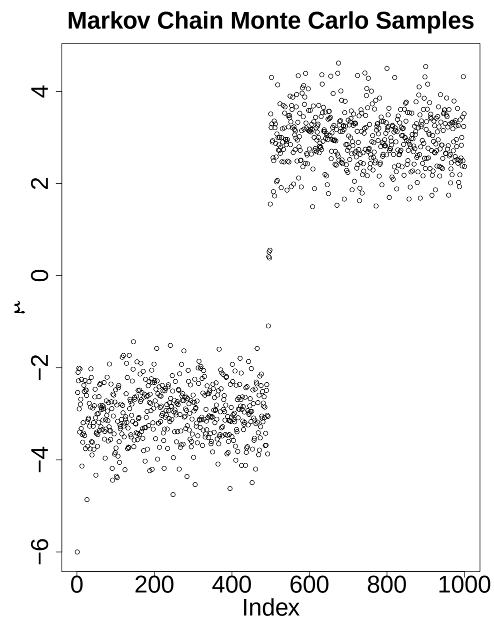


Figure 3: 1000 samples of μ from Monte Carlo approximation.

Figure 4: 1000 samples of μ from MCMC/Gibbs.

(a) Compare Figure 1 with Figure 2, and Figure 3 with Figure 4. What problems do you expect to show up when you perform MCMC diagnostics for μ?

(b) Examine the traceplots of μ from Monte Carlo approximation in Figure 5 and of μ from MCMC/Gibbs in Figure 6. What is the issue with MCMC/Gibbs mixing?

(c) Examine the ACF plots of μ from Monte Carlo approximation in Figure 7 and of μ from MCMC in Figure 8. Also, the effective sample size of MC and MCMC/Gibbs are 1000 and 1.70 respectively. What is the issue with MCMC/Gibbs mixing?

(d) Describe what you would do to improve the MCMC/Gibbs mixing.

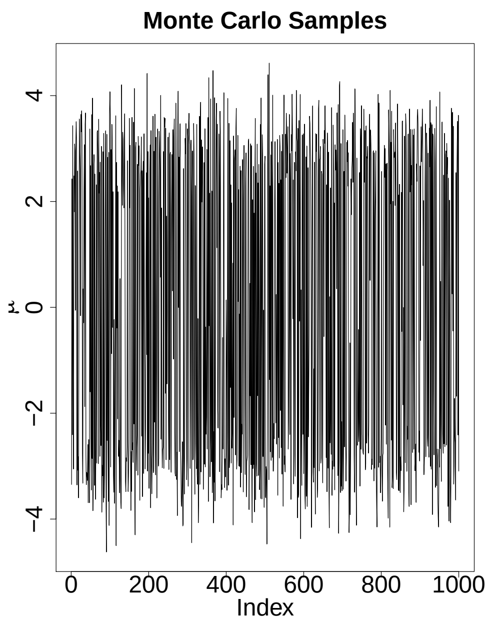


Figure 5: Traceplot of μ from Monte Carlo approximation.

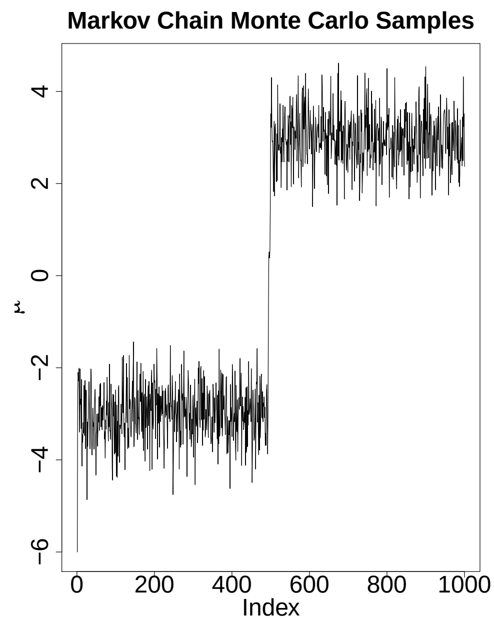


Figure 6: Traceplot of of μ from MCMC/Gibbs.

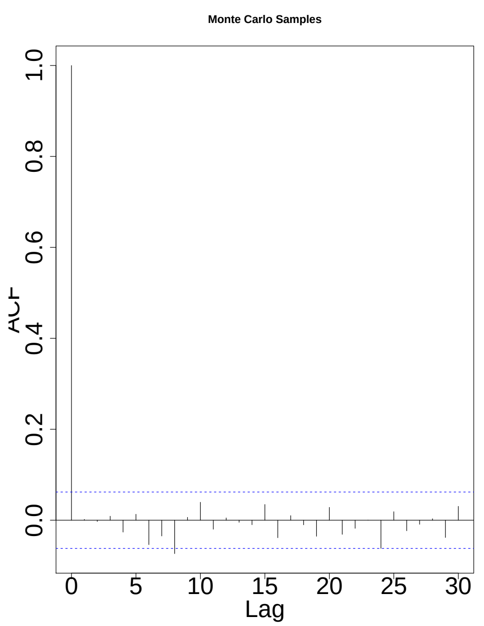


Figure 7: ACF plot of μ from Monte Carlo approximation.

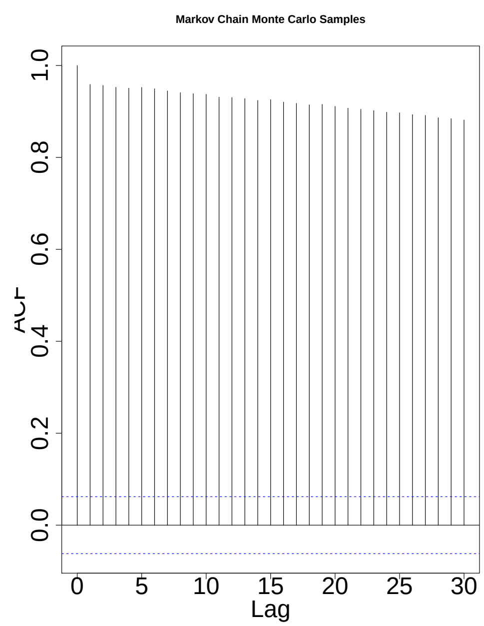


Figure 8: ACF plot of of μ from MCMC/Gibbs.