STAT4CI3/6CI3 Computational Methods for Inference

Assignment 3 — Due at 1:30pm on Monday, March 18, 2019

Instructions:

1. Ensure that all R code is properly commented and attach a print out with your written solution. Also mail your R code as a **single plain text file** to `cantya@mcmaster.ca` using the subject
`S4CI3 Assignment 3:` <*Name*> <*Student ID*>

2. Start each question on a new page and submit questions in the same order as given below.

3. You are expected to show all details of your solution and any results taken from my notes or the textbook must be clearly and properly referenced.

4. No extensions to the due date and time will be given except in extreme circumstances and late assignments will not be accepted.

5. Students are reminded that submitted assignments must be their own work. Submission of someone else's solution (including solutions from the internet or other sources) under your name is academic misconduct and will be dealt with as such. Penalties for academic misconduct can include a 0 for the assignment, an F for the course with an annotation on your transcript and/or dismissal from your program of study.

Q. 1 Suppose we wish to simulate from a gamma distribution with parameters $\alpha \geqslant 1$ and $\beta > 0$.

a) Show that if $\alpha = k$, where $k \geqslant 1$ is an integer, then the required random variable can be written as a sum of k exponential random variables and hence derive a way to generate from the gamma(k, β) distribution.

b) If α is not an integer, we can use the accept-reject method generating from the candidate gamma(k, b) distribution where k is taken to be the largest integer less than α. Show that the optimum value of b to use is $b = \alpha\beta/k$. Generate $10,000$ independent observations from the appropriate candidate density and $10,000$ uniform$(0,1)$ random variates and use these in an accept-reject algorithm to generate a sample from a Gamma distribution with shape parameter $\alpha = 3.2$ and scale parameter $\beta = 2$.

c) Using the same set of candidate gamma variables and the same set of uniforms, run an independence Metropolis-Hastings algorithm to generate a chain of $10,000$ observations from the Gamma distribution with $\alpha = 3.2$ and $\beta = 2$.
Generate the initial value x_0 from the correct target distribution using the `rgamma` *function in R so that the chain is guaranteed to be sampling from the correct target distribution at all iterations.*

d) Compare the two samples in terms of their acceptance probabilities and how well the resulting samples estimate the mean and variance of the Gamma$(\alpha = 3.2, \beta = 2)$ distribution using Monte Carlo.

Q. 2 In his important 1970 Biometrika paper, W. K. Hastings considered the problem of simulating from a standard normal distribution using a random walk Metropolis algorithm with a Uniform$(-\delta, \delta)$ error distribution.

a) Write an R function to return the observed chain for a given initial value $x^{(0)}$ and value of δ.

b) Starting at $x^{(0)} = 0$, take $\delta = 1$ and run the chain for $15,000$ iterations and discard the first $5,000$ iterations. Plot the rest of the chain in order, draw a histogram of the generated values superimposed with the standard normal density and examine the auto-correlation function of the chain.

c) Repeat part (b) for each value of $\delta \in \{0.1,\ 0.5,\ 2,\ 10\}$ and comment on how the performance of the algorithm changes with δ. Which of the 5 values of δ in this question would you suggest using?

Q. 3 The bivariate normal distribution has probability density function $f(x, y \mid \mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$ equal to

$$\frac{1}{2\pi\sigma_1\sigma_2\sqrt{(1-\rho^2)}} \exp\left\{-\frac{1}{2(1-\rho^2)}\left[\frac{(x-\mu_1)^2}{\sigma_1^2} - 2\rho\frac{(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2} + \frac{(y-\mu_2)^2}{\sigma_2^2}\right]\right\}$$

for any $(x, y) \in \mathbb{R}^2$.

a) Write an R function to generate from the bivariate normal distribution using a random walk Metropolis-Hastings algorithm where, at each iteration, the candidate vector is generated as two independent normal random variables with means given by current value of the bivariate chain and with variances given by σ_1^2 and σ_2^2.

b) Prove that

$$X \mid Y = y \ \sim \ \text{normal}\left(\mu_{X|Y=y} = \mu_1 + \frac{\rho\sigma_1}{\sigma_2}(y-\mu_2),\ \sigma^2_{X|Y=y} = (1-\rho^2)\sigma_1^2\right)$$

and deduce the distribution of $Y \mid X = x$.

c) Write an R function to implement the Gibbs sampler for simulating bivariate normal observations.

d) Implement both the Metropolis–Hastings and Gibbs Sampling algorithms to simulate samples of $N = 10,000$ random vectors from the bivariate normal with parameters

$$\mu_1 = -1,\ \mu_2 = 1,\ \sigma_1^2 = 1,\ \sigma_2^2 = 4,\ \rho = -0.5$$

Calculate the means, variances and correlation from each sample and compare them to the true values. Plot the marginal histograms and superimpose the true marginal densities.
You may assume without proof that marginally $X \sim$ normal(μ_1, σ_1^2) and $Y \sim$ normal(μ_2, σ_2^2)

Q. 4 Suppose that the random variable X comes from a binomial distribution with parameters n and θ, both unknown. We will conduct Bayesian inference with the following prior specifications

$$n \sim \text{Poisson}(\mu) \quad \text{independent of} \quad \theta \sim \text{Beta}(\alpha, \beta)$$

where μ, α and β are known.

a) Show that for a given value $X = x$ the posterior distribution satisfies

$$\pi(n, \theta \mid x) \propto \frac{\mu^n}{(n-x)!}\theta^{x+\alpha-1}(1-\theta)^{n-x+\beta-1}$$

b) Show that, for given θ, the conditional posterior distribution for n is of the form $x + Y$ where Y is a Poisson random variable. Give the mean of the random variable Y.

c) Show that, for given n, the conditional posterior distribution for θ is a Beta distribution and give the parameters of the distribution.

d) Write an `R` function to implement a Gibbs Sampler in this case and run your sampler for 10,000 iterations after an initial burn-in of 2000 iterations if we have a prior mean for n of $\mu = 16$ and $\alpha = 2$ $\beta = 4$ in the prior for θ and we observe $x = 2$. Plot the iterations of the Markov Chain and give the Monte Carlo estimates of the posterior means and variances as well as the posterior correlation between n and θ.

e) Now suppose that we do not actually observe X but are interested in its marginal distribution. The binomial model gives the full conditional for X given n and θ. Write and implement the Gibbs sampler with the same prior values of the distributions of n and θ but now considering the random vector (X, n, θ). Use the results of your chain to estimate the mean, median and variance of the random variable X. Have marginal means and variances and the correlation for n and θ changed when x is not observed?