

Lab 3: Gibbs sampler of a Beta-Binomial

Author: _____ (Insert your name here) _____

Total Grade for Lab 3: /10

Comments (optional)

Template for lab report

Instructions: This is the template you will use to type up your responses to the exercises. To produce a document that you can print out and turn in just click on Knit PDF above. All you need to do to complete the lab is to type up your BRIEF answers and the R code (when necessary) in the spaces provided below.

It is strongly recommended that you knit your document regularly (minimally after answering each exercise) for two reasons.

1. Ensure that there are no errors in your code that would prevent the document from knitting.
2. View the instructions and your answers in a more legible, attractive format.

```
# Any text BOTH preceded by a hashtag AND within the ```{r} ``` code chunk is a comment.
# R indicates a comment by turning the text green in the editor, and brown in the knitted
# document.
# Comments are not treated as a command to be interpreted by the computer.
# They normally (briefly!) describe the purpose of your command or chunk in plain English.
# However, for this class, they will have a different goal, as the text above and below
# each chunk should sufficiently describe the chunk's contents.
# For this class, comments will be used to indicate where your code should go, or to give
# hints for what the code should look like.
```

This lab is to practice replicating the results in Example 1 in Casella and George (1992).

Overview

A joint distribution of X and Y is given as

$$f(x, y) \propto \binom{n}{x} y^{x+\alpha-1} (1-y)^{n-x+\beta-1}, \quad x = -1, \dots, n, \quad 0 \leq y \leq 1. \quad (1)$$

Suppose we are interested in calculating some characteristics of the marginal distribution $f(x)$ of X . The Gibbs sampler allows us to generate a sample from this marginal. The full conditional posterior distribution of X and Y are given as

$$x \mid y, n, \alpha, \beta \sim \text{Binomial}(n, y), \quad (2)$$

$$y \mid x, n, \alpha, \beta \sim \text{Beta}(x + \alpha, n - x + \beta). \quad (3)$$

In addition, the marginal distribution of X can be obtained analytically, as

$$f(x) = \binom{n}{x} \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} \frac{\Gamma(x + \alpha)\Gamma(n - x + \beta)}{\Gamma(\alpha + \beta + n)}, x = 0, 1, \dots, n, \quad (4)$$

the Beta-Binomial distribution.

Simulate from the Beta-Binomial distribution

The **VGAM** R package can simulate draws from the Beta-Binomial distribution. Install this package and load the library.

```
install.packages("VGAM")
library(VGAM)
```

To generate a random draw from $X \sim \text{Beta-Binomial}(n, a, b)$:

```
rbetabinom.ab(1, n, a, b, .dontuse.prob = NULL)
```

Simulate from a Gibbs sampler

Following the full conditional posterior distributions for x and y in Equations (2) and (3), one can design a Gibbs sampler to generate draws of x and y iteratively.

Replicate results shown in Figure 1

According to the paper, “Figure 1 displays histograms of two samples, x_1, \dots, x_m of size $m = 500$ from the Beta-Binomial distribution with $n = 16, \alpha = 2$, and $\beta = 4$.”

Exercise 1: Use the `rbetabinom.ab` command to directly generate m draws of x , which follows a Beta-Binomial distribution as shown in Equation (4). Note that $m = 500, n = 16, \alpha = 2, \beta = 4$. Make a histogram of the m draws you have generated.

Grade for Exercise 1: /3

Comments:

According to the caption of Figure 1, “The black histogram sample was obtained using Gibbs sampling with $k = 10$.” Read Section 2 of the paper carefully to see how they implemented the Gibbs sampler.

Exercise 2: According to the Gibbs sampler implementation described in the paper, how many Gibbs samplers/MCMC chains do we need to run?

Grade for Exercise 2: /1

Comments:

Exercise 3: Write a Gibbs sampling scheme to generate m draws of x . Note that $m = 500, n = 16, \alpha = 2, \beta = 4$ and $k = 10$. Make a histogram of the m draws you have generated.

Grade for Exercise 3: /4

Comments:

Exercise 4: Make a side-by-side histogram showing the two samples, one generated in Exercise 1, and the other generated in Exercise 3. (Hint: you can use the following R script to do a side-by-side histogram. The demo below assumes `BetaBinomialDraws` storing the draws from Exercise 1 and `BetaBinomialDrawsGibbs` storing the draws from Exercise 3.)

```
install.packages("plotrix")
library(plotrix)

l <- list(BetaBinomialDraws, BetaBinomialDrawsGibbs)
multhist(l, breaks = seq(-0.5, 16.5, by=1))
```

Grade for Exercise 4: /2

Comments:

Remarks: Most likely you won't get exactly the same results as in the paper, but you will see a similar overall story. A good practice is to `set.seed()` to make your results reproducible.