Portfolio 8

Rachel Wood

2023-04-26

For this portfolio, we use the Pima Indians Diabetes dataset:

```
library(mlbench)
data("PimaIndiansDiabetes")
head(PimaIndiansDiabetes)
```

```
##   pregnant glucose pressure triceps insulin mass pedigree age diabetes
## 1        6     148       72      35       0 33.6    0.627  50      pos
## 2        1      85       66      29       0 26.6    0.351  31      neg
## 3        8     183       64       0       0 23.3    0.672  32      pos
## 4        1      89       66      23      94 28.1    0.167  21      neg
## 5        0     137       40      35     168 43.1    2.288  33      pos
## 6        5     116       74       0       0 25.6    0.201  30      neg
```

where we model the response y_i is the `diabetes` variable and model it using a logistic regression:

$$\mathbb{P}_{\alpha,\beta}(Y_i = 1) = \frac{1}{1 + e^{-\alpha - \beta^T x_i}}$$

which gives the likelihood funcion

$$L_n(\alpha, \beta) = \prod_{i=1}^{n} \mathbb{P}_{\alpha,\beta}(Y_i = y_i)$$

Then the posterior is

$$\pi(\alpha, \beta | y) \propto L_n(\alpha, \beta)\pi(\alpha, \beta)$$

where $\pi(\alpha, \beta)$ is the prior.

1. Choosing a proposal distribution

2. Implementing the MH algorithm

3. Convergence

4. Modifying Q

5. Marginal posterior distributions