Week 7 lab MAST90125: Bayesian Statistical learning

Question One

Consider a Poisson regression,

$$y_i \sim \text{Pois}(\lambda_i) \quad \text{and} \quad \log(\lambda_i) = \mathbf{x}_i'\boldsymbol{\beta}, \quad \boldsymbol{\beta} \in \mathbb{R}^p$$

In lectures we learned various techniques for approximating the posterior distribution. In this lab, attempt as many of these techniques as possible to complete the following tasks.

Consider the dataset `Warpbreaks.csv`, which can be downloaded from Canvas. This dataset contains information of the number of breaks in a consignment of wool. In addition, Wool type (A or B) and tension level (L, M or H) are recorded. To investigate the association between the number of breaks and wool type, various forms of generalised linear model are proposed where Bayesian computing techniques should be used.

As a reminder the following techniques will be considered for approximating the posterior distribution.

- Metropolis-Hastings algorithm.
- Gibbs sampler.

When coding, assume the prior for the coefficients $\boldsymbol{\beta} \sim N(\mathbf{0}, 5\mathbf{I}_p)$.

Some hints:

An initial guess can be determined from fitting a Poisson regression using the function `glm`. Treat wool type as a factor using the function `glm`

```
warpbreak= read.csv(file = './warpbreaks.csv',header=TRUE)
#This line will need to be changed when you run this yourself.
mod<-glm(breaks~as.factor(wool),data=warpbreak,family='poisson')
summary(mod)
```

```
##
## Call:
## glm(formula = breaks ~ as.factor(wool), family = "poisson", data = warpbreak)
##
## Coefficients:
##                   Estimate Std. Error z value Pr(>|z|)
## (Intercept)       3.43518    0.03454  99.443  < 2e-16 ***
## as.factor(wool)B -0.20599    0.05157  -3.994 6.49e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for poisson family taken to be 1)
##
##     Null deviance: 297.37  on 53  degrees of freedom
```

```
## Residual deviance: 281.33  on 52  degrees of freedom
## AIC: 560
##
## Number of Fisher Scoring iterations: 4
```

```
Sigma <-vcov(mod); Sigma
```

```
##                   (Intercept) as.factor(wool)B
## (Intercept)       0.001193293     -0.001193293
## as.factor(wool)B -0.001193293      0.002659566
```

```
X<-model.matrix(mod)
```