Inferencia Bayesiana

Alumno:
Huertas Quispe, Anthony Enrique ***Cod:*** *20173728*

Semestre: *2018-I*

Tema: PC 2

Prof. Victor Giancarlo Sal Y Rosas

Pontificia Universidad Católica del Perú
Escuela de Posgrado
Maestría en Estadística

Problema 1. Solución

Tenemos:

$$f(x,y) \quad \propto \quad \exp\left(-\left[\lambda x + \lambda y + \mu xy\right]\right), \quad x \geq 0, \quad y \geq 0 \tag{1}$$

a) Determinando distribuciones condicionales

$$\begin{aligned} f(x|Y=y) &\propto \exp\left(-\left[\lambda x + \lambda y + \mu xy\right]\right) \\ &\propto \exp\left(-\left[\lambda x + \mu xy\right]\right) \\ &\propto \exp\left(-x\left[\lambda + \mu y\right]\right) \\ \Rightarrow X|Y=y &\sim Exp\left(\lambda + \mu y\right) \end{aligned} \tag{2}$$

Análogamente obtenemos la distribución de $Y|X=x$

$$\begin{aligned} f(y|X=x) &\propto \exp\left(-\left[\lambda x + \lambda y + \mu xy\right]\right) \\ &\propto \exp\left(-\left[\lambda y + \mu xy\right]\right) \\ &\propto \exp\left(-y\left[\lambda + \mu x\right]\right) \\ \Rightarrow Y|X=x &\sim Exp\left(\lambda + \mu x\right) \end{aligned} \tag{3}$$

Listing 1: Algoritmo Gibbs Preg 1

```
##### PREGUNTA 1

I        <- 10000
PARA     <- matrix(0,nrow=I,2)
# Dos columnas: Primera x, segunda y

PARA[1,] <- c(5,5)
lambda <- 1
mu <- 1
cont=numeric(I)

for(i in 2:I){
        #
        # x
        #
        PARA[i,1] <- rexp(1,lambda+mu*PARA[i-1,2])
        #
        # y
        #
        PARA[i,2] <- rexp(1,lambda + mu*PARA[i,1])
        if(PARA[i,1]< 1 & PARA[i,2] < 1 ){cont[i]=1 }
}

x = PARA[1000:I,1]
y = PARA[1000:I,2]
hist(x)
hist(y)

#1.b
sim <- cbind(x,y)
colnames(sim)[1:2] <- c("x","y")
```

```
sim      <- as.mcmc(sim)
summary(sim)

#CONCLUSION:- La media de X es 0.68 aproximadamente
#            - La desviacon estandar de X es 0.7484


#1.c
p = sum(cont[1000:I])/(I-999)
p

#CONCLUSION:- La probabilidad de que X e Y sean menores que 1 es de
# aproximadamente 0.57
}
```

Pregunta 6. Solución

Se tiene una distribución para la v.a. Y como:

$$Y \sim t(\mu, \sigma^2, v) \tag{4}$$

donde $\mu > 0$, $\sigma^2 > 0$ y $v > 0$.

Además,

$$\begin{aligned} Y|W &\sim N(\mu, \sigma^2) & (5) \\ W &\sim Gamma\left(\frac{v}{2}, \frac{v}{2}\right) & (6) \end{aligned}$$

Tenemos $Y_1, \ldots Y_n$ condicionales a $\beta = (\beta_0, \beta_1)^T$ y τ tales que

$$Y_i|\beta.\tau \sim t(x_i^T\beta, \frac{1}{\tau}, 3) \tag{7}$$

y

$$p(\beta.\tau) \propto \frac{1}{\tau} \tag{8}$$

a) Distribución a posteriori aumentada de $\theta = (\beta, \tau)$ y W.

$$\begin{aligned} p(\beta.\tau, W|Datos) &\propto \prod_{i=1}^{n} p(y_i|\beta.\tau, w_i)p(w_i\beta.\tau) \times \frac{1}{\tau} \\ &\propto \prod_{i=1}^{n} (\tau w_i)^{1/2} \exp\left\{-\frac{\tau w_i}{2}\left(y_i - x_i^T\beta\right)^2\right\} & (9) \\ &\quad \times \prod_{i=1}^{n} w_i^{3/2-1} \exp\left\{-\frac{3}{2}w_i\right\} \times \frac{1}{\tau} & (10) \end{aligned}$$

b) Distribuciones condicionales:

$$\begin{aligned} p(w_i|\beta.\tau, Datos) &\propto w_i^{1/2} \exp\left\{-w_i\left(\frac{3}{2}+\frac{\tau}{2}\left(y_i - x_i^T\beta\right)^2\right)w_i^{1/2}\right\} \\ &\propto w_i \exp\left\{-w_i\left(\frac{3}{2}+\frac{\tau}{2}\left(y_i - x_i^T\beta\right)^2\right)\right\} \\ &\Rightarrow w_i \sim Gamma(2, \frac{3}{2}+\frac{\tau}{2}\left(y_i - x_i^T\beta\right)^2) \end{aligned} \tag{11}$$

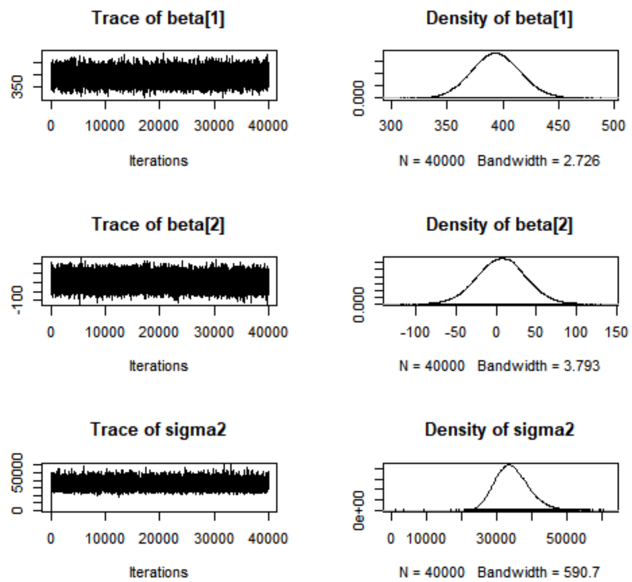
$$\begin{aligned} p(\beta|\tau, W, Datos) &\propto \exp\left\{-\frac{\tau}{2}\left[\beta^T XWX^T\beta - 2\beta^T XWY\right]\right\} \\ &\Rightarrow \beta|\tau, W, Datos \sim N_2\left((XWX^T)^{-1}XWY, \frac{(XWX^T)^{-1}}{\tau}\right) \end{aligned} \tag{12}$$

$$\begin{aligned} p(\tau|\beta, W, Datos) &\propto \tau^{n/2} \exp\left\{-\frac{\tau}{2}\sum_{i=1}^{n} w_i(y_i - x_i^T\beta)^2\right\}\frac{1}{\tau} \\ &\propto \tau^{n/2-1} \exp\left\{-\frac{\tau}{2}\sum_{i=1}^{n} w_i(y_i - x_i^T\beta)^2\right\} \\ &\Rightarrow \tau|\beta, W, Datos \sim Gamma(\frac{n}{2}, \frac{1}{2}\sum_{i=1}^{n} w_i(y_i - x_i^T\beta)^2) \end{aligned} \tag{13}$$

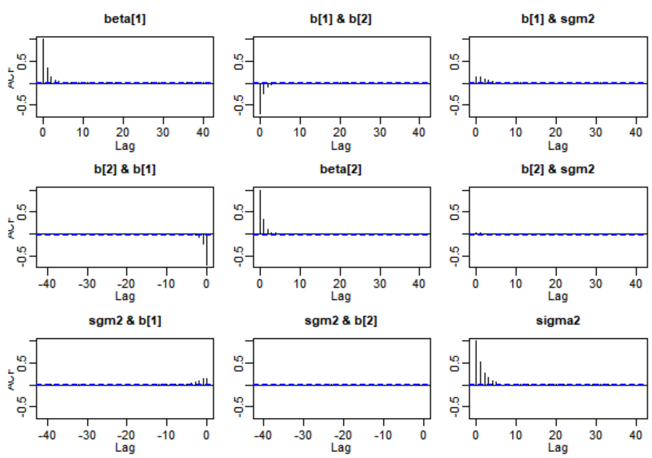
Listing 2: Algoritmo Gibbs Preg 6

```
######PREGUNTA 6

library(LearnBayes)

datos <- read.csv(file.choose())
attach(datos)

y <- datos$sRtimes
x <- datos$Mode

X <- cbind(1,x)
n <- length(y)

# Simulaciones
I <- 40000
beta     <- matrix(0,nrow=I,2)
tau      <- rep(0,I)

beta[1,] <- coef(lm(y~x))
tau[1]   <- 0.001

for(i in 2:I)
{
# W_i
alpha.w <- 2

```

```
eta    <- c(y-X%*%beta[i-1,])
beta.w <- 3/2 + 0.5*tau[i-1]*eta**2

W <- rgamma(n,alpha.w,beta.w)
#

# tau
tau[i] <- rgamma(1,n/2,0.5*sum(W*eta**2))
#
# beta
beta.hat <- coef(lm(y ~ x,weights=W,datos))
#beta.hat=solve(t(X)%*%diag(z)%*%X)%*%t(X)%*%diag(z)%*%y
V        <- solve(t(X)%*%diag(W)%*%X)/tau[i]
beta[i,] <- rmnorm(1,beta.hat,V)
}


#Preg6.c

sim               <- cbind(beta=beta,sigma2=1/tau)
colnames(sim)[1:2] <- c("beta[1]","beta[2]")
sim               <- as.mcmc(sim)


summary(sim)
plot(sim)
acf(sim)

# Preg6.d
w1 = rgamma(I,3/2,3/2)

mu=numeric(I)
for (i in 1:I){
mu[i] = mean(X%*%beta[i,])
}

y.pred=numeric(I)
y.pred <- rnorm(I,mu[i],sqrt(1/(w1*tau)))
hist(y.pred[1000:I])
plot(density(y.pred))
summary(y.pred)
```



Se determinaron los gráficos de distribución para cada parámetro. El gráfico de cadenas nos permite observar la estabilización de nuestros resutlados. Se observa posteriormente que la correlación se vuelve nula entre cada una de ellas.

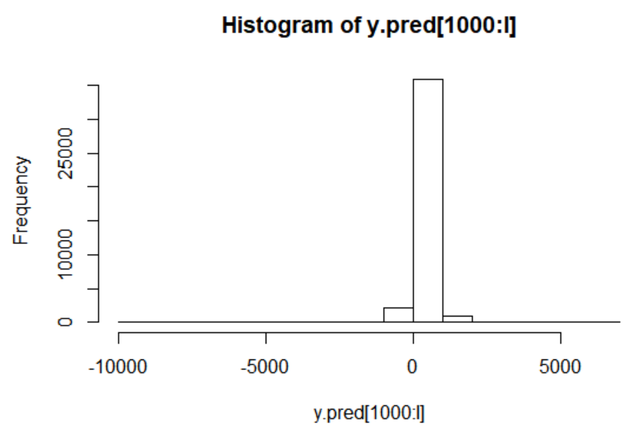


Se obtuvo el siguiente resumen estadístico

```
Iterations = 1:40000
Thinning interval = 1
Number of chains = 1
Sample size per chain = 40000

1. Empirical mean and standard deviation for each variable,
   plus standard error of the mean:

             Mean      SD Naive SE Time-series SE
beta[1]   394.022   21.32   0.1066         0.1541
beta[2]     6.624   30.14   0.1507         0.2143
sigma2  34415.918 4717.06  23.5853        43.0560

2. Quantiles for each variable:

            2.5%      25%       50%      75%   97.5%
beta[1]   352.75   379.61   393.731   408.21   436.9
beta[2]   -52.88   -13.54     6.734    26.75    65.6
sigma2  26205.13 31089.95 34076.382 37375.21 44610.6
```

Notamos que la varianza es un valor muy alto. Considerando el número de simulaciones hechas, iniciales, podemos decir que no existe una fluctuación errónea a causa del algoritmo.

Con respecto a la distribución predictiva se tiene lo siguiente:



Se tiene el siguiente resumen estadístico correspondiente

```
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
-9484.0   268.7   408.2   408.1   547.3  6448.8
```