# Examen Final - Parte 2

Anthony Enrique Huertas Quispe Cod:20173728
Prof: Cristian Bayes

## Pregunta 3

Una variable aleatoria $Y$ tiene distribución *Cauchy* si su función de densidad es dada por

$$f_Y(y) = \frac{1}{\pi\sigma}\left(1 + \left(\frac{y-\mu}{\sigma}\right)^2\right)^{-1} \tag{1}$$

donde $\mu \in \mathbb{R}$ y $\sigma^2 > 0$. Consideremos la siguiente notación $Y \sim Cauchy(\mu, \sigma^2)$ para representar una variable aleatoria que siga esta distribución. Esta distribución tiene la siguiente propiedad:
Sean $Y$ y $W$ dos variables aleatorias tales que

$$Y|W \sim N\left(\mu, \frac{\sigma^2}{W}\right) \tag{2}$$

$$W \sim Gamma\left(\frac{1}{2}, \frac{1}{2}\right) \tag{3}$$

entonces $Y \sim Cauchy(\mu, \sigma^2)$.
Asuma que $Y_1, Y_2, \ldots, Y_n$, condicional a $\mu$ y $\sigma^2$, son variables aleatorias independientes con distribución $Cauchy(\mu, \sigma^2)$.

a) A partir de la propiedad encuentre la distribución de $W_i$ condicional a $Y_i$.

*Solución.* Usando el Teorema de Bayes tenemos que:

$$\begin{aligned} f(w_i|y_i) &\propto f(y_i|w_i)f(w_i) \\ &\propto \frac{w_i^{1/2}}{\sqrt{2\pi}\sigma}\exp\left\{-\frac{1}{2}\left(\frac{y_i-\mu}{\sigma}\right)^2 w_i\right\}\frac{1/2^{1/2}}{\Gamma(1/2)}w_i^{1/2-1}\exp\{-w_i/2\} \\ &\propto \exp\left\{-\left(\frac{1}{2}+\frac{1}{2}\left(\frac{y_i-\mu}{\sigma}\right)^2\right)w_i\right\} \end{aligned} \tag{4}$$

Por tanto,

$$W_i|Y_i \sim Exp\left(\frac{1}{2}+\frac{1}{2}\left(\frac{y_i-\mu}{\sigma}\right)^2\right) \tag{5}$$

□

b) Implemente el algoritmo EM para el conjunto de datos darwin de la librería LearnBayes.

*Solución.* Primero determinemos la función verosimilitud

$$\begin{aligned} \mathcal{L}(\mu, \sigma^2 \mid W, y) &= \prod_{i=1}^{n} f(y_i|w_i)f(w_i) \\ &= \prod_{i=1}^{n}\frac{w_i^{1/2}}{\sqrt{2\pi}\sigma}\exp\left\{-\frac{1}{2}\left(\frac{y_i-\mu}{\sigma}\right)^2 w_i\right\}\frac{1/2^{1/2}}{\Gamma(1/2)}w_i^{1/2-1}\exp\{-w_i/2\} \\ &\propto \frac{1}{(\sigma^2)^{n/2}}\exp\left\{-\frac{1}{2}\sum_{i=1}^{n}\left(\frac{y_i-\mu}{\sigma}\right)^2 w_i\right\} \end{aligned} \tag{6}$$

luego la función log-verosimilitud vendría dada por

$$\ell(\mu, \sigma^2 \mid W, y) \quad \propto \quad -\frac{n}{2}\log(\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i - \mu)^2 w_i \tag{7}$$

Por (5) tenemos además que

$$E_{\mathbf{W}|\mathbf{Y}=y,\mu=\mu^{(j)},\sigma^2=\sigma^{2(j)}}[w_i] \quad = \quad \left(\frac{1}{2} + \frac{1}{2\sigma^{2(j)}}\left(y_i - \mu^{(j)}\right)^2\right)^{-1} \tag{8}$$

Por consiguiente el algoritmo EM, vendría determinado por los siguientes pasos:

- **Paso E**: Tomando $\theta = (\mu, \sigma^2)$.

$$\begin{aligned}
Q(\theta, \theta^{(j)}) &= E_{\mathbf{W}|\mathbf{Y}=\mathbf{y},\mu=\mu^{(j)},\sigma^2=\sigma^{2(j)}}[\ell(\mu, \sigma^2 \mid W, y)] \\
&= E_{\mathbf{W}|\mathbf{Y}=\mathbf{y},\mu=\mu^{(j)},\sigma^2=\sigma^{2(j)}}\left[-\frac{n}{2}\log(\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i - \mu)^2 w_i\right] \\
&= -\frac{n}{2}\log(\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i - \mu)^2 E_{\mathbf{W}|\mathbf{Y}=\mathbf{y},\mu=\mu^{(j)},\sigma^2=\sigma^{2(j)}}[w_i] \\
&= -\frac{n}{2}\log(\sigma^2) - \frac{1}{\sigma^2}\sum_{i=1}^{n}\frac{(y_i - \mu)^2}{1 + \frac{1}{\sigma^{2(j)}}\left(y_i - \mu^{(j)}\right)^2}
\end{aligned} \tag{9}$$

- **Paso M**:

$$\frac{\partial Q}{\partial \mu} \quad = \quad \frac{1}{\sigma^2}\sum_{i=1}^{n}\frac{2(y_i - \mu)}{1 + \frac{1}{\sigma^{2(j)}}\left(y_i - \mu^{(j)}\right)^2} = 0 \tag{10}$$

$$\frac{\partial Q}{\partial(\sigma^2)} \quad = \quad -\frac{n}{2\sigma^2} + \frac{1}{\sigma^4}\sum_{i=1}^{n}\frac{(y_i - \mu)^2}{1 + \frac{1}{\sigma^{2(j)}}\left(y_i - \mu^{(j)}\right)^2} = 0 \tag{11}$$

  Despejando $\mu$ y $\sigma^2$:

$$\mu^{(j+1)} = \frac{\sum_{i=1}^{n} y_i \hat{w}_i^{(j)}}{\sum_{i=1}^{n}\hat{w}_i^{(j)}} \quad \wedge \quad \sigma^{2(j+1)} = 2\sum_{i=1}^{n}\frac{\left(y_i - \mu^{(j+1)}\right)^2 \hat{w}_i^{(j)}}{n} \tag{12}$$

  donde $\hat{w}_i^{(j)} = \left(1 + \frac{1}{\sigma^{2(j)}}\left(y_i - \mu^{(j)}\right)^2\right)^{-1}$.

Listing 1: Algoritmo EM - Darwin (LearnBayes)

```r
library(LearnBayes)

y=data.frame(darwin)[,1]

# Algoritmo EM

res=matrix(c(mean(y),var(y)),1,2) #Inicializando valores

cont=0
h=0
while(cont==0){
        h=h+1
        theta.old=res[h,] #theta=(mu,sigma^2)
        w=(1+((y-theta.old[1])^2)/theta.old[2])^{-1}
        theta.new=c(0,0)
        theta.new[1]=sum(y*w)/sum(w)
        theta.new[2]=mean(((y-theta.new[1])^2)*w*2)
        res=rbind(res,theta.new)
        if(sum((theta.new-theta.old)^2)<1e-20){cont<-1}
}
res
```

Resultado:

| | $\mu$ | $\sigma^2$ |
|---|---|---|
| | 20.93333 | 1424.6381 |
| **theta.new** | 24.93518 | 778.5697 |
| **theta.new** | 25.89387 | 542.9205 |
| | $\vdots$ | $\vdots$ |
| **theta.new** | 24.97054 | 246.6743 |
| **theta.new** | 24.97054 | 246.6743 |

Habiéndose estimado los parámetros $\mu$ y $\sigma^2$, mediante el algoritmo EM, se infiere la distribución de $Y$ como sigue

$$Y \sim Cauchy(24.97054, 246.6743) \tag{13}$$

□

c) Mediante boostrap encuentre un intervalo de confianza del 95 % para los parámetros del modelo.

*Solución.* Haremos uso de la librería boot y de las funciones boot y boot.ci (*para intervalos de confianza*).

Listing 2: Intervalo de Confianza para la Media - Darwin (LearnBayes)

```
library(LearnBayes)
library(boot)

y=data.frame(darwin)[,1]

media=function(data,index){
        mean(data[index])
}
res=boot(y,media,1000) # Funcion Bootstrap

# Intervalo de Confianza Bootstrap
boot.ci(res,conf=0.95,type="all")
```

Resultado:

BOOTSTRAP CONFIDENCE INTERVAL CALCULATIONS
Based on 1000 bootstrap replicates
Intervals :

| Level | Normal | Basic | Percentile | BCa |
|---|---|---|---|---|
| 95 % | $(2.56, 39.05)$ | $(3.20, 39.93)$ | $(1.94, 38.67)$ | $(-1.77, 36.55)$ |

Listing 3: Intervalo de Confianza para la Varianza - Darwin (LearnBayes)

```
library(LearnBayes)
library(boot)

y=data.frame(darwin)[,1]

Var=function(data,index){
        var(data[index])
}
res2=boot(y,Var,1000) # Funcion Bootstrap

# Intervalo de Confianza Bootstrap
boot.ci(res2,conf=0.95,type="all")
```

Resultado:

BOOTSTRAP CONFIDENCE INTERVAL CALCULATIONS
Based on 1000 bootstrap replicates
Intervals :

| Level | Normal | Basic | Percentile | BCa |
|---|---|---|---|---|
| 95 % | $(368, 2622)$ | $(350, 2521)$ | $(328, 2500)$ | $(506, 3115)$ |

□

d) Considere que la distribución a *priori* de $\mu$ y $\sigma^2$ es dada por

$$p(\mu, \sigma^2) \propto \frac{1}{\sigma^2} \tag{14}$$

Halle la distribución a *posteriori* aumentada de $\theta = (\mu, \sigma^2)^T$ y $W = (w_1, \ldots, w_n)^T$ y encuentre las distribuciones condicionales completas de $\mu, \sigma^2$ y $w_i$.

*Solución.* En primera instancia la distribución a *posteriori* aumentada cumple lo siguiente

$$\begin{aligned}
p(\mu, \sigma^2, W \mid y) &\propto p(W, y \mid \mu, \sigma^2) p(\mu, \sigma^2) \\
&\propto \frac{1}{(\sigma^2)^{n/2}} \exp\left\{-\frac{1}{2}\left(\sum_{i=1}^{n}\left(\frac{y_i - \mu}{\sigma}\right)^2 w_i + \sum_{i=1}^{n} w_i\right)\right\} \frac{1}{\sigma^2} \\
&\propto \frac{1}{(\sigma^2)^{n/2+1}} \exp\left\{-\frac{1}{2}\left(\sum_{i=1}^{n}\left(\frac{y_i - \mu}{\sigma}\right)^2 w_i + \sum_{i=1}^{n} w_i\right)\right\}
\end{aligned} \tag{15}$$

Determinemos las siguientes distribuciones condicionales completas

- De $\mu$:

$$\begin{aligned}
p(\mu \mid y, \sigma^2, W) &\propto \exp\left\{-\frac{1}{2}\sum_{i=1}^{n}\left(\frac{y_i - \mu}{\sigma}\right)^2 w_i\right\} \\
&\propto \exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}\left(y_i^2 w_i + \mu^2 w_i - 2\mu y_i w_i\right)\right\} \\
&\propto \exp\left\{-\frac{1}{2\sigma^2}\left(\mu^2 \sum_{i=1}^{n} w_i - 2\mu \sum_{i=1}^{n} y_i w_i\right)\right\} \\
&\propto \exp\left\{-\frac{1}{2}\frac{\sum_{i=1}^{n} w_i}{\sigma^2}\left(\mu^2 - 2\mu\frac{\sum_{i=1}^{n} y_i w_i}{\sum_{i=1}^{n} w_i}\right)\right\} \\
&\propto \exp\left\{-\frac{1}{2}\frac{\sum_{i=1}^{n} w_i}{\sigma^2}\left(\mu - \frac{\sum_{i=1}^{n} y_i w_i}{\sum_{i=1}^{n} w_i}\right)^2\right\}
\end{aligned}$$

Por tanto,

$$\mu | y, \sigma^2, W \quad \sim \quad N\left(\frac{\sum_{i=1}^{n} y_i w_i}{\sum_{i=1}^{n} w_i}, \frac{\sigma^2}{\sum_{i=1}^{n} w_i}\right) \tag{16}$$

- De $\sigma^2$:

$$p(\sigma^2 \mid y, \mu, W) \quad \propto \quad \left(\frac{1}{\sigma^2}\right)^{n/2+1} \exp\left\{-\left(\frac{1}{\sigma^2}\right)\frac{1}{2}\sum_{i=1}^{n}(y_i - \mu)^2 w_i\right\}$$

Por tanto,

$$\frac{1}{\sigma^2} \mid y, \mu, W \quad \sim \quad Gamma\left(\frac{n}{2}, \frac{1}{2}\sum_{i=1}^{n}(y_i - \mu)^2 w_i\right) \tag{17}$$

- De $w_i$:

$$p(w_i \mid y, \sigma^2, w_1, \ldots, w_{i-1}, w_{i+1}, w_n) \quad \propto \quad \exp\left\{-\left(\frac{1}{2}\left(\frac{y_i - \mu}{\sigma}\right)^2 + \frac{1}{2}\right) w_i\right\}$$

Por tanto,

$$w_i \mid y, \mu, \sigma^2, w_1, \ldots, w_{i-1}, w_{i+1}, \ldots, w_n \quad \sim \quad Exp\left(\frac{1}{2} + \frac{1}{2\sigma^2}(y_i - \mu)^2\right) \tag{18}$$

$\square$

e) Implemente el algoritmo de Gibbs en R para el conjunto de datos darwin de la librería LearnBayes. Realice 10,000 simulaciones de la distribución a posteriori, presente gráficos de la cadena, autocorrelación, histogramas y estadísticas resumen. Comente sus resultados.

*Solución.* El algoritmo de Gibbs esta dado por:

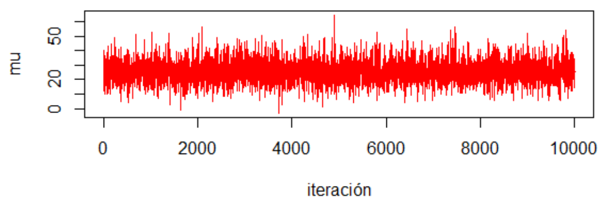
$$\mu^{(j+1)}|y,\sigma^{2(j)},W^{(j)} \sim N\left(\frac{\sum_{i=1}^n y_i w_i^{(j)}}{\sum_{i=1}^n w_i^{(j)}}, \frac{\sigma^{2(j)}}{\sum_{i=1}^n w_i^{(j)}}\right)$$
$$\frac{1}{\sigma^{2(j+1)}} \mid y,\mu^{(j+1)},W^{(j)} \sim Gamma\left(\frac{n}{2},\frac{1}{2}\sum_{i=1}^n \left(y_i-\mu^{(j+1)}\right)^2 w_i^{(j)}\right)$$
$$w_i^{(j+1)} \mid y,\mu^{(j+1)},\sigma^{2(j+1)},w_1^{(j+1)},\ldots,w_{i-1}^{(j+1)},w_{i+1}^{(j)},\ldots,w_n^{(j)} \sim Exp\left(\frac{1}{2}+\frac{1}{2\sigma^{2(j+1)}}\left(y_i-\mu^{(j+1)}\right)^2\right)$$

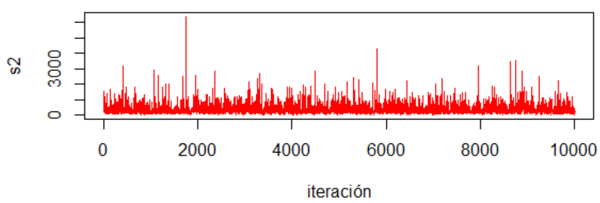
Listing 4: Algoritmo de Gibbs - Darwin (LearnBayes)

```
library(LearnBayes)

n=dim(darwin)[1]
y=data.frame(darwin)[,1]
mu=numeric()
sigma2=numeric()
W=matrix(0,1,n)

N=10000 #Tamaño de la simulación

#Valores iniciales
mu[1]=mean(y)
s2[1]=var(y)
W[1,]=rgamma(n,.5,.5)

#Algoritmo Gibbs
for(h in 1:N){
        mu[h+1]=rnorm(1,sum(y*W[h,])/sum(W[h,]),sqrt(s2[h]/sum(W[h,])))
        s2[h+1]=1/rgamma(1,0.5*n,0.5*sum(((y-mu[h+1])^2)*W[h,]))
        w=c(rep(0,n))
        for(j in 1:n){
                w[j]=rexp(1,0.5+(0.5/s2[h+1])*(y[j]-mu[h+1])^2)
        }
        W=rbind(W,w)
}
```

Listing 5: Algoritmo de Gibbs, Gráfico de cadenas - Darwin (LearnBayes)

```
par(mfrow=c(2,1))
ts.plot(mu,col=2,xlab="iteración")
ts.plot(s2,col=2,xlab="iteración")
```




**Comentarios:**
Se observa una correcta convergencia por parte del algoritmo. Sin embargo los demás detalles lo veremos a continuación.

**Programa WinBugs:** Haremos uso del programa WinBugs para facilitar un estudio más sofisticado, para ello tendremos que tener en cuenta que necesitaremos la función verosmilitud y función a priori, respectivamente, de la siguiente forma:

$$
\begin{aligned}
\mathcal{L}(\mu, \sigma^2, W \mid y) &= p(y \mid \mu, \sigma^2, W) \sim N(\mu, \frac{\sigma^2}{W}) \\
p(\mu, \sigma^2, W) &= p(W \mid \mu, \sigma^2) p(\mu, \sigma^2) = \underbrace{p(W)}_{\sim Gamma(.5,.5)} \quad \underbrace{p(\mu, \sigma^2)}_{\substack{\propto 1/\sigma^2 \\ \text{(no informativa)}}}
\end{aligned} \tag{19}
$$

Al tener una distribución a priori no informativa de $\mu$ y $\sigma$, dado que expresa información general pero no específica por cada variable, entonces consideramos lo siguiente:

$$
\begin{aligned}
\mu &\sim N(0, 100^2) \\
\sigma^2 &\sim Gamma - Inversa(0.01, 0.01)
\end{aligned}
$$

Listing 6: WinBugs, Algoritmo de Gibbs - Darwin (LearnBayes)

```
model{
# Posteriori
y[1]~dnorm(mu,tau1)
y[2]~dnorm(mu,tau2)
y[3]~dnorm(mu,tau3)
y[4]~dnorm(mu,tau4)
y[5]~dnorm(mu,tau5)
y[6]~dnorm(mu,tau6)
y[7]~dnorm(mu,tau7)
y[8]~dnorm(mu,tau8)
y[9]~dnorm(mu,tau9)
y[10]~dnorm(mu,tau10)
y[11]~dnorm(mu,tau11)
y[12]~dnorm(mu,tau12)
y[13]~dnorm(mu,tau13)
y[14]~dnorm(mu,tau14)
y[15]~dnorm(mu,tau15)

# Priori
# Por ser distribuciones no informativas
mu~dnorm(0,0.0001)
tau~dgamma(0.01,0.01)

s2<-1/tau
for(i in 1:n){
w[i]~dgamma(0.5,0.5)
}
tau1<-w[1]/s2
tau2<-w[2]/s2
tau3<-w[3]/s2
tau4<-w[4]/s2
tau5<-w[5]/s2
tau6<-w[6]/s2
tau7<-w[7]/s2
tau8<-w[8]/s2
tau9<-w[9]/s2
tau10<-w[10]/s2
tau11<-w[11]/s2
tau12<-w[12]/s2
tau13<-w[13]/s2
tau14<-w[14]/s2
tau15<-w[15]/s2
}

list(
n=15,
y=c(49,67,8,16,6,23,28,41,14,29,56,24,75,60,-48)
)
```

Obtenemos los siguientes resultados:

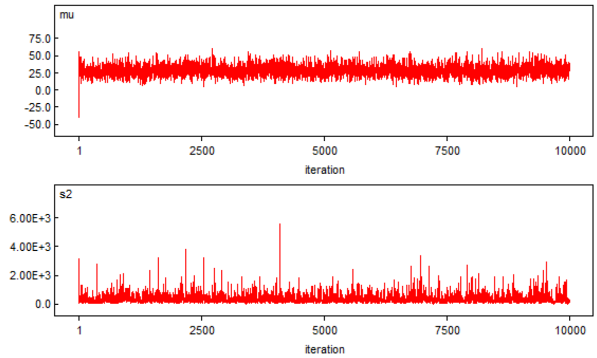



Figura 1: Gráfico de cadenas

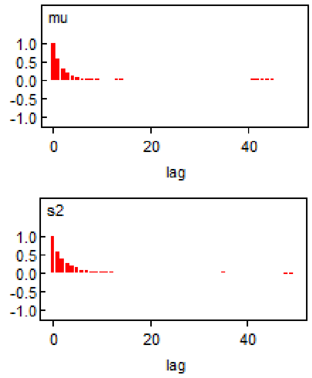



Figura 2: Autocorrelación

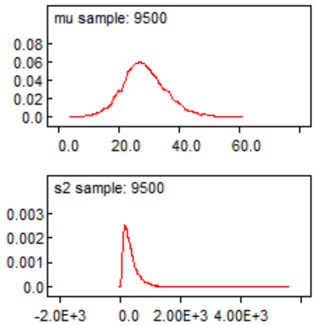



Figura 3: Histogramas

| node | mean | sd | MC error | 5.0% | 95.0% | start | sample |
|---|---|---|---|---|---|---|---|
| mu | 28.45 | 7.342 | 0.1648 | 17.17 | 41.25 | 501 | 9500 |
| s2 | 357.8 | 286.7 | 6.534 | 83.25 | 871.4 | 501 | 9500 |

Figura 4: Estadísticas Resumen

- El gráfico de cadenas nos informa una correcta convergencia del algoritmo, tal como se había observado usando R.
- El gráfico de autocorrelación es adecuado pues nos informa una grado de autocorrelación prácticamente nula.
- El gráfico de histogramas nos muestra una distribución, podría decirse normal para $\mu$ y log-normal para $\sigma^2$.
- Se tiene las siguientes estimaciones de los parámetros

| | $\mu$ | $\sigma^2$ |
|---|---|---|
| **Estimador Puntual** | 28.45 | 357.8 |
| **Intervalo de confianza 90 %** | (17.17,41.21) | (83.25,871.4) |

También podemos implementar un código en Listing 4, para determinar los estimadores puntuales de los parámetros y sus intervalos de confianza

Listing 7: Algoritmo de Gibbs, Intervalo de confianza 90 % - Darwin (LearnBayes)

```
mu=mu[1001:10000] #Eliminando las primeras 1000 iteraciones
s2=s2[1001:10000] #Eliminando las primeras 1000 iteraciones

# Estimadores puntuales

# Estimadores por Intervalo
quantile(mu,probs=c(0.05,0.95))
quantile(s2,probs=c(0.05,0.95))
```

| | $\mu$ | $\sigma^2$ |
|---|---|---|
| **Estimador Puntual** | 25.95765 | 365.1222 |
| **Intervalo de confianza 90 %** | (15.06591,38.25318) | (84.48391,931.79569) |

□

f) Encuentre un intervalo de credibilidad al 90 % para una nueva observación.

*Solución.* En este caso tendremos que evaluar una nueva observacion sobre la distribución de Cauchy con los parámetros puntuales estimados. Haremos uso de los resultados obtenidos por el algoritmo EM (Listing 1) algoritmo de Gibs (usando Listing 4 y Listing 7).

Listing 8: Algoritmo de Gibbs, Intervalo de confianza 90 % - Darwin (LearnBayes)

```
# Para el algoritmo EM
quantile(rcauchy(N−1000,res[dim(res)[1],1],sqrt(res[dim(res)[1],2])),probs=c
    (0.05,0.95))

# Para el algoritmo de Gibbs
quantile(rcauchy(N−1000,mu.est,sqrt(s2.est)),probs=c(0.05,0.95))
```

| Intervalo de confianza al 90 % | Nueva observación $y$ |
|---|---|
| Algoritmo EM | (-71.87799,124.68244) |
| Algoritmo de Gibbs | (-89.5779,140.1819) |

□