



**DEPARTAMENTO
DE COMPUTACION**

Facultad de Ciencias Exactas y Naturales - UBA

Trabajo Práctico

26 de Mayo de 2016

Inferencia Bayesiana

Integrante	LU	Correo electrónico
Costa, Manuel José Joaquín	035/14	manucos94@gmail.com
Gatti, Mathias Nicolás	477/14	mathigatti@gmail.com



Facultad de Ciencias Exactas y Naturales
Universidad de Buenos Aires

Ciudad Universitaria - (Pabellón I/Planta Baja)

Intendente Güiraldes 2160 - C1428EGA

Ciudad Autónoma de Buenos Aires - Rep. Argentina

Tel/Fax: (54 11) 4576-3359

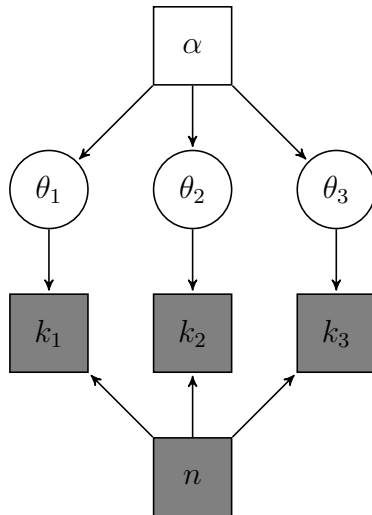
<http://www.fcen.uba.ar>

Índice

1. Ejercicio 1	3
2. Ejercicio 2	4
2.1. Histogramas	5
2.2. Correlaciones	6
2.3. Media y Desvío Standard	7
2.4. ¿Esta cargada?	8
3. Ejercicio 3	9
4. Ejercicio 4	10

1. Ejercicio 1

El modelo que escogimos se describe a continuación, en este entran en juego distintos parametros de distribuciones que modelan el problema y datos obtenidos que nos ayudaran a predecir mejor las distribuciones de los parámetros. Los cálculos de inferencia se realizaran con JAGS. Mas adelante se describe el código utilizado en el código del modelo.



$$\begin{aligned} \beta &\sim \text{Categorical}(p_i) \\ \theta_1 &\sim \begin{cases} \text{Beta}(0,5,0,5) & \text{if } \beta = 1 \\ \text{Beta}(100,100) & \text{if } \beta \neq 1 \end{cases} \\ k_1 &\sim \text{Binomial}(\theta_1, n) \\ k_2 &\sim \text{Binomial}(\theta_2, n) \\ k_3 &\sim \text{Binomial}(\theta_3, n) \\ n &= 10 \end{aligned}$$

```
model{
  # Observed Counts
  k1 ~ dbin(theta1,n)
  k2 ~ dbin(theta2,n)
  k3 ~ dbin(theta3,n)

  # Prior on Rates Theta
  theta1 ~ dbeta(param1, param1)
  theta2 ~ dbeta(param2, param2)
  theta3 ~ dbeta(param3, param3)

  # Auxiliary variables for Theta's distribution
  param1 <- ifelse(alpha== 1, 0.5, 100)
  param2 <- ifelse(alpha== 2, 0.5, 100)
  param3 <- ifelse(alpha== 3, 0.5, 100)

  # Prior on Rate Alpha
  for( i in 1:3){
    pi[i] <- 1/3
  }
  alpha ~ dcat(pi[])
}
```

2. Ejercicio 2

Como valores iniciales de las variables no observadas escogimos distintas opciones para asegurarnos de que no haya diferencia en los valores a los cuales convergen. Para los gráficos expuestos se utilizaron los siguientes parámetros para el algoritmo de muestreo.

```
% Sampling
%MCMC Parameters
nchains = 2; %How Many Chains?
nburnin = 10e2; %How Many Burn-in Samples?
nsamples = 10e4; %How Many Recorded Samples?
nthin = 1; %How Often is a Sample Recorded?
doparallel = 0; % Parallel Option

% Assign Matlab Variables to the Observed Nodes
datastruct = struct('k1',k1,'k2',k2,'k3',k3,'n',n);

% Initialize Unobserved Variables
for i=1:nchains
    S.theta1 = 0.5; % Intial Value
    S.theta2 = 0.5; % Intial Value
    S.theta3 = 0.5; % Intial Value
    S.alpha = randi([1 3]); % Intial Value
    init0(i) = S;
end
```

2.1. Histogramas

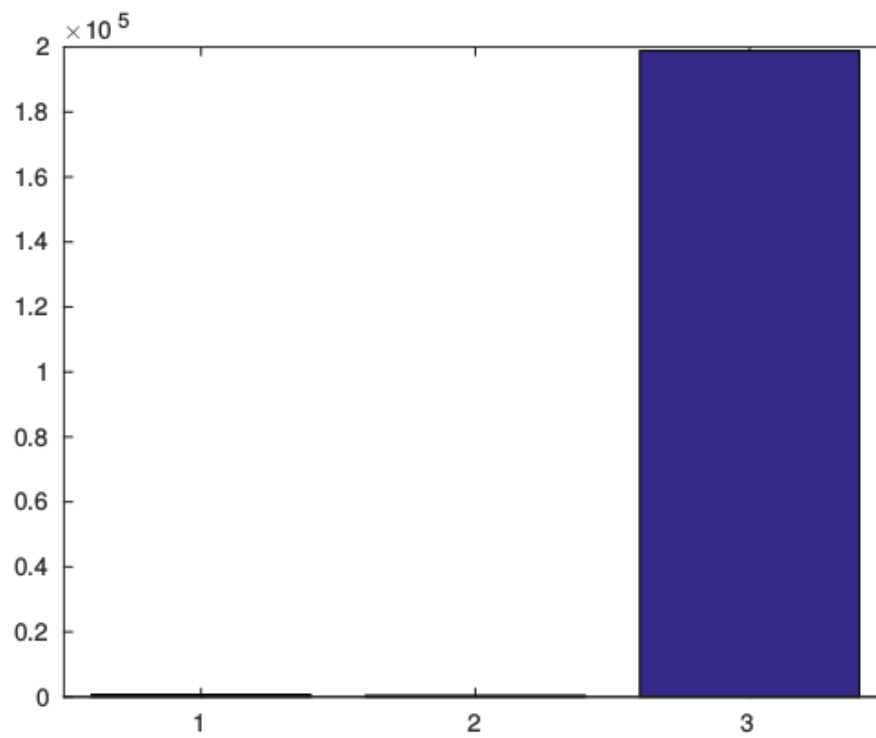


Figura 1: Histograma de la variable α .

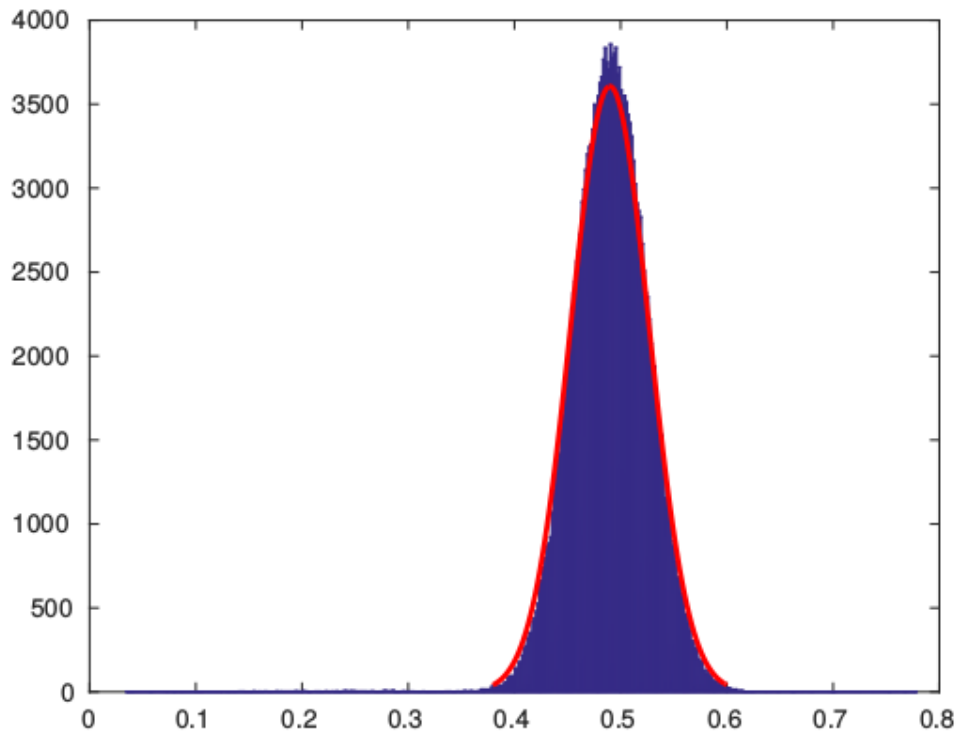


Figura 2: Histograma de la variable θ_1 . La linea azul es una aproximacion con una distribución normal.

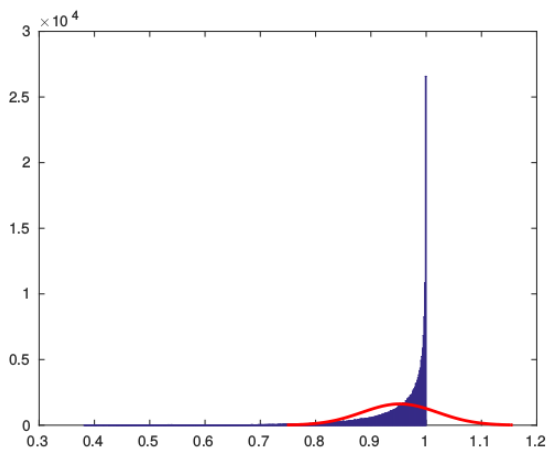


Figura 3: Histograma de la variable θ_3 . La linea azul es una aproximacion (Poco precisa como se podrá observar) con una distribución normal.

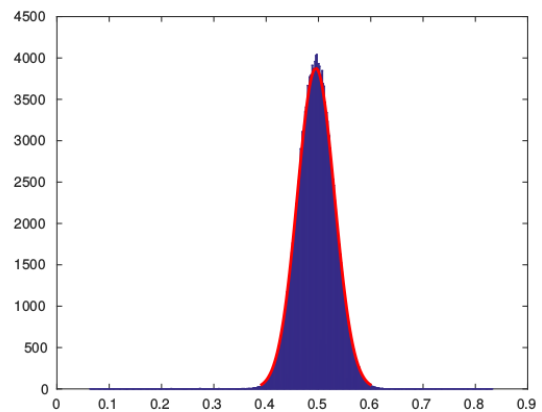
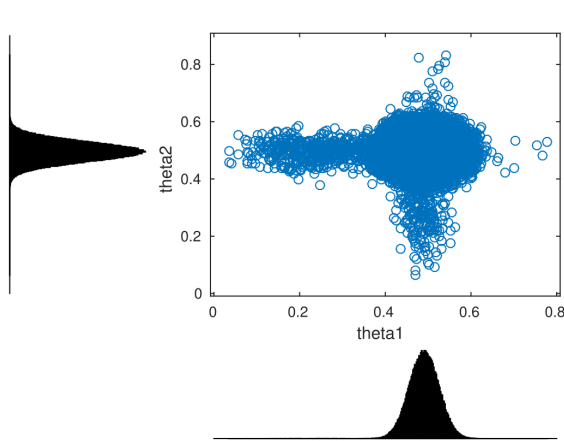
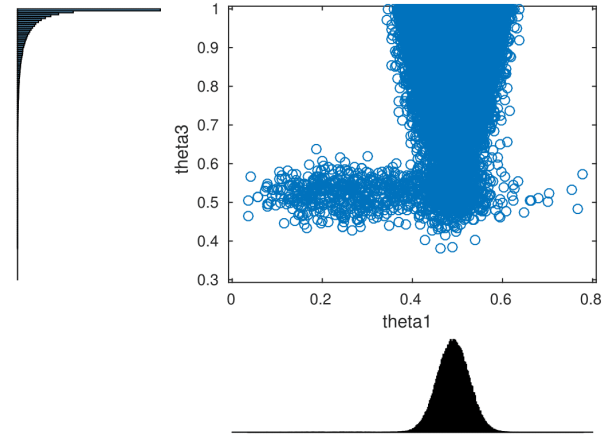
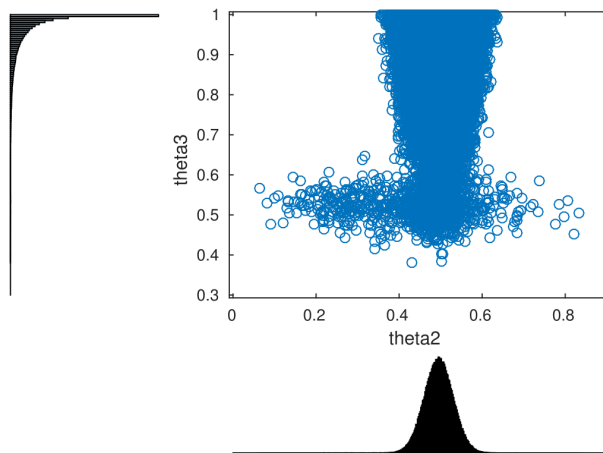


Figura 4: Histograma de la variable θ_2 . La linea azul es una aproximacion con una distribución normal.

2.2. Correlaciones

Creimos que podía ser interesante ver el histograma conjunto entre las distribuciones de las monedas. Con esto pudimos visualizar como si en una distribución en cierto punto la moneda estaba cargada en la distribución de la otra moneda esta no lo estaba, esta situación genera los graficos formados alrededor del par de rectas que pasan por el 0.5.

Figura 5: Correlación entre θ_1 y θ_2 .Figura 6: Correlación entre θ_1 y θ_3 .Figura 7: Correlación entre θ_2 y θ_3 .

2.3. Media y Desvío Standard

En los resultados siguientes se puede ver un poco mas lo que esperabamos y habiamos verificado con los gráficos. Por ejemplo la variable α tiene una esperanza muy cercana a 3, ya que luego de hacer la inferencia con los datos obtenidos todo indica que la moneda número 3 es la cargada.

$$\mathbb{E}(\alpha) = 2,9907$$

$$\sigma(\alpha) = 0,1270$$

$$\mathbb{E}(\theta_1) = 0,4899$$

$$\sigma(\theta_1) = 0,0368$$

$$\mathbb{E}(\theta_2) = 0,4951$$

$$\sigma(\theta_2) = 0,0354$$

$$\mathbb{E}(\theta_3) = 0,9523$$

$$\sigma(\theta_3) = 0,0680$$

Las 2 primeras monedas parecen no estar cargadas con bastante seguridad ya que tienen una medio de casi 0.5 con un desvío standard muy bajo. Esta situación creemos que es inducida por los datos obtenidos y la forma en que es armado el modelo donde solo una de las 3 monedas puede estar cargada.

2.4. ¿Esta cargada?

Para calcular la probabilidad de que cada moneda este cargada, tomamos los 200000 samples de α generados por el algoritmo y calculamos la probabilidad de que salga un 1, 2 o 3. El resultado es el siguiente.

$$P(\alpha = 1) = \frac{694}{200000} = 0.0035$$

$$P(\alpha = 2) = \frac{468}{200000} = 0.0023$$

$$P(\alpha = 3) = \frac{198838}{200000} = 0.9942$$

Algo que nos sorprendió fue que según estos calculos la moneda 1 tiene mas chances de estar cargada que la 2 a pesar de que según los datos esta salió una menor cantidad de veces cara. Quedará como trabajo a futuro verificar si esto fue por el posible error en la aproximación del algoritmo o algun otro aspecto que no tuvimos en cuenta.

3. Ejercicio 3

Esencialmente la diferencia va a tener que estar al nivel de los hiperparámetros, pues antes cada moneda no era indistinta a que las otras estuvieran cargadas o no, pero en este caso sí lo son. Por lo tanto, en lugar de tener un solo α , ahora tendremos tres α_i , con distribución Bernoulli de probabilidad $\frac{1}{2}$ cada una como *prior*. En la figura 8 puede verse el nuevo modelo y la representación gráfica del mismo usando DAGs.

```
model{
  # Observed Counts
  k1 ~ dbin(theta1,n)
  k2 ~ dbin(theta2,n)
  k3 ~ dbin(theta3,n)

  # Prior on Rates Theta
  theta1 ~ dbeta(param1, param1)
  theta2 ~ dbeta(param2, param2)
  theta3 ~ dbeta(param3, param3)

  # Auxiliary variables
  for Theta's distribution
  param1 <- ifelse(alpha1=1, 0.5, 100)
  param2 <- ifelse(alpha2=2, 0.5, 100)
  param3 <- ifelse(alpha3=3, 0.5, 100)

  # Prior on Rates Alpha
  alpha1 ~ dbern(0.5)
  alpha2 ~ dbern(0.5)
  alpha3 ~ dbern(0.5)
}
```

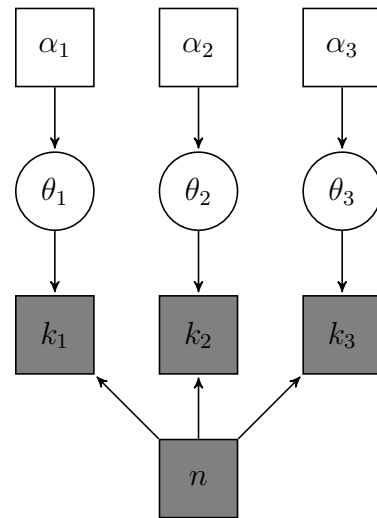


Figura 8: El modelo propuesto junto con su representación gráfica.

4. Ejercicio 4

$$\begin{aligned}
 P(cara_i|D) &= \int P(cara_i, \theta_i|D)d\theta_i \\
 &= \int P(cara_i|D, \theta_i)P(\theta_i|D)d\theta_i \\
 &= \int P(cara_i|\theta_i)P(\theta_i|D)d\theta_i \\
 &= \int \theta_i \sum_j P(\theta_i, \alpha = j|D)d\theta_i \\
 &\propto \int \theta_i \sum_j P(D|\theta_i, \alpha = j)P(\theta_i, \alpha = j)d\theta_i \\
 &= \int \theta_i \sum_j P(D|\theta_i)P(\theta_i|\alpha = j)P(\alpha = j)d\theta_i
 \end{aligned} \tag{1}$$