Tarea 34

Simulación estocástica

Marco Antonio Andrade Barrera

24 de mayo de 2018

Recordando nuestro ejmeplo donde

$$P(\boldsymbol{y}|\theta) = Mult_4(\boldsymbol{y}|1/2, \theta/4, \frac{1}{4}\frac{1}{1-\theta}, \frac{1}{4}\frac{1}{1-\theta}, \theta/4)$$

Donde $\boldsymbol{x} = (x_1 = 125, x_2 = 18, x_3 = 20, x_4 = 34)$, $\boldsymbol{y} = (z_1, z_2, x_2, x_3, x_4)$ y $z_1 + z_2 = 125$.

Supongamos (como $\theta \in (0,1)$) que $P(\theta) = U(\theta|0,1)$

$\implies P(\theta|\boldsymbol{y}) \propto \theta^{z_2+x_4}$ y $P(z_2|x_1,\theta) = Bin(z_2|x_1, \frac{\theta}{\theta+2})$

$\implies P(\theta|\boldsymbol{y}) \propto Be(\theta|z_2 + x_4 + 1, x_2 + x_3 + 1)$

De lo anterior, se propone el siguiente algoritmo para generar valores de la distribución de θ:

- $\theta^0 \sim U(\theta|0,1)$
- $z_2^{(t)} \sim bin(z_2|125, \frac{\theta^{(t-1)}}{2+\theta^{(t-1)}})$
- $\theta^{(t)} \sim Be(\theta|z_2^{(t)} + 35, 39)$

Para probar el algoritmo, se generarán $r = 1,327$ valores de θ y se generará un histograma. Posteriormente se comparará el histograma obtenido con una aproximación de $P(\theta|\boldsymbol{y}) = \frac{1}{r}\sum_{j=1}^{r} Be(\theta|z_2^{(j)} + x_4 + 1, x_2 + x_3 + 1)$.

En el siguiente bloque de código se implementa el algoritmo dado un θ_0 y un número r de valores a generar.

```
iter <- function(theta0,r){
  z <- list()
  theta <- list()
  for(j in 1:r){
    z2 <- rbinom(1,125,theta0/(2+theta0))
    theta0 <- rbeta(1,z2+35,39)
    z[[j]] <- z2
    theta[[j]] <- theta0
  }
  data.frame(z = unlist(z),theta = unlist(theta))
}
```

En el siguiente bloque se fija una semilla y se generan los 1327 parejas (z_2, θ) solicitadas. Además se muestran sólo algunos de los valores generados.

```
set.seed(28)
dta = iter(theta0 = runif(n = 1),r = 1327)
tail(dta)
```

```
##       z     theta
## 1322 31 0.6770850
## 1323 28 0.5441789
## 1324 20 0.6495390
## 1325 28 0.6044722
## 1326 25 0.6070561
## 1327 27 0.6126381
```

En seguida se define la función con la que se aproximará $P(\theta|y)$, continua.

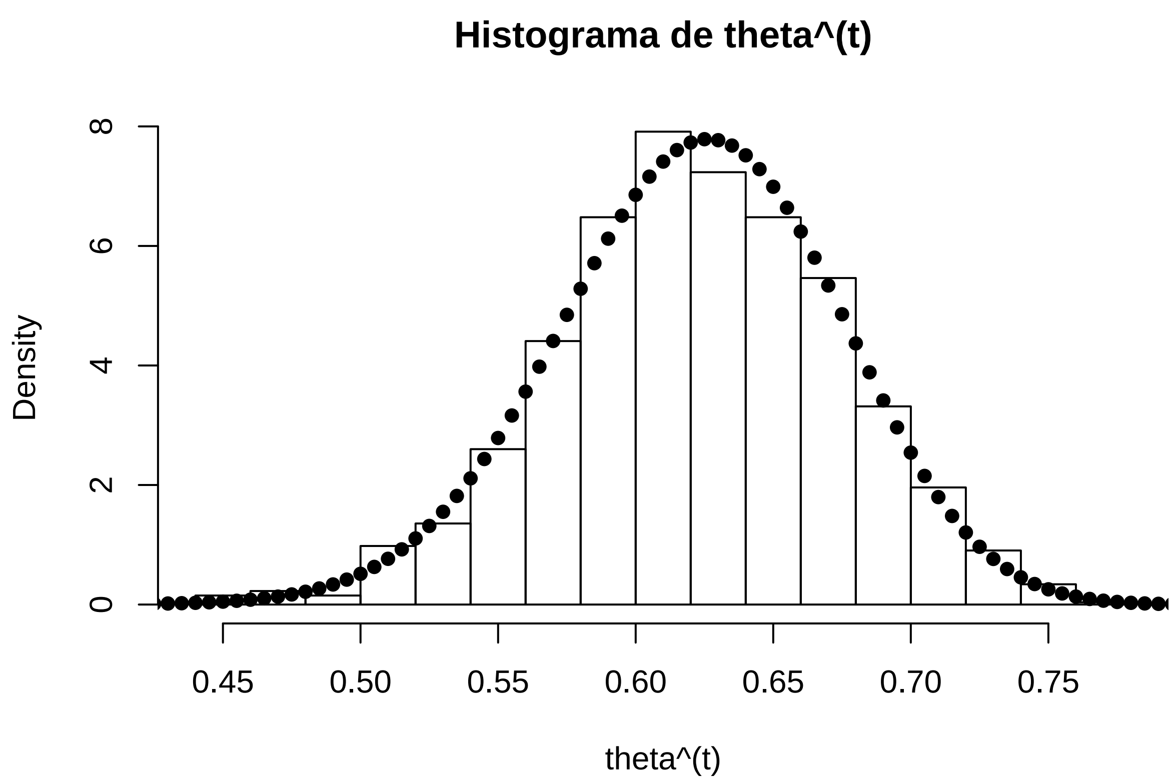
```
p <- function(z,theta){
  mean(unlist(lapply(z, function(z2) dbeta(theta,z2+35,39) )))
}
```

Evaluamos la función anterior en el intervalo en θ para el intervalo $(0, 1)$ con pasos de 0.005.

```
ptheta <- unlist(lapply(seq(0,1,0.005), FUN = function(theta){
  p(dta$z,theta)
}))
```

Finalmente, se muestra un histograma de los 1327 valores generados de θ y se compara con la distribución continua estimada mediante el promedio de betas, los puntos en la gráfica corresponden a esta última distribución aproximada.

```
hist(dta$theta,freq = FALSE,main = "Histograma de theta^(t)",xlab = "theta^(t)")
points(x = seq(0,1,0.005),ptheta,pch=16)
```



Claramente se observa que ambas distribuciones prácticamente coinciden.