Simulación de Montecarlo para la distribución Beta-Binomial

Para completar el contenido de la página 121 de *El detective Salazar y los modelos bayesianos*.

El teorema de Bayes permite actualizar el modelo al incluir la nueva información disponible. Supongamos que nuestro modelo (parámetro) inicial θ sigue una distribución *Uniforme*$(0, 1)$:

$$\theta \sim \textit{Uniforme}(0, 1), \quad \text{a priori.}$$

Recibimos nueva información que indica que de 57 intentos, 19 han sido un éxito. La distribución de esta nueva información sigue una Binomial$(57, \theta)$.

$$X \sim \textit{Binomial}(57, \theta), \quad \text{verosimilitud con } x = 19.$$

Si aplicamos el teorema de Bayes para actualizar nuestro modelo sobre θ,

$$\Pr(\theta \mid X = 19) = \frac{\Pr(X = 19 \mid \theta) \Pr(\theta)}{\Pr(X = 19)},$$

resulta que el nuevo modelo incorpora la información y obtenemos que la distribución sigue una Beta,

$$\theta \mid X = 19 \sim \textit{Beta}(20, 39), \text{a posteriori,}$$

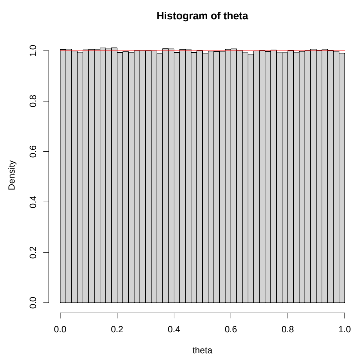
cuya media es 0.3389 y desviación típica 0.0611.

Método de Montecarlo. Simulemos la nueva distribución empleando el método de Montecarlo. Generamos n valores de θ que siguen una distribución Uniforme y para cada uno de ellos generamos un valor x que siga una distribución Binomial$(57, \theta)$.

```r
set.seed(123)
n <- 1e6 # Tamaño de la simulación
theta <- runif(n, 0, 1) # n valores de theta Uniforme(0,1)
```

La densidad de la muestra de θ resulta muy parecida a la densidad teórica de una distribución Uniforme$(0, 1)$.

```r
hist(theta, breaks = 50, prob = TRUE) # distrib. Uniforme
curve(dunif(x, 0, 1), add = TRUE, col = "red") # Teórica
```



Con estos valores de $\theta = \{\theta_1, \theta_2, \ldots, \theta_n\}$, simulamos un valor para cada una de las n distribuciones Binomiales: $x_i \sim \textit{Binomial}(57, \theta_i)$, $i = 1, 2, \ldots, n$.

```r
## Para cada valor de theta, generamos un valor de x
x <- rbinom(n, 57, theta)
```

Los valores x generados varían desde 0 hasta 57. Como la información que disponemos indica que $x = 19$, nos quedamos con los θ_i que han generado ese valor en la simulación de la Binomial.

```r
## Como nos dicen que hemos obtenido 19 éxitos
filtro <- x == 19 # buscamos los theta que lo han obtenido.
thetadadox <- theta[filtro] # los theta con 19 éxitos
```

Comprobamos que la distribución empírica de estos $\theta_i \mid x_i = 19$ se asemeja a la distribución teórica $\text{Beta}(20, 39)$, cotejando la media, desviación típica y una representación gráfica.

```r
## Analizamos y comparamos con la distribución téorica
c(media = mean(thetadadox), sd = sd(thetadadox))

##      media         sd
## 0.33918269 0.06141588

hist(thetadadox,
  breaks = 50, prob = TRUE,
  main = "theta cuando X=19", xlab = "theta"
)
curve(dbeta(x, 20, 39), add = TRUE, col = "red") # Teórica
```