Distribución Uniforme y Teorema Central del Límite

Luz Alba Posse & Martina Monastra

2024-08-26

Ejercicio 1: Distribución Uniforme

Sea $X \sim U(0, 18)$ una variable distribuida uniformemente entre 0 y 18.

1.a Generar la función X_dist(R)

```
set.seed(6803)

# 1.a Generar la función X_dist(R)
X_dist <- function(R) {
  runif(R, min = 0, max = 18)
}
```

1.b Calcular la media y la varianza muestral de los datos para $R \in \{2, 30, 100, 10^4\}$.

```
R_values <- c(2, 30, 100, 10^4)
resultados <- data.frame(R = integer(), Media = numeric(), Varianza = numeric())

for (R in R_values) {
  muestras <- X_dist(R)
  med_m <- mean(muestras)
  var_mue <- var(muestras)
  resultados <- rbind(resultados, data.frame(R = R, Media = med_m, Varianza = var_mue))
}

print(resultados)
```

```
##       R    Media  Varianza
## 1     2 3.377748  2.501266
## 2    30 8.270044 28.293709
## 3   100 8.220039 22.866303
## 4 10000 9.014883 27.267318
```

1.c Calcular el valor teórico de la esperanza de X, $E(X)$, y su varianza, $V(X)$. Comparar estos valores con los obtenidos en 1.b.

Teóricamente, para una distribución uniforme $U(a, b)$, la esperanza $E(X)$ y la varianza $V(X)$ se calculan como:

$$E(X) = \frac{a+b}{2}, \quad V(X) = \frac{(b-a)^2}{12}$$

En nuestro caso:

```r
a <- 0
b <- 18
E_X <- (a + b) / 2
V_X <- (b - a)^2 / 12

print(paste("Esperanza E(X):", E_X))
```

```
## [1] "Esperanza E(X): 9"
```

```r
print(paste("Varianza V(X):", V_X))
```

```
## [1] "Varianza V(X): 27"
```

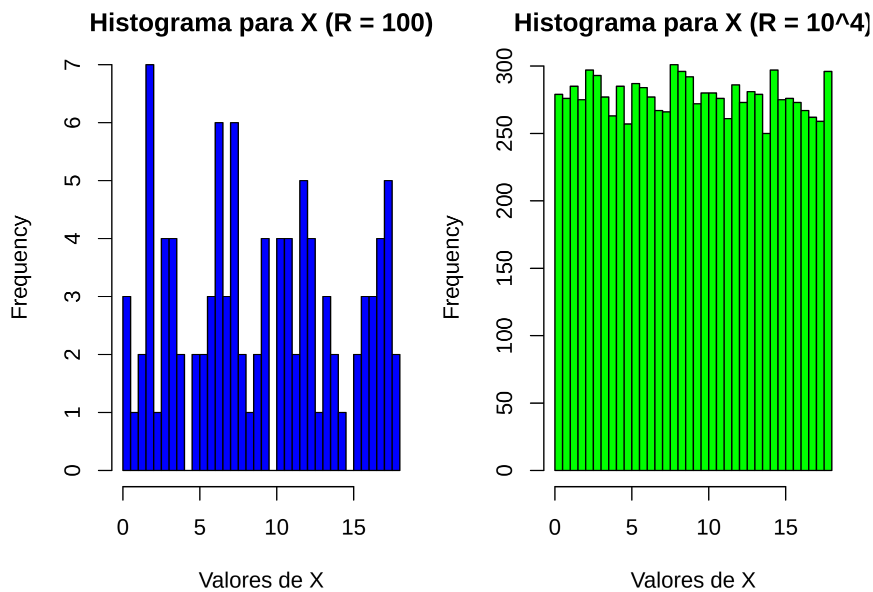
Al comparar con los valores muestrales obtenidos en 1.b, vemos que:

- La media muestral se acerca bastante al valor teórico de $E(X)$.
- La varianza muestral también se aproxima al valor teórico, especialmente cuando R es grande.

1.d Hacer dos histogramas de X, tomando $R \in \{100, 10^4\}$ realizaciones y 30 bines.

```r
par(mfrow = c(1, 2))
par(mar = c(4, 4, 2, 1))

# Histograma para R = 100
```

```r
hist(X_dist(100), breaks = 30, main = "Histograma para X (R = 100)", xlab = "Valores de X'

# Histograma para R = 10^4
hist(X_dist(10^4), breaks = 30, main = "Histograma para X (R = 10^4)", xlab = "Valores de
```



La distribución esperada es una distribución uniforme. Al aumentar el número de realizaciones R, el histograma se hace más suave y más cercano a la densidad teórica de la distribución uniforme.

Ejercicio 2: Otra Variable Aleatoria

Dado que $Y = \bar{X}_{15} = \frac{1}{15}\sum_{i=1}^{15} X_i$, donde cada X_i es independiente e idénticamente distribuido $X_i \sim U(0, 18)$.

2.a Generar una función Y_dist(R) que devuelva un vector con R realizaciones de Y.

```r
Y_dist <- function(R) {
  realizaciones <- replicate(R, mean(runif(15, min = 0, max = 18)))
  return(realizaciones)
}
```

2.b Calcular la media y la varianza muestral de los datos para $R \in \{2, 30, 100, 10^4\}$.

```
resultados_Y <- data.frame(R = integer(), Media = numeric(), Varianza = numeric())

for (R in R_values) {
  muestras_Y <- Y_dist(R)
  med_m_Y <- mean(muestras_Y)
  var_mue_Y <- var(muestras_Y)
  resultados_Y <- rbind(resultados_Y, data.frame(R = R, Media = med_m_Y, Varianza = var_mu
}

print(resultados_Y)
```

```
##       R    Media Varianza
## 1     2 8.427368 2.459774
## 2    30 8.954628 1.182842
## 3   100 9.124603 1.899916
## 4 10000 9.022645 1.766110
```

2.c Comparar las medias empíricas y calcular teóricamente $E(Y)$ y $V(Y)$.

Teóricamente, $E(Y) = E(X)$ y $V(Y) = \frac{V(X)}{15}$.

```
E_Y <- E_X
V_Y <- V_X / 15

print(paste("Esperanza E(Y):", E_Y))
```

```
## [1] "Esperanza E(Y): 9"
```

```
print(paste("Varianza V(Y):", V_Y))
```

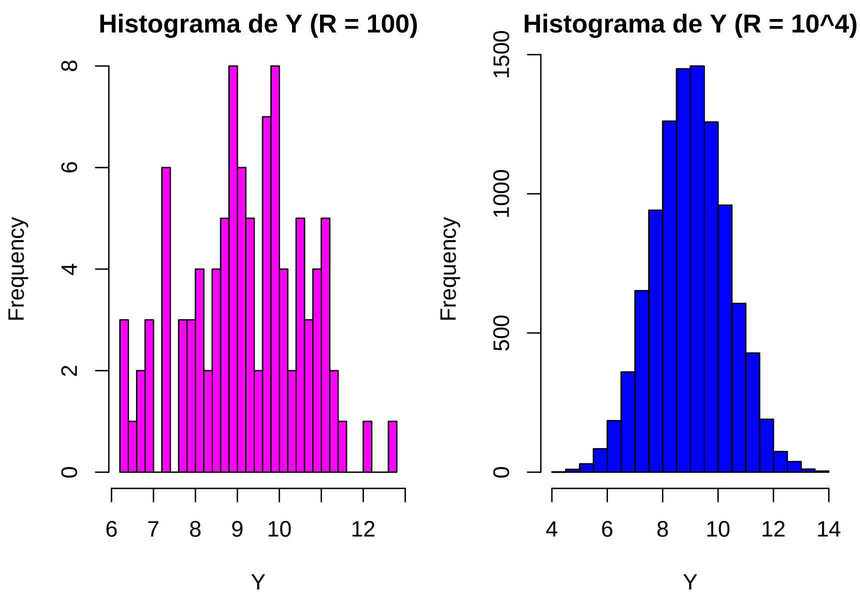
```
## [1] "Varianza V(Y): 1.8"
```

Al comparar, notamos que:

- Las medias empíricas se aproximan bastante al valor esperado teórico $E(Y)$.
- La varianza muestral también se aproxima al valor teórico a medida que R aumenta.

2.d Hacer dos histogramas de Y, tomando $R \in \{100, 10^4\}$ realizaciones y 30 bines.

```
par(mfrow = c(1, 2))
par(mar = c(4, 4, 2, 1))

# Histograma para R = 100
hist(Y_dist(100), breaks = 30, main = "Histograma de Y (R = 100)", xlab = "Y", col = "mage

# Histograma para R = 10^4
hist(Y_dist(10^4), breaks = 30, main = "Histograma de Y (R = 10^4)", xlab = "Y", col = "bl
```



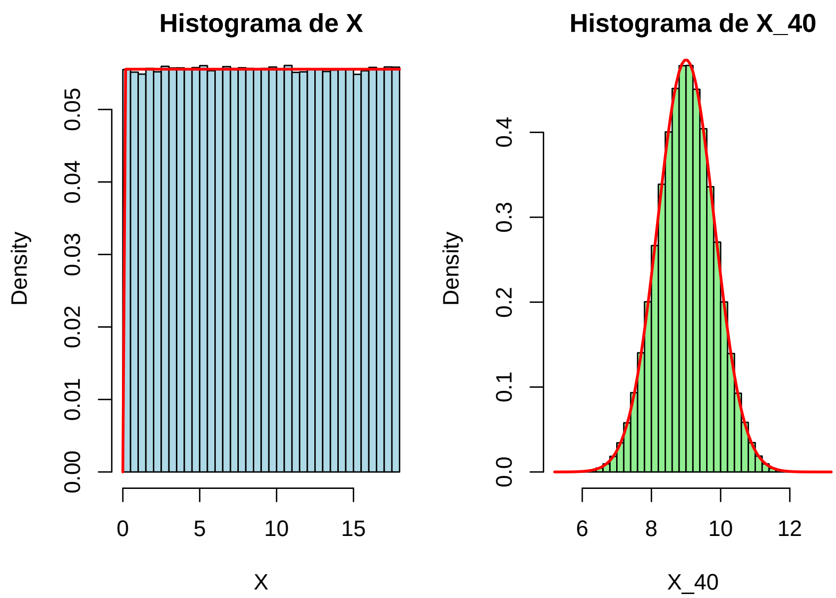
Esperamos ver una normal, tiene mucho menos dispersión que otro tipo de distribuciones dado a que es simétrica. Los histogramas se parecen porque hay menos varianza.

Ejercicio 3: Teorema Central del Límite

3.a Mostrar que se cumple el TCL para un caso particular

Generamos un histograma para X y otro para X_{40}, ambos con $R = 10^6$ realizaciones.

```
R <- 10^6
n <- 40

X <- runif(R, min = 0, max = 18)
X_40 <- replicate(R, mean(runif(n, min = 0, max = 18)))

par(mfrow = c(1, 2))
par(mar = c(4, 4, 2, 1))

# Histograma para X
hist(X, breaks = 50, probability = TRUE, main = "Histograma de X", xlab = "X", col = "ligh
curve(dunif(x, min = 0, max = 18), col = "red", lwd = 2, add = TRUE)

# Histograma para X_40
hist(X_40, breaks = 50, probability = TRUE, main = "Histograma de X_40", xlab = "X_40", co
mu <- 9  # Esperanza de X
sigma <- sqrt((18 - 0)^2 / 12) / sqrt(n)
curve(dnorm(x, mean = mu, sd = sigma), col = "red", lwd = 2, add = TRUE)
```

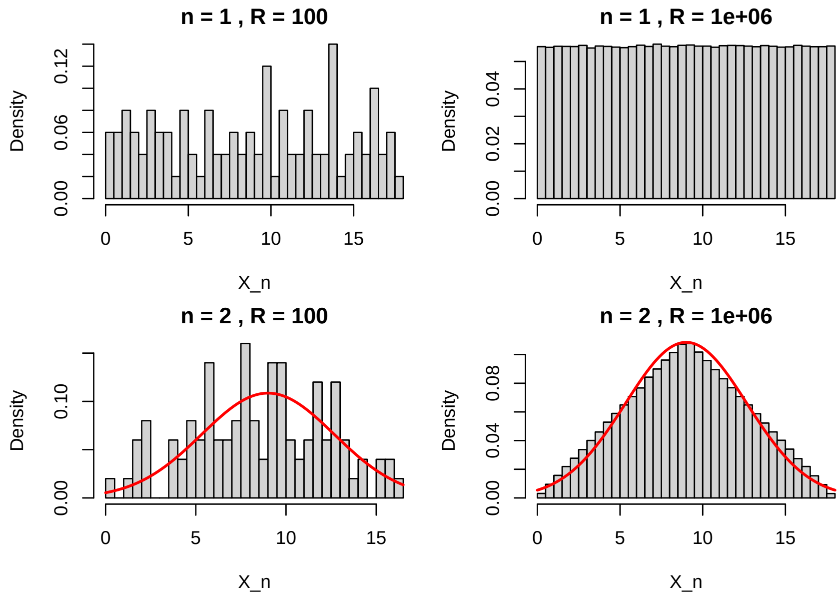


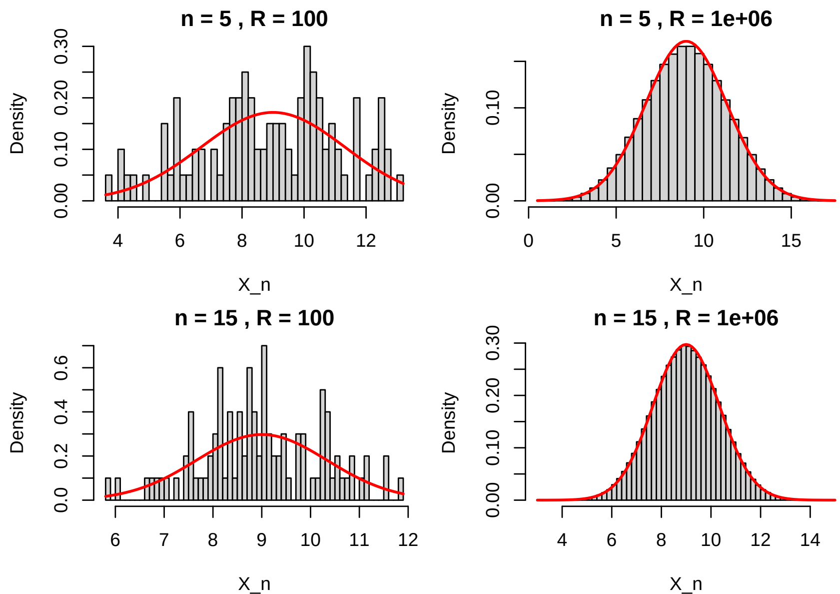
3.b Graficar ocho histogramas en un arreglo de 4x2 paneles

Se varía el número de realizaciones $n \in \{1, 2, 5, 15\}$ y el número de simulaciones $R \in \{10^2, 10^6\}$.

```
n_values <- c(1, 2, 5, 15)
R_values <- c(10^2, 10^6)

par(mfrow = c(2, 2))
par(mar = c(4, 4, 2, 1))

for (n in n_values) {
  for (R in R_values) {
    X_n <- replicate(R

, mean(runif(n, min = 0, max = 18)))
    hist(X_n, breaks = 50, probability = TRUE,
         main = paste("n =", n, ", R =", R),
         xlab = "X_n", col = "lightgray")

    if (n > 1) {
```

```
      mu <- 9  # Esperanza de X
      sigma <- sqrt((18 - 0)^2 / 12) / sqrt(n)
      curve(dnorm(x, mean = mu, sd = sigma), col = "red", lwd = 2, add = TRUE)
    }
  }
}
```





Sabemos que si modificamos la variable R, la distribución no cambiará. Simplemente mostrará más o menos sampleos de la misma distribución, más o menos definida, dependiendo de la cantidad de datos. En cambio, al variar n, el valor promedio de las muestras se acerca cada vez más a la esperanza teórica de la variable aleatoria subyacente, como se espera según el Teorema Central del Límite (TCL) y la Ley de los Grandes Números.