Trabajo Prático 1 en R

Ignacio Pardo & Luca Mazzarello

2022-08-23

Primera parte: Ley de los Grandes Números

Simulación del lanzamiento de un dado

1. Indicar el valor de $\mathrm{P}(X = 5)$ y el de $\mathrm{E}(X)$.

$$
\begin{aligned}
\mathrm{P}(X = 5) &= 1/6 \\
\mathrm{E}(X) &= \sum_{x \in Rx} x * \mathrm{P}(X = x) \\
&= \sum_{x \in Rx} x * 1/6 \\
&= \sum_{x \in Rx} x/6
\end{aligned}
$$

```r
dado <- c(1:6)
sum(dado / 6)
```

```
## [1] 3.5
```

$$\mathrm{E}(X) = 3.5$$

2. Construir un vector `muchosdados` y guardar en él los resultados correspondientes a lanzar `reps=1000` veces el dado

```r
reps <- 1000

muchosdados <- sample(x = dado, size = reps, replace = TRUE)
```

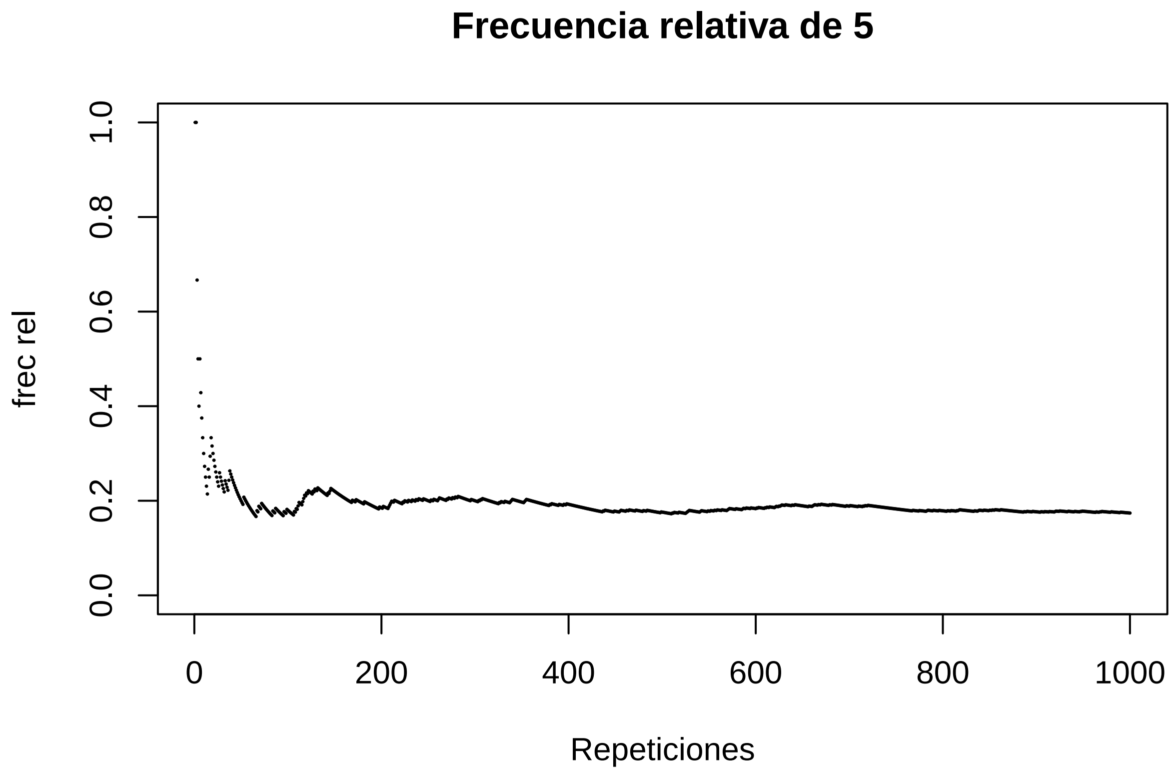
3. Para cada valor $n = 1, ..., reps$, calcular la frecuencia relativa con la que el 5 aparece en los primeros n lanzamientos y guardarla en el vector `frecrelativadado5vec`

```r
frecrelativadado5vec <- rep(NA, reps)

for (n in 1:reps) {
    frecrelativadado5vec[n] <- mean(muchosdados[1:n] == 5)
}
```

4. Graficar en el eje x los valores de n y en el eje y las correspondientes frecuencias relativas. ¿Qué observa? Indicar, si es posible, a qué valor deberían converger esas frecuencias y por qué. ¿Se corresponde lo que observa en la práctica con lo que espera de la teoría?

La frecuencia relativa se aproxima a $\mathrm{P}(X = 5)$ a medida que crece la cantidad de repeticiones.

```r
plot(seq(1:reps), frecrelativadado5vec, cex = 0.1, ylim = c(0, 1),
     xlab = "Repeticiones", ylab = "frec rel",
     main = "Frecuencia relativa de 5")
```



Segunda Parte: Teorema Central del Límite

Distribución Bernoulli

Sea $X \sim \mathrm{Be}(p)$ siendo $p = 0,2$.

```r
p_1 = 0.2
```

- Indicar el valor de $\mathrm{E}(X)$ y de $\mathrm{V}(X)$.

$$\mathrm{E}(X) = 0,2$$

$$\begin{aligned} \mathrm{V}(X) &= 0,2 * 0,8 \\ &= 0,16 \end{aligned}$$

- Guardar $Nrep = 10000$ realizaciones de X en el vector `ber_N_infty` con el comando `rbinom()` (prestar atención a los parámetros que utiliza esta función). Luego, obtener la frecuencia relativa de cada posible valor del rango de X. Se puede utilizar el comando `table()`. ¿Qué valores espera observar en esas frecuencias relativas? ¿Por qué?

```r
Nrep = 10000

ber_N_infty = rbinom(Nrep, 1, p_1)
table(ber_N_infty)
```

```
## ber_N_infty
```

```
##    0    1
## 7963 2037
```

```r
table(ber_N_infty)/Nrep
```

```
## ber_N_infty
##      0      1
## 0.7963 0.2037
```

La frecuencia relativa de X=1 se aproxima a la probabilidad (p) de la Bernoulli, mientras que X=0 se aproxima a 1-p por la Ley de los Grandes Números: El promedio de variables aleatorias converge en probabilidad a su esperanza con $n \to \infty$

- Calcular el promedio y la varianza muestral con los datos del vector `ber_N_infty` y comparar con los obtenidos en el primer ítem. ¿Qué observa?

```r
mean(ber_N_infty)
```

```
## [1] 0.2037
```

El promedio aproxima $\mathrm{E}(X) = 0,2$

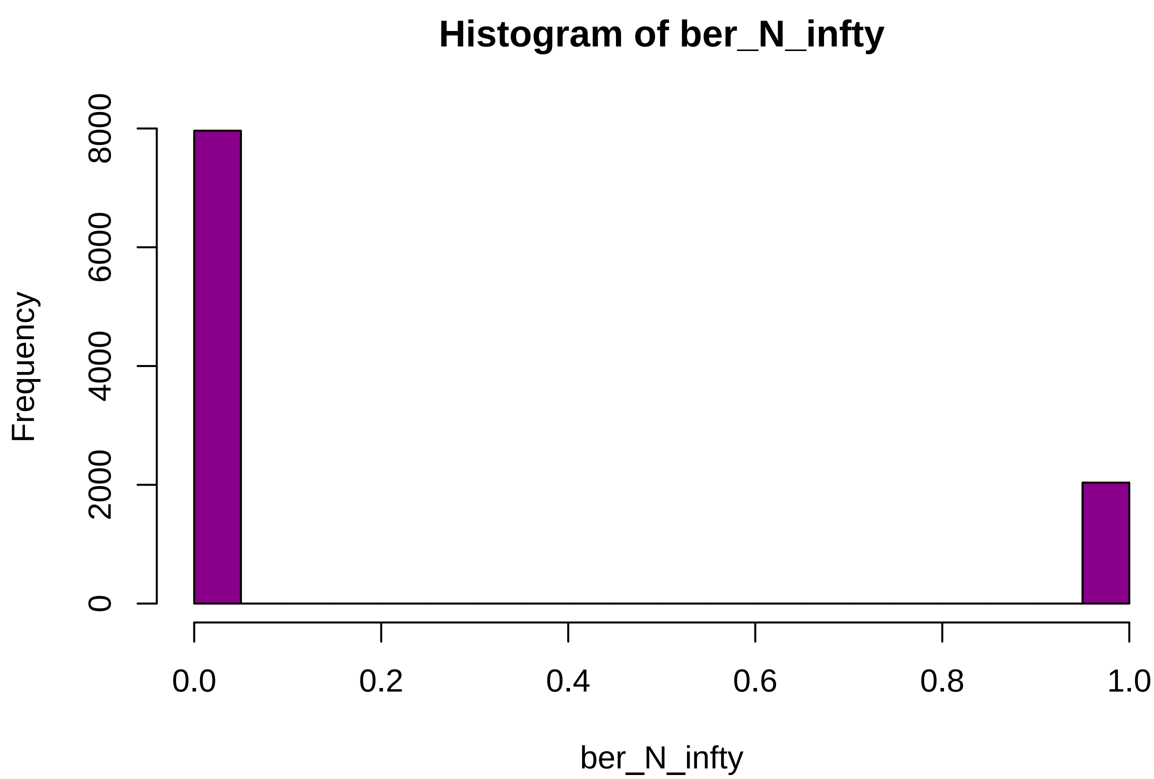
```r
var(ber_N_infty)
```

```
## [1] 0.1622225
```

La varianz aproxima $\mathrm{V}(X) = 0,16$

- Realizar un histograma de los datos. ¿Qué observa? ¿Es consistente con la distribución a partir de la cual fueron generados los datos?

```r
hist(ber_N_infty, col = "darkmagenta")
```



Las ocurrencias de $X = 0$ en la distribución se aproximan a 8000, y las de $X = 1$ a 2000, las cuales corresponden con la probabilidad $p = 0m2$ y la cantidad de repeticiones $Nrep = 10000$

La distribución empírica del promedio

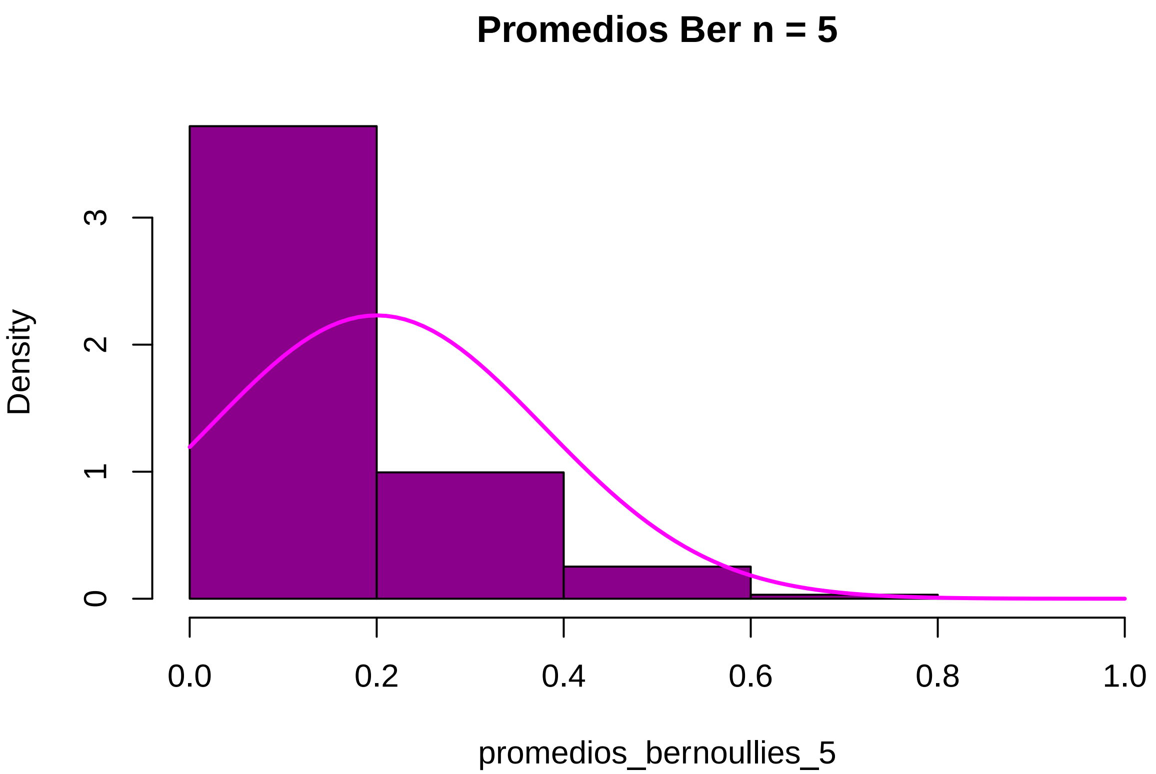
- Guardar en el vector `promedios_bernoullies`, $Nrep = 10000$ promedios utilizando $n = 5$ datos con distribución X.

```
n_dist = 5
promedios_bernoullies_5 = replicate(Nrep, mean(rbinom(n_dist, 1, p_1)))
table(promedios_bernoullies_5)
```

```
## promedios_bernoullies_5
##    0  0.2  0.4  0.6  0.8    1
## 3312 4128 1989  506   63    2
```

- Con los promedios guardados en promedios bernoullies, realizar un histograma de densidad. Para el histograma de densidad, usar el argumento `probability = TRUE` en el comando `hist()`. Además, como n = 5 puede ser pequeño, al usar `hist()` puede ajustar el parámetro `breaks`. Observar qué ocurre al poner (o no) ese parámetro. Luego, superponer el gráfico de la curva de la densidad normal con la media y desvío que considere razonables.

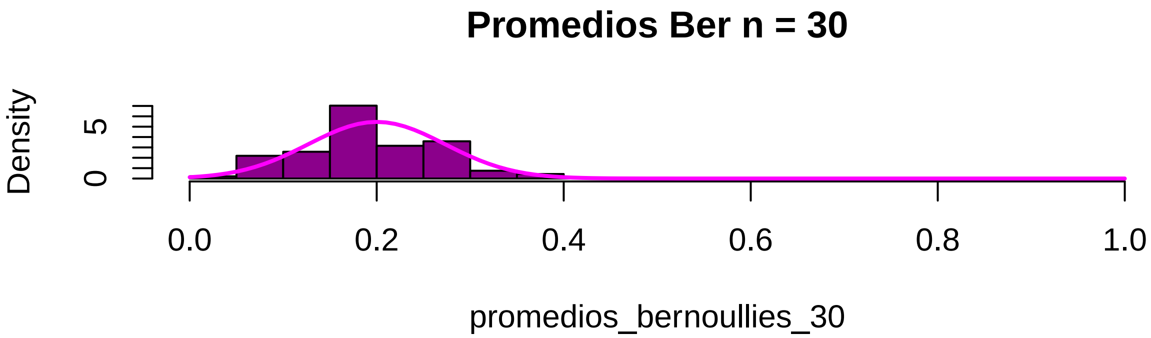
```
hist(main = "Promedios Ber n = 5", promedios_bernoullies_5, probability = TRUE, xlim = c(0,1), col = "da
curve(dnorm(x, p_1, sqrt(p_1 * (1-p_1) / 5)), add=TRUE, lwd=2, col="magenta")
```

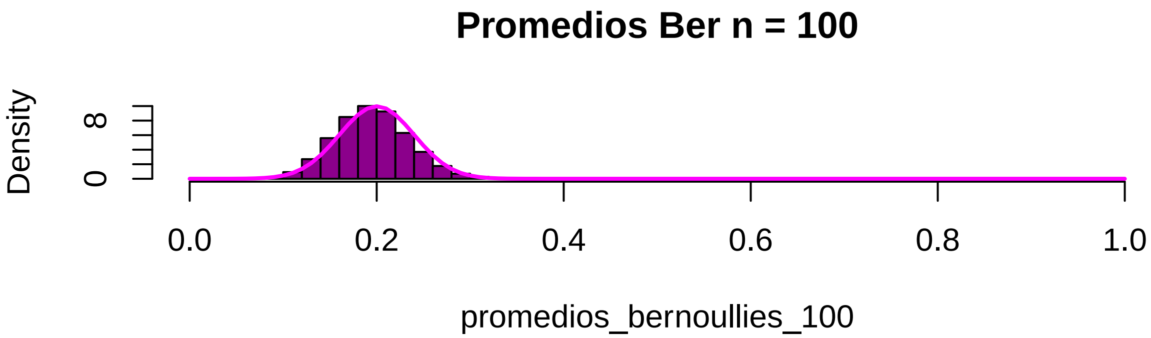


- Repetir los ítems 2 y 3 con $n = 30$ y con $n = 100$. Comparar los tres histogramas obtenidos y explicar qué observa en cada caso y a qué cree que se debe. Tener cuidado con la diferencia en las escalas.

```
n_dist = 30
promedios_bernoullies_30 = replicate(Nrep, mean(rbinom(n_dist, 1, p_1)))

par(mfrow=c(2, 1))
hist(main = "Promedios Ber n = 30", promedios_bernoullies_30, probability = TRUE, xlim = c(0,1), col =
curve(dnorm(x, p_1, sqrt(p_1 * (1-p_1) / 30)), add=TRUE, lwd=2, col="magenta")

n_dist = 100
```

```
promedios_bernoullies_100 = replicate(Nrep, mean(rbinom(n_dist, 1, p_1)))
hist(main = "Promedios Ber n = 100", promedios_bernoullies_100, probability = TRUE, xlim = c(0,1), col =
curve(dnorm(x, p_1, sqrt(p_1 * (1-p_1) / 100)), add=TRUE, lwd=2, col="magenta")
```





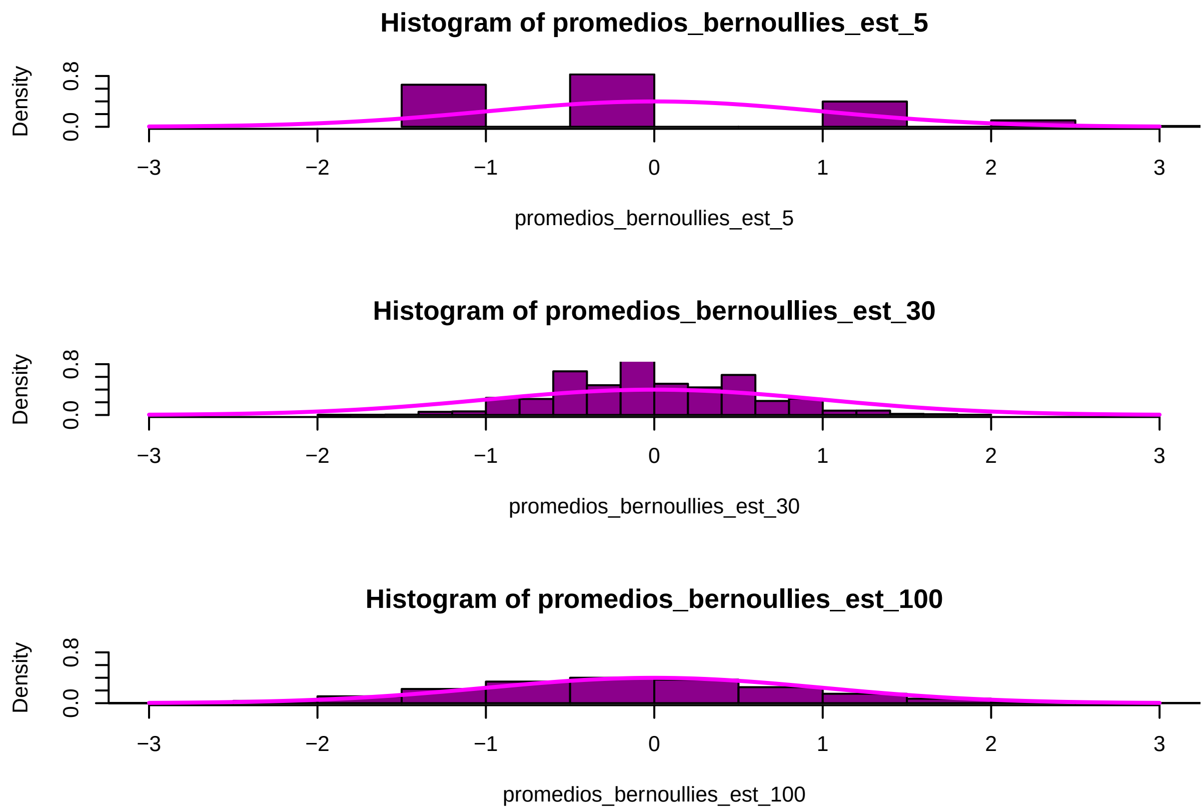

```
par(mfrow=c(1, 1))
```

La distribución empírica del promedio estandarizado

- Realizar ahora histogramas de densidad con los datos de `promedios_bernoullies_est` donde a los valores guardados en `promedios_bernoullies` se les resta μ y se los divide por $\sqrt(\sigma^2/n)$. Utilizar $n \in \{5, 30, 100\}$. ¿Qué observa? Explicar.

  Con n cada vez mayor, la distribución de la muestra se aproxima más a la distribucíon de una $N(0, 1)$.

```
promedios_bernoullies_est_5 = (promedios_bernoullies_5 - p_1) / sqrt(p_1 * (1-p_1) / 5)

par(mfrow=c(3, 1))
hist(promedios_bernoullies_est_5, xlim = c(-3, 3), ylim = c(0, 0.8), probability = TRUE, col = "darkmage
curve(dnorm(x,0,1), add=TRUE, lwd=2, col="magenta")

promedios_bernoullies_est_30 = (promedios_bernoullies_100 - p_1) / sqrt(p_1 * (1-p_1) / 30)
hist(promedios_bernoullies_est_30, xlim = c(-3, 3), ylim = c(0, 0.8), probability = TRUE, col = "darkmag
curve(dnorm(x,0,1), add=TRUE, lwd=2, col="magenta")

promedios_bernoullies_est_100 = (promedios_bernoullies_100 - p_1) / sqrt(p_1 * (1-p_1) / 100)
hist(promedios_bernoullies_est_100, xlim = c(-3, 3), ylim = c(0, 0.8), probability = TRUE, col = "darkma
curve(dnorm(x,0,1), add=TRUE, lwd=2, col="magenta")
```



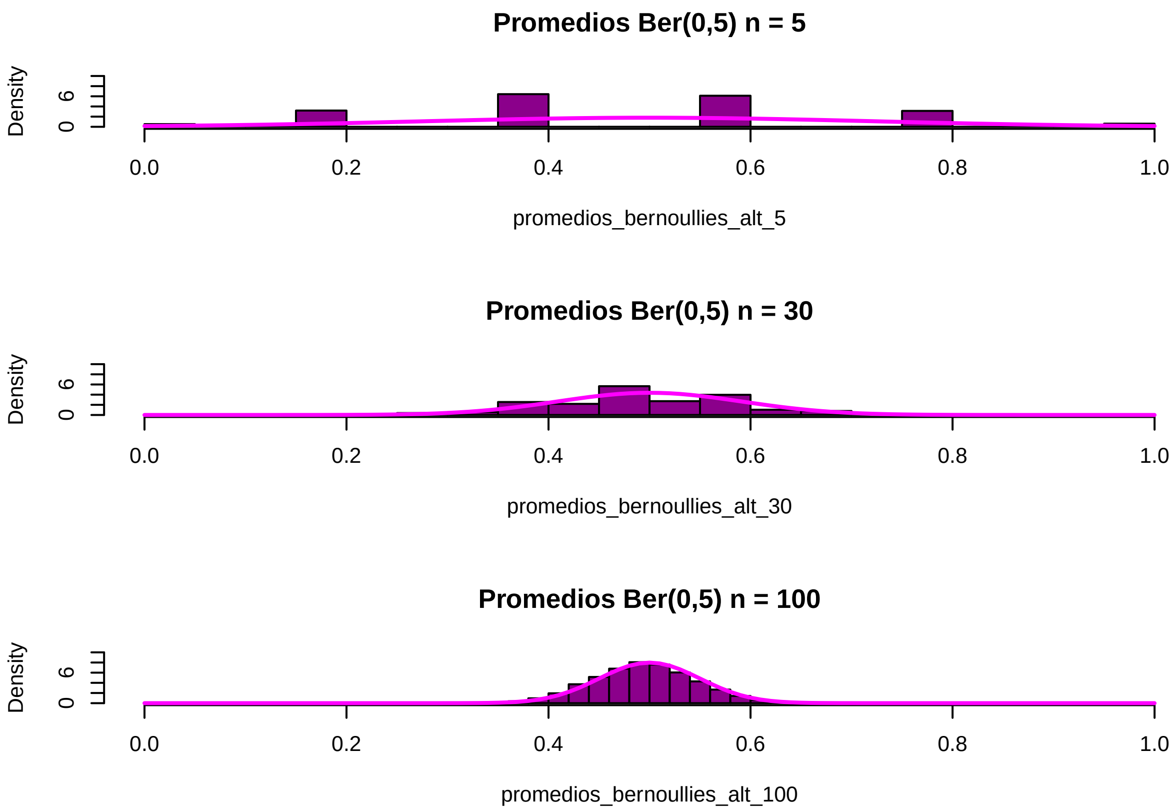

```
par(mfrow=c(1, 1))
```

- A los histogramas de densidad del ítem anterior, superponer las curvas de la densidad normal con la media y desvío que considere adecuados. ¿Qué observa? Explicar.

  Con $n \to \infty$, el histograma de densidad de la estimación de la normal apartir de la muestra con elementos Bernoulli se aproxima a la curva de la normal $N(0,1)$

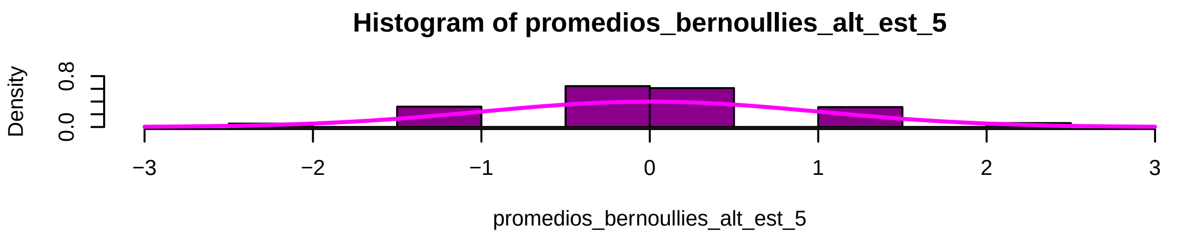
- Repetir todos los ítems anteriores con $X \sim \text{Be}(p)$ siendo $p = 0,5$. ¿Hay alguna diferencia? Comentar.

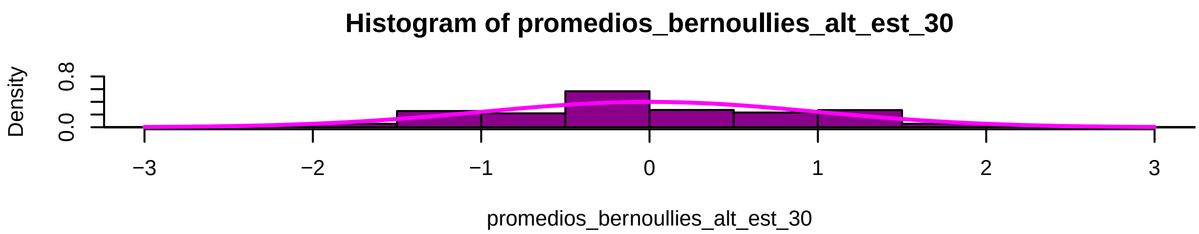
```
p_2 = 0.5

par(mfrow=c(3, 1))

n_dist = 5
promedios_bernoullies_alt_5 = replicate(Nrep, mean(rbinom(n_dist, 1, p_2)))
hist(main = "Promedios Ber(0,5) n = 5", promedios_bernoullies_alt_5, probability = TRUE, xlim = c(0,1),
curve(dnorm(x, p_2, sqrt(p_2 * (1-p_2) / n_dist)), add=TRUE, lwd=2, col="magenta")

n_dist = 30
promedios_bernoullies_alt_30 = replicate(Nrep, mean(rbinom(n_dist, 1, p_2)))
hist(main = "Promedios Ber(0,5) n = 30", promedios_bernoullies_alt_30, probability = TRUE, xlim = c(0,1
curve(dnorm(x, p_2, sqrt(p_2 * (1-p_2) / n_dist)), add=TRUE, lwd=2, col="magenta")

n_dist = 100
promedios_bernoullies_alt_100 = replicate(Nrep, mean(rbinom(n_dist, 1, p_2)))
hist(main = "Promedios Ber(0,5) n = 100", promedios_bernoullies_alt_100, probability = TRUE, xlim = c(0
curve(dnorm(x, p_2, sqrt(p_2 * (1-p_2) / n_dist)), add=TRUE, lwd=2, col="magenta")
```

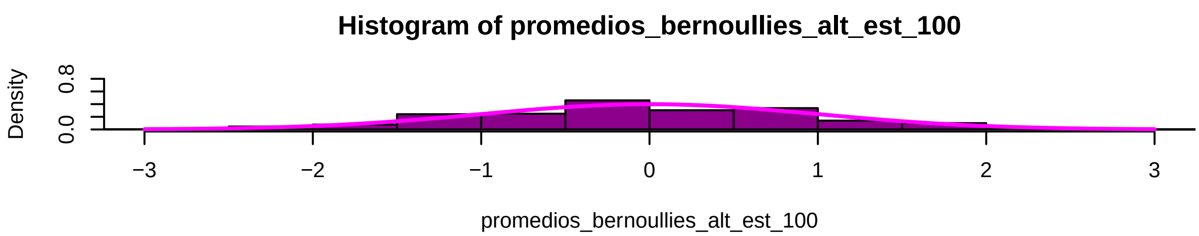



```
par(mfrow=c(1, 1))
```

```
par(mfrow=c(3, 1))

n_dist = 5
promedios_bernoullies_alt_est_5 = (promedios_bernoullies_alt_5 - p_2) / sqrt(p_2 * (1-p_2) / n_dist)
hist(promedios_bernoullies_alt_est_5, xlim = c(-3, 3), ylim = c(0, 0.8), probability = TRUE, col="darkma
curve(dnorm(x,0,1), add=TRUE, lwd=2, col="magenta")

n_dist = 30
promedios_bernoullies_alt_est_30 = (promedios_bernoullies_alt_30 - p_2) / sqrt(p_2 * (1-p_2) / n_dist)
hist(promedios_bernoullies_alt_est_30, xlim = c(-3, 3), ylim = c(0, 0.8), probability = TRUE, col="darkm
curve(dnorm(x,0,1), add=TRUE, lwd=2, col="magenta")

n_dist = 100
promedios_bernoullies_alt_est_100 = (promedios_bernoullies_alt_100 - p_2) / sqrt(p_2 * (1-p_2) / n_dist
hist(promedios_bernoullies_alt_est_100, xlim = c(-3, 3), ylim = c(0, 0.8), probability = TRUE, col="dark
curve(dnorm(x,0,1), add=TRUE, lwd=2, col="magenta")
```








```
par(mfrow=c(1, 1))
```

Distribución Uniforme

Sea $X \sim \mathcal{U}(a, b)$ siendo $a = 67$, $b = 73$.

```
a = 67
b = 73
```

- Indicar el valor de $\mathrm{E}(X)$ y de $\mathrm{V}(X)$.

$$
\begin{aligned}
\mathrm{E}(X) &= \frac{b+a}{2} & \mathrm{V}(X) &= \frac{(b-a)^2}{12} \\
&= \frac{73+67}{2} & &= \frac{(73-67)^2}{12} \\
&= \frac{140}{2} & &= \frac{6^2}{12} \\
&= 70 & &= \frac{36}{12} \\
& & &= 3
\end{aligned}
$$

- Guardar $Nrep = 10000$ realizaciones de X en el vector `unif_N_infty` con el comando `runif()` (prestar atención a los parámetros que utiliza esta función). Luego, obtener la frecuencia relativa de cada posible valor del rango de X. Se puede utilizar el comando `table()`. ¿Qué valores espera observar en esas frecuencias relativas? ¿Por qué?

```
Nrep = 10000

unif_N_infty = runif(Nrep, a, b)
```

- Calcular el promedio y la varianza muestral con los datos del vector `ber_N_infty` y comparar con los obtenidos en el primer ítem. ¿Qué observa?

```
e_unif = (a + b) / 2
e_unif_m = mean(unif_N_infty)
e_unif
```

```
## [1] 70
```

El promedio aproxima

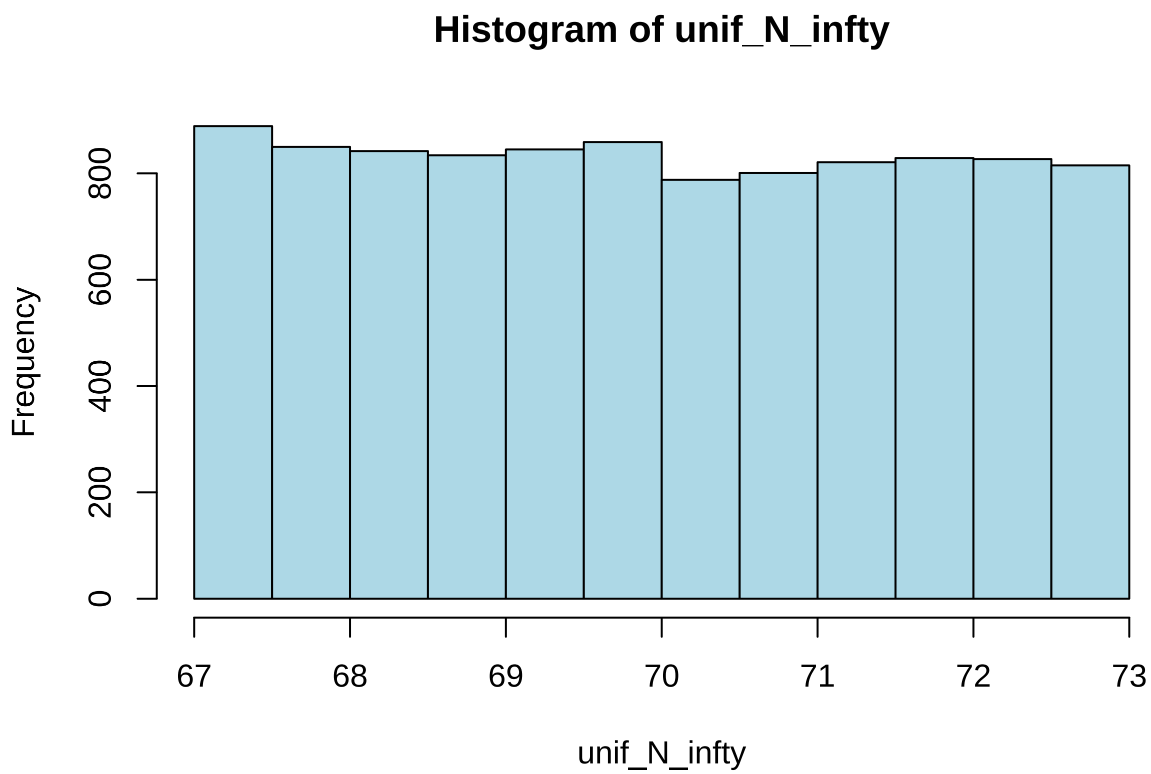
```
var_unif = (b - a)^2 / 12
var_unif_m = var(unif_N_infty)
var_unif
```

```
## [1] 3
```

La varianza aproxima

- Realizar un histograma de los datos. ¿Qué observa? ¿Es consistente con la distribución a partir de la cual fueron generados los datos?

```
hist(unif_N_infty, col = "lightblue")
```



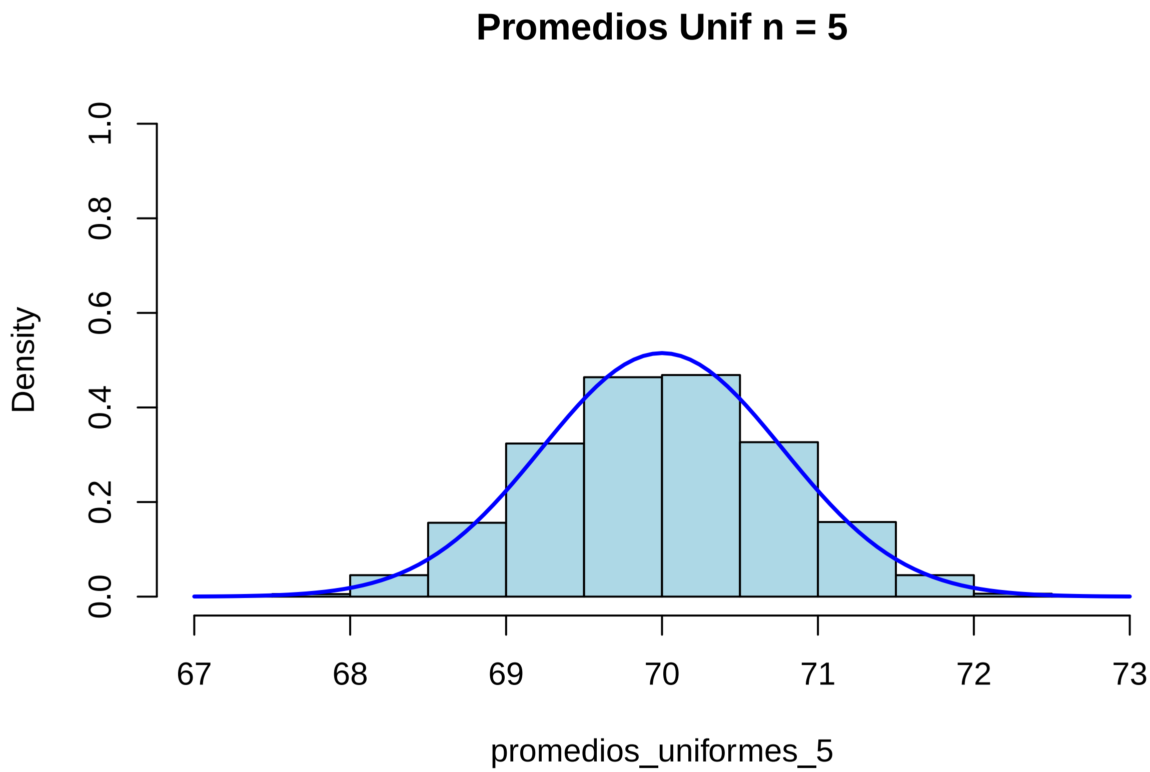
La distribución empírica del promedio

- Guardar en el vector `promedios_uniformes`, $Nrep = 10000$ promedios utilizando $n = 5$ datos con distribución X.

```
n_dist = 5
promedios_uniformes_5 = replicate(Nrep, mean(runif(n_dist, a, b)))
```

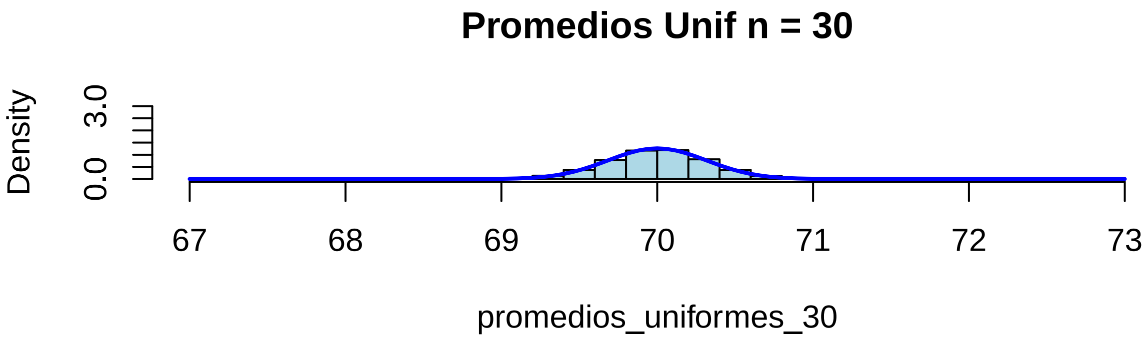
- Con los promedios guardados en promedios bernoullies, realizar un histograma de densidad. Para el histograma de densidad, usar el argumento `probability = TRUE` en el comando `hist()`. Además, como n = 5 puede ser pequeño, al usar `hist()` puede ajustar el parámetro `breaks`. Observar qué ocurre al poner (o no) ese parámetro. Luego, superponer el gráfico de la curva de la densidad normal con la media y desvío que considere razonables.

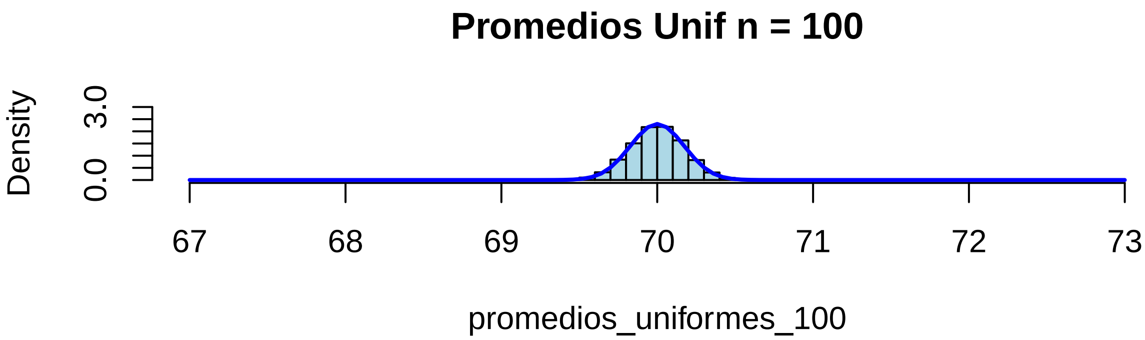
```
n_dist = 5
hist(main = "Promedios Unif n = 5", promedios_uniformes_5, probability = TRUE, xlim = c(a, b), ylim=c(0
curve(dnorm(x, e_unif, sqrt(var_unif/n_dist)), add=TRUE, lwd=2, col="blue")
```



- Repetir los ítems 2 y 3 con $n = 30$ y con $n = 100$. Comparar los tres histogramas obtenidos y explicar qué observa en cada caso y a qué cree que se debe. Tener cuidado con la diferencia en las escalas.

```
n_dist = 30
promedios_uniformes_30 = replicate(Nrep, mean(runif(n_dist, a, b)))

par(mfrow=c(2, 1))
hist(main = "Promedios Unif n = 30", promedios_uniformes_30, probability = TRUE, xlim = c(a, b), ylim =
curve(dnorm(x, e_unif, sqrt( var_unif/ n_dist)), add=TRUE, lwd=2, col="blue")

n_dist = 100
promedios_uniformes_100 = replicate(Nrep, mean(runif(n_dist, a, b)))
hist(main = "Promedios Unif n = 100", promedios_uniformes_100, probability = TRUE, xlim = c(a, b), ylim
curve(dnorm(x, e_unif, sqrt( var_unif/ n_dist)), add=TRUE, lwd=2, col="blue")
```



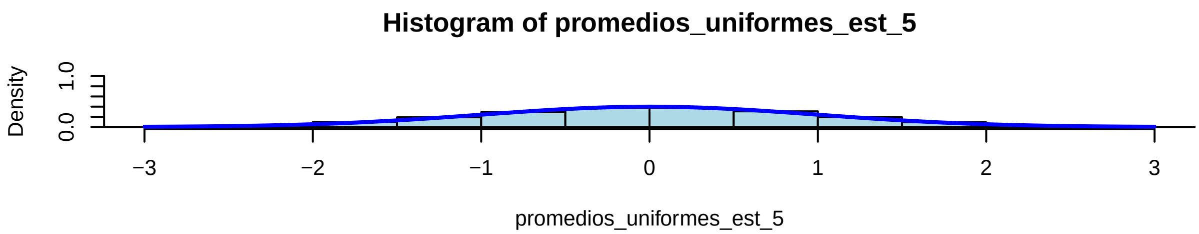


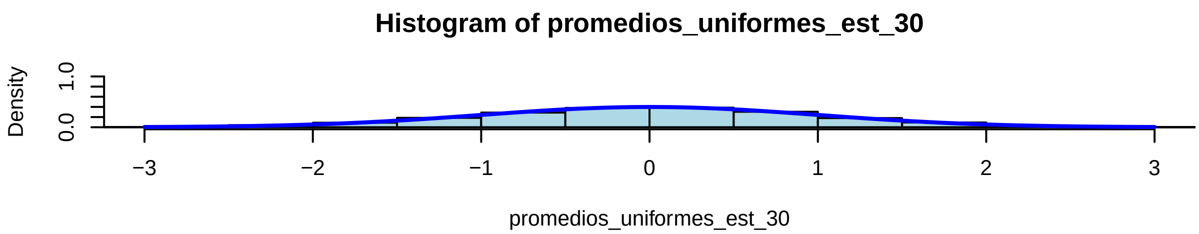

```
par(mfrow=c(1, 1))
```

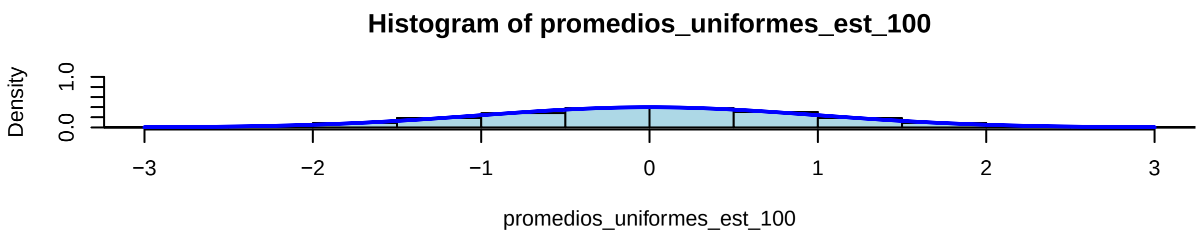
La distribución empírica del promedio estandarizado

- Realizar ahora histogramas de densidad con los datos de `promedios_uniformes_est` donde a los valores guardados en `promedios_uniformes` se les resta μ y se los divide por $\sqrt(\sigma^2/n)$. Utilizar $n \in \{5, 30, 100\}$. ¿Qué observa? Explicar.

```
promedios_uniformes_est_5 = (promedios_uniformes_5 - e_unif) / sqrt(var_unif / 5)

par(mfrow=c(3, 1))
hist(promedios_uniformes_est_5, xlim = c(-3, 3), ylim = c(0, 1), probability = TRUE, col = "lightblue")
curve(dnorm(x,0,1), add=TRUE, lwd=2, col="blue")

promedios_uniformes_est_30 = (promedios_uniformes_30 - e_unif) / sqrt(var_unif / 30)
hist(promedios_uniformes_est_30, xlim = c(-3, 3), ylim = c(0, 1), probability = TRUE, col = "lightblue"
curve(dnorm(x,0,1), add=TRUE, lwd=2, col="blue")

promedios_uniformes_est_100 = (promedios_uniformes_100 - e_unif) / sqrt(var_unif / 100)
hist(promedios_uniformes_est_100, xlim = c(-3, 3), ylim = c(0, 1), probability = TRUE, col = "lightblue
curve(dnorm(x,0,1), add=TRUE, lwd=2, col="blue")
```








```
par(mfrow=c(1, 1))
```

- A los histogramas de densidad del ítem anterior, superponer las curvas de la densidad normal con la media y desvío que considere adecuados. ¿Qué observa? Explicar.

  Con $n \to \infty$, el histograma de densidad de la estimación de la normal apartir de la muestra con elementos Uniforme se aproxima a la curva de la normal $N(0, 1)$

Distribución Exponencial

Sea $X \sim \mathcal{E}(\lambda)$ siendo $\lambda = \frac{1}{10}$.

```
l = 0.1
```

- Indicar el valor de $\mathrm{E}(X)$ y de $\mathrm{V}(X)$.

$$
\begin{aligned}
\mathrm{E}(X) &= \frac{1}{\lambda} & \mathrm{V}(X) &= \frac{1}{\lambda^2} \\
&= \frac{1}{\frac{1}{10}} & &= \frac{1}{(\frac{1}{10})^2} \\
&= 10 & &= \frac{1}{\frac{1}{100}} \\
& & &= 100
\end{aligned}
$$

- Guardar $Nrep = 10000$ realizaciones de X en el vector `exp_N_infty` con el comando `rexp()` (prestar atención a los parámetros que utiliza esta función). Luego, obtener la frecuencia relativa de cada posible valor del rango de X. Se puede utilizar el comando `table()`. ¿Qué valores espera observar en esas frecuencias relativas? ¿Por qué?

```r
Nrep = 10000

exp_N_infty = rexp(Nrep, l)
```

- Calcular el promedio y la varianza muestral con los datos del vector `ber_N_infty` y comparar con los obtenidos en el primer ítem. ¿Qué observa?

```r
e_exp = mean(exp_N_infty)
e_exp = 1/l
e_exp
```

```
## [1] 10
```

El promedio aproxima

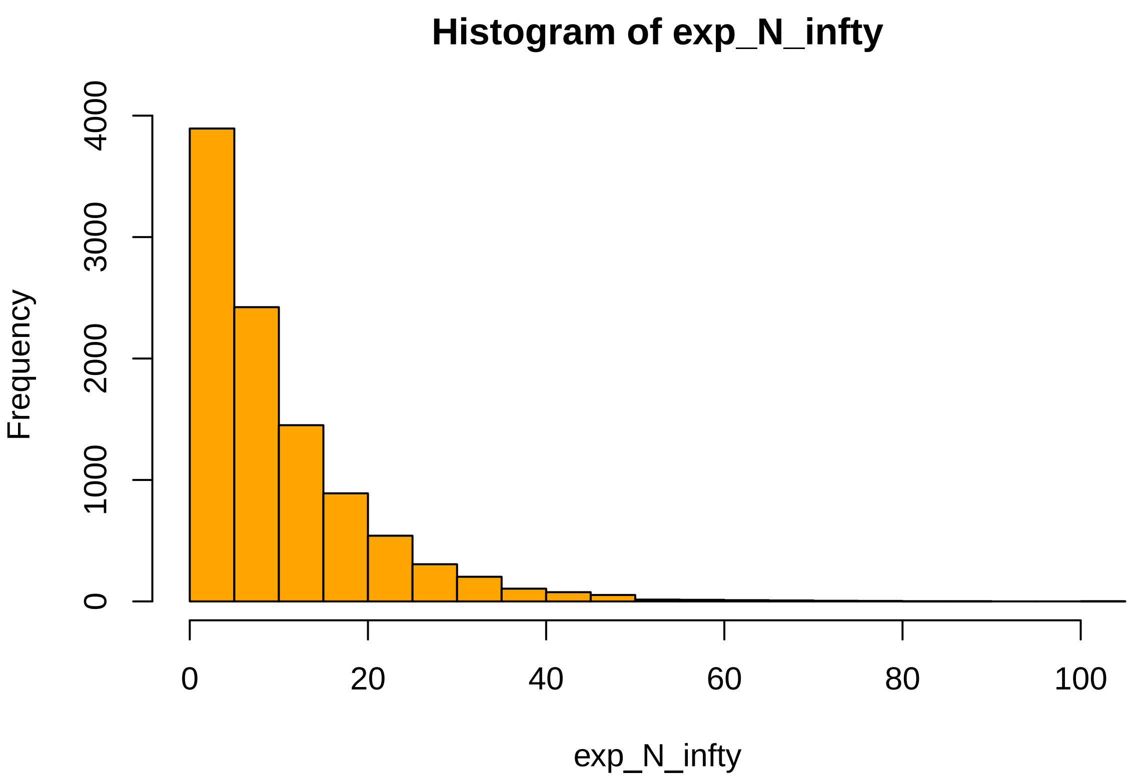
```r
var_exp = var(exp_N_infty)
var_exp = 1/(l^2)
var_exp
```

```
## [1] 100
```

La varianza aproxima

- Realizar un histograma de los datos. ¿Qué observa? ¿Es consistente con la distribución a partir de la cual fueron generados los datos?

```r
hist(exp_N_infty, col = "orange")
```



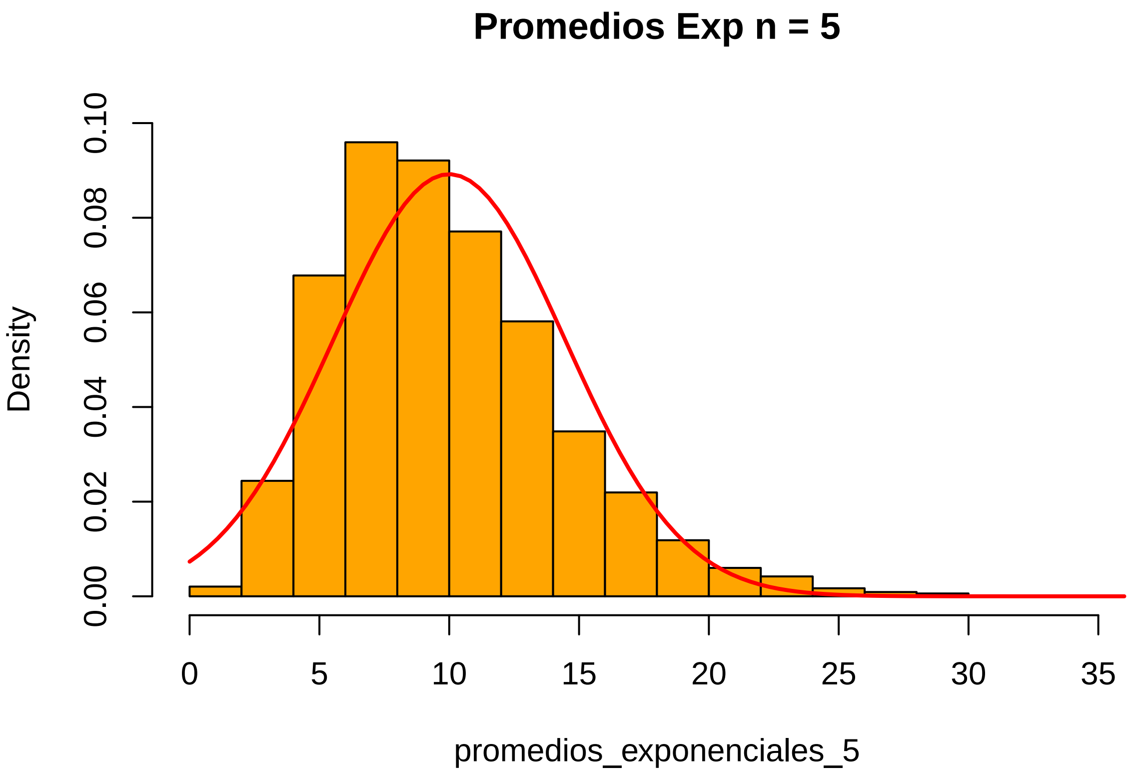
La distribución empírica del promedio

- Guardar en el vector `promedios_exponenciales`, $Nrep = 10000$ promedios utilizando $n = 5$ datos con distribución X.

```r
n_dist = 5
promedios_exponenciales_5 = replicate(Nrep, mean(rexp(n_dist, 1)))
```

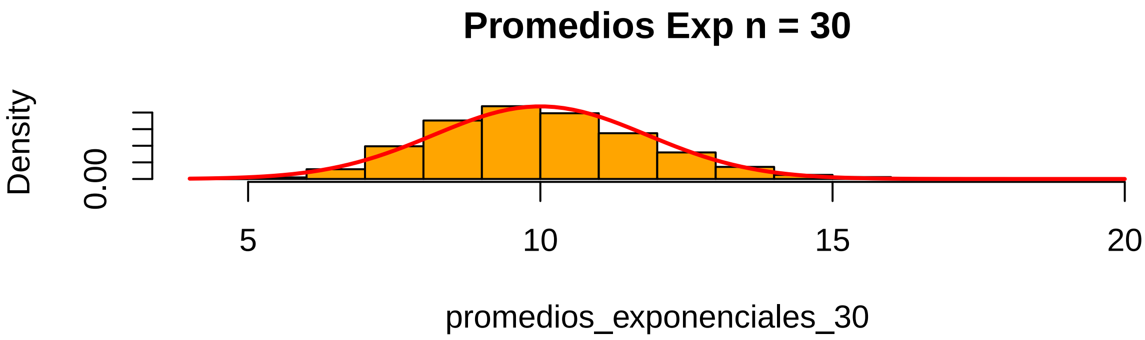
- Con los promedios guardados en promedios bernoullies, realizar un histograma de densidad. Para el histograma de densidad, usar el argumento `probability = TRUE` en el comando `hist()`. Además, como n = 5 puede ser pequeño, al usar `hist()` puede ajustar el parámetro `breaks`. Observar qué ocurre al poner (o no) ese parámetro. Luego, superponer el gráfico de la curva de la densidad normal con la media y desvío que considere razonables.

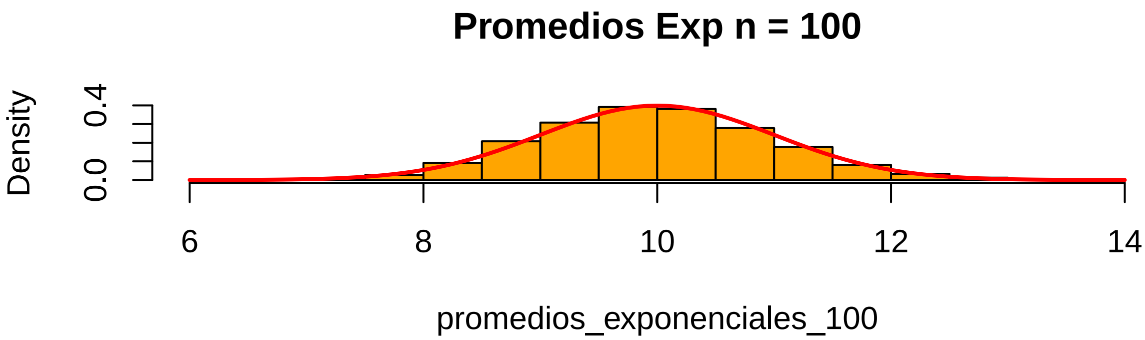
```r
n_dist = 5
hist(main = "Promedios Exp n = 5", promedios_exponenciales_5, probability = TRUE, ylim=c(0, 1), col = "
curve(dnorm(x, e_exp, sqrt(var_exp/ n_dist)), add=TRUE, lwd=2, col="red")
```



- Repetir los ítems 2 y 3 con $n = 30$ y con $n = 100$. Comparar los tres histogramas obtenidos y explicar qué observa en cada caso y a qué cree que se debe. Tener cuidado con la diferencia en las escalas.

```r
n_dist = 30
promedios_exponenciales_30 = replicate(Nrep, mean(rexp(n_dist, 1)))

par(mfrow=c(2, 1))
hist(main = "Promedios Exp n = 30", promedios_exponenciales_30, probability = TRUE, col = "orange")
curve(dnorm(x, e_exp, sqrt( var_exp/ n_dist)), add=TRUE, lwd=2, col="red")

n_dist = 100
promedios_exponenciales_100 = replicate(Nrep, mean(rexp(n_dist, 1)))
hist(main = "Promedios Exp n = 100", promedios_exponenciales_100, probability = TRUE, col = "orange")
curve(dnorm(x, e_exp, sqrt( var_exp/ n_dist)), add=TRUE, lwd=2, col="red")
```



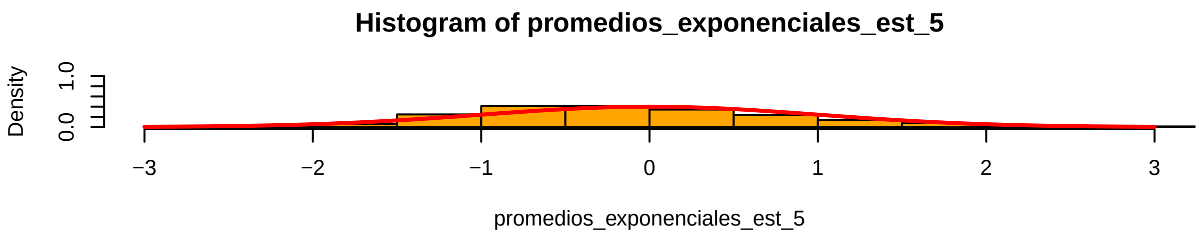


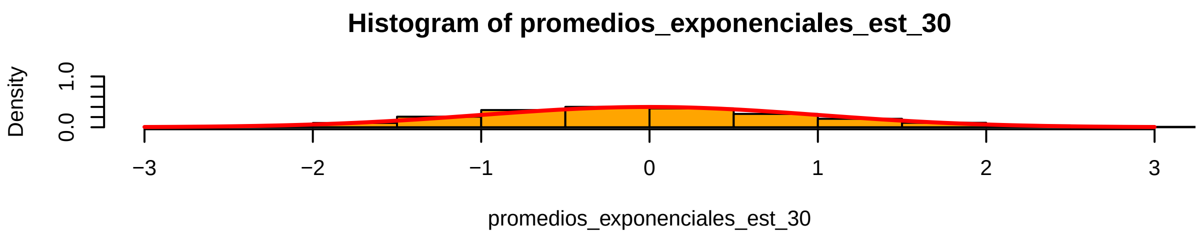

```
par(mfrow=c(1, 1))
```

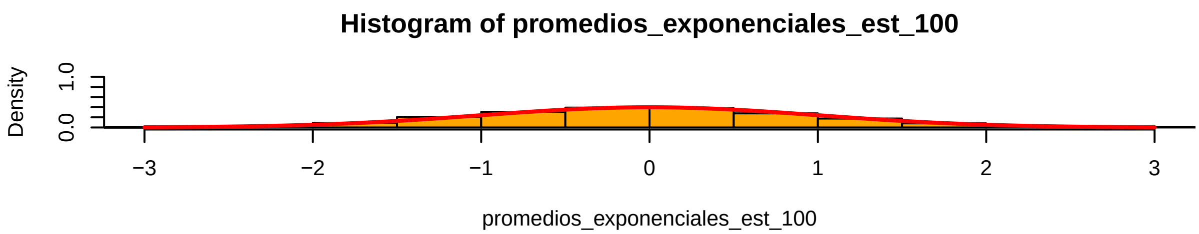
La distribución empírica del promedio estandarizado

- Realizar ahora histogramas de densidad con los datos de `promedios_exponenciales_est` donde a los valores guardados en `promedios_exponenciales` se les resta μ y se los divide por $\sqrt{(\sigma^2/n)}$. Utilizar $n \in \{5, 30, 100\}$. ¿Qué observa? Explicar.

```
promedios_exponenciales_est_5 = (promedios_exponenciales_5 - e_exp) / sqrt(var_exp / 5)

par(mfrow=c(3, 1))
hist(promedios_exponenciales_est_5, xlim = c(-3, 3), ylim = c(0, 1), probability = TRUE, col = "orange")
curve(dnorm(x,0,1), add=TRUE, lwd=2, col="red")

promedios_exponenciales_est_30 = (promedios_exponenciales_30 - e_exp) / sqrt(var_exp / 30)
hist(promedios_exponenciales_est_30, xlim = c(-3, 3), ylim = c(0, 1), probability = TRUE, col = "orange
curve(dnorm(x,0,1), add=TRUE, lwd=2, col="red")

promedios_exponenciales_est_100 = (promedios_exponenciales_100 - e_exp) / sqrt(var_exp / 100)
hist(promedios_exponenciales_est_100, xlim = c(-3, 3), ylim = c(0, 1), probability = TRUE, col = "orang
curve(dnorm(x,0,1), add=TRUE, lwd=2, col="red")
```








```
par(mfrow=c(1, 1))
```

- A los histogramas de densidad del ítem anterior, superponer las curvas de la densidad normal con la media y desvío que considere adecuados. ¿Qué observa? Explicar.

  Con $n \to \infty$, el histograma de densidad de la estimación de la normal apartir de la muestra con elementos Exponencial se aproxima a la curva de la normal $N(0, 1)$

Distribución Normal: un caso particular

Sea $X \sim \mathcal{N}(\mu, \sigma^2)$ siendo $\mu = 70$, $\sigma^2 = 1,2$.

```
mu = 70
sd2 = 1.2
sd = sqrt(sd2)
```

- Indicar el valor de $\text{E}(X)$ y de $\text{V}(X)$.

$$\text{E}(X) = \mu = 70$$
$$\text{V}(X) = \sigma^2 = 1,2$$

- Guardar $Nrep = 10000$ realizaciones de X en el vector `norm_N_infty` con el comando `rnorm()` (prestar atención a los parámetros que utiliza esta función). Luego, obtener la frecuencia relativa de cada posible valor del rango de X. Se puede utilizar el comando `table()`. ¿Qué valores espera observar en esas frecuencias relativas? ¿Por qué?

```
Nrep = 10000

norm_N_infty = rnorm(Nrep, mu, sd)
```

- Calcular el promedio y la varianza muestral con los datos del vector `ber_N_infty` y comparar con los obtenidos en el primer ítem. ¿Qué observa?

```
e_norm = mean(norm_N_infty)
e_norm = mu
e_norm
```

```
## [1] 70
```

El promedio aproxima

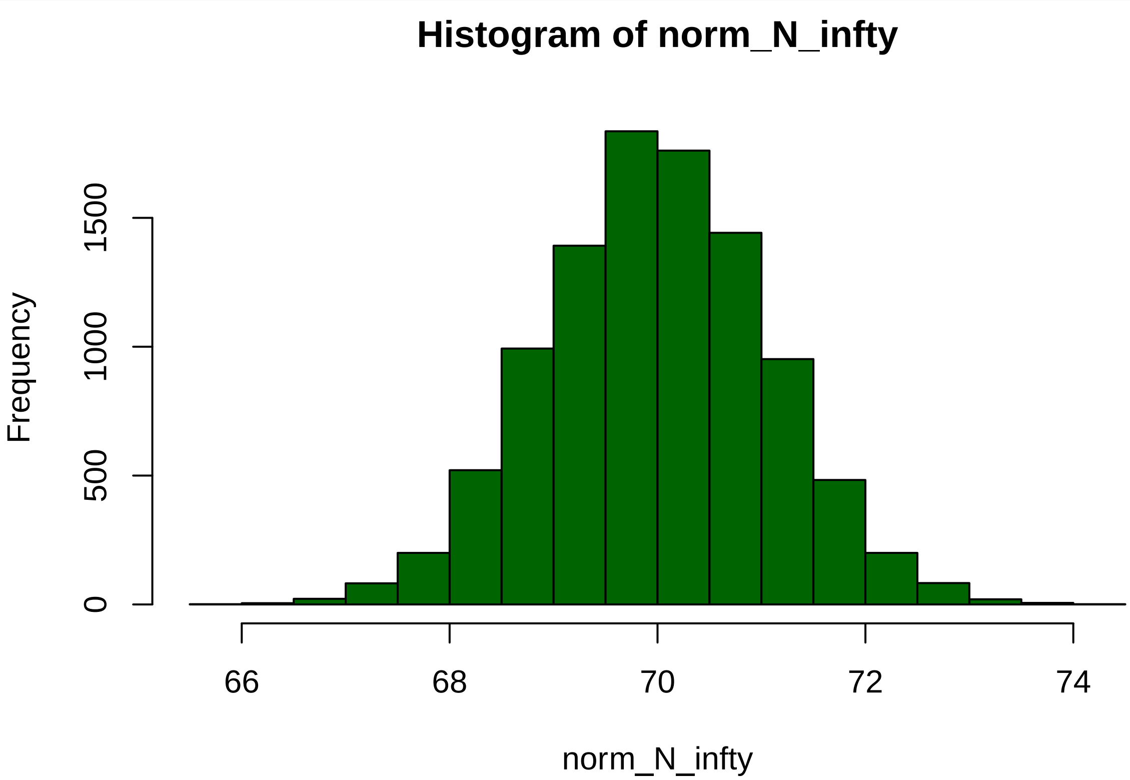
```
var_norm = var(norm_N_infty)
var_norm = sd ^ 2
var_norm
```

```
## [1] 1.2
```

La varianza aproxima

- Realizar un histograma de los datos. ¿Qué observa? ¿Es consistente con la distribución a partir de la cual fueron generados los datos?

```
hist(norm_N_infty, col = "darkgreen")
```



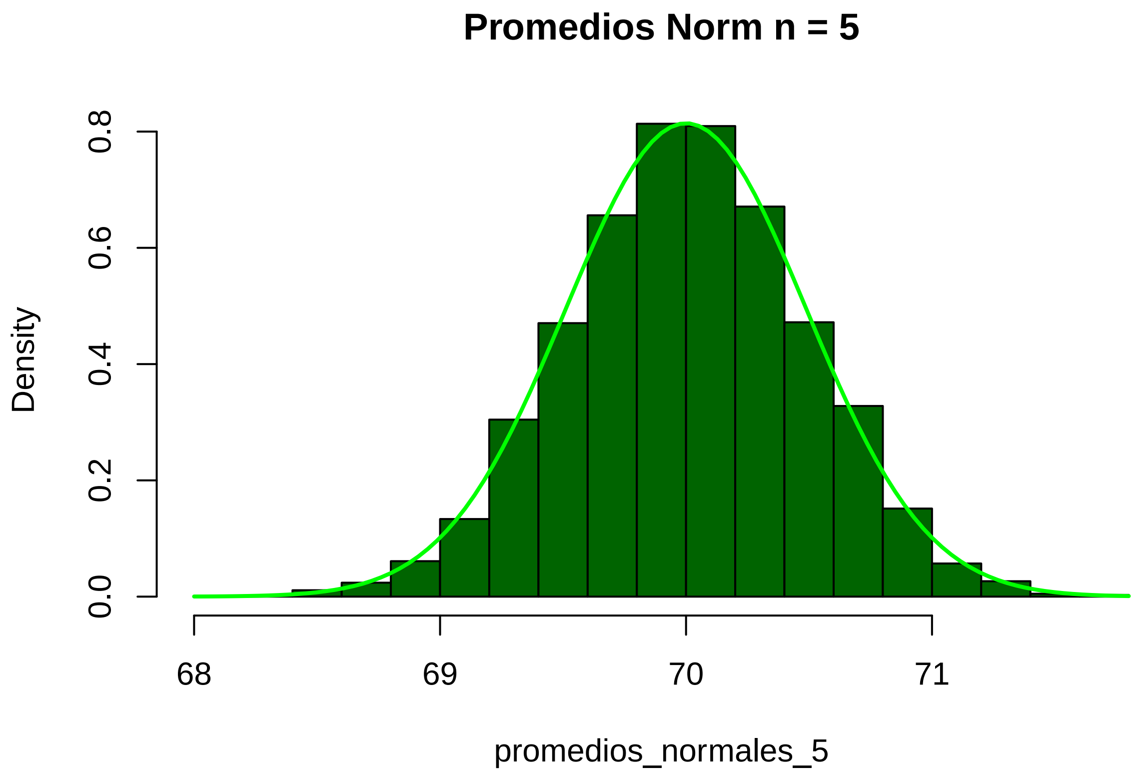
La distribución empírica del promedio

- Guardar en el vector `promedios_normales`, $Nrep = 10000$ promedios utilizando $n = 5$ datos con distribución X.

```
n_dist = 5
promedios_normales_5 = replicate(Nrep, mean(rnorm(n_dist, mu, sd)))
```

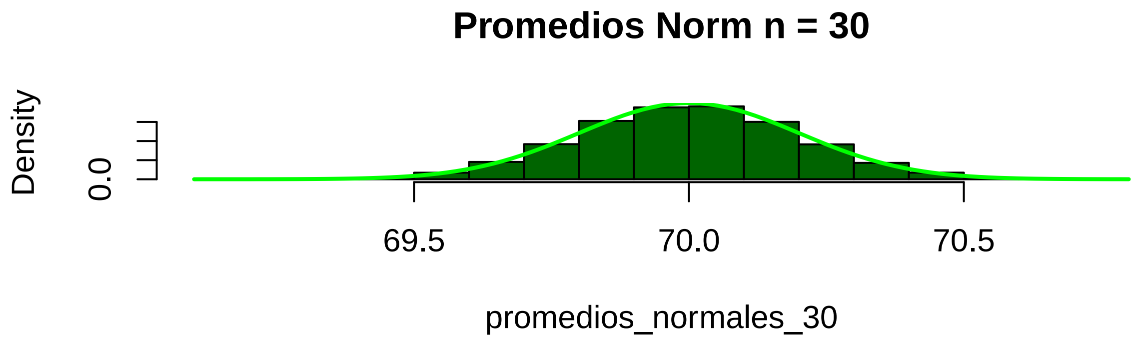
- Con los promedios guardados en promedios bernoullies, realizar un histograma de densidad. Para el histograma de densidad, usar el argumento `probability = TRUE` en el comando `hist()`. Además, como n = 5 puede ser pequeño, al usar `hist()` puede ajustar el parámetro `breaks`. Observar qué ocurre al poner (o no) ese parámetro. Luego, superponer el gráfico de la curva de la densidad normal con la media y desvío que considere razonables.

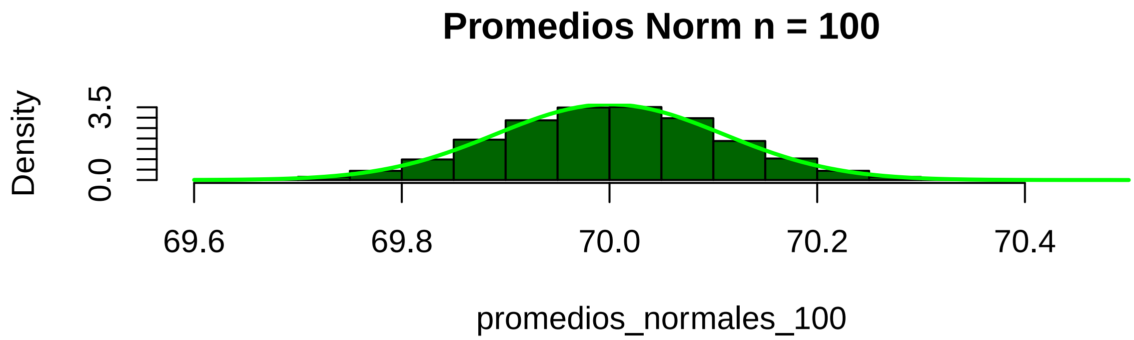
```
n_dist = 5
hist(main = "Promedios Norm n = 5", promedios_normales_5, probability = TRUE, col = "darkgreen")
curve(dnorm(x, e_norm, sqrt( var_norm/ n_dist)), add=TRUE, lwd=2, col="green")
```



- Repetir los ítems 2 y 3 con $n = 30$ y con $n = 100$. Comparar los tres histogramas obtenidos y explicar qué observa en cada caso y a qué cree que se debe. Tener cuidado con la diferencia en las escalas.

```
n_dist = 30
promedios_normales_30 = replicate(Nrep, mean(rnorm(n_dist, mu, sd)))

par(mfrow=c(2, 1))
hist(main = "Promedios Norm n = 30", promedios_normales_30, probability = TRUE, col = "darkgreen")
curve(dnorm(x, e_norm, sqrt( var_norm/ n_dist)), add=TRUE, lwd=2, col="green")

n_dist = 100
promedios_normales_100 = replicate(Nrep, mean(rnorm(n_dist, mu, sd)))
hist(main = "Promedios Norm n = 100", promedios_normales_100, probability = TRUE, col = "darkgreen")
curve(dnorm(x, e_norm, sqrt( var_norm/ n_dist)), add=TRUE, lwd=2, col="green")
```



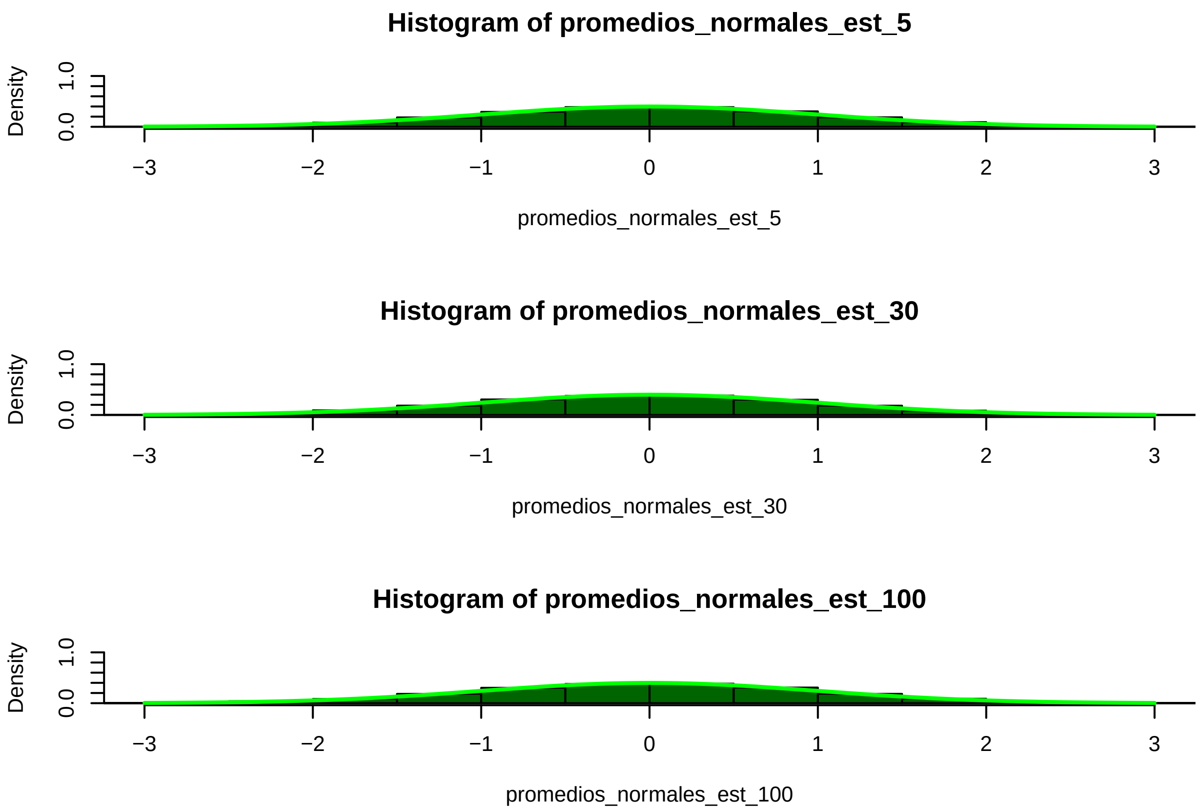



```
par(mfrow=c(1, 1))
```

La distribución empírica del promedio estandarizado

- Realizar ahora histogramas de densidad con los datos de `promedios_normales_est` donde a los valores guardados en `promedios_normales` se les resta μ y se los divide por $\sqrt{(\sigma^2/n)}$. Utilizar $n \in \{5, 30, 100\}$. ¿Qué observa? Explicar.

```
promedios_normales_est_5 = (promedios_normales_5 - e_norm) / sqrt(var_norm / 5)

par(mfrow=c(3, 1))
hist(promedios_normales_est_5, xlim = c(-3, 3), ylim = c(0, 1), probability = TRUE, col = "darkgreen")
curve(dnorm(x,0,1), add=TRUE, lwd=2, col="green")

promedios_normales_est_30 = (promedios_normales_30 - e_norm) / sqrt(var_norm / 30)
hist(promedios_normales_est_30, xlim = c(-3, 3), ylim = c(0, 1), probability = TRUE, col = "darkgreen")
curve(dnorm(x,0,1), add=TRUE, lwd=2, col="green")

promedios_normales_est_100 = (promedios_normales_100 - e_norm) / sqrt(var_norm / 100)
hist(promedios_normales_est_100, xlim = c(-3, 3), ylim = c(0, 1), probability = TRUE, col = "darkgreen"
curve(dnorm(x,0,1), add=TRUE, lwd=2, col="green")
```




```
par(mfrow=c(1, 1))
```

- A los histogramas de densidad del ítem anterior, superponer las curvas de la densidad normal con la media y desvío que considere adecuados. ¿Qué observa? Explicar.

  Con $n \to \infty$, el histograma de densidad de la estimación de la normal apartir de la muestra con elementos Normales es exactamente la curva de la normal $N(0, 1)$

- Explicar las diferencias que puede haber en las conclusiones para este apartado (distribución normal) y las de las distribuciones anteriores (Bernoulli, uniforme y exponencial).