Teorema Central del Límite

Dr. Delfino Vargas Chanes

Facultad de Economía
Universidad Nacional Autónoma de México

14 de abril de 2024

Índice

Convergencia Puntual

Definición 1.1

Sea un espacio de probabilidad $(\Omega, \mathcal{F}, P)$. Diremos que la sucesión de VA (que son funciones) X_n converge puntualmente a X si

$$P\left(\lim_{n\to\infty} X_n = X\right) = 1$$

- Ustedes tienen la posibilidad de hacer un examen de Estadística infinitamente. Deben de hacerlo hasta sacar 10. Sabemos que la sucesión de VA que son sus examenes converge puntualmente a 10.
- Sin embargo, si tomamos un número finito de estos ensayos, existe la posibilidad de que no ocurra la condición final (sacar 10)

Convergencia en Probabilidad

Definición 1.2

Sea un espacio de probabilidad $(\Omega, \mathcal{F}, P)$. Diremos que la sucesión de VA X_n converge a la VA X si $\forall\, \varepsilon > 0$ s.t.q.

$$\lim_{n\to\infty} P\left(|X_n - X| > \varepsilon\right) = 0$$

- Es decir, cuando $n \to \infty$ la probabilidad de que la sucesión de VA X_n este *lejos* de X es nula.
- Podemos denotar la convergencia en probabilidad como

$$X_n \xrightarrow{p} X$$

Ley de Grandes Números

- Sigamos con los dos teoremas más importantes de la probabilidad.
- Para estos dos teoremas vamos a suponer que tenemos $X_1, X_2, \ldots$ variables aleatorias independientes e identicamente distribuidas con la siguiente media muestral:

$$\overline{X}_n = \frac{X_1 + \ldots + X_n}{n} \qquad \forall\, n \in \mathbb{N}$$

- Sabemos que la media poblacional de las VA es μ y su varianza es σ^2.

$$\mathbb{E}[\overline{X}_n] = \frac{\mathbb{E}[X_1 + \ldots + X_n]}{n} = \frac{\mathbb{E}[X_1] + \ldots + \mathbb{E}[X_n]}{n} = \frac{n \cdot \mu}{n} = \mu$$

$$\mathrm{Var}(\overline{X}_n) = \frac{\mathrm{Var}[X_1 + \ldots + X_n]}{n^2} = \frac{\mathrm{Var}[X_1] + \ldots + \mathrm{Var}[X_n]}{n^2}$$

$$= \frac{n \cdot \sigma^2}{n^2} = \frac{\sigma^2}{n}$$

Ley de Grandes Números

- La media muestral es una VA por si misma, y ya vimos que su esperanza es μ y su varianza es $\frac{\sigma^2}{n}$
- La Ley de Grandes Números (LLN) nos dice que la media poblacional $\overline{X}_n$ **converge** puntualmente a la media poblacional *verdadera* μ
- La LLN tiene su versión fuerte y débil.
- A continuación definiremos a ambas y trataremos de dar un ejemplo.

LLN Fuerte

Teorema 2.1 (Ley de Kolmogorov)

La media muestral $\overline{X}_n$ converge puntualmente a la esperanza, o media poblacional $\mathbb{E}[\overline{X}_n] = \mu$. Recordando que las VA son funciones de $\Omega \to \mathbb{R}$

Cuando $n \to \infty$

$$\forall s \in \Omega \Rightarrow \overline{X}_n(s) \to \mu$$

Es decir

$$P\left(\lim_{n\to\infty} \overline{X}_n = \mu\right) = 1$$

LLN **Débil**

Teorema 2.2 (Ley de Khinchin)

Cuando $n \to \infty$

$$\overline{X}_n \xrightarrow{p} X$$

Es decir, $\forall \, \varepsilon > 0$

$$\lim_{n \to \infty} P\left(\left|\overline{X}_n - \mu\right|\right) = 0$$

Desigualdad de Chebyshev

Teorema 2.3

Sea X una VA con media poblacional μ y varianza $\sigma^2 \Rightarrow \forall\, k > 0$

$$P\left(|X - \mu| \geqslant k\right) \leqslant \frac{\sigma^2}{k^2}$$

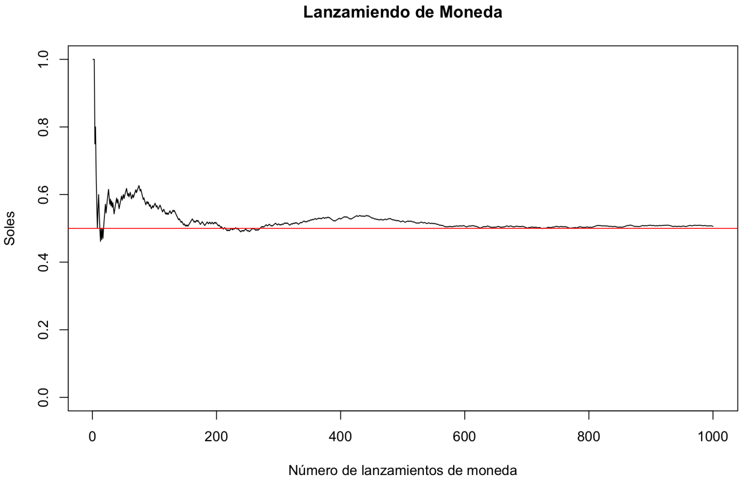
Demostración.

Usando la desigualdad de Chebyshev es muy facil probar la LLN **débil**. Supongamos un $\varepsilon > 0$ fijo pero arbitrario

$$P\left(\left|\overline{X}_n - \mu\right| > \varepsilon\right) \leqslant \frac{\sigma^2}{n \cdot \varepsilon^2}$$

Cuando $n \to \infty$ el lado derecho tiende a cero. $\square$

Ejemplo en R



Lanzamiendo de Moneda
Soles
0.0
0.2
0.4
0.6
0.8
1.0
0
200
400
600
800
1000
Número de lanzamientos de moneda

Código en R

```
set.seed(123)
    flips <- sample(c("heads", "tails"), size = 1000,
          replace = TRUE)
     plot(cumsum(flips == "heads") / (1:length(flips)),
     type = "l", ylim = c(0,1),
     main = "Lanzamiendo de Moneda",
     xlab = "Numero de lanzamientos de moneda", ylab = "
Soles")
abline(h = 0.5, col = "red")
```

Recordatorio Tema 5: La variable estandarizada

Definición 3.1

Si X es una variable aleatoria tal que $X \sim \mathcal{N}(\mu, \sigma^2)$, entonces la variable aleatoria Z se llama variable aleatoria estandarizada

$$X = \frac{X - \mu}{\sigma} \sim \mathcal{N}(0, 1)$$

Teorema Central del Límite

Teorema 3.1 (Teorema Central del Límite)

Sea $(\Omega, \mathcal{F}, P)$ un espacio muestral. Supongamos que tenemos $X_1, X_2, \ldots \in \mathcal{F} \subseteq \Omega$ variables aleatorias independientes e identicamente distribuidas con media poblacional μ y varianza σ^2, y media muestral $\overline{X}_n$

Cuando $n \to \infty$

$$\sqrt{n} \cdot \left(\frac{\overline{X}_n - \mu}{\sigma} \right) \to \mathcal{N}(0, 1)$$

Es decir, cuando n se hace **MUY** grande, cuando estandarizamos la distribución de la VA $\overline{X}_n$ se acerca a una distribución normal estándar.

Teorema Central del Límite (Aproximación)

Teorema 3.2

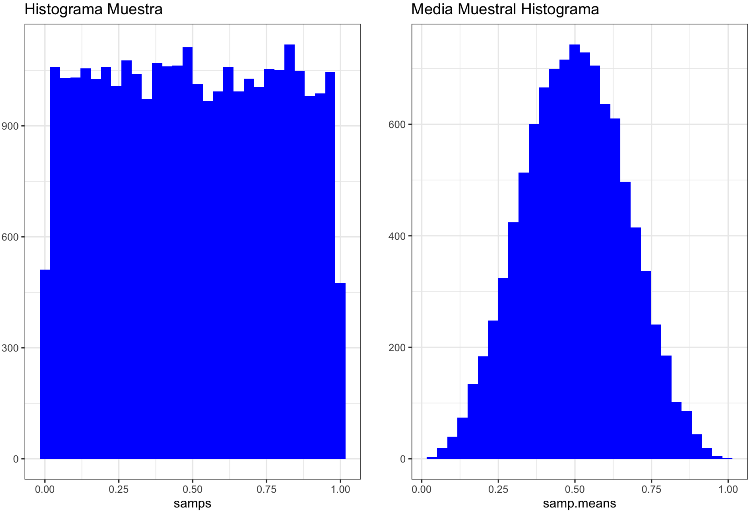
Para grandes n, la distribución de $\overline{X}_n$ es **aproximadamente**

$$\mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right)$$

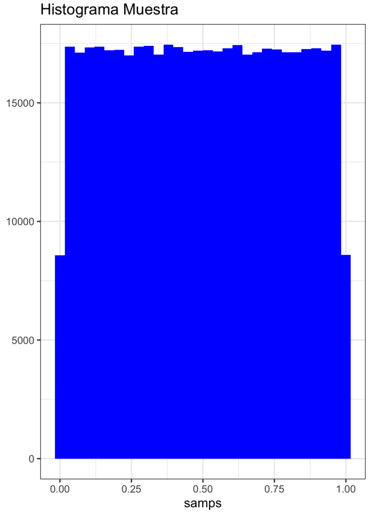
Simulaciones en R

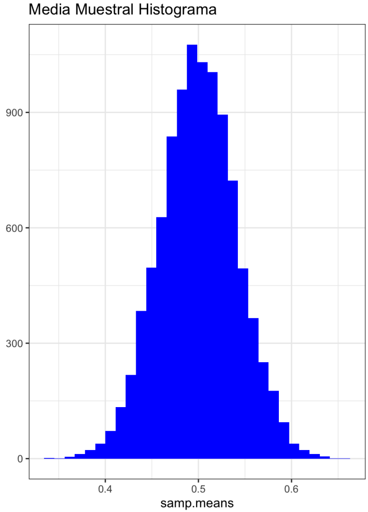
- Veamos que, en general, podemos observar que se cumple el TCL para las distribuciones que hemos estudiado en clase.
- Vamos a graficar el histograma de varias VA $\overline{X}_n$ para distintos valores de $n \in \mathbb{N}$
- Primero, para $n = 1$, se grafica la típica PDF que todos conocemos para cada distribución
- Luego, veremos que pasa con $n = 3$, con $n = 50$ y $n = 300$
- Veremos que se cumple gráficamente el Teorema 3.2
- Es decir, vamos a tener una muestra, y veremos que la Media Muestral como VA tiene una distribución que se acerca a la normal

Distribución Uniforme con $n = 3$

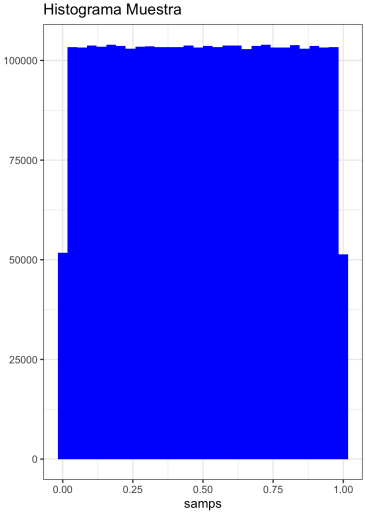


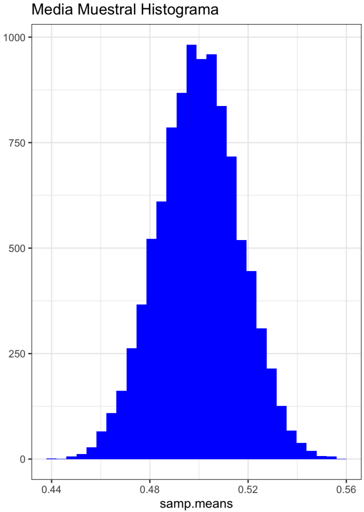
Distribución Uniforme con $n = 50$



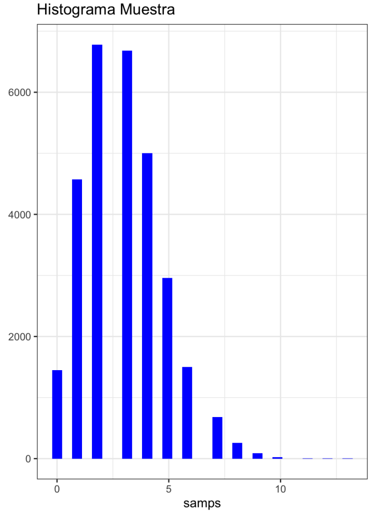


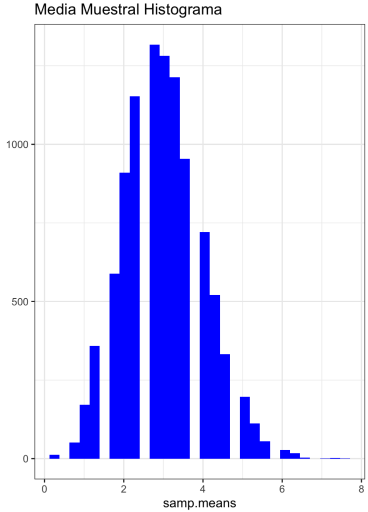
Distribución Uniforme con $n = 500$



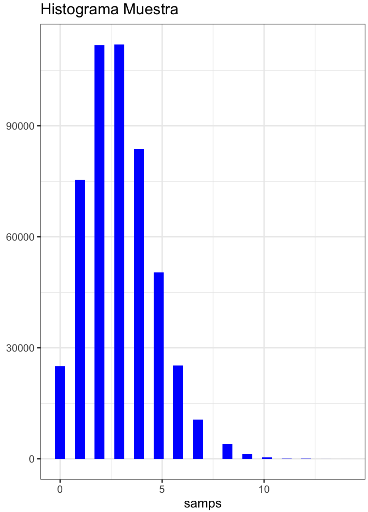


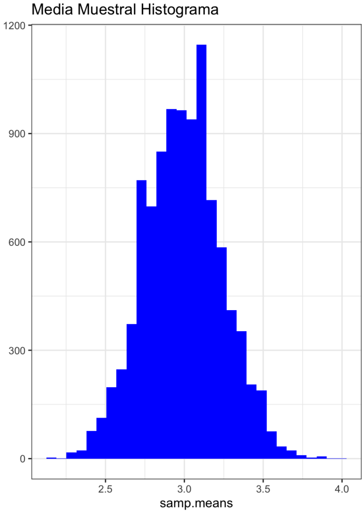
Distribución Poisson con $n = 3$



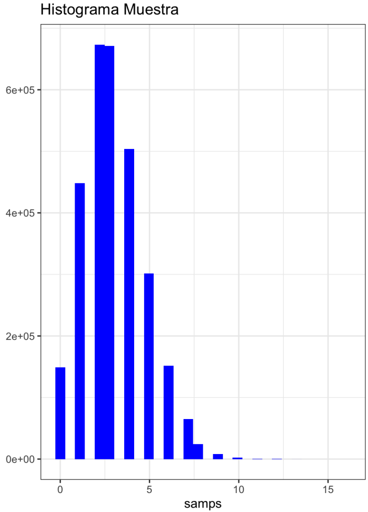


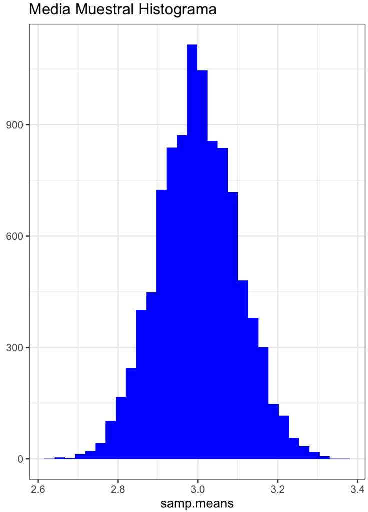
Distribución Poisson con $n = 50$



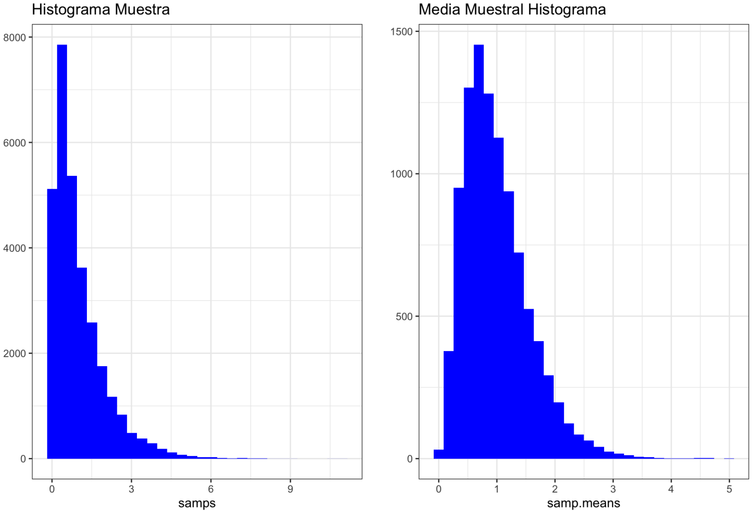


Distribución Poisson con $n = 300$

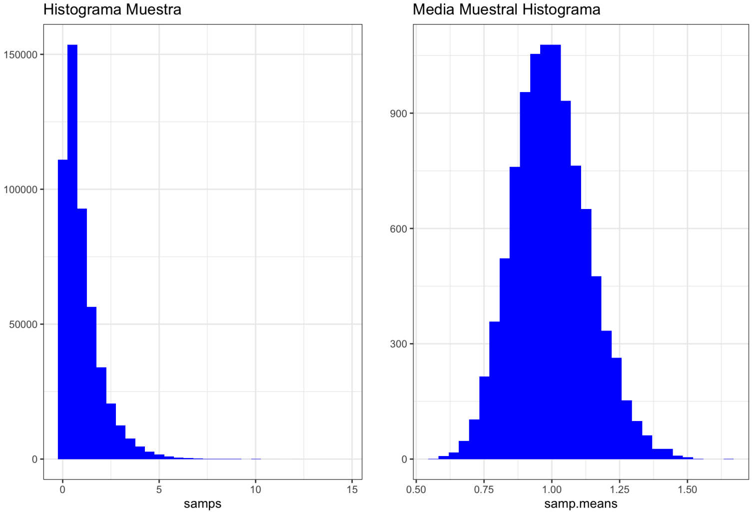




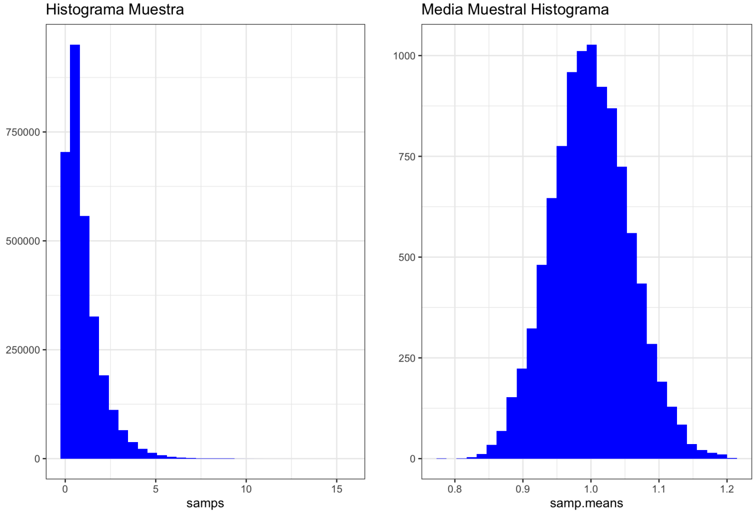
Distribución Exponencial con $n = 3$



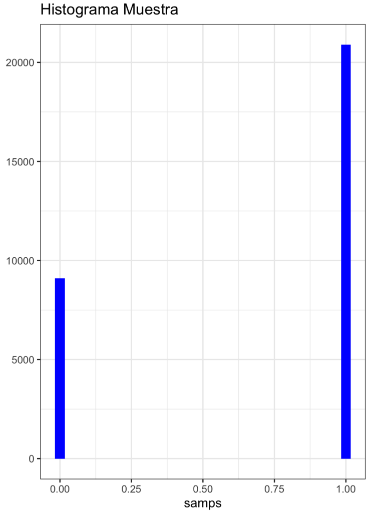
Distribución Exponencial con $n = 50$

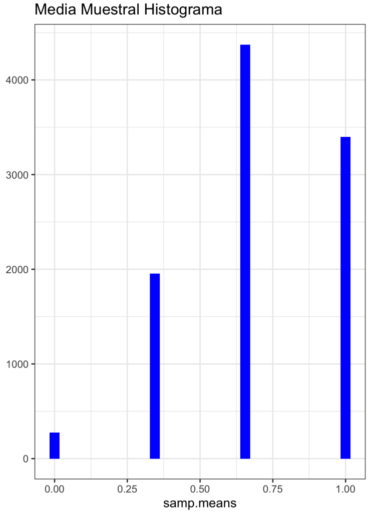


Distribución Exponencial con $n = 300$

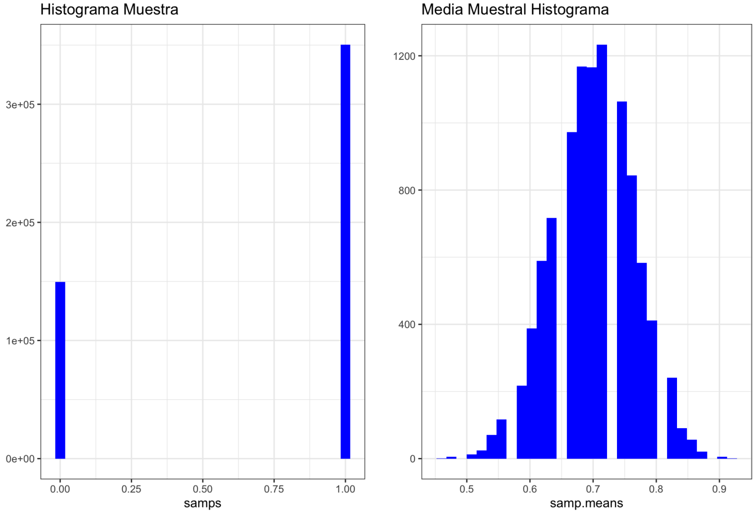


Distribución Binomial con $n = 3$

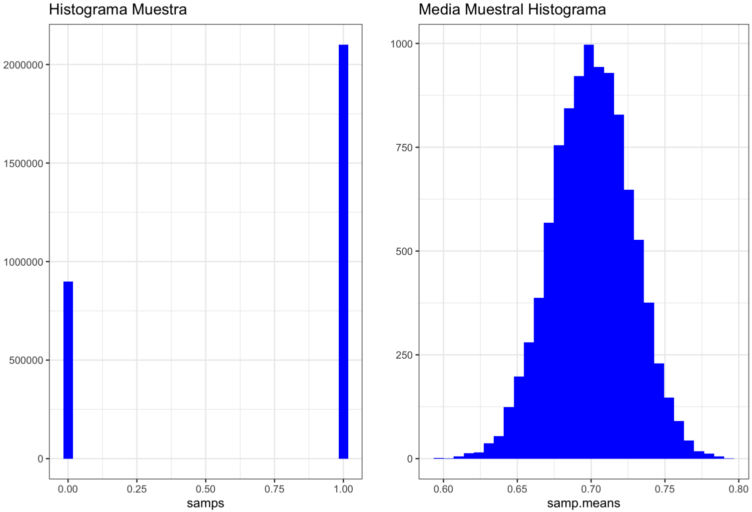




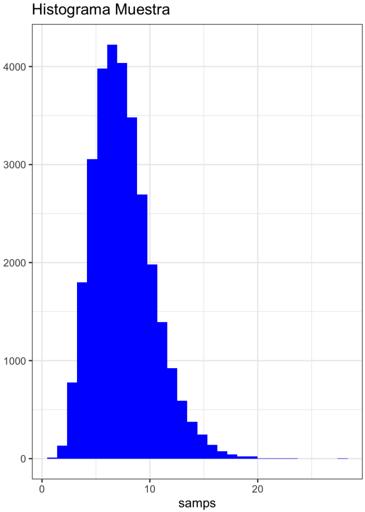
Distribución Binomial con $n = 50$

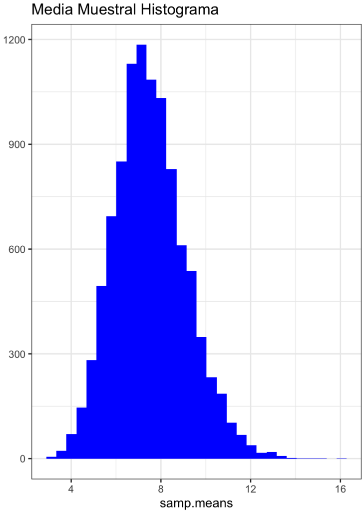


Distribución Binomial con $n = 300$

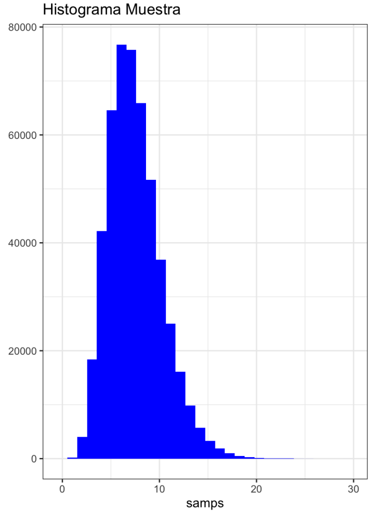


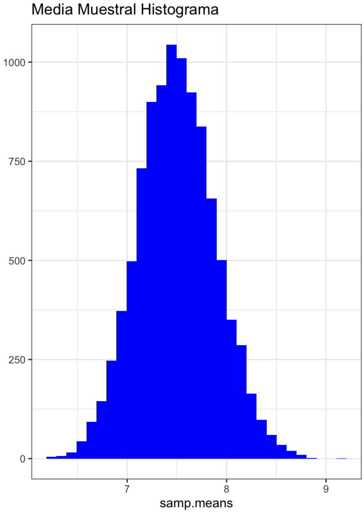
Distribución Gamma con $n = 3$



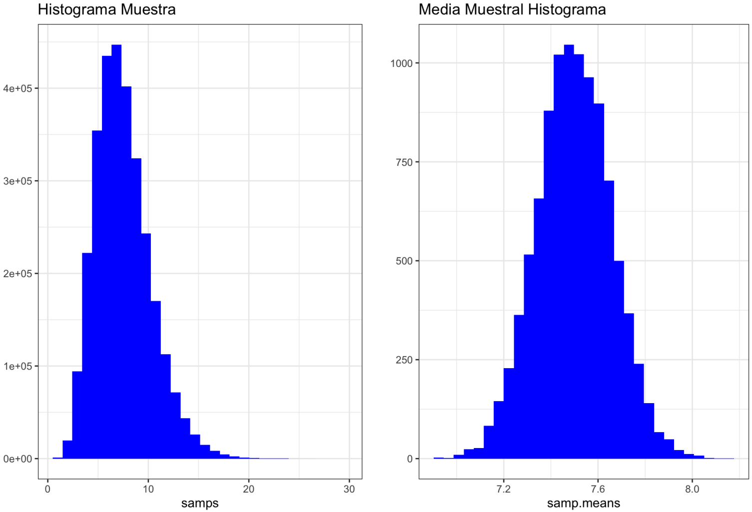


Distribución Gamma con $n = 50$

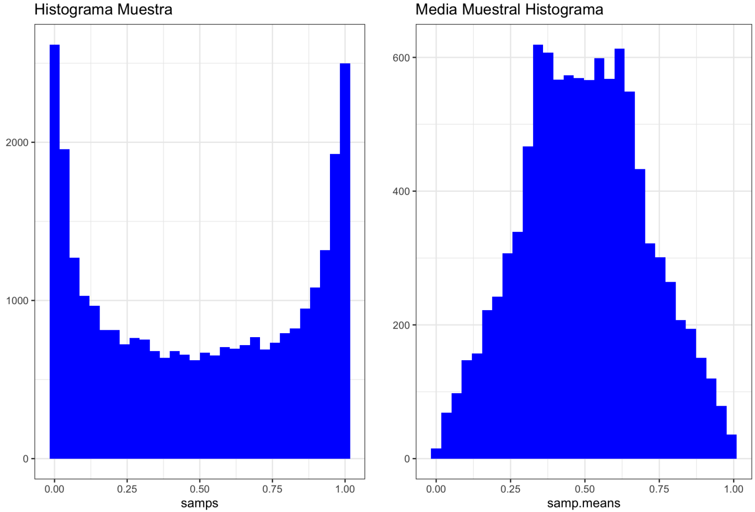




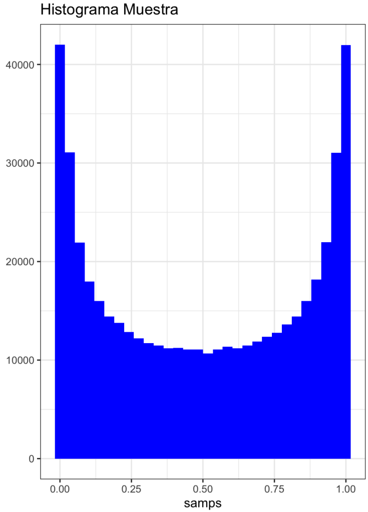
Distribución Gamma con $n = 300$

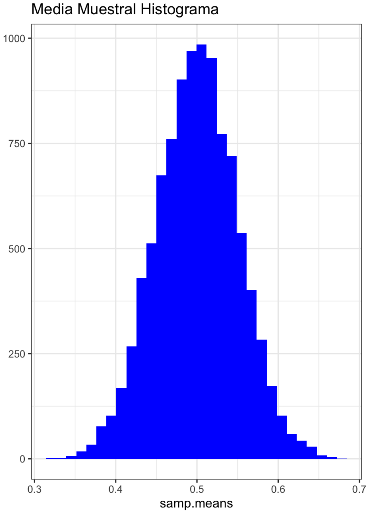


Distribución Beta con $n = 3$

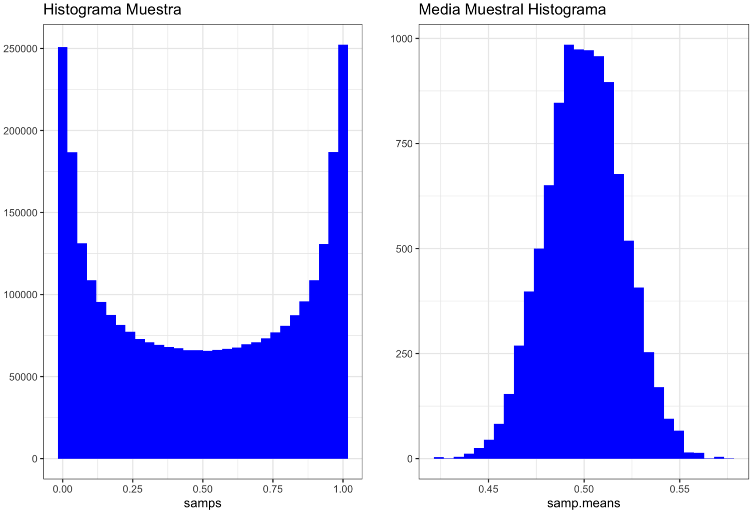


Distribución Beta con $n = 50$





Distribución Beta con $n = 300$



Bibliografía

Bernoulli, J. (1713). *Ars coniectandi*. Impensis Thurnisiorum, fratrum.

Bhattacharyya, G. K., y Johnson, R. A. (1977). *Statistical concepts and methods*. Wiley.

Bickel, P. J., y Doksum, K. A. (2015). *Mathematical statistics: basic ideas and selected topics, volumes i-ii package*. CRC Press.

Fair, R. C. (1978). A theory of extramarital affairs. *Journal of political economy*, *86*(1), 45–61.

Groeneveld, R. A. (1998). A class of quantile measures for kurtosis. *The American Statistician*, *52*(4), 325–329.

Wackerly, D., Mendenhall, W., y Scheaffer, R. L. (2014). *Mathematical statistics with applications*. Cengage Learning.

¡Gracias por su atención!