Ley de los Grandes Números

Mariela Sued, Jemina García y Ana M. Bianco

03 junio 2020

Un caso especial: muestra aleatoria (m.a.)

Un caso especial: muestra aleatoria (m.a.)

1. Repetimos un experimento muchas veces en idénticas condiciones y de manera independiente.
2. X_i: resultado de la i-ésima repetición.
3. $X_i \sim F$: tienen todas la misma distribución porque repetimos en idénticas condiciones.
4. $X_1, \ldots, X_n$ son independientes (estadísticamente) por construcción: las repeticiones se realizan de manera independiente (en sentido de la vida).

Un caso especial: muestra aleatoria (m.a.)

1. Repetimos un experimento muchas veces en idénticas condiciones y de manera independiente.

2. X_i: resultado de la i-ésima repetición.

3. $X_i \sim F$: tienen todas la misma distribución porque repetimos en idénticas condiciones.

4. $X_1, \ldots, X_n$ son independientes (estadísticamente) por construcción: las repeticiones se realizan de manera independiente (en sentido de la vida).

Luego, $X_i \sim F$ para todo i, y por consiguiente:

- $\mathbb{P}(X_i \leq t) = \mathbb{P}(X_j \leq t) = \mathbb{P}(X_1 \leq t)$
- $\mathbb{E}[X_i] = \mathbb{E}[X_j] = \mathbb{E}[X_1]$
- $\mathbb{V}(X_i) = \mathbb{V}(X_j) = \mathbb{V}(X_1)$

- Dicho de otra forma: mongo(X_i)=mongo(X_j)=mongo(X_1).

Un caso especial: muestra aleatoria (m.a.)

1. Repetimos un experimento muchas veces en idénticas condiciones y de manera independiente.
2. X_i: resultado de la i-ésima repetición.
3. $X_i \sim F$: tienen todas la misma distribución porque repetimos en idénticas condiciones.
4. $X_1, \ldots, X_n$ son independientes (estadísticamente) por construcción: las repeticiones se realizan de manera independiente (en sentido de la vida).

Luego, $X_i \sim F$ para todo i, y por consiguiente:

- $\mathbb{P}(X_i \leq t) = \mathbb{P}(X_j \leq t) = \mathbb{P}(X_1 \leq t)$
- $\mathbb{E}[X_i] = \mathbb{E}[X_j] = \mathbb{E}[X_1]$
- $\mathbb{V}(X_i) = \mathbb{V}(X_j) = \mathbb{V}(X_1)$

- Dicho de otra forma:
  mongo(X_i)=mongo(X_j)=mongo(X_1).

Muestra (aleatoria) o variables i.i.d.

$X_1, \dots, X_n$ son una muestra aleatoria si son v. a. independientes, idénticamente distribuídas.

$$X_1, \dots, X_n \text{ , i.i.d.}$$

En tal caso, $X_i \sim F$ para todo i, y por consiguiente,

- mongo(X_i)=mongo(X_j)=mongo(X_1).

Muestra (aleatoria) o variables i.i.d.

¿Qué relación guarda la muestra con nuestros datos?

Una noche en el casino: Ruleta

¿Apostamos una fichita?

- Si acierto: me devuelven 36 fichas (35+la mía)

- Si pierdo: se llevan mi ficha

¿Apostamos una fichita?

- Si acierto: me devuelven 36 fichas (35+la mía)
- Si pierdo: se llevan mi ficha

Dicho de otro modo:

- Si acierto: Gano 35
- Si pierdo: Gano -1

Una larga noche . . .

[1] −1 −1 35 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1

[20] −1 −1 −1 −1 −1 −1 −1 −1 35 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1

[39] −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1

[58] −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1

[77] −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1

[96] −1 −1 −1 −1 35 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1

[115] −1 −1 −1 −1 −1 −1 35 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1

[134] −1 −1 −1 −1 −1 −1 −1 35 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1

[153] −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1

estos son los resultados...

...ésta es la historia de lo que pasó

. . . pensemos. . .

Podríamos decir que los datos que les doy son

- el esultado o lo que observamos en cada jugada:

$$x_1 = -1, x_2 = -1, x_3 = 35, \ldots, x_{27} = -1, \ldots, x_{28} = 35, x_{29} = -1$$

Estos datos, ¿son los resultados o las observaciones de qué?

. . . pensemos. . .

- $x_1 = -1$ es el valor observado de la *ganancia obtenida en la primera jugada* o sea de la v.a. X_1

. . . pensemos. . .

- $x_1 = -1$ es el valor observado de la *ganancia obtenida en la primera jugada* o sea de la v.a. X_1
- $x_2 = -1$ es el valor observado de la *ganancia obtenida en la segunda jugada* o de sea la v.a. X_2

...pensemos...

- $x_1 = -1$ es el valor observado de la *ganancia obtenida en la primera jugada* o sea de la v.a. X_1
- $x_2 = -1$ es el valor observado de la *ganancia obtenida en la segunda jugada* o de sea la v.a. X_2
- $x_3 = 35$ es el valor observado de la *ganancia obtenida en la tercera jugada* o de sea la v.a. X_3
- .....
- $x_i = -1$ es el valor observado de la *ganancia obtenida en la i–ésima jugada* o sea de la v.a. X_i
- ....
- $X_1, X_2, \ldots, X_i \ldots, X_n$ son variables aleatorias idénticamente distribuidas

. . . pensemos. . .

- $x_1 = -1$ es el valor observado de la *ganancia obtenida en la primera jugada* o sea de la v.a. X_1
- $x_2 = -1$ es el valor observado de la *ganancia obtenida en la segunda jugada* o de sea la v.a. X_2
- $x_3 = 35$ es el valor observado de la *ganancia obtenida en la tercera jugada* o de sea la v.a. X_3
- .....
- $x_i = -1$ es el valor observado de la *ganancia obtenida en la i–ésima jugada* o sea de la v.a. X_i
- ....
- $X_1, X_2, \ldots, X_i \ldots, X_n$ son variables aleatorias idénticamente distribuidas
- Por las condiciones de *este experimento* las X_i resultan independientes.

$$X_1, X_2, \ldots, X_i \ldots, X_n \text{ i.i.d.}$$

Para pensar: ¿Siempre es así?

Supongamos que en una caja tenemos 100 bolitas: 90 blancas y 10 rojas. Consideremos las siguientes dos situaciones:

- extraemos 5 bolitas **con reposición**. Para cada extracción definimos $1 \leq i \leq 5$:

$$X_i = \begin{cases} 1 & \text{si la i-ésima bolita extraída es roja} \\ 0 & \text{si la i-ésima bolita extraída es blanca} \end{cases}$$

¿Son las X_i v.a. i.i.d.? ....

Para pensar: ¿Siempre es así?

Supongamos que en una caja tenemos 100 bolitas: 90 blancas y 10 rojas. Consideremos las siguientes dos situaciones:

- extraemos 5 bolitas **con reposición**. Para cada extracción definimos $1 \leq i \leq 5$:

$$X_i = \begin{cases} 1 & \text{si la i-ésima bolita extraída es roja} \\ 0 & \text{si la i-ésima bolita extraída es blanca} \end{cases}$$

¿Son las X_i v.a. i.i.d.? ....

- extraemos 5 bolitas **sin reposición**. Para cada extracción definimos $1 \leq i \leq 5$:

$$Y_i = \begin{cases} 1 & \text{si la i-ésima bolita extraída es roja} \\ 0 & \text{si la i-ésima bolita extraída es blanca} \end{cases}$$

¿Son las Y_i v.a. independientes? ....

Para seguir pensando...

¿Qué relación guarda todo esto con nuestras simulaciones?

Para seguir pensando. . .

¿Qué relación guarda todo esto con nuestras simulaciones?

¿Qué estamos haciendo con el comando `runif(10,0,1)`?

Volviendo a la muestra

$X_1, \ldots, X_n$ i.i.d., con $\mathbb{E}[X_i] = \mu$ y $\mathbb{V}[X_i] = \sigma^2$, para todo i.

- $S_n = \sum_{i=1}^{n} X_i \qquad \Longrightarrow \mathbb{E}[S_n] = n\,\mu \qquad$ y $\qquad \mathbb{V}[S_n] = n\,\sigma^2$

- $\overline{X} = \overline{X}_n = \dfrac{\sum_{i=1}^{n} X_i}{n} \qquad \Longrightarrow \mathbb{E}[\overline{X}_n] = \mu \qquad$ y $\qquad \mathbb{V}[\overline{X}_n] = \dfrac{\sigma^2}{n}$

Caso Particular: Suma y Promedio de normales

- $(X_i)_{1 \leq i \leq n}$ i.i.d., $X_i \sim \mathcal{N}(\mu, \sigma^2)$
- Suma: $S_n = X_1 + \ldots + X_n \sim ?$
- Pomedio: $\overline{X}_n = \frac{1}{n}(X_1 + \ldots + X_n) \sim ?$

Vayamos a la lista de tareas

Simulemos:

Estudiaremos la distribución de S_n y de $\overline{X}_n$ de variables $X_1, \ldots, X_n$ i.i.d. $\mathcal{N}(\mu, \sigma^2)$.

Para ello, fijado n, generaremos datos correspondientes a una muestra $X_1, \ldots, X_n$ i.i.d. de variables aleatorias distribuidas como X, con una distribución $\mathcal{N}(5, 4)$ y luego calcularemos la suma y el promedio de cada conjunto de datos.

Repetimos este procedimiento $Nrep = 1000$ veces. A partir de las $Nrep = 1000$ replicaciones realizaremos un histograma con las sumas y los promedios generados, para obtener una aproximación de la densidad de S_n y de $\bar{X}_n$.

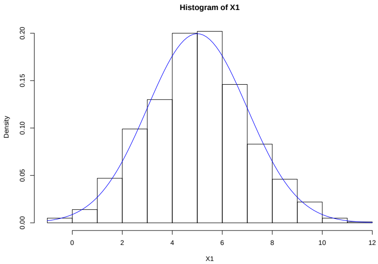
Para ello, fijado n, generaremos datos correspondientes a una muestra $X_1, \ldots, X_n$ i.i.d. de variables aleatorias distribuidas como $X \sim \mathcal{N}(5, 4)$, para $n = 1, 2, 5, 10, 25$. Luego, haremos un histograma para luego tratar de responder a qué densidad se parece el histograma obtenido y superponerle una densidad adecuada.

$n = 1$

- Consideramos $n = 1$ en cuyo caso la variable coincide con la suma y el promedio.

Generamos entonces Nrep =1000 datos correspondientes a $X_1 \sim \mathcal{N}(5, 4)$ y luego hacemos un histograma.

```
Nrep<- 1000
set.seed(123)
X1<-rnorm(Nrep,mean=5,sd=2)    # Genero normales
hist(X1,freq=F)
curve(dnorm(x, mean=5,sd=2), add=T, col="blue")
```



$n = 2$

- Consideramos $n = 2$ y las variables

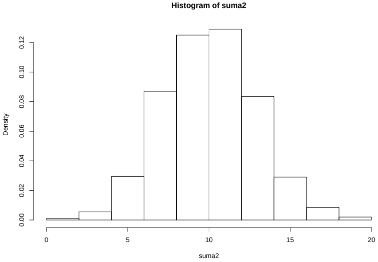
$$S_2 = X_1 + X_2 \qquad \text{y} \qquad \bar{X}_2 = \frac{X_1 + X_2}{2}.$$

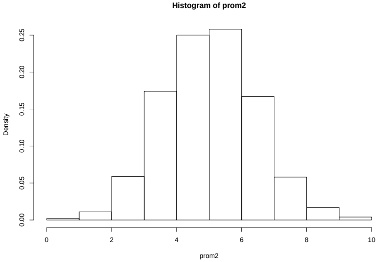
Generamos $n = 2$ datos (independientes) correspondientes a variables aleatorias con distribución $\mathcal{N}(5, 4)$ y computamos la suma y el promedio. Replicamos $Nrep = 1000$ veces y realizamos los dos histograma con los $Nrep = 1000$ promedios obtenidos.

```r
Nrep<- 1000
X1<-rnorm(Nrep,mean=5,sd=2)   # Genero normales
X2<-rnorm(Nrep,mean=5,sd=2)   # Genero normales
suma2<- X1+X2
prom2<- suma2/2
```

Graficamos

```r
hist(suma2,freq=F)
hist(prom2,freq=F)
```



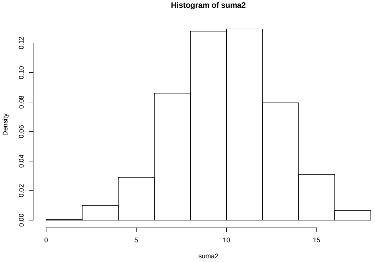


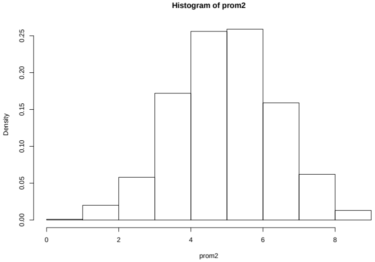
Otra forma

```
ene=2
matriz2<-matrix(rnorm(1000*ene,mean=5,sd=2),nrow=ene, ncol=1000,byrow=T)
suma2<-apply(matriz2,2,sum)
prom2<-apply(matriz2,2,mean)
```

Grafico

```r
hist(suma2,freq=F)
hist(prom2,freq=F)
```





Suma y Promedio de normales

$X_1, \ldots, X_n$ i.i.d., $X_i \sim \mathcal{N}(\mu, \sigma^2)$

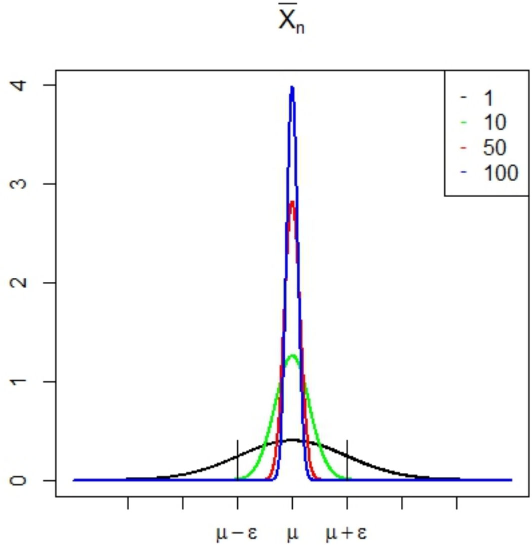
- Suma: $S_n = \sum_{i=1}^{n} X_i \sim \mathcal{N}(n\mu, n\sigma^2)$

Suma y Promedio de normales

$X_1, \ldots, X_n$ i.i.d., $X_i \sim \mathcal{N}(\mu, \sigma^2)$

- Suma: $S_n = \sum_{i=1}^{n} X_i \sim \mathcal{N}(n\mu, n\sigma^2)$
- Pomedio: $\overline{X}_n = \dfrac{\displaystyle\sum_{i=1}^{n} X_i}{n} \sim \mathcal{N}(\mu, \sigma^2/n)$

Comparando las densidades $N(\mu, \sigma^2/n)$



Suma y Promedio de normales

$X_1, \ldots, X_n$ i.i.d., $X_i \sim \mathcal{N}(\mu, \sigma^2)$

- Suma: $S_n = \sum_{i=1}^{n} X_i \sim \mathcal{N}(n\mu, n\sigma^2)$

Suma y Promedio de normales

$X_1, \ldots, X_n$ i.i.d., $X_i \sim \mathcal{N}(\mu, \sigma^2)$

- Suma: $S_n = \sum_{i=1}^{n} X_i \sim \mathcal{N}(n\mu, n\sigma^2)$
- Pomedio: $\overline{X}_n = \dfrac{\displaystyle\sum_{i=1}^{n} X_i}{n} \sim \mathcal{N}(\mu, \sigma^2/n)$

$$\mathbb{P}(\left(|\overline{X}_n - \mu| > \varepsilon\right) = 2(1 - \Phi(\varepsilon\sqrt{n}/\sigma))$$

donde $\phi(z) = \mathbb{P}(Z \leq z),\ Z \sim \mathcal{N}(0, 1)$

En otras palabras, estamos viendo que

Si $X_1, \ldots, X_n$ i.i.d., $X_i \sim \mathcal{N}(\mu, \sigma^2)$, para todo $\varepsilon > 0$, tenemos que

$$\mathbb{P}\left(|\overline{X}_n - \mu| > \varepsilon\right) = 2\left(1 - \phi\left(\frac{\varepsilon \cdot \sqrt{n}}{\sigma}\right)\right).$$

Luego

$$\mathbb{P}\left(|\overline{X}_n - \mu| > \varepsilon\right) \xrightarrow[n\to\infty]{} 0$$

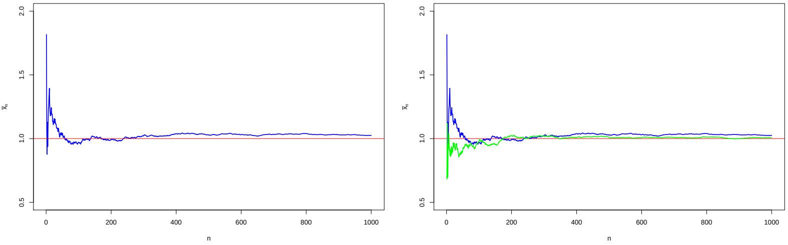
$$\mathbb{P}\left(|\overline{X}_n - \mu| \leq \varepsilon\right) \xrightarrow[n\to\infty]{} 1$$

¿Qué pasará con otras distribuciones?

Simulemos

Generaremos muestras aleatorias de una distribución $\mathcal{U}(0,1)$, de tamaño cada vez mayor, y calcularemos las medias muestrales. Comprobaremos si estos promedios muestrales se acercan al valor verdadero de la media cuando el tamaño muestral crece.

1. Probamos primero con una distribución $\mathcal{U}(0,2)$. Indique cuál es el valor verdadero de la media μ .
2. Generamos una muestra de tamaño $Nrep = 1000$ y para cada n entre 1 y 1000 calculemos el promedio de las primeras n observaciones. Realizamos un scatterplot de n vs. los promedios obtenidos. Incluimos una linea horizontal en el valor de y correspondiente a la verdadera media μ.
3. Repetimos comenzando con otra semilla y superpongamos el nuevo gráfico utilizando un color diferente. Comparamos los gráficos obtenidos. ¿ Qué podemos concluir?



2.0
1.5
1.0
0.5
$\bar{x}_n$
0
200
400
600
800
1000
n
2.0
1.5
1.0
0.5
$\bar{x}_n$
0
200
400
600
800
1000
n

¿Qué pasa con otras distribuciones?

Desigualdad de Tchebycheff

Sea W una v.a. con media $\mathbb{E}(W) = \mu_w$ y $V(W) = \sigma_w^2$. Luego, $\varepsilon > 0$

$$\mathbb{P}\left(|W - \mu_w| \geq \varepsilon\right) \leq \frac{\sigma_w^2}{\varepsilon^2}$$

¿Qué pasa con otras distribuciones?

Desigualdad de Tchebycheff

Sea W una v.a. con media $\mathbb{E}(W) = \mu_w$ y $V(W) = \sigma_w^2$. Luego, $\varepsilon > 0$

$$\mathbb{P}\left(|W - \mu_w| \geq \varepsilon\right) \leq \frac{\sigma_w^2}{\varepsilon^2}$$

Tomando $\varepsilon = k\sigma_w$,

$$\mathbb{P}\left(|W - \mu_w| \geq k\ \sigma_w\right) \leq \frac{\sigma_w^2}{k^2\sigma_w^2} = \frac{1}{k^2}$$

¿Qué pasa con otras distribuciones?

Desigualdad de Tchebycheff

Sea W una v.a. con media $\mathbb{E}(W) = \mu_w$ y $V(W) = \sigma_w^2$. Luego, $\varepsilon > 0$

$$\mathbb{P}\left(|W - \mu_w| \geq \varepsilon\right) \leq \frac{\sigma_w^2}{\varepsilon^2}$$

Tomando $\varepsilon = k\sigma_w$,

$$\mathbb{P}\left(|W - \mu_w| \geq k\ \sigma_w\right) \leq \frac{\sigma_w^2}{k^2\sigma_w^2} = \frac{1}{k^2}$$

En particular, para $k = 1, 2, 3$, obtenemos las las siguientes cotas:

$$\begin{aligned}
\mathbb{P}\left(|W - \mu_w| \geq \ \sigma_w\right) &\leq 1\ , \\
\mathbb{P}\left(|W - \mu_w| \geq \ 2\sigma_w\right) &\leq \frac{1}{4}\ , \\
\mathbb{P}\left(|W - \mu_w| \geq \ 3\sigma_w\right) &\leq \frac{1}{9}
\end{aligned}$$

Desigualdad de Tchebycheff aplicada a promedios

- $(X_i)_{i\geq 1}$ iid, $\mathbb{E}(X_i) = \mu$, $\mathbb{V}(X_i) = \sigma^2$.
- Promedio :

$$\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i \sim ?$$

- Esperanza y Varianza del promedio

$$\mu_{\overline{X}_n} = \mathbb{E}(\overline{X}_n) = \mu \ , \quad \sigma^2_{\overline{X}_n} = V(\overline{X}_n) = \frac{\sigma^2}{n} \ .$$

- Tchebycheff dice:

$$\mathbb{P}\left(|\overline{X}_n - \mu_{\overline{X}_n}| \geq \varepsilon\right) \ \leq \ \frac{\sigma^2_{\overline{X}_n}}{\varepsilon^2}$$

$$\mathbb{P}\left(|\overline{X}_n - \mu| \geq \varepsilon\right) \ \leq \ \frac{\sigma^2/n}{\varepsilon^2} = \frac{\sigma^2}{n\varepsilon^2}$$

Teorema: Ley de los Grandes Números (LGN)

Sean $(X_i)_{i\geq 1}$ i.i.d., con $\mathbb{E}[X_i] = \mu$ y $\mathbb{V}[X_i] = \sigma^2$, para todo i. entonces para todo $\varepsilon > 0$ vale que

$$\lim_{n\to\infty} \mathbb{P}(\Big(|\overline{X}_n - \mu| > \varepsilon\Big) = 0\,.$$

Teorema: Ley de los Grandes Números (LGN)

Sean $(X_i)_{i\geq 1}$ i.i.d., con $\mathbb{E}[X_i] = \mu$ y $\mathbb{V}[X_i] = \sigma^2$, para todo i.
entonces para todo $\varepsilon > 0$ vale que

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

Es decir, el promedio converge a μ en probabilidad:

$$\overline{X}_n \to \mu \,\text{en probabilidad}$$

Ley de los Grandes Números

Demostración: Tchebycheff prueba que

$$\mathbb{P}\left(|W-\mathbb{E}[W]|\geq\varepsilon\right)\leq\frac{\mathbb{V}[W]}{\varepsilon^2}$$

Vamos a invocar la Desigualdad de Tchebicheff, pero con $W=\overline{X}_n$. En tal caso,

$$\mu_{\overline{X}_n}=\mathbb{E}(\overline{X}_n)=\mu$$

$$\sigma^2_{\overline{X}_n}=\mathbb{V}(\overline{X}_n)=\frac{\sigma^2}{n}\,.$$

Tenemos entonces que

$$0\leq\mathbb{P}\left(|\overline{X}_n-\mu|>\varepsilon\right)=\mathbb{P}\left(|\overline{X}_n-\mu_{\overline{X}_n}|>\varepsilon\right)\leq\frac{V(\overline{X}_n)}{\varepsilon^2}=\frac{1}{\varepsilon^2}\frac{\sigma^2}{n}\to 0$$

L.G.N. aplicada a Variables Bernoulli

$\mathbb{E}$: experimento, A evento, $p = P(A)$

Repetimos en forma independiente $\mathbb{E}$ $n >>$ (muchas veces!!)

X_i : 1 si ocurre A 0 si no ocurre A

- $X_i \sim \mathcal{B}(1, p)$
- $\mathbb{E}(X_i) = p$, $\mathbb{V}(X_i) = p(1-p)$.
- $\overline{X}_n$ es la frecuencia relativa de éxitos en las n repeticiones.

$$\lim_{n\to\infty} \mathbb{P}(\left(|\overline{X}_n - p| > \varepsilon\right) = 0\,.$$

L.G.N. aplicada a Variables Bernoulli

$\mathbb{E}$: experimento, A evento, $p = P(A)$

Repetimos en forma independiente $\mathbb{E}$ $n >>$ (muchas veces!!)

X_i : 1 si ocurre A 0 si no ocurre A

- $X_i \sim \mathcal{B}(1, p)$
- $\mathbb{E}(X_i) = p$, $\mathbb{V}(X_i) = p(1-p)$.
- $\overline{X}_n$ es la frecuencia relativa de éxitos en las n repeticiones.

$$\lim_{n\to\infty} \mathbb{P}(\left(|\overline{X}_n - p| > \varepsilon\right) = 0\,.$$

$$\mathbb{P}\left(|\overline{X}_n - p| \geq \varepsilon\right) \;\leq\; \frac{p(1-p)}{n\varepsilon^2} \leq \frac{1/4}{n\varepsilon^2}$$

L.G.N. aplicada a Variables Bernoulli

$\mathbb{E}$: experimento, A evento, $p = P(A)$

Repetimos en forma independiente $\mathbb{E}$ $n >>$ (muchas veces!!)

X_i : 1 si ocurre A 0 si no ocurre A

- $X_i \sim \mathcal{B}(1, p)$
- $\mathbb{E}(X_i) = p$, $\mathbb{V}(X_i) = p(1-p)$.
- $\overline{X}_n$ es la frecuencia relativa de éxitos en las n repeticiones.

$$\lim_{n\to\infty} \mathbb{P}(\left(|\overline{X}_n - p| > \varepsilon\right) = 0\,.$$

$$\mathbb{P}\left(|\overline{X}_n - p| \geq \varepsilon\right) \;\leq\; \frac{p(1-p)}{n\varepsilon^2} \leq \frac{1/4}{n\varepsilon^2}$$

$$\overline{X}_n \to p \quad \text{en probabilidad}$$

Con una mirada estadística.... utilizamos la Ley de los Grandes Números para Estimación

Variables Bernoulli

- $X_i \sim \mathcal{B}(1, p)$, i.i.d.
- $\mathbb{E}(X_i) = p$, $\mathbb{V}(X_i) = p(1-p)$
- Por la Ley de los Grandes Números

$$\overline{X}_n \to \mathbb{E}(X_1) = p \quad \text{en probabilidad}$$

Variables Bernoulli

- $X_i \sim \mathcal{B}(1,p)$, i.i.d.
- $\mathbb{E}(X_i) = p$, $\mathbb{V}(X_i) = p(1-p)$
- Por la Ley de los Grandes Números

$$\overline{X}_n \to \mathbb{E}(X_1) = p \quad \text{en probabilidad}$$

- En tal caso, definimos

$$\widehat{p} = \widehat{p}_n = \overline{X}_n$$

.

- Propiedades de $\widehat{p}_n$:
  - $\mathbb{E}(\widehat{p}_n) = p$
  - $\widehat{p}_n \to p$ en probabilidad.

Estimación de Probabilidades.

- $(X_i)_{i\geq 1}$, i.i.d. $X_i \sim F$.
- Nos interesa estimar $\mathbb{P}(X \in A)$.
- Definimos

$$Y_i = I_{X_i \in A} = I_A(X_i)$$

- $Y_i \sim \mathcal{B}(1, p)$, con

$$p = \mathbb{E}(Y_1) = \mathbb{P}(Y_1 = 1) = \mathbb{P}(X_1 \in A)$$

- Por la Ley de los Grandes Números (caso Bernoulli)

$$\overline{Y}_n \to \mathbb{E}(Y_1) = \mathbb{P}(X_1 \in A) \quad \text{en probabilidad}$$

- Es decir,

$$\frac{1}{n}\sum_{i=1}^{n} I_{\{X_i \in A\}} \longrightarrow \mathbb{P}(X_1 \in A) \text{ en probabilidad}$$

Demostramos que la frecuencia relativa converge a la probabilidad.

Estimación de la Esperanza

- $(X_i)_{i\geq 1}$, i.i.d. $X_i \sim F$.
- $\mathbb{E}(X_i) = \mu$, $\mathbb{V}(X_i) = \sigma^2$
- Por la Ley de los Grandes Números

$$\overline{X}_n \to \mathbb{E}(X_1) = \mu \quad \text{en probabilidad}$$

- En tal caso, definimos

$$\widehat{\mu} = \widehat{\mu}_n = \overline{X}_n$$

.

- Propiedades de $\widehat{\mu}_n$:
  - $\mathbb{E}(\widehat{\mu}_n) = \mu$
  - $\widehat{\mu}_n \to \mu$ en probabilidad.

Estimación de Probabilidades - Ejemplo

- $(X_i)_{i\geq 1}$, i.i.d. $X_i \sim F$.
- Nos interesa estimar $F(3) = \mathbb{P}(X \leq 3)$.
- Definimos
$$Y_i = I_{\{X_i \leq 3\}} = I_{(-\infty,3]}(X_i)$$
- $Y_i \sim \mathcal{B}(1, p)$, con
$$p = \mathbb{E}(Y_1) = \mathbb{P}(Y_1 = 1) = \mathbb{P}(X_1 \leq 3) = F(3)$$
- Por la Ley de los Grandes Números (caso Bernoulli)
$$\overline{Y}_n \to \mathbb{E}(Y_1) = F(3) \quad \text{en probabilidad}$$
- En tal caso, definimos
$$\widehat{F}(3) = \widehat{F}_n(3) = \overline{Y}_n = \frac{1}{n}\sum_{i=1}^{n} I_{\{X_i \leq 3\}}$$

.

Estimación de $F(t)$

- $(X_i)_{i\geq 1}$, i.i.d. $X_i \sim F$.
- Nos interesa estimar $F(t) = \mathbb{P}(X \leq t)$.
- Definimos
$$Y_i = I_{X_i \leq t} = I_{(-\infty,t]}(X_i)$$
- $Y_i \sim \mathcal{B}(1, p)$, con
$$p = \mathbb{E}(Y_1) = \mathbb{P}(Y_1 = 1) = \mathbb{P}(X_1 \leq t) = F(t)$$
- Por la Ley de los Grandes Números (caso Bernoulli)
$$\overline{Y}_n \to \mathbb{E}(Y_1) = F(t) \quad \text{en probabilidad}$$
- En tal caso, definimos
$$\widehat{F}(t) = \widehat{F}_n(t) = \overline{Y}_n = \frac{1}{n}\sum_{i=1}^{n} I_{\{X_i \leq t\}}$$

.

LA empírica

La LGN demuestra que la frecuencia relativa converge a la probabilidad.

$(X_i)_{i\geq 1}$ i.i.d., $X_i \sim F$

$$\text{LGN:} \quad \frac{1}{n}\sum_{i=1}^{n} I_{\{X_i\leq 3\}} \xrightarrow{p} \mathbb{E}(I_{\{X_1\leq 3\}}) = \mathbb{P}(X_1 \leq 3) = F(3).$$

LA empírica

La LGN demuestra que la frecuencia relativa converge a la probabilidad.

$(X_i)_{i\geq 1}$ i.i.d., $X_i \sim F$

$$\text{LGN:} \quad \frac{1}{n}\sum_{i=1}^{n} I_{\{X_i \leq 3\}} \xrightarrow{p} \mathbb{E}(I_{\{X_1 \leq 3\}}) = \mathbb{P}(X_1 \leq 3) = F(3).$$

$$\widehat{F}_n(3) := \frac{1}{n}\sum_{i=1}^{n} I_{\{X_i \leq 3\}}$$

LA empírica

La LGN demuestra que la frecuencia relativa converge a la probabilidad.

$(X_i)_{i\geq 1}$ i.i.d., $X_i \sim F$

$$\text{LGN:}\quad \frac{1}{n}\sum_{i=1}^{n} I_{\{X_i\leq 3\}} \xrightarrow{p} \mathbb{E}(I_{\{X_1\leq 3\}}) = \mathbb{P}(X_1 \leq 3) = F(3).$$

$$\widehat{F}_n(3) := \frac{1}{n}\sum_{i=1}^{n} I_{\{X_i\leq 3\}}$$

$$\widehat{F}_n(t) := \frac{1}{n}\sum_{i=1}^{n} I_{\{X_i\leq t\}}$$

LA empírica

La LGN demuestra que la frecuencia relativa converge a la probabilidad.

$(X_i)_{i\geq 1}$ i.i.d., $X_i \sim F$

$$\text{LGN:} \quad \frac{1}{n}\sum_{i=1}^{n} I_{\{X_i\leq 3\}} \xrightarrow{p} \mathbb{E}(I_{\{X_1\leq 3\}}) = \mathbb{P}(X_1 \leq 3) = F(3).$$

$$\widehat{F}_n(3) := \frac{1}{n}\sum_{i=1}^{n} I_{\{X_i\leq 3\}}$$

$$\widehat{F}_n(t) := \frac{1}{n}\sum_{i=1}^{n} I_{\{X_i\leq t\}}$$

$$\text{LGN:} \quad \widehat{F}_n(t) = \frac{1}{n}\sum_{i=1}^{n} I_{\{X_i\leq t\}} \xrightarrow{p} \mathbb{E}(I_{\{X_1\leq t\}}) = \mathbb{P}(X_1 \leq t) = F(t).$$

LGN y una vuelta más de rosca....

- Promedio de mongo converge a esperanza de mongo:

$$\frac{1}{n}\sum_{i=1}^{n} \bigstar_i \xrightarrow{p} \mathbb{E}(\bigstar)$$

- Ejemplo:

$$\frac{1}{n}\sum_{i=1}^{n} X_i^2 \xrightarrow{p} \mathbb{E}(X_1^2)$$

LGN y una vuelta más de rosca....

- Promedio de mongo converge a esperanza de mongo:

$$\frac{1}{n}\sum_{i=1}^{n}\bigstar_i \xrightarrow{p} \mathbb{E}(\bigstar)$$

- Ejemplo:

$$\frac{1}{n}\sum_{i=1}^{n}X_i^2 \xrightarrow{p} \mathbb{E}(X_1^2)$$

- Funciones g continuas preservan convergencia en probabilidad.

$$g\left(\frac{1}{n}\sum_{i=1}^{n}\bigstar_i\right) \xrightarrow{p} g\left(\mathbb{E}(\bigstar)\right)$$

- Ejemplo:

$$(\overline{X}_n)^2 \xrightarrow{p} \mu^2 \ , \quad \mu = \mathbb{E}(X_1)$$

Estimando $\sigma^2 = \mathbb{V}(X) = \mathbb{E}(X^2) - \mu^2 = \mathbb{E}\{(X - \mu)^2\}$

- Muestra: $X_1, \ldots, X_n$ i.i.d. donde $X_i \sim X$

Estimando $\sigma^2 = \mathbb{V}(X) = \mathbb{E}(X^2) - \mu^2 = \mathbb{E}\{(X-\mu)^2\}$

- Muestra: $X_1, \ldots, X_n$ i.i.d. donde $X_i \sim X$

- ESTIMADOR : cuenta hecha con la muestra

$$\widehat{\sigma^2}_n = \frac{1}{n}\sum_{i=1}^{n} X_i^2 - (\overline{X}_n)^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X}_n)^2$$

- ESTIMACION: Valor del estimador al utilizar mis datos.

- Consistencia: el estimador converge a lo que queremos estimar:

$$\widehat{\sigma^2}_n = \frac{1}{n}\sum_{i=1}^{n} X_i^2 - (\overline{X}_n)^2 \xrightarrow{p} E(X^2) - \mu^2 = \mathbb{V}(X)$$

Estimando $\sigma^2 = \mathbb{V}(X)$ y σ - OTRA VERSION

- Muestra: $X_1, \ldots, X_n$ i.i.d., $X_i \sim F$.

- ESTIMADOR : cuenta hecha con la muestra

$$S^2 = S_n^2 = \frac{n}{n-1}\widehat{\sigma^2}_n = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X}_n)^2$$

Estimando $\sigma^2 = \mathbb{V}(X)$ y σ - OTRA VERSION

- Muestra: $X_1, \ldots, X_n$ i.i.d., $X_i \sim F$.

- ESTIMADOR : cuenta hecha con la muestra

$$S^2 = S_n^2 = \frac{n}{n-1}\widehat{\sigma^2}_n = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X}_n)^2$$

- ESTIMACION: Valor del estimador al utilizar mis datos.

- En R `var(datos)`, `sd(datos)`

- Consistencia: el estimador converge a lo que queremos estimar:

$$S_n^2 = \frac{n}{n-1}\widehat{\sigma^2}_n \xrightarrow{p} \sigma^2 \ , \quad S_n = \sqrt{S_n^2} \xrightarrow{p} \sigma$$