



UNIVERSITAT  
POLITÈCNICA  
DE VALÈNCIA

# El teorema central del límite

<b>Apellidos, nombre</b>	<b>Julio Benítez López<sup>1</sup></b> (jbenitez@mat.upv.es) <b>Alicia Martínez Roca<sup>1</sup></b> (aroca@mat.upv.es)
<b>Departamento</b>	<b><sup>1</sup>Departamento de Matemática Aplicada</b>
<b>Centro</b>	<b>Universitat Politècnica de València</b>

# 1 Resumen de las ideas clave

El teorema central del límite es el teorema más importante de la probabilidad: de forma intuitiva dice que la suma (o el promedio) de muchas variables aleatorias independientes y con la misma distribución es aproximadamente una normal, aunque no sepamos la distribución de los sumandos. Este teorema tiene muchas e importantes aplicaciones en estadística. Por ejemplo, cuando se mide la altura de un número grande de personas, la media de las mediciones sigue aproximadamente una normal. Lo mismo se puede decir si se mide la duración de muchos televisores o cualquier otra variable aleatoria ¡independientemente de cómo se comporte por separado!

En este artículo podrás ver la evolución histórica del teorema central del límite usando los programas de cálculo numérico Octave o Matlab para comprender mejor el significado de este teorema.

A lo largo de este artículo usaremos la siguiente notación: Si  $X$  es una variable aleatoria binomial de parámetros  $n$  y  $p$ , denotaremos esta situación como  $X \sim B(n, p)$ , y si es una variable normal de media  $\mu$  y varianza  $\sigma^2$ , lo denotaremos como  $X \sim N(\mu, \sigma^2)$ .

**¡Cuidado!** Para seguir este documento es conveniente que uses simultáneamente Matlab u Octave. Si optas por Octave, tienes que cargar el paquete `statistics` mediante `pkg load statistics`.

## 2 Objetivos

- Comprender el teorema central del límite y saber cuándo se puede usar.
- Aplicar Octave o Matlab para poder realizar cálculos y dibujar gráficas de funciones relativas a varias distribuciones de probabilidad, y en particular de la distribución normal.

## 3 Abraham de Moivre

El teorema central del límite es el teorema más importante de la teoría de la probabilidad, pero su historia comenzó de forma un poco humilde. Abraham de Moivre (1667-1754, [https://es.wikipedia.org/wiki/Abraham\\_de\\_Moivre](https://es.wikipedia.org/wiki/Abraham_de_Moivre)) publicó una demostración de un caso muy concreto del teorema central del límite, introduciendo por primera vez la distribución normal.

### Actividad 1

Con la sintaxis `binopdf(k,n,p)` calculas la función de cuantía evaluada en  $k$  (este  $k$  puede ser un vector) de una variable binomial de parámetros  $n$  y  $p$ . Así, para calcular  $\text{pr}(X = k)$  para  $k = 0, 1, 2$ , siendo  $X \sim B(2, 0.5)$ , se teclea `binopdf(0:2,2,0.5)`. Crea una función de Octave cuyo argumento de entrada sea  $n$  y produzca una gráfica que muestre  $\text{pr}(X = k)$  para  $k = 0, \dots, n$ , siendo  $X \sim B(n, 0.5)$ . Puedes ver el resultado en la gráfica de la izquierda de la figura 1.

No podemos adivinar lo que de Moivre tenía en su cabeza en sus investigaciones, pero vamos a intentar descubrirlo por nosotros mismos. La gráfica de la función de cuantía de  $X \sim B(n, 0.5)$  para valores altos de  $n$  se le parece mucho a la gráfica de una función de la forma  $f(x) = Ke^{-a(x-b)^2}$  (supondremos  $a > 0$  para que  $f(x)$  tienda a 0 cuando  $x$  tiende a  $\pm\infty$ ). Ahora tenemos que averiguar  $a, b, K$ .

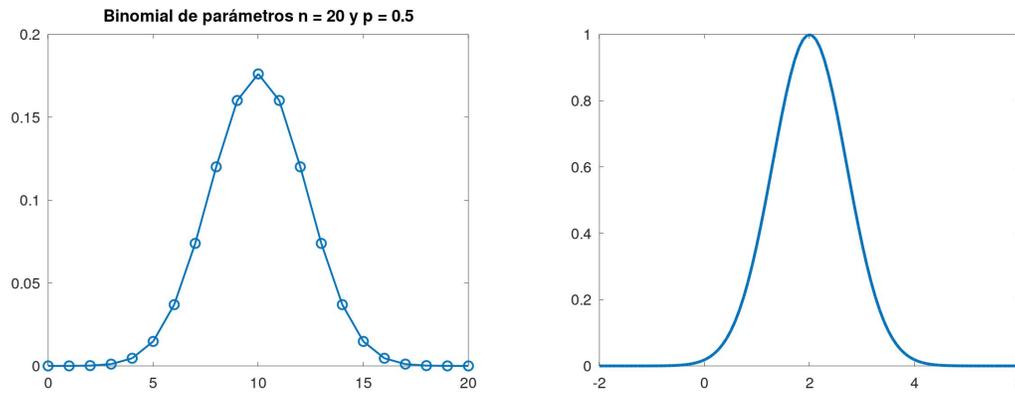


Figura 1: Izquierda: La gráfica de la función de cuantía de una binomial. Derecha: La gráfica de la función  $f(x) = e^{-(x-2)^2}$ . ¡Las formas de las dos gráficas se parecen!

### Actividad 2

En esta actividad, vamos a hallar la constante  $K$  usando que  $1 = \int_{-\infty}^{\infty} f(x) dx$ , ya que  $f$  es una función de densidad. Usa el cálculo simbólico de Matlab o bien una página de internet que calcule integrales online (por ejemplo, <https://www.integral-calculator.com>) para probar  $K = \sqrt{a/\pi}$ .

Ahora vamos a estimar el parámetro  $b$ .

### Actividad 3

Usa la función de la actividad 1 y conjetura el valor de  $k$  (o los valores) que maximiza  $\text{pr}(X = k)$  para algunos valores de  $n$ , siendo  $X \sim B(n, 0.5)$ . ¿Relacionas este valor con la esperanza de  $X$ ?  
Sea  $f(x) = K e^{-a(x-b)^2}$ , con  $a > 0$ . Prueba que  $f$  solo tiene un máximo en  $x_0 = b$ .

Esperamos que hayas conjeturado que si  $n$  es par, entonces el valor de  $k$  que maximiza  $\text{pr}(X = k)$  es  $k = n/2$ , y si  $n$  es impar, los valores de  $k$  que maximizan  $\text{pr}(X = k)$  son  $k = (n-1)/2$  y  $k = (n+1)/2$ . Por tanto, si queremos que  $f(x) = K e^{-a(x-b)^2}$  se ajuste a la función de cuantía de una binomial  $X \sim B(n, 0.5)$ , entonces un valor razonable de  $b$  es  $n/2$ , que es la esperanza de esta binomial.

Para hallar un valor razonable de  $a$ , vamos a igualar el máximo de la función de cuantía de una binomial  $X \sim B(n, 0.5)$  con el máximo de la función  $f(x) = \sqrt{a/\pi} e^{-a(x-b)^2}$ , que es  $f(b) = \sqrt{a/\pi}$ .

### Actividad 4

Sea  $X \sim B(n, 0.5)$  y  $n$  par. Con Octave, calcula  $a$  a partir de la igualdad  $\sqrt{a/\pi} = \text{pr}(X = n/2)$  para varios valores de  $n$  (prueba para  $n \geq 50$ ).

Hemos obtenido la siguiente tabla:

$n$	50	100	150	200
$a$	0.0396020	0.0199003	0.0132890	0.0099750

¿Qué observas? Cuando preparamos estas notas, no observamos ninguna relación. Solo nos dimos cuenta de que a medida que  $n$  crece,  $a$  disminuye. Por esto preparamos otra tabla relacionando  $n$  con  $1/a$  con la idea de descubrir una relación.

### Actividad 5

Repite la actividad anterior, pero calculando  $1/a$  para algunos valores grandes y pares de  $n$ .

Se obtiene esta tabla:

$n$	50	100	150	200
$1/a$	25.251	50.251	75.250	100.250

¿Ves ahora una relación entre  $n$  y  $1/a$ ? ¿Cuál es la relación entre  $a$  y  $n$ ?

### Actividad 6

Escribe una función de Octave cuyo argumento de entrada sea  $n$  y que dibuje en la misma gráfica la función de cuantía de  $X \sim B(n, 0.5)$  y la función  $f(x) = K e^{-a(x-b)^2}$  para los valores de  $K, a, b$  conjeturados previamente. El resultado se debe parecer a las gráficas de la figura 2.

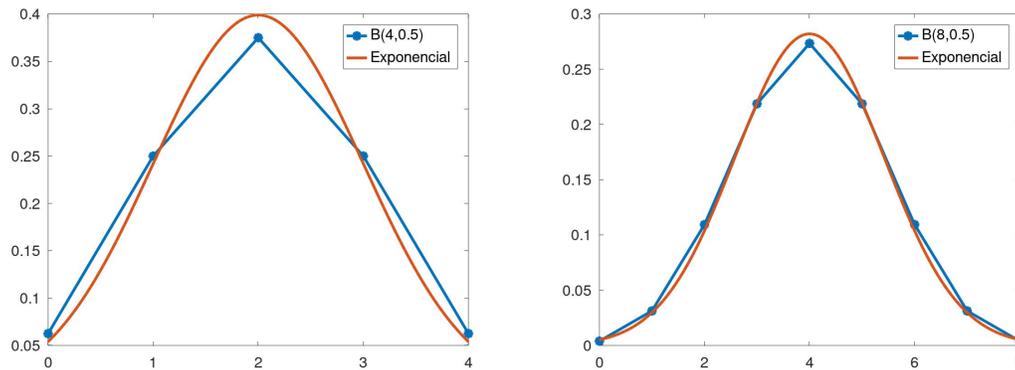


Figura 2: Como puedes observar, la gráfica de la función de cuantía de una binomial se aproxima muy bien a la gráfica de  $f(x) = K e^{-a(x-b)^2}$  para valores adecuados de  $K, a, b$ .

Esto fue lo que de Moivre enunció y demostró:

**Teorema.** Si  $X \sim B(n, 0.5)$ , entonces para valores grandes de  $n$ ,

$$\text{pr}(X = k) \simeq \sqrt{\frac{2}{\pi n}} \exp\left(-\frac{2}{n} \left(k - \frac{n}{2}\right)^2\right),$$

y

$$\text{pr}(k_1 \leq X \leq k_2) \simeq \int_{k_1}^{k_2} \sqrt{\frac{2}{\pi n}} \exp\left(-\frac{2}{n} \left(x - \frac{n}{2}\right)^2\right) dx.$$

## 4 Pierre-Simon Laplace (extensión del resultado de de Moivre)

El siguiente paso fue dado por Pierre-Simon Laplace (1749-1827, [https://es.wikipedia.org/wiki/Pierre-Simon\\_Laplace](https://es.wikipedia.org/wiki/Pierre-Simon_Laplace)). En su libro *Théorie analytique des probabilités* publicado en 1812 generalizó los argumentos dados por de Moivre a una binomial con parámetro  $p$  arbitrario. La generalización de Laplace siguió básicamente los mismos pasos de de Moivre<sup>1</sup>.

### Actividad 7

Generaliza la actividad 1 cuando  $p$  es arbitrario. Puedes ver dos gráficas en la figura 3.

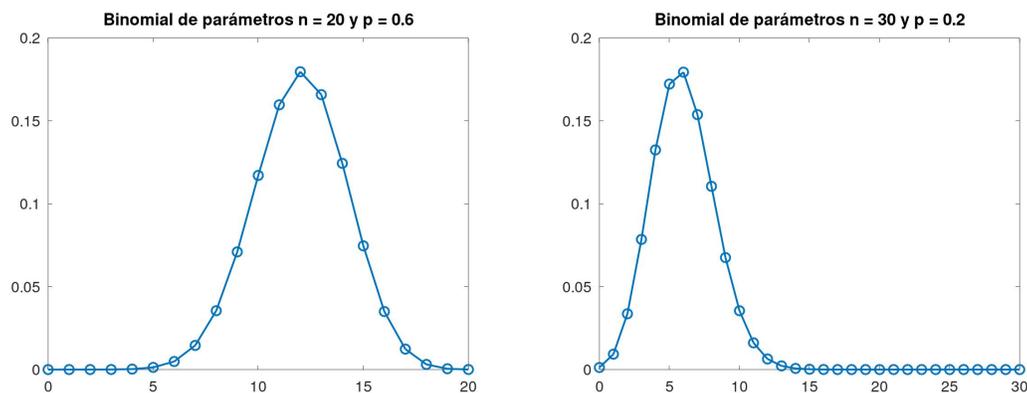


Figura 3: Dos gráficas de la función de cuantía de dos binomiales con  $p \neq 1/2$ .

Observa que de nuevo, las gráficas de las binomiales se parecen a las gráficas de  $f(x) = K e^{-a(x-b)^2}$ . El valor de  $K$  se obtiene de la misma manera que antes y es  $K = \sqrt{a/\pi}$ .

Para averiguar el valor más apropiado de  $b$ , observa de nuevo la actividad 3. Parece que lo más adecuado es igualar el punto donde se alcanza el máximo de  $f(x) = K \exp(-a(x-b)^2)$  al valor de  $k$  que maximiza  $\text{pr}(X = k)$  cuando  $X \sim B(n, p)$ . Mira que  $f$  alcanza el máximo en  $b$  (mira la actividad 3).

Ahora vamos a estimar el valor de  $k$  que maximiza  $\text{pr}(X = k)$ . Observando varias gráficas obtenidas en la actividad 7 podemos intuir este valor de  $k$ ; pero vamos a intuirlo con dos ejemplos. Si lanzas una moneda equilibrada 1000 veces, ¿cuántas caras esperas obtener? Pues más o menos unas 500 caras. Si lanzas un dado equilibrado 600 veces, más o menos esperas obtener unos 100 “seises”. En general, si  $X \sim B(n, p)$ , es razonable obtener  $pn$  éxitos. De hecho, la esperanza de  $X$  es  $pn$ .

Por los párrafos anteriores, parece que el valor más razonable de  $b$  es  $pn$ .

Ahora, para obtener el valor más razonable para  $a$  procederemos de manera parecida a la actividad 4.

### Actividad 8

Toma un valor de  $n$  suficientemente grande (por ejemplo,  $n = 100$ ). Sea  $X \sim B(n, p)$ . Si procedemos como en la actividad 4, nuestro primer intento quizá sea hallar  $a$  tras igualar  $\text{pr}(X = pn)$  y  $f(pn)$ . Trivialmente,  $f(pn) = K = \sqrt{a/\pi}$ . Pero tenemos un problema:  $pn$  no necesariamente es un entero, y si  $pn$  no es entero, entonces  $\text{pr}(X = pn) = 0$  puesto que  $X$  solo toma valores enteros.

<sup>1</sup>Puedes ver el libro “A History of the Central Limit Theorem” escrito por H. Fisher, de nivel algo avanzado.

Por ello, despejaremos  $a$  en la igualdad  $\text{pr}(X = [pn]) = \sqrt{a/\pi}$ .

a) Para  $p = 0.1, 0.2, \dots, 0.9$  y  $n = 100$  halla (usando Octave o Matlab)  $a$  y  $1/a$ . Deberás obtener unos valores parecidos a los de la siguiente tabla

$p$	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
$a$	0.055	0.031	0.024	0.021	0.02	0.021	0.024	0.031	0.055
$1/a$	18.306	32.281	42.264	48.254	50.251	48.254	42.264	32.281	18.306

b) Dibuja los puntos  $(p, 1/a)$ . ¿Qué curva crees que has obtenido? Si no lo sabes, pasa al siguiente apartado.

c) Prueba a dibujar la misma curva, pero ahora para  $p = 0.01, 0.02, \dots, 0.99$ . Puedes observar que es una parábola cuyas raíces son  $p = 0$  y  $p = 1$  y que “mira hacia abajo”. Por tanto, parece razonable que  $1/a = \alpha p(1 - p)$ . Comprueba que  $\alpha$  toma valores cercanos a 200.

d) Prueba para varios valores de  $n$  (por supuesto, grandes). Conjetura una relación entre  $n$  y  $\alpha$ .

¡Ya tenemos una conjetura! Como  $a = 1/(\alpha p(1 - p))$ , parece que la función que estamos buscando es

$$f(x) = K \exp(-a(x - b)^2) = \sqrt{\frac{a}{\pi}} \exp(-a(x - b)^2) = \sqrt{\frac{1}{2\pi np(1 - p)}} \exp\left(-\frac{(x - pn)^2}{2np(1 - p)}\right).$$

### Actividad 9

Generaliza la actividad 6 para  $n$  y  $p$  arbitrarios. Puedes ver dos gráficas en la figura 4.

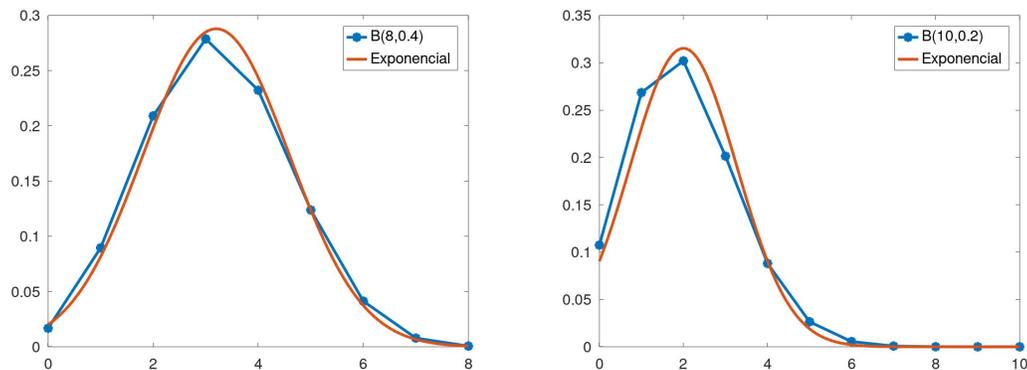


Figura 4: Dos gráficas de dos binomiales junto con sus aproximaciones por  $f(x)$ .

Si recordamos que si  $X \sim B(n, p)$ , entonces la esperanza de  $X$  es  $\mu = np$  y la varianza de  $X$  es  $\sigma^2 = np(1 - p)$ , entonces la función  $f$  anterior se puede escribir como

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x - \mu}{\sigma}\right)^2\right).$$

¡Justo como la función de densidad de una distribución normal! El propio Laplace ya estableció la siguiente aproximación:

**Teorema.** Si  $X \sim B(n, p)$ , entonces para valores grandes de  $n$  se cumple que  $X$  es aproximadamente una normal de media  $np$  y varianza  $np(1 - p)$ .

## 5 Pierre-Simon Laplace (suma de variables aleatorias independientes)

La versión estándar del teorema central del límite, que fue probada sin mucho rigor por Laplace, establece que la suma (o el promedio) de muchas variables aleatorias independientes e idénticamente distribuidas se aproxima a una normal. Laplace estuvo interesado en este teorema debido principalmente a la importancia de medir varias veces la misma cantidad: si se considera que las medidas son independientes e idénticamente distribuidas, entonces su promedio puede ser aproximado por una distribución normal. El enfoque de Laplace fue totalmente distinto al que vamos a ver a continuación.

Vamos a estudiar la suma de dos variables aleatorias independientes discretas  $X$  e  $Y$ . Puesto que

$$\text{pr}(X + Y = k) = \sum_n \text{pr}(X = k - n, Y = n) = \sum_n \text{pr}(X = k - n) \text{pr}(Y = n).$$

si definimos los vectores de cuantía de  $X$  e  $Y$ , respectivamente, como

$$p_x = (\dots, \text{pr}(X = -1), \text{pr}(X = 0), \text{pr}(X = 1), \dots), \quad p_y = (\dots, \text{pr}(Y = -1), \text{pr}(Y = 0), \text{pr}(Y = 1), \dots),$$

entonces

$$\text{pr}(X + Y = k) = \sum_n p_x(k - n) p_y(n).$$

La suma anterior es la convolución de las sucesiones  $p_x$  y  $p_y$  y se denota  $p_x * p_y$ .

**Teorema.** Si  $X$  e  $Y$  son distribuciones discretas independientes cuyas funciones de cuantía son  $p_x$  y  $p_y$ , entonces la función de cuantía de  $X + Y$  es  $p_x * p_y$ .

Veamos un ejemplo usando Octave. Sea  $X$  el número de caras obtenidas cuando lanzamos una moneda equilibrada 10 veces y sea  $Y$  el número de caras cuando se lanza la misma moneda 5 veces. Vamos a calcular  $\text{pr}(X + Y = 7)$ . Primero declaramos el rango de variación de  $X$  e  $Y$ :

```
x = 0:10; y = 0:5;
```

Ahora calculamos el rango de variación de  $X + Y$ :

```
sumaxy = min(x)+min(y):max(x)+max(y);
```

Calculamos las funciones de cuantía de  $X$ ,  $Y$  y  $X + Y$  (el comando que calcula la convolución de dos vectores es `conv`):

```
px = binopdf(x,10,0.5); py = binopdf(y,5,0.5);  
psuma = conv(px,py);
```

Dentro del rango de valores de  $X + Y$  vemos la posición del "7" pues queremos calcular  $\text{pr}(X + Y = 7)$ .

```
posicion = find(sumaxy==7);
```

Y por último, calculamos  $\text{pr}(X + Y = 7)$  con

```
psuma(posicion)
```

Como  $X \sim B(10, 0.5)$  e  $Y \sim B(5, 0.5)$ , entonces  $X + Y \sim B(10 + 5, 0.5) = B(15, 0.5)$  (lanzar una moneda 10 veces y luego la misma moneda 5 veces es lo mismo que lanzar esta moneda 15 veces). Por lo que  $\text{pr}(X + Y = 7)$  también se puede calcular por medio de

```
binopdf(7,15,0.5)
```

Puedes comprobar que ambas maneras conducen al mismo resultado.

## Actividad 10

Se lanza un dado equilibrado dos veces. Calcula la probabilidad de que la suma de las puntuaciones sea 8. La respuesta es 0.13889.

Ahora vamos a estudiar con Octave la suma de variables discretas independientes e idénticamente distribuidas. Si  $X_1, X_2, \dots, X_n$  son variables aleatorias independientes cuyo vector de cuantía común (todas las  $X_i$  tienen el mismo vector de cuantía pues estas variables tienen la misma distribución) es  $\mathbf{p}$ , entonces el vector de cuantía de  $X_1 + X_2 + \dots + X_n$  es  $n$  veces la convolución de  $\mathbf{p}$ , es decir,  $\underbrace{\mathbf{p} * \mathbf{p} * \dots * \mathbf{p}}_{n \text{ veces}}$ .

Aprovechando esto, vamos a ir dibujando las gráficas de los vectores de cuantía de  $X_1 + \dots + X_k$  de la forma siguiente:

$$\mathbf{p} \rightarrow \mathbf{p} * \mathbf{p} \rightarrow (\mathbf{p} * \mathbf{p}) * \mathbf{p} \rightarrow ((\mathbf{p} * \mathbf{p}) * \mathbf{p}) * \mathbf{p} \rightarrow \dots$$

La función de más abajo tiene los siguientes argumentos de entrada:

- $\mathbf{x}$  son los valores que toman las variables  $X_i$ .
- $\mathbf{p}$  son los valores  $\text{pr}(X = x_i)$ .
- $n$  es el número de sumandos en  $X_1 + \dots + X_n$ .

```
function tcl(x,p,n)
m = min(x); M = max(x);
dibujo = plot(x,p,'o-','linewidth',3);
axis([min(m*(1:n)) max(M*(1:n)) 0 max(p)])
title('n = 1. Haz click en la figura','fontsize',16)
w = waitforbuttonpress;
y = p;
hold on
for i = 2: n
    delete(dibujo)
    x = i*m:i*M; y = conv(y,p);
    dibujo=plot(x,y,'o-','linewidth',3);
    title(['n = ' num2str(i)],'fontsize',16)
    pause(0.1)
end
hold off
```

Esta función dibuja los vectores de cuantía de  $X_1, X_1 + X_2, X_1 + X_2 + X_3, \dots$  en la misma gráfica. Primero dibuja la de  $X_1$  y luego espera a un *click* en la figura para dibujar el resto de manera pausada.

Si cada  $X_i$  toma los valores  $x_1 < \dots < x_r$ , entonces el menor valor que puede tomar  $X_1 + X_2$  es  $2x_1$  y el mayor es  $2x_r$ . Por esta razón, el menor y el mayor valor que toma  $X_1 + \dots + X_k$  son  $kx_1$  y  $kx_r$ , respectivamente. Se toma como intervalo de variación de  $x$  el intervalo  $[a, b]$ , siendo  $a$  el mínimo de  $x_1, 2x_1, \dots, nx_1$  y  $b$  el máximo de  $x_r, 2x_r, \dots, nx_r$ , y como intervalo de variación de  $y$  el intervalo que va desde 0 hasta la mayor probabilidad de cada  $X_i$ .

Prueba esta función para varios *input*. Podrás ver que, incluso para valores no grandes de  $n$  (por ejemplo,  $n = 10$ ), la gráfica se le va pareciendo poco a poco a la de una normal. ¿Qué normal? Si  $\mu$  y  $\sigma^2$  es la media y varianza comunes a cada  $X_i$ , entonces la media y varianza de  $X_1 + \dots + X_n$  es  $n\mu$  y  $n\sigma^2$ , respectivamente.

## Actividad 11

Crea una función de Octave de forma que los argumentos de entrada sean los mismos que la función `tc1` y la salida sean las gráficas de la función de cuantía de  $X_1 + \dots + X_n$  y la de la función de densidad de la variable normal que aproxima a  $X_1 + \dots + X_n$ . Puedes ver dos ejemplos en la gráficas de la figura 5.

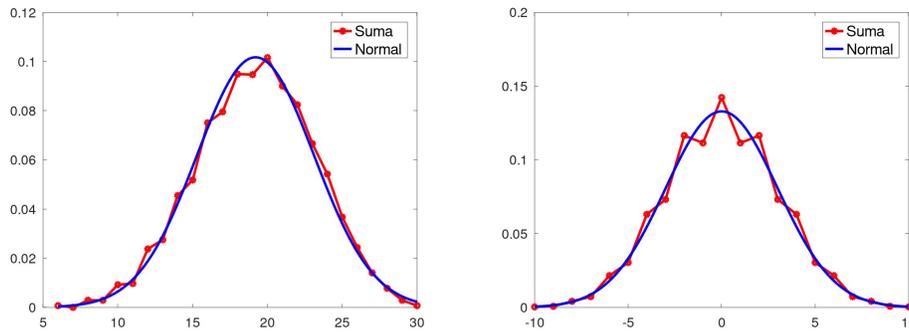


Figura 5: La figura de la izquierda se ha obtenido usando la variable aleatoria  $X$  tal que  $\text{pr}(X = 1) = 0.3$ ,  $\text{pr}(X = 2) = 0.1$ ,  $\text{pr}(X = 3) = \text{pr}(X = 4) = \text{pr}(X = 5) = 0.2$  y  $n = 6$ . La de la derecha cumple  $\text{pr}(X = -1) = \text{pr}(X = 1) = 0.45$ ,  $\text{pr}(X = 0) = 0.1$ .

Ya podemos enunciar (de forma poco rigurosa) el teorema central del límite en una versión muy parecida a lo que enunció Laplace:

**Teorema.** Si  $X_1, \dots, X_n$  son variables aleatorias independientes e idénticamente distribuidas, entonces  $X_1 + \dots + X_n$  es aproximadamente una variable normal si  $n$  es grande.

Para acabar esta sección, observemos que si  $X \sim B(n, p)$ , entonces  $X$  se puede considerar como  $X = X_1 + \dots + X_n$ , siendo cada  $X_i \sim B(1, p)$  (contar el número de caras si lanzo  $n$  monedas trucadas equivale a etiquetar cada moneda con un 1 si sale cara y con un 0 si sale cruz, y sumar todos los unos que salen). Por lo que la aproximación de una binomial para  $n$  grande es un caso particular de la versión del teorema central del límite que acabamos de ver.

## 6 Etapa final del teorema central del límite

Siméon-Denis Poisson (1781-1840) continuó los trabajos de Laplace sobre el teorema central del límite. Este teorema recibió poca importancia más allá de las matemáticas hasta que Adolphe Quetelet (1796-1874) reconoció la importancia de la distribución normal en la biología y en las ciencias sociales. Más adelante, Pafnuti Chebyshev (1821-1894) y Andrey Markov (1856-1922) perfeccionaron el enunciado de este teorema. Más tarde fue enunciado y probado con todo rigor por Aleksandr Liapunov (1857-1918). Posteriormente fueron descubiertas varias generalizaciones y versiones del teorema central del límite.

Veamos cómo se puede enunciar el teorema central del límite mediante actividades de Octave.

Si ejecutas la función `tc1` podrás observar que aunque la gráfica de la función de cuantía de  $X_1 + \dots + X_n$  se va pareciendo cada vez más a una normal, no se van pareciendo entre sí; es decir, no parece que las variables  $X_1, X_1 + X_2, X_1 + X_2 + X_3, \dots$  se vayan pareciendo entre sí ya que el eje horizontal cambia. Puedes ver un ejemplo sencillo con `x = 1:3; p=[0.5 0.1 0.4]; tc1(x,p,15)`.

¿Cómo se arregla esto? Comparando “de tú a tú” todas las variables  $X_1, X_1 + X_2, X_1 + X_2 + X_3, \dots$  y esto se logra con un patrón de medida común, que es la normal estándar. ¿Qué se hace para estandarizar una normal? Restarle la media y dividir por la desviación típica.

### Actividad 12

Sean  $X_1, \dots, X_k$  variables aleatorias independientes idénticamente distribuidas de media  $\mu$  y varianza  $\sigma^2$ . Prueba que la media y la varianza de  $X_1 + \dots + X_k$  son  $k\mu$  y  $k\sigma^2$ , respectivamente.

Por la actividad anterior, la esperanza y la varianza de las variables

$$Z_k = \frac{X_1 + \dots + X_k - k\mu}{\sqrt{k}\sigma},$$

son 0 y 1, respectivamente. Y si  $n$  es grande, entonces  $X_1 + \dots + X_n$  es aproximadamente una normal. Y es de esperar que  $Z_k$  tienda a una normal estándar cuando  $k \rightarrow \infty$ .

La siguiente función es bastante parecida a la función `tcl1`. Los argumentos de entrada son idénticos. La gráfica es algo distinta: En primer lugar, las  $x$  se trasladan (se resta  $\mu$ ) y se escalan (se divide por  $\sigma$ ) para estandarizar  $X_1 + \dots + X_n$ . En segundo lugar, en el eje vertical, se dibujan las funciones de distribución (ni la función de densidad ni la de cuantía); para ello se usan los comandos `cumsum` (que calcula sumas acumuladas) y `normcdf`.

```
function tclbis(x,p,n)
m = min(x); M = max(x);
mu = dot(x,p); a2 = dot(x.^2,p); sigma = sqrt(a2-mu^2);
dibujo = plot((x-mu)/sigma, cumsum(p), 'o-', 'linewidth', 3, 'color', 'b');
axis([sqrt(n)*(m-mu)/sigma sqrt(n)*(M-mu)/sigma 0 1])
title('n = 1. Haz click en la figura', 'fontsize', 16)
w = waitforbuttonpress;
y = p; q = p;
hold on
for i = 2: n
    delete(dibujo)
    y = i*m:i*M; q = conv(q,p);
    dibujo=plot((y-i*mu)/(sqrt(i)*sigma), cumsum(q), 'o-', 'linewidth', 3, 'color', 'b');
    title(['n = ' num2str(i)], 'fontsize', 16)
    pause(0.1)
end
ejex = linspace(sqrt(n)*(m-mu)/sigma, sqrt(n)*(M-mu)/sigma);
ejey = normcdf(ejex);
plot(ejex, ejey, 'linewidth', 3, 'color', 'r')
hold off
```

Puedes ejecutar esta función para los valores que desees de  $x$  y  $p$  y verás que se obtienen aproximaciones buenas incluso para valores no muy altos de  $n$ . Por ejemplo, yo he obtenido las siguientes figuras para  $x=1:5$ ;  $p=[0.2 \ 0.2 \ 0 \ 0.3 \ 0.3]$  y  $n = 3, n = 10$ .

Por lo que parece que se da el siguiente resultado:

**Teorema.** Sean  $X_1, X_2 \dots$  variables aleatorias independientes idénticamente distribuidas de media  $\mu$  y varianza  $\sigma^2$ . Se define

$$Z_k = \frac{X_1 + \dots + X_k - k\mu}{\sqrt{k}\sigma}.$$

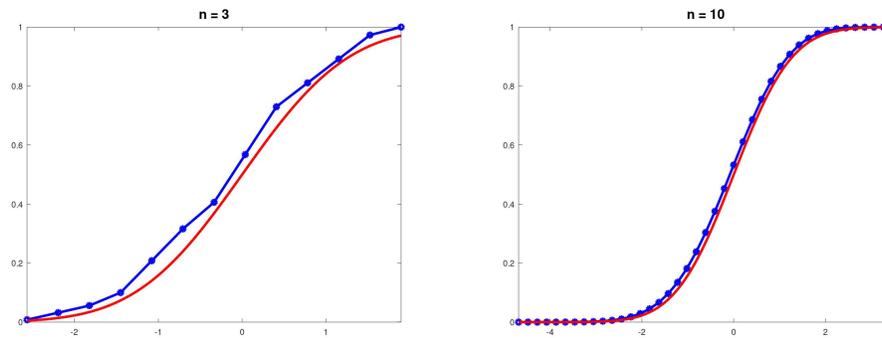


Figura 6: Las curvas azul y roja son las gráficas de las funciones de distribución de la variable  $Z_n$  y de la normal estándar, respectivamente.

Entonces para todo  $t$  se cumple

$$\lim_{n \rightarrow \infty} \text{pr}(Z_n \leq t) = \text{pr}(Z \leq t),$$

siendo  $Z$  una variable normal estándar.

## 7 Cierre

El teorema central del límite es fundamental porque permite estudiar la suma de variables aleatorias independientes e idénticamente distribuidas cuando hay muchos sumandos, por ejemplo cuando se analizan encuestas o datos experimentales donde es importante estudiar la media de los datos empíricos.

En este artículo docente has aprendido, mediante la evolución histórica del teorema central del límite, a comprender el teorema y a usarlo desde el punto de vista práctico. En particular has visto que:

1. Sirve para aproximar  $X_1 + \dots + X_n$ , siendo  $X_i$  variables aleatorias independientes e idénticamente distribuidas.
2. Un uso importante es estudiar la media de una muestra estadística:  $\bar{X} = (X_1 + \dots + X_n)/n$ .
3.  $X_1 + \dots + X_n$  se aproxima cuando  $n$  es grande a una normal.
4. Si la media y varianza de cada  $X_i$  son  $\mu$  y  $\sigma^2$ , respectivamente, entonces la media y varianza de  $X_1 + \dots + X_n$  son  $n\mu$  y  $n\sigma^2$ .
5. Si la media y varianza de cada  $X_i$  es  $\mu$  y  $\sigma^2$ , respectivamente, entonces la media y varianza de  $\bar{X}$  es  $\mu$  y  $\sigma^2/n$ .
6. Como caso particular, y primero que aparece en la historia, si  $X$  sigue una distribución binomial de parámetros  $n$  y  $p$ ,  $X = X_1 + \dots + X_n$  con  $X_i$  dicotómicas e independientes. Entonces cuando  $n$  es grande,  $X$  se aproxima a una normal de media  $pn$  y varianza  $p(1-p)n$ .

Has usado programas numéricos (Matlab u Octave) para hacer cálculos estadísticos y has visto la convolución discreta para estudiar la suma de variables aleatorias.

Además, has aprendido a hacer conjeturas y cómo usar estos programas numéricos para confirmar estas conjeturas, entendiendo que los teoremas no surgen de la nada.