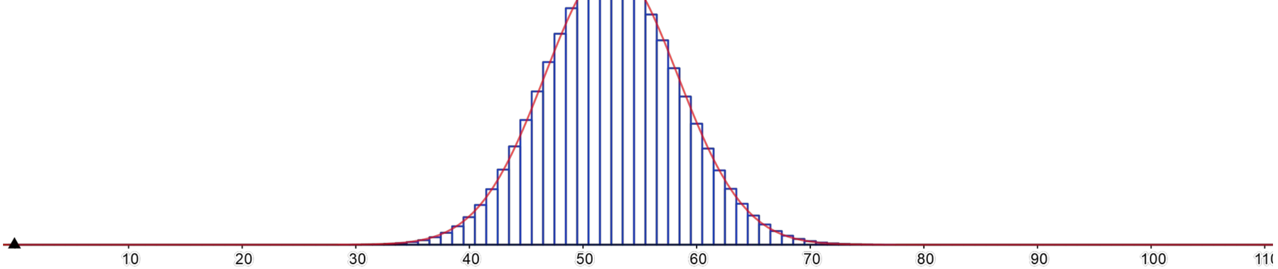
TEMA 5. VARIABLES ALEATORIAS

DISTRIBUCIONES BINOMIAL Y NORMAL

ÍNDICE



1. VARIABLE ALEATORIA DISCRETA

Definición. Una **variable aleatoria discreta X** es una función que asigna valores numéricos a los sucesos elementales de un espacio muestral

$$\begin{aligned} X : E &\rightarrow R \\ w_i &\rightarrow x_i \end{aligned}$$

Ejemplo. Lanzamos dos dados y definimos la variable aleatoria X, suma de los puntos obtenidos. El espacio muestral es = {(1, 1), (1, 2), (1, 3), (1, 4), …, (6, 6)} entonces se tiene

$$\begin{aligned} X : E &\rightarrow R \\ (1, 1) &\rightarrow 2 \\ (1, 2) &\rightarrow 3 \\ (1, 3) &\rightarrow 3 \\ &\dots \\ (6, 6) &\rightarrow 12 \end{aligned}$$

Definición. Función de probabilidad es la aplicación que asigna a cada valor de la variable aleatoria discreta X la probabilidad de que la variable tome dicho valor:

$$\begin{aligned} p : X &\rightarrow [0, 1] \\ x_i &\rightarrow P(X = x_i) \end{aligned}$$

Ejemplo. Lanzamos dos dados y definimos la variable aleatoria X, suma de los puntos obtenidos. La función de probabilidad de X es

$$\begin{aligned} &p(2) = P(X = 2) = P(\ (1, 1)\ ) = 1/36 \\ &p(3) = P(X = 3) = P((1, 2), (2, 1)) = 2/36 \\ &\dots \\ &p(12) = P(X = 12) = P((6,6)) = 1/36 \end{aligned}$$

Espacio Muestral	(1,1)	(1,2) (2,1)	(1,3) (2,2) (3,1)	(1,4) (2,3) (3,2) (4,1)	(1,5) (2,4) (3,3) (4,2) (5,1)	(1,6) (2,5) (3,4) (4,3) (5,2) (6,1)	(2,6) (3,5) (4,4) (5,3) (6,2)	(3,6) (4,5) (5,4) (6,3)	(4,6) (5,5) (6,4)	(5,6) (6,5)	(6,6)
x_i	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
p_i	**1/36**	**2/36**	**3/36**	**4/36**	**5/36**	**6/36**	**5/36**	**4/36**	**3/36**	**2/36**	**1/36**

Es habitual expresar la función de probabilidad en una tabla de la forma:

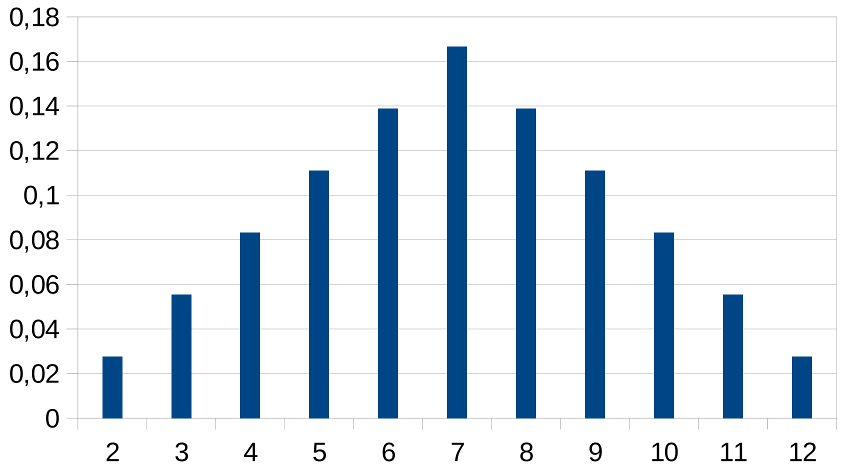
x_i	$p_i = P(X=x_i)$
2	1/36
3	2/36
…	…
12	1/36
	$\sum P(X=x_i)=1$

Definición. Función de distribución es la aplicación $F(x_i)$ que asigna a cada valor x_i de la variable aleatoria discreta X la probabilidad de que la variable tome valores menores o iguales que x_i: $F(x_i) = P(X = x_i)$.

Ejemplo:

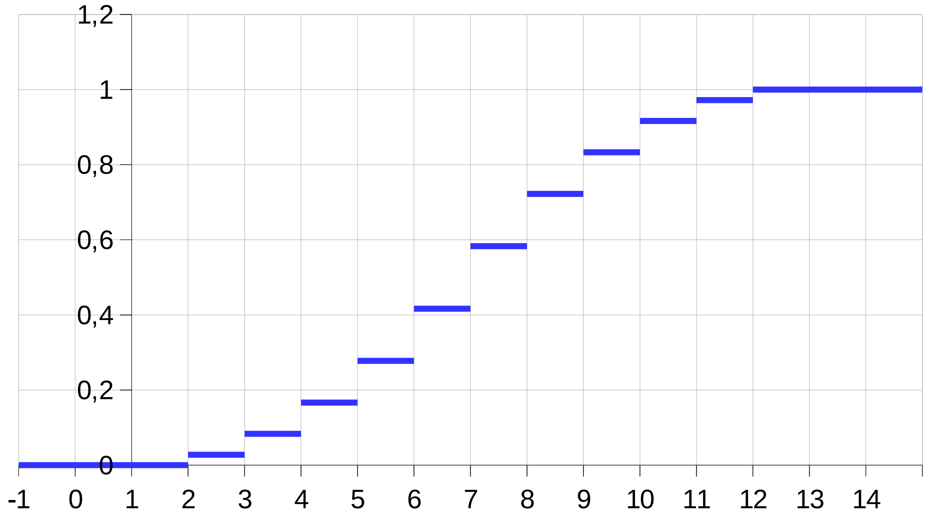
x_i	$p_i = P(X=x_i)$	$F(x_i)=P(X \leq x_i)$
2	1/36	1/36
3	2/36	3/36
...	...	...
12	1/36	36/36=1
	$\sum P(X=x_i)=1$	

Representaremos la función de probabilidad por un diagrama de barras y la función de distribución por una función a saltos:



Función de probabilidad

Función de distribución



Definición. Se define **media o esperanza de una variable aleatoria X y** se representa por **μ** al valor: $\mu = x_1 \cdot p_1 + x_2 \cdot p_2 + ... + x_n \cdot p_n = \sum_{i=1}^{n} x_i \cdot p_i$.

Definición. Se llama **varianza de una variable aleatoria X** y se representa por $\boldsymbol{\sigma^2}$ al valor

$$\sigma^2 = x_1^2 \cdot p_1 + x_2^2 \cdot p_2 + \ldots + x_n^2 \cdot p_n - \mu^2 = \sum_{i=1}^{n} x_i^2 \cdot p_i - \mu^2$$

A partir de la varianza se define la **desviación típica** σ como la raíz cuadrada de la varianza.

Ejemplo:

x_i	$p_i = P(X=x_i)$	$F(x_i)=P(X \le x_i)$	$x_i p_i$	$x_i^2 p_i$
2	1/36	1/36	2/36	4/36
3	2/36	3/36	6/36	18/36
...	...	...	...	...
11	2/36	35/36	22/36	242/36
12	1/36	36/36=1	12/36	144/36
	$\sum P(X=x_i)=1$		$\sum x_i \cdot P(X=x_i)=7$	$\sum x_i^2 \cdot P(X=x_i)=1974/36$

Media: $\mu = x_1 \cdot p_1 + x_2 \cdot p_2 + \ldots + x_n \cdot p_n = \sum_{i=1}^{n} x_i \cdot p_i = 7$

Varianza: $\sigma^2 = \sum_{i=1}^{n} x_i^2 \cdot p_i - \mu^2 = 1974/36 - 7^2 = 35/6$

Desviación típica: $\sigma = \sqrt{\frac{35}{6}}$

EJERCICIOS DE VARIABLE ALEATORIA DISCRETA:

1. Sea X la variable aleatoria que cuenta el número de cruces al lanzar dos monedas. Calcula:
   a) Su función de probabilidad y represéntala.
   b) Su función de distribución y represéntala.
   c) Su media, varianza y desviación típica.

2. En una urna hay 2 bolas blancas y 4 bolas negras. Sea X la variable aleatoria que cuenta el número de bolas blancas al extraer dos bolas de la urna. Calcula:
   a) Su función de probabilidad y represéntala.
   b) Su función de distribución y represéntala.
   c) Su media, varianza y desviación típica.

3. En una urna hay 2 bolas blancas y 4 bolas negras. Sea X la variable aleatoria que cuenta el número de bolas blancas al extraer dos bolas de la urna CON DEVOLUCIÓN. Calcula:
   a) Su función de probabilidad y represéntala.
   b) Su función de distribución y represéntala.
   c) Su media, varianza y desviación típica.

4. Se lanzan dos dados trucados para que la probabilidad de salir 5 sea el doble del resto de las caras. Sea la variable X que cuenta el número de 5 obtenidos. Calcula:
   a) Su función de probabilidad y represéntala.
   b) Su función de distribución y represéntala.
   c) Su media, varianza y desviación típica.

Otros ejemplos en esta página web.

2. VARIABLE ALEATORIA DISCRETA - DISTRIBUCIÓN BINOMIAL

Una variable aleatoria X sigue una **distribución binomial** si cumple:

1) Estudiamos un experimento compuesto por **n experimentos independientes** entre sí.

2) El resultado de cada experimento elemental ha de admitir sólo dos categorías, a las que se denomina **éxito y fracaso**.
El valor de ambas posibilidades ha de ser constante en todos los experimentos, y se denotan como ***p* y *q*** respectivamente, o *p* y 1-*p* de forma alternativa.

3) Se designa por ***X* a la variable que mide el número de éxitos** que se han producido en los *n* experimentos.

Cuando se dan estas circunstancias, se dice que la variable *X* sigue una distribución de probabilidad binomial, y se denota **X ~ *B(n,p)***.

La función de probabilidad de la binomial es: $P(X=k)=\binom{n}{k}p^k q^{n-k}$ donde q = 1-p

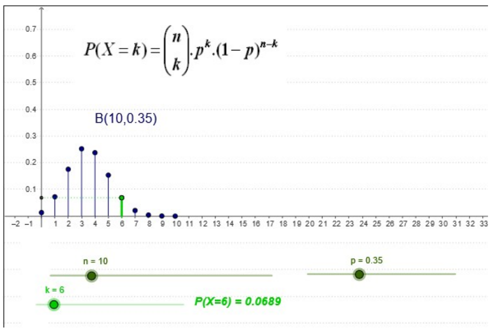
Siendo $\binom{n}{k}$ un número combinatorio: $\binom{n}{k}=\frac{n!}{k!(n-k)!}$ teniendo en cuenta que

$$n!=n\cdot(n-1)\cdot(n-2)\cdot\ldots\cdot 1 \quad \text{y} \quad 0!=1$$

$$\binom{n}{0}=\frac{n!}{0!(n-0)!}=1 \quad , \quad \binom{n}{n}=\frac{n!}{n!(n-n)!}=1 \quad \text{y} \quad \binom{n}{1}=\frac{n!}{1!(n-1)!}=n \quad .$$

Otra propiedad útil es $\binom{n}{k}=\binom{n}{n-k}$ por tanto $\binom{n}{n-1}=\binom{n}{1}=n$.

Representación de la **función de probabilidad de una variable binomial**:



Media: μ = $\sum_{k=0}^{n} k\cdot P(X=k)=\sum_{k=0}^{n} k\cdot\binom{n}{k}p^k q^{n-k}=n\cdot p$

Varianza: $\sigma^2=\sum_{k=0}^{n} k^2\cdot P(X=k)$ - μ ^{2} = $n\cdot p\cdot q$

Desviación típica: $\sigma=\sqrt{\sigma^2}=\sqrt{n\cdot p\cdot q}$

2.1. DISTRIBUCIÓN BINOMIAL - EJERCICIO MODELO:

Ejercicio: La probabilidad de ganar un sorteo es 1/3. Suponiendo que se van a jugar cuatro sorteos:
a) Describe la variable aleatoria que cuenta el número de sorteos ganados. Justifica si se trata o no de una variable binomial.
b) Haz la tabla de la función de probabilidad y de distribución.
c) Representa estas funciones.
d) Calcula a partir de la tabla la media y la varianza de esta variable.
e) Comprueba el cálculo comparando con las fórmulas de la media y la varianza de una binomial.
f) Probabilidad de que gane la mitad de los sorteos.
g) Probabilidad de que gane más de la mitad de los sorteos.
h) Probabilidad de que gane algún sorteo.

SOLUCIÓN:

a) Describe la variable aleatoria que cuenta el número de sorteos ganados. Justifica si se trata o no de una variable binomial:

Sea el suceso elemental G = "ganar un sorteo" entonces el espacio muestral de este experimento compuesto es:
$E = \{ GGGG,\ GGG\overline{G},\ GG\overline{G}G,\ G\overline{G}GG,\ \overline{G}GGG,\ GG\overline{G}\overline{G},\ G\overline{G}\overline{G}G,\ \overline{G}\overline{G}GG,\ \overline{G}G\overline{G}G,\ \overline{G}GG\overline{G},\ G\overline{G}G\overline{G},$
$G\overline{G}\overline{G}\overline{G},\ \overline{G}G\overline{G}\overline{G},\ \overline{G}\overline{G}G\overline{G},\ \overline{G}\overline{G}\overline{G}G,\ \overline{G}\overline{G}\overline{G}\overline{G}\}$

La variable aleatoria X = "nº de sorteos ganados" asigna a cada suceso del espacio muestral un número:

X: **E**	$\to$ $\mathbb{R}$	
$\{GGGG\}$	4	Fíjate que hay 1 suceso que nos llevan al 4: $\binom{4}{4}=1$
$\{GGG\overline{G}\}$	3	Hay 4 sucesos que nos llevan al 3: $\binom{4}{3}=4$
$\{GG\overline{G}G\}$	3	
….		
$\{\overline{G}G\overline{G}G\}$	2	Hay 6 sucesos que nos llevan al 2: $\binom{4}{2}=6$
…		
$\{\overline{G}\overline{G}\overline{G}G\}$	1	Hay 4 sucesos que nos llevan al 1: $\binom{4}{1}=4$
$\{\overline{G}\overline{G}\overline{G}\overline{G}\}$	0	Hay 1 suceso que nos llevan al 0: $\binom{4}{0}=1$

Los sucesos elementales que forman cada suceso de este experimento son independientes con probabilidad de "éxito" de ⅓. Por tanto X sigue una distribución binomial B(4, ⅓).

Para calcular, por ejemplo, la probabilidad de ganar 3 partidos: P(X=3) sumamos las probabilidades de los sucesos "ganar tres partidos":

$P(X=3) = P(\{GGG\overline{G}\}) + P(\{GG\overline{G}G\}) + P(\{G\overline{G}GG\}) + P(\{\overline{G}GGG\}) =$

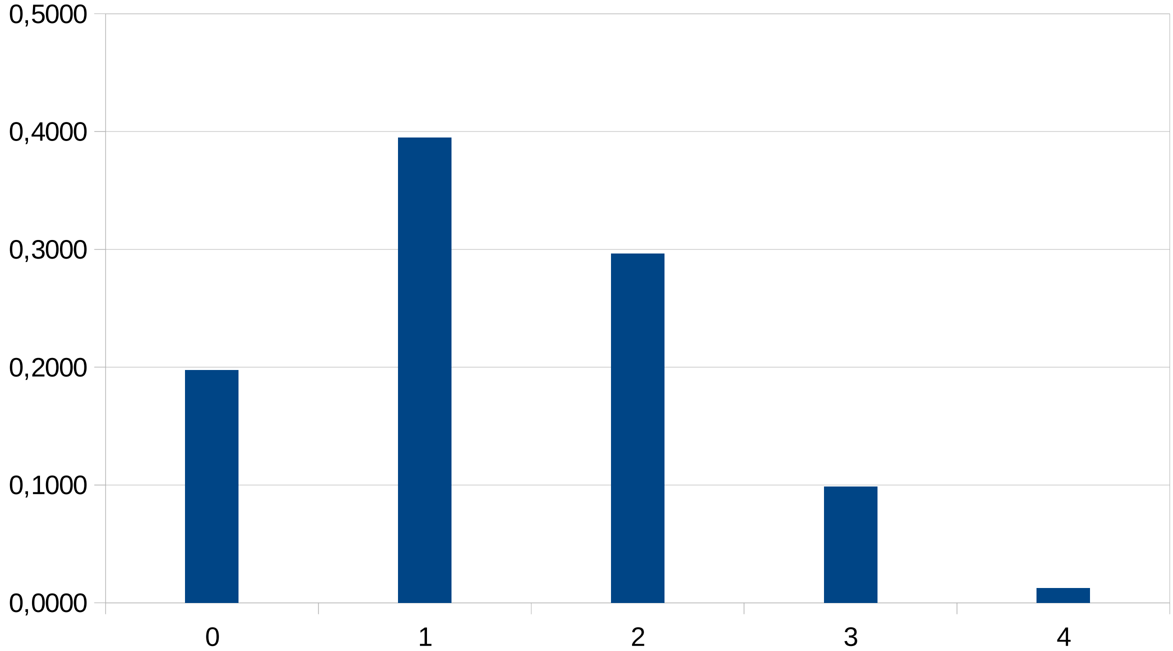
$$\frac{1}{3}\cdot\frac{1}{3}\cdot\frac{1}{3}\cdot\frac{2}{3} + \frac{1}{3}\cdot\frac{1}{3}\cdot\frac{2}{3}\cdot\frac{1}{3} + \frac{1}{3}\cdot\frac{2}{3}\cdot\frac{1}{3}\cdot\frac{1}{3} + \frac{2}{3}\cdot\frac{1}{3}\cdot\frac{1}{3}\cdot\frac{1}{3} = 4\left(\frac{1}{3}\right)^3\cdot\frac{2}{3} = \frac{8}{81}$$

O bien, utilizamos la fórmula de la probabilidad de la binomial: $P(X=3)=\binom{4}{3}\left(\frac{1}{3}\right)^3\left(\frac{2}{3}\right)^{4-3}$

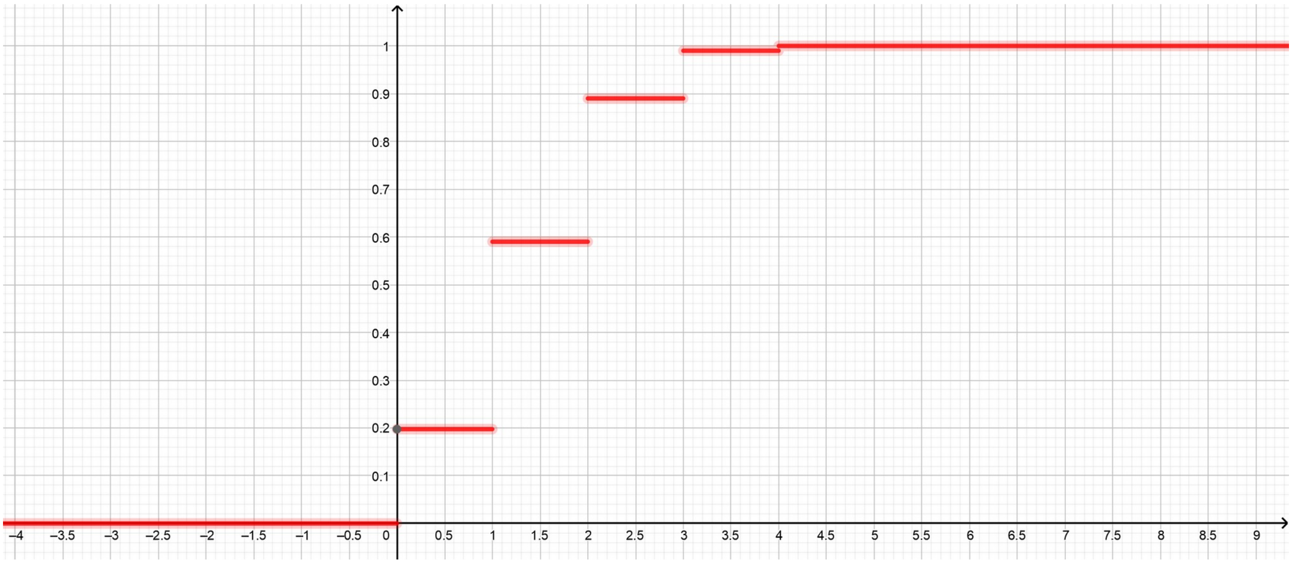
b) Haz la tabla de la función de probabilidad y de distribución:

k	P(X=k)	P(X≤k)	k.P(X=k)	k².P(X=k)
0	$P(X=0)=\binom{4}{0}\left(\frac{1}{3}\right)^0\left(\frac{2}{3}\right)^{4-0}=\frac{16}{81}$	$\frac{16}{81}$	0	0
1	$P(X=1)=\binom{4}{1}\left(\frac{1}{3}\right)^1\left(\frac{2}{3}\right)^{4-1}=\frac{32}{81}$	$\frac{48}{81}$	$\frac{32}{81}$	$\frac{32}{81}$
2	$P(X=2)=\binom{4}{2}\left(\frac{1}{3}\right)^2\left(\frac{2}{3}\right)^{4-2}=\frac{24}{81}$	$\frac{72}{81}$	$\frac{48}{81}$	$\frac{96}{81}$
3	$P(X=3)=\binom{4}{3}\left(\frac{1}{3}\right)^3\left(\frac{2}{3}\right)^{4-3}=\frac{8}{81}$	$\frac{80}{81}$	$\frac{24}{81}$	$\frac{72}{81}$
4	$P(X=4)=\binom{4}{4}\left(\frac{1}{3}\right)^4\left(\frac{2}{3}\right)^{4-4}=\frac{1}{81}$	$\frac{81}{81}=1$	$\frac{4}{81}$	$\frac{16}{81}$
	$\frac{81}{81}=1$		$\frac{108}{81}=\frac{4}{3}$	$\frac{216}{81}=\frac{8}{3}$

c) Representa estas funciones:



Función de probabilidad



Función de distribución

d) Calcula a partir de la tabla la media y la varianza de esta variable:

Media: m= $\sum_{k=0}^{4} k \cdot P(X=k) = \frac{4}{3}$

Varianza: $\sigma^2 = \sum_{k=0}^{4} k^2 \cdot P(X=k) -$ m^2 = $\frac{8}{3} - \left(\frac{4}{3}\right)^2 = \frac{24}{9} - \frac{16}{9} = \frac{8}{9}$

Desviación típica: $\sigma = \sqrt{\sigma^2} = \sqrt{\frac{8}{9}} = \frac{2\sqrt{2}}{3}$

e) Comprueba el cálculo comparando con las fórmulas de la media y la varianza de una binomial:

X ~ Bi (4, ⅓) n = 4 p = ⅓

Media: m= $n \cdot p = 4 \cdot \frac{1}{3} = \frac{4}{3}$

Varianza: $\sigma^2 = n \cdot p \cdot q = 4 \cdot \frac{1}{3} \cdot \frac{2}{3} = \frac{8}{9}$

f) Probabilidad de que gane la mitad de los sorteos:

$$P(X=2) = \binom{4}{2} \left(\frac{1}{3}\right)^2 \left(\frac{2}{3}\right)^{4-2} = \frac{24}{81} = 0{,}2963$$

g) Probabilidad de que gane más de la mitad de los sorteos:

$$P(X>2) = P(X=3) + P(X=4) = \binom{4}{3}\left(\frac{1}{3}\right)^3\left(\frac{2}{3}\right)^{4-3} + \binom{4}{4}\left(\frac{1}{3}\right)^4\left(\frac{2}{3}\right)^{4-4} = \frac{9}{81} = \frac{1}{9} = 0{,}1111$$

h) Probabilidad de que gane algún sorteo:

$$P(X>0) = 1 - P(X=0) = 1 - \binom{4}{0}\left(\frac{1}{3}\right)^0\left(\frac{2}{3}\right)^{4-0} = 1 - \frac{16}{81} = \frac{65}{81} = 0{,}8025$$

2.2. DISTRIBUCIÓN BINOMIAL - EJERCICIOS PROPUESTOS:

1. Se considera una v.a. cuya distribución es binomial del tipo B(3, 1/4). Se pide:
a) Determinar y representar gráficamente su función de probabilidad.
b) Determinar y representar la función de distribución.
c) Calcular $P(2 < X \leq 3)$ y $P(2 \leq X \leq 3)$.

2. De una baraja de 40 cartas se extraen, una a una y con devolución, 5 cartas, y se anota el número de copas obtenidas. Describe este experimento mediante una v.a. X. Calcula la media y la desviación típica, así como la probabilidad de obtener, a lo sumo, tres copas.

3. Al inspeccionar 3240 piezas hechas por una misma máquina, resultó que 108 eran defectuosas. Admitimos que la producción sigue en las mismas condiciones. Si se eligen 10 piezas hechas por esa máquina, ¿cuál es la probabilidad de que por lo menos dos sean defectuosas?

4. En un determinado país, el 90% de sus habitantes son bilingües en francés e inglés. Si se escogen 20 personas al azar:
a) ¿Cuál es la probabilidad de que haya, exactamente, cinco personas bilingües, entre las elegidas?
b) ¿Cuál es la probabilidad de que al menos dos sean bilingües?
c) ¿Cuántos cabe esperar que sean bilingües?

5. Una urna contiene 5 bolas verdes y 7 blancas. Se saca una bola al azar, se apunta el color, y se devuelve a la urna. Suponiendo que esa experiencia se repite cinco veces, hallar
a) la probabilidad de obtener dos bolas verdes
b) la probabilidad de obtener a lo sumo una verde
c) la media y desviación típica de la variable que describe el número de bolas verdes obtenidas.

6. Se ha estudiado que 1/3 de los alumnos de Bachillerato no leen nunca la prensa diaria. Tomando una muestra al azar de 10 alumnos estudiar las probabilidades siguientes:
a) Encontrar dos alumnos que no leen la prensa
b) Más de tres alumnos que no leen la prensa
c) Por lo menos nueve alumnos que no leen la prensa

7. Sea X una v.a. binomial. Sabiendo que m = 3 y $s^2 = 12/5$. Determínese:
a) Los parámetros n y p de la distribución binomial de X.
b) La probabilidad del suceso (X < 2).

8. Considérese una v.a. X cuya distribución es binomial de tipo B(8,p). Sabiendo que su varianza es igual a 2, se pide:
a) Su función de probabilidad y su media.
b) Calcular $P(X \leq 2)$

9. Sea X una v.a. de distribución binomial, con m = 2 y desviación típica $\frac{2\sqrt{3}}{3}$. Determina la probabilidad de que X sea mayor que 0.

10. Un examen tipo “test” consta de cinco preguntas, en cada una de las cuales se adjuntan tres posibles respuestas de las que sólo una es correcta. Para superar el examen, se exige acertar un mínimo de cuatro respuestas. ¿Qué probabilidad hay de que una persona aprueba el examen si responde al azar?

SOLUCIONES:

1. c) $P(2 < X \leq 3) = 1/64$ y $P(2 \leq X \leq 3) = 5/32$

2. X ~ B(5, ¼) $\mu = 5/4$ $\sigma = \frac{\sqrt{15}}{4}$ $P(X \leq 3) = 0{,}9844$

3. 0,0418

4. a) $9{,}155 \cdot 10^{-12} \approx 0$ b) $0{,}99999999999999999819 \approx 1$ c) 18

5. a) 0,3446 b) 0,3088 c) $\mu = 25/12$ y $\sigma = \frac{5\sqrt{7}}{12}$

6. a) 0,1951 b) 0,4407 c) $0{,}0003556 \approx 0{,}0004$

7. a) n = 15 p = 1/5 b) $P(X < 2) = 0{,}1671$

8. b) 37/256

9. 0,9122

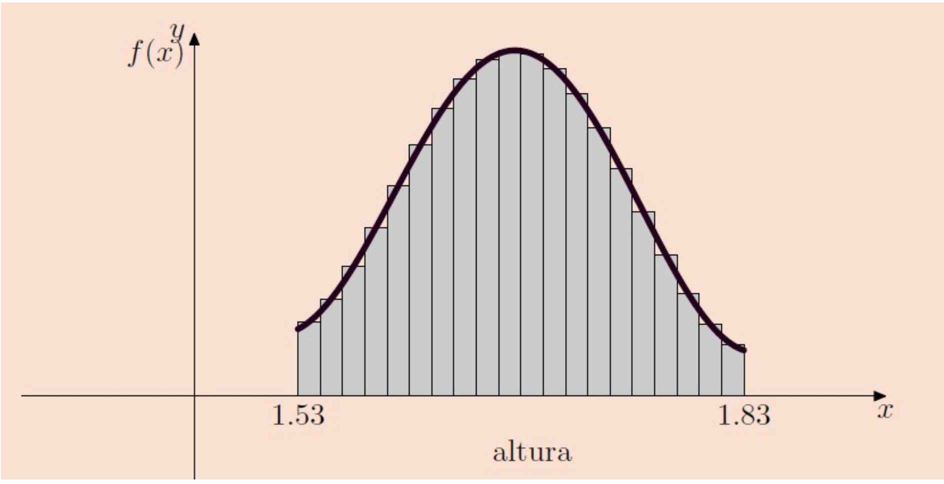
10. 0,0453

3. VARIABLE ALEATORIA CONTINUA

Diremos que una **variable aleatoria es continua cuando puede tomar cualquier valor en un intervalo**. Por ejemplo, el peso o la altura de una persona. Para concretar, consideremos que representamos en un histograma las medida de la alturas X de los chicos de 15 años. Si tomamos más y más observaciones y haciendo clases cada vez más finas, el histograma tenderá a una curva que describirá el comportamiento de la variable estudiada, como muestra la figura.

Eso hace pensar en una función matemática f(x) que modelice la frecuencia relativa de la altura para la población de los chicos de 15 años.

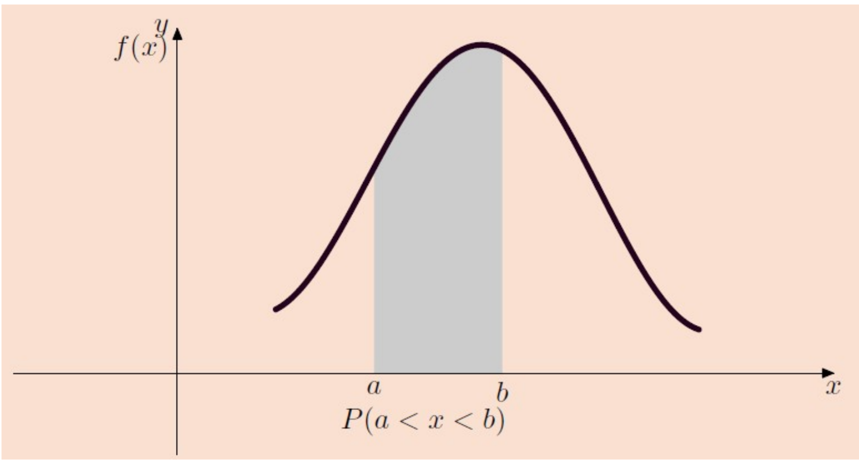
A dicha función **f(x)** se le llama **función de densidad**.



Definición: Una variable aleatoria X se dice que es continua cuando tiene asociada una función de densidad f(x) que cumple:

1. $f(x) \geq 0$ para todos los valores donde está definida.
2. $\int_{-\infty}^{+\infty} f(x)\,dx = 1$ (El área bajo la curva es uno)

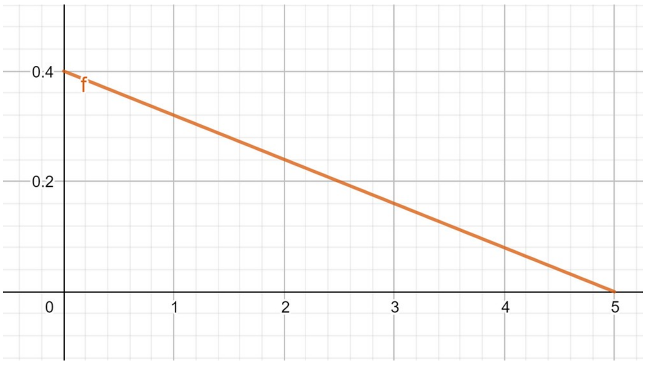
La probabilidad se determina con $P(a < x < b) = \int_a^b f(x)\,dx$ que corresponde al área de la región limitada por f(x), el eje x y las rectas x = a y x = b.



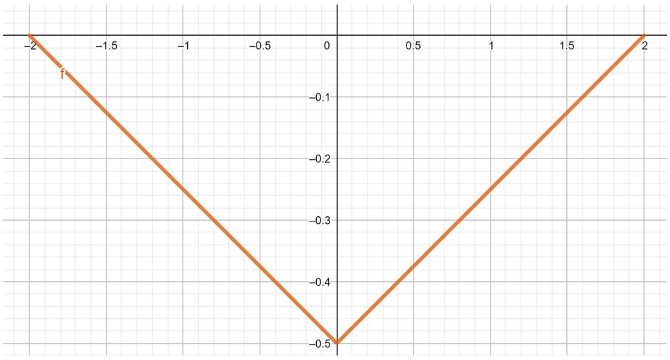
3.1. VARIABLE ALEATORIA CONTINUA - EJERCICIOS PROPUESTOS:

1. ¿Es f(x) una función de densidad?

a)



b)



c) $f(x)=\begin{cases} \dfrac{x+2}{2} & -2\leq x\leq 0 \\ \dfrac{-x+2}{2} & 0\leq x\leq 2 \end{cases}$

d) $f(x)= \dfrac{1}{4} \quad 0 \leq x \leq 4$

2. Calcula $P(X < 2)$, $P(X > 1)$, $P(X = 3)$ y $P(1 < X \leq 2)$ para las funciones del ejercicio anterior que sean funciones de densidad.

3. Sea $f(x) = k$ para $0 \leq x \leq 10$, calcula k para que sea una función de densidad. Calcula $P(X \leq 3)$, $P(X = 2)$ y $P(1 < X \leq 8)$.

SOLUCIONES:

1. a) Si b) No c) No d) Si

2. a) $P(X < 2) = 16/25$, $P(X > 1) = 16/25$, $P(X = 3) = 0$ y $P(1 < X \leq 2)= 7/25$

d) $P(X < 2) = 1/2$, $P(X > 1) = 3/4$, $P(X = 3) = 0$ y $P(1 < X \leq 2) = ¼$

3. k = 1/10 $P(X \leq 3) = 3/10$, $P(X = 2) = 0$ y $P(1 < X \leq 8) = 7/10$.

3.2. VARIABLE ALEATORIA CONTINUA - LA DISTRIBUCIÓN NORMAL

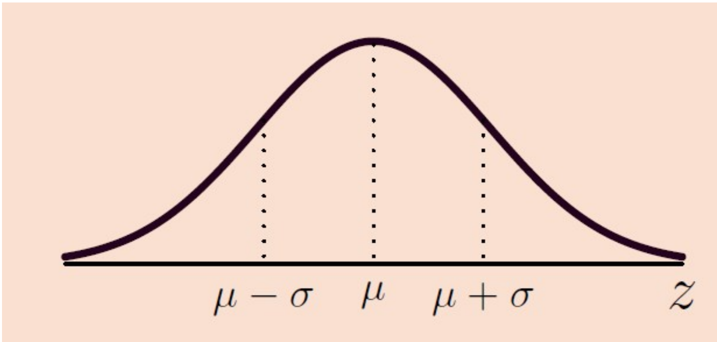
La función de distribución normal juega un papel central en la estadística ya que, además de sus interesantes propiedades de reproductividad y de aproximación de otras distribuciones, sirve para modelizar una gran cantidad de situaciones prácticas.

Definición - Distribución Normal: Una variable aleatoria X se dice que tiene una distribución de probabilidades normal con media μ y varianza σ^2 y se denomina **N(μ, σ^2)** si su función de densidad es de la forma:

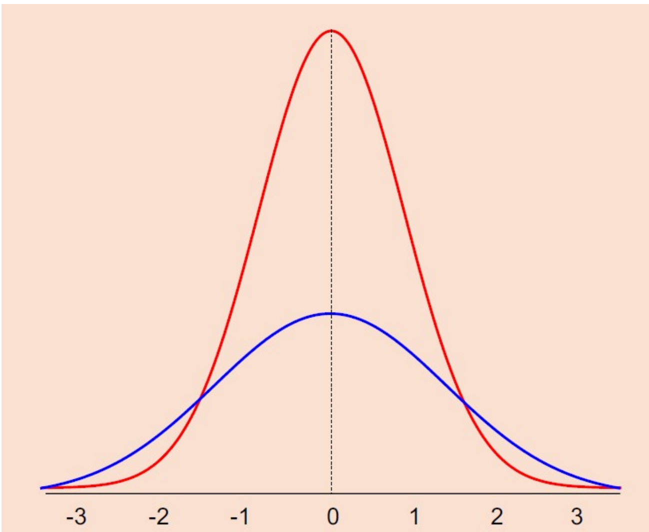
$$f(x)=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-u)^2}{2\sigma^2}} \quad , \quad -\infty < x < +\infty$$

La media y varianza de una variable normal coinciden con los parámetros μ y σ^2 de la misma.

En la figura se muestra la función de densidad y se aprecia que es simétrica respecto de x = μ y los puntos de inflexión de la función de densidad se encuentran a una distancia s de la media.
Su gráfica se llama **campana de Gauss**.



En la figura se muestran dos distribuciones normales, con media μ = 0 y desviaciones típicas = 1 y = 0.5. Como las dos tienen la misma media μ = 0, la función roja es mas estrecha pues tiene menos variabilidad que la azul y los datos se concentran alrededor de la media.



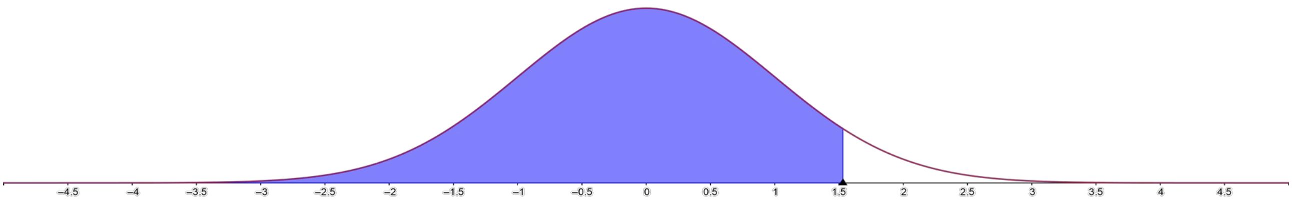
La normal N(0; 1)

Primero estudiaremos la normal típica, que tiene media μ = 0 y desviación típica = 1, N(0; 1) y se designa con la letra Z. Para calcular las probabilidades $P(Z \leq z_0) = f(z_0)$ se diseña una tabla que proporciona las respectivas probabilidades.

3.2.1. USO DE LA TABLA DE LA NORMAL N(0, 1)

Caso I: $P(Z \leq z_0)$ y z_0 es positivo.

Ejemplo: $P(Z \leq 1{,}53)$ buscamos en la tabla la fila del 1,5 y la columna del 0,03 obteniendo $P(Z \leq 1{,}53)=0{,}9370$.



Fíjate que la campana de Gauss es simétrica respecto al eje Y con área total igual a 1, por tanto, $P(Z \leq 0) = 0{,}5$.

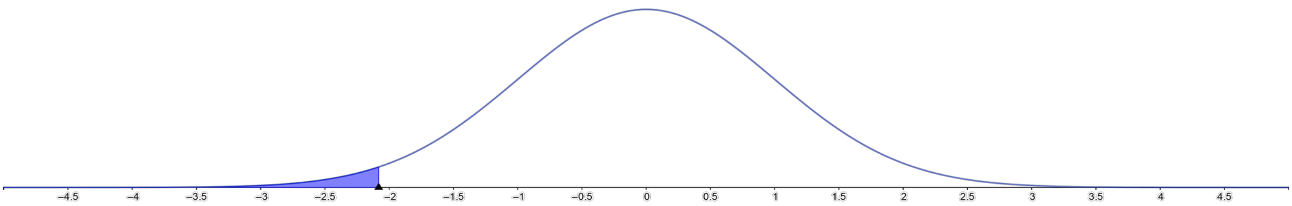
Caso II: $P(Z \leq -z_0)$ y z_0 es positivo. $\boxed{P(Z \leq -z_0) = 1 - P(Z \leq z_0)}$

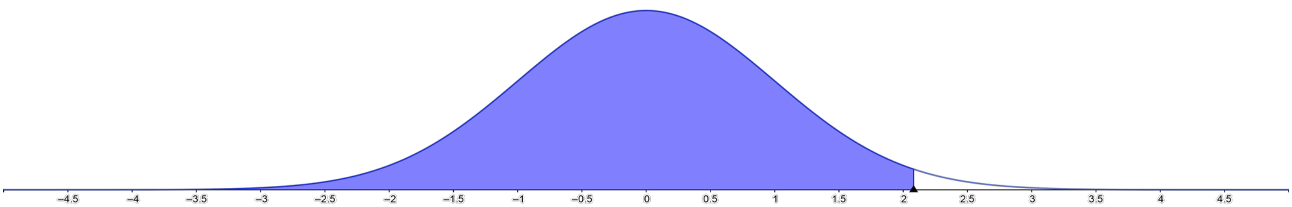
Ejemplo: $P(Z \leq -2{,}08) = ?$

Utilizamos la simetría de la curva: $P(Z \leq -2{,}08) = P(Z \geq 2{,}08) = 1 - P(Z < 2{,}08) = 1 - P(Z \leq 2{,}08)$

Ten en cuenta que $\mathbf{P(Z \leq z_0) = P(Z < z_0) + P(Z = z_0) = P(Z < z_0)}$ ya que por definición

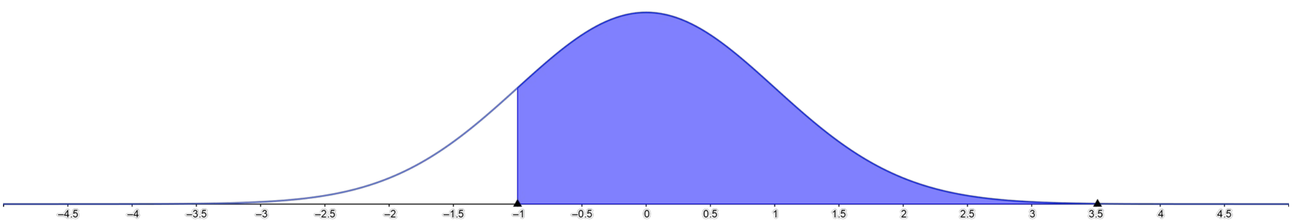
$$P(Z=z_0)=\int_{z_0}^{z_0} f(x)\,dx=0 \quad .$$





Buscamos en la tabla en la fila del 2 y la columna del 0,08 obteniendo $P(Z\leq -2{,}08) = 1 – P(Z\leq 2{,}08) = 1- 0{,}9812 = 0{,}0188$

Caso III: $\boxed{P(z_0 \leq Z \leq z_1) = P(Z \leq z_1) - P(Z \leq z_0)}$



Ejemplo: $P(1 \leq Z \leq 3{,}51) = P(Z \leq 3{,}51) – P(Z \leq 1) = 0{,}9998 – 0{,}8413 = 0{,}1585$

También puedes calcular las probabilidades de la distribución normal en Geogebra: https://www.geogebra.org/classic#probability

3.2.2. USO INVERSO DE LA TABLA N(0,1)

Nos dan la probabilidad y calculamos el valor de la variable z_0 que acumula dicha probabilidad.

Caso I: P(Z ≤ z_0)=p ≥ 0,5 Hallar z_0 de la normal Z ~N(0, 1) tal que P(Z ≤ z_0) = p

Ejemplo: P(Z ≤ z_0) = 0,7019 buscamos en la tabla 0,7019 y nos fijamos que corresponde a la fila 0,5 columna 0,03. Por tanto, z_0 = 0,53.

Caso II: P(Z ≤ z_0)=p < 0,5 Para hallar z_0 debemos tener en cuenta que debe ser negativo, por tanto, P(Z ≤ z_0) = 1- P(Z ≤ -z_0) = p.

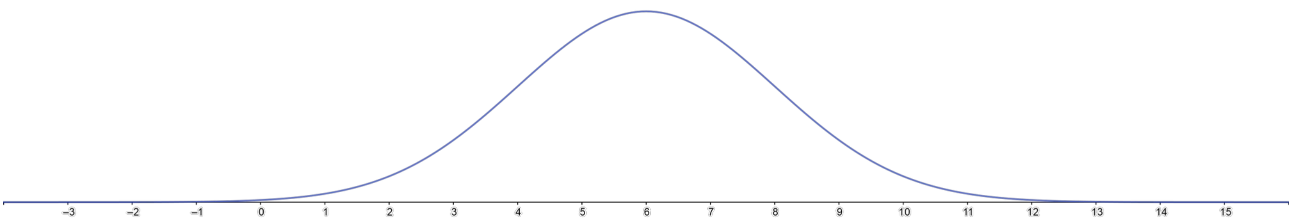
Ejemplo: P(Z ≤ z_0) = 0,2877 → P(Z ≤ -z_0) = 1 – 0,2877 = 0,7123 → Buscamos en la tabla 0,7123 que se corresponde con la fila 0,5 y columna 0,06. Por tanto, -z_0= 0,56 y z_0= -0,56.

3.2.3. TIPIFICAR UNA VARIABLE NORMAL N (μ, σ) - EJERCICIO MODELO

Sea X una variable aleatoria normal $X \sim N(\mu, \sigma)$. Entonces, la variable Z = $(X - \mu) / \sigma$ cumple que $Z \sim N(0,1)$.

Ejemplo: Supongamos que una variable X, que describe la nota media de 2º de Bachillerato del alumnado de un instituto sigue una distribución normal de media 6 y desviación típica 2:

$$X \sim N(6, 2) \rightarrow \frac{X-6}{2} \sim N(0,1)$$



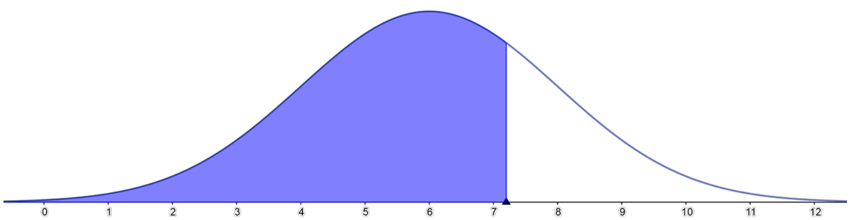
Elegido un alumno al azar, calculemos las probabilidades siguientes:

a) Tenga una nota media de un 7: Evidentemente, P (X = 7) = 0 al ser X una variable continua.

b) Tenga una nota media menor que 7,2:

$$P(X \leq 7{,}2) \underbrace{=}_{tipificamos} P\left(\frac{X-6}{2} \leq \frac{7{,}2-6}{2}\right)$$

$$P(Z \leq 0{,}6) \underbrace{=}_{tabla\, N(0,1)} 0{,}7257$$



c) Tenga una nota media superior a 8:

$$P(X>8)=1-P(X \leq 8) \underbrace{=}_{tipificamos} 1-P\left(\frac{X-6}{2} \leq \frac{8-6}{2}\right)=1-P(Z \leq 1) \underbrace{=}_{tabla\, N(0,1)} 1-0{,}8413=0{,}1587$$

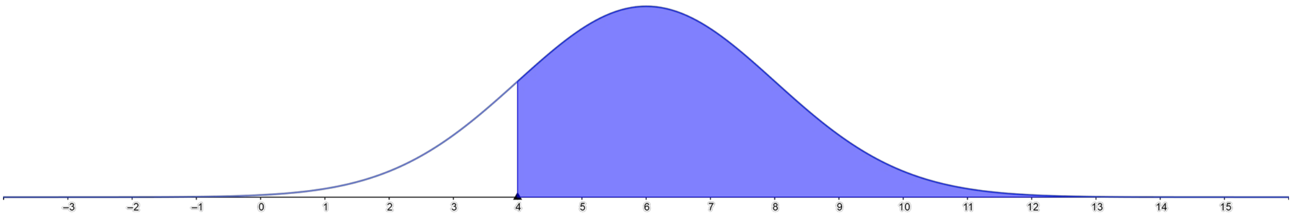
d) Sabiendo que hay 80 alumnos en 2ª de bachillerato, ¿cuántos habrán suspendido?

$$P(X<5)\underbrace{=}_{tipificamos}P\left(\frac{X-6}{2}<\frac{5-6}{2}\right)=P(Z<-0,5)\underbrace{=}_{simetría}1-P(Z\leq 0,5)\underbrace{=}_{tabla\,N(0,1)}1-0,6915=0,3085$$

Por lo tanto, habrá $0,3085\cdot 80=24,68$ aproximadamente 25 alumnos han suspendido.

e) Tenga más de un 4 de nota media:

$$P(X>4)\underbrace{=}_{tipificamos}P\left(\frac{X-6}{2}>\frac{4-6}{2}\right)=P(Z>-1)\underbrace{=}_{simetría}P(Z\leq 1)\underbrace{=}_{tabla\,N(0,1)}0,8413$$



f) Tenga entre un 4 y un 6:

$$P(4\leq X\leq 6)\underbrace{=}_{tipificamos}P\left(\frac{4-6}{2}\leq\frac{X-6}{2}\leq\frac{6-6}{2}\right)=P(-1\leq Z<0)=P(Z\leq 0)-P(Z\leq -1)=$$

$$\underbrace{=}_{simetría}0,5-(1-P(Z\leq 1))\underbrace{=}_{tabla\,N(0,1)}0,5-1+0,8413=0,3413$$

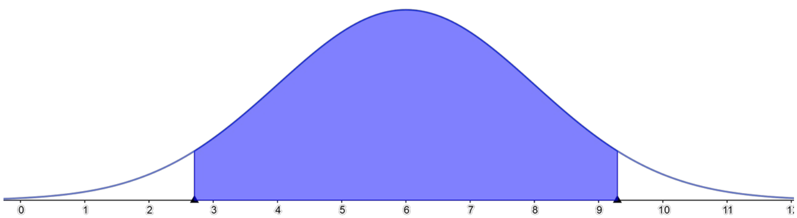
g) Se sabe que sólo un 5% del alumnado puede obtener una matrícula de honor. Halla la nota media que debe alcanzar un alumno/a para que se le conceda cumpla:

$$P(X>z_0)=0,05\Rightarrow P(X\leq z_0)=0,95\underbrace{\Rightarrow}_{tipificamos}P\left(\frac{X-6}{2}\leq\frac{z_0-6}{2}\right)=P\left(Z\leq\frac{z_0-6}{2}\right)=0,95$$

$$\underbrace{\Rightarrow}_{tabla\,N(0,1)}\frac{z_0-6}{2}=1,645\Rightarrow z_0=1,645\cdot 2+6=9,29$$

(0,95 se encuentra entre P(Z≤1,64) = 0,9495 y P(Z≤1,65) = 0,9505)

h) Encuentra un intervalo centrado en la media que encierre el 90% de las notas de los alumnos:
Buscamos un intervalo (z_1, z_0) que deje un 5% de las notas a la izquierda y un 5% a la derecha. El valor z_0 ya lo hemos calculado en el apartado anterior. Hallamos z_1 tal que $P(X\leq z_1)=0,05\Rightarrow$



$$P\left(Z\leq\frac{z_1-6}{2}\right)=0,05\Rightarrow 1-P\left(Z\leq-\frac{z_1-6}{2}\right)=0,05\Rightarrow P\left(Z\leq-\frac{z_1-6}{2}\right)=0,95$$

$$\Rightarrow -\frac{z_1-6}{2}=1,645\Rightarrow z_1=6-2\cdot 1,645=2,71$$

En el intervalo (2,71; 9,29) encontraremos el 90% de las notas de 2º de bachillerato.

3.3. DISTRIBUCIÓN NORMAL - EJERCICIOS PROPUESTOS

1. Suponiendo que Z es la distribución normal tipificada, calcula:

a) $P(Z > 0{,}45)$ b) $P(Z \leq -2{,}5)$ c) $P(1 < Z < 1{,}67)$ d) $P(Z > -1{,}5)$

e) $P(-1{,}5 < Z \leq -0{,}5)$ f) $P(\ |Z|\ < 1)$ g) $P(\ |Z - 2| < 0{,}5)$

2. Si Z es la distribución normal tipificada, halla k en los casos siguientes:

a) $P(Z < k) = 0{,}67$ b) $P(Z \leq k) = 0{,}975$ c) $P(Z > k) = 0{,}0901$

d) $P(1{,}5 \leq Z < k) = 0{,}044$ e) $P(-k < Z < k) = 0{,}9$ f) $P(-0{,}25 < Z < k) = 0{,}44$

g) $P(Z \leq k) = 0{,}1736$ h) $P(k \leq Z < -0{,}5) = 0{,}3023$

3. Las alturas de unos 500 estudiantes se distribuyen normalmente, con una media igual a 165 cm, y una desviación típica de 5 cm. ¿Cuántos de esos estudiantes tienen altura

a) Mayor que 175 cm. b) Menor que 157 cm.

c) Entre 160 y 170 cm. d) Igual a 180 cm.

4. Se ha elegido una muestra de lámparas fabricadas en una fábrica. La media de los diámetros ha resultado ser igual a 35 cm, y la desviación típica 25 mm. Se considera que una lámpara cumple el estándar de calidad si su diámetro es inferior a 360 mm y superior a 340 mm. Sabiendo que los diámetros se distribuyen normalmente, halla el porcentaje de lámparas que no cumplen el estándar mínimo de calidad.

5. Para que una mermelada sea calificada de light, debe contener entre 320 y 450g de azúcar por kilo. El peso medio de azúcar por bote ha resultado ser de 450g, con una desviación típica de 30g. Sabiendo que el contenido de azúcar se distribuye normalmente, halla el porcentaje de producción que no debe llevar la mención de light.

6. Sea X una v.a. normal de media μ y desviación típica σ. Halla μ y σ, sabiendo que $P(X \leq 12) = 0{,}6554$ y $P(X \geq 9) = 0{,}4207$

7. Supóngase que la v.a. que expresa el número de años de vida de las personas es normal del tipo N(82,5 ; 10,5). Halla la probabilidad de que una persona determinada viva más de 90 años.

8. Sea X una v.a. normal tal que $P(X \leq 25) = 0{,}8413$ y $P(X \leq 10) = 0{,}0228$

a) Calcula la media y la desviación típica de X.

b) Halla $P(16{,}5 \leq X \leq 22{,}5)$

c) Calcula el número k tal que $P(X \leq k) = 0{,}95$

9. Se supone que las calificaciones de un examen se distribuyen normalmente, con una media de 8,25 y una **varianza** de 0,25.

a) Sabiendo que se concede sobresaliente al 10% de los estudiantes que hicieron el examen, ¿cuál fue la puntuación mínima exigida para obtener sobresaliente?

b) Si la puntuación mínima exigida para aprobar fue de 6, ¿qué porcentaje de alumnos aprobaron?

10. Los retrasos en la llegada de los trenes de alta velocidad españoles siguen una ley normal con media 5 minutos. Sabemos que el 90% de los trenes llegan con un retraso menor de 8 minutos:

a) ¿Cuál es la desviación típica?

b) ¿Cuál es la probabilidad de que un tren llegue antes de la hora?

c) ¿Cuál es la probabilidad de que un tren se retrase más de 10 minutos?

11. La batería del móvil de una determinada marca dura un promedio de 40 horas, con una desviación típica de 2 horas. Suponiendo que la duración de las baterías es una variable normal:
a) ¿Qué porcentaje de baterías se espera que duren entre 35 y 55 horas?
b) Si una batería lleva funcionando 45 horas, ¿cuál es la probabilidad de que dure menos de 50 horas?

12. En la N(0,1) calcula:
a) El primer cuartil a / $P(Z \leq a) = 0,25$
b) El percentil 10 b / $P(Z \leq b) = 0,10$
c) El intervalo central que encierra el 95% de los datos (k / $P(-k \leq Z \leq k) = 0,95$)
d) Valor que es superado por el 99% de los datos.

13. Sea T la temperatura media del año en Ferrol. Si se distribuye según una normal de media 14,8ºC y desviación típica 3,4ºC, calcula:
a) Probabilidad de que la temperatura media de un año supere los 15ºC.
b) Probabilidad de que la temperatura media se sitúe entre 12ºC y 16ºC.
c) Intervalo que encierra el 90% central: (14,8 – k, 14,8 + k)
d) Percentil 95.

SOLUCIONES:

1. a) 0,3264 b) 0,0062 c) 0,1112 d) 0,9332 e) 0,2417 f) 0,6826 g) 0,0606

2. a) 0,44 b) 1,96 c) 1,34 d) 2 e) 1,645 f) 1 g) -0,94 h) -2,5

3. a) 11 b) 27 c) 341 d) 0

4. 68,92%

5. Aproximadamente un 50%

6. $\mu = 6$ y $\sigma = 15$

7. 0,2389

8. a) $\mu = 20$ y $\sigma = 5$ b) 0,4495 c) 28,225

9. a) 8,89 b) 100%

10. a) 2,34 b) 0,0162 c) 0,0162

11. a) 99,38% b) Aproximadamente 1

12. a) -0,675 b) -1,28 c) 1,96 d) -2,33

13. a) 0,4761 b) 0,4307 c) Entre 9,2ºC y 20,4ºC d) 20,4ºC

4. APROXIMACIÓN DE LA BINOMIAL POR LA NORMAL

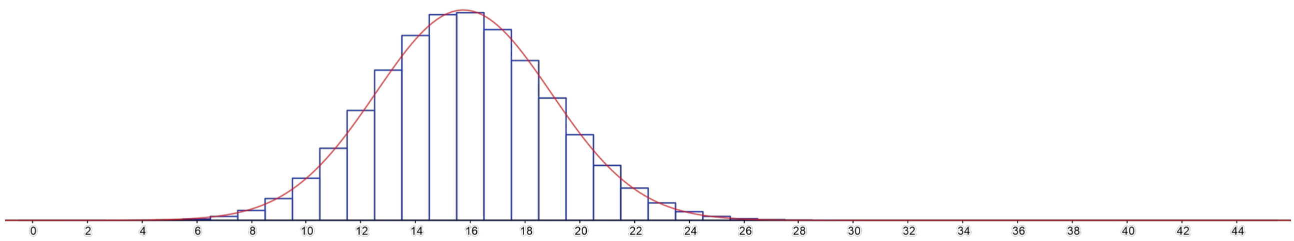
Supongamos que X en una variable aleatoria con distribución de probabilidad binomial B(n, p) y que el parámetro n tiende a infinito mientras que el parámetro p permanece constante. Entonces, puede demostrarse que la función de distribución de la variable aleatoria X tiende a una normal:

Proposición: (Aproximación de la binomial por la normal)

Sea X ~ B(n, p) tal que, n ≥ 30 ; np ≥ 5 y nq ≥ 5. Entonces la variable X se puede aproximar por una distribución normal $\boxed{X' \to N(np, \sqrt{npq})}$.

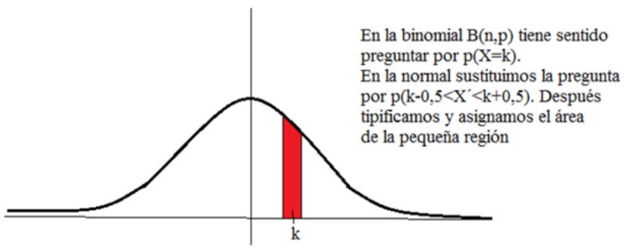
Esta proposición es consecuencia directa de un teorema más general llamado **Teorema Central del Límite.** Esto nos permitirá determinar probabilidades de la distribución binomial utilizando la distribución normal estándar.

Ejemplo: Sea X ~B(45, 0,35), por tanto, cumple las condiciones: n = 45 ≥ 30 ; np = 15,75 ≥ 5 y nq = 29,25 ≥ 5. Podemos aproximarla por una normal X' ~ N(15,75; 3,2).

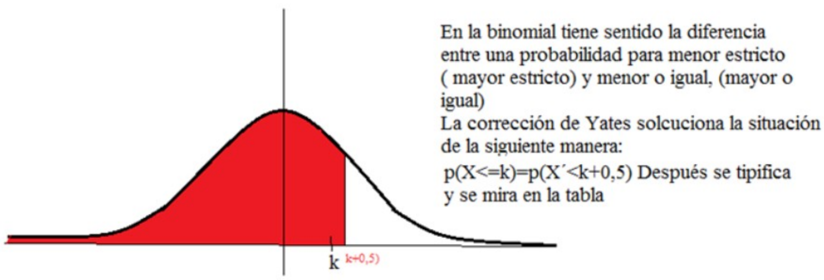


Nota: Como es lógico, el hecho de aproximar una variable discreta como la binomial, por una continua como la normal, da lugar a un problema. El problema es que mientras sí existe la probabilidad de que una binomial tome un valor concreto, en el caso de la normal, esta probabilidad es cero. Para resolver este problema, se introducen las llamadas **correcciones de Yates**, que consiste en considerar los valores de la variable discreta X como marcas de clase de intervalos de la forma siguiente:

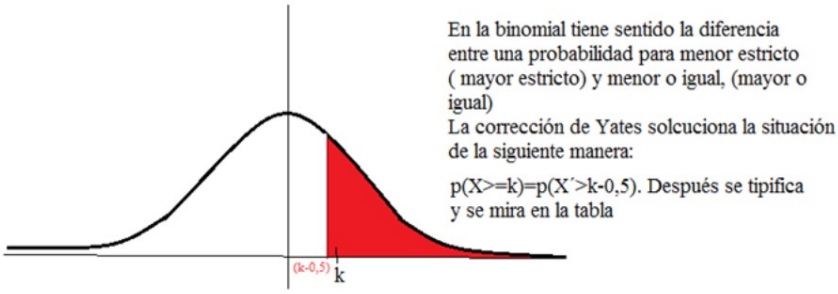
1. P (X = k) ≃ P (k - 0.5 ≤ X' ≤ k + 0.5).



2. P (X ≤ k) ≃ P (X' ≤ k + 0,5)



3. P (X < k) =P (X ≤ k - 1) ≃ P (X' ≤ k - 1 + 0,5) = P (X' ≤ k – 0,5).



4. P (X ≥ k) ≃ P (X' ≥ k – 0,5)

5. P (X > k) = P (X ≥ k + 1) ≃ P (X' ≥ k + 1 - 0,5) = ≃ P (X' ≥> k + 0,5)

Ejemplo: En el ejemplo anterior calculamos P(X = 17):

Sin aproximación por la normal: $P(X=17)=\binom{45}{17}\cdot 0{,}35^{17}\cdot 0{,}65^{28}=0{,}1131$

Aproximando por la normal: $P(X=17) \underbrace{\simeq}_{\substack{aprox.\,normal \\ con\ \ corrección\ \ de\ \ Yates}} P(17-0{,}5\le X'\le 17+0{,}5)=$

$$\underbrace{=}_{tipificación} P\left(\frac{16{,}5-15{,}75}{3{,}2}\le Z\le\frac{17{,}5-15{,}75}{3{,}2}\right)=P(0{,}23\le Z\le 0{,}55)=0{,}7088-0{,}5910=0{,}1178$$

Ejemplo: Durante cierta epidemia de gripe, enferma el 30% de la población. En un aula con 200 estudiantes de Medicina, calcula la probabilidad de que:
a) Ninguno padezca la enfermedad
b) Haya 60 estudiantes con gripe
c) Al menos 40 padezcan la enfermedad

Sea X = "nº de estudiantes enfermos de gripe" es una binomial con n = 200 y p = 0,3

a) $P(X=0)=\binom{200}{0}\,0{,}3^{0}\,0{,}7^{200}=1{,}05\cdot 10^{-31}$

b) $P(X=60)=\binom{200}{60}\,0{,}3^{60}\,0{,}7^{140}=7{,}04\cdot 10^{51}\cdot 4{,}3\cdot 10^{-32}\cdot 2{,}06\cdot 10^{-22}=0{,}0624$

(algunas calculadoras no pueden calcular este número combinatorio)

Podemos aproximar el resultado por la normal si se cumplen las condiciones: n = 200 ≥30, np = 60≥5 y nq=140≥5. Por tanto, X ~ B(200; 0,3) → X' ~N(60; 6,5)

$$P(X=60) \underbrace{\simeq}_{\substack{aprox.\,normal \\ con\ \ corrección\ \ de\ \ Yates}} P(60-0{,}5\le X'\le 60+0{,}5)\underbrace{=}_{tipificación} P\left(\frac{59{,}5-60}{6{,}5}\le Z\le\frac{60{,}5-60}{6{,}5}\right)=$$

$$=P(-0{,}08\le Z\le 0{,}08)=2\cdot P(Z\le 0{,}08)-1=2\cdot 0{,}5319-1=0{,}0638$$

c) $P(X\ge 40) \underbrace{\simeq}_{\substack{aprox.\,normal \\ con\ \ corrección\ \ de\ \ Yates}} P(X'\ge 40-0{,}5)\underbrace{=}_{tipificación} P\left(Z\ge\frac{39{,}5-60}{6{,}5}\right)=P(Z\ge -3{,}15)=$

$P(Z\le 3{,}15)=0{,}9992$

4.1. APROXIMACIÓN BINOMIAL A LA NORMAL - EJERCICIOS PROPUESTOS

1. El 2% de los chips fabricados por una máquina son defectuosos. Si se elige una muestra de 500 chips, hallar la probabilidad de que:

a) Haya 20 chips defectuosos

b) Haya entre 5 y 30 defectuosos

c) Haya menos de 15 defectuosos

2. En cierto país la probabilidad de que nazca un niño es 0,51. Si se producen en un año 25000 nacimientos, ¿cuál es la probabilidad de que el número de niños nacidos sea exactamente de 12500? ¿cuál es la probabilidad de que el número de niños nacidos sean más de 12500 y menos de 12800?

3. En cierta facultad universitaria, 3 de cada 4 alumnos finalizan sus estudios en cuatro años. Si en una promoción comenzaron 600 alumnos, ¿qué probabilidad hay de que más de 470 alumnos finalizaran sus estudios en 4 años?

4. Una fábrica de coches fabrica 1500 al día. Si la probabilidad de que uno sea defectuoso es del 0,5%, ¿cuál es la probabilidad de que en un día el número de coches defectuosos sea menor de 5? ¿Y de que sea igual a 4?

5. Se lanza **n** veces un dado y se define una variable aleatoria que cuenta el número de 5 obtenidos. Halla la probabilidad de obtener un 10% de 5 en los siguientes casos:

a) n = 10 b) n = 100 c) n= 200 d) n = 1000

e) Compara los resultados. ¿Observas alguna tendencia?

6. Se lanza una moneda 600 veces. Hallar la probabilidad de

a) Obtener a lo sumo 250 caras.

b) Obtener más de 335 cruces.

SOLUCIONES:

1. a) 0,0017 b) 0,9608 c) 0,9251

2. a) 0 b) 0,7349

3. 0,0268

4. a) 0,1357 b) 0,0727 (sin aproximación a la normal) 0,0649 (con aproximación a la normal)

5. a) 0,3230 b) 0,0214 c) 0,0025 d) 0

6. a) 0 (aproximado) b) 0,0024 (exacto) 0,0019 (aproximado por la normal)