



# DISTRIBUCIONES DE PROBABILIDAD

## **ÍNDICE TEMA V:**

### **5.1 Introducción**

### **5.2 Definiciones**

### **5.3 Distribución Binomial o de Bernoulli**

#### **5.3.1 Definición**

#### **5.3.2 Función de Probabilidad**

#### **5.3.3 Parámetros de una distribución Binomial**

#### **5.3.4 Uso de tablas**

### **5.4 Distribución Normal**

#### **5.4.1 Definición. Función de densidad. Parámetros**

#### **5.4.2 Distribución $N(0,1)$**

#### **5.4.3 Uso de tablas**

#### **5.4.4 Tipificación**

#### **5.4.5 Aproximación de una Binomial por Normal**

## 5.1 INTRODUCCIÓN

Gracias a las distribuciones de probabilidad podemos modelizar los datos obtenidos al realizar una encuesta o un trabajo estadístico y poder inferir. La probabilidad es el nexo de unión de la Estadística Descriptiva y la Estadística Inferencial.

## 5.2 DEFINICIONES

- Variable Aleatoria: es toda función real  $X$  que asocia a cada elemento del espacio muestral  $E$  un número real.

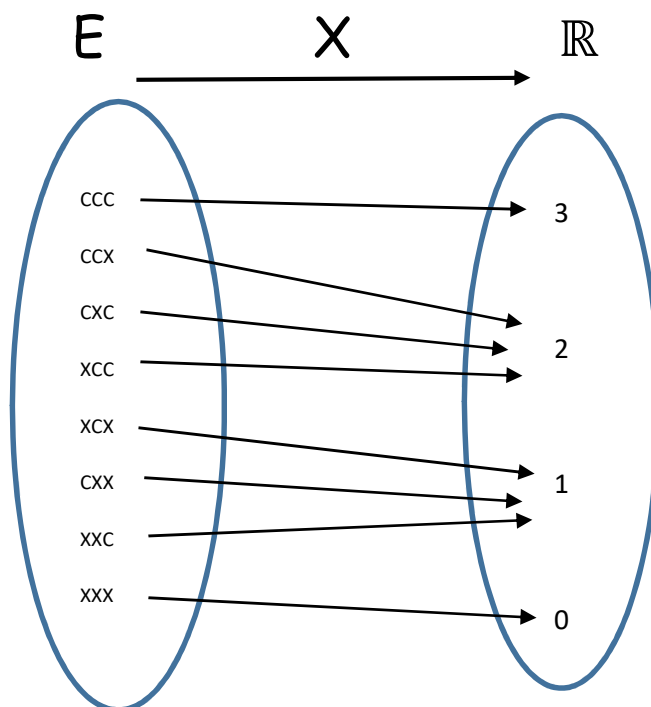


### EJEMPLO

Se considera el experimento aleatorio consistente en lanzar tres monedas diferentes al aire y apuntar su resultado.

El espacio muestral será  $E = \{XXX, CCC, CXC, XCC, CCX, XXC, XCX, CXX\}$

Definimos la función  $X$  que asigna a cada uno de los elementos del espacio muestral el número de caras obtenidas.



➤ Tipos de variables aleatorias:

1. Variable aleatoria discreta: una variable aleatoria es discreta si puede tomar solo valores aislados, en un número finito o infinito. El ejemplo anterior sería una v.a. discreta.

En el caso de v.a. discretas hablamos de función de probabilidad. La función de probabilidad asocia a cada valor de la variable aleatoria discreta, la probabilidad de que suceda.

2. Variable aleatoria continua: una variable aleatoria es continua si puede tomar cualquier valor dentro de uno o varios intervalos. Un ejemplo sería la variable que asocia a cada alumno de 2º de bachillerato su altura.

En el caso de v.a. continua hablamos de función de densidad. Una función  $f(x)$  es una función de densidad de una v.a. continua si cumple dos propiedades: la función es mayor que cero para todo el dominio de definición de la función ( $f(x) \geq 0$ ); y el área total limitada por la gráfica de la función y el eje de abscisas es 1.

## 5.3 Distribución Binomial

### 5.3.1 Definición

Una variable aleatoria  $X$  sigue una distribución Binomial si el experimento, asociado a dicha variable, cumple:

- En cada una de las pruebas solo son posibles dos resultados: el suceso  $A$ , al que llamaremos éxito, y su complementario  $A^c$ , al que llamaremos fracaso.
- El resultado obtenido en cada una de las pruebas es independiente de los obtenidos en las anteriores.
- La probabilidad del suceso  $A$  es constante y, por tanto, no varía de una prueba a otra.

Todo experimento que tenga estas características, decimos que sigue una distribución Binomial de parámetros  $n$  y  $p$ , siendo  $n$  el número de pruebas y  $p$  la probabilidad del suceso  $A$ .



### EJEMPLO

Una marca de bebidas, "COLA", ha determinado que la probabilidad de que una persona consuma su bebida es 0,35. Se escoge al azar una muestra formada por 7 personas. Comprobar si la variable que expresa el número de personas que consume la bebida COLA dentro de esa muestra sigue una distribución binomial e indica, en caso afirmativo los parámetros.

- En cada prueba solo hay dos posibles resultados: o consumes COLA o no la consumes.
- El resultado de cada prueba es independiente: si a una persona le gusta COLA, eso no afecta a que a la siguiente persona le guste o no COLA.
- La probabilidad de que le guste COLA es 0,35 y es constante para todas las personas.

Por todo ello, la variable aleatoria  $X$ = "número de personas que consume COLA" sigue una distribución Binomial de parámetros:  $n=7$ , número de personas que forman la muestra;  $p=0,35$ , probabilidad de consumir COLA.

Eso se indica matemáticamente:  $X \sim B(n, p)$ .

---

## 5.3.2 Función de Probabilidad

La función de probabilidad de una distribución Binomial,  $X \sim B(n, p)$ , viene dada por la siguiente expresión:

$$p[\text{obtener } r \text{ éxitos}] = p[X = r] = \binom{n}{r} \cdot p^r \cdot (1 - p)^{n-r}$$

## 5.3.3 Parámetros distribución Binomial

La media, varianza y desviación típica de una v.a. que sigue una distribución Binomial,  $X \sim B(n, p)$ , es:

$$\begin{aligned}\mu &= n \cdot p \\ \sigma^2 &= n \cdot p \cdot (1 - p) \\ \sigma &= \sqrt{n \cdot p \cdot (1 - p)}\end{aligned}$$

## 5.3.4 Uso de tablas

## 5.3.5 Ajuste de una binomial

Se trata de ajustar una distribución binomial a una distribución dada.



### EJEMPLO

De 500 unidades de un producto farmacéutico, tal que cada unidad lleva dos precintos, se han observado las siguientes frecuencias de precintos rotos.

Nº precintos rotos	0	1	2
Nº de unidades	190	220	90

Ajustar, si es posible, una distribución binomial y calcular las frecuencias absolutas teóricas.

1. Veamos si corresponde a una binomial. En cada prueba solo son posibles dos resultados: precinto roto o no roto. EL resultado obtenido en cada prueba es independiente de los anteriores. La probabilidad del suceso precinto roto, es constante y no varía de una prueba a otra. Por todo ello, la variable aleatoria  $X=$  “número de precintos rotos” sigue una distribución Binomial de parámetros:  $n=2$ , Habría que calcular  $p=$  probabilidad de precinto roto.
2. Calculamos la media a partir de la tabla anterior:

Nº precintos rotos	0	1	2
Nº de unidades	190	220	90
$x_i \cdot n_i$	0	220	180

$$\text{Por tanto, } \bar{x} = \frac{0+2 \cdot 180}{500} = 0,8$$

3. Sabemos que la media de una distribución binomial es  $\mu = n \cdot p$ . Despejando de la fórmula anterior, obtendremos  $p$ .

$$0,8 = 2p \Rightarrow p = \frac{0,8}{2} = 0.4$$

Eso se indica matemáticamente:  $X \sim B(2,0.4)$ .

4. Calcularíamos las frecuencias teóricas usando la binomial:

Nº precintos rotos	Frecuencia empírica	Probabilidad binomial	Frecuencia teórica
0	190	$\binom{2}{0} \cdot 0.4^0 \cdot 0.6^2 = 0.36$	$0.36 \cdot 500 = 180$
1	220	$\binom{2}{1} \cdot 0.4^1 \cdot 0.6^1 = 0.48$	$0.48 \cdot 500 = 240$
2	90	$\binom{2}{2} \cdot 0.4^2 \cdot 0.6^0 = 0.16$	$0.16 \cdot 500 = 80$

Ejercicios:

1. Quinientos opositores han participado en una prueba escrita que consta de tres ejercicios. Los resultados obtenidos son los que figuran en la tabla inferior.

Ajustar a esta distribución empírica una binomial y hallar las frecuencias teóricas esperadas.

Nº ejercicios aprobados	0	1	2	3
Nº opositores	136	223	120	21

2. El departamento de control de calidad de una fábrica de aparatos de televisión realiza cuatro controles. De 600 televisores se han obtenido los resultados que aparecen en la tabla.

Ajustar a esta distribución empírica una binomial y hallar las frecuencias teóricas esperadas.

Nº de fallos	0	1	2	3	4
Nº de TV	136	219	58	6	1

## 5.4 Distribución Normal

### 5.4.1 Definición. Función de densidad. Parámetros

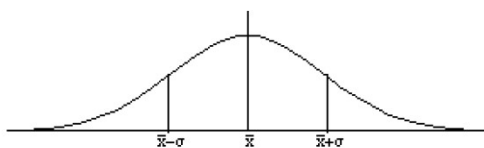
Una v.a. continua  $X$  sigue una distribución normal de media  $\mu$  y desviación típica  $\sigma$ , y se nota  $X \sim N(\mu, \sigma)$ , si se cumplen las siguientes condiciones:

- La variable recorre toda la recta real.
- La función de densidad viene expresada por la ecuación de la curva de Gauss:

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

Donde:

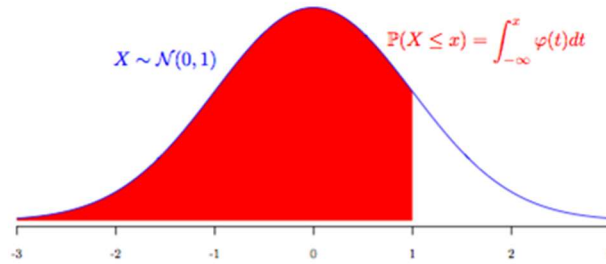
- $e = 2.71 \dots$
- $\pi = 3.1415 \dots$
- $x$  un punto del dominio de definición.
- $\mu$  es la media de la distribución.
- $\sigma$  es la desviación típica de la distribución.
- $f(x)$  es la función de densidad, que es: simétrica respecto a la recta  $x = \mu$ ; tiene una asíntota horizontal en el eje de abscisas, un máximo en  $x = \mu$  y el área encerrada bajo la curva es 1.



### 5.4.2 Distribución N (0,1).

Hay infinitas distribuciones Normales al ir variando los dos parámetros  $\mu$  y  $\sigma$ , pero la que presenta especial interés es aquella que tiene de media 0 y de desviación típica 1, es decir,  $X \sim N(0,1)$ ,





Para obtener la probabilidad de que la v.a. sea menor que un determinado valor, habría que realizar la integral de la función de densidad de la  $N(0,1)$  entre  $-\infty$  y el valor dado. Esto no es algo fácil, por lo que para poder calcular la probabilidad de que la v.a. sea menor que un determinado valor sin tener que usar el cálculo integral, se ha elaborado una tabla de fácil manejo.



### EJEMPLO

Una v.a. sigue una distribución normal de media 0 y desviación típica 1,  $X \sim N(0,1)$ , calcular la probabilidad de que  $X$  tome valores menores que 1,25.

La función de densidad sería

$$f(x) = \frac{1}{1 \cdot \sqrt{2\pi}} \cdot e^{-\frac{1}{2} \left(\frac{x-0}{1}\right)^2} = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{1}{2}x^2}$$

Y, por tanto, para calcular la probabilidad que nos piden, tendríamos que:

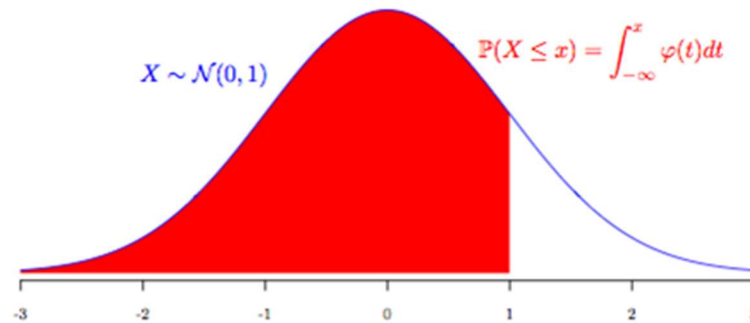
$$P[X \leq 1,25] = \int_{-\infty}^{1,25} \left( \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{1}{2}x^2} \right) dx = 0,8944$$

Por tanto, la probabilidad de que nuestra v.a. sea menor que 1,25 es 0,8944.

Aunque hemos puesto directamente el resultado de la integral, esto no es fácil de hacer. A continuación, mostramos cómo utilizar las tablas de la distribución  $N(0,1)$ .

### 5.4.3 Uso de tablas

La tabla anterior muestra las probabilidades de que una v.a.  $X$ , que sigue una normal de



	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990

parámetros 0 y 1, sea menor que un determinado valor.

Para calcular, en el caso anterior, la probabilidad de que la variable sea menor o igual que 1,25;  $P[X \leq 1,25]$ ; se procede de la siguiente forma:

1. Se busca en la primera columna 1,2.
2. Se busca en la primera fila 0,05.

3. El valor común a la fila y columna seleccionada, será la probabilidad pedida.

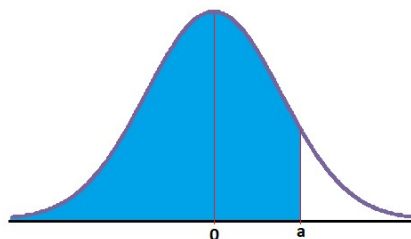
	z										
	-∞	0									+∞
z	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09	
0,0	0,5000	0,5040	0,5080	0,5120	0,5160	0,5199	0,5239	0,5279	0,5319	0,5359	
0,1	0,5398	0,5438	0,5478	0,5517	0,5557	0,5596	0,5636	0,5675	0,5714	0,5753	
0,2	0,5793	0,5832	0,5871	0,5910	0,5948	0,5987	0,6026	0,6064	0,6103	0,6141	
0,3	0,6179	0,6217	0,6255	0,6293	0,6331	0,6368	0,6406	0,6443	0,6480	0,6517	
0,4	0,6554	0,6591	0,6628	0,6664	0,6700	0,6736	0,6772	0,6808	0,6844	0,6879	
0,5	0,6915	0,6950	0,6985	0,7019	0,7054	0,7089	0,7123	0,7157	0,7190	0,7224	
0,6	0,7257	0,7291	0,7324	0,7357	0,7389	0,7421	0,7454	0,7486	0,7517	0,7549	
0,7	0,7580	0,7611	0,7642	0,7673	0,7704	0,7734	0,7764	0,7794	0,7823	0,7852	
0,8	0,7881	0,7910	0,7939	0,7967	0,7995	0,8023	0,8051	0,8078	0,8106	0,8133	
0,9	0,8159	0,8186	0,8212	0,8238	0,8264	0,8289	0,8315	0,8340	0,8365	0,8389	
1,0	0,8413	0,8438	0,8461	0,8485	0,8508	0,8531	0,8554	0,8577	0,8599	0,8621	
1,1	0,8643	0,8665	0,8686	0,8708	0,8729	0,8749	0,8770	0,8790	0,8810	0,8830	
1,2	0,8849	0,8869	0,8888	0,8906	0,8925	0,8944	0,8962	0,8980	0,8997	0,9015	
1,3	0,9032	0,9049	0,9066	0,9082	0,9099	0,9115	0,9131	0,9147	0,9162	0,9177	
1,4	0,9192	0,9207	0,9222	0,9236	0,9251	0,9265	0,9279	0,9292	0,9306	0,9319	
1,5	0,9332	0,9345	0,9357	0,9370	0,9382	0,9394	0,9406	0,9418	0,9429	0,9441	
1,6	0,9452	0,9463	0,9474	0,9484	0,9495	0,9505	0,9515	0,9525	0,9535	0,9545	
1,7	0,9554	0,9564	0,9573	0,9582	0,9591	0,9599	0,9608	0,9616	0,9625	0,9633	
1,8	0,9641	0,9649	0,9656	0,9664	0,9671	0,9678	0,9686	0,9693	0,9699	0,9706	
1,9	0,9713	0,9719	0,9726	0,9732	0,9738	0,9744	0,9750	0,9756	0,9761	0,9767	

En la tabla se pueden apreciar tres cosas fundamentales:

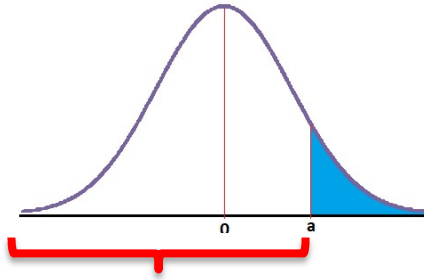
1. Que las probabilidades que nos dan son para valores de la variable menores que un cierto valor. No para otros casos.
2. Que las probabilidades van ordenadas (los valores de dentro de la tabla), algo lógico si se piensa en que el área a integrar será cada vez mayor a medida que aumenta el valor que consideremos.
3. También se aprecia que las probabilidades (valores de dentro de la tabla) son cada vez más próximos a 1. Recordar que, según definimos la función de densidad, el área total que encierra es 1.

Como hemos expresado en el punto 1, la tabla nos daría la probabilidad de ser menor o igual que un valor, pero no la probabilidad de otros casos, así que vamos a explicar cómo transformamos las otras probabilidades para poderlas buscar en la tabla:

1.  $P[X \leq a]$  en este caso se mira directamente la probabilidad como hemos visto en el punto anterior.



2.  $P[X \geq a]$ , para calcular la probabilidad de que la v.a. sea mayor o igual que un valor, hay que hacer unos de dos cosas: la simetría de la gráfica de la normal, y de que el área total que encierra la curva es 1.



Se observa que la probabilidad que nos piden es la complementaria de la que aparece en las tablas (llave roja). Por tanto,  $P[X \geq a] = 1 - P[X \leq a]$ .



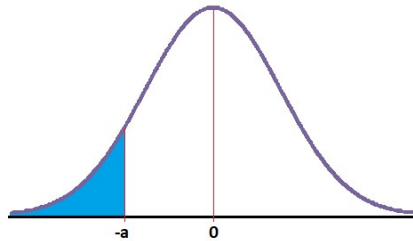
### EJEMPLO

Una v.a. sigue una distribución normal de media 0 y desviación típica 1,  $X \sim N(0,1)$ , calcular la probabilidad de que  $X$  tome valores mayores que 1,25.

$$P[X \geq 1,25] = 1 - P[X \leq 1,25] = 1 - 0,8944 = 0,1056$$

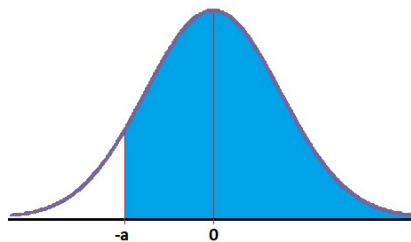
Por tanto, la probabilidad de que nuestra v.a. sea mayor que 1,25 es 0,1056.

3.  $P[X \leq -a]$  tal y como se aprecia en la gráfica, es igual, teniendo en cuenta la simetría, que la probabilidad de  $P[X \geq a]$



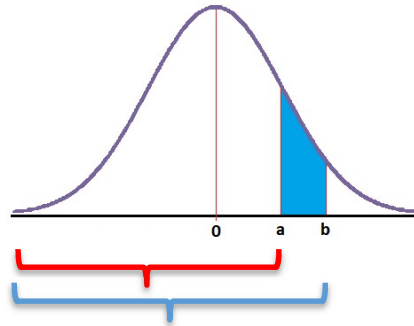
Por tanto,  $P[X \leq -a] = 1 - P[X \leq a]$ .

4.  $P[X \geq -a]$  tal y como se observa en la gráfica, esa área, es igual por simetría a la  $P[X \leq a]$ .



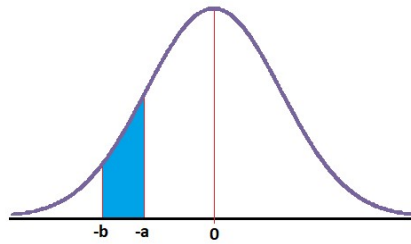
Por tanto,  $P[X \geq -a] = P[X \leq a]$ .

5.  $P[a \leq X \leq b]$  observando la gráfica, se puede ver que el área solicitada es igual a la diferencia entre el área que queda por debajo de  $b$  (llave azul) menos la que queda por debajo de  $a$  (llave roja).



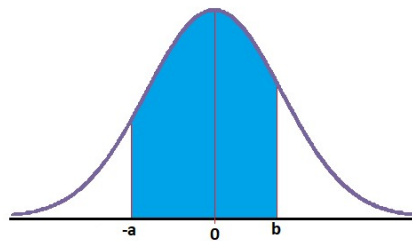
Por tanto,  $P[a \leq X \leq b] = P[X \leq b] - P[X \leq a]$ .

6.  $P[-b \leq X \leq -a]$  tal y como se aprecia en la gráfica, por simetría, el área es la misma que en el apartado anterior. (Recuerda que si  $a < b$  entonces  $-b < -a$ ).



Por tanto,  $P[-b \leq X \leq -a] = P[X \leq b] - P[X \leq a]$ .

7.  $P[-a \leq X \leq b]$  para obtener la probabilidad en este caso, aplicamos lo visto en el caso 5:  $P[-a \leq X \leq b] = P[X \leq b] - P[X \leq -a] = P[X \leq b] - (1 - P[X \leq a]) = P[X \leq b] - 1 + P[X \leq a]$



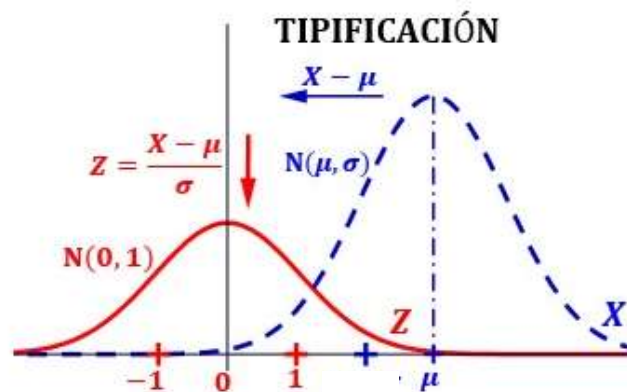
Por tanto,  $P[-a \leq X \leq b] = P[X \leq b] + P[X \leq a] - 1$ .

## 5.4.4 Tipificación

Hasta ahora hemos trabajado con la normal estándar, es decir, con la que tiene por media 1 y desviación típica 0, pero ¿cómo calculamos la probabilidad en el caso de tener otra distribución normal que no sea la estándar, ya que solo disponemos de tablas para ella?

Para ello hay que tipificar la variable. Tipificar una v.a. normal  $X \sim N(\mu, \sigma)$  consiste en transformarla en una nueva variable que siga una distribución normal estándar,  $Z \sim N(0,1)$ . Para ello se haría el cambio de variable  $Z = \frac{X-\mu}{\sigma}$ .

$$X \sim N(\mu, \sigma) \Rightarrow Z \sim N(0,1), \quad \text{siendo } Z = \frac{X - \mu}{\sigma}$$



En el gráfico superior, la distribución azul es la  $X \sim N(\mu, \sigma)$  y la roja es la  $Z \sim N(0,1)$ . Al restarle a todos los valores de la gráfica azul su media,  $\mu$ , se transforma en una normal de media 0, trasladamos la distribución. Y al dividir entre la desviación, estamos transformando su desviación en 1, lo que equivale a dilatar o contraer la gráfica.

### EJEMPLO



La estatura de un grupo de universitarios sigue una distribución normal de media 168cm y desviación típica 8cm. ¿Cuál es la probabilidad de que la estatura sea menor o igual que 170cm? Si hay 700 universitarios, ¿cuántos de ellos miden menos de 170cm? ¿Probabilidad de medir más de 165cm? ¿Probabilidad de medir entre 164 y 173?

**Primero definimos la variable:**

$X = \text{"estatura de universitarios"}$

$X \sim N(168,8)$

**Calculamos la probabilidad pedida:**

$$P[X \leq 170] = P\left[\frac{X - 168}{8} \leq \frac{170 - 168}{8}\right] = P[Z \leq 0,25] = 0,5987$$

Por tanto, la probabilidad de que los universitarios midan menos de 170cm es 0,5987.

**Calculamos cuántos miden menos de 170cm:**

Para ello, multiplicamos la probabilidad por el número de universitarios  $700 \cdot 0,5987 = 419,09 \cong 419$

419 universitarios miden menos de 170cm.

**Calculamos la probabilidad de medir más de 165:**

$$P[X \geq 165] = P\left[\frac{X - 168}{8} \geq \frac{165 - 168}{8}\right] = P[Z \geq -0,38] = (\text{caso 4}) = 0,6480$$

**Calculamos la probabilidad de medir entre 164 y 173:**

$$\begin{aligned} P[164 \leq X \leq 173] &= P\left[\frac{164 - 168}{8} \leq \frac{X - 168}{8} \leq \frac{173 - 168}{8}\right] = P[-0,5 \leq Z \leq 0,63] = \\ &= (\text{caso 7}) = P[X \leq 0,63] + P[X \leq 0,5] - 1 = 0,7357 + 0,6915 - 1 = \\ &= 0,4272 \end{aligned}$$

## 5.4.5 Aproximación de una Binomial por una Normal

Dados los