



### 1. ¿Qué es la estadística?

Es la ciencia que utiliza las matemáticas para conocer y predecir resultados. Alguna vez habrás oído hablar de las encuestas y/o estadísticas. Normalmente se usan para recibir información del comportamiento humano, así las encuestas pueden dar información sobre la intención de voto de una población, sobre el número de accidentes de tráfico en el fin de semana, sobre la edad media de los habitantes de una ciudad...

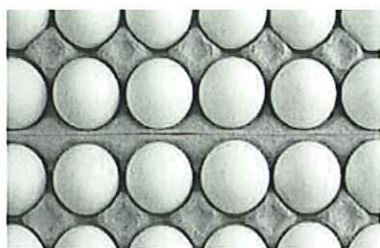
Pero ¡ojo! Hay que saber interpretar los resultados para no sacar conclusiones erróneas, veamos unos ejemplos:

- ✓ Un hombre tenía miedo de tomar un avión por aquello de los secuestros aéreos. Mirando unas estadísticas, encontró que la probabilidad de que hubiese una bomba en su vuelo era de 1 entre 1.000, mientras que la probabilidad de que hubiese dos era 1 entre 100.000. Por lo tanto, tomó el avión llevando él mismo una bomba.



- ✓ En realidad, volar en avión es muy seguro. Prácticamente la totalidad de los fallecidos en accidentes aéreos han muerto al llegar al suelo.
- ✓ En Nueva York un hombre es atropellado cada diez minutos. El pobre tiene que estar hecho polvo.
- ✓ La probabilidad de tener un accidente de tráfico aumenta con el tiempo que te pases en la calle. Por tanto, cuanto más rápido circules, menor es la probabilidad de que tengas un accidente.

Así pues, para aprender bien a manejar datos estadísticos y saber interpretar la información correctamente vamos a aclarar unos conceptos:



• **Población:** conjunto de elementos que queremos conocer.

• **Muestra:** parte de la población que vamos a analizar. En las encuestas la muestra son las personas que han contestado a ella. La muestra debe ser representativa, es decir,

debe haber "un poco de todo". Se representa con la letra **N**.



A las características de la muestra les llamamos **variables** y pueden ser de 2 tipos:

1. **Cualitativos:** son los que **no se pueden medir** (el color de ojos, del pelo...)
2. **Cuantitativos:** son aquellos que **sí podemos medir**, o sea, números (la edad, la altura, la cilindrada de una moto...) Podemos distinguir 2 tipos:
  - ⇒ **Variables discretas:** son siempre números enteros, por ejemplo el número de hijos: 1,2,3..., que nunca podrán ser decimales (0'5, 7'3...).
  - ⇒ **Variables continuas:** son números que pueden ser enteros o decimales: las cantidades, el peso..., así se puede ordenar los pesos de varias personas entre 68 y 69 Kg (68.1, 68.2, 68'3... 68.9).

## 2. ¿Cómo organizamos los datos?

**Media:** refleja el punto medio o más cercano a todos los datos. Es la suma de todos los datos de la variable, dividida por el número de datos. Se expresa con este símbolo  $\bar{x}$ .



$$\bar{x} = \frac{\text{Suma de todos los datos}}{N}$$

$$\bar{x} = \frac{x_1 + x_2 + x_3 + \dots + x_n}{N}$$

Ejemplos:

1. Los pesos de seis amigos son: 84, 91, 72, 68, 87 y 78 kg. Hallar el peso medio.

$$\bar{x} = \frac{84 + 91 + 72 + 68 + 87 + 78}{6} = 80\text{kg}$$

2. En un test realizado a un grupo de 42 personas se han obtenido las puntuaciones que muestra la tabla. **Calcula la puntuación media.**





$X_i$	$f_i$	$X_i \cdot f_i$
15	1	15
25	8	200
35	10	350
45	9	405
55	8	440
65	4	260
75	2	150
	42	1 820

$$\bar{x} = \frac{1820}{42} = 43,33$$

Donde  $X_i$  son las puntuaciones obtenidas en el test y  $f_i$  es la **frecuencia**, es decir el número de personas que han recibido cada una de las puntuaciones.

**Moda:** refleja el dato que más se repite. Solo tendremos que mirar en la frecuencia y ver qué dato se repite más. La moda es el valor que tiene mayor frecuencia absoluta. Se representa por **Mo**. Se puede hallar la moda para variables cualitativas y cuantitativas.

Hallar la **moda** de la distribución:

$$2, 3, 3, 4, 4, 4, 5, 5$$

**Mo = 4**

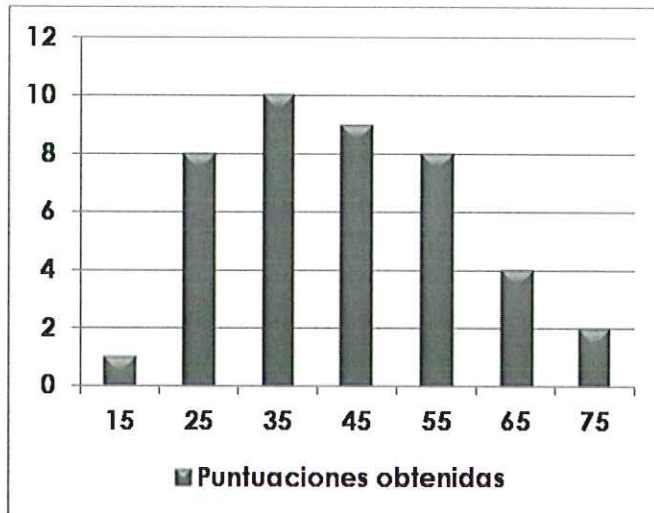
Si **dos puntuaciones adyacentes** tienen la **frecuencia máxima**, la **moda** es el **promedio** de las dos puntuaciones adyacentes.

$$0, 1, 3, 3, 5, 5, 7, 8 \quad Mo = \frac{3+3+5+5}{4} = 4 \quad Mo = 4$$



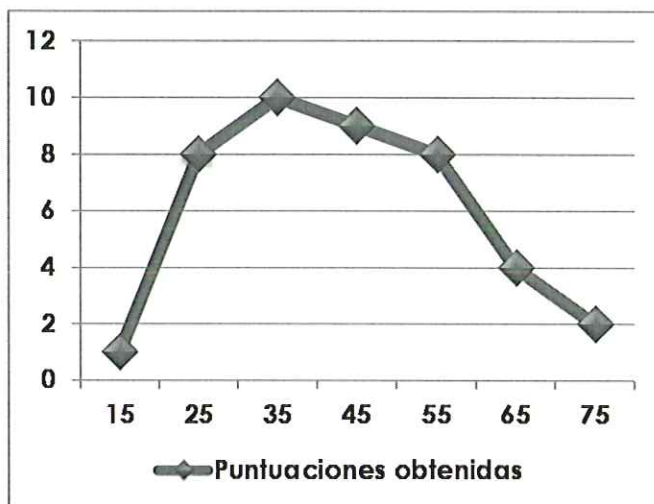
### 3. ¿Cómo representamos gráficamente los datos?

Normalmente cuando vemos informaciones estadísticas es a través de gráficos y no de tablas. En este punto vamos a ver dos tipos. Representan los datos del ejemplo anterior:



#### DIAGRAMA DE BARRAS

El diagrama de barras consiste en dibujar segmentos verticales cuyo pie es el correspondiente al valor de la variable, y cuya altura es la frecuencia de dicho valor.



#### POLÍGONOS DE FRECUENCIAS

Un polígono de frecuencias se forma trazando los puntos que representan las frecuencias y uniéndolos. Observa como aumenta o descende la frecuencia a través de una línea.

Cuando utilizamos variables continuas (2'5, 3'5, 4'5...) puede ocurrir que la variable tiene muchos valores. Entonces es mejor que los agrupemos en intervalos. Un **intervalo** es la acumulación de varios datos entre dos valores.

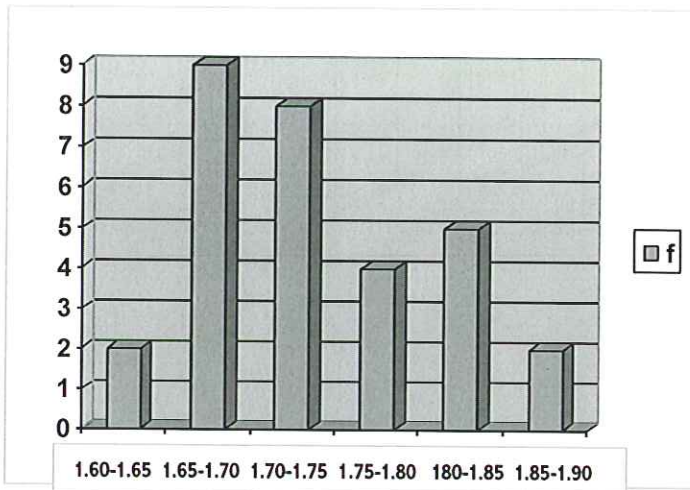




Para agrupar los datos tendremos en cuenta el rango (o recorrido de la variable), que es la diferencia entre el valor máximo y el valor mínimo. Por ejemplo, si seleccionamos las alturas de 30 alumnos, tendríamos que las alturas de todos se distribuyen entre 1,60-1,87 cm. El rango se obtiene restando los dos valores ( $1,87 - 1,60 = 0,27$ ) será entonces 0,27. Parece lógico tomar intervalos desde 1,60 a 1,90, de 5 en 5 centímetros:

## ALTURA ALUMNOS

1,75 1,72 1,67 1,69 1,80  
 1,65 1,79 1,68 1,68 1,70  
 1,63 1,69 1,71 1,85 1,83  
 1,77 1,71 1,74 1,83 1,60  
 1,87 1,74 1,72 1,76 1,69  
 1,81 1,68 1,73 1,69 1,80



1.- Hemos realizado un estudio sobre el número de hijos que tiene una familia. Para ello se tomó como muestra un total de 50 familias, obteniéndose el resultado que aparece en la tabla.

a) El número de hijos, ¿qué tipo de variable es?  
 ¿Por qué?

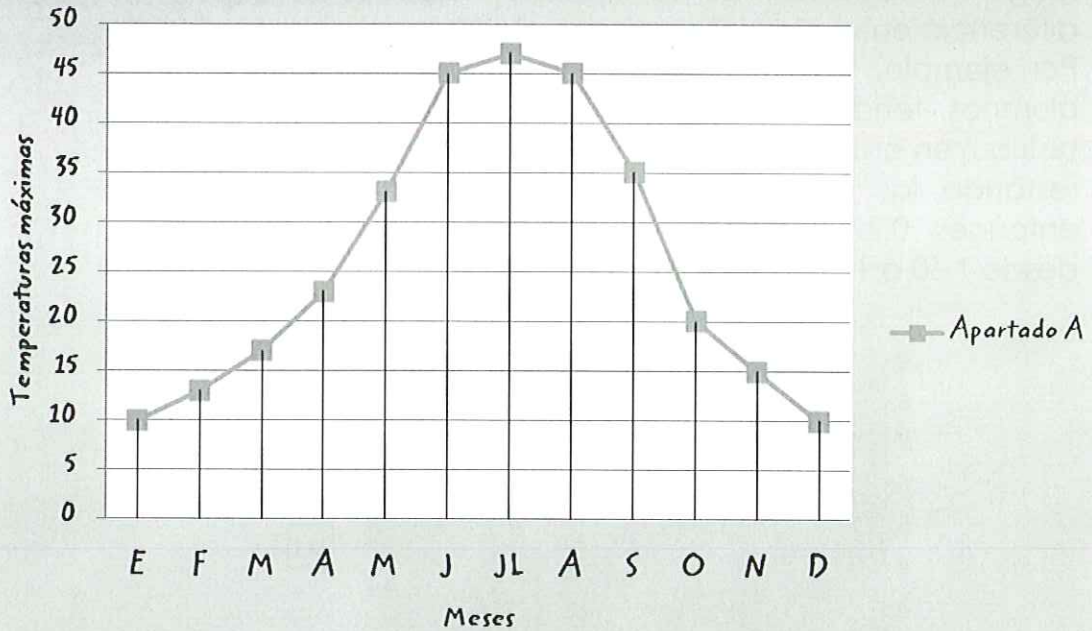
b) ¿Qué tanto por ciento de las 50 familias tienen dos hijos?

c) ¿Cuál es la media? ¿Y la moda?

x (nº hijos)	f
0	9
1	12
2	18
3	6
4	3
5	2



2.- El siguiente gráfico representa las temperaturas máximas mensuales registradas durante un año en una determinada ciudad.



a) Confecciona una tabla con los valores de la gráfica anterior.



Meses												
Temperaturas Máximas												

b) ¿Cuál es la temperatura máxima?

c) Representa en la gráfica anterior con boli azul, la evolución de las temperaturas mínimas registradas en esa misma ciudad.



Meses	E	F	M	A	M	J	JL	A	S	O	N	D
Temperaturas Máximas	1	2	8	10	13	21	23	23	19	16	9	4

3. Un bebe al nacer pesó 3500 gramos. Si gana peso a razón de 40 g semanales, completa la siguiente tabla:

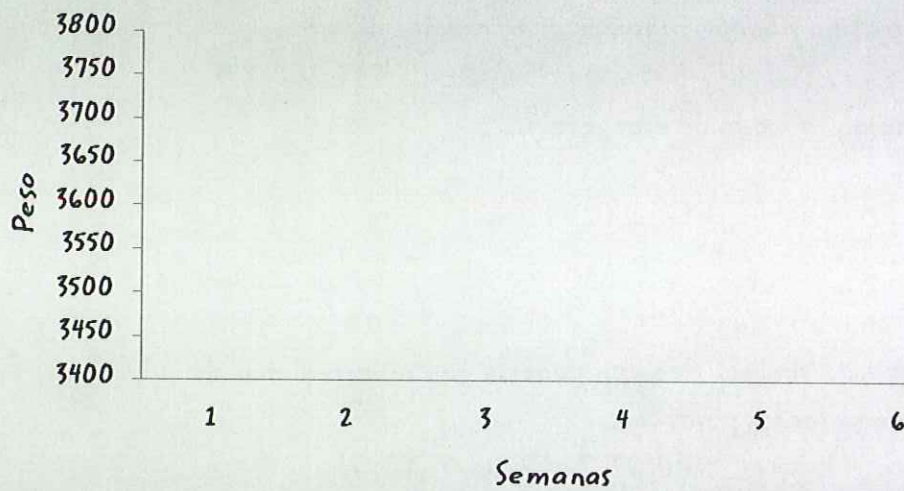


Semana	0	1	2	3	4	5	6
Peso	3500						



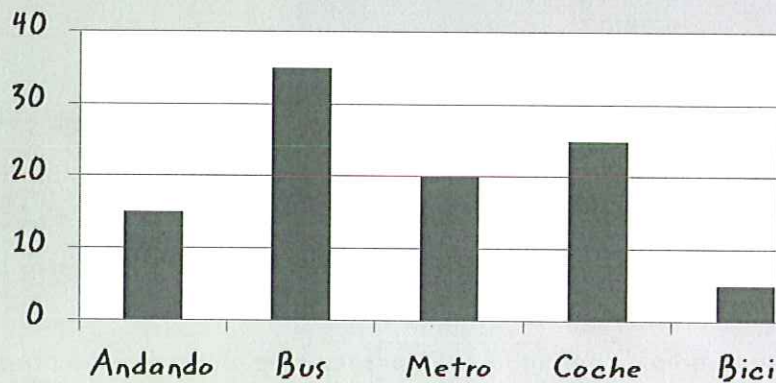


a) Representa graficamente los datos de la tabla anterior:



b) Obtén la fórmula que da el peso del bebé en función del número de semanas de vida.

4. Se ha preguntado a todos los empleados de una empresa el medio de transporte que utilizan asiduamente para ir a trabajar. El resultado es este:



a) ¿Qué nombre recibe el diagrama anterior?

b) ¿Cuántos empleados de la empresa van en bus?

c) ¿Cuántos empleados tiene la empresa?

d) ¿Cuál es el porcentaje de los que utilizan el metro como medio de transporte para ir a trabajar?



5. Hemos consultado en diferentes comercios el precio de una determinada cámara fotográfica y hemos obtenido estos precios:

260-280-305-295-295-280-285-285-280-275

Calcula la media y la moda de esos precios.



6. Un grupo de amigas deciden ponerse juntas a régimen de adelgazar. En el momento de empezar sus pesos son:

80, 77, 76, 62, 68, 71, 67, 71

a. ¿Cuál es el peso medio de las amigas?

b. ¿Y la moda?



7. Cuatro amigos juegan a la lotería todas las semanas. Durante 8 semanas estos han sido los resultados:

-3.000, -3.000, +10.000, +5.000, -6.000, -1.000, +5.000, +3.000

Teniendo en cuenta que los valores negativos representan pérdidas y los positivos ganancias,

a) Calcula el promedio de ganancias entre los cuatro en estas ocho semanas.

b) ¿Y cada uno?



## ESTADÍSTICA

**Medidas de Centralización:** nos sirven para ver la tendencia más común de una población. Hay tres medidas.

1ª- **Media:** la media de una variable estadística es el cociente obtenido al dividir la suma de todos sus valores por el número de datos.

$$\bar{x} = \frac{\sum_{i=1}^N x_i \cdot f_i}{N}$$

$x_i$  ---- valores de la variable x  
 $f_i$  ---- frecuencia (número de veces que se repite  $x_i$ )  
N ---- número de datos en total. =  $\sum_{i=1}^N f_i$

2ª- **Moda:** es el valor que aparece con más frecuencia.

3ª- **Mediana:** es el dato que deja tantos valores por encima como por debajo de él. Es el dato cuya frecuencia supera a la mitad de las frecuencias acumuladas.

**Medidas de Dispersión:** estas medidas nos sirven para medir lo agrupados que están los datos. Hay cuatro medidas.

1ª- **Recorrido o Rango:** es la diferencia entre el dato mayor y el menor de la variable.

2ª- **Desviación Media:** es la media aritmética de las desviaciones en valor absoluto; siendo las desviaciones con respecto a la media las diferencias entre cada valor de la variable y su media ( $x_i - \bar{x}$ ).

$$D_m = \frac{\sum_{i=1}^N |x_i - \bar{x}|}{N}$$

N ---- número de datos totales.

3ª- **Varianza:** es la media aritmética de los cuadrados de las desviaciones de todos los datos respecto a la media. Cuanto mayor es la varianza más dispersos estarán los datos.

$$S^2 = \frac{\sum_{i=1}^N f_i \cdot (x_i - \bar{x})^2}{N} = \frac{\sum_{i=1}^N f_i \cdot (x_i - \bar{x})^2}{N}$$

4ª- **Desviación Típica:** es la raíz cuadrada positiva de la varianza.  $\rightarrow S = +\sqrt{S^2}$

5ª- Coeeficiente de Variación de Pearson  $\rightarrow CV = \frac{S}{\bar{x}}$   
Compara el nivel de dispersión de 2 muestras.

## EJERCICIOS DE ESTADÍSTICA

### Ejercicios de Medidas de Centralización

1º- Supongamos que hay dos alumnos que han sacado las siguientes notas de matemáticas a lo largo del curso. Calcula la nota media de cada uno de ellos.

Alumno A: 3,1,4,5,7,4.

Alumno B: 3,4,6,6,7,4.

2º- En una clínica han nacido 150 bebés con los pesos siguientes en kilos:

[2-2.5)----- 21 bebés

[2.5-3)----- 22 “

[3-3.5)----- 33 “

[3.5-4)----- 31 “

[4-4.5)----- 35 “

[4.5-5)----- 8 “

Halla el peso medio de esos 150 bebés.

3º- En una biblioteca se pregunta qué tipo de libros se han leído en la última semana, obtenemos los siguientes datos:

Policíaco----- 10

Aventura----- 11

Científico----- 12

Novela----- 25

Histórico----- 6

¿Cuál es la moda?.

4º- Se hace un estudio en un barrio sobre el precio del alquiler de 110 viviendas obteniéndose la siguiente tabla:

Precio del alquiler (ptas.)	Número de viviendas
-----------------------------	---------------------

(50.000-60.000) -----	25
-----------------------	----

(60.000-70.000) -----	28
-----------------------	----

(70.000-80.000) -----	38
-----------------------	----

(80.000-90.000) -----	13
-----------------------	----

(90.000-100.000) -----	6
------------------------	---

¿Cuál es la moda?.

5º- En un test aplicado a 9 niños se obtienen los resultados:

3,4,5,2,0,1,7,9,8. ¿Cuál es la moda?.



6º- Al lanzar un dado 70 veces obtenemos la tabla siguiente:

Resultados	1	2	3	4	5	6
Frecuencias	12	10	19	0	15	14
F.Acumuladas	12	22	41	41	56	70

¿ cuál es la mediana?.

7º- En un observatorio meteorológico de una ciudad tenemos las siguientes temperaturas máximas en el mes de diciembre:

12,3,4,10,11,12,4,5,5,6,8,2,5,6,7,7,8,10,12,13,12,6,9,9,6,7,8,10,10,8,8.

Halla la media, la moda y la mediana.

8º- Tomamos 50 centros de trabajo y vemos los accidentes laborales en un determinado año.

Nº de accidentes	Frecuencia
[0 2.000)	5
[2.000 4.000)	5
[4.000 6.000)	7
[6.000 8.000)	7
[8.000 10.000)	6
[10.000 12.000)	4
[12.000 14.000)	2
[14.000 16.000)	8
[16.000 18.000)	3
[18.000 20.000)	3

Calcula la media, la moda y la mediana.

9º- Pesamos a 30 niños y obtenemos la siguiente tabla:

Peso de los niños	Marca de clase	Frecuencia
[10.4, 13.4)	11.9	7
[13.4, 16.4)	14.9	3
[16.4, 19.4)	17.9	8
[19.4, 22.4)	20.9	10
[22.4, 25.4)	23.9	2

Halla la moda, la media y la mediana.

## Ejercicios de Medidas de Dispersión

1º- Si preguntamos a 10 amas de casa cuánto dura una pastilla de jabón de una marca determinada, tenemos las siguientes respuestas en días : 5,6,9,10,6,7,11,8,5,6.

¿Cuál es el recorrido?

2º- Consideramos las medidas de 10 jugadores de baloncesto que son las siguientes: 1,68; 1,70; 1,82; 1,63; 1,71; 1,83; 1,72; 1,80; 1,84; 1,67.

¿Cuál es el recorrido?

3º- Halla la desviación media de las edades de los componentes de un grupo de teatro:

8, 10, 12, 14, 16, 18, 20, 22.

4º- Los jornales de 5 obreros son:

5.000, 6.000, 5.100, 5.300, 5.900.

Halla la desviación media.

5º- Lanzado un dado 50 veces se ha obtenido la siguiente distribución:

$X_i$	1	2	3	4	5	6
$F_i$	6	11	6	7	9	11

Halla la varianza y la desviación típica.

6º- Halla la media aritmética y la desviación típica de las notas obtenidas en un examen de 150 niños que tuvieron las puntuaciones siguientes:

Puntuación	1	2	3	4	5	6	7	8	9	10
Nº alumno	3	4	8	10	28	42	32	16	4	3

7º- Calcula la desviación típica de las alturas de los chicos de una escuela cuya distribución es la que sigue:

Altura	160	161	162	163	164	165	166	167	168	169
nºchicos	1	2	9	11	15	18	10	8	5	3

8º- **(Examen)** Un portero ha encajado en los últimos seis partidos el siguiente número de goles: 1, 0, 1, 3, 0, 1.

Averiguar la media y la desviación típica de la muestra.

9º- **(Examen)**. Las ventas de una fábrica de aceite medidas en litros han sido en los últimos 5 años: 250; 350; 375; 400; 425.

Hallar la media y la desviación típica de la muestra.



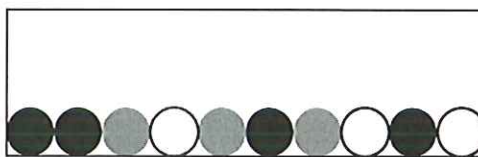


### 4. ¿Qué es la probabilidad?

Lanza un dado 50 veces y fíjate bien en las veces que sale cada cara. Esto nos llevaría a sacar la siguiente conclusión: todos los números del dado al azar tienen las mismas posibilidades de salir. Se dice que tienen la misma **probabilidad** de ocurrir.



Imagínate que tenemos en esta urna 10 bolas del mismo tamaño pero de distintos colores. Realizamos el experimento de sacar una bola al azar (sin mirar):



Para calcular la probabilidad al azar tenemos la **regla de Laplace**.



$$P(a) = \frac{\text{N}^\circ \text{ de casos favorables}}{\text{N}^\circ \text{ de casos posibles}}$$

Si queremos saber la probabilidad del suceso "**sacar bola negra**", al número de bolas negras que hay en la urna se le llama "**número de casos favorables**" (favorables al suceso), y al número total de bolas que hay en la bolsa se le llama "**número de casos posibles**". Por tanto la probabilidad de sacar bola negra en la urna anterior será:

$$P(\text{bola negra}) = \frac{4}{10} = 0,4$$

Al poner la fórmula, siempre ponemos ente paréntesis la probabilidad de lo que tenemos que hallar.



**RECUERDA QUE...**  
La probabilidad de cualquier cosa siempre será un número comprendido entre el 0 y el 1, donde 0 significa que no ocurrirá nunca y 1 que ocurrirá siempre.



Si en un problema nos dicen con o sin reemplazamiento, cambiará el número de casos posibles:

- **con reemplazamiento** el total será siempre el mismo
- **sin reemplazamiento** habrá un caso posible menos

### ¿Cómo se hacen este tipo de problemas?

*Ejemplo: Se extraen sucesivamente 3 esferas de una caja que contiene 6 esferas rojas, 4 blancas y 5 azules. Hallar la probabilidad de que sean extraídas en el orden roja, blanca y azul, Si las extracciones son:*

#### **a) con reemplazamiento:**

Tenemos 15 bolas de las cuales 6 son rojas, por lo que la probabilidad de que la bola sea roja es:  $\frac{6}{15} = 0,4$

Volvemos a meter esta bola. Ahora, para sacar la 2ª bola disponemos otra vez de 15 bolas de las cuales 4 son blancas por lo que la probabilidad de que la segunda bola sea blanca es:  $\frac{4}{15} = 0,2\bar{6}$

Se vuelve a meter la bola. A la hora de sacar la tercera bola, volvemos a disponer de 15 bolas de las cuales 5 son azules, por lo que la probabilidad de que la tercera bola sea azul es:  $\frac{5}{15} = 0,3\bar{3}$

Por lo tanto la probabilidad que las bolas sean extraídas en el orden roja, blanca y azul es:

$$\left(\frac{6}{15}\right) \cdot \left(\frac{4}{15}\right) \cdot \left(\frac{5}{15}\right) = (0,4) \cdot (0,2\bar{6}) \cdot (0,3\bar{3}) = 0,031$$

#### **b) sin reemplazamiento:**

Tenemos 15 bolas de las cuales 6 son rojas, por lo que la probabilidad de que la primera bola que saque sea roja es:  $\frac{6}{15} = 0,4$

A la hora de sacar la segunda bola disponemos de 14 bolas de las cuales 4 son blancas, por lo que la probabilidad de que la segunda bola que saque sea blanca es:  $\frac{4}{14} = 0,28$

Por último, cuando voy a sacar la tercera bola disponemos de 13 bolas de las cuales 5 son azules, por lo que la probabilidad de que la tercera bola que saque sea azul es:  $\frac{5}{13} = 0,38$

Por lo tanto la probabilidad que las bolas sean extraídas en el orden roja, blanca y azul es:

$$\left(\frac{6}{15}\right) \cdot \left(\frac{4}{14}\right) \cdot \left(\frac{5}{13}\right) = (0,4) \cdot (0,28) \cdot (0,38) = 0,042$$





8. La siguiente tabla da el número de alumnos que han acabado sus estudios en la Universidad de Valencia, durante el curso 96/97:

	Mujeres	Hombres	Total
Arquitectura	109	234	343
Informática	84	231	315
Industriales	102	273	375
Agrícola	106	149	255
Totales	401		

a) Si seleccionamos uno al azar del total del alumnado, calcula la probabilidad de que sea mujer.

b) Si seleccionamos uno al azar del total del alumnado de informática, calcula la probabilidad de que sea mujer. ¿y de que sea hombre?

9. ¿Cuál es la probabilidad de que ocurran los siguientes sucesos, al lanzar un dado?

a) Salir el número 3.

b) Salir un número par.

c) Salir un número mayor que 1.

d) Salir el número 8.

e) Salir un número menor que 5.



10. Hemos consultado el precio (en euros) de un determinado lector de DVD en 8 establecimientos diferentes. Los datos obtenidos son: 146, 152, 141, 141, 148, 141, 149, 158

Calcula:

a) La media.

b) La moda.

c) Haz una tabla de frecuencias y dibuja un diagrama de barras.



11. En una bolsa tenemos 5 bolas blancas, 3 bolas rojas y 2 azules. Extraemos 2 bolas,

a) ¿Cuál es la probabilidad de que las dos sean rojas, si después de ver la primera bola se vuelve a introducir en la bolsa? (con reemplazamiento).

b) ¿Cuál es la probabilidad de que las dos sean rojas, si después de ver la primera bola no se vuelve a introducir en la bolsa? (sin reemplazamiento)





12. La siguiente tabla muestra la incidencia de lectura de periódicos en hombres y mujeres. La muestra ha sido tomada a 500 personas.

a) Completa esta tabla:

	Hombres	Mujeres	Total
Leen			203
No leen		175	
Total	241		500

a) Calcula la probabilidad de que una persona elegida sea mujer y lectora.



13. En una clase hay 15 chicas y 10 chicos. De todos ellos, 16 están en el taller de matemáticas, y se sabe que hay 4 chicos que no están en dicho taller. De acuerdo con esta información, completa la tabla:

	Taller	No taller	Total
Chicas			
Chicos			
TOTAL			



14. El número de hijos de 10 familias es el siguiente:

5, 2, 0, 2, 3, 1, 2, 3, 1, 4

a) ¿Cuál es el número medio de hijos?

b) Calcula la moda de esta muestra.

