

La Estadística es la parte de las Matemáticas que estudia numéricamente diferentes características de una población, para extraer conclusiones (Estadística Descriptiva) o para hacer previsiones (Inferencia Estadística)

Aquella característica que es motivo de un estudio estadístico se llama variable estadística, y puede ser de los siguientes tipos:

- Cualitativa (expresa cualidades): color del cabello, llevar gafas, sexo, programas de T.V., ...
- Cuantitativa (expresa cantidades numéricas):
 - Discreta (valores aislados que pueden ser enteros) : habitantes, coches/habitante, hermanos, ...
 - Continua (toma cualquiera de los valores de un intervalo): altura, peso, temperatura, ...

1. Di, en cada caso, cuál la variable que se quiere estudiar y especifica de qué tipo es.
 - a) Tiempo dedicado a las tareas domésticas por parte de los hombres y las mujeres cuando trabajan fuera de casa.
 - b) Estudios que quiere hacer el alumnado de un centro escolar en acabar la E.S.O.
 - c) Intención de voto en unas elecciones autonómicas.
 - d) Horas que dedican a ver la televisión los estudiantes de E.S.O.
 - e) Número de aparatos de radio que hay en las casas españolas.

Población es el conjunto de individuos o elementos a quien va dirigido el estudio estadístico. Cuando por cuestiones de tiempo o económicas no es posible realizar el estudio a toda la población, lo que se toma es una Muestra representativa de esa población.

2. Diferencia entre la población y la muestra, indica cuál es la variable y clasifícala.
 - a) Se observa a 50 alumnos de un instituto para analizar su color de ojos.
 - b) Preguntamos a todos los alumnos de un curso cuántos hermanos tienen.
 - c) Se encuesta a 1500 alumnos de colegios de Alicante sobre el tiempo que tardan en llegar al colegio.

En un proceso estadístico se han de distinguir los siguientes pasos:

- Qué se quiere estudiar
- Seleccionar las variables correspondientes
- Recoger los datos
- Organizar la información en tablas de frecuencias
- Dibujar los gráficos adecuados
- Cálculo de parámetros estadísticos
- Interpretación de los resultados

NOTES OBTINGUDES PER UN GRUP D'ALUMNES	RECOMpte	TAULA DE FREQUÈNCIES	
9 4 8 5 5 4 1	1	x_i	f_i
7 2 2 3 9 6 4	2	1	2
10 8 2 1 6 7 6	3	2	4
10 10 8 8 4 6 5	4	3	3
5 10 6 7 2 5 5	5	4	4
3 5 3 6 8	6	5	7
	7	6	6
	8	7	3
	9	8	5
	10	9	2
		10	4

En una tabla de frecuencias se acostumbra a reflejar toda la información tomada:

- los valores (x_i) de la variable
- la frecuencia absoluta (f_i), que es el número de veces que se repite cada valor
- la frecuencia relativa (f_{r_i}), es la frecuencia absoluta dividida por el número total de valores (N)
- el porcentaje (%) de veces que se repite cada valor
- las columnas de cálculos para parámetros estadísticos ($x_i \cdot f_i$), ($x_i^2 \cdot f_i$), ($|x_i - \bar{x}| \cdot f_i$), ...

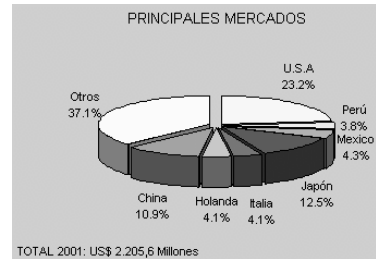
x_i	f_i	$(x_i) \cdot (f_i \cdot x_i)$ $f_i \cdot x_i$	$f_i \cdot x_i^2$
4	1	4	16
5	10	50	250
6	14	84	504
7	5	35	245
8	2	16	128
9	1	9	81
	Σf_i	$\Sigma f_i \cdot x_i$	$\Sigma f_i \cdot x_i^2$
	33	198	1224

Hay muchos tipos de gráficos estadísticos, aunque veremos con más profundidad:

- Diagrama de sectores: se utiliza, preferentemente en variables cualitativas.

Tendremos que dividir todo un círculo (360°) en sectores que sean proporcionales a la frecuencia de cada valor.

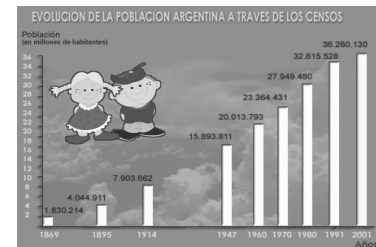
Dentro de cada sector pondremos el nombre de cada valor de la variable, la frecuencia absoluta correspondiente o el porcentaje asociado.



- Diagrama de barras: se utiliza, sobretodo, en variables cuantitativas discretas, nunca continuas.

Las barras han de ser igual de anchas y separadas entre sí, han de estar centradas sobre cada valor.

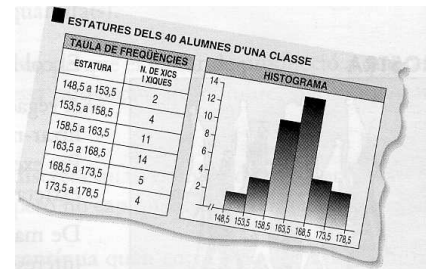
En el eje de abscisas (horizontal) se representan los valores de la variable, mientras que en el eje de ordenadas (vertical) se representan sus frecuencias absolutas o los porcentajes correspondientes.



- Histograma: se utiliza en variables cuantitativas continuas o agrupadas en intervalos.

Las barras son igual de anchas, pero contiguas entre sí.

Las unidades, extremos de los intervalos, están en los extremos de cada barra.



3. Los sesenta estudiantes de 3º de E.S.O. de una localidad toman en su desayuno los siguientes alimentos:

Clasifica la variable motivo de estudio

Representa los datos de la tabla en un diagrama de sectores.

Desayuno	Estud.
Tostada	10
Pan	7
Galletas	24
Cereales	8
Pastel	11

4. En la tabla siguiente se encuentra la distribución de frecuencias del número de asignaturas que han preparado en sus casas un día cualquiera los 30 estudiantes de una clase:

a) ¿De qué tipo de variable se trata?

b) Construye un diagrama de barras para representar los datos contenidos en la tabla.

c) ¿Qué porcentaje de estudiantes preparan exactamente cuatro asignaturas?

d) Obtén el porcentaje de estudiantes que preparan menos de tres asignaturas.

Asign.	f _i
0	1
1	3
2	5
3	10
4	6
5	3
6	2

5. La tabla de la derecha recoge el número de horas semanales que ven la televisión un total de 90 estudiantes.
- ¿De qué tipo es la variable del estudio?
 - Representa gráficamente los datos mediante un histograma.
 - ¿Cuál es la clase más frecuente?
 - Calcula el porcentaje de estudiantes que estudian entre 3 y 9 horas semanales.

Horas	frec.
[0,3[3
[3,6[9
[6,9[20
[9,12[31
[12,15[25
[15,18[2

6. Al contar el número de áreas suspendidas por cada alumno en la primera evaluación de un grupo de 3º de E.S.O., se han obtenido los siguientes datos:

1	1	2	3	2	4	5	0	0	0
1	1	0	1	2	3	2	1	3	1
6	0	0	1	0	0	0	5	4	2

- Haz una tabla de frecuencias absolutas y el diagrama de barras correspondiente.
- Calcula los porcentajes asociados con los siguientes grupos:
 - Estudiantes que no han suspendido ninguna área.
 - Estudiantes que han suspendido una o dos áreas.
 - Estudiantes que han suspendido tres o más áreas.
 - Haz un diagrama de sectores que refleje los porcentajes de estos tres grupos.

Parámetros Centrales son los valores que sirven para representar a todos los de la variable:

- Media aritmética (\bar{x}): se calcula sumando todos los valores de la variable y dividiendo entre el número de valores que haya. $\bar{x} = \frac{\sum x_i}{N}$, si cada valor x_i se repite f_i veces $\rightarrow \bar{x} = \frac{\sum x_i \cdot f_i}{N}$
- Mediana (Me): es el valor central, una vez están ordenados todos los valores de la variable.
- Moda (Mo): es aquel valor que más veces se repite, el que tiene más frecuencia absoluta.

7. El número de coches que pasaron por una carretera local durante las horas de luz de un día determinado fueron: 12, 13, 16, 8, 17, 11, 19, 15, 6, 10, 14, 15, 8, 11
- Calcula la media y la mediana del número de coches.

8. Las visitas de los niños de un barrio a su médico durante el último año fueron las siguientes:
- Calcula la media y la mediana.
 - Sin hacer cálculos, razona cuál de los dos parámetros anteriores se vería más afectado si las visitas del último niño hubieran sido 10 en lugar de 37.

Visitas	Niños
0	2
1	6
2	4
3	3
4	3
5	2
37	1

Parámetros de dispersión son los valores que sirven para medir el grado de dispersión de los datos (si están más o menos separados entre sí):

- Recorrido (Re): nos da la diferencia entre el mayor y el menor valor de la variable.
- Desviación media (Dm): es la media aritmética de las diferencias, en valor absoluto, entre cada valor y la media aritmética. $Dm = \frac{\sum |x_i - \bar{x}| \cdot f_i}{N}$

- Varianza (σ^2): es la media aritmética de los cuadrados de las diferencias entre cada valor y la

media aritmética.
$$\sigma^2 = \frac{\sum (x_i - \bar{x})^2 \cdot f_i}{N} = \frac{\sum x_i^2 \cdot f_i}{N} - \bar{x}^2$$

- Desviación típica (σ): es la raíz cuadrada de la varianza. $\sigma = \sqrt{\text{varianza}}$

Coefficiente de variación: sirve para comprobar si una desviación típica és un valor grande o pequeño. Lo que hace es comparar la desviación típica con la media aritmética y posteriormente, obtener su valor en

forma de porcentaje.
$$C.V. = \frac{\sigma}{\bar{x}}$$

9. Halla los parámetros de dispersión de los valores de la actividad nº 7.
10. Para acudir a su trabajo, Juan Fuster puede atravesar el centro de su ciudad o circular por la circunvalación. Durante 14 días anotó los tiempos empleados en minutos al escoger uno u otro camino:

Por el centro	17	25	20	26	16	28	29
Por la circunvalación	22	25	20	22	25	21	26

- a) Calcula la media de los tiempos utilizados en cada trayecto.
 b) Obtén el recorrido de los tiempos empleados.
 c) ¿Cuál es la varianza y la desviación típica de los tiempos de cada trayecto?
 d) ¿Qué coeficiente de variación tiene cada ruta?
 e) ¿Qué ruta le recomiendas a Juan? ¿Por qué?
11. Los siguientes datos corresponden a los días de trabajo perdidos por 20 trabajadores la última semana, en que ha habido una epidemia de gripe:

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

- a) Elabora una tabla de frecuencias a partir de estos datos.
 b) Representa los datos en el diagrama más adecuado.
 c) Calcula la media, mediana y moda de la distribución.
 d) Obtén el recorrido, la varianza y la desviación típica de los datos.
 e) Deduce el coeficiente de variación.
12. Entre los 538 niños recién nacidos en un hospital a lo largo de un año, los pesos de un grupo representativo de ellos han sido:

Peso (intervalos)	Kg (X_i)	Niños (f_i)					
[2'5,3[6					
[3,3'5[23					
[3'5,4[12					
[4,4'5[9					

- a) Diferencia entre la población y la muestra de este estudio.
 b) Indica cuál es la variable estadística y de qué tipo es.
 c) ¿Qué porcentaje de recién nacidos pesan más de 3 kg?
 d) Representa mediante el diagrama más adecuado el estudio estadístico anterior.
 e) ¿Cuál es el peso medio?
 f) Deduce cuál es el intervalo donde está la mediana y razona cuál es el intervalo modal.

- g) Calcula la desviación media.
- h) Obtén la varianza y la desviación típica.
- i) Indica cuál es el recorrido y el coeficiente de variación.

13. A la pregunta: ¿Cuántas personas forman tu núcleo familiar?, 40 personas respondieron:

5	5	4	7	4	6	4	6	5	6
3	5	5	3	4	4	6	5	5	5
5	4	7	5	6	3	5	6	7	4
5	5	4	3	5	5	4	3	5	6




















- a) Haz la tabla de frecuencias y el diagrama correspondiente.
- b) Calcula la media, la mediana, la moda y la desviación típica.

14. Calcula \bar{x} y σ de esta distribución:

Tiempo utilizado para ir de casa al instituto (los tiempos de algunos alumnos de 1ºD distorsionaban los valores de la tabla y no se incluyeron)

Minutos	0-5	5-10	10-15	15-20	20-25	25-30
Frecuencia	2	11	13	6	3	1

Utilización de la calculadora científica:

Modo estadístico	 
Introducción de datos (x_i)	x_i 
Borrar datos	 
Número de datos (N)	 
Suma de valores ($\sum x_i \cdot f_i$)	 
Suma de cuadrados de los valores ($\sum x_i^2 \cdot f_i$)	 
Media aritmética (\bar{x})	 
Desviación típica (σ)	 
Introducción de datos agrupados ($x_i \cdot f_i$)	x_i  f_i 
Modo normal	 

- 15. Utiliza la calculadora para hacer los cálculos de algunas de las actividades anteriores.
- 16. Los gastos mensuales de una empresa A tienen una media de 60000 € y una desviación típica de 7500 €. En otra empresa más pequeña, B, la media de 9000 € y la desviación típica, 1500 €. Calcula, mediante el coeficiente de variación, cuál de las dos tiene más variación relativa.
- 17. Las estaturas aproximadas de 4350 personas jóvenes son:

(m)	1'52	1'56	1'60	1'64	1'68	1'72	1'76	1'80	1'84	1'88
personas	62	186	530	812	953	860	507	285	126	29

Decimos que las personas que tienen una estatura:

- entre $(\bar{x} + \sigma, \bar{x} + 3\sigma)$, son estadísticamente altas
- entre $(\bar{x} - 3\sigma, \bar{x} - \sigma)$, son bajas
- entre $(\bar{x} - \sigma, \bar{x} + \sigma)$, son normales

Calcula, aproximadamente, qué tanto por ciento de bajos, normales y altos hay.

- 18.** Con los datos de las cuatro actividades siguientes, se ha de efectuar un estudio estadístico completo, teniendo en cuenta los siguientes puntos:
- a) Describir con claridad la variable estadística motivo de cada estudio, población y muestra.
 - b) Clasificar la variable estadística.
 - c) Ordenación de los datos obtenidos.
 - d) Tabla de frecuencias completa.
 - e) Gráficos estadísticos correspondientes, según el tipo de variable.
 - f) Medidas centrales: Media, Mediana y Moda; según el tipo de variable, justificarlas.
 - g) Medidas de dispersión: Recorrido, Desviación media y Desviación típica, según variable.
 - h) Coeficiente de variación e intervalo de normalidad estadística, si corresponde.
 - i) Consecuencias obtenidas.

18.A) Durante la primera evaluación una alumna ha obtenido las siguientes notas en el área de matemáticas: 6'5, 4'25, 5, 7'5 y 5.

18.B) Durante el año 1970, la población española se distribuía en las siguientes zonas: Urbana, 18 millones de personas; Interurbana, 7 millones; Rural, 9 millones.

18.C) Cierta año, durante el mes de julio, en Alicante, se registraron las siguientes temperaturas máximas:

32	31	28	29	33	32	31
30	31	31	27	28	29	29
30	32	31	31	30	30	29
29	30	30	31	30	31	34
33	33	28				

18.D) Las puntuaciones obtenidas por un grupo de niños de una escuela infantil en un test de habilidad psicomotriz, han sido los siguientes:

Entre 5 y 10 puntos, 3 niños; entre 10 y 15 puntos, 6 niños; entre 15 y 20 puntos, 13 niños; entre 20 y 25 puntos, 7 niños; entre 25 y 30 puntos, 2 niños.