# Unitat 2: Paràmetres estadístics.

### 2.1. Continguts.

Quan es vol fer un estudi estadístic, un cop s'han recollitr les dades i s'han posat en forma de taula, com hem vist en la unitat anterior, el següent pas consisteix en determinar uns valors que ens donen una idea de la posició de les dades i com es troben distribuïdes. Els paràmetres que ens donen idea de la posició són els **paràmetres de centralització**, entre els quals tenim: la *mitjana aritmètica*, el *mode* i la *mediana*, mentre que els paràmetres que ens indiquen com de disperses es troben les dades entorn d'un valor central són els **paràmetres de dispersió**, entre els quals tenim: el *rang* i la *desviació estàndard*.

### 2.1.1. Paràmetres de centralització.

 a) La mitjana aritmètica d'un conjunt de dades x<sub>i</sub> amb freqüència absoluta n<sub>i</sub> s'obté aplicant la fórmula:

$$\overline{x} = \frac{\sum x_i \cdot n_i}{N}$$

Exemple 1:

La mitjana aritmètica de les dades que observeu en la taula adjunta és  $\overline{x} = \frac{96}{34} = 2,82$ .

Xi	ni	x <sub>i</sub> :n <sub>i</sub>
1	8	8
2	7	14
3	9	27
4	4	16
5	5	25
6	1	6
	N = 34	96

b) El mode és el valor de la variable que té màxima freqüència absoluta i ve representada per Mo. Si les dades es troben agrupades en intervals de classe, a l'interval que té màxima freqüència, conté al mode i s'anomena interval modal.

Una manera gràfica de trobar el valor de Mo quan les dades estan agrupades és la següent:





3) El mode en la taula adjunta on les dades es troben agrupades és troba en l'interval (4,6].

Intervals	freq. absoluta
(0,2]	5
(2,4]	7
(4,6]	12
(6,8]	3
(8,10]	8
	N = 35

D'aquí que 4 < Mo  $\leq$  6, i tenint en compte el seu histograma, Mo  $\approx$  4,6 (veure el gràfic de la pàgina anterior).

c) La mediana és el valor que deixa per davant seu el 50% de les dades un cop estan ordenades. Si les dades es troben agrupades, el següent mètode gràfic ens dona una bona aproximació de la mediana:



Exemple 4-5:

- 4) La mediana de l'exemple 1 és Me = 3, ja que és el valor que deixa la meitat de les dades per davant seu.
- La mediana de l'exemple 3 es troben en l'interval (4,6]. D'aquí que 4 < Me ≤ 6, i tenint en compte el seu histograma acumulatiu, Me ≈ 5 (veure el gràfic anterior).</li>

## 2.1.2. Paràmetres de dispersió.

- a) El **rang** d'un conjunt de dades és la diferència entre el valor més gran i el valor més petit.
- b) La **desviació estàndard** ens dóna el grau de dispersió dels valors de la variable respecte de la mitjana aritmètica i ve donada per la fórmula:

$$S_x = \sqrt{\frac{\sum x_i^2 \cdot n_i}{N} - \overline{x}^2}$$

Exemple 6:

El rang de la ta	aula adjunta és .	rang = 6 – 1 = 5.
------------------	-------------------	-------------------

Xi	ni	x <sub>i</sub> ·n <sub>i</sub>	$x_i^2 \cdot n_i$
1	8	8	8
2	7	14	28
3	9	27	81
4	4	16	64
5	5	25	125
6	1	6	36
	N = 34	96	342

i com la mitjana aritmètica és  $\overline{x} = \frac{96}{34} = 2,82$ , tenim que la desviació estàndard és:

$$S_x = \sqrt{\frac{342}{34} - 2,82^2} = 1,45$$

Si les dades es troben agrupades en intervals de classe, la fórmula de la desviació estàndard continua essent vàlida. Només cal tenir en compte que els valors  $x_i$  són les marques de classe de cadascun dels intervals.

### 2.1.3. Interpretació conjunta de la mitjana i la desviació estàndard.

Ja sabem calcular la mitjana aritmètica i la desviació estàndard d'una distribució de dades estadístiques. A continuació veurem una de les moltes aplicacions que es poden extreure de considerar conjuntament aquests paràmetres.

El **coeficient de variació** d'una variable estadística és una mesura de dispersió relativa ja que es dóna de forma percentual. Ve donat per la fórmula:

$$CV_x = \frac{S_x}{\overline{r}} \cdot 100$$

Aquest coeficient ens indica fins a quin punt es troben concentrades les dades al voltant de la mitjana aritmètica i ens permet comparar distribucións per a poder treure conclusions.

Exemple 7:

Anem a considerar dos amics, en Joan i la Maria, que estudien 1r de Batxillerat en un mateix centre. Les notes obtingudes per en Joan i La Maria en la darrera avaluació han estat:

	Cat	Cast	Ang	Filo	E.Fís	Mates	Fís	Tecn	D.Tec
Joan	7	6	4	5	7	5	6	4	5
Maria	6	6	5	5	6	6	7	3	5

Quin dels dos amics es pot considerar que ha obtingut millors resultats?

La qualificació mitjana d'en joan és: x = 5,44, mentre que la qualificació mitjana de la Maria és també y = 5,44, ja que:

Xi	ni	x <sub>i</sub> ·n <sub>i</sub>	$x_i^2 \cdot n_i$
4	2	8	32
5	3	15	75
6	2	12	72
7	2	14	98
	N=9	49	277

V:	n:	v.n.	$v_i^2 \cdot n_i$
y,		y/11/	y/ 11/
3	2	6	18
5	2	10	50
6	3	18	108
7	1	7	49
8	1	8	64
	N=9	49	289

Està clar que la mitjana aritmètica no ens dóna la resposta a la nostra pregunta, però si calculem les desviacions estàndards, tenim:

$$S_x = \sqrt{\frac{277}{9} - 5,44^2} = 1,09$$
  $i$   $S_y = \sqrt{\frac{289}{9} - 5,44^2} = 1,59$ 

Això ens porta a la conclusió que les qualificacions d'en Joan són menys disperses que les de la Maria, i que per tant en Joan ha tret millors resultats.

Una manera de presentar aquesta informació és donant-la de forma percentual, és per això que haurem de calcular els respectius coeficients de variació. Així doncs, tenim que:

$$CV_x = \frac{1.09}{5.44} \cdot 100 = 20,04\%$$
 *i*  $CV_y = \frac{1.59}{5.44} \cdot 100 = 29,23\%$ 

Les qualificacions d'en Joan es troben més concentrades al voltant de la mitjana aritmètica perquè té menys dispersió relativa.

## 2.2. Activitats d'aprenentatge amb suport informàtic.

### Activitat 1: ESO\_1.XLS

En aquesta activitat es pretén calcular la mitjana aritmètica, la mediana i el mode de les edats dels alumnes d'una classe de 4t d'ESO. Sabem que l'edat dels 18 alumnes que formen aquesta classe és: 13, 14, 14, 15, 16, 15, 14, 14, 14, 14, 14, 13, 14, 14, 13, 15, 14 i 15.

Es demana:

- a) Construïu la taula de distribucions de freqüències restringida a les columnes: x<sub>i</sub> i n<sub>i</sub>.
- b) Representeu el diagrama de barres.
- c) Afegiu una nova columna formada pels diferents productes x<sub>i</sub>·n<sub>i</sub>. La suma d'aquests productes dividida per la grandària de la mostra ens dóna la mitjana aritmètica.
- d) Com N = 20, la mediana ocupa el punt mig de les posicions 10 i 11. Observeu quines edats es troben en aquestes dues posicions i determineu el seu valor.
- e) Quin valor pren el mode en aquesta distribució?.

Deseu aquesta activitat amb el nom **ESO\_1.XLS**.

### Activitat 2: CATALA\_1.XLS

Com podem calcular els paràmetres de centralització o de posició quan les dades es troben agrupades en intervals de classe?

Quan les dades es troben agrupades, hem de tenir en compte que...

- 1. El càlcul de la mitjana aritmètica passa per utilitzar la seva fórmula considerant els valors de xi com la marca de classe de cada un dels intervals.
- 2. Els càlculs de la mediana i el mode passen per situar aquests paràmetres en intervals de classe i mirar de deduir de manera gràfica una bona aproximació.

En aquesta activitat es pretén determinar els paràmetres de centralització de les notes de 35 alumnes d'una classe de 1r de Batxillerat en un examen de Llengua i literatura catalana agrupant prèviament les dades en intervals de classe. Aquestes notes són:

5,	7,	З,	2,	6,	4,	8,	2,	1,	9,	3,	7,
6,	9,	5,	7,	9,	0,	3,	2,	1,	4,	9,	3,
2,	7,	9,	4,	8,	7,	2,	1,	5,	10,	9 i	8.

 a) Agrupeu aquestes dades en 5 intervals de classe i construïu la taula restringida a x<sub>i</sub>, n<sub>i</sub> i N<sub>i</sub> de les distribucions de freqüències. b) Representeu l'histograma i el polígon de freqüències acumulatiu.

c) Calculeu la mitjana aritmètica i doneu una bona aproximació de la mediana i el mode utilitzant els gràfics.

Deseu aquesta activitat com a CATALA\_1.XLS.

### Activitat 3: ESO\_2.XLS

Hi ha una dita que diu: "Sí *tenim dos pollastres a repartir entre tu i jo, a nivell estadístic, tant és que ens mengem un cadascú o que jo em mengi els dos i tu cap*". Òbviament els dos casos tenen la mateixa mitjana aritmètica, però per diferenciar-los disposem dels paràmetres de dispersió. Aquests paràmetres ens informen si les dades estan molt concentrades o no al voltant dels paràmetres de centralització.

En aquesta activitat volem determinar el rang, la desviació estàndard i el coeficient de variació de les edats observades en l'activitat 1.

Recupereu de l'arxiu ESO\_1.XLS la taula de distribucions de freqüències i determineu:

- a) El rang de les edats d'aquells alumnes.
- b) Completeu la taula amb les columnes que falten per a poder calcular la desviació estàndard.
- c) Quin valor pren el coeficient de variació? Interpreteu-lo.

Deseu aquesta activitat amb el nom **ESO\_2.XLS**.

### Activitat 4: CATALA\_2.XLS

Cal que recupereu de l'arxiu CATALA\_1.XLS la taula de distribució de freqüències i que afegiu les columnes necessàries per a determinar la desviació estàndard i el coeficient de variació de les notes de Llengua i literatura catalana de 1r de Batxillerat.

Deseu aquesta activitat amb el nom CATALA\_2.XLS.

### Activitat 5: ESTADISTICA\_1.XLS

No ens hem d'oblidar que l'EXCEL és una aplicació informàtica que realitza un gran nombre de càlculs. Es pot utilitzar com una simple calculadora, o bé, com una eina molt potent que ens permet fer ús d'unes funcions que l'aplicacio té incorporades.

Per activar les funcions cal fer un clic a la icona  $f_x$  del menú superior de la pantalla i que rep el nom d'**assistent de funcions.** Ens surt una finestra amb totes les possibles funcions que es poden usar en EXCEL.

Pegar función	<u>?</u> ×
Categoría de la función:	Nombre de la función:
Usadas recientemente Todas Financieras Fecha y hora Matemáticas y trigonométricas Estadísticas Búsqueda y referencia Base de datos Texto Lógicas Información	DESVESTP PROMEDIO SUMA SI HIPERVINCULO CONTAR MAX SENO SUMAR.SI PAGO
DESVESTP(número1;número2; Calcula la desviación estándar de la Omite los valores lógicos y el texto	a población total tomada como argumentos.
2	Aceptar Cancelar

A nosaltres només ens interessa les funcions estadístiques. Si les seleccionem, a la part dreta de la finestra ens surten totes les funcions estadístiques que ens ofereix l'aplicacio. Ara només cal escollir la funció que més ens convé.

Pegar función	? ×
Categoría de la función:	Nombre de la función:
Usadas recientemente Todas Financieras Fecha y hora Matemáticas y trigonométricas Estadísticas Búsqueda y referencia Base de datos Texto Lógicas Información	BINOM.CRIT COEF.DE.CORREL COEFICIENTE.ASIMETRIA COEFICIENTE.R2 CONTAR CONTAR.BLANCO CONTAR.SI CONTARA COVAR CRECIMIENTO CUARTIL
BINOM.CRIT(ensayos;prob_é> Devuelve el menor valor cuya distr igual que un valor de criterio.	ibución binomial acumulativa es mayor o

Les funcions que farem servir en les properes activitats fan referència a fórmules que hem editat anteriorment, o bé, d'altres que no hem tingut en compte, però que amb tota seguretat ens haguessin pogut ser útils.

Algunes d'aquestes són:

CONTAR.SI(rang;k)	Ens permet comptar la freqüència d'un valor k del rang de dades.
PROMEDIO(rang),	Ens dóna la mitjana aritmètica d'un rang de dades.
MEDIANA(rang),	Ens dóna la mediana d'un rang de dades.
MODA(rang),	Ens dóna el mode d'un rang de dades.
MAX(rang),	Ens dóna el valor màxim d'un rang de dades.
MIN(rang)	Ens dóna el valor mínim d'un rang de dades.
DESVESTP(rang)	Ens dóna la desviació estàndard poblacional d'un rang de dades.
VARP(rang)	Ens dóna la variància d'un rang de dades.

Observeu que passa si seleccionem PROMEDIO. En el quadre de diàleg se'ns dona una descripció de la funció. En aquest cas ens diu que la funció PROMEDIO retorna la mitjana dels arguments (cal entendre per argument els valors que se li donen a la funció perquè aquesta operi amb ells i ens retorni un resulta!). Si accepten, ens surt el següent quadre:

PROMEDIO	
Número1	🛐 = número
Número2	🗾 = número
	-
Devuelve el promedio (media aritmétic matrices, o referencias que contenga	<ul> <li>a) de los argumentos, los cuales pueden ser números, nombres, n números.</li> </ul>
Número1:	número1;número2; son entre 1 y 30 argumentos numéricos de los que se desea obtener el promedio.
Resultado de la fórmula	= Aceptar Cancelar

En els dos quadres en blanc podem posar els arguments de la funció. Al situar-nos sobre ells, l'EXCEL ens dona una petita descripció de l'argument.

En la part baixa de la finestra, on posa "Resultat de la formula =", un cop introduïts els arguments, es mostra el resultat final obtingut com a mitjana aritmètica.

Utilitzeu L'assistent de funcions per visualitzar els quadres de diàleg que surten quan utilitzem les funcions CONTAR.SI, PROMEDIO, MEDIANA, MODA, DESVESTP i VARP.

En el rang A1:A30 poseu les següents dades que corresponen als resultats d'un examen de Matemàtiques de 3r d'ESO: 3, 2, 4, 1, 5, 9, 4, 7, 6, 5, 8, 10, 3, 2, 5, 7, 4, 8, 9, 3, 4, 3, 7, 6, 10, 0, 2, 7, 8 i 5.

i, apliqueu l'auxiliar de funcions per tal de determinar:

- a) La freqüència absoluta del valor 8
- b) la mitjana aritmètica,
- c) la mediana,

- d) el mode,
- e) el valor màxim,
- f) el valor mínim,
- g) la desviació estàndard,
- h) la variància,

Deseu aquesta activitat amb el nom ESTADISTICA\_1.XLS.

### Activitat 6: ESTADISTICA\_2.XLS

Recolliu el nombre de germans que té cadascun de vosaltres en una columna d'un full de càlcul, i sense editar cap tipus de fórmula (només utilitzant l'auxiliar de funcions), determineu:

- a) La mitjana aritmètica.
- b) la mediana.
- c) el mode.
- d) el valor màxim.
- e) el valor mínim.
- f) la freqüència del valor 2.
- g) la variància.
- h) la desviació estàndard.

Podem calcular el rang i el coeficient de variació d'aquestes dades només fent servir l'assistent de funcions de l'EXCEL? Independentment de quina sigui la resposta, determineu el valor d'ambdós paràmetres.

Deseu aquesta activitat amb el nom ESTADISTICA\_2.XLS.

### Activitat 7: PES\_ALÇADA.XLS

En aquesta activitat compararem els resultats obtinguts si apliquem fórmules per als diferents pàmetres estudiats amb els obtinguts si apliquem l'assistent de funcions d'EXCEL.

Considereu l'alçada i el pes dels 15 alumnes de 3r d'ESO que es troben en una classe de Tecnologia.

Els pesos i les alçades d'aquests alumnes són:

ALUMNE	PES	ALÇADA
1	60	169
2	62	170
3	68	173
4	58	168

5	52	164
6	55	165
7	57	167
8	64	170
9	58	169
10	60	168
11	61	169
12	65	172
13	59	168
14	55	166
15	59	169

- a) Poseu ALUMNE en la cel·la A1, PES en la cel·la B1 i ALÇADA en la cel·la C1. En els rangs A2:A16, B2:B16 i C2:C16 poseu les dades del número d'alumne, del pes i de l'alçada.
- b) En el rang E1:E7 poseu PARÀMETRES, MITJANA, MEDIANA, MODE, VALOR MÀXIM, VALOR MÍNIM i DESVIACIÓ ESTÀNDARD, i en el rang F2:F7 editeu les fórmules sempre que sigui possible fer-ho.
- c) En el rang de cel·les H2:H7, poseu els resultats que obteniu d'utilitzar l'auxiliar de funcions d'EXCEL.
- d) Calculeu en tots dos casos els coeficients de variació i indiqueu quina de les dues distribucions (la del pes i la de l'alçada) tpresenta menys dispersió relativa?

Deseu aquesta activitat amb el nom **PES\_ALÇADA.XLS**.

### Activitat 8: COMPARACIÓ.XLS

Considereu les dades de la taula adjunta:

Notes de Llengua catalana	8	6	5	7	8	4	5	7	8	9	2	5	5	10	3
Notes de Matemàtiques	6	7	5	7	7	5	6	6	8	9	4	6	5	9	4

Es demana:

- a) Determineu el coeficient de variació de les notes de Llengua catalana.
- b) Determineu el coeficient de variació de les notes de Matemàtiques.
- c) Quina de les dues distribucions presenta menys dispersió relativa?

Deseu l'activitat amb el nom **COMPARACIO.XLS**.

### Activitat 9: ESTADISTICA\_3.XLS

 a) Reprodueix en un full de càlcul la fitxa següent, posant en les cel·les de fons blanc les fórmules que corresponen fent servir sempre que sigui possible l'assistent de funcions de l'EXCEL i omplint la columna de dades amb les dades que compte.

DADES	Variable	Freq. Absoluta	PARÀMETRES	
1	1	1	GRANDARIA MOSTRA	20
5	2	2	MITJANA	5,7
6	3	1	MEDIANA	5,5
2	4	1	MODE	Ę
2	5	5	VALOR MÀXIM	10
9	6	3	VALOR MÍNIM	
6	7	2	RANG	ç
5	8	1	DESV. ESTÀNDARD	2,47
6	9	3	VARIÀNCIA	6,11
4	10	1	COEF. VARIACIÓ	43,4
8		20		
9	Gràfic:			
5				
10				
3				
/ 				
5				
5				
/ 9				

- b) Visualitzeu el diagrama de barres que s'obté en el quadre en blanc preparat al respecte.
- c) Utilitza la fitxa que heu creat per saber com varien els valors dels paràmetres quan les notes són: 7, 0, 8, 3, 4, 6, 3, 6, 9, 10, 9, 10, 5, 2, 7, 6, 5, 1, 2 i 8.
  Visualitzeu el diagrama de barres que s'obté.
- d) Digueu en quin dels dos casos es dóna menys dispersió relativa.

Guardeu el llibre amb el nom ESTADISTICA\_3.XLS.

### Activitat 10: ESTADISTICA\_4.XLS

Elaboreu una fitxa semblant a l'obtinguda en l'activitat anterior que us permeti calcular els mateixos paràmetres i que presenti la taula amb dades agrupades. El rang de dades feu-lo per a les 30 notes d'un examen.

- a) Introduïu les dades següents: 7, 4, 5, 6, 7, 2, 3, 4, 6, 5, 4, 1, 2, 8, 9, 6, 9, 5, 4, 3, 7, 6, 4, 5, 1, 2, 2, 3, 4 i 10, i agrupeu-les en 5 intervals d'amplitud 2.
- b) Visualitzeu en aquest cas l'histograma i el poligon de freqüències.

Guardeu el llibre amb el nom **ESTADISTICA\_4.XLS**.

### 2.3. Activitats d'avaluació.

### Activitat 1

Obriu el llibre AVALUACIO\_1.XLS i deseu-lo ara amb el nom AVALUACIO\_2.XLS.

- a) En el full Activitat 1, tenim una taula de freqüències. Completeu-la per a poder calcular la mitjana aritmètica i la desviació estàndard editant la fórmula corresponent.
- b) Quin valor té el mode? Quin valor té la mediana?
- c) Feu servir l'assistent de funcions de l'EXCEL per comprovar que els resultats obtinguts són correctes.

No us oblideu de desar el llibre. Recordeu que s'anomena AVALUACIO\_2.XLS.

#### Activitat 2

Continueu amb AVALUACIO\_2.XLS.

- a) En el full Activitat 2, tenim una taula de freqüències amb dades agrupades.
   Completeu-la per a poder calcular la mitjana aritmètica i la desviació estàndard.
- b) Utilitzeu l'Histograma i el polígon de freqüències per donar una bona aproximació del mode i de la mediana.

No us oblideu de desar el llibre. Recordeu que s'anomena AVALUACIO\_2.XLS.

#### Activitat 3

Continueu amb AVALUACIO\_2.XLS. Elimineu el contingut del full Activitat 3.

En el full **Activitat 3**, situeu-vos en la cel·la A1 i escriviu el rètol **DADES**. En el rang A2:A51 poseu les dades: 5, 6, 7, 5, 3, 5, 6, 2, 1, 2, 4, 9, 3, 4, 5, 6, 2, 2, 9, 10, 2, 1, 4, 3, 5, 4, 9, 2, 5, 5, 7, 6, 8, 6, 5, 1, 2, 3, 4, 5, 4, 4, 8, 9, 10, 2, 3, 5, 5 i 6.

- a) Determineu, fent us de l'assistent de funcions de l'EXCEL: La freqüència absoluta de 7, la mediana, el mode, el valor màxim, el valor mínim, la mitjana aritmètica, la desviació estàndard.
- b) Fent us dels resultats anterior i aplicant fórmules, determineu el rang i el coeficient de variació.

No us oblideu de desar el llibre. Recordeu que s'anomena AVALUACIO\_2.XLS.

### Activitat 4

Continueu amb AVALUACIO\_2.XLS. En el full Activitat 4, teniu la taula adjunta:

Vendes de l'empresa "NET" per trimestres (en €)								
Articles	1r trim.	2n trim.	3r trim.	4t trim.	<b>Total articles</b>			
Sabó	15.200	17.500	12.350	18.756	63.806			
Paper higiènic	9.945	7.890	2.340	9.925	30.100			
Renta vaixelles	12.000	15.345	17.005	12.035	56.385			
Mocadors	4.500	3.456	5.670	8.750	22.376			
Altres productes	5.020	3.450	2.780	5.900	17.150			
Total trimestre	46.665	47.641	40.145	55.366	189.817			

- a) Determineu els ingressos mitjos anuals de cadascun dels articles i les seves desviacions estàndard.
- b) Quin dels articles presenta menys dispersió relativa?
- c) Representeu el corresponent diagrama de sectors per a les vendes del 1r trimestre.

No us oblideu de desar el llibre. Recordeu que s'anomena AVALUACIO\_2.XLS.