

## Objectius

En aquesta quinzena aprendràs a:

- Recordar els conceptes de població, mostra, individu i caràcter.
- Valorar la importància del concepte de variable estadística i distingir entre els diferents tipus.
- Resumir mitjançant una taula de freqüències qualsevol sèrie de dades.
- Associar i interpretar gràfics estadístics valorant el seu ús en diferents àrees de coneixement.
- Calcular, valorar i interpretar la mitjana i moda en variable discreta.

### Abans de començar

1. Vocabulari estadístic ..... pàg. 4  
Població, mostra, individu i caràcter
2. Caràcter. Variable estadística ..... pàg. 6  
Caràcter qualitatiu. Atributs  
Variables discretes  
Variables contínues
3. Ordenació de dades. Tabulació ..... pàg. 10  
Para variable discreta  
Para variable qualitativa
4. Gràfics para variable qualitativa ... pàg. 12  
Diagrama de barres  
Diagrama de sectors
5. Gràfics per a variable discreta ..... pàg. 14  
Diagrama de barres  
Polígons de freqüències  
Diagrama de sectors
6. Mesures de centralització ..... pàg. 17  
Mitjana  
Mediana  
Moda
7. Mesures de dispersió ..... pàg. 20  
Rang  
Desviació mitjana

RESUM

Autoavaluació



## Abans de començar

L'Estadística ha calat en múltiples aspectes de la vida quotidiana fent familiars termes com població, mostra, mitjana, mediana, moda...

Pot assegurar-se que qualsevol persona informada d'avui en dia posseeix un vocabulari bàsic d'estadística, l'entén, l'utilitza i el valora.

Pràcticament totes les ciències, tant científic/tecnològiques com socials utilitzen en aspectes fonamentals de les mateixes a l'estadística.

L'esport no és una excepció. En tots ells i en particular en el bàsquet l'ús de les dades estadístiques constitueix un aspecte a estudiar i manejar tan important a vegades com les tàctiques i la tècnica implícites del propi joc.

En l'exemple següent simula un servei de fons en bàsquet, es representa amb punts vermells els jugadors atacants i amb verds els defensors.

L'estudi que realitzen els cossos tècnics dels equips s'encarrega de calcular quina estadística de tir té cada jugador, d'aquesta manera si es deixa desmarcat el jugador que tingui pitjor estadística; la pilota anirà cap a ell.



### 1. Vocabulari estadístic

#### Població, mostra, individu i caràcter

Las primeres definicions necessàries per a l'inici de qualsevol estudi estadístic són població, individu, mostra i caràcter.

Comencem amb un exemple que ens faci intuir aquests conceptes.

##### Exemple

##### Estudi sobre la possible existència de vida en altres estrelles.

Existeixen sistemes planetaris semblants al nostre que, tal vegada puguin albergar algun tipus de vida?

Fins fa molt poc temps, els astrònoms no tenien proves de l'existència de planetes fora del Sistema Solar.

En l'actualitat s'han descobert alguns centenars de planetes gegants, que difícilment podrien encerrar vida, però que sí serien una espècie de senyal que en aquella estrella pot existir un sistema amb orbitats i grandària més d'acord amb les possibilitats de vida en el sentit que coneixem.

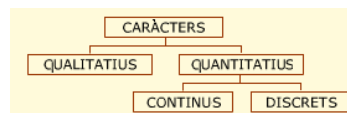
- La població està constituïda per totes les estrelles del univers visible.
- La mostra està constituïda per totes les estrelles escollides i observades en el projecte.
- El individu és cada estrella de l'univers observable.
- El caràcter és la presència o no de perturbacions que indiquen l'existència de

planetes gegants.

Podem definir els conceptes anteriors de la següent forma:

- **Població:** conjunt de tots els elements que verifiquen una característica que serà objecte d'estudi.
- **Individu:** Cada un dels elements de la població.
- **Mostra:** qualsevol subconjunt de la població. Aquest subconjunt és molt important que sigui *representatiu* de la població.
- **Caràcter:** Cada una de les propietats que posseeixen els individus de la població i que poden ser objecte d'estudi.

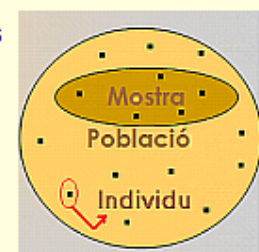
La definició de caràcter ha d'anar acompanyada de la següent classificació:



Recordem llavors que davant qualsevol estudi estadístic hem de tenir en compte la identificació dels elements, d'aquesta forma evitarem errors en les conclusions finals.

##### Identificació d'elements

- Població.
- Mostra
- Individu
- Caràcter



## Exemple

### Estudi sobre l'evolució de la talla en la joventut espanyola

Els espanyols igualen l'estatura a la majoria dels europeus, però evolucionen cap a l'obesitat nord-americana.

Un estudi antropomètric conjunt entre varis hospitals espanyols, revela que l'estatura dels espanyols s'ha igualat en els últims trenta anys respecte a la majoria dels països europeus.

El mateix estudi també alerta sobre la preocupant tendència cap a l'obesitat en nivells similars a la població nord-americana.

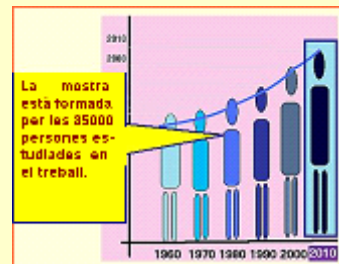
El treball, portat a terme mitjançant la mesura de quasi 35000 subjectes entre els anys 2000 i 2004, també demostrà que les diferències entre les distintes comunitats autònomes dins d'Espanya són quasi inexistent.

Població



La població està formada per totes les persones espanyoles majors de 18 anys.

Mostra



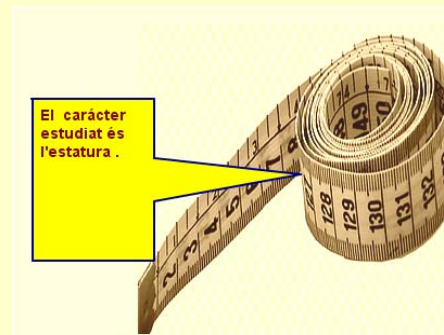
La mostra està formada per les 35000 persones estudiades en el treball.

Individu



L'individu és cada una de les persones que formen la població adulta espanyola.

Caràcter



El caràcter estudiat és l'estatura.

## 2. Caràcter. Variable estadística

### Treball de camp

El treball de camp és l'etapa de la investigació en la qual s'estableix contacte directe amb la població o mostra per recollir les dades que es necessiten.

La planificació és fonamental i el seu desenvolupament depèn del mètode d'obtenció de la informació que s'utilitzi. L'ús de l'ordinador permet una simulació de situacions que fa que realitzem un treball de camp virtual sense desgast físic.

### Caràcter qualitatiu. Atributs

Comencem novament amb un exemple que ens il·lustra.

#### Exemple

##### Afició al futbol

Preguntem a una sèrie de persones sobre les seves preferències respecte a l'afició futbolística.

La mostra que considerem serà de 9 persones de distintes ciutats espanyoles.

Los dades són FC. Barcelona, Sevilla CF., At. Bilbao, R. Madrid, At. Madrid, València CF., FC. Barcelona i Deportivo de la Corunya.

Les característiques d'aquests valors són:

- No són mesurables amb nombres.
- No té sentit l'ordenació.
- Els distints valors s'identifiquen amb el nom de l'equip elegit.

Tots els individus de la població que anem a estudiar tenen una sèrie de propietats o qualitats que en estadística reben el nom de caràcters.

Els caràcters poden ser de dos grans tipus:

- a) QUALITATIUS
- b) QUANTITATIUS

Un caràcter qualitatiu es caracteritza perquè les seves diferents modalitats no poden expressar-se amb nombres.

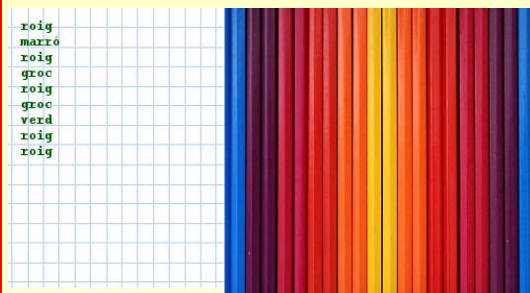
#### Exemple

##### El teu color preferit

Preguntem a una sèrie de persones sobre les seves preferències respecte a colors.

En aquest cas la simulació de la població i color triat es pot realitzar mitjançant l'ordinador, existeixen programes que permeten generar mostres aleatòries que simulen el treball de camp.

La mostra sobre la que actuem serà de 10 persones d'una ciutat qualsevol.



#### CARACTERÍSTIQUES: ▶ ◀

- Els valors que pren no són mesurables numèricament
- No té sentit l'ordenació
- No té sentit parlar de valors CONSECUTIUS
- Les diferents modalitats del caràcter no representables numèricament, s'identifiquen mitjançant el nom del color triat.

## Variables discretes. Caràcter quantitatiu discret

Es denomina així al caràcter les modalitats del qual es poden representar amb nombres.

Dins dels caràcters quantitius es distingeixen dos tipus: discret i continu.

És discret si pren valors aïllats, de manera que entre dos consecutius no n'existeix un altre d'intermedi.

### Exemple

#### Quanta gent hi ha a la platja?

Realitzem una fotografia d'una determinada zona de platja a diferents hores del dia i anotem les persones que hi apareixen.


En aquest cas disposem d'un banc virtual de fotos i un procediment totalment aleatori que simula les diverses situacions.

La mostra sobre la que actuem és de 9 fotografies.

**DADES OBTINGUDES**

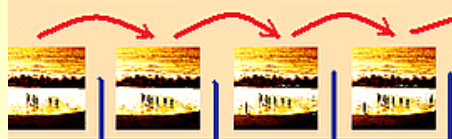
n = 11,00

10 5 5 2 6 3 1 5  
5 2



#### CARACTERÍSTIQUES: ▶ ◀

- Els valors que pren són aïllats; entre 1 i 13.
- Els valors es poden ordenar i comptar
- Entre dos valors CONSECUTIUS no n'hi ha cap altre
- ESQUEMA



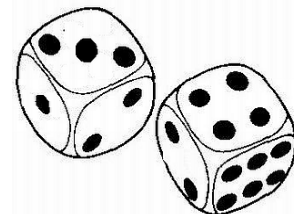
**NO HI HA VALORS INTERMEDIS**

### Exemple

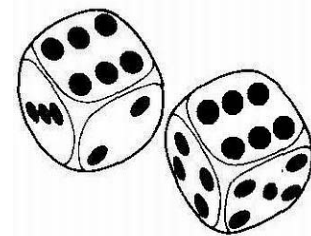
#### Quant sumen les cares superiors de dos daus tirats prèviament?

Tirem dos daus perfectes anotant la suma dels resultats de les cares superiors.

La mostra que considerem serà la suma de 8 parells de tirades.



$$3 + 4 = 7$$



$$6 + 6 = 12$$

Los dados obtinguts són: 7, 6, 9, 2, 8, 1, 8 i 7.

Las características d'aquests valors són:

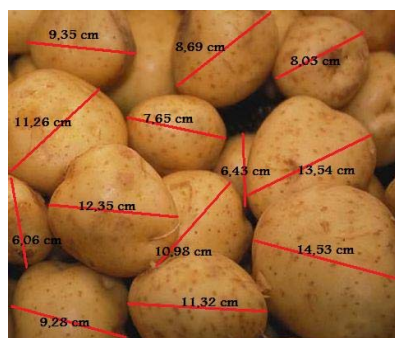
- Els valors que prenen són aïllats; entre 2 i 12.
- Els valors es poden ordenar i comptar.
- Entre dos valors consecutius no existeixen valors intermedis.

## Variables contínues. Caràcter quantitatiu contínu

Quan les modalitats d'un caràcter quantitatiu poden prendre valors d'un conjunt de nombres reals o un interval, al menys teòricament, es diu que estem davant un caràcter quantitatiu.

### Exemple

Mesurant patates. Diàmetre mitjà.



Per a una posterior classificació de qualitat es realitza un estudi sobre el diàmetre mitjà en distintes produccions de patates. La mostra que considerem serà de 8 produccions.

Tot el, en aquest cas, costós procés de recollida de les mostres el substituïm per una simulació per ordinador.

Los dades obtingudes són: 14,6 ; 6,7 ; 9,8 ; 13,2 ; 8,1 ; 9,3; 13,8 i 10,1.

Las característiques que es tenen sobre els valors són:

- Entre dos valors sempre existeix la possibilitat d'un altre.
- No té sentit parlar de valors consecutius.
- Toma valors dins d'un interval.

### Exemple

#### Pesant recent nascuts

Anem a preguntar el pes dels recent nascuts en una determinada ciutat.

Existeixen programes informàtics que permeten generar mostres aleatòries que simulen el treball de camp.

La mostra sobre la que actuem és de 40 nadons.

#### DADES OBTINGUES

2,28	3,22	3,95	2,43	3,59
2,47	3,41	3,50	2,07	2,13
3,55	2,61	3,21	2,29	4,07
2,33	2,51	3,38	3,67	4,10
3,84	2,31	2,63	4,03	2,95
3,86	4,03	3,02	2,99	4,00
3,72	2,98	3,88	3,07	3,28
3,84	3,37	3,55	4,07	2,94

#### CARACTERÍSTIQUES

- Entre dos valors sempre existeix la possibilitat d'un altre
- No té sentit parlar de valors CONSECUTIUS
- Pren valors dintre de l'interval [2,07 , 4,10]
- INTERVAL DE VALORS





## Exercicis resolts

1. Classifica les següents variables: qualitatives, discreta o contínua, escrivint una X en el requadre corresponent.

tipus de música preferida, estatura, marca de refresc favorita, pes de nadons, pel·lícules vistes al més, duració mitjana de bombetes.

SOLUCIÓ:

	QUALITATIVA	DISCRETA	CONTÍNUA
Tipus de música preferida			
Estatura			
Marca de refresc favorita			
Pes de nadons			
Pel·lícules vistes al mes			
Duració mitjana de bombetes			

2. Classifica les següents variables: qualitatives, discreta o contínua, escrivint una X en el requadre corresponent.

Raça de gossos. Nre. de fills, longitud del peu, assignatures pendents, perímetre cranial, cantant favorit.

SOLUCIÓ:

	QUALITATIVA	DISCRETA	CONTÍNUA
Raça de gossos	X		
Nre. de fills		X	
Longitud del peu			X
Assignatures pendents	X		
Perímetre cranial			X
Cantant favorit	X		

Observa:

La variable "assignatures pendents" fa referència al nombre de les assignatures i per això és variable qualitativa, mentre la variable "nre. d' assignatures pendents" seria discreta.

## 3. Ordenació de dades

En finalitzar podràs comprovar que...

### PROPIETATS INTERESSANTS DE LES TAULES ESTADÍSTIQUES

- La suma de totes les freqüències absolutes és igual a la grandària de la població o de la mostra.
- La suma de les freqüències relatives és sempre igual a 1.

*Si s'ha realitzat cap arrodoniment en les freqüències relatives és usual que la suma de les mateixes no sigui exactament igual a un donat als errors comesos.*

## Tabulació per variable discreta

El pas següent al treball de camp és la disposició de les dades de manera ordenada, concisa i visualment atractiva.

En estadística, aquest procés rep el nom de tabulació.

Els valors obtinguts s'ordenen, especificant i agrupant de tal manera que sigui fàcil la informació i la cerca.

Les primeres columnes que han d'aparèixer són:

- Valors de la variable,  $X_i$ .
- Freqüències absolutes,  $f_i$ .
- Freqüències relatives,  $h_i$ .
- Freqüències absolutes, acumulades,  $F_i$ .
- Freqüències relatives acumulades,  $H_i$ .

En alguns casos es pot utilitzar el percentatge en lloc de les freqüències relatives o a més de les freqüències relatives.

### Exemple

#### Edat dels estudiants

Las edats de 30 estudiants d'un institut d'ensenyança secundària dona els valors que posteriorment tabulem a continuació:

16 16 16 13 14 13 13 13 16 16 13 14 12 17 15  
16 15 17 17 15 13 17 12 16 14 16 16 16 14 14

$X_i$	$f_i$	$h_i$	$F_i$	$H_i$
12	2	$\frac{2}{30}$	2	$\frac{2}{30}$
13	6	$\frac{6}{30}$	8	$\frac{8}{30}$
14	5	$\frac{5}{30}$	13	$\frac{13}{30}$
15	3	$\frac{3}{30}$	16	$\frac{16}{30}$
16	10	$\frac{10}{30}$	26	$\frac{26}{30}$
17	4	$\frac{4}{30}$	30	$\frac{30}{30}$

## Tabulació per variable qualitativa

En els casos de caràcter qualitatiu, la tabulació de les dades és molt simple. Les tres columnes que tenen sentit fan referència a:

- El valor dels atributs.
- La freqüència absoluta
- La freqüència relativa

### Exemple

#### La pràctica d'esport

Recollida de dades sobre esports practicats, tabulada:

futbol, tennis, bàsquet, handbol, tennis, voleibol, atletisme, bàsquet, futbol, futbol, handbol, futbol, voleibol, handbol, futbol, handbol, futbol, futbol, tennis, atletisme.

<b>X= esport</b>	<b>f</b>	<b>h</b>
<b>bàsquet</b>	2	2/20
<b>handbol</b>	4	4/20
<b>voleibol</b>	2	2/20
<b>tennis</b>	3	3/20
<b>atletisme</b>	2	2/20
<b>futbol</b>	7	7/20
	20	

### Exercicis resolts

3. Per a un estudi d'accessibilitat, durant 30 dies anotem el nombre de places lliures d'aparcament a les 5 de la tarda.

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

Realitza una tabulació de les dades en la que apareguin les columnes corresponents a les freqüències absolutes, relatives, acumulades absolutes i relatives.

SOLUCIÓ:

Nombre de places d'aparcament	f	h	F	H
0	4	0,14	4	0,14
1	8	0,28	12	0,41
2	8	0,28	20	0,69
3	7	0,24	27	0,93
4	0	0	27	0,93
5	3	0,1	30	1,03
	30			

4. Preguntem a 20 estudiants escollits aleatòriament pel tipus de música que prefereixen escoltar.

Els resultats són: disco, rock, rock, clàssica, rock, llatina, pop, rock, llatina, rock, flamenc, flamenc, flamenc, llatina, rock, clàssica, disco, disco, llatina, rock.

Realitza una tabulació de les dades en la que apareguin les columnes corresponents a les freqüències absolutes i relatives.

SOLUCIÓ:

Tipus de música	f	h
Disco	3	0,1
Rock	7	0,24
Llatina	4	0,14
Clàssica	2	0,07
Flamenc	3	0,1
Pop	1	0,03
	20	

## 4. Gràfics per una variable qualitativa

### Diagrama de barres

El diagrama de barres és, juntament al de sectors, el gràfic més utilitzat per variable qualitativa.

S'utilitza com a complement a la taula de freqüències o substituint aquesta.

A l'eix d'abscisses se situen a igual distància els diferents atributs.

A partir de cada atribut s'aixequen barres d'igual gruix i l'altura de les quals sigui la de la corresponent freqüència absoluta.

### Diagrama de sectors

El diagrama de sectors en variables qualitatives és un dels recursos estadístics més utilitzats.

És especialment útil en els casos en què existeixen poques modalitats del caràcter. Se solen utilitzar juntament a la taula de freqüències o substituint a aquesta.

Per calcular l'angle del sector que correspon a cada valor de la freqüència:

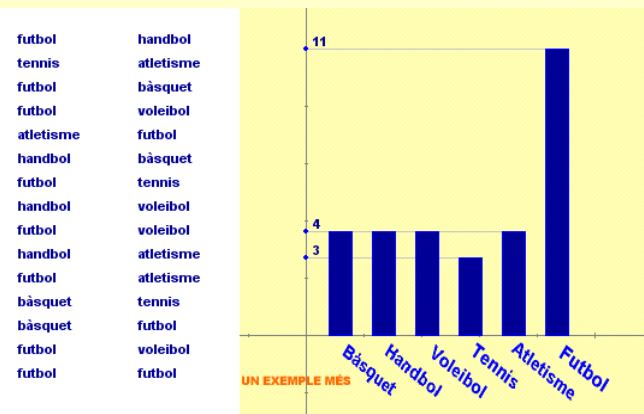
$$x = \frac{f_i \cdot 360}{\text{Total Dades}}$$

### Exemples

#### Esports practicats (Diagrama de barres)

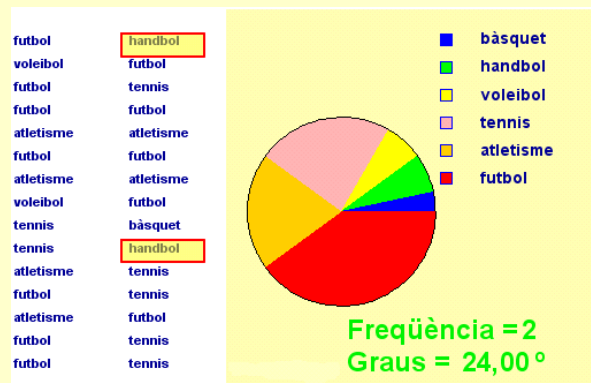
Los dades corresponen a les respostes donades per 30 estudiants sobre l'esport que practicaven amb major freqüència a l'institut.

Si volem tenir una ràpida visió de les dades, una forma d'organitzar-les és a través d'una representació de diagrama de barres. En aquest exemple pots veure la diferència entre fer unes anàlisis sobre el llistat o sobre la gràfica. Quina et resulta més fàcil?



#### Tipus de pel·lícula (Diagrama de sectors)

Hem tornat a preguntar als nostres estudiants sobre el tipus de pel·lícula que els agrada veure. Una altra forma d'organitzar-les de forma més fàcil de veure és el diagrama de sectors. Series capaç de recordar cap altre exemple? (Ajuda: passa cada quatre anys).



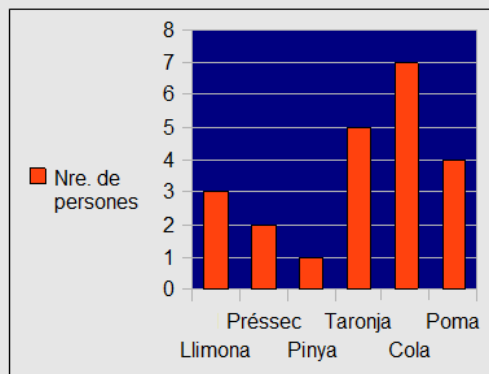
### Exercicis resolts

5. Les dades corresponen a les respostes donades per 22 persones escollides aleatòriament, sobre del sabor preferit en els refrescos d'una determinada marca.

Taronja, poma, cola, taronja, llimona, cola, préssec, cola, llimona, cola, cola, poma, llimona, taronja, cola, pinya, poma, taronja, cola, taronja, poma i préssec.

Dibuixa el diagrama de barres que representa les dades anteriors.

SOLUCIÓ:

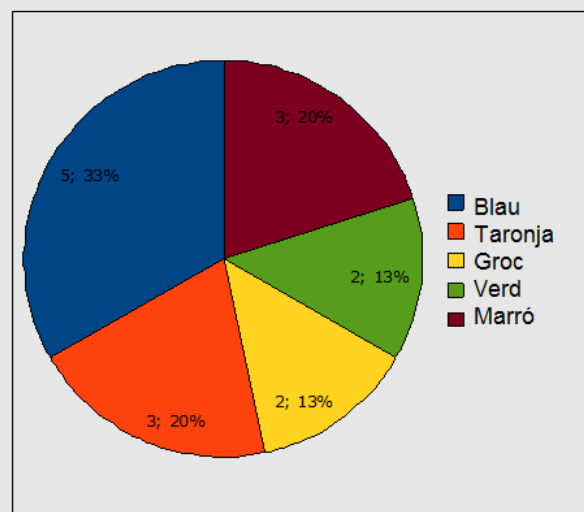


6. Els resultats corresponen a les contestacions realitzades per 15 estudiants sobre quin és el seu color preferit.

Les respostes que donaren són: blau, marró, taronja, groc, blau, taronja, verd, verd, blau, marró, blau, taronja, groc, marró, i blau.

Dibuixa el diagrama de sectors que representa les dades anteriors.

SOLUCIÓ:



## 5. Gràfics per una variable discreta

### Diagrama de barres

Es el gràfic estadístic més utilitzat para variables discretes. Para elaborar el diagrama, se situen en l'eix d'abscisses els valors corresponents de la variable.

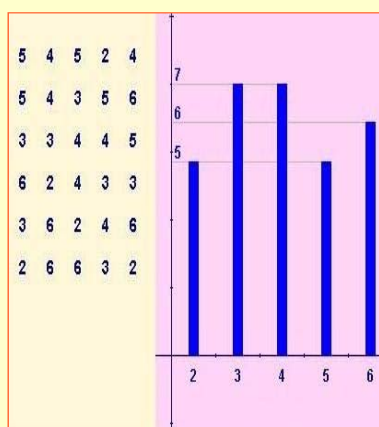
A partir de cada valor s'aixequen barres del mateix gruix i l'altura serà la corresponent a cada freqüència.

#### Exemple

##### Faltes d'ortografia

La professora ha anotat el nombre de faltes d'ortografia dels seus estudiants.

Vol una representació que li permeti veure les dades ràpidament, sabent quants estudiants cometen un nombre determinat d'errades ortogràfiques.



### Polígons de freqüències

El polígon de freqüències es construeix a partir del diagrama de barres, unint els punts mitjos de la base superior dels rectangles que constitueixen les barres.

Si es construeix un diagrama de barres considerant en lloc de les freqüències les freqüències acumulades i unim els punts mitjos de les bases superiors mitjançant segments, obtenim una poligonal creixent que denominem *polígon de freqüències acumulades*.

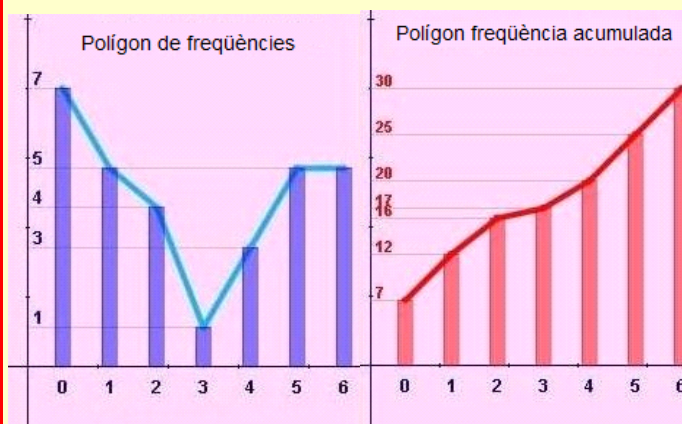
#### Exemple

##### Nombre de trucades

Una empresa de telecomunicacions vol fer un estudi sobre els seus clients, veient el nombre de trucades que rep un grup d'aquests.

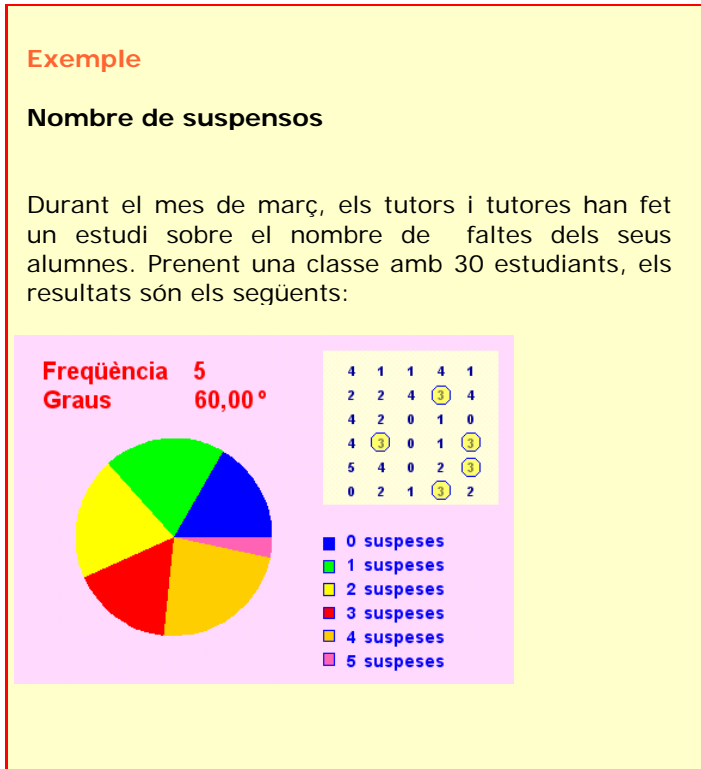
El estudi se realitza sobre 30 persones, anotant el nombre de trucades rebudes en un dia.

$X_i$	$f_i$	$F_i$
0	7	7
1	5	12
2	4	16
3	1	17
4	3	20
5	5	25
6	5	30



## Diagrama de sectors

Veiem primer un exemple.



Com complement a la tabulació i de vegades substituint a aquesta, en Estadística és molt habitual recórrer a gràfics que el seu efecte visual directe capta les primeres característiques d'una distribució estadística.

Per a variables quantitatives discretes, així com per les qualitatives, els gràfics que s'utilitzen amb més freqüència són:

1. El diagrama de barres
2. El diagrama de sectors

Encara que, depèn del tipus d'informació que vulguem obtenir, de vegades, resulta útil realitzar el polígon de freqüències, i el polígon de freqüències acumulades.

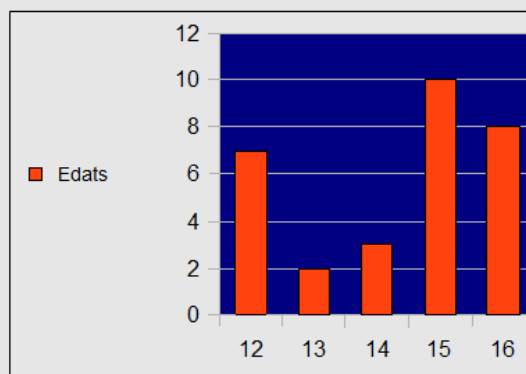
## Exercici resolt

7. Les edats de 30 estudiants d'un institut d'ensenyança secundària són les següents:

15 15 16 15 16 16 16 16 16 12 13 12 15 16 14  
12 14 12 15 13 14 16 15 15 12 15 12 15 15 12

Representa el diagrama de barres corresponent:

SOLUCIÓ:



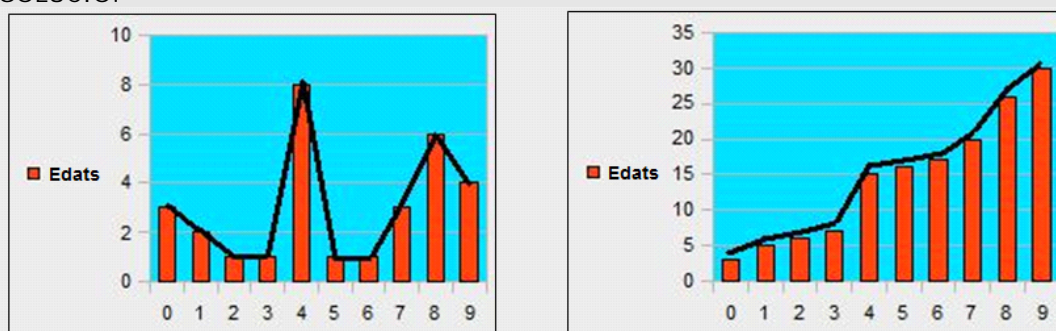
## Exercicis resolts

8. Los dades corresponen al nombre de trucades telefòniques que reben al dia 30 persones.

0 8 8 8 3 9 0 4 4 7 9 7 2 7 4  
4 9 1 4 1 4 5 6 4 9 8 1 8 4 8

Dibuixa el diagrama, els polígons de freqüència i de freqüència acumulades que representa les dades anteriors.

SOLUCIÓ:

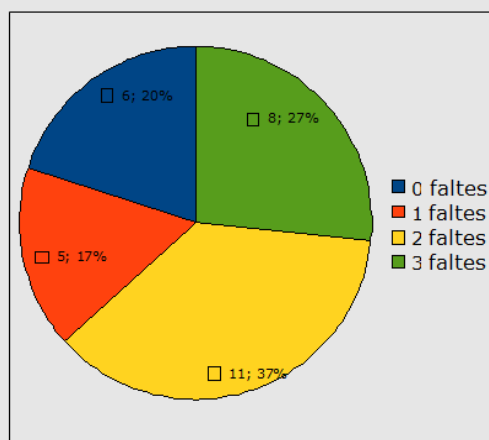


9. Les dades corresponen al nombre d'errades ortogràfiques en el mateix text de 30 estudiants.

2 2 2 1 1 2 3 2 0 0 3 2 1 0 3  
3 3 2 3 0 0 1 2 2 1 3 0 3 2 2

Representa el diagrama de sectors corresponent.

SOLUCIÓ:





## 6. Mesures de centralització

### Mitjana aritmètica

Als paràmetres o mesures estadístiques que informen sobre la tendència habitual o central de les dades d'una distribució se'ls denomina en estadística *mesures de tendència central*. La més utilitzada és la mitjana aritmètica.

La mitjana aritmètica es defineix com la suma de totes les dades dividida entre el nombre total d'aquestes. Com habitualment disposarem d'una taula de dades amb les seves freqüències, aplicarem:

$$\bar{x} = \frac{x_1 \cdot f_1 + x_2 \cdot f_2 + \dots + x_n \cdot f_n}{N}$$

1. La mitjana no té perquè ser un valor propi de la variable.
2. És molt sensible a valors extrems de les dades.
3. Es comporta de forma natural en relació a les operacions aritmètiques.

#### Exemple

##### Faltes d'assistència (Moltes dades)

Quan tenim moltes dades, per evitar realitzar un compte amb gran quantitat de nombres, primer organitzem una taula.

Veiem l'exemple en que es tenen anotades les faltes d'assistència d'un grup de 27 estudiants.

Hi ha 6 estudiants que han faltat 0 vegades, 4 que faltaren 1 vegada,...

$x_i$	$f_i$	$x_i \cdot f_i$
0	0	0
1	6	6
2	1	2
3	4	12
4	2	8
5	5	25
6	3	18
<b>21</b>	<b>71</b>	

Després de tabular les dades, construïm la columna corresponent als productes  $x_i \cdot f_i$

A l'última casella, calculem la suma total de la columna  $\sum_{i=1}^n x_i \cdot f_i$

La mitjana s'obté dividint  $\sum_{i=1}^n x_i \cdot f_i = 71$  entre el total de dades 21,

$$\text{mitjana} = \frac{71}{21} = 3,38$$

#### Exemple

##### Faltes d'assistència (Poques dades)

Las faltes d'assistència de 4 estudiants en un mes venen recollides pels següents valors: 0, 3, 2 i 1.

La mitjana aritmètica es calcula:

$$\frac{0+3+2+1}{4} = 1,5$$

### Mediana

La mediana és aquell valor de la variable estadística que deixa el 50% d'observacions inferiors a ell; així doncs, la mediana divideix en dues parts iguals la distribució estadística.

Dins de les propietats de la mediana es poden destacar:

1. Com a mesura descriptiva no es veu tan afectada com la mitjana per la presència de valors extrems.
2. És de càlcul ràpid i de fàcil interpretació.
3. Té propietats matemàtiques complicades que fan que s'utilitzi poc en inferència estadística.

#### Cas de poques dades i en nombre senar

En aquest cas es procedeix a ordenar les dades de menor a major, se considera el valor de la mediana el que correspon al lloc central.

## Exemple

### La mediana del nombre de suspensos (Moltes dades)

Entrem en una classe de 25 estudiants i preguntem el nombre de suspensos a l'última avaluació, hi ha 4 estudiants amb 0 suspensos, 2 amb 1 suspensos,...

Com tenim moltes dades, les organitzem en la següent taula per calcular la mediana.

Per trobar la mediana, en primer lloc calculem la meitat de les dades:

$$\frac{N}{2} = 12,50$$

Ara busquem en la columna de freqüències acumulades la primera vegada que se supera la meitat de les dades.

El valor corresponent de  $X_i$  és la mediana de la distribució estadística. En aquest cas:

$$Me = 1$$

$X_i$	$f_i$	$F_i$
0	9	9
1	4	13
2	2	15
3	5	20
4	5	25
	25	25

## Moda

Es defineix la moda com el valor de la variable estadística que té la freqüència absoluta més alta.

Si existeixen més d'un valor amb aquesta característica, llavors es diu que la distribució té diverses modes (*plurimodal*).

Aquesta mesura de centralització és sens dubte la de més fàcil càlcul. Se sol utilitzar com a complement de la mitjana aritmètica i mediana ja que per si sola no aporta una informació determinant de la distribució.

No és tan sensible com la mitjana aritmètica a valors extrems.

## Exemple

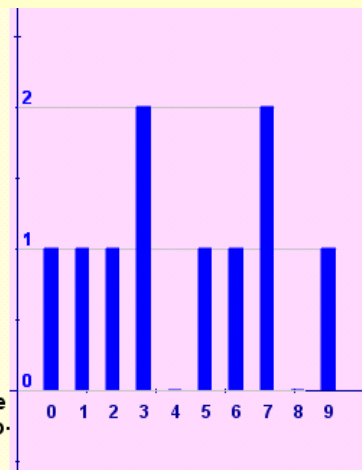
### Nombre de trucades

En un grup de 20 persones se recullen el nombre de trucades que realitzen durant un dia.

Resultant els següents valors: 0 persones fan 1 trucada, 3 persones fan 2 trucades, 2 persones fan 1 trucada...

$X_i$	$f_i$	$F_i$
0	1	1
1	1	2
2	1	3
3	2	5
4	0	5
5	1	6
6	1	7
7	2	9
8	0	9
9	1	10

La moda és el valor de major freqüència absoluta.



### Cas de poques dades i en nombre parell

En aquest cas es procedeix a ordenar de menor a major, es considera el valor de la mediana el corresponent a la semisuma dels dos llocs centrals.

### Exercicis resolts

10. Les edats d'un grup de 9 amigues són: 12, 14, 13, 16, 13, 15, 15, 17 i 13. Calcula la mitjana, mediana i moda.

SOLUCIÓ:

X = Edat	f	F	X·f
12	1	1	12
13	3	4	39
14	1	5	14
15	2	7	30
16	1	8	16
17	1	9	17
			128

Mitjana:  $\frac{128}{9} = 14,22$

Mediana 14 (si ordenem los dades, apareix en la posició 5).

Moda: 13 (apareix 3 vegades).

11. El nombre de trucades telefòniques que reben al dia els 9 integrants d'una família són:

7, 8, 15, 12, 13, 5, 10, 4, 8

Calcula la mitjana, mediana i moda.

SOLUCIÓ:

X = Nre. de trucades	f	F	X·f
4	1	1	4
5	1	2	5
7	1	3	7
8	2	5	16
10	1	6	10
12	1	7	12
13	1	8	13
15	1	9	15
			73

Mitjana:  $\frac{73}{9} = 8,11$

Mediana 8 (si ordenem los dades, apareix en la posició 5).

Moda: 8 (apareix 2 vegades).

La mitjana aritmètica ens informa que si tu t'has menjat dos pollastres i jo cap, de mitjana ens n'hem menjat un. Però és evident que tu has fet un bon àpat i jo he passat gana.

Les mesures de centralització aporten una informació, però és necessari completar-les amb altres mesures que denominarem de dispersió. Així doncs, a l'exemple anterior:

- El rang de valors va de 0 a 2 pollastres, hi ha una amplitud de  $2-0=2$ .
- Tu t'has menjat un pollastre més de la mitjana i jo un de menys. Tots dos ens hem desviat una unitat respecte de la mitjana.

## Exemple

$X_i$	$f_i$	$X_i \cdot f_i$	$ \bar{X} - X_i  \cdot f_i$
5	2	10	29,44
10	1	10	9,72
15	3	45	14,16
20	4	80	1,11
25	6	150	31,66
30	2	60	20,55
18	355	106,66	

$$\text{Rang} = 30 - 5 = 25$$

$$\text{DM} = \frac{106,66}{18} = 5,92$$

## 7. Mesures de dispersió

- Les mesures de **dispersió** indiquen si les dades estan més o menys agrupades respecte de les mesures de centralització. En particular respecte de la mitjana aritmètica.

### Rang

El rang o recorregut és la diferència entre el major i el menor valor de la variable. Indica la longitud de l'interval en el qual hi ha totes les dades.

### Desviación

La desviació respecte de la mitjana és la diferència en valor absolut entre cada valor de la variable estadística i la mitjana aritmètica.

$$D_i = |x_i - \bar{x}|$$

### Desviación media

És la mitjana aritmètica de les desviacions de les diferents dades.

$$D_{\bar{x}} = \frac{|x_1 - \bar{x}|f_1 + |x_2 - \bar{x}|f_2 + \dots + |x_n - \bar{x}|f_n}{N}$$

## EXERCICIS resoltos

12. Calcula el rang i la desviació mitjana de les dades:

8	8	6	10	9	6	7	8	9	7
7	6	6	7	9	5	5	7	10	7

**Solució:**

El rang oscil·la entre 5 i 10 amb una amplitud de 5.

Fem el recompte.

$$\text{La mitjana: } \bar{x} = \frac{147}{20} = 7,4$$

Calculem la desviació de cada dada respecte a la mitjana, en valor absolut. La mitjana de les desviacions:

$$\text{DM} = \frac{24,6}{20} = 1,23$$

$x_i$	$f_i$	$x_i \cdot f_i$	$ x_i - \bar{x}  \cdot f_i$
5	2	10	4,8
6	4	24	5,6
7	6	42	2,4
8	3	24	1,8
9	3	27	4,8
10	2	20	5,2
20	147	24,6	

13. Calcula la desviació mmitjana de les dades tabulades següents:

$x_i$	$f_i$	$x_i \cdot f_i$	$ x_i - \bar{x}  \cdot f_i$
100	7	700	2831,82
300	8	2400	1636,36
500	13	6500	59,09
700	9	6300	1759,09
900	7	6300	2768,18
Total:	44	22200	9054,55

**Solució:**

Calculem la mitjana:

$$\bar{x} = \frac{22200}{44} = 504,55$$

Van completar l'última columna:

$$\text{DM} = \frac{9054,55}{44} = 205,59$$

## Per practicar



### A. Classifica segons el caràcter de la variable les següents situacions:

#### 1. Situacions:

Cantant favorit
Longitud d'espàrrecs
Marca de refresc favorita
Tipus de música preferida
Raça de gossos
Nre. de dies assolats al mes

#### 2. Situacions:

Nre. de dies de vacances
Autor literari favorit
Nre. de germans
Nota mitjana selectivitat
Temperatura mitjana ciutat
Nre. dies faltes a classe

#### 3. Situacions:

Nre. dies plujosos al mes
Temps d'espera de l'autobús
Nre. faltes d'un dictat
Color d'ulls
Pel·lícules vistes al mes
Nota mitjana selectivitat

### B. Realitza una tabulació que inclogui la freqüència absoluta, relativa i les seves acumulades, quan calgui aproxima fins les centèsimes, de les dades que es corresponen amb les situacions següents:

4. El nombre de vegades que han canviat de domicili 23 persones.  
2, 2, 0, 2, 4, 2, 4, 4, 3, 4, 3, 3, 3, 3, 0, 1, 0, 4, 0, 3, 0, 3 i 5.

5. El nombre de germans que tenen 20 estudiants d'un centre.  
1, 4, 0, 2, 3, 1, 0, 3, 4, 1, 1, 3, 3, 3, 4, 1, 1, 2, 1 i 1.

6. El nombre de dormitoris de 28 vivendes d'una ciutat.  
3, 5, 0, 4, 2, 3, 0, 0, 1, 1, 3, 0, 2, 4, 1, 3, 3, 3, 1, 4, 4, 0, 3, 3, 1, 4, 3 i 1.

7. El nombre d'errades ortogràfiques en el mateix text de 30 estudiants són:  
0, 0, 2, 1, 4, 6, 6, 5, 0, 4, 6, 5, 5, 1, 0, 0, 3, 5, 1, 2, 5, 1, 0, 5, 2, 0, 4, 3, 6 i 4.

### C. Efectua una tabulació de les dades en la que apareguin les columnes de freqüències absolutes i relatives. Quan calgui aproxima fins les centèsimes.

8. El sabor preferit en els refrescos d'una determinada marca de 22 persones.  
Taronja, cola, taronja, llimona, cola, préssec, cola, llimona, cola, cola, poma, llimona, taronja, cola, pinya, cola, taronja, poma, taronja, cola, taronja i poma.
9. Las activitats realitzades per 20 estudiants en el seu temps lliure.  
Esport, amics, idiomes, música, idiomes, idiomes, amics, música, esports, ball, ball, música, esports, idiomes, cinema, amics, esports, amics, música, i cinema.

10. El tipus de programa de televisió que prefereixen veure en el seu temps lliure.

Ficció, infantils, esportius, espectacle, documentals, infantils, ficció, culturals, espectacle, infantils, ficció, esportius, esportius, espectacle, ficció, documentals, culturals, ficció, esportius i espectacle.

**D. Dibuixa el diagrama de barres corresponent a les situacions que apareixen.**

11. Preguntem a 25 estudiants escollits aleatòriament pel tipus de música que prefereixen escoltar.

Els resultats són: disco, disco, rock, clàssica, rock, llatina, pop, rock, pop, llatina, rock, flamenc, flamenc, llatina, flamenc, llatina, rock, clàssica, disco, disco, llatina, rock, disco, llatina i rock.

12. Les dades corresponen a les contestacions realitzades per 25 persones escollides aleatòriament, sobre el tipus de pel·lícula que prefereixen veure.

Les dades són les següents: comèdia, terror, suspens, comèdia, aventura, drama, aventura, aventura, comèdia, musical, terror, musical, suspens, aventura, comèdia, terror, musical, terror, terror, comèdia, suspens, suspens, comèdia, aventura i aventura.

13. Els resultats següents corresponen a les contestacions realitzades per 25 estudiants sobre les activitats realitzades en el seu temps lliure.

Esport, amics, amics, idiomes, música, idiomes, esport, música, idiomes, amics, música, esports, ball, música, ball, música, esports, idiomes, cinema, amics, esports, cinema, amics, música, i cinema.

14. Les edats de 30 estudiants d'un institut d'ensenyança secundària són les següents:

12, 13, 12, 15, 12, 15, 13, 14, 15, 12, 12, 12, 15, 15, 13, 14, 14, 16, 13, 12, 13, 14, 15, 16, 15, 13, 14, 15, 15 i 12.

15. Nombre d'assignatures suspenses de 30 estudiants són:

2, 0, 3, 2, 4, 0, 1, 3, 4, 2, 5, 0, 3, 2, 5, 4, 1, 1, 0, 1, 1, 1, 3, 4, 2, 0, 5, 5, 3 i 2.

16. El nombre de trucades telefòniques que reben un dia un grup de 20 amics són:

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

17. Per un estudi d'accessibilitat, durant 30 dies anotem el nombre de places lliures d'aparcament a les 5 de la tarda.

1, 1, 3, 5, 4, 0, 1, 3, 4, 2, 5, 0, 3, 2, 5, 4, 3, 1, 0, 1, 4, 1, 3, 4, 2, 3, 5, 4, 3 i 0.

**E. Dibuixa el diagrama de sectors corresponent a les situacions que apareixen en els exercicis D.11, D.12, D.16 i D.17**

**F. Realitza el polígon de freqüència i el de freqüència acumulada dels exercicis de l'apartat D.14 i D.15**



**G. Troba les mesures de centralització dels exercicis de l'apartat B.6 i B.7**

**H. Troba les mesures de dispersió dels exercicis de l'apartat B.6 i B.7**

## Per saber-ne més



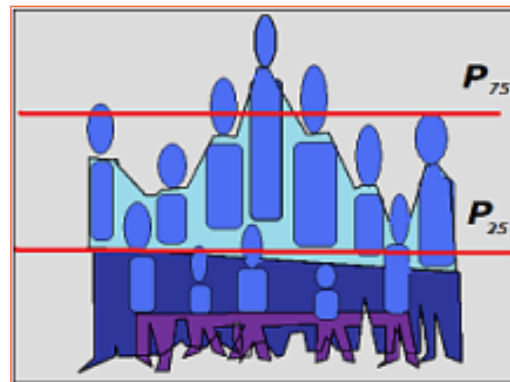
En estadística en moltes ocasions una variable discreta pren tal varietat de valors que per a que la tabulació sigui efectiva, ha de realitzar-se mitjançant intervals. La variable queda d'aquesta manera dividida en classes (interval, generalment de la mateixa amplitud). Aquesta és la tècnica que s'utilitza per a variables contínues.

Intervals	freqüència
...	...
...	...
$[L_i, L_{i+1})$	$f_i$
 VALORS MAJORS O IGUALS QUE $L_i$ I MENORS QUE $L_{i+1}$	 Nre. d'observacions

Com a representant de tot l'interval es considera el valor central (MARCA DE CLASSE)

A més del rang i la desviació mitjana altres mesures de dispersió que s'utilitzen normalment són:

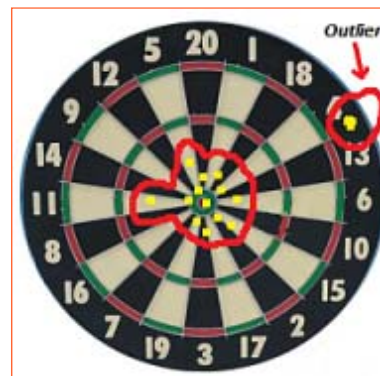
- Variància
- Desviació típica



El principal paràmetre estadístic i el més utilitzat és la mitjana aritmètica, tanmateix una característica important de la mitjana és que es veu molt afectada per valors extrems en la distribució.

La variació d'una simple dada en la distribució afecta a les mesures de tendència central encara que no de la mateixa forma. En estadística a un valor especialment anòmal se'l denomina **Outlier**. Decidir si en un estudi estadístic es depuren aquests valors extrems, és una de les primeres accions que ha de realitzar un investigador.

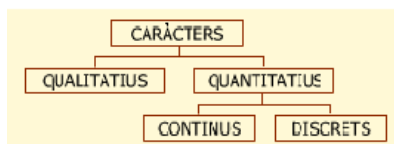
La *mediana* és la mesura de posició que més s'utilitza, tanmateix és molt habitual en la majoria dels estudis estadístics fer referència a altres mesures de posició com els **quartils**, **decils** o **centils**.





## Recorda el més important

Les primeres definicions necessàries per a l'inici de qualsevol estudi estadístic són: **Població, Individu, Mostra i Caràcter.**



$X_i$	$f_i$	$F_i$
0	8	8
1	7	15
2	5	20
3	16	36
4	9	45

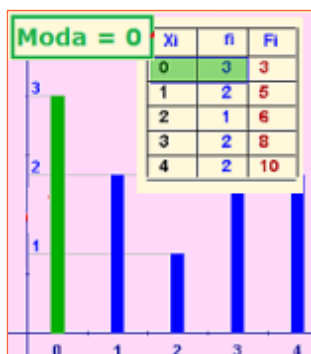
Annotations:  $Me = 3$  (Median),  $22,50$  (Cumulative frequency at  $X_i=2$ ),  $45$  (Total frequency),  $\frac{45}{2}$  (Formula for median calculation).

QUALITATIUS: *No expressables numèricament.*  
 QUANTITATIUS: *Es pot expressar mitjançant un nombre.*

De càlcul ràpid i de fàcil interpretació. No és tan sensible com la mitjana aritmètica a valors extrems.

### Mesures de tendència central

**Moda:** Valor que té la freqüència absoluta més alta.



És l'única que es pot calcular per variable qualitativa.

No és tan sensible com la mitjana aritmètica a valors extrems.

### Diagrama de barres



### Polígon de freqüències



**Mitjana aritmètica:** suma de totes les dades dividida entre el nombre total d'aquets.

$X_i$	$n$	$x_i \cdot n$
1	5	5
2	2	4
3	2	6
4	5	20
5	0	0
6	0	0
7	5	35
8	5	40
9	0	0
	<b>24</b>	<b>110</b>

Molt sensible a valors extrems en les dades

No ha de ser necessàriament un valor propi de la variable

$$\text{mitjana} = \frac{\sum_{i=1}^n x_i \cdot f_i}{\sum_{i=1}^n f_i} = \frac{110}{24} = 4,58$$

### Diagrama de sectors



**Mediana:** divideix en dues parts iguals la distribució estadística.



## Autoavaluació



1. Donades les dades: 7, 5, 7, 5, 6, i 8.  
Calcula la mitjana aritmètica amb dues xifres decimals.
2. La nota mitjana obtinguda en cinc exàmens ha estat 6,8. Si quatre de les notes han estat 4,7; 9,5; 8,3 i 9,2, quina és la cinquena?
3. La nota mitjana de quatre notes és 4,2. Si he tret ara un 8,0, quina nota mitjana tinc ara?
4. En una prova de gimnàstica la puntuació de cada atleta es calcula eliminant la pitjor i la millor nota dels jutges. Si les puntuacions obtingudes han estat: 8,1; 9,0; 9,3; 9,6; 8,2; 8,7 i 9,5, quina nota correspon?
5. Calcula la mediana d'aquestes dades: 9, 15, 19, 22, 31, 38 i 43.
6. Calcula la desviació mitjana d'aquestes dades: 18, 15, 28, 22 i 36.
7. En una distribució de 63 dades, la freqüència absoluta d'un valor de la variable és 21. Quants graus correspondrien a aquest valor en un diagrama de sectors?
8. Per obtenir la nota final de curs ens donen a elegir entre la mitjana, la mediana i la moda de les nou notes obtingudes. Quina triaries? Les notes són: 6, 3, 3, 4, 6, 8, 7, 9 i 3.
9. Calcula la mediana d'aquestes dades: 1, 17, 26, 5, 11 i 24.
10. Indica si la variable és discreta, contínua o qualitativa: Perímetre cranial.

## Solucions dels exercicis per practicar

**A.1** Qualitativa, contínua, qualitativa, qualitativa, qualitativa, discreta.

**A.2** Discreta, qualitativa, discreta, contínua, contínua, discreta.

**A.3** Discreta, contínua, discreta, qualitativa, discreta, contínua.

### B.4

Nre. canvis de domicili	f	h	F	H
0	5	0,17	5	0,17
1	1	0,03	6	0,21
2	4	0,14	10	0,34
3	7	0,24	17	0,59
4	5	0,17	22	0,76
5	1	0,03	23	0,79
	23			

### B.5

Nre. de germans	f	h	F	H
0	2	0,07	2	0,07
1	8	0,28	10	0,34
2	2	0,07	12	0,41
3	5	0,17	17	0,59
4	3	0,1	20	0,69
	20			

### B.6

Nre. de dormitoris	f	h	F	H
0	5	0,17	5	0,17
1	6	0,21	11	0,38
2	2	0,07	13	0,45
3	9	0,31	22	0,76
4	5	0,17	27	0,93
5	1	0,03	28	0,97
	28			

### B.7

Nre. de faltes d'ortografia	f	h	F	H
0	7	0,23	7	0,23
1	4	0,13	11	0,37
2	3	0,1	14	0,47
3	2	0,07	16	0,53
4	4	0,13	20	0,67
5	6	0,2	26	0,87
6	4	0,13	30	1
	30			

Activitats temps lliure	f	h
Esport	4	0,14
Amics	4	0,14
Idiomes	4	0,14
Ball	2	0,07
Cine	2	0,07
Música	4	0,14
	20	

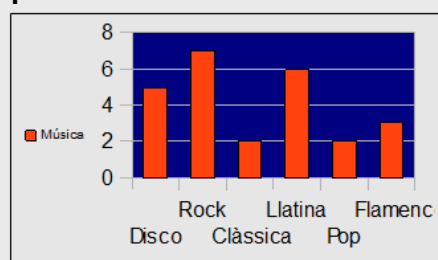
### C.9

Sabor preferit	f	h
Taronja	6	0,21
Cola	8	0,28
Llmona	3	0,1
Pinya	1	0,03
Préssec	1	0,03
Poma	3	0,1
	22	

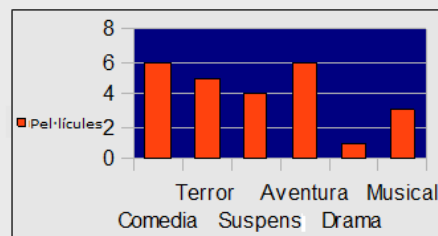
### C.10

Programa de televisió	f	h
Ficció	5	0,17
Infantils	3	0,1
Esportius	4	0,14
Espectacle	4	0,14
Documentals	2	0,07
Culturals	2	0,07
	20	

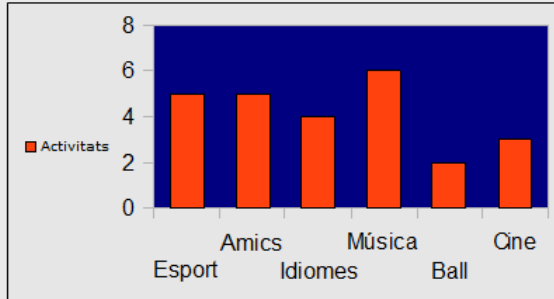
### D.11



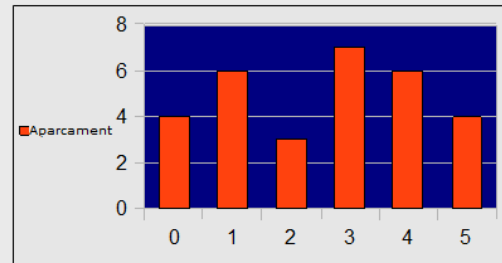
### D.12



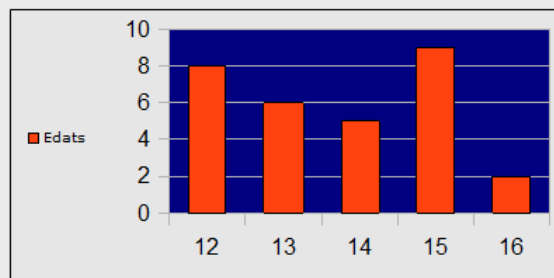
D.13



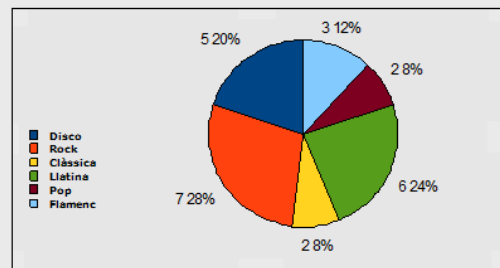
D.17



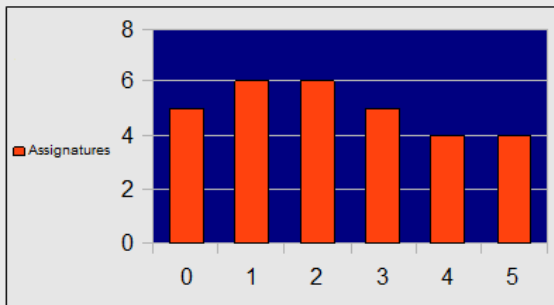
D.14



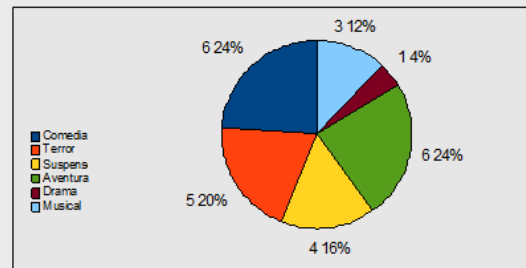
E.11



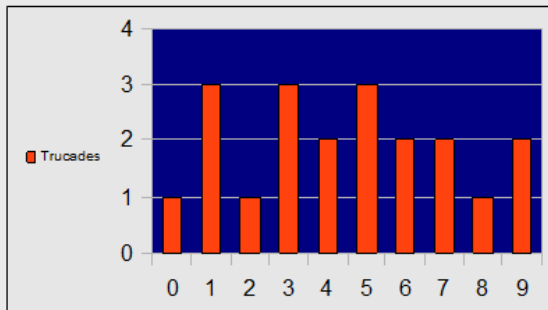
D.15



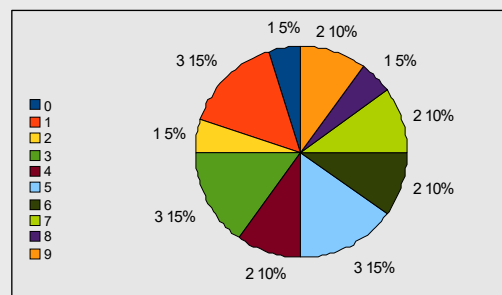
E.12



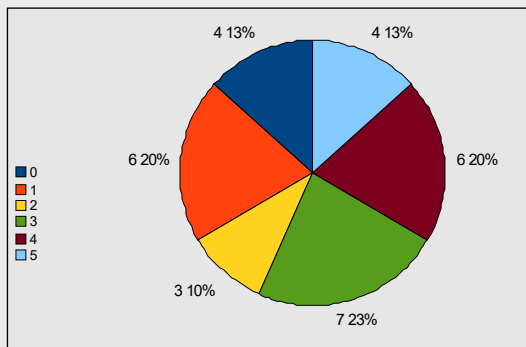
D.16



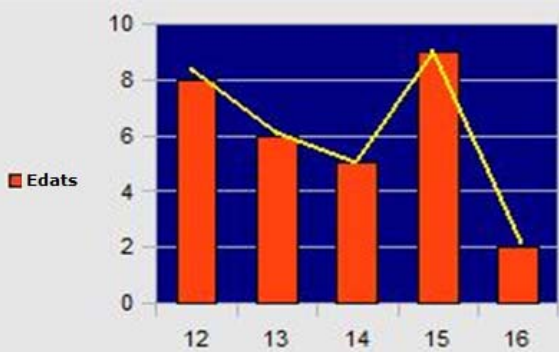
E.16



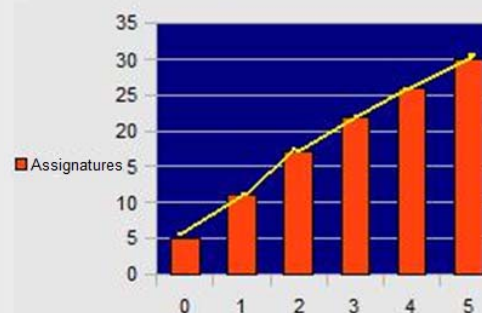
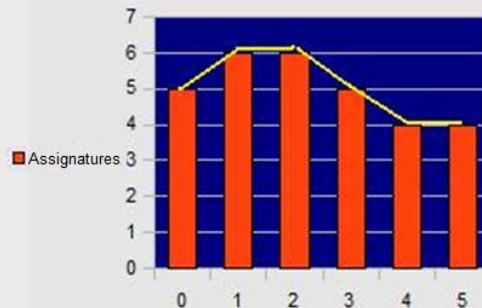
**E.17**



**G.14**



**G.15**



**G.6**

Mitjana = 2,21  
 Mediana = 3  
 Moda = 3

**G.7**

Mitjana = 2,86  
 Mediana = 3  
 Moda = 0

**H.6**

Rang = 5  
 Desviació mitjana = 1,34

**H.7**

Rang = 6  
 Desviació mitjana = 2,01

## Solucions AUTOAVALUACIÓ

1. 6'33
2. 2'3
3. 5'8
4. 8'94
5. 22
6. 6,56
7. 120°
8. La mediana
9. 14
10. Qualitativa