

12

Estadística

Continguts

1. Estadística descriptiva
 - Població i mostra
 - Variabls estadístiques
 - Gràfics variables qualitatives
 - Gràfics variables quantitatives discretes
 - Gràfics variables quantitatives contínues
2. Mesures de centralització
 - Mitjana, moda i mediana
 - Evolució de la mitjana
 - Evolució de la mediana
 - Mitjana i mediana comparades
 - Mesures de posició
3. Mesures de dispersió
 - Desviació típica i recorregut
 - Càlcul de les mesures de dispersió
 - La mitjana i la desviació típica
4. Representativitat de les mostres
 - Mostreig estratificat
 - Mostreig aleatori. Biaix
5. Estadística bidimensional
 - Distribucions bidimensionals
 - Correlació lineal
 - Rectes de regressió

Objectius

- Distingir els conceptes de població i mostra.
- Diferenciar els tres tipus de variables estadístiques.
- Fer recomptes i gràfics.
- Calcular i interpretar les mesures estadístiques de centralització més importants.
- Calcular les principals mesures de dispersió.
- Entendre la importància de l'elecció de la mostra perquè sigui representativa.
- Calcular el centre de gravetat, la covariància, el coeficient de correlació i la recta de regressió en una distribució bidimensional.



Abans de començar

Un joc per començar

Ves fent clic, durant una estona, en peces adossades al buit per desplaçar-les i desfer el puzzle. Després el reconstrueixes.



Clica per anar a la pàgina següent.

1. Estadística descriptiva

1.a. Població i mostra.

Població és _____ sobre el qual és fa un estudi estadístic.

La **mostra** és _____, per això la propietat més important de les mostres és la seva _____.

El procés seguit en l'extracció de la mostra s'anomena _____.

En l'escena adjunta tenim 625 petits quadrats que representen als alumnes d'un institut fictici. Si vas fent clic en els quadrats, aniràs seleccionant part dels alumnes.

Contesta:

- a. Quina és la població? _____
- b. Quina és la mostra? _____
- c. Com s'anomena el procés en el qual es pregunta a tota la població? _____

Clica per anar a la pàgina següent.

1.b. Variables estadístiques.

La característica que s'estudia en una població és la variable estadística.

Completa la taula següent amb les característiques dels diferents tipus de variables estadístiques:

Tipus de variables estadístiques		
Qualitatives		
	Disretes	Contínues
Quantitatives		

A l'escena de la dreta tens exemples de cada tipus de variable estadística.

Clica en el botó per fer un exercici.

Completa la taula amb els exemples:

Qualitatives	Quantitatives Disretes	Quantitatives Contínues

Clica per anar a la pàgina següent.


1.c. Gràfics en variables qualitatives

El **diagrama de sectors** és el més indicat per aquest tipus d'informació. El percentatge de dades de cada valor en una mostra es correspon amb el mateix percentatge de sector d'un cercle.

$$\frac{\text{freqüència}}{\text{freqüència total}} = \frac{\text{angle}}{360^\circ}$$

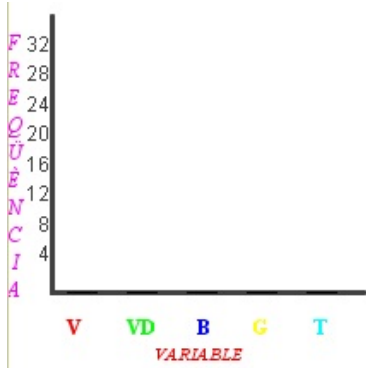
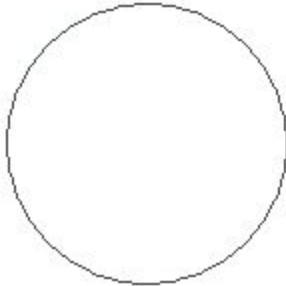
Així, per exemple, si les dades són A, A, A, A, A, B, B, B, C, C, completa la taula amb les dades corresponents:


x_i	Freqüència	Percentatge	Angle
A			
B			
C			

Fes clic a  per veure un vídeo sobre gràfics.

Amb l'ajuda de l'escena de la dreta pots fer un exercici sobre representació gràfica de variables estadístiques qualitatives.

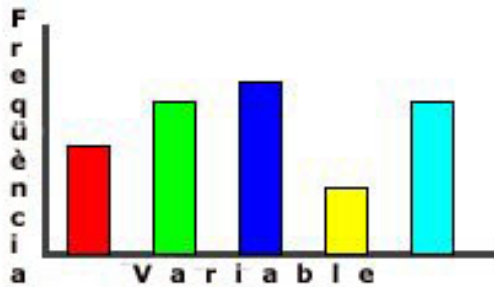
L'exercici simula que tenim una població de 30 alumnes i cada un d'ells tria un color. Clicant a **Genera** tindrem els 30 colors escollits aleatòriament, clica **ajuda** i llegeix com l'escena et facilita el recompte i completa la taula, comprovant que és correcte el teu recompte. A continuació clica el botó de diagrames per veure els gràfics, i dibuixa'ls en el lloc corresponent.

Color _i	Freqüència	D. de columnes	D. de sectors
Vermell			<p>D. de sectors</p> 
Verd			
Blau			
Groc			
Turquesa			

Clica  per anar a la pàgina següent.

1.d. Gràfics en variables quantitatives discretes

Diagrama de barres. N'hi haurà prou que observis exemples fets de l'escena de la dreta per comprendre com es fan i el seu significat. Aquest és el gràfic més indicat per a les variables quantitatives discretes.




Pots llegir un **article** de l'Institut Nacional d'Estadística, sobre el comportament o actuacions del nostre país amb el medi ambient i les energies renovables, en el qual es mostren diversos tipus de diagrames.

Amb l'ajuda de l'escena de la dreta pots fer uns exercicis sobre representació gràfica de variables estadístiques quantitatives discretes. A la taula següent copia'n un d'ells.

L'exercici simula que tenim una població de 30 alumnes i cada un d'ells ens diu el nombre de germans que té. Clicant a **Genera** tindrem les 30 dades generades aleatòriament, clica **ajuda** i llegeix com l'escena et facilita el recompte i completa la taula, comprovant que és correcte el teu recompte. A continuació clica el botó de diagrames per veure els gràfics, i dibuixa'ls en el lloc corresponent.

Variable	Freqüència	D. de columnes	D. de sectors
0			
1			
2			
3			
4			

Clica  per anar a la pàgina següent.

1.e. Gràfics en variables quantitatives contínues

Histograma.

Llegeix l'explicació d'aquest tipus de gràfic estadístic.

Contesta.

	RESPOSTA
Quina figura s'utilitza per representar les dades?	
Si tots els intervals són de la mateixa amplitud, què ens indica l'altura?	
Si tots els intervals <u>no</u> són de la mateixa amplitud, quina magnitud és proporcional a la freqüència?	

Clica a l'enllaç: **Exemple**. Fixa't en l'exemple resolt que apareix.

Polígon de freqüències. Unirem els centres de la part superior de tots els rectangles per obtenir-lo.

També es pot dibuixar l'histograma de les **freqüències acumulades**: en cada dada s'acumula la freqüència de les dades anteriors.

Amb l'ajuda de l'escena de la dreta pots fer uns exercicis sobre representació gràfica de variables estadístiques quantitatives contínues.

En la taula següent copia'n un d'ells:

L'exercici simula que tenim una població de 30 alumnes i mesurem l'altura de cada un d'ells. Clicant a **Clica per començar** tindrem les 30 dades generades aleatòriament, clica **ajuda** i llegeix com l'escena et facilita el recompte i completa la taula, comprovant que és correcte el teu recompte. A continuació clica el botó de diagrames per veure els gràfics, i dibuixa'ls en el lloc corresponent.

Interval	Freqüència	Histograma	D. de freq. acumulades
[150, 160)			
[160, 170)			
[170, 180)			
[180, 190)			
[190, 200)			

Clica per anar a la pàgina següent.

EXERCICIS

1. Classifica els següents exemples de variables estadístiques: longitud d'un camió, càrrega màxima, nre. de rodes, nre. d'eixos, tipus de camió, marques de neumàtics, tipus de tapisseria, nre. de portes, altura màxima.

Qualitatives:

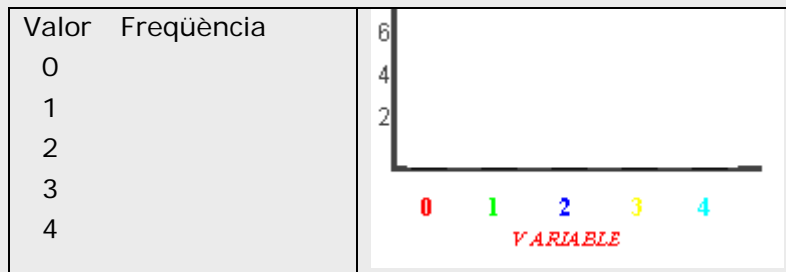
Q. discretes:

Q. contínues:

2. Calcula els graus que corresponen a cada valor en un gràfic de sectors fet a partir de les dades: R, R, V, V, V, V, V, A, A, A

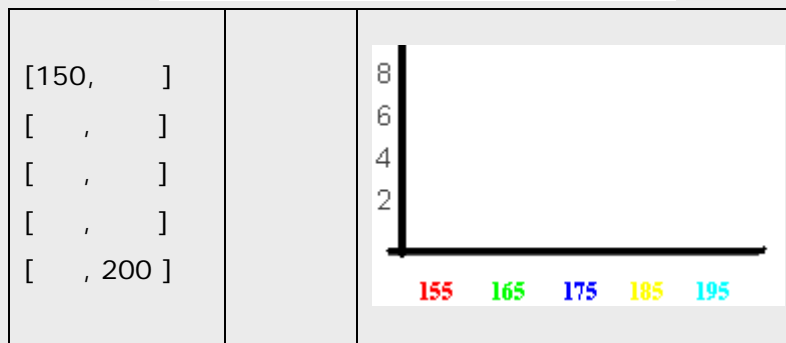
3. Agrupa les dades següents i fes un diagrama de barres adequat.

Dades = { 0 1 0 2 3 4 1 2 2 1 2 2 3 4 3 2 1 3 }



4. Classifica les dades en intervals i dibuixa un histograma adequat.

180 197 154 181 189 162 152 162 167 190
 189 160 166 197 187 194 152 181 173 154
 177 184 186 174 177 159 158 189 160 150



Clica per anar a la pàgina següent.

2. Mesures de centralització

2.a. Mitjana, mediana i moda

Un conjunt N d'observacions, N nombres, pot ser que per si sol no ens digui res. En canvi, si a més ens diuen que estan situats al voltant d'un o diversos valors centrals, ja tenim una referència que sintetitza la informació. Per això, es defineixen els següents **paràmetres de centralització** (perquè ens indiquen el **centre** de la distribució)

Mitjana: _____



Moda: _____

En el cas d'una variable contínua, considerarem per moda _____.

També pot passar que hi hagi dues modes o que no n'hi hagi cap que destaquí.

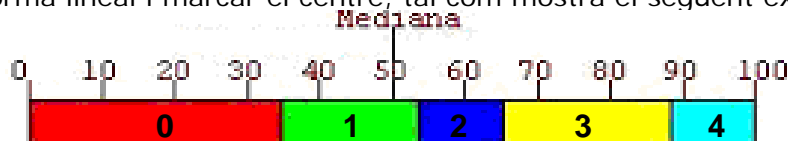
Mediana: _____

A l'escena de la dreta veiem exemples de com calcular aquests paràmetres. Copia a continuació dos dels exemples:

<i>Dades</i>	<i>Mitjana</i>	<i>Moda</i>	<i>Mediana</i>

En el cas de la mediana, si hi ha poques dades, el millor és procedir segons l'exemple de l'escena, segons sigui una quantitat parell o senar.

Per calcular la mediana si la quantitat d'observacions és gran, s'hauran d'agrupar les dades primer en una taula. Després dibuixar segments de longitud proporcional a la seva freqüència, i disposar-los de forma lineal i marcar el centre, tal com mostra el següent exemple:



Clica en el botó



per veure un exercici resolt.

Clica



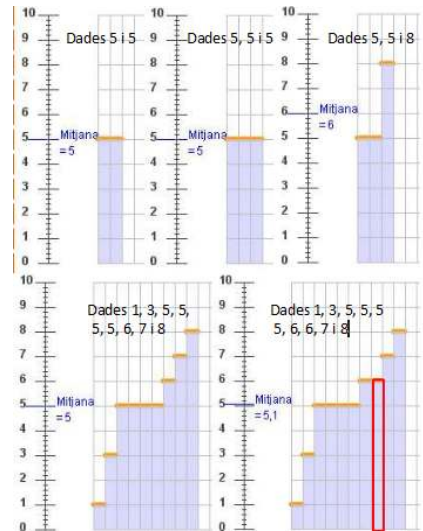
per anar a la pàgina següent.

2.b Evolució de la mitjana.

1 Per a les dades 5 i 5 la mitjana és ____.
 Si afegim un 5 _____.
 Si afegim un 8 _____.

2 Si tenim 9 dades amb mitjana 5,
 necessitem afegir un 6 perquè la mitjana passi a ser ____
 Si tenim 19 dades amb mitjana 5,
 necessitem una dada de valor 7 perquè la mitjana passi a ____

3 Per a un conjunt de dades amb mitjana 5,
 si n'afegim un altre amb mitjana 5, per exemple 6 i 4,



A l'escena de la dreta de la pàgina pots comprovar com es modifica la mitjana en diversos exemples.

Tria el número de l'exemple:



A continuació pots modificar el nombre de vegades que apareix una dada clicant a




i observa com varia en cada cas la mitjana.

Clica en el botó  per fer uns exercicis.

En aquests exercicis has de calcular la mitjana, pots triar si la variable és discreta o contínua. El recompte ja surt fet. Fes-ne uns quants i a continuació copia un exercici de cada tipus a la taula següent:

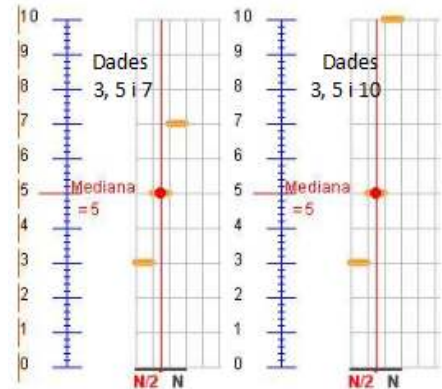
Variable quantitativa discreta			Variable quantitativa contínua		
Marca	Freqüència		Marca	Freqüència	
x_i	f_i	$x_i \cdot f_i$	x_i	f_i	$x_i \cdot f_i$
Total			Total		
Mitjana \bar{x}			Mitjana \bar{x}		

Clica  per anar a la pàgina següent.

2.c Evolució de la mediana

1 La mediana, per a les dades 2, 3 i 4 és $Me = \underline{\hspace{2cm}}$.
 Si canviem el 4 per 5 o per 6 o per qualsevol altre valor més gran, $\underline{\hspace{2cm}}$

2 Si afegim una altra dada i tenim 2,3,4 i 4, per exemple, la $Me = \underline{\hspace{2cm}}$
 I si afegim un cinquè valor, un 4 o un 5 o un 6 o qualsevol altre major que 4, la mediana en 2,3, 4, 4 i ?? passa a ser $\underline{\hspace{2cm}}$
 És igual que el valor ?? sigui 5, 10 o 25.



Per a veure la mediana es traça una vertical des de l'eix horitzontal en $N/2$

A l'escena de la dreta tens exemples on la mediana canvia i on no. A més, tu mateix pots variar el valor o valors que vulguis per observar com evoluciona. També tens la possibilitat de realitzar exercicis de càlcul de la mediana, a la mateixa escena.

Clicant en els botons **Nombre parell de dades** i **Nombre senar de dades** obtens exemples de dades i de com calcular la mediana. Si cliques **canviar** pots veure com calcular la mediana:

Tria el número de l'exemple:



A continuació pots modificar el nombre de vegades que apareix una dada clicant sobre



i observar com varia en cada cas la mediana.

Clica en el botó



per fer uns exercicis.

En aquests exercicis has de calcular la mediana. Pots triar si la variable és discreta o contínua i ja apareix fet el recompte. Fes-ne uns quants i a continuació copia un exercici de cada tipus. Pots consultar l'ajuda per resoldre'ls

Variable quantitativa discreta			Variable quantitativa contínua		
Marca	Freqüència		Marca	Freqüència	
x_i	f_i	F_i acumulada	x_i	f_i	F_i acumulada
<input style="width: 100%; height: 20px;" type="text"/>			<input style="width: 100%; height: 20px;" type="text"/>		

Clica per anar a la pàgina següent.

2.d Mitjana i mediana comparades

Llegeix el text i completa els valors de la mitjana i la mediana en cada cas:

Dades	Mitjana	Mediana
4, 6		
4, 6, 8		
4, 6, 11		
Els valors 8 i 11 es consideren observacions _____.		
Si les dades estiguessin repartides _____ respecte a un valor, aquest valor seria _____.		
Si els valors a un costat de la mediana n'estan més allunyats que els de l'altre costat, la mitjana _____.		
_____ . Hi ha una _____.		

Per exemple, si tenim les observacions:

- 20, 24 i 28 **Me=24**
- per 20, 24, 28 i 30
Me=(24+28)/2=26
- per 20, 24, 28 i 100
Me=(24+28)/2=26

En canvi, la mitjana no es comporta de la mateixa forma amb les mateixes dades:

- $\bar{x}=24$
- $\bar{x}=25,5$
- $\bar{x}=43$

Juga amb l'escena de la dreta.

Hi ha tres grups d'exemples: simètrics, asimètrics i atípics.

Pots observar l'evolució de la mediana i la mitjana

Tria el número de l'exemple:

Simètrics			Asimètrics						Atípics		
1	2	3	4	5	6	7	8	9	10	11	12

Si vols

pots modificar el nombre de vegades que apareix una dada clicant a sobre



Clica per anar a la pàgina següent.

2.e Mesures de posició: quartils i percentils

Donat un conjunt de dades numèriques a més de la mediana podem considerar altres mesures de posició.

- El primer valor que supera el 25% és el _____ **Q₁**
- El primer valor que supera el 75% és el _____ **Q₃**
- Per altres valors com el 10%, o el 80% parlem de _____ **P₁₀ i P₈₀**.

A l'escena de la dreta tens un exemple resolt, si cliques la fletxa i a sobre del botó **genera** pots obtenir molts exemples resolts, triant si vols que la variable sigui discreta o contínua.

Clica en el botó



per practicar el càlcul de les mesures de posició.

Clicant el botó **genera** obtens noves dades, i amb el botó **Discreta/Contínua** canvies el tipus de dades. Copia dos exercicis a la taula següent:

Variable quantitativa discreta			Variable quantitativa contínua		
Marca	Freqüència	Mediana	Marca	Freqüència	Mediana
x_i	f_i		x_i	f_i	
		Quartil Q_1			Quartil Q_1
		Quartil Q_3			Quartil Q_3
		Percentil			Percentil
Total			Total		

EXERCICIS

5. Calcula la mitjana en cada cas:

- a) 4, 6, 8
- b) 4, 6, 8, 6
- c) 100, 120, 180, 200

6. Calcula la mitjana en cada cas:

<p>a)</p> <table border="1" style="border-collapse: collapse; text-align: center;"> <tr><th>Marca</th><th>Fr</th></tr> <tr><td>10</td><td>2</td></tr> <tr><td>20</td><td>4</td></tr> <tr><td>30</td><td>3</td></tr> <tr><td>40</td><td>2</td></tr> </table>	Marca	Fr	10	2	20	4	30	3	40	2	<p>b)</p> <table border="1" style="border-collapse: collapse; text-align: center;"> <tr><th>Marca</th><th>Fr</th></tr> <tr><td>100</td><td>2</td></tr> <tr><td>200</td><td>4</td></tr> <tr><td>300</td><td>3</td></tr> <tr><td>400</td><td>2</td></tr> </table>	Marca	Fr	100	2	200	4	300	3	400	2
Marca	Fr																				
10	2																				
20	4																				
30	3																				
40	2																				
Marca	Fr																				
100	2																				
200	4																				
300	3																				
400	2																				

7. Determina la moda i la mediana:

- a) 5,6,6
- b) 1,1,2,3
- c) 1,2,3,4,2
- d) 3,2,3,2,2,2

8. Calcula la moda i la mediana en cada cas:

<p>a)</p> <table border="1" style="border-collapse: collapse; text-align: center;"> <tr><th>Marca</th><th>Fr</th></tr> <tr><td>10</td><td>2</td></tr> <tr><td>20</td><td>4</td></tr> <tr><td>30</td><td>3</td></tr> <tr><td>40</td><td>2</td></tr> </table>	Marca	Fr	10	2	20	4	30	3	40	2	<p>b)</p> <table border="1" style="border-collapse: collapse; text-align: center;"> <tr><th>Marca</th><th>Fr</th></tr> <tr><td>100</td><td>2</td></tr> <tr><td>200</td><td>3</td></tr> <tr><td>300</td><td>4</td></tr> <tr><td>400</td><td>1</td></tr> </table>	Marca	Fr	100	2	200	3	300	4	400	1
Marca	Fr																				
10	2																				
20	4																				
30	3																				
40	2																				
Marca	Fr																				
100	2																				
200	3																				
300	4																				
400	1																				

9. Calcula la mediana, quartils primer i tercer, i els percentils 30 60 i 90 de les dades:

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

Clica per anar a la pàgina següent.

3. Mesures de dispersió

3.a Variància, desviació típica i rang

“L'estadística és una ciència segons la qual, si jo em menjo un pollastre i tu no te'n menges cap, de mitjana, ens hem menjat mig pollastre cada un”.

L'estadística ens dirà que tots mengen el mateix quan les mesures de dispersió siguin totes nul·les.

Rang:

L'interval definit per _____.

També s'anomena rang a _____.

Variància:

La mitjana aritmètica dels _____

Si cliques a l'enllaç [Fórmules](#), s'obre una finestra en la qual pots veure les dues fórmules que ens permeten calcular la variància i de com són equivalents entre sí.

Escriu en els quadres aquestes dues fórmules:

Desviació típica:

Com més grans són la variància o la desviació típica, les dades es separen més de la mitjana, és a dir, hi ha més dispersió.

Si cliques a l'enllaç [Càlcul en diferents exemples](#) pots generar exemples de variables discretes o contínues en els quals veuràs dos mètodes diferents pel càlcul de la variància.

Quin és el mètode més adient per al càlcul? _____ Per què?


A l'escena de la dreta tens diversos exemples de les mesures de dispersió i del seu significat, llegeix-los amb atenció.

Clica en el botó



per comparar distribucions amb mesures de centralització iguals,

en les quals canvia la desviació típica. Copia'n a continuació dues d'elles:

Clica  per anar a la pàgina següent.

3.b Càlcul de les mesures de dispersió.

Recorregut

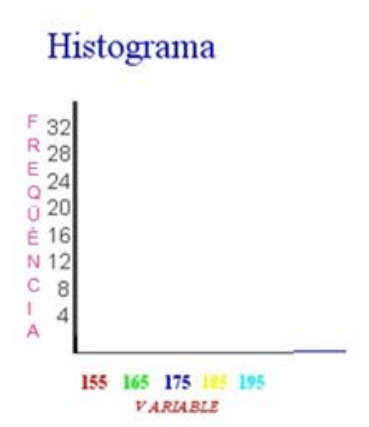
Clicant en el botó **genera** obtens noves dades, i amb el botó **Discreta/Contínua** canvies el tipus de dades. Copia aquí dos exercicis de cada tipus:

Variable estadística discreta				Variable estadística contínua			
Màxim		Màxim		Màxim		Màxim	
Mínim		Mínim		Mínim		Mínim	
Recorregut		Recorregut		Recorregut		Recorregut	

Desviació típica

A l'escena de la dreta pots generar unes dades, calcular la desviació típica i veure l'histograma. Copia a continuació dos exercicis

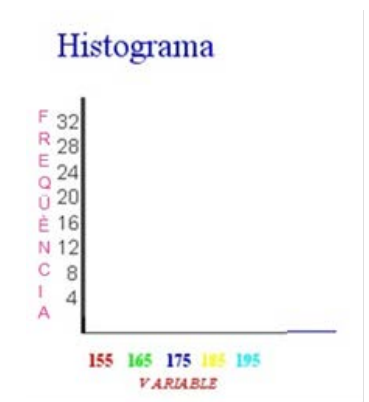
Interval	Marca	Frequència		
	x_i	f_i	$x_i \cdot f_i$	$f_i \cdot (\bar{x} - x_i)^2$
Total				



Mitjana Desviació típica

Mínim Màxim Recorregut

Interval	Marca	Frequència		
	x_i	f_i	$x_i \cdot f_i$	$f_i \cdot (\bar{x} - x_i)^2$
Total				



Mitjana Desviació típica

Mínim Màxim Recorregut

Clica en el botó  per fer uns exercicis.


Clicant en el botó **Genera** obtens noves dades, i amb el botó **Discreta/Contínua** canvies el tipus de dades. Copia aquí dos exercicis de cada tipus:

Variable discreta					Variable discreta				
Marca	Freqüència				Marca	Freqüència			
x_i	f_i	$x_i \cdot f_i$	$f_i \cdot (\bar{x} - x_i)^2$	$f_i \cdot x_i^2$	x_i	f_i	$x_i \cdot f_i$	$f_i \cdot (\bar{x} - x_i)^2$	$f_i \cdot x_i^2$
Total					Total				

Mitjana Desviació típica Mitjana Desviació típica

Variable contínua					Variable contínua				
Marca	Freqüència				Marca	Freqüència			
x_i	f_i	$x_i \cdot f_i$	$f_i \cdot (\bar{x} - x_i)^2$	$f_i \cdot x_i^2$	x_i	f_i	$x_i \cdot f_i$	$f_i \cdot (\bar{x} - x_i)^2$	$f_i \cdot x_i^2$
Total					Total				

Mitjana Desviació típica Mitjana Desviació típica

Clica  per anar a la pàgina següent.

3.c Mitjana i desviació típica.

Per a mostres unimodals (una sola moda) i gairebé simètriques, al voltant de la mitjana podem considerar un interval que contingui la majoria de les dades.

Per exemple, per a una mostra amb mitjana 100 i desviació típica 10, la major part de les dades estaran entre 90 i 110, aproximadament el 68%; entre 80 i 120 hi estaran el 95% aproximadament. I gairebé totes entre 70 i 130.

Hi ha una forma de distribució de dades anomenada **normal** que compleix amb l'anterior, i que, de tota manera, de totes les poblacions grans es poden extreure dades que s'hi ajusten. En cursos superiors veuràs la importància d'aquestes distribucions.

A l'escena de la dreta tens uns exemples on apareix la mitjana i unes franges de color al seu voltant. Tria el número de l'exemple:

A continuació pots modificar el nombre de vegades que apareix una dada clicant sobre i observa com varia en cada cas la mitjana i les franges del seu voltant.

Clica en el botó per fer uns exercicis.

Clicant en el botó **genera** obtens noves dades. A continuació, escriu-ne a les taules dos d'ells, i després comprova el resultat a l'escena:

Marca	Freqüència				
x_i	f_i	$x_i \cdot f_i$	$f_i \cdot (\bar{x} - x_i)^2$	$f_i \cdot x_i^2$	Mitjana
					Desviació típica
					$[\bar{x} + \sigma, \bar{x} - \sigma] = [\quad , \quad]$
					Nre. de dades
					$[\bar{x} + 2\sigma, \bar{x} - 2\sigma] = [\quad , \quad]$
					Nre. de dades
Total					

Marca	Freqüència				
x_i	f_i	$x_i \cdot f_i$	$f_i \cdot (\bar{x} - x_i)^2$	$f_i \cdot x_i^2$	Mitjana
					Desviació típica
					$[\bar{x} + \sigma, \bar{x} - \sigma] = [\quad , \quad]$
					Nre. de dades
					$[\bar{x} + 2\sigma, \bar{x} - 2\sigma] = [\quad , \quad]$
					Nre. de dades
Total					


Clica per anar a la pàgina següent.

EXERCICIS

10. Calcula la mitjana i la desviació típica de:
- a) 200, 250
 - b) 175, 275
 - c) 250, 250
11. Calcula la mitjana i la desviació típica de:
- a) 7, 5, 3, 2, 4, 5
 - b) 20, 25, 20, 22, 21
12. Organitza les dades següents en intervals de 10 cm des de 150 a 200. Amplia la taula amb dues columnes, una per al producte de las marques amb les freqüències i l'altra per al producte de les freqüències amb els quadrats de les diferències amb la mitjana. Calcula la mitjana i la desviació típica.

174 158 150 185 186 178 166 185 199
183 175 173 175 164 173 178 179 164
176 159 190 173 189 163 156 169

Interval	Marca	Freqüència			
	x_i	f_i	$x_i \cdot f_i$	$f_i \cdot (\bar{x} - x_i)^2$	
					Mitjana=
					Desviació típica=
	Total				

Clica  per anar a la pàgina següent.

4. Representativitat

4.a Representativitat. Mostreig estratificat

Una mostra és **representativa** de la població quan _____

De què depèn que l'estudi d'una població sigui o no representatiu? _____

Per exemple, si volem estudiar el poder adquisitiu d'una població, i només triem individus d'una determinada zona, o principalment d'una determinada zona, com serà la mostra?

Si hi ha tres zones amb 12.000, 18.000 i 20.000 habitants, escriu amb quin percentatge hem de escollir els individus de cada zona per elaborar una mostra representativa. _____

Un **mostreig estratificat** és _____

A l'escena tens 625 quadrats que representen els alumnes d'un institut fictici. Seguint les instruccions pots observar la diferència entre un mostreig representatiu i un altre que no ho és.



Si comparem els gràfics en ambdós exemples de mostra, en quin tipus de mostra s'assemblen més a les de la població total? _____

Per què? _____

Clica  per anar a la pàgina següent.


4.b Biaix. Mostreig aleatori

Quan diem que la mostra és esbiaixada? _____

Explica en què consisteix un mostreig aleatori total: _____

A l'escena pots animar una elecció totalment aleatòria o realitzar mostrejors simulant enquestes en fer clic.



Clica en el botó  per fer un exercici sobre representativitat.

Copia en aquest quadern un exercici i comprova-ho després a l'escena.

D'una població volem extreure una mostra de mida _____.

Si procedeixen de 5 àrees diferents, A, B, C, D i E amb percentatges del total de la població de _____%, _____%, _____%, _____% i _____%


A quants de cada zona s'han d'entrevistar?

A= _____, B= _____, C= _____, D= _____ i E= _____

EXERCICI

13. Una gran empresa té treballadors en quatre àrees: Operaris, Representants, Administració i Direcció. Les condicions de treball són força diferents en cada àrea, per la qual cosa, el grau de satisfacció no és igual en cada una d'elles. Per esbrinar-ho, si hi ha 1000, 500, 300 i 200 treballadors en les àrees d'operaris, representants, administratius i directius, quants se n'han de seleccionar de cada àrea per una mostra de mida ...?

- a) 200
- b) 100
- c) 300

Clica  per anar a la pàgina següent.

5. Estadística bidimensional

5.a Distribucions bidimensionals

Llegeix a la pantalla l'explicació sobre aquest tipus de distribucions i respon:

Què és una **distribució bidimensional**?

Les variables que intervenen, han d'estar relacionades? _____


Què és un diagrama de dispersió? _____

Amb quin altre nom es coneix el diagrama de dispersió? _____

Quan hi ha moltes dades i els parells de valors es repeteixen, utilitzem una _____.

Quins tipus de gràfics s'utilitzen en el cas de moltes dades o de parells repetits? _____

A l'escena de la dreta pots veure els diferents tipus de gràfics. Clica en "Veure gràfic següent" per anar d'un tipus a l'altre.

Clica  Per anar a la pàgina següent.

5.b Correlació lineal

Llegeix a la pantalla l'explicació sobre aquest concepte. A l'escena de la dreta pots moure els punts per modificar així la correlació i observar com canvia el seu valor i com va canviant el núvol de punts. Respon les preguntes següents després de llegir i entendre les explicacions:

Quin és l'objectiu d'un **estudi bidimensional**? _____

Què és la correlació? _____

Com s'aprecia la correlació? _____

Quan diem que la correlació és lineal? _____

Com se sap si la correlació lineal és forta o no? _____

Si la recta és creixent, es diu que la correlació lineal és _____

Si la recta es decreixent, es diu que la correlació lineal és _____

Com s'anomena el paràmetre que s'utilitza per quantificar la correlació? _____

Quins valors pot prendre aquest coeficient? _____

En quins casos la correlació és més forta? _____

Completa la següent taula amb els tipus de correlació en funció dels valors de "r":

Si $r = 1$ ó $r = -1$	
Si $0,5 < r < 1$	
Si $-1 < r < 0,5$	
Si $r = 0$ ó pròxim	

Clica en el botó

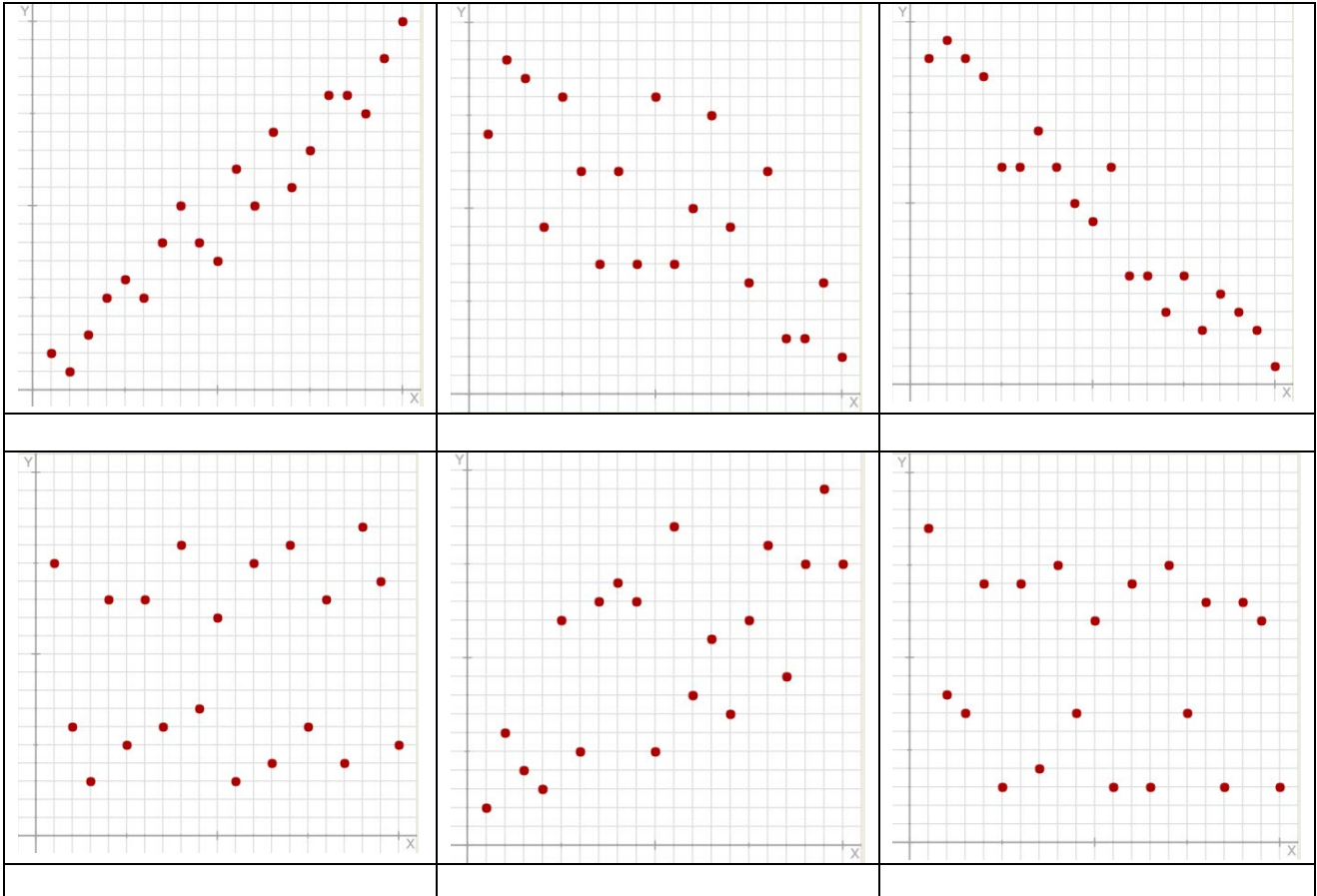


Per fer exercicis.

Apareix un núvol de punts d'una distribució bidimensional. Has de triar la solució correcta entre les proposades indicant com creus que és la correlació.

EXERCICIS:

Indica de quin tipus de correlació és cadascun dels següents núvols de punts, entre les opcions: Molt Forta i Directa, Molt Forta i Inversa, Forta i Directa, Forta i Inversa, Dèbil, Molt Dèbil Quasi Nul·la.



Clica



Per anar a la pàgina següent.

5.c Rectes de regressió

Llegeix a la pantalla l'explicació sobre aquest concepte. A l'escena de la dreta pots veure un núvol de punts que podràs modificar si vols fent lliscar els punts.

Quan hagi situat els punts on vulguis, modifica el valor del pendent amb el control:



Quan hagi trobat el seu valor sortirà dibuixada una recta de color blau, hauràs trobat la **recta de regressió** que s'ajusta a aquest núvol de punts.

Si cliques a **Veure càlculs** podràs observar com s'ha calculat l'equació d'aquesta recta. I en aquesta nova escena, a la part inferior pots clicar **Recta de regressió de X sobre Y** per veure aquesta altra, que com podràs observar pot ser diferent de l'anterior.

Contesta les preguntes següents:

Per quin punt passa la Recta de regressió de Y sobre X? _____

Quina és la seva equació? _____

Quin és el valor del pendent d'aquesta recta? _____

Quin nom rep el pendent d'aquesta recta? _____

Per a què serveix la recta de regressió? _____

En quins casos es més fiable el valor de l'estimació?

- _____
- _____

Clica en el botó



Per fer exercicis.

Apareix un núvol de punts d'una distribució bidimensional. Has de triar la solució correcta entre les proposades indicant quina creus que és l'equació de la recta de regressió de Y sobre X.

EXERCICIS:

A cadascun dels núvols de punts següents indica quina és l'equació que correspon a la recta de regressió de Y sobre X:

<p>Centre de gravetat: (5,5 , 5,8)</p> <p><input type="checkbox"/> $y = -0,5x + 3$</p> <p><input type="checkbox"/> $y = 0,5x + 8,6$</p> <p><input type="checkbox"/> $y = -0,5x + 8,6$</p> <p><input type="checkbox"/> $y = 0,5x + 3$</p>	<p>Centre de gravetat: (5,5 , 4,9)</p> <p><input type="checkbox"/> $y = 0,7x + 8,8$</p> <p><input type="checkbox"/> $y = -0,7x + 1$</p> <p><input type="checkbox"/> $y = 0,7x + 1$</p> <p><input type="checkbox"/> $y = -0,7x + 8,8$</p>
<p>Centre de gravetat: (5,5 , 6,3)</p> <p><input type="checkbox"/> $y = 0,7x + 2,4$</p> <p><input type="checkbox"/> $y = -0,7x + 10,2$</p> <p><input type="checkbox"/> $y = 0,7x + 10,2$</p> <p><input type="checkbox"/> $y = -0,7x + 2,4$</p>	<p>Centre de gravetat: (5,5 , 4,9)</p> <p><input type="checkbox"/> $y = -0,6x + 8,2$</p> <p><input type="checkbox"/> $y = 0,6x + 1,6$</p> <p><input type="checkbox"/> $y = -0,6x + 1,6$</p> <p><input type="checkbox"/> $y = 0,6x + 8,2$</p>

Clica Per anar a la pàgina següent.



Recorda el més important – RESUM

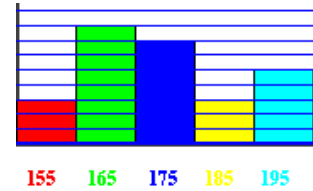
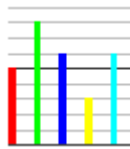
Població.

Mostra

Variabls estadístiques

Tipus

Tipus de gràfics



Mitjana, moda i desviació típica

Mitjana

Moda

Desviació típica

$$\bar{X} =$$

$$M_o =$$

$$\sigma = \sqrt{\quad}$$

Quartil, mediana, percentil

Quartils

Mediana

Percentils

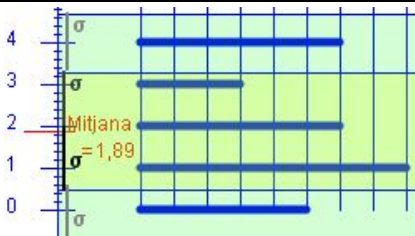
$$Q_1 =$$

$$M_e =$$

$$P_i =$$

$$Q_3 =$$

Mitjana i desviació típica: Observa l'exemple



$$[\bar{x} + \sigma, \bar{x} - \sigma] = [\quad , \quad]$$


% de dades

$$[\bar{x} + 2\sigma, \bar{x} - 2\sigma] = [\quad , \quad]$$

% de dades

Representativitat

Una mostra és representativa de la població quan _____

Clica  per anar a la pàgina següent.



Per practicar

Ara practicaràs resolent diferents EXERCICIS. A les següents pàgines trobaràs EXERCICIS de:

**Mesures de centralització i dispersió. Representativitat
Interpretació de gràfics de l'INE**

Completa l'enunciat amb les dades de cada EXERCICI de la pantalla i després resol-lo. És important que primer el resolguis tu i després comprovis a l'ordinador si ho has fet bé.



Mesures de centralització i dispersió. Representativitat.

1. Tipus de variable (fes dos exercicis)

Classifica les següents variables:	Velocitat → Acceleració → nre. de vàlvules → nre. de places → tipus de vehicle → nre. de rodes → carga neta → tipus de tapisseria →	
<table border="1" style="width: 100%;"> <tr> <td>nre. de fills → flor preferida → pes → temperatura mitjana → sabor → altura →</td> </tr> </table>	nre. de fills → flor preferida → pes → temperatura mitjana → sabor → altura →	
nre. de fills → flor preferida → pes → temperatura mitjana → sabor → altura →		

Classifica les següents variables estadístiques d'un partit de futbol:	jugador preferit → nre. de gols → temps transcorregut →	
<table border="1" style="width: 100%;"> <tr> <td>nre. d'espectadors en el camp →</td> </tr> </table>	nre. d'espectadors en el camp →	
nre. d'espectadors en el camp →		

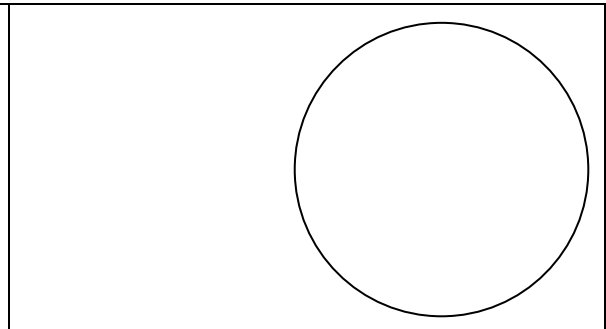
2. Recompte de dades (fes dos exercicis)

Fes un recompte de les dades en una taula.	
 <p>Nre. de germans</p>	
Fes un recompte de les dades en una taula	
 <p>Llargada de la tela</p>	

3. Diagrama de sectors

Fes un diagrama de sectors per a les dades del color preferit de la taula

Marca	Freqüència
x_i	f_i
Total	



4. Diagrama de barres

Fes un diagrama de barres per a les dades de la taula.

Marca	Freqüència
x_i	f_i
Total	

5. Histograma

Amb les dades de la taula fes un histograma

Interval	Marca	Freqüència
	x_i	f_i
	Total	

6. Moda

Quina és la moda en cada grup?

A = {vermell, blau, verd, blau}

B = {blanc, negre, blau}

C = {vermell, verd, groc, vermell, blau, vermell, blau, blau}

A →

B →

C →

Quina és la moda en cada grup?

A = {1, 2, 3, 4, 1, 1, 2, 3}

B = {1, 1, 2, 2, 3, 4, 4}

C = {1, 2, 3, 3, 3, 7, 7, 7, 4}

A →

B →

C →

7. Mediana

Quina és la mediana en cada cas?

A = {1, 2, 3, 4, 5}

B = {1, 2, 3, 4}

C = {1, 2, 3, 4, 5, 6}

D = {1, 2, 3, 3}

E = {1, 2, 3, 3, 3}

A →

B →

C →

D →

E →

Quina és la mediana en cada cas?

A = {1, 2, 7, 10}

B = {3, 6, 7}

C = {1, 2, 3, 4, 5, 6, 1, 2, 1}

A →

B →

C →

8. Igual mitjana

Quina és la mediana en cada cas?

A = { , }; B = { , }; C = { , }

A →

B →

C →

9. Concepte de mitjana

Calcula la mitjana per a les dades:

$x_1 =$ $f_1 =$

$x_2 =$ $f_2 =$

$x_3 =$ $f_3 =$

10. Càlcul de la mitjana

Calcula la mitjana: Distribució discreta			
Marca	Freqüència		
x_i	f_i		
Total			
Calcula la mitjana: Distribució contínua			
Interval	Marca	Freqüència	
	x_i	f_i	
	Total		

11. Cas simple de desviació típica

Quina és la desviació típica en cada cas?			
$A = \{ \quad , \quad \}$; $B = \{ \quad , \quad \}$; $C = \{ \quad , \quad \}$	A →	B →	C →

12. Concepte de desviació típica

Calcula la desviació típica per a les dades:		
$x_1 =$	$f_1 =$	
$x_2 =$	$f_2 =$	
$x_3 =$	$f_3 =$	

13. Càlcul de desviació típica

Calcula la desviació típica: Distribució discreta			
Marca	Freqüència		
x_i	f_i		
Total			
Calcula la desviació típica: Distribució contínua			
Interval	Marca	Freqüència	
	x_i	f_i	
	Total		

14. Representativitat

Prenem una mostra de mida 2000 d'una població on l'edat influeix en la característica de l'estudi. El ___% de la població és major, el ___% jove i el ___% mitjana. A quants entrevistare de cada grup d'edat?	Joves → Mitjans → Majors →
----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	----------------------------------

Clica per anar a la pàgina següent.

Interpretació de gràfics de l'INE.

(En cada apartat apareix una imatge i, en el text, preguntes sobre ella. Clicant en UN ALTRE EXERCICI apareixen més preguntes sobre la mateixa imatge)

1. Què fem? Observa el gràfic de sectors de l'INE i respon a les preguntes:

Quina és la variable estudiada?
I la freqüència?

A quin grup d'activitats dediquem més temps els espanyols?
Quina és la moda?

Calcula quant temps dediquem a la llar i la família.

Quants graus ocupa aquest sector en el diagrama?



2. Quant passegem? En el gràfic és fàcil veure que som els europeus que més passegem.

En quins països passegen més les dones que els homes?

Calcula el temps mitjà que es dedica a cada país a passejar.

Quin país hi ha en el percentil 50?

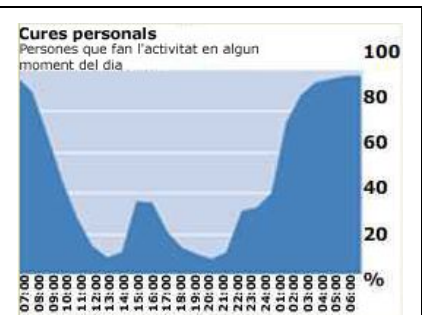


3. Cura personal. Observa el gràfic i respon a les preguntes:

Creus que el son s'ha comptat com a activitat de cura personal?

A les 15:00 hi ha un màxim local en la gràfica, quina és la causa?

A l'hora del dinar, el 38% de les persones es dedica a la cura personal. Significa això que un 62% de les persones no menja?



4. Vida social. Observa el gràfic i respon a les preguntes:

Quines són les comunitats en les quals es dedica menys temps a la vida social i a la diversió?

Quant temps dediquen a la diversió o a la vida social la major part de les comunitats?

Quin és el temps mitjà que es dedica a Espanya a aquesta activitat?



Clica per anar a la pàgina següent.

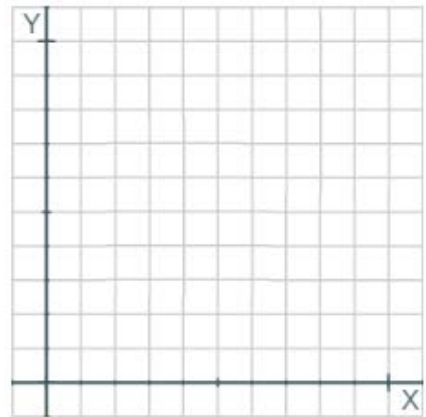
DISTRIBUCIONS BIDIMENSIONALS.

1. Notes

Les notes de 8 alumnes en Llengua i Anglès van ser:

LLENGUA (X)									
ANGLÈS (Y)									

Dibuixa el núvol de punts i calcula el coeficient de correlació lineal:

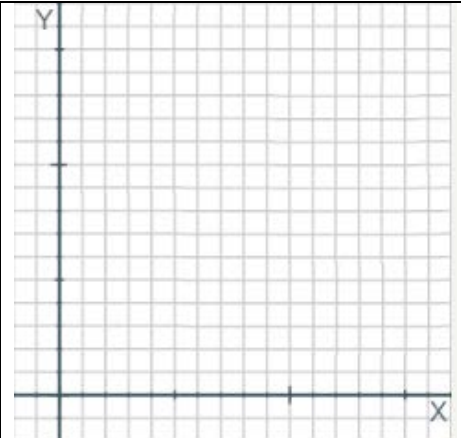


2. Esport o TV?

Les hores setmanals que dediquen 10 persones a fer esport i a veure la TV són:

ESPORT (X)									
TV (Y)									

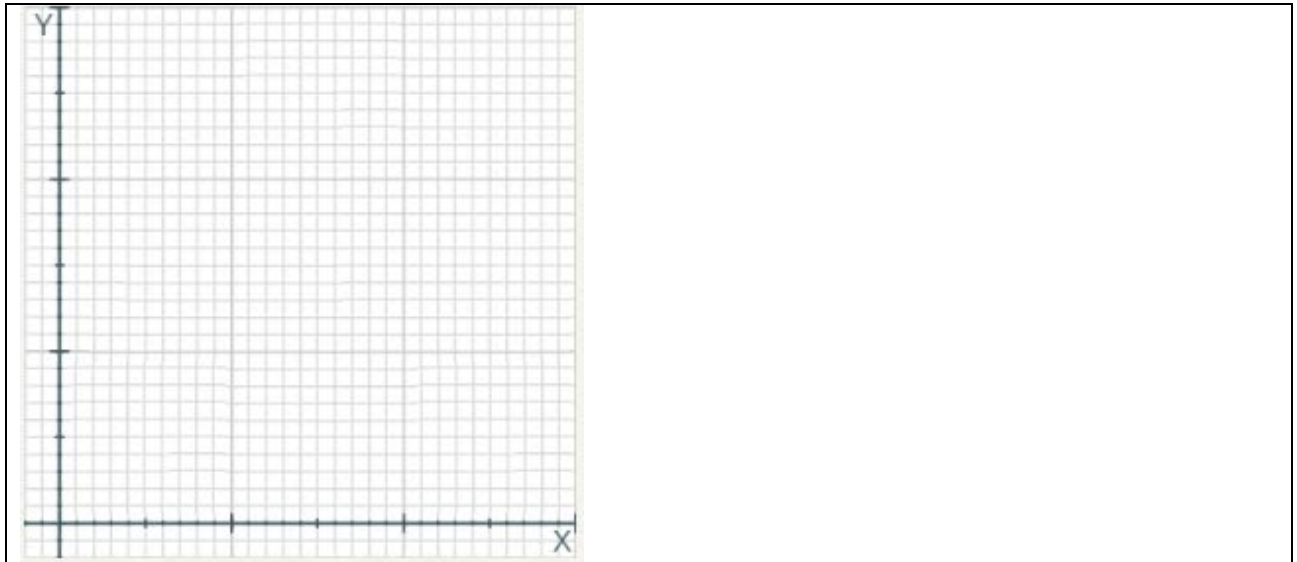
Dibuixa el núvol de punts i calcula el coeficient de correlació lineal:



3. Climatologia

En un lloc s'han mesurat les temperatures mitjanes i precipitacions cada mes durant un any. Dibuixa un núvol de punts i calcula el coeficient de correlació lineal:

°C (X)												
Litres/m ² (Y)												



4. Rectes de regressió

D'una distribució bidimensional (X,Y) coneixem:

$X =$ $Y =$ $\sigma_x =$ $\sigma_y =$ $\sigma_{xy} =$

- a) Calcula el coeficient de correlació lineal
- b) Calcula la recta de regressió de Y sobre X i el valor estimat de y per a $x = \underline{\hspace{2cm}}$

5. Amb els comptes fets

D'una distribució bidimensional (X,Y) coneixem:

$X =$ $Y =$ $\sigma_x =$ $\sigma_y =$ $\sigma_{xy} =$

- a) Calcula la recta de regressió de Y sobre X
- b) Calcula el valor estimat de y per a $x = \underline{\hspace{2cm}}$
- c) És fiable aquesta predicció?

Clica Per anar a la pàgina següent.

Autoavaluació



Completa aquí cada un dels enunciats que van apareixent a l'ordinador i resol-lo, després introdueix el resultat per comprovar si la solució és correcta.

1 Quants graus corresponen en el diagrama de sectors al valor de freqüència_____?	
2 Quina és la mediana?	
3 Quina és la moda?	
4 Quin és el percentatge de la mostra que correspon a les dues primeres marques?	
5 Quin és el percentil més petit que deixa per sota els valors menors a 3?	
6 Quina és la mitjana?	
7 Calcula la desviació típica.	
8 Associa cada núvol de punts amb el seu corresponent coeficient de correlació, r. 1) _____ 2) _____ 3) _____	
9 Calcula la covariància.	
10 El centre de gravetat de una distribució bidimensional és (____ , ____) i el pendent de la recta de regressió de Y sobre X és _____. Estima el valor de y para x=_____.	