



## Reconocimiento-No comercial-Compartir bajo la misma licencia 3.0 España

### Usted es libre de:



copiar, distribuir y comunicar públicamente la obra



hacer obras derivadas

### Bajo las condiciones siguientes:



**Reconocimiento.** You must attribute this work to [Departament de Matemàtiques de l'IES el SUJ](#) (with link).

Attribute this work:

```
<div xmlns:cc="http://creativecommons.org/ns#" about="http://www.xtec.cat/ieselsuj" data-bbox="263 373 739 387" type="code">
```



**No comercial.** No puede utilizar esta obra para fines comerciales.



**Compartir bajo la misma licencia.** Si altera o transforma esta obra, o genera una obra derivada, sólo puede distribuir la obra generada bajo una licencia idéntica a ésta.

- Al reutilizar o distribuir la obra, tiene que dejar bien claro los términos de la licencia de esta obra.
- Alguna de estas condiciones puede no aplicarse si se obtiene el permiso del titular de los derechos de autor
- Nada en esta licencia menoscaba o restringe los derechos morales del autor.

Advertencia

Los derechos derivados de usos legítimos u otras limitaciones reconocidas por ley no se ven afectados por lo anterior.

## Full de treball A.

### *Distribucions bidimensionals*

**A.1** Es vol buscar la relació experimental que hi ha entre les variables velocitat i temps que triga un mòbil en fer un determinat recorregut

Després de fer l'experiment 5 cops s'han obtingut els següents resultats:

<b>Temps</b>	0.5	2	2.5	3	4
<b>Velocitat</b>	2	8	10	12	16

A partir de les dades obtingudes es vol fer la predicció de quina velocitat anava el mòbil si ha trigat 5 segons en fer el recorregut.

**A.2** Una fàbrica de medicaments ha desenvolupat un nou analgèsic i vol determinar la dosi òptima per un adult. Per a fer-ho subministra diferents quantitats d'analgèsic a 7 pacients i observa la variació de temperatura corporal, en valor absolut, al cap d'una hora. Els resultats són els següents:

<b>Miligrams</b>	350	400	500	550	600	700	800
<b>Variació temperatura</b>	0.2	0.4	0.7	1.1	1.1	1.6	1.9

Si es vol aconseguir una variació de temperatura d'1° ¿Quina dosi caldrà subministrar?. Justifica la teva resposta.

### **UNA MICA DE VOCABULARI:**

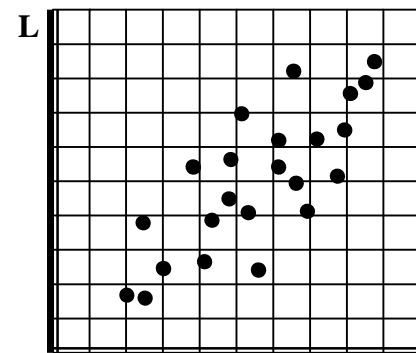
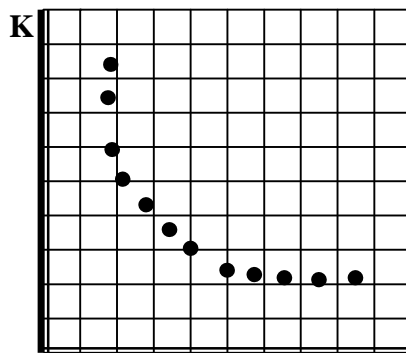
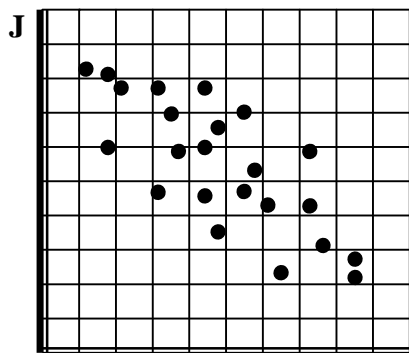
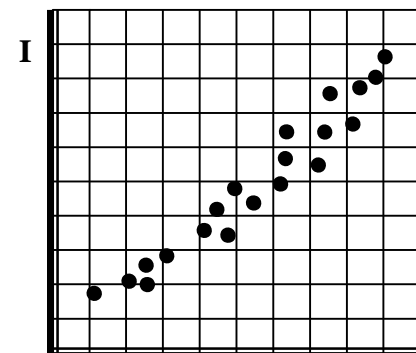
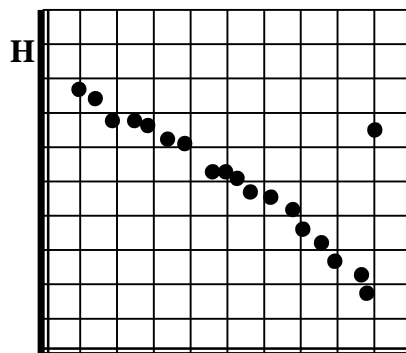
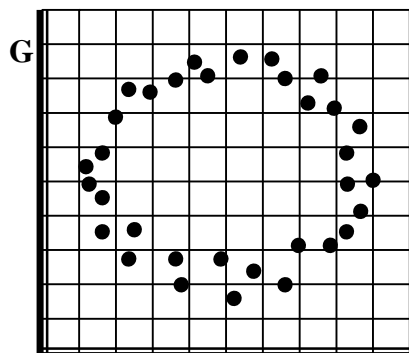
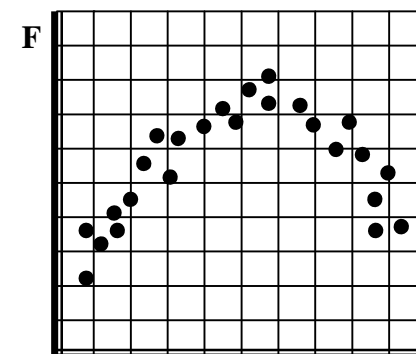
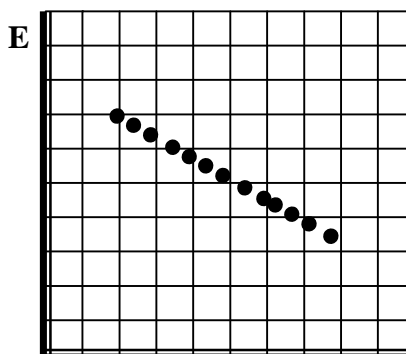
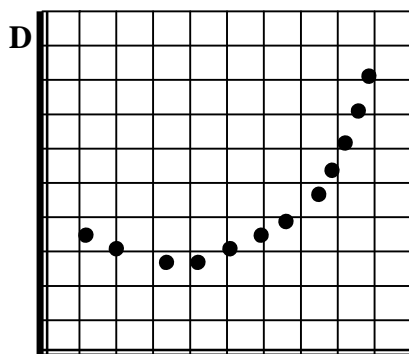
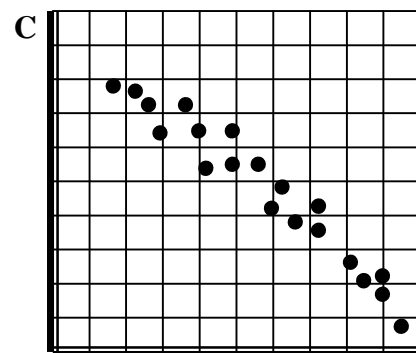
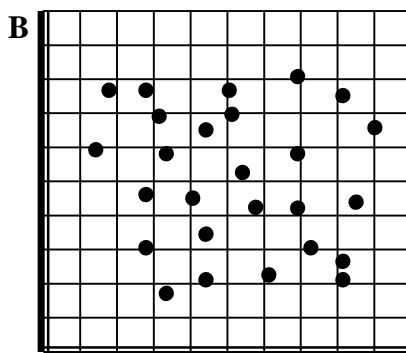
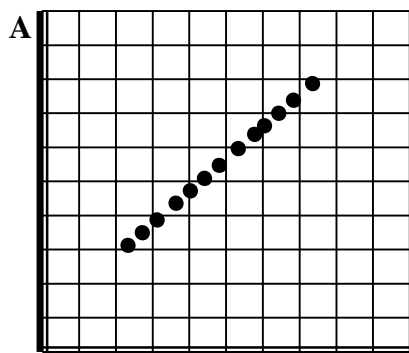
Algunes vegades ens pot interessar conèixer si dues variables estadístiques estan relacionades. Per a fer-ho, es pot començar considerant que les variables estadístiques són les variables  $X, Y$  d'uns eixos de coordenades i dibuixar tots els punts. Aquest dibuix s'anomena **diagrama de dispersió o núvol de punts** i ens permet fer una primera estimació de si la suposada relació existeix, de si és forta o fluixa i de quin tipus de relació hi ha.

Si existeix relació entre dues variables, aquesta pot ser de molts tipus. Si es tracta, per exemple, de dos variables físiques relacionades per alguna fórmula matemàtica (com el primer exercici) direm hi ha una **relació funcional** doncs existeix una funció (fórmula) que les relaciona. En aquest cas, els punts del diagrama de dispersió estan perfectament ordenats seguint una línia recta o corba.

Quan dues variables estan relacionades però no existeix una relació funcional entre elles els punts del diagrama de dispersió envolten una recta o corba invisible. En aquests casos direm que existeix una **relació lineal o curvilínia** segons el cas.

En aquest tema estudiarem els casos en que hi hagi relació lineal. En aquests casos, la recta a la que s'aproximen tots els punts s'anomena **recta de regressió**. El problema matemàtic al que ens enfrontem consistirà, en primer lloc en determinar la intensitat de la relació i en segon lloc en trobar l'equació de la recta de regressió per a poder fer prediccions fiables.

**A.3** Observa els diagrames de dispersió següents:



Respon a les preguntes següents per als diagrames anteriors:

- Existeix alguna relació entre les variables X i Y?
- Si existeix, ¿És lineal, corbada,...?
- Es la relació directa (positiva) o inversa (negativa)?
- És la relació funcional? Si és que no, ¿És forta o fluixa?

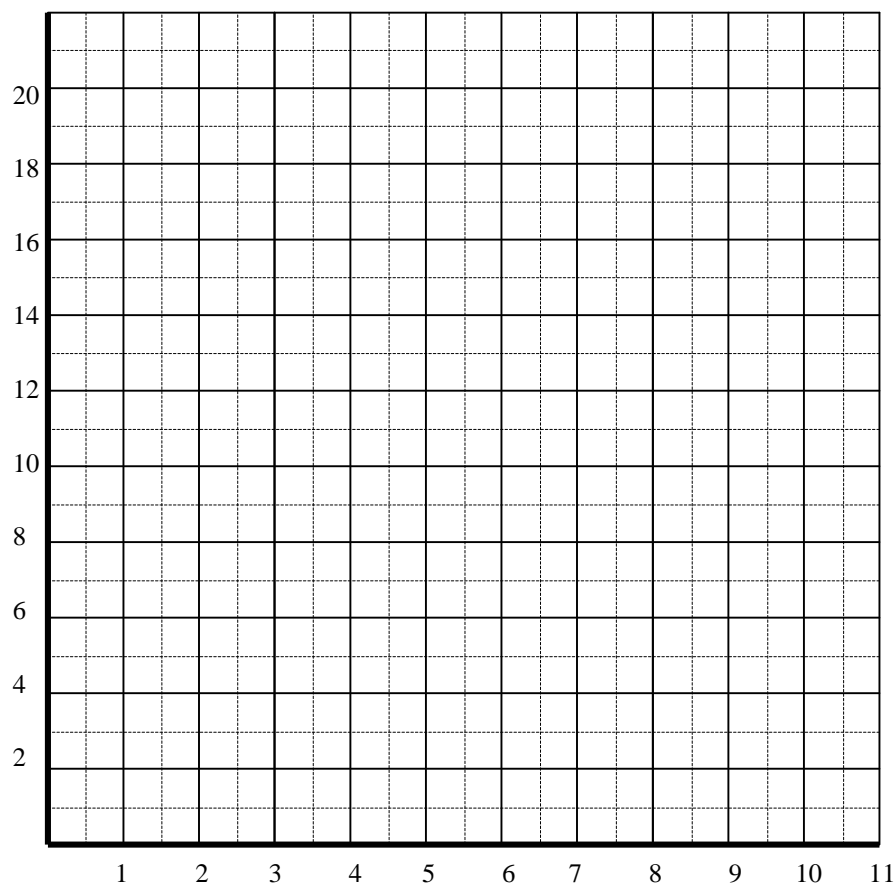
## Full de treball B.

### COEFICIENT DE CORRELACIÓ LINEAL

**B.1** Per tal de quantificar la possible relació que hi ha entre dos variables estadístiques necessitem un valor numèric que determini d'una manera matemàticament rigorosa quina és la intensitat d'aquesta possible relació. Aquest valor s'anomena **coeficient de correlació lineal**. Per entendre bé d'on surt aquest coeficient ho veurem en un exemple numèric molt senzill.

Pots utilitzar la següent taula i el següent gràfic per a fer els exercicis:

$x_i$	$y_i$	$x_i^2$	$y_i^2$	$x_i \cdot y_i$
5	9			
10	16			
5	5			
7	11			
11	21			
8	16			



- Dibuixa els punts en el diagrama de dispersió
- Calcula la mitjana de les dues variables:  $\bar{x}$  i  $\bar{y}$
- Si considerem aquests valors  $\bar{x}, \bar{y}$  com a coordenades podem dibuixar un nou punt  $(\bar{x}, \bar{y})$  en el núvol de punts. Dibuixa'l. Aquest punt s'anomena **punt mitjà**.

Observa que aquest punt constitueix una espècie de centre de gravetat del núvol de punts. Sembla, doncs, necessari que qualsevol recta que vulgui ajustar-se bé al núvol de punts tingui que passar, necessàriament, per aquest punt.

- Calcula la **desviació tipus** de les dues variables:  $\sigma_x$  i  $\sigma_y$  (utilitza les columnes

$$x_i^2 \text{ i } y_i^2 \text{ de la taula). Recorda que } \sigma_x = \sqrt{\frac{\sum_{i=1}^n x_i^2}{N} - \bar{x}^2}$$

- Calcula (utilitzant la columna  $x_i \cdot y_i$ ) el resultat de fer:

$$\sigma_{xy} = \frac{\sum_{i=1}^n x_i y_i}{N} - \bar{x} \bar{y}$$

Aquest valor s'anomena **covariància**

- Definim el **coeficient de correlació lineal** com  $r = \frac{\sigma_{xy}}{\sigma_x \sigma_y}$ . Calcula'l

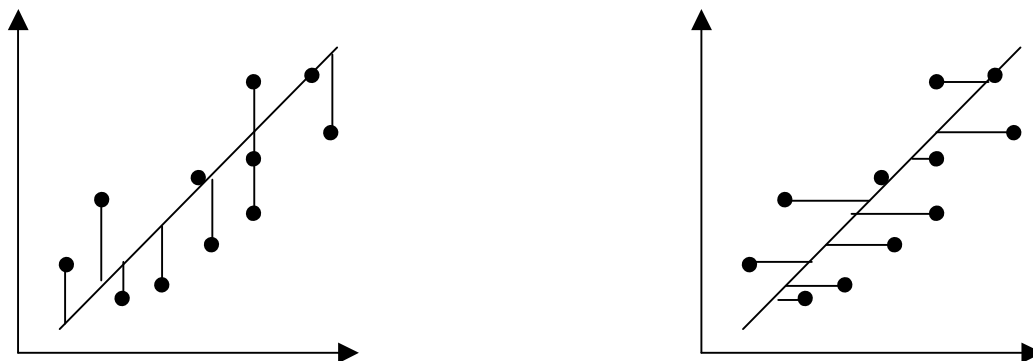
### **INTERPRETACIÓ DEL COEFICIENT DE CORRELACIÓ LINEAL:**

El coeficient de correlació lineal és un valor que oscil·la entre  $-1$  i  $1$ .

- Si  $r = -1$  es tracta d'una relació funcional inversa (la recta de regressió tindrà pendent negatiu)
- Si  $r$  és prop de  $-1$  existeix una relació inversa no funcional entre les variables que serà més forta contra més prop de  $-1$  estigui el valor.
- Si  $r$  és prop de  $0$  no existeix relació entre les variables.
- Si  $r$  és prop de  $1$  existeix una relació directa no funcional entre les variables que serà més forta contra més prop de  $1$  estigui el valor.
- Si  $r = 1$  es tracta d'una relació funcional directa (la recta de regressió tindrà pendent positiu)

## RECTA DE REGRESIÓ

Un dels objectius fonamentals d'aquest tema és fer prediccions. Es a dir, una vegada determinat que sí existeix una determinada correlació lineal entre dues variables ens plantejem la possibilitat de, coneixen una nova dada  $x$ , endevinar la nova dada  $y$ . Per a fer-ho, tan sols ens cal trobar l'equació matemàtica de la recta que més s'aproxima a la majoria de punts. La tècnica matemàtica que permet la recerca d'aquesta equació s'anomena mètode dels *mínims quadrats* i consisteix en trobar la recta que fa que el quadrat de la longitud dels segments que va dels punts a la recta sigui mínima. Aquest mètode es pot aplicar de dues maneres: Considerant el segment en vertical, o considerant-lo en horitzontal:



Aquesta diferència de mètodes produeix dues rectes **diferents** una serveix per aïllar la  $x$  en funció de la  $y$  i l'altra serveix per aïllar la  $y$  en funció de la  $x$ .

**Recta de regressió de  $y$  en funció de la  $x$ :**  $y = Ax + B$

$$A = \frac{\sigma_{xy}}{\sigma_x^2} \quad B = \bar{y} - A\bar{x}$$

**Recta de regressió de  $x$  en funció de la  $y$ :**  $x = A'y + B'$

$$A' = \frac{\sigma_{xy}}{\sigma_y^2} \quad B' = \bar{x} - A'\bar{y}$$

### B.2 Utilitzant l'exercici B.1

- Calcula la recta de regressió de  $y$  en funció de  $x$
- Dibuixa-la en el diagrama de dispersió
- Calcula la recta de regressió de  $x$  sobre  $y$
- Dibuixa-la en el diagrama de dispersió
- ¿Coincideixen les dues rectes?

Com has comprovat el càlcul de la recta de regressió és llarg i una mica feixuc. Ara veurem com tots aquests càlculs es poden simplificar utilitzant el modus estadístic de la teva calculadora. Qualsevol calculadora científica (excepte l'antiga *Casio Fx 82*) calcula tots els paràmetres estadístics de la correlació lineal. A tall d'exemple, explicarem com fer-ho amb la *Casio fx-350MS* (si tens dubtes consulta el manual de la teva calculadora).

Primer haurem de posar la calculadora a punt per fer càlculs estadístics. Per això hem d'activar el modus estadístic. Prem la tecla mode i apareixerà en pantalla tres opcions: COMP, SD i REG. La opció SD (mode 2) en permetrà calcular la mitjana i la desviació tipus d'una variable sola, però com que volem treballar en dues variables triarem l'opció **REG (mode3)** que ens permet calcular la **regressió**. Ara

podem triar entre diferents tipus de regressió. Volem la regressió lineal, per això triem la **opció 1 (lin)**.

Abans de res netegem les memòries fent  $\text{SHIFT}$   $\text{CLR}$   $\text{1}$  (Scl)  $\text{SHIFT}$   $\text{C}$  (Stat clear)

Ara ja podem entrar les dades **x** i **y** separades per una **coma** i després prem la tecla **DT**. Fes-ho tantes vegades com dades hi hagi. Una vegada acabada la introducció de dades anem a obtenir resultats. Per això aquí teniu un quadre amb tots els càlculs que podem fer:

Para llamar este tipo de valor:	Realice esta operación de tecla:
$\Sigma x^2$	$\text{SHIFT}$ $\text{S-SUM}$ $\text{1}$
$\Sigma x$	$\text{SHIFT}$ $\text{S-SUM}$ $\text{2}$
$n$	$\text{SHIFT}$ $\text{S-SUM}$ $\text{3}$
$\Sigma y^2$	$\text{SHIFT}$ $\text{S-SUM}$ $\text{▶ 1}$
$\Sigma y$	$\text{SHIFT}$ $\text{S-SUM}$ $\text{▶ 2}$
$\Sigma xy$	$\text{SHIFT}$ $\text{S-SUM}$ $\text{▶ 3}$
$\Sigma x^3$	$\text{SHIFT}$ $\text{S-SUM}$ $\text{▶ ▶ 1}$
$\Sigma x^2 y$	$\text{SHIFT}$ $\text{S-SUM}$ $\text{▶ ▶ 2}$
$\Sigma x^4$	$\text{SHIFT}$ $\text{S-SUM}$ $\text{▶ ▶ 3}$
$\bar{x}$	$\text{SHIFT}$ $\text{S-VAR}$ $\text{1}$
$x\sigma_{ii}$	$\text{SHIFT}$ $\text{S-VAR}$ $\text{2}$
$x\sigma_{ii-1}$	$\text{SHIFT}$ $\text{S-VAR}$ $\text{3}$
$\bar{y}$	$\text{SHIFT}$ $\text{S-VAR}$ $\text{▶ 1}$
$y\sigma_{ii}$	$\text{SHIFT}$ $\text{S-VAR}$ $\text{▶ 2}$
$y\sigma_{ii-1}$	$\text{SHIFT}$ $\text{S-VAR}$ $\text{▶ 3}$
Coefficiente de regresión A	$\text{SHIFT}$ $\text{S-VAR}$ $\text{▶ ▶ 1}$
Coefficiente de regresión B	$\text{SHIFT}$ $\text{S-VAR}$ $\text{▶ ▶ 2}$
Solamente una regresión no cuadrática	
Coefficiente de correlación $r$	$\text{SHIFT}$ $\text{S-VAR}$ $\text{▶ ▶ 3}$
$\hat{x}$	$\text{SHIFT}$ $\text{S-VAR}$ $\text{▶ ▶ ▶ 1}$
$\hat{y}$	$\text{SHIFT}$ $\text{S-VAR}$ $\text{▶ ▶ ▶ 2}$

### Exemple:

Amb les dades de l'exercici A1 anem a contestar la pregunta que es feia. Comencem entrant totes les dades. Ara calcularem el coeficient de correlació lineal  $r$  fent la combinació de tecles  $\text{SHIFT}$   $\text{S-VAR}$   $\text{▶ ▶ 3}$  i igual i ens dona 1. Ara podem calcular la velocitat si ha trigat 5 segons. Per fer-ho només cal trobar la  $\hat{y}$  tot fent  $\text{5}$   $\text{SHIFT}$   $\text{S-VAR}$   $\text{▶ ▶ ▶ 2}$

Si hem de calcular la recta de regressió cal trobar la A i la B. Però atenció a la calculadora la fórmula és  $y = A + Bx$  i per tant la A i la B estan al revés del que estem acostumats.

En l'exemple anterior dóna  $A=0$  i  $B=4$  per tant la recta és  $y = 4x$

**B.3** Torna a fer tots els càlculs de l'exemple B1 amb la teva calculadora i comprova que els resultats els tens correctes.

**B.4** Els alumnes de 1er de batxillerat d'un institut han tret les següents notes

Física	4.5	6	3.4	7	9
Matemàtiques	5	7	3	6.5	8

Volen saber si hi ha relació entre les notes de matemàtiques i les de física. Per aquesta raó:

- a) Troba el coeficient de correlació lineal i interpreta-ho.
- b) Troba les dues rectes de regressió.
- c) Intenta predir quina nota tindrà de matemàtiques un alumne que hagi tret un 5 de física.
- d) I un alumne que ha tret un 4 de matemàtiques, quina nota pot treure de física.



## Full de treball C.

**C.1** Les despeses en publicitat d'una empresa, en milers d'euros, i les seves corresponents vendes, també en milers d'euros, són les de la taula:

<b>Publicitat</b>	1	2	3	4	5	6	7	8
<b>Vendes</b>	15	16	14	17	20	18	18	19

- ¿Creus que es podria fer una bona predicció de vendes a partir d'una partida pressupostària en publicitat? Raona la resposta.
- Quants diners ens hem de gastar en publicitat si desitgem 20 milers d'euros en vendes?

**C.2** El nombre de bacteris per unitat de volum, presents en un cultiu després d'un cert nombre d'hores, és expressat en la següent taula:

<b>Hores</b>	0	1	2	3	4	5
<b>Bacteris per unitat de volum</b>	12	19	23	34	56	62

- Fes el diagrama de dispersió.
- ¿És pot fer una predicció fiable del bacteris que hi haurà a les 6 hores? Justifica la resposta i fes, si es cau, la predicció
- ¿I a les 20 hores?. Justifica la resposta i fes, si es cau, la predicció.

**C.3** Dibuixeu el núvol de punts de les dades següents:

<b>x</b>	1	1	2	3	5	8	10	10
<b>y</b>	9	8	7	5	3	3	10	1

Estimeu i calculeu el coeficient de correlació lineal i la recta de regressió. Hi ha una dada causant del baix valor que té  $r$ ; traieu-la i calculeu novament el coeficient de correlació i la recta de regressió. Dibuixeu les dues rectes i compareu els resultats.

**C.4** L'encarregat d'un bar ha triat una mostra a l'atzar de 10 setmanes durant un any i n'ha mesurat la temperatura mitjana així com també, el nombre de caixes de begudes refrescants venudes, i ha obtingut la taula següent:

<b>Temperatura</b>	10	28	11	32	30	18	25	5	8	15
<b>Nº de caixes</b>	21	65	18	72	75	40	65	10	12	25

Utilitzant aquestes dades, ¿Es possible planificar la compra de begudes refrescants en funció de la temperatura?. ¿De quina manera?

**C.5** S'ha observat que existeix una forta correlació entre el nombre d'hectàrees cremades en incendis a Espanya i el nombre d'hectàrees cremades a la resta d'Europa. ¿Significa això que Espanya té la culpa de que es cremin els boscos a Europa?.

Quan existeix correlació entre dues variables pot ser que:

- Que una sigui causa de l'altre
- Que hi hagi influència mútua
- Que les dues estiguin influïdes per terceres variables.
- Que la correlació sigui fruit de l'atzar

**C.6** Suposem que entre les següents variables s'ha observat una forta correlació. ¿A quina de les causes anteriors creus que es deu la correlació:

- a) Gruix de la neu i nombre d'esquiadors
- b) Sou de funcionaris i consum d'aliments precuinats
- c) Preu del tabac i grau de contaminació industrial
- d) Pluja i collita agrícola
- e) Oferta de serveis i demanda de serveis.
- f) Nivell cultural i dieta alimentària
- g) Absentisme laboral i nivells de lectura.
- h) Consum i sou
- i) Hores de treball i unitats produïdes

**C.7** Un individu vol obtenir un benefici de 1 500 euros per aconseguir-ho vol invertir en Borsa amb totes les garanties de guany. Investiga alguns experts inversionistes i recull la següent informació d'ells: diners guanyats després de fer certa inversió

<b>Inversió (en milers)</b>	20	30	5	40	7	13	25	3
<b>Beneficis (en milers)</b>	2	0.5	1	2	0	3	0.2	1

¿Quants diners cal invertir en borsa per tal d'obtenir el milió i mig de benefici desitjat?

**C.8** Les qualificacions en matemàtiques i física de 20 alumnes de segon de Batxillerat han estat les següents:

<b>Nota Matemàtiques (<math>x_i</math>)</b>	2	4	5	6	6	7	7	8	10
<b>Nota Física (<math>y_i</math>)</b>	3	5	4	6	5	6	7	9	10
<b>Nombre alumnes (<math>f_i</math>)</b>	2	3	6	2	2	2	1	1	1

- a) Calculeu el coeficient de correlació lineal entre  $x$  i  $y$ .
- b) Si un alumne no ha pogut assistir a l'examen de Matemàtiques però sí al de Física, i té un 5 en aquesta darrera assignatura, ¿Quina nota és d'esperar que hagués tret de Matemàtiques?

- c) Si un alumne té un 4 de Matemàtiques ¿Quina nota és previsible que tregui a física?
- d) Seria just posar la nota d'un alumne utilitzant aquest mètode? ¿Per què?

## ANNEX

Estimat Alumne d'estadística:

Soc el president del tribunal electoral i he rebut una greu denúncia del partit progressista respecte a presumptes irregularitats en el vot per correu en les últimes eleccions. M'he pres la confiança de demanar-li un **informe tècnic** en el que exposi, **justificadament**, la seva opinió al respecte, doncs és de tot hom reconegut el seu alt prestigi com alumne del Institut.

La denúncia presentada pel partit progressista afirma que el partit conservador ha comès certes irregularitats en el vot per correu en les recents eleccions. Les proves presentades pels denunciants són ambigües i, per contra, el partit conservador afirma que el bon resultat en el vot per correu és fruit d'una campanya electoral exhaustiva dedicada, específicament, a aquest sector de la població.

En un intent d'esbrinar la veritat he recollit les estadístiques d'eleccions passades per tal de comparar els següents valors:

$x$  = Diferència (positiva), en percentatge, entre els vots obtinguts pel partit progressista i els vots obtinguts pel partit conservador **en les urnes**

$y$  = Diferència (positiva), en percentatge, entre els vots obtinguts pel partit progressista i els vots obtinguts pel partit conservador **en el vot per correu**

$x$	2.5	3.7	1.3	1.6	7.6	9.4	6.2	4.9	4.5	3.9
$y$	4.2	6	1.4	3.1	9.9	14.6	9.8	7.5	6.3	5.3

En les recents eleccions en que s'ha produït el conflicte les diferències en percentatge de vots han estat les següents:  $x = 5.1$      $y = 15.9$

Li prego m'envii amb prestesa les conclusions a les seves investigacions el més justificades possibles per tal d'ajudar-me a dictar la sentència que anul·li o no el vot per correu.

Atentament i agraint-li el seu inestimable ajut:  
Sr. Just Deltot i Massa

( Aquesta història està basada en uns fets reals succeïts amb els vots de Filadèlfia a les eleccions senatorials de l'estat de Pensilvània de l'any 95. En aquest cas real el president del tribunal va utilitzar l'informe d'un prestigiós matemàtic per dictar sentència)