



PRELIMINARS DEL CÀLCUL ESTADÍSTIC

Durant un mes s'han observat les despeses de manutenció i l'ingrés total de 6 famílies. Els resultats obtinguts són:

<i>Despeses en milers d'€ (Y)</i>	<i>Ingressos en milers d'€ (X)</i>
1,5	2,4
1,8	3,2
2,4	3,6
2,8	4,2
3,2	4,4
3,2	4,5

Es demana:

- Calculeu el coeficient de correlació lineal. Segons aquest valor, penseu si la relació lineal entre X i Y és feble o forta?
- Determineu la recta de regressió que expressa les despeses en funció dels ingressos.
- Quines despeses tindria una família amb un ingrés de 3 800 € al mes?

Per al càlcul del coeficient de correlació lineal o de Pearson,  $r = \frac{\sigma_{xy}}{\sigma_x \cdot \sigma_y}$ , hem de

saber la covariància i les desviacions típiques per a cada variable:

$$\sigma_{xy} = \frac{\sum_{i=1}^6 (x_i - \bar{x})(y_i - \bar{y})n_i}{6}; \quad \sigma_x = \sqrt{\frac{\sum_{i=1}^6 (x_i - \bar{x})^2 n_i}{6}}; \quad \sigma_y = \sqrt{\frac{\sum_{i=1}^6 (y_i - \bar{y})^2 n_i}{6}}$$

El significat dels sumatoris és:

$$\sum_{i=1}^6 (x_i - \bar{x})(y_i - \bar{y})n_i = (x_1 - \bar{x})(y_1 - \bar{y})n_1 + \dots + (x_6 - \bar{x})(y_6 - \bar{y})n_6$$

$$\sum_{i=1}^6 (x_i - \bar{x})^2 n_i = (x_1 - \bar{x})^2 n_1 + \dots + (x_6 - \bar{x})^2 n_6$$

$$\sum_{i=1}^6 (y_i - \bar{y})^2 n_i = (y_1 - \bar{y})^2 n_1 + \dots + (y_6 - \bar{y})^2 n_6$$

On és clar que  $\bar{x} = \frac{\sum_{i=1}^6 x_i n_i}{6}$ ;  $\bar{y} = \frac{\sum_{i=1}^6 y_i n_i}{6}$

Per realitzar i visualitzar més fàcilment aquestes operacions farem una taula, anomenada Taula de freqüències que, per aquests càlculs, haurem d'organitzar així:



$x_i$	$y_i$	$n_i$	$x_i n_i$	$y_i n_i$	$x_i - \bar{x}$	$y_i - \bar{y}$	$(x_i - \bar{x})^2 \cdot n_i$	$(y_i - \bar{y})^2 \cdot n_i$	$(x_i - \bar{x})(y_i - \bar{y})n_i$
2,4	1,5	1	2,4	1,5	-1,32	-0,98	1,7424	0,9604	1,2936
3,2	1,8	1	3,2	1,8	-0,52	-0,68	0,2704	0,4624	0,3536
3,6	2,4	1	3,6	2,4	-0,12	-0,08	0,0144	0,0064	0,0096
4,2	2,8	1	4,2	2,8	0,48	0,32	0,2304	0,1024	0,1536
4,4	3,2	1	4,4	3,2	0,68	0,72	0,4624	0,5184	0,4896
4,5	3,2	1	4,5	3,2	0,78	0,72	0,6084	0,5184	0,5616

SUMES  
NECESSÀRIES:

22,3 14,9

3,3284

2,5684

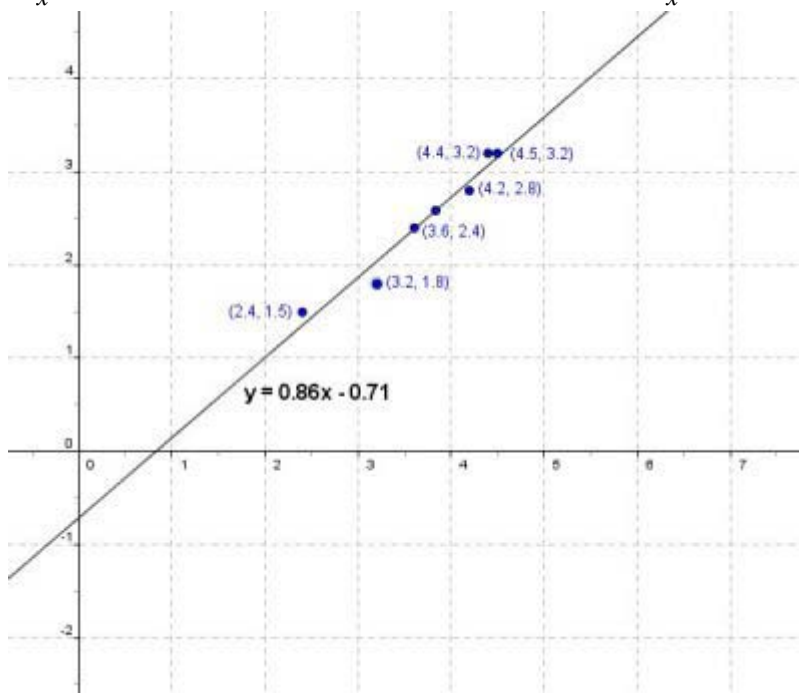
2,8616

S'han calculat les mitjanes aritmètiques:

$$\bar{x} = \frac{\sum_{i=1}^6 x_i n_i}{6} = \frac{22,3}{6} = 3,71\hat{6}; \quad \bar{y} = \frac{\sum_{i=1}^6 y_i n_i}{6} = \frac{14,9}{6} = 2,48\hat{3}.$$

Amb les fórmules donades abans i aquests resultats, també es calculen les desviacions típiques i la covariància. **Calculem el pendent de la recta de regressió,**

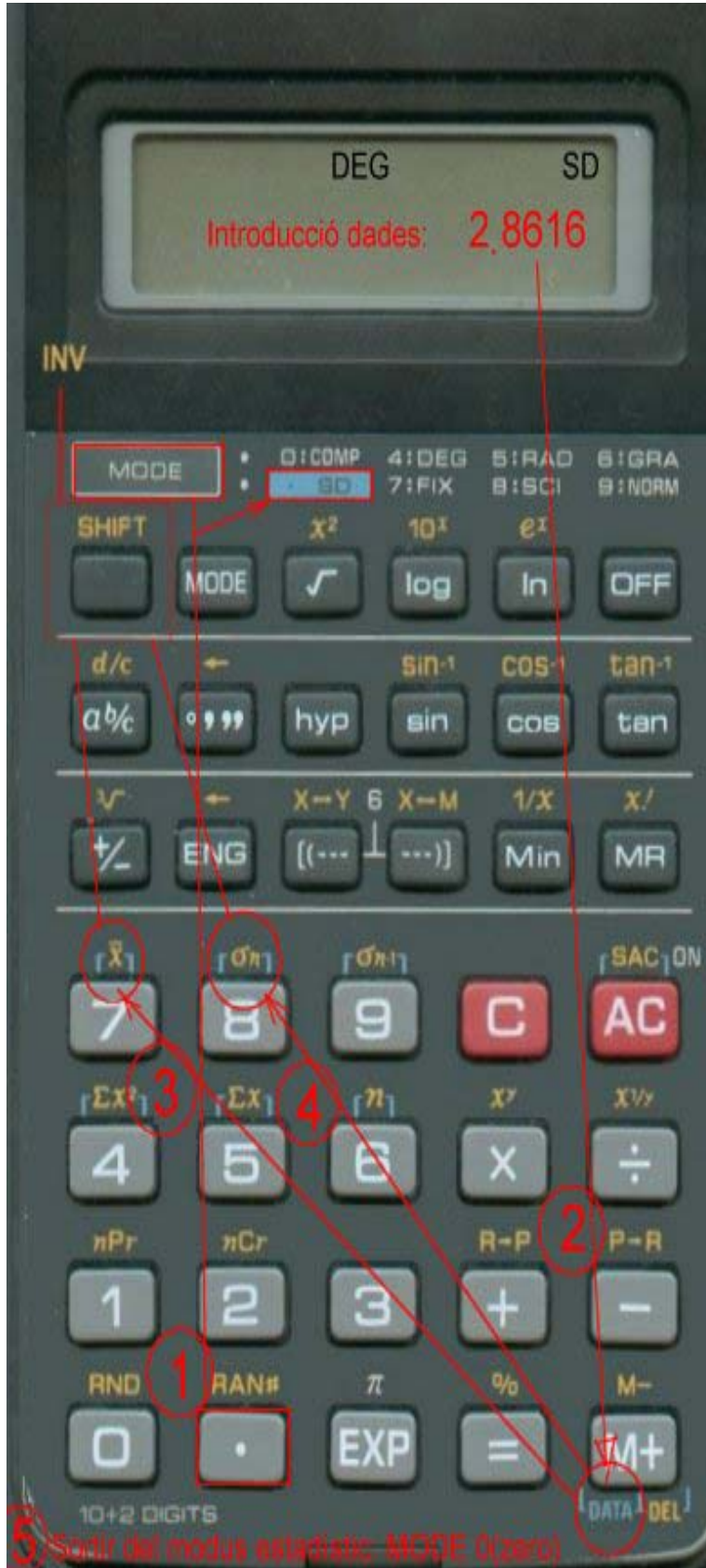
$$\frac{\sigma_{xy}}{\sigma_x^2}, \text{ i l'equació d'aquesta recta: } y - \bar{y} = \frac{\sigma_{xy}}{\sigma_x^2} (x - \bar{x}) \Rightarrow y = \frac{\sigma_{xy}}{\sigma_x^2} x - \frac{\sigma_{xy}}{\sigma_x^2} \bar{x} + \bar{y}$$



Com es pot observar en la figura, la correlació lineal és positiva i forta, d'acord amb el coeficient de Pearson  $r = 0,9786$ , positiu i pròxim a 1.

Situeu el punt (3,8, 2,6); observareu que pertany a la recta. De fet, el teniu dibuixat.

Comentem que amb una calculadora científica CASIO fx-82SUPER FRACTION (molt antiga, però eficient per a principiants) podeu trobar la mitjana aritmètica i la desviació típica.



A continuació t'expliquem el funcionament d'una calculadora estàndard preparada per a efectuar càlculs estadístics.

### Introducció de dades

- Posem la calculadora en mode estadístic **SD**.
- Esborrem de la memòria els càlculs anteriors.
- Hi introduïm les dades, l'una a continuació de l'altra, separades només per la tecla **DATA**:

$$x_1 \text{ DATA } x_2 \text{ DATA } \dots x_n \text{ DATA}$$

Si les dades estan tabulades en una taula de freqüències, introduïm les dades de la manera següent:

$$x_1 \text{ X } n_1 \text{ DATA } x_2 \text{ X } n_2 \text{ DATA } \dots \\ \dots x_k \text{ X } n_k \text{ DATA}$$

### Obtenció de paràmetres estadístics

- Per a obtenir la mitjana aritmètica, premem la tecla  **$\bar{x}$** .
- Per a obtenir la desviació estàndard, premem la tecla  **$\sigma_n$** .
- També podem obtenir altres resultats parcials:

Nombre de dades: **n**

Suma de totes les dades:  **$\Sigma x$**

Suma dels quadrats de les dades:  **$\Sigma x^2$**



Uns detalls:

1. Entrar en el modus estadístic SD: **MODE** .(punt).
1. Introduir les dades: Des de teclat i prement **DATA** per a cada dada.
2. Obtenció de la mitjana aritmètica: Prement INV(SHIFT) i 7.
3. Obtenció de la desviació típica  $\sigma$ : Prement INV(SHIFT) i 8.
4. Sortir del modus SD: **MODE** 0(zero).

Aquesta calculadora és molt elemental. Els altres models -més moderns- de la mateixa marca tenen seqüències de tecles diferents per a l'obtenció d'aquestes dades. També hi ha altres marques (Texas Instruments, Hewlett Packard,...) que fan aquests càlculs amb la mateixa o superior eficiència.