

Statistique à deux variables : L'essentiel du cours

Pour une série :

Caractère x_i	x_1	x_2	\dots	x_n
Caractère y_i	y_1	y_2	\dots	y_n

a) Point moyen

Le point moyen G a pour abscisse la moyenne des x_i et pour ordonnée la moyenne des y_i .

b) Droite des moindres carrés

La droite des moindres carrés de y en x :

- a pour équation $y = ax + b$ où a et b sont donnés par la calculatrice
 - ▶ pour CASIO :
 -
 - Entrée des données : rentrer les valeurs x_i dans la liste 1 et les valeurs y_i dans la liste 2.
 - Affichage des résultats : →
 - Pour 2Var XList, choisir List 1
 - Pour 2Var YList, choisir List 2
 - Pour 2VarFreq, taper 1
 - Choisir , puis
 - On peut lire a et b dans la liste des résultats
 - ▶ pour TI :
 - Entrée des données : ; rentrer les valeurs x_i dans L1 et les valeurs y_i dans L2.
 - Affichage des résultats :
 - Xlist : L1 et Ylist : L2, puis Calculs
 - doit passer par le point moyen.

▶ Exemple :

x_i	1	2	3	4	5
y_i	8	9	12	12	14

- Point moyen : $\bar{x} = \frac{1+2+3+4+5}{5} = 3$; $\bar{y} = \frac{8+9+12+12+14}{5} = 11$

Donc, on a $G(3; 11)$

- Droite des moindres carrés : La calculatrice donne $a = 1,5$ et $b = 6,5$. Une équation de la droite des moindres carrés est donc : $y = 1,5x + 6,5$
- Estimation de la valeur de y pour $x = 7$: $y = 1,5 \times 7 + 6,5 = 17$

c) Coefficient de corrélation

Le coefficient de corrélation linéaire r :

- s'obtient à la calculatrice ;
- est toujours compris entre -1 et 1 ;
- plus les points sont « regroupés » autour d'une même droite et plus il est proche de 1 pour des données croissantes et de -1 pour des données décroissantes ;
- ne prouve en rien un lien de cause à effet entre les deux variables de la série statistique.