

Chapitre 13

Statistiques à deux variables

Table des matières

1	Nuage de points d'une série statistique à deux variables	2
2	Ajustement affine d'un nuage de points	2
3	Ajustement affine et changements de variables	3

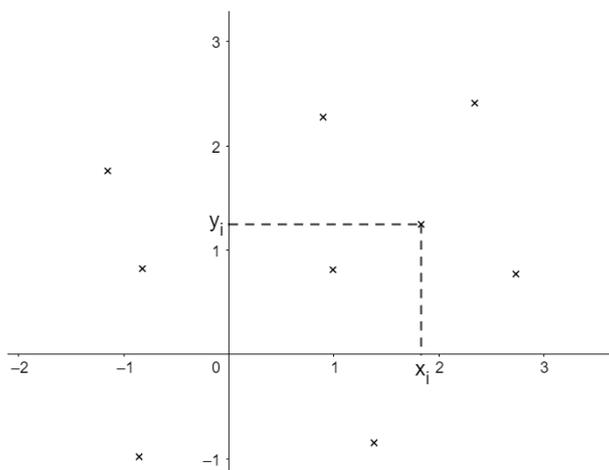
1 Nuage de points d'une série statistique à deux variables

Définition 1

Soient deux variables X et Y étudiées sur un même ensemble de N individus. On représente les relevés de valeurs dans un tableau.

X	x_1	x_2	\dots	x_N
Y	y_1	y_2	\dots	y_N

- Dans un repère du plan, le **nuage de points** associé à cette série statistique est l'ensemble des points de coordonnées (x_i, y_i) avec l'entier i compris entre 1 et N .
- Si la variable X correspond à des dates, la série est dite chronologique.



Exemple.

Dans un échantillon de N personnes, on peut considérer la variable X correspondant à la taille des individus et la variable Y correspondant à leur poids. On représente alors le poids (en ordonnée) en fonction de la taille (en abscisse).

2 Ajustement affine d'un nuage de points

Définition 2

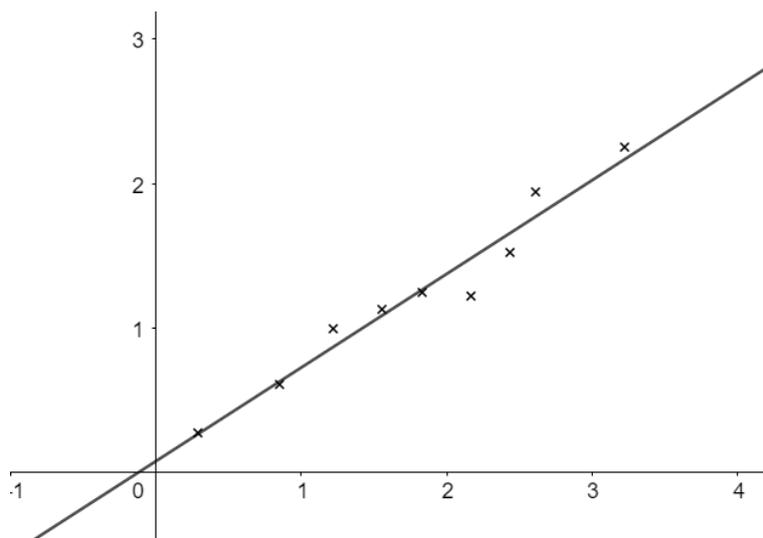
Déterminer un **ajustement affine** d'une série statistique de deux variables revient à trouver une fonction affine qui exprime de façon approchée les valeurs prises par Y en fonction de celles prises par X .

Méthode – Méthode pour déterminer un ajustement affine d'une série statistiques

1. On cherche la droite passant « au plus près » du nuage de points.
2. On détermine ensuite graphiquement l'équation de la droite (de la forme $y = ax + b$).

Exemple.

Sur le graphique ci-dessous, on a tracé une droite passant « au plus près du nuage de points ».



Graphiquement, on voit que l'ajustement affine correspond à la droite d'équation $y = ax + b$ avec $a \simeq 1,5$ et $b \simeq 0,1$

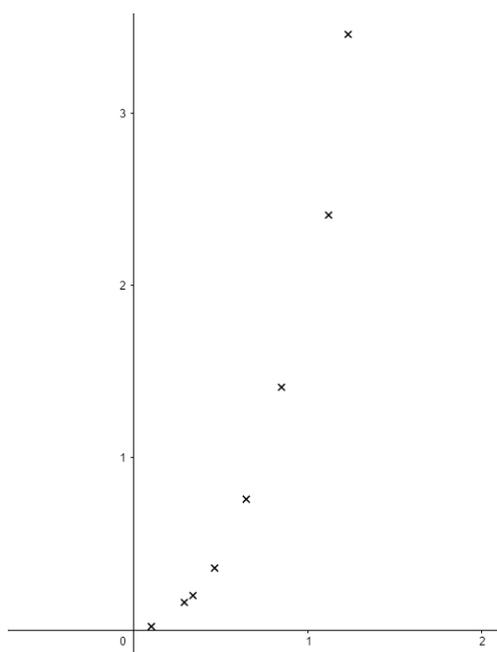
3 Ajustement affine et changements de variables

Parfois, il n'est pas pertinent de modéliser un nuage de points par une droite. On peut alors chercher à approcher le nuage par d'autres types de courbes (paraboles, exponentielles, logarithmes, etc.). On effectue alors un changement de variable afin de se ramener à une situation d'ajustement affine.

Exemple.

On considère les variables X et Y suivantes dont on donne le relevé de valeurs suivant.

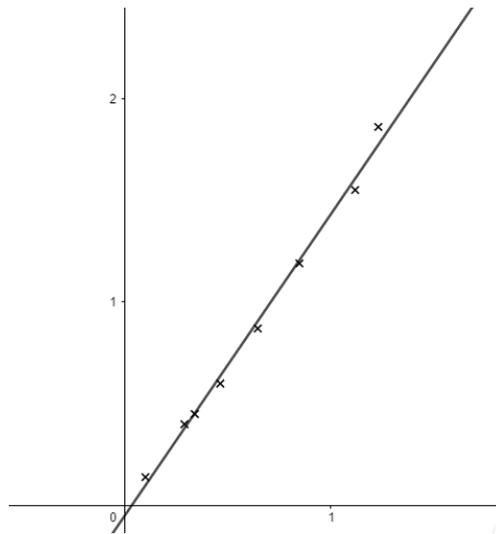
X	0,29	0,34	0,85	0,1	1,12	0,65	1,23	0,4
Y	0,16	0,2	1,41	0,02	2,4	0,76	3,46	0,36



Sur le graphique précédent, on voit qu'il n'est pas possible d'approcher le nuage de points par une droite mais plutôt par une parabole. On définit la variable Z telle que $Y = Z^2$ (autrement dit, $Z = \sqrt{Y}$). On a alors le tableau de valeurs suivant :

X	0,29	0,34	0,85	0,1	1,12	0,65	1,23	0,4
Z	0,4	0,45	1,19	0,14	1,55	0,87	1,86	0,6

On obtient alors le nuage de points suivant (Z en fonction de X).



On voit alors que l'on peut effectuer un ajustement affine. Par lecture graphique, on obtient l'équation suivante :

$$z \simeq 1,45x.$$

Autrement dit :

$$\begin{aligned}\sqrt{y} &\simeq 1,45x \\ y &\simeq (1,45x)^2 \\ y &\simeq 2,1x^2\end{aligned}$$