



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

Définitions

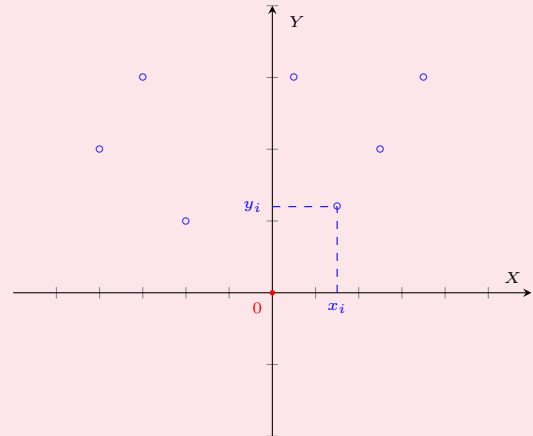
On considère deux variables quantitatives X et Y étudiées sur un même ensemble d'individus. On les représente dans un tableau dans lequel sont consignées N séries de relevés de valeurs.

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

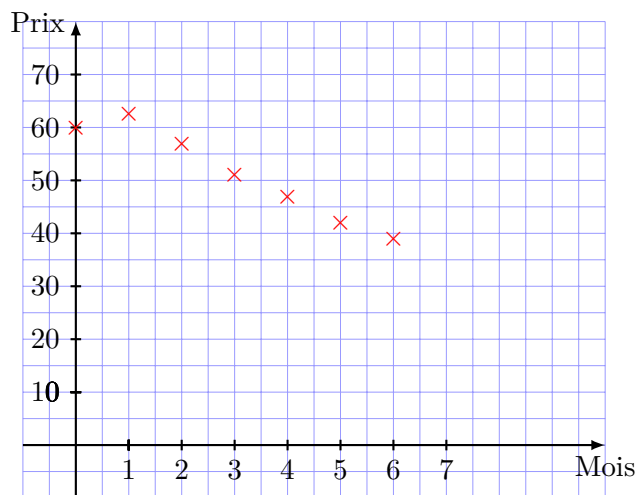
On appelle série statistique à deux variables, ou encore série statistique double, un ensemble de couples $(x_1; y_1)$, $(x_2; y_2)$, ...

Le nuage de points de cette série est l'ensemble des points de coordonnées $(x_i; y_i)$, avec $1 \leq i \leq N$.

Si la variable X correspond à des dates, la série est dite chronologique.



Exemple : Un comparateur de prix sur Internet a étudié l'évolution du prix d'un jeu vidéo dans une grande enseigne suivant le temps écoulé (en mois) après sa sortie. Le résultat de l'enquête peut être visualisé à l'aide du graphique ci-dessous.

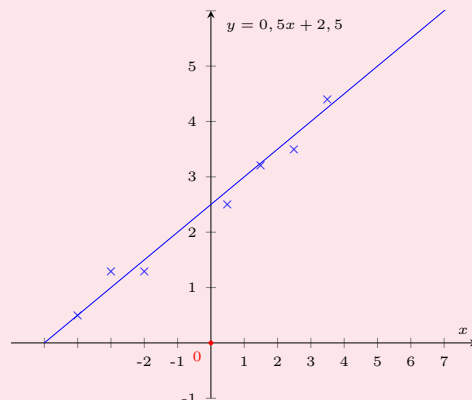


1. Quel est le prix ce jeu vidéo à sa sortie ?
.....
.....
2. Au cours des six premiers mois, quel est approximativement son prix le plus haut et quel est le plus bas ?
.....
.....
3. Donner la tendance générale d'évolution de ce prix. Si elle se poursuit, à quel prix peut-on estimer le jeu au 7e mois de sa sortie ?
.....
.....

2 Ajustement affine d'un nuage de points

Définitions

- Déterminer un ajustement affine d'une série statistique de deux variables revient à trouver une fonction affine qui exprime de façon approchée les valeurs y en fonction de celles de x .
- Une droite d'ajustement d'un nuage de points est une droite passant « au plus près » du nuage de points associé à cette série.



Propriété

Parmi toutes les droites d'ajustement que l'on peut tracer, il existe une unique droite qui est « le plus proche » des points du nuage.

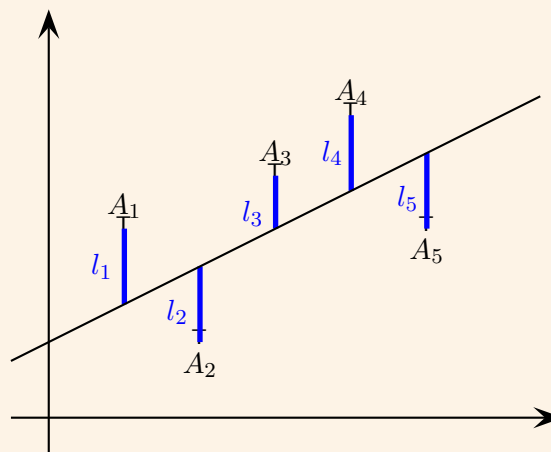
Plus précisément, pour le nuage de points $A_1(x_1; y_1)$, $A_2(x_2; y_2), \dots$ cette droite est celle telle que la somme des carrés des longueurs

$$l_1^2 + l_2^2 + l_3^2 + \dots$$

séparant les points du nuage de la droite soit la plus petite.

Cette droite s'appelle **la droite des moindres carrés**.

On détermine l'équation de la droite des moindres carrés à l'aide de la calculatrice (ou ordinateur, tableur, ...).



Méthodes : l'équation de la droite d'ajustement par moindres carrés en utilisant la calculatrice

- **TI** : `[Stat]` puis `[Edit]`, puis saisir les données dans L_1 et L_2 .
Pour le calcul de l'équation de la droite, `[Stat]`, puis `[Calc]` et enfin `RegLin(ax+b)`.
- **Casio** : Menu `[Stat]`, puis saisir les données dans $List1$ et $List2$.
Pour le calcul, `[Stat]` puis `[Set]`, et enfin choisir $List1$ pour 2Var Xlist et $List1$ pour 2Var Ylist.
Choisir alors `[Reg]`, puis `[X]` et enfin `aX+b`.

Exemple : On considère la série statistique à deux variables :

x_i	5	10	15	20	25	30	35	40
y_i	13	23	34	44	50	65	75	90

1. En utilisant la calculatrice, donner l'équation de la droite d'ajustement.
.....
2. Estimer la valeur de y pour $x = 45$ (*extrapolation*).
.....
3. Estimer la valeur de x pour $y = 70$ (*interpolation*).
.....

3 Interpolation, extrapolation, prévision

À partir de l'équation de la droite d'ajustement (par moindres carrés), on peut estimer une valeur supplémentaire à l'intérieur de la plage de données : on **interpole** une donnée.

On peut aussi chercher une valeur à l'extérieur de la plage de données : on **extrapole** alors, ou fait une **prévision**.

