Cours de mathématiques

TGRH

Statistiques à deux variables

Année scolaire 2011-2012

mise à jour 21 septembre 2011

Table des matières

1 Tableaux de données, nuages de points

1.1 Tableaux de données

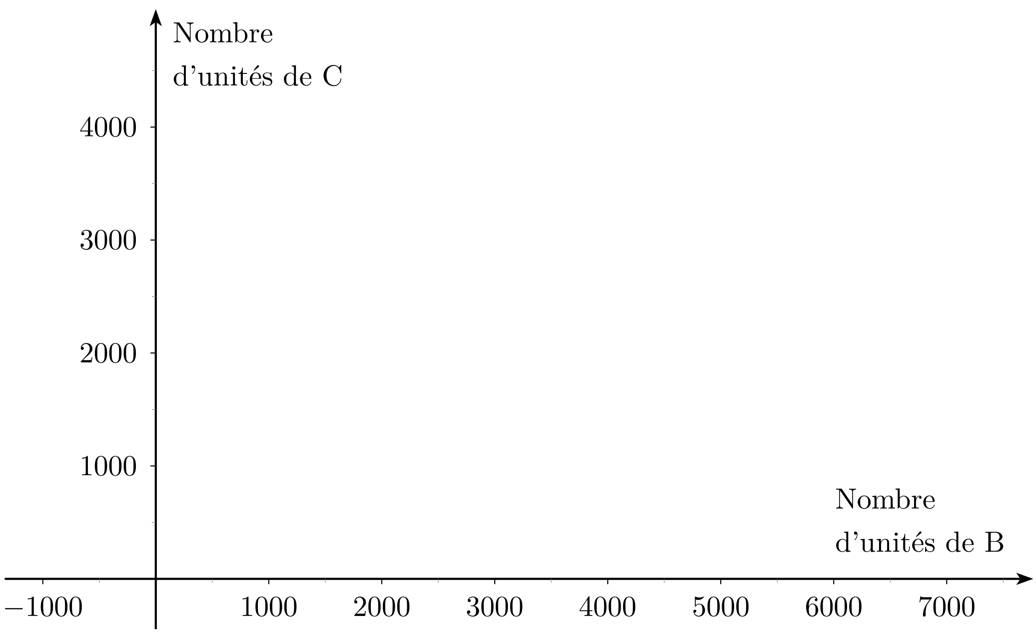
Exemple : Au cours du premier trimestre de cette année, une entreprise a lancé la commercialisation d'un accessoire C nécessaire à la pose de son produit B. On dispose des quantités vendues par zones de ventes :

Zones	Nombre d'unités de B vendues : x_i	Nombre d'unités de C vendues : y_i
1	4000	2400
2	2000	1200
3	6000	3000
4	3000	1500
5	3000	1200
6	6000	2700

1.2 Nuages de points

Le plan étant muni d'un repère, nous pouvons associer au couple $(x_i; y_i)$ de la série statistique double le point M_i de coordonnées $(x_i; y_i)$. L'ensemble des points M_i obtenus constitue re-présentant la série statistique.

Exemple :



Dans cet exemple, on peut essayer de trouver une fonction f tel que la courbe d'équation $y = f(x)$ passe le plus près possible des points du nuage. C'est **le problème de l'ajustement**. Ici, on peut penser qu'une droite $\mathcal{D}$ peut-être tracée au voisinage de ces six points. On dit que l'on fait un .

Remarque 1.1. *L'ajustement affine n'est pas toujours possible, il n'y a pas toujours un lien entre les x_i et les y_i. Un contre exemple peut-être donné en prenant la taille de plusieurs personnes et leurs nombres de frère et soeurs, par exemple...*

1.3 Points moyen

Définition 1 :
On appelle **point moyen** d'un nuage de n points M_i de coordonnées $(x_i; y_i)$ le point G de coordonnées :

$$\begin{cases} x_G & = \\ \\ y_G & = \end{cases}$$

Exemple : Dans l'exemple précédent :

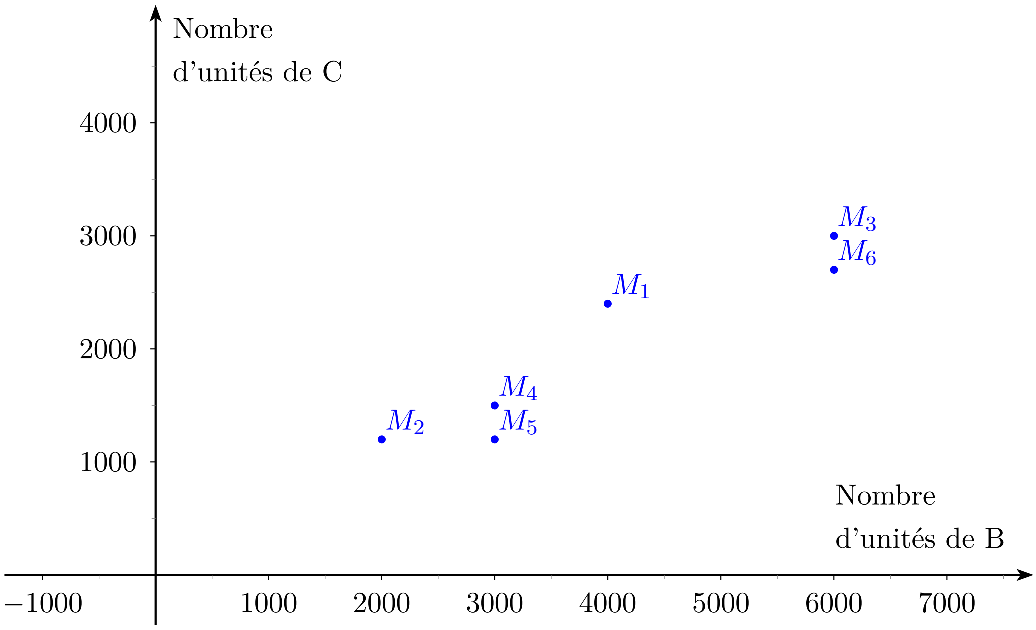
$$\begin{cases} x_G & = \\ \\ y_G & = \end{cases}$$

Donc

2 Ajustement affine

2.1 Méthode graphique

On reprend l'exemple précédent. On se propose, à partir des quantités vendues, de faire des prévisions de ventes de produit C pour d'autres chiffres de vente du produit B. Un moyen d'y parvenir est de tracer "au jugé" une droite $\mathcal{D}$ passant par le plus près possible des points et d'admettre que les chiffres de vente y_i du produit C et x_i du produit B sont liés par l'équation $y = ax + b$ de $\mathcal{D}$.

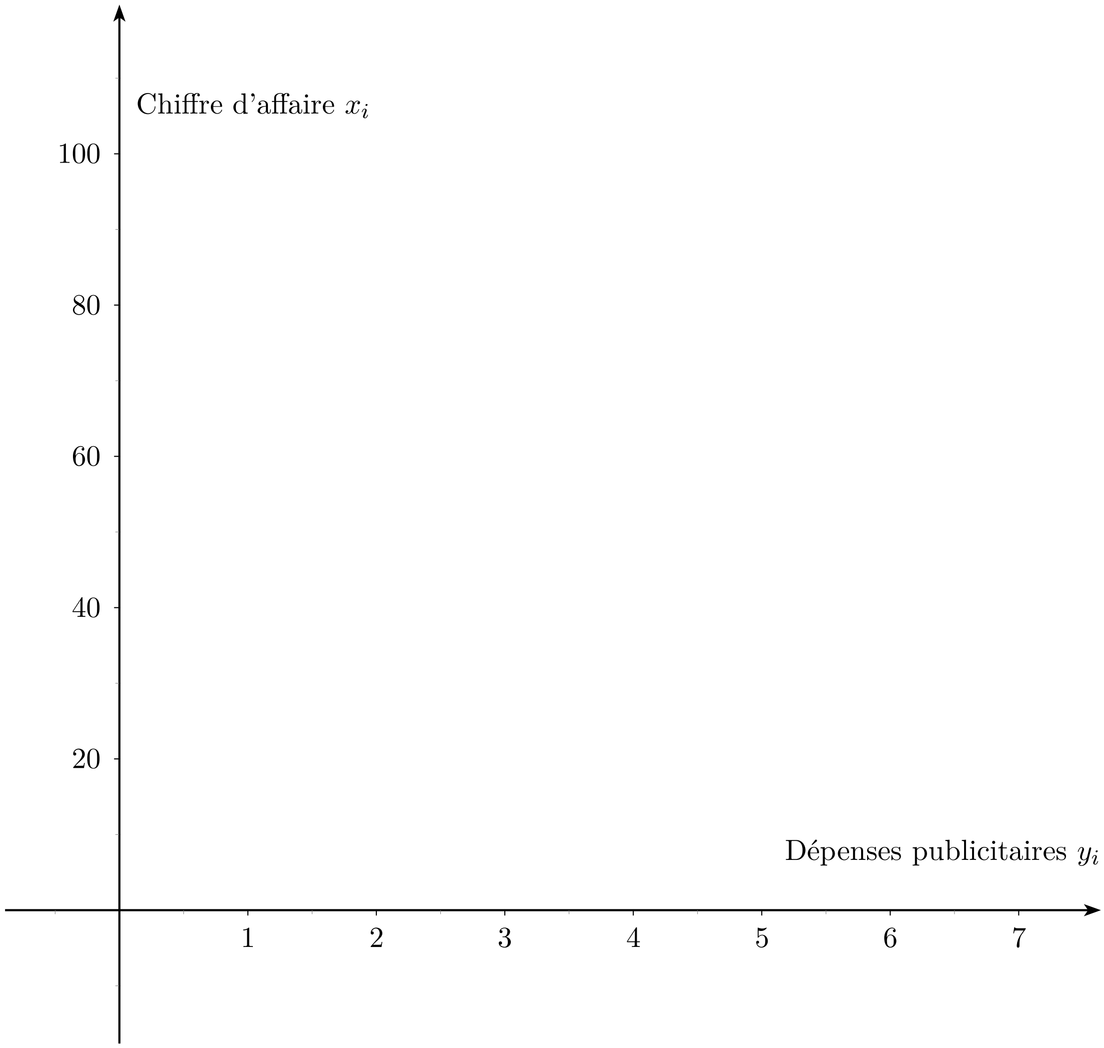


Remarque 2.1. *Aucun calcul n'est nécessaire mais chaque personne a une droite différente.*

2.2 Méthode des moindres carrés

Une entreprise s'intéresse au lien entre ses dépenses publicitaires et son chiffre d'affaires. Elle recueille les données suivantes exprimées en millions d'euros, portant sur 5 périodes où les dépenses publicitaires sont notées $x_1, \ldots, x_5$ et les chiffres d'affaires $y_1, \ldots, y_5$.

Dépenses publicitaires : x_i	0,5	2,0	2,9	4,5	5,6
Chiffre d'affaires : y_i	35	37	75	92	90



A l'aide d'une calculatrice ou d'un tableur, on détermine l'équation de la droite qui minimise les écarts du nuage de points à cette droite que l'on appelle .

Voici un bref descriptif de l'emploi de certaines calculatrices, pour plus d'informations consulter le manuel de l'utilisateur.

Avec une calcultrice TI 82 :

- **Effacement des colonnes précédentes :** Appuyer sur la touche $\boxed{\text{Stats}}$ puis choisir 4 : EffListe $\boxed{4}$ $\boxed{\text{2nde}}$ $\boxed{\text{L1}}$, $\boxed{\text{2nd}}$ $\boxed{\text{L2}}$ $\boxed{\text{Entrer}}$
- **Entrée des données :** Appuyer sur la touche $\boxed{\text{Stats}}$ puis choisir 1 : Edite... saisir les valeurs x_i et y_i dans les colonnes $L1$ et $L2$.
- **Regression linéaire :** Appuyer sur la touche $\boxed{\text{Stats}}$ puis dans $\boxed{\text{CALC}}$ choisir 4 : RegLin(ax+b) $\boxed{4}$ $\boxed{\text{L1}}$, $\boxed{\text{L2}}$ puis appuyer sur $\boxed{\text{Entrer}}$

Vous obtenez alors les valeurs de a et b telle que la droite d'équation $y = ax+b$ soit la droite de régression.

Avec une calcultrice Casio 25 :

- **Effacement des colonnes précédentes :** Selectionner le menu STAT puis choisir DEL-A $\boxed{>}$ $\boxed{\text{F2}}$ puis YES $\boxed{\text{F1}}$ et le refaire pour les autres colonnes.
- **Entrée des données :** Saisir les valeurs x_i et y_i dans les colonnes $L1$ et $L2$.
- **Paramétrage (uniquement pour la première fois) :** Aller dans le menu CALC $\boxed{\text{F2}}$ puis SET $\boxed{\text{F4}}$ 2 VAR X : List 1 2 VAR Y : List 2 $\boxed{\text{EXE}}$
- **Regression linéaire :** Aller dans le menu CALC $\boxed{\text{F2}}$ puis choisir REG $\boxed{\text{F3}}$ et X $\boxed{\text{F1}}$

Vous obtenez alors les valeurs de a et b telle que la droite d'équation $y = ax+b$ soit la droite de régression.

Avec l'exemple précédent, on obtient :

$$\begin{cases} a & = \\ b & = \end{cases}$$

et donc la droite de régression a pour équation :

$$y =$$

2.3 Estimations et prévisions

Cette relation nous permet d'estimer, par exemple, le montant du chiffre d'affaires associé à des dépenses publicitaires $y = 3,5$.

On trouve : $c =$

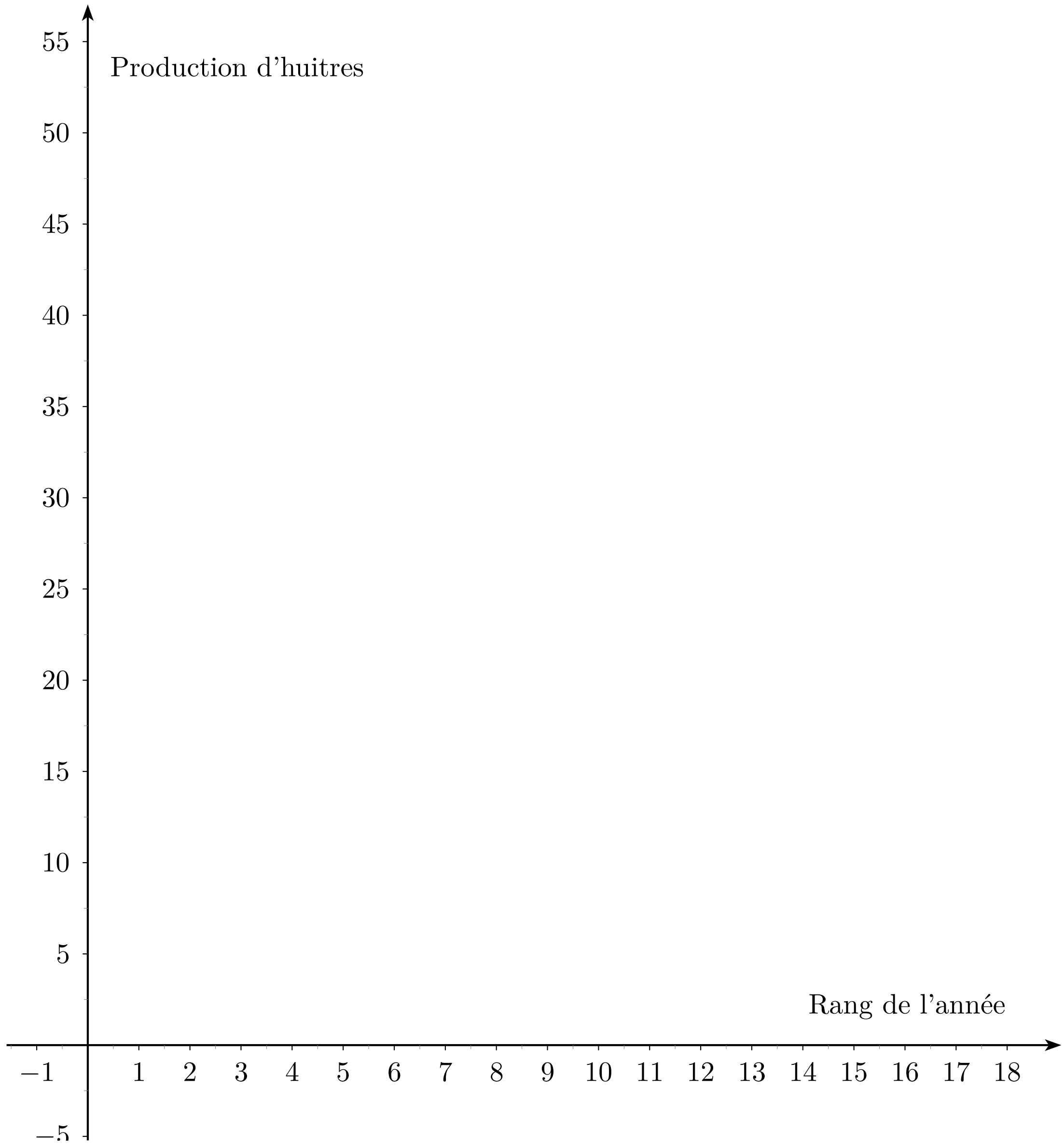
> **Définition 2 :**
> On appelle **série chronologique** une série où l'une des variables est le temps.

3 Exemple type bac

Le tableau ci-dessous récapitule la production en tonne, sur la côte atlantique, d'un ostréiculteur.

Année	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006
Rang de l'année : x_i	1	2	3	4	5	6	7	8	9	10
Production d'huitres : y_i	20	21,5	23,1	28	31,4	32,4	35	37,4	39,5	42

1. (a) Représenter le nuage de points de cette série. On prendra comme échelle 1 cm pour 1 an en abscisse et 1 cm pour 2 tonnes en ordonnées
   (b) Un ajustement affine est-il envisageable ?
2. Déterminer les coordonnées du point moyen G du nuage. Placer G sur le graphique.
3. Placer le point A de coordonnées $A(0; 16)$.
   On prendra alors comme droite d'ajustement la droite (AG).
4. Montrer que (AG) a pour équation réduite :
$$y = 2,73x + 16.$$
5. On suppose que la tendance observée se poursuit.
   (a) Calculer une estimation de la production en 2007.
   (b) Calculer l'année à partir de laquelle la production dépassera 57 tonnes.
6. Retrouver graphiquement les résultats de la question 5.


Production d'huitres
55
50
45
40
35
30
25
20
15
10
5
−5
Rang de l'année
−1
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18