

Chapitre 8 : Statistiques

Table des matières

Chapitre 8 : Statistiques	1
Axel CARPENTIER	
Contenu	2
1 Statistiques à deux variables	3
2 Exercice bilan	4

Contenu

- Nuage de points associé à une série statistique à deux variables quantitatives.
- Ajustement affine.

1 Statistiques à deux variables

Pour étudier une population selon deux caractères quantitatifs, on considère une série statistique à deux variables x et y composée de n observations $(x_1; y_1), (x_2; y_2), \dots, (x_n; y_n)$, chaque x_i correspondant au caractère x et chaque y_i correspondant au caractère y .

Exemple:

On interroge 8 personnes en leur demandant chacune leur taille et leur poids. On a donc une série statistique x relative à la taille et une série statistique y relative au poids.

Personne	A	B	C	D	E	F	G	H
Taille	165	167	169	171	173	174	175	178
Poids	68	73	71	72	70	75	82	85

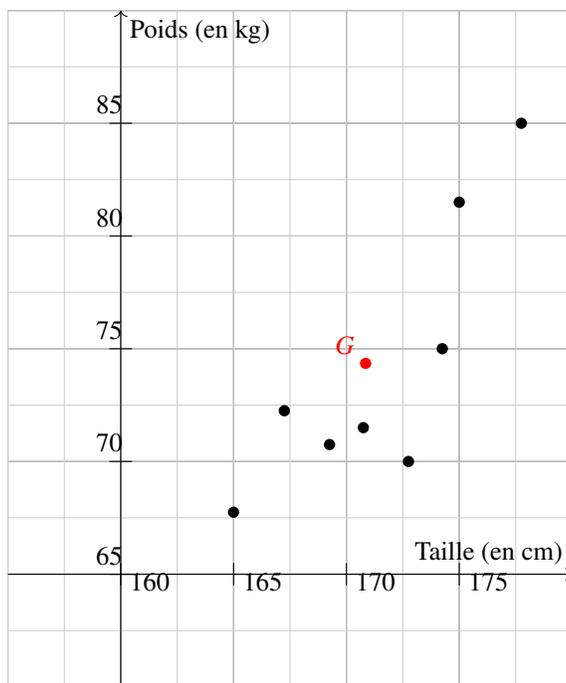
Définition:

A une série statistique on associe :

- Un nuage de points, qui est l'ensemble des n points $M_i(x_i; y_i)$ dans un repère du plan.
- Un point moyen dans le même repère : $G(\bar{x}; \bar{y})$ où \bar{x} et \bar{y} sont respectivement les moyennes des séries x et y .

Exemple:

En reprenant l'exemple précédent on a donc le nuage de point suivant :



En calculant les moyennes des séries statistiques taille (x) et poids (y), on obtient respectivement $\bar{x} = 171,5$ et $\bar{y} = 74,5$. On a donc le point moyen $G(171,5; 74,5)$.

Définition:

Le principe de l'ajustement est de chercher un lien éventuel et simple entre x et y .

Dans le cadre d'un ajustement affine, on cherche à lier x et y par une relation de la forme $y = ax + b$. On obtient alors une droite d'ajustement censée représenter le nuage de points.

Méthode:

Il y a plusieurs méthodes pour tracer la droite d'ajustement :

- "Au jugé", c'est-à-dire de la faire passer "au milieu" du nuage de points.
- En utilisant deux points données, appartenant ou non au nuage de points ou dont l'un des deux est le point G .
- A partir de l'équation de la droite d'ajustement, si elle est donnée.
- Si le coefficient directeur ou l'ordonnée à l'origine est donné et la droite d'ajustement passe par G ou un point du nuage.
- Par la méthode des moindres carrés via la calculatrice.

! Remarque

Si les points ne sont pas à peu près alignés, un ajustement affine n'a pas d'intérêt. Cependant, on peut faire un changement de variable (c'est-à-dire transformer les valeurs de x ou y) pour obtenir un nuage de points à peu près alignés.

Exemple:

En reprenant l'exemple précédent on a donc diverses méthodes pour obtenir l'ajustement affine de cette série statistique.

- **Méthode 1:** On suppose que la droite passe deux points G et $A(159; 60, 5)$.
On a alors son coefficient directeur $m = \frac{y_G - y_A}{x_G - x_A} = 1,12$.
On a son ordonnée à l'origine donné par $y_G = 1,12x_G + b \iff b = -117,58$.
- **Méthode 2:** On suppose que la droite passe par G en connaissant son ordonnée à l'origine.
On a $y_G = ax_G - 117,58 \iff a \approx 1,12$.
- **Méthode 3:** On suppose que la droite passe par G en connaissant son coefficient directeur.
On a $y_G = 1,12x_G + b \iff b \approx -117,58$.
- **Méthode 4:** Méthode des moindres carrés.
A la calculatrice, on a $a \approx 1,1212$ et $b \approx -117,7878$.

! Remarque

L'objectif d'un ajustement est de faire une estimation à partir de la relation trouvée entre x et y .

2 Exercice bilan

Le tableau suivant donne l'évolution des ventes de miel, en kg, dans une région pendant cinq années consécutives.

Rang de l'année : x_i	1	2	3	4	5
Masse des ventes : y_i	1 671	1 772	1 394	1 621	1 321

1. On considère le nuage de points $M_i(x_i; y_i)$ dans un repère orthogonal. Déterminer les coordonnées $(\bar{x}; \bar{y})$ du point moyen de ce nuage.
2. On ajuste le nuage de points précédent par la droite Δ d'équation $y = -87,5x + p$. Déterminer p pour que la droite Δ passe par le point G .
3. A l'aide de l'équation obtenue précédemment, estimer la masse des ventes l'année de rang 9. Arrondir à l'unité.
4. En utilisant la méthode des moindres carrés, déterminer la droite d'ajustement de cette série statistique.