Chapitre 1

Statistiques

Activité Cp38 (utilisation des listes avec la calculatrice)

A Série statistique à deux variables

Plutôt que simplement étudier un seul caractère sur une population, on peut s'intéresser à l'étude de deux caractères en même temps. On obtient alors deux listes de valeurs : x_i et y_i .

L'ensemble des couples $(x_i; y_i)$ forme une **série statistique double**.

1 Nuage de points et point moyen

On peut représenter dans un repère orthogonal l'ensemble des points de coordonnées $(x_i; y_i)$. On obtient alors ce que l'on appelle un **nuage de points**.

Le **point moyen** de ce nuage est le point $G(\bar{x}; \bar{y})$ dont les coordonnées sont les moyennes des deux séries :

$$\bar{x} = \frac{x_1 + \dots + x_n}{n} \quad \bar{y} = \frac{y_1 + \dots + y_n}{n}$$

Exemple : voir livre page 40.

2 Transformation affine

On peut changer chacune des deux séries en leur appliquant une fonction affine.

Soit $f : x \mapsto ax + b$ et $g : y \mapsto cy + d$.

Soit alors $z_i = f(x_i)$ et $t_i = g(y_i)$.

La double série des $(z_i; t_i)$ a alors pour point moyen le point G' de coordonnées $(a\bar{x} + b; c\bar{y} + d)$.

Exemple voir livre page 40.

→ **Exercices** 1,3,4p47 (moyennes, utilisation de la calculatrice ou non)

→ **Exercices** 11,12,13p48, 9p48

→ **Exercices** 15,17p49

B Ajustement affine

Activité 1 et 2 p39 (choix entre deux droites)

Dans certains cas, le nuage de points semble être formé de points presque alignés. On cherche alors à exprimer cela en effectuant un **ajustement affine**, c'est à dire à déterminer une fonction affine f dont la droite représentative est « la plus proche possible » de l'ensemble des points.

L'idée consiste à trouver la droite qui minimise les carrés des distances (verticales) des points à cette droite.

Dessin avec les distances verticales

On admet qu'il existe une unique droite vérifiant cette propriété et qu'elle passe par le point moyen du nuage.

On appelle cette droite la **droite de régression de y en x** , et on dit qu'elle est obtenue par la méthode des **moindres carrés**.

Son équation est de la forme $y = ax + b$, où a et b sont obtenus grâce à la calculatrice.

On donne les formules suivantes :

$$a = \frac{\text{cov}(x, y)}{V(x)} = \frac{\frac{1}{n} \sum (x_i - \bar{x})(y_i - \bar{y})}{\frac{1}{n} \sum (x_i - \bar{x})^2} \quad b = \bar{y} - a\bar{x}$$

L'intérêt d'une telle droite est par exemple de prévoir une valeur future (si les x_i sont des années).

→ **Exercices** 22,23,24p50

→ **Exercices** 26,27,28p51 (lire page 43 des méthodes)

→ **Exercice** 1p44 (ajustement par une parabole, utilisant une régression affine)