

Chapitre 2 – Cours

Statistiques bivariées : nuages de points et ajustements

Dans ce chapitre, tu vas apprendre à :

- Construire et interpréter un nuage de points.
- Calculer le point moyen et déterminer un ajustement affine simple.
- Distinguer interpolation, extrapolation et causalité.

Les mots pour comprendre

Définition – Un nuage de points

Chaque couple de mesures $(x_i; y_i)$ est représenté par un point. Le nuage permet d'observer une tendance sans prouver une causalité.

Définition – Le point moyen

Le point moyen possède pour coordonnées les deux moyennes : $G(\bar{x}; \bar{y})$. Il résume la position du nuage.

1. Observer deux caractères

Une série statistique bivariable associe deux informations à chaque individu : une durée et un coût, une année et un effectif, par exemple. Lorsque les caractères sont quantitatifs, on représente les couples $(x_i; y_i)$ par un nuage de points.

Les unités et les échelles des deux axes sont précisées ; l'ordre d'association entre les données doit être conservé. Pour deux caractères qualitatifs, un tableau croisé et des diagrammes adaptés permettent de comparer les groupes.

Dans ce chapitre, les données numériques sont fictives pour l'entraînement. Une étude réelle doit documenter les sources, la méthode de collecte et les limites des mesures.

Des données nombreuses ne garantissent pas l'absence de biais.

2. Point moyen

Le point moyen G a pour coordonnées les moyennes de chaque variable : $G(\bar{x}; \bar{y})$. On calcule séparément la moyenne des abscisses et celle des ordonnées, en tenant compte d'éventuels effectifs.

Pour les couples $(1; 3)$, $(2; 5)$, $(3; 7)$, le point moyen est $(2; 5)$. Le point moyen résume la position du nuage mais pas son étalement ni sa forme.

Deux nuages très différents peuvent posséder le même point moyen. Avant de calculer un modèle, il est donc utile de représenter les données pour vérifier si une tendance affine semble plausible ou si une courbure importante apparaît.

3. Ajustement affine

Un ajustement affine cherche une droite $y = ax + b$ qui approche les observations.

Une droite au jugé, une droite obtenue à partir de deux points représentatifs ou un ajustement fourni par un logiciel peuvent être utilisés. Pour une droite passant par deux points d'abscisses distinctes, $a = (y_2 - y_1)/(x_2 - x_1)$ puis $b = y_1 - ax_1$.

Le coefficient a mesure une variation de y par unité de x . Le terme b représente la valeur prévue pour $x = 0$, seulement si cette interprétation est pertinente dans le domaine.

Pour une droite de Mayer, on partage les observations ordonnées en deux groupes et on relie leurs points moyens. Un logiciel peut proposer la méthode des moindres carrés ; sa théorie n'est pas exigée ici.

4. Interpolation et extrapolation

L'interpolation estime une valeur pour une abscisse située entre les abscisses observées.

L'extrapolation prévoit une valeur en dehors de cet intervalle. Cette dernière est plus fragile : la tendance peut se modifier et un modèle affine peut conduire à des effectifs négatifs ou à des valeurs irréalistes à long terme.

Un résidu est la différence entre une valeur observée et la valeur prévue à la même abscisse. Examiner les résidus permet de discuter l'ajustement : une organisation courbe des écarts suggère qu'une droite ne décrit pas bien le phénomène.

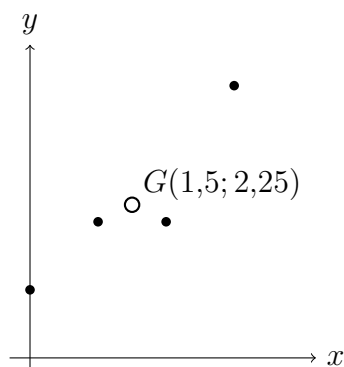
5. Décrire n'est pas expliquer

Une association statistique entre deux variables ne prouve pas une relation causale.

Une troisième variable, un changement de mesure ou une sélection des individus peut expliquer une partie de la liaison. On formule donc une conclusion limitée : le modèle décrit une tendance dans les données et sur un domaine précis.

Il ne transforme ni une prévision en certitude ni une corrélation en démonstration causale.

Voir la notion



Le point moyen G est entouré. Il ne représente pas nécessairement une observation réelle.

Deux exemples, pas à pas

Exemple résolu – 1 – Équation d'ajustement

Question. Déterminer la droite passant par $(2; 7)$ et $(5; 16)$.

Résolution. Avec $a = 3$, $7 = 3 \times 2 + b$ donne $b = 1$. L'équation est $y = 3x + 1$.

Exemple résolu – 2 – Point sur la droite

Question. Le point $(4; 11)$ appartient-il à la droite $y = 2x + 3$?

Résolution. Oui : $2 \times 4 + 3 = 11$.
