MATHS COMPLÉMENTAIRES / COURS

Statistiques à deux variables et ajustements

Astrolab Cours · astrolabcours.fr

01 Le problème

Une chaîne vidéo compte ses abonnés (en milliers) sur cinq années :

Année x_i	1	2	3	4	5
Abonnés y_i	12	15	17	20	21

Si la tendance continue, combien d'abonnés peut-on prévoir pour l'année 8 ?

L'idée : la ficelle tendue

Plante un clou à chaque point du nuage. Tends une **ficelle** droite au milieu des clous, pour qu'elle passe le plus près possible de tous à la fois. Elle ne touche peut-être aucun clou, mais elle résume la **tendance**.

La méthode des **moindres carrés** choisit cette ficelle avec un critère précis : rendre la plus petite possible la somme des carrés des écarts verticaux entre les points et la droite.

01 Nuage et point moyen

Une série statistique à deux variables $(x_i\,;y_i)$ se représente par un **nuage de points**.

Le **point moyen** est $G(\bar{x}\,;\bar{y})$, où $\bar{x}$ et $\bar{y}$ sont les moyennes des x_i et des y_i.

01 Droite des moindres carrés

La droite des moindres carrés $y = ax + b$ a pour coefficients

$$a = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2} \quad \text{et} \quad b = \bar{y} - a\bar{x}.$$

Elle **passe par le point moyen** G. En pratique, la calculatrice donne a, b et le **coefficient de corrélation** r : plus $|r|$ est proche de 1, meilleur est l'ajustement affine.

- **Interpoler** : estimer une valeur à l'intérieur des données (fiable).
- **Extrapoler** : prolonger hors des données (à prendre avec prudence).

Changement de variable : si les y_i semblent croître de façon exponentielle, on pose $z_i = \ln y_i$ et on ajuste z par une droite $z = ax + b$. On en déduit $y = e^b e^{ax}$.

© Astrolab Cours · astrolabcours.fr/maths/maths-complementaires/statistiques-a-deux-variables/ · Tous droits réservés : usage personnel uniquement, republication et usage commercial interdits

Retour au problème

$\bar{x} = 3$ et $\bar{y} = 17$. Les écarts donnent $\sum (x_i - \bar{x})(y_i - \bar{y}) = 10 + 2 + 0 + 3 + 8 = 23$ et $\sum (x_i - \bar{x})^2 = 10$.

Donc $a = 2{,}3$ et $b = 17 - 2{,}3 \times 3 = 10{,}1$: la droite est $y = 2{,}3x + 10{,}1$. Pour $x = 8$: $y = 28{,}5$.

On peut prévoir environ 28 500 abonnés l'année 8, à condition que la tendance se maintienne.

Pour revoir les bases : les statistiques de 2de. Pour les modèles exponentiels : exponentielle et logarithme.

© Astrolab Cours · astrolabcours.fr/maths/maths-complementaires/statistiques-a-deux-variables/ · Tous droits réservés : usage personnel uniquement, republication et usage commercial interdits