Statistique à deux variables : corrélation et causalité — Terminale maths complémentaires

20 exercices corrigés · mathlvl.fr

Énoncés et corrigés détaillés

Présentation

Deux grandeurs mesurées ensemble forment un nuage de points. Ces 20 exercices corrigés de maths complémentaires apprennent à calculer le point moyen, à déterminer et utiliser la droite des moindres carrés $y = ax + b$ (avec $a = \dfrac{\operatorname{cov}(x,y)}{V(x)}$ et $b = \overline{y} - a\overline{x}$), à juger un ajustement avec le coefficient de corrélation, à se ramener à un ajustement affine par le changement de variable $z = \ln y$, et surtout à ne pas confondre corrélation et causalité.

Version en ligne, avec les corrections dépliables : mathlvl.fr/exercices-corriges/terminale-maths-complementaires/statistiques-deux-variables

Exercice 1 – Point moyen

Série : $x = 1, 2, 3, 4, 5$ et $y = 3, 5, 6, 8, 9$. Calcule les coordonnées du point moyen G.

Exercice 2 – Coefficient directeur

Pour cette série, $\operatorname{cov}(x,y) = 3$ et $V(x) = 2$. Détermine la droite des moindres carrés.

Exercice 3 – Interpoler

Avec $y = 1{,}5x + 1{,}7$, estime y pour $x = 3{,}5$.

Exercice 4 – Signe de r

Pour chaque cas, donne le signe probable de r : a) taille et masse d'adultes ; b) prix d'un article et quantité vendue.

Exercice 5 – Dépenses de publicité

Une entreprise relève ses dépenses publicitaires x (en milliers d'euros) et son chiffre d'affaires y (en dizaines de milliers d'euros) : $x = 1$ à 6, $y = 12, 15, 17, 21, 22, 26$. La calculatrice donne $y = 2{,}714x + 9{,}333$ et $r \approx 0{,}993$.

1. Vérifie que la droite passe par le point moyen.

2. Estime le chiffre d'affaires pour 8 milliers d'euros de publicité. Est-ce une interpolation ?

Exercice 6 – Loi d'Ohm

Mesures d'un conducteur : $I = 1, 2, 3, 4$ (A) et $U = 2{,}1$; $3{,}9$; $6{,}2$; $7{,}8$ (V). La régression donne $U = 1{,}94I + 0{,}15$ avec $r \approx 0{,}998$. Quelle est la résistance ? Estime U pour $I = 2{,}5$ A.

mathlvl.fr — Tous droits réservés. Toute reproduction, diffusion, revente ou exploitation commerciale, même partielle, est strictement interdite.
Tout contrevenant s'expose à des poursuites en vertu du Code de la propriété intellectuelle (art. L.335-2 et suivants).

Exercice 7 – Lire la qualité d'un ajustement

Trois séries ont pour coefficients de corrélation 0,35, −0,97 et 0,82. Laquelle se prête le mieux à un ajustement affine ?

Exercice 8 – Retrouver b

La droite des moindres carrés a pour pente −0,8 et le point moyen est $G(5\,;12)$. Détermine son équation.

Exercice 9 – Résidus

Avec $y = 1{,}5x + 1{,}7$, calcule le résidu du point $(4\,;8)$.

Exercice 10 – Corrélation sans lien affine

Les points $(-2\,;4)$, $(-1\,;1)$, $(0\,;0)$, $(1\,;1)$, $(2\,;4)$ donnent $r = 0$. Peut-on dire que x et y n'ont aucun lien ?

Exercice 11 – Changement de variable

La concentration y d'un médicament est mesurée toutes les heures : 200, 150, 112, 85, 63, 47 (mg/L) pour $t = 0$ à 5.

1. Pourquoi poser $z = \ln y$?

2. L'ajustement donne $z = -0{,}289t + 5{,}300$. Exprime y en fonction de t.

3. Par quel facteur la concentration est-elle multipliée chaque heure ?

Exercice 12 – Prévoir avec un modèle exponentiel

Avec $y \approx 200 \times 0{,}75^{t}$, estime la concentration après 8 h et dis quand elle passe sous 20 mg/L.

Exercice 13 – Une relation en carré

On mesure $x = 1, 2, 3, 4$ et $y = 2, 8, 18, 32$. Quel changement de variable rend le nuage aligné ? Quelle relation obtient-on ?

Exercice 14 – Glaces et coups de soleil

Le nombre de glaces vendues et le nombre de coups de soleil soignés en pharmacie sont fortement corrélés ($r = 0{,}9$). Les glaces provoquent-elles des coups de soleil ?

Exercice 15 – Causalité inversée

On constate que les personnes qui consultent souvent un médecin sont plus souvent malades. Faut-il en conclure que consulter rend malade ?

mathlvl.fr — Tous droits réservés. Toute reproduction, diffusion, revente ou exploitation commerciale, même partielle, est strictement interdite.
Tout contrevenant s'expose à des poursuites en vertu du Code de la propriété intellectuelle (art. L.335-2 et suivants).

Exercice 16 – Extrapolation risquée

Le record du 100 m a baissé à peu près linéairement de 1912 à 2009. Une droite d'ajustement prévoit un temps nul vers l'an 3000. Qu'en penser ?

Exercice 17 – Établir une causalité

Comment démontrer qu'un engrais augmente le rendement d'un champ, au-delà d'une simple corrélation ?

Exercice 18 – Changement d'unité

On exprime y en grammes au lieu de kilogrammes (multiplication par 1 000). Que deviennent la pente a et le coefficient r ?

Exercice 19 – Pourquoi la droite passe par G

À pente a fixée, on pose $z_i = y_i - a x_i$ et $S(b) = \sum (z_i - b)^2$. Montre que S est minimale pour $b = \overline{y} - a\overline{x}$.

Exercice 20 – Problème complet : une application

Le nombre d'utilisateurs y (en milliers) d'une application est relevé chaque mois x. L'ajustement de $z = \ln y$ donne $z = 0{,}18x + 1{,}2$ avec $r = 0{,}995$.

1. Exprime y en fonction de x.

2. Quel est le pourcentage d'augmentation mensuel ?

3. Estime y pour $x = 12$. Peut-on prévoir le nombre d'utilisateurs dans 5 ans ?

mathlvl.fr — Tous droits réservés. Toute reproduction, diffusion, revente ou exploitation commerciale, même partielle, est strictement interdite.
Tout contrevenant s'expose à des poursuites en vertu du Code de la propriété intellectuelle (art. L.335-2 et suivants).

Corrigés détaillés

Les corrections des 20 exercices, dans l'ordre des énoncés.

Correction de l'exercice 1 – Point moyen

$\overline{x} = \frac{15}{5} = 3$.

$\overline{y} = \frac{31}{5} = 6{,}2$.

Réponse. $G(3\,;6{,}2)$.

Correction de l'exercice 2 – Coefficient directeur

$a = \frac{3}{2} = 1{,}5$.

$b = \overline{y} - a\overline{x} = 6{,}2 - 1{,}5 \times 3 = 1{,}7$.

Réponse. $y = 1{,}5x + 1{,}7$.

Correction de l'exercice 3 – Interpoler

$1{,}5 \times 3{,}5 + 1{,}7 = 5{,}25 + 1{,}7$.

Réponse. $y \approx 6{,}95$.

Correction de l'exercice 4 – Signe de r

a) Plus on est grand, plus on est lourd en moyenne : $r > 0$.

b) Plus le prix est élevé, moins on vend : $r < 0$.

Réponse. a) positif ; b) négatif.

Correction de l'exercice 5 – Dépenses de publicité

$\overline{x} = 3{,}5$, $\overline{y} = \frac{113}{6} \approx 18{,}83$; $2{,}714 \times 3{,}5 + 9{,}333 \approx 18{,}83$.

$2{,}714 \times 8 + 9{,}333 \approx 31{,}0$, soit environ 310 000 €. C'est une extrapolation ($8 > 6$).

Réponse. Passe par G ; environ 310 000 € (extrapolation).

Correction de l'exercice 6 – Loi d'Ohm

La pente est la résistance : $R \approx 1{,}94\ \Omega$, presque 2 Ω (l'ordonnée à l'origine vient des erreurs de mesure).

$U \approx 1{,}94 \times 2{,}5 + 0{,}15 = 5$ V.

Réponse. $R \approx 1{,}94\ \Omega$; $U \approx 5$ V.

Correction de l'exercice 7 – Lire la qualité d'un ajustement

On compare $|r|$ à 1 : 0,35, 0,97, 0,82.

Réponse. La série avec $r = -0{,}97$ (tendance décroissante).

Correction de l'exercice 8 – Retrouver b

$b = 12 - (-0{,}8) \times 5 = 12 + 4 = 16$.

Réponse. $y = -0{,}8x + 16$.

mathlvl.fr — Tous droits réservés. Toute reproduction, diffusion, revente ou exploitation commerciale, même partielle, est strictement interdite.
Tout contrevenant s'expose à des poursuites en vertu du Code de la propriété intellectuelle (art. L.335-2 et suivants).

Correction de l'exercice 9 – Résidus

Valeur prédite : $1{,}5 \times 4 + 1{,}7 = 7{,}7$.

Résidu $= 8 - 7{,}7 = 0{,}3$.

Réponse. Résidu 0,3.

Correction de l'exercice 10 – Corrélation sans lien affine

Les points vérifient exactement $y = x^2$.

r ne mesure que la tendance affine.

Réponse. Non : il y a un lien parfait, mais non affine.

Correction de l'exercice 11 – Changement de variable

Les valeurs semblent multipliées par un facteur constant : décroissance exponentielle, que le logarithme rend affine.

$y = \mathrm{e}^{5{,}300}\mathrm{e}^{-0{,}289t} \approx 200\,\mathrm{e}^{-0{,}289t}$.

$\mathrm{e}^{-0{,}289} \approx 0{,}749$: environ 25 % éliminés par heure.

Réponse. $y \approx 200\,\mathrm{e}^{-0{,}289t}$; facteur $\approx 0{,}75$ par heure.

Correction de l'exercice 12 – Prévoir avec un modèle exponentiel

$y(8) \approx 200 \times 0{,}75^8 \approx 20{,}0$ mg/L.

$200 \times 0{,}75^t < 20 \Leftrightarrow 0{,}75^t < 0{,}1 \Leftrightarrow t > \dfrac{\ln 0{,}1}{\ln 0{,}75} \approx 8{,}0$.

Réponse. Environ 20 mg/L après 8 h ; sous 20 mg/L juste après 8 h.

Correction de l'exercice 13 – Une relation en carré

Avec $X = x^2$: $X = 1, 4, 9, 16$ et $y = 2, 8, 18, 32$.

Les points $(X\,;y)$ sont alignés : $y = 2X$.

Réponse. $X = x^2$; $y = 2x^2$.

Correction de l'exercice 14 – Glaces et coups de soleil

Non : un facteur commun (l'ensoleillement, l'été) fait augmenter les deux.

La corrélation ne prouve pas une causalité.

Réponse. Non : facteur commun, l'ensoleillement.

Correction de l'exercice 15 – Causalité inversée

C'est l'inverse : on consulte parce qu'on est malade.

Réponse. Non : causalité inversée.

Correction de l'exercice 16 – Extrapolation risquée

Le modèle affine n'est valable que sur la période observée.

Un temps nul est absurde : les performances humaines ont des limites, la tendance finira par se stabiliser.

Réponse. Extrapolation absurde : le modèle ne vaut pas hors des données.

mathlvl.fr — Tous droits réservés. Toute reproduction, diffusion, revente ou exploitation commerciale, même partielle, est strictement interdite.
Tout contrevenant s'expose à des poursuites en vertu du Code de la propriété intellectuelle (art. L.335-2 et suivants).

Correction de l’exercice 17 – Établir une causalité

On réalise une expérience contrôlée : des parcelles comparables, tirées au hasard, reçoivent ou non l’engrais.

Le tirage au hasard équilibre les autres facteurs (sol, ensoleillement).

Réponse. Par une expérience avec groupe témoin tiré au hasard.

Correction de l’exercice 18 – Changement d’unité

$\text{cov}(x, y)$ est multipliée par 1 000 et $V(x)$ ne change pas : a est multipliée par 1 000.

r est sans unité : il ne change pas.

Réponse. $a \times 1\,000$; r inchangé.

Correction de l’exercice 19 – Pourquoi la droite passe par G

$S(b) = nb^2 - 2b\sum z_i + \sum z_i^2$: trinôme en b de coefficient $n > 0$.

Minimum en $b = \dfrac{\sum z_i}{n} = \overline{z} = \overline{y} - a\overline{x}$.

La droite passe alors par G : $a\overline{x} + b = \overline{y}$.

Réponse. $b = \overline{y} - a\overline{x}$: la droite passe par G.

Correction de l’exercice 20 – Problème complet : une application

$y = \text{e}^{1,2}\text{e}^{0,18x} \approx 3{,}32\,\text{e}^{0,18x}$.

$\text{e}^{0,18} \approx 1{,}197$: environ +19,7 % par mois.

$y(12) \approx 3{,}32\,\text{e}^{2,16} \approx 28{,}8$ milliers. Dans 5 ans, l’extrapolation n’a pas de sens : la croissance finira par plafonner.

Réponse. $y \approx 3{,}32\text{e}^{0,18x}$; +19,7 % par mois ; 28,8 milliers à 12 mois ; pas de prévision fiable à 5 ans.

mathlvl.fr — Tous droits réservés. Toute reproduction, diffusion, revente ou exploitation commerciale, même partielle, est strictement interdite.
Tout contrevenant s’expose à des poursuites en vertu du Code de la propriété intellectuelle (art. L.335-2 et suivants).