

Plaquette 1 - Nuage de points, ajustement affine et prévisions

Statistiques à deux variables — Terminale STMG

Corrections détaillées

Boîte à outils

Méthodes et formules du cours.

- Un **nuage de points** représente les couples $(x_i; y_i)$ dans un repère.
- **Point moyen** G : ses coordonnées sont les moyennes des deux séries,
$$G(\bar{x}; \bar{y}) \quad \text{avec} \quad \bar{x} = \frac{x_1 + \dots + x_n}{n}$$
- La **droite d'ajustement affine** obtenue par la méthode des moindres carrés a une équation de la forme $y = ax + b$ (coefficients donnés par la calculatrice). Elle passe **toujours** par le point moyen G .
- **Interprétation de a** : quand x augmente de 1, y varie de a en moyenne.
- **Prévision** : on remplace x par la valeur souhaitée dans l'équation.
- **Interpolation** : la valeur de x est à l'intérieur de la plage observée (prévision fiable) ;
- **extrapolation** : elle est en dehors (prévision à prendre avec prudence).
- Un **coefficient de corrélation** r proche de 1 ou de -1 indique que l'ajustement affine est pertinent ; proche de 0, il ne l'est pas.
- Ajustement affine ne veut pas dire **causalité** : la droite décrit une tendance observée, elle n'explique pas.

Correction de l'exercice 1 - Calculer le point moyen

On calcule séparément les deux moyennes.

- $\bar{x} = \frac{10 + 12 + 15 + 18 + 20}{5} = \frac{75}{5} = 15$.
- $\bar{y} = \frac{122 + 128 + 145 + 158 + 172}{5} = \frac{725}{5} = 145$.
- Le point moyen est donc $G(15; 145)$.
- Remarque : ce point appartiendra nécessairement à la droite d'ajustement.

Réponse. $G(15; 145)$.

Correction de l'exercice 2 - Lire une équation d'ajustement

a) On remplace x par 15 :

$$5 \times 15 + 70 = 75 + 70 = 145$$

- On obtient bien $\bar{y} = 145$: la droite passe par $G(15; 145)$, comme l'affirme le cours.

b) Le coefficient directeur 5 signifie que **chaque millier d'euros investi en publicité correspond, en moyenne, à 5 milliers d'euros de chiffre d'affaires supplémentaires** sur la période observée.

- L'ordonnée à l'origine 70 représente le chiffre d'affaires « théorique » sans publicité
 - à interpréter avec prudence, car $x = 0$ est loin des valeurs observées.

Réponse. a) $5 \times 15 + 70 = 145 = \bar{y}$ · b) +1 millier de publicité → +5 milliers de CA en moyenne.

Correction de l'exercice 3 - Faire une prévision

- a) $y = 5 \times 17 + 70 = 85 + 70 = 155$ milliers d'euros.
b) $y = 5 \times 30 + 70 = 150 + 70 = 220$ milliers d'euros.
c) La première : 17 est à **l'intérieur** de la plage observée (10 à 20), c'est une **interpolation**.
— La seconde est une **extrapolation** : 30 est bien au-delà des données. Rien ne garantit que la tendance se poursuive (saturation du marché, budget publicitaire moins efficace...). La prévision est donc peu fiable.

Réponse. a) 155 · b) 220 · c) la première (interpolation).

Correction de l'exercice 4 - Prévision inverse

On résout l'équation $5x + 70 = 180$.

- $5x = 180 - 70 = 110$.
— $x = \frac{110}{5} = 22$ milliers d'euros.
— Cette valeur est juste au-delà de la plage observée (10 à 20) : l'estimation reste plausible mais il s'agit déjà d'une légère extrapolation.

Réponse. 22 milliers d'euros.

Correction de l'exercice 5 - Nuage et pertinence de l'ajustement

- a) $r_1 = 0,98$ et $r_2 = -0,91$ sont proches de 1 en valeur absolue : l'ajustement affine est **pertinent** pour les séries 1 et 2. Avec $r_3 = 0,12$, les points ne s'alignent pas : l'ajustement affine n'a **pas de sens**.
b) $r_1 > 0$: quand x augmente, y augmente (relation croissante). $r_2 < 0$: quand x augmente, y diminue (relation décroissante). Pour la série 3, aucune tendance affine ne se dégage.

Réponse. a) séries 1 et 2 · b) croissante pour la 1, décroissante pour la 2, pas de tendance pour la 3.

Correction de l'exercice 6 - Ajustement d'une série chronologique

- a) $\bar{x} = \frac{1 + 2 + 3 + 4 + 5 + 6}{6} = \frac{21}{6} = 3,5$.
— $\bar{y} = \frac{210 + 231 + 246 + 266 + 281 + 302}{6} = \frac{1536}{6} = 256$.
— Point moyen : $G(3,5; 256)$. Contrôle : $18 \times 3,5 + 193 = 256$, la droite passe bien par G .
b) 2026 correspond au rang $x = 7$.
— $y = 18 \times 7 + 193 = 126 + 193 = 319$.
— On peut estimer environ 319 **adhérents** en 2026.
— Contrôle de cohérence : la tendance est de +18 adhérents par an et le dernier relevé vaut 302, donc $302 + 18 = 320$. L'ordre de grandeur est confirmé.

Réponse. a) $G(3,5; 256)$ · b) environ 319 adhérents.

Correction de l'exercice 7 - Changement d'unité

- a) Passer des milliers d'euros aux euros revient à multiplier toutes les valeurs de y par 1 000 : chaque coefficient est donc multiplié par 1 000.

$$y = 2\,400x + 15\,000$$

b) Non : le coefficient de corrélation mesure l'**alignement** des points, qui ne dépend pas des unités choisies. Il reste inchangé.

— À retenir : changer d'unité change les coefficients de la droite, jamais la qualité de l'ajustement.

Réponse. a) $y = 2\,400x + 15\,000$ · b) non, r est inchangé.

Correction de l'exercice 8 - Critiquer un modèle

a) Oui : $r = 0,95$ est très proche de 1, les points sont bien alignés.

b) $y = 32 \times 3 - 180 = 96 - 180 = -84$ glaces. Un nombre de glaces **négalif** est absurde : le modèle affine n'est valable que sur la plage de températures observée, vraisemblablement des températures estivales. En dehors, il perd tout sens.

c) Non. Une forte corrélation n'établit **pas** une causalité : ici le lien est plausible, mais la corrélation seule ne le démontre pas. D'autres facteurs (vacances scolaires, jour de la semaine) varient en même temps que la température.

Réponse. a) oui · b) -84 glaces : le modèle sort de son domaine de validité · c) non, corrélation n'est pas causalité.

Correction de l'exercice 9 - Comparer deux prévisions

a) $\bar{x} = \frac{1 + 2 + 3 + 4 + 5}{5} = 3.$

— $\bar{y} = \frac{45 + 52 + 56 + 63 + 69}{5} = \frac{285}{5} = 57.$

— Point moyen : $G(3; 57)$. Contrôle : $5,9 \times 3 + 39,3 = 17,7 + 39,3 = 57$. La droite passe bien par G .

b) Rang 6 : $5,9 \times 6 + 39,3 = 35,4 + 39,3 = 74,7$ milliers d'unités.

— Rang 12 : $5,9 \times 12 + 39,3 = 70,8 + 39,3 = 110,1$ milliers d'unités.

c) Celle de rang 6 : elle prolonge d'un an seulement la période observée.

— La prévision de rang 12 suppose que la tendance linéaire se maintienne pendant 7 ans de plus, ce qui est très incertain (cycle de vie du produit, concurrence). À mentionner, mais avec une réserve explicite.

Réponse. a) $G(3; 57)$ · b) 74,7 puis 110,1 · c) celle de rang 6.

Correction de l'exercice 10 - Ajustement après transformation

— a) Une croissance de 10 % par an signifie que y est **multiplié** par 1,10 chaque année : les écarts augmentent d'année en année. Un nuage de points prendrait une allure incurvée vers le haut, et une droite le décrirait mal : elle sous-estimerait les valeurs récentes et surestimerait les anciennes.

— b) Il faut un modèle **exponentiel** : $y = y_0 \times 1,10^x$, cohérent avec une évolution en pourcentage constant.

— Indicateur de contrôle : si les quotients z d'une année sur l'autre sont à peu près constants, le modèle exponentiel convient ; si ce sont les **différences** qui sont constantes, le modèle affine convient.

Réponse. a) les écarts croissent, le nuage est incurvé · b) un modèle exponentiel $y = y_0 \times 1,10^x$.

Correction de l'exercice 11 - Synthèse : étude complète

a) $\bar{x} = \frac{21}{6} = 3,5$.

— $\bar{y} = \frac{84 + 79 + 77 + 72 + 68 + 65}{6} = \frac{445}{6} \approx 74,17$ €.

— Point moyen : $G(3,5; \approx 74,17)$.

b) Le coefficient $-3,8$ est **négalif** : le prix baisse en moyenne de 3,80 € par semestre sur la période observée.

c) Rang 8 : $y = -3,8 \times 8 + 87,47 = -30,4 + 87,47 = 57,07$ €.

— On résout $-3,8x + 87,47 < 50$: $-3,8x < -37,47$, donc $x > \frac{37,47}{3,8} \approx 9,86$.

— Le modèle prévoit un prix inférieur à 50 € à partir du rang 10.

— Réserve : au rang 10, on extrapole de quatre semestres ; par ailleurs un prix ne peut pas décroître indéfiniment (coût de production plancher).

Réponse. a) $G(3,5; \approx 74,17)$ · b) $-3,80$ € par semestre · c) 57,07 € au rang 8 ; sous 50 € à partir du rang 10.