

Ajustement linéaire.

Partiel mai 2023 (durée : 1h)

Consigne : La présentation, la clarté et la qualité de vos raisonnements tiendront une part importante dans la notation. Justifier vos calculs et ne pas hésiter à écrire la formule du cours que vous utilisez. Calculatrices et une feuille A4 de notes de cours autorisées.

Exercice 1 Le responsable d'un site internet s'intéresse au nombre de pages visitées sur son site chaque semaine. Le tableau ci-dessous donne le nombre de pages visitées, exprimé en milliers, durant chacune des quatre semaines suivant l'ouverture du site.

Semaine, X_i	1	2	3	4
Nb de pages visitées (en milliers), N_i	40	45	55	70

1. Tracer le nuage de points (X_i, N_i) . Préciser l'échelle choisie.
2. Calculer le coefficient de corrélation entre N et X . On s'attachera à détailler les calculs des quantités nécessaires (moyennes, covariance, variance...)
3. Effectuer l'ajustement linéaire de N en fonction de X . Tracer alors sur votre graphe la droite obtenue.
4. En supposant que cet ajustement reste valable pendant les deux mois qui suivent l'ouverture du site, donner une estimation du nombre de pages visitées au cours de la huitième semaine suivant l'ouverture du site.

Le responsable décide de mettre en place, au cours de la quatrième semaine suivant l'ouverture du site, une vaste campagne publicitaire afin d'augmenter le nombre de visiteurs du site. Il étudie ensuite l'évolution du nombre de pages du site visitées au cours des trois semaines suivant cette opération publicitaire. Le tableau ci-dessous donne le nombre de pages visitées au cours des sept semaines suivant l'ouverture du site.

Semaine, X_i	1	2	3	4	5	6	7
Nb de pages visitées (en milliers), N_i	40	45	55	70	95	125	175

5. Calculer le nouveau coefficient de corrélation linéaire entre N et X . Un ajustement linéaire est-il envisageable? Expliquer.
6. Compte tenu de l'allure du nuage, on se tourne vers un ajustement de la forme $N \approx \alpha e^{\beta X}$. Justifier que l'on doit alors considérer les variables $Y := \ln(N)$ et X pour répondre à cette question par un ajustement linéaire.

7. Calculer alors le coefficient de corrélation entre Y et X , puis effectuer cet ajustement (ie : déterminer α et β , garder au moins 2 décimales dans vos calculs)
8. Donner une estimation du nombre de pages visitées au cours de la huitième semaine suivant l'ouverture du site.
9. Combien de semaines auraient été nécessaires pour atteindre ce résultat sans campagne publicitaire?

Exercice 2 (Les questions sont indépendantes)

1. Un étudiant de gea en stage dans une agence immobilière, étudie la corrélation entre la superficie S et le prix P des appartements. Après la collecte de ces données sur des dizaines de transactions, il trouve un coefficient de corrélation $r(S, P)$ négatif. Que pensez vous de son calcul? Justifier avec soin votre argument!
2. Douze personnes sont inscrites à une formation. Au début de la formation, ces stagiaires subissent une épreuve A notée sur 20. A la fin de la formation, elles subissent une épreuve B de niveau identique, notée également sur 20. Connaissant toutes ces notes, quelle quantité peut on calculer pour examiner l'impact de cette formation par rapport au niveau initial des stagiaires?
3. Soient X, Y deux séries, et a, b des réels strictement positifs. Comment est transformé le coefficient de corrélation lorsqu'on multiplie les valeurs de la série X par a et celles de Y par b ? On pourra exprimer $cov(aX, bY)$ en fonction de $cov(X, Y)$, puis $r(aX, bY)$ en fonction de $r(X, Y)$.

Bonus : Traiter le cas où a et b sont de signes quelconques (mais non nuls).

4. On considère la série suivante :

Temps t	1	2	3	4	5	6	7
X_t	10	45	30	40	24	44	72

Calculer la moyenne mobile $M_3^{(4)}$ au temps $t = 3$ d'ordre 4, puis celle au temps $t = 5$ d'ordre 3, soit $M_5^{(3)}$.

5. Soit (Y_t) une série chronologique. La tendance générale (g_t) se calcule à l'aide des moyennes mobiles. Quel doit être l'ordre de la moyenne mobile?

