

Corrélation et régression linéaire – Exercices - Devoirs

QCM 1 corrigé disponible

A propos de la corrélation et de la régression linéaire, indiquer si les propositions suivantes sont vraies ou fausses.

- A. La corrélation et la régression linéaire permettent d'étudier l'association entre deux variables qualitatives binaires
- B. Si X et Y sont exprimées en kg, le coefficient de corrélation ρ est également exprimé en kg
- C. Dans la régression linéaire, le test de la pente permet de tester l'existence d'une association linéaire significative entre X et Y
- D. L'hypothèse nulle du test de significativité du coefficient de corrélation linéaire de Pearson est que le coefficient de corrélation $\rho=0$ dans la population dont est issu l'échantillon
- E. Si le coefficient de corrélation entre X et Y est égal à 0, alors les variables X et Y sont indépendantes

QCM 2

On s'intéresse à l'évolution de la concentration en bactéries dans le temps pour une population au cours des 1000 prochaines heures ; on réalise une mesure sur les 5 premières heures

Temps T en heures	0	1	2	3	4	5
Concentration C en mmol.L^{-1}	5	7	9	11	13	15

On donne $t_{(4; \alpha=0,025)} = 2,776$

- A. Le coefficient de corrélation de l'échantillon vaut $r \approx -0,5$
- B. On peut estimer avec un risque $\alpha = 5\%$ qu'il y a corrélation entre T et C dans la population générale

- C. L'équation de la droite de régression de l'échantillon peut s'écrire $C = -0,5T + 5$
- D. L'équation de la droite de régression estimée sur la population peut s'écrire $C = 2T + 5$ (coefficients arrondis à l'unité)
- E. La concordance des variables C et T permet d'extrapoler $C = 40 \text{ mmol.L}^{-1}$ dans la population générale au bout de 20 heures

QCM 3 corrigé disponible

Une étude s'intéresse à comprendre ce qui influence la satisfaction du patient lors d'un entretien médical, mesurée grâce à un score quantitatif, le MISS (Medical Interview Satisfaction Scale). On souhaite savoir si la satisfaction est liée au nombre de gestes du médecin. L'analyse de régression linéaire du score MISS en fonction du nombre de gestes du médecin a montré une association entre les deux variables avec une valeur estimée de la pente de 0,11 et une p-value associée au test de la pente de $p=0,018$.

Indiquer si les propositions suivantes sont vraies ou fausses.

- A. Dans cette étude, on pourrait calculer la valeur du coefficient de corrélation entre le score MISS et le nombre de gestes du médecin
- B. Si le coefficient de corrélation entre le score MISS et le nombre de gestes du médecin est égal à 0, cela signifie que les deux variables sont indépendantes
- C. Dans l'analyse de régression linéaire évoquée dans l'énoncé, la satisfaction du patient (score MISS) serait la variable explicative et le nombre de gestes du médecin la variable réponse
- D. Pour chaque geste supplémentaire du médecin, le score MISS diminue en moyenne de 0,11 point

- E. Les résultats du test de la pente indiquent qu'il existe une relation linéaire positive significative entre la satisfaction du patient (score MISS) et le nombre de gestes du médecin au risque d'erreur 5%

QCM 4 corrigé disponible

Sur un échantillon de 8 femmes, tirées au sort dans une maternité, on a dosé une hormone dans le sang et dans le liquide amniotique.

Les résultats (en ng/mL) sont présentés dans le tableau ci-dessous :

Liquide amniotique	11	8	15	13	10	11	14	9
Sang	4,8	3,9	6,3	6,7	5,1	5,4	6,5	4,3

On souhaite savoir s'il existe une association entre les concentrations d'hormone dans le sang et dans le liquide amniotique.

Indiquer si les propositions suivantes sont vraies ou fausses.

- A. Pour répondre à la question posée, le test le plus approprié est un test du chi 2
- B. Pour évaluer l'existence d'une relation linéaire, on pourrait effectuer une régression linéaire
- C. Si le coefficient de corrélation linéaire entre les concentrations d'hormone dans le sang et dans le liquide amniotique est nul, cela signifie que les concentrations sont indépendantes
- D. L'hypothèse nulle du test du coefficient de corrélation linéaire s'écrit $H_0 : \rho = 0$
- E. Le coefficient de corrélation linéaire entre deux variables X et Y dépend de la variabilité de ces deux variables

QCM 5 corrigé disponible

A propos de corrélation et de la régression linéaire, indiquer si les propositions suivantes sont vraies ou fausses

- A. La corrélation linéaire permet de mesurer l'association entre deux variables continues
- B. Dans la régression linéaire, on suppose que la variable explicative est mesurée sans erreur
- C. Si le coefficient de corrélation linéaire entre X et Y est égal à 0, alors il se peut qu'il existe une relation non-linéaire entre X et Y
- D. On s'intéresse à la relation entre l'indice de masse corporelle (Y) et l'âge (X) que l'on souhaite décrire à l'aide d'une régression linéaire. On dispose de 3 échantillons différents de 30 individus chacun. La valeur de la pente estimée par la méthode des moindres carrés est la même dans chaque échantillon
- E. Pour savoir s'il y a une relation linéaire significative entre X et Y , l'hypothèse nulle à tester est que l'ordonnée à l'origine soit nulle