

RÉGRÉSSION ET CORRÉLATION

UE6

2022 - 2023

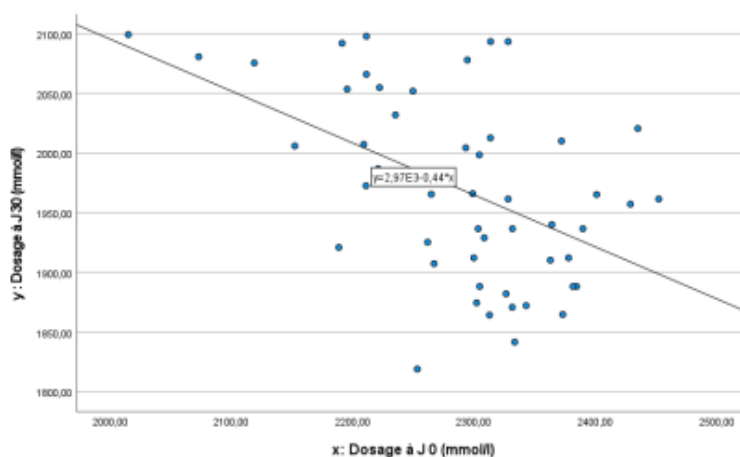
Quelle(s) est (sont) l'(les) affirmation(s) exacte(s) ?

- A. La régression linéaire simple concerne deux variables qui doivent être appariées et qui peuvent ne pas avoir la même unité
- B. La corrélation exprime une variable à expliquer en fonction d'un prédicteur
- C. Dans le modèle linéaire simple, l'erreur aléatoire correspond à l'écart au modèle
- D. La covariance est utilisée dans la corrélation mais pas dans la régression linéaire simple
- E. L'homoscédasticité des résidus doit être vérifiée dans la régression linéaire simple

2021 - 2022

Le dosage d'une substance sérique a été réalisé, chez 50 sujets atteints d'une maladie M, d'une part à J0 (dosage à J0 noté x représenté en x) et d'autre part à J30 (dosage à J30 noté y représenté en y). Entre J0 et J30, les sujets ont été traités par le traitement A.

La figure jointe représente le nuage de points correspondant aux valeurs recueillies et la droite de régression de y en fonction de x .



Quelle(s) est (sont) parmi les suivantes l'(les) affirmation(s) exacte(s) ?

- A. Le calcul du coefficient de corrélation linéaire associé au nuage de points donnera une valeur négative.
- B. La droite de régression permet de formaliser la comparaison entre les valeurs du dosage à J0 et les valeurs du dosage à J30.
- C. Le dosage à J0 et le dosage à J30 sont appariés.
- D. Le calcul de la régression de y en fonction de x utilise la covariance (x,y) .
- E. La droite de régression de x en fonction de y donnera la même droite que celle représentée sur la figure.

2020 - 2021

Cochez pour chaque affirmation si elle est vraie ou fausse.

- A. Dans le modèle linéaire simple, l'erreur aléatoire correspond à l'écart au modèle.
- B. Dans une droite de régression, la valeur prédite par le modèle est unique pour chaque valeur du prédicteur.
- C. L'ordonnée à l'origine dans la droite de régression linéaire s'exprime avec l'unité du prédicteur.
- D. Un coefficient de corrélation linéaire négatif correspond à une covariance négative.
- E. La variance de la pente utilise la variance résiduelle de la régression.

2022 - 2023

A,C et E sont vrai

A. Vrai : Effectivement, une régression linéaire simple concerne deux variables appariées ayant ou non la même unité.

B. Faux : C'est le cas de la régression linéaire.

C. Vrai : Par définition.

D. Faux : La covariance est utilisée dans le cadre de la régression linéaire pour le calcul du "a" reliant $Y = aX + B$: $a = \text{covariance}(X,Y) / \text{variance}(X)$.

E. Vrai : Cf. cours.

2021 - 2022

A, B, C et D sont vrais

A. Vrai : Le coefficient de corrélation linéaire a le même signe que la pente de la droite, qui est donc négatif dans cet item.

B. Vrai : Avec un test de la pente à 0, on pourra dire si la relation observée avec la droite est significative ou non.

C. Vrai : En effet, ce sont les mêmes patients qui ont deux valeurs. Chaque patient a deux valeurs chacun, un à J0 et un à J30 (après le traitement), c'est donc des valeurs appariées.

D. Vrai : Pour calculer la droite de régression (y en fonction de x), on a besoin du calcul de la pente de la droite. Cette pente (appelée "a" tel que $y=ax+b$) se calcule $a = \text{Covariance}(X,Y) / \text{Variance}(X)$

E. Faux : La régression n'est pas interchangeable, x en fonction de y est différent de y en fonction de x

2020 - 2021

A, B, C et D sont justes.

A. Vrai : l'erreur aléatoire, qui correspond à ϵ , correspond à l'écart au modèle.

B. Vrai : par définition, la droite de régression est composée de valeurs uniques obtenues pour chaque valeur du prédicteur. Si un même prédicteur pouvait donner plusieurs valeurs prédites, on ne pourrait pas tracer de droite.

C. Vrai : l'ordonnée à l'origine est la valeur β , c'est la valeur qu'on obtient à l'origine de la droite, c'est-à-dire lorsque X vaut zéro (pour rappel, l'équation d'une droite de régression linéaire correspond à $Y = \alpha \cdot X + \beta$). Elle est de la même unité que le prédicteur, X .

D. Vrai : il suffit d'analyser la formule du coefficient de corrélation linéaire, pour rappel :

$$\rho = \text{COV}(X, Y) / \sqrt{\text{VAR}(X) \times \text{VAR}(Y)}$$

Une racine carrée étant dans tous les cas positive, c'est la covariance qui va définir le signe du coefficient de corrélation linéaire.

E. Faux : la variance résiduelle correspond à la variance due à des facteurs autres que ceux pris en compte. Elle ne rentre pas dans le calcul de la variance de la pente.