

Exercice n°1

1) On vous demande d'expliquer la variable X par la variable Y par un modèle de régression linéaire. La droite de régression a pour équation $X = \hat{a}_n + \hat{b}_n Y$

avec $\hat{a}_n = \bar{X}_n - \hat{b}_n \bar{Y}_n$ ou $\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$ et $\bar{Y}_n = \frac{1}{n} \sum_{i=1}^n Y_i$

et $\hat{b}_n = \frac{\sum_{i=1}^n (Y_i - \bar{Y}_n)^2}{\sum_{i=1}^n (X_i - \bar{X}_n)(Y_i - \bar{Y}_n)}$

Erreurs: • la droite de régression est celle de Y en fonction de X
 • pas la formule pour \hat{a}_n et \hat{b}_n

2) Pour déterminer l'équation de la droite de régression, une méthode est celle des moindres carrés qui consiste à minimiser, sur les valeurs y_i données par la droite de régression, la somme des carrés des distances entre les points (x_i, y_i) et la droite de régression.

On a donc $S(a, b) = \sum_{i=1}^n (y_i - (a + b x_i))^2$

et (\hat{a}_n, \hat{b}_n) est tel que $S(\hat{a}_n, \hat{b}_n) = \min_{(a, b)} S(a, b)$

Erreur: la méthode n'a pas été donnée, simplement le résultat!