

Exercices 001 - Solutions

GSF-6053

Hiver 2025

Énoncé

Un chercheur, en utilisant des données sur la taille des classes (CS) et les scores moyens de test provenant de 100 classes de troisième, estime la régression OLS suivante :

$$\widehat{TestScore} = 520.4 - 5.82 \times CS, \quad R^2 = 0.08, \quad SER = 11.5$$

Questions

- i. Une classe a 22 élèves. Quelle est la prédiction de la régression pour le score moyen de test de cette classe ?
- ii. L'année dernière, une classe avait 19 élèves, et cette année elle en a 23. Quelle est la prédiction de la régression pour le changement dans le score moyen de test de la classe ?
- iii. La taille moyenne des classes parmi les 100 classes est de 21.4. Quel est le score moyen des tests parmi les 100 classes ?
- iv. Quelle est l'écart-type des scores de test parmi les 100 classes ?

Solutions

- i. La prédiction de la régression est :

$$Test\hat{Score}|_{CS=22} = 520.4 - 5.82 \times 22 = 520.4 - 128.04 = 392.36$$

- ii. La prédiction du changement dans le score moyen est :

$$\begin{aligned}\Delta Test\hat{Score}|_{CS=19 \rightarrow 23} &= Test\hat{Score}|_{CS=23} - Test\hat{Score}|_{CS=19} \\ &= (520.4 - 5.82 \times 23) - (520.4 - 5.82 \times 19) \\ &= (520.4 - 133.86) - (520.4 - 110.58) \\ &= 386.54 - 409.82 = -23.28\end{aligned}$$

La taille de la classe a été augmentée de 4 élèves supplémentaires, donc on s'attend à ce que le score moyen diminue d'environ 23,28 points.

- iii. Le score moyen parmi les 100 classes est :

$$Test\hat{Score} = \hat{\beta}_0 + \hat{\beta}_1 \times CS = 520.4 - 5.82 \times 21.4 = 520.4 - 124.308 = 396.092$$

- iv. Pour déterminer l'écart-type des scores de test (s_Y), nous devons d'abord calculer la somme des carrés totaux (TSS) et la somme des carrés résiduels (RSS).

Étape 1 : Calcul de RSS

L'erreur standard de la régression (SER) est donnée par :

$$SER = \sqrt{\frac{RSS}{n - k}}$$

où :

— RSS est la somme des carrés résiduels,



- n est le nombre d'observations (100),
- k est le nombre de paramètres estimés (ici, 2 : β_0 et β_1).

Ainsi,

$$RSS = (n - k) \times SER^2 = (100 - 2) \times 11.5^2 = 98 \times 132.25 = 12\,961$$

Étape 2 : Calcul de TSS

Le coefficient de détermination (R^2) est lié à RSS et TSS par la formule :

$$R^2 = 1 - \frac{RSS}{TSS} \Rightarrow TSS = \frac{RSS}{1 - R^2} = \frac{12\,961}{1 - 0.08} = \frac{12\,961}{0.92} \approx 14\,113$$

Étape 3 : Calcul de la Variance des Scores de Test

La variance des scores de test (s_Y^2) est donnée par :

$$s_Y^2 = \frac{TSS}{n - 1} = \frac{14\,113}{99} \approx 142.56$$

L'écart-type des scores de test est donc :

$$s_Y = \sqrt{142.56} \approx 11.94$$

