

# Exercices 003 - Solution

GSF-6053

Hiver 2025

## Énoncé

Une régression des gains hebdomadaires moyens (AWE, mesurés en dollars) sur l'âge (mesuré en années) utilisant un échantillon aléatoire de travailleurs à temps plein diplômés universitaires âgés de 25 à 65 ans donne la régression suivante :

$$\widehat{AWE} = 696.7 + 9.6 \times \text{Age}, \quad R^2 = 0.023, \quad SER = 624.1$$

## Questions

- i. Expliquez ce que signifient les valeurs des coefficients 696.7 et 9.6.
- ii. L'erreur standard de la régression (SER) est 624.1. Quelles sont les unités de mesure pour la SER ?
- iii. Le  $R^2$  de la régression est 0.023. Quelles sont les unités de mesure pour  $R^2$  ?
- iv. Quelle sera la prédiction des gains pour un travailleur de 25 ans ? Pour un travailleur de 45 ans ?
- v. La régression donnera-t-elle des prédictions fiables pour un travailleur de 99 ans ? Pourquoi ou pourquoi pas ?
- vi. D'après ce que vous savez sur la distribution des gains, pensez-vous qu'il est plausible que la distribution des erreurs dans la régression soit normale ?
- vii. L'âge moyen dans cet échantillon est de 41.6 ans. Quelle est la valeur moyenne des AWE dans l'échantillon ?

## Solutions

- i. Le coefficient 696.7 est l'ordonnée à l'origine de la droite de régression. Plus l'ordonnée à l'origine est élevée, plus la position générale de la droite de régression est haute. Ce coefficient donne la valeur moyenne de AWE lorsque l'âge est égal à 0. Le coefficient 9.6 mesure l'effet marginal de l'âge sur AWE. En moyenne, chaque augmentation d'une année de l'âge entraîne une augmentation de 9.6 dollars dans AWE.
- ii. La SER est mesurée dans les mêmes unités que la variable  $y$ , c'est-à-dire en dollars par semaine.
- iii. Le  $R^2$  est défini comme un ratio où le numérateur et le dénominateur sont tous deux mesurés en dollars carrés par semaine. Les unités se simplifient donc, et le  $R^2$  est un nombre pur, sans unité.
- iv. La régression prédit :

$$696.7 + 9.6 \times 25 = 936.7 \quad \text{pour un travailleur de 25 ans}$$

$$696.7 + 9.6 \times 45 = 1,128.7 \quad \text{pour un travailleur de 45 ans}$$

- v. Une régression peut ne pas donner de prédictions fiables pour des valeurs de  $X$  en dehors de l'échantillon. Dans ce cas, 99 ans est bien au-delà de l'échantillon, puisque la personne la plus âgée dans l'échantillon est âgée de 65 ans.
- vi. La distribution des gains elle-même n'est pas normalement distribuée ; elle est fortement biaisée positivement.
- vii. L'âge moyen dans l'échantillon étant de 41.6, on peut calculer la valeur moyenne des AWE dans l'échantillon :

$$696.7 + 9.6 \times 41.6 = 1,096.06$$

