

Exercice 1: Analyse de la relation entre l'âge et le revenu d'une population

On considère le modèle linéaire suivant :

$$y_i = \beta_0 + \beta_1 x_i + u_i$$

où y_i représente le revenu annuel (en milliers d'euros) de l'individu i , x_i est l'âge de l'individu, et u_i est le terme d'erreur non observable. Vous disposez de données pour un échantillon de $n = 15$ individus et des informations suivantes :

$$\sum_{i=1}^{15} x_i = 550, \quad \sum_{i=1}^{15} y_i = 300, \quad \sum_{i=1}^{15} (x_i - \bar{x})(y_i - \bar{y}) = 900, \quad \sum_{i=1}^{15} (x_i - \bar{x})^2 = 400$$

et

$$\sum_{i=1}^{15} u_i^2 = 40, \quad \sum_{i=1}^{15} (y_i - \bar{y})^2 = 120$$

Répondez aux questions suivantes :

1. Écrivez les expressions des estimateurs $\hat{\beta}_1$ et $\hat{\beta}_0$ obtenus par la méthode des moindres carrés ordinaires (MCO).

Les formules des moindres carrés ordinaires pour $\hat{\beta}_1$ et $\hat{\beta}_0$ sont :

$$\hat{\beta}_1 = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2}$$
$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$$

2. Calculez les valeurs de $\hat{\beta}_1$ et $\hat{\beta}_0$ en utilisant les données fournies.

$$\hat{\beta}_1 = \frac{900}{400} = 2.25$$

$$\bar{x} = \frac{550}{15} = 36.67, \quad \bar{y} = \frac{300}{15} = 20$$

$$\hat{\beta}_0 = 20 - 2.25 \cdot 36.67 = 20 - 82.5 = -62.5$$

3. Quelle est l'interprétation économique de $\hat{\beta}_1$ dans le contexte de cette étude ?

Chaque année supplémentaire d'âge augmente le revenu annuel de 2 250 euros (2.25 milliers d'euros).

4. Un individu a 50 ans et un revenu observé de 25 000 euros. Calculez le résidu associé à cet individu.

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 \cdot 50 = -62.5 + 2.25 \cdot 50 = 50$$

Le résidu est donné par :

$$\hat{u}_i = y_i - \hat{y}_i = 25 - 50 = -25$$

Exercice 2 : Étude de la consommation énergétique des ménages

Vous disposez d'un échantillon de 500 ménages et vous étudiez leur consommation annuelle d'énergie (en milliers de kWh). Vous avez les variables suivantes :

- *ConsoEnerg* : Consommation d'énergie annuelle (en milliers de kWh),
- *RevMen* : Revenu annuel du ménage (en milliers d'euros),
- *TailLog* : Taille du logement en m²,
- *ChaufElec* : Variable binaire indiquant si le ménage utilise un chauffage électrique (1 = oui, 0 = non).

Le modèle estimé est le suivant :

$$\widehat{\text{ConsoEnerg}} = 200 + 1.5 \cdot \text{RevMen} + 0.3 \cdot \text{TailLog} + 50 \cdot \text{ChaufElec}$$

Répondez aux questions suivantes :

1. Interprétez les coefficients estimés associés à *RevMen* et à *TailLog* dans ce modèle. Que nous apprennent-ils sur la consommation énergétique ?
 - 1.5 (RevenuMenage) : Cela signifie qu'une augmentation du revenu du ménage de 1 000 euros est associée à une augmentation de la consommation énergétique de 1 500 kWh par an, toutes les autres variables étant constantes.
 - 0.3 (TailleLogement) : Chaque mètre carré supplémentaire du logement augmente la consommation énergétique de 0.3 millier de kWh par an, toutes choses étant égales par ailleurs.
2. Estimez la consommation d'énergie pour un ménage dont le revenu est de 40 000 euros, avec un logement de 100 m² et utilisant un chauffage électrique.

$$\widehat{\text{ConsoEnerg}} = 200 + 1.5 \cdot 40 + 0.3 \cdot 100 + 50 \cdot 1$$

$$\widehat{\text{ConsoEnerg}} = 200 + 60 + 30 + 50 = 340 \text{ milliers de kWh}$$

Ce ménage consommera environ 340 milliers de kWh d'énergie par an.

3. Effets directs

- (a) Quel est l'impact marginal d'une augmentation de 10 m² de la taille du logement sur la consommation d'énergie ?

$$\Delta \widehat{\text{ConsoEnerg}} = 0.3 \cdot 10 = 3 \text{ milliers de kWh}$$

Une augmentation de 10 m² entraîne une augmentation de 3 milliers de kWh de la consommation d'énergie.

- (b) Quel est l'impact marginal d'une augmentation de 1 000 euros du revenu du ménage sur la consommation d'énergie.

$$\Delta \widehat{\text{ConsoEnerg}} = 1.5 \cdot 1 = 1.5 \text{ milliers de kWh}$$

Chaque augmentation de 1 000 euros du revenu augmente la consommation d'énergie de 1.5 milliers de kWh.

4. Quel est l'effet conjoint d'une augmentation de 5 000 euros du revenu et d'une augmentation de 20 m² de la taille du logement sur la consommation d'énergie ?

$$\widehat{\Delta \text{ConsoEnerg}} = 1.5 \cdot 5 + 0.3 \cdot 20 = 7.5 + 6 = 13.5 \text{ milliers de kWh}$$

L'augmentation conjointe entraînera une hausse de 13.5 milliers de kWh de la consommation d'énergie.

5. Que se passera-t-il si l'on multiplie la variable *ConsoEnerg* par 1 000 pour l'exprimer en Wh au lieu de kWh ? Le R^2 sera-t-il impacté par ce changement d'unité ? Justifiez votre réponse.

Si *ConsoEnerg* est exprimée en Wh (plutôt qu'en kWh), tous les coefficients seront multipliés par 1 000. Par exemple, le coefficient de *RevMen* deviendra 1500 (au lieu de 1.5).

Supposons que nous multiplions Y par une constante c , c'est-à-dire $Y'_i = cY_i$.

Effet sur la SST :

$$\text{SST}' = \sum_{i=1}^n (Y'_i - \bar{Y}')^2 = \sum_{i=1}^n (cY_i - c\bar{Y})^2 = c^2 \sum_{i=1}^n (Y_i - \bar{Y})^2 = c^2 \text{SST}$$

Effet sur la SSR :

$$\text{SSR}' = \sum_{i=1}^n (\hat{Y}'_i - \bar{Y}')^2 = \sum_{i=1}^n (c\hat{Y}_i - c\bar{Y})^2 = c^2 \sum_{i=1}^n (\hat{Y}_i - \bar{Y})^2 = c^2 \text{SSR}$$

Effet sur le R^2 :

$$R'^2 = \frac{\text{SSR}'}{\text{SST}'} = \frac{c^2 \text{SSR}}{c^2 \text{SST}} = \frac{\text{SSR}}{\text{SST}} = R^2$$

Donc, un changement dans les unités de la variable dépendante Y ne modifie pas le R^2 .

6. Attendez-vous que le coefficient du *RevMen* change si l'on omet la variable *TailLog* du modèle? Si oui, dans quelle direction et justifier votre réponse?

Omettre la variable *TailLog* pourrait entraîner un biais de spécification, car la taille du logement est probablement corrélée avec d'autres variables comme le revenu, ce qui fausserait les coefficients estimés des autres variables. Si le revenu du ménage est positivement corrélé à la taille du logement, donc le biais serait positif étant donné $\hat{\beta}_2 > 0$. Si le revenu du ménage est négativement corrélé à la taille du logement, donc le biais serait négatif étant donné $\hat{\beta}_2 > 0$.

7. Selon vous, quelles sont les autres variables contenues dans le terme d'erreur? Pensez-vous que ces variables sont corrélées avec *RevMen*?

- Équipements et appareils : La présence ou le nombre d'appareils énergivores comme des climatiseurs, des chauffages d'appoint, ou d'autres équipements peut influencer la consommation totale. Les ménages plus aisés sont plus susceptibles de posséder une gamme plus large d'appareils électroménagers (climatiseurs, chauffages d'appoint, sèche-linge, lave-vaisselle, etc.) qui consomment de l'énergie.

- Efficacité énergétique du logement : L'isolation du logement, les types de fenêtres, la qualité de l'équipement de chauffage, etc., peuvent influencer la consommation d'énergie. Les ménages à plus haut revenu pourraient avoir accès à de meilleurs équipements de chauffage, d'isolation, ou à des appareils électroménagers plus économes en énergie. Cela pourrait réduire leur consommation énergétique pour le même niveau de confort.
- Conditions climatiques locales : Par exemple, des températures hivernales plus froides pourraient entraîner une consommation d'énergie plus élevée. La variable conditions climatiques locales n'est généralement pas directement corrélée avec le revenu annuel, car le climat d'une région ne dépend pas des revenus des individus qui y vivent.