

Contrôle Continu d'Econométrie Appliquée-1h30min

Ulysse Lawogni & Olivier Donni
Université de Cergy-Pontoise
2018-2019

Exercice

On considère le modèle linéaire suivant :

$$y_i = \beta_0 + \beta_1 x_i + u_i, \text{ avec } u_i \sim \mathcal{N}(0, \sigma^2)$$

où u_i est l'erreur non observable sur l'individu i . L'économètre observe $n = 8$ individus et obtient les informations suivantes :

$$\sum_{i=1}^n x_i = 207, \sum_{i=1}^n y_i = 25.7, \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) = 5.813$$

$$\sum_{i=1}^n (x_i - \bar{x})^2 = 56.875, \sum_{i=1}^n \hat{u}_i^2 = 0.4347, \sum_{i=1}^n (y_i - \bar{y})^2 = 1.0288$$

On donne les quantiles de Student : $t_6^{0.975} = 2.447$, $t_7^{0.975} = 2.365$, $t_8^{0.975} = 2.306$, $t_6^{0.95} = 1.943$ et $t_7^{0.95} = 1.895$.

1. Donner les expressions des estimateurs $\hat{\beta}_1$ et $\hat{\beta}_0$ par moindres carrés ordinaires de β_1 et β_0 . Calculer leurs valeurs.
2. Donner l'expression de la variance de $\hat{\beta}_1$ puis en donner une estimation.
3. Déduisez une estimation $\hat{\sigma}_{\hat{\beta}_1}$ pour l'écart-type de l'estimateur de β_1 .
4. Le paramètre β_1 est-il significatif? Pourquoi?
5. Donner une estimation pour y quand $x = 28$.

Problème

Rappel : Les quantiles d'ordre 0.95, 0.975 et 0.995 d'une loi normale centrée réduite sont égaux à 1.63, 1.96 et 2.58. On rappelle aussi que la loi de Student converge vers la loi normale centrée réduite lorsque son degré de liberté tend vers l'infini ; ainsi pour un degré de liberté suffisamment grand, ses quantiles sont très proches de ceux d'une loi normale centrée réduite. On donne également les quantiles de la loi de Fischer : $F_{1,\infty}^{0.95} = 3.85$, $F_{2,\infty}^{0.95} = 3.01$, $F_{3,\infty}^{0.95} = 2.61$, $F_{4,\infty}^{0.95} = 2.38$ et $F_{5,\infty}^{0.95} = 2.22$.

Dans un échantillon de 1126 personnes de 25 à 55 ans, on estime le modèle linéaire liant le logarithme du taux de salaire ($ltsal$) en fonction du niveau d'éducation $educ$, de la variable d'âge age et de son carré $age2$ ainsi que d'une

variable indicatrice de la région parisienne, *paris* égale à 1 si l'individu observé habite la région parisienne et 0 si celui-ci n'y habite pas.

$$ltsal = 0.86 + 0.054 \cdot educ + 0.017 \cdot age - 0.00012 \cdot age^2 + 0.068 \cdot paris + \hat{u}$$

(0.20)
(0.0033)
(0.010)
(0.000012)
(0.025)

où les écarts-types des coefficients estimés figurent entre parenthèses en dessous des valeurs estimées. La somme des carrés des résidus *SCR* est égale à 91.72 ; la somme totale des carrés *SCT* est égale à 150.

1. Interprétez le coefficient de *educ* et de *paris*.
2. Donner l'expression du R^2 en fonction de *SCT* et de *SCR*. Calculer sa valeur.
3. Tester la significativité globale de ce modèle au seuil 5%.
4. Tester au seuil 5%, l'hypothèse $H_0 : \beta_{paris} = 0.10$ contre l'alternative $H_a : \beta_{paris} < 0.10$. Interprétez le résultat.
5. Est-ce que le coefficient de la variable *educ* est significatif ? Pourquoi ?
6. Tester au seuil 5%, l'hypothèse $H_0 : \beta_{age} = 0.02$ contre l'alternative $H_a : \beta_{age} \neq 0.02$.
7. Tester l'hypothèse que le fait de résider en région parisienne n'a aucun impact sur le taux de salaire contre l'alternative que cela l'impacte positivement.
8. Construire un intervalle de confiance de 95% pour l'estimateur du paramètre de la variable *educ*.
9. Donner une estimation de *ltsal* pour un individu de 30 ans, ayant 10 ans d'années d'éducation et résidant à Cergy-Pontoise.
10. Donner une estimation de la différence en *ltsal* entre un Parisien de 37 ans et un Marseillais de 42 ans, tous deux ayant le même niveau d'éducation.
11. Donner une estimation de la différence moyenne en niveau d'éducation entre un Francilien et un non Francilien ayant le même taux de salaire et le même âge. Précisez celui qui a le plus grand niveau d'éducation.