

Exercices pour le cours Econométrie et évaluations statistiques (Ch. VII).

Inférence II: Intervalle de confiance et probabilité critique – Test de Fischer

Exercice 1. Une régression peut être utilisée pour tester si le marché utilise efficacement toutes les informations dans la valorisation des actifs financiers. Soit return le taux de rendement d'un actif sur une période de 4 ans, de la fin de l'année 1990 à la fin de l'année 1994. L'hypothèse de marchés efficients dit que les rendements ne doivent pas être systématiquement liés aux informations connues en 1990. Si les caractéristiques de l'entreprise connues au début de la période permettaient de prédire les rendements des actifs, alors il serait possible d'utiliser ces informations pour investir dans certains actifs et faire des profits anormalement élevés. Soit dkr le ratio d'endettement sur fonds propres d'une entreprise en 1990, eps son résultat par action, netinc son résultat net, et salary la rémunération totale de son PDG. Les résultats de l'estimation sont les suivants:

$$\begin{aligned} \widehat{\text{return}} = & -14,37 + 0,321 \text{ dkr} + 0,043 \text{ eps} \\ & (6,89) \quad (0,201) \quad (0,078) \\ & - 0,005 \text{ netinc} + 0,003 \text{ salary} \\ & (0,005) \quad (0,002) \end{aligned}$$

$$N = 142, \quad R^2 = 0,039$$

1. Vérifiez si les variables explicatives sont conjointement significatives au seuil de 5%. Est-ce que certaines variables explicatives sont significatives séparément?
2. Réestimez le modèle en utilisant la forme logarithmique pour netinc et salary .

Les résultats sont les suivants:

$$\widehat{\text{return}} = - \begin{matrix} 36,30 \\ (39,37) \end{matrix} + \begin{matrix} 0,327 \text{ dkr} \\ (0,203) \end{matrix} + \begin{matrix} 0,069 \text{ eps} \\ (0,080) \end{matrix} \\ - \begin{matrix} 4,740 \text{ log(netinc)} \\ (3,395) \end{matrix} + \begin{matrix} 7,240 \text{ log(salary)} \\ (6,315) \end{matrix}$$

$$N = 142, \quad R^2 = 0,033$$

Certaines des conclusions de la partie (a) changent-elles? Dans l'ensemble, les preuves de la prévisibilité des rendements boursiers sont-elles fortes ou faibles?

3. Selon vous, pourquoi ne peut-on pas prendre la forme logarithmique pour les variables dkr et eps?

Exercice 2. Soit les données HPRICE1.xls sur les caractéristiques de 88 maisons. Pour tester la rationalité des évaluations des prix des maisons, le modèle suivant est utilisé:

$$\text{price} = \beta_0 + \beta_1 \text{assess} + u$$

où price représente le prix de la maison et assess l'évaluation faite par un expert (avant que la maison ne soit vendue). Si $\beta_0 = 0$ et $\beta_1 = 1$, l'évaluation de l'expert est dite "rationnelle".

1. L'équation estimée est la suivante:

$$\widehat{\text{price}} = -14,47 + 0,976 \text{ assess} \\ (16,27) \quad (0,049)$$

$$N = 88, \quad \text{SCR} = 165.644,51 \quad R^2 = 0,820.$$

Testez l'hypothèse $H_0: \beta_0 = 0$ par rapport à l'alternative bilatérale, et l'hypothèse $H_0: \beta_1 = 1$ par rapport à l'alternative bilatérale. Que concluez-vous?

2. Pour tester l'hypothèse jointe $H_0: \beta_0 = 0$ et $\beta_1 = 1$, le SCR du modèle restreint est simplement égal à:

$$\sum_{i=1}^N (\text{price}_i - \text{assess}_i)^2,$$

où $N = 88$. Pourquoi? Expliquez. Sachant que la SCR du modèle restreint est égale à 209.448,99, faites le test de Fischer de $H_0: \beta_0 = 0$ et $\beta_1 = 1$.

3. Testez $H_0: \beta_2 = 0, \beta_3 = 0, \beta_4 = 0$ dans le modèle suivant:

$$\text{price} = \beta_0 + \beta_1 \text{assess} + \beta_2 \text{lotsize} + \beta_3 \text{sqrft} + \beta_4 \text{bdrms} + u,$$

où *lotsize* représente la surface du terrain en pieds carrés, *sqrft* la surface habitable de la maison en pieds carrés, *bdrms* le nombre de chambres. Pour cela sachez que le R-carré de cette régression est 0,829.

Exercice 3. Justifiez les propositions suivantes.

1. Soit un test de Fischer (ou de Student) avec un seuil de signification α . Si le seuil de signification α est relevé par l'économètre, la probabilité de rejeter l'hypothèse nulle, qu'elle soit vraie ou fausse, augmente.
2. Soit le modèle suivant:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + u.$$

et l'intervalle de confiance $[\underline{\beta}_1, \overline{\beta}_1]$ pour le paramètre β_1 au seuil de confiance de 95%. Soit le test de Student de l'hypothèse

$$H_0: \beta_1 = b_1$$

où b_1 est un scalaire quelconque. Si b_1 appartient à l'intervalle $[\underline{\beta}_1, \overline{\beta}_1]$, alors l'hypothèse H_0 n'est pas rejetée au seuil de 5% par un test de bilatéral Student.

3. Soit le modèle suivant

$$y = \beta_0 + \beta_1 x_1 + u.$$

Si l'hypothèse nulle $H_0 : \beta_1 = 0$ est rejetée par un test de Student unilatéral au seuil de 5% en faveur de l'hypothèse alternative $H_1 : \beta_1 > 0$, alors elle sera nécessairement rejetée par un test de Fischer au seuil de 10%.

Exercice 4. Considérons l'équation estimée présentée en cours, qui peut être utilisée pour étudier les effets de l'absence des cours au sur les résultats à l'université:

$$\begin{aligned} \widehat{\text{colGPA}} = & 1,39 + 0,412 \text{ hsGPA} + 0,015 \text{ ACT} \\ & (0,33) \quad (0,094) \quad (0,011) \\ & - 0,083 \text{ skipped} \\ & (0,026) \end{aligned}$$

$$N = 141, \quad R^2 = 0.234.$$

où colGPA est la note obtenue en première année de collège (université de 1er cycle), hsGPA la note obtenue au lycée, ACT la note obtenu au American College Test, et skipped le nombre de cours auxquels l'étudiant a été absent.

1. Calculez l'intervalle de confiance au seuil de 95% pour le paramètre β_3 associé à skipped.
2. Soit β_1 le paramètre associé à hsGPA. Testez l'hypothèse $H_0 : \beta_1 = 0,5$ contre l'alternative bilatérale au seuil de 5%? Sans utiliser de table ou d'ordinateur, quelle est, selon vous, et approximativement, la valeur de la probabilité critique? Donnez un ordre de grandeur et justifiez.
3. Testez l'hypothèse $H_0 : \beta_1 = 1,0$ contre l'alternative bilatérale au seuil de 1%? Sans utiliser de table ou d'ordinateur, quelle est, selon vous, et approximativement, la valeur de la probabilité critique? Donnez un ordre de grandeur et justifiez.

Exercice 5. Soit les résultats d'estimation obtenus à l'aide des données wage1.xls:

$$\begin{aligned}\widehat{\log(wage)} = & 0.392 + 0.082 \text{ educ} + 0.036 \text{ exper} - 0.001 \text{ exper}^2 \\ & (0.102) \quad (0.007) \quad (0.007) \quad (0.0001) \\ & - 0.329 \text{ female} + 0.064 \text{ married} \\ & (0.037) \quad (0.042)\end{aligned}$$

$$R^2 = 0.402, \quad N = 526, \quad SSR_U = 88.657$$

1. Calculez l'intervalle de confiance pour β_4 (correspondant à "female") au seuil de confiance de 95%.
2. En utilisant cet intervalle de confiance, vous semble-t-il possible que, une fois que l'on contrôle pour l'éducation, l'expérience et le statut marital, les femmes ont le même salaire que les hommes?
3. Considérez l'hypothèse suivante: "si on contrôle pour le niveau d'éducation, le sexe et le statut marital, l'expérience n'a pas d'effet sur le logarithme du salaire". Formulez précisément l'hypothèse nulle correspondante à tester.
4. Considérez les résultats d'estimation suivants:

$$\begin{aligned}\widehat{\log(wage)} = & 0.708 + 0.075 \text{ educ} - 0.327 \text{ female} + 0.209 \text{ married} \\ & (0.094) \quad (0.007) \quad (0.039) \quad (0.039)\end{aligned}$$

$$R^2 = 0.336, \quad N = 526, \quad SSR_R = 98.472$$

A l'aide d'un test de Fisher, testez l'hypothèse nulle présentée en (c) au seuil de 1%.