

Econométrie et Applications
2ème session

2013-14

La durée de l'examen est fixée à 2h00. Vous devez répondre aux questions dans l'ordre de l'énoncé. Chaque question nécessite une réponse concise et précise. Une courte explication est généralement nécessaire. Si possible, la notation adoptée au cours doit être utilisée. L'examen est évalué sur 20 points.

1. [5 points] Vrai ou faux. Une justification est toujours nécessaire.
 - (a) Le \bar{R}^2 (ou R^2 ajusté) est nécessairement compris entre 0 et 1.
 - (b) Sous les hypothèses de Gauss-Markov, l'estimateur des MCO est celui qui a la plus petite variance dans la classe des estimateurs non biaisés.
 - (c) Dans le modèle de régression linéaire simple, la somme du produit de la variable expliquée et du résidu est toujours égale à zéro: $\sum_{i=1}^N y_i \hat{u}_i = 0$.
 - (d) Le test de Brausch-Pagan est un test d'hétéroscédasticité.
 - (e) Dans le modèle de régression simple, la convergence des estimateurs des MCO peut être démontrée en utilisant la condition $\text{cov}(x, y) = 0$ à la place de $E(u|x) = 0$.

2. [4.5 points] Un chercheur utilise des données sur 27 individus pour étudier la relation entre une variable expliquée y_i et une variable explicative

x_i . Une analyse préliminaire donne les informations suivantes :

$$\begin{aligned} \sum_{i=1}^N y_i &= 405; & \sum_{i=1}^N x_i &= 270; \\ \sum_{i=1}^N (x_i - \bar{x})^2 &= 240; & \sum_{i=1}^N (y_i - \bar{y})^2 &= 300; \\ \sum_{i=1}^N (y_i - \bar{y})(x_i - \bar{x}) &= 120; & \sum_{i=1}^N \hat{u}_i^2 &= 54. \end{aligned}$$

avec $N = 27$. Utilisez ces informations afin de répondre à toutes les questions ci-dessous. Montrez tous vos calculs.

1. Calculez les estimateurs des MCO de la constante β_0 et de la pente β_1 .
2. Calculez la valeur du R^2 , le coefficient de détermination pour les estimateurs des MCO, et interprétez le.
3. Calculez l'estimateur de la variance pour l'estimateur du paramètre de la pente β_1 .

3. [6 points] Considérez la relation entre le poids à la naissance d'un enfant en kilogrammes (pdnais) et différentes variables explicatives :

$$\begin{aligned} \text{pdnais} &= 3,154 - 0,0091 \cdot \text{cigs} + 0,0091 \cdot \text{rang} \\ &\quad \begin{matrix} (0,267) & (0,0027) & (0,0057) \end{matrix} \\ &\quad + 0,031 \cdot \text{edufem} + 0,067 \cdot \text{eduhom} \\ &\quad \begin{matrix} (0,024) & (0,055) \end{matrix} \\ N &= 31, R^2 = 0,316 \end{aligned}$$

où cigs désigne le nombre moyen de cigarettes fumées quotidiennement par la mère, rang l'ordre de l'enfant à la naissance (premier, deuxième,...), edufem l'éducation de la mère, eduhom celle du père.

1. Comment se modifieraient les estimations des paramètres si le poids de l'enfant était mesuré en grammes plutôt qu'en kilogrammes? Comment se modifierait le R^2 ?
2. Quel est le poids moyen d'un premier enfant dont les parents ont tous les deux douze années d'éducation, et dont la mère fume 25 cigarettes par jour?

3. Pensez-vous que l'estimateur β_{cigs} mesure l'effet "ceteris paribus" du nombre de cigarettes sur le poids de l'enfant? En d'autres termes, si les pouvoirs publics décidaient d'interdire la consommation de cigarettes, le poids des enfants à la naissance augmenterait-il nécessairement en conséquence? Argumentez votre réponse.
4. Calculez l'intervalle de confiance à 95 % du paramètre β_{cigs} . Précisez, à l'aide de cet intervalle de confiance, si l'effet des cigarettes est significativement différent de zéro. Expliquez la démarche.
5. Testez l'hypothèse que $\beta_{\text{rang}} = 0$ (au seuil de 1%) sous l'hypothèse alternative que $\beta_{\text{rang}} \neq 0$. Expliquez la démarche.
6. Supposez que l'on teste l'hypothèse $\beta_{\text{edufem}} = \beta_{\text{eduhom}} = 0$ à l'aide du test de Fisher. Pour cela, le modèle contraint est estimé, le R^2 de ce modèle et la statistique de Fisher sont calculés. Cette dernière est égale à 7,35. Quel est le R^2 du modèle contraint?

4. [1.5 points] Considérez la droite de régression suivante:

$$\widehat{\log \text{wage}} = 1,53 + 0,07 \cdot \text{edu} - 0,35 \cdot \text{fem}$$

où edu est le nombre d'années d'éducation, wage est le salaire et fem est une variable dichotomique qui prend la valeur 1 pour une femme et 0 pour un homme. Un économètre veut tester l'hypothèse que l'écart entre le salaire des femmes et celui des hommes dépend du niveau d'éducation. Dans ce cas, comment doit-il réécrire le modèle afin de réaliser ce test?

5. [3 points] Considérez le modèle de régression simple (avec une seule variable explicative) et supposez que le terme d'erreur est hétéroscédastique.

1. Qu'est-ce que cela signifie? Quelles sont les implications de l'hétéroscédasticité sur le biais des estimateurs des MCO et la validité de l'inférence (la validité des procédures de tests)?
2. Supposez que la variance du terme d'erreur puisse s'écrire:

$$\text{var}(u|x) = \sigma^2 \times h(x)$$

où $h(x)$ est une fonction positive, et considérez le cas où la fonction $h(x)$ est inconnue. Décrivez les différentes étapes permettant d'estimer ce modèle par les Moindres Carrés Pondérés (MCP) ou Moindres Carrés Généralisés "Faisables". Quels sont les avantages des MCP sur les MCO? Quels sont leurs inconvénients?