UGA: L3 Miash

Économétrie 1

Examen de contrôle continu

Michal Urdanivia, Université de Grenoble Alpes, Faculté d'Économie, GAEL

Courriel: michal.wong-urdanivia@univ-grenoble-alpes.fr

15 décembre 2019

Remarques

(1) Durée : 90 minutes.
(2) Les calculatrices ne sont pas autorisées(seulement de quoi écrire à la main-stylo,...- si vous avez besoin de faire des calculs sur les feuilles distribuées).
(3) Pour vos réponses choisisser parmi les propositions de réponses qui vous sont proposées. Pour cela barrez celle que vous jugez fausses.
(4) Pour chaque question et donc ensemble de choix de réponses proposées, il peut y avoir plusieurs propositions correctes.
(5) Pour avoir l'ensemble des points à une questions tous les choix doivent être justes.
(6) Si vous ne répondez pas à une question vous serez noté pour celle-ci 0. Si vous répondez "faux" vous serez noté "-0.5". Une réponsé pouvant être jugée "fausse" dès lors que toutes les propositions de réponses correctes n'ont pas été retenues.

1. Modèle de régression linéaire et estimateur des moindres carrés

On considère un modèle de régression ayant pour équation,

$$Y_i = \beta_1 + X_{2i}\beta_2 + X_{3i}\beta_3 + U_i. \tag{1}$$

où Y_i est la variable dépendante, X_{2i}, X_{3i} deux régresseurs(en plus du régresseur constant $X_{1i} \equiv 1$) définis sur $\mathbb{R}$, et U_i l'erreur du modèle. L'indice "i" est celui d'une observation dans un échantillon i.i.d., $\{(Y_i, X_i^\top)\}_{i=1}^n$, avec $X_i := [1, X_{2i}, X_{3i}]^\top$.

(1) On cherche à calculer l'estimateur des MCO de $\beta := [\beta_1, \beta_2, \beta_3]$. On le note $\hat\beta := [\hat\beta_1, \hat\beta_2, \hat\beta_3]$. Nous savons qu'il est obtenu à partir de :

$$\hat\beta = \left[n^{-1}\sum_{i=1}^n X_i X_i^\top\right]^{-1}\left[n^{-1}\sum_{i=1}^n X_i Y_i\right]$$

Après quelques calculs nous obtenons que :

$$n^{-1}\sum_{i=1}^n X_{2i} = 0,\ n^{-1}\sum_{i=1}^n X_{3i} = 0,\ n^{-1}\sum_{i=1}^n X_{2i}^2 = 4,\ n^{-1}\sum_{i=1}^n X_{3i}^2 = 2,$$

$$n^{-1}\sum_{i=1}^n X_{2i}X_{3i} = 0,\ n^{-1}\sum_{i=1}^n Y_i = 2,\ n^{-1}\sum_{i=1}^n X_{2i}Y_i = 4,\ n^{-1}\sum_{i=1}^n X_{3i}Y_i = 6.$$

Parmi les deux résulats suivants lequel est correct ?

(a) $\hat{\beta} := [2, 1, 3]^\top$.

(b) $\hat{\beta} := [1, 3, 0.5]^\top$.

(2) Pour que cet estimateur soit sans biais pour β nous devons supposer que :

(a) $\mathbf{E}[U_i^2|X_i] = \sigma^2 > 0$, où σ est un scalaire.

(b) $\mathbf{E}[X_i U_i] = 0$.

(c) $\mathbf{E}[U_i|X_i] = 0$.

(d) $\mathbf{E}[X_i X_i^\top]^{-1}$ existe.

(3) Pour que cet estimateur soit convergent pour β nous devons supposer que :

(a) $\mathbf{E}[U_i^2|X_i] = \sigma^2 > 0$, où σ est un scalaire.

(b) $\mathbf{E}[X_i U_i] = 0$.

(c) $\mathbf{E}[U_i|X_i] = 0$.

(d) $\mathbf{E}[X_i X_i^\top]^{-1}$ existe.

(4) On considère la spécification suivante pour l'équation du modèle :

$$Y_i = \beta_1 + \sum_{j=1}^{J} X_{2i}^j \beta_{2j} + \sum_{j=1}^{J} X_{3i}^j \beta_3 j + U_i.$$

c'est donc un modèle avec des polynomes sur X_{2i} et X_{3i}((par exemple pour $J = 1$ on retrouve le modèle des questions précédentes). Les paramètres à estimer sont donc $\beta := [\beta_1, \beta_{21}, \ldots, \beta_{2J}, \beta_{31}, \ldots, \beta_{3J}]^\top$. Est-ce que cette équation est celle d'un modèle modèle linéaire ?

(a) Oui.

(b) Non.

(5) On considère à présent la spécification suivante :

$$Y_i = \beta_1 + X_{2i}\beta_2 + X_{3i}\beta_3 + \beta_4 X_{4i} + U_i$$

où $X_{4i} = X_{2i} + X_{3i}$. Peut-on calculer l'estimateur des MCO(d'après la formule classique rappelée à la question 1) dans ce cas ?

(a) Oui

(b) Non

(6) Enfin nous voulons calculer l'estimateur des MCO pour un modèle ayant pour équation :

$$Y_i = \beta_1 + X_{2i}\beta_2 + U_i$$

Lequel des deux résultat suivants est correct ?

(a) $\hat{\beta} := [3, 1]$.

(b) $\hat{\beta} := [2, 1]$.

(7) On cherche à estimer l'ecart-type de $\hat{\beta}2$ pour le modèle de la question 1, et cela sous l'hypothèse d'homoscédasticité des erreurs. Cela signifie qu'on suppose que :

(a) $\mathbf{E}[U_i^2|X_i] = \sigma_i^2 > 0$.

(b) $\mathbf{E}[U_i^2|X_i] = \sigma^2 > 0$.

où σ et σ_i sont des scalaires.

(8) Après quelques calcul nous obtenons que $n^{-1}\sum_{i=1}^{n}\widehat{U}_i^2 = 2$, avec $\widehat{U}_i := Y_i - X_i^\top\hat{\beta}$. Par consquent l'écart-type estimé de $\hat{\beta}_2$ est égal à :

(a) $\sqrt{\widehat{\mathbf{V}}(\hat{\beta})_{(2,2)}} = \sqrt{0.5}$.

(b) $\sqrt{\widehat{\mathbf{V}}(\hat{\beta})_{(2,2)}} = \sqrt{2}$.

$\widehat{\mathbf{V}}(\hat{\beta})_{(2,2)}$ est l'élément $(2,2)$ de la matrice des variances-covariances estimées de $\hat{\beta}$ sous l'hypothèse d'homoscédasticité.

(9) On soupçonne X_{3i} d'être endogène. Cela signifie donc que :

(a) $\mathbf{E}[U_i] \neq 0$.

(b) $\mathbf{E}[U_i X_i] \neq 0$.

(10) Dans ce cas on ne peut pas utiliser l'estimateur des MCO car :

(a) On ne peut plus le calculer.

(b) Il n'est pas convergent.

(11) Nous avons une variable Z_i qui peut être employé comme instrument de X_{3i}. Autrement dit cette variable est supposée vérifier :

(a) $\mathbf{Cov}[Z_i, X_{3i}] = 0$.

(b) $\mathbf{E}[Z_i U_i] = 0$.

(c) $\mathbf{Cov}[Z_i, X_{3i}] \neq 0$.

(d) $\mathbf{E}[Z_i U_i] = 0$.

2. Card (1993)

(1) On rappelle que la problèmatique de l'article est de :

(a) Mesurer empiriquement les déterminants du niveau d'études.

(b) Mesurer empiriquement l'effet causal en termes de salaire sur le marché du travail de la proximité du lieu de résidence avec les universités(ou autres institutions d'enseignement supérieur).

(c) Mesurer empiriquement l'effet causal en termes de salaire sur le marché du travai du plus haut niveau d'études atteint avant l'entrée sur le marché du travail.

(2) Les données employées concernent notamment :

(a) Des hommes et des femmes de toute classe d'âge aux USA.

(b) Des femmes de moins de 30 ans aux USA.

(c) Des hommes de moins de 30 ans aux USA.

(3) Le niveau d'études est mesuré en termes de :

(a) Diplôme le plus élevé ayant été obtenu.

(b) Nombre d'années d'étude.

(4) La variable dépendante dans les régressions estimées est une mesure :

(a) Du revenu du ménage.

(b) Du salaire.

(5) La variable dont l'auteur traite l'endogénéité potentielle est

(a) La proximité du lieu de résidence.

(b) Le niveau d'études.

(6) Le principal instrument pour traiter cette endogénéité est :

(a) La proximité du lieu de résidence.

(b) Le niveau d'études.

Références

Card, David. 1993. "Using geographic variation in college proximity to estimate the return to schooling." Tech. rep., National Bureau of Economic Research. URL http://www.nber.org/papers/w4483.