Université Paris 7 **Économétrie**
Licence 3, 2009-2010 **Pierre Alquier - Joseph Salmon**

TD N° 9 : Introduction à l'utilisation des variables instrumentales

---

Dans ce TD, on s'intéresse au cas où l'hypothèse d'indépendance entre les variables explicatives et le bruit est violée. Une méthode très connue permet de résoudre ce problème, la méthode des *variables instrumentales*, mais elle n'est pas facile. On l'étudie dans l'Exercice 1 dans un cas théorique simple, et dans l'Exercice 2 sur un exemple.

L'Exercice 3 est complètement différent et tente de généraliser les modèles à retards finis pour obtenir un modèle à retards infinis.

Exercice 1. On suppose que l'on observe deux variables économiques x et y reliées par l'équation

$$y_i = bx_i + \varepsilon_i \tag{1}$$

pour $i = 1, ..., n$ où les couples (x_i, ε_i) sont i.i.d. de loi

$$\begin{pmatrix} x_i \\ \varepsilon_i \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} m_x \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_x^2 & \gamma \\ \gamma & \sigma^2 \end{pmatrix} \right)$$

avec $\gamma \neq 0$. On rappelle que dans le DM 1, on a obtenu l'expression suivante pour l'estimateur des moindres carrés de b :

$$\hat{b} = \frac{\frac{1}{n}\sum_{i=1}^n y_i x_i}{\frac{1}{n}\sum_{i=1}^n x_i^2}.$$

1) Quelle hypothèse usuelle est ici violée ?
2) Montrer que
$$\hat{b} = b + \frac{\frac{1}{n}\sum_{i=1}^n \varepsilon_i x_i}{\frac{1}{n}\sum_{i=1}^n x_i^2}.$$
3) L'estimateur $\hat{b}$ est-il sans biais ?
4) Déterminer $\lim_{n\to\infty} \hat{b}$. L'estimateur $\hat{b}$ est-il consistant ?
5) Un économètre astucieux observe qu'une autre variable économique z vérifie pour tout i,
$$x_i = az_i + \eta_i \tag{2}$$
où les couples (z_i, η_i) sont i.i.d. de loi
$$\begin{pmatrix} z_i \\ \eta_i \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} m_z \\ 0 \end{pmatrix}, \begin{pmatrix} s_z^2 & 0 \\ 0 & s^2 \end{pmatrix} \right).$$
De plus, tous les z_i sont indépendants de tous les ε_i. Question technique (sans conséquence pour la suite) : déterminer la loi des couples (ε_i, η_i).
6) On estime le paramètre a par moindres carrés, rappeler l'expression de $\hat{a}$.
7) On pose comme d'habitude $\hat{x}_i = \hat{a}z_i$ et on propose alors l'estimateur suivant de b :
$$\hat{b}^{\mathrm{VI}} = \frac{\frac{1}{n}\sum_{i=1}^n y_i \hat{x}_i}{\frac{1}{n}\sum_{i=1}^n \hat{x}_i^2}.$$
Montrer que
$$\hat{b}^{\mathrm{VI}} = b + \frac{\frac{1}{n}\sum_{i=1}^n \varepsilon_i z_i}{\frac{1}{n}\sum_{i=1}^n x_i z_i} = b + \frac{\frac{1}{n}\sum_{i=1}^n \varepsilon_i z_i}{\frac{1}{n}\sum_{i=1}^n \hat{x}_i z_i}.$$

TAB. 1 – Données pour l'Exercice 2.

i	prev_i	crois_i	z_i
1	-0.22	0.23	-1.41
2	-0.89	0.45	-0.59
3	0.09	-1.74	-1.32
4	-0.17	-0.48	1.40
5	-0.81	-0.10	-0.06
6	-0.35	-2.24	-0.56
7	2.12	2.16	1.41
8	2.92	2.14	0.84
9	-0.57	0.04	-0.61
10	-0.35	-1.04	-0.79

8) L'estimateur $\hat{b}^{\text{VI}}$ est-il sans biais ? Est-il consistant ?

Remarque : $\hat{b}^{\text{VI}}$ s'appelle estimateur à variable instrumentale, la variable z s'appelle la variable instrumentale. Il est fréquent d'utiliser plusieurs variables instrumentales, mais c'est plus difficile.

EXERCICE 2. On s'intéresse ici à un problème d'économie (avec de fausses données). Des données sont fournies dans la Table 1. Ici crois_i est la croissance d'un pays i en 2009, et prev_i est la prévision de croissance qu'un célèbre magazine économique avait attribué à ce pays pour 2009.

Un économètre part du principe que les prédictions trouvées dans la presse ont une influence (idée de la "prophétie auto-réalisatrice" en économie) et postule le modèle

$$\text{crois}_i = b\,\text{prev}_i + \varepsilon_i. \tag{3}$$

Cependant, un autre économètre se rend compte que les notes attribuées par le magazine sont presque exactement les croissance de l'année précédente (2008), que l'on notera z_i.

1) Expliquer pourquoi, dans le modèle donné par l'Equation (3), on attend $\text{Cov}(\text{prev}_i, \varepsilon_i) \neq 0$. Quel signe attend-on pour cette quantité ?
2) D'après l'exercice précédent, quelle en sera la conséquence sur l'estimation de b ?
3) On obtient les estimateurs suivants, en utilisant la variable instrumentale z_i :

$$\hat{b} = 0.76 \text{ et } \hat{b}^{\text{VI}} = 1.13.$$

Ceci est-il conforme à vos attentes ?

EXERCICE 3. On souhaite étudier un modèle à retards infinis (pour simplifier, on ne met pas de terme d'erreur, on suppose donc pour le moment que le modèle est exact).

$$y_t = b_0 + \sum_{j=0}^{+\infty} b_j z_{t-j}.$$

1) Quel(s) problème(s) se pose(nt) si l'on souhaite déterminer les coefficients b_j ?
2) Souvent, il est justifié de postuler une relation de la forme $b_j = b^j$ pour un certain $b \in]-1, 1[$. Montrer qu'on a alors :

$$y_t = a_0 + a_1 z_t + a_2 y_{t-1},$$

où a_0, a_1 et a_2 seront à exprimer en fonction de b_0 et b. Ceci motive l'introduction de modèles "auto-regressifs", c'est-à-dire où l'on autorise y_{t-1} comme variable explicative (on étudiera ces modèles en cours).

3) Quel est dans ce cas l'effet sur la variable y d'un choc sur la variable z ? On pourra supposer que pour tout t, $z_t = c$ sauf pour la date t_0 où $z_{t_0} = c + \delta$ avec $\delta > 0$, et calculer l'évolution de y_t.