Econométrie

puf

René Giraud
Nicole Chaix

ÉCONOMIE

Econométrie

« ÉCONOMIE »
COLLECTION DIRIGÉE
PAR CLAUDE JESSUA, CHRISTIAN LABROUSSE
ET DANIEL VITRY

Econométrie

RENÉ GIRAUD
Professeur à l'Université de Poitiers chargé d'un enseignement d'Econométrie en Doctorat à Paris 2 - Panthéon et à l'Institut Français du Pétrole

NICOLE CHAIX
Maître de conférences à l'Université Paris 2

PRESSES UNIVERSITAIRES DE FRANCE

ISBN 2 13 046017 8

Dépôt légal — 1re édition : 1989, janvier
2e édition mise à jour : 1994, mars

108, boulevard Saint-Germain, 75006 Paris

Sommaire

Présentation de l'ouvrage

Le cours d'économétrie* qui est présenté aux étudiants de Sciences Economiques a un double objectif :

1 / permettre de bien comprendre les méthodes statistiques utilisées pour l'estimation des modèles économiques;

2 / appliquer ces méthodes à l'aide d'exercices dont les énoncés et les solutions sont donnés à la fin de chaque chapitre.

En effet la publication de deux ouvrages séparés, cours et exercices, pose souvent des problèmes aux utilisateurs. La rédaction des exercices venant juste après l'exposé du cours doit simplifier le travail des étudiants; ceux-ci peuvent contrôler aussitôt les méthodes qu'ils ont appliquées à la résolution d'un modèle.

Nous avons abordé les problèmes classiques mais les démonstrations délicates sont reportées en note pour ne pas alourdir l'exposé.

Nous nous sommes proposés de donner au lecteur une bonne formation de base et de développer chez lui l'*aptitude au raisonnement économétrique*.

Une bibliographie suffisamment étendue permet à ceux qui veulent approfondir une méthode de se reporter aux ouvrages spécialisés correspondants.

* Le cours est de René Giraud, professeur à l'Université de Poitiers. Les exercices et les corrigés sont de Nicole Chaix, maître de conférences à l'Université de Paris 2.

Les premiers chapitres traitent de la notion de *modèle aléatoire* et du *modèle linéaire de la régression simple.*

Ce dernier est étudié très complètement. Les calculs concernant l'estimation, la qualité des estimateurs obtenus par la *méthode des Moindres Carrés Ordinaires*, leur distribution de probabilité sont mis en évidence.

Les tests et régions de confiance qui en résultent pour les coefficients du modèle sont explicités. Par ailleurs l'erreur de prévision, son espérance et sa variance ainsi que la fourchette prévisionnelle pour la variable endogène du modèle sont étudiées avec soin.

Une note en fin de chapitre rappelle les définitions des lois de probabilité utilisées : loi du χ^2, loi de Student-Fisher, loi de Fisher-Snedecor.

Ces premiers chapitres vus de façon très détaillée permettent d'aborder l'étude du *modèle linéaire général* à *erreurs indépendantes* d'abord, à *erreurs corrélées* ensuite, avec les méthodes pratiques d'estimation qui en découlent.

Les expressions des estimateurs sont maintenant données sous forme matricielle. Les méthodes élémentaires d'algèbre linéaire utilisées sont précisées au fur et à mesure des difficultés rencontrées.

Les processus *autorégressifs* et à *retards échelonnés* sont ensuite étudiés. Les démonstrations délicates concernant les autorégressifs ont été évitées. Quelques méthodes d'estimation concernant les deux types de modèles sont proposées.

Un chapitre très simple est consacré aux *modèles non linéaires.*

Une dernière partie enfin concerne l'étude des *modèles à équations simultanées.*

Les définitions et théorèmes ayant trait à l'*identification* ont été exposés sans entrer dans les démonstrations souvent très longues et peu utilisables pour les économistes.

Les méthodes de *régression indirecte*, des *Doubles Moindres Carrés* et des *Triples Moindres Carrés* sont exposées. La méthode du *Maximum de Vraisemblance à information limitée* (Cowles Commission) est présentée sans entrer dans le détail des calculs. Quelques précisions sur les *modèles récursifs* achèvent le cours.

Par contre l'analyse des séries temporelles, modèles ARMA et ARIMA ainsi que l'introduction à l'analyse spectrale n'ont pas été abordées dans ce livre. Il nous a semblé que le traitement de ces modèles pouvait faire l'objet à lui seul d'un travail spécifique.

*
* *

Nous tenons à remercier Jean-Pierre Berdot, maître de conférences à Poitiers, pour les suggestions qu'il nous a faites à la lecture de cet ouvrage et les conseils avisés qu'il nous a donnés, ainsi que M. Régis Bourbonnais, assistant d'Economie à Paris 2, pour l'intérêt qu'il a bien voulu porter à la lecture du manuscrit. Nous remercions enfin MM. les professeurs Jessua, Labrousse et Vitry de nous avoir confié l'élaboration et la rédaction de ce travail.

R. Giraud,
Professeur
à l'Université de Poitiers

N. Chaix,
Maître de conférences
à l'Université de Paris 2

Préface de la deuxième édition

L'introduction dans cette nouvelle édition du livre *Econométrie* de deux nouveaux chapitres appelle quelques commentaires.

Le manuel ne traitant dans sa première édition que de processus autorégressifs (AR) devait, pour être plus complet, s'étendre à des processus autorégressifs et de moyennes mobiles (ARMA) et à des modèles comportant une variable non stationnaire, mais qui, différenciée plusieurs fois, pouvait atteindre la stationnarité. Ce sont les modèles ARIMA ou « Autoregressive Integrated Moving Average Models ».

Ces notions étant bien établies, le rôle de l'opérateur de décalage étant bien précisé, il est alors relativement simple d'aborder la notion de variable intégrée d'ordre *d*, puisqu'une telle variable, stationnaire après *d* différenciations, possède une représentation ARMA inversible. Cette propriété d'inversibilité, liée à celle d'opérateur convergent, est précisée lors de l'étude des processus ARMA et ARIMA. On peut ensuite aborder la notion de co-intégration, combinaison de variables non stationnaires dans un modèle.

On imagine tout l'intérêt de ces notions de variables intégrées et co-intégrées puisque, dans un modèle économétrique, il est rare, pour représenter l'évolution d'un phénomène, de n'avoir affaire qu'à des variables stationnaires.

Le mélange de variables stationnaires et non stationnaires est abordé généralement avec les modèles à correction d'erreurs de Hendry (modèles ECM), mais il nous a semblé intéressant d'introduire cette notion à l'aide de séries temporelles, d'autant plus que les tests présentés à la suite reposent sur la notion de variable purement aléatoire — appelée encore

« bruit blanc » — présentée, dès la première édition, lors de l'étude des modèles de régression classiques ainsi qu'au cours de celle des autorégressifs.

Exposés simplement, ces deux chapitres complémentaires doivent éclairer d'un jour nouveau la présentation générale de l'ouvrage.

René GIRAUD. Nicole CHAIX.

Chapitre 1

Notion de modèles économétriques Généralités

1 / Introduction

Pour étudier un phénomène économique on essaie de représenter celui-ci par le comportement d'une variable.

Cette variable économique dépend elle-même d'autres variables que l'on relie entre elles par une relation mathématique.

— Par exemple, si l'on étudie l'Offre O et la Demande D d'un certain bien sur un marché, on sait que D et O dépendent du prix p du bien. On peut écrire que O est une certaine fonction du prix, D une autre fonction de ce même prix et que l'équilibre sur le marché se traduit par $O = D$.

On vient donc de construire un modèle élémentaire sous la forme :

$$(1) \quad \begin{cases} O = f(p) \\ D = g(p) \\ O = D. \end{cases}$$

Cependant, Offre et Demande dépendent d'autres variables que le prix.

— La demande de certains produits alimentaires par exemple dépend du revenu des ménages, du prix de produits analogues, etc.

S'il s'agit de denrées agricoles l'offre dépend des prix de l'année précédente.

Il n'est donc pas inutile de préciser la relation donnée à un instant t et d'indexer les variables utilisées de la façon suivante :

$$(2) \quad \begin{cases} O_t = f(p_{t-1}, x_{1t}, x_{2t}, \dots, x_{nt}) \\ D_t = g(p_t, x_{1t}, \dots, x_{pt}) \\ O_t = D_t. \end{cases}$$

Dans le deuxième modèle ainsi obtenu, on a introduit différentes variables explicatives $X_1, X_2, \ldots, X_n$ et on a considéré les réalisations de ces variables aux instants t ou $t - 1$.

On remarquera ici que le modèle proposé comporte plusieurs équations : nous avons un modèle à équations simultanées.

Mais pour commencer il est plus simple de raisonner sur un modèle comportant une seule équation.

2 / Notion de modèle aléatoire

Proposons-nous d'étudier la consommation C_i d'un certain bien par un ménage i.

Cette consommation dépend entre autres du revenu r_i du ménage.

Le modèle le plus élémentaire consiste à expliciter C_i en fonction de r_i.

D'autres facteurs, dont certains sont partiellement inconnus, déterminent également C_i.

On peut condenser les effets de ces autres facteurs en un facteur aléatoire ε_i.

On obtiendra alors le modèle *aléatoire* :

(1) $$C_i = f(r_i) + \varepsilon_i.$$

Les ε_i obéissent à une loi de probabilité P qu'il faudra préciser au cours des hypothèses faites sur le modèle. Très souvent, ces hypothèses qui portent sur les premiers moments de ε_i suffiront.

En présence du modèle (1) il faudra maintenant s'assurer que la classe de la fonction choisie pour f n'est pas en contradiction avec les résultats de l'expérience.

Par exemple, si on décide de choisir pour f une fonction du premier degré :

(2) $$C_i = ar_i + b + \varepsilon_i$$

on s'assurera en faisant varier i selon les différents ménages considérés que la relation (2) est bien satisfaisante.

On dit que l'on « *teste* » le modèle.

Si le résultat obtenu est convenable on « *estimera* » alors les paramètres a et b. Enfin on définira une *règle de prévision* permettant, connaissant r_i, de déterminer C_i.

3 / Nature des variables figurant dans le modèle. Spécification et structure du modèle

a / On distingue deux types de variables dans un modèle économétrique.

1. Les variables *exogènes*.

Ce sont des variables explicatives de la variable étudiée. Elles sont considérées comme des données autonomes. Ainsi dans le modèle

(1) $$C_i = ar_i + b + \varepsilon_i$$

la grandeur r_i est la variable explicative ou variable exogène. C'est sa valeur pour i donné qui permet de déterminer C_i, à l'aléatoire près ε_i. Parfois même elles sont considérées comme prédéterminées. Rappelons que dans une fonction keynésienne r_i serait le revenu de l'individu i sur la période qui précède la période de référence de la consommation.

2. Les variables *endogènes* ou variables à expliquer.

C_i est la variable endogène du modèle précédent.

C_i est maintenant devenue une variable aléatoire par l'intermédiaire de ε_i.

Elle ne sera pas traitée comme une variable exogène; cette distinction entre nature des variables est très importante et devra toujours être précisée avant l'étude d'un modèle.

b / Qu'entend-on maintenant par spécification du modèle ?

On dit que l'on *spécifie* un modèle quand on donne à celui-ci sa formulation mathématique définitive. Le modèle (1) ci-dessus est spécifié.

On connaît la forme de la fonction f dans l'expression

(2) $$C_i = f(r_i) + \varepsilon_i$$

(3) $$f(r_i) = ar_i + b.$$

L'adjonction de la variable aléatoire ε_i donne au modèle sa formulation définitive. La *structure* du modèle, elle, est constituée par l'ensemble des paramètres qui définissent complètement le modèle.

Par exemple, supposons que :

$$a = 0{,}3$$

$$b = 15$$

et que les ε suivent une loi normale d'espérance mathématique nulle et de variance égale à 3; alors l'ensemble

$$\begin{cases} a = 0{,}3 \\ b = 15 \\ \sigma = \sqrt{3} \end{cases}$$

constitue la structure du modèle.

Le but de l'économètre sera alors, connaissant les couples (C_i, r_i) associés aux différents ménages i, de déterminer la structure vraie du modèle; c'est-à-dire, à partir d'un *espace échantillon* défini par l'ensemble des (C_i, r_i), de *déterminer* la *structure vraie* du modèle *dans l'espace* à 3 dimensions *des structures* (a, b, σ).

C'est ici qu'intervient alors l' « induction statistique ».

4 / L'induction statistique

L'objet de l'*induction statistique* est de déterminer une procédure qui, à partir des seules observations dont dispose l'économètre, permette de *passer* de *l'espace échantillon* à *l'espace des structures*.

Le modèle étant choisi une fois pour toutes on admet qu'il existe un triplet (a, b, σ) qui permet de représenter exactement le processus par lequel les valeurs des variables observées ont été déterminées.

Au cours de l'induction statistique, on ne modifiera plus le modèle, qui ne sera plus remis en question.

La procédure choisie, comme nous le verrons ultérieurement, consistera à obtenir des estimateurs des paramètres a et b permettant de déterminer au mieux les valeurs réelles de ces paramètres. On les appréciera en général à l'aide d'intervalles de confiance choisis à un niveau de signification donné α.

Par exemple dans le modèle :

$$C_i = ar_i + b + \varepsilon_i$$

on trouvera que a appartient à l'intervalle [0,27; 0,33] avec une probabilité de 95 %.

De même b appartiendra à l'intervalle [14; 16] avec la même probabilité.

Il sera possible également d'estimer l'écart-type σ de la variable aléatoire ε. On va voir maintenant le rôle important joué par cette variable aléatoire dans le modèle.

5 / L'identification du modèle. Structures non identifiables - structures équivalentes

Considérons toujours notre modèle :

(1) $C_i = ar_i + b + \varepsilon_i$

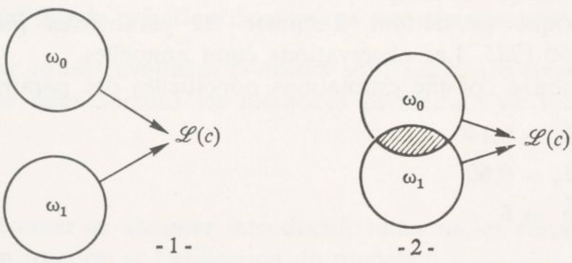
et imaginons que la procédure utilisée en partant des informations recueillies sur (C_i, r_i) nous conduise non pas à une solution unique mais à deux structures distinctes :

$$\omega_0 = \begin{Bmatrix} a_0 \\ b_0 \\ \sigma_0 \end{Bmatrix} \quad \text{et} \quad \omega_1 = \begin{Bmatrix} a_1 \\ b_1 \\ \sigma_1 \end{Bmatrix}$$

Il est évident que la loi de probabilité définie sur ε précise la loi de la variable endogène C.

Donc chaque structure, compte tenu des valeurs données aux exogènes et de la loi de ε, conduit à une loi de probabilité de C.

Imaginons que ω_0 et ω_1 conduisent à la même loi $\mathscr{L}$ de C :



-1- -2-

Deux cas de figure sont possibles :

— Ou ω_0 et ω_1 sont distinctes et on ne peut choisir ω_0 plutôt que ω_1. Dans ce cas les structures considérées ne sont pas identifiables et, par là même, le modèle n'est *pas identifiable*. On ne peut déterminer les valeurs des paramètres qui figurent dans le modèle.

— Ou ω_0 et ω_1 ne sont pas distinctes, leur intersection n'est pas vide. Les structures ω_0 et ω_1 permettront d'identifier une partie des paramètres du modèle, ceux précisément qui appartiennent à l'intersection de ω_0 et ω_1.

Ces deux structures seront dites *équivalentes*, mais ne permettront pas une identification complète du modèle.

Ce problème très important de l'identification se posera en particulier lors de l'étude de modèles à équations multiples; nous y reviendrons plus complètement lors de l'étude du chapitre concernant ces modèles.

6 / Prévision des variables endogènes à l'aide des valeurs fixées pour les exogènes

L'intérêt d'un modèle dont la structure est déterminée consiste à l'utiliser pour prévoir, à une époque future ou dans une circonstance fixée s'il s'agit d'observations prises au même instant, les valeurs des variables endogènes lorsque celles des exogènes sont fixées.

Soit le modèle suivant. On étudie l'évolution des importations Y en fonction de la production intérieure brute X_1 et de la formation des stocks X_2. On aura :

$$(1) \qquad y_t = a_1 x_{1t} + a_2 x_{2t} + b + \varepsilon_t \qquad t = 1, 2, \ldots, T.$$

t représente ici le temps.

L'historique permettant d'estimer les paramètres porte sur la période 1960-1985. Les observations étant annuelles.

On a trouvé comme estimations ponctuelles des paramètres :

$$\begin{cases} \hat{a}_1 = 0{,}140 \\ \hat{a}_2 = 0{,}60 \\ \hat{b} = 6. \end{cases}$$

Le *modèle estimé* s'écrit alors :

$$(2) \qquad \begin{array}{l} \hat{y}_t = (0{,}140)\, x_{1t} + (0{,}60)\, x_{2t} + 6 \\ \qquad\;\; (0{,}004) \qquad\quad (0{,}10) \qquad\; (1{,}1). \end{array}$$

Les chiffres entre parenthèses sous les valeurs estimées représentent ici les écarts-types des estimateurs $\hat{a}_1$, $\hat{a}_2$ et $\hat{b}$ afin de déterminer des intervalles de confiances pour a_1, a_2, b.

Supposons que l'on veuille faire une prévision sur l'année 1987. Il faudra alors fixer la production intérieure brute X_1, et la formation des stocks X_2 en 1987.

Si on impose $x_1 = 1\,030$ en milliards de francs au prix de 1960, et $x_2 = 12{,}7$ on aura comme prévision pour Y :

$$y^p_{1987} = (0{,}140)\,(1\,030) + (0{,}60)\,(12{,}7) + 6 \tag{3}$$

ou d'une façon générale :

$$y^p_\theta = \hat{a}_1\, x_{1\theta} + \hat{a}_2\, x_{2\theta} + \hat{b} \tag{4}$$

θ étant l'époque de la prévision.

Il est évident que la valeur prévue pour Y en 1987 appelle des remarques.

1. On a choisi arbitrairement les valeurs $x_{1\theta}$ et $x_{2\theta}$ compte tenu de leur évolution passée. Première source d'erreur.
2. La spécification du modèle n'est peut-être pas parfaite. La forme de la fonction choisie pour expliquer l'évolution de Y n'est pas suffisamment précise.
3. Il est possible enfin que les variables qui expliquent, d'après notre modèle, l'évolution de Y n'interviennent plus de la même façon que lorsque l'on a étudié notre phénomène sur la période 1960-1985. Autrement dit, il peut y avoir rupture d'équilibre entre les variables qui expliquent le phénomène au moment de la prévision. Les parts respectives qu'elles représentent dans la variation de Y ne sont plus les mêmes. Il est évident que la prévision dans ce cas sera sérieusement biaisée.

Les trois causes évoquées ci-dessus sont sources d'erreurs pour la prévision et nous verrons les méthodes permettant de minimiser ces erreurs.

Pour résumer ce chapitre introductif, retenons les étapes suivantes dans la construction et l'utilisation du modèle :

1 / spécification du modèle;
2 / estimation des paramètres et test du modèle à l'aide de statistiques déjà connues;
3 / prévision de la variable endogène.

Nous allons dans les chapitres qui vont suivre nous intéresser plus précisément aux méthodes d'estimation des paramètres et aux propriétés des estimateurs obtenus.

Chapitre 2

Modèle linéaire de la régression simple

Nous nous proposons d'étudier ici le modèle le plus simple. Une variable endogène représente l'évolution du phénomène étudié et cette évolution est expliquée par une seule variable exogène.

Nous examinerons de façon très complète les propriétés des estimateurs obtenus et il nous arrivera par la suite de généraliser pour des modèles plus complexes les résultats de cette analyse.

Cette étude relativement dense pour une première présentation sera décomposée en deux parties.

La première traitera de l'obtention des estimateurs des paramètres du modèle et de leurs propriétés.

La deuxième partie traitera de l'interprétation géométrique de la méthode utilisée, des tests et régions de confiance concernant les paramètres et enfin de la prévision qui peut être faite avec un tel modèle.

I / LE MODÈLE. LES ESTIMATEURS DES PARAMÈTRES OBTENUS PAR LA MÉTHODE DES MOINDRES CARRÉS ORDINAIRES. PROPRIÉTÉS DE CES ESTIMATEURS

1 / Le modèle linéaire de régression simple

Considérons le modèle :

(1) $$y_t = ax_t + b + \varepsilon_t \qquad t = 1, 2 \ldots, \mathrm{T}$$

dans lequel :

Y représente une variable endogène,
X une variable exogène,
ε une variable aléatoire dont les caractéristiques seront précisées au cours des hypothèses.

On dispose de T observations sur Y et X, c'est-à-dire de T couples (x_t, y_t) qui sont les réalisations de X et Y.

a et *b* sont des paramètres réels inconnus que l'on se propose d'estimer à l'aide des observations (x_t, y_t) connues.

2 / Hypothèses fondamentales

- H_1 | x_t et y_t représentent des grandeurs numériques observées sans erreur.
  Y est aléatoire par l'intermédiaire de ε.
  X, variable explicative, est considérée comme une donnée dans le modèle.

- H_2 | Hypothèse concernant la distribution des erreurs ε

1) ε *est distribuée selon une loi indépendante du temps.*

$$E(\varepsilon_t) = 0 \qquad \forall\, t = 1, 2, \ldots, T$$

$V(\varepsilon_t) = \sigma_\varepsilon^2$ (1), quantité finie quel que soit t.

Les réalisations de ε sont indépendantes des valeurs prises par X au cours du temps.

Lorsque cette hypothèse est réalisée on dit qu'il y a : *homoscédasticité.* Dans le cas contraire on aura *hétéroscédasticité.*

2) *Indépendance des erreurs.*

Deux erreurs relatives à deux observations différentes t et t' sont indépendantes entre elles, ce qui entraîne :

$$\mathrm{Cov}\,(\varepsilon_t, \varepsilon_{t'}) = 0$$

soit : $\quad E(\varepsilon_t . \varepsilon_{t'}) = 0 \quad$ car $E(\varepsilon_t) = E(\varepsilon_{t'}) = 0.$

3) *On pourra supposer de plus que la loi de* ε *est une loi normale.*

- H_3 | Hypothèse concernant la variable explicative X

On suppose lorsque T devient très grand que les premiers moments empiriques de X sont finis.

$$\frac{1}{T}\sum_{t=1}^{T} x_t \xrightarrow[T\to\infty]{} \bar{x}_0 \text{ quantité finie}$$

(moyenne empirique)

$$\frac{1}{T}\sum_{t=1}^{T} (x_t - \bar{x})^2 \xrightarrow[T\to\infty]{} s^2 \text{ quantité finie}$$

(variance empirique).

On utilisera cette hypothèse pour préciser les propriétés asymptotiques des estimateurs de a et b.

Ces hypothèses peuvent sembler très restrictives. En fait, nous verrons ultérieurement les conséquences de l'abandon de certaines d'entre elles sur les propriétés des estimateurs des coefficients a et b.

(1) $V(\varepsilon_t) = E[\varepsilon_t - E(\varepsilon_t)]^2 = E(\varepsilon_t^2)$.

3 / Détermination des estimateurs de a et b par la méthode des Moindres Carrés Ordinaires (MCO)

On se propose de déterminer $\hat{a}$ et $\hat{b}$ dont les réalisations minimisent la somme des carrés des erreurs soit :

$$(1) \qquad \sum_{t=1}^{T} \varepsilon_t^2 = \sum_{t=1}^{T} [y_t - ax_t - b]^2 = \varphi(a, b).$$

Cette expression dépend de 2 paramètres a et b.
Pour que $\varphi(a, b)$ soit minimale il faut que :

1° La condition *nécessaire* suivante soit réalisée :

$$(2) \qquad \begin{cases} \dfrac{\partial \varphi}{\partial a} = 0 \\[2ex] \dfrac{\partial \varphi}{\partial b} = 0. \end{cases}$$

Traduisons cette condition.

$$(3) \qquad \begin{cases} \sum\limits_{t=1}^{T} (y_t - ax_t - b)\, x_t = 0 \\[2ex] \sum\limits_{t=1}^{T} (y_t - ax_t - b) = 0. \end{cases}$$

En sommant par rapport à t et en divisant chaque terme des expressions (3) par T il vient :

$$(4) \qquad \begin{cases} \dfrac{1}{T} \sum\limits_{t=1}^{T} y_t . x_t - a \dfrac{1}{T} \sum\limits_{t=1}^{T} x_t^2 - b\bar{x} = 0 \\[2ex] \bar{y} \quad - a\bar{x} \qquad - b \quad = 0. \end{cases}$$

En appelant $\hat{a}$ et $\hat{b}$ les solutions de (4) on obtient :

$$(5) \qquad \begin{cases} \hat{a} = \dfrac{\sum\limits_{t=1}^{T} y_t . x_t - T\bar{y}\bar{x}}{\sum\limits_{t=1}^{T} x_t^2 - T\bar{x}^2} = \dfrac{\sum\limits_{t=1}^{T} (y_t - \bar{y})(x_t - \bar{x})}{\sum\limits_{t=1}^{T} (x_t - \bar{x})^2} \\[3ex] \hat{b} = \bar{y} - \hat{a}\bar{x}. \end{cases}$$

$\hat{a}$ est une variable aléatoire puisque fonction de y_t, elle-même aléatoire. Il en est de même de $\hat{b}$.

2° La condition *suffisante* soit satisfaite :

$$(6) \qquad \frac{\partial^2 \varphi}{\partial a^2} > 0; \qquad \begin{vmatrix} \dfrac{\partial^2 \varphi}{\partial a^2} & \dfrac{\partial^2 \varphi}{\partial a\, \partial b} \\ \dfrac{\partial^2 \varphi}{\partial b\, \partial a} & \dfrac{\partial^2 \varphi}{\partial b^2} \end{vmatrix} > 0.$$

Cette condition est réalisée.

Les estimateurs de a et b sont donc donnés par les expressions (5).

On notera dorénavant pour simplifier $\sum_{t=1}^{T}$ par $\sum_{t}$.

4 / Propriétés des estimateurs $\hat{a}$ et $\hat{b}$

Nous allons montrer compte tenu des hypothèses posées au § 2 que les estimateurs obtenus par les MCO sont des estimateurs sans biais et convergents. Nous allons au préalable transformer les expressions de $\hat{a}$ et $\hat{b}$ en les exprimant en fonction des coefficients a et b.

1 | Transformation des expressions (5) obtenues au § 3

Nous avons considéré le modèle :

$$(1) \qquad y_t = ax_t + b + \varepsilon_t.$$

Soit en sommant par rapport à t et en divisant par T :

$$(2) \qquad \bar{y} = a\bar{x} + b + \bar{\varepsilon}.$$

Retranchons membre à membre (2) de (1) :

$$(3) \qquad y_t - \bar{y} = a(x_t - \bar{x}) + (\varepsilon_t - \bar{\varepsilon}).$$

En remplaçant $y_t - \bar{y}$ par l'expression ci-dessus dans $\hat{a}$ il vient :

$$\hat{a} = \frac{\sum_t [a(x_t - \bar{x}) + (\varepsilon_t - \bar{\varepsilon})](x_t - \bar{x})}{\sum_t (x_t - \bar{x})^2}$$

$$\hat{a} = a + \frac{\sum_t (x_t - \bar{x})(\varepsilon_t - \bar{\varepsilon})}{\sum_t (x_t - \bar{x})^2}$$

ou encore :

$$(4) \qquad \hat{a} = a + \frac{\sum_t (x_t - \bar{x})\,\varepsilon_t}{\sum_t (x_t - \bar{x})^2}$$

car $\bar{\varepsilon} \sum_t (x_t - \bar{x}) = 0.$

De même on sait que :

$$(5) \qquad \begin{cases} \bar{y} = \hat{a}\bar{x} + \hat{b} \\ \bar{y} = a\bar{x} + b + \bar{\varepsilon}. \end{cases}$$

On en déduit :

$$(6) \qquad \hat{b} = b + \bar{\varepsilon} - (\hat{a} - a)\,\bar{x}.$$

2 | $\hat{a}$ ET $\hat{b}$ SONT DES ESTIMATEURS SANS BIAIS DE a ET b

D'après (4) :

$$(7) \qquad \hat{a} = a + \sum_t \omega_t \,.\, \varepsilon_t \,.$$

En posant :

$$\omega_t = \frac{x_t - \bar{x}}{\sum_t (x_t - \bar{x})^2} \text{ quantité exogène.}$$

D'où :

$$\mathrm{E}(\hat{a}) = a + \sum_t \omega_t \,\mathrm{E}(\varepsilon_t).$$

$$(8) \qquad \mathrm{E}(\hat{a}) = a \quad \text{puisque } \mathrm{E}(\varepsilon_t) = 0.$$

De même :

$$E(\hat{b}) = b + E(\bar{\varepsilon}) - \bar{x}E(\hat{a} - a)$$

$$E(\bar{\varepsilon}) = \frac{1}{T}\sum_t E(\varepsilon_t) = 0$$

$$E(\hat{a} - a) = 0.$$

Donc :

$$E(\hat{b}) = b. \tag{9}$$

$\hat{a}$ et $\hat{b}$ sont des estimateurs sans biais de a et b.

3 | $\hat{a}$ ET $\hat{b}$ SONT DES ESTIMATEURS CONVERGENTS DE a ET b

On sait que lorsque :

$$E(\hat{a}) = a$$

$$E(\hat{b}) = b$$

il suffit pour que $\hat{a}$ et $\hat{b}$ soient convergents que :

$$V(\hat{a}) \underset{T \to \infty}{\to} 0$$

et :

$$V(\hat{b}) \underset{T \to \infty}{\to} 0.$$

— *Calcul de* $V(\hat{a})$. — D'après (7) :

$$V(\hat{a}) = E(\hat{a} - a)^2 = E[(\sum_t \omega_t . \varepsilon_t)^2]$$

$$V(\hat{a}) = E[\sum_t \omega_t^2 . \varepsilon_t^2 + 2 \sum_{t < t'} \omega_t\, \omega_{t'}\, \varepsilon_t . \varepsilon_{t'}]$$

$$V(\hat{a}) = \sum_t \omega_t^2\, E(\varepsilon_t^2) + 2 \sum_{t < t'} \omega_t\, \omega_{t'}\, E(\varepsilon_t . \varepsilon_{t'}).$$

Or d'après l'hypothèse H_2 du § 2 :

$$E(\varepsilon_t^2) = \sigma_\varepsilon^2 \qquad \forall\, t = 1, 2, \ldots, T$$

$$E(\varepsilon_t . \varepsilon_{t'}) = 0 \qquad t \neq t'.$$

Donc :

(10) $$V(\hat{a}) = \sigma_\varepsilon^2 \sum_t \omega_t^2.$$

Or :

$$\sum_t \omega_t^2 = \sum_t \left[\frac{(x_t - \bar{x})}{\sum_t (x_t - \bar{x})^2} \right]^2$$

$$= \frac{1}{\sum_t (x_t - \bar{x})^2}.$$

D'où :

(11) $$\boxed{V(\hat{a}) = \frac{\sigma_\varepsilon^2}{\sum_t (x_t - \bar{x})^2}.}$$

On sait que d'après l'hypothèse H_3 :

$$\frac{1}{T} \sum_t (x_t - \bar{x})^2 \xrightarrow[T \to \infty]{} s^2.$$

Donc :

(12) $$V(\hat{a}) \xrightarrow[T \to \infty]{} \frac{\sigma_\varepsilon^2}{Ts^2} \to 0.$$

$\hat{a}$ est bien un estimateur qui *converge en probabilité* vers a et on notera :

$$\hat{a} \xrightarrow[T \to \infty]{P} a.$$

— *Calcul de* $V(\hat{b})$. — D'après (6) :

$$V(\hat{b}) = E(\hat{b} - b)^2$$

(13) $$V(\hat{b}) = E[\bar{\varepsilon}^2 - 2\bar{x}(\hat{a} - a)\, \bar{\varepsilon} + \bar{x}^2(\hat{a} - a)^2].$$

On sait que :

$$\bar{\varepsilon} = \frac{1}{T} \sum_t \varepsilon_t;$$

donc :

$$(14)\quad \begin{cases} E(\bar{\varepsilon}) = 0 \\ V(\bar{\varepsilon}) = \dfrac{TV(\varepsilon)}{T^2} = \dfrac{\sigma_\varepsilon^2}{T} \end{cases}$$

(car les ε_t sont des variables aléatoires indépendantes).

$$E[(\hat{a} - a)\,\bar{\varepsilon}] = E\left[(\sum_t \omega_t\, \varepsilon_t)\left(\frac{1}{T}\sum_t \varepsilon_t\right)\right]$$

$$= \frac{1}{T}\, E(\sum_t \omega_t\, \varepsilon_t^2 + \sum_{t \neq t'} \omega_t\, \varepsilon_t\, \varepsilon_{t'})$$

$$= \frac{1}{T}\, \sigma_\varepsilon^2 \sum_t \omega_t + \sum_{t \neq t'} \omega_t\, E(\varepsilon_t\, \varepsilon_{t'}).$$

Mais :

$$\sum_t \omega_t = \sum_t \left[\frac{x_t - \bar{x}}{\sum_t (x_t - \bar{x})^2}\right] = \frac{1}{\sum_t (x_t - \bar{x})^2} \sum_t (x_t - \bar{x}) = 0.$$

Donc :

$$(15)\quad \left.\begin{cases} \dfrac{1}{T}\, \sigma_\varepsilon^2 \sum_t \omega_t = 0 \\ \text{et} \\ \sum_{t \neq t'} \omega_t\, E(\varepsilon_t\, \varepsilon_{t'}) = 0 \end{cases}\right\} \Rightarrow E[(\hat{a} - a)\,\bar{\varepsilon}] = 0.$$

Il vient alors d'après (13), (14) et (15) :

$$V(\hat{b}) = \frac{\sigma_\varepsilon^2}{T} + \bar{x}^2\, V(\hat{a})$$

$$(16)\quad V(\hat{b}) = \sigma_\varepsilon^2 \left[\frac{1}{T} + \frac{\bar{x}^2}{\sum_t (x_t - \bar{x})^2}\right].$$

Lorsque $T \to \infty$, $\dfrac{1}{T} \to 0$, $\dfrac{1}{\sum_t (x_t - \bar{x})^2} \to \dfrac{1}{Ts^2} \to 0.$

Donc $\hat{b}$ *converge en probabilité* vers b. On notera :

$$V(\hat{b}) \xrightarrow[T\to\infty]{} 0 \Longrightarrow \hat{b} \xrightarrow[T\to\infty]{P} b.$$

Avant de conclure ce paragraphe il n'est pas inutile de calculer la covariance de $(\hat{a}, \hat{b})$ ce qui nous permettra d'obtenir la matrice des variances et covariances des estimateurs $\hat{a}$ et $\hat{b}$.

4 | Calcul de la covariance de $(\hat{a}, \hat{b})$

$$\begin{aligned}
\mathrm{Cov}\,(\hat{a}, \hat{b}) &= \mathrm{E}[(\hat{a} - a)(\hat{b} - b)] \\
&= \mathrm{E}[(\hat{a} - a)(\bar{\varepsilon} - \bar{x}(\hat{a} - a))] \\
&= \mathrm{E}[(\hat{a} - a)\,\bar{\varepsilon} - \bar{x}(\hat{a} - a)^2].
\end{aligned}$$

D'après (15), $\mathrm{E}[(\hat{a} - a)\,\bar{\varepsilon}] = 0$.
Donc :

$$\begin{aligned}
\mathrm{Cov}\,(\hat{a}, \hat{b}) &= -\,\bar{x}\,\mathrm{E}(\hat{a} - a)^2 \\
&= -\,\bar{x}\,\mathrm{V}(\hat{a})
\end{aligned}$$

$$\mathrm{Cov}\,(\hat{a}, \hat{b}) = -\,\frac{\sigma_\varepsilon^2\,\bar{x}}{\sum_t (x_t - \bar{x})^2}.$$

En appelant $\Omega_{(\hat{a}, \hat{b})}$ *la matrice des variances et covariances* de $\hat{a}$ et $\hat{b}$ nous avons, puisque $\mathrm{cov}\,(\hat{a}, \hat{b}) = \mathrm{cov}\,(\hat{b}, \hat{a})$:

$$\Omega_{(\hat{a}, \hat{b})} = \begin{pmatrix} \mathrm{V}(\hat{a}) & \mathrm{cov}\,(\hat{a}, \hat{b}) \\ \mathrm{cov}\,(\hat{a}, \hat{b}) & \mathrm{V}(\hat{b}) \end{pmatrix}$$

$$(18) \qquad \Omega_{(\hat{a}, \hat{b})} = \sigma_\varepsilon^2 \begin{pmatrix} \dfrac{1}{\sum_t (x_t - \bar{x})^2} & -\,\dfrac{\bar{x}}{\sum_t (x_t - \bar{x})^2} \\[2ex] -\,\dfrac{\bar{x}}{\sum_t (x_t - \bar{x})^2} & \dfrac{1}{\mathrm{T}} + \dfrac{\bar{x}^2}{\sum_t (x_t - \bar{x})^2} \end{pmatrix}.$$

Remarquons que les expressions qui constituent la matrice des variances et covariances de $\widehat{a}$ et $\widehat{b}$ contiennent σ_ε^2 variance des ε_t. Or cette variance est inconnue.

Pour obtenir une estimation de $\Omega_{(\widehat{a},\widehat{b})}$ il est donc maintenant nécessaire de déterminer une estimation de σ_ε^2, soit $\widehat{\sigma}_\varepsilon^2$.

5 / Détermination d'un estimateur sans biais de σ_ε^2

Posons :

$$\widehat{y}_t = \widehat{a}x_t + \widehat{b}$$

et soit $\widehat{\varepsilon}_t = y_t - \widehat{y}_t$ un estimateur de ε_t.

On peut écrire :

$$\widehat{\varepsilon}_t = y_t - \widehat{a}x_t - \widehat{b}$$

$$\widehat{\varepsilon}_t = ax_t + b + \varepsilon_t - \widehat{a}x_t - \widehat{b}$$

$$\widehat{\varepsilon}_t = \varepsilon_t - (\widehat{a} - a)\, x_t - (\widehat{b} - b). \tag{1}$$

Remarquons que si l'on tient compte de la convergence en probabilité de $\widehat{a}$ et $\widehat{b}$ vers a et b la distribution de $\widehat{\varepsilon}_t$ converge en probabilité vers celle de ε_t d'après (1) ci-dessus.

Remplaçons $\widehat{b} - b$ par son expression en fonction de $(\widehat{a} - a)$ dans (1), il vient :

$$\widehat{\varepsilon}_t = \varepsilon_t - (\widehat{a} - a)\, x_t - [\bar{\varepsilon} - (\widehat{a} - a)\, \bar{x}]$$

$$\widehat{\varepsilon}_t = \varepsilon_t - \bar{\varepsilon} - (\widehat{a} - a)\,(x_t - \bar{x}). \tag{2}$$

On en déduit :

$$\frac{1}{T}\sum_t \widehat{\varepsilon}_t^2 = \frac{1}{T}\sum_t (\varepsilon_t - \bar{\varepsilon})^2 - 2(\widehat{a} - a)\,\frac{1}{T}\sum_t (x_t - \bar{x})\,(\varepsilon_t - \bar{\varepsilon}) + (\widehat{a} - a)^2\,\frac{1}{T}\sum_t (x_t - \bar{x})^2.$$

Or, si l'on tient compte de l'expression de $(\widehat{a} - a)$ obtenue antérieurement, il vient :

$$\frac{1}{T}\sum_t \widehat{\varepsilon}_t^2 = \frac{1}{T}\sum_t (\varepsilon_t - \bar{\varepsilon})^2 - (\widehat{a} - a)^2\,\frac{1}{T}\sum_t (x_t - \bar{x})^2. \tag{3}$$

Posons :

(4) $$\sigma'^2 = \frac{1}{T} \sum_t (\varepsilon_t - \bar{\varepsilon})^2.$$

On démontre en statistique que :

(5) $$E(\sigma'^2) = \sigma_\varepsilon^2 \left(1 - \frac{1}{T}\right).$$

D'après (3) on peut écrire :

$$E\left(\frac{1}{T} \sum_t \hat{\varepsilon}_t^2\right) = E(\sigma'^2) - \frac{1}{T} \sum_t (x_t - \bar{x})^2 \, V(\hat{a})$$

$$= \sigma_\varepsilon^2 \left(1 - \frac{1}{T}\right) - \frac{\sigma_\varepsilon^2}{T}.$$

(6) $$E\left(\frac{1}{T} \sum_t \hat{\varepsilon}_t^2\right) = \sigma_\varepsilon^2 \left(1 - \frac{2}{T}\right).$$

On en déduit que :

(7) $$E\left(\frac{\sum_t \hat{\varepsilon}_t^2}{T - 2}\right) = \sigma_\varepsilon^2.$$

Soit en posant :

(8) $$\hat{\sigma}_\varepsilon^2 = \frac{1}{T - 2} \sum_t \hat{\varepsilon}_t^2$$

$$E(\hat{\sigma}_\varepsilon^2) = \sigma_\varepsilon^2.$$

$\hat{\sigma}_\varepsilon^2$ est donc un estimateur sans biais de la variance des résidus.

Remarquons que le modèle $y_t = ax_t + b + \varepsilon_t$ comporte deux coefficients a et b à estimer. Le dénominateur de $\hat{\sigma}_\varepsilon^2$ est alors égal à $T - 2$. $(T - 2)$ constitue ce que l'on appellera le *nombre de degrés de liberté*.

Nous reviendrons ultérieurement sur cette remarque.

Nous pouvons donc maintenant donner un résumé des formules des estimateurs des paramètres inconnus du modèle :

Ce cours d'Économétrie est destiné aux étudiants de Maîtrise de Sciences Économiques. Il représente un enseignement actuellement dispensé à l'Université PARIS 2, sous des formes spécifiques, aux étudiants des Maîtrises d'Économie, de Gestion et d'Économétrie.

A ce jour, les livres d'Économétrie existants étaient soit trop spécialisés, soit d'un niveau théorique difficilement abordable par des étudiants de Maîtrise.

Ce nouveau manuel, où chaque chapitre est suivi d'exercices corrigés, concilie une approche formelle rigoureuse correspondant à la formation de mathématiciens des deux auteurs, avec un souci pédagogique de faire acquérir aux étudiants une méthodologie rigoureuse ainsi qu'une maîtrise intelligente des développements informatiques actuels de l'Économétrie. Ces deux aspects, réunis pour la première fois, doivent permettre aux étudiants de développer l'approfondissement de leurs connaissances dans un domaine actuellement en grand développement.

9 782130 460176

188 FF

22405923/3/94

Didier Thimonier