

Économétrie appliquée des séries temporelles

Guide pratique

**

*

I. INTRODUCTION

Qu'est ce que l'économétrie ?

L'économétrie peut être définie comme l'application d'un ensemble de techniques économiques, mathématiques et statistiques à l'étude du domaine économique.

Elle constitue aujourd'hui le principal outil de validation des modèles théoriques et d'analyse des phénomènes économiques.

L'économétrie a deux objectifs :

- 1) valider (ou invalider) les relations entre des variables économiques mises en avant par les modèles théoriques
- 2) mettre au jour des phénomènes économiques laissés dans l'ombre de la théorie (recherche de relations de causalité lorsque les modèles atteignent leurs limites)

Les 10 commandements de l'économétrie appliquée :

- Règle n° 1 : Utiliser la théorie économique et son bon sens
- Règle n° 2 : Ne pas chercher à répondre juste à une mauvaise question
- Règle n° 3 : S'informer de la qualité des données
- Règle n° 4 : Examiner les données
- Règle n° 5 : Rester simple dans l'analyse.
- Règle n° 6 : Proposer des résultats d'analyse compréhensibles
- Règle n° 7 : Être rigoureux et ne pas chercher à retenir une technique et/ou un échantillon de données en fonction du résultat recherché.
- Règle n° 8 : Être préparé aux compromis.
- Règle n° 9 : Ne pas confondre significativité statistique et pertinence économique.
- Règle n° 10 : Procéder à des tests de robustesse.

II. PRISE EN MAIN DE EVIEWS

Présentation de Eviews :

Eviews est un logiciel d'analyse économétrique qui s'appuie sur une interface conviviale et intuitive de gestion et de traitement des données. La plupart des tâches courantes de l'économétrie (régressions, tests de significativité, de violations des hypothèses, etc.) sont pré-programmées.

Il fonctionne avec des fenêtres de dialogue.

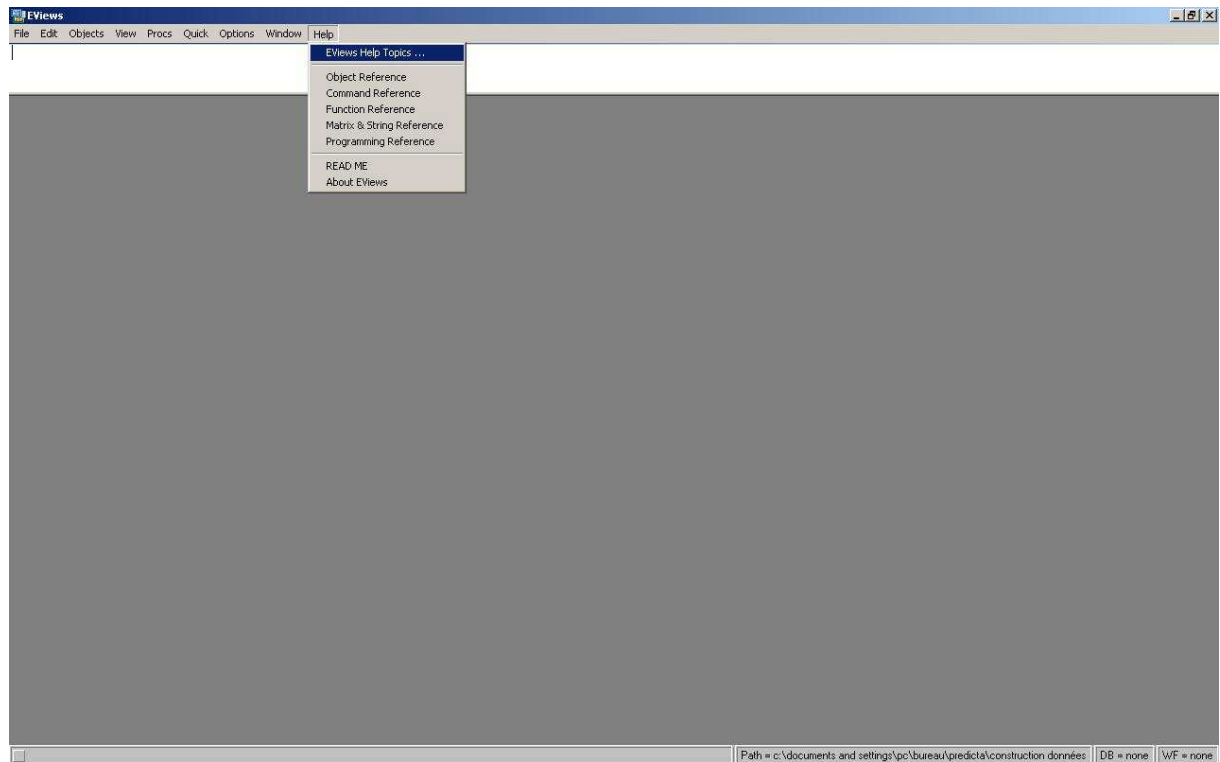
Après avoir double-cliqué sur l'icône « Eviews », la fenêtre suivante apparaît :



Sous la barre de menu se situe la barre de commande. Toutes les commandes d'Eviews peuvent être décrites et exécutées dans cette barre.

Tout en bas de la fenêtre apparaît la ligne de statut, qui fournit plusieurs renseignements relatifs au chemin d'accès emprunté par le logiciel, à la base de données et au workfile par défaut.

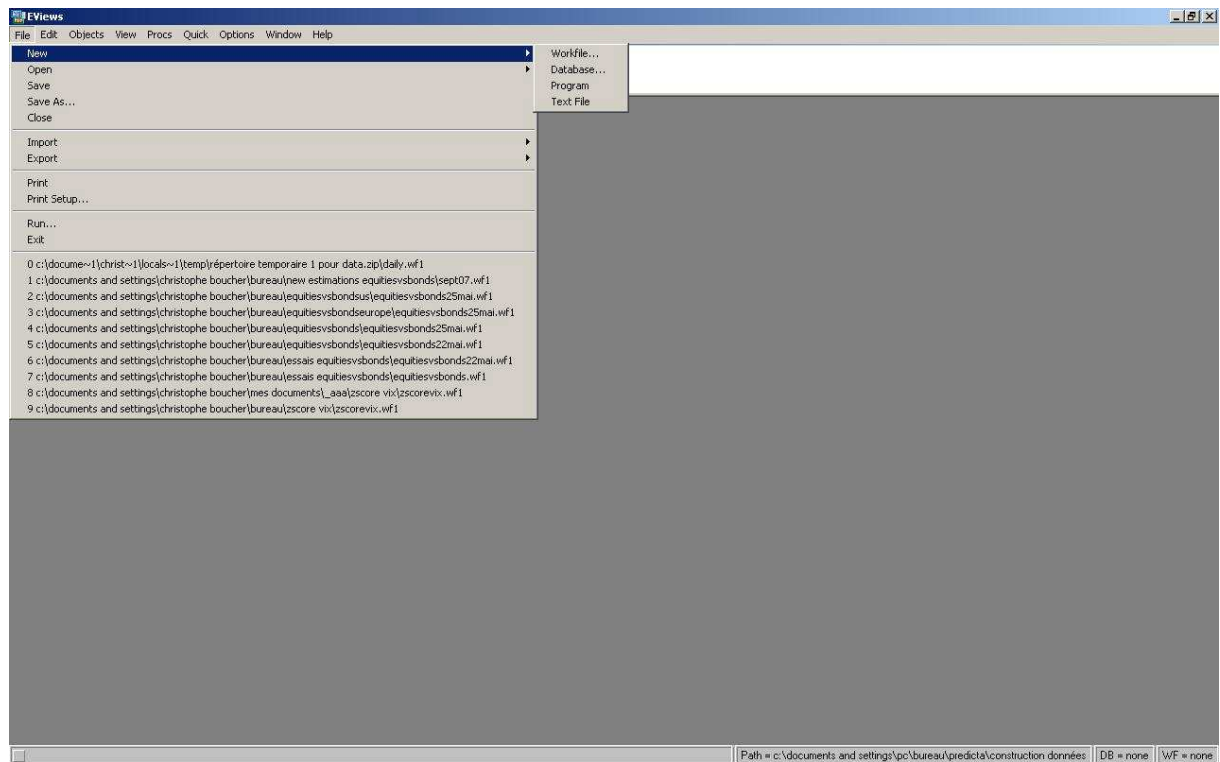
Eviews propose une aide très complète intégrée au logiciel (Menu help). L'essentiel de trouve dans *Eviews Help Topics*, rubrique *index*.



L'essentiel du langage de programmation est regroupé sous différentes rubriques dans *Programming Reference*.

La particularité de ce logiciel : effectuer les tâches (estimations, prévisions, etc.) directement dans l'interface par le biais des menus déroulants.

Exemple : Aller dans *File*, rubrique *New* pour créer une base de données, un workfile, un programme.



Pourtant, il peut être préférable de conserver l'ensemble de son travail sous forme de programme (.prg), afin d'éviter les redites, les oublis contraignant à réitérer des manipulations effectuées auparavant, etc.

Importation et gestion des bases de données

Il peut importer et gérer plusieurs types de fichiers contenant des données, notamment les fichiers Excel ou text.

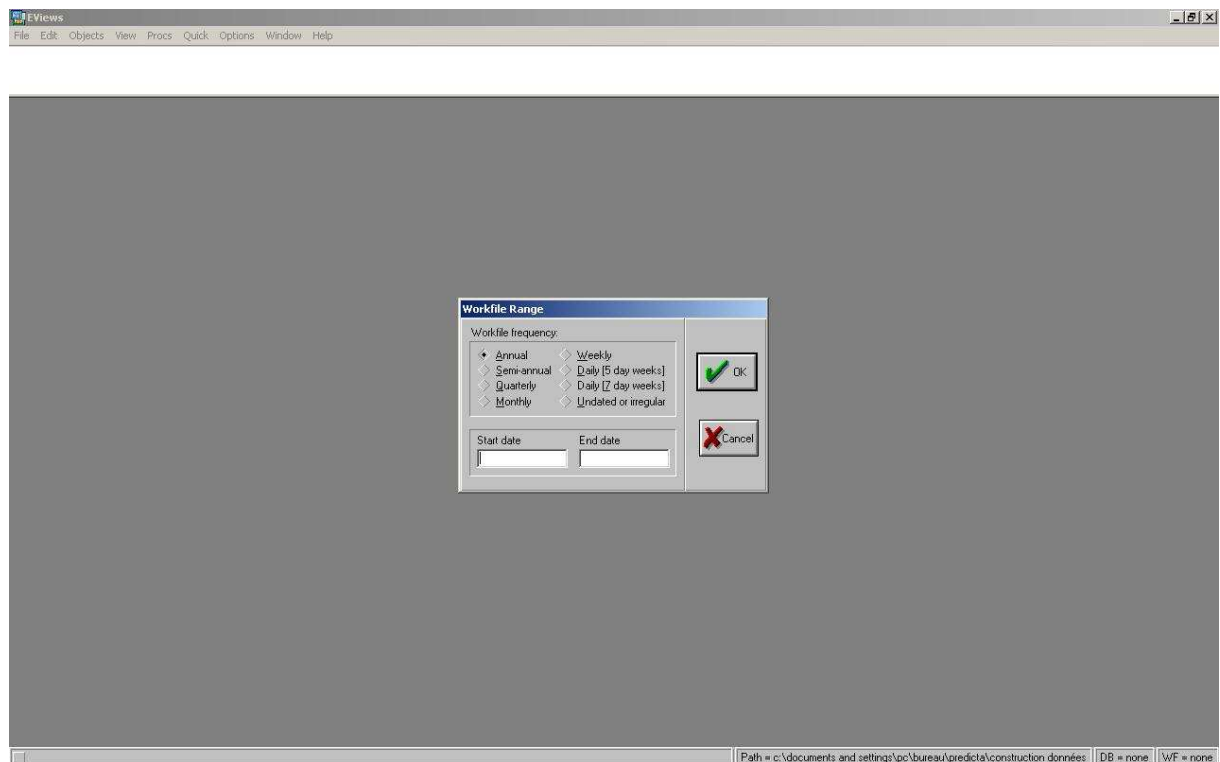
Avant cela, il faut créer un fichier destiné à les recevoir qui peut être de deux types :

- soit une base de données (.edb)
- soit un workfile (.wfl).

Pour créer un workfile :

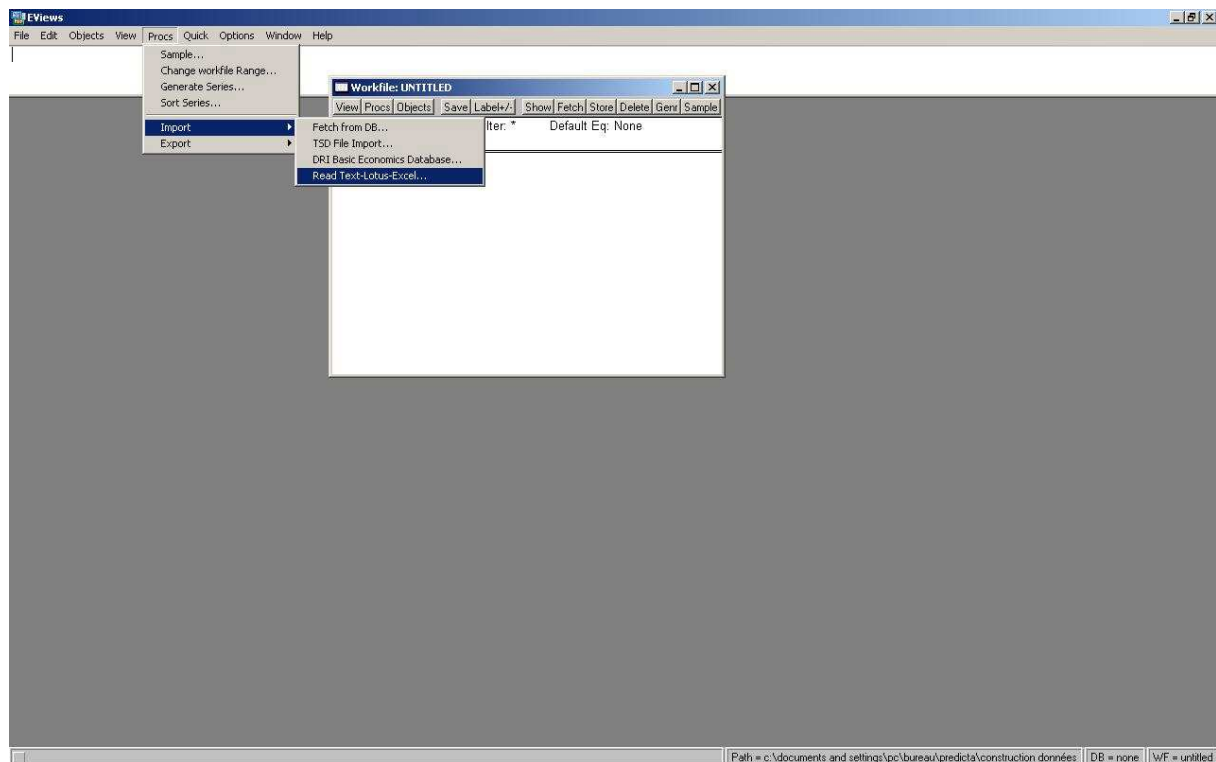
File/New/Workfile et ensuite remplir la fréquence ainsi que les dates de début et de fin.

Exemple : pour des données trimestrielles, 1968:04 et 2007:02



Ce workfile peut être enregistré et renommé en cliquant sur l'option *Save* du menu du worfile.

Il faut ensuite importer les données en cliquant sur *Procs/Import/Read Text-Lotus-Excel...*



Il faut choisir le sens d'importation (colonne ou ligne), le nombre de séries à importer et l'intervalle temporel d'importation.

Les séries apparaissent ensuite dans le Workfile. Chacune porte le nom indiqué dans la première ligne ou première colonne du fichier d'origine.

OU

Créer un programme en allant dans *File/New/Program*

Taper dans la fenêtre de saisie (blanche) : *workfile nom frequence debut fin*

La fréquence se spécifie par *a*, *m* ou *q*.

Exemple : *workfile demo q 1952:01 1996:04*

Pour importer les données, il faut indiquer le chemin d'accès du fichier sous la forme :

cd C:\EViews3\Example files

La commande *cd* signifie change directory (changer le répertoire). Il suffit maintenant d'indiquer à Eviews les données à importer et à lire avec la commande suivante :

Read(première case, s = nom de la feuille excel) nomdufichier.xls nombre de colonnes

Soit par exemple :

read(B2,s=demo) demo.xls 4

Afin de renommer, copier ou détruire une série, un simple clic droit suffit.

Les opérations d'addition, soustraction, etc. des séries s'effectuent à partir du programme ou de la barre de commande en utilisant la commande *genr*.

Exemple :

genr serie3 = serie1/serie2

Pour un taux de croissance :

*Genr croissance = ((serie-serie(-1))/serie(-1))*100*

On peut aussi utiliser la fonction proposée par Eviews :

Genr croissance = @pch(serie)

III. RAPPELS DE STATISTIQUES

Notions simples de statistiques descriptives

Une analyse statistique descriptive est un préalable à toute étude économétrique. Une fois vos données récupérées, représentez-les graphiquement afin d'avoir une première idée de la distribution des séries et d'identifier d'éventuels *outliers*.

Nous allons caractériser ensuite les distributions de ces variables à l'aide de :

- La moyenne, la somme des observations divisée par le nombre d'observations.

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$$

- La variance, la somme des écart à la moyenne au carré divisé par le nombre d'observations. La variance représente la dispersion de l'échantillon autour de sa moyenne.

$$V(X) = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 = \frac{1}{n} \sum_{i=1}^n (x_i)^2 - \bar{x}^2$$

- L'écart type, la racine de la variance.

$$\sigma_x = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2} = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i)^2 - \bar{x}^2}$$

- Le coefficient de variation, l'écart type divisé par la moyenne. Dans certaines situations, on désire comparer le taux de dispersion de distributions alors que leur échelle de mesure respective ne sont pas comparables. L'objectif du coefficient de variation est de fournir un indice quantitatif permettant cette comparaison compris entre 0 et 1.

$$c = \frac{\mu}{\sigma}$$

- La skewness, la somme des écarts à la moyenne (divisée par l'écart type) au cube divisée par le nombre d'observations. Le signe de ce coefficient indique le type de dissymétrie (positive ou négative) : étalée vers la droite ou vers la gauche. Ce coefficient est nul lorsque la distribution est symétrique.

$$Sk = \frac{\left(\frac{1}{n}\right) \sum_{i=1}^n (R_i - \mu)^3}{\sigma(R)^3}$$

- La kurtosis, la somme des écarts à la moyenne (divisée par l'écart type) à la puissance quatre divisée par le nombre d'observations. Elle mesure le degré d'aplatissement (hors effet de dispersion) en indiquant si la distribution est très plate (platikurtique pour un coefficient faible) ou très pointue (leptokurtique pour un coefficient élevé).

$$K = \frac{\left(\frac{1}{n}\right) \sum_{i=1}^n (R_i - \mu)^4}{\sigma(R)^4}$$

NB : Pour une distribution normale $K = 3$

De manière générale, on appelle quantile les valeurs prises par une variable lorsqu'une série est classée. Le quartile classe la distribution par bloc de 25%, le décile par bloc de 10% et le centile par bloc de 1%.

Analyse de la normalité

L'analyse de la normalité peut être réalisée suivant plusieurs procédures. Nous en présentons ici quelques un seulement.

1) L'analyse graphique

L'analyse graphique consiste à comparer la distribution de la variable concernée avec celle d'une loi normale. Cette comparaison peut se faire par quantiles (QQ plot), par centiles (PP plot) ou via un histogramme.

⇒ Après sélection de la série : *View\Distribution\Kernel Density* ou *QQplot* ou *CDF*

2) Le test de normalité de Jarque Bera

Le test de Jarque Bera cherche à déterminer si les données suivent une loi normale suivant l'hypothèse :

H0 : les données suivent une loi normale

H1 : les données ne suivent pas une loi normale

La statistique du test est égale à :

$$JB = \frac{n-2}{6} (Sk^2 + \frac{1}{4} (K-3)^2)$$

où Sk et K sont respectivement la *skewness* et la *kurtosis*. Sous l'hypothèse de normalité, la statistique JB est distribuée selon un χ^2 à deux degrés de liberté.

Analyse de la corrélation

L'analyse de la corrélation entre deux variables X et Y permet de tester leur indépendance.

Deux variables sont dites indépendantes si leur covariance est nulle.

La covariance est positive pour chaque couple de valeurs qui diffèrent de leur moyenne dans le même sens, et plus négative pour chaque couple de valeur qui diffère de leur moyenne dans le sens opposé.

$$\text{Cov}(X, Y) = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})$$

On peut également calculer le coefficient de corrélation :

$$\text{Corr}(X, Y) = \frac{\text{Cov}(X, Y)}{\sigma(X)\sigma(Y)} .$$

Ses valeurs sont comprises entre -1 et $+1$.

Une valeur de -1 signifie que les titres évoluent en sens inverse. Une valeur de $+1$ qu'ils évoluent dans le même sens, une valeur de 0 qu'il n'y a pas de relation linéaire entre leurs évolutions.

⇒ Après sélection des séries en groupe : *View/Correlation/Common Sample*

Le test d'autocorrélation

Une variable est dite indépendante lorsque sa réalisation en t ne dépend pas de sa réalisation en $t - n$. Pour tester l'indépendance d'une série, il convient de s'intéresser à son autocorrélation.

Définition 1 : La fonction d'autocorrélation est la fonction notée ρ_k qui mesure la corrélation de la série avec elle-même décalée de k périodes :

$$\rho_k = \frac{\text{cov}(x_t, x_{t-k})}{\sigma_{y_t} \sigma_{y_{t-k}}} = \frac{\sum_{i=k+1}^n (x_i - \bar{x})(x_{i-k} - \bar{x})}{\sqrt{\frac{1}{n} \sum_{i=k+1}^n (x_i)^2 - \bar{x}^2} \cdot \sqrt{\frac{1}{n} \sum_{i=k+1}^n (x_{i-k})^2 - \bar{x}^2}}$$

Le graphique de la fonction d'autocorrélation est appelé *corrélogramme*.

Définition 2 : La fonction d'autocorrélation partielle mesure la corrélation entre x_t et x_{t-k} , l'influence des autres variables ($x_{t-1}, x_{t-2}, \dots, x_{t-k+1}$) ayant été retirée.

⇒ Après sélection de la série : *View/Correlogramm*

La stationnarité

Une série aléatoire (dite encore *processus stochastique*), z_t , est *stationnaire* si :

- $E(y_t) = \mu \quad \forall t$ (moyenne constante qui ne dépend pas du temps)
- $\text{Var}(y_t) = \sigma^2 < \infty \quad \forall t$ (variance finie et qui ne dépend pas du temps)
- $\text{Corr}(y_t, y_{t+k}) = \gamma_k$ (son autocorrélation ne dépend pas du temps mais de l'écart entre les observations)

III. LE MODÈLE LINÉAIRE SIMPLE & MULTIPLE ET LES TESTS D'HYPOTHÈSES

Le modèle linéaire simple :

Une variable endogène est expliquée par une variable exogène

Exemple : la fonction de consommation keynésienne :

$$C = a_0 + a_1 Y$$

C = consommation

Y = revenu

a_0 = consommation autonome ou incompressible

a_1 = propension marginale à consommer

Nous pouvons distinguer trois types de spécification :

- le modèle en série temporelle où les variables représentent des phénomènes observés à intervalles de temps réguliers (un individu)
- le modèle en coupe instantanée où les variables représentent des phénomènes observés au même instant mais sur plusieurs individus (pays, entreprises)
- le modèle en panel qui associe la dimension temporelle à la pluralité des individus.

Spécification du modèle	série temporelle	coupe instantanée	panel
Forme du modèle	$C_t = a_0 + a_1 Y_t$	$C_i = a_0 + a_1 Y_i$	$C_{i,t} = a_0 + a_1 Y_{i,t}$

Quel que soit le modèle, le modèle spécifié n'est qu'une caricature de la réalité.

Expliquer la consommation par le revenu est insuffisant

On regroupe tous ces facteurs omis par le modèle théorique par un terme synthétique dénommé résidu qui va également tenir compte des erreurs de mesure liées aux données.

L'introduction de ce résidu permet de passer du modèle théorique au modèle économétrique :

$$C_t = a_0 + a_1 Y_t + \varepsilon_t$$

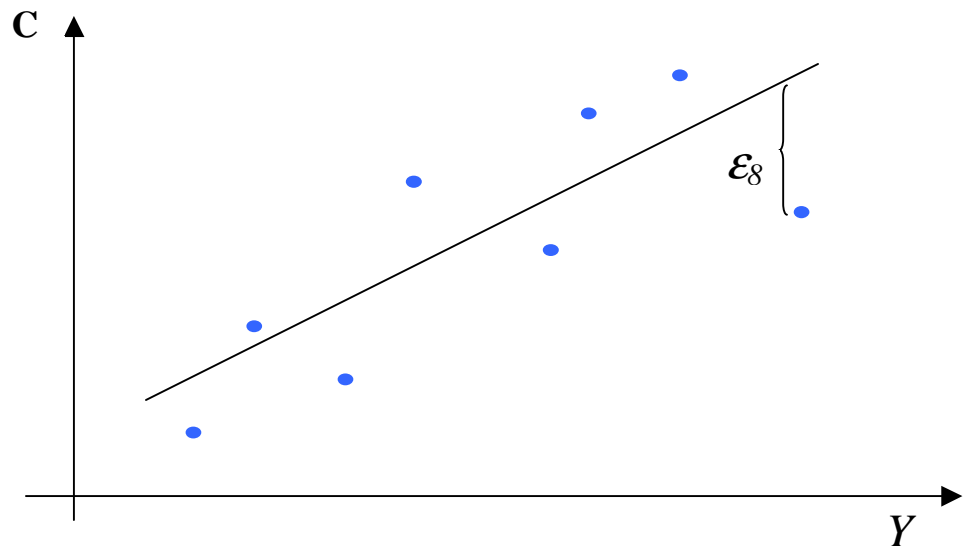
La tâche de l'économètre : estimer les paramètres du modèle (a_0 et a_1)

Les estimateurs des paramètres notés \hat{a}_0 et \hat{a}_1 sont des variables aléatoires qui étant fonction d' ε , suivent la même loi de probabilité que ce dernier.

La moyenne et l'écart type de ces coefficients permettent de construire des tests de validité du modèle estimé.

Les estimateurs des Moindres Carrés ordinaires (MCO)

L'estimateur des MCO permet d'obtenir les paramètres en minimisant la distance au carré entre chaque observation et la droite de régression.



On minimise l'écart entre les valeurs observées du modèle et les valeurs du modèle empirique utilisé, c'est à dire le résidu.

$$\text{Min} \sum_{t=1}^n \varepsilon_t^2$$

$$\text{où } \varepsilon_t = C_t - \hat{a}_0 - \hat{a}_1 Y_t = C_t - \hat{C}_t$$

Principales hypothèses :

H1 : $E(\varepsilon_t) = 0$; le modèle est en moyenne bien spécifié et donc l'erreur moyenne est nulle.

H2 : $E(\varepsilon_t^2) = \sigma_\varepsilon^2$; la variance du terme d'erreur est constante (homoscédasticité contre hétéroscédasticité) : le risque d'amplitude de l'erreur est le même quelle que soit la période.

H3 : $E(\varepsilon_t \varepsilon_{t'}) = 0$; absence d'autocorrélation des erreurs. Les erreurs doivent être indépendantes : l'erreur en t ne doit pas influencer l'erreur en t' et inversement.

H4 : $\text{Cov}(Y_t, \varepsilon_t) = 0$; l'erreur est indépendante de la variable explicative : exogénéité des variables.

H5 : $\varepsilon_t \rightarrow N(0, \sigma_\varepsilon^2)$; le terme d'erreur suit une loi normale de moyenne nulle et de variance σ_ε^2 . Hypothèse d'absence de point aberrants au sein de l'échantillon (« outliers »).

H1, H2 et H3 sont nécessaires pour les estimateurs des MCO affichent les bonnes propriétés statistiques.

H4 rappelle que la variable explicative doit être indépendante du terme d'erreur pour que l'estimation soit sans biais (sinon biais d'endogénéité → recours à des variables instrumentales).

H5 est indispensable pour pouvoir discuter de la significativité des estimations obtenues.

Qu'attend-on d'un bon estimateur ?

- 1) sans biais : estimations en moyenne égales à la vraie valeur des paramètres, sans écart systématique.
- 2) Convergent : la variance de l'estimateur décroît avec le nombre d'observations.

Les MCO présentent ces deux propriétés permettant d'aboutir au meilleur ajustement possible. Ces estimateurs sont dit BLUE (*Best Linear Unbiased Estimators*), c'est à dire meilleurs estimateurs linéaires sans biais.

Tests de significativité

Vérifier la significativité statistique des estimateurs (pas significativement différents de 0 ; variabilité des estimateurs).

Juger l'intérêt du modèle

H0 : $a_i = 0$

H1 : $a_i \neq 0$

La loi de probabilité des estimateurs \hat{a}_0 et \hat{a}_1 est déduite de celle du terme d'erreur.

Comme H5: $\varepsilon_t \rightarrow N(0, \sigma_\varepsilon^2)$

L'hypothèse de normalité des erreurs implique :

$$\frac{\hat{a}_i - a_i}{\sigma_{\hat{a}_i}} \rightarrow N(0,1)$$

et

on peut montrer que :

$$\frac{\hat{a}_i - a_i}{\hat{\sigma}_{\hat{a}_i}} \rightarrow T(n-2)$$

où $T(n-2)$ désigne une loi de Student à $n-2$ degrés de liberté

Sous l'hypothèse H_0 , $\frac{\hat{a}_i - a_i}{\hat{\sigma}_{\hat{a}_i}}$ devient $\frac{\hat{a}_i}{\hat{\sigma}_{\hat{a}_i}} = t^*$ appelé ratio de student ou « *t-stat* ».

La loi de student est symétrique et permet de déterminer directement la valeur critique de la statistique à partir de laquelle on rejette H_0 .

Cette valeur dépend du seuil de significativité α retenu.

Ce seuil renvoie au risque de rejeter H_0 alors que H_0 est vraie (risque de première espèce).

On retient généralement de façon arbitraire un seuil de 5% ou 1% voire 10%.

Pour un échantillon d'une centaine d'observations, nous pouvons lire dans la table de Student, les valeurs critiques suivantes :

α	valeur critique
1%	2,58
5%	1,96
10%	1,65

Si t^* est supérieur à sa valeur critique, cela signifie qu'on rejette H_0 au seuil α .

Exemple : si $t^* = 1,70$: on rejette H_0 au seuil de 10% mais pas au seuil de 5% ou 1%.

Les logiciels économétriques présentent quelquefois la *p-value* qui se lit comme la probabilité de faire une erreur en rejetant H_0 (erreur de première espèce).

Pour $0 \leq p\text{-value} < 1\%$, l'hypothèse nulle est rejetée au seuil de 1%.

Pour $1\% \leq p\text{-value} < 5\%$, l'hypothèse nulle est rejetée au seuil de 5%.

Pour $5\% \leq p\text{-value} < 10\%$, l'hypothèse nulle est rejetée au seuil de 10%.

Pour $10\% \leq p\text{-value}$, l'hypothèse nulle n'est pas rejetée au seuil de 10% (le coefficient n'est pas significativement différent de zéro).

NB : En décidant de rejeter ou non H_0 , on commet deux types d'erreurs :

1. rejeter H_0 alors qu'elle est vraie, c'est l'erreur de 1ère espèce ou l'erreur de type I ;
2. ne pas rejeter H_0 (ou rejeter H_1) alors qu'elle est fausse (ou H_1 est vraie), c'est l'erreur de 2ème espèce ou l'erreur de type II.

2. Le modèle linéaire multiple

Le modèle linéaire multiple se présente comme une extension du modèle linéaire simple, contenant plusieurs variables explicatives :

$$Y_t = \beta_0 + \beta_1 X_{1t} + \beta_2 X_{2t} + \dots + \beta_k X_{kt} + \varepsilon_t$$

où :

Y_t est la variable expliquée à la date t ;

$X_{1t}, X_{2t}, \dots, X_{kt}$ sont les k variables explicatives à la date t ;

$\beta_0, \beta_1, \beta_2, \dots, \beta_k$ sont $k+1$ paramètres du modèle ;

ε_t est le terme d'erreur.

Ce modèle peut se réécrire sous forme matricielle :

$$Y = X\beta + \varepsilon$$

où Y est un vecteur colonne de dimension T contenant les observations de la variable dépendante, X est une matrice ($T \times (k+1)$) de variables indépendantes, β est un vecteur de dimension $(k+1)$ de coefficients, et ε est un vecteur de dimension T des aléas. T est le nombre d'observations et k le nombre de variables explicatives ou indépendantes (constante exclue).

Notons que la première colonne de la matrice X est composée de 1, ce qui correspond au coefficient β_0 .

Les MCO

Les coefficients de régression par les MCO sont calculés par :

$$\hat{\beta} = (X'X)^{-1} X'Y$$

La matrice des covariances des coefficients estimés est définie par :

$$Var(\hat{\beta}) = \hat{\sigma}_\varepsilon^2 (X'X)^{-1}$$

$$\hat{\sigma}_\varepsilon^2 = \hat{\varepsilon}'\hat{\varepsilon} / (T - k - 1)$$

$$\hat{\varepsilon} = Y - \hat{Y} = Y - X\hat{\beta}$$

La t-statistique permet de tester l'hypothèse nulle (H_0) qu'un coefficient est égal à zéro, et est calculée par :

$$t-stat = \frac{\hat{\beta} - \beta}{\hat{\sigma}_\beta} = \frac{\hat{\beta}}{\hat{\sigma}_\beta} \sim \text{Student}(T - k - 1)$$

3. Diagnostic et tests

La statistique du R^2 (coefficient de détermination) est un indicateur de la qualité de l'ajustement du modèle. Il mesure le rapport entre la variance expliquée par le modèle et la variance totale. Il donne le pourcentage de la variance totale de la variable dépendante (Y) expliqué par les variables indépendantes (X).

$$R^2 = 1 - \frac{\hat{\varepsilon}'\hat{\varepsilon}}{(Y - \bar{Y})'(Y - \bar{Y})}$$

$$\bar{Y} = T^{-1} \sum_{t=1}^T Y_t$$

Dans un modèle de régression linéaire avec constante, si $R^2 = 1$ alors le modèle ajuste parfaitement ; si $R^2 = 0$ alors le modèle n'ajuste pas mieux qu'une simple moyenne de la variable dépendante.

NB: $R^2 < 0$ peut être dû à un modèle sans constante ou bien à un modèle comportant des restrictions de coefficients.

Un problème avec le R^2 est qu'il a tendance à augmenter lorsqu'on ajoute d'autres variables indépendantes. On calcule donc la statistique \bar{R}^2 (R^2 corrigé) qui pénalise le R^2 pour l'ajout de variables indépendantes qui ne contribuent pas à la puissance explicative du modèle ($\bar{R}^2 \leq R^2$):

$$\bar{R}^2 = 1 - (1 - R^2) \frac{T-1}{T-k}$$

L'écart type de la régression est une mesure basée sur la variance estimée des résidus

$$\hat{\sigma} = \sqrt{\frac{\hat{\varepsilon}'\hat{\varepsilon}}{T-k}}$$

La somme des carrés des résidus peut être utilisée dans divers calculs statistiques (i.e. test de Dickey-Fuller)

$$\hat{\varepsilon}'\hat{\varepsilon} = \sum_{t=1}^T (Y_t - X_t' \hat{\beta})^2$$

La log vraisemblance (en supposant que les erreurs sont normalement distribuées) peut être utilisée pour calculer des tests de rapport de vraisemblance (LR, *likelihood ratio*). Lors d'une sélection de modèle, cette valeur doit être maximisée.

$$\ell = -\frac{1}{2} (1 + \log(2\pi) + \log(\hat{\varepsilon}'\hat{\varepsilon}/T))$$

La statistique de Durbin-Watson permet de détecter une autocorrélation (corrélation sérielle) uniquement d'ordre 1 dans les résidus, i.e.

$$\varepsilon_t = \rho \varepsilon_{t-1} + u_t \text{ avec } u_t \sim N(0, \sigma_u^2)$$

Le test d'hypothèse est le suivant :

$$H_0 : \rho = 0$$

$$H_1 : \rho \neq 0$$

Si la statistique DW est proche de 2 alors on accepte l'hypothèse de non autocorrélation d'ordre 1.

$$DW = \frac{\sum_{t=2}^T (\hat{\varepsilon}_t - \hat{\varepsilon}_{t-1})^2}{\sum_{t=1}^T \hat{\varepsilon}_t^2}$$

Par construction, cette statistique varie entre 0 et 4. Il faut consulter les valeurs critiques du test de DW tabulées par les auteurs en fonction de la taille de l'échantillon n et du nombre de variables explicatives (k).

La lecture de la table permet de déterminer deux valeurs d_1 et d_2 comprises entre 0 et 2 qui délimitent l'espace entre 0 et 4 :

$0 < DW < d_1$: autocorrélation positive (rejet de H_0)

$d_1 < DW < d_2$ ou $4 - d_2 < DW < 4 - d_1$: impossible de conclure

$d_2 < DW < 4 - d_2$: non rejet de H_0

$4 - d_1 < DW < 4$: autocorrélation négative (rejet de H_0)

La moyenne et l'écart type de la variable expliquée :

$$\bar{Y} = T^{-1} \sum_{t=1}^T Y_t$$

$$\hat{\sigma}_Y = \sqrt{\frac{\sum_{t=1}^T (Y_t - \bar{Y})^2}{T - 1}}$$

Le critère d'information de Akaike (AIC) est utilisé pour la sélection de modèle et doit être minimiser (par exemple, pour le choix de la longueur de retard des variables)

$$AIC = -\frac{2\ell}{T} + \frac{2k}{T}$$

Le critère d'information de Schwarz (BIC ou SC) est une alternative au AIC en imposant une plus grande sanction pour l'ajout de coefficients :

$$BIC = -\frac{2\ell}{T} + \frac{k \log T}{T}$$

La F-statistique est un test sur l'hypothèse nulle H_0 que tous les coefficients (sauf la constante) du modèle sont nuls :

$$F = \frac{R^2 / k}{(1 - R^2) / (T - k - 1)}$$

Sous l'hypothèse nulle avec des erreurs normalement distribuées, on a :

$$F \sim F_{1-\alpha}(k; T - k - 1)$$

4. Tests sur les résidus

4.1 La normalité du résidu

Calculer la skewness et la kurtosis et pratiquer un tests de Jarque-Berra (voir *supra*).

4.2 Les tests d'autocorrélation

Ces tests sont souvent utilisés comme des tests pour savoir si les résidus suivent un bruit blanc (moyenne nulle, variance constante, non autocorrélé). En effet, les résidus doivent être non autocorrélés, i.e. leur covariance est nulle.

Le test de Durbin-Watson (DW) est basé sur les hypothèses suivantes :

- les variables explicatives ne sont pas stochastiques
- le terme d'erreur suit la distribution normale
- les modèles de régression ne comportent pas de valeurs retardées de la variable dépendante (expliquée)

Le test de DW teste l'absence d'autocorrélation d'ordre 1 uniquement. Pour ces raisons, les tests de Breusch-Godfrey et de Ljung-Box sont préconisés.

Le test de Breusch et Godfrey (1978) ou test du multiplicateur de Lagrange (LM) de corrélation sérielle est un test de non autocorrélation au sens large car il prend en compte :

- les valeurs retardées de la variable dépendante
- des autocorrélations d'ordre supérieur à 1
- la possibilité que les résidus soient autocorrélés

Soit le modèle de régression linéaire :

$$Y_t = \beta_0 + \beta_1 X_{1t} + \dots + \varepsilon_t$$

La statistique de test pour des retards d'ordre p est basée sur la régression intermédiaire pour les résidus :

$$\hat{\varepsilon}_t = \delta_0 + \delta_1 X_{1t} + \dots + \sum_{i=1}^p (\alpha_i \hat{\varepsilon}_{t-i}) + u_t$$

L'hypothèse nulle de non autocorrélation des résidus à l'ordre p testée est :

$$H_0 : \alpha_1 = \alpha_2 = \dots = \alpha_p = 0$$

La statistique de test LM de Breusch-Godfrey est donnée par :

$$(T - p)R^2 \sim \chi^2_{1-p}(p)$$

où R^2 est le coefficient de détermination de la régression intermédiaire.

Le test de Ljung et Box (1978) permet de tester l'hypothèse nulle qu'il n'y a pas d'autocorrélation jusqu'à l'ordre k . La statistique de test est :

$$Q(k) = T(T+2) \sum_{j=1}^k \frac{\tau_j^2}{T-j}$$

où τ_j est la j -ème autocorrélation des résidus du modèle, et T est le nombre d'observations.

Sous l'hypothèse nulle, la Q-statistique est asymptotiquement distribuée comme un χ^2 , i.e.

$$Q(k) \sim \chi^2_{1-p-q}(k-p-q)$$

où p et q sont les ordres d'un processus ARMA(p ; q).

Le test de Ljung-Box peut aussi être appliqué sur les résidus aux carrés et dans ce cas la Q-statistique permet de tester la présence d'effet ARCH (*Autoregressive conditional heteroscedasticity*) dans les résidus.

Correction de l'autocorrélation

En présence d'autocorrélation :

- application de la méthode des moindres carrés généralisés (MCG) lorsque le coefficient d'autocorrélation est connu ;
- application de la méthode des moindres carrés quasi-généralisés (MCQG) lorsque le coefficient d'autocorrélation est inconnu. Ce coefficient peut être estimé par la méthode itérative de Cochrane-Orcutt ou celle du « balayage » de Hildreth-Lu
- utilisation de la correction de Newey-West (1987) qui propose un estimateur convergent de la matrice des variances-covariances pour l'autocorrélation. On parle également de matrice des variances/covariances cohérents avec l'hétéroscédasticité et l'autocorrélation (HAC, *heteroskedasticity and autocorrelation consistent*).

4.3 Les test d'homoscédasticité

Les tests statistiques d'homoscédasticité portent sur l'hypothèse d'une variance des résidus identique à chaque individu (H_0).

Le test de White (1980) est un test d'homoscédasticité fondé sur l'existence d'une relation entre le carré du résidu et une ou plusieurs variables explicatives en niveau ou au carré.

Pour effectuer ce test, on utilise la statistique $LM\ TR^2$, où T est le nombre d'observation et R^2 est le coefficient de détermination de la régression précédente.

Sous l'hypothèse nulle d'homoscédasticité, on a :

$$TR^2 \sim \chi^2_{1-\alpha}(2p)$$

où p est le nombre de régresseurs excluant la constante.

Le test d'effet ARCH

Il s'agit d'examiner s'il existe des phénomènes de grappe ou de cluster sur la volatilité des résidus.

Pour cela, on régresse les résidus au carré sur une constante et les résidus au carré retardés à jusqu'à l'ordre q .

Pour effectuer ce test, on utilise la statistique $LM\ TR^2$, où T est le nombre d'observation et R^2 est le coefficient de détermination de la régression précédente.

Sous l'hypothèse nulle d'homoscédasticité, on a :

$$TR^2 \sim \chi^2_{1-\alpha}(k)$$

Correction de l'hétéroscédasticité

En présence d'hétéroscédasticité :

- si la variance du modèle est connue alors on utilise la méthode des moindres carrés pondérés ;
- si la variance du modèle est inconnue alors on utilise la correction de White (1980) qui propose un estimateur convergent de la matrice des variances-covariances, quelle que soit la forme d'hétéroscédasticité.

5. Tests sur les coefficients

Ces tests permettent d'évaluer des restrictions sur les coefficients estimés d'un modèle, avec notamment le test de Wald qui est basé sur l'estimateur du maximum de vraisemblance (EMV).

Soit le modèle de régression linéaire :

$$Y = X\beta + \varepsilon$$

Supposons que le modèle a été estimé par la méthode du maximum de vraisemblance et que l'on souhaite tester J contraintes linéaires :

$$H_0 : R\beta = q \Leftrightarrow R\beta - q = 0$$

$$H_1 : R\beta \neq q \Leftrightarrow R\beta - q \neq 0$$

où R est une matrice d'ordre $(J \times K)$ et q un vecteur composé de J lignes.

Le test de Wald a pour but de déterminer si l'écart entre X et q est significativement égal à 0. La statistique de ce test est définie de la manière suivante :

$$\text{Wald} = (R\hat{\beta} - q)'(\hat{\sigma}^2 R(X'X)^{-1}R')^{-1}(R\hat{\beta} - q)$$

Sous H_0 cette statistique suit la loi suivante :

$$\text{Wald} \sim \chi^2_{1-\alpha}(J)$$

Il existe une autre statistique définie par

$$F = \frac{\text{Wald}}{J} = \frac{(\tilde{u}'\tilde{u} - u'u)/J}{(u'u)/(T-k-1)} = \frac{(SCR_c - SCR_{nc})/J}{(SCR_{nc})/(T-k-1)}$$

Sous l'hypothèse de normalité des erreurs, on a

$$F = \frac{\text{Wald}}{J} \sim F_{1-\alpha}(J, T-k-1)$$

où \tilde{u} est le vecteur des résidus issus du modèle restreint ; u celui des résidus issus du modèle non restreint ; SCR_c est la somme des carrés des résidus du modèle contraint et SCR_{nc} est la somme des carrés des résidus du modèle non contraint.

IV. SÉRIES TEMPORELLES ET COINTÉGRATION

1. Les processus ARMA

Les processus ARMA (AutoRegressive Moving Average) ont été introduits par Box et Jenkins (1970). L'objet est de modéliser une série temporelle en fonction de ses valeurs passées, mais aussi en fonction des valeurs présente et passées d'un bruit. Afin de déterminer le processus ARMA adéquat, Box et Jenkins ont suggéré une procédure en quatre étapes :

- identification du modèle
- estimation des paramètres
- validation du modèle (test de diagnostic)
- Prévision à l'aide du modèle validé

1.1 Définition des processus ARMA

1.1.1 Les processus autorégressifs

On appelle un processus autorégressif d'ordre p , noté AR(p), un processus X_t stationnaire, vérifiant une relation du type :

$$X_t - \Phi_1 X_{t-1} - \dots - \Phi_p X_{t-p} = \varepsilon_t$$

où les Φ_p sont des réels et $\varepsilon_t \sim BB(0, \sigma_\varepsilon^2)$

Exemple : Le processus AR(1)

$$X_t = \Phi_1 X_{t-1} + \varepsilon_t$$

avec $|\Phi_p| < 1$ (condition de stationnarité) et $\varepsilon_t \sim BB(0, \sigma_\varepsilon^2)$

On distingue deux cas :

- Si $\Phi_1 > 0$ alors $\rho_k > 0$ (la fonction d'autocorrélation est positive) et décroît de façon exponentielle quand k augmente.
- Si $\Phi_1 < 0$ alors $\rho_k > 0$ alterne de signe et diminue de façon sinusoïdale.

Dans tous les cas :

- $\rho_0 = 1$
- $\Phi_1 = \rho_1 = \varphi_1$ (fonction d'autocorrélation partielle)
- $\varphi_{kk} = 0 \quad \forall k > 1$

1.1.2 Les processus moyenne mobile

On appelle un processus moyenne mobile d'ordre q , noté MA(q), un processus X_t stationnaire, vérifiant une relation du type :

$$X_t = \varepsilon_t - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q}$$

où les θ_q sont des réels et $\varepsilon_t \sim BB(0, \sigma_\varepsilon^2)$

Exemple : Le processus MA(1)

$$X_t = \varepsilon_t - \theta_1 \varepsilon_{t-1}$$

avec $|\theta_1| < 1$ (condition d'inversibilité) et $\varepsilon_t \sim BB(0, \sigma_\varepsilon^2)$

On observe que $\rho_k = 0 \quad \forall k > 1$ et que $\rho_0 = 1$

On distingue deux cas :

- si $\theta_1 > 0$, toutes les valeurs de ρ_{kk} sont négatives,
- si $\theta_1 < 0$, les autocorrélations partielles alternent de signe

1.1.3 Les processus ARMA

Ces processus constituent une extension naturelle des processus AR et MA. Ce sont des processus mixtes (à la fois AR et MA).

Un processus stationnaire

On appelle un processus autorégressif d'ordre p , noté ARMA(p, q), un processus X_t stationnaire, vérifiant une relation du type :

$$X_t - \Phi_1 X_{t-1} - \dots - \Phi_p X_{t-p} = \varepsilon_t - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q}$$

où les coefficients Φ_p et θ_q sont des réels et $\varepsilon_t \sim BB(0, \sigma_\varepsilon^2)$

On peut représenter les propriétés des fonctions d'autocorrélation et d'autocorrélation partielle de processus ARMA(p, q) dans le tableau suivant :

Processus	Fonction d'autocorrélation		Fonction d'autocorrélation partielle	
	$\Phi_1 > 0$	$\Phi_1 < 0$	$\theta_1 > 0$	$\theta_1 < 0$
AR(1)	diminution exponentielle	diminution sinusoïdale	$\varphi_{kk} = 0 \quad \forall k > 1$	
AR(p)	diminution expo. ou sinus. selon les coeff.		$\varphi_{kk} = 0 \quad \forall k > p$	
MA(1)	$\varphi_{kk} = 0 \quad \forall k > 1$		diminution exponentielle	diminution sinusoïdale
MA(q)	$\varphi_{kk} = 0 \quad \forall k > q$		diminution expo. ou sinus. selon les coeff.	
ARMA(1,1)	diminution exponentielle	diminution sinusoïdale	diminution exponentielle	diminution sinusoïdale
ARMA(p,q)	diminution expo. ou sinus. selon les coeff.		diminution expo. ou sinus. selon les coeff.	

1.2 Identification et estimation des processus ARMA

Les 4 étapes sont les suivantes :

- identification du modèle
- estimation des paramètres
- validation du modèle (test de diagnostic)
- Prévision à l'aide du modèle validé

1) Pour identifier le modèle : examen des fonctions d'autocorrélation et d'autocorrélation partielle

2) Estimation des processus et examen de la significativité des coefficients

3) validation du modèle : tests sur les résidus (Ljung-Box, White, etc)

4) Prévision (+ éventuelle sélection des modèles concurrents sur la base des *Mean Absolute Error*, *Mean Squared Error*, *Root Mean Squared Error*, critères d'information, etc.).

2. Non-stationnarité et tests de racine unitaire

De nombreuses des variables économiques et financières sont non-stationnaires. La non stationnarité peut concerner l'espérance comme les moments de second ordre.

Les cas de non stationnarité en moyenne sont analysés à partir de deux types de processus :

- Processus TS (*Trend Stationary*) qui représentent les processus caractérisés par une non stationnarité de type déterministe,
- Processus DS (*Difference Stationary*) qui représentent les processus dont la non stationnarité est de nature stochastique.

On distingue les processus TS des processus DS car suite à un choc, un processus TS revient à son niveau pre-choc, un processus DS n'y revient jamais.

La non stationnarité a des conséquences fondamentales sur le plan économétrique :

- propriétés asymptotiques usuelles des estimateurs ne sont plus valables
- régressions fallacieuses (régression statistiquement correctes entre des variables sans relation en réalité)

2.1 Définition générale

Une série aléatoire (dite encore *processus stochastique*), z_t , est *stationnaire* au sens faible ou de second ordre si :

- $E(y_t) = \mu \forall t$ (moyenne constante qui ne dépend pas du temps)
- $\text{Var}(y_t) = \sigma^2 < \infty \forall t$ (variance finie et qui ne dépend pas du temps)
- $\text{Corr}(y_t, y_{t+k}) = \gamma_k$ (son autocorrélation ne dépend pas du temps mais de l'écart entre les observations)

Le processus TS (*Trend Stationary*) s'écrit :

$$y_t = \alpha + \beta t + \varepsilon_t$$

où $\varepsilon_t \sim BB(0, \sigma_\varepsilon^2)$.

Il présente une non stationnarité de nature déterministe. Le processus TS est non stationnaire car $E(y_t) = \alpha + \beta t$ dépend du temps t .

Le processus y_t peut être stationnarisé en retranchant à y_t la valeur estimée par la méthode des Moindres Carrés Ordinaires de $\hat{\beta}t$.

Le processus DS (*Differency Stationary*) s'exprime comme suit :

$$y_t = y_{t-1} + \beta + \varepsilon_t$$

Ce processus est connu sous le nom de marche au hasard (ou marche aléatoire) avec dérive lorsque $\beta \neq 0$ et sans dérive lorsque $\beta = 0$.

Ce processus présente une non stationnarité de nature stochastique.

Par récurrence, on obtient (dans le cas avec dérive) :

$$y_1 = y_0 + \beta + \varepsilon_1$$

$$y_2 = y_0 + 2\beta + \varepsilon_1 + \varepsilon_2$$

...

$$y_t = y_0 + t\beta + \sum_{j=1}^t \varepsilon_j$$

Le processus DS avec dérive est non stationnaire car $E(y_t) = \alpha + \beta t$. Lorsque $t \rightarrow \infty$, $E(y_t) \rightarrow \infty$.

Le processus DS sans dérive est non stationnaire car la variance du processus DS sans dérive dépend du temps t :

$$\text{Var}(y_t) = \text{Var}\left(\sum_{j=1}^t \varepsilon_j\right) = \sum_{j=1}^t \text{Var}(\varepsilon_j) = t\sigma_\varepsilon^2$$

Pour stationnariser le processus DS (avec ou sans dérive), il suffit de le passer en différence première.

Ainsi, une série est dite intégrée d'ordre d ($y_t \sim I(d)$) s'il convient de la différencier d fois afin de la stationnariser. La série stationnarisée est alors intégrée d'ordre 0.

2.2 Les tests de stationnarité

Il existe plusieurs tests de racine unitaire : tests de Dickey-Fuller simple et Dickey-Fuller Augmenté, test de Phillips et Perron, test de Kwiatkowski, Phillips, Schmidt et Shin (test de KPSS). EvIEWS propose par défaut les tests de DF augmenté et de Phillips-Perron ainsi que d'autres test pour les versions les plus récentes d'EvIEWS. Stata après installation des procédures spécifiques permet de mener les tests ADF et KPSS.

Les tests DF et ADF

Le test de Dickey-Fuller permet de savoir si une série est stationnaire ou non et permet aussi de déterminer la bonne manière de stationnariser la série.

Les hypothèses du test sont les suivantes :

H_0 : processus non stationnaire, il correspond à une de ces formes de non stationnarité :

$$(1) X_t = \Phi_1 X_{t-1} + \varepsilon_t$$

$$(2) X_t = \Phi_1 X_{t-1} + \alpha + \varepsilon_t$$

$$(3) X_t = \Phi_1 X_{t-1} + \alpha + \beta t + \varepsilon_t$$

avec $\Phi_1 = 1$ et $\varepsilon_t \sim BB(0, \sigma_\varepsilon^2)$.

$$H_1 : \Phi_1 < 1$$

On peut écrire aussi les hypothèses sous la forme suivante :

H_0 : processus non stationnaire, il correspond à une de ces formes de non stationnarité :

$$(1) \Delta X_t = (\Phi_1 - 1) X_{t-1} + \varepsilon_t$$

$$(2) \Delta X_t = (\Phi_1 - 1) X_{t-1} + \alpha + \varepsilon_t$$

$$(3) \Delta X_t = (\Phi_1 - 1) X_{t-1} + \alpha + \beta t + \varepsilon_t$$

avec $\Phi_1 - 1 = 0$ et $\varepsilon_t \sim BB(0, \sigma_\varepsilon^2)$.

$$H_1 : \Phi_1 - 1 < 0.$$

On commence par étudier le modèle général (3). On regarde si β est significativement différent de 0 ou non. Si β est significativement non différent de 0, on passe à l'étude du modèle (2) et on cherche à savoir si α est significativement différent de 0 ou pas. Si α est significativement non différent de 0, on étudie le modèle (1).

La règle est la suivante :

Si la valeur absolue de la statistique du test est supérieure à sa valeur critique (en valeur absolue) à un certain seuil : on rejette H_0 (non-stationnarité) à ce seuil.

Le test de Dickey-Fuller suppose que $\varepsilon_t \sim BB(0, \sigma_\varepsilon^2)$. Or il n'y a aucune raison pour que, à priori, l'erreur soit non corrélée. Le tests Dickey Fuller Augmenté (ADF) consiste à ajouter des termes retardés des différences premières de variable dépendante du côté droit de la régression.

Par exemple pour le modèle (3) :

$$\Delta X_t = (\Phi_1 - 1) X_{t-1} - \sum_{k=2}^p \Delta X_{t-k+1} + \alpha + \beta t + \varepsilon_t$$

Il existe plusieurs façons de déterminer l'ordre du retard :

- utilisation du corrélogramme partiel de la série différenciée
- méthode du général au spécifique (on commence avec un grand nombre de retards et on les diminue en s'arrêtant dès que le dernier retard de la variable différenciée est significatif)
- utilisation des critères d'information.

La règle de décision est ensuite la même que pour le test de Dickey-Fuller simple.

Le test de Phillips-Perron

Le test de Phillips et Perron permet de prendre en compte à la fois l'autocorrélation et l'hétéroscédasticité des erreurs. Il s'appuie sur les mêmes modèles que ceux du test de Dickey et Fuller simple mais propose une correction non-paramétrique de la statistique.

Notez que certains tests comme celui de KPSS posent la stationnarité sous l'hypothèse nulle. Le non-rejet de H_0 signifie contrairement aux tests ADF et PP que l'on valide l'hypothèse de stationnarité.

3. La cointégration

Le point de départ de la théorie de la cointégration réside dans le fait que la plupart de séries macroéconomiques et financières sont non stationnaires.

Or si on applique les méthodes habituelles, deux problèmes se posent :

- régression fallacieuses (tests usuels de significativité biaisés)
- certaines lois asymptotiques ne sont plus valables

Aussi, les modèles ARMA ne sont valables que pour de séries stationnaires.

La solution pourrait être de différencier les séries afin de les rendre stationnaires mais cela fait perdre de l'information sur les relations à long terme entre les variables.

3.1 Définition de la cointégration

Si X_t et Y_t sont deux séries $I(d)$ alors en général la combinaison linéaire, z_t :

$$z_t = X_t - aY_t$$

Est aussi $I(d)$.

Cependant il est possible que z_t ne soit pas $I(d)$ mais $I(d - b)$ où b est un entier positif. Dans ce cas X_t et Y_t sont dits cointégrées. a est le paramètre de cointégration et le vecteur $[1, -a]$ est le vecteur de cointégration.

Le cas le plus étudié correspond à : $d = b = 1$.

Ainsi deux séries $I(1)$ sont cointégrées si leur combinaison linéaire est stationnaire.

Intuition :

X_t et Y_t peuvent diverger à court terme mais une force de rappel les fait évoluer ensemble à long terme.

Si cette relation existe, elle est appelée relation de long terme ou relation de cointégration.

z_t mesure l'ampleur de ce déséquilibre entre les deux séries et est appelé « erreur d'équilibre »

Exemples : Consommation-Revenu, taux d'intérêt court et long ; indices boursiers, salaires et prix.

3.2 Représentation sous forme de modèle à correction d'erreurs

Les modèles à correction d'erreur sont des modèles dynamiques qui intègrent à la fois les évolution de court terme et les évolution de long terme.

Soient X_t et Y_t , deux variables cointégrées. Le modèle à correction d'erreurs s'écrit comme suit :

$$\begin{cases} \Delta X_t = \gamma_1 z_{t-1} + \sum_i \beta_i \Delta X_{t-i} + \sum_i \delta_i \Delta Y_{t-i} + \varepsilon_{Xt} \\ \Delta Y_t = \gamma_2 z_{t-1} + \sum_i \beta'_i \Delta X_{t-i} + \sum_i \delta'_i \Delta Y_{t-i} + \varepsilon_{Yt} \end{cases}$$

où ε_{Xt} et ε_{Yt} sont deux bruit blancs ; $z_t = X_t - aY_t$ est le résidu de la relation de long terme entre X_t et Y_t .

La différence par rapport à un modèle VAR classique réside dans la présence du terme à correction d'erreur.

Les coefficients γ_1 et γ_2 représentent les « forces de rappel » vers la cible de long terme. On doit avoir $\gamma_1 < 0$ sinon il n'existe pas de phénomène de retour à l'équilibre.

3.3 Estimation des modèles à correction d'erreur et tests de cointégration

Deux méthodes très courantes :

- la méthode en deux étapes de Engle et Granger (1987)
- l'approche multivariée de Johansen

a) La méthode de Engle-Granger

Méthode simple mais valable que pour deux séries.

1^{ère} étape : Estimation de la relation de long terme (MCO)

$$Y_t = \alpha + \beta X_t + z_t$$

z_t doit être stationnaire pour conclure à la relation de cointégration :

! valeurs critiques différentes des tests de stationnarité classiques ! \Rightarrow test de cointégration

2^{ème} étape : Estimation du modèle à correction d'erreur (MCO)

$$\Delta Y_t = \gamma \hat{z}_{t-1} + \sum_i a_i \Delta X_{t-i} + \sum_j b_j \Delta Y_{t-j} + \varepsilon_{Yt}$$

b) L'approche de Johansen

Considérons un processus autorégressif de dimension p avec des erreurs gaussiennes :

$$z_t = A_1 z_{t-1} + \dots + A_k z_{t-k} + Bx_t + \mu + \varepsilon_t, \quad t = 1, \dots, T$$

où z_t est un vecteur de variables aléatoires $I(1)$ de dimension $p \times 1$; x_t est un vecteur de variables $I(0)$ exogènes de dimension $s \times 1$ et $\varepsilon \sim (0, \Sigma)$.

Ce modèle peut être réécrit sous la forme d'un modèle à correction d'erreur (MVCE) :

$$\Delta Z_t = \Pi Z_{t-1} + \sum_{i=1}^{k-1} \Gamma_i \Delta Z_{t-i} + Bx_t + \mu + \varepsilon_t, \quad (2)$$

où Z_t est un vecteur contenant N variables $I(1)$; x_t est un vecteur de s variables exogènes stationnaires et $\varepsilon \sim (0, \Sigma)$. Les matrices Π et Γ_i sont de taille $(N \times N)$.

S'il y a r relations de cointégration dans le système, la matrice Π est de rang r . Elle peut alors s'écrire $\Pi = \alpha\beta'$, où α et β sont deux matrices de dimensions $(N \times r)$. Ces deux matrices représentent respectivement les facteurs des termes de correction d'erreur et les combinaisons

cointégrantes. Les différentes matrices sont estimées par la méthode du maximum de vraisemblance.

Le nombre de vecteurs de cointégration r est déterminé à l'aide de deux tests :

- le test de la Trace
- le test de la valeur propre maximale.