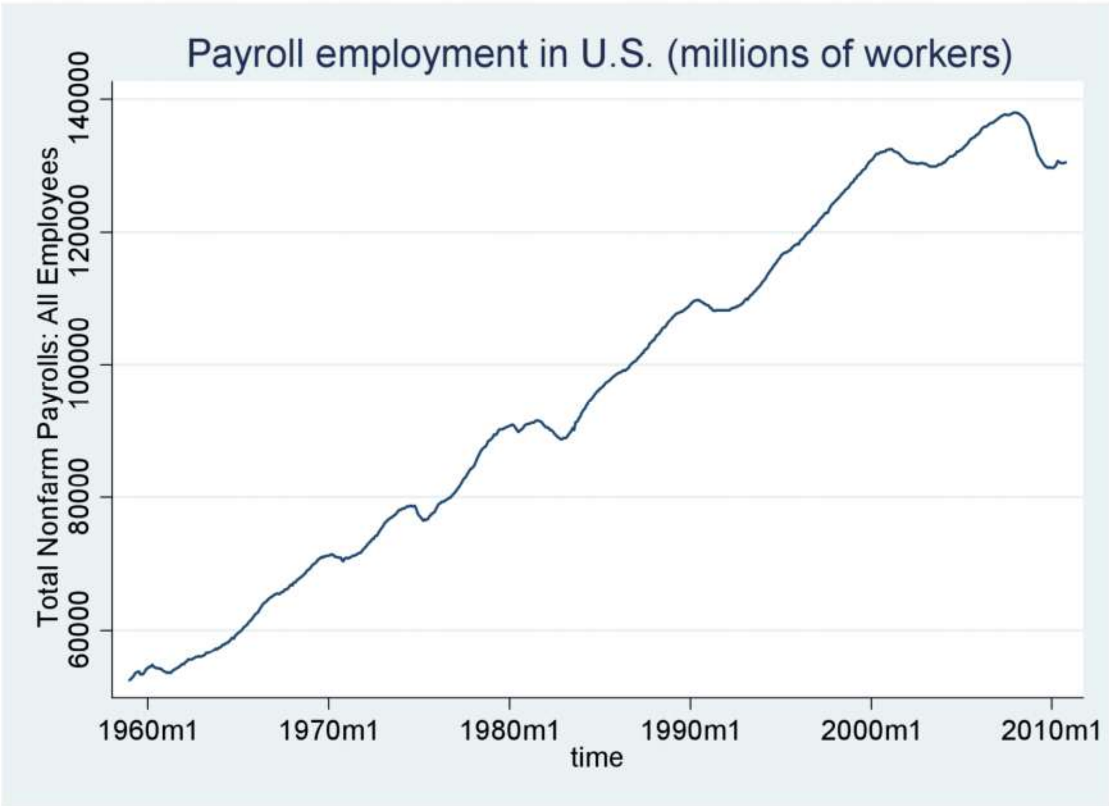
Séries temporelles 1

Patrick Waelbroeck, ENST
waelbroe@enst.fr

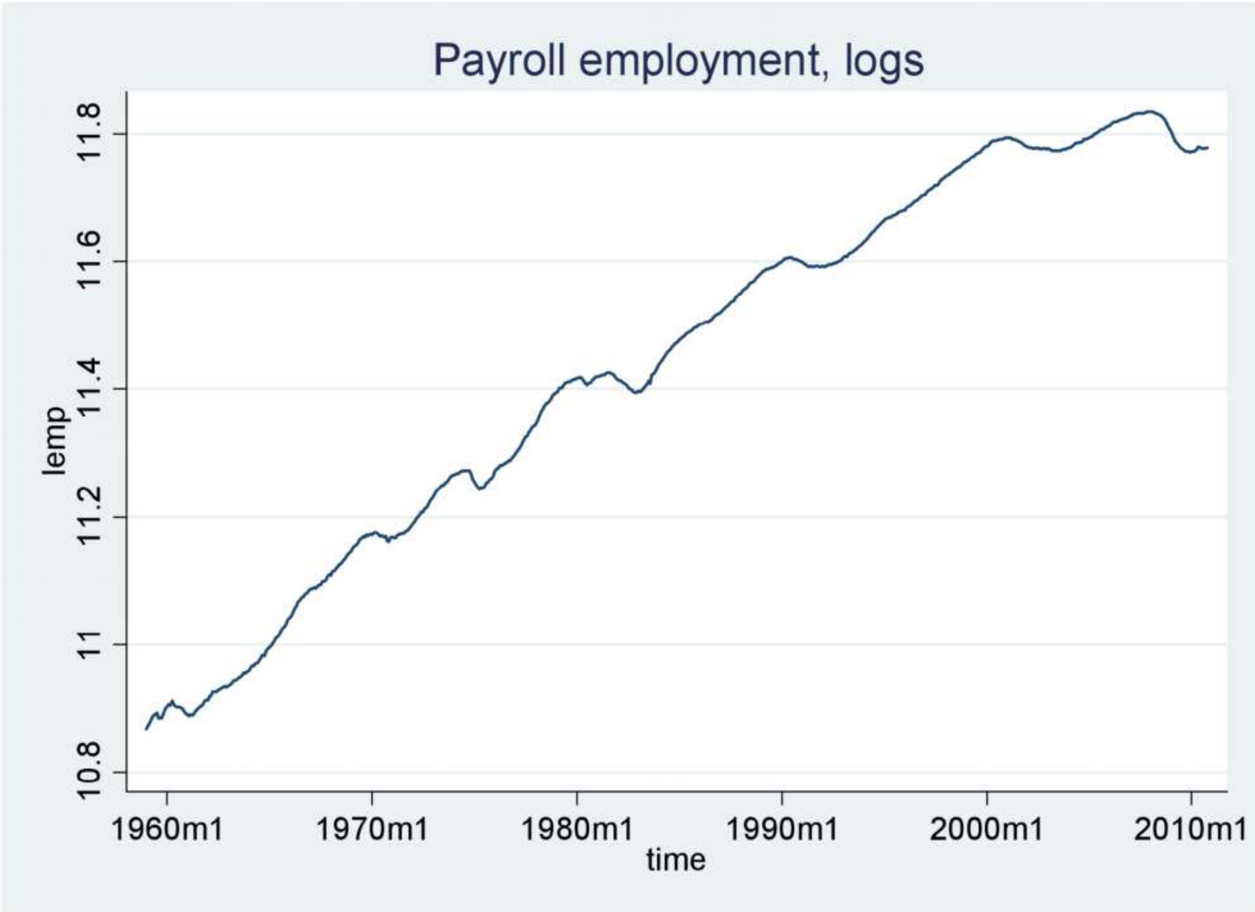
Introduction

- Les séries temporelles sont collectées pour la même unité observée sur plusieurs périodes
  - Consommation agrégée, PNB pour un pays
  - Taux de change Yen/$, EUR/$
  - Consommation de cigarettes par habitant pour un état, par an
- L'ordre est important

Exemples

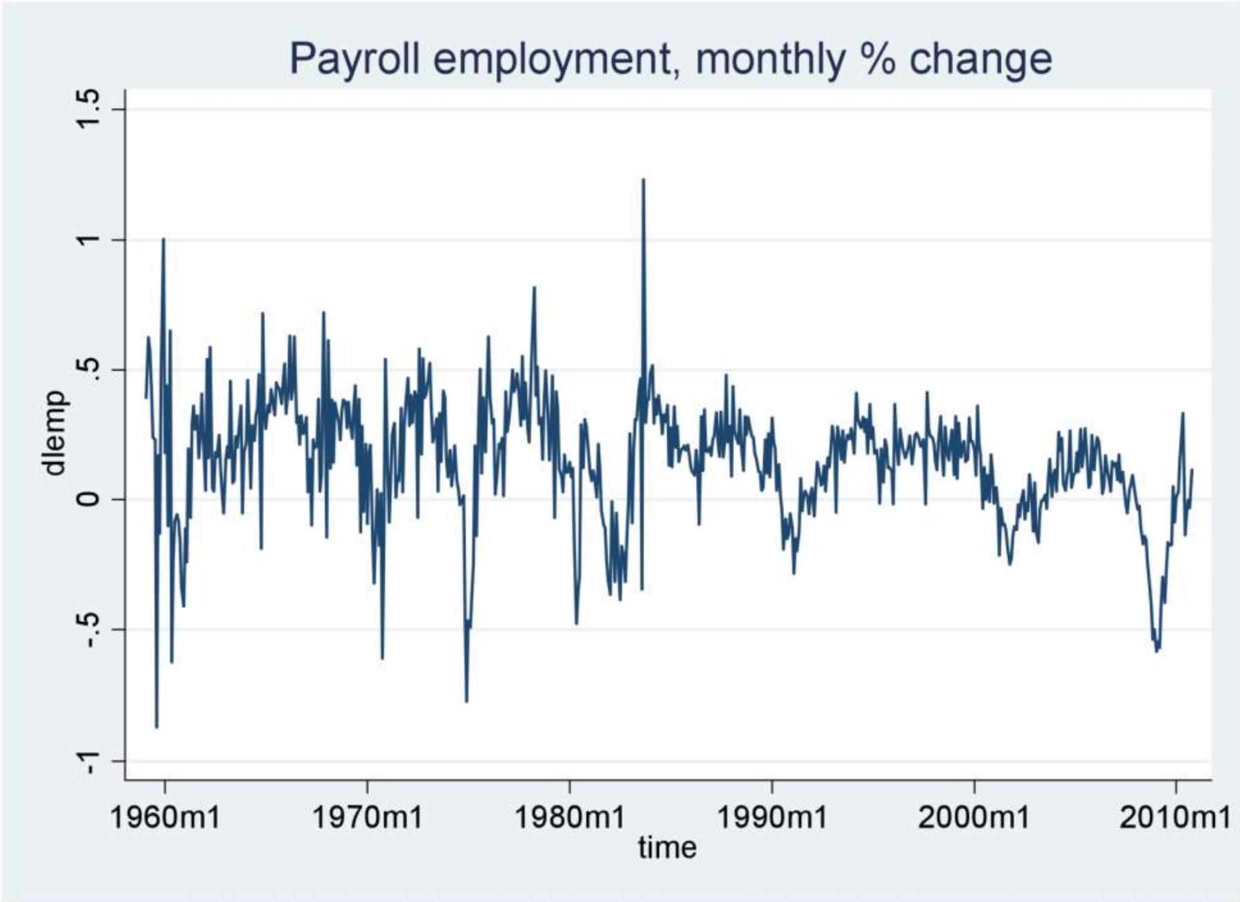

Payroll employment in U.S. (millions of workers)
Total Nonfarm Payrolls: All Employees
60000
80000
100000
120000
140000
1960m1
1970m1
1980m1
1990m1
2000m1
2010m1
time

Exemples

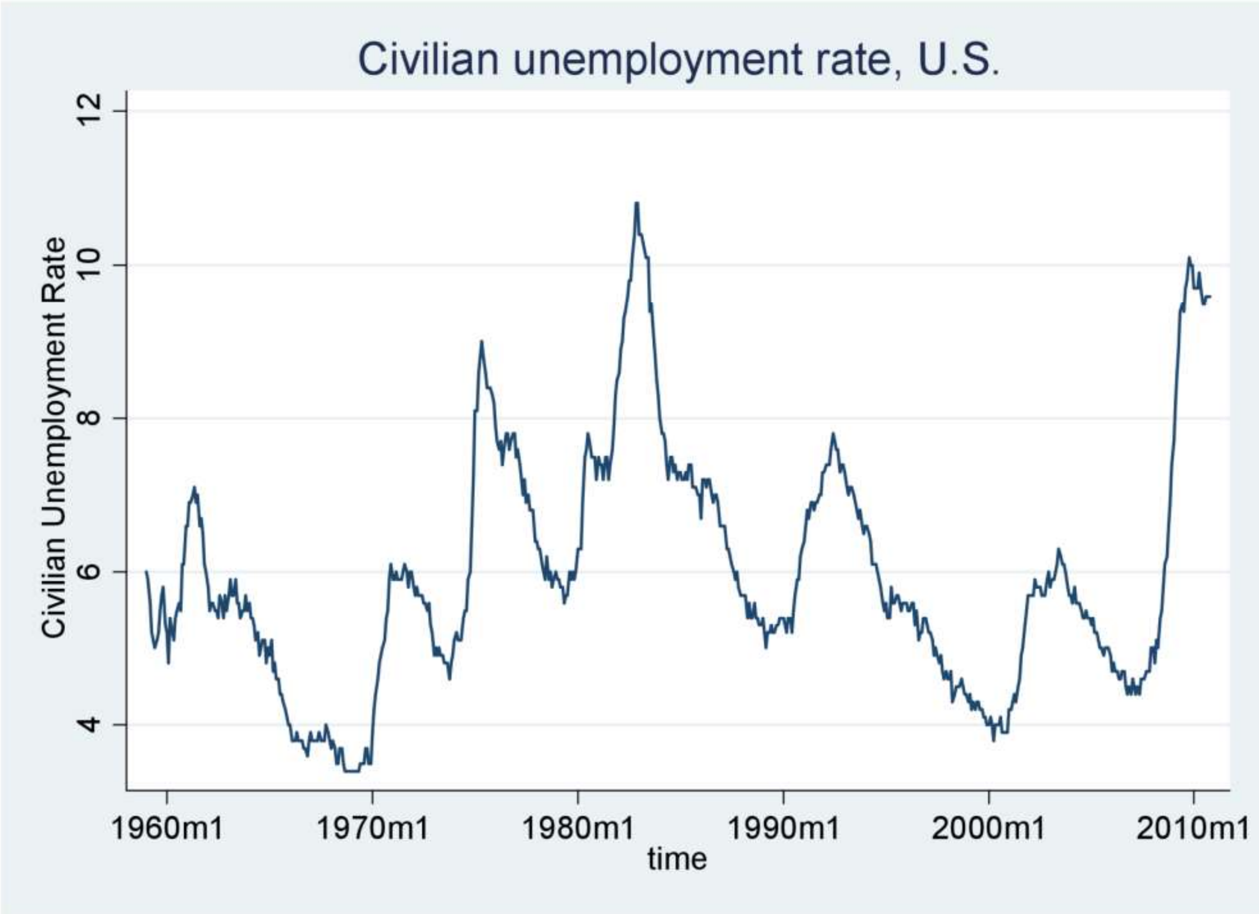


Payroll employment, logs
lemp
10.8
11
11.2
11.4
11.6
11.8
1960m1
1970m1
1980m1
1990m1
2000m1
2010m1
time

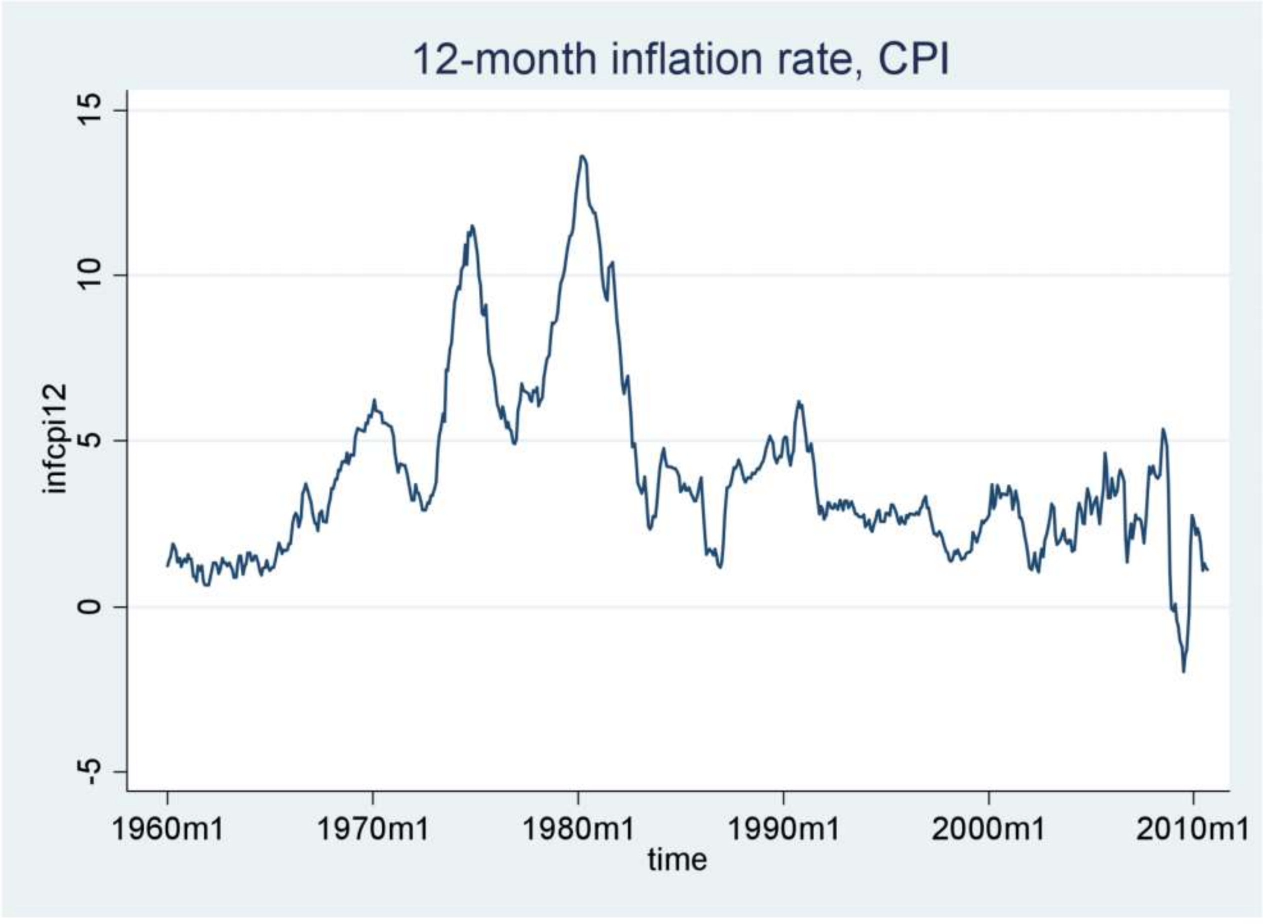
Exemples


Payroll employment, monthly % change
dlemp
1.5
1
.5
0
-.5
-1
1960m1
1970m1
1980m1
1990m1
2000m1
2010m1
time

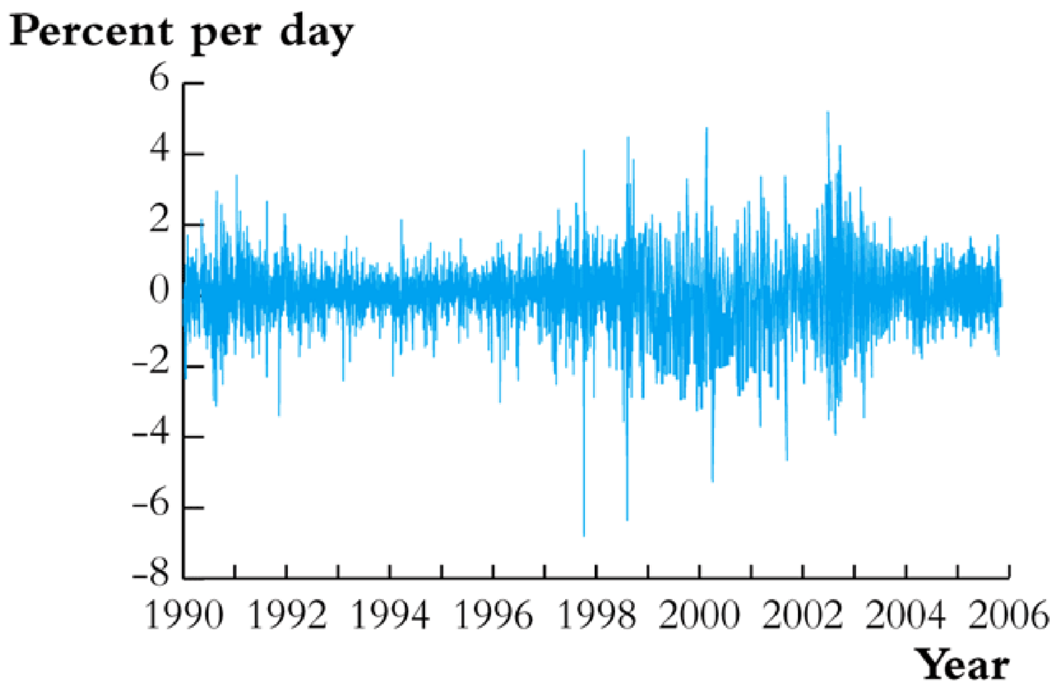
Exemples


Civilian unemployment rate, U.S.
Civilian Unemployment Rate
12
10
8
6
4
1960m1
1970m1
1980m1
1990m1
2000m1
2010m1
time

Exemples


12-month inflation rate, CPI
infcpi12
15
10
5
0
-5
1960m1
1970m1
1980m1
1990m1
2000m1
2010m1
time

Pourquoi utiliser les séries temporelles ?



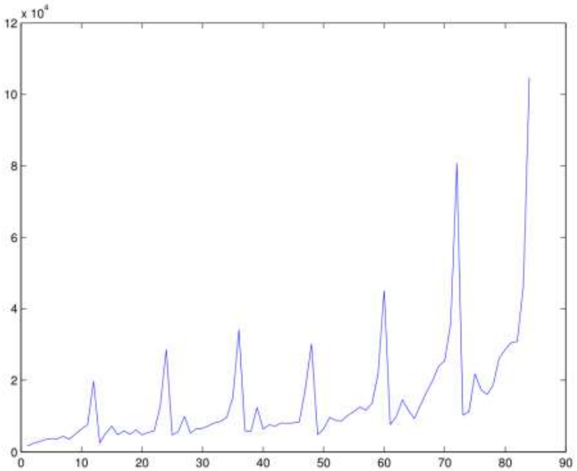
(d) Percentage Changes in Daily Values of the NYSE Composite Stock Index

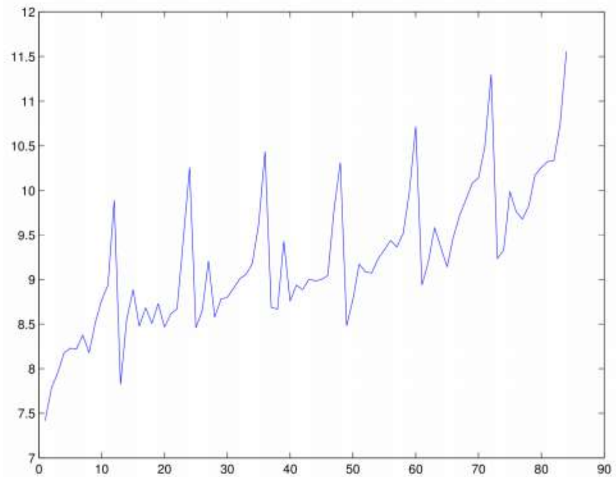
Pourquoi utiliser les séries temporelles ?

- Faire des prévisions

Monthly sales for a souvenir shop at a beach resort town in Queensland.

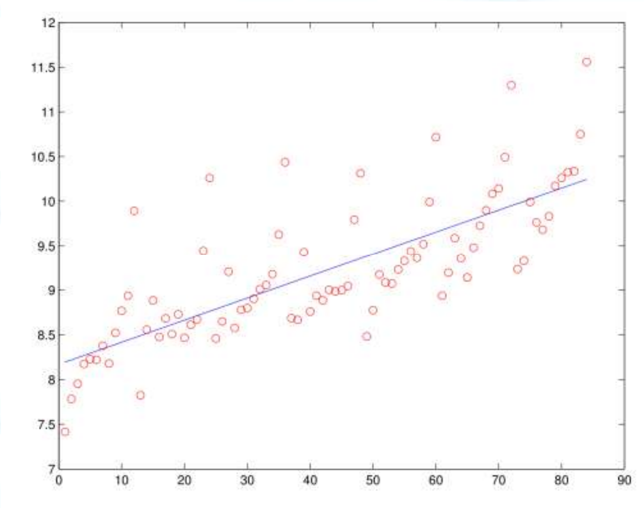
(Makridakis, Wheelwright and Hyndman, 1998)

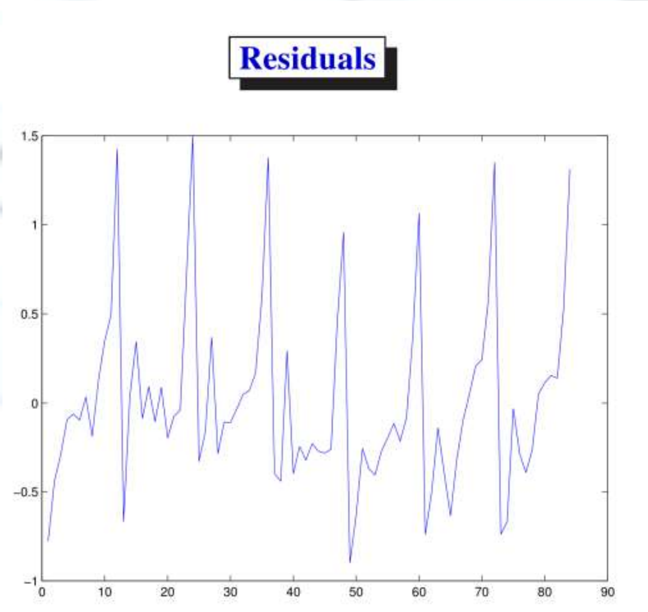


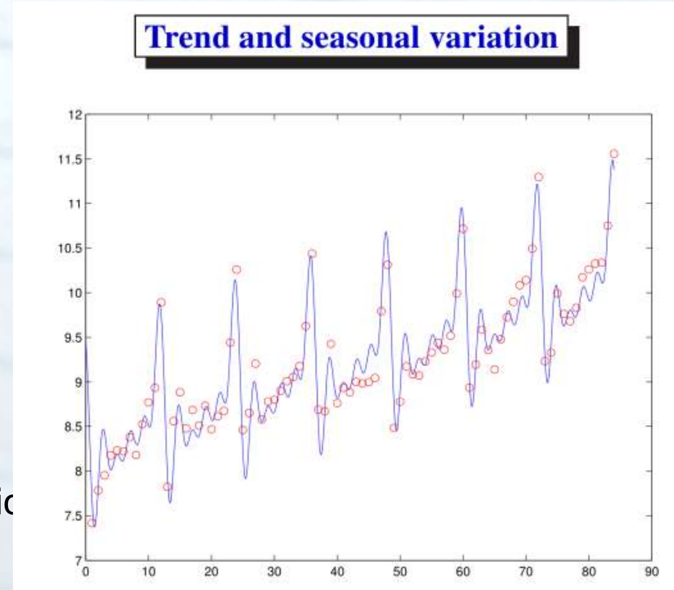


Pourquoi utiliser les séries temporelles ?

- Faire des prévisions







Pourquoi utiliser les séries temporelles ?

- Faire des prévisions
- Objectif : faire une décomposition

$$X_t = T_t + S_t + Y_t.$$

Pourquoi utiliser les séries temporelles ?

- Faire des prévisions
- Débat forme réduite/forme structurelle
  - La critique de Lucas (1976)
    - Les comportement des agents économiques sont optimaux par rapport aux paramètres du modèle, si l'on change ces paramètres, par exemple à travers une politique économique, alors les comportements changent également; nécessité d'avoir un modèle structurel.

Pourquoi utiliser les séries temporelles ?

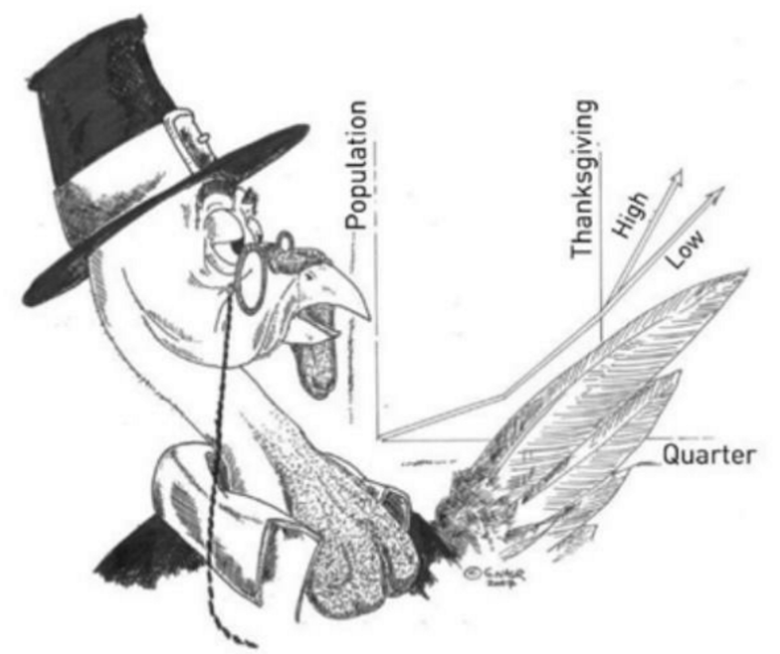


FIGURE 4. A turkey using "evidence"; unaware of Thanksgiving, it is making "rigorous" future projections based on the past. Credit: George Nasr

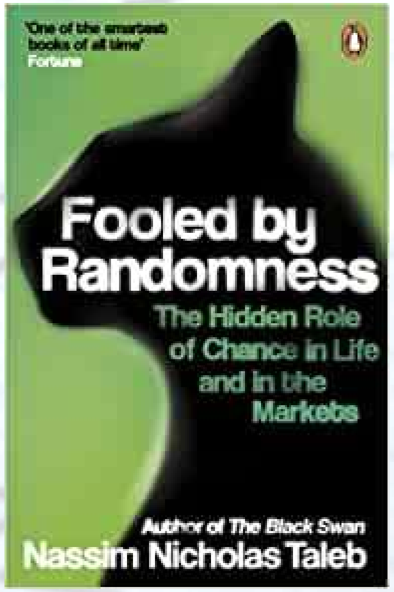
Pourquoi utiliser les séries temporelles ?

- Faire des prévisions
- Rôle des anticipations
  - L'équivalence Ricardienne revisitée par Robert Barro (1974)
    - Une politique de relance par la demande (Keynésienne) n'a pas d'effet si les agents anticipent les hausses futures d'impôt et épargnent en conséquence.
  - Permanent income hypothesis < Friedman (1957)
    - Les gens consomment en fonction de leur revenu permanent qui inclut des anticipations de revenus futurs et non pas en fonction de leur revenu courant (= propension marginale à consommer de Keynes)
    - Exemple : Prix de l'immobilier

Pourquoi utiliser les séries temporelles ?

- Faire des prévisions
- La théorie des marchés efficients
  - Samuelson (1965) et Fama (1963-1970)
    - Les prix des marché financiers reflètent toute l'information des investisseurs; on ne peut pas faire des prévisions profitables sur le long terme; les prix fluctuent de manière aléatoire.
  - Différentes formes d'efficience (faible, semi-forte, forte)
  - Deux prix Nobel (1970, 2013), mais résultat très débattu < Robert Shiller (Prix Nobel, 2013)
  - Enjeux pour le trading algorithmique

Pourquoi utiliser les séries temporelles ?



Pourquoi utiliser les séries temporelles ?

- Estimer des effets (causaux) dynamiques
  - Si la BCE augmente les taux directeurs, de combien augmenteront l'inflation et le chomage dans 3 mois ? 1 an ?
  - Quel est l'effet à long terme sur la consommation de cigarette d'une taxe sur les cigarettes ?

Nouveaux enjeux liés aux séries temporelles

- Délais
- Corrélation dans le temps (autocorrélation)
- Méthodes de prévision basés sur des modèles de régression
  - AR
  - Délais distribués
- Stationnarité, dépendance, ergodicité

Les différents approches

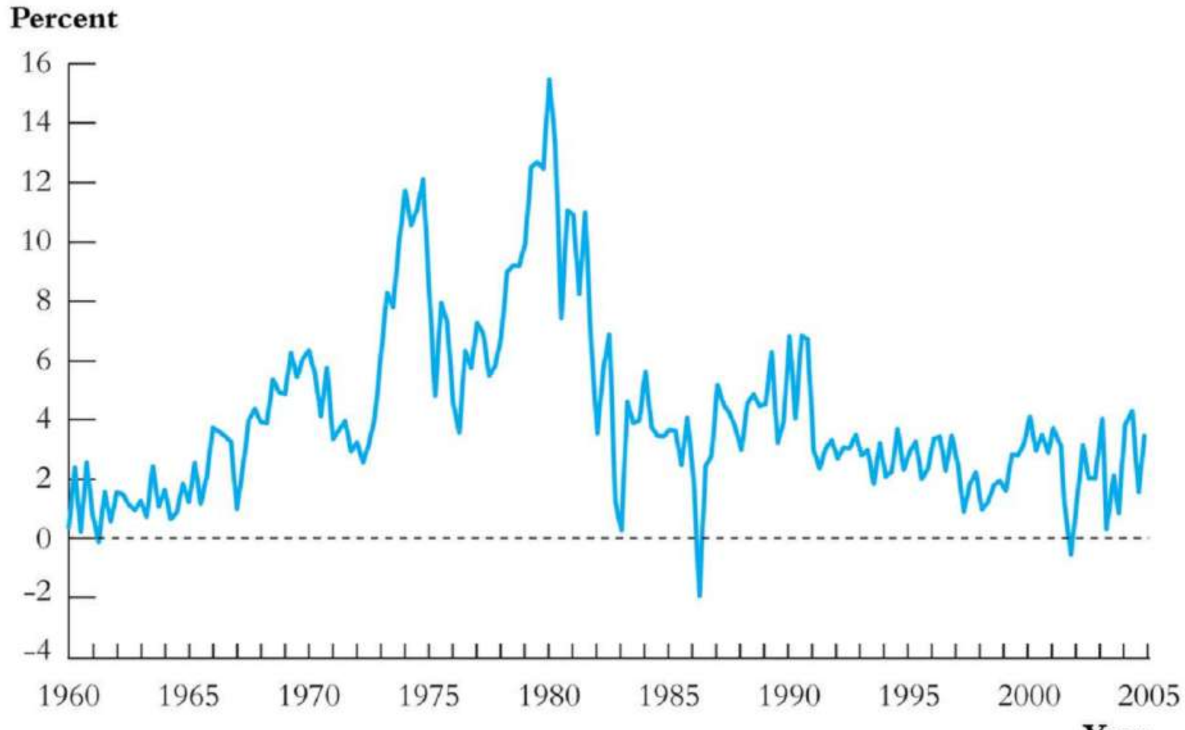
- Domaine temporel
  - Box-Jenkins (ARMA)
  - Vecteur Autorégression (VAR) : plusieures séries, Chris Sims (Prix Nobel, 2011)
- Domaine des fréquences
  - Analyse spectrale : Clive Granger (Prix Nobel, 2003)
- Modèles avec variables explicatives
  - Autocorrélation
  - Structural VAR

Notations

- y_t = valeur de y à la période t
- Données = $y_1, \ldots, y_T$
- Observations régulièrement espacées, pas d'observations manquantes
- Remarque : économétrie spatiale
- Le délai d'ordre 1 est y_{t-1}, celui d'ordre j, y_{t-j}
- La différence première $\Delta y_t = y_t - y_{t-1}$
- La différence première en log est $\Delta \ln(y_t) = \ln(y_t) - \ln(y_{t-1})$

Exemple

- CPI inflation rate



(a) U.S. CPI Inflation Rate

Exemple

- Si le CPI passe de 186.57 à 188.6d'un trimestre à un autre, le taux d'inflation trimestriel est de 1.088 %
- Le taux annuel est égal à 4 fois cette valeur, soit 4.4%

Exemple : US CPI

TABLE 14.1 Inflation in the United States in 2004 and the First Quarter of 2005

Quarter	U.S. CPI	Rate of Inflation at an Annual Rate (Inf_t)	First Lag (Inf_{t-1})	Change in Inflation (ΔInf_t)
2004:I	186.57	3.8	0.9	2.9
2004:II	188.60	4.4	3.8	0.6
2004:III	189.37	1.6	4.4	−2.8
2004:IV	191.03	3.5	1.6	1.9
2005:I	192.17	2.4	3.5	−1.1

The annualized rate of inflation is the percentage change in the CPI from the previous quarter to the current quarter, times four. The first lag of inflation is its value in the previous quarter, and the change in inflation is the current inflation rate minus its first lag. All entries are rounded to the nearest decimal.

Auto-corrélation

- L'auto-corrélation d'ordre 1 est corr(y_t, y_{t-1})
- L'auto-covariance d'ordre 1 est cov(y_t, y_{t-1})
- On utilise la formule

$$\text{corr}(Y_t, Y_{t-1}) = \frac{\text{cov}(Y_t, Y_{t-1})}{\sqrt{\text{var}(Y_t)\,\text{var}(Y_{t-1})}} = \rho_1$$

- L'auto-corrélation et l'auto-covariance d'ordre j sont définies de la même manière :
- Corr(y_t, y_{t-j}), cov(y_t, y_{t-j})

Auto-corrélation empirique

- L'auto-corrélatoin empirique (sample auto-correlation) : $r_j = cov^e(y_t, y_{t-j})/var^e(y_t)$
- $Cov^e(y_t, y_{t-j}) =$

$$\frac{1}{T}\sum_{t=j+1}^{T}(Y_t - \overline{Y}_{j+1,T})(Y_{t-j} - \overline{Y}_{1,T-j})$$

Exemple d’auto-corrélation

TABLE 14.2 First Four Sample Autocorrelations of the U.S. Inflation Rate and Its Change, 1960:I–2004:IV

	Autocorrelation of:	
Lag	Inflation Rate (Inf_t)	Change of Inflation Rate (ΔInf_t)
1	0.84	−0.26
2	0.76	−0.25
3	0.76	0.29
4	0.67	−0.06

14-14

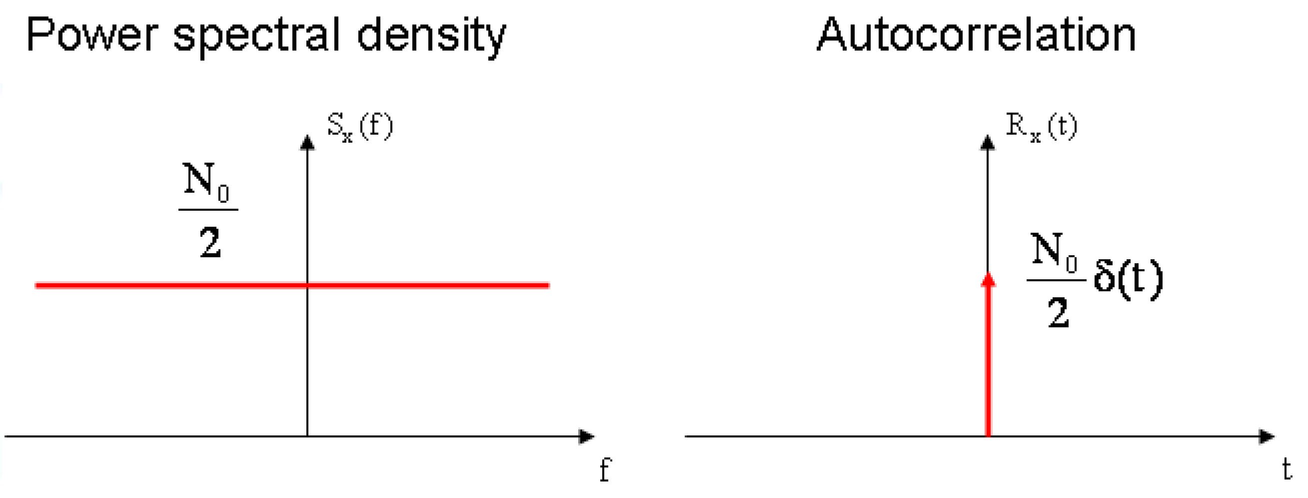
Analyse dans le domaine des fréquences

- Il existe une dualité entre l'analyse dans le domaine du temps et l'analyse dans le domaine des fréquences
- On peut représenter une série par un superposition de sinus et de cosinus. On cherche la périodicité des cycles = fréquence.
- On utilise la transformée de Fourier des autocorrélations pour estimer les fréquences des cycles

Analyse dans le domaine des fréquences

- Exemple : Bruit blanc (white noise) // lumière blanche : toutes les fréquences contribuent de manière égale à la variance (énergie)
- Problèmes de l’approche :
  - Fonctionne sur des séries univariées (bi-variées)
  - Ne permet pas d’ajouter de variables explicatives

Bruit blanc



Power spectral density
Autocorrelation
$S_x(f)$
$\frac{N_0}{2}$
f
$R_x(t)$
$\frac{N_0}{2}\delta(t)$
t

La notion de stationnarité et d'ergodicité

- Stationnarité au strict = La distribution jointe de $(y_{s+1}, y_{s+2}, \ldots, y_{s+T})$ ne dépend pas de s; le futur et le passé se ressemblent
- Ergocité = le processus oublie les conditions initiales : l'auto-corrélation d'ordre k tend vers 0 lorsque k tend vers l'infini
- Théorème ergodique : si {yt} est strictement stationnaire et ergodique et E(yT) < ∞ quand T → ∞, alors la moyenne temporelle converge vers la moyenne spatiale (l'espérance)

Stationnaire mais pas ergodique

Example 3 *Stationary but not ergodic process (White, 1984)*

Let $\{Y_t\}$ be an iid sequence with $E[Y_t] = \mu$ and let $X \sim N(0,1)$ independent of $\{Y_t\}$. Let $Z_t = Y_t + X$. Note that $E[Z_t] = \mu$.

Ergodicité : Chaîne de Markov discrète

- Est caractérisée par les probabilité de transition d'un état *i* à un état *j*.
- Trois conditions pour l'existence de la distribution invariante (ergodique)
  - Absence d'états absorbants
  - Absence de cycles
  - tout état peut être visité à partir de n'importe quel autre état
  - Résultat : la châine est positive récurrente et admet une distribution stationnaire invariante, on peut appliquer le théorème ergodique
  - Remarque : une chaine de Markov avec des états absorbants est stationnaire mais pas ergodique.

Exogénéité

- Considérons le modèle

$$y_t = \beta_0 + \beta_1 x_{t1} + \ldots + \beta_k x_{tk} + u_t,$$

- Stricte exogénéité:

$$\mathrm{E}(u_t|X) = 0, t = 1,2, \ldots, n.$$

- Exogénéité contemporaine:

$$\mathrm{E}(u_t|x_{t1}, \ldots, x_{tk}) = \mathrm{E}(u_t|x_t) = 0.$$

- Moins fort (suffisant pour consistence)

Problème d'exogénéité

- Considérons le modèle

$$y_t = \beta_0 + \beta_1 z_t + u_t,$$

- Ut doit être non corrélé avec les valeurs passées et futures de z_t !
- Si effet retardé : utiliser délais distribués
- Si effet de feedback de y sur z ?

Propriétés estimateurs MCO

- Considérons le modèle

$$y_t = \beta_0 + \beta_1 x_{t1} + \ldots + \beta_k x_{tk} + u_t,$$

- Stricte exogénéité
- Absence de collinéarité (délais distribués?)

- Estimateurs MCO non biaisés

$$\mathrm{E}(\hat{\beta}_j) = \beta_j,\ j = 0, 1, \ldots, k.$$

Efficacité

- Homoscédasticité : pour tout t,

$$\text{Var}(u_t|\boldsymbol{X}) = \text{Var}(u_t) = \sigma^2,$$

- discussion (politiques monétaires)

$$i3_t = \beta_0 + \beta_1 inf_t + \beta_2 def_t + u_t.$$

- Pas d'autocorrélation: Corr(u_t, u_s | X) = 0 pout tout s$\neq$t
- Interprétation: $\text{Corr}(u_t,u_s) = 0, \text{ for all } t \neq s.$

- Pas de conditions sur la corrélation entre x_s, x_t !

Inférence

- Si on suppose la normalité des résidus
  - Ut ~iid N(0, σ^2)
- On peut appliquer les test que nous avons développé dans les cours précédents
- On peut généralisé sous des hypothèses moins fortes : stationarité, ergodicité et variance finie (application d'un TCL)

Exemple

- Considérons la relation entre les taux à court termet et l'inflation et le déficit public

$$\hat{i3}_t = 1.25 + .613\ inf_t + .700\ def_t$$
$$(0.44)\quad (.076) \qquad (.118)$$
$$n = 49,\ R^2 = .697,\ \bar{R}^2 = .683.$$

- Influence positives

Délais distribués

- Considérons le modèle

$$y_t = \alpha_0 + \delta_0 z_t + \delta_1 z_{t-1} + \delta_2 z_{t-2} + u_t,$$

- Interprétation

$$\ldots, z_{t-2} = c, z_{t-1} = c, z_t = c + 1, z_{t+1} = c, z_{t+2} = c, \ldots.$$

- Calcul de l'effet ceteris paribus

Calcul de l'effet ceteris paribus

$$y_{t-1} = \alpha_0 + \delta_0 c + \delta_1 c + \delta_2 c,$$

$$y_t = \alpha_0 + \delta_0 (c + 1) + \delta_1 c + \delta_2 c,$$

$$y_{t+1} = \alpha_0 + \delta_0 c + \delta_1 (c + 1) + \delta_2 c,$$

$$y_{t+2} = \alpha_0 + \delta_0 c + \delta_1 c + \delta_2 (c + 1),$$

$$y_{t+3} = \alpha_0 + \delta_0 c + \delta_1 c + \delta_2 c,$$

- δ_0 = effet immédiat, δ_1 = effet à la période suivante, δ_2 = effet dans deux périodes

Effet de long terme

- Avant t, z_t = c; après t, z_s = c+1, s≥t

$$y_{t-1} = \alpha_0 + \delta_0 c + \delta_1 c + \delta_2 c,$$
$$y_t = \alpha_0 + \delta_0(c+1) + \delta_1 c + \delta_2 c,$$
$$y_{t+1} = \alpha_0 + \delta_0(c+1) + \delta_1(c+1) + \delta_2 c,$$
$$y_{t+2} = \alpha_0 + \delta_0(c+1) + \delta_1(c+1) + \delta_2(c+1),$$

- Somme des coefficients δ_0 , δ_1 et δ_2
- De manière générale:

$$\text{LRP} = \delta_0 + \delta_1 + \ldots + \delta_q.$$

La notion de "causalité" au sens de Granger

- Test l'hypothèse qu'aucune des valeurs passées de X ne peuvent prédire la valeur de Y
- Il s'agit donc d'un test d'hypothèses de Fisher.
- Notion utilisée dans un système VAR

Etude d'événements

- Utilisation de variable binaires pour tester l'effet d'un événement
- Ex : Effet d'une régulation de la circulation des camions (d) sur l'action des compagnies produisant des camions (R^f)

$$R_t^f = \beta_0 + \beta_1 R_t^m + \beta_2 d_t + u_t,$$

- R^m = rendement du marché
- Autre exemple : effet d'un accident d'avion sur la cotation boursière

Autocorrélation

- Supposons

$$u_t = \rho u_{t-1} + e_t, t = 1,2, \ldots, n \qquad |\rho| < 1,$$

- Propriétés MCO ?

$$y_t = \beta_0 + \beta_1 x_t + u_t,$$

- Estimateur MCO

$$\hat{\beta}_1 = \beta_1 + \mathrm{SST}_x^{-1} \sum_{t=1}^{n} x_t u_t,$$

- Non biaisé

Ecart-type

$$\begin{aligned}\mathrm{Var}(\hat{\beta}_1) &= \mathrm{SST}_x^{-2}\mathrm{Var}\left(\sum_{t=1}^{n} x_t u_t\right) = \mathrm{SST}_x^{-2}\left(\sum_{t=1}^{n} x_t^2 \mathrm{Var}(u_t)\right. \\ &\quad \left. + 2\sum_{t=1}^{n-1}\sum_{j=1}^{n-t} x_t x_{t+j}\mathrm{E}(u_t u_{t+j})\right) \\ &= \sigma^2/\mathrm{SST}_x + 2(\sigma^2/\mathrm{SST}_x^2)\sum_{t=1}^{n-1}\sum_{j=1}^{n-t}\rho^j x_t x_{t+j},\end{aligned}$$

- Biais si $\rho \neq 0$

Variable dépendante retardée

- Considérons un modèle AR(1)

$$y_t = \beta_0 + \beta_1 y_{t-1} + u_t,$$

- Avec des erreurs AR(1)

$$u_t = \rho u_{t-1} + e_t, t = 1,2, \ldots, n \qquad |\rho| < 1,$$

- Biais si $\rho \neq 0$:

$$\mathrm{Cov}(y_{t-1}, u_t) = \rho \mathrm{Cov}(y_{t-1}, u_{t-1}),$$

Reformulation AR(2)

- On peut écrire

$$u_{t-1} = y_{t-1} - \beta_0 - \beta_1 y_{t-2}$$

- Et remplacer dans

$$y_t = \beta_0 + \beta_1 y_{t-1} + u_t, \qquad u_t = \rho u_{t-1} + e_t.$$

- On obtient un AR(2)

$$\begin{aligned} y_t &= \beta_0 + \beta_1 y_{t-1} + \rho(y_{t-1} - \beta_0 - \beta_1 y_{t-2}) + e_t \\ &= \beta_0(1-\rho) + (\beta_1 + \rho)y_{t-1} - \rho\beta_1 y_{t-2} + e_t \\ &= \alpha_0 + \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + e_t, \end{aligned}$$

Test d'autocorrélation

- On veut tester

$$H_0\colon \rho = 0 \qquad H_1\colon \rho \neq 0$$

- Procédure simple:
- Obtenir les résidus par MCO, puis tester H_0

Test de Durbin-Watson

- Utilise la statistique

$$DW = \frac{\sum_{t=2}^{n} (\hat{u}_t - \hat{u}_{t-1})^2}{\sum_{t=1}^{n} \hat{u}_t^2}.$$

- Relation avec test précédent ?

$$DW \approx 2(1 - \hat{\rho}).$$

- Mais problèmes pratiques

Correction autocorrélation

- Considérons le modèle simple

$$y_{t-1} = \beta_0 + \beta_1 x_{t-1} + u_{t-1}$$

$$y_t = \beta_0 + \beta_1 x_t + u_t.$$

- Différence entre les deux équations

$$y_t - \rho y_{t-1} = (1-\rho)\beta_0 + \beta_1(x_t - \rho x_{t-1}) + e_t,\ t \geq 2,$$

- Régression sans autocorrélation

$$\tilde{y}_t = (1-\rho)\beta_0 + \beta_1 \tilde{x}_t + e_t,\ t \geq 2,$$

Estimation MCG

- Séries temporelles : corrélation d'une période à une autre des facteurs non observés
- $\varepsilon_i = \varepsilon_{i-1} + u_i$
- Matrice de corrélation des erreurs

$$\sigma^2 \mathbf{\Omega} = \sigma^2 \begin{bmatrix} 1 & \rho_1 & \cdots & \rho_{n-1} \\ \rho_1 & 1 & \cdots & \rho_{n-2} \\ & & \vdots & \\ \rho_{n-1} & \rho_{n-2} & \cdots & 1 \end{bmatrix}.$$

Interprétation MCG faisable

- Obternir ρ estimé par MCO
- Transformer les variables
- Faire la régression MCO avec les variables transformées; Revient à MCG avec covariance

$$\sigma^2\Omega = \frac{\sigma_u^2}{1-\rho^2}\begin{bmatrix} 1 & \rho & \rho^2 & \rho^3 & \cdots & \rho^{T-1} \\ \rho & 1 & \rho & \rho^2 & \cdots & \rho^{T-2} \\ \rho^2 & \rho & 1 & \rho & \cdots & \rho^{T-3} \\ \vdots & \vdots & \vdots & \vdots & \cdots & \rho \\ \rho^{T-1} & \rho^{T-2} & \rho^{T-3} & \cdots & \rho & 1 \end{bmatrix}.$$

Interprétation MCG faisable

- La matrice de transformation = quasi-différence:

$$\mathbf{y}_* = \begin{bmatrix} \sqrt{1-\rho^2}\, y_1 \\ y_2 - \rho y_1 \\ y_3 - \rho y_2 \\ \vdots \\ y_T - \rho y_{T-1} \end{bmatrix}, \quad \mathbf{X}_* = \begin{bmatrix} \sqrt{1-\rho^2}\, \mathbf{x}_1 \\ \mathbf{x}_2 - \rho \mathbf{x}_1 \\ \mathbf{x}_3 - \rho \mathbf{x}_2 \\ \vdots \\ \mathbf{x}_T - \rho \mathbf{x}_{T-1} \end{bmatrix}.$$

- Estimation de ρ par MCO; utilisée ensuite dans la formule ci-dessus

Résumé

- Stationnarité
- Autocorrélation
- Biais si erreur AR(1)
- Sinon perte efficacité
- Test
- Correction par MCG faisable