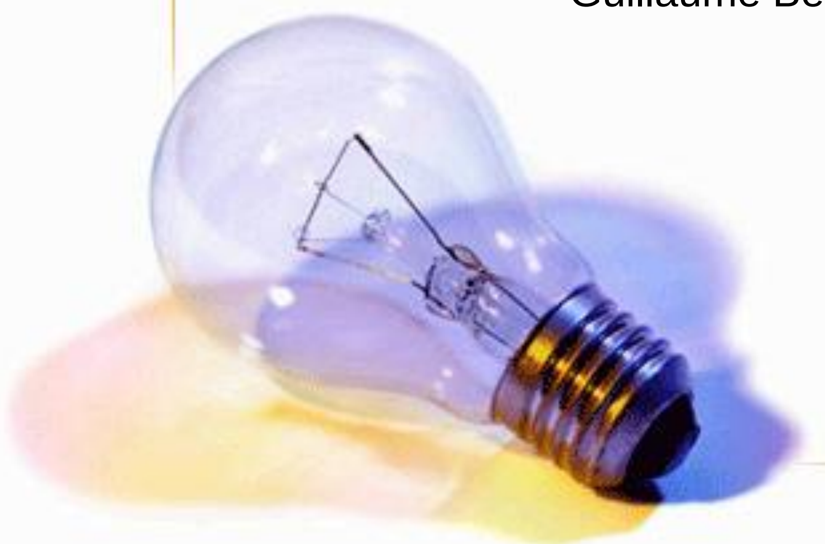
Le Modèle ARMAX

21 Mars 2013

Guillaume Beaurain – Vincent Gallmann



UNIVERSITÉ DE RENNES 1

Intérêts du modèle

Double but du modèle à fonction de transfert

Sert à la modélisation

Sert à la prévision

Plan

Panorama historique

En théorie:

Présentation du modèle (sous forme générale)

- Dans un but de modélisation
  i. Description des différentes étapes
  ii. Exemple illustratif:
  iii. Application

- Dans un but prédictif
  i. Description des différentes étapes
  ii. Exemple illustratif:
  iii. Application

Généralisation au cas de plusieurs input

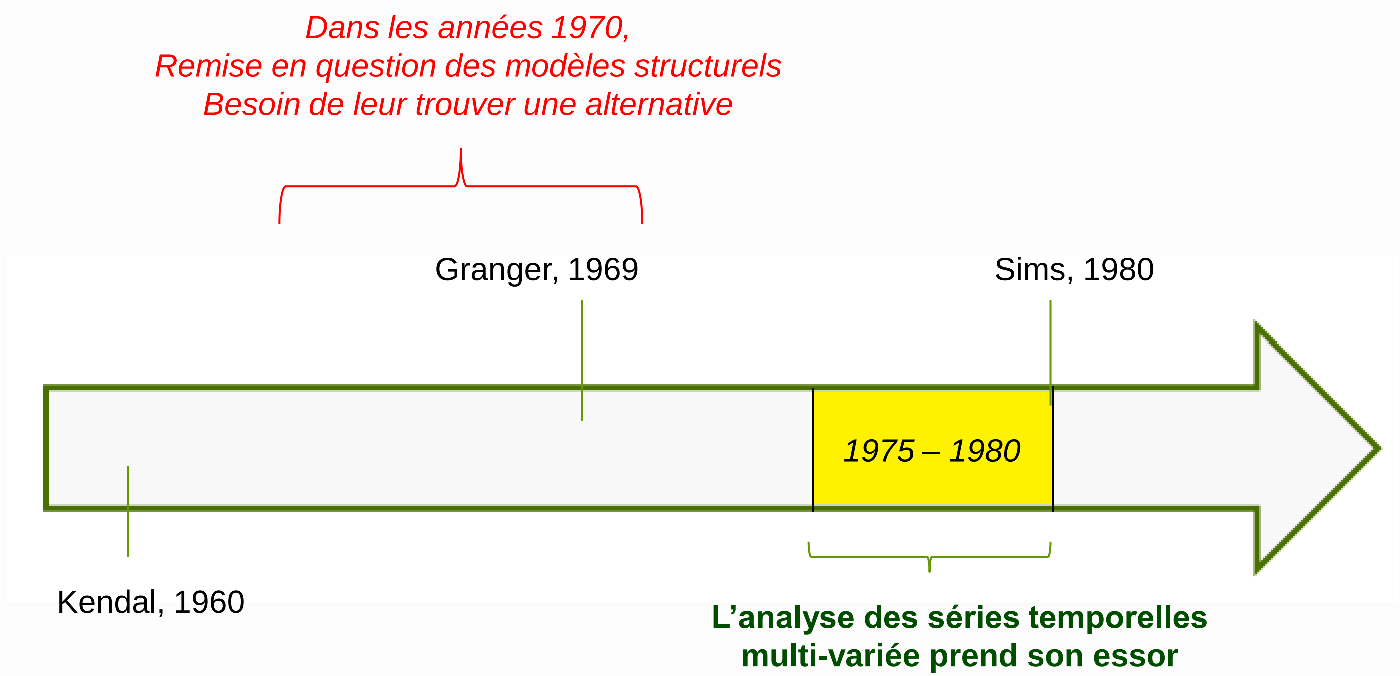
Pour aller plus loin:

Les modèles d'intervention

Discussion

Repères bibliographiques

Historiquement



Forme générale du modèle

Soit $\tilde{Y}_t$ la série endogène et $\tilde{X}_t$ la variable explicative

! Il est **NECESSAIRE** de stationnariser les séries !

En stationnarisant,on obtient:

$\mathbf{Y_t}$ la série endogène stationnarisée
$\mathbf{X_t}$ la série exogène stationnarisée

Forme générale du modèle

Le modèle à fonction de transfert:

$$Y_t = a_0 + A(L)Y_{t-1} + B(L)X_t + C(L)\varepsilon_t$$

où A(L), B(L), C(L) sont des polynômes retards

a_0 constante d'ajustement
A(L) correspond à la composante AR de Y_t
C(L) correspond à la composante MA de Y_t
B(L) est appelée FONCTION DE TRANSFERT

Dans un but de modélisation

$$Y_t = a_0 + A(L)Y_{t-1} + B(L)X_t + C(L)\varepsilon_t$$

5 étapes sont à respecter:

- **Etape 1:** Déterminer le processus ARMA de X_t
- **Etape 2:** Filtrer la série Y_t
- **Etape 3:** Déterminer l'ordre *(le retard)* de la fonction de transfert
- **Etape 4:** Déterminer C(L) *(s'il reste de la dynamique)*
- **Etape 5:** Estimation finale

Exemple illustratif:

Modélisation de l'impact d'une augmentation du prix du carburant à la pompe sur le nombre de nouvelles immatriculations de voitures

VOITURESCVS: Nombre d'immatriculations de voitures neuves

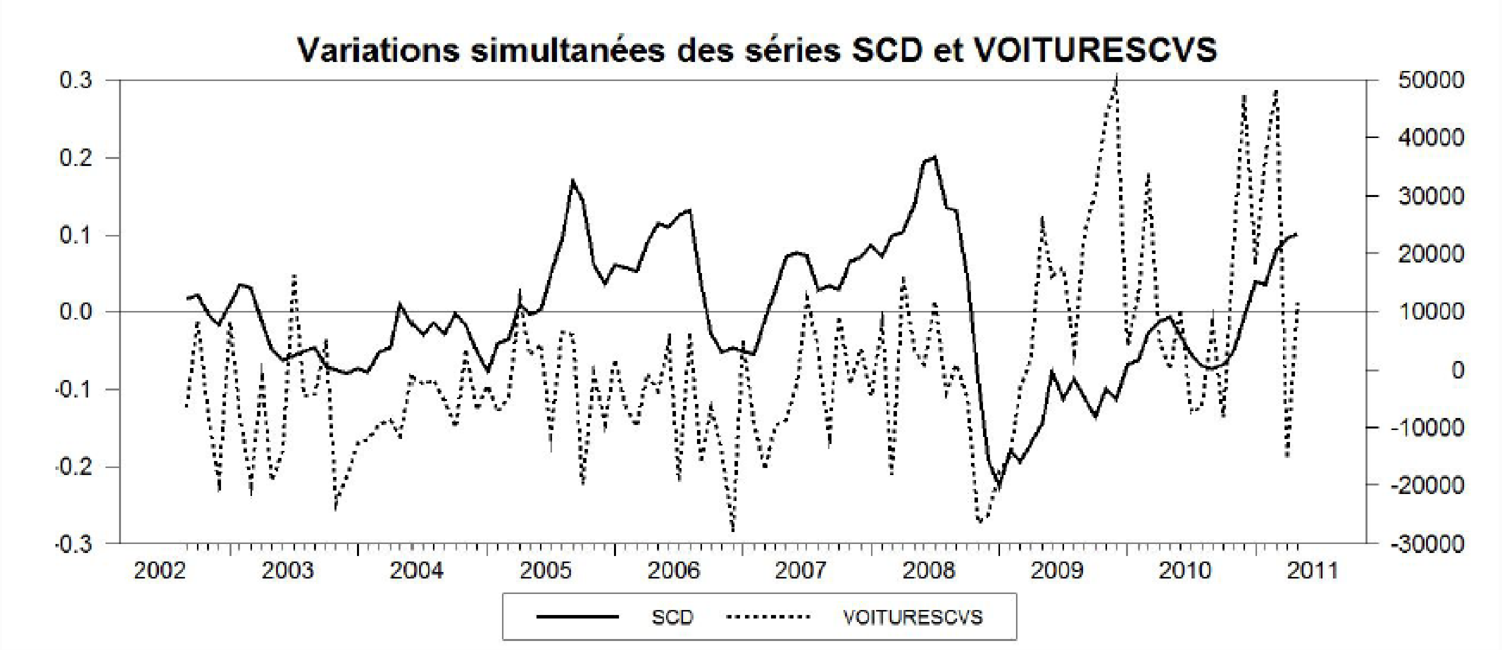
Source INSEE, de septembre 2002 à décembre 2011

SCD: Prix du litre de super carburant en euros

Source INSEE, de septembre 2002 à décembre 2011

Les motivations:

Pourquoi utiliser un modèle à fonction de transfert?



$$Y_t = a_0 + A(L)Y_{t-1} + B(L)X_t + C(L)\varepsilon_t$$

Etape 1 →

Pré-blanchiment de X_t

↓

Filtrage de Y_t par D(L)/E(L)

↓

Estimation des polynômes A(L) et B(L) ← Calcul des corrélations pour déterminer l'ordre de la fonction de transfert

↓

Eventuellement s'il reste de la dynamique, on estime CL)

↓

Estimation finale

Exemple illustratif:

Adaptation du modèle général à notre exemple

$$\begin{cases} VOITURESCVS_t = a_0 + A(L)\,VOITURESCVS_{t-1} + B(L)SCD_t + C(L)\,\varepsilon_t \\ D(L)SCD_t = E(L)\,\mu_t \end{cases}$$

Etape 1 → Pré-blanchiment de X_t

Exemple illustratif:

Etape 1: Pré-blanchiment de X_t

$$\begin{cases} VOITURESCVS_t = a_0 + A(L)\,VOITURESCVS_{t-1} + B(L)SCD_t + C(L)\,\varepsilon_t \\ D(L)SCD_t = E(L)\,\mu_t \end{cases}$$

On détermine le processus ARMA de la série X_t *(SCD_t)* :

➡ On utilise la **Méthodologie de Box & Jenkins**

But de la méthode: Modéliser les séries univariées au moyen de processus ARMA

Méthodologie de Box & Jenkins

Etape 1: **Identification**

Etude des fonctions d'autocorrélation et d'autocorrélation partielle

Etape 2: **Estimation**

Estimation des modèles candidats de l'étape 1

Etape 3: **Validation**

On vérifie que les résidus des estimations faites à l'étape 2 suivent un processus bruit blanc

Etape 4: **Prévision**

On effectue les prévisions de la série en partant de sa modélisation ARMA retenue

Etape 1 → Pré-blanchiment de X_t

$$\begin{cases} VOITURESCVS_t = a_0 + A(L)\,VOITURESCVS_{t-1} + B(L)SCD_t + C(L)\,\varepsilon_t \\ D(L)SCD_t = E(L)\,\mu_t \end{cases}$$

On retient le processus suivant pour représenter la dynamique de la série exogène:

$$\mathbf{SCD_t \sim AR\ (2)}$$

$$Y_t = a_0 + A(L)Y_{t-1} + B(L)X_t + C(L)\varepsilon_t$$

Pré-blanchiment de X_t

↓

Etape 2 → Filtrage de Y_t par D(L)/E(L)

↓

Estimation des polynômes A(L) et B(L) ← Calcul des corrélations pour déterminer l'ordre de la fonction de transfert

↓

Eventuellement s'il reste de la dynamique, on estime CL)

↓

Estimation finale

Dans un but de modélisation

Il faut partir du système suivant:

$$\begin{cases} Y_t = a_0 + A(L)Y_{t-1} + B(L)X_t + C(L)\varepsilon_t & (1) \\ D(L)X_t = E(L)\mu_t & (2) \end{cases}$$

Avec:

$$D(L) = 1 - d_1 L - d_2 L^2 - \ \dots$$
$$E(L) = 1 + e_1 L + e_2 L^2 + \ \dots$$

Dans un but de modélisation

Etape 2: Filtrage de Y_t par D(L)/E(L)

$$\frac{D(L)}{E(L)}Y_t = A(L)\frac{D(L)}{E(L)}Y_{t-1} + B(L)\frac{D(L)}{E(L)}X_t + C(L)\frac{D(L)}{E(L)}\varepsilon_t$$

On note:

$$\begin{cases} f_{y_t} = \dfrac{D(L)}{E(L)}Y_t \\ e_t = \dfrac{D(L)}{E(L)}\varepsilon_t \end{cases}$$

$$f_{y_t} = A(L)f_{y_{t-1}} + B(L)\mu_t + C(L)e_t$$

Exemple illustratif:

Etape 2: Filtrage de Yt par D(L)/E(L)

$$\begin{cases} VOITURESCVS_t = a_0 + A(L)\, VOITURESCVS_{t-1} + B(L) SCD_t + C(L)\, \varepsilon_t \\ D(L) SCD_t = E(L)\, \mu_t \end{cases}$$

$$f_{VOITURESCVS_t} = a_0 + A(L) f_{VoituresCVS_{t-1}} + B(L)\mu_t + C(L) e_t$$

Sur la série filtrée

$$Y_t = a_0 + A(L)Y_{t-1} + B(L)X_t + C(L)\varepsilon_t$$

Pré-blanchiment de X_t

↓

Filtrage de Y_t par D(L)/E(L)

↓

Etape 3 → Estimation des polynômes A(L) et B(L) ← Calcul des corrélations pour déterminer l'ordre de la fonction de transfert

↓

Eventuellement s'il reste de la dynamique, on estime C(L)

↓

Estimation finale

Exemple illustratif:

Etape 3: Estimation des polynômes A(L) et B(L)

Etape 3 → Estimation des polynômes A(L) et B(L)

Calcul des corrélations pour déterminer l'ordre de la fonction de transfert

$$f_{VOITURESCVS_t} = a_0 + A(L) f_{VoituresCVS_{t-1}} + B(L)\mu_t + C(L)e_t$$

On s'intéresse au **corrélogramme croisé** de la série filtrée $f_{VOITURESCVS}$

Zoom sur la *Cross Correlation Function*

$$E[y_t x_t] = 0$$
$$E[y_t x_{t-1}] = 0$$
$$\vdots$$
$$E[y_t x_{t-d}] = b_d \sigma_x^2$$
$$E[y_t x_{t-d-1}] = b_d a_1 \sigma_x^2$$
$$E[y_t x_{t-d-2}] = b_d a_1^2 \sigma_x^2$$

Sous sa forme général:

$$E[y_t x_{t-i}] = \begin{cases} 0 \; pour \; tout \; i < d \\ b_d a_i^{i-d} \sigma_x^2 \; pour \; tout \; i \geq d \end{cases}$$

Exemple illustratif:

Etape 3: Estimation des polynômes A(L) et B(L)

- Estimation de B(L):

lag	**-10**	**-9**	**-8**	**-7**	**-6**	**-5**	**-4**	**-3**	**-2**	**-1**
	0.05443	-0.06238	0.03483	-0.02718	-0.08050	0.10495	-0.00213	-0.03854	0.01408	0.09178
Lead	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
	-0.02985	0.05996	**-0.10555**	0.01235	-0.02872	-0.05647	-0.02101	0.02076	-0.04892	0.03366
	10									
	-0.00842									

Le polynôme B(L) sera d’ordre **2**

Exemple illustratif:

Etape 3: Estimation des polynômes A(L) et B(L)

- Estimation de A(L):

Quel est l'ordre du processus autorégressif de la série filtrée?

Via la méthode de Box & Jenkins, on obtient:

➡ Le polynôme A(L) sera d'ordre **2**

Sur la série filtrée

$$Y_t = a_0 + A(L) Y_{t-1} + B(L) X_t + C(L) \varepsilon_t$$

Pré-blanchiment de X_t

↓

Filtrage de Y_t par D(L)/E(L)

↓

Estimation des polynômes A(L) et B(L) ← Calcul des corrélations pour déterminer l'ordre de la fonction de transfert

↓

Etape 4 → *Eventuellement* s'il reste de la dynamique, on estime C(L)

↓

Estimation finale

Exemple illustratif:

Etape 4: Estimation éventuelle de C(L)

Test de Ljung Box:

$$Q = T(T+2)\sum_{k=1}^{T/4}\frac{\hat{\rho}_k^2(\hat{\varepsilon}_t)}{T-k}$$

On teste H_0 l'hypothèse d'absence d'autocorrélations,

$$H_0 \; : \; \hat{\rho}_1^2(\hat{\varepsilon}_t) \; = \; \hat{\rho}_2^2(\hat{\varepsilon}_t) \; = \; ... \; = \; 0$$

Q suit une loi du Khi 2 à (T/4 – p – q)

Dans notre exemple,

Q(27-2) = 23,64
p-value = 0,54

Pas besoin d'estimer C(L)

Exemple illustratif:

Etape 5: Estimation finale

Etape 5 → Estimation finale

$$VOITURESCVS_t = a_0 + A(L)\,VOITURESCVS_{t-1} + B(L)SCD_t + C(L)\,\varepsilon_t$$

Variable	Coefficient	Ecart-type	p-value
a_0	52,23	478,56	0,91
a_1	-0,98	0,08	0
a_2	-0,59	0,08	0
b_2	-45707,3	26052,93	0,082

→ Le coefficient de la fonction de transfert est significatif à **10%**

Exemple illustratif:

Etape 5: Estimation finale

$$VOITCVS_t = 52{,}23 - 0{,}98\ VOITCVS_{t-1} - 0{,}59\ VOITCVS_{t-2} - 45707\ SCD_{t-2} + \varepsilon_t$$

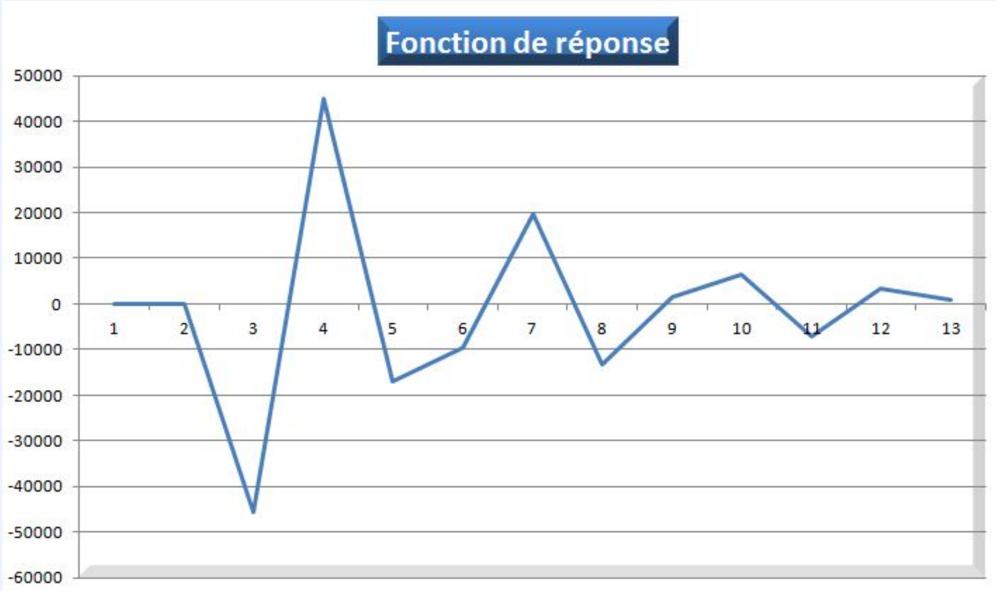
✓ On accepte que les **résidus sont normalement distribués** *(Test de Jarque-Bera, p-value>0,5)*

Une hausse de **0,01€** du prix du super-carburant engendre après **2 mois** une diminutionde **457** nouvelles immatriculations.

Exemple illustratif:

Conclusion

Fonction de réponse:



Effet de long-terme:

$$\frac{B(L)}{1 - A(L)}$$

$$Y_t = a_0 + A(L)Y_{t-1} + B(L)X_t + C(L)\varepsilon_t$$

Pré-blanchiment de X_t

Filtrage de Y_t par D(L)/E(L)

Estimation des polynômes A(L) et B(L)

Calcul des corrélations pour déterminer l'ordre de la fonction de transfert

Eventuellement s'il reste de la dynamique, on estime CL)

Estimation finale

Le modèle ARMAX

Dans un but prédictif

Dans un but prédictif

$$Y_t = a_0 + A(L)Y_{t-1} + B(L)X_t + C(L)\varepsilon_t$$

Le modèle devient:

$$Y_t = H(L).X_t + \mu_t$$

Où **H(L)** est la fonction de transfert

Y_t la série endogène stationnarisée
X_t la série exogène stationnarisée

NB: **μ_t** **n'est pas nécessairement** un bruit blanc

Zoom sur la fonction de transfert

$$Y_t = H(L).X_t + \mu_t$$

Forme générale: $$H(L) = \sum_{\tau \geq 0} H_\tau L^\tau$$

Poids de la fonction de transfert

Rapport de deux polynômes d'ordre fini:

$$H(L) = H_0 + H_1 L + H_2 L^2 + \ \dots = \frac{\omega(L)}{\delta(L)} L^b$$

$$H(L) = \frac{\omega_0 - \omega_1 L - \cdots - \omega_s L^s}{1 - \delta_1 L - \cdots - \delta_r L^r} L^b$$

$$Y_t = \left(\frac{\omega_0 - \omega_1 L - \cdots - \omega_s L^s}{1 - \delta_1 L - \cdots - \delta_r L^r} L^b\right) X_t + \left(\frac{1 - \theta_1 L - \cdots - \theta_q L^q}{1 - \varphi_1 L - \cdots - \varphi_p L^p}\right) \mu_t$$

Dans un but prédictif: 6 étapes à suivre

Construction d'un modèle ARMA pour l'input X_t

↓

Filtrage de l'output Y_t par les résidus estimés du modèle ARMA de X_t

↓

Calcul des corrélations pour déterminer la forme de la fonction de transfert

↓

Estimation des poids de la fonction de transfert et construction du modèle

↓

Test de diagnostic

NON → Estimation des poids de la fonction de transfert et construction du modèle

↓ OUI

Prévision

Exemple illustratif:

Un exemple classique

La relation entre la concentration en CO2 et le taux d'alimentation en air et en méthane

(Travaux de Box et Jenkins)

$\mathbf{Y_t}$ **:** Concentration de CO_2 *(l'endogène)*

$\mathbf{X_t}$ **:** Taux d'alimentation en air et méthane *(l'exogène)*

Dans un but prédictif

Les deux premières étapes sont **similaires** au modèle de la première partie *(Dans un but de modélisation)*

Etape 3 → **Calcul des corrélations pour déterminer la forme de la fonction de transfert**

$$Y_t = H(L)X_{t-3} + \mu_t = \frac{\omega(L)}{\delta(L)} X_{t-3} + \mu_t$$

Exemple illustratif:

Etape 4 → Estimation des poids de la fonction de transfert et construction du modèle

$$Y_t = \frac{-0{,}53 + 0{,}38\,L + 0{,}52\,L^2}{1 - 0{,}55\,L} X_{t-3} + \frac{1}{1 - 1{,}53L + 0{,}63L^2}\,\mu_t$$

Exemple illustratif:

Etape 5 → Test de diagnostic

↓ **OUI**

Prévision

Test de diagnostic:

✓ On vérifie si les résidus estimés contiennent encore de l'informartion

✓Si les résidus sont non corrélés, l'estimation *(la spécification)* est bonne ***(O.K pour la prévision)***

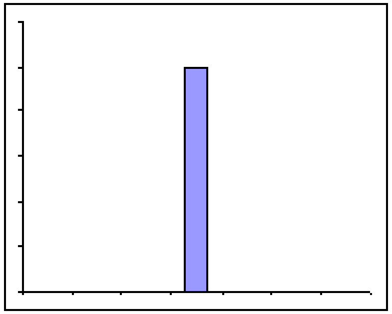
✓ On peut aussi utiliser la « statistique porte-manteau »:

$$Q = \beta \sum_{\tau=0}^{M} \widehat{\rho_{\varepsilon}}(\tau) \qquad M = \min\left(\frac{T}{2}, 3\sqrt{T}\right)$$

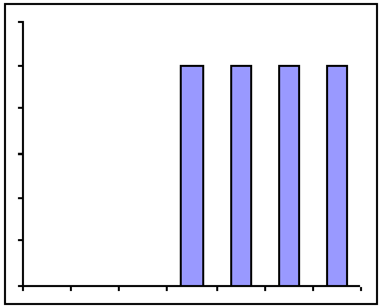
Les modèles d'intervention

$$Y_t = a_0 + A(L)Y_{t-1} + B(L)X_t + C(L)\varepsilon_t$$

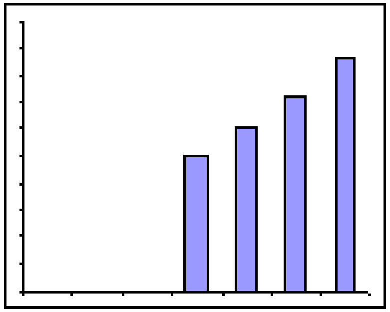
$$Y_t = a_0 + A(L)Y_{t-1} + b_0 X_t + C(L)\varepsilon_t$$

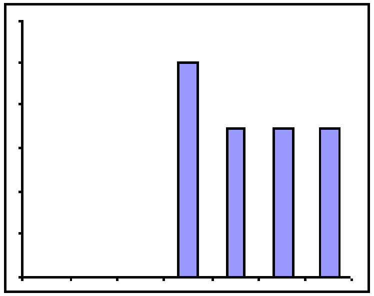


Effet ponctuel $X_t = \begin{cases} 0 \; si \; t \neq t' \\ 1 \; si \; t = t' \end{cases}$



Effet palier $X_t = \begin{cases} 0 \; si \; t < t_1 \\ 1 \; si \; t_1 \leq \; t \leq t_2 \\ 0 \; si \; t > t_2 \end{cases}$





Pour résumer

Pré-blanchiment de X_t

↓

Filtrage de Y_t par D(L)/E(L)

↓

Estimation des polynômes A(L) et B(L)

↓

Eventuellement s'il reste de la dynamique, on estime C(L)

↓

Estimation finale

Modélisation

Construction d'un modèle ARMA pour l'input X_t

↓

Filtrage de l'output Y_t par les résidus estimés du modèle ARMA de X_t

↓

Calcul des corrélations pour déterminer la forme de la fonction de transfert

↓

Estimation des poids de la fonction de transfert et construction du modèle

↓

Test de diagnostic

↓ **OUI**

Prévision

NON

Prévision

Discussion

- ✓ Nécessité de travailler avec des séries stationnaires
- ✓ Difficilement généralisable pour plusieurs inputs
- ✓ Problème potentiel de *spurious correlation*

Repères Bibliographiques

Time series techniques for economists

Terence C. Mills

Econométrie des séries temporelles

G. Bresson / A. Pirotte

Applied econometric time series

W. Enders