



## 5 - Modèles ARIMA

ALAIN QUARTIER-LA-TENTE  
Insee

1. Stationnarité et différenciation
2. Construction du modèle ARIMA
3. Détermination du modèle ARIMA
4. ARIMA et ETS
5. Retour sur TRAMO-SEATS
6. Conclusion

# Objectifs de cette séquence

---

- Notions sur la stationnarisation et la différenciation
- Modèles (S)ARIMA

## Questions de positionnement

---

Qu'est-ce qu'un processus stationnaire ?

Tendance, cycle, saisonnalité sont-ils des processus stationnaires ?

Que signifie "ARIMA" et que reflète un tel modèle ?

Comment se comportent les erreurs de prévision d'un modèle ARIMA ?

Qu'est-ce qu'un SARMA ?

Y a-t-il un lien entre ARIMA et ETS ?

# Sommaire

---

## 1. Stationnarité et différenciation

1.1 Notion de stationnarité

1.2 Repérer la stationnarité

1.3 Stationnariser une série

1.4 Tests

2. Construction du modèle ARIMA

3. Détermination du modèle ARIMA

4. ARIMA et ETS

5. Retour sur TRAMO-SEATS

## Quelques définitions (1/2)

---

*Série temporelle* : suite de variables aléatoires  $(X_t)_t$  dont on observe une réalisation  $(X_t(\omega))_t$

La suite  $(X_t)_t$  est appelée *processus stochastique*

## Quelques définitions (1/2)

---

*Série temporelle* : suite de variables aléatoires  $(X_t)_t$  dont on observe une réalisation  $(X_t(\omega))_t$

La suite  $(X_t)_t$  est appelée *processus stochastique*

Un processus est dit *stationnaire* lorsque la loi de  $X_t$  n'évolue pas dans le temps : distribution  $\forall s, (X_t, \dots, X_{t+s})$  indépendante du temps

$\implies$  série plus ou moins horizontale et de variance constante

➔ Notion pour faire l'inférence et construire un modèle ARIMA

## Quelques définitions (2/2)

---

Stationnarité, hypothèse invérifiable  $\Rightarrow$  en pratique processus *faiblement stationnaire* :

- les moments d'ordre 2 existent
- espérance constante
- covariance entre  $t$  et  $t - h$  ne dépend pas du temps, mais de la distance  $h$   
 $\implies$  variance constante



## Quelques définitions (2/2)

---

Stationnarité, hypothèse invérifiable  $\Rightarrow$  en pratique processus *faiblement stationnaire* :

- les moments d'ordre 2 existent
- espérance constante
- covariance entre  $t$  et  $t - h$  ne dépend pas du temps, mais de la distance  $h$   
 $\implies$  variance constante

Exemple : un bruit blanc, i.e. :

- espérance nulle
- covariance entre  $t$  et  $t - h$  nulle, pour tout  $h \neq 0$
- variance non nulle et constante

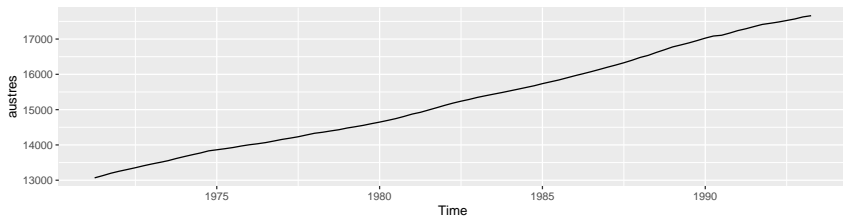
# Comment identifier une série non-stationnaire (en niveau) ?

---

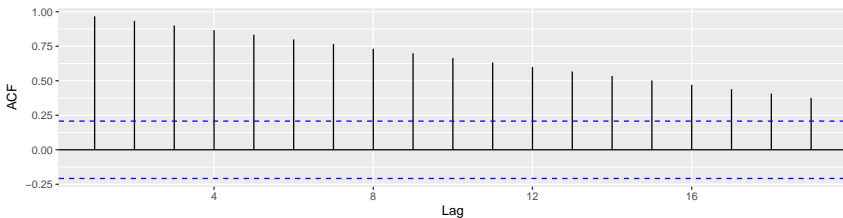
- Tracer le chronogramme
- Etudier l'ACF :
  - Série non-stationnaire : tend lentement vers 0 et  $\hat{\rho}(1)$  souvent positif et élevé
  - Série stationnaire : tend rapidement vers 0

# Exemple

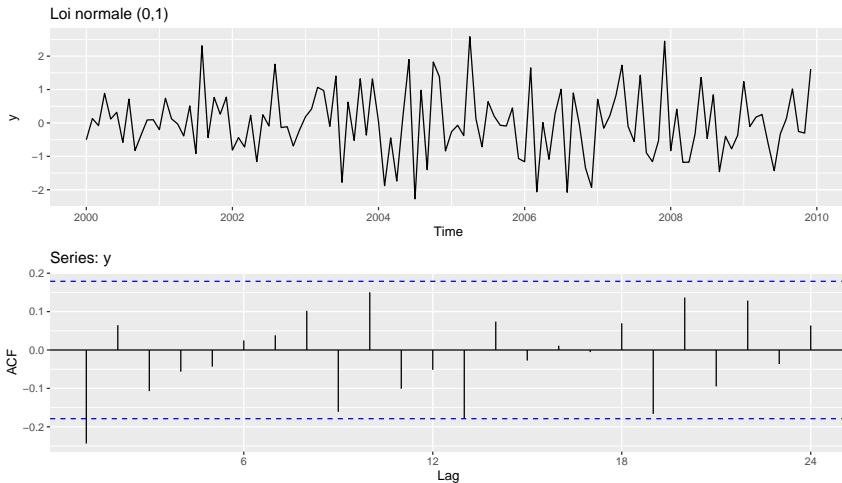
Nombre de résidents australiens



Series: austres



# Exemple



## La différenciation pour stabiliser le niveau

- Si la série différenciée est un bruit blanc de moyenne nulle (marche aléatoire) :

$$(I - B)y_t = y_t - y_{t-1} = \varepsilon_t \implies y_t = y_0 + \sum_{i=1}^t \varepsilon_i$$

➔ Modèle naïf

Généralement mouvement à la hausse ou à la baisse aléatoire,

## La différenciation pour stabiliser le niveau

- Si la série différenciée est un bruit blanc de moyenne nulle (marche aléatoire) :

$$(I - B)y_t = y_t - y_{t-1} = \varepsilon_t \implies y_t = y_0 + \sum_{i=1}^t \varepsilon_i$$

### ➔ Modèle naïf

Généralement mouvement à la hausse ou à la baisse aléatoire,

- Si la série différenciée est un bruit blanc de moyenne non nulle (marche aléatoire avec dérive / *drift*) :

$$(I - B)y_t = c + \varepsilon_t \implies y_t = y_0 + ct + \sum_{i=1}^t \varepsilon_i$$

## La différenciation pour stabiliser le niveau

- Si la série différenciée est un bruit blanc de moyenne nulle (marche aléatoire) :

$$(I - B)y_t = y_t - y_{t-1} = \varepsilon_t \implies y_t = y_0 + \sum_{i=1}^t \varepsilon_i$$

### ➔ Modèle naïf

Généralement mouvement à la hausse ou à la baisse aléatoire,

- Si la série différenciée est un bruit blanc de moyenne non nulle (marche aléatoire avec dérive / *drift*) :

$$(I - B)y_t = c + \varepsilon_t \implies y_t = y_0 + ct + \sum_{i=1}^t \varepsilon_i$$

- Parfois on a besoin de différencier plusieurs fois

$(I - B)^2 y_t = (y_t - y_{t-1}) - (y_{t-1} - y_{t-2})$  ou de faire une différenciation saisonnière  $(I - B^m)y_t = y_t - y_m$

## La différenciation pour stabiliser le niveau

- Si la série différenciée est un bruit blanc de moyenne nulle (marche aléatoire) :

$$(I - B)y_t = y_t - y_{t-1} = \varepsilon_t \implies y_t = y_0 + \sum_{i=1}^t \varepsilon_i$$

### ➔ Modèle naïf

Généralement mouvement à la hausse ou à la baisse aléatoire,

- Si la série différenciée est un bruit blanc de moyenne non nulle (marche aléatoire avec dérive / *drift*) :

$$(I - B)y_t = c + \varepsilon_t \implies y_t = y_0 + ct + \sum_{i=1}^t \varepsilon_i$$

- Parfois on a besoin de différencier plusieurs fois  
 $(I - B)^2 y_t = (y_t - y_{t-1}) - (y_{t-1} - y_{t-2})$  ou de faire une différenciation saisonnière  
 $(I - B^m)y_t = y_t - y_m$
- Si saisonnalité importante, commencer par la différenciation saisonnière



# Modèles Intégrés (1/3)

---

Soit  $X$ , processus « tendance linéaire » :

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

Calculer l'espérance et la variance de la v.a.  $X_t$  ?  
 $X$  est stationnaire ?

## Modèles Intégrés (1/3)

---

Soit  $X$ , processus « tendance linéaire » :

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

Calculer l'espérance et la variance de la v.a.  $X_t$  ?

$X$  est stationnaire ?

Différence d'ordre 1 :

$$(I - B)X_t = ?$$

La série obtenue est-elle stationnaire ?

Si  $X$  est un processus « tendance polynomiale d'ordre 2 », comment stationnariser la série ?

## Modèles Intégrés (2/3)

---

Soit  $X$ , processus « saisonnier stable » :

$$X_t = S_t + \varepsilon_t \quad \text{avec} \quad \forall t, S_t = S_{t+s}$$

$X$  stationnaire ?

## Modèles Intégrés (2/3)

---

Soit  $X$ , processus « saisonnier stable » :

$$X_t = S_t + \varepsilon_t \quad \text{avec} \quad \forall t, S_t = S_{t+s}$$

$X$  stationnaire ?

Différence d'ordre 1, avec retard d'ordre  $s$  :

$$(I - B^s)X_t = ?$$

La série obtenue est-elle stationnaire ?

Si  $X_t = a + bt + S_t + \varepsilon_t$ , que donnerait cette différenciation ?

## Modèles Intégrés (3/3)

---

Une différenciation « simple » d'ordre  $d$  supprime les tendances polynomiales d'ordre  $d$  :

$$(I - B)^d X_t$$

Une différenciation « saisonnière » supprime aussi les tendances linéaires :

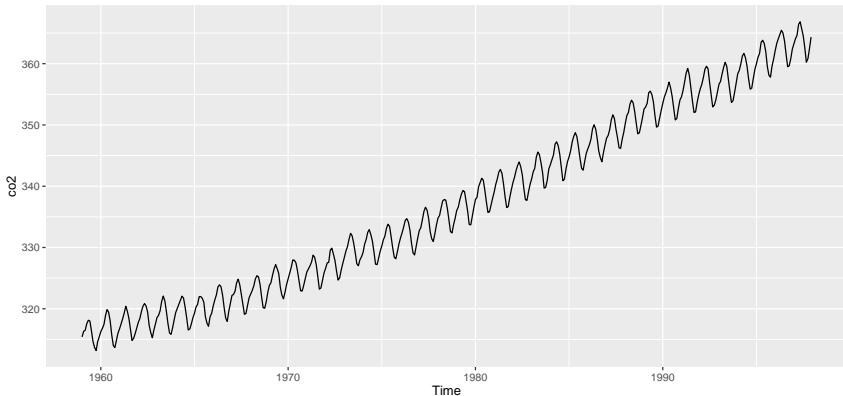
$$(I - B^s) X_t$$

Une différenciation « saisonnière » d'ordre  $D$  plus grand que 1 est rare :

$$(I - B^s)^D X_t$$

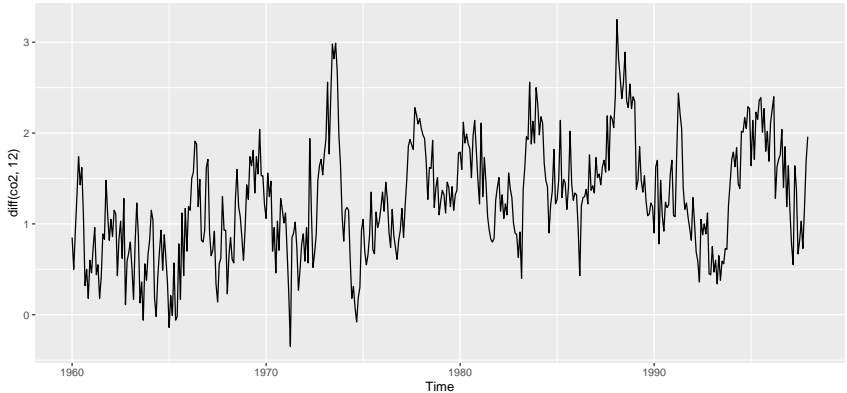
# Exemple

```
autoplot(co2)
```



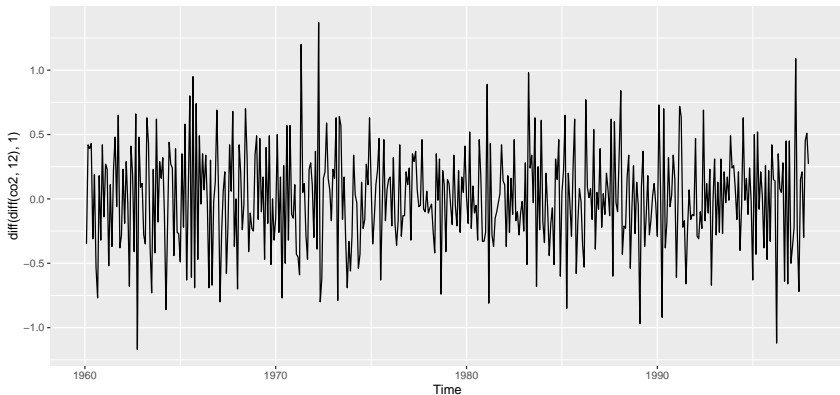
# Exemple

```
autoplot(diff(co2, 12))
```



# Exemple

```
autoplot(diff(diff(co2, 12), 1))
```





# Faut-il toujours différencier ?

Pour modéliser une série avec tendance on peut distinguer deux types de non-stationnarité :

1. Modèle trend-stationnaire :

$$X_t = a + bt + \varepsilon_t$$

2. Modèle avec racine unité

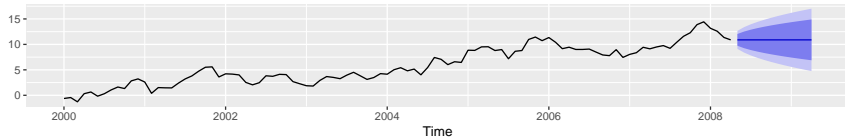
$$(1 - B)Y_t = b + \eta \implies Y_t = a + bt + \underbrace{\sum_{i=1}^t \eta_i}_{\text{tend. stochastique}}$$

On a  $\mathbb{V}[X_t] = \mathbb{V}[\varepsilon_t] = \text{cst}$  indépendante du temps mais  $\mathbb{V}[Y_t] = t\mathbb{V}[\eta_t]$

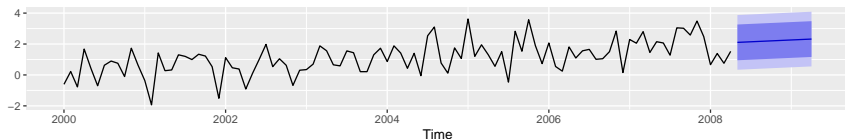
# Exemple

```
set.seed(1); e = rnorm(100)
u1 = ts(cumsum(e), start = 2000, frequency = 12)
u2 = ts(e + seq_along(e)/50, start = 2000, frequency = 12)
m1 = Arima(u1, order = c(0, 1, 0))
m2 = Arima(u2, include.drift = TRUE)
(autoplot(u1, y = NULL, main = "Marche aléatoire") +
  autolayer(forecast(m1, h = 12))) /
(autoplot(u2, y = NULL, main = "Trend-stationnaire") +
  autolayer(forecast(m2, h = 12)))
```

Marche aléatoire



Trend-stationnaire



# Tests de racine unitaire

---

Plusieurs tests existent pour déterminer l'ordre de différenciation :

- Sur séries non-saisonnnières :
  - Test de Dickey-Fuller augmenté (ADF) `fUnitRoots::adfTest()`, `tseries::adf.test()`, `urca::ur.df()`. ( $H_0$ ) racine unitaire (avec ou non tendance linéaire)
  - Test de Phillips-Perron `tseries::pp.test`, `urca::ur.pp` ou `feasts::unitroot_pp()`. ( $H_0$ ) racine unitaire (avec ou non tendance linéaire)
  - Test KPSS `tseries::kpss.test()` ou `urca::ur.kpss()`. ( $H_0$ ) série stationnaire (avec ou non tendance linéaire)
- Pour les séries saisonnières : d'autres tests du type Canova-Hansen (`uroot::ch.test`), F-Tests, etc.

## Tests de racine unitaire

---

Plusieurs tests existent pour déterminer l'ordre de différenciation :

- Sur séries non-saisonnnières :
  - Test de Dickey-Fuller augmenté (ADF) `fUnitRoots::adfTest()`, `tseries::adf.test()`, `urca::ur.df()`. ( $H_0$ ) racine unitaire (avec ou non tendance linéaire)
  - Test de Phillips-Perron `tseries::pp.test`, `urca::ur.pp` ou `feasts::unitroot_pp()`. ( $H_0$ ) racine unitaire (avec ou non tendance linéaire)
  - Test KPSS `tseries::kpss.test()` ou `urca::ur.kpss()`. ( $H_0$ ) série stationnaire (avec ou non tendance linéaire)
- Pour les séries saisonnières : d'autres tests du type Canova-Hansen (`uroot::ch.test`), F-Tests, etc.

`forecast::ndiffs()` (ou `feasts::unitroot_ndiffs`) et `forecast::nsdiffs()` (ou `feasts::unitroot_nsdiffs`) permettent de déterminer les ordres de différenciation en utilisant ces tests.

# Exemple (1)

```
uroot::ch.test(co2)
```

```
##  
## Canova and Hansen test for seasonal stability  
##  
## data:  co2  
##  
##      statistic pvalue  
## Jan      2.3112      0 ***  
## Feb      2.3104      0 ***  
## Mar      2.3137      0 ***  
## Apr      2.3156      0 ***  
## May      2.3158      0 ***  
## Jun      2.3152      0 ***  
## Jul      2.3142      0 ***  
## Aug      2.3155      0 ***  
## Sep      2.3128      0 ***  
## Oct      2.3135      0 ***  
## Nov      2.3136      0 ***  
## Dec      2.3148      0 ***  
## joint    2.9814 0.0476 *
```

## Exemple (2)

---

```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Test type: seasonal dummies
## NW covariance matrix lag order: 18
## First order lag: no
## Other regressors: no
## P-values: based on response surface regressions
uroot::ch.test(diff(co2, 12))
```

```
##
## Canova and Hansen test for seasonal stability
##
## data:  diff(co2, 12)
##
##      statistic pvalue
## Jan      0.8938 0.003 **
## Feb      0.7737 0.0067 **
## Mar      0.5276 0.0327 *
## Apr      0.6513 0.0149 *
## May      0.6378 0.0162 *
```

## Exemple (3)

```
## Jun      0.6103 0.0193 *
## Jul      0.5519 0.0281 *
## Aug      0.5907 0.0219 *
## Sep      0.6174 0.0185 *
## Oct      0.7576 0.0075 **
## Nov      0.8991 0.0029 **
## Dec      1.0098 0.0014 **
## joint    1.3838 0.5406
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Test type: seasonal dummies
## NW covariance matrix lag order: 18
## First order lag: no
## Other regressors: no
## P-values: based on response surface regressions
forecast::nsdiffs(co2) # Autre test que Canova-Hansen est utilisé

## [1] 1
```

## Exemple (4)

---

```
tseries::kpss.test(diff(co2, 12))
```

```
## Warning in tseries::kpss.test(diff(co2, 12)): p-value smaller than printed p-  
## value
```

```
##
```

```
## KPSS Test for Level Stationarity
```

```
##
```

```
## data: diff(co2, 12)
```

```
## KPSS Level = 1.9439, Truncation lag parameter = 5, p-value = 0.01
```

```
tseries::kpss.test(diff(diff(co2, 12), 1))
```

```
## Warning in tseries::kpss.test(diff(diff(co2, 12), 1)): p-value greater than  
## printed p-value
```

```
##
```

```
## KPSS Test for Level Stationarity
```

```
##
```

```
## data: diff(diff(co2, 12), 1)
```

```
## KPSS Level = 0.011476, Truncation lag parameter = 5, p-value = 0.1
```



## Exemple (5)

---

```
forecast::ndiffs(diff(co2, 12))
```

```
## [1] 1
```

# Sommaire

---

## 1. Stationnarité et différenciation

## 2. Construction du modèle ARIMA

### 2.1 Modèles AR et MA

### 2.2 Modèle ARMA

### 2.3 Modèles SARMA et modèles intégrés

## 3. Détermination du modèle ARIMA

## 4. ARIMA et ETS

## 5. Retour sur TRAMO-SEATS

## 6. Conclusion

## La partie « modélisation ARIMA »

---

ARIMA, modèle auto-projectif :

$$X_t = f(X_{t-1}, X_{t-2}, X_{t-3}, \dots, \varepsilon_t, \varepsilon_{t-1}, \varepsilon_{t-2} \dots)$$

Trouver  $f$  ?

Sous hypothèse de stationnarité, il existe un « modèle ARMA » qui approche la série.

Conséquence (th de Wold) : erreurs de prévision se comportent comme le résidu du modèle (bruit blanc)

On privilégie les modèles avec faible nombre de paramètres.

Méthode de Box et Jenkins pour estimer et juger de la qualité des modèles.

# Modèles Autorégressifs (AR)

---

Modèle *autorégressif* d'ordre  $p$ ,  $AR(p)$  :

$$\begin{aligned} X_t &= \phi_1 X_{t-1} + \phi_2 X_{t-2} + \cdots + \phi_p X_{t-p} + \varepsilon_t \\ \Leftrightarrow (1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p) X_t &= \varepsilon_t \\ \Leftrightarrow \Phi(B) X_t &= \varepsilon_t \end{aligned}$$

# Modèles Autorégressifs (AR)

---

Modèle *autorégressif* d'ordre  $p$ ,  $AR(p)$  :

$$\begin{aligned} X_t &= \phi_1 X_{t-1} + \phi_2 X_{t-2} + \dots + \phi_p X_{t-p} + \varepsilon_t \\ \Leftrightarrow (1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p) X_t &= \varepsilon_t \\ \Leftrightarrow \Phi(B) X_t &= \varepsilon_t \end{aligned}$$

On retrouve les marches aléatoires (sans ou avec dérive)

# Modèles Autorégressifs (AR)

---

Modèle *autorégressif* d'ordre  $p$ ,  $AR(p)$  :

$$\begin{aligned}X_t &= \phi_1 X_{t-1} + \phi_2 X_{t-2} + \cdots + \phi_p X_{t-p} + \varepsilon_t \\ \Leftrightarrow (1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p) X_t &= \varepsilon_t \\ \Leftrightarrow \Phi(B) X_t &= \varepsilon_t\end{aligned}$$

On retrouve les marches aléatoires (sans ou avec dérive)

Un AR modélise l'influence des  $p$  réalisations passées sur la réalisation courante : effet mémoire

# Modèles Autorégressifs (AR)

---

Modèle *autorégressif* d'ordre  $p$ ,  $AR(p)$  :

$$\begin{aligned} X_t &= \phi_1 X_{t-1} + \phi_2 X_{t-2} + \dots + \phi_p X_{t-p} + \varepsilon_t \\ \iff (1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p) X_t &= \varepsilon_t \\ \iff \Phi(B) X_t &= \varepsilon_t \end{aligned}$$

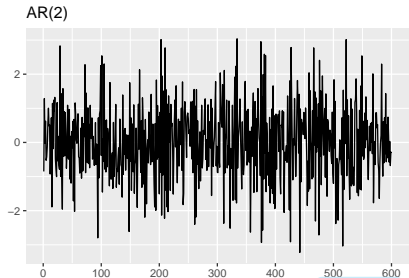
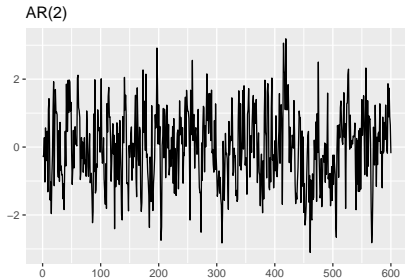
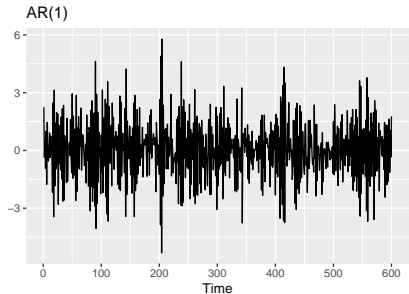
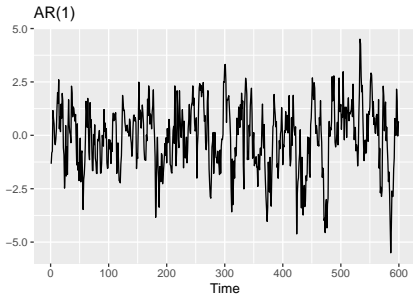
On retrouve les marches aléatoires (sans ou avec dérive)

Un AR modélise l'influence des  $p$  réalisations passées sur la réalisation courante : effet mémoire

Exemples classiques (voir TP)

- Le niveau du lac Huron peut être modélisé par un  $AR(1)$  ou un  $AR(2)$  ;
- $AR(2)$  : nombre de tâches solaires - Yules

# Exemples : que dire sur $\phi_1$ ?





## Exemples : réponse

---

Dans les graphiques de droite on observe une alternance entre périodes positives et négatives :  $\phi_1 < 0$

```
set.seed(100)
ar1_pos = arima.sim(n=600, list(ar=0.8))
ar1_neg = arima.sim(n=600, list(ar=-0.8))
ar2_pos = arima.sim(n=600, list(ar=c(0.3, 0.2)))
ar2_neg = arima.sim(n=600, list(ar=c(-0.3, 0.2)))
(autoplot(ar1_pos, main = "AR(1)", y = NULL) +
  autoplot(ar1_neg, main = "AR(1)", y = NULL)) /
(autoplot(ar2_pos, main = "AR(2)", y = NULL) +
  autoplot(ar2_neg, main = "AR(2)", y = NULL)
)
```

## Reconnaitre un modèle $AR(p)$

---

Pour reconnaître un  $AR(p)$  on peut analyser l'autocorrélogramme partiel (PACF) :  $r(k)$  mesure relation entre  $y_t$  et  $y_{t-k}$  en enlevant les effets de  $y_{t-1}, \dots, y_{t-k-1}$

$\alpha_k$  est le coefficient  $\phi_k$  dans la régression

$$y_t = c + \phi_1 y_{t-1} + \dots + \phi_k y_{t-k} + \varepsilon_t$$

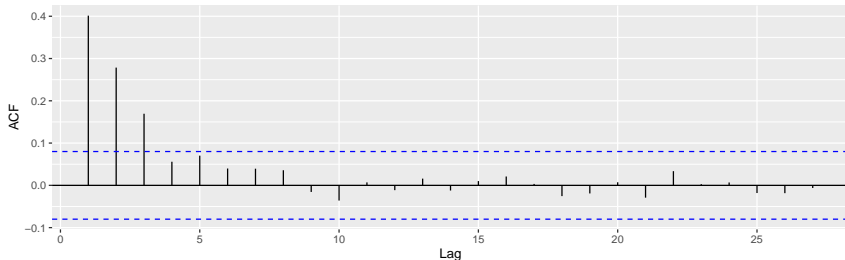
On a  $r(1) = \rho(1)$

Pour un  $AR(p)$  :

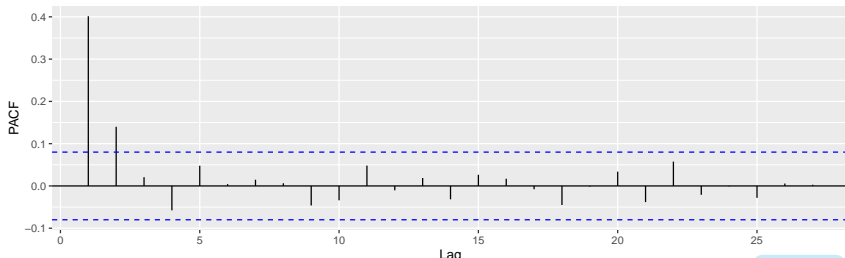
- **ACF** :  $\rho(h)$  décroît exponentiellement vers 0 (ou de manière sinusoïdale si  $\phi_1 < 0$ )
- **PACF** :  $r(h) = 0$  pour  $h > p$

# Exemple $AR(2)$ $y_t = 0.3y_{t-1} + 0.2y_{t-2} + \varepsilon_t$

$$AR(2) \ y_t = 0.3y_{t-1} + 0.2y_{t-2} + \varepsilon_t$$



$$AR(2) \ y_t = 0.3y_{t-1} + 0.2y_{t-2} + \varepsilon_t$$



# Conditions de stationnarité

---

On restreint généralement les modèles autorégressifs aux modèles stationnaires.

Pour cela il faut que les racines de  $\Phi$  soient en dehors du cercle unité (sinon un choc pourrait avoir un effet permanent)

# Modèles « Moving Average » (MA)

Modèle *moyenne mobile* d'ordre  $q$ ,  $MA(q)$  :

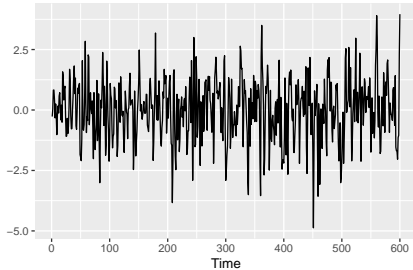
$$\begin{aligned} X_t &= \varepsilon_t + \theta_1\varepsilon_{t-1} + \theta_2\varepsilon_{t-2} + \cdots + \theta_q\varepsilon_{t-q} \\ \iff X_t &= (1 + \theta_1B + \theta_2B^2 + \cdots + \theta_qB^q)\varepsilon_t \\ \iff X_t &= \Theta(B)\varepsilon_t \end{aligned}$$

Rmq : dans certains logiciels/packages (e.g. R, JDemetra) on utilise plutôt la notation  $X_t = \varepsilon_t - \theta_1\varepsilon_{t-1} - \theta_2\varepsilon_{t-2} - \cdots - \theta_q\varepsilon_{t-q}$

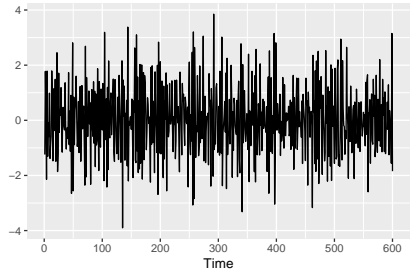
Résulte d'une accumulation non persistante de  $q$  chocs indépendants

# Exemples

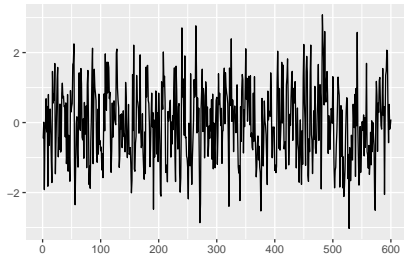
MA(1)



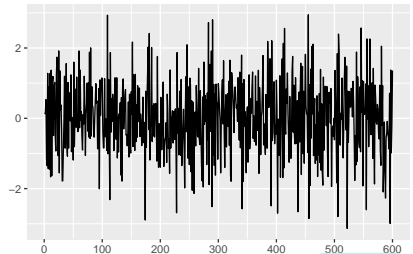
MA(1)



MA(2)



MA(2)



## Reconnaitre un modèle $MA(q)$

---

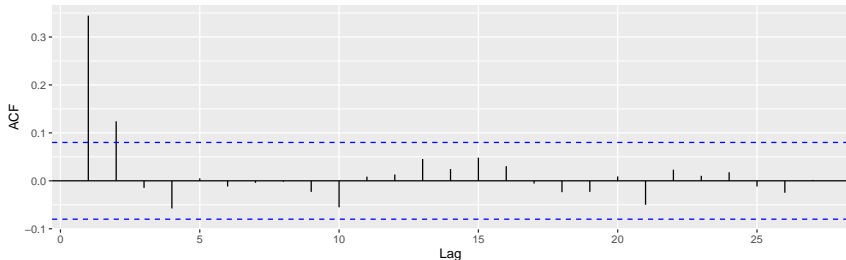
Pour reconnaître un  $MA(q)$  on peut analyser l'autocorrélogramme (ACF)

Pour un  $MA(q)$  :

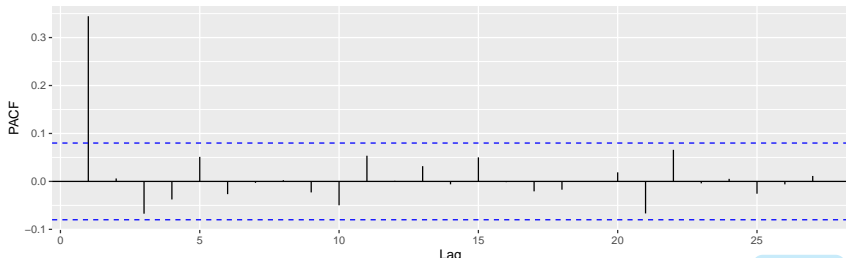
- **ACF** :  $\rho(h) = 0$  pour  $h > q$
- **PACF** :  $r(h) = 0$  décroît exponentiellement vers 0

# Exemple $MA(2) y_t = \varepsilon_t + 0.3\varepsilon_{t-1} + 0.2\varepsilon_{t-2}$

$$MA(2) y_t = \varepsilon_t + 0.3\varepsilon_{t-1} + 0.2\varepsilon_{t-2}$$



$$MA(2) y_t = \varepsilon_t + 0.3\varepsilon_{t-1} + 0.2\varepsilon_{t-2}$$





# Modèles ARMA

---

Modèles  $ARMA(p, q)$  : combine  $AR(p)$  et  $MA(q)$ , sans ou avec constante

$$\Phi(B)X_t = \Theta(B)\varepsilon_t$$

$$\Phi(B)X_t = \mu + \Theta(B)\varepsilon_t$$

# Modèles SARMA

---

Modèle  $SARMA(P, Q)$  :  $ARMA$  avec polynôme d'ordre  $s$  (4 séries trimestrielles, 12 séries mensuelles) :

$$\Phi(B^s)X_t = \Theta(B^s)\varepsilon_t \text{ ou } \Phi_s(B)X_t = \Theta_s(B)\varepsilon_t$$

Intérêt :

- montrer autocorrélations d'ordre  $s$
- simplifier l'écriture par factorisation

# Modèles SARMA

---

Modèle  $SARMA(P, Q)$  :  $ARMA$  avec polynôme d'ordre  $s$  (4 séries trimestrielles, 12 séries mensuelles) :

$$\Phi(B^s)X_t = \Theta(B^s)\varepsilon_t \text{ ou } \Phi_s(B)X_t = \Theta_s(B)\varepsilon_t$$

Intérêt :

- montrer autocorrélations d'ordre  $s$
- simplifier l'écriture par factorisation

$ARMA(p, q)(P, Q)$  combine parties régulière et saisonnière :  
 $ARMA(p, q) \times SARMA(P, Q)$ .

Identique à  $ARMA(p + P * s, q + Q * s)$

Exemple série mensuelle :  $ARMA(1, 1)(1, 1) = ARMA(13, 13)$

# Modèles ARIMA

---

$ARIMA(p, d, q)$  modélise les séries non stationnaires avec tendance

$$\Phi(B)(I - B)^d X_t = \Theta(B)\varepsilon_t$$

$ARIMA(p, d, q)(P, D, Q)$  modélise les séries avec tendance et saisonnalité

$$\Phi(B)\Phi_s(B)(I - B)^d(I - B^s)^D X_t = \Theta(B)\Theta_s(B)\varepsilon_t$$

Factorisation des polynômes en  $B$  de la partie *régulière* et de la partie *saisonnière*

# Modèles ARIMA

---

$ARIMA(p, d, q)$  modélise les séries non stationnaires avec tendance

$$\Phi(B)(I - B)^d X_t = \Theta(B)\varepsilon_t$$

$ARIMA(p, d, q)(P, D, Q)$  modélise les séries avec tendance et saisonnalité

$$\Phi(B)\Phi_s(B)(I - B)^d(I - B^s)^D X_t = \Theta(B)\Theta_s(B)\varepsilon_t$$

Factorisation des polynômes en  $B$  de la partie *régulière* et de la partie *saisonnière*

Modèle le plus souvent observé :  $ARIMA(0, 1, 1)(0, 1, 1)$  appelé modèle Airline (voir TP)

# Sommaire

---

1. Stationnarité et différenciation
2. Construction du modèle ARIMA
- 3. Détermination du modèle ARIMA**
  - 3.1 Méthode de Box-Jenkins
  - 3.2 Méthode générale
4. ARIMA et ETS
5. Retour sur TRAMO-SEATS
6. Conclusion

# Méthode de Box-Jenkins

---

1. Stationnariser le processus :  $d, D$
2. Identifier les ordres ARMA :  $p, P, q, Q$  ➔ structure d'autocorrélation de la série
3. Estimer les coefficients ARMA ➔ degré de variabilité de la structure d'autocorrélation  
➔ Peut-on simplifier le modèle ?
4. Valider le modèle ➔ résidus = bruit blanc ?
5. Choix du modèle (si plusieurs modèles valides) ➔ critères d'information
6. Prévision

# Méthode générale

---

- Tracer la série : transformation nécessaire ? Points atypiques ?



# Méthode générale

---

- Tracer la série : transformation nécessaire ? Points atypiques ?
- Stationnariser la série pour déterminer  $D$  puis  $d$  (analyse graphique et/ou tests)

# Méthode générale

---

- Tracer la série : transformation nécessaire ? Points atypiques ?
- Stationnariser la série pour déterminer  $D$  puis  $d$  (analyse graphique et/ou tests)
- Examen des ACF/PACF pour déterminer des ordres  $P, Q, p, q$  plausibles

# Méthode générale

---

- Tracer la série : transformation nécessaire ? Points atypiques ?
- Stationnariser la série pour déterminer  $D$  puis  $d$  (analyse graphique et/ou tests)
- Examen des ACF/PACF pour déterminer des ordres  $P, Q, p, q$  plausibles
- Sélection des modèles par minimisation AICc

# Méthode générale

---

- Tracer la série : transformation nécessaire ? Points atypiques ?
- Stationnariser la série pour déterminer  $D$  puis  $d$  (analyse graphique et/ou tests)
- Examen des ACF/PACF pour déterminer des ordres  $P, Q, p, q$  plausibles
- Sélection des modèles par minimisation AICc
- Vérifier la qualité des résidus : si ne ressemble pas à un bruit blanc changer de modèle

# Méthode générale

---

- Tracer la série : transformation nécessaire ? Points atypiques ?
- Stationnariser la série pour déterminer  $D$  puis  $d$  (analyse graphique et/ou tests)
- Examen des ACF/PACF pour déterminer des ordres  $P, Q, p, q$  plausibles
- Sélection des modèles par minimisation AICc
- Vérifier la qualité des résidus : si ne ressemble pas à un bruit blanc changer de modèle
- Prévission

## Méthode utilisée dans `forecast::auto.arima()` (1)

1. On choisit  $D$  (Canova-Hansen STL) puis  $d$  déterminé en utilisant des tests successifs de KPSS.

## Méthode utilisée dans `forecast::auto.arima()` (1)

1. On choisit  $D$  (Canova-Hansen STL) puis  $d$  déterminé en utilisant des tests successifs de KPSS.
2. Sélection d'un des 4 modèles en minimisant l'AICc :  
 $ARIMA(2, d, 2)(1, D, 1)$ ,  $ARIMA(2, d, 2)(0, D, 0)$ ,  $ARIMA(1, d, 0)(1, D, 0)$   
et  $ARIMA(0, d, 1)(0, D, 1)$

## Méthode utilisée dans `forecast::auto.arima()` (1)

1. On choisit  $D$  (Canova-Hansen STL) puis  $d$  déterminé en utilisant des tests successifs de KPSS.
2. Sélection d'un des 4 modèles en minimisant l'AICc :  
 $ARIMA(2, d, 2)(1, D, 1)$ ,  $ARIMA(2, d, 2)(0, D, 0)$ ,  $ARIMA(1, d, 0)(1, D, 0)$   
et  $ARIMA(0, d, 1)(0, D, 1)$
3. On considère 30 variations du modèle retenu :
  - En faisant varier un seul des paramètres  $p$ ,  $q$ ,  $P$  ou  $Q$  de  $\pm 1$  ;
  - En faisant varier  $p$  et  $q$  en même temps de  $\pm 1$  ;
  - En faisant varier  $P$  et  $Q$  en même temps de  $\pm 1$  ;
  - En incluant ou non la constante.
  - Si un modèle minimise l'AICc on recommence.



## Méthode utilisée dans `forecast::auto.arima()` (2)

Si on n'a qu'une série, utiliser `forecast::auto.arima(., stepwise = FALSE, approximation = FALSE)` pour étudier tous les modèles (plus lent).

D'autres algorithmes (TRAMO ou pickmdl) utilisent également d'autres tests : autocorrélation (Ljung-Box à l'ordre  $2m$ ), tests de sur-différenciation, de passage au log.

# Sommaire

---

1. Stationnarité et différenciation
2. Construction du modèle ARIMA
3. Détermination du modèle ARIMA
- 4. ARIMA et ETS**
5. Retour sur TRAMO-SEATS
6. Conclusion

# Équivalence entre ARIMA et ETS

---

Il y a plusieurs équivalences entre ETS et ARIMA :

- Modèles exponentiels linéaires sont tous des cas particulier de modèle ARIMA
- Modèles exponentiels non linéaires n'ont pas d'équivalent ARIMA
- De nombreux ARIMA n'ont pas de modèle ETS équivalent
- Les modèles ETS sont non-stationnaires

## Exemple 1 : $ETS(A, N, N)$

---

$$\begin{cases} y_t = I_{t-1} + \varepsilon_t \\ I_t = I_{t-1} + \alpha \varepsilon_t \end{cases} \implies y_t = (y_{t-1} - \varepsilon_{t-1} + \alpha \varepsilon_{t-1}) + \varepsilon_t$$

Donc  $y_t \sim ARIMA(0, 1, 1)$  :

$$(1 - B)y_t = (1 + (\alpha - 1)B)\varepsilon_t$$

## Exemple 2 : $ETS(A, A, N)$

$$\begin{cases} y_t = l_{t-1} + b_{t-1} + \varepsilon_t \\ l_t = l_{t-1} + b_{t-1} + \alpha\varepsilon_t \\ b_t = b_{t-1} + \beta\varepsilon_t \end{cases} \iff \begin{cases} y_t = l_{t-1} + b_{t-1} + \varepsilon_t \\ (1-B)l_t = b_{t-1} + \alpha\varepsilon_t \\ (1-B)b_t = \beta\varepsilon_t \end{cases}$$

D'où

$$\begin{aligned} (1-B)^2 y_t &= (1-B)B(Bb_t + \alpha\varepsilon_t) + (1-B)B\beta\varepsilon_t + (1-B)^2\varepsilon_t \\ &= B^2\beta\varepsilon_t + (1-B)B(\beta + \alpha)\varepsilon_t + (1-2B+B^2)\varepsilon_t \\ &= [1 + (\alpha + \beta - 2)B + (1-\alpha)B^2] \varepsilon_t \end{aligned}$$

et  $y_t \sim ARIMA(0, 2, 2)$

## Exemple 3 : $ETS(A, A, A)$

$$\begin{cases} y_t = l_{t-1} + b_{t-1} + s_{t-m} + \varepsilon_t \\ l_t = l_{t-1} + b_{t-1} + \alpha\varepsilon_t \\ b_t = b_{t-1} + \beta\varepsilon_t \\ s_t = s_{t-m} + \gamma\varepsilon_t \end{cases} \iff \begin{cases} y_t = Bl_t + Bb_t + B^m s_t + \varepsilon_t \\ (1-B)l_t = Bb_t + \alpha\varepsilon_t \\ (1-B)b_t = \beta\varepsilon_t \\ (1-B^m)s_t = \gamma\varepsilon_t \end{cases}$$

$$\begin{aligned} (1-B)(1-B^m)y_t &= (1-B^m)B^2b_t + (1-B^m)B\alpha\varepsilon_t + (1-B^m)B\beta\varepsilon_t + \\ &\quad (1-B)B^m\gamma\varepsilon_t + (1-B)(1-B^m)\varepsilon_t \\ &= \left[ (1+B+\dots+B^{m-1})B^2\beta + (1-B^m)B(\alpha+\beta) + \right. \\ &\quad \left. (1-B)B^m(\gamma-1) + (1-B) \right] \varepsilon_t \end{aligned}$$

Donc  $y_t \sim ARIMA(0, 1, m+1)(0, 1, 0)$

# Comparaison des modèles ETS et ARIMA

---

Peut-on comparer les modèles ETS et ARIMA en utilisant un critère d'information ?

## Comparaison des modèles ETS et ARIMA

---

Peut-on comparer les modèles ETS et ARIMA en utilisant un critère d'information ?



**NON!** Attention aux ordres de différenciation, sur un modèle ARIMA ne comparer par exemple que les modèles qui ont le même ordre de différenciation



## Comparaison des modèles ETS et ARIMA

---

Peut-on comparer les modèles ETS et ARIMA en utilisant un critère d'information ?



**NON!** Attention aux ordres de différenciation, sur un modèle ARIMA ne comparer par exemple que les modèles qui ont le même ordre de différenciation

... Mais quid lorsque les modèles sont équivalents : peut-on comparer  $ETS(A, A, A)$  avec un  $ARIMA(0, 1, 1)(0, 1, 1)$  ?

## Comparaison des modèles ETS et ARIMA

---

Peut-on comparer les modèles ETS et ARIMA en utilisant un critère d'information ?



NON ! Attention aux ordres de différenciation, sur un modèle ARIMA ne comparer par exemple que les modèles qui ont le même ordre de différenciation

... Mais quid lorsque les modèles sont équivalents : peut-on comparer  $ETS(A, A, A)$  avec un  $ARIMA(0, 1, 1)(0, 1, 1)$  ?



Encore NON ! Avec un modèle ARIMA, lorsqu'il y a différenciation on perd les premières données alors que pour un ETS on a des prévisions sur toutes les données.

# Sommaire

---

1. Stationnarité et différenciation
2. Construction du modèle ARIMA
3. Détermination du modèle ARIMA
4. ARIMA et ETS
- 5. Retour sur TRAMO-SEATS**
  - 5.1 TRAMO
  - 5.2 SEATS
6. Conclusion

# Principe de TRAMO

---

TRAMO = Time series Regression with ARIMA noise, Missing values and Outliers

Objectifs de TRAMO :

- corriger la série de points atypiques, des effets de calendrier et imputation des valeurs manquantes
- prolonger la série
- fournir à SEATS le modèle ARIMA à la base de la décomposition

$$Y_t = \sum \hat{\alpha}_i O_{it} + \sum \hat{\beta}_j C_{jt} + \varepsilon_t$$

Et on modélise  $(Y_t - \sum \hat{\alpha}_i O_{it} + \sum \hat{\beta}_j C_{jt}) = \varepsilon_t$  comme un modèle ARIMA

## Principe de SEATS (1/3)

---

SEATS = Signal Extraction in ARIMA Time Series

SEATS utilise le modèle ARIMA de la série linéarisée TRAMO :

$$\underbrace{\Phi(B)\Phi_s(B)(I-B)^d(I-B^s)^D}_{\Phi(B)} X_t = \underbrace{\Theta(B)\Theta_s(B)}_{\Theta(B)} \varepsilon_t$$

Hypothèses :

1. La série linéarisée peut être modélisée par un modèle ARIMA
2. Les différentes composantes sont décorréées et chaque composante peut être modélisée par un modèle ARIMA
3. Les polynômes AR des composantes n'ont pas de racine commune

## Principe de SEATS (2/3)

---

On factorise le polynôme AR  $\Phi(B)$  :

$$\Phi(B) = \phi_T(B)\phi_S(B)\phi_C(B)$$

- $\phi_T(B)$  racines correspondant à la tendance
- $\phi_S(B)$  racines correspondant à la saisonnalité
- $\phi_C(B)$  racines correspondant au cycle

## Principe de SEATS (3/3)

$X_t$  est exprimé sous la forme :

$$X_t = \frac{\Theta(B)}{\Phi(B)} \varepsilon_t = \underbrace{\frac{\theta_T(B)}{\phi_T(B)} \varepsilon_{T,t}}_{\text{Tendance}} + \underbrace{\frac{\theta_S(B)}{\phi_S(B)} \varepsilon_{S,t}}_{\text{Saisonnalité}} + \underbrace{\frac{\theta_C(B)}{\phi_C(B)} \varepsilon_{C,t}}_{\text{Cycle}} + \underbrace{\nu_t}_{\substack{\text{Irrégulier} \\ \text{(bruit} \\ \text{blanc)}}$$

Un modèle ARIMA est associé à chaque composante.

Infinité de solutions : on retient celle qui minimise la variance de l'irrégulier

➔ Estimation par filtre de Wiener-Kolmogorov

# Sommaire

---

1. Stationnarité et différenciation
2. Construction du modèle ARIMA
3. Détermination du modèle ARIMA
4. ARIMA et ETS
5. Retour sur TRAMO-SEATS
- 6. Conclusion**



# Les essentiels

---

Les séries économiques ne sont pas stationnaires, ni leur niveau, ni leurs fluctuations ne sont constants dans le temps

Intégrer un processus permet de le stationnariser

Un MA capte les fluctuations non persistantes autour d'un niveau constant - processus stationnaire

Un AR met en évidence l'influence des réalisations passées sur la réalisation courante

Un ARIMA reflète la structure des autocorrélations de la série, ainsi que le degré de sa variabilité dans le temps

L'examen des résidus permet de valider les modèles, le choix "optimal" se fait grâce aux critères d'information

# Bibliographie

---

Hyndman, R.J., & Athanasopoulos, G. (2018) *Forecasting : principles and practice*, 2nd edition, OTexts : Melbourne, Australia. [OTexts.com/fpp2](https://otexts.com/fpp2). Accessed on nov. 2022.

Hyndman, R.J., & Athanasopoulos, G. (2021) *Forecasting : principles and practice*, 3rd edition, OTexts : Melbourne, Australia. [OTexts.com/fpp3](https://otexts.com/fpp3). Accessed on nov. 2022.