

# Processus ARMA stationnaires

## Chapitre 2

Économétrie des séries temporelles



# Structure du cours

Processus linéaires

Décomposition de Wold

Modèles de séries temporelles univariées : AR(p), MA(q)

Équations de Yule Walker

Décomposition de prédiction

Estimation du maximum de vraisemblance

Exemple empirique

# 1. Processus linéaires

Supposons que  $\{y_t\}$  est un processus stochastique :

## Opérateur retard

$$Ly_t = y_{t-1}$$

$$L^j y_t = y_{t-j} \quad \forall j \in \mathbb{N}$$

## Opérateur différence

$$\Delta y_t = (1 - L) y_t = y_t - y_{t-1}$$

$$\Delta^j y_t = (1 - L)^j y_t \quad \forall j \in \mathbb{N}_+$$

$$\Delta_s y_t = (1 - L^s) y_t = y_t - y_{t-s}$$

## Processus linéaires (suite)

**Filtre linéaire** transforme une série d'entrée  $\{x_t\}$  en une série de sortie  $\{y_t\}$  en utilisant un polynôme de retard  $A(L)$  :

$$\begin{aligned} y_t = A(L)x_t &= \left( \sum_{-n}^m a_j L^j \right) x_t = \sum_{j=-n}^m a_j x_{t-j} \\ &= a_{-n} x_{t+n} + \cdots + a_0 x_0 + \cdots + a_m x_{t-m} \end{aligned}$$

### Processus linéaire

$$y_t = A(L)\varepsilon_t = \left( \sum_{-\infty}^{\infty} a_j L^j \right) \varepsilon_t = \sum_{j=-\infty}^{\infty} a_j \varepsilon_{t-j} \quad (1)$$

où  $\varepsilon_t \sim WN(0, \sigma^2)$

Note - pour  $|x| < 1$  :

$$1 + x + x^2 + \cdots + x^n = \sum_{i=0}^n x^i = \frac{1 - x^{n+1}}{1 - x} \rightarrow \frac{1}{1 - x} = \sum_{i=0}^{\infty} x^i \quad (2)$$

## 2. Décomposition de Wold

**La décomposition de Wold** - tout processus stationnaire de covariance à moyenne nulle  $\{y_t\}$  peut être représenté sous la forme :

$$y_t = \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j} + \kappa_t$$

où  $\psi_0 = 1$  et  $\sum_{j=0}^{\infty} \psi_j^2 < \infty$ .  $\varepsilon_t$  est un bruit blanc ; il représente l'erreur commise dans la prévision de  $y_t$  sur la base d'une fonction linéaire de son passé  $Y_{t-1} = y_{t-j}$  pour  $j \geq 1$  :

$$\varepsilon_t \equiv y_t - \widehat{E}(y_t | Y_{t-1})$$

Notez que  $\text{corr}(\kappa_t, \varepsilon_{t-j}) = 0, \forall j$ , mais :

$$\kappa_t = \widehat{E}(\kappa_t | Y_{t-1})$$

$\widehat{E}$  indique une projection linéaire sur un vecteur de variables aléatoires  $Y_t$ .

## Approche de Box-Jenkins

Approximation du polynôme à retard infini avec le rapport de deux polynômes d'ordre fini  $\alpha(L)$  et  $\beta(L)$  :

$$\Psi(L) = \sum_{j=0}^{\infty} \psi_j L^j \simeq \frac{\beta(L)}{\alpha(L)} = \frac{1 + \beta_1 L + \cdots + \beta_q L^q}{1 - \alpha_1 L - \cdots - \alpha_p L^p}$$

## Types de modèles de séries temporelles

Type	Modèle	$p$	$q$
AR( $p$ )	$\alpha(L)y_t = \varepsilon_t$	$p > 0$	$q = 0$
MA( $q$ )	$y_t = \beta(L)\varepsilon_t$	$p = 0$	$q > 0$
ARMA( $p, q$ )	$\alpha(L)y_t = \beta(L)\varepsilon_t$	$p > 0$	$q > 0$

### 3. Processus AR(1)

Le processus AR(1) satisfait l'équation différentielle :

$$y_t = \nu + \alpha y_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim WN(0, \sigma^2) \quad (3)$$

En utilisant l'opérateur de retard :

$$(1 - \alpha L)y_t = \nu + \varepsilon_t. \quad (4)$$

Lorsque  $|\alpha| < 1$  :

$$(1 - \alpha L)^{-1} = \lim_{j \rightarrow \infty} (1 + \alpha L + \alpha^2 L^2 + \cdots + \alpha^j L^j) \quad (5)$$

Par conséquent :

$$y_t = (1 + \alpha + \alpha^2 + \cdots) \nu + (\varepsilon_t + \alpha \varepsilon_{t-1} + \alpha^2 \varepsilon_{t-2} + \cdots) \quad (6)$$

En prenant les espérances avec  $|\alpha| < 1$  :

$$E(y_t) = \frac{\nu}{1 - \alpha} = \mu \quad (7)$$

## Analyse de stabilité basée sur une équation différentielle linéaire non homogène :

$$(y_t - \mu) = \alpha(y_{t-1} - \mu) + \varepsilon_t \quad (8)$$

Par conséquent :

$$\mathbb{E}(y_t - \mu)^2 = \alpha^2 \mathbb{E}(y_{t-1} - \mu)^2 + \mathbb{E}(\varepsilon_t^2) + 2\alpha \mathbb{E}[(y_{t-1} - \mu)\varepsilon_t] \quad (9)$$

Sous la condition  $|\alpha| < 1$  :

$$\mathbb{E}(y_t - \mu)^2 = \mathbb{E}(y_{t-1} - \mu)^2 = \gamma(0) \quad (10)$$

de sorte que :

$$\mathbb{V}(y_t) = \gamma(0) = \frac{\sigma^2}{1 - \alpha^2} \quad (11)$$

Fonction d'autocovariance :

$$\mathbb{E}[(y_t - \mu)(y_{t-1} - \mu)] = \alpha \mathbb{E}(y_{t-1} - \mu)^2 + \mathbb{E}[(y_{t-1} - \mu)\varepsilon_t] \quad (12)$$

Par conséquent :

$$\gamma(1) = \alpha\gamma(0). \quad (13)$$

Aussi :

$$\mathbb{E}[(y_t - \mu)(y_{t-2} - \mu)] = \alpha \mathbb{E}[(y_{t-1} - \mu)(y_{t-2} - \mu)] + \mathbb{E}[(y_{t-2} - \mu)\varepsilon_t] \quad (14)$$

pour obtenir :

$$\gamma(2) = \alpha\gamma(1). \quad (15)$$

En général :

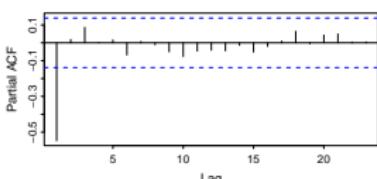
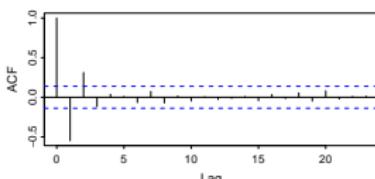
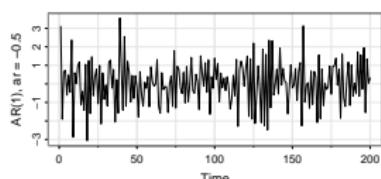
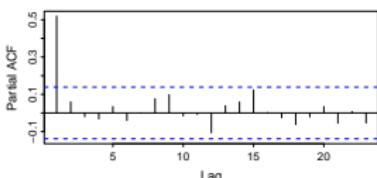
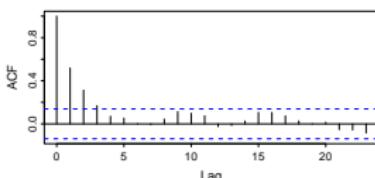
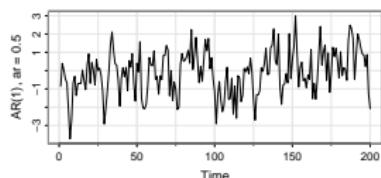
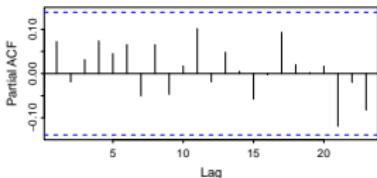
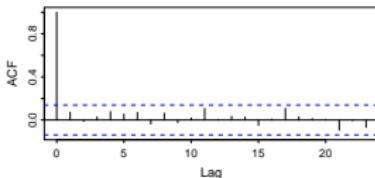
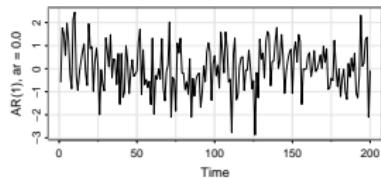
$$\gamma(h) = \alpha\gamma(h-1) = \alpha^h\gamma(0) \quad (16)$$

pour  $h \neq 0$ .

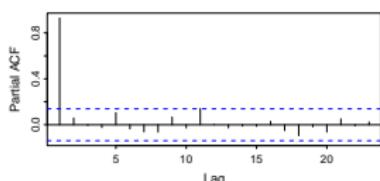
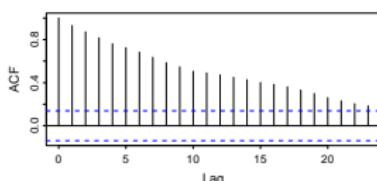
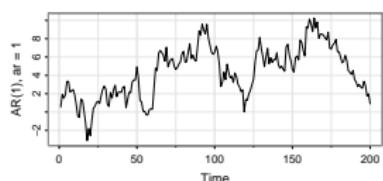
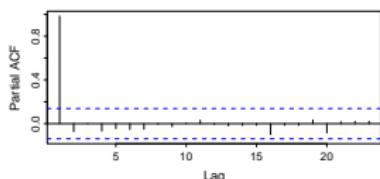
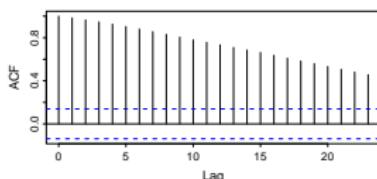
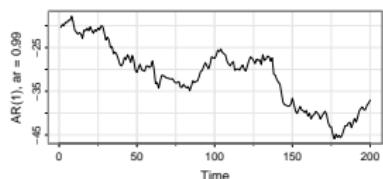
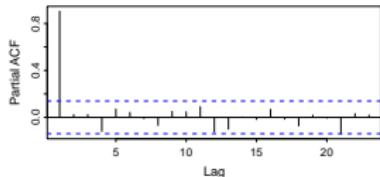
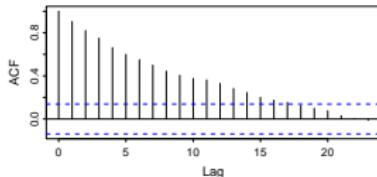
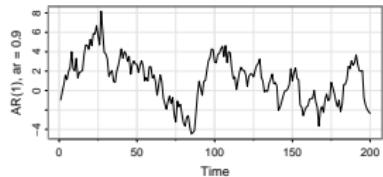
Coefficient d'autocorrélation :

$$\rho(h) = \frac{\gamma(h)}{\gamma(0)} = \frac{\alpha^h \gamma(0)}{\gamma(0)} = \alpha^h \quad (17)$$

## Processus AR(1) stationnaires



## Processus AR(1) persistents et non stationnaires



## 4. Processus AR(2)

$$y_t = \nu + \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \varepsilon_t, \quad \varepsilon_t \sim \text{WN}(0, \sigma^2) \quad (18)$$

En supposant la stationnarité :

$$\mathbb{E}(y_t) = \frac{\nu}{1 - \alpha_1 - \alpha_2} \quad (19)$$

Aussi :

$$\mathbb{V}(y_t) = \gamma(0) = \alpha_1 \gamma(1) + \alpha_2 \gamma(2) + \sigma^2 \quad (20)$$

Et :

$$\gamma(1) = \alpha_1 \gamma(0) + \alpha_2 \gamma(1)$$

$$\gamma(2) = \alpha_1 \gamma(1) + \alpha_2 \gamma(0)$$

En résolvant pour  $\gamma(0)$  :

$$\gamma(0) = \frac{(1 - \alpha_2)\sigma^2}{(1 + \alpha_2)(1 - \alpha_1 - \alpha_2)(1 + \alpha_1 - \alpha_2)} \quad (21)$$

Conditions pour la stationnarité :

$$\alpha_2 + \alpha_1 < 1; \quad \alpha_2 - \alpha_1 < 1; \quad |\alpha_2| < 1 \quad (22)$$

Les racines complexes apparaissent si :

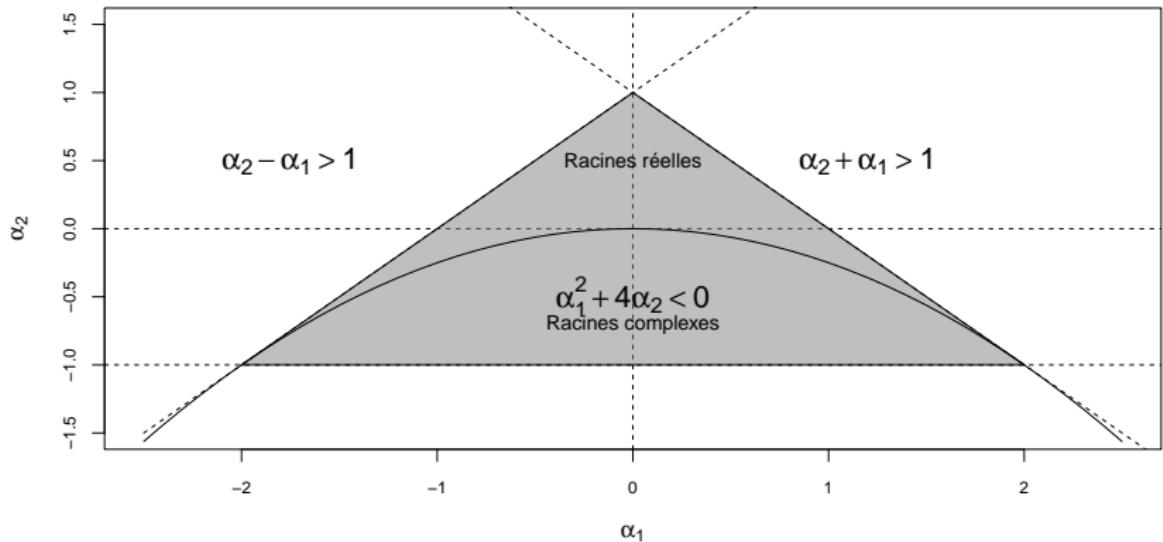
$$\alpha_1^2 + 4\alpha_2 < 0 \quad (23)$$

Équations de Yule-Walker :

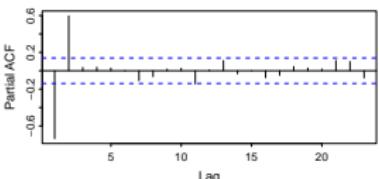
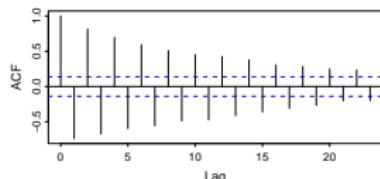
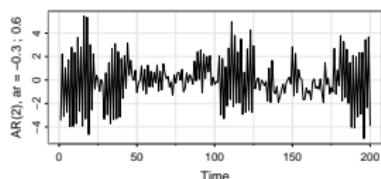
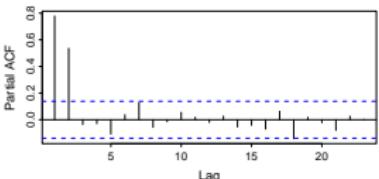
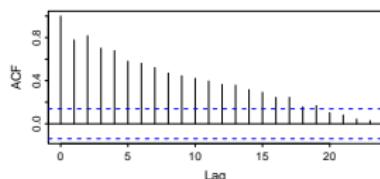
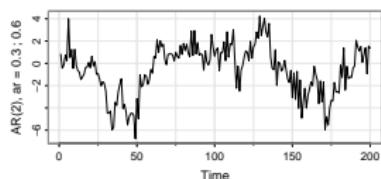
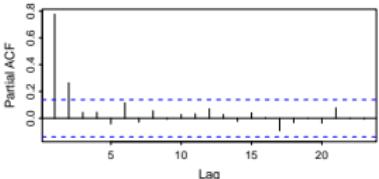
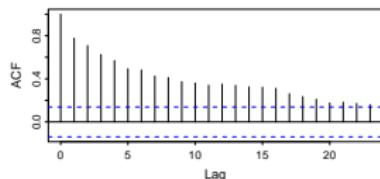
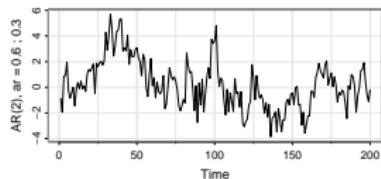
$$\rho(1) = \alpha_1 + \alpha_2 \rho(1)$$

$$\rho(2) = \alpha_1 \rho(1) + \alpha_2$$

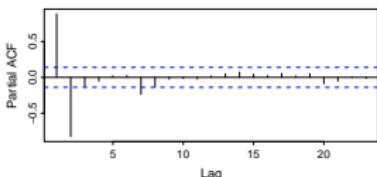
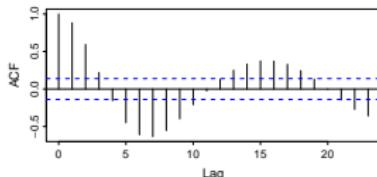
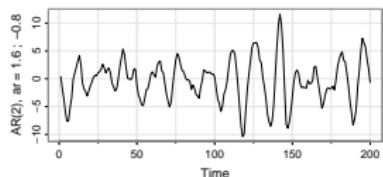
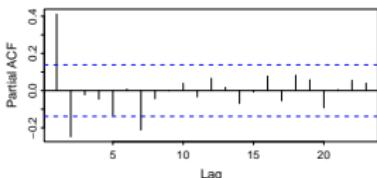
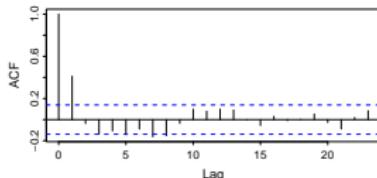
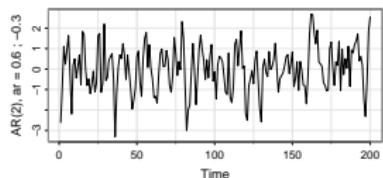
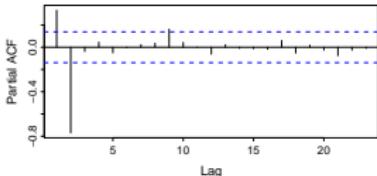
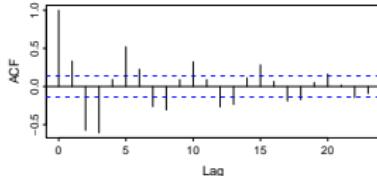
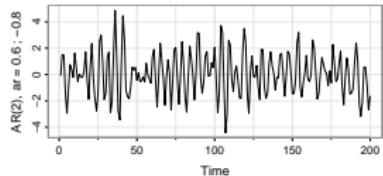
## Triangle de stationnarité avec séparation des racines complexes et réelles



## Processus AR(2) avec racines réelles



## Processus AR(2) avec racines complexes



## 5. Processus AR( $p$ )

$$y_t = \nu + \sum_{j=1}^p \alpha_j y_{t-j} + \varepsilon_t, \quad \varepsilon_t \sim \text{WN}(0, \sigma^2) \quad (1) \quad (24)$$

**Stabilité :**  $\alpha(z) = 0 \Rightarrow |z| > 1$  garantit la stationnarité et la représentation MA( $\infty$ ) :

$$y_t = \alpha(1)^{-1}\nu + \alpha(L)^{-1}\varepsilon_t$$

$$y_t = \mu + \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j}$$

où  $\mu = \frac{\nu}{\alpha(1)}$  et  $\Psi(L) = \alpha(L)^{-1}$  avec  $\sum_{j=0}^{\infty} |\psi_j| < \infty$ .

En prenant les espérances de l'équation (1) :

$$\mathbb{E}(y_t) = \mu = \frac{\nu}{(1 - \alpha_1 - \cdots - \alpha_p)} \quad (25)$$

## Fonction d'autocovariance :

$$\begin{aligned}\gamma(h) &= E(y_t y_{t-h}) = E[(\alpha_1 y_{t-1} + \cdots + \alpha_p y_{t-p} + \varepsilon_t) y_{t-h}] \\ &= \alpha_1 E(y_{t-1} y_{t-h}) + \cdots + \alpha_p E(y_{t-p} y_{t-h}) + E(\varepsilon_t y_{t-h}) \\ &= \alpha_1 \gamma(h-1) + \cdots + \alpha_p \gamma(h-p)\end{aligned}$$

## Équations de Yule-Walker :

$$\rho(1) = \alpha_1 + \alpha_2 \rho(1) + \cdots + \alpha_p \rho(p-1)$$

$$\rho(2) = \alpha_1 \rho(1) + \alpha_2 + \cdots + \alpha_p \rho(p-2)$$

⋮

$$\rho(p) = \alpha_1 \rho(p-1) + \alpha_2 \rho(p-2) + \cdots + \alpha_p$$

pour  $\rho_1, \dots, \rho_p$

$$\rho(k) = \alpha_1 \rho(k-1) + \alpha_2 \rho(k-2) + \cdots + \alpha_p \rho(k-p)$$

pour  $k > p$

## 6. Processus de moyenne mobile

$$y_t = \mu + \beta(L)\varepsilon_t = \mu + \sum_{i=1}^q \beta_i \varepsilon_{t-i} + \varepsilon_t \quad (26)$$

où  $\varepsilon_t \sim WN(0, \sigma^2)$  et  $\beta(L) = 1 + \beta_1 L + \dots + \beta_q L^q$ ,  $\beta_q \neq 0$ .

Le processus MA( $q$ ) est stationnaire pour tout  $(\beta_1, \beta_2, \dots, \beta_q)$ .

**Inversibilité** :  $\beta(z) = 0 \Rightarrow |z| > 1$  garantit la représentation AR( $\infty$ ) :

$$\beta(L)^{-1}y_t = \beta(1)^{-1}\mu + \varepsilon_t$$

$$y_t = \mu + \sum_{j=1}^{\infty} \phi_j (y_{t-j} - \mu) + \varepsilon_t$$

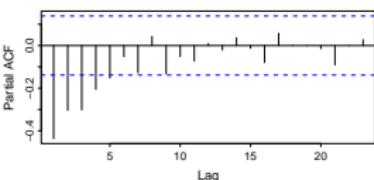
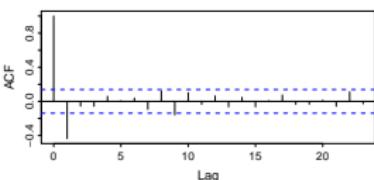
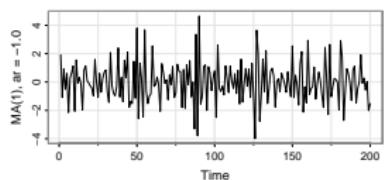
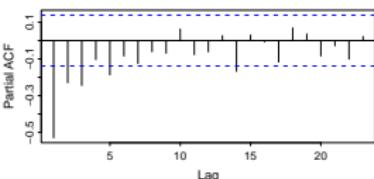
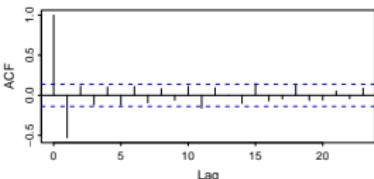
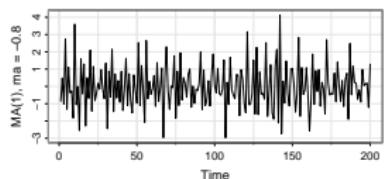
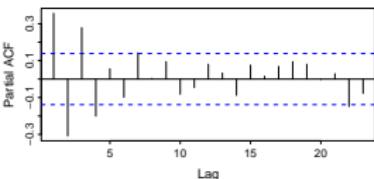
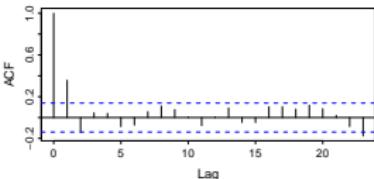
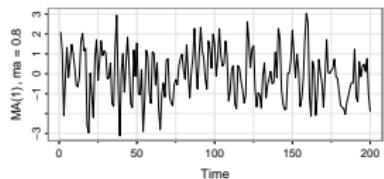
où  $\phi(L) = 1 - \sum_{j=1}^{\infty} \phi_j L^j = 1 - \phi_1 L - \phi_2 L^2 - \dots = \beta(L)^{-1}$ .

Fonction d'autocovariance :

$$\gamma(0) = \left( \sum_{i=0}^q \beta_i^2 \right) \sigma^2$$
$$\gamma(k) = \left( \sum_{i=0}^{q-k} \beta_i \beta_{i+k} \right) \sigma^2 \quad \text{pour } k = 1, 2, \dots, q$$
$$\gamma(k) = 0 \quad \text{pour } k > q$$

**Les processus MA sont stationnaires et ergodiques.**

## Processus MA(1)



## 7. Processus ARMA( $p, q$ )

$$\alpha(L)y_t = \nu + \beta(L)\varepsilon_t$$

$$y_t = \nu + \sum_{j=1}^p \alpha_j y_{t-j} + \varepsilon_t + \sum_{i=1}^q \beta_i \varepsilon_{t-i}$$

où  $\varepsilon_t \sim \text{WN}(0, \sigma^2)$  ;

$$\alpha(L) = 1 - \alpha_1 L - \cdots - \alpha_p L^p, \quad \alpha_p \neq 0 ;$$

$$\beta(L) = 1 + \beta_1 L + \cdots + \beta_q L^q, \quad \beta_q \neq 0.$$

### 1. Stabilité

$\alpha(z) = 0 \Rightarrow |z| > 1$  garantit la stationnarité et la représentation MA( $\infty$ ) :

$$y_t = \alpha(1)^{-1}\nu + \alpha(L)^{-1}\beta(L)\varepsilon_t$$

$$y_t = \mu + \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j}$$

## 2. Inversibilité

$\beta(z) = 0 \Rightarrow |z| > 1$  permet la représentation AR( $\infty$ ) :

$$\begin{aligned}\beta(L)^{-1} \alpha(L)(y_t - \mu) &= \varepsilon_t \\ y_t &= \mu + \sum_{j=1}^{\infty} \phi_j(y_{t-j} - \mu) + \varepsilon_t\end{aligned}$$

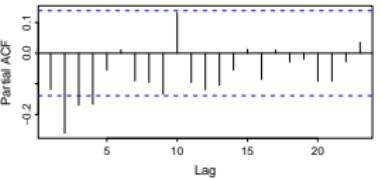
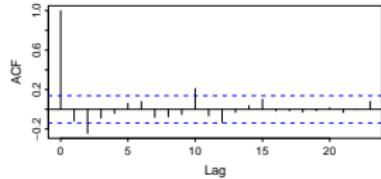
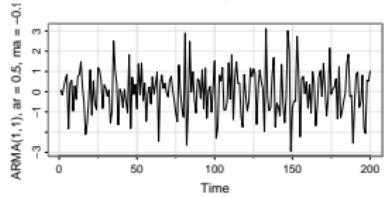
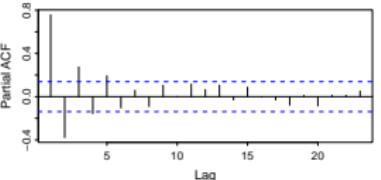
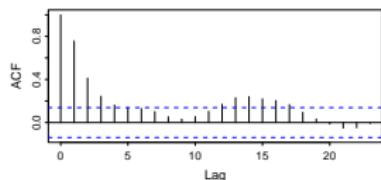
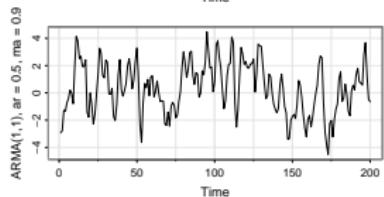
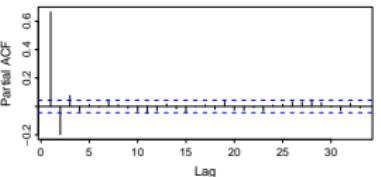
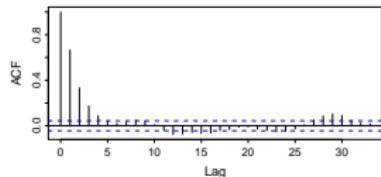
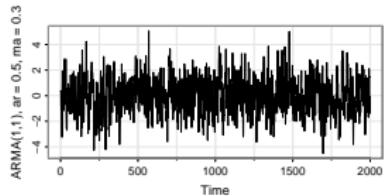
## 3. Pas de racines communes dans $\alpha(L)$ et $\beta(L)$

$$\alpha(L) = \prod_{j=1}^p (1 - \lambda_j L)$$

$$\beta(L) = \prod_{i=1}^q (1 - \mu_i L)$$

$$\Rightarrow \lambda_j \neq \mu_i \quad \forall i, j$$

## Processus ARMA(1,1)



## 8. Formulation statistique du modèle AR(1)

**Indépendance conditionnelle** :  $(y_t|y_0, \dots, y_{t-1}) \stackrel{d}{=} (y_t|y_{t-1})$  ;

**Distribution conditionnelle** :  $(y_t|y_{t-1}) \stackrel{d}{=} \mathcal{N}(\nu + \alpha y_{t-1}, \sigma^2)$  ;  $t \geq 1$  ;

**Espace des paramètres** :  $\nu, \alpha, \sigma^2 \in \mathbb{R} \times \mathbb{R} \times \mathbb{R}^+$ .

L'observation initiale  $y_0$  n'est pas modélisée – conditionnelle à  $y_0$ .

Le régresseur est la variable dépendante retardée.

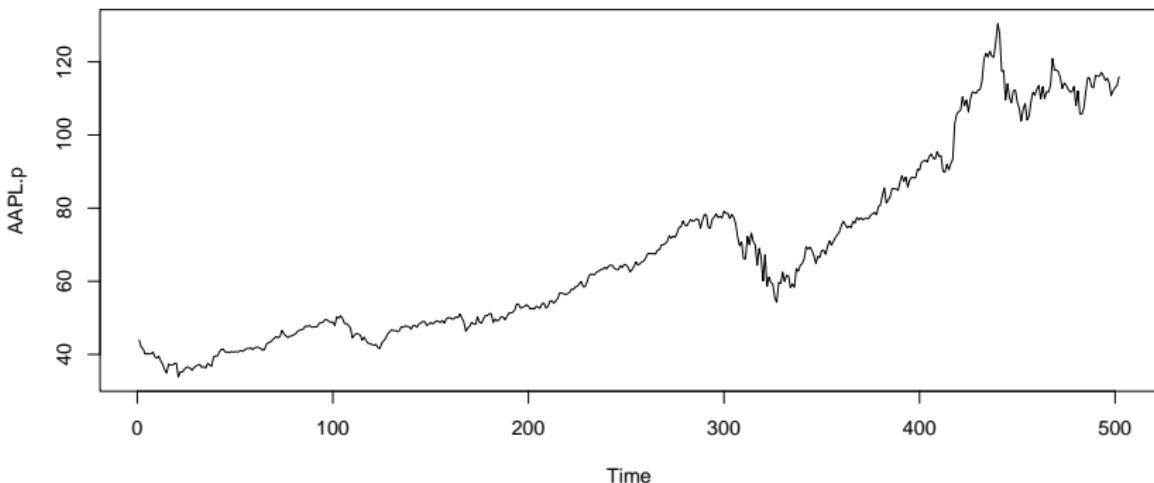
$$y_t = \nu + \alpha y_{t-1} + \varepsilon_t \quad \text{pour } t = 1, \dots, T.$$

où  $\varepsilon_t \sim \text{NID}(0, \sigma^2)$  sont des innovations.

## Dans la pratique : modélisation - données

$y_t$  sont les rendements de l'action Apple sur trois années autour de la période covid.

```
AAPL <- getSymbols(Symbols = 'AAPL', src = 'yahoo',
                     from='2018-12-01', to='2020-12-01', auto.assign = FALSE)
AAPL$date = as.Date.character(AAPL$date, format = "%Y/%m/%d") #changement
# du format dans la dataframe
AAPL.date <- AAPL$date #création de la variable date (non utilisée ici)
AAPL.p <- AAPL$AAPL.Adjusted #création de la variable prix
ts.plot(AAPL.p)
```



```
AAPL.r = na.omit(diff(log(AAPL.p))*100) #création de la variable rendements
# stationnaires
summary(cbind(AAPL.p[-1],AAPL.r))
```

## Dans la pratique : modélisation - stationnarité, méthode *a la {tidyverse}*

Test KPSS sur les prix -  $H_0$  : données stationnaires

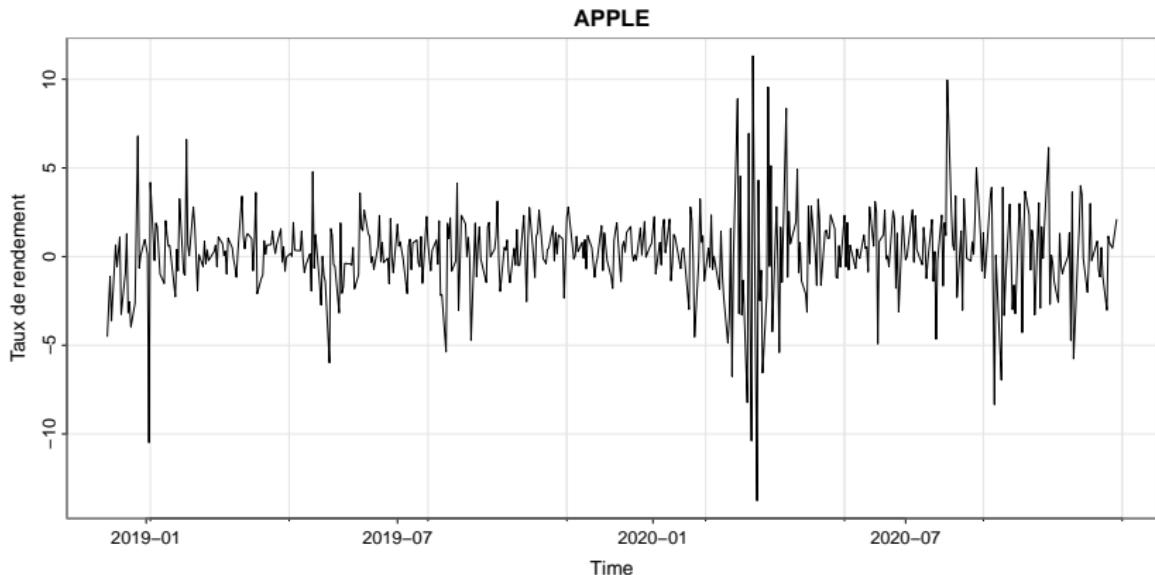
```
## $test_stat  
## [1] 7.460074  
##  
## $critical_values  
##          10pct  5pct 2.5pct 1pct  
## critical values 0.347 0.463  0.574 0.739
```

Test KPSS sur les rendements -  $H_0$  : données stationnaires

```
## $test_stat  
## [1] 0.1184969  
##  
## $critical_values  
##          10pct  5pct 2.5pct 1pct  
## critical values 0.347 0.463  0.574 0.739
```

## Dans la pratique : modélisation - graphique

```
tsplot(index(AAPL.r), AAPL.r, main = "APPLE",  
       ylab = "Taux de rendement")
```



## 9. Vraisemblance autorégressive

**Densité conditionnelle :**

$$f(y_T, y_{T-1}, \dots, y_1 | y_0) = f(y_T | y_{T-1}, \dots, y_1, y_0) \times f(y_{T-1}, \dots, y_1 | y_0)$$

$$= f(y_T | y_{T-1}, \dots, y_1, y_0) \times f(y_{T-1} | y_{T-2}, \dots, y_1, y_0) \times f(y_{T-2}, \dots, y_1 | y_0)$$

$$= \dots$$

$$= \prod_{t=1}^T f(y_t | y_{t-1}, \dots, y_1, y_0)$$

Cette formule est généralement valable - décomposition de prédition.

En utilisant la propriété de Markov :

$$f_{\nu, \alpha, \sigma^2}(y_T, \dots, y_1 | y_0) = \prod_{t=1}^T f_{\nu, \alpha, \sigma^2}(y_t | y_{t-1})$$

## Vraisemblance autorégressive (suite)

Étant donné l'hypothèse de normalité conditionnelle :

$$f_{\nu, \alpha, \sigma^2}(y_T, \dots, y_1 | y_0) = \prod_{t=1}^T (2\pi\sigma^2)^{-1/2} \exp\left(-\frac{1}{2\sigma^2}(y_t - \nu - \alpha y_{t-1})^2\right)$$
$$= (2\pi\sigma^2)^{-T/2} \exp\left(-\frac{1}{2\sigma^2} \sum_{t=1}^T (y_t - \nu - \alpha y_{t-1})^2\right)$$

Et la fonction de vraisemblance est :

$$L(y_1, \dots, y_T | y_0; \nu, \alpha, \sigma^2) = (2\pi\sigma^2)^{-T/2} \exp\left(-\frac{1}{2\sigma^2} \sum_{t=1}^T (y_t - \nu - \alpha y_{t-1})^2\right)$$

La vraisemblance exacte inclut la condition initiale. La vraisemblance conditionnelle est conditionnée par  $y_0$ .

## MLE

L'estimation par maximum de vraisemblance (MLE) de  $(\nu, \alpha)$  est obtenue en minimisant la somme des carrés des résidus :

$$\arg \max_{(\nu, \alpha)} I(\nu, \alpha) = \arg \min_{(\nu, \alpha)} \sum_{t=1}^T \varepsilon_t^2(\nu, \alpha)$$

La MLE de  $(\tilde{\nu}, \tilde{\alpha})$  est équivalente à l'estimation par les moindres carrés ordinaires (OLS) de  $(\hat{\nu}, \hat{\alpha})$ .

$(\tilde{\nu}, \tilde{\alpha})$  sont des estimateurs convergents si  $y_t$  est stationnaire et  $\sqrt{T}(\tilde{\delta} - \delta)$  est asymptotiquement normalement distribué, où  $\delta = (\nu, \alpha)'$ .

Les MLE basées sur la vraisemblance exacte et la vraisemblance conditionnelle sont asymptotiquement équivalentes.

L'estimation par la méthode des moments (en utilisant les équations de Yule-Walker) est également équivalente pour  $\alpha$ .

## Dans la pratique : estimation - {stats}

On suppose une distribution conditionnelle gaussienne des  $\varepsilon_t$ .

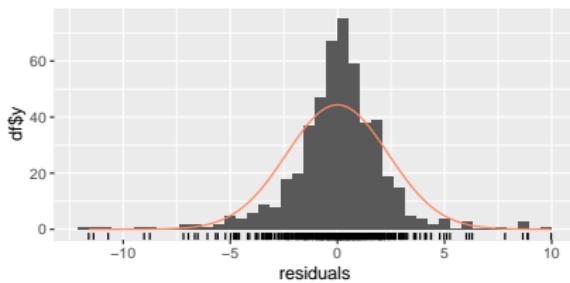
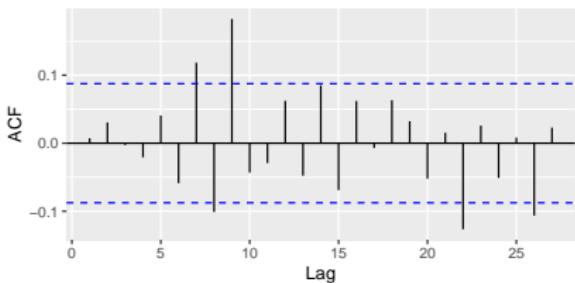
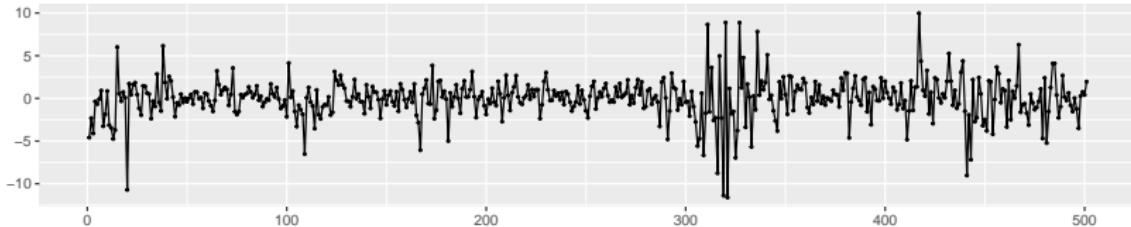
```
AAPL.estim = arima(AAPL.r, order = c(1, 0, 0))
coeftest(AAPL.estim)

##
## z test of coefficients:
##
##           Estimate Std. Error z value Pr(>|z|)
## ar1      -0.211688   0.043821 -4.8307 1.361e-06 ***
## intercept 0.194581   0.087407  2.2262     0.026 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

## Dans la pratique : estimation - {forecast}

```
fit = Arima(AAPL.r, order = c(1, 0, 0))
checkresiduals(fit)
```

Residuals from ARIMA(1,0,0) with non-zero mean



```
##
## Ljung-Box test
##
## data: Residuals from ARIMA(1,0,0) with non-zero mean
## Q* = 33.733, df = 9, p-value = 9.947e-05
##
## Model df: 1. Total lags used: 10
```

## Dans la pratique : estimation - autoarima{forecast}

```
fit = auto.arima(AAPL.r)
fit

## Series: AAPL.r
## ARIMA(4,0,1) with non-zero mean
##
## Coefficients:
##      ar1      ar2      ar3      ar4      ma1      mean
##     -1.0701  -0.1448  0.0191  -0.0724  0.8901  0.1948
## s.e.   0.0859   0.0676  0.0662   0.0504  0.0748  0.0870
##
## sigma^2 = 5.517: log likelihood = -1135.81
## AIC=2285.62  AICc=2285.85  BIC=2315.14
```

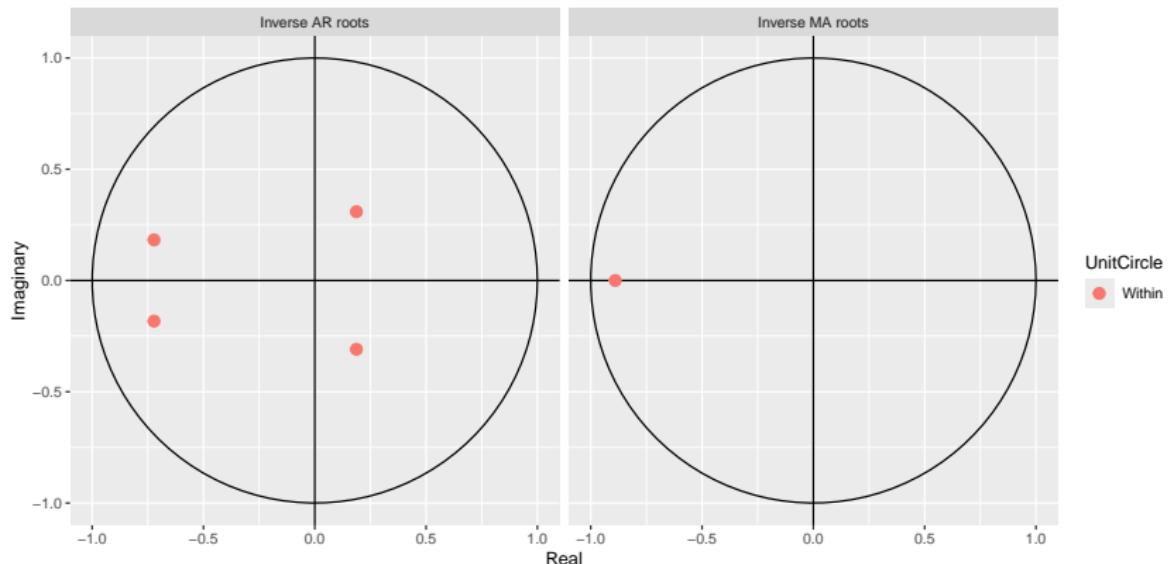
Dans la pratique : estimation - autoarima{forecast}, pvalues

```
coeftest(fit)

##
## z test of coefficients:
##
##             Estimate Std. Error   z value Pr(>|z|)
## ar1      -1.070091  0.085929 -12.4532 < 2e-16 ***
## ar2      -0.144838  0.067607  -2.1423  0.03217 *
## ar3       0.019084  0.066192   0.2883  0.77311
## ar4      -0.072443  0.050358  -1.4385  0.15028
## ma1       0.890073  0.074819  11.8963 < 2e-16 ***
## intercept 0.194787  0.086957   2.2400  0.02509 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Dans la pratique : estimation - autoarima{forecast}, racines unitaires

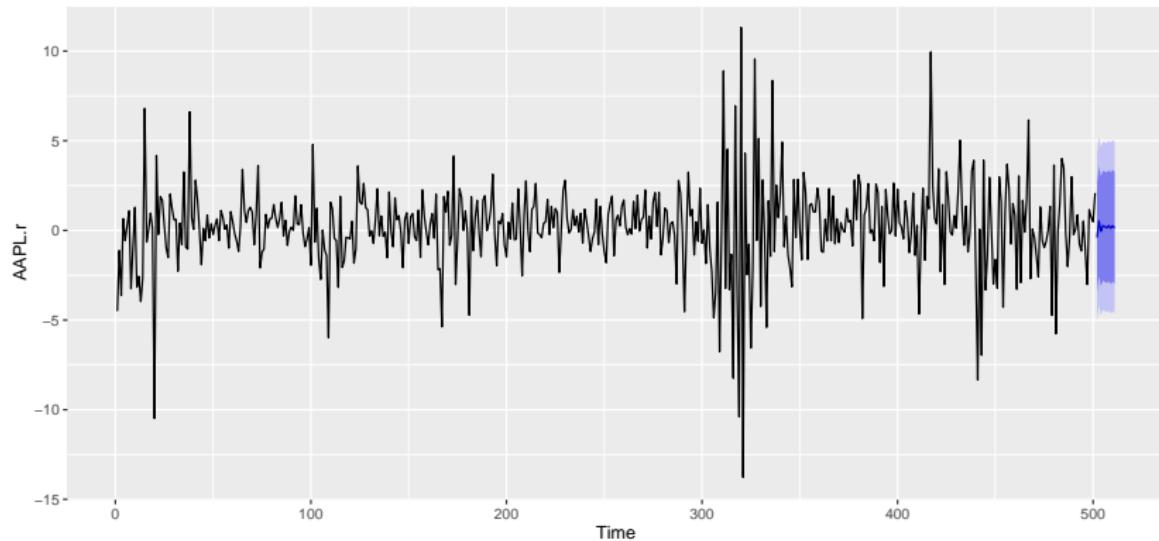
```
autoplot(fit)
```



## Dans la pratique : forecast - {forecast}

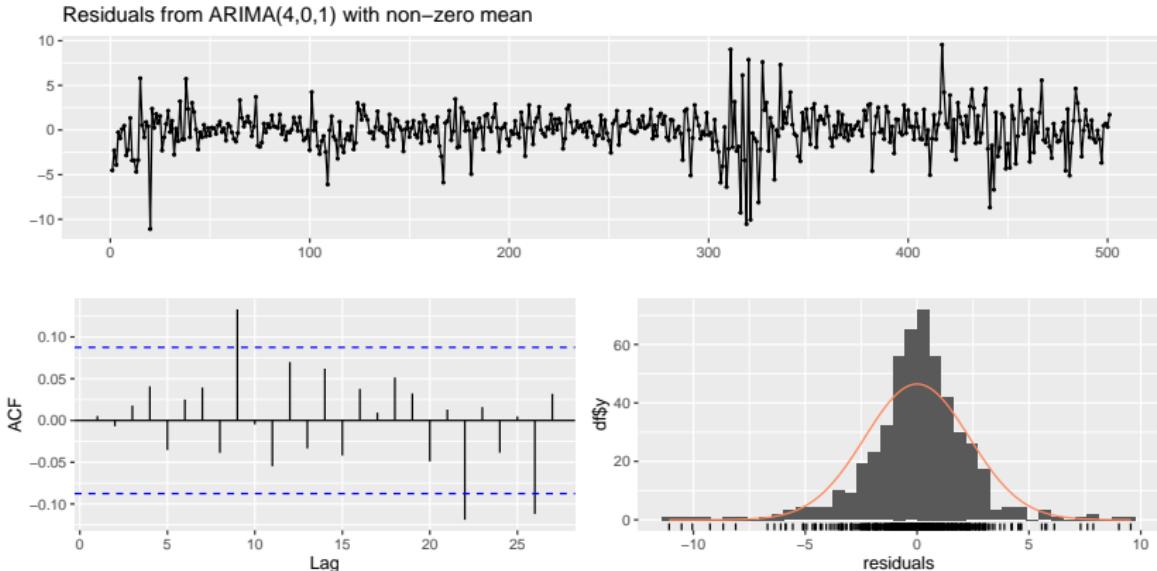
```
autoplot(forecast(fit))
```

Forecasts from ARIMA(4,0,1) with non-zero mean



## Dans la pratique : residuals check - autoarima{forecast}

```
checkresiduals(fit)
```



```
##  
## Ljung-Box test  
##  
## data: Residuals from ARIMA(4,0,1) with non-zero mean  
## Q* = 12.629, df = 5, p-value = 0.02711  
##  
## Model df: 5. Total lags used: 10
```

## AIC et BIC

Pour effectuer une sélection de modèles avec des critères d'information.

AIC est la méthode de sélection de modèles la plus connue et utilisée.

**AIC** Akaike (1973) :

$$AIC = -2 \times LogLik + 2 \times p$$

**BIC** Schwarz (1978) :

$$BIC = -2 \times LogLik + \log(n) \times p$$

avec  $n$  le nombre d'observations dans l'échantillon étudié et  $p$  le nombre de paramètres.

Plus petit est AIC/BIC, meilleur est le modèle.

```
AIC(AAPL.estim) #AR(1)
```

```
## [1] 2292.332
```

```
AIC(fit)          #ARMA(4,1)
```

```
## [1] 2285.623
```

## 10. Analyse de la mauvaise spécification et des résidus :

### Tests de mauvaise spécification et Tests de diagnostic

Les hypothèses faites lors de la construction de modèles économétriques.

Exemple – Modèle AR(1) :

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

Le modèle statistique conditionnel de  $y_1, y_2, \dots, y_T$  étant donné  $y_0$  est défini par les hypothèses suivantes :

1. **Indépendance** :  $\varepsilon_1, \dots, \varepsilon_T$  sont indépendants ;
2. **Normalité conditionnelle** :  $\varepsilon_t \stackrel{D}{=} N(0, \sigma^2)$  ;
3. **Espace des paramètres** :  $\nu, \alpha, \sigma^2 \in \mathbb{R} \times \mathbb{R} \times \mathbb{R}_+$ .

Il est important de tester si ces hypothèses sont valides.

## Normalité

Tester si l'asymétrie et l'aplatissement correspondent à une distribution normale.

Soit  $x_t \sim D(\mu, \sigma^2)$  tel que :

$$y_t = \frac{x_t - \mu}{\sigma} \sim D(0, 1)$$

Définir :

$$\kappa_3 = E(y_t^3) \quad (\text{asymétrie})$$

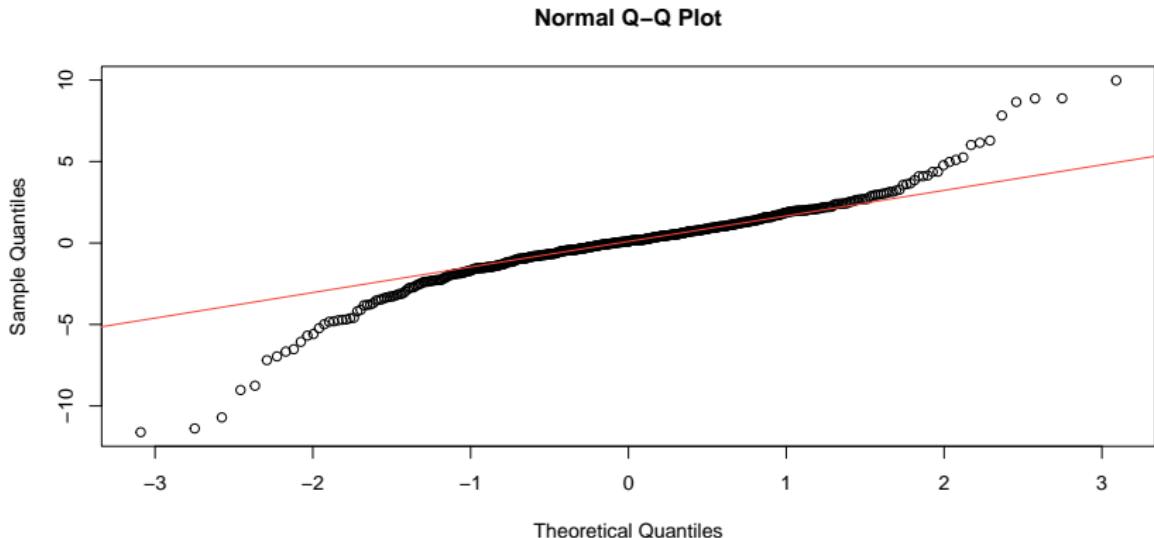
$$\kappa_4 = E(y_t^4) - 3 \quad (\text{excès d'aplatissement})$$

Le test de **Jarque-Bera** -  $H_0 : \kappa_3 = \kappa_4 = 0$  pour la normalité.

$$JB = \frac{T-k}{6} \left( \kappa_3^2 + \frac{\kappa_4^2}{4} \right) \sim \chi^2_2 \text{ et } k \text{ le nombre de variables explicatives.}$$

## Dans la pratique : normalité des résidus

```
AAPL.resid = AAPL.estim$residuals  
qqnorm(AAPL.resid)#, ylim = c(-10,10), xlim = c(-10,10))  
qqline(AAPL.resid, col=2)
```



```
jarque.bera.test(AAPL.resid)
```

```
##  
## Jarque Bera Test  
##  
## data: AAPL.resid  
## X-squared = 527.32, df = 2, p-value < 2.2e-16
```

## Hétéroscédasticité

Hétéroscédasticité - variance non constante

**Test de White (1980)** : La variance  $V(\varepsilon_t | Y_{t-1})$  varie-t-elle avec  $Y_{t-1}$  ?

1. Obtenir les résidus  $\widehat{\varepsilon}_t$  du modèle original.
2. Obtenir  $R_{het}^2$  à partir de la régression auxiliaire :

$$\widehat{\varepsilon}_t^2 = \beta_1 + \beta_2 y_{t-1} + \beta_3 y_{t-1}^2 + \eta_t$$

3. Tester  $\beta_2 = \beta_3 = 0$  en utilisant :

$$TR_{het}^2 \approx F_{het} = \frac{R_{het}^2 / m}{(1 - R_{het}^2) / (T - k - m - 1)}$$

où  $k$  est le nombre de régresseurs et  $m$  est le nombre de termes quadratiques et de produits croisés.

Notez que le  $TR_{het}^2 \sim \chi^2_2$

## Dans la pratique : Test de White - hétéroscédasticité

```
# Étape 1 : Extraire les résidus
residuals <- resid(AAPL.estim)
# Étape 2 : Construire les termes quadratiques et croisés
residuals_squared <- residuals^2
x1 = lag(AAPL.r)
x2 = lag(AAPL.r)^2
auxiliary_data <- data.frame(x1, x2)
# Supprimer les valeurs NA introduites par le décalage
auxiliary_data <- na.omit(auxiliary_data)
# Étape 3 : Régression auxiliaire pour le test de White
auxiliary_model <- lm(residuals_squared ~ x1 + x2, data = auxiliary_data)
# Étape 4 : Test de significativité globale (statistique de White)
white_stat <- summary(auxiliary_model)$r.squared * nrow(auxiliary_data) # Stat
p_value <- 1 - pchisq(white_stat, df = 2) # Degré de liberté = nb de variables
# Résultat : H0 - hétéroscédasticité
cat("Statistique de White :", white_stat, "\n")

## Statistique de White : 22.61791
cat("P-valeur :", p_value, "\n")

## P-valeur : 1.22626e-05
```

## Forme fonctionnelle

La forme fonctionnelle log-linéaire est souvent supposée.

**Test RESET de Ramsey (1969)** : Erreur de spécification de la régression.

Les polynômes de la variable prédictive  $\hat{y}_t$  aident-ils à expliquer  $y_t$  ? Sous l'hypothèse nulle, les variables de la forme fonctionnelle correcte sont irrelevantes.

1. Obtenir les variables prédictives  $\hat{y}_t$  du modèle original.
2. Obtenir la corrélation partielle  $R_{reset}^2$  de  $y_t$  et  $\hat{y}_t^2$  étant donné  $y_{t-1}$  à partir de la régression auxiliaire :  
$$y_t = \beta_1 + \beta_2 y_{t-1} + \beta_3 \hat{y}_t^2 + \eta_t$$
3. Tester  $\beta_3 = 0$  en utilisant :

$$TR_{reset}^2 \approx F_{reset} = \frac{R_{reset}^2 / m}{(1 - R_{reset}^2) / (T - k - m - 1)}$$

où  $k$  est le nombre de régresseurs et  $m$  est le nombre de restrictions.

## Dans la pratique : Test de Ramsey - erreur de spécification

```
# Étape 1 : Extraire la série ajustée
y_hat <- fitted(AAPL.estim)
# Étape 2 : Construire la dataframe
y1 = ts(lag((AAPL.r)))
y_hat2 = y_hat^2
aux_data <- data.frame(y1, y_hat2)
# Étape 3 : Régression auxiliaire pour le test de Ramsey
aux_model <- lm(AAPL.r ~ y1 + y_hat2, data = aux_data)
# Étape 4 : Test de significativité globale (statistique de Ramsey)
ramsey_stat <- summary(aux_model)$r.squared * nrow(aux_data) # Statistique du
p_value <- 1 - pchisq(ramsey_stat, df = 1) # Degré de liberté = nb de variables
# Résultat : H0 - erreur de spécification
cat("Statistique de Ramsey :", ramsey_stat, "\n")

## Statistique de Ramsey : 22.34384
cat("P-valeur :", p_value, "\n")

## P-valeur : 2.279428e-06
```

## Erreurs autocorrélées

L'indépendance des innovations dans un AR(1) implique  $H_0 : \alpha_2 = 0$  dans :

$$y_t = \nu + \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \varepsilon_t$$

où  $\varepsilon_t \sim \text{NID}(0, \sigma^2)$ .

Test du rapport de vraisemblance pour  $\alpha_2 = 0$  :

$$LR = -T \log (1 - r_{02.1}^2)$$

où  $r_{02.1}^2$  est le coefficient de détermination partiel entre  $y_t$  et  $y_{t-2}$  conditionnellement à  $y_{t-1}$ .

Test LM : régression auxiliaire – obtenir  $R_{ar}^2$  à partir de :

$$\widehat{\varepsilon}_t = \beta_1 + \beta_2 y_{t-1} + \beta_3 \widehat{\varepsilon}_{t-1} + \eta_t$$

Tester  $\beta_2 = \beta_3 = 0$  en utilisant :

$$TR_{ar}^2 \approx F_{ar} = \frac{R_{ar}^2 / m}{(1 - R_{ar}^2) / (T - k - m - 1)}$$

où  $k$  est le nombre de régresseurs et  $m$  est le nombre de restrictions.

## Autocorrélation

### Box-Pierce et Ljung-Box

$H_0$  : absence d'autocorrélation

$H_1$  : autocorrélation

Idée principale : On estime la fonction d'autocorrélation et on teste si elle est nulle jusqu'à un certain nombre de retards  $m$ .

$$H_0 : \rho(1) = \rho(2) = \cdots = \rho(m) = 0$$

La statistique du test Ljung-Box :

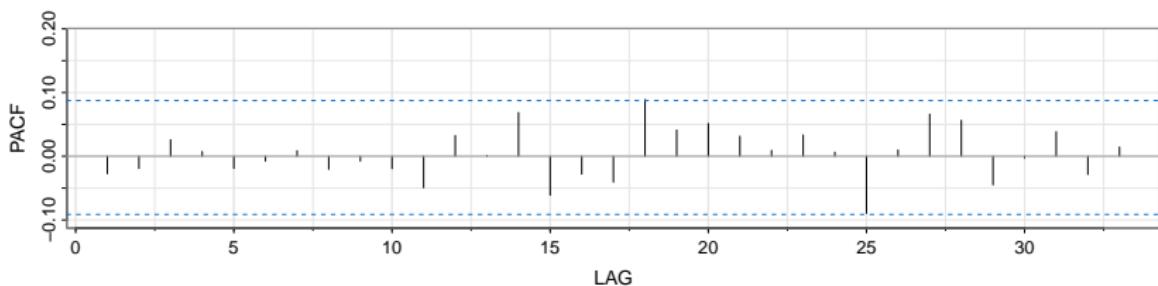
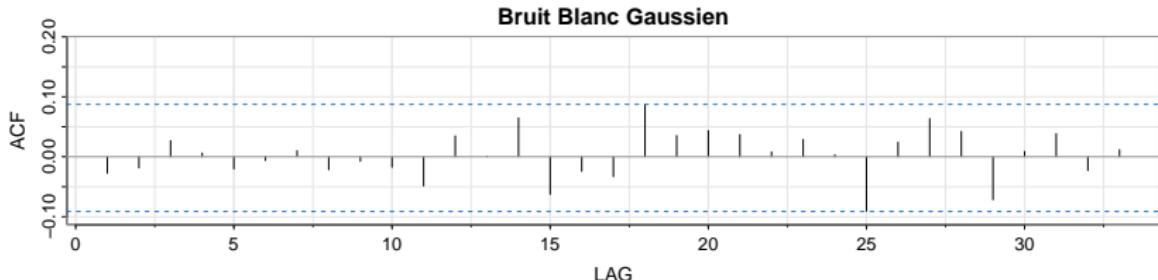
$$Q(m) = T(T+2) \sum_{h=1}^m \frac{\widehat{\rho}(h)^2}{T-h}$$

$$\text{avec } \widehat{\rho}(h) = \frac{\sum_{t=h+1}^T (\varepsilon_t - \bar{\varepsilon})(\varepsilon_{t-h} - \bar{\varepsilon})}{(\varepsilon_t - \bar{\varepsilon})^2}$$

On rejette  $H_0$  si  $Q(m) > q_\alpha$  avec  $q_\alpha$  le  $100(1-\alpha)$  quantile d'une  $\chi^2_{m-p-q}$

## Dans la pratique : ACF et PACF d'un bruit blanc gaussien

```
z <- rnorm(501)
AAPL_ACF_PACF=acf2(z, # 501 observations : comme les résidus
                     main = "Bruit Blanc Gaussien")
```



Dans la pratique : Test de Ljung-Box - absence d'autocorrélation sur GWN

```
Box.test(z, lag = 1, type = "Ljung")
```

```
##  
##  Box-Ljung test  
##  
## data: z  
## X-squared = 0.37993, df = 1, p-value = 0.5376
```

```
Box.test(z, lag = 6, type = "Ljung")
```

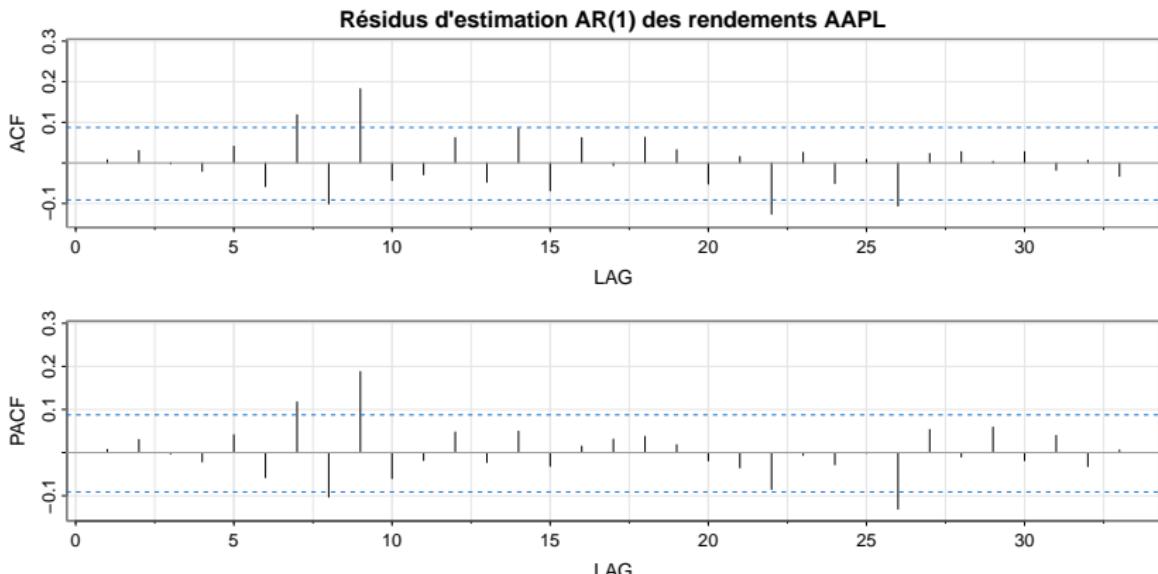
```
##  
##  Box-Ljung test  
##  
## data: z  
## X-squared = 1.1703, df = 6, p-value = 0.9783
```

```
Box.test(z, lag = 12, type = "Ljung")
```

```
##  
##  Box-Ljung test  
##  
## data: z  
## X-squared = 3.4908, df = 12, p-value = 0.991
```

## Dans la pratique : ACF et PACF des résidus

```
AAPL_ACF_PACF=acf2(resid(AAPL.estim),  
main = "Résidus d'estimation AR(1) des rendements AAPL")
```



## Dans la pratique : Test de Ljung-Box - absence d'autocorrélation

```
Box.test(resid(AAPL.estim), lag = 1, type = "Ljung")
```

```
##  
## Box-Ljung test  
##  
## data: resid(AAPL.estim)  
## X-squared = 0.027691, df = 1, p-value = 0.8678
```

```
Box.test(resid(AAPL.estim), lag = 6, type = "Ljung")
```

```
##  
## Box-Ljung test  
##  
## data: resid(AAPL.estim)  
## X-squared = 3.3109, df = 6, p-value = 0.7689
```

```
Box.test(resid(AAPL.estim), lag = 12, type = "Ljung")
```

```
##  
## Box-Ljung test  
##  
## data: resid(AAPL.estim)  
## X-squared = 36.169, df = 12, p-value = 0.0003043
```

## Dans la pratique : Test de Ljung-Box - absence d'autocorrélation V2

```
Box.test(resid(AAPL.estim), lag = 20, type = "Ljung")
```

```
##  
## Box-Ljung test  
##  
## data: resid(AAPL.estim)  
## X-squared = 49.636, df = 20, p-value = 0.0002497
```

```
Box.test(resid(AAPL.estim), lag = 250, type = "Ljung")
```

```
##  
## Box-Ljung test  
##  
## data: resid(AAPL.estim)  
## X-squared = 275.07, df = 250, p-value = 0.1324
```

```
Box.test(resid(AAPL.estim), lag = 8, type = "Ljung")
```

```
##  
## Box-Ljung test  
##  
## data: resid(AAPL.estim)  
## X-squared = 15.684, df = 8, p-value = 0.04714
```

## Dans la pratique : Test de Box-Pierce - absence d'autocorrélation

```
Box.test(resid(AAPL.estim), lag = 1, type = "Box")
```

```
##  
##  Box-Pierce test  
##  
## data:  resid(AAPL.estim)  
## X-squared = 0.027526, df = 1, p-value = 0.8682  
Box.test(resid(AAPL.estim), lag = 6, type = "Box")
```

```
##  
##  Box-Pierce test  
##  
## data:  resid(AAPL.estim)  
## X-squared = 3.2648, df = 6, p-value = 0.775  
Box.test(resid(AAPL.estim), lag = 12, type = "Box")
```

```
##  
##  Box-Pierce test  
##  
## data:  resid(AAPL.estim)  
## X-squared = 35.428, df = 12, p-value = 0.0004002
```

## Erreurs ARCH

**Engle (1982)** : Hétéroscédasticité conditionnelle autorégressive

ARCH - les variances des innovations sont une fonction du temps :

$$V(\varepsilon_t | \varepsilon_{t-1}) = \beta_1 + \beta_2 \varepsilon_{t-1}^2$$

1. Obtenir les résidus  $\hat{\varepsilon}_t$  du modèle original.
2. Obtenir  $R_{arch}^2$  à partir de la régression auxiliaire :

$$\hat{\varepsilon}_t^2 = \beta_1 + \beta_2 \hat{\varepsilon}_{t-1}^2 + \eta_t$$

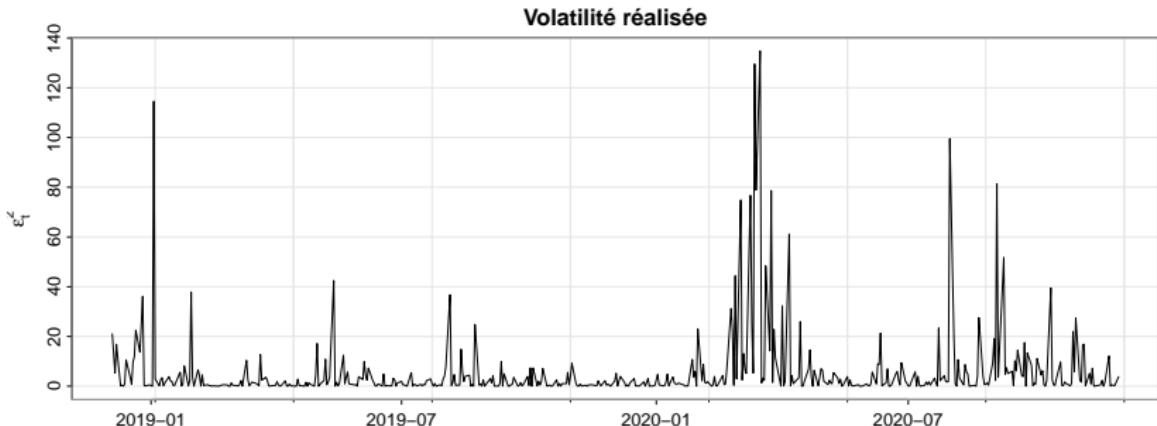
3. Tester  $H_0 : \beta_2 = 0$  en utilisant :

$$TR_{arch}^2 \approx F_{arch} = \frac{R_{arch}^2 / m}{(1 - R_{arch}^2) / (T - k - m - 1)}$$

où  $k$  est le nombre de régresseurs et  $m$  est le nombre de restrictions.

## Dans la pratique : Test d'Engle - ARCH

```
# L'un des proxies de la volatilité et l'élévation de la série au carré :  
tsplot(index(AAPL.r), resid(AAPL.estim)^2, main = "Volatilité réalisée",  
       ylab = expression(epsilon[t]^2), xlab = '')
```



```
# Le test statistique, quant à lui, existe dans le package {FinTS}  
ArchTest(resid(AAPL.estim), lags = 1)
```

```
##  
##  ARCH LM-test; Null hypothesis: no ARCH effects  
##  
##  data:  resid(AAPL.estim)  
##  Chi-squared = 20.499, df = 1, p-value = 5.966e-06
```

## 11. Application

### Taux de chômage au USA (1948-2024)

Le modèle est donné par :

$$x_t = \nu + \alpha x_{t-1} + \varepsilon_t$$

Les résidus  $\{\varepsilon_t\}$  sont obtenus à partir de  $\varepsilon_t = x_t - \nu - \alpha x_{t-1}$ .

Les résultats sont les suivants :

$$x_t = 5.5867 + 0.9704x_{t-1} + \varepsilon_t$$

(0.0000) (0.0000)

(pvalues)

## Les données

```
USunrate = fredr(series_id = "UNRATE",
                  observation_start = as.Date("1948-01-01"),
                  observation_end = as.Date("2024-10-01"),
                  frequency = "m")
x = USunrate$value
xDate = USunrate$date

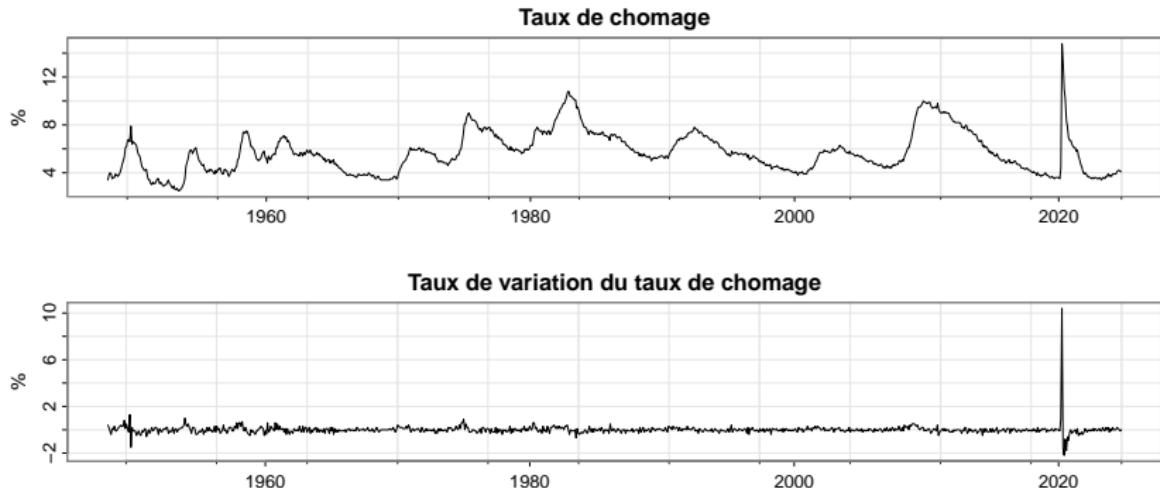
# Tester la stationarité
adf_test = adf.test(x)
print(adf_test)

## 
##  Augmented Dickey-Fuller Test
## 
## data: x
## Dickey-Fuller = -3.9534, Lag order = 9, p-value = 0.01132
## alternative hypothesis: stationary
#kpss.test(x)

# Si la série n'est pas stationnaire, appliquer une différenciation
# x = diff(x)
```

## Graph de la série

```
par(mfrow=c(2,1))
tsplot(USunrate$date, USunrate$value,
main="Taux de chomage", ylab = "%", xlab=' ')
# Si on avait dû différencier la série :
tsplot(xDate[-1],diff(USunrate$value),
main= "Taux de variation du taux de chomage", ylab = "%", xlab=' ')
```



## Estimation d'un AR(1)

```
fit = arima(x, order=c(1,0,0))
coefs = coef(fit)
print(coefs)

##          ar1 intercept
## 0.9704851 5.5866513

# Significativité statistique
std_errors = sqrt(diag(fit$var.coef))
p_values = 2 * (1 - pnorm(abs(coefs / std_errors)))
print(p_values)

##          ar1 intercept
## 0             0

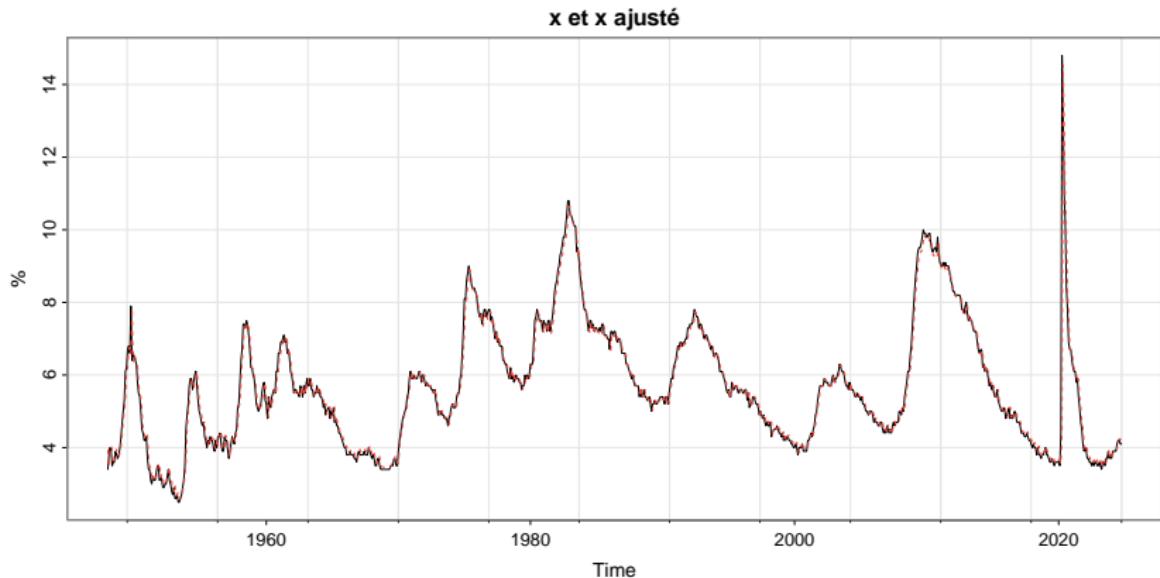
# Ou simplement

#coefest(fit)

#Résidus et serie ajustée
xres = residuals(fit)
xfit = x - xres
# print(head(xres))
# print(head(xfit))
```

## Graph de $y$ et $\hat{y}$

```
par(mfrow=c(1,1))
tsplot(xDate, x, main="x et x ajusté", ylab = "%")
points(xDate, xfit, type = "l", col = 2, lty = 2)
```



## Tests de diagnostic

```
Box.test(xres, lag=2, type="Ljung-Box") #H0: absence d'autocorrelation

##
##  Box-Ljung test
##
## data: xres
## X-squared = 2.8639, df = 2, p-value = 0.2388

shapiro.test(xres) #H0: normalité

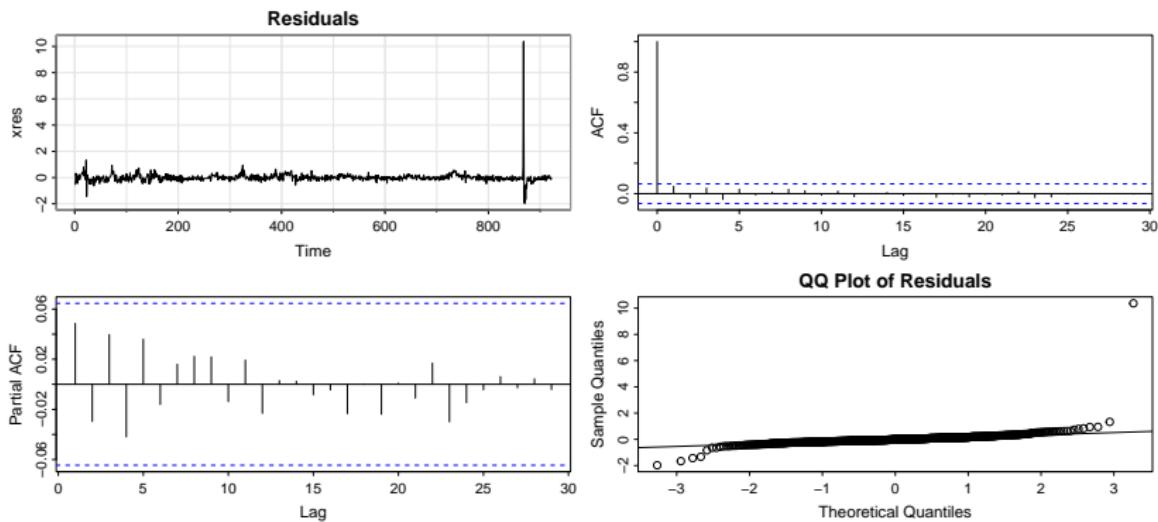
##
##  Shapiro-Wilk normality test
##
## data: xres
## W = 0.38604, p-value < 2.2e-16

bptest(xres ~ xfit) #H0: homoscédasticité

##
##  studentized Breusch-Pagan test
##
## data: xres ~ xfit
## BP = 0.11146, df = 1, p-value = 0.7385
```

## Graph des résidus

```
par(mfrow = c(2, 2))
tsplot(xres, main = "Residuals")
acf(xres, main = "ACF of Residuals")
pacf(xres, main = "PACF of Residuals")
qqnorm(xres, main = "QQ Plot of Residuals")
qqline(xres)
```



## Conclusion

Le modèle est non congruent - échoue au test de normalité.

Un modèle AR( $p$ ) ( $p > 1$ ) pourrait être essayé.

L'explication macroéconomique suggère que des covariables sont nécessaires.

Pas de preuves statistiques de non-stationnarité, mais forte persistence

$$\alpha = 0.9704^{***}$$

Par conséquent, examiner les données comme non stationnaires ensuite.

Puis considérer les modèles dynamiques multivariés.