

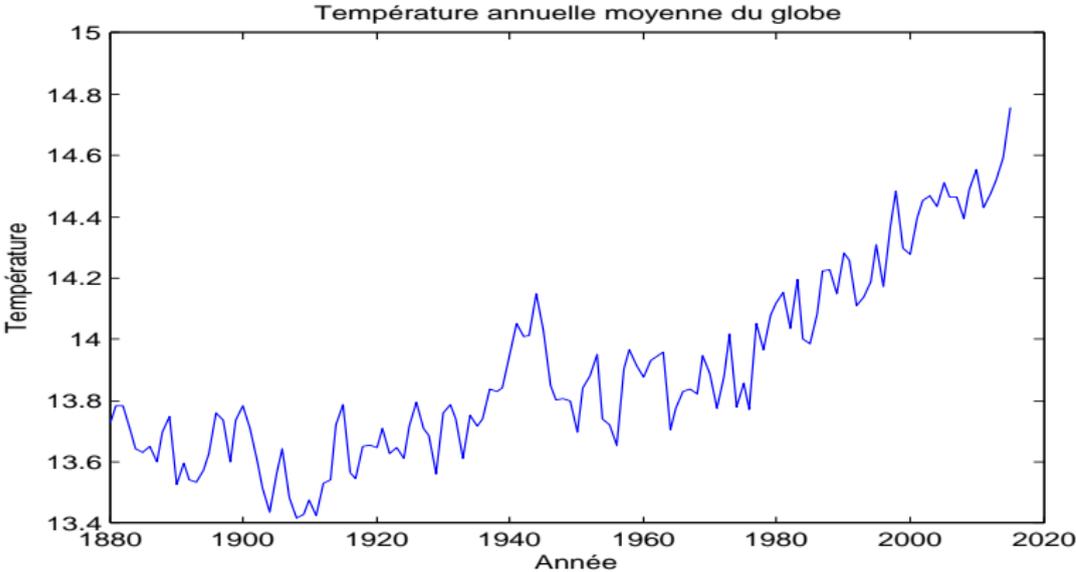
# Cours d'Économétrie des séries chronologiques

Master 2 T.I.D.E.

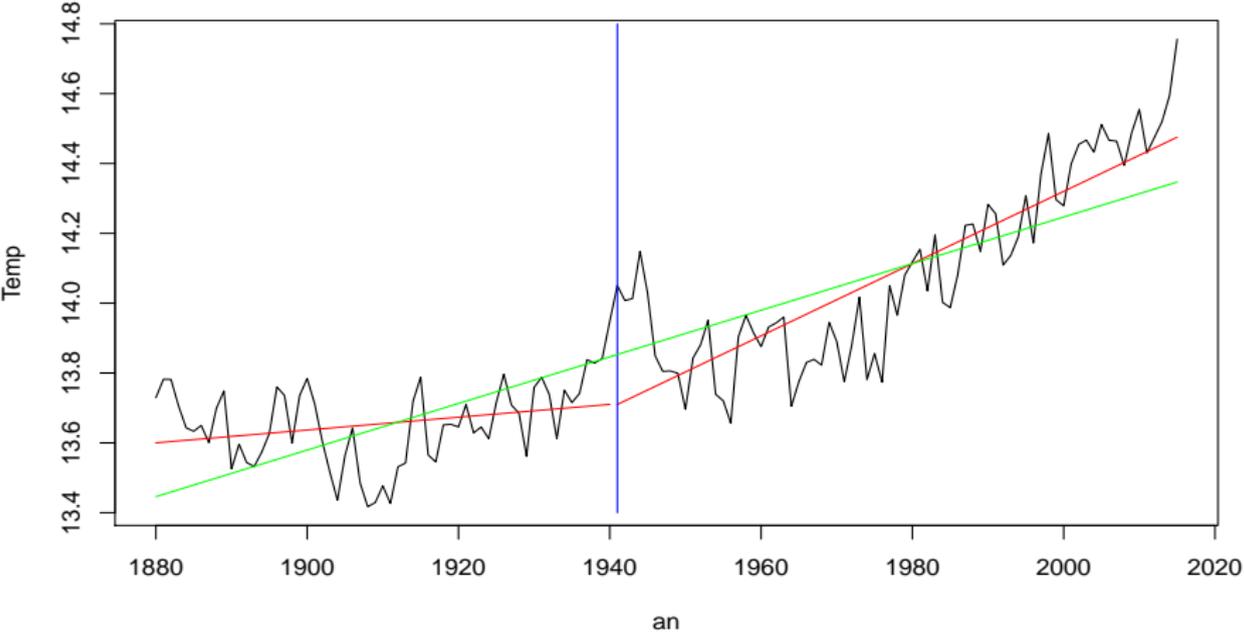


Année 2024-2025

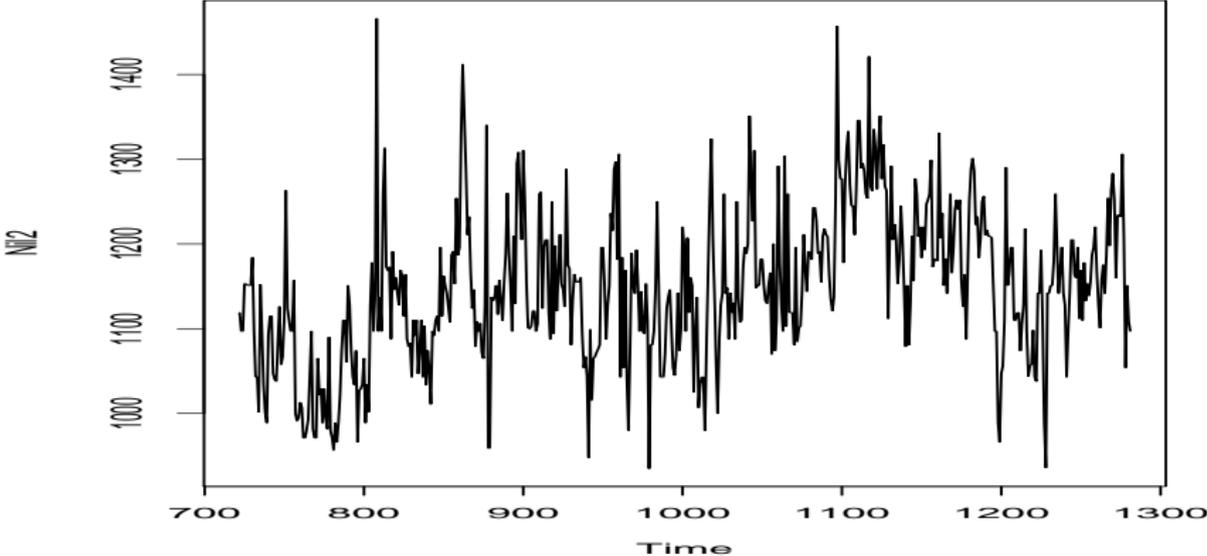
# Un exemple



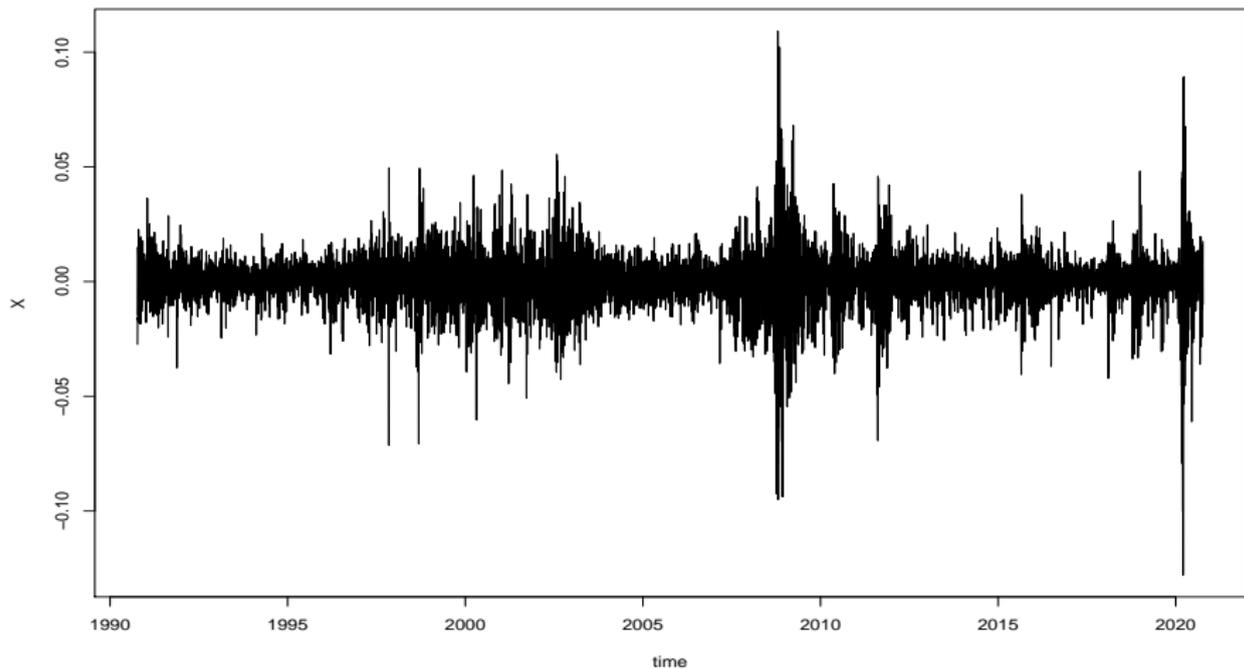
# L'exemple travaillé



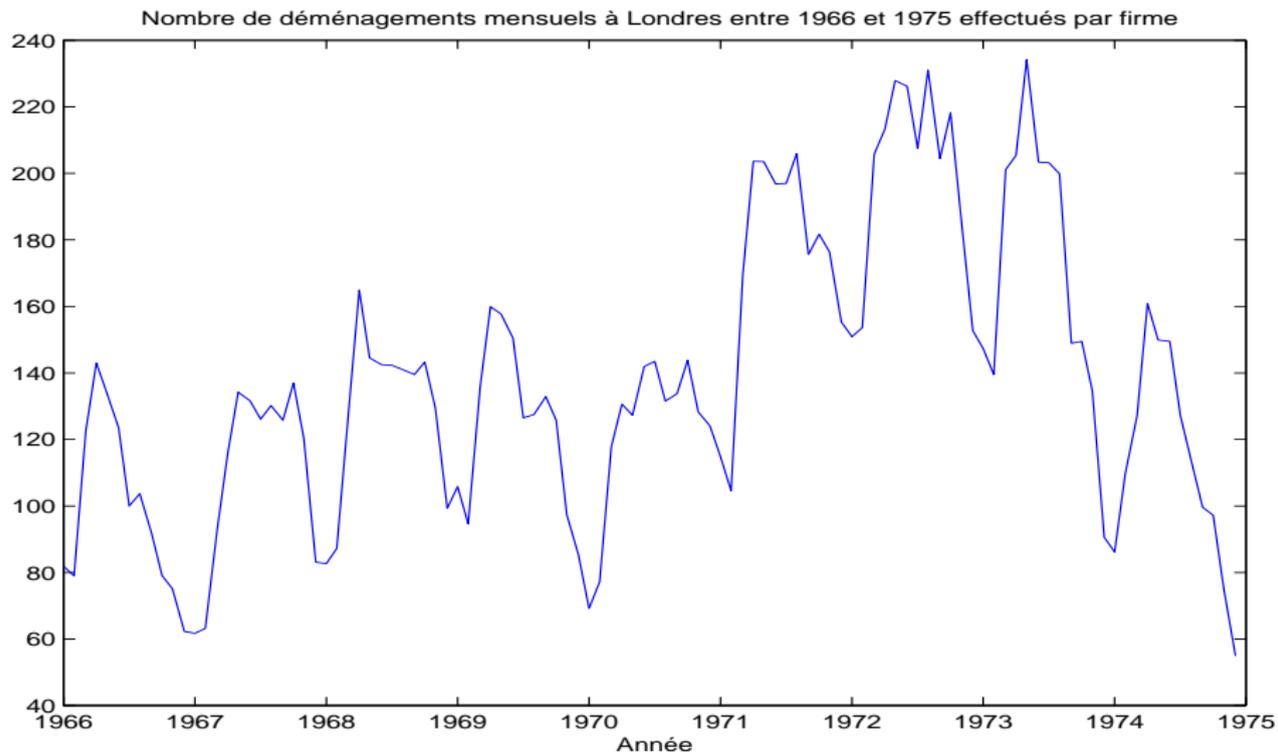
# Un autre exemple : étiages du Nil



# Log-rendements SP500 d'octobre 1990 à octobre 2020



# Un autre autre exemple



# Séries chronologiques

On s'intéresse à  $n$  valeurs observées en des **temps régulièrement espacés**

⇒ On essaie de décrire la série, de la **modéliser**

⇒ On pourra en déduire des **estimations**, des **tests**

⇒ On pourra effectuer des **prédictions**

**Remarque** : Dans ce cours, on ne considère pas :

- Des phénomènes observés en temps continu ou irréguliers;
- Des phénomènes multidimensionnels.

# Séries chronologiques

On s'intéresse à  $n$  valeurs observées en des **temps régulièrement espacés**

⇒ On essaie de décrire la série, de la **modéliser**

⇒ On pourra en déduire des **estimations**, des **tests**

⇒ On pourra effectuer des **prédictions**

**Remarque** : Dans ce cours, on ne considère pas :

- Des phénomènes observés en temps continu ou irréguliers;
- Des phénomènes multidimensionnels.

# Séries chronologiques

On s'intéresse à  $n$  valeurs observées en des **temps régulièrement espacés**

⇒ On essaie de décrire la série, de la **modéliser**

⇒ On pourra en déduire des **estimations**, des **tests**

⇒ On pourra effectuer des **prédictions**

**Remarque** : Dans ce cours, on ne considère pas :

- Des phénomènes observés en temps continu ou irréguliers;
- Des phénomènes multidimensionnels.

# Séries chronologiques

On s'intéresse à  $n$  valeurs observées en des **temps régulièrement espacés**

⇒ On essaie de décrire la série, de la **modéliser**

⇒ On pourra en déduire des **estimations**, des **tests**

⇒ On pourra effectuer des **prédictions**

**Remarque** : Dans ce cours, on ne considère pas :

- Des phénomènes observés en temps continu ou irréguliers;
- Des phénomènes multidimensionnels.

# Séries chronologiques

On s'intéresse à  $n$  valeurs observées en des **temps régulièrement espacés**

⇒ On essaie de décrire la série, de la **modéliser**

⇒ On pourra en déduire des **estimations**, des **tests**

⇒ On pourra effectuer des **prédictions**

**Remarque** : Dans ce cours, on ne considère pas :

- Des phénomènes observés en temps continu ou irréguliers;
- Des phénomènes multidimensionnels.

# Séries chronologiques

On s'intéresse à  $n$  valeurs observées en des **temps régulièrement espacés**

⇒ On essaie de décrire la série, de la **modéliser**

⇒ On pourra en déduire des **estimations**, des **tests**

⇒ On pourra effectuer des **prédictions**

**Remarque** : Dans ce cours, on ne considère pas :

- Des phénomènes observés en temps continu ou irréguliers;
- Des phénomènes multidimensionnels.

# Séries chronologiques

On s'intéresse à  $n$  valeurs observées en des **temps régulièrement espacés**

⇒ On essaie de décrire la série, de la **modéliser**

⇒ On pourra en déduire des **estimations**, des **tests**

⇒ On pourra effectuer des **prédictions**

**Remarque** : Dans ce cours, on ne considère pas :

- Des phénomènes observés en temps continu ou irréguliers;
- Des phénomènes multidimensionnels.

# Plan du cours

- 1 Définitions et premières propriétés
  - Séries chronologiques
  - Stationnarité
  - Première analyse statistique de la dépendance
- 2 Tendances et saisonnalités
  - Définition et propriétés
  - Estimation semi-paramétrique de la tendance et de la saisonnalité
  - Estimation non-paramétrique de la tendance et de la saisonnalité
- 3 Exemples de modèles de séries stationnaires
  - Processus ARMA
  - Processus GARCH
  - Autres exemples de séries temporelles
- 4 Estimation, sélection de modèle, test et prédiction
  - Estimation semi-paramétrique
  - Sélection de modèles et test d'adéquation
  - Prédiction

# Séries chronologiques

Dans toute la suite, on considérera  $n$  valeurs réelles observées en des temps régulièrement espacés

⇒ Modélisées par  $(X_1, \dots, X_n)$  où  $X_i$  variables aléatoires

(plus exactement par  $(X_1(\omega), \dots, X_n(\omega))$  pour une trajectoire...)

**Remarque** : On ne peut plus espérer en général que  $(X_i)$  v.a.i.i.d.

- Le phénomène observé évolue en général au fil du temps;
- Une donnée et sa suivante sont rarement indépendantes.

# Séries chronologiques

Dans toute la suite, on considérera  $n$  valeurs réelles observées en des temps régulièrement espacés

⇒ Modélisées par  $(X_1, \dots, X_n)$  où  $X_i$  variables aléatoires

(plus exactement par  $(X_1(\omega), \dots, X_n(\omega))$  pour une trajectoire...)

**Remarque** : On ne peut plus espérer en général que  $(X_i)$  v.a.i.i.d.

- Le phénomène observé évolue en général au fil du temps;
- Une donnée et sa suivante sont rarement indépendantes.

# Séries chronologiques

Dans toute la suite, on considérera  $n$  valeurs réelles observées en des temps régulièrement espacés

⇒ Modélisées par  $(X_1, \dots, X_n)$  où  $X_i$  variables aléatoires

(plus exactement par  $(X_1(\omega), \dots, X_n(\omega))$  pour une trajectoire...)

**Remarque :** On ne peut plus espérer en général que  $(X_i)$  v.a.i.i.d.

- Le phénomène observé évolue en général au fil du temps;
- Une donnée et sa suivante sont rarement indépendantes.

# Séries chronologiques

Dans toute la suite, on considérera  $n$  valeurs réelles observées en des temps régulièrement espacés

⇒ Modélisées par  $(X_1, \dots, X_n)$  où  $X_i$  variables aléatoires

(plus exactement par  $(X_1(\omega), \dots, X_n(\omega))$  pour une trajectoire...)

**Remarque** : On ne peut plus espérer en général que  $(X_i)$  v.a.i.i.d.

- Le phénomène observé évolue en général au fil du temps ;
- Une donnée et sa suivante sont rarement indépendantes.

# Séries chronologiques

Dans toute la suite, on considérera  $n$  valeurs réelles observées en des temps régulièrement espacés

⇒ Modélisées par  $(X_1, \dots, X_n)$  où  $X_i$  variables aléatoires

(plus exactement par  $(X_1(\omega), \dots, X_n(\omega))$  pour une trajectoire...)

**Remarque** : On ne peut plus espérer en général que  $(X_i)$  v.a.i.i.d.

- Le phénomène observé évolue en général au fil du temps;
- Une donnée et sa suivante sont rarement indépendantes.

# Séries chronologiques

Dans toute la suite, on considérera  $n$  valeurs réelles observées en des temps régulièrement espacés

⇒ Modélisées par  $(X_1, \dots, X_n)$  où  $X_i$  variables aléatoires

(plus exactement par  $(X_1(\omega), \dots, X_n(\omega))$  pour une trajectoire...)

**Remarque** : On ne peut plus espérer en général que  $(X_i)$  v.a.i.i.d.

- Le phénomène observé évolue en général au fil du temps;
- Une donnée et sa suivante sont rarement indépendantes.

# Séries chronologiques

## Définition

Soit  $(\Omega, \mathcal{A}, P)$  un espace de probabilité.

- $X = (X_t, t \in \mathbb{Z})$  série chronologique si pour tout  $t \in \mathbb{Z}$ ,  $X_t$  v.a. sur  $(\Omega, \mathcal{A})$  à valeurs dans  $\mathbb{R}$ .
- Pour  $\omega \in \Omega$ ,  $\{X_t(\omega), 1 \leq t \leq n\}$  appelé une trajectoire de la série.

Exemples :  $(X_t)$  peut être par exemple :

- 1 une fonction réelle (pas d'aléa);
- 2 une suite de v.a.i.i.d;
- 3 une marche aléatoire ( $X_t = Y_1 + \dots + Y_t$  où  $(Y_t)$  v.a.i.i.d. centrées);
- 4 une chaîne de Markov ( $\mathbb{P}(X_t = x \mid X_{t-1} = y)$  défini la série).

# Séries chronologiques

## Définition

Soit  $(\Omega, \mathcal{A}, P)$  un espace de probabilité.

- $X = (X_t, t \in \mathbb{Z})$  **série chronologique** si pour tout  $t \in \mathbb{Z}$ ,  $X_t$  v.a. sur  $(\Omega, \mathcal{A})$  à valeurs dans  $\mathbb{R}$ .
- Pour  $\omega \in \Omega$ ,  $\{X_t(\omega), 1 \leq t \leq n\}$  appelé une **trajectoire de la série**.

Exemples :  $(X_t)$  peut être par exemple :

- 1 une fonction réelle (pas d'aléa);
- 2 une suite de v.a.i.i.d;
- 3 une marche aléatoire ( $X_t = Y_1 + \dots + Y_t$  où  $(Y_t)$  v.a.i.i.d. centrées);
- 4 une chaîne de Markov ( $\mathbb{P}(X_t = x \mid X_{t-1} = y)$  définit la série).

# Séries chronologiques

## Définition

Soit  $(\Omega, \mathcal{A}, P)$  un espace de probabilité.

- $X = (X_t, t \in \mathbb{Z})$  **série chronologique** si pour tout  $t \in \mathbb{Z}$ ,  $X_t$  v.a. sur  $(\Omega, \mathcal{A})$  à valeurs dans  $\mathbb{R}$ .
- Pour  $\omega \in \Omega$ ,  $\{X_t(\omega), 1 \leq t \leq n\}$  appelé une **trajectoire** de la série.

Exemples :  $(X_t)$  peut être par exemple :

- 1 une fonction réelle (pas d'aléa);
- 2 une suite de v.a.i.i.d;
- 3 une marche aléatoire ( $X_t = Y_1 + \dots + Y_t$  où  $(Y_t)$  v.a.i.i.d. centrées);
- 4 une chaîne de Markov ( $\mathbb{P}(X_t = x \mid X_{t-1} = y)$  défini la série).

# Séries chronologiques

## Définition

Soit  $(\Omega, \mathcal{A}, P)$  un espace de probabilité.

- $X = (X_t, t \in \mathbb{Z})$  **série chronologique** si pour tout  $t \in \mathbb{Z}$ ,  $X_t$  v.a. sur  $(\Omega, \mathcal{A})$  à valeurs dans  $\mathbb{R}$ .
- Pour  $\omega \in \Omega$ ,  $\{X_t(\omega), 1 \leq t \leq n\}$  appelé une **trajectoire** de la série.

**Exemples** :  $(X_t)$  peut être par exemple :

- 1 une fonction réelle (pas d'aléa) ;
- 2 une suite de v.a.i.i.d ;
- 3 une marche aléatoire ( $X_t = Y_1 + \dots + Y_t$  où  $(Y_t)$  v.a.i.i.d. centrées) ;
- 4 une chaîne de Markov ( $\mathbb{P}(X_t = x \mid X_{t-1} = y)$  défini la série).

# Séries chronologiques

## Définition

Soit  $(\Omega, \mathcal{A}, P)$  un espace de probabilité.

- $X = (X_t, t \in \mathbb{Z})$  **série chronologique** si pour tout  $t \in \mathbb{Z}$ ,  $X_t$  v.a. sur  $(\Omega, \mathcal{A})$  à valeurs dans  $\mathbb{R}$ .
- Pour  $\omega \in \Omega$ ,  $\{X_t(\omega), 1 \leq t \leq n\}$  appelé une **trajectoire** de la série.

**Exemples** :  $(X_t)$  peut être par exemple :

- 1 une fonction réelle (pas d'aléa) ;
- 2 une suite de v.a.i.i.d ;
- 3 une marche aléatoire ( $X_t = Y_1 + \dots + Y_t$  où  $(Y_t)$  v.a.i.i.d. centrées) ;
- 4 une chaîne de Markov ( $\mathbb{P}(X_t = x \mid X_{t-1} = y)$  défini la série).

# Séries chronologiques

## Définition

Soit  $(\Omega, \mathcal{A}, P)$  un espace de probabilité.

- $X = (X_t, t \in \mathbb{Z})$  **série chronologique** si pour tout  $t \in \mathbb{Z}$ ,  $X_t$  v.a. sur  $(\Omega, \mathcal{A})$  à valeurs dans  $\mathbb{R}$ .
- Pour  $\omega \in \Omega$ ,  $\{X_t(\omega), 1 \leq t \leq n\}$  appelé une **trajectoire** de la série.

**Exemples** :  $(X_t)$  peut être par exemple :

- 1 une fonction réelle (pas d'aléa) ;
- 2 une suite de v.a.i.i.d ;
- 3 une marche aléatoire ( $X_t = Y_1 + \dots + Y_t$  où  $(Y_t)$  v.a.i.i.d. centrées) ;
- 4 une chaîne de Markov ( $\mathbb{P}(X_t = x \mid X_{t-1} = y)$  définit la série).

# Séries chronologiques

## Définition

Soit  $(\Omega, \mathcal{A}, P)$  un espace de probabilité.

- $X = (X_t, t \in \mathbb{Z})$  **série chronologique** si pour tout  $t \in \mathbb{Z}$ ,  $X_t$  v.a. sur  $(\Omega, \mathcal{A})$  à valeurs dans  $\mathbb{R}$ .
- Pour  $\omega \in \Omega$ ,  $\{X_t(\omega), 1 \leq t \leq n\}$  appelé une **trajectoire** de la série.

**Exemples** :  $(X_t)$  peut être par exemple :

- 1 une fonction réelle (pas d'aléa) ;
- 2 une suite de v.a.i.i.d ;
- 3 une marche aléatoire ( $X_t = Y_1 + \dots + Y_t$  où  $(Y_t)$  v.a.i.i.d. centrées) ;
- 4 une chaîne de Markov ( $\mathbb{P}(X_t = x \mid X_{t-1} = y)$  défini la série).

# Séries chronologiques

## Définition

Soit  $(\Omega, \mathcal{A}, P)$  un espace de probabilité.

- $X = (X_t, t \in \mathbb{Z})$  **série chronologique** si pour tout  $t \in \mathbb{Z}$ ,  $X_t$  v.a. sur  $(\Omega, \mathcal{A})$  à valeurs dans  $\mathbb{R}$ .
- Pour  $\omega \in \Omega$ ,  $\{X_t(\omega), 1 \leq t \leq n\}$  appelé une **trajectoire** de la série.

**Exemples** :  $(X_t)$  peut être par exemple :

- 1 une fonction réelle (pas d'aléa) ;
- 2 une suite de v.a.i.i.d ;
- 3 une marche aléatoire ( $X_t = Y_1 + \dots + Y_t$  où  $(Y_t)$  v.a.i.i.d. centrées) ;
- 4 une chaîne de Markov ( $\mathbf{P}(X_t = x \mid X_{t-1} = y)$  définit la série).

# Séries chronologiques

## Définition

Soit  $(X_t)_{t \in \mathbb{Z}}$  série chronologique telle que  $\mathbb{E}[X_t^2] < \infty$  pour tout  $t \in \mathbb{Z}$

Pour  $(s, t) \in \mathbb{Z}^2$ , on appelle :

- fonction espérance :  $m(t) = \mathbb{E}[X_t]$  ;
- fonction variance :  $\sigma^2(t) = \mathbb{E}[X_t^2] - m^2(t) = \mathbb{E}[(X_t - m(t))^2]$  ;
- fonction covariance :  
 $\gamma(s, t) = \mathbb{E}[X_s X_t] - m(s)m(t) = \mathbb{E}[(X_s - m(s))(X_t - m(t))]$  ;
- fonction corrélation :  $\rho(s, t) = \frac{\gamma(s, t)}{\sigma(s)\sigma(t)}$  et  $|\rho(s, t)| \leq 1$ .

Exemples : Si  $(X_t)$  est :

- 1 une fonction réelle :  $m(t) = X_t$ ,  $\sigma^2(t) = 0$ ,  $\gamma(s, t) = r(s, t) = 0$
- 2 une suite de v.a.i.i.d. :  $m(t) = m$ ,  $\sigma^2(t) = \sigma^2$ ,  $\gamma(s, t) = r(s, t) = 0$
- 3 une marche aléatoire :  $m(t) = 0$ ,  $\sigma^2(t) = \sigma^2 t$ ,  $\gamma(s, t) = \sigma^2 \min(s, t)$ .

# Séries chronologiques

## Définition

Soit  $(X_t)_{t \in \mathbb{Z}}$  série chronologique telle que  $\mathbb{E}[X_t^2] < \infty$  pour tout  $t \in \mathbb{Z}$   
Pour  $(s, t) \in \mathbb{Z}^2$ , on appelle :

- fonction espérance :  $m(t) = \mathbb{E}[X_t]$  ;
- fonction variance :  $\sigma^2(t) = \mathbb{E}[X_t^2] - m^2(t) = \mathbb{E}[(X_t - m(t))^2]$  ;
- fonction covariance :  
 $\gamma(s, t) = \mathbb{E}[X_s X_t] - m(s)m(t) = \mathbb{E}[(X_s - m(s))(X_t - m(t))]$  ;
- fonction corrélation :  $\rho(s, t) = \frac{\gamma(s, t)}{\sigma(s)\sigma(t)}$  et  $|\rho(s, t)| \leq 1$ .

Exemples : Si  $(X_t)$  est :

- 1 une fonction réelle :  $m(t) = X_t$ ,  $\sigma^2(t) = 0$ ,  $\gamma(s, t) = r(s, t) = 0$
- 2 une suite de v.a.i.i.d. :  $m(t) = m$ ,  $\sigma^2(t) = \sigma^2$ ,  $\gamma(s, t) = r(s, t) = 0$
- 3 une marche aléatoire :  $m(t) = 0$ ,  $\sigma^2(t) = \sigma^2 t$ ,  $\gamma(s, t) = \sigma^2 \min(s, t)$ .

# Séries chronologiques

## Définition

Soit  $(X_t)_{t \in \mathbb{Z}}$  série chronologique telle que  $\mathbb{E}[X_t^2] < \infty$  pour tout  $t \in \mathbb{Z}$   
Pour  $(s, t) \in \mathbb{Z}^2$ , on appelle :

- **fonction espérance** :  $m(t) = \mathbb{E}[X_t]$  ;
- **fonction variance** :  $\sigma^2(t) = \mathbb{E}[X_t^2] - m^2(t) = \mathbb{E}[(X_t - m(t))^2]$  ;
- **fonction covariance** :  
 $\gamma(s, t) = \mathbb{E}[X_s X_t] - m(s)m(t) = \mathbb{E}[(X_s - m(s))(X_t - m(t))]$  ;
- **fonction corrélation** :  $\rho(s, t) = \frac{\gamma(s, t)}{\sigma(s)\sigma(t)}$  et  $|\rho(s, t)| \leq 1$ .

Exemples : Si  $(X_t)$  est :

- 1 une fonction réelle :  $m(t) = X_t$ ,  $\sigma^2(t) = 0$ ,  $\gamma(s, t) = r(s, t) = 0$
- 2 une suite de v.a.i.i.d. :  $m(t) = m$ ,  $\sigma^2(t) = \sigma^2$ ,  $\gamma(s, t) = r(s, t) = 0$
- 3 une marche aléatoire :  $m(t) = 0$ ,  $\sigma^2(t) = \sigma^2 t$ ,  $\gamma(s, t) = \sigma^2 \min(s, t)$ .

# Séries chronologiques

## Définition

Soit  $(X_t)_{t \in \mathbb{Z}}$  série chronologique telle que  $\mathbb{E}[X_t^2] < \infty$  pour tout  $t \in \mathbb{Z}$   
Pour  $(s, t) \in \mathbb{Z}^2$ , on appelle :

- **fonction espérance** :  $m(t) = \mathbb{E}[X_t]$  ;
- **fonction variance** :  $\sigma^2(t) = \mathbb{E}[X_t^2] - m^2(t) = \mathbb{E}[(X_t - m(t))^2]$  ;
- **fonction covariance** :  
 $\gamma(s, t) = \mathbb{E}[X_s X_t] - m(s)m(t) = \mathbb{E}[(X_s - m(s))(X_t - m(t))]$  ;
- **fonction corrélation** :  $\rho(s, t) = \frac{\gamma(s, t)}{\sigma(s)\sigma(t)}$  et  $|\rho(s, t)| \leq 1$ .

Exemples : Si  $(X_t)$  est :

- ① une fonction réelle :  $m(t) = X_t$ ,  $\sigma^2(t) = 0$ ,  $\gamma(s, t) = r(s, t) = 0$
- ② une suite de v.a.i.i.d. :  $m(t) = m$ ,  $\sigma^2(t) = \sigma^2$ ,  $\gamma(s, t) = r(s, t) = 0$
- ③ une marche aléatoire :  $m(t) = 0$ ,  $\sigma^2(t) = \sigma^2 t$ ,  $\gamma(s, t) = \sigma^2 \min(s, t)$ .

# Séries chronologiques

## Définition

Soit  $(X_t)_{t \in \mathbb{Z}}$  série chronologique telle que  $\mathbb{E}[X_t^2] < \infty$  pour tout  $t \in \mathbb{Z}$   
Pour  $(s, t) \in \mathbb{Z}^2$ , on appelle :

- **fonction espérance** :  $m(t) = \mathbb{E}[X_t]$  ;
- **fonction variance** :  $\sigma^2(t) = \mathbb{E}[X_t^2] - m^2(t) = \mathbb{E}[(X_t - m(t))^2]$  ;
- **fonction covariance** :  
 $\gamma(s, t) = \mathbb{E}[X_s X_t] - m(s)m(t) = \mathbb{E}[(X_s - m(s))(X_t - m(t))]$  ;
- **fonction corrélation** :  $\rho(s, t) = \frac{\gamma(s, t)}{\sigma(s)\sigma(t)}$  et  $|\rho(s, t)| \leq 1$ .

Exemples : Si  $(X_t)$  est :

- 1 une fonction réelle :  $m(t) = X_t$ ,  $\sigma^2(t) = 0$ ,  $\gamma(s, t) = r(s, t) = 0$
- 2 une suite de v.a.i.i.d. :  $m(t) = m$ ,  $\sigma^2(t) = \sigma^2$ ,  $\gamma(s, t) = r(s, t) = 0$
- 3 une marche aléatoire :  $m(t) = 0$ ,  $\sigma^2(t) = \sigma^2 t$ ,  $\gamma(s, t) = \sigma^2 \min(s, t)$ .

# Séries chronologiques

## Définition

Soit  $(X_t)_{t \in \mathbb{Z}}$  série chronologique telle que  $\mathbb{E}[X_t^2] < \infty$  pour tout  $t \in \mathbb{Z}$   
Pour  $(s, t) \in \mathbb{Z}^2$ , on appelle :

- **fonction espérance** :  $m(t) = \mathbb{E}[X_t]$  ;
- **fonction variance** :  $\sigma^2(t) = \mathbb{E}[X_t^2] - m^2(t) = \mathbb{E}[(X_t - m(t))^2]$  ;
- **fonction covariance** :  
 $\gamma(s, t) = \mathbb{E}[X_s X_t] - m(s)m(t) = \mathbb{E}[(X_s - m(s))(X_t - m(t))]$  ;
- **fonction corrélation** :  $\rho(s, t) = \frac{\gamma(s, t)}{\sigma(s)\sigma(t)}$  et  $|\rho(s, t)| \leq 1$ .

**Exemples** : Si  $(X_t)$  est :

- 1 une fonction réelle :  $m(t) = X_t$ ,  $\sigma^2(t) = 0$ ,  $\gamma(s, t) = r(s, t) = 0$
- 2 une suite de v.a.i.i.d. :  $m(t) = m$ ,  $\sigma^2(t) = \sigma^2$ ,  $\gamma(s, t) = r(s, t) = 0$
- 3 une marche aléatoire :  $m(t) = 0$ ,  $\sigma^2(t) = \sigma^2 t$ ,  $\gamma(s, t) = \sigma^2 \min(s, t)$ .

# Séries chronologiques

## Définition

Soit  $(X_t)_{t \in \mathbb{Z}}$  série chronologique telle que  $\mathbb{E}[X_t^2] < \infty$  pour tout  $t \in \mathbb{Z}$   
Pour  $(s, t) \in \mathbb{Z}^2$ , on appelle :

- **fonction espérance** :  $m(t) = \mathbb{E}[X_t]$  ;
- **fonction variance** :  $\sigma^2(t) = \mathbb{E}[X_t^2] - m^2(t) = \mathbb{E}[(X_t - m(t))^2]$  ;
- **fonction covariance** :  
 $\gamma(s, t) = \mathbb{E}[X_s X_t] - m(s)m(t) = \mathbb{E}[(X_s - m(s))(X_t - m(t))]$  ;
- **fonction corrélation** :  $\rho(s, t) = \frac{\gamma(s, t)}{\sigma(s)\sigma(t)}$  et  $|\rho(s, t)| \leq 1$ .

**Exemples** : Si  $(X_t)$  est :

- 1 une fonction réelle :  $m(t) = X_t$ ,  $\sigma^2(t) = 0$ ,  $\gamma(s, t) = r(s, t) = 0$
- 2 une suite de v.a.i.i.d. :  $m(t) = m$ ,  $\sigma^2(t) = \sigma^2$ ,  $\gamma(s, t) = r(s, t) = 0$
- 3 une marche aléatoire :  $m(t) = 0$ ,  $\sigma^2(t) = \sigma^2 t$ ,  $\gamma(s, t) = \sigma^2 \min(s, t)$ .

# Séries chronologiques

## Définition

Soit  $(X_t)_{t \in \mathbb{Z}}$  série chronologique telle que  $\mathbb{E}[X_t^2] < \infty$  pour tout  $t \in \mathbb{Z}$   
Pour  $(s, t) \in \mathbb{Z}^2$ , on appelle :

- **fonction espérance** :  $m(t) = \mathbb{E}[X_t]$  ;
- **fonction variance** :  $\sigma^2(t) = \mathbb{E}[X_t^2] - m^2(t) = \mathbb{E}[(X_t - m(t))^2]$  ;
- **fonction covariance** :  
 $\gamma(s, t) = \mathbb{E}[X_s X_t] - m(s)m(t) = \mathbb{E}[(X_s - m(s))(X_t - m(t))]$  ;
- **fonction corrélation** :  $\rho(s, t) = \frac{\gamma(s, t)}{\sigma(s)\sigma(t)}$  et  $|\rho(s, t)| \leq 1$ .

**Exemples** : Si  $(X_t)$  est :

- 1 une fonction réelle :  $m(t) = X_t$ ,  $\sigma^2(t) = 0$ ,  $\gamma(s, t) = r(s, t) = 0$
- 2 une suite de v.a.i.i.d. :  $m(t) = m$ ,  $\sigma^2(t) = \sigma^2$ ,  $\gamma(s, t) = r(s, t) = 0$
- 3 une marche aléatoire :  $m(t) = 0$ ,  $\sigma^2(t) = \sigma^2 t$ ,  $\gamma(s, t) = \sigma^2 \min(s, t)$ .

# Séries chronologiques

## Définition

Soit  $(X_t)_{t \in \mathbb{Z}}$  série chronologique telle que  $\mathbb{E}[X_t^2] < \infty$  pour tout  $t \in \mathbb{Z}$   
Pour  $(s, t) \in \mathbb{Z}^2$ , on appelle :

- **fonction espérance** :  $m(t) = \mathbb{E}[X_t]$  ;
- **fonction variance** :  $\sigma^2(t) = \mathbb{E}[X_t^2] - m^2(t) = \mathbb{E}[(X_t - m(t))^2]$  ;
- **fonction covariance** :  
 $\gamma(s, t) = \mathbb{E}[X_s X_t] - m(s)m(t) = \mathbb{E}[(X_s - m(s))(X_t - m(t))]$  ;
- **fonction corrélation** :  $\rho(s, t) = \frac{\gamma(s, t)}{\sigma(s)\sigma(t)}$  et  $|\rho(s, t)| \leq 1$ .

**Exemples** : Si  $(X_t)$  est :

- 1 une fonction réelle :  $m(t) = X_t$ ,  $\sigma^2(t) = 0$ ,  $\gamma(s, t) = r(s, t) = 0$
- 2 une suite de v.a.i.i.d. :  $m(t) = m$ ,  $\sigma^2(t) = \sigma^2$ ,  $\gamma(s, t) = r(s, t) = 0$
- 3 une marche aléatoire :  $m(t) = 0$ ,  $\sigma^2(t) = \sigma^2 t$ ,  $\gamma(s, t) = \sigma^2 \min(s, t)$ .

# Séries chronologiques

## Définition

Une série  $(X_t)_{t \in \mathbb{Z}}$  est **gaussienne** si  $\sum_{i=1}^n \alpha_i X_{t_i}$  est une variable gaussienne pour tout  $n \in \mathbb{N}^*$ ,  $(t_1, \dots, t_n) \in \mathbb{Z}^n$ ,  $(\alpha_1, \dots, \alpha_n) \in \mathbb{R}^n$ .

## Propriété

Pour une série chronologique  $(X_t)_{t \in \mathbb{Z}}$  gaussienne :

$(\gamma(s, t) = 0, \text{ pour tout } s \neq t) \iff ((X_t)_{t \in \mathbb{Z}} \text{ suite de v.a. indépendantes } )$ .

## Définition

- Un **bruit blanc** est une suite v.a.i.i.d. centrées.
- Un **bruit blanc gaussien** est une suite de v.a.i.i.d. gaussiennes centrées.

# Séries chronologiques

## Définition

Une série  $(X_t)_{t \in \mathbb{Z}}$  est **gaussienne** si  $\sum_{i=1}^n \alpha_i X_{t_i}$  est une variable gaussienne pour tout  $n \in \mathbb{N}^*$ ,  $(t_1, \dots, t_n) \in \mathbb{Z}^n$ ,  $(\alpha_1, \dots, \alpha_n) \in \mathbb{R}^n$ .

## Propriété

Pour une série chronologique  $(X_t)_{t \in \mathbb{Z}}$  gaussienne :

$(\gamma(s, t) = 0, \text{ pour tout } s \neq t) \iff ((X_t)_{t \in \mathbb{Z}} \text{ suite de v.a. indépendantes } )$ .

## Définition

- Un **bruit blanc** est une suite v.a.i.i.d. centrées.
- Un **bruit blanc gaussien** est une suite de v.a.i.i.d. gaussiennes centrées.

# Séries chronologiques

## Définition

Une série  $(X_t)_{t \in \mathbb{Z}}$  est **gaussienne** si  $\sum_{i=1}^n \alpha_i X_{t_i}$  est une variable gaussienne pour tout  $n \in \mathbb{N}^*$ ,  $(t_1, \dots, t_n) \in \mathbb{Z}^n$ ,  $(\alpha_1, \dots, \alpha_n) \in \mathbb{R}^n$ .

## Propriété

Pour une série chronologique  $(X_t)_{t \in \mathbb{Z}}$  gaussienne :  
 $(\gamma(s, t) = 0, \text{ pour tout } s \neq t) \iff ((X_t)_{t \in \mathbb{Z}} \text{ suite de v.a. indépendantes } )$ .

## Définition

- Un **bruit blanc** est une suite v.a.i.i.d. centrées.
- Un **bruit blanc gaussien** est une suite de v.a.i.i.d. gaussiennes centrées.

# Séries chronologiques

## Définition

Une série  $(X_t)_{t \in \mathbb{Z}}$  est **gaussienne** si  $\sum_{i=1}^n \alpha_i X_{t_i}$  est une variable gaussienne pour tout  $n \in \mathbb{N}^*$ ,  $(t_1, \dots, t_n) \in \mathbb{Z}^n$ ,  $(\alpha_1, \dots, \alpha_n) \in \mathbb{R}^n$ .

## Propriété

Pour une série chronologique  $(X_t)_{t \in \mathbb{Z}}$  gaussienne :  
 $(\gamma(s, t) = 0, \text{ pour tout } s \neq t) \iff ((X_t)_{t \in \mathbb{Z}} \text{ suite de v.a. indépendantes } )$ .

## Définition

- Un **bruit blanc** est une suite v.a.i.i.d. centrées.
- Un **bruit blanc gaussien** est une suite de v.a.i.i.d. gaussiennes centrées.

# Plan du cours

- 1 Définitions et premières propriétés
  - Séries chronologiques
  - Stationnarité
  - Première analyse statistique de la dépendance
- 2 Tendances et saisonnalités
  - Définition et propriétés
  - Estimation semi-paramétrique de la tendance et de la saisonnalité
  - Estimation non-paramétrique de la tendance et de la saisonnalité
- 3 Exemples de modèles de séries stationnaires
  - Processus ARMA
  - Processus GARCH
  - Autres exemples de séries temporelles
- 4 Estimation, sélection de modèle, test et prédiction
  - Estimation semi-paramétrique
  - Sélection de modèles et test d'adéquation
  - Prédiction

# Stationnarité

## Définition

Une série chronologique  $(X_t)_{t \in \mathbb{Z}}$  est (strictement) **stationnaire** lorsque  $\forall n \in \mathbb{N}^*, \forall (t_1, \dots, t_n) \in \mathbb{Z}^n, \forall c \in \mathbb{Z},$

$$(X_{t_1}, \dots, X_{t_n}) \stackrel{\mathcal{L}}{\sim} (X_{t_1+c}, \dots, X_{t_n+c}).$$

$\implies$  Si  $(X_t)$  stationnaire,  $(X_t)$  suite identiquement distribuée

## Propriété

Si  $(X_t)$  stationnaire

- les fonctions  $m(t) = \mathbb{E}[X_t]$  et  $\sigma^2(t) = \text{var}(X_t)$  sont constantes
- $\gamma(s, t) = \text{cov}(X_s, X_t) = r(|t - s|)$  et  $\rho(s, t) = \text{Cor}(X_s, X_t) = \frac{r(|t-s|)}{r(0)}$ .

Exemple :  $(X_k)$  suite de v.a.i.i.d. est stationnaire.

# Stationnarité

## Définition

Une série chronologique  $(X_t)_{t \in \mathbb{Z}}$  est (strictement) **stationnaire** lorsque  $\forall n \in \mathbb{N}^*, \forall (t_1, \dots, t_n) \in \mathbb{Z}^n, \forall c \in \mathbb{Z}$ ,

$$(X_{t_1}, \dots, X_{t_n}) \stackrel{\mathcal{L}}{\sim} (X_{t_1+c}, \dots, X_{t_n+c}).$$

$\implies$  Si  $(X_t)$  stationnaire,  $(X_t)$  suite identiquement distribuée

## Propriété

Si  $(X_t)$  stationnaire

- les fonctions  $m(t) = \mathbb{E}[X_t]$  et  $\sigma^2(t) = \text{var}(X_t)$  sont constantes
- $\gamma(s, t) = \text{cov}(X_s, X_t) = r(|t - s|)$  et  $\rho(s, t) = \text{Cor}(X_s, X_t) = \frac{r(|t-s|)}{r(0)}$ .

Exemple :  $(X_k)$  suite de v.a.i.i.d. est stationnaire.

# Stationnarité

## Définition

Une série chronologique  $(X_t)_{t \in \mathbb{Z}}$  est (strictement) **stationnaire** lorsque  $\forall n \in \mathbb{N}^*, \forall (t_1, \dots, t_n) \in \mathbb{Z}^n, \forall c \in \mathbb{Z}$ ,

$$(X_{t_1}, \dots, X_{t_n}) \stackrel{\mathcal{L}}{\sim} (X_{t_1+c}, \dots, X_{t_n+c}).$$

$\implies$  Si  $(X_t)$  stationnaire,  $(X_t)$  suite identiquement distribuée

## Propriété

Si  $(X_t)$  stationnaire

- les fonctions  $m(t) = \mathbb{E}[X_t]$  et  $\sigma^2(t) = \text{var}(X_t)$  sont constantes
- $\gamma(s, t) = \text{cov}(X_s, X_t) = r(|t - s|)$  et  $\rho(s, t) = \text{Cor}(X_s, X_t) = \frac{r(|t-s|)}{r(0)}$ .

Exemple :  $(X_k)$  suite de v.a.i.i.d. est stationnaire.

# Stationnarité

## Définition

Une série chronologique  $(X_t)_{t \in \mathbb{Z}}$  est (strictement) **stationnaire** lorsque  $\forall n \in \mathbb{N}^*, \forall (t_1, \dots, t_n) \in \mathbb{Z}^n, \forall c \in \mathbb{Z},$

$$(X_{t_1}, \dots, X_{t_n}) \stackrel{\mathcal{L}}{\sim} (X_{t_1+c}, \dots, X_{t_n+c}).$$

$\implies$  Si  $(X_t)$  stationnaire,  $(X_t)$  suite identiquement distribuée

## Propriété

Si  $(X_t)$  stationnaire

- les fonctions  $m(t) = \mathbb{E}[X_t]$  et  $\sigma^2(t) = \text{var}(X_t)$  sont constantes
- $\gamma(s, t) = \text{cov}(X_s, X_t) = r(|t - s|)$  et  $\rho(s, t) = \text{Cor}(X_s, X_t) = \frac{r(|t-s|)}{r(0)}$ .

Exemple :  $(X_k)$  suite de v.a.i.i.d. est stationnaire.

# Stationnarité

## Définition

Une série chronologique  $(X_t)_{t \in \mathbb{Z}}$  est (strictement) **stationnaire** lorsque  $\forall n \in \mathbb{N}^*$ ,  $\forall (t_1, \dots, t_n) \in \mathbb{Z}^n$ ,  $\forall c \in \mathbb{Z}$ ,

$$(X_{t_1}, \dots, X_{t_n}) \stackrel{\mathcal{L}}{\sim} (X_{t_1+c}, \dots, X_{t_n+c}).$$

$\implies$  Si  $(X_t)$  stationnaire,  $(X_t)$  suite identiquement distribuée

## Propriété

Si  $(X_t)$  stationnaire

- les fonctions  $m(t) = \mathbb{E}[X_t]$  et  $\sigma^2(t) = \text{var}(X_t)$  sont constantes
- $\gamma(s, t) = \text{cov}(X_s, X_t) = r(|t - s|)$  et  $\rho(s, t) = \text{Cor}(X_s, X_t) = \frac{r(|t-s|)}{r(0)}$ .

Exemple :  $(X_k)$  suite de v.a.i.i.d. est stationnaire.

# Stationnarité

## Définition

Une série chronologique  $(X_t)_{t \in \mathbb{Z}}$  est (strictement) **stationnaire** lorsque  $\forall n \in \mathbb{N}^*, \forall (t_1, \dots, t_n) \in \mathbb{Z}^n, \forall c \in \mathbb{Z},$

$$(X_{t_1}, \dots, X_{t_n}) \stackrel{\mathcal{L}}{\sim} (X_{t_1+c}, \dots, X_{t_n+c}).$$

$\implies$  Si  $(X_t)$  stationnaire,  $(X_t)$  suite identiquement distribuée

## Propriété

Si  $(X_t)$  stationnaire

- les fonctions  $m(t) = \mathbb{E}[X_t]$  et  $\sigma^2(t) = \text{var}(X_t)$  sont constantes
- $\gamma(s, t) = \text{cov}(X_s, X_t) = r(|t - s|)$  et  $\rho(s, t) = \text{Cor}(X_s, X_t) = \frac{r(|t-s|)}{r(0)}$ .

**Exemple** :  $(X_k)$  suite de v.a.i.i.d. est stationnaire.

# Stationnarité

## Définition

On dit que  $(X_t)_{t \in \mathbb{Z}}$  série stationnaire d'ordre 2 lorsque :

- 1 son espérance  $m(t)$  est constante ;
- 2 sa covariance  $\gamma(s, t) = \text{cov}(X_s, X_t)$  est une fonction de  $|t - s|$ .

**Conséquence** : Si  $m(t)$  non constante, alors  $(X_t)$  admet une tendance et/ou une saisonnalité et est non stationnaire.

## Propriété

- (Stationnarité  $\implies$  Stationnarité d'ordre 2), mais réciproque fausse.
- Si  $(X_t)$  série gaussienne, (Stationnarité  $\iff$  Stationnarité d'ordre 2).

# Stationnarité

## Définition

On dit que  $(X_t)_{t \in \mathbb{Z}}$  série stationnaire d'ordre 2 lorsque :

- 1 son espérance  $m(t)$  est constante ;
- 2 sa covariance  $\gamma(s, t) = \text{cov}(X_s, X_t)$  est une fonction de  $|t - s|$ .

Conséquence : Si  $m(t)$  non constante, alors  $(X_t)$  admet une tendance et/ou une saisonnalité et est non stationnaire.

## Propriété

- (Stationnarité  $\implies$  Stationnarité d'ordre 2), mais réciproque fausse.
- Si  $(X_t)$  série gaussienne, (Stationnarité  $\iff$  Stationnarité d'ordre 2).

# Stationnarité

## Définition

On dit que  $(X_t)_{t \in \mathbb{Z}}$  série stationnaire d'ordre 2 lorsque :

- 1 son espérance  $m(t)$  est constante ;
- 2 sa covariance  $\gamma(s, t) = \text{cov}(X_s, X_t)$  est une fonction de  $|t - s|$ .

**Conséquence** : Si  $m(t)$  non constante, alors  $(X_t)$  admet une tendance et/ou une saisonnalité et est non stationnaire.

## Propriété

- (Stationnarité  $\implies$  Stationnarité d'ordre 2), mais réciproque fausse.
- Si  $(X_t)$  série gaussienne, (Stationnarité  $\iff$  Stationnarité d'ordre 2).

# Stationnarité

## Définition

On dit que  $(X_t)_{t \in \mathbb{Z}}$  série stationnaire d'ordre 2 lorsque :

- 1 son espérance  $m(t)$  est constante ;
- 2 sa covariance  $\gamma(s, t) = \text{cov}(X_s, X_t)$  est une fonction de  $|t - s|$ .

**Conséquence** : Si  $m(t)$  non constante, alors  $(X_t)$  admet une tendance et/ou une saisonnalité et est non stationnaire.

## Propriété

- (Stationnarité  $\implies$  Stationnarité d'ordre 2), mais réciproque fausse.
- Si  $(X_t)$  série gaussienne, (Stationnarité  $\iff$  Stationnarité d'ordre 2).

# Stationnarité

## Définition

On dit que  $(X_t)_{t \in \mathbb{Z}}$  série stationnaire d'ordre 2 lorsque :

- 1 son espérance  $m(t)$  est constante ;
- 2 sa covariance  $\gamma(s, t) = \text{cov}(X_s, X_t)$  est une fonction de  $|t - s|$ .

**Conséquence** : Si  $m(t)$  non constante, alors  $(X_t)$  admet une tendance et/ou une saisonnalité et est non stationnaire.

## Propriété

- (Stationnarité  $\implies$  Stationnarité d'ordre 2), mais réciproque fausse.
- Si  $(X_t)$  série gaussienne, (Stationnarité  $\iff$  Stationnarité d'ordre 2).

# Plan du cours

- 1 Définitions et premières propriétés
  - Séries chronologiques
  - Stationnarité
  - Première analyse statistique de la dépendance
- 2 Tendances et saisonnalités
  - Définition et propriétés
  - Estimation semi-paramétrique de la tendance et de la saisonnalité
  - Estimation non-paramétrique de la tendance et de la saisonnalité
- 3 Exemples de modèles de séries stationnaires
  - Processus ARMA
  - Processus GARCH
  - Autres exemples de séries temporelles
- 4 Estimation, sélection de modèle, test et prédiction
  - Estimation semi-paramétrique
  - Sélection de modèles et test d'adéquation
  - Prédiction

# Autocorrélation

## Définition

Soit  $(X_k)_{k \in \mathbb{Z}}$  série chronologique du second ordre stationnaire.

- $r(k) = \text{cov}(X_0, X_k)$  pour  $k \in \mathbb{Z}$  autocovariance de  $X$ ,  $r(0)$  variance.
- $\rho(k) = r(k)/r(0)$  autocorrélation de  $X$  et  $-1 \leq \rho(k) \leq 1$ ,  $k \in \mathbb{Z}$ .

**Conséquence** : Si  $\rho(k) \neq 0$  pour un  $k \neq 0$  alors  $(X_k)$  suite de v.a. non indépendantes.

Si  $(X_1, \dots, X_n)$  trajectoire observée, comment estimer  $r(k)$  et  $\rho(k)$  ?

# Autocorrélation

## Définition

Soit  $(X_k)_{k \in \mathbb{Z}}$  série chronologique du second ordre stationnaire.

- $r(k) = \text{cov}(X_0, X_k)$  pour  $k \in \mathbb{Z}$  **autocovariance** de  $X$ ,  $r(0)$  variance.
- $\rho(k) = r(k)/r(0)$  **autocorrélation** de  $X$  et  $-1 \leq \rho(k) \leq 1$ ,  $k \in \mathbb{Z}$ .

**Conséquence** : Si  $\rho(k) \neq 0$  pour un  $k \neq 0$  alors  $(X_k)$  suite de v.a. non indépendantes.

Si  $(X_1, \dots, X_n)$  trajectoire observée, comment estimer  $r(k)$  et  $\rho(k)$  ?

# Autocorrélation

## Définition

Soit  $(X_k)_{k \in \mathbb{Z}}$  série chronologique du second ordre stationnaire.

- $r(k) = \text{cov}(X_0, X_k)$  pour  $k \in \mathbb{Z}$  **autocovariance** de  $X$ ,  $r(0)$  variance.
- $\rho(k) = r(k)/r(0)$  **autocorrélation** de  $X$  et  $-1 \leq \rho(k) \leq 1$ ,  $k \in \mathbb{Z}$ .

**Conséquence** : Si  $\rho(k) \neq 0$  pour un  $k \neq 0$  alors  $(X_k)$  suite de v.a. non indépendantes.

Si  $(X_1, \dots, X_n)$  trajectoire observée, comment estimer  $r(k)$  et  $\rho(k)$  ?

# Autocorrélation

## Définition

Soit  $(X_k)_{k \in \mathbb{Z}}$  série chronologique du second ordre stationnaire.

- $r(k) = \text{cov}(X_0, X_k)$  pour  $k \in \mathbb{Z}$  **autocovariance** de  $X$ ,  $r(0)$  variance.
- $\rho(k) = r(k)/r(0)$  **autocorrélation** de  $X$  et  $-1 \leq \rho(k) \leq 1$ ,  $k \in \mathbb{Z}$ .

**Conséquence** : Si  $\rho(k) \neq 0$  pour un  $k \neq 0$  alors  $(X_k)$  suite de v.a. non indépendantes.

Si  $(X_1, \dots, X_n)$  trajectoire observée, comment estimer  $r(k)$  et  $\rho(k)$  ?

# Autocorrélation

## Définition

Soit  $(X_k)_{k \in \mathbb{Z}}$  série chronologique du second ordre stationnaire.

- $r(k) = \text{cov}(X_0, X_k)$  pour  $k \in \mathbb{Z}$  **autocovariance** de  $X$ ,  $r(0)$  variance.
- $\rho(k) = r(k)/r(0)$  **autocorrélation** de  $X$  et  $-1 \leq \rho(k) \leq 1$ ,  $k \in \mathbb{Z}$ .

**Conséquence** : Si  $\rho(k) \neq 0$  pour un  $k \neq 0$  alors  $(X_k)$  suite de v.a. non indépendantes.

Si  $(X_1, \dots, X_n)$  trajectoire observée, comment estimer  $r(k)$  et  $\rho(k)$  ?

# Moments empiriques

## Définition

Si  $(X_1, \dots, X_n)$  trajectoire observée de  $(X_t)$  série chronologique

- La moyenne empirique  $\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$  ;

- La variance empirique  $\overline{\sigma}_n^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X}_n)^2$  ;

Autre estimateur (biaisé) :  $\widehat{\sigma}_n^2 = \frac{1}{n} \sum_{i=1}^n X_i^2 - (\bar{X}_n)^2$  ;

- L'autocovariance empirique  $\hat{r}(k) = \frac{1}{n} \sum_{i=1}^{n-|k|} (X_i - \bar{X}_n)(X_{i+k} - \bar{X}_n)$  ;

- L'autocorrélation empirique (correlogram)  $\hat{\rho}(k) = \frac{\hat{r}(k)}{\hat{r}(0)} = \frac{\hat{r}(k)}{\widehat{\sigma}_n^2}$  ;

# Moments empiriques

## Définition

Si  $(X_1, \dots, X_n)$  trajectoire observée de  $(X_t)$  série chronologique

- La **moyenne empirique**  $\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$  ;

- La **variance empirique**  $\overline{\sigma}_n^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X}_n)^2$  ;

*Autre estimateur (biaisé) :  $\widehat{\sigma}_n^2 = \frac{1}{n} \sum_{i=1}^n X_i^2 - (\bar{X}_n)^2$  ;*

- L'**autocovariance empirique**  $\hat{r}(k) = \frac{1}{n} \sum_{i=1}^{n-|k|} (X_i - \bar{X}_n)(X_{i+k} - \bar{X}_n)$  ;

- L'**autocorrélation empirique (correlogram)**  $\hat{\rho}(k) = \frac{\hat{r}(k)}{\hat{r}(0)} = \frac{\hat{r}(k)}{\widehat{\sigma}_n^2}$  ;

# Moments empiriques

## Définition

Si  $(X_1, \dots, X_n)$  trajectoire observée de  $(X_t)$  série chronologique

- La **moyenne empirique**  $\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$  ;

- La **variance empirique**  $\overline{\sigma}_n^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X}_n)^2$  ;

*Autre estimateur (biaisé) :  $\widehat{\sigma}_n^2 = \frac{1}{n} \sum_{i=1}^n X_i^2 - (\bar{X}_n)^2$  ;*

- L'**autocovariance empirique**  $\hat{r}(k) = \frac{1}{n} \sum_{i=1}^{n-|k|} (X_i - \bar{X}_n)(X_{i+k} - \bar{X}_n)$  ;

- L'**autocorrélation empirique (correlogram)**  $\hat{\rho}(k) = \frac{\hat{r}(k)}{\hat{r}(0)} = \frac{\hat{r}(k)}{\widehat{\sigma}_n^2}$  ;

# Moments empiriques

## Définition

Si  $(X_1, \dots, X_n)$  trajectoire observée de  $(X_t)$  série chronologique

- La **moyenne empirique**  $\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$  ;

- La **variance empirique**  $\overline{\sigma}_n^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X}_n)^2$  ;

Autre estimateur (biaisé) :  $\widehat{\sigma}_n^2 = \frac{1}{n} \sum_{i=1}^n X_i^2 - (\bar{X}_n)^2$  ;

- L'autocovariance empirique  $\hat{r}(k) = \frac{1}{n} \sum_{i=1}^{n-|k|} (X_i - \bar{X}_n)(X_{i+k} - \bar{X}_n)$  ;

- L'autocorrélation empirique (correlogram)  $\hat{\rho}(k) = \frac{\hat{r}(k)}{\hat{r}(0)} = \frac{\hat{r}(k)}{\widehat{\sigma}_n^2}$  ;

# Moments empiriques

## Définition

Si  $(X_1, \dots, X_n)$  trajectoire observée de  $(X_t)$  série chronologique

- La **moyenne empirique**  $\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$  ;

- La **variance empirique**  $\overline{\sigma}_n^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X}_n)^2$  ;

Autre estimateur (biaisé) :  $\widehat{\sigma}_n^2 = \frac{1}{n} \sum_{i=1}^n X_i^2 - (\bar{X}_n)^2$  ;

- L'**autocovariance empirique**  $\hat{r}(k) = \frac{1}{n} \sum_{i=1}^{n-|k|} (X_i - \bar{X}_n)(X_{i+k} - \bar{X}_n)$  ;

- L'**autocorrélation empirique (correlogram)**  $\hat{\rho}(k) = \frac{\hat{r}(k)}{\hat{r}(0)} = \frac{\hat{r}(k)}{\widehat{\sigma}_n^2}$  ;

# Moments empiriques

## Définition

Si  $(X_1, \dots, X_n)$  trajectoire observée de  $(X_t)$  série chronologique

- La **moyenne empirique**  $\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$  ;

- La **variance empirique**  $\overline{\sigma}_n^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X}_n)^2$  ;

*Autre estimateur (biaisé) :  $\widehat{\sigma}_n^2 = \frac{1}{n} \sum_{i=1}^n X_i^2 - (\bar{X}_n)^2$  ;*

- L'**autocovariance empirique**  $\hat{r}(k) = \frac{1}{n} \sum_{i=1}^{n-|k|} (X_i - \bar{X}_n)(X_{i+k} - \bar{X}_n)$  ;

- L'**autocorrélation empirique (correlogram)**  $\hat{\rho}(k) = \frac{\hat{r}(k)}{\hat{r}(0)} = \frac{\hat{r}(k)}{\widehat{\sigma}_n^2}$  ;

# Convergence pour une suite de v.a.i.i.d.

## Propriété

Si  $(X_k)_{k \in \mathbb{Z}}$  est une suite de v.a.i.i.d. avec  $\mathbb{E}[X_0^2] < \infty$ ,

①  $\bar{X}_n \xrightarrow[n \rightarrow +\infty]{\mathcal{P}} m = \mathbb{E}[X_0]$  et  $\overline{\sigma_n^2}$  ou  $\widehat{\sigma_n^2} \xrightarrow[n \rightarrow +\infty]{\mathcal{P}} \sigma^2 = \text{var}(X_0)$ ;

②  $\widehat{r}(0) = \widehat{\sigma_n^2} \xrightarrow[n \rightarrow +\infty]{\mathcal{P}} \sigma^2$  et  $\widehat{r}(k) \xrightarrow[n \rightarrow +\infty]{\mathcal{P}} 0$  pour tout  $k \neq 0$ ;

③  $\widehat{\rho}(0) = 1$  et  $\widehat{\rho}(k) \xrightarrow[n \rightarrow +\infty]{\mathcal{P}} 0$  pour tout  $k \neq 0$ .

## Attention !

- Si  $(X_k)$  série stationnaire avec v.a. non indépendantes, les convergences peuvent être vraies ... ou fausses (exemple  $X_k = X_0$  pour tout  $k$ )!
- Si  $(X_k)$  série non stationnaire, les convergences peuvent être vraies ... ou fausses, même si indépendance !

# Convergence pour une suite de v.a.i.i.d.

## Propriété

Si  $(X_k)_{k \in \mathbb{Z}}$  est une suite de v.a.i.i.d. avec  $\mathbb{E}[X_0^2] < \infty$ ,

①  $\bar{X}_n \xrightarrow[n \rightarrow +\infty]{\mathcal{P}} m = \mathbb{E}[X_0]$  et  $\overline{\sigma_n^2}$  ou  $\widehat{\sigma_n^2} \xrightarrow[n \rightarrow +\infty]{\mathcal{P}} \sigma^2 = \text{var}(X_0)$  ;

②  $\widehat{r}(0) = \widehat{\sigma_n^2} \xrightarrow[n \rightarrow +\infty]{\mathcal{P}} \sigma^2$  et  $\widehat{r}(k) \xrightarrow[n \rightarrow +\infty]{\mathcal{P}} 0$  pour tout  $k \neq 0$  ;

③  $\widehat{\rho}(0) = 1$  et  $\widehat{\rho}(k) \xrightarrow[n \rightarrow +\infty]{\mathcal{P}} 0$  pour tout  $k \neq 0$ .

## Attention !

- Si  $(X_k)$  série stationnaire avec v.a. non indépendantes, les convergences peuvent être vraies ... ou fausses (exemple  $X_k = X_0$  pour tout  $k$ ) !
- Si  $(X_k)$  série non stationnaire, les convergences peuvent être vraies ... ou fausses, même si indépendance !

# Convergence pour une suite de v.a.i.i.d.

## Propriété

Si  $(X_k)_{k \in \mathbb{Z}}$  est une suite de v.a.i.i.d. avec  $\mathbb{E}[X_0^2] < \infty$ ,

$$\textcircled{1} \quad \bar{X}_n \xrightarrow[n \rightarrow +\infty]{\mathcal{P}} m = \mathbb{E}[X_0] \quad \text{et} \quad \overline{\sigma_n^2} \text{ ou } \widehat{\sigma_n^2} \xrightarrow[n \rightarrow +\infty]{\mathcal{P}} \sigma^2 = \text{var}(X_0);$$

$$\textcircled{2} \quad \widehat{r}(0) = \widehat{\sigma_n^2} \xrightarrow[n \rightarrow +\infty]{\mathcal{P}} \sigma^2 \quad \text{et} \quad \widehat{r}(k) \xrightarrow[n \rightarrow +\infty]{\mathcal{P}} 0 \text{ pour tout } k \neq 0;$$

$$\textcircled{3} \quad \widehat{\rho}(0) = 1 \quad \text{et} \quad \widehat{\rho}(k) \xrightarrow[n \rightarrow +\infty]{\mathcal{P}} 0 \text{ pour tout } k \neq 0.$$

## Attention !

- Si  $(X_k)$  série stationnaire avec v.a. non indépendantes, les convergences peuvent être vraies ... ou fausses (exemple  $X_k = X_0$  pour tout  $k$ )!
- Si  $(X_k)$  série non stationnaire, les convergences peuvent être vraies ... ou fausses, même si indépendance !

# Convergence pour une suite de v.a.i.i.d.

## Propriété

Si  $(X_k)_{k \in \mathbb{Z}}$  est une suite de v.a.i.i.d. avec  $\mathbb{E}[X_0^2] < \infty$ ,

$$\textcircled{1} \quad \bar{X}_n \xrightarrow[n \rightarrow +\infty]{\mathcal{P}} m = \mathbb{E}[X_0] \quad \text{et} \quad \overline{\sigma_n^2} \text{ ou } \widehat{\sigma_n^2} \xrightarrow[n \rightarrow +\infty]{\mathcal{P}} \sigma^2 = \text{var}(X_0);$$

$$\textcircled{2} \quad \widehat{r}(0) = \widehat{\sigma_n^2} \xrightarrow[n \rightarrow +\infty]{\mathcal{P}} \sigma^2 \quad \text{et} \quad \widehat{r}(k) \xrightarrow[n \rightarrow +\infty]{\mathcal{P}} 0 \text{ pour tout } k \neq 0;$$

$$\textcircled{3} \quad \widehat{\rho}(0) = 1 \quad \text{et} \quad \widehat{\rho}(k) \xrightarrow[n \rightarrow +\infty]{\mathcal{P}} 0 \text{ pour tout } k \neq 0.$$

## Attention !

- Si  $(X_k)$  série stationnaire avec v.a. non indépendantes, les convergences peuvent être vraies ... ou fausses (exemple  $X_k = X_0$  pour tout  $k$ )!
- Si  $(X_k)$  série non stationnaire, les convergences peuvent être vraies ... ou fausses, même si indépendance !

# Convergence pour une suite de v.a.i.i.d.

## Propriété

Si  $(X_k)_{k \in \mathbb{Z}}$  est une suite de v.a.i.i.d. avec  $\mathbb{E}[X_0^2] < \infty$ ,

$$\textcircled{1} \quad \bar{X}_n \xrightarrow[n \rightarrow +\infty]{\mathcal{P}} m = \mathbb{E}[X_0] \quad \text{et} \quad \overline{\sigma_n^2} \text{ ou } \widehat{\sigma_n^2} \xrightarrow[n \rightarrow +\infty]{\mathcal{P}} \sigma^2 = \text{var}(X_0);$$

$$\textcircled{2} \quad \widehat{r}(0) = \widehat{\sigma_n^2} \xrightarrow[n \rightarrow +\infty]{\mathcal{P}} \sigma^2 \quad \text{et} \quad \widehat{r}(k) \xrightarrow[n \rightarrow +\infty]{\mathcal{P}} 0 \text{ pour tout } k \neq 0;$$

$$\textcircled{3} \quad \widehat{\rho}(0) = 1 \quad \text{et} \quad \widehat{\rho}(k) \xrightarrow[n \rightarrow +\infty]{\mathcal{P}} 0 \text{ pour tout } k \neq 0.$$

## Attention !

- Si  $(X_k)$  série stationnaire avec v.a. non indépendantes, les convergences peuvent être vraies ... ou fausses (exemple  $X_k = X_0$  pour tout  $k$ )!

- Si  $(X_k)$  série non stationnaire, les convergences peuvent être vraies ... ou fausses, même si indépendance!

# Convergence pour une suite de v.a.i.i.d.

## Propriété

Si  $(X_k)_{k \in \mathbb{Z}}$  est une suite de v.a.i.i.d. avec  $\mathbb{E}[X_0^2] < \infty$ ,

$$\textcircled{1} \quad \bar{X}_n \xrightarrow[n \rightarrow +\infty]{\mathcal{P}} m = \mathbb{E}[X_0] \quad \text{et} \quad \overline{\sigma_n^2} \text{ ou } \widehat{\sigma_n^2} \xrightarrow[n \rightarrow +\infty]{\mathcal{P}} \sigma^2 = \text{var}(X_0);$$

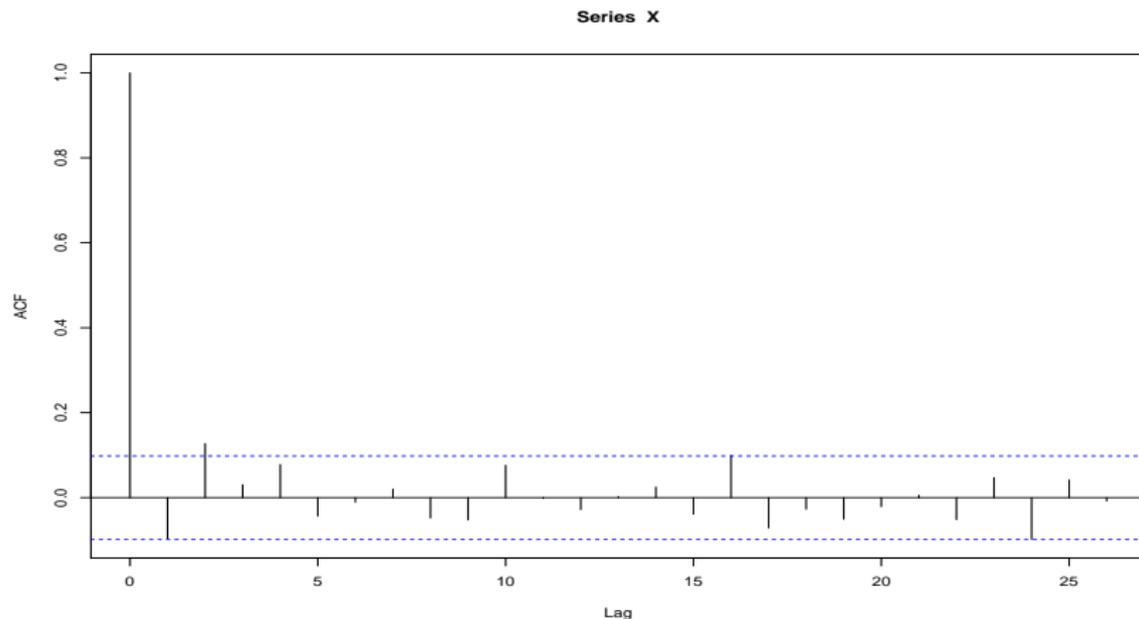
$$\textcircled{2} \quad \widehat{r}(0) = \widehat{\sigma_n^2} \xrightarrow[n \rightarrow +\infty]{\mathcal{P}} \sigma^2 \quad \text{et} \quad \widehat{r}(k) \xrightarrow[n \rightarrow +\infty]{\mathcal{P}} 0 \text{ pour tout } k \neq 0;$$

$$\textcircled{3} \quad \widehat{\rho}(0) = 1 \quad \text{et} \quad \widehat{\rho}(k) \xrightarrow[n \rightarrow +\infty]{\mathcal{P}} 0 \text{ pour tout } k \neq 0.$$

## Attention !

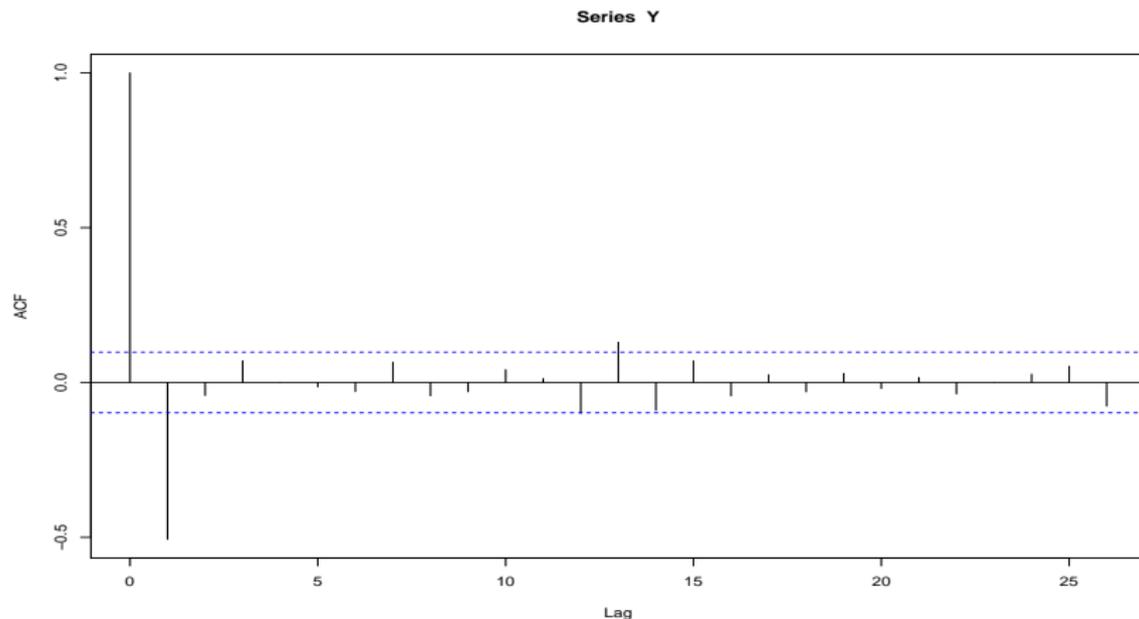
- Si  $(X_k)$  série stationnaire avec v.a. non indépendantes, les convergences peuvent être vraies ... ou fausses (exemple  $X_k = X_0$  pour tout  $k$ )!
- Si  $(X_k)$  série non stationnaire, les convergences peuvent être vraies ... ou fausses, même si indépendance !

# ACF



ACF (Auto-Correlogram Function) d'une suite  $(\varepsilon_k)$  de v.a.i.i.d. pour  $n = 400$

# ACF



ACF d'une suite de v.a.  $(X_k)$  telle que  $X_k = \varepsilon_k - \varepsilon_{k-1}$  pour  $n = 400$

## Vers un test de "blancheur"

ACF intéressant pour visualiser une dépendance mais difficile de conclure.

### Définition

- Un bruit blanc (fort) est une suite de v.a.i.i.d. centrées.
- Un bruit blanc faible est une suite de v.a.i.i.d. centrées non corrélées.

On va tester si  $(X_k)$  est un bruit blanc (ou même une suite de v.a.i.i.d.)

### Théorème

Soit  $X = (X_t)_{t \in \mathbb{Z}}$  suite de v.a.i.i.d. de variance  $0 < \sigma_X^2 < \infty$ . Pour  $K_{\max}$  fixé,

$$\begin{aligned} \sqrt{n} (\hat{r}_X(k))_{1 \leq k \leq K_{\max}} &\xrightarrow[n \rightarrow \infty]{\mathcal{L}} \mathcal{N}_{K_{\max}}(0, \sigma_X^4 I_{K_{\max}}) \\ \implies \sqrt{n} (\hat{\rho}_X(k))_{1 \leq k \leq K_{\max}} &\xrightarrow[n \rightarrow \infty]{\mathcal{L}} \mathcal{N}_{K_{\max}}(0, I_{K_{\max}}). \end{aligned}$$

## Vers un test de "blancheur"

ACF intéressant pour visualiser une dépendance mais difficile de conclure.

### Définition

- *Un bruit blanc (fort) est une suite de v.a.i.i.d. centrées.*
- *Un bruit blanc faible est une suite de v.a.i.i.d. centrées non corrélées.*

On va tester si  $(X_k)$  est un bruit blanc (ou même une suite de v.a.i.i.d.)

### Théorème

Soit  $X = (X_t)_{t \in \mathbb{Z}}$  suite de v.a.i.i.d. de variance  $0 < \sigma_X^2 < \infty$ . Pour  $K_{\max}$  fixé,

$$\begin{aligned} \sqrt{n} (\hat{r}_X(k))_{1 \leq k \leq K_{\max}} &\xrightarrow[n \rightarrow \infty]{\mathcal{L}} \mathcal{N}_{K_{\max}}(0, \sigma_X^4 I_{K_{\max}}) \\ \implies \sqrt{n} (\hat{\rho}_X(k))_{1 \leq k \leq K_{\max}} &\xrightarrow[n \rightarrow \infty]{\mathcal{L}} \mathcal{N}_{K_{\max}}(0, I_{K_{\max}}). \end{aligned}$$

## Vers un test de "blancheur"

ACF intéressant pour visualiser une dépendance mais difficile de conclure.

### Définition

- *Un bruit blanc (fort) est une suite de v.a.i.i.d. centrées.*
- *Un bruit blanc faible est une suite de v.a.i.i.d. centrées non corrélées.*

On va tester si  $(X_k)$  est un bruit blanc (ou même une suite de v.a.i.i.d.)

### Théorème

Soit  $X = (X_t)_{t \in \mathbb{Z}}$  suite de v.a.i.i.d. de variance  $0 < \sigma_X^2 < \infty$ . Pour  $K_{\max}$  fixé,

$$\begin{aligned} \sqrt{n} (\hat{r}_X(k))_{1 \leq k \leq K_{\max}} &\xrightarrow[n \rightarrow \infty]{\mathcal{L}} \mathcal{N}_{K_{\max}}(0, \sigma_X^4 I_{K_{\max}}) \\ \implies \sqrt{n} (\hat{\rho}_X(k))_{1 \leq k \leq K_{\max}} &\xrightarrow[n \rightarrow \infty]{\mathcal{L}} \mathcal{N}_{K_{\max}}(0, I_{K_{\max}}). \end{aligned}$$

## Vers un test de "blancheur"

ACF intéressant pour visualiser une dépendance mais difficile de conclure.

### Définition

- Un bruit blanc (fort) est une suite de v.a.i.i.d. centrées.
- Un bruit blanc faible est une suite de v.a.i.i.d. centrées non corrélées.

On va tester si  $(X_k)$  est un bruit blanc (ou même une suite de v.a.i.i.d.)

### Théorème

Soit  $X = (X_t)_{t \in \mathbb{Z}}$  suite de v.a.i.i.d. de variance  $0 < \sigma_X^2 < \infty$ . Pour  $K_{\max}$  fixé,

$$\begin{aligned} \sqrt{n} (\hat{r}_X(k))_{1 \leq k \leq K_{\max}} &\xrightarrow[n \rightarrow \infty]{\mathcal{L}} \mathcal{N}_{K_{\max}}(0, \sigma_X^4 I_{K_{\max}}) \\ \implies \sqrt{n} (\hat{\rho}_X(k))_{1 \leq k \leq K_{\max}} &\xrightarrow[n \rightarrow \infty]{\mathcal{L}} \mathcal{N}_{K_{\max}}(0, I_{K_{\max}}). \end{aligned}$$

## Vers un test de "blancheur"

ACF intéressant pour visualiser une dépendance mais difficile de conclure.

### Définition

- Un bruit blanc (fort) est une suite de v.a.i.i.d. centrées.
- Un bruit blanc faible est une suite de v.a.i.i.d. centrées non corrélées.

On va tester si  $(X_k)$  est un bruit blanc (ou même une suite de v.a.i.i.d.)

### Théorème

Soit  $X = (X_t)_{t \in \mathbb{Z}}$  suite de v.a.i.i.d. de variance  $0 < \sigma_X^2 < \infty$ . Pour  $K_{\max}$  fixé,

$$\begin{aligned} \sqrt{n} (\hat{r}_X(k))_{1 \leq k \leq K_{\max}} &\xrightarrow[n \rightarrow \infty]{\mathcal{L}} \mathcal{N}_{K_{\max}}(0, \sigma_X^4 I_{K_{\max}}) \\ \implies \sqrt{n} (\hat{\rho}_X(k))_{1 \leq k \leq K_{\max}} &\xrightarrow[n \rightarrow \infty]{\mathcal{L}} \mathcal{N}_{K_{\max}}(0, I_{K_{\max}}). \end{aligned}$$

## Vers un test de "blancheur"

ACF intéressant pour visualiser une dépendance mais difficile de conclure.

### Définition

- Un bruit blanc (fort) est une suite de v.a.i.i.d. centrées.
- Un bruit blanc faible est une suite de v.a.i.i.d. centrées non corrélées.

On va tester si  $(X_k)$  est un bruit blanc (ou même une suite de v.a.i.i.d.)

### Théorème

Soit  $X = (X_t)_{t \in \mathbb{Z}}$  suite de v.a.i.i.d. de variance  $0 < \sigma_X^2 < \infty$ . Pour  $K_{\max}$  fixé,

$$\begin{aligned} \sqrt{n} (\widehat{r}_X(k))_{1 \leq k \leq K_{\max}} &\xrightarrow[n \rightarrow \infty]{\mathcal{L}} \mathcal{N}_{K_{\max}}(0, \sigma_X^4 I_{K_{\max}}) \\ \implies \sqrt{n} (\widehat{\rho}_X(k))_{1 \leq k \leq K_{\max}} &\xrightarrow[n \rightarrow \infty]{\mathcal{L}} \mathcal{N}_{K_{\max}}(0, I_{K_{\max}}). \end{aligned}$$

## Vers un test de "blancheur"

ACF intéressant pour visualiser une dépendance mais difficile de conclure.

### Définition

- Un bruit blanc (fort) est une suite de v.a.i.i.d. centrées.
- Un bruit blanc faible est une suite de v.a.i.i.d. centrées non corrélées.

On va tester si  $(X_k)$  est un bruit blanc (ou même une suite de v.a.i.i.d.)

### Théorème

Soit  $X = (X_t)_{t \in \mathbb{Z}}$  suite de v.a.i.i.d. de variance  $0 < \sigma_X^2 < \infty$ . Pour  $K_{\max}$  fixé,

$$\begin{aligned} \sqrt{n} (\widehat{r}_X(k))_{1 \leq k \leq K_{\max}} &\xrightarrow[n \rightarrow \infty]{\mathcal{L}} \mathcal{N}_{K_{\max}}(0, \sigma_X^4 I_{K_{\max}}) \\ \implies \sqrt{n} (\widehat{\rho}_X(k))_{1 \leq k \leq K_{\max}} &\xrightarrow[n \rightarrow \infty]{\mathcal{L}} \mathcal{N}_{K_{\max}}(0, I_{K_{\max}}). \end{aligned}$$

# Test portemanteau

## Théorème

Soit  $X = (X_t)_{t \in \mathbb{Z}}$  série stationnaire du second ordre. On désire tester :

- $$\begin{cases} H_0 : (X_t) \text{ forme une suite de variables indépendantes} \\ H_1 : (X_t) \text{ ne forme pas une suite de variables indépendantes} \end{cases}$$

On définit deux statistique de test :

- La statistique de Box-Pierce :  $\hat{T}_{BP} = n \sum_{k=1}^{K_{\max}} \hat{\rho}_X^2(k)$  ;
- La statistique de Ljung-Box :  $\hat{T}_{LB} = n(n+2) \sum_{k=1}^{K_{\max}} \frac{\hat{\rho}_X^2(k)}{n-k}$ .

Dans les deux cas, on a

$$\begin{matrix} \hat{T}_{BP} \\ \hat{T}_{LB} \end{matrix} \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \chi^2_{K_{\max}}$$

# Test portemanteau

## Théorème

Soit  $X = (X_t)_{t \in \mathbb{Z}}$  série stationnaire du second ordre. On désire tester :

$$\begin{cases} H_0 : (X_t) \text{ forme une suite de variables indépendantes} \\ H_1 : (X_t) \text{ ne forme pas une suite de variables indépendantes} \end{cases}$$

On définit deux statistique de test :

- La statistique de Box-Pierce :  $\hat{T}_{BP} = n \sum_{k=1}^{K_{\max}} \hat{\rho}_X^2(k)$  ;
- La statistique de Ljung-Box :  $\hat{T}_{LB} = n(n+2) \sum_{k=1}^{K_{\max}} \frac{\hat{\rho}_X^2(k)}{n-k}$ .

Dans les deux cas, on a

$$\begin{matrix} \hat{T}_{BP} \\ \hat{T}_{LB} \end{matrix} \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \chi_{K_{\max}}^2$$

# Test portemanteau

## Théorème

Soit  $X = (X_t)_{t \in \mathbb{Z}}$  série stationnaire du second ordre. On désire tester :

$$\begin{cases} H_0 : (X_t) \text{ forme une suite de variables indépendantes} \\ H_1 : (X_t) \text{ ne forme pas une suite de variables indépendantes} \end{cases}$$

On définit deux statistique de test :

- La statistique de Box-Pierce :  $\hat{T}_{BP} = n \sum_{k=1}^{K_{\max}} \hat{\rho}_X^2(k)$  ;
- La statistique de Ljung-Box :  $\hat{T}_{LB} = n(n+2) \sum_{k=1}^{K_{\max}} \frac{\hat{\rho}_X^2(k)}{n-k}$ .

Dans les deux cas, on a

$$\begin{matrix} \hat{T}_{BP} \\ \hat{T}_{LB} \end{matrix} \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \chi_{K_{\max}}^2$$

# Test portemanteau

## Théorème

Soit  $X = (X_t)_{t \in \mathbb{Z}}$  série stationnaire du second ordre. On désire tester :

$$\begin{cases} H_0 : (X_t) \text{ forme une suite de variables indépendantes} \\ H_1 : (X_t) \text{ ne forme pas une suite de variables indépendantes} \end{cases}$$

On définit deux statistique de test :

- La statistique de Box-Pierce :  $\hat{T}_{BP} = n \sum_{k=1}^{K_{\max}} \hat{\rho}_X^2(k)$ ;

- La statistique de Ljung-Box :  $\hat{T}_{LB} = n(n+2) \sum_{k=1}^{K_{\max}} \frac{\hat{\rho}_X^2(k)}{n-k}$ .

Dans les deux cas, on a

$$\begin{matrix} \hat{T}_{BP} \\ \hat{T}_{LB} \end{matrix} \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \chi_{K_{\max}}^2$$

# Test portemanteau

## Théorème

Soit  $X = (X_t)_{t \in \mathbb{Z}}$  série stationnaire du second ordre. On désire tester :

$$\begin{cases} H_0 : (X_t) \text{ forme une suite de variables indépendantes} \\ H_1 : (X_t) \text{ ne forme pas une suite de variables indépendantes} \end{cases}$$

On définit deux statistique de test :

- La statistique de Box-Pierce :  $\hat{T}_{BP} = n \sum_{k=1}^{K_{\max}} \hat{\rho}_X^2(k)$  ;
- La statistique de Ljung-Box :  $\hat{T}_{LB} = n(n+2) \sum_{k=1}^{K_{\max}} \frac{\hat{\rho}_X^2(k)}{n-k}$ .

Dans les deux cas, on a

$$\begin{matrix} \hat{T}_{BP} \\ \hat{T}_{LB} \end{matrix} \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \chi_{K_{\max}}^2$$

# Test portemanteau

## Théorème

Soit  $X = (X_t)_{t \in \mathbb{Z}}$  série stationnaire du second ordre. On désire tester :

- $$\begin{cases} H_0 : (X_t) \text{ forme une suite de variables indépendantes} \\ H_1 : (X_t) \text{ ne forme pas une suite de variables indépendantes} \end{cases}$$

On définit deux statistique de test :

- La statistique de Box-Pierce :  $\hat{T}_{BP} = n \sum_{k=1}^{K_{\max}} \hat{\rho}_X^2(k)$  ;
- La statistique de Ljung-Box :  $\hat{T}_{LB} = n(n+2) \sum_{k=1}^{K_{\max}} \frac{\hat{\rho}_X^2(k)}{n-k}$ .

Dans les deux cas, on a

$$\begin{matrix} \hat{T}_{BP} \\ \hat{T}_{LB} \end{matrix} \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \chi_{K_{\max}}^2$$

# Test portemanteau

## Théorème

Soit  $X = (X_t)_{t \in \mathbb{Z}}$  série stationnaire du second ordre. On désire tester :

- $$\begin{cases} H_0 : (X_t) \text{ forme une suite de variables indépendantes} \\ H_1 : (X_t) \text{ ne forme pas une suite de variables indépendantes} \end{cases}$$

On définit deux statistique de test :

- La statistique de Box-Pierce :  $\hat{T}_{BP} = n \sum_{k=1}^{K_{\max}} \hat{\rho}_X^2(k)$  ;
- La statistique de Ljung-Box :  $\hat{T}_{LB} = n(n+2) \sum_{k=1}^{K_{\max}} \frac{\hat{\rho}_X^2(k)}{n-k}$ .

Dans les deux cas, on a

$$\begin{matrix} \hat{T}_{BP} \\ \hat{T}_{LB} \end{matrix} \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \chi_{K_{\max}}^2.$$

# Test portemanteau

## Choix de $K_{\max}$ :

- Théoriquement il faut  $K_{\max} = o(\sqrt{n})$ .
- Empiriquement, on utilisera  $K_{\max} = 5$  si  $n$  ordre de la centaine,  $K_{\max} = 10$  pour ordre du millier

**Attention !** Si une tendance ou saisonnalité est présente, aucune conclusion n'est possible !

$\implies$  On commence par estimer et retirer une éventuelle tendance ou saisonnalité

# Test portemanteau

## Choix de $K_{\max}$ :

- Théoriquement il faut  $K_{\max} = o(\sqrt{n})$ .
- Empiriquement, on utilisera  $K_{\max} = 5$  si  $n$  ordre de la centaine,  $K_{\max} = 10$  pour ordre du millier

**Attention !** Si une tendance ou saisonnalité est présente, aucune conclusion n'est possible !

$\implies$  On commence par estimer et retirer une éventuelle tendance ou saisonnalité

# Test portemanteau

## Choix de $K_{\max}$ :

- Théoriquement il faut  $K_{\max} = o(\sqrt{n})$ .
- Empiriquement, on utilisera  $K_{\max} = 5$  si  $n$  ordre de la centaine,  $K_{\max} = 10$  pour ordre du millier

**Attention !** Si une tendance ou saisonnalité est présente, aucune conclusion n'est possible !

$\implies$  On commence par estimer et retirer une éventuelle tendance ou saisonnalité

# Test portemanteau

## Choix de $K_{\max}$ :

- Théoriquement il faut  $K_{\max} = o(\sqrt{n})$ .
- Empiriquement, on utilisera  $K_{\max} = 5$  si  $n$  ordre de la centaine,  $K_{\max} = 10$  pour ordre du millier

**Attention !** Si une tendance ou saisonnalité est présente, aucune conclusion n'est possible !

⇒ On commence par estimer et retirer une éventuelle tendance ou saisonnalité

# Test portemanteau

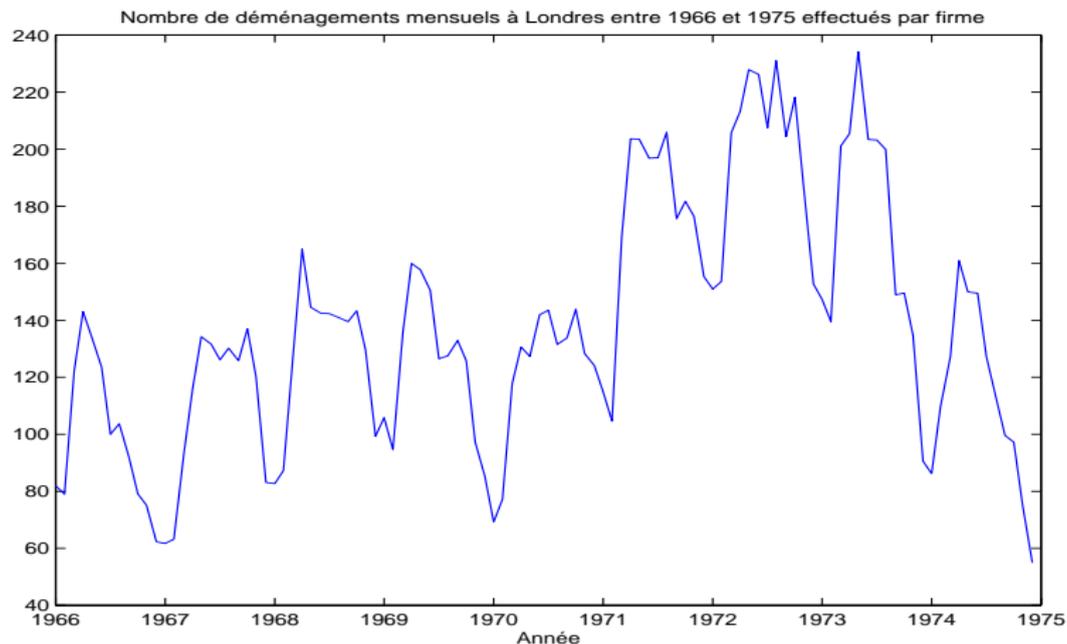
## Choix de $K_{\max}$ :

- Théoriquement il faut  $K_{\max} = o(\sqrt{n})$ .
- Empiriquement, on utilisera  $K_{\max} = 5$  si  $n$  ordre de la centaine,  $K_{\max} = 10$  pour ordre du millier

**Attention !** Si une tendance ou saisonnalité est présente, aucune conclusion n'est possible !

$\implies$  On commence par estimer et retirer une éventuelle tendance ou saisonnalité

# Séries non stationnaires



⇒ On commence par estimer et retirer tendance ou saisonnalité

# Plan du cours

- 1 Définitions et premières propriétés
  - Séries chronologiques
  - Stationnarité
  - Première analyse statistique de la dépendance
- 2 Tendances et saisonnalités
  - Définition et propriétés
  - Estimation semi-paramétrique de la tendance et de la saisonnalité
  - Estimation non-paramétrique de la tendance et de la saisonnalité
- 3 Exemples de modèles de séries stationnaires
  - Processus ARMA
  - Processus GARCH
  - Autres exemples de séries temporelles
- 4 Estimation, sélection de modèle, test et prédiction
  - Estimation semi-paramétrique
  - Sélection de modèles et test d'adéquation
  - Prédiction

## Définition

Une série temporelle  $X = (X_t)_{t \in \mathbb{Z}}$  peut toujours s'écrire sous la forme

$$X_t = a(t) + s(t) + u_t, \quad \text{pour } t \in \mathbb{Z},$$

$t \mapsto a(t)$  et  $t \mapsto s(t)$  fonctions déterministes avec  $\mathbb{E}[X_t] = a(t) + s(t)$ , et

- 1  $t \mapsto s(t)$  fonction périodique de période  $r > 0$  telle que  $\sum_{i=1}^r s(i) = 0$ ,  
alors, si  $s$  non nulle,  $s$  est la **saisonnalité** de  $X$ ,
- 2 si  $t \mapsto a(t)$  est non nulle,  $a$  est la **tendance** de  $X$ ,
- 3  $u = (u_t)_t$  tel que  $\mathbb{E}[u_t] = 0$ , **bruit** de  $X$ .

## Exemple

Tendance polynômiale :  $a(t) = a_0 + a_1 t + \dots + a_k t^k$ , saisonnalité annuelle :  
 $s(t) = \sum_{i=1}^{12} s_i \mathbb{1}_{t=i} \text{ [12]}$  ou  $s(t) = \alpha \sin(2\pi t/12) + \beta \cos(2\pi t/12), \dots$

## Définition

Une série temporelle  $X = (X_t)_{t \in \mathbb{Z}}$  peut toujours s'écrire sous la forme

$$X_t = a(t) + s(t) + u_t, \quad \text{pour } t \in \mathbb{Z},$$

$t \mapsto a(t)$  et  $t \mapsto s(t)$  fonctions déterministes avec  $\mathbf{E}[X_t] = a(t) + s(t)$ , et

- 1  $t \mapsto s(t)$  fonction périodique de période  $r > 0$  telle que  $\sum_{i=1}^r s(i) = 0$ ,  
alors, si  $s$  non nulle,  $s$  est la **saisonnalité** de  $X$ ,
- 2 si  $t \mapsto a(t)$  est non nulle,  $a$  est la **tendance** de  $X$ ,
- 3  $u = (u_t)_t$  tel que  $\mathbf{E}[u_t] = 0$ , **bruit** de  $X$ .

## Exemple

Tendance polynômiale :  $a(t) = a_0 + a_1 t + \dots + a_k t^k$ , saisonnalité annuelle :  
 $s(t) = \sum_{i=1}^{12} s_i \mathbf{1}_{t=i} \text{ [12]}$  ou  $s(t) = \alpha \sin(2\pi t/12) + \beta \cos(2\pi t/12), \dots$

## Définition

Une série temporelle  $X = (X_t)_{t \in \mathbb{Z}}$  peut toujours s'écrire sous la forme

$$X_t = a(t) + s(t) + u_t, \quad \text{pour } t \in \mathbb{Z},$$

$t \mapsto a(t)$  et  $t \mapsto s(t)$  fonctions déterministes avec  $\mathbf{E}[X_t] = a(t) + s(t)$ , et

- 1  $t \mapsto s(t)$  fonction périodique de période  $r > 0$  telle que  $\sum_{i=1}^r s(i) = 0$ ,  
alors, si  $s$  non nulle,  $s$  est la **saisonnalité** de  $X$ ,
- 2 si  $t \mapsto a(t)$  est non nulle,  $a$  est la **tendance** de  $X$ ,
- 3  $u = (u_t)_t$  tel que  $\mathbf{E}[u_t] = 0$ , **bruit** de  $X$ .

## Exemple

Tendance polynômiale :  $a(t) = a_0 + a_1 t + \dots + a_k t^k$ , saisonnalité annuelle :  
 $s(t) = \sum_{i=1}^{12} s_i \mathbf{1}_{t=i} \text{ [12]}$  ou  $s(t) = \alpha \sin(2\pi t/12) + \beta \cos(2\pi t/12), \dots$

## Définition

Une série temporelle  $X = (X_t)_{t \in \mathbb{Z}}$  peut toujours s'écrire sous la forme

$$X_t = a(t) + s(t) + u_t, \quad \text{pour } t \in \mathbb{Z},$$

$t \mapsto a(t)$  et  $t \mapsto s(t)$  fonctions déterministes avec  $\mathbf{E}[X_t] = a(t) + s(t)$ , et

- 1  $t \mapsto s(t)$  fonction périodique de période  $r > 0$  telle que  $\sum_{i=1}^r s(i) = 0$ ,  
alors, si  $s$  non nulle,  $s$  est la **saisonnalité** de  $X$ ,
- 2 si  $t \mapsto a(t)$  est non nulle,  $a$  est la **tendance** de  $X$ ,
- 3  $u = (u_t)_t$  tel que  $\mathbf{E}[u_t] = 0$ , **bruit** de  $X$ .

## Exemple

Tendance polynômiale :  $a(t) = a_0 + a_1 t + \dots + a_k t^k$ , saisonnalité annuelle :  
 $s(t) = \sum_{i=1}^{12} s_i \mathbf{1}_{t=i} \text{ [12]}$  ou  $s(t) = \alpha \sin(2\pi t/12) + \beta \cos(2\pi t/12), \dots$

## Définition

Une série temporelle  $X = (X_t)_{t \in \mathbb{Z}}$  peut toujours s'écrire sous la forme

$$X_t = a(t) + s(t) + u_t, \quad \text{pour } t \in \mathbb{Z},$$

$t \mapsto a(t)$  et  $t \mapsto s(t)$  fonctions déterministes avec  $\mathbf{E}[X_t] = a(t) + s(t)$ , et

- 1  $t \mapsto s(t)$  fonction périodique de période  $r > 0$  telle que  $\sum_{i=1}^r s(i) = 0$ ,  
alors, si  $s$  non nulle,  $s$  est la **saisonnalité** de  $X$ ,
- 2 si  $t \mapsto a(t)$  est non nulle,  $a$  est la **tendance** de  $X$ ,
- 3  $u = (u_t)_t$  tel que  $\mathbf{E}[u_t] = 0$ , **bruit** de  $X$ .

## Exemple

Tendance polynômiale :  $a(t) = a_0 + a_1 t + \dots + a_k t^k$ , saisonnalité annuelle :  
 $s(t) = \sum_{i=1}^{12} s_i \mathbf{1}_{t=i} \text{ [12]}$  ou  $s(t) = \alpha \sin(2\pi t/12) + \beta \cos(2\pi t/12), \dots$

## Définition

Une série temporelle  $X = (X_t)_{t \in \mathbb{Z}}$  peut toujours s'écrire sous la forme

$$X_t = a(t) + s(t) + u_t, \quad \text{pour } t \in \mathbb{Z},$$

$t \mapsto a(t)$  et  $t \mapsto s(t)$  fonctions déterministes avec  $\mathbf{E}[X_t] = a(t) + s(t)$ , et

- 1  $t \mapsto s(t)$  fonction périodique de période  $r > 0$  telle que  $\sum_{i=1}^r s(i) = 0$ ,  
alors, si  $s$  non nulle,  $s$  est la **saisonnalité** de  $X$ ,
- 2 si  $t \mapsto a(t)$  est non nulle,  $a$  est la **tendance** de  $X$ ,
- 3  $u = (u_t)_t$  tel que  $\mathbf{E}[u_t] = 0$ , **bruit** de  $X$ .

## Exemple

*Tendance polynômiale* :  $a(t) = a_0 + a_1 t + \dots + a_k t^k$ , *saisonnalité annuelle* :  
 $s(t) = \sum_{i=1}^{12} s_i \mathbb{1}_{t=i} \text{ [12]}$  ou  $s(t) = \alpha \sin(2\pi t/12) + \beta \cos(2\pi t/12), \dots$

## Définition

Une série temporelle  $X = (X_t)_{t \in \mathbb{Z}}$  peut toujours s'écrire sous la forme

$$X_t = a(t) + s(t) + u_t, \quad \text{pour } t \in \mathbb{Z},$$

$t \mapsto a(t)$  et  $t \mapsto s(t)$  fonctions déterministes avec  $\mathbf{E}[X_t] = a(t) + s(t)$ , et

- 1  $t \mapsto s(t)$  fonction périodique de période  $r > 0$  telle que  $\sum_{i=1}^r s(i) = 0$ ,  
alors, si  $s$  non nulle,  $s$  est la **saisonnalité** de  $X$ ,
- 2 si  $t \mapsto a(t)$  est non nulle,  $a$  est la **tendance** de  $X$ ,
- 3  $u = (u_t)_t$  tel que  $\mathbf{E}[u_t] = 0$ , **bruit** de  $X$ .

## Exemple

Tendance polynômiale :  $a(t) = a_0 + a_1 t + \dots + a_k t^k$ , saisonnalité annuelle :  
 $s(t) = \sum_{i=1}^{12} s_i \mathbf{1}_{t=i} \text{ [12]}$  ou  $s(t) = \alpha \sin(2\pi t/12) + \beta \cos(2\pi t/12), \dots$

# Définitions (encore)

## Propriété

- Si  $(X_t)$  série avec  $a(t)$  non constante ou  $s \neq 0$ , alors  $(X_t)$  non stationnaire.
- Décomposition non unique :  $X_t = a(t) + s(t) + u_t = (a(t) - s(t)) + 2s(t) + u_t$ .

## Définition

$(X_t)_t$  série avec tendance  $a(t)$  et/ou saisonnalité  $s(t)$ . On appelle :

- $(X_t - a(t))_t$  ou  $(X_t - \hat{a}(t))_t$ , où  $\hat{a}(t)$  estimateur de  $a(t)$ , série détendancialisée.
- $(X_t - s(t))_t$  ou  $(X_t - \hat{s}(t))_t$ , où  $\hat{s}(t)$  estimateur de  $s(t)$  série désaisonnalisée ou corrigée des variations saisonnières.

$a$  et  $s$  : tendances **additives**. Autre type de tendance :

## Définition

$(X_t)_t$  possède une **tendance multiplicative** si  $X_t - \mathbb{E}[X_t] = u(t)$  où  $(u_t)_{t \in \mathbb{Z}}$  bruit de variance  $\sigma(\cdot)$  non constante.

# Définitions (encore)

## Propriété

- Si  $(X_t)$  série avec  $a(t)$  non constante ou  $s \neq 0$ , alors  $(X_t)$  non stationnaire.
- Décomposition non unique :  $X_t = a(t) + s(t) + u_t = (a(t) - s(t)) + 2s(t) + u_t$ .

## Définition

$(X_t)_t$  série avec tendance  $a(t)$  et/ou saisonnalité  $s(t)$ . On appelle :

- $(X_t - a(t))_t$  ou  $(X_t - \hat{a}(t))_t$ , où  $\hat{a}(t)$  estimateur de  $a(t)$ , série **détendancialisée**.
- $(X_t - s(t))_t$  ou  $(X_t - \hat{s}(t))_t$ , où  $\hat{s}(t)$  estimateur de  $s(t)$  série **désaisonnalisée** ou **corrigée des variations saisonnières**.

$a$  et  $s$  : tendances **additives**. Autre type de tendance :

## Définition

$(X_t)_t$  possède une **tendance multiplicative** si  $X_t - \mathbb{E}[X_t] = u(t)$  où  $(u_t)_{t \in \mathbb{Z}}$  bruit de variance  $\sigma(\cdot)$  non constante.

# Définitions (encore)

## Propriété

- Si  $(X_t)$  série avec  $a(t)$  non constante ou  $s \neq 0$ , alors  $(X_t)$  non stationnaire.
- Décomposition non unique :  $X_t = a(t) + s(t) + u_t = (a(t) - s(t)) + 2s(t) + u_t$ .

## Définition

$(X_t)_t$  série avec tendance  $a(t)$  et/ou saisonnalité  $s(t)$ . On appelle :

- $(X_t - a(t))_t$  ou  $(X_t - \hat{a}(t))_t$ , où  $\hat{a}(t)$  estimateur de  $a(t)$ , série détendancialisée.
- $(X_t - s(t))_t$  ou  $(X_t - \hat{s}(t))_t$ , où  $\hat{s}(t)$  estimateur de  $s(t)$  série désaisonnalisée ou corrigée des variations saisonnières.

$a$  et  $s$  : tendances **additives**. Autre type de tendance :

## Définition

$(X_t)_t$  possède une **tendance multiplicative** si  $X_t - \mathbb{E}[X_t] = u(t)$  où  $(u_t)_{t \in \mathbb{Z}}$  bruit de variance  $\sigma(\cdot)$  non constante.

# Définitions (encore)

## Propriété

- Si  $(X_t)$  série avec  $a(t)$  non constante ou  $s \neq 0$ , alors  $(X_t)$  non stationnaire.
- Décomposition non unique :  $X_t = a(t) + s(t) + u_t = (a(t) - s(t)) + 2s(t) + u_t$ .

## Définition

$(X_t)_t$  série avec tendance  $a(t)$  et/ou saisonnalité  $s(t)$ . On appelle :

- $(X_t - a(t))_t$  ou  $(X_t - \hat{a}(t))_t$ , où  $\hat{a}(t)$  estimateur de  $a(t)$ , **série détendancialisée**.
- $(X_t - s(t))_t$  ou  $(X_t - \hat{s}(t))_t$ , où  $\hat{s}(t)$  estimateur de  $s(t)$  série désaisonnalisée ou corrigée des variations saisonnières.

$a$  et  $s$  : tendances **additives**. Autre type de tendance :

## Définition

$(X_t)_t$  possède une **tendance multiplicative** si  $X_t - \mathbb{E}[X_t] = u(t)$  où  $(u_t)_{t \in \mathbb{Z}}$  bruit de variance  $\sigma(\cdot)$  non constante.

# Définitions (encore)

## Propriété

- Si  $(X_t)$  série avec  $a(t)$  non constante ou  $s \neq 0$ , alors  $(X_t)$  non stationnaire.
- Décomposition non unique :  $X_t = a(t) + s(t) + u_t = (a(t) - s(t)) + 2s(t) + u_t$ .

## Définition

$(X_t)_t$  série avec tendance  $a(t)$  et/ou saisonnalité  $s(t)$ . On appelle :

- $(X_t - a(t))_t$  ou  $(X_t - \hat{a}(t))_t$ , où  $\hat{a}(t)$  estimateur de  $a(t)$ , série **détendancialisée**.
- $(X_t - s(t))_t$  ou  $(X_t - \hat{s}(t))_t$ , où  $\hat{s}(t)$  estimateur de  $s(t)$  série **désaisonnalisée** ou **corrigée des variations saisonnières**.

$a$  et  $s$  : tendances **additives**. Autre type de tendance :

## Définition

$(X_t)_t$  possède une **tendance multiplicative** si  $X_t - \mathbb{E}[X_t] = u(t)$  où  $(u_t)_{t \in \mathbb{Z}}$  bruit de variance  $\sigma(\cdot)$  non constante.

# Définitions (encore)

## Propriété

- Si  $(X_t)$  série avec  $a(t)$  non constante ou  $s \neq 0$ , alors  $(X_t)$  non stationnaire.
- Décomposition non unique :  $X_t = a(t) + s(t) + u_t = (a(t) - s(t)) + 2s(t) + u_t$ .

## Définition

$(X_t)_t$  série avec tendance  $a(t)$  et/ou saisonnalité  $s(t)$ . On appelle :

- $(X_t - a(t))_t$  ou  $(X_t - \hat{a}(t))_t$ , où  $\hat{a}(t)$  estimateur de  $a(t)$ , série **détendancialisée**.
- $(X_t - s(t))_t$  ou  $(X_t - \hat{s}(t))_t$ , où  $\hat{s}(t)$  estimateur de  $s(t)$  série **désaisonnalisée** ou **corrigée des variations saisonnières**.

$a$  et  $s$  : tendances **additives**. Autre type de tendance :

## Définition

$(X_t)_t$  possède une **tendance multiplicative** si  $X_t - \mathbb{E}[X_t] = u(t)$  où  $(u_t)_{t \in \mathbb{Z}}$  bruit de variance  $\sigma(\cdot)$  non constante.

# Définitions (encore)

## Propriété

- Si  $(X_t)$  série avec  $a(t)$  non constante ou  $s \neq 0$ , alors  $(X_t)$  non stationnaire.
- Décomposition non unique :  $X_t = a(t) + s(t) + u_t = (a(t) - s(t)) + 2s(t) + u_t$ .

## Définition

$(X_t)_t$  série avec tendance  $a(t)$  et/ou saisonnalité  $s(t)$ . On appelle :

- $(X_t - a(t))_t$  ou  $(X_t - \hat{a}(t))_t$ , où  $\hat{a}(t)$  estimateur de  $a(t)$ , série **détendancialisée**.
- $(X_t - s(t))_t$  ou  $(X_t - \hat{s}(t))_t$ , où  $\hat{s}(t)$  estimateur de  $s(t)$  série **désaisonnalisée** ou **corrigée des variations saisonnières**.

$a$  et  $s$  : tendances **additives**. Autre type de tendance :

## Définition

$(X_t)_t$  possède une **tendance multiplicative** si  $X_t - \mathbb{E}[X_t] = u(t)$  où  $(u_t)_{t \in \mathbb{Z}}$  bruit de variance  $\sigma(\cdot)$  non constante.

# Estimation de la tendance et de la saisonnalité

Soit  $(X_1, \dots, X_n)$  trajectoire observée de  $(X_t)_t$ .

⇒ Estimation de  $a(\cdot)$  et/ou  $s(\cdot)$

⇒ A partir de  $(\hat{u}_t)_{1 \leq t \leq n} = (X_t - \hat{a}(t) - \hat{s}(t))_{1 \leq t \leq n}$ , étude du bruit

⇒ Utilisation de  $\hat{a}(t)$ ,  $\hat{s}(t)$  et  $(\hat{u}_t)$  pour prédire

# Estimation de la tendance et de la saisonnalité

Soit  $(X_1, \dots, X_n)$  trajectoire observée de  $(X_t)_t$ .

⇒ Estimation de  $a(\cdot)$  et/ou  $s(\cdot)$

⇒ A partir de  $(\hat{u}_t)_{1 \leq t \leq n} = (X_t - \hat{a}(t) - \hat{s}(t))_{1 \leq t \leq n}$ , étude du bruit

⇒ Utilisation de  $\hat{a}(t)$ ,  $\hat{s}(t)$  et  $(\hat{u}_t)$  pour prédire

# Estimation de la tendance et de la saisonnalité

Soit  $(X_1, \dots, X_n)$  trajectoire observée de  $(X_t)_t$ .

⇒ Estimation de  $a(\cdot)$  et/ou  $s(\cdot)$

⇒ A partir de  $(\hat{u}_t)_{1 \leq t \leq n} = (X_t - \hat{a}(t) - \hat{s}(t))_{1 \leq t \leq n}$ , étude du bruit

⇒ Utilisation de  $\hat{a}(t)$ ,  $\hat{s}(t)$  et  $(\hat{u}_t)$  pour prédire

# Estimation de la tendance et de la saisonnalité

Soit  $(X_1, \dots, X_n)$  trajectoire observée de  $(X_t)_t$ .

⇒ Estimation de  $a(\cdot)$  et/ou  $s(\cdot)$

⇒ A partir de  $(\hat{u}_t)_{1 \leq t \leq n} = (X_t - \hat{a}(t) - \hat{s}(t))_{1 \leq t \leq n}$ , étude du bruit

⇒ Utilisation de  $\hat{a}(t)$ ,  $\hat{s}(t)$  et  $(\hat{u}_t)$  pour prédire

# Plan du cours

- 1 Définitions et premières propriétés
  - Séries chronologiques
  - Stationnarité
  - Première analyse statistique de la dépendance
- 2 Tendances et saisonnalités
  - Définition et propriétés
  - Estimation semi-paramétrique de la tendance et de la saisonnalité
  - Estimation non-paramétrique de la tendance et de la saisonnalité
- 3 Exemples de modèles de séries stationnaires
  - Processus ARMA
  - Processus GARCH
  - Autres exemples de séries temporelles
- 4 Estimation, sélection de modèle, test et prédiction
  - Estimation semi-paramétrique
  - Sélection de modèles et test d'adéquation
  - Prédiction

# Estimation semi-paramétrique de la saisonnalité seule

## Définition

On peut toujours écrire  $s(t)$  de période  $T$  sous la forme :

$$s(t) = \sum_{i=1}^T s_i \mathbb{1}_{t=i} [T] \quad \text{avec} \quad \sum_{i=1}^T s_i = 0.$$

$\Rightarrow s(t) = \sum_{i=1}^{T-1} s_i (\mathbb{1}_{t=i} [T] - \mathbb{1}_{t=T} [T])$  combinaison linéaire

## Propriété

Si  $X_t = a + s(t) + u_t$  pour  $t \in \{1, \dots, nT\}$  (tendance constante), par régression par moindres carrés :

$$\hat{s}_i = \frac{1}{n} \sum_{k=1}^n X_{i+(k-1)T} - \frac{1}{nT} \sum_{k=1}^{nT} X_k \quad \text{pour } i = 1, \dots, T$$

Interprétation : la saisonnalité estimée de janvier est la moyenne des mois de janvier moins la moyenne globale

# Estimation semi-paramétrique de la saisonnalité seule

## Définition

On peut toujours écrire  $s(t)$  de période  $T$  sous la forme :

$$s(t) = \sum_{i=1}^T s_i \mathbb{1}_{t=i} [T] \quad \text{avec} \quad \sum_{i=1}^T s_i = 0.$$

$\Rightarrow s(t) = \sum_{i=1}^{T-1} s_i (\mathbb{1}_{t=i} [T] - \mathbb{1}_{t=T} [T])$  combinaison linéaire

## Propriété

Si  $X_t = a + s(t) + u_t$  pour  $t \in \{1, \dots, nT\}$  (tendance constante), par régression par moindres carrés :

$$\hat{s}_i = \frac{1}{n} \sum_{k=1}^n X_{i+(k-1)T} - \frac{1}{nT} \sum_{k=1}^{nT} X_k \quad \text{pour } i = 1, \dots, T$$

Interprétation : la saisonnalité estimée de janvier est la moyenne des mois de janvier moins la moyenne globale

# Estimation semi-paramétrique de la saisonnalité seule

## Définition

On peut toujours écrire  $s(t)$  de période  $T$  sous la forme :

$$s(t) = \sum_{i=1}^T s_i \mathbb{1}_{t=i} [T] \quad \text{avec} \quad \sum_{i=1}^T s_i = 0.$$

$\Rightarrow s(t) = \sum_{i=1}^{T-1} s_i (\mathbb{1}_{t=i} [T] - \mathbb{1}_{t=T} [T])$  combinaison linéaire

## Propriété

Si  $X_t = a + s(t) + u_t$  pour  $t \in \{1, \dots, nT\}$  (tendance constante), par régression par moindres carrés :

$$\hat{s}_i = \frac{1}{n} \sum_{k=1}^n X_{i+(k-1)T} - \frac{1}{nT} \sum_{k=1}^{nT} X_k \quad \text{pour } i = 1, \dots, T$$

Interprétation : la saisonnalité estimée de janvier est la moyenne des mois de janvier moins la moyenne globale

# Estimation semi-paramétrique de la saisonnalité seule

## Définition

On peut toujours écrire  $s(t)$  de période  $T$  sous la forme :

$$s(t) = \sum_{i=1}^T s_i \mathbb{1}_{t=i} [T] \quad \text{avec} \quad \sum_{i=1}^T s_i = 0.$$

$\Rightarrow s(t) = \sum_{i=1}^{T-1} s_i (\mathbb{1}_{t=i} [T] - \mathbb{1}_{t=T} [T])$  combinaison linéaire

## Propriété

Si  $X_t = a + s(t) + u_t$  pour  $t \in \{1, \dots, nT\}$  (tendance constante), par régression par moindres carrés :

$$\hat{s}_i = \frac{1}{n} \sum_{k=1}^n X_{i+(k-1)T} - \frac{1}{nT} \sum_{k=1}^{nT} X_k \quad \text{pour } i = 1, \dots, T$$

Interprétation : la saisonnalité estimée de janvier est la moyenne des mois de janvier moins la moyenne globale

# Estimation semi-paramétrique de la saisonnalité seule

## Définition

On peut toujours écrire  $s(t)$  de période  $T$  sous la forme :

$$s(t) = \sum_{i=1}^T s_i \mathbb{1}_{t=i} [T] \quad \text{avec} \quad \sum_{i=1}^T s_i = 0.$$

$\Rightarrow s(t) = \sum_{i=1}^{T-1} s_i (\mathbb{1}_{t=i} [T] - \mathbb{1}_{t=T} [T])$  combinaison linéaire

## Propriété

Si  $X_t = a + s(t) + u_t$  pour  $t \in \{1, \dots, nT\}$  (tendance constante), par régression par moindres carrés :

$$\hat{s}_i = \frac{1}{n} \sum_{k=1}^n X_{i+(k-1)T} - \frac{1}{nT} \sum_{k=1}^{nT} X_k \quad \text{pour } i = 1, \dots, T$$

Interprétation : la saisonnalité estimée de janvier est la moyenne des mois de janvier moins la moyenne globale

## Estimation semi-paramétrique de la tendance seule

On fait la supposition que  $a(t) = a_0 + a_1 g_1(t) + \dots + a_p g_p(t)$  pour  $t \in \mathbb{Z}$ .

### Exemple

①  $a(t) = a_0 + a_1 t$  : tendance linéaire

②  $a(t) = a_0 + a_1 t + \dots + a_p t^p$  : tendance polynomiale

③  $a(t) = \sum_{k=0}^p a_k \cos\left(2\pi \frac{kt}{n}\right) + \sum_{k=1}^p b_k \sin\left(2\pi \frac{kt}{n}\right)$  : tendance trigonométrique

⇒ Régression linéaire par moindres carrés!

### Propriété

Si  $X_t = a(t) + u_t$ , avec  $Z = \begin{bmatrix} 1 & G_1 & \dots & G_p \end{bmatrix}$  où  $G_i = {}^t(g_i(1), \dots, g_i(n))$ , on a :

$${}^t(\hat{a}_0, \dots, \hat{a}_p) = ({}^t Z Z)^{-1} {}^t Z {}^t(X_1, \dots, X_n).$$

⇒ Estimation de la tendance :  $\hat{a}(t) = \hat{a}_0 + \hat{a}_1 g_1(t) + \dots + \hat{a}_p g_p(t)$

## Estimation semi-paramétrique de la tendance seule

On fait la supposition que  $a(t) = a_0 + a_1 g_1(t) + \dots + a_p g_p(t)$  pour  $t \in \mathbb{Z}$ .

### Exemple

❶  $a(t) = a_0 + a_1 t$  : *tendance linéaire*

❷  $a(t) = a_0 + a_1 t + \dots + a_p t^p$  : *tendance polynomiale*

❸  $a(t) = \sum_{k=0}^p a_k \cos\left(2\pi \frac{kt}{n}\right) + \sum_{k=1}^p b_k \sin\left(2\pi \frac{kt}{n}\right)$  : *tendance trigonométrique*

⇒ Régression linéaire par moindres carrés!

### Propriété

Si  $X_t = a(t) + u_t$ , avec  $Z = \begin{bmatrix} 1 & G_1 & \dots & G_p \end{bmatrix}$  où  $G_i = {}^t(g_i(1), \dots, g_i(n))$ , on a :

$${}^t(\hat{a}_0, \dots, \hat{a}_p) = ({}^t Z Z)^{-1} {}^t Z {}^t(X_1, \dots, X_n).$$

⇒ Estimation de la tendance :  $\hat{a}(t) = \hat{a}_0 + \hat{a}_1 g_1(t) + \dots + \hat{a}_p g_p(t)$

## Estimation semi-paramétrique de la tendance seule

On fait la supposition que  $a(t) = a_0 + a_1 g_1(t) + \dots + a_p g_p(t)$  pour  $t \in \mathbb{Z}$ .

### Exemple

①  $a(t) = a_0 + a_1 t$  : *tendance linéaire*

②  $a(t) = a_0 + a_1 t + \dots + a_p t^p$  : *tendance polynomiale*

③  $a(t) = \sum_{k=0}^p a_k \cos\left(2\pi \frac{kt}{n}\right) + \sum_{k=1}^p b_k \sin\left(2\pi \frac{kt}{n}\right)$  : *tendance trigonométrique*

⇒ Régression linéaire par moindres carrés!

### Propriété

Si  $X_t = a(t) + u_t$ , avec  $Z = \begin{bmatrix} 1 & G_1 & \dots & G_p \end{bmatrix}$  où  $G_i = {}^t(g_i(1), \dots, g_i(n))$ , on a :

$${}^t(\hat{a}_0, \dots, \hat{a}_p) = ({}^t Z Z)^{-1} {}^t Z {}^t(X_1, \dots, X_n).$$

⇒ Estimation de la tendance :  $\hat{a}(t) = \hat{a}_0 + \hat{a}_1 g_1(t) + \dots + \hat{a}_p g_p(t)$

## Estimation semi-paramétrique de la tendance seule

On fait la supposition que  $a(t) = a_0 + a_1 g_1(t) + \dots + a_p g_p(t)$  pour  $t \in \mathbb{Z}$ .

### Exemple

❶  $a(t) = a_0 + a_1 t$  : *tendance linéaire*

❷  $a(t) = a_0 + a_1 t + \dots + a_p t^p$  : *tendance polynomiale*

❸  $a(t) = \sum_{k=0}^p a_k \cos\left(2\pi \frac{kt}{n}\right) + \sum_{k=1}^p b_k \sin\left(2\pi \frac{kt}{n}\right)$  : *tendance trigonométrique*

⇒ Régression linéaire par moindres carrés!

### Propriété

Si  $X_t = a(t) + u_t$ , avec  $Z = \begin{bmatrix} 1 & G_1 & \dots & G_p \end{bmatrix}$  où  $G_i = {}^t(g_i(1), \dots, g_i(n))$ , on a :

$${}^t(\hat{a}_0, \dots, \hat{a}_p) = ({}^t Z Z)^{-1} {}^t Z {}^t(X_1, \dots, X_n).$$

⇒ Estimation de la tendance :  $\hat{a}(t) = \hat{a}_0 + \hat{a}_1 g_1(t) + \dots + \hat{a}_p g_p(t)$

## Estimation semi-paramétrique de la tendance seule

On fait la supposition que  $a(t) = a_0 + a_1 g_1(t) + \dots + a_p g_p(t)$  pour  $t \in \mathbb{Z}$ .

### Exemple

❶  $a(t) = a_0 + a_1 t$  : tendance linéaire

❷  $a(t) = a_0 + a_1 t + \dots + a_p t^p$  : tendance polynomiale

❸  $a(t) = \sum_{k=0}^p a_k \cos\left(2\pi \frac{kt}{n}\right) + \sum_{k=1}^p b_k \sin\left(2\pi \frac{kt}{n}\right)$  : tendance trigonométrique

⇒ Régression linéaire par moindres carrés!

### Propriété

Si  $X_t = a(t) + u_t$ , avec  $Z = \begin{bmatrix} 1 & G_1 & \dots & G_p \end{bmatrix}$  où  $G_i = {}^t(g_i(1), \dots, g_i(n))$ , on a :

$${}^t(\hat{a}_0, \dots, \hat{a}_p) = ({}^t Z Z)^{-1} {}^t Z {}^t(X_1, \dots, X_n).$$

⇒ Estimation de la tendance :  $\hat{a}(t) = \hat{a}_0 + \hat{a}_1 g_1(t) + \dots + \hat{a}_p g_p(t)$

## Estimation semi-paramétrique de la tendance seule

On fait la supposition que  $a(t) = a_0 + a_1 g_1(t) + \dots + a_p g_p(t)$  pour  $t \in \mathbb{Z}$ .

### Exemple

①  $a(t) = a_0 + a_1 t$  : tendance linéaire

②  $a(t) = a_0 + a_1 t + \dots + a_p t^p$  : tendance polynomiale

③  $a(t) = \sum_{k=0}^p a_k \cos\left(2\pi \frac{kt}{n}\right) + \sum_{k=1}^p b_k \sin\left(2\pi \frac{kt}{n}\right)$  : tendance trigonométrique

⇒ Régression linéaire par moindres carrés!

### Propriété

Si  $X_t = a(t) + u_t$ , avec  $Z = \begin{bmatrix} 1 & G_1 & \dots & G_p \end{bmatrix}$  où  $G_i = {}^t(g_i(1), \dots, g_i(n))$ , on

a :

$${}^t(\hat{a}_0, \dots, \hat{a}_p) = ({}^t Z Z)^{-1} {}^t Z {}^t(X_1, \dots, X_n).$$

⇒ Estimation de la tendance :  $\hat{a}(t) = \hat{a}_0 + \hat{a}_1 g_1(t) + \dots + \hat{a}_p g_p(t)$

## Estimation semi-paramétrique de la tendance seule

On fait la supposition que  $a(t) = a_0 + a_1 g_1(t) + \dots + a_p g_p(t)$  pour  $t \in \mathbb{Z}$ .

### Exemple

①  $a(t) = a_0 + a_1 t$  : tendance linéaire

②  $a(t) = a_0 + a_1 t + \dots + a_p t^p$  : tendance polynomiale

③  $a(t) = \sum_{k=0}^p a_k \cos\left(2\pi \frac{kt}{n}\right) + \sum_{k=1}^p b_k \sin\left(2\pi \frac{kt}{n}\right)$  : tendance trigonométrique

⇒ Régression linéaire par moindres carrés!

### Propriété

Si  $X_t = a(t) + u_t$ , avec  $Z = \begin{bmatrix} 1 & G_1 & \dots & G_p \end{bmatrix}$  où  $G_i = {}^t(g_i(1), \dots, g_i(n))$ , on a :

$${}^t(\hat{a}_0, \dots, \hat{a}_p) = ({}^t Z Z)^{-1} {}^t Z {}^t(X_1, \dots, X_n).$$

⇒ Estimation de la tendance :  $\hat{a}(t) = \hat{a}_0 + \hat{a}_1 g_1(t) + \dots + \hat{a}_p g_p(t)$

## Estimation semi-paramétrique de la tendance seule

On fait la supposition que  $a(t) = a_0 + a_1 g_1(t) + \dots + a_p g_p(t)$  pour  $t \in \mathbb{Z}$ .

### Exemple

①  $a(t) = a_0 + a_1 t$  : tendance linéaire

②  $a(t) = a_0 + a_1 t + \dots + a_p t^p$  : tendance polynomiale

③  $a(t) = \sum_{k=0}^p a_k \cos\left(2\pi \frac{kt}{n}\right) + \sum_{k=1}^p b_k \sin\left(2\pi \frac{kt}{n}\right)$  : tendance trigonométrique

⇒ Régression linéaire par moindres carrés!

### Propriété

Si  $X_t = a(t) + u_t$ , avec  $Z = \begin{bmatrix} 1 & G_1 & \dots & G_p \end{bmatrix}$  où  $G_i = {}^t(g_i(1), \dots, g_i(n))$ , on a :

$${}^t(\hat{a}_0, \dots, \hat{a}_p) = ({}^t Z Z)^{-1} {}^t Z {}^t(X_1, \dots, X_n).$$

⇒ Estimation de la tendance :  $\hat{a}(t) = \hat{a}_0 + \hat{a}_1 g_1(t) + \dots + \hat{a}_p g_p(t)$

# Estimation conjointe de la tendance et de la saisonnalité

On suppose que

$$X_t = (a_0 + a_1 g_1(t) + \dots + a_p g_p(t)) + (s_1 h_1(t) + \dots + s_{T-1} h_{T-1}(t)) + u_t$$

$$\text{où } h_i(t) = \mathbb{1}_{t=i} [T] - \mathbb{1}_{t=T} [T] \text{ pour } t \in \mathbb{Z}.$$

## Propriété

$$\text{Avec } Z = \begin{bmatrix} 1 & G_1 & \dots & G_p & H_1 & \dots & H_{T-1} \end{bmatrix} \text{ où } \begin{cases} G_i = {}^t(g_i(1), \dots, g_i(n)) \\ H_i = {}^t(h_i(1), \dots, h_i(n)) \end{cases},$$

$$\text{on a : } {}^t(\hat{a}_0, \dots, \hat{a}_p, \hat{s}_1, \dots, \hat{s}_{T-1}) = ({}^t Z Z)^{-1} {}^t Z {}^t(X_1, \dots, X_n).$$

## Conséquence :

- Estimation de la tendance :  $\hat{a}(t) = \hat{a}_0 + \hat{a}_1 g_1(t) + \dots + \hat{a}_p g_p(t)$
- Estimation de la saisonnalité :  $\hat{s}(t) = \hat{s}_1 h_1(t) + \dots + \hat{s}_{T-1} h_{T-1}(t)$

# Estimation conjointe de la tendance et de la saisonnalité

On suppose que

$$X_t = (a_0 + a_1 g_1(t) + \cdots + a_p g_p(t)) + (s_1 h_1(t) + \cdots + s_{T-1} h_{T-1}(t)) + u_t$$

$$\text{où } h_i(t) = \mathbb{1}_{t=i} [T] - \mathbb{1}_{t=T} [T] \text{ pour } t \in \mathbb{Z}.$$

## Propriété

$$\text{Avec } Z = \begin{bmatrix} 1 & G_1 & \cdots & G_p & H_1 & \cdots & H_{T-1} \end{bmatrix} \text{ où } \begin{cases} G_i = {}^t(g_i(1), \dots, g_i(n)) \\ H_i = {}^t(h_i(1), \dots, h_i(n)) \end{cases},$$

$$\text{on a : } {}^t(\hat{a}_0, \dots, \hat{a}_p, \hat{s}_1, \dots, \hat{s}_{T-1}) = ({}^t Z Z)^{-1} {}^t Z {}^t(X_1, \dots, X_n).$$

## Conséquence :

- Estimation de la tendance :  $\hat{a}(t) = \hat{a}_0 + \hat{a}_1 g_1(t) + \cdots + \hat{a}_p g_p(t)$
- Estimation de la saisonnalité :  $\hat{s}(t) = \hat{s}_1 h_1(t) + \cdots + \hat{s}_{T-1} h_{T-1}(t)$

# Estimation conjointe de la tendance et de la saisonnalité

On suppose que

$$X_t = (a_0 + a_1 g_1(t) + \cdots + a_p g_p(t)) + (s_1 h_1(t) + \cdots + s_{T-1} h_{T-1}(t)) + u_t$$

$$\text{où } h_i(t) = \mathbb{1}_{t=i} [T] - \mathbb{1}_{t=T} [T] \text{ pour } t \in \mathbb{Z}.$$

## Propriété

$$\text{Avec } Z = \begin{bmatrix} 1 & G_1 & \cdots & G_p & H_1 & \cdots & H_{T-1} \end{bmatrix} \text{ où } \begin{cases} G_i = {}^t(g_i(1), \dots, g_i(n)) \\ H_i = {}^t(h_i(1), \dots, h_i(n)) \end{cases},$$

$$\text{on a : } {}^t(\hat{a}_0, \dots, \hat{a}_p, \hat{s}_1, \dots, \hat{s}_{T-1}) = ({}^t Z Z)^{-1} {}^t Z {}^t(X_1, \dots, X_n).$$

## Conséquence :

- Estimation de la tendance :  $\hat{a}(t) = \hat{a}_0 + \hat{a}_1 g_1(t) + \cdots + \hat{a}_p g_p(t)$
- Estimation de la saisonnalité :  $\hat{s}(t) = \hat{s}_1 h_1(t) + \cdots + \hat{s}_{T-1} h_{T-1}(t)$

# Estimation conjointe de la tendance et de la saisonnalité

On suppose que

$$X_t = (a_0 + a_1 g_1(t) + \dots + a_p g_p(t)) + (s_1 h_1(t) + \dots + s_{T-1} h_{T-1}(t)) + u_t$$

$$\text{où } h_i(t) = \mathbb{1}_{t=i} [T] - \mathbb{1}_{t=T} [T] \text{ pour } t \in \mathbb{Z}.$$

## Propriété

$$\text{Avec } Z = \begin{bmatrix} 1 & G_1 & \dots & G_p & H_1 & \dots & H_{T-1} \end{bmatrix} \text{ où } \begin{cases} G_i = {}^t(g_i(1), \dots, g_i(n)) \\ H_i = {}^t(h_i(1), \dots, h_i(n)) \end{cases},$$

$$\text{on a : } {}^t(\hat{a}_0, \dots, \hat{a}_p, \hat{s}_1, \dots, \hat{s}_{T-1}) = ({}^t Z Z)^{-1} {}^t Z {}^t(X_1, \dots, X_n).$$

## Conséquence :

- Estimation de la tendance :  $\hat{a}(t) = \hat{a}_0 + \hat{a}_1 g_1(t) + \dots + \hat{a}_p g_p(t)$
- Estimation de la saisonnalité :  $\hat{s}(t) = \hat{s}_1 h_1(t) + \dots + \hat{s}_{T-1} h_{T-1}(t)$

# Estimation conjointe de la tendance et de la saisonnalité

On suppose que

$$X_t = (a_0 + a_1 g_1(t) + \dots + a_p g_p(t)) + (s_1 h_1(t) + \dots + s_{T-1} h_{T-1}(t)) + u_t$$

$$\text{où } h_i(t) = \mathbb{1}_{t=i} [T] - \mathbb{1}_{t=T} [T] \text{ pour } t \in \mathbb{Z}.$$

## Propriété

$$\text{Avec } Z = \begin{bmatrix} 1 & G_1 & \dots & G_p & H_1 & \dots & H_{T-1} \end{bmatrix} \text{ où } \begin{cases} G_i = {}^t(g_i(1), \dots, g_i(n)) \\ H_i = {}^t(h_i(1), \dots, h_i(n)) \end{cases},$$

$$\text{on a : } {}^t(\hat{a}_0, \dots, \hat{a}_p, \hat{s}_1, \dots, \hat{s}_{T-1}) = ({}^t Z Z)^{-1} {}^t Z {}^t(X_1, \dots, X_n).$$

## Conséquence :

• Estimation de la tendance :  $\hat{a}(t) = \hat{a}_0 + \hat{a}_1 g_1(t) + \dots + \hat{a}_p g_p(t)$

• Estimation de la saisonnalité :  $\hat{s}(t) = \hat{s}_1 h_1(t) + \dots + \hat{s}_{T-1} h_{T-1}(t)$

# Estimation conjointe de la tendance et de la saisonnalité

On suppose que

$$X_t = (a_0 + a_1 g_1(t) + \dots + a_p g_p(t)) + (s_1 h_1(t) + \dots + s_{T-1} h_{T-1}(t)) + u_t$$

$$\text{où } h_i(t) = \mathbb{1}_{t=i} [T] - \mathbb{1}_{t=T} [T] \text{ pour } t \in \mathbb{Z}.$$

## Propriété

$$\text{Avec } Z = \begin{bmatrix} 1 & G_1 & \dots & G_p & H_1 & \dots & H_{T-1} \end{bmatrix} \text{ où } \begin{cases} G_i = {}^t(g_i(1), \dots, g_i(n)) \\ H_i = {}^t(h_i(1), \dots, h_i(n)) \end{cases},$$

$$\text{on a : } {}^t(\hat{a}_0, \dots, \hat{a}_p, \hat{s}_1, \dots, \hat{s}_{T-1}) = ({}^t Z Z)^{-1} {}^t Z {}^t(X_1, \dots, X_n).$$

## Conséquence :

- Estimation de la tendance :  $\hat{a}(t) = \hat{a}_0 + \hat{a}_1 g_1(t) + \dots + \hat{a}_p g_p(t)$
- Estimation de la saisonnalité :  $\hat{s}(t) = \hat{s}_1 h_1(t) + \dots + \hat{s}_{T-1} h_{T-1}(t)$

# Sélection de modèle

## Détermination de $p$

### Exemple

*Typiquement le degré d'une régression polynomiale...*

**Remarque** : Non nécessaire pour la saisonnalité, sauf écrite en cos et sin

⇒ On utilise le **critère BIC**

$$BIC(p) = \sum_{i=1}^n (X_t - \hat{a}_0 - \hat{a}_1 g_1(t) - \dots - \hat{a}_p g_p(t) - \hat{s}(t))^2 + (p + T) \ln(n)$$

$$\Rightarrow \hat{p} = \arg \min_{0 \leq p \leq p_{\max}} BIC(p)$$

# Sélection de modèle

## Détermination de $p$

### Exemple

*Typiquement le degré d'une régression polynomiale...*

**Remarque** : Non nécessaire pour la saisonnalité, sauf écrite en cos et sin

⇒ On utilise le **critère BIC**

$$BIC(p) = \sum_{i=1}^n (X_t - \hat{a}_0 - \hat{a}_1 g_1(t) - \dots - \hat{a}_p g_p(t) - \hat{s}(t))^2 + (p + T) \ln(n)$$

$$\Rightarrow \hat{p} = \arg \min_{0 \leq p \leq p_{\max}} BIC(p)$$

# Sélection de modèle

## Détermination de $p$

### Exemple

*Typiquement le degré d'une régression polynomiale...*

**Remarque** : Non nécessaire pour la saisonnalité, sauf écrite en cos et sin

⇒ On utilise le critère BIC

$$BIC(p) = \sum_{i=1}^n (X_t - \hat{a}_0 - \hat{a}_1 g_1(t) - \dots - \hat{a}_p g_p(t) - \hat{s}(t))^2 + (p + T) \ln(n)$$

$$\Rightarrow \hat{p} = \arg \min_{0 \leq p \leq p_{\max}} BIC(p)$$

# Sélection de modèle

## Détermination de $p$

### Exemple

*Typiquement le degré d'une régression polynomiale...*

**Remarque** : Non nécessaire pour la saisonnalité, sauf écrite en cos et sin

⇒ On utilise le **critère BIC**

$$BIC(p) = \sum_{i=1}^n (X_t - \hat{a}_0 - \hat{a}_1 g_1(t) - \dots - \hat{a}_p g_p(t) - \hat{s}(t))^2 + (p + T) \ln(n)$$

$$\Rightarrow \hat{p} = \arg \min_{0 \leq p \leq p_{\max}} BIC(p)$$

## Détermination de $p$

### Exemple

*Typiquement le degré d'une régression polynomiale...*

**Remarque** : Non nécessaire pour la saisonnalité, sauf écrite en cos et sin

⇒ On utilise le **critère BIC**

$$BIC(p) = \sum_{i=1}^n (X_t - \hat{a}_0 - \hat{a}_1 g_1(t) - \dots - \hat{a}_p g_p(t) - \hat{s}(t))^2 + (p + T) \ln(n)$$

$$\Rightarrow \hat{p} = \arg \min_{0 \leq p \leq p_{\max}} BIC(p)$$

## Détermination de $p$

### Exemple

*Typiquement le degré d'une régression polynomiale...*

**Remarque** : Non nécessaire pour la saisonnalité, sauf écrite en cos et sin

⇒ On utilise le **critère BIC**

$$BIC(p) = \sum_{i=1}^n (X_t - \hat{a}_0 - \hat{a}_1 g_1(t) - \dots - \hat{a}_p g_p(t) - \hat{s}(t))^2 + (p + T) \ln(n)$$

$$\Rightarrow \hat{p} = \arg \min_{0 \leq p \leq p_{\max}} BIC(p)$$

# Convergence des estimateurs

## Quelques résultats théoriques

### Propriété

On suppose que l'on a

$$X_t = (a_0 + a_1 g_1(t) + \dots + a_p g_p(t)) + (s_1 h_1(t) + \dots + s_{T-1} h_{T-1}(t)) + u_t$$

$$\text{et } Z = \begin{bmatrix} 1 & G_1 & \dots & G_p & H_1 & \dots & H_{T-1} \end{bmatrix}.$$

Si  $(u_t)$  bruit blanc et si  $\max_{1 \leq i \leq n} (Z' Z Z)^{-1} Z'_{ii} \xrightarrow{n \rightarrow +\infty} 0$ , alors

$$\hat{\sigma}^{-1} (Z' Z)^{-1} (Z' (\hat{a}_0, \dots, \hat{s}_{T-1}) - (a_0, \dots, s_{T-1})) \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \mathcal{N}(0, 1)$$

$$\text{et } \hat{p} \xrightarrow[n \rightarrow +\infty]{\mathcal{P}} p.$$

**Remarque :** Vrai pour tendance polynomiale, trigonométrique,...

A adapter si  $(u_t)$  n'est pas un bruit blanc,...

# Convergence des estimateurs

## Quelques résultats théoriques

### Propriété

On suppose que l'on a

$$X_t = (a_0 + a_1 g_1(t) + \dots + a_p g_p(t)) + (s_1 h_1(t) + \dots + s_{T-1} h_{T-1}(t)) + u_t$$

$$\text{et } Z = \begin{bmatrix} 1 & G_1 & \dots & G_p & H_1 & \dots & H_{T-1} \end{bmatrix}.$$

Si  $(u_t)$  bruit blanc et si  $\max_{1 \leq i \leq n} (Z' Z Z^{-1} Z)_{ii} \xrightarrow{n \rightarrow +\infty} 0$ , alors

$$\hat{\sigma}^{-1} (Z' Z)^{-1} (Z' (\hat{a}_0, \dots, \hat{s}_{T-1}) - (a_0, \dots, s_{T-1})) \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \mathcal{N}(0, 1)$$

$$\text{et } \hat{p} \xrightarrow[n \rightarrow +\infty]{\mathcal{P}} p.$$

**Remarque :** Vrai pour tendance polynomiale, trigonométrique,...

A adapter si  $(u_t)$  n'est pas un bruit blanc,...

# Convergence des estimateurs

## Quelques résultats théoriques

### Propriété

On suppose que l'on a

$$X_t = (a_0 + a_1 g_1(t) + \dots + a_p g_p(t)) + (s_1 h_1(t) + \dots + s_{T-1} h_{T-1}(t)) + u_t$$

$$\text{et } Z = \begin{bmatrix} 1 & G_1 & \dots & G_p & H_1 & \dots & H_{T-1} \end{bmatrix}.$$

Si  $(u_t)$  bruit blanc et si  $\max_{1 \leq i \leq n} (Z' Z Z^{-1} Z)'_{ii} \xrightarrow{n \rightarrow +\infty} 0$ , alors

$$\hat{\sigma}^{-1} (Z' Z)^{-1} (Z' (\hat{a}_0, \dots, \hat{s}_{T-1}) - (a_0, \dots, s_{T-1})) \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \mathcal{N}(0, 1)$$

$$\text{et } \hat{p} \xrightarrow[n \rightarrow +\infty]{\mathcal{P}} p.$$

Remarque : Vrai pour tendance polynomiale, trigonométrique,...

A adapter si  $(u_t)$  n'est pas un bruit blanc,...

# Convergence des estimateurs

## Quelques résultats théoriques

### Propriété

On suppose que l'on a

$$X_t = (a_0 + a_1 g_1(t) + \dots + a_p g_p(t)) + (s_1 h_1(t) + \dots + s_{T-1} h_{T-1}(t)) + u_t$$

$$\text{et } Z = \begin{bmatrix} 1 & G_1 & \dots & G_p & H_1 & \dots & H_{T-1} \end{bmatrix}.$$

Si  $(u_t)$  bruit blanc et si  $\max_{1 \leq i \leq n} (Z' Z Z^{-1} Z)_{ii} \xrightarrow{n \rightarrow +\infty} 0$ , alors

$$\hat{\sigma}^{-1} (Z' Z)^{-1} (Z' (\hat{a}_0, \dots, \hat{s}_{T-1}) - (a_0, \dots, s_{T-1})) \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \mathcal{N}(0, 1)$$

$$\text{et } \hat{p} \xrightarrow[n \rightarrow +\infty]{\mathcal{P}} p.$$

Remarque : Vrai pour tendance polynomiale, trigonométrique,...

A adapter si  $(u_t)$  n'est pas un bruit blanc,...

# Convergence des estimateurs

## Quelques résultats théoriques

### Propriété

On suppose que l'on a

$$X_t = (a_0 + a_1 g_1(t) + \dots + a_p g_p(t)) + (s_1 h_1(t) + \dots + s_{T-1} h_{T-1}(t)) + u_t$$

$$\text{et } Z = \begin{bmatrix} 1 & G_1 & \dots & G_p & H_1 & \dots & H_{T-1} \end{bmatrix}.$$

Si  $(u_t)$  bruit blanc et si  $\max_{1 \leq i \leq n} (Z' Z Z^{-1} Z)_{ii} \xrightarrow{n \rightarrow +\infty} 0$ , alors

$$\hat{\sigma}^{-1} (Z' Z)^{-1} (Z' (\hat{a}_0, \dots, \hat{s}_{T-1}) - (a_0, \dots, s_{T-1})) \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \mathcal{N}(0, 1)$$

$$\text{et } \hat{p} \xrightarrow[n \rightarrow +\infty]{\mathcal{P}} p.$$

Remarque : Vrai pour tendance polynomiale, trigonométrique,...

A adapter si  $(u_t)$  n'est pas un bruit blanc,...

# Convergence des estimateurs

## Quelques résultats théoriques

### Propriété

On suppose que l'on a

$$X_t = (a_0 + a_1 g_1(t) + \dots + a_p g_p(t)) + (s_1 h_1(t) + \dots + s_{T-1} h_{T-1}(t)) + u_t$$

$$\text{et } Z = \begin{bmatrix} 1 & G_1 & \dots & G_p & H_1 & \dots & H_{T-1} \end{bmatrix}.$$

Si  $(u_t)$  bruit blanc et si  $\max_{1 \leq i \leq n} (Z' Z Z^{-1} Z)_{ii} \xrightarrow{n \rightarrow +\infty} 0$ , alors

$$\hat{\sigma}^{-1} (Z' Z)^{-1} (Z' (\hat{a}_0, \dots, \hat{s}_{T-1}) - (a_0, \dots, s_{T-1})) \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \mathcal{N}(0, 1)$$

$$\text{et } \hat{p} \xrightarrow[n \rightarrow +\infty]{\mathcal{P}} p.$$

**Remarque :** Vrai pour tendance polynomiale, trigonométrique,...

A adapter si  $(u_t)$  n'est pas un bruit blanc,...

# Convergence des estimateurs

## Quelques résultats théoriques

### Propriété

On suppose que l'on a

$$X_t = (a_0 + a_1 g_1(t) + \dots + a_p g_p(t)) + (s_1 h_1(t) + \dots + s_{T-1} h_{T-1}(t)) + u_t$$

$$\text{et } Z = \begin{bmatrix} 1 & G_1 & \dots & G_p & H_1 & \dots & H_{T-1} \end{bmatrix}.$$

Si  $(u_t)$  bruit blanc et si  $\max_{1 \leq i \leq n} (Z' Z Z^{-1} Z)_{ii} \xrightarrow{n \rightarrow +\infty} 0$ , alors

$$\hat{\sigma}^{-1} (Z' Z)^{-1} (Z' (\hat{a}_0, \dots, \hat{s}_{T-1}) - (a_0, \dots, s_{T-1})) \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \mathcal{N}(0, 1)$$

$$\text{et } \hat{p} \xrightarrow[n \rightarrow +\infty]{\mathcal{P}} p.$$

**Remarque :** Vrai pour tendance polynomiale, trigonométrique,...

A adapter si  $(u_t)$  n'est pas un bruit blanc,...

# Extension en cas de variables exogènes observées

Tendance et saisonnalité estimées jusqu'ici comme fonctions du temps

⇒ Extension en cas de **variables exogènes** observées aux mêmes instants

⇒ On suppose observés  $(X_t)_{1 \leq t \leq n}$  mais aussi  $(Y_t)_{1 \leq t \leq n}$  où  $Y_t = (Y_t^{(1)}, \dots, Y_t^{(d)})$ , où  $Y^{(1)}, \dots, Y^{(d)}$  variables exogènes

⇒ **Modèle linéaire** comme première modélisation :

$$X_t = a(t) + s(t) + \sum_{j=1}^d \mu_j Y_t^{(j)} + u_t \quad \text{pour tout } t$$

où  $a(t) = a_0 + a_1 g_1(t) + \dots + a_p g_p(t)$ ,  $s(t) = s_1 h_1(t) + \dots + s_{T-1} h_{T-1}(t)$

## Extension en cas de variables exogènes observées

Tendance et saisonnalité estimées jusqu'ici comme fonctions du temps

⇒ Extension en cas de **variables exogènes** observées aux mêmes instants

⇒ On suppose observés  $(X_t)_{1 \leq t \leq n}$  mais aussi  $(Y_t)_{1 \leq t \leq n}$  où  $Y_t = (Y_t^{(1)}, \dots, Y_t^{(d)})$ , où  $Y^{(1)}, \dots, Y^{(d)}$  variables exogènes

⇒ **Modèle linéaire** comme première modélisation :

$$X_t = a(t) + s(t) + \sum_{j=1}^d \mu_j Y_t^{(j)} + u_t \quad \text{pour tout } t$$

où  $a(t) = a_0 + a_1 g_1(t) + \dots + a_p g_p(t)$ ,  $s(t) = s_1 h_1(t) + \dots + s_{T-1} h_{T-1}(t)$

## Extension en cas de variables exogènes observées

Tendance et saisonnalité estimées jusqu'ici comme fonctions du temps

⇒ Extension en cas de **variables exogènes** observées aux mêmes instants

⇒ On suppose observés  $(X_t)_{1 \leq t \leq n}$  mais aussi  $(Y_t)_{1 \leq t \leq n}$  où  $Y_t = (Y_t^{(1)}, \dots, Y_t^{(d)})$ , où  $Y^{(1)}, \dots, Y^{(d)}$  variables exogènes

⇒ **Modèle linéaire** comme première modélisation :

$$X_t = a(t) + s(t) + \sum_{j=1}^d \mu_j Y_t^{(j)} + u_t \quad \text{pour tout } t$$

où  $a(t) = a_0 + a_1 g_1(t) + \dots + a_p g_p(t)$ ,  $s(t) = s_1 h_1(t) + \dots + s_{T-1} h_{T-1}(t)$

## Extension en cas de variables exogènes observées

Tendance et saisonnalité estimées jusqu'ici comme fonctions du temps

⇒ Extension en cas de **variables exogènes** observées aux mêmes instants

⇒ On suppose observés  $(X_t)_{1 \leq t \leq n}$  mais aussi  $(Y_t)_{1 \leq t \leq n}$  où  $Y_t = (Y_t^{(1)}, \dots, Y_t^{(d)})$ , où  $Y^{(1)}, \dots, Y^{(d)}$  variables exogènes

⇒ **Modèle linéaire** comme première modélisation :

$$X_t = a(t) + s(t) + \sum_{j=1}^d \mu_j Y_t^{(j)} + u_t \quad \text{pour tout } t$$

où  $a(t) = a_0 + a_1 g_1(t) + \dots + a_p g_p(t)$ ,  $s(t) = s_1 h_1(t) + \dots + s_{T-1} h_{T-1}(t)$

## Extension en cas de variables exogènes observées

Tendance et saisonnalité estimées jusqu'ici comme fonctions du temps

⇒ Extension en cas de **variables exogènes** observées aux mêmes instants

⇒ On suppose observés  $(X_t)_{1 \leq t \leq n}$  mais aussi  $(Y_t)_{1 \leq t \leq n}$  où  $Y_t = (Y_t^{(1)}, \dots, Y_t^{(d)})$ , où  $Y^{(1)}, \dots, Y^{(d)}$  variables exogènes

⇒ **Modèle linéaire** comme première modélisation :

$$X_t = a(t) + s(t) + \sum_{j=1}^d \mu_j Y_t^{(j)} + u_t \quad \text{pour tout } t$$

où  $a(t) = a_0 + a_1 g_1(t) + \dots + a_p g_p(t)$ ,  $s(t) = s_1 h_1(t) + \dots + s_{T-1} h_{T-1}(t)$

## Extension en cas de variables exogènes observées

Tendance et saisonnalité estimées jusqu'ici comme fonctions du temps

⇒ Extension en cas de **variables exogènes** observées aux mêmes instants

⇒ On suppose observés  $(X_t)_{1 \leq t \leq n}$  mais aussi  $(Y_t)_{1 \leq t \leq n}$  où  $Y_t = (Y_t^{(1)}, \dots, Y_t^{(d)})$ , où  $Y^{(1)}, \dots, Y^{(d)}$  variables exogènes

⇒ **Modèle linéaire** comme première modélisation :

$$X_t = a(t) + s(t) + \sum_{j=1}^d \mu_j Y_t^{(j)} + u_t \quad \text{pour tout } t$$

où  $a(t) = a_0 + a_1 g_1(t) + \dots + a_p g_p(t)$ ,  $s(t) = s_1 h_1(t) + \dots + s_{T-1} h_{T-1}(t)$

## Extension en cas de variables exogènes observées (2)

Du modèle linéaire :

- 1 Estimateur des paramètres par moindres carrés ;
- 2 Sélection des variables (et fonctions) par critère BIC
- 3 Intervalles de confiance grâce au TLC
- 4 **Prédictions :**

$$\hat{X}_{n+h} = \hat{a}(n+h) + \hat{s}(n+h) + \sum_{j=1}^d \hat{\mu}_j \hat{Y}_{n+h}^{(j)}$$

où  $\hat{Y}_{n+h}^{(j)}$  prédicteurs des  $Y_{n+h}^{(j)}$ .

## Extension en cas de variables exogènes observées (2)

Du modèle linéaire :

- 1 Estimateur des paramètres par moindres carrés ;
- 2 Sélection des variables (et fonctions) par critère BIC
- 3 Intervalles de confiance grâce au TLC
- 4 **Prédictions :**

$$\hat{X}_{n+h} = \hat{a}(n+h) + \hat{s}(n+h) + \sum_{j=1}^d \hat{\mu}_j \hat{Y}_{n+h}^{(j)}$$

où  $\hat{Y}_{n+h}^{(j)}$  prédicteurs des  $Y_{n+h}^{(j)}$ .

## Extension en cas de variables exogènes observées (2)

Du modèle linéaire :

- 1 Estimateur des paramètres par moindres carrés ;
- 2 Sélection des variables (et fonctions) par critère BIC
- 3 Intervalles de confiance grâce au TLC
- 4 Prédiction :

$$\hat{X}_{n+h} = \hat{a}(n+h) + \hat{s}(n+h) + \sum_{j=1}^d \hat{\mu}_j \hat{Y}_{n+h}^{(j)}$$

où  $\hat{Y}_{n+h}^{(j)}$  prédicteurs des  $Y_{n+h}^{(j)}$ .

## Extension en cas de variables exogènes observées (2)

Du modèle linéaire :

- 1 Estimateur des paramètres par moindres carrés ;
- 2 Sélection des variables (et fonctions) par critère BIC
- 3 Intervalles de confiance grâce au TLC
- 4 Prédiction :

$$\hat{X}_{n+h} = \hat{a}(n+h) + \hat{s}(n+h) + \sum_{j=1}^d \hat{\mu}_j \hat{Y}_{n+h}^{(j)}$$

où  $\hat{Y}_{n+h}^{(j)}$  prédicteurs des  $Y_{n+h}^{(j)}$ .

## Extension en cas de variables exogènes observées (2)

Du modèle linéaire :

- 1 Estimateur des paramètres par moindres carrés ;
- 2 Sélection des variables (et fonctions) par critère BIC
- 3 Intervalles de confiance grâce au TLC
- 4 **Prédictions :**

$$\widehat{X}_{n+h} = \widehat{a}(n+h) + \widehat{s}(n+h) + \sum_{j=1}^d \widehat{\mu}_j \widehat{Y}_{n+h}^{(j)}$$

où  $\widehat{Y}_{n+h}^{(j)}$  prédicteurs des  $Y_{n+h}^{(j)}$ .

# Plan du cours

- 1 Définitions et premières propriétés
  - Séries chronologiques
  - Stationnarité
  - Première analyse statistique de la dépendance
- 2 Tendances et saisonnalités
  - Définition et propriétés
  - Estimation semi-paramétrique de la tendance et de la saisonnalité
  - Estimation non-paramétrique de la tendance et de la saisonnalité
- 3 Exemples de modèles de séries stationnaires
  - Processus ARMA
  - Processus GARCH
  - Autres exemples de séries temporelles
- 4 Estimation, sélection de modèle, test et prédiction
  - Estimation semi-paramétrique
  - Sélection de modèles et test d'adéquation
  - Prédiction

# Semi-paramétrique $\implies$ non-paramétrique

Semi-paramétrique :

- On fixe un modèle à l'avance dépendant de paramètres pour la tendance;
- La saisonnalité peut être estimée dans toute sa généralité.

Non-paramétrique :

- Pas de modèle fixé pour la tendance, mais sa "régularité" à choisir;
- La "régularité" de la saisonnalité peut aussi être choisie.

# Semi-paramétrique $\implies$ non-paramétrique

## Semi-paramétrique :

- On fixe un modèle à l'avance dépendant de paramètres pour la tendance ;
- La saisonnalité peut être estimée dans toute sa généralité.

## Non-paramétrique :

- Pas de modèle fixé pour la tendance, mais sa "régularité" à choisir ;
- La "régularité" de la saisonnalité peut aussi être choisie.

# Semi-paramétrique $\implies$ non-paramétrique

Semi-paramétrique :

- On fixe un modèle à l'avance dépendant de paramètres pour la tendance;
- La saisonnalité peut être estimée dans toute sa généralité.

Non-paramétrique :

- Pas de modèle fixé pour la tendance, mais sa "régularité" à choisir;
- La "régularité" de la saisonnalité peut aussi être choisie.

## Semi-paramétrique $\implies$ non-paramétrique

Semi-paramétrique :

- On fixe un modèle à l'avance dépendant de paramètres pour la tendance;
- La saisonnalité peut être estimée dans toute sa généralité.

Non-paramétrique :

- Pas de modèle fixé pour la tendance, mais sa "régularité" à choisir;
- La "régularité" de la saisonnalité peut aussi être choisie.

## Semi-paramétrique $\implies$ non-paramétrique

Semi-paramétrique :

- On fixe un modèle à l'avance dépendant de paramètres pour la tendance;
- La saisonnalité peut être estimée dans toute sa généralité.

Non-paramétrique :

- Pas de modèle fixé pour la tendance, mais sa "régularité" à choisir;
- La "régularité" de la saisonnalité peut aussi être choisie.

# Semi-paramétrique $\implies$ non-paramétrique

Semi-paramétrique :

- On fixe un modèle à l'avance dépendant de paramètres pour la tendance;
- La saisonnalité peut être estimée dans toute sa généralité.

Non-paramétrique :

- Pas de modèle fixé pour la tendance, mais sa "régularité" à choisir;
- La "régularité" de la saisonnalité peut aussi être choisie.

# Semi-paramétrique $\implies$ non-paramétrique

Semi-paramétrique :

- On fixe un modèle à l'avance dépendant de paramètres pour la tendance;
- La saisonnalité peut être estimée dans toute sa généralité.

Non-paramétrique :

- Pas de modèle fixé pour la tendance, mais sa "régularité" à choisir;
- La "régularité" de la saisonnalité peut aussi être choisie.

# Semi-paramétrique $\implies$ non-paramétrique

Semi-paramétrique :

- On fixe un modèle à l'avance dépendant de paramètres pour la tendance;
- La saisonnalité peut être estimée dans toute sa généralité.

Non-paramétrique :

- Pas de modèle fixé pour la tendance, mais sa "régularité" à choisir;
- La "régularité" de la saisonnalité peut aussi être choisie.

# Moyennes mobiles

## Définition

Soit  $(X_t)$  une série temporelle.

Pour  $m \in \mathbb{N}^*$ , l'estimation par moyenne mobile de la tendance de  $X$  est :

$$\begin{cases} \hat{a}_m(t) = \frac{1}{m} \sum_{i=-[m/2]}^{[m/2]} X_{t+i} & \text{si } m \text{ impair} \\ \hat{a}_m(t) = \frac{1}{2m} (X_{t-m} + X_{t+m}) + \frac{1}{m} \sum_{i=-m/2+1}^{m/2-1} X_{t+i} & \text{si } m \text{ pair} \end{cases}$$

## Propriété

Soit  $X_t = a(t) + \varepsilon_t$  série avec tendance et  $(\varepsilon_t)$  bruit blanc. Alors

$$\text{var}(\hat{a}_m(t)) = \frac{\sigma_\varepsilon^2}{m} \quad \text{ou} \quad \frac{(m-1/2)}{m^2} \sigma_\varepsilon^2$$

$\implies$  si  $a$  varie peu,  $\hat{a}_m(t) \simeq a(t)$  et  $\text{var}(\hat{a}_m(t)) \rightarrow 0$  quand  $m$  grand

# Moyennes mobiles

## Définition

Soit  $(X_t)$  une série temporelle.

Pour  $m \in \mathbb{N}^*$ , l'estimation par **moyenne mobile** de la tendance de  $X$  est :

$$\begin{cases} \hat{a}_m(t) = \frac{1}{m} \sum_{i=-[m/2]}^{[m/2]} X_{t+i} & \text{si } m \text{ impair} \\ \hat{a}_m(t) = \frac{1}{2m} (X_{t-m} + X_{t+m}) + \frac{1}{m} \sum_{i=-m/2+1}^{m/2-1} X_{t+i} & \text{si } m \text{ pair} \end{cases}$$

## Propriété

Soit  $X_t = a(t) + \varepsilon_t$  série avec tendance et  $(\varepsilon_t)$  bruit blanc. Alors

$$\text{var}(\hat{a}_m(t)) = \frac{\sigma_\varepsilon^2}{m} \quad \text{ou} \quad \frac{(m-1/2)}{m^2} \sigma_\varepsilon^2$$

$\implies$  si  $a$  varie peu,  $\hat{a}_m(t) \simeq a(t)$  et  $\text{var}(\hat{a}_m(t)) \rightarrow 0$  quand  $m$  grand

# Moyennes mobiles

## Définition

Soit  $(X_t)$  une série temporelle.

Pour  $m \in \mathbb{N}^*$ , l'estimation par **moyenne mobile** de la tendance de  $X$  est :

$$\begin{cases} \hat{a}_m(t) = \frac{1}{m} \sum_{i=-[m/2]}^{[m/2]} X_{t+i} & \text{si } m \text{ impair} \\ \hat{a}_m(t) = \frac{1}{2m} (X_{t-m} + X_{t+m}) + \frac{1}{m} \sum_{i=-m/2+1}^{m/2-1} X_{t+i} & \text{si } m \text{ pair} \end{cases}$$

## Propriété

Soit  $X_t = a(t) + \varepsilon_t$  série avec tendance et  $(\varepsilon_t)$  bruit blanc. Alors

$$\text{var}(\hat{a}_m(t)) = \frac{\sigma_\varepsilon^2}{m} \quad \text{ou} \quad \frac{(m-1/2)}{m^2} \sigma_\varepsilon^2$$

$\implies$  si  $a$  varie peu,  $\hat{a}_m(t) \simeq a(t)$  et  $\text{var}(\hat{a}_m(t)) \rightarrow 0$  quand  $m$  grand

# Moyennes mobiles

## Définition

Soit  $(X_t)$  une série temporelle.

Pour  $m \in \mathbb{N}^*$ , l'estimation par **moyenne mobile** de la tendance de  $X$  est :

$$\begin{cases} \hat{a}_m(t) = \frac{1}{m} \sum_{i=-[m/2]}^{[m/2]} X_{t+i} & \text{si } m \text{ impair} \\ \hat{a}_m(t) = \frac{1}{2m} (X_{t-m} + X_{t+m}) + \frac{1}{m} \sum_{i=-m/2+1}^{m/2-1} X_{t+i} & \text{si } m \text{ pair} \end{cases}$$

## Propriété

Soit  $X_t = a(t) + \varepsilon_t$  série avec tendance et  $(\varepsilon_t)$  bruit blanc. Alors

$$\text{var}(\hat{a}_m(t)) = \frac{\sigma_\varepsilon^2}{m} \quad \text{ou} \quad \frac{(m-1/2)}{m^2} \sigma_\varepsilon^2$$

$\implies$  si  $a$  varie peu,  $\hat{a}_m(t) \simeq a(t)$  et  $\text{var}(\hat{a}_m(t)) \rightarrow 0$  quand  $m$  grand

# Moyennes mobiles

## Définition

Soit  $(X_t)$  une série temporelle.

Pour  $m \in \mathbb{N}^*$ , l'estimation par **moyenne mobile** de la tendance de  $X$  est :

$$\begin{cases} \hat{a}_m(t) = \frac{1}{m} \sum_{i=-\lfloor m/2 \rfloor}^{\lfloor m/2 \rfloor} X_{t+i} & \text{si } m \text{ impair} \\ \hat{a}_m(t) = \frac{1}{2m} (X_{t-m} + X_{t+m}) + \frac{1}{m} \sum_{i=-m/2+1}^{m/2-1} X_{t+i} & \text{si } m \text{ pair} \end{cases}$$

## Propriété

Soit  $X_t = a(t) + \varepsilon_t$  série avec tendance et  $(\varepsilon_t)$  bruit blanc. Alors

$$\text{var}(\hat{a}_m(t)) = \frac{\sigma_\varepsilon^2}{m} \quad \text{ou} \quad \frac{(m-1/2)}{m^2} \sigma_\varepsilon^2$$

$\implies$  si  $a$  varie peu,  $\hat{a}_m(t) \simeq a(t)$  et  $\text{var}(\hat{a}_m(t)) \rightarrow 0$  quand  $m$  grand

# Moyennes mobiles

## Définition

Soit  $(X_t)$  une série temporelle.

Pour  $m \in \mathbb{N}^*$ , l'estimation par **moyenne mobile** de la tendance de  $X$  est :

$$\begin{cases} \hat{a}_m(t) = \frac{1}{m} \sum_{i=-\lfloor m/2 \rfloor}^{\lfloor m/2 \rfloor} X_{t+i} & \text{si } m \text{ impair} \\ \hat{a}_m(t) = \frac{1}{2m} (X_{t-m} + X_{t+m}) + \frac{1}{m} \sum_{i=-m/2+1}^{m/2-1} X_{t+i} & \text{si } m \text{ pair} \end{cases}$$

## Propriété

Soit  $X_t = a(t) + \varepsilon_t$  série avec tendance et  $(\varepsilon_t)$  bruit blanc. Alors

$$\text{var}(\hat{a}_m(t)) = \frac{\sigma_\varepsilon^2}{m} \quad \text{ou} \quad \frac{(m-1/2)}{m^2} \sigma_\varepsilon^2$$

$\implies$  si  $a$  varie peu,  $\hat{a}_m(t) \simeq a(t)$  et  $\text{var}(\hat{a}_m(t)) \rightarrow 0$  quand  $m$  grand

# Moyennes mobiles

## Définition

Soit  $(X_t)$  une série temporelle.

Pour  $m \in \mathbb{N}^*$ , l'estimation par **moyenne mobile** de la tendance de  $X$  est :

$$\begin{cases} \hat{a}_m(t) = \frac{1}{m} \sum_{i=-[m/2]}^{[m/2]} X_{t+i} & \text{si } m \text{ impair} \\ \hat{a}_m(t) = \frac{1}{2m} (X_{t-m} + X_{t+m}) + \frac{1}{m} \sum_{i=-m/2+1}^{m/2-1} X_{t+i} & \text{si } m \text{ pair} \end{cases}$$

## Propriété

Soit  $X_t = a(t) + \varepsilon_t$  série avec tendance et  $(\varepsilon_t)$  bruit blanc. Alors

$$\text{var}(\hat{a}_m(t)) = \frac{\sigma_\varepsilon^2}{m} \quad \text{ou} \quad \frac{(m-1/2)}{m^2} \sigma_\varepsilon^2$$

$\implies$  si  $a$  varie peu,  $\hat{a}_m(t) \simeq a(t)$  et  $\text{var}(\hat{a}_m(t)) \rightarrow 0$  quand  $m$  grand

## Moyennes mobiles (2)

Rappel : pour  $s$  saisonnalité de période  $T$ ,  $\sum_{i=k}^{k+T-1} s(i) = 0$ .

⇒ Moyenne mobile de  $s$  sur une (ou plusieurs) période(s) s'annule.

⇒ Si  $X_t = a(t) + s(t) + u_t$ , procédure d'estimation :

- 1 On estime  $a(t)$  par  $\hat{a}_m(t)$  avec  $m = k T$  sur  $(X_t)_t$ ;
- 2 On estime  $s(t)$  avec  $\hat{s}(t)$  par LSE sur  $(X_t - \hat{a}_m(t))_t$ ;
- 3 On estime  $a(t)$  par  $\hat{a}_m(t)$  avec  $m = k T$  sur  $(X_t - \hat{s}(t))_t$ .

⇒ Choix de  $k$  ?

## Moyennes mobiles (2)

Rappel : pour  $s$  saisonnalité de période  $T$ ,  $\sum_{i=k}^{k+T-1} s(i) = 0$ .

⇒ Moyenne mobile de  $s$  sur une (ou plusieurs) période(s) s'annule.

⇒ Si  $X_t = a(t) + s(t) + u_t$ , procédure d'estimation :

- 1 On estime  $a(t)$  par  $\hat{a}_m(t)$  avec  $m = k T$  sur  $(X_t)_t$ ;
- 2 On estime  $s(t)$  avec  $\hat{s}(t)$  par LSE sur  $(X_t - \hat{a}_m(t))_t$ ;
- 3 On estime  $a(t)$  par  $\hat{a}_m(t)$  avec  $m = k T$  sur  $(X_t - \hat{s}(t))_t$ .

⇒ Choix de  $k$  ?

## Moyennes mobiles (2)

Rappel : pour  $s$  saisonnalité de période  $T$ ,  $\sum_{i=k}^{k+T-1} s(i) = 0$ .

⇒ Moyenne mobile de  $s$  sur une (ou plusieurs) période(s) s'annule.

⇒ Si  $X_t = a(t) + s(t) + u_t$ , procédure d'estimation :

- 1 On estime  $a(t)$  par  $\hat{a}_m(t)$  avec  $m = k T$  sur  $(X_t)_t$ ;
- 2 On estime  $s(t)$  avec  $\hat{s}(t)$  par LSE sur  $(X_t - \hat{a}_m(t))_t$ ;
- 3 On estime  $a(t)$  par  $\hat{a}_m(t)$  avec  $m = k T$  sur  $(X_t - \hat{s}(t))_t$ .

⇒ Choix de  $k$  ?

## Moyennes mobiles (2)

Rappel : pour  $s$  saisonnalité de période  $T$ ,  $\sum_{i=k}^{k+T-1} s(i) = 0$ .

⇒ Moyenne mobile de  $s$  sur une (ou plusieurs) période(s) s'annule.

⇒ Si  $X_t = a(t) + s(t) + u_t$ , procédure d'estimation :

- 1 On estime  $a(t)$  par  $\hat{a}_m(t)$  avec  $m = k T$  sur  $(X_t)_t$ ;
- 2 On estime  $s(t)$  avec  $\hat{s}(t)$  par LSE sur  $(X_t - \hat{a}_m(t))_t$ ;
- 3 On estime  $a(t)$  par  $\hat{a}_m(t)$  avec  $m = k T$  sur  $(X_t - \hat{s}(t))_t$ .

⇒ Choix de  $k$  ?

## Moyennes mobiles (2)

Rappel : pour  $s$  saisonnalité de période  $T$ ,  $\sum_{i=k}^{k+T-1} s(i) = 0$ .

⇒ Moyenne mobile de  $s$  sur une (ou plusieurs) période(s) s'annule.

⇒ Si  $X_t = a(t) + s(t) + u_t$ , procédure d'estimation :

- 1 On estime  $a(t)$  par  $\hat{a}_m(t)$  avec  $m = k T$  sur  $(X_t)_t$ ;
- 2 On estime  $s(t)$  avec  $\hat{s}(t)$  par LSE sur  $(X_t - \hat{a}_m(t))_t$ ;
- 3 On estime  $a(t)$  par  $\hat{a}_m(t)$  avec  $m = k T$  sur  $(X_t - \hat{s}(t))_t$ .

⇒ Choix de  $k$  ?

## Moyennes mobiles (2)

Rappel : pour  $s$  saisonnalité de période  $T$ ,  $\sum_{i=k}^{k+T-1} s(i) = 0$ .

⇒ Moyenne mobile de  $s$  sur une (ou plusieurs) période(s) s'annule.

⇒ Si  $X_t = a(t) + s(t) + u_t$ , procédure d'estimation :

- 1 On estime  $a(t)$  par  $\hat{a}_m(t)$  avec  $m = k T$  sur  $(X_t)_t$ ;
- 2 On estime  $s(t)$  avec  $\hat{s}(t)$  par LSE sur  $(X_t - \hat{a}_m(t))_t$ ;
- 3 On estime  $a(t)$  par  $\hat{a}_m(t)$  avec  $m = k T$  sur  $(X_t - \hat{s}(t))_t$ .

⇒ Choix de  $k$  ?

## Moyennes mobiles (2)

Rappel : pour  $s$  saisonnalité de période  $T$ ,  $\sum_{i=k}^{k+T-1} s(i) = 0$ .

$\implies$  Moyenne mobile de  $s$  sur une (ou plusieurs) période(s) s'annule.

$\implies$  Si  $X_t = a(t) + s(t) + u_t$ , procédure d'estimation :

- 1 On estime  $a(t)$  par  $\hat{a}_m(t)$  avec  $m = k T$  sur  $(X_t)_t$ ;
- 2 On estime  $s(t)$  avec  $\hat{s}(t)$  par LSE sur  $(X_t - \hat{a}_m(t))_t$ ;
- 3 On estime  $a(t)$  par  $\hat{a}_m(t)$  avec  $m = k T$  sur  $(X_t - \hat{s}(t))_t$ .

$\implies$  Choix de  $k$  ?

# Procédure de validation croisée

Soit  $(X_t)$  série temporelle et  $(X_1, \dots, X_n)$  trajectoire observée

$\widehat{S}_h$  stat. non-paramétrique de  $(X_1, \dots, X_n)$  dépendant de  $h > 0$ . Choix de  $h$  ?

**Remarque** : Séries temporelles : attention aux bases apprentissage/test !

## Définition

*Procédure de validation croisée (cross-validation) :*

- 1 On définit une grille  $\{h_1, \dots, h_m\}$  de valeurs de  $h$  ;
- 2 Pour  $h$  et  $i = 1, \dots, n$ , calcul  $\widehat{S}_h^{(i)}$  sur  $(X_1, \dots, X_{i-1}, X_{i+1}, \dots, X_n)$
- 3 Pour chaque  $h$ , calcul de  $\widehat{SCR}(h) = \sum_{i=1}^n (X_i - \widehat{S}_h^{(i)})^2$
- 4 On choisit  $\widehat{h} = \underset{h \in \{h_1, \dots, h_m\}}{\operatorname{argmin}} \{ \widehat{SCR}(h) \}$

## Procédure de validation croisée

Soit  $(X_t)$  série temporelle et  $(X_1, \dots, X_n)$  trajectoire observée  
 $\widehat{S}_h$  stat. non-paramétrique de  $(X_1, \dots, X_n)$  dépendant de  $h > 0$ . Choix de  $h$  ?

Remarque : Séries temporelles : attention aux bases apprentissage/test !

### Définition

Procédure de validation croisée (cross-validation) :

- 1 On définit une grille  $\{h_1, \dots, h_m\}$  de valeurs de  $h$  ;
- 2 Pour  $h$  et  $i = 1, \dots, n$ , calcul  $\widehat{S}_h^{(i)}$  sur  $(X_1, \dots, X_{i-1}, X_{i+1}, \dots, X_n)$
- 3 Pour chaque  $h$ , calcul de  $\widehat{SCR}(h) = \sum_{i=1}^n (X_i - \widehat{S}_h^{(i)})^2$
- 4 On choisit  $\widehat{h} = \underset{h \in \{h_1, \dots, h_m\}}{\operatorname{argmin}} \{ \widehat{SCR}(h) \}$

## Procédure de validation croisée

Soit  $(X_t)$  série temporelle et  $(X_1, \dots, X_n)$  trajectoire observée  
 $\widehat{S}_h$  stat. non-paramétrique de  $(X_1, \dots, X_n)$  dépendant de  $h > 0$ . Choix de  $h$  ?

**Remarque** : Séries temporelles : attention aux bases apprentissage/test !

### Définition

*Procédure de validation croisée (cross-validation) :*

- 1 On définit une grille  $\{h_1, \dots, h_m\}$  de valeurs de  $h$  ;
- 2 Pour  $h$  et  $i = 1, \dots, n$ , calcul  $\widehat{S}_h^{(i)}$  sur  $(X_1, \dots, X_{i-1}, X_{i+1}, \dots, X_n)$
- 3 Pour chaque  $h$ , calcul de  $\widehat{SCR}(h) = \sum_{i=1}^n (X_i - \widehat{S}_h^{(i)})^2$
- 4 On choisit  $\widehat{h} = \underset{h \in \{h_1, \dots, h_m\}}{\operatorname{argmin}} \{ \widehat{SCR}(h) \}$

## Procédure de validation croisée

Soit  $(X_t)$  série temporelle et  $(X_1, \dots, X_n)$  trajectoire observée  
 $\widehat{S}_h$  stat. non-paramétrique de  $(X_1, \dots, X_n)$  dépendant de  $h > 0$ . Choix de  $h$  ?

**Remarque** : Séries temporelles : attention aux bases apprentissage/test !

### Définition

Procédure de **validation croisée** (*cross-validation*) :

- 1 On définit une grille  $\{h_1, \dots, h_m\}$  de valeurs de  $h$  ;
- 2 Pour  $h$  et  $i = 1, \dots, n$ , calcul  $\widehat{S}_h^{(i)}$  sur  $(X_1, \dots, X_{i-1}, X_{i+1}, \dots, X_n)$
- 3 Pour chaque  $h$ , calcul de  $\widehat{SCR}(h) = \sum_{i=1}^n (X_i - \widehat{S}_h^{(i)})^2$
- 4 On choisit  $\widehat{h} = \underset{h \in \{h_1, \dots, h_m\}}{\operatorname{argmin}} \{ \widehat{SCR}(h) \}$

## Procédure de validation croisée

Soit  $(X_t)$  série temporelle et  $(X_1, \dots, X_n)$  trajectoire observée  
 $\widehat{S}_h$  stat. non-paramétrique de  $(X_1, \dots, X_n)$  dépendant de  $h > 0$ . Choix de  $h$ ?

**Remarque** : Séries temporelles : attention aux bases apprentissage/test !

### Définition

Procédure de **validation croisée** (*cross-validation*) :

- 1 On définit une grille  $\{h_1, \dots, h_m\}$  de valeurs de  $h$  ;
- 2 Pour  $h$  et  $i = 1, \dots, n$ , calcul  $\widehat{S}_h^{(i)}$  sur  $(X_1, \dots, X_{i-1}, X_{i+1}, \dots, X_n)$
- 3 Pour chaque  $h$ , calcul de  $\widehat{SCR}(h) = \sum_{i=1}^n (X_i - \widehat{S}_h^{(i)})^2$
- 4 On choisit  $\widehat{h} = \underset{h \in \{h_1, \dots, h_m\}}{\operatorname{argmin}} \{ \widehat{SCR}(h) \}$

## Procédure de validation croisée

Soit  $(X_t)$  série temporelle et  $(X_1, \dots, X_n)$  trajectoire observée  
 $\widehat{S}_h$  stat. non-paramétrique de  $(X_1, \dots, X_n)$  dépendant de  $h > 0$ . Choix de  $h$  ?

**Remarque** : Séries temporelles : attention aux bases apprentissage/test !

### Définition

Procédure de **validation croisée** (*cross-validation*) :

- 1 On définit une grille  $\{h_1, \dots, h_m\}$  de valeurs de  $h$  ;
- 2 Pour  $h$  et  $i = 1, \dots, n$ , calcul  $\widehat{S}_h^{(i)}$  sur  $(X_1, \dots, X_{i-1}, X_{i+1}, \dots, X_n)$
- 3 Pour chaque  $h$ , calcul de  $\widehat{SCR}(h) = \sum_{i=1}^n (X_i - \widehat{S}_h^{(i)})^2$
- 4 On choisit  $\widehat{h} = \underset{h \in \{h_1, \dots, h_m\}}{\operatorname{argmin}} \{ \widehat{SCR}(h) \}$

## Procédure de validation croisée

Soit  $(X_t)$  série temporelle et  $(X_1, \dots, X_n)$  trajectoire observée  
 $\widehat{S}_h$  stat. non-paramétrique de  $(X_1, \dots, X_n)$  dépendant de  $h > 0$ . Choix de  $h$  ?

**Remarque** : Séries temporelles : attention aux bases apprentissage/test !

### Définition

Procédure de **validation croisée** (*cross-validation*) :

- 1 On définit une grille  $\{h_1, \dots, h_m\}$  de valeurs de  $h$  ;
- 2 Pour  $h$  et  $i = 1, \dots, n$ , calcul  $\widehat{S}_h^{(i)}$  sur  $(X_1, \dots, X_{i-1}, X_{i+1}, \dots, X_n)$
- 3 Pour chaque  $h$ , calcul de  $\widehat{SCR}(h) = \sum_{i=1}^n (X_i - \widehat{S}_h^{(i)})^2$
- 4 On choisit  $\widehat{h} = \underset{h \in \{h_1, \dots, h_m\}}{\operatorname{argmin}} \{ \widehat{SCR}(h) \}$

## Procédure de validation croisée

Soit  $(X_t)$  série temporelle et  $(X_1, \dots, X_n)$  trajectoire observée  
 $\widehat{S}_h$  stat. non-paramétrique de  $(X_1, \dots, X_n)$  dépendant de  $h > 0$ . Choix de  $h$  ?

**Remarque** : Séries temporelles : attention aux bases apprentissage/test !

### Définition

Procédure de **validation croisée** (*cross-validation*) :

- 1 On définit une grille  $\{h_1, \dots, h_m\}$  de valeurs de  $h$  ;
- 2 Pour  $h$  et  $i = 1, \dots, n$ , calcul  $\widehat{S}_h^{(i)}$  sur  $(X_1, \dots, X_{i-1}, X_{i+1}, \dots, X_n)$
- 3 Pour chaque  $h$ , calcul de  $\widehat{SCR}(h) = \sum_{i=1}^n (X_i - \widehat{S}_h^{(i)})^2$
- 4 On choisit  $\widehat{h} = \underset{h \in \{h_1, \dots, h_m\}}{\operatorname{argmin}} \{ \widehat{SCR}(h) \}$

# Estimateur par noyau de la tendance

## Définition

On appelle noyau  $K : \mathbb{R} \rightarrow \mathbb{R}$  fonction telle que  $\int_{\mathbb{R}} K(t)dt = 1$  et  $K(0) > 0$ .

**Exemple :**  $K_1(x) = \mathbb{1}_{x \in [-\frac{1}{2}, \frac{1}{2}]}$ ,  $K_2(x) = \frac{e^{-x^2}}{\sqrt{2\pi}}$ ,  $K_3(x) = \frac{3}{4}(1 - x^2)\mathbb{1}_{x \in [-1, 1]}$

Moyenne mobile : Pour  $0 < h < 1$ ,  $\hat{a}_{nh}(t) \simeq \frac{\sum_{j=1}^n X_j K_1\left(\frac{t-j}{nh}\right)}{\sum_{j=1}^n K_1\left(\frac{t-j}{nh}\right)}$

$\implies$  On remplace  $K_1$  par un autre noyau  $K$

$$\hat{a}_{n,h}^{ker}(t) = \frac{\sum_{j=1}^n X_j K\left(\frac{t-j}{nh}\right)}{\sum_{j=1}^n K\left(\frac{t-j}{nh}\right)}$$

$\implies$  Avec  $K_2$  ou  $K_3$ , donne plus de poids aux valeurs proches de  $X_t$

**Remarque :** Choix de  $h$  par validation croisée!

# Estimateur par noyau de la tendance

## Définition

On appelle noyau  $K : \mathbb{R} \rightarrow \mathbb{R}$  fonction telle que  $\int_{\mathbb{R}} K(t)dt = 1$  et  $K(0) > 0$ .

Exemple :  $K_1(x) = \mathbb{1}_{x \in [-\frac{1}{2}, \frac{1}{2}]}$ ,  $K_2(x) = \frac{e^{-x^2}}{\sqrt{2\pi}}$ ,  $K_3(x) = \frac{3}{4}(1-x^2)\mathbb{1}_{x \in [-1,1]}$

Moyenne mobile : Pour  $0 < h < 1$ ,  $\hat{a}_{nh}(t) \simeq \frac{\sum_{j=1}^n X_j K_1\left(\frac{t-j}{nh}\right)}{\sum_{j=1}^n K_1\left(\frac{t-j}{nh}\right)}$

$\implies$  On remplace  $K_1$  par un autre noyau  $K$

$$\hat{a}_{n,h}^{ker}(t) = \frac{\sum_{j=1}^n X_j K\left(\frac{t-j}{nh}\right)}{\sum_{j=1}^n K\left(\frac{t-j}{nh}\right)}$$

$\implies$  Avec  $K_2$  ou  $K_3$ , donne plus de poids aux valeurs proches de  $X_t$

Remarque : Choix de  $h$  par validation croisée!

# Estimateur par noyau de la tendance

## Définition

On appelle noyau  $K : \mathbb{R} \rightarrow \mathbb{R}$  fonction telle que  $\int_{\mathbb{R}} K(t)dt = 1$  et  $K(0) > 0$ .

**Exemple :**  $K_1(x) = \mathbb{1}_{x \in [-\frac{1}{2}, \frac{1}{2}]}$ ,  $K_2(x) = \frac{e^{-x^2}}{\sqrt{2\pi}}$ ,  $K_3(x) = \frac{3}{4}(1 - x^2)\mathbb{1}_{x \in [-1, 1]}$

Moyenne mobile : Pour  $0 < h < 1$ ,  $\hat{a}_{nh}(t) \simeq \frac{\sum_{j=1}^n X_j K_1\left(\frac{t-j}{nh}\right)}{\sum_{j=1}^n K_1\left(\frac{t-j}{nh}\right)}$

$\implies$  On remplace  $K_1$  par un autre noyau  $K$

$$\hat{a}_{n,h}^{ker}(t) = \frac{\sum_{j=1}^n X_j K\left(\frac{t-j}{nh}\right)}{\sum_{j=1}^n K\left(\frac{t-j}{nh}\right)}$$

$\implies$  Avec  $K_2$  ou  $K_3$ , donne plus de poids aux valeurs proches de  $X_t$

Remarque : Choix de  $h$  par validation croisée!

# Estimateur par noyau de la tendance

## Définition

On appelle noyau  $K : \mathbb{R} \rightarrow \mathbb{R}$  fonction telle que  $\int_{\mathbb{R}} K(t)dt = 1$  et  $K(0) > 0$ .

**Exemple :**  $K_1(x) = \mathbb{1}_{x \in [-\frac{1}{2}, \frac{1}{2}]}$ ,  $K_2(x) = \frac{e^{-x^2}}{\sqrt{2\pi}}$ ,  $K_3(x) = \frac{3}{4}(1 - x^2)\mathbb{1}_{x \in [-1, 1]}$

Moyenne mobile : Pour  $0 < h < 1$ ,  $\hat{a}_{nh}(t) \simeq \frac{\sum_{j=1}^n X_j K_1\left(\frac{t-j}{nh}\right)}{\sum_{j=1}^n K_1\left(\frac{t-j}{nh}\right)}$

$\implies$  On remplace  $K_1$  par un autre noyau  $K$

$$\hat{a}_{n,h}^{ker}(t) = \frac{\sum_{j=1}^n X_j K\left(\frac{t-j}{nh}\right)}{\sum_{j=1}^n K\left(\frac{t-j}{nh}\right)}$$

$\implies$  Avec  $K_2$  ou  $K_3$ , donne plus de poids aux valeurs proches de  $X_t$

Remarque : Choix de  $h$  par validation croisée!

# Estimateur par noyau de la tendance

## Définition

On appelle noyau  $K : \mathbb{R} \rightarrow \mathbb{R}$  fonction telle que  $\int_{\mathbb{R}} K(t)dt = 1$  et  $K(0) > 0$ .

**Exemple :**  $K_1(x) = \mathbb{1}_{x \in [-\frac{1}{2}, \frac{1}{2}]}$ ,  $K_2(x) = \frac{e^{-x^2}}{\sqrt{2\pi}}$ ,  $K_3(x) = \frac{3}{4}(1 - x^2)\mathbb{1}_{x \in [-1, 1]}$

Moyenne mobile : Pour  $0 < h < 1$ ,  $\hat{a}_{nh}(t) \simeq \frac{\sum_{j=1}^n X_j K_1\left(\frac{t-j}{nh}\right)}{\sum_{j=1}^n K_1\left(\frac{t-j}{nh}\right)}$

$\implies$  On remplace  $K_1$  par un autre noyau  $K$

$$\hat{a}_{n,h}^{ker}(t) = \frac{\sum_{j=1}^n X_j K\left(\frac{t-j}{nh}\right)}{\sum_{j=1}^n K\left(\frac{t-j}{nh}\right)}$$

$\implies$  Avec  $K_2$  ou  $K_3$ , donne plus de poids aux valeurs proches de  $X_t$

Remarque : Choix de  $h$  par validation croisée!

# Estimateur par noyau de la tendance

## Définition

On appelle noyau  $K : \mathbb{R} \rightarrow \mathbb{R}$  fonction telle que  $\int_{\mathbb{R}} K(t)dt = 1$  et  $K(0) > 0$ .

**Exemple :**  $K_1(x) = \mathbb{1}_{x \in [-\frac{1}{2}, \frac{1}{2}]}$ ,  $K_2(x) = \frac{e^{-x^2}}{\sqrt{2\pi}}$ ,  $K_3(x) = \frac{3}{4}(1 - x^2)\mathbb{1}_{x \in [-1, 1]}$

Moyenne mobile : Pour  $0 < h < 1$ ,  $\hat{a}_{nh}(t) \simeq \frac{\sum_{j=1}^n X_j K_1\left(\frac{t-j}{nh}\right)}{\sum_{j=1}^n K_1\left(\frac{t-j}{nh}\right)}$

$\implies$  On remplace  $K_1$  par un autre noyau  $K$

$$\hat{a}_{n,h}^{ker}(t) = \frac{\sum_{j=1}^n X_j K\left(\frac{t-j}{nh}\right)}{\sum_{j=1}^n K\left(\frac{t-j}{nh}\right)}$$

$\implies$  Avec  $K_2$  ou  $K_3$ , donne plus de poids aux valeurs proches de  $X_t$

Remarque : Choix de  $h$  par validation croisée!

# Estimateur par noyau de la tendance

## Définition

On appelle noyau  $K : \mathbb{R} \rightarrow \mathbb{R}$  fonction telle que  $\int_{\mathbb{R}} K(t)dt = 1$  et  $K(0) > 0$ .

**Exemple :**  $K_1(x) = \mathbb{1}_{x \in [-\frac{1}{2}, \frac{1}{2}]}$ ,  $K_2(x) = \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}}$ ,  $K_3(x) = \frac{3}{4}(1 - x^2)\mathbb{1}_{x \in [-1, 1]}$

Moyenne mobile : Pour  $0 < h < 1$ ,  $\hat{a}_{nh}(t) \simeq \frac{\sum_{j=1}^n X_j K_1\left(\frac{t-j}{nh}\right)}{\sum_{j=1}^n K_1\left(\frac{t-j}{nh}\right)}$

$\implies$  On remplace  $K_1$  par un autre noyau  $K$

$$\hat{a}_{n,h}^{ker}(t) = \frac{\sum_{j=1}^n X_j K\left(\frac{t-j}{nh}\right)}{\sum_{j=1}^n K\left(\frac{t-j}{nh}\right)}$$

$\implies$  Avec  $K_2$  ou  $K_3$ , donne plus de poids aux valeurs proches de  $X_t$

Remarque : Choix de  $h$  par validation croisée!

# Estimateur par noyau de la tendance

## Définition

On appelle noyau  $K : \mathbb{R} \rightarrow \mathbb{R}$  fonction telle que  $\int_{\mathbb{R}} K(t)dt = 1$  et  $K(0) > 0$ .

**Exemple :**  $K_1(x) = \mathbb{1}_{x \in [-\frac{1}{2}, \frac{1}{2}]}$ ,  $K_2(x) = \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}}$ ,  $K_3(x) = \frac{3}{4}(1 - x^2)\mathbb{1}_{x \in [-1, 1]}$

Moyenne mobile : Pour  $0 < h < 1$ ,  $\hat{a}_{nh}(t) \simeq \frac{\sum_{j=1}^n X_j K_1\left(\frac{t-j}{nh}\right)}{\sum_{j=1}^n K_1\left(\frac{t-j}{nh}\right)}$

$\implies$  On remplace  $K_1$  par un autre noyau  $K$

$$\hat{a}_{n,h}^{ker}(t) = \frac{\sum_{j=1}^n X_j K\left(\frac{t-j}{nh}\right)}{\sum_{j=1}^n K\left(\frac{t-j}{nh}\right)}$$

$\implies$  Avec  $K_2$  ou  $K_3$ , donne plus de poids aux valeurs proches de  $X_t$

**Remarque :** Choix de  $h$  par validation croisée !

# Estimateur par lissage local de la tendance

Autre estimateur non-paramétrique possible :

$$\hat{a}_\beta^{(0)}(t) = \frac{\sum_{i=1}^n W_\beta(|t - i|) X_i}{\sum_{i=1}^n W_\beta(|t - i|)},$$

$W_\beta$  fonction "poids" (à valeurs dans  $[0, \infty[$  et décroissants), par exemple :

- Poids exponentiels :  $W_\beta(x) = \beta^x$  et  $\beta \in [0, 1]$  (si  $\beta = 1$ ,  $\hat{a}_\beta(t) = \bar{X}_n$ );
- Poids tricubes :  $W_\beta(x) = \left(1 - \left(\frac{x}{\beta n}\right)^3\right)^3 \mathbb{1}_{x \leq \beta n}$  et  $\beta \in [0, 1]$ ,  $\beta n$  fenêtre.

⇒ estimateur LOESS (LOcally Estimated Scatterplot Smoothing).

⇒ Extension à une régression pondérée de degré 1 ou 2.

**Remarque** : Choix de  $\beta$  par validation croisée !

# Estimateur par lissage local de la tendance

Autre estimateur non-paramétrique possible :

$$\hat{a}_\beta^{(0)}(t) = \frac{\sum_{i=1}^n W_\beta(|t - i|) X_i}{\sum_{i=1}^n W_\beta(|t - i|)},$$

$W_\beta$  fonction "poids" (à valeurs dans  $[0, \infty[$  et décroissants), par exemple :

- Poids exponentiels :  $W_\beta(x) = \beta^x$  et  $\beta \in [0, 1]$  (si  $\beta = 1$ ,  $\hat{a}_\beta(t) = \bar{X}_n$ );
- Poids tricubes :  $W_\beta(x) = \left(1 - \left(\frac{x}{\beta n}\right)^3\right)^3 \mathbb{1}_{x \leq \beta n}$  et  $\beta \in [0, 1]$ ,  $\beta n$  fenêtre.

⇒ estimateur LOESS (LOcally Estimated Scatterplot Smoothing).

⇒ Extension à une régression pondérée de degré 1 ou 2.

Remarque : Choix de  $\beta$  par validation croisée !

# Estimateur par lissage local de la tendance

Autre estimateur non-paramétrique possible :

$$\hat{a}_\beta^{(0)}(t) = \frac{\sum_{i=1}^n W_\beta(|t - i|) X_i}{\sum_{i=1}^n W_\beta(|t - i|)},$$

$W_\beta$  fonction "poids" (à valeurs dans  $[0, \infty[$  et décroissants), par exemple :

- Poids exponentiels :  $W_\beta(x) = \beta^x$  et  $\beta \in [0, 1]$  (si  $\beta = 1$ ,  $\hat{a}_\beta(t) = \bar{X}_n$ );
- Poids tricubes :  $W_\beta(x) = \left(1 - \left(\frac{x}{\beta n}\right)^3\right)^3 \mathbb{1}_{x \leq \beta n}$  et  $\beta \in [0, 1]$ ,  $\beta n$  fenêtre.

⇒ estimateur LOESS (LOcally Estimated Scatterplot Smoothing).

⇒ Extension à une régression pondérée de degré 1 ou 2.

Remarque : Choix de  $\beta$  par validation croisée !

# Estimateur par lissage local de la tendance

Autre estimateur non-paramétrique possible :

$$\hat{a}_\beta^{(0)}(t) = \frac{\sum_{i=1}^n W_\beta(|t - i|) X_i}{\sum_{i=1}^n W_\beta(|t - i|)},$$

$W_\beta$  fonction "poids" (à valeurs dans  $[0, \infty[$  et décroissants), par exemple :

- Poids exponentiels :  $W_\beta(x) = \beta^x$  et  $\beta \in [0, 1]$  (si  $\beta = 1$ ,  $\hat{a}_\beta(t) = \bar{X}_n$ );
- Poids tricubes :  $W_\beta(x) = \left(1 - \left(\frac{x}{\beta n}\right)^3\right)^3 \mathbb{1}_{x \leq \beta n}$  et  $\beta \in [0, 1]$ ,  $\beta n$  fenêtre.

⇒ estimateur LOESS (LOcally Estimated Scatterplot Smoothing).

⇒ Extension à une régression pondérée de degré 1 ou 2.

Remarque : Choix de  $\beta$  par validation croisée !

# Estimateur par lissage local de la tendance

Autre estimateur non-paramétrique possible :

$$\hat{a}_\beta^{(0)}(t) = \frac{\sum_{i=1}^n W_\beta(|t - i|) X_i}{\sum_{i=1}^n W_\beta(|t - i|)},$$

$W_\beta$  fonction "poids" (à valeurs dans  $[0, \infty[$  et décroissants), par exemple :

- Poids exponentiels :  $W_\beta(x) = \beta^x$  et  $\beta \in [0, 1]$  (si  $\beta = 1$ ,  $\hat{a}_\beta(t) = \bar{X}_n$ );
- Poids tricubes :  $W_\beta(x) = \left(1 - \left(\frac{x}{\beta n}\right)^3\right)^3 \mathbb{1}_{x \leq \beta n}$  et  $\beta \in [0, 1]$ ,  $\beta n$  fenêtre.

⇒ estimateur LOESS (LOcally Estimated Scatterplot Smoothing).

⇒ Extension à une régression pondérée de degré 1 ou 2.

Remarque : Choix de  $\beta$  par validation croisée !

# Estimateur par lissage local de la tendance

Autre estimateur non-paramétrique possible :

$$\hat{a}_\beta^{(0)}(t) = \frac{\sum_{i=1}^n W_\beta(|t - i|) X_i}{\sum_{i=1}^n W_\beta(|t - i|)},$$

$W_\beta$  fonction "poids" (à valeurs dans  $[0, \infty[$  et décroissants), par exemple :

- Poids exponentiels :  $W_\beta(x) = \beta^x$  et  $\beta \in [0, 1]$  (si  $\beta = 1$ ,  $\hat{a}_\beta(t) = \bar{X}_n$ );
- Poids tricubes :  $W_\beta(x) = \left(1 - \left(\frac{x}{\beta n}\right)^3\right)^3 \mathbb{1}_{x \leq \beta n}$  et  $\beta \in [0, 1]$ ,  $\beta n$  fenêtre.

⇒ estimateur LOESS (LOcally Estimated Scatterplot Smoothing).

⇒ Extension à une régression pondérée de degré 1 ou 2.

Remarque : Choix de  $\beta$  par validation croisée !

# Estimateur par lissage local de la tendance

Autre estimateur non-paramétrique possible :

$$\hat{a}_\beta^{(0)}(t) = \frac{\sum_{i=1}^n W_\beta(|t - i|) X_i}{\sum_{i=1}^n W_\beta(|t - i|)},$$

$W_\beta$  fonction "poids" (à valeurs dans  $[0, \infty[$  et décroissants), par exemple :

- Poids exponentiels :  $W_\beta(x) = \beta^x$  et  $\beta \in [0, 1]$  (si  $\beta = 1$ ,  $\hat{a}_\beta(t) = \bar{X}_n$ );
- Poids tricubes :  $W_\beta(x) = \left(1 - \left(\frac{x}{\beta n}\right)^3\right)^3 \mathbb{1}_{x \leq \beta n}$  et  $\beta \in [0, 1]$ ,  $\beta n$  fenêtre.

⇒ estimateur LOESS (LOcally Estimated Scatterplot Smoothing).

⇒ Extension à une régression pondérée de degré 1 ou 2.

Remarque : Choix de  $\beta$  par validation croisée !

# Estimateur par lissage local de la tendance

Autre estimateur non-paramétrique possible :

$$\hat{a}_\beta^{(0)}(t) = \frac{\sum_{i=1}^n W_\beta(|t - i|) X_i}{\sum_{i=1}^n W_\beta(|t - i|)},$$

$W_\beta$  fonction "poids" (à valeurs dans  $[0, \infty[$  et décroissants), par exemple :

- Poids exponentiels :  $W_\beta(x) = \beta^x$  et  $\beta \in [0, 1]$  (si  $\beta = 1$ ,  $\hat{a}_\beta(t) = \bar{X}_n$ );
- Poids tricubes :  $W_\beta(x) = \left(1 - \left(\frac{x}{\beta n}\right)^3\right)^3 \mathbb{1}_{x \leq \beta n}$  et  $\beta \in [0, 1]$ ,  $\beta n$  fenêtre.

⇒ estimateur LOESS (LOcally Estimated Scatterplot Smoothing).

⇒ Extension à une régression pondérée de degré 1 ou 2.

Remarque : Choix de  $\beta$  par validation croisée !

# Estimateur par lissage local de la tendance

Autre estimateur non-paramétrique possible :

$$\hat{a}_\beta^{(0)}(t) = \frac{\sum_{i=1}^n W_\beta(|t - i|) X_i}{\sum_{i=1}^n W_\beta(|t - i|)},$$

$W_\beta$  fonction "poids" (à valeurs dans  $[0, \infty[$  et décroissants), par exemple :

- Poids exponentiels :  $W_\beta(x) = \beta^x$  et  $\beta \in [0, 1]$  (si  $\beta = 1$ ,  $\hat{a}_\beta(t) = \bar{X}_n$ );
- Poids tricubes :  $W_\beta(x) = \left(1 - \left(\frac{x}{\beta n}\right)^3\right)^3 \mathbb{1}_{x \leq \beta n}$  et  $\beta \in [0, 1]$ ,  $\beta n$  fenêtre.

⇒ estimateur LOESS (LOcally Estimated Scatterplot Smoothing).

⇒ Extension à une régression pondérée de degré 1 ou 2.

**Remarque** : Choix de  $\beta$  par validation croisée !

## LOESS de degré $d$

Extension à un modèle de régression locale de degré  $d$  :

$$Z_d^t = \begin{pmatrix} 1 & t-1 & \dots & (t-1)^d \\ 1 & t-2 & \dots & (t-2)^d \\ \vdots & \vdots & \ddots & \vdots \\ 1 & t-n & \dots & (t-n)^d \end{pmatrix} \quad W_\beta^t = \begin{pmatrix} w_\beta(|t-1|) & 0 & \dots & 0 \\ 0 & w_\beta(|t-2|) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & w_\beta(|t-n|) \end{pmatrix}$$

$$\implies \hat{a}_\beta^{(d)}(t) = ((Z_d^t)^\top \times W_\beta^t \times Z_d^t)^{-1} (Z_d^t)^\top \times W_\beta^t \times \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{pmatrix}$$

$\implies$  Moindres carrés pondérés : LOESS de degré  $d$

Remarque : On retrouve  $\hat{a}_\beta^{(0)}(t)$  avec  $d = 0$

## LOESS de degré $d$

Extension à un modèle de régression locale de degré  $d$  :

$$Z_d^t = \begin{pmatrix} 1 & t-1 & \cdots & (t-1)^d \\ 1 & t-2 & \cdots & (t-2)^d \\ \vdots & \vdots & & \vdots \\ 1 & t-n & \cdots & (t-n)^d \end{pmatrix} \quad W_\beta^t = \begin{pmatrix} w_\beta(|t-1|) & 0 & \cdots & 0 \\ 0 & w_\beta(|t-2|) & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & w_\beta(|t-n|) \end{pmatrix}$$

$$\implies \hat{a}_\beta^{(d)}(t) = ((Z_d^t)^\top \times W_\beta^t \times Z_d^t)^{-1} (Z_d^t)^\top \times W_\beta^t \times \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{pmatrix}$$

$\implies$  Moindres carrés pondérés : LOESS de degré  $d$

Remarque : On retrouve  $\hat{a}_\beta^{(0)}(t)$  avec  $d = 0$

## LOESS de degré $d$

Extension à un modèle de régression locale de degré  $d$  :

$$Z_d^t = \begin{pmatrix} 1 & t-1 & \cdots & (t-1)^d \\ 1 & t-2 & \cdots & (t-2)^d \\ \vdots & \vdots & & \vdots \\ 1 & t-n & \cdots & (t-n)^d \end{pmatrix} \quad W_\beta^t = \begin{pmatrix} w_\beta(|t-1|) & 0 & \cdots & 0 \\ 0 & w_\beta(|t-2|) & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & w_\beta(|t-n|) \end{pmatrix}$$

$$\implies \hat{a}_\beta^{(d)}(t) = ((Z_d^t)^\top \times W_\beta^t \times Z_d^t)^{-1} (Z_d^t)^\top \times W_\beta^t \times \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{pmatrix}$$

$\implies$  Moindres carrés pondérés : LOESS de degré  $d$

Remarque : On retrouve  $\hat{a}_\beta^{(0)}(t)$  avec  $d = 0$

## LOESS de degré $d$

Extension à un modèle de régression locale de degré  $d$  :

$$Z_d^t = \begin{pmatrix} 1 & t-1 & \cdots & (t-1)^d \\ 1 & t-2 & \cdots & (t-2)^d \\ \vdots & \vdots & & \vdots \\ 1 & t-n & \cdots & (t-n)^d \end{pmatrix} \quad W_\beta^t = \begin{pmatrix} w_\beta(|t-1|) & 0 & \cdots & 0 \\ 0 & w_\beta(|t-2|) & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & w_\beta(|t-n|) \end{pmatrix}$$

$$\implies \hat{a}_\beta^{(d)}(t) = ((Z_d^t)^\top \times W_\beta^t \times Z_d^t)^{-1} (Z_d^t)^\top \times W_\beta^t \times \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{pmatrix}$$

$\implies$  Moindres carrés pondérés : LOESS de degré  $d$

Remarque : On retrouve  $\hat{a}_\beta^{(0)}(t)$  avec  $d = 0$

## LOESS de degré $d$

Extension à un modèle de régression locale de degré  $d$  :

$$Z_d^t = \begin{pmatrix} 1 & t-1 & \cdots & (t-1)^d \\ 1 & t-2 & \cdots & (t-2)^d \\ \vdots & \vdots & & \vdots \\ 1 & t-n & \cdots & (t-n)^d \end{pmatrix} \quad W_\beta^t = \begin{pmatrix} w_\beta(|t-1|) & 0 & \cdots & 0 \\ 0 & w_\beta(|t-2|) & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & w_\beta(|t-n|) \end{pmatrix}$$

$$\implies \hat{a}_\beta^{(d)}(t) = ((Z_d^t)^\top \times W_\beta^t \times Z_d^t)^{-1} (Z_d^t)^\top \times W_\beta^t \times \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{pmatrix}$$

$\implies$  Moindres carrés pondérés : LOESS de degré  $d$

**Remarque** : On retrouve  $\hat{a}_\beta^{(0)}(t)$  avec  $d = 0$

# Extension en cas de variables exogènes observées

Tendance et saisonnalité estimées jusqu'ici comme fonctions du temps

⇒ Extension en cas de **variables exogènes** observées aux mêmes instants

⇒ On suppose observés  $(X_t)_{1 \leq t \leq n}$  mais aussi  $(Y_t)_{1 \leq t \leq n}$  où

$$Y_t = (Y_t^{(1)}, \dots, Y_t^{(d)}), \text{ où } Y^{(1)}, \dots, Y^{(d)} \text{ variables exogènes}$$

⇒ Méthodes de l'**apprentissage statistique** :

$$X_t = h(t, s(t), Y_t^{(1)}, \dots, Y_t^{(d)}, u_t) \quad \text{pour tout } t$$

$$\text{où } s(t) = s_1 h_1(t) + \dots + s_{T-1} h_{T-1}(t)$$

⇒  $\hat{h}$  par forêts aléatoires, XGBoost, Deep learning...

## Extension en cas de variables exogènes observées

Tendance et saisonnalité estimées jusqu'ici comme fonctions du temps

⇒ Extension en cas de **variables exogènes** observées aux mêmes instants

⇒ On suppose observés  $(X_t)_{1 \leq t \leq n}$  mais aussi  $(Y_t)_{1 \leq t \leq n}$  où

$$Y_t = (Y_t^{(1)}, \dots, Y_t^{(d)}), \text{ où } Y^{(1)}, \dots, Y^{(d)} \text{ variables exogènes}$$

⇒ Méthodes de l'**apprentissage statistique** :

$$X_t = h(t, s(t), Y_t^{(1)}, \dots, Y_t^{(d)}, u_t) \quad \text{pour tout } t$$

$$\text{où } s(t) = s_1 h_1(t) + \dots + s_{T-1} h_{T-1}(t)$$

⇒  $\hat{h}$  par forêts aléatoires, XGBoost, Deep learning...

## Extension en cas de variables exogènes observées

Tendance et saisonnalité estimées jusqu'ici comme fonctions du temps

⇒ Extension en cas de **variables exogènes** observées aux mêmes instants

⇒ On suppose observés  $(X_t)_{1 \leq t \leq n}$  mais aussi  $(Y_t)_{1 \leq t \leq n}$  où

$$Y_t = (Y_t^{(1)}, \dots, Y_t^{(d)}), \text{ où } Y^{(1)}, \dots, Y^{(d)} \text{ variables exogènes}$$

⇒ Méthodes de l'**apprentissage statistique** :

$$X_t = h(t, s(t), Y_t^{(1)}, \dots, Y_t^{(d)}, u_t) \quad \text{pour tout } t$$

$$\text{où } s(t) = s_1 h_1(t) + \dots + s_{T-1} h_{T-1}(t)$$

⇒  $\hat{h}$  par forêts aléatoires, XGBoost, Deep learning...

## Extension en cas de variables exogènes observées

Tendance et saisonnalité estimées jusqu'ici comme fonctions du temps

⇒ Extension en cas de **variables exogènes** observées aux mêmes instants

⇒ On suppose observés  $(X_t)_{1 \leq t \leq n}$  mais aussi  $(Y_t)_{1 \leq t \leq n}$  où

$$Y_t = (Y_t^{(1)}, \dots, Y_t^{(d)}), \text{ où } Y^{(1)}, \dots, Y^{(d)} \text{ variables exogènes}$$

⇒ Méthodes de l'**apprentissage statistique** :

$$X_t = h(t, s(t), Y_t^{(1)}, \dots, Y_t^{(d)}, u_t) \quad \text{pour tout } t$$

$$\text{où } s(t) = s_1 h_1(t) + \dots + s_{T-1} h_{T-1}(t)$$

⇒  $\hat{h}$  par forêts aléatoires, XGBoost, Deep learning...

## Extension en cas de variables exogènes observées

Tendance et saisonnalité estimées jusqu'ici comme fonctions du temps

⇒ Extension en cas de **variables exogènes** observées aux mêmes instants

⇒ On suppose observés  $(X_t)_{1 \leq t \leq n}$  mais aussi  $(Y_t)_{1 \leq t \leq n}$  où

$$Y_t = (Y_t^{(1)}, \dots, Y_t^{(d)}), \text{ où } Y^{(1)}, \dots, Y^{(d)} \text{ variables exogènes}$$

⇒ Méthodes de l'**apprentissage statistique** :

$$X_t = h(t, s(t), Y_t^{(1)}, \dots, Y_t^{(d)}, u_t) \quad \text{pour tout } t$$

$$\text{où } s(t) = s_1 h_1(t) + \dots + s_{T-1} h_{T-1}(t)$$

⇒  $\hat{h}$  par forêts aléatoires, XGBoost, Deep learning...

## Extension en cas de variables exogènes observées

Tendance et saisonnalité estimées jusqu'ici comme fonctions du temps

⇒ Extension en cas de **variables exogènes** observées aux mêmes instants

⇒ On suppose observés  $(X_t)_{1 \leq t \leq n}$  mais aussi  $(Y_t)_{1 \leq t \leq n}$  où

$$Y_t = (Y_t^{(1)}, \dots, Y_t^{(d)}), \text{ où } Y^{(1)}, \dots, Y^{(d)} \text{ variables exogènes}$$

⇒ Méthodes de l'**apprentissage statistique** :

$$X_t = h(t, s(t), Y_t^{(1)}, \dots, Y_t^{(d)}, u_t) \quad \text{pour tout } t$$

$$\text{où } s(t) = s_1 h_1(t) + \dots + s_{T-1} h_{T-1}(t)$$

⇒  $\hat{h}$  par forêts aléatoires, XGBoost, Deep learning...

## Extension en cas de variables exogènes observées

Tendance et saisonnalité estimées jusqu'ici comme fonctions du temps

⇒ Extension en cas de **variables exogènes** observées aux mêmes instants

⇒ On suppose observés  $(X_t)_{1 \leq t \leq n}$  mais aussi  $(Y_t)_{1 \leq t \leq n}$  où

$$Y_t = (Y_t^{(1)}, \dots, Y_t^{(d)}), \text{ où } Y^{(1)}, \dots, Y^{(d)} \text{ variables exogènes}$$

⇒ Méthodes de l'**apprentissage statistique** :

$$X_t = h(t, s(t), Y_t^{(1)}, \dots, Y_t^{(d)}, u_t) \quad \text{pour tout } t$$

$$\text{où } s(t) = s_1 h_1(t) + \dots + s_{T-1} h_{T-1}(t)$$

⇒  $\hat{h}$  par forêts aléatoires, XGBoost, Deep learning...

# Conclusions estimation de la tendance et de la saisonnalité

- Prédire dans le cas linéaire : estimation semi-paramétrique de  $a$  et  $s$  ;
- Sinon :
  - ▶ 1ère démarche par moyenne mobile pour estimer  $a$  et  $s$
  - ▶ Amélioration par méthodes non paramétriques sur série désaisonnalisée

Pour prédire :

- ▶ Utiliser un modèle pour la partie bruit estimée par  $(\hat{u}_t)$  ;
- ▶ Garder  $\hat{s}$  pour prédire ;
- ▶ Prédire la tendance par filtrage Holt-Winter ;
- ▶ Prédire la partie bruit grâce au modèle obtenu pour  $(\hat{u}_t)$ .
- ▶ Sommer le tout :  $\hat{X}_{n+h} = \hat{a}(n+h) + \hat{s}(n+h) + \hat{u}_{n+h}$

# Conclusions estimation de la tendance et de la saisonnalité

- Prédire dans le cas linéaire : estimation semi-paramétrique de  $a$  et  $s$  ;
- Sinon :
  - ▶ 1ère démarche par moyenne mobile pour estimer  $a$  et  $s$
  - ▶ Amélioration par méthodes non paramétriques sur série désaisonnalisée

Pour prédire :

- ▶ Utiliser un modèle pour la partie bruit estimée par  $(\hat{u}_t)$  ;
- ▶ Garder  $\hat{s}$  pour prédire ;
- ▶ Prédire la tendance par filtrage Holt-Winter ;
- ▶ Prédire la partie bruit grâce au modèle obtenu pour  $(\hat{u}_t)$ .
- ▶ Sommer le tout :  $\hat{X}_{n+h} = \hat{a}(n+h) + \hat{s}(n+h) + \hat{u}_{n+h}$

# Conclusions estimation de la tendance et de la saisonnalité

- Prédire dans le cas linéaire : estimation semi-paramétrique de  $a$  et  $s$  ;
- Sinon :
  - ▶ 1ère démarche par moyenne mobile pour estimer  $a$  et  $s$
  - ▶ Amélioration par méthodes non paramétriques sur série désaisonnalisée

Pour prédire :

- ▶ Utiliser un modèle pour la partie bruit estimée par  $(\hat{u}_t)$  ;
- ▶ Garder  $\hat{s}$  pour prédire ;
- ▶ Prédire la tendance par filtrage Holt-Winter ;
- ▶ Prédire la partie bruit grâce au modèle obtenu pour  $(\hat{u}_t)$ .
- ▶ Sommer le tout :  $\hat{X}_{n+h} = \hat{a}(n+h) + \hat{s}(n+h) + \hat{u}_{n+h}$

# Conclusions estimation de la tendance et de la saisonnalité

- Prédire dans le cas linéaire : estimation semi-paramétrique de  $a$  et  $s$  ;
- Sinon :
  - ▶ 1ère démarche par moyenne mobile pour estimer  $a$  et  $s$
  - ▶ Amélioration par méthodes non paramétriques sur série désaisonnalisée

Pour prédire :

- ▶ Utiliser un modèle pour la partie bruit estimée par  $(\hat{u}_t)$  ;
- ▶ Garder  $\hat{s}$  pour prédire ;
- ▶ Prédire la tendance par filtrage Holt-Winter ;
- ▶ Prédire la partie bruit grâce au modèle obtenu pour  $(\hat{u}_t)$ .
- ▶ Sommer le tout :  $\hat{X}_{n+h} = \hat{a}(n+h) + \hat{s}(n+h) + \hat{u}_{n+h}$

# Conclusions estimation de la tendance et de la saisonnalité

- Prédire dans le cas linéaire : estimation semi-paramétrique de  $a$  et  $s$  ;
- Sinon :
  - ▶ 1ère démarche par moyenne mobile pour estimer  $a$  et  $s$
  - ▶ Amélioration par méthodes non paramétriques sur série désaisonnalisée

Pour prédire :

- ▶ Utiliser un modèle pour la partie bruit estimée par  $(\hat{u}_t)$  ;
- ▶ Garder  $\hat{s}$  pour prédire ;
- ▶ Prédire la tendance par filtrage Holt-Winter ;
- ▶ Prédire la partie bruit grâce au modèle obtenu pour  $(\hat{u}_t)$ .
- ▶ Sommer le tout :  $\hat{X}_{n+h} = \hat{a}(n+h) + \hat{s}(n+h) + \hat{u}_{n+h}$

# Conclusions estimation de la tendance et de la saisonnalité

- Prédire dans le cas linéaire : estimation semi-paramétrique de  $a$  et  $s$  ;
- Sinon :
  - ▶ 1ère démarche par moyenne mobile pour estimer  $a$  et  $s$
  - ▶ Amélioration par méthodes non paramétriques sur série désaisonnalisée

Pour prédire :

- ▶ Utiliser un modèle pour la partie bruit estimée par  $(\hat{u}_t)$  ;
- ▶ Garder  $\hat{s}$  pour prédire ;
- ▶ Prédire la tendance par filtrage Holt-Winter ;
- ▶ Prédire la partie bruit grâce au modèle obtenu pour  $(\hat{u}_t)$ .
- ▶ Sommer le tout :  $\hat{X}_{n+h} = \hat{a}(n+h) + \hat{s}(n+h) + \hat{u}_{n+h}$

# Conclusions estimation de la tendance et de la saisonnalité

- Prédire dans le cas linéaire : estimation semi-paramétrique de  $a$  et  $s$  ;
- Sinon :
  - ▶ 1ère démarche par moyenne mobile pour estimer  $a$  et  $s$
  - ▶ Amélioration par méthodes non paramétriques sur série désaisonnalisée

Pour prédire :

- ▶ Utiliser un modèle pour la partie bruit estimée par  $(\hat{u}_t)$  ;
- ▶ Garder  $\hat{s}$  pour prédire ;
- ▶ Prédire la tendance par filtrage Holt-Winter ;
- ▶ Prédire la partie bruit grâce au modèle obtenu pour  $(\hat{u}_t)$ .
- ▶ Sommer le tout :  $\hat{X}_{n+h} = \hat{a}(n+h) + \hat{s}(n+h) + \hat{u}_{n+h}$

# Conclusions estimation de la tendance et de la saisonnalité

- Prédire dans le cas linéaire : estimation semi-paramétrique de  $a$  et  $s$  ;
- Sinon :
  - ▶ 1ère démarche par moyenne mobile pour estimer  $a$  et  $s$
  - ▶ Amélioration par méthodes non paramétriques sur série désaisonnalisée

Pour prédire :

- ▶ Utiliser un modèle pour la partie bruit estimée par  $(\hat{u}_t)$  ;
- ▶ Garder  $\hat{s}$  pour prédire ;
- ▶ Prédire la tendance par filtrage Holt-Winter ;
- ▶ Prédire la partie bruit grâce au modèle obtenu pour  $(\hat{u}_t)$ .
- ▶ Sommer le tout :  $\hat{X}_{n+h} = \hat{a}(n+h) + \hat{s}(n+h) + \hat{u}_{n+h}$

# Conclusions estimation de la tendance et de la saisonnalité

- Prédire dans le cas linéaire : estimation semi-paramétrique de  $a$  et  $s$  ;
- Sinon :
  - ▶ 1ère démarche par moyenne mobile pour estimer  $a$  et  $s$
  - ▶ Amélioration par méthodes non paramétriques sur série désaisonnalisée

Pour prédire :

- ▶ Utiliser un modèle pour la partie bruit estimée par  $(\hat{u}_t)$  ;
- ▶ Garder  $\hat{s}$  pour prédire ;
- ▶ Prédire la tendance par filtrage Holt-Winter ;
- ▶ Prédire la partie bruit grâce au modèle obtenu pour  $(\hat{u}_t)$ .
- ▶ Sommer le tout :  $\hat{X}_{n+h} = \hat{a}(n+h) + \hat{s}(n+h) + \hat{u}_{n+h}$

# Position du problème

Soit  $(X_1, \dots, X_n)$  une trajectoire observée de  $(X_t)_{t \in \mathbb{Z}}$ .

- Si  $(X_t)$  ne possède ni tendance et saisonnalité

⇒ On veut directement un modèle pour  $(X_t)$

- Si  $(X_t)$  possède une tendance et/ou saisonnalité additive ;

⇒ On obtient  $\hat{a}(\cdot)$  et  $\hat{s}(\cdot)$

⇒ On cherche un modèle pour  $(\hat{u}(t))_{t \in \mathbb{Z}}$  où

$$\hat{u}(t) = X_t - \hat{a}(t) - \hat{s}(t)$$

- Un modèle pour quoi faire ?

⇒ Pour améliorer la prédiction, évaluer un risque...

## Position du problème

Soit  $(X_1, \dots, X_n)$  une trajectoire observée de  $(X_t)_{t \in \mathbb{Z}}$ .

- Si  $(X_t)$  ne possède ni tendance et saisonnalité

⇒ On veut directement un modèle pour  $(X_t)$

- Si  $(X_t)$  possède une tendance et/ou saisonnalité additive ;

⇒ On obtient  $\hat{a}(\cdot)$  et  $\hat{s}(\cdot)$

⇒ On cherche un modèle pour  $(\hat{u}(t))_{t \in \mathbb{Z}}$  où

$$\hat{u}(t) = X_t - \hat{a}(t) - \hat{s}(t)$$

- Un modèle pour quoi faire ?

⇒ Pour améliorer la prédiction, évaluer un risque...

## Position du problème

Soit  $(X_1, \dots, X_n)$  une trajectoire observée de  $(X_t)_{t \in \mathbb{Z}}$ .

- Si  $(X_t)$  ne possède ni tendance et saisonnalité

⇒ On veut directement un modèle pour  $(X_t)$

- Si  $(X_t)$  possède une tendance et/ou saisonnalité additive ;

⇒ On obtient  $\hat{a}(\cdot)$  et  $\hat{s}(\cdot)$

⇒ On cherche un modèle pour  $(\hat{u}(t))_{t \in \mathbb{Z}}$  où

$$\hat{u}(t) = X_t - \hat{a}(t) - \hat{s}(t)$$

- Un modèle pour quoi faire ?

⇒ Pour améliorer la prédiction, évaluer un risque...

## Position du problème

Soit  $(X_1, \dots, X_n)$  une trajectoire observée de  $(X_t)_{t \in \mathbb{Z}}$ .

- Si  $(X_t)$  ne possède ni tendance et saisonnalité

⇒ On veut directement un modèle pour  $(X_t)$

- Si  $(X_t)$  possède une tendance et/ou saisonnalité additive ;

⇒ On obtient  $\hat{a}(\cdot)$  et  $\hat{s}(\cdot)$

⇒ On cherche un modèle pour  $(\hat{u}(t))_{t \in \mathbb{Z}}$  où

$$\hat{u}(t) = X_t - \hat{a}(t) - \hat{s}(t)$$

- Un modèle pour quoi faire ?

⇒ Pour améliorer la prédiction, évaluer un risque...

## Position du problème

Soit  $(X_1, \dots, X_n)$  une trajectoire observée de  $(X_t)_{t \in \mathbb{Z}}$ .

- Si  $(X_t)$  ne possède ni tendance et saisonnalité

⇒ On veut directement un modèle pour  $(X_t)$

- Si  $(X_t)$  possède une tendance et/ou saisonnalité additive ;

⇒ On obtient  $\hat{a}(\cdot)$  et  $\hat{s}(\cdot)$

⇒ On cherche un modèle pour  $(\hat{u}(t))_{t \in \mathbb{Z}}$  où

$$\hat{u}(t) = X_t - \hat{a}(t) - \hat{s}(t)$$

- Un modèle pour quoi faire ?

⇒ Pour améliorer la prédiction, évaluer un risque...

## Position du problème

Soit  $(X_1, \dots, X_n)$  une trajectoire observée de  $(X_t)_{t \in \mathbb{Z}}$ .

- Si  $(X_t)$  ne possède ni tendance et saisonnalité

⇒ On veut directement un modèle pour  $(X_t)$

- Si  $(X_t)$  possède une tendance et/ou saisonnalité additive ;

⇒ On obtient  $\hat{a}(\cdot)$  et  $\hat{s}(\cdot)$

⇒ On cherche un modèle pour  $(\hat{u}(t))_{t \in \mathbb{Z}}$  où

$$\hat{u}(t) = X_t - \hat{a}(t) - \hat{s}(t)$$

- Un modèle pour quoi faire ?

⇒ Pour améliorer la prédiction, évaluer un risque...

## Position du problème

Soit  $(X_1, \dots, X_n)$  une trajectoire observée de  $(X_t)_{t \in \mathbb{Z}}$ .

- Si  $(X_t)$  ne possède ni tendance et saisonnalité

⇒ On veut directement un modèle pour  $(X_t)$

- Si  $(X_t)$  possède une tendance et/ou saisonnalité additive ;

⇒ On obtient  $\hat{a}(\cdot)$  et  $\hat{s}(\cdot)$

⇒ On cherche un modèle pour  $(\hat{u}(t))_{t \in \mathbb{Z}}$  où

$$\hat{u}(t) = X_t - \hat{a}(t) - \hat{s}(t)$$

- Un modèle pour quoi faire ?

⇒ Pour améliorer la prédiction, évaluer un risque...

## Position du problème

Soit  $(X_1, \dots, X_n)$  une trajectoire observée de  $(X_t)_{t \in \mathbb{Z}}$ .

- Si  $(X_t)$  ne possède ni tendance et saisonnalité

$\implies$  On veut directement un modèle pour  $(X_t)$

- Si  $(X_t)$  possède une tendance et/ou saisonnalité additive ;

$\implies$  On obtient  $\hat{a}(\cdot)$  et  $\hat{s}(\cdot)$

$\implies$  On cherche un modèle pour  $(\hat{u}(t))_{t \in \mathbb{Z}}$  où

$$\hat{u}(t) = X_t - \hat{a}(t) - \hat{s}(t)$$

- Un modèle pour quoi faire ?

$\implies$  Pour améliorer la prédiction, évaluer un risque...

# Plan du cours

- 1 Définitions et premières propriétés
  - Séries chronologiques
  - Stationnarité
  - Première analyse statistique de la dépendance
- 2 Tendances et saisonnalités
  - Définition et propriétés
  - Estimation semi-paramétrique de la tendance et de la saisonnalité
  - Estimation non-paramétrique de la tendance et de la saisonnalité
- 3 Exemples de modèles de séries stationnaires
  - Processus ARMA
  - Processus GARCH
  - Autres exemples de séries temporelles
- 4 Estimation, sélection de modèle, test et prédiction
  - Estimation semi-paramétrique
  - Sélection de modèles et test d'adéquation
  - Prédiction

# Processus ARMA

Définition (Yule, 1921 et Slutsky, 1927)

Un processus ARMA( $p, q$ ) où  $(p, q) \in \mathbb{N}^2$  est une série  $(X_t)_{t \in \mathbb{Z}}$  stationnaire telle que pour  $t \in \mathbb{Z}$  :

$$X_t + a_1 X_{t-1} + \dots + a_p X_{t-p} = \varepsilon_t + b_1 \varepsilon_{t-1} + \dots + b_q \varepsilon_{t-q}$$

où :

- $(\varepsilon_t)$  bruit blanc de variance  $\sigma_\varepsilon^2$  ;
- $(a_1, \dots, a_p, b_1, \dots, b_q) \in \mathbb{R}^{p+q}$  avec  $a_p \neq 0$  et  $b_q \neq 0$ .

Cas particuliers :

- Si  $q = 0$ , un ARMA( $p, q$ ) est un AR( $p$ ) (Auto-Regressive) ;
- Si  $p = 0$ , un ARMA( $p, q$ ) est un MA( $q$ ) (Moving Average) ;

# Processus ARMA

Définition (Yule, 1921 et Slutsky, 1927)

Un processus ARMA( $p, q$ ) où  $(p, q) \in \mathbb{N}^2$  est une série  $(X_t)_{t \in \mathbb{Z}}$  stationnaire telle que pour  $t \in \mathbb{Z}$  :

$$X_t + a_1 X_{t-1} + \dots + a_p X_{t-p} = \varepsilon_t + b_1 \varepsilon_{t-1} + \dots + b_q \varepsilon_{t-q}$$

où :

- $(\varepsilon_t)$  bruit blanc de variance  $\sigma_\varepsilon^2$  ;
- $(a_1, \dots, a_p, b_1, \dots, b_q) \in \mathbb{R}^{p+q}$  avec  $a_p \neq 0$  et  $b_q \neq 0$ .

Cas particuliers :

- Si  $q = 0$ , un ARMA( $p, q$ ) est un AR( $p$ ) (Auto-Regressive) ;
- Si  $p = 0$ , un ARMA( $p, q$ ) est un MA( $q$ ) (Moving Average) ;

# Processus ARMA

Définition (Yule, 1921 et Slutsky, 1927)

Un processus ARMA( $p, q$ ) où  $(p, q) \in \mathbb{N}^2$  est une série  $(X_t)_{t \in \mathbb{Z}}$  stationnaire telle que pour  $t \in \mathbb{Z}$  :

$$X_t + a_1 X_{t-1} + \dots + a_p X_{t-p} = \varepsilon_t + b_1 \varepsilon_{t-1} + \dots + b_q \varepsilon_{t-q}$$

où :

- $(\varepsilon_t)$  bruit blanc de variance  $\sigma_\varepsilon^2$  ;
- $(a_1, \dots, a_p, b_1, \dots, b_q) \in \mathbb{R}^{p+q}$  avec  $a_p \neq 0$  et  $b_q \neq 0$ .

Cas particuliers :

- Si  $q = 0$ , un ARMA( $p, q$ ) est un AR( $p$ ) (Auto-Regressive) ;
- Si  $p = 0$ , un ARMA( $p, q$ ) est un MA( $q$ ) (Moving Average) ;

# Processus ARMA

Définition (Yule, 1921 et Slutsky, 1927)

Un processus ARMA( $p, q$ ) où  $(p, q) \in \mathbb{N}^2$  est une série  $(X_t)_{t \in \mathbb{Z}}$  stationnaire telle que pour  $t \in \mathbb{Z}$  :

$$X_t + a_1 X_{t-1} + \dots + a_p X_{t-p} = \varepsilon_t + b_1 \varepsilon_{t-1} + \dots + b_q \varepsilon_{t-q}$$

où :

- $(\varepsilon_t)$  bruit blanc de variance  $\sigma_\varepsilon^2$  ;
- $(a_1, \dots, a_p, b_1, \dots, b_q) \in \mathbb{R}^{p+q}$  avec  $a_p \neq 0$  et  $b_q \neq 0$ .

Cas particuliers :

- Si  $q = 0$ , un ARMA( $p, q$ ) est un AR( $p$ ) (Auto-Regressive) ;
- Si  $p = 0$ , un ARMA( $p, q$ ) est un MA( $q$ ) (Moving Average) ;

# Processus ARMA

Définition (Yule, 1921 et Slutsky, 1927)

Un processus ARMA( $p, q$ ) où  $(p, q) \in \mathbb{N}^2$  est une série  $(X_t)_{t \in \mathbb{Z}}$  stationnaire telle que pour  $t \in \mathbb{Z}$  :

$$X_t + a_1 X_{t-1} + \dots + a_p X_{t-p} = \varepsilon_t + b_1 \varepsilon_{t-1} + \dots + b_q \varepsilon_{t-q}$$

où :

- $(\varepsilon_t)$  bruit blanc de variance  $\sigma_\varepsilon^2$  ;
- $(a_1, \dots, a_p, b_1, \dots, b_q) \in \mathbb{R}^{p+q}$  avec  $a_p \neq 0$  et  $b_q \neq 0$ .

Cas particuliers :

- Si  $q = 0$ , un ARMA( $p, q$ ) est un AR( $p$ ) (Auto-Regressive) ;
- Si  $p = 0$ , un ARMA( $p, q$ ) est un MA( $q$ ) (Moving Average) ;

# Processus ARMA

Définition (Yule, 1921 et Slutsky, 1927)

Un processus ARMA( $p, q$ ) où  $(p, q) \in \mathbb{N}^2$  est une série  $(X_t)_{t \in \mathbb{Z}}$  stationnaire telle que pour  $t \in \mathbb{Z}$  :

$$X_t + a_1 X_{t-1} + \dots + a_p X_{t-p} = \varepsilon_t + b_1 \varepsilon_{t-1} + \dots + b_q \varepsilon_{t-q}$$

où :

- $(\varepsilon_t)$  bruit blanc de variance  $\sigma_\varepsilon^2$  ;
- $(a_1, \dots, a_p, b_1, \dots, b_q) \in \mathbb{R}^{p+q}$  avec  $a_p \neq 0$  et  $b_q \neq 0$ .

Cas particuliers :

- Si  $q = 0$ , un ARMA( $p, q$ ) est un AR( $p$ ) (Auto-Regressive) ;
- Si  $p = 0$ , un ARMA( $p, q$ ) est un MA( $q$ ) (Moving Average) ;

# Processus ARMA

Définition (Yule, 1921 et Slutsky, 1927)

Un processus ARMA( $p, q$ ) où  $(p, q) \in \mathbb{N}^2$  est une série  $(X_t)_{t \in \mathbb{Z}}$  stationnaire telle que pour  $t \in \mathbb{Z}$  :

$$X_t + a_1 X_{t-1} + \dots + a_p X_{t-p} = \varepsilon_t + b_1 \varepsilon_{t-1} + \dots + b_q \varepsilon_{t-q}$$

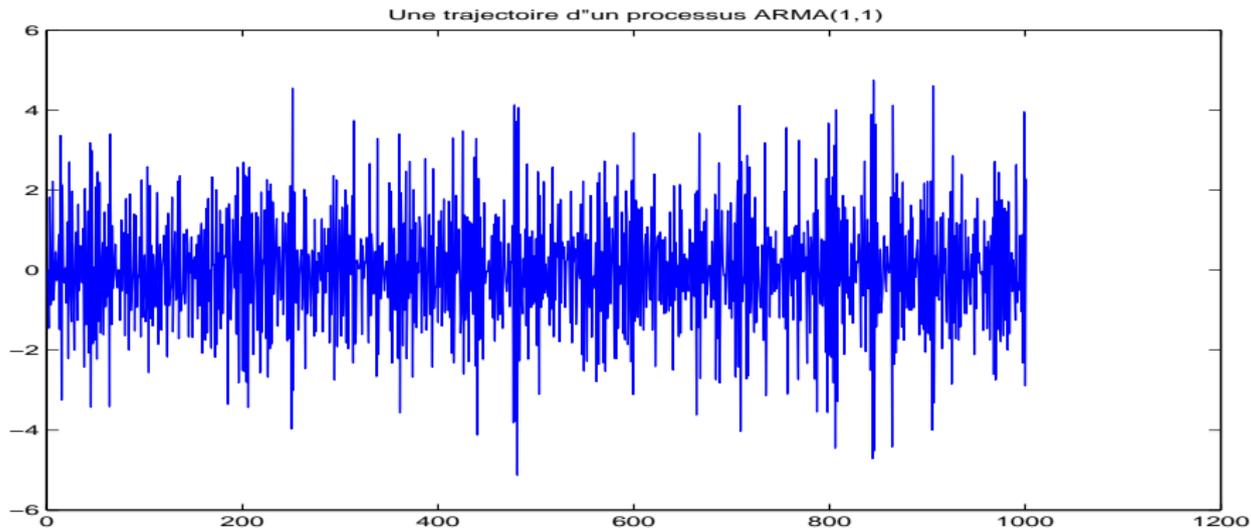
où :

- $(\varepsilon_t)$  bruit blanc de variance  $\sigma_\varepsilon^2$  ;
- $(a_1, \dots, a_p, b_1, \dots, b_q) \in \mathbb{R}^{p+q}$  avec  $a_p \neq 0$  et  $b_q \neq 0$ .

Cas particuliers :

- Si  $q = 0$ , un ARMA( $p, q$ ) est un AR( $p$ ) (Auto-Regressive) ;
- Si  $p = 0$ , un ARMA( $p, q$ ) est un MA( $q$ ) (Moving Average) ;

# Une trajectoire de processus ARMA(1,1)



# Processus ARMA : deux exemples

- ① Soit le processus AR(1) avec  $X_t = \alpha X_{t-1} + \varepsilon_t$  et  $|\alpha| < 1$ . Alors :

$$\begin{aligned} X_t &= \alpha(\alpha X_{t-2} + \varepsilon_{t-1}) + \varepsilon_t \\ &= \alpha^2 X_{t-2} + \alpha \varepsilon_{t-1} + \varepsilon_t \\ &= \alpha^2(\alpha X_{t-3} + \varepsilon_{t-2}) + \alpha \varepsilon_{t-1} + \varepsilon_t \\ &= \alpha^3 X_{t-3} + \alpha^2 \varepsilon_{t-2} + \alpha \varepsilon_{t-1} + \varepsilon_t \\ &\vdots \quad \quad \quad \vdots \quad \quad \quad \vdots \quad \quad \quad \vdots \\ &= \sum_{k=0}^{\infty} \alpha^k \varepsilon_{t-k} \quad (\text{existe car } |\alpha| < 1) \end{aligned}$$

- ② Soit le processus AR(1) avec  $X_t = \alpha X_{t-1} + \varepsilon_t$  et  $|\alpha| > 1$ . Alors :

$$\begin{aligned} X_t &= \alpha^{-1} X_{t+1} - \alpha^{-1} \varepsilon_{t+1} \\ &= \alpha^{-1}(\alpha^{-1} X_{t+2} - \alpha^{-1} \varepsilon_{t+2}) - \alpha^{-1} \varepsilon_{t+1} \\ &= \alpha^{-2} X_{t+2} - \alpha^{-2} \varepsilon_{t+2} - \alpha^{-1} \varepsilon_{t+1} \\ &\vdots \quad \quad \quad \vdots \quad \quad \quad \vdots \quad \quad \quad \vdots \\ &= - \sum_{k=1}^{\infty} (\alpha^{-1})^k \varepsilon_{t+k} \quad (\text{existe car } |\alpha^{-1}| < 1) \end{aligned}$$

## Processus ARMA : deux exemples

- ❶ Soit le processus AR(1) avec  $X_t = \alpha X_{t-1} + \varepsilon_t$  et  $|\alpha| < 1$ . Alors :

$$\begin{aligned}X_t &= \alpha(\alpha X_{t-2} + \varepsilon_{t-1}) + \varepsilon_t \\&= \alpha^2 X_{t-2} + \alpha \varepsilon_{t-1} + \varepsilon_t \\&= \alpha^2 (\alpha X_{t-3} + \varepsilon_{t-2}) + \alpha \varepsilon_{t-1} + \varepsilon_t \\&= \alpha^3 X_{t-3} + \alpha^2 \varepsilon_{t-2} + \alpha \varepsilon_{t-1} + \varepsilon_t \\&\vdots \quad \quad \quad \vdots \quad \quad \quad \vdots \quad \quad \quad \vdots \\&= \sum_{k=0}^{\infty} \alpha^k \varepsilon_{t-k} \quad (\text{existe car } |\alpha| < 1)\end{aligned}$$

- ❷ Soit le processus AR(1) avec  $X_t = \alpha X_{t-1} + \varepsilon_t$  et  $|\alpha| > 1$ . Alors :

$$\begin{aligned}X_t &= \alpha^{-1} X_{t+1} - \alpha^{-1} \varepsilon_{t+1} \\&= \alpha^{-1} (\alpha^{-1} X_{t+2} - \alpha^{-1} \varepsilon_{t+2}) - \alpha^{-1} \varepsilon_{t+1} \\&= \alpha^{-2} X_{t+2} - \alpha^{-2} \varepsilon_{t+2} - \alpha^{-1} \varepsilon_{t+1} \\&\vdots \quad \quad \quad \vdots \quad \quad \quad \vdots \quad \quad \quad \vdots \\&= - \sum_{k=1}^{\infty} (\alpha^{-1})^k \varepsilon_{t+k} \quad (\text{existe car } |\alpha^{-1}| < 1)\end{aligned}$$

# Processus ARMA : deux exemples

- ① Soit le processus AR(1) avec  $X_t = \alpha X_{t-1} + \varepsilon_t$  et  $|\alpha| < 1$ . Alors :

$$\begin{aligned} X_t &= \alpha(\alpha X_{t-2} + \varepsilon_{t-1}) + \varepsilon_t \\ &= \alpha^2 X_{t-2} + \alpha \varepsilon_{t-1} + \varepsilon_t \\ &= \alpha^2 (\alpha X_{t-3} + \varepsilon_{t-2}) + \alpha \varepsilon_{t-1} + \varepsilon_t \\ &= \alpha^3 X_{t-3} + \alpha^2 \varepsilon_{t-2} + \alpha \varepsilon_{t-1} + \varepsilon_t \\ &\vdots \quad \quad \quad \vdots \quad \quad \quad \vdots \quad \quad \quad \vdots \\ &= \sum_{k=0}^{\infty} \alpha^k \varepsilon_{t-k} \quad (\text{existe car } |\alpha| < 1) \end{aligned}$$

- ② Soit le processus AR(1) avec  $X_t = \alpha X_{t-1} + \varepsilon_t$  et  $|\alpha| > 1$ . Alors :

$$\begin{aligned} X_t &= \alpha^{-1} X_{t+1} - \alpha^{-1} \varepsilon_{t+1} \\ &= \alpha^{-1} (\alpha^{-1} X_{t+2} - \alpha^{-1} \varepsilon_{t+2}) - \alpha^{-1} \varepsilon_{t+1} \\ &= \alpha^{-2} X_{t+2} - \alpha^{-2} \varepsilon_{t+2} - \alpha^{-1} \varepsilon_{t+1} \\ &\vdots \quad \quad \quad \vdots \quad \quad \quad \vdots \quad \quad \quad \vdots \\ &= - \sum_{k=1}^{\infty} (\alpha^{-1})^k \varepsilon_{t+k} \quad (\text{existe car } |\alpha^{-1}| < 1) \end{aligned}$$

## Processus ARMA : deux exemples

- ① Soit le processus AR(1) avec  $X_t = \alpha X_{t-1} + \varepsilon_t$  et  $|\alpha| < 1$ . Alors :

$$\begin{aligned} X_t &= \alpha(\alpha X_{t-2} + \varepsilon_{t-1}) + \varepsilon_t \\ &= \alpha^2 X_{t-2} + \alpha \varepsilon_{t-1} + \varepsilon_t \\ &= \alpha^2 (\alpha X_{t-3} + \varepsilon_{t-2}) + \alpha \varepsilon_{t-1} + \varepsilon_t \\ &= \alpha^3 X_{t-3} + \alpha^2 \varepsilon_{t-2} + \alpha \varepsilon_{t-1} + \varepsilon_t \\ &\vdots \quad \quad \quad \vdots \quad \quad \quad \vdots \quad \quad \quad \vdots \\ &= \sum_{k=0}^{\infty} \alpha^k \varepsilon_{t-k} \quad (\text{existe car } |\alpha| < 1) \end{aligned}$$

- ② Soit le processus AR(1) avec  $X_t = \alpha X_{t-1} + \varepsilon_t$  et  $|\alpha| > 1$ . Alors :

$$\begin{aligned} X_t &= \alpha^{-1} X_{t+1} - \alpha^{-1} \varepsilon_{t+1} \\ &= \alpha^{-1} (\alpha^{-1} X_{t+2} - \alpha^{-1} \varepsilon_{t+2}) - \alpha^{-1} \varepsilon_{t+1} \\ &= \alpha^{-2} X_{t+2} - \alpha^{-2} \varepsilon_{t+2} - \alpha^{-1} \varepsilon_{t+1} \\ &\vdots \quad \quad \quad \vdots \quad \quad \quad \vdots \quad \quad \quad \vdots \\ &= - \sum_{k=1}^{\infty} (\alpha^{-1})^k \varepsilon_{t+k} \quad (\text{existe car } |\alpha^{-1}| < 1) \end{aligned}$$

## Processus ARMA : deux exemples

- ❶ Soit le processus AR(1) avec  $X_t = \alpha X_{t-1} + \varepsilon_t$  et  $|\alpha| < 1$ . Alors :

$$\begin{aligned} X_t &= \alpha(\alpha X_{t-2} + \varepsilon_{t-1}) + \varepsilon_t \\ &= \alpha^2 X_{t-2} + \alpha \varepsilon_{t-1} + \varepsilon_t \\ &= \alpha^2(\alpha X_{t-3} + \varepsilon_{t-2}) + \alpha \varepsilon_{t-1} + \varepsilon_t \\ &= \alpha^3 X_{t-3} + \alpha^2 \varepsilon_{t-2} + \alpha \varepsilon_{t-1} + \varepsilon_t \\ &\vdots \quad \quad \quad \vdots \quad \quad \quad \vdots \quad \quad \quad \vdots \\ &= \sum_{k=0}^{\infty} \alpha^k \varepsilon_{t-k} \quad (\text{existe car } |\alpha| < 1) \end{aligned}$$

- ❷ Soit le processus AR(1) avec  $X_t = \alpha X_{t-1} + \varepsilon_t$  et  $|\alpha| > 1$ . Alors :

$$\begin{aligned} X_t &= \alpha^{-1} X_{t+1} - \alpha^{-1} \varepsilon_{t+1} \\ &= \alpha^{-1}(\alpha^{-1} X_{t+2} - \alpha^{-1} \varepsilon_{t+2}) - \alpha^{-1} \varepsilon_{t+1} \\ &= \alpha^{-2} X_{t+2} - \alpha^{-2} \varepsilon_{t+2} - \alpha^{-1} \varepsilon_{t+1} \\ &\vdots \quad \quad \quad \vdots \quad \quad \quad \vdots \quad \quad \quad \vdots \\ &= - \sum_{k=1}^{\infty} (\alpha^{-1})^k \varepsilon_{t+k} \quad (\text{existe car } |\alpha^{-1}| < 1) \end{aligned}$$



## Processus ARMA : deux exemples

- ❶ Soit le processus AR(1) avec  $X_t = \alpha X_{t-1} + \varepsilon_t$  et  $|\alpha| < 1$ . Alors :

$$\begin{aligned}X_t &= \alpha(\alpha X_{t-2} + \varepsilon_{t-1}) + \varepsilon_t \\&= \alpha^2 X_{t-2} + \alpha \varepsilon_{t-1} + \varepsilon_t \\&= \alpha^2(\alpha X_{t-3} + \varepsilon_{t-2}) + \alpha \varepsilon_{t-1} + \varepsilon_t \\&= \alpha^3 X_{t-3} + \alpha^2 \varepsilon_{t-2} + \alpha \varepsilon_{t-1} + \varepsilon_t \\&\vdots \quad \quad \quad \vdots \quad \quad \quad \vdots \quad \quad \quad \vdots \\&= \sum_{k=0}^{\infty} \alpha^k \varepsilon_{t-k} \quad (\text{existe car } |\alpha| < 1)\end{aligned}$$

- ❷ Soit le processus AR(1) avec  $X_t = \alpha X_{t-1} + \varepsilon_t$  et  $|\alpha| > 1$ . Alors :

$$\begin{aligned}X_t &= \alpha^{-1} X_{t+1} - \alpha^{-1} \varepsilon_{t+1} \\&= \alpha^{-1}(\alpha^{-1} X_{t+2} - \alpha^{-1} \varepsilon_{t+2}) - \alpha^{-1} \varepsilon_{t+1} \\&= \alpha^{-2} X_{t+2} - \alpha^{-2} \varepsilon_{t+2} - \alpha^{-1} \varepsilon_{t+1} \\&\vdots \quad \quad \quad \vdots \quad \quad \quad \vdots \quad \quad \quad \vdots \\&= -\sum_{k=1}^{\infty} (\alpha^{-1})^k \varepsilon_{t+k} \quad (\text{existe car } |\alpha^{-1}| < 1)\end{aligned}$$

## Processus ARMA : deux exemples

- ❶ Soit le processus AR(1) avec  $X_t = \alpha X_{t-1} + \varepsilon_t$  et  $|\alpha| < 1$ . Alors :

$$\begin{aligned} X_t &= \alpha(\alpha X_{t-2} + \varepsilon_{t-1}) + \varepsilon_t \\ &= \alpha^2 X_{t-2} + \alpha \varepsilon_{t-1} + \varepsilon_t \\ &= \alpha^2(\alpha X_{t-3} + \varepsilon_{t-2}) + \alpha \varepsilon_{t-1} + \varepsilon_t \\ &= \alpha^3 X_{t-3} + \alpha^2 \varepsilon_{t-2} + \alpha \varepsilon_{t-1} + \varepsilon_t \\ &\vdots \quad \quad \quad \vdots \quad \quad \quad \vdots \quad \quad \quad \vdots \\ &= \sum_{k=0}^{\infty} \alpha^k \varepsilon_{t-k} \quad (\text{existe car } |\alpha| < 1) \end{aligned}$$

- ❷ Soit le processus AR(1) avec  $X_t = \alpha X_{t-1} + \varepsilon_t$  et  $|\alpha| > 1$ . Alors :

$$\begin{aligned} X_t &= \alpha^{-1} X_{t+1} - \alpha^{-1} \varepsilon_{t+1} \\ &= \alpha^{-1}(\alpha^{-1} X_{t+2} - \alpha^{-1} \varepsilon_{t+2}) - \alpha^{-1} \varepsilon_{t+1} \\ &= \alpha^{-2} X_{t+2} - \alpha^{-2} \varepsilon_{t+2} - \alpha^{-1} \varepsilon_{t+1} \\ &\vdots \quad \quad \quad \vdots \quad \quad \quad \vdots \quad \quad \quad \vdots \\ &= - \sum_{k=1}^{\infty} (\alpha^{-1})^k \varepsilon_{t+k} \quad (\text{existe car } |\alpha^{-1}| < 1) \end{aligned}$$

## Processus ARMA : deux exemples

- ❶ Soit le processus AR(1) avec  $X_t = \alpha X_{t-1} + \varepsilon_t$  et  $|\alpha| < 1$ . Alors :

$$\begin{aligned}X_t &= \alpha(\alpha X_{t-2} + \varepsilon_{t-1}) + \varepsilon_t \\&= \alpha^2 X_{t-2} + \alpha \varepsilon_{t-1} + \varepsilon_t \\&= \alpha^2(\alpha X_{t-3} + \varepsilon_{t-2}) + \alpha \varepsilon_{t-1} + \varepsilon_t \\&= \alpha^3 X_{t-3} + \alpha^2 \varepsilon_{t-2} + \alpha \varepsilon_{t-1} + \varepsilon_t \\&\vdots \quad \quad \quad \vdots \quad \quad \quad \vdots \quad \quad \quad \vdots \\&= \sum_{k=0}^{\infty} \alpha^k \varepsilon_{t-k} \quad (\text{existe car } |\alpha| < 1)\end{aligned}$$

- ❷ Soit le processus AR(1) avec  $X_t = \alpha X_{t-1} + \varepsilon_t$  et  $|\alpha| > 1$ . Alors :

$$\begin{aligned}X_t &= \alpha^{-1} X_{t+1} - \alpha^{-1} \varepsilon_{t+1} \\&= \alpha^{-1}(\alpha^{-1} X_{t+2} - \alpha^{-1} \varepsilon_{t+2}) - \alpha^{-1} \varepsilon_{t+1} \\&= \alpha^{-2} X_{t+2} - \alpha^{-2} \varepsilon_{t+2} - \alpha^{-1} \varepsilon_{t+1} \\&\vdots \quad \quad \quad \vdots \quad \quad \quad \vdots \quad \quad \quad \vdots \\&= - \sum_{k=1}^{\infty} (\alpha^{-1})^k \varepsilon_{t+k} \quad (\text{existe car } |\alpha^{-1}| < 1)\end{aligned}$$

## Processus ARMA : deux exemples

- ❶ Soit le processus AR(1) avec  $X_t = \alpha X_{t-1} + \varepsilon_t$  et  $|\alpha| < 1$ . Alors :

$$\begin{aligned} X_t &= \alpha(\alpha X_{t-2} + \varepsilon_{t-1}) + \varepsilon_t \\ &= \alpha^2 X_{t-2} + \alpha \varepsilon_{t-1} + \varepsilon_t \\ &= \alpha^2(\alpha X_{t-3} + \varepsilon_{t-2}) + \alpha \varepsilon_{t-1} + \varepsilon_t \\ &= \alpha^3 X_{t-3} + \alpha^2 \varepsilon_{t-2} + \alpha \varepsilon_{t-1} + \varepsilon_t \\ &\vdots \quad \quad \quad \vdots \quad \quad \quad \vdots \quad \quad \quad \vdots \\ &= \sum_{k=0}^{\infty} \alpha^k \varepsilon_{t-k} \quad (\text{existe car } |\alpha| < 1) \end{aligned}$$

- ❷ Soit le processus AR(1) avec  $X_t = \alpha X_{t-1} + \varepsilon_t$  et  $|\alpha| > 1$ . Alors :

$$\begin{aligned} X_t &= \alpha^{-1} X_{t+1} - \alpha^{-1} \varepsilon_{t+1} \\ &= \alpha^{-1}(\alpha^{-1} X_{t+2} - \alpha^{-1} \varepsilon_{t+2}) - \alpha^{-1} \varepsilon_{t+1} \\ &= \alpha^{-2} X_{t+2} - \alpha^{-2} \varepsilon_{t+2} - \alpha^{-1} \varepsilon_{t+1} \\ &\vdots \quad \quad \quad \vdots \quad \quad \quad \vdots \quad \quad \quad \vdots \\ &= -\sum_{k=1}^{\infty} (\alpha^{-1})^k \varepsilon_{t+k} \quad (\text{existe car } |\alpha^{-1}| < 1) \end{aligned}$$

## Processus ARMA : deux exemples

- ❶ Soit le processus AR(1) avec  $X_t = \alpha X_{t-1} + \varepsilon_t$  et  $|\alpha| < 1$ . Alors :

$$\begin{aligned} X_t &= \alpha(\alpha X_{t-2} + \varepsilon_{t-1}) + \varepsilon_t \\ &= \alpha^2 X_{t-2} + \alpha \varepsilon_{t-1} + \varepsilon_t \\ &= \alpha^2(\alpha X_{t-3} + \varepsilon_{t-2}) + \alpha \varepsilon_{t-1} + \varepsilon_t \\ &= \alpha^3 X_{t-3} + \alpha^2 \varepsilon_{t-2} + \alpha \varepsilon_{t-1} + \varepsilon_t \\ &\vdots \quad \quad \quad \vdots \quad \quad \quad \vdots \quad \quad \quad \vdots \\ &= \sum_{k=0}^{\infty} \alpha^k \varepsilon_{t-k} \quad (\text{existe car } |\alpha| < 1) \end{aligned}$$

- ❷ Soit le processus AR(1) avec  $X_t = \alpha X_{t-1} + \varepsilon_t$  et  $|\alpha| > 1$ . Alors :

$$\begin{aligned} X_t &= \alpha^{-1} X_{t+1} - \alpha^{-1} \varepsilon_{t+1} \\ &= \alpha^{-1}(\alpha^{-1} X_{t+2} - \alpha^{-1} \varepsilon_{t+2}) - \alpha^{-1} \varepsilon_{t+1} \\ &= \alpha^{-2} X_{t+2} - \alpha^{-2} \varepsilon_{t+2} - \alpha^{-1} \varepsilon_{t+1} \\ &\vdots \quad \quad \quad \vdots \quad \quad \quad \vdots \\ &= - \sum_{k=1}^{\infty} (\alpha^{-1})^k \varepsilon_{t+k} \quad (\text{existe car } |\alpha^{-1}| < 1) \end{aligned}$$

## Processus ARMA : deux exemples

- ❶ Soit le processus AR(1) avec  $X_t = \alpha X_{t-1} + \varepsilon_t$  et  $|\alpha| < 1$ . Alors :

$$\begin{aligned} X_t &= \alpha(\alpha X_{t-2} + \varepsilon_{t-1}) + \varepsilon_t \\ &= \alpha^2 X_{t-2} + \alpha \varepsilon_{t-1} + \varepsilon_t \\ &= \alpha^2(\alpha X_{t-3} + \varepsilon_{t-2}) + \alpha \varepsilon_{t-1} + \varepsilon_t \\ &= \alpha^3 X_{t-3} + \alpha^2 \varepsilon_{t-2} + \alpha \varepsilon_{t-1} + \varepsilon_t \\ &\vdots \quad \quad \quad \vdots \quad \quad \quad \vdots \quad \quad \quad \vdots \\ &= \sum_{k=0}^{\infty} \alpha^k \varepsilon_{t-k} \quad (\text{existe car } |\alpha| < 1) \end{aligned}$$

- ❷ Soit le processus AR(1) avec  $X_t = \alpha X_{t-1} + \varepsilon_t$  et  $|\alpha| > 1$ . Alors :

$$\begin{aligned} X_t &= \alpha^{-1} X_{t+1} - \alpha^{-1} \varepsilon_{t+1} \\ &= \alpha^{-1}(\alpha^{-1} X_{t+2} - \alpha^{-1} \varepsilon_{t+2}) - \alpha^{-1} \varepsilon_{t+1} \\ &= \alpha^{-2} X_{t+2} - \alpha^{-2} \varepsilon_{t+2} - \alpha^{-1} \varepsilon_{t+1} \\ &\vdots \quad \quad \quad \vdots \quad \quad \quad \vdots \quad \quad \quad \vdots \\ &= - \sum_{k=1}^{\infty} (\alpha^{-1})^k \varepsilon_{t+k} \quad (\text{existe car } |\alpha^{-1}| < 1) \end{aligned}$$

## Processus ARMA : deux exemples

- ❶ Soit le processus AR(1) avec  $X_t = \alpha X_{t-1} + \varepsilon_t$  et  $|\alpha| < 1$ . Alors :

$$\begin{aligned}X_t &= \alpha(\alpha X_{t-2} + \varepsilon_{t-1}) + \varepsilon_t \\&= \alpha^2 X_{t-2} + \alpha \varepsilon_{t-1} + \varepsilon_t \\&= \alpha^2(\alpha X_{t-3} + \varepsilon_{t-2}) + \alpha \varepsilon_{t-1} + \varepsilon_t \\&= \alpha^3 X_{t-3} + \alpha^2 \varepsilon_{t-2} + \alpha \varepsilon_{t-1} + \varepsilon_t \\&\vdots \quad \quad \quad \vdots \quad \quad \quad \vdots \quad \quad \quad \vdots \\&= \sum_{k=0}^{\infty} \alpha^k \varepsilon_{t-k} \quad (\text{existe car } |\alpha| < 1)\end{aligned}$$

- ❷ Soit le processus AR(1) avec  $X_t = \alpha X_{t-1} + \varepsilon_t$  et  $|\alpha| > 1$ . Alors :

$$\begin{aligned}X_t &= \alpha^{-1} X_{t+1} - \alpha^{-1} \varepsilon_{t+1} \\&= \alpha^{-1}(\alpha^{-1} X_{t+2} - \alpha^{-1} \varepsilon_{t+2}) - \alpha^{-1} \varepsilon_{t+1} \\&= \alpha^{-2} X_{t+2} - \alpha^{-2} \varepsilon_{t+2} - \alpha^{-1} \varepsilon_{t+1} \\&\vdots \quad \quad \quad \vdots \quad \quad \quad \vdots \quad \quad \quad \vdots \\&= -\sum_{k=1}^{\infty} (\alpha^{-1})^k \varepsilon_{t+k} \quad (\text{existe car } |\alpha^{-1}| < 1)\end{aligned}$$

## Processus ARMA : deux exemples

- ① Soit le processus AR(1) avec  $X_t = \alpha X_{t-1} + \varepsilon_t$  et  $|\alpha| < 1$ . Alors :

$$\begin{aligned} X_t &= \alpha(\alpha X_{t-2} + \varepsilon_{t-1}) + \varepsilon_t \\ &= \alpha^2 X_{t-2} + \alpha \varepsilon_{t-1} + \varepsilon_t \\ &= \alpha^2(\alpha X_{t-3} + \varepsilon_{t-2}) + \alpha \varepsilon_{t-1} + \varepsilon_t \\ &= \alpha^3 X_{t-3} + \alpha^2 \varepsilon_{t-2} + \alpha \varepsilon_{t-1} + \varepsilon_t \\ &\vdots \quad \quad \quad \vdots \quad \quad \quad \vdots \quad \quad \quad \vdots \\ &= \sum_{k=0}^{\infty} \alpha^k \varepsilon_{t-k} \quad (\text{existe car } |\alpha| < 1) \end{aligned}$$

- ② Soit le processus AR(1) avec  $X_t = \alpha X_{t-1} + \varepsilon_t$  et  $|\alpha| > 1$ . Alors :

$$\begin{aligned} X_t &= \alpha^{-1} X_{t+1} - \alpha^{-1} \varepsilon_{t+1} \\ &= \alpha^{-1}(\alpha^{-1} X_{t+2} - \alpha^{-1} \varepsilon_{t+2}) - \alpha^{-1} \varepsilon_{t+1} \\ &= \alpha^{-2} X_{t+2} - \alpha^{-2} \varepsilon_{t+2} - \alpha^{-1} \varepsilon_{t+1} \\ &\vdots \quad \quad \quad \vdots \quad \quad \quad \vdots \quad \quad \quad \vdots \\ &= - \sum_{k=1}^{\infty} (\alpha^{-1})^k \varepsilon_{t+k} \quad (\text{existe car } |\alpha^{-1}| < 1) \end{aligned}$$

# Processus ARMA : condition de stationnarité

## Proposition

Soit  $(X_t)_{t \in \mathbb{Z}}$  défini pour  $t \in \mathbb{Z}$  par

$$X_t + a_1 X_{t-1} + \cdots + a_p X_{t-p} = \varepsilon_t + b_1 \varepsilon_{t-1} + \cdots + b_q \varepsilon_{t-q}$$

Soit le polynôme  $P(x) = 1 + a_1 x + \cdots + a_p x^p$ .

Alors  $(X_t)$  est un processus ARMA( $p, q$ ) si les racines de  $P$  ne sont pas sur le cercle unité, soit  $P(z) \neq 0$  pour tout  $|z| = 1$

Cas considéré : On ne considérera désormais que le cas **causal** :

Racines de  $P$  en dehors du disque unité, soit  $P(z) \neq 0$  pour tout  $|z| \leq 1$

$$\implies X_t = \sum_{k=0}^{\infty} \alpha_k \varepsilon_{t-k} \text{ dépend linéairement des valeurs passées de } \varepsilon$$

# Processus ARMA : condition de stationnarité

## Proposition

Soit  $(X_t)_{t \in \mathbb{Z}}$  défini pour  $t \in \mathbb{Z}$  par

$$X_t + a_1 X_{t-1} + \cdots + a_p X_{t-p} = \varepsilon_t + b_1 \varepsilon_{t-1} + \cdots + b_q \varepsilon_{t-q}$$

Soit le polynôme  $P(x) = 1 + a_1 x + \cdots + a_p x^p$ .

Alors  $(X_t)$  est un processus ARMA( $p, q$ ) si les racines de  $P$  ne sont pas sur le cercle unité, soit  $P(z) \neq 0$  pour tout  $|z| = 1$

Cas considéré : On ne considérera désormais que le cas **causal** :

Racines de  $P$  en dehors du disque unité, soit  $P(z) \neq 0$  pour tout  $|z| \leq 1$

$$\implies X_t = \sum_{k=0}^{\infty} \alpha_k \varepsilon_{t-k} \text{ dépend linéairement des valeurs passées de } \varepsilon$$

## Processus ARMA : condition de stationnarité

### Proposition

Soit  $(X_t)_{t \in \mathbb{Z}}$  défini pour  $t \in \mathbb{Z}$  par

$$X_t + a_1 X_{t-1} + \cdots + a_p X_{t-p} = \varepsilon_t + b_1 \varepsilon_{t-1} + \cdots + b_q \varepsilon_{t-q}$$

Soit le polynôme  $P(x) = 1 + a_1 x + \cdots + a_p x^p$ .

Alors  $(X_t)$  est un processus ARMA( $p, q$ ) si les racines de  $P$  ne sont pas sur le cercle unité, soit  $P(z) \neq 0$  pour tout  $|z| = 1$

Cas considéré : On ne considérera désormais que le cas causal :

Racines de  $P$  en dehors du disque unité, soit  $P(z) \neq 0$  pour tout  $|z| \leq 1$

$$\implies X_t = \sum_{k=0}^{\infty} \alpha_k \varepsilon_{t-k} \text{ dépend linéairement des valeurs passées de } \varepsilon$$

## Processus ARMA : condition de stationnarité

### Proposition

Soit  $(X_t)_{t \in \mathbb{Z}}$  défini pour  $t \in \mathbb{Z}$  par

$$X_t + a_1 X_{t-1} + \cdots + a_p X_{t-p} = \varepsilon_t + b_1 \varepsilon_{t-1} + \cdots + b_q \varepsilon_{t-q}$$

Soit le polynôme  $P(x) = 1 + a_1 x + \cdots + a_p x^p$ .

Alors  $(X_t)$  est un processus ARMA( $p, q$ ) si les racines de  $P$  ne sont pas sur le cercle unité, soit  $P(z) \neq 0$  pour tout  $|z| = 1$

Cas considéré : On ne considérera désormais que le cas causal :

Racines de  $P$  en dehors du disque unité, soit  $P(z) \neq 0$  pour tout  $|z| \leq 1$

$$\implies X_t = \sum_{k=0}^{\infty} \alpha_k \varepsilon_{t-k} \text{ dépend linéairement des valeurs passés de } \varepsilon$$

# Processus ARMA : condition de stationnarité

## Proposition

Soit  $(X_t)_{t \in \mathbb{Z}}$  défini pour  $t \in \mathbb{Z}$  par

$$X_t + a_1 X_{t-1} + \cdots + a_p X_{t-p} = \varepsilon_t + b_1 \varepsilon_{t-1} + \cdots + b_q \varepsilon_{t-q}$$

Soit le polynôme  $P(x) = 1 + a_1 x + \cdots + a_p x^p$ .

Alors  $(X_t)$  est un processus ARMA( $p, q$ ) si les racines de  $P$  ne sont pas sur le cercle unité, soit  $P(z) \neq 0$  pour tout  $|z| = 1$

**Cas considéré** : On ne considérera désormais que le cas **causal** :

Racines de  $P$  en dehors du disque unité, soit  $P(z) \neq 0$  pour tout  $|z| \leq 1$

$$\implies X_t = \sum_{k=0}^{\infty} \alpha_k \varepsilon_{t-k} \text{ dépend linéairement des valeurs passées de } \varepsilon$$

# Processus ARMA : condition de stationnarité

## Proposition

Soit  $(X_t)_{t \in \mathbb{Z}}$  défini pour  $t \in \mathbb{Z}$  par

$$X_t + a_1 X_{t-1} + \cdots + a_p X_{t-p} = \varepsilon_t + b_1 \varepsilon_{t-1} + \cdots + b_q \varepsilon_{t-q}$$

Soit le polynôme  $P(x) = 1 + a_1 x + \cdots + a_p x^p$ .

Alors  $(X_t)$  est un processus ARMA( $p, q$ ) si les racines de  $P$  ne sont pas sur le cercle unité, soit  $P(z) \neq 0$  pour tout  $|z| = 1$

**Cas considéré** : On ne considérera désormais que le cas **causal** :

Racines de  $P$  en dehors du disque unité, soit  $P(z) \neq 0$  pour tout  $|z| \leq 1$

$$\implies X_t = \sum_{k=0}^{\infty} \alpha_k \varepsilon_{t-k} \text{ dépend linéairement des valeurs passées de } \varepsilon$$

## Processus ARMA : condition de stationnarité

### Proposition

Soit  $(X_t)_{t \in \mathbb{Z}}$  défini pour  $t \in \mathbb{Z}$  par

$$X_t + a_1 X_{t-1} + \cdots + a_p X_{t-p} = \varepsilon_t + b_1 \varepsilon_{t-1} + \cdots + b_q \varepsilon_{t-q}$$

Soit le polynôme  $P(x) = 1 + a_1 x + \cdots + a_p x^p$ .

Alors  $(X_t)$  est un processus ARMA( $p, q$ ) si les racines de  $P$  ne sont pas sur le cercle unité, soit  $P(z) \neq 0$  pour tout  $|z| = 1$

**Cas considéré** : On ne considérera désormais que le cas **causal** :

Racines de  $P$  en dehors du disque unité, soit  $P(z) \neq 0$  pour tout  $|z| \leq 1$

$$\implies X_t = \sum_{k=0}^{\infty} \alpha_k \varepsilon_{t-k} \text{ dépend linéairement des valeurs passées de } \varepsilon$$

# Propriétés des processus ARMA

## Proposition

Soit le processus ARMA( $p, q$ ) causal  $(X_t)_{t \in \mathbb{Z}}$  défini pour  $t \in \mathbb{Z}$  par

$$X_t + a_1 X_{t-1} + \dots + a_p X_{t-p} = \varepsilon_t + b_1 \varepsilon_{t-1} + \dots + b_q \varepsilon_{t-q}$$

Alors :

- 1  $X_t = \sum_{k=0}^{\infty} \alpha_j \varepsilon_{t-j}$  et  $\exists 0 \leq \rho < 1, c \geq 0, |\alpha_j| \leq c \rho^j$  pour tout  $j \geq 0$  ;
- 2  $r_X(k) = \text{cov}(X_0, X_k) : \exists 0 \leq \rho' < 1, c' \geq 0, |r_X(k)| \leq c' \rho'^k \forall k \geq 0$  ;

Cas particulier : Si  $(X_t)$  est un processus MA( $q$ ),

- 1  $\alpha_j = b_j$  pour  $j = 1, \dots, q$ , et  $\alpha_j = 0$  si  $j \geq q + 1$  ;
- 2  $r_X(k) = 0$  pour  $k \geq q + 1$  et avec  $b_0 = 1$ , pour  $0 \leq k \leq q$ ,

$$r_X(k) = \text{cov}\left(\sum_{j=0}^q b_j \varepsilon_{-j}, \sum_{\ell=0}^q b_\ell \varepsilon_{k-\ell}\right) = \sum_{j=0}^q \sum_{\ell=0}^q b_j b_\ell \mathbb{E}[\varepsilon_{-j} \varepsilon_{k-\ell}] = \sigma_\varepsilon^2 \sum_{j=0}^{q-k} b_j b_{j+k}$$

# Propriétés des processus ARMA

## Proposition

Soit le processus ARMA( $p, q$ ) causal  $(X_t)_{t \in \mathbb{Z}}$  défini pour  $t \in \mathbb{Z}$  par

$$X_t + a_1 X_{t-1} + \dots + a_p X_{t-p} = \varepsilon_t + b_1 \varepsilon_{t-1} + \dots + b_q \varepsilon_{t-q}$$

Alors :

- 1  $X_t = \sum_{k=0}^{\infty} \alpha_j \varepsilon_{t-j}$  et  $\exists 0 \leq \rho < 1, c \geq 0, |\alpha_j| \leq c \rho^j$  pour tout  $j \geq 0$  ;
- 2  $r_X(k) = \text{cov}(X_0, X_k) : \exists 0 \leq \rho' < 1, c' \geq 0, |r_X(k)| \leq c' \rho'^k \forall k \geq 0$  ;

Cas particulier : Si  $(X_t)$  est un processus MA( $q$ ),

- 1  $\alpha_j = b_j$  pour  $j = 1, \dots, q$ , et  $\alpha_j = 0$  si  $j \geq q + 1$  ;
- 2  $r_X(k) = 0$  pour  $k \geq q + 1$  et avec  $b_0 = 1$ , pour  $0 \leq k \leq q$ ,

$$r_X(k) = \text{cov}\left(\sum_{j=0}^q b_j \varepsilon_{-j}, \sum_{\ell=0}^q b_\ell \varepsilon_{k-\ell}\right) = \sum_{j=0}^q \sum_{\ell=0}^q b_j b_\ell \mathbb{E}[\varepsilon_{-j} \varepsilon_{k-\ell}] = \sigma_\varepsilon^2 \sum_{j=0}^{q-k} b_j b_{j+k}$$

# Propriétés des processus ARMA

## Proposition

Soit le processus ARMA( $p, q$ ) causal  $(X_t)_{t \in \mathbb{Z}}$  défini pour  $t \in \mathbb{Z}$  par

$$X_t + a_1 X_{t-1} + \dots + a_p X_{t-p} = \varepsilon_t + b_1 \varepsilon_{t-1} + \dots + b_q \varepsilon_{t-q}$$

Alors :

- 1  $X_t = \sum_{k=0}^{\infty} \alpha_j \varepsilon_{t-j}$  et  $\exists 0 \leq \rho < 1, c \geq 0, |\alpha_j| \leq c \rho^j$  pour tout  $j \geq 0$  ;
- 2  $r_X(k) = \text{cov}(X_0, X_k) : \exists 0 \leq \rho' < 1, c' \geq 0, |r_X(k)| \leq c' \rho'^k \forall k \geq 0$  ;

Cas particulier : Si  $(X_t)$  est un processus MA( $q$ ),

- 1  $\alpha_j = b_j$  pour  $j = 1, \dots, q$ , et  $\alpha_j = 0$  si  $j \geq q + 1$  ;
- 2  $r_X(k) = 0$  pour  $k \geq q + 1$  et avec  $b_0 = 1$ , pour  $0 \leq k \leq q$ ,

$$r_X(k) = \text{cov}\left(\sum_{j=0}^q b_j \varepsilon_{-j}, \sum_{\ell=0}^q b_\ell \varepsilon_{k-\ell}\right) = \sum_{j=0}^q \sum_{\ell=0}^q b_j b_\ell \mathbb{E}[\varepsilon_{-j} \varepsilon_{k-\ell}] = \sigma_\varepsilon^2 \sum_{j=0}^{q-k} b_j b_{j+k}$$

# Propriétés des processus ARMA

## Proposition

Soit le processus ARMA( $p, q$ ) causal  $(X_t)_{t \in \mathbb{Z}}$  défini pour  $t \in \mathbb{Z}$  par

$$X_t + a_1 X_{t-1} + \dots + a_p X_{t-p} = \varepsilon_t + b_1 \varepsilon_{t-1} + \dots + b_q \varepsilon_{t-q}$$

Alors :

- 1  $X_t = \sum_{k=0}^{\infty} \alpha_j \varepsilon_{t-j}$  et  $\exists 0 \leq \rho < 1, c \geq 0, |\alpha_j| \leq c \rho^j$  pour tout  $j \geq 0$  ;
- 2  $r_X(k) = \text{cov}(X_0, X_k) : \exists 0 \leq \rho' < 1, c' \geq 0, |r_X(k)| \leq c' \rho'^k \forall k \geq 0$  ;

Cas particulier : Si  $(X_t)$  est un processus MA( $q$ ),

- 1  $\alpha_j = b_j$  pour  $j = 1, \dots, q$ , et  $\alpha_j = 0$  si  $j \geq q + 1$  ;
- 2  $r_X(k) = 0$  pour  $k \geq q + 1$  et avec  $b_0 = 1$ , pour  $0 \leq k \leq q$ ,

$$r_X(k) = \text{cov}\left(\sum_{j=0}^q b_j \varepsilon_{-j}, \sum_{\ell=0}^q b_\ell \varepsilon_{k-\ell}\right) = \sum_{j=0}^q \sum_{\ell=0}^q b_j b_\ell \mathbb{E}[\varepsilon_{-j} \varepsilon_{k-\ell}] = \sigma_\varepsilon^2 \sum_{j=0}^{q-k} b_j b_{j+k}$$

# Propriétés des processus ARMA

## Proposition

Soit le processus ARMA( $p, q$ ) causal  $(X_t)_{t \in \mathbb{Z}}$  défini pour  $t \in \mathbb{Z}$  par

$$X_t + a_1 X_{t-1} + \dots + a_p X_{t-p} = \varepsilon_t + b_1 \varepsilon_{t-1} + \dots + b_q \varepsilon_{t-q}$$

Alors :

- 1  $X_t = \sum_{k=0}^{\infty} \alpha_j \varepsilon_{t-j}$  et  $\exists 0 \leq \rho < 1, c \geq 0, |\alpha_j| \leq c \rho^j$  pour tout  $j \geq 0$  ;
- 2  $r_X(k) = \text{cov}(X_0, X_k) : \exists 0 \leq \rho' < 1, c' \geq 0, |r_X(k)| \leq c' \rho'^k \forall k \geq 0$  ;

**Cas particulier** : Si  $(X_t)$  est un processus MA( $q$ ),

- 1  $\alpha_j = b_j$  pour  $j = 1, \dots, q$ , et  $\alpha_j = 0$  si  $j \geq q + 1$  ;
- 2  $r_X(k) = 0$  pour  $k \geq q + 1$  et avec  $b_0 = 1$ , pour  $0 \leq k \leq q$ ,

$$r_X(k) = \text{cov}\left(\sum_{j=0}^q b_j \varepsilon_{-j}, \sum_{\ell=0}^q b_\ell \varepsilon_{k-\ell}\right) = \sum_{j=0}^q \sum_{\ell=0}^q b_j b_\ell \mathbb{E}[\varepsilon_{-j} \varepsilon_{k-\ell}] = \sigma_\varepsilon^2 \sum_{j=0}^{q-k} b_j b_{j+k}$$

# Propriétés des processus ARMA

## Proposition

Soit le processus ARMA( $p, q$ ) causal  $(X_t)_{t \in \mathbb{Z}}$  défini pour  $t \in \mathbb{Z}$  par

$$X_t + a_1 X_{t-1} + \dots + a_p X_{t-p} = \varepsilon_t + b_1 \varepsilon_{t-1} + \dots + b_q \varepsilon_{t-q}$$

Alors :

- 1  $X_t = \sum_{k=0}^{\infty} \alpha_j \varepsilon_{t-j}$  et  $\exists 0 \leq \rho < 1, c \geq 0, |\alpha_j| \leq c \rho^j$  pour tout  $j \geq 0$  ;
- 2  $r_X(k) = \text{cov}(X_0, X_k) : \exists 0 \leq \rho' < 1, c' \geq 0, |r_X(k)| \leq c' \rho'^k \forall k \geq 0$  ;

**Cas particulier** : Si  $(X_t)$  est un processus MA( $q$ ),

- 1  $\alpha_j = b_j$  pour  $j = 1, \dots, q$ , et  $\alpha_j = 0$  si  $j \geq q + 1$  ;
- 2  $r_X(k) = 0$  pour  $k \geq q + 1$  et avec  $b_0 = 1$ , pour  $0 \leq k \leq q$ ,

$$r_X(k) = \text{cov}\left(\sum_{j=0}^q b_j \varepsilon_{-j}, \sum_{\ell=0}^q b_\ell \varepsilon_{k-\ell}\right) = \sum_{j=0}^q \sum_{\ell=0}^q b_j b_\ell \mathbb{E}[\varepsilon_{-j} \varepsilon_{k-\ell}] = \sigma_\varepsilon^2 \sum_{j=0}^{q-k} b_j b_{j+k}$$

# Propriétés des processus ARMA

## Proposition

Soit le processus ARMA( $p, q$ ) causal  $(X_t)_{t \in \mathbb{Z}}$  défini pour  $t \in \mathbb{Z}$  par

$$X_t + a_1 X_{t-1} + \dots + a_p X_{t-p} = \varepsilon_t + b_1 \varepsilon_{t-1} + \dots + b_q \varepsilon_{t-q}$$

Alors :

- 1  $X_t = \sum_{k=0}^{\infty} \alpha_j \varepsilon_{t-j}$  et  $\exists 0 \leq \rho < 1, c \geq 0, |\alpha_j| \leq c \rho^j$  pour tout  $j \geq 0$  ;
- 2  $r_X(k) = \text{cov}(X_0, X_k) : \exists 0 \leq \rho' < 1, c' \geq 0, |r_X(k)| \leq c' \rho'^k \forall k \geq 0$  ;

**Cas particulier** : Si  $(X_t)$  est un processus MA( $q$ ),

- 1  $\alpha_j = b_j$  pour  $j = 1, \dots, q$ , et  $\alpha_j = 0$  si  $j \geq q + 1$  ;
- 2  $r_X(k) = 0$  pour  $k \geq q + 1$  et avec  $b_0 = 1$ , pour  $0 \leq k \leq q$ ,

$$r_X(k) = \text{cov}\left(\sum_{j=0}^q b_j \varepsilon_{-j}, \sum_{\ell=0}^q b_\ell \varepsilon_{k-\ell}\right) = \sum_{j=0}^q \sum_{\ell=0}^q b_j b_\ell \mathbb{E}[\varepsilon_{-j} \varepsilon_{k-\ell}] = \sigma_\varepsilon^2 \sum_{j=0}^{q-k} b_j b_{j+k}$$

# Propriétés des processus ARMA

## Proposition

Soit le processus ARMA( $p, q$ ) causal  $(X_t)_{t \in \mathbb{Z}}$  défini pour  $t \in \mathbb{Z}$  par

$$X_t + a_1 X_{t-1} + \dots + a_p X_{t-p} = \varepsilon_t + b_1 \varepsilon_{t-1} + \dots + b_q \varepsilon_{t-q}$$

Alors :

- 1  $X_t = \sum_{k=0}^{\infty} \alpha_j \varepsilon_{t-j}$  et  $\exists 0 \leq \rho < 1, c \geq 0, |\alpha_j| \leq c \rho^j$  pour tout  $j \geq 0$  ;
- 2  $r_X(k) = \text{cov}(X_0, X_k) : \exists 0 \leq \rho' < 1, c' \geq 0, |r_X(k)| \leq c' \rho'^k \forall k \geq 0$  ;

**Cas particulier** : Si  $(X_t)$  est un processus MA( $q$ ),

- 1  $\alpha_j = b_j$  pour  $j = 1, \dots, q$ , et  $\alpha_j = 0$  si  $j \geq q + 1$  ;
- 2  $r_X(k) = 0$  pour  $k \geq q + 1$  et avec  $b_0 = 1$ , pour  $0 \leq k \leq q$ ,

$$r_X(k) = \text{cov}\left(\sum_{j=0}^q b_j \varepsilon_{-j}, \sum_{\ell=0}^q b_\ell \varepsilon_{k-\ell}\right) = \sum_{j=0}^q \sum_{\ell=0}^q b_j b_\ell \mathbf{E}[\varepsilon_{-j} \varepsilon_{k-\ell}] = \sigma_\varepsilon^2 \sum_{j=0}^{q-k} b_j b_{j+k}$$

# Propriétés des processus ARMA (2)

## Propriété

Soit le processus ARMA( $p, q$ ) causal  $(X_t)_{t \in \mathbb{Z}}$  défini pour  $t \in \mathbb{Z}$  par

$$X_t + a_1 X_{t-1} + \cdots + a_p X_{t-p} = \omega_0 + \varepsilon_t + b_1 \varepsilon_{t-1} + \cdots + b_q \varepsilon_{t-q}$$

Alors :

- 1 Si  $\omega_0 = 0$ , on a  $\mathbb{E}(X_t) = 0$  ;
- 2 Si  $\omega \neq 0$ , alors  $(X'_t)$  tel que  $X'_t = X_t - \frac{\omega_0}{1+a_1+\cdots+a_p}$  vérifie

$$X'_t + a_1 X'_{t-1} + \cdots + a_p X'_{t-p} = \varepsilon_t + b_1 \varepsilon_{t-1} + \cdots + b_q \varepsilon_{t-q}$$

$$\implies \mathbb{E}(X_t) = \frac{\omega_0}{1 + a_1 + \cdots + a_p} : \text{ARMA}(p, q) \text{ décentré}$$

## Propriété

Un processus ARMA( $p, q$ ) causal de bruit blanc  $(\varepsilon_t)$  gaussien est gaussien.

## Propriétés des processus ARMA (2)

### Propriété

Soit le processus ARMA( $p, q$ ) causal  $(X_t)_{t \in \mathbb{Z}}$  défini pour  $t \in \mathbb{Z}$  par

$$X_t + a_1 X_{t-1} + \cdots + a_p X_{t-p} = \omega_0 + \varepsilon_t + b_1 \varepsilon_{t-1} + \cdots + b_q \varepsilon_{t-q}$$

Alors :

① Si  $\omega_0 = 0$ , on a  $\mathbb{E}(X_t) = 0$  ;

② Si  $\omega \neq 0$ , alors  $(X'_t)$  tel que  $X'_t = X_t - \frac{\omega_0}{1+a_1+\cdots+a_p}$  vérifie

$$X'_t + a_1 X'_{t-1} + \cdots + a_p X'_{t-p} = \varepsilon_t + b_1 \varepsilon_{t-1} + \cdots + b_q \varepsilon_{t-q}$$

$$\implies \mathbb{E}(X_t) = \frac{\omega_0}{1 + a_1 + \cdots + a_p} : \text{ARMA}(p, q) \text{ décentré}$$

### Propriété

Un processus ARMA( $p, q$ ) causal de bruit blanc  $(\varepsilon_t)$  gaussien est gaussien.

## Propriétés des processus ARMA (2)

### Propriété

Soit le processus ARMA( $p, q$ ) causal  $(X_t)_{t \in \mathbb{Z}}$  défini pour  $t \in \mathbb{Z}$  par

$$X_t + a_1 X_{t-1} + \cdots + a_p X_{t-p} = \omega_0 + \varepsilon_t + b_1 \varepsilon_{t-1} + \cdots + b_q \varepsilon_{t-q}$$

Alors :

- 1 Si  $\omega_0 = 0$ , on a  $\mathbb{E}(X_t) = 0$  ;
- 2 Si  $\omega \neq 0$ , alors  $(X'_t)$  tel que  $X'_t = X_t - \frac{\omega_0}{1+a_1+\cdots+a_p}$  vérifie

$$X'_t + a_1 X'_{t-1} + \cdots + a_p X'_{t-p} = \varepsilon_t + b_1 \varepsilon_{t-1} + \cdots + b_q \varepsilon_{t-q}$$

$$\implies \mathbb{E}(X_t) = \frac{\omega_0}{1 + a_1 + \cdots + a_p} : \text{ARMA}(p, q) \text{ décentré}$$

### Propriété

Un processus ARMA( $p, q$ ) causal de bruit blanc  $(\varepsilon_t)$  gaussien est gaussien.

## Propriétés des processus ARMA (2)

### Propriété

Soit le processus ARMA( $p, q$ ) causal  $(X_t)_{t \in \mathbb{Z}}$  défini pour  $t \in \mathbb{Z}$  par

$$X_t + a_1 X_{t-1} + \cdots + a_p X_{t-p} = \omega_0 + \varepsilon_t + b_1 \varepsilon_{t-1} + \cdots + b_q \varepsilon_{t-q}$$

Alors :

- 1 Si  $\omega_0 = 0$ , on a  $\mathbf{E}(X_t) = 0$  ;
- 2 Si  $\omega \neq 0$ , alors  $(X'_t)$  tel que  $X'_t = X_t - \frac{\omega_0}{1+a_1+\cdots+a_p}$  vérifie

$$X'_t + a_1 X'_{t-1} + \cdots + a_p X'_{t-p} = \varepsilon_t + b_1 \varepsilon_{t-1} + \cdots + b_q \varepsilon_{t-q}$$

$$\implies \mathbf{E}(X_t) = \frac{\omega_0}{1 + a_1 + \cdots + a_p} : \text{ARMA}(p, q) \text{ décentré}$$

### Propriété

Un processus ARMA( $p, q$ ) causal de bruit blanc  $(\varepsilon_t)$  gaussien est gaussien.

## Propriétés des processus ARMA (2)

### Propriété

Soit le processus ARMA( $p, q$ ) causal  $(X_t)_{t \in \mathbb{Z}}$  défini pour  $t \in \mathbb{Z}$  par

$$X_t + a_1 X_{t-1} + \dots + a_p X_{t-p} = \omega_0 + \varepsilon_t + b_1 \varepsilon_{t-1} + \dots + b_q \varepsilon_{t-q}$$

Alors :

- 1 Si  $\omega_0 = 0$ , on a  $\mathbf{E}(X_t) = 0$  ;
- 2 Si  $\omega_0 \neq 0$ , alors  $(X'_t)$  tel que  $X'_t = X_t - \frac{\omega_0}{1+a_1+\dots+a_p}$  vérifie

$$X'_t + a_1 X'_{t-1} + \dots + a_p X'_{t-p} = \varepsilon_t + b_1 \varepsilon_{t-1} + \dots + b_q \varepsilon_{t-q}$$

$$\implies \mathbf{E}(X_t) = \frac{\omega_0}{1 + a_1 + \dots + a_p} : \text{ARMA}(p, q) \text{ décentré}$$

### Propriété

Un processus ARMA( $p, q$ ) causal de bruit blanc  $(\varepsilon_t)$  gaussien est gaussien.

# Processus ARMA : corrélogramme

## Proposition

Soit  $(X_t)_{t \in \mathbb{Z}}$  un processus ARMA( $p, q$ ) stationnaire causal. Soit  $\rho(\cdot)$  son autocorrélation et  $\hat{\rho}(\cdot)$  son autocorrélation empirique calculée sur  $(X_1, \dots, X_n)$ . Alors :

$$\hat{\rho}(k) \xrightarrow[n \rightarrow +\infty]{\mathcal{P}} \rho(k) \quad \text{pour tout } k \in \mathbb{N}$$

Conséquence : l'ACF converge vers les autocorrélations théoriques

# Processus ARMA : corrélogramme

## Proposition

Soit  $(X_t)_{t \in \mathbb{Z}}$  un processus ARMA( $p, q$ ) stationnaire causal. Soit  $\rho(\cdot)$  son autocorrélation et  $\hat{\rho}(\cdot)$  son autocorrélation empirique calculée sur  $(X_1, \dots, X_n)$ . Alors :

$$\hat{\rho}(k) \xrightarrow[n \rightarrow +\infty]{\mathcal{P}} \rho(k) \quad \text{pour tout } k \in \mathbb{N}$$

Conséquence : l'ACF converge vers les autocorrélations théoriques

# Processus ARMA : corrélogramme

## Proposition

Soit  $(X_t)_{t \in \mathbb{Z}}$  un processus ARMA( $p, q$ ) stationnaire causal. Soit  $\rho(\cdot)$  son autocorrélation et  $\hat{\rho}(\cdot)$  son autocorrélation empirique calculée sur  $(X_1, \dots, X_n)$ . Alors :

$$\hat{\rho}(k) \xrightarrow[n \rightarrow +\infty]{\mathcal{P}} \rho(k) \quad \text{pour tout } k \in \mathbb{N}$$

Conséquence : l'ACF converge vers les autocorrélations théoriques

# Processus ARMA : corrélogramme

## Proposition

Soit  $(X_t)_{t \in \mathbb{Z}}$  un processus ARMA( $p, q$ ) stationnaire causal. Soit  $\rho(\cdot)$  son autocorrélation et  $\hat{\rho}(\cdot)$  son autocorrélation empirique calculée sur  $(X_1, \dots, X_n)$ . Alors :

$$\hat{\rho}(k) \xrightarrow[n \rightarrow +\infty]{\mathcal{P}} \rho(k) \quad \text{pour tout } k \in \mathbb{N}$$

**Conséquence** : l'ACF converge vers les autocorrélations théoriques

# Plan du cours

- 1 Définitions et premières propriétés
  - Séries chronologiques
  - Stationnarité
  - Première analyse statistique de la dépendance
- 2 Tendances et saisonnalités
  - Définition et propriétés
  - Estimation semi-paramétrique de la tendance et de la saisonnalité
  - Estimation non-paramétrique de la tendance et de la saisonnalité
- 3 Exemples de modèles de séries stationnaires
  - Processus ARMA
  - Processus GARCH
  - Autres exemples de séries temporelles
- 4 Estimation, sélection de modèle, test et prédiction
  - Estimation semi-paramétrique
  - Sélection de modèles et test d'adéquation
  - Prédiction

# Processus GARCH

Définition (Engle, 1982 et Bollerssev, 1986)

Un processus GARCH( $p, q$ ) où  $(p, q) \in \mathbb{N}^2$  est une série  $(X_t)_{t \in \mathbb{Z}}$  stationnaire telle que pour  $t \in \mathbb{Z}$  :

$$X_t = \varepsilon_t \sigma_t \quad \text{et} \quad \sigma_t^2 = \omega_0 + a_1 X_{t-1}^2 + \cdots + a_p X_{t-p}^2 + b_1 \sigma_{t-1}^2 + \cdots + b_q \sigma_{t-q}^2$$

où :

- $(\varepsilon_t)$  bruit blanc de variance 1 ;
- $\omega_0 > 0$ ,  $(a_1, \dots, a_p, b_1, \dots, b_q) \in [0, \infty)^{p+q}$  avec  $a_p \neq 0$  et  $b_q \neq 0$ .

Cas particulier :

- Si  $q = 0$ , un GARCH( $p, q$ ) est un ARCH( $p$ ) (Conditionnaly Heteroskedastic Auto-Regressive) ;

# Processus GARCH

Définition (Engle, 1982 et Bollerssev, 1986 )

Un processus GARCH( $p, q$ ) où  $(p, q) \in \mathbb{N}^2$  est une série  $(X_t)_{t \in \mathbb{Z}}$  stationnaire telle que pour  $t \in \mathbb{Z}$  :

$$X_t = \varepsilon_t \sigma_t \quad \text{et} \quad \sigma_t^2 = \omega_0 + a_1 X_{t-1}^2 + \cdots + a_p X_{t-p}^2 + b_1 \sigma_{t-1}^2 + \cdots + b_q \sigma_{t-q}^2$$

où :

- $(\varepsilon_t)$  bruit blanc de variance 1 ;
- $\omega_0 > 0$ ,  $(a_1, \dots, a_p, b_1, \dots, b_q) \in [0, \infty)^{p+q}$  avec  $a_p \neq 0$  et  $b_q \neq 0$ .

Cas particulier :

- Si  $q = 0$ , un GARCH( $p, q$ ) est un ARCH( $p$ ) (Conditionnaly Heteroskedastic Auto-Regressive) ;

# Processus GARCH

Définition (Engle, 1982 et Bollerssev, 1986 )

Un processus GARCH( $p, q$ ) où  $(p, q) \in \mathbb{N}^2$  est une série  $(X_t)_{t \in \mathbb{Z}}$  stationnaire telle que pour  $t \in \mathbb{Z}$  :

$$X_t = \varepsilon_t \sigma_t \quad \text{et} \quad \sigma_t^2 = \omega_0 + a_1 X_{t-1}^2 + \cdots + a_p X_{t-p}^2 + b_1 \sigma_{t-1}^2 + \cdots + b_q \sigma_{t-q}^2$$

où :

- $(\varepsilon_t)$  bruit blanc de variance 1 ;
- $\omega_0 > 0$ ,  $(a_1, \dots, a_p, b_1, \dots, b_q) \in [0, \infty)^{p+q}$  avec  $a_p \neq 0$  et  $b_q \neq 0$ .

Cas particulier :

- Si  $q = 0$ , un GARCH( $p, q$ ) est un ARCH( $p$ ) (Conditionnaly Heteroskedastic Auto-Regressive) ;

# Processus GARCH

Définition (Engle, 1982 et Bollerssev, 1986 )

Un processus GARCH( $p, q$ ) où  $(p, q) \in \mathbb{N}^2$  est une série  $(X_t)_{t \in \mathbb{Z}}$  stationnaire telle que pour  $t \in \mathbb{Z}$  :

$$X_t = \varepsilon_t \sigma_t \quad \text{et} \quad \sigma_t^2 = \omega_0 + a_1 X_{t-1}^2 + \cdots + a_p X_{t-p}^2 + b_1 \sigma_{t-1}^2 + \cdots + b_q \sigma_{t-q}^2$$

où :

- $(\varepsilon_t)$  bruit blanc de variance 1 ;
- $\omega_0 > 0$ ,  $(a_1, \dots, a_p, b_1, \dots, b_q) \in [0, \infty)^{p+q}$  avec  $a_p \neq 0$  et  $b_q \neq 0$ .

Cas particulier :

- Si  $q = 0$ , un GARCH( $p, q$ ) est un ARCH( $p$ ) (Conditionnaly Heteroskedastic Auto-Regressive) ;

# Processus GARCH

Définition (Engle, 1982 et Bollerssev, 1986 )

Un processus GARCH( $p, q$ ) où  $(p, q) \in \mathbb{N}^2$  est une série  $(X_t)_{t \in \mathbb{Z}}$  stationnaire telle que pour  $t \in \mathbb{Z}$  :

$$X_t = \varepsilon_t \sigma_t \quad \text{et} \quad \sigma_t^2 = \omega_0 + a_1 X_{t-1}^2 + \dots + a_p X_{t-p}^2 + b_1 \sigma_{t-1}^2 + \dots + b_q \sigma_{t-q}^2$$

où :

- $(\varepsilon_t)$  bruit blanc de variance 1 ;
- $\omega_0 > 0, (a_1, \dots, a_p, b_1, \dots, b_q) \in [0, \infty)^{p+q}$  avec  $a_p \neq 0$  et  $b_q \neq 0$ .

Cas particulier :

- Si  $q = 0$ , un GARCH( $p, q$ ) est un ARCH( $p$ ) (Conditionnaly Heteroskedastic Auto-Regressive) ;

# Processus GARCH

Définition (Engle, 1982 et Bollerssev, 1986 )

Un processus GARCH( $p, q$ ) où  $(p, q) \in \mathbb{N}^2$  est une série  $(X_t)_{t \in \mathbb{Z}}$  stationnaire telle que pour  $t \in \mathbb{Z}$  :

$$X_t = \varepsilon_t \sigma_t \quad \text{et} \quad \sigma_t^2 = \omega_0 + a_1 X_{t-1}^2 + \cdots + a_p X_{t-p}^2 + b_1 \sigma_{t-1}^2 + \cdots + b_q \sigma_{t-q}^2$$

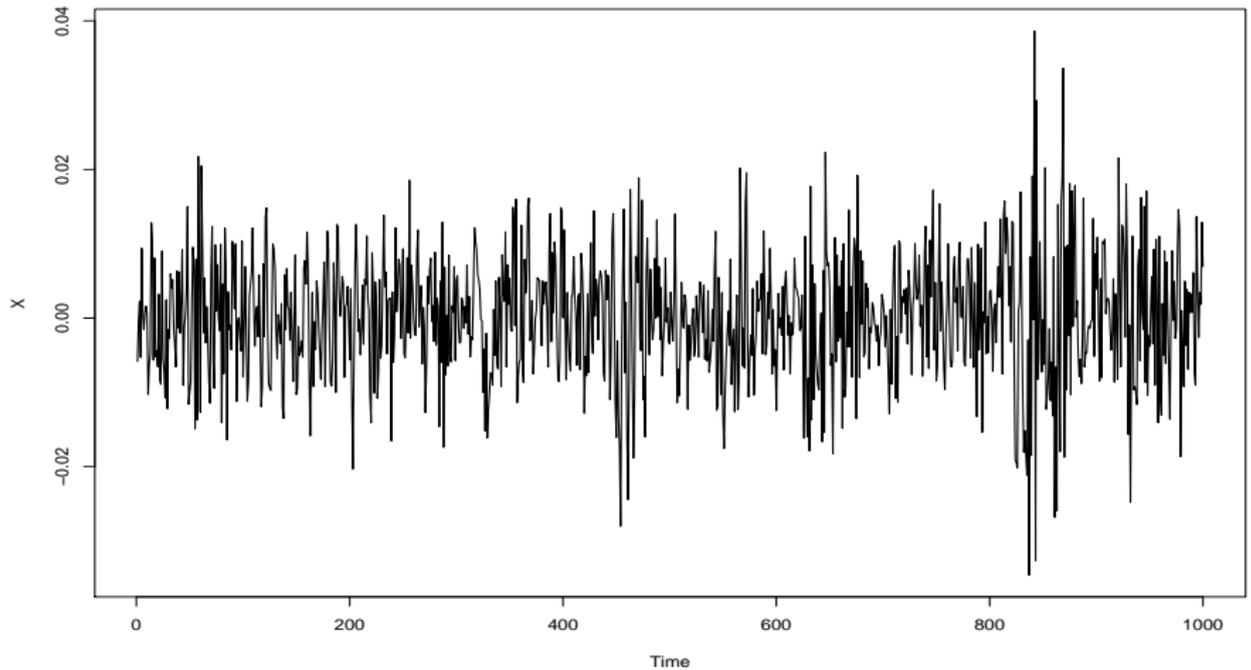
où :

- $(\varepsilon_t)$  bruit blanc de variance 1 ;
- $\omega_0 > 0$ ,  $(a_1, \dots, a_p, b_1, \dots, b_q) \in [0, \infty)^{p+q}$  avec  $a_p \neq 0$  et  $b_q \neq 0$ .

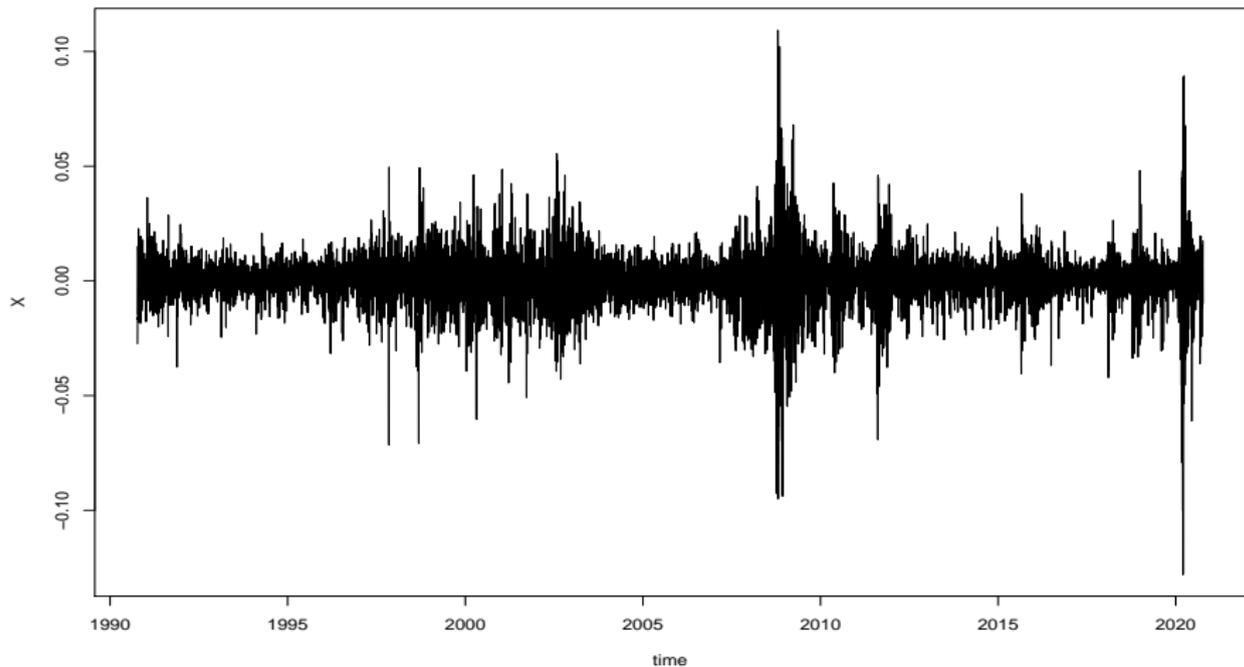
Cas particulier :

- Si  $q = 0$ , un GARCH( $p, q$ ) est un ARCH( $p$ ) (Conditionnaly Heteroskedastic Auto-Regressive) ;

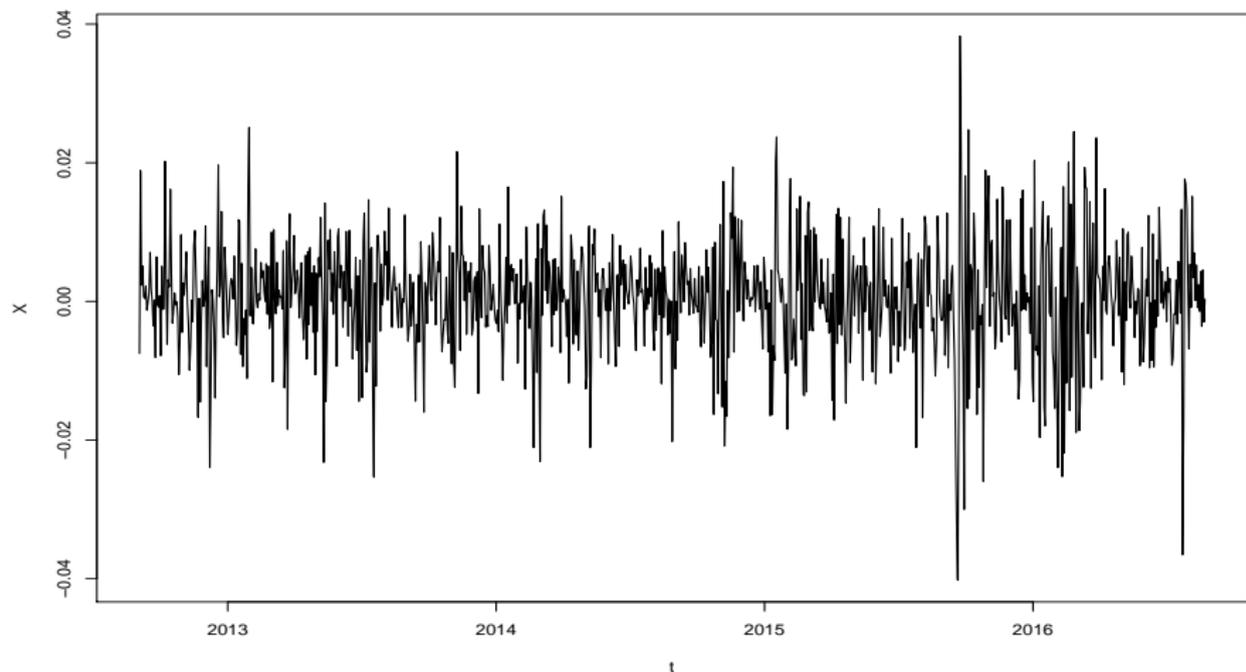
# Trajectoire d'un processus GARCH(1,1)



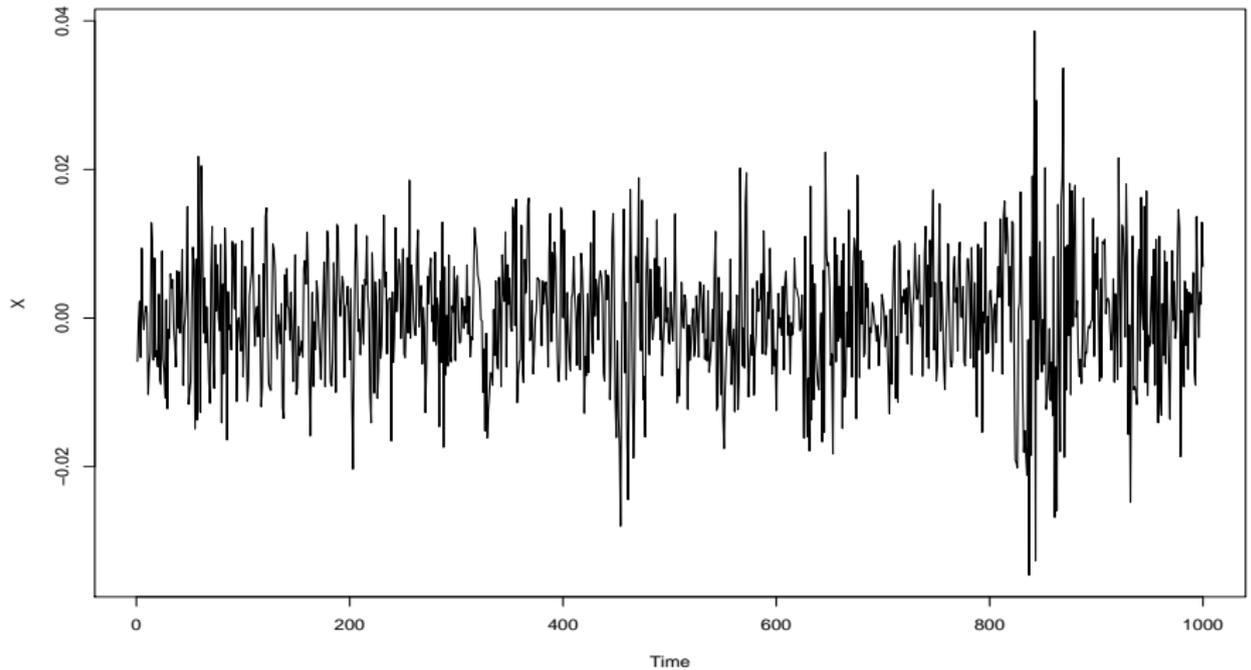
# Trajectoire Log-rendement SP500 d'octobre 1990 à octobre 2020



# Trajectoire Log-rendement SP500 09/2012 -> 09/2016



# Trajectoire d'un processus GARCH(1,1)



# Conditions de stationnarité d'un processus GARCH

## Propriété

Un processus GARCH( $p, q$ ) est stationnaire d'ordre 2 si et seulement si

$$\sum_{i=1}^p a_i + \sum_{j=1}^q b_j < 1.$$

De ceci, on peut déduire les propriétés suivantes :

## Propriété

Soit  $X = (X_k)_{k \in \mathbb{Z}}$  un processus GARCH( $p, q$ ) stationnaire d'ordre 2. Alors

- 1  $\mathbb{E}[X_t] = 0$ ,  $\text{var}(X_t) = \frac{\omega_0}{1 - \sum_{i=1}^p a_i - \sum_{j=1}^q b_j}$  et  $\text{cov}(X_0, X_k) = 0$  si  $k \neq 0$ .
- 2  $\mathbb{E}[X_t | X_{t-1}, X_{t-2}, \dots] = 0$  et  $\text{var}(X_t | X_{t-1}, X_{t-2}, \dots) = \sigma_t^2 \neq C^{te}$  :  
Conditionnellement Hétéroscédastique

# Conditions de stationnarité d'un processus GARCH

## Propriété

Un processus GARCH( $p, q$ ) est stationnaire d'ordre 2 si et seulement si

$$\sum_{i=1}^p a_i + \sum_{j=1}^q b_j < 1.$$

De ceci, on peut déduire les propriétés suivantes :

## Propriété

Soit  $X = (X_k)_{k \in \mathbb{Z}}$  un processus GARCH( $p, q$ ) stationnaire d'ordre 2. Alors

- 1  $\mathbb{E}[X_t] = 0$ ,  $\text{var}(X_t) = \frac{\omega_0}{1 - \sum_{i=1}^p a_i - \sum_{j=1}^q b_j}$  et  $\text{cov}(X_0, X_k) = 0$  si  $k \neq 0$ .
- 2  $\mathbb{E}[X_t | X_{t-1}, X_{t-2}, \dots] = 0$  et  $\text{var}(X_t | X_{t-1}, X_{t-2}, \dots) = \sigma_t^2 \neq C^{te}$  :  
Conditionnellement Hétéroscédastique

# Conditions de stationnarité d'un processus GARCH

## Propriété

Un processus GARCH( $p, q$ ) est stationnaire d'ordre 2 si et seulement si

$$\sum_{i=1}^p a_i + \sum_{j=1}^q b_j < 1.$$

De ceci, on peut déduire les propriétés suivantes :

## Propriété

Soit  $X = (X_k)_{k \in \mathbb{Z}}$  un processus GARCH( $p, q$ ) stationnaire d'ordre 2. Alors

- 1  $\mathbb{E}[X_t] = 0$ ,  $\text{var}(X_t) = \frac{\omega_0}{1 - \sum_{i=1}^p a_i - \sum_{j=1}^q b_j}$  et  $\text{cov}(X_0, X_k) = 0$  si  $k \neq 0$ .
- 2  $\mathbb{E}[X_t | X_{t-1}, X_{t-2}, \dots] = 0$  et  $\text{var}(X_t | X_{t-1}, X_{t-2}, \dots) = \sigma_t^2 \neq C^{te}$  :  
Conditionnellement Hétéroscédastique

# Conditions de stationnarité d'un processus GARCH

## Propriété

Un processus GARCH( $p, q$ ) est stationnaire d'ordre 2 si et seulement si

$$\sum_{i=1}^p a_i + \sum_{j=1}^q b_j < 1.$$

De ceci, on peut déduire les propriétés suivantes :

## Propriété

Soit  $X = (X_k)_{k \in \mathbb{Z}}$  un processus GARCH( $p, q$ ) stationnaire d'ordre 2. Alors

①  $\mathbb{E}[X_t] = 0$ ,  $\text{var}(X_t) = \frac{\omega_0}{1 - \sum_{i=1}^p a_i - \sum_{j=1}^q b_j}$  et  $\text{cov}(X_0, X_k) = 0$  si  $k \neq 0$ .

②  $\mathbb{E}[X_t | X_{t-1}, X_{t-2}, \dots] = 0$  et  $\text{var}(X_t | X_{t-1}, X_{t-2}, \dots) = \sigma_t^2 \neq C^{te}$  :  
Conditionnellement Hétéroscédastique

# Conditions de stationnarité d'un processus GARCH

## Propriété

Un processus GARCH( $p, q$ ) est stationnaire d'ordre 2 si et seulement si

$$\sum_{i=1}^p a_i + \sum_{j=1}^q b_j < 1.$$

De ceci, on peut déduire les propriétés suivantes :

## Propriété

Soit  $X = (X_k)_{k \in \mathbb{Z}}$  un processus GARCH( $p, q$ ) stationnaire d'ordre 2. Alors

①  $\mathbb{E}[X_t] = 0$ ,  $\text{var}(X_t) = \frac{\omega_0}{1 - \sum_{i=1}^p a_i - \sum_{j=1}^q b_j}$  et  $\text{cov}(X_0, X_k) = 0$  si  $k \neq 0$ .

②  $\mathbb{E}[X_t | X_{t-1}, X_{t-2}, \dots] = 0$  et  $\text{var}(X_t | X_{t-1}, X_{t-2}, \dots) = \sigma_t^2 \neq C^{te}$  :

*Conditionnellement Hétéroscédastique*

# Conditions de stationnarité d'un processus GARCH

## Propriété

Un processus GARCH( $p, q$ ) est stationnaire d'ordre 2 si et seulement si

$$\sum_{i=1}^p a_i + \sum_{j=1}^q b_j < 1.$$

De ceci, on peut déduire les propriétés suivantes :

## Propriété

Soit  $X = (X_k)_{k \in \mathbb{Z}}$  un processus GARCH( $p, q$ ) stationnaire d'ordre 2. Alors

- 1  $\mathbb{E}[X_t] = 0$ ,  $\text{var}(X_t) = \frac{\omega_0}{1 - \sum_{i=1}^p a_i - \sum_{j=1}^q b_j}$  et  $\text{cov}(X_0, X_k) = 0$  si  $k \neq 0$ .
- 2  $\mathbb{E}[X_t | X_{t-1}, X_{t-2}, \dots] = 0$  et  $\text{var}(X_t | X_{t-1}, X_{t-2}, \dots) = \sigma_t^2 \neq C^{te}$  :  
Conditionnellement Hétéroscédastique

# Propriétés d'un processus GARCH

## Propriété

Soit  $X = (X_k)_{k \in \mathbb{Z}}$  un processus GARCH( $p, q$ ) stationnaire d'ordre 2 tel que  $\mathbb{E}(\varepsilon^4) < \infty$ . Alors  $Y = (X_k^2)_{k \in \mathbb{Z}}$  est un processus ARMA( $\max(p, q), p$ ) faible non centré.

## Démonstration.

On considère  $v_t = X_t^2 - \sigma_t^2 = (\varepsilon_t^2 - 1)\sigma_t^2$ . La série  $(v_t)$  est un bruit blanc faible car  $(v_t)$  est stationnaire,  $\mathbb{E}(v_t) = 0$  et pour  $t > 0$ ,  $\text{cov}(v_0, v_t) = \text{cov}((\varepsilon_0^2 - 1)\sigma_0^2, (\varepsilon_t^2 - 1)\sigma_t^2) = 0$  car  $(\varepsilon_t^2 - 1)$  est indépendant des 3 autres termes et d'espérance nulle. De plus, on montre que  $X_t^2 = v_t + (a_0 + \sum_{i=1}^p a_i X_{t-i}^2) + \sum_{j=1}^q b_j (X_{t-j}^2 - v_{t-j}) = a_0 + \sum_{i=1}^p a_i X_{t-i}^2 + \sum_{j=1}^q b_j (X_{t-j}^2 + v_t - \sum_{j=1}^q v_{t-j})$ , d'où  $(X_t^2)$  ARMA faible non centré.  $\square$

Conséquences : • Processus GARCH : bruit blanc faible mais pas fort

• Un processus GARCH ne peut jamais être un processus gaussien

# Propriétés d'un processus GARCH

## Propriété

Soit  $X = (X_k)_{k \in \mathbb{Z}}$  un processus GARCH( $p, q$ ) stationnaire d'ordre 2 tel que  $\mathbb{E}(\varepsilon^4) < \infty$ . Alors  $Y = (X_k^2)_{k \in \mathbb{Z}}$  est un processus ARMA( $\max(p, q), p$ ) faible non centré.

## Démonstration.

On considère  $v_t = X_t^2 - \sigma_t^2 = (\varepsilon_t^2 - 1)\sigma_t^2$ . La série  $(v_t)$  est un bruit blanc faible car  $(v_t)$  est stationnaire,  $\mathbb{E}(v_t) = 0$  et pour  $t > 0$ ,  $\text{cov}(v_0, v_t) = \text{cov}((\varepsilon_0^2 - 1)\sigma_0^2, (\varepsilon_t^2 - 1)\sigma_t^2) = 0$  car  $(\varepsilon_t^2 - 1)$  est indépendant des 3 autres termes et d'espérance nulle. De plus, on montre que

$$X_t^2 = v_t + (a_0 + \sum_{i=1}^p a_i X_{t-i}^2) + \sum_{j=1}^q b_j (X_{t-j}^2 - v_{t-j}) =$$
$$a_0 + \sum_{i=1}^p a_i X_{t-i}^2 + \sum_{j=1}^q b_j (X_{t-j}^2 + v_t - \sum_{j=1}^q v_{t-j}), \text{ d'où } (X_t^2) \text{ ARMA faible non centré. } \square$$

Conséquences : • Processus GARCH : bruit blanc faible mais pas fort

• Un processus GARCH ne peut jamais être un processus gaussien

# Propriétés d'un processus GARCH

## Propriété

Soit  $X = (X_k)_{k \in \mathbb{Z}}$  un processus GARCH( $p, q$ ) stationnaire d'ordre 2 tel que  $\mathbb{E}(\varepsilon^4) < \infty$ . Alors  $Y = (X_k^2)_{k \in \mathbb{Z}}$  est un processus ARMA( $\max(p, q), p$ ) faible non centré.

## Démonstration.

On considère  $v_t = X_t^2 - \sigma_t^2 = (\varepsilon_t^2 - 1)\sigma_t^2$ . La série  $(v_t)$  est un bruit blanc faible car  $(v_t)$  est stationnaire,  $\mathbb{E}(v_t) = 0$  et pour  $t > 0$ ,  $\text{cov}(v_0, v_t) = \text{cov}((\varepsilon_0^2 - 1)\sigma_0^2, (\varepsilon_t^2 - 1)\sigma_t^2) = 0$  car  $(\varepsilon_t^2 - 1)$  est indépendant des 3 autres termes et d'espérance nulle. De plus, on montre que

$$X_t^2 = v_t + (a_0 + \sum_{i=1}^p a_i X_{t-i}^2) + \sum_{j=1}^q b_j (X_{t-j}^2 - v_{t-j}) =$$
$$a_0 + \sum_{i=1}^p a_i X_{t-i}^2 + \sum_{j=1}^q b_j (X_{t-j}^2 + v_t - \sum_{j=1}^q v_{t-j}), \text{ d'où } (X_t^2) \text{ ARMA faible non centré. } \square$$

**Conséquences :** • Processus GARCH : bruit blanc faible mais pas fort

• Un processus GARCH ne peut jamais être un processus gaussien

# Propriétés d'un processus GARCH

## Propriété

Soit  $X = (X_k)_{k \in \mathbb{Z}}$  un processus GARCH( $p, q$ ) stationnaire d'ordre 2 tel que  $\mathbb{E}(\varepsilon^4) < \infty$ . Alors  $Y = (X_k^2)_{k \in \mathbb{Z}}$  est un processus ARMA( $\max(p, q), p$ ) faible non centré.

## Démonstration.

On considère  $v_t = X_t^2 - \sigma_t^2 = (\varepsilon_t^2 - 1)\sigma_t^2$ . La série  $(v_t)$  est un bruit blanc faible car  $(v_t)$  est stationnaire,  $\mathbb{E}(v_t) = 0$  et pour  $t > 0$ ,  $\text{cov}(v_0, v_t) = \text{cov}((\varepsilon_0^2 - 1)\sigma_0^2, (\varepsilon_t^2 - 1)\sigma_t^2) = 0$  car  $(\varepsilon_t^2 - 1)$  est indépendant des 3 autres termes et d'espérance nulle. De plus, on montre que

$$X_t^2 = v_t + (a_0 + \sum_{i=1}^p a_i X_{t-i}^2) + \sum_{j=1}^q b_j (X_{t-j}^2 - v_{t-j}) =$$
$$a_0 + \sum_{i=1}^p a_i X_{t-i}^2 + \sum_{j=1}^q b_j (X_{t-j}^2 + v_t - \sum_{j=1}^q v_{t-j}), \text{ d'où } (X_t^2) \text{ ARMA faible non centré. } \square$$

**Conséquences :** • Processus GARCH : bruit blanc faible mais pas fort

• Un processus GARCH ne peut jamais être un processus gaussien

# Plan du cours

- 1 Définitions et premières propriétés
  - Séries chronologiques
  - Stationnarité
  - Première analyse statistique de la dépendance
- 2 Tendances et saisonnalités
  - Définition et propriétés
  - Estimation semi-paramétrique de la tendance et de la saisonnalité
  - Estimation non-paramétrique de la tendance et de la saisonnalité
- 3 Exemples de modèles de séries stationnaires
  - Processus ARMA
  - Processus GARCH
  - Autres exemples de séries temporelles
- 4 Estimation, sélection de modèle, test et prédiction
  - Estimation semi-paramétrique
  - Sélection de modèles et test d'adéquation
  - Prédiction

# Processus ARMA-GARCH [Ding, Granger et Engle, 1993]

## Définition

$(X_k)_{k \in \mathbb{Z}}$  est un processus ARMA( $p, q$ )-GARCH( $p', q'$ ) si

$$X_t + a_1 X_{t-1} + \dots + a_p X_{t-p} = \varepsilon_t + b_1 \varepsilon_{t-1} + \dots + b_q \varepsilon_{t-q}$$

où  $\varepsilon_t = \xi_t \sigma_t$  et  $\sigma_t^2 = \omega_0 + c_1 \varepsilon_{t-1}^2 + \dots + c_{p'} \varepsilon_{t-p'}^2 + d_1 \sigma_{t-1}^2 + \dots + d_{q'} \sigma_{t-q'}^2$

avec  $(\xi_t)$  un bruit blanc de variance 1,  $\omega_0 > 0$ ,  $a_p, b_q, c_{p'}, d_{q'}$  non nuls.

**Remarque :** • On doit avoir  $P(z) \neq 0$  pour  $|z| \leq 1$  où

$$P(x) = 1 + a_1 x + \dots + a_p x^p \text{ et } \sum_{i=1}^{p'} c_i + \sum_{j=1}^{q'} d_j < 1$$

- ARMA-GARCH=ARMA avec bruit GARCH : même ACF qu'un ARMA
- ARMA-GARCH : espérance et variance conditionnelles non constantes

# Processus ARMA-GARCH [Ding, Granger et Engle, 1993]

## Définition

$(X_k)_{k \in \mathbb{Z}}$  est un processus ARMA( $p, q$ )-GARCH( $p', q'$ ) si

$$X_t + a_1 X_{t-1} + \dots + a_p X_{t-p} = \varepsilon_t + b_1 \varepsilon_{t-1} + \dots + b_q \varepsilon_{t-q}$$

où  $\varepsilon_t = \xi_t \sigma_t$  et  $\sigma_t^2 = \omega_0 + c_1 \varepsilon_{t-1}^2 + \dots + c_{p'} \varepsilon_{t-p'}^2 + d_1 \sigma_{t-1}^2 + \dots + d_{q'} \sigma_{t-q'}^2$

avec  $(\xi_t)$  un bruit blanc de variance 1,  $\omega_0 > 0$ ,  $a_p, b_q, c_{p'}, d_{q'}$  non nuls.

**Remarque :** • On doit avoir  $P(z) \neq 0$  pour  $|z| \leq 1$  où  
 $P(x) = 1 + a_1 x + \dots + a_p x^p$  et  $\sum_{i=1}^{p'} c_i + \sum_{j=1}^{q'} d_j < 1$

- ARMA-GARCH=ARMA avec bruit GARCH : même ACF qu'un ARMA
- ARMA-GARCH : espérance et variance conditionnelles non constantes

## Définition

$(X_k)_{k \in \mathbb{Z}}$  est un processus ARMA( $p, q$ )-GARCH( $p', q'$ ) si

$$X_t + a_1 X_{t-1} + \dots + a_p X_{t-p} = \varepsilon_t + b_1 \varepsilon_{t-1} + \dots + b_q \varepsilon_{t-q}$$

où  $\varepsilon_t = \xi_t \sigma_t$  et  $\sigma_t^2 = \omega_0 + c_1 \varepsilon_{t-1}^2 + \dots + c_{p'} \varepsilon_{t-p'}^2 + d_1 \sigma_{t-1}^2 + \dots + d_{q'} \sigma_{t-q'}^2$

avec  $(\xi_t)$  un bruit blanc de variance 1,  $\omega_0 > 0$ ,  $a_p, b_q, c_{p'}, d_{q'}$  non nuls.

Remarque : • On doit avoir  $P(z) \neq 0$  pour  $|z| \leq 1$  où  
 $P(x) = 1 + a_1 x + \dots + a_p x^p$  et  $\sum_{i=1}^{p'} c_i + \sum_{j=1}^{q'} d_j < 1$

- ARMA-GARCH=ARMA avec bruit GARCH : même ACF qu'un ARMA
- ARMA-GARCH : espérance et variance conditionnelles non constantes

## Définition

$(X_k)_{k \in \mathbb{Z}}$  est un processus ARMA( $p, q$ )-GARCH( $p', q'$ ) si

$$X_t + a_1 X_{t-1} + \dots + a_p X_{t-p} = \varepsilon_t + b_1 \varepsilon_{t-1} + \dots + b_q \varepsilon_{t-q}$$

où  $\varepsilon_t = \xi_t \sigma_t$  et  $\sigma_t^2 = \omega_0 + c_1 \varepsilon_{t-1}^2 + \dots + c_{p'} \varepsilon_{t-p'}^2 + d_1 \sigma_{t-1}^2 + \dots + d_{q'} \sigma_{t-q'}^2$

avec  $(\xi_t)$  un bruit blanc de variance 1,  $\omega_0 > 0$ ,  $a_p, b_q, c_{p'}, d_{q'}$  non nuls.

**Remarque :** • On doit avoir  $P(z) \neq 0$  pour  $|z| \leq 1$  où

$$P(x) = 1 + a_1 x + \dots + a_p x^p \text{ et } \sum_{i=1}^{p'} c_i + \sum_{j=1}^{q'} d_j < 1$$

- ARMA-GARCH=ARMA avec bruit GARCH : même ACF qu'un ARMA
- ARMA-GARCH : espérance et variance conditionnelles non constantes

## Définition

$(X_k)_{k \in \mathbb{Z}}$  est un processus ARMA( $p, q$ )-GARCH( $p', q'$ ) si

$$X_t + a_1 X_{t-1} + \dots + a_p X_{t-p} = \varepsilon_t + b_1 \varepsilon_{t-1} + \dots + b_q \varepsilon_{t-q}$$

où  $\varepsilon_t = \xi_t \sigma_t$  et  $\sigma_t^2 = \omega_0 + c_1 \varepsilon_{t-1}^2 + \dots + c_{p'} \varepsilon_{t-p'}^2 + d_1 \sigma_{t-1}^2 + \dots + d_{q'} \sigma_{t-q'}^2$

avec  $(\xi_t)$  un bruit blanc de variance 1,  $\omega_0 > 0$ ,  $a_p, b_q, c_{p'}, d_{q'}$  non nuls.

**Remarque :** • On doit avoir  $P(z) \neq 0$  pour  $|z| \leq 1$  où

$$P(x) = 1 + a_1 x + \dots + a_p x^p \text{ et } \sum_{i=1}^{p'} c_i + \sum_{j=1}^{q'} d_j < 1$$

- ARMA-GARCH=ARMA avec bruit GARCH : même ACF qu'un ARMA
- ARMA-GARCH : espérance et variance conditionnelles non constantes

## Définition

$(X_k)_{k \in \mathbb{Z}}$  est un processus ARMA( $p, q$ )-GARCH( $p', q'$ ) si

$$X_t + a_1 X_{t-1} + \dots + a_p X_{t-p} = \varepsilon_t + b_1 \varepsilon_{t-1} + \dots + b_q \varepsilon_{t-q}$$

où  $\varepsilon_t = \xi_t \sigma_t$  et  $\sigma_t^2 = \omega_0 + c_1 \varepsilon_{t-1}^2 + \dots + c_{p'} \varepsilon_{t-p'}^2 + d_1 \sigma_{t-1}^2 + \dots + d_{q'} \sigma_{t-q'}^2$

avec  $(\xi_t)$  un bruit blanc de variance 1,  $\omega_0 > 0$ ,  $a_p, b_q, c_{p'}, d_{q'}$  non nuls.

**Remarque :** • On doit avoir  $P(z) \neq 0$  pour  $|z| \leq 1$  où

$$P(x) = 1 + a_1 x + \dots + a_p x^p \text{ et } \sum_{i=1}^{p'} c_i + \sum_{j=1}^{q'} d_j < 1$$

- ARMA-GARCH=ARMA avec bruit GARCH : même ACF qu'un ARMA
- ARMA-GARCH : espérance et variance conditionnelles non constantes

# Processus APARCH [Ding, Granger et Engle, 1993]

## Définition

$(X_k)_{k \in \mathbb{Z}}$  est un processus APARCH( $\delta, p, q$ ) si  $X_t = \varepsilon_t \sigma_t$  et

$$\sigma_t^\delta = \omega_0 + a_1(|X_{t-1}| - \gamma_1 X_{t-1})^\delta + \dots + a_p(|X_{t-p}| - \gamma_p X_{t-p})^\delta \\ + b_1 \sigma_{t-1}^\delta + \dots + b_q \sigma_{t-q}^\delta$$

où :

- $(\varepsilon_t)$  bruit blanc de variance 1 ;
- $\omega_0 > 0, \delta > 0, a_i \geq 0, b_j \geq 0, \gamma_i \in [-1, 1], a_p \neq 0$  et  $b_q \neq 0$

**Remarque :** • Condition de stationnarité compliquée...

- Bruit blanc faible :  $r_X(k) = 0$  pour  $|k| \geq 1$ .
- Variance conditionnelle non constante et dépendant du signe des valeurs

## Définition

$(X_k)_{k \in \mathbb{Z}}$  est un processus APARCH( $\delta, p, q$ ) si  $X_t = \varepsilon_t \sigma_t$  et

$$\sigma_t^\delta = \omega_0 + a_1(|X_{t-1}| - \gamma_1 X_{t-1})^\delta + \dots + a_p(|X_{t-p}| - \gamma_p X_{t-p})^\delta + b_1 \sigma_{t-1}^\delta + \dots + b_q \sigma_{t-q}^\delta$$

où :

- $(\varepsilon_t)$  bruit blanc de variance 1 ;
- $\omega_0 > 0, \delta > 0, a_i \geq 0, b_j \geq 0, \gamma_i \in [-1, 1], a_p \neq 0$  et  $b_q \neq 0$

Remarque : • Condition de stationnarité compliquée...

- Bruit blanc faible :  $r_X(k) = 0$  pour  $|k| \geq 1$ .
- Variance conditionnelle non constante et dépendant du signe des valeurs

# Processus APARCH [Ding, Granger et Engle, 1993]

## Définition

$(X_k)_{k \in \mathbb{Z}}$  est un processus APARCH( $\delta, p, q$ ) si  $X_t = \varepsilon_t \sigma_t$  et

$$\sigma_t^\delta = \omega_0 + a_1(|X_{t-1}| - \gamma_1 X_{t-1})^\delta + \dots + a_p(|X_{t-p}| - \gamma_p X_{t-p})^\delta \\ + b_1 \sigma_{t-1}^\delta + \dots + b_q \sigma_{t-q}^\delta$$

où :

- $(\varepsilon_t)$  bruit blanc de variance 1 ;
- $\omega_0 > 0, \delta > 0, a_i \geq 0, b_j \geq 0, \gamma_i \in [-1, 1], a_p \neq 0$  et  $b_q \neq 0$

Remarque : • Condition de stationnarité compliquée...

- Bruit blanc faible :  $r_X(k) = 0$  pour  $|k| \geq 1$ .
- Variance conditionnelle non constante et dépendant du signe des valeurs

## Définition

$(X_k)_{k \in \mathbb{Z}}$  est un processus APARCH( $\delta, p, q$ ) si  $X_t = \varepsilon_t \sigma_t$  et

$$\sigma_t^\delta = \omega_0 + a_1(|X_{t-1}| - \gamma_1 X_{t-1})^\delta + \dots + a_p(|X_{t-p}| - \gamma_p X_{t-p})^\delta \\ + b_1 \sigma_{t-1}^\delta + \dots + b_q \sigma_{t-q}^\delta$$

où :

- $(\varepsilon_t)$  bruit blanc de variance 1 ;
- $\omega_0 > 0, \delta > 0, a_i \geq 0, b_j \geq 0, \gamma_i \in [-1, 1], a_p \neq 0$  et  $b_q \neq 0$

Remarque : • Condition de stationnarité compliquée...

- Bruit blanc faible :  $r_X(k) = 0$  pour  $|k| \geq 1$ .
- Variance conditionnelle non constante et dépendant du signe des valeurs

## Définition

$(X_k)_{k \in \mathbb{Z}}$  est un processus APARCH( $\delta, p, q$ ) si  $X_t = \varepsilon_t \sigma_t$  et

$$\sigma_t^\delta = \omega_0 + a_1(|X_{t-1}| - \gamma_1 X_{t-1})^\delta + \dots + a_p(|X_{t-p}| - \gamma_p X_{t-p})^\delta \\ + b_1 \sigma_{t-1}^\delta + \dots + b_q \sigma_{t-q}^\delta$$

où :

- $(\varepsilon_t)$  bruit blanc de variance 1 ;
- $\omega_0 > 0, \delta > 0, a_i \geq 0, b_j \geq 0, \gamma_i \in [-1, 1], a_p \neq 0$  et  $b_q \neq 0$

Remarque : • Condition de stationnarité compliquée...

• Bruit blanc faible :  $r_X(k) = 0$  pour  $|k| \geq 1$ .

• Variance conditionnelle non constante et dépendant du signe des valeurs

## Définition

$(X_k)_{k \in \mathbb{Z}}$  est un processus APARCH( $\delta, p, q$ ) si  $X_t = \varepsilon_t \sigma_t$  et

$$\sigma_t^\delta = \omega_0 + a_1(|X_{t-1}| - \gamma_1 X_{t-1})^\delta + \dots + a_p(|X_{t-p}| - \gamma_p X_{t-p})^\delta \\ + b_1 \sigma_{t-1}^\delta + \dots + b_q \sigma_{t-q}^\delta$$

où :

- $(\varepsilon_t)$  bruit blanc de variance 1 ;
- $\omega_0 > 0, \delta > 0, a_i \geq 0, b_j \geq 0, \gamma_i \in [-1, 1], a_p \neq 0$  et  $b_q \neq 0$

**Remarque :** • Condition de stationnarité compliquée...

- Bruit blanc faible :  $r_X(k) = 0$  pour  $|k| \geq 1$ .
- Variance conditionnelle non constante et dépendant du signe des valeurs

## Définition

$(X_k)_{k \in \mathbb{Z}}$  est un processus APARCH( $\delta, p, q$ ) si  $X_t = \varepsilon_t \sigma_t$  et

$$\sigma_t^\delta = \omega_0 + a_1(|X_{t-1}| - \gamma_1 X_{t-1})^\delta + \dots + a_p(|X_{t-p}| - \gamma_p X_{t-p})^\delta + b_1 \sigma_{t-1}^\delta + \dots + b_q \sigma_{t-q}^\delta$$

où :

- $(\varepsilon_t)$  bruit blanc de variance 1 ;
- $\omega_0 > 0, \delta > 0, a_i \geq 0, b_j \geq 0, \gamma_i \in [-1, 1], a_p \neq 0$  et  $b_q \neq 0$

**Remarque :** • Condition de stationnarité compliquée...

- Bruit blanc faible :  $r_X(k) = 0$  pour  $|k| \geq 1$ .

- Variance conditionnelle non constante et dépendant du signe des valeurs

## Définition

$(X_k)_{k \in \mathbb{Z}}$  est un processus APARCH( $\delta, p, q$ ) si  $X_t = \varepsilon_t \sigma_t$  et

$$\sigma_t^\delta = \omega_0 + a_1(|X_{t-1}| - \gamma_1 X_{t-1})^\delta + \cdots + a_p(|X_{t-p}| - \gamma_p X_{t-p})^\delta \\ + b_1 \sigma_{t-1}^\delta + \cdots + b_q \sigma_{t-q}^\delta$$

où :

- $(\varepsilon_t)$  bruit blanc de variance 1 ;
- $\omega_0 > 0, \delta > 0, a_i \geq 0, b_j \geq 0, \gamma_i \in [-1, 1], a_p \neq 0$  et  $b_q \neq 0$

**Remarque :** • Condition de stationnarité compliquée...

- Bruit blanc faible :  $r_X(k) = 0$  pour  $|k| \geq 1$ .
- Variance conditionnelle non constante et dépendant du signe des valeurs

# Processus FARIMA [Granger et Joyeux, 1980]

## Définition

$(X_k)_{k \in \mathbb{Z}}$  est un processus FARIMA( $p, d, q$ ) où  $-0.5 \leq d \leq 0.5$  et

$$X_t + a_1 X_{t-1} + \dots + a_p X_{t-p} = \varepsilon_t + b_1 \varepsilon_{t-1} + \dots + b_q \varepsilon_{t-q}$$

$$\text{où } \varepsilon_t = \xi_t + \sum_{j=1}^{\infty} \left( \frac{\Gamma(j+d)}{\Gamma(d)\Gamma(j+1)} \right) \xi_{t-j}$$

avec  $(\xi_t)$  bruit blanc de variance  $\sigma_\xi^2$ ,  $a_p, b_q$  non nuls,  $\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt$ .

Commentaire : • Avec  $P(x) = 1 + a_1 x + \dots + a_p x^p$ ,  $P(z) \neq 0$  pour  $|z| \leq 1$

• Si  $0 < d < 0.5$ ,  $r_X(k) \simeq C |k|^{2d-1}$  pour  $k \rightarrow \infty$  : processus longue mémoire

• On a  $\hat{\rho}(\cdot) \xrightarrow[n \rightarrow +\infty]{\mathcal{P}} \rho(\cdot)$  : ACF estime  $\rho$

# Processus FARIMA [Granger et Joyeux, 1980]

## Définition

$(X_k)_{k \in \mathbb{Z}}$  est un processus FARIMA( $p, d, q$ ) où  $-0.5 \leq d \leq 0.5$  et

$$X_t + a_1 X_{t-1} + \dots + a_p X_{t-p} = \varepsilon_t + b_1 \varepsilon_{t-1} + \dots + b_q \varepsilon_{t-q}$$

$$\text{où } \varepsilon_t = \xi_t + \sum_{j=1}^{\infty} \left( \frac{\Gamma(j+d)}{\Gamma(d)\Gamma(j+1)} \right) \xi_{t-j}$$

avec  $(\xi_t)$  bruit blanc de variance  $\sigma_\xi^2$ ,  $a_p, b_q$  non nuls,  $\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt$ .

Commentaire : • Avec  $P(x) = 1 + a_1 x + \dots + a_p x^p$ ,  $P(z) \neq 0$  pour  $|z| \leq 1$

• Si  $0 < d < 0.5$ ,  $r_X(k) \simeq C |k|^{2d-1}$  pour  $k \rightarrow \infty$  : processus longue mémoire

• On a  $\hat{\rho}(\cdot) \xrightarrow[n \rightarrow +\infty]{\mathcal{P}} \rho(\cdot)$  : ACF estime  $\rho$

# Processus FARIMA [Granger et Joyeux, 1980]

## Définition

$(X_k)_{k \in \mathbb{Z}}$  est un processus FARIMA( $p, d, q$ ) où  $-0.5 \leq d \leq 0.5$  et

$$X_t + a_1 X_{t-1} + \dots + a_p X_{t-p} = \varepsilon_t + b_1 \varepsilon_{t-1} + \dots + b_q \varepsilon_{t-q}$$

$$\text{où } \varepsilon_t = \xi_t + \sum_{j=1}^{\infty} \left( \frac{\Gamma(j+d)}{\Gamma(d)\Gamma(j+1)} \right) \xi_{t-j}$$

avec  $(\xi_t)$  bruit blanc de variance  $\sigma_\xi^2$ ,  $a_p, b_q$  non nuls,  $\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt$ .

Commentaire : • Avec  $P(x) = 1 + a_1 x + \dots + a_p x^p$ ,  $P(z) \neq 0$  pour  $|z| \leq 1$

• Si  $0 < d < 0.5$ ,  $r_X(k) \simeq C |k|^{2d-1}$  pour  $k \rightarrow \infty$  : processus longue mémoire

• On a  $\hat{\rho}(\cdot) \xrightarrow[n \rightarrow +\infty]{\mathcal{P}} \rho(\cdot)$  : ACF estime  $\rho$

# Processus FARIMA [Granger et Joyeux, 1980]

## Définition

$(X_k)_{k \in \mathbb{Z}}$  est un processus FARIMA( $p, d, q$ ) où  $-0.5 \leq d \leq 0.5$  et

$$X_t + a_1 X_{t-1} + \dots + a_p X_{t-p} = \varepsilon_t + b_1 \varepsilon_{t-1} + \dots + b_q \varepsilon_{t-q}$$

$$\text{où } \varepsilon_t = \xi_t + \sum_{j=1}^{\infty} \left( \frac{\Gamma(j+d)}{\Gamma(d)\Gamma(j+1)} \right) \xi_{t-j}$$

avec  $(\xi_t)$  bruit blanc de variance  $\sigma_{\xi}^2$ ,  $a_p, b_q$  non nuls,  $\Gamma(x) = \int_0^{\infty} t^{x-1} e^{-t} dt$ .

Commentaire : • Avec  $P(x) = 1 + a_1 x + \dots + a_p x^p$ ,  $P(z) \neq 0$  pour  $|z| \leq 1$

• Si  $0 < d < 0.5$ ,  $r_X(k) \simeq C |k|^{2d-1}$  pour  $k \rightarrow \infty$  : processus longue mémoire

• On a  $\hat{\rho}(\cdot) \xrightarrow[n \rightarrow +\infty]{\mathcal{P}} \rho(\cdot)$  : ACF estime  $\rho$

# Processus FARIMA [Granger et Joyeux, 1980]

## Définition

$(X_k)_{k \in \mathbb{Z}}$  est un processus FARIMA( $p, d, q$ ) où  $-0.5 \leq d \leq 0.5$  et

$$X_t + a_1 X_{t-1} + \dots + a_p X_{t-p} = \varepsilon_t + b_1 \varepsilon_{t-1} + \dots + b_q \varepsilon_{t-q}$$

$$\text{où } \varepsilon_t = \xi_t + \sum_{j=1}^{\infty} \left( \frac{\Gamma(j+d)}{\Gamma(d)\Gamma(j+1)} \right) \xi_{t-j}$$

avec  $(\xi_t)$  bruit blanc de variance  $\sigma_\xi^2$ ,  $a_p, b_q$  non nuls,  $\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt$ .

Commentaire : • Avec  $P(x) = 1 + a_1 x + \dots + a_p x^p$ ,  $P(z) \neq 0$  pour  $|z| \leq 1$

• Si  $0 < d < 0.5$ ,  $r_X(k) \simeq C |k|^{2d-1}$  pour  $k \rightarrow \infty$  : processus longue mémoire

• On a  $\hat{\rho}(\cdot) \xrightarrow[n \rightarrow +\infty]{\mathcal{P}} \rho(\cdot)$  : ACF estime  $\rho$

# Processus FARIMA [Granger et Joyeux, 1980]

## Définition

$(X_k)_{k \in \mathbb{Z}}$  est un processus FARIMA( $p, d, q$ ) où  $-0.5 \leq d \leq 0.5$  et

$$X_t + a_1 X_{t-1} + \dots + a_p X_{t-p} = \varepsilon_t + b_1 \varepsilon_{t-1} + \dots + b_q \varepsilon_{t-q}$$

$$\text{où } \varepsilon_t = \xi_t + \sum_{j=1}^{\infty} \left( \frac{\Gamma(j+d)}{\Gamma(d)\Gamma(j+1)} \right) \xi_{t-j}$$

avec  $(\xi_t)$  bruit blanc de variance  $\sigma_\xi^2$ ,  $a_p, b_q$  non nuls,  $\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt$ .

**Commentaire :** • Avec  $P(x) = 1 + a_1 x + \dots + a_p x^p$ ,  $P(z) \neq 0$  pour  $|z| \leq 1$

• Si  $0 < d < 0.5$ ,  $r_X(k) \simeq C |k|^{2d-1}$  pour  $k \rightarrow \infty$  : processus longue mémoire

• On a  $\hat{\rho}(\cdot) \xrightarrow[n \rightarrow +\infty]{\mathcal{P}} \rho(\cdot)$  : ACF estime  $\rho$

# Processus FARIMA [Granger et Joyeux, 1980]

## Définition

$(X_k)_{k \in \mathbb{Z}}$  est un processus FARIMA( $p, d, q$ ) où  $-0.5 \leq d \leq 0.5$  et

$$X_t + a_1 X_{t-1} + \dots + a_p X_{t-p} = \varepsilon_t + b_1 \varepsilon_{t-1} + \dots + b_q \varepsilon_{t-q}$$

$$\text{où } \varepsilon_t = \xi_t + \sum_{j=1}^{\infty} \left( \frac{\Gamma(j+d)}{\Gamma(d)\Gamma(j+1)} \right) \xi_{t-j}$$

avec  $(\xi_t)$  bruit blanc de variance  $\sigma_\xi^2$ ,  $a_p, b_q$  non nuls,  $\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt$ .

**Commentaire :** • Avec  $P(x) = 1 + a_1 x + \dots + a_p x^p$ ,  $P(z) \neq 0$  pour  $|z| \leq 1$

• Si  $0 < d < 0.5$ ,  $r_X(k) \simeq C |k|^{2d-1}$  pour  $k \rightarrow \infty$  : processus longue mémoire

• On a  $\hat{\rho}(\cdot) \xrightarrow[n \rightarrow +\infty]{\mathcal{P}} \rho(\cdot)$  : ACF estime  $\rho$

## Définition

$(X_k)_{k \in \mathbb{Z}}$  est un processus FARIMA( $p, d, q$ ) où  $-0.5 \leq d \leq 0.5$  et

$$X_t + a_1 X_{t-1} + \dots + a_p X_{t-p} = \varepsilon_t + b_1 \varepsilon_{t-1} + \dots + b_q \varepsilon_{t-q}$$

$$\text{où } \varepsilon_t = \xi_t + \sum_{j=1}^{\infty} \left( \frac{\Gamma(j+d)}{\Gamma(d)\Gamma(j+1)} \right) \xi_{t-j}$$

avec  $(\xi_t)$  bruit blanc de variance  $\sigma_\xi^2$ ,  $a_p, b_q$  non nuls,  $\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt$ .

**Commentaire :** • Avec  $P(x) = 1 + a_1 x + \dots + a_p x^p$ ,  $P(z) \neq 0$  pour  $|z| \leq 1$

• Si  $0 < d < 0.5$ ,  $r_X(k) \simeq C |k|^{2d-1}$  pour  $k \rightarrow \infty$  : processus longue mémoire

• On a  $\hat{\rho}(\cdot) \xrightarrow[n \rightarrow +\infty]{\mathcal{P}} \rho(\cdot)$  : ACF estime  $\rho$

# Plan du cours

- 1 Définitions et premières propriétés
  - Séries chronologiques
  - Stationnarité
  - Première analyse statistique de la dépendance
- 2 Tendances et saisonnalités
  - Définition et propriétés
  - Estimation semi-paramétrique de la tendance et de la saisonnalité
  - Estimation non-paramétrique de la tendance et de la saisonnalité
- 3 Exemples de modèles de séries stationnaires
  - Processus ARMA
  - Processus GARCH
  - Autres exemples de séries temporelles
- 4 Estimation, sélection de modèle, test et prédiction
  - **Estimation semi-paramétrique**
  - Sélection de modèles et test d'adéquation
  - Prédiction

# Nouvelle écriture des processus

Tous les processus vus dans la section précédente peuvent s'écrire :

$$X_t = M_\theta(X_{t-1}, X_{t-2}, \dots) \xi_t + f_\theta(X_{t-1}, X_{t-2}, \dots)$$

( $\xi_t$ ) bruit blanc de variance 1,  $\theta$  vecteur de paramètres,  $M_\theta$  et  $f_\theta$  fonctions.

**Remarque** : pour les ARMA, ARMA-GARCH, FARIMA, il faut supposer que  $Q(z) \neq 0$  pour  $|z| \leq 1$   $Q(x) = 1 + b_1x + \dots + b_qx^q$

**Exemples** : • Si ( $X_t$ ) AR( $p$ ),  $X_t = \xi_t - a_1 X_{t-1} - \dots - a_p X_{t-p}$   
 $\implies M_\theta(X_{t-1}, \dots) = \sigma_\xi$ ,  $f_\theta(X_{t-1}, \dots) = -a_1 X_{t-1} - \dots - a_p X_{t-p}$

• Si ( $X_t$ ) ARCH( $p$ ),  $X_t = \xi_t \sqrt{\omega_0 + a_1 X_{t-1}^2 + \dots + a_p X_{t-p}^2}$   
 $\implies M_\theta(X_{t-1}, \dots) = \sqrt{\omega_0 + a_1 X_{t-1}^2 + \dots + a_p X_{t-p}^2}$ ,  $f_\theta(X_{t-1}, \dots) = 0$

## Nouvelle écriture des processus

Tous les processus vus dans la section précédente peuvent s'écrire :

$$X_t = M_\theta(X_{t-1}, X_{t-2}, \dots) \xi_t + f_\theta(X_{t-1}, X_{t-2}, \dots)$$

( $\xi_t$ ) bruit blanc de variance 1,  $\theta$  vecteur de paramètres,  $M_\theta$  et  $f_\theta$  fonctions.

**Remarque :** pour les ARMA, ARMA-GARCH, FARIMA, il faut supposer que  $Q(z) \neq 0$  pour  $|z| \leq 1$   $Q(x) = 1 + b_1x + \dots + b_qx^q$

**Exemples :** • Si ( $X_t$ ) AR( $p$ ),  $X_t = \xi_t - a_1 X_{t-1} - \dots - a_p X_{t-p}$   
 $\implies M_\theta(X_{t-1}, \dots) = \sigma_\xi$ ,  $f_\theta(X_{t-1}, \dots) = -a_1 X_{t-1} - \dots - a_p X_{t-p}$

• Si ( $X_t$ ) ARCH( $p$ ),  $X_t = \xi_t \sqrt{\omega_0 + a_1 X_{t-1}^2 + \dots + a_p X_{t-p}^2}$   
 $\implies M_\theta(X_{t-1}, \dots) = \sqrt{\omega_0 + a_1 X_{t-1}^2 + \dots + a_p X_{t-p}^2}$ ,  $f_\theta(X_{t-1}, \dots) = 0$

## Nouvelle écriture des processus

Tous les processus vus dans la section précédente peuvent s'écrire :

$$X_t = M_\theta(X_{t-1}, X_{t-2}, \dots) \xi_t + f_\theta(X_{t-1}, X_{t-2}, \dots)$$

( $\xi_t$ ) bruit blanc de variance 1,  $\theta$  vecteur de paramètres,  $M_\theta$  et  $f_\theta$  fonctions.

Remarque : pour les ARMA, ARMA-GARCH, FARIMA, il faut supposer que  $Q(z) \neq 0$  pour  $|z| \leq 1$   $Q(x) = 1 + b_1x + \dots + b_qx^q$

Exemples : • Si ( $X_t$ ) AR( $p$ ),  $X_t = \xi_t - a_1 X_{t-1} - \dots - a_p X_{t-p}$   
 $\implies M_\theta(X_{t-1}, \dots) = \sigma_\xi$ ,  $f_\theta(X_{t-1}, \dots) = -a_1 X_{t-1} - \dots - a_p X_{t-p}$

• Si ( $X_t$ ) ARCH( $p$ ),  $X_t = \xi_t \sqrt{\omega_0 + a_1 X_{t-1}^2 + \dots + a_p X_{t-p}^2}$   
 $\implies M_\theta(X_{t-1}, \dots) = \sqrt{\omega_0 + a_1 X_{t-1}^2 + \dots + a_p X_{t-p}^2}$ ,  $f_\theta(X_{t-1}, \dots) = 0$

## Nouvelle écriture des processus

Tous les processus vus dans la section précédente peuvent s'écrire :

$$X_t = M_\theta(X_{t-1}, X_{t-2}, \dots) \xi_t + f_\theta(X_{t-1}, X_{t-2}, \dots)$$

$(\xi_t)$  bruit blanc de variance 1,  $\theta$  vecteur de paramètres,  $M_\theta$  et  $f_\theta$  fonctions.

**Remarque** : pour les ARMA, ARMA-GARCH, FARIMA, il faut supposer que  $Q(z) \neq 0$  pour  $|z| \leq 1$   $Q(x) = 1 + b_1x + \dots + b_qx^q$

Exemples : • Si  $(X_t)$  AR( $p$ ),  $X_t = \xi_t - a_1 X_{t-1} - \dots - a_p X_{t-p}$   
 $\implies M_\theta(X_{t-1}, \dots) = \sigma_\xi$ ,  $f_\theta(X_{t-1}, \dots) = -a_1 X_{t-1} - \dots - a_p X_{t-p}$

• Si  $(X_t)$  ARCH( $p$ ),  $X_t = \xi_t \sqrt{\omega_0 + a_1 X_{t-1}^2 + \dots + a_p X_{t-p}^2}$   
 $\implies M_\theta(X_{t-1}, \dots) = \sqrt{\omega_0 + a_1 X_{t-1}^2 + \dots + a_p X_{t-p}^2}$ ,  $f_\theta(X_{t-1}, \dots) = 0$

## Nouvelle écriture des processus

Tous les processus vus dans la section précédente peuvent s'écrire :

$$X_t = M_\theta(X_{t-1}, X_{t-2}, \dots) \xi_t + f_\theta(X_{t-1}, X_{t-2}, \dots)$$

$(\xi_t)$  bruit blanc de variance 1,  $\theta$  vecteur de paramètres,  $M_\theta$  et  $f_\theta$  fonctions.

**Remarque** : pour les ARMA, ARMA-GARCH, FARIMA, il faut supposer que  $Q(z) \neq 0$  pour  $|z| \leq 1$   $Q(x) = 1 + b_1x + \dots + b_qx^q$

**Exemples** : • Si  $(X_t)$  AR( $p$ ),  $X_t = \xi_t - a_1 X_{t-1} - \dots - a_p X_{t-p}$   
 $\implies M_\theta(X_{t-1}, \dots) = \sigma_\xi, f_\theta(X_{t-1}, \dots) = -a_1 X_{t-1} - \dots - a_p X_{t-p}$

• Si  $(X_t)$  ARCH( $p$ ),  $X_t = \xi_t \sqrt{\omega_0 + a_1 X_{t-1}^2 + \dots + a_p X_{t-p}^2}$   
 $\implies M_\theta(X_{t-1}, \dots) = \sqrt{\omega_0 + a_1 X_{t-1}^2 + \dots + a_p X_{t-p}^2}, f_\theta(X_{t-1}, \dots) = 0$

## Nouvelle écriture des processus

Tous les processus vus dans la section précédente peuvent s'écrire :

$$X_t = M_\theta(X_{t-1}, X_{t-2}, \dots) \xi_t + f_\theta(X_{t-1}, X_{t-2}, \dots)$$

$(\xi_t)$  bruit blanc de variance 1,  $\theta$  vecteur de paramètres,  $M_\theta$  et  $f_\theta$  fonctions.

**Remarque** : pour les ARMA, ARMA-GARCH, FARIMA, il faut supposer que  $Q(z) \neq 0$  pour  $|z| \leq 1$   $Q(x) = 1 + b_1x + \dots + b_qx^q$

**Exemples** : • Si  $(X_t)$  AR( $p$ ),  $X_t = \xi_t - a_1 X_{t-1} - \dots - a_p X_{t-p}$   
 $\implies M_\theta(X_{t-1}, \dots) = \sigma_\xi$ ,  $f_\theta(X_{t-1}, \dots) = -a_1 X_{t-1} - \dots - a_p X_{t-p}$

• Si  $(X_t)$  ARCH( $p$ ),  $X_t = \xi_t \sqrt{\omega_0 + a_1 X_{t-1}^2 + \dots + a_p X_{t-p}^2}$   
 $\implies M_\theta(X_{t-1}, \dots) = \sqrt{\omega_0 + a_1 X_{t-1}^2 + \dots + a_p X_{t-p}^2}$ ,  $f_\theta(X_{t-1}, \dots) = 0$

## Nouvelle écriture des processus

Tous les processus vus dans la section précédente peuvent s'écrire :

$$X_t = M_\theta(X_{t-1}, X_{t-2}, \dots) \xi_t + f_\theta(X_{t-1}, X_{t-2}, \dots)$$

$(\xi_t)$  bruit blanc de variance 1,  $\theta$  vecteur de paramètres,  $M_\theta$  et  $f_\theta$  fonctions.

**Remarque** : pour les ARMA, ARMA-GARCH, FARIMA, il faut supposer que  $Q(z) \neq 0$  pour  $|z| \leq 1$   $Q(x) = 1 + b_1x + \dots + b_qx^q$

**Exemples** : • Si  $(X_t)$  AR( $p$ ),  $X_t = \xi_t - a_1 X_{t-1} - \dots - a_p X_{t-p}$   
 $\implies M_\theta(X_{t-1}, \dots) = \sigma_\xi$ ,  $f_\theta(X_{t-1}, \dots) = -a_1 X_{t-1} - \dots - a_p X_{t-p}$

• Si  $(X_t)$  ARCH( $p$ ),  $X_t = \xi_t \sqrt{\omega_0 + a_1 X_{t-1}^2 + \dots + a_p X_{t-p}^2}$   
 $\implies M_\theta(X_{t-1}, \dots) = \sqrt{\omega_0 + a_1 X_{t-1}^2 + \dots + a_p X_{t-p}^2}$ ,  $f_\theta(X_{t-1}, \dots) = 0$

## Nouvelle écriture des processus

Tous les processus vus dans la section précédente peuvent s'écrire :

$$X_t = M_\theta(X_{t-1}, X_{t-2}, \dots) \xi_t + f_\theta(X_{t-1}, X_{t-2}, \dots)$$

$(\xi_t)$  bruit blanc de variance 1,  $\theta$  vecteur de paramètres,  $M_\theta$  et  $f_\theta$  fonctions.

**Remarque** : pour les ARMA, ARMA-GARCH, FARIMA, il faut supposer que  $Q(z) \neq 0$  pour  $|z| \leq 1$   $Q(x) = 1 + b_1x + \dots + b_qx^q$

**Exemples** : • Si  $(X_t)$  AR( $p$ ),  $X_t = \xi_t - a_1 X_{t-1} - \dots - a_p X_{t-p}$   
 $\implies M_\theta(X_{t-1}, \dots) = \sigma_\xi$ ,  $f_\theta(X_{t-1}, \dots) = -a_1 X_{t-1} - \dots - a_p X_{t-p}$

• Si  $(X_t)$  ARCH( $p$ ),  $X_t = \xi_t \sqrt{\omega_0 + a_1 X_{t-1}^2 + \dots + a_p X_{t-p}^2}$   
 $\implies M_\theta(X_{t-1}, \dots) = \sqrt{\omega_0 + a_1 X_{t-1}^2 + \dots + a_p X_{t-p}^2}$ ,  $f_\theta(X_{t-1}, \dots) = 0$

## Nouvelle écriture des processus (2)

**Exemples :** • Si  $(X_t)$  MA(1),  $X_t = \xi_t + b_1\xi_{t-1}$  avec  $|b_1| < 1$  alors

$$\xi_t = \sum_{k=0}^{\infty} (-b_1)^k X_{t-k} \quad \implies X_t = \xi_t + b_1 \sum_{k=0}^{\infty} (-b_1)^k X_{t-1-k}$$

$$\implies M_\theta(X_{t-1}, \dots) = \sigma_\xi, \quad f_\theta(X_{t-1}, \dots) = b_1 \sum_{k=0}^{\infty} (-b_1)^k X_{t-1-k}$$

• Par extension si  $(X_t)$  ARMA( $p, q$ ),  $X_t = \xi_t + \sum_{k=1}^{\infty} u_k X_{t-k}$  quand  $Q(x) = 1 + b_1x + \dots + b_q x^q$  a ses racines en dehors du disque unité

$$\implies M_\theta(X_{t-1}, \dots) = \sigma_\xi, \quad f_\theta(X_{t-1}, \dots) = \sum_{k=1}^{\infty} u_k X_{t-k}$$

• Si  $(X_t)$  GARCH( $p, q$ ), on montre que  $X_t = \xi_t \left( \omega_0 + \sum_{i=1}^{\infty} u_i X_{t-i}^2 \right)^{1/2}$

$$\implies M_\theta(X_{t-1}, \dots) = \left( \omega_0 + \sum_{i=1}^{\infty} u_i X_{t-i}^2 \right)^{1/2}, \quad f_\theta(X_{t-1}, \dots) = 0$$

## Nouvelle écriture des processus (2)

**Exemples :** • Si  $(X_t)$  MA(1),  $X_t = \xi_t + b_1\xi_{t-1}$  avec  $|b_1| < 1$  alors

$$\xi_t = \sum_{k=0}^{\infty} (-b_1)^k X_{t-k} \quad \Rightarrow \quad X_t = \xi_t + b_1 \sum_{k=0}^{\infty} (-b_1)^k X_{t-1-k}$$

$$\Rightarrow M_\theta(X_{t-1}, \dots) = \sigma_\xi, \quad f_\theta(X_{t-1}, \dots) = b_1 \sum_{k=0}^{\infty} (-b_1)^k X_{t-1-k}$$

• Par extension si  $(X_t)$  ARMA( $p, q$ ),  $X_t = \xi_t + \sum_{k=1}^{\infty} u_k X_{t-k}$  quand  $Q(x) = 1 + b_1x + \dots + b_q x^q$  a ses racines en dehors du disque unité

$$\Rightarrow M_\theta(X_{t-1}, \dots) = \sigma_\xi, \quad f_\theta(X_{t-1}, \dots) = \sum_{k=1}^{\infty} u_k X_{t-k}$$

• Si  $(X_t)$  GARCH( $p, q$ ), on montre que  $X_t = \xi_t \left( \omega_0 + \sum_{i=1}^{\infty} u_i X_{t-i}^2 \right)^{1/2}$

$$\Rightarrow M_\theta(X_{t-1}, \dots) = \left( \omega_0 + \sum_{i=1}^{\infty} u_i X_{t-i}^2 \right)^{1/2}, \quad f_\theta(X_{t-1}, \dots) = 0$$

## Nouvelle écriture des processus (2)

**Exemples :** • Si  $(X_t)$  MA(1),  $X_t = \xi_t + b_1\xi_{t-1}$  avec  $|b_1| < 1$  alors

$$\xi_t = \sum_{k=0}^{\infty} (-b_1)^k X_{t-k} \quad \implies X_t = \xi_t + b_1 \sum_{k=0}^{\infty} (-b_1)^k X_{t-1-k}$$

$$\implies M_\theta(X_{t-1}, \dots) = \sigma_\xi, \quad f_\theta(X_{t-1}, \dots) = b_1 \sum_{k=0}^{\infty} (-b_1)^k X_{t-1-k}$$

• Par extension si  $(X_t)$  ARMA( $p, q$ ),  $X_t = \xi_t + \sum_{k=1}^{\infty} u_k X_{t-k}$  quand  $Q(x) = 1 + b_1x + \dots + b_q x^q$  a ses racines en dehors du disque unité

$$\implies M_\theta(X_{t-1}, \dots) = \sigma_\xi, \quad f_\theta(X_{t-1}, \dots) = \sum_{k=1}^{\infty} u_k X_{t-k}$$

• Si  $(X_t)$  GARCH( $p, q$ ), on montre que  $X_t = \xi_t \left( \omega_0 + \sum_{i=1}^{\infty} u_i X_{t-i}^2 \right)^{1/2}$

$$\implies M_\theta(X_{t-1}, \dots) = \left( \omega_0 + \sum_{i=1}^{\infty} u_i X_{t-i}^2 \right)^{1/2}, \quad f_\theta(X_{t-1}, \dots) = 0$$

## Nouvelle écriture des processus (2)

**Exemples :** • Si  $(X_t)$  MA(1),  $X_t = \xi_t + b_1 \xi_{t-1}$  avec  $|b_1| < 1$  alors

$$\xi_t = \sum_{k=0}^{\infty} (-b_1)^k X_{t-k} \quad \implies X_t = \xi_t + b_1 \sum_{k=0}^{\infty} (-b_1)^k X_{t-1-k}$$

$$\implies M_{\theta}(X_{t-1}, \dots) = \sigma_{\xi}, \quad f_{\theta}(X_{t-1}, \dots) = b_1 \sum_{k=0}^{\infty} (-b_1)^k X_{t-1-k}$$

• Par extension si  $(X_t)$  ARMA( $p, q$ ),  $X_t = \xi_t + \sum_{k=1}^{\infty} u_k X_{t-k}$  quand  $Q(x) = 1 + b_1 x + \dots + b_q x^q$  a ses racines en dehors du disque unité

$$\implies M_{\theta}(X_{t-1}, \dots) = \sigma_{\xi}, \quad f_{\theta}(X_{t-1}, \dots) = \sum_{k=1}^{\infty} u_k X_{t-k}$$

• Si  $(X_t)$  GARCH( $p, q$ ), on montre que  $X_t = \xi_t \left( \omega_0 + \sum_{i=1}^{\infty} u_i X_{t-i}^2 \right)^{1/2}$

$$\implies M_{\theta}(X_{t-1}, \dots) = \left( \omega_0 + \sum_{i=1}^{\infty} u_i X_{t-i}^2 \right)^{1/2}, \quad f_{\theta}(X_{t-1}, \dots) = 0$$

## Nouvelle écriture des processus (2)

**Exemples :** • Si  $(X_t)$  MA(1),  $X_t = \xi_t + b_1 \xi_{t-1}$  avec  $|b_1| < 1$  alors

$$\xi_t = \sum_{k=0}^{\infty} (-b_1)^k X_{t-k} \quad \implies X_t = \xi_t + b_1 \sum_{k=0}^{\infty} (-b_1)^k X_{t-1-k}$$

$$\implies M_\theta(X_{t-1}, \dots) = \sigma_\xi, \quad f_\theta(X_{t-1}, \dots) = b_1 \sum_{k=0}^{\infty} (-b_1)^k X_{t-1-k}$$

• Par extension si  $(X_t)$  ARMA( $p, q$ ),  $X_t = \xi_t + \sum_{k=1}^{\infty} u_k X_{t-k}$  quand  $Q(x) = 1 + b_1 x + \dots + b_q x^q$  a ses racines en dehors du disque unité

$$\implies M_\theta(X_{t-1}, \dots) = \sigma_\xi, \quad f_\theta(X_{t-1}, \dots) = \sum_{k=1}^{\infty} u_k X_{t-k}$$

• Si  $(X_t)$  GARCH( $p, q$ ), on montre que  $X_t = \xi_t \left( \omega_0 + \sum_{i=1}^{\infty} u_i X_{t-i}^2 \right)^{1/2}$

$$\implies M_\theta(X_{t-1}, \dots) = \left( \omega_0 + \sum_{i=1}^{\infty} u_i X_{t-i}^2 \right)^{1/2}, \quad f_\theta(X_{t-1}, \dots) = 0$$

## Nouvelle écriture des processus (2)

**Exemples :** • Si  $(X_t)$  MA(1),  $X_t = \xi_t + b_1 \xi_{t-1}$  avec  $|b_1| < 1$  alors

$$\xi_t = \sum_{k=0}^{\infty} (-b_1)^k X_{t-k} \quad \implies X_t = \xi_t + b_1 \sum_{k=0}^{\infty} (-b_1)^k X_{t-1-k}$$

$$\implies M_\theta(X_{t-1}, \dots) = \sigma_\xi, \quad f_\theta(X_{t-1}, \dots) = b_1 \sum_{k=0}^{\infty} (-b_1)^k X_{t-1-k}$$

• Par extension si  $(X_t)$  ARMA( $p, q$ ),  $X_t = \xi_t + \sum_{k=1}^{\infty} u_k X_{t-k}$  quand  $Q(x) = 1 + b_1 x + \dots + b_q x^q$  a ses racines en dehors du disque unité

$$\implies M_\theta(X_{t-1}, \dots) = \sigma_\xi, \quad f_\theta(X_{t-1}, \dots) = \sum_{k=1}^{\infty} u_k X_{t-k}$$

• Si  $(X_t)$  GARCH( $p, q$ ), on montre que  $X_t = \xi_t \left( \omega_0 + \sum_{i=1}^{\infty} u_i X_{t-i}^2 \right)^{1/2}$

$$\implies M_\theta(X_{t-1}, \dots) = \left( \omega_0 + \sum_{i=1}^{\infty} u_i X_{t-i}^2 \right)^{1/2}, \quad f_\theta(X_{t-1}, \dots) = 0$$

## Nouvelle écriture des processus (2)

**Exemples :** • Si  $(X_t)$  MA(1),  $X_t = \xi_t + b_1 \xi_{t-1}$  avec  $|b_1| < 1$  alors

$$\xi_t = \sum_{k=0}^{\infty} (-b_1)^k X_{t-k} \quad \implies X_t = \xi_t + b_1 \sum_{k=0}^{\infty} (-b_1)^k X_{t-1-k}$$

$$\implies M_\theta(X_{t-1}, \dots) = \sigma_\xi, \quad f_\theta(X_{t-1}, \dots) = b_1 \sum_{k=0}^{\infty} (-b_1)^k X_{t-1-k}$$

• Par extension si  $(X_t)$  ARMA( $p, q$ ),  $X_t = \xi_t + \sum_{k=1}^{\infty} u_k X_{t-k}$  quand  $Q(x) = 1 + b_1 x + \dots + b_q x^q$  a ses racines en dehors du disque unité

$$\implies M_\theta(X_{t-1}, \dots) = \sigma_\xi, \quad f_\theta(X_{t-1}, \dots) = \sum_{k=1}^{\infty} u_k X_{t-k}$$

• Si  $(X_t)$  GARCH( $p, q$ ), on montre que  $X_t = \xi_t \left( \omega_0 + \sum_{i=1}^{\infty} u_i X_{t-i}^2 \right)^{1/2}$

$$\implies M_\theta(X_{t-1}, \dots) = \left( \omega_0 + \sum_{i=1}^{\infty} u_i X_{t-i}^2 \right)^{1/2}, \quad f_\theta(X_{t-1}, \dots) = 0$$

## Nouvelle écriture des processus (2)

**Exemples :** • Si  $(X_t)$  MA(1),  $X_t = \xi_t + b_1 \xi_{t-1}$  avec  $|b_1| < 1$  alors

$$\xi_t = \sum_{k=0}^{\infty} (-b_1)^k X_{t-k} \quad \implies X_t = \xi_t + b_1 \sum_{k=0}^{\infty} (-b_1)^k X_{t-1-k}$$

$$\implies M_\theta(X_{t-1}, \dots) = \sigma_\xi, \quad f_\theta(X_{t-1}, \dots) = b_1 \sum_{k=0}^{\infty} (-b_1)^k X_{t-1-k}$$

• Par extension si  $(X_t)$  ARMA( $p, q$ ),  $X_t = \xi_t + \sum_{k=1}^{\infty} u_k X_{t-k}$  quand  $Q(x) = 1 + b_1 x + \dots + b_q x^q$  a ses racines en dehors du disque unité

$$\implies M_\theta(X_{t-1}, \dots) = \sigma_\xi, \quad f_\theta(X_{t-1}, \dots) = \sum_{k=1}^{\infty} u_k X_{t-k}$$

• Si  $(X_t)$  GARCH( $p, q$ ), on montre que  $X_t = \xi_t \left( \omega_0 + \sum_{i=1}^{\infty} u_i X_{t-i}^2 \right)^{1/2}$

$$\implies M_\theta(X_{t-1}, \dots) = \left( \omega_0 + \sum_{i=1}^{\infty} u_i X_{t-i}^2 \right)^{1/2}, \quad f_\theta(X_{t-1}, \dots) = 0$$

# Estimation par maximum de vraisemblance

$$X_t = M_\theta(X_{t-1}, X_{t-2}, \dots) \xi_t + f_\theta(X_{t-1}, X_{t-2}, \dots)$$

On veut estimer  $\theta \in \mathbb{R}^d$  à partir de  $(X_1, \dots, X_n)$  trajectoire observée.

Première idée : Utiliser le maximum de vraisemblance

- Requierit la connaissance de la loi de  $(X_1, \dots, X_n)$
- En dehors du cas gaussien, quasiment impossible car dépendance !
- Et les GARCH( $p, q$ ) ne sont jamais gaussiens !

# Estimation par maximum de vraisemblance

$$X_t = M_\theta(X_{t-1}, X_{t-2}, \dots) \xi_t + f_\theta(X_{t-1}, X_{t-2}, \dots)$$

On veut estimer  $\theta \in \mathbb{R}^d$  à partir de  $(X_1, \dots, X_n)$  trajectoire observée.

Première idée : Utiliser le maximum de vraisemblance

- Requiert la connaissance de la loi de  $(X_1, \dots, X_n)$
- En dehors du cas gaussien, quasiment impossible car dépendance !
- Et les GARCH( $p, q$ ) ne sont jamais gaussiens !

# Estimation par maximum de vraisemblance

$$X_t = M_\theta(X_{t-1}, X_{t-2}, \dots) \xi_t + f_\theta(X_{t-1}, X_{t-2}, \dots)$$

On veut estimer  $\theta \in \mathbb{R}^d$  à partir de  $(X_1, \dots, X_n)$  trajectoire observée.

**Première idée** : Utiliser le maximum de vraisemblance

- Requierit la connaissance de la loi de  $(X_1, \dots, X_n)$
- En dehors du cas gaussien, quasiment impossible car dépendance !
- Et les GARCH( $p, q$ ) ne sont jamais gaussiens !

# Estimation par maximum de vraisemblance

$$X_t = M_\theta(X_{t-1}, X_{t-2}, \dots) \xi_t + f_\theta(X_{t-1}, X_{t-2}, \dots)$$

On veut estimer  $\theta \in \mathbb{R}^d$  à partir de  $(X_1, \dots, X_n)$  trajectoire observée.

**Première idée** : Utiliser le maximum de vraisemblance

- Requiert la connaissance de la loi de  $(X_1, \dots, X_n)$
- En dehors du cas gaussien, quasiment impossible car dépendance !
- Et les GARCH( $p, q$ ) ne sont jamais gaussiens !

# Estimation par maximum de vraisemblance

$$X_t = M_\theta(X_{t-1}, X_{t-2}, \dots) \xi_t + f_\theta(X_{t-1}, X_{t-2}, \dots)$$

On veut estimer  $\theta \in \mathbb{R}^d$  à partir de  $(X_1, \dots, X_n)$  trajectoire observée.

**Première idée** : Utiliser le maximum de vraisemblance

- Requiert la connaissance de la loi de  $(X_1, \dots, X_n)$
- En dehors du cas gaussien, quasiment impossible car dépendance !
- Et les GARCH( $p, q$ ) ne sont jamais gaussiens !

# Estimation par maximum de vraisemblance

$$X_t = M_\theta(X_{t-1}, X_{t-2}, \dots) \xi_t + f_\theta(X_{t-1}, X_{t-2}, \dots)$$

On veut estimer  $\theta \in \mathbb{R}^d$  à partir de  $(X_1, \dots, X_n)$  trajectoire observée.

**Première idée** : Utiliser le maximum de vraisemblance

- Requiert la connaissance de la loi de  $(X_1, \dots, X_n)$
- En dehors du cas gaussien, quasiment impossible car dépendance !
- Et les GARCH( $p, q$ ) ne sont jamais gaussiens !

# Estimation par maximum de vraisemblance conditionnelle

$$X_t = M_\theta(X_{t-1}, X_{t-2}, \dots) \xi_t + f_\theta(X_{t-1}, X_{t-2}, \dots)$$

Seconde idée : Utiliser le maximum de vraisemblance conditionnelle

⇒ On calcule la densité de  $(X_1, \dots, X_n)$  sachant  $(X_0, X_{-1}, \dots)$ .

On obtient alors la vraisemblance conditionnelle :

$$L_\theta(X_1, \dots, X_n) = f(X_n | X_{n-1}, X_{n-2}, \dots) \\ \times f(X_{n-1} | X_{n-2}, X_{n-3}, \dots) \times \dots \times f(X_1 | X_0, X_{-1}, \dots)$$

Si  $(\xi_t)$  gaussien  $\mathcal{N}(0, 1)$ ,  $X_k$  est gaussien sachant  $X_{k-1}, X_{k-2}, \dots$  d'où :

$$f(X_k | X_{k-1}, X_{k-2}, \dots) = \frac{1}{\sqrt{2\pi}} \frac{1}{M_\theta(X_{k-1}, X_{k-2}, \dots)} \exp\left(-\frac{1}{2} \frac{(X_k - f_\theta(X_{k-1}, X_{k-2}, \dots))^2}{M_\theta^2(X_{k-1}, X_{k-2}, \dots)}\right)$$

# Estimation par maximum de vraisemblance conditionnelle

$$X_t = M_\theta(X_{t-1}, X_{t-2}, \dots) \xi_t + f_\theta(X_{t-1}, X_{t-2}, \dots)$$

**Seconde idée :** Utiliser le maximum de vraisemblance conditionnelle

⇒ On calcule la densité de  $(X_1, \dots, X_n)$  sachant  $(X_0, X_{-1}, \dots)$ .

On obtient alors la vraisemblance conditionnelle :

$$L_\theta(X_1, \dots, X_n) = f(X_n | X_{n-1}, X_{n-2}, \dots) \\ \times f(X_{n-1} | X_{n-2}, X_{n-3}, \dots) \times \dots \times f(X_1 | X_0, X_{-1}, \dots)$$

Si  $(\xi_t)$  gaussien  $\mathcal{N}(0, 1)$ ,  $X_k$  est gaussien sachant  $X_{k-1}, X_{k-2}, \dots$  d'où :

$$f(X_k | X_{k-1}, X_{k-2}, \dots) = \frac{1}{\sqrt{2\pi}} \frac{1}{M_\theta(X_{k-1}, X_{k-2}, \dots)} \exp\left(-\frac{1}{2} \frac{(X_k - f_\theta(X_{k-1}, X_{k-2}, \dots))^2}{M_\theta^2(X_{k-1}, X_{k-2}, \dots)}\right)$$

# Estimation par maximum de vraisemblance conditionnelle

$$X_t = M_\theta(X_{t-1}, X_{t-2}, \dots) \xi_t + f_\theta(X_{t-1}, X_{t-2}, \dots)$$

**Seconde idée :** Utiliser le maximum de vraisemblance conditionnelle

⇒ On calcule la densité de  $(X_1, \dots, X_n)$  sachant  $(X_0, X_{-1}, \dots)$ .

On obtient alors la vraisemblance conditionnelle :

$$L_\theta(X_1, \dots, X_n) = f(X_n | X_{n-1}, X_{n-2}, \dots) \\ \times f(X_{n-1} | X_{n-2}, X_{n-3}, \dots) \times \dots \times f(X_1 | X_0, X_{-1}, \dots)$$

Si  $(\xi_t)$  gaussien  $\mathcal{N}(0, 1)$ ,  $X_k$  est gaussien sachant  $X_{k-1}, X_{k-2}, \dots$  d'où :

$$f(X_k | X_{k-1}, X_{k-2}, \dots) = \frac{1}{\sqrt{2\pi}} \frac{1}{M_\theta(X_{k-1}, X_{k-2}, \dots)} \exp\left(-\frac{1}{2} \frac{(X_k - f_\theta(X_{k-1}, X_{k-2}, \dots))^2}{M_\theta^2(X_{k-1}, X_{k-2}, \dots)}\right)$$

# Estimation par maximum de vraisemblance conditionnelle

$$X_t = M_\theta(X_{t-1}, X_{t-2}, \dots) \xi_t + f_\theta(X_{t-1}, X_{t-2}, \dots)$$

**Seconde idée :** Utiliser le maximum de vraisemblance conditionnelle

⇒ On calcule la densité de  $(X_1, \dots, X_n)$  sachant  $(X_0, X_{-1}, \dots)$ .

On obtient alors la vraisemblance conditionnelle :

$$L_\theta(X_1, \dots, X_n) = f(X_n | X_{n-1}, X_{n-2}, \dots) \\ \times f(X_{n-1} | X_{n-2}, X_{n-3}, \dots) \times \dots \times f(X_1 | X_0, X_{-1}, \dots)$$

Si  $(\xi_t)$  gaussien  $\mathcal{N}(0, 1)$ ,  $X_k$  est gaussien sachant  $X_{k-1}, X_{k-2}, \dots$  d'où :

$$f(X_k | X_{k-1}, X_{k-2}, \dots) = \frac{1}{\sqrt{2\pi}} \frac{1}{M_\theta(X_{k-1}, X_{k-2}, \dots)} \exp\left(-\frac{1}{2} \frac{(X_k - f_\theta(X_{k-1}, X_{k-2}, \dots))^2}{M_\theta^2(X_{k-1}, X_{k-2}, \dots)}\right)$$

## Estimation par maximum de vraisemblance conditionnelle

$$X_t = M_\theta(X_{t-1}, X_{t-2}, \dots) \xi_t + f_\theta(X_{t-1}, X_{t-2}, \dots)$$

**Seconde idée :** Utiliser le maximum de vraisemblance conditionnelle

⇒ On calcule la densité de  $(X_1, \dots, X_n)$  sachant  $(X_0, X_{-1}, \dots)$ .

On obtient alors la vraisemblance conditionnelle :

$$L_\theta(X_1, \dots, X_n) = f(X_n | X_{n-1}, X_{n-2}, \dots) \\ \times f(X_{n-1} | X_{n-2}, X_{n-3}, \dots) \times \dots \times f(X_1 | X_0, X_{-1}, \dots)$$

Si  $(\xi_t)$  gaussien  $\mathcal{N}(0, 1)$ ,  $X_k$  est gaussien sachant  $X_{k-1}, X_{k-2}, \dots$  d'où :

$$f(X_k | X_{k-1}, X_{k-2}, \dots) = \frac{1}{\sqrt{2\pi}} \frac{1}{M_\theta(X_{k-1}, X_{k-2}, \dots)} \exp\left(-\frac{1}{2} \frac{(X_k - f_\theta(X_{k-1}, X_{k-2}, \dots))^2}{M_\theta^2(X_{k-1}, X_{k-2}, \dots)}\right)$$

## Estimation par maximum de vraisemblance conditionnelle

$$X_t = M_\theta(X_{t-1}, X_{t-2}, \dots) \xi_t + f_\theta(X_{t-1}, X_{t-2}, \dots)$$

**Seconde idée :** Utiliser le maximum de vraisemblance conditionnelle

⇒ On calcule la densité de  $(X_1, \dots, X_n)$  sachant  $(X_0, X_{-1}, \dots)$ .

On obtient alors la vraisemblance conditionnelle :

$$L_\theta(X_1, \dots, X_n) = f(X_n | X_{n-1}, X_{n-2}, \dots) \\ \times f(X_{n-1} | X_{n-2}, X_{n-3}, \dots) \times \dots \times f(X_1 | X_0, X_{-1}, \dots)$$

Si  $(\xi_t)$  gaussien  $\mathcal{N}(0, 1)$ ,  $X_k$  est gaussien sachant  $X_{k-1}, X_{k-2}, \dots$  d'où :

$$f(X_k | X_{k-1}, X_{k-2}, \dots) = \frac{1}{\sqrt{2\pi}} \frac{1}{M_\theta(X_{k-1}, X_{k-2}, \dots)} \exp\left(-\frac{1}{2} \frac{(X_k - f_\theta(X_{k-1}, X_{k-2}, \dots))^2}{M_\theta^2(X_{k-1}, X_{k-2}, \dots)}\right)$$

## Estimation par maximum de vraisemblance conditionnelle

$$X_t = M_\theta(X_{t-1}, X_{t-2}, \dots) \xi_t + f_\theta(X_{t-1}, X_{t-2}, \dots)$$

**Seconde idée :** Utiliser le maximum de vraisemblance conditionnelle

⇒ On calcule la densité de  $(X_1, \dots, X_n)$  sachant  $(X_0, X_{-1}, \dots)$ .

On obtient alors la vraisemblance conditionnelle :

$$L_\theta(X_1, \dots, X_n) = f(X_n | X_{n-1}, X_{n-2}, \dots) \\ \times f(X_{n-1} | X_{n-2}, X_{n-3}, \dots) \times \dots \times f(X_1 | X_0, X_{-1}, \dots)$$

Si  $(\xi_t)$  gaussien  $\mathcal{N}(0, 1)$ ,  $X_k$  est gaussien sachant  $X_{k-1}, X_{k-2}, \dots$  d'où :

$$f(X_k | X_{k-1}, X_{k-2}, \dots) = \frac{1}{\sqrt{2\pi}} \frac{1}{M_\theta(X_{k-1}, X_{k-2}, \dots)} \exp\left(-\frac{1}{2} \frac{(X_k - f_\theta(X_{k-1}, X_{k-2}, \dots))^2}{M_\theta^2(X_{k-1}, X_{k-2}, \dots)}\right)$$

# Estimation par quasi-maximum de vraisemblance

On obtient alors la log-vraisemblance conditionnelle :

$$\log(L_{\theta}(X_1, \dots, X_n)) = -\frac{1}{2} \left( n \log(2\pi) + \sum_{t=1}^n \log(M_{\theta}^2(X_{t-1}, X_{t-2}, \dots)) + \frac{(X_t - f_{\theta}(X_{t-1}, X_{t-2}, \dots))^2}{M_{\theta}^2(X_{t-1}, X_{t-2}, \dots)} \right)$$

Mais  $X_0, X_{-1}, \dots$  non observées alors que par exemple pour un MA(1) on a

$$f_{\theta}(X_{t-1}, \dots) = b_1 \sum_{k=0}^{\infty} (-b_1)^k X_{t-1-k}$$

$$\Rightarrow \begin{cases} \hat{M}_{\theta}^t = M_{\theta}(X_{t-1}, \dots, X_1, 0, 0, \dots) \\ \hat{f}_{\theta}^t = f_{\theta}(X_{t-1}, \dots, X_1, 0, 0, \dots) \end{cases} \text{ remplace } \begin{cases} M_{\theta}(X_{t-1}, X_{t-2}, \dots) \\ f_{\theta}(X_{t-1}, X_{t-2}, \dots) \end{cases}$$

## Estimation par quasi-maximum de vraisemblance

On obtient alors la log-vraisemblance conditionnelle :

$$\log(L_{\theta}(X_1, \dots, X_n)) = -\frac{1}{2} \left( n \log(2\pi) + \sum_{t=1}^n \log(M_{\theta}^2(X_{t-1}, X_{t-2}, \dots)) + \frac{(X_t - f_{\theta}(X_{t-1}, X_{t-2}, \dots))^2}{M_{\theta}^2(X_{t-1}, X_{t-2}, \dots)} \right)$$

Mais  $X_0, X_{-1}, \dots$  non observées alors que par exemple pour un MA(1) on a

$$f_{\theta}(X_{t-1}, \dots) = b_1 \sum_{k=0}^{\infty} (-b_1)^k X_{t-1-k}$$

$$\Rightarrow \begin{cases} \hat{M}_{\theta}^t = M_{\theta}(X_{t-1}, \dots, X_1, 0, 0, \dots) \\ \hat{f}_{\theta}^t = f_{\theta}(X_{t-1}, \dots, X_1, 0, 0, \dots) \end{cases} \text{ remplace } \begin{cases} M_{\theta}(X_{t-1}, X_{t-2}, \dots) \\ f_{\theta}(X_{t-1}, X_{t-2}, \dots) \end{cases}$$

## Estimation par quasi-maximum de vraisemblance

On obtient alors la log-vraisemblance conditionnelle :

$$\log(L_{\theta}(X_1, \dots, X_n)) = -\frac{1}{2} \left( n \log(2\pi) + \sum_{t=1}^n \log(M_{\theta}^2(X_{t-1}, X_{t-2}, \dots)) + \frac{(X_t - f_{\theta}(X_{t-1}, X_{t-2}, \dots))^2}{M_{\theta}^2(X_{t-1}, X_{t-2}, \dots)} \right)$$

**Mais**  $X_0, X_{-1}, \dots$  non observées alors que par exemple pour un MA(1) on a

$$f_{\theta}(X_{t-1}, \dots) = b_1 \sum_{k=0}^{\infty} (-b_1)^k X_{t-1-k}$$

$$\Rightarrow \begin{cases} \hat{M}_{\theta}^t = M_{\theta}(X_{t-1}, \dots, X_1, 0, 0, \dots) \\ \hat{f}_{\theta}^t = f_{\theta}(X_{t-1}, \dots, X_1, 0, 0, \dots) \end{cases} \text{ remplace } \begin{cases} M_{\theta}(X_{t-1}, X_{t-2}, \dots) \\ f_{\theta}(X_{t-1}, X_{t-2}, \dots) \end{cases}$$

## Estimation par quasi-maximum de vraisemblance

On obtient alors la log-vraisemblance conditionnelle :

$$\log(L_{\theta}(X_1, \dots, X_n)) = -\frac{1}{2} \left( n \log(2\pi) + \sum_{t=1}^n \log(M_{\theta}^2(X_{t-1}, X_{t-2}, \dots)) + \frac{(X_t - f_{\theta}(X_{t-1}, X_{t-2}, \dots))^2}{M_{\theta}^2(X_{t-1}, X_{t-2}, \dots)} \right)$$

**Mais**  $X_0, X_{-1}, \dots$  non observées alors que par exemple pour un MA(1) on a

$$f_{\theta}(X_{t-1}, \dots) = b_1 \sum_{k=0}^{\infty} (-b_1)^k X_{t-1-k}$$

$$\Rightarrow \begin{cases} \hat{M}_{\theta}^t = M_{\theta}(X_{t-1}, \dots, X_1, 0, 0, \dots) \\ \hat{f}_{\theta}^t = f_{\theta}(X_{t-1}, \dots, X_1, 0, 0, \dots) \end{cases} \text{ remplace } \begin{cases} M_{\theta}(X_{t-1}, X_{t-2}, \dots) \\ f_{\theta}(X_{t-1}, X_{t-2}, \dots) \end{cases}$$

# Estimation par quasi-maximum de vraisemblance

On obtient alors la quasi-log-vraisemblance conditionnelle :

$$\log(\widehat{L}_\theta(X_1, \dots, X_n)) = -\frac{1}{2} \left( \sum_{t=1}^n \log((\widehat{M}_\theta^t)^2) + \frac{(X_t - \widehat{f}_\theta^t)^2}{(\widehat{M}_\theta^t)^2} \right)$$

On définit l'estimateur par quasi-maximum de vraisemblance QMLE :

$$\widehat{\theta}_n = \underset{\theta \in \Theta}{\text{Argmax}} \log(\widehat{L}_\theta(X_1, \dots, X_n))$$

Exemples :

- Pour un processus AR( $p$ ),

$$(\widehat{a}_1, \dots, \widehat{a}_p) \simeq \underset{(a_1, \dots, a_p) \in \text{ACRP}}{\text{Argmin}} \sum_{k=p+1}^n (X_k - (a_1 X_{k-1} + \dots + a_p X_{k-p}))^2$$

- Pour un processus ARCH( $p$ ),

$$(\widehat{\omega}_0, \widehat{a}_1, \dots, \widehat{a}_p) \simeq \underset{(\omega_0, a_1, \dots, a_p) \in \text{ACRP}^{p+1}}{\text{Argmin}} \sum_{k=p+1}^n \log(\omega_0 + a_1 X_{k-1}^2 + \dots + a_p X_{k-p}^2)$$

$$+ \frac{X_k^2}{\omega_0 + a_1 X_{k-1}^2 + \dots + a_p X_{k-p}^2}$$

# Estimation par quasi-maximum de vraisemblance

On obtient alors la quasi-log-vraisemblance conditionnelle :

$$\log(\widehat{L}_\theta(X_1, \dots, X_n)) = -\frac{1}{2} \left( \sum_{t=1}^n \log((\widehat{M}_\theta^t)^2) + \frac{(X_t - \widehat{f}_\theta^t)^2}{(\widehat{M}_\theta^t)^2} \right)$$

On définit l'estimateur par quasi-maximum de vraisemblance QMLE :

$$\widehat{\theta}_n = \underset{\theta \in \Theta}{\text{Argmax}} \log(\widehat{L}_\theta(X_1, \dots, X_n))$$

Exemples :

- Pour un processus AR(p),

$$(\widehat{a}_1, \dots, \widehat{a}_p) \simeq \underset{(a_1, \dots, a_p) \in \text{ACRP}}{\text{Argmin}} \sum_{k=p+1}^n (X_k - (a_1 X_{k-1} + \dots + a_p X_{k-p}))^2$$

- Pour un processus ARCH(p),

$$(\widehat{\omega}_0, \widehat{a}_1, \dots, \widehat{a}_p) \simeq \underset{(\omega_0, a_1, \dots, a_p) \in \text{ACRP}^{p+1}}{\text{Argmin}} \sum_{k=p+1}^n \log(\omega_0 + a_1 X_{k-1}^2 + \dots + a_p X_{k-p}^2)$$

$$+ \frac{X_k^2}{\omega_0 + a_1 X_{k-1}^2 + \dots + a_p X_{k-p}^2}$$

## Estimation par quasi-maximum de vraisemblance

On obtient alors la quasi-log-vraisemblance conditionnelle :

$$\log(\widehat{L}_\theta(X_1, \dots, X_n)) = -\frac{1}{2} \left( \sum_{t=1}^n \log((\widehat{M}_\theta^t)^2) + \frac{(X_t - \widehat{f}_\theta^t)^2}{(\widehat{M}_\theta^t)^2} \right)$$

On définit l'estimateur par quasi-maximum de vraisemblance QMLE :

$$\widehat{\theta}_n = \underset{\theta \in \Theta}{\text{Argmax}} \log(\widehat{L}_\theta(X_1, \dots, X_n))$$

Exemples :

- Pour un processus AR( $p$ ),

$$(\widehat{a}_1, \dots, \widehat{a}_p) \simeq \underset{(a_1, \dots, a_p) \in \text{ACRP}}{\text{Argmin}} \sum_{k=p+1}^n (X_k - (a_1 X_{k-1} + \dots + a_p X_{k-p}))^2$$

- Pour un processus ARCH( $p$ ),

$$(\widehat{\omega}_0, \widehat{a}_1, \dots, \widehat{a}_p) \simeq \underset{(\omega_0, a_1, \dots, a_p) \in \text{ACRP}^{p+1}}{\text{Argmin}} \sum_{k=p+1}^n \log(\omega_0 + a_1 X_{k-1}^2 + \dots + a_p X_{k-p}^2)$$

$$+ \frac{X_k^2}{\omega_0 + a_1 X_{k-1}^2 + \dots + a_p X_{k-p}^2}$$

## Estimation par quasi-maximum de vraisemblance

On obtient alors la quasi-log-vraisemblance conditionnelle :

$$\log(\widehat{L}_\theta(X_1, \dots, X_n)) = -\frac{1}{2} \left( \sum_{t=1}^n \log((\widehat{M}_\theta^t)^2) + \frac{(X_t - \widehat{f}_\theta^t)^2}{(\widehat{M}_\theta^t)^2} \right)$$

On définit l'estimateur par quasi-maximum de vraisemblance QMLE :

$$\widehat{\theta}_n = \underset{\theta \in \Theta}{\text{Argmax}} \log(\widehat{L}_\theta(X_1, \dots, X_n))$$

Exemples :

- Pour un processus AR(p),

$$(\widehat{a}_1, \dots, \widehat{a}_p) \simeq \underset{(a_1, \dots, a_p) \in \text{ACRP}}{\text{Argmin}} \sum_{k=p+1}^n (X_k - (a_1 X_{k-1} + \dots + a_p X_{k-p}))^2$$

- Pour un processus ARCH(p),

$$(\widehat{\omega}_0, \widehat{a}_1, \dots, \widehat{a}_p) \simeq \underset{(\omega_0, a_1, \dots, a_p) \in \text{ACRP}^{p+1}}{\text{Argmin}} \sum_{k=p+1}^n \log(\omega_0 + a_1 X_{k-1}^2 + \dots + a_p X_{k-p}^2)$$

$$+ \frac{X_k^2}{\omega_0 + a_1 X_{k-1}^2 + \dots + a_p X_{k-p}^2}$$

# Estimation par quasi-maximum de vraisemblance

On obtient alors la quasi-log-vraisemblance conditionnelle :

$$\log(\widehat{L}_\theta(X_1, \dots, X_n)) = -\frac{1}{2} \left( \sum_{t=1}^n \log((\widehat{M}_\theta^t)^2) + \frac{(X_t - \widehat{f}_\theta^t)^2}{(\widehat{M}_\theta^t)^2} \right)$$

On définit l'estimateur par quasi-maximum de vraisemblance QMLE :

$$\widehat{\theta}_n = \underset{\theta \in \Theta}{\text{Argmax}} \log(\widehat{L}_\theta(X_1, \dots, X_n))$$

**Exemples :**

- Pour un processus AR(p),

$$(\widehat{a}_1, \dots, \widehat{a}_p) \simeq \underset{(a_1, \dots, a_p) \in \mathbb{A} \subset \mathbb{R}^p}{\text{Argmin}} \sum_{k=p+1}^n (X_k - (a_1 X_{k-1} + \dots + a_p X_{k-p}))^2$$

- Pour un processus ARCH(p),

$$(\widehat{\omega}_0, \widehat{a}_1, \dots, \widehat{a}_p) \simeq \underset{(\omega_0, a_1, \dots, a_p) \in \mathbb{A} \subset \mathbb{R}^{p+1}}{\text{Argmin}} \sum_{k=p+1}^n \log(\omega_0 + a_1 X_{k-1}^2 + \dots + a_p X_{k-p}^2)$$

$$+ \frac{X_k^2}{\omega_0 + a_1 X_{k-1}^2 + \dots + a_p X_{k-p}^2}$$

## Estimation par quasi-maximum de vraisemblance

On obtient alors la quasi-log-vraisemblance conditionnelle :

$$\log(\widehat{L}_\theta(X_1, \dots, X_n)) = -\frac{1}{2} \left( \sum_{t=1}^n \log((\widehat{M}_\theta^t)^2) + \frac{(X_t - \widehat{f}_\theta^t)^2}{(\widehat{M}_\theta^t)^2} \right)$$

On définit l'estimateur par quasi-maximum de vraisemblance QMLE :

$$\widehat{\theta}_n = \underset{\theta \in \Theta}{\text{Argmax}} \log(\widehat{L}_\theta(X_1, \dots, X_n))$$

**Exemples :**

- Pour un processus AR( $p$ ),

$$(\widehat{a}_1, \dots, \widehat{a}_p) \simeq \underset{(a_1, \dots, a_p) \in \mathbb{A} \subset \mathbb{R}^p}{\text{Argmin}} \sum_{k=p+1}^n (X_k - (a_1 X_{k-1} + \dots + a_p X_{k-p}))^2$$

- Pour un processus ARCH( $p$ ),

$$(\widehat{\omega}_0, \widehat{a}_1, \dots, \widehat{a}_p) \simeq \underset{(\omega_0, a_1, \dots, a_p) \in \mathbb{A} \subset \mathbb{R}^{p+1}}{\text{Argmin}} \sum_{k=p+1}^n \log(\omega_0 + a_1 X_{k-1}^2 + \dots + a_p X_{k-p}^2)$$

$$+ \frac{X_k^2}{\omega_0 + a_1 X_{k-1}^2 + \dots + a_p X_{k-p}^2}$$

# Convergence de l'estimateur QML

On peut montrer que :

## Théorème

Sous certaines conditions, et  $\mathbb{E}[\xi_0^4] < \infty$  même si  $(\xi_t)$  non gaussien on a

$$\sqrt{n}(\hat{\theta}_n - \theta^*) \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \mathcal{N}_d(0, F(\theta^*)^{-1}G(\theta^*)F(\theta^*)^{-1}),$$

$$\text{avec } \begin{cases} F(\theta^*) &= \lim_{n \rightarrow \infty} \frac{1}{n} \left[ \frac{\partial^2}{\partial \theta_i \partial \theta_j} \log(\hat{L}_{\theta^*}(X_1, \dots, X_n)) \right]_{i,j} \\ G(\theta^*) &= \lim_{n \rightarrow \infty} \frac{1}{n} \left[ \frac{\partial}{\partial \theta_i} \log(\hat{L}_{\theta^*}(X_1, \dots, X_n)) \frac{\partial}{\partial \theta_j} \log(\hat{L}_{\theta^*}(X_1, \dots, X_n)) \right]_{i,j} \end{cases}$$

**Exemples** : Le théorème est valable pour les processus ARMA, GARCH, APARCH, FARIMA, ARMA-GARCH,...

**Remarque** : Avec le Lemme de Slutsky et  $\hat{F}(\hat{\theta}_n)$  et  $\hat{G}(\hat{\theta}_n)$  au lieu de  $F(\theta^*)$  et  $G(\theta^*)$  (sans lim), on obtient des intervalles de confiance, tests de Wald,

# Convergence de l'estimateur QML

On peut montrer que :

## Théorème

Sous certaines conditions, et  $\mathbf{E}[\xi_0^4] < \infty$  même si  $(\xi_t)$  non gaussien on a

$$\sqrt{n}(\hat{\theta}_n - \theta^*) \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \mathcal{N}_d(0, F(\theta^*)^{-1} G(\theta^*) F(\theta^*)^{-1}),$$

$$\text{avec } \begin{cases} F(\theta^*) &= \lim_{n \rightarrow \infty} \frac{1}{n} \left[ \frac{\partial^2}{\partial \theta_i \partial \theta_j} \log(\hat{L}_{\theta^*}(X_1, \dots, X_n)) \right]_{i,j} \\ G(\theta^*) &= \lim_{n \rightarrow \infty} \frac{1}{n} \left[ \frac{\partial}{\partial \theta_i} \log(\hat{L}_{\theta^*}(X_1, \dots, X_n)) \frac{\partial}{\partial \theta_j} \log(\hat{L}_{\theta^*}(X_1, \dots, X_n)) \right]_{i,j} \end{cases}$$

**Exemples** : Le théorème est valable pour les processus ARMA, GARCH, APARCH, FARIMA, ARMA-GARCH,...

**Remarque** : Avec le Lemme de Slutsky et  $\hat{F}(\hat{\theta}_n)$  et  $\hat{G}(\hat{\theta}_n)$  au lieu de  $F(\theta^*)$  et  $G(\theta^*)$  (sans lim), on obtient des intervalles de confiance, tests de Wald,

# Convergence de l'estimateur QML

On peut montrer que :

## Théorème

Sous certaines conditions, et  $\mathbf{E}[\xi_0^4] < \infty$  même si  $(\xi_t)$  non gaussien on a

$$\sqrt{n}(\hat{\theta}_n - \theta^*) \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \mathcal{N}_d(0, F(\theta^*)^{-1}G(\theta^*)F(\theta^*)^{-1}),$$

$$\text{avec } \begin{cases} F(\theta^*) &= \lim_{n \rightarrow \infty} \frac{1}{n} \left[ \frac{\partial^2}{\partial \theta_i \partial \theta_j} \log(\hat{L}_{\theta^*}(X_1, \dots, X_n)) \right]_{i,j} \\ G(\theta^*) &= \lim_{n \rightarrow \infty} \frac{1}{n} \left[ \frac{\partial}{\partial \theta_i} \log(\hat{L}_{\theta^*}(X_1, \dots, X_n)) \frac{\partial}{\partial \theta_j} \log(\hat{L}_{\theta^*}(X_1, \dots, X_n)) \right]_{i,j} \end{cases}$$

**Exemples** : Le théorème est valable pour les processus ARMA, GARCH, APARCH, FARIMA, ARMA-GARCH,...

**Remarque** : Avec le Lemme de Slutsky et  $\hat{F}(\hat{\theta}_n)$  et  $\hat{G}(\hat{\theta}_n)$  au lieu de  $F(\theta^*)$  et  $G(\theta^*)$  (sans lim), on obtient des intervalles de confiance, tests de Wald,

# Convergence de l'estimateur QML

On peut montrer que :

## Théorème

Sous certaines conditions, et  $\mathbf{E}[\xi_0^4] < \infty$  même si  $(\xi_t)$  non gaussien on a

$$\sqrt{n}(\hat{\theta}_n - \theta^*) \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \mathcal{N}_d(0, F(\theta^*)^{-1} G(\theta^*) F(\theta^*)^{-1}),$$

$$\text{avec } \begin{cases} F(\theta^*) &= \lim_{n \rightarrow \infty} \frac{1}{n} \left[ \frac{\partial^2}{\partial \theta_i \partial \theta_j} \log(\hat{L}_{\theta^*}(X_1, \dots, X_n)) \right]_{i,j} \\ G(\theta^*) &= \lim_{n \rightarrow \infty} \frac{1}{n} \left[ \frac{\partial}{\partial \theta_i} \log(\hat{L}_{\theta^*}(X_1, \dots, X_n)) \frac{\partial}{\partial \theta_j} \log(\hat{L}_{\theta^*}(X_1, \dots, X_n)) \right]_{i,j} \end{cases}$$

**Exemples** : Le théorème est valable pour les processus ARMA, GARCH, APARCH, FARIMA, ARMA-GARCH,...

**Remarque** : Avec le Lemme de Slutsky et  $\hat{F}(\hat{\theta}_n)$  et  $\hat{G}(\hat{\theta}_n)$  au lieu de  $F(\theta^*)$  et  $G(\theta^*)$  (sans lim), on obtient des intervalles de confiance, tests de Wald,

# Convergence de l'estimateur QML

On peut montrer que :

## Théorème

Sous certaines conditions, et  $\mathbf{E}[\xi_0^4] < \infty$  même si  $(\xi_t)$  non gaussien on a

$$\sqrt{n}(\hat{\theta}_n - \theta^*) \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \mathcal{N}_d(0, F(\theta^*)^{-1} G(\theta^*) F(\theta^*)^{-1}),$$

$$\text{avec } \begin{cases} F(\theta^*) &= \lim_{n \rightarrow \infty} \frac{1}{n} \left[ \frac{\partial^2}{\partial \theta_i \partial \theta_j} \log(\hat{L}_{\theta^*}(X_1, \dots, X_n)) \right]_{i,j} \\ G(\theta^*) &= \lim_{n \rightarrow \infty} \frac{1}{n} \left[ \frac{\partial}{\partial \theta_i} \log(\hat{L}_{\theta^*}(X_1, \dots, X_n)) \frac{\partial}{\partial \theta_j} \log(\hat{L}_{\theta^*}(X_1, \dots, X_n)) \right]_{i,j} \end{cases}$$

**Exemples** : Le théorème est valable pour les processus ARMA, GARCH, APARCH, FARIMA, ARMA-GARCH,...

**Remarque** : Avec le Lemme de Slutsky et  $\hat{F}(\hat{\theta}_n)$  et  $\hat{G}(\hat{\theta}_n)$  au lieu de  $F(\theta^*)$  et  $G(\theta^*)$  (sans lim), on obtient des intervalles de confiance, tests de Wald,

# Convergence de l'estimateur QML

On peut montrer que :

## Théorème

Sous certaines conditions, et  $\mathbf{E}[\xi_0^4] < \infty$  même si  $(\xi_t)$  non gaussien on a

$$\sqrt{n}(\hat{\theta}_n - \theta^*) \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \mathcal{N}_d(0, F(\theta^*)^{-1} G(\theta^*) F(\theta^*)^{-1}),$$

$$\text{avec } \begin{cases} F(\theta^*) &= \lim_{n \rightarrow \infty} \frac{1}{n} \left[ \frac{\partial^2}{\partial \theta_i \partial \theta_j} \log(\hat{L}_{\theta^*}(X_1, \dots, X_n)) \right]_{i,j} \\ G(\theta^*) &= \lim_{n \rightarrow \infty} \frac{1}{n} \left[ \frac{\partial}{\partial \theta_i} \log(\hat{L}_{\theta^*}(X_1, \dots, X_n)) \frac{\partial}{\partial \theta_j} \log(\hat{L}_{\theta^*}(X_1, \dots, X_n)) \right]_{i,j} \end{cases}$$

**Exemples** : Le théorème est valable pour les processus ARMA, GARCH, APARCH, FARIMA, ARMA-GARCH,...

**Remarque** : Avec le Lemme de Slutsky et  $\hat{F}(\hat{\theta}_n)$  et  $\hat{G}(\hat{\theta}_n)$  au lieu de  $F(\theta^*)$  et  $G(\theta^*)$  (sans lim), on obtient des intervalles de confiance, tests de Wald, ...

# Plan du cours

- 1 Définitions et premières propriétés
  - Séries chronologiques
  - Stationnarité
  - Première analyse statistique de la dépendance
- 2 Tendances et saisonnalités
  - Définition et propriétés
  - Estimation semi-paramétrique de la tendance et de la saisonnalité
  - Estimation non-paramétrique de la tendance et de la saisonnalité
- 3 Exemples de modèles de séries stationnaires
  - Processus ARMA
  - Processus GARCH
  - Autres exemples de séries temporelles
- 4 Estimation, sélection de modèle, test et prédiction
  - Estimation semi-paramétrique
  - Sélection de modèles et test d'adéquation
  - Prédiction

# Sélection de modèles

**Problématique** : On suppose  $(X_1, \dots, X_n)$  trajectoire observée. Quel modèle choisir pour modéliser cette trajectoire ?

**Exemple** : Soit la trajectoire des log-rendements d'un indice financier. Est-il préférable de la modéliser par un GARCH(1, 1) ou par un ARCH(2) ? Ou bien un ARMA(3, 2) ?

## Définition

On considère  $\mathcal{M}$  une famille de modèles affines causaux et on notera  $m \in \mathcal{M}$  un modèle, vérifiant, avec  $\theta(m) \in \mathbb{R}^d$

$$X_t = M_{\theta(m)}(X_{t-1}, X_{t-2}, \dots) \xi_t + f_{\theta(m)}(X_{t-1}, X_{t-2}, \dots)$$

**Exemple** :  $\mathcal{M}$  famille des  $AR(p)$  pour  $0 \leq p \leq 10$  :  $\theta(m) = (a_1, \dots, a_{10}, \sigma_\xi^2)$

## Sélection de modèles

**Problématique** : On suppose  $(X_1, \dots, X_n)$  trajectoire observée. Quel modèle choisir pour modéliser cette trajectoire ?

**Exemple** : Soit la trajectoire des log-rendements d'un indice financier. Est-il préférable de la modéliser par un GARCH(1, 1) ou par un ARCH(2) ? Ou bien un ARMA(3, 2) ?

### Définition

On considère  $\mathcal{M}$  une famille de modèles affines causaux et on notera  $m \in \mathcal{M}$  un modèle, vérifiant, avec  $\theta(m) \in \mathbb{R}^d$

$$X_t = M_{\theta(m)}(X_{t-1}, X_{t-2}, \dots) \xi_t + f_{\theta(m)}(X_{t-1}, X_{t-2}, \dots)$$

**Exemple** :  $\mathcal{M}$  famille des  $AR(p)$  pour  $0 \leq p \leq 10$  :  $\theta(m) = (a_1, \dots, a_{10}, \sigma_\xi^2)$

## Sélection de modèles

**Problématique** : On suppose  $(X_1, \dots, X_n)$  trajectoire observée. Quel modèle choisir pour modéliser cette trajectoire ?

**Exemple** : Soit la trajectoire des log-rendements d'un indice financier. Est-il préférable de la modéliser par un GARCH(1, 1) ou par un ARCH(2) ? Ou bien un ARMA(3, 2) ?

### Définition

*On considère  $\mathcal{M}$  une famille de modèles affines causaux et on notera  $m \in \mathcal{M}$  un modèle, vérifiant, avec  $\theta(m) \in \mathbb{R}^d$*

$$X_t = M_{\theta(m)}(X_{t-1}, X_{t-2}, \dots) \xi_t + f_{\theta(m)}(X_{t-1}, X_{t-2}, \dots)$$

**Exemple** :  $\mathcal{M}$  famille des  $AR(p)$  pour  $0 \leq p \leq 10$  :  $\theta(m) = (a_1, \dots, a_{10}, \sigma_\xi^2)$

## Sélection de modèles

**Problématique** : On suppose  $(X_1, \dots, X_n)$  trajectoire observée. Quel modèle choisir pour modéliser cette trajectoire ?

**Exemple** : Soit la trajectoire des log-rendements d'un indice financier. Est-il préférable de la modéliser par un GARCH(1, 1) ou par un ARCH(2) ? Ou bien un ARMA(3, 2) ?

### Définition

On considère  $\mathcal{M}$  une famille de modèles affines causaux et on notera  $m \in \mathcal{M}$  un modèle, vérifiant, avec  $\theta(m) \in \mathbb{R}^d$

$$X_t = M_{\theta(m)}(X_{t-1}, X_{t-2}, \dots) \xi_t + f_{\theta(m)}(X_{t-1}, X_{t-2}, \dots)$$

Exemple :  $\mathcal{M}$  famille des  $AR(p)$  pour  $0 \leq p \leq 10$  :  $\theta(m) = (a_1, \dots, a_{10}, \sigma_\xi^2)$

## Sélection de modèles

**Problématique** : On suppose  $(X_1, \dots, X_n)$  trajectoire observée. Quel modèle choisir pour modéliser cette trajectoire ?

**Exemple** : Soit la trajectoire des log-rendements d'un indice financier. Est-il préférable de la modéliser par un GARCH(1, 1) ou par un ARCH(2) ? Ou bien un ARMA(3, 2) ?

### Définition

On considère  $\mathcal{M}$  une famille de modèles affines causaux et on notera  $m \in \mathcal{M}$  un modèle, vérifiant, avec  $\theta(m) \in \mathbb{R}^d$

$$X_t = M_{\theta(m)}(X_{t-1}, X_{t-2}, \dots) \xi_t + f_{\theta(m)}(X_{t-1}, X_{t-2}, \dots)$$

**Exemple** :  $\mathcal{M}$  famille des AR( $p$ ) pour  $0 \leq p \leq 10$  :  $\theta(m) = (a_1, \dots, a_{10}, \sigma_\xi^2)$

## Sélection de modèles (2)

### Définition

Pour une famille de modèles affines causaux  $\mathcal{M}$ , on peut définir :

- le critère AIC pour ce modèle :  $\widehat{AIC}(m) = -2 \widehat{L}_N(\widehat{\theta}(m)) + 2 |m|$  ;
- le critère BIC pour ce modèle :  $\widehat{BIC}(m) = -2 \widehat{L}_N(\widehat{\theta}(m)) + \log(n) |m|$

où  $|m|$  est le nombre composantes non nulles de  $\theta(m)$ .

La famille de modèle peut-être hiérarchique (exemple : AR( $p$ )) ou non (exemple : ARMA( $p, q$ ) ou ARMA( $p, q$ )+GARCH( $p', q'$ )).

### Définition

Pour une famille de modèles affines causaux  $\mathcal{M}$ , on peut définir :

$$\widehat{m}_{AIC} = \underset{m \in \mathcal{M}}{\operatorname{Argmin}} \widehat{AIC}(m) \quad \text{et} \quad \widehat{m}_{BIC} = \underset{m \in \mathcal{M}}{\operatorname{Argmin}} \widehat{BIC}(m).$$

## Sélection de modèles (2)

### Définition

Pour une famille de modèles affines causaux  $\mathcal{M}$ , on peut définir :

- le critère AIC pour ce modèle :  $\widehat{AIC}(m) = -2 \widehat{L}_N(\widehat{\theta}(m)) + 2 |m|$  ;
- le critère BIC pour ce modèle :  $\widehat{BIC}(m) = -2 \widehat{L}_N(\widehat{\theta}(m)) + \log(n) |m|$

où  $|m|$  est le nombre composantes non nulles de  $\theta(m)$ .

La famille de modèle peut-être hiérarchique (exemple : AR( $p$ )) ou non (exemple : ARMA( $p, q$ ) ou ARMA( $p, q$ )+GARCH( $p', q'$ )).

### Définition

Pour une famille de modèles affines causaux  $\mathcal{M}$ , on peut définir :

$$\widehat{m}_{AIC} = \underset{m \in \mathcal{M}}{\operatorname{Argmin}} \widehat{AIC}(m) \quad \text{et} \quad \widehat{m}_{BIC} = \underset{m \in \mathcal{M}}{\operatorname{Argmin}} \widehat{BIC}(m).$$

## Sélection de modèles (2)

### Définition

Pour une famille de modèles affines causaux  $\mathcal{M}$ , on peut définir :

- le critère AIC pour ce modèle :  $\widehat{AIC}(m) = -2 \widehat{L}_N(\widehat{\theta}(m)) + 2 |m|$  ;
- le critère BIC pour ce modèle :  $\widehat{BIC}(m) = -2 \widehat{L}_N(\widehat{\theta}(m)) + \log(n) |m|$

où  $|m|$  est le nombre composantes non nulles de  $\theta(m)$ .

La famille de modèle peut-être hiérarchique (exemple : AR( $p$ )) ou non (exemple : ARMA( $p, q$ ) ou ARMA( $p, q$ )+GARCH( $p', q'$ )).

### Définition

Pour une famille de modèles affines causaux  $\mathcal{M}$ , on peut définir :

$$\widehat{m}_{AIC} = \underset{m \in \mathcal{M}}{\operatorname{Argmin}} \widehat{AIC}(m) \quad \text{et} \quad \widehat{m}_{BIC} = \underset{m \in \mathcal{M}}{\operatorname{Argmin}} \widehat{BIC}(m).$$

## Sélection de modèles (2)

### Définition

Pour une famille de modèles affines causaux  $\mathcal{M}$ , on peut définir :

- le critère AIC pour ce modèle :  $\widehat{AIC}(m) = -2 \widehat{L}_N(\widehat{\theta}(m)) + 2 |m|$  ;
- le critère BIC pour ce modèle :  $\widehat{BIC}(m) = -2 \widehat{L}_N(\widehat{\theta}(m)) + \log(n) |m|$

où  $|m|$  est le nombre composantes non nulles de  $\theta(m)$ .

La famille de modèle peut-être hiérarchique (exemple : AR( $p$ )) ou non (exemple : ARMA( $p, q$ ) ou ARMA( $p, q$ )+GARCH( $p', q'$ )).

### Définition

Pour une famille de modèles affines causaux  $\mathcal{M}$ , on peut définir :

$$\widehat{m}_{AIC} = \underset{m \in \mathcal{M}}{\operatorname{Argmin}} \widehat{AIC}(m) \quad \text{et} \quad \widehat{m}_{BIC} = \underset{m \in \mathcal{M}}{\operatorname{Argmin}} \widehat{BIC}(m).$$

## Sélection de modèles (2)

### Définition

Pour une famille de modèles affines causaux  $\mathcal{M}$ , on peut définir :

- le critère AIC pour ce modèle :  $\widehat{AIC}(m) = -2 \widehat{L}_N(\widehat{\theta}(m)) + 2 |m|$  ;
- le critère BIC pour ce modèle :  $\widehat{BIC}(m) = -2 \widehat{L}_N(\widehat{\theta}(m)) + \log(n) |m|$

où  $|m|$  est le nombre composantes non nulles de  $\theta(m)$ .

La famille de modèle peut-être hiérarchique (exemple : AR( $p$ )) ou non (exemple : ARMA( $p, q$ ) ou ARMA( $p, q$ )+GARCH( $p', q'$ )).

### Définition

Pour une famille de modèles affines causaux  $\mathcal{M}$ , on peut définir :

$$\widehat{m}_{AIC} = \underset{m \in \mathcal{M}}{\operatorname{Argmin}} \widehat{AIC}(m) \quad \text{et} \quad \widehat{m}_{BIC} = \underset{m \in \mathcal{M}}{\operatorname{Argmin}} \widehat{BIC}(m).$$

## Sélection de modèles (2)

### Définition

Pour une famille de modèles affines causaux  $\mathcal{M}$ , on peut définir :

- le critère AIC pour ce modèle :  $\widehat{AIC}(m) = -2 \widehat{L}_N(\widehat{\theta}(m)) + 2 |m|$  ;
- le critère BIC pour ce modèle :  $\widehat{BIC}(m) = -2 \widehat{L}_N(\widehat{\theta}(m)) + \log(n) |m|$

où  $|m|$  est le nombre composantes non nulles de  $\theta(m)$ .

La famille de modèle peut-être hiérarchique (exemple : AR( $p$ )) ou non (exemple : ARMA( $p, q$ ) ou ARMA( $p, q$ )+GARCH( $p', q'$ )).

### Définition

Pour une famille de modèles affines causaux  $\mathcal{M}$ , on peut définir :

$$\widehat{m}_{AIC} = \underset{m \in \mathcal{M}}{\operatorname{Argmin}} \widehat{AIC}(m) \quad \text{et} \quad \widehat{m}_{BIC} = \underset{m \in \mathcal{M}}{\operatorname{Argmin}} \widehat{BIC}(m).$$

## Sélection de modèles (2)

### Définition

Pour une famille de modèles affines causaux  $\mathcal{M}$ , on peut définir :

- le critère AIC pour ce modèle :  $\widehat{AIC}(m) = -2 \widehat{L}_N(\widehat{\theta}(m)) + 2 |m|$  ;
- le critère BIC pour ce modèle :  $\widehat{BIC}(m) = -2 \widehat{L}_N(\widehat{\theta}(m)) + \log(n) |m|$

où  $|m|$  est le nombre composantes non nulles de  $\theta(m)$ .

La famille de modèle peut-être hiérarchique (exemple : AR( $p$ )) ou non (exemple : ARMA( $p, q$ ) ou ARMA( $p, q$ )+GARCH( $p', q'$ )).

### Définition

Pour une famille de modèles affines causaux  $\mathcal{M}$ , on peut définir :

$$\widehat{m}_{AIC} = \underset{m \in \mathcal{M}}{\operatorname{Argmin}} \widehat{AIC}(m) \quad \text{et} \quad \widehat{m}_{BIC} = \underset{m \in \mathcal{M}}{\operatorname{Argmin}} \widehat{BIC}(m).$$

## Sélection de modèles (3)

### Théorème

Soit  $\mathcal{M}$  famille de taille finie de modèles affines causaux. S'il existe un "vrai" modèle  $m^* \in \mathcal{M}$  tel que  $(X_1, \dots, X_n)$  trajectoire observée issue de  $m^*$  et sous certaines conditions,

$$\mathbb{P}(\hat{m}_{BIC} = m^*) \xrightarrow{n \rightarrow +\infty} 1 \quad \text{et} \quad \mathbb{P}(m^* \subset \hat{m}_{AIC}) \xrightarrow{n \rightarrow +\infty} 1 \quad \text{et}$$

$$\sqrt{n} (\hat{\theta}_n(\hat{m}_{BIC}) - \theta(m^*)) \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \mathcal{N}_{|m^*|}(0, F(\theta(m^*))^{-1} G(\theta(m^*)) F(\theta(m^*))^{-1})$$

**Exemples :** Le théorème est valable pour les processus ARMA, GARCH, APARCH, FARIMA, ARMA-GARCH,...

**Remarque :** AIC reste intéressant quand  $\nexists m^* \in \mathcal{M}$

## Sélection de modèles (3)

### Théorème

Soit  $\mathcal{M}$  famille de taille finie de modèles affines causaux. S'il existe un "vrai" modèle  $m^* \in \mathcal{M}$  tel que  $(X_1, \dots, X_n)$  trajectoire observée issue de  $m^*$  et sous certaines conditions,

$$\mathbb{P}(\hat{m}_{BIC} = m^*) \xrightarrow{n \rightarrow +\infty} 1 \quad \text{et} \quad \mathbb{P}(m^* \subset \hat{m}_{AIC}) \xrightarrow{n \rightarrow +\infty} 1 \quad \text{et}$$

$$\sqrt{n}(\hat{\theta}_n(\hat{m}_{BIC}) - \theta(m^*)) \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \mathcal{N}_{|m^*|}(0, F(\theta(m^*))^{-1} G(\theta(m^*)) F(\theta(m^*))^{-1})$$

**Exemples** : Le théorème est valable pour les processus ARMA, GARCH, APARCH, FARIMA, ARMA-GARCH,...

**Remarque** : AIC reste intéressant quand  $\nexists m^* \in \mathcal{M}$

## Sélection de modèles (3)

### Théorème

Soit  $\mathcal{M}$  famille de taille finie de modèles affines causaux. S'il existe un "vrai" modèle  $m^* \in \mathcal{M}$  tel que  $(X_1, \dots, X_n)$  trajectoire observée issue de  $m^*$  et sous certaines conditions,

$$\mathbb{P}(\hat{m}_{BIC} = m^*) \xrightarrow{n \rightarrow +\infty} 1 \quad \text{et} \quad \mathbb{P}(m^* \subset \hat{m}_{AIC}) \xrightarrow{n \rightarrow +\infty} 1 \quad \text{et}$$

$$\sqrt{n}(\hat{\theta}_n(\hat{m}_{BIC}) - \theta(m^*)) \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \mathcal{N}_{|m^*|}(0, F(\theta(m^*))^{-1} G(\theta(m^*)) F(\theta(m^*))^{-1})$$

**Exemples :** Le théorème est valable pour les processus ARMA, GARCH, APARCH, FARIMA, ARMA-GARCH,...

**Remarque :** AIC reste intéressant quand  $\nexists m^* \in \mathcal{M}$

## Sélection de modèles (3)

### Théorème

Soit  $\mathcal{M}$  famille de taille finie de modèles affines causaux. S'il existe un "vrai" modèle  $m^* \in \mathcal{M}$  tel que  $(X_1, \dots, X_n)$  trajectoire observée issue de  $m^*$  et sous certaines conditions,

$$\mathbb{P}(\hat{m}_{BIC} = m^*) \xrightarrow{n \rightarrow +\infty} 1 \quad \text{et} \quad \mathbb{P}(m^* \subset \hat{m}_{AIC}) \xrightarrow{n \rightarrow +\infty} 1 \quad \text{et}$$

$$\sqrt{n}(\hat{\theta}_n(\hat{m}_{BIC}) - \theta(m^*)) \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \mathcal{N}_{|m^*|}(0, F(\theta(m^*))^{-1} G(\theta(m^*)) F(\theta(m^*))^{-1})$$

**Exemples** : Le théorème est valable pour les processus ARMA, GARCH, APARCH, FARIMA, ARMA-GARCH,...

Remarque : AIC reste intéressant quand  $\nexists m^* \in \mathcal{M}$

## Sélection de modèles (3)

### Théorème

Soit  $\mathcal{M}$  famille de taille finie de modèles affines causaux. S'il existe un "vrai" modèle  $m^* \in \mathcal{M}$  tel que  $(X_1, \dots, X_n)$  trajectoire observée issue de  $m^*$  et sous certaines conditions,

$$\mathbb{P}(\hat{m}_{BIC} = m^*) \xrightarrow{n \rightarrow +\infty} 1 \quad \text{et} \quad \mathbb{P}(m^* \subset \hat{m}_{AIC}) \xrightarrow{n \rightarrow +\infty} 1 \quad \text{et}$$

$$\sqrt{n}(\hat{\theta}_n(\hat{m}_{BIC}) - \theta(m^*)) \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \mathcal{N}_{|m^*|}(0, F(\theta(m^*))^{-1} G(\theta(m^*)) F(\theta(m^*))^{-1})$$

**Exemples** : Le théorème est valable pour les processus ARMA, GARCH, APARCH, FARIMA, ARMA-GARCH,...

**Remarque** : AIC reste intéressant quand  $\nexists m^* \in \mathcal{M}$

# Test d'adéquation

**Problématique** : Une fois un modèle  $\hat{m}_{BIC}$  obtenu pour modéliser  $X_1, \dots, X_n$  comment s'assurer que ce modèle convient ?

⇒ Test d'adéquation de type portemanteau

$$\begin{cases} H_0 : \exists m^* \in \mathcal{M}, \text{ tel que } (X_1, \dots, X_n) \text{ trajectoire de } X \text{ avec } \theta(m^*) \\ H_1 : \nexists m^* \in \mathcal{M}, \text{ tel que } (X_1, \dots, X_n) \text{ trajectoire de } X \text{ avec } \theta(m^*) \end{cases}$$

# Test d'adéquation

**Problématique** : Une fois un modèle  $\hat{m}_{BIC}$  obtenu pour modéliser  $(X_1, \dots, X_n)$  comment s'assurer que ce modèle convient ?

⇒ Test d'adéquation de type portemanteau

$$\begin{cases} H_0 : \exists m^* \in \mathcal{M}, \text{ tel que } (X_1, \dots, X_n) \text{ trajectoire de } X \text{ avec } \theta(m^*) \\ H_1 : \nexists m^* \in \mathcal{M}, \text{ tel que } (X_1, \dots, X_n) \text{ trajectoire de } X \text{ avec } \theta(m^*) \end{cases}$$

# Test d'adéquation

**Problématique** : Une fois un modèle  $\hat{m}_{BIC}$  obtenu pour modéliser  $X_1, \dots, X_n$  comment s'assurer que ce modèle convient ?

⇒ Test d'adéquation de type portemanteau

$$\begin{cases} H_0 : \exists m^* \in \mathcal{M}, \text{ tel que } (X_1, \dots, X_n) \text{ trajectoire de } X \text{ avec } \theta(m^*) \\ H_1 : \nexists m^* \in \mathcal{M}, \text{ tel que } (X_1, \dots, X_n) \text{ trajectoire de } X \text{ avec } \theta(m^*) \end{cases}$$

# Test d'adéquation

**Problématique** : Une fois un modèle  $\hat{m}_{BIC}$  obtenu pour modéliser  $(X_1, \dots, X_n)$  comment s'assurer que ce modèle convient ?

⇒ Test d'adéquation de type portemanteau

$$\begin{cases} H_0 : \exists m^* \in \mathcal{M}, \text{ tel que } (X_1, \dots, X_n) \text{ trajectoire de } X \text{ avec } \theta(m^*) \\ H_1 : \nexists m^* \in \mathcal{M}, \text{ tel que } (X_1, \dots, X_n) \text{ trajectoire de } X \text{ avec } \theta(m^*) \end{cases} .$$

## Test d'adéquation (2)

On définit les résidus pour le modèle  $m$  :

$$\hat{e}_t(m) := (\hat{M}_{\hat{\theta}(m)}^t)^{-1} (X_t - \hat{f}_{\hat{\theta}(m)}^t).$$

Pour  $K \in \mathbb{N}^*$  fixé, vecteur d'autocorrélations des carrés des résidus :

$$\hat{\rho}(m) := (\hat{\rho}_1(m), \dots, \hat{\rho}_K(m))',$$

$$\text{où } \hat{\rho}_k(m) := \frac{\hat{\gamma}_k(m)}{\hat{\gamma}_0(m)} \text{ et } \hat{\gamma}_k(m) := \frac{1}{n} \sum_{t=k+1}^n (\hat{e}_t^2(m) - 1)(\hat{e}_{t-k}^2(m) - 1)$$

### Théorème

*Sous certaines hypothèses, il existe une matrice  $V(\theta(m^*))$*

$$\sqrt{n} \hat{\rho}(\hat{m}_{BIC}) \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \mathcal{N}_K(0, V(\theta(m^*))).$$

## Test d'adéquation (2)

On définit les résidus pour le modèle  $m$  :

$$\hat{e}_t(m) := (\hat{M}_{\hat{\theta}(m)}^t)^{-1} (X_t - \hat{f}_{\hat{\theta}(m)}^t).$$

Pour  $K \in \mathbb{N}^*$  fixé, vecteur d'autocorrélations des carrés des résidus :

$$\hat{\rho}(m) := (\hat{\rho}_1(m), \dots, \hat{\rho}_K(m))',$$

$$\text{où } \hat{\rho}_k(m) := \frac{\hat{\gamma}_k(m)}{\hat{\gamma}_0(m)} \text{ et } \hat{\gamma}_k(m) := \frac{1}{n} \sum_{t=k+1}^n (\hat{e}_t^2(m) - 1)(\hat{e}_{t-k}^2(m) - 1)$$

### Théorème

*Sous certaines hypothèses, il existe une matrice  $V(\theta(m^*))$*

$$\sqrt{n} \hat{\rho}(\hat{m}_{BIC}) \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \mathcal{N}_K(0, V(\theta(m^*))).$$

## Test d'adéquation (2)

On définit les résidus pour le modèle  $m$  :

$$\hat{e}_t(m) := (\hat{M}_{\hat{\theta}(m)}^t)^{-1} (X_t - \hat{f}_{\hat{\theta}(m)}^t).$$

Pour  $K \in \mathbb{N}^*$  fixé, vecteur d'autocorrélations des carrés des résidus :

$$\hat{\rho}(m) := (\hat{\rho}_1(m), \dots, \hat{\rho}_K(m))',$$

$$\text{où } \hat{\rho}_k(m) := \frac{\hat{\gamma}_k(m)}{\hat{\gamma}_0(m)} \text{ et } \hat{\gamma}_k(m) := \frac{1}{n} \sum_{t=k+1}^n (\hat{e}_t^2(m) - 1)(\hat{e}_{t-k}^2(m) - 1)$$

### Théorème

*Sous certaines hypothèses, il existe une matrice  $V(\theta(m^*))$*

$$\sqrt{n} \hat{\rho}(\hat{m}_{BIC}) \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \mathcal{N}_K(0, V(\theta(m^*))).$$

## Test d'adéquation (2)

On définit les résidus pour le modèle  $m$  :

$$\hat{e}_t(m) := (\hat{M}_{\hat{\theta}(m)}^t)^{-1} (X_t - \hat{f}_{\hat{\theta}(m)}^t).$$

Pour  $K \in \mathbb{N}^*$  fixé, vecteur d'autocorrélations des carrés des résidus :

$$\hat{\rho}(m) := (\hat{\rho}_1(m), \dots, \hat{\rho}_K(m))',$$

$$\text{où } \hat{\rho}_k(m) := \frac{\hat{\gamma}_k(m)}{\hat{\gamma}_0(m)} \text{ et } \hat{\gamma}_k(m) := \frac{1}{n} \sum_{t=k+1}^n (\hat{e}_t^2(m) - 1)(\hat{e}_{t-k}^2(m) - 1)$$

### Théorème

*Sous certaines hypothèses, il existe une matrice  $V(\theta(m^*))$*

$$\sqrt{n} \hat{\rho}(\hat{m}_{BIC}) \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \mathcal{N}_K(0, V(\theta(m^*))).$$

## Test d'adéquation (2)

On définit les résidus pour le modèle  $m$  :

$$\hat{e}_t(m) := (\hat{M}_{\hat{\theta}(m)}^t)^{-1} (X_t - \hat{f}_{\hat{\theta}(m)}^t).$$

Pour  $K \in \mathbb{N}^*$  fixé, vecteur d'autocorrélations des carrés des résidus :

$$\hat{\rho}(m) := (\hat{\rho}_1(m), \dots, \hat{\rho}_K(m))',$$

$$\text{où } \hat{\rho}_k(m) := \frac{\hat{\gamma}_k(m)}{\hat{\gamma}_0(m)} \text{ et } \hat{\gamma}_k(m) := \frac{1}{n} \sum_{t=k+1}^n (\hat{e}_t^2(m) - 1)(\hat{e}_{t-k}^2(m) - 1)$$

### Théorème

*Sous certaines hypothèses, il existe une matrice  $V(\theta(m^*))$*

$$\sqrt{n} \hat{\rho}(\hat{m}_{BIC}) \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \mathcal{N}_K(0, V(\theta(m^*))).$$

## Test d'adéquation (2)

On définit les résidus pour le modèle  $m$  :

$$\hat{e}_t(m) := (\hat{M}_{\hat{\theta}(m)}^t)^{-1} (X_t - \hat{f}_{\hat{\theta}(m)}^t).$$

Pour  $K \in \mathbb{N}^*$  fixé, vecteur d'autocorrélations des carrés des résidus :

$$\hat{\rho}(m) := (\hat{\rho}_1(m), \dots, \hat{\rho}_K(m))',$$

$$\text{où } \hat{\rho}_k(m) := \frac{\hat{\gamma}_k(m)}{\hat{\gamma}_0(m)} \text{ et } \hat{\gamma}_k(m) := \frac{1}{n} \sum_{t=k+1}^n (\hat{e}_t^2(m) - 1)(\hat{e}_{t-k}^2(m) - 1)$$

### Théorème

*Sous certaines hypothèses, il existe une matrice  $V(\theta(m^*))$*

$$\sqrt{n} \hat{\rho}(\hat{m}_{BIC}) \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \mathcal{N}_K(0, V(\theta(m^*))).$$

## Test d'adéquation (2)

On définit les résidus pour le modèle  $m$  :

$$\hat{e}_t(m) := (\hat{M}_{\hat{\theta}(m)}^t)^{-1} (X_t - \hat{f}_{\hat{\theta}(m)}^t).$$

Pour  $K \in \mathbb{N}^*$  fixé, vecteur d'autocorrélations des carrés des résidus :

$$\hat{\rho}(m) := (\hat{\rho}_1(m), \dots, \hat{\rho}_K(m))',$$

$$\text{où } \hat{\rho}_k(m) := \frac{\hat{\gamma}_k(m)}{\hat{\gamma}_0(m)} \text{ et } \hat{\gamma}_k(m) := \frac{1}{n} \sum_{t=k+1}^n (\hat{e}_t^2(m) - 1)(\hat{e}_{t-k}^2(m) - 1)$$

### Théorème

*Sous certaines hypothèses, il existe une matrice  $V(\theta(m^*))$*

$$\sqrt{n} \hat{\rho}(\hat{m}_{BIC}) \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \mathcal{N}_K(0, V(\theta(m^*))).$$

## Test d'adéquation (3)

On en déduit la statistique de test :

$$\widehat{Q}_K(\widehat{m}_{BIC}) := n^t \widehat{\rho}(\widehat{m}_{BIC}) \widehat{V}^{-1} \widehat{\rho}(\widehat{m}_{BIC})$$

$$\text{où } \begin{cases} \widehat{V} = I_K + \frac{1}{(\widehat{\mu}_4 - 1)^2} \widehat{J}_K F(\theta(\widehat{m}_{BIC}))^{-1} \left( G(\theta(\widehat{m}_{BIC})) F(\theta(\widehat{m}_{BIC}))^{-1} + (\widehat{\mu}_4 - 1) I_K \right)^t \widehat{J}_K \\ \widehat{J}_K = \frac{1}{n} \sum_{t=1}^{n-i} (\widehat{e}_t^2(\widehat{m}_{BIC}) - 1) \left( \frac{\partial \log(M_{\widehat{\theta}(\widehat{m}_{BIC})}^{t+i})}{\partial \theta_j} \right)_{1 \leq i \leq K, j}, \quad \widehat{\mu}_4 = \frac{1}{n} \sum_{t=1}^n \widehat{e}_t^4(\widehat{m}_{BIC}) \end{cases}$$

### Théorème

Sous l'hypothèse  $H_0$  et sous certaines conditions,  $\widehat{Q}_K(\widehat{m}_{BIC}) \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \chi^2(K)$ .

**Remarque :** • Le test est utilisable pour les processus ARMA, GARCH, APARCH, FARIMA, ARMA-GARCH,...

- Pour un ARMA,  $\widehat{Q}_K(\widehat{m}_{BIC})$  peut être remplacé par  $n \|\widehat{\rho}(\widehat{m}_{BIC})\|^2$
- Pour un GARCH,  $\widehat{V} \simeq I_K - \frac{1}{4} \widehat{J}_K (\widehat{G}(\widehat{\theta}(\widehat{m}_{BIC})))^{-1} \widehat{J}_K$

## Test d'adéquation (3)

On en déduit la statistique de test :

$$\widehat{Q}_K(\widehat{m}_{BIC}) := n^t \widehat{\rho}(\widehat{m}_{BIC}) \widehat{V}^{-1} \widehat{\rho}(\widehat{m}_{BIC})$$

$$\text{où } \begin{cases} \widehat{V} = I_K + \frac{1}{(\widehat{\mu}_4 - 1)^2} \widehat{J}_K F(\theta(\widehat{m}_{BIC}))^{-1} \left( G(\theta(\widehat{m}_{BIC})) F(\theta(\widehat{m}_{BIC}))^{-1} + (\widehat{\mu}_4 - 1) I_K \right)^t \widehat{J}_K \\ \widehat{J}_K = \frac{1}{n} \sum_{t=1}^{n-i} (\widehat{e}_t^2(\widehat{m}_{BIC}) - 1) \left( \frac{\partial \log(M^t)}{\partial \theta_j} \right)_{1 \leq i \leq K, j}, \quad \widehat{\mu}_4 = \frac{1}{n} \sum_{t=1}^n \widehat{e}_t^4(\widehat{m}_{BIC}) \end{cases}$$

### Théorème

Sous l'hypothèse  $H_0$  et sous certaines conditions,  $\widehat{Q}_K(\widehat{m}_{BIC}) \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \chi^2(K)$ .

**Remarque :** • Le test est utilisable pour les processus ARMA, GARCH, APARCH, FARIMA, ARMA-GARCH,...

- Pour un ARMA,  $\widehat{Q}_K(\widehat{m}_{BIC})$  peut être remplacé par  $n \|\widehat{\rho}(\widehat{m}_{BIC})\|^2$
- Pour un GARCH,  $\widehat{V} \simeq I_K - \frac{1}{4} \widehat{J}_K (\widehat{G}(\widehat{\theta}(\widehat{m}_{BIC})))^{-1} \widehat{J}_K$

## Test d'adéquation (3)

On en déduit la statistique de test :

$$\widehat{Q}_K(\widehat{m}_{BIC}) := n^t \widehat{\rho}(\widehat{m}_{BIC}) \widehat{V}^{-1} \widehat{\rho}(\widehat{m}_{BIC})$$

$$\text{où } \begin{cases} \widehat{V} = I_K + \frac{1}{(\widehat{\mu}_4 - 1)^2} \widehat{J}_K F(\theta(\widehat{m}_{BIC}))^{-1} \left( G(\theta(\widehat{m}_{BIC})) F(\theta(\widehat{m}_{BIC}))^{-1} + (\widehat{\mu}_4 - 1) I_K \right)^t \widehat{J}_K \\ \widehat{J}_K = \frac{1}{n} \sum_{t=1}^{n-i} (\widehat{e}_t^2(\widehat{m}_{BIC}) - 1) \left( \frac{\partial \log(M_{\widehat{\theta}(\widehat{m}_{BIC})}^{t+i})}{\partial \theta_j} \right)_{1 \leq i \leq K, j}, \quad \widehat{\mu}_4 = \frac{1}{n} \sum_{t=1}^n \widehat{e}_t^4(\widehat{m}_{BIC}) \end{cases}$$

### Théorème

Sous l'hypothèse  $H_0$  et sous certaines conditions,  $\widehat{Q}_K(\widehat{m}_{BIC}) \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \chi^2(K)$ .

**Remarque :** • Le test est utilisable pour les processus ARMA, GARCH, APARCH, FARIMA, ARMA-GARCH,...

- Pour un ARMA,  $\widehat{Q}_K(\widehat{m}_{BIC})$  peut être remplacé par  $n \|\widehat{\rho}(\widehat{m}_{BIC})\|^2$
- Pour un GARCH,  $\widehat{V} \simeq I_K - \frac{1}{4} \widehat{J}_K (\widehat{G}(\widehat{\theta}(\widehat{m}_{BIC})))^{-1} \widehat{J}_K$

## Test d'adéquation (3)

On en déduit la statistique de test :

$$\hat{Q}_K(\hat{m}_{BIC}) := n^t \hat{\rho}(\hat{m}_{BIC}) \hat{V}^{-1} \hat{\rho}(\hat{m}_{BIC})$$

$$\text{où } \begin{cases} \hat{V} = I_K + \frac{1}{(\hat{\mu}_4 - 1)^2} \hat{J}_K F(\theta(\hat{m}_{BIC}))^{-1} \left( G(\theta(\hat{m}_{BIC})) F(\theta(\hat{m}_{BIC}))^{-1} + (\hat{\mu}_4 - 1) I_K \right)^t \hat{J}_K \\ \hat{J}_K = \frac{1}{n} \sum_{t=1}^{n-i} (\hat{e}_t^2(\hat{m}_{BIC}) - 1) \left( \frac{\partial \log(M_{\hat{\theta}(\hat{m}_{BIC})}^{t+i})}{\partial \theta_j} \right)_{1 \leq i \leq K, j}, \quad \hat{\mu}_4 = \frac{1}{n} \sum_{t=1}^n \hat{e}_t^4(\hat{m}_{BIC}) \end{cases}$$

### Théorème

Sous l'hypothèse  $H_0$  et sous certaines conditions,  $\hat{Q}_K(\hat{m}_{BIC}) \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \chi^2(K)$ .

Remarque : • Le test est utilisable pour les processus ARMA, GARCH, APARCH, FARIMA, ARMA-GARCH,...

- Pour un ARMA,  $\hat{Q}_K(\hat{m}_{BIC})$  peut être remplacé par  $n \|\hat{\rho}(\hat{m}_{BIC})\|^2$
- Pour un GARCH,  $\hat{V} \simeq I_K - \frac{1}{4} \hat{J}_K (\hat{G}(\hat{\theta}(\hat{m}_{BIC})))^{-1} \hat{J}_K$

## Test d'adéquation (3)

On en déduit la statistique de test :

$$\widehat{Q}_K(\widehat{m}_{BIC}) := n {}^t \widehat{\rho}(\widehat{m}_{BIC}) \widehat{V}^{-1} \widehat{\rho}(\widehat{m}_{BIC})$$

$$\text{où } \begin{cases} \widehat{V} = I_K + \frac{1}{(\widehat{\mu}_4 - 1)^2} \widehat{J}_K F(\theta(\widehat{m}_{BIC}))^{-1} \left( G(\theta(\widehat{m}_{BIC})) F(\theta(\widehat{m}_{BIC}))^{-1} + (\widehat{\mu}_4 - 1) I_K \right) {}^t \widehat{J}_K \\ \widehat{J}_K = \frac{1}{n} \sum_{t=1}^{n-i} (\widehat{e}_t^2(\widehat{m}_{BIC}) - 1) \left( \frac{\partial \log(M_{\widehat{\theta}(\widehat{m}_{BIC})}^{t+i})}{\partial \theta_j} \right)_{1 \leq i \leq K, j}, \quad \widehat{\mu}_4 = \frac{1}{n} \sum_{t=1}^n \widehat{e}_t^4(\widehat{m}_{BIC}) \end{cases}$$

### Théorème

Sous l'hypothèse  $H_0$  et sous certaines conditions,  $\widehat{Q}_K(\widehat{m}_{BIC}) \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \chi^2(K)$ .

**Remarque :** • Le test est utilisable pour les processus ARMA, GARCH, APARCH, FARIMA, ARMA-GARCH,...

- Pour un ARMA,  $\widehat{Q}_K(\widehat{m}_{BIC})$  peut être remplacé par  $n \|\widehat{\rho}(\widehat{m}_{BIC})\|^2$
- Pour un GARCH,  $\widehat{V} \simeq I_K - \frac{1}{4} \widehat{J}_K (\widehat{G}(\widehat{\theta}(\widehat{m}_{BIC})))^{-1} {}^t \widehat{J}_K$

## Test d'adéquation (3)

On en déduit la statistique de test :

$$\hat{Q}_K(\hat{m}_{BIC}) := n {}^t \hat{\rho}(\hat{m}_{BIC}) \hat{V}^{-1} \hat{\rho}(\hat{m}_{BIC})$$

$$\text{où } \begin{cases} \hat{V} = I_K + \frac{1}{(\hat{\mu}_4 - 1)^2} \hat{J}_K F(\theta(\hat{m}_{BIC}))^{-1} (G(\theta(\hat{m}_{BIC})) F(\theta(\hat{m}_{BIC}))^{-1} + (\hat{\mu}_4 - 1) I_K) {}^t \hat{J}_K \\ \hat{J}_K = \frac{1}{n} \sum_{t=1}^{n-1} (\hat{e}_t^2(\hat{m}_{BIC}) - 1) \left( \frac{\partial \log(M_{\hat{\theta}(\hat{m}_{BIC})}^{t+i})}{\partial \theta_j} \right)_{1 \leq i \leq K, j}, \quad \hat{\mu}_4 = \frac{1}{n} \sum_{t=1}^n \hat{e}_t^4(\hat{m}_{BIC}) \end{cases}$$

### Théorème

Sous l'hypothèse  $H_0$  et sous certaines conditions,  $\hat{Q}_K(\hat{m}_{BIC}) \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \chi^2(K)$ .

**Remarque :** • Le test est utilisable pour les processus ARMA, GARCH, APARCH, FARIMA, ARMA-GARCH,...

- Pour un ARMA,  $\hat{Q}_K(\hat{m}_{BIC})$  peut être remplacé par  $n \|\hat{\rho}(\hat{m}_{BIC})\|^2$
- Pour un GARCH,  $\hat{V} \simeq I_K - \frac{1}{4} \hat{J}_K (\hat{G}(\hat{\theta}(\hat{m}_{BIC})))^{-1} {}^t \hat{J}_K$

## Test d'adéquation (3)

On en déduit la statistique de test :

$$\hat{Q}_K(\hat{m}_{BIC}) := n^t \hat{\rho}(\hat{m}_{BIC}) \hat{V}^{-1} \hat{\rho}(\hat{m}_{BIC})$$

$$\text{où } \begin{cases} \hat{V} = I_K + \frac{1}{(\hat{\mu}_4 - 1)^2} \hat{J}_K F(\theta(\hat{m}_{BIC}))^{-1} \left( G(\theta(\hat{m}_{BIC})) F(\theta(\hat{m}_{BIC}))^{-1} + (\hat{\mu}_4 - 1) I_K \right)^t \hat{J}_K \\ \hat{J}_K = \frac{1}{n} \sum_{t=1}^{n-i} (\hat{e}_t^2(\hat{m}_{BIC}) - 1) \left( \frac{\partial \log(M_{\hat{\theta}(\hat{m}_{BIC})}^{t+i})}{\partial \theta_j} \right)_{1 \leq i \leq K, j}, \quad \hat{\mu}_4 = \frac{1}{n} \sum_{t=1}^n \hat{e}_t^4(\hat{m}_{BIC}) \end{cases}$$

### Théorème

Sous l'hypothèse  $H_0$  et sous certaines conditions,  $\hat{Q}_K(\hat{m}_{BIC}) \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \chi^2(K)$ .

**Remarque :** • Le test est utilisable pour les processus ARMA, GARCH, APARCH, FARIMA, ARMA-GARCH,...

- Pour un ARMA,  $\hat{Q}_K(\hat{m}_{BIC})$  peut être remplacé par  $n \|\hat{\rho}(\hat{m}_{BIC})\|^2$
- Pour un GARCH,  $\hat{V} \simeq I_K - \frac{1}{4} \hat{J}_K (\hat{G}(\hat{\theta}(\hat{m}_{BIC})))^{-1} \hat{J}_K$

## Test d'adéquation (3)

On en déduit la statistique de test :

$$\hat{Q}_K(\hat{m}_{BIC}) := n^t \hat{\rho}(\hat{m}_{BIC}) \hat{V}^{-1} \hat{\rho}(\hat{m}_{BIC})$$

$$\text{où } \begin{cases} \hat{V} = I_K + \frac{1}{(\hat{\mu}_4 - 1)^2} \hat{J}_K F(\theta(\hat{m}_{BIC}))^{-1} \left( G(\theta(\hat{m}_{BIC})) F(\theta(\hat{m}_{BIC}))^{-1} + (\hat{\mu}_4 - 1) I_K \right)^t \hat{J}_K \\ \hat{J}_K = \frac{1}{n} \sum_{t=1}^{n-i} (\hat{e}_t^2(\hat{m}_{BIC}) - 1) \left( \frac{\partial \log(M_{\hat{\theta}(\hat{m}_{BIC})}^{t+i})}{\partial \theta_j} \right)_{1 \leq i \leq K, j}, \quad \hat{\mu}_4 = \frac{1}{n} \sum_{t=1}^n \hat{e}_t^4(\hat{m}_{BIC}) \end{cases}$$

### Théorème

Sous l'hypothèse  $H_0$  et sous certaines conditions,  $\hat{Q}_K(\hat{m}_{BIC}) \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \chi^2(K)$ .

**Remarque :** • Le test est utilisable pour les processus ARMA, GARCH, APARCH, FARIMA, ARMA-GARCH,...

- Pour un ARMA,  $\hat{Q}_K(\hat{m}_{BIC})$  peut être remplacé par  $n \|\hat{\rho}(\hat{m}_{BIC})\|^2$
- Pour un GARCH,  $\hat{V} \simeq I_K - \frac{1}{4} \hat{J}_K (\hat{G}(\hat{\theta}(\hat{m}_{BIC})))^{-1} \hat{J}_K$

# Plan du cours

- 1 Définitions et premières propriétés
  - Séries chronologiques
  - Stationnarité
  - Première analyse statistique de la dépendance
- 2 Tendances et saisonnalités
  - Définition et propriétés
  - Estimation semi-paramétrique de la tendance et de la saisonnalité
  - Estimation non-paramétrique de la tendance et de la saisonnalité
- 3 Exemples de modèles de séries stationnaires
  - Processus ARMA
  - Processus GARCH
  - Autres exemples de séries temporelles
- 4 Estimation, sélection de modèle, test et prédiction
  - Estimation semi-paramétrique
  - Sélection de modèles et test d'adéquation
  - Prédiction

# Prédiction à partir d'un modèle

On suppose observée  $(X_1, \dots, X_n)$ .

**Objectif** : On veut prédire  $X_{n+h}$ , soit  $\hat{X}_{n+h}$  ( $h$  horizon)

**Méthode 1** : On suppose que  $X_t = a(t) + s(t) + u_t$  pour  $t \in \mathbb{Z}$  et :

- On a obtenu  $\hat{a}_n(t)$  et  $\hat{s}_n(t)$  à partir de  $(X_1, \dots, X_n)$
- Pour  $\hat{u}_t = X_t - \hat{a}_n(t) - \hat{s}_n(t)$  on a obtenu et validé un modèle

$$\hat{u}_t = M_{\hat{\theta}}(\hat{u}_{t-1}, \hat{u}_{t-2}, \dots, \hat{u}_1, 0, \dots) \xi_t + f_{\hat{\theta}}(\hat{u}_{t-1}, \hat{u}_{t-2}, \dots, \hat{u}_1, 0, \dots)$$

# Prédiction à partir d'un modèle

On suppose observée  $(X_1, \dots, X_n)$ .

Objectif : On veut prédire  $X_{n+h}$ , soit  $\hat{X}_{n+h}$  ( $h$  horizon)

Méthode 1 : On suppose que  $X_t = a(t) + s(t) + u_t$  pour  $t \in \mathbb{Z}$  et :

- On a obtenu  $\hat{a}_n(t)$  et  $\hat{s}_n(t)$  à partir de  $(X_1, \dots, X_n)$
- Pour  $\hat{u}_t = X_t - \hat{a}_n(t) - \hat{s}_n(t)$  on a obtenu et validé un modèle

$$\hat{u}_t = M_{\hat{\theta}}(\hat{u}_{t-1}, \hat{u}_{t-2}, \dots, \hat{u}_1, 0, \dots) \xi_t + f_{\hat{\theta}}(\hat{u}_{t-1}, \hat{u}_{t-2}, \dots, \hat{u}_1, 0, \dots)$$

# Prédiction à partir d'un modèle

On suppose observée  $(X_1, \dots, X_n)$ .

**Objectif** : On veut prédire  $X_{n+h}$ , soit  $\hat{X}_{n+h}$  ( $h$  horizon)

**Méthode 1** : On suppose que  $X_t = a(t) + s(t) + u_t$  pour  $t \in \mathbb{Z}$  et :

- On a obtenu  $\hat{a}_n(t)$  et  $\hat{s}_n(t)$  à partir de  $(X_1, \dots, X_n)$
- Pour  $\hat{u}_t = X_t - \hat{a}_n(t) - \hat{s}_n(t)$  on a obtenu et validé un modèle

$$\hat{u}_t = M_{\hat{\theta}}(\hat{u}_{t-1}, \hat{u}_{t-2}, \dots, \hat{u}_1, 0, \dots) \xi_t + f_{\hat{\theta}}(\hat{u}_{t-1}, \hat{u}_{t-2}, \dots, \hat{u}_1, 0, \dots)$$

# Prédiction à partir d'un modèle

On suppose observée  $(X_1, \dots, X_n)$ .

**Objectif** : On veut prédire  $X_{n+h}$ , soit  $\hat{X}_{n+h}$  ( $h$  horizon)

**Méthode 1** : On suppose que  $X_t = a(t) + s(t) + u_t$  pour  $t \in \mathbb{Z}$  et :

- On a obtenu  $\hat{a}_n(t)$  et  $\hat{s}_n(t)$  à partir de  $(X_1, \dots, X_n)$
- Pour  $\hat{u}_t = X_t - \hat{a}_n(t) - \hat{s}_n(t)$  on a obtenu et validé un modèle

$$\hat{u}_t = M_{\hat{\theta}}(\hat{u}_{t-1}, \hat{u}_{t-2}, \dots, \hat{u}_1, 0, \dots) \xi_t + f_{\hat{\theta}}(\hat{u}_{t-1}, \hat{u}_{t-2}, \dots, \hat{u}_1, 0, \dots)$$

# Prédiction à partir d'un modèle

On suppose observée  $(X_1, \dots, X_n)$ .

**Objectif** : On veut prédire  $X_{n+h}$ , soit  $\hat{X}_{n+h}$  ( $h$  horizon)

**Méthode 1** : On suppose que  $X_t = a(t) + s(t) + u_t$  pour  $t \in \mathbb{Z}$  et :

- On a obtenu  $\hat{a}_n(t)$  et  $\hat{s}_n(t)$  à partir de  $(X_1, \dots, X_n)$
- Pour  $\hat{u}_t = X_t - \hat{a}_n(t) - \hat{s}_n(t)$  on a obtenu et validé un modèle

$$\hat{u}_t = M_{\hat{\theta}}(\hat{u}_{t-1}, \hat{u}_{t-2}, \dots, \hat{u}_1, 0, \dots) \xi_t + f_{\hat{\theta}}(\hat{u}_{t-1}, \hat{u}_{t-2}, \dots, \hat{u}_1, 0, \dots)$$

# Prédiction à partir d'un modèle

On suppose observée  $(X_1, \dots, X_n)$ .

**Objectif** : On veut prédire  $X_{n+h}$ , soit  $\hat{X}_{n+h}$  ( $h$  horizon)

**Méthode 1** : On suppose que  $X_t = a(t) + s(t) + u_t$  pour  $t \in \mathbb{Z}$  et :

- On a obtenu  $\hat{a}_n(t)$  et  $\hat{s}_n(t)$  à partir de  $(X_1, \dots, X_n)$
- Pour  $\hat{u}_t = X_t - \hat{a}_n(t) - \hat{s}_n(t)$  on a obtenu et validé un modèle

$$\hat{u}_t = M_{\hat{\theta}}(\hat{u}_{t-1}, \hat{u}_{t-2}, \dots, \hat{u}_1, 0, \dots) \xi_t + f_{\hat{\theta}}(\hat{u}_{t-1}, \hat{u}_{t-2}, \dots, \hat{u}_1, 0, \dots)$$

## Prédiction à partir d'un modèle

On suppose observée  $(X_1, \dots, X_n)$ .

**Objectif** : On veut prédire  $X_{n+h}$ , soit  $\hat{X}_{n+h}$  ( $h$  horizon)

**Méthode 1** : On suppose que  $X_t = a(t) + s(t) + u_t$  pour  $t \in \mathbb{Z}$  et :

- On a obtenu  $\hat{a}_n(t)$  et  $\hat{s}_n(t)$  à partir de  $(X_1, \dots, X_n)$
- Pour  $\hat{u}_t = X_t - \hat{a}_n(t) - \hat{s}_n(t)$  on a obtenu et validé un modèle

$$\hat{u}_t = M_{\hat{\theta}}(\hat{u}_{t-1}, \hat{u}_{t-2}, \dots, \hat{u}_1, 0, \dots) \xi_t + f_{\hat{\theta}}(\hat{u}_{t-1}, \hat{u}_{t-2}, \dots, \hat{u}_1, 0, \dots)$$

## Prédiction à partir d'un modèle (2)

**Procédure :** On veut écrire  $\hat{X}_{n+h} = g(X_1, \dots, X_n)$  et on cherche

$$\hat{g} = \text{Arg min}_{g \in \mathbf{L}^2} \mathbb{E}[(X_{n+h} - g(X_1, \dots, X_n))^2]$$

$$\Rightarrow \hat{X}_{n+h} = \mathbb{E}[X_{n+h} \mid (X_1, \dots, X_n)]$$

$$\Rightarrow \hat{X}_{n+h} \simeq \hat{a}_n(n+h) + \hat{s}_n(n+h) + \mathbb{E}[\hat{u}_{n+h} \mid (\hat{u}_1, \dots, \hat{u}_n)]$$

## Prédiction à partir d'un modèle (2)

**Procédure :** On veut écrire  $\hat{X}_{n+h} = g(X_1, \dots, X_n)$  et on cherche

$$\hat{g} = \text{Arg min}_{g \in \mathbf{L}^2} \mathbf{E}[(X_{n+h} - g(X_1, \dots, X_n))^2]$$

$$\Rightarrow \hat{X}_{n+h} = \mathbf{E}[X_{n+h} \mid (X_1, \dots, X_n)]$$

$$\Rightarrow \hat{X}_{n+h} \simeq \hat{a}_n(n+h) + \hat{s}_n(n+h) + \mathbf{E}[\hat{u}_{n+h} \mid (\hat{u}_1, \dots, \hat{u}_n)]$$

## Prédiction à partir d'un modèle (2)

**Procédure :** On veut écrire  $\hat{X}_{n+h} = g(X_1, \dots, X_n)$  et on cherche

$$\hat{g} = \text{Arg min}_{g \in \mathbf{L}^2} \mathbf{E}[(X_{n+h} - g(X_1, \dots, X_n))^2]$$

$$\Rightarrow \hat{X}_{n+h} = \mathbf{E}[X_{n+h} \mid (X_1, \dots, X_n)]$$

$$\Rightarrow \hat{X}_{n+h} \simeq \hat{a}_n(n+h) + \hat{s}_n(n+h) + \mathbf{E}[\hat{u}_{n+h} \mid (\hat{u}_1, \dots, \hat{u}_n)]$$

## Prédiction à partir d'un modèle (2)

**Procédure** : On veut écrire  $\hat{X}_{n+h} = g(X_1, \dots, X_n)$  et on cherche

$$\hat{g} = \text{Arg min}_{g \in \mathbf{L}^2} \mathbf{E}[(X_{n+h} - g(X_1, \dots, X_n))^2]$$

$$\Rightarrow \hat{X}_{n+h} = \mathbf{E}[X_{n+h} \mid (X_1, \dots, X_n)]$$

$$\Rightarrow \hat{X}_{n+h} \simeq \hat{a}_n(n+h) + \hat{s}_n(n+h) + \mathbf{E}[\hat{u}_{n+h} \mid (\hat{u}_1, \dots, \hat{u}_n)]$$

# Prédiction à partir d'un modèle (3)

**Exemples :** On veut calculer  $\mathbb{E}[\hat{u}_{n+h} \mid (\hat{u}_1, \dots, \hat{u}_n)]$

① Si  $(\hat{u}_t)$  AR(1) causal  $\hat{u}_t = \hat{\alpha} \hat{u}_{t-1} + \xi_t$  d'où :

▶  $\hat{u}_{n+1} = \mathbb{E}[\hat{\alpha} \hat{u}_n + \xi_{n+1} \mid (\hat{u}_1, \dots, \hat{u}_n)] = \hat{\alpha} \hat{u}_n$

▶  $\hat{u}_{n+2} = \mathbb{E}[\hat{\alpha}^2 \hat{u}_n + \hat{\alpha} \xi_{n+1} + \xi_{n+2} \mid (\hat{u}_1, \dots, \hat{u}_n)] = \hat{\alpha}^2 \hat{u}_n$

▶  $\hat{u}_{n+h} = \mathbb{E}[\hat{\alpha}^h \hat{u}_n + \sum_{i=1}^h \hat{\alpha}^{h-i} \xi_{n+i} \mid (\hat{u}_1, \dots, \hat{u}_n)] = \hat{\alpha}^h \hat{u}_n$

② Si  $(\hat{u}_t)$  MA(1) inversible  $\hat{u}_t = \xi_t + \hat{\alpha} \xi_{t-1}$ , soit

$$\hat{u}_t = \xi_t - \sum_{k=1}^{\infty} (-\hat{\alpha})^k \hat{u}_{t-k}$$

▶  $\hat{u}_{n+1} = \mathbb{E}[\xi_{n+1} - \sum_{k=0}^{n-1} (-\hat{\alpha})^{k+1} \hat{u}_{n-k} \mid (\hat{u}_1, \dots, \hat{u}_n)] = - \sum_{k=0}^{n-1} (-\hat{\alpha})^{k+1} \hat{u}_{n-k}$

▶  $\hat{u}_{n+h} = \mathbb{E}[\xi_{n+h} + \hat{\alpha} \xi_{n+h-1} \mid (\hat{u}_1, \dots, \hat{u}_n)] = 0$  pour  $h \geq 2$

③ Si  $(\hat{u}_t)$  GARCH( $p, q$ ) alors  $\hat{u}_t = \hat{\sigma}_t \xi_t$  d'où  $\hat{X}_{n+h} = 0$  si  $h \geq 1$

# Prédiction à partir d'un modèle (3)

**Exemples :** On veut calculer  $\mathbb{E}[\hat{u}_{n+h} \mid (\hat{u}_1, \dots, \hat{u}_n)]$

❶ Si  $(\hat{u}_t)$  AR(1) causal  $\hat{u}_t = \hat{\alpha} \hat{u}_{t-1} + \xi_t$  d'où :

▶  $\hat{u}_{n+1} = \mathbb{E}[\hat{\alpha} \hat{u}_n + \xi_{n+1} \mid (\hat{u}_1, \dots, \hat{u}_n)] = \hat{\alpha} \hat{u}_n$

▶  $\hat{u}_{n+2} = \mathbb{E}[\hat{\alpha}^2 \hat{u}_n + \hat{\alpha} \xi_{n+1} + \xi_{n+2} \mid (\hat{u}_1, \dots, \hat{u}_n)] = \hat{\alpha}^2 \hat{u}_n$

▶  $\hat{u}_{n+h} = \mathbb{E}[\hat{\alpha}^h \hat{u}_n + \sum_{i=1}^h \hat{\alpha}^{h-i} \xi_{n+i} \mid (\hat{u}_1, \dots, \hat{u}_n)] = \hat{\alpha}^h \hat{u}_n$

❷ Si  $(\hat{u}_t)$  MA(1) inversible  $\hat{u}_t = \xi_t + \hat{\alpha} \xi_{t-1}$ , soit

$$\hat{u}_t = \xi_t - \sum_{k=1}^{\infty} (-\hat{\alpha})^k \hat{u}_{t-k}$$

▶  $\hat{u}_{n+1} = \mathbb{E}[\xi_{n+1} - \sum_{k=0}^{n-1} (-\hat{\alpha})^{k+1} \hat{u}_{n-k} \mid (\hat{u}_1, \dots, \hat{u}_n)] = - \sum_{k=0}^{n-1} (-\hat{\alpha})^{k+1} \hat{u}_{n-k}$

▶  $\hat{u}_{n+h} = \mathbb{E}[\xi_{n+h} + \hat{\alpha} \xi_{n+h-1} \mid (\hat{u}_1, \dots, \hat{u}_n)] = 0$  pour  $h \geq 2$

❸ Si  $(\hat{u}_t)$  GARCH( $p, q$ ) alors  $\hat{u}_t = \hat{\sigma}_t \xi_t$  d'où  $\hat{X}_{n+h} = 0$  si  $h \geq 1$

# Prédiction à partir d'un modèle (3)

**Exemples :** On veut calculer  $\mathbb{E}[\hat{u}_{n+h} \mid (\hat{u}_1, \dots, \hat{u}_n)]$

❶ Si  $(\hat{u}_t)$  AR(1) causal  $\hat{u}_t = \hat{\alpha} \hat{u}_{t-1} + \xi_t$  d'où :

▶  $\hat{u}_{n+1} = \mathbb{E}[\hat{\alpha} \hat{u}_n + \xi_{n+1} \mid (\hat{u}_1, \dots, \hat{u}_n)] = \hat{\alpha} \hat{u}_n$

▶  $\hat{u}_{n+2} = \mathbb{E}[\hat{\alpha}^2 \hat{u}_n + \hat{\alpha} \xi_{n+1} + \xi_{n+2} \mid (\hat{u}_1, \dots, \hat{u}_n)] = \hat{\alpha}^2 \hat{u}_n$

▶  $\hat{u}_{n+h} = \mathbb{E}[\hat{\alpha}^h \hat{u}_n + \sum_{i=1}^h \hat{\alpha}^{h-i} \xi_{n+i} \mid (\hat{u}_1, \dots, \hat{u}_n)] = \hat{\alpha}^h \hat{u}_n$

❷ Si  $(\hat{u}_t)$  MA(1) inversible  $\hat{u}_t = \xi_t + \hat{\alpha} \xi_{t-1}$ , soit

$$\hat{u}_t = \xi_t - \sum_{k=1}^{\infty} (-\hat{\alpha})^k \hat{u}_{t-k}$$

▶  $\hat{u}_{n+1} = \mathbb{E}[\xi_{n+1} - \sum_{k=0}^{n-1} (-\hat{\alpha})^{k+1} \hat{u}_{n-k} \mid (\hat{u}_1, \dots, \hat{u}_n)] = - \sum_{k=0}^{n-1} (-\hat{\alpha})^{k+1} \hat{u}_{n-k}$

▶  $\hat{u}_{n+h} = \mathbb{E}[\xi_{n+h} + \hat{\alpha} \xi_{n+h-1} \mid (\hat{u}_1, \dots, \hat{u}_n)] = 0$  pour  $h \geq 2$

❸ Si  $(\hat{u}_t)$  GARCH( $p, q$ ) alors  $\hat{u}_t = \hat{\sigma}_t \xi_t$  d'où  $\hat{X}_{n+h} = 0$  si  $h \geq 1$

# Prédiction à partir d'un modèle (3)

**Exemples :** On veut calculer  $\mathbb{E}[\hat{u}_{n+h} \mid (\hat{u}_1, \dots, \hat{u}_n)]$

❶ Si  $(\hat{u}_t)$  AR(1) causal  $\hat{u}_t = \hat{\alpha} \hat{u}_{t-1} + \xi_t$  d'où :

▶  $\hat{u}_{n+1} = \mathbb{E}[\hat{\alpha} \hat{u}_n + \xi_{n+1} \mid (\hat{u}_1, \dots, \hat{u}_n)] = \hat{\alpha} \hat{u}_n$

▶  $\hat{u}_{n+2} = \mathbb{E}[\hat{\alpha}^2 \hat{u}_n + \hat{\alpha} \xi_{n+1} + \xi_{n+2} \mid (\hat{u}_1, \dots, \hat{u}_n)] = \hat{\alpha}^2 \hat{u}_n$

▶  $\hat{u}_{n+h} = \mathbb{E}[\hat{\alpha}^h \hat{u}_n + \sum_{i=1}^h \hat{\alpha}^{h-i} \xi_{n+i} \mid (\hat{u}_1, \dots, \hat{u}_n)] = \hat{\alpha}^h \hat{u}_n$

❷ Si  $(\hat{u}_t)$  MA(1) inversible  $\hat{u}_t = \xi_t + \hat{\alpha} \xi_{t-1}$ , soit

$$\hat{u}_t = \xi_t - \sum_{k=1}^{\infty} (-\hat{\alpha})^k \hat{u}_{t-k}$$

▶  $\hat{u}_{n+1} = \mathbb{E}[\xi_{n+1} - \sum_{k=0}^{n-1} (-\hat{\alpha})^{k+1} \hat{u}_{n-k} \mid (\hat{u}_1, \dots, \hat{u}_n)] = - \sum_{k=0}^{n-1} (-\hat{\alpha})^{k+1} \hat{u}_{n-k}$

▶  $\hat{u}_{n+h} = \mathbb{E}[\xi_{n+h} + \hat{\alpha} \xi_{n+h-1} \mid (\hat{u}_1, \dots, \hat{u}_n)] = 0$  pour  $h \geq 2$

❸ Si  $(\hat{u}_t)$  GARCH( $p, q$ ) alors  $\hat{u}_t = \hat{\sigma}_t \xi_t$  d'où  $\hat{X}_{n+h} = 0$  si  $h \geq 1$

# Prédiction à partir d'un modèle (3)

**Exemples :** On veut calculer  $\mathbb{E}[\hat{u}_{n+h} \mid (\hat{u}_1, \dots, \hat{u}_n)]$

❶ Si  $(\hat{u}_t)$  AR(1) causal  $\hat{u}_t = \hat{\alpha} \hat{u}_{t-1} + \xi_t$  d'où :

▶  $\hat{u}_{n+1} = \mathbb{E}[\hat{\alpha} \hat{u}_n + \xi_{n+1} \mid (\hat{u}_1, \dots, \hat{u}_n)] = \hat{\alpha} \hat{u}_n$

▶  $\hat{u}_{n+2} = \mathbb{E}[\hat{\alpha}^2 \hat{u}_n + \hat{\alpha} \xi_{n+1} + \xi_{n+2} \mid (\hat{u}_1, \dots, \hat{u}_n)] = \hat{\alpha}^2 \hat{u}_n$

▶  $\hat{u}_{n+h} = \mathbb{E}[\hat{\alpha}^h \hat{u}_n + \sum_{i=1}^h \hat{\alpha}^{h-i} \xi_{n+i} \mid (\hat{u}_1, \dots, \hat{u}_n)] = \hat{\alpha}^h \hat{u}_n$

❷ Si  $(\hat{u}_t)$  MA(1) inversible  $\hat{u}_t = \xi_t + \hat{\alpha} \xi_{t-1}$ , soit

$$\hat{u}_t = \xi_t - \sum_{k=1}^{\infty} (-\hat{\alpha})^k \hat{u}_{t-k}$$

▶  $\hat{u}_{n+1} = \mathbb{E}[\xi_{n+1} - \sum_{k=0}^{n-1} (-\hat{\alpha})^{k+1} \hat{u}_{n-k} \mid (\hat{u}_1, \dots, \hat{u}_n)] = - \sum_{k=0}^{n-1} (-\hat{\alpha})^{k+1} \hat{u}_{n-k}$

▶  $\hat{u}_{n+h} = \mathbb{E}[\xi_{n+h} + \hat{\alpha} \xi_{n+h-1} \mid (\hat{u}_1, \dots, \hat{u}_n)] = 0$  pour  $h \geq 2$

❸ Si  $(\hat{u}_t)$  GARCH( $p, q$ ) alors  $\hat{u}_t = \hat{\sigma}_t \xi_t$  d'où  $\hat{X}_{n+h} = 0$  si  $h \geq 1$

## Prédiction à partir d'un modèle (3)

**Exemples :** On veut calculer  $\mathbb{E}[\hat{u}_{n+h} \mid (\hat{u}_1, \dots, \hat{u}_n)]$

❶ Si  $(\hat{u}_t)$  AR(1) causal  $\hat{u}_t = \hat{\alpha} \hat{u}_{t-1} + \xi_t$  d'où :

▶  $\hat{u}_{n+1} = \mathbb{E}[\hat{\alpha} \hat{u}_n + \xi_{n+1} \mid (\hat{u}_1, \dots, \hat{u}_n)] = \hat{\alpha} \hat{u}_n$

▶  $\hat{u}_{n+2} = \mathbb{E}[\hat{\alpha}^2 \hat{u}_n + \hat{\alpha} \xi_{n+1} + \xi_{n+2} \mid (\hat{u}_1, \dots, \hat{u}_n)] = \hat{\alpha}^2 \hat{u}_n$

▶  $\hat{u}_{n+h} = \mathbb{E}[\hat{\alpha}^h \hat{u}_n + \sum_{i=1}^h \hat{\alpha}^{h-i} \xi_{n+i} \mid (\hat{u}_1, \dots, \hat{u}_n)] = \hat{\alpha}^h \hat{u}_n$

❷ Si  $(\hat{u}_t)$  MA(1) inversible  $\hat{u}_t = \xi_t + \hat{\alpha} \xi_{t-1}$ , soit

$$\hat{u}_t = \xi_t - \sum_{k=1}^{\infty} (-\hat{\alpha})^k \hat{u}_{t-k}$$

▶  $\hat{u}_{n+1} = \mathbb{E}[\xi_{n+1} - \sum_{k=0}^{n-1} (-\hat{\alpha})^{k+1} \hat{u}_{n-k} \mid (\hat{u}_1, \dots, \hat{u}_n)] = - \sum_{k=0}^{n-1} (-\hat{\alpha})^{k+1} \hat{u}_{n-k}$

▶  $\hat{u}_{n+h} = \mathbb{E}[\xi_{n+h} + \hat{\alpha} \xi_{n+h-1} \mid (\hat{u}_1, \dots, \hat{u}_n)] = 0$  pour  $h \geq 2$

❸ Si  $(\hat{u}_t)$  GARCH( $p, q$ ) alors  $\hat{u}_t = \hat{\sigma}_t \xi_t$  d'où  $\hat{X}_{n+h} = 0$  si  $h \geq 1$

## Prédiction à partir d'un modèle (3)

**Exemples :** On veut calculer  $\mathbb{E}[\hat{u}_{n+h} \mid (\hat{u}_1, \dots, \hat{u}_n)]$

❶ Si  $(\hat{u}_t)$  AR(1) causal  $\hat{u}_t = \hat{\alpha} \hat{u}_{t-1} + \xi_t$  d'où :

▶  $\hat{u}_{n+1} = \mathbb{E}[\hat{\alpha} \hat{u}_n + \xi_{n+1} \mid (\hat{u}_1, \dots, \hat{u}_n)] = \hat{\alpha} \hat{u}_n$

▶  $\hat{u}_{n+2} = \mathbb{E}[\hat{\alpha}^2 \hat{u}_n + \hat{\alpha} \xi_{n+1} + \xi_{n+2} \mid (\hat{u}_1, \dots, \hat{u}_n)] = \hat{\alpha}^2 \hat{u}_n$

▶  $\hat{u}_{n+h} = \mathbb{E}[\hat{\alpha}^h \hat{u}_n + \sum_{i=1}^h \hat{\alpha}^{h-i} \xi_{n+i} \mid (\hat{u}_1, \dots, \hat{u}_n)] = \hat{\alpha}^h \hat{u}_n$

❷ Si  $(\hat{u}_t)$  MA(1) inversible  $\hat{u}_t = \xi_t + \hat{\alpha} \xi_{t-1}$ , soit

$$\hat{u}_t = \xi_t - \sum_{k=1}^{\infty} (-\hat{\alpha})^k \hat{u}_{t-k}$$

▶  $\hat{u}_{n+1} = \mathbb{E}[\xi_{n+1} - \sum_{k=0}^{n-1} (-\hat{\alpha})^{k+1} \hat{u}_{n-k} \mid (\hat{u}_1, \dots, \hat{u}_n)] = - \sum_{k=0}^{n-1} (-\hat{\alpha})^{k+1} \hat{u}_{n-k}$

▶  $\hat{u}_{n+h} = \mathbb{E}[\xi_{n+h} + \hat{\alpha} \xi_{n+h-1} \mid (\hat{u}_1, \dots, \hat{u}_n)] = 0$  pour  $h \geq 2$

❸ Si  $(\hat{u}_t)$  GARCH( $p, q$ ) alors  $\hat{u}_t = \hat{\sigma}_t \xi_t$  d'où  $\hat{X}_{n+h} = 0$  si  $h \geq 1$

# Prédiction à partir d'un modèle (3)

**Exemples :** On veut calculer  $\mathbb{E}[\hat{u}_{n+h} \mid (\hat{u}_1, \dots, \hat{u}_n)]$

❶ Si  $(\hat{u}_t)$  AR(1) causal  $\hat{u}_t = \hat{\alpha} \hat{u}_{t-1} + \xi_t$  d'où :

▶  $\hat{u}_{n+1} = \mathbb{E}[\hat{\alpha} \hat{u}_n + \xi_{n+1} \mid (\hat{u}_1, \dots, \hat{u}_n)] = \hat{\alpha} \hat{u}_n$

▶  $\hat{u}_{n+2} = \mathbb{E}[\hat{\alpha}^2 \hat{u}_n + \hat{\alpha} \xi_{n+1} + \xi_{n+2} \mid (\hat{u}_1, \dots, \hat{u}_n)] = \hat{\alpha}^2 \hat{u}_n$

▶  $\hat{u}_{n+h} = \mathbb{E}[\hat{\alpha}^h \hat{u}_n + \sum_{i=1}^h \hat{\alpha}^{h-i} \xi_{n+i} \mid (\hat{u}_1, \dots, \hat{u}_n)] = \hat{\alpha}^h \hat{u}_n$

❷ Si  $(\hat{u}_t)$  MA(1) inversible  $\hat{u}_t = \xi_t + \hat{\alpha} \xi_{t-1}$ , soit

$$\hat{u}_t = \xi_t - \sum_{k=1}^{\infty} (-\hat{\alpha})^k \hat{u}_{t-k}$$

▶  $\hat{u}_{n+1} = \mathbb{E}[\xi_{n+1} - \sum_{k=0}^{n-1} (-\hat{\alpha})^{k+1} \hat{u}_{n-k} \mid (\hat{u}_1, \dots, \hat{u}_n)] = - \sum_{k=0}^{n-1} (-\hat{\alpha})^{k+1} \hat{u}_{n-k}$

▶  $\hat{u}_{n+h} = \mathbb{E}[\xi_{n+h} + \hat{\alpha} \xi_{n+h-1} \mid (\hat{u}_1, \dots, \hat{u}_n)] = 0$  pour  $h \geq 2$

❸ Si  $(\hat{u}_t)$  GARCH( $p, q$ ) alors  $\hat{u}_t = \hat{\sigma}_t \xi_t$  d'où  $\hat{X}_{n+h} = 0$  si  $h \geq 1$

## Prédiction à partir d'un modèle (3)

**Exemples :** On veut calculer  $\mathbb{E}[\hat{u}_{n+h} \mid (\hat{u}_1, \dots, \hat{u}_n)]$

❶ Si  $(\hat{u}_t)$  AR(1) causal  $\hat{u}_t = \hat{\alpha} \hat{u}_{t-1} + \xi_t$  d'où :

▶  $\hat{u}_{n+1} = \mathbb{E}[\hat{\alpha} \hat{u}_n + \xi_{n+1} \mid (\hat{u}_1, \dots, \hat{u}_n)] = \hat{\alpha} \hat{u}_n$

▶  $\hat{u}_{n+2} = \mathbb{E}[\hat{\alpha}^2 \hat{u}_n + \hat{\alpha} \xi_{n+1} + \xi_{n+2} \mid (\hat{u}_1, \dots, \hat{u}_n)] = \hat{\alpha}^2 \hat{u}_n$

▶  $\hat{u}_{n+h} = \mathbb{E}[\hat{\alpha}^h \hat{u}_n + \sum_{i=1}^h \hat{\alpha}^{h-i} \xi_{n+i} \mid (\hat{u}_1, \dots, \hat{u}_n)] = \hat{\alpha}^h \hat{u}_n$

❷ Si  $(\hat{u}_t)$  MA(1) inversible  $\hat{u}_t = \xi_t + \hat{\alpha} \xi_{t-1}$ , soit

$$\hat{u}_t = \xi_t - \sum_{k=1}^{\infty} (-\hat{\alpha})^k \hat{u}_{t-k}$$

▶  $\hat{u}_{n+1} = \mathbb{E}[\xi_{n+1} - \sum_{k=0}^{n-1} (-\hat{\alpha})^{k+1} \hat{u}_{n-k} \mid (\hat{u}_1, \dots, \hat{u}_n)] = - \sum_{k=0}^{n-1} (-\hat{\alpha})^{k+1} \hat{u}_{n-k}$

▶  $\hat{u}_{n+h} = \mathbb{E}[\xi_{n+h} + \hat{\alpha} \xi_{n+h-1} \mid (\hat{u}_1, \dots, \hat{u}_n)] = 0$  pour  $h \geq 2$

❸ Si  $(\hat{u}_t)$  GARCH( $p, q$ ) alors  $\hat{u}_t = \hat{\sigma}_t \xi_t$  d'où  $\hat{X}_{n+h} = 0$  si  $h \geq 1$

# Prédiction sans modèle

**Méthode 2** : On détermine  $\hat{X}_{n+h}$  sans modéliser  $(X_1, \dots, X_n)$

⇒ Influence décroissante (exponentielle) aux données lointaines

⇒ Filtrage exponentiel, moindres carrés pondérés

① Filtrage simple de Holt-Winters simple : si  $\beta_h \in [0, 1]$ , on cherche :

$$\hat{X}_{n+h} = \hat{a}_h = \underset{a_h \in \mathbb{R}}{\operatorname{Argmin}} \sum_{j=1}^n \beta_h^{n-j} (X_j - a_h)^2$$

$$\Rightarrow \hat{a}_h = ({}^t J \Omega(\beta_h) J)^{-1} {}^t J \Omega(\beta_h) \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix}$$

$$\text{avec } \Omega(\beta_h) = \begin{pmatrix} \beta_h^{n-1} & 0 & 0 & \dots & 0 \\ 0 & \beta_h^{n-2} & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix} \quad \text{et } J = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$$

# Prédiction sans modèle

**Méthode 2** : On détermine  $\hat{X}_{n+h}$  sans modéliser  $(X_1, \dots, X_n)$

⇒ Influence décroissante (exponentielle) aux données lointaines

⇒ Filtrage exponentiel, moindres carrés pondérés

① Filtrage simple de Holt-Winters simple : si  $\beta_h \in [0, 1]$ , on cherche :

$$\hat{X}_{n+h} = \hat{a}_h = \underset{a_h \in \mathbb{R}}{\text{Argmin}} \sum_{j=1}^n \beta_h^{n-j} (X_j - a_h)^2$$

$$\Rightarrow \hat{a}_h = ({}^t J \Omega(\beta_h) J)^{-1} {}^t J \Omega(\beta_h) \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix}$$

$$\text{avec } \Omega(\beta_h) = \begin{pmatrix} \beta_h^{n-1} & 0 & 0 & \dots & 0 \\ 0 & \beta_h^{n-2} & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix} \quad \text{et } J = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$$

# Prédiction sans modèle

**Méthode 2** : On détermine  $\hat{X}_{n+h}$  sans modéliser  $(X_1, \dots, X_n)$

⇒ Influence décroissante (exponentielle) aux données lointaines

⇒ Filtrage exponentiel, moindres carrés pondérés

① Filtrage simple de Holt-Winters simple : si  $\beta_h \in [0, 1]$ , on cherche :

$$\hat{X}_{n+h} = \hat{a}_h = \underset{a_h \in \mathbb{R}}{\text{Argmin}} \sum_{j=1}^n \beta_h^{n-j} (X_j - a_h)^2$$

$$\Rightarrow \hat{a}_h = ({}^t J \Omega(\beta_h) J)^{-1} {}^t J \Omega(\beta_h) \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix}$$

$$\text{avec } \Omega(\beta_h) = \begin{pmatrix} \beta_h^{n-1} & 0 & 0 & \dots & 0 \\ 0 & \beta_h^{n-2} & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix} \quad \text{et } J = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$$

# Prédiction sans modèle

**Méthode 2** : On détermine  $\hat{X}_{n+h}$  sans modéliser  $(X_1, \dots, X_n)$

⇒ Influence décroissante (exponentielle) aux données lointaines

⇒ Filtrage exponentiel, moindres carrés pondérés

① Filtrage simple de Holt-Winters simple : si  $\beta_h \in [0, 1]$ , on cherche :

$$\hat{X}_{n+h} = \hat{a}_h = \underset{a_h \in \mathbb{R}}{\text{Argmin}} \sum_{j=1}^n \beta_h^{n-j} (X_j - a_h)^2$$

$$\Rightarrow \hat{a}_h = ({}^t J \Omega(\beta_h) J)^{-1} {}^t J \Omega(\beta_h) \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix}$$

$$\text{avec } \Omega(\beta_h) = \begin{pmatrix} \beta_h^{n-1} & 0 & 0 & \dots & 0 \\ 0 & \beta_h^{n-2} & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix} \quad \text{et } J = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$$

# Prédiction sans modèle

**Méthode 2** : On détermine  $\hat{X}_{n+h}$  sans modéliser  $(X_1, \dots, X_n)$

⇒ Influence décroissante (exponentielle) aux données lointaines

⇒ Filtrage exponentiel, moindres carrés pondérés

① Filtrage simple de Holt-Winters simple : si  $\beta_h \in [0, 1]$ , on cherche :

$$\hat{X}_{n+h} = \hat{a}_h = \underset{a_h \in \mathbb{R}}{\text{Argmin}} \sum_{j=1}^n \beta_h^{n-j} (X_j - a_h)^2$$

$$\Rightarrow \hat{a}_h = ({}^t J \Omega(\beta_h) J)^{-1} {}^t J \Omega(\beta_h) \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix}$$

$$\text{avec } \Omega(\beta_h) = \begin{pmatrix} \beta_h^{n-1} & 0 & 0 & \dots & 0 \\ 0 & \beta_h^{n-2} & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix} \quad \text{et } J = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$$

# Prédiction sans modèle

**Méthode 2** : On détermine  $\hat{X}_{n+h}$  sans modéliser  $(X_1, \dots, X_n)$

⇒ Influence décroissante (exponentielle) aux données lointaines

⇒ Filtrage exponentiel, moindres carrés pondérés

① Filtrage simple de Holt-Winters simple : si  $\beta_h \in [0, 1]$ , on cherche :

$$\hat{X}_{n+h} = \hat{a}_h = \underset{a_h \in \mathbb{R}}{\text{Argmin}} \sum_{j=1}^n \beta_h^{n-j} (X_j - a_h)^2$$

$$\Rightarrow \hat{a}_h = ({}^t J \Omega(\beta_h) J)^{-1} {}^t J \Omega(\beta_h) \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix}$$

$$\text{avec } \Omega(\beta_h) = \begin{pmatrix} \beta_h^{n-1} & 0 & 0 & \dots & 0 \\ 0 & \beta_h^{n-2} & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix} \quad \text{et } J = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$$

## Prédiction sans modèle (2)

- 2 Possibilité d'extension : filtrage de Holt-Winters double
- 3 Possibilité d'extension avec saisonnalité  $s$

## Prédiction sans modèle (2)

2 Possibilité d'extension : filtrage de Holt-Winters double

3 Possibilité d'extension avec saisonnalité  $s$

## Prédiction sans modèle (2)

- 2 Possibilité d'extension : filtrage de Holt-Winters double
- 3 Possibilité d'extension avec saisonnalité  $s$

## Prédiction sans modèle (2)

- 2 Possibilité d'extension : filtrage de Holt-Winters double
- 3 Possibilité d'extension avec saisonnalité  $s$

## Prédiction sans modèle (3)

**Problème** : Comment choisir  $\beta_h$  ?

Utilisation d'une **validation croisée** particulière :

- Pour  $m$  grand, on choisit une grille  $B_m = \{0, 1/m, 2/m, \dots, 1\}$  pour  $\beta_h$
- Pour  $M$  grand, et  $n - M \leq t \leq n - h$ , pour chaque  $\beta_h$  choisi,

$\implies \hat{X}_{t+h}(\beta_h)$  calculé à partir de  $(X_1, \dots, X_t)$

- On choisit  $\hat{\beta}_h = \operatorname{Argmin}_{\beta_h \in B_m} \sum_{t=n-M}^{n-h} (X_{t+h} - \hat{X}_{t+h}(\beta_h))^2$

**Remarque** :  $\beta_h \rightarrow 1$  même poids pour tous,  $\beta_h \rightarrow 0$  dernière valeur

## Prédiction sans modèle (3)

**Problème** : Comment choisir  $\beta_h$  ?

Utilisation d'une **validation croisée** particulière :

- Pour  $m$  grand, on choisit une grille  $B_m = \{0, 1/m, 2/m, \dots, 1\}$  pour  $\beta_h$
- Pour  $M$  grand, et  $n - M \leq t \leq n - h$ , pour chaque  $\beta_h$  choisi,

$\implies \hat{X}_{t+h}(\beta_h)$  calculé à partir de  $(X_1, \dots, X_t)$

- On choisit  $\hat{\beta}_h = \text{Argmin}_{\beta_h \in B_m} \sum_{t=n-M}^{n-h} (X_{t+h} - \hat{X}_{t+h}(\beta_h))^2$

**Remarque** :  $\beta_h \rightarrow 1$  même poids pour tous,  $\beta_h \rightarrow 0$  dernière valeur

## Prédiction sans modèle (3)

**Problème** : Comment choisir  $\beta_h$  ?

Utilisation d'une **validation croisée** particulière :

- Pour  $m$  grand, on choisit une grille  $B_m = \{0, 1/m, 2/m, \dots, 1\}$  pour  $\beta_h$

- Pour  $M$  grand, et  $n - M \leq t \leq n - h$ , pour chaque  $\beta_h$  choisi,

$$\implies \hat{X}_{t+h}(\beta_h) \text{ calculé à partir de } (X_1, \dots, X_t)$$

- On choisit  $\hat{\beta}_h = \text{Argmin}_{\beta_h \in B_m} \sum_{t=n-M}^{n-h} (X_{t+h} - \hat{X}_{t+h}(\beta_h))^2$

**Remarque** :  $\beta_h \rightarrow 1$  même poids pour tous,  $\beta_h \rightarrow 0$  dernière valeur

## Prédiction sans modèle (3)

**Problème** : Comment choisir  $\beta_h$  ?

Utilisation d'une **validation croisée** particulière :

- Pour  $m$  grand, on choisit une grille  $B_m = \{0, 1/m, 2/m, \dots, 1\}$  pour  $\beta_h$
- Pour  $M$  grand, et  $n - M \leq t \leq n - h$ , pour chaque  $\beta_h$  choisi,

$\implies \hat{X}_{t+h}(\beta_h)$  calculé à partir de  $(X_1, \dots, X_t)$

- On choisit  $\hat{\beta}_h = \text{Argmin}_{\beta_h \in B_m} \sum_{t=n-M}^{n-h} (X_{t+h} - \hat{X}_{t+h}(\beta_h))^2$

**Remarque** :  $\beta_h \rightarrow 1$  même poids pour tous,  $\beta_h \rightarrow 0$  dernière valeur

## Prédiction sans modèle (3)

**Problème** : Comment choisir  $\beta_h$  ?

Utilisation d'une **validation croisée** particulière :

- Pour  $m$  grand, on choisit une grille  $B_m = \{0, 1/m, 2/m, \dots, 1\}$  pour  $\beta_h$
- Pour  $M$  grand, et  $n - M \leq t \leq n - h$ , pour chaque  $\beta_h$  choisi,

$\implies \hat{X}_{t+h}(\beta_h)$  calculé à partir de  $(X_1, \dots, X_t)$

- On choisit  $\hat{\beta}_h = \operatorname{Argmin}_{\beta_h \in B_m} \sum_{t=n-M}^{n-h} (X_{t+h} - \hat{X}_{t+h}(\beta_h))^2$

**Remarque** :  $\beta_h \rightarrow 1$  même poids pour tous,  $\beta_h \rightarrow 0$  dernière valeur

## Prédiction sans modèle (3)

**Problème** : Comment choisir  $\beta_h$  ?

Utilisation d'une **validation croisée** particulière :

- Pour  $m$  grand, on choisit une grille  $B_m = \{0, 1/m, 2/m, \dots, 1\}$  pour  $\beta_h$
- Pour  $M$  grand, et  $n - M \leq t \leq n - h$ , pour chaque  $\beta_h$  choisi,

$\implies \hat{X}_{t+h}(\beta_h)$  calculé à partir de  $(X_1, \dots, X_t)$

- On choisit  $\hat{\beta}_h = \text{Argmin}_{\beta_h \in B_m} \sum_{t=n-M}^{n-h} (X_{t+h} - \hat{X}_{t+h}(\beta_h))^2$

**Remarque** :  $\beta_h \rightarrow 1$  même poids pour tous,  $\beta_h \rightarrow 0$  dernière valeur

## Prédiction sans modèle (3)

**Problème** : Comment choisir  $\beta_h$  ?

Utilisation d'une **validation croisée** particulière :

- Pour  $m$  grand, on choisit une grille  $B_m = \{0, 1/m, 2/m, \dots, 1\}$  pour  $\beta_h$
- Pour  $M$  grand, et  $n - M \leq t \leq n - h$ , pour chaque  $\beta_h$  choisi,

$\implies \hat{X}_{t+h}(\beta_h)$  calculé à partir de  $(X_1, \dots, X_t)$

- On choisit  $\hat{\beta}_h = \text{Argmin}_{\beta_h \in B_m} \sum_{t=n-M}^{n-h} (X_{t+h} - \hat{X}_{t+h}(\beta_h))^2$

**Remarque** :  $\beta_h \rightarrow 1$  même poids pour tous,  $\beta_h \rightarrow 0$  dernière valeur

# Prédiction à partir de méthodes d'apprentissage statistique

**Cadre** : On dispose de  $X = (X_1, \dots, X_n)$ ,  $T = (1, \dots, n)$

+ éventuellement des  $S^{(i)} = (\mathbb{1}_{t=i [T]} - \mathbb{1}_{t=T [T]})_t$  si **saisonnalité** période  $T$

+ éventuellement de variables exogènes  $Z^{(j)} = (Z_1^{(j)}, \dots, Z_n^{(j)})$ ,  $1 \leq j \leq p$   
(variables dont on connaît des prédictions à horizon  $h$ )

But : Prédire  $\hat{X}_{n+h}$  à partir des  $\left\{ \begin{array}{l} (X_1, \dots, X_n) \\ (1, \dots, n+h) \\ (S_1^{(i)}, \dots, S_{n+h}^{(i)}), \quad i = 1, \dots, T-1 \\ (Z_1^{(j)}, \dots, Z_{n+h}^{(j)}), \quad j = 1, \dots, p \end{array} \right.$

# Prédiction à partir de méthodes d'apprentissage statistique

**Cadre** : On dispose de  $X = (X_1, \dots, X_n)$ ,  $T = (1, \dots, n)$

+ éventuellement des  $S^{(i)} = (\mathbb{1}_{t=i} [T] - \mathbb{1}_{t=T} [T])_t$  si **saisonnalité** période  $T$

+ éventuellement de variables exogènes  $Z^{(j)} = (Z_1^{(j)}, \dots, Z_n^{(j)})$ ,  $1 \leq j \leq p$   
(variables dont on connaît des prédictions à horizon  $h$ )

But : Prédire  $\hat{X}_{n+h}$  à partir des  $\left\{ \begin{array}{l} (X_1, \dots, X_n) \\ (1, \dots, n+h) \\ (S_1^{(i)}, \dots, S_{n+h}^{(i)}), \quad i = 1, \dots, T-1 \\ (Z_1^{(j)}, \dots, Z_{n+h}^{(j)}), \quad j = 1, \dots, p \end{array} \right.$

# Prédiction à partir de méthodes d'apprentissage statistique

**Cadre** : On dispose de  $X = (X_1, \dots, X_n)$ ,  $T = (1, \dots, n)$

+ éventuellement des  $S^{(i)} = (\mathbb{1}_{t=i} [T] - \mathbb{1}_{t=T} [T])_t$  **si saisonnalité** période  $T$

+ éventuellement de variables exogènes  $Z^{(j)} = (Z_1^{(j)}, \dots, Z_n^{(j)})$ ,  $1 \leq j \leq p$   
(variables dont on connaît des prédictions à horizon  $h$ )

But : Prédire  $\hat{X}_{n+h}$  à partir des  $\left\{ \begin{array}{l} (X_1, \dots, X_n) \\ (1, \dots, n+h) \\ (S_1^{(i)}, \dots, S_{n+h}^{(i)}), \quad i = 1, \dots, T-1 \\ (Z_1^{(j)}, \dots, Z_{n+h}^{(j)}), \quad j = 1, \dots, p \end{array} \right.$

# Prédiction à partir de méthodes d'apprentissage statistique

**Cadre** : On dispose de  $X = (X_1, \dots, X_n)$ ,  $T = (1, \dots, n)$

+ éventuellement des  $S^{(i)} = (\mathbb{1}_{t=i} [T] - \mathbb{1}_{t=T} [T])_t$  si **saisonnalité** période  $T$

+ éventuellement de variables exogènes  $Z^{(j)} = (Z_1^{(j)}, \dots, Z_n^{(j)})$ ,  $1 \leq j \leq p$   
(variables dont on connaît des prédictions à horizon  $h$ )

But : Prédire  $\hat{X}_{n+h}$  à partir des

$$\begin{cases} (X_1, \dots, X_n) \\ (1, \dots, n+h) \\ (S_1^{(i)}, \dots, S_{n+h}^{(i)}), & i = 1, \dots, T-1 \\ (Z_1^{(j)}, \dots, Z_{n+h}^{(j)}), & j = 1, \dots, p \end{cases}$$

# Prédiction à partir de méthodes d'apprentissage statistique

**Cadre** : On dispose de  $X = (X_1, \dots, X_n)$ ,  $T = (1, \dots, n)$

+ éventuellement des  $S^{(i)} = (\mathbb{1}_{t=i} [T] - \mathbb{1}_{t=T} [T])_t$  si **saisonnalité** période  $T$

+ éventuellement de variables exogènes  $Z^{(j)} = (Z_1^{(j)}, \dots, Z_n^{(j)})$ ,  $1 \leq j \leq p$   
(variables dont on connaît des prédictions à horizon  $h$ )

**But** : Prédire  $\hat{X}_{n+h}$  à partir des

$$\left\{ \begin{array}{l} (X_1, \dots, X_n) \\ (1, \dots, n+h) \\ (S_1^{(i)}, \dots, S_{n+h}^{(i)}), \quad i = 1, \dots, T-1 \\ (Z_1^{(j)}, \dots, Z_{n+h}^{(j)}), \quad j = 1, \dots, p \end{array} \right.$$

# Prédiction à partir de méthodes d'apprentissage statistique

## Méthode "Hold out" ou "Out of sample" :

- Choisir l'ordre de la mémoire  $m$  (i.e.  $m = k T$  avec  $k$  grand);
- Découper en base d'apprentissage  $1 \rightarrow n_a$  et base test  $n_a + 1 \rightarrow n$
- Pour  $m \leq t \leq n_a - h$ , entraînement "machine statistique"  $G$  :

$$\left. \begin{array}{l} (X_{t-m}, \dots, X_t) \\ (t-m, \dots, t+h) \\ (S_t^{(i)}, \dots, S_{t+h}^{(i)}), \quad i = 1, \dots, T-1 \\ (Z_t^{(j)}, \dots, Z_{t+h}^{(j)}), \quad j = 1, \dots, p \end{array} \right\} \xrightarrow{G} X_{t+h}$$

- Comparaison des  $\hat{G}$  sur base test pour optimiser (méta)-paramètres

# Prédiction à partir de méthodes d'apprentissage statistique

## Méthode "Hold out" ou "Out of sample" :

- Choisir l'ordre de la mémoire  $m$  (i.e.  $m = k T$  avec  $k$  grand);
- Découper en base d'apprentissage  $1 \rightarrow n_a$  et base test  $n_a + 1 \rightarrow n$
- Pour  $m \leq t \leq n_a - h$ , entraînement "machine statistique"  $G$  :

$$\left. \begin{array}{l} (X_{t-m}, \dots, X_t) \\ (t-m, \dots, t+h) \\ (S_t^{(i)}, \dots, S_{t+h}^{(i)}), \quad i = 1, \dots, T-1 \\ (Z_t^{(j)}, \dots, Z_{t+h}^{(j)}), \quad j = 1, \dots, p \end{array} \right\} \xrightarrow{G} X_{t+h}$$

- Comparaison des  $\hat{G}$  sur base test pour optimiser (méta)-paramètres

# Prédiction à partir de méthodes d'apprentissage statistique

## Méthode "Hold out" ou "Out of sample" :

- Choisir l'ordre de la mémoire  $m$  (i.e.  $m = k T$  avec  $k$  grand);
- Découper en base d'apprentissage  $1 \rightarrow n_a$  et base test  $n_a + 1 \rightarrow n$
- Pour  $m \leq t \leq n_a - h$ , entraînement "machine statistique"  $G$  :

$$\left. \begin{array}{l} (X_{t-m}, \dots, X_t) \\ (t-m, \dots, t+h) \\ (S_t^{(i)}, \dots, S_{t+h}^{(i)}), \quad i = 1, \dots, T-1 \\ (Z_t^{(j)}, \dots, Z_{t+h}^{(j)}), \quad j = 1, \dots, p \end{array} \right\} \xrightarrow{G} X_{t+h}$$

- Comparaison des  $\hat{G}$  sur base test pour optimiser (méta)-paramètres

# Prédiction à partir de méthodes d'apprentissage statistique

## Méthode "Hold out" ou "Out of sample" :

- Choisir l'ordre de la mémoire  $m$  (i.e.  $m = k T$  avec  $k$  grand);
- Découper en base d'apprentissage  $1 \rightarrow n_a$  et base test  $n_a + 1 \rightarrow n$
- Pour  $m \leq t \leq n_a - h$ , entraînement "machine statistique"  $G$  :

$$\left. \begin{array}{l} (X_{t-m}, \dots, X_t) \\ (t-m, \dots, t+h) \\ (S_t^{(i)}, \dots, S_{t+h}^{(i)}), \quad i = 1, \dots, T-1 \\ (Z_t^{(j)}, \dots, Z_{t+h}^{(j)}), \quad j = 1, \dots, p \end{array} \right\} \xrightarrow{G} X_{t+h}$$

- Comparaison des  $\hat{G}$  sur base test pour optimiser (méta)-paramètres

# Prédiction à partir de méthodes d'apprentissage statistique

## Méthode "Hold out" ou "Out of sample" :

- Choisir l'ordre de la mémoire  $m$  (i.e.  $m = k T$  avec  $k$  grand);
- Découper en base d'apprentissage  $1 \rightarrow n_a$  et base test  $n_a + 1 \rightarrow n$
- Pour  $m \leq t \leq n_a - h$ , entraînement "machine statistique"  $G$  :

$$\left. \begin{array}{l} (X_{t-m}, \dots, X_t) \\ (t-m, \dots, t+h) \\ (S_t^{(i)}, \dots, S_{t+h}^{(i)}), \quad i = 1, \dots, T-1 \\ (Z_t^{(j)}, \dots, Z_{t+h}^{(j)}), \quad j = 1, \dots, p \end{array} \right\} \xrightarrow{G} X_{t+h}$$

- Comparaison des  $\hat{G}$  sur base test pour optimiser (méta)-paramètres

# Prédiction à partir de méthodes d'apprentissage statistique

## Méthode "Hold out" ou "Out of sample" :

- Choisir l'ordre de la mémoire  $m$  (i.e.  $m = k T$  avec  $k$  grand);
- Découper en base d'apprentissage  $1 \rightarrow n_a$  et base test  $n_a + 1 \rightarrow n$
- Pour  $m \leq t \leq n_a - h$ , entraînement "machine statistique"  $G$  :

$$\left. \begin{array}{l} (X_{t-m}, \dots, X_t) \\ (t-m, \dots, t+h) \\ (S_t^{(i)}, \dots, S_{t+h}^{(i)}), \quad i = 1, \dots, T-1 \\ (Z_t^{(j)}, \dots, Z_{t+h}^{(j)}), \quad j = 1, \dots, p \end{array} \right\} \xrightarrow{G} X_{t+h}$$

- Comparaison des  $\hat{G}$  sur base test pour optimiser (méta)-paramètres