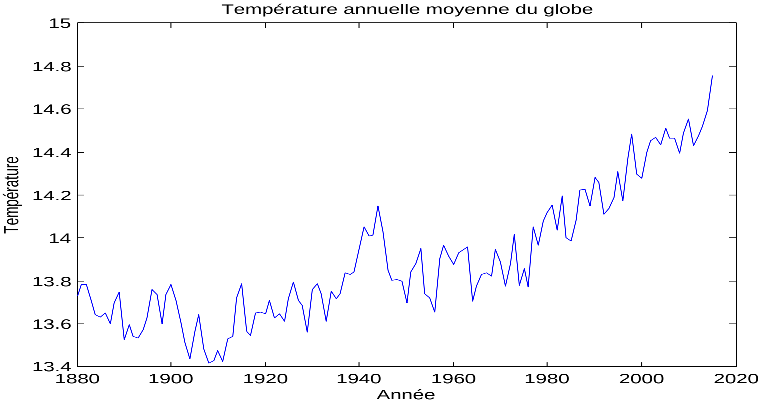
Cours d'Économétrie des séries chronologiques

Master 2 T.I.D.E.

UNIVERSITÉ PARIS 1
PANTHÉON SORBONNE

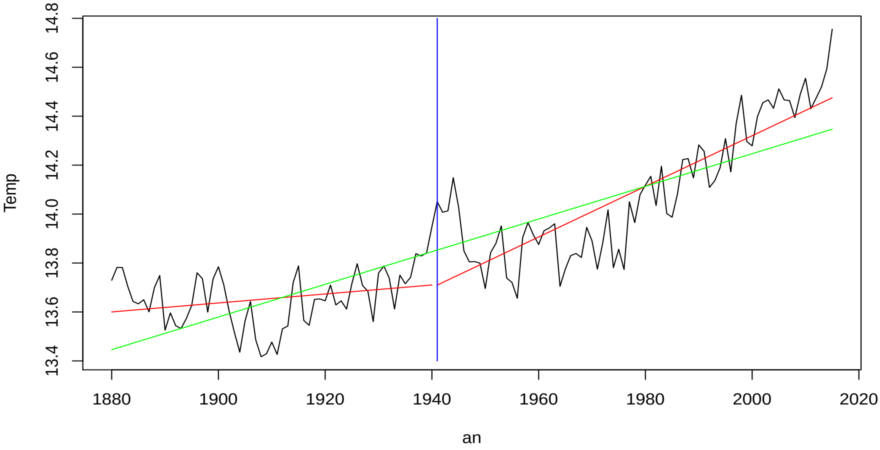
Année 2024-2025

Un exemple

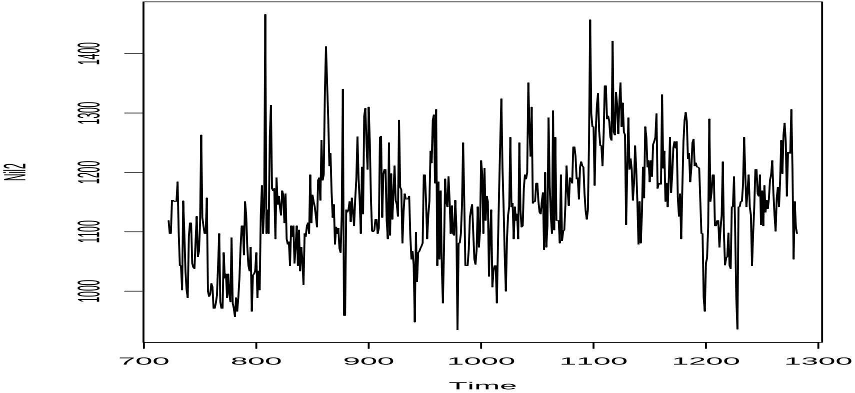


Température annuelle moyenne du globe
Température
15
14.8
14.6
14.4
14.2
14
13.8
13.6
13.4
1880
1900
1920
1940
1960
1980
2000
2020
Année

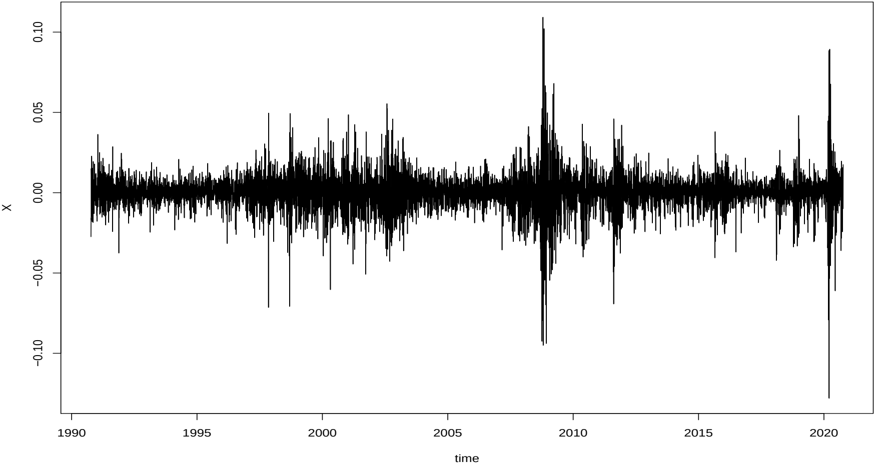
L'exemple travaillé



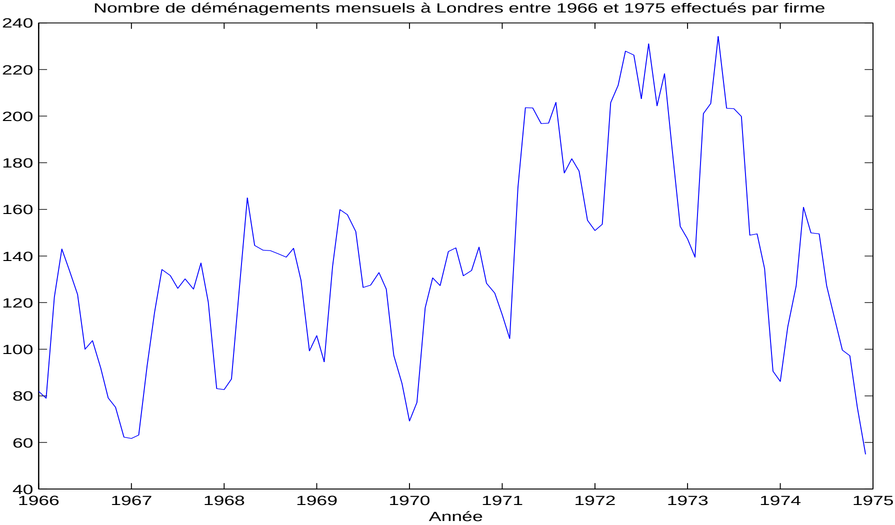
Un autre exemple : étiages du Nil



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**

$\Longrightarrow$ On essaie de décrire la série, de la **modéliser**

$\Longrightarrow$ On pourra en déduire des **estimations**, des **tests**

$\Longrightarrow$ 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**

$\Longrightarrow$ On essaie de décrire la série, de la **modéliser**

$\Longrightarrow$ On pourra en déduire des **estimations**, des **tests**

$\Longrightarrow$ 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**

$\Longrightarrow$ On essaie de décrire la série, de la **modéliser**

$\Longrightarrow$ On pourra en déduire des **estimations**, des **tests**

$\Longrightarrow$ 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**

$\Longrightarrow$ On essaie de décrire la série, de la **modéliser**

$\Longrightarrow$ On pourra en déduire des **estimations**, des **tests**

$\Longrightarrow$ 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**

$\Longrightarrow$ On essaie de décrire la série, de la **modéliser**

$\Longrightarrow$ On pourra en déduire des **estimations**, des **tests**

$\Longrightarrow$ 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**

$\Longrightarrow$ On essaie de décrire la série, de la **modéliser**

$\Longrightarrow$ On pourra en déduire des **estimations**, des **tests**

$\Longrightarrow$ 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**

$\Longrightarrow$ On essaie de décrire la série, de la **modéliser**

$\Longrightarrow$ On pourra en déduire des **estimations**, des **tests**

$\Longrightarrow$ 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

Séries chronologiques

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

$\Longrightarrow$ Modélisées par $(X_1, \ldots, X_n)$ où X_i variables aléatoires

(plus exactement par $(X_1(\omega), \ldots, X_n(\omega))$ 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

$\Longrightarrow$ Modélisées par $(X_1, \ldots, X_n)$ où X_i variables aléatoires

(plus exactement par $(X_1(\omega), \ldots, X_n(\omega))$ 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

$\Longrightarrow$ Modélisées par $(X_1, \ldots, X_n)$ où X_i variables aléatoires

(plus exactement par $(X_1(\omega), \ldots, X_n(\omega))$ 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

$\Longrightarrow$ Modélisées par $(X_1, \ldots, X_n)$ où X_i variables aléatoires

(plus exactement par $(X_1(\omega), \ldots, X_n(\omega))$ 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

$\Longrightarrow$ Modélisées par $(X_1, \ldots, X_n)$ où X_i variables aléatoires

(plus exactement par $(X_1(\omega), \ldots, X_n(\omega))$ 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

$\Longrightarrow$ Modélisées par $(X_1, \ldots, X_n)$ où X_i variables aléatoires

(plus exactement par $(X_1(\omega), \ldots, X_n(\omega))$ 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 ($X_t = Y_1 + \ldots + Y_t$ 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 ($X_t = Y_1 + \ldots + Y_t$ 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 ($X_t = Y_1 + \ldots + Y_t$ 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 ($X_t = Y_1 + \ldots + Y_t$ 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 ($X_t = Y_1 + \ldots + Y_t$ 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 ($X_t = Y_1 + \ldots + Y_t$ 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).

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 ($X_t = Y_1 + \ldots + Y_t$ 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 ($X_t = Y_1 + \ldots + Y_t$ 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 $(X_t)_{t\in\mathbb{Z}}$ série chronologique telle que $\mathbb{E}\left[X_t^2\right] < \infty$ pour tout $t \in \mathbb{Z}$

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

- **fonction espérance** : $m(t) = \mathbb{E}\left[X_t\right]$;
- **fonction variance** : $\sigma^2(t) = \mathbb{E}\left[X_t^2\right] - m^2(t) = \mathbb{E}\left[(X_t - m(t))^2\right]$;
- **fonction covariance** : $\gamma(s,t) = \mathbb{E}\left[X_s X_t\right] - m(s)m(t) = \mathbb{E}\left[(X_s - m(s))(X_t - m(t))\right]$;
- **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 $\mathbb{E}\left[X_t^2\right] < \infty$ pour tout $t \in \mathbb{Z}$
Pour $(s,t) \in \mathbb{Z}^2$, on appelle :

- **fonction espérance** : $m(t) = \mathbb{E}\left[X_t\right]$;
- **fonction variance** : $\sigma^2(t) = \mathbb{E}\left[X_t^2\right] - m^2(t) = \mathbb{E}\left[(X_t - m(t))^2\right]$;
- **fonction covariance** : $\gamma(s,t) = \mathbb{E}\left[X_s X_t\right] - m(s)m(t) = \mathbb{E}\left[(X_s - m(s))(X_t - m(t))\right]$;
- **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* $\mathbb{E}\left[X_t^2\right] < \infty$ *pour tout* $t \in \mathbb{Z}$
Pour $(s,t) \in \mathbb{Z}^2$, *on appelle :*

- **fonction espérance** : $m(t) = \mathbb{E}\left[X_t\right]$;
- **fonction variance** : $\sigma^2(t) = \mathbb{E}\left[X_t^2\right] - m^2(t) = \mathbb{E}\left[(X_t - m(t))^2\right]$;
- **fonction covariance** :
  $\gamma(s,t) = \mathbb{E}\left[X_s X_t\right] - m(s)m(t) = \mathbb{E}\left[(X_s - m(s))(X_t - m(t))\right]$;
- **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* $\mathbb{E}\left[X_t^2\right] < \infty$ *pour tout* $t \in \mathbb{Z}$
Pour $(s,t) \in \mathbb{Z}^2$*, on appelle :*

- **fonction espérance** : $m(t) = \mathbb{E}\left[X_t\right]$;
- **fonction variance** : $\sigma^2(t) = \mathbb{E}\left[X_t^2\right] - m^2(t) = \mathbb{E}\left[(X_t - m(t))^2\right]$;
- **fonction covariance** : $\gamma(s,t) = \mathbb{E}\left[X_s X_t\right] - m(s)m(t) = \mathbb{E}\left[(X_s - m(s))(X_t - m(t))\right]$;
- **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 $\mathbb{E}\left[X_t^2\right] < \infty$ pour tout $t \in \mathbb{Z}$
Pour $(s,t) \in \mathbb{Z}^2$, on appelle :

- **fonction espérance** : $m(t) = \mathbb{E}\left[X_t\right]$;
- **fonction variance** : $\sigma^2(t) = \mathbb{E}\left[X_t^2\right] - m^2(t) = \mathbb{E}\left[(X_t - m(t))^2\right]$;
- **fonction covariance** :
  $\gamma(s,t) = \mathbb{E}\left[X_s X_t\right] - m(s)m(t) = \mathbb{E}\left[(X_s - m(s))(X_t - m(t))\right]$;
- **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 $\mathbb{E}\left[X_t^2\right] < \infty$ pour tout $t \in \mathbb{Z}$
Pour $(s,t) \in \mathbb{Z}^2$, on appelle :

- **fonction espérance** : $m(t) = \mathbb{E}\left[X_t\right]$;
- **fonction variance** : $\sigma^2(t) = \mathbb{E}\left[X_t^2\right] - m^2(t) = \mathbb{E}\left[(X_t - m(t))^2\right]$;
- **fonction covariance** :
  $\gamma(s,t) = \mathbb{E}\left[X_s X_t\right] - m(s)m(t) = \mathbb{E}\left[(X_s - m(s))(X_t - m(t))\right]$;
- **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 $\mathbb{E}\left[X_t^2\right] < \infty$ pour tout $t \in \mathbb{Z}$
Pour $(s,t) \in \mathbb{Z}^2$, on appelle :

- **fonction espérance** : $m(t) = \mathbb{E}\left[X_t\right]$;
- **fonction variance** : $\sigma^2(t) = \mathbb{E}\left[X_t^2\right] - m^2(t) = \mathbb{E}\left[(X_t - m(t))^2\right]$;
- **fonction covariance** :
  $\gamma(s,t) = \mathbb{E}\left[X_s X_t\right] - m(s)m(t) = \mathbb{E}\left[(X_s - m(s))(X_t - m(t))\right]$;
- **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 $\mathbb{E}\left[X_t^2\right] < \infty$ pour tout $t \in \mathbb{Z}$
Pour $(s,t) \in \mathbb{Z}^2$, on appelle :

- **fonction espérance** : $m(t) = \mathbb{E}\left[X_t\right]$;
- **fonction variance** : $\sigma^2(t) = \mathbb{E}\left[X_t^2\right] - m^2(t) = \mathbb{E}\left[(X_t - m(t))^2\right]$;
- **fonction covariance** :
  $\gamma(s,t) = \mathbb{E}\left[X_s X_t\right] - m(s)m(t) = \mathbb{E}\left[(X_s - m(s))(X_t - m(t))\right]$;
- **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 $\mathbb{E}\left[X_t^2\right] < \infty$ pour tout $t \in \mathbb{Z}$
Pour $(s,t) \in \mathbb{Z}^2$, on appelle :

- **fonction espérance** : $m(t) = \mathbb{E}\left[X_t\right]$;
- **fonction variance** : $\sigma^2(t) = \mathbb{E}\left[X_t^2\right] - m^2(t) = \mathbb{E}\left[(X_t - m(t))^2\right]$;
- **fonction covariance** :
  $\gamma(s,t) = \mathbb{E}\left[X_s X_t\right] - m(s)m(t) = \mathbb{E}\left[(X_s - m(s))(X_t - m(t))\right]$;
- **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* $n \in \mathbb{N}^*$, $(t_1, \ldots, t_n) \in \mathbb{Z}^n$, $(\alpha_1, \ldots, \alpha_n) \in \mathbb{R}^n$.

Propriété

Pour une série chronologique $(X_t)_{t\in\mathbb{Z}}$ *gaussienne :*
$(\gamma(s,t) = 0$, *pour tout* $s \neq t) \iff ((X_t)_{t\in\mathbb{Z}}$ *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* $n \in \mathbb{N}^*$, $(t_1, \ldots, t_n) \in \mathbb{Z}^n$, $(\alpha_1, \ldots, \alpha_n) \in \mathbb{R}^n$.

Propriété

Pour une série chronologique $(X_t)_{t\in\mathbb{Z}}$ *gaussienne :*
$(\gamma(s,t) = 0$, *pour tout* $s \neq t) \iff ((X_t)_{t\in\mathbb{Z}}$ *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* $n \in \mathbb{N}^*$, $(t_1, \ldots, t_n) \in \mathbb{Z}^n$, $(\alpha_1, \ldots, \alpha_n) \in \mathbb{R}^n$.

Propriété

Pour une série chronologique $(X_t)_{t\in\mathbb{Z}}$ *gaussienne :*
$(\gamma(s,t) = 0$, *pour tout* $s \neq t) \iff ((X_t)_{t\in\mathbb{Z}}$ *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* $n \in \mathbb{N}^*$, $(t_1, \ldots, t_n) \in \mathbb{Z}^n$, $(\alpha_1, \ldots, \alpha_n) \in \mathbb{R}^n$.

Propriété

Pour une série chronologique $(X_t)_{t\in\mathbb{Z}}$ *gaussienne :*
$(\gamma(s,t) = 0,$ *pour tout* $s \neq t) \iff ((X_t)_{t\in\mathbb{Z}}$ *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

Stationnarité

Définition

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

$$(X_{t_1}, \ldots, X_{t_n}) \overset{\mathcal{L}}{\sim} (X_{t_1+c}, \cdots, X_{t_n+c}).$$

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

Propriété

Si (X_t) *stationnaire*

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

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* $\forall n \in \mathbb{N}^*$, $\forall (t_1, \ldots, t_n) \in \mathbb{Z}^n$, $\forall c \in \mathbb{Z}$,

$$(X_{t_1}, \ldots, X_{t_n}) \overset{\mathcal{L}}{\sim} (X_{t_1+c}, \cdots, X_{t_n+c}).$$

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

Propriété

Si (X_t) *stationnaire*

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

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* $\forall n \in \mathbb{N}^*$, $\forall (t_1,\dots,t_n)\in\mathbb{Z}^n$, $\forall c\in\mathbb{Z}$,

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

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

Propriété

Si (X_t) *stationnaire*

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

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* $\forall n \in \mathbb{N}^*$, $\forall (t_1, \ldots, t_n) \in \mathbb{Z}^n$, $\forall c \in \mathbb{Z}$,

$$(X_{t_1}, \ldots, X_{t_n}) \overset{\mathcal{L}}{\sim} (X_{t_1+c}, \cdots, X_{t_n+c}).$$

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

Propriété

Si (X_t) *stationnaire*

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

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* $\forall n \in \mathbb{N}^*$, $\forall (t_1,\ldots,t_n) \in \mathbb{Z}^n$, $\forall c \in \mathbb{Z}$,

$$(X_{t_1},\ldots,X_{t_n}) \overset{\mathcal{L}}{\sim} (X_{t_1+c},\cdots,X_{t_n+c}).$$

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

Propriété

Si (X_t) *stationnaire*

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

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* $\forall n \in \mathbb{N}^*$, $\forall (t_1, \ldots, t_n) \in \mathbb{Z}^n$, $\forall c \in \mathbb{Z}$,

$$(X_{t_1}, \ldots, X_{t_n}) \overset{\mathcal{L}}{\sim} (X_{t_1+c}, \cdots, X_{t_n+c}).$$

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

Propriété

Si (X_t) *stationnaire*

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

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 $\gamma(s,t) = cov(X_s, X_t)$ 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 $\Longrightarrow$ Stationnarité d'ordre 2), mais réciproque fausse.*
- *Si (X_t) série gaussienne, (Stationnarité $\Longleftrightarrow$ 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 $\gamma(s,t) = cov(X_s, X_t)$ 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 $\Longrightarrow$ Stationnarité d'ordre 2), mais réciproque fausse.*
- *Si (X_t) série gaussienne, (Stationnarité $\Longleftrightarrow$ 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 $\gamma(s,t) = cov(X_s, X_t)$ 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 $\Longrightarrow$ Stationnarité d'ordre 2), mais réciproque fausse.*
- *Si (X_t) série gaussienne, (Stationnarité $\Longleftrightarrow$ 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 $\gamma(s,t) = cov(X_s, X_t)$ 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 $\Longrightarrow$ Stationnarité d'ordre 2), mais réciproque fausse.*
- *Si (X_t) série gaussienne, (Stationnarité $\Longleftrightarrow$ 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 $\gamma(s,t) = cov(X_s, X_t)$ 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 $\Longrightarrow$ Stationnarité d'ordre 2), mais réciproque fausse.*
- *Si (X_t) série gaussienne, (Stationnarité $\Longleftrightarrow$ Stationnarité d'ordre 2).*

Plan du cours

Autocorrélation

Définition

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

- $r(k) = cov(X_0, X_k)$ *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 $(X_1, \ldots, X_n)$ trajectoire observée, comment estimer $r(k)$ et $\rho(k)$?

Autocorrélation

Définition

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

- $r(k) = cov(X_0, X_k)$ *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 $(X_1, \ldots, X_n)$ trajectoire observée, comment estimer $r(k)$ et $\rho(k)$?

Autocorrélation

Définition

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

- $r(k) = cov(X_0, X_k)$ *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 $(X_1, \ldots, X_n)$ trajectoire observée, comment estimer $r(k)$ et $\rho(k)$?

Autocorrélation

Définition

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

- $r(k) = cov(X_0, X_k)$ *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 $(X_1, \ldots, X_n)$ trajectoire observée, comment estimer $r(k)$ et $\rho(k)$?

Autocorrélation

Définition

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

- $r(k) = cov(X_0, X_k)$ *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 $(X_1, \ldots, X_n)$ trajectoire observée, comment estimer $r(k)$ et $\rho(k)$?

Moments empiriques

Définition

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

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

  Autre estimateur (biaisé) : $\widehat{\sigma_n^2} = \frac{1}{n}\sum_{i=1}^{n} X_i^2 - \left(\overline{X}_n\right)^2$;
- *L'*autocovariance empirique $\widehat{r}(k) = \frac{1}{n}\sum_{i=1}^{n-|k|}\left(X_i - \overline{X}_n\right)\left(X_{i+k} - \overline{X}_n\right)$;
- *L'*autocorrélation empirique (correlogram) $\widehat{\rho}(k) = \frac{\widehat{r}(k)}{\widehat{r}(0)} = \frac{\widehat{r}(k)}{\widehat{\sigma_n^2}}$;

Moments empiriques

Définition

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

- *La* **moyenne empirique** $\overline{X}_n = \dfrac{1}{n} \displaystyle\sum_{i=1}^{n} X_i$;
- *La* **variance empirique** $\overline{\sigma_n^2} = \dfrac{1}{n-1} \displaystyle\sum_{i=1}^{n} \left(X_i - \overline{X}_n\right)^2$;

  Autre estimateur (biaisé) : $\widehat{\sigma_n^2} = \dfrac{1}{n} \displaystyle\sum_{i=1}^{n} X_i^2 - \left(\overline{X}_n\right)^2$;
- *L'***autocovariance empirique** $\widehat{r}(k) = \dfrac{1}{n} \displaystyle\sum_{i=1}^{n-|k|} \left(X_i - \overline{X}_n\right)\left(X_{i+k} - \overline{X}_n\right)$;
- *L'***autocorrélation empirique (correlogram)** $\widehat{\rho}(k) = \dfrac{\widehat{r}(k)}{\widehat{r}(0)} = \dfrac{\widehat{r}(k)}{\widehat{\sigma_n^2}}$;

Moments empiriques

Définition

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

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

  Autre estimateur (biaisé) : $\widehat{\sigma_n^2} = \frac{1}{n}\sum_{i=1}^{n} X_i^2 - \left(\overline{X}_n\right)^2$;
- *L'* **autocovariance empirique** $\widehat{r}(k) = \frac{1}{n}\sum_{i=1}^{n-|k|}\left(X_i - \overline{X}_n\right)\left(X_{i+k} - \overline{X}_n\right)$;
- *L'* **autocorrélation empirique (correlogram)** $\widehat{\rho}(k) = \frac{\widehat{r}(k)}{\widehat{r}(0)} = \frac{\widehat{r}(k)}{\widehat{\sigma_n^2}}$;

Moments empiriques

Définition

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

- *La* **moyenne empirique** $\overline{X}_n = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} X_i$;
- *La* **variance empirique** $\overline{\sigma_n^2} = \dfrac{1}{n-1}\displaystyle\sum_{i=1}^{n}\left(X_i - \overline{X}_n\right)^2$;

  Autre estimateur (biaisé) : $\widehat{\sigma_n^2} = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} X_i^2 - \left(\overline{X}_n\right)^2$;
- *L'***autocovariance empirique** $\widehat{r}(k) = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n-|k|}\left(X_i - \overline{X}_n\right)\left(X_{i+k} - \overline{X}_n\right)$;
- *L'***autocorrélation empirique (correlogram)** $\widehat{\rho}(k) = \dfrac{\widehat{r}(k)}{\widehat{r}(0)} = \dfrac{\widehat{r}(k)}{\widehat{\sigma_n^2}}$;

Moments empiriques

Définition

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

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

  Autre estimateur (biaisé) : $\widehat{\sigma_n^2} = \frac{1}{n}\sum_{i=1}^{n} X_i^2 - (\overline{X}_n)^2$;
- *L'***autocovariance empirique** $\widehat{r}(k) = \frac{1}{n}\sum_{i=1}^{n-|k|} (X_i - \overline{X}_n)(X_{i+k} - \overline{X}_n)$;
- *L'***autocorrélation empirique (correlogram)** $\widehat{\rho}(k) = \frac{\widehat{r}(k)}{\widehat{r}(0)} = \frac{\widehat{r}(k)}{\widehat{\sigma_n^2}}$;

Moments empiriques

Définition

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

- *La* **moyenne empirique** $\overline{X}_n = \dfrac{1}{n}\sum_{i=1}^{n} X_i$;
- *La* **variance empirique** $\overline{\sigma_n^2} = \dfrac{1}{n-1}\sum_{i=1}^{n} \left(X_i - \overline{X}_n\right)^2$;

  Autre estimateur (biaisé) : $\widehat{\sigma_n^2} = \dfrac{1}{n}\sum_{i=1}^{n} X_i^2 - \left(\overline{X}_n\right)^2$;
- *L'***autocovariance empirique** $\widehat{r}(k) = \dfrac{1}{n}\sum_{i=1}^{n-|k|} \left(X_i - \overline{X}_n\right)\left(X_{i+k} - \overline{X}_n\right)$;
- *L'***autocorrélation empirique (correlogram)** $\widehat{\rho}(k) = \dfrac{\widehat{r}(k)}{\widehat{r}(0)} = \dfrac{\widehat{r}(k)}{\widehat{\sigma_n^2}}$;

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. $\overline{X}_n \xrightarrow[n\to+\infty]{\mathcal{P}} m = \mathbb{E}[X_0]$ *et* $\overline{\sigma_n^2}$ *ou* $\widehat{\sigma_n^2} \xrightarrow[n\to+\infty]{\mathcal{P}} \sigma^2 = var(X_0)$;
2. $\widehat{r}(0) = \widehat{\sigma_n^2} \xrightarrow[n\to+\infty]{\mathcal{P}} \sigma^2$ *et* $\widehat{r}(k) \xrightarrow[n\to+\infty]{\mathcal{P}} 0$ *pour tout* $k \neq 0$;
3. $\widehat{\rho}(0) = 1$ *et* $\widehat{\rho}(k) \xrightarrow[n\to+\infty]{\mathcal{P}} 0$ *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. $\overline{X}_n \xrightarrow[n\to+\infty]{\mathcal{P}} m = \mathbb{E}[X_0]$ *et* $\overline{\sigma_n^2}$ *ou* $\widehat{\sigma_n^2} \xrightarrow[n\to+\infty]{\mathcal{P}} \sigma^2 = var(X_0)$;
2. $\widehat{r}(0) = \widehat{\sigma_n^2} \xrightarrow[n\to+\infty]{\mathcal{P}} \sigma^2$ *et* $\widehat{r}(k) \xrightarrow[n\to+\infty]{\mathcal{P}} 0$ *pour tout* $k \neq 0$;
3. $\widehat{\rho}(0) = 1$ *et* $\widehat{\rho}(k) \xrightarrow[n\to+\infty]{\mathcal{P}} 0$ *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. $\overline{X}_n \xrightarrow[n\to+\infty]{\mathcal{P}} m = \mathbb{E}[X_0]$ *et* $\overline{\sigma_n^2}$ *ou* $\widehat{\sigma_n^2} \xrightarrow[n\to+\infty]{\mathcal{P}} \sigma^2 = var(X_0)$;
2. $\widehat{r}(0) = \widehat{\sigma_n^2} \xrightarrow[n\to+\infty]{\mathcal{P}} \sigma^2$ *et* $\widehat{r}(k) \xrightarrow[n\to+\infty]{\mathcal{P}} 0$ *pour tout* $k \neq 0$;
3. $\widehat{\rho}(0) = 1$ *et* $\widehat{\rho}(k) \xrightarrow[n\to+\infty]{\mathcal{P}} 0$ *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. $\overline{X}_n \xrightarrow[n\to+\infty]{\mathcal{P}} m = \mathbb{E}[X_0]$ *et* $\overline{\sigma_n^2}$ *ou* $\widehat{\sigma_n^2} \xrightarrow[n\to+\infty]{\mathcal{P}} \sigma^2 = var(X_0)$;
2. $\widehat{r}(0) = \widehat{\sigma_n^2} \xrightarrow[n\to+\infty]{\mathcal{P}} \sigma^2$ *et* $\widehat{r}(k) \xrightarrow[n\to+\infty]{\mathcal{P}} 0$ *pour tout* $k \neq 0$;
3. $\widehat{\rho}(0) = 1$ *et* $\widehat{\rho}(k) \xrightarrow[n\to+\infty]{\mathcal{P}} 0$ *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. $\overline{X}_n \xrightarrow[n\to+\infty]{\mathcal{P}} m = \mathbb{E}[X_0]$ *et* $\overline{\sigma_n^2}$ *ou* $\widehat{\sigma_n^2} \xrightarrow[n\to+\infty]{\mathcal{P}} \sigma^2 = var(X_0)$;
2. $\widehat{r}(0) = \widehat{\sigma_n^2} \xrightarrow[n\to+\infty]{\mathcal{P}} \sigma^2$ *et* $\widehat{r}(k) \xrightarrow[n\to+\infty]{\mathcal{P}} 0$ *pour tout* $k \neq 0$;
3. $\widehat{\rho}(0) = 1$ *et* $\widehat{\rho}(k) \xrightarrow[n\to+\infty]{\mathcal{P}} 0$ *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 $\mathbb{E}[X_0^2] < \infty$,

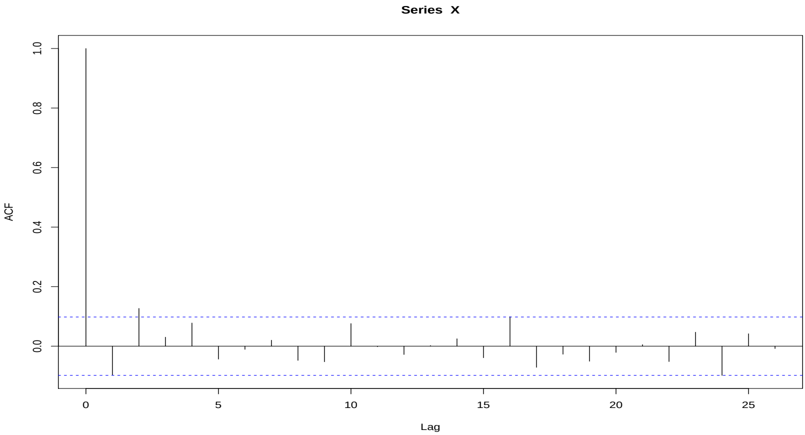
1. $\overline{X}_n \xrightarrow[n\to+\infty]{\mathcal{P}} m = \mathbb{E}[X_0]$ *et* $\overline{\sigma_n^2}$ *ou* $\widehat{\sigma_n^2} \xrightarrow[n\to+\infty]{\mathcal{P}} \sigma^2 = var(X_0)$;
2. $\widehat{r}(0) = \widehat{\sigma_n^2} \xrightarrow[n\to+\infty]{\mathcal{P}} \sigma^2$ *et* $\widehat{r}(k) \xrightarrow[n\to+\infty]{\mathcal{P}} 0$ *pour tout* $k \neq 0$;
3. $\widehat{\rho}(0) = 1$ *et* $\widehat{\rho}(k) \xrightarrow[n\to+\infty]{\mathcal{P}} 0$ *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) !

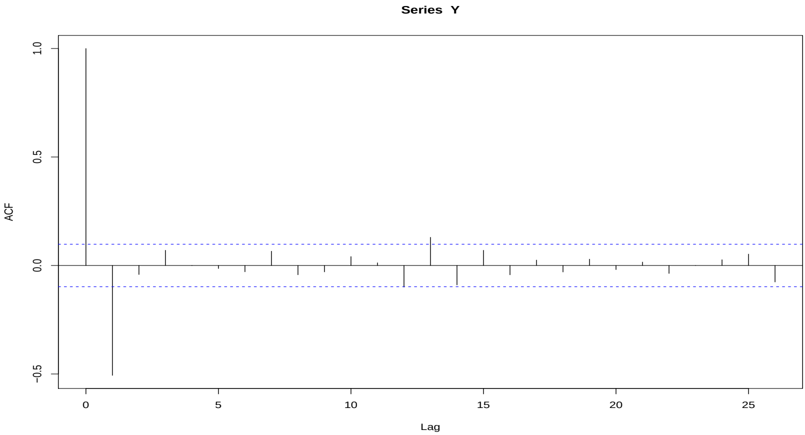
• Si (X_k) série non stationnaire, les convergences peuvent être vraies ... ou fausses, même si indépendance !

ACF



ACF (Auto-Correlogram Function) d'une suite (ε_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.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,

$$\sqrt{n}\,\left(\widehat{r}_X(k)\right)_{1\le k\le K_{\max}} \xrightarrow[n\to\infty]{\mathcal{L}} \mathcal{N}_{K_{\max}}\left(0\,,\,\sigma_X^4\,I_{K_{\max}}\right)$$

$$\Longrightarrow \qquad \sqrt{n}\,\left(\widehat{\rho}_X(k)\right)_{1\le k\le K_{\max}} \xrightarrow[n\to\infty]{\mathcal{L}} \mathcal{N}_{K_{\max}}\left(0\,,\,I_{K_{\max}}\right).$$

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.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,*

$$\sqrt{n}\left(\widehat{r}_X(k)\right)_{1\le k\le K_{\max}} \xrightarrow[n\to\infty]{\mathcal{L}} \mathcal{N}_{K_{\max}}\left(0\,,\,\sigma_X^4\, I_{K_{\max}}\right)$$

$$\Longrightarrow \qquad \sqrt{n}\left(\widehat{\rho}_X(k)\right)_{1\le k\le K_{\max}} \xrightarrow[n\to\infty]{\mathcal{L}} \mathcal{N}_{K_{\max}}\left(0\,,\, I_{K_{\max}}\right).$$

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.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,*

$$\sqrt{n}\left(\widehat{r}_X(k)\right)_{1\le k\le K_{\max}} \xrightarrow[n\to\infty]{\mathcal{L}} \mathcal{N}_{K_{\max}}\left(0\,,\,\sigma_X^4\, I_{K_{\max}}\right)$$

$$\Longrightarrow \qquad \sqrt{n}\left(\widehat{\rho}_X(k)\right)_{1\le k\le K_{\max}} \xrightarrow[n\to\infty]{\mathcal{L}} \mathcal{N}_{K_{\max}}\left(0\,,\, I_{K_{\max}}\right).$$

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.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,

$$\sqrt{n}\left(\widehat{r}_X(k)\right)_{1\leq k\leq K_{\max}} \xrightarrow[n\to\infty]{\mathcal{L}} \mathcal{N}_{K_{\max}}\left(0\,,\,\sigma_X^4\, I_{K_{\max}}\right)$$

$$\Longrightarrow \qquad \sqrt{n}\left(\widehat{\rho}_X(k)\right)_{1\leq k\leq K_{\max}} \xrightarrow[n\to\infty]{\mathcal{L}} \mathcal{N}_{K_{\max}}\left(0\,,\,I_{K_{\max}}\right).$$

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.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,

$$\sqrt{n}\left(\widehat{r}_X(k)\right)_{1\le k\le K_{\max}} \xrightarrow[n\to\infty]{\mathcal{L}} \mathcal{N}_{K_{\max}}\left(0\,,\,\sigma_X^4\, I_{K_{\max}}\right)$$

$$\Longrightarrow \qquad \sqrt{n}\left(\widehat{\rho}_X(k)\right)_{1\le k\le K_{\max}} \xrightarrow[n\to\infty]{\mathcal{L}} \mathcal{N}_{K_{\max}}\left(0\,,\,I_{K_{\max}}\right).$$

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.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,*

$$\sqrt{n}\left(\widehat{r}_X(k)\right)_{1\leq k\leq K_{\max}} \xrightarrow[n\to\infty]{\mathcal{L}} \mathcal{N}_{K_{\max}}\left(0\,,\,\sigma_X^4\, I_{K_{\max}}\right)$$

$$\Longrightarrow \qquad \sqrt{n}\left(\widehat{\rho}_X(k)\right)_{1\leq k\leq K_{\max}} \xrightarrow[n\to\infty]{\mathcal{L}} \mathcal{N}_{K_{\max}}\left(0\,,\,I_{K_{\max}}\right).$$

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.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,*

$$\sqrt{n}\left(\widehat{r}_X(k)\right)_{1\le k\le K_{\max}} \xrightarrow[n\to\infty]{\mathcal{L}} \mathcal{N}_{K_{\max}}\left(0\,,\,\sigma_X^4\, I_{K_{\max}}\right)$$

$$\Longrightarrow \qquad \sqrt{n}\left(\widehat{\rho}_X(k)\right)_{1\le k\le K_{\max}} \xrightarrow[n\to\infty]{\mathcal{L}} \mathcal{N}_{K_{\max}}\left(0\,,\,I_{K_{\max}}\right).$$

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

Dans les deux cas, on a

$$\begin{matrix} \widehat{T}_{BP} \\ \widehat{T}_{LB} \end{matrix} \quad \xrightarrow[n \to \infty]{\mathcal{L}} \chi^2_{K_{\max}}.$$

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

Dans les deux cas, on a

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

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

Dans les deux cas, on a

$$\begin{matrix} \widehat{T}_{BP} \\ \widehat{T}_{LB} \end{matrix} \quad \xrightarrow[n\to\infty]{\mathcal{L}} \chi^2_{K_{\max}}.$$

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

Dans les deux cas, on a

$$\begin{matrix} \widehat{T}_{BP} \\ \widehat{T}_{LB} \end{matrix} \quad \xrightarrow[n\to\infty]{\mathcal{L}} \chi^2_{K_{\max}}.$$

Test portemanteau

Théorème

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

$$\left\{\begin{array}{l} H_0 : (X_t) \textit{ forme une suite de variables indépendantes} \\ H_1 : (X_t) \textit{ ne forme pas une suite de variables indépendantes} \end{array}\right. .$$

On définit deux statistique de test :

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

Dans les deux cas, on a

$$\begin{array}{l} \widehat{T}_{BP} \\ \widehat{T}_{LB} \end{array} \xrightarrow[n\to\infty]{\mathcal{L}} \chi^2_{K_{\max}}.$$

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

Dans les deux cas, on a

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

Test portemanteau

Théorème

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

$$\left\{\begin{array}{l} H_0 : (X_t) \text{ forme une suite de variables indépendantes} \\ H_1 : (X_t) \text{ ne forme pas une suite de variables indépendantes} \end{array}\right. .$$

On définit deux statistique de test :

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

Dans les deux cas, on a

$$\begin{array}{c} \widehat{T}_{BP} \\ \widehat{T}_{LB} \end{array} \xrightarrow[n\to\infty]{\mathcal{L}} \chi^2_{K_{\max}}.$$

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 !

$\Longrightarrow$ 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 !

$\Longrightarrow$ 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 !

$\Longrightarrow$ 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 !

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

Test portemanteau

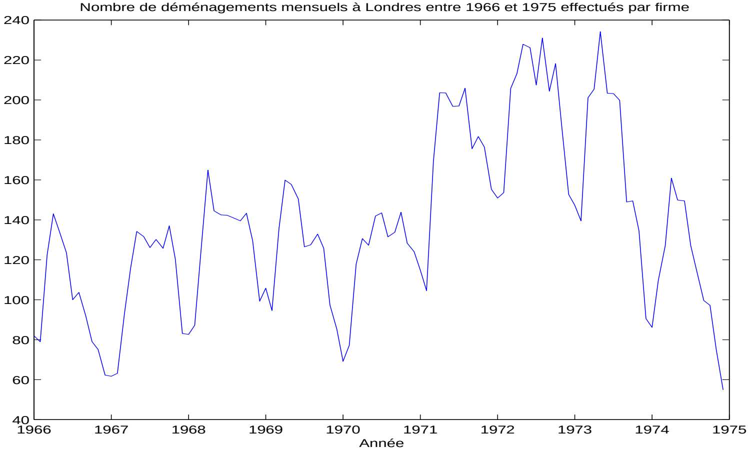
Choix de $K_{\max}$:

- Théoriquement il faut $K_{\max} = o\left(\sqrt{n}\right)$.
- 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 !

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

Séries non stationnaires



$\Longrightarrow$ On commence par estimer et retirer tendance ou saisonnalité

Plan du cours

Définition

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

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

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

Définition

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

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

$t \mapsto a(t)$ 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 $\displaystyle\sum_{i=1}^{r} s(i) = 0$, 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 : $a(t) = a_0 + a_1 t + \cdots + a_k t^k$, saisonnalité annuelle : $s(t) = \sum_{i=1}^{12} s_i \mathbb{1}_{t=i\ [12]}$ ou $s(t) = \alpha \sin(2\pi t/12) + \beta \cos(2\pi t/12)$,...

Définition

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

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

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

Définition

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

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

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

Définition

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

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

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

Définition

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

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

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

Définition

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

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

$t \mapsto a(t)$ *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* $\displaystyle\sum_{i=1}^{r} s(i) = 0$, *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 : $a(t) = a_0 + a_1 t + \cdots + a_k t^k$, *saisonnalité annuelle :* $s(t) = \sum_{i=1}^{12} s_i \mathbb{I}_{t=i\ [12]}$ *ou* $s(t) = \alpha \sin(2\pi t/12) + \beta \cos(2\pi t/12)$,...

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* $(X_t - \widehat{a}(t))_t$, *où* $\widehat{a}(t)$ *estimateur de* $a(t)$, **série détendancialisée**.
- $(X_t - s(t))_t$ *ou* $(X_t - \widehat{s}(t))_t$, *où* $\widehat{s}(t)$ *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* $(u_t)_{t \in Z}$ *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 :* $X_t = a(t) + s(t) + u_t = \big(a(t) - s(t)\big) + 2s(t) + u_t$.

Définition

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

- *$(X_t - a(t))_t$ ou $(X_t - \widehat{a}(t))_t$, où $\widehat{a}(t)$ estimateur de $a(t)$,* **série détendancialisée**.
- *$(X_t - s(t))_t$ ou $(X_t - \widehat{s}(t))_t$, où $\widehat{s}(t)$ 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 $(u_t)_{t \in Z}$ 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 :* $X_t = a(t) + s(t) + u_t = \big(a(t) - s(t)\big) + 2s(t) + u_t$.

Définition

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

- *$(X_t - a(t))_t$ ou $(X_t - \widehat{a}(t))_t$, où $\widehat{a}(t)$ estimateur de $a(t)$,* **série détendancialisée**.
- *$(X_t - s(t))_t$ ou $(X_t - \widehat{s}(t))_t$, où $\widehat{s}(t)$ 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 $(u_t)_{t \in Z}$ 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 :* $X_t = a(t) + s(t) + u_t = \big(a(t) - s(t)\big) + 2s(t) + u_t$.

Définition

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

- $(X_t - a(t))_t$ *ou* $(X_t - \widehat{a}(t))_t$, *où* $\widehat{a}(t)$ *estimateur de* $a(t)$, **série détendancialisée**.
- $(X_t - s(t))_t$ *ou* $(X_t - \widehat{s}(t))_t$, *où* $\widehat{s}(t)$ *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* $(u_t)_{t \in Z}$ *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 : $X_t = a(t) + s(t) + u_t = \big(a(t) - s(t)\big) + 2s(t) + u_t$.*

Définition

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

- *$(X_t - a(t))_t$ ou $(X_t - \widehat{a}(t))_t$, où $\widehat{a}(t)$ estimateur de $a(t)$,* **série détendancialisée**.
- *$(X_t - s(t))_t$ ou $(X_t - \widehat{s}(t))_t$, où $\widehat{s}(t)$ 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 $(u_t)_{t \in Z}$ 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 :* $X_t = a(t) + s(t) + u_t = \big(a(t) - s(t)\big) + 2s(t) + u_t$.

Définition

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

- $(X_t - a(t))_t$ *ou* $(X_t - \widehat{a}(t))_t$, *où* $\widehat{a}(t)$ *estimateur de* $a(t)$, **série détendancialisée**.
- $(X_t - s(t))_t$ *ou* $(X_t - \widehat{s}(t))_t$, *où* $\widehat{s}(t)$ *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* $(u_t)_{t \in Z}$ *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 :* $X_t = a(t) + s(t) + u_t = \big(a(t) - s(t)\big) + 2s(t) + u_t$.

Définition

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

- $(X_t - a(t))_t$ *ou* $(X_t - \widehat{a}(t))_t$, *où* $\widehat{a}(t)$ *estimateur de* $a(t)$, **série détendancialisée**.
- $(X_t - s(t))_t$ *ou* $(X_t - \widehat{s}(t))_t$, *où* $\widehat{s}(t)$ *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* $(u_t)_{t \in Z}$ *bruit de variance* $\sigma(\cdot)$ *non constante.*

Estimation de la tendance et de la saisonnalité

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

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

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

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

Estimation de la tendance et de la saisonnalité

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

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

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

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

Estimation de la tendance et de la saisonnalité

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

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

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

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

Estimation de la tendance et de la saisonnalité

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

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

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

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

Plan du cours

Estimation semi-paramétrique de la saisonnalité seule

Définition

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

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

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

Propriété

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

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

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 :

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

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

Propriété

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

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

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 :

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

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

Propriété

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

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

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 :

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

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

Propriété

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

$$\widehat{s_i} = \frac{1}{n} \sum_{k=1}^{n} X_{i+(k-1)T} - \frac{1}{nT} \sum_{k=1}^{nT} X_k \quad pour\ i = 1, \ldots, T$$

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 :

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

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

Propriété

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

$$\widehat{s_i} = \frac{1}{n} \sum_{k=1}^{n} X_{i+(k-1)T} - \frac{1}{nT} \sum_{k=1}^{nT} X_k \quad \textit{pour } i = 1, \ldots, T$$

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 $a(t) = a_0 + a_1\, g_1(t) + \cdots + a_p\, g_p(t)$ pour $t \in \mathbb{Z}$.

Exemple

1. $a(t) = a_0 + a_1\, t$: *tendance linéaire*
2. $a(t) = a_0 + a_1\, t + \cdots + a_p\, t^p$: *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*

$\Longrightarrow$ Régression linéaire par moindres carrés !

Propriété

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

$$ {}^t\left(\widehat{a}_0, \ldots, \widehat{a}_p\right) = \left({}^tZ\, Z\right)^{-1}\, {}^tZ\, {}^t\left(X_1, \ldots, X_n\right). $$

$\Longrightarrow$ Estimation de la tendance : $\widehat{a}(t) = \widehat{a}_0 + \widehat{a}_1\, g_1(t) + \cdots + \widehat{a}_p\, g_p(t)$

Estimation semi-paramétrique de la tendance seule

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

Exemple

1. $a(t) = a_0 + a_1\, t$: *tendance linéaire*
2. $a(t) = a_0 + a_1\, t + \cdots + a_p\, t^p$: *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*

$\Longrightarrow$ Régression linéaire par moindres carrés !

Propriété

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

$$ {}^t(\widehat{a}_0, \ldots, \widehat{a}_p) = \left({}^tZ\, Z\right)^{-1}\, {}^tZ\, {}^t(X_1, \ldots, X_n). $$

$\Longrightarrow$ Estimation de la tendance : $\widehat{a}(t) = \widehat{a}_0 + \widehat{a}_1\, g_1(t) + \cdots + \widehat{a}_p\, g_p(t)$

Estimation semi-paramétrique de la tendance seule

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

Exemple

1. $a(t) = a_0 + a_1\, t$: *tendance linéaire*
2. $a(t) = a_0 + a_1\, t + \cdots + a_p\, t^p$: *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*

$\Longrightarrow$ Régression linéaire par moindres carrés !

Propriété

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

$$ {}^t(\widehat{a}_0, \ldots, \widehat{a}_p) = \left({}^tZ\, Z\right)^{-1}\, {}^tZ\, {}^t(X_1, \ldots, X_n). $$

$\Longrightarrow$ Estimation de la tendance : $\widehat{a}(t) = \widehat{a}_0 + \widehat{a}_1\, g_1(t) + \cdots + \widehat{a}_p\, g_p(t)$

Estimation semi-paramétrique de la tendance seule

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

Exemple

1. $a(t) = a_0 + a_1\, t$: *tendance linéaire*
2. $a(t) = a_0 + a_1\, t + \cdots + a_p\, t^p$: *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*

$\Longrightarrow$ Régression linéaire par moindres carrés !

Propriété

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

$${}^t(\widehat{a}_0, \ldots, \widehat{a}_p) = \left({}^t Z\, Z\right)^{-1}\, {}^t Z\, {}^t(X_1, \ldots, X_n).$$

$\Longrightarrow$ Estimation de la tendance : $\widehat{a}(t) = \widehat{a}_0 + \widehat{a}_1\, g_1(t) + \cdots + \widehat{a}_p\, g_p(t)$

Estimation semi-paramétrique de la tendance seule

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

Exemple

1. $a(t) = a_0 + a_1\, t$: *tendance linéaire*
2. $a(t) = a_0 + a_1\, t + \cdots + a_p\, t^p$: *tendance polynomiale*
3. $a(t) = \displaystyle\sum_{k=0}^{p} a_k \cos\left(2\pi \frac{kt}{n}\right) + \sum_{k=1}^{p} b_k \sin\left(2\pi \frac{kt}{n}\right)$: *tendance trigonométrique*

$\Longrightarrow$ Régression linéaire par moindres carrés !

Propriété

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

$${}^t(\widehat{a}_0, \ldots, \widehat{a}_p) = \left({}^tZ\,Z\right)^{-1}\, {}^tZ\, {}^t(X_1, \ldots, X_n).$$

$\Longrightarrow$ Estimation de la tendance : $\widehat{a}(t) = \widehat{a}_0 + \widehat{a}_1\, g_1(t) + \cdots + \widehat{a}_p\, g_p(t)$

Estimation semi-paramétrique de la tendance seule

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

Exemple

1. $a(t) = a_0 + a_1\, t$: *tendance linéaire*
2. $a(t) = a_0 + a_1\, t + \cdots + a_p\, t^p$: *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*

$\Longrightarrow$ Régression linéaire par moindres carrés !

Propriété

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

$$ {}^t(\widehat{a}_0, \ldots, \widehat{a}_p) = \left({}^tZ\, Z\right)^{-1}\, {}^tZ\, {}^t(X_1, \ldots, X_n). $$

$\Longrightarrow$ Estimation de la tendance : $\widehat{a}(t) = \widehat{a}_0 + \widehat{a}_1\, g_1(t) + \cdots + \widehat{a}_p\, g_p(t)$

Estimation semi-paramétrique de la tendance seule

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

Exemple

1. $a(t) = a_0 + a_1\, t$: *tendance linéaire*
2. $a(t) = a_0 + a_1\, t + \cdots + a_p\, t^p$: *tendance polynomiale*
3. $a(t) = \displaystyle\sum_{k=0}^{p} a_k \cos\left(2\pi \frac{kt}{n}\right) + \sum_{k=1}^{p} b_k \sin\left(2\pi \frac{kt}{n}\right)$: *tendance trigonométrique*

$\Longrightarrow$ Régression linéaire par moindres carrés !

Propriété

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

$$ {}^t(\widehat{a}_0, \ldots, \widehat{a}_p) = \left({}^tZ\, Z\right)^{-1}\, {}^tZ\; {}^t(X_1, \ldots, X_n).$$

$\Longrightarrow$ Estimation de la tendance : $\widehat{a}(t) = \widehat{a}_0 + \widehat{a}_1\, g_1(t) + \cdots + \widehat{a}_p\, g_p(t)$

Estimation semi-paramétrique de la tendance seule

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

Exemple

1. $a(t) = a_0 + a_1\; t$: *tendance linéaire*
2. $a(t) = a_0 + a_1\; t + \cdots + a_p\; t^p$: *tendance polynomiale*
3. $a(t) = \sum_{k=0}^{p} a_k \cos\big(2\pi \frac{kt}{n}\big) + \sum_{k=1}^{p} b_k \sin\big(2\pi \frac{kt}{n}\big)$: *tendance trigonométrique*

$\Longrightarrow$ Régression linéaire par moindres carrés !

Propriété

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

$${}^t\big(\widehat{a}_0, \ldots, \widehat{a}_p\big) = \big({}^tZ\, Z\big)^{-1}\, {}^tZ\, {}^t\big(X_1, \ldots, X_n\big).$$

$\Longrightarrow$ Estimation de la tendance : $\widehat{a}(t) = \widehat{a}_0 + \widehat{a}_1\, g_1(t) + \cdots + \widehat{a}_p\, g_p(t)$

Estimation conjointe de la tendance et de la saisonnalité

On suppose que

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

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

Propriété

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

on a : $${}^t\big(\widehat{a}_0, \ldots, \widehat{a}_p, \widehat{s}_1, \ldots, \widehat{s}_{T-1}\big) = \big({}^tZ\,Z\big)^{-1}\,{}^tZ\,{}^t(X_1, \ldots, X_n).$$

Conséquence :

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

Estimation conjointe de la tendance et de la saisonnalité

On suppose que

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

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

Propriété

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

on a : $\quad {}^t\big(\widehat{a}_0, \ldots, \widehat{a}_p, \widehat{s}_1, \ldots, \widehat{s}_{T-1}\big) = \big({}^tZ\, Z\big)^{-1}\, {}^tZ\, {}^t(X_1, \ldots, X_n).$

Conséquence :

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

Estimation conjointe de la tendance et de la saisonnalité

On suppose que

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

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

Propriété

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

on a : $\quad {}^t\big(\widehat{a}_0, \ldots, \widehat{a}_p, \widehat{s}_1, \ldots, \widehat{s}_{T-1}\big) = \big({}^tZ\,Z\big)^{-1}\,{}^tZ\,{}^t(X_1, \ldots, X_n).$

Conséquence :

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

Estimation conjointe de la tendance et de la saisonnalité

On suppose que

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

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

Propriété

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

on a : $\quad {}^t\big(\widehat{a}_0, \ldots, \widehat{a}_p, \widehat{s}_1, \ldots, \widehat{s}_{T-1}\big) = \big({}^tZ\, Z\big)^{-1}\, {}^tZ\, {}^t\big(X_1, \ldots, X_n\big).$

Conséquence :

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

Estimation conjointe de la tendance et de la saisonnalité

On suppose que

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

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

Propriété

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

on a : $\quad {}^t\big(\widehat{a}_0, \ldots, \widehat{a}_p, \widehat{s}_1, \ldots, \widehat{s}_{T-1}\big) = \big({}^tZ\, Z\big)^{-1}\, {}^tZ\, {}^t\big(X_1, \ldots, X_n\big).$

Conséquence :

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

Estimation conjointe de la tendance et de la saisonnalité

On suppose que

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

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

Propriété

Avec $Z = \begin{bmatrix} 1 & G_1 & \cdots & G_p & H_1 & \cdots & H_{T-1} \end{bmatrix}$ *où* $\left\{ \begin{array}{l} G_i = {}^t\big(g_i(1), \ldots, g_i(n)\big) \\ H_i = {}^t\big(h_i(1), \ldots, h_i(n)\big) \end{array} \right.$,

on a : $\quad {}^t\big(\widehat{a}_0, \ldots, \widehat{a}_p, \widehat{s}_1, \ldots, \widehat{s}_{T-1}\big) = \big({}^tZ\, Z\big)^{-1}\, {}^tZ\, {}^t\big(X_1, \ldots, X_n\big).$

Conséquence :

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

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

$\Longrightarrow$ On utilise le **critère BIC**

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

$$\Longrightarrow \widehat{p} = \underset{0 \le p \le p_{\max}}{\arg\min}\ BIC(p)$$

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

$\Longrightarrow$ On utilise le **critère BIC**

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

$$\Longrightarrow \widehat{p} = \underset{0 \leq p \leq p_{\max}}{\arg\min}\ BIC(p)$$

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

$\Longrightarrow$ On utilise le **critère BIC**

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

$$\Longrightarrow \widehat{p} = \underset{0 \le p \le p_{\max}}{\arg\min}\ BIC(p)$$

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

$\Longrightarrow$ On utilise le **critère BIC**

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

$$\Longrightarrow \widehat{p} = \underset{0 \le p \le p_{\max}}{\arg\min}\ BIC(p)$$

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

$\Longrightarrow$ On utilise le **critère BIC**

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

$$\Longrightarrow \widehat{p} = \underset{0 \le p \le p_{\max}}{\arg\min}\ BIC(p)$$

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

$\Longrightarrow$ On utilise le **critère BIC**

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

$$\Longrightarrow \widehat{p} = \underset{0 \leq p \leq p_{\max}}{\arg\min}\ BIC(p)$$

Convergence des estimateurs

Quelques résultats théoriques

Propriété

On suppose que l'on a

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

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

Si (u_t) *bruit blanc* *et si* $\max_{1 \le i \le n} \left(Z \left({}^t Z\, Z \right)^{-1} {}^t Z \right)_{ii} \underset{n \to +\infty}{\longrightarrow} 0$, *alors*

$$\widehat{\sigma}^{-1} \left({}^t Z\, Z \right) \left({}^t(\widehat{a}_0, \ldots, \widehat{s}_{T-1}) - (a_0, \ldots, s_{T-1}) \right) \underset{n \to \infty}{\overset{\mathcal{L}}{\longrightarrow}} \mathcal{N}(0,1)$$

$$\text{et} \qquad \widehat{p} \underset{n \to +\infty}{\overset{\mathcal{P}}{\longrightarrow}} p.$$

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

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

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

Si (u_t) *bruit blanc et si* $\max\limits_{1\leq i\leq n} \left(Z\left({}^tZ\,Z\right)^{-1}\,{}^tZ\right)_{ii} \underset{n\to+\infty}{\longrightarrow} 0$, *alors*

$$\widehat{\sigma}^{-1}\left({}^tZ\,Z\right)\left({}^t(\widehat{a}_0,\ldots,\widehat{s}_{T-1}) - (a_0,\ldots,s_{T-1})\right) \underset{n\to\infty}{\overset{\mathcal{L}}{\longrightarrow}} \mathcal{N}(0,1)$$

$$\text{et} \qquad \widehat{p} \underset{n\to+\infty}{\overset{\mathcal{P}}{\longrightarrow}} p.$$

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

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

et $Z = \left[1 \;\; G_1 \;\; \cdots \;\; G_p \;\; H_1 \cdots \;\; H_{T-1}\right]$.

Si (u_t) *bruit blanc* *et si* $\max\limits_{1 \leq i \leq n} \left(Z\left({}^tZ\,Z\right)^{-1}\,{}^tZ\right)_{ii} \underset{n \to +\infty}{\longrightarrow} 0$, *alors*

$$\widehat{\sigma}^{-1}\left({}^tZ\,Z\right)\left({}^t(\widehat{a}_0, \ldots, \widehat{s}_{T-1}) - (a_0, \ldots, s_{T-1})\right) \xrightarrow[n \to \infty]{\mathcal{L}} \mathcal{N}(0,1)$$

et $\widehat{p} \xrightarrow[n \to +\infty]{\mathcal{P}} p$.

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

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

et $Z = \left[1 \;\; G_1 \;\; \cdots \;\; G_p \;\; H_1 \cdots \;\; H_{T-1}\right]$.

Si (u_t) *bruit blanc et si* $\max_{1 \le i \le n} \left(Z \left({}^t Z\, Z\right)^{-1} {}^t Z\right)_{ii} \underset{n \to +\infty}{\longrightarrow} 0$, *alors*

$$\widehat{\sigma}^{-1} \left({}^t Z\, Z\right) \left({}^t(\widehat{a}_0, \ldots, \widehat{s}_{T-1}) - (a_0, \ldots, s_{T-1})\right) \xrightarrow[n \to \infty]{\mathcal{L}} \mathcal{N}(0, 1)$$

et $\widehat{p} \xrightarrow[n \to +\infty]{\mathcal{P}} p$.

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

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

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

Si (u_t) *bruit blanc et si* $\max_{1 \le i \le n} \left(Z\left({}^tZ\,Z\right)^{-1}\, {}^tZ\right)_{ii} \xrightarrow[n \to +\infty]{} 0$*, alors*

$$\widehat{\sigma}^{-1}\left({}^tZ\,Z\right)\left({}^t(\widehat{a}_0, \ldots, \widehat{s}_{T-1}) - (a_0, \ldots, s_{T-1})\right) \xrightarrow[n \to \infty]{\mathcal{L}} \mathcal{N}(0,1)$$

et $\widehat{p} \xrightarrow[n \to +\infty]{\mathcal{P}} p$.

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

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

Quelques résultats théoriques

Propriété

On suppose que l'on a

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

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

Si (u_t) *bruit blanc et si* $\max\limits_{1\leq i\leq n} \left(Z\left({}^tZ\,Z\right)^{-1}\,{}^tZ\right)_{ii} \xrightarrow[n\to+\infty]{} 0$, *alors*

$$\widehat{\sigma}^{-1}\left({}^tZ\,Z\right)\left({}^t(\widehat{a}_0,\ldots,\widehat{s}_{T-1}) - (a_0,\ldots,s_{T-1})\right) \xrightarrow[n\to\infty]{\mathcal{L}} \mathcal{N}(0,1)$$

$$\text{et} \qquad \widehat{p} \xrightarrow[n\to+\infty]{\mathcal{P}} p.$$

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

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

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

Si (u_t) *bruit blanc et si* $\max_{1 \leq i \leq n} \left(Z \left({}^t Z\, Z \right)^{-1} {}^t Z \right)_{ii} \underset{n \to +\infty}{\longrightarrow} 0$*, alors*

$$\widehat{\sigma}^{-1} \left({}^t Z\, Z \right) \left({}^t (\widehat{a}_0, \ldots, \widehat{s}_{T-1}) - (a_0, \ldots, s_{T-1}) \right) \xrightarrow[n \to \infty]{\mathcal{L}} \mathcal{N}(0, 1)$$

$$\text{et} \qquad \widehat{p} \xrightarrow[n \to +\infty]{\mathcal{P}} p.$$

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

$\Longrightarrow$ Extension en cas de **variables exogènes** observées aux mêmes instants

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

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

Extension en cas de variables exogènes observées

Tendance et saisonnalité estimées jusqu'ici comme fonctions du temps

$\Longrightarrow$ Extension en cas de **variables exogènes** observées aux mêmes instants

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

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

Extension en cas de variables exogènes observées

Tendance et saisonnalité estimées jusqu'ici comme fonctions du temps

$\Longrightarrow$ Extension en cas de **variables exogènes** observées aux mêmes instants

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

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

Extension en cas de variables exogènes observées

Tendance et saisonnalité estimées jusqu'ici comme fonctions du temps

$\Longrightarrow$ Extension en cas de **variables exogènes** observées aux mêmes instants

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

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

Extension en cas de variables exogènes observées

Tendance et saisonnalité estimées jusqu'ici comme fonctions du temps

$\Longrightarrow$ Extension en cas de **variables exogènes** observées aux mêmes instants

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

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

Extension en cas de variables exogènes observées

Tendance et saisonnalité estimées jusqu'ici comme fonctions du temps

$\Longrightarrow$ Extension en cas de **variables exogènes** observées aux mêmes instants

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

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

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 :**

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

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

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 :**

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

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

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 :**

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

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

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 :**

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

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

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 :**

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

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

Plan du cours

Semi-paramétrique $\Longrightarrow$ 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 $\Longrightarrow$ 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 $\Longrightarrow$ 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 $\Longrightarrow$ 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 $\Longrightarrow$ 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 $\Longrightarrow$ 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 $\Longrightarrow$ 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 $\Longrightarrow$ 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 :*

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

Propriété

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

$$var\left(\widehat{a}_m(t)\right) = \frac{\sigma_\varepsilon^2}{m} \quad ou \quad \frac{(m-1/2)}{m^2}\,\sigma_\varepsilon^2$$

$$\Longrightarrow \text{si } a \text{ varie peu, } \widehat{a}_m(t) \simeq a(t) \text{ et } var\left(\widehat{a}_m(t)\right) \to 0 \text{ quand } m \text{ grand}$$

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 :*

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

Propriété

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

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

$$\Longrightarrow \text{si } a \text{ varie peu, } \widehat{a}_m(t) \simeq a(t) \text{ et } var(\widehat{a}_m(t)) \to 0 \text{ quand } m \text{ grand}$$

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 :*

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

Propriété

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

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

$$\Longrightarrow \text{si } a \text{ varie peu, } \widehat{a}_m(t) \simeq a(t) \text{ et } var(\widehat{a}_m(t)) \to 0 \text{ quand } m \text{ grand}$$

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 :*

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

Propriété

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

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

$$\Longrightarrow \text{si } a \text{ varie peu, } \widehat{a}_m(t) \simeq a(t) \text{ et } var\big(\widehat{a}_m(t)\big) \to 0 \text{ quand } m \text{ grand}$$

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 :*

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

Propriété

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

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

$$\Longrightarrow \text{si } a \text{ varie peu, } \widehat{a}_m(t) \simeq a(t) \text{ et } var(\widehat{a}_m(t)) \to 0 \text{ quand } m \text{ grand}$$

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 :*

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

Propriété

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

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

$\Longrightarrow$ *si a varie peu, $\widehat{a}_m(t) \simeq a(t)$ et $var\big(\widehat{a}_m(t)\big) \to 0$ 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 :*

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

Propriété

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

$$var\left(\widehat{a}_m(t)\right) = \frac{\sigma_\varepsilon^2}{m} \quad ou \quad \frac{(m-1/2)}{m^2}\,\sigma_\varepsilon^2$$

$$\Longrightarrow \text{si } a \text{ varie peu, } \widehat{a}_m(t) \simeq a(t) \text{ et } var\left(\widehat{a}_m(t)\right) \to 0 \text{ quand } m \text{ grand}$$

Moyennes mobiles (2)

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

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

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

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

$\Longrightarrow$ Choix de k ?

Moyennes mobiles (2)

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

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

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

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

$\Longrightarrow$ Choix de k ?

Moyennes mobiles (2)

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

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

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

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

$\Longrightarrow$ Choix de k ?

Moyennes mobiles (2)

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

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

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

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

$\Longrightarrow$ Choix de k ?

Moyennes mobiles (2)

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

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

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

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

$\Longrightarrow$ Choix de k ?

Moyennes mobiles (2)

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

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

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

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

$\Longrightarrow$ Choix de k ?

Moyennes mobiles (2)

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

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

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

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

$\Longrightarrow$ Choix de k ?

Procédure de validation croisée

Soit (X_t) série temporelle et $(X_1, \ldots, X_n)$ trajectoire observée
$\widehat{S}_h$ stat. non-paramétrique de $(X_1, \ldots, X_n)$ 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* $\{h_1, \ldots, h_m\}$ *de valeurs de* h *;*
2. *Pour* h *et* $i = 1, \ldots, n$*, calcul* $\widehat{S}_h^{(i)}$ *sur* $(X_1, \ldots, X_{i-1}, X_{i+1}, \ldots, X_n)$
3. *Pour chaque* h*, calcul de* $\widehat{SCR}(h) = \sum_{i=1}^{n} \left(X_i - \widehat{S}_h^{(i)}\right)^2$
4. *On choisit* $\widehat{h} = \underset{h \in \{h_1, \ldots, h_m\}}{\arg\min} \left\{\widehat{SCR}(h)\right\}$

Procédure de validation croisée

Soit (X_t) série temporelle et $(X_1, \ldots, X_n)$ trajectoire observée
$\widehat{S}_h$ stat. non-paramétrique de $(X_1, \ldots, X_n)$ 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* $\{h_1, \ldots, h_m\}$ *de valeurs de* h ;
2. *Pour* h *et* $i = 1, \ldots, n$, *calcul* $\widehat{S}_h^{(i)}$ *sur* $(X_1, \ldots, X_{i-1}, X_{i+1}, \ldots, X_n)$
3. *Pour chaque* h, *calcul de* $\widehat{SCR}(h) = \sum_{i=1}^{n} \left(X_i - \widehat{S}_h^{(i)}\right)^2$
4. *On choisit* $\widehat{h} = \underset{h \in \{h_1, \ldots, h_m\}}{\arg\min} \left\{\widehat{SCR}(h)\right\}$

Procédure de validation croisée

Soit (X_t) série temporelle et $(X_1, \ldots, X_n)$ trajectoire observée
$\widehat{S}_h$ stat. non-paramétrique de $(X_1, \ldots, X_n)$ 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 $\{h_1, \ldots, h_m\}$ de valeurs de h ;*
2. *Pour h et $i = 1, \ldots, n$, calcul $\widehat{S}_h^{(i)}$ sur $(X_1, \ldots, X_{i-1}, X_{i+1}, \ldots, X_n)$*
3. *Pour chaque h, calcul de $\widehat{SCR}(h) = \sum_{i=1}^{n} \left(X_i - \widehat{S}_h^{(i)}\right)^2$*
4. *On choisit $\widehat{h} = \underset{h \in \{h_1, \ldots, h_m\}}{\operatorname{argmin}} \left\{\widehat{SCR}(h)\right\}$*

Procédure de validation croisée

Soit (X_t) série temporelle et $(X_1, \ldots, X_n)$ trajectoire observée
$\widehat{S}_h$ stat. non-paramétrique de $(X_1, \ldots, X_n)$ 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* $\{h_1, \ldots, h_m\}$ *de valeurs de* h *;*
2. *Pour* h *et* $i = 1, \ldots, n$*, calcul* $\widehat{S}_h^{(i)}$ *sur* $(X_1, \ldots, X_{i-1}, X_{i+1}, \ldots, X_n)$
3. *Pour chaque* h*, calcul de* $\widehat{SCR}(h) = \sum_{i=1}^{n} (X_i - \widehat{S}_h^{(i)})^2$
4. *On choisit* $\widehat{h} = \underset{h \in \{h_1, \ldots, h_m\}}{\operatorname{argmin}} \{\widehat{SCR}(h)\}$

Procédure de validation croisée

Soit (X_t) série temporelle et $(X_1, \ldots, X_n)$ trajectoire observée
$\widehat{S}_h$ stat. non-paramétrique de $(X_1, \ldots, X_n)$ 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* $\{h_1, \ldots, h_m\}$ *de valeurs de* h ;
2. *Pour* h *et* $i = 1, \ldots, n$*, calcul* $\widehat{S}_h^{(i)}$ *sur* $(X_1, \ldots, X_{i-1}, X_{i+1}, \ldots, X_n)$
3. *Pour chaque* h*, calcul de* $\widehat{SCR}(h) = \sum_{i=1}^n (X_i - \widehat{S}_h^{(i)})^2$
4. *On choisit* $\widehat{h} = \underset{h \in \{h_1, \ldots, h_m\}}{\operatorname{argmin}} \{\widehat{SCR}(h)\}$

Procédure de validation croisée

Soit (X_t) série temporelle et $(X_1, \ldots, X_n)$ trajectoire observée
$\widehat{S}_h$ stat. non-paramétrique de $(X_1, \ldots, X_n)$ 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 $\{h_1, \ldots, h_m\}$ de valeurs de h ;*
2. *Pour h et $i = 1, \ldots, n$, calcul $\widehat{S}_h^{(i)}$ sur $(X_1, \ldots, X_{i-1}, X_{i+1}, \ldots, X_n)$*
3. *Pour chaque h, calcul de $\widehat{SCR}(h) = \sum_{i=1}^n \left(X_i - \widehat{S}_h^{(i)}\right)^2$*
4. *On choisit $\widehat{h} = \underset{h \in \{h_1, \ldots, h_m\}}{\operatorname{argmin}} \left\{\widehat{SCR}(h)\right\}$*

Procédure de validation croisée

Soit (X_t) série temporelle et $(X_1, \ldots, X_n)$ trajectoire observée
$\widehat{S}_h$ stat. non-paramétrique de $(X_1, \ldots, X_n)$ 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* $\{h_1, \ldots, h_m\}$ *de valeurs de* h *;*
2. *Pour* h *et* $i = 1, \ldots, n$, *calcul* $\widehat{S}_h^{(i)}$ *sur* $(X_1, \ldots, X_{i-1}, X_{i+1}, \ldots, X_n)$
3. *Pour chaque* h, *calcul de* $\widehat{SCR}(h) = \sum_{i=1}^{n} (X_i - \widehat{S}_h^{(i)})^2$
4. *On choisit* $\widehat{h} = \underset{h \in \{h_1, \ldots, h_m\}}{\operatorname{argmin}} \{\widehat{SCR}(h)\}$

Procédure de validation croisée

Soit (X_t) série temporelle et $(X_1, \ldots, X_n)$ trajectoire observée
$\widehat{S}_h$ stat. non-paramétrique de $(X_1, \ldots, X_n)$ 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 $\{h_1, \ldots, h_m\}$ de valeurs de h ;*
2. *Pour h et $i = 1, \ldots, n$, calcul $\widehat{S}_h^{(i)}$ sur $(X_1, \ldots, X_{i-1}, X_{i+1}, \ldots, X_n)$*
3. *Pour chaque h, calcul de $\widehat{SCR}(h) = \sum_{i=1}^n \left(X_i - \widehat{S}_h^{(i)}\right)^2$*
4. *On choisit $\widehat{h} = \underset{h \in \{h_1, \ldots, h_m\}}{\operatorname{argmin}} \left\{\widehat{SCR}(h)\right\}$*

Estimateur par noyau de la tendance

Définition

On appelle noyau $K : \mathbb{R} \to \mathbb{R}$ fonction telle que $\int_{\mathbb{R}} K(t)dt = 1$ et $K(0) > 0$.

Exemple : $K_1(x) = \mathbb{I}_{x \in [-\frac{1}{2}, \frac{1}{2}]}$, $K_2(x) = \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}}$, $K_3(x) = \frac{3}{4}(1 - x^2)\mathbb{I}_{x \in [-1,1]}$

Moyenne mobile : Pour $0 < h < 1$, $\widehat{a}_{nh}(t) \simeq \frac{\sum_{j=1}^{n} X_j K_1\left(\frac{t-j}{nh}\right)}{\sum_{j=1}^{n} K_1\left(\frac{t-j}{nh}\right)}$

$\Longrightarrow$ On remplace K_1 par un autre noyau K

$$\widehat{a}_{n,h}^{ker}(t) = \frac{\sum_{j=1}^{n} X_j K\left(\frac{t-j}{nh}\right)}{\sum_{j=1}^{n} K\left(\frac{t-j}{nh}\right)}$$

$\Longrightarrow$ 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 $K : \mathbb{R} \to \mathbb{R}$ fonction telle que $\int_{\mathbb{R}} K(t)dt = 1$ et $K(0) > 0$.

Exemple : $K_1(x) = \mathbb{I}_{x \in [-\frac{1}{2}, \frac{1}{2}]}$, $K_2(x) = \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}}$, $K_3(x) = \frac{3}{4}(1 - x^2)\mathbb{I}_{x \in [-1,1]}$

Moyenne mobile : Pour $0 < h < 1$, $\widehat{a}_{nh}(t) \simeq \frac{\sum_{j=1}^{n} X_j K_1\left(\frac{t-j}{nh}\right)}{\sum_{j=1}^{n} K_1\left(\frac{t-j}{nh}\right)}$

$\Longrightarrow$ On remplace K_1 par un autre noyau K

$$\widehat{a}_{n,h}^{ker}(t) = \frac{\sum_{j=1}^{n} X_j K\left(\frac{t-j}{nh}\right)}{\sum_{j=1}^{n} K\left(\frac{t-j}{nh}\right)}$$

$\Longrightarrow$ 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 $K : \mathbb{R} \to \mathbb{R}$ fonction telle que $\int_{\mathbb{R}} K(t)dt = 1$ et $K(0) > 0$.

Exemple : $K_1(x) = \mathbb{I}_{x \in [-\frac{1}{2}, \frac{1}{2}]}$, $K_2(x) = \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}}$, $K_3(x) = \frac{3}{4}(1 - x^2)\mathbb{I}_{x \in [-1,1]}$

Moyenne mobile : Pour $0 < h < 1$, $\widehat{a}_{nh}(t) \simeq \frac{\sum_{j=1}^{n} X_j K_1\left(\frac{t-j}{nh}\right)}{\sum_{j=1}^{n} K_1\left(\frac{t-j}{nh}\right)}$

$\Longrightarrow$ On remplace K_1 par un autre noyau K

$$\widehat{a}_{n,h}^{ker}(t) = \frac{\sum_{j=1}^{n} X_j K\left(\frac{t-j}{nh}\right)}{\sum_{j=1}^{n} K\left(\frac{t-j}{nh}\right)}$$

$\Longrightarrow$ 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 $K : \mathbb{R} \to \mathbb{R}$ fonction telle que $\int_{\mathbb{R}} K(t)dt = 1$ et $K(0) > 0$.

Exemple : $K_1(x) = \mathbb{I}_{x \in [-\frac{1}{2}, \frac{1}{2}]}$, $K_2(x) = \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}}$, $K_3(x) = \frac{3}{4}(1 - x^2)\mathbb{I}_{x \in [-1,1]}$

Moyenne mobile : Pour $0 < h < 1$, $\widehat{a}_{nh}(t) \simeq \frac{\sum_{j=1}^n X_j K_1\left(\frac{t-j}{nh}\right)}{\sum_{j=1}^n K_1\left(\frac{t-j}{nh}\right)}$

$\Longrightarrow$ On remplace K_1 par un autre noyau K

$$\widehat{a}_{n,h}^{ker}(t) = \frac{\sum_{j=1}^n X_j K\left(\frac{t-j}{nh}\right)}{\sum_{j=1}^n K\left(\frac{t-j}{nh}\right)}$$

$\Longrightarrow$ 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 $K : \mathbb{R} \to \mathbb{R}$ fonction telle que $\int_{\mathbb{R}} K(t)dt = 1$ et $K(0) > 0$.

Exemple : $K_1(x) = \mathbb{I}_{x \in [-\frac{1}{2}, \frac{1}{2}]}$, $K_2(x) = \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}}$, $K_3(x) = \frac{3}{4}(1 - x^2)\mathbb{I}_{x \in [-1,1]}$

Moyenne mobile : Pour $0 < h < 1$, $\widehat{a}_{nh}(t) \simeq \frac{\sum_{j=1}^n X_j K_1\left(\frac{t-j}{nh}\right)}{\sum_{j=1}^n K_1\left(\frac{t-j}{nh}\right)}$

$\Longrightarrow$ On remplace K_1 par un autre noyau K

$$\widehat{a}_{n,h}^{ker}(t) = \frac{\sum_{j=1}^n X_j K\left(\frac{t-j}{nh}\right)}{\sum_{j=1}^n K\left(\frac{t-j}{nh}\right)}$$

$\Longrightarrow$ 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 $K : \mathbb{R} \to \mathbb{R}$ 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 $0 < h < 1$, $\widehat{a}_{nh}(t) \simeq \frac{\sum_{j=1}^{n} X_j K_1\left(\frac{t-j}{nh}\right)}{\sum_{j=1}^{n} K_1\left(\frac{t-j}{nh}\right)}$

$\Longrightarrow$ On remplace K_1 par un autre noyau K

$$\widehat{a}_{n,h}^{ker}(t) = \frac{\sum_{j=1}^{n} X_j K\left(\frac{t-j}{nh}\right)}{\sum_{j=1}^{n} K\left(\frac{t-j}{nh}\right)}$$

$\Longrightarrow$ 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 $K : \mathbb{R} \to \mathbb{R}$ fonction telle que $\int_{\mathbb{R}} K(t)dt = 1$ et $K(0) > 0$.

Exemple : $K_1(x) = \mathbb{I}_{x \in [-\frac{1}{2}, \frac{1}{2}]}$, $K_2(x) = \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}}$, $K_3(x) = \frac{3}{4}(1 - x^2)\mathbb{I}_{x \in [-1,1]}$

Moyenne mobile : Pour $0 < h < 1$, $\widehat{a}_{nh}(t) \simeq \frac{\sum_{j=1}^{n} X_j K_1\left(\frac{t-j}{nh}\right)}{\sum_{j=1}^{n} K_1\left(\frac{t-j}{nh}\right)}$

$\Longrightarrow$ On remplace K_1 par un autre noyau K

$$\widehat{a}_{n,h}^{ker}(t) = \frac{\sum_{j=1}^{n} X_j K\left(\frac{t-j}{nh}\right)}{\sum_{j=1}^{n} K\left(\frac{t-j}{nh}\right)}$$

$\Longrightarrow$ 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 $K : \mathbb{R} \to \mathbb{R}$ *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 $0 < h < 1$, $\widehat{a}_{nh}(t) \simeq \frac{\sum_{j=1}^n X_j K_1\left(\frac{t-j}{nh}\right)}{\sum_{j=1}^n K_1\left(\frac{t-j}{nh}\right)}$

$\Longrightarrow$ On remplace K_1 par un autre noyau K

$$\widehat{a}_{n,h}^{ker}(t) = \frac{\sum_{j=1}^n X_j K\left(\frac{t-j}{nh}\right)}{\sum_{j=1}^n K\left(\frac{t-j}{nh}\right)}$$

$\Longrightarrow$ 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 :

$$\widehat{a}_\beta^{(0)}(t) = \frac{\sum_{i=1}^n W_\beta(|t-i|)\, X_i}{\sum_{i=1}^n W_\beta(|t-i|)},$$

W_β fonction "poids" (à valeurs dans $[0,\infty[$ et décroissants), par exemple :

- Poids exponentiels : $W_\beta(x) = \beta^x$ et $\beta \in [0,1]$ (si $\beta = 1$, $\widehat{a}_\beta(t) = \overline{X}_n$) ;
- Poids tricubes : $W_\beta(x) = \left(1 - \left(\frac{x}{\beta n}\right)^3\right)^3 \mathbb{I}_{x \leq \beta n}$ et $\beta \in [0,1]$, βn fenêtre.

$\Longrightarrow$ estimateur LOESS (LOcally Estimated Scatterplot Smoothing).

$\Longrightarrow$ Extension à une régression pondérée de degré 1 ou 2.

Remarque : Choix de β par validation croisée !

Estimateur par lissage local de la tendance

Autre estimateur non-paramétrique possible :

$$\widehat{a}_{\beta}^{(0)}(t) = \frac{\sum_{i=1}^{n} W_\beta(|t-i|)\,X_i}{\sum_{i=1}^{n} W_\beta(|t-i|)},$$

W_β fonction "poids" (à valeurs dans $[0,\infty[$ et décroissants), par exemple :

- Poids exponentiels : $W_\beta(x) = \beta^x$ et $\beta \in [0,1]$ (si $\beta = 1$, $\widehat{a}_\beta(t) = \overline{X}_n$) ;
- Poids tricubes : $W_\beta(x) = \left(1 - \left(\frac{x}{\beta n}\right)^3\right)^3 \mathbb{I}_{x \leq \beta n}$ et $\beta \in [0,1]$, βn fenêtre.

$\Longrightarrow$ estimateur LOESS (LOcally Estimated Scatterplot Smoothing).

$\Longrightarrow$ Extension à une régression pondérée de degré 1 ou 2.

Remarque : Choix de β par validation croisée !

Estimateur par lissage local de la tendance

Autre estimateur non-paramétrique possible :

$$\widehat{a}_\beta^{(0)}(t) = \frac{\sum_{i=1}^n W_\beta(|t-i|)\, X_i}{\sum_{i=1}^n W_\beta(|t-i|)},$$

W_β fonction "poids" (à valeurs dans $[0,\infty[$ et décroissants), par exemple :

- Poids exponentiels : $W_\beta(x) = \beta^x$ et $\beta \in [0,1]$ (si $\beta = 1$, $\widehat{a}_\beta(t) = \overline{X}_n$) ;
- Poids tricubes : $W_\beta(x) = \left(1 - \left(\frac{x}{\beta n}\right)^3\right)^3 \mathbb{I}_{x \leq \beta n}$ et $\beta \in [0,1]$, βn fenêtre.

$\Longrightarrow$ estimateur LOESS (LOcally Estimated Scatterplot Smoothing).

$\Longrightarrow$ Extension à une régression pondérée de degré 1 ou 2.

Remarque : Choix de β par validation croisée !

Estimateur par lissage local de la tendance

Autre estimateur non-paramétrique possible :

$$\widehat{a}_\beta^{(0)}(t) = \frac{\sum_{i=1}^n W_\beta(|t-i|)\, X_i}{\sum_{i=1}^n W_\beta(|t-i|)},$$

W_β fonction "poids" (à valeurs dans $[0,\infty[$ et décroissants), par exemple :

- Poids exponentiels : $W_\beta(x) = \beta^x$ et $\beta \in [0,1]$ (si $\beta = 1$, $\widehat{a}_\beta(t) = \overline{X}_n$) ;
- Poids tricubes : $W_\beta(x) = \left(1 - \left(\frac{x}{\beta n}\right)^3\right)^3 \mathbb{I}_{x \leq \beta n}$ et $\beta \in [0,1]$, βn fenêtre.

$\Longrightarrow$ estimateur LOESS (LOcally Estimated Scatterplot Smoothing).

$\Longrightarrow$ Extension à une régression pondérée de degré 1 ou 2.

Remarque : Choix de β par validation croisée !

Estimateur par lissage local de la tendance

Autre estimateur non-paramétrique possible :

$$\widehat{a}_\beta^{(0)}(t) = \frac{\sum_{i=1}^n W_\beta(|t-i|)\, X_i}{\sum_{i=1}^n W_\beta(|t-i|)},$$

W_β fonction "poids" (à valeurs dans $[0,\infty[$ et décroissants), par exemple :

- Poids exponentiels : $W_\beta(x) = \beta^x$ et $\beta \in [0,1]$ (si $\beta = 1$, $\widehat{a}_\beta(t) = \overline{X}_n$) ;
- Poids tricubes : $W_\beta(x) = \left(1 - \left(\frac{x}{\beta n}\right)^3\right)^3 \mathbb{I}_{x \leq \beta n}$ et $\beta \in [0,1]$, βn fenêtre.

$\Longrightarrow$ estimateur LOESS (LOcally Estimated Scatterplot Smoothing).

$\Longrightarrow$ Extension à une régression pondérée de degré 1 ou 2.

Remarque : Choix de β par validation croisée !

Estimateur par lissage local de la tendance

Autre estimateur non-paramétrique possible :

$$\widehat{a}_\beta^{(0)}(t) = \frac{\sum_{i=1}^n W_\beta(|t-i|)\, X_i}{\sum_{i=1}^n W_\beta(|t-i|)},$$

W_β fonction "poids" (à valeurs dans $[0,\infty[$ et décroissants), par exemple :

- Poids exponentiels : $W_\beta(x) = \beta^x$ et $\beta \in [0,1]$ (si $\beta = 1$, $\widehat{a}_\beta(t) = \overline{X}_n$) ;
- Poids tricubes : $W_\beta(x) = \left(1 - \left(\frac{x}{\beta n}\right)^3\right)^3 \mathbb{I}_{x \leq \beta n}$ et $\beta \in [0,1]$, βn fenêtre.

$\Longrightarrow$ estimateur LOESS (LOcally Estimated Scatterplot Smoothing).

$\Longrightarrow$ Extension à une régression pondérée de degré 1 ou 2.

Remarque : Choix de β par validation croisée !

Estimateur par lissage local de la tendance

Autre estimateur non-paramétrique possible :

$$\widehat{a}_\beta^{(0)}(t) = \frac{\sum_{i=1}^n W_\beta(|t-i|)\, X_i}{\sum_{i=1}^n W_\beta(|t-i|)},$$

W_β fonction "poids" (à valeurs dans $[0,\infty[$ et décroissants), par exemple :

- Poids exponentiels : $W_\beta(x) = \beta^x$ et $\beta \in [0,1]$ (si $\beta = 1$, $\widehat{a}_\beta(t) = \overline{X}_n$) ;
- Poids tricubes : $W_\beta(x) = \left(1 - \left(\frac{x}{\beta n}\right)^3\right)^3 \mathbb{I}_{x \leq \beta n}$ et $\beta \in [0,1]$, βn fenêtre.

$\Longrightarrow$ estimateur LOESS (LOcally Estimated Scatterplot Smoothing).

$\Longrightarrow$ Extension à une régression pondérée de degré 1 ou 2.

Remarque : Choix de β par validation croisée !

Estimateur par lissage local de la tendance

Autre estimateur non-paramétrique possible :

$$\widehat{a}_\beta^{(0)}(t) = \frac{\sum_{i=1}^n W_\beta(|t-i|)\, X_i}{\sum_{i=1}^n W_\beta(|t-i|)},$$

W_β fonction "poids" (à valeurs dans $[0,\infty[$ et décroissants), par exemple :

- Poids exponentiels : $W_\beta(x) = \beta^x$ et $\beta \in [0,1]$ (si $\beta = 1$, $\widehat{a}_\beta(t) = \overline{X}_n$) ;
- Poids tricubes : $W_\beta(x) = \left(1 - \left(\frac{x}{\beta n}\right)^3\right)^3 \mathbb{I}_{x \leq \beta n}$ et $\beta \in [0,1]$, βn fenêtre.

$\Longrightarrow$ estimateur LOESS (LOcally Estimated Scatterplot Smoothing).

$\Longrightarrow$ Extension à une régression pondérée de degré 1 ou 2.

Remarque : Choix de β par validation croisée !

Estimateur par lissage local de la tendance

Autre estimateur non-paramétrique possible :

$$\widehat{a}_\beta^{(0)}(t) = \frac{\sum_{i=1}^n W_\beta(|t-i|)\, X_i}{\sum_{i=1}^n W_\beta(|t-i|)},$$

W_β fonction "poids" (à valeurs dans $[0,\infty[$ et décroissants), par exemple :

- Poids exponentiels : $W_\beta(x) = \beta^x$ et $\beta \in [0,1]$ (si $\beta = 1$, $\widehat{a}_\beta(t) = \overline{X}_n$) ;
- Poids tricubes : $W_\beta(x) = \left(1 - \left(\frac{x}{\beta n}\right)^3\right)^3 \mathbb{I}_{x \leq \beta n}$ et $\beta \in [0,1]$, βn fenêtre.

$\Longrightarrow$ estimateur LOESS (LOcally Estimated Scatterplot Smoothing).

$\Longrightarrow$ Extension à une régression pondérée de degré 1 ou 2.

Remarque : Choix de β par validation croisée !

LOESS de degré d

Extension à un modèle de régression locale de degré d :

$$Z_d^t = \begin{pmatrix} 1 & t-1 & \cdots & (t-1)^d \\ 1 & t-2 & \cdots & (t-2)^d \\ \vdots & \vdots & & \vdots \\ 1 & t-n & \cdots & (t-n)^d \end{pmatrix} \quad W_\beta^t = \begin{pmatrix} W_\beta(|t-1|) & 0 & \cdots & 0 \\ 0 & W_\beta(|t-2|) & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & W_\beta(|t-n|) \end{pmatrix}$$

$$\Longrightarrow \widehat{a}_\beta^{(d)}(t) = \left((Z_d^t)^\top \times W_\beta^t \times Z_d^t\right)^{-1} (Z_d^t)^\top \times W_\beta^t \times \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{pmatrix}$$

$\Longrightarrow$ Moindres carrés pondérés : LOESS de degré d

Remarque : On retrouve $\widehat{a}_\beta^{(0)}(t)$ avec $d = 0$

LOESS de degré d

Extension à un modèle de régression locale de degré d :

$$Z_d^t = \begin{pmatrix} 1 & t-1 & \cdots & (t-1)^d \\ 1 & t-2 & \cdots & (t-2)^d \\ \vdots & \vdots & & \vdots \\ 1 & t-n & \cdots & (t-n)^d \end{pmatrix} \quad W_\beta^t = \begin{pmatrix} W_\beta(|t-1|) & 0 & \cdots & 0 \\ 0 & W_\beta(|t-2|) & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & W_\beta(|t-n|) \end{pmatrix}$$

$$\Longrightarrow \widehat{a}_\beta^{(d)}(t) = \left((Z_d^t)^\top \times W_\beta^t \times Z_d^t\right)^{-1} (Z_d^t)^\top \times W_\beta^t \times \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{pmatrix}$$

$\Longrightarrow$ Moindres carrés pondérés : LOESS de degré d

Remarque : On retrouve $\widehat{a}_\beta^{(0)}(t)$ avec $d = 0$

LOESS de degré d

Extension à un modèle de régression locale de degré d :

$$Z_d^t = \begin{pmatrix} 1 & t-1 & \cdots & (t-1)^d \\ 1 & t-2 & \cdots & (t-2)^d \\ \vdots & \vdots & & \vdots \\ 1 & t-n & \cdots & (t-n)^d \end{pmatrix} \quad W_\beta^t = \begin{pmatrix} W_\beta(|t-1|) & 0 & \cdots & 0 \\ 0 & W_\beta(|t-2|) & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & W_\beta(|t-n|) \end{pmatrix}$$

$$\Longrightarrow \widehat{a}_\beta^{(d)}(t) = \left((Z_d^t)^\top \times W_\beta^t \times Z_d^t\right)^{-1} (Z_d^t)^\top \times W_\beta^t \times \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{pmatrix}$$

$\Longrightarrow$ Moindres carrés pondérés : LOESS de degré d

Remarque : On retrouve $\widehat{a}_\beta^{(0)}(t)$ avec $d = 0$

LOESS de degré d

Extension à un modèle de régression locale de degré d :

$$Z_d^t = \begin{pmatrix} 1 & t-1 & \cdots & (t-1)^d \\ 1 & t-2 & \cdots & (t-2)^d \\ \vdots & \vdots & & \vdots \\ 1 & t-n & \cdots & (t-n)^d \end{pmatrix} \quad W_\beta^t = \begin{pmatrix} W_\beta(|t-1|) & 0 & \cdots & 0 \\ 0 & W_\beta(|t-2|) & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & W_\beta(|t-n|) \end{pmatrix}$$

$$\Longrightarrow \widehat{a}_\beta^{(d)}(t) = \left((Z_d^t)^\top \times W_\beta^t \times Z_d^t\right)^{-1} (Z_d^t)^\top \times W_\beta^t \times \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{pmatrix}$$

$\Longrightarrow$ Moindres carrés pondérés : LOESS de degré d

Remarque : On retrouve $\widehat{a}_\beta^{(0)}(t)$ avec $d = 0$

LOESS de degré d

Extension à un modèle de régression locale de degré d :

$$Z_d^t = \begin{pmatrix} 1 & t-1 & \cdots & (t-1)^d \\ 1 & t-2 & \cdots & (t-2)^d \\ \vdots & \vdots & & \vdots \\ 1 & t-n & \cdots & (t-n)^d \end{pmatrix} \quad W_\beta^t = \begin{pmatrix} W_\beta(|t-1|) & 0 & \cdots & 0 \\ 0 & W_\beta(|t-2|) & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & W_\beta(|t-n|) \end{pmatrix}$$

$$\Longrightarrow \widehat{a}_\beta^{(d)}(t) = \left((Z_d^t)^\top \times W_\beta^t \times Z_d^t\right)^{-1} (Z_d^t)^\top \times W_\beta^t \times \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{pmatrix}$$

$\Longrightarrow$ Moindres carrés pondérés : LOESS de degré d

Remarque : On retrouve $\widehat{a}_\beta^{(0)}(t)$ avec $d = 0$

Extension en cas de variables exogènes observées

Tendance et saisonnalité estimées jusqu'ici comme fonctions du temps

$\Longrightarrow$ Extension en cas de **variables exogènes** observées aux mêmes instants

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

$\Longrightarrow$ Méthodes de l'**apprentissage statistique** :

$$X_t = h(t, s(t), Y_t^{(1)}, \ldots, Y_t^{(d)}, u_t) \quad \text{pour tout } t$$

$$\text{où } s(t) = s_1\, h_1(t) + \cdots + s_{T-1}\, h_{T-1}(t)$$

$\Longrightarrow$ $\widehat{h}$ par forêts aléatoires, XGBoost, Deep learning...

Extension en cas de variables exogènes observées

Tendance et saisonnalité estimées jusqu'ici comme fonctions du temps

$\Longrightarrow$ Extension en cas de **variables exogènes** observées aux mêmes instants

$\Longrightarrow$ On suppose observés $(X_t)_{1\leq t\leq n}$ mais aussi $(Y_t)_{1\leq t\leq n}$ où

$$Y_t = (Y_t^{(1)}, \ldots, Y^{(d)}), \text{ où } Y^{(1)}, \ldots, Y^{(d)} \text{ variables exogènes}$$

$\Longrightarrow$ Méthodes de l'**apprentissage statistique** :

$$X_t = h(t, s(t), Y_t^{(1)}, \ldots, Y_t^{(d)}, u_t) \quad \text{pour tout } t$$

$$\text{où } s(t) = s_1\, h_1(t) + \cdots + s_{T-1}\, h_{T-1}(t)$$

$\Longrightarrow$ $\widehat{h}$ par forêts aléatoires, XGBoost, Deep learning...

Extension en cas de variables exogènes observées

Tendance et saisonnalité estimées jusqu'ici comme fonctions du temps

$\Longrightarrow$ Extension en cas de **variables exogènes** observées aux mêmes instants

$\Longrightarrow$ On suppose observés $(X_t)_{1\leq t\leq n}$ mais aussi $(Y_t)_{1\leq t\leq n}$ où

$$Y_t = (Y_t^{(1)}, \dots, Y^{(d)}), \text{ où } Y^{(1)}, \dots, Y^{(d)} \text{ variables exogènes}$$

$\Longrightarrow$ Méthodes de l'**apprentissage statistique** :

$$X_t = h(t, s(t), Y_t^{(1)}, \dots, Y_t^{(d)}, u_t) \quad \text{pour tout } t$$

$$\text{où } s(t) = s_1\, h_1(t) + \cdots + s_{T-1}\, h_{T-1}(t)$$

$\Longrightarrow$ $\widehat{h}$ par forêts aléatoires, XGBoost, Deep learning...

Extension en cas de variables exogènes observées

Tendance et saisonnalité estimées jusqu'ici comme fonctions du temps

$\Longrightarrow$ Extension en cas de **variables exogènes** observées aux mêmes instants

$\Longrightarrow$ On suppose observés $(X_t)_{1\leq t\leq n}$ mais aussi $(Y_t)_{1\leq t\leq n}$ où

$$Y_t = (Y_t^{(1)}, \dots, Y^{(d)}),\ \text{où}\ Y^{(1)}, \dots, Y^{(d)}\ \text{variables exogènes}$$

$\Longrightarrow$ Méthodes de l'**apprentissage statistique** :

$$X_t = h(t, s(t), Y_t^{(1)}, \dots, Y_t^{(d)}, u_t) \quad \text{pour tout } t$$

$$\text{où } s(t) = s_1\, h_1(t) + \cdots + s_{T-1}\, h_{T-1}(t)$$

$\Longrightarrow$ $\widehat{h}$ par forêts aléatoires, XGBoost, Deep learning...

Extension en cas de variables exogènes observées

Tendance et saisonnalité estimées jusqu'ici comme fonctions du temps

$\Longrightarrow$ Extension en cas de **variables exogènes** observées aux mêmes instants

$\Longrightarrow$ On suppose observés $(X_t)_{1\leq t\leq n}$ mais aussi $(Y_t)_{1\leq t\leq n}$ où

$$Y_t = (Y_t^{(1)}, \dots, Y^{(d)}), \text{ où } Y^{(1)}, \dots, Y^{(d)} \text{ variables exogènes}$$

$\Longrightarrow$ Méthodes de l'**apprentissage statistique** :

$$X_t = h(t, s(t), Y_t^{(1)}, \dots, Y_t^{(d)}, u_t) \quad \text{pour tout } t$$

$$\text{où } s(t) = s_1\, h_1(t) + \cdots + s_{T-1}\, h_{T-1}(t)$$

$\Longrightarrow$ $\widehat{h}$ par forêts aléatoires, XGBoost, Deep learning...

Extension en cas de variables exogènes observées

Tendance et saisonnalité estimées jusqu'ici comme fonctions du temps

$\Longrightarrow$ Extension en cas de **variables exogènes** observées aux mêmes instants

$\Longrightarrow$ On suppose observés $(X_t)_{1\leq t\leq n}$ mais aussi $(Y_t)_{1\leq t\leq n}$ où

$$Y_t = (Y_t^{(1)}, \dots, Y^{(d)}), \text{ où } Y^{(1)}, \dots, Y^{(d)} \text{ variables exogènes}$$

$\Longrightarrow$ Méthodes de l'**apprentissage statistique** :

$$X_t = h(t, s(t), Y_t^{(1)}, \dots, Y_t^{(d)}, u_t) \quad \text{pour tout } t$$

$$\text{où } s(t) = s_1\, h_1(t) + \cdots + s_{T-1}\, h_{T-1}(t)$$

$\Longrightarrow$ $\widehat{h}$ par forêts aléatoires, XGBoost, Deep learning...

Extension en cas de variables exogènes observées

Tendance et saisonnalité estimées jusqu'ici comme fonctions du temps

$\Longrightarrow$ Extension en cas de **variables exogènes** observées aux mêmes instants

$\Longrightarrow$ On suppose observés $(X_t)_{1\le t\le n}$ mais aussi $(Y_t)_{1\le t\le n}$ où

$$Y_t = (Y_t^{(1)}, \ldots, Y^{(d)}), \text{ où } Y^{(1)}, \ldots, Y^{(d)} \text{ variables exogènes}$$

$\Longrightarrow$ Méthodes de l'**apprentissage statistique** :

$$X_t = h(t, s(t), Y_t^{(1)}, \ldots, Y_t^{(d)}, u_t) \quad \text{pour tout } t$$

$$\text{où } s(t) = s_1\, h_1(t) + \cdots + s_{T-1}\, h_{T-1}(t)$$

$\Longrightarrow$ $\widehat{h}$ 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 $(\widehat{u}_t)$;
  - Garder $\widehat{s}$ pour prédire ;
  - Prédire la tendance par filtrage Holt-Winter ;
  - Prédire la partie bruit grâce au modèle obtenu pour $(\widehat{u}_t)$.
  - Sommer le tout : $\widehat{X}_{n+h} = \widehat{a}(n+h) + \widehat{s}(n+h) + \widehat{u}_{n+h}$

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 $(\widehat{u}_t)$;
  - Garder $\widehat{s}$ pour prédire ;
  - Prédire la tendance par filtrage Holt-Winter ;
  - Prédire la partie bruit grâce au modèle obtenu pour $(\widehat{u}_t)$.
  - Sommer le tout : $\widehat{X}_{n+h} = \widehat{a}(n+h) + \widehat{s}(n+h) + \widehat{u}_{n+h}$

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 $(\widehat{u}_t)$;
- Garder $\widehat{s}$ pour prédire ;
- Prédire la tendance par filtrage Holt-Winter ;
- Prédire la partie bruit grâce au modèle obtenu pour $(\widehat{u}_t)$.
- Sommer le tout : $\widehat{X}_{n+h} = \widehat{a}(n+h) + \widehat{s}(n+h) + \widehat{u}_{n+h}$

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 $(\widehat{u}_t)$;
- Garder $\widehat{s}$ pour prédire ;
- Prédire la tendance par filtrage Holt-Winter ;
- Prédire la partie bruit grâce au modèle obtenu pour $(\widehat{u}_t)$.
- Sommer le tout : $\widehat{X}_{n+h} = \widehat{a}(n+h) + \widehat{s}(n+h) + \widehat{u}_{n+h}$

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 $(\widehat{u}_t)$;
  - Garder $\widehat{s}$ pour prédire ;
  - Prédire la tendance par filtrage Holt-Winter ;
  - Prédire la partie bruit grâce au modèle obtenu pour $(\widehat{u}_t)$.
  - Sommer le tout : $\widehat{X}_{n+h} = \widehat{a}(n+h) + \widehat{s}(n+h) + \widehat{u}_{n+h}$

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 $(\widehat{u}_t)$;
  - Garder $\widehat{s}$ pour prédire ;
  - Prédire la tendance par filtrage Holt-Winter ;
  - Prédire la partie bruit grâce au modèle obtenu pour $(\widehat{u}_t)$.
  - Sommer le tout : $\widehat{X}_{n+h} = \widehat{a}(n+h) + \widehat{s}(n+h) + \widehat{u}_{n+h}$

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 $(\widehat{u}_t)$;
  - Garder $\widehat{s}$ pour prédire ;
  - Prédire la tendance par filtrage Holt-Winter ;
  - Prédire la partie bruit grâce au modèle obtenu pour $(\widehat{u}_t)$.
  - Sommer le tout : $\widehat{X}_{n+h} = \widehat{a}(n+h) + \widehat{s}(n+h) + \widehat{u}_{n+h}$

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 $(\widehat{u}_t)$;
  - Garder $\widehat{s}$ pour prédire ;
  - Prédire la tendance par filtrage Holt-Winter ;
  - Prédire la partie bruit grâce au modèle obtenu pour $(\widehat{u}_t)$.
  - Sommer le tout : $\widehat{X}_{n+h} = \widehat{a}(n+h) + \widehat{s}(n+h) + \widehat{u}_{n+h}$

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 $(\widehat{u}_t)$;
  - Garder $\widehat{s}$ pour prédire ;
  - Prédire la tendance par filtrage Holt-Winter ;
  - Prédire la partie bruit grâce au modèle obtenu pour $(\widehat{u}_t)$.
  - Sommer le tout : $\widehat{X}_{n+h} = \widehat{a}(n+h) + \widehat{s}(n+h) + \widehat{u}_{n+h}$

Position du problème

Soit $(X_1, \ldots, X_n)$ une trajectoire observée de $(X_t)_{t\in\mathbb{Z}}$.

- Si (X_t) ne possède ni tendance et saisonnalité

$\Longrightarrow$ On veut directement un modèle pour (X_t)

- Si (X_t) possède une tendance et/ou saisonnalité additive ;

$\Longrightarrow$ On obtient $\widehat{a}(\cdot)$ et $\widehat{s}(\cdot)$
$\Longrightarrow$ On cherche un modèle pour $(\widehat{u}(t))_{t\in\mathbb{Z}}$ où $\widehat{u}(t) = X_t - \widehat{a}(t) - \widehat{s}(t)$

- Un modèle pour quoi faire ?

$\Longrightarrow$ Pour améliorer la prédiction, évaluer un risque...

Position du problème

Soit $(X_1, \ldots, X_n)$ une trajectoire observée de $(X_t)_{t\in\mathbb{Z}}$.

- Si (X_t) ne possède ni tendance et saisonnalité

  $\Longrightarrow$ On veut directement un modèle pour (X_t)

- Si (X_t) possède une tendance et/ou saisonnalité additive ;

  $\Longrightarrow$ On obtient $\widehat{a}(\cdot)$ et $\widehat{s}(\cdot)$
  $\Longrightarrow$ On cherche un modèle pour $(\widehat{u}(t))_{t\in\mathbb{Z}}$ où $\widehat{u}(t) = X_t - \widehat{a}(t) - \widehat{s}(t)$

- Un modèle pour quoi faire ?

  $\Longrightarrow$ Pour améliorer la prédiction, évaluer un risque...

Position du problème

Soit $(X_1, \ldots, X_n)$ une trajectoire observée de $(X_t)_{t\in\mathbb{Z}}$.

- Si (X_t) ne possède ni tendance et saisonnalité

$\Longrightarrow$ On veut directement un modèle pour (X_t)

- Si (X_t) possède une tendance et/ou saisonnalité additive ;

$\Longrightarrow$ On obtient $\widehat{a}(\cdot)$ et $\widehat{s}(\cdot)$
$\Longrightarrow$ On cherche un modèle pour $(\widehat{u}(t))_{t\in\mathbb{Z}}$ où $\widehat{u}(t) = X_t - \widehat{a}(t) - \widehat{s}(t)$

- Un modèle pour quoi faire ?

$\Longrightarrow$ Pour améliorer la prédiction, évaluer un risque...

Position du problème

Soit $(X_1, \ldots, X_n)$ une trajectoire observée de $(X_t)_{t\in\mathbb{Z}}$.

- Si (X_t) ne possède ni tendance et saisonnalité

$\Longrightarrow$ On veut directement un modèle pour (X_t)

- Si (X_t) possède une tendance et/ou saisonnalité additive ;

$\Longrightarrow$ On obtient $\widehat{a}(\cdot)$ et $\widehat{s}(\cdot)$
$\Longrightarrow$ On cherche un modèle pour $(\widehat{u}(t))_{t\in\mathbb{Z}}$ où $\widehat{u}(t) = X_t - \widehat{a}(t) - \widehat{s}(t)$

- Un modèle pour quoi faire ?

$\Longrightarrow$ Pour améliorer la prédiction, évaluer un risque...

Position du problème

Soit $(X_1, \ldots, X_n)$ une trajectoire observée de $(X_t)_{t\in\mathbb{Z}}$.

- Si (X_t) ne possède ni tendance et saisonnalité

  $\Longrightarrow$ On veut directement un modèle pour (X_t)

- Si (X_t) possède une tendance et/ou saisonnalité additive ;

  $\Longrightarrow$ On obtient $\widehat{a}(\cdot)$ et $\widehat{s}(\cdot)$

  $\Longrightarrow$ On cherche un modèle pour $(\widehat{u}(t))_{t\in\mathbb{Z}}$ où $\widehat{u}(t) = X_t - \widehat{a}(t) - \widehat{s}(t)$

- Un modèle pour quoi faire ?

  $\Longrightarrow$ Pour améliorer la prédiction, évaluer un risque...

Position du problème

Soit $(X_1, \ldots, X_n)$ une trajectoire observée de $(X_t)_{t \in \mathbb{Z}}$.

- Si (X_t) ne possède ni tendance et saisonnalité

  $\Longrightarrow$ On veut directement un modèle pour (X_t)

- Si (X_t) possède une tendance et/ou saisonnalité additive ;

  $\Longrightarrow$ On obtient $\widehat{a}(\cdot)$ et $\widehat{s}(\cdot)$
  $\Longrightarrow$ On cherche un modèle pour $(\widehat{u}(t))_{t \in \mathbb{Z}}$ où $\widehat{u}(t) = X_t - \widehat{a}(t) - \widehat{s}(t)$

- Un modèle pour quoi faire ?

  $\Longrightarrow$ 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

  $\Longrightarrow$ On veut directement un modèle pour (X_t)

- Si (X_t) possède une tendance et/ou saisonnalité additive ;

  $\Longrightarrow$ On obtient $\widehat{a}(\cdot)$ et $\widehat{s}(\cdot)$
  $\Longrightarrow$ On cherche un modèle pour $(\widehat{u}(t))_{t\in\mathbb{Z}}$ où $\widehat{u}(t) = X_t - \widehat{a}(t) - \widehat{s}(t)$

- Un modèle pour quoi faire ?

  $\Longrightarrow$ Pour améliorer la prédiction, évaluer un risque...

Position du problème

Soit $(X_1, \ldots, X_n)$ une trajectoire observée de $(X_t)_{t\in\mathbb{Z}}$.

- Si (X_t) ne possède ni tendance et saisonnalité

$\Longrightarrow$ On veut directement un modèle pour (X_t)

- Si (X_t) possède une tendance et/ou saisonnalité additive ;

$\Longrightarrow$ On obtient $\widehat{a}(\cdot)$ et $\widehat{s}(\cdot)$

$\Longrightarrow$ On cherche un modèle pour $(\widehat{u}(t))_{t\in\mathbb{Z}}$ où $\widehat{u}(t) = X_t - \widehat{a}(t) - \widehat{s}(t)$

- Un modèle pour quoi faire ?

$\Longrightarrow$ 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 $t \in \mathbb{Z}$:*

$$X_t + a_1 X_{t-1} + \cdots + a_p X_{t-p} = \varepsilon_t + b_1 \varepsilon_{t-1} + \cdots + b_q \varepsilon_{t-q}$$

où :

- *(ε_t) bruit blanc de variance σ_ε^2 ;*
- *$(a_1, \ldots, a_p, b_1, \ldots, b_q) \in \mathbb{R}^{p+q}$ 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 $t \in \mathbb{Z}$:*

$$X_t + a_1 X_{t-1} + \cdots + a_p X_{t-p} = \varepsilon_t + b_1 \varepsilon_{t-1} + \cdots + b_q \varepsilon_{t-q}$$

où :

- *(ε_t) bruit blanc de variance σ_ε^2 ;*
- *$(a_1,\ldots,a_p,b_1,\ldots,b_q) \in \mathbb{R}^{p+q}$ 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 $t \in \mathbb{Z}$:*

$$X_t + a_1 X_{t-1} + \cdots + a_p X_{t-p} = \varepsilon_t + b_1 \varepsilon_{t-1} + \cdots + b_q \varepsilon_{t-q}$$

où :

- (ε_t) *bruit blanc de variance* σ_ε^2 ;
- $(a_1, \ldots, a_p, b_1, \ldots, b_q) \in \mathbb{R}^{p+q}$ *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 $t \in \mathbb{Z}$:*

$$X_t + a_1 X_{t-1} + \cdots + a_p X_{t-p} = \varepsilon_t + b_1 \varepsilon_{t-1} + \cdots + b_q \varepsilon_{t-q}$$

où :

- (ε_t) *bruit blanc de variance* σ_ε^2 ;
- $(a_1, \ldots, a_p, b_1, \ldots, b_q) \in \mathbb{R}^{p+q}$ *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* $t \in \mathbb{Z}$ *:*

$$X_t + a_1\, X_{t-1} + \cdots + a_p\, X_{t-p} = \varepsilon_t + b_1\, \varepsilon_{t-1} + \cdots + b_q\, \varepsilon_{t-q}$$

où :

- (ε_t) *bruit blanc de variance* σ_ε^2 *;*
- $(a_1, \ldots, a_p, b_1, \ldots, b_q) \in \mathbb{R}^{p+q}$ *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* $t\in\mathbb{Z}$ *:*

$$X_t + a_1\,X_{t-1} + \cdots + a_p\,X_{t-p} = \varepsilon_t + b_1\,\varepsilon_{t-1} + \cdots + b_q\,\varepsilon_{t-q}$$

où :

- (ε_t) *bruit blanc de variance* σ_ε^2 *;*
- $(a_1,\ldots,a_p,b_1,\ldots,b_q)\in\mathbb{R}^{p+q}$ *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 $t \in \mathbb{Z}$:*

$$X_t + a_1\,X_{t-1} + \cdots + a_p\,X_{t-p} = \varepsilon_t + b_1\,\varepsilon_{t-1} + \cdots + b_q\,\varepsilon_{t-q}$$

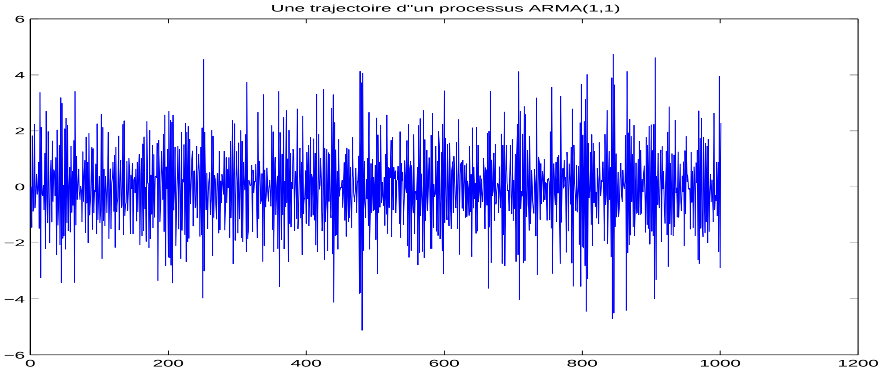
où :

- *(ε_t) bruit blanc de variance σ_ε^2 ;*
- *$(a_1, \ldots, a_p, b_1, \ldots, b_q) \in \mathbb{R}^{p+q}$ 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 :

$$\begin{aligned}
X_t &= \alpha(\alpha X_{t-2} + \varepsilon_{t-1}) + \varepsilon_t \\
&= \alpha^2 X_{t-2} + \alpha \varepsilon_{t-1} + \varepsilon_t \\
&= \alpha^2 (\alpha X_{t-3} + \varepsilon_{t-2}) + \alpha \varepsilon_{t-1} + \varepsilon_t \\
&= \alpha^3 X_{t-3} + \alpha^2 \varepsilon_{t-2} + \alpha \varepsilon_{t-1} + \varepsilon_t \\
&\vdots \\
&= \sum_{k=0}^{\infty} \alpha^k \varepsilon_{t-k} \quad \text{(existe car } |\alpha| < 1)
\end{aligned}$$

2. Soit le processus AR(1) avec $X_t = \alpha X_{t-1} + \varepsilon_t$ et $|\alpha| > 1$. Alors :

$$\begin{aligned}
X_t &= \alpha^{-1} X_{t+1} - \alpha^{-1} \varepsilon_{t+1} \\
&= \alpha^{-1} (\alpha^{-1} X_{t+2} - \alpha^{-1} \varepsilon_{t+2}) - \alpha^{-1} \varepsilon_{t+1} \\
&= \alpha^{-2} X_{t+2} - \alpha^{-2} \varepsilon_{t+2} - \alpha^{-1} \varepsilon_{t+1} \\
&\vdots \\
&= -\sum_{k=1}^{\infty} (\alpha^{-1})^k \varepsilon_{t+k} \quad \text{(existe car } |\alpha^{-1}| < 1)
\end{aligned}$$

Processus ARMA : deux exemples

1. Soit le processus AR(1) avec $X_t = \alpha\, X_{t-1} + \varepsilon_t$ et $|\alpha| < 1$. Alors :

$$\begin{aligned} X_t &= \alpha\,(\alpha X_{t-2} + \varepsilon_{t-1}) + \varepsilon_t \\ &= \alpha^2\, X_{t-2} + \alpha\,\varepsilon_{t-1} + \varepsilon_t \\ &= \alpha^2\,(\alpha X_{t-3} + \varepsilon_{t-2}) + \alpha\,\varepsilon_{t-1} + \varepsilon_t \\ &= \alpha^3\, X_{t-3} + \alpha^2\,\varepsilon_{t-2} + \alpha\,\varepsilon_{t-1} + \varepsilon_t \\ &\vdots \\ &= \sum_{k=0}^{\infty} \alpha^k\,\varepsilon_{t-k} \quad (\text{existe car } |\alpha| < 1) \end{aligned}$$

2. Soit le processus AR(1) avec $X_t = \alpha\, X_{t-1} + \varepsilon_t$ et $|\alpha| > 1$. Alors :

$$\begin{aligned} X_t &= \alpha^{-1}\, X_{t+1} - \alpha^{-1}\,\varepsilon_{t+1} \\ &= \alpha^{-1}\,(\alpha^{-1}\, X_{t+2} - \alpha^{-1}\,\varepsilon_{t+2}) - \alpha^{-1}\,\varepsilon_{t+1} \\ &= \alpha^{-2}\, X_{t+2} - \alpha^{-2}\,\varepsilon_{t+2} - \alpha^{-1}\,\varepsilon_{t+1} \\ &\vdots \\ &= -\sum_{k=1}^{\infty} (\alpha^{-1})^k\,\varepsilon_{t+k} \quad (\text{existe car } |\alpha^{-1}| < 1) \end{aligned}$$

Processus ARMA : deux exemples

1. Soit le processus AR(1) avec $X_t = \alpha\, X_{t-1} + \varepsilon_t$ et $|\alpha| < 1$. Alors :

$$
\begin{aligned}
X_t &= \alpha\,(\alpha X_{t-2} + \varepsilon_{t-1}) + \varepsilon_t \\
&= \alpha^2\, X_{t-2} + \alpha\,\varepsilon_{t-1} + \varepsilon_t \\
&= \alpha^2\,(\alpha X_{t-3} + \varepsilon_{t-2}) + \alpha\,\varepsilon_{t-1} + \varepsilon_t \\
&= \alpha^3\, X_{t-3} + \alpha^2\,\varepsilon_{t-2} + \alpha\,\varepsilon_{t-1} + \varepsilon_t \\
&\vdots \qquad \vdots \qquad\qquad \vdots \qquad\qquad \vdots \\
&= \sum_{k=0}^{\infty} \alpha^k\, \varepsilon_{t-k} \quad (\text{existe car } |\alpha| < 1)
\end{aligned}
$$

2. Soit le processus AR(1) avec $X_t = \alpha\, X_{t-1} + \varepsilon_t$ et $|\alpha| > 1$. Alors :

$$
\begin{aligned}
X_t &= \alpha^{-1}\, X_{t+1} - \alpha^{-1}\,\varepsilon_{t+1} \\
&= \alpha^{-1}\,(\alpha^{-1}\, X_{t+2} - \alpha^{-1}\,\varepsilon_{t+2}) - \alpha^{-1}\,\varepsilon_{t+1} \\
&= \alpha^{-2}\, X_{t+2} - \alpha^{-2}\,\varepsilon_{t+2} - \alpha^{-1}\,\varepsilon_{t+1} \\
&\vdots \qquad \vdots \qquad\qquad \vdots \qquad\qquad \vdots \\
&= -\sum_{k=1}^{\infty} (\alpha^{-1})^k\, \varepsilon_{t+k} \quad (\text{existe car } |\alpha^{-1}| < 1)
\end{aligned}
$$

Processus ARMA : deux exemples

1. Soit le processus AR(1) avec $X_t = \alpha\, X_{t-1} + \varepsilon_t$ et $|\alpha| < 1$. Alors :

$$
\begin{aligned}
X_t &= \alpha\,(\alpha X_{t-2} + \varepsilon_{t-1}) + \varepsilon_t \\
&= \alpha^2\, X_{t-2} + \alpha\,\varepsilon_{t-1} + \varepsilon_t \\
&= \alpha^2\,(\alpha X_{t-3} + \varepsilon_{t-2}) + \alpha\,\varepsilon_{t-1} + \varepsilon_t \\
&= \alpha^3\, X_{t-3} + \alpha^2\,\varepsilon_{t-2} + \alpha\,\varepsilon_{t-1} + \varepsilon_t \\
\vdots \quad & \quad \vdots \qquad\qquad \vdots \qquad\qquad \vdots \\
&= \sum_{k=0}^{\infty} \alpha^k\,\varepsilon_{t-k} \quad (\text{existe car } |\alpha| < 1)
\end{aligned}
$$

2. Soit le processus AR(1) avec $X_t = \alpha\, X_{t-1} + \varepsilon_t$ et $|\alpha| > 1$. Alors :

$$
\begin{aligned}
X_t &= \alpha^{-1}\, X_{t+1} - \alpha^{-1}\,\varepsilon_{t+1} \\
&= \alpha^{-1}\,(\alpha^{-1}\, X_{t+2} - \alpha^{-1}\,\varepsilon_{t+2}) - \alpha^{-1}\,\varepsilon_{t+1} \\
&= \alpha^{-2}\, X_{t+2} - \alpha^{-2}\,\varepsilon_{t+2} - \alpha^{-1}\,\varepsilon_{t+1} \\
\vdots \quad & \quad \vdots \qquad\qquad \vdots \qquad\qquad \vdots \\
&= -\sum_{k=1}^{\infty} (\alpha^{-1})^k\,\varepsilon_{t+k} \quad (\text{existe car } |\alpha^{-1}| < 1)
\end{aligned}
$$

Processus ARMA : deux exemples

1. Soit le processus AR(1) avec $X_t = \alpha\, X_{t-1} + \varepsilon_t$ et $|\alpha| < 1$. Alors :

$$
\begin{aligned}
X_t &= \alpha\,(\alpha X_{t-2} + \varepsilon_{t-1}) + \varepsilon_t \\
&= \alpha^2\, X_{t-2} + \alpha\,\varepsilon_{t-1} + \varepsilon_t \\
&= \alpha^2\,(\alpha X_{t-3} + \varepsilon_{t-2}) + \alpha\,\varepsilon_{t-1} + \varepsilon_t \\
&= \alpha^3\, X_{t-3} + \alpha^2\,\varepsilon_{t-2} + \alpha\,\varepsilon_{t-1} + \varepsilon_t \\
&\vdots \quad \vdots \qquad \vdots \qquad \vdots \\
&= \sum_{k=0}^{\infty} \alpha^k\, \varepsilon_{t-k} \quad \text{(existe car } |\alpha| < 1)
\end{aligned}
$$

2. Soit le processus AR(1) avec $X_t = \alpha\, X_{t-1} + \varepsilon_t$ et $|\alpha| > 1$. Alors :

$$
\begin{aligned}
X_t &= \alpha^{-1}\, X_{t+1} - \alpha^{-1}\,\varepsilon_{t+1} \\
&= \alpha^{-1}\,(\alpha^{-1}\, X_{t+2} - \alpha^{-1}\,\varepsilon_{t+2}) - \alpha^{-1}\,\varepsilon_{t+1} \\
&= \alpha^{-2}\, X_{t+2} - \alpha^{-2}\,\varepsilon_{t+2} - \alpha^{-1}\,\varepsilon_{t+1} \\
&\vdots \quad \vdots \qquad \vdots \qquad \vdots \\
&= -\sum_{k=1}^{\infty} (\alpha^{-1})^k\, \varepsilon_{t+k} \quad \text{(existe car } |\alpha^{-1}| < 1)
\end{aligned}
$$

Processus ARMA : deux exemples

1. Soit le processus AR(1) avec $X_t = \alpha\, X_{t-1} + \varepsilon_t$ et $|\alpha| < 1$. Alors :

$$\begin{aligned}
X_t &= \alpha\,(\alpha X_{t-2} + \varepsilon_{t-1}) + \varepsilon_t \\
&= \alpha^2\, X_{t-2} + \alpha\,\varepsilon_{t-1} + \varepsilon_t \\
&= \alpha^2\,(\alpha X_{t-3} + \varepsilon_{t-2}) + \alpha\,\varepsilon_{t-1} + \varepsilon_t \\
&= \alpha^3\, X_{t-3} + \alpha^2\,\varepsilon_{t-2} + \alpha\,\varepsilon_{t-1} + \varepsilon_t \\
&\;\;\vdots \qquad \vdots \qquad\qquad \vdots \qquad\qquad \vdots \\
&= \sum_{k=0}^{\infty} \alpha^k\, \varepsilon_{t-k} \quad \text{(existe car } |\alpha| < 1\text{)}
\end{aligned}$$

2. Soit le processus AR(1) avec $X_t = \alpha\, X_{t-1} + \varepsilon_t$ et $|\alpha| > 1$. Alors :

$$\begin{aligned}
X_t &= \alpha^{-1}\, X_{t+1} - \alpha^{-1}\,\varepsilon_{t+1} \\
&= \alpha^{-1}\,(\alpha^{-1}\, X_{t+2} - \alpha^{-1}\,\varepsilon_{t+2}) - \alpha^{-1}\,\varepsilon_{t+1} \\
&= \alpha^{-2}\, X_{t+2} - \alpha^{-2}\,\varepsilon_{t+2} - \alpha^{-1}\,\varepsilon_{t+1} \\
&\;\;\vdots \qquad \vdots \qquad\qquad \vdots \qquad\qquad \vdots \\
&= -\sum_{k=1}^{\infty} \left(\alpha^{-1}\right)^k \varepsilon_{t+k} \quad \text{(existe car } |\alpha^{-1}| < 1\text{)}
\end{aligned}$$

Processus ARMA : deux exemples

1. Soit le processus AR(1) avec $X_t = \alpha\, X_{t-1} + \varepsilon_t$ et $|\alpha| < 1$. Alors :

$$
\begin{array}{rcl}
X_t & = & \alpha\,(\alpha X_{t-2} + \varepsilon_{t-1}) + \varepsilon_t \\
 & = & \alpha^2\, X_{t-2} + \alpha\,\varepsilon_{t-1} + \varepsilon_t \\
 & = & \alpha^2\,(\alpha X_{t-3} + \varepsilon_{t-2}) + \alpha\,\varepsilon_{t-1} + \varepsilon_t \\
 & = & \alpha^3\, X_{t-3} + \alpha^2\,\varepsilon_{t-2} + \alpha\,\varepsilon_{t-1} + \varepsilon_t \\
 & \vdots & \\
 & = & \sum_{k=0}^{\infty} \alpha^k\,\varepsilon_{t-k} \quad (\text{existe car } |\alpha| < 1)
\end{array}
$$

2. Soit le processus AR(1) avec $X_t = \alpha\, X_{t-1} + \varepsilon_t$ et $|\alpha| > 1$. Alors :

$$
\begin{array}{rcl}
X_t & = & \alpha^{-1}\, X_{t+1} - \alpha^{-1}\,\varepsilon_{t+1} \\
 & = & \alpha^{-1}\,(\alpha^{-1}\, X_{t+2} - \alpha^{-1}\,\varepsilon_{t+2}) - \alpha^{-1}\,\varepsilon_{t+1} \\
 & = & \alpha^{-2}\, X_{t+2} - \alpha^{-2}\,\varepsilon_{t+2} - \alpha^{-1}\,\varepsilon_{t+1} \\
 & \vdots & \\
 & = & -\sum_{k=1}^{\infty} (\alpha^{-1})^k\,\varepsilon_{t+k} \quad (\text{existe car } |\alpha^{-1}| < 1)
\end{array}
$$

Processus ARMA : deux exemples

1. Soit le processus AR(1) avec $X_t = \alpha\, X_{t-1} + \varepsilon_t$ et $|\alpha| < 1$. Alors :

$$\begin{aligned}
X_t &= \alpha\,(\alpha X_{t-2} + \varepsilon_{t-1}) + \varepsilon_t \\
&= \alpha^2\, X_{t-2} + \alpha\,\varepsilon_{t-1} + \varepsilon_t \\
&= \alpha^2\,(\alpha X_{t-3} + \varepsilon_{t-2}) + \alpha\,\varepsilon_{t-1} + \varepsilon_t \\
&= \alpha^3\, X_{t-3} + \alpha^2\,\varepsilon_{t-2} + \alpha\,\varepsilon_{t-1} + \varepsilon_t \\
&\vdots \qquad \vdots \qquad \vdots \qquad \vdots \\
&= \sum_{k=0}^{\infty} \alpha^k\,\varepsilon_{t-k} \quad \text{(existe car } |\alpha| < 1)
\end{aligned}$$

2. Soit le processus AR(1) avec $X_t = \alpha\, X_{t-1} + \varepsilon_t$ et $|\alpha| > 1$. Alors :

$$\begin{aligned}
X_t &= \alpha^{-1}\, X_{t+1} - \alpha^{-1}\,\varepsilon_{t+1} \\
&= \alpha^{-1}\,(\alpha^{-1}\, X_{t+2} - \alpha^{-1}\,\varepsilon_{t+2}) - \alpha^{-1}\,\varepsilon_{t+1} \\
&= \alpha^{-2}\, X_{t+2} - \alpha^{-2}\,\varepsilon_{t+2} - \alpha^{-1}\,\varepsilon_{t+1} \\
&\vdots \qquad \vdots \qquad \vdots \qquad \vdots \\
&= -\sum_{k=1}^{\infty} (\alpha^{-1})^k\,\varepsilon_{t+k} \quad \text{(existe car } |\alpha^{-1}| < 1)
\end{aligned}$$

Processus ARMA : deux exemples

1. Soit le processus AR(1) avec $X_t = \alpha\, X_{t-1} + \varepsilon_t$ et $|\alpha| < 1$. Alors :

$$\begin{aligned}
X_t &= \alpha\,(\alpha X_{t-2} + \varepsilon_{t-1}) + \varepsilon_t \\
&= \alpha^2\, X_{t-2} + \alpha\,\varepsilon_{t-1} + \varepsilon_t \\
&= \alpha^2\,(\alpha X_{t-3} + \varepsilon_{t-2}) + \alpha\,\varepsilon_{t-1} + \varepsilon_t \\
&= \alpha^3\, X_{t-3} + \alpha^2\,\varepsilon_{t-2} + \alpha\,\varepsilon_{t-1} + \varepsilon_t \\
&\vdots \qquad \vdots \qquad \vdots \qquad \vdots \\
&= \sum_{k=0}^{\infty} \alpha^k\,\varepsilon_{t-k} \quad \text{(existe car } |\alpha| < 1\text{)}
\end{aligned}$$

2. Soit le processus AR(1) avec $X_t = \alpha\, X_{t-1} + \varepsilon_t$ et $|\alpha| > 1$. Alors :

$$\begin{aligned}
X_t &= \alpha^{-1}\, X_{t+1} - \alpha^{-1}\,\varepsilon_{t+1} \\
&= \alpha^{-1}\,(\alpha^{-1}\, X_{t+2} - \alpha^{-1}\,\varepsilon_{t+2}) - \alpha^{-1}\,\varepsilon_{t+1} \\
&= \alpha^{-2}\, X_{t+2} - \alpha^{-2}\,\varepsilon_{t+2} - \alpha^{-1}\,\varepsilon_{t+1} \\
&\vdots \qquad \vdots \qquad \vdots \qquad \vdots \\
&= -\sum_{k=1}^{\infty} (\alpha^{-1})^k\,\varepsilon_{t+k} \quad \text{(existe car } |\alpha^{-1}| < 1\text{)}
\end{aligned}$$

Processus ARMA : deux exemples

1. Soit le processus AR(1) avec $X_t = \alpha\, X_{t-1} + \varepsilon_t$ et $|\alpha| < 1$. Alors :

$$\begin{aligned}
X_t &= \alpha\,(\alpha X_{t-2} + \varepsilon_{t-1}) + \varepsilon_t \\
&= \alpha^2\, X_{t-2} + \alpha\, \varepsilon_{t-1} + \varepsilon_t \\
&= \alpha^2\,(\alpha X_{t-3} + \varepsilon_{t-2}) + \alpha\, \varepsilon_{t-1} + \varepsilon_t \\
&= \alpha^3\, X_{t-3} + \alpha^2\, \varepsilon_{t-2} + \alpha\, \varepsilon_{t-1} + \varepsilon_t \\
&\vdots \qquad \vdots \qquad \vdots \qquad \vdots \\
&= \sum_{k=0}^{\infty} \alpha^k\, \varepsilon_{t-k} \quad \text{(existe car } |\alpha| < 1\text{)}
\end{aligned}$$

2. Soit le processus AR(1) avec $X_t = \alpha\, X_{t-1} + \varepsilon_t$ et $|\alpha| > 1$. Alors :

$$\begin{aligned}
X_t &= \alpha^{-1}\, X_{t+1} - \alpha^{-1}\, \varepsilon_{t+1} \\
&= \alpha^{-1}\,(\alpha^{-1}\, X_{t+2} - \alpha^{-1}\, \varepsilon_{t+2}) - \alpha^{-1}\, \varepsilon_{t+1} \\
&= \alpha^{-2}\, X_{t+2} - \alpha^{-2}\, \varepsilon_{t+2} - \alpha^{-1}\, \varepsilon_{t+1} \\
&\vdots \qquad \vdots \qquad \vdots \qquad \vdots \\
&= -\sum_{k=1}^{\infty} (\alpha^{-1})^k\, \varepsilon_{t+k} \quad \text{(existe car } |\alpha^{-1}| < 1\text{)}
\end{aligned}$$

Processus ARMA : deux exemples

1. Soit le processus AR(1) avec $X_t = \alpha\, X_{t-1} + \varepsilon_t$ et $|\alpha| < 1$. Alors :

$$\begin{aligned}
X_t &= \alpha\,(\alpha X_{t-2} + \varepsilon_{t-1}) + \varepsilon_t \\
&= \alpha^2\, X_{t-2} + \alpha\,\varepsilon_{t-1} + \varepsilon_t \\
&= \alpha^2\,(\alpha X_{t-3} + \varepsilon_{t-2}) + \alpha\,\varepsilon_{t-1} + \varepsilon_t \\
&= \alpha^3\, X_{t-3} + \alpha^2\,\varepsilon_{t-2} + \alpha\,\varepsilon_{t-1} + \varepsilon_t \\
&\vdots \qquad \vdots \qquad \vdots \qquad \vdots \\
&= \sum_{k=0}^{\infty} \alpha^k\,\varepsilon_{t-k} \quad (\text{existe car } |\alpha| < 1)
\end{aligned}$$

2. Soit le processus AR(1) avec $X_t = \alpha\, X_{t-1} + \varepsilon_t$ et $|\alpha| > 1$. Alors :

$$\begin{aligned}
X_t &= \alpha^{-1}\, X_{t+1} - \alpha^{-1}\,\varepsilon_{t+1} \\
&= \alpha^{-1}\,(\alpha^{-1}\, X_{t+2} - \alpha^{-1}\,\varepsilon_{t+2}) - \alpha^{-1}\,\varepsilon_{t+1} \\
&= \alpha^{-2}\, X_{t+2} - \alpha^{-2}\,\varepsilon_{t+2} - \alpha^{-1}\,\varepsilon_{t+1} \\
&\vdots \qquad \vdots \qquad \vdots \qquad \vdots \\
&= -\sum_{k=1}^{\infty} (\alpha^{-1})^k\,\varepsilon_{t+k} \quad (\text{existe car } |\alpha^{-1}| < 1)
\end{aligned}$$

Processus ARMA : deux exemples

1. Soit le processus AR(1) avec $X_t = \alpha\, X_{t-1} + \varepsilon_t$ et $|\alpha| < 1$. Alors :

$$
\begin{aligned}
X_t &= \alpha\,(\alpha X_{t-2} + \varepsilon_{t-1}) + \varepsilon_t \\
&= \alpha^2\, X_{t-2} + \alpha\,\varepsilon_{t-1} + \varepsilon_t \\
&= \alpha^2\,(\alpha X_{t-3} + \varepsilon_{t-2}) + \alpha\,\varepsilon_{t-1} + \varepsilon_t \\
&= \alpha^3\, X_{t-3} + \alpha^2\,\varepsilon_{t-2} + \alpha\,\varepsilon_{t-1} + \varepsilon_t \\
&\vdots \qquad \vdots \qquad\qquad \vdots \qquad\qquad \vdots \\
&= \sum_{k=0}^{\infty} \alpha^k\,\varepsilon_{t-k} \quad \text{(existe car } |\alpha| < 1)
\end{aligned}
$$

2. Soit le processus AR(1) avec $X_t = \alpha\, X_{t-1} + \varepsilon_t$ et $|\alpha| > 1$. Alors :

$$
\begin{aligned}
X_t &= \alpha^{-1}\, X_{t+1} - \alpha^{-1}\,\varepsilon_{t+1} \\
&= \alpha^{-1}\,(\alpha^{-1}\, X_{t+2} - \alpha^{-1}\,\varepsilon_{t+2}) - \alpha^{-1}\,\varepsilon_{t+1} \\
&= \alpha^{-2}\, X_{t+2} - \alpha^{-2}\,\varepsilon_{t+2} - \alpha^{-1}\,\varepsilon_{t+1} \\
&\vdots \qquad \vdots \qquad\qquad \vdots \qquad\qquad \vdots \\
&= -\sum_{k=1}^{\infty} (\alpha^{-1})^k\,\varepsilon_{t+k} \quad \text{(existe car } |\alpha^{-1}| < 1)
\end{aligned}
$$

Processus ARMA : deux exemples

1. Soit le processus AR(1) avec $X_t = \alpha\, X_{t-1} + \varepsilon_t$ et $|\alpha| < 1$. Alors :

$$
\begin{aligned}
X_t &= \alpha\,(\alpha X_{t-2} + \varepsilon_{t-1}) + \varepsilon_t \\
&= \alpha^2\, X_{t-2} + \alpha\,\varepsilon_{t-1} + \varepsilon_t \\
&= \alpha^2\,(\alpha X_{t-3} + \varepsilon_{t-2}) + \alpha\,\varepsilon_{t-1} + \varepsilon_t \\
&= \alpha^3\, X_{t-3} + \alpha^2\,\varepsilon_{t-2} + \alpha\,\varepsilon_{t-1} + \varepsilon_t \\
&\vdots \qquad \vdots \qquad \vdots \qquad \vdots \\
&= \sum_{k=0}^{\infty} \alpha^k\,\varepsilon_{t-k} \quad (\text{existe car } |\alpha| < 1)
\end{aligned}
$$

2. Soit le processus AR(1) avec $X_t = \alpha\, X_{t-1} + \varepsilon_t$ et $|\alpha| > 1$. Alors :

$$
\begin{aligned}
X_t &= \alpha^{-1}\, X_{t+1} - \alpha^{-1}\,\varepsilon_{t+1} \\
&= \alpha^{-1}\,(\alpha^{-1}\, X_{t+2} - \alpha^{-1}\,\varepsilon_{t+2}) - \alpha^{-1}\,\varepsilon_{t+1} \\
&= \alpha^{-2}\, X_{t+2} - \alpha^{-2}\,\varepsilon_{t+2} - \alpha^{-1}\,\varepsilon_{t+1} \\
&\vdots \qquad \vdots \qquad \vdots \qquad \vdots \\
&= -\sum_{k=1}^{\infty} (\alpha^{-1})^k\,\varepsilon_{t+k} \quad (\text{existe car } |\alpha^{-1}| < 1)
\end{aligned}
$$

Processus ARMA : deux exemples

1. Soit le processus AR(1) avec $X_t = \alpha X_{t-1} + \varepsilon_t$ et $|\alpha| < 1$. Alors :

$$\begin{aligned} X_t &= \alpha(\alpha X_{t-2} + \varepsilon_{t-1}) + \varepsilon_t \\ &= \alpha^2 X_{t-2} + \alpha\varepsilon_{t-1} + \varepsilon_t \\ &= \alpha^2(\alpha X_{t-3} + \varepsilon_{t-2}) + \alpha\varepsilon_{t-1} + \varepsilon_t \\ &= \alpha^3 X_{t-3} + \alpha^2\varepsilon_{t-2} + \alpha\varepsilon_{t-1} + \varepsilon_t \\ &\vdots \\ &= \sum_{k=0}^{\infty} \alpha^k \varepsilon_{t-k} \quad (\text{existe car } |\alpha| < 1) \end{aligned}$$

2. Soit le processus AR(1) avec $X_t = \alpha X_{t-1} + \varepsilon_t$ et $|\alpha| > 1$. Alors :

$$\begin{aligned} X_t &= \alpha^{-1} X_{t+1} - \alpha^{-1}\varepsilon_{t+1} \\ &= \alpha^{-1}(\alpha^{-1} X_{t+2} - \alpha^{-1}\varepsilon_{t+2}) - \alpha^{-1}\varepsilon_{t+1} \\ &= \alpha^{-2} X_{t+2} - \alpha^{-2}\varepsilon_{t+2} - \alpha^{-1}\varepsilon_{t+1} \\ &\vdots \\ &= -\sum_{k=1}^{\infty} (\alpha^{-1})^k \varepsilon_{t+k} \quad (\text{existe car } |\alpha^{-1}| < 1) \end{aligned}$$

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 $|z| \leq 1$

$\Longrightarrow X_t = \sum_{k=0}^{\infty} \alpha_k\, \varepsilon_{t-k}$ dépend linéairement des valeurs passés de ε

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 $|z| \leq 1$

$$\Longrightarrow X_t = \sum_{k=0}^{\infty} \alpha_k\, \varepsilon_{t-k} \text{ dépend linéairement des valeurs passés de } \varepsilon$$

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 $|z| \leq 1$

$$\Longrightarrow X_t = \sum_{k=0}^{\infty} \alpha_k \varepsilon_{t-k} \text{ dépend linéairement des valeurs passés de } \varepsilon$$

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 $|z| \leq 1$

$$\Longrightarrow X_t = \sum_{k=0}^{\infty} \alpha_k\, \varepsilon_{t-k} \text{ dépend linéairement des valeurs passés de } \varepsilon$$

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 $|z| \leq 1$

$$\Longrightarrow X_t = \sum_{k=0}^{\infty} \alpha_k\, \varepsilon_{t-k} \text{ dépend linéairement des valeurs passés de } \varepsilon$$

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 $|z| \leq 1$

$\Longrightarrow X_t = \sum_{k=0}^{\infty} \alpha_k\, \varepsilon_{t-k}$ dépend linéairement des valeurs passés de ε

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 $|z| \leq 1$

$\Longrightarrow X_t = \sum_{k=0}^{\infty} \alpha_k\, \varepsilon_{t-k}$ dépend linéairement des valeurs passés de ε

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

$$X_t + a_1\,X_{t-1} + \cdots + a_p\,X_{t-p} = \varepsilon_t + b_1\,\varepsilon_{t-1} + \cdots + b_q\,\varepsilon_{t-q}$$

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. $r_X(k) = cov(X_0, X_k)$ *:* $\exists 0 \leq \rho' < 1,\ c' \geq 0,\ |r_X(k)| \leq c'\,{\rho'}^k\ \forall k \geq 0$ *;*

Cas particulier : Si (X_t) est un processus MA(q),

1. $\alpha_j = b_j$ pour $j = 1, \ldots, q$, 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 $0 \leq k \leq q$,

$$r_X(k) = \text{cov}\Big(\sum_{j=0}^{q} b_j\varepsilon_{-j}, \sum_{\ell=0}^{q} b_\ell\varepsilon_{k-\ell}\Big) = \sum_{j=0}^{q}\sum_{\ell=0}^{q} b_j b_\ell \mathbb{E}[\varepsilon_{-j}\varepsilon_{k-\ell}] = \sigma_\varepsilon^2 \sum_{j=0}^{q-k} b_j b_{j+k}$$

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

$$X_t + a_1 X_{t-1} + \cdots + a_p X_{t-p} = \varepsilon_t + b_1 \varepsilon_{t-1} + \cdots + b_q \varepsilon_{t-q}$$

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. $r_X(k) = cov(X_0, X_k)$ *:* $\exists 0 \leq \rho' < 1,\ c' \geq 0,\ |r_X(k)| \leq c'\,{\rho'}^k\ \forall k \geq 0$ *;*

Cas particulier : Si (X_t) est un processus MA(q),

1. $\alpha_j = b_j$ pour $j = 1, \ldots, q$, 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 $0 \leq k \leq q$,

$$r_X(k) = \text{cov}\Big(\sum_{j=0}^{q} b_j \varepsilon_{-j}, \sum_{\ell=0}^{q} b_\ell \varepsilon_{k-\ell}\Big) = \sum_{j=0}^{q}\sum_{\ell=0}^{q} b_j b_\ell \mathbb{E}[\varepsilon_{-j}\varepsilon_{k-\ell}] = \sigma_\varepsilon^2 \sum_{j=0}^{q-k} b_j b_{j+k}$$

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

$$X_t + a_1 X_{t-1} + \cdots + a_p X_{t-p} = \varepsilon_t + b_1 \varepsilon_{t-1} + \cdots + b_q \varepsilon_{t-q}$$

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. $r_X(k) = cov(X_0, X_k)$: $\exists 0 \leq \rho' < 1,\ c' \geq 0,\ |r_X(k)| \leq c'\,{\rho'}^k\ \forall k \geq 0$;

Cas particulier : Si (X_t) est un processus MA(q),

1. $\alpha_j = b_j$ pour $j = 1, \ldots, q$, 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 $0 \leq k \leq q$,

$$r_X(k) = \mathrm{cov}\Big(\sum_{j=0}^{q} b_j \varepsilon_{-j}, \sum_{\ell=0}^{q} b_\ell \varepsilon_{k-\ell}\Big) = \sum_{j=0}^{q}\sum_{\ell=0}^{q} b_j b_\ell \mathbb{E}[\varepsilon_{-j}\varepsilon_{k-\ell}] = \sigma_\varepsilon^2 \sum_{j=0}^{q-k} b_j b_{j+k}$$

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

$$X_t + a_1\, X_{t-1} + \cdots + a_p\, X_{t-p} = \varepsilon_t + b_1\, \varepsilon_{t-1} + \cdots + b_q\, \varepsilon_{t-q}$$

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. $r_X(k) = cov(X_0, X_k)$ *:* $\exists 0 \leq \rho' < 1,\ c' \geq 0,\ |r_X(k)| \leq c'\, {\rho'}^k\ \forall k \geq 0$ *;*

Cas particulier : Si (X_t) est un processus MA(q),

1. $\alpha_j = b_j$ pour $j = 1, \ldots, q$, 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 $0 \leq k \leq q$,

$$r_X(k) = \operatorname{cov}\Big(\sum_{j=0}^{q} b_j \varepsilon_{-j}, \sum_{\ell=0}^{q} b_\ell \varepsilon_{k-\ell}\Big) = \sum_{j=0}^{q}\sum_{\ell=0}^{q} b_j b_\ell \mathbb{E}[\varepsilon_{-j}\varepsilon_{k-\ell}] = \sigma_\varepsilon^2 \sum_{j=0}^{q-k} b_j b_{j+k}$$

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

$$X_t + a_1\,X_{t-1} + \cdots + a_p\,X_{t-p} = \varepsilon_t + b_1\,\varepsilon_{t-1} + \cdots + b_q\,\varepsilon_{t-q}$$

Alors :

1. $X_t = \sum_{k=0}^{\infty} \alpha_j\,\varepsilon_{t-j}$ *et* $\exists 0 \le \rho < 1,\ c \ge 0,\ |\alpha_j| \le c\,\rho^j$ *pour tout* $j \ge 0$ *;*
2. $r_X(k) = cov(X_0, X_k)$ *:* $\exists 0 \le \rho' < 1,\ c' \ge 0,\ |r_X(k)| \le c'\,{\rho'}^k\ \forall k \ge 0$ *;*

Cas particulier : Si (X_t) est un processus MA(q),

1. $\alpha_j = b_j$ pour $j = 1, \ldots, q$, et $\alpha_j = 0$ si $j \ge q+1$;
2. $r_X(k) = 0$ pour $k \ge q+1$ et avec $b_0 = 1$, pour $0 \le k \le q$,

$$r_X(k) = \operatorname{cov}\Big(\sum_{j=0}^{q} b_j\varepsilon_{-j}, \sum_{\ell=0}^{q} b_\ell\varepsilon_{k-\ell}\Big) = \sum_{j=0}^{q}\sum_{\ell=0}^{q} b_j b_\ell \mathbb{E}[\varepsilon_{-j}\varepsilon_{k-\ell}] = \sigma_\varepsilon^2 \sum_{j=0}^{q-k} b_j b_{j+k}$$

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

$$X_t + a_1\, X_{t-1} + \cdots + a_p\, X_{t-p} = \varepsilon_t + b_1\, \varepsilon_{t-1} + \cdots + b_q\, \varepsilon_{t-q}$$

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. $r_X(k) = cov(X_0, X_k)$ *:* $\exists 0 \leq \rho' < 1,\ c' \geq 0,\ |r_X(k)| \leq c'\, {\rho'}^k\ \forall k \geq 0$ *;*

Cas particulier : Si (X_t) est un processus MA(q),

1. $\alpha_j = b_j$ pour $j = 1, \ldots, q$, 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 $0 \leq k \leq q$,

$$r_X(k) = \operatorname{cov}\Big(\sum_{j=0}^{q} b_j \varepsilon_{-j}, \sum_{\ell=0}^{q} b_\ell \varepsilon_{k-\ell}\Big) = \sum_{j=0}^{q}\sum_{\ell=0}^{q} b_j b_\ell \mathbb{E}[\varepsilon_{-j}\varepsilon_{k-\ell}] = \sigma_\varepsilon^2 \sum_{j=0}^{q-k} b_j b_{j+k}$$

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

$$X_t + a_1\, X_{t-1} + \cdots + a_p\, X_{t-p} = \varepsilon_t + b_1\, \varepsilon_{t-1} + \cdots + b_q\, \varepsilon_{t-q}$$

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. $r_X(k) = cov(X_0, X_k)$ *:* $\exists 0 \leq \rho' < 1,\ c' \geq 0,\ |r_X(k)| \leq c'\, {\rho'}^k\ \forall k \geq 0$ *;*

Cas particulier : Si (X_t) est un processus MA(q),

1. $\alpha_j = b_j$ pour $j = 1, \ldots, q$, 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 $0 \leq k \leq q$,

$$r_X(k) = \operatorname{cov}\Big(\sum_{j=0}^{q} b_j \varepsilon_{-j}, \sum_{\ell=0}^{q} b_\ell \varepsilon_{k-\ell}\Big) = \sum_{j=0}^{q}\sum_{\ell=0}^{q} b_j b_\ell \mathbb{E}[\varepsilon_{-j}\varepsilon_{k-\ell}] = \sigma_\varepsilon^2 \sum_{j=0}^{q-k} b_j b_{j+k}$$

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

$$X_t + a_1\, X_{t-1} + \cdots + a_p\, X_{t-p} = \varepsilon_t + b_1\, \varepsilon_{t-1} + \cdots + b_q\, \varepsilon_{t-q}$$

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. $r_X(k) = cov(X_0, X_k)$: $\exists 0 \leq \rho' < 1,\ c' \geq 0,\ |r_X(k)| \leq c'\, {\rho'}^k\ \forall k \geq 0$;

Cas particulier : Si (X_t) est un processus MA(q),

1. $\alpha_j = b_j$ pour $j = 1, \ldots, q$, 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 $0 \leq k \leq q$,

$$r_X(k) = \mathrm{cov}\Big(\sum_{j=0}^{q} b_j \varepsilon_{-j}, \sum_{\ell=0}^{q} b_\ell \varepsilon_{k-\ell}\Big) = \sum_{j=0}^{q}\sum_{\ell=0}^{q} b_j b_\ell \mathbb{E}[\varepsilon_{-j}\varepsilon_{k-\ell}] = \sigma_\varepsilon^2 \sum_{j=0}^{q-k} b_j b_{j+k}$$

Propriétés des processus ARMA (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*

$$X'_t + a_1\, X'_{t-1} + \cdots + a_p\, X'_{t-p} = \varepsilon_t + b_1\, \varepsilon_{t-1} + \cdots + b_q\, \varepsilon_{t-q}$$

$$\Longrightarrow \quad \mathbb{E}(X_t) = \frac{\omega_0}{1+a_1+\cdots+a_p} \ : \quad ARMA(p,q) \text{ décentré}$$

Propriété

Un processus ARMA(p,q) causal de bruit blanc (ε_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*

$$X'_t + a_1\,X'_{t-1} + \cdots + a_p\,X'_{t-p} = \varepsilon_t + b_1\,\varepsilon_{t-1} + \cdots + b_q\,\varepsilon_{t-q}$$

$$\Longrightarrow \quad \mathbb{E}(X_t) = \frac{\omega_0}{1 + a_1 + \cdots + a_p} : \quad \text{ARMA}(p,q) \text{ décentré}$$

Propriété

Un processus ARMA(p,q) causal de bruit blanc (ε_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*

$$X'_t + a_1\,X'_{t-1} + \cdots + a_p\,X'_{t-p} = \varepsilon_t + b_1\,\varepsilon_{t-1} + \cdots + b_q\,\varepsilon_{t-q}$$

$$\Longrightarrow \quad \mathbb{E}(X_t) = \frac{\omega_0}{1 + a_1 + \cdots + a_p} : \quad \textit{ARMA}(p,q) \textit{ décentré}$$

Propriété

Un processus ARMA(p,q) causal de bruit blanc (ε_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*

$$X'_t + a_1\, X'_{t-1} + \cdots + a_p\, X'_{t-p} = \varepsilon_t + b_1\, \varepsilon_{t-1} + \cdots + b_q\, \varepsilon_{t-q}$$

$$\Longrightarrow \quad \mathbb{E}(X_t) = \frac{\omega_0}{1 + a_1 + \cdots + a_p} : \quad ARMA(p,q) \text{ décentré}$$

Propriété

Un processus ARMA(p,q) causal de bruit blanc (ε_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*

$$X'_t + a_1\,X'_{t-1} + \cdots + a_p\,X'_{t-p} = \varepsilon_t + b_1\,\varepsilon_{t-1} + \cdots + b_q\,\varepsilon_{t-q}$$

$$\Longrightarrow \quad \mathbb{E}(X_t) = \frac{\omega_0}{1 + a_1 + \cdots + a_p} : \quad \textit{ARMA}(p,q) \textit{ décentré}$$

Propriété

Un processus ARMA(p,q) causal de bruit blanc (ε_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 $\widehat{\rho}(\cdot)$ son autocorrélation empirique calculée sur $(X_1,\ldots,X_n)$. Alors :

$$\widehat{\rho}(k) \xrightarrow[n\to+\infty]{\mathcal{P}} \rho(k) \qquad \text{pour tout } k \in \mathbb{N}$$

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 $\widehat{\rho}(\cdot)$ son autocorrélation empirique calculée sur $(X_1,\dots,X_n)$. Alors :

$$\widehat{\rho}(k) \xrightarrow[n\to+\infty]{\mathcal{P}} \rho(k) \qquad \text{pour tout } k \in \mathbb{N}$$

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 $\widehat{\rho}(\cdot)$ son autocorrélation empirique calculée sur $(X_1,\ldots,X_n)$. Alors :

$$\widehat{\rho}(k) \xrightarrow[n\to+\infty]{\mathcal{P}} \rho(k) \qquad \textit{pour tout } k\in\mathbb{N}$$

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* $\widehat{\rho}(\cdot)$ *son autocorrélation empirique calculée sur* $(X_1,\ldots,X_n)$*. Alors :*

$$\widehat{\rho}(k) \xrightarrow[n\to+\infty]{\mathcal{P}} \rho(k) \qquad \textit{pour tout } k \in \mathbb{N}$$

Conséquence : l'ACF converge vers les autocorrélations théoriques

Plan du cours

Processus GARCH

Définition (Engle, 1982 et Bollersev, 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 :

- (ε_t) *bruit blanc de variance* 1 ;
- $\omega_0 > 0$, $(a_1, \ldots, a_p, b_1, \ldots, b_q) \in [0,\infty)^{p+q}$ *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 Bollersev, 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 $t \in \mathbb{Z}$:*

$X_t = \varepsilon_t \, \sigma_t \quad$ *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 :

- (ε_t) *bruit blanc de variance* 1 *;*
- $\omega_0 > 0$, $(a_1, \ldots, a_p, b_1, \ldots, b_q) \in [0, \infty)^{p+q}$ *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 Bollersev, 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 :

- (ε_t) *bruit blanc de variance* 1 ;
- $\omega_0 > 0$, $(a_1,\ldots,a_p,b_1,\ldots,b_q) \in [0,\infty)^{p+q}$ *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 Bollersev, 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 :

- (ε_t) *bruit blanc de variance* 1 *;*
- $\omega_0 > 0$, $(a_1,\ldots,a_p,b_1,\ldots,b_q) \in [0,\infty)^{p+q}$ *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 Bollersev, 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 :

- (ε_t) *bruit blanc de variance* 1 *;*
- $\omega_0 > 0$, $(a_1, \ldots, a_p, b_1, \ldots, b_q) \in [0, \infty)^{p+q}$ *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 Bollersev, 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 $t \in \mathbb{Z}$:*

$$X_t = \varepsilon_t \, \sigma_t \quad \text{et} \quad \sigma_t^2 = \omega_0 + a_1 \, X_{t-1}^2 + \cdots + a_p \, X_{t-p}^2 + b_1 \, \sigma_{t-1}^2 + \cdots + b_q \, \sigma_{t-q}^2$$

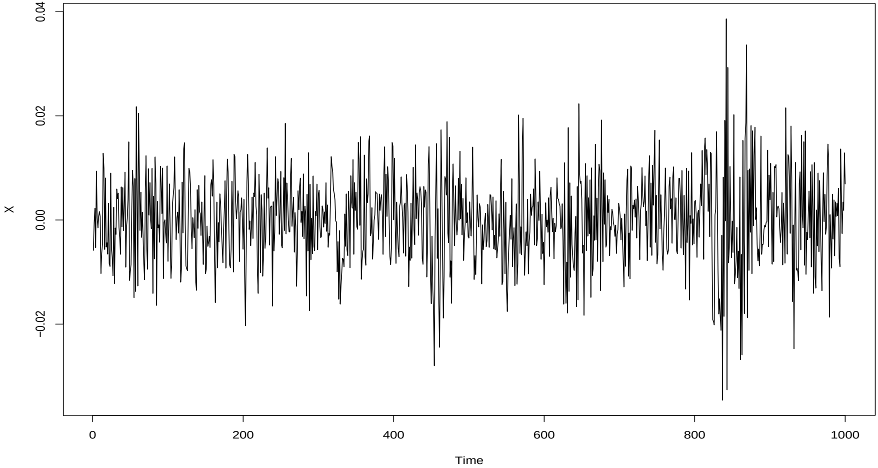
où :

- (ε_t) *bruit blanc de variance* 1 ;
- $\omega_0 > 0$, $(a_1, \ldots, a_p, b_1, \ldots, b_q) \in [0, \infty)^{p+q}$ *avec* $a_p \neq 0$ *et* $b_q \neq 0$.

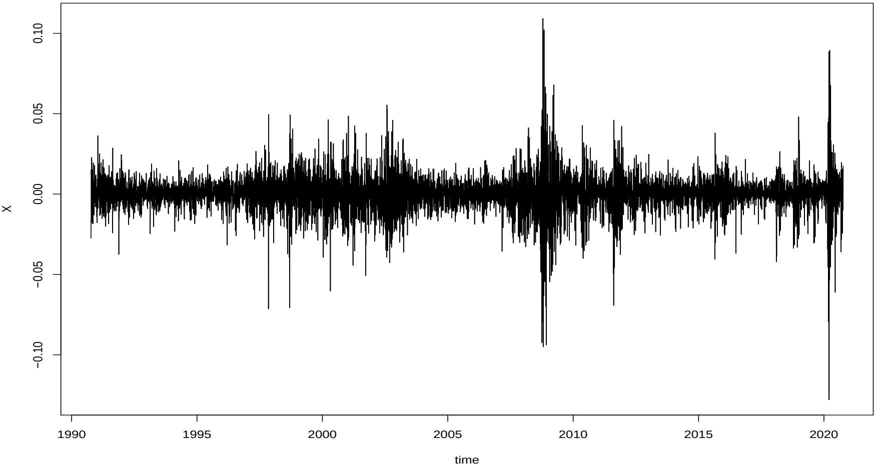
Cas particulier :

- Si $q = 0$, un GARCH(p,q) est un ARCH(p) (*Conditionnaly Heteroskedastic Auto-Regressive*) ;

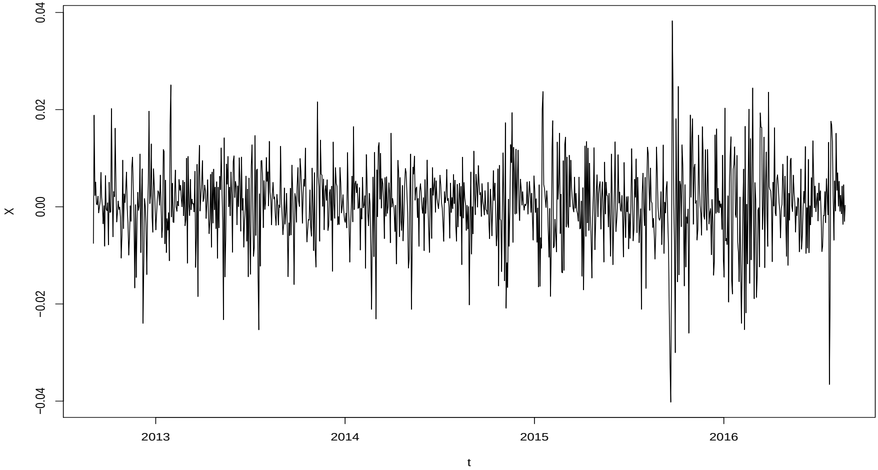
Trajectoire d'un processus GARCH$(1,1)$



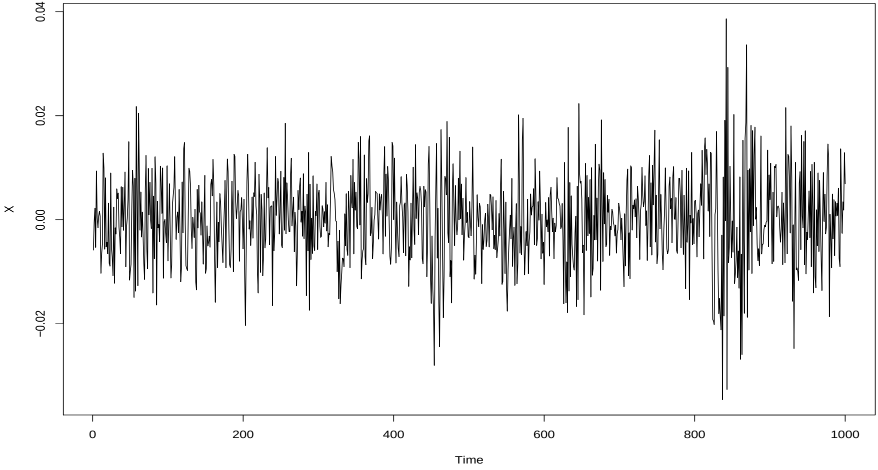
Trajectoire Log-rendement SP500 d'octobre 1990 à octobre 2020



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. $\mathbb{E}[X_t] = 0$, $var(X_t) = \frac{\sigma_\varepsilon^2}{1-\sum_{i=1}^{p} a_i - \sum_{j=1}^{q} b_j}$ *et* $cov(X_0, X_k) = 0$ *si* $k \neq 0$.
2. $\mathbb{E}[X_t \mid X_{t-1}, X_{t-2}, \ldots] = 0$ *et* $var(X_t \mid X_{t-1}, X_{t-2}, \ldots) = \sigma_t^2 \neq C^{te}$: *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. $\mathbb{E}[X_t] = 0$, $var(X_t) = \frac{\sigma_\varepsilon^2}{1-\sum_{i=1}^{p} a_i - \sum_{j=1}^{q} b_j}$ *et* $cov(X_0, X_k) = 0$ *si* $k \neq 0$.
2. $\mathbb{E}[X_t \,|\, X_{t-1}, X_{t-2}, \ldots] = 0$ *et* $var(X_t \,|\, X_{t-1}, X_{t-2}, \ldots) = \sigma_t^2 \neq C^{te}$: *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. $\mathbb{E}[X_t] = 0$, $var(X_t) = \frac{\sigma_\varepsilon^2}{1-\sum_{i=1}^{p} a_i - \sum_{j=1}^{q} b_j}$ *et* $cov(X_0, X_k) = 0$ *si* $k \neq 0$.
2. $\mathbb{E}[X_t \mid X_{t-1}, X_{t-2}, \ldots] = 0$ *et* $var(X_t \mid X_{t-1}, X_{t-2}, \ldots) = \sigma_t^2 \neq C^{te}$: *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. $\mathbb{E}[X_t] = 0$, $var(X_t) = \frac{\sigma_\varepsilon^2}{1-\sum_{i=1}^{p} a_i - \sum_{j=1}^{q} b_j}$ *et* $cov(X_0, X_k) = 0$ *si* $k \neq 0$.
2. $\mathbb{E}[X_t \mid X_{t-1}, X_{t-2}, \ldots] = 0$ *et* $var(X_t \mid X_{t-1}, X_{t-2}, \ldots) = \sigma_t^2 \neq C^{te}$ *: 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. $\mathbb{E}[X_t] = 0$, $var(X_t) = \frac{\sigma_\varepsilon^2}{1-\sum_{i=1}^p a_i - \sum_{j=1}^q b_j}$ *et* $cov(X_0, X_k) = 0$ *si* $k \neq 0$.
2. $\mathbb{E}[X_t \mid X_{t-1}, X_{t-2}, \ldots] = 0$ *et* $var(X_t \mid X_{t-1}, X_{t-2}, \ldots) = \sigma_t^2 \neq C^{te}$:
   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. $\mathbb{E}[X_t] = 0$, $var(X_t) = \frac{\sigma_\varepsilon^2}{1-\sum_{i=1}^{p} a_i - \sum_{j=1}^{q} b_j}$ *et* $cov(X_0, X_k) = 0$ *si* $k \neq 0$.
2. $\mathbb{E}[X_t \mid X_{t-1}, X_{t-2}, \dots] = 0$ *et* $var(X_t \mid X_{t-1}, X_{t-2}, \dots) = \sigma_t^2 \neq C^{te}$: *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 $t > 0$, $\mathrm{cov}(v_0, v_t) = \mathrm{cov}((\varepsilon_0^2 - 1)\sigma_0^2(\varepsilon_t^2 - 1)\sigma_t^2) = 0$ 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 $t > 0$, $\mathrm{cov}(v_0, v_t) = \mathrm{cov}((\varepsilon_0^2 - 1)\sigma_0^2(\varepsilon_t^2 - 1)\sigma_t^2) = 0$ 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. □

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 $t > 0$, $\mathrm{cov}(v_0, v_t) = \mathrm{cov}((\varepsilon_0^2 - 1)\sigma_0^2(\varepsilon_t^2 - 1)\sigma_t^2) = 0$ 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. □

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 $t > 0$, $\mathrm{cov}(v_0, v_t) = \mathrm{cov}((\varepsilon_0^2 - 1)\sigma_0^2(\varepsilon_t^2 - 1)\sigma_t^2) = 0$ 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

Plan du cours

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

$$X_t + a_1\, X_{t-1} + \cdots + a_p\, X_{t-p} = \varepsilon_t + b_1\, \varepsilon_{t-1} + \cdots + b_q\, \varepsilon_{t-q}$$

où $\quad \varepsilon_t = \xi_t\, \sigma_t \quad$ *et* $\quad \sigma_t^2 = \omega_0 + c_1\, \varepsilon_{t-1}^2 + \cdots + c_{p'}\, \varepsilon_{t-p'}^2 + d_1\, \sigma_{t-1}^2 + \cdots + d_{q'}\, \sigma_{t-q'}^2$

avec (ξ_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 $P(x) = 1 + a_1 x + \ldots + a_p x^p$ 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

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

$$X_t + a_1 X_{t-1} + \cdots + a_p X_{t-p} = \varepsilon_t + b_1 \varepsilon_{t-1} + \cdots + b_q \varepsilon_{t-q}$$

où $\varepsilon_t = \xi_t \sigma_t$ *et* $\sigma_t^2 = \omega_0 + c_1 \varepsilon_{t-1}^2 + \cdots + c_{p'} \varepsilon_{t-p'}^2 + d_1 \sigma_{t-1}^2 + \cdots + d_{q'} \sigma_{t-q'}^2$

avec (ξ_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 $P(x) = 1 + a_1 x + \ldots + a_p x^p$ 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

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

$$X_t + a_1\, X_{t-1} + \cdots + a_p\, X_{t-p} = \varepsilon_t + b_1\, \varepsilon_{t-1} + \cdots + b_q\, \varepsilon_{t-q}$$

où $\quad \varepsilon_t = \xi_t\, \sigma_t \quad$ *et* $\quad \sigma_t^2 = \omega_0 + c_1\, \varepsilon_{t-1}^2 + \cdots + c_{p'}\, \varepsilon_{t-p'}^2 + d_1\, \sigma_{t-1}^2 + \cdots + d_{q'}\, \sigma_{t-q'}^2$

avec (ξ_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 $P(x) = 1 + a_1 x + \ldots + a_p x^p$ 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

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

$$X_t + a_1\,X_{t-1} + \cdots + a_p\,X_{t-p} = \varepsilon_t + b_1\,\varepsilon_{t-1} + \cdots + b_q\,\varepsilon_{t-q}$$

où $\varepsilon_t = \xi_t\,\sigma_t$ *et* $\sigma_t^2 = \omega_0 + c_1\,\varepsilon_{t-1}^2 + \cdots + c_{p'}\,\varepsilon_{t-p'}^2 + d_1\,\sigma_{t-1}^2 + \cdots + d_{q'}\,\sigma_{t-q'}^2$

avec (ξ_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 $P(x) = 1 + a_1 x + \ldots + a_p x^p$ 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

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*

$$X_t + a_1\, X_{t-1} + \cdots + a_p\, X_{t-p} = \varepsilon_t + b_1\, \varepsilon_{t-1} + \cdots + b_q\, \varepsilon_{t-q}$$

où $\varepsilon_t = \xi_t\, \sigma_t$ *et* $\sigma_t^2 = \omega_0 + c_1\, \varepsilon_{t-1}^2 + \cdots + c_{p'}\, \varepsilon_{t-p'}^2 + d_1\, \sigma_{t-1}^2 + \cdots + d_{q'}\, \sigma_{t-q'}^2$

avec (ξ_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 $P(x) = 1 + a_1 x + \ldots + a_p x^p$ 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

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

$$X_t + a_1 X_{t-1} + \cdots + a_p X_{t-p} = \varepsilon_t + b_1 \varepsilon_{t-1} + \cdots + b_q \varepsilon_{t-q}$$

où $\varepsilon_t = \xi_t \sigma_t$ *et* $\sigma_t^2 = \omega_0 + c_1 \varepsilon_{t-1}^2 + \cdots + c_{p'} \varepsilon_{t-p'}^2 + d_1 \sigma_{t-1}^2 + \cdots + d_{q'} \sigma_{t-q'}^2$

avec (ξ_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 $P(x) = 1 + a_1 x + \ldots + a_p x^p$ 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

Processus APARCH [Ding, Granger et Engle, 1993]

Définition

$(X_k)_{k\in\mathbb{Z}}$ *est un processus APARCH*(δ, 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 :

- (ε_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*(δ, 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 :

- (ε_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*(δ, 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 :

- (ε_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(δ, 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 :

- *(ε_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(δ,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 :

- *(ε_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*(δ,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 :

- (ε_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*(δ, 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 :

- (ε_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(δ, 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 :

- *(ε_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

$$X_t + a_1\, X_{t-1} + \cdots + a_p\, X_{t-p} = \varepsilon_t + b_1\, \varepsilon_{t-1} + \cdots + b_q\, \varepsilon_{t-q}$$

$$\text{où} \quad \varepsilon_t = \xi_t + \sum_{j=1}^{\infty} \Big(\frac{\Gamma(j+d)}{\Gamma(d)\Gamma(j+1)} \Big)\, \xi_{t-j}$$

avec (ξ_t) bruit blanc de variance σ_ξ^2, a_p, b_q non nuls, $\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt$.

Commentaire : • Avec $P(x) = 1 + a_1 x + \ldots + a_p x^p$, $P(z) \neq 0$ pour $|z| \leq 1$

• Si $0 < d < 0.5$, $r_X(k) \simeq C\,|k|^{2d-1}$ pour $k \to \infty$: processus longue mémoire

• On a $\widehat{\rho}(\cdot) \xrightarrow[n\to+\infty]{\mathcal{P}} \rho(\cdot)$: ACF estime ρ

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*

$$X_t + a_1\,X_{t-1} + \cdots + a_p\,X_{t-p} = \varepsilon_t + b_1\,\varepsilon_{t-1} + \cdots + b_q\,\varepsilon_{t-q}$$

$$\text{où} \quad \varepsilon_t = \xi_t + \sum_{j=1}^{\infty}\left(\frac{\Gamma(j+d)}{\Gamma(d)\Gamma(j+1)}\right)\xi_{t-j}$$

avec (ξ_t) *bruit blanc de variance* σ_ξ^2*,* a_p*,* b_q *non nuls,* $\Gamma(x) = \int_0^\infty t^{x-1}e^{-t}dt$*.*

Commentaire : • Avec $P(x) = 1 + a_1 x + \ldots + a_p x^p$, $P(z) \neq 0$ pour $|z| \leq 1$

• Si $0 < d < 0.5$, $r_X(k) \simeq C\,|k|^{2d-1}$ pour $k \to \infty$: processus longue mémoire

• On a $\widehat{\rho}(\cdot) \xrightarrow[n\to+\infty]{\mathcal{P}} \rho(\cdot)$: ACF estime ρ

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

$$X_t + a_1\, X_{t-1} + \cdots + a_p\, X_{t-p} = \varepsilon_t + b_1\, \varepsilon_{t-1} + \cdots + b_q\, \varepsilon_{t-q}$$

$$\text{où} \quad \varepsilon_t = \xi_t + \sum_{j=1}^{\infty} \left(\frac{\Gamma(j+d)}{\Gamma(d)\Gamma(j+1)} \right) \xi_{t-j}$$

avec (ξ_t) bruit blanc de variance σ_ξ^2, a_p, b_q non nuls, $\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt$.

Commentaire : • Avec $P(x) = 1 + a_1 x + \ldots + a_p x^p$, $P(z) \neq 0$ pour $|z| \leq 1$

• Si $0 < d < 0.5$, $r_X(k) \simeq C\,|k|^{2d-1}$ pour $k \to \infty$: processus longue mémoire

• On a $\widehat{\rho}(\cdot) \xrightarrow[n\to+\infty]{\mathcal{P}} \rho(\cdot)$: ACF estime ρ

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

$$X_t + a_1\, X_{t-1} + \cdots + a_p\, X_{t-p} = \varepsilon_t + b_1\, \varepsilon_{t-1} + \cdots + b_q\, \varepsilon_{t-q}$$

$$\text{où} \quad \varepsilon_t = \xi_t + \sum_{j=1}^{\infty} \Big(\frac{\Gamma(j+d)}{\Gamma(d)\Gamma(j+1)} \Big)\, \xi_{t-j}$$

avec (ξ_t) bruit blanc de variance σ_ξ^2, a_p, b_q non nuls, $\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt$.

Commentaire : • Avec $P(x) = 1 + a_1 x + \ldots + a_p x^p$, $P(z) \neq 0$ pour $|z| \leq 1$

• Si $0 < d < 0.5$, $r_X(k) \simeq C\, |k|^{2d-1}$ pour $k \to \infty$: processus longue mémoire

• On a $\widehat{\rho}(\cdot) \xrightarrow[n\to+\infty]{\mathcal{P}} \rho(\cdot)$: ACF estime ρ

Processus FARIMA [Granger et Joyeux, 1980]

Définition

$(X_k)_{k\in\mathbb{Z}}$ est un processus FARIMA(p,d,q) où $-0.5 \le d \le 0.5$ et

$$X_t + a_1\, X_{t-1} + \cdots + a_p\, X_{t-p} = \varepsilon_t + b_1\, \varepsilon_{t-1} + \cdots + b_q\, \varepsilon_{t-q}$$

$$\text{où} \quad \varepsilon_t = \xi_t + \sum_{j=1}^{\infty} \left(\frac{\Gamma(j+d)}{\Gamma(d)\Gamma(j+1)} \right) \xi_{t-j}$$

avec (ξ_t) bruit blanc de variance σ_ξ^2, a_p, b_q non nuls, $\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt$.

Commentaire : • Avec $P(x) = 1 + a_1 x + \ldots + a_p x^p$, $P(z) \neq 0$ pour $|z| \le 1$

• Si $0 < d < 0.5$, $r_X(k) \simeq C\,|k|^{2d-1}$ pour $k \to \infty$: processus longue mémoire

• On a $\widehat{\rho}(\cdot) \xrightarrow[n\to+\infty]{\mathcal{P}} \rho(\cdot)$: ACF estime ρ

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

$$X_t + a_1\, X_{t-1} + \cdots + a_p\, X_{t-p} = \varepsilon_t + b_1\, \varepsilon_{t-1} + \cdots + b_q\, \varepsilon_{t-q}$$

$$\text{où} \quad \varepsilon_t = \xi_t + \sum_{j=1}^{\infty} \Big(\frac{\Gamma(j+d)}{\Gamma(d)\Gamma(j+1)}\Big)\, \xi_{t-j}$$

avec (ξ_t) bruit blanc de variance σ_ξ^2, a_p, b_q non nuls, $\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt$.

Commentaire : • Avec $P(x) = 1 + a_1 x + \ldots + a_p x^p$, $P(z) \neq 0$ pour $|z| \leq 1$

• Si $0 < d < 0.5$, $r_X(k) \simeq C\,|k|^{2d-1}$ pour $k \to \infty$: processus longue mémoire

• On a $\widehat{\rho}(\cdot) \xrightarrow[n\to+\infty]{\mathcal{P}} \rho(\cdot)$: ACF estime ρ

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

$$X_t + a_1\,X_{t-1} + \cdots + a_p\,X_{t-p} = \varepsilon_t + b_1\,\varepsilon_{t-1} + \cdots + b_q\,\varepsilon_{t-q}$$

$$\text{où} \quad \varepsilon_t = \xi_t + \sum_{j=1}^{\infty}\left(\frac{\Gamma(j+d)}{\Gamma(d)\Gamma(j+1)}\right)\xi_{t-j}$$

avec (ξ_t) bruit blanc de variance σ_ξ^2, a_p, b_q non nuls, $\Gamma(x) = \int_0^\infty t^{x-1}e^{-t}dt$.

Commentaire : • Avec $P(x) = 1 + a_1 x + \ldots + a_p x^p$, $P(z) \neq 0$ pour $|z| \leq 1$

• Si $0 < d < 0.5$, $r_X(k) \simeq C\,|k|^{2d-1}$ pour $k \to \infty$: processus longue mémoire

• On a $\widehat{\rho}(\cdot) \xrightarrow[n\to+\infty]{\mathcal{P}} \rho(\cdot)$: ACF estime ρ

Processus FARIMA [Granger et Joyeux, 1980]

Définition

$(X_k)_{k\in\mathbb{Z}}$ est un processus FARIMA(p,d,q) où $-0.5 \le d \le 0.5$ et

$$X_t + a_1 X_{t-1} + \cdots + a_p X_{t-p} = \varepsilon_t + b_1 \varepsilon_{t-1} + \cdots + b_q \varepsilon_{t-q}$$

$$\text{où} \quad \varepsilon_t = \xi_t + \sum_{j=1}^{\infty} \Big(\frac{\Gamma(j+d)}{\Gamma(d)\Gamma(j+1)} \Big) \xi_{t-j}$$

avec (ξ_t) bruit blanc de variance σ_ξ^2, a_p, b_q non nuls, $\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt$.

Commentaire : • Avec $P(x) = 1 + a_1 x + \ldots + a_p x^p$, $P(z) \neq 0$ pour $|z| \le 1$

• Si $0 < d < 0.5$, $r_X(k) \simeq C\,|k|^{2d-1}$ pour $k \to \infty$: processus longue mémoire

• On a $\widehat{\rho}(\cdot) \xrightarrow[n\to+\infty]{\mathcal{P}} \rho(\cdot)$: ACF estime ρ

Plan du cours

Nouvelle écriture des processus

Tous les processus vus dans la section précédente peuvent s'écrire :

$$X_t = M_\theta\big(X_{t-1}, X_{t-2}, \dots\big)\,\xi_t + f_\theta\big(X_{t-1}, X_{t-2}, \dots\big)$$

(ξ_t) bruit blanc de variance 1, θ vecteur de paramètres, M_θ et f_θ fonctions.

Remarque : pour les ARMA, ARMA-GARCH, FARIMA, il faut supposer que $Q(z) \neq 0$ pour $|z| \leq 1$ $Q(x) = 1 + b_1 x + \ldots + b_q x^q$

Exemples : • Si (X_t) AR(p), $X_t = \xi_t - a_1\, X_{t-1} - \cdots - a_p\, X_{t-p}$

$$\Longrightarrow M_\theta\big(X_{t-1}, \dots\big) = \sigma_\xi, \quad f_\theta\big(X_{t-1}, \dots\big) = -a_1\, X_{t-1} - \cdots - a_p\, X_{t-p}$$

• Si (X_t) ARCH(p), $X_t = \xi_t\,\sqrt{\omega_0 + a_1\, X_{t-1}^2 + \cdots + a_p\, X_{t-p}^2}$

$$\Longrightarrow M_\theta\big(X_{t-1}, \dots\big) = \sqrt{\omega_0 + a_1 X_{t-1}^2 + \cdots + a_p X_{t-p}^2}, \quad f_\theta\big(X_{t-1}, \dots\big) = 0$$

Nouvelle écriture des processus

Tous les processus vus dans la section précédente peuvent s'écrire :

$$X_t = M_\theta\big(X_{t-1}, X_{t-2}, \ldots\big)\,\xi_t + f_\theta\big(X_{t-1}, X_{t-2}, \ldots\big)$$

(ξ_t) bruit blanc de variance 1, θ vecteur de paramètres, M_θ et f_θ fonctions.

Remarque : pour les ARMA, ARMA-GARCH, FARIMA, il faut supposer que $Q(z) \neq 0$ pour $|z| \leq 1$ $Q(x) = 1 + b_1 x + \ldots + b_q x^q$

Exemples : • Si (X_t) AR(p), $X_t = \xi_t - a_1\,X_{t-1} - \cdots - a_p\,X_{t-p}$

$\Longrightarrow M_\theta\big(X_{t-1}, \ldots\big) = \sigma_\xi, \;\; f_\theta\big(X_{t-1}, \ldots\big) = -a_1\,X_{t-1} - \cdots - a_p\,X_{t-p}$

• Si (X_t) ARCH(p), $X_t = \xi_t\,\sqrt{\omega_0 + a_1\,X_{t-1}^2 + \cdots + a_p\,X_{t-p}^2}$

$\Longrightarrow M_\theta\big(X_{t-1}, \ldots\big) = \sqrt{\omega_0 + a_1 X_{t-1}^2 + \cdots + a_p X_{t-p}^2}, \;\; f_\theta\big(X_{t-1}, \ldots\big) = 0$

Nouvelle écriture des processus

Tous les processus vus dans la section précédente peuvent s'écrire :

$$X_t = M_\theta\big(X_{t-1}, X_{t-2}, \dots\big)\,\xi_t + f_\theta\big(X_{t-1}, X_{t-2}, \dots\big)$$

(ξ_t) bruit blanc de variance 1, θ vecteur de paramètres, M_θ et f_θ fonctions.

Remarque : pour les ARMA, ARMA-GARCH, FARIMA, il faut supposer que $Q(z) \neq 0$ pour $|z| \leq 1$ $Q(x) = 1 + b_1 x + \ldots + b_q x^q$

Exemples : • Si (X_t) AR(p), $X_t = \xi_t - a_1\,X_{t-1} - \cdots - a_p\,X_{t-p}$

$\Longrightarrow M_\theta\big(X_{t-1}, \dots\big) = \sigma_\xi, \;\; f_\theta\big(X_{t-1}, \dots\big) = -a_1\,X_{t-1} - \cdots - a_p\,X_{t-p}$

• Si (X_t) ARCH(p), $X_t = \xi_t\,\sqrt{\omega_0 + a_1\,X_{t-1}^2 + \cdots + a_p\,X_{t-p}^2}$

$\Longrightarrow M_\theta\big(X_{t-1}, \dots\big) = \sqrt{\omega_0 + a_1 X_{t-1}^2 + \cdots + a_p X_{t-p}^2}, \;\; f_\theta\big(X_{t-1}, \dots\big) = 0$

Nouvelle écriture des processus

Tous les processus vus dans la section précédente peuvent s'écrire :

$$X_t = M_\theta\big(X_{t-1}, X_{t-2}, \ldots\big)\, \xi_t + f_\theta\big(X_{t-1}, X_{t-2}, \ldots\big)$$

(ξ_t) bruit blanc de variance 1, θ vecteur de paramètres, M_θ et f_θ fonctions.

Remarque : pour les ARMA, ARMA-GARCH, FARIMA, il faut supposer que $Q(z) \neq 0$ pour $|z| \leq 1$ $Q(x) = 1 + b_1 x + \ldots + b_q x^q$

Exemples : • Si (X_t) AR(p), $X_t = \xi_t - a_1\, X_{t-1} - \cdots - a_p\, X_{t-p}$
$\Longrightarrow M_\theta\big(X_{t-1}, \ldots\big) = \sigma_\xi, \;\; f_\theta\big(X_{t-1}, \ldots\big) = -a_1\, X_{t-1} - \cdots - a_p\, X_{t-p}$

• Si (X_t) ARCH(p), $X_t = \xi_t \sqrt{\omega_0 + a_1\, X_{t-1}^2 + \cdots + a_p\, X_{t-p}^2}$
$\Longrightarrow M_\theta\big(X_{t-1}, \ldots\big) = \sqrt{\omega_0 + a_1 X_{t-1}^2 + \cdots + a_p X_{t-p}^2}, \;\; f_\theta\big(X_{t-1}, \ldots\big) = 0$

Nouvelle écriture des processus

Tous les processus vus dans la section précédente peuvent s'écrire :

$$X_t = M_\theta\big(X_{t-1}, X_{t-2}, \dots\big)\, \xi_t + f_\theta\big(X_{t-1}, X_{t-2}, \dots\big)$$

(ξ_t) bruit blanc de variance 1, θ vecteur de paramètres, M_θ et f_θ fonctions.

Remarque : pour les ARMA, ARMA-GARCH, FARIMA, il faut supposer que $Q(z) \neq 0$ pour $|z| \leq 1$ $Q(x) = 1 + b_1 x + \dots + b_q x^q$

Exemples : • Si (X_t) AR(p), $X_t = \xi_t - a_1\, X_{t-1} - \cdots - a_p\, X_{t-p}$

$\Longrightarrow M_\theta\big(X_{t-1}, \dots\big) = \sigma_\xi,\ \ f_\theta\big(X_{t-1}, \dots\big) = -a_1\, X_{t-1} - \cdots - a_p\, X_{t-p}$

• Si (X_t) ARCH(p), $X_t = \xi_t \sqrt{\omega_0 + a_1\, X_{t-1}^2 + \cdots + a_p\, X_{t-p}^2}$

$\Longrightarrow M_\theta\big(X_{t-1}, \dots\big) = \sqrt{\omega_0 + a_1 X_{t-1}^2 + \cdots + a_p X_{t-p}^2},\ \ f_\theta\big(X_{t-1}, \dots\big) = 0$

Nouvelle écriture des processus

Tous les processus vus dans la section précédente peuvent s'écrire :

$$X_t = M_\theta\big(X_{t-1}, X_{t-2}, \ldots\big)\, \xi_t + f_\theta\big(X_{t-1}, X_{t-2}, \ldots\big)$$

(ξ_t) bruit blanc de variance 1, θ vecteur de paramètres, M_θ et f_θ fonctions.

Remarque : pour les ARMA, ARMA-GARCH, FARIMA, il faut supposer que $Q(z) \neq 0$ pour $|z| \leq 1$ $Q(x) = 1 + b_1 x + \ldots + b_q x^q$

Exemples : • Si (X_t) AR(p), $X_t = \xi_t - a_1 X_{t-1} - \cdots - a_p X_{t-p}$

$\Longrightarrow M_\theta\big(X_{t-1}, \ldots\big) = \sigma_\xi, \;\; f_\theta\big(X_{t-1}, \ldots\big) = -a_1 X_{t-1} - \cdots - a_p X_{t-p}$

• Si (X_t) ARCH(p), $X_t = \xi_t \sqrt{\omega_0 + a_1 X_{t-1}^2 + \cdots + a_p X_{t-p}^2}$

$\Longrightarrow M_\theta\big(X_{t-1}, \ldots\big) = \sqrt{\omega_0 + a_1 X_{t-1}^2 + \cdots + a_p X_{t-p}^2}, \;\; f_\theta\big(X_{t-1}, \ldots\big) = 0$

Nouvelle écriture des processus

Tous les processus vus dans la section précédente peuvent s'écrire :

$$X_t = M_\theta\big(X_{t-1}, X_{t-2}, \dots\big)\, \xi_t + f_\theta\big(X_{t-1}, X_{t-2}, \dots\big)$$

(ξ_t) bruit blanc de variance 1, θ vecteur de paramètres, M_θ et f_θ fonctions.

Remarque : pour les ARMA, ARMA-GARCH, FARIMA, il faut supposer que $Q(z) \neq 0$ pour $|z| \leq 1$ $Q(x) = 1 + b_1 x + \ldots + b_q x^q$

Exemples : • Si (X_t) AR(p), $X_t = \xi_t - a_1\, X_{t-1} - \cdots - a_p\, X_{t-p}$
$\Longrightarrow M_\theta\big(X_{t-1}, \dots\big) = \sigma_\xi, \;\; f_\theta\big(X_{t-1}, \dots\big) = -a_1\, X_{t-1} - \cdots - a_p\, X_{t-p}$

• Si (X_t) ARCH(p), $X_t = \xi_t \sqrt{\omega_0 + a_1\, X_{t-1}^2 + \cdots + a_p\, X_{t-p}^2}$

$\Longrightarrow M_\theta\big(X_{t-1}, \dots\big) = \sqrt{\omega_0 + a_1 X_{t-1}^2 + \cdots + a_p X_{t-p}^2}, \;\; f_\theta\big(X_{t-1}, \dots\big) = 0$

Nouvelle écriture des processus

Tous les processus vus dans la section précédente peuvent s'écrire :

$$X_t = M_\theta\big(X_{t-1}, X_{t-2}, \dots\big)\, \xi_t + f_\theta\big(X_{t-1}, X_{t-2}, \dots\big)$$

(ξ_t) bruit blanc de variance 1, θ vecteur de paramètres, M_θ et f_θ fonctions.

Remarque : pour les ARMA, ARMA-GARCH, FARIMA, il faut supposer que $Q(z) \neq 0$ pour $|z| \leq 1$ $Q(x) = 1 + b_1 x + \ldots + b_q x^q$

Exemples : • Si (X_t) AR(p), $X_t = \xi_t - a_1 X_{t-1} - \cdots - a_p X_{t-p}$
$\Longrightarrow M_\theta\big(X_{t-1}, \dots\big) = \sigma_\xi, \quad f_\theta\big(X_{t-1}, \dots\big) = -a_1 X_{t-1} - \cdots - a_p X_{t-p}$

• Si (X_t) ARCH(p), $X_t = \xi_t \sqrt{\omega_0 + a_1 X_{t-1}^2 + \cdots + a_p X_{t-p}^2}$
$\Longrightarrow M_\theta\big(X_{t-1}, \dots\big) = \sqrt{\omega_0 + a_1 X_{t-1}^2 + \cdots + a_p X_{t-p}^2}, \quad f_\theta\big(X_{t-1}, \dots\big) = 0$

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
$\xi_t = \sum_{k=0}^{\infty} (-b_1)^k X_{t-k} \quad \Longrightarrow X_t = \xi_t + b_1 \sum_{k=0}^{\infty} (-b_1)^k X_{t-1-k}$

$$\Longrightarrow M_\theta(X_{t-1}, \ldots) = \sigma_\xi, \;\; f_\theta(X_{t-1}, \ldots) = b_1 \sum_{k=0}^{\infty} (-b_1)^k X_{t-1-k}$$

• Par extension si (X_t) ARMA(p,q), $X_t = \xi_t + \sum_{k=1}^{\infty} u_k X_{t-k}$ quand $Q(x) = 1 + b_1 x + \cdots + b_q x^q$ a ses racines en dehors du disque unité

$$\Longrightarrow M_\theta(X_{t-1}, \ldots) = \sigma_\xi, \;\; f_\theta(X_{t-1}, \ldots) = \sum_{k=1}^{\infty} u_k X_{t-k}$$

• Si (X_t) GARCH(p,q), on montre que $X_t = \xi_t \left(\omega_0 + \sum_{i=1}^{\infty} u_i X_{t-i}^2\right)^{1/2}$

$$\Longrightarrow M_\theta(X_{t-1}, \ldots) = \left(\omega_0 + \sum_{i=1}^{\infty} u_i X_{t-i}^2\right)^{1/2}, \;\; f_\theta(X_{t-1}, \ldots) = 0$$

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 $\xi_t = \sum_{k=0}^{\infty} (-b_1)^k X_{t-k}$ $\Longrightarrow X_t = \xi_t + b_1 \sum_{k=0}^{\infty} (-b_1)^k X_{t-1-k}$

$$\Longrightarrow M_\theta(X_{t-1}, \dots) = \sigma_\xi, \;\; f_\theta(X_{t-1}, \dots) = b_1 \sum_{k=0}^{\infty} (-b_1)^k X_{t-1-k}$$

• Par extension si (X_t) ARMA(p,q), $X_t = \xi_t + \sum_{k=1}^{\infty} u_k X_{t-k}$ quand $Q(x) = 1 + b_1 x + \cdots + b_q x^q$ a ses racines en dehors du disque unité

$$\Longrightarrow M_\theta(X_{t-1}, \dots) = \sigma_\xi, \;\; f_\theta(X_{t-1}, \dots) = \sum_{k=1}^{\infty} u_k X_{t-k}$$

• Si (X_t) GARCH(p,q), on montre que $X_t = \xi_t \left(\omega_0 + \sum_{i=1}^{\infty} u_i X_{t-i}^2 \right)^{1/2}$

$$\Longrightarrow M_\theta(X_{t-1}, \dots) = \left(\omega_0 + \sum_{i=1}^{\infty} u_i X_{t-i}^2 \right)^{1/2}, \;\; f_\theta(X_{t-1}, \dots) = 0$$

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 $\xi_t = \sum_{k=0}^{\infty} (-b_1)^k X_{t-k} \quad \Longrightarrow X_t = \xi_t + b_1 \sum_{k=0}^{\infty} (-b_1)^k X_{t-1-k}$

$$\Longrightarrow M_\theta(X_{t-1}, \dots) = \sigma_\xi, \ \ f_\theta(X_{t-1}, \dots) = b_1 \sum_{k=0}^{\infty} (-b_1)^k X_{t-1-k}$$

• 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

$$\Longrightarrow M_\theta(X_{t-1}, \dots) = \sigma_\xi, \ \ f_\theta(X_{t-1}, \dots) = \sum_{k=1}^{\infty} u_k X_{t-k}$$

• Si (X_t) GARCH(p,q), on montre que $X_t = \xi_t \left(\omega_0 + \sum_{i=1}^{\infty} u_i X_{t-i}^2 \right)^{1/2}$

$$\Longrightarrow M_\theta(X_{t-1}, \dots) = \left(\omega_0 + \sum_{i=1}^{\infty} u_i X_{t-i}^2 \right)^{1/2}, \ \ f_\theta(X_{t-1}, \dots) = 0$$

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
$\xi_t = \sum_{k=0}^{\infty} (-b_1)^k X_{t-k} \quad \Longrightarrow X_t = \xi_t + b_1 \sum_{k=0}^{\infty} (-b_1)^k X_{t-1-k}$

$$\Longrightarrow M_\theta\big(X_{t-1}, \dots\big) = \sigma_\xi, \;\; f_\theta\big(X_{t-1}, \dots\big) = b_1 \sum_{k=0}^{\infty} (-b_1)^k X_{t-1-k}$$

• Par extension si (X_t) ARMA(p, q), $X_t = \xi_t + \sum_{k=1}^{\infty} u_k X_{t-k}$ quand $Q(x) = 1 + b_1 x + \cdots + b_q x^q$ a ses racines en dehors du disque unité

$$\Longrightarrow M_\theta\big(X_{t-1}, \dots\big) = \sigma_\xi, \;\; f_\theta\big(X_{t-1}, \dots\big) = \sum_{k=1}^{\infty} u_k X_{t-k}$$

• Si (X_t) GARCH(p, q), on montre que $X_t = \xi_t \Big(\omega_0 + \sum_{i=1}^{\infty} u_i X_{t-i}^2\Big)^{1/2}$

$$\Longrightarrow M_\theta\big(X_{t-1}, \dots\big) = \Big(\omega_0 + \sum_{i=1}^{\infty} u_i X_{t-i}^2\Big)^{1/2}, \;\; f_\theta\big(X_{t-1}, \dots\big) = 0$$

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 $\xi_t = \sum_{k=0}^{\infty} (-b_1)^k X_{t-k} \quad \Longrightarrow X_t = \xi_t + b_1 \sum_{k=0}^{\infty} (-b_1)^k X_{t-1-k}$

$$\Longrightarrow M_\theta(X_{t-1}, \dots) = \sigma_\xi, \;\; f_\theta(X_{t-1}, \dots) = b_1 \sum_{k=0}^{\infty} (-b_1)^k X_{t-1-k}$$

• 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

$$\Longrightarrow M_\theta(X_{t-1}, \dots) = \sigma_\xi, \;\; f_\theta(X_{t-1}, \dots) = \sum_{k=1}^{\infty} u_k X_{t-k}$$

• Si (X_t) GARCH(p,q), on montre que $X_t = \xi_t \left(\omega_0 + \sum_{i=1}^{\infty} u_i X_{t-i}^2\right)^{1/2}$

$$\Longrightarrow M_\theta(X_{t-1}, \dots) = \left(\omega_0 + \sum_{i=1}^{\infty} u_i X_{t-i}^2\right)^{1/2}, \;\; f_\theta(X_{t-1}, \dots) = 0$$

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 $\xi_t = \sum_{k=0}^{\infty} (-b_1)^k X_{t-k} \quad \Longrightarrow X_t = \xi_t + b_1 \sum_{k=0}^{\infty} (-b_1)^k X_{t-1-k}$

$$\Longrightarrow M_\theta\big(X_{t-1}, \dots\big) = \sigma_\xi, \;\; f_\theta\big(X_{t-1}, \dots\big) = b_1 \sum_{k=0}^{\infty} (-b_1)^k X_{t-1-k}$$

• 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

$$\Longrightarrow M_\theta\big(X_{t-1}, \dots\big) = \sigma_\xi, \;\; f_\theta\big(X_{t-1}, \dots\big) = \sum_{k=1}^{\infty} u_k X_{t-k}$$

• Si (X_t) GARCH(p,q), on montre que $X_t = \xi_t \left(\omega_0 + \sum_{i=1}^{\infty} u_i X_{t-i}^2\right)^{1/2}$

$$\Longrightarrow M_\theta\big(X_{t-1}, \dots\big) = \left(\omega_0 + \sum_{i=1}^{\infty} u_i X_{t-i}^2\right)^{1/2}, \;\; f_\theta\big(X_{t-1}, \dots\big) = 0$$

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
$\xi_t = \sum_{k=0}^{\infty} (-b_1)^k X_{t-k} \quad \Longrightarrow X_t = \xi_t + b_1 \sum_{k=0}^{\infty} (-b_1)^k X_{t-1-k}$

$$\Longrightarrow M_\theta\big(X_{t-1},\dots\big) = \sigma_\xi, \;\; f_\theta\big(X_{t-1},\dots\big) = b_1 \sum_{k=0}^{\infty} (-b_1)^k X_{t-1-k}$$

• Par extension si (X_t) ARMA(p,q), $X_t = \xi_t + \sum_{k=1}^{\infty} u_k X_{t-k}$ quand $Q(x) = 1 + b_1 x + \cdots + b_q x^q$ a ses racines en dehors du disque unité

$$\Longrightarrow M_\theta\big(X_{t-1},\dots\big) = \sigma_\xi, \;\; f_\theta\big(X_{t-1},\dots\big) = \sum_{k=1}^{\infty} u_k X_{t-k}$$

• Si (X_t) GARCH(p,q), on montre que $X_t = \xi_t \Big(\omega_0 + \sum_{i=1}^{\infty} u_i X_{t-i}^2\Big)^{1/2}$

$$\Longrightarrow M_\theta\big(X_{t-1},\dots\big) = \Big(\omega_0 + \sum_{i=1}^{\infty} u_i X_{t-i}^2\Big)^{1/2}, \;\; f_\theta\big(X_{t-1},\dots\big) = 0$$

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
$\xi_t = \sum_{k=0}^{\infty} (-b_1)^k X_{t-k} \quad \Longrightarrow X_t = \xi_t + b_1 \sum_{k=0}^{\infty} (-b_1)^k X_{t-1-k}$

$$\Longrightarrow M_\theta\big(X_{t-1}, \dots\big) = \sigma_\xi, \;\; f_\theta\big(X_{t-1}, \dots\big) = b_1 \sum_{k=0}^{\infty} (-b_1)^k X_{t-1-k}$$

• Par extension si (X_t) ARMA(p, q), $X_t = \xi_t + \sum_{k=1}^{\infty} u_k X_{t-k}$ quand $Q(x) = 1 + b_1 x + \cdots + b_q x^q$ a ses racines en dehors du disque unité

$$\Longrightarrow M_\theta\big(X_{t-1}, \dots\big) = \sigma_\xi, \;\; f_\theta\big(X_{t-1}, \dots\big) = \sum_{k=1}^{\infty} u_k X_{t-k}$$

• Si (X_t) GARCH(p, q), on montre que $X_t = \xi_t \Big(\omega_0 + \sum_{i=1}^{\infty} u_i X_{t-i}^2\Big)^{1/2}$

$$\Longrightarrow M_\theta\big(X_{t-1}, \dots\big) = \Big(\omega_0 + \sum_{i=1}^{\infty} u_i X_{t-i}^2\Big)^{1/2}, \;\; f_\theta\big(X_{t-1}, \dots\big) = 0$$

Estimation par maximum de vraisemblance

$$X_t = M_\theta\big(X_{t-1}, X_{t-2}, \dots\big)\,\xi_t + f_\theta\big(X_{t-1}, X_{t-2}, \dots\big)$$

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

$$X_t = M_\theta\big(X_{t-1}, X_{t-2}, \dots\big)\,\xi_t + f_\theta\big(X_{t-1}, X_{t-2}, \dots\big)$$

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

$$X_t = M_\theta\big(X_{t-1}, X_{t-2}, \dots\big)\, \xi_t + f_\theta\big(X_{t-1}, X_{t-2}, \dots\big)$$

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

$$X_t = M_\theta\big(X_{t-1}, X_{t-2}, \dots\big)\, \xi_t + f_\theta\big(X_{t-1}, X_{t-2}, \dots\big)$$

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

$$X_t = M_\theta\big(X_{t-1}, X_{t-2}, \dots\big)\, \xi_t + f_\theta\big(X_{t-1}, X_{t-2}, \dots\big)$$

Seconde idée : Utiliser le maximum de vraisemblance conditionnelle

$\Longrightarrow$ On calcule la densité de $(X_1, \dots, X_n)$ sachant $(X_0, X_{-1}, \dots)$.

On obtient alors la vraisemblance conditionnelle :

$$\begin{aligned} L_\theta(X_1, \dots, X_n) = & f(X_n \mid X_{n-1}, X_{n-2}, \dots) \\ & \times f(X_{n-1} \mid X_{n-2}, X_{n-3}, \dots) \times \cdots \times f(X_1 \mid X_0, X_{-1}, \dots) \end{aligned}$$

Si (ξ_t) gaussien $\mathcal{N}(0,1)$, X_k est gaussien sachant $X_{k-1}, X_{k-2}, \dots$ d'où :

$$f(X_k \mid X_{k-1}, X_{k-2}, \dots) = \frac{1}{\sqrt{2\pi}} \frac{1}{M_\theta\big(X_{k-1}, X_{k-2}, \dots\big)} \exp\Big(-\frac{1}{2} \frac{\big(X_k - f_\theta\big(X_{k-1}, X_{k-2}, \dots\big)\big)^2}{M_\theta^2\big(X_{k-1}, X_{k-2}, \dots\big)}\Big)$$

Estimation par maximum de vraisemblance conditionnelle

$$X_t = M_\theta\big(X_{t-1}, X_{t-2}, \ldots\big)\, \xi_t + f_\theta\big(X_{t-1}, X_{t-2}, \ldots\big)$$

Seconde idée : Utiliser le maximum de vraisemblance conditionnelle

$\Longrightarrow$ On calcule la densité de $(X_1, \ldots, X_n)$ sachant $(X_0, X_{-1}, \ldots)$.

On obtient alors la vraisemblance conditionnelle :

$$\begin{aligned} L_\theta(X_1, \ldots, X_n) = {} & f(X_n \mid X_{n-1}, X_{n-2}, \ldots) \\ & \times f(X_{n-1} \mid X_{n-2}, X_{n-3}, \ldots) \times \cdots \times f(X_1 \mid X_0, X_{-1}, \ldots) \end{aligned}$$

Si (ξ_t) gaussien $\mathcal{N}(0,1)$, X_k est gaussien sachant $X_{k-1}, X_{k-2}, \ldots$ d'où :

$$f(X_k \mid X_{k-1}, X_{k-2}, \ldots) = \frac{1}{\sqrt{2\pi}} \frac{1}{M_\theta\big(X_{k-1}, X_{k-2}, \ldots\big)} \exp\Big(- \frac{1}{2} \frac{\big(X_k - f_\theta\big(X_{k-1}, X_{k-2}, \ldots\big)\big)^2}{M_\theta^2\big(X_{k-1}, X_{k-2}, \ldots\big)}\Big)$$

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

$\Longrightarrow$ On calcule la densité de $(X_1, \dots, X_n)$ sachant $(X_0, X_{-1}, \dots)$.

On obtient alors la vraisemblance conditionnelle :

$$L_\theta(X_1, \dots, X_n) = f(X_n \mid X_{n-1}, X_{n-2}, \dots) \times f(X_{n-1} \mid X_{n-2}, X_{n-3}, \dots) \times \dots \times f(X_1 \mid X_0, X_{-1}, \dots)$$

Si (ξ_t) gaussien $\mathcal{N}(0,1)$, X_k est gaussien sachant $X_{k-1}, X_{k-2}, \dots$ d'où :

$$f(X_k \mid X_{k-1}, X_{k-2}, \dots) = \frac{1}{\sqrt{2\pi}} \frac{1}{M_\theta(X_{k-1}, X_{k-2}, \dots)} \exp\Big(-\frac{1}{2} \frac{\big(X_k - f_\theta(X_{k-1}, X_{k-2}, \dots)\big)^2}{M_\theta^2(X_{k-1}, X_{k-2}, \dots)}\Big)$$

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

$\Longrightarrow$ On calcule la densité de $(X_1, \dots, X_n)$ sachant $(X_0, X_{-1}, \dots)$.

On obtient alors la vraisemblance conditionnelle :

$$\begin{aligned} L_\theta(X_1, \dots, X_n) = f(X_n \mid X_{n-1}, X_{n-2}, \dots) \\ \times f(X_{n-1} \mid X_{n-2}, X_{n-3}, \dots) \times \cdots \times f(X_1 \mid X_0, X_{-1}, \dots) \end{aligned}$$

Si (ξ_t) gaussien $\mathcal{N}(0,1)$, X_k est gaussien sachant $X_{k-1}, X_{k-2}, \dots$ d'où :

$$f(X_k \mid X_{k-1}, X_{k-2}, \dots) = \frac{1}{\sqrt{2\pi}} \frac{1}{M_\theta(X_{k-1}, X_{k-2}, \dots)} \exp\Big(-\frac{1}{2} \frac{\big(X_k - f_\theta(X_{k-1}, X_{k-2}, \dots)\big)^2}{M_\theta^2(X_{k-1}, X_{k-2}, \dots)} \Big)$$

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

$\Longrightarrow$ On calcule la densité de $(X_1, \dots, X_n)$ sachant $(X_0, X_{-1}, \dots)$.

On obtient alors la vraisemblance conditionnelle :

$$\begin{aligned} L_\theta(X_1, \dots, X_n) = f(X_n \mid X_{n-1}, X_{n-2}, \dots) \\ \times f(X_{n-1} \mid X_{n-2}, X_{n-3}, \dots) \times \cdots \times f(X_1 \mid X_0, X_{-1}, \dots) \end{aligned}$$

Si (ξ_t) gaussien $\mathcal{N}(0,1)$, X_k est gaussien sachant $X_{k-1}, X_{k-2}, \dots$ d'où :

$$f(X_k \mid X_{k-1}, X_{k-2}, \dots) = \frac{1}{\sqrt{2\pi}} \frac{1}{M_\theta(X_{k-1}, X_{k-2}, \dots)} \exp\left(-\frac{1}{2} \frac{\left(X_k - f_\theta(X_{k-1}, X_{k-2}, \dots)\right)^2}{M_\theta^2(X_{k-1}, X_{k-2}, \dots)}\right)$$

Estimation par maximum de vraisemblance conditionnelle

$$X_t = M_\theta\big(X_{t-1}, X_{t-2}, \dots\big)\, \xi_t + f_\theta\big(X_{t-1}, X_{t-2}, \dots\big)$$

Seconde idée : Utiliser le maximum de vraisemblance conditionnelle

$\Longrightarrow$ On calcule la densité de $(X_1, \dots, X_n)$ sachant $(X_0, X_{-1}, \dots)$.

On obtient alors la vraisemblance conditionnelle :

$$\begin{aligned} L_\theta(X_1, \dots, X_n) &= f(X_n \mid X_{n-1}, X_{n-2}, \dots) \\ &\quad \times f(X_{n-1} \mid X_{n-2}, X_{n-3}, \dots) \times \dots \times f(X_1 \mid X_0, X_{-1}, \dots) \end{aligned}$$

Si (ξ_t) gaussien $\mathcal{N}(0,1)$, X_k est gaussien sachant $X_{k-1}, X_{k-2}, \dots$ d'où :

$$f(X_k \mid X_{k-1}, X_{k-2}, \dots) = \frac{1}{\sqrt{2\pi}} \frac{1}{M_\theta\big(X_{k-1}, X_{k-2}, \dots\big)} \exp\Big(-\frac{1}{2} \frac{\big(X_k - f_\theta\big(X_{k-1}, X_{k-2}, \dots\big)\big)^2}{M_\theta^2\big(X_{k-1}, X_{k-2}, \dots\big)}\Big)$$

Estimation par maximum de vraisemblance conditionnelle

$$X_t = M_\theta\big(X_{t-1}, X_{t-2}, \dots\big)\, \xi_t + f_\theta\big(X_{t-1}, X_{t-2}, \dots\big)$$

Seconde idée : Utiliser le maximum de vraisemblance conditionnelle

$\Longrightarrow$ On calcule la densité de $(X_1, \dots, X_n)$ sachant $(X_0, X_{-1}, \dots)$.

On obtient alors la vraisemblance conditionnelle :

$$\begin{aligned} L_\theta(X_1, \dots, X_n) = {} & f(X_n \mid X_{n-1}, X_{n-2}, \dots) \\ & \times f(X_{n-1} \mid X_{n-2}, X_{n-3}, \dots) \times \cdots \times f(X_1 \mid X_0, X_{-1}, \dots) \end{aligned}$$

Si (ξ_t) gaussien $\mathcal{N}(0,1)$, X_k est gaussien sachant $X_{k-1}, X_{k-2}, \dots$ d'où :

$$f(X_k \mid X_{k-1}, X_{k-2}, \dots) = \frac{1}{\sqrt{2\pi}} \frac{1}{M_\theta\big(X_{k-1}, X_{k-2}, \dots\big)} \exp\Big(- \frac{1}{2} \frac{\big(X_k - f_\theta\big(X_{k-1}, X_{k-2}, \dots\big)\big)^2}{M_\theta^2\big(X_{k-1}, X_{k-2}, \dots\big)} \Big)$$

Estimation par quasi-maximum de vraisemblance

On obtient alors la log-vraisemblance conditionnelle :

$$\log\left(L_\theta(X_1,\ldots,X_n)\right) = -\frac{1}{2}\Big(n\log(2\pi) + \sum_{t=1}^{n}\log\left(M_\theta^2(X_{t-1},X_{t-2},\ldots)\right) + \frac{\left(X_t - f_\theta(X_{t-1},X_{t-2},\ldots)\right)^2}{M_\theta^2(X_{t-1},X_{t-2},\ldots)}\Big)$$

Mais $X_0, X_{-1},\ldots$ non observées alors que par exemple pour un MA(1) on a

$$f_\theta(X_{t-1},\ldots) = b_1\sum_{k=0}^{\infty}(-b_1)^k X_{t-1-k}$$

$$\Longrightarrow \left\{\begin{array}{l}\widehat{M}_\theta^t = M_\theta(X_{t-1},\ldots,X_1,0,0,\ldots)\\ \widehat{f}_\theta^t = f_\theta(X_{t-1},\ldots,X_1,0,0,\ldots)\end{array}\right. \quad \text{remplace} \left\{\begin{array}{l}M_\theta(X_{t-1},X_{t-2},\ldots)\\ f_\theta(X_{t-1},X_{t-2},\ldots)\end{array}\right.$$

Estimation par quasi-maximum de vraisemblance

On obtient alors la log-vraisemblance conditionnelle :

$$\log\left(L_\theta(X_1,\ldots,X_n)\right) = -\frac{1}{2}\Big(n\log(2\pi) + \sum_{t=1}^{n}\log\left(M_\theta^2\left(X_{t-1},X_{t-2},\ldots\right)\right) + \frac{\left(X_t - f_\theta\left(X_{t-1},X_{t-2},\ldots\right)\right)^2}{M_\theta^2\left(X_{t-1},X_{t-2},\ldots\right)}\Big)$$

Mais $X_0, X_{-1}, \ldots$ non observées alors que par exemple pour un MA(1) on a

$$f_\theta\left(X_{t-1},\ldots\right) = b_1\sum_{k=0}^{\infty}(-b_1)^k X_{t-1-k}$$

$$\Longrightarrow \begin{cases} \widehat{M}_\theta^t = M_\theta\left(X_{t-1},\ldots,X_1,0,0,\ldots\right) \\ \widehat{f}_\theta^t = f_\theta\left(X_{t-1},\ldots,X_1,0,0,\ldots\right) \end{cases} \quad \text{remplace} \quad \begin{cases} M_\theta\left(X_{t-1},X_{t-2},\ldots\right) \\ f_\theta\left(X_{t-1},X_{t-2},\ldots\right) \end{cases}$$

Estimation par quasi-maximum de vraisemblance

On obtient alors la log-vraisemblance conditionnelle :

$$\log\big(L_\theta(X_1,\dots,X_n)\big) = -\frac{1}{2}\Big(n\log(2\pi) + \sum_{t=1}^{n} \log\big(M_\theta^2(X_{t-1},X_{t-2},\dots)\big) + \frac{\big(X_t - f_\theta(X_{t-1},X_{t-2},\dots)\big)^2}{M_\theta^2(X_{t-1},X_{t-2},\dots)}\Big)$$

Mais $X_0, X_{-1},\dots$ non observées alors que par exemple pour un MA(1) on a

$$f_\theta(X_{t-1},\dots) = b_1 \sum_{k=0}^{\infty} (-b_1)^k X_{t-1-k}$$

$$\Longrightarrow \begin{cases} \widehat{M}_\theta^t = M_\theta(X_{t-1},\dots,X_1,0,0,\dots) \\ \widehat{f}_\theta^t = f_\theta(X_{t-1},\dots,X_1,0,0,\dots) \end{cases} \text{ remplace } \begin{cases} M_\theta(X_{t-1},X_{t-2},\dots) \\ f_\theta(X_{t-1},X_{t-2},\dots) \end{cases}$$

Estimation par quasi-maximum de vraisemblance

On obtient alors la log-vraisemblance conditionnelle :

$$\log\left(L_\theta(X_1,\dots,X_n)\right) = -\frac{1}{2}\Big(n\log(2\pi) + \sum_{t=1}^{n}\log\left(M_\theta^2\left(X_{t-1},X_{t-2},\dots\right)\right) + \frac{\left(X_t - f_\theta\left(X_{t-1},X_{t-2},\dots\right)\right)^2}{M_\theta^2\left(X_{t-1},X_{t-2},\dots\right)}\Big)$$

Mais $X_0, X_{-1},\dots$ non observées alors que par exemple pour un MA(1) on a

$$f_\theta\left(X_{t-1},\dots\right) = b_1\sum_{k=0}^{\infty}(-b_1)^k X_{t-1-k}$$

$$\Longrightarrow \begin{cases} \widehat{M}_\theta^t = M_\theta\left(X_{t-1},\dots,X_1,0,0,\dots\right) \\ \widehat{f}_\theta^t = f_\theta\left(X_{t-1},\dots,X_1,0,0,\dots\right) \end{cases} \text{remplace} \begin{cases} M_\theta\left(X_{t-1},X_{t-2},\dots\right) \\ f_\theta\left(X_{t-1},X_{t-2},\dots\right) \end{cases}$$

Estimation par quasi-maximum de vraisemblance

On obtient alors la quasi-log-vraisemblance conditionnelle :

$$\log\left(\widehat{L}_\theta(X_1,\dots,X_n)\right) = -\frac{1}{2}\left(\sum_{t=1}^{n}\log\left((\widehat{M}_\theta^t)^2\right) + \frac{\left(X_t-\widehat{f}_\theta^t\right)^2}{\left(\widehat{M}_\theta^t\right)^2}\right)$$

On définit **l'estimateur par quasi-maximum de vraisemblance** QMLE :

$$\widehat{\theta}_n = \operatorname*{Argmax}_{\theta\in\Theta}\log\left(\widehat{L}_\theta(X_1,\dots,X_n)\right)$$

Exemples :

- Pour un processus AR(p),

$$(\widehat{a}_1,\dots,\widehat{a}_p) \simeq \operatorname*{Argmin}_{(a_1,\dots,a_p)\in A\subset\mathbb{R}^p}\sum_{k=p+1}^{n}\left(X_k-(a_1X_{k-1}+\cdots+a_pX_{k-p})\right)^2$$

- Pour un processus ARCH(p),

$$(\widehat{\omega}_0,\widehat{a}_1,\dots,\widehat{a}_p) \simeq \operatorname*{Argmin}_{(\omega_0,a_1,\dots,a_p)\in A\subset\mathbb{R}^{p+1}}\sum_{k=p+1}^{n}\log\left(\omega_0+a_1X_{k-1}^2+\cdots+a_pX_{k-p}^2\right)$$

$$+\frac{X_k^2}{\omega_0+a_1X_{k-1}^2+\cdots+a_pX_{k-p}^2}$$

Estimation par quasi-maximum de vraisemblance

On obtient alors la quasi-log-vraisemblance conditionnelle :

$$\log\left(\widehat{L}_\theta(X_1,\dots,X_n)\right) = -\frac{1}{2}\left(\sum_{t=1}^{n}\log\left((\widehat{M}_\theta^t)^2\right) + \frac{\left(X_t - \widehat{f}_\theta^t\right)^2}{\left(\widehat{M}_\theta^t\right)^2}\right)$$

On définit **l'estimateur par quasi-maximum de vraisemblance** QMLE :

$$\widehat{\theta}_n = \operatorname*{Argmax}_{\theta\in\Theta}\log\left(\widehat{L}_\theta(X_1,\dots,X_n)\right)$$

Exemples :

- Pour un processus AR(p),

$$(\widehat{a}_1,\dots,\widehat{a}_p) \simeq \operatorname*{Argmin}_{(a_1,\dots,a_p)\in A\subset\mathbb{R}^p}\sum_{k=p+1}^{n}\left(X_k - (a_1X_{k-1}+\cdots+a_pX_{k-p})\right)^2$$

- Pour un processus ARCH(p),

$$(\widehat{\omega}_0,\widehat{a}_1,\dots,\widehat{a}_p) \simeq \operatorname*{Argmin}_{(\omega_0,a_1,\dots,a_p)\in A\subset\mathbb{R}^{p+1}}\sum_{k=p+1}^{n}\log\left(\omega_0 + a_1X_{k-1}^2+\cdots+a_pX_{k-p}^2\right)$$

$$+\frac{X_k^2}{\omega_0 + a_1X_{k-1}^2+\cdots+a_pX_{k-p}^2}$$

Estimation par quasi-maximum de vraisemblance

On obtient alors la quasi-log-vraisemblance conditionnelle :

$$\log\left(\widehat{L}_\theta(X_1,\dots,X_n)\right) = -\frac{1}{2}\left(\sum_{t=1}^{n}\log\left((\widehat{M}_\theta^t)^2\right) + \frac{\left(X_t - \widehat{f}_\theta^t\right)^2}{\left(\widehat{M}_\theta^t\right)^2}\right)$$

On définit **l'estimateur par quasi-maximum de vraisemblance** QMLE :

$$\widehat{\theta}_n = \operatorname*{Argmax}_{\theta\in\Theta}\log\left(\widehat{L}_\theta(X_1,\dots,X_n)\right)$$

Exemples :

- Pour un processus AR(p),

$$(\widehat{a}_1,\dots,\widehat{a}_p) \simeq \operatorname*{Argmin}_{(a_1,\dots,a_p)\in A\subset\mathbb{R}^p}\sum_{k=p+1}^{n}\left(X_k-(a_1X_{k-1}+\dots+a_pX_{k-p})\right)^2$$

- Pour un processus ARCH(p),

$$(\widehat{\omega}_0,\widehat{a}_1,\dots,\widehat{a}_p) \simeq \operatorname*{Argmin}_{(\omega_0,a_1,\dots,a_p)\in A\subset\mathbb{R}^{p+1}}\sum_{k=p+1}^{n}\log\left(\omega_0+a_1X_{k-1}^2+\dots+a_pX_{k-p}^2\right)$$

$$+\frac{X_k^2}{\omega_0+a_1X_{k-1}^2+\dots+a_pX_{k-p}^2}$$

Estimation par quasi-maximum de vraisemblance

On obtient alors la quasi-log-vraisemblance conditionnelle :

$$\log\left(\widehat{L}_\theta(X_1,\dots,X_n)\right) = -\frac{1}{2}\left(\sum_{t=1}^{n}\log\left((\widehat{M}_\theta^t)^2\right) + \frac{\left(X_t - \widehat{f}_\theta^t\right)^2}{\left(\widehat{M}_\theta^t\right)^2}\right)$$

On définit **l'estimateur par quasi-maximum de vraisemblance** QMLE :

$$\widehat{\theta}_n = \underset{\theta\in\Theta}{\operatorname{Argmax}}\log\left(\widehat{L}_\theta(X_1,\dots,X_n)\right)$$

Exemples :

- Pour un processus AR(p),

$$(\widehat{a}_1,\dots,\widehat{a}_p) \simeq \underset{(a_1,\dots,a_p)\in A\subset\mathbb{R}^p}{\operatorname{Argmin}}\sum_{k=p+1}^{n}\left(X_k - (a_1X_{k-1}+\cdots+a_pX_{k-p})\right)^2$$

- Pour un processus ARCH(p),

$$(\widehat{\omega}_0,\widehat{a}_1,\dots,\widehat{a}_p) \simeq \underset{(\omega_0,a_1,\dots,a_p)\in A\subset\mathbb{R}^{p+1}}{\operatorname{Argmin}}\sum_{k=p+1}^{n}\log\left(\omega_0 + a_1X_{k-1}^2+\cdots+a_pX_{k-p}^2\right) + \frac{X_k^2}{\omega_0 + a_1X_{k-1}^2+\cdots+a_pX_{k-p}^2}$$

Estimation par quasi-maximum de vraisemblance

On obtient alors la quasi-log-vraisemblance conditionnelle :

$$\log\left(\widehat{L}_\theta(X_1,\ldots,X_n)\right) = -\frac{1}{2}\left(\sum_{t=1}^{n}\log\left((\widehat{M}_\theta^t)^2\right) + \frac{\left(X_t - \widehat{f}_\theta^t\right)^2}{\left(\widehat{M}_\theta^t\right)^2}\right)$$

On définit **l'estimateur par quasi-maximum de vraisemblance** QMLE :

$$\widehat{\theta}_n = \underset{\theta\in\Theta}{\mathrm{Argmax}}\,\log\left(\widehat{L}_\theta(X_1,\ldots,X_n)\right)$$

Exemples :

- Pour un processus AR(p),

$$(\widehat{a}_1,\ldots,\widehat{a}_p) \simeq \underset{(a_1,\ldots,a_p)\in A\subset\mathbb{R}^p}{\mathrm{Argmin}}\sum_{k=p+1}^{n}\left(X_k - (a_1X_{k-1}+\cdots+a_pX_{k-p})\right)^2$$

- Pour un processus ARCH(p),

$$(\widehat{\omega}_0,\widehat{a}_1,\ldots,\widehat{a}_p) \simeq \underset{(\omega_0,a_1,\ldots,a_p)\in A\subset\mathbb{R}^{p+1}}{\mathrm{Argmin}}\sum_{k=p+1}^{n}\log\left(\omega_0 + a_1X_{k-1}^2+\cdots+a_pX_{k-p}^2\right)$$

$$+\frac{X_k^2}{\omega_0 + a_1X_{k-1}^2+\cdots+a_pX_{k-p}^2}$$

Estimation par quasi-maximum de vraisemblance

On obtient alors la quasi-log-vraisemblance conditionnelle :

$$\log\left(\widehat{L}_\theta(X_1,\ldots,X_n)\right) = -\frac{1}{2}\left(\sum_{t=1}^{n}\log\left((\widehat{M}_\theta^t)^2\right) + \frac{\left(X_t - \widehat{f}_\theta^t\right)^2}{\left(\widehat{M}_\theta^t\right)^2}\right)$$

On définit **l'estimateur par quasi-maximum de vraisemblance** QMLE :

$$\widehat{\theta}_n = \operatorname*{Argmax}_{\theta\in\Theta}\log\left(\widehat{L}_\theta(X_1,\ldots,X_n)\right)$$

Exemples :

- Pour un processus AR(p),

$$(\widehat{a}_1,\ldots,\widehat{a}_p) \simeq \operatorname*{Argmin}_{(a_1,\ldots,a_p)\in A\subset\mathbb{R}^p}\sum_{k=p+1}^{n}\left(X_k - (a_1X_{k-1}+\cdots+a_pX_{k-p})\right)^2$$

- Pour un processus ARCH(p),

$$(\widehat{\omega}_0,\widehat{a}_1,\ldots,\widehat{a}_p) \simeq \operatorname*{Argmin}_{(\omega_0,a_1,\ldots,a_p)\in A\subset\mathbb{R}^{p+1}}\sum_{k=p+1}^{n}\log\left(\omega_0 + a_1X_{k-1}^2+\cdots+a_pX_{k-p}^2\right)$$

$$+\frac{X_k^2}{\omega_0 + a_1X_{k-1}^2+\cdots+a_pX_{k-p}^2}$$

Convergence de l'estimateur QML

On peut montrer que :

Théorème

Sous certaines conditions, et $\mathbb{E}[\xi_0^4] < \infty$ **même si** (ξ_t) **non gaussien** *on a*

$$\sqrt{n}\big(\widehat{\theta}_n - \theta^*\big) \underset{n\to\infty}{\overset{\mathcal{L}}{\longrightarrow}} \mathcal{N}_d\big(0\ ,\ F(\theta^*)^{-1}G(\theta^*)F(\theta^*)^{-1}\big),$$

$$avec \quad \begin{cases} F(\theta^*) &= \lim_{n\to\infty} \frac{1}{n}\left[\frac{\partial^2}{\partial\theta_i\partial\theta_j}\log\left(\widehat{L}_{\theta^*}(X_1,\dots,X_n)\right)\right]_{i,j} \\ G(\theta^*) &= \lim_{n\to\infty} \frac{1}{n}\left[\frac{\partial}{\partial\theta_i}\log\left(\widehat{L}_{\theta^*}(X_1,\dots,X_n)\right)\frac{\partial}{\partial\theta_j}\log\left(\widehat{L}_{\theta^*}(X_1,\dots,X_n)\right)\right]_{i,j} \end{cases}$$

Exemples : Le théorème est valable pour les processus ARMA, GARCH, APARCH, FARIMA, ARMA-GARCH,...

Remarque : Avec le Lemme de Slutsky et $\widehat{F}(\widehat{\theta}_n)$ et $\widehat{G}(\widehat{\theta}_n)$ au lieu de $F(\theta^*)$ et $G(\theta^*)$ (sans lim), on obtient des intervalles de confiances, tests de Wald.

Convergence de l'estimateur QML

On peut montrer que :

Théorème

Sous certaines conditions, et $\mathbb{E}[\xi_0^4] < \infty$ **même si** (ξ_t) **non gaussien** *on a*

$$\sqrt{n}\left(\widehat{\theta}_n - \theta^*\right) \xrightarrow[n\to\infty]{\mathcal{L}} \mathcal{N}_d\left(0\ ,\ F(\theta^*)^{-1} G(\theta^*) F(\theta^*)^{-1}\right),$$

$$\text{avec} \quad \begin{cases} F(\theta^*) &= \lim_{n\to\infty} \frac{1}{n} \left[\frac{\partial^2}{\partial\theta_i \partial\theta_j} \log\left(\widehat{L}_{\theta^*}(X_1,\ldots,X_n)\right)\right]_{i,j} \\ G(\theta^*) &= \lim_{n\to\infty} \frac{1}{n} \left[\frac{\partial}{\partial\theta_i} \log\left(\widehat{L}_{\theta^*}(X_1,\ldots,X_n)\right) \frac{\partial}{\partial\theta_j} \log\left(\widehat{L}_{\theta^*}(X_1,\ldots,X_n)\right)\right]_{i,j} \end{cases}$$

Exemples : Le théorème est valable pour les processus ARMA, GARCH, APARCH, FARIMA, ARMA-GARCH,...

Remarque : Avec le Lemme de Slutsky et $\widehat{F}(\widehat{\theta}_n)$ et $\widehat{G}(\widehat{\theta}_n)$ au lieu de $F(\theta^*)$ et $G(\theta^*)$ (sans lim), on obtient des intervalles de confiances, tests de Wald,...

Convergence de l'estimateur QML

On peut montrer que :

Théorème

Sous certaines conditions, et $\mathbb{E}[\xi_0^4] < \infty$ **même si** (ξ_t) **non gaussien** *on a*

$$\sqrt{n}\big(\widehat{\theta}_n - \theta^*\big) \ \xrightarrow[n\to\infty]{\mathcal{L}} \ \mathcal{N}_d\big(0\ ,\ F(\theta^*)^{-1} G(\theta^*) F(\theta^*)^{-1}\big),$$

$$\text{avec} \quad \begin{cases} F(\theta^*) & = & \lim_{n\to\infty} \frac{1}{n}\left[\frac{\partial^2}{\partial\theta_i\partial\theta_j} \log\left(\widehat{L}_{\theta^*}(X_1,\ldots,X_n)\right)\right]_{i,j} \\ G(\theta^*) & = & \lim_{n\to\infty} \frac{1}{n}\left[\frac{\partial}{\partial\theta_i} \log\left(\widehat{L}_{\theta^*}(X_1,\ldots,X_n)\right) \frac{\partial}{\partial\theta_j} \log\left(\widehat{L}_{\theta^*}(X_1,\ldots,X_n)\right)\right]_{i,j} \end{cases}$$

Exemples : Le théorème est valable pour les processus ARMA, GARCH, APARCH, FARIMA, ARMA-GARCH,...

Remarque : Avec le Lemme de Slutsky et $\widehat{F}(\widehat{\theta}_n)$ et $\widehat{G}(\widehat{\theta}_n)$ au lieu de $F(\theta^*)$ et $G(\theta^*)$ (sans lim), on obtient des intervalles de confiances, tests de Wald,...

Convergence de l'estimateur QML

On peut montrer que :

Théorème

Sous certaines conditions, et $\mathbb{E}[\xi_0^4] < \infty$ **même si** (ξ_t) **non gaussien** *on a*

$$\sqrt{n}\big(\widehat{\theta}_n - \theta^*\big) \xrightarrow[n\to\infty]{\mathcal{L}} \mathcal{N}_d\Big(0\ ,\ F(\theta^*)^{-1} G(\theta^*) F(\theta^*)^{-1}\Big),$$

$$\textit{avec}\quad \begin{cases} F(\theta^*) &= \lim_{n\to\infty} \frac{1}{n}\left[\frac{\partial^2}{\partial\theta_i\partial\theta_j}\log\left(\widehat{L}_{\theta^*}(X_1,\ldots,X_n)\right)\right]_{i,j} \\ G(\theta^*) &= \lim_{n\to\infty} \frac{1}{n}\left[\frac{\partial}{\partial\theta_i}\log\left(\widehat{L}_{\theta^*}(X_1,\ldots,X_n)\right)\frac{\partial}{\partial\theta_j}\log\left(\widehat{L}_{\theta^*}(X_1,\ldots,X_n)\right)\right]_{i,j} \end{cases}$$

Exemples : Le théorème est valable pour les processus ARMA, GARCH, APARCH, FARIMA, ARMA-GARCH,...

Remarque : Avec le Lemme de Slutsky et $\widehat{F}(\widehat{\theta}_n)$ et $\widehat{G}(\widehat{\theta}_n)$ au lieu de $F(\theta^*)$ et $G(\theta^*)$ (sans lim), on obtient des intervalles de confiances, tests de Wald

Convergence de l'estimateur QML

On peut montrer que :

Théorème

Sous certaines conditions, et $\mathbb{E}[\xi_0^4] < \infty$ **même si** (ξ_t) **non gaussien** *on a*

$$\sqrt{n}\big(\widehat{\theta}_n - \theta^*\big) \xrightarrow[n\to\infty]{\mathcal{L}} \mathcal{N}_d\Big(0\ ,\ F(\theta^*)^{-1} G(\theta^*) F(\theta^*)^{-1}\Big),$$

$$\textit{avec} \quad \left\{ \begin{array}{rcl} F(\theta^*) & = & \lim_{n\to\infty} \frac{1}{n} \Big[\frac{\partial^2}{\partial\theta_i \partial\theta_j} \log\big(\widehat{L}_{\theta^*}(X_1,\ldots,X_n)\big)\Big]_{i,j} \\ G(\theta^*) & = & \lim_{n\to\infty} \frac{1}{n} \Big[\frac{\partial}{\partial\theta_i} \log\big(\widehat{L}_{\theta^*}(X_1,\ldots,X_n)\big) \frac{\partial}{\partial\theta_j} \log\big(\widehat{L}_{\theta^*}(X_1,\ldots,X_n)\big)\Big]_{i,j} \end{array} \right.$$

Exemples : Le théorème est valable pour les processus ARMA, GARCH, APARCH, FARIMA, ARMA-GARCH,...

Remarque : Avec le Lemme de Slutsky et $\widehat{F}(\widehat{\theta}_n)$ et $\widehat{G}(\widehat{\theta}_n)$ au lieu de $F(\theta^*)$ et $G(\theta^*)$ (sans lim), on obtient des intervalles de confiances, tests de Wald...

Convergence de l'estimateur QML

On peut montrer que :

Théorème

Sous certaines conditions, et $\mathbb{E}[\xi_0^4] < \infty$ **même si** (ξ_t) **non gaussien** *on a*

$$\sqrt{n}\big(\widehat{\theta}_n - \theta^*\big) \xrightarrow[n\to\infty]{\mathcal{L}} \mathcal{N}_d\big(0\ ,\ F(\theta^*)^{-1}G(\theta^*)F(\theta^*)^{-1}\big),$$

$$\textit{avec} \quad \left\{ \begin{array}{lll} F(\theta^*) & = & \lim_{n\to\infty} \frac{1}{n} \left[\frac{\partial^2}{\partial\theta_i\partial\theta_j} \log\left(\widehat{L}_{\theta^*}(X_1,\ldots,X_n)\right)\right]_{i,j} \\ G(\theta^*) & = & \lim_{n\to\infty} \frac{1}{n} \left[\frac{\partial}{\partial\theta_i} \log\left(\widehat{L}_{\theta^*}(X_1,\ldots,X_n)\right) \frac{\partial}{\partial\theta_j} \log\left(\widehat{L}_{\theta^*}(X_1,\ldots,X_n)\right)\right]_{i,j} \end{array} \right.$$

Exemples : Le théorème est valable pour les processus ARMA, GARCH, APARCH, FARIMA, ARMA-GARCH,...

Remarque : Avec le Lemme de Slutsky et $\widehat{F}(\widehat{\theta}_n)$ et $\widehat{G}(\widehat{\theta}_n)$ au lieu de $F(\theta^*)$ et $G(\theta^*)$ (sans lim), on obtient des intervalles de confiances, tests de Wald,...

Plan du cours

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 $\theta(m) \in \mathbb{R}^d$

$$X_t = M_{\theta(m)}\big(X_{t-1}, X_{t-2}, \dots\big)\, \xi_t + f_{\theta(m)}\big(X_{t-1}, X_{t-2}, \dots\big)$$

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 $(X_1, \ldots, X_n)$ 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 $\theta(m) \in \mathbb{R}^d$

$$X_t = M_{\theta(m)}\big(X_{t-1}, X_{t-2}, \ldots\big)\,\xi_t + f_{\theta(m)}\big(X_{t-1}, X_{t-2}, \ldots\big)$$

Exemple : $\mathcal{M}$ famille des AR(p) pour $0 \leq p \leq 10$: $\theta(m) = (a_1, \ldots, a_{10}, \sigma_\xi^2)$

Sélection de modèles

Problématique : On suppose $(X_1, \ldots, X_n)$ 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 $\theta(m) \in \mathbb{R}^d$

$$X_t = M_{\theta(m)}\big(X_{t-1}, X_{t-2}, \ldots\big)\,\xi_t + f_{\theta(m)}\big(X_{t-1}, X_{t-2}, \ldots\big)$$

Exemple : $\mathcal{M}$ famille des AR(p) pour $0 \leq p \leq 10$: $\theta(m) = (a_1, \ldots, a_{10}, \sigma_\xi^2)$

Sélection de modèles

Problématique : On suppose $(X_1, \ldots, X_n)$ 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 $\theta(m) \in \mathbb{R}^d$

$$X_t = M_{\theta(m)}\big(X_{t-1}, X_{t-2}, \ldots\big)\,\xi_t + f_{\theta(m)}\big(X_{t-1}, X_{t-2}, \ldots\big)$$

Exemple : $\mathcal{M}$ famille des AR(p) pour $0 \le p \le 10$: $\theta(m) = (a_1, \ldots, a_{10}, \sigma_\xi^2)$

Sélection de modèles

Problématique : On suppose $(X_1, \ldots, X_n)$ 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 $\mathrm{GARCH}(1,1)$ ou par un $\mathrm{ARCH}(2)$? Ou bien un $\mathrm{ARMA}(3,2)$?

Définition

On considère $\mathcal{M}$ une famille de modèles affines causaux et on notera $m \in \mathcal{M}$ un modèle, vérifiant, avec $\theta(m) \in \mathbb{R}^d$

$$X_t = M_{\theta(m)}\big(X_{t-1}, X_{t-2}, \ldots\big)\,\xi_t + f_{\theta(m)}\big(X_{t-1}, X_{t-2}, \ldots\big)$$

Exemple : $\mathcal{M}$ famille des $\mathrm{AR}(p)$ pour $0 \leq p \leq 10$: $\theta(m) = (a_1, \ldots, a_{10}, \sigma_\xi^2)$

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 :

$$\widehat{m}_{AIC} = \operatorname*{Argmin}_{m\in\mathcal{M}} \widehat{AIC}(m) \quad et \quad \widehat{m}_{BIC} = \operatorname*{Argmin}_{m\in\mathcal{M}} \widehat{BIC}(m).$$

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 :

$$\widehat{m}_{AIC} = \operatorname*{Argmin}_{m \in \mathcal{M}} \widehat{AIC}(m) \quad et \quad \widehat{m}_{BIC} = \operatorname*{Argmin}_{m \in \mathcal{M}} \widehat{BIC}(m).$$

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 :

$$\widehat{m}_{AIC} = \operatorname*{Argmin}_{m\in\mathcal{M}} \widehat{AIC}(m) \quad \text{et} \quad \widehat{m}_{BIC} = \operatorname*{Argmin}_{m\in\mathcal{M}} \widehat{BIC}(m).$$

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 :

$$\widehat{m}_{AIC} = \operatorname*{Argmin}_{m\in\mathcal{M}} \widehat{AIC}(m) \quad \text{et} \quad \widehat{m}_{BIC} = \operatorname*{Argmin}_{m\in\mathcal{M}} \widehat{BIC}(m).$$

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 :

$$\widehat{m}_{AIC} = \operatorname*{Argmin}_{m \in \mathcal{M}} \widehat{AIC}(m) \quad \text{et} \quad \widehat{m}_{BIC} = \operatorname*{Argmin}_{m \in \mathcal{M}} \widehat{BIC}(m).$$

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 :

$$\widehat{m}_{AIC} = \operatorname*{Argmin}_{m\in\mathcal{M}} \widehat{AIC}(m) \quad \text{et} \quad \widehat{m}_{BIC} = \operatorname*{Argmin}_{m\in\mathcal{M}} \widehat{BIC}(m).$$

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 :

$$\widehat{m}_{AIC} = \operatorname*{Argmin}_{m \in \mathcal{M}} \widehat{AIC}(m) \quad \text{et} \quad \widehat{m}_{BIC} = \operatorname*{Argmin}_{m \in \mathcal{M}} \widehat{BIC}(m).$$

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 $(X_1, \cdots, X_n)$ trajectoire observée issue de m^* et sous certaines conditions,

$$\mathbb{P}(\widehat{m}_{BIC} = m^*) \underset{n\to+\infty}{\longrightarrow} 1 \quad \text{et} \quad \mathbb{P}(m^* \subset \widehat{m}_{AIC}) \underset{n\to+\infty}{\longrightarrow} 1 \quad \text{et}$$

$$\sqrt{n}\left(\widehat{\theta}_n(\widehat{m}_{BIC}) - \theta(m^*)\right) \underset{n\to\infty}{\overset{\mathcal{L}}{\longrightarrow}} \mathcal{N}_{|m^*|}\left(0\,,\, F(\theta(m^*))^{-1} G(\theta(m^*)) F(\theta(m^*))^{-1}\right)$$

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 $(X_1, \cdots, X_n)$ trajectoire observée issue de m^* et sous certaines conditions,

$$\mathbb{P}(\widehat{m}_{BIC} = m^*) \underset{n\to+\infty}{\longrightarrow} 1 \quad et \quad \mathbb{P}(m^* \subset \widehat{m}_{AIC}) \underset{n\to+\infty}{\longrightarrow} 1 \quad et$$

$$\sqrt{n}\left(\widehat{\theta}_n(\widehat{m}_{BIC}) - \theta(m^*)\right) \underset{n\to\infty}{\overset{\mathcal{L}}{\longrightarrow}} \mathcal{N}_{|m^*|}\left(0\,,\; F(\theta(m^*))^{-1} G(\theta(m^*)) F(\theta(m^*))^{-1}\right)$$

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 $(X_1, \cdots, X_n)$ trajectoire observée issue de m^* et sous certaines conditions,

$$\mathbb{P}(\widehat{m}_{BIC} = m^*) \underset{n\to+\infty}{\longrightarrow} 1 \quad \text{et} \quad \mathbb{P}(m^* \subset \widehat{m}_{AIC}) \underset{n\to+\infty}{\longrightarrow} 1 \quad \text{et}$$

$$\sqrt{n}\,\big(\widehat{\theta}_n(\widehat{m}_{BIC}) - \theta(m^*)\big) \underset{n\to\infty}{\xrightarrow{\mathcal{L}}} \mathcal{N}_{|m^*|}\big(0\,,\ F(\theta(m^*))^{-1} G(\theta(m^*)) F(\theta(m^*))^{-1}\big)$$

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 $(X_1, \cdots, X_n)$ trajectoire observée issue de m^* et sous certaines conditions,

$$\mathbb{P}(\widehat{m}_{BIC} = m^*) \underset{n\to+\infty}{\longrightarrow} 1 \quad et \quad \mathbb{P}(m^* \subset \widehat{m}_{AIC}) \underset{n\to+\infty}{\longrightarrow} 1 \quad et$$

$$\sqrt{n}\,\big(\widehat{\theta}_n(\widehat{m}_{BIC}) - \theta(m^*)\big) \underset{n\to\infty}{\overset{\mathcal{L}}{\longrightarrow}} \mathcal{N}_{|m^*|}\big(0\,,\ F(\theta(m^*))^{-1}G(\theta(m^*))F(\theta(m^*))^{-1}\big)$$

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 $(X_1, \cdots, X_n)$ trajectoire observée issue de m^* et sous certaines conditions,

$$\mathbb{P}(\widehat{m}_{BIC} = m^*) \underset{n \to +\infty}{\longrightarrow} 1 \quad \text{et} \quad \mathbb{P}(m^* \subset \widehat{m}_{AIC}) \underset{n \to +\infty}{\longrightarrow} 1 \quad \text{et}$$

$$\sqrt{n}\left(\widehat{\theta}_n(\widehat{m}_{BIC}) - \theta(m^*)\right) \underset{n \to \infty}{\xrightarrow{\mathcal{L}}} \mathcal{N}_{|m^*|}\left(0\,,\, F(\theta(m^*))^{-1} G(\theta(m^*)) F(\theta(m^*))^{-1}\right)$$

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 $\widehat{m}_{BIC}$ obtenu pour modéliser $X_1, \ldots, X_n)$ comment s'assurer que ce modèle convient ?

$\Longrightarrow$ Test d'adéquation de type portemanteau

$$\left\{ \begin{array}{l} H_0 : \ \exists m^* \in \mathcal{M}, \text{ tel que } (X_1, \ldots, X_n) \text{ trajectoire de } X \text{ avec } \theta(m^*) \\ \\ H_1 : \ \nexists m^* \in \mathcal{M}, \text{ tel que } (X_1, \ldots, X_n) \text{ trajectoire de } X \text{ avec } \theta(m^*) \end{array} \right. .$$

Test d'adéquation

Problématique : Une fois un modèle $\widehat{m}_{BIC}$ obtenu pour modéliser $X_1, \ldots, X_n$) comment s'assurer que ce modèle convient ?

$\Longrightarrow$ Test d'adéquation de type portemanteau

$$\begin{cases} H_0 : \; \exists m^* \in \mathcal{M}, \text{ tel que } (X_1, \ldots, X_n) \text{ trajectoire de } X \text{ avec } \theta(m^*) \\ \\ H_1 : \; \nexists m^* \in \mathcal{M}, \text{ tel que } (X_1, \ldots, X_n) \text{ trajectoire de } X \text{ avec } \theta(m^*) \end{cases} .$$

Test d'adéquation

Problématique : Une fois un modèle $\widehat{m}_{BIC}$ obtenu pour modéliser $X_1, \ldots, X_n)$ comment s'assurer que ce modèle convient ?

$\Longrightarrow$ Test d'adéquation de type portemanteau

$$\begin{cases} H_0 : \ \exists m^* \in \mathcal{M}, \text{ tel que } (X_1, \ldots, X_n) \text{ trajectoire de } X \text{ avec } \theta(m^*) \\ \\ H_1 : \ \nexists m^* \in \mathcal{M}, \text{ tel que } (X_1, \ldots, X_n) \text{ trajectoire de } X \text{ avec } \theta(m^*) \end{cases}.$$

Test d'adéquation

Problématique : Une fois un modèle $\widehat{m}_{BIC}$ obtenu pour modéliser $X_1, \ldots, X_n)$ comment s'assurer que ce modèle convient ?

$\Longrightarrow$ Test d'adéquation de type portemanteau

$$\left\{ \begin{array}{ll} H_0 : & \exists m^* \in \mathcal{M}, \text{ tel que } (X_1, \ldots, X_n) \text{ trajectoire de } X \text{ avec } \theta(m^*) \\ & \\ H_1 : & \nexists m^* \in \mathcal{M}, \text{ tel que } (X_1, \ldots, X_n) \text{ trajectoire de } X \text{ avec } \theta(m^*) \end{array} \right. .$$

Test d'adéquation (2)

On définit les résidus pour le modèle m :

$$\widehat{e}_t(m) := \left(\widehat{M}^t_{\widehat{\theta}(m)}\right)^{-1}\left(X_t - \widehat{f}^t_{\widehat{\theta}(m)}\right).$$

Pour $K \in \mathbb{N}^*$ fixé, vecteur d'autocorrélations des carrés des résidus :

$$\widehat{\rho}(m) := \left(\widehat{\rho}_1(m), \ldots, \widehat{\rho}_K(m)\right)',$$

où $\widehat{\rho}_k(m) := \dfrac{\widehat{\gamma}_k(m)}{\widehat{\gamma}_0(m)}$ et $\widehat{\gamma}_k(m) := \dfrac{1}{n}\displaystyle\sum_{t=k+1}^{n}\left(\widehat{e}_t^2(m) - 1\right)\left(\widehat{e}_{t-k}^2(m) - 1\right)$

Théorème

Sous certaines hypothèses, il existe une matrice $V(\theta(m^*))$

$$\sqrt{n}\,\widehat{\rho}(\widehat{m}_{BIC}) \xrightarrow[n\to\infty]{\mathcal{L}} \mathcal{N}_K\left(0\,,\, V(\theta(m^*))\right).$$

Test d'adéquation (2)

On définit les résidus pour le modèle m :

$$\widehat{e}_t(m) := \left(\widehat{M}^t_{\widehat{\theta}(m)}\right)^{-1}\left(X_t - \widehat{f}^t_{\widehat{\theta}(m)}\right).$$

Pour $K \in \mathbb{N}^*$ fixé, vecteur d'autocorrélations des carrés des résidus :

$$\widehat{\rho}(m) := \left(\widehat{\rho}_1(m), \ldots, \widehat{\rho}_K(m)\right)',$$

où $\widehat{\rho}_k(m) := \dfrac{\widehat{\gamma}_k(m)}{\widehat{\gamma}_0(m)}$ et $\widehat{\gamma}_k(m) := \dfrac{1}{n}\displaystyle\sum_{t=k+1}^{n}\left(\widehat{e}_t^2(m)-1\right)\left(\widehat{e}_{t-k}^2(m)-1\right)$

Théorème

Sous certaines hypothèses, il existe une matrice $V(\theta(m^*))$

$$\sqrt{n}\,\widehat{\rho}(\widehat{m}_{BIC}) \xrightarrow[n\to\infty]{\mathcal{L}} \mathcal{N}_K\left(0\,,\,V(\theta(m^*))\right).$$

Test d'adéquation (2)

On définit les résidus pour le modèle m :

$$\widehat{e}_t(m) := \left(\widehat{M}^t_{\widehat{\theta}(m)}\right)^{-1}\left(X_t - \widehat{f}^t_{\widehat{\theta}(m)}\right).$$

Pour $K \in \mathbb{N}^*$ fixé, vecteur d'autocorrélations des carrés des résidus :

$$\widehat{\rho}(m) := \left(\widehat{\rho}_1(m), \dots, \widehat{\rho}_K(m)\right)',$$

où $\widehat{\rho}_k(m) := \dfrac{\widehat{\gamma}_k(m)}{\widehat{\gamma}_0(m)}$ et $\widehat{\gamma}_k(m) := \dfrac{1}{n} \displaystyle\sum_{t=k+1}^{n} \left(\widehat{e}^2_t(m) - 1\right)\left(\widehat{e}^2_{t-k}(m) - 1\right)$

Théorème

Sous certaines hypothèses, il existe une matrice $V(\theta(m^*))$

$$\sqrt{n}\,\widehat{\rho}(\widehat{m}_{BIC}) \xrightarrow[n\to\infty]{\mathcal{L}} \mathcal{N}_K\left(0\,,\,V(\theta(m^*))\right).$$

Test d'adéquation (2)

On définit les résidus pour le modèle m :

$$\widehat{e}_t(m) := \left(\widehat{M}^t_{\widehat{\theta}(m)}\right)^{-1}\left(X_t - \widehat{f}^t_{\widehat{\theta}(m)}\right).$$

Pour $K \in \mathbb{N}^*$ fixé, vecteur d'autocorrélations des carrés des résidus :

$$\widehat{\rho}(m) := \left(\widehat{\rho}_1(m), \ldots, \widehat{\rho}_K(m)\right)',$$

où $\widehat{\rho}_k(m) := \dfrac{\widehat{\gamma}_k(m)}{\widehat{\gamma}_0(m)}$ et $\widehat{\gamma}_k(m) := \dfrac{1}{n}\displaystyle\sum_{t=k+1}^{n}\left(\widehat{e}^2_t(m) - 1\right)\left(\widehat{e}^2_{t-k}(m) - 1\right)$

Théorème

Sous certaines hypothèses, il existe une matrice $V(\theta(m^*))$

$$\sqrt{n}\,\widehat{\rho}(\widehat{m}_{BIC}) \xrightarrow[n\to\infty]{\mathcal{L}} \mathcal{N}_K\left(0\,,\,V(\theta(m^*))\right).$$

Test d'adéquation (2)

On définit les résidus pour le modèle m :

$$\widehat{e}_t(m) := \left(\widehat{M}^t_{\widehat{\theta}(m)}\right)^{-1}\left(X_t - \widehat{f}^t_{\widehat{\theta}(m)}\right).$$

Pour $K \in \mathbb{N}^*$ fixé, vecteur d'autocorrélations des carrés des résidus :

$$\widehat{\rho}(m) := \left(\widehat{\rho}_1(m), \ldots, \widehat{\rho}_K(m)\right)',$$

où $\widehat{\rho}_k(m) := \dfrac{\widehat{\gamma}_k(m)}{\widehat{\gamma}_0(m)}$ et $\widehat{\gamma}_k(m) := \dfrac{1}{n} \displaystyle\sum_{t=k+1}^{n} \left(\widehat{e}_t^2(m) - 1\right)\left(\widehat{e}_{t-k}^2(m) - 1\right)$

Théorème

Sous certaines hypothèses, il existe une matrice $V(\theta(m^*))$

$$\sqrt{n}\,\widehat{\rho}(\widehat{m}_{BIC}) \xrightarrow[n\to\infty]{\mathcal{L}} \mathcal{N}_K\left(0\,,\, V(\theta(m^*))\right).$$

Test d'adéquation (2)

On définit les résidus pour le modèle m :

$$\widehat{e}_t(m) := \left(\widehat{M}^t_{\widehat{\theta}(m)}\right)^{-1}\left(X_t - \widehat{f}^t_{\widehat{\theta}(m)}\right).$$

Pour $K \in \mathbb{N}^*$ fixé, vecteur d'autocorrélations des carrés des résidus :

$$\widehat{\rho}(m) := \left(\widehat{\rho}_1(m), \ldots, \widehat{\rho}_K(m)\right)',$$

où $\widehat{\rho}_k(m) := \dfrac{\widehat{\gamma}_k(m)}{\widehat{\gamma}_0(m)}$ et $\widehat{\gamma}_k(m) := \dfrac{1}{n}\displaystyle\sum_{t=k+1}^{n}\left(\widehat{e}^2_t(m) - 1\right)\left(\widehat{e}^2_{t-k}(m) - 1\right)$

Théorème

Sous certaines hypothèses, il existe une matrice $V(\theta(m^*))$

$$\sqrt{n}\,\widehat{\rho}(\widehat{m}_{BIC}) \xrightarrow[n\to\infty]{\mathcal{L}} \mathcal{N}_K\left(0\,,\,V(\theta(m^*))\right).$$

Test d'adéquation (2)

On définit les résidus pour le modèle m :

$$\widehat{e}_t(m) := \left(\widehat{M}^t_{\widehat{\theta}(m)}\right)^{-1}\left(X_t - \widehat{f}^t_{\widehat{\theta}(m)}\right).$$

Pour $K \in \mathbb{N}^*$ fixé, vecteur d'autocorrélations des carrés des résidus :

$$\widehat{\rho}(m) := \left(\widehat{\rho}_1(m), \ldots, \widehat{\rho}_K(m)\right)',$$

où $\widehat{\rho}_k(m) := \dfrac{\widehat{\gamma}_k(m)}{\widehat{\gamma}_0(m)}$ et $\widehat{\gamma}_k(m) := \dfrac{1}{n} \displaystyle\sum_{t=k+1}^{n} \left(\widehat{e}^2_t(m) - 1\right)\left(\widehat{e}^2_{t-k}(m) - 1\right)$

Théorème

Sous certaines hypothèses, il existe une matrice $V(\theta(m^*))$

$$\sqrt{n}\,\widehat{\rho}(\widehat{m}_{BIC}) \xrightarrow[n\to\infty]{\mathcal{L}} \mathcal{N}_K\left(0\,,\ V(\theta(m^*))\right).$$

Test d'adéquation (3)

On en déduit la statistique de test :

$$\widehat{Q}_K(\widehat{m}_{BIC}) := n\, {}^t\widehat{\rho}(\widehat{m}_{BIC})\,\widehat{V}^{-1}\,\widehat{\rho}(\widehat{m}_{BIC})$$

où
$$\begin{cases} \widehat{V} = I_K + \frac{1}{(\widehat{\mu}_4-1)^2}\,\widehat{J}_K\,F(\theta(\widehat{m}_{BIC}))^{-1}\Big(G(\theta(\widehat{m}_{BIC}))F(\theta(\widehat{m}_{BIC}))^{-1} + (\widehat{\mu}_4-1)\,I_K\Big)\,{}^t\widehat{J}_K \\ \widehat{J}_K = \frac{1}{n}\sum_{t=1}^{n-i}\left(\widehat{e}_t^2(\widehat{m}_{BIC})-1\right)\left(\frac{\partial \log\left(M^{t+i}_{\widehat{\theta}(\widehat{m}_{BIC})}\right)}{\partial\theta_j}\right)_{1\le i\le K,j}, \qquad \widehat{\mu}_4 = \frac{1}{n}\sum_{t=1}^{n}\widehat{e}_t^4(\widehat{m}_{BIC}) \end{cases}$$

Théorème

Sous l'hypothèse H_0 et sous certaines conditions, $\widehat{Q}_K(\widehat{m}_{BIC}) \xrightarrow[n\to\infty]{\mathcal{L}} \chi^2(K)$.

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 $n\left\|\widehat{\rho}(\widehat{m}_{BIC})\right\|^2$

• Pour un GARCH, $\widehat{V} \simeq I_K - \frac{1}{4}\,\widehat{J}_K\left(\widehat{G}(\widehat{\theta}(\widehat{m}_{BIC}))\right)^{-1}\,{}^t\widehat{J}_K$

Test d'adéquation (3)

On en déduit la statistique de test :

$$\widehat{Q}_K(\widehat{m}_{BIC}) := n\,{}^t\widehat{\rho}(\widehat{m}_{BIC})\,\widehat{V}^{-1}\,\widehat{\rho}(\widehat{m}_{BIC})$$

où $\begin{cases} \widehat{V} = I_K + \frac{1}{(\widehat{\mu}_4-1)^2}\,\widehat{J}_K\,F(\theta(\widehat{m}_{BIC}))^{-1}\Big(G(\theta(\widehat{m}_{BIC}))F(\theta(\widehat{m}_{BIC}))^{-1} + (\widehat{\mu}_4-1)\,I_K\Big)^t\widehat{J}_K \\ \widehat{J}_K = \frac{1}{n}\sum_{t=1}^{n-i}\big(\widehat{e}_t^2(\widehat{m}_{BIC})-1\big)\Big(\frac{\partial \log\big(M^{t+i}_{\widehat{\theta}(\widehat{m}_{BIC})}\big)}{\partial\theta_j}\Big)_{1\leq i\leq K,j}, \qquad \widehat{\mu}_4 = \frac{1}{n}\sum_{t=1}^{n}\widehat{e}_t^4(\widehat{m}_{BIC}) \end{cases}$

Théorème

Sous l'hypothèse H_0 et sous certaines conditions, $\widehat{Q}_K(\widehat{m}_{BIC}) \xrightarrow[n\to\infty]{\mathcal{L}} \chi^2(K)$.

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 $n\left\|\widehat{\rho}(\widehat{m}_{BIC})\right\|^2$

• Pour un GARCH, $\widehat{V} \simeq I_K - \frac{1}{4}\,\widehat{J}_K\left(\widehat{G}(\widehat{\theta}(\widehat{m}_{BIC}))\right)^{-1}{}^t\widehat{J}_K$

Test d'adéquation (3)

On en déduit la statistique de test :

$$\widehat{Q}_K(\widehat{m}_{BIC}) := n \, {}^t\widehat{\rho}(\widehat{m}_{BIC}) \, \widehat{V}^{-1} \, \widehat{\rho}(\widehat{m}_{BIC})$$

où
$$\begin{cases} \widehat{V} = I_K + \frac{1}{(\widehat{\mu}_4-1)^2} \widehat{J}_K \, F(\theta(\widehat{m}_{BIC}))^{-1} \Big(G(\theta(\widehat{m}_{BIC})) F(\theta(\widehat{m}_{BIC}))^{-1} + (\widehat{\mu}_4 - 1) I_K \Big) {}^t\widehat{J}_K \\ \widehat{J}_K = \frac{1}{n} \sum_{t=1}^{n-i} \left(\widehat{e}_t^2(\widehat{m}_{BIC}) - 1\right) \left(\frac{\partial \log \left(M^{t+i}_{\widehat{\theta}(\widehat{m}_{BIC})} \right)}{\partial \theta_j} \right)_{1 \leq i \leq K, j}, \qquad \widehat{\mu}_4 = \frac{1}{n} \sum_{t=1}^{n} \widehat{e}_t^4(\widehat{m}_{BIC}) \end{cases}$$

Théorème

Sous l'hypothèse H_0 et sous certaines conditions, $\widehat{Q}_K(\widehat{m}_{BIC}) \xrightarrow[n \to \infty]{\mathcal{L}} \chi^2(K)$.

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 $n \left\| \widehat{\rho}(\widehat{m}_{BIC}) \right\|^2$

• Pour un GARCH, $\widehat{V} \simeq I_K - \frac{1}{4} \widehat{J}_K \left(\widehat{G}(\widehat{\theta}(\widehat{m}_{BIC})) \right)^{-1} {}^t\widehat{J}_K$

Test d'adéquation (3)

On en déduit la statistique de test :

$$\widehat{Q}_K(\widehat{m}_{BIC}) := n\,{}^t\widehat{\rho}(\widehat{m}_{BIC})\,\widehat{V}^{-1}\,\widehat{\rho}(\widehat{m}_{BIC})$$

où
$$\begin{cases} \widehat{V} = I_K + \frac{1}{(\widehat{\mu}_4-1)^2}\,\widehat{J}_K\,F(\theta(\widehat{m}_{BIC}))^{-1}\Big(G(\theta(\widehat{m}_{BIC}))F(\theta(\widehat{m}_{BIC}))^{-1} + (\widehat{\mu}_4-1)\,I_K\Big)\,{}^t\widehat{J}_K \\ \widehat{J}_K = \frac{1}{n}\sum_{t=1}^{n-i}\left(\widehat{e}_t^2(\widehat{m}_{BIC})-1\right)\left(\frac{\partial \log\left(M^{t+i}_{\widehat{\theta}(\widehat{m}_{BIC})}\right)}{\partial\theta_j}\right)_{1\le i\le K,j}, \qquad \widehat{\mu}_4 = \frac{1}{n}\sum_{t=1}^{n}\widehat{e}_t^4(\widehat{m}_{BIC}) \end{cases}$$

Théorème

Sous l'hypothèse H_0 et sous certaines conditions, $\widehat{Q}_K(\widehat{m}_{BIC}) \xrightarrow[n\to\infty]{\mathcal{L}} \chi^2(K)$.

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 $n\left\|\widehat{\rho}(\widehat{m}_{BIC})\right\|^2$

• Pour un GARCH, $\widehat{V} \simeq I_K - \frac{1}{4}\,\widehat{J}_K\left(\widehat{G}(\widehat{\theta}(\widehat{m}_{BIC}))\right)^{-1}{}^t\widehat{J}_K$

Test d'adéquation (3)

On en déduit la statistique de test :

$$\widehat{Q}_K(\widehat{m}_{BIC}) := n\,{}^t\widehat{\rho}(\widehat{m}_{BIC})\,\widehat{V}^{-1}\,\widehat{\rho}(\widehat{m}_{BIC})$$

où
$$\begin{cases} \widehat{V} = I_K + \frac{1}{(\widehat{\mu}_4-1)^2}\,\widehat{J}_K\,F(\theta(\widehat{m}_{BIC}))^{-1}\Big(G(\theta(\widehat{m}_{BIC}))F(\theta(\widehat{m}_{BIC}))^{-1} + (\widehat{\mu}_4-1)\,I_K\Big)\,{}^t\widehat{J}_K \\ \widehat{J}_K = \frac{1}{n}\sum_{t=1}^{n-i}\left(\widehat{e}_t^2(\widehat{m}_{BIC})-1\right)\left(\frac{\partial \log\left(M_{\widehat{\theta}(\widehat{m}_{BIC})}^{t+i}\right)}{\partial\theta_j}\right)_{1\leq i\leq K,j}, \qquad \widehat{\mu}_4 = \frac{1}{n}\sum_{t=1}^{n}\widehat{e}_t^4(\widehat{m}_{BIC}) \end{cases}$$

Théorème

Sous l'hypothèse H_0 et sous certaines conditions, $\widehat{Q}_K(\widehat{m}_{BIC}) \xrightarrow[n\to\infty]{\mathcal{L}} \chi^2(K)$.

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 $n\left\|\widehat{\rho}(\widehat{m}_{BIC})\right\|^2$

• Pour un GARCH, $\widehat{V} \simeq I_K - \frac{1}{4}\,\widehat{J}_K\left(\widehat{G}(\widehat{\theta}(\widehat{m}_{BIC}))\right)^{-1}\,{}^t\widehat{J}_K$

Test d'adéquation (3)

On en déduit la statistique de test :

$$\widehat{Q}_K(\widehat{m}_{BIC}) := n\,{}^t\widehat{\rho}(\widehat{m}_{BIC})\,\widehat{V}^{-1}\,\widehat{\rho}(\widehat{m}_{BIC})$$

où
$$\begin{cases} \widehat{V} = I_K + \frac{1}{(\widehat{\mu}_4-1)^2}\,\widehat{J}_K\,F(\theta(\widehat{m}_{BIC}))^{-1}\Big(G(\theta(\widehat{m}_{BIC}))F(\theta(\widehat{m}_{BIC}))^{-1} + (\widehat{\mu}_4-1)\,I_K\Big)\,{}^t\widehat{J}_K \\ \widehat{J}_K = \frac{1}{n}\sum_{t=1}^{n-i}\left(\widehat{e}_t^2(\widehat{m}_{BIC})-1\right)\left(\frac{\partial\log\left(M^{t+i}_{\widehat{\theta}(\widehat{m}_{BIC})}\right)}{\partial\theta_j}\right)_{1\le i\le K,j}, \qquad \widehat{\mu}_4 = \frac{1}{n}\sum_{t=1}^{n}\widehat{e}_t^4(\widehat{m}_{BIC}) \end{cases}$$

Théorème

Sous l'hypothèse H_0 et sous certaines conditions, $\widehat{Q}_K(\widehat{m}_{BIC}) \xrightarrow[n\to\infty]{\mathcal{L}} \chi^2(K)$.

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 $n\left\|\widehat{\rho}(\widehat{m}_{BIC})\right\|^2$

• Pour un GARCH, $\widehat{V} \simeq I_K - \frac{1}{4}\widehat{J}_K\left(\widehat{G}(\widehat{\theta}(\widehat{m}_{BIC}))\right)^{-1}\,{}^t\widehat{J}_K$

Test d'adéquation (3)

On en déduit la statistique de test :

$$\widehat{Q}_K(\widehat{m}_{BIC}) := n\,{}^t\widehat{\rho}(\widehat{m}_{BIC})\,\widehat{V}^{-1}\,\widehat{\rho}(\widehat{m}_{BIC})$$

où $\begin{cases} \widehat{V} = I_K + \frac{1}{(\widehat{\mu}_4-1)^2}\,\widehat{J}_K\,F(\theta(\widehat{m}_{BIC}))^{-1}\Big(G(\theta(\widehat{m}_{BIC}))F(\theta(\widehat{m}_{BIC}))^{-1} + (\widehat{\mu}_4-1)\,I_K\Big)\,{}^t\widehat{J}_K \\ \widehat{J}_K = \frac{1}{n}\sum_{t=1}^{n-i}\left(\widehat{e}_t^2(\widehat{m}_{BIC})-1\right)\left(\frac{\partial \log\left(M^{t+i}_{\widehat{\theta}(\widehat{m}_{BIC})}\right)}{\partial\theta_j}\right)_{1\leq i\leq K,j}, \qquad \widehat{\mu}_4 = \frac{1}{n}\sum_{t=1}^{n}\widehat{e}_t^4(\widehat{m}_{BIC}) \end{cases}$

Théorème

Sous l'hypothèse H_0 et sous certaines conditions, $\widehat{Q}_K(\widehat{m}_{BIC}) \xrightarrow[n\to\infty]{\mathcal{L}} \chi^2(K)$.

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 $n\left\|\widehat{\rho}(\widehat{m}_{BIC})\right\|^2$

• Pour un GARCH, $\widehat{V} \simeq I_K - \frac{1}{4}\,\widehat{J}_K\left(\widehat{G}(\widehat{\theta}(\widehat{m}_{BIC}))\right)^{-1}\,{}^t\widehat{J}_K$

Test d'adéquation (3)

On en déduit la statistique de test :

$$\widehat{Q}_K(\widehat{m}_{BIC}) := n\,{}^t\widehat{\rho}(\widehat{m}_{BIC})\,\widehat{V}^{-1}\,\widehat{\rho}(\widehat{m}_{BIC})$$

où
$$\begin{cases} \widehat{V} = I_K + \frac{1}{(\widehat{\mu}_4-1)^2}\,\widehat{J}_K\,F(\theta(\widehat{m}_{BIC}))^{-1}\Big(G(\theta(\widehat{m}_{BIC}))F(\theta(\widehat{m}_{BIC}))^{-1} + (\widehat{\mu}_4 - 1)\,I_K\Big)\,{}^t\widehat{J}_K \\ \widehat{J}_K = \frac{1}{n}\sum_{t=1}^{n-i}\left(\widehat{e}_t^2(\widehat{m}_{BIC})-1\right)\left(\frac{\partial \log\left(M^{t+i}_{\widehat{\theta}(\widehat{m}_{BIC})}\right)}{\partial\theta_j}\right)_{1\le i\le K,j}, \qquad \widehat{\mu}_4 = \frac{1}{n}\sum_{t=1}^{n}\widehat{e}_t^4(\widehat{m}_{BIC}) \end{cases}$$

Théorème

Sous l'hypothèse H_0 et sous certaines conditions, $\widehat{Q}_K(\widehat{m}_{BIC}) \xrightarrow[n\to\infty]{\mathcal{L}} \chi^2(K)$.

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 $n\left\|\widehat{\rho}(\widehat{m}_{BIC})\right\|^2$

• Pour un GARCH, $\widehat{V} \simeq I_K - \frac{1}{4}\,\widehat{J}_K\left(\widehat{G}(\widehat{\theta}(\widehat{m}_{BIC}))\right)^{-1}\,{}^t\widehat{J}_K$

Plan du cours

Prédiction à partir d'un modèle

On suppose observée $(X_1, \ldots, X_n)$.

Objectif : On veut prédire X_{n+h}, soit $\widehat{X}_{n+h}$ (h horizon)

Méthode 1 : On suppose que $X_t = a(t) + s(t) + u_t$ pour $t \in \mathbb{Z}$ et :

- On a obtenu $\widehat{a}_n(t)$ et $\widehat{s}_n(t)$ à partir de $(X_1, \ldots, X_n)$
- Pour $\widehat{u}_t = X_t - \widehat{a}_n(t) - \widehat{s}_n(t)$ on a obtenu et validé un modèle

$$\widehat{u}_t = M_{\widehat{\theta}}(\widehat{u}_{t-1}, \widehat{u}_{t-2}, \ldots, \widehat{u}_1, 0, \ldots)\,\xi_t + f_{\widehat{\theta}}(\widehat{u}_{t-1}, \widehat{u}_{t-2}, \ldots, \widehat{u}_1, 0, \ldots)$$

Prédiction à partir d'un modèle

On suppose observée $(X_1, \dots, X_n)$.

Objectif : On veut prédire X_{n+h}, soit $\widehat{X}_{n+h}$ (h horizon)

Méthode 1 : On suppose que $X_t = a(t) + s(t) + u_t$ pour $t \in \mathbb{Z}$ et :

- On a obtenu $\widehat{a}_n(t)$ et $\widehat{s}_n(t)$ à partir de $(X_1, \dots, X_n)$
- Pour $\widehat{u}_t = X_t - \widehat{a}_n(t) - \widehat{s}_n(t)$ on a obtenu et validé un modèle

$$\widehat{u}_t = M_{\widehat{\theta}}(\widehat{u}_{t-1}, \widehat{u}_{t-2}, \dots, \widehat{u}_1, 0, \dots)\,\xi_t + f_{\widehat{\theta}}(\widehat{u}_{t-1}, \widehat{u}_{t-2}, \dots, \widehat{u}_1, 0, \dots)$$

Prédiction à partir d'un modèle

On suppose observée $(X_1, \ldots, X_n)$.

Objectif : On veut prédire X_{n+h}, soit $\widehat{X}_{n+h}$ (h horizon)

Méthode 1 : On suppose que $X_t = a(t) + s(t) + u_t$ pour $t \in \mathbb{Z}$ et :

- On a obtenu $\widehat{a}_n(t)$ et $\widehat{s}_n(t)$ à partir de $(X_1, \ldots, X_n)$
- Pour $\widehat{u}_t = X_t - \widehat{a}_n(t) - \widehat{s}_n(t)$ on a obtenu et validé un modèle

$$\widehat{u}_t = M_{\widehat{\theta}}(\widehat{u}_{t-1}, \widehat{u}_{t-2}, \ldots, \widehat{u}_1, 0, \ldots)\,\xi_t + f_{\widehat{\theta}}(\widehat{u}_{t-1}, \widehat{u}_{t-2}, \ldots, \widehat{u}_1, 0, \ldots)$$

Prédiction à partir d'un modèle

On suppose observée $(X_1, \ldots, X_n)$.

Objectif : On veut prédire X_{n+h}, soit $\widehat{X}_{n+h}$ (h horizon)

Méthode 1 : On suppose que $X_t = a(t) + s(t) + u_t$ pour $t \in \mathbb{Z}$ et :

- On a obtenu $\widehat{a}_n(t)$ et $\widehat{s}_n(t)$ à partir de $(X_1, \ldots, X_n)$
- Pour $\widehat{u}_t = X_t - \widehat{a}_n(t) - \widehat{s}_n(t)$ on a obtenu et validé un modèle

$$\widehat{u}_t = M_{\widehat{\theta}}(\widehat{u}_{t-1}, \widehat{u}_{t-2}, \ldots, \widehat{u}_1, 0, \ldots)\, \xi_t + f_{\widehat{\theta}}(\widehat{u}_{t-1}, \widehat{u}_{t-2}, \ldots, \widehat{u}_1, 0, \ldots)$$

Prédiction à partir d'un modèle

On suppose observée $(X_1, \ldots, X_n)$.

Objectif : On veut prédire X_{n+h}, soit $\widehat{X}_{n+h}$ (h horizon)

Méthode 1 : On suppose que $X_t = a(t) + s(t) + u_t$ pour $t \in \mathbb{Z}$ et :

- On a obtenu $\widehat{a}_n(t)$ et $\widehat{s}_n(t)$ à partir de $(X_1, \ldots, X_n)$
- Pour $\widehat{u}_t = X_t - \widehat{a}_n(t) - \widehat{s}_n(t)$ on a obtenu et validé un modèle

$$\widehat{u}_t = M_{\widehat{\theta}}(\widehat{u}_{t-1}, \widehat{u}_{t-2}, \ldots, \widehat{u}_1, 0, \ldots)\, \xi_t + f_{\widehat{\theta}}(\widehat{u}_{t-1}, \widehat{u}_{t-2}, \ldots, \widehat{u}_1, 0, \ldots)$$

Prédiction à partir d'un modèle

On suppose observée $(X_1, \ldots, X_n)$.

Objectif : On veut prédire X_{n+h}, soit $\widehat{X}_{n+h}$ (h horizon)

Méthode 1 : On suppose que $X_t = a(t) + s(t) + u_t$ pour $t \in \mathbb{Z}$ et :

- On a obtenu $\widehat{a}_n(t)$ et $\widehat{s}_n(t)$ à partir de $(X_1, \ldots, X_n)$
- Pour $\widehat{u}_t = X_t - \widehat{a}_n(t) - \widehat{s}_n(t)$ on a obtenu et validé un modèle

$$\widehat{u}_t = M_{\widehat{\theta}}(\widehat{u}_{t-1}, \widehat{u}_{t-2}, \ldots, \widehat{u}_1, 0, \ldots)\,\xi_t + f_{\widehat{\theta}}(\widehat{u}_{t-1}, \widehat{u}_{t-2}, \ldots, \widehat{u}_1, 0, \ldots)$$

Prédiction à partir d'un modèle

On suppose observée $(X_1, \ldots, X_n)$.

Objectif : On veut prédire X_{n+h}, soit $\widehat{X}_{n+h}$ (h horizon)

Méthode 1 : On suppose que $X_t = a(t) + s(t) + u_t$ pour $t \in \mathbb{Z}$ et :

- On a obtenu $\widehat{a}_n(t)$ et $\widehat{s}_n(t)$ à partir de $(X_1, \ldots, X_n)$
- Pour $\widehat{u}_t = X_t - \widehat{a}_n(t) - \widehat{s}_n(t)$ on a obtenu et validé un modèle

$$\widehat{u}_t = M_{\widehat{\theta}}(\widehat{u}_{t-1}, \widehat{u}_{t-2}, \ldots, \widehat{u}_1, 0, \ldots)\,\xi_t + f_{\widehat{\theta}}(\widehat{u}_{t-1}, \widehat{u}_{t-2}, \ldots, \widehat{u}_1, 0, \ldots)$$

Prédiction à partir d'un modèle (2)

Procédure : On veut écrire $\widehat{X}_{n+h} = g(X_1, \ldots, X_n)$ et on cherche

$$\widehat{g} = \text{Arg}\min_{g \in \mathbb{L}^2} \mathbb{E}\left[\left(X_{n+h} - g(X_1, \ldots, X_n)\right)^2\right]$$

$$\Longrightarrow \widehat{X}_{n+h} = \mathbb{E}\left[X_{n+h} \mid (X_1, \ldots, X_n)\right]$$

$$\Longrightarrow \boldsymbol{\widehat{X}_{n+h} \simeq \widehat{a}_n(n+h) + \widehat{s}_n(n+h) + \mathbb{E}\left[\widehat{u}_{n+h} \mid (\widehat{u}_1, \ldots, \widehat{u}_n)\right]}$$

Prédiction à partir d'un modèle (2)

Procédure : On veut écrire $\widehat{X}_{n+h} = g(X_1, \ldots, X_n)$ et on cherche

$$\widehat{g} = \operatorname{Arg}\min_{g \in \mathbb{L}^2} \mathbb{E}\left[\left(X_{n+h} - g(X_1, \ldots, X_n)\right)^2\right]$$

$$\Longrightarrow \widehat{X}_{n+h} = \mathbb{E}\left[X_{n+h} \mid (X_1, \ldots, X_n)\right]$$

$$\Longrightarrow \widehat{X}_{n+h} \simeq \widehat{a}_n(n+h) + \widehat{s}_n(n+h) + \mathbb{E}\left[\widehat{u}_{n+h} \mid (\widehat{u}_1, \ldots, \widehat{u}_n)\right]$$

Prédiction à partir d'un modèle (2)

Procédure : On veut écrire $\widehat{X}_{n+h} = g(X_1, \ldots, X_n)$ et on cherche

$$\widehat{g} = \operatorname{Arg}\min_{g \in \mathbb{L}^2} \mathbb{E}\left[\left(X_{n+h} - g(X_1, \ldots, X_n)\right)^2\right]$$

$$\Longrightarrow \widehat{X}_{n+h} = \mathbb{E}\left[X_{n+h} \mid (X_1, \ldots, X_n)\right]$$

$$\Longrightarrow \widehat{X}_{n+h} \simeq \widehat{a}_n(n+h) + \widehat{s}_n(n+h) + \mathbb{E}\left[\widehat{u}_{n+h} \mid (\widehat{u}_1, \ldots, \widehat{u}_n)\right]$$

Prédiction à partir d'un modèle (2)

Procédure : On veut écrire $\widehat{X}_{n+h} = g(X_1, \ldots, X_n)$ et on cherche

$$\widehat{g} = \text{Arg}\min_{g \in \mathbb{L}^2} \mathbb{E}\left[\left(X_{n+h} - g(X_1, \ldots, X_n)\right)^2\right]$$

$$\Longrightarrow \widehat{X}_{n+h} = \mathbb{E}\left[X_{n+h} \mid (X_1, \ldots, X_n)\right]$$

$$\Longrightarrow \boldsymbol{\widehat{X}_{n+h} \simeq \widehat{a}_n(n+h) + \widehat{s}_n(n+h) + \mathbb{E}\left[\widehat{u}_{n+h} \mid (\widehat{u}_1, \ldots, \widehat{u}_n)\right]}$$

Prédiction à partir d'un modèle (3)

Exemples : On veut calculer $\mathbb{E}\left[\widehat{u}_{n+h} \mid (\widehat{u}_1, \ldots, \widehat{u}_n)\right]$

1. Si $(\widehat{u}_t)$ AR(1) causal $\widehat{u}_t = \widehat{\alpha}\,\widehat{u}_{t-1} + \xi_t$ d'où :
   - $\widehat{u}_{n+1} = \mathbb{E}\left[\widehat{\alpha}\,\widehat{u}_n + \xi_{n+1} \mid (\widehat{u}_1, \ldots, \widehat{u}_n)\right] = \widehat{\alpha}\,\widehat{u}_n$
   - $\widehat{u}_{n+2} = \mathbb{E}\left[\widehat{\alpha}^2\,\widehat{u}_n + \widehat{\alpha}\,\xi_{n+1} + \xi_{n+2} \mid (\widehat{u}_1, \ldots, \widehat{u}_n)\right] = \widehat{\alpha}^2\,\widehat{u}_n$
   - $\widehat{u}_{n+h} = \mathbb{E}\left[\widehat{\alpha}^h\,\widehat{u}_n + \sum_{i=1}^h \widehat{\alpha}^{h-i}\,\xi_{n+i} \mid (\widehat{u}_1, \ldots, \widehat{u}_n)\right] = \widehat{\alpha}^h\,\widehat{u}_n$
2. Si $(\widehat{u}_t)$ MA(1) inversible $\widehat{u}_t = \xi_t + \widehat{\alpha}\,\xi_{t-1}$, soit
   $$\widehat{u}_t = \xi_t - \sum_{k=1}^{\infty}(-\widehat{\alpha})^k\,\widehat{u}_{t-k}$$
   - $\widehat{u}_{n+1} = \mathbb{E}\left[\xi_{n+1} - \sum_{k=0}^{n-1}(-\widehat{\alpha})^{k+1}\,\widehat{u}_{n-k} \mid (\widehat{u}_1, \ldots, \widehat{u}_n)\right] = -\sum_{k=0}^{n-1}(-\widehat{\alpha})^{k+1}\,\widehat{u}_{n-k}$
   - $\widehat{u}_{n+h} = \mathbb{E}\left[\xi_{n+h} + \widehat{\alpha}\,\xi_{n+h-1} \mid (\widehat{u}_1, \ldots, \widehat{u}_n)\right] = 0$ pour $h \geq 2$
3. Si $(\widehat{u}_t)$ GARCH(p,q) alors $\widehat{u}_t = \widehat{\sigma}_t\,\xi_t$ d'où $\widehat{X}_{n+h} = 0$ si $h \geq 1$

Prédiction à partir d'un modèle (3)

Exemples : On veut calculer $\mathbb{E}\left[\widehat{u}_{n+h} \mid (\widehat{u}_1, \ldots, \widehat{u}_n)\right]$

1. Si $(\widehat{u}_t)$ AR(1) causal $\widehat{u}_t = \widehat{\alpha}\, \widehat{u}_{t-1} + \xi_t$ d'où :
   - $\widehat{u}_{n+1} = \mathbb{E}\left[\widehat{\alpha}\, \widehat{u}_n + \xi_{n+1} \mid (\widehat{u}_1, \ldots, \widehat{u}_n)\right] = \widehat{\alpha}\, \widehat{u}_n$
   - $\widehat{u}_{n+2} = \mathbb{E}\left[\widehat{\alpha}^2\, \widehat{u}_n + \widehat{\alpha}\, \xi_{n+1} + \xi_{n+2} \mid (\widehat{u}_1, \ldots, \widehat{u}_n)\right] = \widehat{\alpha}^2\, \widehat{u}_n$
   - $\widehat{u}_{n+h} = \mathbb{E}\left[\widehat{\alpha}^h\, \widehat{u}_n + \sum_{i=1}^{h} \widehat{\alpha}^{h-i}\, \xi_{n+i} \mid (\widehat{u}_1, \ldots, \widehat{u}_n)\right] = \widehat{\alpha}^h\, \widehat{u}_n$
2. Si $(\widehat{u}_t)$ MA(1) inversible $\widehat{u}_t = \xi_t + \widehat{\alpha}\, \xi_{t-1}$, soit

   $$\widehat{u}_t = \xi_t - \sum_{k=1}^{\infty} (-\widehat{\alpha})^k\, \widehat{u}_{t-k}$$

   - $\widehat{u}_{n+1} = \mathbb{E}\left[\xi_{n+1} - \sum_{k=0}^{n-1} (-\widehat{\alpha})^{k+1}\, \widehat{u}_{n-k} \mid (\widehat{u}_1, \ldots, \widehat{u}_n)\right] = -\sum_{k=0}^{n-1} (-\widehat{\alpha})^{k+1}\, \widehat{u}_{n-k}$
   - $\widehat{u}_{n+h} = \mathbb{E}\left[\xi_{n+h} + \widehat{\alpha}\, \xi_{n+h-1} \mid (\widehat{u}_1, \ldots, \widehat{u}_n)\right] = 0$ pour $h \geq 2$
3. Si $(\widehat{u}_t)$ GARCH(p, q) alors $\widehat{u}_t = \widehat{\sigma}_t\, \xi_t$ d'où $\widehat{X}_{n+h} = 0$ si $h \geq 1$

Prédiction à partir d'un modèle (3)

Exemples : On veut calculer $\mathbb{E}\big[\widehat{u}_{n+h} \mid (\widehat{u}_1, \ldots, \widehat{u}_n)\big]$

1. Si $(\widehat{u}_t)$ AR(1) causal $\widehat{u}_t = \widehat{\alpha}\,\widehat{u}_{t-1} + \xi_t$ d'où :
   - $\widehat{u}_{n+1} = \mathbb{E}\big[\widehat{\alpha}\,\widehat{u}_n + \xi_{n+1} \mid (\widehat{u}_1, \ldots, \widehat{u}_n)\big] = \widehat{\alpha}\,\widehat{u}_n$
   - $\widehat{u}_{n+2} = \mathbb{E}\big[\widehat{\alpha}^2\,\widehat{u}_n + \widehat{\alpha}\,\xi_{n+1} + \xi_{n+2} \mid (\widehat{u}_1, \ldots, \widehat{u}_n)\big] = \widehat{\alpha}^2\,\widehat{u}_n$
   - $\widehat{u}_{n+h} = \mathbb{E}\big[\widehat{\alpha}^h\,\widehat{u}_n + \sum_{i=1}^h \widehat{\alpha}^{h-i}\,\xi_{n+i} \mid (\widehat{u}_1, \ldots, \widehat{u}_n)\big] = \widehat{\alpha}^h\,\widehat{u}_n$
2. Si $(\widehat{u}_t)$ MA(1) inversible $\widehat{u}_t = \xi_t + \widehat{\alpha}\,\xi_{t-1}$, soit

   $$\widehat{u}_t = \xi_t - \sum_{k=1}^{\infty}(-\widehat{\alpha})^k\,\widehat{u}_{t-k}$$

   - $\widehat{u}_{n+1} = \mathbb{E}\Big[\xi_{n+1} - \sum_{k=0}^{n-1}(-\widehat{\alpha})^{k+1}\,\widehat{u}_{n-k} \mid (\widehat{u}_1, \ldots, \widehat{u}_n)\Big] = -\sum_{k=0}^{n-1}(-\widehat{\alpha})^{k+1}\,\widehat{u}_{n-k}$
   - $\widehat{u}_{n+h} = \mathbb{E}\big[\xi_{n+h} + \widehat{\alpha}\,\xi_{n+h-1} \mid (\widehat{u}_1, \ldots, \widehat{u}_n)\big] = 0$ pour $h \geq 2$
3. Si $(\widehat{u}_t)$ GARCH(p,q) alors $\widehat{u}_t = \widehat{\sigma}_t\,\xi_t$ d'où $\widehat{X}_{n+h} = 0$ si $h \geq 1$

Prédiction à partir d'un modèle (3)

Exemples : On veut calculer $\mathbb{E}\big[\widehat{u}_{n+h} \mid (\widehat{u}_1, \ldots, \widehat{u}_n)\big]$

1. Si $(\widehat{u}_t)$ AR(1) causal $\widehat{u}_t = \widehat{\alpha}\, \widehat{u}_{t-1} + \xi_t$ d'où :
   - $\widehat{u}_{n+1} = \mathbb{E}\big[\widehat{\alpha}\, \widehat{u}_n + \xi_{n+1} \mid (\widehat{u}_1, \ldots, \widehat{u}_n)\big] = \widehat{\alpha}\, \widehat{u}_n$
   - $\widehat{u}_{n+2} = \mathbb{E}\big[\widehat{\alpha}^2\, \widehat{u}_n + \widehat{\alpha}\, \xi_{n+1} + \xi_{n+2} \mid (\widehat{u}_1, \ldots, \widehat{u}_n)\big] = \widehat{\alpha}^2\, \widehat{u}_n$
   - $\widehat{u}_{n+h} = \mathbb{E}\big[\widehat{\alpha}^h\, \widehat{u}_n + \sum_{i=1}^{h} \widehat{\alpha}^{h-i}\, \xi_{n+i} \mid (\widehat{u}_1, \ldots, \widehat{u}_n)\big] = \widehat{\alpha}^h\, \widehat{u}_n$
2. Si $(\widehat{u}_t)$ MA(1) inversible $\widehat{u}_t = \xi_t + \widehat{\alpha}\, \xi_{t-1}$, soit

   $$\widehat{u}_t = \xi_t - \sum_{k=1}^{\infty} (-\widehat{\alpha})^k\, \widehat{u}_{t-k}$$

   - $\widehat{u}_{n+1} = \mathbb{E}\big[\xi_{n+1} - \sum_{k=0}^{n-1} (-\widehat{\alpha})^{k+1}\, \widehat{u}_{n-k} \mid (\widehat{u}_1, \ldots, \widehat{u}_n)\big] = -\sum_{k=0}^{n-1} (-\widehat{\alpha})^{k+1}\, \widehat{u}_{n-k}$
   - $\widehat{u}_{n+h} = \mathbb{E}\big[\xi_{n+h} + \widehat{\alpha}\, \xi_{n+h-1} \mid (\widehat{u}_1, \ldots, \widehat{u}_n)\big] = 0$ pour $h \geq 2$
3. Si $(\widehat{u}_t)$ GARCH(p, q) alors $\widehat{u}_t = \widehat{\sigma}_t\, \xi_t$ d'où $\widehat{X}_{n+h} = 0$ si $h \geq 1$

Prédiction à partir d'un modèle (3)

Exemples : On veut calculer $\mathbb{E}\big[\widehat{u}_{n+h} \mid (\widehat{u}_1, \ldots, \widehat{u}_n)\big]$

1. Si $(\widehat{u}_t)$ AR(1) causal $\widehat{u}_t = \widehat{\alpha}\, \widehat{u}_{t-1} + \xi_t$ d'où :
   - $\widehat{u}_{n+1} = \mathbb{E}\big[\widehat{\alpha}\, \widehat{u}_n + \xi_{n+1} \mid (\widehat{u}_1, \ldots, \widehat{u}_n)\big] = \widehat{\alpha}\, \widehat{u}_n$
   - $\widehat{u}_{n+2} = \mathbb{E}\big[\widehat{\alpha}^2\, \widehat{u}_n + \widehat{\alpha}\, \xi_{n+1} + \xi_{n+2} \mid (\widehat{u}_1, \ldots, \widehat{u}_n)\big] = \widehat{\alpha}^2\, \widehat{u}_n$
   - $\widehat{u}_{n+h} = \mathbb{E}\big[\widehat{\alpha}^h\, \widehat{u}_n + \sum_{i=1}^{h} \widehat{\alpha}^{h-i}\, \xi_{n+i} \mid (\widehat{u}_1, \ldots, \widehat{u}_n)\big] = \widehat{\alpha}^h\, \widehat{u}_n$
2. Si $(\widehat{u}_t)$ MA(1) inversible $\widehat{u}_t = \xi_t + \widehat{\alpha}\, \xi_{t-1}$, soit
$$\widehat{u}_t = \xi_t - \sum_{k=1}^{\infty} (-\widehat{\alpha})^k\, \widehat{u}_{t-k}$$
   - $\widehat{u}_{n+1} = \mathbb{E}\big[\xi_{n+1} - \sum_{k=0}^{n-1} (-\widehat{\alpha})^{k+1}\, \widehat{u}_{n-k} \mid (\widehat{u}_1, \ldots, \widehat{u}_n)\big] = -\sum_{k=0}^{n-1} (-\widehat{\alpha})^{k+1}\, \widehat{u}_{n-k}$
   - $\widehat{u}_{n+h} = \mathbb{E}\big[\xi_{n+h} + \widehat{\alpha}\, \xi_{n+h-1} \mid (\widehat{u}_1, \ldots, \widehat{u}_n)\big] = 0$ pour $h \geq 2$
3. Si $(\widehat{u}_t)$ GARCH(p, q) alors $\widehat{u}_t = \widehat{\sigma}_t\, \xi_t$ d'où $\widehat{X}_{n+h} = 0$ si $h \geq 1$

Prédiction à partir d'un modèle (3)

Exemples : On veut calculer $\mathbb{E}\left[\widehat{u}_{n+h} \mid (\widehat{u}_1, \ldots, \widehat{u}_n)\right]$

1. Si $(\widehat{u}_t)$ AR(1) causal $\widehat{u}_t = \widehat{\alpha}\,\widehat{u}_{t-1} + \xi_t$ d'où :
   - $\widehat{u}_{n+1} = \mathbb{E}\left[\widehat{\alpha}\,\widehat{u}_n + \xi_{n+1} \mid (\widehat{u}_1, \ldots, \widehat{u}_n)\right] = \widehat{\alpha}\,\widehat{u}_n$
   - $\widehat{u}_{n+2} = \mathbb{E}\left[\widehat{\alpha}^2\,\widehat{u}_n + \widehat{\alpha}\,\xi_{n+1} + \xi_{n+2} \mid (\widehat{u}_1, \ldots, \widehat{u}_n)\right] = \widehat{\alpha}^2\,\widehat{u}_n$
   - $\widehat{u}_{n+h} = \mathbb{E}\left[\widehat{\alpha}^h\,\widehat{u}_n + \sum_{i=1}^h \widehat{\alpha}^{h-i}\,\xi_{n+i} \mid (\widehat{u}_1, \ldots, \widehat{u}_n)\right] = \widehat{\alpha}^h\,\widehat{u}_n$
2. Si $(\widehat{u}_t)$ MA(1) inversible $\widehat{u}_t = \xi_t + \widehat{\alpha}\,\xi_{t-1}$, soit

$$\widehat{u}_t = \xi_t - \sum_{k=1}^{\infty} (-\widehat{\alpha})^k\,\widehat{u}_{t-k}$$

   - $\widehat{u}_{n+1} = \mathbb{E}\left[\xi_{n+1} - \sum_{k=0}^{n-1} (-\widehat{\alpha})^{k+1}\,\widehat{u}_{n-k} \mid (\widehat{u}_1, \ldots, \widehat{u}_n)\right] = -\sum_{k=0}^{n-1} (-\widehat{\alpha})^{k+1}\,\widehat{u}_{n-k}$
   - $\widehat{u}_{n+h} = \mathbb{E}\left[\xi_{n+h} + \widehat{\alpha}\,\xi_{n+h-1} \mid (\widehat{u}_1, \ldots, \widehat{u}_n)\right] = 0$ pour $h \geq 2$

3. Si $(\widehat{u}_t)$ GARCH(p, q) alors $\widehat{u}_t = \widehat{\sigma}_t\,\xi_t$ d'où $\widehat{X}_{n+h} = 0$ si $h \geq 1$

Prédiction à partir d'un modèle (3)

Exemples : On veut calculer $\mathbb{E}\big[\widehat{u}_{n+h} \mid (\widehat{u}_1, \ldots, \widehat{u}_n)\big]$

1. Si $(\widehat{u}_t)$ AR(1) causal $\widehat{u}_t = \widehat{\alpha}\, \widehat{u}_{t-1} + \xi_t$ d'où :
   - $\widehat{u}_{n+1} = \mathbb{E}\big[\widehat{\alpha}\, \widehat{u}_n + \xi_{n+1} \mid (\widehat{u}_1, \ldots, \widehat{u}_n)\big] = \widehat{\alpha}\, \widehat{u}_n$
   - $\widehat{u}_{n+2} = \mathbb{E}\big[\widehat{\alpha}^2\, \widehat{u}_n + \widehat{\alpha}\, \xi_{n+1} + \xi_{n+2} \mid (\widehat{u}_1, \ldots, \widehat{u}_n)\big] = \widehat{\alpha}^2\, \widehat{u}_n$
   - $\widehat{u}_{n+h} = \mathbb{E}\big[\widehat{\alpha}^h\, \widehat{u}_n + \sum_{i=1}^h \widehat{\alpha}^{h-i}\, \xi_{n+i} \mid (\widehat{u}_1, \ldots, \widehat{u}_n)\big] = \widehat{\alpha}^h\, \widehat{u}_n$
2. Si $(\widehat{u}_t)$ MA(1) inversible $\widehat{u}_t = \xi_t + \widehat{\alpha}\, \xi_{t-1}$, soit
   $$\widehat{u}_t = \xi_t - \sum_{k=1}^{\infty} (-\widehat{\alpha})^k\, \widehat{u}_{t-k}$$
   - $\widehat{u}_{n+1} = \mathbb{E}\Big[\xi_{n+1} - \sum_{k=0}^{n-1} (-\widehat{\alpha})^{k+1}\, \widehat{u}_{n-k} \mid (\widehat{u}_1, \ldots, \widehat{u}_n)\Big] = -\sum_{k=0}^{n-1} (-\widehat{\alpha})^{k+1}\, \widehat{u}_{n-k}$
   - $\widehat{u}_{n+h} = \mathbb{E}\big[\xi_{n+h} + \widehat{\alpha}\, \xi_{n+h-1} \mid (\widehat{u}_1, \ldots, \widehat{u}_n)\big] = 0$ pour $h \geq 2$
3. Si $(\widehat{u}_t)$ GARCH(p, q) alors $\widehat{u}_t = \widehat{\sigma}_t\, \xi_t$ d'où $\widehat{X}_{n+h} = 0$ si $h \geq 1$

Prédiction à partir d'un modèle (3)

Exemples : On veut calculer $\mathbb{E}\left[\widehat{u}_{n+h} \mid (\widehat{u}_1, \ldots, \widehat{u}_n)\right]$

1. Si $(\widehat{u}_t)$ AR(1) causal $\widehat{u}_t = \widehat{\alpha}\, \widehat{u}_{t-1} + \xi_t$ d'où :
   - $\widehat{u}_{n+1} = \mathbb{E}\left[\widehat{\alpha}\, \widehat{u}_n + \xi_{n+1} \mid (\widehat{u}_1, \ldots, \widehat{u}_n)\right] = \widehat{\alpha}\, \widehat{u}_n$
   - $\widehat{u}_{n+2} = \mathbb{E}\left[\widehat{\alpha}^2\, \widehat{u}_n + \widehat{\alpha}\, \xi_{n+1} + \xi_{n+2} \mid (\widehat{u}_1, \ldots, \widehat{u}_n)\right] = \widehat{\alpha}^2\, \widehat{u}_n$
   - $\widehat{u}_{n+h} = \mathbb{E}\left[\widehat{\alpha}^h\, \widehat{u}_n + \sum_{i=1}^{h} \widehat{\alpha}^{h-i}\, \xi_{n+i} \mid (\widehat{u}_1, \ldots, \widehat{u}_n)\right] = \widehat{\alpha}^h\, \widehat{u}_n$
2. Si $(\widehat{u}_t)$ MA(1) inversible $\widehat{u}_t = \xi_t + \widehat{\alpha}\, \xi_{t-1}$, soit

   $$\widehat{u}_t = \xi_t - \sum_{k=1}^{\infty} (-\widehat{\alpha})^k\, \widehat{u}_{t-k}$$

   - $\widehat{u}_{n+1} = \mathbb{E}\left[\xi_{n+1} - \sum_{k=0}^{n-1} (-\widehat{\alpha})^{k+1}\, \widehat{u}_{n-k} \mid (\widehat{u}_1, \ldots, \widehat{u}_n)\right] = -\sum_{k=0}^{n-1} (-\widehat{\alpha})^{k+1}\, \widehat{u}_{n-k}$
   - $\widehat{u}_{n+h} = \mathbb{E}\left[\xi_{n+h} + \widehat{\alpha}\, \xi_{n+h-1} \mid (\widehat{u}_1, \ldots, \widehat{u}_n)\right] = 0$ pour $h \geq 2$
3. Si $(\widehat{u}_t)$ GARCH(p,q) alors $\widehat{u}_t = \widehat{\sigma}_t\, \xi_t$ d'où $\widehat{X}_{n+h} = 0$ si $h \geq 1$

Prédiction à partir d'un modèle (3)

Exemples : On veut calculer $\mathbb{E}\big[\widehat{u}_{n+h} \mid (\widehat{u}_1, \ldots, \widehat{u}_n)\big]$

1. Si $(\widehat{u}_t)$ AR(1) causal $\widehat{u}_t = \widehat{\alpha}\, \widehat{u}_{t-1} + \xi_t$ d'où :
   - $\widehat{u}_{n+1} = \mathbb{E}\big[\widehat{\alpha}\, \widehat{u}_n + \xi_{n+1} \mid (\widehat{u}_1, \ldots, \widehat{u}_n)\big] = \widehat{\alpha}\, \widehat{u}_n$
   - $\widehat{u}_{n+2} = \mathbb{E}\big[\widehat{\alpha}^2\, \widehat{u}_n + \widehat{\alpha}\, \xi_{n+1} + \xi_{n+2} \mid (\widehat{u}_1, \ldots, \widehat{u}_n)\big] = \widehat{\alpha}^2\, \widehat{u}_n$
   - $\widehat{u}_{n+h} = \mathbb{E}\big[\widehat{\alpha}^h\, \widehat{u}_n + \sum_{i=1}^h \widehat{\alpha}^{h-i}\, \xi_{n+i} \mid (\widehat{u}_1, \ldots, \widehat{u}_n)\big] = \widehat{\alpha}^h\, \widehat{u}_n$
2. Si $(\widehat{u}_t)$ MA(1) inversible $\widehat{u}_t = \xi_t + \widehat{\alpha}\, \xi_{t-1}$, soit

   $$\widehat{u}_t = \xi_t - \sum_{k=1}^{\infty} (-\widehat{\alpha})^k\, \widehat{u}_{t-k}$$

   - $\widehat{u}_{n+1} = \mathbb{E}\big[\xi_{n+1} - \sum_{k=0}^{n-1} (-\widehat{\alpha})^{k+1}\, \widehat{u}_{n-k} \mid (\widehat{u}_1, \ldots, \widehat{u}_n)\big] = -\sum_{k=0}^{n-1} (-\widehat{\alpha})^{k+1}\, \widehat{u}_{n-k}$
   - $\widehat{u}_{n+h} = \mathbb{E}\big[\xi_{n+h} + \widehat{\alpha}\, \xi_{n+h-1} \mid (\widehat{u}_1, \ldots, \widehat{u}_n)\big] = 0$ pour $h \geq 2$
3. Si $(\widehat{u}_t)$ GARCH(p, q) alors $\widehat{u}_t = \widehat{\sigma}_t\, \xi_t$ d'où $\widehat{X}_{n+h} = 0$ si $h \geq 1$

Prédiction sans modèle

Méthode 2 : On détermine $\widehat{X}_{n+h}$ sans modéliser $(X_1, \ldots, X_n)$

$\Longrightarrow$ Influence décroissante (exponentielle) aux données lointaines

$\Longrightarrow$ Filtrage exponentiel, moindres carrés pondérés

1. Filtrage simple de Holt-Winters simple : si $\beta_h \in [0,1]$, on cherche :

$$\widehat{X}_{n+h} = \widehat{a}_h = \operatorname*{Argmin}_{a_h \in \mathbb{R}} \sum_{j=1}^{n} \beta_h^{n-j} \left(X_j - a_h\right)^2$$

$$\Longrightarrow \quad \widehat{a}_h = \left({}^t J \,\Omega(\beta_h)\, J\right)^{-1} {}^t J \,\Omega(\beta_h) \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix}$$

$$\text{avec} \quad \Omega(\beta_h) = \begin{pmatrix} \beta_h^{n-1} & 0 & 0 & \cdots & 0 \\ 0 & \beta_h^{n-2} & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix} \quad \text{et} \quad J = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$$

Prédiction sans modèle

Méthode 2 : On détermine $\widehat{X}_{n+h}$ sans modéliser $(X_1, \ldots, X_n)$

$\Longrightarrow$ Influence décroissante (exponentielle) aux données lointaines

$\Longrightarrow$ Filtrage exponentiel, moindres carrés pondérés

1. Filtrage simple de Holt-Winters simple : si $\beta_h \in [0,1]$, on cherche :

$$\widehat{X}_{n+h} = \widehat{a}_h = \operatorname*{Argmin}_{a_h \in \mathbb{R}} \sum_{j=1}^{n} \beta_h^{n-j} \left(X_j - a_h\right)^2$$

$$\Longrightarrow \quad \widehat{a}_h = \left({}^t J\, \Omega(\beta_h)\, J\right)^{-1} {}^t J\, \Omega(\beta_h) \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix}$$

$$\text{avec} \quad \Omega(\beta_h) = \begin{pmatrix} \beta_h^{n-1} & 0 & 0 & \cdots & 0 \\ 0 & \beta_h^{n-2} & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix} \quad \text{et} \quad J = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$$

Prédiction sans modèle

Méthode 2 : On détermine $\widehat{X}_{n+h}$ sans modéliser $(X_1, \ldots, X_n)$

$\Longrightarrow$ Influence décroissante (exponentielle) aux données lointaines

$\Longrightarrow$ Filtrage exponentiel, moindres carrés pondérés

1. Filtrage simple de Holt-Winters simple : si $\beta_h \in [0,1]$, on cherche :

$$\widehat{X}_{n+h} = \widehat{a}_h = \underset{a_h \in \mathbb{R}}{\text{Argmin}} \sum_{j=1}^{n} \beta_h^{n-j} \left(X_j - a_h\right)^2$$

$$\Longrightarrow \quad \widehat{a}_h = \left({}^t J\, \Omega(\beta_h)\, J\right)^{-1}\, {}^t J\, \Omega(\beta_h) \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix}$$

$$\text{avec} \quad \Omega(\beta_h) = \begin{pmatrix} \beta_h^{n-1} & 0 & 0 & \cdots & 0 \\ 0 & \beta_h^{n-2} & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix} \quad \text{et} \quad J = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$$

Prédiction sans modèle

Méthode 2 : On détermine $\widehat{X}_{n+h}$ sans modéliser $(X_1, \ldots, X_n)$

$\Longrightarrow$ Influence décroissante (exponentielle) aux données lointaines

$\Longrightarrow$ Filtrage exponentiel, moindres carrés pondérés

1. Filtrage simple de Holt-Winters simple : si $\beta_h \in [0,1]$, on cherche :

$$\widehat{X}_{n+h} = \widehat{a}_h = \operatorname*{Argmin}_{a_h \in \mathbb{R}} \sum_{j=1}^{n} \beta_h^{n-j} \left(X_j - a_h\right)^2$$

$$\Longrightarrow \quad \widehat{a}_h = \left({}^t J\, \Omega(\beta_h)\, J\right)^{-1}\, {}^t J\, \Omega(\beta_h) \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix}$$

$$\text{avec} \quad \Omega(\beta_h) = \begin{pmatrix} \beta_h^{n-1} & 0 & 0 & \cdots & 0 \\ 0 & \beta_h^{n-2} & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix} \quad \text{et} \quad J = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$$

Prédiction sans modèle

Méthode 2 : On détermine $\widehat{X}_{n+h}$ sans modéliser $(X_1, \ldots, X_n)$

$\Longrightarrow$ Influence décroissante (exponentielle) aux données lointaines

$\Longrightarrow$ Filtrage exponentiel, moindres carrés pondérés

1. Filtrage simple de Holt-Winters simple : si $\beta_h \in [0,1]$, on cherche :

$$\widehat{X}_{n+h} = \widehat{a}_h = \operatorname*{Argmin}_{a_h \in \mathbb{R}} \sum_{j=1}^{n} \beta_h^{n-j} \left(X_j - a_h\right)^2$$

$$\Longrightarrow \quad \widehat{a}_h = \left({}^t J \, \Omega(\beta_h) \, J\right)^{-1} {}^t J \, \Omega(\beta_h) \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix}$$

avec $\Omega(\beta_h) = \begin{pmatrix} \beta_h^{n-1} & 0 & 0 & \cdots & 0 \\ 0 & \beta_h^{n-2} & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix}$ et $J = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$

Prédiction sans modèle

Méthode 2 : On détermine $\widehat{X}_{n+h}$ sans modéliser $(X_1, \ldots, X_n)$

$\Longrightarrow$ Influence décroissante (exponentielle) aux données lointaines

$\Longrightarrow$ Filtrage exponentiel, moindres carrés pondérés

1. Filtrage simple de Holt-Winters simple : si $\beta_h \in [0,1]$, on cherche :

$$\widehat{X}_{n+h} = \widehat{a}_h = \operatorname*{Argmin}_{a_h \in \mathbb{R}} \sum_{j=1}^{n} \beta_h^{n-j} \left(X_j - a_h\right)^2$$

$$\Longrightarrow \quad \widehat{a}_h = \left({}^t J\, \Omega(\beta_h)\, J\right)^{-1}\, {}^t J\, \Omega(\beta_h) \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix}$$

$$\text{avec} \quad \Omega(\beta_h) = \begin{pmatrix} \beta_h^{n-1} & 0 & 0 & \cdots & 0 \\ 0 & \beta_h^{n-2} & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix} \quad \text{et} \quad J = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$$

Prédiction sans modèle (2)

2. Filtrage de Holt-Winters double : pour $\beta_h \in [0,1]$, on cherche :

$$\widehat{X}_{n+h} = \widehat{a}_h\,(n+h) + \widehat{b}_h \quad \text{où} \quad (\widehat{a}_h, \widehat{b}_h)^t = \operatorname*{Arg\,min}_{a_h, b_h \in \mathbb{R}} \sum_{j=1}^{n} \beta_h^{n-j} \left(X_j - (a_h\, j + b_h)\right)^2$$

$$\Longrightarrow \quad \widehat{a}_h = \left({}^t J\, \Omega(\beta_h)\, J\right)^{-1}\, {}^t J\, \Omega(\beta_h) \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix} \text{avec} \quad J = \begin{pmatrix} 1 & 1 \\ \vdots & \\ 1 & n \end{pmatrix}$$

3. Possibilité d'extension avec saisonnalité s :

$$\widehat{X}_{n+h} = \widehat{a}_h\,(n+h) + \widehat{b}_h + \sum_{k=1}^{T} s_k\, \mathbb{I}_{n+h=k} \quad \text{où}$$

$$(\widehat{a}_h, \widehat{b}_h, \widehat{s}_1, \ldots, \widehat{s}_T)^t = \operatorname*{Arg\,min}_{a_h, b_h, s_1, \ldots, s_T \in \mathbb{R}} \sum_{j=1}^{n} \beta_h^{n-j} \left(X_j - (a_h\, j + b_h) - \sum_{k=1}^{T} s_k\, \mathbb{I}_{j=k}\right)^2$$

Prédiction sans modèle (2)

2. Filtrage de Holt-Winters double : pour $\beta_h \in [0,1]$, on cherche :

$$\widehat{X}_{n+h} = \widehat{a}_h\,(n+h) + \widehat{b}_h \quad \text{où} \quad (\widehat{a}_h, \widehat{b}_h)^t = \operatorname*{Arg\,min}_{a_h, b_h \in \mathbb{R}} \sum_{j=1}^{n} \beta_h^{n-j}\,(X_j - (a_h\,j + b_h))^2$$

$$\Longrightarrow \quad \widehat{a}_h = \left({}^t J\,\Omega(\beta_h)\,J\right)^{-1}\,{}^t J\,\Omega(\beta_h) \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix} \text{avec} \quad J = \begin{pmatrix} 1 & 1 \\ \vdots & \\ 1 & n \end{pmatrix}$$

3. Possibilité d'extension avec saisonnalité s :

$$\widehat{X}_{n+h} = \widehat{a}_h\,(n+h) + \widehat{b}_h + \sum_{k=1}^{T} s_k\,\mathbb{I}_{n+h=k} \quad \text{où}$$

$$(\widehat{a}_h, \widehat{b}_h, \widehat{s}_1, \ldots, \widehat{s}_T)^t = \operatorname*{Arg\,min}_{a_h, b_h, s_1, \ldots, s_T \in \mathbb{R}} \sum_{j=1}^{n} \beta_h^{n-j} \Big(X_j - (a_h\,j + b_h) - \sum_{k=1}^{T} s_k\,\mathbb{I}_{j=k}\Big)^2$$

Prédiction sans modèle (2)

2. Filtrage de Holt-Winters double : pour $\beta_h \in [0,1]$, on cherche :

$$\widehat{X}_{n+h} = \widehat{a}_h\,(n+h) + \widehat{b}_h \quad \text{où} \quad (\widehat{a}_h, \widehat{b}_h)^t = \operatorname{Arg}\min_{a_h, b_h \in \mathbb{R}} \sum_{j=1}^{n} \beta_h^{n-j}\,\big(X_j - (a_h\, j + b_h)\big)^2$$

$$\Longrightarrow \quad \widehat{a}_h = \left({}^t J\,\Omega(\beta_h)\,J\right)^{-1}\,{}^t J\,\Omega(\beta_h) \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix} \text{avec} \quad J = \begin{pmatrix} 1 & 1 \\ \vdots & \\ 1 & n \end{pmatrix}$$

3. Possibilité d'extension avec saisonnalité s :

$$\widehat{X}_{n+h} = \widehat{a}_h\,(n+h) + \widehat{b}_h + \sum_{k=1}^{T} s_k\, \mathbb{I}_{n+h=k} \quad \text{où}$$

$$(\widehat{a}_h, \widehat{b}_h, \widehat{s}_1, \ldots, \widehat{s}_T)^t = \operatorname{Arg}\min_{a_h, b_h, s_1, \ldots, s_T \in \mathbb{R}} \sum_{j=1}^{n} \beta_h^{n-j} \Big(X_j - (a_h\, j + b_h) - \sum_{k=1}^{T} s_k\, \mathbb{I}_{j=k}\Big)^2$$

Prédiction sans modèle (2)

2. Filtrage de Holt-Winters double : pour $\beta_h \in [0,1]$, on cherche :

$$\widehat{X}_{n+h} = \widehat{a}_h\,(n+h) + \widehat{b}_h \quad \text{où} \quad (\widehat{a}_h, \widehat{b}_h)^t = \operatorname*{Arg\,min}_{a_h, b_h \in \mathbb{R}} \sum_{j=1}^{n} \beta_h^{n-j} \left(X_j - (a_h\, j + b_h)\right)^2$$

$$\Longrightarrow \quad \widehat{a}_h = \left({}^t J\,\Omega(\beta_h)\,J\right)^{-1}\, {}^t J\,\Omega(\beta_h) \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix} \text{avec} \quad J = \begin{pmatrix} 1 & 1 \\ \vdots & \\ 1 & n \end{pmatrix}$$

3. Possibilité d'extension avec saisonnalité s :

$$\widehat{X}_{n+h} = \widehat{a}_h\,(n+h) + \widehat{b}_h + \sum_{k=1}^{T} s_k\, \mathbb{I}_{n+h=k} \quad \text{où}$$

$$(\widehat{a}_h, \widehat{b}_h, \widehat{s}_1, \ldots, \widehat{s}_T)^t = \operatorname*{Arg\,min}_{a_h, b_h, s_1, \ldots, s_T \in \mathbb{R}} \sum_{j=1}^{n} \beta_h^{n-j} \left(X_j - (a_h\, j + b_h) - \sum_{k=1}^{T} s_k\, \mathbb{I}_{j=k}\right)^2$$

Prédiction sans modèle (2)

2. Filtrage de Holt-Winters double : pour $\beta_h \in [0,1]$, on cherche :

$$\widehat{X}_{n+h} = \widehat{a}_h\,(n+h) + \widehat{b}_h \quad \text{où} \quad (\widehat{a}_h, \widehat{b}_h)^t = \operatorname*{Arg\,min}_{a_h, b_h \in \mathbb{R}} \sum_{j=1}^{n} \beta_h^{n-j}\,\big(X_j - (a_h\, j + b_h)\big)^2$$

$$\Longrightarrow \quad \widehat{a}_h = \left({}^t J\,\Omega(\beta_h)\,J\right)^{-1}\,{}^t J\,\Omega(\beta_h) \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix} \text{avec} \quad J = \begin{pmatrix} 1 & 1 \\ \vdots & \\ 1 & n \end{pmatrix}$$

3. Possibilité d'extension avec saisonnalité s :

$$\widehat{X}_{n+h} = \widehat{a}_h\,(n+h) + \widehat{b}_h + \sum_{k=1}^{T} s_k\, \mathbb{I}_{n+h=k} \quad \text{où}$$

$$(\widehat{a}_h, \widehat{b}_h, \widehat{s}_1, \dots, \widehat{s}_T)^t = \operatorname*{Arg\,min}_{a_h, b_h, s_1, \dots, s_T \in \mathbb{R}} \sum_{j=1}^{n} \beta_h^{n-j} \Big(X_j - (a_h\, j + b_h) - \sum_{k=1}^{T} s_k\, \mathbb{I}_{j=k}\Big)^2$$

Prédiction sans modèle (3)

Problème : Comment choisir β_h ?

Utilisation d'une **validation croisée** particulière :

- Pour m grand, on choisit une grille $B_m = \{0, 1/m, 2/m, \ldots, 1\}$ pour β_h
- Pour M grand, et $n - M \leq t \leq n - h$, pour chaque β_h choisi,

$$\Longrightarrow \widehat{X}_{t+h}(\beta_h) \text{ calculé à partir de } (X_1, \ldots, X_t)$$

- On choisit $\widehat{\beta}_h = \mathrm{Argmin}_{\beta_h \in B_m} \sum_{t=n-M}^{n-h} \left(X_{t+h} - \widehat{X}_{t+h}(\beta_h) \right)^2$

Remarque : $\beta_h \to 1$ même poids pour tous, $\beta_h \to 0$ dernière valeur

Prédiction sans modèle (3)

Problème : Comment choisir β_h ?

Utilisation d'une **validation croisée** particulière :

- Pour m grand, on choisit une grille $B_m = \{0, 1/m, 2/m, \ldots, 1\}$ pour β_h
- Pour M grand, et $n - M \leq t \leq n - h$, pour chaque β_h choisi,

$$\Longrightarrow \widehat{X}_{t+h}(\beta_h) \text{ calculé à partir de } (X_1, \ldots, X_t)$$

- On choisit $\widehat{\beta}_h = \text{Argmin}_{\beta_h \in B_m} \sum_{t=n-M}^{n-h} \left(X_{t+h} - \widehat{X}_{t+h}(\beta_h)\right)^2$

Remarque : $\beta_h \to 1$ même poids pour tous, $\beta_h \to 0$ dernière valeur

Prédiction sans modèle (3)

Problème : Comment choisir β_h ?

Utilisation d'une **validation croisée** particulière :

- Pour m grand, on choisit une grille $B_m = \{0, 1/m, 2/m, \ldots, 1\}$ pour β_h
- Pour M grand, et $n - M \leq t \leq n - h$, pour chaque β_h choisi,

$$\Longrightarrow \widehat{X}_{t+h}(\beta_h) \text{ calculé à partir de } (X_1, \ldots, X_t)$$

- On choisit $\widehat{\beta}_h = \text{Argmin}_{\beta_h \in B_m} \sum_{t=n-M}^{n-h} \left(X_{t+h} - \widehat{X}_{t+h}(\beta_h)\right)^2$

Remarque : $\beta_h \to 1$ même poids pour tous, $\beta_h \to 0$ dernière valeur

Prédiction sans modèle (3)

Problème : Comment choisir β_h ?

Utilisation d'une **validation croisée** particulière :

- Pour m grand, on choisit une grille $B_m = \{0, 1/m, 2/m, \dots, 1\}$ pour β_h
- Pour M grand, et $n - M \leq t \leq n - h$, pour chaque β_h choisi,

$$\Longrightarrow \widehat{X}_{t+h}(\beta_h) \text{ calculé à partir de } (X_1, \dots, X_t)$$

- On choisit $\widehat{\beta}_h = \text{Argmin}_{\beta_h \in B_m} \sum_{t=n-M}^{n-h} \left(X_{t+h} - \widehat{X}_{t+h}(\beta_h)\right)^2$

Remarque : $\beta_h \to 1$ même poids pour tous, $\beta_h \to 0$ dernière valeur

Prédiction sans modèle (3)

Problème : Comment choisir β_h ?

Utilisation d'une **validation croisée** particulière :

- Pour m grand, on choisit une grille $B_m = \{0, 1/m, 2/m, \ldots, 1\}$ pour β_h
- Pour M grand, et $n - M \leq t \leq n - h$, pour chaque β_h choisi,

$$\Longrightarrow \widehat{X}_{t+h}(\beta_h) \text{ calculé à partir de } (X_1, \ldots, X_t)$$

- On choisit $\widehat{\beta}_h = \mathrm{Argmin}_{\beta_h \in B_m} \sum_{t=n-M}^{n-h} \left(X_{t+h} - \widehat{X}_{t+h}(\beta_h)\right)^2$

Remarque : $\beta_h \to 1$ même poids pour tous, $\beta_h \to 0$ dernière valeur

Prédiction sans modèle (3)

Problème : Comment choisir β_h ?

Utilisation d'une **validation croisée** particulière :

- Pour m grand, on choisit une grille $B_m = \{0, 1/m, 2/m, \ldots, 1\}$ pour β_h
- Pour M grand, et $n - M \leq t \leq n - h$, pour chaque β_h choisi,

$$\Longrightarrow \widehat{X}_{t+h}(\beta_h) \text{ calculé à partir de } (X_1, \ldots, X_t)$$

- On choisit $\widehat{\beta}_h = \text{Argmin}_{\beta_h \in B_m} \sum_{t=n-M}^{n-h} \left(X_{t+h} - \widehat{X}_{t+h}(\beta_h) \right)^2$

Remarque : $\beta_h \to 1$ même poids pour tous, $\beta_h \to 0$ dernière valeur

Prédiction sans modèle (3)

Problème : Comment choisir β_h ?

Utilisation d'une **validation croisée** particulière :

- Pour m grand, on choisit une grille $B_m = \{0, 1/m, 2/m, \dots, 1\}$ pour β_h
- Pour M grand, et $n - M \leq t \leq n - h$, pour chaque β_h choisi,

$$\Longrightarrow \widehat{X}_{t+h}(\beta_h) \text{ calculé à partir de } (X_1, \dots, X_t)$$

- On choisit $\widehat{\beta}_h = \mathrm{Argmin}_{\beta_h \in B_m} \sum_{t=n-M}^{n-h} \left(X_{t+h} - \widehat{X}_{t+h}(\beta_h)\right)^2$

Remarque : $\beta_h \to 1$ même poids pour tous, $\beta_h \to 0$ dernière valeur