

RÉPUBLIQUE ALGÉRIENNE DÉMOCRATIQUE ET POPULAIRE
MINISTÈRE DE L'ENSEIGNEMENT SUPÉRIEURE
ET DE LA RECHERCHE SCIENTIFIQUE

UNIVERSITÉ MUSTAPHA BEN BOULAIID BATNA 2
FACULTÉ DES MATHÉMATIQUES ET DE L'INFORMATIQUE
DÉPARTEMENT DE MATHÉMATIQUES

Filière : Statistique et Analyse des Données (SAD)



Méthodes chronologiques et prévision (Cours-TP)

Par

FATEH MERAHI

Première année Master Mathématiques (SAD)

2022-2023

Table des matières

Introduction	3
1 Rappels sur les séries chronologiques	5
1.1 Les séries chronologiques	5
1.2 Tendances et composantes saisonnières	6
1.3 Indices descriptifs d'une série temporelle	6
1.4 Représentation graphique	7
1.5 Décomposition des séries temporelles	12
1.6 Modèles de décomposition	13
1.7 Le bruit blanc	14
1.8 Suites stationnaires de variables aléatoires	14
1.9 Ajustement de la tendance	14
1.9.1 Ajustement linéaire	15
1.9.2 Ajustement polynômial	16
1.9.3 Ajustement non linéaire	16
1.10 Estimation et élimination de la tendance et de la saisonnalité	17
1.10.1 Moyenne mobile	17
1.10.2 Différenciation	17
1.10.3 Estimation par moyenne mobile	18
1.11 Test sur la série résiduelle	19
Travaux Pratiques	20

2	Lissage et prévision	35
2.1	Lissage par moyennes mobiles	35
2.1.1	Moyenne mobile	35
2.1.2	Lissage par moyennes mobiles	36
2.1.3	La décomposition additive	38
2.2	Lissages exponentiels	40
2.2.1	Lissage exponentiel simple	41
2.2.2	Lissage exponentiel double	42
2.2.3	Méthode de Holt-Winters	44
	Travaux Pratiques	45
3	Modélisation stochastique des séries temporelles	58
3.1	Processus stochastiques stationnaires	58
3.2	Les processus stationnaires ARMA	62
3.2.1	Cas particuliers	63
3.2.2	Espérance, Covariance, corrélation	64
3.2.3	Modèle d'état pour un processus ARMA	65
3.2.4	Identification des processus ARMA	66
3.3	Les processus non stationnaires ARIMA et SARIMA	68
3.4	Les processus ARCH et GARCH	69
3.4.1	Rappels de probabilité	69
3.4.2	Le processus ARCH	69
3.4.3	Le processus GARCH	71
	Travaux Pratiques	73

Introduction

Le présent polycopié reprend un cours de première année Master statistique et analyse des données (SAD) donné à l'Université de Mustapha Ben Boulaid Batn2, Batna, pendant les années universitaires 2021-2022 et 2022-2023. Le but de ce cours était de présenter aux étudiants les notions concernant les méthodes chronologiques et prévision.

Nous supposons que le lecteur a une bonne connaissance des statistiques descriptives, la théorie des probabilités et les notions de base concernant les séries chronologiques. Dans chaque chapitre, nous avons inclus un nombre considérable d'exemples pratiques tels qu'ils ont été testés dans le cadre de travaux pratiques.

Venons-en à une description plus précise du contenu ce polycopié. Dans le premier chapitre, nous avons donné des rappels sur les séries chronologiques concernant la définition, l'objectif de l'étude des séries chronologiques et domaines d'application, tendances et composantes saisonnières, les indices descriptifs, la représentation graphique et la décomposition de ces séries temporelles et les différents modèles de décomposition. On trouve, par la suite, la notion de stationnarités des variables aléatoires et la définition du bruit blanc. Nous avons présenté aussi l'ajustement linéaire et non linéaire de la tendance et nous avons offert une étude concernant l'estimation et l'élimination de la tendance en utilisant la moyenne mobile et la différenciation.

Le second chapitre est consacré pour clarifier le concept de lissage et de la prévision dans le domaine des séries chronologiques. Nous avons donné les propriétés de lissage par les moyennes mobiles et par la méthode de Holt-Winters, de plus nous avons présenté des exemples pratiques traités par Rstudio pour illustrer l'efficacité de ces techniques pour faire des prédictions sur l'évolution de la série et prévoir les valeurs futures de ces séries chronologiques.

Dans le troisième chapitre, nous avons abordé et traité la modélisation stochastique des séries chronologiques dont nous avons donné la définition des processus stochastiques stationnaires au sens strict et au sens large. Nous avons étudié les processus stationnaires ARMA et les processus non stationnaires ARIMA et SARIMA. Nous avons présenté les processus ARCH et GARCH qui permettent de présenter et modéliser les séries financières. L'approche ARCH-GARCH est un outil puissant pour traiter les phénomènes non linéaires et les mouvements imprévisibles ou volatils du marché financier.

Rappels sur les séries chronologiques

1.1 Les séries chronologiques

Definition 1. On appelle série chronologique, ou bien encore chronique ou série temporelle, une suite réelle finie $(x_t)_{1 \leq t \leq n}$ ($n \in \mathbb{N}^*$) de données quantitatives indexée par le temps. L'indice t représente une unité de temps qui peut être selon les cas, la seconde, la minute, l'heure, le jour, le mois, l'année, . . .

Objectif : L'objectif de l'étude des séries temporelles est de faire des prédictions sur l'évolution de la série.

- Analyser un phénomène : Décrire-comprendre-juger l'évolution de la série.
- Prévision : faire des prédictions sur l'évolution de la série et prévoir les valeurs futures de x_t .

Domaines d'application : Séries temporelles, appliquées en :

- Finance : Cours de bourse quotidiens.
- Epidémiologie : Décès mensuels dus à des certaines maladies.
- Econométrie.
- Economie.
- Météorologie, pollution.
- Médecine.
- Biologie.
- Traitement du signal.
- Imagerie médicale.
- Sciences environnementales : océanographie, météorologie, géologie, ...

- Statistiques spatiales appliquées à la prospection et exploitation pétrolière.

Exemple : On suppose que x_t est polynomial en t , par exemple

$$x_t = \alpha_2 t^2 + \alpha_1 t + \alpha_0 + \epsilon_t.$$

avec ϵ_t un bruit aléatoire. On estime les coefficients par $\hat{\alpha}_2, \hat{\alpha}_1, \hat{\alpha}_0$ à partir des valeurs x_1, x_2, \dots, x_n . Ainsi, avec la donnée de x_1, x_2, \dots, x_n on fera la prédiction

$$\hat{x}_{n+1} = \hat{\alpha}_2 (n+1)^2 + \hat{\alpha}_1 (n+1) + \hat{\alpha}_0.$$

1.2 Tendances et composantes saisonnières

On dit que la série $(x_t)_{1 \leq t \leq n}$ admet une **tendance** si on peut écrire

$$x_t = f(t) + \epsilon_t.$$

avec f une fonction fixée et ϵ_t un bruit aléatoire.

- Si $f(t) = \alpha t + \beta$, on dit que la tendance est **linéaire**.
- Si $f(t) = \sum_{i=0}^p \alpha_i t^i$, on dit que la tendance est **polynômiale**.
- Si $f(t)$ est périodique, on dit que la tendance est **périodique**.
- Si $f(t) = s(t) + \alpha t + \beta$, avec s une fonction périodique, on dit que la série a une tendance **linéaire** et **une composante périodique (saisonnière)**.

1.3 Indices descriptifs d'une série temporelle

Indice de tendance centrale : Moyenne empirique,

$$\bar{x}_n = \frac{1}{n} \sum_{t=1}^n x_t.$$

Indices de dispersion : Variance empirique,

$$\hat{V}_n(0) = \frac{1}{n} \sum_{t=1}^n (x_t - \bar{x}_n)^2.$$

La racine carrée de $\hat{V}_n(0)$ est l'**écart-type empirique** $\hat{\sigma}_n(0)$,

$$\hat{\sigma}_n(0) = \sqrt{\hat{V}_n(0)}.$$

Indices de dépendance : Auto-covariance empirique d'ordre h ($h \in \mathbb{N}^*$),

$$\widehat{V}_n(h) = \frac{1}{n-h} \sum_{t=1}^{n-h} (x_t - \bar{x}_n)(x_{t+h} - \bar{x}_n), \quad (h < n).$$

Les indices de dispersion renseignent sur la dépendance entre les données x_t .

- Fonction d'auto-covariance empirique :

$$h \mapsto \widehat{V}_n(h).$$

- Auto-corrélation empirique : prend ses valeurs dans $[0, 1]$,

$$\widehat{\rho}_n(h) = \frac{\widehat{V}_n(h)}{\widehat{V}_n(0)}.$$

- Fonction auto-corrélation empirique :

$$h \mapsto \widehat{\rho}_n(h).$$

1.4 Représentation graphique

On représente graphiquement la série chronologique $(x_t)_{1 \leq t \leq n}$:

- en dessinant le nuage formé par les points $(t; x_t)_{1 \leq t \leq n}$.
- en reliant les points entre eux par des segments de droite, pour indiquer la chronologie.

Remarque 2. La représentation graphique d'un nuage de points $(x_t, x_{t+h})_{1 \leq t \leq n-h}$ ($h < n$) illustre la valeur de $\widehat{\rho}_n(h)$. Plus le nuage est **arrondi**, plus $\widehat{\rho}_n(h)$ est proche de 0. Plus le nuage est **allongé**, plus $\widehat{\rho}_n(h)$ est proche de 1.

Exemple 3. Série airpass : nombre mensuel de passagers aériens, en milliers, de janvier 1949 à décembre 1960.

Proposition 4. Supposons $x_t = at + b + \epsilon_t$, avec $(\epsilon_t)_{t \geq 1}$ une suite de variables aléatoires i.i.d. (indépendantes et identiquement distribuées) et $a \neq 0$. Supposons que $\mathbb{E}\{\epsilon_1^2\} < \infty$. Alors, pour h fixé dans \mathbb{N}^* ,

$$\widehat{\rho}_n(h) \xrightarrow{p.s.} 1.$$

Démonstration. Notons

$$\bar{\epsilon}_n = \frac{1}{n} \sum_{t=1}^n \epsilon_t.$$

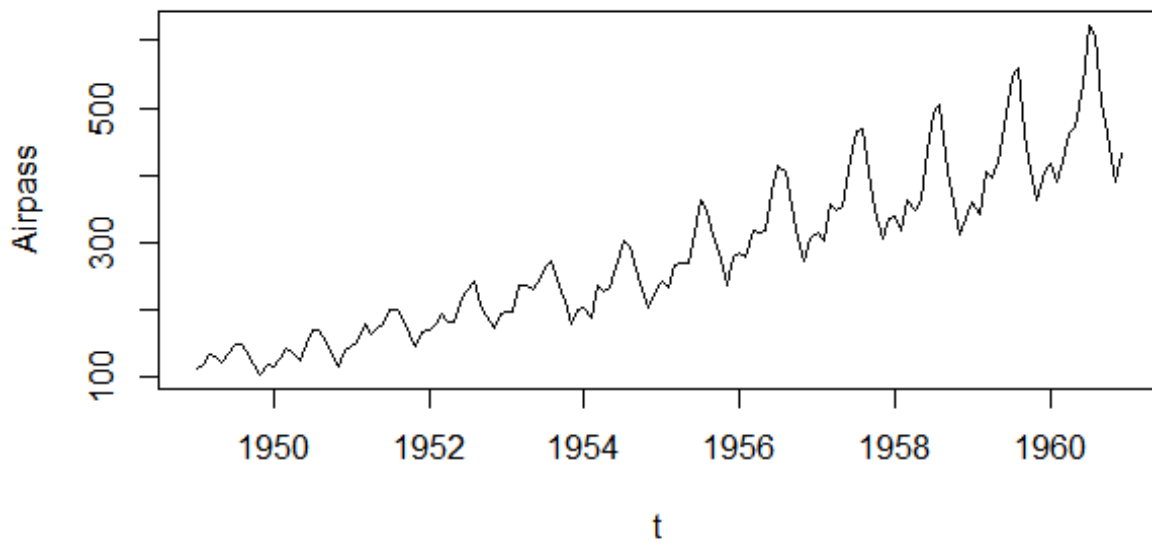


FIGURE 1.1 – Nombre mensuel de passagers aériens, en milliers, de janvier 1949 à décembre 1960.

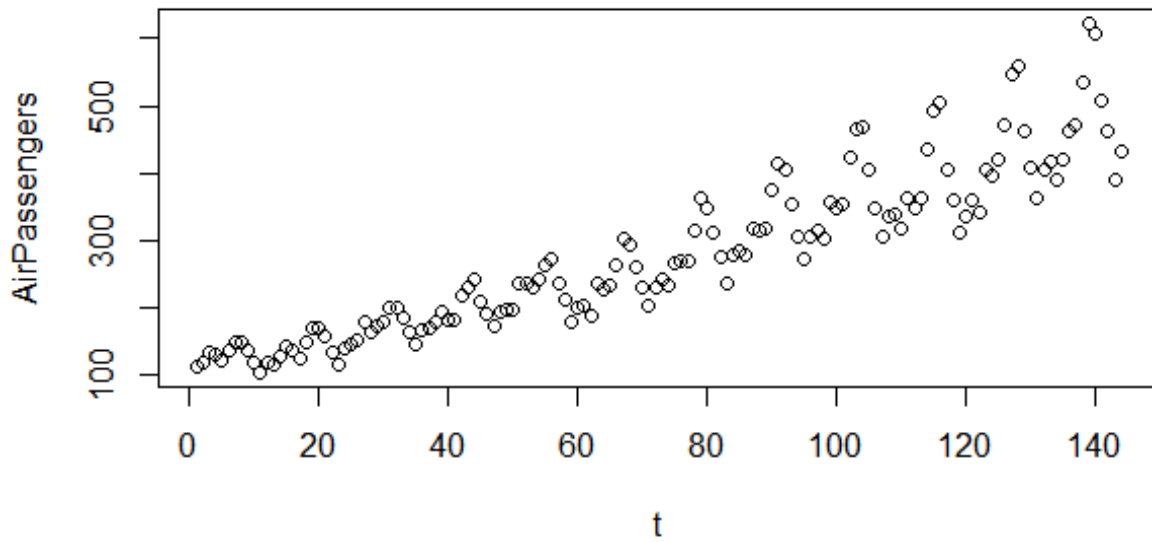


FIGURE 1.2 – Nuage de points simple

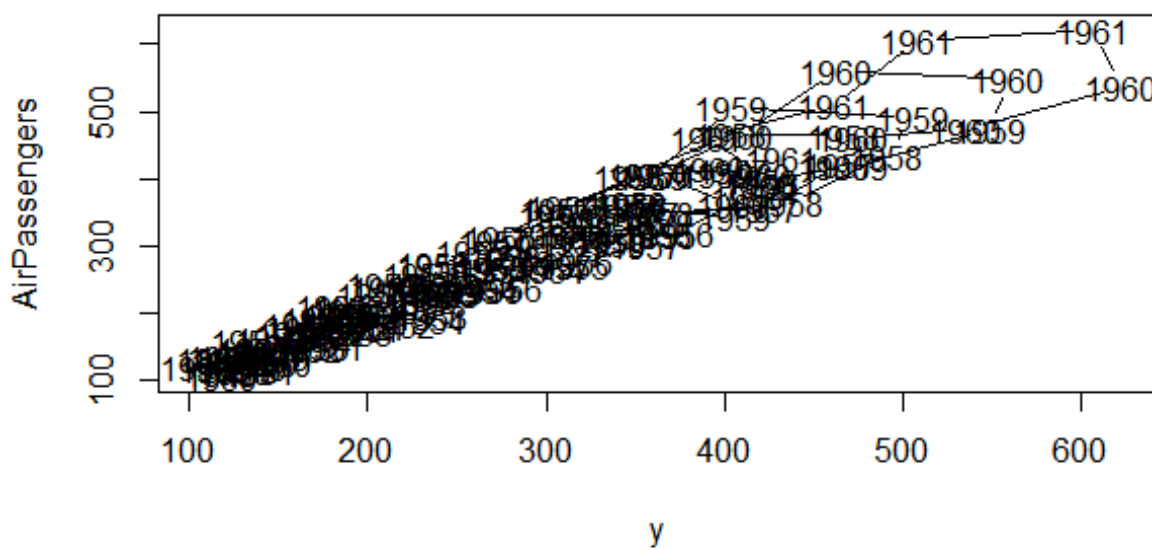


FIGURE 1.3 – Deux nuages de points (x_t, x_{t+1}) .

Nous avons

$$\bar{x}_n = \frac{1}{n} \sum_{t=1}^n (at + b + \epsilon_t) = \frac{a(n+1)}{2} + b + \bar{\epsilon}_n.$$

Fixons $h \in \{1, 2, n-1\}$. Nous avons

$$\begin{aligned} \widehat{V}_n(h) &= \frac{1}{n-h} \sum_{t=1}^{n-h} (x_t - \bar{x}_n)(x_{t+h} - \bar{x}_n) \\ &= \frac{1}{n-h} \sum_{t=1}^{n-h} \left(at + b + \epsilon_t - \frac{a(n+1)}{2} - b - \bar{\epsilon}_n \right) \left(a(t+h) + b + \epsilon_{t+h} - \frac{a(n+1)}{2} - b - \bar{\epsilon}_n \right) \\ &= \frac{1}{n-h} \sum_{t=1}^{n-h} \left(a \left(t - \frac{(n+1)}{2} \right) + \epsilon_t - \bar{\epsilon}_n \right) \left(a \left(t+h - \frac{(n+1)}{2} \right) + \epsilon_{t+h} - \bar{\epsilon}_n \right) \\ &= \frac{1}{n-h} \sum_{t=1}^{n-h} \left[a^2 \left(t - \frac{(n+1)}{2} \right) \left(t+h - \frac{(n+1)}{2} \right) + (\epsilon_t - \bar{\epsilon}_n)(\epsilon_{t+h} - \bar{\epsilon}_n) \right. \\ &\quad \left. + (\epsilon_t - \bar{\epsilon}_n) a \left(t+h - \frac{(n+1)}{2} \right) + (\epsilon_{t+h} - \bar{\epsilon}_n) a \left(t - \frac{(n+1)}{2} \right) \right]. \end{aligned}$$

Nous avons

$$\frac{1}{n-h} \sum_{t=1}^{n-h} (\epsilon_t - \bar{\epsilon}_n)(\epsilon_{t+h} - \bar{\epsilon}_n) = \frac{1}{n-h} \sum_{t=1}^{n-h} (\epsilon_t \epsilon_{t+h} + \bar{\epsilon}_n^2 - \bar{\epsilon}_n \epsilon_{t+h} - \epsilon_t \bar{\epsilon}_n),$$

par application de la loi de grands nombres, on a

$$\begin{aligned} \frac{1}{n-h} \sum_{t=1}^{n-h} \epsilon_t \epsilon_{t+h} &\xrightarrow{p.s} \mathbb{E} \{ \epsilon_1 \epsilon_{1+h} \}, \\ \frac{1}{n-h} \sum_{t=1}^{n-h} \bar{\epsilon}_n^2 &= \bar{\epsilon}_n^2 \xrightarrow{p.s} \mathbb{E} \{ \epsilon_1^2 \}, \\ \frac{1}{n-h} \sum_{t=1}^{n-h} \bar{\epsilon}_n \epsilon_{t+h} &= \bar{\epsilon}_n \frac{1}{n-h} \sum_{t=1}^{n-h} \epsilon_{t+h} \xrightarrow{p.s} (\mathbb{E} \{ \epsilon_1 \})^2, \\ \frac{1}{n-h} \sum_{t=1}^{n-h} \bar{\epsilon}_n \epsilon_t &= \bar{\epsilon}_n \frac{1}{n-h} \sum_{t=1}^{n-h} \epsilon_t \xrightarrow{p.s} (\mathbb{E} \{ \epsilon_1 \})^2. \end{aligned}$$

Par Cauchy-Schwartz,

$$\begin{aligned} \left| \frac{1}{n-h} \sum_{t=1}^{n-h} (\epsilon_t - \bar{\epsilon}_n) a \left(t+h - \frac{(n+1)}{2} \right) \right| &\leq |a| \left(\frac{1}{n-h} \sum_{t=1}^{n-h} (\epsilon_t - \bar{\epsilon}_n)^2 \right)^{1/2} \left(\frac{1}{n-h} \sum_{t=1}^{n-h} \left(t+h - \frac{(n+1)}{2} \right)^2 \right)^{1/2} \\ &\leq |a| \left(\frac{1}{n-h} \sum_{t=1}^{n-h} (\epsilon_t - \bar{\epsilon}_n)^2 \right)^{1/2} \left(\frac{1}{n-h} \left(n+h - \frac{(n+1)}{2} \right)^2 \right)^{1/2} \end{aligned}$$

Nous avons (par application de la loi de grands nombres)

$$\frac{1}{n-h} \sum_{t=1}^{n-h} (\epsilon_t - \bar{\epsilon}_n)^2 \xrightarrow{p.s} \text{Var} \{ \epsilon_1 \}.$$

Donc, p.s.,

$$\frac{1}{n-h} \sum_{t=1}^{n-h} (\epsilon_t - \bar{\epsilon}_n) a \left(t + h - \frac{(n+1)}{2} \right) = O(n).$$

De même

$$\frac{1}{n-h} \sum_{t=1}^{n-h} (\epsilon_{t+h} - \bar{\epsilon}_n) a \left(t - \frac{(n+1)}{2} \right) = O(n).$$

Nous avons

$$\begin{aligned} & \frac{1}{n-h} \sum_{t=1}^{n-h} a^2 \left(t - \frac{(n+1)}{2} \right) \left(t + h - \frac{(n+1)}{2} \right) \\ &= \frac{a^2}{n-h} \sum_{t=1}^{n-h} \left[t^2 + \frac{(n+1)^2}{4} + \left((h - (n+1))t + \frac{(n+1)h}{2} \right) \right] \end{aligned}$$

Par la formule pour la somme des carrés, on obtient

$$\begin{aligned} & \frac{1}{n-h} \sum_{t=1}^{n-h} a^2 \left(t - \frac{(n+1)}{2} \right) \left(t + h - \frac{(n+1)}{2} \right) \\ &= \frac{a^2}{n-h} \left(\frac{(2(n-h)+1)(n-h+1)(n-h)}{6} \right) + a^2 \frac{(n+1)^2}{4} \\ & \quad - \frac{a^2(h-n-1)n(n+1)}{n-h} \frac{1}{2} + a^2 \frac{(n+1)}{2} \\ &= \frac{a^2}{3} n^2 + \frac{a^2}{4} n^2 + o(n^2). \end{aligned}$$

Donc, p.s.,

$$\widehat{V}_n(h) \underset{n \rightarrow \infty}{\sim} \frac{7a^2}{12} n^2.$$

Donc quand $n \rightarrow \infty$,

$$\widehat{\rho}_n(h) \xrightarrow{p.s.} 1.$$

□

Proposition 5. Si $x_t = a \cos(2\pi t/p) + \epsilon_t$, où $(\epsilon_t)_{t \geq 1}$ une suite de variable aléatoires i.i.d. (indépendantes et identiquement distribuées) avec $a \neq 0$ et $p \in \mathbb{N}^*$. Supposons que $\mathbb{E}\{\epsilon_1^2\} < \infty$. Alors, pour h fixé dans \mathbb{N}^* ,

$$\widehat{\rho}_n(h) \xrightarrow{p.s.} \cos(2\pi h/p) \quad \text{quand } n \rightarrow \infty$$

Remarque :

- si l'auto-corrélation d'une série est constante, c'est qu'elle a une tendance linéaire.
- si l'auto-corrélation est périodique de période p , alors la série a une composante périodique de période p .

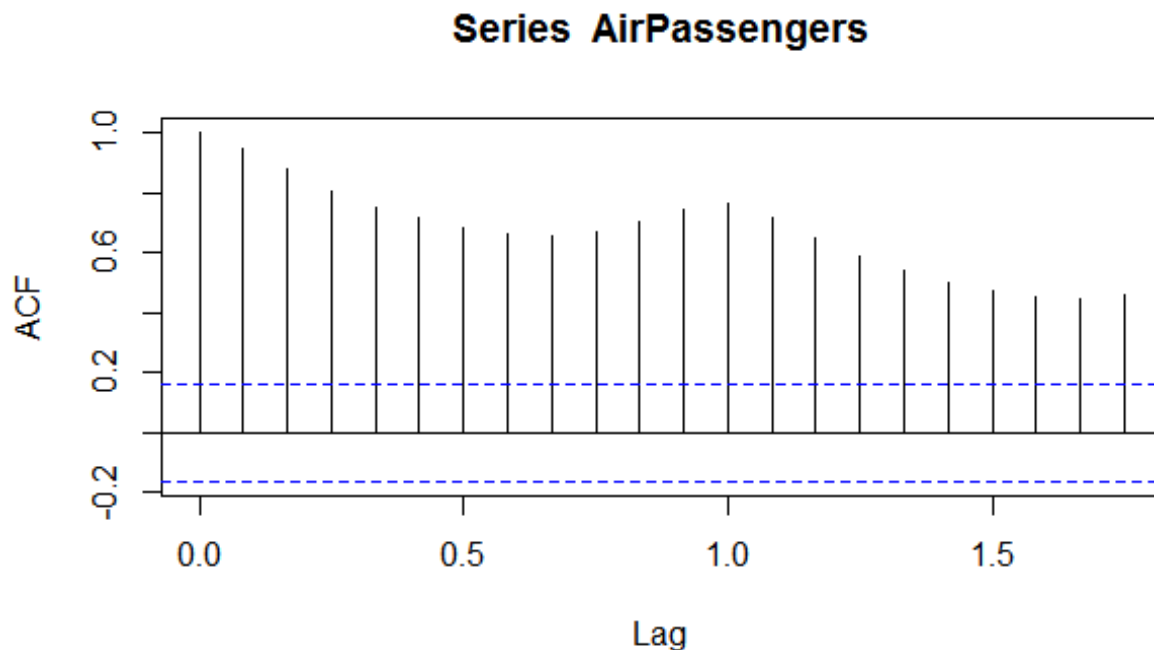


FIGURE 1.4 – La fonction d'autocorrelation (ACF) de la série "AirPassengers".

1.5 Décomposition des séries temporelles

La série chronologique (x_t) peut se décomposer en :

- **une tendance** $(f_t)_{1 \leq t \leq n}$: représente l'évolution à long terme de la grandeur étudiée, et traduit l'aspect général de la série. C'est une fonction monotone, qui peut contenir par exemple des composantes :

- linéaire : $f_t = at + b$.
- quadratique : $f_t = at^2 + bt + c$.
- polynomiale : $f_t = a_m t^m + a_{m-1} t^{m-1} + \dots + a_1 t + a_0$, ($a_m \neq 0$, $a_m \in \mathbb{R}$, $\forall m = 2, 3, \dots$)
- exponentielles : $f_t = f_0 \exp(at)$.
- . . .

avec les paramètres a, b, c, f_0 . . . à déterminer.

- **une composante saisonnière (périodique)** $(s_t)_{1 \leq t \leq n}$: sont liées au rythme imposé par les saisons météorologiques ou encore par des activités économiques et sociales. Mathématiquement, il s'agit des fonctions périodiques, c'est-à-dire qu'il existe un entier p , appelé

période, tel que

$$s_t = s_{t+p}, \quad \forall t = 1, 2, \dots$$

pour des raisons d'unicité dans la décomposition proposée on impose que la somme des facteurs saisonniers soit nulle :

$$\sum_{t=1}^p s_t = 0.$$

• **une composante résiduelle** $(\epsilon_t)_{1 \leq t \leq n}$: sont des variations de faible intensité et de courte durée, et de nature aléatoire.

1.6 Modèles de décomposition

◦ Modèle additive :

$$x_t = f_t + s_t + \epsilon_t \quad \text{pour } t = 1, 2, \dots, n,$$

avec

$$\sum_{t=1}^p s_t = 0 \quad \text{et} \quad \sum_{t=1}^n \epsilon_t = 0.$$

◦ Modèle multiplicative :

$$x_t = f_t \times s_t \times \epsilon_t \quad \text{pour } t = 1, 2, \dots, n,$$

avec

$$\sum_{t=1}^p s_t = 0 \quad \text{et} \quad \sum_{t=1}^n \epsilon_t = 0.$$

On remarque qu'on se ramène naturellement à un modèle additif par passage au log.

◦ Modèle mixte :

$$x_t = f_t (1 + s_t)(1 + \epsilon_t) \quad \text{pour } t = 1, 2, \dots, n,$$

avec

$$\sum_{t=1}^p s_t = 0 \quad \text{et} \quad \sum_{t=1}^n \epsilon_t = 0.$$

1.7 Le bruit blanc

Une suite de variables aléatoires $(e_t)_{t \in \mathbb{N}}$ est dite un processus de bruit blanc si :

- L'espérance et la variance sont constante, c'est-à-dire $\mathbb{E}\{e_t\}$ et $\text{Var}\{e_t\}$ ne dépendent pas de t .

- $\text{Cov}(e_t, e_{t+h}) = 0$, si $h \neq 0$.

- Si l'espérance est nulle, le bruit blanc est dit **centré**.

- Si les variables sont gaussiennes, le bruit blanc est dit **gaussien**.

- Si les variables sont *i.i.d* (**indépendents et identiquement distribués**), le bruit blanc est **stationnaire**.

- Si l'espérance est nulle $\mathbb{E}\{e_t\} = 0$, la variance constante $\text{Var}\{e_t\} = \sigma^2$, et covariance $C(s, t) = \mathbb{E}\{e_t e_s\} = 0, \forall t \neq s$, $(e_t)_{t \in \mathbb{N}}$ est un bruit blanc de second ordre.

1.8 Suites stationnaires de variables aléatoires

Un processus aléatoire $(X_t)_{t \in \mathbb{N}}$ est stationnaire si :

- L'espérance constante : $\mathbb{E}\{X_t\}$ ne dépendent pas de t .

- les covariances sont stables par translation dans le temps : $\text{Cov}(X_t, X_{t+h})$

1.9 Ajustement de la tendance

On dispose d'une série chronologique $(x_t)_{t=1, \dots, n}$ où les seules composantes présentes sont :

- la tendance f_t .

- la composante résiduelle ϵ_t .

$$x_t = f_t + \epsilon_t.$$

Objectif : trouver une fonction simple du temps qui modélise au mieux la tendance de la série $(x_t)_{t=1, \dots, n}$.

1.9.1 Ajustement linéaire

On veut ajuster les données x_1, x_2, \dots, x_n par une droite et donc modéliser la tendance par une fonction linéaire de la forme :

$$f_t = at + b,$$

$$\theta = (a, b)'$$

Deux méthodes :

• **la méthode des moindres carrés :**

On peut estimer $\theta = (a, b)'$ par

$$\hat{\theta}' = (\hat{a}, \hat{b}) = \arg \min_{a, b} \sum_{t=1}^n (x_t - at - b)^2.$$

La solution (\hat{a}, \hat{b}) est donnée par

$$\begin{aligned} \hat{a} &= \frac{\sum_{t=1}^n tx_t - n\bar{x}_n\bar{t}_n}{\sum_{t=1}^n t^2 - n(\bar{t}_n)^2}, \\ \hat{b} &= \bar{x}_n - \hat{a}\bar{t}_n \end{aligned}$$

avec $\bar{t}_n = \frac{1}{n} \sum_{t=1}^n t$, $\sum_{t=1}^n t = \frac{n(n+1)}{2}$ et $\sum_{t=1}^n t^2 = \frac{n(n+1)(2n+1)}{6}$.

- La droite d'équation $x_t = \hat{a}t + \hat{b}$ s'appelle la droite des moindres carrés (**DMC**).
- Le point moyen de coordonnées (\bar{t}_n, \bar{x}_n) appartient à la **DMC**.
- On peut apprécier la qualité de l'ajustement linéaire à l'aide du **coefficient de corrélation linéaire** noter et défini par :

$$r = \frac{\sum_{t=1}^n (x_t - \bar{x}_n)(t - \bar{t}_n)}{\sqrt{\sum_{t=1}^n (x_t - \bar{x}_n)^2} \sqrt{\sum_{t=1}^n (t - \bar{t}_n)^2}}.$$

- La variance de la série $(x_t)_{t=1, \dots, n}$ se décompose en

$$\underbrace{\frac{1}{n} \sum_{t=1}^n (x_t - \bar{x}_n)^2}_{\text{Variance totale de } y} = \underbrace{\frac{1}{n} \sum_{t=1}^n ((\hat{a}t + \hat{b}) - \bar{x}_n)^2}_{\text{Variance expliquée par la DMC}} + \underbrace{\frac{1}{n} \sum_{t=1}^n (x_t - (\hat{a}t + \hat{b}))^2}_{\text{Variance résiduelle}}$$

1.9.2 Ajustement polynômial

On veut ajuster les données par un polynôme de degré d et donc modéliser la tendance par une fonction de la forme :

$$f_t = a_d t^d + a_{d-1} t^{d-1} + \dots + a_1 t^1 + a_0, (a_d \neq 0, a_j \in \mathbb{R}, \forall j = 1, 2, \dots, d)$$

avec

$$\theta = (a_0, a_1, \dots, a_d)'$$

- Le cas $d = 0$ correspond à la tendance constante et le cas $d = 1$ à la tendance linéaire.
- On estime les paramètres par la méthode des moindres carrés

$$\hat{\theta}' = \arg \min_{a_0, a_1, \dots, a_d} \sum_{t=1}^n \left(x_t - (a_d t^d + a_{d-1} t^{d-1} + \dots + a_1 t^1 + a_0) \right)^2,$$

ou encore

$$\hat{\theta}' = \arg \min_{a_0, a_1, \dots, a_d} \|X - T\theta\|.$$

- La solution de ce problème de minimisation est bien connue :

$$\hat{\theta}' = (T'T)^{-1} T'X,$$

où T' désigne la matrice transposée de T et

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \quad \text{et} \quad T = \begin{pmatrix} 1 & 1^1 & 1^2 & \dots & 1^d \\ 1 & 2^1 & 2^2 & \dots & 2^d \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & n^1 & n^2 & \dots & n^d \end{pmatrix}.$$

- Le choix du degré du polynôme est basé sur l'examen du graphe de la série $(x_t)_{t=1, \dots, n}$ qui permet de se faire une première idée du degré et des résidus obtenus après l'ajustement.

1.9.3 Ajustement non linéaire

On souhaite ajuster à la série $(x_t)_{t=1, \dots, n}$ une tendance non linéaire paramétrique $(f_t)_{t=1, \dots, n}$. Plusieurs approches, en particulier :

- **le changement de variable** : Se ramener à des tendances linéaires ou polynômiales grâce à un simple changement de variables. Cette approche est possible pour des tendances de certaine forme particulière :

- $f(t) = \frac{1}{at+b}$.
- $f(t) = b \exp(at)$.
- $f(t) = \exp(at + b)$.
- $f(t) = \ln(a + bt + ct^2)$.
-

1.10 Estimation et élimination de la tendance et de la saisonnalité

1.10.1 Moyenne mobile

Si $(x_t)_{t=1,\dots,n}$ est une série temporelle, alors la moyenne mobile d'ordre t associée est la série temporelle définie pour par

$$\hat{x}_t = \frac{1}{2q+1} \sum_{k=-q}^q x_{t+k},$$

et on a par convention $x_t = x_1$ si $t \leq 0$ et $x_t = x_n$ si $t \geq n$.

1.10.2 Différenciation

On peut éliminer la tendance par différenciation.

- Notons Δ l'opérateur de différenciation, c'est à dire l'opérateur défini par

$$\Delta x_t = x_t - x_{t-1}$$

pour $t \geq 2$.

- On définit l'opérateur de différenciation d'ordre k par la formule de récurrence

$$\Delta^{(k)} x_t = \Delta \left(\Delta^{(k-1)} x_t \right)$$

- Propriétés :

- Soient u_t, v_t deux séries temporelles,

$$\Delta(u_t + v_t) = \Delta u_t + \Delta v_t.$$

- Soit x_t une série temporelle et $\lambda \in \mathbb{R}$, $\Delta(\lambda x_t) = \lambda \Delta x_t$.
- Soit $x_t = a_0 + a_1 t + \dots + a_d t^d$, alors $\Delta^{(d)} x_t = d! a_d$ et $\Delta^{(d+1)} x_t = 0$.

1.10.3 Estimation par moyenne mobile

- **Tendance linéaire** : On suppose que

$$x_t = f_t + \epsilon_t$$

avec f_t déterministe et ϵ_t aléatoire. Si f_t est une fonction affine (donc à croissance linéaire) dans un intervalle autour de t , on peut estimer la moyenne par

$$\hat{f}_t = \frac{1}{2q+1} \sum_{k=-q}^q x_{t+k}$$

- **Tendance linéaire et composante périodique** : On suppose que

$$x_t = f_t + s_t + \epsilon_t$$

avec f_t et s_t déterministe et ϵ_t aléatoire. La fonction f_t est supposée affine et la fonction s_t est supposée p -périodique. On peut toujours supposer que

$$\sum_{t=1}^p s_t = 0.$$

On peut estimer f_t par la formule

$$\hat{f}_t = \frac{1}{2q+1} \sum_{k=-q}^q x_{t+k}.$$

Alors, on estime la composante périodique (saisonnière) en faisant la moyenne sur toutes les périodes :

$$\forall t \in \{1, 2, \dots, p\}, \hat{s}_t = \frac{1}{k+1} \sum_{i=0}^k (x_{t+ip} - \hat{f}_{t+ip}),$$

avec $k = \sup \{i : t + ip \leq n\}$. Puis pour $t \in \{p+1, p+2, \dots, n\}$

$$\hat{s}_t = \hat{s}_{t[p]},$$

où $t[p]$ est le reste de la division euclidienne de t par p . Cette dernière formule fournit un estimateur raisonnable puisque s est p -périodique.

- **Élimination de la tendance et de la saisonnalité par la méthode des différences** :

On peut éliminer la partie déterministe sans l'estimer. Fixons p dans \mathbb{N}^* et notons

$$\Delta_p x_t = x_t - x_{t-p}, \quad \forall t \in \mathbb{Z}.$$

Proposition 6. Si $(x_t)_{t=1,\dots,n}$ est une série temporelle admettant une tendance polynômiale de degré $d \geq 1$ alors $\Delta_p x_t$ admet une tendance polynômiale de degré $\leq d - 1$. Si $(x_t)_{t=1,\dots,n}$ est une série temporelle admettant une tendance polynômiale de degré 0 (constante) alors $\Delta_p x_t$ une tendance égale à 0.

Proposition 7. Dans le cas $x_t = f_t + s_t + \epsilon_t$ avec f_t et s_t déterministe et ϵ_t un bruit et s_t est p -périodique. le processus $\Delta_p x_t$ est un processus qui ne contient plus de partie périodique. De plus, si la tendance est linéaire, elle est également supprimée.

1.11 Test sur la série résiduelle

- On veut tester la présence d'un bruit blanc :

Lemme 8. Si $(\epsilon_t)_{t \geq 1}$ un bruit est un bruit blanc, alors ses fonctions d'auto-corrélation empiriques vérifient quand $n \rightarrow \infty$,

$$\sqrt{n} \hat{\rho}_n(h) \xrightarrow{loi} \mathcal{N}(0, 1),$$

pour tout $h \in \mathbb{N}^*$, De plus, les $(\hat{\rho}_n(h))_{h \geq 1}$ sont asymptotiquement indépendants quand $n \rightarrow \infty$.

- Test d'hypothèses simple : On veut tester l'hypothèse nulle H_0 suivante :

H_0 : Les (ϵ_t) forment un bruit blanc,

contre

H_1 : Les (ϵ_t) ne forment pas un bruit blanc.

- on fixe le niveau $\alpha = 0.05$.
- D'après le lemme précédent ; si H_0 est vraie,

$$P\left(\sqrt{n} \hat{\rho}_n(1) \in [-1.96, 1.96]\right) \xrightarrow{n \rightarrow \infty} P(Z \in [-1.96, 1.96]) = 0.05,$$

avec $Z \rightsquigarrow \mathcal{N}(0, 1)$.

- La procédure de test :
 - si $\sqrt{n} \hat{\rho}_n(1) \in \left[-\frac{1.96}{\sqrt{n}}, \frac{1.96}{\sqrt{n}}\right]$ alors on considère que H_0 est vraie.
 - La probabilité de rejeter à tort H_0 est donc inférieure ou égale à α .

- Test d'hypothèses plus sophistiqués :

◦ soit

$$Q_n = \sum_{j=1}^h \hat{\rho}_n^2(j) \xrightarrow{loi} \mathcal{X}_h^2, \quad \text{quand } n \rightarrow \infty$$

\mathcal{X}_h^2 une loi du **chi-deux** à **h** degrés de liberté.

◦ on rejette H_0 quand

$$Q_n \geq \mathcal{X}_{h,1-\alpha}^2.$$

$\mathcal{X}_{h,1-\alpha}^2$: α -quantile de la loi limite.

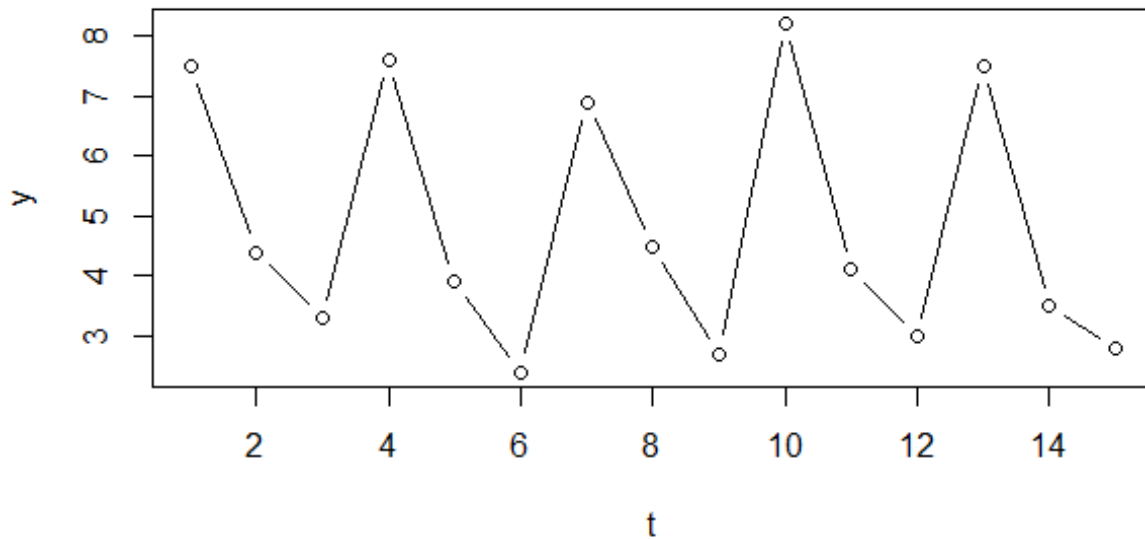
◦ on se limitera à un **lag** inférieur 10 dans l'utilisation de **Box.test**.

Travaux Pratiques

Séries temporelles en R

Quelques fonctions R utiles à l'étude des séries temporelles :

- **serie <- ts (data,start,end,frequency)** : créer un objet de type série temporelle.
 - **data** contient le vecteur des données (un fichier contenant les données peut être mentionné en remplaçant **data** par **file="donnees.dat"**),
 - **start** et **end** mentionne les dates de début et de fin de la série : par exemple **start=c(1990,1)** et **end=c(1999,6)** pour des données allant de janvier 90 à juin 99.
 - **frequency** mentionne le nombre de données par unité de temps : par exemple si les dates de début et de fin sont des années, et que les données sont mensuelles, il faudra indiquer **frequency=12**.
- **plot.ts(serie)** : Représenter graphiquement un objet de type série temporelle.
- **x <- rnorm(n,mean=0,sd=1)** : Simulation de variables normales indépendantes de moyenne 0 et d'écart-type 1.
- **acf(serie,lag.max=25,type=c('correlation'))** : Représenter graphiquement la fonction d'autocorrélation. La ligne pointillée bleue indique que la corrélation n'est plus statistiquement significative.
- **diff.ts(serie,lag=T,difference=k)** : la fonction **diff** permet d'appliquer l'opérateur de différenciation Δ_p^k .
- **serie_decomp <- decompose(serie,type=c("additive","multiplicative"))** : La fonction **decompose** permet d'extraire d'une série temporelle (via la méthode de la moyenne mobile) choisir "additive", "multiplicative" mais pas les deux en même temps).
 - **serie_decomp\$seasonal** : la composante saisonnière.



- `serie_decomp$trend` : la tendance
- `serie_decomp$random` : la partie aléatoire stationnaire de la série.

.

Exemple 1 :

```
#Transformation d'un vecteur de données en un objet « série chronologique »
```

```
y=c(7.5,4.4,3.3,7.6,3.9,2.4,6.9,4.5,2.7,8.2,4.1,3.0,7.5,3.5,2.8)
```

```
t=1 :15
```

```
plot(t,y,type="b")
```

```
# Créer la série temporelle z :
```

```
z=ts(y, frequency = 3)
```

```
#Décomposition d'une série chronologique en une tendance, un effet saisonnier, et un résidu
```

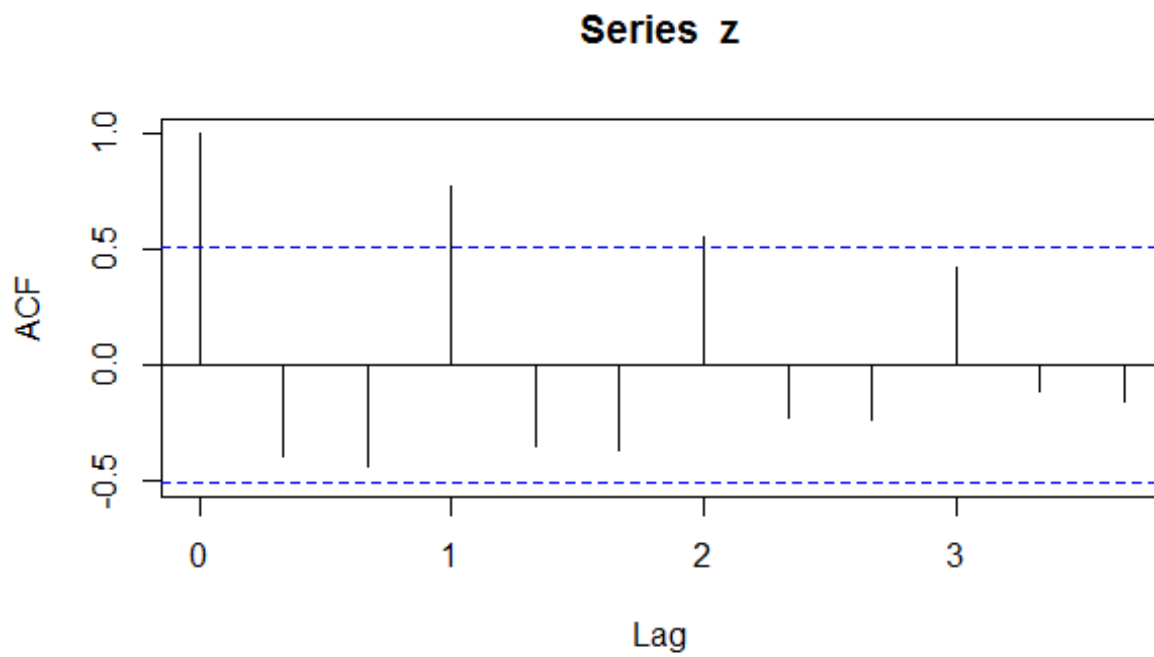
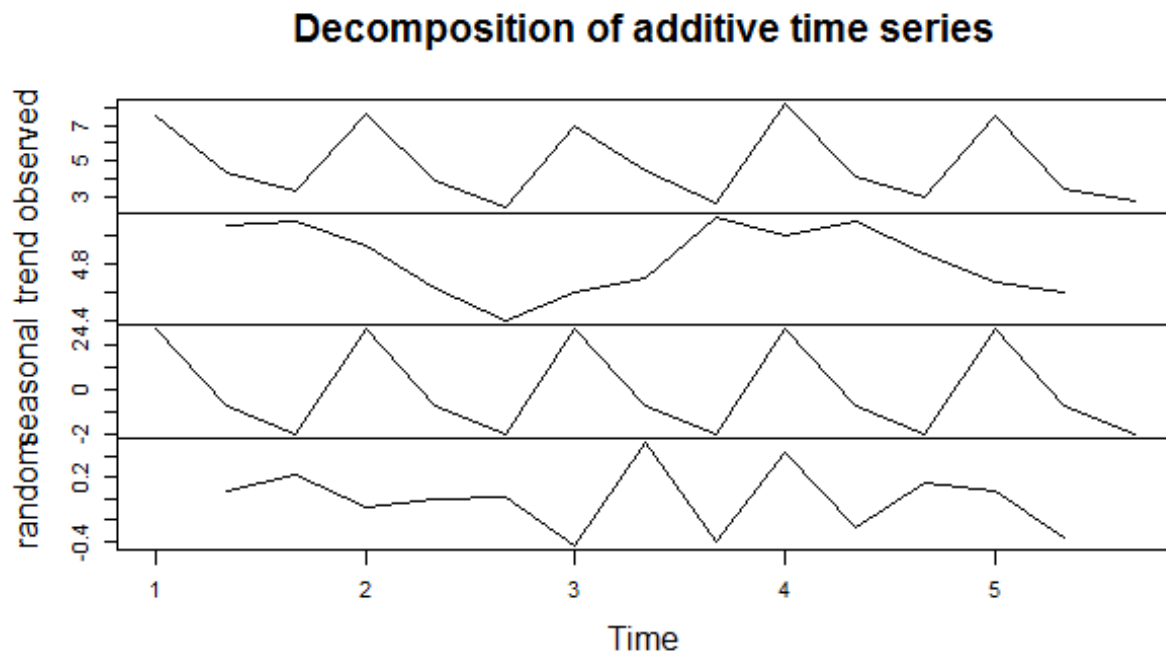
```
z.dec <-decompose(z)
```

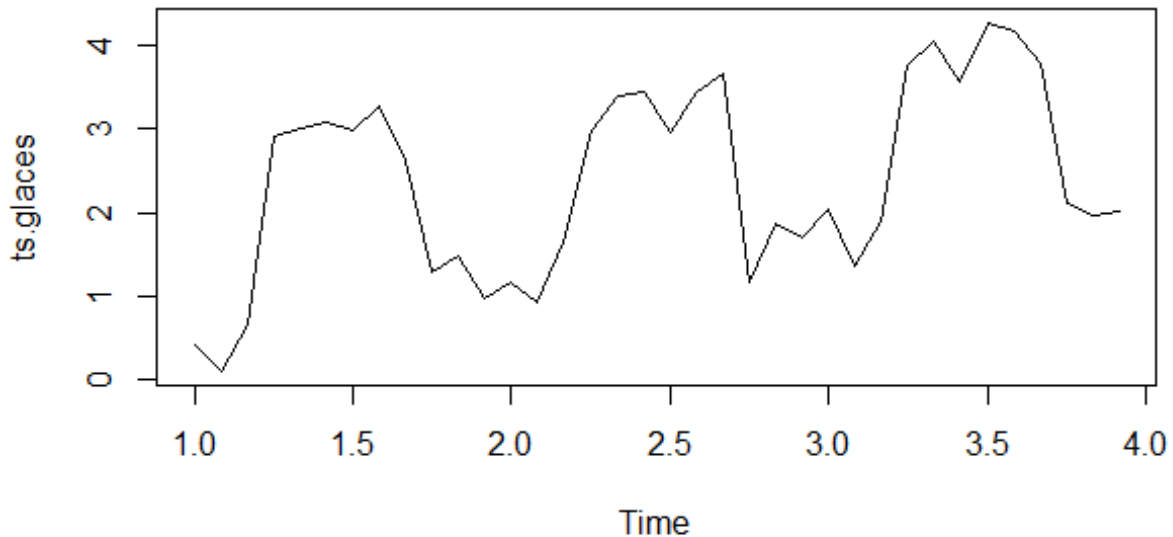
```
plot(z.dec)
```

```
# La fonction d'autocorrélation de la série (ACF) :
```

```
acf(z)
```

Exemple 2 :

FIGURE 1.5 – La fonction d'autocorrélation (ACF) de la série z .



```

#Transformation d'un vecteur de données en un objet « série chronologique »
ts(data, start, end, frequency...)
#Décomposition d'une série chronologique en une tendance, un effet saisonnier,
et un résidu
decompose(x, type = c("additive", "multiplicative"), filter = NULL)

Glaces <- c(0.42, 0.11, 0.65, 2.92, 3.01, 3.08, 2.98, 3.29, 2.63, 1.29, 1.49, 0.98, 1.16,
           0.93, 1.65, 2.98, 3.40, 3.46, 2.96, 3.45, 3.66, 1.17, 1.87, 1.71, 2.05, 1.36, 1.89,
           3.77, 4.06, 3.56, 4.27, 4.17, 3.79, 2.11, 1.96, 2.02)
ts.glaces<-ts(Glaces, start=1, frequency=12)
plot(ts.glaces)
glaces.dec <-decompose(ts.glaces)
plot(glaces.dec)
# La fonction d'autocorrélation de la série :
acf(ts.glaces)

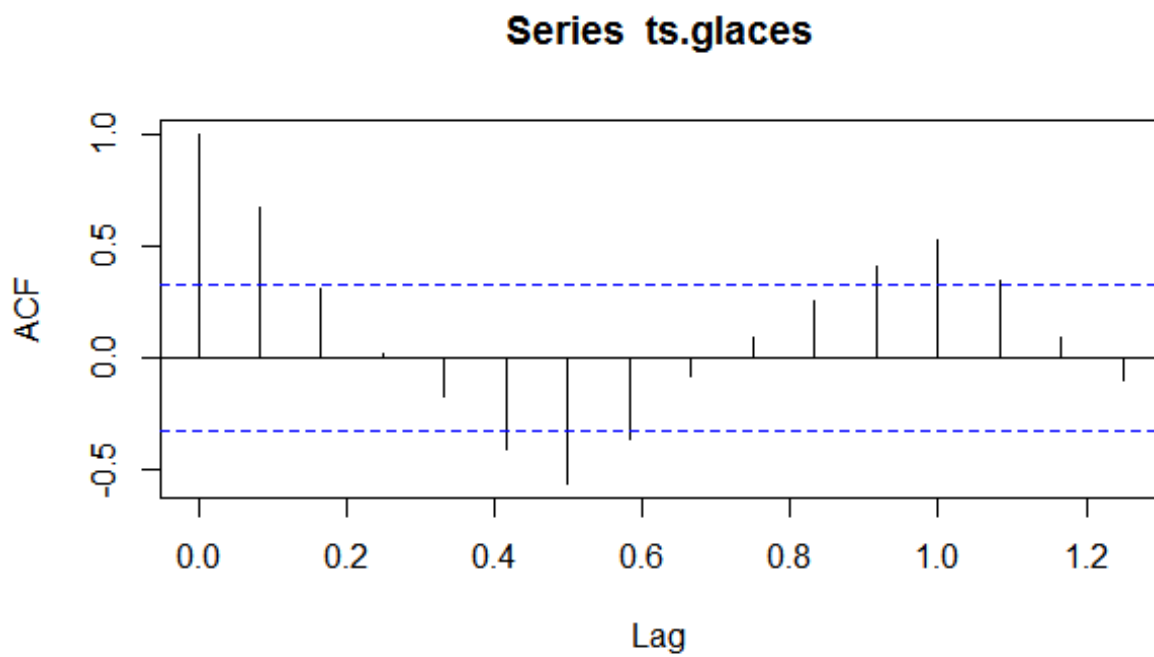
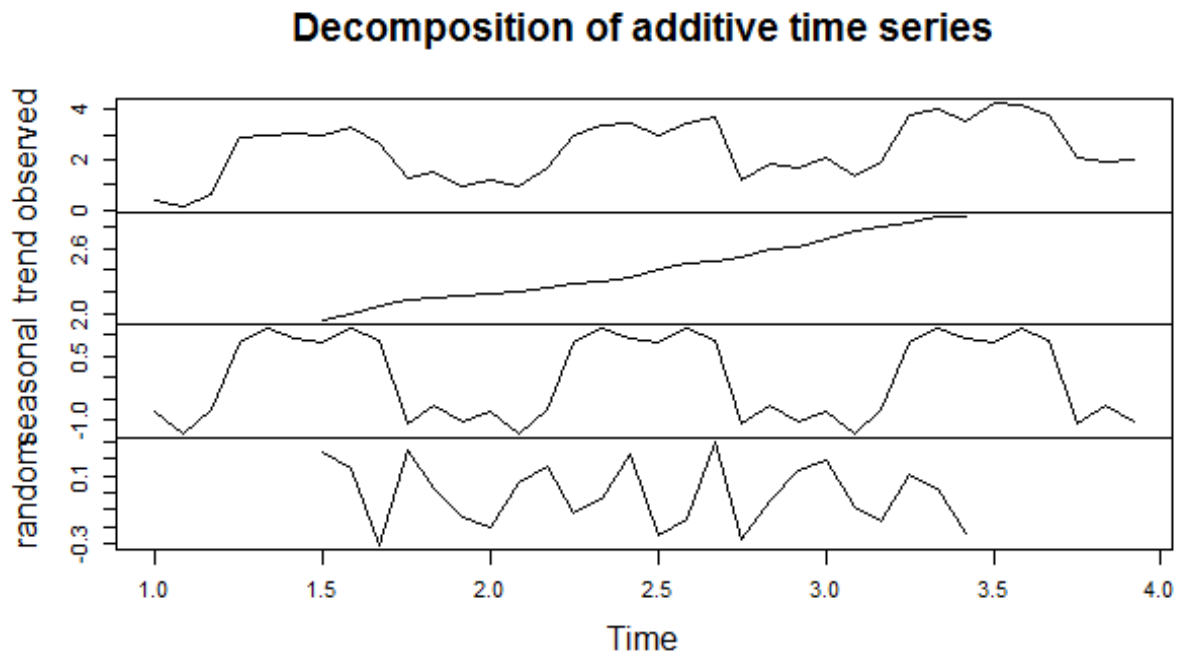
```

Exemple 3 :

```

# Simuler un bruit blanc gaussien  $\mathcal{N}(0, 1)$  de taille 100 :
v=rnorm(100, mean = 0, sd = 1)

```



```
plot(1 :100,v)

# la fonction d'autocorrélation :
acf(v,lag.max=24,type=c("correlation"))

# Simuler maintenant la série temporelle
```

$$X(t) = 0.5t + 2\epsilon_t$$

,avec $\epsilon_t \rightsquigarrow \mathcal{N}(0, 1)$

```
t=1 :100 ; x=0.5*t+2*v
plot(x)
plot(t,x)
# convertissons x en série temporelle
ts.x <- ts(x)
plot(ts.x)

# Simuler la série temporelle
```

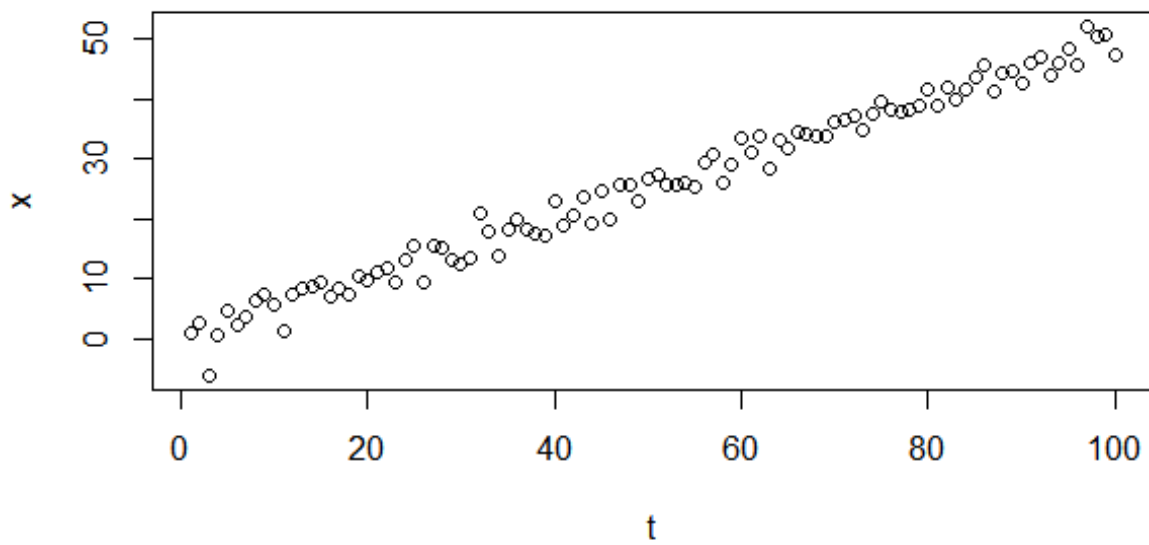
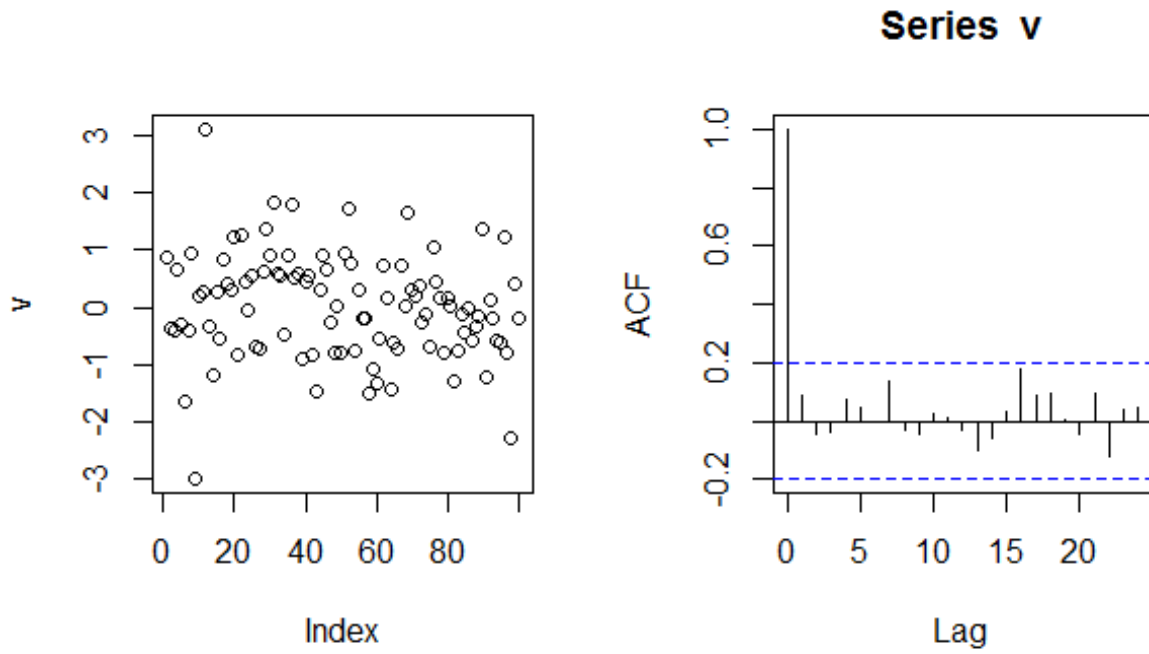
$$Y(t) = 0.5t + \epsilon_t + 2\cos(\pi t)$$

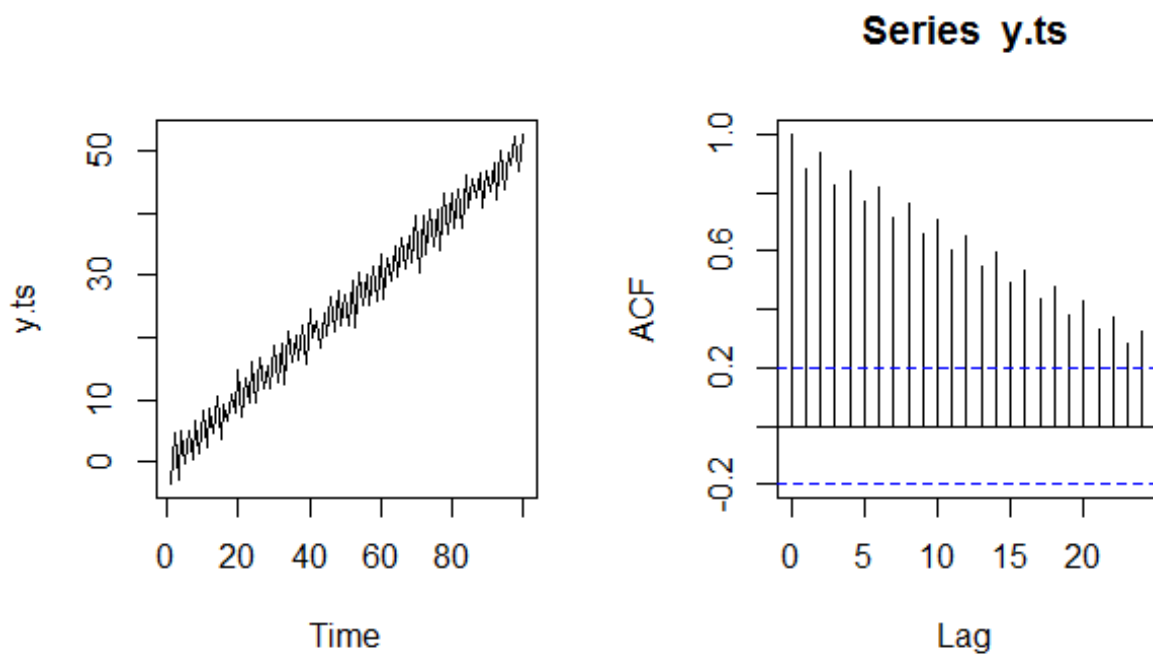
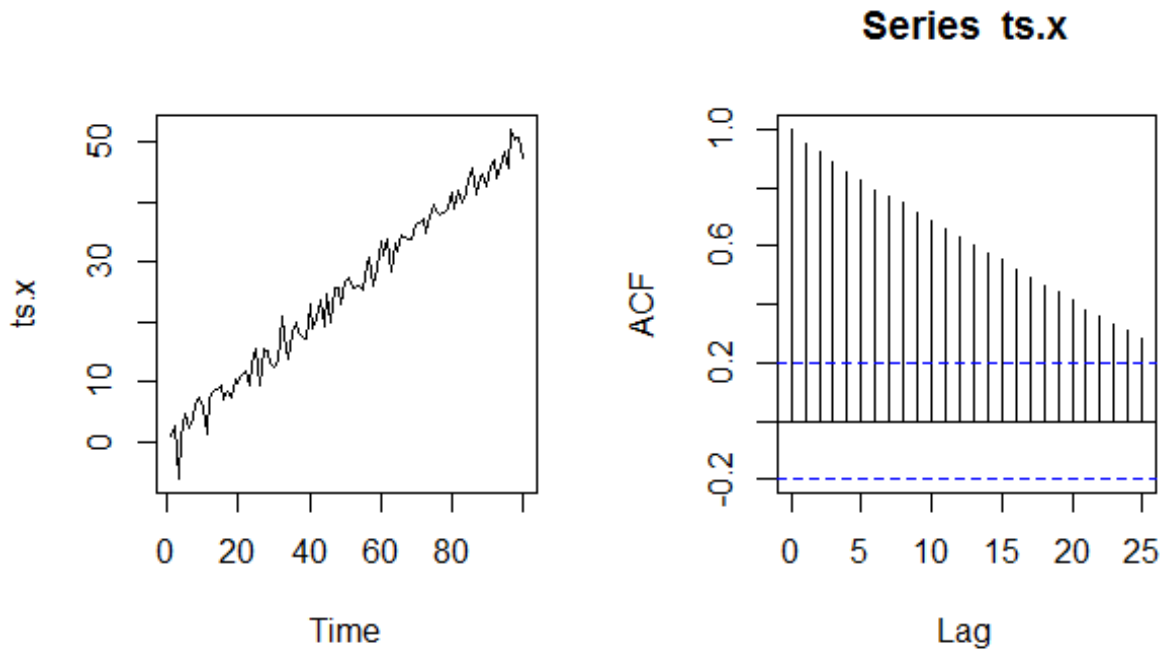
```
t=1 :100
v=rnorm(100, mean = 0, sd = 1)
y=0.5*t+v+3*cos(pi*t)
plot(y)
y.ts <- ts(y)
plot(y.ts)
acf(y.ts,lag.max=24,type=c("correlation"))
```

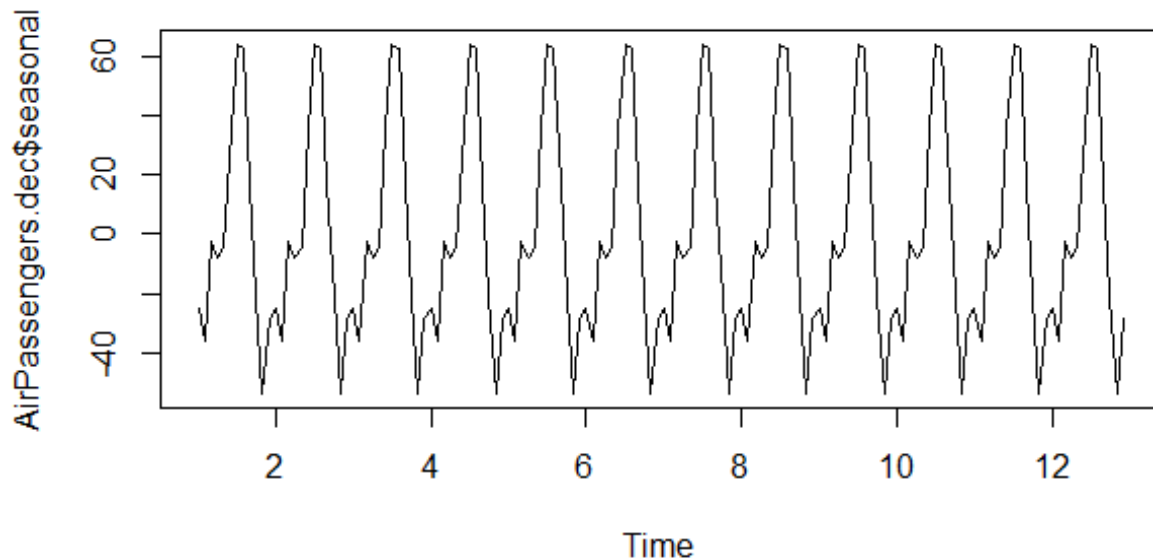
Exemple 4 :

• Données AirPassengers :

```
ts.AirPassengers <-ts(AirPassengers, start=1, frequency=12)
# serie_decomp <-decompose(serie,type=c("additive")) :
AirPassengers.dec <-decompose(ts.AirPassengers,type=c("additive"))
# la composante saisonnière :
AirPassengers.dec$seasonal
plot(AirPassengers.dec$seasonal)
```







la tendance :

```
AirPassengers.dec$trend
```

```
plot(AirPassengers.dec$trend)
```

la partie aléatoire stationnaire de la série :

```
AirPassengers.dec$random
```

```
plot(AirPassengers.dec$random)
```

Nuage de points (x_t, x_{t+h}) , $h = 1, 2, 3, 4$:

```
par(mfrow=c(2,2))
```

```
plot(stats : :lag(ts.AirPassengers,1),ts.AirPassengers,pch=20,cex=0.8)
```

```
plot(stats : :lag(ts.AirPassengers,2),ts.AirPassengers,pch=20,cex=0.8)
```

```
plot(stats : :lag(ts.AirPassengers,3),ts.AirPassengers,pch=20,cex=0.8)
```

```
plot(stats : :lag(ts.AirPassengers,4),ts.AirPassengers,pch=20,cex=0.8)
```

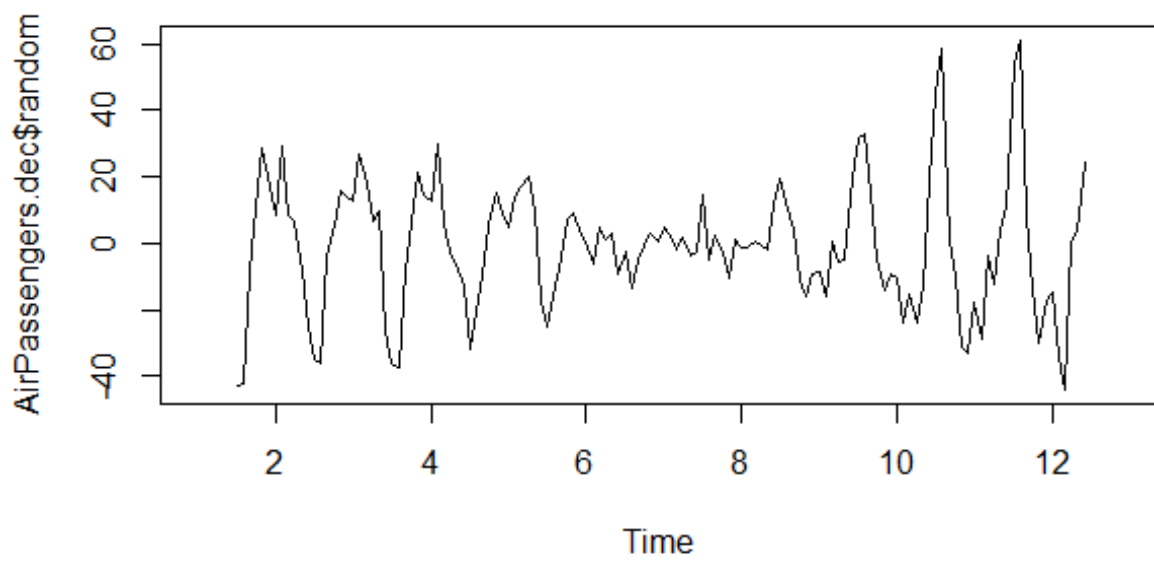
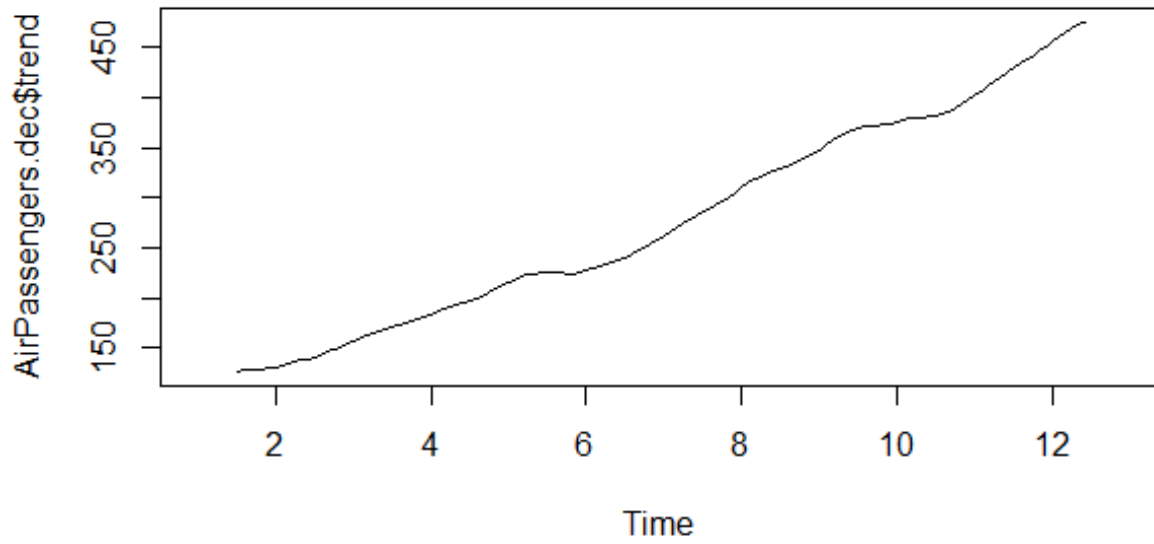
Méthode des différences :

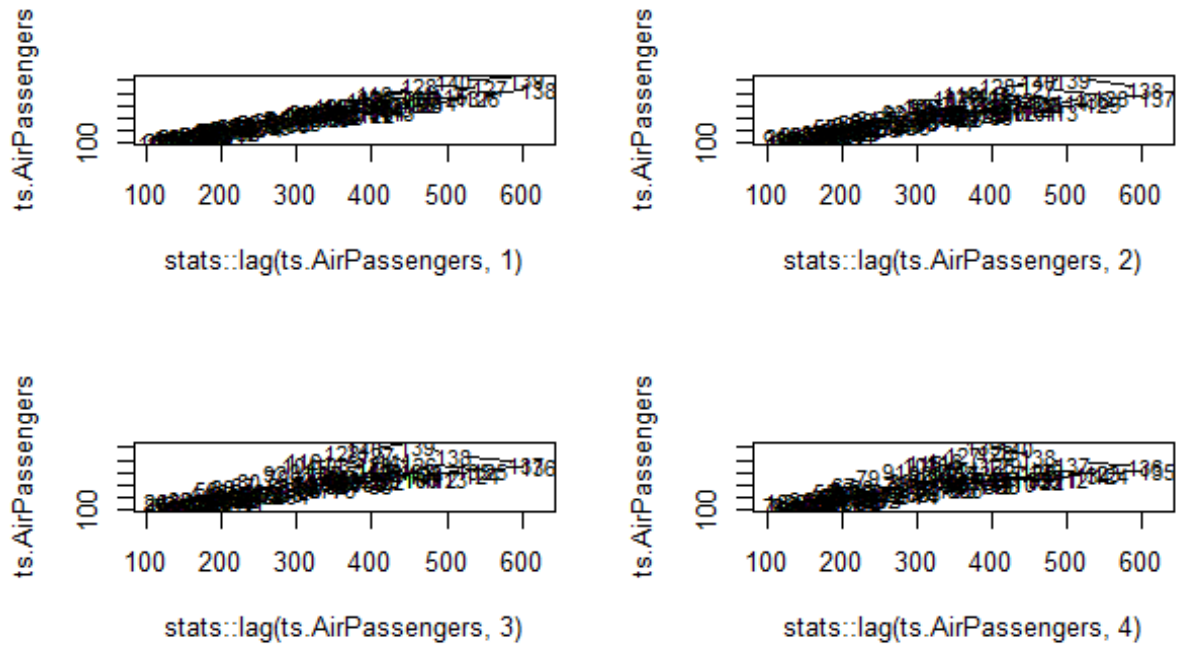
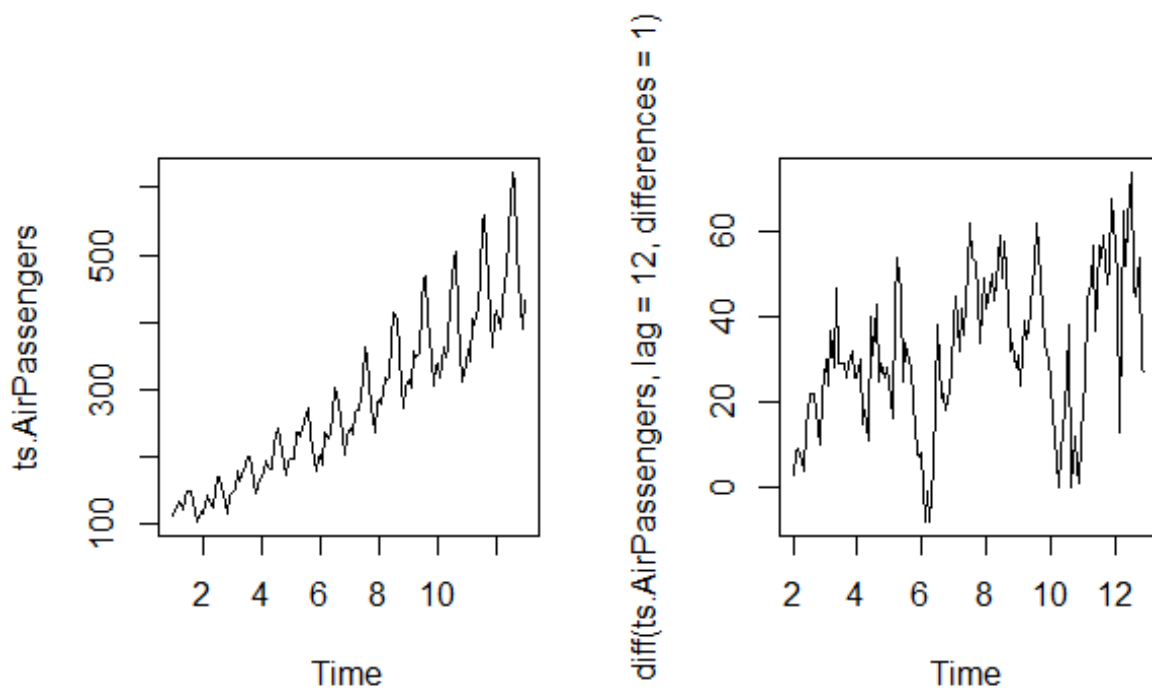
```
par(mfrow=c(1,2))
```

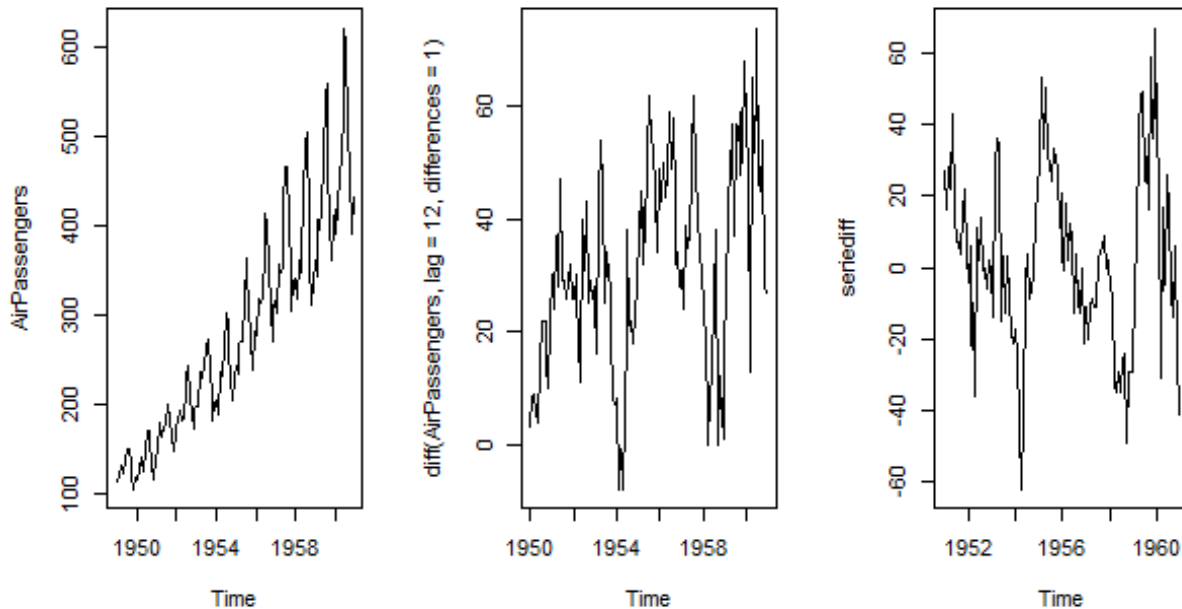
```
plot(ts.AirPassengers)
```

```
plot(diff(ts.AirPassengers,lag=12,differences=1))
```

- Appliquer l'opérateur de différence $\text{lag}=p=12$ et une différence de 1 (tendance linéaire) :



FIGURE 1.6 – Nuage de points (x_t, x_{t+h}) , pour $h = 1, 2, 3, 4$ 



```
data("AirPassengers")
par(mfrow=c(1,3))
plot(AirPassengers)
plot(diff(AirPassengers,lag = 12,differences = 1))
```

◦ Il semble qu'il reste une tendance légèrement croissante.

• Appliquons l'opérateur de différenciation avec une différence de 2 cette fois (tendance quadratique) :

```
seriediff=diff(AirPassengers,lag = 12,differences = 2)
plot(seriediff)
```

◦ Cela semble à l'oeil approximativement stationnaire.

• **Test sur la série résiduelle :**

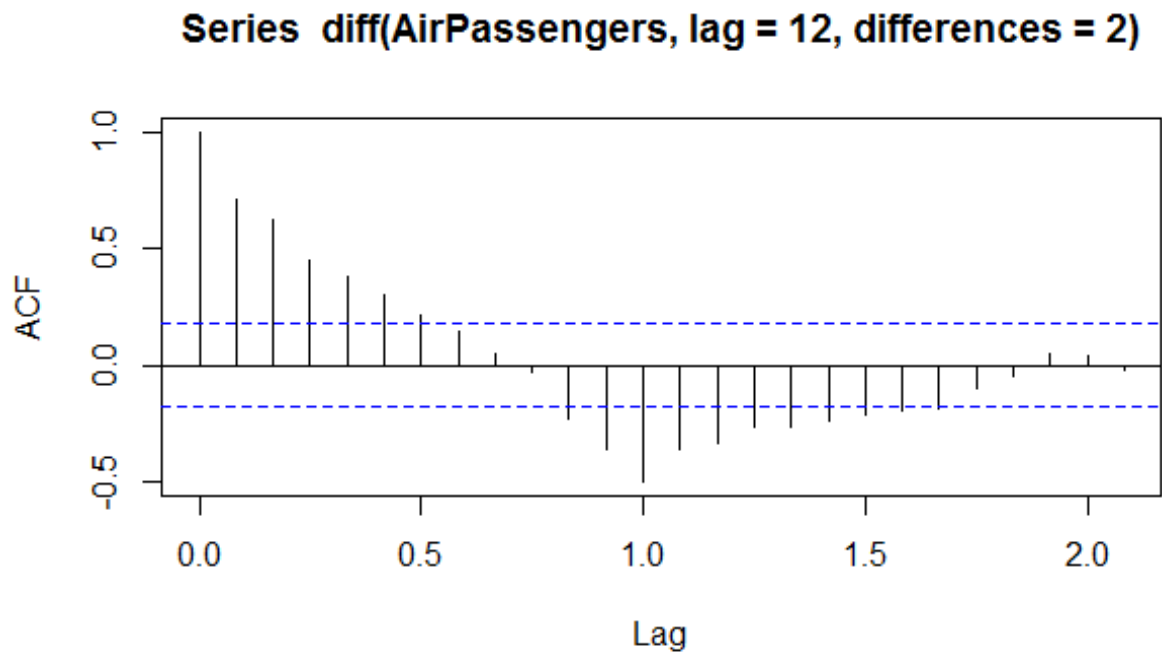
◦ La fonction **Box.test(serie,lag=**H**)** examine l'hypothèse nulle H_0 de nullité des **H** première auto-covariance, par défaut **H** est fixé à 1.

◦ si $p\text{-value} < \alpha$: nous rejetons l'hypothèse nulle.

```
> Box.test(seriediff,lag=1)
```

Box-Pierce test

data : seriediff



X-squared = 61.752, df = 1, p-value = 3.886e-15

◦ $p\text{-value} = 3.886e-15 < \alpha = 0.05$, on rejette l'hypothèse H_0 que la série est assimilable à un bruit blanc

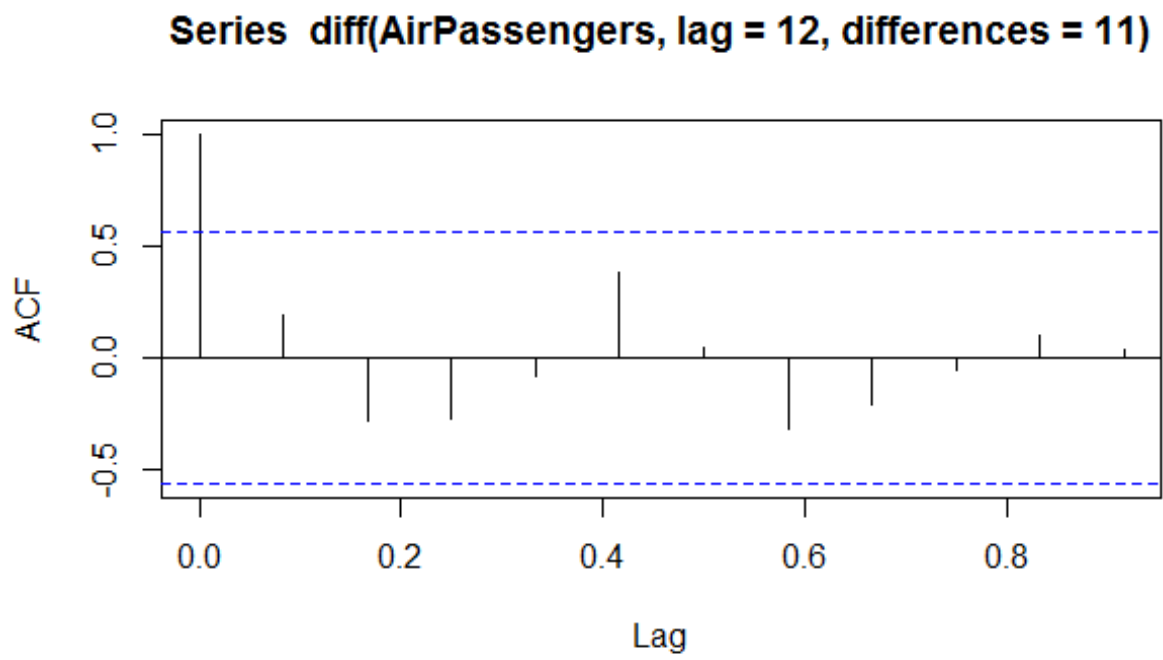
```
> Box.test(diff(AirPassengers,lag = 12,differences = 11),lag=1)
```

Box-Pierce test

```
data : diff(AirPassengers, lag = 12, differences = 11)
```

X-squared = 0.41423, df = 1, p-value = 0.5198

◦ ◦ $p\text{-value} = 0.5198 > \alpha = 0.05$, on accepte l'hypothèse H_0 que la série est assimilable à un bruit blanc.



Lissage et prévision

2.1 Lissage par moyennes mobiles

2.1.1 Moyenne mobile

• La moyenne mobile permet de décomposer la série initiale en fonction d'une tendance et d'une composante saisonnière.

- Moyenne mobile $k = 2m + 1$: (k impaire)

$$Z_t = \frac{x_{t-m} + \dots + x_t + \dots + x_{t+m}}{2m + 1}, \quad t = m + 1, \dots, n - m$$

et on note $MM(k)$ la série des moyennes mobiles d'ordre k de la série $(x_t)_{t=1, \dots, n}$.

- Moyenne mobile $k = 2m$: (k paire)

$$Z_t^a = \frac{x_{t-m} + \dots + x_t + \dots + x_{t+m-1}}{2m}, \quad t = m + 1, \dots, n - m + 1$$

et

$$Z_t^b = \frac{x_{t-m+1} + \dots + x_t + \dots + x_{t+m}}{2m}, \quad t = m, \dots, n - m$$

alors la **moyenne mobile centrée** $MMC(k)$ est donnée par :

$$Z_t = \frac{Z_t^a + Z_t^b}{2} = \frac{0.5 * x_{t-m} + x_{t-m+1} + \dots + x_{t+m-1} + 0.5 * x_{t+m}}{2}, \quad t = m + 1, \dots, n - m$$

et on note $MMC(k)$ la série des moyennes mobiles d'ordre $k = 2m$ de la série $(x_t)_{t=1, \dots, n}$.

Exemple 9. Soit la série (y_t) . Le tableau suivant présente les moyennes mobiles d'ordre 2, 3 et 4 de la série $(y_t)_{t=1, \dots, 16}$.

y_i	$MM(2)$	$MMC(2)$	$MM(3)$	$MM(4)$	$MMC(4)$
30	22.5				
15	10	16.25	16.67	20	
5	17.5	13.75	16.67	21.5	20.75
30	33	25.25	23.67	22.25	21.875
36	27	30	28	23.25	22.75
18	13.5	20.25	21	24.75	24
9	22.5	18	21	27	25.875
36	40.5	31.5	30	26.25	26.625
45	30	35.25	32	26.50	26.375
15	12.5	21.25	23.33	32.5	29.5
10	35	23.75	28.33	33.25	32.875
60	54	44.50	39.33	33.50	33.375
48	32	43	41.33	33	33.25
16	12	22	24	36	34.5
8	40	26	32		
72					

2.1.2 Lissage par moyennes mobiles

• **Objectif :** L'objectif d'un lissage par moyenne mobile est de faire apparaître l'allure de la tendance :

- en éliminant la composante saisonnière.
- en atténuant les fluctuations irrégulières.

Propriétés :

- Élimination de la composante saisonnière :
 - Si la série chronologique $(x_t)_{t=1,\dots,n}$ possède une composante saisonnière de période p , alors l'application d'une moyenne mobile d'ordre p supprime cette saisonnalité.
 - La série $MM(p)$ ou $MMC(p)$ ne possède plus de composante saisonnière de période p .
 - On se servira donc d'une moyenne mobile d'ordre p pour éliminer une composante saisonnière de période p .

· Soit $(s_t)_{t=1,\dots,n}$ une série purement périodique p . On a

$$s_1 + s_2 + \dots + s_p = 0,$$

et

$$\forall t, s_t = s_{t+p}.$$

La moyenne mobile d'ordre p est donnée par

$$MM(p)_t = \frac{s_t + s_{t+1} + \dots + s_{t+p-1}}{p}, \quad \text{pour } t = 1, \dots, n-p$$

et puisque la série est périodique de période p , a une permutation circulaire près des p termes

$$MM(p)_t = \frac{s_t + s_{t+1} + \dots + s_{t+p-1}}{p} = 0.$$

• Atténuation des fluctuations irrégulières :

◦ Une moyenne mobile atténuée (réduire le plus possible) l'amplitude des fluctuations irrégulières d'une chronique.

◦ Plus l'ordre de la moyenne mobile est élevé, et plus cette atténuation est importante.

◦ Donc en appliquant une moyenne mobile sur une série chronologique on obtient un lissage de la série.

• Effet d'une moyenne mobile sur la tendance :

◦ Une moyenne mobile (d'ordre quelconque) ne modifie pas une tendance constante.

◦ Une moyenne mobile simple (d'ordre impair quelconque) ou centrée (d'ordre quelconque) ne modifie pas une tendance linéaire.

◦ Lorsque la tendance n'est pas aussi régulière, les détails disparaîtront d'autant plus que l'ordre de la moyenne mobile sera grand.

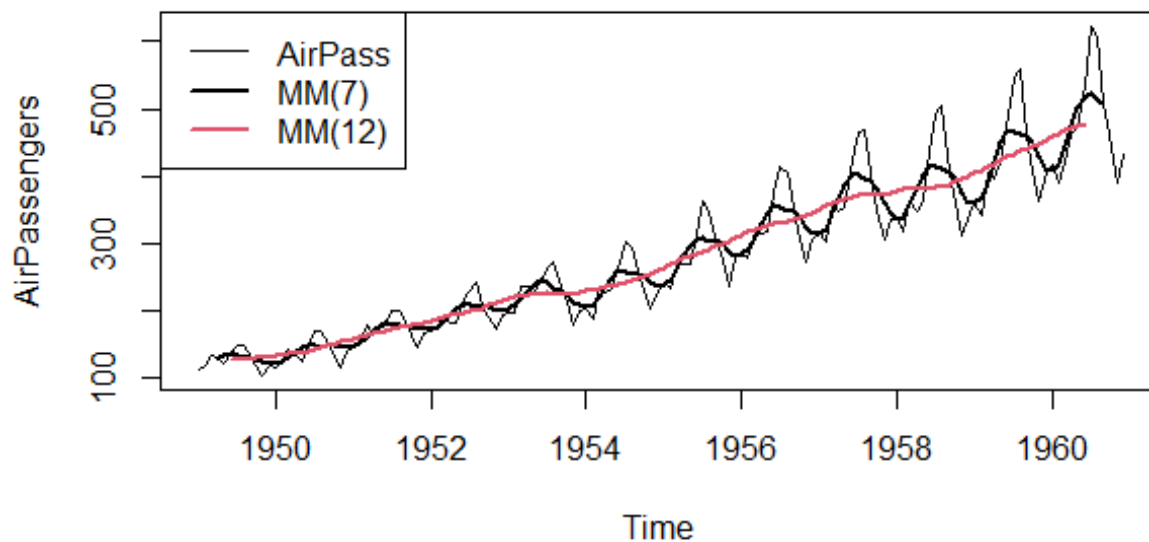
· Soit $(\epsilon_t)_{t=1,\dots,n}$ une suite de variables aléatoires indépendantes de moyenne nulle et de variance σ^2 . On a $MM(k)$ de cette suite est donnée par :

$$MM(k)_t = \frac{1}{k} \sum_{i=1}^k \epsilon_{t+i},$$

alors

$$Var \{MM(k)_t\} = \frac{1}{k^2} \sum_{i=1}^k Var \{\epsilon_{t+i}\} = \frac{1}{k^2} \sum_{i=1}^k \sigma^2 = \frac{\sigma^2}{k},$$

La variance est donc divisée par un facteur k , c'est-à-dire que l'écart-type est divisée par un facteur \sqrt{k} , alors $MM(k)$ réduit l'amplitude de ces fluctuations irrégulières d'un facteur \sqrt{k} .



Exemple 10. *Série airpass : nombre mensuel de passagers aériens, en milliers, de janvier 1949 à décembre 1960*

2.1.3 La décomposition additive

- La désaisonnalisation

- Lissage par moyennes mobiles : On notera (\tilde{f}_t) la série obtenue par un lissage par la méthode des Moyennes Mobiles :

$$\tilde{f}_t = \begin{cases} MMC(p)_t & \text{si } p \text{ est pair} \\ MM(p)_t & \text{si } p \text{ est impair} \end{cases}$$

p désigne la période de la composante saisonnière.

- Construction de la série des différences : On calcule la série des différences, soit

$$D_t = x_t - \tilde{f}_t \text{ pour } t = m + 1, \dots, n - m.$$

avec $p = 2m$ (p est supposée pair).

- Calcul des coefficients saisonniers non centrés :

· supposons que $n = p * K_0$, on calcule les p coefficients saisonniers non centrés (\tilde{s}_t) :

$$\tilde{s}_t = \begin{cases} \frac{1}{K_0-1} \sum_{k=1}^{K_0-1} D_{t+p(k-1)}, & \text{pour } j > m \\ \frac{1}{K_0-1} \sum_{k=1}^{K_0-1} D_{t+pk}, & \text{pour } j \leq m \end{cases}$$

○ Centrage des coefficients saisonniers :

· On calcule la moyenne arithmétique des p coefficients saisonniers :

$$\bar{s} = \frac{1}{p} \sum_{t=1}^p \tilde{s}_t.$$

· on centre les coefficients saisonniers (\tilde{s}_t)_{t=1,...,p} on obtient :

$$\hat{s}_t = \tilde{s}_t - \bar{s}, \text{ pour } t = 1, \dots, p.$$

○ Construction de la série corrigée des variations saisonnières : La série corrigée des variations saisonnières (CVS_t)_{t=1,...,n} ou la série **désaisonnalisée** est donnée par

$$CVS_t = x_t - \hat{s}_t, \text{ pour } t = 1, \dots, n.$$

● La série lissée des prévisions :

○ Ajustement d'une courbe de tendance :

· à partir de la série est corrigée des variations saisonnières (CVS_t)_{t=1,...,n}, on peut ajuster une courbe de tendance de la forme

$$x_t = f_\theta(t),$$

○ Construction de la série lissée des prévisions :

· La série lissée des prévisions (\hat{x}_t)_{t=1,...,n} est utile pour **la prévision**, est donnée par

$$\hat{x}_t = \hat{f}_t + \hat{s}_t, \text{ pour } t = 1, \dots, n.$$

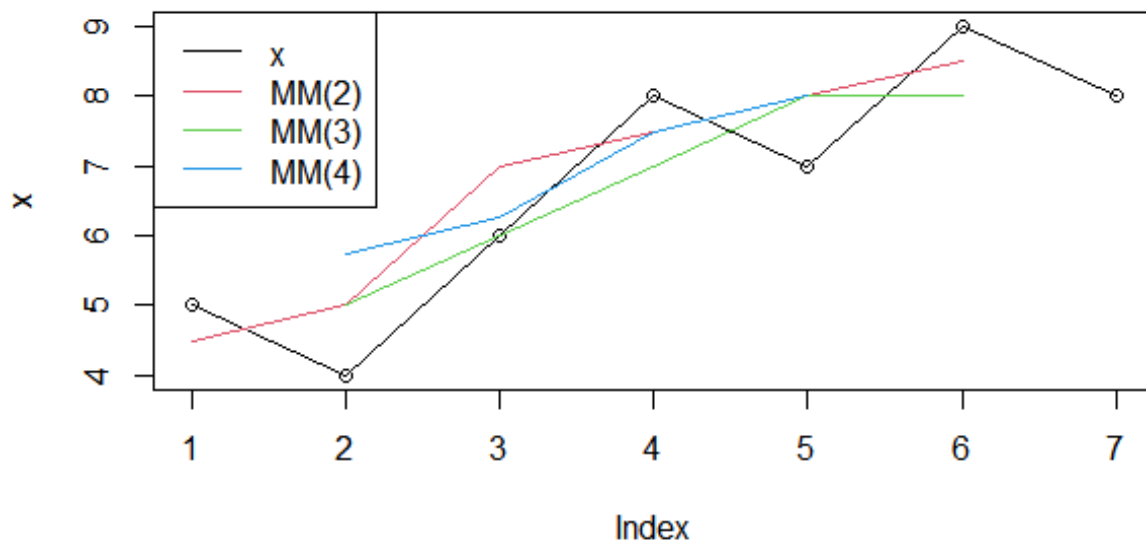
avec $\hat{f}_t = f_{\hat{\theta}}(t)$.

· l'erreur de prévision associée (\hat{e}_t)_{t=1,...,n} est définie par

$$\hat{e}_t = x_t - \hat{x}_t \text{ pour } t = 1, \dots, n.$$

Remarque 11. La qualité de la modélisation de la chronique est étudié a l'aide de représentations graphiques et à l'aide de critères numériques, comme l'erreur quadratique moyenne (MSE)

$$MSE = \frac{1}{n} \sum_{t=1}^n (x_t - \hat{x}_t)^2,$$



l'erreur absolue moyenne (MAE)

$$MSE = \frac{1}{n} \sum_{t=1}^n |x_t - \hat{x}_t|.$$

Exemple 12. Soit la série $x = (5, 4, 6, 8, 7, 9, 8)$ Le tableau suivant présente les moyennes mobiles d'ordre 2, 3 et 4 de la série $(x_t)_{t=1, \dots, 7}$.

x_t	$MM(2)_t$	$MM(3)_t$	$MM(4)$
5			
4	$\frac{5+4}{2} = 4.5$	$\frac{5+4+6}{3} = 5$	$\frac{5+4+6+8}{4} = 5.75$
6	$\frac{4+6}{2} = 5.0$	$\frac{4+6+8}{3} = 6$	$\frac{4+6+8+7}{4} = 6.25$
8	$\frac{6+8}{2} = 7.0$	$\frac{6+8+7}{3} = 7$	$\frac{6+8+7+9}{4} = 7.50$
7	$\frac{8+7}{2} = 7.5$	$\frac{8+7+9}{3} = 8$	$\frac{8+7+9+8}{4} = 8.00$
9	$\frac{7+9}{2} = 8.0$	$\frac{7+9+8}{3} = 8$	
8	$\frac{9+8}{2} = 8.5$		

2.2 Lissages exponentiels

$(x_t)_{t=1, \dots, n}$ une série chronologique on se situe à la date n (base de la prévision) et on souhaite prévoir la valeur $x_{n,h}$ non encore observée, cette prévision est notée $\hat{x}_{t,h}$. il s'agit

d'ajuster à la chronique, localement, une fonction simple :

- **une constante** dans le **lissage exponentiel simple**,
- **une droite** dans le **lissage exponentiel double**,
- **des fonctions polynomiales** ou **périodiques** dans les **lissages plus généraux**.

2.2.1 Lissage exponentiel simple

• On se donne α appelé **constante de lissage**, avec $0 < \alpha < 1$, et on définit (la prévision ne dépend pas de h) la prévision $\hat{x}_{t,h}$ est basée sur toute l'histoire de la chronique, de sorte que, plus on s'éloigne de la base n de la prévision, moins l'influence des observations correspondantes est importante, cette décroissance de l'influence est de type exponentiel, et on définit (la prévision ne dépend pas de h)

$$\hat{x}_{n,h} = (1 - \alpha) \sum_{j=0}^{n-1} \alpha^j x_{n-j}.$$

- α proche de 1 : prise en compte de tout le passé.
 - α proche de 0 : prise en compte d'avantage des valeurs récentes (plus sensible aux fluctuations).
- La définition précédente vérifie les formules de récurrences suivantes :

$$\begin{aligned} \hat{x}_{n,h} &= (1 - \alpha) \sum_{j=0}^{n-1} \alpha^j x_{n-j} \\ &= (1 - \alpha) x_n + (1 - \alpha) \sum_{j=1}^{n-1} \alpha^j x_{n-j} \\ &= (1 - \alpha) x_n + (1 - \alpha) \sum_{j=0}^{n-2} \alpha^{j+1} x_{n-j-1} \\ &= (1 - \alpha) x_n + \underbrace{\alpha (1 - \alpha) \sum_{j=0}^{n-1-1} \alpha^j x_{n-1-j}}_{\alpha \hat{x}_{(n-1),h}} \\ &= (1 - \alpha) x_n + \alpha \hat{x}_{(n-1),h} \\ &= \hat{x}_{(n-1),h} + (1 - \alpha) (x_n - \hat{x}_{(n-1),h}), \end{aligned}$$

en particulier on peut choisir $\hat{x}_{1,h} = x_1$ comme valeur initiale de cette formule récursive.

- On peut remarquer que valeur initiale a peu d'influence pour n assez grand.
- Interprétation :

On ajuste au voisinage de n une constante a , en effet la valeur de a qui minimise

$$\sum_{j=0}^{n-1} \alpha^j (x_{n-j} - a)^2$$

est fournie par

$$\hat{a}_n = \frac{1 - \alpha}{1 - \alpha^n} \sum_{j=0}^{n-1} \alpha^j x_{n-j},$$

et si on cherche la valeur de a qui minimise

$$\sum_{j=0}^{\infty} \alpha^j (x_{n-j} - a)^2,$$

la solution est fournie par

$$\hat{x}_{n,h} = (1 - \alpha) \sum_{j=0}^{n-1} \alpha^j x_{n-j}$$

• **Choix de la constante de lissage :**

◦ En pratique le lissage exponentiel simple est assez peu utilisé parce cette méthode n'est pas bonne en présence de tendance, de composante saisonnière ou de fluctuations de hautes fréquences.

◦ On peut aussi utiliser le critère des moindres carrés pour déterminer un α adapté à la prévision d'une série donnée, alors pour un horizon h on cherche la valeur de α qui minimise

$$\sum_{t=1}^{n-h} \left(x_{t+h} - (1 - \alpha) \sum_{j=0}^{n-1} \alpha^j x_{n-j} \right)^2$$

2.2.2 Lissage exponentiel double

• On ajuste au voisinage de n une droite d'équation

$$y_t = a_1 + a_2 (t - n).$$

• Le prédicteur sera

$$\hat{x}_{n,h} = \hat{a}_1 + \hat{a}_2 h.$$

Les coefficients \hat{a}_1, \hat{a}_2 sont solution de

$$\begin{aligned} & \min_{a_1, a_2} \sum_{j=0}^{n-1} \alpha^j (x_{n-j} - (a_1 + a_2 (n - j - n)))^2 \\ &= \min_{a_1, a_2} \sum_{j=0}^{n-1} \alpha^j (x_{n-j} - (a_1 - a_2 j))^2 = \min_{a_1, a_2} \mathcal{C}(a_1, a_2). \end{aligned}$$

En remplaçant, comme dans le lissage simple, la sommation finie par une somme de 0 à ∞ , le problème précédent devient

$$\min_{a_1, a_2} \sum_{j=0}^{\infty} \alpha^j (x_{n-j} - (a_1 - a_2 j))^2.$$

• On annule les dérivées partielles de la fonction $\mathcal{C}(a_1, a_2)$ par rapport à a_1, a_2 et on vérifie à l'aide des dérivées secondes que l'on a bien un minimum, on obtient

$$\begin{aligned} \frac{\partial \mathcal{C}(a_1, a_2)}{\partial a_1} &= -2 \sum_{j=0}^{n-1} \alpha^j (x_{n-j} - (a_1 - a_2 j))^2, \\ \frac{\partial \mathcal{C}(a_1, a_2)}{\partial a_2} &= 2 \sum_{j=0}^{n-1} j \alpha^j (x_{n-j} - (a_1 - a_2 j))^2. \end{aligned}$$

Comme $\sum_{j=0}^{\infty} \alpha^j = \frac{1}{1-\alpha}$, $\sum_{j=0}^{\infty} j \alpha^j = \frac{\alpha}{(1-\alpha)^2}$ et $\sum_{j=0}^{\infty} j^2 \alpha^j = \frac{\alpha(1+\alpha)}{(1-\alpha)^3}$, en annulant les dérivées, on obtient

$$\begin{aligned} (1-\alpha) \sum_{j=0}^{\infty} \alpha^j x_{n-j} - a_1 + a_2 \frac{\alpha}{(1-\alpha)} &= 0, \\ (1-\alpha)^2 \sum_{j=0}^{\infty} j \alpha^j x_{n-j} - a_1 \alpha + a_2 \frac{\alpha(1+\alpha)}{(1-\alpha)} &= 0, \end{aligned}$$

et donc les solutions sont données par :

$$\begin{cases} \hat{a}_1(n) = 2L_1(n) - L_2(n) \\ \hat{a}_2(n) = \frac{(1-\alpha)}{\alpha} (L_1(n) - L_2(n)) \end{cases},$$

avec

$$\begin{cases} L_1(n) = (1-\alpha) \sum_{j=0}^{n-1} \alpha^j x_{n-j} = (1-\alpha) x_n + \alpha L_1(n-1), & (\text{série lissée}) \\ L_2(n) = (1-\alpha) \sum_{j=0}^{n-1} \alpha^j L_1(n-j) = (1-\alpha) L_1(n) + \alpha L_2(n-1), & (\text{série lissée 2 fois}) \\ = (1-\alpha)^2 \sum_{j=0}^{n-1} \sum_{i=0}^{n-j-1} \alpha^{j+i} x_{n-(i+j)}. \end{cases}$$

Soit encore

$$\begin{cases} L_1(n) = \hat{a}_1(n) - \frac{\alpha}{1-\alpha} \hat{a}_2(n), \\ L_2(n) = \hat{a}_1(n) - \frac{2\alpha}{1-\alpha} \hat{a}_2(n). \end{cases}$$

• Formules récursives de mises à jour :

On peut en déduire des formules récursives de mise à jour

$$\begin{aligned} \hat{a}_1(n) &= \hat{a}_1(n-1) + \hat{a}_2(n-1) + (1-\alpha^2) (x_n - \hat{x}_{n-1,1}), \\ \hat{a}_2(n) &= \hat{a}_1(n-1) + (1-\alpha^2) (x_n - \hat{x}_{n-1,1}). \end{aligned}$$

On peut prendre comme valeurs initiales $\hat{a}_1(0) = x_1$ et $\hat{a}_2(0) = x_2 - x_1$.

2.2.3 Méthode de Holt-Winters

- Cette approche a pour but d'améliorer et de généraliser le lissage exponentiel simple.

Nous étudions plusieurs cas particuliers de cette méthode :

- ajustement d'une droite affine (sans saisonnalité).
- ajustement d'une droite affine + une composante saisonnière.
- ajustement d'une constante + une composante saisonnière.

- **Méthode non saisonnière :**

On ajuste au voisinage de l'instant n une droite

$$x_t = a_1 + (t - n) a_2.$$

De la même façon que pour le lissage exponentiel double, l'ajustement se fait de façon linéaire au voisinage de n , Soit $0 < \beta < 1$ et $0 < \gamma < 1$ deux constantes fixées et les formules de mise à jour

$$\begin{aligned}\hat{a}_1(n) &= (1 - \beta) (\hat{a}_1(n-1) + \hat{a}_2(n-1)) + \beta x_n, \\ \hat{a}_2(n) &= \gamma (\hat{a}_1(n) - \hat{a}_1(n-1)) + (1 - \gamma) \hat{a}_2(n-1).\end{aligned}$$

La prévision prend la forme

$$\hat{x}_{n,h} = \hat{a}_1(n) + h\hat{a}_2(n).$$

- **Méthode saisonnière additive :**

On ajuste au voisinage de n , la chronique

$$x_t = a_1 + (t - n) a_2 + s_t ,$$

où s_t est une composante périodique de période p , et soit $0 < \beta < 1$, $0 < \gamma < 1$, $0 < \delta < 1$, alors les formules récursives de mise à jour sont :

$$\begin{aligned}\hat{a}_1(n) &= \beta (x_n - \hat{s}_{n-p}) + (1 - \beta) (\hat{a}_1(n-1) + \hat{a}_2(n-1)) , \\ \hat{a}_2(n) &= \gamma (\hat{a}_1(n) - \hat{a}_1(n-1)) + (1 - \gamma) \hat{a}_2(n-1) , \\ \hat{s}_n &= \delta (x_n - \hat{a}_1(n)) + (1 - \delta) \hat{s}_{n-p}.\end{aligned}$$

La prévision prend la forme

$$\hat{x}_{n,h} = \hat{a}_1(n) + h\hat{a}_2(n) + \hat{s}_{n+h-jp} , \text{ pour } (j-1)p + 1 \leq h \leq jp \quad (j = 1, 2, \dots).$$

Les valeurs initiales suivantes permettent de commencer le calcul récursif des prévisions à partir de l'instant $p + 2$:

$$\begin{aligned}\hat{a}_1(p+1) &= x_{p+1} , \\ \hat{a}_2(p+1) &= \frac{x_{p+1} - x_1}{p} , \\ \hat{s}_j &= x_j - (x_1 + (p-1)\hat{a}_2(p+1)) \quad \text{pour } j = 1, \dots, p.\end{aligned}$$

Travaux Pratiques

- **Data Frames** :

L'objet permettant de stocker des tableaux de données en R est la data frame. Il s'agit d'une liste de vecteur

de même taille, potentiellement de type (numeric, character) différent.

```
x1 <-c(1 :5)
```

```
x2 <-c("a", "b", "c", "d", "e")
```

```
x3 <-c(rep("TRUE", 3), rep("FALSE", 2))
```

```
data_test <-data.frame(x1, x2, x3)
```

○ Par défaut le nom des éléments de la liste est celui des vecteurs la constituant. Il est possible de les modifier

via la fonction **names()** :

```
names(data_test) <-c("var1", "var2", "var3")
```

```
data_test
```

	<i>var1</i>	<i>var2</i>	<i>var3</i>
1	1	<i>a</i>	<i>TRUE</i>
2	2	<i>b</i>	<i>TRUE</i>
3	3	<i>c</i>	<i>TRUE</i>
4	4	<i>d</i>	<i>FALSE</i>
5	5	<i>e</i>	<i>FALSE</i>

Lorsqu'on affiche la data frame, la première ligne de la table, appelée header, contient le nom des colonnes.

Les observations constituent les lignes de la table.

Pour connaître les dimensions de la data frame, utiliser **dim()**, **ncol()** ou **nrow()**.

Il est souvent pratique de visualiser les premières lignes de la table à l'aide de `head()`.

```
head(data_test, 2)
```

Une fonction très utile, notamment lorsqu'on manipule des data frame est la fonction `str`, elle permet

d'obtenir le type de chacune des variables de la data frame :

```
str(data_test)
```

La fonction `summary()` permet de calculer des statistiques de base de la data frame :

```
summary(data_test)
```

Le type d'objet obtenu lorsqu'on sélectionne une variable de la data frame dépend de l'instruction :

```
x <-data_test$var1
```

Exemple 1 :

```
filt1.nit <-filter(AirPassengers,filter=rep(1/7,7))
```

```
filt2.nit <-filter(AirPassengers,filter=rep(1/12,12))
```

```
par(mfrow=c(1,2))
```

```
ts.plot(AirPassengers,filt1.nit,col=c(1,2),lwd=c(1,2))
```

```
legend("topleft",legend=c("AirPass","MM(7)"),col=c(1,2),lwd=c(1,2))
```

```
ts.plot(AirPassengers,filt2.nit,col=c(1,2),lwd=c(1,2))
```

```
legend("topleft",legend=c("AirPass","MM(12)"),col=c(1,2),lwd=c(1,2))
```

Exemple 2 :

```
# Calculer les séries des moyennes-mobiles d'ordre 2, 3 et 4 de la série  $(y_t)_{t=1,\dots,8}$ .
```

```
> y=c(30,15,5,30,36,18,9,36,45,15,10,60,48,16,8,72)
```

```
> MM2 <- filter(y,filter=rep(1/2,2))
```

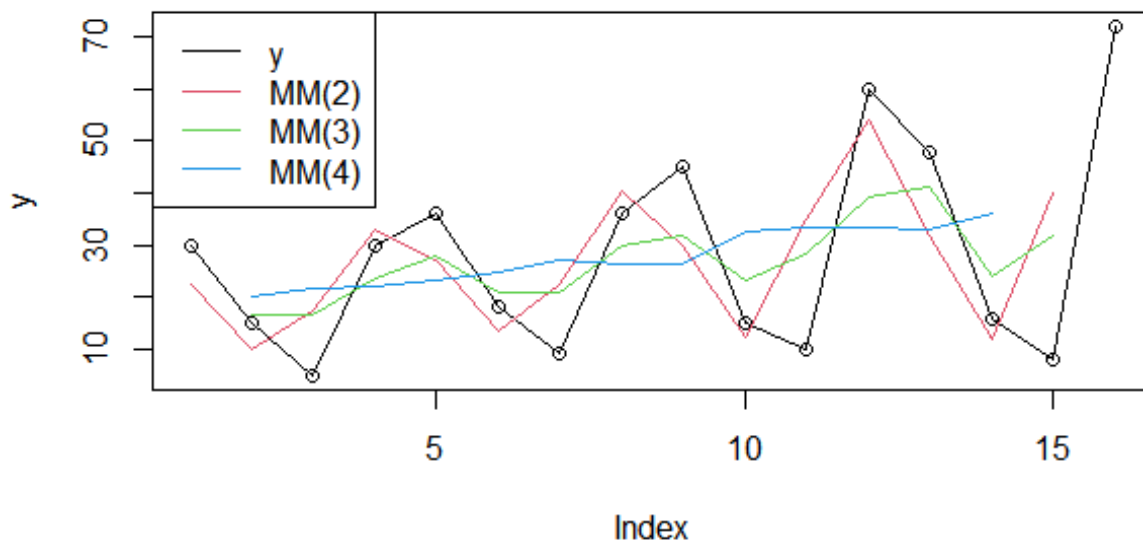
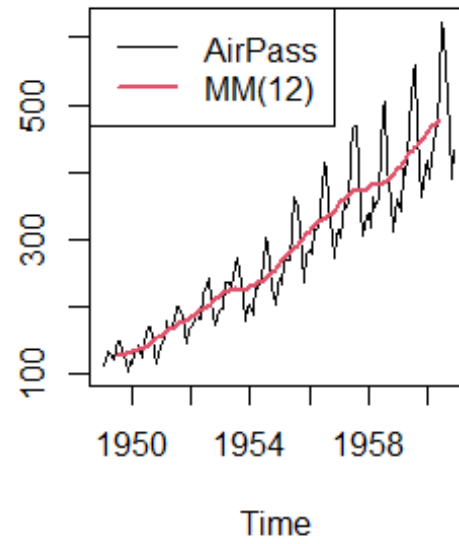
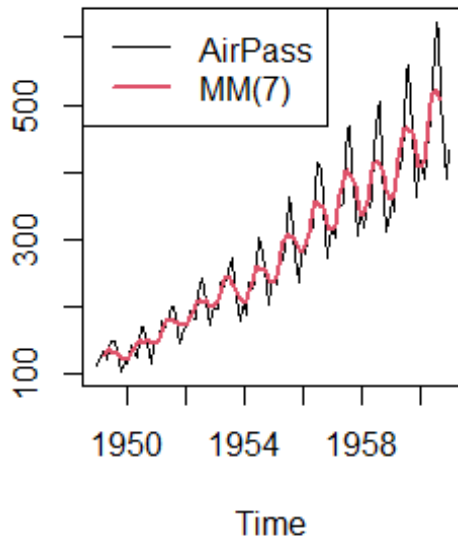
```
> MM3 <- filter(y,filter=rep(1/3,3))
```

```
> MM4 <- filter(y,filter=rep(1/4,4))
```

```
# Les moyennes mobiles centrés d'ordre  $k = 2, 3, 4$  :
```

$$MMC(k)_t = \frac{MM(k)_{t-1} + MM(k)_t}{2} \text{ pour } t = m + 1, \dots, n - m$$

Lissage par moyennes mobiles : la (\tilde{f}_t) est obtenue par lissage par la méthode des moyennes mobiles :



$$\tilde{f}_t = MMC(p)_t,$$

avec $p = 4$. (la période est pair).

calculer MMC(4) : n=16

MMC(4) : $k = 2 * m, m = 2, j = m + 1 = 3, \dots, n - m = 14$.

> MM4<-c(20,21.5,22.25,23.25,24.75,27,26.25,26.5,32.5,33.25,33.5,33,36)

> MMC4 <- c(1 :12) > MMC4 <- (MM4[2 :13]+MM4[3 :14])/2

$$\tilde{f}_t = MMC(4)_t$$

MMC(4)=(20.75,21.875,22.75,24,25.875,26.625,26.375,29.5,32.875, 33.375,33.25,34.5)

Construction de la série des différences (D_t) :

$$D_t = y_t - \tilde{f}_t \quad \text{pour } t = m + 1, \dots, n - m$$

$n = 16, m = 2$.

> D <- c(1 :12)

> for (t in 1 :12) { D[t] = y[t]-MMC4[t] }

D=(9.250,-6.875,-17.750,6.000,10.125,-8.625,-17.375,6.500,12.125,-18.375,-23.250, 25.500)

Calculs des facteurs saisonnières :

$n = 16,$

$p = 2m,$

$K_0 * p = n, (p = 4, K_0 = 4)$

$m = 2.$

$$\tilde{s}_1 = \frac{1}{3} (D_5 + D_9 + D_{13}) = -19.45833$$

$$\tilde{s}_2 = \frac{1}{3} (D_6 + D_{10} + D_{14}) = 12.66667$$

$$\tilde{s}_3 = \frac{1}{3} (D_3 + D_7 + D_{11}) = 10.5$$

$$\tilde{s}_4 = \frac{1}{3} (D_4 + D_8 + D_{12}) = -11.29167$$

Centrage des coefficients saisonniers :

$$\hat{s}_t = \tilde{s}_t - \bar{s}, \quad \text{pour } t = 1, \dots, p = 4$$

```

> s <- c(1 :4)
> s <- c(-19.45833,12.66667,10.5,-11.29167)
> s.center <- s-mean(s)
> s.hat=c(-17.562497,14.562503,12.395833,-9.395837)

```

$$\bar{s} = \frac{1}{4} \sum_{t=1}^4 \tilde{s}_t = -1.895833.$$

et on trouve :

$$\begin{aligned} \hat{s}_1 &= -17.562497 \\ \hat{s}_2 &= 14.562503 \\ \hat{s}_3 &= 12.395833 \\ \hat{s}_4 &= -9.395837 \end{aligned}$$

Construction de la série corrigée des variations saisonnières (CSV_t)_{t=1,...,n} (**La désaisonnalisation**)

$$CSV_t = y_t - \hat{s}_t \text{ pour } t = 1, \dots, n.$$

```

> for (t in 1 :4) for (j in 1 :3) { S.hat[t+4*j] = s.hat[t] } > CSV <- c(1 :16) > for (t in
1 :16) { CSV[t] = y[t]-s.hat[t] }
y=(30,15,5,30,36,18,9,36,45,15,10,60,48,16,8 72)
CSV=(47.56,0.43,-7.39,39.39,53.56,3.43,-3.39,45.39,62.56, 0.43,-2.39,69.39,65.56,1.43,-4.39,81.39)

```

Construction de la série lissée des prédictions :

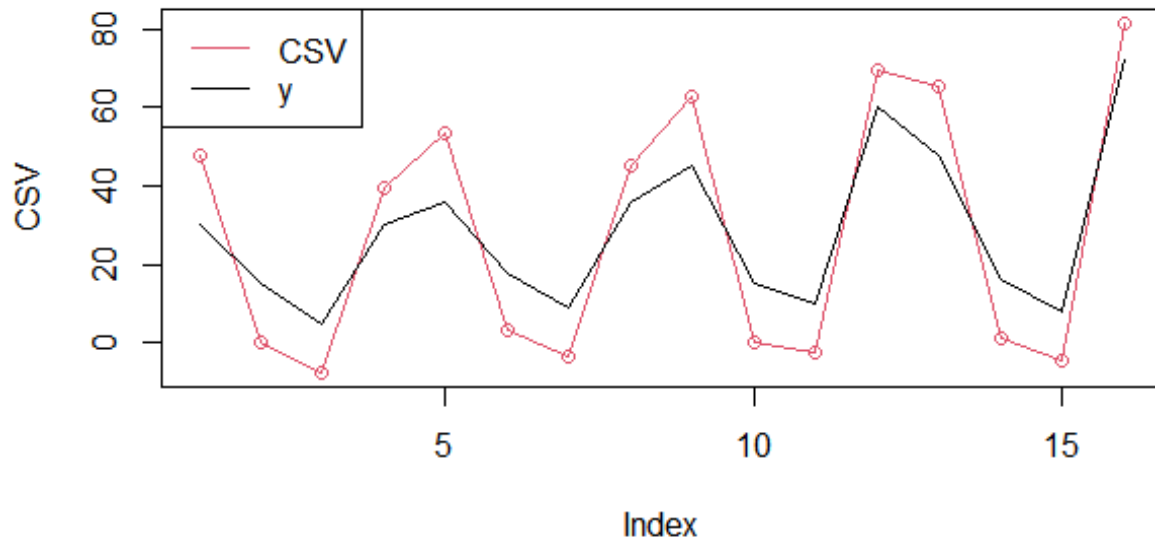
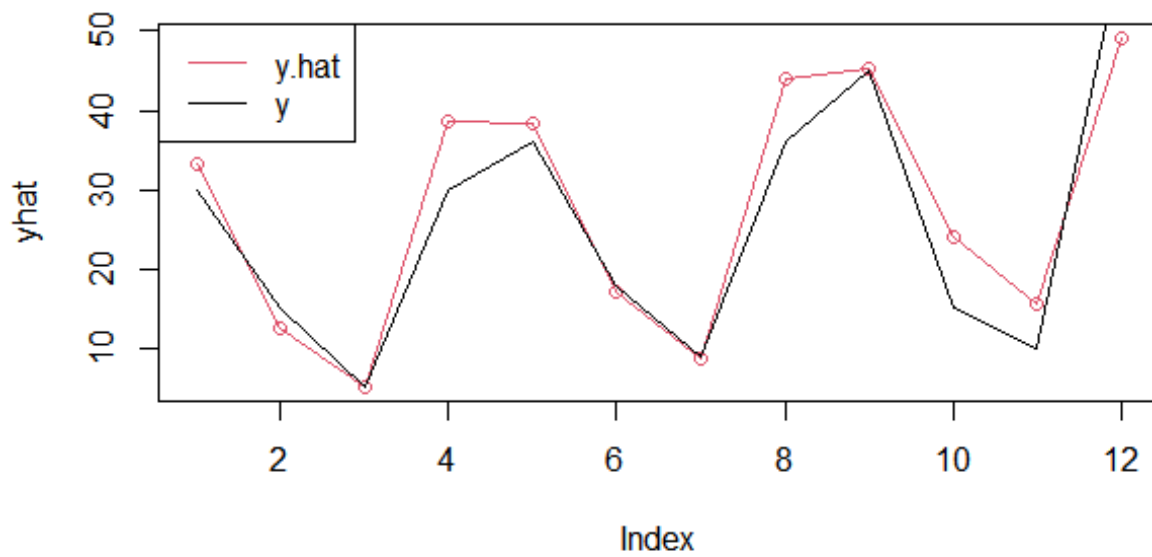
$$\hat{y}_t = \hat{f}_t + \hat{s}_t \text{ pour } t = m + 1 = 3, \dots, n - m = 14.$$

```

>f.hat <- c(20.75,21.875,22.75,24,25.875,26.625,26.375,29.5,32.875,33.375,33.25,34.5)
> y.hat <- c(1 :12) > yhat <- S[3 :14]+f.fit[1 :12] > y.hat > yhat=c(33.14, 12.47, 5.18,
38.56, 38.27, 17.22, 8.81, 44.06, 45.27, 23.97, 15.68, 49.06)
# l'erreur de prévision associée
> < -c(1 :12)
> e< - y[1 :12]-y.hat

```

Exemple 3 : (*Lissage exponentiel simple*)

FIGURE 2.1 – La série corrigée des variations saisonnières $(CSV_t)_{t=1,\dots,n}$ FIGURE 2.2 – La série lissée des prédictions (\hat{y}_t) .

$$\begin{aligned}
 \hat{x}_{t+1} &= \alpha x_t + (1 - \alpha) \hat{x}_t \\
 &= \underbrace{\hat{x}_t}_{\text{prédiction}} + \alpha \underbrace{(x_t - \hat{x}_t)}_{\text{erreur}} \\
 &= \text{prédiction} + \text{erreur} \\
 &\quad \text{au temps } t
 \end{aligned}$$

$0 < \alpha < 1$ et $\hat{x}_1 = x_1$.

Algorithme 1 :

```

>x=c(0.42,0.11,0.65,2.92,3.01,3.08,2.98,3.29,2.63,1.29,1.49,0.98,1.16,0.93,1.65,2.98,3.40,
3.46,2.96,3.45,3.66)
> y=ts(x)
>alpha=0.9
> # alpha=0.75
> n=21
> y.hat <- c(1 : n+1)
> y.hat[1] <- y[1]
> for (i in 2 :n+1) { y.hat[i] <- y.hat[i-1]+alpha*(y[i]-y.hat[i-1]) }
> print(y.hat)
> par(mfrow = c(1,1))
> plot(st(y.hat),col=1,lwd=2)
> lines(y.hat,col=1,lwd=1)
> lines(y,col=2,lwd=2)
> legend("topleft",legend=c("y.hat", "y",col=c(1,2),lwd=c(2,2)))
> # mean square error : MSE :
> e <- c(1 :n)
> for (i in 2 :n) { e[i] <- y[i]-y.hat[i] }
> var(e)
> #Root Mean Square Error "RMSE" :
> sd(e)

```

• On peut utiliser la commande suivante :

```

> xlisser <- HoltWinters(x,alpha=alpha ,beta=FALSE,gamma=FALSE) :

```

Algorithme2 :

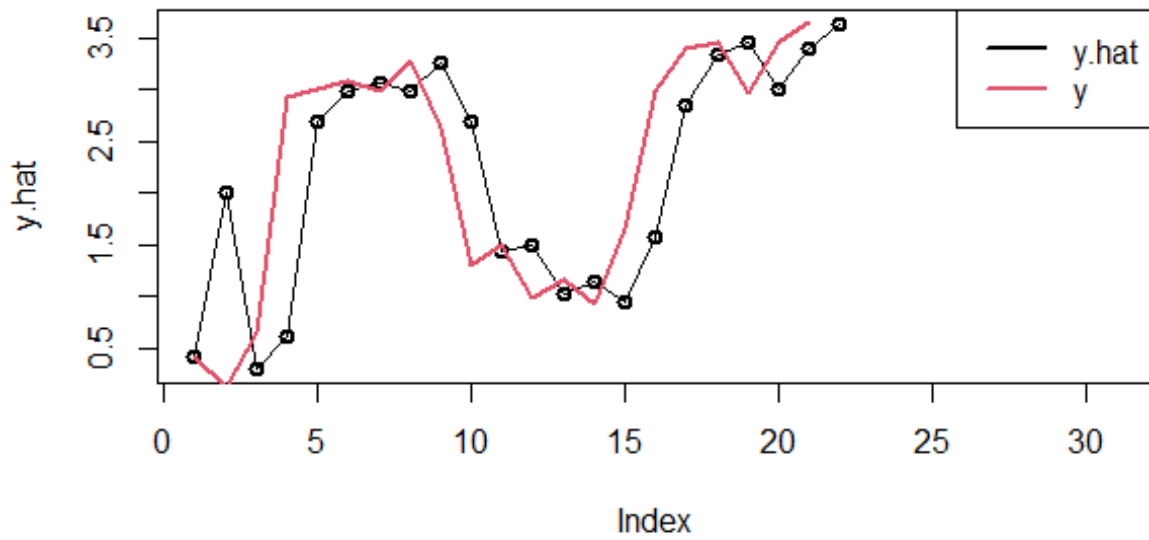


FIGURE 2.3 – Lissage exponentiel simple $\alpha = 0.9$. et $MSE = 0.79007$, $RMSE = 0.88886$

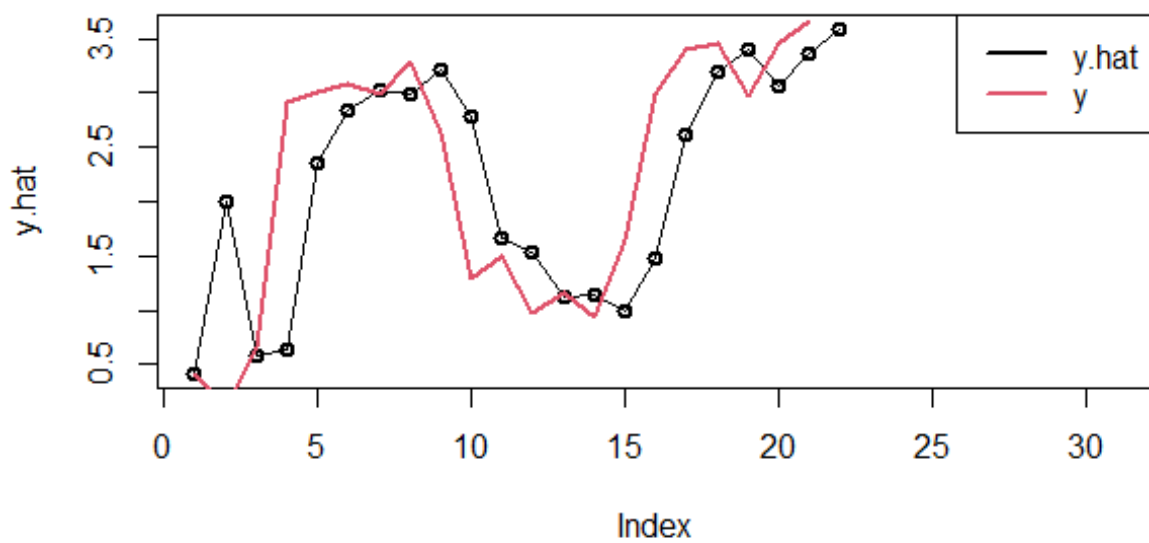


FIGURE 2.4 – Lissage exponentiel simple : $\alpha = 0.75$, et $MSE = 0.83366$, $RMSE = 0.91305$

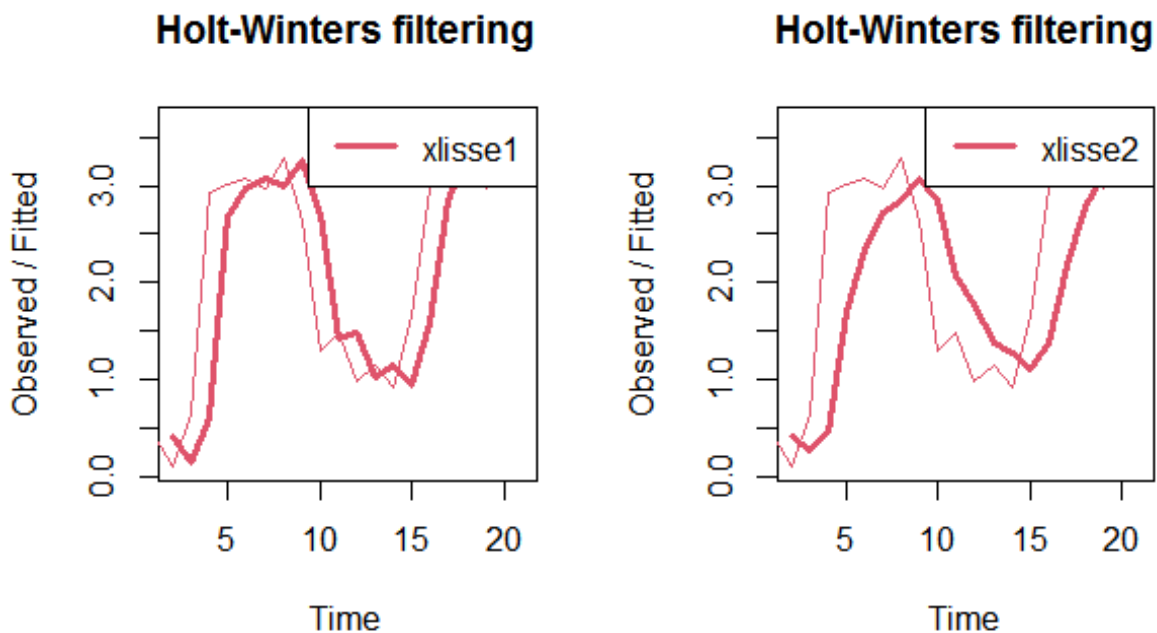


FIGURE 2.5 – Lissage exponentiel simple : $\alpha = 0.9$ (left), $\alpha = 0.5$ (right)

```
xlisce1 <- HoltWinters(y,alpha=0.9 ,beta=FALSE,gamma=FALSE)
xlisce2 <- HoltWinters(y,alpha=0.5,beta=FALSE,gamma=FALSE)
par(mfrow = c(1,2))
plot(xlisce1,col=1,lwd=2)
legend("topright",legend=c("xlisce1"),col=1,lwd=2)
plot(xlisce2,col=1,lwd=#lines(y,col=2,lwd=1)
legend("topright",legend=c("xlisce2"),col=1,lwd=2)
```

Exemple 4 : (*Lissage exponentiel double*)

- On utilise la commande :

```
xlisce <- HoltWinters(x,alpha= $\alpha$  , beta= $\beta$  , gamma=FALSE),
avec
```

$$\alpha = 1 - (\alpha')^2,$$

$$\beta = \frac{1 - \alpha'}{1 + \alpha'}.$$

Algorithme :

```
alpha0=0.5
```

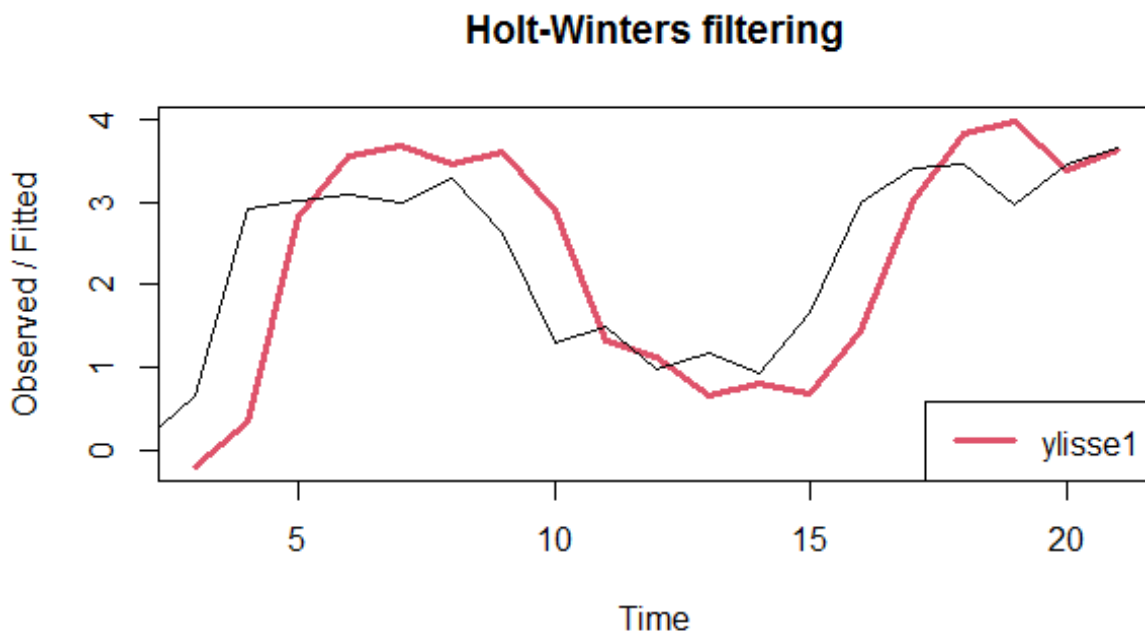


FIGURE 2.6 – Lissage exponentiel double, $\alpha = 0.75$, $\beta = 0.33$ (*Exemple 4*)

```
alpha <- 1-alpha0*alpha0
print(alpha1)
beta1 <- (1-alpha0)/(1+alpha0)
print(beta1)
ylisse1 <- HoltWinters(y,alpha=0.75 ,beta=0.33,gamma=FALSE)
plot(ylisse1,col=2,lwd=3)
legend("topright",legend=c("ylisse1"),col=2,lwd=3)
```

Exemple 5 : (*Lissage exponentiel-HotWinters*)

#un lissage de Holt-Winters sans composante saisonnière :

```
ylisse2 <- HoltWinters(y,alpha= $\alpha$  ,beta= $\beta$  ,gamma=FALSE)
```

un lissage Holt-Winters additif :

```
ylisse3 <- HoltWinters(y,alpha= $\alpha$ ,beta= $\beta$ ,gamma= $\gamma$ , seasonal="add")
```

un lissage Holt-Winters multiplicatif :

```
ylisse4 <- HoltWinters(y,alpha= $\alpha$ ,beta= $\beta$ ,gamma= $\gamma$  ,seasonal="mul").
```

Les prévisions à l'horizon h sont réalisées à l'aide de la fonction predict :


```
p2 <-predict(ylisse2, n.ahead=h)
p3 <-predict(ylisse3, n.ahead=h)
p4 <-predict(ylisse4, n.ahead=h)
Algorithme :
library(ggplot2)
ts.AirPass <-ts(AirPassengers, start=1, frequency=12)
# Lissage exponentiel simple
a <- HoltWinters(ts.AirPass,alpha=0.5 ,beta=FALSE,gamma=FALSE)
#un lissage de Holt-Winters sans composante saisonnière :
b <- HoltWinters(ts.AirPass,alpha=0.5 ,beta=0.9 ,gamma=FALSE)
# un lissage Holt-Winters additif :
c <- HoltWinters(ts.AirPass,alpha=0.5,beta=0.9,gamma=0.25, seasonal="add")
# un lissage Holt-Winters multiplicatif :
d <- HoltWinters(ts.AirPass,alpha=0.5,beta=0.9,gamma=0.25 ,seasonal="mul")
par(mfrow=c(2,2))
plot(a,col=1,lwd=2)
legend("topleft",legend=c("a"),col=1,lwd=2)
plot(b,col=1,lwd=2)
legend("topleft",legend=c("b"),col=1,lwd=2)
plot(c,col=1,lwd=2)
legend("topleft",legend=c("c"),col=1,lwd=2)
plot(d,col=1,lwd=2)
legend("topleft",legend=c("d"),col=1,lwd=2)
# Les prévisions à l'horizon h sont réalisées à l'aide de la fonction predict :
p1 <-predict(a, n.ahead=10)
p2 <-predict(b, n.ahead=10)
p3 <-predict(c, n.ahead=10)
p4 <-predict(d, n.ahead=10)
```

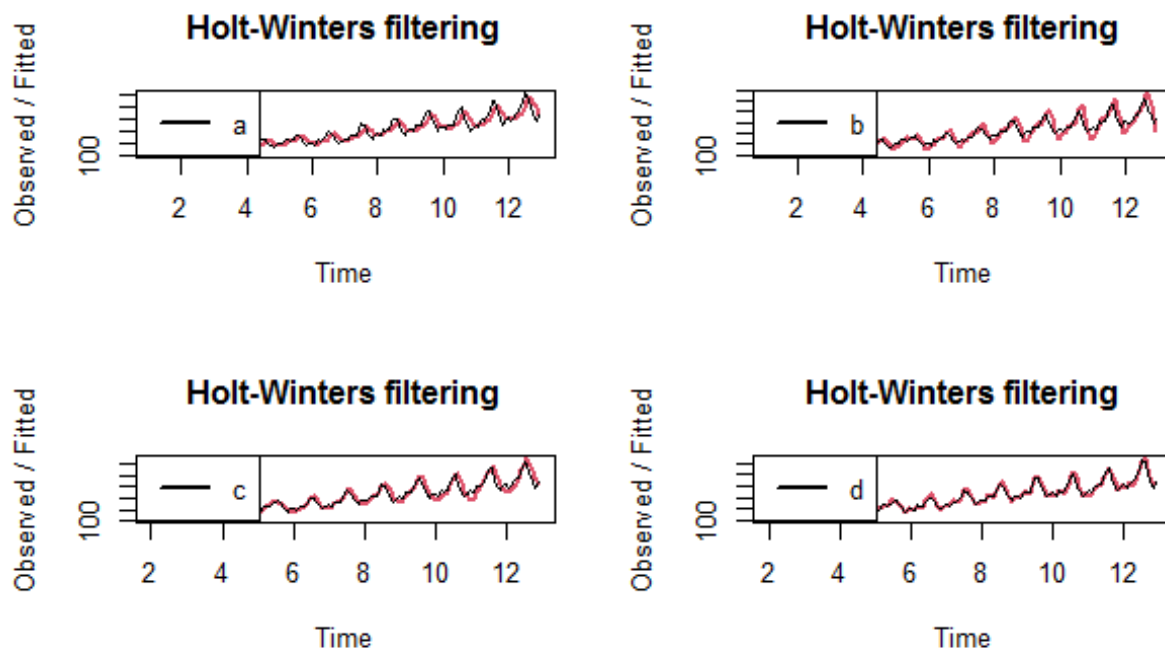


FIGURE 2.7 – Lissage exponentiel simple, double, Holt-Winters appliqué sur "AirPassengers Data"

Modélisation stochastique des séries temporelles

3.1 Processus stochastiques stationnaires

- On dit que le processus $(X_t)_{t \in \mathbb{Z}}$ est **stationnaire au sens strict** si pour toute famille finie d'instants t_1, t_2, \dots, t_k et tout entier s , les lois jointes de $(X_{t_1}, X_{t_2}, \dots, X_{t_k})$ et de $(X_{t_1+s}, X_{t_2+s}, \dots, X_{t_k+s})$ sont les mêmes.

- On dit que le processus $(X_t)_{t \in \mathbb{Z}}$ est **stationnaire à l'ordre 2** si la moyenne $m(t) = E\{X(t)\}$ et la covariance $C(t, s) = Cov(X(t), X(s))$ sont invariantes par translation dans le temps, c'est-à-dire si la moyenne est constante :

$$m(t) = m, \quad \forall t \in \mathbb{Z},$$

et si la covariance dépend seulement de l'écart de temps $h = t - s$, c'est-à-dire il existe une fonction d'une variable $V(h)$, **paire**, telle que :

$$\begin{aligned} Cov(X(t), X(s)) &= C(t, s) \\ &= V(t - s) \\ &= V(h), \quad \forall h \in \mathbb{Z}, \end{aligned}$$

en particulier $V(0) = Cov(X(t), X(t)) = Var\{X(t)\}$.

La fonction $V(h)$ est dite la fonction **d'auto-covariance** de X .

- On définit la fonction **d'auto-covariance** de X par

$$\rho(h) = \frac{V(h)}{V(0)}.$$

- Auto-corrélation partielle :

• Pour un processus stationnaire, la moyenne empirique, la variance empirique, l'auto-covariance empirique, l'auto-covariance empirique sont estimées à partir de n observations convergent vers la moyenne théorique, la variance théorique, l'auto-covariance théorique, l'auto-corrélation théorique, respectivement :

- **La moyenne empirique :**

$$\bar{X} = \frac{1}{n} \sum_{t=1}^n X_t.$$

- **La variance empirique :**

$$\begin{aligned} \hat{V}(0) &= \frac{1}{n} \sum_{t=1}^n (X_t - \bar{X})^2 \\ &= \frac{1}{n} \sum_{t=1}^n X_t^2 - (\bar{X})^2. \end{aligned}$$

- **L'auto-covariance empirique :**

$$\hat{V}(h) = \frac{1}{n-h} \sum_{t=1}^{n-h} (X_t - \bar{X})(X_{t+h} - \bar{X}).$$

- **L'auto-corrélation empirique :**

$$\hat{\rho}(h) = \frac{\hat{V}(h)}{\hat{V}(0)}.$$

- Exemples :

- Le processus de moyenne mobile :

Soit $(\epsilon_t)_{t \in \mathbb{Z}}$ un bruit blanc centré de variance σ^2 , c'est-à-dire : $(\epsilon_t)_{t \in \mathbb{Z}}$ est une suite de variables aléatoires non corrélées et d'espérance constante nulle et de variance constante, $E\{\epsilon_t\} = 0$, $Var\{\epsilon_t\} = \sigma^2$ et $Cov(\epsilon_t, \epsilon_{t+h}) = 0$ pour $h \neq 0$. On appelle **processus de moyenne mobile** un processus $(X_t)_{t \in \mathbb{Z}}$ de la forme

$$\begin{aligned} X_t &= \epsilon_t + a_1 \epsilon_{t-1} + \dots + a_q \epsilon_{t-q} \\ &= \sum_{j=0}^q a_j \epsilon_{t-j}, \quad (a_0 = 1). \end{aligned}$$

La moyenne de X_t est donnée par

$$\begin{aligned} E\{X_t\} &= \sum_{j=0}^q a_j E\{\epsilon_{t-j}\} \\ &= 0. \end{aligned}$$

La variance de X_t est donnée par

$$\begin{aligned}
 \gamma(0) &= \text{Var} \{X_t\} = \text{Cov}(X_t, X_t) \\
 &= \text{Cov} \left(\sum_{i=0}^q a_i \epsilon_{t-i}, \sum_{j=0}^q a_j \epsilon_{t-j} \right) \\
 &= \sum_{i=0}^q \sum_{j=0}^q a_i a_j \text{Cov}(\epsilon_{t-i}, \epsilon_{t-j}) \\
 &= \sigma^2 \sum_{i=0}^q a_i^2,
 \end{aligned}$$

car

$$\text{Cov}(\epsilon_{t-i}, \epsilon_{t-j}) = \begin{cases} \sigma^2 & \text{si } i = j \\ 0 & \text{si } i \neq j \end{cases}$$

La covariance de X_t est donnée par

$$\begin{aligned}
 \gamma(h) &= \text{Cov}(X_t, X_{t+h}) = \text{Cov} \left(\sum_{i=0}^q a_i \epsilon_{t-i}, \sum_{j=0}^q a_j \epsilon_{t+h-j} \right) \\
 &= \sum_{i=0}^q \sum_{j=0}^q a_i a_j \text{Cov}(\epsilon_{t-i}, \epsilon_{t+h-j}) \\
 &= \sum_{i=0}^q \sum_{j=-h}^q a_i a_j \text{Cov}(\epsilon_{t-i}, \epsilon_{t-(j-h)}) \quad (k = j - h) \\
 &= \sum_{i=0}^q \sum_{k=-h}^{q-h} a_i a_{k+h} \text{Cov}(\epsilon_{t-i}, \epsilon_{t-k}) \\
 &= \sigma^2 \sum_{i=0}^q \sum_{k=-h}^{q-h} a_i a_{k+h} \\
 &= \sigma^2 \sum_{k=0}^{q-h} a_k a_{k+h} \\
 &= \begin{cases} \sigma^2 \sum_{k=0}^{q-h} a_k a_{k+h} & \text{si } h = 0, 1, \dots, q. \\ 0 & \text{si } h = q+1, q+2, \dots \end{cases}
 \end{aligned}$$

La fonction de covariance $\gamma(h)$ est paire, i.e., ($\gamma(h) = \gamma(-h)$), alors elle est donnée par

$$\gamma(h) = \begin{cases} \sigma^2 \sum_{k=0}^{q-|h|} a_k a_{k+|h|} & \text{si } h = -q, \dots, -1, 0, 1, \dots, q. \\ 0 & \text{si } h = \pm(q+1), \pm(q+2), \dots \end{cases}$$

• Processus autorégressif d'ordre 1 :

Soit $(\epsilon_t)_{t \in \mathbb{Z}}$ un bruit blanc centré de variance σ^2 et $a \in]-1, 1[$ un paramètre fixé.

Un processus **autorégressif d'ordre 1** est un processus stationnaire centré qui vérifie les

équations de récurrence suivantes :

$$X_t = aX_{t-1} + \epsilon_t, \forall t \in \mathbb{Z}.$$

On peut déduire

$$\begin{aligned} X_t &= a(aX_{t-2} + \epsilon_{t-1}) + \epsilon_t \\ &= a^2X_{t-2} + a\epsilon_{t-1} + \epsilon_t \\ &= a^2(aX_{t-3} + \epsilon_{t-2}) + a\epsilon_{t-1} + \epsilon_t \\ &= a^3X_{t-3} + a^2\epsilon_{t-2} + a\epsilon_{t-1} + \epsilon_t \\ &\vdots \\ &= \epsilon_t + a\epsilon_{t-1} + a^2\epsilon_{t-2} + a^3\epsilon_{t-3} + \dots \\ &= \sum_{j=0}^{\infty} a^j \epsilon_{t-j}. \end{aligned}$$

On a

$$E\{X_t\} = \sum_{j=0}^{\infty} a^j E\{\epsilon_{t-j}\} = 0.$$

car $E\{\epsilon_{t-j}\} = 0$.

$$\begin{aligned} Var\{X_t\} &= V(0) = Cov(X_t, X_t) \\ &= Cov\left(\sum_{i=0}^{\infty} a^i \epsilon_{t-i}, \sum_{j=0}^{\infty} a^j \epsilon_{t-j}\right) \\ &= \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} a^i a^j Cov(\epsilon_{t-i}, \epsilon_{t-j}) \\ &= \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} a^{i+j} \begin{cases} \sigma^2 & \text{si } i = j \\ 0 & \text{si } i \neq j \end{cases} \\ &= \sigma^2 \sum_{i=0}^{\infty} a^{2i} \\ &= \sigma^2 \frac{1}{1-a^2} \quad \text{car } |a| < 1. \end{aligned}$$

On a aussi

$$\begin{aligned}
V(h) &= \text{Cov}(X_t, X_{t+h}) \\
&= \text{Cov}\left(\sum_{i=0}^{\infty} a^i \epsilon_{t-i}, \sum_{j=0}^{\infty} a^j \epsilon_{t+h-j}\right) \\
&= \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} a^i a^j \text{Cov}(\epsilon_{t-i}, \epsilon_{t+h-j}) \\
&= \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} a^i a^j \text{Cov}(\epsilon_{t-i}, \epsilon_{t-(j-h)}) \quad \text{on pose } k = j - h \\
&= \sum_{i=0}^{\infty} \sum_{k=-h}^{\infty} a^i a^{k+h} \text{Cov}(\epsilon_{t-i}, \epsilon_{t-k}) \\
&= \sum_{i=0}^{\infty} \sum_{k=-h}^{\infty} a^{i+k+h} \begin{cases} \sigma^2 & \text{si } i = k \\ 0 & \text{si } i \neq k \end{cases} \\
&= \sigma^2 \sum_{i=0}^{\infty} a^{2i+h} \\
&= \sigma^2 \frac{a^{|h|}}{1-a^2}, \quad (\text{la fonction } V(h) \text{ est } \mathbf{paire})
\end{aligned}$$

alors, l'auto-corrélation est donnée par

$$\rho(h) = \frac{V(h)}{V(0)} = a^{|h|} \rightarrow 0 \text{ quand } h \rightarrow \infty.$$

3.2 Les processus stationnaires ARMA

- Le processus $(X_t)_{t \in \mathbb{Z}}$ est un $ARMA(p, q)$ de moyenne nulle si :

$$\begin{cases} (X_t) \text{ est stationnaire} \\ X_t - \sum_{i=1}^p \alpha_i X_{t-i} = \sum_{i=0}^q \beta_i \epsilon_{t-i} \end{cases},$$

avec $\beta_0 = 1$ et $(\epsilon_t)_{t \in \mathbb{Z}}$ un bruit blanc **faible** (si ϵ_t i.i.d, $(\epsilon_t)_{t \in \mathbb{Z}}$ un bruit blanc **fort**).

On peut écrire

$$X_t - \alpha_1 X_{t-1} - \dots - \alpha_p X_{t-p} = \epsilon_t + \beta_1 \epsilon_{t-1} + \dots + \beta_q \epsilon_{t-q}.$$

Une représentation équivalente du processus $(X_t)_{t \in \mathbb{Z}}$ est donnée par :

$$\Phi(B) X_t = \Theta(B) \epsilon_t,$$

où Φ et Θ sont des polynômes de degrés respectifs p et q :

$$\Phi(z) = 1 - \alpha_1 z - \dots - \alpha_p z^p,$$

et

$$\Theta(z) = 1 + \beta_1 z + \dots + \beta_q z^q,$$

et B est l'opérateur "**retard**" défini par :

$$B^k X_t = X_{t-k}.$$

3.2.1 Cas particuliers

- Si $\Phi(z) = 1$, (X_t) est un processus *MA* d'ordre q :

$$X_t = \epsilon_t + \beta_1 \epsilon_{t-1} + \dots + \beta_q \epsilon_{t-q}.$$

- Si $\Theta(z) = 1$, (X_t) est un processus *AR* d'ordre p :

$$X_t - \alpha_1 X_{t-1} - \dots - \alpha_p X_{t-p} = \epsilon_t.$$

- (X_t) est un processus *MA* d'ordre infini, *MA*(∞), si

$$X_t = \sum_{i=0}^{\infty} \theta_i \epsilon_{t-i},$$

avec

$$\sum_{i=0}^{\infty} |\theta_i| < \infty.$$

- Un processus *ARMA* est dit **causal** s'il peut être écrit sous la forme d'un *MA*(∞), soit

$$X_t = \sum_{i=0}^{\infty} \theta_i \epsilon_{t-i}.$$

- X_t ne dépend que de ϵ_t et des valeurs précédentes de ϵ .
- Un processus *ARMA* est dit **inversible** si l'on peut écrire :

$$\epsilon_t = \sum_{i=0}^{\infty} \pi_i X_{t-i},$$

avec

$$\sum_{i=0}^{\infty} |\pi_i| < \infty.$$

3.2.2 Espérance, Covariance, corrélation

- **Espérance** : l'espérance mathématique d'un processus *ARMA causal* est :

$$\begin{aligned} E\{X_t\} &= E\left\{\sum_{i=0}^{\infty} \theta_i \epsilon_{t-i}\right\} \\ &= \sum_{i=0}^{\infty} \theta_i E\{\epsilon_{t-i}\} \\ &= 0 \end{aligned}$$

- **Covariance** : Il est important, pour étudier les processus *ARMA*, d'établir les liens entre leurs paramètres de définition et leur fonction de covariance, la fonction de covariance d'un processus *ARMA causal* est :

$$\begin{aligned} V(h) &= Cov(X_t, X_{t+h}) \\ &= Cov\left(\sum_{i=0}^{\infty} \theta_i \epsilon_{t-i}, \sum_{j=0}^{\infty} \theta_j \epsilon_{t+h-j}\right) \\ &= \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \theta_i \theta_j Cov(\epsilon_{t-i}, \epsilon_{t-(j-h)}) \quad \text{on pose } k = j - h \\ &= \sum_{i=0}^{\infty} \sum_{k=-h}^{\infty} \theta_i \theta_{k+h} Cov(\epsilon_{t-i}, \epsilon_{t-k}) \\ &= \sum_{i=0}^{\infty} \sum_{k=-h}^{\infty} \theta_i \theta_{k+h} \begin{cases} \sigma^2 & \text{si } i = k \\ 0 & \text{si } i \neq k \end{cases} \\ &= \sigma^2 \sum_{i=0}^{\infty} \theta_i \theta_{i+h} \quad \text{pour } h = 0, 1, \dots \end{aligned}$$

La fonction $V(h)$ est paire, i.e., $V(h) = V(-h)$, alors $\forall h \in \mathbb{Z}$,

$$V(h) = \sigma^2 \sum_{i=0}^{\infty} \theta_i \theta_{i+|h|}.$$

$$\sigma^2 = Var\{\epsilon_t\} = Cov(\epsilon_t, \epsilon_t).$$

- **corrélation** : la corrélation d'un processus *ARMA causal* est :

$$\begin{aligned} \rho(h) &= \frac{V(h)}{V(0)} \\ &= \frac{\sigma^2 \sum_{i=0}^{\infty} \theta_i \theta_{i+|h|}}{\sigma^2 \sum_{i=0}^{\infty} \theta_i^2} \\ &= \frac{\sum_{i=0}^{\infty} \theta_i \theta_{i+|h|}}{\sum_{i=0}^{\infty} \theta_i^2}, \end{aligned}$$

avec $V(0) = Var\{X_t\} = Cov(X_t, X_t) = \sigma^2 \sum_{i=0}^{\infty} \theta_i^2$.

3.2.3 Modèle d'état pour un processus ARMA

Les processus ARMA peuvent être représentés par des modèles d'état, on donne ici la représentation par modèle d'état d'un processus AR et d'un processus ARMA.

• Modèle d'état pour un AR(p) :

X est un processus AR d'ordre p :

$$X_{t+1} = \epsilon_{t+1} + \alpha_1 X_t + \alpha_2 X_{t-1} + \dots + \alpha_{p-1} X_{t-p+2} + \alpha_p X_{t-p+1}, \quad t \in \mathbb{Z}.$$

Soit le vecteur d'état \mathbf{Y} de taille p défini par :

$$\mathbf{Y}_{t+1} = (X_{t-p+1}, X_{t-p+2}, \dots, X_{t-1}, X_t)', \quad t \in \mathbb{Z}.$$

On peut écrire le modèle d'état pour un AR(p) comme suit :

$$\mathbf{Y}_{t+1} = F\mathbf{Y}_t + V_t,$$

avec

$$F = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ \alpha_p & \alpha_{p-1} & \alpha_{p-1} & \dots & \alpha_1 \end{pmatrix}_{(p,p)}, \quad G = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}_{(p,1)} \quad \text{et} \quad V_t = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ \epsilon_{t+1} \end{pmatrix}_{(p,1)}$$

• Modèle d'état pour un ARMA(p, q) :

X est un processus ARMA(p, q) :

$$X_t = \alpha_1 X_{t-1} + \alpha_2 X_{t-2} + \dots + \alpha_{p-1} X_{t-p+1} + \alpha_p X_{t-p} + \epsilon_t + \beta_1 \epsilon_{t-1} + \dots + \beta_q \epsilon_{t-q}.$$

Soit le vecteur d'état \mathbf{Y} de taille p défini par :

$$\mathbf{Y}_{t+1} = (X_{t-p+1}, X_{t-p+2}, \dots, X_{t-1}, X_t)', \quad t \in \mathbb{Z}.$$

Le modèle d'état pour un ARMA(p, q) est défini par :

$$\mathbf{Y}_{t+1} = F\mathbf{Y}_t + B V_t,$$

avec

$$F = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ \alpha_p & \alpha_{p-1} & \alpha_{p-1} & \dots & \alpha_1 \end{pmatrix}_{(p,p)}, \quad B = \begin{pmatrix} 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 0 & 0 & 0 & \dots & 0 \\ \beta_q & \beta_{q-1} & \dots & \beta_1 & 1 \end{pmatrix}_{(p,q+1)} \quad \text{et} \quad V_t = \begin{pmatrix} \epsilon_{t-q} \\ \epsilon_{t-q+1} \\ \vdots \\ \epsilon_{t-1} \\ \epsilon_t \end{pmatrix}_{(q+1,1)}$$

3.2.4 Identification des processus ARMA

• L'identification d'un processus *ARMA* de moyenne nulle est l'opération qui consiste à estimer au mieux les paramètres qui définissent ses propriétés du deuxième ordre, à savoir :

- les ordres p et q ,
- les p coefficients α et les q coefficients β ,
- l'écart-type de $\epsilon : \sigma$

• Critères de présomption de processus ARMA :

- Il y a présomption de processus *ARMA* si les conditions suivantes sont satisfaites :
 - le processus est stationnaire à l'analyse visuelle : (pas de tendance, pas de saisonnalité, variance constante).
 - la fonction de corrélation empirique est : (à décroissance pas trop lente, sans pics périodiques).

• Principes d'identification :

L'identification se fait selon un principe global, illustré par la figure suivante :

hypothèse $p, q \rightarrow$ estimation des paramètres $\alpha, \beta, \sigma^2 \rightarrow$ prédiction à un pas \rightarrow calcul du résidu \rightarrow test sur le résidu

Exemple 13. *On suppose que*

$$X_t = a_1 X_{t-1} + a_2 X_{t-2} + \epsilon_t,$$

ϵ est un bruit blanc centré, de variance σ^2 et que (X_t) est stationnaire. i.e.,

$$E \{X_t\} = m, \quad Var \{X_t\} = V(0), \quad Cov(X_t, X_{t+h}) = V(h),$$

On calcule la moyenne

$$\begin{aligned} E \{X_t\} &= a_1 E \{X_{t-1}\} + a_2 E \{X_{t-2}\} + E \{\epsilon_t\} \\ m &= a_1 m + a_2 m + 0, \end{aligned}$$

donc

$$m(1 - a_1 - a_2) = 0,$$

alors

$$m = 0.$$

Pour la variance, on a

$$\begin{aligned} V(0) &= \text{Cov}(X_t, X_t) \\ &= \text{Cov}(a_1 X_{t-1} + a_2 X_{t-2} + \epsilon_t, X_t) \\ &= a_1 \text{Cov}(X_t, X_{t-1}) + a_2 \text{Cov}(X_t, X_{t-2}) + \text{Cov}(X_t, \epsilon_t) \\ &= a_1 V(1) + a_2 V(2) + 0, \end{aligned}$$

et obtient

$$a_1 V(1) + a_2 V(2) = V(0).$$

On a aussi

$$\begin{aligned} V(1) &= \text{Cov}(X_t, X_{t-1}) \\ &= \text{Cov}(a_1 X_{t-1} + a_2 X_{t-2} + \epsilon_t, X_{t-1}) \\ &= a_1 \text{Cov}(X_{t-1}, X_{t-1}) + a_2 \text{Cov}(X_{t-2}, X_{t-1}) + \text{Cov}(\epsilon_t, X_{t-1}) \\ &= a_1 V(0) + a_2 V(1) + 0, \end{aligned}$$

donc

$$(1 - a_2) V(1) = a_1 V(0).$$

et on a

$$\begin{aligned} V(2) &= \text{Cov}(X_t, X_{t-2}) \\ &= \text{Cov}(a_1 X_{t-1} + a_2 X_{t-2} + \epsilon_t, X_{t-2}) \\ &= a_1 \text{Cov}(X_{t-1}, X_{t-2}) + a_2 \text{Cov}(X_{t-2}, X_{t-2}) + \text{Cov}(\epsilon_t, X_{t-2}) \\ &= a_1 V(1) + a_2 V(0) + 0, \end{aligned}$$

donc

$$V(2) = a_1 V(1) + a_2 V(0).$$

(a_1, a_2) est la solution du système

$$\begin{cases} a_1 V(0) + a_2 V(1) = V(1), \\ a_1 V(1) + a_2 V(0) = V(2), \end{cases}$$

en remplaçant $V(0), V(1), V(2)$ par leurs estimateurs empiriques dans les formules ci-dessus on peut estimer a_1, a_2 .

3.3 Les processus non stationnaires ARIMA et SARIMA

Definition 14. Le processus $(X_t)_{t \geq 0}$ est un processus ARIMA (p, d, q) si $Y_t = \Delta_1^{(d)} X_t$ est un processus ARMA (p, q) .

Definition 15. Le processus $(X_t)_{t \geq 0}$ est un processus SARIMA (p, d, q, T) si $Y_t = \Delta_T \circ \Delta_1^{(d)} X_t$ est un processus ARMA (p, q)

Remarque :

- Les processus ARIMA (p, d, q) sont donc bien adaptés à l'étude des séries temporelles présentant une tendance polynômiale de degré $d - 1$.

- Les processus SARIMA (p, d, q, T) sont donc bien adaptés à l'étude des séries temporelles qui présentent une saisonnalité de période T et qui ont une tendance polynômiale de degré $d - 1$.

Exemple 16. Soit la série temporelle $(X_t)_{t=1, \dots, n}$ qui a une saisonnalité de période $T = 12$ ($n > 12$), pour supprimer la saisonnalité on définit la série temporelle (Y_t) par

$$Y_t = \Delta_{12} X_t = X_t - X_{t-12}.$$

Pour $t = n + h$, ($h = 1, \dots, 11$), on peut écrire

$$\begin{aligned} X_t &= Y_t + X_{t-12} \\ &= Y_t + (Y_{t-12} + X_{t-24}) \\ &= Y_t + Y_{t-12} + (Y_{t-24} + X_{t-36}) \\ &= \dots \\ &= Y_t + Y_{t-12} + Y_{t-24} + \dots + Y_{r+12} + X_r. \end{aligned}$$

$r = t[12]$ le reste de la division euclidienne de t par 12.

On en déduit les prévisions $\widehat{X}_{n,h}$ pour $h = 1, \dots, 11$, on utilise l'égalité ci-dessus avec $t = n + h$ pour calculer la prévision

$$\widehat{X}_{n,h} = \widehat{Y}_{n,h} + Y_{t-12} + Y_{t-24} + \dots + Y_{r+12} + X_r.$$

3.4 Les processus ARCH et GARCH

3.4.1 Rappels de probabilité

• Si (X, Y) est un couple de variables aléatoires de densité $f_{(X,Y)}$ dans \mathbb{R}^2 , la densité conditionnelles de X sachant $Y = y$ est définie par

$$f_{X|Y}(x) = \frac{f_{(X,Y)}(x, y)}{f_Y(y)}, \text{ si } f_Y(y) > 0,$$

où $f_Y(y)$ est la densité marginale de Y . On peut alors calculer l'espérance conditionnelle de n'importe quelle fonction φ de X :

$$E\{\varphi(X) | Y = y\} = \int_{\mathbb{R}} \varphi(x) f_{X|Y}(x) dx.$$

On note $\mathcal{D}_s = \{X_{s-1}, X_{s-2}, \dots\}$ pour $s \geq 1$. On a alors, pour tout processus (X_t) ,

$$\text{Var}\{X_t | \mathcal{D}_{t-1}\} = E\{X_t^2 | \mathcal{D}_{t-1}\} - (E\{X_t | \mathcal{D}_{t-1}\})^2.$$

3.4.2 Le processus ARCH

Definition 17. *Un processus ARCH(p) est un processus stationnaire $(X_t)_{t \in \mathbb{Z}}$ vérifiant une relation de récurrence du type*

$$\begin{cases} X_t = \epsilon_t \sigma_t \text{ avec } \epsilon_t | X_{t-1}, X_{t-2}, \dots \rightsquigarrow \mathcal{N}(0, \sigma_t), \\ \sigma_t^2 = \alpha_0 + \alpha_1 X_{t-1}^2 + \dots + \alpha_p X_{t-p}^2, \end{cases}$$

avec des coefficients $\alpha_0, \alpha_1, \dots, \alpha_p \geq 0$.

Proposition 18. *Soit $(X_t)_{t \in \mathbb{Z}}$ un processus ARCH(p). On note, pour tout t , $\mathcal{D}_t = \{X_{t-1}, X_{t-2}, \dots\}$.*

On a les propriétés suivantes pour tout t et $h \geq 1$:

- $E\{X_t | \mathcal{D}_{t-1}\} = 0$,
- $E\{X_t\} = 0$,
- $\text{Var}\{X_t | \mathcal{D}_{t-1}\} = \alpha_0 + \alpha_1 X_{t-1}^2 + \dots + \alpha_p X_{t-p}^2$,
- $\text{Var}\{X_t\} = \frac{\alpha_0}{1 - \sum_{i=1}^p \alpha_i}$, si $\sum_{i=1}^p \alpha_i < 1$,
- $\text{Cov}(X_t, X_{t+h} | \mathcal{D}_{t-1}) = 0$, pour $h \geq 1$,
- $\text{Cov}(X_t, X_{t+h}) = 0$, pour $h \geq 1$.

Démonstration. Par définition, $E\{X_t|\mathcal{D}_{t-1}\} = 0$, donc $E\{X_t\} = 0$. On a par définition

$$\text{Var}\{X_t|\mathcal{D}_{t-1}\} = \alpha_0 + \alpha_1 X_{t-1}^2 + \dots + \alpha_p X_{t-p}^2.$$

Si $\sum_{i=1}^p \alpha_i < 1$, puisque X_t est centré, nous avons

$$\begin{aligned} \text{Var}\{X_t\} &= E\{X_t^2\} \\ &= E\{E\{X_t^2|\mathcal{D}_{t-1}\}\} \\ &= E\{\sigma_t^2\} \\ &= E\{\alpha_0 + \alpha_1 X_{t-1}^2 + \dots + \alpha_p X_{t-p}^2\} \\ &= \alpha_0 + \alpha_1 E\{X_{t-1}^2\} + \dots + \alpha_p E\{X_{t-p}^2\} \\ &= \alpha_0 + \alpha_1 \text{Var}\{X_{t-1}\} + \dots + \alpha_p \text{Var}\{X_{t-p}\} \quad \text{car } X \text{ est centré,} \end{aligned}$$

X étant stationnaire, i.e., $V(0) = \text{Var}\{X_t\} = \text{Var}\{X_{t-1}\} = \dots = \text{Var}\{X_{t-p}\}$, d'où

$$\begin{aligned} \text{Var}\{X_t\} &= \alpha_0 + \alpha_1 \text{Var}\{X_t\} + \dots + \alpha_p \text{Var}\{X_t\} \\ &= \alpha_0 + \left(\sum_{i=1}^p \alpha_i\right) \text{Var}\{X_t\}, \end{aligned}$$

alors

$$\left(1 - \sum_{i=1}^p \alpha_i\right) \text{Var}\{X_t\} = \alpha_0,$$

et

$$V(0) = \text{Var}\{X_t\} = \frac{\alpha_0}{1 - \sum_{i=1}^p \alpha_i}.$$

On a

$$\begin{aligned} V(h) &= \text{Cov}(X_t, X_{t+h}) \\ &= E\{X_t X_{t+h}\} - E\{X_t\} E\{X_{t+h}\} \\ &= E\{E\{X_t X_{t+h}|\mathcal{D}_{t+h-1}\}\} - 0 \\ &= E\{X_t E\{X_{t+h}|\mathcal{D}_{t+h-1}\}\} \\ &= E\{0\} \quad \text{car } E\{X_{t+h}|\mathcal{D}_{t+h-1}\} = 0 \\ &= 0, \end{aligned}$$

et

$$\begin{aligned}
\text{Cov}(X_t, X_{t+h} | \mathcal{D}_{t-1}) &= E\{X_t X_{t+h} | \mathcal{D}_{t-1}\} - E\{X_t | \mathcal{D}_{t-1}\} E\{X_{t+h} | \mathcal{D}_{t-1}\} \\
&= E\{E\{X_t X_{t+h} | \mathcal{D}_{t+h-1}\} | \mathcal{D}_{t-1}\} - 0 \\
&= E\{X_t E\{X_{t+h} | \mathcal{D}_{t+h-1}\} | \mathcal{D}_{t-1}\} \\
&= E\{X_t \cdot 0 | \mathcal{D}_{t-1}\} \text{ car } E\{X_{t+h} | \mathcal{D}_{t+h-1}\} = 0 \\
&= E\{0 | \mathcal{D}_{t-1}\} \\
&= 0.
\end{aligned}$$

□

3.4.3 Le processus GARCH

Definition 19. Un processus GARCH (p, q) est un processus $(X_t)_{t \in \mathbb{Z}}$ vérifiant une relation de récurrence du type

$$\begin{cases} X_t = \epsilon_t \sigma_t \text{ avec } (\epsilon_t)_{t \in \mathbb{Z}} \text{ sont des variables aléatoires i.i.d., } \epsilon_t \rightsquigarrow \mathcal{N}(0, 1), \\ \sigma_t^2 = \alpha_0 + \alpha_1 X_{t-1}^2 + \dots + \alpha_p X_{t-p}^2 + \beta_1 \sigma_{t-1}^2 + \dots + \beta_q \sigma_{t-q}^2, \end{cases}$$

avec $\alpha_0 > 0$, $\alpha_i \geq 0$ pour $i = 1, \dots, p$ et $\beta_i \geq 0$ pour $i = 1, \dots, q$.

Si $q = 0$, Un GARCH $(p, 0)$ est un ARCH (p) .

Remarque 20. Dans le cas d'un processus GARCH Si on pose $Y = (X_{t-1}, X_{t-2}, \dots)$, $X = X_t$, on peut écrire

$$f_{X|Y}(x) = \frac{1}{\sigma_t \sqrt{2\pi}} \exp\left(-\frac{x^2}{2\sigma_t^2}\right),$$

avec

$$\text{Var}\{X_t | Y\} = \sigma_t^2.$$

Proposition 21. Soit $(X_t)_{t \in \mathbb{Z}}$ un processus GARCH (p, q) , supposé stationnaire. On note, pour tout t , $\mathcal{D}_t = \{X_t, \sigma_t, X_{t-1}, \sigma_{t-1}, \dots\}$. On a les propriétés suivantes pour tout t et $h \geq 1$:

- $E\{X_t | \mathcal{D}_{t-1}\} = 0$,

Proposition 22. • $E\{X_t\} = 0$,

- $\text{Cov}(X_t, X_{t+h} | \mathcal{D}_{t-1}) = 0$, pour tout $h \geq 1$,
- $\text{Cov}(X_t, X_{t+h}) = 0$, pour tout $h \geq 1$.

Démonstration. Soit $t \geq 1$. Par définition, $E\{X_t|\mathcal{D}_{t-1}\} = 0$, donc $E\{X_t\} = 0$. Si $h > 0$, on a

$$\begin{aligned}
Cov(X_t, X_{t+h}|\mathcal{D}_{t-1}) &= E\{X_t X_{t+h}|\mathcal{D}_{t-1}\} - E\{X_t|\mathcal{D}_{t-1}\} E\{X_{t+h}|\mathcal{D}_{t-1}\} \\
&= E\{E\{X_t X_{t+h}|\mathcal{D}_{t+h-1}\}|\mathcal{D}_{t-1}\} - 0 \\
&= E\{X_t E\{X_{t+h}|\mathcal{D}_{t+h-1}\}|\mathcal{D}_{t-1}\} \\
&= E\{X_t 0|\mathcal{D}_{t-1}\} \text{ car } E\{X_{t+h}|\mathcal{D}_{t+h-1}\} = 0 \\
&= E\{0|\mathcal{D}_{t-1}\} \\
&= 0,
\end{aligned}$$

et donc

$$\begin{aligned}
Cov(X_t, X_{t+h}) &= E\{X_t X_{t+h}\} - E\{X_t\} E\{X_{t+h}\} \\
&= E\{X_t X_{t+h}\} - 0 \\
&= E\{E\{X_t X_{t+h}|\mathcal{D}_{t+h-1}\}\} \\
&= E\{X_t E\{X_{t+h}|\mathcal{D}_{t+h-1}\}\} \\
&= E\{X_t 0\} \text{ car } E\{X_{t+h}|\mathcal{D}_{t+h-1}\} = 0 \\
&= E\{0\} \\
&= 0.
\end{aligned}$$

□

Proposition 23. *Si $(X_t)_{t \in \mathbb{Z}}$ un processus GARCH (p, q) alors $(X_t^2)_{t \in \mathbb{Z}}$ est un processus ARMA (m, q) avec $m = \sup(p, q)$.*

Démonstration. On a, pour tout t ,

$$\sigma_t^2 = \alpha_0 + \alpha_1 X_{t-1}^2 + \dots + \alpha_p X_{t-p}^2 + \beta_1 \sigma_{t-1}^2 + \dots + \beta_q \sigma_{t-q}^2,$$

$$\begin{aligned}
\sigma_t^2 + X_t^2 &= \alpha_0 + \alpha_1 X_{t-1}^2 + \dots + \alpha_p X_{t-p}^2 + X_t^2 + \beta_1 \sigma_{t-1}^2 + \dots + \beta_q \sigma_{t-q}^2 \\
&= \alpha_0 + \alpha_1 X_{t-1}^2 + \dots + \alpha_p X_{t-p}^2 + X_t^2 + \beta_1 X_{t-1}^2 + \dots + \beta_q X_{t-q}^2 \\
&\quad - \beta_1 (X_{t-1}^2 - \sigma_{t-1}^2) - \dots - \beta_q (X_{t-q}^2 - \sigma_{t-q}^2).
\end{aligned}$$

On pose $W_t = X_t^2 - \sigma_t^2$, on obtient

$$X_t^2 = \alpha_0 + \sum_{i=1}^m (\alpha_i + \beta_i) X_{t-i}^2 + W_t - \sum_{i=1}^q \beta_i W_{t-i},$$

avec $\alpha_i = 0$ si $i > p$, $\beta_i = 0$ si $i > q$ et les (W_t) forment un bruit blanc, donc $(X_t^2)_{t \in \mathbb{Z}}$ est un processus $ARMA(m, q)$. \square

Travaux Pratiques

- Le fichier : **UKinterestrates.dat** contient le spread des taux d'intérêts (différence entre taux d'intérêt à long terme et à court terme) pour le Royaume-Uni entre mars 1953 et décembre 1995.

- La fonction : **arima.sim(modele,n)** permet de simuler un processus $ARMA(p, q)$ défini par

$$X_t - \sum_{i=1}^p \alpha_i X_{t-i} = \epsilon_t + \sum_{j=1}^q \beta_j \epsilon_{t-j}.$$

- Les paramètres $(\alpha_i)_{i=1,p}$, $(\beta_j)_{j=1,q}$ du processus sont précisés dans le paramètre **model** de la fonction

```
modele <- list(ar=c(a1,...,ap),ma=c(b1,...,bq))
```

- (choisir **ma** ou **ar** pour un processus MA ou AR , si $p = 1$, écrire $ar = a1$).

- Pour simuler un modèle $ARIMA(p, d, q)$ il faut ajouter le composant **order=c(p,d,q)** dans le paramètre **modele** de la fonction **arima.sim**.

- La fonction **ar** permet d'estimer les paramètres d'un processus $AR(p)$:

```
out <- ar(data,aic=TRUE,order.max=NULL)
```

- L'ordre p du processus auto-régressif est choisi (inférieur à **order.max**) à l'aide du critère **AIC** (si l'option **aic** est validée). La fonction **arima** permet d'estimer les paramètres :

- d'un $ARMA(p, q)$:

```
out <- arima(serie,order=c(p,0,q))
```

- d'un $ARIMA(p, d, q)$:

```
out <- arima(serie,order=c(p,d,q))
```

◦ d'un $SARIMA(p, d, q, T)$:

```
out <-arima(serie,order=c(p,d,q),seasonal=list(order=c(P,D,Q),period=T)
```

· Les paramètres P, D, Q du modèle $SARIMA$ ne sont pas abordés dans ce cours, nous leur donnerons par défaut la valeur des paramètres p, d, q (pas trop grands).

· Parmi les sorties de cette fonction, on peut obtenir :

* **out\$coef** : estimation des coefficients,

* **out\$aic** : valeur du critère **AIC**,

* **out\$resid** : estimation des résidus.

• La fonction **p=predict(out,h)** permet d'effectuer une prévision à l'horizon **h**. Parmi les sorties de cette fonction :

· **p\$pred** : contient les prévisions.

· **p\$se** : contient l'écart-type de l'erreur de prévision.

• **EuStockMarket** : valeurs à la fermeture de divers indices des marchés européens.

• Charger les logarithmes des accroissement de l'indice **DAX** en exécutant :

```
dax <-diff(log(EuStockMarkets)) [,"DAX"]
```

· **dax2=dax^2** :

• **NYSE** : New-York stock exchange daily return.

· Charger le fichier **nyse.dat** à l'adresse : <http://math.unice.fr/~rubentha/enseignement/nyse.dat>

• Pour utiliser les fonctions spécifiques à l'étude des modèles $ARCH$ et $GARCH$, il faut avant tout charger le package **tseries** à l'aide de la commande **library(tseries)**.

• La fonction **garch** permet d'estimer un $GARCH(p,q)$:

```
serie <- garch(data,order=c(p,q))
```

· Parmi les sorties de cette fonction : **coef**, **residuals**, **fitted.values**.

◦ La sortie **serie\$fitted.values** contient pour chaque temps **t** les valeurs $-\sigma_t$ et σ_t , avec la notation standard : $X_t = \sigma_t \epsilon_t$ et $\epsilon_t \rightsquigarrow \mathcal{N}(0, 1)$ et σ_t^2 est la variance calculée à partir de $X_{t-1}, X_{t-2}, \dots, \sigma_{t-1}, \sigma_{t-2}, \dots$

• si **A** est un tableau à deux colonnes alors **A[,j]** est la j-ème colonne de **A**.

• Pour fixer la taille de la fenêtre graphique en ordonnée, par exemple : **plot(x,ylim=c(-0.7,0.9))**.

Bibliographie

- [1] Brockwell P.J., David R.A. 2002. Introduction to time series and forecasting. Springer
- [2] Chatfield C. 2004. The analysis of time series. An introduction. (6th edition). Chapman & Hall/CRC.
- [3] Harvey A.C. 1991. Forecasting, structural time series models and the Kalman filter. Cambridge University Press.
- [4] D, Monod H. 2011. Analyse statistique des risques agro-environnementaux. Springer.
- [5] L., Makowski D. 2013. Comparison of statistical models for analyzing wheat yield time series. Plos one 8(10).
- [6] Petris G. 2010. A R package for dynamic linear models. Journal of Statistical Software 36, 1-14.
- [7] W.N., Ripley B.D. 2002. Modern applied statistics with S. Springer. Chapitre 14.
- [8] David Makowski. 2014. Introduction aux séries chronologiques et à la prévision. Agro-ParisTech, novembre 2014.
- [9] Julien Ah-Pine. 2012. Séries chronologiques saisonnières et prévision. Université Lyon 2 - FSEG.
- [10] M.-C. Viano., 2004. Maîtrise d'Économétrie. Cours de Séries Temporelles. Années 1999 à 2004. Université des Sciences et Technologies de Lille.
- [11] V. Monbet. 2017. Modélisation des séries temporelles. Master Statistique et Économétrie. Notes de cours.
- [12] A. Godichon-Baggioni. 2018. Analyse descriptive des S eries Chronologiques. (INSA de Rouen).

- [13] Sylvain Rubenthaler. 2018. Séries chronologiques (avec R). Université Nice.
- [14] Frédéric Sur. Modélisation et prévision : séries chronologiques. Cours GIMAS8AD (département GIMA, 2ème année École des Mines de Nancy). https://members.loria.fr/FSur/enseignement/modprev/modprev_chrono_FSur_1.pdf
- [15] J.J. Daudin, C. Duby, S. Robin & P. Trécourt. 1996. Analyse de Séries Chronologiques. INA-PG, Mathématiques.