Analyse des séries temporelles avec R

Patrick Ilunga

2023-12-27

Contents

1 Introduction **5**

1.1 Usage 5

1.2 Objectif de l'analyse d'une série temporelle 5

2 Description **7**

2.1 Définitions 7

2.2 Notation 7

2.3 Exemple 8

2.4 Statistiques descriptives usuelles 10

2.5 Représentation graphique 14

2.6 Structure ou caracteristiques 19

2.7 Etude de la normalité 22

2.8 Test d'indépendance 28

2.9 Performances d'une série temporelle 29

3 Modélisation **31**

3.1 Stationnarité 31

3.2 Modèle sans variable explicative 38

3.3 Modèle statique 38

3.4 Modèle dynamique 38

4 Prédiction **39**

4.1 Partant d'un modèle 39

4.2 Sans reference à un modèle 39

5 Footnotes and citations **41**
5.1 Footnotes 41
5.2 Citations 41

6 Blocks **43**
6.1 Equations 43
6.2 Theorems and proofs 43
6.3 Callout blocks 43

7 Sharing your book **45**
7.1 Publishing 45
7.2 404 pages 45
7.3 Metadata for sharing 45

Chapter 1

Introduction

This is a *sample* book written in **Markdown**. You can use anything that Pandoc's Markdown supports; for example, a math equation $a^2 + b^2 = c^2$.

1.1 Usage

Each **bookdown** chapter is an .Rmd file, and each .Rmd file can contain one (and only one) chapter. A chapter *must* start with a first-level heading: `# A good chapter`, and can contain one (and only one) first-level heading.

Use second-level and higher headings within chapters like: `## A short section` or `### An even shorter section`.

The `index.Rmd` file is required, and is also your first book chapter. It will be the homepage when you render the book.

1.2 Objectif de l'analyse d'une série temporelle

Quoique l'on dit, le but ultime d'analyse d'une série temporelle est d'arriver à comprendre son passé et son present pour prédire son futur. Ceci passe par la description et la modélisation de la série.

- Décrire la série :
- Modéliser la série :
- Prédire les valeurs futures :

Chapter 2

Description

All chapters start with a first-level heading followed by your chapter title, like the line above. There should be only one first-level heading (`#`) per .Rmd file.

2.1 Définitions

2.1.1 Définition 1 : Série temporelle

On appelle par série temporelle, une suite ordonnée des valeurs numériques décrivant l'évolution d'un phénomène spécifique au cours du temps. On l'appelle aussi série chronologique. Mathématiquement, on parle d'une suite finie des données indexée par le temps.

2.2 Notation

On note une série temporelle par :

$$\{x_t\}_{t \in T}$$

Où :

- $t_1, t_2, ..., t_n$ sont les n instants d'observation du phénomène, ca peut etre le jour, le mois, l'année …
- T est l'ensemble ordonné des instants, i.e : $T = \{t_1, t_2, ..., t_n\}$,
- x_{t_i} est la valeur du phénomène observé à l'instant t_i, avec $1 \leq i \leq n$,

On peut aussi la noter par l'ensemble des couples ci-après :

$$(t_i, x_i)_{1 \leq i \leq n}$$

Où :

- la première composante l'instant d'observation t_i,
- la deuxième composante est la valeur du phénomène x observée à l'instant t_i.

ou encore x_t.

La serie x_t est la realisation d'un processus aléatoire X_t. Les X_t sont des variables aléatoires (v.a), et x_t est la valeur prise par X_t à l'instant t. Ces v.a à des instants differents sont normalement corrélées.

2.2.1 Définition 2 : Le bruit blanc

Un bruit blanc est une suite z_t de v.a. non corrélées de moyenne nulle et de variance constante. Elle est considérée comme la série de reference.

On note $z_t \sim BB(0, \sigma_z^2)$ pour designer un bruit blanc.

2.3 Exemple

- Le chiffre d'affaire mensuel (en million de dollars) du service Voix mobile en République Démocratique du Congo pour l'ensemble des opérateurs de téléphonie (Fig. 1.1). La croissance de l'activité a brusquement changé suite à l'inflation du marché.

- Le nombre de minutes consommées[1] par jour en 2021 par l'ensemble des clients ayant passé un appel vocal (Fig. 1.2). La consommation baisse les jours week-ends, suite du fait de l'arrêt des activités économiques.

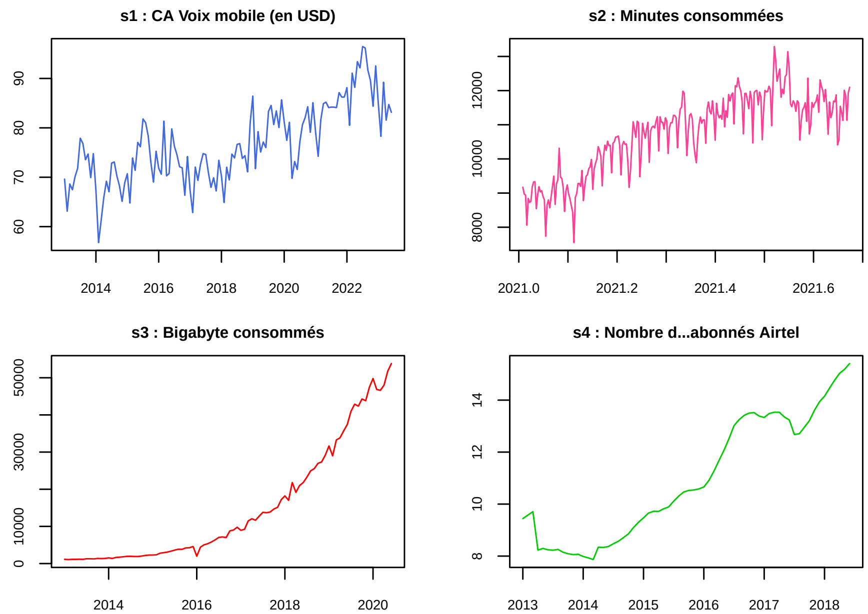
- Le nombre de Terabyte consommé par l'ensemble des clients ayant naviguer sur internet (Fig. 1.3). La croissance est exponentielle, favorisée plus par l'augmentation d'une part du nombre des smartphones et de la baisse du prix de service internet mobile sur le marché congolais.

- Le nombre de clients actifs sur une periode de 90 jours de l'operateur Airtel (Fig. 1.4).

- On simule une marche aléatoire avec les fonctions R *cumsum()* et *rnorm()* (Fig. 1.6).

[1] Données simulées

```
## Warning in title(main = main, xlab = xlab, ylab = ylab, ...): conversion
## failure on 's4 : Nombre d'abonnés Airtel' in 'mbcsToSbcs': dot substituted for
## <e2>

## Warning in title(main = main, xlab = xlab, ylab = ylab, ...): conversion
## failure on 's4 : Nombre d'abonnés Airtel' in 'mbcsToSbcs': dot substituted for
## <80>

## Warning in title(main = main, xlab = xlab, ylab = ylab, ...): conversion
## failure on 's4 : Nombre d'abonnés Airtel' in 'mbcsToSbcs': dot substituted for
## <99>
```

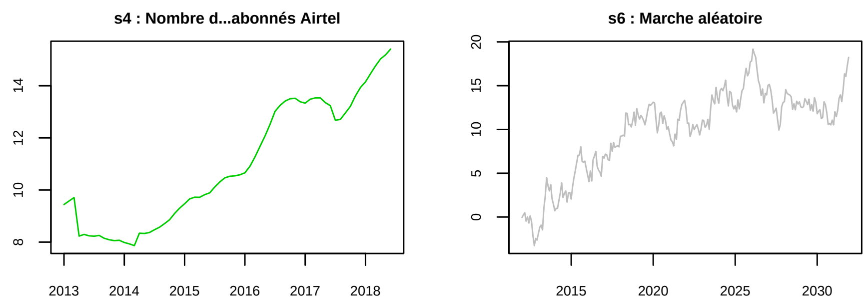



```
par(mfrow = c(2,2), oma = c(0, 0, 0, 0), mar = c(2.5, 2.5, 2,2), cex.main = 0.8, cex.lab = 0.7, c
plot(s4, col = "green3", xlab = "", main = "s4 : Nombre d'abonnés Airtel")
```

```
## Warning in title(main = main, xlab = xlab, ylab = ylab, ...): conversion
## failure on 's4 : Nombre d'abonnés Airtel' in 'mbcsToSbcs': dot substituted for
## <e2>

## Warning in title(main = main, xlab = xlab, ylab = ylab, ...): conversion
## failure on 's4 : Nombre d'abonnés Airtel' in 'mbcsToSbcs': dot substituted for
## <80>
```

```
## Warning in title(main = main, xlab = xlab, ylab = ylab, ...): conversion
## failure on 's4 : Nombre d'abonnés Airtel' in 'mbcsToSbcs': dot substituted for
## <99>
```

```
plot(s6, col = "grey", xlab = "", main = "s6 : Marche aléatoire")
```



2.4 Statistiques descriptives usuelles

Soit x_t une série temporelle avec $t \in T$,

2.4.1 Les statistiques de tendance centrale

- La moyenne empirique

La moyenne empirique de la série est donnée par :

$$\overline{x} = \frac{1}{n} \sum_{t=1}^{n} x_t$$

2.4.2 Les statistiques de dispersion

- La variance empirique

La variance empirique est donnée par :

$$\sigma^2 = \frac{1}{n}\sum_{t=1}^{n}(x_t - \overline{x})^2$$

L'ecart-type est definit par la racine carrée de la variance empirique : σ.

- Le coefficient de variation

Le coefficient de variation est donné par :

$$cv = \frac{\sigma}{\overline{x}}$$

Permet de determiner l'homogénéité ou la dispersion et est exprimé en pourcentage.

- Pour $cv < 15$, on dit que la série est homogène.
- Pour $cv > 15$, on dit que les valeurs de la série sont relativement dispersées.

```
round(sd(s1)/mean(s1)*100, 2)
```

```
## [1] 10.23
```

```
round(sd(s2)/mean(s2)*100, 2)
```

```
## [1] 10.67
```

```
round(sd(s3)/mean(s3)*100, 2)
```

```
## [1] 105.46
```

```
round(sd(s4)/mean(s4)*100, 2)
```

```
## [1] 21.47
```

Ainsi, on peut observer qu'il y a de forte dispersion avec la $s3$ et $s4$ contrairement aux observations de série $s1$ et $s2$ qui elles sont homogènes.

2.4.3 Statistiques de dépendance

Celles qui renseignent sur la dépendance entre les observations de la série x_t :

- L'auto_covariance

L'auto-covariance d'ordre h de la série est donnée par :

$$\hat{\sigma}(h) = \frac{1}{n-h} \sum_{t=1}^{n-h} (x_t - \overline{x})(x_{t+h} - \overline{x})$$

Avec $h \in N^*$ et $h < n$. Dans le cas où $h = 0$, c'est-à-dire $\hat{\sigma}(0)$, on a la variance empirique σ^2.

- L'auto-correlation

L'auto-correlation de la série est donnée par :

$$\hat{\gamma}(h) = \frac{\hat{\sigma}(h)}{\hat{\sigma}(0)}$$

2.4.4 Les statistiques des formes

La loi normale se caractérise par un coefficient d'asymétrie et un coefficient d'aplatissement nuls. Ainsi les calculs de ces coefficients peuvent nous aider à avoir une idée sur la normalité.

- Les coefficients d'aplatissement

Le coefficient d'aplatissemnt (ou Kurtosis) permet de mesurer l'aplatissement de la distribution de la série. On s'en sert aussi pour vérifier si une distribution est normale. Le coefficient est donné par :

$$P = \frac{n(n+1)}{(n-1)(n-2)(n-3)} \sum_{t=1}^{n} \left(\frac{x_t - \overline{x}}{\sigma} \right)^4 - \frac{3(n-1)^2}{(n-2)(n-3)}$$

ou

$$K = \frac{1}{n\sigma^4} \sum_{t=1}^{n} (x_t - \overline{x})^4$$

Pour : - Si $K = 3$, la distribution est normale, - Si $K > 3$, la distribution est plus applatie, c'est-à-dire, une distribution relativement hétérogène, avec beaucoup de valeurs éloignées de la moyenne. - Si $K < 3$, la distribution est moins applatie, c'est-à-dire, il y a peu de variations dans les valeurs observées, une distribution relativement homogène, avec beaucoup de valeurs égales ou proches de la moyenne.

```
library(moments)
kurtosis(s1)
```

```
## [1] 2.665041
```

```
kurtosis(s2)
```

```
## [1] 2.419901
```

```
kurtosis(s3)
```

```
## [1] 2.753559
```

```
kurtosis(s4)
```

```
## [1] 1.619768
```

```
kurtosis(s6)
```

```
## [1] 2.990772
```

- Le coefficients d'asymétrie

Le coefficient d'asymetrie (de Fisher) est donner par :

$$A = \frac{n}{(n-1)(n-2)} \sum_{t=1}^{n} \left(\frac{x_t - \overline{x}}{\sigma} \right)^3$$

ou

$$F = \frac{1}{n\sigma^3} \sum_{t=1}^{n} (x_t - \mu)^3$$

L'interprétation de ces coefficients est directe :

- Si le coefficient est nul, la distribution est symétrique
- Si le coefficient est négatif, la distribution est déformée à gauche de la médiane (sur-représentation de valeurs faibles, à gauche)
- Si le coefficient est positif, la distribution est déformée à droite de la médiane (sur-représentation de valeurs fortes, à droite)

```r
skewness(s1)
```

```
## [1] 0.307306
```

```r
skewness(s2)
```

```
## [1] -0.4777813
```

```r
skewness(s3)
```

```
## [1] 1.035407
```

```r
skewness(s4)
```

```
## [1] 0.25069
```

```r
skewness(s6)
```

```
## [1] -0.8200721
```

2.5 Représentation graphique

2.5.1 Le chronogramme

Le chronogramme (ou time plot) est la représentation graphique d'une série temporelle. Les figures de l'exemple sont des chronogrammes. Sous R on peut utiliser d'autres paramètres pour ajouter les éléments dans les différentes figures.

2.5.2 L'histogramme

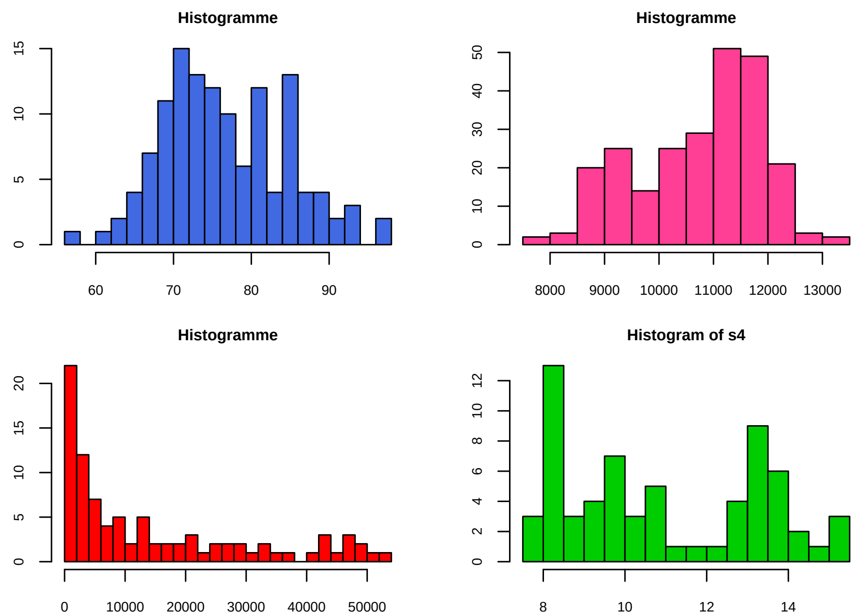
La représentation qui permet d'avoir une idée sur la distribution des valeurs de la série x_t.

```
## Warning in plot.window(xlim, ylim, "", ...): "min" is not a graphical parameter
```

```
## Warning in title(main = main, sub = sub, xlab = xlab, ylab = ylab, ...): "min"
## is not a graphical parameter
```

```
## Warning in axis(1, ...): "min" is not a graphical parameter
```

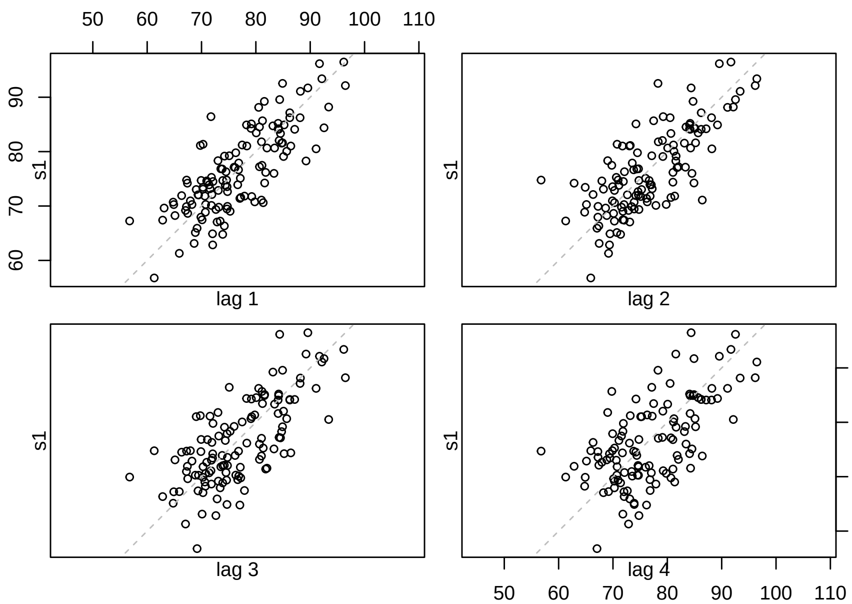
```
## Warning in axis(2, at = yt, ...): "min" is not a graphical parameter
```



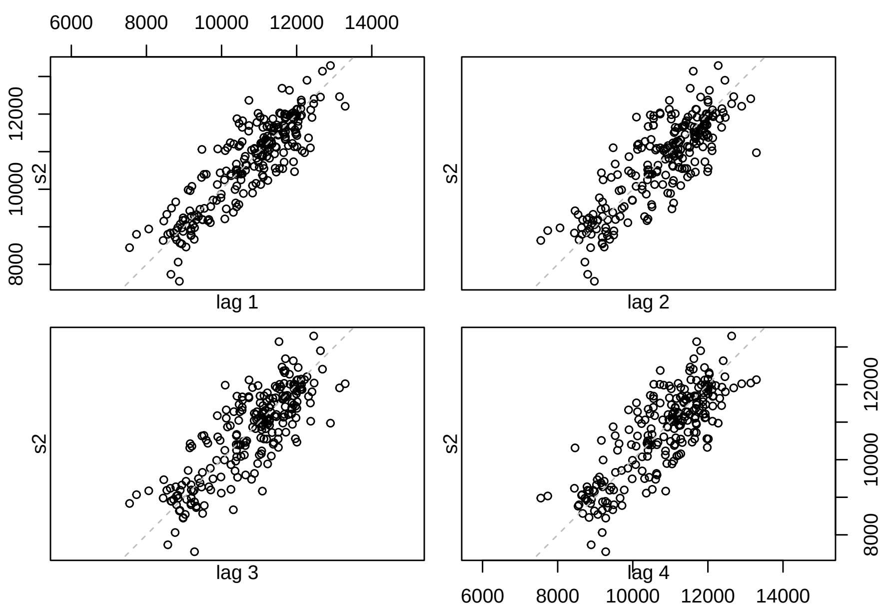
2.5.3 Le diagramme retardé

Le diagramme retardé (ou Lagplot) est une représentation qui permet de comprendre la dépendance des observations de la série x_t obtenu avec la fonction *lagplot()*.

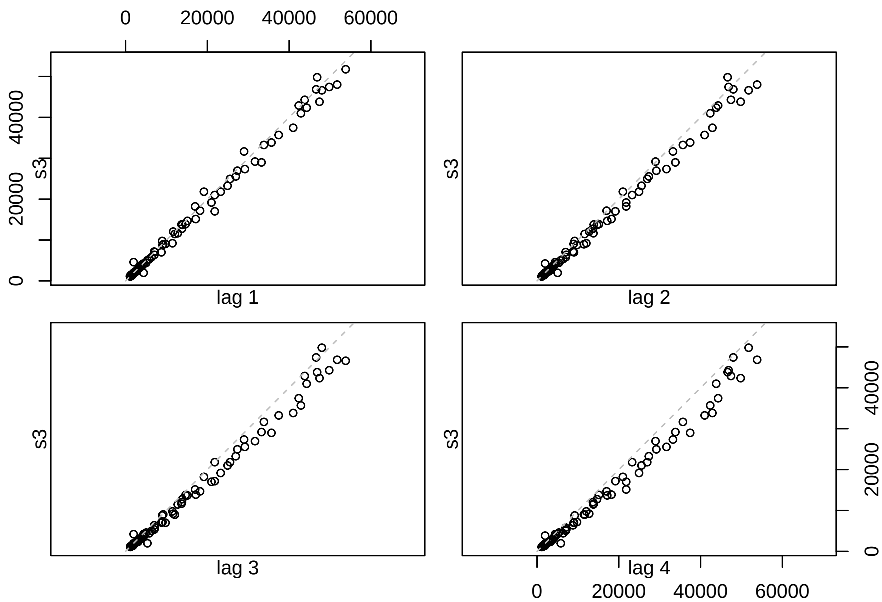
```
lag.plot(s1, 4, do.lines = FALSE)
```



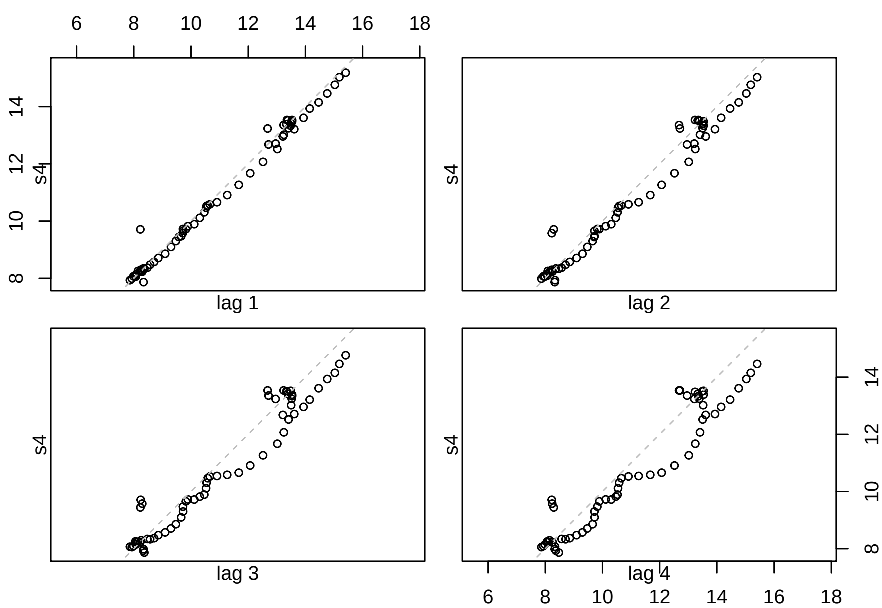

```
lag.plot(s2, 4, do.lines = FALSE)
```



```
lag.plot(s3, 4, do.lines = FALSE)
```

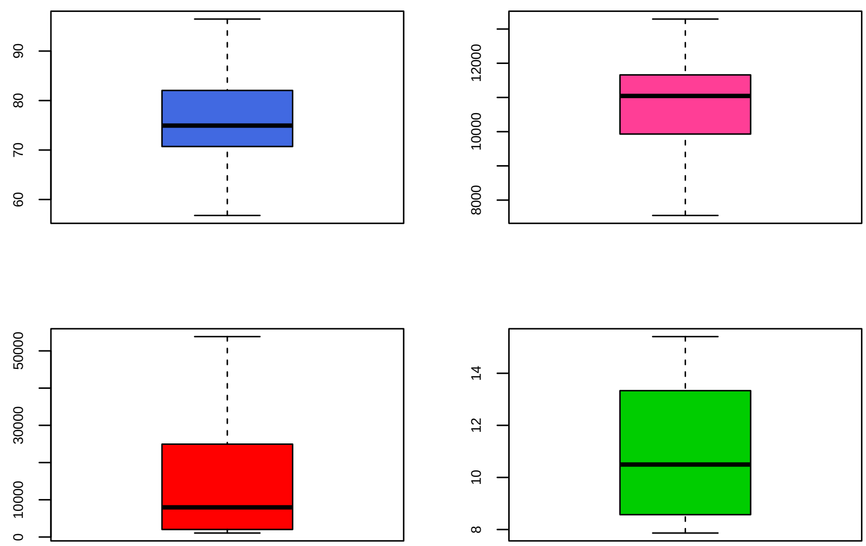



```
lag.plot(s4, 4, do.lines = FALSE)
```



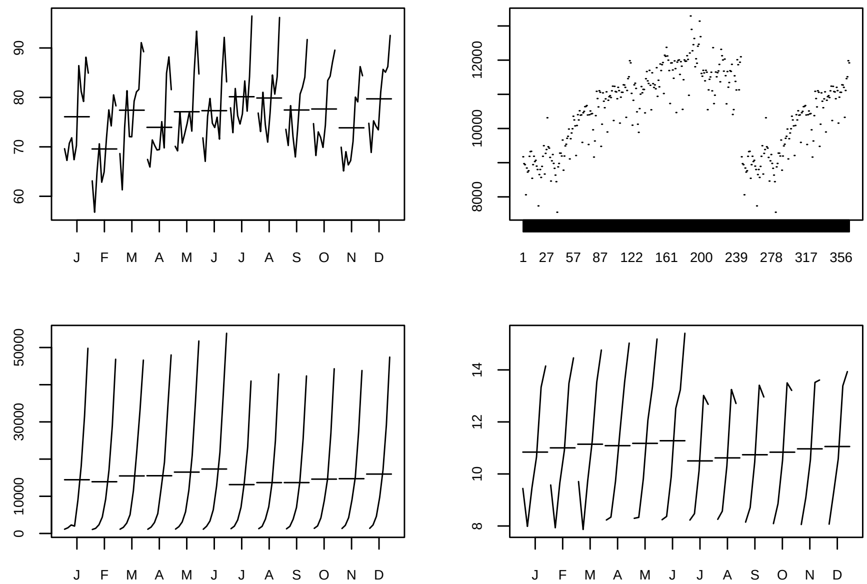
2.5.4 La boîte à moustache

La boîte à moustache (ou boxplot) est une représentation qui permet de repérer les valeurs atypiques de la série x_t. Elle permet aussi d'avoir une idée sur la symetrie de la distribution quoique cette symetrie n'affirme pas la normalité. La boite est symetrique lorsque la mediane se situe au milieu de la boîte à moustache et qu'il y a symetrie des moustaches.



2.5.5 Le month plot

```
par(mfrow = c(2,2), oma = c(0, 0, 0, 0), mar = c(2.5, 2.5, 2,2), cex.main = 0.8, cex.la
monthplot(s1)
monthplot(s2)
monthplot(s3)
monthplot(s4)
```



2.6 Structure ou caracteristiques

Une série temporelle x_t est caractérisée par une tendance, une saisonnalité et résidus.

2.6.1 La tendance

Une tendance (ou trend) est l'orientation à la hausse ou à la baisse des valeurs observées d'une série temporelle sur une periode assez longue. On note par m_t. Une tendance est soit :

- locale ou globale :
- linéaire ou non linéaire :

2.6.2 La saisonnalité

La saisonnalité permet de détecter les éléments repetitifs sur les observations d'une série après une période. La répétition peut etre mensuelle ou hebdomadaire. On note par s_t la composante saisonnière d'une série.

2.6.3 Erreur (Le bruit)

Les 3 composants ci-dessus permet de decomposer la série x_t à un modèle additif :

$$x_t = m_t + s_t + u_t$$

avec $E(u_t) = 0$, ou soit à un modèle multiplicatif :

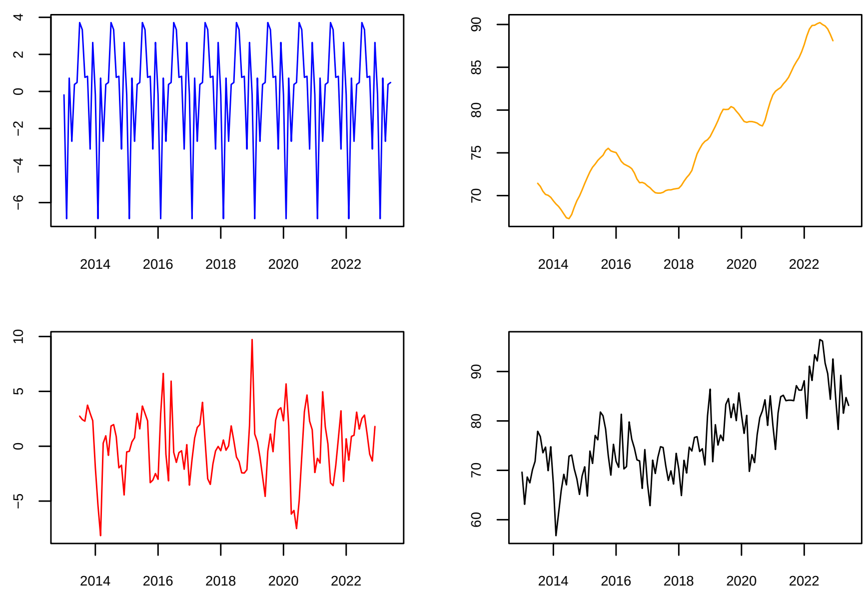
$$x_t = m_t s_t u_t$$

avec $E(u_t) = 1$.

Avec le logiciel R, on peut décomposer la série avec la fonction *decompose()*.

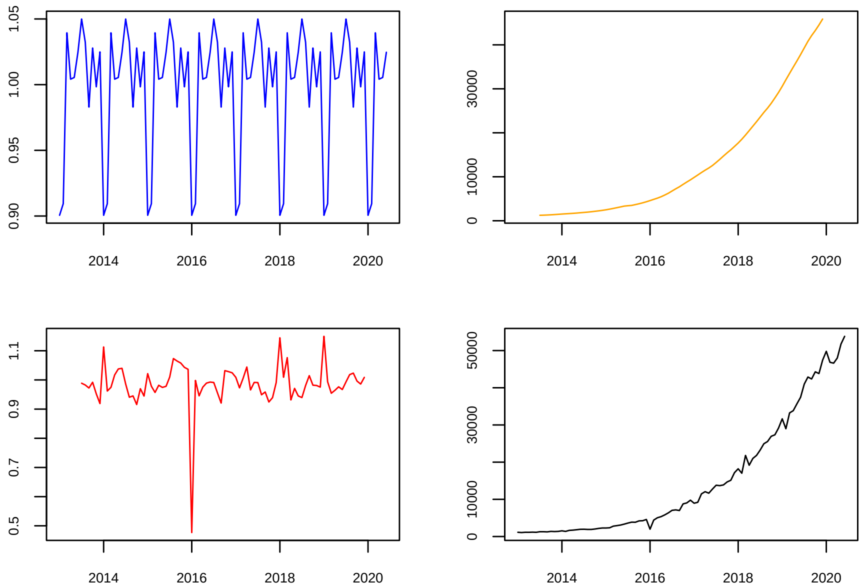
Pour la série $s1$ on a :

```
m1 = decompose(s1, type = "additive")

par(mfrow = c(2,2), oma = c(0, 0, 0, 0), mar = c(2.5, 2.5, 2,2), cex.main = 0.8, cex.la
plot(m1$seasonal, col = "blue", ylab = "Saisonnalité")
plot(m1$trend, col = "orange", ylab = "Tendance")
plot(m1$random, col = "red", ylab = "Random")
plot(m1$x, col = "black", ylab = "Valeur")
```



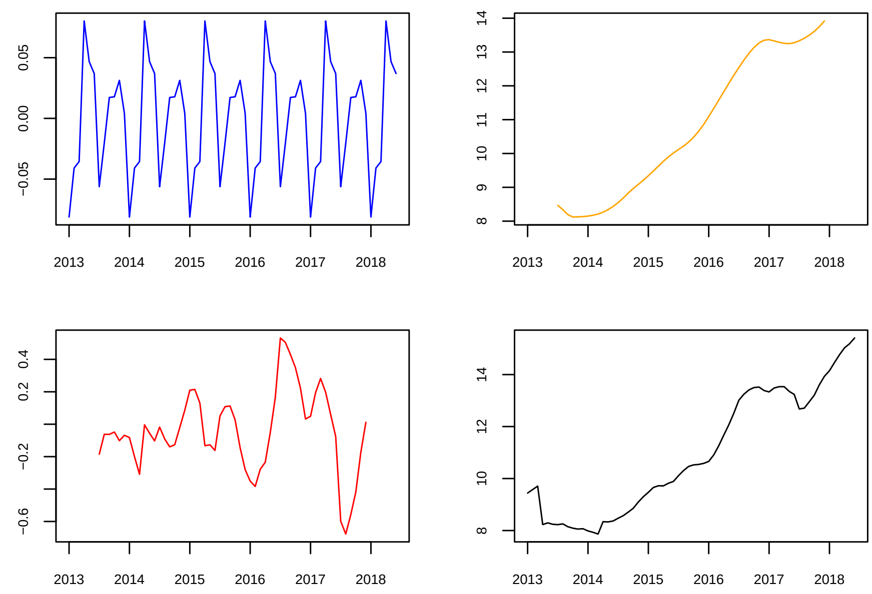
Pour la série $s3$ on a :

```
m3 = decompose(s3, type = "multiplicative")

par(mfrow = c(2,2), oma = c(0, 0, 0, 0), mar = c(2.5, 2.5, 2,2), cex.main = 0.8, cex.lab = 0.7, c
plot(m3$seasonal, col = "blue", ylab = "Saisonnalité")
plot(m3$trend, col = "orange", ylab = "Tendance")
plot(m3$random, col = "red", ylab = "Random")
plot(m3$x, col = "black", ylab = "Valeur")
```



Pour la série $s4$ on a :

```
m4 = decompose(s4, type = "additive")

par(mfrow = c(2,2), oma = c(0, 0, 0, 0), mar = c(2.5, 2.5, 2,2), cex.main = 0.8, cex.lab = 0.7, c
plot(m4$seasonal, col = "blue", ylab = "Saisonnalité")
plot(m4$trend, col = "orange", ylab = "Tendance")
plot(m4$random, col = "red", ylab = "Random")
plot(m4$x, col = "black", ylab = "Valeur")
```

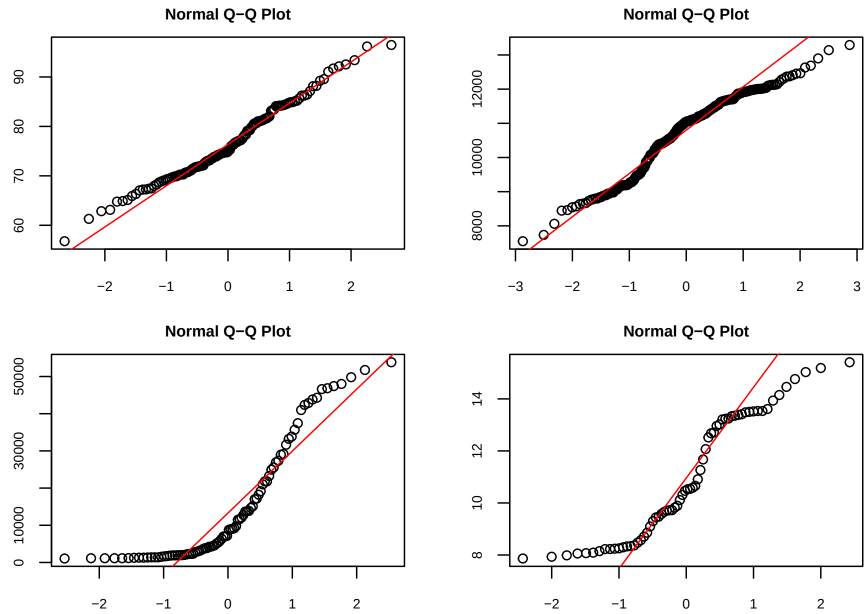


2.7 Etude de la normalité

2.7.1 Le diagramme Quantile-Quantile

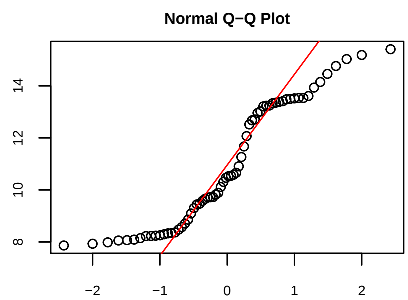
Le diagramme Quantile-Quantile (ou QQ-Plot) est une représentation d'évaluer la pertinence de l'ajustement d'une distribution donnée à un modèle théorique. De ce fait, elle permet aussi de reconnaître un bruit blanc.

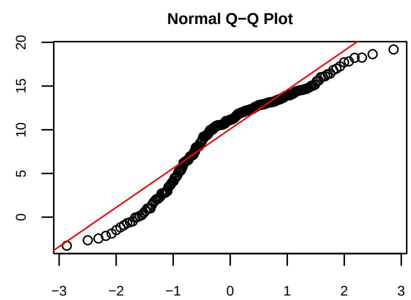
Plus les points se rapprochent de la droite, plus la distribution empirique est dite normale.



Interprétation :

```
par(mfrow = c(2,2), oma = c(0, 0, 0, 0), mar = c(2.5, 2.5, 2,2), cex.main = 0.8, cex.lab = 0.7, c

qqnorm(s4)
qqline(s4, col = "red")

qqnorm(s6)
qqline(s6, col = "red")
```





2.7.2 Test de normalité

2.7.2.1 Test de Shapiro-Wilk

Il est appliquable pour des échantillons allant jusqu'à 50 valeurs.

```r
shapiro.test(s1)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  s1
## W = 0.98342, p-value = 0.1267
```

```r
shapiro.test(s2)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  s2
## W = 0.95935, p-value = 2.185e-06
```

```r
shapiro.test(s3)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  s3
## W = 0.81534, p-value = 3.025e-09
```

```r
shapiro.test(s4)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  s4
## W = 0.90233, p-value = 7.706e-05
```

```r
shapiro.test(s6)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  s6
## W = 0.92795, p-value = 2.001e-09
```

A $n = 30$ *(ddl)*, $W_{0.05}$ vaut 0.927. On a ce qui suit :

- Pour $s1$, $W_{obs} > W_{0,05}$, avec le *p-value* est superieur à 0.05, l'hypothèse de normalité est donc tolérée. C'est-à-dire les données suivent une distribution normale.
- Pour $s2$, $W_{obs} > W_{0,05}$, avec le *p-value* est inférieur à 0.05, …..
- Pour $s3$ et $s4$, $W_{obs} < W_{0,05}$, avec les *p-value* toutes inférieures à 0.05, l'hypothse de normalité est donc rejetée. C'est-à-dire les données ne suivent pas une distribution normale.

2.7.2.2 Test de Kolmogorov-Smirnov

Le test de Kolmogorov-Smirnov permet de tester l'ajustement des données x à n'importe quelle loi, dont la loi normale. Il est intéressant d'opter pour ce test plutôt que celui de Shapiro-Wilk en cas de très grands échantillons.

```r
# Test de normalité de Kolmogorov S.
ks.test(s1, y = "pnorm", mean(s1), sd(s1))
```

```
## Warning in ks.test.default(s1, y = "pnorm", mean(s1), sd(s1)): ties should not
## be present for the Kolmogorov-Smirnov test
```

```
##
##  Asymptotic one-sample Kolmogorov-Smirnov test
##
## data:  s1
## D = 0.092237, p-value = 0.234
## alternative hypothesis: two-sided
```

```r
ks.test(s2, y = "pnorm", mean(s2), sd(s2))
```

```
##
##  Asymptotic one-sample Kolmogorov-Smirnov test
##
## data:  s2
## D = 0.10765, p-value = 0.007001
## alternative hypothesis: two-sided
```

```r
ks.test(s3, y = "pnorm", mean(s3), sd(s3))
```

```
## Warning in ks.test.default(s3, y = "pnorm", mean(s3), sd(s3)): ties should not
## be present for the Kolmogorov-Smirnov test
```

```
##
##  Asymptotic one-sample Kolmogorov-Smirnov test
##
## data:  s3
## D = 0.18943, p-value = 0.003133
## alternative hypothesis: two-sided
```

```r
ks.test(s4, y = "pnorm", mean(s4), sd(s4))
```

```
##
##  Exact one-sample Kolmogorov-Smirnov test
##
## data:  s4
## D = 0.13672, p-value = 0.1544
## alternative hypothesis: two-sided
```

```r
ks.test(s6, y = "pnorm", mean(s6), sd(s6))
```

```
##
##  Asymptotic one-sample Kolmogorov-Smirnov test
##
## data:  s6
## D = 0.14847, p-value = 5.078e-05
## alternative hypothesis: two-sided
```

Pour $D_{0.05} = 0.240$, (pour n = 30) :

- Si $D_{max} < D_{0.05}$, l'H_0 est tolérée: la distribution empirique semble correspondre à une distribution normale.
- Si $D_{max} > D_{0.05}$, l'H_0 est rejetée: la distribution empirique ne provient pas d'une distribution normale.

Ainsi, avec ce test, on tolère l'hypothèse H_0, les distributions des séries semblent correspondre à une distribution normale.

```r
# Test de normalité
library(tseries)
```

```
## Registered S3 method overwritten by 'quantmod':
##   method            from
##   as.zoo.data.frame zoo
```

```r
jarque.bera.test(s1)
```

```
##
##  Jarque Bera Test
##
## data:  s1
## X-squared = 2.5722, df = 2, p-value = 0.2763
```

```r
jarque.bera.test(s2)
```

```
##
##  Jarque Bera Test
##
## data:  s2
## X-squared = 12.704, df = 2, p-value = 0.001743
```

```r
jarque.bera.test(s3)
```

```
##
##  Jarque Bera Test
##
## data:  s3
## X-squared = 16.309, df = 2, p-value = 0.0002875
```

```r
jarque.bera.test(s4)
```

```
##
##  Jarque Bera Test
##
## data:  s4
## X-squared = 5.9302, df = 2, p-value = 0.05156
```

```r
jarque.bera.test(s6)
```

```
##
##  Jarque Bera Test
##
## data:  s6
## X-squared = 26.902, df = 2, p-value = 1.44e-06
```

2.8 Test d'indépendance

On peut utiliser la fonction *Box.test()* pour examiner l'hypothèse nulle d'indépendance dans une série temporelle donnée.

```r
Box.test(s1, lag = 1, type = "Ljung-Box")
```

```
##
##  Box-Ljung test
##
## data:  s1
## X-squared = 79.334, df = 1, p-value < 2.2e-16
```

```r
Box.test(s1, lag = 2, type = "Ljung-Box")
```

```
##
##  Box-Ljung test
##
## data:  s1
## X-squared = 149.46, df = 2, p-value < 2.2e-16
```

```
Box.test(s1, lag = 3, type = "Ljung-Box")
```

```
##
##  Box-Ljung test
##
## data:  s1
## X-squared = 212.03, df = 3, p-value < 2.2e-16
```

2.9 Performances d'une série temporelle

Dans les domaines de finance, de comptabilité ou de marketing opérationnel, les experts suivent au quotidien l'évolution de leurs activités en comparant à chaque fois les résultats d'une date (ou d'une période) par rapport à une date (ou une période) antérieure, par rapport aux résultats du marché global ou par rapport aux résultats de leurs concurrents directs. On parle là de concept de *performance*.

2.9.1 La performance Year-to-date

La performance *Year-To-Date* conrespond au résultat cumulé commençant le premier jour de l'année civile ou de l'exercice en cours jusqu'à la date actuelle,

$$YTD_x = \sum_{t=1}^{k} x_t$$

x_1 correspond à la valeur de la série au 1er janvier, et x_k à la valeur de la série à la date actuelle.

2.9.2 La performance Month-to-date

La performance *Month-To-Date* conrespond au résultat cumulé commençant le premier jour du mois en cours jusqu'à la date actuelle au cours du même mois,

$$MTD_x = \sum_{t=1}^{k} x_t$$

x_1 correspond à la valeur de la série au 1er jour du mois, et x_k à la valeur de la série à la date actuelle au cours du même mois.

2.9.3 La performance Quater-to-date

La performance *Quater-To-Date* conrespond au résultat cumulé commençant le permier jour du trimestre jusqu'à la date actuelle au cours du même trimestre.

$$QTD_x = \sum_{t=1}^{k} x_t$$

x_1 correspond à la valeur de la série au 1er jour du trimestre, et x_k à la valeur de la série à la date actuelle au cours du même trimestre.

2.9.4 La performance Week-to-date

La performance *Week-To-Date* conrespond au résultat cumulé commençant le permier jour de la semaine jusqu'à la date actuelle au cours de la même semaine.

$$WTD_x = \sum_{t=1}^{k} x_t$$

x_1 correspond à la valeur de la série au 1er jour de la semaine, et x_k à la valeur de la série à la date actuelle au cours de la même semaine.

Chapter 3

Modélisation

3.1 Stationnarité

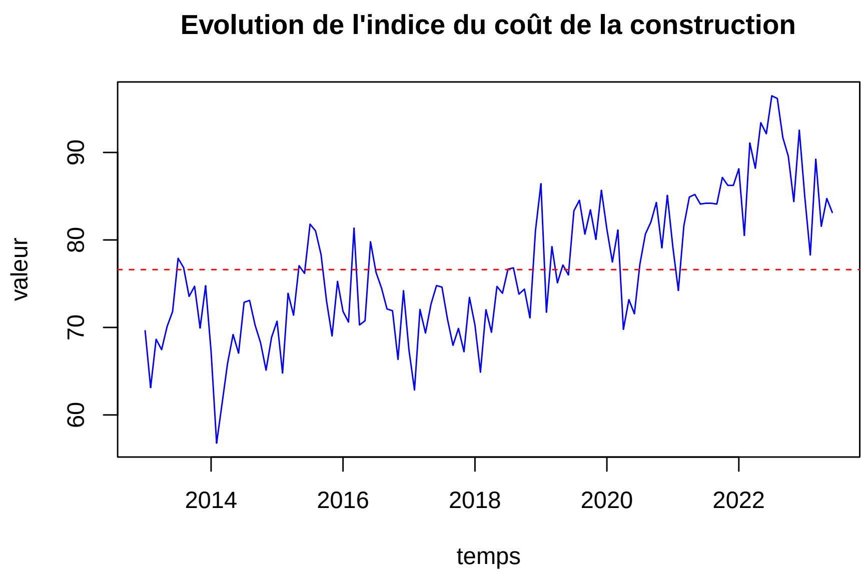
3.1.1 Definition 2.1.

xxxx

3.1.2 Graph

La série est-elle stationnaire?

```
plot(s1, xlab = "temps", ylab = "valeur", col = "blue",
     main = "Evolution de l'indice du coût de la construction")
abline(h = mean(s1), col = "red", lty = 2)
```



On voit clairement sur le chronogramme de la série qu'elle ne pourrait pas être stationnaire : sa moyenne depend de l'instant auquel on se trouve et le niveau de la série ne fluctue pas autour d'une valeur quelconque.

3.1.3 Test de

Test de stationnarité

```
library(aTSA)
```

```
##
## Attaching package: 'aTSA'

## The following objects are masked from 'package:tseries':
##
##     adf.test, kpss.test, pp.test

## The following object is masked from 'package:graphics':
##
##     identify
```

```r
adf.test(s1)
```

```
## Augmented Dickey-Fuller Test
## alternative: stationary
##
## Type 1: no drift no trend
##      lag    ADF p.value
## [1,]   0 -0.132   0.605
## [2,]   1  0.256   0.716
## [3,]   2  0.339   0.740
## [4,]   3  0.301   0.729
## [5,]   4  0.330   0.737
## Type 2: with drift no trend
##      lag   ADF p.value
## [1,]   0 -3.82   0.010
## [2,]   1 -2.66   0.089
## [3,]   2 -2.17   0.262
## [4,]   3 -2.30   0.214
## [5,]   4 -1.98   0.338
## Type 3: with drift and trend
##      lag   ADF p.value
## [1,]   0 -6.27  0.0100
## [2,]   1 -4.24  0.0100
## [3,]   2 -3.61  0.0354
## [4,]   3 -3.96  0.0138
## [5,]   4 -3.63  0.0332
## ----
## Note: in fact, p.value = 0.01 means p.value <= 0.01
```

```r
adf.test(s2)
```

```
## Augmented Dickey-Fuller Test
## alternative: stationary
##
## Type 1: no drift no trend
##      lag     ADF p.value
## [1,]   0 -0.0801   0.621
## [2,]   1  0.1562   0.689
## [3,]   2  0.3774   0.752
## [4,]   3  0.6151   0.821
## [5,]   4  0.6733   0.838
## Type 2: with drift no trend
##      lag   ADF p.value
## [1,]   0 -3.99  0.0100
```

```
## [2,]   1 -3.06  0.0327
## [3,]   2 -2.34  0.1940
## [4,]   3 -2.27  0.2209
## [5,]   4 -2.04  0.3109
## Type 3: with drift and trend
##      lag   ADF p.value
## [1,]   0 -7.78  0.0100
## [2,]   1 -6.15  0.0100
## [3,]   2 -4.58  0.0100
## [4,]   3 -4.01  0.0100
## [5,]   4 -3.64  0.0295
## ----
## Note: in fact, p.value = 0.01 means p.value <= 0.01
```

```
adf.test(s3)
```

```
## Augmented Dickey-Fuller Test
## alternative: stationary
##
## Type 1: no drift no trend
##      lag  ADF p.value
## [1,]   0 5.41    0.99
## [2,]   1 6.72    0.99
## [3,]   2 6.28    0.99
## [4,]   3 5.70    0.99
## Type 2: with drift no trend
##      lag  ADF p.value
## [1,]   0 3.25    0.99
## [2,]   1 4.54    0.99
## [3,]   2 4.76    0.99
## [4,]   3 4.75    0.99
## Type 3: with drift and trend
##      lag    ADF p.value
## [1,]   0 -0.271    0.99
## [2,]   1  0.257    0.99
## [3,]   2  0.397    0.99
## [4,]   3  0.481    0.99
## ----
## Note: in fact, p.value = 0.01 means p.value <= 0.01
```

```
adf.test(s4)
```

```
## Augmented Dickey-Fuller Test
## alternative: stationary
```

```
##
## Type 1: no drift no trend
##      lag  ADF p.value
## [1,]   0 3.03   0.990
## [2,]   1 2.08   0.990
## [3,]   2 1.62   0.972
## [4,]   3 3.25   0.990
## Type 2: with drift no trend
##      lag    ADF p.value
## [1,]   0  1.368   0.990
## [2,]   1  0.827   0.990
## [3,]   2  0.518   0.985
## [4,]   3 -0.113   0.941
## Type 3: with drift and trend
##      lag   ADF p.value
## [1,]   0 -2.98  0.1742
## [2,]   1 -3.60  0.0403
## [3,]   2 -4.61  0.0100
## [4,]   3 -1.81  0.6464
## ----
## Note: in fact, p.value = 0.01 means p.value <= 0.01
```

Au seul de 5%, on observe par le résultat du test de Dickey-Fuller augmenté que cette série n'est pas stationnaire.

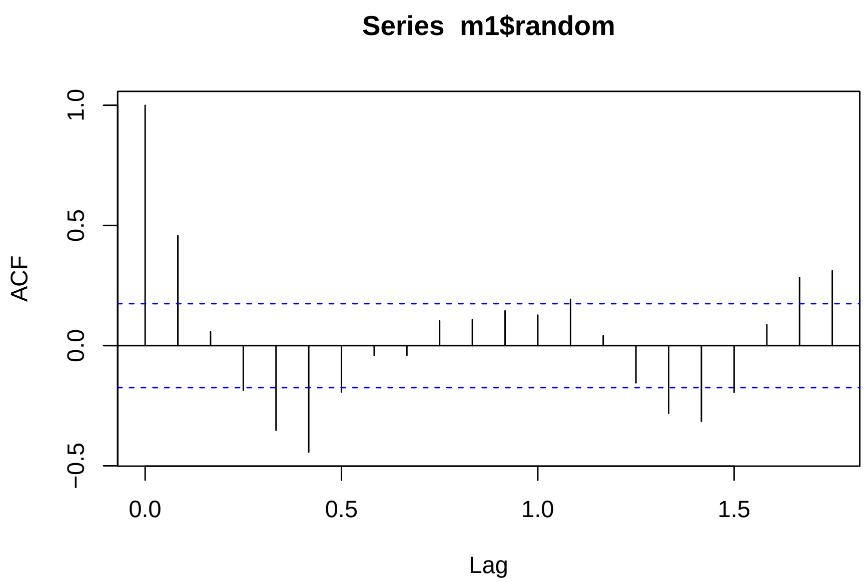
```r
library(caschrono)
```

```
## Loading required package: zoo
```

```
##
## Attaching package: 'zoo'
```

```
## The following objects are masked from 'package:base':
##
##     as.Date, as.Date.numeric
```

```r
library(zoo)

acf(m1$random, na.action = na.pass)
```



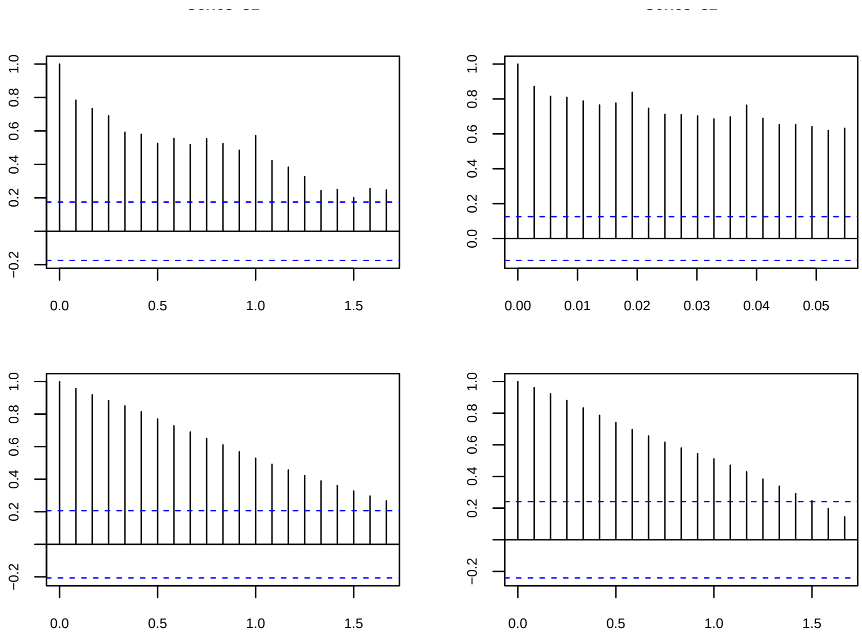

```
Box.test.2(m1$random, nlag = c(3, 6, 9, 12), type = "Ljung-Box", decim = 4)
```

```
##      Retard p-value
## [1,]      3       0
## [2,]      6       0
## [3,]      9       0
## [4,]     12       0
```

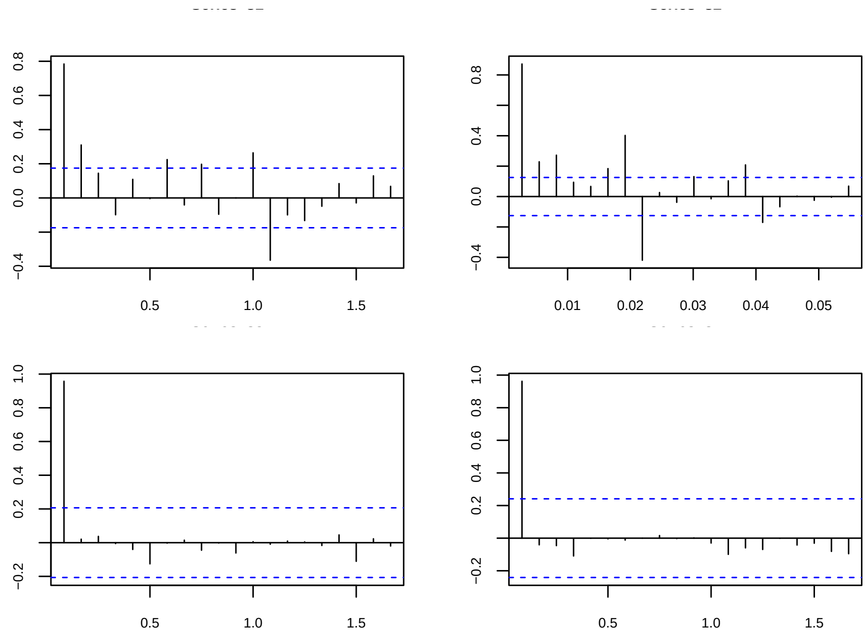
Du résultat ci-dessus, la p-value > 0,05 indiquant que la distribution des données n'est pas significativement différente de la distribution normale. En d'autres termes, nous pouvons supposer que la normalité.

fonction d'auto-corrélation

```
par(mfrow = c(2,2), oma = c(0, 0, 0, 0), mar = c(2.5, 2.5, 2,2), cex.main = 0.8, cex.la
acf(s1, lag.max = 20)
acf(s2, lag.max = 20)
acf(s3, lag.max = 20)
acf(s4, lag.max = 20)
```




```
par(mfrow = c(2,2), oma = c(0, 0, 0, 0), mar = c(2.5, 2.5, 2,2), cex.main = 0.8, cex.lab = 0.7, c
pacf(s1, lag.max = 20)
pacf(s2, lag.max = 20)
pacf(s3, lag.max = 20)
pacf(s4, lag.max = 20)
```



3.2 Modèle sans variable explicative

3.3 Modèle statique

3.4 Modèle dynamique

Chapter 4

Prédiction

You can add parts to organize one or more book chapters together. Parts can be inserted at the top of an .Rmd file, before the first-level chapter heading in that same file.

Add a numbered part: `# (PART) Act one {-}` (followed by `# A chapter`)

Add an unnumbered part: `# (PART\*) Act one {-}` (followed by `# A chapter`)

Add an appendix as a special kind of un-numbered part: `# (APPENDIX) Other stuff {-}` (followed by `# A chapter`). Chapters in an appendix are prepended with letters instead of numbers.

4.1 Partant d'un modèle

4.2 Sans reference à un modèle

Chapter 5

Footnotes and citations

5.1 Footnotes

Footnotes are put inside the square brackets after a caret `^[]`. Like this one [1].

5.2 Citations

Reference items in your bibliography file(s) using `@key`.

For example, we are using the **bookdown** package [Xie, 2023] (check out the last code chunk in index.Rmd to see how this citation key was added) in this sample book, which was built on top of R Markdown and **knitr** [Xie, 2015] (this citation was added manually in an external file book.bib). Note that the `.bib` files need to be listed in the index.Rmd with the YAML `bibliography` key.

The RStudio Visual Markdown Editor can also make it easier to insert citations: https://rstudio.github.io/visual-markdown-editing/#/citations

[1] This is a footnote.

Chapter 6

Blocks

6.1 Equations

Here is an equation.

$$f\left(k\right) = \binom{n}{k} p^k \left(1-p\right)^{n-k} \tag{6.1}$$

You may refer to using `\@ref(eq:binom)`, like see Equation (6.1).

6.2 Theorems and proofs

Labeled theorems can be referenced in text using `\@ref(thm:tri)`, for example, check out this smart theorem 6.1.

Theorem 6.1. *For a right triangle, if c denotes the* length *of the hypotenuse and a and b denote the lengths of the **other** two sides, we have*

$$a^2 + b^2 = c^2$$

Read more here https://bookdown.org/yihui/bookdown/markdown-extensions-by-bookdown.html.

6.3 Callout blocks

The R Markdown Cookbook provides more help on how to use custom blocks to design your own callouts: https://bookdown.org/yihui/rmarkdown-cookbook/custom-blocks.html

Chapter 7

Sharing your book

7.1 Publishing

HTML books can be published online, see: https://bookdown.org/yihui/bookdown/publishing.html

7.2 404 pages

By default, users will be directed to a 404 page if they try to access a webpage that cannot be found. If you'd like to customize your 404 page instead of using the default, you may add either a `_404.Rmd` or `_404.md` file to your project root and use code and/or Markdown syntax.

7.3 Metadata for sharing

Bookdown HTML books will provide HTML metadata for social sharing on platforms like Twitter, Facebook, and LinkedIn, using information you provide in the `index.Rmd` YAML. To setup, set the `url` for your book and the path to your `cover-image` file. Your book's `title` and `description` are also used.

This `gitbook` uses the same social sharing data across all chapters in your book-all links shared will look the same.

Specify your book's source repository on GitHub using the `edit` key under the configuration options in the `_output.yml` file, which allows users to suggest an edit by linking to a chapter's source file.

Read more about the features of this output format here:

https://pkgs.rstudio.com/bookdown/reference/gitbook.html

Or use:

```
?bookdown::gitbook
```

Bibliography

Yihui Xie. *Dynamic Documents with R and knitr*. Chapman and Hall/CRC, Boca Raton, Florida, 2nd edition, 2015. URL http://yihui.org/knitr/. ISBN 978-1498716963.

Yihui Xie. *bookdown: Authoring Books and Technical Documents with R Markdown*, 2023. URL https://github.com/rstudio/bookdown. R package version 0.36.