SÉRIES CHRONOLOGIQUES

Karine Chalvet-Monfray

VetAgro Sup

Université de Lyon

ENSV– 4 février2015

Objectifs pédagogiques

- Comprendre les particularités des données chronologiques
- Savoir réaliser des analyses simples sur des données chronologiques
- Connaître les intérêts et limites de ses analyses.

Plan

- Introduction
- Etude préliminaire
- Estimation de la tendance
- Estimation de la variation saisonnière
- Conclusion

Plan

- Introduction
- Etude préliminaire
- Estimation de la tendance
- Estimation de la variation saisonnière
- Conclusion

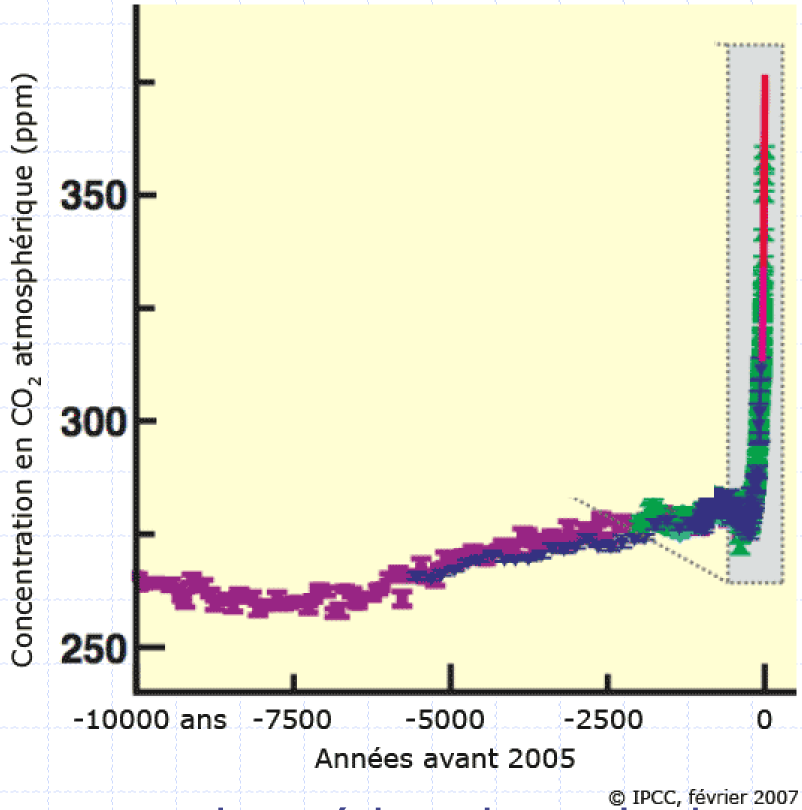
Introduction

- Définition :
  - D'une manière générale :
    - Données mesurées à des intervalles de temps réguliers (ex. en économie (indices boursiers), en agriculture (production laitière),…)
  - En épidémiologie :
    - Données de morbidités, de mortalités,… à des intervalles de temps réguliers, jours, semaine, mois, … -> séries quotidiennes, hebdomadaires, mensuelles, …

Introduction

- Intérêts :
  - Comprendre le passé :
    - Analyser les données et estimer les tendances
  - Prédire l’avenir :
    - Nombreuses méthodes empiriques ou non -> modèle déterministe

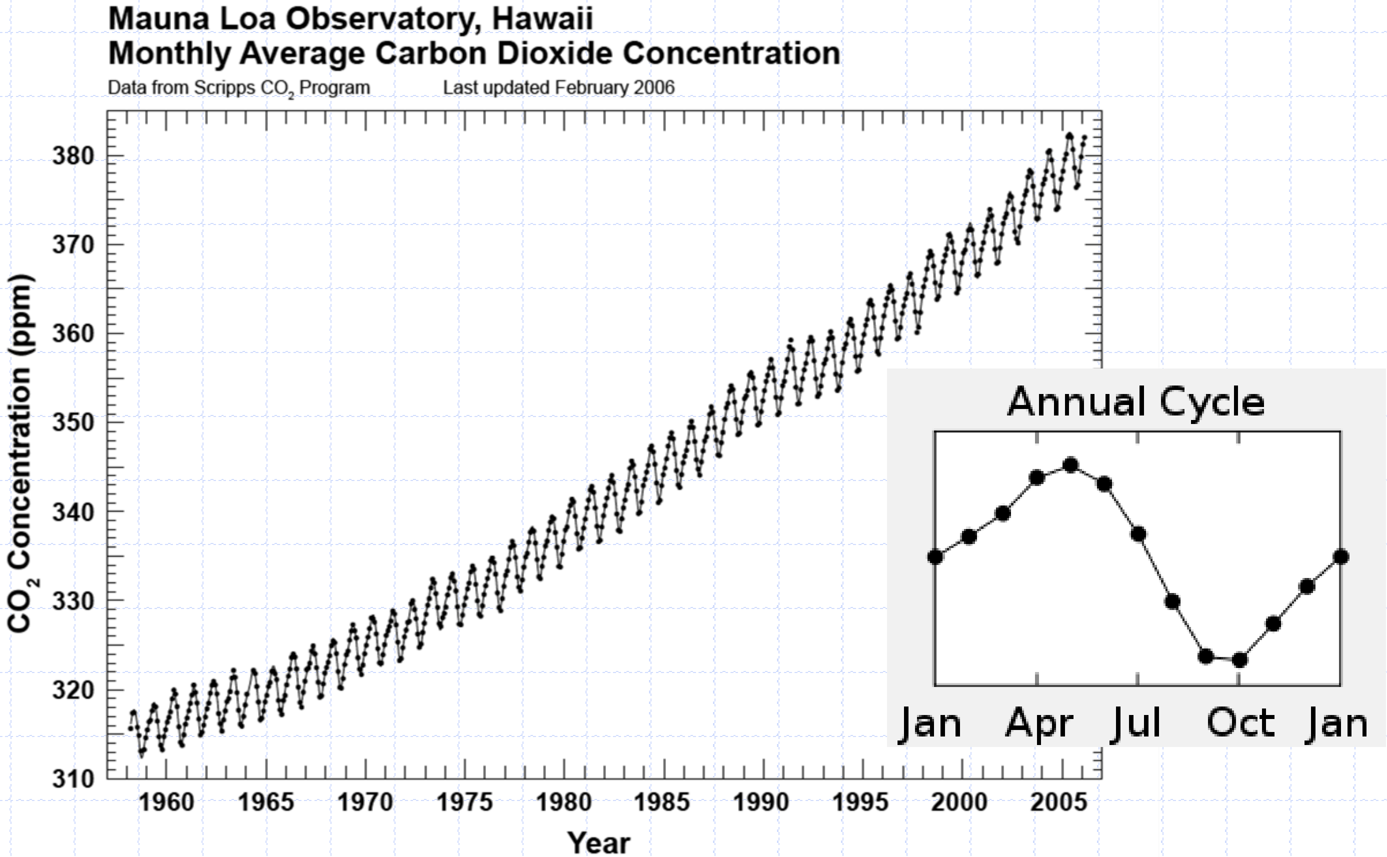
Introduction
Evolution du CO_2 atmosphérique



Ici nous n'avons pas que des séries chronologiques. Pourquoi ?…

… pas d'intervalles réguliers

Séries chronologiques
Evolution du CO_2 atmosphérique



On n'observe pas la même chose selon l'échelle de temps où l'on se situe.

Qu'observez- vous ?

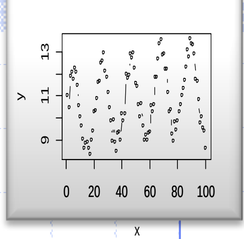
Evolution de l’espérance de vie à la naissance en France de 1740 à 2005

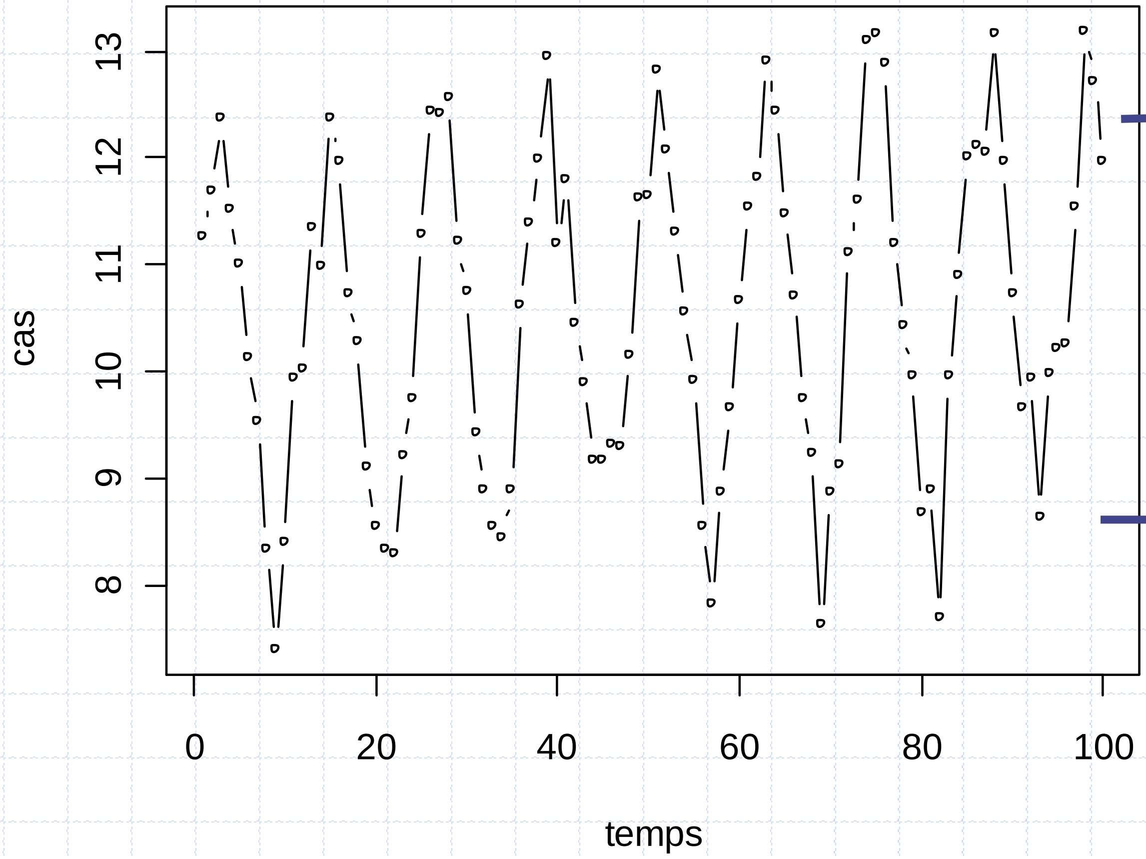
- Qu’observez-vous ?

Plan

- Introduction
- Etude préliminaire
- Estimation de la tendance
- Estimation de la variation saisonnière
- Conclusion

Etude préliminaire
Représentation graphique

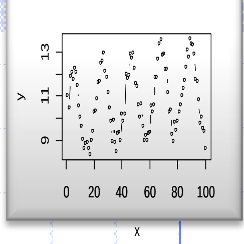




Vue d'ensemble
- Tendance
- Variation périodique

Valeur surprenante
- Erreur de saisie
- Phénomène exceptionnel

Etude préliminaire
Test d'une valeur surprenante

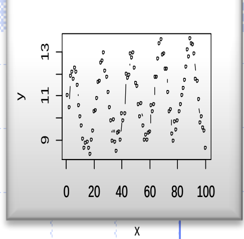


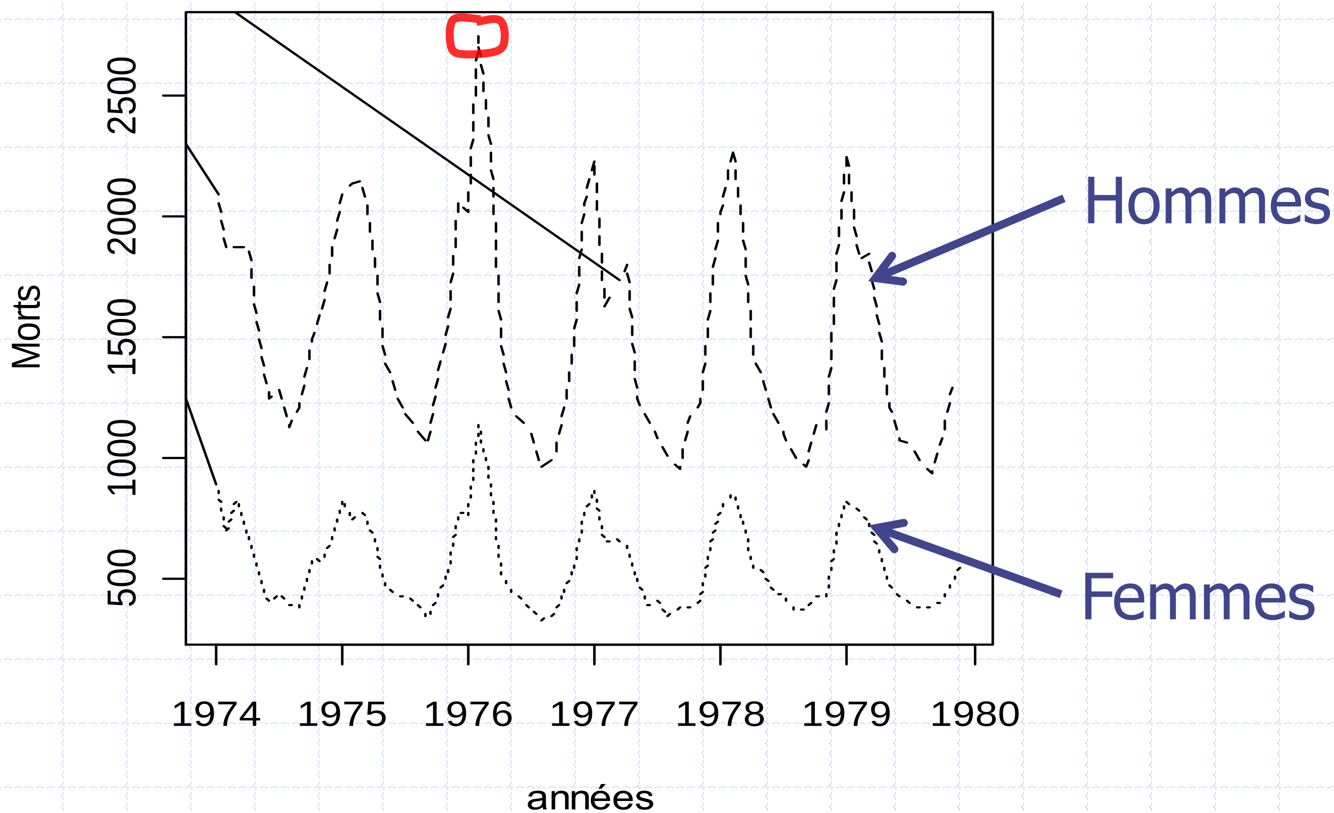
- Pour un problème de santé assez rare survenant dans une vaste population, on admet que des variations au cours du temps suit une loi de poisson.
- On peut estimer l'intervalle de confiance autours de la valeur observée. Si la valeur attendue n'est pas comprise dans l'intervalle de confiance à 95%, il y a alors une différence significative entre les deux.

Application

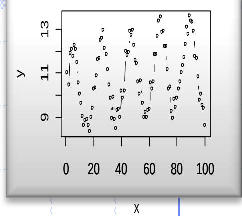
Etude préliminaire

Décès par bronchite et asthme au Royaume Uni entre 1974 et 1979





Application



Etude préliminaire

Test d'une valeur surprenante

Avec R

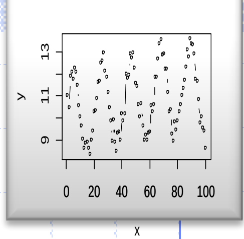
Nom de la fonction qui charge des package spécifique- A n'utiliser que si nécessaire

Nom du « package » qui offre des fonctions spécifiques à l'épidémiologie

Nombre de cas observés
Ici 2750

```
library(epitools)
pois.exact(2750)
```

Nom de la fonction qui estime les intervalles de confiance pour une loi de poisson



Etude préliminaire

Décès par bronchite et asthme au Royaume Uni entre 1974 et 1979

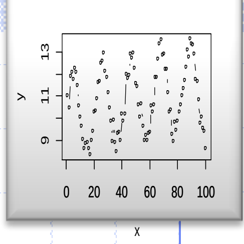
```
> library(epitools)
> pois.exact(2750)
     x pt rate    lower    upper conf.level
1 2750  1 2750 2648.169 2854.744       0.95
```

IC=[2648.2; 2854.7]

Or moyenne est de la mortalité mensuelle chez les hommes est de 1496.

Attention, en concluant à la différence significative, les données ne sont pas indépendantes. ->Autocorrélation

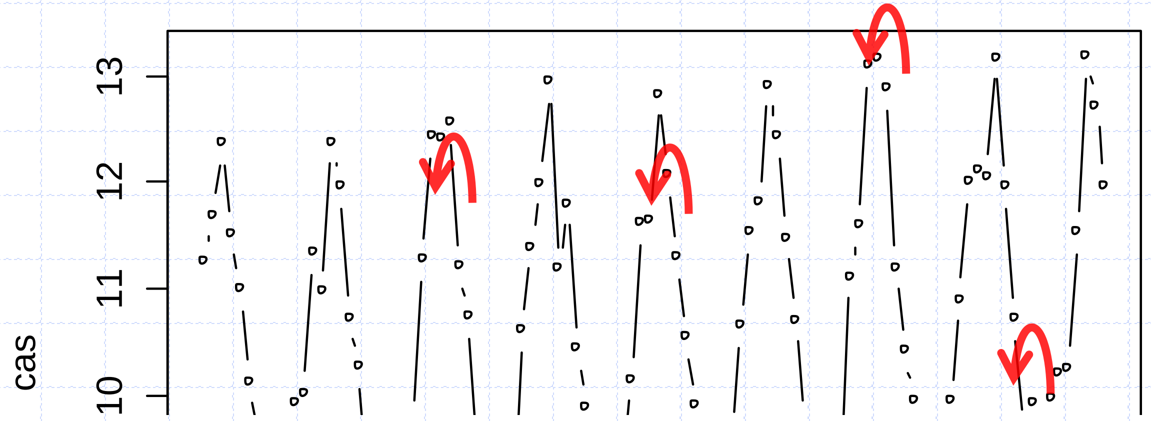
Etude préliminaire



Autocorrélation

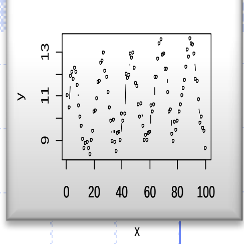
Autocorrélation : corrélation entre les valeurs successives.
Ex. Autocorrélation d'ordre k, c'est la corrélation entre les valeur de X_t et X_{t+k}.

=> Ex. Conséquence dans les régressions car les données ne sont pas indépendantes



Etude préliminaire

Autocorrélation



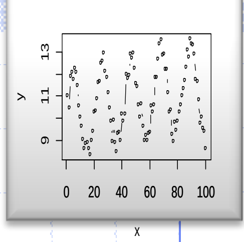
Rappel de la corrélation : coefficient de corrélation r

Covariance de xy

$$r = \frac{\sigma_{xy}}{\sigma_x \sigma_y} = \frac{\frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})^2}\sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \bar{y})^2}}$$

Écart-type de x

Écart-type de y



Etude préliminaire

Autocorrélation

C'est simplement la corrélation du processus par rapport à une version décalée dans le temps de lui-même.

autocovariance d'ordre k

Coefficient d'autocorrélation d'ordre k

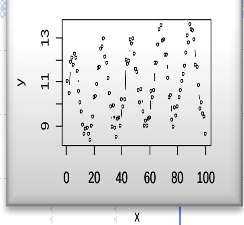
$$r_k = \frac{C_k}{C_0} = \frac{\frac{1}{N-k}\sum_{i=1}^{N-k}\left(x_i - \bar{x}\right)\left(x_{i+k} - \bar{x}\right)}{\frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})^2}$$

autocovariance d'ordre 0
=variance de x

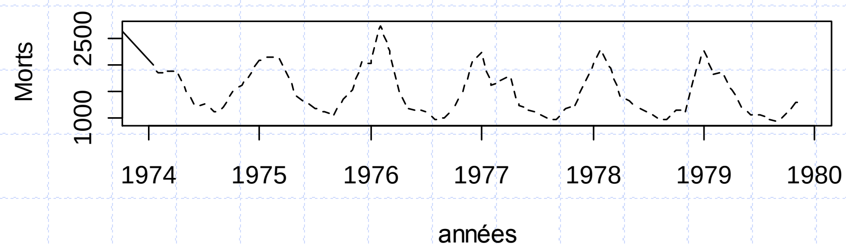
Application

Etude préliminaire

Autocorrélation



- Pour les données sur les décès par bronchite et asthme au Royaume Uni entre 1974 et 1979 chez les hommes.



- Avec R

Colonne correspondant à la série : colonne « mort » du jeu de donnée « d ».

```
acf(d$mort)
```

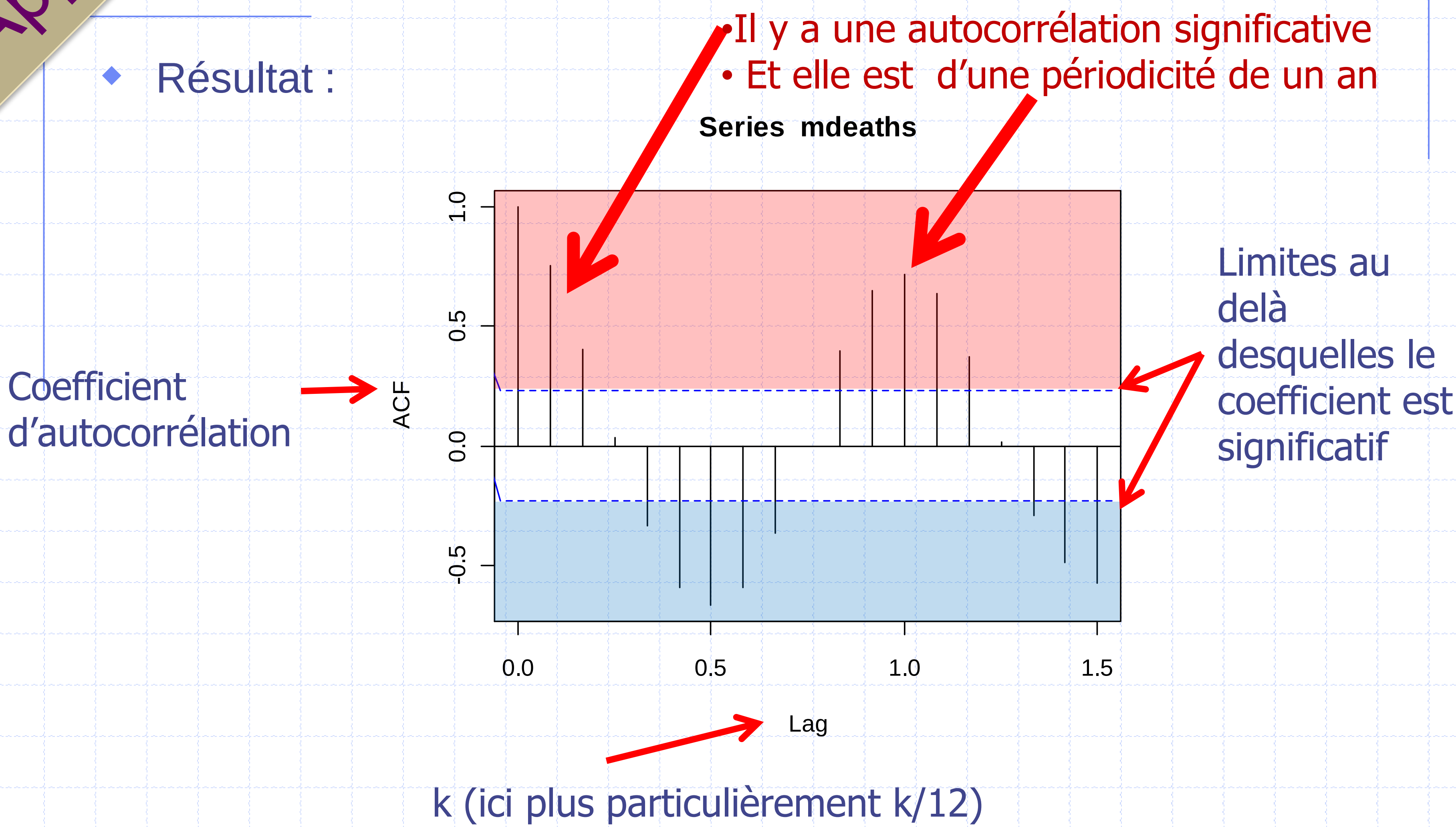
Nom de la fonction qui calcule les estimations des autocorélations et des autocovariances pour différentes valeurs de k

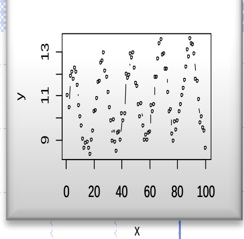
Application

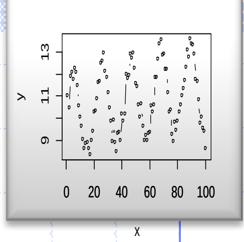
Etude préliminaire

Autocorrélation

- Résultat :







Application

Etude préliminaire

Autocorrélation

- Résultat complémentaire :

```
acf(d$mort)$acf
```

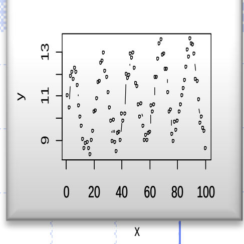
Nom de la valeur qui conserve les estimations des autocorélations

```
 [1,]  1.000000000
 [2,]  0.757059088
 [3,]  0.402052538
 [4,]  0.037835423
 [5,] -0.334653832
 [6,] -0.588929104
 [7,] -0.663957769
 [8,] -0.591538927
 [9,] -0.364128377
[10,] -0.002048197
[11,]  0.395954646
[12,]  0.650274675
[13,]  0.717108918
[14,]  0.637540770
[15,]  0.373228840
[16,]  0.013931354
[17,] -0.288680880
[18,] -0.488659376
[19,] -0.572098334
```

- Il y a une autocorrélation significative
- Et elle est d'une périodicité de un an

- -> Saisonnalité

Exercice 1



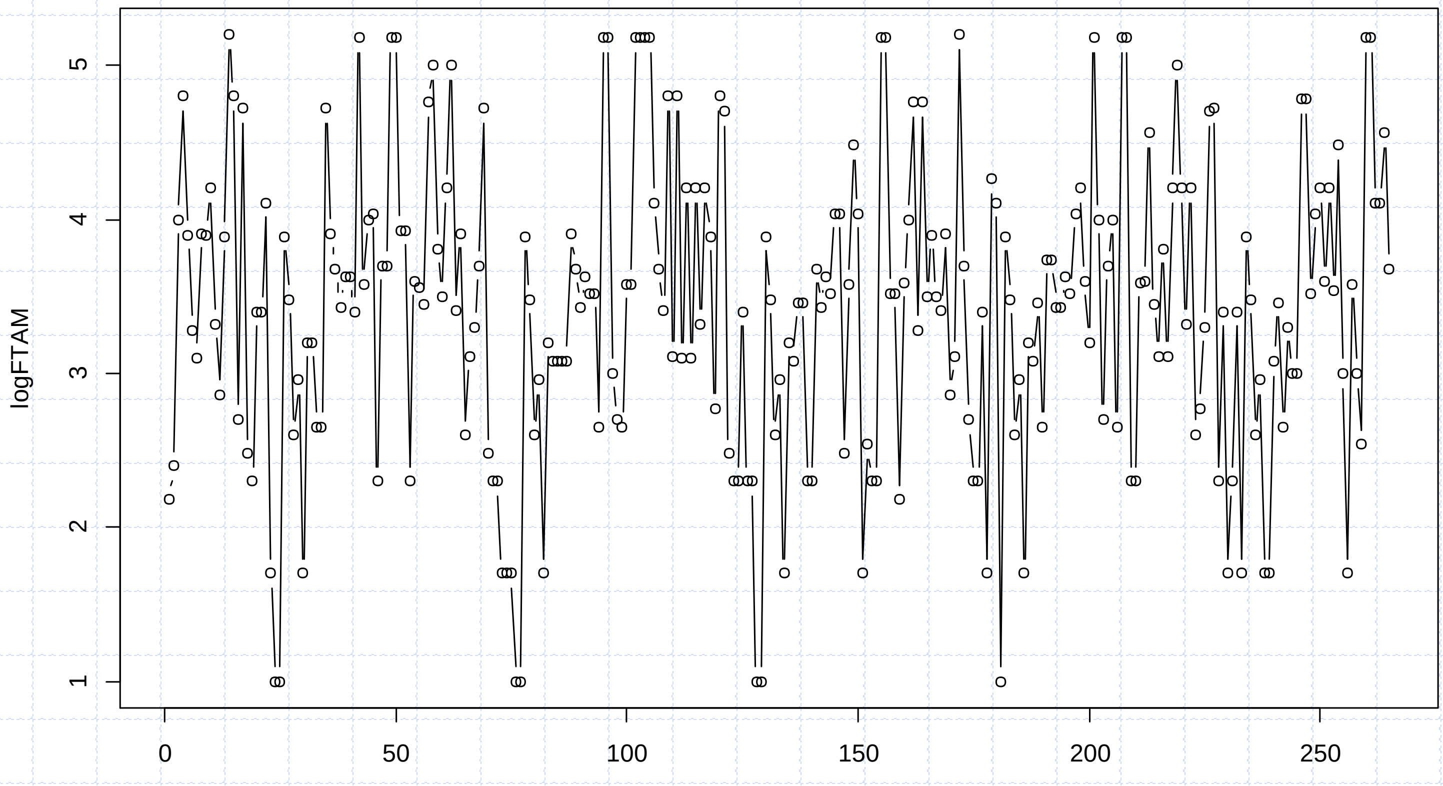
Suivi du Log(FTAM) pendant 5 ans avec une mesure hebdomadaire

Fichier Classeur2.txt

Représentez les données et décrivez ce que vous observez, étudiez son autocorrélation.

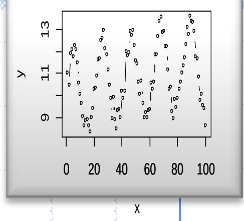
Exercice 1

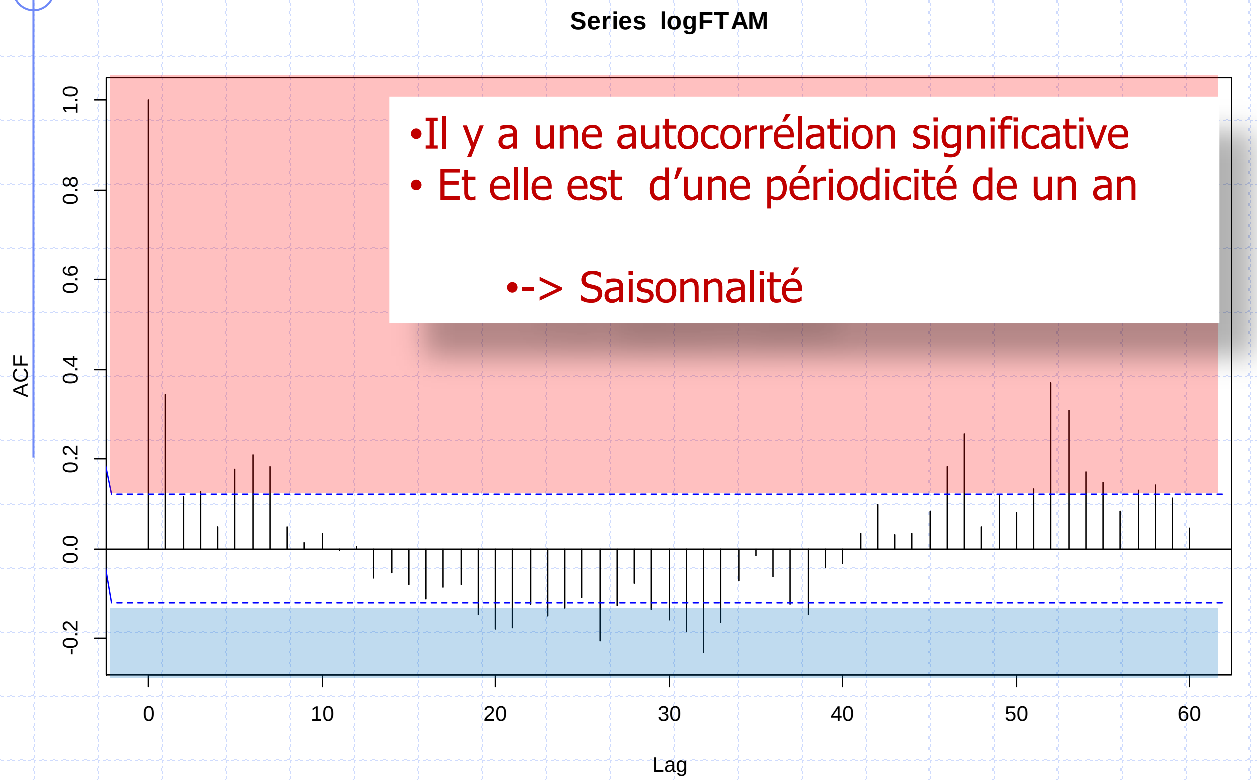
Suivi du Log(FTAM) pendant 5 ans avec une mesure hebdomadaire



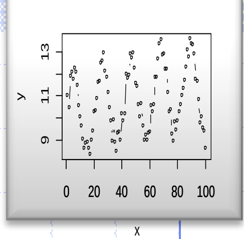
Représentez les données et décrivez ce que vous observez, étudiez son autocorrélation.

Suivi du Log(FTAM) pendant 5 ans avec une mesure hebdomadaire



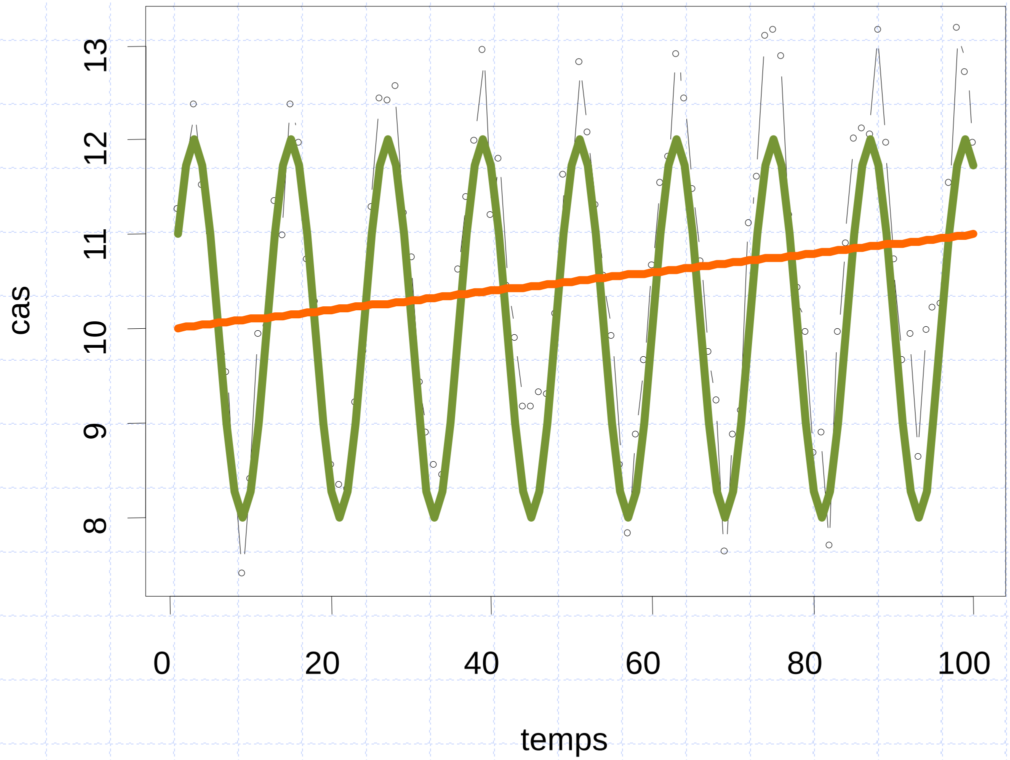


Représentez les données et décrivez ce que vous observez, étudiez son autocorrélation.

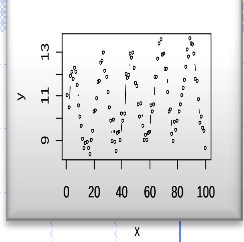


Etude préliminaire

Composantes d'une série chronologique



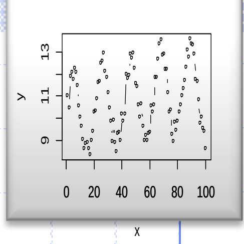
Tendance
Variation périodique (saisonnière ou cyclique)
Variations irrégulières ou accidentelles



Etude préliminaire

Composantes d'une série chronologique

- La tendance T :
Évolution « moyenne » à long terme
- La composante cyclique C :
Comportement périodique dont la période est suffisamment importante pour être qualifiée de mouvement à moyen ou à long terme (ex. cycle pluriannuelle…)
- La composante saisonnière S :
Composante périodique dépendante de l'environnement (nature, économique, sociale,…)
- La composante irrégulière I :
Ce qui n'est pas pris en compte par T, C et S



Etude préliminaire

Composantes d’une série chronologique

Cas des modèles d’ajustement

1. Le modèle additif

$$x = TC + S + I$$

Si l’amplitude des variation périodique semble à peu près constante-> additif

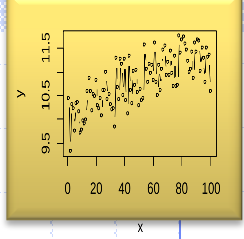
2. Le modèle multiplicatif

$$x = TC.S.\ I$$

Si l’amplitude des variation périodique semble augmenter ou diminuer -> multiplicatif

Plan

- Introduction
- Etude préliminaire
- Estimation de la tendance
- Estimation de la variation saisonnière
- Conclusion

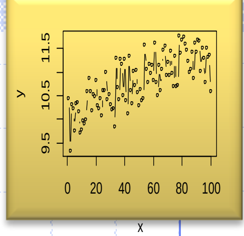


Estimation de la tendance

Coefficient de corrélation :

- Un coefficient de corrélation entre le nombre de cas et le temps peut tester une tendance.
- Le test de coefficient de corrélation de Spearman réalisé sur les rang ne nécessite pas d'hypothèse préalable.
- Attention, il peut ne pas détecter de tendance si la séquence est trop courte et le phénomène cyclique.

Application



Estimation de la tendance

Coefficient de corrélation de Spearman

Avec R

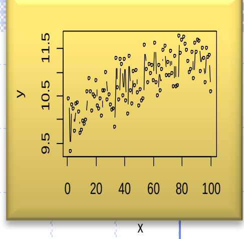
Nom de la fonction qui test d'association entre deux séries appariées

Colonne correspondant à la série : colonne « mort » du jeu de donnée « d ».

```
cor.test(d$mort,c(1:n),method='spearman')
```

Vecteur représentant le temps et n étant le nombre total de mesure

Argument spécifiant la méthode utilisée



Application

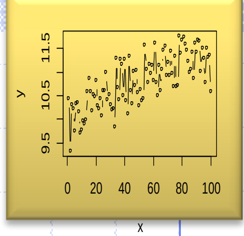
Estimation de la tendance

Coefficient de corrélation de Spearman

```
cor.test(d$mort,c(1:72),method='spearman')
Spearman's rank correlation rho

data:  mdeaths and c(1:72)
S = 82662.66, p-value = 0.004766
alternative hypothesis: true rho is not equal
to 0
sample estimates:
       rho
-0.3290671
```

P<0.05 => corrélation significative

Corrélation négative
Diminution avec le temps



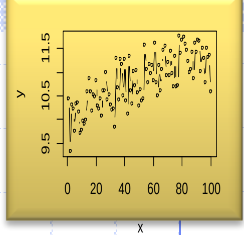
Estimation de la tendance

Méthode de la moyenne mobile :

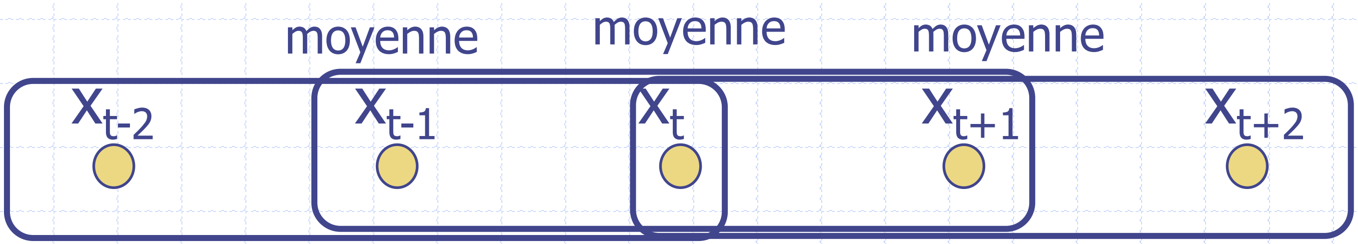
- Elle permet de lisser la série chronologique.
- Elle atténue les fluctuations saisonnières et irrégulières.
- Le principe est de remplacer les valeurs de la série par une moyenne mobile sur p valeurs consécutives.

Estimation de la tendance

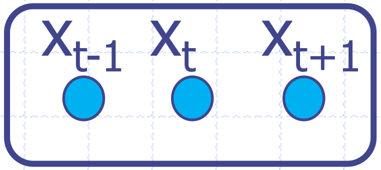
Méthode de la moyenne mobile



Principe :

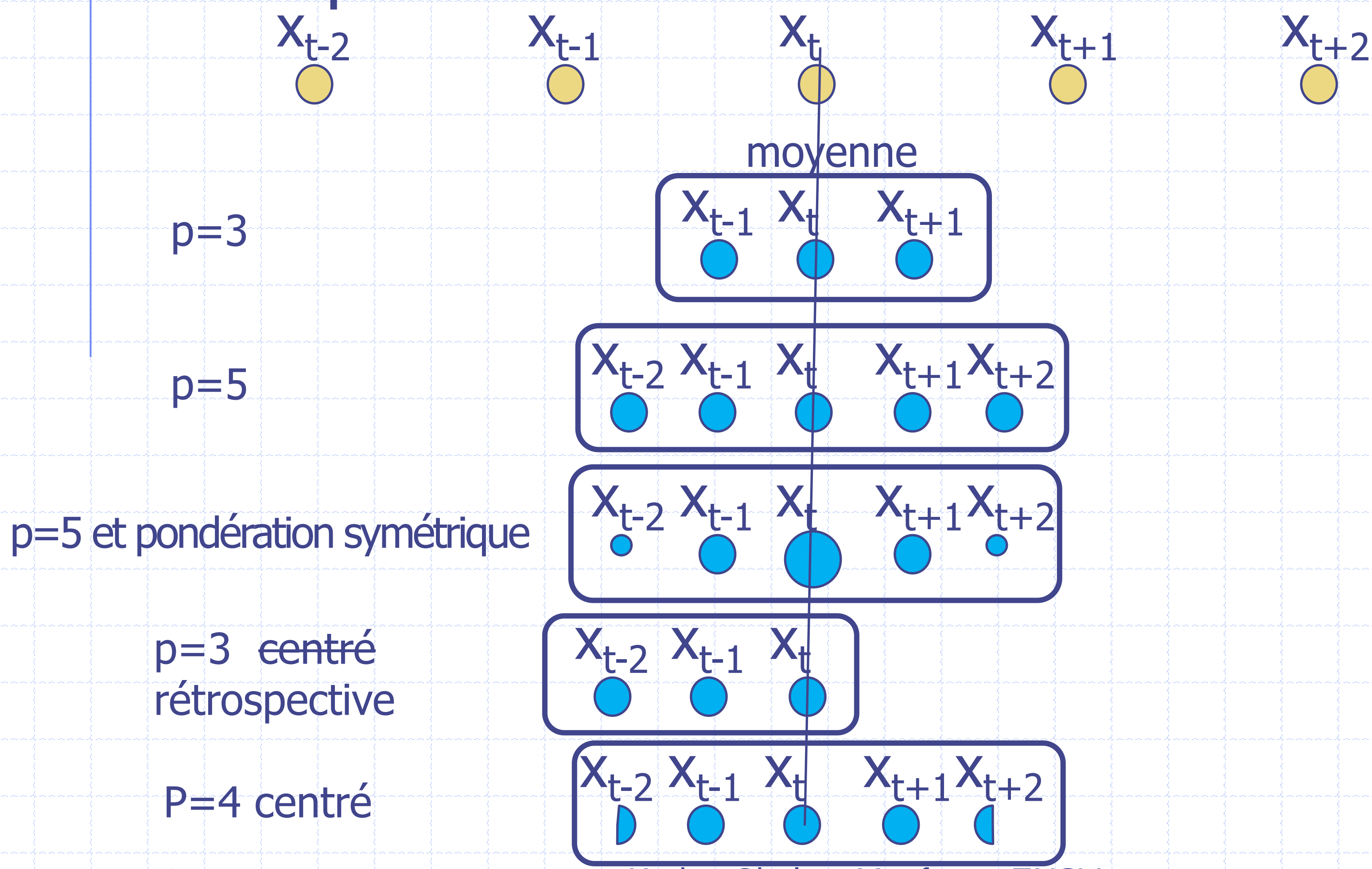


Moyenne mobile d'ordre 3 (p=3)
Centrée symétrique uniforme

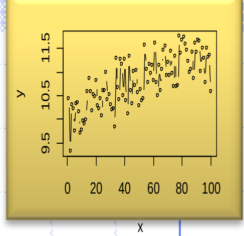


Estimation de la tendance
Méthode de la moyenne mobile

Principe :



Estimation de la tendance
Méthode de la moyenne mobile



Formule :

Centré symétrique uniforme

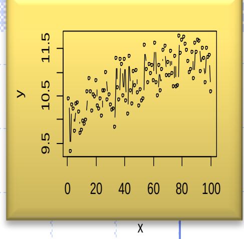
p impaire (p=2k+1):

$$m_t = \frac{1}{2k+1}\sum_{j=-k}^{k} x_{t+j}$$

p paire (p=2k):

$$m_t = \frac{1}{2k}\left(\sum_{j=-k+1}^{k-1} x_{t+j} + \frac{1}{2}\left(x_{t-k} + x_{t+k}\right)\right)$$

Application



Estimation de la tendance

Méthode de la moyenne mobile

Avec R, cas d'une moyenne centrée uniforme d'ordre impair

Nom de la fonction qui calcule la moyenne glissante

Colonne correspondant à la série : colonne « mort » du jeu de donnée « d ».

```
filter(d$mort,1/p*rep(1,p),side=2)
```

Vecteur représentant la pondération (ici elle est uniforme) et p étant l'ordre et impair

Argument spécifiant si la moyenne est centrée (side =2) ou rétrospective (side=1)

Application

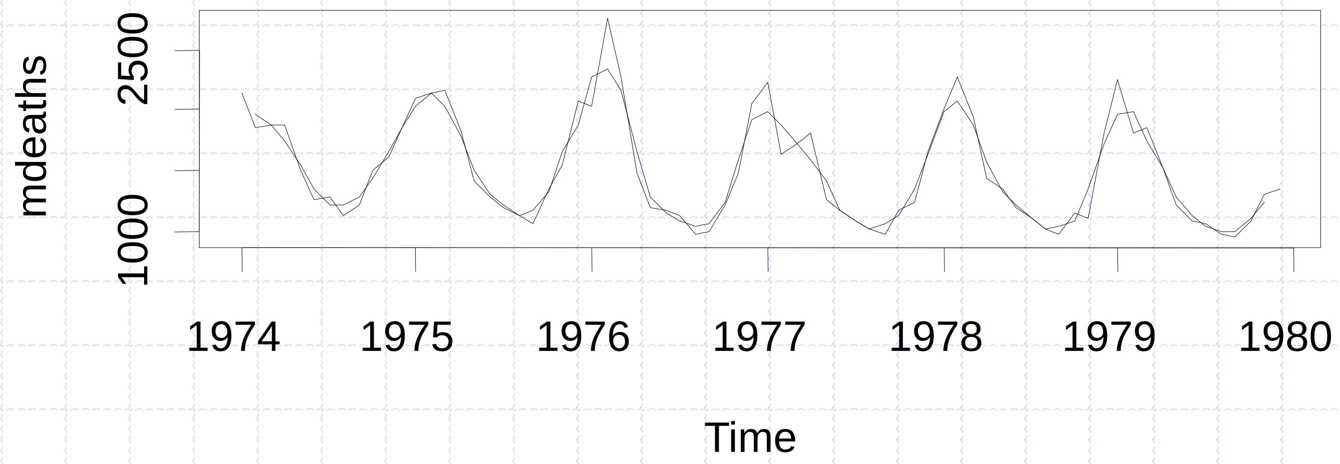
Estimation de la tendance

- Méthode de la moyenne mobile

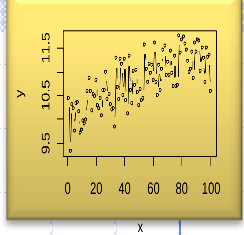
```
plot(d$mort)
lines(filter(d$mort,1/3*rep(1,3)))
```

Fonction qui rajoute une ligne à un plot

Moyenne sur 3 mois



Application



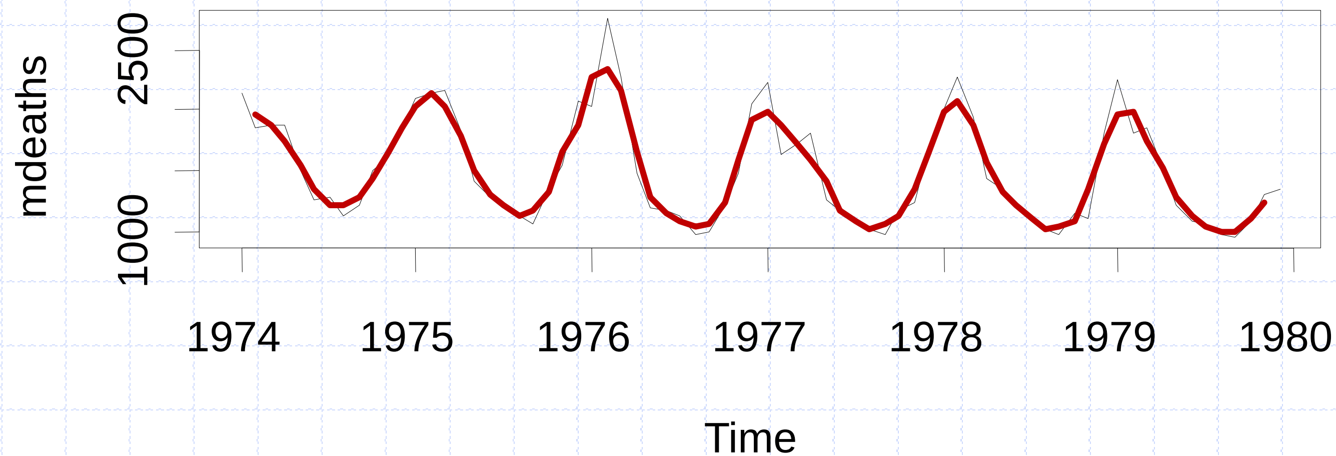
Estimation de la tendance

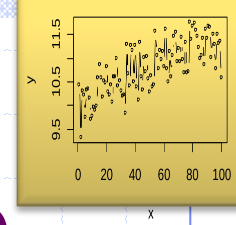
- Méthode de la moyenne mobile

```
plot(d$mort)
lines(filter(d$mort,1/3*rep(1,3)))
```

Fonction qui rajoute une ligne à un plot

Moyenne sur 3 mois





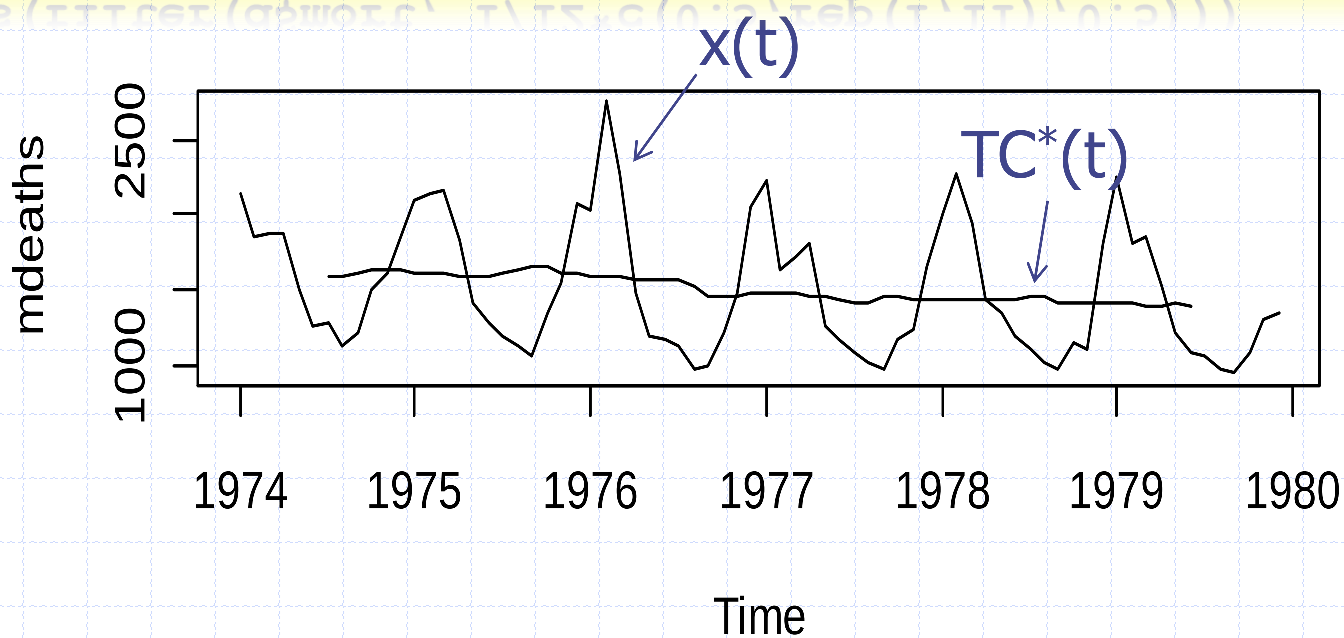
Application

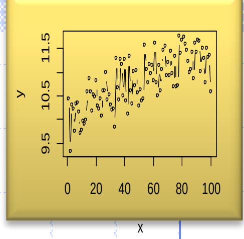
Estimation de la tendance

- Méthode de la moyenne mobile

```
plot(d$mort)
lines(filter(d$mort, 1/12*c(0.5,rep(1,11),0.5)))
```

Moyenne sur 1 an

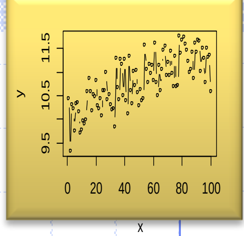




Estimation de la tendance

Méthode de Régression :

- Elle fait partie des modèles mathématiques
- Il en existe plusieurs formes de régression selon la relation entre le temps et le paramètre observé.
- Un modèle linéaire ne peut être appliqué que sur une courte période mais supérieur à celle des fluctuations saisonnières -> résultats aberrants.



Estimation de la tendance

Méthode de Régression :

$$TC(t) = f(t) + \varepsilon(t)$$

- Linéaire :

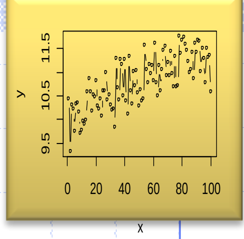
$$TC(t) = a + bt + \varepsilon(t)$$

- Polynomial :

$$TC(t) = a + bt + bt^2 + ... + \varepsilon(t)$$

- Exponentiel :

$$TC(t) = ab^t + \varepsilon(t)$$



Estimation de la tendance

Méthode régression

Avec R, cas d'une régression linéaire simple

Nom de la fonction qui estime les paramètres d'une régression linéaire

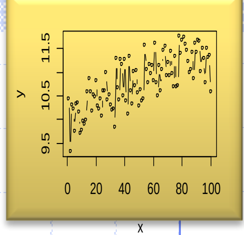
Vecteur représentant le temps et n étant le nombre d'observation

```
time<-c(1:n)
lm(d$mort ~time)
```

Colonne correspondant à la série : colonne « mort » du jeu de donnée « d ».

Signe utilisé dans les modèles pour indiquer « en fonction de »

Application



Estimation de la tendance

Méthode régression

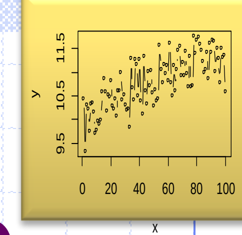
Avec R , cas d'une régression linéaire simple

```
Time<-c(1:n)
lm(d$mort ~time)

Call:
lm(formula = mdeaths ~ time)

Coefficients:
(Intercept)          time
   1709.152        -5.841
```

Pente =b

Ordonnées à l'origine =a

Application



Estimation de la tendance
Méthode régression

```
summary(lm(d$mort ~time))
Call:
lm(formula = d$mort ~ time)
Residuals:
   Min      1Q Median      3Q     Max
-552.2 -340.4 -136.6  261.9 1192.7
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 1709.152     99.678  17.147   <2e-16 ***
time          -5.841      2.373  -2.461   0.0163 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0
1
Residual standard error: 418.5 on 70 degrees of freedom
Multiple R-squared: 0.07966,    Adjusted R-squared: 0.06651
F-statistic: 6.058 on 1 and 70 DF,  p-value: 0.01631
```

p value pour le test de la pente

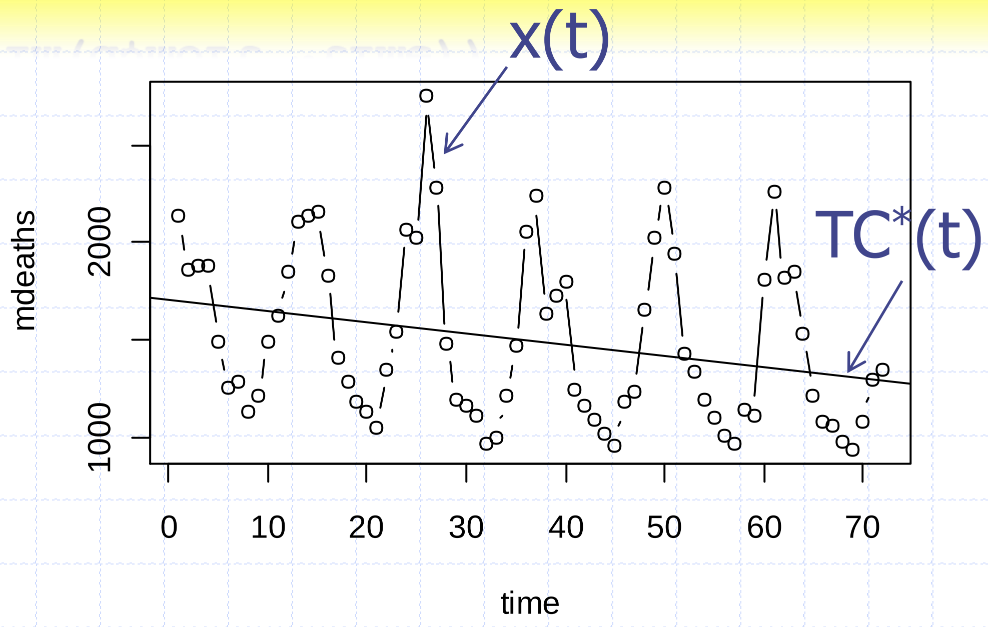
Application

Estimation de la tendance

Méthode régression

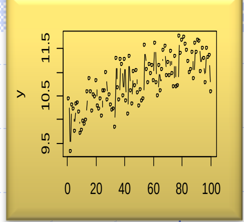
Avec R, cas d'une régression linéaire simple

```
plot(d$mort)
abline(lm(d$mort ~time))
```

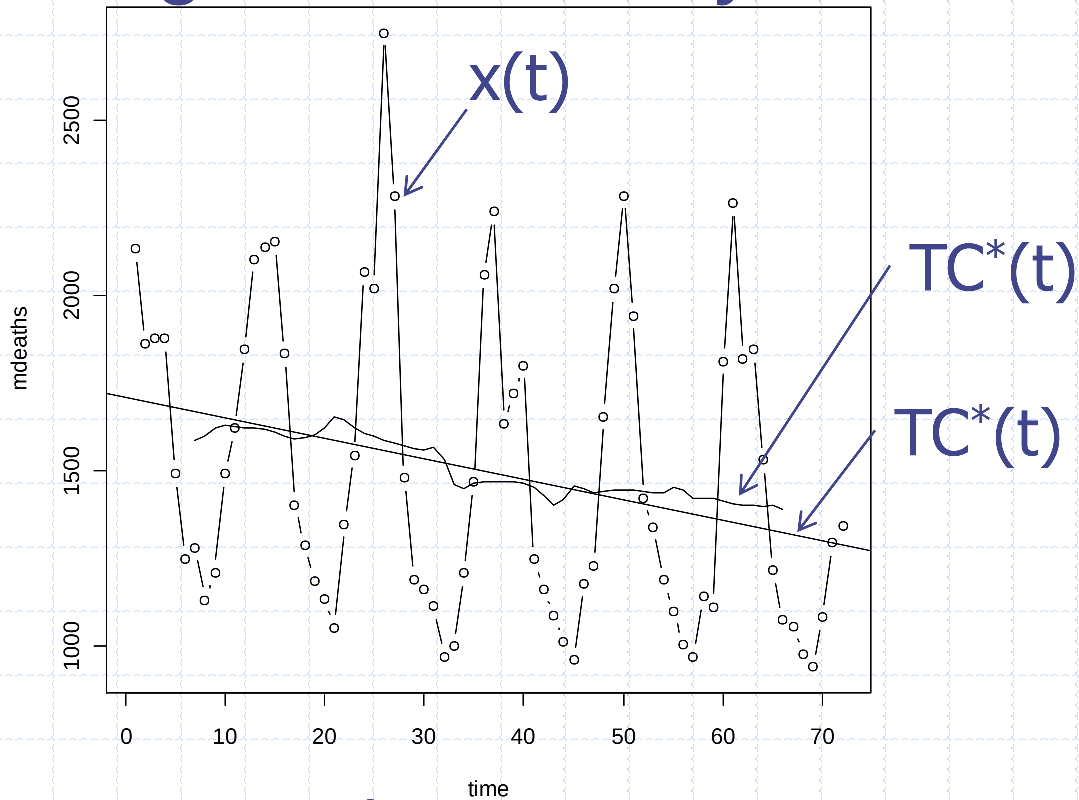


Application

Estimation de la tendance



Méthode régression et moyenne mobile

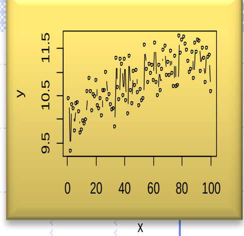


Estimation différente

On peut aussi faire une régression sur la moyenne mobile ayant gommé la variation saisonnière.

Exercice 2

Suivi du Log(FTAM) pendant 5 ans avec une mesure hebdomadaire

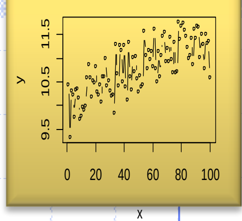


En reprenant les données de l'exo 1, estimer la tendance :

1. par un test de corrélation de Spearman
2. par une moyenne mobile d'une période de 52 semaine)
3. par une régression linéaire du temps

Puis représenter graphiquement la série chronologique avec la moyenne mobile et la droite de régression.

Exercice 2



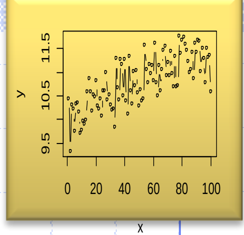
Suivi du Log(FTAM) pendant 5 ans avec une mesure hebdomadaire

En reprenant les données de l’exo 1, estimer la tendance :

1. par un test de corrélation de Spearman

```
cor.test(temps,logFTAM,method='spearman')

         Spearman's rank correlation rho

data:  temps and logFTAM
S = 3033016, p-value = 0.7203
alternative hypothesis: true rho is not equal to 0
sample estimates:
       rho
0.02209978
```

Exercice 2

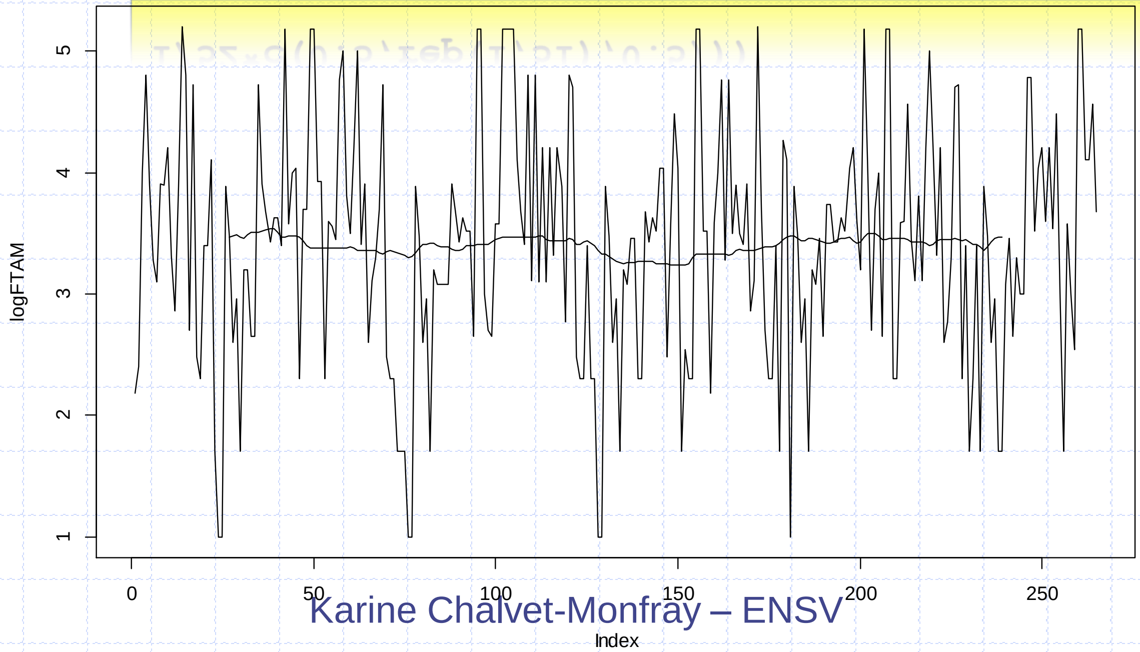


Suivi du Log(FTAM) pendant 5 ans avec une mesure hebdomadaire

En reprenant les données de l'exo 1, estimer la tendance :

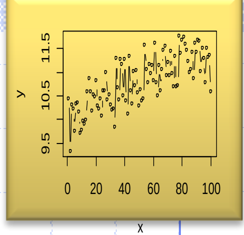
2. par une moyenne mobile d'une période de 12 mois)

```
plot(logFTAM,type="l")
lines(filter(logFTAM,
1/52*c(0.5,rep(1,51),0.5)))
```



Exercice 2

Suivi du Log(FTAM) pendant 5 ans avec une mesure hebdomadaire

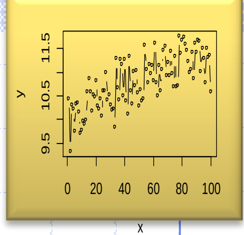


En reprenant les données de l'exo 1, estimer la tendance :

- par une régression linéaire du temps

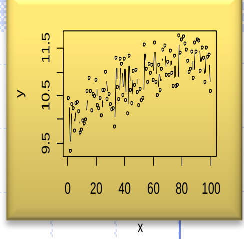
```
 > lm(logFTAM~temps)

Call:
lm(formula = logFTAM ~ temps)

Coefficients:
(Intercept)        temps
  3.3600755    0.0003813
```

Exercice 2



Suivi du Log(FTAM) pendant 5 ans avec une mesure hebdomadaire

```
> summary(lm(logFTAM~temps))

Call:
lm(formula = logFTAM ~ temps)

Residuals:
     Min       1Q   Median       3Q      Max 
-2.42910 -0.69745  0.05543  0.57853  1.83459 

Coefficients:
             Estimate Std. Error t value Pr(>|t|)    
(Intercept) 3.3600755  0.1203710  27.914   <2e-16 ***
temps       0.0003813  0.0007845   0.486    0.627    
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1 

Residual standard error: 0.977 on 263 degrees of freedom
Multiple R-squared: 0.0008975,          Adjusted R-squared: -0.002901 
F-statistic: 0.2363 on 1 and 263 DF,  p-value: 0.6273
```

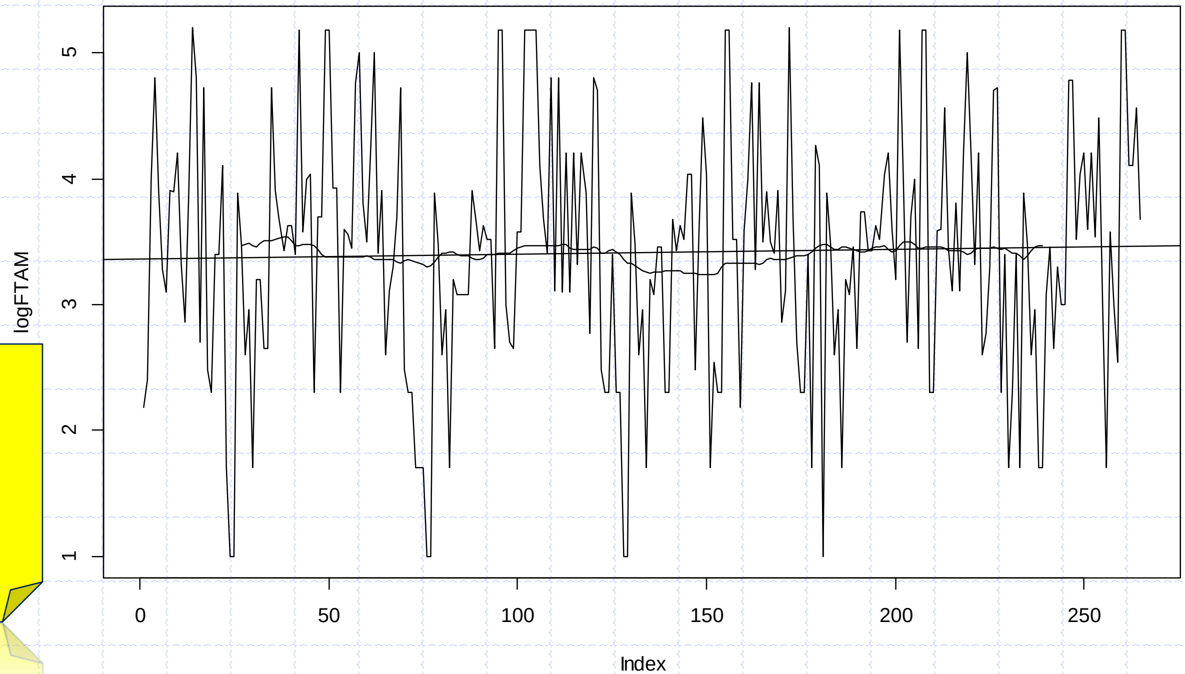


Exercice 2

Suivi du Log(FTAM) pendant 5 ans avec une mesure hebdomadaire

Puis représenter graphiquement la série chronologique avec la moyenne mobile et la droite de régression.

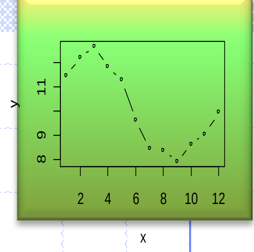
```
plot(logFTAM,type="l")
lines(filter(logFTAM,
1/52*c(0.5,rep(1,51),0.5)))
abline(lm(logFTAM~temps))
lm(logFTAM~temps)
```



Plan

- Introduction
- Etude préliminaire
- Estimation de la tendance
- Estimation de la variation saisonnière
- Conclusion

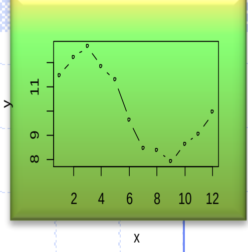
Estimation de la variation saisonnière



Il existe classiquement deux méthodes :

- Effet saisonnier -> effet moyen par mois (coefficient saisonnier) qui va être vu.
- Régression saisonnière -> analyse spectral (complexe) qui ne sera pas vu.

Estimation de la variation saisonnière

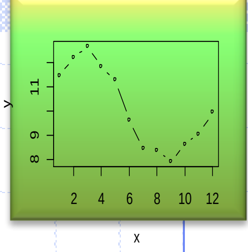


Par la suite, la tendance et la saisonnalité ayant été estimé il est possible :

-> analyse des résidus

Par ailleurs, il existe des outils simples qui permettent en utilisant la méthode des moyennes mobiles la décomposition d’une série chronologique

Estimation de la variation saisonnière



Méthode des coefficients saisonniers

Principe :

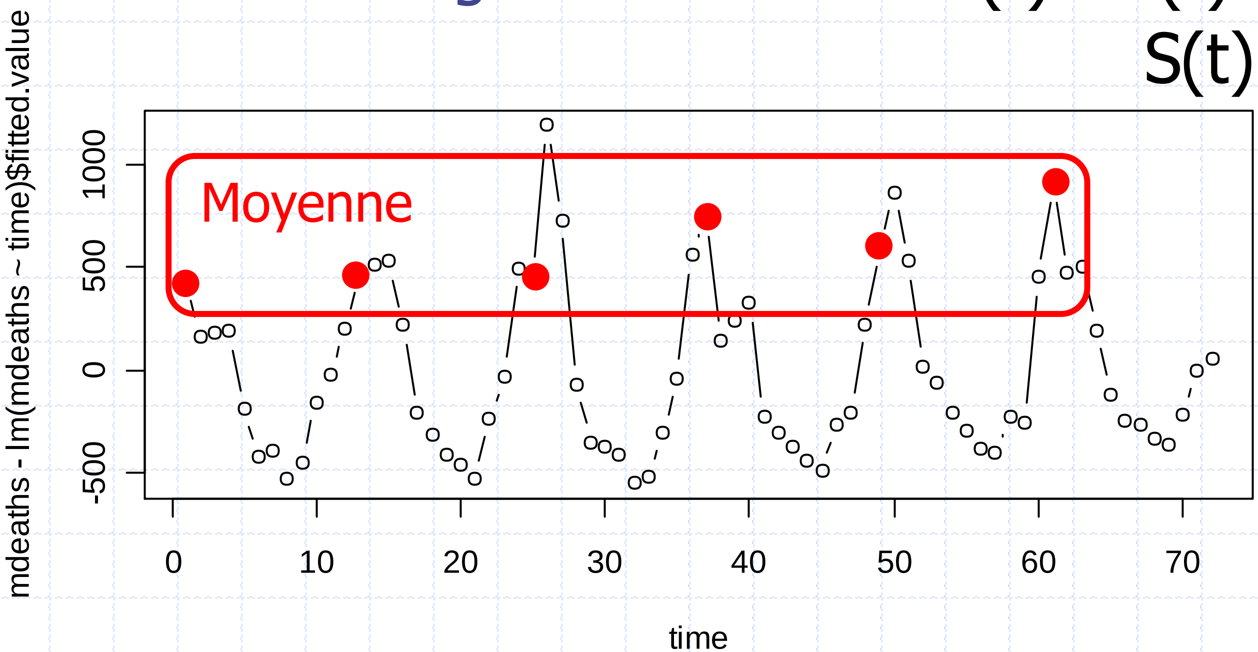
A partir d'une série chronologie stabilisée (la composante tendance à été enlevée), une valeur moyenne est calculée pour les 12 mois de l'année.

Estimation de la variation saisonnière

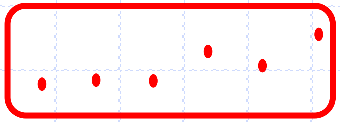
Méthode des coefficients saisonniers

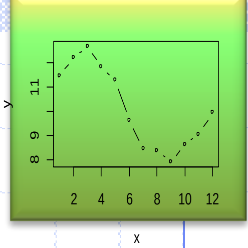
série chronologie **stabilisée** $x'(t) = x(t) - TC^*(t)$

$S(t) + I(t)$



Moyenne pour le mois de janvier
= coefficient saisonnier de janvier





Estimation de la variation saisonnière

Méthode des coefficients saisonniers

$$x(t) = TC(t) + S(t) + I(t)$$

Pour le mois i de l'année j

$$x_{ij} = TC_{ij} + S_i + I_{ij}$$

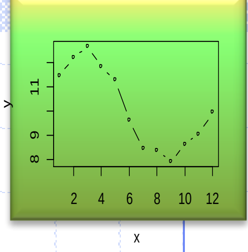
$$TC_{ij} -> \text{estimation } TC^*_{ij}$$ Par moyenne mobile ou régression

$$S_i = \frac{1}{m} \sum_{\text{années}} \underbrace{\left(x_{ij} - TC^*_{ij} \right)}_{\text{série chronologie stabilisée}}$$

m est le nombre d'années de suivi

Application

Estimation de la variation saisonnière



Méthode des coefficients saisonniers

Avec R ,

Nom de la valeur qui conserve les valeurs simulées (a+bt)

$$x(t) - TC^*(t) = x'(t)$$

$$x'_{ij}$$

```
stab<-d$mort-lm(d$mort ~time)$fitted.value
nstab<-ts(stab, frequency = 12, start = c(1974, 1))
```

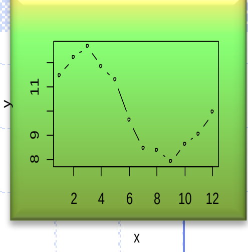
Nom de la fonction qui crée des objet de séries temporelles (chronologiques)

Données mensuelles

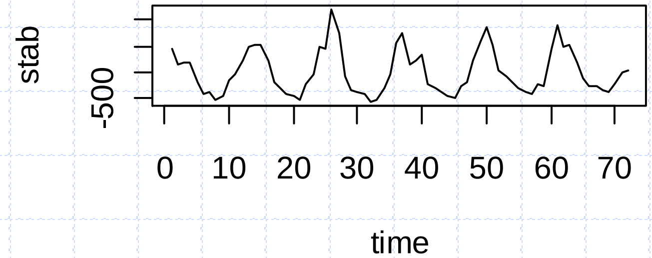
Date de début : janvier 1974

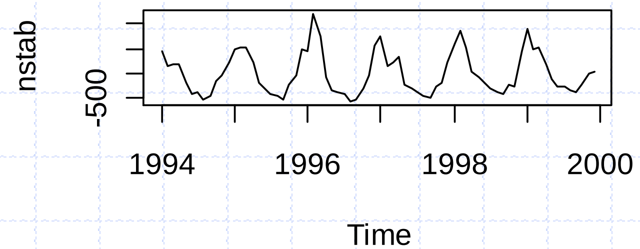
Application

Estimation de la variation saisonnière



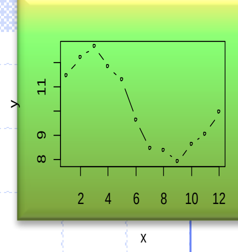

```
> stab
          1            2            3            4            5 
 64.14795918 -24.94344333  -6.03484585 -32.12624837 -42.21765089 
          6            7            8            9           10 
-47.30905341  -5.40045593 -37.49185845  -8.58326096 -16.67466348 
         11           12           13           14           15 
-18.76606600  15.14253148  78.05112896  -1.04027356  19.86832393 
         16           17           18           19           20 
 15.77692141   9.68551889   0.59411637  16.50271385  23.41131133 
         21           22           23           24           25 
 -3.68009119   5.22850630 -17.86289622  73.04570126  34.95429874 
         26           27           28           29           30 
  2.86289622   4.77149370 -23.31990881 -30.41131133 -13.50271385 
         31           32           33           34           35 
 10.40588363   5.31448111 -38.77692141 -51.86832393  -8.95972644 
         36           37           38           39           40 
 81.94887104  86.85746852   1.76606600  -0.32533652  10.58326096 
         41           42           43           44           45 
-10.50814155 -34.59954407 -42.69094659 -32.78234911 -35.87375163 
         46           47           48 
 -7.96515415  -0.05655667  32.85204082 
> nstab
             Jan          Feb          Mar          Apr          May
1994  64.14795918 -24.94344333  -6.03484585 -32.12624837 -42.21765089
1995  78.05112896  -1.04027356  19.86832393  15.77692141   9.68551889
1996  34.95429874   2.86289622   4.77149370 -23.31990881 -30.41131133
1997  86.85746852   1.76606600  -0.32533652  10.58326096 -10.50814155
             Jun          Jul          Aug          Sep          Oct
1994 -47.30905341  -5.40045593 -37.49185845  -8.58326096 -16.67466348
1995   0.59411637  16.50271385  23.41131133  -3.68009119   5.22850630
1996 -13.50271385  10.40588363   5.31448111 -38.77692141 -51.86832393
1997 -34.59954407 -42.69094659 -32.78234911 -35.87375163  -7.96515415
             Nov          Dec
1994 -18.76606600  15.14253148
1995 -17.86289622  73.04570126
1996  -8.95972644  81.94887104
1997  -0.05655667  32.85204082
```





Application

Estimation de la variation saisonnière



Méthode des coefficients saisonniers

Avec R ,

Nom de la fonction qui crée un vecteur

Paramètre permettant d'indiquer la longueur du vecteur. Ici =12

```
sais<-vector(length=12)
for (i in 1:12) mean(window(nstab,  start= c(1974,i ), end=
c(1979,4), frequency =1))->sais[i]
```

Nom de la fonction qui calcule les moyennes

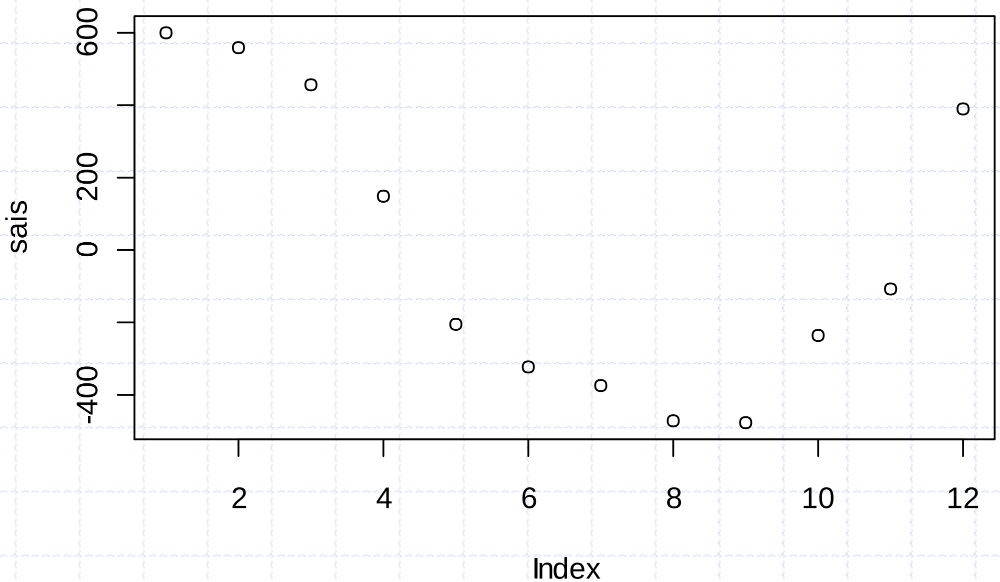
Nom de la fonction qui extrait des données d'une série temporelle

$$S_i = \frac{1}{m} \sum_{\text{années}} \left(x_{ij} - TC^*_{ij} \right)$$

Application

Estimation de la variation saisonnière

```
> sais
 [1]  601.7617  559.1030  454.1110  146.7856 -205.7543 -324.7129
 [7] -375.2716 -472.0303 -478.3890 -237.3477 -110.3064  388.7349
```



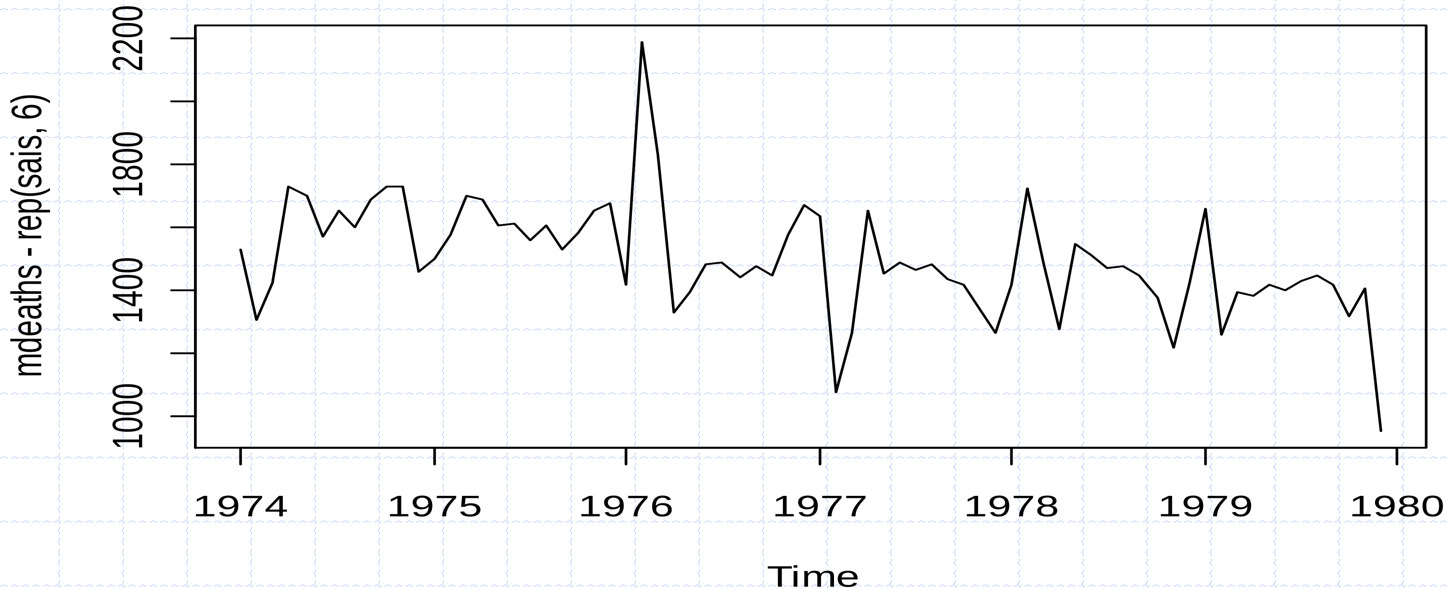
Estimation de la variation saisonnière

série chronologie **désaisonnalisée**

```
plot(d$mort-rep(sais,6))
```

$$x''(t) = x(t) - S^*(t)$$

$$TC(t) + I(t)$$



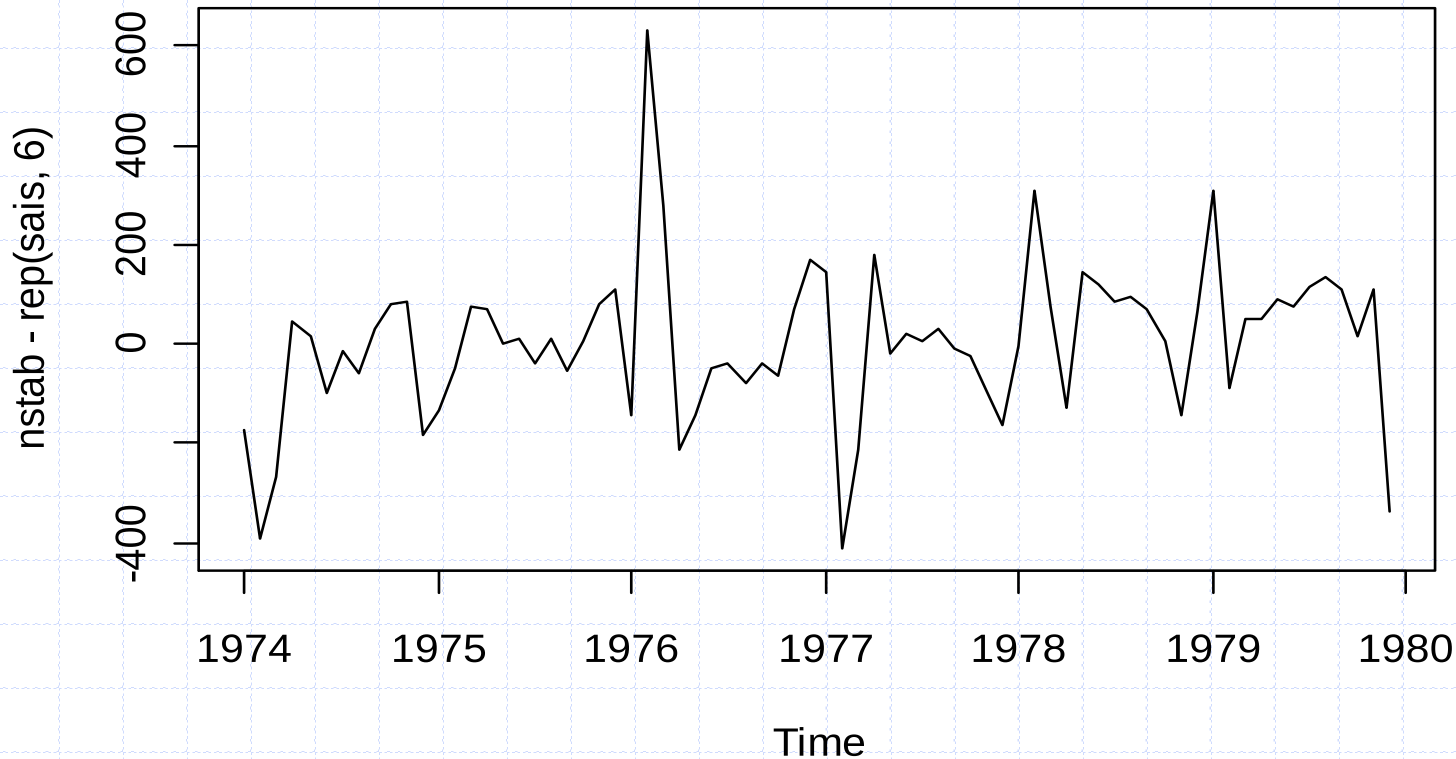
Estimation de la variation saisonnière

Composante irrégulière

```
plot(nstab-rep(sais,6))
```

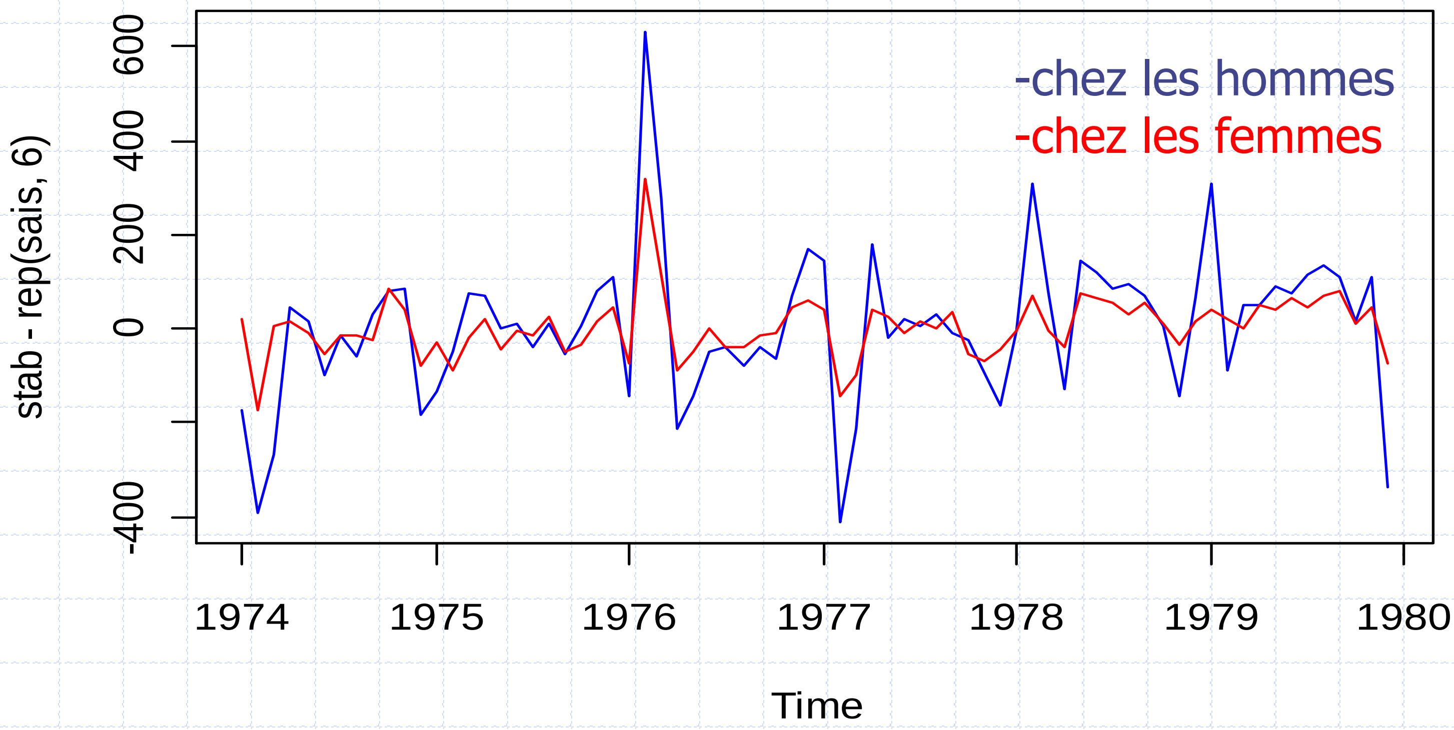
$$x'''(t) = x(t) - S^{*}(t) - TC^{*}(t)$$

$$I(t)$$

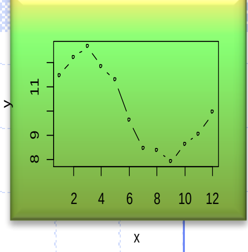


Estimation de la variation saisonnière

Composante irrégulière

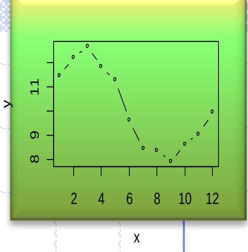


Estimation de la variation saisonnière



Composante irrégulière

Attention !!!
La composante irrégulière peut être structurée…
Analyse de la composante irrégulière est possible…



Décomposition d'une série chronologique

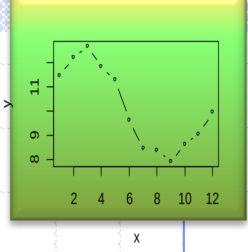
Elle consiste à identifier les différentes composantes d'une série chronologique :

$$x(t) = TC(t) + S(t) + I(t)$$

Plusieurs méthodes existent dont celle par la moyenne mobile.

Application

Décomposition d’une série chronologique



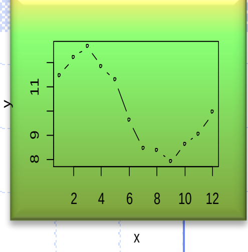

```
nmort<-ts(d$mort, frequency = 12, start = c(1974, 1))
m<-decompose(nmort)
```

nom de la fonction qui décompose une série chronologique en 3 termes (Trend=TC, Seasonal=S, random=I)

série chronologique (temporal serie)

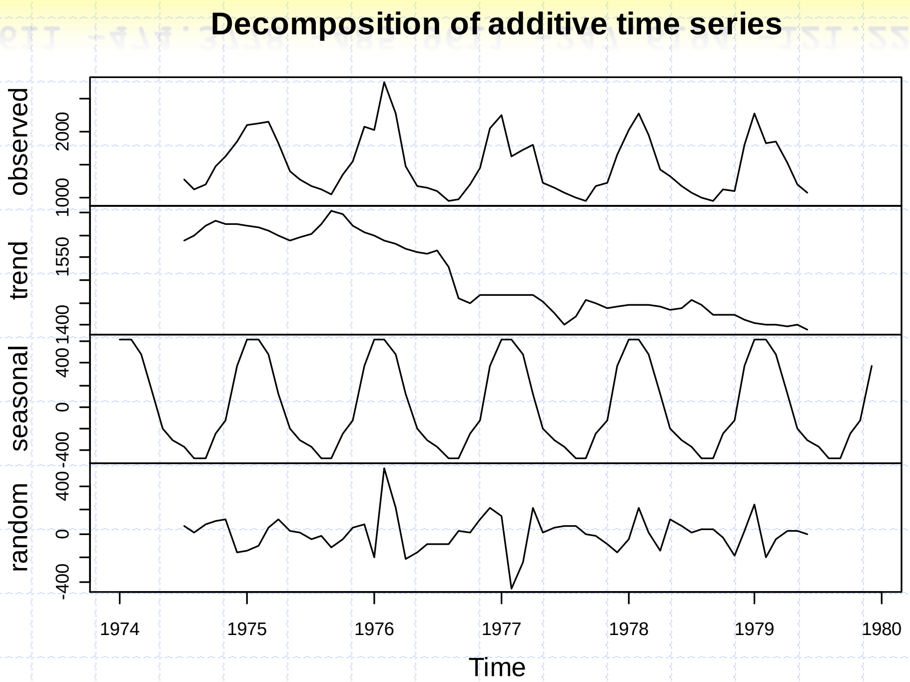
Application

Décomposition d’une série chronologique



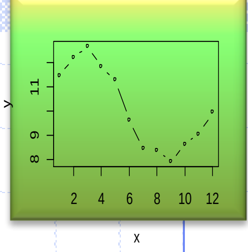

```
plot(m)
m$figure
 [1] 620.4472  619.6139  487.3556  117.2222 -211.4278 -308.6944
 [7] -371.0611 -474.3778 -485.9611 -247.6194 -121.2278  375.7306
```

valeur des S_i



Exercice 3

Suivi du Log(FTAM) pendant 5 ans avec une mesure hebdomadaire



En reprenant les données de l'exo 1 :

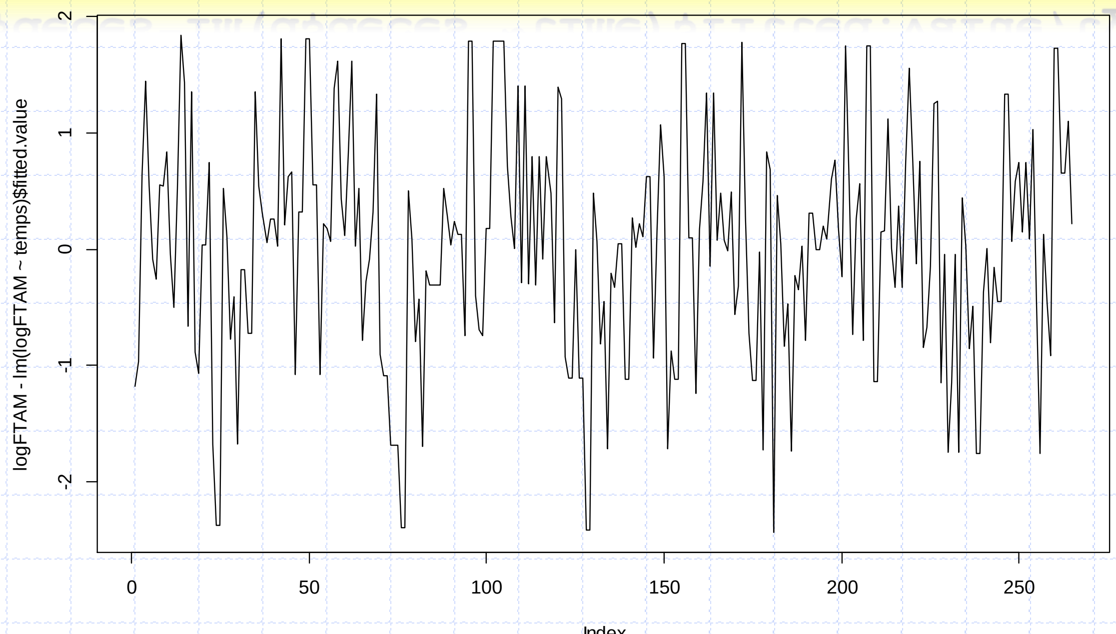
1. Représenter la série chronologique stationnarisée
2. Estimer et représenter la variation saisonnière
3. Représenter la série chronologique désaisonnalisé
4. Représenter la composante irrégulière
5. Décomposer la série

Exercice 3

Suivi du Log(FTAM) pendant 5 ans avec une mesure hebdomadaire

- Représenter la série chronologique stationnarisée

```
plot(d$décès-lm(d$décès ~time)$fitted.value,type='l')
```



Exercice 3

Suivi du Log(FTAM) pendant 5 ans avec une mesure hebdomadaire

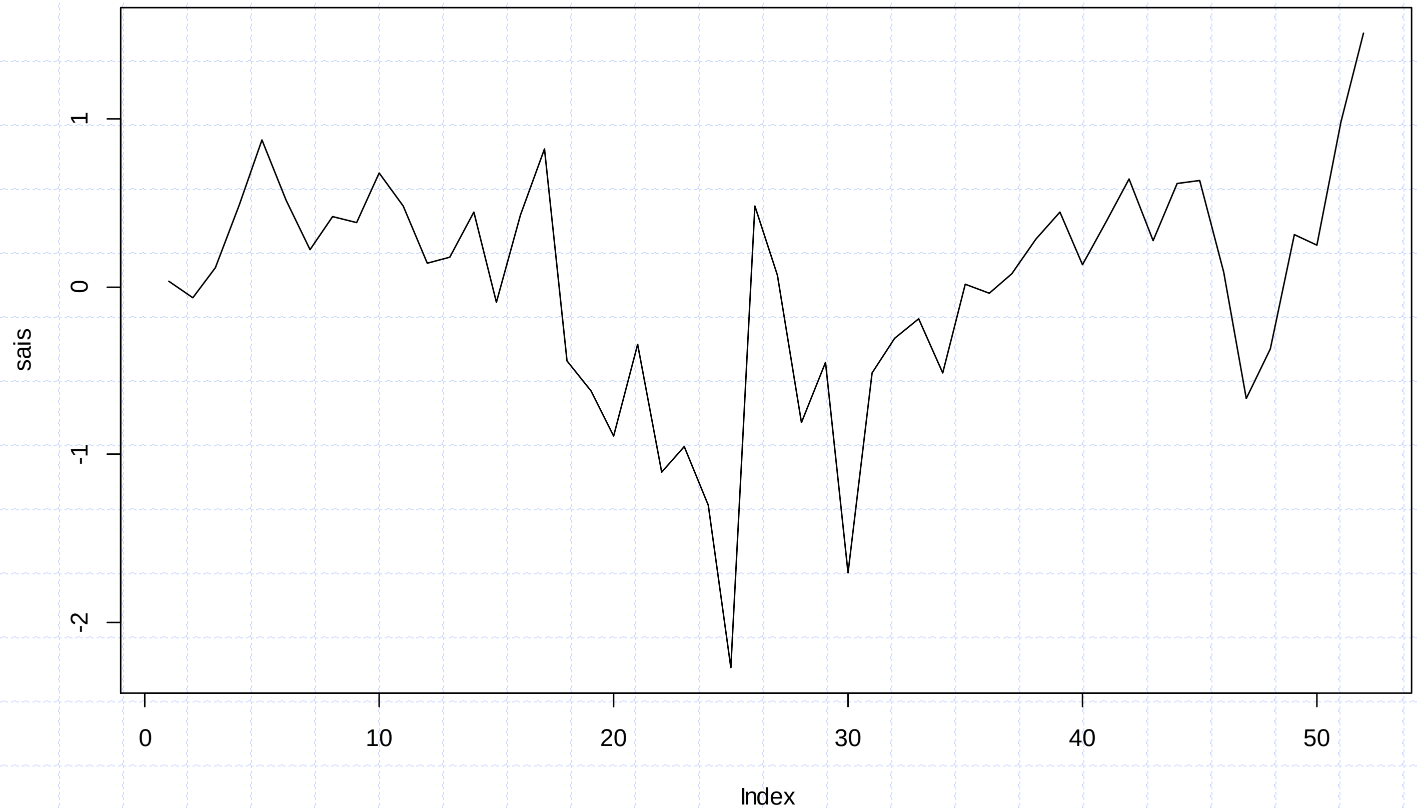
- Estimer et représenter la variation saisonnière

```
> stab<-logFTAM-lm(logFTAM~temps)$fitted.value
> nstab<-ts(stab, frequency = 52, start = c(2008, 1))
> sais<-vector(length=52)
> for (i in 1:52) mean(window(nstab,  start= c(2008,i ), end= c(2013,5), frequency =1))->sais[i]
> sais
 [1]  0.03330354 -0.07041112  0.10920755  0.49049288  0.87177822 0.51797815  0.21759682  0.41321549  0.37883416
[10]  0.67645283  0.48207150  …
> plot(sais,type='l')
```

Exercice 3

Suivi du Log(FTAM) pendant 5 ans avec une mesure hebdomadaire

- Estimer et représenter la variation saisonnière

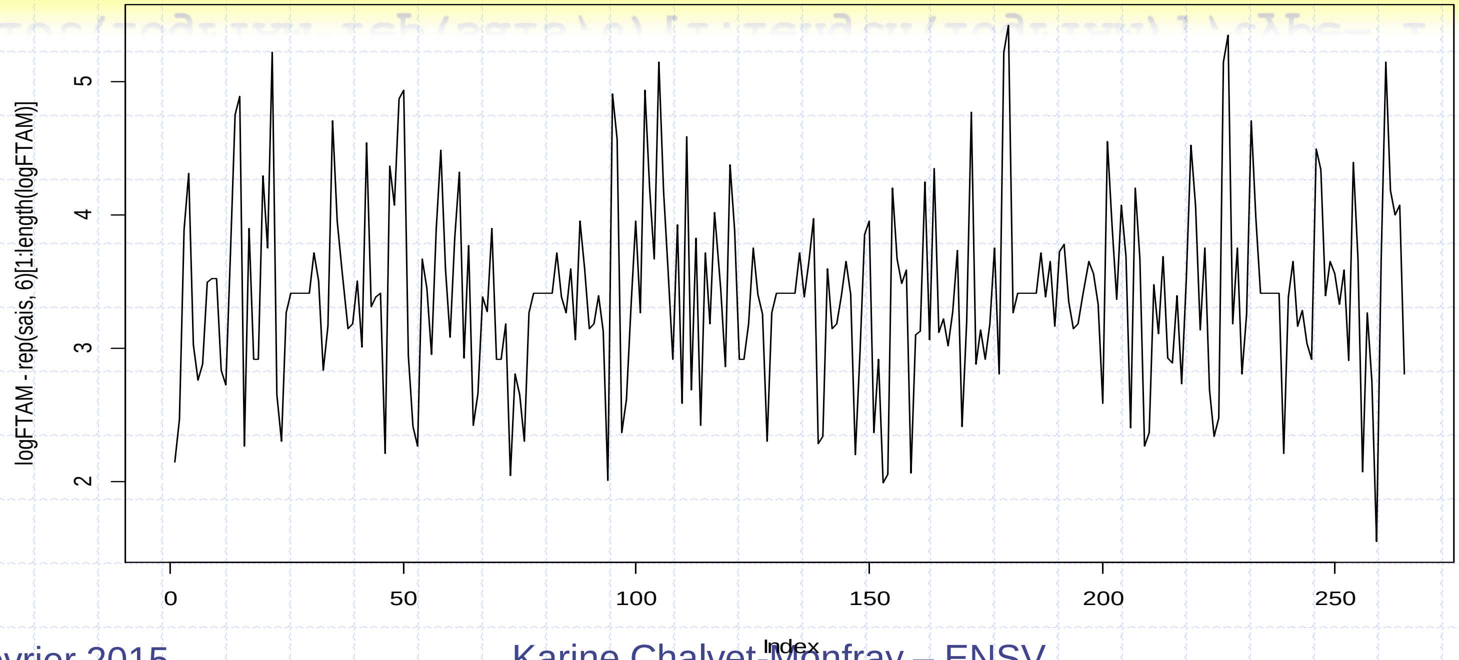


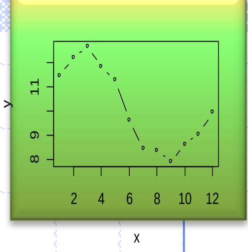
Exercice 3

Suivi du Log(FTAM) pendant 5 ans avec une mesure hebdomadaire

- Représenter la série chronologique désaisonnalisé

```
plot(logFTAM-rep(sais,6)[1:length(logFTAM)],type='l')
```



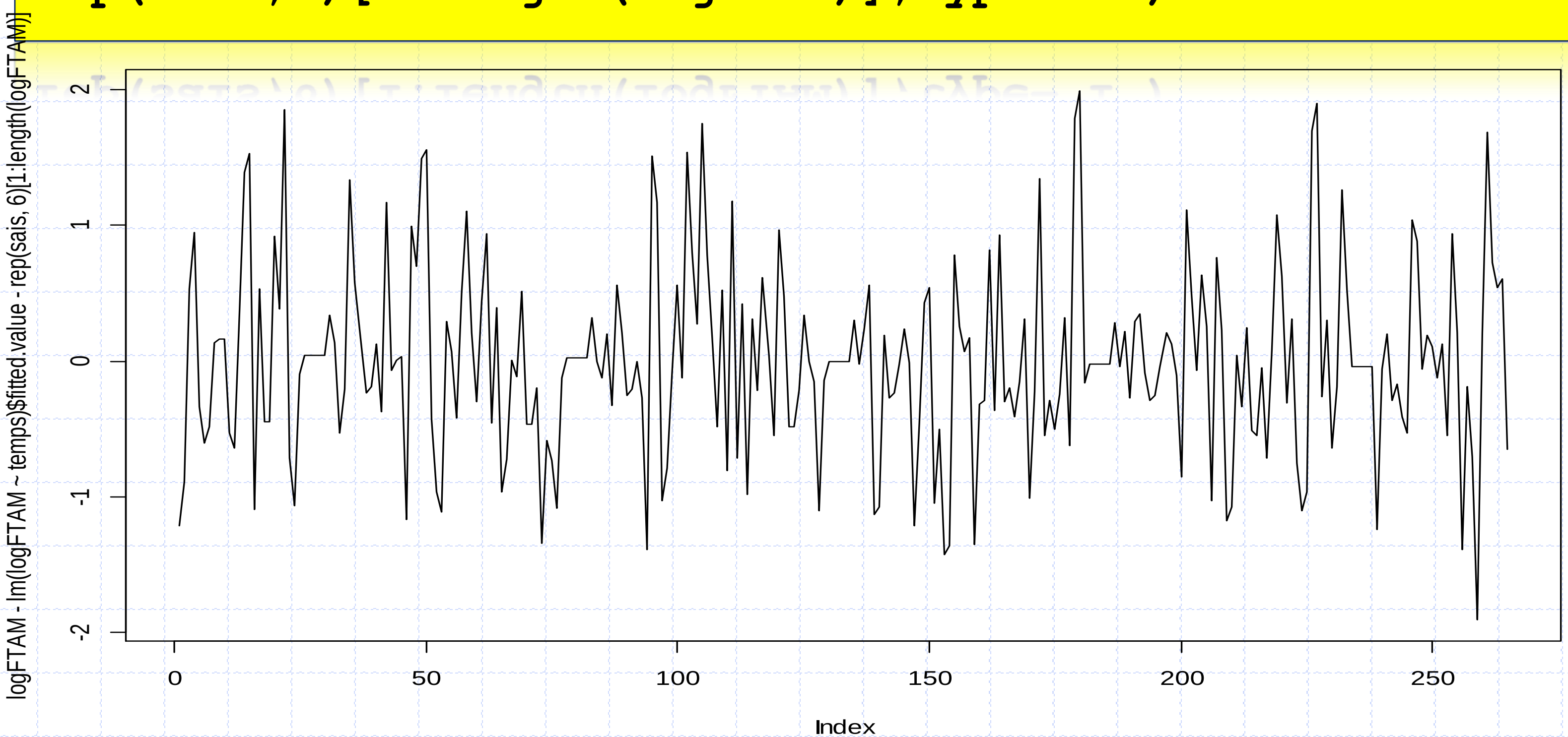


Exercice 3

Suivi du Log(FTAM) pendant 5 ans avec une mesure hebdomadaire

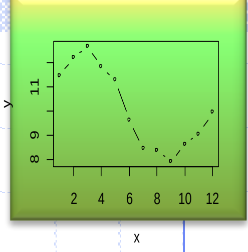
- Représenter la composante irrégulière

```
plot(logFTAM-lm(logFTAM~temps)$fitted.value-
rep(sais,6)[1:length(logFTAM)],type='l')
```



Exercice 3

Suivi du Log(FTAM) pendant 5 ans avec une mesure hebdomadaire



- Décomposer la série

```
> nlog<-ts(logFTAM,frequency = 52, start = c(2008, 1))
> m<-decompose(nlog)
> plot(m)
> m$figure
 [1] -0.08479595 -0.02729595 -0.15729595  0.10270405  1.12380982
0.67770405  0.34604540  0.38659828  0.35001174
[10]  0.65311270  0.63443482  0.31532424 …
```

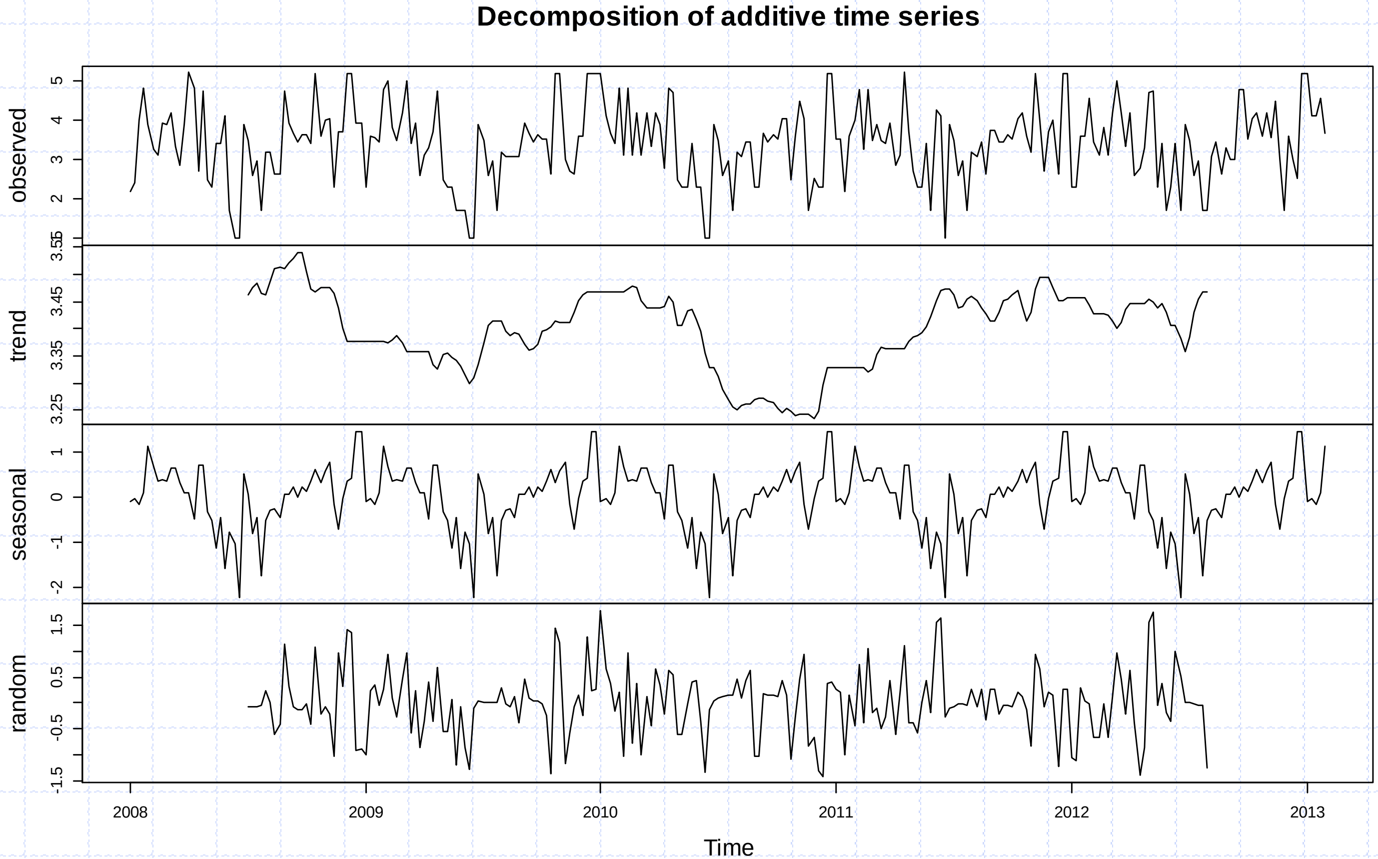
```
sais
 [1]  0.03330354 -0.07041112  0.10920755  0.49049288  0.87177822
0.51797815  0.21759682  0.41321549  0.37883416
[10]  0.67645283  0.48207150  0.13569017…
```

Résultats un peu différents car pas la même estimation de TC

Exercice 3

Suivi du Log(FTAM) pendant 5 ans avec une mesure hebdomadaire

- Décomposer la série



On ne pouvait pas voir tous ceci avant l'analyse.

Plan

- Introduction
- Etude préliminaire
- Estimation de la tendance
- Estimation de la variation saisonnière
- Conclusion

Conclusion
Pour aller plus loin

- Mieux expliquer (l'ensemble ou la composante irrégulière)
  - Méthodes
    - Modélisation statistiques multifactorielle
    - ex. mortalité en fonction de la pollution, de la température,…
- Alerte précoce (composante irrégulière)
  - Différent méthodes selon ce que l'on recherche :
    - Saut brusque
    - Lente dérive

Conclusion
Pour aller plus loin

- Prédire l'avenir d'une série chronologique
  - Intérêt : Anticiper (ex. consommation électrique)
  - Méthodes
    - Lissage Exponentielle Simple ou Double
    - Modèles ARMA et ses dérivés basés sur des processus aléatoires
    - Modèles mécanistes (déterministes ou stochastique)
    - …

Conclusion

- Ceci est juste une introduction,

  -> visualisation de l'intérêt
- Il existe de nombreux outils qui ont été développés pour les séries chronologiques :
  - pour les analyser ;
  - mais aussi pour aussi les prédire.
- Ces outils sont très différents selon les hypothèses posées et les objectifs visés

Merci pour votre attention