Université Paris I, Panthéon - Sorbonne

Deuxième Année Master T.I.D.E. 2022 – 2023

Econométrie des séries chronologiques

Examen final, janvier 2023

Examen de 2h00. Tout document ou calculatrice est interdit.

1. **(13 points)** Avec le logiciel `R`, on simule la consommation semestrielle d'électricité d'une municipalité depuis 1990.

   (a) On commence par simuler une tendance et une saisonnalité:

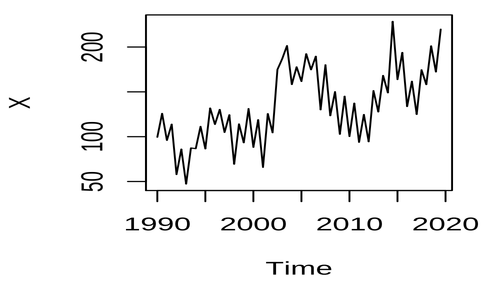
```
n=60
t=c(1:n)
a=(2*t^2-5*t)/(t-0.8)-20*log(0.001*t)
s=20*cos(t*pi)
```

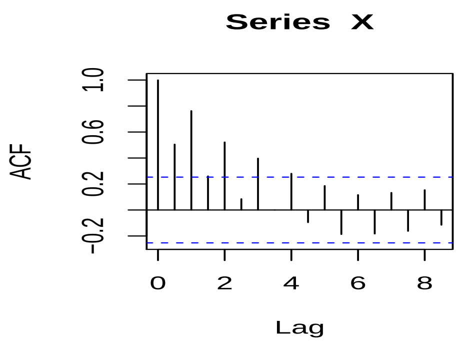
   Question 1: *Vérifier que* `s` *est bien une saisonnalité dont on précisera la période. Comment se comporte la tendance* `a` *quand* t *devient grand?*

   (b) On a ensuite tapé les commandes:

```
u=0; m=100
for (k in c(1:(n+m)))
u[k+1]=0.8*u[k]+20*(rexp(1,1)-1)
X=ts(a+s+u[m:(n+m-1)],1990,frequency = 2)
ts.plot(X)
acf(X)
```

   On a ainsi obtenu:





   Question 2: *Montrer que* `u[m:(n+m-1)]` *est une trajectoire d'un processus que l'on précisera. Pouvait-on s'attendre à un tel corrélogramme?*

   (c) On travaille maintenant à partir de la trajectoire `X` sans rien connaître d'autre et on essaie de prédire la consommation des années 2021 et 2022. On commence par les commandes:

```
Z=(-1)^t
reg1=lm(X~t+Z)
summary(reg1)
reg2=lm(X~t+I(t^2)+Z)
BIC(reg1,reg2)
```

On a ainsi obtenu:

```
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  90.5947     7.3235  12.370  < 2e-16 ***
t             1.4240     0.2088   6.819 6.36e-09 ***
Z            17.7245     3.6164   4.901 8.27e-06 ***

Residual standard error: 28 on 57 degrees of freedom
Multiple R-squared:  0.5599,Adjusted R-squared:  0.5445
F-statistic: 36.26 on 2 and 57 DF,  p-value: 6.94e-11

> BIC(reg1,reg2)
     df      BIC
reg1  4 583.4413
reg2  5 587.0427
```

Question 3: *Expliquer pourquoi les valeurs* `1.4240`, `17.7245` *ne sont pas surprenantes. Qu'a-t-on fait et que conclut-on après la dernière commande? Est-ce surprenant?*

(d) D'autres traitements sont effectués:

```
rho=acf(reg1$residuals)
(rho1=rho$acf[2])
library(forecast)
uchap=auto.arima(reg1$residuals,max.p=5,max.q=5,ic="bic")
uchap
Box.test(uchap$residuals^2,lag = 5,type = c("Box-Pierce"))
```

On a ainsi obtenu:

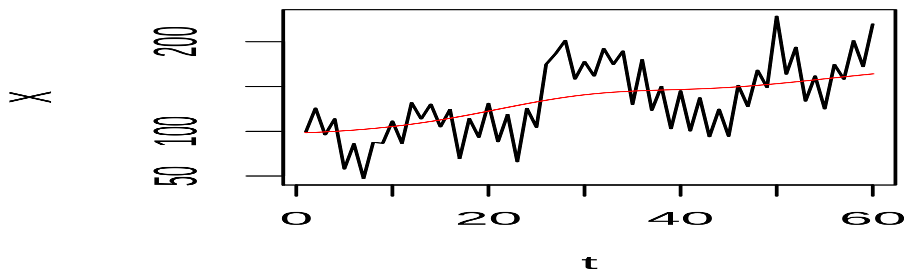
```
> (rho1=rho$acf[2])
[1] 0.7895965

> uchap
Series: reg1$residuals
ARIMA(1,0,0) with zero mean

Coefficients:
         ar1
      0.7998
s.e.  0.0755

> Box.test(uchap$residuals^2,lag = 5,type = c("Box-Pierce"))

Box-Pierce test
X-squared = 4.7198, df = 5, p-value = 0.451
```

Question 4: *Expliquer ce qu'est la valeur* `0.7895965` *et pourquoi elle n'est pas surprenante. Comment est obtenu* `uchap` *? Le résultat était-il attendu? Quelle conclusion tirer de la valeur* `0.451`*? Est-ce normal?*

(e) Par la suite on essaie autre chose:

```
hn=10; Num=0; Denom=0; Tend=0
for (x in c(1:n))
{ Num[x]=sum(dnorm((t-x)/hn)*X)/hn
Denom[x]=sum(dnorm((t-x)/hn))/hn
Tend[x]=Num[x]/Denom[x]
}
plot(t,X,'l')
lines(t,Tend,col='red')
```

d'où le graphe:



Question 5: *Expliquer ce qui a été fait. Que pensez-vous du choix de* `hn`*? Qu'aurait-on également pu faire?*

(f) Voici les commandes qui suivent:

```
resn=X-Tend
reg3=lm(resn~Z)
library(forecast)
uchap2=auto.arima(reg3$residuals,max.p=5,max.q=5,ic="bic")
Box.test(uchap2$residuals^2,lag = 5,type = c("Box-Pierce"))
```

et voici les résultats numériques obtenus:

```
> uchap2
ARIMA(1,0,0) with zero mean

Coefficients:
         ar1
      0.7891
s.e.  0.0787

> Box.test(uchap2$residuals^2,lag = 5,type = c("Box-Pierce"))

data:  uchap2$residuals^2
X-squared = 4.9605, df = 5, p-value = 0.4207
```

Question 6: *Expliquer pourquoi l'on a fait une régression et les résultats numériques obtenus. Quelle conclusion pouvez-vous en tirer?*

(g) Enfin, on effectue:

```
tt=c((n+1):(n+4))
new=data.frame(t=tt,Z=(-1)^tt)
pred1=predict(reg1,new)
u1=predict(uchap,n.ahead=4); u1
Xpred=pred1+u1$pred
Xpred
```

et voici les résultats numériques obtenus:

```
> u1
$pred
Time Series:
Start = 61
End = 64
Frequency = 1
[1] 20.85359 16.67935 13.34066 10.67028

> Xpred
Time Series:
Start = 61
End = 64
Frequency = 1
```

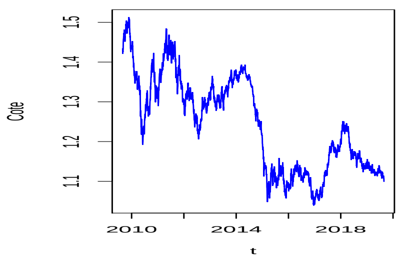
```
       1         2         3         4
180.5899 213.2886 175.9251 210.1276
```

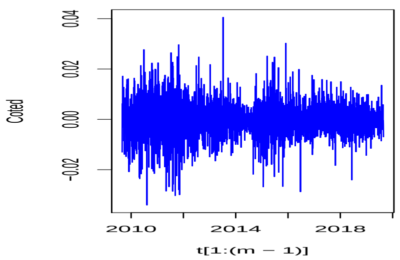
Question 7: *Expliquer ce qui a été fait. Que représentnt les valeurs de* `pred1`*,* `u1` *et* `Xpred`*? Comment sont-elles calculées? Si on avait prédit les 2 semestres de 2031, qu'aurait-on environ obtenu?*

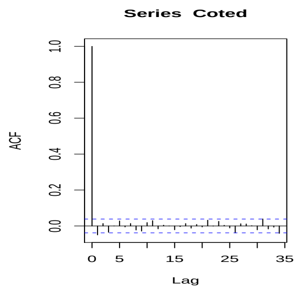
2. **(8 points)** On s'intéresse à modéliser la cotation quotidienne en clôture du taux de change Euro/Dollar (variable `Cote`) depuis le 01/09/2009 jusqu'au 01/09/2019. On utilise le logiciel `R` à cet effet.

   (a) Voici les premières commandes effectuées:

```
Cote=as.numeric(ED$Close)
t=as.numeric(ED$Time)
m=length(Cote)
plot(t,Cote,'l',col='blue')
Coted=Cote[2:m]-Cote[1:(m-1)]
ts.plot(Coted)
acf(Coted)
Box.test(Coted, lag = 10, type = c("Ljung-Box"))
```







Avec pour résultats numériques:

```
Box-Ljung test
data:  Coted
X-squared = 17.663, df = 10, p-value = 0.06092
```

Question 1: *Que représente* `Coded` *par rapport à* `Cote`*? Pourquoi peut-on utliser* `Coded` *plutôt que les log-returns dans ce cadre? Qu'en conclure quant à la série* `Coded`*? Qu'en déduire, dans un premier temps, quant au type de processus qu'est* `Cote`*?*

   (b) On exécute alors:

```
Fit1=garchFit(~garch(2,2),data=Coted,trace=FALSE)
BIC1=Fit1@fit$ics[2];
Fit2=garchFit(~garch(1,1),data=Coted,trace=FALSE)
summary(Fit2)
BIC2=Fit2@fit$ics[2];
c(BIC1, BIC2)
```

Avec pour résultats numériques:

```
        Estimate  Std. Error  t value Pr(>|t|)
mu     -8.813e-05   1.194e-04   -0.738   0.4606
omega   1.111e-07   4.771e-08    2.328   0.0199 *
alpha1  2.585e-02   3.704e-03    6.977 3.01e-12 ***
beta1   9.718e-01   3.824e-03  254.135  < 2e-16 ***

> c(BIC1, BIC2)
      BIC       BIC
-7.164414 -7.167115
```

Question 2: *Expliquer ce qui a été fait. Que conclure de la valeur* `< 2e-16`*? Comment sont calculées concrètement les valeurs* `3.704e-03` *et* `6.977`*? Que conclure finalement des valeurs* `-7.164414` *et* `-7.167115`*?*

(c) On continue alors par:

```
Fit3=arima(Coted^2,c(1,0,1))
Fit3
```

On obtient alors:

```
Coefficients:
         ar1      ma1  intercept
      0.9968  -0.9732      1e-04
s.e.  0.0020   0.0056      0e+00
```

Question 3: *: Qu'a-t-on fait ici et pourquoi? Ces résultats vous semblent-ils cohérents avec les précédents? En particulier, expliquez comment les coefficients obtenus pouvaient être calculés à partir de ceux de la question II.2.*

(d) On tape enfin:

```
AC=as.numeric(unlist(acf(Fit3$residuals^2)))
B=m*sum(AC[2:11]^2); B
1-pchisq(B,10)
```

Et on obtient les réultats:

```
B=m*sum(AC[2:11]^2); B
[1] 6.741783
> 1-pchisq(B,10)
[1] 0.7495766
```

Question 4: *Pourquoi a-t-on exécuté ces commandes? Que représentent les 2 valeurs numériques obtenues? Qu'en déduire?*

(e) On cherche à calculer la Value-at-Risk (perte maximale) à horizon 1 jour d'un euro au 01/09/2019 à risque α, c'est-à-dire valeur $V_m(\alpha)$ telle que $\mathbb{P}\big(Coted(m) \leq V_m(\alpha) \mid Coted(m-1), Coted(m-2), \dots\big) = \alpha$. Pour $\alpha = 0.05$, et en supposant gaussienne la distribution conditionnelle, donner les commandes permettant de calculer $V_m(\alpha)$.