

## Deuxième Année Master T.I.D.E. 2020 – 2021

# Econométrie des séries chronologiques

Examen final, janvier 2021

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

1. (13 points) Avec le logiciel R, on simule une tendance et une saisonnalité:

```
t=c(1:180)
a=150-5*t*exp(-0.03*t)
s=20*cos(t*2*pi/12)-10*sin(t*2*pi/12)
plot(t,a+s)
```

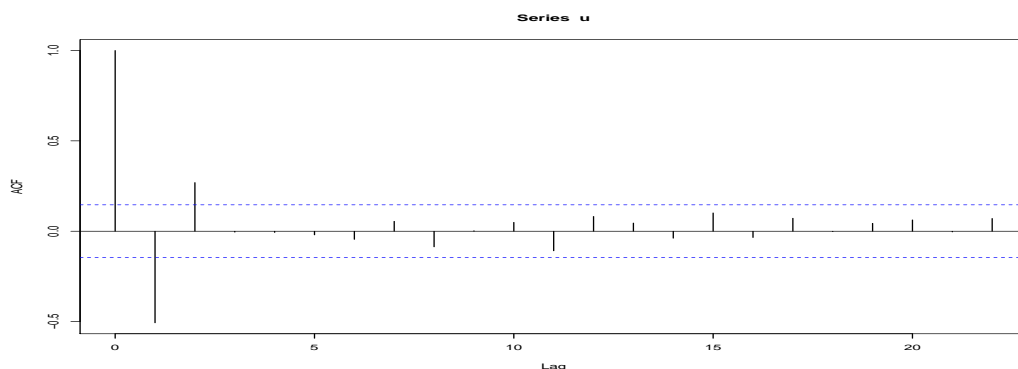
**Question 1:** Vérifier que  $s$  est bien une saisonnalité dont on précisera la période. Comment se comporte la tendance quand  $t$  devient grand?

On a ensuite tapé les commandes:

```
epsi=runif(182,-40,40)
u=epsi[3:182]-0.5*epsi[2:181]+0.4*epsi[1:180]
acf(u)
Box.test(u,lag = 5,type = c("Box-Pierce"))
```

On a ainsi obtenu:

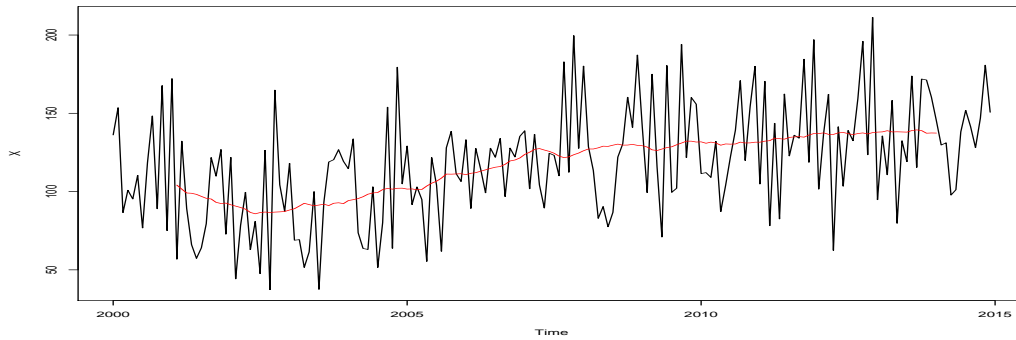
```
Box-Pierce test
X-squared = 59.382, df = 5, p-value = 1.631e-11
```



**Question 2:** Donner une formule théorique pour  $u$ . Pouvait-on s'attendre à un tel corrélogramme? Comment expliquer la  $p$ -value très faible?

On a ensuite tapé des commandes et obtenu le graphe suivant:

```
X=ts(a+s+u,2000,frequency = 12)
ts.plot(X)
m=24; tt=c(13:168)
am=filter(X,c(1/(2*m),rep(1/m,m-1),1/(2*m)))
lines(2000+tt/12,am[tt],col='red')
```



**Question 3:** On considère désormais que  $X$  a été observée, mais que  $a$ ,  $s$  et  $u$  ne sont pas connus. Donner la formule générale en fonction des composantes de  $X$ . Expliquer ce qui a été fait en calculant  $am$ . Pourquoi le choix  $m=24$ ?

D'autres traitements sont effectués:

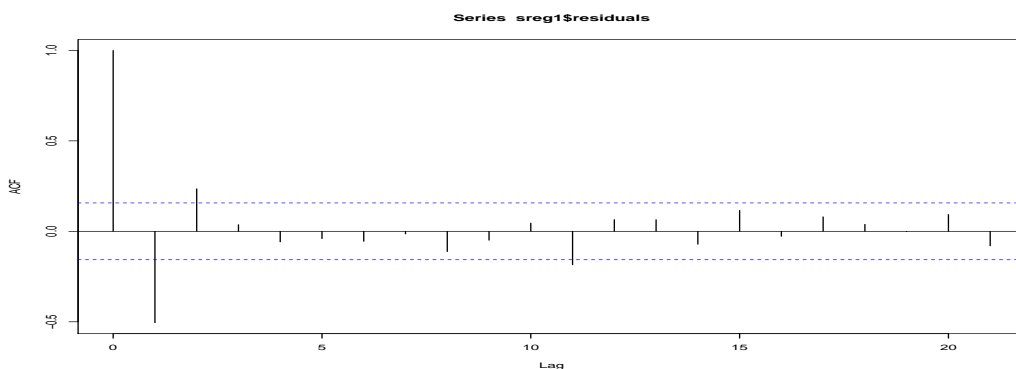
```
res=X[tt]-am[tt]
co=cos(2*pi*tt/12); si=sin(2*pi*tt/12)
sreg1=lm(res~co+si)
summary(sreg1)
acf(sreg1$residuals)
```

On a ainsi obtenu:

Coefficients:

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	0.04694	2.28177	0.021	0.983613
co	20.12463	3.22691	6.237	4.17e-09 ***
si	-11.70100	3.22691	-3.626	0.000391 ***

Residual standard error: 28.5 on 153 degrees of freedom  
 Multiple R-squared: 0.2538, Adjusted R-squared: 0.2441  
 F-statistic: 26.02 on 2 and 153 DF, p-value: 1.874e-10



**Question 4:** Expliquer ce qui a été fait. Que représente  $res$  par rapport à  $X$ ? Pourquoi la régression effectuée et que penser des résultats obtenus? Pouvait-on s'attendre à la valeur  $-11.70100$ ? Pouvait-on s'attendre au corrélogramme obtenu?

Par la suite on a tapé les commandes:

```
long=length(res); T=12; Part=floor(long/T)
Z=matrix(nrow=long,ncol=(T-1))
for (i in 1:long)
  for (j in 1:(T-1))
    Z[i,j]=floor(abs(cos(pi*(i-j)/T)))
```

```

for (i in 1:Part)
  for (j in 1:(T-1))
    {ii=i*(T)
      Z[ii,j]==-1}
sreg2=lm(res~Z)
BIC(sreg1,sreg2)

```

d'où les résultats numériques obtenus:

```

      df      BIC
sreg1  4 1505.041
sreg2 13 1536.895

```

**Question 5:** *Expliquer ce qui a été fait. Comment expliquer le nombre 13? Quel choix en conclure? Pouvait-on s'attendre à de tels résultats?*

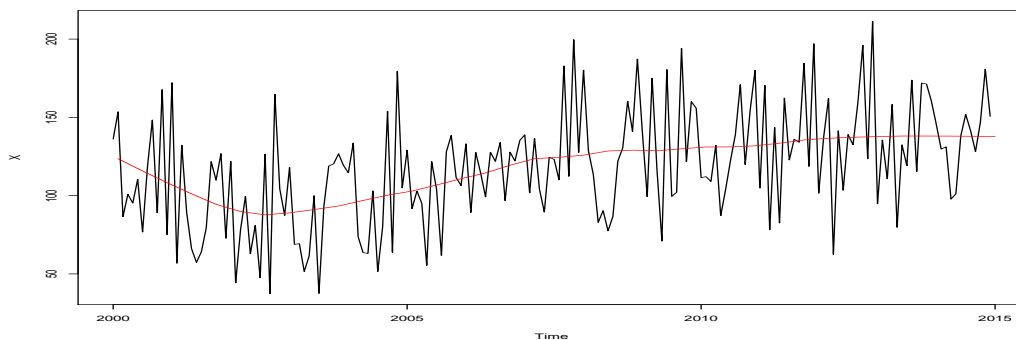
Voici les commandes qui suivent:

```

st=stl(X,s.win="perio",t.win=43)
ts.plot(X)
lines(2000+t/12,st$time[,"trend"],col='red')
mean(abs(a[tt]-am[tt])); mean(abs(a[tt]-st$time[,"trend"][tt]))
mean(abs(s[tt]-sreg1$fitted.values)); mean(abs(s-st$time[,"seasonal"]))
mean(abs(u[tt]-sreg1$residuals)); mean(abs(u[tt]-st$time[,"remainder"][tt]))

```

et voici la figure et les résultats numériques obtenus:



```

> [1] 3.210916 [1] 2.976393
> [1] 1.078806 [1] 5.614236
> [1] 3.355175 [1] 6.469319

```

**Question 6:** *Expliquer ce qui a été fait. Quelle conclusion pouvez-vous tirer?*

Enfin, on exécute:

```

library(forecast)
auto.arima(sreg1$residuals,max.p=5,max.q=5,ic="bic")

```

et voici les résultats numériques obtenus:

```
ARIMA(0,0,2) with zero mean
```

```

Coefficients:
      ma1      ma2
    -0.5291  0.4286
s.e.    0.0699  0.0796

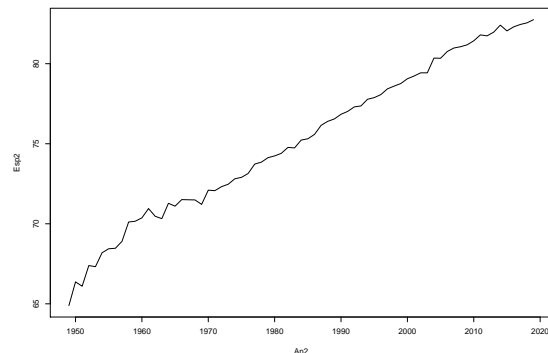
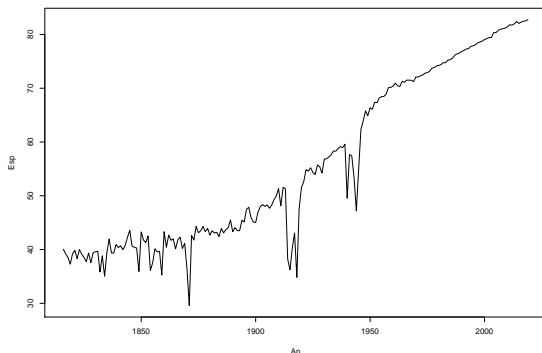
```

**Question 7:** *Expliquer ce qui a été fait. Pouvait-on s'attendre à un tel résultat?*

2. (8 points) On s'intéresse à l'évolution annuelle de l'espérance de vie moyenne (hommes et femmes confondus) en France depuis plus de 200 ans. On désire prédire cette espérance de vie pour de 2020 à 2025. Pour cela on commence par écrire:

```
Esp=as.numeric(EsperanceVie$Total)
An=EsperanceVie$Year
plot(An,Esp,'l')
Esp2=as.numeric(EsperanceVie$Total[134:204])
An2=EsperanceVie$Year[134:204]
plot(An2,Esp2,'l')
```

Voici les graphes obtenus:

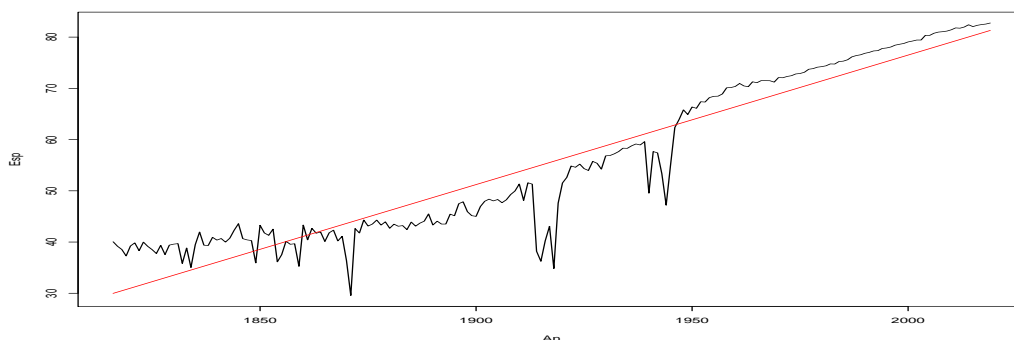


**Question 1:** Décrire ce qui a été fait. Quel est l'avantage de `Esp2` par rapport à `Esp1` pour prédire de 2020 à 2025?

Par la suite, on tape les commandes:

```
reg1=lm(Esp~An)
summary(reg1)
plot(An,Esp,'l')
lines(An,reg1$fitted.values,col='red')
new=data.frame(An=c(2020:2025))
(pred1=predict.lm(reg1,newdata=new))
```

Voici le graphe et les résultats numériques obtenus:



	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	-429.05464	11.29982	-37.97	<2e-16 ***
An	0.25278	0.00589	42.91	<2e-16 ***

Residual standard error: 4.954 on 202 degrees of freedom  
 Multiple R-squared: 0.9012, Adjusted R-squared: 0.9007  
 F-statistic: 1842 on 1 and 202 DF, p-value: < 2.2e-16

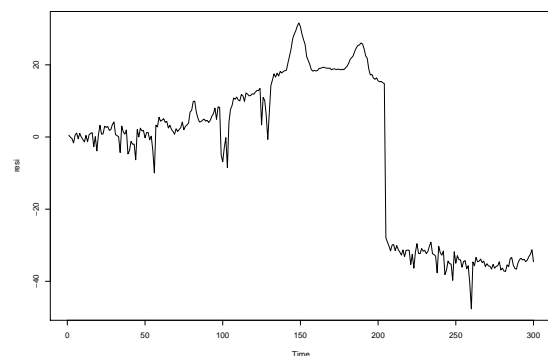
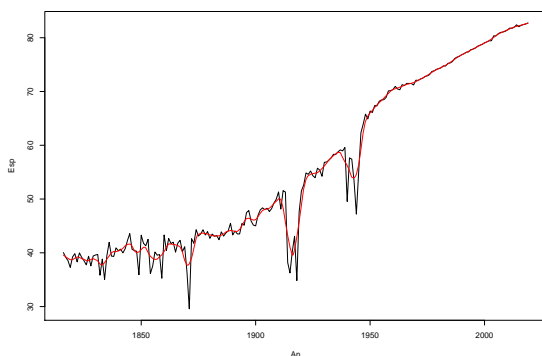
```
>      1      2      3      4      5      6
> 81.55081 81.80358 82.05636 82.30913 82.56191 82.81468
```

**Question 2:** *Décrire ce qui a été fait. Que pensez vous de la pertinence de ces prédictions?*

Par la suite, on tape les commandes:

```
library(lokern)
kerada=glkerns(An,Esp)
plot(An,Esp,'l')
lines(kerada,col='red')
resi=Esp-kerada$est
ts.plot(resi)
```

Voici les graphes obtenus:



**Question 3:** *Décrire ce qui a été fait. Que représente resi? Que conclure de sa représentation?*

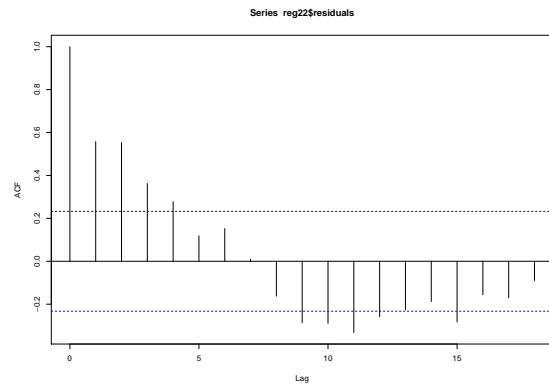
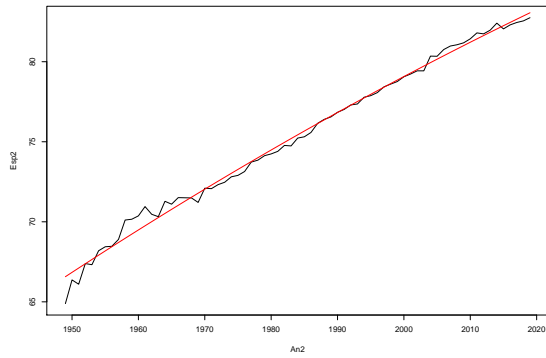
On s'intéresse ensuite au vecteur Esp2:

```
reg21=lm(Esp2~poly(An2,1,raw='T'))
reg22=lm(Esp2~poly(An2,2,raw='T'))
reg23=lm(Esp2~poly(An2,3,raw='T'))
BIC(reg21,reg22,reg23)
plot(An2,Esp2,'l')
lines(An2,reg22$fitted.values,col='red')
new=data.frame(An2=c(2020:2025))
(pred2=predict.lm(reg22,newdata=new))
```

Voici les résultats numériques et les graphes obtenus:

	df	BIC
reg21	3	106.11126
reg22	4	97.51639
reg23	5	98.64133

```
>      1      2      3      4      5      6
> 83.25779 83.45639 83.65398 83.85056 84.04612 84.24068
```

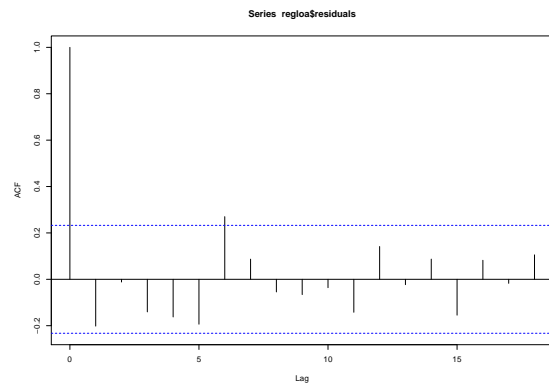
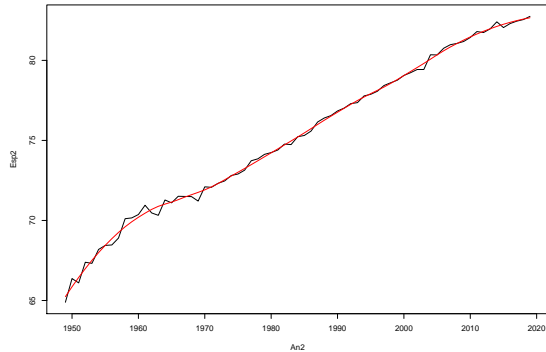


**Question 4:** *Qu'a-t-il été fait par ces commandes? Que suggère le corrélogramme? Que pensez vous des prédictions obtenues?*

Pour terminer, on exécute les commandes suivantes:

```
library(fANCOVA,quietly = T)
regloa=loess.as(An2,Esp2,degree = 2,criterion=c("aicc","gcv")[2],family ="gaussian",plot="FALSE")
plot(An2,Esp2,'l')
lines(An2,regloa$fitted,col='red')
plot(An2,regloa$residuals,'l')
acf(regloa$residuals)
```

Voici les graphes obtenus:



**Question 5:** *Qu'a-t-il été fait par ces commandes? Que suggère le corrélogramme? Comment expliquer son allure par rapport au précédent? Que suggérez vous pour améliorer encore les prédictions?*