

Deuxième Année Master T.I.D.E. 2018 – 2019

Econométrie des séries chronologiques

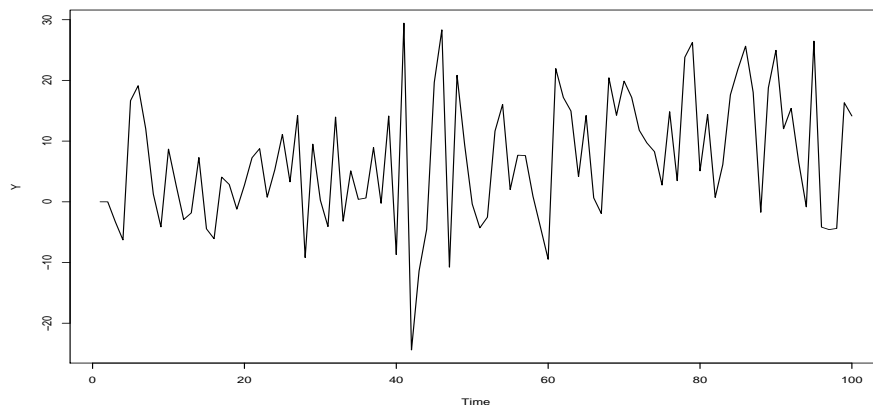
Examen final, janvier 2020

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

1. (7 points) On a tapé les commandes suivantes en R:

```
u=runif(100,-2,2)
X[1:100]=0; Y[1:100]=0
for (i in c(3:100))
{X[i]=sqrt(0.6*X[i-1]^2+1)*u[i]
Y[i]=2+0.1*log(1+2*exp(i))+5*X[i]}
ts.plot(Y)
```

Voici le graphe produit:



Question 1: *Qu'est ce que u ? Quel processus est sous-jacent à la simulation de la trajectoire X (formaliser...)? Est-il stationnaire? Mêmes questions pour Y .*

On a ensuite tapé les commandes:

```
t=c(1:100)
reg=lm(Y ~ t)
summary(reg)
```

On a ainsi obtenu:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	0.91377	2.00650	0.455	0.64983
t	0.11644	0.03449	3.376	0.00106 **

Residual standard error: 9.957 on 98 degrees of freedom

Multiple R-squared: 0.1042, Adjusted R-squared: 0.09502

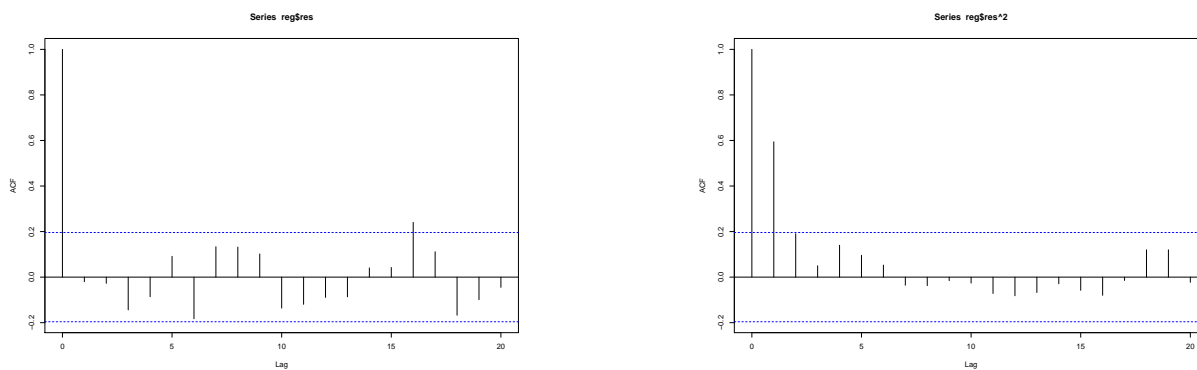
F-statistic: 11.39 on 1 and 98 DF, p-value: 0.001057

Question 2: *Que pensez vous des valeurs 0.91377 et 0.11644? Pourquoi sont-elles de bonnes estimations? Comment expliquer la p -value très faible?*

Enfin on a tapé les commandes:

```
acf(reg$res)
acf(reg$res^2)
```

et obtenu les graphes suivants:



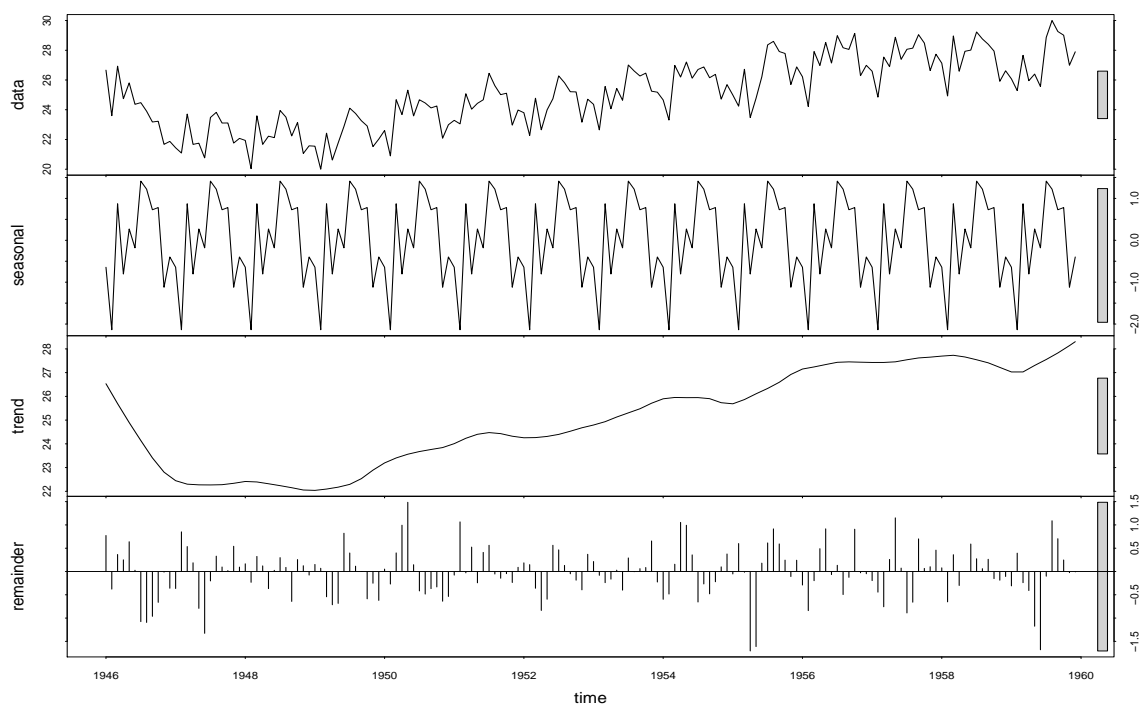
Question 3: *Qu'est-ce qui a été fait et que représentent ces deux graphes? Pouvait-on s'attendre à de tels résultats?*

2. (13 points) On s'intéresse à l'évolution mensuelle des naissances (en milliers) à New-York entre 1946 et 1959. Voici les premières commandes pour visionner et traiter ces données:

```
births <- scan("http://robjhyndman.com/tsdldata/data/nybirths.dat")
tsbirths <- ts(births, frequency=12, start=c(1946,1))
s=stl(tsbirths,s.win="perio")
plot(s)
Box.test(s$time[, "remainder"], lag = 10, type = c("Ljung-Box"))
```

Voici les résultats numériques et les graphes qui ont été obtenus:

X-squared = 65.052, df = 10, p-value = 3.962e-10



Question 1: *Décrire ce qui a été fait. Sur quoi porte le test effectué, quel est-il et que peut-on conclure du résultat?*

Par la suite, on tape les commandes:

```
library(forecast)
fit=auto.arima(as.numeric(s$time[,"remainder"]),max.p=5,max.q=5,ic="bic")
fit
Box.test(fit$res, lag = 10, type = c("Ljung-Box"))
1-pchisq(19.579,9)
```

Voici les résultats numériques obtenus:

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

Coefficients:
      ma1
      0.4711
s.e.  0.0624

> Box.test(fit$res, lag = 10, type = c("Ljung-Box"))
X-squared = 19.579, df = 10, p-value = 0.03349

> 1-pchisq(19.579,9)
[1] 0.02069593
```

Question 2: *Décrire ce qui a été fait. Quelles conclusions peut-on déduire de ces résultats? Expliquer ce que représente la valeur 0.02069593.*

Par la suite, on tape les commandes:

```
t=c(1:156)/12
reg2=lm(tsbirths[13:168]~t)
summary(reg2)
Tend=ts(reg2$fitted.values,frequency = 12, c(1947,1))
ts.plot(tsbirths)
lines(Tend,col="red")
res2=ts(reg2$residuals,frequency = 12, c(1947,1))
s2=stl(res2,s.win="perio",t.win=500)
fit2=auto.arima(as.numeric(s2$time[,"remainder"]),max.p=5,max.q=5,ic="bic")
fit2
Box.test(fit2$res, lag = 10, type = c("Ljung-Box"))
1-pchisq(7.8242,8)
```

Voici les résultats numériques obtenus:

```
> summary(reg2)

Call:
lm(formula = tsbirths[13:168] ~ t)

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  21.68391    0.20487  105.84  <2e-16 ***
t              0.52607    0.02717   19.36  <2e-16 ***

Residual standard error: 1.273 on 154 degrees of freedom
Multiple R-squared:  0.7089, Adjusted R-squared:  0.707
F-statistic:  375 on 1 and 154 DF, p-value: < 2.2e-16
> fit2
```

```
Series: as.numeric(s2$time[, "remainder"])
ARIMA(2,0,0) with zero mean
```

```
Coefficients:
```

```
      ar1      ar2
      0.7451 -0.1452
s.e.  0.0791  0.0800
```

```
sigma^2 estimated as 0.2961: log likelihood=-125.72
```

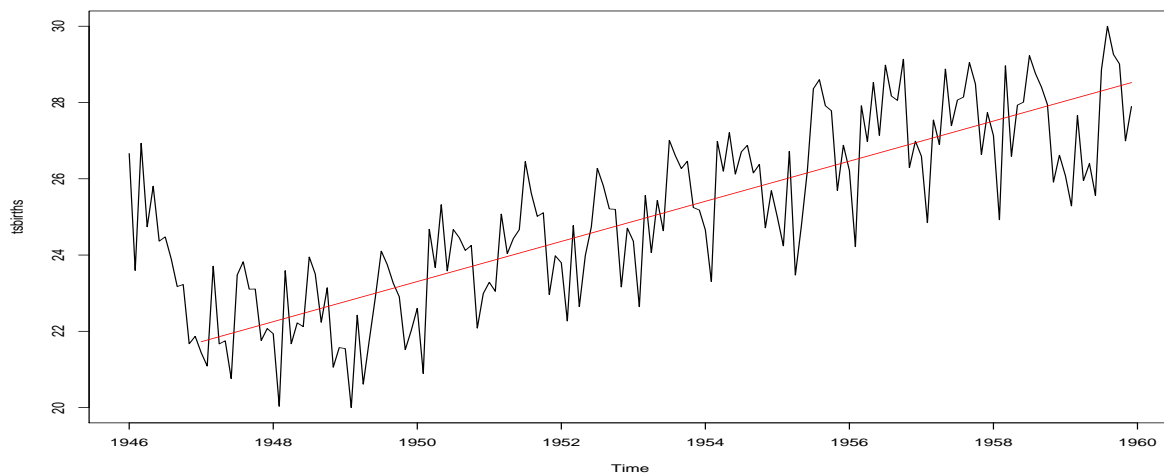
```
AIC=257.43 AICc=257.59 BIC=266.58
```

```
> Box.test(fit2$res, lag = 10, type = c("Ljung-Box"))
```

```
X-squared = 7.8242, df = 10, p-value = 0.646
```

```
> 1-pchisq(7.8242,8)
```

```
[1] 0.4508287
```



Question 3: *Décrire ce qui a été fait. Expliquer ce que représentent les valeurs numériques 19.36 et -0.1452. Quelles conclusions peut-on tirer de ces résultats?*

Question 4: *Expliquer votre démarche puis écrire les commandes en R qui permettraient de prédire le nombre de naissances pour tous les mois de l'année 1960.*