Université Paris VII - Paris I

Master 2 M0 2017 – 2018

Analyse des séries financières

Final exam, February 2018

3h00, without any documents.

In all the sequel, $(\xi_t)_{t\in\mathbf{Z}}$ is a sequence of centered independent and identically distributed random variables with a symmetric distribution (the distributions of ξ_0 and $-\xi_0$ are the same) continuous with respect to Lebesgue measure, such as $\mathbb{E}(\xi_0^2)=1$.

1. **Preliminaries:**

   (a) Let $(X_t)_{t\in\mathbf{Z}}$ be a stationary time series. Prove that $(X_t^2)_{t\in\mathbf{Z}}$ is a stationary time series.

   (b) Assume that $(Y_t^2)_{t\in\mathbf{Z}}$ is a stationary time series. Prove that $(Y_t)_{t\in\mathbf{Z}}$ is not necessary a stationary time series.

   (c) Let $(u_t)_{t\in\mathbf{Z}}$ be a stationary time series, independent to $(\xi_t)_{t\in\mathbf{Z}}$. Prove that $(\xi_t\, u_t)_{t\in\mathbf{Z}}$ is a stationary time series.

   (d) For a random variable Z, define $\operatorname{sign}(Z)=\mathbb{I}_{Z>0}-\mathbb{I}_{Z<0}$. Prove that $(\operatorname{sign}(\xi_t))_{t\in\mathbf{Z}}$ is a white noise, independent of $(|\xi_t|)_{t\in\mathbf{Z}}$. Let $(Y_t)_{t\in\mathbf{Z}}$ be a time series defined by $Y_t=\xi_t\, G\big((|Y_{t-i}|)_{i\in\mathbf{N}^*}\big)$ for any $t\in\mathbf{Z}$ where $G:\mathbf{R}^{\mathbf{N}}\to(0,\infty)$ is a fixed function. Assume that $(Y_t^2)_{t\in\mathbf{Z}}$ is a causal (with respect to $((\xi_s)_{s\le t})_{t\in\mathbf{Z}}$) stationary process. Prove that $(Y_t)_{t\in\mathbf{Z}}$ is also a causal stationary time series.

2. **Main theoretical part:** If it exists, we consider a sequence $(X_t)_{t\in\mathbf{Z}}$ such as:

$$X_t=\alpha\, X_{t-1}+\varepsilon_t \quad \text{with} \quad \varepsilon_t=\xi_t\,\sqrt{a_0+a_1X_{t-1}^2} \quad \text{for any } t\in\mathbf{Z} \tag{1}$$

   where $(\alpha,a_0,a_1)\in]=\mathbf{R}\times(0,\infty)\times[0,\infty)$ are unknown parameters.

   (a) In this question, we assume $\alpha=a_1=0$. Which kind of process is $(X_t)_{t\in\mathbf{Z}}$ and provide a condition of the existence of a stationary causal second order solution. Compute $\mathbb{E}(X_0)$ and $r_X(k)=\operatorname{cov}(X_0,X_k)$ for $k\in\mathbf{N}$.

   (b) In this question, we assume $a_1=0$ and $\alpha\neq 0$. Which kind of process is $(X_t)_{t\in\mathbf{Z}}$ and provide a condition of the existence of a stationary causal second order solution. Compute $\mathbb{E}(X_0)$ and $r_X(k)$ for $k\in\mathbf{N}$.

   (c) In this question, we assume $\alpha=0$ and $a_1>0$. Which kind of process is $(X_t)_{t\in\mathbf{Z}}$ and provide condition of the existence of a stationary causal second order solution. Compute $\mathbb{E}(X_0)$ and $r_X(k)$ for $k\in\mathbf{N}$.

   (d) Now and until the end, we study the general case $(\alpha,a_0,a_1)\in]=\mathbf{R}\times(0,\infty)\times[0,\infty)$. Prove that $(X_t)_{t\in\mathbf{Z}}$ is an affine causal process. Prove that the function $x\to\sqrt{1+x^2}$ is Lipshitzian and deduce that a sufficient condition for $(X_t)_{t\in\mathbf{Z}}$ to be a causal stationary second order process is:

$$|\alpha|+\sqrt{a_1}<1. \tag{2}$$

   (e) For $|\alpha|<1$, prove that if $(X_t)_{t\in\mathbf{Z}}$ is a causal stationary second order process then $(\varepsilon_t)_{t\in\mathbf{Z}}$ defined by

$$\varepsilon_t=\xi_t\sqrt{a_0+a_1\Big(\sum_{i=0}^{\infty}\alpha^i\,\varepsilon_{t-1-i}\Big)^2} \quad \text{for any } t\in\mathbf{Z}, \tag{3}$$

   is a causal stationary second order process and a weak white noise. Show that if $(\varepsilon_t)_{t\in\mathbf{Z}}$ is a causal stationary second order process then

$$a_1+\alpha^2<1. \tag{4}$$

   Compare this condition with (2).

   (f) In Doukhan *et al.* (2016), it was established that under (4), then $(\varepsilon_t)_{t\in\mathbf{Z}}$ is a causal stationary second order process. Extend this property to $(X_t)_{t\in\mathbf{Z}}$. Under (4), compute $\mathbb{E}(X_0)$ and $r_X(k)$ for $k\in\mathbf{N}$.

   (g) Deduce also $\mathbb{E}(X_t \mid (X_{t-s})_{s\in\mathbf{N}^*})$ and $\operatorname{var}(X_t \mid (X_{t-s})_{s\in\mathbf{N}^*})$. Is $(X_t)_{t\in\mathbf{Z}}$ a conditionaly heteroskedastic process?

   (h) Assume now that $(X_1,\cdots,X_N)$ is observed and let $\theta={}^t(\alpha,a_0,a_1)$. Provide the expression of the quasi-maximum likelihood estimator $\widehat{\theta}$ of θ. Is $\widehat{\theta}$ a consistent estimator? What is its convergence rate?

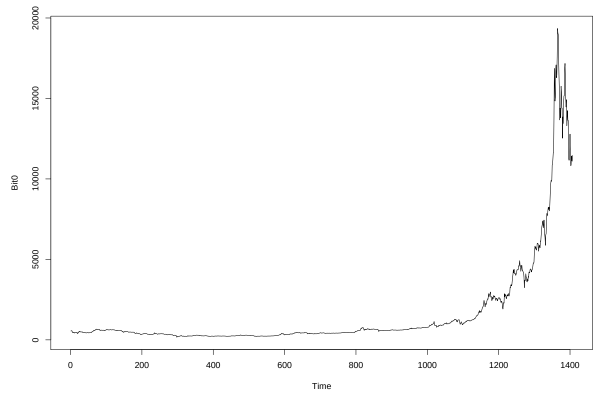
   (i) Provide forecasting of X_{N+1} and X_{N+1}^2.

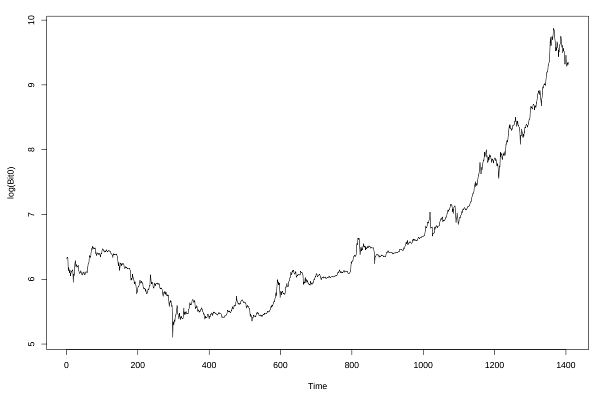
3. **Numerical part:** We study with `R` software the open daily historical data of Bitcoin from January 28 2014 to January 28 2018.

   (a) First the following commands have been executed with figures exhibited below:

```
Bit=read.csv("C:/Users/Admin/Dropbox/Enseignement/M2 MO/TP/BTC-USD.csv")
Bit0=Bit$Open; n=length(Bit0)
plot.ts(Bit0); plot.ts(log(Bit0))
Y=log(Bit0); X1=c(1:n); X2=X1^2
Y.lm=lm(Y~X1+X2); summary(Y.lm)
```

   Command `lm` realizes a least squares linear regression. Here there are the graphs and main numerical results:





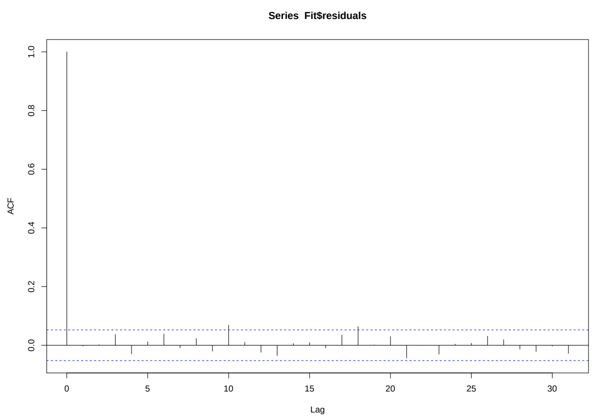

```
Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  6.541e+00  1.645e-02  397.61   <2e-16 ***
X1          -4.034e-03  5.396e-05  -74.76   <2e-16 ***
X2           4.332e-06  3.711e-08  116.73   <2e-16 ***

Residual standard error: 0.2054 on 1404 degrees of freedom
Multiple R-squared:  0.9635,Adjusted R-squared:  0.9635
F-statistic: 1.854e+04 on 2 and 1404 DF,  p-value: < 2.2e-16
```

   Question II.1: Explain what is done.

   (b) New commands are then executed:

```
plot.ts(Y.lm$residuals)
Fit=arima(Y.lm$residuals, order = c(1,0,2))
acf(Fit$residuals)
Box.test(Fit$residuals, lag = 20,"Ljung-Box", fitdf=3)
```

   Here there are graphs and numerical results:




```
Box-Ljung test
data:  Fit$residuals
X-squared = 25.557, df = 17, p-value = 0.08292
```

Question II.2: Explain what is done (notably why we use `fitdf=3`*) and explain which conclusions you deduce.*

(c) The following sequence of commands is then executed:

```
pred=predict(Fit,n.ahead=1); pred[1]
exp(pred$pred[1]+sum(Y.lm$coeff*c(1,(n+1),(n+1)^2)))
```

The results are the following:

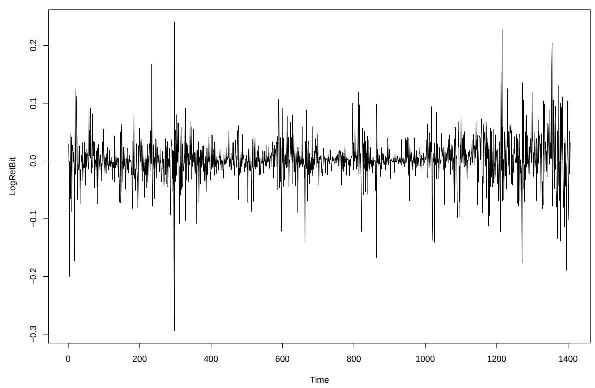
```
>-0.09092553
> 11582.19
```

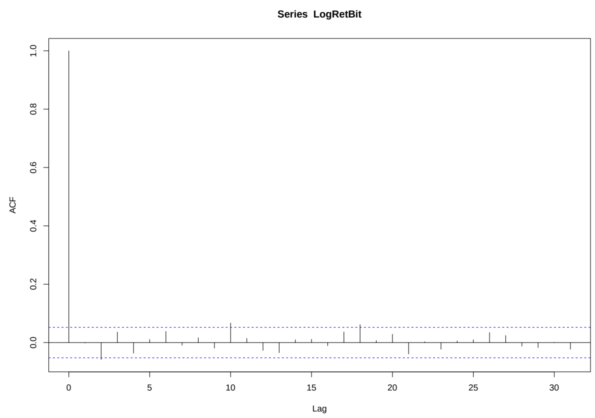
Question II.3: What is done here and what are your conclusions?

(d) Finally, the following sequence of commands is executed:

```
LogRetBit=log(Bit0[2:n]/Bit0[1:(n-1)])
plot.ts(LogRetBit); acf(LogRetBit)
library(fGarch)
FitLogRet1=garchFit(~garch(1,1),data=LogRetBit,trace=FALSE)
summary(FitLogRet1)
FitLogRet2=garchFit(~garch(1,2),data=LogRetBit,trace=FALSE)
summary(FitLogRet2)
```

The graphs and results are the following:






```
> Error Analysis:
        Estimate  Std. Error  t value Pr(>|t|)
mu      1.646e-03   7.012e-04    2.347 0.018911 *
omega   2.894e-05   7.565e-06    3.825 0.000131 ***
alpha1  1.690e-01   2.155e-02    7.845 4.44e-15 ***
beta1   8.365e-01   1.774e-02   47.139  < 2e-16 ***

Standardised Residuals Tests:
                                Statistic p-Value
 Jarque-Bera Test   R    Chi^2  2253.533  0
 Shapiro-Wilk Test  R    W      0.9142998 0
 Ljung-Box Test     R    Q(10)  31.40408  0.0005030119
 Ljung-Box Test     R    Q(15)  34.67742  0.002732705
 Ljung-Box Test     R    Q(20)  42.10945  0.002676158
 Ljung-Box Test     R^2  Q(10)  7.152754  0.7109496
 Ljung-Box Test     R^2  Q(15)  11.73218  0.6991771
 Ljung-Box Test     R^2  Q(20)  16.21253  0.7033551
 LM Arch Test       R    TR^2   9.629902  0.648393

Information Criterion Statistics:
      AIC       BIC       SIC      HQIC
-3.971264 -3.956332 -3.971280 -3.965683

> Error Analysis:
```

```
       Estimate  Std. Error  t value Pr(>|t|)
mu     1.711e-03   7.038e-04    2.431 0.015041 *
omega  3.529e-05   1.068e-05    3.305 0.000949 ***
alpha1 2.098e-01   3.277e-02    6.401 1.54e-10 ***
beta1  4.464e-01   1.945e-01    2.295 0.021745 *
beta2  3.485e-01   1.755e-01    1.985 0.047127 *

Standardised Residuals Tests:
                                Statistic p-Value
 Jarque-Bera Test   R    Chi^2  2360.269  0
 Shapiro-Wilk Test  R    W      0.9149907 0
 Ljung-Box Test     R    Q(10)  30.51412  0.000705495
 Ljung-Box Test     R    Q(15)  33.53633  0.003954055
 Ljung-Box Test     R    Q(20)  40.94799  0.003782937
 Ljung-Box Test     R^2  Q(10)  5.451614  0.8590436
 Ljung-Box Test     R^2  Q(15)  8.699528  0.8926969
 Ljung-Box Test     R^2  Q(20)  13.33315  0.8626366
 LM Arch Test       R    TR^2   7.509703  0.8221771

Information Criterion Statistics:
      AIC       BIC       SIC      HQIC
-3.974214 -3.955549 -3.974239 -3.967238
```

Question II.4: What is done here and which model could you chose?