Université Paris VII and I

Master 2 M0 2019 – 2020

Analyse des séries financières

Examen final, Février 2020

3h00, sans aucun document

1. Soit $(\varepsilon_t)_{t\in\mathbb{Z}}$ une suite de variables aléatoires centrées indépendantes et identiquement distribuées telles que $\mathbb{E}\big[|\varepsilon_0|^r\big] < \infty$. On s'intéresse à la suite de variables $(X_t)_{t\in\mathbb{Z}}$ définies par récurrence:

$$X_t = \varepsilon_t\,\sigma_t \quad \text{avec} \quad \sigma_t = a_0^* + a_1^*\,X_{t-1} \quad \text{pour tout } t \in \mathbb{Z} \tag{1}$$

avec $(a_0^*, a_1^*) \in]0,\infty[\times\mathbb{R}$ une vecteur de paramètres inconnus. Ce modèle lorsqu'il existe est appelé modèle LARCH(1).

(a) On considère $(Y_t)_{t\in\mathbb{Z}}$ un bruit blanc non nul ayant une loi symétrique (Y_0 et $-Y_0$ ont la même loi) et $(Z_t)_{t\in\mathbb{Z}}$ tel que $Z_t = (-1)^t Y_t$ pour tout $t \in \mathbb{Z}$. Montrer que la somme ou le produit de deux séries stationnaires n'est pas nécessairement une série stationnaire.

(b) On suppose que $(\sigma_t)_{t\in\mathbb{Z}}$ est une série telle que $\sigma_t = F\big((\varepsilon_{t-k})_{k\geq 1}\big)$ pour tout $t \in \mathbb{Z}$, où $F : \mathbb{R}^\infty \to \mathbb{R}$ est une fonction mesurable vérifiant $F\big((\varepsilon_{t-k})_{k\geq 1}\big)$ existe presque sûrement pour tout $t \in \mathbb{Z}$. Montrer que $(\sigma_t)_{t\in\mathbb{Z}}$ est une série stationnaire. En déduire qu'alors $(X_t)_{t\in\mathbb{Z}}$ est une série stationnaire.

(c) Montrer que pour tout $n \in \mathbb{N}^*$ et tout $t \in \mathbb{Z}$,

$$\sigma_t = a_0^*\Big(1 + \sum_{k=1}^{n}\prod_{j=1}^{k} a_1^*\,\varepsilon_{t-j}\Big) + \sigma_{t-n-1}\prod_{j=1}^{n+1} a_1^*\,\varepsilon_{t-j}.$$

(d) Pour une variable aléatoire Z et $r \geq 1$, on note $\|Z\|_r = \big(\mathbb{E}\big[|Z|^r\big]\big)^{1/r}$ et $\mathbb{L}^r$ l'espace vectoriel associé à cette norme. Montrer que si $|a_1^*|\,\big\|\varepsilon_0\big\|_r < 1$ alors

$$a_0^*\Big(1 + \sum_{k=1}^{n}\prod_{j=1}^{k} a_1^*\,\varepsilon_{t-j}\Big) \quad \text{converge dans } \mathbb{L}^r.$$

En déduire que $X_t = a_0^*\,\varepsilon_t\Big(1 + \sum_{k=1}^{\infty}\prod_{j=1}^{k} a_1^*\,\varepsilon_{t-j}\Big)$ est une solution stationnaire de (1) telle que $\|X_0\|_r < \infty$. Montrer que cette solution est unique.

(e) Montrer que la condition d'existence d'une solution stationnaire ayant un moment d'ordre r pouvait être obtenue par une autre méthode bien plus rapide.

(f) Pour $r \geq 2$, déterminer l'espérance, la variance et la covariance de cette solution. De même donner son espérance et sa variance conditionnelles. Quelles sont les différences avec un ARCH(1)? Est-ce appropriée pour des données financières?

(g) On suppose désormais que $\mathrm{var}\,(\varepsilon_0)=1$ et on désire estimer $\theta^*=(a_0^*,a_1^*)$. Expliquer pourquoi la méthode d'estimation par maximum de quasi-vraisemblance conditionnelle gaussienne ne peut pas être utilisée.

(h) Pour tout $\theta={}^t(a_0,a_1)\in\Theta(r)=]0,\infty[\times\big]-\frac{1}{\|\varepsilon_0\|_r},\frac{1}{\|\varepsilon_0\|_r}\big[$ et $t\in\mathbb{Z}$, on pose

$$M_\theta^t=\big(X_t^2-(a_0+a_1\,X_{t-1})^2\big)^2.$$

En supposant désormais que $r\geq 4$ et $\theta^*\in\Theta(r)$, montrer que pour tout $\theta\in\Theta(r)$ et $t\in\mathbb{Z}$,

$$\mathbb{E}\left[M_\theta^t\right]=\mathrm{var}\,(\varepsilon_0^2)\,\mathbb{E}\left[(a_0^*+a_1^*X_{t-1})^4\right]+\mathbb{E}\left[\big((a_0^*+a_1^*X_{t-1})^2-(a_0+a_1X_{t-1})^2\big)^2\right].$$

En déduire que $\mathbb{E}\left[M_{\theta^*}^t\right]<\mathbb{E}\left[M_\theta^t\right]$ pour tout $\theta\in\Theta(r)\setminus\{\theta^*\}$.

(i) Par la suite, on admet que si $(v_t(\phi))_{t\in\mathbb{Z}}$ est une suite de variables aléatoires telle que $v_t(\phi)=F_\phi\big((\varepsilon_{t-k})_{k\geq 0}\big)$ avec pour tout $\phi\in\Phi\subset\mathbb{R}^d$, $F_\phi:\mathbb{R}^\infty\to\mathbb{R}$ et $\mathbb{E}\left[\sup_{\phi\in\Phi}\big|v_0(\phi)\big|\right]<\infty$, alors

$$\sup_{\phi\in\Phi}\Big|\frac{1}{n}\sum_{t=1}^n v_t(\phi)-\mathbb{E}\left[v_0(\phi)\right]\Big|\xrightarrow[n\to+\infty]{p.s.}0.$$

On note $\widehat{C}_n(\theta)=\frac{1}{n-1}\sum_{t=2}^n M_\theta^t$. Montrer que pour tout ensemble compact $\Theta\subset\Theta(4)$,

$$\sup_{\theta\in\Theta}\Big|\widehat{C}_n(\theta)-\mathbb{E}\left[M_\theta^0\right]\Big|\xrightarrow[n\to+\infty]{p.s.}0.$$

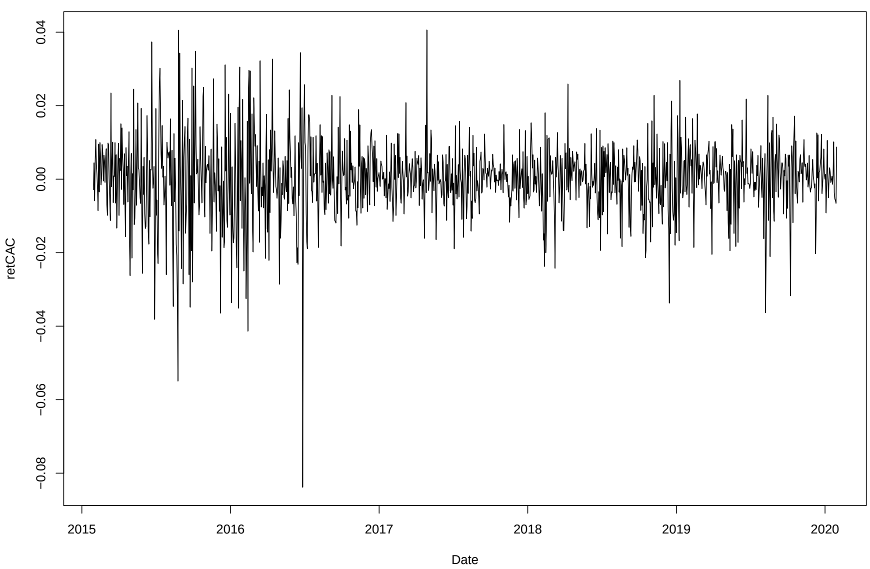
En déduire que $\widehat{\theta}_n=\operatorname*{Argmin}_{\theta\in\Theta}\widehat{C}_n(\theta)\xrightarrow[n\to+\infty]{p.s.}\theta^*$.

(j) Déterminer un prédicteur pour X_{n+1} and X_{n+1}^2 lorsque la trajectoire $(X_1,\cdots,X_n)$ a été observée.

2. On s'intéresse à modéliser la cotation quotidienne en cloture du CAC40 depuis le 27 janvier 2015. On utilise le logiciel `R` à cet effet.

(a) Voici les premières commandes effectuées:

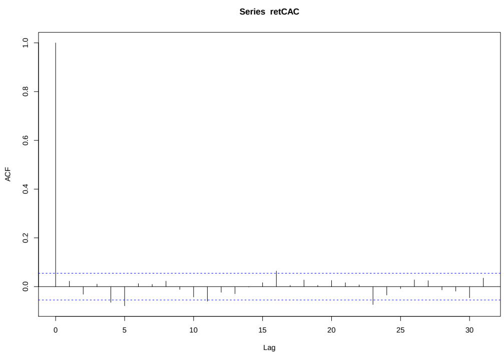
```
CACc=read_excel("CAC.xlsx")
n=length(CACc)
retCAC=log(CACc[2:n])-log(CACc[1:(n-1)])
Date=2015+27/365+c(1:(n-1))/255
plot(Date,retCAC,"l")
```

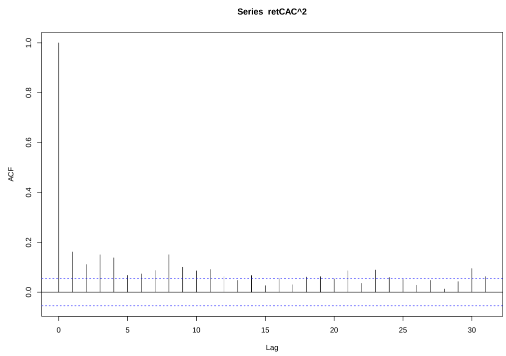


Question II.1: Expliquer ce qui a été fait? Le graphe vous semble-t-il mettre en évidence une hétéroscédasticité conditionnelle?

(b) On exécute alors:

```
acf(retCAC)
Box.test(retCAC,lag=9,type=c("Ljung-Box"))
acf(retCAC^2)
Box.test(retCAC^2,lag=9,type=c("Ljung-Box"))
```





Avec pour résultats numériques:

```
Box-Ljung test

data:  retCAC
X-squared = 16.484, df = 9, p-value = 0.05743

Box-Ljung test

data:  retCAC^2
X-squared = 167.04, df = 9, p-value < 2.2e-16
```

Question II.2: Expliquer ce qui a été fait (qu'est-ce qu'effectuent les commandes `ACF` *et quelles hypothèses testent les commandes* `Box.test`*?) et ce que l'on a obtenu. Que montre le second test de Ljung-Box?*

(c) On continue alors par:

```
library(forecast)
fit=auto.arima(retCAC,max.p=5,max.q=5,ic="bic")
fit
```

On obtient alors:

```
Series: retCAC
ARIMA(0,0,0) with zero mean

sigma^2 estimated as 0.0001089:  log likelihood=4010.98
AIC=-8019.95   AICc=-8019.95   BIC=-8014.8
```

Question II.3: Qu'est-ce qui est fait et quelles sont vos conclusions?

```
library(fGarch)
FitRetCAC0=garchFit(~garch(1,0),data=retCAC,trace=FALSE)
BIC0=FitRetCAC0@fit$ics[2]; BIC0
FitRetCAC1=garchFit(~garch(1,1),data=retCAC,trace=FALSE)
BIC1=FitRetCAC1@fit$ics[2]; BIC1
FitRetCAC2=garchFit(~garch(2,1),data=retCAC,trace=FALSE)
BIC2=FitRetCAC2@fit$ics[2]; BIC2
```

Et les réultats:

```
>  BIC0
-6.362328
>  BIC1
-6.494399
> BIC2
-6.488426
```

Question II.4: Qu'est-ce qui a été fait et quelles sont vos conclusions?

(d) Voici les dernières commandes tapées:

```
Cn=function(the){sum((retCAC[2:(n-1)]^2-(the[1]+the[2]*retCAC[1:(n-2)])^2)^2)}
ECn=optim(c(0.1,-0.1),Cn,method = "L-BFGS-B",lower = c(0,-0.99999), upper = c(100,0.999))
(a0=ECn$par[1]); (a1=ECn$par[2])
Ln=-0.5*(sum(retCAC[2:(n-1)]^2/(a0+a1*retCAC[1:(n-2)])^2+log((a0+a1*retCAC[1:(n-2)])^2))+(n-2)*log(2*pi))
omega=FitRetCAC0@fit$par[2]; alpha1=FitRetCAC0@fit$par[3]
An=-0.5*(sum(retCAC[2:(n-1)]^2/(omega+alpha1*retCAC[1:(n-2)]^2)+log((omega+alpha1*retCAC[1:(n-2)]^2)))+(n-2)*log(2*pi)
Ln; An
```

Avec pour résultat:

```
> (a0=ECn$par[1]); (a1=ECn$par[2])
[1] 0.01033321
[1] -0.098922
> Ln; An
[1] 3998.841
[1] 4065.412
```

Question II.5: Qu'est-ce qui a été fait et quelles sont vos conclusions?