

Université Paris VII and I

Master 2 M0 2019 – 2020

# Analyse des séries financières

Correction de l'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  $\mathbb{E} [|\varepsilon_0|^r] < \infty$ . On s'intéresse à la suite de variables  $(X_t)_{t \in \mathbb{Z}}$  définies par récurrence:

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

avec  $(a_0^*, a_1^*) \in ]0, \infty[ \times \mathbb{R}$  un 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  $\sigma_t = F((\varepsilon_{t-k})_{k \geq 1})$  pour tout  $t \in \mathbb{Z}$ , où  $F : \mathbb{R}^\infty \rightarrow \mathbb{R}$  est une fonction mesurable vérifiant  $F((\varepsilon_{t-k})_{k \geq 1})$  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  $t \in \mathbb{Z}$ ,

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

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

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

En déduire que  $X_t = a_0^* \varepsilon_t \left( 1 + \sum_{k=1}^{\infty} \prod_{j=1}^k a_1^* \varepsilon_{t-j} \right)$  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  $\text{var}(\varepsilon_0) = 1$  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  $\theta = {}^t(a_0, a_1) \in \Theta(r) = ]0, \infty[ \times ] -\frac{1}{\|\varepsilon_0\|_r}, \frac{1}{\|\varepsilon_0\|_r}[$  et  $t \in \mathbb{Z}$ , on pose

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

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

$$\mathbb{E}[M_\theta^t] = \text{var}(\varepsilon_0^2) \mathbb{E}[(a_0^* + a_1^* X_{t-1})^4] + \mathbb{E}[(a_0^* + a_1^* X_{t-1})^2 - (a_0 + a_1 X_{t-1})^2]^2.$$

En déduire que  $\mathbb{E}[M_{\theta^*}^t] < \mathbb{E}[M_\theta^t]$  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  $v_t(\phi) = F_\phi((\varepsilon_{t-k})_{k \geq 0})$  avec pour tout  $\phi \in \Phi \subset \mathbb{R}^d$ ,  $F_\phi : \mathbb{R}^\infty \rightarrow \mathbb{R}$  et  $\mathbb{E}[\sup_{\phi \in \Phi} |v_0(\phi)|] < \infty$ , alors

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

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

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

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

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

*Proof.* (a) Pour tout  $t \in \mathbb{Z}$  et  $u \in \mathbb{R}$ ,  $\mathbb{E}[e^{iuZ_t}] = \mathbb{E}[e^{iu(-1)^t Y_t}] = \mathbb{E}[e^{iuY_t}]$  car  $Y_t$  a une loi symétrique, donc  $Z_t$  a la même loi que  $Y_t$  pour tout  $t \in \mathbb{Z}$ . Ainsi  $(Z_t)_{t \in \mathbb{Z}}$  est aussi un bruit blanc à loi symétrique. Mais pour tout  $t \in \mathbb{Z}$ ,  $Y_t + Z_t = (1 + (-1)^t)Y_t = 0$  si  $t$  impair et  $2Y_t$  si  $t$  pair: la loi de  $(Y_t + Z_t)$  n'est pas la même suivant  $t$ , donc  $(Y_t + Z_t)$  n'est pas stationnaire. De même pour  $Y_t Z_t = (-1)^t Y_t^2$ , qui est positive pour  $t$  pair et négative sinon. Comme un bruit blanc est une série stationnaire on en déduit que la somme ou le produit de 2 séries stationnaires n'est pas nécessairement stationnaire.

- (b) On pose  $m \in \mathbb{N}^*$  et  $\sigma_t^{(m)} = F(\varepsilon_{t-1}, \dots, \varepsilon_{t-m}, 0, \dots)$ . Il est clair que comme  $(\varepsilon_t)$  est un bruit blanc, donc une série stationnaire, pour tout  $k \in \mathbb{N}^*$ ,  $(t_1, \dots, t_k) \in \mathbb{R}^k$  et tout  $c \in \mathbb{Z}$ , on a  $x = (\varepsilon_{t_1-1}, \dots, \varepsilon_{t_1-m}, \varepsilon_{t_2-1}, \dots, \varepsilon_{t_k-m})$  a la même loi que  $y = (\varepsilon_{t_1+c-1}, \dots, \varepsilon_{t_1+c-m}, \varepsilon_{t_2+c-1}, \dots, \varepsilon_{t_k+c-m})$ . Ainsi pour la fonction  $g((x_i)_{1 \leq i \leq km}) = (F((x_i)_{1 \leq i \leq m}, 0, \dots), F((x_i)_{m+1 \leq i \leq 2m}, 0, \dots), \dots, F((x_i)_{(k-1)m+1 \leq i \leq km}, 0, \dots))$  alors  $g(x)$  a la même loi que  $g(y)$ : on en déduit que  $(\sigma_t^{(m)})_{t \in \mathbb{Z}}$  est une série stationnaire.

Il ne reste plus qu'à considérer la limite quand  $m \rightarrow \infty$ , ce qui a été déjà fait dans le cours, et ainsi  $(\sigma_t)$  est bien une série stationnaire.

Pour montrer que  $(X_t)$  il suffit reprendre la preuve précédente avec une fonction  $G((\varepsilon_{t-j})_{j \geq 0}) = \varepsilon_t F((\varepsilon_{t-j})_{j \geq 1})$  qui existe également presque sûrement.

- (c) On sait que  $\sigma_t = a_0^* + a_1^* X_{t-1} = a_0^* + a_1^* \varepsilon_{t-1} \sigma_{t-1}$  pour tout  $t \in \mathbb{Z}$ . On montre la propriété par récurrence. Elle est vraie au rang  $n = 1$ , puisque  $\sigma_t = a_0^* + a_1^* \varepsilon_{t-1} \sigma_{t-1} = a_0^* + a_1^* \varepsilon_{t-1} (a_0^* + a_1^* \varepsilon_{t-2} \sigma_{t-2}) = a_0^* + a_0^* a_1^* \varepsilon_{t-1} + (a_1^*)^2 \varepsilon_{t-1} \varepsilon_{t-2} \sigma_{t-2}$ . Supposons qu'elle soit vraie au rang  $n$ . Alors:

$$\begin{aligned} \sigma_t &= a_0^* \left( 1 + \sum_{k=1}^n \prod_{j=1}^k a_1^* \varepsilon_{t-j} \right) + (a_0^* + a_1^* \varepsilon_{t-n-2} \sigma_{t-n-2}) \prod_{j=1}^{n+1} a_1^* \varepsilon_{t-j} \\ &= a_0^* \left( 1 + \sum_{k=1}^{n+1} \prod_{j=1}^k a_1^* \varepsilon_{t-j} \right) + a_1^* \sigma_{t-n-2} \prod_{j=1}^{n+2} a_1^* \varepsilon_{t-j}. \end{aligned}$$

Elle est donc vraie au rang  $n + 1$ , donc vraie pour tout  $n \in \mathbb{N}^*$ .

- (d) Soit  $Z_t^{(n)} = a_0^* \left(1 + \sum_{k=1}^n \prod_{j=1}^k a_1^* \varepsilon_{t-j}\right)$ . Pour montrer que  $(Z_n)$  admet bien une limite dans  $\mathbb{L}^r$  qui ne peut être que  $Z_t^{(\infty)} = a_0^* \left(1 + \sum_{k=1}^{\infty} \prod_{j=1}^k a_1^* \varepsilon_{t-j}\right)$ , il suffit de montrer que  $(Z_n)$  est bien une suite de Cauchy dans  $\mathbb{L}^r$ . Pour  $1 \leq p < q$ , on a :

$$\begin{aligned} \|Z_t^{(q)} - Z_t^{(p)}\|_r &= \left\| a_0^* \sum_{k=p+1}^q \prod_{j=1}^k a_1^* \varepsilon_{t-j} \right\|_r \\ &\leq |a_0^*| \sum_{k=p+1}^q |a_1^*|^k \left\| \prod_{j=1}^k \varepsilon_{t-j} \right\|_r \end{aligned}$$

par l'inégalité triangulaire (valable dès que  $r \geq 1$ ). Les  $(\varepsilon_{t-j})$  étant indépendantes, on a  $\left\| \prod_{j=1}^k \varepsilon_{t-j} \right\|_r = \prod_{j=1}^k \|\varepsilon_{t-j}\|_r$ , d'où

$$\|Z_t^{(q)} - Z_t^{(p)}\|_r \leq |a_0^*| \sum_{k=p+1}^q (|a_1^*| \|\varepsilon_0\|_r)^k.$$

Donc si  $|a_1^*| \|\varepsilon_0\|_r < 1$ , on a bien  $\|Z_t^{(q)} - Z_t^{(p)}\|_r \rightarrow 0$  quand  $p \rightarrow \infty$  : la suite est bien de Cauchy, donc elle converge bien dans  $\mathbb{L}^r$ .

Si on considère  $X_t = \varepsilon_t Z_t^{(\infty)}$ , alors  $(X_t)$  est bien stationnaire d'après la question (b). De plus, il est clair que cette solution vérifie l'équation (1) puisque  $\varepsilon_t Z_t^{(\infty)} = \varepsilon_t (a_0^* + a_1^* \varepsilon_{t-1} Z_{t-1}^{(\infty)})$ . Et on a bien  $\|\varepsilon_t Z_t^{(\infty)}\|_r = \|\varepsilon_t\|_r \|\varepsilon_t Z_t^{(\infty)}\|_r < \infty$  d'après ce qui précède.

Enfin, si on suppose une autre solution de (1) avec  $X_t = \varepsilon_t \sigma'_t$  avec  $\sigma'_t \neq Z_t^{(\infty)}$ , alors ces deux suites vérifient  $\sigma_t = a_0^* \left(1 + \sum_{k=1}^n \prod_{j=1}^k a_1^* \varepsilon_{t-j}\right) + \sigma_{t-n-1} \prod_{j=1}^{n+1} a_1^* \varepsilon_{t-j}$ , d'où

$$\sigma'_t - \sigma'_{t-n-1} \prod_{j=1}^{n+1} a_1^* \varepsilon_{t-j} = Z_t^{(\infty)} - Z_{t-n-1}^{(\infty)} \prod_{j=1}^{n+1} a_1^* \varepsilon_{t-j},$$

pour tout  $n \in \mathbb{N}^*$ . Si  $n$  tend vers l'infini, alors les produits tendent vers 0 dans  $\mathbb{L}^r$  et ainsi nécessairement  $\sigma'_t = Z_t^{(\infty)}$  pour tout  $t \in \mathbb{Z}$  : la solution est unique.

- (e)  $(X_t)$  est un processus affine causal et la fonction  $M((X_{t-k})_{k \geq 1}) = a_0^* + a_1^* X_{t-1}$  et  $\frac{\partial}{\partial x_1} M((x_k)_{k \geq 1}) = a_1^*$  et on trouve ainsi la même condition de stationnarité pour avoir une solution dans  $\mathbb{L}^r$ .
- (f) On a  $\mathbb{E}[X_t] = 0$  et  $\text{var}(X_t) = \text{var}(\varepsilon_0)((a_0^*)^2 + (a_1^*)^2 \text{var}(X_t))$ , donc  $\text{var}(X_t) = (a_0^*)^2 \text{var}(\varepsilon_0) (1 - (a_1^*)^2 \text{var}(\varepsilon_0))^{-1}$  (on utilise l'indépendance de  $\varepsilon_t$  et  $X_{t-1}$ ). Enfin, comme pour un GARCH,  $\text{cov}(X_t, X_{t+k}) = 0$  si  $k \geq 1$ . Pour les moments conditionnels, on a  $\mathbb{E}[X_t | X_{t-1}, \dots] = 0$  et  $\text{var}(X_t | X_{t-1}, \dots) = \mathbb{E}[X_t^2 | X_{t-1}, \dots] = \text{var}(\varepsilon_0)(a_0^* + a_1^* X_{t-1})^2$ . Comme un ARCH(1), le processus LARCH(1) est un processus conditionnellement hétéroscédastique. Mais à la différence d'un ARCH(1) pour lequel la variance conditionnelle est toujours plus grande qu'une constante strictement positive, la variance conditionnelle (et donc la volatilité instantannée) peut s'annuler.
- (g) La dernière propriété rend difficile l'utilisation du QMLE gaussien, puisque la variance conditionnelle peut s'annuler.
- (h) On a :

$$\begin{aligned} \mathbb{E}[M_\theta^t] &= \mathbb{E} \left[ ((\varepsilon_t^2 - 1)(a_0^* + a_1^* X_{t-1})^2 + (a_0^* + a_1^* X_{t-1})^2 - (a_0 + a_1 X_{t-1})^2)^2 \right] \\ &= \mathbb{E} \left[ (\varepsilon_t^2 - 1)^2 (a_0^* + a_1^* X_{t-1})^4 \right] + \mathbb{E} \left[ ((a_0^* + a_1^* X_{t-1})^2 - (a_0 + a_1 X_{t-1})^2)^2 \right] \\ &\quad + 2 \mathbb{E} \left[ (\varepsilon_t^2 - 1)(a_0^* + a_1^* X_{t-1})^2 ((a_0^* + a_1^* X_{t-1})^2 - (a_0 + a_1 X_{t-1})^2) \right] \\ &= \text{var}(\varepsilon_0^2) \mathbb{E} \left[ (a_0^* + a_1^* X_{t-1})^4 \right] + \mathbb{E} \left[ ((a_0^* + a_1^* X_{t-1})^2 - (a_0 + a_1 X_{t-1})^2)^2 \right], \end{aligned}$$

car  $\varepsilon_t$  est indépendante de  $X_{t-1}$  et  $\mathbb{E}(\varepsilon_t^2) = 1$ .

On a  $\mathbb{E} \left[ ((a_0^* + a_1^* X_{t-1})^2 - (a_0 + a_1 X_{t-1})^2)^2 \right] \geq 0$  comme carré et = 0 pour  $\theta = \theta^*$ . En développant on voit que dès que  $a_0 \neq a_0^*$  ou  $a_1 \neq a_1^*$ , alors ce terme devient  $> 0$ . On en déduit le résultat.

- (i) Comme  $X_t$  et  $X_{t-1}$  s'écrivent comme des fonctions mesurables de  $(\varepsilon_{t-k})_{k \geq 0}$  et de  $(\varepsilon_{t-k})_{k \geq 1}$ , alors il en est de même pour  $M_\theta^t$ . Il suffit donc de démontrer que  $\mathbb{E}[\sup_{\theta \in \Theta} |M_\theta^0|] < \infty$ . On sait que  $\mathbb{E}[X_t^4] < \infty$  pour

tout  $t \in \mathbb{Z}$  car on a supposé que  $\Theta \subset \Theta(4)$ . Or,

$$\begin{aligned} \sup_{\theta \in \Theta} |M_\theta^0| &\leq 2 \sup_{\theta \in \Theta} (X_t^4 + (a_0 + a_1 X_{t-1})^4) \\ &\leq 2 \sup_{\theta \in \Theta} (X_t^4 + 4a_0^4 + 4a_1^4 X_{t-1}^4) \\ &\leq 2X_t^4 + 8 \sup_{\theta \in \Theta} a_0^4 + 8X_{t-1}^4 \sup_{\theta \in \Theta} a_1^4. \end{aligned}$$

Comme  $\Theta$  est un compact, les 2 suprema existent et on en déduit que  $\mathbb{E} [\sup_{\theta \in \Theta} |M_\theta^0|] < \infty$  ainsi que le résultat demandé.

D'après le résultat précédent, pour tout  $\eta > 0$ , il existe  $n_0(\eta)$  tel que pour  $n \geq n_0(\eta)$ ,

$$|\widehat{C}_n(\theta) - \mathbb{E}[M_\theta^0]| \leq \varepsilon \quad \text{p.s. pour tout } \theta \in \Theta.$$

Ainsi pour  $n \geq n_0(\eta)$ ,  $\min_{\theta \in \Theta} \mathbb{E}[M_\theta^0] - \eta \leq \min_{\theta \in \Theta} \widehat{C}_n(\theta) \leq \min_{\theta \in \Theta} \mathbb{E}[M_\theta^0] + \eta$ . Comme la fonction  $\theta \in \Theta \mapsto \widehat{C}_n(\theta)$  est p.s. continue, que  $\min_{\theta \in \Theta} \mathbb{E}[M_\theta^0]$  est obtenu de manière unique en  $\theta^*$ , alors  $\min_{\theta \in \Theta} \widehat{C}_n(\theta)$  est obtenu en  $\widehat{\theta}_n$  qui est d'autant plus proche de  $\theta^*$  que  $\eta$  est petit. D'où le résultat  $\widehat{\theta}_n = \underset{\theta \in \Theta}{\text{Argmin}} \widehat{C}_n(\theta) \xrightarrow[n \rightarrow +\infty]{\text{p.s.}} \theta^*$ .

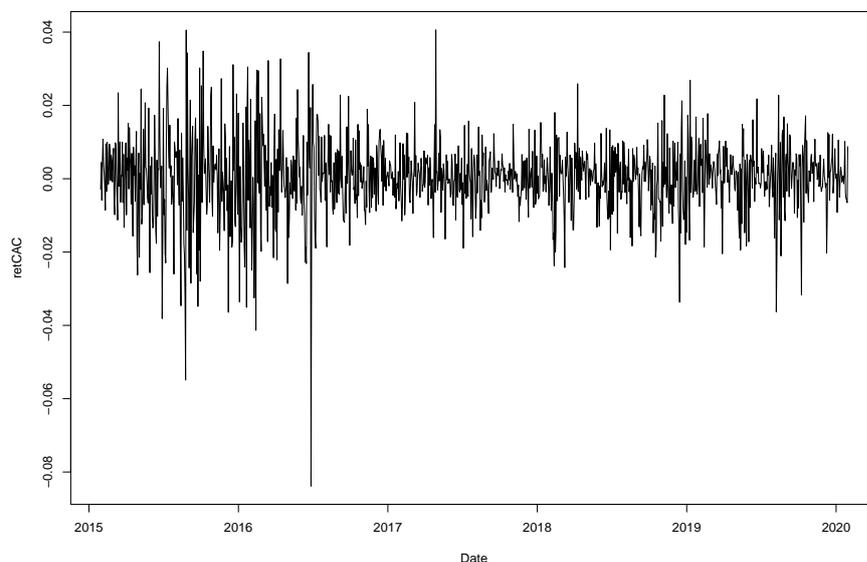
- (j) Il est clair que comme pour l'ARCH(1) on a  $\mathbb{E}[X_{n+1} | (X_1, \dots, X_n)] = 0$ . On en déduit que  $\widehat{X}_{n+1} = 0$ . On a également  $\mathbb{E}[X_{n+1}^2 | (X_1, \dots, X_n)] = \mathbb{E}[\varepsilon_{n+1}^2 (a_0^* + a_1^* X_n)^2 | X_n] = (a_0^* + a_1^* X_n)^2$ . On en déduit que  $\widehat{X}_{n+1}^2 = (\widehat{a}_0 + \widehat{a}_1 X_n)^2$ .

□

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")
```

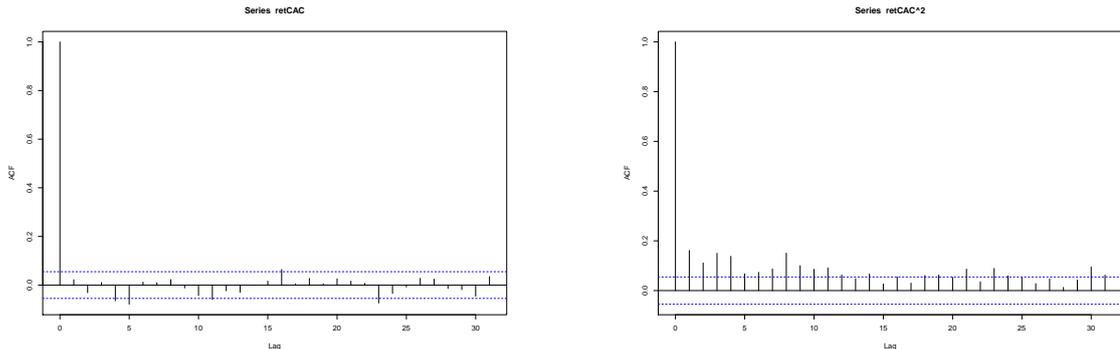


*Question II.1: Expliquer ce qui a été fait? Le graphe vous semble-t-il mettre en évidence une hétéroscédasticité conditionnelle?*

*Proof.* On a calculé les log-rendements du CAC40 pendant toute cette période, puis tracé en fonction du temps. On s'aperçoit de périodes pendant lesquelles les fluctuations sont fortes et d'autres pendant lesquelles elles sont faibles: cela met en évidence une hétéroscédasticité conditionnelle.  $\square$

(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?*

*Proof.* On a tracé les corrélogrammes des log-rendements et des log-rendements au carré, c'est-à-dire l'ensemble des autocorrélations de ces séries pour des lags allant de 0 à 31. Dans le premier cas, presque toutes ces autocorrélations sont dans l'intervalle de confiance à 95%, plusieurs sont en dehors en ce qui concerne les autocorrélations des carrés.

Les commandes `Box.test` proposent des tests portemanteaux de type Ljung-Box sur chacune des 2 séries, pour tester:  $H_0$ : la série est une suite de v.a.i.i.d., contre  $H_1$ : elle ne l'est pas. Si le test accepte (de peu)  $H_0$  pour les log-rendements, celle-ci est très nettement rejetée pour les carrés de ces log-rendements. On en conclut que si les log-rendements peuvent apparaître comme un bruit blanc, ce ne peut être qu'un bruit blanc faible (s'il était fort le test ne serait pas rejeté pour les carrés).  $\square$

(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?*

*Proof.* En choisissant comme critère le BIC, on a choisi le meilleur modèle ARMA( $p, q$ ) où  $0 \leq p, q \leq 5$ , pour représenter les log-rendements. Et finalement ce critère choisit  $p = 0$  et  $q = 0$ , donc plutôt un bruit blanc qu'un modèle ARMA. On n'est guère surpris car les modèles ARMA ne sont pas apte à modéliser des processus conditionnellement hétéroscédastiques.  $\square$

(d) Voici les commandes suivantes:

```
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 résultats:

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

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

*Proof.* Toujours en utilisant le critère BIC, on cherche parmi 3 modèles GARCH celui qui modéliserait le mieux les log-rendements. Et on choisirait plutôt le GARCH(1, 1).  $\square$

(e) 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?*

*Proof.* Cette fois-ci, on recalcule à la main la log-vraisemblance conditionnelle gaussienne pour le modèle ARCH(1) et pour le modèle LARCH(1) défini dans la partie théorique. Ce sont deux modèles qui ont même nombre de paramètres estimés (2) et la log-vraisemblance suffit à choisir entre les 2 celui qui aurait la meilleure adéquation. Et clairement on choisira plutôt le modèle ARCH(1) que le modèle LARCH(1).  $\square$