

## EXERCICE 1

- Soit le processus stochastique suivant:  $Y_t = \sum_{j=1}^{\infty} \psi_j \varepsilon_{t-j}$  où  $\psi_j = \frac{(-1)^j}{j}$  et  $\{\varepsilon_t\}$  est un bruit blanc de variance  $\text{IE}[\varepsilon_t^2] = \sigma_\varepsilon^2$ .

- a) Montrer que  $\{Y_t\}$  est stationnaire et que la fonction d'auto-covariance  $\gamma_h$  est:  $\gamma_h = (-1)^h \frac{\sigma_\varepsilon^2}{h} \sum_{j=1}^h \frac{1}{j}$ .

PARTIE 1: STATIONNARITÉ DU PROCESSUS  $\{Y_t\}$ 

- 1)  $\text{IE}[Y_t^2] < \infty, \forall t \in \mathbb{T} \Leftrightarrow$  la variance est finie:

$$\begin{aligned} V_{\text{var}}(Y_t) &= V_{\text{var}}\left(\sum_{j=1}^{\infty} \psi_j \varepsilon_{t-j}\right) \\ &= \sum_{j=1}^{\infty} \psi_j^2 V_{\text{var}}(\varepsilon_{t-j}) \quad \text{puisque } \{\varepsilon_t\} \text{ est bruit blanc.} \\ &= \sigma_\varepsilon^2 \sum_{j=1}^{\infty} \psi_j^2 = \sigma_\varepsilon^2 \sum_{j=1}^{\infty} \frac{1}{j^2} < \infty, \text{ car } \sum_{j=1}^{\infty} \frac{1}{j^2} \text{ est une} \\ &\quad \text{série-p où } p > 2, \text{ donc converge.} \end{aligned}$$

2)  $\text{IE}[Y_t] = \text{IE}\left[\sum_{j=1}^{\infty} \psi_j \varepsilon_{t-j}\right]$   
 $= \sum_{j=1}^{\infty} \text{IE}[\psi_j \varepsilon_{t-j}] = \sum_{j=1}^{\infty} \psi_j \text{IE}[\varepsilon_{t-j}] = 0, \text{ donc constante.}$

3)  $\text{Cov}\{Y_t, Y_{t+h}\} = \text{Cov}\left\{\sum_{j \geq 1} \psi_j \varepsilon_{t-j}, \sum_{l \geq 1} \psi_l \varepsilon_{t+h-l}\right\}$   
 $= \sum_{j \geq 1} \sum_{l \geq 1} \psi_j \psi_l \text{Cov}\{\varepsilon_{t-j}, \varepsilon_{t+h-l}\}$

$\Rightarrow$  lorsque  $t-j = t+h-l \Rightarrow \text{Cov}\{\varepsilon_{t-j}, \varepsilon_{t+h-l}\} = \sigma_\varepsilon^2$ .  
sinon  $\text{Cov}\{\varepsilon_{t-j}, \varepsilon_{t+h-l}\} = 0$ .

$$\begin{aligned} &= \sigma_\varepsilon^2 \sum_{j \geq 1} \psi_j \psi_{j+h} \quad \text{puisque } \begin{matrix} t-j = t+h-l \\ l = h+j \end{matrix} \\ &= \gamma_h \quad \text{puisque cela dépend uniquement du délai } h. \\ &\quad (\neq \text{dépend pas de } t). \end{aligned}$$

..  $\{Y_t\}$  est stationnaire

## PARTIE #2: FONCTION D'AUTO-COVARIAНCE

- On sait que  $\chi_n = \sigma_{\varepsilon}^2 \sum_{j \geq 1} \psi_j \psi_{j+n}$  où  $\psi_j = \frac{(-1)^j}{j}$ , donc :

$$\sigma_{\varepsilon}^2 \sum_{j=1}^{\infty} \frac{(-1)^j}{j} \cdot \frac{(-1)^{j+n}}{j+n} = \sigma_{\varepsilon}^2 \sum_{j=1}^{\infty} \frac{(-1)^{2j+n}}{j^2 + nj} = \boxed{\sigma_{\varepsilon}^2 (-1)^n} \sum_{j=1}^{\infty} \frac{1}{j^2 + nj}$$

puisque  $(-1)^{2j+n} = (-1)^{2j} (-1)^n = (-1)^n$

$$\text{À l'aside des fractions partielles: } \sum_{j=1}^{\infty} \frac{1}{j^2 + nj} = \sum_{j=1}^{\infty} \left( \frac{1}{nj} - \frac{1}{n(j+n)} \right)$$

### DÉMARCHE (FRACTIONS PARTIELLES):

$$\frac{1}{j^2 + nj} = \frac{1}{j(j+n)} = \frac{A}{j} + \frac{B}{j+n} = \frac{(j+n)A + Bj}{j(j+n)}$$

$$\text{On souhaite que } (j+n)A + Bj = 1 \Leftrightarrow Aj + Ah + Bj = 1 \\ j(A+B) + Ah = 1$$

On résoud le sys. d'équations:

$$\begin{cases} A+B=0 \\ hA=1 \end{cases} \Leftrightarrow A = \frac{1}{h} \text{ et } B = -\frac{1}{h}$$

$$\therefore \frac{1}{j(j+n)} = \frac{\frac{1}{h}}{j} + \frac{-\frac{1}{h}}{j+n} = \boxed{\frac{1}{jh} - \frac{1}{h(j+n)}}$$

Examinons la somme de plus près:

$$\sum_{j=1}^{\infty} \left[ \frac{1}{jh} - \frac{1}{h(j+n)} \right] = \frac{1}{h} - \cancel{\frac{1}{h(h+1)}} + \frac{1}{2h} - \cancel{\frac{1}{h(h+2)}} + \frac{1}{3h} - \cancel{\frac{1}{h(h+3)}} + \dots + \frac{1}{nh} - \cancel{\frac{1}{h(h+n)}} +$$

$\frac{1}{h(h+1)} - \cancel{\frac{1}{h(2h+1)}} + \cancel{\frac{1}{h(2h+2)}} - \cancel{\frac{1}{h(2h+3)}} + \dots$

$\downarrow$   
s'élime lorsque  $j=2h+1$        $\downarrow$   
s'élime lorsque  $j=2h+2$

$$= \frac{1}{h} + \frac{1}{2h} + \dots + \frac{1}{nh} = \frac{1}{h} \sum_{j=1}^n \frac{1}{j}$$

∴ Puisque  $\sum_{j=1}^{\infty} \frac{1}{j(j+n)} = \sum_{j=1}^{\infty} \left[ \frac{1}{jh} - \frac{1}{h(j+n)} \right] = \frac{1}{h} \sum_{j=1}^h \frac{1}{j}$ , on peut conclure que:

$$\chi_n = \sigma_{\varepsilon}^2 \sum_{j \geq 1} \psi_j \psi_{j+n} = \boxed{(-1)^n \frac{\sigma_{\varepsilon}^2}{h} \sum_{j=1}^h \frac{1}{j}}$$

b) Montrer que  $\sum_{n=1}^{\infty} \gamma_n^2 < \infty$ .

ÉTAPE H1 : BORNES POUR  $\sum_{n=1}^{\infty} \gamma_n^2$

- Posons  $\sum_{j=1}^n \frac{1}{j} = H_n$ , un nombre harmonique.
- On sait que  $\sum_{n=1}^{\infty} \gamma_n = \sum_{n=1}^{\infty} (-1)^n \frac{\sigma_{\varepsilon}^2}{n} H_n$
- Donc :  $\sum_{n=1}^{\infty} \gamma_n^2 = \sum_{n=1}^{\infty} \frac{\sigma_{\varepsilon}^4}{n^2} H_n^2 = \sigma_{\varepsilon}^4 \sum_{n=1}^{\infty} \frac{1}{n^2} H_n^2$
- L'idée est de trouver une série qui borne supérieurement  $\sum_{n=1}^{\infty} \gamma_n^2$ . Si cette borne est une série qui converge, nécessairement,  $\sum_{n=1}^{\infty} \gamma_n^2$  va converger aussi.
- Simplifier la tâche : commençons par trouver une borne supérieure et inférieure pour  $H_n$ .

BORNES POUR  $H_n$

Ici, posons  $k = \lceil \log_2(n) \rceil - 1$

↳ Borne supérieure :

$$\begin{aligned}
 H_n &= 1 + \frac{1}{2} + \frac{1}{3} + \dots + \frac{1}{n} \\
 &\leq 1 + \left[ \frac{1}{2} + \frac{1}{3} \right] + \left[ \frac{1}{4} + \frac{1}{5} + \frac{1}{6} + \frac{1}{7} \right] + \dots + \left[ \frac{1}{2^{k-1}} + \dots + \frac{1}{2^k-1} \right] + \left[ \frac{1}{2^k} + \frac{1}{2^{k+1}} + \dots + \frac{1}{2^{k+1}-1} \right] + \left[ \frac{1}{2^{k+2}} \right] \\
 &< 1 + \left[ \frac{1}{2} + \frac{1}{2} \right] + \left[ \frac{1}{4} + \frac{1}{4} + \frac{1}{4} + \frac{1}{4} \right] + \dots + \left[ \frac{1}{2^{k-1}} + \dots + \frac{1}{2^{k-1}} \right] + \left[ \frac{1}{2^k} + \dots + \frac{1}{2^k} \right] + \left[ \frac{1}{2^{k+1}} \right] \\
 &\quad \textcircled{1} \quad \textcircled{1} \quad \textcircled{1} \quad \textcircled{1} \\
 &= k+1 + \frac{1}{2^{k+1}} \\
 &< k+2 \\
 &= \lceil \log_2(n) \rceil - 1 + 2 \\
 &\leq \boxed{\lceil \log_2(n) \rceil + 2} \quad \text{puisque } \lceil x \rceil \leq x+1
 \end{aligned}$$

↳ Borne inférieure :

$$\begin{aligned}
 H_n &= 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \dots + \frac{1}{n} \\
 &\geq 1 + \frac{1}{2} + \left[ \frac{1}{3} + \frac{1}{4} \right] + \left[ \frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8} \right] + \dots + \left[ \frac{1}{2^{k+1}} + \dots + \frac{1}{2^{k+2}} \right] \\
 &> 1 + \frac{1}{2} + \left[ \frac{1}{2} + \frac{1}{2} \right] + \left[ \frac{1}{8} + \frac{1}{8} + \frac{1}{8} + \frac{1}{8} \right] + \dots + \left[ \frac{1}{2^{k+1}} + \dots + \frac{1}{2^{k+1}} \right] \\
 &= 1 + \frac{1}{2} + \frac{k}{2} \\
 &= 1 + \frac{1}{2} + \frac{\lfloor \log_2(n) \rfloor - 1}{2} \\
 &= \frac{\lfloor \log_2(n) \rfloor}{2} + \frac{1}{2} \\
 &\geq \boxed{\frac{\log_2(n)}{2}} \quad \text{puisque } \lfloor x \rfloor \geq x-1
 \end{aligned}$$

Ici, posons  $k = \lfloor \log_2(n) \rfloor - 1$

• Par conséquent, nous avons réussi à trouver les bornes pour  $H_h$ :

$$\frac{\lg(h)}{2} < H_h < \lg(h) + 2$$

où  $\lg(h) = \log_2(h)$

• Donc lorsqu'on veut trouver les bornes pour  $H_h^2$ , on obtient cela:

$$\frac{\lg^2(h)}{4} < H_h^2 < \lg^2(h) + 4\lg(h) + 4$$

⇒ cela est valide puisque  $0 \leq a < b \Rightarrow a^2 < b^2$

• Pour les bornes de  $\sum_{h=1}^{\infty} \chi_h^2$ , on obtient:

$$\sum_{h=1}^{\infty} \chi_h^2 = \sigma_E^4 \sum_{h=1}^{\infty} \frac{1}{h^2} H_h^2 \quad \text{ET} \quad \sigma_E^4 \sum_{h=1}^{\infty} \frac{\lg^2(h) + 4\lg(h) + 4}{h^2}$$

$$\text{où } \frac{H_h^2}{h^2} < \frac{\lg^2(h) + 4\lg(h) + 4}{h^2} \text{ pour tout } h.$$

### ÉTAPE H2: ANALYSE DE LA SÉRIE A

• Séparons la série en trois séries différentes:

$$A = \sum_{h=1}^{\infty} \frac{\lg^2(h)}{h^2} + \sum_{h=1}^{\infty} \frac{4\lg(h)}{h^2} + \sum_{h=1}^{\infty} \frac{4}{h^2} \quad \therefore A = B + C + D$$

B

C

D

si

• Nous allons faire le test d'intégrales pour B et C afin de vérifier ces dernières convergent.

RAPPEL TEST D'INTÉGRALE: Soit la fonction est continue, positive et décroissante sur  $[1, \infty$ . t.q.  $a_n = f(n) \Rightarrow \sum a_n$  converge  $\Leftrightarrow \int_1^{\infty} f(x) dx$  est convergente.

Série C:  $\sum_{h=1}^{\infty} \frac{\lg(h)}{h^2}$  où on néglige la constante 4.

$$\frac{\lg(x)}{x^2} = \frac{\ln(x)}{\ln(2)x^2} : \text{Dérivée de cette expression: } -\frac{2\ln(x) - 1}{\ln(2)x^3} = 0$$

$$\ln x = \frac{1}{2}$$

$$x = \sqrt{e}$$

Fonction positive et décroissante sur  $[\sqrt{e}, +\infty$

continue

On calcule l'intégrale impropre de  $\int_{\sqrt{e}}^{\infty} \frac{\ln(x)}{\ln(2)x^2} dx$

$$\frac{1}{\ln(2)} \left[ \int_{\sqrt{e}}^{\infty} \frac{\ln(x)}{x^2} dx \right] \quad \begin{bmatrix} u = \ln(x) & v = -\frac{1}{x} \\ du = \frac{1}{x} dx & dv = x^{-2} dx \end{bmatrix} \quad \int u dv = uv - \int v du$$

$$\int_{\sqrt{e}}^{\infty} \frac{\ln(x)}{x^2} dx = \left[ \frac{\ln(x)}{x} \right]_{\sqrt{e}}^{\infty} + \int_{\sqrt{e}}^{\infty} \frac{1}{x^2} dx = \left[ \frac{-\ln(x)}{x} \right]_{\sqrt{e}}^{\infty} - \left[ \frac{1}{x} \right]_{\sqrt{e}}^{\infty}$$

$$= \left[ \frac{-\ln(x) - 1}{x} \right]_{\sqrt{e}}^{\infty} = \left[ \frac{-\ln(x) - 1}{x} \right]_{\sqrt{e}}^{\infty} - \left\{ \frac{-\ln(\sqrt{e}) - 1}{\sqrt{e}} \right\}$$

$$= + \left[ \frac{1}{2} + \frac{1}{2} \right] = \frac{3}{2\sqrt{e}}$$

$$\therefore \text{On obtient } \int_e^\infty \frac{\ln(x)}{\ln^2(2)x^2} dx = \frac{3}{2\ln^2(2)e} < \infty$$

$\Rightarrow \sum_{h=e}^{\infty} \frac{\ln(h)}{h^2}$  converge. En outre, puisque  $\frac{\ln(x)}{x^2}$  est continue entre  $[1, \sqrt{e}]$ ,  $\sum_{h=1}^{\infty} \frac{\ln(h)}{h^2}$  converge aussi.

Série ⑧ :  $\sum_{h=1}^{\infty} \frac{\ln^2(h)}{h^2}$

$$f(x) = \frac{\ln^2(x)}{x^2} = \frac{\ln^2(x)}{\ln^2(2)x^2} : \text{Dérivée de cette expression: } -\frac{2(\ln(x)-1)\ln(x)}{\ln^2(2)x^3} = 0$$

$\Downarrow$   
 $x=e$

- Fonction continue, positive et décroissante sur  $[e, +\infty)$

$$\frac{1}{\ln^2(2)} \left[ \int_e^\infty \frac{\ln^2(x)}{x^2} dx \right]$$

$\Downarrow$

$$\begin{aligned} u &= \ln^2(x) & v &= -\frac{1}{x} \\ du &= 2\frac{\ln(x)}{x} dx & dv &= x^{-2} dx \end{aligned}$$

$$\begin{aligned} \int_e^\infty \frac{\ln^2(x)}{x^2} dx &= -\frac{\ln^2(x)}{x} \Big|_e^\infty + \int_e^\infty \frac{2\ln(x)}{x^2} dx \\ &= \left[ -\frac{\ln^2(x)}{x} + 2 \left\{ -\frac{\ln(x) - 1}{x} \right\} \right]_e^\infty \\ &= \left[ \frac{-\ln^2(x) - 2\ln(x) - 2}{x} \right]_e^\infty \\ &= -\cancel{\ln^2(\infty) - 2\ln(\infty) - 2} + \left\{ \frac{\ln^2(e) + 2\ln(e) + 2}{e} \right\} \\ &= \frac{1+2+2}{e} = \frac{5}{e} \end{aligned}$$

$$\therefore \text{On obtient } \int_e^\infty \frac{\ln^2(x)}{x^2} dx = \frac{5}{\ln^2(2)e} < \infty$$

$\Rightarrow \sum_{h=e}^{\infty} \frac{\ln^2(h)}{h^2}$  converge. En outre, puisque  $\frac{\ln^2(x)}{x^2}$  est continue entre  $[1, e]$ ,  $\sum_{h=1}^{\infty} \frac{\ln^2(h)}{h^2}$  converge aussi.

Série ⑨ :  $\sum_{h=1}^{\infty} \frac{1}{h^2}$  où on néglige la constante 4.

$\hookrightarrow$  Série - p où  $p > 1$  ( $p=2$ ), donc automatiquement, cette série converge.

## ÉTAPE #3: CONVERGENCE de $\sum_{h=1}^{\infty} \gamma_h^2$

- Puisque  $\sum_{h=1}^{\infty} \frac{\lg^2(h)}{h^2}$ ,  $\sum_{h=1}^{\infty} \frac{4\lg(h)}{h^2}$  et  $\sum_{h=1}^{\infty} \frac{4}{h^2}$  convergent, la somme des trois séries convergent aussi, donc:  $\sum_{h=1}^{\infty} \frac{\lg^2(h) + 4\lg(h) + 4}{h^2} < \infty$

- Retournons à la question:

$$\sum_{h=1}^{\infty} \gamma_h^2 = \sum_{h=1}^{\infty} \sigma_E^2 \frac{1}{h^2} H_h^2 \quad \text{ET} \quad \sigma_E^2 \sum_{h=1}^{\infty} \frac{\lg^2(h) + 4\lg(h) + 4}{h^2}$$

- Pour le test de comparaison, étant donné que  $0 \leq \frac{H_h^2}{h^2} < \frac{\lg^2(h) + 4\lg(h) + 4}{h^2}$  et que (A) converge, nécessairement  $\sum_{h=1}^{\infty} \gamma_h^2$  converge aussi.

$$\therefore \sum_{h=1}^{\infty} \gamma_h^2 < \infty$$

■

- c)** Qu'en est-il de  $\sum_{h=1}^{\infty} |\gamma_h|$ ? Discuter en rapport avec b).

$$\sum_{h=1}^{\infty} |\gamma_h| = \sigma_E^2 \sum_{h=1}^{\infty} \frac{H_h}{h}.$$

⇒ Pour simplifier la preuve, on va se servir encore une fois une borne pour  $H_h$  trouvée en b) et faire un test de comparaison.

- Prenons la borne inférieure de  $H_h$  qui est  $\frac{\lg(h)}{2}$ .

$$\sum_{h=1}^{\infty} |\gamma_h| = \sigma_E^2 \sum_{h=1}^{\infty} \frac{H_h}{h} \quad \text{ET} \quad \frac{\sigma_E^2}{2} \sum_{h=1}^{\infty} \frac{\lg(h)}{h} \quad \text{(E)}$$

(E)

où  $\frac{H_h}{h} > \frac{\lg(h)}{h}$  pour tout  $h$ .

- Pour  $\sum_{h=1}^{\infty} \frac{\lg(h)}{h}$ , pour  $h=2$ , on constate que  $\frac{\lg(h)}{h} > \frac{1}{h}$
- Par conséquent  $\sum_{h=1}^{\infty} \frac{\lg(h)}{h}$  (ET)  $\sum_{h=1}^{\infty} \frac{1}{h}$ : puisque  $\sum_{h=1}^{\infty} \frac{1}{h}$  diverge,  $\sum_{h=1}^{\infty} \frac{\lg(h)}{h}$  diverge aussi.
- Donc (E) est une série divergente. Pour le test de comparaison, puisque  $\frac{H_h}{h} > \frac{\lg(h)}{h}$  et que (E) diverge, donc  $\sum_{h=1}^{\infty} |\gamma_h|$  diverge aussi.

$$\therefore \sum_{h=1}^{\infty} |\gamma_h| \text{ diverge.}$$

## EXERCICE 2

a)  $X_t = -0,2X_{t-1} + 0,48X_{t-2} + \varepsilon_t$

$$\underbrace{X_t + 0,2X_{t-1} - 0,48X_{t-2}}_{\phi(B)} = \varepsilon_t$$

$$(I + 0,2B - 0,48B^2)X_t = \varepsilon_t$$

$$(I + 0,8B)(I - 0,6B)X_t = \varepsilon_t$$

$$\begin{aligned}\phi(z) &= -0,48z^2 + 0,2z + 1 & -0,6 \times 0,8 &= -0,48 \\ &= -0,48z^2 - 0,6z + 0,8z + 1 & -0,6 \times 0,8 &= 0,2 \\ &= (0,8z + 1) - 0,6z + (0,8z + 1) \\ &= (0,8z + 1)(1 - 0,6z)\end{aligned}$$

• Nous avons AR(2), donc inversible.

• Racines de  $\phi(z)$ :  $1 + 0,8z = 0$  IET  $1 - 0,6z = 0$

$$z = \frac{-1}{0,8} = -1,25$$

$$z = \frac{5}{3}$$

$$|z| > 1$$

$$|z| > 1$$

• Les racines de  $\phi(z)$  sont supérieures à 1 en module, donc stationnaire.

$$\begin{cases} \phi_1 = -0,2 \\ \phi_2 = 0,48 \end{cases}$$

- ARMA(2,0) = AR(2)
- Inversible
- Stationnaire

b)  $X_t = 0,80X_{t-1} - 0,15X_{t-2} + \varepsilon_t - 0,30\varepsilon_{t-1}$

$$X_t - 0,8X_{t-1} + 0,15X_{t-2} = \varepsilon_t - 0,3\varepsilon_{t-1}$$

$$(I - 0,8B + 0,15B^2)X_t = (I - 0,3B)\varepsilon_t$$

$$(I - 0,5B)(I - 0,3B)X_t = (I - 0,3B)\varepsilon_t$$

$$(I - 0,5B)X_t = \varepsilon_t \iff X_t = 0,5X_{t-1} + \varepsilon_t$$

$$\begin{aligned}\phi(z) &= 0,15z^2 - 0,8z + 1 & -0,3 \times 0,5 &= 0,15 \\ &= 0,15z^2 - 0,3z - 0,5z + 1 & -0,3 + 0,5 &= -0,8\end{aligned}$$

$$\begin{aligned}&= (-0,5z + 1) - 0,3z - 0,5z + 1 \\ &= (1 - 0,5z)(1 - 0,3z)\end{aligned}$$

• Nous avons AR(1), donc inversible

• Racines de  $\phi(z)$ :  $1 - 0,5z = 0$

$$z = \frac{+1}{0,5} = +2 \Rightarrow |z| > 1$$

• La racine de  $\phi(z)$  est > 1 en module, donc stationnaire.

$$\begin{cases} \phi_1 = 0,5 \end{cases}$$

- ARMA(1,0) = AR(1)
- Inversible
- Stationnaire

$$c) X_t = -1,9 X_{t-1} - 0,88 X_{t-2} + \varepsilon_t + 0,2 \varepsilon_{t-1} + 0,7 \varepsilon_{t-2}$$

$$X_t + 1,9 X_{t-1} + 0,88 X_{t-2} = \varepsilon_t + 0,2 \varepsilon_{t-1} + 0,7 \varepsilon_{t-2}$$

$$\underbrace{(I + 1,9B + 0,88B^2)}_{\phi(B)} X_t = \underbrace{(I + 0,2 B + 0,7 B^2)}_{\Theta(B)} \varepsilon_t$$

$$\begin{aligned} \cdot \phi(z) &= 0,88z^2 + 1,9z + 1 \\ &= 0,88z^2 + 0,8 + 1,1z + 1 \\ &= (1,1z + 1) 0,8z + 1,1z + 1 \\ &= (1 + 1,1z)(1 + 0,8z) \end{aligned}$$

$$\frac{0,8}{0,8} \times \frac{1,1}{+1,1} = 0,88 \quad \Theta(z) = 0,7z^2 + 0,2z + 1$$

$$\begin{aligned} z &= \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} = \frac{-0,2 \pm \sqrt{0,04 - 4 \cdot 0,7}}{2 \cdot 0,7} \\ &= \frac{-0,2 \pm \sqrt{2,76}i}{1,4} = \frac{-0,2 \pm \sqrt{\frac{69}{25}}i}{1,4} \\ &= \frac{-0,2 \pm \frac{\sqrt{69}}{5}i}{1,4} = \frac{-1 \pm \frac{\sqrt{69}}{7}i}{1,4} \end{aligned}$$

$\therefore$  Aucun facteur en commun: ARMA(2,2)

Racines de  $\Theta(z)$ :

Module d'un nombre complexe:  $|z| = \sqrt{a^2 + b^2}$  où  $a$  est partie réelle  
 $b$  est partie imaginaire.

$$\begin{aligned} \text{Pour } -\frac{1}{7} + \frac{\sqrt{69}}{7}i: \quad |z| &= \sqrt{\left(-\frac{1}{7}\right)^2 + \left(\frac{\sqrt{69}}{7}\right)^2} \\ &= \sqrt{\frac{1}{49} + \frac{69}{49}} = \sqrt{\frac{70}{49}} = 1,1952 > 1 \end{aligned}$$

$$\text{Pour } -\frac{1}{7} - \frac{\sqrt{69}}{7}i: \quad |z| = \sqrt{\left(-\frac{1}{7}\right)^2 + \left(-\frac{\sqrt{69}}{7}\right)^2} = \sqrt{\frac{70}{49}} = 1,1952 > 1$$

$\therefore$  Les racines de  $\Theta(z)$  sont  $> 1$  en module, donc inversible.

Racines de  $\phi(z)$ :

$$1 + 1,1z = 0$$

ET

$$1 + 0,8z = 0$$

$$z = \frac{-1}{1,1} = -\frac{10}{11}$$

$$z = \frac{-1}{0,8} = -1,25$$

$$|z| > 1$$

$$|z| > 1$$

$\therefore$  Il existe une racine où son module est  $< 1$ , donc non stationnaire.

$$\begin{cases} \phi_1 = -1,9 & \theta_1 = 0,2 \\ \phi_2 = -0,88 & \theta_2 = 0,7 \end{cases} \Leftarrow$$

- ARMA(2,2)
- Inversible
- Non Stationnaire

$$d) X_t = X_{t-1} - 0,5X_{t-2} + \varepsilon_t - \varepsilon_{t-1}$$

$$\phi(z) = 0,5z^2 - z + 1$$

$$X_t - X_{t-1} + 0,5X_{t-2} = \varepsilon_t - \varepsilon_{t-1}$$

$$z = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} = \frac{1 \pm \sqrt{1-2}}{1}$$

$$\underbrace{(I - B + 0,5B^2)}_{\phi(B)} X_t = \underbrace{(I - B)}_{\Theta(B)} \varepsilon_t$$

$$= 1 \pm i$$

• Racines de  $\Theta(z)$ :  $1-z=0$   
 $z=1$        $|z| \neq 1$

∴ Aucun facteur en commun  $\Rightarrow$  ARMA(2,1)

∴  $|z| \neq 1$ , donc non inversible.

• Racines de  $\phi(z)$ :

Module d'un nombre complexe:

$$|z| = \sqrt{(1)^2 + (1)^2} = \sqrt{2} > 1$$

∴ Toutes les racines de  $\phi(z)$  sont  $> 1$  en module, donc stationnaire.

$$\begin{cases} \phi_1 = 1 \\ \phi_2 = -0,5 \\ \theta_1 = -1 \end{cases} \Leftrightarrow \boxed{\begin{array}{l} \cdot \text{ARMA}(2,1) \\ \cdot \text{Non inversible} \\ \cdot \text{Stationnaire} \end{array}}$$

$$e) X_t + 0,6X_{t-1} = \varepsilon_t + 1,2\varepsilon_{t-1}$$

$$\underbrace{(I + 0,6B)}_{\phi(B)} X_t = \underbrace{(I + 1,2B)}_{\Theta(B)} \varepsilon_t \quad \therefore \text{Aucun facteur en commun, donc ARMA}(1,1)$$

• Racines de  $\Theta(z)$ :  $1 + 1,2z = 0$

$$z = -\frac{1}{1,2} = -\frac{5}{6} \quad \therefore |z| \neq 1 \quad \therefore \text{Non inversible}$$

• Racines de  $\phi(z)$ :  $1 + 0,6z = 0$

$$z = -\frac{1}{0,6} = -\frac{5}{3} \quad \therefore |z| > 1 \quad \therefore \text{Stationnaire.}$$

$$\begin{cases} \phi_1 = -0,6 \\ \theta_1 = 1,2 \end{cases} \Leftrightarrow \boxed{\begin{array}{l} \cdot \text{ARMA}(1,1) \\ \cdot \text{Non inversible} \\ \cdot \text{Stationnaire} \end{array}}$$

$$f) X_t + 1,8X_{t-1} + 0,81X_{t-2} = \varepsilon_t$$

$$\phi(z) = 0,81z^2 + 1,8z + 1$$

$$\frac{0,9}{0,9} \times \frac{0,9}{0,9} = 0,81$$

$$\underbrace{(I + 1,8B + 0,81B^2)}_{\phi(B)} X_t = \varepsilon_t \quad = (1 + 0,9z)^2$$

$$(I + 0,9B)^2 X_t = \varepsilon_t \quad \rightarrow \text{AR}(2), \text{ donc inversible}$$

• Racines de  $\phi(z)$ :  $1 + 0,9z = 0$

$$z = -\frac{1}{0,9} = -\frac{10}{9}, \text{ donc } |z| > 1 \quad \therefore \text{Stationnaire}$$

$$\begin{cases} \phi_1 = -1,8 \\ \phi_2 = -0,81 \end{cases} \Leftrightarrow$$

$$\boxed{\begin{array}{l} \cdot \text{ARMA}(2,0) = \text{AR}(2) \\ \cdot \text{Inversible} \\ \cdot \text{Stationnaire} \end{array}}$$

$$g) X_t + 1,6X_{t-1} = \varepsilon_t - 0,4\varepsilon_{t-1} + 0,04\varepsilon_{t-2} \quad \Theta(z) = 0,04z^2 - 0,4z + 1 \quad \frac{-0,2}{-0,2} \times \frac{-0,2}{-0,2} = 0,04$$

$$(I + 1,6B)X_t = (I - 0,4B + 0,04B^2)\varepsilon_t \quad = (1 - 0,2z)^2$$

$$(I + 1,6B)X_t = (I - 0,2B)^2 \varepsilon_t \quad \therefore \text{Aucun facteur en commun, donc ARMA(1,2)}$$

• Racines de  $\Theta(z)$ :  $(1 - 0,2z) = 0$

$$z = \frac{1}{0,2} = 5 > 1, \text{ donc inversible.}$$

• Racines de  $\phi(z)$ :  $(1 + 1,6z) = 0$

$$z = \frac{-1}{1,6} = -\frac{5}{8}, \text{ donc } |z| > 1 \quad \therefore \text{Non stationnaire.}$$

$$\begin{cases} \phi_1 = -1,6 \\ \theta_1 = -0,4 \\ \theta_2 = 0,04 \end{cases} \leftarrow \boxed{\begin{array}{l} \cdot \text{ARMA(1,2)} \\ \cdot \text{Inversible} \\ \cdot \text{Non stationnaire} \end{array}}$$



### EXERCICE 3

Soit les processus: ①  $X_t = \phi X_{t-1} + \varepsilon_t + \theta \varepsilon_{t-1} \rightarrow \text{ARMA}(1,1)$

où  $\{\varepsilon_t\}$  est le bruit blanc.

$$\textcircled{2} \quad Y_t = \phi Y_{t-1} + \varepsilon_t \rightarrow \text{ARMA}(1,0) = \text{AR}(1)$$

$$\textcircled{3} \quad Z_t = \varepsilon_t + \theta \varepsilon_{t-1} \rightarrow \text{ARMA}(0,1) = \text{MA}(1)$$

a) Calculer les fonctions d'auto-corrélation des processus  $X_t, Y_t$  et  $Z_t$ .

$\Rightarrow$  Pour les prochaines questions, on pose  $\phi = 0,6$  et  $\theta = 0,9$ , donc vérifier si ces trois processus sont inversibles / stationnaires.

•  $\phi(z) = 1 - 0,6z = 0$       •  $\Theta(z) = 1 + 0,9z = 0$

$$z = \frac{1}{0,6} = \frac{5}{3} > 1 \quad z = \frac{-1}{0,9} = -\frac{10}{9}$$

$$|z| > 1$$

$\therefore$  Tous les processus sont stationnaires (et inversibles).

$\Rightarrow$  On peut donc utiliser les formules suivantes (p.10 des notes de cours - chap.4):

$$\left\{ \begin{array}{l} \textcircled{1} \quad \gamma_k = \sigma_\varepsilon^2 \sum_{j \geq 0} \psi_j \psi_{j+k} \\ \textcircled{2} \quad \gamma_0 = \sigma_\varepsilon^2 \sum_{j \geq 0} \psi_j^2 \\ \textcircled{3} \quad \rho_k = \frac{\gamma_k}{\gamma_0} = \frac{\sum_{j \geq 0} \psi_j \psi_{j+k}}{\sum_{j \geq 0} \psi_j^2} \end{array} \right.$$

PARTIE #1 : PROCESSUS ARMA(1,1) - X<sub>t</sub>

$$X_t = \phi X_{t-1} + \varepsilon_t + \theta \varepsilon_{t-1}$$

$$(I - \phi B) X_t = (I + \theta B) \varepsilon_t$$

$$X_t = \left\{ (I + \theta B) \cdot \frac{1}{I - \phi B} \right\} \varepsilon_t$$

$$= (I + \theta B) (1 + \phi B + \phi^2 B^2 + \phi^3 B^3 + \phi^4 B^4 + \dots) \varepsilon_t$$

$$= [1 + \phi B + \phi^2 B^2 + \phi^3 B^3 + \phi^4 B^4 + \dots + \theta B + \phi \theta B^2 + \phi^2 \theta B^3 + \phi^3 \theta B^4 + \dots] \varepsilon_t$$

$$= [1 + (\phi + \theta) B + \phi(\phi + \theta) B^2 + \phi^2(\phi + \theta) B^3 + \dots] \varepsilon_t$$

Donc, finalement, on obtient :  $\rightarrow \psi_0 = 1$   
 $\rightarrow \psi_j = \phi^{j-1}(\phi + \theta)$

• Fonction de variance ( $\gamma_0$ ) : et d'auto-covariance ( $\gamma_k$ ) :

$$\begin{aligned} \gamma_0 &= \sigma_\varepsilon^2 \sum_{j \geq 0} \psi_j^2 \\ &= \sigma_\varepsilon^2 \{ \psi_0^2 + \psi_1^2 + \psi_2^2 + \psi_3^2 + \dots \} \\ &= \sigma_\varepsilon^2 \{ 1 + (\phi + \theta)^2 + \phi^2(\phi + \theta)^2 + \phi^4(\phi + \theta)^2 + \dots \} \\ &= \sigma_\varepsilon^2 \{ 1 + \frac{(\phi + \theta)^2}{1 - \phi^2} \} \end{aligned}$$

$$\begin{aligned} \gamma_k &= \sigma_\varepsilon^2 \sum_{j \geq 0} \psi_j \psi_{j+k} \\ &= \sigma_\varepsilon^2 \{ \psi_0 \psi_k + \psi_1 \psi_{k+1} + \psi_2 \psi_{k+2} + \psi_3 \psi_{k+3} + \dots \} \\ &= \sigma_\varepsilon^2 \{ 1 \cdot \phi^{k-1}(\phi + \theta) + (\phi + \theta) \phi^k(\phi + \theta) + \phi(\phi + \theta) \phi^{k+1}(\phi + \theta) + \phi^2(\phi + \theta) \phi^{k+2}(\phi + \theta) + \dots \} \\ &= \phi^{k-1} \sigma_\varepsilon^2 \{ (\phi + \theta) + \phi(\phi + \theta)^2 + \phi^2 \phi(\phi + \theta)^2 + \phi^4 \phi(\phi + \theta)^2 + \dots \} \\ &= \sigma_\varepsilon^2 \phi^{k-1} \{ (\phi + \theta) + \frac{\phi(\phi + \theta)^2}{1 - \phi^2} \} \quad \hookrightarrow \phi^5 \phi^{k-1} = \phi^{5+k-1} = \phi^{k+4} \end{aligned}$$

• Fonction d'Auto-corrélation : ①

$$\rho_k = \frac{\gamma_k}{\gamma_0} = \frac{\phi^{k-1} \{ (\phi + \theta) + \frac{\phi(\phi + \theta)^2}{1 - \phi^2} \}}{1 + \frac{(\phi + \theta)^2}{1 - \phi^2}} \quad ②$$

$$\begin{aligned} ① \quad \phi + \theta + \frac{\phi(\phi + \theta)^2}{1 - \phi^2} &= \frac{(\phi + \theta)(1 - \phi^2)}{1 - \phi^2} + \phi \{ \theta^2 + 2\phi\theta + \phi^2 \} \\ &= \frac{\phi - \phi^3 + \theta - \phi^2\theta + \theta^2\phi + 2\phi^2\theta + \phi^3\theta}{1 - \phi^2} = \frac{\phi + \theta + \phi^2\theta + \theta^2\phi}{1 - \phi^2} \\ &= \frac{(\phi + \theta)(1 + \phi\theta)}{1 - \phi^2} \end{aligned}$$

$$② \quad 1 + \frac{(\phi + \theta)^2}{1 - \phi^2} = \frac{1 - \phi^2 + \phi^2 + 2\phi\theta + \theta^2}{1 - \phi^2} = \frac{1 + 2\phi\theta + \theta^2}{1 - \phi^2} = \frac{1 + \theta^2 + 2\phi\theta}{1 - \phi^2}$$

$$\therefore \text{Finlement: } \rho_k = \frac{\gamma_k}{\gamma_0} = \left\{ \frac{(\phi + \theta)(1 + \phi\theta)}{1 + \theta^2 + 2\phi\theta} \right\} \phi^{k-1}$$

{ on suppose que les délays sont possibles. }

$$\hookrightarrow \text{Pour ARMA}(1,1): \boxed{\rho_k = \phi^{1k-1} \left\{ \frac{(\phi + \theta)(1 + \phi\theta)}{1 + \theta^2 + 2\phi\theta} \right\}}$$

.. De façon générale:

$$\rho_k = \begin{cases} 1 & \text{si } k=0 \\ \frac{(\phi+\theta)(1+\phi\theta)}{1+\theta^2+2\phi\theta} & \text{si } k=1 \\ \phi^{|k|} \left\{ \frac{(\phi+\theta)+(1+\phi\theta)}{1+\theta^2+2\phi\theta} \right\} & \text{si } |k| > 1 \\ \downarrow \\ \text{ou } \phi^{|k|-1} \rho_1 \end{cases}$$

## PARTIE #2: PROCESSUS AR(1) - $\gamma_t$ :

$$Y_t - \phi Y_{t-1} = \varepsilon_t$$

$$(1 - \phi B) Y_t = \varepsilon_t$$

$$Y_t = \frac{1}{1 - \phi B} \varepsilon_t$$

$$Y_t = \{1 + \phi B + \phi^2 B^2 + \phi^3 B^3 + \dots\} \varepsilon_t$$

$$\therefore \psi_0 = 1 \text{ et } \psi_j = \phi^j$$

• Calcul de  $\gamma_0$  et de  $\gamma_k$ :

$$\begin{aligned} \boxed{\gamma_0} &= \sigma_\varepsilon^2 \sum_{j \geq 0} \psi_j^2 \\ &= \sigma_\varepsilon^2 \{ \psi_0^2 + \psi_1^2 + \psi_2^2 + \psi_3^2 + \psi_4^2 + \dots \} \\ &= \sigma_\varepsilon^2 \{ 1 + \phi^2 + \phi^4 + \phi^6 + \phi^8 + \dots \} \\ &= \frac{\sigma_\varepsilon^2}{1 - \phi^2} \end{aligned}$$

$$\boxed{\gamma_k} = \sigma_\varepsilon^2 \sum_{j \geq 0} \psi_j \psi_{j+k}$$

$$= \sigma_\varepsilon^2 \{ \psi_0 \psi_k + \psi_1 \psi_{k+1} + \psi_2 \psi_{k+2} + \psi_3 \psi_{k+3} + \dots \}$$

$$= \sigma_\varepsilon^2 \{ 1 \cdot \phi^k + \phi \cdot \phi^{k+1} + \phi^2 \cdot \phi^{k+2} + \phi^3 \cdot \phi^{k+3} + \dots \}$$

$$= \phi^k \sigma_\varepsilon^2 \{ 1 + \phi^2 + \phi^4 + \phi^6 + \dots \}$$

$$= \frac{\sigma_\varepsilon^2}{1 - \phi^2} \phi^k$$

$$\rho_k = \begin{cases} 1 & \text{si } k=0 \\ \phi^{|k|} & \text{si } |k| > 0 \end{cases}$$

• Calcul de  $\rho_k$ :

$$\rho_k = \frac{\gamma_k}{\gamma_0} = \frac{\sigma_\varepsilon^2 \phi^k}{1 - \phi^2} \cdot \frac{1 - \phi^2}{\sigma_\varepsilon^2} = \phi^k$$



PARTIE #3: PROCESSUS MA(1) -  $Z_t$ :

$$Z_t = \varepsilon_t + \theta \varepsilon_{t-1}$$

$$Z_t = (1 + \theta B) \varepsilon_t$$

$$\hookrightarrow \text{Ici: } \begin{cases} \psi_0 = 1 \\ \psi_1 = \theta \\ \psi_j = 0 \text{ pour } j \geq 2 \end{cases}$$

• Calcul  $\gamma_0$  et  $\gamma_k$ :

$$|\gamma_0| = \sigma_\varepsilon^2 \sum_{j \geq 0} \psi_j^2 = \sigma_\varepsilon^2 \{1 + \theta^2\}$$

$$|\gamma_k| = \sigma_\varepsilon^2 \sum_{j \geq 0} \psi_j \psi_{j+k} = \begin{cases} \sigma_\varepsilon^2 \theta & \text{si } k = \pm 1 \\ 0 & \text{si } |k| > 1 \end{cases}$$

• Calcul de  $\rho_k$ :

$$\rho_k = \frac{\gamma_k}{\gamma_0} = \frac{\sigma_\varepsilon^2 \theta}{\sigma_\varepsilon^2 (1 + \theta^2)} = \frac{\theta}{(1 + \theta^2)} \quad \text{si } k = \pm 1$$

$$= 0 \quad \text{si } |k| > 1$$

$$\therefore \rho_k = \begin{cases} \frac{\theta}{1 + \theta^2} & \text{si } k = 0 \\ 0 & \text{si } k = \pm 1 \\ 0 & \text{si } |k| > 1 \end{cases}$$

Pour les autres questions, consulter la version papier du devoir  
(tous les résultats sont obtenus sous R).



```

# PROJET FINAL
# Auteur: Ronnie Liu
# Date: Mercredi 27 juillet 2021
# Objet: Fichier R pour QUESTIONS 3 et 4
n <- 100
phi <- 0.6
theta <- 0.9

#####
# a) Fonctions pour calculer les fonctions d'auto-corrélation

# Voir la démarche théorique de l'obtention des fonctions d'auto-corrélation
# pour chaque processus (p.13).

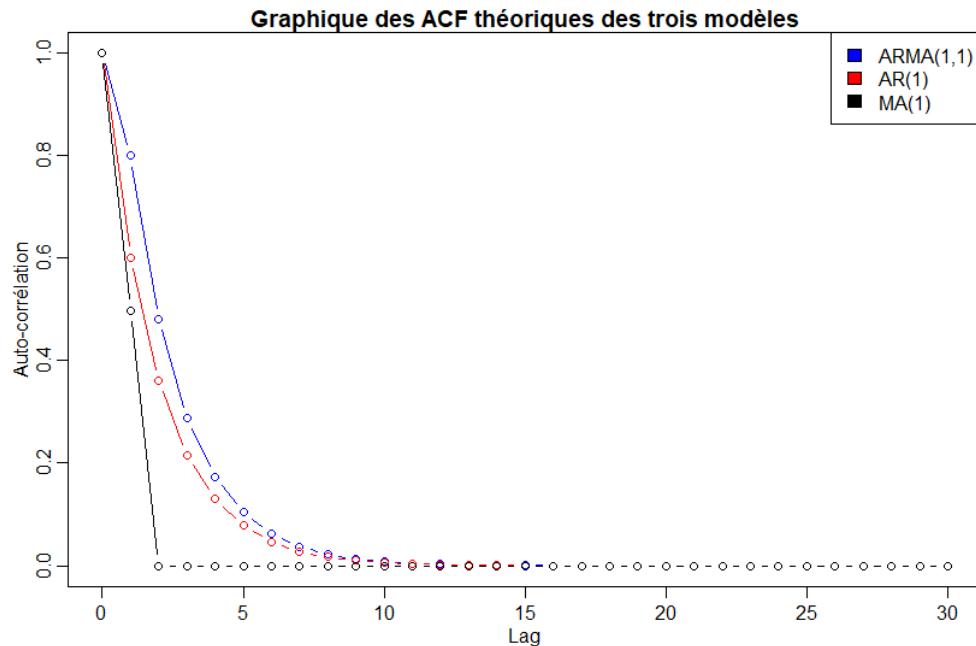
rho_1 <- ((phi+theta)*(1+phi*theta))/(1+2*phi*theta+theta^2)
# Pour le modèle ARMA(1,1):
#   | = 1           si k = 0
# rho_k |
#   | = phi^(k-1)*rho_1 sinon

# Pour le modèle AR(1):
#   | = 1           si k = 0
# rho_k |
#   | = phi^|k|   sinon

# Pour le modèle MA(1):
#   | = 1           si k = 0
# rho_k | = theta/(1+theta^2) si k = +/- 1
#   | = 0           sinon

#####
# b) Traçage des ACF des trois séries

```



```

# Capacités de diagnostic pour chaque processus:
# Tous les processus sont inversibles et stationnaires dans cette question
# (voir 3a).

# Processus Xt: ARMA(1,1):
# 1. ACF: Décroissance exponentielle moins rapide que celle d'un modèle AR(1).
# 2. PACF: Décroissance exponentielle aussi, donc nombre infini de pics.
# Bref, détecter s'il y a un nombre infini de pics dans les deux graphiques
# d'auto-corrélation.

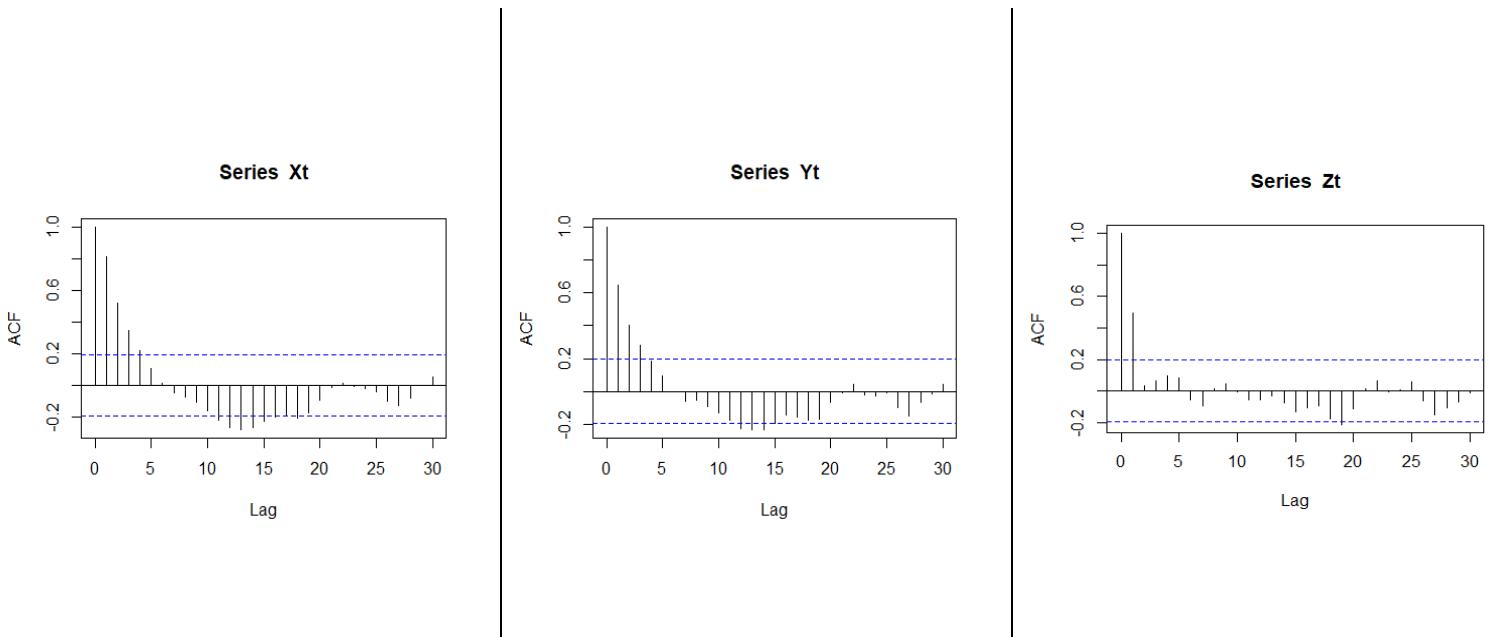
# Processus Yt: ARMA(1,0) ou AR(1):
# 1. ACF: Décroissance exponentielle plus rapide que celle d'un ARMA(1,1).
# 2. PACF: Présence d'un seul pic significatif.
# Bref, ACF avec nombre infini de pics qui décroît et PACF avec un seul pic.

# Processus Zt: ARMA(0,1) ou MA(1):
# 1. ACF: Présence d'un seul pic significatif. On voit qu'après le 2e délai,
# la valeur d'auto-corrélation devient nulle.
# 2. PACF: Décroissance exponentielle quant à la longueur des pics.
# Bref, ACF avec un seul pic et PACF avec nombre infini de pics qui décroît.

# En conclusion, si on se fie uniquement des graphiques d'ACF:
# ARMA(1,1): décroissance exponentielle
# AR(1): décroissance exponentielle plus rapide qu'ARMA(1,1)
# MA(1): valeurs non nulles dans les délais  $k = \{0,1\}$  seulement.

#####
# c) ACF Échantillonale des séries générées et comparaison aux valeurs théo.

```



```

## [1] "RÉSULTATS POUR ARMA(1,1)"
EXPÉIMENTALE
##
## Autocorrelations of series 'Xt', by lag
##
##      0      1      2      3      4      5      6      7      8      9      10
## 1.000  0.814  0.522  0.343  0.218  0.106  0.011 -0.046 -0.072 -0.105 -0.159
##      11     12     13     14     15     16     17     18     19     20     21
## -0.219 -0.266 -0.280 -0.266 -0.228 -0.198 -0.197 -0.207 -0.174 -0.092 -0.017
##      22     23     24     25     26     27     28     29     30
##  0.011 -0.008 -0.019 -0.039 -0.104 -0.130 -0.078 -0.001  0.050

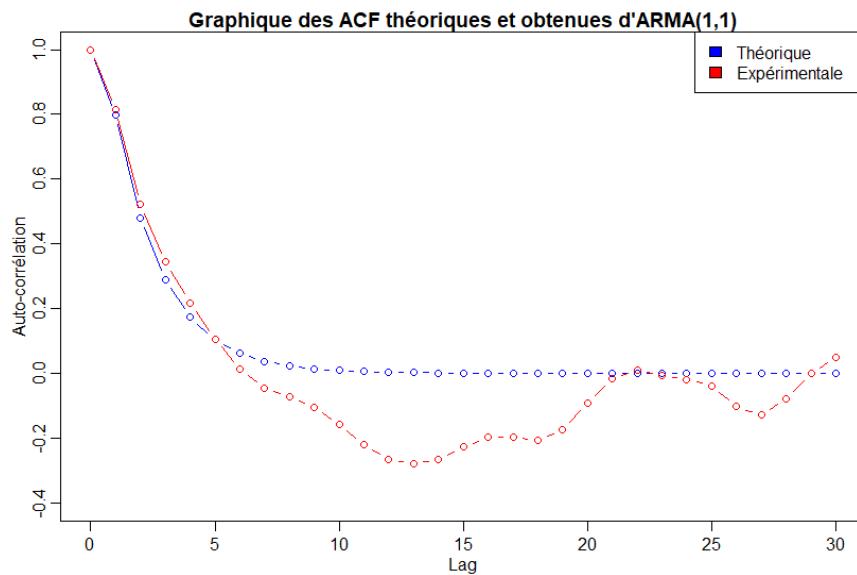
```

### THÉORIQUE

```

## [1] 1.000 0.799 0.480 0.288 0.173 0.104 0.062 0.037 0.022 0.013 0.008 0.005
## [13] 0.003 0.002 0.001 0.001 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000
## [25] 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000

```



```

## [1] "RÉSULTATS POUR AR(1)"
EXPÉIMENTALE
##
## Autocorrelations of series 'Yt', by lag
##
##      0      1      2      3      4      5      6      7      8      9      10
## 1.000  0.646  0.404  0.281  0.186  0.097  0.001 -0.060 -0.051 -0.087 -0.128
##      11     12     13     14     15     16     17     18     19     20     21
## -0.174 -0.224 -0.229 -0.231 -0.194 -0.144 -0.155 -0.177 -0.168 -0.064 -0.006
##      22     23     24     25     26     27     28     29     30
##  0.042 -0.019 -0.023 -0.004 -0.095 -0.146 -0.062 -0.015  0.046

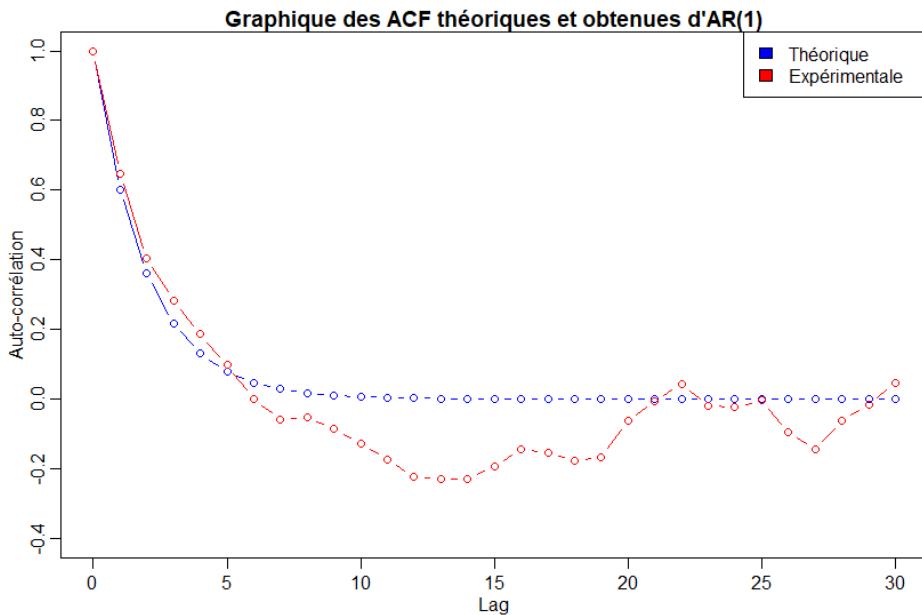
```

### THÉORIQUE

```

## [1] 1.000 0.600 0.360 0.216 0.130 0.078 0.047 0.028 0.017 0.010 0.006 0.004
## [13] 0.002 0.001 0.001 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000
## [25] 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000

```



```
## [1] "RÉSULTATS POUR MA(1)"
```

### EXPÉIMENTALE

```
##
```

```
## Autocorrelations of series 'Zt', by lag
```

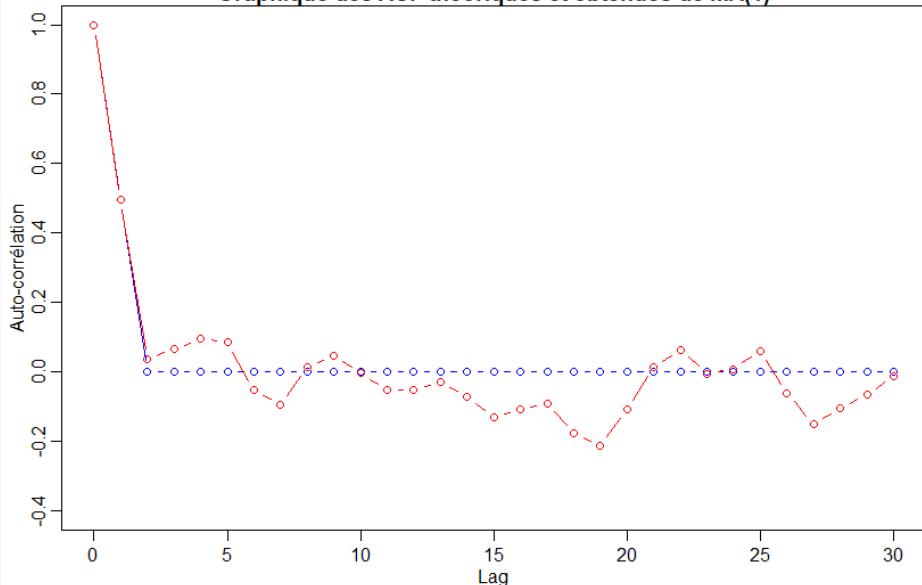
```
##
```

```
##   0    1    2    3    4    5    6    7    8    9    10
## 1.000 0.495 0.035 0.066 0.094 0.085 -0.053 -0.094 0.013 0.046 -0.005
##   11   12   13   14   15   16   17   18   19   20   21
## -0.054 -0.054 -0.029 -0.073 -0.131 -0.108 -0.092 -0.177 -0.213 -0.110  0.013
##   22   23   24   25   26   27   28   29   30
##  0.063 -0.006  0.008  0.058 -0.061 -0.151 -0.105 -0.066 -0.014
```

### THÉORIQUE

```
## [1] 1.000 0.497 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000
## [13] 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000
## [25] 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000
```

Graphique des ACF théoriques et obtenues de MA(1)



```
# Les calculs théoriques représentent quand même les données expérimentales
# à l'exception des derniers délais. Cela est causé par le fait qu'il n'y a pas
# assez de données. Si par exemple n = 10000, on verra que les courbes
# théoriques et expérimentales sont exactement identiques.
```

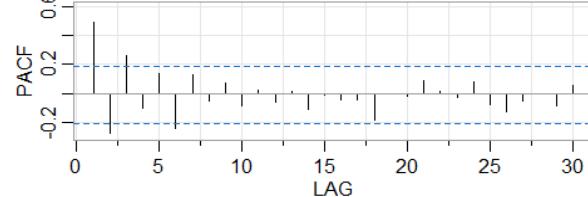
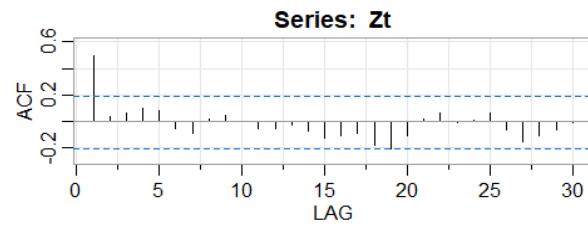
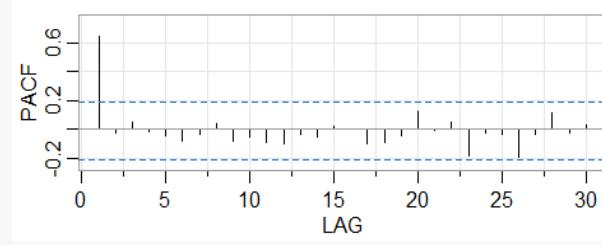
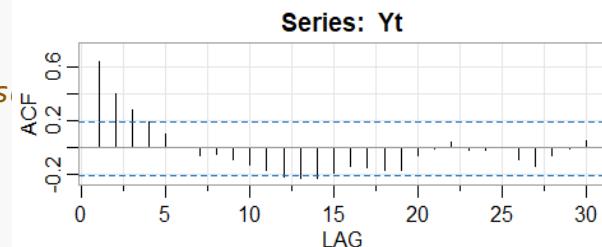
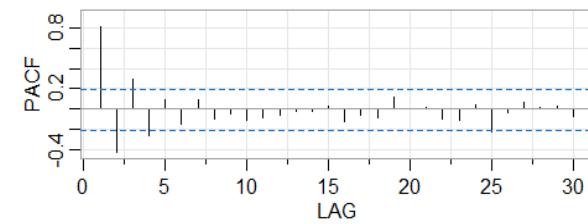
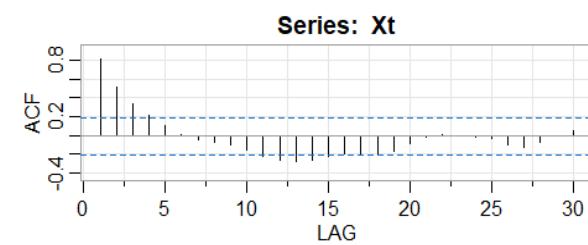
```
#####
# d) ACF, PACF, comparaisons et conclusion
```

```
# Modèle ARMA(1,1)
library(astsa)
acf2(Xt, max.lag=30)
# ACF: Nombre infini d'ACF
# PACF: Nombre infini de PACF
```

```
# Modèle AR(1)
acf2(Yt, max.lag=30)
# ACF: Nombre infini d'ACF (décroissant)
# PACF: Nombre fini (un seul pic)
```

```
# Modèle MA(1)
acf2(Zt, max.lag = 30)
# ACF: Nombre fini (un seul pic)
# PACF: Nombre infini de PACF
```

```
# En conclusion, toutes les observations
# confirment les résultats
# d'identification des modèles ARMA vus en
# classe.
```



## EXERCICE 4

```
#####
# ÉTAPES PRÉLIMINAIRES POUR L'ANALYSE DES DONNÉES
# 1) CUEILLETTE D'INFORMATIONS

# Les données sont tirées du site: https://www150.statcan.gc.ca/n1/fr/type/donnees
# Quelques informations pertinentes par rapport aux données cueillies:

# Numéro d'identification: 24-10-0002-01 (anciennement CANSIM 427-0002)
# Nom d'étude: Nombre de véhicules voyageant entre Le Canada et Les États-Unis
#           (section: Total des véhicules entrant au Canada)
# Période considérée: Janvier 2009 - Avril 2018

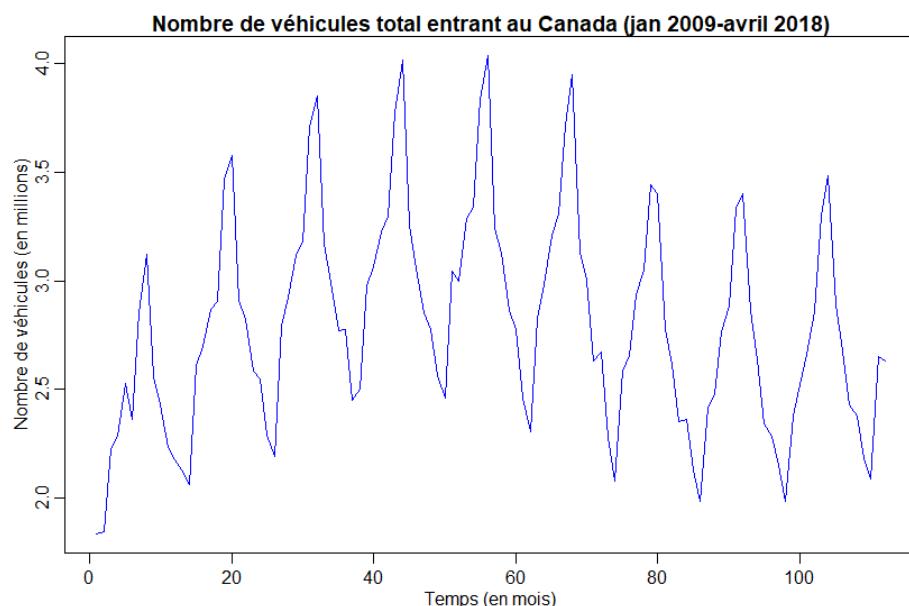
# On veut étudier le nombre total des véhicules entrant au Canada entre janvier
# 2009 et avril 2018. Étant donné qu'on reçoit des millions de véhicules par
# mois, l'analyse sera abordée selon les données recueillies, mais divisées par
# un million afin de travailler avec des nombres plus petits. Les conclusions
# demeurent identiques; le but est uniquement de faciliter la tâche à visualiser
# Les résultats en réduisant le nombre de chiffres dans une donnée.

# Nous avons créé un fichier nommé Data_Q4.csv qui contient toutes les données
# nécessaires. L'annexe A montrera le contenu de ce dernier.

data <- read.csv('Data_Q4.csv', sep=';')
mois <- data$Mois
y <- data$Qty
y_new <- data$Qty/1000000

write.csv(cbind(mois, y_new), 'Data_Q4_modified.csv')
data <- read.csv('Data_Q4_modified.csv', sep=',')
y <- data$y_new

# Premier aperçu des données recueillies
```



```

# Commentaires: La série chronologique est saisonnière de fréquence mensuelle.
# On peut voir que chaque année, le mois qui a le plus de véhicules qui entre
# au Canada est le mois d'août, ce qui a du sens étant donné que c'est la
# période des vacances.

# Selon le graphique, nous n'étions pas certains si la tendance est linéaire
# dans cette série chronologique, mais il y a une tendance peu importante.
# Nous avons essayé plusieurs tentatives pour éliminer cette tendance:
# La racine carrée, l'inverse, le logarithme, mais la tendance demeure
# présente. L'hypothèse sera de faire une différenciation deux fois puisqu'on
# voit une tendance qui ressemble à une fonction quadratique. Ensuite,
# on fait une différenciation saisonnière afin que la série soit plus
# stationnaire.

# 2) AJUSTEMENT DE CETTE SÉRIE AVEC LES MÉTHODES ARIMA

# Ajustement 1: Diviser les données par un million pour simplifier la tâche
#                  d'analyse
# Ajustement 2: Différenciation première (et double?) pour éliminer la tendance
# Ajustement 3: Différenciation saisonnière où s = 12 mois

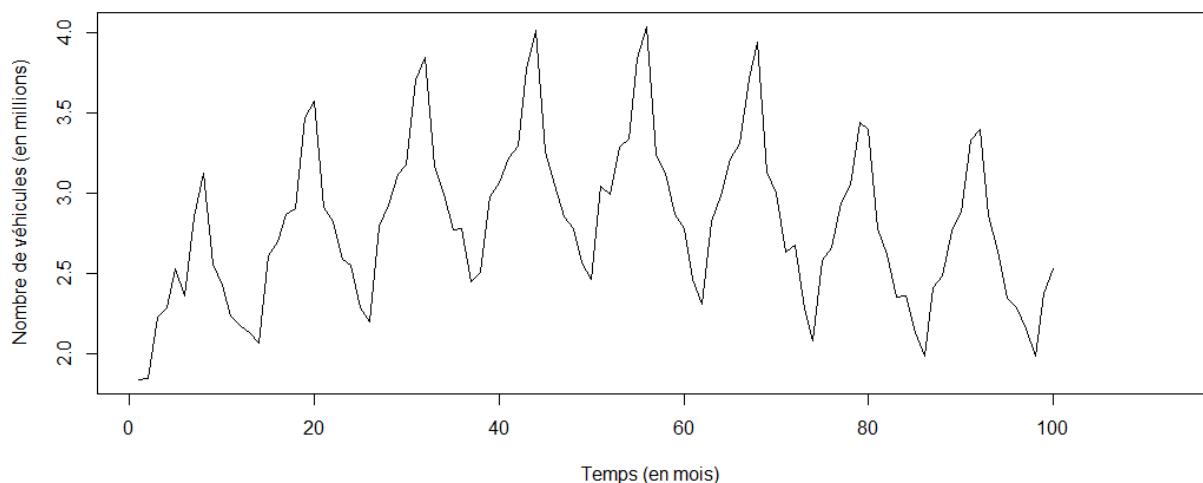
# 3) NOMBRE D'OBSERVATIONS

# Troncage des 12 dernières données, afin de faire l'analyse avec les 110
# premières observations

n <- 100
t <- seq(1, 100, 1)
y <- data$y_new[-seq(101, 112, 1)]

```

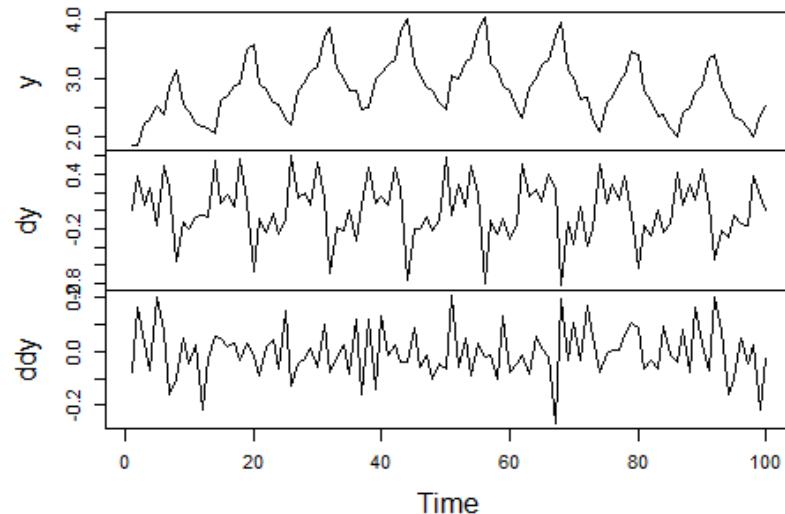
Nombre de véhicules total entrant au Canada (jan 2009-avril 2017)



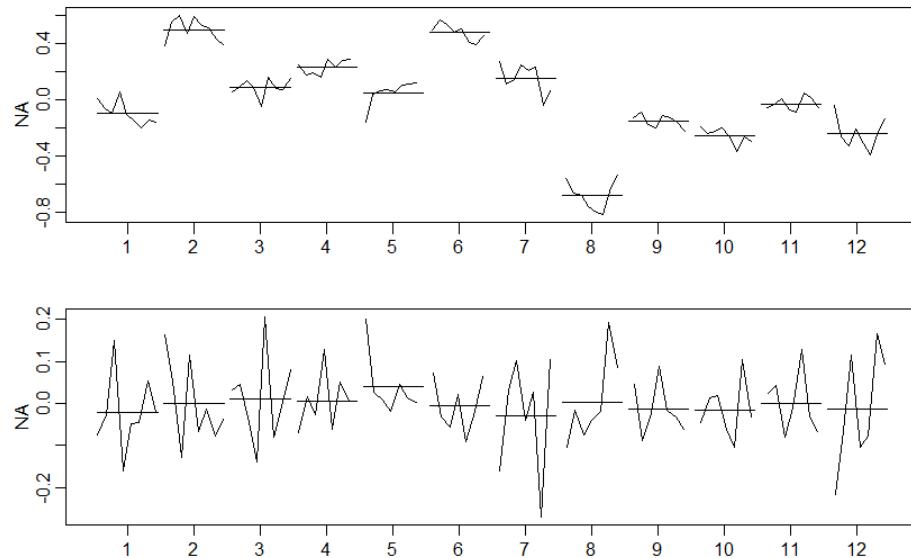
```
#####
# a) Transformation de La série

# En faisant plusieurs expérimentations, on rejette notre hypothèse initiale,
# c'est-à-dire de faire deux différenciations non saisonnières.
# Avec une seule différenciation, on observe que Les données semblent plus
# "stationnaires" et visuellement, nous avons éliminé la tendance.

# dy: données brutes différencierées une seule fois
# ddy: dy différencierées de façon saisonnière une seule fois
```



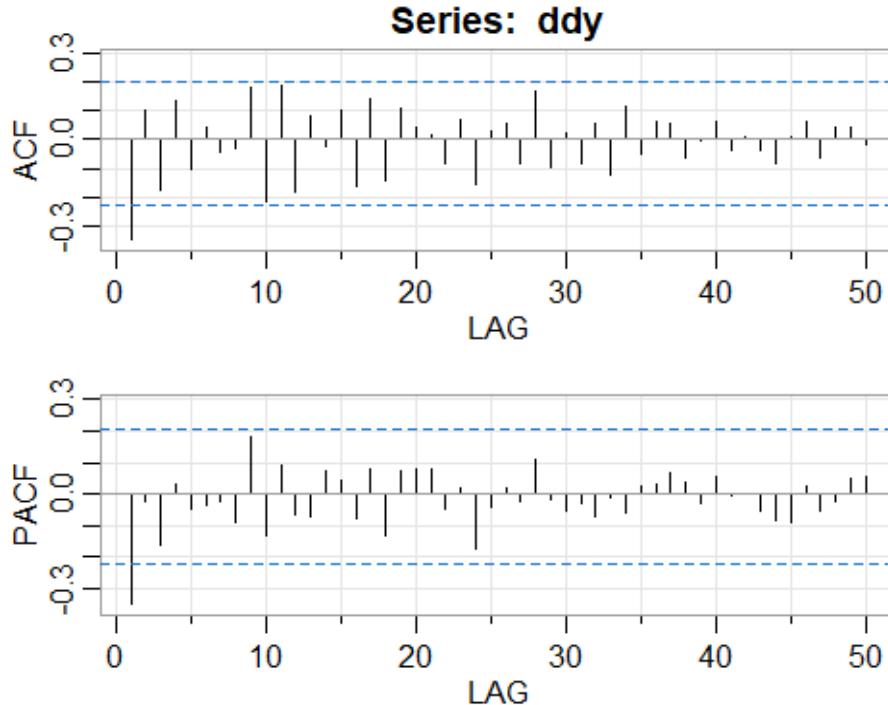
```
# On observe que La présence d'un motif pour chaque mois, donc il faut
# l'éliminer à L'aide de la différenciation saisonnière afin que les données
# semblent être stationnaires.
```



```
# Conclusion, on prend les données brutes qui sont différencierées
# de façon linéaire et saisonnière une fois chacune.
```

```
#####
# b) Les ordres du modèle choisi

# Commençons par analyser les ACF et les PACF du modèle
acf2(ddy, max.lag=50)
```



```
# COMPOSANTE NON SAISONNIÈRE (délais faibles entre 1 et 11):
# ACF: Observation d'un seul pic (ou) décroissance exponentielle des auto-
#       corrélations
# PACF: Observation d'un seul pic aussi (ou) décroissance exponentielle des
#       auto-corrélations

# Nous sommes certains la valeur de p ou de q ne dépassera pas de 1, donc nous
# avons trois modèles possibles à tester:
# ARIMA(1,1,0) - ARIMA(0,1,1) - ARIMA(1,1,1)

# COMPOSANTE SAISONNIÈRE (délais 12, 24, 36...):
# Dans le graphique des ACF et des PACF, c'est la partie la plus compliquée
# à remarquer puisque plusieurs phénomènes peuvent y arriver.

# Voici quelques hypothèses par rapport à la composante saisonnière du modèle:
# Le délai 12 dans ACF n'atteint pas vraiment la ligne "bleue", mais est proche,
# tandis que ce dernier n'est pas du tout atteint dans PACF.

# 1. Il se peut que le délai 12 est ignoré dans les deux graphiques, donc ça va
#       être ARIMA(0,1,0)_12.
# 2. Sinon, puisque le délai 12 a une auto-corrélation plus élevée dans ACF, il
#       se peut que ça soit un pic à considérer, donc ARIMA(0,1,1)_12 ou
#       ARIMA(1,1,1)_12 (peut considérer comme une décroissance exponentielle
#       dans le graphique ACF).
```

```

# Noter que les valeurs à côté de la ligne "sarima" sont les valeurs
# de log(vraisemblance) correspondante.

# Test 1: Avec ARMA(1,1) non saisonnier (modèles 1 à 3)
1. sarima(y, 1,1,1, 0,1,0, 12) #92.00 -- non-significatif
2. sarima(y, 1,1,1, 0,1,1, 12) #93.88 -- non-significatif
3. sarima(y, 1,1,1, 1,1,1, 12) #94.82 -- non-significatif

# Test 2: Avec AR(1) non saisonnier (modèles 4 à 6)
4. sarima(y, 1,1,0, 0,1,0, 12) #91.93 -- significatif
5. sarima(y, 1,1,0, 0,1,1, 12) #93.76 -- significatif
6. sarima(y, 1,1,0, 1,1,1, 12) #94.76 -- non significatif

# Test 3: Avec MA(1) non saisonnier (modèles 7 à 9)
7. sarima(y, 0,1,1, 0,1,0, 12) #91.96 -- significatif
8. sarima(y, 0,1,1, 0,1,1, 12) #93.86 -- significatif
9. sarima(y, 0,1,1, 1,1,1, 12) #94.80 -- non-significatif

# Dès le départ les modèles 1, 2, 3, 6 et 9 ne sont pas considérées puisqu'il
# existe des paramètres dont la p-valeur n'est pas du tout significative.

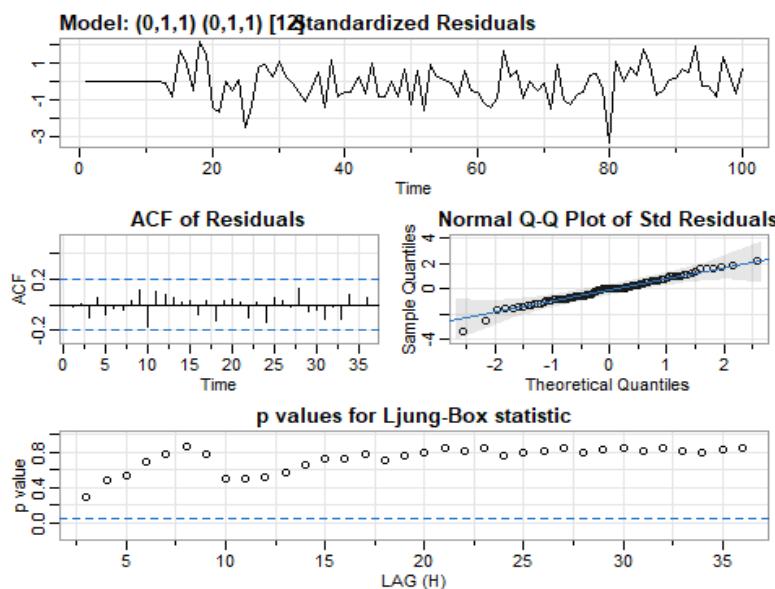
# Il reste à choisir un modèle parmi 4, 5, 7 et 8.
# La valeur log(vraisemblance) la plus élevée (AIC la moins élevée) est le
# modèle 8, donc on choisira ce dernier et l'analyser pour confirmer si
# c'est un bon choix.

# Modèle choisi: ARIMA(0,1,1) x (0,1,1)_12

#####
# c) Statistiques diagnostics et autres diagnostics
# Observations des tests et des résidus

# MODÈLE ARIMA(0,1,1) X (0,1,1)_12
sarima(y, 0,1,1, 0,1,1, 12)
## converged

```



```

## $fit
##
## Call:
## arima(x = xdata, order = c(p, d, q), seasonal = list(order = c(P, D, Q), period =
S),
##       include.mean = !no.constant, transform.pars = trans, fixed = fixed,
optim.control = list(trace = trc,
##                      REPORT = 1, reltol = tol))
##
## Coefficients:
##             ma1      sma1
##           -0.3226  -0.3164
## s.e.    0.1037   0.1554
##
## sigma^2 estimated as 0.006662:  log likelihood = 93.86,  aic = -181.72
##
## $degrees_of_freedom
## [1] 85
##
## $ttable
##     Estimate      SE t.value p.value
## ma1   -0.3226 0.1037 -3.1100  0.0025
## sma1  -0.3164 0.1554 -2.0362  0.0448
##
## $AIC
## [1] -1.854237
##
## $AICc
## [1] -1.852948
##
## $BIC
## [1] -1.77875

# AIC de -181.72 qui est le plus petit parmi les autres modèles testés.
# Log Likelihood de 93.86.

# Standardized Residuals: Les résidus sont majoritairement entre -3 et 3 comme
# dans la loi normale centrée réduite (sauf au temps t=80 où ça dépasse un peu
# -3). On peut dire que les résidus se comportent selon la loi normale.

# ACF of Residuals: Aucune présence d'auto-corrélation entre les résidus, donc
# le modèle est accepté.

# Ici, les résidus sont selon la loi normale de façon majoritaire. On peut
# voir un peu de déviations lors des extrémités, mais la plupart des données
# demeurent "normales".

# Ljung-Box: les p-valeurs se trouvent au-dessus de la ligne bleue, donc on ne
# rejette pas l'hypothèse nulle pour tous les délais. Donc, on peut prendre
# ce modèle pour faire des prévisions des 12 prochains mois.

```

```

# DISCUSSION DES P-VALEURS DES ESTIMATEURS ET DE LA VALEUR ALPHA:
# Il est vrai que les estimateurs ont des p-valeurs significatives, mais en
# ce qui concerne de smal: La p-valeur est 0.0448, ce qui est très proche de 5%.
# Par conséquent, le modèle choisi va dépendre de l'erreur du type I tolérée.
# Si par exemple on choisit un modèle avec un alpha de 2.5%, ce modèle ne sera
# plus valide. Dans ce cas-ci, on prendra le modèle 7 (car le modèle 5 a aussi
# un estimateur avec une p-valeur proche de 5%) qui est ARIMA(0,1,1)x(0,1,0)_12.

# Bref, ici, alpha = 5% et c'est pour cela qu'on prend le modèle choisi en b).
# Cependant, le modèle 5 doit être réservé si jamais on veut des prévisions plus
# précises (alpha < 5%).

#####
# d) Performance prévisionnelle

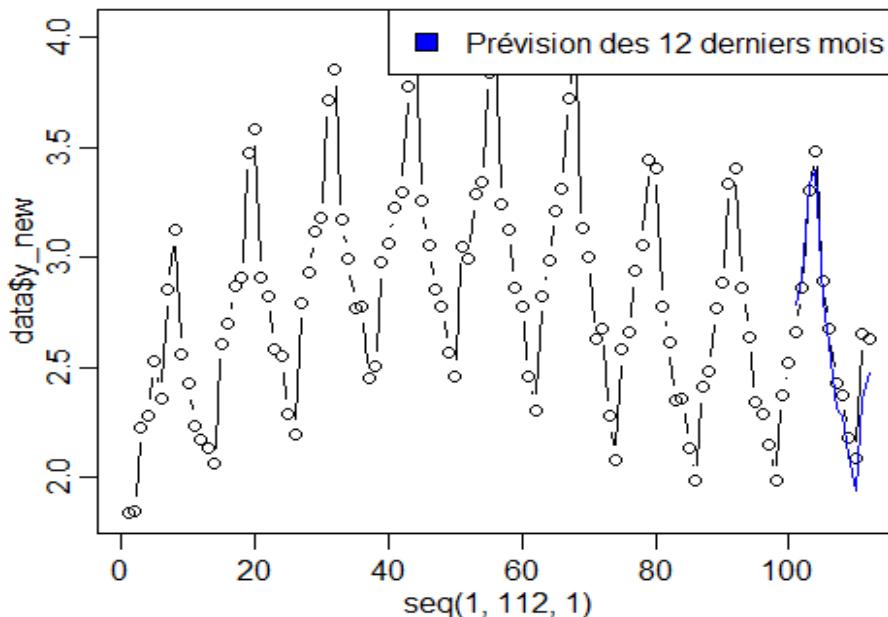
# Vraies valeurs des 12 derniers mois
last_y <- data$y_new[seq(101,112,1)]

# Valeurs prédites des 12 derniers mois
predict_model <- sarima.for(y, 12, 0,1,1, 0,1,1, 12)
predict_y <- predict_model$pred

# EQMP
EQMP <- mean((predict_y - last_y)^2)
EQMP
## [1] 0.01698084

# Ligne bleue: les prévisions
# Ligne noire: les données brutes

```



```

# Nous avons une EQMP de 0.01698084, ce qui est proche de zéro. Cela montre
# que le modèle choisi prédit assez bien les 12 prochains mois.

```

```

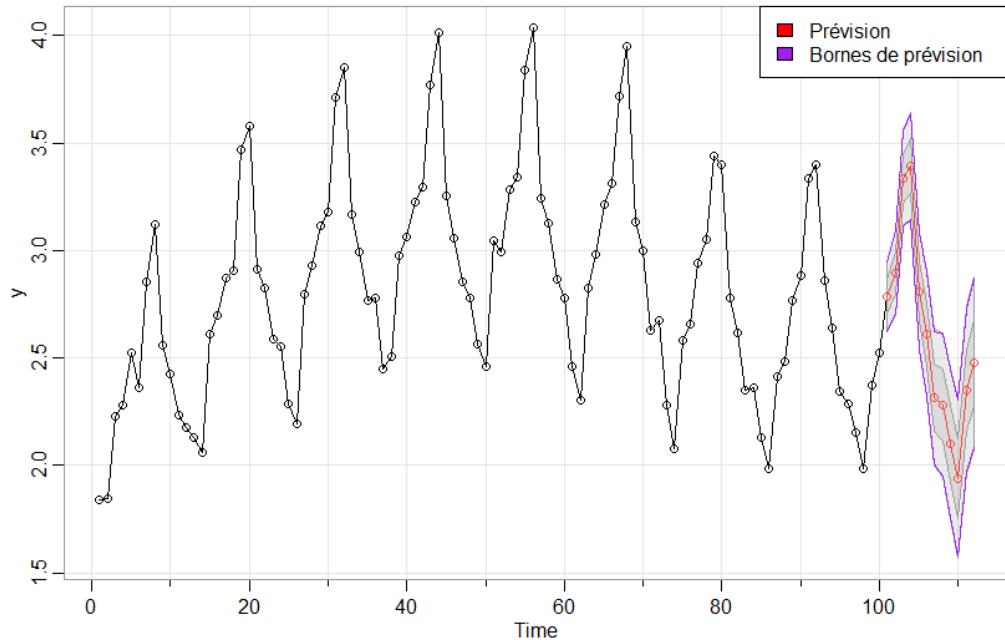
#####
# e) Intervalles de prévision

# En se servant de la formule  $y_{\text{hat}} \pm 1.96 * \text{se}$ , on obtient les intervalles
# de prévision suivants pour les 12 dernières données
borneInf <- predict_y - 1.96*predict_model$se
borneSup <- predict_y + 1.96*predict_model$se

cbind(borneInf, predict_y, borneSup)
## Time Series:
## Start = 101
## End = 112
## Frequency = 1
##      borneInf predict_y borneSup
## 101 2.624813 2.784791 2.944768
## 102 2.704049 2.897273 3.090497
## 103 3.112857 3.334393 3.555929
## 104 3.145944 3.392562 3.639181
## 105 2.540047 2.809423 3.078799
## 106 2.318592 2.608947 2.899302
## 107 2.005358 2.315275 2.625193
## 108 1.951728 2.280045 2.608361
## 109 1.756602 2.102340 2.448077
## 110 1.577291 1.939613 2.301934
## 111 1.972231 2.350411 2.728590
## 112 2.081813 2.475211 2.868609
#####
# f) Représentation graphique des prévisions

# Les lignes violettes représentent les bornes inférieures et supérieures
# des intervalles de prévisions.
sarima.for(y, 12, 0, 1, 1, 0, 1, 1, 12)

```



```

# Modèle choisi: ARIMA(0,1,1) X (0,1,1)12
#  $(1 - B^{12})(1 - B) Y_t = (1 - 0.3226B)(1 - 0.3164B^{12}) e_t$ 
# ou
#  $Y_t = Y_{t-1} + Y_{t-12} - Y_{t-13} + e_t - 0.3226 e_{t-1} - 0.3164 e_{t-12} + 0.10207064 e_{t-13}$ 
# où  $e = epsilon$ 

```

**ANNEXE A – CONTENU DU FICHIER *Data\_Q4.csv***

Mois	Qty		12-juil	3772442		16-févr	1985152
09-janv	1837530		12-août	4014665		16-mars	2413193
09-févr	1845623		12-sept	3256274		16-avr	2481273
09-mars	2227909		12-oct	3055174		16-mai	2765665
09-avr	2283439		12-nov	2851501		16-juin	2881952
09-mai	2525737		12-déc	2776945		16-juil	3334402
09-juin	2361955		13-janv	2566371		16-août	3399740
09-juil	2853322		13-févr	2459058		16-sept	2861710
09-août	3122978		13-mars	3045758		16-oct	2638655
09-sept	2560721		13-avr	2996044		16-nov	2341600
09-oct	2427607		13-mai	3285259		16-déc	2285715
09-nov	2235203		13-juin	3338322		17-janv	2151134
09-déc	2175527		13-juil	3837298		17-févr	1985971
10-janv	2130795		13-août	4038831		17-mars	2376086
10-févr	2063053		13-sept	3241223		17-avr	2524673
10-mars	2609264		13-oct	3128149		17-mai	2661197
10-avr	2696851		13-nov	2863926		17-juin	2859996
10-mai	2868900		13-déc	2777557		17-juil	3302694
10-juin	2905292		14-janv	2461195		17-août	3483738
10-juil	3470077		14-févr	2305347		17-sept	2890900
10-août	3578790		14-mars	2826228		17-oct	2675920
10-sept	2911655		14-avr	2982675		17-nov	2428578
10-oct	2824254		14-mai	3210645		17-déc	2376122
10-nov	2586615		14-juin	3310302		18-janv	2181782
10-déc	2551168		14-juil	3717591		18-févr	2087613
11-janv	2288391		14-août	3947167		18-mars	2653130
11-févr	2194255		14-sept	3130659		18-avr	2632503
11-mars	2793839		14-oct	3000814			
11-avr	2928143		14-nov	2630768			
11-mai	3115116		14-déc	2674248			
11-juin	3177884		15-janv	2279015			
11-juil	3711376		15-févr	2077112			
11-août	3851780		15-mars	2583410			
11-sept	3169819		15-avr	2658525			
11-oct	2993192		15-mai	2938559			
11-nov	2769486		15-juin	3051461			
11-déc	2776098		15-juil	3440653			
12-janv	2449101		15-août	3400486			
12-févr	2503953		15-sept	2776848			
12-mars	2974933		15-oct	2615971			
12-avr	3064577		15-nov	2350337			
12-mai	3224985		15-déc	2361467			
12-juin	3295533		16-janv	2132795			

Site consultée :

<https://www150.statcan.gc.ca/t1/tbl1/fr/tv.action?pid=2410000201&pickMembers%5B0%5D=1.1&pickMembers%5B1%5D=3.1&pickMembers%5B2%5D=4.1&cubeTimeFrame.startMonth=01&cubeTimeFrame.startYear=2009&cubeTimeFrame.endMonth=04&cubeTimeFrame.endYear=2018&referencePeriods=20090101%2C20180401>