ENSAE - 2ème année

Séries temporelles linéaires
Examen 2018-2019

Durée : 2 heures. Sans document et calculatrice.

Les exercices sont indépendants. Il est demandé de justifier les réponses de façon concise.

Exercice 1 (graphiques)

1. Pour chacune des trajectoires (a), (b), (c) et (d) de la figure 1, quelle est la transformation la plus adaptée pour rendre la série stationnaire :
   - la différentiation ordinaire $\nabla X_t = X_t - X_{t-1}$?
   - la différentiation saisonnière $\nabla_s X_t = X_t - X_{t-s}$ avec un s à préciser ?
   - aucune car la série semble déjà stationnaire ?

   La série (a) semble stationnaire, pour les séries (b) et (d) qui présentent une tendance linéaire, la differentiation ordinaire semble la plus adaptée, pour la série (c) la différentiation saisonnière avec $s = 12$ devrait faire disparaitre la saisonnalité.

2. Quelle est la définition de la fonction d'autocorrélation ? de la fonction d'autocorrélation partielle ? Sous quelles hypothèses ces fonctions sont-elles définies ? Quel modèle ARMA est compatible avec les autocorrélogrammes de la figure 2 ? Pour un processus stationnaire au second ordre (X_t), l'autocorrelation de retard h est la correlation entre X_t et X_{t-h}, l'autocorrélation partielle de retard $h > 0$ est la correlation entre X_t et X_{t-h} sachant les variable intermédiaires $X_{t-h+1}, \ldots, X_{t-1}$. Comme les autocorrélation partielles d'un AR(p) s'annulent à partir du retard $h > p$, la figure est compatible avec un AR(2).

3. La figure 3 représente les autocorrélations et autocorrélations partielles empiriques d'une série $X_1, \ldots, X_n$ avec $n = 60$. Comment sont calculées

les autocorrélations empiriques ? Que représentent les pointillés sur la figure ? L'autocorrélation empirique de retard $0 \leq h < n$ est définie par

$$\hat{\rho}_n(h) = \frac{1}{n} \sum_{t=1+h}^{n} (X_t - \overline{X}_n)(X_{t-h} - \overline{X}_n), \quad \overline{X}_n = \frac{1}{n} \sum_{t=1}^{n} X_t.$$

Si les X_t sont iid de variance finie, d'après le TCL on sait que

$$\boxed{P(\hat{\rho}_n(h) \in \pm 1.96/\sqrt{n}) \simeq 0.95.}$$

Les pointillés figurent les limites $1.96/\sqrt{n}$ de cette zone dans laquelle les autocorrélations empiriques de données iid se trouvent avec un probabilité de 95% lorsque n est grand.

4. Quel modèle ou quels modèles ARMA semblent le plus plausibles pour la série dont les autocorrélogrammes sont donnés par la figure 3 ?D'après ce que l'on vient de dire, il est plausible que la série soit réalisation d'un AR(2).

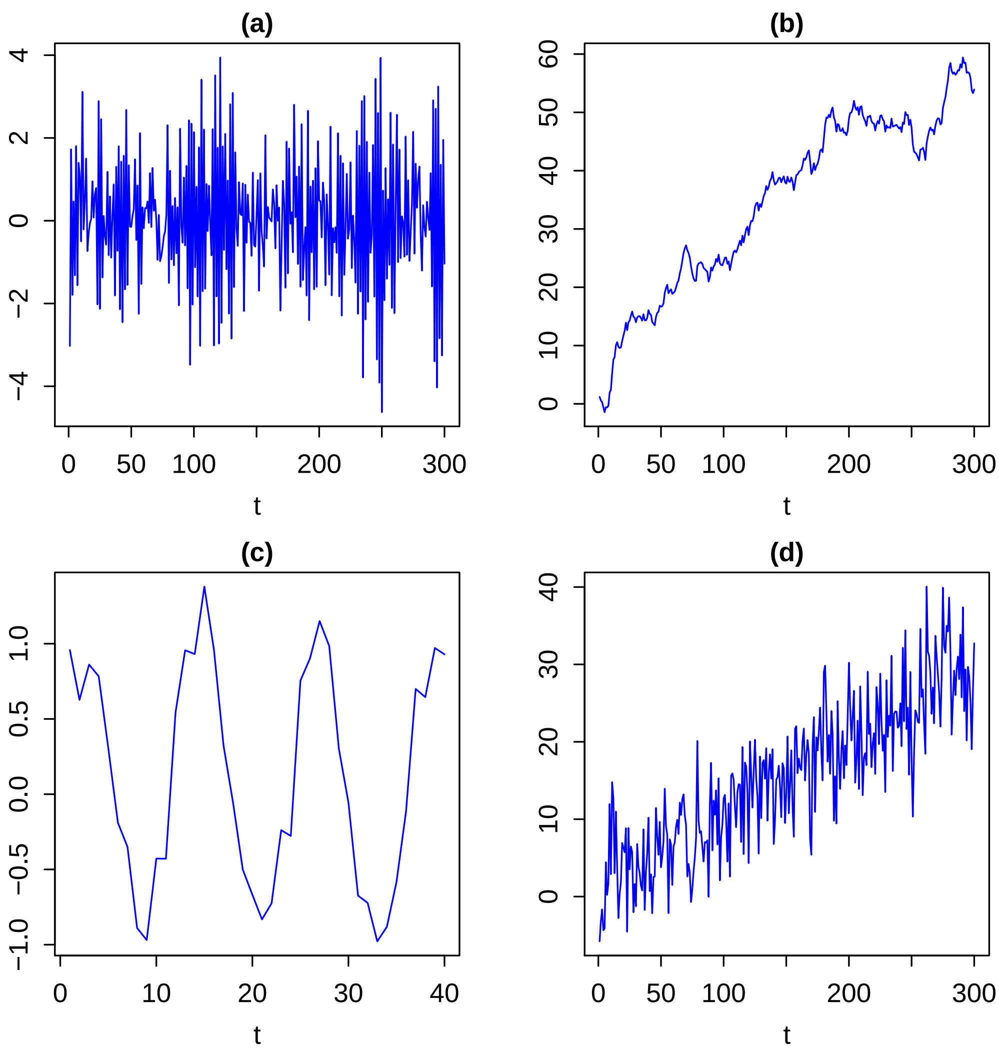

FIGURE 1 – Trajectoire de 4 séries.

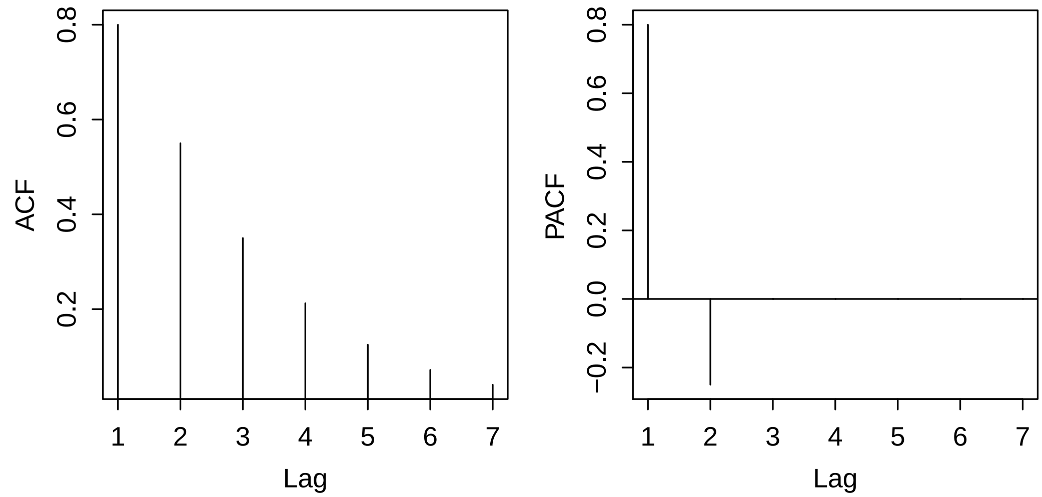

Figure 2 – Autocorrelations (ACF) et autocorrélations partielles (PACF) théoriques d'un processus.

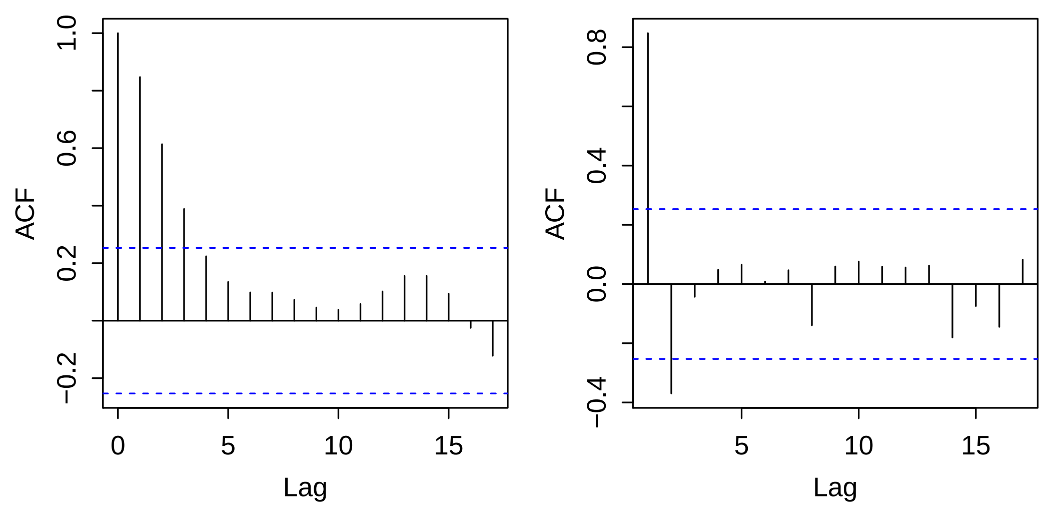

Figure 3 – Autocorrelations (ACF) et autocorrélations partielles (PACF) empiriques d'une série.

Exercice 2 (Causalité au sens de Granger) Soit $\epsilon_t = (\epsilon_{1t}, \epsilon_{2t}, \epsilon_{3t})'$ un bruit blanc de matrice de variance

$$\Sigma = \begin{pmatrix} \times & 0 & \\ & \times & \\ & & \times \end{pmatrix}.$$

où "0" représente un zéro, "$\times$" représente un nombre non nul et certains éléments de la matrice ne sont pas spécifiés. Soit $X_t = (X_{1t}, X_{2t}, X_{3t})'$ satisfaisant un modèle VAR(1) stationnaire de la forme

$$X_t = AX_{t-1} + \epsilon_t, \quad A = \begin{pmatrix} \times & & 0 \\ & \times & \\ & & \times \end{pmatrix}.$$

1. Compléter la matrice A et/ou Σ par des "0" et/ou des "$\times$" pour que, au sens de Granger,

   — X_{1t} ne cause pas (X_{2t}, X_{3t}) ;
   — X_{2t} cause (X_{1t}, X_{3t}) ;
   — X_{3t} cause (X_{1t}, X_{2t}) ;
   — (X_{2t}, X_{3t}) cause X_{1t}.

   D'après l'écriture VAR, à la date $t-1$ la prévision optimale de (X_{2t}, X_{3t}) est
$$\begin{pmatrix} \times & a_{23} \\ a_{32} & \times \end{pmatrix} \begin{pmatrix} X_{2,t-1} \\ X_{3,t-1} \end{pmatrix} + \begin{pmatrix} a_{21} \\ a_{31} \end{pmatrix} X_{1t-1}.$$
   Pour que $X_{1,t-1}$ n'intervienne pas dans cette précision, c'est-à-dire pour que X_{1t} ne cause pas (X_{2t}, X_{3t}), il faut donc que $a_{21} = a_{31} = 0$. Par le même raisonnement, pour que X_{2t} cause (X_{1t}, X_{3t}) il faut que a_{12} ou a_{32} soit non nul. Pour que X_{3t} cause (X_{1t}, X_{2t}) il faut que a_{13} ou a_{23} soit non nul, mais comme $a_{13} = 0$, on doit avoir $a_{23} = \times$. Pour que (X_{2t}, X_{3t}) cause X_{1t}, il faut $a_{12} = \times$, puisque $a_{13} = 0$. Finalement

$$A = \begin{pmatrix} \times & \times & 0 \\ 0 & \times & \times \\ 0 & & \times \end{pmatrix}$$

   avec a_{32} quelconque.

2. Compléter la matrice A et/ou Σ par des "0" et/ou des "$\times$" pour que, au sens de Granger,
   - il y ait causalité instantanée entre X_{1t} et (X_{2t}, X_{3t}) ;
   - il y ait pas de causalité instantanée entre X_{2t} et (X_{1t}, X_{3t}).

   Pour qu'il n'y ait pas de causalité instantanée entre X_{2t} et (X_{1t}, X_{3t}), il faut $\Sigma(2,1) = \Sigma(2,3) = 0$ et par symétrie $\Sigma(1,2) = \Sigma(3,2) = 0$. Pour qu'il y ait causalité instantanée entre X_{1t} et (X_{2t}, X_{3t}), il faut $\Sigma(1,3) = \times$, car $\Sigma(1,2) = 0$. On a aussi $\Sigma(1,3) = \times$. Finalement

$$\Sigma = \begin{pmatrix} \times & 0 & \times \\ 0 & \times & 0 \\ \times & 0 & \times \end{pmatrix}.$$

Exercice 3 (Cointégration ou spurious régression) Soit (ϵ_t) et (η_t) deux bruits blancs forts indépendants et le système

$$\begin{cases} Y_t = aX_t + b + \eta_t \\ X_t = cX_{t-1} + d + \epsilon_t. \end{cases}$$

1. Donner des valeurs à a, b, c, et d pour que le système soit cointégré.
   Si $c = 1$ alors X_t est I(1) et si $a \neq 0$ la première équation est bien une relation de cointégration.
2. Fabriquer un exemple semblable de système ou la régression de Y_t sur X_t soit spurious (fallacieuse). Avec un système de la forme

$$\begin{cases} Y_t = Y_{t-1} + b + \eta_t \\ X_t = X_{t-1} + d + \epsilon_t, \end{cases}$$

   la régression de Y_t sur X_t n'aurait aucun sens. Même si les variables sont indépendantes, ~~comme elles sont I(1), le beta de cette régression pourrait sembler significativement non nul.~~

Exercice 4 (Processus stationnaire ou à racine unité) Soit (ϵ_t) et (η_t) deux bruits blancs forts indépendants et deux processus (X_t) et (Y_t) satisfaisant

— $X_t = aX_{t-1} + b + \epsilon_t$,

— $Y_t = X_t + \eta_t$ pour tout t.

1. Donner une condition sur a et/ou b pour que (X_t) soit un AR(1) stationnaire causal. Quel est alors le modèle suivi par (Y_t) ? Si $|a| < 1$, on sait que (X_t) est un AR(1) stationnaire causal, dont la fonction d'autocorrélation satisfait $\rho_X(h) = a\rho_X(h-1)$ pour tout $h > 0$. La fonction d'autocorrélation de (Y_t) satisfait $\rho_Y(h) = \rho_X(h) = a\rho_X(h-1)$ pour tout $h > 0$, donc $\rho_Y(h) = a\rho_Y(h-1)$ pour tout $h > 1$, ce qui montre que (Y_t) est un ARMA(1,1).

2. On suppose maintenant que l'on observe $Y_1, \ldots, Y_n$, mais pas les X_t. Peut-on tester l'hypothèse nulle $H_0 : a = 1$? par quel type de test ? est-il plus facile de le faire quand $b \neq 0$ ou quand quand $b = 0$? Sous H_0 on a $Y_t = X_{t-1} + b + \epsilon_t + \eta_t = Y_{t-1} + b + u_t$ avec $u_t = \epsilon_t + \eta_t - \eta_{t-1}$ qui suis une MA(1) (puisque $\rho_u(h) = 0$ pour $h > 1$). On peut donc faire le test de racine unité sur les Y_t (en utilisant en principe Perron-Phillips plutôt que ADF car le terme d'erreur n'est pas un AR). Si $b \neq 0$, sous la nulle on un modèle a racine unité avec une tendance déterministe et sous l'alternative un ARMA(1,1) stationnaire, ce qui doit être facile à distinguer sur la trajectoire, même sans test.