Feuille d'exercices

Exercice 1 (Autocorrélation et tendance linéaire)**.**
On s'intéresse à l'impact que peut avoir l'estimation de l'autocorrélation à partir d'une série temporelle ayant une tendance linéaire marquée.

a) Rappelez l'expression de l'autocorrélation empirique.

b) On considère la série temporelle (purement déterministe) suivante

$$X_t = at + b, \qquad t \in \mathbb{Z},$$

que l'on observe en $t = 0, \ldots, n$.
Des calculs simples (mais un peu pénible pour les faignants comme moi !) conduisent à montrer que pour $h \in \mathbb{N}$ fixé, on a $\hat{\rho}(h) \to 1$ lorsque $n \to \infty$. Comment interprétez vous ce résultat ?
Conseil pour les courageux : $\sum_{j=1}^{n} j^2 = n(n+1)(2n+1)/6$.

c) Lisez la documentation de la fonction `acf` de `R`. Illustrez numériquement la convergence de la première question.

d) Reprenez la question précédente mais en rajoutant à la série purement déterministe un bruit blanc gaussien de variance que vous ferez varier. Que constatez vous ?

e) On considère désormais la série

$$X_t = \sin\left(\frac{2\pi t}{\omega}\right), \qquad t \in \mathbb{Z}.$$

A l'aide de `R`, générez cette série temporelle sur une grande période et tracez sa fonction d'autocorrélation. Vous choisirez une fréquence ω de votre choix. Que constatez vous ? Est en accord avec l'ACF des données $'aaaa \cdots hhhh'$ vues en cours ?

✎ ✎ ✎

Exercice 2 (Causalité et inversibilité)**.** On s'intéresse à la série temporelle suivante

$$X_t = 0.4X_{t-1} + 0.45X_{t-2} + \varepsilon_t + \varepsilon_{t-1} + 0.25\varepsilon_{t-2}.$$

a) Est ce un processus $ARMA$? Si oui, déterminez ses polynômes $\theta(B)$ et $\phi(B)$.

b) Ce processus est-il causal ?

c) Ce processus est-il inversible ?

d) Donnez la représentation causale de cette série temporelle. Comparez vos résultats à ceux obtenus via `R` à l'aide de la fonction `ARMAtoMA`

e) Donnez la représentation inversible de cette série temporelle. Comparez vos résultats à ceux obtenus via `R` à l'aide de la fonction `ARMAtoMA`. *Astuce : il suffit juste d'intervertir les polynômes dans l'utilisation de la fonction ;-)*

✎ ✎ ✎

Exercice 3 (Vraisemblance d'un $AR(1)$)**.** On s'intéresse ici à la série temporelle

$$X_{t+1} = \mu + \phi(X_t - \mu) + \varepsilon_{t+1}, \qquad t \in \mathbb{Z}, \quad |\phi| < 1,$$

où $\varepsilon_1, \varepsilon_2, \dots$ sont des v.a.i.i.d. $N(0, \sigma^2)$. On suppose de plus que l'on dispose de n observations $x_1, \dots, x_n$ de cette série temporelle.

a) Quels sont les paramètres du modèle statistique considéré ici ?

b) Montrez que pour tout $t \in \mathbb{Z}$ on a la propriété Markovienne

$$\Pr(X_{t+1} \leq a_{t+1} \mid X_t = a_t, \dots, X_0 = a_0) = \Pr(X_{t+1} \leq a_{t+1} \mid X_t = a_t).$$

c) En déduire que la vraisemblance issue des observations $x_1, \dots, x_n$ est donnée par

$$f(x_1) f(x_2 \mid x_1) f(x_3 \mid x_2) \cdots f(x_n \mid x_{n-1}).$$

d) Donnez la loi de $X_{t+1} \mid X_t = x_t$.

e) En utilisant la réprésentation causale de cet $AR(1)$, montrez que x_1 est une réalisation d'une $N\{\mu, \sigma^2/(1-\phi^2)\}$. Faites le lien avec le slide du cours portant sur la modélisation de la température corporelle d'un castor.

f) Ecrivez un code `R` permettant d'ajuster cet $AR(1)$ par maximum de vraisemblance (via optimisation numérique) et comparez vos résultats à ceux de la fonction `(s)arima`. Vous ajusterez ce modèle sur le jeu de données `lh` inclus dans `R`.

✎ ✎ ✎

Exercice 4 (Un peu de lecture)**.** Allez sur la page web `https://www.stat.pitt.edu/stoffer/tsa4/Rissues.htm` et lisez là. C'est très instructif. Forcez vous à bien comprendre les incohérences que l'auteur pointe.

✎ ✎ ✎

Exercice 5 (Modélisation de la température globale)**.** Dans cet exercice, on va tenter de modéliser la température globale terrestre.

a) Installez la librairie `astsa` de `R` fournissant le jeu de données `gtemp` que nous allons utiliser. Renseignez vous sur ces données en lisant sa documentation.

b) Commencez par tenter d'identifier l'ordre de votre modèle $ARIMA$.

c) Ajustez votre modèle mais considérez également d'autres modèles $ARIMA$ concurrents.

d) Faites un tableau (synthétique) reportant l'AIC pour chacun des modèles ajustés considérés comme par exemple le Tableau 1.

e) Analysez les résidus de votre meilleur modèle et en déduire si ce dernier est satisfaisant.

f) Faites une prédiction pour les 50 prochaines années pour votre meilleur modèle.

✎ ✎ ✎

Table 1 – AIC pour différents modèles $ARIMA$, d fixé à ? ? ?, sur les données de température globalte.

$p \setminus q$	0	1	2	3	4
0					
1					
2					
3					
4					