Examen du 23/01/2018

MAP-STA2 : Séries chronologiques

Yannig Goude - yannig.goude@edf.fr; M. Zaouche - mounia.zaouche@gmail.com;

Exercice 1

1. Soit X_t un processus stationnaire du second ordre, ε_t suite de v.a. iid centrée de variance σ^2. On fait l'hypothèse H_a que ce processus est un AR(1):

$$X_t = aX_{t-1} + \varepsilon_t$$

avec $|a| < 1$.

1. calculer l'autocorrélation d'ordre 1 : $\rho(1)$.
2. supposons qu'on observe une trajectoire de taille n de ce processus. Définir l'estimateur empirique a_n de a. Quelle est la limite de a_n quand n tend vers l'infini?
3. suggérer un test pour valider l'hypothèse H_a.
4. le processus est en fait un processus AR(2):

$$X_t = \alpha X_{t-1} + \beta X_{t-2} + \varepsilon_t$$

montrer que pour que ce processus puisse se décomposer en une somme de valeur passées de ε_t il est nécessaire d'avoir $|\beta| < 1$.

5. écrire les équations de Yule-Walker reliant $\rho(1)$, $\rho(2)$, α et β. Exprimer $\rho(1)$ et $\rho(2)$ en fonction de α et β.

Exercice 2

1. Etudiez la stationnarité des processus suivants (ε_t est un bruit blanc de variance σ^2):
2. $y_t = at + \varepsilon_t$
3. $x_t = \Delta y_t$
4. $z_t = at^2 + bt + \varepsilon_t$
5. $z_t = \Delta^2 yt$

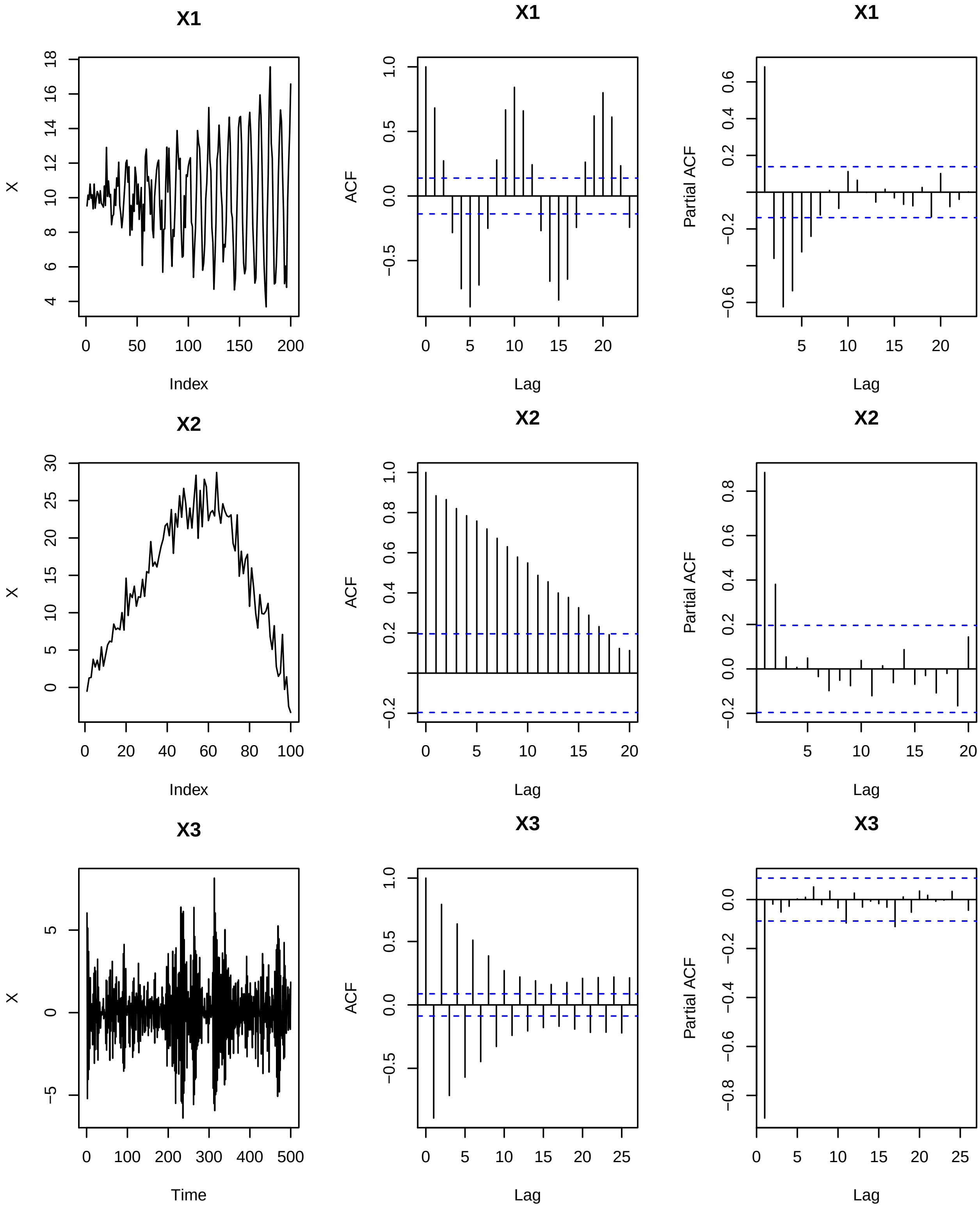
Exercice 3

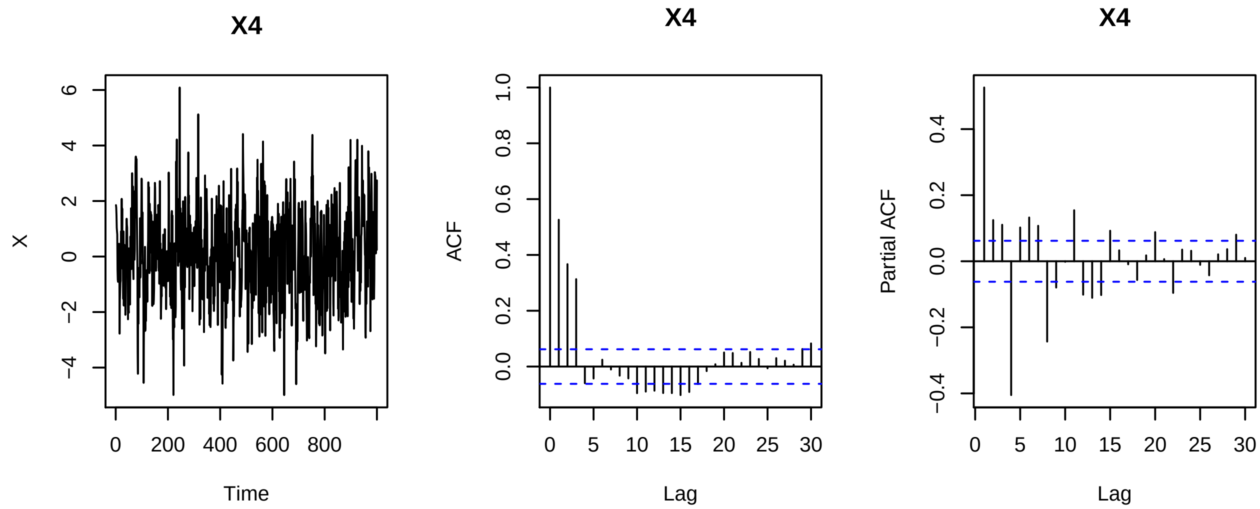
Soit Z_t une suite de variables aléatoires iid de loi normale de moyenne 0 et de variance σ^2. a, b et c des constantes. Déterminer lequel ou lesquels des processus ci-dessous sont stationnaires. Pour chaque processus stationnaire, calculer la moyenne et la fonction d'autocovariance.

1. $X_t = a + Z_t + cZ_{t-2}$
2. $X_t = Z_1 \cos(ct) + Z_2 \sin(ct)$
3. $X_t = Z_t \cos(ct) + Z_{t-1} \sin(ct)$
4. $X_t = Z_t Z_{t-1}$

Exercice 4

Un statisticien étudie un jeux de données composé de 4 séries temporelles pour lesquelles il a représenté/calculé les statistiques suivantes:





1. Proposer une démarche de modélisation pour chacune de ces séries, justifier.

Notre statisticien s'intéresse ensuite à une autre série X_t^5. Il propose de la modéliser par un ARMA et cherche ensuite à valider son modèle. Il obtient les sorties suivantes:

```
##
## Call:
## arima(x = X5, order = c(3, 0, 4), include.mean = T, method = c("ML"))
##
## Coefficients:
##           ar1     ar2     ar3     ma1      ma2     ma3     ma4  intercept
##       -0.2995  0.4580  0.2501  0.9064  -0.0054  0.0914  0.2061     0.2204
## s.e.   0.2213  0.1404  0.1866  0.2183   0.1672  0.1907  0.1047     0.4837
##
## sigma^2 estimated as 5.183:  log likelihood = -673.01,  aic = 1364.02
```

```
## pvalue student-test:
```

```
##       ar1       ar2       ar3       ma1       ma2       ma3       ma4
##      0.18      0.00      0.18      0.00      0.97      0.63      0.05
## intercept
##      0.65
```

2. Expliquer comment sont estimés les coefficients du modèle ARMA?
3. Expliquer ce que signifie la sortie "aic=1364.02".
4. A quoi correspond la ligne "s.e"? Comment est calculée la p-value du test de student de nullité de chacun des coefficients?
5. Au vu de ces résultats que doit faire notre statisticien?
6. Il s'intéresse ensuite à une série X_t^6 et effectue un test de Box-Pierce sur $X6_t$. Il obtient le résultat suivant:

```
##  Lags Statistic df     pvalue
##    20  29.13311 20 0.08517126
```

à quoi correspond "Lags"? "df"? Quelle est la conclusion de ce test?