Examen du 03/04/2025

MAP-STA2 : Séries chronologiques

Yannig Goude [yannig.goude@edf.fr]
Guillaume Principato [guillaume.principato@edf.fr]

Exercice 1

1. Donner la définition du lissage exponentiel simple. De quel problème d'optimisation l'estimateur obtenu est-il la solution (à justifier)?
2. Donner la définition du lissage exponentiel double. De quel problème d'optimisation l'estimateur obtenu est-il la solution, (à justifier)?
3. Quel type de lissage exponentiel doit-on utiliser pour modéliser une série temporelle possédant une tendance et une saisonnalité?

Exercice 2

1. Donner la définition de la stationnarité forte et faible. Expliquer en pratique comment on vérifie cette dernière.
2. Donner 3 exemples de processus stationnaires et justifier.
3. Soit $(X_t)_{t\in\mathbf{Z}}$ un processus stationnaire d'espérance μ et de fonction d'autocovariance $\gamma(h)$, $(a_i)_{i\in\mathbf{Z}}$ une suite de nombres réels absolument sommable $\sum_{i\in\mathbf{Z}}|a_i| < \infty$, montrer que $Y_t = \sum_{i\in\mathbf{Z}} a_i X_{t-i}$ est un processus stationnaire.

Exercice 3

Soit X un AR(1) stationnaire défini ainsi:

$$\forall t \in Z, X_t = \phi X_{t-1} + \varepsilon_t$$

avec ε_t un bruit blanc centré de variance σ^2. et $|\phi| < 1$.

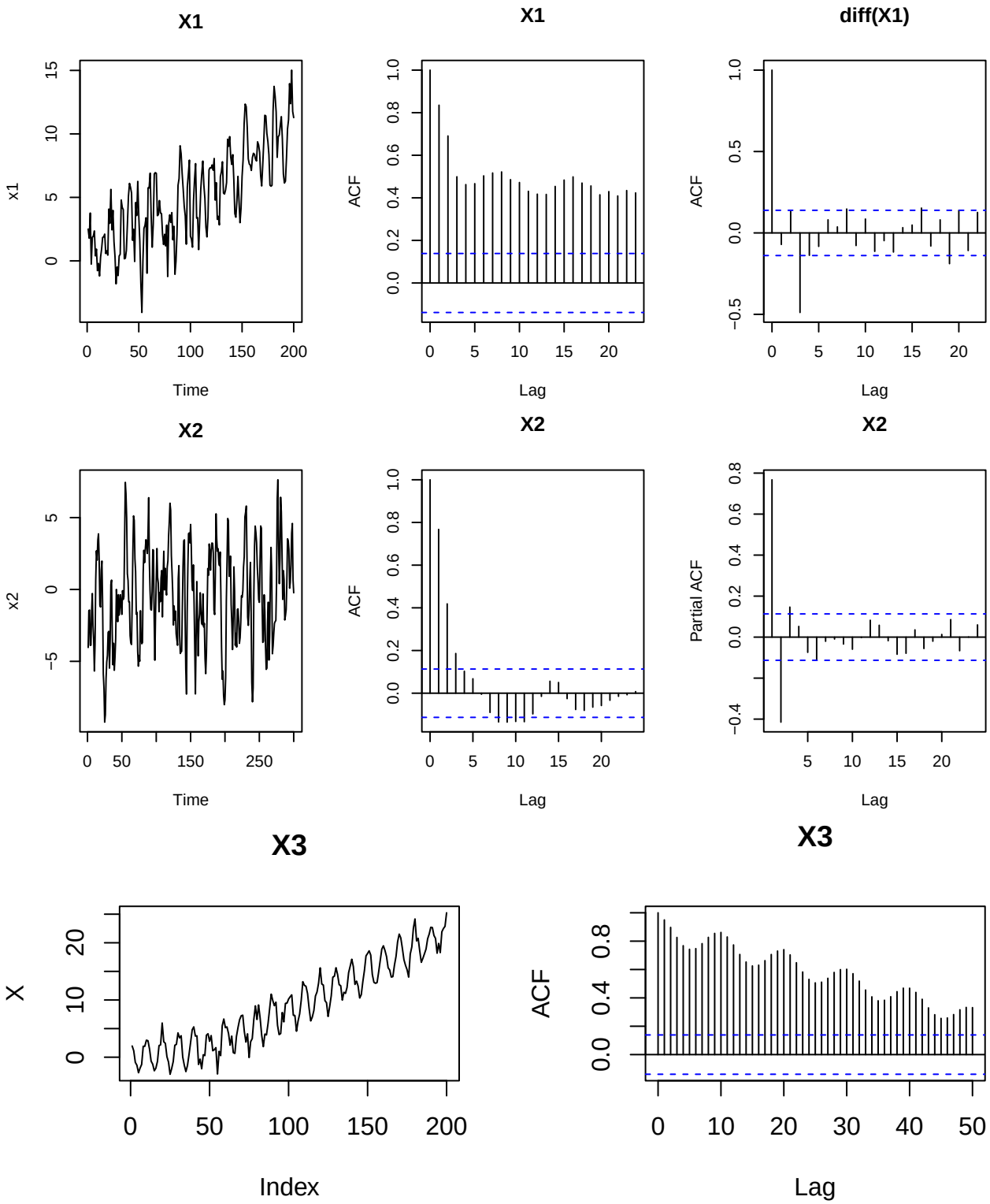
Soit ensuite le processus Y:

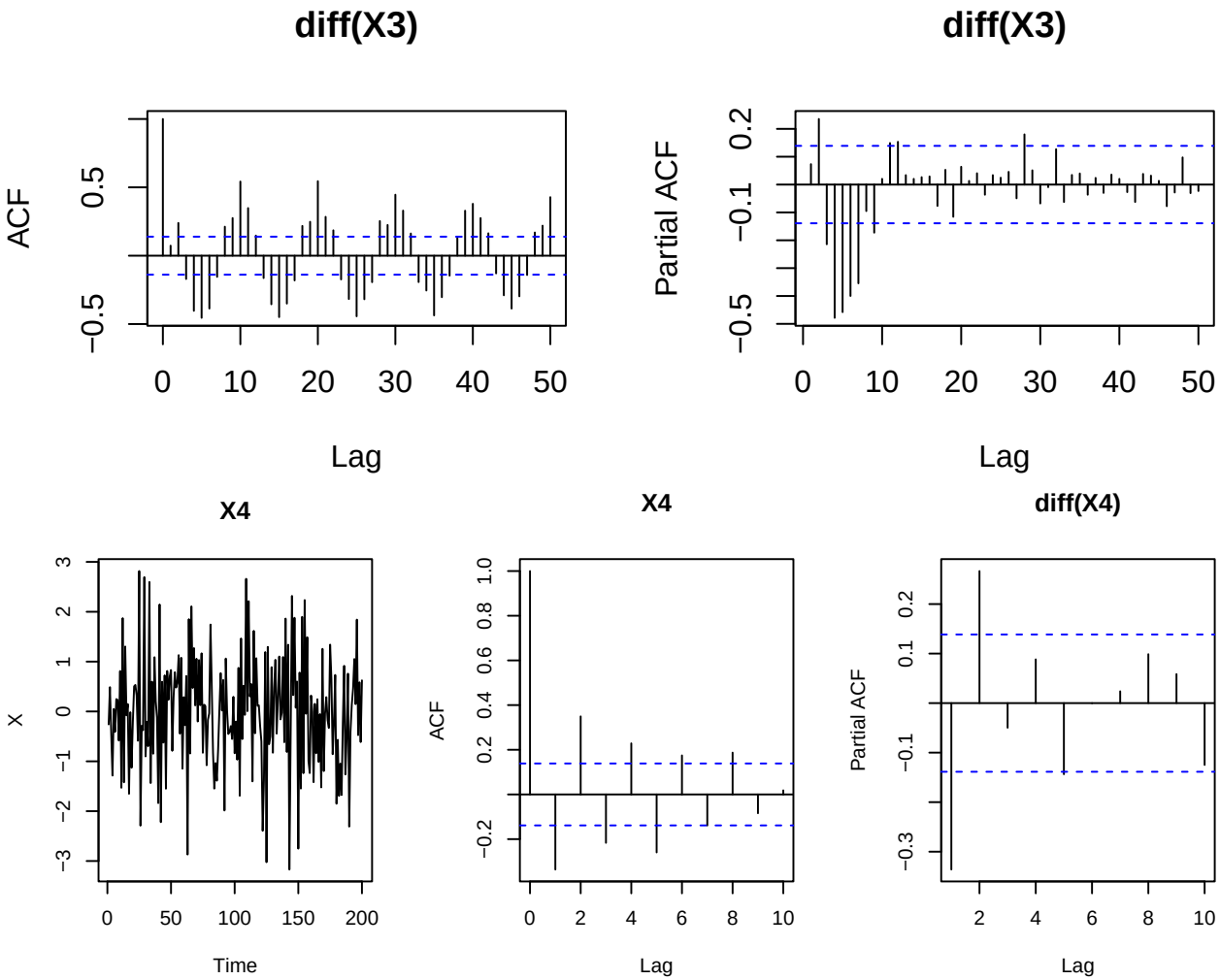
$$\forall t \in Z, Y_t = X_t - X_{t-1}$$

1. Montrer que Y est un processus stationnaire centré. Donner l'équation de récurrence vérifiée par Y et le nom exact de ce processus?
2. Calculer la variance de Y_t.

Exercice 4

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


X1
x1
Time
X1
ACF
Lag
diff(X1)
ACF
Lag
X2
x2
Time
X2
ACF
Lag
X2
Partial ACF
Lag
X3
X
Index
X3
ACF
Lag



1. Proposer une démarche de modélisation pour chacune de ces séries, justifier.
2. Notre statisticien s'intéresse ensuite à une autre série $X5_t$. Il propose de la modéliser par un ARMA et cherche ensuite à valider son modèle. Il obtient les sorties suivantes, à quoi correspond la ligne "s.e"? Quelle est l'instruction (en R) qui lui a permis d'obtenir cette sortie?

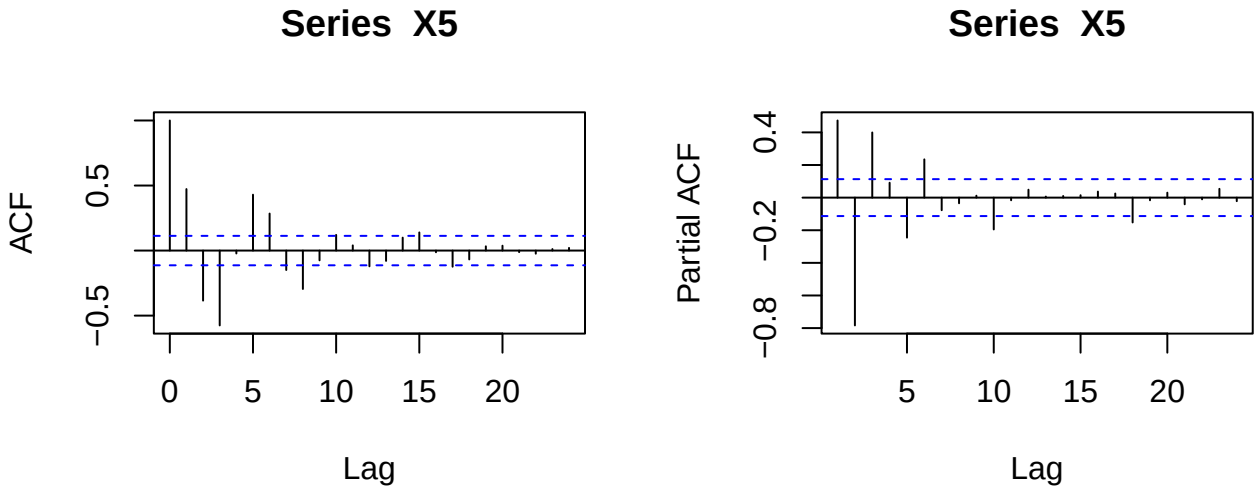
```
##
## Call:
## arima(x = X5, order = c(2, 0, 5), include.mean = F, method = c("ML"))
##
## Coefficients:
##          ar1      ar2     ma1     ma2      ma3     ma4      ma5
##       0.5778  -0.7899  0.8302  0.6687  -0.0654  0.1970  -0.1158
## s.e.  0.0678   0.0543  0.0879  0.1283   0.1127  0.0992   0.0761
##
## sigma^2 estimated as 8.85:  log likelihood = -755.99,  aic = 1527.97
```

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

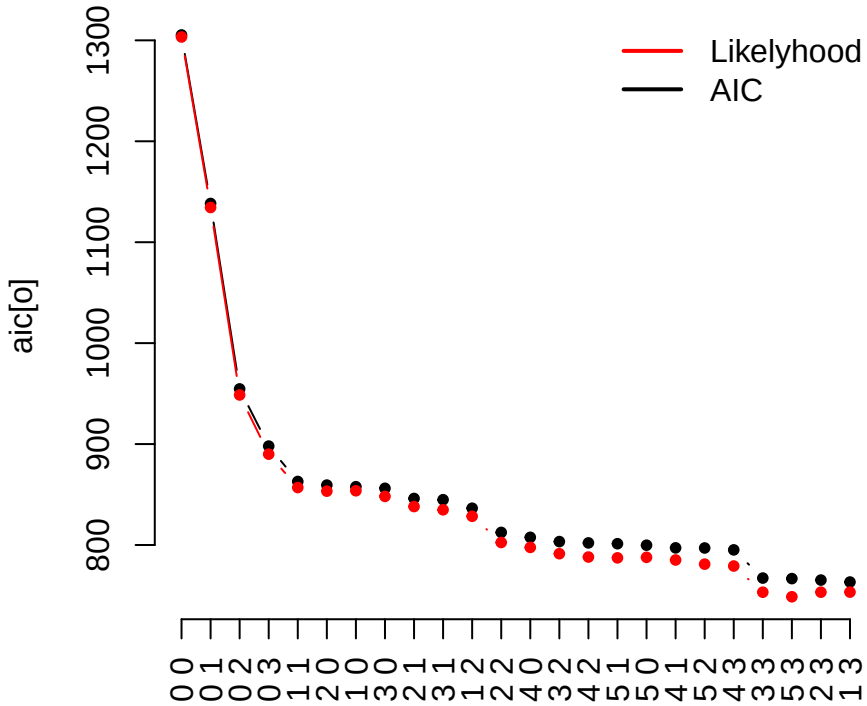
```
##  ar1  ar2  ma1  ma2  ma3  ma4  ma5
## 0.00 0.00 0.00 0.00 0.56 0.05 0.13
```

3. Au vu de ces résultats que doit faire notre statisticien?
4. Il trace ensuite l'ACF et le PACF de $X5$. Au de ces graphiques, il décide d'effectuer une procédure de

sélection de modèle via le critère AIC. Expliquer comment est calculé ce critère et quel est le but de cette approche.



Commenter le graphique ci-dessous? Quelles conclusions en tirer?

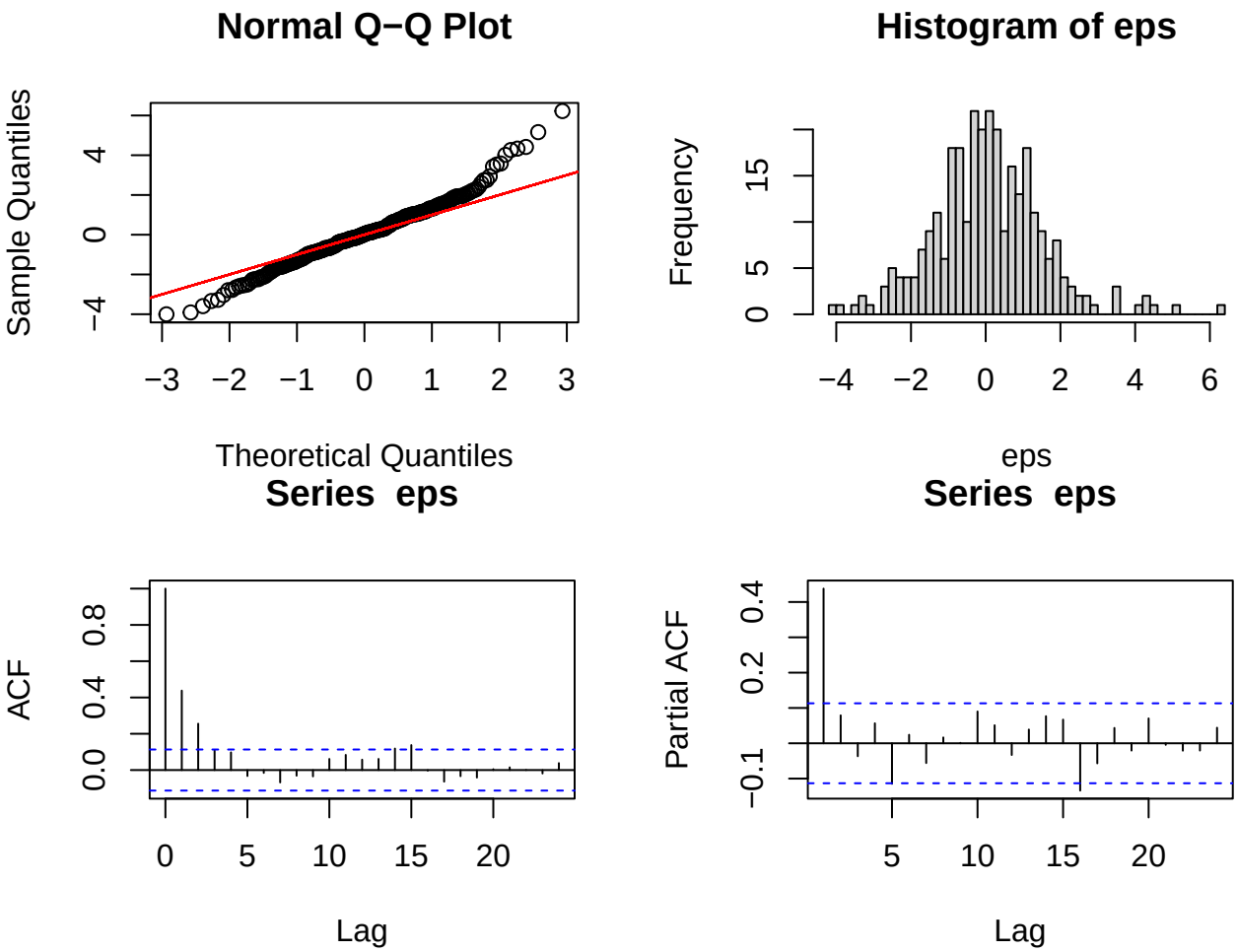


5. Il s'intéresse ensuite à une série $X6_t$ et effectue un test de Box-Pierce d'ordre 20 sur $X6_t$. Il obtient le résultat suivant:

```
##  lags statistic df    p-value
##    20  35.71442 20 0.01661317
```

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

5. Un de ces collègues cherche à diagnostiquer un modèle de prévision. Il étudie les résidus d'apprentissage de son modèle (un ARIMA Gaussien). Que doit il dire à son collègue?


Normal Q−Q Plot
Sample Quantiles
Theoretical Quantiles
Histogram of eps
Frequency
eps
Series eps
ACF
Lag
Series eps
Partial ACF
Lag