

TP1: Introduction à la manipulation de série temporelle avec R

MAP-STA2 : Séries chronologiques

Yannig Goude [yannig.goude@edf.fr]
Bianca Marin-Moreno [bianca.marin-moreno@edf.fr]

Exercice 1

- importer les données `beer2.csv` (production mensuelle de bière en Australie).
- à l'aide de la fonction `strptime` fabriquer une variable date correspondante.
- à l'aide des classes `ts` et `zoo` convertir les données en séries temporelles et les représenter graphiquement.
- effectuer une analyse descriptive de base de la série: moyenne, écart-type, boxplot, histogramme.
- calculer la moyenne de la série par mois de l'année, que concluez-vous? Etait-ce attendu?

Exercice 2

- importer les données `conso_2015.csv` correspondant à la consommation électrique française par demi-heure en MW.
- sous quelle forme sont fournies les données? Est ce un format adéquate pour l'analyse d'une série temporelle?
- construire une variable `Date` adaptée puis à l'aide de la classe `xts` convertir les données en séries temporelles et les représenter graphiquement.
- en exploitant les fonctions `.indexmon`, `.indexyday` et `.indexhour` du package `xts`, calculer:
 - la consommation moyenne par mois
 - le profil horaire (conso. moyenne par heure) de la consommation française
 - le profil horaire par type de jour (lundi, mardi...), puis le représenter graphiquement.
- représenter les boxplots de la consommation à 20h par mois.
- calculer les auto-corrélations et auto-corrélation partielles empiriques d'ordre $h = 1, 2, \dots, 336$. Ne pas utiliser les fonctions `acf` et `pacf` mais proposer votre propre implémentation.

Exercice 3

Simuler une série chronologique X au pas mensuel de mars 1986 à avril 2002, telle que:

$$X_t = T_t + S_t + \varepsilon_t$$

t étant donc le temps en mois et:

- T_t la tendance logarithmique: $T_t = \log(t/10 + 1)$
- S_t la saisonnalité de période 1 an: $S_t = \cos(2t\pi/12)$
- ε un bruit blanc Gaussien de variance 1

Utiliser la classe `xts` et les fonctions de date de R pour construire X_t . Représenter graphiquement la série et la superposer à sa tendance et sa période.

Sélectionner et superposer sur un même graphique les années 1987, 1990 et 1993.

Effectuer le même exercice en supposant que X_t suit un modèle multiplicatif $X_t = T_t S_t \varepsilon_t$

Exercice 4

Importer les données `Solar1.RDS` (le format `.RDS` est un format de compression de données sous R, utiliser `readRDS` pour importer les données) de production photovoltaïque.

- quel est le pas de temps des observations?
- calculer et représenter la production (variable `Z1`) moyenne par heure, par mois, puis par mois et par heure.
- à partir de la variable `Ssrd` (rayonnement solaire) construire une variable valant `Ssrd` à 12h et sa différenciation d'ordre 1 autrement.
- à l'aide du package `dygraphs`, représenter la variable ainsi obtenu et la variable `Z1`.