redaction time series

Wilfried Ehounou

November 11‘ 2017

Contents

Chapter 1

Mesures : Des Series Temporelles

1.1 Series temporelles

Définition 1 *Une série temporelle est une suite chronologique de valeurs réelles x_t à des instants de temps écartés régulièrement.*

$$(x_t)_{1\leq t\leq n} \tag{1.1}$$

avec n la dimension de la série.

L'intervalle de temps entre deux mesures successives dépend de la série. Il peut s'agir d'un jour, d'une semaine, d'une minute ...
Généralement, les séries temporelles s'utilisent dans les problèmes suivants :

- modélisation : la representation de la serie sous la forme d'une fonction du temps.
- prévision : prédire les données futures à partir de valeurs prédécentes.
- changement de régime : la série change-t-elle significativement à un instant t.
- comparaison : determiner la relation existante entre une série observée et d'autres séries candidates.

Dans ce document, nous nous concentrerons à la dernière categorie des problèmes.
Selon le but de l'etude à realiser, deux definitions s'imposent.

1.1.1 Particularités des series temporelles

Bien que la série temporelle $(x_t)_{1\leq t\leq n}$ est l'observation des n premières réalisations d'un processus stochastique $(X_t)_t$, elle se modelise sous la forme suivante:

Définition 2 *la série temporelle se décompose en une somme de trois composantes*

- *deterministe ou tendance non nulle m*
- *périodique ou saisonnière s*
- *aléatoire ou le bruit ϵ*

$$X_t = m + s + \epsilon = f(t) + \epsilon \tag{1.2}$$

Définition 3 *La fonction $f(t)$ est decomposée en deux termes:*

$$f(t) = m(t) + S(t) \tag{1.3}$$

òu $m(t)$ et $S(t)$ sont deux fonctions avec

1. *$S(t)$ est une fonction périodique non nulle de période T telle que $\sum_{i=1}^{T} S(i) = 0$ alors s est la composante saisonnière de X.*

2. *si $m(t)$ est non nulle alors m est la tendance de X.*

avec $S(i)$ les coefficients de la fonction S.

La tendance est une fonction d'une famille simple décrite par peu de paramètres : fonction affine (droite), polynôme de degré faible, ou fonction exponentielle. Le choix des familles de fonctions parmi lesquelles on cherche la fonction f se fait souvent à vue, en regardant les graphiques des données sans autre justifcation théorique : certaines saisonalités sont visibles, d'autres apparaissent plus nettement après certaines transformations des données brutes (transformée de Fourier appelée périodogramme). Une tendance croissante régulière est modélisée par une tendance affine, par un polynôme d'ordre 2 si on observe un certain creusement... Il n'existe pas de critère objectif pour faire le choix de cette famille.
Le choix de la période de la saisonalité est orienté par des connaissances a priori sur la série. Par exemple, les phénomènes liés aux climats, ainsi qu'une grande partie des séries économiques ont une périodicité annuelle. Mais une série de consommation électrique présente une saisonnalité hebdomadaire en plus.
Certaines séries se modélisent par un produit des trois composantes. Cependant, l'application d'un logarithme nous permet de revenir à la propriété précédente.

$$Y_t = log(X_t) = log(m) + log(s) + log(\epsilon) \tag{1.4}$$

Le modèle multiplicatif est approché au modèle additif.

1.1.2 Estimation de la tendance et la saisonnalité

Les methodes presentées s'utilisent generalement dans la prévision de données. Nous utilisons ces méthodes dans le but de déterminer des paramêtres qui permettront de s'approcher, au pire des cas, de la série X_t.
la sêrie temporelle peut être modélisé par un processus *stationnaire* et *non stationnaire.*

Processus stationnaire

Processus non stationnaire

Connaître ces paramêtres, en particulier la période T, nous permet de selectionner des échantillon représentatifs des séries dans le but de les comparer entre eux

1.2 application des méthodes sur les séries temporelles de Champlan