

TP1 : simulation d'une série simple - désaisonnalisation - modélisation de la tendance par régression

1 But du TP

Dans ce TP, on va simuler une série suivant le modèle :

$$X_t = at + b + \lambda \cos(\omega t) + \sigma \epsilon_t, \quad t = 1, \dots, n, \quad (1)$$

avec $n = 72$, $a = 3$, $b = 4$, $\lambda = 10$, $\sigma = 2$, $\omega = \frac{2\pi}{12}$, et $\{\epsilon_t, t = 1, \dots, n\}$ i.i.d. de distribution $\mathcal{N}(0, 1)$.

1. Décomposer la série en tendance + composante saisonnière + bruit. Quelle est la période de la composante saisonnière ?
2. Quelle est la particularité des résidus de cette série temporelle ?

2 Préliminaire : simulation de données normales

Rappel : le code suivant permet de simuler un échantillon i.i.d. de taille 5 suivant la loi normale $\mathcal{N}(0, 1)$.

```
/* Simulation de 5 observations iid N(0,1) sans choisir la graine */  
data simuleessai; do i = 1 to 5;  
z = rannor(0);  
output; end;  
run;  
/* Impression */  
proc print data=simuleessai; run;
```

Dans ce cas la graine a été calculée à partir de l'horloge interne de l'ordinateur. On peut également choisir la graine (entier positif inférieur à $2^{31} - 1$). Cela permet d'obtenir toujours le même échantillon pseudo-aléatoire.

3 Simulation des données

Dans une étape `data`, simuler les données en créant une variable `ddate` qui contiendra le numéro de mois (de 1 à 72), les différentes composantes de la série et enfin la série x_t . On nommera le tableau de données `simulx`. L'afficher.

Tracer le chronogramme de x_t et de chacune de ses composantes.

La suite du TP va consister à étudier la série x_t et à retrouver, par plusieurs méthodes, la décomposition de x_t .

4 Modélisation

4.1 Etude des diagrammes retardés

1. Grâce à une étape `data`, créer la série des observations retardées et des différences à l'ordre 1, 6, 12.
2. Tracer les diagrammes retardés à l'ordre 1, 6, 12. Commenter.

4.2 Tendance de la série des différences par lissage exponentiel

1. Appliquer la méthode de lissage exponentiel sur la série des différences à l'ordre 12.
2. Tracer sur un même graphique la série lissée et la série brute (des différences).
3. Calculer la série des résidus et la tracer.
4. La méthode de lissage exponentiel semble-t-elle bien adaptée ?

4.3 Tendance de la série des différences par régression sur le temps

1. Calculer la corrélation de la série des différences à l'ordre 12 avec le temps.
2. Faire une régression simple de la série des différences à l'ordre 12 sur le temps. Comparer les coefficients de la régression avec les coefficients de l'équation du modèle. Comprendre.

4.4 Modélisation par `proc x11`

1. Appliquer la `proc x11` sur la série brute. Attention, il faut ajouter l'option `additive` pour préciser qu'il s'agit d'un modèle additif (par défaut la procédure considère un modèle multiplicatif). De plus, la `proc x11` a besoin d'une date au format date de SAS pour fonctionner. Rajouter donc dans les options (par exemple en faisant démarrer la série en janvier 1966) :

```
monthly additive;
start=jan1966
```

2. Représenter la série lissée, la série saisonnière et la série des résidus.