

Les **objectifs** de ce TD sont

- d'apprendre à simuler les séries chronologiques simples en utilisant le logiciel **R** ;
- d'apprendre à calculer et interpréter une fonction d'autocorrélation ;
- de faire de la prévision à l'aide des techniques de lissage exponentiel.

1 Simulation de séries temporelles

On appelle *bruit blanc gaussien* une suite de variables aléatoires indépendantes et identiquement distribuées ϵ_t de loi normale centrée réduite.

1. Quelle est la fonction d'auto-corrélation d'un bruit blanc ?

Définition : Soit $X = \{X_t\}_{t \in \{1, \dots, T\}}$ un processus. La fonction d'auto-covariance de X est donnée par

$$C(t, s) = E(X_t X_s) - E(X_t)E(X_s).$$

La fonction d'auto-corrélation est donnée par

$$\rho(t, s) = \frac{C(t, s)}{\sqrt{\text{Var}(X_t)\text{Var}(X_s)}}$$

2. Simulez un bruit blanc gaussien de taille 100, et représenter le graphiquement.
3. Tracez la fonction d'auto-corrélation théorique et la comparer à la fonction d'auto-corrélation empirique (utiliser la fonction **acf**).
4. Recommencez les deux questions précédentes et observez la variabilité des résultats. Jouez sur la longueur de la série.
5. Simulez maintenant la série temporelle $X(t) = 0.5t + 2\epsilon_t$ avec $\epsilon_t \simeq N(0, 1)$ (taille 100).
6. Représentez graphiquement la série et interprétez la qualitativement.
7. Faites de même pour $X(t) = 0.5t + \epsilon_t + 3 \cos(\frac{\pi}{6}t)$ avec $\epsilon_t \simeq N(0, 1)$.

2 Lissage et prévision de données simulées

Simuler les trois séries temporelles suivantes, de taille 100 :

- $X_1(t) = \epsilon_t$
- $X_2(t) = 0.5t + 2\epsilon_t$
- $X_3(t) = 0.5t + \epsilon_t + 3 \cos(\frac{\pi}{6}t)$

avec $\epsilon_t \simeq N(0, 1)$. Les 70 premières données de chaque série vont être utilisés pour prédire les 30 dernières.

1. Pour chaque série, effectuer la prévision par lissage exponentiel simple et double. Tester différentes valeurs du paramètre de lissage α (4 à 5 valeurs), et représenter graphiquement la série ainsi que la prévision. Commenter chaque résultat, et essayer de déterminer graphiquement le lissage le plus adapté pour chaque série.
2. Calculer pour chaque prévision effectuée la somme des carrés des erreurs, et sélectionner le meilleur modèle à l'aide de cette quantité.
3. Tester maintenant le lissage exponentiel de Holt-Winters avec composante saisonnière additive puis multiplicative. Voir fonction `HoltWinters`
4. Les prédictions obtenues sont-elles meilleures ?

3 Lissage et prévision de la concentration en CO2

Le fichier de données `co2` contenu dans R contient les concentrations en CO2 à proximité du volcan Mauna Loa (Hawaï) de 1959 à 1997. Après avoir représenté graphiquement ces données, quel modèle de lissage exponentiel vous semble le mieux approprié ? Afin de valider ce modèle, tester la prédiction des données de 1990 à 1997 en utilisant celles de 1959 à 1989. Si cela vous semble graphiquement correct, utilisez cette méthode pour prédire les concentrations en `co2` de 1997 à 2007. Sinon, tester d'autres méthodes de lissage exponentiel.

4 Lissage et prévision du CAC40

Récupérer le fichier contenant les valeurs de clôture journalière du CAC40 de 1991 à 1998 (données R `EuStockMarkets`). Essayer de prédire par lissage exponentiel les valeurs de clôture de 1998 en utilisant les valeurs de clôture de 1991 à 1997.