Séance 1

Définition 1 (Série temporelle) *Une série temporelles pour les données observées* $\{x_t\}$ *est une spécification des distributions (ou seulement des moyennes et covariances) d'une suite de variables aléatoires* $\{X_t\}$ *pour lesquelles* $\{x_t\}$ *représente une réalisation.*

On utilisera le terme de série temporelle à la fois pour les données et le processus aléatoire dont elles sont une réalisation. On spécifiera uniquement les espérances $\mathbb{E}[X_t]$ et $\mathbb{E}[X_t X_{t+h}]$ pour $t = 1, 2, ...$ et $h = 0, 1, ...$ en se concentrant uniquement sur les propriétés de la série $\{X_t\}$ dépendant de ces dernières.

Définition 2 (Série temporelle stationnaire) *Une série temporelle* $\{X_t\}$ *est appellée stationnaire si elle a des propriétés statistiques similaires à celles correspondant à la série décalée en temps* $\{X_{t+h}\}$ *pour chaque entier* h.

Définition 3 (Moyenne et covariance) *Soit* $\{X_t\}$ *une série temporelle avec* $\mathbb{E}[X_t^2] < \infty$. *La moyenne de* $\{X_t\}$ *est*

$$\mu_X(t) = \mathbb{E}[X_t]$$

et la fonction covariance est

$$\gamma_X(r,s) = cov(X_r, X_s) = \mathbb{E}[(X_r - \mu_X(r))(X_s - \mu_X(s))].$$

Définition 4 (Série stationnaire) $\{X_t\}$ *est appellée faiblement stationnaire si*

(i) $\mu_X(t)$ *est indépendant de* t.

(ii) $\gamma_X(t+h,t)$ *est indépendant de* t *pour chaque* h.

$\{X_t\}$ *est strictement stationnaire si* $(X_1, X_2, ..., X_n)$ *et* $(X_{1+h}, X_{2+h}, ..., X_{n+h})$ *ont les mêmes distributions pour tous les entiers* h *et* $n > 0$.

Définition 5 (Auto-covariance et auto-corrélation) *Soit* $\{X_t\}$ *une série temporelle stationnaire. Sa fonction d'autocovariance est définie par*

$$\gamma_X(h) := \gamma_X(h,0) = \gamma_X(t+h,t).$$

Sa fonction d'auto-corrélation est

$$\rho_X(h) = \frac{\gamma_X(h)}{\gamma_X(0)}.$$

On rappelle la propriété de linéarité de la covariance $\mathrm{cov}(aX+bY+c, Z) = a\,\mathrm{cov}(X,Z) + b\,\mathrm{cov}(Y,Z)$ qui sera très utile par la suite.

En pratique on a souvent accès à des données observées $\{x_1, x_2, ..., x_n\}$ à partir desquelles on pourra reconstituer le modèle après avoir analysé la dépendance entre elles.

Exercices

1. Soit X et Y deux variables aléatoires avec $\mathbb{E}[Y] = \mu$ et $\mathbb{E}[Y^2] < \infty$.
   (i) Montrer que la constante qui minimise $\mathbb{E}[(Y-c)^2]$ est $c = \mu$.
   (ii) En déduire que la variable aléatoire $f(X)$ qui minimise $\mathbb{E}[(Y - f(X))^2|X]$ est
   $$f(X) = \mathbb{E}[Y|X].$$

2. Soit Z_t une suite de variables aléatoires normales, indépendantes de moyenne 0 et variance σ^2 et soit a, b, c des constantes. Est-ce que les processus suivants sont stationnaires? Spécifier pour chacun la moyenne et la fonction d'auto-covariance.
   (i) $X_t = a + bZ_t + cZ_{t-2}$.
   (ii) $X_t = Z_1 \cos(ct) + Z_2 \sin(ct)$.
   (iii) $X_t = Z_t \cos(ct) + Z_{t-1} \sin(ct)$.
   (iv) $X_t = a + bZ_0$.
   (v) $X_t = Z_0 \cos(ct)$.
   (vi) $X_t = Z_t Z_{t-1}$.

3. Soit X_t un processus aléatoire donné par
   $$X_t = Z_t + \theta Z_{t-2}$$
   où $\{Z_t\}$ est une suite de variables aléatoires appelée "bruit blanc" (en anglais "white noise"), c.a.d. des variables non correlées de moyenne nulle et variance σ^2, $\{Z_t\} \sim WN(0, \sigma^2)$.
   (i) Trouver l'autocovariance et l'autocorrélation pour $\theta = 0.8$.
   (ii) Calculer la variance de l'échantillon $(X_1 + X_2 + X_3 + X_4)/4$ quand $\theta = 0.8$.
   (iii) Répéter (ii) avec $\theta = -0.8$ et comparer la réponse avec (ii).

4. Soit $\{Z_t\}$ des variable IID (indépendantes identiquement distribuées) en suivant la loi normale N(0,1). On définit le processus aléatoire suivant:
   $$X_t = \begin{cases} Z_t, \ t \text{ pair}, \\ \frac{Z_{t-1}^2 - 1}{\sqrt{2}}, \ t \text{ impair} \end{cases}$$
   (i) Montrer que $\{X_t\}$ est un bruit blanc, $\{X_t\} \sim WN(0, 1)$ mais pas IID.
   (ii) Trouver $\mathbb{E}(X_{n+1}|X_1, X_2, ..., X_n)$ pour n pair et impair.