FICHE TP 2

Exercice 1. Simulation d'un vecteur gaussien bidimensionnel

Soit $X = (X_1 , X_2)^T$ un vecteur gaussien bidimensionnel avec $E(X) = \mu = (\mu_1, \mu_2)^T$, $Var(X_1) = \sigma_1^2 > 0$, $Var(X_2) = \sigma_2^2 > 0$ et $\rho = Cor(X_1, X_2) \in]-1, 1[$.

Partie 1. Dans toute cette partie, on considère le cas $\boldsymbol{\mu = (1, 2)^T}$, $\boldsymbol{\sigma_1^2 = \sigma_2^2 = 1}$ et $\boldsymbol{\rho = 0.8}$.
1. Simuler un échantillon de taille n = 100 de la loi du vecteur X en écrivant que

$$X_1 = 1 + \varepsilon_1$$
$$X_2 = 2 + 0.8\times\varepsilon_1 + 0.6\times\varepsilon_2$$

où ε_1, ε_2 sont **indépendantes** de même loi **N(0, 1)**.
Représenter le nuage de points correspondant. Superposer (en rouge) les deux droites d'équations respectives $y = x + 1$ et $y = -x + 3$.

Fonctions R utiles : matrix, plot, abline

2. Obtenir la décomposition spectrale de la matrice de covariance $\Gamma = Cov(X)$. Que valent les deux valeurs propres $\lambda_1 \geq \lambda_2$ et les vecteurs propres associés ? En déduire l'interprétation des deux axes précédents et vérifier expérimentalement (sur votre jeu de données simulées) que les nouvelles coordonnées sont indépendantes et gaussiennes de variances respectives λ_1 et λ_2 (on pourra augmenter la taille de n pour de meilleures estimations des variances λ_1 et λ_2).

Fonctions R utiles : eigen, var, qqnorm, qqline

Partie 2. Ecrire une fonction qui simule un échantillon de taille n de la loi du vecteur X en fonction de n et des différents paramètres à partir de la décomposition de Cholesky de la matrice de covariance $\boldsymbol{\Gamma = LU}$ avec $U=L^T$ (fonction **chol**) et du fait que $\boldsymbol{X = LZ}$ est alors de loi $N(0, \Gamma)$ si Z est de loi N(0, Id) (voir prochain cours pour ce résultat). On pourra prendre $\mu = 0$ qui est simplement un vecteur de décentrage. En jouant sur les autres valeurs des paramètres (σ_1, σ_2 et ρ), étudier les différentes formes de nuages de points que l'on peut obtenir (mêmes variances individuelles mais $\rho > 0$, $\rho < 0$ ou $\rho = 0$ puis variances différentes). Que se passe-il si $\rho = \pm 1$? Que peut-on dire du VG correspondant à ce cas extrême ? Utiliser la fonction chol pour la décomposition de Cholesky

Exercice 2. Loi conditionnelle dans le cas d'un vecteur gaussien

On reprend le jeu de données simulées exercice 1, partie 1, question 1. L'objectif de cet exercice est d'étudier empiriquement la loi conditionnelle de X_2 sachant $X_1 = x_1$ où x_1 est donné.

1. Pour h « petit » devant $\sigma_1 = 1$, extraire le jeu de données correspondant à $X_1 \in [x_1 - h, x_1 + h]$ puis étudier la distribution empirique des valeurs associées de la variable X_2. Qu'observez-vous ? (**fonctions R utiles** : **mean**, **var**, **qqnorm**, **qqline**) On augmentera la taille de n pour de meilleurs résultats.

2. Choisir d'autres valeurs de x_1. Quelle est la nature de la fonction $x_1 \to E(X_2 \mid X_1 = x_1)$? Que peut-on dire de $Var(X_2 \mid X_1 = x_1)$? Et de la loi conditionnelle de X_2 sachant $X_1 = x_1$?

Exercice 3. Un jeu de données financières

On met à disposition les données de marché du 24/09/2018 au 24/09/2019 de quarante actions cotées en continu sur le premier marché et qui constituent l'indice CAC40. Les sociétés correspondantes, représentatives de différentes branches d'activités, reflètent la tendance globale de l'économie des grandes entreprises françaises et leur liste est revue régulièrement pour maintenir cette représentativité.

On s'intéressera au prix S_t à la clôture journalière (jour t) de ces actions, voir le code R fourni pour extraire par exemple l'évolution de l'action RENAULT.

Pour analyser l'évolution du prix S_t au jour le jour, on considère la série des taux de hausse ou de baisse logarithmiques (ou taux de rendement logarithmiques) :

$$R_t = \log(\frac{S_t}{S_{t-1}}) \approx \frac{S_t - S_{t-1}}{S_t}$$

L'hypothèse classique est de considérer que les variables $R_1, \ldots, R_n$ sont indépendantes (attendre le cours de « Séries Temporelles » de l'UP3 pour une analyse de cette hypothèse). Comme l'historique de marché est de 1 an, on supposera également que ces taux sont de même loi, en particulier de même écart-type (appelé volatilité).

1. En sélectionnant par exemple les deux actions SOCIETE GENERALE et BNP PARIBAS, tester empiriquement si la loi du vecteur bidimensionnel des taux de hausse ou de baisse de ces deux actions est un vecteur gaussien.

2. Tracer sur un même graphique l'évolution du cours S_t de ces deux actions et interpréter la corrélation entre les deux courbes en liaison avec la question précédente.

3. Sélectionner un couple d'actions avec une corrélation la plus faible possible. Faire la même analyse. Quel peut être l'intérêt en pratique pour des investisseurs ?