Modéle de régression linéaire - Feuille 1

Introduction

Exercice 1

1. Soient X et ε deux variables aléatoires réelles indépendantes, telles que $P(\varepsilon = +1) = P(\varepsilon = -1) = 1/2$ et X de loi $\mathcal{N}(0, 1)$. On pose $Y = \varepsilon X$.
   (a) Calculer la fonction de répartition de Y. En déduire sa loi.
   (b) Les variables aléatoires réelles X et Y sont-elles non corrélées ?
   (c) Calculer $P(X + Y = 0)$. Le vecteur (X, Y) est-il gaussien ? Les variables aléatoires réelles X et Y sont-elles indépendantes ?
2. Soient X, Y, Z trois variables aléatoires indépendantes de même loi $\mathcal{N}(0, 1)$. On pose $U = X - Y$ et $V = X + Y - 2Z$.
   (a) Quelles sont les lois de U et de V ?
   (b) Ces variables aléatoires sont-elles indépendantes ?

Exercice 2 Soit $y_1, \cdots, y_n$ des réels.

1. Déterminer le réel $\widehat{m}$ qui minimise $SCR(m) = \sum_{i=1}^{n} |y_i - m|^2$.
2. Que représente $\widehat{m}$ par rapport à $y_1, \cdots, y_n$?

Exercice 3 Soit $y_1, \cdots, y_n$ des réels.

Montrer que le réel $\tilde{m}$ qui minimise la somme des valeurs absolues $SVA(m) = \sum_{i=1}^{n} |y_i - m|$ est la médiane de la série.

Indication : on pourra traiter d'abord le cas où $n = 2p + 1$ et classer les réels (y_i) sous la forme $y_{(1)} \leq y_{(1)} \leq \ldots \leq y_{(n-1)} \leq y_{(n)}$, puis passer au cas où $n = 2p$.

Exercice 4 Soit $X = (X_1, X_2, X_3)$ un vecteur gaussien tel que $\forall i \in \{1, 2, 3\}, \mathbb{E}X_i = 0, \mathbb{E}X_i^2 = 1$ et $\mathbb{E}(X_i X_j) = 1/2, i \neq j$. Écrire la fonction caractéristique de ce vecteur et chercher sa densité.

Exercice 5 Soit (X, Y) un vecteur aléatoire dans $\mathbb{R}^2$, dont la loi admet la densité

$$f_{(X,Y)}(x, y) = \frac{1}{2\pi} e^{-\frac{x^2+y^2}{2}}.$$

Déterminer les lois de X, Y, $X + Y$, $X^2 + Y^2$.

Rappel : $\int_{\mathbb{R}} e^{-u^2/2} du = \sqrt{2\pi}$ (on le retrouve facilement en utilisant la densité d'une normale unidimensionnelle centrée réduite dont l'intégrale fait 1).

Exercice 6 Soit $(X_1, \ldots, X_n)$ un n-échantillon de loi $\mathcal{N}(m, \sigma^2)$. On pose

$$Y_i = X_i - X_1,\ 2 \leq i \leq n,$$

$$\overline{X}_n = n^{-1} \sum_{i=1}^{n} X_i \qquad \text{et} \qquad S^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \overline{X}_n)^2.$$

1. Quelle est la loi de $(Y_2, \ldots, Y_n)$?
2. Montrer que $(Y_2, \ldots, Y_n)$ est indépendant de $\overline{X}_n$. En déduire que $\overline{X}_n$ et S^2 sont indépendants.

Exercice 7 Soit le vecteur aléatoire X de loi $\mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 3 & -1 & 0 \\ -1 & 3 & 0 \\ 0 & 0 & 2 \end{pmatrix}\right)$. Trouver une matrice A de dimension (3,3) telle que AX soit un vecteur de composantes indépendantes.

Exercice 8 Une variable aléatoire X est dite de loi Gamma de paramètres α et λ $(\alpha > 0, \lambda > 0)$, notée $\gamma(\alpha, \lambda)$, si sa loi a la densité :

$$f(x) = \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x} \mathbb{1}_{(0,+\infty)}(x),$$

où $\forall \alpha > 0,\ \Gamma(\alpha) = \displaystyle\int_0^\infty x^{\alpha-1} e^{-x} dx.$

1. Verifier que la loi $\gamma(\alpha, \lambda)$ est bien une loi de probabilité.
2. Soit X une variable aléatoire réelle de loi $\gamma(\alpha, \lambda)$. Calculer sa transformée de Laplace. Calculer sa moyenne et sa variance.
3. Soit X une variable aléatoire réelle de loi $\mathcal{N}(0, 1)$. Donner la loi de X^2. En déduire que $\gamma(1/2, 1/2) = \chi^2(1)$.
4. Soient X et Y deux variables aléatoires indépendantes de lois $\gamma(\alpha_1, \lambda)$ et $\gamma(\alpha_2, \lambda)$.
   (a) Donner la loi de $X + Y$.
   (b) Montrer que $X + Y$ et $\dfrac{X}{X+Y}$ sont indépendantes et calculer leur loi de probabilité.
5. Si $X_1, ..., X_n$ sont n variables aléatoires réelles indépendantes et de même loi exponentielle de paramètre λ, donner la loi de la somme $S_n = X_1 + ... + X_n$. Calculer $\mathbb{E}(S_n)$ et $\mathrm{Var}(S_n)$.
6. Si $Y_1, ..., Y_n$ sont n variables aléatoires indépendantes de loi $\mathcal{N}(0, 1)$, donner la loi de $Z = Y_1^2 + ... + Y_n^2$ et calculer sa transformée de Laplace, sa moyenne et variance.
   En déduire que $\gamma(n/2, 1/2) = \chi^2(n)$.

Rappels :

$\Gamma(1/2) = \sqrt{\pi}$,

$\Gamma(\alpha + 1) = \alpha\Gamma(\alpha)$, $\forall \alpha > 0$,

$B(\alpha_1, \alpha_2) = \displaystyle\int_0^1 u^{\alpha_1 - 1}(1-u)^{\alpha_2 - 1} du = \frac{\Gamma(\alpha_1)\Gamma(\alpha_2)}{\Gamma(\alpha_1 + \alpha_2)}$, $\forall \alpha_1, \alpha_2 > 0$,

La densité d'une loi $\beta(a, b)$, $a, b > 0$, est $f(x) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x^{a-1}(1-x)^{b-1}\mathbf{1}_{0 \leq x \leq 1}$.