TD DE MODÈLES LINÉAIRES I - SÉRIE 3

Exercice 1. Le salaire désiré d'un individu s'écrit $Y^* = Xb + \sigma\epsilon$, où $\sigma > 0$, $b \in \mathbb{R}$, X une variable aléatoire admettant des moments d'ordre 2 mesurant la capacité de l'individu, ϵ est variable aléatoire indépendante de X et de loi $\mathcal{N}(0,1)$. Si Y^* est plus grand que le SMIC S alors le salaire reçu Y est gale au salaire dsir Y^*, sinon le salaire reçu Y vaut S. Calculer $E[Y/X]$. Cette espérance est-elle linéaire?

Exercice 2. Soient (X_i, ϵ_i), $i = 1, \cdots, n$, des couples de variables i.i.d. On suppose que X_i et ϵ_i admettent des moments d'ordre 2 finis, et que $\mathbb{E}(\epsilon_1) = 0$, $\mathbb{E}(X_1^2) > 0$. Pour un réel b, on pose $Y_i = bX_i + \epsilon_i$, on note

$$\hat{b} = \frac{\sum_{i=1}^n Y_i X_i}{\sum_{i=1}^n X_i^2}.$$

1) En observant que

$$\hat{b} = b + \frac{\sum_{i=1}^n \epsilon_i X_i / n}{\sum_{i=1}^n X_i^2 / n}$$

déduire que $\hat{b}$ converge presque sûrement vers b.

2) Trouver la loi limite de $\sqrt{n}(\hat{b} - b)$.

Exercice 2. Dans de nombreux problèmes pratiques, lorsque l'on étudie la relation qui peut exister entre 2 quantités Y et X, nous savons que si $X = 0$ alors $Y = 0$. C'est le cas, par exemple, si nous étudions le volume Y occupé par un gaz en fonction temps X, temps égal à 0 au moment où nous avons actionné un vaporisateur contenant le gaz.
Le modèle de régression linéaire simple n'est pas adapté pour modéliser ce type de lien. En effet, nous devons ajuster une droite dont l'ordonnée à l'origine est nulle. Nous considérons n valeurs de X que nous notons x_i. Pour chaque x_i, nous observons une valeur de Y notée Y_i.
Pour tout $i \in \{1, \ldots, n\}$, nous proposons le modèle

$$Y_i = \beta x_i + \varepsilon_i$$

où β est un paramètre inconnu et ε_i est une variable aléatoire appelée résidu telle que $\mathbb{E}(\varepsilon_i) = 0$, $Var(\varepsilon_i) = \sigma^2$ et $Cov(\varepsilon_i, \varepsilon_j) = 0$ lorsque $i \neq j$.
Il s'agit du modèle de régression linéaire simple sans coefficient constant. Nous considérons les 2 estimateurs de β suivants :

$$\hat{\beta} = \frac{\sum_{i=1}^n x_i Y_i}{\sum_{i=1}^n x_i^2} \quad et \quad \tilde{\beta} = \frac{\sum_{i=1}^n Y_i}{\sum_{i=1}^n x_i}$$

1) Déterminer la logique de construction de ces deux estimateurs.

2) Montrer que $\hat{\beta}$ et $\tilde{\beta}$ sont des estimateurs non biaisés de β.

3) Montrer que $Var(\hat{\beta}) < Var(\tilde{\beta})$ sauf dans le cas où les x_i sont égaux. Ce résultat était-il prévisible?

Exercice 4. Soient m et λ deux paramètres réels. Soit le modèle linéaire gaussien défini par les équations

$\mathbb{E}(X_1) = m + \lambda$ et $\forall i \in \{2, ..., n\}, \mathbb{E}(X_i) = m$.

Calculer l'estimateur sans biais de variance minimale par deux méthodes différentes.