

Statistique non paramétrique - ISFA - Devoirs (sujet au choix)

Sujet 1. On considère le jeu de données INSURANCE1.csv, qui contient les montants de $n = 1000$ sinistres IARD au bout de deux ans. On sait que la population des assurés se décompose en un mélange de deux sous-populations. Essayez d'estimer la loi de chaque sous-population.

Sujet 2. On s'intéresse au modèle de Lee-Carter (1992) qui modélise l'évolution du taux de mortalité de la manière suivante :

$$\log(q_{(x,t)}) = \alpha_x + \beta_x * \kappa_t + \epsilon_{(x,t)}$$

avec x l'âge et t l'année, et

- κ = vitesse moyenne de baisse de la mortalité
- α = moyenne de la mortalité en échelle logarithmique
- β = correction de la vitesse par âge
- ϵ = innovations Gaussiennes centrées réduites

κ correspond à la vitesse de baisse moyenne de la mortalité pour la population concernée. C'est une composante très importante dans la modélisation de la longévité. Le fichier *LeeCarter* contient les valeurs de κ de 1965 à 2017. On vous demande de réaliser une régression non paramétrique sur ces données.

Sujet 3. Le jeu de données Loss-ALAE (source : US Insurance Services Office) contient 1500 réclamations de responsabilité civile, chacune composée du paiement de l'indemnité (*Loss*) et de la charge de compensation allouée (*ALAE*). Ces données sont étudiées dans de nombreux articles et on s'intéresse à la loi du couple (X, Y) , avec $X = \text{Loss}$ et $Y = \text{ALAE}$. Etudier la loi jointe et les lois marginales du couple.

Sujet 4. Le jeu de données Loss-ALAE (source : US Insurance Services Office) contient 1500 réclamations de responsabilité civile, chacune composée du paiement de l'indemnité (*Loss*) et de la charge de compensation allouée (*ALAE*). Ces données sont étudiées dans de nombreux articles et on s'intéresse à la loi du couple (X, Y) , avec $X = \text{Loss}$ et $Y = \text{ALAE}$. Réalisez une régression non paramétrique.