
APPLICATIONS STATISTIQUES

1. Estimation de densité

Le fichier *ex1cusp1* contient un échantillon de taille 1000 d'une variable aléatoire X de densité inconnue f . Représenter les données.

Calculer l'histogramme des données en nb classes, nb étant la taille de l'échantillon divisée par 4 (fonction *hist*).

Effectuer la décomposition en ondelettes de la série des fréquences.

Débruiter et reconstituer la série débruitée.

Interpoler (fonction *interp1*) et représenter l'estimation de la densité.

2. Estimation de densité par noyau

Estimer la densité pour l'exemple précédent en utilisant la méthode par noyau, en prenant un noyau Gaussien et en faisant varier la fenêtre.

3. Estimation d'une fonction de régression

Le fichier *ex1nst0* contient un échantillon de taille 500 d'un couple de variables aléatoires (X, Y) suivant le modèle

$$Y = f(X) + \varepsilon$$

La fonction f et la densité h de la variable X sont inconnues.

Représenter les données.

Calculer l'histogramme des données en nb classes, nb étant la taille de l'échantillon divisée par 4.

Calculer la moyenne par classe de la variable Y .

Effectuer la décomposition en ondelettes de la série des Y moyennés.

Débruiter et reconstituer la série débruitée.

Interpoler et représenter l'estimation de la fonction f .

4. Régression de données réelles

Le fichier *pollu* contient des couples d'observations (x_i, y_i) où x_i est la température en degrés du jour i et y_i est la concentration moyenne mesurée en $\mu\text{g}/\text{m}^3$ sur la région Parisienne le jour i . Estimer la fonction de régression de la pollution sur la température à l'aide de la méthode vue ci-dessus.

5. Régression de données réelles par noyau

Estimer la fonction de regression précédente en utilisant la méthode par noyau.