

Modèles linéaires et généralisés

ECTS crédits

Composante Sciences Fondamentales et Appliquées

Présentation

Description

Confronté à des données réparties entre différentes classes, il est important de savoir discriminer si le fait d'appartenir à une classe particulière détermine un comportement différent, le cadre du modèle linéaire gaussien permet de fournir une représentation de cette dépendance, ainsi que les outils pertinents pour tester différentes hypothèses sur cette dépendance. En particulier les problèmes d'analyse de la variance à un et deux facteurs permettent de tester l'influence de l'appartenance à une classe ou de l'interaction entre classes croisées pour un paramètre de moyenne dans les observations.

Face à ces différentes façons de modéliser la dépendance en fonction de paramètres, on étudiera la problématique de la sélection du modèle pour en réduire la complexité.

Lorsque les observations sont discrètes cela nous amènera à des modèles différents, régression logistique pour des données de type 0-1, ou régression de Poisson pour des données de type comptage.

Objectifs

- * comprendre et appliquer le modèle de régression linéaire simple, qualité des estimateurs (biais, variance), les tests simples et multiples, les modèles d'analyse de variance à 1 et 2 facteurs ;

- * comprendre et appliquer le principe de la sélection de modèle dans le modèle linéaire
- * comprendre et appliquer la régression logistique et de Poisson.

Heures d'enseignement

Modèles linéaires et généralisés - TD	TD	30h
Modèles linéaires et généralisés - CM	CM	20h

Pré-requis nécessaires

Connaissances en probabilités (indépendance, vecteurs gaussiens, variables de Bernoulli, variables de Poisson).

Théorie de l'estimation statistique et des tests.

Réduction des matrices symétriques, algèbre linéaire.

Compétences en programmation en R.

Compétences visées

A l'issue de ce cours l'étudiant devra savoir poser un modèle pertinent en fonction du type de données. Il saura le résoudre numériquement (à l'aide de programmes en R ou de procédures issues de bibliothèques existantes) et théoriquement, et en déduire les paramètres estimés du modèle et leur précision numérique.

Il saura appliquer les méthodes de sélection de modèle et les critères Cp de Mallows, AIC/BIC.

Il saura mettre en œuvre la régression logistique et la régression de Poisson.

L'étudiant saura présenter les résultats de ces procédures de modélisation et de tests dans un contexte socio-économique comme un retour de consulting statistique.

Liste des enseignements

Modèles généralisés

Estimation et tests