

Modèles linéaires et généralisés

ECTS
6 crédits

Composante
**Sciences Fondamentales
et Appliquées**

Période de l'année
Semestre 2

En bref

- # **Langue(s) d'enseignement:** Français
- # **Organisation de l'enseignement:** Formation initiale
- # **Ouvert aux étudiants en échange:** Oui
- # **Référentiel ERASMUS:** Mathématiques et statistiques

Présentation

Description

Dans des jeux de données il est souvent pertinent d'essayer d'interpréter une variable, dite variable d'intérêt, par les influences que peuvent avoir d'autres variables, dites explicatives. Cette explication passe par la création de modèles, dépendant de paramètres décrivant l'influence quantitative de ces variables explicatives. Selon la typologie de ces variables évidemment les modèles que l'on construira sont évidemment très variés : le cadre du modèle linéaire en est le premier exemple, à partir de la simple régression linéaire permettant d'expliquer une relation affine entre deux variables continues, jusqu'à des modèles plus riches comportant plus de variables, en passant par les différents cas d'analyse de la variance lorsque la (les) variable(s) explicative(s) est (sont) factorielle(s), voire l'analyse de la covariance. Si par contre on considère des réponses qui ne sont pas continues mais discrètes, voire factorielles, les modèles vont se composer différemment : l'aspect linéaire n'est plus aussi explicite mais reste présent par le biais de fonctions de lien. Ces modèles comprennent la régression logistique, la régression de Poisson (éventuellement avec inflation de zéro), ainsi que les régressions polytomique ou multinomiale. Dans tous ces modèles on pourra construire des procédures de sélection de modèles conciliant parcimonie et qualité de l'approximation par le modèle.

Objectifs

Les objectifs de ce cours sont :

- savoir quel type de modèle proposer et l'interpréter ;

- savoir calculer et interpréter les résultats du modèle : estimations, tests, étude des résidus, tant théoriquement que sur les sorties d'un logiciel de statistiques ;
- savoir procéder à une sélection de modèle ;
- maîtriser les différentes étapes de construction d'un modèle pour pouvoir proposer un modèle dans un cadre non étudié en cours, et le résoudre.

Heures d'enseignement

TD	TD	16h
CM	CM	20h
P-Proj	Pédagogie par projet	12h

Pré-requis obligatoires

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

	Nature	CM	TD	TP	Crédits
Modèles linéaires	EC	10h	8h		
Modèles généralisés	EC	10h	8h		

UE = Unité d'enseignement

EC = Élément Constitutif

Infos pratiques

Contacts

Responsable de la mention

Pol Vanhaecke

+33 5 49 49 68 87

pol.vanhaecke@univ-poitiers.fr

Lieu(x)

Futuroscope