



Certificat d'université en Biostatistique 2022-2023

Parcours commun

Modèles linéaires I

Intervenant(s) : Caroline Verhoeven

Nombre d'heures (applications et exercices compris) : 15

Remarque : ce nombre ne tient pas compte des exercices à préparer à domicile, ni de la préparation de l'examen.

Langue : Français

Contenu du module

En statistique, nous voulons souvent modéliser une variable Y quantitative en fonction de variables prédictives X_1, X_2, X_3, \dots qui peuvent être quantitatives ou qualitatives. Si nous considérons que

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \dots,$$

Il s'agit d'un modèle linéaire où $\beta_0, \beta_1, \beta_2, \beta_3, \dots$ sont des inconnues.

Les modèles linéaires sont un outil important en statistique. Les contraintes pour pouvoir appliquer les modèles linéaires sont très restrictives : linéarité, normalité, indépendance et variances constantes, pourtant les méthodes associées sont robustes et versatiles.

Le but de ce module est de maîtriser les différentes méthodes aux modèles linéaires. Ce module se composera d'une partie théorique, mais également de cas pratiques pour lesquels nous utiliserons le logiciel R. A la fin de ce module l'étudiant devra connaître les différentes méthodes, savoir les utiliser, savoir interpréter les résultats obtenus à l'aide du logiciel et comprendre les contraintes et limitations.

Table des matières

1. Introduction : Rappels sur la régression simple et l'ANOVA à 1 facteur
2. Quelques mots sur les matrices
3. Régression simple : Modèle, Contraintes, Inférence
4. Régression multiple : Modèle, Contraintes, Inférence, Sélection de variables
5. ANOVA à 1 facteur : Modèle, Contraintes, Tests.
6. ANOVA à 2 facteurs (cas équilibrés ou non) : Modèle, Contraintes, Tests
7. ANCOVA : Modèle, Contraintes, Tests

Pré-requis

Module d'initiation au logiciel R

Evaluation :

Une évaluation unique portera sur l'ensemble du parcours obligatoire.

Modalités pédagogiques : A distance