



Ecole d'été de Santé Publique et d'Epidémiologie de Paris-Saclay



O.Lambert et A.Nambiema

1. Régression multiple en épidémiologie : régression linéaire, logistique et modèle de Cox

1 semaine temps complet du 22 juin au 26 juin 2026

Objectifs et contenu

L'objectif du cours est de présenter deux des principaux modèles de régression multivariée les plus fréquemment utilisés en épidémiologie et en recherche biomédicale : la régression logistique et le modèle de Cox. Il s'agira de développer une démarche d'analyse, depuis la vérification des hypothèses inhérentes à chaque modèle, le choix et le codage des facteurs de confusion, jusqu'à l'interprétation des résultats. Certains aspects moins usuels seront abordés (variable dépendante du temps dans le modèle de Cox). A côté des bases théoriques sous-tendant l'utilisation de ces modèles, de nombreux exemples et situations concrètes seront présentés, notamment sur la base de lectures critiques d'articles. A l'issue de cet enseignement, les étudiants auront acquis les connaissances leur permettant d'utiliser les principaux modèles de régression en fonction de leur problématique, et d'avoir un regard critique dans l'interprétation des résultats.

Pré-requis

Enseignement de base ou cours de statistique et d'épidémiologie de niveau CESAM, M1 de sciences biomédicales, de santé publique ou de mathématiques. Avoir des connaissances de base concernant le modèle de régression linéaire.

Plan

I. Régression linéaire et logistique (5 demi-journées) :

Rappels sur la régression linéaire (2h)

La régression logistique :

- Présentation du modèle, estimation et tests des paramètres
- Codage des variables et interprétation des coefficients - Evaluation du modèle
- Sélection des variables selon l'objectif de la modélisation (étiologique ou prédictif)
- Introduction à la méthode de sélection de variables par pénalisation LASSO

II. Données de survie (5 demi-journées) :

Ce module est consacré à l'analyse des données de survie, largement utilisées en épidémiologie et en santé publique pour l'étude du temps jusqu'à la survenue d'un événement (décès, maladie, rechute, etc.). Il aborde les méthodes d'estimation et de comparaison des courbes de survie, ainsi que l'utilisation du modèle de Cox pour l'analyse multivariée. Contenu du module :

- Problématique des données de survie (censure, temps à l'événement)
- Construction des courbes de survie :
 - méthode de Kaplan-Meier
 - méthode actuarielle
- Comparaison des courbes de survie : test du logrank
- Modèle de Cox