

# 2EL1750 - Statistiques avancées

Responsables : **Sarah LEMLER**

Département de rattachement : **DÉPARTEMENT MATHÉMATIQUES**

Langues d'enseignement : **FRANCAIS**

Type de cours : **Electif 2A**

Campus où le cours est proposé : **CAMPUS DE PARIS - SACLAY**

Nombre d'heures d'études élèves (HEE) : **60**

Nombre d'heures présentielles d'enseignement (HPE) : **30**

Année académique : **2024-2025**

Catégorie d'électif : **Sciences fondamentales**

Niveau avancé : **oui**

## Présentation, objectifs généraux du cours :

Le cours de Statistiques Avancées a pour objectif de présenter diverses méthodes statistiques permettant de faire de l'estimation, de la prédiction et de déterminer les propriétés des estimateurs proposés. Nous développerons deux grands thèmes :

- la régression linéaire multivariée
- les statistiques non-paramétriques.

Le cours aborde à la fois les aspects théoriques de ces notions mais propose également une mise en pratique des modèles et méthodes considérées à l'aide de TP avec le logiciel R sur des jeux de données provenant de différents domaines.

A télécharger avant le premier TP :

- le logiciel R <https://www.r-project.org/>
- l'interface RStudio <https://www.rstudio.com/>

## Période(s) du cours (n° de séquence ou hors séquence) :

SG8

## Prérequis :

Pour ce cours les prérequis sont les suivants :

Il faut connaître les notions vues en cours de Statistiques et Apprentissage en 1ère année, en particulier :

- les estimateurs,
- les intervalles de confiance,
- les tests,
- les premières notions sur le modèle linéaire

## Plan détaillé du cours (contenu) :

Nous développerons deux thèmes :

- la régression linéaire multivariée
- les statistiques non-paramétriques

## Déroulement, organisation du cours :

Le cours se fera au tableau (pour la partie la plus théorique) ou à partir de slides projetées lorsqu'il s'agira par exemple de présenter une application faite à partir du logiciel R.

Il y a 28,5 heures prévues pour cet électif, dont 2 heures pour l'examen final.

## Organisation de l'évaluation :

Le cours comporte deux évaluations : un devoir maison (DM) obligatoire à composer par binôme à rendre à mi parcours (la date sera précisée lors du premier cours) et un examen final (EX) sur table de 2 heures qui porte sur l'ensemble du programme traité en cours. La note de cet électif sera alors la moyenne pondérée des notes des deux évaluations précédentes  $((1/3)*DM+(2/3)*EX)$ , arrondie au demi point le plus proche. Le rattrapage est un examen EX2 de même modalités que EX. La note de l'électif après le rattrapage est celle de l'examen EX2 uniquement (sans le DM).

## Moyens :

A l'issue de chaque cours, des exercices de TDs ou des TPs seront proposés pour se familiariser avec les notions théoriques vues en cours et les mettre en pratique pour répondre à des problèmes concrets éventuellement à partir de jeux de données.

## Acquis d'apprentissage visés dans le cours :

- Utiliser des techniques d'estimation statistiques paramétriques et non-paramétriques
- Valider une modélisation et comprendre les limites d'un modèle statistique
- Proposer, implémenter et calibrer un modèle prédictif
- Utiliser le logiciel R et interpréter des résultats

## Description des compétences acquises à l'issue du cours

:

- Use of parametric and non-parametric statistical estimation techniques
- Validate a model and understand the limits of a statistical model
- Propose, implement and calibrate a predictive model
- Use the R software and interpret results

## Bibliographie :

Le cours de Statistiques et apprentissage de première année de Arthur Tenenhaus