

MODELISATION STATISTIQUE AVEC R

Code UE : HAA708V

Crédits : 5 ECTS

Semestre : 1

Heures enseignées : TD 42h

RESPONSABLE(S) PEDAGOGIQUE(S)

François BOUTIN, francois.boutin@umontpellier.fr

Guilhem DESBROSSES, guilhem.desbrosses@umontpellier.fr

OBJECTIFS

1. Savoir identifier une « bonne » problématique de recherche et utiliser R pour y répondre.
2. Savoir réaliser des tests « simples » de comparaison de moyennes et pourcentages.
3. Savoir utiliser un modèle adapté (linéaire, linéaire généralisé, mixte, survie, non linéaire...).
4. Savoir construire un plan d'expérience (complet ou incomplet).
5. Savoir optimiser un modèle (pas à pas, ou par « pénalisation »).
6. Savoir valider un modèle (validation simple ou croisée sur jeu de test).
7. Savoir utiliser RMarkdown pour construire un rapport dynamique et reproductible.
8. Savoir analyser et critiquer un article scientifique de A à Z (méthodologie et résultats).

DESCRIPTION DU CONTENU DE L'UE

Un modèle statistique cherche à « expliquer » une réponse biologique par un ou plusieurs facteurs.

Ce cours présente les diverses étapes de la modélisation statistique en agronomie : du choix d'un modèle, à son optimisation et sa validation, avec le logiciel libre et gratuit R.

Le choix d'un « bon » modèle dépend notamment de la nature des variables étudiées (quantitative ou qualitative) et des relations entre elles :

- Une analyse de variance (modèle linéaire) pourra ainsi expliquer le rendement d'une serre, par le niveau d'éclairage (modéré ou élevé) et le type d'arrosage (continu ou non).
- Une régression logistique (modèle linéaire généralisé) pourra expliquer le risque de maladie, par le traitement utilisé (A, B ou C), sa dose (simple ou double) et l'âge du « sujet ».
- D'autres modèles (mixte, survie, non linéaire...) seront adaptés à des situations particulières.

Un plan d'expérience pourra être utilisé pour optimiser les ressources disponibles.

L'analyse statistique d'un modèle permet d'identifier les facteurs « explicatifs », d'évaluer la qualité prédictive du modèle et d'effectuer des prédictions à partir de nouvelles données.

PRE-REQUIS

Néant

CONTROLE DES CONNAISSANCES :

Contrôlé continu (travail à réaliser seul, en binôme ou trinôme) :

- Mener une lecture critique d'article scientifique au format IMReD.
- Mener une étude statistique et construire un « mini » article associé.