

R5.06 Modélisation statistique avancée

Composante
Institut universitaire de technologie de Poitiers-Châtelleraut-Niort

Présentation

Description

L'objectif de cette ressource est de présenter les méthodes de base en statistique paramétrique et non paramétrique, ainsi que leurs applications essentielles (études de modèles multi-variés, critères de sélection de variables, critères de sélection de fenêtre ou de choix de partition, ...).

Contenus :

Partie 1 : Statistique paramétrique

- Principe de la méthode des moments et de la méthode delta (cas d'un modèle à un paramètre)
- Principe de la méthode du maximum de vraisemblance (cas d'un modèle à un paramètre)
- Estimation par maximum de vraisemblance dans le modèle exponentiel à 1 paramètre (modèle canonique et cas général). Propriétés des estimateurs du maximum de vraisemblance (variance asymptotique minimale).
- Exemples d'estimation par maximum de vraisemblance dans les modèles multi-variés (loi Gamma et loi Gaussienne).
- Estimation par maximum de vraisemblance dans le modèle linéaire Gaussien : calcul des estimateurs du maximum de vraisemblance et tests du rapport de vraisemblance.
- Modèle de régression logistique multi-varié : définition, interprétation, tests de Wald.
- Critères de sélection basés sur la vraisemblance : AIC et BIC. Application au modèle linéaire Gaussien et au modèle de régression logistique.

Partie 2 : Introduction à la statistique non paramétrique

- Estimateurs à noyau de la densité.
- Estimateurs par Histogramme de la densité sur un intervalle : calcul du biais de la variance des estimateurs (cas d'une densité à dérivée bornée et d'une partition régulière).
- Estimateur par Régressogramme de la fonction de régression sur un intervalle : calcul du biais et de la variance des estimateurs (cas d'un design fixe, partition régulière et fonction de régression à dérivée bornée).
- Présentation de quelques règles de sélection de fenêtre et de sélection de partition (validation croisée, méthode de pénalisation, ...)

Partie 3

- Statistiques avancées pour un domaine d'application donné (par exemple données de survie en stat pour la santé), plan d'expérience (anova plusieurs facteurs), ...

Ces notions permettent aux étudiants du parcours d'appréhender les principales méthodes d'estimation paramétriques et non paramétriques, ainsi que les principales méthodes de sélection automatique (sélection de variables, sélection de fenêtre, sélection de partition, ...)

Heures d'enseignement

TD	TD	48h
----	----	-----