

Statistique inférentielle et analyse de données



Présentation

Code interne : CO6SFMA0

Description

Ce module vise à la maîtrise des notions fondamentales des outils de statistique inférentielle (estimation ponctuelle, intervalles de confiance, tests d'hypothèses) et d'analyse des données d'un point de vue théorique et appliqué (avec l'utilisation du logiciel R).

Mots-clés :

Estimation ponctuelle, intervalle de confiance, test d'hypothèses.

Analyse en composantes principales (ACP), données multidimensionnelles.

Pré-requis obligatoires

Les prérequis de ce module concernent l'algèbre linéaire, l'analyse, ainsi que le calcul des probabilités et la statistique descriptive (module CO5SFMA0).

Syllabus

Partie 1 : Statistique inférentielle (7h CM, 11h TD)

Chapitre 1 : Introduction générale à la statistique inférentielle

Quelques généralités

Modèle statistique

Définition d'une statistique

Notions d'estimation et de tests d'hypothèses

Chapitre 2 : Estimation ponctuelle (cas d'un paramètre unidimensionnel)

Quelques notions de base sur les estimateurs

Propriétés des estimateurs d'une espérance et d'une variance

Inégalité de Cramer-Rao

Quelques méthodes usuelles d'estimation

Chapitre 3 : Estimation par intervalles de confiance

Définition d'un intervalle de confiance



Quelques Remarques importantes

Liens entre les bornes de l'intervalle de confiance

Intervalle de confiance d'une moyenne, d'une variance et d'une proportion

Chapitre 4 : Quelques tests d'hypothèses paramétriques

Introduction aux tests d'hypothèses

Tests d'hypothèses portant sur la moyenne ou sur la variance

Analyse statistique de deux échantillons (indépendants ou appariés)

Chapitre 5 : Quelques tests d'hypothèses non paramétriques

Test du khi-deux d'indépendance

Test du khi-deux d'ajustement

Test de Wilcoxon pour deux échantillons indépendants

Test (des rangs signés) de Wilcoxon pour deux échantillons appariés

Partie 2 : Analyse des données (5h CM, 5h TD)

Chapitre 1 : Introduction générale à l'analyse des données

Chapitre 2 : Analyse en composantes principales (ACP)

Présentation et objectifs de l'ACP

L'espace des variables

L'espace des individus

Calculs des axes principaux, des facteurs principaux et des composantes principales

ACP centrée réduite

Qualité de représentation sur les plans principaux

Interprétation des résultats d'une ACP

Variables et individus supplémentaires

Documents autorisés (da) pour l'épreuve d'examen des parties 1 et 2 :

Seule une feuille de format A4 recto-verso de notes personnelles et manuscrites par l'étudiant est autorisée lors des épreuves d'examen (en sessions 1 et 2). Les photocopies sont donc proscrites. Ce document sera rendu avec la copie de l'épreuve et sera restitué à la demande de l'élève à l'issue de la correction.

"Philosophie des TD" :

Partie 1 : Statistique inférentielle

L'objectif des TD est la compréhension des outils de statistique inférentielle présentés en cours.

Une première partie des TD porte sur l'étude théorique de certains estimateurs qui n'ont pas été étudiés en cours.

Une seconde partie des TD concerne l'interprétation de sorties numériques et graphiques issues du logiciel de statistique R illustrant les notions de statistique inférentielle présentées en cours (estimation ponctuelle et par intervalle de confiance, test d'hypothèses) dans le cadre de problématiques réelles.

Partie 2 : Analyse des données

Les TD portent sur l'interprétation de sorties numériques et graphiques issues du logiciel de statistique R illustrant l'application de la méthode d'ACP associées à des problématiques réelles (environnement, biologie, ...).

"Philosophie des TP" :

Partie 1 : Statistique inférentielle

L'objectif des TP est la mise en œuvre pratique avec le logiciel R des notions de statistique descriptive (vues au module CO5SFMA0) et de statistique inférentielle traitées en cours/TD (dans le cadre de ce module) en utilisant des données des données réelles associées à des problématiques diverses (environnement, biologie, ...).

Partie 2 : Analyse des données

L'objectif des TP est la mise en œuvre pratique avec le logiciel R des notions de la méthode d'ACP en utilisant des données des données réelles associées à des problématiques réelles (environnement, biologie, ...).

"Philosophie du projet" :



Parties 1 et 2 : Statistique inférentielle et Analyse des données

Les élèves-ingénieurs doivent faire un projet sous R dans le cadre de leur contrôle continu (CC). L'objectif du projet est la mise en œuvre sur des données et problématiques réelles des notions de statistique descriptive et inférentielle et d'analyse des données traitées en cours/TD/TP dans le cadre de ce module CO6SFMA0 et du module CO5SFMA0.

Informations complémentaires

Mathématiques appliquées

Modalités de contrôle des connaissances

Évaluation initiale / Session principale - Épreuves

| Type d'évaluation | Nature de l'épreuve | Durée (en minutes) | Nombre d'épreuves | Coefficient de l'épreuve | Note éliminatoire de l'épreuve | Remarques |
|-------------------|---------------------|--------------------|-------------------|--------------------------|--------------------------------|-----------|
| Projet | Rapport | | | 1 | 6.0 | |
| Epreuve Terminale | Ecrit | 90 | | 1 | 6.0 | |

Seconde chance / Session de rattrapage - Épreuves

| Type d'évaluation | Nature de l'épreuve | Durée (en minutes) | Nombre d'épreuves | Coefficient de l'épreuve | Note éliminatoire de l'épreuve | Remarques |
|-------------------|---------------------|--------------------|-------------------|--------------------------|--------------------------------|---|
| Epreuve terminale | Ecrit | 60 | | 1 | 6.0 | documents autorisés calculatrice autorisée |

Infos pratiques



Contacts

Responsable module

Jerome Saracco

✉ Jerome.Saracco@bordeaux-inp.fr