UCLouvain

linge1222

2019

Analyse statistique multivariée

Au vu du contexte sanitaire lié à la propagation du coronavirus, les modalités d'organisation et d'évaluation des unités d'enseignement ont pu, dans différentes situations, être adaptées ; ces éventuelles nouvelles modalités ont été -ou seront- communiquées par les enseignant·es aux étudiant·es.

4 crédits	30.0 h + 15.0 h	Q2

Enseignants	Segers Johan ;Uyttendaele Nathan (supplée Segers Johan) ;
Langue d'enseignement	Français
Lieu du cours	Louvain-la-Neuve
Préalables	*Le(s) prérequis de cette Unité d'enseignement (UE) sont précisés à la fin de cette fiche, en regard des programmes/formations qui proposent cette UE.*
Thèmes abordés	Partie 1 : Méthodes descriptives de base et notations de base. Il s'agit ici de montrer comment la notation matricielle permet un traitement aisé des données multidimensionnelles mais aussi des propriétés de base des vecteurs aléatoires. On voit aussi que les outils descriptifs de base (uni- et bivariés) ont leur utilité mais aussi leur limites. Partie 2 : Techniques d'analyse de données multivariées. On aborde ici les techniques de base de réduction de dimension pour variables continues mais aussi pour variables qualitatives (composantes principales, analyse des correspondances). Les techniques de base de la classification sont également présentées. Les différentes méthodes sont motivées et illustrées par de nombreux exemples. Partie 3 : Modèles d'analyse multivariées. Dans cette partie, on montre comment on peut modéliser les relations pouvant exister entre des variables : modèles linéaires (y compris l'analyse de la variance et l'analyse de variance-covariance) qui permettent d'expliquer les variations d'une variable " réponse " en fonction de variables explicatives. Les modèles adaptés à des variables réponses catégorielles sont également présentés, modèles log-linéaires pour tables de contingence, modèle logit mais aussi les modèles de l'analyse discriminante. Les différentes méthodes sont ici aussi, motivées et illustrées par de nombreux exemples
Acquis d'apprentissage	1 Développer les éléments introduits dans les cours de base de probabilités et statistique dans un cadre multivarié ; le but étant de donner les outils permettant d'analyser des ensembles de données multidimensionnels. A la sortie du cours l'étudiant devrait être capable de mettre en #uvre, avec des données réelles, les outils les plus classiques de traitement. Le cours sera donc centré sur la bonne compréhension des méthodes et leur mise en application, y compris la maîtrise d'un logiciel approprié. - - - - *La contribution de cette UE au développement et à la maîtrise des compétences et acquis du (des) programme(s) est accessible à la fin de cette fiche, dans la partie « Programmes/formations proposant cette unité d'enseignement (UE) ».*
Modes d'évaluation des acquis des étudiants	**En raison de la crise du COVID-19, les informations de cette rubrique sont particulièrement susceptibles d'être modifiées.** • Test informatique: vers la fin du cours, les étudiants répondent à des questions à choix multiple liées à de vrais jeux de données en s'assistant du logiciel R. Ce test est à livre ouvert. • Examen écrit, à livre fermé, à l'aide d'un formulaire et d'une calculatrice. L'examen comprend à la fois des questions de théorie et des exercices liés à l'interprétation et la reconstruction de sorties du logiciel R.
Méthodes d'enseignement	**En raison de la crise du COVID-19, les informations de cette rubrique sont particulièrement susceptibles d'être modifiées.** • Exposés magistraux : l'enseignant introduit les concepts à partir d'une application et dégage leur forme abstraite • Séances d'exercices en salle informatique : l'enseignant y soumet des applications/problèmes aux étudiants sur base de données réelles. La résolution se fait à l'aide du logiciel R.
Contenu	• Introduction aux méthodes multivariées • Algèbre linéaire et géométrie euclidienne • Statistiques descriptives pour matrices de données • Analyse en composantes principales • Classification non supervisée : centre mobiles et classification hiérarchique • Analyse discriminante linéaire • Théorie de distributions • Régression linéaire multiple, y compris AN(C)OVA • Régression logistique

Ressources en ligne	Le formulaire, les slides projetés lors des cours magistraux et des séances d'exercices, de la documentation du logiciel R et des liens vers des ressources web externes (vidéos, cours en ligne, documents) sont disponibles sur la page MoodleUCL du cours.
Bibliographie	• Härdle, W. and L. Simar (2007): Applied Multivariate Statistical Analysis, 2nd Edition, Springer-Verlag, Berlin. • James, G., Witten, D., Hastie, T. and R. Tibshirani (2013): An Introduction to Statistical Learning, Springer, New York. • Saporta, G. (2011): Probabilités, analyse des données et statistique, 3e édition révisée, Editions TECHNIP, Paris.
Faculté ou entité en charge:	ESPO

Programmes / formations proposant cette unité d'enseignement (UE)

Intitulé du programme	Sigle	Crédits	Prérequis	Acquis d'apprentissage
Certificat d'université : Statistique et sciences des données (15/30 crédits)	STAT2FC	4		
Bachelier en sciences mathématiques	MATH1BA	4		
Bachelier : ingénieur de gestion	INGE1BA	4	LINGE1113 ET LINGE1121	