UCLouvain

lstat2130

2023

Introduction to Bayesian statistics

5.00 crédits	22.5 h + 7.5 h	Q2

Enseignants	Lambert Philippe ;
Langue d'enseignement	Anglais
Lieu du cours	Louvain-la-Neuve
Préalables	Concepts et outils équivalents à ceux enseignés dans les UE LSTAT2020 Logiciels et programmation statistique de base LSTAT2120 Linear models LSTAT2190 Concepts et traitement de vecteurs aléatoires
Thèmes abordés	- Le modèle bayesien: principes généraux. - La fonction de vraisemblance et spécification a priori. - Modèles à un paramètre: choix de la distribution a priori, calcul de la distribution a posteriori, résumer la distribution a posteriori. - Modèles multiparamètres: choix des distributions a priori et calcul des distributions a posteriori, paramètres de nuisance. Cas des modèles multinomial et gaussien multivarié. - Inférence en grand échantillon et relation avec l'inférence fréquentiste. - Méthodes de calcul en analyse bayesienne.
Acquis d'apprentissage	**A la fin de cette unité d'enseignement, l'étudiant est capable de :** 1 A. Eu égard au référentiel AA du programme de master en statistique, orientation générale, cette activité contribue au développement et à l'acquisition des AA suivants, de manière prioritaire : 1.1, 1.3, 1.4, 2.2, 2.3, 2.5, 3.1, 3.2, 3.3, 3.4, 4.1, 4.2, 6.3 Eu égard au référentiel AA du programme de master en statistique, orientation biostatistique, cette activité contribue au développement et à l'acquisition des AA suivants, de manière prioritaire : 1.1, 1.3, 1.4, 2.2, 2.3, 2.5, 3.1, 3.2, 3.3, 3.4, 4.1, 4.2, 4.3, 4.5, 6.3 B. Au terme du cours l'étudiant aura acquis les principes et les techniques de base de la statistique bayesienne, et sera capable de les utiliser et de mettre en évidence leurs avantages et inconvénients dans des problèmes simples.
Modes d'évaluation des acquis des étudiants	L'évaluation de ce cours combinera un examen écrit (sur 15 points) et un travail (sur 5 points). Ce travail prendra la forme d'un rapport écrit remis par chaque étudiant avant la fin du 2e quadrimestre, sans possibilité de le représenter en 2e session. Cependant, la note obtenue pour ce travail interviendra de manière identique durant les 2 sessions pour le calcul de la note finale.
Méthodes d'enseignement	Le cours comprend des exposés magistraux, éventuellement complétés par des podcasts mis à disposition des étudiants sur Moodle, et des séances d'exercices. Il se donne sur base hebdomadaire et se répartit sur 11 semaines dès le début du 2e quadrimestre.
Contenu	Ce cours propose une introduction à la statistique bayesienne. Après avoir précisé la notion de probabilité subjective, les principes de base de l'inférence bayesienne sont exposés avec l'estimation d'une proportion. Ces mêmes principes sont ensuite mis en oeuvre pour comparer des proportions et des taux. L'estimation de la moyenne (variance) d'une distribution normale lorsque la variance (moyenne) est connue est également envisagée. Après une présentation du principe du maximum d'entropie, le paradigme bayesien est utilisé pour faire de l'inférence dans des modèles multi-paramètres. Les notions de distributions marginales et conditionnelles a posteriori, de région de crédibilité et de distribution prédictive sont définies. Une première illustration de l'utilité des ces concepts est apportée avec l'estimation conjointe de la moyenne et de la variance d'une distribution normale. Le problème de la comparaison des moyennes de deux distributions normales de variance(s) connue(s) ou inconnue(s) est ensuite abordé. La simulation d'un échantillon aléatoire de la distribution conjointe a posteriori est la piste suivie lorsque les variances ne peuvent pas être supposées identiques. Les modèles de régression multiple et d'analyse de la variance à un critère sont également étudiés dans un cadre bayesien. La production d'échantillons aléatoires à partir d'une distribution a posteriori est une étape essentielle en inférence bayesienne lorsque les modèles impliqués sont complexes. Les algorithmes fondamentaux utilisés à cette fin sont présentés et illustrés. Le cours se termine par une introduction aux modèles hiérarchiques.
Ressources en ligne	Transparents et podcasts sont à disposition des étudiants sur Moodle

Bibliographie	Gelman, A., Carlin, J.B., Stern, H.S., Dunson, D.B., Vehtari, A. and Rubin, D.B. (2013,3nd edition) Bayesian Data Analysis. Chapman and Hall - CRC Press. Bolstad, W.M. and J.M. Curran (2016) Introduction to Bayesian Statistics. Wiley.
Autres infos	Prérequis: il est supposé que les étudiants disposent d'une formation de base en théorie des probabilités, en inférence statistique et à l'utilisation du logiciel R.
Faculté ou entité en charge:	LSBA

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

Intitulé du programme	Sigle	Crédits	Prérequis	Acquis d'apprentissage
Master [120] en science des données, orientation statistique	DATS2M	5		
Master [120] : ingénieur civil biomédical	GBIO2M	5		
Master [120] en statistique, orientation biostatistiques	BSTA2M	5		
Master [120] en sciences mathématiques	MATH2M	4		
Master [120] en statistique, orientation générale	STAT2M	5		
Approfondissement en statistique et sciences des données	APPSTAT	5		
Master [120] : ingénieur civil en mathématiques appliquées	MAP2M	5		
Master [120] en sciences économiques, orientation générale	ECON2M	5		
Master [120] : ingénieur civil en science des données	DATE2M	5		
Certificat d'université : Statistique et science des données (15/30 crédits)	STAT2FC	5		
Master [120] en science des données, orientation technologies de l'information	DATI2M	5		