

STT 3910 - Introduction au modèle général linéaire

Automne 2020

Plan de cours

Chargé de cours : Serge Vicente
Bureau : 4245, pavillon André Aisenstadt
Courriel : vicentes@dms.umontreal.ca

I Horaire du cours

Lundi de 8h30 à 11h30 - à distance (du 14/09/2020 au 05/10/2020 et du 02/11/2020 au 07/12/2020).

II Objectifs du cours

L'objectif du cours est de familiariser les étudiants avec les principaux outils d'analyse de données provenant d'études dans le domaine des neurosciences. Le cours PSY1004 a permis de découvrir les statistiques descriptives, qui permettent de décrire les données en main et d'en faire un sommaire utile pour pouvoir généraliser les résultats à l'ensemble de la population (inférence statistique). Pour cela, plusieurs techniques ont été abordées, comme les tests d'hypothèses, la régression linéaire et l'analyse de la variance (ANOVA). Cependant, ces techniques sont applicables à des études relativement simples et, lorsque qu'une étude plus complexe est développée, celles-ci se révèlent insuffisantes ou impraticables pour analyser les données. Le cours STT3910 reviendra sur ces méthodes de façon à élargir leur champ d'application et introduira d'autres méthodes d'inférence comme la régression logistique, la régression de Poisson, le modèle général linéaire et autres. Comme la plupart de ces méthodes sont basées sur des matrices et des vecteurs, les notions vues en MAT1600 seront utiles. Finalement, comme de nos jours l'analyse de données implique nécessairement un contact avec des logiciels statistiques, le cours permettra aux étudiants d'avoir un contact permanent avec eux.

III Plan du cours

1. Introduction
2. Régression linéaire et Analyse de la variance (ANOVA). Sélection de modèles.
3. Modèle de régression linéaire généralisé : cas de la régression logistique et de la régression de Poisson.
4. Modèles linéaires mixtes.
5. Techniques avancées d'analyse en neurosciences :
 - Analyse en composantes principales ;
 - Analyse canonique des corrélations ;
 - Partitionnement de données (*clustering*).
6. Analyse de données d'imagerie par résonance magnétique fonctionnelle (IRMf).

IV Évaluation

La note finale du cours est déterminée de la façon suivante :

1. **Examens** : comptent pour 40% de la note finale, répartis de la façon suivante :
 - *Examen intra* : compte pour 15% de la note finale et aura lieu le lundi 26 octobre 2020, de 8h30 à 10h30.
 - *Examen final* : compte pour 25% de la note finale et aura lieu le lundi 14 décembre 2020, de 9h à 12h.
2. **Études de cas** : comptent pour 60% de la note finale, répartis en 3 études, à titre de 20% chacune. Les examens sont composés par des questions à choix multiples et par des questions à développement, qui portent sur les aspects théoriques vus pendant le cours. Les 3 études de cas consistent en une analyse de données réelles provenant du domaine des neurosciences et faisant intervenir les méthodes d'analyse vues pendant le cours. Les études de cas sont à remettre sous forme de rapport écrit. Pour réussir le cours, la note globale doit être égale ou supérieure à 50%.

V Dates limites

Les dates d'annulation et abandon du cours sont les suivantes :

1. Dernier jour pour modifier un choix de cours et annuler un cours sans frais : 17 septembre 2020
2. Dernier jour pour abandonner un cours avec frais : 6 novembre 2020

VI Plagiat

L'Université de Montréal possède un Règlement disciplinaire sur le plagiat ou la fraude concernant les étudiants. Il est du devoir de l'étudiant d'en prendre connaissance. Les sanctions visant un étudiant reconnu coupable varient selon la gravité de l'infraction : l'attribution de la note F pour l'examen ou le cours en cause, la suspension du programme pour un trimestre, etc. Plusieurs documents ont été préparés par la Faculté des arts et des sciences pour initier les étudiants aux responsabilités sous-jacentes au travail intellectuel et prévenir la malhonnêteté académique. Ces documents ainsi que le Règlement disciplinaire sont accessibles sur la page <http://www.integrite.umontreal.ca/>

VII Bibliographie

Les ouvrages de référence **obligatoires** sont les suivants :

Kass, R.E., Eden, U., & Brown, E.N. (2014). Analysis of Neural Data. Springer.

Durstewitz, Daniel (2017). Advanced Data Analysis in Neuroscience - Integrating Statistical and Computational Models. Springer.

Les deux autres ouvrages suivants sont recommandés comme compléments :

Madsen, H. and Thyregod, P. (2012). Introduction to General and Generalized Linear Models. CRC Press - Chapman & Hall.

Lafaye de Micheaux, P., Drouilhet, R. & Liquet, B. (2011). Le logiciel R : Maîtriser le langage Effectuer des analyses statistiques. Deuxième Édition. Springer.