UQÀM | Département de mathématiques
FACULTÉ DES SCIENCES
Université du Québec à Montréal

PLAN DE COURS
MAT 998K - GÉOSTATISTIQUE ET TESTS MULTIPLES - 3 CRÉDITS (AUTOMNE 2019)

Professeur: Jean-François Coeurjolly
coeurjolly.jean-francois@uqam.ca
Département de Mathématiques, Pavillon Kennedy,
Faculté des Sciences, UQAM
Bureau 5525.
Disponibilités: Jeudi 10:00-12:00.

1 Description du cours

1.1 Objectifs

Ce cours est divisé en deux parties. La première partie a pour objectif d'introduire des concepts et méthodologies statistiques adéquates en présence de données spatiales et en particulier d'un champ aléatoire. La seconde partie, totalement distincte vise à appréhender la problématique de tests multiples, problématique intervenant dans tous les champs de la statistique.

Les modèles et méthodologies seront discutés mathématiquement et appliqués à des jeux de données simulées et réelles en utilisant le logiciel R.

1.2 Mots-clés

Données spatiales; modélisation; estimation; prédiction; champ aléatoire; tests multiples; FWER/FDR.

1.3 Objectifs visés pour l'étudiant

A la fin du cours, l'étudiant aura acquis une certaine expérience face à des données ayant des composantes spatiales, saura identifier la nature de ces données et sera à même d'utiliser la méthodologie statistique adéquate lorsque les données seront la réalisation d'un champ

aléatoire. D'autre part, l'étudiant sera sensibilisé au problème des tests multiples notamment dans un cadre de grande dimension et connaître quelques procédures standard permettant de corriger ce problème.

Le cours étant également basé sur l'utilisation de paquets R spécifiques, l'étudiant saura analyser un jeu de données spatiales (rentrant dans le cadre du cours) par lui-même. D'un point de vue plus théorique et mathématique, ce cours se veut être une porte d'entrée pour lire en profondeur et avec une vision plus recherche les ouvrages cités ci-dessous.

1.4 Prérequis

Cours de statistique inférentielle de niveau 1 et 2 (STT-1000 et STT-2000 à l'UQAM), de probabilités (MAT-1700, éventuellement MAT-2710 à l'UQAM), de Régression (par exemple STT-2120 ou MAT7381 à l'UQAM); connaissances du logiciel R obligatoires.

1.5 Horaire et format du cours

Les cours débutent le vendredi 6 septembre et se terminent le vendredi 13 décembre. Ils auront lieu les jeudi de 10h00 à 13h00 SH-3760.

Chaque séance de cours prendra la forme d'un exposé magistral assuré par le professeur.

Dans la mesure du possible d'une semaine sur l'autre, l'étudiant se verra proposer des séries d'exercices en lien avec le cours (démonstrations de résultats du cours, applications de résultats ou applications numériques). Il pourra s'agir d'exercices théoriques ou pratiques et dans ce cas ils seront à réaliser avec le logiciel R. L'étudiant aura à charge d'imprimer les notes de cours (copie partielle des acétates présentées en cours) qu'il pourra compléter pendant la présentation du professeur.

La premère partie (géostatistique) devrait prendre approximativement 60-70% du temps tandis que la seconde partie (tests multiples) 30-40% du temps.

2 Evaluation des apprentissages

- *Examen écrit (30% de la note finale)*: un examen écrit de 3h sur la partie I du cours.
- *Projet de géostatistique (20% de la note finale)*: projet décrit lors du cours d'introduction; à préparer par équipe de deux et à rendre pour mi-novembre.
- *Devoirs (25% de la note finale)*: Au cours de la session, plusieurs devoirs seront à rendre (sur les deux parties du cours). L'étudiant aura en général une semaine pour préparer ce devoir. Il pourra s'agir de démonstrations de cours, d'applications en R des méthodologies vues en cours, ou de résumés d'articles en lien avec le cours et qui seront discutés le cours d'après. Les devoirs devront être rendus sous format .pdf et

utiliser l'environnement `Rmarkdown`. Le sujet sera déposé une semaine avant sa date de remise sur le moodle associé au cours. La non remise d'un devoir dans le temps imparti impliquera nécessairement une note 0 au devoir correspondant.

- *Présentation article de recherche (25% de la note finale)*: présentation par équipe d'un article de recherche (fourni par le professeur) en lien avec la partie tests multiples.

Remarques importantes: La présence en classe est fortement recommandée pour bien réussir le cours. Tout acte de tricherie pendant les examens ou tout plagiat entraînera de lourdes sanctions.

Calendrier prévisionnel: (le calendrier est susceptible d'être modifié légèrement)

- Examen géostatistique : date à définir; 3h date précisée ultérieurement (fin octobtre approximativement)
- Mi-novembre (date précisée ultérieurement): dépôt d'un rapport pour le projet de géostatistique.
- Vendredi 13 décembre: présentation d'un article de recherche.

3 Plan du cours

1. **Introduction générale au cours d'introduction à la statistique spatiale**: exemples et objectifs, établissement des prérequis, présentation du matériel et des logiciels utilisés en particulier R/RStudio/Rmarkdown si cela s'avère nécessaire.
2. **Partie I: Géostatistique**
   (a) Processus stochastiques, champs aléatoires: définitions, notion de stationarité, fonction de covariance, représentation spectrale, notion de variogramme, stationarité intrinsèque.
   (b) Champs aléatoires gaussiens: rappels sur les vecteurs gaussiens, définitions, propriétés, simulation d'un champ gaussien, extensions.
   (c) Estimation de covariances, variogramme.
   (d) Krigeage ou plus simplement interpolation spatiale.
3. **Partie II: Tests multiples**
   (a) Définitions générales et appréhension du problème.
   (b) Contrôle d'erreurs de type I: FWER, k-FWER, correction de Bonferonni, procédure de Holm, procédures step-up, step-down.
   (c) Contrôle du False Discovery rate: procédure de Benjamini-Hochberg et extenstions.

4 Bibliographie

Le cours ne suit pas d'ouvrages particulier. Aussi, il n'est pas obligatoire de se procurer et lire les ouvrages suivants pour réussir avec succés le cours. Les notes de cours préparés par le professeur devraient être suffisantes. Néanmoins, elles pourraient parfois pour des lectures complémentaires renvoyer à certains chapitres de ces ouvrages.

1. Noel A.C. Cressie. *Statistics for Spatial Data*. Wiley Series in Probability and Mathematical Statistics, 1993.

2. Carlo Gaetan et Xavier Guyon. *Modélisation et Statistique Spatiale*. Vol.63 de Mathématiques et Applications, Springer 2008.

3. Jesper Møller and Rasmus Waagepetersen. *Statistical Inference for Spatial Point Processes*. Chapman and Hall, 2004.

4. S Dudoit and MJ Van Der Laan Multiple testing procedures with applications to genomics. Springer Science & Business Media, 2007.

Politique 16 sur le harcèlement sexuel
Le harcèlement sexuel se définit comme étant un comportement à connotation sexuelle unilatéral et non désiré ayant pour effet de compromettre le droit à des conditions de travail et d'études justes et raisonnables ou le droit à la dignité.
La Politique 16 identifie les comportements suivants comme du harcèlement sexuel : 1. Manifestations persistantes ou abusives d'un intérêt sexuel non désirées. 2. Remarques, commentaires, allusions, plaisanteries ou insultes persistants à caractère sexuel portant atteinte à un environnement propice au travail ou à l'étude. 3. Avances verbales ou propositions insistantes à caractère sexuel non désirées. 4. Avances physiques, attouchements, frôlements, pincements, baisers non désirés. 5. Promesses de récompense ou menaces de représailles, implicites ou explicites, représailles liées à l'acceptation ou au refus d'une demande d'ordre sexuel. 6. Actes de voyeurisme ou d'exhibitionnisme. 7. Manifestations de violence physique à caractère sexuel ou imposition d'une intimité sexuelle non voulue. 8. Toute autre manifestation à caractère sexuel offensante ou non désirée. Pour plus d'information : http://www.instances.uqam.ca/ReglementsPolitiquesDocuments/Documents/Politique_no_16.pdf Pour rencontrer une personne ou faire un signalement : Bureau d'intervention et de prévention en matière de harcèlement : 514-987-3000, poste 0886 http://www.harcelement.uqam.ca

RÈGLEMENT

SUR LES INFRACTIONS DE NATURE ACADÉMIQUE

UQÀM

TOLÉRANCE ZÉRO r18.uqam.ca

Tout acte de plagiat, fraude, copiage, tricherie ou falsification de document commis par une étudiante, un étudiant, de même que toute participation à ces actes ou tentative de les commettre, à l'occasion d'un examen ou d'un travail faisant l'objet d'une évaluation ou dans toute autre circonstance, constituent une infraction au sens de ce règlement.

La liste non limitative des infractions est définie comme suit :

- la substitution de personnes ;
- l'utilisation totale ou partielle du texte d'autrui en le faisant passer pour sien ou sans indication de référence ;
- la transmission d'un travail aux fins d'évaluation alors qu'il constitue essentiellement un travail qui a déjà été transmis aux fins d'évaluation académique à l'Université ou dans une autre institution d'enseignement, sauf avec l'accord préalable de l'enseignante, l'enseignant ;
- l'obtention par vol, manœuvre ou corruption de questions ou de réponses d'examen ou de tout autre document ou matériel non autorisés, ou encore d'une évaluation non méritée ;
- la possession ou l'utilisation, avant ou pendant un examen, de tout document non autorisé ;
- l'utilisation pendant un examen de la copie d'examen d'une autre personne ;
- l'obtention de toute aide non autorisée, qu'elle soit collective ou individuelle ;
- la falsification d'un document, notamment d'un document transmis par l'Université ou d'un document de l'Université transmis ou non à une tierce personne, quelles que soient les circonstances ;
- la falsification de données de recherche dans un travail, notamment une thèse, un mémoire, un mémoire-création, un rapport de stage ou un rapport de recherche.

Les sanctions reliées à ces infractions sont précisées à l'article 3 du Règlement no 18

Vous pouvez consulter sur le site r18.uqam.ca des capsules vidéos qui vous en apprendront davantage sur l'intégrité académique et le R18, tout en vous orientant vers les ressources mises à votre disposition par l'UQAM pour vous aider à éliminer le plagiat de vos travaux.