

PLAN DE COURS

Données catégorielles

STT6516

Mathématiques et statistique

Automne 2026 | Groupe A | 3 crédits

Équipe enseignante

ALEXANDRO MURUA

alejandro.murua@umontreal.ca

Bureau 4221 Pav. André-Aisenstadt

J 12h-13h, V 11h30 - 12h30

Apprentissages visés

Présentation du cours

Ce cours est une introduction aux modèles statistiques utilisés pour l'étude et l'analyse de données catégorielles nominales ou ordinales, ainsi que les données de comptage. Le cours couvre des sujets telles que les tableaux de contingence à plusieurs dimensions, les modèles de Rasch, le modèle log-linéaire, les modèles de régression logistique, de Poisson et multinomiale. Il adresse aussi des méthodes d'estimation et d'inférence basés sur la théorie asymptotique ou bayésienne. Les notions théoriques seront appliquées sur des exemples pratiques avec l'aide de l'ordinateur et du logiciel R.

Ce cours est une introduction aux concepts fondamentaux de la discipline. Il vise à fournir aux personnes étudiantes les outils théoriques et pratiques nécessaires pour aborder des problèmes complexes dans leur domaine. L'accent est mis à la fois sur la compréhension conceptuelle et sur l'application concrète des méthodes présentées.

Objectifs généraux

Ce cours vise à développer une maîtrise solide des concepts fondamentaux et de leurs applications. Les personnes étudiantes acquerront les outils nécessaires pour analyser des problèmes, formuler des solutions rigoureuses et communiquer leurs résultats de manière claire et structurée.

Objectifs d'apprentissage spécifiques

1. Maîtriser les concepts théoriques fondamentaux du cours et les appliquer dans des contextes variés.
2. Analyser des problèmes complexes en mobilisant les outils et méthodes appropriés.
3. Démontrer des résultats en produisant des arguments rigoureux et bien structurés.
4. Communiquer des solutions de façon claire, précise et adaptée au contexte.

Contenu et matériel didactique

Contenus

- Semaine 1 : introduction, tableaux de contingence
- Semaine 2 : tableaux de contingence à deux variables catégorielles ; distribution des données (Poisson, multinomiale, multinomiale par lignes ou colonnes) ; estimation par maximum de vraisemblance ; distribution asymptotique ; intervalles de confiance pour une probabilité
- Semaine 3 : tests d'association
- Semaine 4 : inférence conditionnelle ; test exact de Fisher ;
- Semaine 5 : modèle de Rasch ; valeur- p exacte approximative par Monte-Carlo ou Monte-Carlo par chaînes de Markov
- Semaine 6 : modèles log-linéaires (deux ou plusieurs variables catégorielles) ; modèles hiérarchiques
- Semaine 7 : semaine d'activités libres
- Semaine 8 : modèles log-linéaires, modèles hiérarchiques
- Semaine 9 : modèles graphiques
- Semaine 10 : modèle linéaire par linéaire ; théorie asymptotique ; sélection d'un modèle (critères d'information AIC, BIC) ; sélection par lasso, lasso de fusion
- Semaine 12 : régression logistique ; théorie asymptotique ; rapport de cotes
- Semaine 13 : régression logistique ; sélection de variables ; test d'adéquation
- Semaine 14 : régression multi-logit ; modèle multinomiale logistique cumulatif (variables ordinales)
- Semaine 15 : modèles à effets aléatoires ou mixtes.

Matériel didactique

- [AA] Alan Agresti (2013), *Categorical Data Analysis*, 3ème édition, John Wiley & Sons (*fortement recommandé*).

AAI Alan Agresti (2007), An Introduction to categorical Data Analysis, 2ème édition. John Wiley & Sons. (lecture supplémentaire. Non requis.)

BHF Yvonne M. Bishop, Stephen E. Fienberg et Paul W. Holland (2007), Discrete Multivariate Analysis : theory and applications, Springer. (lecture supplémentaire. Non requis.)

Déroulement et charge de travail

Format du cours

Le cours STT6516 a été conçu comme un cours traditionnel, ce qui signifie qu'il comprend à la fois des activités en classe (cours magistraux) et des activités hors classe (devoirs, examens à emporter et projets de trimestre).

Les séances en classe se déroulent les mardis et les jeudis (cours théoriques).

Méthodes d'enseignement

- Lectures (livre [AA])
- Exposés magistraux avec approfondissement de la matière, preuve de théorèmes (en classe)
- Forum de discussion pour des questions et rétroactions (Studium)

Charge de travail recommandée

Activité	Par semaine	Total
Lecture (notes de cours, livre [AA] ou autre)	3h/sem.	Total 45h
Travaux pratiques (Maison)	3h/sem.	Total 45h
Cours théorique (En classe)	3h/sem.	Total 45h

Évaluations et règles

Activité d'évaluation	Date	Pondération	
		description	Évaluation
Devoir 1	annonce : 17 / 09 / 2026	échéance : 01 / 10 / 2026	30%
Devoir 2	annonce : 01 / 10 / 2026	échéance : 15 / 10 / 2026	
Devoir 3	annonce : 12 / 11 / 2026	échéance : 26 / 11 / 2026	
Examen Intra	29 / 10 / 2026	annonce	30%
	03 / 11 / 2026	rapport	
Projet Final	16 / 10 / 2026	date limite pour la proposition de projet	40%
	Dernière semaine avant examens	présentation et rapport	

Consignes

- Complétion des devoirs : Aucun retard ne sera accepté, sauf en cas de situation exceptionnelle dûment justifiée.
- Seuil de réussite exigé : La note de passage est de 60% (C).
- L'examen intra : consistera à des applications particulières de l'analyse de données catégorielles à des problèmes pratiques réels. Il exigera l'exploration de données et une compréhension claire des matières présentées dans la classe et des tâches de lecture. Chaque étudiant(e) doit répondre aux questions de l'examen, incluant l'analyse des données, de manière individuelle. Toute collaboration de n'importe quelle sorte est strictement interdite. Soumettez votre rapport d'analyse des données au format pdf sur StudiUM.
- Le projet final :
 - Chaque étudiant(e) devra présenter un projet sur un des sujets associés à l'analyse de données catégorielles (qui peut provenir d'un sujet vu ou à voir dans le cours, ou d'un article). Chaque projet consistera à
 - travailler avec une application aux données réelles ;
 - faire un résumé écrit du projet (dans un maximum de six à huit pages).
 - faire une présentation orale devant la classe pour exposer le projet.

- Remarques :
- avant de choisir un projet, il faudra présenter une proposition écrite (d'environ une page) du sujet et l'envoyer au professeur. Les dates des présentations seront fixées pour les dernières deux semaines du trimestre.
- Des copies numériques du rapport et de la présentation devront être soumis au professeur.
- Ce projet implique une étude très détaillée du sujet choisi, une compréhension élevée du sujet, et une étude qui va au delà d'une présentation et d'une application aux données : il faut que vous fassiez votre propre recherche dans le sujet choisi, c'est-à-dire, regarder d'autres alternatives, et proposer et tester vos propres idées.
- Si nécessaire, je vous donnerai de travail extra requis pour finir le projet après la présentation et le rapport initial. Un rapport final sera aussi requis avant la fin du trimestre.
- Présentation des devoirs :
- Les devoirs ne sont pas facultatives. Si vous manquez la date-limite pour soumettre le devoir, votre devoir recevra zéro (0) points.
- Les conditions suivantes simplifieront considérablement l'évaluation des devoirs et sont obligatoires.
- Nommez votre devoir `_rapport.pdf`, et votre script R `_script.txt` où est remplacé par votre nom.
- Les sorties d'un logiciel sans aucune annotation ne sont pas acceptable. Vous devez clarifier quels aspects des sorties d'ordinateur sont appropriées et vous devez montrer comment ils répondent aux questions posées dans le devoir. Des parties non pertinentes ou incorrectes des sorties d'ordinateur devraient être supprimé ou bien clairement biffées.
- Un rapport composé uniquement des sorties de logiciel ne sera pas noté ; c'est-à-dire que ce type de rapport recevra une note de zéro (0).
- Soumettez les problèmes dans l'ordre donné.
- Organisez chaque devoir de sorte que les graphiques et leur discussion soient ensemble. **NE METTEZ PAS tous le graphiques à la fin du devoir.** Marquez quels graphiques sont assortis à quels problèmes.
- Ces conditions aident à s'assurer que votre devoir soit évalué efficacement et dans le meilleur délai. Les règles qui ne sont pas suivies peuvent vous faire perdre des points.
- Remarques importants :
- Le plagiat : attention, c'est sérieux ! Vous êtes invité à consulter le site <https://www.integrite.umontreal.ca>
- ATTENTION : Les copies de solutions de devoirs réalisées par une autre personne, y compris l'enseignant, ou par l'IA sont considérées comme du plagiat : si l'une de vos solutions est une copie d'une solution réalisée par une autre personne (y compris l'enseignant) ou par l'IA, vous recevrez zéro point pour le problème.
- Chaque étudiant(e) est requis(e) de préparer et soumettre ses propres solutions (travail d'ordinateur y compris) aux devoirs. Des solutions préparées "en comité" ne sont pas acceptables. **La duplication des solutions des devoirs et des sorties d'ordinateur préparé entièrement ou partiellement par quelqu'un d'autre ne sont pas acceptables et sont considérées plagiat.** Si vous recevez l'aide de n'importe qui, vous devez dûment lui (leur) rendre reconnaissance dans votre rapport (exemple : "puisque les données sont toutes positives et leur distribution est asymétrique, une transformation logarithmique est clairement appropriée dans la prochaine étape. Je remercie David Cox de m'indiquer ceci.").
- **La collaboration de n'importe quelle sorte sur des examens ou des projets est interdite.**
- Nous faisons bon accueil à des commentaires ou à des suggestions au sujet du cours à tout

moment, soit en personne, par courriel, ou par visioconférence.

- Ce programme est prévu pour fournir une vue d'ensemble de STT6516. Vous ne pouvez revendiquer aucun droit de lui. En particulier, les dates d'examen peuvent changer. Tandis que le programme devrait être un guide assez fiable pour la session présente, les annonces officielles sont toujours ceux qu'on fait dans la classe.

Dates importantes, soutien à la réussite et autres ressources

Dates importantes

- Début session : 2026-08-31
- Congé universitaire (férié) : 2026-09-07
- Modification de l'inscription : 2026-09-16
- Abandon sans frais : 2026-09-16
- Congé universitaire (férié) : 2026-09-30
- Journée des élections provinciales : 2026-10-05
- Abandon avec frais : 2026-10-06
- Congé universitaire (férié) : 2026-10-12
- Début de la période d'activités libres : 2026-10-19
- Dernier jour de cours : 2026-12-08
- Fin du trimestre : 2026-12-23

Soutien à la réussite étudiante (DMS)

RESSOURCES PÉDAGOGIQUES

- [Soutien à l'apprentissage](#)
- [Centre d'aide \(Tutorat gratuit\)](#)

ACCOMPAGNEMENT, CONSEIL ET SOUTIEN ADMINISTRATIF

- [Rendez-vous avec la conseillère à la réussite du DMS](#)
- [Technicienne à la gestion des dossiers étudiants 1^{er} cycle \(TGDE\) du DMS](#)
- [Technicienne à la gestion des dossiers étudiants 2^e et 3^e cycle \(TGDE\) du DMS](#)
- [Responsable des programmes de premier cycle du DMS](#)
- [Responsable des programmes des cycles supérieurs en mathématiques](#)
- [Responsable des programmes des cycles supérieurs en statistique](#)

ASSOCIATIONS ÉTUDIANTES EN MATHÉMATIQUES ET STATISTIQUE

- [1er cycle](#)
- [Cycles supérieurs](#)
- [Club d'actuariat de l'Université de Montréal](#)

Soutien aux études et à la réussite (UdeM)

- Soutien à la vie étudiante
- Centre de communication écrite
- Services des bibliothèques UdeM

Politiques institutionnelles

Règlement des études

Que vous soyez une personne étudiante régulière, libre ou visiteuse, connaître le règlement qui encadre les études est tout à votre avantage. Consultez le [Règlement des études \(1er cycle\)](#) ou des [Cycles supérieurs](#).

Intégrité, fraude et plagiat

Le plagiat ne se limite pas à faire passer un texte d'autrui pour sien. Soumettre le même travail dans deux cours, utiliser des sources non autorisées, ou obtenir de l'aide illicite sont des manquements à l'intégrité pouvant entraîner l'échec, la suspension ou le renvoi. Détails sur le site [Intégrité UdeM](#).

Personnes étudiantes en situation de handicap

Renseignez-vous sur les ressources adaptées à votre situation auprès du [Bureau de soutien \(BSESH\)](#). Consultez la [Politique-cadre sur l'intégration](#) pour les accommodements.

Intelligence Artificielle

L'utilisation de l'intelligence artificielle générative n'est pas autorisée pour la production des travaux, examens, ou rapports, et pourrait être reconnue comme une fraude.