

Université de
Sherbrooke

Département de mathématiques
STT 418 – Statistique appliquée

Plan d'activité pédagogique

Hiver 2022

Enseignant

Jean-Philippe Morissette

Courriel : jean-philippe.morissette@usherbrooke.ca

Local : D3-1031-1

Téléphone : +1 819 821-8000 x65569

Disponibilités : Une période de disponibilité sera fixée au cours de la première semaine du cours.

Responsable(s) : Direction du département

Site web du cours : <https://moodle.usherbrooke.ca>

Horaire

Exposé magistral :	Lundi	13h30 à 15h20	salle D3-2041
	Jeudi	10h30 à 12h20	salle D3-2035
	Vendredi	10h30 à 11h20	salle D3-2040
	Exercices/laboratoires :	Lundi	13h30 à 15h20
			salle À venir

Description officielle de l'activité pédagogique¹

Cibles de formation :	Acquérir les notions de probabilités et de statistique indispensables à l'analyse des données.
Contenu :	Éléments de statistique descriptive. Notions fondamentales de probabilités. Notions d'échantillonnage. Estimation ponctuelle. Généralités sur les tests d'hypothèses. Tests usuels. Ajustement de données par des lois. Modèles de régression et tests associés. Étude de cas tirés des milieux des affaires et de l'économie.
Crédits	3
Organisation	3 heures d'exposé magistral par semaine 2 heures d'exercices par semaine 4 heures de travail personnel par semaine
Particularités	Aucune

¹<https://www.usherbrooke.ca/admission/fiches-cours/stt418>

1 Présentation

Cette section présente les cibles de formation spécifiques et le contenu détaillé de l'activité pédagogique. Cette section, non modifiable sans l'approbation du comité de programme du Département d'informatique, constitue la version officielle.

1.1 Mise en contexte

Le cours STT418 - Statistique appliquée est un cours obligatoire des programmes de baccalauréats en informatique, de baccalauréat en informatique de gestion et de baccalauréat en sciences du multimédia et du jeu vidéo. Il vise à initier l'étudiante et l'étudiant à la modélisation probabiliste et au traitement statistique des données. Nous proposons de construire une théorie de la statistique et des probabilités qui associe la rigueur mathématique et l'intuition populaire.

La plupart des gens possèdent une certaine conception des notions de probabilité et statistique. On pourrait même dire qu'elles sont indispensables dans la vie courante compte tenu du grand nombre de messages transmis à la population par le biais des statistiques : prévisions météorologiques, sondages, pronostics médicaux, etc. Dans ce cours, nous allons définir et étudier ces notions avec plus de rigueur et de précision que l'idée intuitive qui se développe dans notre esprit, par l'interprétation des sondages populaires, des prévisions météorologiques, etc.

Nous verrons comment utiliser les probabilités pour modéliser des données provenant de sondages ou autres enquêtes sur une population, pour inférer des résultats sur cette population et même dans certains cas pour émettre une prévision (météo, élections, niveau d'eau dans un bassin hydrographique, coût de production, profit, vente, . . .) ou prendre des décisions concernant l'avenir de la population.

Plusieurs exemples et exercices seront faits au cours, en mettant l'accent (autant que possible) sur des applications de la vie courante.

1.2 Cibles de formation spécifiques

À la fin de cette activité pédagogique, l'étudiante ou l'étudiant devrait connaître, comprendre et être capable de manipuler les concepts suivants :

1. Statistique descriptive ;
2. Théorie des probabilités ;
3. Loi de probabilité (une seule variable) ;
4. Loi de probabilité (variables simultanées) ;
5. Échantillonnage et estimation ;
6. Tests d'hypothèses.

1.3 Contenu détaillé

Thème	Contenu	Nbr. d'heures	Objectifs
1	Statistique descriptive : Définitions, types de variables et échelles de mesure, mesures de tendance centrale (moyenne, médiane, mode), mesures de dispersion (étendue, variance, écart-type), mesures de position (quantiles), mesures de forme (coefficient d'aplatissement, coefficient d'asymétrie), minimum et maximum, covariance, causalité et corrélation. Exemples d'application : Algorithmes de tri et de sélection. Filtre médian. Utilisation du coefficient d'aplatissement pour la détection de photomontage. Régression linéaire entre deux variables (par exemple le nombre de lignes de code d'un programme et le temps de développement).	6	1
2	Théorie des probabilités : Définitions, axiomes et propriétés, probabilité conditionnelle, indépendance d'événements, probabilités totales, règle de Bayes. Exemples d'application : Problèmes de dénombrements (choix des mots de passes). Détection de pourriels.	8	2
3	Loi de probabilité (une seule variable) : Variable aléatoire, fonction de répartition, espérance, variance, covariance, distributions discrètes (binomiale et Poisson) et continues (normale). Exemples d'application : Bruit et loi de poisson	9	3
4	Loi de probabilité (variables simultanées) : Probabilité marginale, probabilité conditionnelle, indépendance, espérance et variance conditionnelle, vecteurs aléatoires, matrice de covariance, loi normale multidimensionnelle. Modélisation de file d'attente. Exemples d'application : Mesure de similitude entre deux images. Application de la loi normale multidimensionnelle à la prédiction linéaire (recommandation de films ou produits). Représentation de la densité de probabilités des pixels dans une image via un mélange de gaussiennes.	5	4
5	Échantillonnage et estimation : Distribution d'échantillonnage, théorème de la limite centrale, estimation ponctuelle, estimation par intervalle de confiance. Exemples d'application : Contrôle de qualité de la durée de vie d'un ordinateur.	5	5
6	Tests d'hypothèses : Une et deux moyennes, proportion et variance. Exemples d'application : Comparaison du temps d'exécution de deux algorithmes. Comparaison des prédictions (malade vs sain) obtenues par une méthode d'analyse d'images médicales par opposition aux prédictions obtenues par un expert. Segmentation en deux classes d'une image.	6	6

1. Le cours doit comprendre au moins deux travaux pratiques couvrant tous les sujets marqués «✓» dans le tableau.

2 Organisation

Cette section propre à l'approche pédagogique de chaque enseignante ou enseignant présente la méthode pédagogique, le calendrier, le barème et la procédure d'évaluation ainsi que l'échéancier des travaux. Cette section doit être cohérente avec le contenu de la section précédente.

2.1 Méthode pédagogique

L'approche pédagogique privilégiée sera les cours magistraux, et ce à raison de trois heures par semaine. Évidemment, des exercices seront disponibles afin de pouvoir concrétiser les différents concepts vus en classe. Le but est de permettre aux étudiants et aux étudiantes de bien acquérir, mais surtout de bien comprendre, les méthodes et notions présentées en classe. Certains seront résolus en détail lors des périodes d'exercices, qui ont lieu à raison de deux heures par semaine.

La participation en classe est fortement encouragée. Elle permet aux étudiants et étudiantes de consolider leurs acquis. Les étudiants et les étudiantes seront d'ailleurs incités à poser des questions et à émettre des commentaires, constructifs, par rapport à l'enseignement. Évidemment, la présence aux cours est un facteur clé de la réussite.

L'entre-aide entre confrère et consœur est grandement encouragée. Par contre, puisque les évaluations en classe seront faites individuellement, une bonne partie du travail demandé devrait être faite de manière individuelle afin de bien préparer les étudiants et les étudiantes.

Des documents seront disponibles sur le [site Moodle](#) du cours et les résultats aux évaluations seront disponibles à partir de l'interface [Genote](#). Les messages en lien avec le cours seront transmis par courriel à la liste des étudiantes et des étudiants inscrits à STT418 – Statistique appliquée. Assurez-vous d'[activer](#) votre compte de courrier électronique.

2.2 Calendrier

Semaine	Date	Thème	Devoirs	QCM
1	2022-01-03	1		Remise QCM1
2	2022-01-10	1		Remise QCM2
3	2022-01-17	2		Remise QCM3
4	2022-01-24	2		Remise QCM4
5	2022-01-31	2 et 3	Remise Devoir 1	Remise QCM5
6	2022-02-07	3		Remise QCM6
7	2022-02-14	3	Remise Devoir 2	
8	2022-02-21	Examen périodique		
9	2022-02-28	Relâche		
10	2022-03-07	4		
11	2022-03-14	5		Remise QCM7
12	2022-03-21	5	Remise Devoir 3	Remise QCM8
13	2022-03-28	6		Remise QCM9
14	2022-04-04	6	Remise Devoir 4	Remise QCM10
15	2022-04-11	Révision		
16	2022-04-18	Examen final		

2.3 Évaluation

Devoirs (4)	24 %
QCM (10)	10 %
Examen intra	30 %
Examen final	36 %

Tout le contenu présenté en cours, que ce soit lors des séances magistrales, lors des séances d'exercices ou lors des devoirs, est susceptible d'être évalué lors des examens intra et final. Il est à noter que le raisonnement probabiliste et statistique, ainsi que les compétences acquises dans ce cours sont essentielles en informatique et seront exploitées dans d'autres cours qui suivront.

Après chaque évaluation, il y aura une période de rétroaction permettant aux étudiantes et aux étudiants d'identifier leurs lacunes et de vérifier leur copie.

Le cours STT418 – Statistique appliquée porte sur un contenu qui se construit et s'élabore de plus en plus tout au long de la session. Les notions apprises auparavant seront donc reprises à l'examen final.

Les examens comportent des questions « théoriques » qui visent à vérifier l'acquisition des connaissances et leur compréhension et des questions « pratiques » qui visent à vérifier si vous êtes en mesure d'appliquer ces connaissances dans des cas concrets.

Les critères de correction seront la pertinence et la cohérence de la démarche, la rigueur des raisonnements, la clarté, l'exactitude et la précision des solutions aux problèmes et la justesse des calculs.

De plus, il demeure incontestable que le succès aux évaluations est directement relié aux efforts qui ont été consacrés aux exercices. Le travail investi dans la résolution de ceux-ci est donc très important pour l'atteinte des objectifs du cours et pour la réussite aux évaluations.

Enfin, il n'y a pas de balises de pré-établies pour les cotes finales; ce cours est dit à cotes variables.

2.3.1 Qualité de la langue et de la présentation

Conformément à l'article 17 du règlement facultaire d'évaluation des apprentissages² l'enseignante ou l'enseignant peut retourner à l'étudiante ou à l'étudiant tout travail non conforme aux exigences quant à la qualité de la langue et aux normes de présentation.

2.3.2 Plagiat

Le plagiat consiste à utiliser des résultats obtenus par d'autres personnes afin de les faire passer pour sien et dans le dessein de tromper l'enseignante ou l'enseignant. Vous trouverez en annexe un document d'information relatif à l'intégrité intellectuelle qui fait état de l'article 9.4.1 du Règlement des études³. Lors de la correction de tout travail individuel ou de groupe une attention spéciale sera portée au plagiat. Si une preuve de plagiat est attestée, elle sera traitée en conformité, entre autres, avec l'article 9.4.1 du Règlement des études de l'Université de Sherbrooke. L'étudiante ou l'étudiant peut s'exposer à de graves sanctions qui peuvent être soit l'attribution de la note E ou de la note zéro (0) pour un travail, un examen ou une activité évaluée, soit de reprendre un travail, un examen ou une activité pédagogique. Tout travail suspecté de plagiat sera transmis au Secrétaire de la Faculté des sciences. Ceci n'indique pas que vous n'avez pas le droit de coopérer entre deux équipes, tant que la rédaction finale des documents et la création du programme restent le fait de votre équipe. En cas de doute de plagiat, l'enseignante ou l'enseignant peut demander à l'équipe d'expliquer les notions ou le fonctionnement du code qu'elle ou qu'il considère comme étant plagié. En cas d'incertitude, ne pas hésiter à demander conseil et assistance à l'enseignante ou l'enseignant afin d'éviter toute situation délicate par la suite.

2.4 Échéancier des travaux

Les dates de remise des travaux seront indiquées sur les énoncés.

2.4.1 Directives particulières

L'évaluation se fera au moyen de six épreuves écrites (quatre devoirs et deux examens) et de questions à choix multiple (hebdomadaire). Tout le contenu présenté en classe, que ce soit lors des séances théoriques, des séances d'exercices, lors de travaux ou de devoirs, est susceptible d'être évalué lors des examens intra et final.

²https://www.usherbrooke.ca/sciences/fileadmin/sites/sciences/Etudiants_actuels/Informations_academiques_et_reglements/2017-10-27_Reglement_facultaire_-_evaluation_des_apprentissages.pdf

³<https://www.usherbrooke.ca/registraire/droits-et-responsabilites/reglement-des-etudes/>

Pour toute absence à un examen, tout travail remis en retard ou toute tentative de plagiat, l'étudiante ou l'étudiant recevra une note de zéro pour l'évaluation en question. Sous réserve d'application du règlement facultaire d'évaluation des apprentissages des étudiantes et des étudiants, il n'y aura aucun examen de reprise. En cas d'absence motivée à un des tests, le poids de cette évaluation sera reporté sur l'examen final.

Détails des évaluations

- Les questions à choix multiples (QCM) ont pour but de garder à jour les étudiants et les étudiantes. Ainsi, presque chaque jeudi, le cours commencera avec quelques QCM qui seront évaluées sur 1 point. Vous aurez donc la possibilité d'accumuler jusqu'à 10 points. Afin d'être conciliant, vous n'aurez qu'à obtenir 8 points sur un maximum de 10 points pour avoir la note maximale dans cette catégorie d'évaluation.
- Les devoirs doivent être remis en version papier avant l'heure inscrite sur l'énoncé du devoirs. Les devoirs seront remis en équipe de deux à quatre personnes. Tout travail qui ne respectera pas ces règles sera refusé et l'étudiante ou l'étudiant recevra une note de zéro pour l'évaluation en question.

2.5 Utilisation d'appareils électroniques et du courriel

Selon le règlement complémentaire des études, section 4.2.3⁴, l'utilisation d'ordinateurs, de cellulaires ou de tablettes pendant une prestation est interdite à condition que leur usage soit explicitement permise dans le plan de cours.

Dans ce cours, l'usage de téléphones cellulaires, de tablettes ou d'ordinateurs est autorisées. Cette permission peut être retirée en tout temps si leur usage entraîne des abus.

Tel qu'indiqué dans le règlement universitaire des études, section 4.2.3⁵, toute utilisation d'appareils de captation de la voix ou de l'image exige la permission de la personne enseignante.

Note : L'utilisation du courriel est recommandée pour poser vos questions.

3 Matériel nécessaire pour l'activité pédagogique

- Des notes de cours et d'autres documents seront disponibles sur le [site Moodle](#) du cours.
- À l'occasion, le recours à l'environnement RStudio et au langage R pourrait être nécessaire, notamment lors des devoirs, de manière à travailler sur des cas réels non triviaux.

4 Références

- [1] The R Project for Statistical Computing. <https://www.r-project.org/>.
- [2] ALLEN, ARNOLD O : *Probability, Statistics and Queuing Theory*, Academic Press édition, 1990.
- [3] ANDERSON, DAVID R. ET DENNIS J. SWEENEY : *Statistiques pour l'économie et la gestion*, De Boeck supérieur édition, 2013.
- [4] GOODMAN, ROE : *Introduction to Stochastic Models*, Dover Publications édition, 2006.
- [5] MATLOFF, NORM : *From Algorithms to Z-Scores : Probabilistic and Statistical Modeling in Computer Science*. University of California.
- [6] MITZENMACHER, MICHAEL ET ELI, UPFAL : *Probability and Computing : Randomized Algorithms and Probabilistic Analysis*, Cambridge University Press édition, 2005.
- [7] ROSS, SHELDON M : *A First Course in Probability*, Presses polytechniques et universitaires romandes édition, 2009.

⁴https://www.usherbrooke.ca/sciences/fileadmin/sites/sciences/documents/Intranet/Informations_academiques/Sciences_Reglement_complementaire_2017-05-09.pdf

⁵<https://www.usherbrooke.ca/registraire/droits-et-responsabilites/reglement-des-etudes/>

L'intégrité intellectuelle passe, notamment, par la reconnaissance des sources utilisées. À l'Université de Sherbrooke, on y veille!

Extrait du Règlement des études (Règlement 2575-009)

9.4.1 DÉLITS RELATIFS AUX ÉTUDES

Un délit relatif aux études désigne tout acte trompeur ou toute tentative de commettre un tel acte, quant au rendement scolaire ou une exigence relative à une activité pédagogique, à un programme ou à un parcours libre.

Sont notamment considérés comme un délit relatif aux études les faits suivants :

- a) commettre un plagiat, soit faire passer ou tenter de faire passer pour sien, dans une production évaluée, le travail d'une autre personne ou des passages ou des idées tirés de l'œuvre d'autrui (ce qui inclut notamment le fait de ne pas indiquer la source d'une production, d'un passage ou d'une idée tirée de l'œuvre d'autrui);
 - b) commettre un autoplagiat, soit soumettre, sans autorisation préalable, une même production, en tout ou en partie, à plus d'une activité pédagogique ou dans une même activité pédagogique (notamment en cas de reprise);
 - c) usurper l'identité d'une autre personne ou procéder à une substitution de personne lors d'une production évaluée ou de toute autre prestation obligatoire;
 - d) fournir ou obtenir toute aide non autorisée, qu'elle soit collective ou individuelle, pour une production faisant l'objet d'une évaluation;
 - e) obtenir par vol ou toute autre manœuvre frauduleuse, posséder ou utiliser du matériel de toute forme (incluant le numérique) non autorisé avant ou pendant une production faisant l'objet d'une évaluation;
 - f) copier, contrefaire ou falsifier un document pour l'évaluation d'une activité pédagogique;
- [...]

Par plagiat, on entend notamment :

- Copier intégralement une phrase ou un passage d'un livre, d'un article de journal ou de revue, d'une page Web ou de tout autre document en omettant d'en mentionner la source ou de le mettre entre guillemets;
- reproduire des présentations, des dessins, des photographies, des graphiques, des données... sans en préciser la provenance et, dans certains cas, sans en avoir obtenu la permission de reproduire;
- utiliser, en tout ou en partie, du matériel sonore, graphique ou visuel, des pages Internet, du code de programme informatique ou des éléments de logiciel, des données ou résultats d'expérimentation ou toute autre information en provenance d'autrui en le faisant passer pour sien ou sans en citer les sources;
- résumer ou paraphraser l'idée d'un auteur sans en indiquer la source;
- traduire en partie ou en totalité un texte en omettant d'en mentionner la source ou de le mettre entre guillemets ;
- utiliser le travail d'un autre et le présenter comme sien (et ce, même si cette personne a donné son accord);
- acheter un travail sur le Web ou ailleurs et le faire passer pour sien;
- utiliser sans autorisation le même travail pour deux activités différentes (autoplagiat).

Autrement dit : mentionnez vos sources
