

## Programme court de premier cycle avancé en statistique et science des données

Code	Titre	Crédits
9070	Programme court de premier cycle avancé en statistique et science des données	12

<b>Trimestre(s) d'admission</b>	Automne Hiver
<b>Contingent</b>	Programme non contingenté
<b>Régime et durée des études</b>	Offert à temps complet et à temps partiel
<b>Campus</b>	Campus de Montréal

**OBJECTIFS**

- 1) Permettre aux étudiantes, étudiants ayant une formation universitaire de niveau intermédiaire en statistique ou dans une discipline connexe de compléter celle-ci à travers des cours pertinents en statistique et science des données. Les cours offerts sont de niveau intermédiaire et avancé, permettant ainsi un apprentissage spécialisé et sur mesure à travers les cours choisis.
- 2) Contribuer à la formation académique des étudiantes, étudiants, notamment sur le plan de l'acquisition des connaissances en statistique et science des données et des capacités d'analyse et de synthèse.

**CONDITIONS D'ADMISSION****Trimestre d'admission (information complémentaire)**

Admission aux trimestres d'automne et d'hiver.

**Base DEC**

Il n'y a pas d'admission sur cette base.

**Base expérience**

Il n'y a pas d'admission sur cette base.

**Base études universitaires**

Au moment du dépôt de la demande d'admission, avoir réussi au moins quinze cours (quarante-cinq crédits ou l'équivalent) de niveau universitaire dans un programme d'études de premier cycle en statistique ou dans un domaine connexe.

Préalable requis (ou l'équivalent) : MAT1115 Calcul I, MAT1250 Algèbre linéaire I, MAT1700 Probabilités I, STT1000 Statistique I et INF1035 Informatique pour les sciences : programmation simulation et exploitation de données.

OU

Avoir réussi le programme court de premier cycle en statistique et sciences des données.

**Base études hors Québec**

Il n'y a pas d'admission sur cette base.

**Régime et durée des études**

Offert à temps complet et à temps partiel

**COURS À SUIVRE**

(Sauf indication contraire, les cours comportent 3 crédits.)

**Entre zéro (0) et deux (2) cours parmi la liste des cours suivants :**

MAT2720 Processus stochastiques  
STT2000 Statistique II  
STT2010 Échantillonnage  
STT2100 Laboratoire de statistique  
STT2110 Plans d'expérience et ANOVA  
STT2120 Régression  
STT5100 Modèles linéaires appliqués

**Entre deux (2) et quatre (4) cours parmi la liste des cours suivants :**

ACT6100 Analyse de données en actuariat  
STT3000 Statistique III  
STT3010 Statistique informatique  
STT3020 Sujets spéciaux de statistique  
STT3030 Apprentissage statistique  
STT3100 Analyse multivariée appliquée  
STT3120 Biostatistique  
STT3200 Synthèse  
STT3300 Compléments de statistique I  
STT3320 Compléments de statistique II  
STT3340 Compléments de statistique III

Ou tout autre cours en adéquation avec les objectifs du programme choisi avec l'accord de la direction du programme.

**CHAMPS DE RECHERCHE**

Statistique et science des données

**REMARQUES PARTICULIÈRES**

Un accompagnement personnalisé par la direction de programme est prévu pour le choix des cours. Cet accompagnement permettra d'identifier les cours les plus pertinents en fonction des objectifs académiques et professionnels de l'étudiante, de l'étudiant et de s'assurer que celui-ci possède les préalables requis pour les cours choisis.

**DESCRIPTION DES COURS**

**ACT6100 Analyse de données en actuariat**

## Objectifs

Comprendre les fondements des méthodes d'apprentissage statistique supervisées et non-supervisées sous un angle actuariel. Applications de différentes méthodes pour l'analyse de données actuarielles à l'aide d'un langage de programmation comme R ou SAS. Applications actuarielles typiques en assurance-vie, assurance IARD, assurance collective, régimes de retraite ainsi qu'en finance.

## Sommaire du contenu

Exploration de données, validation et vérifications; Méthodes d'apprentissage statistique supervisées et non-supervisées; Analyse aux composantes principales; Arbres de décision et de classification; Groupement et partitionnement de données (cluster analysis); Analyse de données temporelles; Ce cours comporte une séance de travaux pratiques (TP) de deux heures par semaine. Le cours est une composante de l'examen Statistics for Risk Modeling de la Society of Actuaries et de l'examen Modern Actuarial Statistics II de la Casualty Actuarial Society.

## Préalables académiques

STT5100 Modèles linéaires appliqués ou STT2120 Régression

**MAT2720 Processus stochastiques**

## Objectifs

Familiariser l'étudiant avec les principaux modèles mathématiques pertinents à l'étude des processus stochastiques.

## Sommaire du contenu

Processus stochastiques. Chaînes de Markov: matrice de transition, équations de Chapman-Kolmogorov, classification des états, analyse des premiers pas, probabilités limites, chaînes de Markov réductibles, promenades aléatoires et autres problèmes. Processus de branchement : distribution du nombre de descendants et probabilité d'extinction. Processus de Poisson : loi exponentielle, processus de comptage, temps d'attente, autres propriétés et généralisation du processus de Poisson (non-homogène, composé). Chaînes de Markov à temps continu : probabilité de transition, générateur infinitésimal, équations de Chapman-Kolmogorov, probabilités limites et système M/M/n. Introduction au mouvement brownien : temps d'atteinte et variables aléatoires maximales, pont brownien. Ce cours comporte une séance de travaux pratiques (TP) de deux heures par semaine. Quelques séances de TP pourront être évaluées.

## Préalables académiques

MAT1700 Probabilités I

**STT2000 Statistique II**

## Objectifs

Ce cours développe les bases mathématiques de l'inférence statistique et introduit des concepts nouveaux portant sur des propriétés d'optimalité en estimation et tests d'hypothèses.

## Sommaire du contenu

Espérance conditionnelle, loi normale bivariable, statistiques exhaustives. Estimation ponctuelle et par intervalle. Approche bayésienne et fréquentiste. Méthodes d'estimation et propriétés des estimateurs, théorème de Rao-Blackwell, inégalité de Cramér-Rao. Tests d'hypothèses et intervalle de confiance. Lemme de Neyman-Pearson, tests uniformément les plus puissants, test du rapport de vraisemblance. Méthodes non paramétriques. Applications diverses. Ce cours comporte une séance de travaux pratiques (TP) de deux heures par semaine.

## Préalables académiques

STT1000 Statistique I

**STT2010 Échantillonnage**

## Objectifs

Développer les aptitudes nécessaires à l'élaboration des plans de sondage et à l'analyse des résultats.

## Sommaire du contenu

Échantillon aléatoire simple: estimation d'une moyenne et d'un total, variance des estimateurs, estimation de la variance, intervalle de confiance, détermination de la taille d'un échantillon; estimation d'une proportion; estimation d'un quotient. Estimation d'une moyenne par le quotient et par régression. Échantillonnage par stratification: estimation d'une moyenne, d'un total, d'une proportion, d'un quotient. Quelques méthodes d'échantillonnage par grappes.

## Préalables académiques

STT1000 Statistique I

**STT2100 Laboratoire de statistique**

## Objectifs

Introduire l'analyse des données par le biais de méthodes simples mais éprouvées, en cultivant le raisonnement statistique et en favorisant une approche globale de l'analyse.

## Sommaire du contenu

Tendance centrale, dispersion, concentration, représentation graphique, moustache, ajustement, t-test, ANOVA à un facteur, série chronologique, corrélation, régression linéaire, données catégorielles, indépendance. Logiciel SAS : données, tableau, saisie et importation, sélection, recodage, étiquetage, étape DATA, étape PROC, graphisme. Logiciel R : types, sélection, entrée/sortie, graphisme, programmation de base. Ce cours comporte une séance de travaux pratiques (TP) de deux heures par semaine.

## Préalables académiques

STT1000 Statistique I, INF1035 Informatique pour les sciences : programmation simulation et exploitation de données ou INF1120 Programmation I

**STT2110 Plans d'expérience et ANOVA**

## Objectifs

À la fin du cours, l'étudiant devrait être en mesure de choisir un plan d'expérience approprié, de faire son analyse et de présenter ses résultats clairement.

## Sommaire du contenu

Ce cours présente les principaux modèles et techniques employés dans la planification et dans l'analyse des expériences. Les résultats des expériences sont analysés avec le logiciel SAS. Expériences avec un facteur sans contraintes sur la randomisation : suppositions, diagnostiques, transformations, tests sur des moyennes, contrastes ; expériences avec des contraintes sur la randomisation: plans en blocs aléatoires et carré latin; expériences factorielles à effets fixes et aléatoires, calcul des espérances des carrés moyens; plans d'expériences emboîtés, à parcelles partagées, croisés et à mesures répétées; expériences avec 2f et avec 3f facteurs; expériences avec des effets confondus. Ce cours comporte une séance de travaux pratiques (TP) de deux heures par semaine.

## Préalables académiques

STT2000 Statistique II, STT2100 Laboratoire de statistique

**STT2120 Régression**

## Objectifs

Ce cours présente de façon concrète certains des modèles utilisés dans l'analyse de la dépendance entre variables : la régression linéaire simple et multiple, avec variables exogènes quantitatives, qualitatives, ou mixtes. L'objectif du cours est de développer l'aptitude à utiliser ces techniques correctement. Introduit en un premier temps comme techniques indépendantes de manière à développer l'intuition et sensibiliser l'étudiant(e) aux difficultés d'interprétation qui surgissent en pratique, ces modèles seront ensuite rassemblés et traités comme cas particuliers du modèle linéaire général. Si le cours accorde une importance particulière aux applications, la base théorique ne sera pas pour autant négligée.

## Sommaire du contenu

Rappel de certains préalables mathématiques et statistiques : matrices,

loi normale multidimensionnelle, distribution de formes quadratiques. Estimation et inférence dans des modèles linéaires simples : estimateur des moindres carrés, distribution des estimateurs, intervalles de confiance, hypothèse linéaire générale, comparaisons multiples. Régression linéaire simple : estimation, tests d'hypothèses et techniques diagnostiques. Le modèle linéaire général : estimateurs des paramètres et leur distribution. Régression multiple et régression polynomiale. Sélection de modèles. Applications aux modèles à variables exogènes qualitatives et mixtes: ANOVA et ANCOVA. Ce cours comporte une séance de travaux pratiques (TP) de deux heures par semaine.

Préalables académiques

STT2000 Statistique II , STT2100 Laboratoire de statistique , MAT1250 Algèbre linéaire I

### **STT3000 Statistique III**

Objectifs

Introduire l'étudiant à la statistique mathématique.

Sommaire du contenu

Le cours se concentre essentiellement sur l'estimation ponctuelle, autant dans le cadre classique que bayésien. Notions couvertes: Familles exponentielles, estimation sans biais à variance minimale et estimation par la méthode du maximum de vraisemblance. Information de Fisher et entropie. Fonctions de perte et risque fréquentiste. L'approche bayésienne, lois a priori conjuguées et lois non-informatives. Risque moyen optimal et règles de Bayes. Admissibilité. L'approche minimax et la loi a priori la moins favorable. Calcul bayésien (si le temps le permet).

Conditions d'accès

Pour les étudiants du baccalauréat en mathématiques : avoir complété 60 crédits du programme

Préalables académiques

STT2000 Statistique II

### **STT3010 Statistique informatique**

Objectifs

Familiariser l'étudiant avec certaines techniques statistiques nécessitant l'informatique et utiliser l'ordinateur et les simulations pour favoriser la compréhension des concepts et méthodes statistiques.

Sommaire du contenu

Nombres pseudo-aléatoires, génération de variables aléatoires discrètes et continues, méthodes Monte-Carlo, techniques de réduction de variance, bootstrap, optimisation, algorithmes d'apprentissage, sujets et applications choisis. Laboratoires informatiques et utilisation d'un langage de programmation comme R.

Préalables académiques

STT2000 Statistique II, STT2120 Régression

### **STT3020 Sujets spéciaux de statistique**

Objectifs

Pour la formation d'un statisticien, ce cours présente des sujets importants qui ne sont pas traités dans d'autres cours.

Sommaire du contenu

Dans une session donnée, le cours peut porter sur deux thèmes ; les sujets seront choisis parmi: (a) Statistique non paramétrique, (b) Séries chronologiques, (c) Méthodes asymptotiques, (d) Fiabilité, (e) Lissage, (f) Analyse de données catégorielles, (g) Statistique séquentielle, ou tout autre sujet d'importance choisi par le professeur.

Préalables académiques

STT2000 Statistique II

### **STT3030 Apprentissage statistique**

Objectifs

Familiariser l'étudiant avec les méthodes d'apprentissage statistique et

le rendre apte à appliquer ces méthodes à de vrais jeux de données.

Sommaire du contenu

Méthodes linéaires pour la régression et la classification. Choix et évaluation de modèles. Méthodes de rétrécissement. Méthodes de lissages par noyau. Arbres de décision et méthodes d'agrégation. Machines à vecteurs de support. Techniques d'apprentissage non supervisé. Réseaux neuronaux.

Préalables académiques

STT2120 Régression

### **STT3100 Analyse multivariée appliquée**

Objectifs

Présenter diverses méthodes statistiques applicables à des données multidimensionnelles, et illustrer les techniques abordées à l'aide du logiciel statistique SAS.

Sommaire du contenu

Rappels sur les notions essentielles d'algèbre linéaire. La loi normale multivariée, la loi de Wishart et la loi de Hotelling. Tests multivariés à un et à deux échantillons. Tests multivariés pour des observations appariées. Régions de confiance et intervalles de confiance simultanés. Analyse en composantes principales. Analyse factorielle. Analyse de corrélation canonique. Analyse discriminante et classification.

Préalables académiques

STT2110 Plans d'expérience et ANOVA, STT2120 Régression, MAT1250 Algèbre linéaire I

### **STT3120 Biostatistique**

Objectifs

Initier l'étudiant aux concepts et méthodes de base en biostatistique.

Sommaire du contenu

Ce cours traite de sujets dans les domaines de l'épidémiologie, l'analyse de survie et l'inférence causale. L'étudiant apprendra des techniques pour analyser des données provenant d'études observationnelles et d'études planifiées. Certains sujets spécifiques incluent : prévalence et incidence, études prospectives et rétrospectives, estimation et inférence pour les mesures d'association, confusion et interaction, variables d'exposition à plusieurs niveaux, modèles de régression, tests d'ajustement, études paires, estimateur de Kaplan-Meier et modèle de Cox. Des sujets liés à d'autres domaines de la biostatistique pourraient être examinés à la fin du cours.

Préalables académiques

STT2100 Laboratoire de statistique

### **STT3200 Synthèse**

Objectifs

Ce séminaire de synthèse mène l'étudiant à une compréhension plus approfondie et plus large de la théorie statistique et ses applications en se basant d'abord sur des concepts maîtrisés dans des cours antérieurs.

Sommaire du contenu

Les thèmes spécifiques dépendent de l'enseignant, mais comprennent généralement des méthodes statistiques avancées utilisées pour traiter des problèmes récents de recherche appliquée. L'objectif du cours est de mener l'étudiant à faire une synthèse des techniques statistiques apprises dans des cours précédents, à explorer de nouvelles techniques et à apprendre à communiquer ses résultats statistiques.

Préalables académiques

STT2110 Plans d'expérience et ANOVA, STT2120 Régression

### **STT3300 Compléments de statistique I**

Objectifs

Compléter la formation de premier cycle en statistique.

Sommaire du contenu

On y traite de sujets ne faisant pas l'objet d'un cours particulier, ou de sujets propres à faire une synthèse des cours précédents, ou encore de sujets mixtes se rapportant à plusieurs disciplines particulières.

Conditions d'accès

Pour les étudiants du baccalauréat en mathématiques : avoir complété 60 crédits du programme

Préalables académiques

STT2000 Statistique II

### **STT3320 Compléments de statistique II**

Objectifs

Compléter la formation de premier cycle en statistique.

Sommaire du contenu

On y traite de sujets ne faisant pas l'objet d'un cours particulier, ou de sujets propres à faire une synthèse des cours précédents, ou encore de sujets mixtes se rapportant à plusieurs disciplines particulières.

Préalables académiques

STT2000 Statistique II

### **STT3340 Compléments de statistique III**

Objectifs

Compléter la formation de premier cycle en statistique.

Sommaire du contenu

On y traite de sujets ne faisant pas l'objet d'un cours particulier, ou de sujets propres à faire une synthèse des cours précédents, ou encore de sujets mixtes se rapportant à plusieurs disciplines particulières.

Conditions d'accès

Avoir complété 60 crédits du programme

Préalables académiques

STT2000 Statistique II

### **STT5100 Modèles linéaires appliqués**

Objectifs

Ce cours introduit l'étudiant aux modèles de régression linéaire simple et multiple, et introduit aux modèles linéaires généralisés.

Sommaire du contenu

Méthode d'estimation par moindres carrés, Régression linéaire simple et multiple, Tests d'hypothèses et intervalles de confiance, tests d'ajustement, introduction aux modèles linéaires généralisés (GLM): régression logistique, régression Poisson, régression Gamma. Applications numériques à l'aide de SAS ou R. Ce cours comporte une séance de travaux pratiques (TP) de deux heures par semaine. Le cours est une composante du VEE - Applied Statistics de la Society of Actuaries.

Préalables académiques

Soit le cours ACT4400 Modèles de survie ou bien les trois cours suivants : STT2000 Statistique II, STT2100 Laboratoire de statistique et MAT1250 Algèbre linéaire I.

N.B. : Le masculin désigne à la fois les hommes et les femmes sans aucune discrimination et dans le seul but d'alléger le texte.  
Cet imprimé est publié par le Registrariat. Basé sur les renseignements disponibles le 14/07/22, son contenu est sujet à changement sans préavis.  
Version Automne 2022