

Grade: Philosophiae doctor (Ph.D.)**Crédits: 90****Note**

Ce programme est offert en extension à l'UQTR en vertu d'un protocole entre l'Université Laval et l'Université du Québec à Trois-Rivières.

Présentation**En bref**

Ce programme fera de vous un chercheur de haut niveau dans les sciences statistiques pures ou appliquées. Vous serez préparé à travailler dans des organismes de recherche gouvernementaux, publics et privés ou encore à entreprendre une carrière universitaire.

Vous aiguiserez votre esprit critique et vous cultiverez l'originalité de votre pensée. Vous exercerez votre créativité de façon à contribuer à l'avancement de la science. Voici quelques domaines et expertises qui vous seront accessibles: probabilités et processus stochastiques, analyse de données et méthodologie d'enquête, statistique mathématiques.

Votre environnement de recherche sera à la hauteur de vos ambitions: nombreux colloques et séminaires, professeurs aux expertises diverses offrant un encadrement de qualité, service de consultation statistique, possibilités d'emploi étudiant au Centre de dépannage et d'assistance (CDA) en mathématiques et au Service de consultation statistique (SCS) et accès à un laboratoire informatique très bien équipé.

Objectifs du programme

Ce programme vise à conduire l'étudiant à la fine pointe de la recherche dans une spécialité donnée et à en faire un chercheur autonome. Chemin faisant, celui-ci aiguisera son esprit critique, cultivera son originalité de pensée et exercera ses talents de créativité, afin de contribuer par son travail à l'avancement de la science.

Concentrations, profils, cheminements

Le programme de troisième cycle en statistique est destiné à conduire l'étudiant à la pointe d'une certaine spécialité en statistique et à en faire un chercheur autonome. L'essentiel des études dans ce programme est donc la rédaction d'une thèse qui doit comporter des résultats nouveaux. Il faut auparavant avoir réussi trois cours et deux examens de synthèse.

Avenir : Carrière et débouchés

Ce programme fera de vous un chercheur de pointe autonome dans une spécialité des sciences statistiques.

Atouts UQTR

Ce programme permet aux étudiants d'obtenir une bourse Universalis Causa.

Passage accéléré au doctorat

Un étudiant inscrit à la maîtrise en statistique avec mémoire peut être admis au doctorat sans être tenu de franchir toutes les étapes habituelles de la maîtrise, c'est-à-dire ne pas rédiger le mémoire (voir le Règlement des études) de l'Université Laval.

Admission**Trimestre d'admission et rythme des études**

Automne, hiver et été

Temps complet et temps partiel (voir note suivante)

Note : L'étudiant doit s'inscrire à temps complet durant au moins trois sessions. Cette exigence peut être satisfaite à tout moment en cours d'études, y compris pendant les sessions d'été.

L'étudiant doit terminer les cours propres au programme dans les six sessions qui suivent sa première inscription comme étudiant régulier.

Conditions d'admission

Études au Québec

Grade et discipline

Le candidat détient une maîtrise en statistique ou en mathématiques, ou un diplôme jugé équivalent.

Directeur de recherche

Avant de faire sa demande d'admission, le candidat doit prendre contact avec l'un des professeurs du programme. La direction de programme ne peut admettre un candidat que si un professeur a accepté de diriger ses travaux de recherche.

Exigences linguistiques

L'enseignement à l'Université Laval se fait en français. La maîtrise du français écrit et parlé est donc essentielle.

Même si la connaissance du français et de l'anglais n'est pas une condition d'admission, la réussite de ce programme d'études est liée à la capacité de l'étudiant de lire et de comprendre des textes en français et en anglais. L'étudiant qui ne maîtrise pas suffisamment ces langues pourrait éprouver des difficultés dans ses études. L'étudiant doit s'assurer d'avoir une bonne compréhension du français et de l'anglais et, si nécessaire, prendre des mesures pour développer ses compétences linguistiques en cours de formation (par exemple à l'École internationale de Français. En cas de lacunes importantes, la direction de programme peut imposer des correctifs.

Études hors Québec

Grade et discipline

Le candidat détient une maîtrise en statistique ou en mathématiques, ou un diplôme jugé équivalent.

Pour valider les équivalences de diplômes.

Directeur de recherche

Avant de faire sa demande d'admission, le candidat doit prendre contact avec l'un des professeurs du programme. La direction de programme ne peut admettre un candidat que si un professeur a accepté de diriger ses travaux de recherche.

Exigences linguistiques

L'enseignement à l'Université Laval se fait en français. La maîtrise du français écrit et parlé est donc essentielle.

Même si la connaissance du français et de l'anglais n'est pas une condition d'admission, la réussite de ce programme d'études est liée à la capacité de l'étudiant de lire et de comprendre des textes en français et en anglais. L'étudiant qui ne maîtrise pas suffisamment ces langues pourrait éprouver des difficultés dans ses études. L'étudiant doit s'assurer d'avoir une bonne compréhension du français et de l'anglais et, si nécessaire, prendre des mesures pour développer ses compétences linguistiques en cours de formation (par exemple à l'École internationale de Français. En cas de lacunes importantes, la direction de programme peut imposer des correctifs.

Modalités de sélection des candidatures

Le fait de satisfaire aux exigences d'admission à un programme n'entraîne pas automatiquement l'admission d'un candidat. Chaque demande est étudiée par la direction de programme qui tient compte, dans son évaluation, de la préparation antérieure du candidat, de son dossier scolaire, de son aptitude à la recherche et de l'ensemble de son dossier, ainsi que des ressources du département d'accueil.

Structure du programme et liste des cours

A moins d'indication contraire, un cours comporte trois (3) crédits.

Cours obligatoires (6 crédits)

STT8001 Examen doctoral : volet rétrospectif
STT8002 Examen doctoral : volet prospectif

Cours optionnels (12 crédits)

L'étudiant suit de 6 à 12 crédits parmi les cours suivants:

EPM7026 Introduction aux méthodes statistiques d'inférence causale
MAT7005 Probabilités avancées
STT7115 Statistique mathématique avancée
STT7125 Théorie et applications des méthodes de régression
STT7135 Analyse de durées de vie
STT7145 Statistique bayésienne
STT7325 Statistique computationnelle
STT7335 Méthodes d'analyse de données
STT7345 Sondages : modèles et techniques
STT7615 Introduction à la statistique génétique
STT7635 Séries chronologiques
STT7645 Statistique génétique (thèmes choisis)

L'étudiant suit de 0 à 6 crédits parmi les cours suivants:

EPM7000 Concepts et méthodes en épidémiologie
EPM7002 Epidémiologie appliquée
EPM7010 Essais cliniques et d'interventions
EPM7028 Analyse de survie
ETH7900 Conduite responsable de la recherche : cadres normatifs (1 crédit)
GIF7005 Introduction à l'apprentissage automatique
GLO7027 Analyse et traitement de données massives
GLO7030 Apprentissage par réseaux de neurones profonds
IFT7002 Fondements de l'apprentissage machine
STT7710 Sujets spéciaux 1 (1 crédit)
STT7720 Sujets spéciaux II (2 crédits)
STT7730 Sujets spéciaux III
STT7740 Sujets spéciaux IV (4 crédits)

Crédits de recherche (72 crédits)

Pour réussir son programme l'étudiant doit réaliser un travail de recherche comptant pour 72 crédits.

Travail de recherche

Avant la fin de sa deuxième session d'inscription comme étudiant régulier, il doit soumettre, pour entérinement par la direction de programme, un exposé écrit de son projet de recherche. Cet exposé, approuvé par le directeur de recherche, doit comporter une définition du problème posé et un calendrier pour la réalisation du projet.

À l'UQTR, le projet de thèse pourra être joint au Plan global d'études.

Autres renseignements

Règlements pédagogiques particuliers

Au plus tard durant la quatrième session de son inscription au programme, l'étudiant est soumis à un ensemble d'examen écrits et oraux, appelés examens prédoctoraux, destinés à vérifier le niveau de ses connaissances dans la spécialisation choisie et les domaines connexes. En cas d'échec, l'étudiant n'a droit qu'à une reprise.

L'étudiant doit s'inscrire à temps complet durant au moins trois sessions. Cette exigence peut être satisfaite à tout moment en cours d'études, y compris pendant les sessions d'été.

L'étudiant doit terminer les cours propres au programme dans les six sessions qui suivent sa première inscription comme étudiant régulier.

Description des activités

EPM7000 Concepts et méthodes en épidémiologie

Cette initiation complète à l'épidémiologie permet d'acquérir les connaissances indispensables à la compréhension des études épidémiologiques et les habiletés nécessaires à la réalisation d'analyses simples de données épidémiologiques.

Les thèmes suivants sont étudiés en profondeur : types de populations, mesures de fréquence, types d'études, mesures d'association, mesures d'exposition, validité des mesures, biais de sélection, d'information et de confusion, causalité, mesures d'effet, précision des mesures, intervalles de confiance, interactions, modification des mesures d'effet, introduction à la modélisation, stratégie d'analyse, lecture critique d'articles en épidémiologie, rédaction d'un rapport, dépistage.

EPM7002 Épidémiologie appliquée

Ce cours fait suite au cours EPM-7000 Concepts et méthodes en épidémiologie. La perspective épidémiologique de l'analyse de la relation dose-réponse et de l'interaction est présentée. Les thèmes suivants sont aussi étudiés : surveillance, études écologiques, statistiques vitales, dépistage, épidémiologie clinique et épidémiologie des maladies infectieuses.

EPM7010 Essais cliniques et d'interventions

Ce cours porte sur la méthodologie des essais cliniques et d'interventions. Dans la première partie du cours, les questions suivantes sont étudiées en détail : concepts statistiques, devis d'étude, interaction, surveillance et analyses intérimaires, gestion des données et contrôle de la qualité, analyses statistiques, événements intermédiaires, facteurs pronostiques.

Dans la deuxième partie du cours, l'étudiant fait un exposé sur un aspect particulier des essais cliniques et d'interventions.

EPM7026 Introduction aux méthodes statistiques d'inférence causale

Ce cours présente les méthodes statistiques d'inférence causale ainsi que leurs applications dans le domaine de la santé. L'étudiant apprend à connaître les paradigmes graphiques et contrefactuels à l'inférence causale et à obtenir des estimations de l'effet causal d'une variable d'exposition sur une variable réponse avec différentes méthodes statistiques (régression, pondération, score de propension, variables instrumentales).

Il apprend également à estimer l'effet d'une exposition variant dans le temps sur une variable réponse (g-estimation, modèles structurels marginaux). Il apprend enfin à discerner les hypothèses selon lesquelles les estimations obtenues ont réellement une interprétation causale, à apprécier la crédibilité de ces hypothèses et à effectuer des études de sensibilité pour mesurer l'impact d'un non-respect des hypothèses.

EPM7028 Analyse de survie

Au terme de ce cours, l'étudiant aura acquis les connaissances théoriques et pratiques des principales techniques d'analyse de survie utilisées en épidémiologie. Données censurées, temps de survie, fonction de survie, taux conditionnel de décès (fonction de hasard), taux non conditionnel de décès.

Estimation d'une fonction de survie par Kaplan-Meier, comparaison de deux ou plusieurs courbes de survie. Modèle de Cox, ajustement du modèle, vérification des hypothèses de base du modèle, analyses stratifiées, variables liées au temps, modèles paramétriques, événements récurrents, risques compétitifs, survie attendue et survie relative. Tailles d'échantillon. Application de la méthodologie à l'aide de SAS.

ETH7900 Conduite responsable de la recherche : cadres normatifs (1 crédit)

Par une approche multidisciplinaire, ce cours vise à favoriser l'acquisition de comportements responsables en recherche et à développer une pensée transversale et globale.

Le chercheur en formation reçoit un enseignement sur l'encadrement normatif de la recherche. Il évalue les enjeux soulevés par les développements biomédicaux ou par la recherche en santé des populations.

GIF7005 Introduction à l'apprentissage automatique

Ce cours porte sur les méthodes permettant l'inférence à partir d'observations de modèles de classement, de régression, d'analyse de données ou de prise de décision. Ces méthodes se caractérisent par une phase d'entraînement à partir de données ou d'expériences, afin d'effectuer des tâches qui seraient difficiles ou impossibles à faire par des moyens algorithmiques plus classiques.

Le cours aborde différentes approches actives en apprentissage et cherche à expliquer leurs mécanismes de base. Une perspective applicative de ces différentes techniques est également présentée, avec un accent particulier sur l'utilisation d'outils logiciels modernes.

L'étudiant qui a réussi le cours GIF-4101 ne peut s'inscrire à ce cours.

GLO7027 Analyse et traitement de données massives

Description des défis du traitement de données massives (big data). Prétraitement des données. Recherche de relations fréquentes. Algorithmes d'apprentissage supervisé et non supervisé pour données massives. Algorithmes de traitement des flux de données. Algorithmes de traitement de données sur internet : recherche d'information et systèmes de recommandation. Anonymisation des données et vie privée. Etudes de cas d'utilisation des données massives dans la pratique.

L'étudiant qui a réussi le cours GLO-4027 ne peut s'inscrire à ce cours.

Ce cours est parfois offert en comodal.

GLO7030 Apprentissage par réseaux de neurones profonds

Fournir des connaissances permettant la réalisation pratique de réseaux de neurones profonds pour la classification et donnant un aperçu des concepts et développements sur les architectures profondes, afin de stimuler les idées de recherche. Présentation des concepts de base : historique, fonctions d'activation et coûts, calcul des gradients et rétro-propagation, régularisation, techniques d'entraînements et architectures à convolution type CNN (AlexNet, ResNet, VGG).

Survol horizontal et rapide des concepts plus avancés (réseaux génératifs GAN, réseaux récurrents (RNN, LSTM, GRU), auto-encodeurs, parallélisation et modèles d'attention).

L'étudiant qui a déjà suivi le cours de premier cycle GLO-4030 ne peut s'inscrire à ce cours.

IFT7002 Fondements de l'apprentissage machine

Apprentissage probablement approximativement correct, dimension de Vapnik et Chervonenkis et minimisation du risque empirique. La minimisation du risque structurel et la longueur de description minimale. Les arbres de décision, les classificateurs premiers voisins, les classificateurs linéaires, le « Boosting », la régression logistique et les moindres carrés. La minimisation du risque régularisé et la régularisation de Tikhonov.

Les problèmes d'apprentissage convexes et la descente de gradient stochastique. Les méthodes à noyaux et les machines à vecteurs de support (SVM). Les classificateurs multiclassés, la prédiction de sorties structurées et l'apprentissage d'ordonnement. Les réseaux de neurones et l'algorithme de la rétropropagation du gradient de l'erreur.

MAT7005 Probabilités avancées

Rappel sur la théorie de la mesure. Variable aléatoire, distribution, espérance. Indépendance. Loi du tout ou rien de Kolmogorov. Convergence forte et loi des grands nombres. Convergence faible et théorème limite central. Espérance conditionnelle et théorie des martingales.

STT7115 Statistique mathématique avancée

Notions de probabilité : variables aléatoires, lois de probabilité, méthodes de calcul multidimensionnel, notions de convergence et théorèmes limites. Estimation ponctuelle et par intervalles de confiance : définitions, inférence par la méthode du maximum de vraisemblance. Test d'hypothèses : lemme de Neyman-Pearson, tests uniformément les plus puissants, tests basés sur la théorie de la vraisemblance.

STT7125 Théorie et applications des méthodes de régression

Régression linéaire. Modèles linéaires généralisés. Méthodes de sélection de variables. Validation de modèles. Modèles mixtes. Équations d'estimation généralisées. Couverture des aspects théoriques et mise en oeuvre pratique avec un logiciel statistique de tous ces modèles et méthodes.

STT7135 Analyse de durées de vie

Révision des lois usuelles en durée de vie, censure et troncation, processus de dénombrement et martingales, estimateurs non paramétriques : Nelson-Aalen et Kaplan-Meier, tests d'hypothèse à un, deux ou k échantillons, régression et modèle des risques proportionnels de Cox, fonctions de vraisemblance marginale et partielle, modèles de durée de vie accélérée log-linéaires : estimation et tests d'adéquation.

STT7145 Statistique bayésienne

Introduction à la philosophie de l'inférence bayésienne : théorème de Bayes, loi a priori, vraisemblance, loi a posteriori. Particularité des lois a priori. Inférence : choix du prior, calcul de la loi a posteriori, intervalle de crédibilité. Famille exponentielle et lois a priori conjuguées. Modèles hiérarchiques.

Méthodes de calcul des lois a posteriori : échantillonneur de Gibbs et autres techniques de Monte Carlo de simulations. Application à des problèmes précis, tels que les modèles linéaires, de régression mixtes, de classification et de mélanges. Apprentissage des méthodes de Monte Carlo par chaînes de Markov (MCMC).

L'étudiant qui a réussi le cours STT-4140 ne peut s'inscrire à ce cours.

STT7335 Méthodes d'analyse de données

Réduction de la dimensionnalité : par exemple, analyse en composantes principales et analyse canonique des corrélations. Classification non supervisée : classification hiérarchique, non hiérarchique et sur la base de modèles, évaluation de la qualité et choix du nombre de groupes. Classification supervisée : classifieurs linéaires et non linéaires, évaluation de la qualité des classifieurs.

STT7345 Sondages : modèles et techniques

Composantes d'une enquête complexe : strates, grappes et poids d'échantillonnage. Modification des poids d'échantillonnage pour

tenir compte de la non-réponse, possiblement par calage. Estimation de la variance dans des enquêtes complexes : méthodes de linéarisation, du « jackknife » et du « bootstrap ».

Modélisation des données recueillies à l'aide d'un plan d'échantillonnage complexe : tests d'association, modèles de régression et de régression logistique. Estimation dans des petits domaines : modèles au niveau des régions et modèles au niveau des unités.

STT7615 Introduction à la statistique génétique

Rappel des principes de base en génétique. Analyse de lien pour traits simples (méthodes non paramétriques). Analyse de lien pour traits complexes (méthodes paramétriques). Etude d'associations génétiques (études cas-témoins). Application : introduction à différents logiciels de statistique génétique. Rappel des principes de génétique moléculaire. Analyse de données micropuces.

STT7635 Séries chronologiques

Décomposition d'une série chronologique, tendance, saisonnalité, innovations, lissages, processus stationnaires, moyennes mobiles, processus autorégressifs, autocorrélations, bruit blanc, densité spectrale, prévisions optimales, modélisation SARIMA.

STT7645 Statistique génétique (thèmes choisis)

Brève introduction aux concepts génétiques. Sélection de sujets parmi les suivants : épidémiologie génétique, concepts et introduction; études d'agrégation familiale; analyse de liaison paramétrique; analyse de liaison non paramétrique, méthodes de partage d'allèles; études d'association de population; test de déséquilibre de transmission (test TDT); traits quantitatifs; données de puce à ADN (Microarray).

STT7710 Sujets spéciaux 1 (1 crédit)

Le sujet traité varie d'une session à l'autre et est annoncé durant la période d'inscription.

STT7720 Sujets spéciaux II (2 crédits)

Le sujet traité varie d'une session à l'autre et est annoncé durant la période d'inscription.

STT7730 Sujets spéciaux III

Le sujet traité varie d'une session à l'autre et est annoncé durant la période d'inscription.

STT7740 Sujets spéciaux IV (4 crédits)

Le sujet traité varie d'une session à l'autre et est annoncé durant la période d'inscription.

STT8001 Examen doctoral : volet rétrospectif

Cette épreuve, obligatoire pour tout étudiant inscrit au doctorat en statistique, vise à s'assurer que l'étudiant possède les connaissances requises pour entreprendre un doctorat dans le champ d'études de son projet de recherche.

STT8002 Examen doctoral : volet prospectif

Cette épreuve, obligatoire pour tout étudiant inscrit au doctorat en statistique, vise à s'assurer que l'étudiant a une vision claire du projet de recherche qu'il se propose d'entreprendre, et qu'il possède les aptitudes requises et l'autonomie nécessaire pour poursuivre des recherches et mener à bien son projet.