



Canadian Statistical Sciences Institute
Institut canadien des sciences statistiques

Simon Fraser University
8888 University Drive
Burnaby BC Canada V5A 1S6

canssi.ca | incass.ca

Updated October 27, 2025

CANSSI QUEBEC RECRUITMENT SCHOLARSHIP 2025 BOURSE DE RECRUTEMENT DE L'INCASS QUÉBEC 2025

List of Potential Supervisors | Liste des superviseurs potentiels

Funding amounts are in Canadian dollars. | Les montants de financement sont indiqués en dollars canadiens.

1. **Archer Yi Yang** (archer.yang@mcgill.ca)

Associate Professor, Department of Mathematics and Statistics, McGill University

Research interests: Statistical machine learning; AI for science; AI safety research; high-dimensional inference; computational statistics; biomedical, biochemical and industrial data science; applications in drug discovery. Please see my recent publications for topics of potential projects:

www.math.mcgill.ca/yyang/pub.html

Standard support (funding): The 4-year funding package has an overall value of \$135,024. Specifically, the Department of Math and Stats will provide an amount equal to the total expected university tuition (\$36K for years 1, 2, 3 and \$21K for year 4) for four years plus a guaranteed base amount (\$2K) for the first two years. Students who do not receive external scholarships (NSERC, FRQNT, etc.) will also receive the supplement amount (\$12K for years 1, 2, 3, 4). Students are responsible for paying all university student fees.

Typically, our PhD students augment this guaranteed funding with one Teaching Assistantship per year that is worth an additional \$4,954.50. The Department can promise this level of support for four years, contingent on satisfactory progress towards your degree.

2. **Josée Dupuis** (josee.dupuis3@mcgill.ca)

Professor, Department of Epidemiology, Biostatistics and Occupational Health, McGill University

Research interests: My research focuses on the development and application of statistical and computational methods for genome-wide association studies, gene-by-environment interaction investigation, genetic meta-analysis, rare

variant analysis, and omics data integration, with special emphasis on developing innovative statistical and machine learning methods for large-scale genetic and genomics summary statistics when individual-level data are inaccessible.

Despite this wealth of data, practical applications face constraints due to privacy concerns and the restricted availability of individual-level data. Researchers increasingly turn to summary statistics to overcome these challenges, yet a gap persists in the development of methodologies designed to leverage these datasets effectively.

One of my current projects involves the development of methods for multi-ancestry co-localization of genome-wide association signals. Co-localization is often performed to help in the interpretation of genome-wide association signals. However, because co-localization relies on linkage disequilibrium, or correlation between genetic variants, and that linkage disequilibrium varies by ancestry, most approaches have been designed for single-ancestry cohorts. With collaborators, I am working on extending these approaches to be valid when used in multi-ancestry cohort settings.

Another project involves the development of a framework to integrate multi-omics summary data, handling the complexity and dimensionality of these diverse data types. The goal is to develop statistical methods to derive insights across multiple omics platforms, identifying key pathways, leading to a better understanding of complex diseases such as type-2 diabetes.

Standard support (funding): All admitted PhD students in the Department of Epidemiology, Biostatistics and Occupational Health are offered a minimum funding package, as set by the Faculty of Medicine and Health Sciences. For the academic year 2025-2026, the yearly amount was set to \$29,731 for Canadian students and \$46,077 (\$29,731 + international fee differential) for international students. See www.mcgill.ca/medhealthsci-gradstudies/fmhs-graduate-student-funding-policy/minimum-funding-packages-fmhs-graduate-programs-2025-2026 for details.

3. **Juliana Schulz** (Juliana.schulz@hec.ca)

Associate Professor, Department of Decision Sciences, HEC Montréal

Research interests: My work primarily focuses on dependence modelling, with the goal of developing novel multivariate statistical models appropriate for various types of data, including multi-dimensional discrete and mixed outcomes. With a background in actuarial mathematics, I am particularly interested in developing statistical methods appropriate for the analysis of multivariate claims data stemming from non-life insurance. I am also interested in biostatistical approaches for precision medicine wherein the objective is to

develop robust statistical methods for estimating optimal personalized treatment strategies.

Currently, I have several ongoing projects that focus on dependence modelling, particularly in developing flexible multivariate models appropriate for discrete and mixed-type data. Such models are particularly relevant in the context of ratemaking in non-life insurance. Ratemaking involves a comprehensive evaluation of past claims, which are typically multidimensional as they reflect the multi-peril nature of insurance products. In order to effectively manage and price the underlying risk of an insurance portfolio, the analysis of claims data must be carried out on a multivariate scale. In analyzing claims, it is often helpful to consider both components driving the claims process, namely, the frequency (number of claims) and severity (loss amount per claim). A main goal of my current research program is to develop a two-part modelling framework involving multivariate frequency and multivariate severity models, as well as mixed-outcome models allowing to capture the association between claim counts and amounts. These multivariate models must allow to adequately capture the complex dependence structures inherent in claims data. In order for the models to be particularly appropriate for ratemaking purposes, covariate effects must also be incorporated into the modelling framework. As such, regression-based versions of the multivariate models will also be explored, both in a more classical GLM type setting as well as through more flexible tree-based methods.

Standard support (funding): Students admitted to the PhD program at HEC Montréal receive a funding offer for a period of four years. The standard offer is a minimum of \$120,000 (typically \$30,000 per year), in addition to a tuition waiver and \$5,000 in financial assistance for participation in scientific activities. (Note that in the case where a student receives significant external funding, the offer may be adjusted.) The offer is personalized for each student and may include several sources of funding, including teaching and assistantship, and funding in the form of scholarships. The detailed offer is communicated to students at the time of the admission offer.

4. **Karim Oualkacha** (oualkacha.karim@uqam.ca)

Professeur, Département de mathématiques, Université du Québec à Montréal

Domaines de recherche : Professeur de statistique au Département de mathématiques de l'Université du Québec à Montréal (UQAM) depuis juin 2012. Mes recherches portent sur le développement de méthodologies et d'outils d'inférence statistique pour l'analyse de données complexes et de grande dimension, avec un intérêt marqué pour les applications en biostatistique, en particulier dans les études génétiques et pangénomiques. Je m'intéresse à la

régression quantile et expectile, notamment dans leurs versions pénalisées et robustes, aux modèles linéaires généralisés mixtes pénalisés, à l'analyse multidimensionnelle, à la modélisation de la dépendance par copules et aux approches de réduction de dimension. Je développe également des méthodes pour la complétion de matrices de données de grandes dimensions en présence de covariables informatives, un domaine présentant un fort potentiel méthodologique et applicatif. Mes travaux incluent aussi le développement d'approches adaptées à l'inférence statistique en environnements distribués et à l'apprentissage fédéré. Ces contributions méthodologiques trouvent des applications directes dans l'analyse de données génomiques, épigénétiques et transcriptomiques, dans la modélisation des interactions gène-environnement, dans l'étude de données familiales corrélées et longitudinales, ainsi que dans l'exploration de données hétérogènes présentant des structures de dépendance complexes.

Je propose un projet potentiel, en collaboration avec mon collègue Cédric Beaulac (dépt. mathématiques, UQAM). Nous proposons de développer de nouvelles méthodologies pour la complétion de matrices de grande dimension en présence de covariables informatives, en tirant parti d'architectures neuronales profondes. L'approche envisagée repose sur des autoencodeurs enrichis par des réseaux de neurones élément-par-élément, capables d'apprendre de manière adaptative des fonctions d'activation complexes pour améliorer la reconstruction des données manquantes. L'intégration explicite des covariables (par exemple cliniques, environnementales ou omiques) dans ces modèles permettra de guider l'imputation en exploitant des sources d'information externes pertinentes. Ce projet vise ainsi à fournir un cadre théorique et computationnel robuste pour la complétion de matrices au-delà des hypothèses classiques de faible rang (low-rank matrix completion), avec des applications directes en biostatistique et en génomique, où la structure des données est souvent hétérogène, parcellaire et fortement corrélée.

Soutien standard (financement) : Je propose à mes doctorants un soutien financier stable et compétitif de l'ordre de 30 000 à 35 000 \$ par année pendant quatre ans, combinant plusieurs sources institutionnelles, externes et personnelles. Ce plan vise à garantir la sécurité financière tout en favorisant un environnement académique riche et stimulant.

Le financement inclut d'abord la bourse INCASS-BRIQ (7 000 \$ par année pendant deux ans) ainsi qu'une bourse de recrutement doctoral offerte par le centre STATQAM (5 000 \$ une fois). À cela s'ajoutent les bourses de l'Institut des sciences mathématiques (ISM) (6 000 \$ par année pendant deux ans), qui permettent également de participer à des écoles d'été et activités collaboratives. Les étudiants internationaux bénéficient par ailleurs d'un soutien universel de

l'UQAM sous forme d'exemption des frais supplémentaires de scolarité.

Le département de mathématiques offre aussi des charges d'auxiliaire d'enseignement (environ 100 heures par session, soit près de 6 000 \$ par année), qui assurent à la fois un revenu complémentaire et une expérience pédagogique formatrice. Enfin, je complète systématiquement ces sources par mes fonds de recherche (subvention à la découverte du CRSNG), afin de garantir un financement global se situant entre 30 000 \$ et 35 000 \$ par année, indépendamment de l'obtention des bourses externes.

En résumé, mes doctorants reçoivent un soutien complet, structuré et garanti, leur permettant de se consacrer pleinement à leurs recherches et à leur développement académique.

5. **Denis Talbot** (denis.talbot@fmed.ulaval.ca)

Professeur titulaire, Département de médecine sociale et préventive, Université Laval

Domaines de recherche : Mes travaux de recherche portent globalement sur le développement, l'adaptation et l'évaluation des méthodes statistiques d'inférence causale à partir de données de santé épidémiologiques. L'inférence causale vise à déterminer l'impact qu'aurait une intervention, par exemple un traitement médical ou une politique publique, sur un résultat.

Le projet de recherche pour lequel je souhaite recruter une personne étudiante au doctorat vise à développer de nouvelles méthodes d'inférence causale pour la médecine personnalisée lorsque l'issue d'intérêt est un temps jusqu'à l'événement (c'est-à-dire, pour les données de survie). La médecine personnalisée vise à identifier des règles de décision prenant en compte les caractéristiques individuelles des personnes afin d'optimiser leurs résultats de santé. Bien que plusieurs méthodes ont été développées au cours des dernières années, il demeure des défis importants dans le cas de données de survie. La personne recrutée travaillera spécifiquement aux deux objectifs suivants : 1) développer une méthode adéquate lorsque les moments de prise de décision de traitement sont informatifs concernant la réponse d'intérêt; 2) développer une méthode de sélection de variables de personnalisation. Nous utiliserons la théorie de l'efficacité semiparamétrique ou nonparamétrique afin de développer des estimateurs qui permettent d'intégrer l'apprentissage automatique tout en conservant des propriétés théoriques désirables (par exemple, convergence à vitesse racine-n).

Ce projet est motivé par le besoin de mieux personnaliser les recommandations de traitement par hormonothérapie pour les femmes atteintes du cancer du sein. Ce type de traitement est actuellement recommandé à la majorité des femmes ayant un cancer du sein hormonodépendant non

métastatique, mais ne produit pas toujours les effets désirés. Afin d'éviter des effets secondaires inutiles aux femmes qui n'auront pas de bénéfices, il est primordial de mieux identifier ces femmes. Le projet sera réalisé en collaboration étroite avec des experts en épidémiologie et une patiente partenaire. Le projet est financé par les Instituts de recherche en santé du Canada.

Soutien standard (financement) : Le programme de doctorat en biostatistique est un programme bi-facultaire géré conjointement par la Faculté de médecine et la Faculté de sciences et de génie de l'Université Laval. La Faculté de médecine recommande fortement un financement minimal de 25 000 \$ par année en bourses pour les personnes au doctorat. Je m'engage à m'assurer que la personne recrutée obtiendra un soutien à ce niveau pour une durée minimale de quatre ans.

6. **Frédéric Ouimet** (frederic.ouimet2@uqtr.ca)

Professeur régulier en statistique, Département de mathématiques et informatique, Université du Québec à Trois-Rivières

Research interests: My research interests are in probability and mathematical statistics. My group develops theory and methods for matrix-valued data, with three connected themes: (i) nonparametric estimation on the cone of symmetric positive-definite (SPD) matrices (Wishart kernel density estimation, regression, and classification) under temporal/spatial dependence; (ii) Stein's method for matrix-variate laws (matrix-normal, Wishart, etc.) to obtain explicit approximation bounds and asymptotic distributions for matrix statistics; and (iii) Gaussian and Wishart moments and inequalities (Wilks' problem; the Gaussian Product Inequality and its Wishart analogue). Applications include diffusion tensor imaging (segmentation and denoising, tractography) and financial econometrics (realized-covariance modeling and risk).

Standard support (funding): I aim to give at least \$35,000 CAD per year after the first year (i.e., for Years 2, 3, and 4), using a combination of teaching duties (such as teaching assistantships), institutional funding, my research grants, the research grants of my collaborators, and various scholarships.

7. **Alexandre Bureau** (alexandre.bureau@fmed.ulaval.ca)

Professeur, Département de médecine sociale et préventive, Université Laval

Domaines de recherche : Plusieurs maladies psychiatriques ont des causes complexes, qui incluent des facteurs génétiques parfois nombreux, et des contributions environnementales importantes. De plus, les symptômes associés aux différentes maladies psychiatriques sont parfois variables d'un individu à

l'autre. Mes travaux visent à développer et à appliquer des méthodes statistiques pour identifier les gènes impliqués dans ces maladies complexes.

Les approches statistiques que je développe permettent d'analyser des génomes entiers de personnes affectées par des maladies psychiatriques et des membres de leur famille. Ce type d'analyse génétique complexe permet de détecter des variations génétiques rares qui sont associées aux maladies.

Des variants génétiques rares et fréquents interagissent pour causer les maladies complexes. Il est reconnu que prendre en compte un facteur modifiant l'effet d'un variant génétique augmente la puissance de détection de ce variant. Un projet potentiel pour un étudiant au doctorat consisterait à développer un test statistique pour détecter des variants rares dans des familles en conditionnant sur l'effet des variants génétiques fréquents capturé par un score de risque polygénique, qui agit alors comme facteur modifiant l'effet du variant rare. Les enjeux à traiter incluent la dérivation des variances et covariances des termes entrant dans la statistique de test et la prise en compte de la dépendance entre les variants rares et fréquents qui survient dans une population structurée en sous-populations. Le test développé sera appliqué aux données de séquence génomique de diverses études familiales de traits complexes.

Soutien standard (financement) : Financement sous forme de bourse d'un montant de 27 000 \$ par année, renouvelable deux fois pour un total de trois ans.

À ce montant s'ajoutent :

- Bourse d'admission à la suite d'un passage accéléré : D'une valeur de 500 \$, cette bourse récompense la réussite du passage accéléré.
- Bourse pour la remise du plan de collaboration : D'une valeur de 500 \$, cette bourse récompense le dépôt du Plan de collaboration.
- Bourse de réussite de l'examen de doctorat: D'une valeur de 2 500 \$, cette bourse récompense la réussite de l'examen de doctorat.
- Bourse de progression: D'une valeur de 1 750 \$, cette bourse récompense la progression satisfaisante de l'étudiante ou l'étudiant à la 7e session d'inscription.
- Bourse de rédaction 1: D'une valeur de 1 750 \$, cette bourse récompense la progression satisfaisante de l'étudiante ou l'étudiant à la 9e session d'inscription.

8. **Geneviève Gauthier** (genevieve.gauthier@hec.ca)

Professeure titulaire, Département de sciences de la décision, HEC Montréal

Domaines de recherche : Mes recherches portent principalement sur la volatilité, que j'aborde sous plusieurs angles. Je m'intéresse à la modélisation de la surface de volatilité implicite et à ses applications pour la gestion des risques,

ainsi qu'à l'information contenue dans les produits dérivés liés à la volatilité comme le VIX et le VVIX. J'étudie aussi des modèles à composantes permettant de décomposer la structure de volatilité et d'identifier les facteurs qui gouvernent son évolution. Ces travaux mobilisent à la fois des méthodes de filtrage, des approches d'apprentissage statistique et des techniques d'apprentissage profond adaptées aux données financières.

Soutien standard (financement) : À HEC Montréal, les étudiants inscrits au programme de doctorat reçoivent un financement minimal de 30 000 \$ par année ainsi qu'une exonération des frais de scolarité pour les 4 premières années d'études.

9. **Pierre Duchesne** (pierre.duchesne@umontreal)

Professeur titulaire, Département de mathématiques et de statistique, Université de Montréal

Domaines de recherche : Je travaille principalement en statistique appliquée, notamment en séries chronologiques et en échantillonnage.

Soutien standard (financement) : Le plan de financement intégré de mon établissement prévoit un financement de l'ordre de 20 000 \$/an.

10. **Cédric Beaulac** (beaulac.cedric@uqam.ca)

Professeur, Département de mathématiques, Université du Québec à Montréal

Domaines de recherche : Je mène un programme de recherche dédié au développement de nouveaux modèles pour l'analyse d'images, où celles-ci sont représentées par les formes, les textures et les couleurs des objets qu'elles contiennent. Ce programme vise également à explorer les applications de ces modèles dans divers domaines, tels que la sécurité routière, la santé et l'écologie. Mon programme de recherche est tout jeune et de nombreux problèmes restent à résoudre, par exemple la paramétrisation de l'évolution de la forme d'un objet à travers le temps, la caractérisation de la texture d'un objet ou encore la classification d'objets observés partiellement.

Research interests: I lead a research program dedicated to the development of new models for image analysis, where images are represented through the shapes, textures, and colors of the objects they contain. This program also aims to explore the applications of these models in various fields such as road safety, health, and ecology. As the program is still in its early stages, several important methodological challenges remain to be addressed, including the parameterization of shape evolution over time, the characterization of object texture, and the classification of partially observed objects.

Soutien standard (financement) : Je prévois offrir un soutien financier annuel

variant entre 35 000 \$ et 40 000 \$ aux étudiant·e·s de doctorat que je superviserai. Ce soutien sera assuré grâce à un ensemble de sources complémentaires : un financement de base de l'Université du Québec à Montréal (UQAM), une contribution du centre de recherche STATQAM, mes fonds de recherche à la Découverte du CRSNG, ainsi que des emplois étudiants offerts à l'UQAM.

Standard support (funding): I plan to provide annual financial support ranging from \$35,000 to \$40,000 to the doctoral students under my supervision. This support will be secured through a combination of complementary sources: a base funding from the Université du Québec à Montréal (UQAM), a contribution from the STATQAM research centre, my NSERC Discovery research funds, as well as student employment opportunities at UQAM.

11. Léo Belzile (leo.belzile@hec.ca)

Professeur agrégé, Département de sciences de la décision, HEC Montréal

Domaines de recherche : valeurs extrêmes: méthodes fréquentistes et Bayésiennes, modèles spatio-temporels en haute-dimension

Projets potentiels:

1. adaptation de la méthodologie pour les extrêmes géométriques au cas spatio-temporel. Méthodologie pour les seuils radiaux-angulaires.
2. Sélection de seuils avec exploration de la pondération
3. Méthodes hiérarchiques Bayésiennes pour champs spatiaux combinant données de réanalyse tirées de modèles climatiques et données de station avec des aprioris Gaussiens multidimensionnels.
4. Modèles semiparamétriques pour la fonction de risque avec données de survie dans le cas d'excès de seuils.

Soutien standard (financement) : Les étudiant(e)s au doctorat à HEC bénéficient de CAD 30K par an sous forme de bourse; ce montant combine le financement de l'école et du superviseur, garanti pendant quatre ans. Un supplément de CAD 5K est réservé pour la diffusion des résultats de recherche (congrès, etc.).

12. Arthur Charpentier (charpentier.arthur@uqam.ca)

Professeur, Département de mathématiques, Université du Québec à Montréal

Domaines de recherche : Travaux sur l'équité algorithmique, l'interprétabilité des modèles prédictifs d'apprentissage machine et la calibration, sur les jeux coopératifs et non coopératifs, sur les risques de catastrophes naturelles. Fort intérêt pour la modélisation des risques dans un contexte d'assurance.

Publication dans des conférences d'apprentissage machine (AAAI, Neurips,

IJCAI, AiStats, ICML) ou des journaux de statistiques / OR / actuariat.

Soutien standard (financement) : Financement sur 4 ans, 36 000 CAD par année, de bourse. Nombreuses possibilités de compléments pour aller présenter en conférence (2 ou 3 conférences par an), et possibilité de séjour à l'étranger à partir de la 3ème année (3–6 mois). Possibilité de supervisions de stagiaires d'été.

13. Mireille Schnitzer (mireille.schnitzer@umontreal.ca)

Professeure de biostatistique, Faculté de pharmacie, Université de Montréal

Research interests: My research program is on causal inference with complex epidemiological designs. Our statistical approaches typically fall under the popular semiparametric efficiency framework in causal inference and incorporate machine learning for nuisance mode estimation.

One axis of my current program is causal inference with longitudinal data and time-dependent treatments. Machine learning methods, supported by semiparametric theory, exist to estimate low-dimensional parameters that represent causal effects. However, in longitudinal settings, estimation may require a very large sample size and substantial exposure variability over a high-dimensional probability space, which is not available in standard epidemiological applications. We are thus developing new approaches to variable (confounder) selection, nonparametric smoothing, and statistical inference in settings with partially identified causal effects. Potential projects suitable for PhD students include:

- Variance estimation under propensity score smoothing over instrumental variables,
- Extensions of longitudinal outcome-adaptive LASSO to repeated measures and survival outcomes, and
- Applications of longitudinal outcome-adaptive LASSO in pharmacoepidemiology.

Another axis is causal inference under biased sampling designs, and in particular the test-negative design (TND). The TND is used almost exclusively to estimate the effectiveness of vaccines to prevent morbidity and healthcare utilization due to an infectious disease of interest. Due to its relative ease of implementation with administrative health data and cost-effectiveness compared to other designs, this design was instrumental in evaluating real-world vaccine effectiveness during the COVID-19 pandemic. We are interested in the identifiability of causal effects under variations of this design, including both novel variations that we propose for bias reduction, and popular existing variations. Potential projects suitable for PhD students are:

- Matching for confounding control under the test-negative design,

- Matching for confounding control under the case-control design, and
- Extensions of instrumental variable approaches to test-negative design sampling under a randomized vaccination encouragement intervention.

Standard support (funding): Typically, PhD students in my lab receive \$32K in yearly stipends. Competitive bursaries and scholarships replace the supervisor's contribution. Teaching assistant, statistics lab supervision, and even lectureship positions are available depending on experience and linguistic competencies, and any income from these is in addition to the bursary amount. Graduate students have access to offices, high-performance computing resources, and dynamic collaborations in health science. Multiple PhD programs are available at Université de Montréal which could support a student following a biostatistics-themed PhD (see programs in statistics, epidemiology, and pharmaceutical sciences) and professors can supervise students across departments. Students also have access to a network of Montreal-area graduate courses in statistics, biostatistics, computer science, etc. My students typically meet with me once a week and I follow their projects closely. Students with research results are financially supported in presenting their results at local, national, and international conferences; all students are supported to attend to local conferences and workshops as funds permit.

14. **Marzia Angela Cremona** (marzia.cremona@fsa.ulaval.ca)

Professeure agrégée, Département d'opérations et systèmes de décision, Université Laval

Domaines de recherche : Ma recherche vise à développer de méthodes d'apprentissage statistique/automatique pour les données complexes dans les sciences biomédicales et sociales. En particulier, elle se concentre sur les données fonctionnelles, c'est-à-dire les données qui varient sur un continuum et peuvent être représentées sous forme de courbes ou de surfaces.

Mes projets potentiels de doctorat ont comme objectif d'étendre le champ des techniques d'apprentissage statistique à l'analyse de données fonctionnelles. Ces projets concernent le développement de nouvelles méthodes non supervisées (clustering, découverte de motifs, analyse en composantes principales, etc.) et/ou supervisées (sélection de variables dans un modèle de régression, test d'hypothèses, etc.) pour l'analyse de données fonctionnelles. Des aspects importants sont l'implémentation des méthodes développées dans de logiciels conviviaux et efficaces, et l'application à l'analyse de données réelles issues des sciences biomédicales et sociales.

Soutien standard (financement) : Les doctorants vont recevoir au moins 25 000 CAD/an pendant 4 ans (incluant les bourses de la faculté, les bourses de département, ainsi que les bourses du superviseur). Ils peuvent également

effectuer de travaux de corrections en tant qu'auxiliaires d'enseignement.

15. Mélina Mailhot (melina.mailhot@concordia.ca)

Associate Professor, Department of Mathematics and Statistics, Concordia University

Research interests: dependence modeling, risk measures, geometric representation, uncertainty quantification. Potential project: The candidate will work in the development and study of properties of multivariate stochastic processes. Developing uncertainty frameworks and risk measures in contexts of multivariate probability space.

Standard support (funding): At Concordia University, every PhD student is offered a financial package worth nearly \$30,000/year.

16. Simone Brugiapaglia (simone.brugiapaglia@concordia.ca)

Associate Professor, Department of Mathematics and Statistics, Concordia University

Research interests: My current research lies at the intersection of machine learning and numerical analysis. Although I am not a statistician by training, my recent work has developed a strong statistical component, encompassing areas such as the theory of deep learning, high-dimensional regression and approximation, compressive sensing, signal reconstruction, and scientific machine learning.

I currently have projects available in deep learning theory, with a focus on topics such as convergence guarantees, stability, and robustness. Another area of particular interest is the theoretical study of black-box machine learning auditing methods, such as those based on Shapley values. Additional topics include efficient algorithms for sparse PCA in large-scale datasets, the theory and practical implementation of generative models, and the application of machine learning techniques to challenging problems in computational chemistry.

Standard support (funding): Typical financial support for PhD students in my department (after tuition and other fees) is approximately \$21,000 per year. This amount includes a combination of teaching assistant (TA) duties and a research bursary provided by the supervisor. I also support students in preparing and submitting additional scholarship applications, such as those offered by the Natural Sciences and Engineering Research Council of Canada (NSERC) and the Fonds de recherche du Québec (FRQ).

17. Arusharka Sen (arusharka.sen@concordia.ca)

Associate Professor, Department of Mathematics and Statistics, Concordia

University

Research interests: My current research interests are as follows:

(1) Multivariate Survival Analysis: Recently my co-author and I have constructed a multivariate version of the Kaplan-Meier estimator under general multivariate random censoring using an eigenfunction approach. I now propose to use a similar approach to develop a multivariate version of the Vardi estimator under multivariate multiplicative censoring, which has potential application in prevalent cohort data analysis. Further, I also propose to construct an alternative to the non-parametric maximum likelihood estimator (NPMLE) under interval censoring with $k \geq 2$ intervals. The construction as well as the asymptotic distribution theory of the estimator is expected to be easier than that of the NPMLE.

(2) Statistical Learning Methods: Recently my doctoral student and I developed a distance-based approach to learning the mixing matrix in Independent Components Analysis, an important statistical learning problem, where one observes only a linear mixing of independent random variables (components). I now propose to extend this approach to smooth mixing of components, i.e., where observed data is a smooth function of independent components. In addition, I propose to develop a unified, non-parametric, likelihood-ratio-based approach to several other statistical learning problems, such as Support Vector Machines (SVM) and Sufficient Dimension Reduction. This approach is intuitive and avoids the unnecessarily stringent assumptions currently employed in these problems, for instance, linear or non-linear separability of data vectors from different samples, as is assumed in SVM.

Standard support (funding): At this time I have no research grant, so I will have to rely on the CANSSI support as well as TA support from my department for my PhD student.

18. **Aurélie Labbe** (aurelie.labbe@hec.ca)

Professeure, Département de sciences de la décision, HEC Montréal

Domaines de recherche : Je travaille sur des sujets variés en statistique, motivés principalement par des applications dans les domaines de la santé et des transports, et nécessitant un développement méthodologique statistique innovateur. Biostatisticienne de formation, j'ai développé une expertise en statistique génétique et plus spécifiquement dans l'analyse de données d'expression génétique, de méthylation, ou de microbiome. De façon générale, je m'intéresse à la modélisation de la structure de corrélation complexe dans ces données en utilisant le concept de facteurs latents. Beaucoup de mes travaux se déclinent dans un cadre Bayésien. Dans les dernières années, j'ai élargi mes intérêts à des applications dans le domaine du transport, et spécifiquement la

sécurité routière ou la surveillance du réseau des transports à l'aide de différents capteurs mobiles ou fixes. La structure de corrélation des données se modélise souvent dans un cadre spatio-temporel et je travaille à l'adaptation de certaines méthodes en apprentissage machine ou apprentissage statistique à ce contexte.

Voici quelques exemples de projets de recherche réalisés par mes étudiants dans les dernières années :

- Régression matricielle pour données incomplètes et imputation en utilisant la factorisation à faible rang. Application aux données de comptage de vélos BIXI.
- Agrégation temporelle et prédiction. Application au problème de prédiction de la consommation de CO₂ dans les voitures hybrides.
- Forêts aléatoires dans un contexte spatio-temporel.
- Prédiction de la propagation des délais dans le métro lors d'incidents.
- Modèles de régression linéaire généralisée robuste pour les données de microbiome.
- Méthodes d'intégration de données de capteurs fixes et mobiles dans un contexte de prédiction de la demande en transport : modèles dynamiques Bayésiens pour données spatio-temporelles.
- Prédiction des accidents routiers.

Soutien standard (financement) : À HEC Montréal, nous finançons les étudiants de doctorat pour un montant minimum de 30,000 \$ par an pour 4 ans. Des bourses d'exemption de frais de scolarité sont disponibles pour les étudiants étrangers.