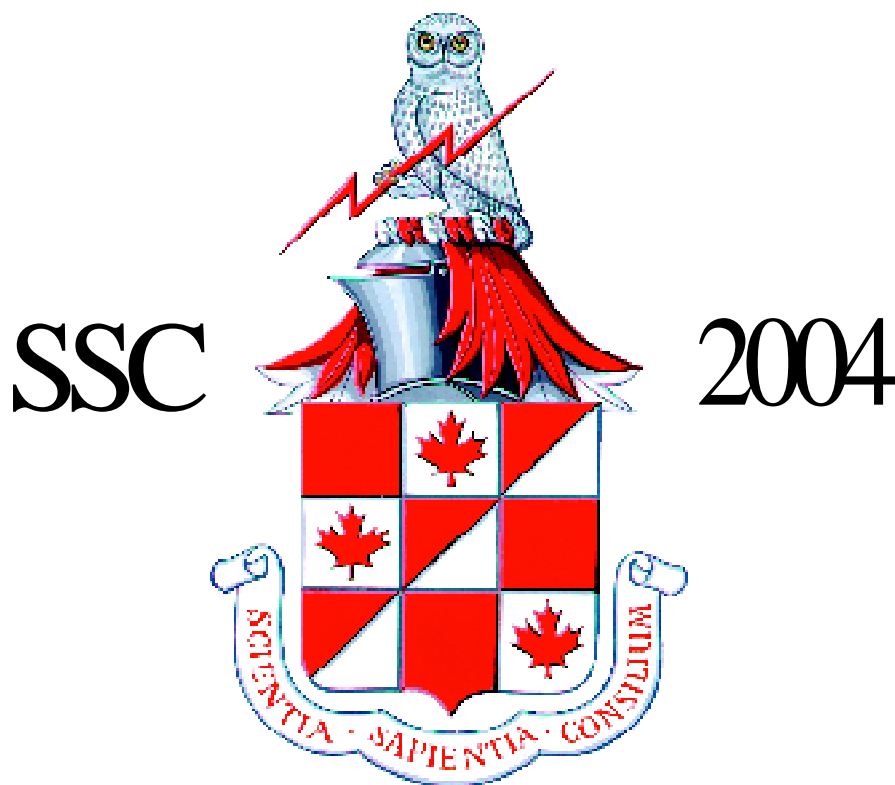


32<sup>E</sup> CONGRÈS ANNUEL DE LA  
SOCIÉTÉ STATISTIQUE DU CANADA

32<sup>ND</sup> ANNUAL MEETING OF THE  
STATISTICAL SOCIETY OF CANADA

30 MAI - 2 JUIN • MAY 30 - JUNE 2, 2004



CENTRE DE RECHERCHES MATHÉMATIQUES  
& DÉPARTEMENT DE MATHÉMATIQUES ET DE STATISTIQUE,  
UNIVERSITÉ DE MONTRÉAL, MONTRÉAL, QUÉBEC

# Table of Contents • Table des matières

Table of Contents • Table des matières	3
Welcome • Bienvenue	4
Sponsors • Commanditaires	5
Organizers • Organisateurs	6
Exhibitors • Exposants	7
General Information • Information générale	7
Committees and Meetings • Comités et réunions	13
List of Sessions • Liste des sessions	16
Program • Programme	21
Abstracts • Résumés	57
Author Index • Index des auteurs	163
Author Index (per session) • Index des auteurs (par session)	168

# Welcome • Bienvenue

The Centre de recherches mathématiques (CRM) and the Department of Mathematics and Statistics of Université de Montréal are proud to welcome you to Montréal for the 32nd Annual Meeting of the Statistical Society of Canada. As Université de Montréal is celebrating its 125th anniversary, the timing could not be better for a first meeting of the SSC on our campus. We are also glad to welcome the members of the *Association des statisticiennes et statisticiens du Québec* who have agreed to meet with us.

Many people have worked hard to prepare a meeting that should be intellectually challenging. Moreover, we have planned some very exciting social events, but we have to confess that the competition from the city itself will be hard to beat! The participants wearing blue nametags are all very familiar with the campus of Université de Montréal and will be willing to help you with any questions you may have.

We hope that you will have some time to explore Montréal and discover the *joie de vivre* of the Montréalaises and Montréalais. Whether you wander around the Crescent Street bars downtown, close to the conference hotels, discover the charm of Old Montreal (Champ-de-Mars metro station) or experience the French culture in the *Quartier latin* (Berri-UQAM metro station), you can count on having a good time. If you prefer nature, the campus is on the northern slope of Mont Royal. It is one of the nicest city parks and is well worth exploring (Bus 11 from the Mont-Royal metro station). You will have a chance to discover it in part as the Banquet will be held at the *Chalet du Mont-Royal*. Finally, art lovers may be interested in an exposition presented on campus by the CRM and featuring science (see the description below).

Le Centre de recherches mathématiques (CRM) et le Département de mathématiques et de statistique de l'Université de Montréal sont fiers de vous souhaiter la bienvenue à Montréal dans le cadre du 32<sup>e</sup> Congrès annuel de la Société statistique du Canada. L'Université de Montréal célèbre cette année son 125<sup>e</sup> anniversaire et nous ne pouvons que nous réjouir d'être les hôtes de notre premier congrès de la SSC. Nous voulons également souhaiter la bienvenue aux membres de l'Association des statisticiennes et des statisticiens du Québec, qui ont bien voulu accepter notre invitation de tenir leur rencontre avec nous.

Nous avons été nombreux à participer à la préparation de ce congrès, qui saura vous poser des défis intellectuels stimulants. Nous avons aussi prévu divers événements sociaux d'intérêt, bien qu'à ce chapitre, il faut admettre que la ville elle-même nous fera une rude compétition ! Les personnes portant un macaron bleu sont toutes très familières avec le campus et répondront volontiers à vos questions.

Nous espérons que vous aurez du temps pour explorer Montréal et découvrir ainsi la joie de vivre des Montréalaises et des Montréalais. Que vous flâniez dans les bars de la rue Crescent au centre-ville, près des établissements de nos partenaires hôteliers, que vous découvriez le charme du Vieux-Montréal (via la station de métro Champ-de-Mars) ou que vous fassiez l'expérience de la culture française sur les terrasses du Quartier latin (via la station Berri-UQAM), vous êtes assurés de passer du bon temps. Si la nature vous attire, le campus est sur le versant nord du Mont Royal, l'un des plus beaux parcs urbains qui soit ; il vaut le déplacement (autobus n<sup>o</sup> 11 à partir de la station de métro Mont-Royal). Vous aurez la chance de le découvrir en partie lors du banquet qui aura lieu au Chalet du Mont-Royal. Finalement, les amateurs d'art apprécieront une exposition fort intéressante présentée par le CRM et mettant en vedette la science (voir la description ci-dessous).

# Sponsors • Commanditaires

SSC 2004 thanks the following sponsors for their extremely generous contributions for the success of this meeting:

SSC 2004 remercie les commanditaires suivants pour leurs très généreuses contributions qui ont mené au succès du congrès :

Centre de recherches mathématiques, the Fields Institute, the Pacific Institute for the Mathematical Sciences, MITACS, Hydro-Québec, Université de Montréal, Pfizer, Laboratoires universitaires Bell University Laboratories & Institut de la statistique du Québec.



[www.crm.umontreal.ca](http://www.crm.umontreal.ca)



[www.fields.utoronto.ca](http://www.fields.utoronto.ca)



[www.mitacs.math.ca](http://www.mitacs.math.ca)



[www.pims.math.ca](http://www.pims.math.ca)

présenté en collaboration avec



[www.pfizer.ca](http://www.pfizer.ca)



[www.umontreal.ca](http://www.umontreal.ca)

Laboratoires universitaires Bell  
Bell University Laboratories

[www.bce.ca](http://www.bce.ca)



[www.stat.gouv.qc.ca](http://www.stat.gouv.qc.ca)

# Organizers • Organisateurs

## Local Arrangements Committee • Comité organisateur local

**Chair • Président:** Christian Léger

The Chair would like to express his sincerest thanks to his colleagues Jean-François Angers, Pierre Duchesne, and Denis Larocque for their precious and generous help in various ways in the last few months, as well as the support received from Martin Bilodeau, Yves Lepage, François Perron, and Roch Roy. Thanks also go to Delija Geca, Pina D'Angelo, and Louise Bourque whose help has secured some of our sponsorships. But mostly, thanks go to all of CRM's staff who have invested numerous hours in the preparation of this meeting under the able leadership of Louis Pelletier, in particular Jasmin Lapalme, Louise Letendre, Josée Laferrière, André Montpetit, Daniel Ouimet, and Suzette Paradis. Finally, many graduate students have been very helpful with many little tasks.

Le président souhaiterait offrir ses plus sincères remerciements à ses collègues Jean-François Angers, Pierre Duchesne et Denis Larocque pour leur généreuse aide sous plusieurs formes au cours des derniers mois. Il veut également souligner l'appui qu'il a reçu de Martin Bilodeau, Yves Lepage, François Perron et Roch Roy. Il est également important de remercier chaleureusement Delija Geca, Pina D'Angelo et Louise Bourque sans qui certaines des commandites ne se seraient pas matérialisées. Mais ses remerciements les plus importants vont au personnel du CRM qui, sous la supervision efficace de Louis Pelletier, a investi un nombre incalculable d'heures dans la préparation de ce congrès, en particulier Jasmin Lapalme, Louise Letendre, Josée Laferrière, André Montpetit, Daniel Ouimet et Suzette Paradis. Finalement, plusieurs étudiants des cycles supérieurs ont été fort utiles en accomplissant plusieurs petites tâches.

## Program Committee • Comité du programme

**Chair • Président:** Christian Genest, *Université Laval*

**Members • Membres:**

### **Biostatistics Section • Groupe de biostatistique**

Salomon Minkin, *Ontario Cancer Institute*, Gordon Fick, *University of Calgary*

### **Business and Industrial Statistics Section • Groupe de statistique industrielle et de gestion**

Fernando Camacho, *Damos*

### **Survey Methods Section • Groupe de méthodologie d'enquête**

Georgia Roberts, *Statistics Canada/Statistique Canada*, Patricia Whitridge, *Royal Canadian Mounted Police/Gendarmerie royale du Canada*

## Translation • Traduction

Louis-François Poirier, a graduate student at Université de Montréal, has translated most abstracts under the supervision of Pierre Duchesne (Université de Montréal). Christian Léger has also participated.

La traduction de la plupart des résumés a été réalisée par Louis-François Poirier, étudiant de deuxième cycle à l'Université de Montréal. Pierre Duchesne (Université de Montréal) a supervisé le travail. Christian Léger y a également participé.

## Web Page Maintenance • Gestion de page web

The Christian(s) wish to thank Peter Macdonald (McMaster University) and Jamie Macdonald for their great help and patience in keeping the meeting's web site up to date.

Les "Christian(s)" souhaitent remercier Peter Macdonald (McMaster University) et Jamie Macdonald pour leur aide précieuse et la patience dont ils ont fait preuve dans l'entretien du site web du congrès.

## Exhibitors • Exposants

The book displays are located in room AA-1221 of the Pavillon André-Aisenstadt (building 15 of the map on the back cover) and will be available for viewing and purchasing from 10:30 to 17:00 Monday and Tuesday and from 10:00 to 12:00 on Wednesday.

Le salon des exposants est situé dans le local AA-1221 du Pavillon André-Aisenstadt (édifice 15 sur la carte de la quatrième de couverture); consultation et achats possibles entre 10h00 à 17h00 les lundi et mardi, ainsi que de 10h00 à 12h00 le mercredi.

## General Information • Information générale

The Registration Desk will be open as follows in the Hall of the Pavillon André-Aisenstadt:

L'inscription aura lieu dans le Hall du Pavillon André-Aisenstadt aux heures suivantes :

Date	Time/Heures
Sunday, May 30 • dimanche 30 mai	8:00–18:30
Monday, May 31 • lundi 31 mai	8:00–17:00
Tuesday, June 1 • mardi 1 <sup>er</sup> juin	8:30–17:00
Wednesday, June 2 • mercredi 2 juin	8:30–12:00

If you need further help, you may want to check with the CRM Secretariat at AA-5357.

Pour obtenir de l'aide supplémentaire, vous pouvez contacter le secrétariat du CRM au local AA-5357.

### Directions

From downtown Montréal to the Université de Montréal campus, you can take a cab (about \$15.00) or the Métro for \$2.50 a ticket. Exit the Métro at the Université de Montréal Station (blue line). Then take the moving ramp to the campus, located at the top of the hill. For those who will stay in the conference hotels, it is recommended to get to the campus by Bus 165 or 535 (these buses ride on Guy street which turns into Côte-des-Neiges). By exiting at the intersection of "Queen Mary Road" and "Côte-des-Neiges," a nice five-minute walk will take you to campus. The building "Pavillon André-Aisenstadt" is located west of

the main building “Pavillon Roger-Gaudry” (formerly called "Pavillon Principal"). Finally, there is a lot of construction on campus and parking is somewhat limited and therefore not recommended. Nevertheless, it is possible to park for \$10 per day in the Garage Louis-Collin (building 17 on the map). You can access the campus from either the entrance at the corner of Chemin Queen-Mary Road and Decelles or that on Louis-Collin. Parking in the surrounding streets is not recommended.

Du centre-ville de Montréal au campus de l'Université de Montréal, vous pouvez prendre un taxi (environ 15 \$) ou le métro (2,50 \$) jusqu'à la station « Université de Montréal » (ligne bleue). À la sortie du métro, empruntez la rampe mobile qui vous mènera au sommet de la colline. Pour ceux et celles logeant chez l'un ou l'autre des partenaires hôteliers, le transport par autobus est recommandé ; prendre les circuits n° 165 ou 535, qui empruntent la rue Guy, laquelle devient Côte-des-Neiges. Descendre à l'intersection des rues « Chemin Queen Mary » et « Chemin Côte-des-Neiges », située à environ 5 minutes de marche du campus. Le Pavillon André-Aisenstadt est situé à l'ouest du Pavillon Roger-Gaudry (anciennement le Pavillon principal). Finalement, il y a beaucoup de construction sur le campus ce qui limite considérablement le stationnement. Bien que ce ne soit pas recommandé, il est toutefois possible de stationner dans le garage Louis-Collin (édifice 17 sur la carte) pour 10 \$ par jour. Vous pouvez accéder au campus via l'entrée au coin du Chemin Queen-Mary et avenue Decelles ou via l'entrée de la rue Louis-Collin. Le stationnement dans les rues environnantes n'est définitivement pas recommandé.

## Rooms • Salles

The rooms used for the conference are in three buildings: Pavillon André-Aisenstadt (AA; number 15 on the map), Pavillon Claire-McNicoll (Z; number 11a on the map) and Pavillon 3200, rue Jean-Brillant (B; number 21 on the map). In all buildings, the first number indicates the floor.

Les locaux utilisés pour le congrès se retrouvent dans trois édifices, soit le Pavillon André-Aisenstadt (AA ; numéro 15 sur la carte), le Pavillon Claire-McNicoll (Z ; numéro 11) et le Pavillon 3200, rue Jean-Brillant (B ; numéro 21). Le premier numéro d'un local indique l'étage et ce dans tous les pavillons.

## Workshops • Ateliers

The workshops organized by the three sections will take place in the Pavillon André-Aisenstadt (building 15 on the back cover) on Sunday. Note that the Survey Methods Section workshop starts at 8:30 while the other two start at 9:00. See the scientific programme for further details.

Les ateliers organisés par les trois groupes auront tous lieu le dimanche dans le Pavillon André-Aisenstadt (édifice 15 en quatrième de couverture). Notez bien que l'atelier du groupe de méthodologie d'enquête débutera à **8h30** alors que les deux autres débuteront à **9h00**. Voir le programme scientifique pour plus de détails.

## Cocktail and Poster Session • Coquetel et séance d'affichage

The opening cocktail will take place in the *Hall d'honneur* of the Pavillon Roger-Gaudry (formerly known as Pavillon Principal) from 18:30 to 22:00 on Sunday May 30 immediately following the Late Breaking Session. The Pavillon Roger-Gaudry is building 11 on the map and the entrance to the *Hall d'honneur* is through the doors L-4 (level 4); see the smaller map at the bottom right-hand side of the back cover. Hors d'oeuvres and a cash bar will be available. Two tickets are included in your registration package. Soft drinks and juices cost one ticket while beer and wine cost two tickets. All conference attendees and

companions are warmly invited to come and meet old friends and make new acquaintances. Moreover, you will be able to interact with those presenting posters at the same time and the same place.

Le coquetel d'accueil aura lieu dans le Hall d'honneur du Pavillon Roger-Gaudry (l'ancien pavillon principal) de 18h30 à 22h00 le dimanche 30 mai immédiatement après la séance « dernières nouvelles ». Le Pavillon Roger-Gaudry est l'édifice 11 sur la carte et l'accès au Hall d'honneur se fait par l'entrée au 4<sup>e</sup> étage via les portes L-4 (voir la petite carte dans le coin inférieur droit de la quatrième de couverture). Des hors-d'œuvre seront servis et un bar payant sera disponible. Votre trousse d'inscription contient deux billets pour boissons. Les boissons gazeuses et les jus nécessiteront un seul coupon alors que deux coupons vous permettront d'obtenir de la bière ou du vin. Les participants au congrès et leurs compagnons sont tous invités à venir retrouver de vieux amis et faire de nouvelles connaissances. De plus, il vous sera possible d'interagir avec les personnes qui feront une présentation par affichage qui aura lieu au même endroit et à la même période.

## Food Services on Campus • Restauration sur le campus

The recommended option on campus is the cafeteria in Pavillon 3200, rue Jean-Brillant (number 21 on the map) in room B-2246. All morning and lunch committee meetings from Monday through Wednesday will be held there. Service begins at 7:30 in the morning for breakfast (but the room itself opens earlier). Lunch is served between 11:30 and 13:30. The cafeteria is not open for dinner.

La seule véritable option sur le campus est la cafétéria du Pavillon 3200, rue Jean-Brillant (numéro 21 sur la carte) dans le local B-2246. Toutes les rencontres de comités qui auront lieu lors du déjeuner ou du lunch entre les lundi et mercredi auront lieu dans la cafétéria. Notez que le service pour le déjeuner ne débute qu'à 7h30, bien que le local ouvre avant cela. Le dîner est servi entre 11h30 et 13h30 alors qu'il n'y a pas de service pour le souper.

## Restaurant off but close to Campus • Restaurants près du campus

Montréal is well known for its restaurants. You will find plenty of restaurant guides on the web or in your hotel for the downtown area. The following list includes only restaurants in the vicinity of the campus.

Montréal est reconnu pour ses restaurants. Pour vos sorties au centre-ville, le web ou votre hôtel sauront vous renseigner adéquatement. Les restaurants de la liste qui suit sont tous à proximité du campus.

### Lacombe (Gatineau–Côte-des-Neiges)

**Il Galateo** (5315 Gatineau) Italian cuisine • cuisine italienne

**La Maisonnée** (Lacombe & Gatineau) Brasserie

**Chez Lien** (3527 Lacombe) Vietnamese cuisine • cuisine vietnamienne

**Le Figuier** (3515 Lacombe) Couscous

**Pizzafiore** (3518 Lacombe) pizzas, pastas • pâtes, pizzas

**Le Crocodile** (Lacombe & Gatineau) Sandwiches, salads • sandwiches, salades

### Gatineau (Lacombe–Queen Mary)

**Kalimera** (5188 Gatineau) Greek cuisine • cuisine grecque

### Queen Mary (Côte-des-Neiges–Decelles)

**Raga** (3533 Queen Mary) Indian cuisine (Buffet) • buffet indien

**Second Cup** (3333 Queen Mary) Sandwiches, salads and good coffee • sandwiches, salades, cafés

### Côte-des-Neiges (Queen Mary–Côte-Ste-Catherine)

**La Fiesta** (5179 Côte-des-Neiges) Mexican cuisine • cuisine mexicaine



**Monsieur Falafel** (5171 Côte-des-Neiges) Lebanese cuisine • cuisine libanaise

**Rôtisserie St-Hubert** (5235 Côte-des-Neiges) Bar-B-Q

**Nickels** (5252 Côte-des-Neiges) Delicatessen • sandwichs à la viande fumée

**La Brûlerie** (5252 Côte-des-Neiges) Sandwiches, croissants and good coffee • sandwichs, croissants, café

**Pacini** (5520 Côte-des-Neiges) Italian cuisine • cuisine italienne

**Atami** (5499 Côte-des-Neiges) Sushis

**Le Béarn** (5613 Côte-des-Neiges) French cuisine • cuisine française

### **Côte-des-Neiges & Queen Mary**

**Le Commensal** (3715 Queen Mary) Vegetarian cuisine • cuisine végétarienne

**Subway** (5185 Côte-des-Neiges) Fast food • cuisine rapide

**Pizzédélic** (515 3 Côte-des-Neiges) Pizzas

**Masako** (5050 Côte-des-Neiges) Sushis

### **Côte-des-Neiges & Côte-Ste-Catherine**

**Café la République** (5693 Côte-des-Neiges), Sandwiches, salads, good coffee • sandwichs, salades, cafés

**Taqueria Sol y Luna** (5701 Côte-des-Neiges) Mexican cuisine • cuisine mexicaine

## **Exposition**

Science and art don't often meet, but when they do, the results can be spectacular! As part of the 125th anniversary of Université de Montréal, CRM is proud to present an Exposition entitled: "Visions from the Quantum Frontier" at the Centre d'exposition de l'Université de Montréal in the Pavillon de la Faculté d'aménagement, 2940, chemin de la Côte-Sainte-Catherine, room 0056 (building 29 on the map). You may visit the exposition on Sunday May 30 and Tuesday June 1st from 12:00 to 18:00.

This exposition presents the works of Eric J. Heller, a physicist from Harvard University, who studies the quantum world of molecules, atoms, and electrons through numerical simulations on computers. In the mid 80s, he imagined created works of art from graphs created from these simulations. His works, which evoke the mysteries of the quantum world, quickly gained wide dissemination through exposition centres as well as in the press.

Il est plutôt rare que la science et l'art décident de se rencontrer, mais lorsque c'est le cas, les résultats peuvent être assez surprenants ! Dans le cadre des activités du 125<sup>e</sup> anniversaire de l'Université de Montréal, c'est avec beaucoup de plaisir que le CRM présente l'exposition « Aux confins de l'univers quantique » au Centre d'exposition de l'Université de Montréal dans le Pavillon de la Faculté d'aménagement au 2940, chemin de la Côte-Sainte-Catherine, local 0056 (édifice 29 sur la carte). Vous pourrez visiter l'exposition le dimanche 30 mai ainsi que le mardi 1er juin de 12h00 à 18h00.

Cette exposition présente les œuvres du physicien Eric J. Heller de Harvard, qui étudie l'univers quantique des molécules, des atomes et des électrons en effectuant des simulations numériques sur ordinateur. Au milieu des années 80, il a imaginé de créer des œuvres d'art à partir de graphiques tirés de ces simulations. Ses œuvres, qui évoquent les mystères de l'univers quantique, ont vite connu une très vaste diffusion tant dans les centres d'exposition que dans la presse.

## **Barbecue**

The Annual Barbecue will be held in Place Laurentienne on campus (the outside area between buildings 19 to 21). We will also have access to the cafeteria if the weather does not cooperate. All students will

have received a barbecue ticket with their registration. All other participants who paid for this option when they registered will also have a ticket. The meal will feature various sausages, exquisite salads, and a mousse. There will also be a cash bar.

Le Barbecue annuel aura lieu sur la Place Laurentienne au coeur du campus (la place extérieure entre les édifices 19 à 21). Nous aurons également accès à la cafétéria si le temps n'est pas clément. Tous les étudiants trouveront un billet pour le barbecue dans leur trousse d'admission. Tous les autres participants qui ont payé pour cette option lorsqu'ils se sont inscrits, ont également un billet dans leur trousse. Le menu consiste en des saucisses variées, des salades élaborées, ainsi qu'une mousse. Un bar payant sera disponible sur place.

## **Women in Statistics Reception • Réception pour les femmes en statistique**

The Caucus for Women in Statistics in collaboration with the Committee on Women in Statistics of the SSC will host an informal reception on Sunday, May 30 from 17:00 to 18:00 in room AA-1221 of the Pavillon André-Aisenstadt. Interested women and men are welcome.

The Committee on Women in Statistics invites all members of the Caucus for Women in Statistics to join the executive meeting to discuss Caucus business during the conference. The meeting will take place on Tuesday June 1 at 12:15 in the cafeteria (room B-2246 of the Pavillon Jean-Brillant).

Le Caucus canadien des statisticiennes, en collaboration avec le Comité sur les femmes en statistique de la SSC, organise une réception informelle dimanche 30 mai de 17h00 à 18h00 au local AA-1221 du Pavillon André-Aisenstadt. Toutes les femmes — et tous les hommes — intéressé(e)s sont les bienvenu(e)s.

Le Comité de la SSC pour la promotion de la femme en statistique invite tous les membres du Caucus canadien des statisticiennes à participer à la réunion de leur exécutif pour discuter des affaires du Caucus durant le congrès. La réunion aura lieu le mardi 1<sup>er</sup> juin à 12h15 dans la cafétéria (local B-2246 du Pavillon Jean-Brillant).

## **Waterloo Statistics Alumni Reception • Réception des anciens diplômés en statistique de Waterloo**

University of Waterloo Math alumni attending the 2004 SSC Meeting in Montréal are warmly invited to a reception on Monday, May 31. Grads can re-connect and enjoy light refreshments (cash bar) at Le Commensal, 3715 Chemin Queen Mary from 16:30–18:00 as you mingle with alumni and UW faculty. Le Commensal is on the upper right-end corner of the map on the back cover.

Les anciens étudiants en mathématiques de l'Université de Waterloo qui participent au Congrès 2004 de la SSC sont invités à une réception le lundi 31 mai. Les diplômés peuvent se rappeler de bons souvenirs et profiter d'un repas léger ainsi que d'un bar payant au restaurant Le Commensal, 3715 Chemin Queen Mary de 16h30 à 18h00. Venez retrouver d'autres anciens étudiants et des professeurs de l'Université de Waterloo. Le Commensal est situé dans le coin supérieur droit de la carte en quatrième de couverture.

## **Banquet**

All participants are cordially invited to attend the Conference Banquet on Tuesday evening, June 1. Banquet tickets have been included in your registration material and are colour coded to indicate your food preference. The Banquet will be held at the Chalet du Mont-Royal, located on a natural promontory dominating a large belvedere affording a breathtaking view of the downtown district.

The Chalet is in the Parc du Mont-Royal. To get there, buses will take participants from the Pavillon André-Aisenstadt at 18:15 and at 18:45 to the parking of the park. From that point, the Chalet is within a very nice five minute walk in the park; a small bus will help those with reduced mobility. It is also possible to drive to the park; don't forget to pay for the parking. If the weather is nice and you are in good shape, it is also possible to walk from the campus to the park; it could take 45 minutes.

A (non-alcoholic) cocktail will be served on the steps of the Chalet and a cash bar will be available inside. Note that alcoholic beverages will not be allowed outside the Chalet. A string quartet will entertain the patrons during the evening. Buses will take you back from the parking of the park to the downtown hotels or to the residences of Université de Montréal. Departures will take place at 22:30 and 23:00.

Tous les participants sont cordialement invités à prendre part au Banquet du congrès qui aura lieu le mardi 1er juin en soirée. Vous trouverez vos billets pour le banquet dans votre trousse d'inscription. La couleur de votre billet indique votre préférence de menu. Le Banquet aura lieu au Chalet du Mont-Royal, situé sur un promontoire naturel dominant un immense belvédère offrant une vue imprenable sur le centre-ville.

Le Chalet est situé dans le Parc du Mont-Royal. Les participants pourront s'y rendre en prenant les autobus qui quitteront le Pavillon André-Aisenstadt à 18h15 et à 18h45. Arrivé au stationnement du parc, une agréable petite marche de cinq minutes vous mènera au Chalet. Pour les gens à mobilité restreinte, un petit autobus fera la navette entre le stationnement et le Chalet. Pour les gens en forme, si le beau temps est au rendez-vous, il est possible de marcher jusqu'au Chalet, mais ça pourrait prendre 45 minutes.

Une boisson non alcoolisée vous sera servie à l'entrée du Chalet alors qu'un bar payant sera disponible à l'intérieur. Notez qu'il ne sera pas permis de sortir du Chalet avec des boissons alcoolisées. Un quatuor à cordes égayera la soirée. Le retour se fera à partir du stationnement du parc, là où les autobus attendront pour vous ramener soit aux hôtels du centre-ville, soit aux résidences de l'Université de Montréal. Les départs auront lieu à 22h30 et à 23h00.

## **Job Fair • Salon de l'emploi**

Interviews for those participating in the Job Fair will be conducted in various rooms located in the Pavillon André-Aisenstadt (building 15 on the map). Check with the Job Fair booth near the registration area for schedules and locations for interviews.

Les entrevues pour le Salon de l'emploi se tiendront dans différents locaux du Pavillon André-Aisenstadt (édifice 15 sur la carte). Nous vous invitons à consulter le kiosque du Salon de l'emploi près de l'aire d'inscription pour l'horaire et le lieu des entretiens.

## **Internet Access • Accès à l'internet**

Your conference bag contains a page with a userid, a password, and instructions for computers and terminals located in rooms AA-1340 and AA-4191 of the Pavillon André-Aisenstadt. Note that it will not be possible to print.

Votre trousse d'inscription contient une page avec un numéro d'usager, un mot de passe, ainsi que des instructions pour accéder aux ordinateurs et aux terminaux des locaux AA-1340 et AA-4191 du Pavillon André-Aisenstadt. Notez qu'il ne sera pas possible d'imprimer.

# Committees and Meetings

## Comités et réunions

### Calendar Listing • Calendrier journalier

#### Saturday May 29 • Samedi 29 mai

18:00–22:00 Executive Committee • Comité exécutif Hôtel Versailles

#### Sunday, May 30 • Dimanche 30 mai

9:00–11:00 Finance Committee • Comité des finances AA-6245  
 9:00–13:00 Statistics Chairs • Directeurs de statistique AA-5340  
 11:00–12:00 Publications Committee • Comité des publications AA-6245  
 12:00–16:00 Board of Directors • Conseil d'administration AA-6245

#### Monday, May 31 • Lundi 31 mai

7:15–8:15 Biostatistics Section (Section Officers) • Groupe de biostatistique (directeurs du groupe) B-2246  
 Business and Industrial Statistics Section (Section Officers) • Groupe de statistique industrielle et de gestion (directeurs du groupe) B-2246  
 Survey Methods Section (Section Officers) • Groupe des méthodes d'enquête (directeurs du groupe) B-2246  
 12:15–13:15 Accreditation Appeals Committee • Comité des appels d'accréditation B-2246  
 CJS Editorial Board • Comité de rédaction de la RCS B-2246  
 Public Relations Committee • Comité des relations publiques B-2246  
 17:00–1800 Biostatistics Section (General Meeting) • Groupe de biostatistique (assemblée générale) AA-1140  
 Business and Industrial Statistics Section (General Meeting) • Groupe de statistique industrielle et de gestion (assemblée générale) AA-1355  
 Survey Methods Section (General Meeting) • Groupe des méthodes d'enquête (assemblée générale) AA-1360  
 Assemblée générale de l'Association des statisticiennes et statisticiens du Québec Z-317  
 18:00–18:30 Business and Industrial Statistics Section (Section Officers) • Groupe de statistique industrielle et de gestion (directeurs du groupe) AA-1355

#### Tuesday, June 1 • Mardi 1<sup>er</sup> juin

7:15–8:15 Research Committee • Comité de la recherche B-2246  
 Accreditation Committee • Comité sur l'accréditation B-2246  
 12:15–13:15 NPCDS • PNSDC B-2246  
 Bilingualism Committee • Comité du bilinguisme B-2246  
 Committee on Women in Statistics • Comité sur les femmes en statistique B-2246  
 Statistical Education Committee • Comité d'éducation en statistique B-2246

17:00–18:30 SSC General Meeting • Assemblée générale de la SSC AA-1360

### **Wednesday, June 2 • Mercredi 2 juin**

7:15–8:15 Professional Development Committee • Comité sur le perfectionnement professionnel B-2246  
 12:15–13:15 Program Committee • Comité du programme B-2246  
 17:30–19:30 Board of Directors • Conseil d'administration AA-6245  
 19:30 Executive Committee • Comité exécutif AA-6245

## **Alphabetical Listing • liste alphabétique**

### **Accreditation Committee • Comité sur l'accréditation**

Tuesday, June 1 • mardi 1<sup>er</sup> juin, 7:15–8:15, B-2246

### **Accreditation Appeals Committee • Comité des appels d'accréditation**

Monday, May 31 • lundi 31 mai, 12:15–13:15, B-2246

### **Assemblée générale de l'Association des statisticiennes et statisticiens du Québec**

Monday, May 31 • lundi 31 mai, 17:00–18:00, Z-317

### **Bilingualism Committee • Comité du bilinguisme**

Tuesday, June 1 • mardi 1<sup>er</sup> juin, 12:15–13:15, B-2246

### **Biostatistics Section (General Meeting) • Groupe de biostatistique (assemblée générale)**

Monday, May 31 • lundi 31 mai, 17:00–18:00, AA-1140

### **Biostatistics Section (Section Officers) • Groupe de biostatistique (directeurs du groupe)**

Monday, May 31 • lundi 31 mai, 7:15–8:15, B-2246

### **Board of Directors • Conseil d'administration**

Sunday, May 30 • dimanche 30 mai, 12:00–16:00, AA-6245

Wednesday June 2 • mercredi 2 juin, 17:30–19:30, AA-6245

### **Business and Industrial Statistics Section (General Meeting) • Groupe de statistique industrielle et de gestion (assemblée générale)**

Monday, May 31 • lundi 31 mai, 17:00–18:00, AA-1355

### **Business and Industrial Statistics Section (Section Officers) • Groupe de statistique industrielle et de gestion (directeurs du groupe)**

Monday, May 31 • lundi 31 mai, 7:15–8:15, B-2246

Monday, May 31 • lundi 31 mai, 18:00–18:30, AA-1355

### **CJS Editorial Board • Comité de rédaction de la RCS**

Monday, May 31 • lundi 31 mai, 12:15–13:15, B-2246

### **Committee on Women in Statistics • Comité sur les femmes en statistique**

Tuesday, June 1 • mardi 1<sup>er</sup> juin, 12:15–13:15, B-2246

### **Executive Committee • Comité exécutif**

Saturday May 29 • samedi 29 mai, 18:30–22:00, Hôtel Versailles

Wednesday June 2 • mercredi 2 juin, 19:30, AA-6245

### **Finance Committee • Comité des finances**

Sunday, May 30 • dimanche 30 mai, 9:00–11:00, AA-6245

### **NPCDS • PNSDC**

Tuesday, June 1 • mardi 1<sup>er</sup> juin, 12:15–13:15, B-2246

**Professional Development Committee • Comité sur le perfectionnement professionnel**

Wednesday June 2 • mercredi 2 juin, 7:15–8:15, B-2246

**Program Committee • Comité du programme**

Wednesday June 2 • mercredi 2 juin, 12:15–13:15, B-2246

**Public Relations Committee • Comité des relations publiques**

Monday, May 31 • lundi 31 mai, 12:15–13:15, B-2246

**Publications Committee • Comité des publications**

Sunday, May 30 • dimanche 30 mai, 11:00–12:00, AA-6245

**Research Committee • Comité de la recherche**Tuesday, June 1 • mardi 1<sup>er</sup> juin, 7:15–8:15, B-2246**SSC General Meeting • Assemblée générale de la SSC**Tuesday, June 1 • mardi 1<sup>er</sup> juin, 17:00–18:30, AA-1360**Statistical Education Committee • Comité d'éducation en statistique**Tuesday, June 1 • mardi 1<sup>er</sup> juin, 12:15–13:15, B-2246**Statistics Chairs • Directeurs de statistique**

Sunday, May 30 • dimanche 30 mai, 9:00–13:00, AA-5340

**Survey Methods Section (General Meeting) • Groupe des méthodes d'enquête (assemblée générale)**

Monday, May 31 • lundi 31 mai, 17:00–18:00, AA-1360

**Survey Methods Section (Section Officers) • Groupe des méthodes d'enquête (directeurs du groupe)**

Monday, May 31 • lundi 31 mai, 7:15–8:15, B-2246

# List of Sessions • Liste des sessions

## Sunday, May 30 • Dimanche 30 mai

### 8:30–17:00

AA-1360 Workshop of the Survey Methods Section • Atelier du Groupe de méthodologie d'enquête

### 9:00–17:00

AA-1140 Workshop of the Biostatistics Section • Atelier du Groupe de biostatistique

AA-1355 Workshop of the Business and Industrial Statistics Section • Atelier du Groupe de statistique industrielle et de gestion

### 17:30–18:30

AA-1140 Session 0A Accreditation of Professional Statisticians in Canada  
L'accréditation des statisticiens professionnels au Canada

### 18:30–22:00

Hall Session 0B Contributed Poster Session and Cocktail (evening)  
d'honneur Présentation d'affiches et coquetel (soirée)

## Monday, May 31 • Lundi 31 mai

### 8:30–10:00

B-2285 Session 1 Inaugural Session and SSC Presidential Invited Address  
Séance inaugurale et allocution de l'invitée de la présidente

### 10:30–12:00

Z-310 Session 2A Gender-Related Statistics in Education  
Statistiques reliées au genre dans le domaine de l'éducation

Z-330 Session 2B The Escalation of Non-Response and its Impact on Surveys  
La croissance de la non-réponse et ses impacts sur les enquêtes

AA-1140 Session 2C Current Trends in Statistical Genetics  
Tendances actuelles en statistique génétique

AA-1360 Session 2D Case Study I: Asthma  
Étude de cas I : Asthme

AA-1355 Session 2E Finance

Z-317 Session 2F GEE and Related Methods  
EGE et méthodes connexes

### 13:30–15:00

Z-310 Session 3A Special Invited Session of the Survey Methods Section  
Allocution de l'invité d'honneur du Groupe de méthodologie d'enquête

Z-330 Session 3B Application of Statistical Methods to Actuarial Problems  
Application de méthodes statistiques à des problèmes actuariels

AA-1140 Session 3C Biostatistics Applications: The Dynamics of Collaboration  
Applications biostatistiques : La dynamique de la collaboration

- AA-1360 Session 3D Capture-Recapture Models  
Modèles de capture-recapture
- AA-1355 Session 3E Time Series Analysis  
Analyse de séries chronologiques
- Z-317 Session 3F Survival Analysis I  
Analyse de durées de vie I

**15:30–17:00**

- AA-1140 Session 4A In Honour of Jim Zidek, on his 65th Birthday  
En hommage à Jim Zidek, pour ses 65 ans
- Z-310 Session 4B Causal Inference in Biostatistics and Epidemiology  
Inférence causale en biostatistique et en épidémiologie
- AA-1360 Session 4C New Methodologies for Census 2006  
Nouvelles méthodologies pour le recensement de 2006
- Z-330 Session 4D Applied Probability  
Probabilité appliquée
- Z-317 Session 4E Regression  
Régression
- AA-1355 Session 4F Statistical Genetics  
Statistique génétique

**17:00–18:30**

- Z-330 Session 4G NSERC Workshop: Tips to Prepare your Next Discovery Grant Application  
Atelier du CRSNG : Conseils pour l'élaboration de votre prochaine demande de subvention à la découverte

**Tuesday, June 1 • Mardi 1er juin****8:30–10:00**

- B-2285 Session 5 Gold Medal Address  
Allocution de la récipiendaire de la médaille d'or

**10:30–12:00**

- AA-1140 Session 6A Special Invited Session of the Biostatistics Section  
Allocution de l'invité d'honneur du Groupe de biostatistique
- AA-1360 Session 6B Analysis on Complex Surveys  
L'analyse de données d'enquêtes complexes
- Z-310 Session 6C Applications of Wavelets to Statistical Problems  
Applications des ondelettes à des problèmes statistiques
- Z-330 Session 6D Monte Carlo Methods: Theory and Applications  
Méthodes de Monte-Carlo : Théorie et applications
- Z-317 Session 6E Survey Sampling I  
Méthodologie d'enquête I
- AA-1355 Session 6F Biostatistics I  
Biostatistique I



**13:30–15:00**

AA-1140	Session 7A	Isobel Loutit Invited Address on Business and Industrial Statistics Allocution Isobel-Loutit en statistique industrielle et en gestion
Z-310	Session 7B	Shape Analysis of 3D Brain Mapping Data Analyse de forme de données de cartographie cérébrale 3D
Z-330	Session 7C	Multi-level Modelling Modélisation multi-niveau
AA-1360	Session 7D	Nonlinear Time Series Chroniques non linéaires
AA-1355	Session 7E	Survey Sampling II Méthodologie d'enquête II
Z-317	Session 7F	Environmental Statistics Statistique environnementale

**15:30–17:00**

Z-310	Session 8A	The Canadian Initiative in Social Statistics and Cross-Disciplinary Collaboration L'initiative canadienne en matière de statistiques sociales et la collaboration inter-disciplinaire
AA-1140	Session 8B	Statistical Analysis of Human Growth Data Analyse statistique de données sur la croissance humaine
AA-1355	Session 8C	Empirical Likelihood Methodology Méthodes empirico-vraisemblantistes
Z-330	Session 8D	Teaching Statistics to Non-statisticians: Spreading the Good Word to Achieve Recognition for our Profession L'enseignement de la statistique à des non-statisticiens : Prêcher la bonne parole pour faire reconnaître notre profession
Z-220	Session 8E	Survey Sampling III Méthodologie d'enquête III
Z-317	Session 8F	Design and Analysis of Health Studies Planification et analyse d'études sanitaires

**Wednesday, June 2 • Mercredi 2 juin****8:30–10:00**

AA-1140	Session 9A1	Pierre Robillard Award Address (8:30–9:15) Présentation de la lauréate du Prix Pierre-Robillard
AA-1140	Session 9A2	<i>Canadian Journal of Statistics</i> Award Address (9:15–10:00) Allocution des récipiendaires du Prix de <i>La revue canadienne de statistique</i>
AA-1360	Session 9B	Statistics and Science La statistique et la science
Z-310	Session 9C	New Researchers and New Directions for Research in Biostatistics Nouveaux chercheurs et nouvelles perspectives de recherche en biostatistique

- Z-330      Session 9D      Non-traditional Survey Methods  
Méthodes d'enquête non traditionnelles
- AA-1355    Session 9E      Stochastic Processes and Their Applications  
Les processus stochastiques et leurs applications
- Z-317      Session 9F      Survival Analysis II  
Analyse de durées de vie II

**10:30–12:00**

- AA-1360    Session 10A    Statistical Hydrology  
Hydrologie statistique
- AA-1140    Session 10B    Empirical Likelihood Methods in Survey Sampling  
Méthodes empirico-vraisemblantistes en méthodologie d'enquête
- Z-310      Session 10C    Parametric Survival Analysis  
Analyse paramétrique de durées de vie
- Z-330      Session 10D    Case Study II: Anorexia Nervosa  
Étude de cas II : Anorexie mentale
- AA-1355    Session 10E    Statistical Inference  
L'inférence statistique
- Z-317      Session 10F    Biostatistics II  
Biostatistique II

**13:30–15:00**

- Z-310      Session 11A    Health Surveys  
Les enquêtes dans le domaine de la santé
- AA-1360    Session 11B    Non-IID Sampling Data  
Données d'échantillonnage non iid
- AA-1140    Session 11C    Analysis of Longitudinal Data  
Analyse de données longitudinales
- Z-330      Session 11D    Statistics and the Physical Sciences  
La statistique et les sciences physiques
- Z-317      Session 11E    Testing Procedures  
Procédures de test
- AA-1355    Session 11F    Epidemiology  
Épidémiologie
- Z-240      Session 11G    Panel Discussion on the Mathematical Sciences Institutes and Programs:  
As a Statistician, Should I be Concerned?  
Débat sur les instituts et les programmes en sciences mathématiques : En tant que statisticien, suis-je concerné ?

**15:30–17:00**

- AA-1140    Session 12A    Address of the 2004 CRM–SSC Prize Winner  
Allocution du lauréat du Prix CRM-SSC 2004
- Z-330      Session 12B    Nonparametric and Bayesian Methods  
Méthodologie bayésienne et non paramétrique
- AA-1360    Session 12C    Estimation

Z-317	Session 12D	Experimental Design and Related Issues Planification d'expérience et sujets connexes
AA-1355	Session 12E	Various Topics Thèmes divers
Z-310	Session 12F	Biostatistics III Biostatistique III

# Program • Programme

**Sunday, May 30 • Dimanche 30 mai**

---

**8:30–17:00 Workshop of the Survey Methods Section** **AA-1360**  
**Atelier du Groupe de méthodologie d'enquête**

An Introduction to Best Survey Practices  
 Une introduction aux meilleures pratiques d'enquête

Leaders • Animateurs : Éric Rancourt, Allen Gower, John Kovar, Jack Gambino, Jean Dumais & Jean-Louis Tambay, *Statistics Canada/Statistique Canada*

**9:00–17:00 Workshop of the Biostatistics Section** **AA-1140**  
**Atelier du Groupe de biostatistique**

Statistical Learning and Data Mining  
 Apprentissage statistique et forage des données

Leaders • Animateurs : Rob Tibshirani, *Stanford University*, Hugh Chipman, *University of Waterloo*

**9:00–17:00 Workshop of the Business and Industrial Statistics Section** **AA-1355**  
**Atelier du Groupe de statistique industrielle et de gestion**

Data Mining Tools for Business Applications  
 Outils de forage de données pour applications en entreprise

Leaders • Animateurs : Shirley Mills & Ted Norminton, *Carleton University*

**17:30 – 18:30 Session 0A** **AA-1140**

Accreditation of Professional Statisticians in Canada  
 L'accréditation des statisticiens professionnels au Canada

Special Late Breaking Session • Conférence spéciale dernière nouvelle

Organizer and Chair • Responsable et présidente : Judy-Anne CHAPMAN, *University of Toronto*

17:30 Sylvain VÉGIARD, *Institut de la statistique du Québec*

ASSQ, Accreditation and “Statistician” Membership Category • ASSQ, membre statisticien et accréditation [MS-1]

17:45 Carl SCHWARZ, *Simon Fraser University*

Educational Requirements for A.Stat. and P.Stat. • Formation académique requise pour les titres A.Stat et P.Stat [MS-2]

18:00 Ken MCRAE, *Agriculture and Agri-Food Canada*, Fernando CAMACHO, *DAMOS Consulting*  
Professional Statistical Experience • Expérience professionnelle en statistique [MS-3]

18:15 Judy-Anne CHAPMAN, *University of Toronto*  
Professional Development • Développement professionnel [MS-4]

### 18:30 – 22:00 Session 0B

### Hall d'honneur

Contributed Poster Session and Cocktail (evening)  
Présentation d'affiches et coquetel (soirée)

18:30 Keith WORSLEY, *McGill University*  
Statistics, Geometry and Brain Mapping • Statistique, géométrie et cartographie cérébrale [MS-5]

18:30 Julie ZHOU, Peilin SHI & Jane YE, *University of Victoria*  
Discrete Minimax Robust Designs for Regression Models with Autocorrelated MA Errors • Plans minimax robustes discrets pour des modèles de régression avec des erreurs MA autocorrélées [MS-6]

18:30 Zilin WANG & David BELLHOUSE, *University of Western Ontario*  
Shift Function Plots for Regression Fitting • Graphique de la fonction de décalage pour l'ajustement de régressions [MS-7]

18:30 Jason LOEPPKY, Derek BINGHAM & Randy R. SITTER, *Simon Fraser University*  
Constructing Nonregular Robust Parameter Designs • Construction de design non réguliers à paramètre robuste [MS-8]

18:30 David BRILLINGER, *University of California, Berkeley*  
An Analysis of a Bivariate Time Series in Which the Components Are Sampled at Different Instants • Analyse d'une série chronologique bivariée dans laquelle les composantes sont prélevées à différents instants [MS-9]

18:30 Philip MOORHEAD, *University of Western Ontario*  
Small Sample Inference for Detecting Operational Confounders • Inférences sur de petits échantillons pour détecter des facteurs parasites opérationnels [MS-10]

18:30 Jeremy HAMM, Carl SCHWARZ & Ian WILLIAMSON, *Simon Fraser University*  
Estimating the Rate of Concussions in British Columbia Minor Hockey Using Community Volunteer Collected Data • Estimation du taux de commotions cérébrales dans le hockey mineur en Colombie-Britannique en utilisant des volontaires dans la communauté pour cueillir les données [MS-11]

18:30 Azadeh MOGHTADERI, Glen TAKAHARA & David THOMSON, *Queen's University*  
Spectral Analysis of Solar Flare Activity • Analyse spectrale des éruptions solaires [MS-12]

18:30 Zi JIN, *McMaster University*  
On a Mixture Generalized Autoregressive Conditional Heteroscedastic Model • Un modèle de mélange autorégressif conditionnellement hétéroscédastique généralisé [MS-13]

- 18:30 Ian MOORE & David J. THOMSON, *Queens University*  
Efficient Data Smoothers Using Slepian Functions • Lissage efficace de données en utilisant des fonctions de type slepian [MS-14]
- 18:30 Gemai CHEN, *University of Calgary*, Lingyun ZHANG, *Massey University, New Zealand*  
EWMA Charts for Monitoring Mean Changes in Weibull Processes with Censored Lifetime Data • Diagramme EWMA pour détecter des changements dans la moyenne de processus de Weibull avec des données de survie censurées [MS-15]
- 18:30 Srimantoorao APPADOO & A. THAVANESWARAN, *University of Manitoba*  
Some Properties of GARCH Processes • Quelques propriétés des processus GARCH [MS-16]
- 18:30 Shesh RAI, *St. Jude Children's Research Hospital, Memphis, TN*, Daniel HUNT, *St. Jude Children's Research Hospital, Memphis, TN*  
Modeling Random Effects in a Dose-Response Model with Application to Developmental Toxicity Data • Modélisation de l'effet aléatoire dans un modèle dose-réponse avec des applications pour des données sur la toxicité de développement [MS-17]
- 18:30 Scott AMISS & David J. THOMSON, *Queen's University*  
Detection of Long-Range Dependence and the Nile River • Détection de dépendance à long terme et le fleuve du Nil [MS-18]
- 18:30 Ying ZHANG & Ian MCLEOD, *University of Western Ontario*  
Computer Algebra Derivation of the Bias of Burg Estimators • Dérivation par l'algèbre computationnelle du biais des estimateurs de Burg [MS-19]
- 18:30 Mylène BÉDARD, *University of Toronto*  
On the Robustness of Optimal Scaling for Metropolis-Hastings Algorithms • Robustesse de l'échellonnage optimal pour les algorithmes de type Metropolis-Hastings [MS-20]
- 18:30 María Inés URRUTIA, Helena LUNAZZI, Marta DE LA FUENTE, Diana ELIAS, Favia FERNANDEZ & Soledad DE LA FUENTE, *Universidad Nacional de La Plata, Argentina*  
Construction of scales for the Rorschach technique (SC) in Argentina • Construction de barèmes pour la technique Rorschach (SC) en Argentine [MS-21]

## Monday, May 31 • Lundi 31 mai

---

### 08:30 – 10:00 Session 1

**B-2285**

Inaugural Session and SSC Presidential Invited Address

Séance inaugurale et allocution de l'invitée de la présidente

Special Session • Conférence spéciale

Organizer and Chair • Responsable et présidente : Mary E. THOMPSON, *University of Waterloo*

08:30 Kathryn ROEDER, A. RINALDO, B. DEVLIN & L. WASSERMAN, *Carnegie Mellon University*

Discovering Haplotype Blocks in the Human Genome • Découvrir des blocs haplotypes dans le génome humain [MS-22]

### 10:00 – 10:30 Coffee Break • Pause-café

**Pavillon André-Aisenstadt**

### 10:30 – 12:00 Session 2A

**Z-310**

Gender-Related Statistics in Education

Statistiques liées au genre dans le domaine de l'éducation

Invited Paper Session • Présentations sur invitation

Session sponsored by the *Canadian Caucus for Women in Statistics* and by the *Women in Statistics Committee of the SSC*

Séance parrainée par le *Caucus canadien des statisticiennes* et par le *Comité de la SSC pour la promotion de la femme en statistique*

Organizer and Chair • Responsable et présidente : Susan RUBIN-BLEUER, *Statistics Canada/Statistique Canada*

10:30 Denise LIEVESLEY, *UNESCO*

Engendering Education Statistics: The International Perspective • Intégration du genre dans les statistiques de l'éducation : Perspective internationale [MS-23]

11:00 Jennifer CONNOLLY & Trish WILLIAMS, *York University*

Sexual Harassment and Bullying in High Schools: Implications • Harcèlement sexuel et l'intimidation dans les écoles secondaires : implications [MS-24]

11:30 Sophia HUYER, *Gender Advisory Board—UN Commission on Science and Technology for Development*

Developing Indicators to Measure Participation of Girls and Women in Education, Science and Technology • Développer des indicateurs pour mesurer la participation des filles et des femmes dans l'éducation de la science et de la technologie [MS-25]

**10:30 – 12:00 Session 2B****Z-330**

The Escalation of Non-response and its Impacts on Surveys

La croissance de la non-réponse et ses impacts sur les enquêtes

Invited Paper Session • Présentations sur invitation

Session sponsored by the *Association des statisticiennes et des statisticiens du Québec*Séance parrainée par l'*Association des statisticiennes et des statisticiens du Québec*Organizer and Chair • Responsable et président : Michel FLUET, *SOM*10:30 Sylvain VÉGIARD, *Institut de la statistique du Québec*

Evolution of the Response Rate in the Studies of the Institut de la statistique du Québec • Évolution des taux de réponse dans les enquêtes à l'Institut de la statistique du Québec [MS-26]

11:00 Richard SAINT-PIERRE, *Ad hoc recherche*

Nonresponse : A Survey Firm's Point of View • La non-réponse — le point de vue d'une firme de sondage [MS-27]

11:30 Michel SAULNIER, *Bell Canada*

Nonresponse... What Is the Real Problem? • La non-réponse... quel est le vrai problème ? [MS-28]

**10:30 – 12:00 Session 2C****AA-1140**

Current Trends in Statistical Genetics

Tendances actuelles en statistique génétique

Invited Paper Session • Présentations sur invitation

Session sponsored by the *Institute of Mathematical Statistics*Séance parrainée par l'*Institut de statistique mathématique*Organizer and Chair • Responsable et président : Sabin LESSARD, *Université de Montréal*10:30 Robert GRIFFITHS, *University of Oxford*, Maria De IORIO, *Imperial College*

Importance Sampling on Evolutionary Histories • Échantillonnage avec fonction d'importance pour des histoires évolutionnistes [MS-29]

11:15 Rasmus NIELSEN, Rick DURRETT & Thomas YORK, *Cornell University*

Bayesian Estimation of the Evolutionary History of Two Genomes • Estimation bayésienne de l'histoire évolutionniste de deux génomes [MS-30]

**10:30 – 12:00 Session 2D****AA-1360**

Case Study I: Asthma

Étude de cas I : Asthme

Organizer and Chair • Responsable et présidente : Peggy T. NG, *York University*10:30 Edmee FRANSSSEN, *GSK*

Presentation of the Data Set • Présentation du jeu de données [MS-31]

10:40 Binod NEUPANE, Miao FANG & Afisi ISMAILA, *McMaster University*



- 10:55 Yongxu LAN, Meijie GUO, Xian ZHOU & Rui WANG, *University of Calgary*  
 11:10 Nikolai SLOBODIANIK, Tao SUN, Ye SUN & Noemi TOIBER TEMIN, *York University*  
 11:25 Ana Paola SANCHEZ LEZAMA, Jinnan LIU & Eraste NKUNDUMUKIZA, *York University*  
 11:40 Meng DU, *University of Toronto*, Jing WANG & Shufang WANG, *York University*

**10:30 – 12:00 Session 2E****AA-1355**

Finance

Finance

Contributed Paper Session • Séance de communications libres

Chair • Président : Pierre DUCHESNE, *Université de Montréal*

- 10:30 Yanqing YI & Xikui WANG, *University of Manitoba*

Portfolio Selection and Markov Decision Processes • Sélection du portefeuille et processus de décision de Markov [MS-32]

- 10:45 Yan WANG & Xikui WANG, *University of Manitoba*

Optimal Portfolio Selection Among Risky Assets • Sélection optimale d'un portefeuille parmi des titres risqués [MS-33]

- 11:00 Alex BADESCU & Reg KULPERGER, *University of Western Ontario*

Discrete Time Finance and Option Pricing: A Semiparametric Approach • Finance à temps discret et évaluation du prix d'une option : une approche semi-paramétrique [MS-34]

- 11:15 Zengjing CHEN & Reg KULPERGER, *University of Western Ontario*

Nonlinear Expectations with Application to Incomplete Market • Espérances non linéaires avec application au marché incomplet [MS-35]

- 11:30 Long YU, Haitao LI & Martin WELLS, *Cornell University*

Monte Carlo Markov Chain Analysis of Time-changed Levy Processes of Stock Return Dynamics • Analyse de Monte Carlo avec chaîne de Markov de processus de Levy à changement temporel pour la dynamique de rendement d'une action [MS-36]

- 11:45 Mahmoud ZAREPOUR & H. ISHWARAN, *University of Ottawa/Université d'Ottawa*

On Some Extension to Gamma Processes • Sur le prolongement des processus gamma [MS-37]

**10:30 – 12:00 Session 2F****Z-317**

GEE and Related Methods

EGE et méthodes connexes

Contributed Paper Session • Séance de communications libres

Chair • Présidente : Rachel MACKAY ALTMAN, *University of Washington*

- 10:30 Ye TAN & Robert PLATT, *McGill University*

Analysis of Large Cluster Randomized Trials with Binary Outcomes and Weak Intra-class Correlation • Analyse d'essais randomisés à grandes grappes avec des résultats binaires et une corrélation intra-groupe faible [MS-38]

10:45 Leilei ZENG & Richard J. COOK, *University of Waterloo*

Misspecification of the Drop-out Process in Inverse Probability Weighted GEE • Mauvaise spécification du processus de retrait dans des EEG à probabilité inverse pondérées [MS-39]

11:00 Lang WU, *University of British Columbia*

Approximate Inference for Nonlinear Mixed-effects Models with Nonignorably Missing Covariates • Inférence approximative pour les modèles non linéaires à effet mixte avec covariables manquantes non ignorables [MS-40]

11:15 Ioana SCHIOPU-KRATINA, *Statistics Canada/Statistique Canada*, Raluca M. BALAN, *University of Ottawa/Université d'Ottawa*

Asymptotic Theory with Generalized Estimation Equations for Longitudinal Data • Théorie asymptotique avec des équations d'estimation généralisées pour des données longitudinales [MS-41]

11:30 Ali Reza FOTOUHI, *University College of the Fraser Valley*

Two-stage Models to Control for Overdispersion in Longitudinal Count Data • Modèles en deux étapes pour contrôler la sur-dispersion dans des données de dénombrement longitudinales [MS-42]

## 12:00 – 13:30 Lunch • Dîner

### 13:30 – 15:00 Session 3A

**Z-310**

Special Invited Session of the Survey Methods Section

Allocution de l'invité d'honneur du Groupe de méthodologie d'enquête

Special Session • Conférence spéciale

Session sponsored by the *Survey Methods Section*

Séance parrainée par le *Groupe de méthodologie d'enquête*

Organizer and Chair • Responsable et présidente : Georgia R. ROBERTS, *Statistics Canada/Statistique Canada*

13:30 Raymond CHAMBERS, *University of Southampton*

Informative Sampling and Sampling Information • Échantillonnage informatif et information d'échantillonnage [MS-43]

14:20 Discussant • Modérateur : David BINDER, *Statistics Canada/Statistique Canada*

14:35 Floor discussion • Discussion plénière

### 13:30 – 15:00 Session 3B

**Z-330**

Application of Statistical Methods to Actuarial Problems

Application de méthodes statistiques à des problèmes actuariels

Invited Paper Session • Présentations sur invitation

Session sponsored by the *Business and Industrial Statistics Section*

Séance parrainée par le *Groupe de statistique industrielle et de gestion*

Organizer and Chair • Responsable et président : Thierry DUCHESNE, *Université Laval*

13:30 José GARRIDO, *Concordia University*

Actuarial Science in the City • L'actuariat dans la cité [MS-44]

14:00 Charles DUGAS, *Université de Montréal & ApSTAT Technologies Inc.*

Asymmetric Heavy-tailed Distributions of Insurance Claims and Neural Networks • Les distributions asymétriques à aile lourde de réclamations d'assurance et les réseaux de neurones [MS-45]

14:30 Jean-Denis ROY, *AXA, Montréal*

Can Advanced Statistical Models be Applied in the Real World? • Les modèles statistiques avancés peuvent-ils vraiment être mis en pratique ? [MS-46]

### 13:30 – 15:00 Session 3C

**AA-1140**

Biostatistics Applications: The Dynamics of Collaboration

Applications biostatistiques : La dynamique de la collaboration

Invited Paper Session • Présentations sur invitation

Session sponsored by the *Biostatistics Section*

Séance parrainée par le *Groupe de biostatistique*

Chair • Président : Penny BRASHER, *Alberta Cancer Board*

Organizers • Responsables : Rollin F. BRANT, *University of Calgary*

13:30 Rick ROUTLEDGE, *Simon Fraser University*

Ecosystem Research in an Unstable Environment • Recherche sur les écosystèmes dans un environnement instable [MS-47]

14:00 Rob TIBSHIRANI, *Stanford University*

Using a Car Phone: Like Driving Drunk? • Utiliser un cellulaire au volant : est-ce comme conduire en état d'ivresse ? [MS-48]

14:30 Rollin BRANT, *University of Calgary*

Randomized Trials (and Tribulations) in the Intensive Care Unit • Essais randomisés (et tribulations) dans l'unité des soins intensifs [MS-49]

### 13:30 – 15:00 Session 3D

**AA-1360**

Capture-recapture Models

Modèles de capture-recapture

Invited Paper Session • Présentations sur invitation

Organizer and Chair • Responsable et président : Louis-Paul RIVEST, *Université Laval*

13:30 Carl SCHWARZ, *Simon Fraser University*

Stratification in Capture-recapture Modeling • Stratification et les modèles capture-recapture [MS-50]

14:00 Kenneth POLLOCK, Mathew ALLDREDGE & Ted SIMONS, *North Carolina State University*

Application of Capture-recapture Type Models in Point Count Surveys • Application de modèles de type capture-recapture dans des sondages de dénombrement ponctuel [MS-51]

14:30 Roger PRADEL, *Centre d'écologie fonctionnelle et évolutive, Montpellier*

Use of Hidden Markov Models in Capture-recapture • Utilisation des modèles de Markov cachés en capture-recapture [MS-52]

### 13:30 – 15:00 Session 3E

AA-1355

Time Series Analysis

Analyse de séries chronologiques

Contributed Paper Session • Séance de communications libres

Chair • Présidente : Anne-Catherine FAVRE, *Institut national de la recherche scientifique du Québec*

13:30 Rolf TURNER, *University of New Brunswick*, Patrick CHAREKA, *University of Missouri*, Florance MATARISE, *University of Zimbabwe*

A Test for Additive Outliers in Time Series • Un test pour des valeurs aberrantes additives dans les séries chronologiques [MS-53]

13:45 Abdeljelil FARHAT & Jean-Marie DUFOUR, *Université de Montréal & Centre interuniversitaire de recherche en analyse des organisations & Centre interuniversitaire de recherche en économie quantitative*, Abdelwahed TRABELSI, *Université de Tunis*

Exact Randomized Permutation Serial Dependence Tests for Continuous and Discrete Distributions • Tests exacts d'indépendance sérielle dans les cas de distributions continues et discrètes [MS-54]

14:00 Chafik BOUHADDIOUI, *Centre interuniversitaire de recherche en analyse des organisations*, Jean-Marie DUFOUR, *Université de Montréal & Centre interuniversitaire de recherche en analyse des organisations & Centre interuniversitaire de recherche en économie quantitative*

Test for Noncorrelation of Two Infinite Order Cointegrated Vector Autoregressive Series • Test de non corrélation de deux séries cointégrées autorégressives d'ordre infini. [MS-55]

14:15 Abdessamad SAIDI & Roch ROY, *Université de Montréal*, Guy MÉLARD, *Université Libre de Bruxelles*

Exact Maximum Likelihood Estimation of Structured or Unit Root Multivariate Time Series Models • Estimation de vraisemblance maximale exacte de modèles structurés ou à racine unitaire pour des séries chronologiques multivariées [MS-56]

14:30 David CAMPBELL, *McGill University*, Hong GU, *Dalhousie University*

Wavelets and the Clustering of Time Series • Ondelettes et classification de séries chronologiques [MS-57]

14:45 Yogendra CHAUBEY, S. K. GOYAL & F. NEBEBE, *Concordia University*

Reorder Levels of Products with Autocorrelated Demand and Resource Restriction • Seuil de commande de produits avec une demande autocorrélée et des restrictions de ressources [MS-58]

**13:30 – 15:00 Session 3F****Z-317**

## Survival Analysis I

## Analyse de durées de vie I

Contributed Paper Session • Séance de communications libres

Chair • Président : Alain VANDAL, *McGill University*13:30 Denise BABINEAU, *University of Waterloo*

Goodness of Fit Tests for General Interval Censored Data • Test d'adéquation pour des données d'un modèle général de censure par intervalles [MS-59]

13:45 Uditha BALASOORIYA & C.-K. LOW, *Nanyang Technological University, Singapore*

Design of Test Plans for Weibull Lifetimes under Competing Causes of Failure • Design de plans de tests pour des temps de vie Weibull sous des causes d'échecs concurrentes [MS-60]

14:00 Katherine DAVIES, *McMaster University*

Inference for Type II Right Progressively Censored Samples from the Generalized Extreme Value Distribution • Inférence pour des échantillons censurés progressivement à droite de type II de la distribution de valeurs extrêmes généralisée [MS-61]

14:15 Lajmi LAKHAL CHAIEB, Louis-Paul RIVEST & Belkacem ABDIOUS, *Université Laval*

Estimating Survival Function Under Dependent Truncation • Estimation de la fonction de survie sous une troncation dépendante [MS-62]

14:30 Yun Hee CHOI & David E. MATTHEWS, *University of Waterloo*

Accelerated Life Regression Modeling of Dependent Bivariate Time-to-event Data • Modélisation par régression du temps de vie accéléré pour des données bivariées dépendantes du temps jusqu'à un événement [MS-63]

**15:00 – 15:30 Coffee Break • Pause-café****Pavillon André-Aisenstadt****15:30 – 17:00 Session 4A****AA-1140**

## In Honour of Jim Zidek, on his 65th Birthday

## En hommage à Jim Zidek, pour ses 65 ans

Special Session • Conférence spéciale

Organizer and Chair • Responsable et président : Harry JOE, *University of British Columbia*15:30 Nhu LE, *BC Cancer Agency & University of British Columbia*

Bayesian Approaches to Environmental Network Design and Spatial Prediction • Approche bayésienne pour le design de réseaux environnementaux et la prédiction spatiale [MS-64]

16:00 Feifang HU, *University of Virginia*, Jim ZIDEK, *University of British Columbia*

Weighted Likelihood and its Applications • Vraisemblance pondérée et ses applications [MS-65]

16:30 Constance VAN EEDEN, *University of British Columbia*

Group-Bayes Estimation • Estimation bayésienne basée sur un consensus [MS-66]

**15:30 – 17:00 Session 4B****Z-310**

Causal Inference in Biostatistics and Epidemiology  
L'inférence causale en biostatistique et en épidémiologie

Invited Paper Session • Présentations sur invitation

Session sponsored by the *Biostatistics Section*  
Séance parrainée par le *Groupe de biostatistique*

Organizer and Chair • Responsable et président : Robert W. PLATT, *McGill University*

15:30 Robert PLATT & Tong GUO, *McGill University*

Causal Inference in Randomized Trials in the Presence of Partial Compliance • Inférence causale dans les études cliniques randomisées en présence de conformité partielle [MS-67]

16:00 Miguel HERNÁN, *Harvard University*

Causal Inference for Time-varying Exposures • Inférence causale pour des expositions variables dans le temps [MS-68]

**15:30 – 17:00 Session 4C****AA-1360**

New Methodologies for Census 2006  
Nouvelles méthodologies pour le recensement de 2006

Invited Paper Session • Présentations sur invitation

Session sponsored by the *Survey Methods Section*  
Séance parrainée par le *Groupe de méthodologie d'enquête*

Organizer and Chair • Responsable et président : Dave DOLSON, *Statistics Canada/Statistique Canada*

15:30 Kathy MCCLEAN & Kim CHARLAND, *Statistics Canada/Statistique Canada*

Using an Address Register for a Mail-out Census • Utilisation d'un registre d'adresse pour un recensement par voie postale [MS-69]

15:52 Laurent ROY & Manon MONETTE, *Statistics Canada/Statistique Canada*

Census Questionnaire Design for Automated Data Capture • La conception de questionnaire du recensement dans le cadre d'une saisie automatisée des données [MS-70]

16:14 Danielle LAROCHE & Graeme GILMOUR, *Statistics Canada/Statistique Canada*

Census of Population Data Collection via Internet • La collecte des données du recensement de la population par Internet [MS-71]

16:36 Yves SAINT-PIERRE & Joel YAN, *Statistique Canada/Statistique Canada*

Statistical Literacy: How Statistics Canada Promotes It • Comment Statistique Canada fait la promotion de l'alphabétisme statistique [MS-72]

**15:30 – 17:00 Session 4D****Z-330**Applied Probability  
Probabilité appliquée

Invited Paper Session • Présentations sur invitation

Chair • Président : Jean-François ANGERS, *Université de Montréal*Organizers • Responsables : Liqun WANG, *University of Manitoba*15:30 Antonio DI CRESCENZO, V. GIORNO & A. G. NOBILE, *Università di Salerno*, L. M. RICCIARDI, *Università Federico II, Napoli*

On the First-visit-time Problem for Birth and Death Processes with Catastrophes • Sur le problème du premier temps de visite pour des processus de naissance-mort avec catastrophes [MS-73]

16:00 Alexander NOVIKOV, *University of Technology, Sydney*, K. BOROVKOV, *The Univeristy of Melbourne*

Boundary Crossing Problems with Applications to Option Pricing • Problème de croisement de frontière avec application en tarification d'options [MS-74]

16:30 Liqun WANG, *University of Manitoba*

Boundary crossing probability for diffusion processes and piecewise continuous boundaries • Probabilité de croisement de frontière pour des processus de diffusion et des frontières continues par morceaux [MS-75]

**15:30 – 17:00 Session 4E****Z-317**Regression  
Régression

Contributed Paper Session • Séance de communications libres

Chair • Président : Jean-Pierre CARMICHAEL, *Université Laval*15:30 Shahjahan KHAN, *University of Southern Queensland*

Estimation of the Intercept Parameters of Two Simple Regression Models When the Slopes Are Suspected to Be Equal • Estimation du paramètre d'ordonnée à l'origine de deux modèles de régression simple quand les pentes sont soupçonnées être égales [MS-76]

15:45 Hyang Mi KIM & A. K. Md. E. SALEH, *Carleton University*

Improved Estimation of Regression Parameters in Measurement Error Models • Estimation améliorée des paramètres de régression dans des modèles d'erreur de mesure [MS-77]

16:00 Bashir KHAN, *Saint Mary's University*

Estimation of Regression Parameters: A Two-samples Problem • Estimation des paramètres de régression : un problème à deux échantillons [MS-78]

16:15 Pierre LAFAYE DE MICHEAUX, *Université Pierre Mendès France*, Christian LÉGER, *Université de Montréal*

Multiple Regression with Model Selection: A Bootstrap Approach • Régression linéaire adaptative multiple : une approche bootstrap [MS-79]

16:30 Dianliang DENG, *University of Regina*

Goodness of Fit of Multinomial Regression Models to Sparse Data • Adéquation d'un modèle de régression multinomial à des données rares [MS-80]

16:45 Kamanzi BINYAVANGA, *University of Transkei, South Africa*, Sarel J. STEEL, *University of Stellenbosch, South Africa*

Edgeworth-corrected Small-sample Confidence Intervals for Nonlinear Functions of Parameters in Linear Regression • Intervalles de confiance corrigés par Edgeworth sur de petits échantillons pour des fonctions non linéaires des paramètres d'une régression linéaire [MS-81]

### 15:30 – 17:00 Session 4F

AA-1355

Statistical Genetics

Statistique en génétique

Contributed Paper Session • Séance de communications libres

Chair • Président : Robert NADON, *McGill University*

15:30 André DABROWSKI, *University of Ottawa/Université d'Ottawa*

Comparing Two Independent cDNA Microarrays • Une comparaison de deux microréseaux indépendants [MS-82]

15:45 Gary SNEDDON & Brajendra C. SUTRADHAR, *Memorial University of Newfoundland*

On Semiparametric Familial-longitudinal Models • Sur les modèles familiaux longitudinaux semi-paramétriques [MS-83]

16:00 Zeny FENG, Jiahua CHEN & Mary E. THOMPSON, *University of Waterloo*

The Universal Validity of the Possible Triangle Constraints of the Affected-Sib-Pairs. • La validité universelle de la contrainte triangulaire possible de la paire de fratrie affectée [MS-84]

16:15 Karen KOPCIUK, *Alberta Cancer Board*, L. BRIOLLAIS & W. HE, *Samuel Lunenfeld Research Institute*, E. PARKHOMENKO, *University of Toronto*, J. GREEN, *Memorial University of Newfoundland*, J MCLAUGHLIN, *Samuel Lunenfeld Research Institute*

Effects of Familial Correlation and Ascertainment Bias on Estimating the Penetrance of a Major Gene • Effets des corrélations familiales et du biais de sélection sur l'estimation de la pénétrance d'un gène majeur [MS-85]

16:30 Cyr Emile M'LAN & Celia GREENWOOD, *Hospital for Sick Children, Toronto*, Shelley BULL, *Mount Sinai Hospital, Toronto*

Detecting Differentially Expressed Genes among Human Population • Détecter les différentes expressions des gènes à travers la population humaine [MS-86]

### 17:00 – 18:30 Session 4G

Z-330

NSERC Workshop: Tips to Prepare your Next Discovery Grant Application

Atelier du CRSNG : Conseils pour l'élaboration de votre prochaine demande de subvention à la découverte

Special Session • Conférence spéciale

Organizer and Chair • Responsable et président : Jamie STAFFORD, *University of Toronto*



## Tuesday, June 1 • Mardi 1<sup>er</sup> juin

---

### 08:30 – 10:00 Session 5

**B-2285**

Gold Medal Address

Allocution de la récipiendaire de la médaille d'or

Special Session • Conférence spéciale

Organizer and Chair • Responsable et président : Louis-Paul RIVEST, *Université Laval*

08:30 Mary THOMPSON, *University of Waterloo*

Understanding Associations: Implications for the Design and Analysis of Longitudinal Surveys •  
Comprendre les associations : Implications pour la planification et l'analyse d'enquêtes longitudinales [MS-87]

### 10:00 – 10:30 Coffee Break • Pause-café

**Pavillon André-Aisenstadt**

### 10:30 – 12:00 Session 6A

**AA-1140**

Special Invited Session of the Biostatistics Section

Allocution de l'invité d'honneur du Groupe de biostatistique

Special Session • Conférence spéciale

Session sponsored by the *Biostatistics Section*

Séance parrainée par le *Groupe de biostatistique*

Organizer and Chair • Responsable et président : Salomon MINKIN, *Ontario Cancer Institute*

10:30 Richard SIMON, *National Cancer Institute, Rockville, MD*

Key Features in the Design and Analysis of DNA Microarray Studies • Caractéristiques importantes dans la planification et l'analyse d'expériences microréseaux d'ADN [MS-88]

11:30 Discussant • Modératrice : Celia GREENWOOD, *Hospital for Sick Children, Toronto*

11:45 Floor discussion • Discussion plénière

### 10:30 – 12:00 Session 6B

**AA-1360**

Analysis on Complex Surveys

L'analyse de données d'enquêtes complexes

Invited Paper Session • Présentations sur invitation

Session sponsored by the *National Program on Complex Data Structures*

Séance parrainée par le *Programme national sur les structures de données complexes*

Organizer and Chair • Responsable et président : D. Roland THOMAS, *Carleton University*

10:30 J. N. K. RAO, *Carleton University*

Analysis of Complex Sample Survey Data: An Appraisal • Analyses de données d'enquêtes complexes : une évaluation [MS-89]

11:00 Emmanuel BENHIN, *Statistics Canada/Statistique Canada*, J. N. K. RAO, *Carleton University*  
 Analysis of Categorical Data from Complex Sample Surveys Using Inverse Sampling • Analyse de données catégoriques provenant de sondages complexes en utilisant l'échantillonnage inverse [MS-90]

11:30 Danny PFEFFERMANN, *Hebrew University & University of Southampton*  
 Multi-level Modelling of Complex Survey Data • Modélisation multi-niveaux de données de sondages complexes [MS-91]

### 10:30 – 12:00 Session 6C

**Z-310**

#### Applications of Wavelets to Statistical Problems

#### Applications des ondelettes à des problèmes statistiques

Invited Paper Session • Présentations sur invitation

Organizer and Chair • Responsable et président : Alwell J. OYET, *Memorial University*

10:30 Guy NASON, *University of Bristol*, Piotr FRYZLEWICZ, *Imperial College*

How Statistics Can Help You See in the Dark! • Comment la statistique peut-elle vous aider à voir dans le noir ? [MS-92]

11:00 Anestis ANTONIADIS, *Université Joseph Fourier, Grenoble*, U. AMATO, *Consiglio Nazionale delle Ricerche, Napoli*, M. PENSKY, *University of Central Florida*

Wavelet Kernel Penalized Estimation for Random Design Regression • Noyaux autoreproduisants à base d'ondelettes et débruitage de signaux sur plans déterministes non équidistants [MS-93]

11:30 R. OGDEN, *Columbia University*

Some Applications of Wavelets in Functional Linear Models Analysis • Quelques utilisations des ondelettes dans l'analyse de modèles linéaires fonctionnels [MS-94]

### 10:30 – 12:00 Session 6D

**Z-330**

#### Monte Carlo Methods: Theory and Applications

#### Méthodes de Monte-Carlo : Théorie et applications

Invited Paper Session • Présentations sur invitation

Session sponsored by the *Bernoulli Society*

Séance parrainée par la *Société Bernoulli*

Organizer and Chair • Responsable et président : Bruno RÉMILLARD, *HEC Montréal*

10:30 Marcel RINDISBACHER, *University of Toronto*, J. DETEMPLE, *Boston University*, R. GARCIA, *Université de Montréal*

Asymptotic Properties of Monte Carlo Estimators for Diffusions • Propriétés asymptotiques des estimateurs pour les diffusions [MS-95]

11:00 Neal MADRAS, *York University*, John Wai Kong YUEN, *Brock University*

Decomposition Bounds for MCMC Convergence Rates • Limites de décomposition pour les taux de convergence de MCMC [MS-96]

11:30 Ivan GENTIL, *Université Paris Dauphine*, Bruno RÉMILLARD, *HEC Montréal*

New Selection in a Branching and Interacting Particle Systems • Nouvelle sélection pour des systèmes de particules en interaction [MS-97]

### 10:30 – 12:00 Session 6E

**Z-317**

Survey Sampling I

Méthodologie d'enquête I

Contributed Paper Session • Séance de communications libres

Chair • Présidente : Diane STUKEL, *UNESCO*

10:30 Edward CHEN, J. LINDEYER & G. LAFLAMME, *Statistics Canada/Statistique Canada*

Issues with High Vacancy Areas and Remote Areas in the Canadian Labour Force Survey Sample Redesign • Problèmes avec les secteurs à faible densité et les régions éloignées dans la révision du design de l'Enquête sur la population active [MS-98]

10:45 Fritz PIERRE & Yves BÉLAND, *Statistics Canada/Statistique Canada*

Person-level Sampling in the Canadian Nutrition Survey • Sélection des personnes dans l'enquête canadienne sur la nutrition [MS-99]

11:00 Marco GRENIER & François BRISEBOIS, *Statistics Canada/Statistique Canada*

The Weighting Strategy for Cycle 2.1 of the Canadian Community Health Survey • Pondération du cycle 2.1 de l'Enquête sur la santé dans les collectivités canadiennes [MS-100]

11:15 Julie HORROCKS, *University of Guelph*, Mieko YOSHIHAMA & Kristin DUNKLE, *University of Michigan*, Saori KAMANO, *National Institute of Population and Social Security Research, Tokyo*

Estimating Lifetime Risk • Estimation du risque à vie [MS-101]

11:30 Ante ROZGA, *University of Split/Université de Split*

Statistical Problems with Inflation Measures in Transition Countries • Problèmes statistiques avec les mesures d'inflation dans les pays en transition [MS-102]

### 10:30 – 12:00 Session 6F

**AA-1355**

Biostatistics I

Biostatistique I

Contributed Paper Session • Séance de communications libres

Chair • Présidente : Wendy LOU, *University of Toronto*

10:30 Howard WAINER, *National Board of Medical Examiners, Philadelphia*, Xiaohui WANG, *University of North Carolina*, William P. SKORUPSKI, *University of Massachusetts*, Eric T. BRADLOW, *University of Pennsylvania*

A Bayesian Advantage in Setting Passing Scores: The PPOp Curve • Un avantage bayésien pour introduire des scores de passage : la courbe PPOp [MS-103]

10:45 Paul COREY, Malcolm KOO, Raymond TAM & Andrew WHITE, *University of Toronto*

Using Random Datasets and Automatic Marking in a Web Supported Biostatistics Module • Utilisation de jeux de données aléatoires et notation automatique dans un module de biostatistique sur internet [MS-104]

11:00 Victor NAWA & K. S. BROWN, *University of Waterloo*

Analysis of Developmental Trajectories and Binary Longitudinal Data • Analyses de trajectoires de développement mental et données longitudinales binaires [MS-105]

11:15 Moh'd ALODAT, *McGill University*

Detecting Conjunctions in Images • Détection des conjonctions dans les images [MS-106]

11:30 Alexander DE LEON, *University of Calgary*, C. RUDNISKY, *University of Alberta*, G. SINGH, *University of Calgary*

Estimating Sensitivity & Specificity for Binocular Data: Application to Ophthalmology • Estimation de la sensibilité et de la spécificité pour des données binoculaires : applications en ophtalmologie [MS-107]

## 12:00 – 13:30 Lunch • Dîner

### 13:30 – 15:00 Session 7A

**AA-1140**

Isobel Loutit Invited Address on Business and Industrial Statistics

Allocution Isobel-Loutit en statistique industrielle et en gestion

Special Session • Conférence spéciale

Session sponsored by the *Business and Industrial Statistics Section*

Séance parrainée par le *Groupe de statistique industrielle et de gestion*

Organizer and Chair • Responsable et président : Fernando CAMACHO, *Damos*

13:30 Daryl PREGIBON, *Google, Inc., New York*, Corinna CORTES & Chris VOLINSKY, *Google, Inc. & AT&T Labs*

Graph Mining • Exploitation de graphique [MS-108]

### 13:30 – 15:00 Session 7B

**Z-310**

Shape Analysis of 3D Brain Mapping Data

Analyse de forme de données de cartographie cérébrale 3D

Invited Paper Session • Présentations sur invitation

Organizer and Chair • Responsable et président : Keith J. WORSLEY, *McGill University*

13:30 Jason LERCH, Alex ZIJDENBOS & Alan C. EVANS, *McGill University*

Cortical Thickness from MRI • Épaisseur du cortex à partir d'images par résonance magnétique [MS-109]

14:00 Jean-François MANGIN, D. RIVIÈRE & A. CACHIA, *Service Hospitalier Frédéric Joliot, Orsay*,  
D. L. COLLINS & Alan C. EVANS, *McGill University*, J. REGIS, *CHU La Timone, Marseille*

A Structural Strategy for Morphometry of the Cerebral Cortex • Une stratégie structurelle pour la morphométrie du cortex cérébral [MS-110]

14:30 Paul THOMPSON & Arthur W. TOGA, *University of California, Los Angeles*

Mathematical/Computational Challenges in Brain Mapping • Défis mathématiques et informatiques dans la cartographie du cerveau [MS-111]

### 13:30 – 15:00 Session 7C

**Z-330**

#### Multi-level Modelling

#### Modélisation multi-niveau

Invited Paper Session • Présentations sur invitation

Session sponsored by the *National Program on Complex Data Structures*

Séance parrainée par le *Programme national sur les structures de données complexes*

Organizer and Chair • Responsable et président : Milorad KOVACEVIC, *Statistics Canada/Statistique Canada*

13:30 Fiona STEELE, Constantinos KALLIS, Harvey GOLDSTEIN & Heather JOSHI, *University of London*

Multilevel Multiprocess Modelling of Partnership Transitions and Fertility in Britain • Modélisation multi-états multi-niveaux de la transition d'union et de la fertilité en Angleterre [MS-112]

14:00 Renjun MA, *University of New Brunswick*

Some Challenging Issues in the Analysis of the National Longitudinal Survey of Children and Youth • Quelques problèmes intéressants dans l'analyse du sondage longitudinal des enfants et de la jeunesse [MS-113]

14:30 James BROWN & Nyovani MADISE, *University of Southampton*, David STEEL, *University of Wollongong, Australia*

Accounting for Survey Design in Multilevel Models: An Empirical Study • Tenir compte du plan d'échantillonnage dans les modèles multi-niveaux : une étude empirique [MS-114]

### 13:30 – 15:00 Session 7D

**AA-1360**

#### Nonlinear Time Series

#### Chroniques non linéaires

Invited Paper Session • Présentations sur invitation

Organizer and Chair • Responsable et président : Gemai CHEN, *University of Calgary*

13:30 Jianqing FAN, *Princeton University*, Jiancheng JIANG, *Peking University*

Dynamic Integration of Time- and State-domain Methods for Volatility Estimation • Méthode d'intégration dynamique des domaines du temps et de l'état pour l'estimation de la volatilité [MS-115]

14:00 Wolfgang SCHMID & Stefan SCHIPPER, *Europa-Universität Viadrina, Frankfurt*

Statistical Process Control for Nonlinear Time Series • Processus de contrôle statistique pour des séries chronologiques non linéaires [MS-116]

14:30 Min CHEN, *Chinese Academy of Sciences, Beijing/Académie des sciences de Chine, Beijing*, Gemai CHEN, *University of Calgary*

A Bootstrap Test of Conditional Heteroscedasticity in Nonlinear Time Series • Un test bootstrap pour l'hétéroscédasticité pour des séries chronologiques non linéaires [MS-117]

### 13:30 – 15:00 Session 7E

AA-1355

#### Survey Sampling II

#### Méthodologie d'enquête II

Contributed Paper Session • Séance de communications libres

Chair • Président : Sylvain VÉGIARD, *Institut de la statistique du Québec*

13:30 Hansheng XIE & Sanping CHEN, *Statistics Canada/Statistique Canada*

Collection follow-up score function and response bias • Fonction de score de suivi de collecte et biais de réponse [MS-118]

13:45 Caroline PELLETIER, *Statistics Canada/Statistique Canada*

Importance of a good imputation method for the education questions in the LFS • Importance d'une bonne méthode d'imputation aux questions d'éducation dans le cadre de l'EPA [MS-119]

14:00 Sat GUPTA, Bill THORNTON & Sunil SINGHAL, *University of Southern Maine*, Javid SHABBIR, *Quad-I-Azam University, Islamabad*

Use of Randomized Response Techniques in Dealing with Social Desirability Response Bias • L'utilisation de techniques de randomisation des réponses pour traiter le biais dans les réponses dû à la désirabilité sociale [MS-120]

14:15 Martin PANTEL & Yves LAFORTUNE, *Statistics Canada/Statistique Canada*

Graphical Analysis of Complex Survey Data • Analyse graphique de données provenant d'enquêtes à plan complexe [MS-121]

### 13:30 – 15:00 Session 7F

Z-317

#### Environmental Statistics

#### Statistique environnementale

Contributed Paper Session • Séance de communications libres

Chair • Président : James V. ZIDEK, *University of British Columbia*

13:30 Sabit ÇAKMAK, *Health Canada/Santé Canada*, Robert DALES, *University of Ottawa/Université d'Ottawa*

Does Socio-Demographic Status Influence the Effect of Ambient Air Pollution on Hospitalisation for Respiratory Conditions? Results from a Time Series Study in 10 Canadian Cities • Est-ce que le statut socio-démographique influence l'effet de la pollution atmosphérique ambiante sur

l'hospitalisation pour des problèmes respiratoires ? Résultats d'une étude de séries chronologiques dans dix villes canadiennes [MS-122]

13:45 Zhong LIU, Dou YIPING, Howard CHANG, Nhu LE & Jim ZIDEK, *University of British Columbia*

Evaluating the Performance of Regional Photochemical Ozone Modeling System • Évaluation de la performance de systèmes photochimiques de modélisation régionale de l'ozone [MS-123]

14:00 Dan GILLIS, *University of Guelph*

Logistic Analysis of Space-Time Data Investigating Potential Climatic Risk Factors of Waterborne Enteric Illness • Analyse logistique de données espace-temps pour investiguer des facteurs de risque potentiels pour les maladies entériques d'origine hydrique [MS-124]

14:15 Manon FLEURY, *Health Canada/Santé Canada*, John HOLT, *University of Guelph*, Dominique CHARRON, *Health Canada/Santé Canada*, Brian ALLEN, *University of Guelph*, Abdel MAAROUF, *Environment Canada/Environnement Canada*

Time Series Analyses for the Seasonality of Foodborne Diseases and Temperature • Analyses de séries chronologiques pour la saisonnalité des maladies d'origine alimentaire et la température [MS-125]

14:30 Alexandre BUREAU, *University of Lethbridge*

Efficiency of Study Designs to Detect Interaction Effects Between Genetic and Environmental Factors on an Additive Risk Scale • Efficacité de plans d'échantillonnage pour détecter un effet d'interaction entre des facteurs génétiques et environnementaux sur une échelle additive [MS-126]

**15:00 – 15:30 Coffee Break • Pause-café**

**Pavillon André-Aisenstadt**

**15:30 – 17:00 Session 8A**

**Z-310**

The Canadian Initiative in Social Statistics and Cross-disciplinary Collaboration

L'initiative canadienne en matière de statistiques sociales et la collaboration interdisciplinaire

Invited Paper Session • Présentations sur invitation

Organizer and Chair • Responsable et président : Jerald F. LAWLESS, *University of Waterloo*

15:30 Michael SMITH & Steven RYTINA, *McGill University*

Pushing the Limits: Using Statistics with Varying Amounts of Expertise • Pousser les limites : utiliser la statistique avec des niveaux d'expertise variables [MS-127]

16:00 Patricia NEWCOMBE-WELCH, *University of Waterloo*

Sharing Statistical and Social Science Methodologies: Mixed Blessings • Échange de méthodologies statistiques et de sciences sociales : approbation mixte [MS-128]

16:30 Renjun MA & Jon Douglas WILLMS, *University of New Brunswick*

Growth Trajectories in Childhood and Adolescent Overweight and Obesity in Canada • Courbes de croissance de l'embonpoint et de l'obésité chez les enfants et les adolescents au Canada [MS-129]

**15:30 – 17:00 Session 8B**

**AA-1140**

Statistical Analysis of Human Growth Data

Analyse statistique de données sur la croissance humaine

Invited Paper Session • Présentations sur invitation

Organizer and Chair • Responsable et président : James O. RAMSAY, *McGill University*

15:30 Michael HERMANUSSEN, *Aschauhof, Altenhof, Germany*, Christof MEIGEN, *Universität Leipzig*

A Web-based System For Monitoring the Growth of Children • Un système basé sur Internet pour suivre la croissance des enfants [MS-130]

16:00 James HANLEY, *McGill University*

“Transmuting” Women Into Men: Galton’s Family Data on Human Stature • Transformer les femmes en hommes : Les données de Galton sur la taille [MS-131]

16:30 James RAMSAY, *McGill University*

Human Growth Data: Multiresolution Challenges and Phase/Amplitude Variation • Données de croissance humaine : défis multi-résolutions et variation de phase et d’amplitude [MS-132]

**15:30 – 17:00 Session 8C**

**AA-1355**

Empirical Likelihood Methodology

Méthodes empirico-vraisemblantistes

Invited Paper Session • Présentations sur invitation

Organizer and Chair • Responsable et président : Jiahua CHEN, *University of Waterloo*

15:30 Jian-Jian REN, *University of Central Florida*

Weighted Empirical Likelihood • Vraisemblance empirique pondérée [MS-133]

16:00 Jason FINE, *University of Wisconsin-Madison*, Fei ZOU, *The University of North Carolina at Chapel Hill*

Empirical Likelihood Analysis of a Semiparametric Mixture Model, with Application to QTL Mapping • Analyse de vraisemblance empirique d’un modèle de mélange semi-paramétrique avec des applications au mapping de locus quantitatif [MS-134]

16:30 Min TSAO, *University of Victoria*

Some Finite Sample Results on Empirical Likelihood • Quelques résultats d’échantillons finis portant sur la vraisemblance empirique [MS-135]

**15:30 – 17:00 Session 8D**

**Z-330**

Teaching Statistics to Non-statisticians: Spreading the Good Word to Achieve Recognition for our Profession

L’enseignement de la statistique à des non-statisticiens : prêcher la bonne parole pour faire reconnaître notre profession

Invited Paper Session • Présentations sur invitation

Session sponsored by the *Association des statisticiennes et des statisticiens du Québec*

Séance parrainée par l’*Association des statisticiennes et des statisticiens du Québec*



Organizer and Chair • Responsable et président : Michel GUILLET, *Creascience*

15:30 François BELLAVANCE, *HEC Montréal*

Teaching Statistics to Future Managers: What Should Be the Main Expected Goal? • L'enseignement de la statistique à de futurs gestionnaires : quel doit être le principal objectif à atteindre? [MS-136]

16:00 Jean HARDY, *Services Conseils Hardy Inc., Québec*

Challenges Regarding the Use of Data Analysis Software in the Teaching of Statistics • Les défis posés par l'utilisation de logiciels d'analyse de données dans l'enseignement de la statistique [MS-137]

16:30 Natalie RODRIGUE, *Creascience Inc., Montréal*

Teaching Statistics in Companies and Research Centers: Replacing the Toolbox by a Better Understanding of Concepts • L'enseignement de la statistique en entreprise et dans les centres de recherche : de la boîte à outils à la maîtrise des concepts [MS-138]

## 15:30 – 17:00 Session 8E

**Z-220**

### Survey Sampling III

### Méthodologie d'enquête III

Contributed Paper Session • Séance de communications libres

Chair • Président : David R. BELLHOUSE, *University of Western Ontario*

15:30 Torsten HARMS & Pierre DUCHESNE, *Université de Montréal*

Calibration Estimators for Quantiles • Estimateurs de calage pour les quantiles [MS-139]

15:45 Raghunath ARNAB, *University of Durban-Westville*

Estimation of Variance of a Generalized Regression Predictor • Estimation de la variance d'un estimateur par la régression généralisé [MS-140]

16:00 Wilson LU & Randy R. SITTER, *Simon Fraser University*

Replicate Weights and Disclosure Control • Réplication des poids et contrôle de la divulgation [MS-141]

16:15 Yong YOU, *Statistics Canada/Statistique Canada*

Hierarchical Bayes Small Area Estimation with Estimated Sampling Variances • Estimation de la variance échantillonnale dans un petit domaine par un modèle hiérarchique bayésien [MS-142]

16:30 Sarjinder SINGH, *Saint Cloud State University*

Repair of Two-phase Calibration Methodology in Survey Sampling • Réparation de la méthodologie de calibration en deux phases en échantillonnage [MS-143]

**15:30 – 17:00 Session 8F****Z-317****Design and Analysis of Health Studies****Planification et analyse d'études sanitaires**

Contributed Paper Session • Séance de communications libres

Chair • Président : Peter D.M. MACDONALD, *McMaster University*15:30 Jie HUANG, *Novartis Pharmaceuticals, East Hanover, NJ*, Ming-Chun HSU, *MedFocus*Analysis of Binary Data in  $2 \times 2$  Crossover Designs: A Kappa Based Approach • Analyse de données binaires dans des designs croisés  $2 \times 2$  : une approche basée sur la mesure Kappa [MS-144]15:45 Jie ZHANG, *Novartis Pharmaceutical, East Hanover, NJ*

Evaluation of Sample Size in Thorough QT/QTc Study • Évaluation de la taille échantillonnale dans une étude QT/QTc complète [MS-145]

16:00 Jianan PENG, *Acadia University*, Charles Chu-In LEE, *Memorial University of Newfoundland*, Lin LIU, *University of California, San Diego*

Simultaneous Confidence Intervals for Umbrella Alternative • Intervalles de confiance simultanés pour une alternative parapluie [MS-146]

16:15 Lehana THABANE, *McMaster University*, K. O'ROURKE, *Ottawa Health Research Institute*, L. CHAMBERS, *University of Ottawa/Université d'Ottawa*, J. KACZOROWSKI, T. KARWALAJTYS & C. LEVITT, *McMaster University*, W. HOGG, *University of Ottawa/Université d'Ottawa*, L. DOLOVICH & C. BELLEY, *McMaster University*

Composite Outcomes in Cluster-randomized Clinical Trials: Experiences from the Community Hypertension Assessment Trial (CHAT) • Issues multiples dans des études cliniques à grappes randomisées : expérience de l'étude sur l'évaluation de l'hypertension dans la communauté (CHAT) [MS-147]

16:30 William ROSS & Jagadish RANGREJ, *Health Canada/Santé Canada*

Designing Dose Response Experiments for Microbial Pathogens • Design d'expériences dose-réponse pour des microbes pathogènes [MS-148]

## Wednesday, June 2 • Mercredi 2 juin

---

### 08:30 – 09:15 Session 9A1

AA-1140

Pierre Robillard Award Address

Présentation de la lauréate du Prix Pierre-Robillard

Special Session • Conférence spéciale

Organizer and Chair • Responsable et président : Bruno RÉMILLARD, *HEC Montréal*

### 09:15 – 10:00 Session 9A2

AA-1140

Canadian Journal of Statistics Award Address

Allocution des récipiendaires du Prix de La revue canadienne de statistique

Special Session • Conférence spéciale

Organizer and Chair • Responsable et président : David R. BRILLINGER, *University of California at Berkeley*

### 08:30 – 10:00 Session 9B

AA-1360

Statistics and Science

La statistique et la science

Invited Paper Session • Présentations sur invitation

Organizer and Chair • Responsable et présidente : Agnes M. HERZBERG, *Queen's University*08:30 Lawrence MYSAK, *McGill University*

Climate Change: Where on Earth Are We Going? • Les changements climatiques : Mais où donc la Terre s'en va-t-elle ? [MS-149]

09:00 William ALLEN, *Michigan State University*

Political Arithmetic: Social Science, Scientific Revolution, and Political Founding • Le calcul politique : sciences sociales, la révolution scientifique et fondations politiques [MS-150]

09:30 Murray MCQUIGGE, *Public Health Consulting & University of Western Ontario*

Statistical Sleuthing in Walkerton • Investigation statistique à Walkerton [MS-151]

### 08:30 – 10:00 Session 9C

Z-310

New Researchers and New Directions for Research in Biostatistics

Nouveaux chercheurs et nouvelles perspectives de recherche en biostatistique

Invited Paper Session • Présentations sur invitation

Session sponsored by the *Biostatistics Section*Séance parrainée par le *Groupe de biostatistique*Organizer and Chair • Responsable et président : Rafal KUSTRA, *University of Toronto*

08:30 Timothy RAMSAY, *University of Ottawa/Université d'Ottawa*, George DOUGLAS & Carole YAUK, *Health Canada/Santé Canada*

New Challenges from Smaller Microarrays • Nouveaux défis provenant de microréseaux plus petits [MS-152]

09:00 Jonathan TAYLOR, Rob TIBSHIRANI, Trevor HASTIE & Guenther WALTHER, *Stanford University*

Connections Between Stagewise Algorithms and the LASSO • Connexions entre quelques algorithmes infinitésimaux et le LASSO [MS-153]

09:30 Rafal KUSTRA, *University of Toronto*

Multivariate Analysis of Microarray Expression Data • Analyse multivariée de données d'expression microréseaux [MS-154]

### 08:30 – 10:00 Session 9D

**Z-330**

#### Non-traditional Survey Methods

#### Méthodes d'enquête non traditionnelles

Invited Paper Session • Présentations sur invitation

Session sponsored by the *Association des statisticiennes et des statisticiens du Québec*

Séance parrainée par l'*Association des statisticiennes et des statisticiens du Québec*

Organizer and Chair • Responsable et président : Mike SIROIS, *Statistics Canada/Statistique Canada*

08:30 Diane LEROUX, *Ministère des Transports du Québec*

Sampling of Images Collected by Digital Camera for Estimating the Level of Cracking on the Quebec Roadways • Échantillonnage d'images captées par caméra vidéo pour l'estimation de l'étendue de la fissuration sur les chaussées du Québec [MS-155]

09:00 Éric PELLETIER, *Statistics Canada/Statistique Canada*

Modeling and Increase of Use of the Tax Data for the Unified Enterprise Survey • Modélisation et utilisation accrue des données fiscales dans le cadre de l'Enquête unifiée sur les entreprises [MS-156]

09:30 Tony LABILLOIS, *Statistics Canada/Statistique Canada*, Judy LEE, *Statistique Canada/Statistique Canada*

Research on our Data Collection Activities; a Way to Better Understand and Improve Our Work • La recherche sur nos activités de collecte ; une façon de mieux comprendre et d'améliorer notre travail [MS-157]

### 08:30 – 10:00 Session 9E

**AA-1355**

#### Stochastic Processes and Their Applications

#### Processus stochastiques et leurs applications

Contributed Paper Session • Séance de communications libres

Chair • Président : Jean-François QUESSY, *Université du Québec à Trois-Rivières*

08:30 Nikolaos LIMNIOS, *Université de Technologie de Compiègne*, Valérie GIRARDIN, *Université de Caen*

Entropy and Semi-Markov Processes • Entropie et processus semi-markoviens [MS-158]

08:45 Rafal KULIK, *University of Ottawa/Université d'Ottawa*, Hans DADUNA & Cornelia SAUER, *Universität Hamburg*, Ryszard SZEKLI, *Uniwersytet Wrocławski*

Dependence Ordering of Markov Processes and Queueing Networks • Dépendance d'ordre de processus de Markov et de réseaux de files d'attentes [MS-159]

09:00 Jerzy FILUS, *Oakton Community College, DesPlaines, IL*, Lidia Z. FILUS, *Northeastern Illinois University*

Construction of New Classes of Pseudonormal Markovian Stochastic Processes • Construction de nouvelles classes de processus stochastiques markoviens pseudo-normaux [MS-160]

09:15 Joykrishna SARKAR & Arusharka SEN, *Concordia University*

Weak Convergence Approach to Compound Poisson Processes Perturbed by Diffusion • Convergence faible pour des processus de Poisson composés perturbés par un facteur de diffusion [MS-161]

09:30 David THOMSON, *Queen's University*

A Test for “Long-Memory” Processes • Un test pour les processus à longue mémoire [MS-162]

09:45 Andrei VOLODIN, *University of Regina*

On the Rate of Convergence of the Dependent Bootstrapped Means • Sur le taux de convergence des moyennes bootstrap dépendantes [MS-163]

## 08:30 – 10:00 Session 9F

**Z-317**

### Survival Analysis II

### Analyse de durées de vie II

Contributed Paper Session • Séance de communications libres

Chair • Président : David E. MATTHEWS, *University of Waterloo*

08:30 Xuecheng LIU & Alain C. VANDAL, *McGill University*, Robert GENTLEMAN, *Harvard University*

Bounds on Self-consistent Estimates of Survival with Multivariate Interval Censored Data • Bornes sur les estimations autoconvergentes de la survie avec des données censurées par intervalle multivariées [MS-164]

08:45 Marc FREDETTE & Jerry LAWLESS, *University of Waterloo*

Prediction of Recurrent Events • Prédiction d'événements récurrents [MS-165]

09:00 Jean-Marie BOHER & Richard COOK, *University of Waterloo*

Issues in Testing and Model Misspecification Involving Multiple Event Data • Problèmes des méthodes de comparaison et de l'énoncé erroné d'un modèle d'analyse des délais d'apparition d'un évènement multiple [MS-166]

09:15 Rinku SUTRADHAR & Richard COOK, *University of Waterloo*

Bivariate Truncated Failure Time Data • Données de temps de bris tronquées bivariées [MS-167]

## 10:00 – 10:30 Coffee Break • Pause-café

**Pavillon André-Aisenstadt**

**10:30 – 12:00 Session 10A****AA-1360**Statistical Hydrology  
Hydrologie statistique

Public Session • Conférence publique

Chair • Président : Jean-François ANGERS, *Université de Montréal*Organizers • Responsables : Jean-François ANGERS, *Université de Montréal*, LUC PERRAULT, *Institut de recherche d'Hydro-Québec*10:30 Vincent FORTIN, *Institut de recherche d'Hydro-Québec*

On the Contribution of Statistics to Water Resources Management for Production • Sur l'apport de la statistique à la gestion des ressources en eau à des fins de production hydroélectrique [MS-168]

10:45 James MERLEAU & Jean-François ANGERS, *Université de Montréal*, Luc PERREAULT, *Institut de recherche d'Hydro-Québec*

Bayesian Modelisation of Hydrographs • Modélisation bayésienne d'hydrogrammes [MS-169]

11:10 Luc PERREAULT, Vincent FORTIN & Jocelyn GAUDET, *Institut de recherche d'Hydro-Québec*, Oli SVEINSSON, *Columbia University*

Seasonal Hydrological Forecast: Some Statistical Models Currently Used at Hydro-Québec • Prévission hydrologique saisonnière : quelques modèles statistiques employés à Hydro-Québec [MS-170]

11:35 Anne-Catherine FAVRE & Salaheddine EL ADLOUNI, *Institut national de la recherche scientifique—Eau, Terre & Environnement*, Luc PERREAULT, *Institut de recherche d'Hydro-Québec*

What is the Future of Copulas in Hydrology? • Quel futur pour les copules en hydrologie? [MS-171]

**10:30 – 12:00 Session 10B****AA-1140**Empirical Likelihood Methods in Survey Sampling  
Méthodes empirico-vraisemblantistes en méthodologie d'enquête

Invited Paper Session • Présentations sur invitation

Session sponsored by the *Survey Methods Section*Séance parrainée par le *Groupe de méthodologie d'enquête*Chair • Président : Éric RANCOURT, *Statistics Canada/Statistique Canada*Organizers • Responsables : Changbao WU, *University of Waterloo*10:30 J. N. K. RAO, *Carleton University*

Empirical Likelihood Methods for Sample Survey Data: An Overview • Méthodes de vraisemblance empirique pour des données de sondage : un survol [MS-172]

11:00 Jiahua CHEN, *University of Waterloo*

Confidence Intervals for the Mean of a Population Containing Many Zero Values Under Unequal Probability Sampling • Intervalles de confiance pour la moyenne d'une population contenant plusieurs zéros sous un plan d'échantillonnage à probabilités inégales [MS-173]

11:30 Changbao WU, *University of Waterloo*

Some Algorithmic Aspects of the Empirical Likelihood Method in Survey Sampling • Quelques aspects algorithmiques de la méthode de vraisemblance empirique en échantillonnage [MS-174]

**10:30 – 12:00 Session 10C**

**Z-310**

Parametric Survival Analysis

Analyse paramétrique de durées de vie

Invited Paper Session • Présentations sur invitation

Organizer and Chair • Responsable et présidente : Judy-Anne CHAPMAN, *University of Toronto*

10:30 Jerry LAWLESS, *University of Waterloo*, W. HE, *Samuel Lunenfeld Research Institute*

Multivariate Accelerated Failure Time Regression Analysis • Analyse de régression multivariée de temps de défaut accéléré [MS-175]

11:00 Judy-Anne CHAPMAN, *Women's College Hospital & University of Waterloo*

Over 50 Years of Log-normal Model use in Breast Cancer Investigations • Plus de 50 ans d'utilisation du modèle log-normal dans les recherches sur le cancer du sein [MS-176]

11:30 Patricia TAI, *University of Saskatchewan*

Evidence of Applicability of the Log-normal to Other Cancers, and Clinical Usefulness • Évidence de l'applicabilité du modèle log-normal à d'autres types de cancer et utilité clinique [MS-177]

**10:30 – 12:00 Session 10D**

**Z-330**

Case Study II: Anorexia Nervosa

Étude de cas II : Anorexie mentale

Organizer and Chair • Responsable et présidente : Peggy T. NG, *York University*

10:30 Peggy NG, *York University*

Presentation of the Data Set • Présentation du jeu de données [MS-178]

10:40 Omar EL BIDANI, Jean-Philippe GATARAYIHA & Rose Carline EVRA, *Université de Montréal*

**10:30 – 12:00 Session 10E**

**AA-1355**

Statistical Inference

L'inférence

Contributed Paper Session • Séance de communications libres

Chair • Président : Martin BILODEAU, *Université de Montréal*

10:30 Shahadut HOSSAIN & Paul G. GUSTAFSON, *University of British Columbia*

Conservative Priors for Hierarchical Models • Densités *a priori* conservatrices pour les modèles hiérarchiques [MS-179]

10:45 Matias SALIBIAN-BARRERA & Azadeh ALIMADAD, *Carleton University*

A Robust Fit for Generalised Additive Models • Un ajustement robuste pour des modèles additifs généralisés [MS-180]

- 11:00 Debbie DUPUIS, *University of Western Ontario*, Joanna MILLS-FLEMMING, *Université of Genève*  
Fitting Nonlinear Mixed Effects Models Robustly • Ajustement robuste de modèles non linéaires à effets mixtes [MS-181]
- 11:15 Ka Lok CHU, *McGill University*, J. ISOTALO & S. PUNTANEN, *University of Tampere, Finland/Université de Tampere, Finlande*, G. P. H. STYAN, *McGill University*  
On Decomposing the Watson Efficiency of OLS in a Partitioned Weakly Singular Linear Model  
• Sur la décomposition de l'efficacité de Watson des MCO dans un modèle linéaire partitionné faiblement singulier [MS-182]
- 11:30 Thu PHAM-GIA & N. TURKKAN, *Université de Moncton*  
Predictive Analysis of Stress-strength Reliability Models • Analyse prédictive du modèle de fiabilité contrainte-force [MS-183]

**10:30 – 12:00 Session 10F****Z-317****Biostatistics II****Biostatistique II**

Contributed Paper Session • Séance de communications libres

Chair • Présidente : Jeanette O'HARA-HINES, *University of Waterloo*

- 10:30 Mu ZHU & Fuping HUANG, *University of Waterloo*

The Evaluation of Information Retrieval Systems with Applications to Drug Discovery • Un paradigme de filtrage séquentiel pour détecter des items rares dans une grande base de données avec des application en découverte de médicaments [MS-184]

- 10:45 Xiaoqiang LI, *University of Toronto*

Using Multilevel Modeling to Study the Patient Care Delivery System • Modélisation multi-niveaux pour étudier le système de soins de santé [MS-185]

- 11:00 Cathlin MCNALLY, Lehana THABANE & Anne HOLBROOK, *McMaster University*, Kathy GAEBEL, *St. Joseph's Healthcare, Hamilton*

Data Mining Electronic Medical Records • Forage de données de dossiers médicaux électroniques [MS-186]

- 11:30 Elizabeth TURNER & Alain C. VANDAL, *McGill University*

A Covariate-based Coefficient of Source Dependence for Capture-recapture Models Applied to Epidemiologic Studies • Un coefficient de dépendance de source basé sur les covariables pour des modèles de capture-libération appliqués aux études épidémiologiques [MS-187]

- 11:45 J. LOREDO-OSTI, N. ROSLIN, & K. MORGAN, *McGill University*, A. BONNARDEAUX, *Université de Montréal*

Segregation of Quantitative Traits with Covariates • Ségrégation de traits quantitatifs avec des covariables [MS-188]

**12:00 – 13:30 Lunch • Dîner**



**13:30 – 15:00 Session 11A****Z-310**

## Health Surveys

## Les enquêtes dans le domaine de la santé

Invited Paper Session • Présentations sur invitation

Session sponsored by the *Biostatistics and Survey Methods Sections*Séance parrainée par les *Groupes de biostatistique et de méthodologie d'enquête*Organizer and Chair • Responsable et présidente : Sue BONDY, *University of Toronto*

13:30 Elizabeth POTTER, *University of Ottawa/Université d'Ottawa*, Kathy N. SPEECHLEY, *Child Health Research Institute, London, ON*, John J. KOVAL, *University of Western Ontario*, Iris A. GUTMANIS, *Southwest Region Health Information Partnership, London, ON*, M. Karen CAMPBELL, *University of Western Ontario*, Douglas MANUEL, *Institute for Clinical Evaluative Sciences, Toronto*

Socioeconomic Status and Nonfatal Injuries Among Adolescents: Variations Across SES and Injury Measures • Statut socio-économique et les blessures non fatales chez les adolescents : variations à travers les mesures de SSE et les mesures de blessures [MS-189]

14:15 Anne RHODES, *St. Michael's Hospital, Toronto*

The Linkage of Household Survey Data with Administrative Records — A Mental Health Services Example from Ontario • Le jumelage de données de sondage sur les ménages avec des dossiers administratifs : un exemple sur les services de santé mentale en Ontario [MS-190]

**13:30 – 15:00 Session 11B****AA-1360**

## Non-iid Sampling Data

## Données d'échantillonnage non iid

Invited Paper Session • Présentations sur invitation

Organizer and Chair • Responsable et président : Pierre DUTILLEUL, *McGill University*

13:30 Gülhan ALPARGU, *University of Massachusetts*, Pierre DUTILLEUL, *McGill University*

About Efficient Estimation and Valid and Powerful Testing in Quantitative Linear Models with Autocorrelated Errors • A propos de l'efficacité d'estimateurs et de la validité et la puissance de tests dans les modèles linéaires quantitatifs avec erreurs autocorrélées [MS-191]

13:52 Pierre DUTILLEUL, *McGill University*

To Be or not to Be i.i.d.? A Very Good Question for a Sample of Data • Être ou ne pas être i.i.d. ? Une très bonne question pour un échantillon de données [MS-192]

14:14 Bernard PELLETIER, *McGill University*

Assessing Relationships Between Multivariate Spatial Data at Multiple Scales • L'évaluation de relations multiéchelles avec données spatiales multivariées [MS-193]

14:36 Hans-Peter PIEPHO & K. EMRICH, *Universität Hohenheim*

Accounting for Genetic and Environmental Sources of Correlation in QTL Mapping Studies • Tenir compte des sources de corrélations génétiques et environnementales dans les études de mapping du LTQ [MS-194]

**13:30 – 15:00 Session 11C****AA-1140**

## Analysis of Longitudinal Data

## Analyse de données longitudinales

Invited Paper Session • Présentations sur invitation

Organizer and Chair • Responsable et président : Lang WU, *University of British Columbia*13:30 Richard COOK & Leilei ZENG, *University of Waterloo*

Transition Models for Multivariate Longitudinal Binary Data • Modèles de transition pour des données binaires longitudinales multivariées [MS-195]

14:00 Hulin WU, *University of Rochester*, Jeong-Gun PARK, *Frontier Science and Technology Research Foundation, Inc.*

Backfitting and Local Likelihood Methods for Nonparametric Mixed-effects Models with Longitudinal Data • Backfitting et méthodes de vraisemblance locale pour des modèles non paramétriques à effets mixtes avec des données longitudinales [MS-196]

14:30 Rachel MacKay ALTMAN & Brian LEROUX, *University of Washington*

Robust Estimation of Parameter-driven Models for Longitudinal Count Data • Estimation robuste de modèles guidés par des paramètres pour des données de dénombrement longitudinales [MS-197]

**13:30 – 15:00 Session 11D****Z-330**

## Statistics and the Physical Sciences

## La statistique et les sciences physiques

Invited Paper Session • Présentations sur invitation

Session sponsored by the *Business and Industrial Statistics Section*Séance parrainée par le *Groupe de statistique industrielle et de gestion*Organizer and Chair • Responsable et président : Duncan J. MURDOCH, *University of Western Ontario*13:30 Kim MCAULEY, *Queen's University*

Fundamental Modeling of Chemical Processes—Statistical Challenges and Approaches • Modélisation fondamentale de processus chimiques, défis et approches statistiques [MS-198]

14:00 William JEFFERYS & Thomas R. JEFFERYS, *University of Texas at Austin*, Thomas G. BARNES, *McDonald Observatory*

Bayesian Analysis of RR Lyrae Distances and Kinematics • Analyse bayésienne des distances et cinématiques des étoiles de RR Lyrae [MS-199]

14:30 Peter KIM, *University of Guelph*

Multivariate Deconvolution and Applications • Déconvolution multivariée et applications [MS-200]

**13:30 – 15:00 Session 11E****Z-317**

## Testing Procedures

## Procédures de test

Contributed Paper Session • Séance de communications libres

Chair • Président : Donald L. MCLEISH, *University of Waterloo*13:30 Richard LOCKHART & Michael A. STEPHENS, *Simon Fraser University*

Power Similarity of Two Groups of Tests for the Exponential Distribution • Similarité des puissances de deux groupes de tests pour la distribution exponentielle [MS-201]

13:45 Steve OVERDUIN & Michael A. STEPHENS, *Simon Fraser University*

Clustering of Goodness-of-fit Tests • Groupement des tests d'adéquation [MS-202]

14:00 Aurélie LABBE & Mary E. THOMPSON, *University of Waterloo*

Multiple Testing in a Hierarchical Framework • Tests multiples dans un contexte bayésien [MS-203]

14:15 Ana-Maria STAIKU & Nancy REID, *University of Toronto*

Accuracy of Different Methods when Testing a Binomial Proportion • Précisions de différentes méthodes pour tester une proportion binomiale [MS-204]

14:30 Keyue DING, *Queen's University*

Point Estimation Following a Group Sequential Test • Estimation ponctuelle suivant un test séquentiel par groupes [MS-205]

14:45 Debaraj SEN & Yogendra CHAUBEY, *Concordia University*

Preliminary Test Estimation in Inverse Gaussian Data • Estimation préliminaire pour des données gaussiennes inverses [MS-206]

**13:30 – 15:00 Session 11F****AA-1355**

## Epidemiology

## Épidémiologie

Contributed Paper Session • Séance de communications libres

Chair • Président : Alexandre BUREAU, *University of Lethbridge*13:30 Lehana THABANE, D. A. MARSHALL & J. K. MARSHALL, *McMaster University*, F. R. JOHNSON, *Research Triangle Institute, NC*, K. A. PHILLIPS, *University of California, San Francisco*, G. FOSTER, *McMaster University*, B. J. O'BRIEN, *McMaster University*

Modeling Patient Preferences for Colorectal Cancer (CRC) Screening Strategies • Modélisation des préférences des patients pour des stratégies de dépistage du cancer colorectal (CCR) [MS-207]

13:45 Yuanyuan LIANG & Keumhee Chough CARRIÈRE, *University of Alberta*

Stratified and Randomized Play-the-winner Rule for Designs in Clinical Trials • Règle miser-sur-le-gagnant stratifiée et randomisée pour des designs d'études cliniques [MS-208]

14:00 Aicha GOUBAR & A. E. ADES, *Medical Research Council—Health Services Research Collaboration, Bristol*, D. DE ANGELIS, *Medical Research Council—Biostatistics Unit, Cambridge*, O. N. GILL, *Health Protection Agency—Communicable Disease Surveillance Centre, London*

Bayesian Synthesis of Information in a Complex Evidence Structure: HIV Surveillance in England and Wales, 2001 • Analyse bayésienne de données à structure complexe : surveillance du VIH en Angleterre et au pays de Galles, 2001 [MS-209]

14:15 Lawrence MCCANDLESS & Paul G. GUSTAFSON, *University of British Columbia*, A. R. LEVY, *University of British Columbia & St. Paul's Hospital*

Assessing Sensitivity to Unobserved Confounding in Observational Studies: A Bayesian Approach • Évaluation de la sensibilité des facteurs parasites non observés dans des études d'observation : une approche bayésienne [MS-210]

**13:30 – 15:00 Session 11G**

**Z-240**

Panel Discussion on the Mathematical Sciences Institutes and Programs: As a Statistician, Should I be Concerned?

Débat sur les instituts et les programmes en sciences mathématiques : En tant que statisticien, suis-je concerné ?

Special Session • Conférence spéciale

Organizer and Chair • Responsable et président : Christian LÉGER, *Université de Montréal*

13:30 Christian LÉGER, *Centre de recherches mathématiques*

13:40 Thomas SALISBURY, *Fields Institute*

13:50 Manfred TRUMMER, *Pims*

14:00 Arvind GUPTA, *MITACS*

14:10 Jamie STAFFORD, *NPCDS*

14:20 Floor discussion • Discussion plénière

**15:00 – 15:30 Coffee Break • Pause-café**

**Pavillon André-Aisenstadt**

**15:30 – 17:00 Session 12A**

**AA-1140**

Address of the 2004 CRM-SSC Prize Winner

Allocution du lauréat du Prix CRM-SSC 2004

Special Session • Conférence spéciale

Organizer and Chair • Responsable et président : Christian GENEST, *Université Laval*

**15:30 – 17:00 Session 12B**

**Z-330**

Nonparametric and Bayesian Methods

Méthodologie bayésienne et non paramétrique

Contributed Paper Session • Séance de communications libres

Chair • Président : Belkacem ABDOUS, *Université Laval*

- 15:30 Xiaobin YUAN & D. A. S. FRASER, *University of Toronto*  
 Noninformative Priors Based on Asymptotic Likelihood Methods • Densités *a priori* non informatives basées sur des méthodes de vraisemblance asymptotiques [MS-211]
- 15:45 Raluca BALAN, *University of Ottawa/Université d'Ottawa*  
 A New Class of Prior Distributions in Bayesian Nonparametric Statistics • Une nouvelle classe de distributions *a priori* en statistique bayésienne non paramétrique [MS-212]
- 16:00 Denis LAROCQUE, *HEC Montréal*  
 The Wilcoxon Signed-rank Test for Cluster Correlated Data • Le test de Wilcoxon des rangs signés pour données corrélées en grappes [MS-213]
- 16:15 Jean-François PLANTE, *University of British Columbia*, Christian GENEST, *Université Laval*  
 On Blest's Correlation Coefficient • À propos du coefficient de corrélation de Blest [MS-214]
- 16:30 Jeffrey BAKAL, Glen TAKAHARA, J. Terry SMITH & Joan STEVENSON, *Queen's University*  
 Clustering for Functional Biomechanical Data Using a Model-based Distance Measure • Classification de données fonctionnelles biomécaniques en utilisant une mesure de distance basée sur un modèle [MS-215]
- 16:45 Pierre DUCHESNE, *Université de Montréal*, Maria PACURAR, *HEC Montréal*  
 Tests for Duration Clustering and Diagnostic Checking of ACD Models Using Kernel-based Spectral Density Estimators • Tests pour effets ACD et évaluation de l'adéquation des modèles ACD avec des estimateurs à noyau de la densité spectrale [MS-216]

**15:30 – 17:00 Session 12C****AA-1360**

## Estimation

## Estimation

Contributed Paper Session • Séance de communications libres

Chair • Présidente : Brenda MACGIBBON, *Université du Québec à Montréal*

- 15:30 Sévérien NKURUNZIZA & Sorana FRODA, *Université du Québec à Montréal*  
 Estimating the Parameters of Some Ecological Systems • Estimation des paramètres de certains systèmes écologiques [MS-217]
- 15:45 Éric MARCHAND, *University of New Brunswick*  
 Estimating a Bounded Location Parameter: Recent Results • Estimation d'un paramètre de position borné : résultats récents [MS-218]
- 16:00 Mohammed HADDOU & François PERRON, *Université de Montréal*  
 Estimation of a cdf and its Derivative • Estimation de la fonction de répartition et de sa dérivée [MS-219]
- 16:15 Saumen MANDAL, *University of Manitoba*  
 Construction of Optimizing Distributions with Applications in Estimation • Construction de distributions d'optimisation avec des applications en estimation [MS-220]

16:30 Shoja'eddin CHENOURI, *University of Waterloo*

Multivariate Generalization of R-estimators Based on Data Depth • Généralisation multivariée des estimateurs R basée sur la profondeur des données [MS-221]

16:45 Ali Reza FOTOUHI, *University College of the Fraser Valley*

Comparisons of Estimation Procedures for Nonlinear Multilevel Models • Comparaisons des procédures d'estimation pour des modèles multi-niveaux non linéaires [MS-222]

### 15:30 – 17:00 Session 12D

**Z-317**

Experimental Design and Related Issues

Planification d'expérience et sujets connexes

Contributed Paper Session • Séance de communications libres

Chair • Président : Douglas P. WIENS, *University of Alberta*

15:30 Stefan STEINER & R. Jock MACKAY, *University of Waterloo*

Statistical Engineering: A New Algorithm for Reducing Variation in Manufacturing Processes • L'ingénierie statistique : un nouvel algorithme pour réduire la variation dans les processus de manufacture [MS-223]

15:45 Asokan MULAYATH VARIYATH, Bovas ABRAHAMN & Jiahua CHEN, *University of Waterloo*

Analysis of Factorial Experiments with Concurrent Measures Using Jackknife • Analyse d'expériences à plan factoriel avec des mesures simultanées en utilisant le jackknife [MS-224]

16:00 Mikelis BICKIS, *University of Saskatchewan*

Eigenvalues in Optimal Design: Information Matrix vs Information Metric • Les valeurs propres dans les plans optimaux : matrice d'information vs métrique d'information [MS-225]

16:15 Adeniyi ADEWALE & Douglas P. WIENS, *University of Alberta*

Robust Designs for Approximate Generalized Linear Models • Designs robustes pour des modèles linéaires généralisés approximatifs [MS-226]

16:30 Yongge TIAN & A. M. HERZBERG, *Queen's University*

Optimal Designs for Linear Haar-Wavelet Models • Designs optimaux pour les modèles linéaires d'ondelettes de Haar [MS-227]

16:45 G. Arthur MIHRAM, *Princeton, NJ*, Danielle MIHRAM, *University of Southern California*

Correction of an Erroneous Assertion of Margolin et al. • Correction d'une affirmation erronée de Margolin et co-auteurs [MS-228]

### 15:30 – 17:00 Session 12E

**AA-1355**

Various Topics

Thèmes divers

Contributed Paper Session • Séance de communications libres

Chair • Président : Mhamed MESFIOUI, *Université du Québec à Trois-Rivières*

15:30 Jianmin PAN & Jiahua CHEN, *University of Waterloo*

Information Criteria and Change Point Problems • Critères d'information et problème du point de changement [MS-229]

15:45 Henrik STRYHN, *University of Prince Edward Island*, Keith E. BAPTISTE, *University of Liverpool*

A Hierarchical Latent Class Model for Comparing Individual and Composite Tests • Model hiérarchique de classe latente pour comparer des tests individuels et composés [MS-230]

16:00 Jeffrey PICKA, *University of New Brunswick*

Descriptors of Large-scale Structures in Random Sets • Descripteurs de la structure à grande échelle dans des ensembles aléatoires [MS-231]

16:15 Carles CUADRAS & Daniel CUADRAS, *Universitat de Barcelona*

Some Aspects of the First Principal Dimension of a Random Variable • Quelques aspects de la première dimension principale d'une variable aléatoire [MS-232]

## 15:30 – 17:00 Session 12F

**Z-310**

### Biostatistics III

### Biostatistique III

Contributed Paper Session • Séance de communications libres

Chair • Présidente : Pascale ROUSSEAU, *Université du Québec à Montréal*

15:30 Nikolaos LIMNIOS, *Université de Technologie de Compiègne*, Brahim OUHBI, *École Nationale Supérieure d'Arts et Métiers, Meknès, Maroc*

Nonparametric Estimation of Reliability for Semi-Markov Processus Based on  $N$  Histories • Estimation non paramétrique de la fiabilité de processus semi-markoviens basée sur  $N$  histoires [MS-233]

15:45 Jin QIAN & Richard COOK, *University of Waterloo*

Parametric and Nonparametric Estimation of Current Status Data with a Nonsusceptible Fraction • Estimation paramétrique et non paramétrique de données d'état actuel avec une fraction non susceptible [MS-234]

16:00 Chul PARK & Patrick J. FARREL, *Carleton University*

Testing for Ordered Group Effects in Bivariate Dichotomous and Continuous Outcomes • Tester l'effet d'ordre de groupe dans des résultats bivariés dichotomiques et continus [MS-235]

16:15 Narasimha PRASAD, *University of Alberta*

Bootstrap Method for Classification Using Random Effects Model • Méthode bootstrap pour la classification en utilisant des modèles à effets aléatoires [MS-236]

# Abstracts • Résumés

The translation team did its best. Please forgive it for any misinterpretation; its members are neither professional translators nor specialists in all areas of statistics.

L'équipe de traduction a fait de son mieux. Elle s'excuse pour toute erreur de sens ; ses membres ne sont ni des traducteurs professionnels ni des spécialistes des différents domaines de la statistique.

**Session 0A      Sunday, May 30 • Dimanche 30 mai, 17:30 – 18:30      AA-1140**

**Accreditation of Professional Statisticians in Canada**

**L'accréditation des statisticiens professionnels au Canada**

[MS-1]

**ASSQ, Accreditation and “Statistician” Membership Category**

**ASSQ, membre statisticien et accréditation**

Sylvain VÉGIARD, *Institut de la statistique du Québec*

Since its foundation in 1995, the Association des statisticiennes et des statisticiens du Québec (ASSQ) has two categories of members: the statistician member and the affiliated member. The affiliated member is any person interested in statistics to the point of becoming a member of ASSQ without necessarily having training or a diploma in statistics. This category is thus opened to the entire population. On the other hand, the statistician member must have at least a bachelor's degree in statistics or any other diploma from a recognized three year university program (normal duration of undergraduate studies in Québec) having at least a minimum of 24 units of statistics or probability courses among courses recognized for this purpose. To obtain membership as a statistician member within the ASSQ, the person must prove that he/she has the necessary educational background. If a person has at least a bachelor but does not fill the criteria associated with the statistician membership, it is also possible for him/her to become a statistician member through an accreditation process. The applicant must demonstrate to the accreditation committee that the professional work carried out during his/her career compensates for the lack of formal training in university. Either through academic training or through the accreditation process, the statistician member is finally recognized as having a professional level equivalent

Depuis sa fondation en 1995, l'Association des statisticiennes et statisticiens du Québec (ASSQ) compte, entre autres, deux catégories de membres : le membre statisticien et le membre affilié. Le membre affilié est ni plus ni moins toute personne intéressée à la statistique au point de vouloir faire partie de l'ASSQ, sans toutefois avoir nécessairement un diplôme ou une certaine formation en statistique. Cette catégorie est donc ouverte à l'ensemble de la population. Pour sa part, le membre statisticien doit posséder au moins un baccalauréat en statistique ou tout autre diplôme d'un programme universitaire reconnu d'au moins trois années et comportant un minimum de vingt-quatre crédits de cours de statistique ou de probabilités parmi les cours reconnus à cet effet. Lors de l'adhésion comme membre statisticien au sein de l'ASSQ, la personne doit faire la preuve qu'elle possède la scolarité requise. Si une personne possède au moins un baccalauréat mais ne remplit pas les critères associés au membre statisticien, il lui est aussi possible de devenir membre statisticien en suivant un processus d'accréditation. Selon ce processus, l'aspirant doit démontrer au comité d'accréditation que le travail professionnel effectué au cours de sa carrière compense les connaissances acquises lors de la formation académique. Que ce soit suite à sa formation académique, ou suite à un processus d'accréditation, le membre statisticien est finalement reconnu comme ayant un niveau professionnel équivalent au niveau A.Stat. développé par la Société statistique du Canada avec la participation de l'ASSQ. Durant cette présentation, nous ferons un bref historique de l'ASSQ. Nous décrirons par la suite ses



lent to the A.Stat. level developed by the Statistical Society of Canada with the participation of the ASSQ. During this presentation, we will give a short history of the ASSQ. We will then describe its different categories of membership and give a detailed description of the accreditation process of the ASSQ.

catégories de membres pour finalement donner une description détaillée du processus d'accréditation de l'ASSQ.

### [MS-2]

#### **Educational Requirements for A.Stat. and P.Stat** **Formation académique requise pour les titres A.Stat et P.Stat**

Carl SCHWARZ, *Simon Fraser University*

An accredited statistician has completed a course of study equivalent to a major or honours degree in statistics, or in exceptional instances, has otherwise demonstrated an advanced understanding of statistical theory and its application. There is a plan to eventually accredit university programs. In the meantime, using the substantive ASA work on guidelines for an undergraduate program in statistics, effective statisticians at any level display a combination of skills that are not exclusively mathematical. Programs should provide some background in the following areas: statistics, mathematics, computation, non-mathematical communication skills, and substantive area(s) of application.

Un statisticien accrédité a complété une formation académique équivalente à une majeure ou un baccalauréat en statistique ou, dans certains cas exceptionnels, a démontré une compréhension avancée de la statistique et de ses applications acquise d'une autre façon. Il est également prévu d'éventuellement accréditer des programmes universitaires. Par ailleurs, en utilisant le travail substantiel de l'ASA sur l'élaboration de directives pour un programme de premier cycle en statistique, il est bien connu que, quelque soit le niveau d'éducation, des statisticiens efficaces doivent détenir une combinaison de qualifications qui ne sont pas exclusivement mathématiques. Les programmes devraient fournir une certaine formation dans les secteurs suivants : statistique, mathématiques, informatique, communication non mathématique et plusieurs domaines d'applications.

### [MS-3]

#### **Professional Statistical Experience** **Expérience professionnelle en statistique**

Ken MCRAE, *Agriculture and Agri-Food Canada*, Fernando CAMACHO, *DAMOS Consulting*

A professional statistician (P.Stat.) assumes duties and responsibilities beyond that expected of a recent graduate or a junior statistician (A.Stat.). Simply stated, the professional statistician requires practical knowledge, experience, and expertise to assume the responsibility for all aspects of statistical work on projects and studies. While the specific details for accreditation will depend on the particular job, the following aspects are central to the work of professional statisticians. Accredited statisticians are expected to ensure good statistical practices for applications throughout: planning of projects and studies; data collection and management; analyses of data; synthesis of results; interpretation of data and results; and oral and written

Un statisticien professionnel (P.Stat.) s'acquitte de fonctions et de responsabilités au-delà de celles prévues d'un diplômé récent ou d'un statisticien junior (A.Stat.). De manière simplifiée, le statisticien professionnel doit conjuguer les connaissances pratiques, l'expérience et l'expertise pour assumer la responsabilité de tous les aspects du travail statistique dans les projets et les études. Alors que les détails spécifiques de l'accréditation dépendront du travail particulier, les aspects suivants sont fondamentaux au travail des statisticiens professionnels. On s'attend à ce que les statisticiens accrédités assurent de bonnes pratiques statistiques pour des applications en : planification des projets et des études ; collecte et gestion de données ; analyse des données ; synthèse des résultats ; interprétation des données et des résultats ; et communication orale et écrite des résultats des projets et des

communication about projects and studies. Statistical Accreditation will establish national standards for good statistical practice, and a code for ethical conduct. Specific examples of accruing professional experience will be provided in two application domains. Kenneth McRae will discuss statistical collaboration in agricultural and agri-food research within a governmental environment. Statistical consulting within industry and for pharmaceutical companies will be discussed by Fernando Camacho.

**[MS-4]**

**Professional Development**

**Développement professionnel**

Judy-Anne CHAPMAN, *University of Toronto*

An integral part of SSC Accreditation is integration of statistical methods to areas of professional practice, e.g., biostatistics, financial statistics, industrial statistics, sample survey statistics. Individual practitioners will take responsibility for deciding what activities are most timely and relevant for maintaining competence. Canadian society will benefit from this integration. An example of specialization will be provided in the area of clinical breast cancer research, with the demonstration of a survival benefit from the routine clinical use of mammography for women less than 40, 40-49, 50-69, and those of more than 69 years of age.

études. L'accréditation statistique établira des normes nationales pour la bonne pratique statistique et un code d'éthique. Des exemples spécifiques d'une expérience professionnelle seront fournis dans deux domaines d'application. Kenneth McRae discutera de l'apport de la statistique en recherche agricole et agroalimentaire dans un environnement gouvernemental. Par la suite, Fernando Camacho discutera de la consultation statistique dans l'industrie et pour les compagnies pharmaceutiques.

Une partie intégrale de l'accréditation à la SSC est l'intégration des méthodes statistiques dans des secteurs de pratique professionnelle, par exemple la biostatistique, la statistique en finance, la statistique industrielle et l'échantillonnage. Les différents praticiens prendront la responsabilité de décider quelles activités sont les plus opportunes et appropriées pour maintenir la compétence. La société canadienne tirera bénéfice de cette intégration. Un exemple de spécialisation sera fourni dans le domaine de la recherche clinique sur le cancer du sein en montrant l'avantage sur la survie de l'utilisation clinique courante de la mammographie pour les femmes de moins de 40 ans, 40-49 ans, 50-69 ans et celles de plus de 69 ans.

**Session 0B Sunday, May 30 • Dimanche 30 mai, 18:30 – 22:00 Hall d'honneur**

Contributed Poster Session and Cocktail (evening)

Présentation d'affiches et coquetel (soirée)

**[MS-5]**

**Statistics, Geometry and Brain Mapping**

**Statistique, géométrie et cartographie cérébrale**

Keith WORSLEY, *McGill University*

The geometry in the title is not the geometry of lines and angles but the geometry of topology, shape and knots. The Euler characteristic (EC) of the set of high regions is used to measure the topology of such shapes; it counts the number of connected components of the set, minus the number of 'holes,' plus the number of 'hollows.' Despite its complex definition, the exact expectation of the EC

La géométrie dans le titre n'est pas la géométrie de lignes et angles mais la géométrie de topologie, de forme et des noeuds. La caractéristique de Euler (EC) de l'ensemble des régions élevées est utilisée pour mesurer la topologie de telles formes; elle compte le nombre de pièces connectées de l'ensemble, moins le nombre de trous, plus le nombre de trous intérieurs. Malgré sa définition complexe, l'espérance exacte du EC peut être obtenue. Ceci permet aux chercheurs

can be found for some simple models. This allows researchers in human brain mapping to find the regions of the brain that are 'activated' by a stimulus.

utilisant la cartographie cérébrale de trouver les régions du cerveau qui sont 'activées' par un stimuli.

#### [MS-6]

### **Discrete Minimax Robust Designs for Regression Models with Autocorrelated MA Errors** **Plans minimax robustes discrets pour des modèles de régression avec des erreurs MA autocorrélées**

Julie ZHOU, Peilin SHI & Jane YE, *University of Victoria*

Minimax robust designs for regression models with possible misspecification in the response and possible autocorrelated errors are investigated on discrete design spaces. These designs minimize the maximum value of the trace of the mean squared error matrix, where the maximum is obtained over a class of departure functions from the regression response and a class of autocorrelated moving average error processes. Several examples are discussed to show robust designs for various regression models.

Des plans minimax robustes pour des modèles de régression contenant possiblement des erreurs de spécification dans la réponse et possiblement des erreurs autocorrélées sont étudiés sur des espaces de design discrets. Ces plans minimisent la valeur maximale de la trace de la matrice de l'erreur quadratique moyenne où le maximum est obtenu par rapport à une classe de fonctions d'éloignement par rapport à la réponse de régression et une classe de processus d'erreurs moyennes mobiles autocorrélées. Plusieurs exemples sont discutés pour montrer des plans robustes pour différents modèles de régression.

#### [MS-7]

### **Shift Function Plots for Regression Fitting** **Graphique de la fonction de décalage pour l'ajustement de régressions**

Zilin WANG & David BELLHOUSE, *University of Western Ontario*

Although nonparametric regression models allow us to obtain graphical displays of the relationship between response and independent variables, the exact form of the regression function is not evident. This weakness can be overcome by parametric models since the relationship between the dependent and the independent variables is specified mathematically. Here we introduce a new graphical approach called shift function plotting with which a hypothesis test is constructed to evaluate the goodness of fit of a parametric regression model. Inferential issues relating to the shift function are resolved. A chi-squared test statistic will also be developed with the aid of bootstrapping.

Bien que les modèles de régression non paramétrique nous permettent d'obtenir des représentations graphiques de la relation entre la variable réponse et les variables indépendantes, la forme exacte de la fonction de régression n'est pas toujours évidente. Cette faiblesse peut être surmontée par les modèles paramétriques puisque la relation entre les variables dépendantes et indépendantes est spécifiée mathématiquement. Ici nous présentons une nouvelle approche graphique appelée le graphique de la fonction de décalage à partir duquel un test d'hypothèse est construit pour évaluer la qualité de l'ajustement d'un modèle de régression paramétrique. Des problèmes d'inférence concernant la fonction de décalage sont résolus. Une statistique de test du khi-deux sera également développée à l'aide du bootstrap.

#### [MS-8]

### **Constructing Nonregular Robust Parameter Designs** **Construction de design non réguliers à paramètre robuste**

Jason LOEPPKY, Derek BINGHAM & Randy R. SITTE, *Simon Fraser University*

There has been considerable recent attention paid to robust parameter design as a strategy for variance reduction. Of particular concern is the selec-

Récemment, beaucoup d'attention a été portée aux designs à paramètre robuste comme stratégie pour réduire la variance. En particulier, le choix d'un bon plan d'expérience à

tion of a good experiment plan in light of the two different types of factors in the experiment (control and noise factors) and the asymmetric manner in which effects are treated. We consider the construction and selection of optimal nonregular experiment plans for robust parameter design. We demonstrate how one can greatly reduce the computational effort in obtaining minimum aberration robust parameter designs by using existing orthogonal arrays. Results are presented for 12, 16 and 20-run orthogonal arrays.

la lumière des deux types de facteurs dans l'expérience (les facteurs de contrôle et de bruit) et la manière asymétrique dont les effets sont traités ont reçu beaucoup d'intérêt. Nous considérons la construction et la sélection de plans d'expérience non réguliers optimaux pour le design de paramètre robuste. Nous démontrons comment nous pouvons considérablement réduire l'effort informatique en obtenant des designs à paramètre robuste avec un minimum d'aberrations en utilisant des tableaux orthogonaux existants. Des résultats sont présentés pour des tableaux orthogonaux à 12, 16 et 20 essais.

### [MS-9]

#### **An Analysis of a Bivariate Time Series in Which the Components Are Sampled at Different Instants**

#### **Analyse d'une série chronologique bivariée dans laquelle les composantes sont prélevées à différents instants**

David BRILLINGER, *University of California, Berkeley*

It is desired to express the relationship between the components of a bivariate time series. What is unusual is that the components are observed at different times and that the observation times are irregularly distributed. The research is motivated by data collected at a station along the Solimoes River in central Brazil and also at a second station along a branch of the Solimoes. Of interest to geographers is the possible change in the proportion of the Solimoes waters entering the branch. This is because an increased flow of Solimoes water into the branch might lead to the branch's widening and becoming the main stream. This could have substantial environmental effects.

Nous désirons exprimer la relation entre les composantes d'une série chronologique bivariée. Particularité peu commune, les composantes sont observées à des temps différents et les temps d'observation sont distribués de manière irrégulière. La recherche est motivée par des données recueillies à une station le long du fleuve Solimoes dans le centre du Brésil et aussi à une deuxième station le long d'un embranchement du Solimoes. Un des intérêts des géographes est le changement possible de la proportion des eaux du Solimoes qui entre dans l'embranchement car un plus grand débit d'eau du Solimoes entrant dans l'embranchement pourrait causer l'élargissement de ce dernier et devenir le courant principal. Ceci pourrait avoir des effets substantiels sur l'environnement.

### [MS-10]

#### **Small Sample Inference for Detecting Operational Confounders**

#### **Inférences sur de petits échantillons pour détecter des facteurs parasites opérationnels**

Philip MOORHEAD, *University of Western Ontario*

In epidemiologic studies concerned with causality, confounding is an important phenomena. An operational confounder is measured by the change in the effect of the risk factor when the confounder is removed. While large sample inference under a continuously measured outcome is well established, the methodology for detection in small samples is yet undeveloped. Assuming that both the outcome and the confounder are at least approximately normally distributed, inference for detection in small samples is developed. The methodology is based on a product distribution function that

Dans les études épidémiologiques concernées par la causalité, la présence de facteurs parasites est un phénomène important. Un facteur parasite opérationnel est mesuré par le changement dans l'effet du facteur de risque quand le facteur parasite est enlevé. Tandis que l'inférence avec de grands échantillons lorsque les résultats sont mesurés continuellement est bien établie, la méthodologie basée sur de petits échantillons est peu développée. En supposant que les résultats et le facteur parasite sont au moins approximativement normalement distribués, on développe l'inférence pour la détection avec de petits échantillons. La méthodologie est basée sur une fonction de distribution de produit qui se rap-

closely approximates the true distribution. A test procedure is proposed and evaluated through intensive simulation.

proche étroitement de la vraie distribution. Une méthode de test est proposée et évaluée par des simulations intensives.

[MS-11]

**Estimating the Rate of Concussions in British Columbia Minor Hockey Using Community Volunteer Collected Data**

**Estimation du taux de commotions cérébrales dans le hockey mineur en Colombie-Britannique en utilisant des volontaires dans la communauté pour cueillir les données**

Jeremy HAMM, Carl SCHWARZ & Ian WILLIAMSON, *Simon Fraser University*

The rate of concussions has been examined in elite levels of ice hockey but has yet to be studied in community, youth hockey where they are also thought to occur. Due to the relative rarity of the concussion, a large number of games need to be observed in order to gain a reliable estimate. This can become very costly if hired people are used to collect this data. This project assesses the use of community volunteers as a valid source of data collection while examining the effect on concussion rate due to player age and ability.

Le taux de commotions cérébrales au hockey sur glace a été étudié dans les niveaux élite mais pas chez les jeunes dans la communauté où l'on pense également que ça se produise. En raison de la rareté relative des commotions cérébrales, un grand nombre de parties doivent être observées afin d'obtenir une estimation fiable. Ceci peut devenir très coûteux si nous payons des personnes pour recueillir les données. Ce projet évalue l'utilisation de volontaires de la communauté comme source valide de collecte de données tout en étudiant l'effet de l'âge et des capacités du joueur sur le taux de commotions.

[MS-12]

**Spectral Analysis of Solar Flare Activity**  
**Analyse spectrale des éruptions solaires**

Azadeh MOGHADERI, Glen TAKAHARA & David THOMSON, *Queen's University*

We consider 43 years of NOAA solar radio burst data over 3 different nonoverlapping frequency bands. We compute yearly power spectra using the Thomson's multitaper method, apply a harmonic F-test to isolate significant line components in the data, and compute multitaper coherences between selected spectra. With this approach we are able to identify many coincident frequencies across frequency bands and observation stations, which appear to be associated either with solar rotation or with discrete solar frequencies. The identification of these coincident frequencies, happening as a result of random occurrence, is exceedingly low.

Nous considérons 43 années de données solaires d'éclat radio NOAA sur 3 bandes de fréquences différentes non superposées. Nous calculons des spectres de puissance annuels en utilisant la méthode dite multitaper de Thomson et appliquons un test F harmonique pour isoler la composante linéaire significative dans les données. De plus, nous calculons les cohérences de type multitaper entre les spectres choisis. Avec cette approche, nous pouvons identifier beaucoup de fréquences coïncidentes à travers les bandes de fréquences et les stations d'observation qui semblent être associées à la rotation solaire ou aux fréquences solaires discrètes. L'identification de ces fréquences coïncidentes, qui se produisent en raison de l'occurrence aléatoire, est excessivement basse.

[MS-13]

**On a Mixture Generalized Autoregressive Conditional Heteroscedastic Model**  
**Un modèle de mélange autorégressif conditionnellement hétéroscédastique généralisé**

Zi JIN, *McMaster University*

We propose a mixture generalized autoregressive conditional heteroscedastic (MGARCH) model. The models consist of a mixture of  $K$  autoregressive components with generalized autoregressive conditional heteroscedasticity. The shape-changing features of the conditional distributions make it capable of modeling time series with multimodal conditional distributions and heteroscedasticity. The stationary conditions and autocorrelation function are derived. The estimation is done via EM algorithm, and the model selection problem is addressed. Construction of multiple step predictive distributions is discussed. The models are applied to real datasets, IBM common stock closing prices, and compared to the other competing models. It appears to be the best.

Nous proposons un modèle de mélange autorégressif conditionnellement hétéroscédastique généralisé (MGARCH). Les modèles se composent d'un mélange de  $K$  composantes autorégressives admettant une hétéroscédasticité conditionnellement autorégressive généralisée. Les caractéristiques de changement de forme des distributions conditionnelles permettent de modéliser une série chronologique hétéroscédastique avec des distributions conditionnelles multimodales. Les conditions de stationnarité et la fonction d'autocorrélation sont obtenues. L'estimation se fait par l'intermédiaire de l'algorithme EM et le problème de sélection de modèle est également considéré. Nous discutons de la construction des distributions prédictives à multiples étapes. Les modèles sont appliqués à de vrais jeux de données, les prix de fermeture de l'action ordinaire d'IBM, et comparés aux autres modèles concurrents. Le modèle proposé semble être le meilleur.

[MS-14]

**Efficient Data Smoothers Using Slepian Functions**  
**Lissage efficace de données en utilisant des fonctions de type « slepian »**

Ian MOORE & David J. THOMSON, *Queens University*

In this investigation we introduce a near optimal data smoother using Prolate Spheroidal Wave Functions (PSWFs) in the form of a series expansion. We show that there is a broad optimum in the number of terms in the series and the bandwidth of the smoother gives a minimum mean-squared error in the smoothed output data. Using climate data as an example, we demonstrate that this method is better than many commonly used data smoothing techniques.

Dans cette recherche, nous présentons un lisseur de données optimal qui utilise un développement en séries de fonctions d'ondes sphéroïdales allongées. Nous prouvons qu'il y a un large optimum dans le nombre de termes de la série et la largeur de bande du lisseur donne une erreur quadratique minimale dans les données lissées résultantes. En utilisant des données climatiques comme exemple, nous démontrons que cette méthode est meilleure que plusieurs techniques de lissage généralement utilisées.

[MS-15]

**EWMA Charts for Monitoring Mean Changes in Weibull Processes with Censored Lifetime Data**  
**Diagramme EWMA pour détecter des changements dans la moyenne de processus de Weibull avec des données de survie censurées**

Gemai CHEN, *University of Calgary*, Lingyun ZHANG, *Massey University, New Zealand*

In this talk, we develop a lower-sided and an upper-sided EWMA charts for detecting mean changes (decreases and increases) in processes characterized by Weibull distributions, when censoring occurs at a fixed level. It is shown that the lower-

Dans cette présentation, nous développons des diagrammes EWMA unilatéraux supérieurs et inférieurs pour détecter des changements dans la moyenne (diminutions et augmentations) de processus caractérisés par la distribution de Weibull lorsque la censure se produit à un niveau fixe. Nous

sided EWMA chart performs better than its counterpart, the Shewhart-type chart, that a Shewhart-type chart is not appropriate to detect mean increases when the censoring rate is high, and that the upper-sided EWMA chart works well in detecting mean increases.

montrons que le diagramme EWMA unilatéral inférieur performe mieux que le diagramme correspondant de type Shewhart, que le diagramme de type Shewhart n'est pas approprié pour détecter des augmentations dans la moyenne quand le taux de censure est élevé et que le diagramme EWMA unilatéral supérieur fonctionne bien pour détecter des augmentations dans la moyenne.

#### [MS-16]

##### **Some Properties of GARCH Processes**

##### **Quelques propriétés des processus GARCH**

Srimantoora APPADOO & A. THAVANESWARAN, *University of Manitoba*

Both volatility clustering and conditional nonnormality induce the leptokurtosis typically observed in financial data. An ARMA representation is used to derive the kurtosis of the various class of GARCH models such as power GARCH, non-Gaussian GARCH, non-stationary and random coefficient GARCH. The kurtosis is also derived for random normal mixture GARCH process and for normal mixture IGARCH processes.

Les amas de volatilité et la non normalité conditionnelle induisent le phénomène de leptokurtose typiquement observé dans des données financières. Une représentation ARMA est utilisée pour obtenir la kurtose de diverses classes de modèles GARCH tels que le GARCH puissance, le GARCH non gaussien et le GARCH non stationnaire à coefficients aléatoires. La kurtose est également dérivée pour un mélange normal aléatoire de processus GARCH et pour des mélanges normaux de processus IGARCH.

#### [MS-17]

##### **Modeling Random Effects in a Dose-Response Model with Application to Developmental Toxicity Data**

##### **Modélisation de l'effet aléatoire dans un modèle dose-réponse avec des applications pour des données sur la toxicité de développement**

Shesh RAI, *St. Jude Children's Research Hospital, Memphis, TN*, Daniel HUNT, *St. Jude Children's Research Hospital, Memphis, TN*

Teratological experiments are dose-response studies in which animals are randomly assigned to various exposure of a toxic substance. Subsequently, both continuous and discrete responses are recorded on the litters of fetuses that these animals produce. This clustered binary data usually exhibits overdispersion. To model the correlation and/or variation, the beta-binomial distribution has been assumed. We propose to incorporate a random effect into a general threshold dose-response model to account for the variation in responses, while at the same time estimating the threshold effect. Simulation studies are performed to assess the validity of the random effects threshold model.

Les expériences tératologiques sont des études de dose-réponse dans lesquelles des animaux sont aléatoirement assignés à diverses expositions d'une substance toxique. En conséquence, des réponses continues et discrètes sont enregistrées sur les portées de foetus que ces animaux produisent. Ces données binaires groupées démontrent généralement une sur-dispersion. Pour modéliser la corrélation et/ou la variation, la distribution bêta-binomiale est supposée. Nous proposons d'incorporer un effet aléatoire à un modèle général de dose-réponse avec seuil pour expliquer la variation des réponses et en même temps estimer l'effet du seuil. Des études de simulation sont réalisées pour évaluer la validité du modèle à effet aléatoire avec seuil.

**[MS-18]****Detection of Long-Range Dependence and the Nile River  
Détection de dépendance à long terme et le fleuve du Nil**Scott AMISS & David J. THOMSON, *Queen's University*

A famous example of a time series which supposedly displays “long memory” (LM) is the Nile River yearly minima, 622–1284 AD. The LM phenomenon is measured by a single parameter  $H$ , also known as the Hurst exponent. Traditional estimates indicate ‘strong’ LM. We test for LM using multitaper spectral analysis, yielding the lowest known estimates of  $H$  for the Nile data. We also present a simple test for LM based on the ratio of the first frequency derivative of the spectrum to the spectrum itself. Results from this test indicate classifying the Nile data as LM may be in error.

Le niveau minimal du fleuve du Nil de 622 à 1284 est un exemple célèbre de séries chronologiques qui est supposé illustrer la dépendance à long terme (DLT). Le phénomène DLT est mesuré à l’aide d’un paramètre  $H$  connu sous le vocable d’exposant de Hurst. Les estimateurs traditionnels sont indicateurs d’une DLT « forte ». Nous testons l’hypothèse de DLT en utilisant l’analyse spectrale de type « multitaper », ce qui mène à la plus petite valeur estimée connue de  $H$  pour les données du Nil. Nous présentons également un test simple pour la DLT basé sur le rapport de la dérivée de la première fréquence du spectre sur le spectre lui-même. Les résultats de ce test indiquent que cataloguer les données du Nil comme étant DLT pourrait être une erreur.

**[MS-19]****Computer Algebra Derivation of the Bias of Burg Estimators  
Dérivation par l’algèbre computationnelle du biais des estimateurs de Burg**Ying ZHANG & Ian MCLEOD, *University of Western Ontario*

Using computer algebra the bias to  $O(1/n)$  of the Burg estimator in AR(1) and AR(2) models is shown to be equal to that of the least squares estimators in both the known and unknown mean cases. Previous researchers have only been able to obtain simulation results for this bias because this problem is too intractable without using computer algebra.

En utilisant l’algèbre computationnelle, nous montrons que le biais d’ordre  $O(1/n)$  de l’estimateur de Burg dans les modèles AR(1) et AR(2) est égal à celui des estimateurs des moindres carrés dans le cas où la moyenne est connue et celui où la moyenne est inconnue. D’autres chercheurs avaient seulement pu obtenir des résultats de simulations pour ce biais parce que ce problème est trop difficile sans l’utilisation de l’algèbre computationnelle.

**[MS-20]****On the Robustness of Optimal Scaling for Metropolis-Hastings Algorithms  
Robustesse de l’échelonnage optimal pour les algorithmes de type Metropolis-Hastings**Mylène BÉDARD, *University of Toronto*

Metropolis-Hastings algorithms, an important class of MCMC algorithms, allow for data generation from highly complex distributions. In applying Metropolis-Hastings algorithms, it is necessary to choose a proposal density and to select the “scaling” of this proposal density, in order to have some level of optimality in the performance of the algorithm. Results about the optimal scaling of the proposal distribution and the optimal acceptance rate of the algorithms are available in the literature for multidimensional target distributions with i.i.d.

Les algorithmes de type Metropolis-Hastings, une classe importante d’algorithmes MCMC, permettent de générer des échantillons provenant de distributions complexes. En appliquant ces algorithmes, il est nécessaire de choisir une distribution instrumentale et de sélectionner le paramètre d’échelle de cette distribution afin d’optimiser la performance de l’algorithme. Des résultats sur l’échelonnage optimal de la distribution instrumentale et sur l’optimisation du taux d’acceptation des algorithmes sont disponibles dans la littérature pour les distributions cibles multidimensionnelles aux composantes i.i.d.. Je considère dans quelle mesure ces



components. I consider the extent to which these results can be generalised to more complicated, realistic target distributions.

résultats peuvent être généralisés à des distributions cibles plus complexes et réalistes.

**[MS-21]**

**Construction of scales for the Rorschach technique (SC) in Argentina**

**Construction de barèmes pour la technique Rorschach (SC) en Argentine**

María Inés URRUTIA, Helena LUNAZZI, Marta DE LA FUENTE, Diana ELIAS, Favia FERNANDEZ & Soledad DE LA FUENTE, *Universidad Nacional de La Plata, Argentina*

The 10 ink spots which constitute the stimuli of the Rorschach technique date back to a 1921 publication. A complete system for the Exner Rorschach (1974) was gradually adopted internationally. In Argentina, between 1999 and 2003, we established local standards on a sample of 506 non-patients individuals representative of the demographic diversity of the population. These scales, which did not exist before, allow on the one hand psychologists to evaluate their scores on the basis of their own cultural context and on the other hand, to make intercultural and international studies. This is what makes them original. From a statistical point of view, the work consists in evaluating the demographic composition of the sample and to establish the sections in order to go deeper in the study of the 116 variables provided by the technique for each subject.

Les 10 taches d'encre qui constituent les stimuli de la technique de Rorschach remontent à une publication datant de 1921. Un système complet pour le Rorschach d'Exner (1974) a été graduellement adopté au plan international. En Argentine, nous avons établi entre 1999 et 2003 les normes locales sur un échantillon de 506 individus non patients représentatifs de la diversité démographique de la population. Ces barèmes, auparavant inexistantes, permettent d'une part aux psychologues d'évaluer leurs pointages par rapport à leur propre contexte culturel et, d'autre part, de réaliser des études interculturelles et internationales ; c'est ce qui fait leur originalité. Au plan statistique, le travail consiste surtout à évaluer la composition démographique de l'échantillon et à établir en conséquence les coupes en vue d'approfondir l'étude des 116 variables que fournit la technique sur chaque sujet.

**Session 1**

**Monday, May 31 • Lundi 31 mai, 08:30 – 10:00**

**B-2285**

**Inaugural Session and SSC Presidential Invited Address**

**Séance inaugurale et allocution de l'invitée de la présidente**

**[MS-22]**

**Discovering Haplotype Blocks in the Human Genome**

**Découvrir des blocs haplotypes dans le génome humain**

Kathryn ROEDER, A. RINALDO, B. DEVLIN & L. WASSERMAN, *Carnegie Mellon University*

The HapMap project aims to assess the statistical dependence of local variants in the DNA of the human genome. Strong local dependence results in "haplotype blocks," or segments of chromosomes apparently transmitted intact through many generations. Their identification aids substantially the discovery of genetic variants associated with increased disease risk. Given a set of  $k$  binary variants (SNPs) defining a haplotype, its distribution can be described in a  $2^k$ -dimensional contingency

Le projet HapMap vise à évaluer la dépendance statistique des variantes locales dans l'ADN du génome humain. La forte dépendance locale mène à des « blocs haplotypes », ou des segments de chromosomes apparemment transmis intacts à travers plusieurs générations. Leur identification facilite substantiellement la découverte des variantes génétiques liées au risque accru de la maladie. Étant donné un ensemble de  $k$  variantes binaires (SNP) définissant un haplotype, sa distribution peut être décrite dans un tableau de contingence à  $2^k$ -dimensions qui sera extrêmement dispersé si un sous-

table, which will be extremely sparse if a subset of the SNPs defines a block. Based on the distribution of the haplotypes, we develop a new method for block discovery.

ensemble de SNP définit un bloc. Sur la base de la distribution des haplotypes, nous développons une nouvelle méthode pour la découverte de blocs.

**Session 2A**                      **Monday, May 31 • Lundi 31 mai, 10:30 – 12:00**                      **Z-310**

**Gender-Related Statistics in Education**

**Statistiques liées au genre dans le domaine de l'éducation**

**[MS-23]**

**Engendering Education Statistics: The International Perspective**

**Intégration du genre dans les statistiques de l'éducation : Perspective internationale**

Denise LIEVESLEY, *UNESCO*

Almost ten years ago the Beijing World Summit on the Status of Women highlighted the importance of the disaggregation of statistics by gender. Without such data the situation of women relative to that of men cannot be studied. Statistics give visibility and allow policies to be informed by data leading to the reduction of inequalities in our societies. But it is not sufficient to disaggregate data, it is also essential to explore the bias inherent in the selection of issues to be addressed. Denise Lievesley will discuss these issues drawing on the experience of UNESCO in education and literacy.

Il y a près de 10 ans, le Sommet mondial de Beijing sur le Statut de la femme soulève l'importance de désagréger les statistiques par sexe. La situation de la femme comparativement à celle de l'homme ne peut être étudiée en l'absence de ces données. Les statistiques assurent la visibilité et permettent aux décideurs de prendre des décisions informées à travers des données favorisant ainsi la réduction des inégalités dans nos sociétés. Cependant, désagréger les données n'est pas suffisant. Il est essentiel d'explorer le biais inhérent à la sélection des questions à remédier. Ces questions seront abordées par Denise Lievesley avec des illustrations basées sur l'expérience de l'UNESCO dans les domaines de l'éducation et de l'alphabétisme.

**[MS-24]**

**Sexual Harassment and Bullying in High Schools: Implications**

**Harcèlement sexuel et l'intimidation dans les écoles secondaires : implications**

Jennifer CONNOLLY & Trish WILLIAMS, *York University*

The academic and mental health links of high school bullying and sexual harassment are examined. 1,230 boys and 1,247 girls reported their academic achievement, suspensions, substance use, depression, self-esteem, social-verbal bullying, sexual harassment, and physical aggression. Victimization experiences were common, with sexual harassment occurring most frequently. For girls, sexual harassment and physical victimization were strongly associated with impaired academic and social adjustment. Similar effects were found for boys, although the links were less strong. Discussion focuses on the negative effects of hostile learning environments.

Nous examinons les liens académiques et de santé mentale de l'intimidation et du harcèlement sexuel à l'école secondaire. Une cohorte de 1230 garçons et 1247 filles ont fourni de l'information sur leur réussite scolaire, suspensions, utilisation de substances illicites, dépression, estime de soi, intimidation sociale et verbale, harcèlement sexuel et agression physique. Les expériences de victimisation étaient fréquentes et le harcèlement sexuel se produisait le plus fréquemment. Pour les filles, le harcèlement sexuel et la victimisation physique sont fortement associés à une adaptation scolaire et sociale inadéquate. Des effets semblables ont été trouvés chez les garçons bien que les liens aient été moins forts. La discussion se concentre sur les effets négatifs des environnements d'étude hostiles.

**[MS-25]****Developing Indicators to Measure Participation of Girls and Women in Education, Science and Technology  
Développer des indicateurs pour mesurer la participation des filles et des femmes dans l'éducation de la science et de la technologie**Sophia HUYER, *Gender Advisory Board—UN Commission on Science and Technology for Development*

Generally women are under-represented in almost every area of recognized scientific activity. We also know that the numbers of girls and women in science education and research decreases as one moves up the science hierarchy. Even in areas where women are statistically well represented at lower and medium levels—university faculties, for instance—they are under-represented at the higher levels of decision-making and influence. This is a phenomenon known as the leaky pipeline. Despite calls for the collection of sex-disaggregated data, there have been few efforts to collect, analyze or disseminate such data in a systematic manner. Although sex-disaggregated education statistics (beyond literacy) are collected in many countries, they present just one aspect of the picture, and further are not collected in most developing countries. The presentation will identify the areas where some data exists on the participation of women and girls in science education and research, and where it doesn't. It will make suggestions as to where it is most useful to collect data and develop indicators, bringing in examples such as the EU Women in Science Statistics initiative.

Les femmes sont généralement sous-représentées dans presque tous les secteurs d'activités des sciences reconnues. Nous savons aussi que le nombre de filles et femmes en science de l'éducation et en recherche décroît au fur et à mesure que l'on monte dans la hiérarchie de la science. Même dans les secteurs où les femmes sont bien représentées statistiquement aux niveaux inférieurs et médians—les facultés universitaires, par exemple—elles sont sous-représentées aux niveaux supérieurs de décision et d'influence. C'est ce qu'on appelle le phénomène du "pipeline percé." Malgré les appels pour des données désagrégées par sexe, il y a eu peu d'effort pour collecter, analyser et diffuser de telles données de façon systématique. Même si les statistiques en éducation désagrégées par sexe (au-delà de l'alphabétisation) sont collectées dans plusieurs pays, elles ne présentent qu'un aspect du portrait et de plus, elles ne sont pas collectées dans la plupart des pays en développement. La présentation identifiera les secteurs où des données existent sur la participation de femmes et filles en science de l'éducation et en recherche, et où il n'y en a pas. Ceci suggèrera où il est le plus utile de collecter des données et de développer des indicateurs, avec des exemples tels que le programme de Statistiques sur les Femmes en Sciences de l'UE.

**Session 2B****Monday, May 31 • Lundi 31 mai, 10:30 – 12:00****Z-330****The Escalation of Non-response and its Impacts on Surveys****La croissance de la non-réponse et ses impacts sur les enquêtes****[MS-26]****Evolution of the Response Rate in the Studies of the Institut de la statistique du Québec  
Évolution des taux de réponse dans les enquêtes à l'Institut de la statistique du Québec**Sylvain VÉGIARD, *Institut de la statistique du Québec*

The response rate will be examined from a historical point of view at the Institut de la statistique du Québec (formerly Bureau de la statistique du Québec) for a variety of study, either punctual or recurring, transversal or longitudinal and for both survey or census. The households surveys as well as the industrial surveys will be examined in this presentation. The various modes of data collection

Les taux de réponse seront examinés d'un point de vue historique à l'Institut de la statistique (autrefois le Bureau de la statistique du Québec), pour une variété d'enquêtes, qu'elles soient ponctuelles ou récurrentes, transversales ou longitudinales, ou encore selon qu'il s'agisse d'une enquête par échantillon ou d'un recensement. Les enquêtes ménages de même que les enquêtes auprès des entreprises seront examinées dans cette présentation. Les différents modes de col-

will also be compared (postal, telephone or in person).

lecte seront aussi comparés entre eux (poste, téléphone ou en face à face).

**[MS-27]**

**Nonresponse : A Survey Firm's Point of View**

**La non-réponse — le point de vue d'une firme de sondage**

Richard SAINT-PIERRE, *Ad hoc recherche*

Since 1995, Ad Hoc Research makes a quarterly study on Internet usage. By analyzing the calling process results of these surveys, we note a progressive reduction in the response rate. This report is the illustration of a major stake of our industry. This reduction increases the costs and makes it more difficult to obtain valid results. We will see in particular that the profile of the people most difficult to reach is different from that of the others. In the same way, we note that a weighting based on socio-demographic variables cannot entirely correct the bias caused by the nonresponse.

Depuis 1995, Ad Hoc Recherche réalise trimestriellement une étude portant sur l'utilisation d'internet. En analysant les résultats d'appel de ces sondages, on constate une diminution progressive du taux de réponse. Ce constat est l'illustration d'un enjeu majeur de notre industrie. Cette baisse, en plus d'amener une augmentation des coûts, rend plus ardue l'obtention de résultats valides. Nous verrons notamment que le profil des personnes les plus difficiles à joindre est différent de celui des autres répondants. De même, on constatera qu'une pondération basée sur des variables socio-démographiques ne peut corriger entièrement les biais causés par la non-réponse.

**[MS-28]**

**Nonresponse... What Is the Real Problem?**

**La non-réponse... quel est le vrai problème ?**

Michel SAULNIER, *Bell Canada*

The research industry and particularly the service suppliers have been alarmed for a certain time by the decline of the response rates. But do we really have a problem of nonresponse? One should rather ask: is this the real problem? Data show that when studies are done by the book, the response rate does not fall. With data coming from a follow-up study over a few years, we note that the response rate did not change. But then, why do we say that there is a fall of the response rate? Is it our way of making things that has an impact on the response rate? Is the response rate simply a consequence of the desire to always go faster?

L'industrie de la recherche, et en particulier les fournisseurs de service, sonne l'alarme sur la chute des taux de réponse depuis déjà un certain temps. Mais avons-nous un problème réel de non-réponse ? On devrait plutôt demander, est-ce le vrai problème ? Des données portent à croire que si nous faisons les choses selon les règles de l'art, le taux de réponse ne chute pas. À l'aide de données provenant d'études de suivi couvrant une période de quelques années, on constate que le taux de réponse n'a pas vraiment changé. Mais alors, pourquoi dit-on qu'il y a une chute du taux de réponse ? Est-ce notre façon de faire les choses qui a un impact sur le taux de réponse ? Le taux de réponse est-il simplement une conséquence du désir d'aller toujours plus vite ?

**Session 2C**

**Monday, May 31 • Lundi 31 mai, 10:30 – 12:00**

**AA-1140**

**Current Trends in Statistical Genetics**

**Tendances actuelles en statistique génétique**

**[MS-29]****Importance Sampling on Evolutionary Histories****Échantillonnage avec fonction d'importance pour des histoires évolutionnistes**Robert GRIFFITHS, *University of Oxford*, Maria De IORIO, *Imperial College*

The likelihood of a sample configuration of gene types depends on their ancestral history. Ancestral inference about parameters in a coalescent model of ancestry and questions about past history of the genes can be answered by likelihood methods. A computationally intensive method of simulating the sample likelihood is to use importance sampling on the stochastic ancestral history process of the genes. This talk will give an overview of the technique, explaining the connection with diffusion process models of the population gene frequencies. An example application is to samples of genes from red foxes in Australia.

La probabilité de la configuration des types génétiques d'un échantillon dépend de l'histoire des ses ancêtres. L'inférence ancestrale sur des paramètres dans un modèle de coalescence des ancêtres et certaines questions sur l'histoire passée des gènes peuvent être résolues par des méthodes de vraisemblance. Une méthode numériquement intensive de simuler la vraisemblance échantillonnale consiste à utiliser l'échantillonnage avec fonction d'importance pour le processus stochastique de l'histoire ancestrale des gènes. Cette présentation donnera une vue d'ensemble de la technique, expliquant les liens avec des modèles de processus de diffusion de la fréquence des gènes d'une population. On présentera un exemple d'application à des échantillons de gènes des renards rouges en Australie.

**[MS-30]****Bayesian Estimation of the Evolutionary History of Two Genomes****Estimation bayésienne de l'histoire évolutionniste de deux génomes**Rasmus NIELSEN, Rick DURRETT & Thomas YORK, *Cornell University*

The order of genes in the genomes of two different species will differ due to certain mutational mechanisms, inversions and translocations, which permute the marker order. The problem of estimating the number of evolutionary events has previously been considered a problem of identifying the minimum number of evolutionary events that can transform the genome of one species into that of another. We present a Markov chain model of evolution that will allow statistical inferences regarding the history of evolutionary events. Using MCMC we obtain estimates of the rates at which inversions and translocations occur and make other evolutionary inferences.

L'ordre des gènes dans le génome de deux espèces différentes diffère en raison de certains mécanismes de mutation, d'inversions et de translocations qui permutent l'ordre des marqueurs. Le problème d'estimation du nombre d'événements évolutionnistes a déjà été considéré comme un problème d'identification du nombre minimum d'événements évolutionnistes pouvant transformer le génome d'une espèce en celui d'une autre. Nous présentons un modèle de chaîne de Markov de l'évolution qui permettra des inférences statistiques sur l'histoire des événements évolutionnistes. En utilisant la méthode MCMC, nous obtenons des estimations des taux auxquels les inversions et les translocations se produisent et nous faisons d'autres inférences évolutionnistes.

**Session 2D****Monday, May 31 • Lundi 31 mai, 10:30 – 12:00****AA-1360****Case Study I: Asthma****Étude de cas I : Asthme**

**[MS-31]****Presentation of the Data Set****Présentation du jeu de données**Edmee FRANSSSEN, *GSK*

Asthma is a chronic condition and one measure of how well a subject's asthma is controlled is by daily scoring of various asthma symptoms over a given time period. Common asthma symptoms recorded are cough, wheeze and shortness of breath. Each symptom is usually rated from 0-3: no symptoms, mild, moderate and severe. Symptoms during the day can be different than night symptoms. Also, use of rescue medication is a sign of poor control. However, this symptom score data may be summarised in several different ways and then also analysed in several different ways. This project is designed, via use of the data from two asthma trials, to assess whether differing methods of summary and analysis are more or less sensitive to detecting treatment differences, and whether one method provides summary statistics that are easier to understand and interpret. We have data from two 12 week large, multi-centre asthma trials available for this analysis.

L'asthme est une condition chronique. L'une des façons de mesurer si le sujet contrôle son asthme est de lui demander de juger ses symptômes pendant une période donnée. Les symptômes les plus communément signalés sont la toux, une respiration sifflante et le manque de souffle. Chaque symptôme est noté de 0 à 3 : symptôme inexistant, bénin, modéré et grave. Les symptômes de jour sont parfois différents de ceux de nuit. De plus, l'utilisation de médicaments bronchodilatateurs est un signe de mauvais contrôle. Cependant, ces données de score peuvent être résumées et analysées de différentes manières.

Ce projet est destiné à évaluer, sur la base de données provenant de deux essais d'asthme, quelles sont les méthodes de résumé et d'analyse qui permettent de détecter des différences de traitement de la façon la plus fiable et si l'une de ces méthodes produit des statistiques sommaires qui sont plus faciles à comprendre et à interpréter. Nous disposons pour cette analyse de données tirées de deux grands essais multi-centres de 12 semaines.

**Session 2E****Monday, May 31 • Lundi 31 mai, 10:30 – 12:00****AA-1355**

Finance

Finance

**[MS-32]****Portfolio Selection and Markov Decision Processes****Sélection du portefeuille et processus de décision de Markov**Yanqing YI & Xikui WANG, *University of Manitoba*

We model the problem of optimal portfolio selection by Markov decision processes. After establishing the existence of an optimal strategy, we discuss various properties of the model and the optimal strategy. In particular, we examine the portfolio selection problems with constant relative risk aversion and absolute risk aversion.

Nous modélisons le problème du choix optimal d'un portefeuille par des processus de décision de Markov. Après avoir établi l'existence d'une stratégie optimale, nous discutons de diverses propriétés du modèle et de la stratégie optimale. En particulier, nous examinons les problèmes lors de la sélection du portefeuille avec une aversion au risque relative constante et une aversion absolue au risque.

**[MS-33]****Optimal Portfolio Selection Among Risky Assets  
Sélection optimale d'un portefeuille parmi des titres risqués**Yan WANG & Xikui WANG, *University of Manitoba*

We examine the problem of optimal portfolio selection in a financial market consisting of a number of risky assets. Expected returns are uncertain and the Bayesian approach is applied. With a suitable model and under appropriate conditions, each risky asset is characterized with an index value similar to the Gittins index in the theory of bandit processes. The number of assets invested can be either fixed or random, and the portfolio selection is based on comparing the index values for all risky assets. We also discuss properties of the index value function.

Nous examinons le problème du choix optimal d'un portefeuille sur un marché financier se composant d'un certain nombre de titres risqués. Les rendements prévus sont incertains et nous utilisons une approche bayésienne. Avec un modèle convenable et sous des conditions appropriées, chaque titre risqué est caractérisé par une valeur d'index semblable à l'index de Gittins dans la théorie du processus de bandit. Le nombre de titres investis peut être fixe ou aléatoire et la sélection du portefeuille est basée sur la comparaison des valeurs des index pour tous les titres risqués. Nous discutons également des propriétés de la fonction de valeur d'index.

**[MS-34]****Discrete Time Finance and Option Pricing: A Semiparametric Approach  
Finance à temps discret et évaluation du prix d'une option : une approche semi-paramétrique**Alex BADESCU & Reg KULPERGER, *University of Western Ontario*

Financial returns data can be modeled by nonlinear time series. Option pricing also uses a risk neutral measure, but these are not unique for discrete time series. Duan uses a local risk neutral valuation method when the returns model is GARCH with Gaussian noise. We consider GARCH models with Gaussian mixture (GM) noise. A RN measure is then obtained by an Esscher of the historical GARCH model. When fitting such GARCH models, one can use a semiparametric bootstrap method to approximate option prices. A numerical comparison is made between these methods and Duan's model.

Les données de rendements financiers peuvent être modélisées par des séries chronologiques non linéaires. L'évaluation du prix d'un option utilise également des mesures neutres au risque, mais celles-ci ne sont pas uniques pour des séries chronologiques discrètes. Duan utilise une méthode d'évaluation locale neutre au risque quand le modèle de rendements est GARCH avec un bruit gaussien. Nous considérons les modèles GARCH avec un bruit de mélange gaussien (MG). Une mesure neutre au risque est alors obtenue par un Esscher du modèle historique GARCH. Quand nous ajustons un tel modèle GARCH, nous pouvons utiliser une méthode de bootstrap semi-paramétrique afin d'estimer le prix des options. Une comparaison numérique est faite entre ces méthodes et le modèle de Duan.

**[MS-35]****Nonlinear Expectations with Application to Incomplete Market  
Espérances non linéaires avec application au marché incomplet**Zengjing CHEN & Reg KULPERGER, *University of Western Ontario*

In a complete market model in mathematical finance the Black–Scholes theory gives the price of a contingent claim as an expectation in terms of the risk neutral measure. Incomplete market models have many risk neutral measures. In the case of stochastic differential equation (SDE) models of

Dans un modèle de marché complet en finance mathématique, la théorie de Black Scholes donne le prix d'un droit contingent comme une espérance en termes de la mesure neutre au risque. Les modèles de marché incomplet ont plusieurs mesures neutres au risque. Dans le cas des modèles d'équations différentielles stochastiques (EDS) du cours des

stock prices, the price of a contingent claim can be defined in terms of a backward SDE process, that is through a control SDE with the control at the expiration time. In an incomplete market model this leads to the notion of nonlinear or Choquet expectation, which is explored in this talk.

actions, le prix d'un droit contingent peut être défini en termes d'un processus d'EDS renversé, c.-à-d. par une EDS de contrôle avec le contrôle sur le temps d'expiration. Dans un modèle de marché incomplet, ceci mène à la notion de l'espérance non linéaire ou de Choquet, notion que nous explorons dans cette présentation.

### [MS-36]

#### **Monte Carlo Markov Chain Analysis of Time-changed Levy Processes of Stock Return Dynamics Analyse de Monte Carlo avec chaîne de Markov de processus de Levy à changement temporel pour la dynamique de rendement d'une action**

Long YU, Haitao LI & Martin WELLS, *Cornell University*

We develop MCMC techniques for estimating time-changed Levy processes of stock return dynamics. The models exhibit stochastic volatility and jumps. Unlike Poisson jumps considered in most existing studies, jumps in our models follow Levy-type of distributions, such as Variance-Gamma and Log-Stable distribution. Special techniques are needed for estimating Levy processes because for certain models the probability density does not have analytic form and certain moments do not exist. The MCMC methods developed in our paper have excellent performance in estimating Levy processes. Empirically we show that stochastic volatility models with Levy-type jumps perform much better than models with Poisson jumps.

Nous développons des techniques MCMC pour estimer des processus de Levy à changement temporel de la dynamique de rendement d'un cours boursier. Les modèles montrent la volatilité et les sauts stochastiques. Contrairement aux sauts de Poisson considérés dans les études existantes, les sauts dans notre modèle suivent une distribution de type Levy, tels que la Gamma-variance et la loi stable logarithmique. Des techniques spéciales sont nécessaires pour estimer des processus de Levy car pour certains modèles, la densité de probabilité n'a pas de forme analytique et certains moments n'existent pas. Les méthodes MCMC développées dans notre article estiment très bien les processus de Levy. Empiriquement, nous prouvons que les modèles de volatilité stochastique avec des sauts de type Levy performant mieux que les modèles avec des sauts de Poisson.

### [MS-37]

#### **On Some Extension to Gamma Processes Sur le prolongement des processus gamma**

Mahmoud ZAREPOUR & H. ISHWARAN, *University of Ottawa/Université d'Ottawa*

We first introduce a generalized gamma process which covers the gamma processes and the stable process. We extend this to the generalized signed gamma processes. The important special case is the well known variance gamma process which is used to model stock prices. Then we develop a technique that can generalize this to a multidimensional process which in our opinion gives a natural definition for the multidimensional gamma and generalized gamma processes. Since gamma processes and their counterparts Dirichlet processes are widely used in the Bayesian inference, our theory can be used to construct a Bayesian framework in finance.

Nous présentons tout d'abord un processus gamma généralisé qui couvre les processus gamma et le processus stable. Nous prolongeons ceci aux processus gamma signés généralisés. Le cas particulier important est le processus de variance gamma bien connu qui est utilisé pour modéliser le cours d'actions boursières. Nous développons ensuite une technique qui peut généraliser ceci à un processus multidimensionnel, qui à notre avis donne une définition naturelle pour le processus gamma multidimensionnel et le processus gamma généralisé. Puisque les processus gamma et leurs contreparties les processus de Dirichlet sont largement répandus dans l'inférence bayésienne, notre théorie peut être utilisée pour construire un cadre bayésien en finance.



**Session 2F****Monday, May 31 • Lundi 31 mai, 10:30 – 12:00****Z-317****GEE and Related Methods  
EGE et méthodes connexes****[MS-38]****Analysis of Large Cluster Randomized Trials with Binary Outcomes and Weak Intra-class Correlation  
Analyse d'essais randomisés à grandes grappes avec des résultats binaires et une corrélation intra-groupe faible**Ye TAN & Robert PLATT, *McGill University*

Several methods can be used in cluster-randomized studies when the outcome is binary, including GLMM, GEE, and ordinary logistic regression. To compare these methods, we performed a simulation study. Mean parameter estimates were close to true values; however, GLMM gives the highest 95% confidence interval coverage of the true coefficients, while the GEE approaches performed slightly less well. Logistic regression models performed well when the correlation was weak, but fared poorly when the correlation was stronger. When correlations are quite low, logistic models may be acceptable for clustered data; however, they give inappropriate inference when correlation is elevated.

Plusieurs méthodes peuvent être utilisées dans des études à grappes-randomisées quand les résultats sont binaires, y compris GLMM, GEE et la régression logistique ordinaire. Pour comparer ces méthodes, nous avons réalisé une étude de simulation. Les estimations des moyennes étaient près des vraies valeurs, cependant GLMM donne l'intervalle de confiance à 95% avec la couverture des vrais coefficients la plus élevée, alors que les approches GEE performant légèrement moins bien. Les modèles de régression logistique performant bien quand la corrélation est faible et moins bien quand la corrélation est plus forte. Quand les corrélations sont plutôt faibles, les modèles logistiques peuvent être acceptables pour des données en grappes, cependant ils donnent des inférences inadéquates quand la corrélation est élevée.

**[MS-39]****Misspecification of the Drop-out Process in Inverse Probability Weighted GEE  
Mauvaise spécification du processus de retrait dans des EEG à probabilité inverse pondérées**Leilei ZENG & Richard J. COOK, *University of Waterloo*

Inverse probability weighted estimating equations are often used to obtain consistent parameter estimates from the analysis of incomplete longitudinal data with “random drop-outs” (RD). Such methods are sensitive to misspecification of the model for the drop-out process. We study the impact of misspecification of drop-out model on the asymptotic and finite sample frequency properties of regression coefficients in the response model. Particular attention is given to two cases of model misspecification: the case that important covariates are omitted and, the case that there are competing reasons for drop-outs but a single cause drop-out process is assumed at the analysis stage.

Les équations d'estimation à probabilité inverse pondérée sont souvent utilisées pour obtenir des estimations convergentes de paramètre à partir de l'analyse de données longitudinales incomplètes avec retrait aléatoire. De telles méthodes sont sensibles à la mauvaise spécification du modèle pour le processus de retrait. Nous étudions l'impact de la spécification erronée du modèle sur les propriétés asymptotiques et d'échantillon fini des coefficients de régression dans le modèle de réponse. Une attention particulière est portée à deux cas de mauvaise spécification du modèle : le cas où des covariables importantes sont omises et le cas où il y a plusieurs raisons de retrait, mais qu'on postule un seul processus de retrait lors de l'analyse.

**[MS-40]****Approximate Inference for Nonlinear Mixed-effects Models with Nonignorably Missing Covariates**  
**Inférence approximative pour les modèles non linéaires à effet mixte avec covariables manquantes non ignorables**Lang WU, *University of British Columbia*

Nonlinear mixed-effects models (NLME) are popular in many longitudinal studies such as HIV viral dynamics, pharmacokinetic analyses, and studies of growth and decay. We consider NLME models with missing covariates when the missing data mechanism may be nonignorable. We propose an approximate method based on Taylor expansions, implemented by a Monte Carlo EM algorithm along with Gibbs sampler and rejection/importance sampling methods. An AIDS dataset is used for illustration.

Les modèles non linéaires à effet mixte (NLEM) sont populaires dans beaucoup d'études longitudinales telles que la dynamique virale du VIH, les analyses pharmacocinétiques et les études sur la croissance et l'affaiblissement. Nous considérons des modèles NLEM avec covariables manquantes quand le mécanisme de données manquantes peut être non ignorable. Nous proposons une méthode approximative basée sur un développement en séries de Taylor mis en oeuvre par un algorithme EM de Monte Carlo ainsi qu'avec l'échantillonneur de Gibbs et les méthodes d'échantillonnage de rejet/importance. Un ensemble de données sur le SIDA est utilisé pour illustrer la méthode.

**[MS-41]****Asymptotic Theory with Generalized Estimation Equations for Longitudinal Data**  
**Théorie asymptotique avec des équations d'estimation généralisées pour des données longitudinales**Ioana SCHIOPU-KRATINA, *Statistics Canada/Statistique Canada*, Raluca M. BALAN, *University of Ottawa/Université d'Ottawa*

We develop a complete theory of statistical inference for the marginal models introduced by Liang and Zeger (1986). We prove the existence, consistency (weak and strong) and asymptotic normality of sequences of estimators defined as roots of quasi-likelihood or pseudolikelihood equations. As in Rao (1998), we first obtain preliminary consistent estimators of the main regression parameter. Our results are presented in a very general setting but we give simple conditions for our theory to hold, which can be easily verified by analysts.

Nous développons une théorie complète d'inférence statistique pour les modèles marginaux introduits par Liang et Zeger (1986). Nous prouvons l'existence, la convergence (faible et forte) et la normalité asymptotique de suites d'estimateurs définis comme les racines des équations de quasi-vraisemblance ou de pseudo-vraisemblance. Comme dans Rao (1998), nous obtenons premièrement des estimateurs préliminaires convergents du paramètre principal de régression. Nos résultats sont présentés dans un contexte très général, mais nous donnons des conditions simples facilement vérifiables pour que notre théorie soit valable.

**[MS-42]****Two-stage Models to Control for Overdispersion in Longitudinal Count Data**  
**Modèles en deux étapes pour contrôler la sur-dispersion dans des données de dénombrement longitudinales**Ali Reza FOTOUHI, *University College of the Fraser Valley*

We introduced some models with different types of error terms in their linear predictor to control for omitted variables and consequently to control for overdispersion in longitudinal count data analysis. We have shown, through application to epileptic seizure and polyp data and simulation, that the type of error term is important to overcome the problem

Nous introduisons quelques modèles avec différents types de termes d'erreur dans leur prédicteur linéaire pour contrôler les variables omises et par conséquent pour contrôler la sur-dispersion dans l'analyse de données de dénombrement longitudinales. Nous montrons par l'application et la simulation sur des données sur les crises et les polypes épileptiques que le type de terme d'erreur est important pour surmonter

of overdispersion. We have also shown that the link function and the outliers are important factors. As expected we have shown that the standard error of estimate increases as variance inflation factor decreases.

le problème de la sur-dispersion. Nous montrons également que la fonction de lien et les données aberrantes sont des facteurs importants. Comme prévu, nous avons prouvé que l'écart type de l'estimateur augmente quand le facteur d'inflation de la variance diminue.

**Session 3A**                      **Monday, May 31 • Lundi 31 mai, 13:30 – 15:00**                      **Z-310**

**Special Invited Session of the Survey Methods Section**

**Allocution de l'invité d'honneur du Groupe de méthodologie d'enquête**

[MS-43]

**Informative Sampling and Sampling Information**

**Échantillonnage informatif et information d'échantillonnage**

Raymond CHAMBERS, *University of Southampton*

Model-based analyses of sample surveys typically assume that the population represented by the sample is the same as the sampled population. This is often not the case. A sample where these two population concepts differ can be said to be informative. Unfortunately, how to measure this information, and how to use it if it is available, is unclear. In this talk I describe a likelihood-based framework for analysis of sample survey data that explicitly allows for informative sampling. The relationship of this approach to alternative methods in current use will be explored and illustrated using pedagogical (rather than practical) examples.

Les analyses basées sur des modèles dans des enquêtes par échantillonnage supposent typiquement que la population représentée par l'échantillon est la même que la population échantillonnée, mais ce n'est souvent pas le cas. Un échantillon où ces deux concepts de population diffèrent est dit informatif. Malheureusement, il n'est pas évident de mesurer cette information et, si elle est disponible, de l'utiliser. Dans cette présentation, je décris un cadre probabiliste pour l'analyse des données de sondage qui tient compte explicitement de l'échantillonnage informatif. La relation entre cette approche et les méthodes alternatives d'utilisation courante sera explorée et illustrée en utilisant des exemples pédagogiques (plutôt que pratiques).

**Session 3B**                      **Monday, May 31 • Lundi 31 mai, 13:30 – 15:00**                      **Z-330**

**Application of Statistical Methods to Actuarial Problems**

**Application de méthodes statistiques à des problèmes actuariels**

[MS-44]

**Actuarial Science in the City**

**L'actuariat dans la cité**

José GARRIDO, *Concordia University*

We review recent advances in risk theory, putting the emphasis on the probabilistic and statistical modeling aspects. In particular, we present the celebrated Gerber–Shiu function, showing its unifying role in solving multivariate ruin problems. Other illustrations include seasonal frequency model fitting to hurricane data as well as robust risk classification methods.

Nous présentons une revue des avancées récentes en théorie du risque, mettant l'accent sur la modélisation probabiliste et statistique. En particulier, nous présentons la fonction de Gerber–Shiu, illustrant son rôle unificateur dans la solution de problèmes de ruine multivariés. D'autres illustrations incluent l'ajustement de modèles saisonniers aux données d'ouragans ainsi que des méthodes robustes de classification du risque.

**[MS-45]****Asymmetric Heavy-tailed Distributions of Insurance Claims and Neural Networks****Les distributions asymétriques à aile lourde de réclamations d'assurance et les réseaux de neurones**Charles DUGAS, *Université de Montréal & ApSTAT Technologies Inc.*

Modeling the distribution of automobile insurance claims level, conditional on a policy holder's profile, is a difficult task. The unconditional distribution is asymmetric with a long tail: the average claim level may be around \$2000 whereas the largest claims often reach hundreds of thousands, sometimes millions of dollars. Training neural networks using gradient descent and a mean-squared error criterion with such noisy data leads to results similar to those of other techniques. We present techniques for robust regression with neural networks when dealing with asymmetric heavy-tailed distributions.

La modélisation de la distribution des réclamations d'assurance automobile, conditionnelle au profil de l'assuré, est difficile. La distribution inconditionnelle est asymétrique avec une aile lourde : la réclamation moyenne est d'environ 2000\$ alors que les plus grandes réclamations coûtent souvent plusieurs centaines de milliers, parfois des millions de dollars. L'entraînement de réseaux de neurones par descente de gradient avec un critère d'erreur quadratique moyenne lorsque les données sont si bruitées mène à des résultats similaires à ceux d'autres techniques. Nous présentons des techniques pour la régression robuste avec réseaux de neurones lorsque nous sommes en présence de distributions asymétriques avec une aile lourde.

**[MS-46]****Can Advanced Statistical Models be Applied in the Real World?****Les modèles statistiques avancés peuvent-ils vraiment être mis en pratique ?**Jean-Denis ROY, *AXA, Montréal*

The point of view of an actuary working for a P&C insurance company on the process of applying advanced models to practical situations. The specific difficulties of predicting P&C losses. The potential communication problems between academic statisticians and industry actuaries. The incredible slowness of the insurance industry in applying anything new. How can all these difficulties be overcome?

Le point de vue d'un actuaire travaillant pour une compagnie d'assurances I.A.R.D. sur le processus d'application de modèles complexes à des situations pratiques. Les difficultés spécifiques de la prédiction des pertes en assurances I.A.R.D.. Les difficultés de communication entre les statisticiens du milieu académique et les actuaires oeuvrant dans l'industrie. L'étonnante lenteur de l'industrie de l'assurance à intégrer de nouveaux outils. Comment toutes ces difficultés peuvent être surmontées ?

**Session 3C****Monday, May 31 • Lundi 31 mai, 13:30 – 15:00****AA-1140****Biostatistics Applications: The Dynamics of Collaboration****Applications biostatistiques : La dynamique de la collaboration****[MS-47]****Ecosystem Research in an Unstable Environment****Recherche sur les écosystèmes dans un environnement instable**Rick ROUTLEDGE, *Simon Fraser University*

Over a million sockeye salmon (*Oncorhynchus nerka*) used to return annually to Rivers Inlet on the British Columbia Central Coast. By the late 1990's, returns had dwindled to a few thousand. In 2001, I was approached by a local stewardship group with

Plus d'un million de saumons sockeyes (*Oncorhynchus nerka*) retournaient annuellement à Rivers Inlet sur la côte centrale de la Colombie-Britannique. Vers la fin des années 90, les retours ont diminué à quelques milliers. En 2001, j'ai été approché par un groupe de pêcheurs locaux pour

a request to help develop a research plan. This request eventually led to a collaborative ecosystem research project. The collaboration has been extensive, involving high-flying entrepreneurs, a very isolated aboriginal community, and a cumbersome federal bureaucracy. At times, some of them have not even been on speaking terms. Yet we're making solid progress, albeit in an unstable environment.

aider à développer un plan de recherche. Cette demande a mené par la suite à un projet de recherche collaborative sur l'écosystème. La collaboration s'est étendue, impliquant des entrepreneurs de haut niveau, une communauté d'autochtones très isolée et une bureaucratie fédérale encombrante. Parfois, certains d'entre eux n'étaient même pas capables de se parler. Pourtant nous accomplissons de gros progrès, quoique dans un environnement instable.

**[MS-48]**

**Using a Car Phone: Like Driving Drunk?**

**Utiliser un cellulaire au volant : est-ce comme conduire en état d'ivresse ?**

Rob TIBSHIRANI, *Stanford University*

I describe the collaboration with Don Redelemeir leading to our 1997 NEJM paper on cell phones and automobile collisions.

Je décris la collaboration que j'ai eue avec Don Redelemeir et qui a mené à notre article sur les téléphones cellulaires et les collisions en automobiles paru dans NEJM en 1997.

**[MS-49]**

**Randomized Trials (and Tribulations) in the Intensive Care Unit**

**Essais randomisés (et tribulations) dans l'unité des soins intensifs**

Rollin BRANT, *University of Calgary*

In 1990 investigators at the University of Calgary Faculty of Medicine initiated research that culminated in 2002 with the completion of a 2000 patient multicenter clinical trial designed to evaluate the clinical utility of pulmonary-hemodynamic monitoring using the Swan–Ganz catheter, a mainstay of practice in North American critical care units since the late 1970's. The implied challenge to prevailing norms generated widespread controversy and charges that the undertaking was inherently unethical. Retrenchment in research funding in the early 1990's coupled with widespread health spending cuts presented further obstacles. The challenges and rewards inherent in such high-stakes research are considered.

En 1990, des chercheurs à la Faculté de médecine de l'Université de Calgary ont lancé une recherche qui a abouti en 2002 avec la fin d'un essai clinique multi-centres sur 2000 patients. Cette recherche a été conçue pour évaluer l'utilité clinique de la surveillance pulmonaire-hémodynamique à l'aide du cathéter de Swan-Ganz, une pratique courante dans les unités de soins intensifs en Amérique du Nord depuis la fin des années 70. La mise au défi implicite par rapport à la norme représentée par cette procédure a engendré une grande polémique et des accusations que cette recherche n'était pas éthique en soi. Les suppressions dans le financement de la recherche au début des années 90 et les coupures générales en santé ont présenté d'autres obstacles. Les défis et les récompenses inhérents de recherches avec de tels enjeux sont considérés.

**Session 3D**

**Monday, May 31 • Lundi 31 mai, 13:30 – 15:00**

**AA-1360**

**Capture-recapture Models**

**Modèles de capture-recapture**

**[MS-50]****Stratification in Capture-recapture Modeling  
Stratification et les modèles capture-recapture**Carl SCHWARZ, *Simon Fraser University*

Stratification is a well known technique for improving precision in survey sampling. Stratification in capture-recapture modeling can lead to some difficult modeling problems. This talk will review the recent advances in the use stratification in capture-recapture to estimate survival, abundance, and movement with examples from fish, fowl, and forest fires. It will conclude with an overview of areas requiring further research.

La stratification est une technique bien connue pour améliorer la précision dans les sondages. La stratification dans la modélisation capture-libération peut mener à quelques problèmes difficiles de modélisation. Cette présentation passera en revue les avancées récentes dans les études de capture-recapture avec stratification pour estimer la survie, l'abondance et le mouvement avec des exemples sur des poissons, des volailles et des feux de forêt. Nous concluons avec une vue d'ensemble des secteurs exigeant davantage de recherche.

**[MS-51]****Application of Capture-recapture Type Models in Point Count Surveys  
Application de modèles de type capture-recapture dans des sondages de dénombrement ponctuel**Kenneth POLLOCK, Mathew ALLDREDGE & Ted SIMONS, *North Carolina State University*

Point counts are used to estimate density of birds. We compare several methodological approaches for dealing with detection probability (distance sampling, multiple observers, and a new time of detection method). The later two methods are based on capture-recapture type models. We develop a conceptual model for detection probability where first a bird must have a probability of being available for detection and then, conditional on availability, a probability of detection. We show how the three current methods fit into this conceptual framework and discuss their strengths and weaknesses. Combinations of these methods may lead to stronger inferences.

Les dénombrements ponctuels sont utilisés pour estimer la densité des oiseaux. Nous comparons plusieurs approches méthodologiques pour aborder la probabilité de détection (échantillonnage de distance, observateurs multiples et une nouvelle méthode de détection du temps). Les deux dernières méthodes sont basées sur des modèles de type capture-recapture. Nous développons un modèle conceptuel pour la probabilité de détection où tout d'abord un oiseau doit avoir une probabilité d'être disponible pour la détection et puis, conditionnellement sur la disponibilité, doit avoir une probabilité de détection. Nous montrons comment les trois méthodes courantes entrent dans ce cadre conceptuel et nous discutons de leurs forces et faiblesses. Des combinaisons de ces méthodes peuvent mener à de meilleures inférences.

**[MS-52]****Use of Hidden Markov Models in Capture-recapture  
Utilisation des modèles de Markov cachés en capture-recapture**Roger PRADEL, *Centre d'écologie fonctionnelle et évolutive, Montpellier*

Capture-recapture models are discrete time models for the following of individually recognizable animals. Recognition is possible due to the fitting of uniquely coded tags. These models describe the transitions among states (e.g., breeding location or status) despite the fact that the individuals are not systematically encountered. Their relation-

Les modèles de capture-libération sont des modèles de suivi en temps discret d'animaux reconnaissables individuellement suite à la pose de marques portant un code unique. Ils décrivent les transitions entre des états (lieu ou statut de reproduction par exemple) alors même que les individus ne sont pas systématiquement observés. Leur parenté avec les modèles d'épidémiologie dont ils se distinguent par une pro-

ship with the epidemiological models, from which they differ by a lower than 1 detection probability, is well known. I examine here how they relate to hidden Markov models. Indeed, transitions are 1- or 2-order Markovian but the state is not directly observed.

babilité de détection inférieure à 1 est bien connue. J'explore ici leur relation avec les modèles de Markov cachés. En effet, les transitions entre états sont markoviennes d'ordre 1 ou 2 mais l'état n'est pas systématiquement observé.

**Session 3E**                      **Monday, May 31 • Lundi 31 mai, 13:30 – 15:00**                      **AA-1355**

**Time Series Analysis**

**Analyse de séries chronologiques**

**[MS-53]**

**A Test for Additive Outliers in Time Series**

**Un test pour des valeurs aberrantes additives dans les séries chronologiques**

Rolf TURNER, *University of New Brunswick*, Patrick CHAREKA, *University of Missouri*, Florance MATARISE, *University of Zimbabwe*

We propose a test for “additive outliers” valid for a wide class of underlying stationary series, including long memory (e.g., fractionally integrated) series. The test statistic is calculated very simply, without fitting any underlying model. It converges in distribution to the Gumbel distribution. In contrast to existing tests, ours may thus be conducted at a specified significance level.

To accelerate convergence we introduce modified normalization constants. The proof that these “work” rests on a generalization of Mills’ Ratio, to Student’s  $t$  distribution, which is interesting in its own right. Simulation studies indicate that the test has good performance characteristics.

Nous proposons un test pour les « valeurs aberrantes additives » valide pour une grande classe de séries stationnaires incluant les séries à longue mémoire (par exemple intégré fractionnellement). La statistique du test est calculée très simplement sans ajuster aucun modèle sous-jacent. Elle converge en distribution vers la distribution de Gumbel. Contrairement aux tests existants, le nôtre peut être effectué au niveau de signification spécifié.

Pour accélérer la convergence, nous introduisons des constantes de normalisation modifiées. La preuve que celles-ci « onctionnent » repose sur une généralisation du ratio de Mills à la distribution  $t$  de Student, qui est intéressante en soi. Des études de simulations indiquent que le test performe bien.

**[MS-54]**

**Exact Randomized Permutation Serial Dependence Tests for Continuous and Discrete Distributions**

**Tests exacts d’indépendance sérielle dans les cas de distributions continues et discrètes**

Abdeljelil FARHAT & Jean-Marie DUFOUR, *Université de Montréal & Centre interuniversitaire de recherche en analyse des organisations & Centre interuniversitaire de recherche en économie quantitative*, Abdelwahed TRABELSI, *Université de Tunis*

We study the problem of testing the nonparametric hypothesis of randomness, with both continuous and discrete distributions, against linear and nonlinear serial dependence. We consider tests based on ranks, sample autocorrelations, goodness-of-fit statistics [Kolmogorov–Smirnov, Cramér–von Mises], and measures of distance between kernel-based density estimators. Many of these tests raise difficult distributional problems. We

Nous étudions le problème de tester l’hypothèse non paramétrique de l’aspect aléatoire avec des distributions continues et discrètes, contre la dépendance sérielle linéaire et non linéaire. Nous considérons des tests basés sur les rangs, les autocorrélations échantillonnales, les statistiques d’adéquation (Kolmogorov–Smirnov, Cramér–von Mises) et des mesures de distance entre les estimateurs par le noyau de densité. Plusieurs de ces tests soulèvent des problèmes distributionnels difficiles. Nous montrons que ces problèmes

show that these problems can easily be solved by using permutational versions of the procedures jointly with the method of Monte Carlo tests. We also suggest test combinations which yield power improvements. The performance of the proposed procedures is studied in a simulation experiment.

peuvent facilement être résolus en utilisant des versions de permutation de ces procédures conjointement avec la méthode de Monte-Carlo. Nous suggérons également une combinaison de tests qui améliore la puissance. La performance des procédures proposées est étudiée par des simulations.

**[MS-55]**

**Test for Noncorrelation of Two Infinite Order Cointegrated Vector Autoregressive Series  
Test de non corrélation de deux séries cointégrées autorégressives d'ordre infini.**

Chafik BOUHADDIOUI, *Centre interuniversitaire de recherche en analyse des organisations*, Jean-Marie DUFOUR, *Université de Montréal & Centre interuniversitaire de recherche en analyse des organisations & Centre interuniversitaire de recherche en économie quantitative*

We propose two approaches for testing noncorrelation between the innovations of two nonstationary possibly cointegrated vector processes, in the general case where the processes have infinite-order autoregressive representations. In the first approach, we show that, under the hypothesis of noncorrelation, residual cross-correlation matrices follow the same asymptotic distribution as the corresponding cross-correlation matrices based on the true innovations. In the second approach, we propose a statistic test based on partial cross-correlation matrices, which are computed as parameters of multivariate regressions between the two residuals series. A simulation study is presented to investigate the finite sample properties of the proposed tests.

Dans cette présentation, nous proposons deux approches pour tester la non corrélation des innovations de deux séries vectorielles, non stationnaires et possiblement cointégrées, dans le cas général où celles-ci possèdent une représentation autorégressive d'ordre infini. Dans la première approche, nous montrons que, sous l'hypothèse de non corrélation, un vecteur arbitraire de matrices de corrélations croisées résiduelles suit la même loi asymptotique que le vecteur correspondant des corrélations croisées entre les deux séries d'innovations. Dans la deuxième approche, nous proposons une statistique basée sur les corrélations croisées partielles qui sont calculées simplement comme des paramètres d'une régression multivariée entre les séries résiduelles. Une étude de simulation est aussi présentée afin d'évaluer la qualité de l'approximation de la loi exacte des statistiques de tests proposés.

**[MS-56]**

**Exact Maximum Likelihood Estimation of Structured or Unit Root Multivariate Time Series Models  
Estimation de vraisemblance maximale exacte de modèles structurés ou à racine unitaire pour des séries chronologiques multivariées**

Abdessamad SAIDI & Roch ROY, *Université de Montréal*, Guy MÉLARD, *Université Libre de Bruxelles*

The exact likelihood function of a Gaussian vector autoregressive-moving average (VARMA) model is evaluated in two nonstandard cases: (a) a parsimonious structured form, such as obtained in the echelon form structure or the scalar component model (SCM) structure; (b) a partially nonstationary (unit root) model in error-correction form. The starting point is Shea's algorithm (1987, 1989) for standard stationary and invertible VARMA models. Our algorithm also provides the parameter estimates and their standard errors. The small sample properties of our algorithm were studied by Monte

La fonction de vraisemblance gaussienne exacte d'un modèle autorégressif-moyenne mobile vectoriel (VARMA) est évaluée dans deux cas non standard : (a) une représentation parcimonieuse structurée comme la représentation ARMA échelon ou celle des modèles à composantes scalaires (SCM) ; (b) la représentation à correction d'erreur d'un modèle partiellement non stationnaire (à racine unitaire). Le point de départ est l'algorithme de Shea (1987, 1989) pour l'estimation d'un modèle VARMA standard, stationnaire et inversible. Notre algorithme fournit aussi les estimations des paramètres et leurs écarts types. Les propriétés de notre algorithme pour de courtes séries ont été étudiées par des



Carlo methods. Examples with real data are provided.

méthodes de Monte-Carlo. Des exemples avec des données réelles sont présentées.

**[MS-57]**

**Wavelets and the Clustering of Time Series  
Ondelettes et classification de séries chronologiques**

David CAMPBELL, *McGill University*, Hong GU, *Dalhousie University*

The extension of wavelet methods, such as basis selection and signal de-noising, to a series of signals is presented with the goal of making them easily comparable in terms of their time and frequency localizations. Signal comparisons are continued through a modification of a  $k$ -means clustering algorithm which incorporates a time shifting step. This step moves the features of a signal to align with those of other signals. Shifting and subsequent clustering of the signals are performed directly on the wavelet coefficients. An application of these methods is presented using a series of luminescent signals.

Le prolongement de méthodes d'ondelettes, telles que la sélection de base et le débruitage de signaux, à une série de signaux est présenté dans le but de les rendre facilement comparables en terme de la localisation de leur temps et fréquence. Les comparaisons des signaux se continuent par une modification d'un algorithme de clustering de  $k$ -moyennes qui incorpore une étape de décalage du temps. Cette étape déplace les caractéristiques d'un signal pour les aligner avec celles d'autres signaux. Le décalage et clustering subséquent des signaux sont exécutés directement sur les coefficients des ondelettes. Une application de ces méthodes est présentée en utilisant une série de signaux luminescents.

**[MS-58]**

**Reorder Levels of Products with Autocorrelated Demand and Resource Restriction  
Seuil de commande de produits avec une demande autocorrélée et des restrictions de ressources**

Yogendra CHAUBEY, S. K. GOYAL & F. NEBEBE, *Concordia University*

In this paper we consider the problem of determining economic reorder quantity of several products having autocorrelated demand. The products share common limited resource such as storage capacity, manufacturing facility, and the capital required to finance the investment in stocks. We assume that inventory levels are reviewed at fixed time intervals and replenished with negligible lead time.

Dans cette présentation nous considérons le problème de déterminer le seuil de commande économique de plusieurs produits ayant une demande autocorrélée. Les produits partagent des ressources limitées communes telles que la capacité de stockage, le service de fabrication et le capital nécessaire pour financer l'investissement en stock. Nous supposons que le niveau des inventaires est analysé à des intervalles de temps fixes et regarnis avec un délai d'exécution négligeable.

**Session 3F**

**Monday, May 31 • Lundi 31 mai, 13:30 – 15:00**

**Z-317**

Survival Analysis I

Analyse de durées de vie I

**[MS-59]****Goodness of Fit Tests for General Interval Censored Data****Test d'adéquation pour des données d'un modèle général de censure par intervalles**Denise BABINEAU, *University of Waterloo*

This talk discusses problems encountered when assessing the fit of a hypothesized parametric model for general interval censored data. Two situations discussed involve subjects seen at the same times (but may possibly miss some of these times) or subjects seen at varying inspection times. In the latter case, the times are not necessarily the same for all subjects. Two Pearson type goodness of fit tests are then discussed. Due to the incomplete nature of the data, simulation methods are used to investigate the tests' distribution and power. Depending on the failure and inspection process assumed, different simulation methods are developed.

Dans cette présentation nous discutons des problèmes encourus lors de l'ajustement d'un modèle paramétrique hypothétique pour des données d'un modèle général de censure par intervalles. Deux situations sont discutées, l'une impliquant des sujets observés aux mêmes temps (avec possiblement des données manquantes à certaines de ces périodes) et l'autre des sujets observés à différents temps d'évaluation. Dans le dernier cas, les temps ne sont pas nécessairement les mêmes pour tous les sujets. Deux tests d'adéquation de type Pearson sont discutés. En raison de la nature incomplète des données, des méthodes de simulation sont utilisées pour trouver la distribution et la puissance des tests. Selon les processus d'échec et d'inspection supposés, différentes méthodes de simulation sont développées.

**[MS-60]****Design of Test Plans for Weibull Lifetimes under Competing Causes of Failure****Design de plans de tests pour des temps de vie Weibull sous des causes d'échecs concurrentes**Uditha BALASOORIYA & C.-K. LOW, *Nanyang Technological University, Singapore*

We consider the design of Type I progressively censored variable-sampling plans for Weibull lifetime distributions under competing causes of failure. Though various sampling plans under Type II censoring have been discussed in the literature, for high reliability products, Type I censoring is more practical, in particular, when a test is conducted under severe time constraint. Progressive censoring allows valuable degradation-related information to be collected for improving product quality. Further, in situations where there are competing uses for test facilities, the proposed procedure releases scarce test facilities for other use.

Nous considérons le design de plan d'échantillonnage de variables progressivement censurées de Type I pour des distributions de temps de vie de Weibull sous des causes d'échecs concurrentes. Bien que divers plans d'échantillonnage sous la censure de type II ont été discutés dans la littérature pour des produits de haute fiabilité, la censure de type I est plus pratique, en particulier quand un test est effectué sous de fortes contraintes de temps. La censure progressive permet de collecter de l'information valable relativement à la dégradation pour améliorer la qualité du produit. De plus, dans les situations où il y a de la concurrence pour l'utilisation des équipements de test, la procédure proposée libère les équipements de test en grande demande pour d'autres usages.

**[MS-61]****Inference for Type II Right Progressively Censored Samples from the Generalized Extreme Value Distribution****Inférence pour des échantillons censurés progressivement à droite de type II de la distribution de valeurs extrêmes généralisée**Katherine DAVIES, *McMaster University*

In lifetime and reliability studies, there are many statistical techniques available to analyze the available data. Under certain circumstances, the data

Dans des études sur la durée de vie et la fiabilité, il y a beaucoup de techniques statistiques disponibles pour analyser les données disponibles. Dans certaines circonstances, les don-

may be censored and in particular, the data may be progressively censored. With this in mind, a common distribution used in lifetime and reliability studies is the Generalized Extreme Value distribution. In this talk, I will combine the concepts of progressive censoring and the Generalized Extreme Value distribution. Inference procedures will be discussed and exemplified with respect to these two concepts. In order to understand the performance of these methods, examples and Monte Carlo simulation results will be discussed.

nées peuvent être censurées et en particulier les données peuvent être progressivement censurées. À cet effet, une distribution souvent utilisée dans les études sur la durée de vie et la fiabilité est la distribution de valeurs extrêmes généralisée. Dans cette présentation, je combinerai les concepts de la censure progressive et de la distribution de valeurs extrêmes généralisée. Des procédures d'inférence seront discutées et exemplifiées par rapport à ces deux concepts. Afin de comprendre la performance de ces méthodes, des exemples et des résultats de simulations de Monte Carlo seront discutés.

### [MS-62]

#### Estimating Survival Function Under Dependent Truncation

#### Estimation de la fonction de survie sous une troncation dépendante

Lajmi LAKHAL CHAIEB, Louis-Paul RIVEST & Belkacem ABDOUS, *Université Laval*

The product-limit estimator for data subject to random left-truncation rely on the testable assumption of quasi-independence between failure time  $X$  and truncation time  $T$ . In the talk, we present a method to estimate the failure survival function  $S_X$  when it is related to the truncation time cumulative distribution function  $F_T$  by a known Archimedean semisurvival copula. After defining the semisurvival copulas, we will show that the survival estimator can be expressed in terms of the unknown probability  $c = P(X > T)$  estimator. This estimator as well as its asymptotic properties are presented. Finally, we will investigate our survival estimator properties.

L'estimateur classique pour des données tronquées à gauche est basé sur l'hypothèse testable de quasi-indépendance entre le temps de survie  $X$  et la variable de troncation  $T$ . Dans cet exposé, on présente une méthode d'estimation de la fonction de survie lorsqu'on suppose qu'elle est liée à la fonction de répartition de la troncation par une copule archimédienne de semi-survie. Après avoir défini les copules de semi-survie, on montre que l'estimateur de survie peut s'écrire en termes de l'estimateur de la probabilité inconnue  $c = P(X > T)$ . On présentera ensuite les propriétés asymptotiques de cet estimateur. Enfin, on étudiera les propriétés de notre estimateur de survie.

### [MS-63]

#### Accelerated Life Regression Modeling of Dependent Bivariate Time-to-event Data

#### Modélisation par régression du temps de vie accéléré pour des données bivariées dépendantes du temps jusqu'à un événement

Yun Hee CHOI & David E. MATTHEWS, *University of Waterloo*

To analyze bivariate time-to-event data from paired study designs, researchers frequently rely on a random effect called the frailty, which represents a risk factor between response measurements in a pair and induces a within-pair dependence. We developed a computational framework for fitting accelerated life regression models of dependent bivariate time-to-event data. Our framework allows the user to combine various choices of frailty distributions with different conditional distributions of the time-to-event of interest within a pair, given the frailty. To illustrate the flexibility of our frame-

Pour analyser des données bivariées du temps jusqu'à un événement provenant d'études appariées, les chercheurs comptent souvent sur un effet aléatoire appelé la « faiblesse » qui représente un facteur de risque entre les mesures de réponse dans une paire et induit une dépendance intra-paire. Nous avons développé un cadre informatique pour ajuster des modèles de régression de temps de vie accéléré à des données bivariées dépendantes de temps jusqu'à un événement. Notre cadre permet à l'utilisateur de combiner divers choix de distributions de faiblesse avec différentes distributions conditionnelles du temps jusqu'à l'événement d'intérêt dans une paire, sachant la faiblesse. Pour illustrer la flexibi-

work, we will describe results obtained using various model combinations for the diabetic retinopathy study data.

lité de notre cadre, nous décrivons les résultats obtenus en utilisant diverses combinaisons de modèles sur les données de l'étude sur la rétinopathie diabétique.

**Session 4A                      Monday, May 31 • Lundi 31 mai, 15:30 – 17:00                      AA-1140**

In Honour of Jim Zidek, on his 65th Birthday

En hommage à Jim Zidek, pour ses 65 ans

[MS-64]

**Bayesian Approaches to Environmental Network Design and Spatial Prediction**

**Approche bayésienne pour le design de réseaux environnementaux et la prédiction spatiale**

Nhu LE, *BC Cancer Agency & University of British Columbia*

In this talk, I will describe important challenges associated with re/designing environmental networks, as well as with the related-area of spatial prediction. One of the key challenges is the selection of design criteria. One approach pioneered by Caselton and Zidek (1984) overcomes this selection problem. This fundamental work has led to many subsequent developments relevant for specific environmental applications. One requirement in these applications is the predictive distribution of the random fields, often exhibiting nonstationarity characteristics. That presents another important challenge. Bayesian approaches for dealing with such problems will be discussed, along with their limitations and some future research directions.

Dans cette présentation, je vais décrire les défis importants associés à la planification des réseaux environnementaux ainsi qu'avec le contexte relié de la prédiction spatiale. Un des principaux problèmes est la sélection des critères de planification. Une approche instaurée par Caselton et Zidek (1984) résout ce problème de sélection. Ce travail fondamental a mené à plusieurs développements subséquents utiles pour des applications environnementales spécifiques. Un des besoins de ces applications est la distribution prédictive du champ aléatoire qui possède souvent des caractéristiques non stationnaires. Ceci présente un autre défi important. L'approche bayésienne pour travailler avec de tels problèmes sera discutée avec ses limitations et quelques directions pour des recherches futures.

[MS-65]

**Weighted Likelihood and its Applications**

**Vraisemblance pondérée et ses applications**

Feifang HU, *University of Virginia*, Jim ZIDEK, *University of British Columbia*

In this talk, we consider a weighted version of the classical likelihood that applies when the need is felt to diminish the role of some of the data in order to trade bias for precision. This weighted likelihood has been applied to predict the NBA basketball playoff outcomes. The method proves to be quite accurate.

Dans cette présentation, nous considérons une version pondérée de la vraisemblance classique qui s'applique lorsque nous voulons diminuer le rôle de certaines des données afin de compenser le biais pour la précision. Nous avons appliqué cette vraisemblance pondérée pour prédire les résultats des éliminatoires du basket-ball de la NBA. La méthode s'avère très précise.

[MS-66]

**Group-Bayes Estimation****Estimation bayésienne basée sur un consensus**Constance VAN EEDEN, *University of British Columbia*

In this talk I will present some of the results of Jim and myself on Group-Bayes estimation of an exponential mean. For the special case where the Bayesians use the same utility function, but different (conjugate) priors, G-(in)admissible, as well as G-minimax estimators are discussed. Here, the “assessment profile” of an estimator, i.e., the Bayesian’s expected utilities of that estimator, is used as an analogue of the risk function in Wald’s decision theory. The results include the case where the parameter indexing the family of priors is restricted.

Dans cette présentation, j’exposerai certains des résultats obtenus par Jim et moi-même sur l’estimation bayésienne basée sur un consensus d’une moyenne exponentielle. Pour le cas spécial où les bayésiens utilisent la même fonction d’utilité, mais différentes densités a priori (conjuguées), des estimateurs G-admissible (inadmissible), ainsi que des estimateurs G-minimax sont discutés. Ici, le « profil d’évaluation » d’un estimateur, c.-à-d. l’utilité espérée bayésienne de cet estimateur, est utilisé comme l’analogue de la fonction de risque dans la théorie de la décision de Wald. Les résultats incluent le cas où le paramètre indexant la famille des lois *a priori* est restreint.

**Session 4B****Monday, May 31 • Lundi 31 mai, 15:30 – 17:00****Z-310****Causal Inference in Biostatistics and Epidemiology****L’inférence causale en biostatistique et en épidémiologie**

[MS-67]

**Causal Inference in Randomized Trials in the Presence of Partial Compliance****Inférence causale dans les études cliniques randomisées en présence de conformité partielle**Robert PLATT & Tong GUO, *McGill University*

Noncompliance in randomized clinical trials is at the heart of the intention-to-treat (ITT) framework, and adjustment for noncompliance has received considerable attention. However, the focus is limited to dichotomous classification of compliance behavior. We extend the description of compliance behavior from all-or-none compliance defined by Imbens and Rubin (1997) to full-partial-none compliance. In addition to complier average causal effect (CACE), we define average causal effects for partial compliance using the framework of the Rubin causal model. We apply our new methodology to a breastfeeding promotion intervention trial to evaluate the impact of full and partial breastfeeding on infant infectious outcomes.

La non conformité dans des études cliniques randomisées est au coeur du cadre d’analyse du projet thérapeutique (APT), et l’ajustement pour la non conformité a suscité une attention considérable. Cependant, cette attention est limitée à la classification dichotomique des comportements de conformité. Nous prolongeons la description du comportement de conformité de celle tout ou rien définie par Imbens et Rubin (1997) vers une description de conformité totale, partielle et aucune. En plus de l’effet causal moyen du conformant (ECMC), nous définissons les effets causaux moyens pour la conformité partielle en utilisant le cadre du modèle causal de Rubin. Nous appliquons notre nouvelle méthodologie à une étude d’intervention sur la promotion de l’allaitement au sein pour évaluer l’impact de l’allaitement complet et partiel sur les résultats infectieux infantiles.

**[MS-68]****Causal Inference for Time-varying Exposures****Inférence causale pour des expositions variables dans le temps**Miguel HERNÁN, *Harvard University*

Even in the absence of unmeasured confounding factors or model misspecification, standard methods for estimating the causal effect of time-varying exposures on survival are biased when (i) there exists a time-dependent risk factor for survival that also predicts subsequent exposure, and (ii) past exposure history predicts subsequent risk factor level. In contrast, marginal structural models (MSMs) can provide consistent estimates of causal effects when unmeasured confounding and model misspecification are absent. We will describe the application of a marginal structural Cox proportional hazards model to estimate the effect of highly active antiretroviral therapy on time to AIDS or death.

Même en l'absence de facteurs parasites non mesurés ou de mauvaise spécification du modèle, les méthodes usuelles pour estimer l'effet causal de l'exposition variable dans le temps sur la survie sont biaisées quand (i) il existe un facteur de risque qui dépend du temps pour la survie qui prédit également l'exposition ultérieure, et (ii) quand l'exposition antérieure prédit le niveau ultérieur du facteur de risque. En revanche, les modèles structuraux marginaux (MSM) peuvent donner des estimations convergentes de l'effet causal quand des facteurs parasites non mesurés et la mauvaise spécification du modèle sont absents. Nous décrirons l'application d'un modèle structurel marginal de taux de panne proportionnel de Cox pour estimer l'effet de la thérapie antirétrovirale très active sur le temps jusqu'au SIDA ou la mort.

**Session 4C****Monday, May 31 • Lundi 31 mai, 15:30 – 17:00****AA-1360****New Methodologies for Census 2006****Nouvelles méthodologies pour le recensement de 2006****[MS-69]****Using an Address Register for a Mail-out Census****Utilisation d'un registre d'adresse pour un recensement par voie postale**Kathy MCCLEAN & Kim CHARLAND, *Statistics Canada/Statistique Canada*

Statistics Canada is planning on conducting its first mail-out Census of Population in 2006. Questionnaires will be mailed to two-thirds of dwellings. An Address Register is used and must be updated and geocoded to the geographic base. A complete field check of the list will be conducted in the fall of 2005. Each address verified in the field must also meet Canada Post Corporation addressing standards. This paper will describe the various steps and problems encountered in using an Address Register as a mailing list.

Statistique Canada projette d'effectuer son premier recensement de la population par voie postale en 2006. Des questionnaires seront expédiés aux deux-tiers des logements. Un registre d'adresse est utilisé et doit être mis à jour et géocodé sur une base géographique. Un contrôle complet sur le terrain de la liste sera effectué à l'automne 2005. Chaque adresse vérifiée sur le terrain doit également rencontrer les normes postales de Poste Canada. Cette présentation décrit les diverses étapes et problèmes rencontrés en utilisant un registre d'adresse comme liste d'expédition.

**[MS-70]****Census Questionnaire Design for Automated Data Capture****La conception de questionnaire du recensement dans le cadre d'une saisie automatisée des données**Laurent ROY & Manon MONETTE, *Statistics Canada/Statistique Canada*

In the 2006 Census of population, optical scanning and character recognition will be used to capture data on the long and short paper forms. How-

Lors du Recensement de la population de 2006, les techniques de lecture optique et de reconnaissance de texte seront utilisées pour la saisie des données des questionnaires

ever, such techniques cannot be used on the current forms, mainly because the response spaces for open questions cannot be readily scanned. A number of qualitative and quantitative surveys have been conducted with the aim of improving the forms so that the data can be scanned in. The objective was to recommend the best questionnaire formats based on certain criteria. The presentation will focus on the results of the key findings.

papiers longs et courts. Les questionnaires actuels ne permettent pas de faire cette opération : la principale lacune étant les espaces réponses des questions ouvertes qui ne permettent pas de faire cette opération convenablement. Plusieurs études qualitatives et quantitatives ont eu lieu pour améliorer les questionnaires et permettre la lecture optique de ceux-ci pour la saisie des données. L'objectif était de déterminer le format de questionnaire qui convenait le mieux selon certains critères. La présentation portera sur les résultats des différentes études.

### [MS-71]

#### **Census of Population Data Collection via Internet**

#### **La collecte des données du recensement de la population par Internet**

Danielle LAROCHE & Graeme GILMOUR, *Statistics Canada/Statistique Canada*

In the 2006 Census, a new data collection mode will be offered to all households in Canada. Respondents will have the option of completing and returning their questionnaire via the Internet. This new mode of collection will complement traditional census data collection modes. The presentation will give a brief overview of some of the requirements that must be satisfied in order to implement an Internet census, and of the main characteristics of Statistics Canada's electronic questionnaire. The discussion will be followed by a presentation of the electronic questionnaire that will be tested in May 2004, during the Census Test.

Dans le cadre du Recensement de 2006, un nouveau mode de collecte des données sera offert à tous les ménages au Canada. Les répondants auront la possibilité de remplir et de retourner leur questionnaire par Internet. Ce nouveau mode de collecte constitue un complément aux modes traditionnels de collecte des données du recensement. La présentation donnera un bref aperçu de certaines des exigences auxquelles il faut satisfaire pour assurer la mise en oeuvre d'un recensement par Internet, ainsi que des principales caractéristiques du questionnaire électronique de Statistique Canada. Cette discussion sera suivie de la présentation du questionnaire électronique qui sera mis à l'essai en mai 2004, au cours du Test du recensement.

### [MS-72]

#### **Statistical Literacy: How Statistics Canada Promotes It**

#### **Comment Statistique Canada fait la promotion de l'alphabétisme statistique**

Yves SAINT-PIERRE & Joel YAN, *Statistique Canada/Statistique Canada*

As the nation's statistical agency, Statistics Canada's role is not only to produce reams of data about our country's social and economic life, but also to help Canadians of all ages make sense of that information. Statistics Canada promotes statistical literacy in many ways in its communications with the public (its website in particular), and also by its Education Outreach program. More than access to data, the program gives numerous opportunities for learning from grade school to university, as well as tools to manipulate data. We will present some of these initiatives.

À titre de bureau national de la statistique, le rôle de Statistique Canada ne consiste pas seulement à produire une foule de données sur la vie sociale et économique du pays, mais également à aider les citoyens de tous âges à comprendre cette information. Statistique Canada fait de nombreux efforts dans ce sens dans ses communications avec le public (son site Internet en particulier) et par son programme de soutien à l'éducation. Plus qu'un accès aux données, le programme offre de nombreuses occasions d'apprentissage de l'élémentaire à l'université, et des outils de manipulation des données. Nous présenterons certaines de ces initiatives.

**Session 4D****Monday, May 31 • Lundi 31 mai, 15:30 – 17:00****Z-330**Applied Probability  
Probabilité appliquée**[MS-73]****On the First-visit-time Problem for Birth and Death Processes with Catastrophes****Sur le problème du premier temps de visite pour des processus de naissance-mort avec catastrophes**Antonio DI CRESCENZO, V. GIORNO & A. G. NOBILE, *Università di Salerno*, L. M. RICCIARDI, *Università Federico II, Napoli*

For a birth-death process subject to catastrophes, defined on the state-space  $\{r, r+1, r+2, \dots\}$ , with  $r$  a nonnegative integer, the first-visit time to a state  $k$  is considered and the Laplace transform of its probability density function is determined, use of which is then made to obtain mean and variance. The Laplace transform of the probability density function of the first effective catastrophe occurrence time and its expected value are also obtained. Some extensions to time-nonhomogeneous processes are then provided. Finally, certain additional results concerning the determination of the steady-state distribution and the representation of the transition probabilities are worked out.

Pour un processus de naissance-mort sujet aux catastrophes définies sur l'espace  $\{r, r+1, r+2, \dots\}$ , avec  $r$  un nombre entier non négatif, le temps de la première visite à un état  $k$  est considéré et la transformation de Laplace de sa fonction de densité est déterminée, laquelle est alors utilisée pour obtenir la moyenne et la variance. La transformation de Laplace de la fonction de densité du premier temps de catastrophe efficace et sa valeur espérée sont également obtenues. Quelques prolongements aux processus non homogènes par rapport au temps sont alors fournis. En conclusion, quelques résultats additionnels au sujet de la détermination de la distribution à l'équilibre et la représentation des probabilités de transition sont établis.

**[MS-74]****Boundary Crossing Problems with Applications to Option Pricing****Problème de croisement de frontière avec application en tarification d'options**Alexander NOVIKOV, *University of Technology, Sydney*, K. BOROVKOV, *The Univeristy of Melbourne*

Boundary crossing problems will be examined in the framework of two models frequently exploited in option pricing theory. Firstly, we consider the standard Brownian motion and present bounds for convergence rates of approximating curvilinear boundary crossing probabilities to similar probabilities. Applications to pricing of continuously or discretely monitored barrier options are discussed as well. Secondly, we consider the generalized Ornstein–Uhlenbeck process and describe cases when the Laplace transform of the level crossing can be obtained in an explicit form. As an application we shall consider the problem of path-dependent option pricing in the framework of the affine term structure model.

Les problèmes de croisement de frontière seront examinés dans le cadre de deux modèles fréquemment exploités dans la théorie de la tarification d'options. Premièrement, nous considérons le mouvement brownien standard et les bornes actuelles pour les taux de convergence pour l'approximation des probabilités de croisement de frontières curvilignes aux probabilités semblables. Les applications à la tarification d'options à frontière suivie continuellement ou discrètement sont également discutées. Deuxièmement, nous considérons le processus généralisé d'Ornstein-Uhlenbeck et décrivons les cas où la transformation de Laplace au niveau de frontière peut être obtenue de manière explicite. Comme application, nous considérerons le problème de l'évaluation du prix d'une option dépendant du chemin parcouru dans le cadre du modèle du terme de structure affine.



[MS-75]

**Boundary crossing probability for diffusion processes and piecewise continuous boundaries**  
**Probabilité de croisement de frontière pour des processus de diffusion et des frontières continues par morceaux**

Liqun WANG, *University of Manitoba*

I propose a direct approach to compute the boundary crossing probability, or first passage time distribution, for a class of diffusion processes which can be expressed as piecewise monotone (not necessarily one-to-one) functionals of the standard Brownian motion. This class includes many interesting processes in real applications, e.g., Ornstein-Uhlenbeck processes, growth processes and geometric Brownian motion with time-dependent drift. The proposed method applies to both one-sided and two-sided boundaries which may be nonlinear and/or discontinuous. Using this method, it is possible to derive closed-form formulas for certain nonlinear barriers, which are useful in evaluation and comparison of different computational methods and algorithms.

Je propose une approche directe afin de calculer les probabilités de croisement de frontière ou les distributions du temps de premier croisement pour une classe de processus de diffusion qui peuvent être exprimés comme des fonctionnelles monotones par morceaux (sans être nécessairement biunivoques) du mouvement Brownien standard. Cette classe comprend plusieurs processus d'intérêt pour de vraies applications, par exemple les processus de Ornstein-Uhlenbeck, les processus de croissance et le mouvement Brownien géométrique avec une déviation dépendante du temps. La méthode proposée peut être appliquée autant aux frontières unilatérales ou bilatérales qui peuvent être non linéaires et/ou discontinues. En utilisant cette méthode, il est possible d'obtenir des expressions sous forme analytique pour certaines barrières non linéaires qui sont utiles pour l'évaluation et la comparaison de différentes méthodes et algorithmes de calcul.

**Session 4E****Monday, May 31 • Lundi 31 mai, 15:30 – 17:00****Z-317**

Regression  
 Régression

[MS-76]

**Estimation of the Intercept Parameters of Two Simple Regression Models When the Slopes Are Suspected to Be Equal**

**Estimation du paramètre d'ordonnée à l'origine de deux modèles de régression simple quand les pentes sont soupçonnées être égales**

Shahjahan KHAN, *University of Southern Queensland*

Estimation of intercept parameter of a simple regression model involves the estimation of the slope. This presentation considers the estimation of the intercepts of two regression models, when the slopes are a priori suspected to be equal, but in doubt. Three different estimators are defined by using the sample data, nonsample uncertain prior information, an appropriate test statistic, and the coefficient of distrust. The relative performances of the unrestricted, shrinkage restricted and shrinkage preliminary test estimators are investigated based on the analysis of the bias and risk functions. An example based on a medical study is used to illustrate the method.

L'estimation du paramètre d'ordonnée à l'origine d'un modèle de régression simple implique l'estimation de la pente. Cette présentation considère l'estimation de ce paramètre dans deux modèles de régression, quand les pentes sont a priori soupçonnées d'être égales, mais un doute persiste. Trois estimateurs différents sont définis en utilisant l'échantillon, l'information incertaine *a priori* non échantillonnale, une statistique de test appropriée et le coefficient de méfiance. Les performances relatives des estimateurs du test sans restriction, restreint par rétrécissement et avec rétrécissement sont étudiées sur la base de l'analyse des fonctions de biais et de risque. Un exemple basé sur une étude médicale est utilisé pour illustrer la méthode.

[MS-77]

**Improved Estimation of Regression Parameters in Measurement Error Models****Estimation améliorée des paramètres de régression dans des modèles d'erreur de mesure**Hyang Mi KIM & A. K. Md. E. SALEH, *Carleton University*

This paper deals with improved estimation of the regression parameters in measurement error models when the reliability ratio matrix is known. We consider empirical Bayes estimators which include (i) the unrestricted unbiased (UE) estimator, (ii) the preliminary test estimator (PTE), (iii) the James–Stein estimator (SE), and (iv) the Positive-Rule Stein estimator (PRSE). The risks under the squared loss function are evaluated for all the four estimators and compared.

Cette présentation traite de l'estimation améliorée des paramètres de régression dans des modèles d'erreur de mesure quand la matrice du rapport de fiabilité est connue. Nous considérons les estimateurs empiriques de Bayes qui incluent : (i) l'estimateur sans biais sans restriction (UE), (ii) l'estimateur de test préliminaire (PTE), (iii) l'estimateur de James-Stein (SE), et (iv) l'estimateur de la règle positive de Stein (PRSE). Les risques sous la fonction de perte quadratique sont évalués pour les quatre estimateurs et sont comparés.

[MS-78]

**Estimation of Regression Parameters: A Two-samples Problem****Estimation des paramètres de régression : un problème à deux échantillons**Bashir KHAN, *Saint Mary's University*

Consider a system of two regression models  $y_\alpha = X_\alpha \text{Beta}_\alpha + e_\alpha$  ( $\alpha = 1, 2$ ). Our problem is the estimation of  $(p_\alpha \times 1)$  vector of unknown regression parameters,  $\text{Beta}_\alpha = (\beta_{\alpha 1}, \beta_{\alpha 2}, \dots, \beta_{\alpha p_\alpha})$ , when it is suspected but not sure that  $H_1 \text{Beta}_1 = H_2 \text{Beta}_2$  may hold where  $H_\alpha$  is a  $(q \times p_\alpha)$  matrix of known constants. Five estimators namely; unrestricted, restricted, preliminary test (PTE), James–Stein (SE) and its positive version (PPSE) are proposed and their properties are studied via risk matrices and quadratic risk criteria.

Considérez un système de deux modèles de régression  $y_\alpha = X_\alpha \text{Beta}_\alpha + e_\alpha$  ( $\alpha = 1, 2$ ). Notre problème est l'estimation d'un vecteur de  $(p_\alpha \times 1)$  paramètres de régression inconnus,  $\text{Beta}_\alpha = (\beta_{\alpha 1}, \beta_{\alpha 2}, \dots, \beta_{\alpha p_\alpha})$ , quand on suspecte que  $H_1 \text{Beta}_1 = H_2 \text{Beta}_2$  soit valide où  $H_\alpha$  est une matrice  $(q \times p_\alpha)$  de constantes connues. Cinq estimateurs notamment ; sans restriction, restreint, test préliminaire (PTE), le James-Stein (SE) et sa version positive (PPSE) sont proposés et leurs propriétés sont étudiées par l'intermédiaire des matrices de risque et du critère de risque quadratique.

[MS-79]

**Multiple Regression with Model Selection: A Bootstrap Approach****Régression linéaire adaptative multiple : une approche bootstrap**Pierre LAFAYE DE MICHEAUX, *Université Pierre Mendès France*, Christian LÉGER, *Université de Montréal*

It is well known (Hurvich and Tsai (1984) or Zhang (1992)) that confidence regions taking into account model selection in multiple regression have a bigger size than if the model is supposed to be known. In this work, we introduce bootstrap methods in order to construct confidence intervals that take into account all the original variables and the selection made. We only consider consistent selection methods. This strategy has already been numerically evaluated (Garriguet, 1999). Here, we are interested in theoretical aspects: validity of the bootstrap is established.

Il est bien connu (Hurvich et Tsai(1984) ou Zhang(1992)) que les régions de confiance tenant compte de la sélection de modèle en régression multiple ont une taille plus importante que si le modèle est supposé connu. Dans ce travail, nous introduisons des méthodes bootstrap de façon à construire des intervalles de confiance qui tiennent compte de toutes les variables d'origine et de la sélection qui en a été faite. Nous ne considérons que des méthodes de sélection convergentes. Cette stratégie a déjà été évaluée numériquement (Garriguet, 1999). Ici, nous nous intéressons à l'aspect théorique en établissant la validité du bootstrap.

**[MS-80]****Goodness of Fit of Multinomial Regression Models to Sparse Data  
Adéquation d'un modèle de régression multinomial à des données rares**Dianliang DENG, *University of Regina*

We derive approximations to the first three moments of the distribution of the Pearson chi-squared statistic, for testing goodness of fit of the multinomial models with non-canonical links, using an estimating-equations approach, for data that are extensive but sparse. A supplementary estimating equation is proposed from which the modified Pearson statistic is obtained. We also conduct a performance study of the usual Pearson statistic, the modified Pearson statistic and Edgeworth statistic derived in this paper, in terms of size and power, through a small-scale simulation experiment. One example and a discussion are given.

Nous dérivons des approximations pour les trois premiers moments de la distribution de la statistique du khi-carré de Pearson pour tester la qualité de l'ajustement des modèles multinomiaux avec des liens non canoniques, en utilisant une approche par équations estimantes pour des données qui sont considérables mais rares. Nous proposons une équation estimante supplémentaire à partir de laquelle la statistique de Pearson modifiée est obtenue. Nous entreprenons également une étude de performance de la statistique usuelle de Pearson, de la statistique de Pearson modifiée et de la statistique d'Edgeworth dérivée dans cette recherche, en termes de taille et de puissance, par une expérience de simulation à petite échelle. Nous donnons un exemple et une discussion.

**[MS-81]****Edgeworth-corrected Small-sample Confidence Intervals for Nonlinear Functions of Parameters in Linear Regression****Intervalles de confiance corrigés par Edgeworth sur de petits échantillons pour des fonctions non linéaires des paramètres d'une régression linéaire**Kamanzi BINYAVANGA, *University of Transkei, South Africa*, Sarel J. STEEL, *University of Stellenbosch, South Africa*

In this paper, we construct a central approximate confidence interval for a smooth scalar nonlinear function of the parameter-vector in a general linear regression model. We do this by first developing an Edgeworth expansion for the distribution function of a standardized equivariant point estimator. The confidence interval is then constructed in the manner discussed. A summary of simulation results reported at the end show that the interval has the potential to perform well in many small-sample situations.

Dans cette présentation, nous construisons un intervalle de confiance approximatif central pour une fonction non-linéaire scalaire lisse du vecteur des paramètres dans un modèle de régression linéaire général. Nous faisons ceci en développant d'abord une série d'Edgeworth pour la fonction de répartition d'un estimateur ponctuel équivariant normalisé. L'intervalle de confiance est alors construit de la manière discutée. Un résumé des résultats de simulations rapportés à la fin montre que l'intervalle a le potentiel de bien performer dans plusieurs situations où les échantillons sont de petites tailles.

**Session 4F****Monday, May 31 • Lundi 31 mai, 15:30 – 17:00****AA-1355**

Statistical Genetics

Statistique en génétique

**[MS-82]****Comparing Two Independent cDNA Microarrays  
Une comparaison de deux microréseaux indépendants**André DABROWSKI, *University of Ottawa/Université d'Ottawa*

The great interest in gene expression in molecular biology has led to the development of the microarray experiment. A very large number of genes can be tested at once for their level of expression through these DNA “chips.” This leads to data with relatively few observations per gene but a large number of simultaneous comparisons. Here we develop a parametric approach to establishing how many of the thousands of tests are indicative of a change in gene expression, and apply it to the case of two independent microarrays.

Le grand intérêt pour l'expression des gènes en biologie moléculaire a mené au développement de l'expérience microréseaux. Un très grand nombre de gènes peuvent être examinés d'un seul coup pour leur niveau d'expression par ces puces à ADN. Ceci donne des données avec relativement peu d'observations par gène, mais avec un grand nombre de comparaisons simultanées. Ici nous développons une approche paramétrique pour établir lesquels de ces milliers de tests sont indicatifs d'un changement dans l'expression des gènes et nous l'appliquons au cas de deux microréseaux indépendants.

**[MS-83]****On Semiparametric Familial-longitudinal Models  
Sur les modèles familiaux longitudinaux semi-paramétriques**Gary SNEDDON & Brajendra C. SUTRADHAR, *Memorial University of Newfoundland*

Familial-longitudinal data are collected from a large number of groups or families over a small period of time. This type of data exhibit correlation among family members, as well as among the repeated observations. We propose a semiparametric linear model with a two-way correlated error structure that accommodates the longitudinal and familial correlations. Estimation of the parametric regression function, the time trend function and the correlation parameters will be discussed. Simulation results will be presented to show that the procedure performs well when the errors are both normally distributed and arise from heavy-tailed distributions.

Des données familiales longitudinales sont rassemblées parmi un grand nombre de groupes ou de familles sur une petite période de temps. Ce type de données montre une corrélation entre les membres d'une même famille, ainsi qu'entre les observations répétées. Nous proposons un modèle linéaire semi-paramétrique avec une structure d'erreur corrélée bi-directionnelle qui s'ajuste à la corrélation longitudinale et familiale. L'estimation de la fonction de régression paramétrique, de la fonction de tendance temporelle et des paramètres de corrélation sera discutée. Des résultats de simulation seront présentés pour montrer que la méthode performe bien lorsque les erreurs sont normalement distribuées et proviennent d'une distribution à queue épaisse.

**[MS-84]****The Universal Validity of the Possible Triangle Constraints of the Affected-Sib-Pairs.  
La validité universelle de la contrainte triangulaire possible de la paire de fratrie affectée**Zeny FENG, Jiahua CHEN & Mary E. THOMPSON, *University of Waterloo*

In affected-sib-pair linkage analysis, under the hypothesis of no linkage, the probabilities of a sib pair sharing 0, 1 and 2 alleles IBD are 0.25, 0.5 and 0.25 respectively. Holmans (1993) showed that, when a disease is Mendelian and the recombination rates are same for male and female, the IBD distribution of a marker satisfies the “possible triangle

Dans l'analyse de liens de la paire dans la fratrie se trouvant affectée, sous l'hypothèse qu'il n'y a pas de liens, les probabilités qu'une paire de fratrie partagent 0, 1 et 2 allèles IPD (identité par descendance) sont 0.25, 0.5 et 0.25 respectivement. Holmans (1993) a montré que lorsqu'une maladie est mendélienne et que les taux de recombinaisons sont les mêmes pour les mâles et les femelles, la distribu-

constraint.” By a simple approach, we show this conclusion is true in general regardless of genetic models behind it. We also investigate the asymptotic properties of the likelihood ratio test under the possible triangle constraint.

tion IPD d'un marqueur satisfait la « contrainte triangulaire possible ». Par une approche simple, nous montrons que cette conclusion est vraie en général, indépendamment des modèles génétiques derrière elle. Nous étudions également les propriétés asymptotiques du test du rapport de vraisemblance sous la contrainte triangulaire possible.

[MS-85]

**Effects of Familial Correlation and Ascertainment Bias on Estimating the Penetrance of a Major Gene**  
**Effets des corrélations familiales et du biais de sélection sur l'estimation de la pénétrance d'un gène majeur**

Karen KOPCIUK, *Alberta Cancer Board*, L. BRIOLLAIS & W. HE, *Samuel Lunenfeld Research Institute*, E. PARKHOMENKO, *University of Toronto*, J. GREEN, *Memorial University of Newfoundland*, J MCLAUGHLIN, *Samuel Lunenfeld Research Institute*

Advances in the identification and treatment of genetically transmitted diseases have led to an increased need for reliable estimates of genetic susceptibility risk (penetrance). Most current methodologies available to estimate penetrance of genes involved in complex diseases are based on simplifying assumptions to model the familial correlation. Violations of such assumptions can lead to an underestimation of the variability of the penetrance estimate and a biased estimate of the penetrance. We will illustrate these problems using simulated and real data from an on-going study of a rare mutation and suggest appropriate methods for family-based studies.

Les avances sur l'identification et le traitement des maladies génétiquement transmissibles a entraîné un besoin croissant d'estimations fiables du risque de susceptibilité génétique (la pénétrance). La plupart des méthodes courantes disponibles pour estimer la pénétrance des gènes impliqués dans les maladies complexes sont basées sur des hypothèses simplificatrices pour modéliser les corrélations familiales. La violation de telles hypothèses peut entraîner une sous-estimation de la variabilité de l'estimation de la pénétrance et une estimation biaisée de la pénétrance. Nous illustrerons ces problèmes à partir de données simulées et réelles portant sur une étude en cours de pénétrance d'une rare mutation et proposerons des méthodes appropriées pour les études familiales.

[MS-86]

**Detecting Differentially Expressed Genes among Human Population**  
**Détecter les différentes expressions des gènes à travers la population humaine**

Cyr Emile M'LAN & Celia GREENWOOD, *Hospital for Sick Children, Toronto*, Shelley BULL, *Mount Sinai Hospital, Toronto*

The ability to measure thousands of mRNA transcript expressions simultaneously using high-throughput genomic technology has revolutionized the field of Genetics. In this talk, we will describe how microarray gene expression studies can be used to detect genes that may differentiate among human population subgroups. In our study, RNA from a collection of Lymphoblastoid Cell Lines were hybridized onto Affymetrix genechip arrays. ANOVA was performed to estimate effects due to ChipLot, Operator and Population group. We describe permutation methods for obtaining empirical significance levels in unbalanced designs, adjusting for multiple comparison. We also discuss com-

La capacité de mesurer des milliers d'expressions de transcription de mARN simultanément en utilisant des technologies génomiques qui traitent plusieurs données a révolutionné la génétique. Dans cette présentation, nous décrirons comment des études microréseaux d'expression génétique peuvent être utilisées pour détecter les gènes qui peuvent différencier des sous-groupes de la population humaine. Dans notre étude, l'ARN d'une collection de lignes de cellules lymphoblastoïdes a été hybridé sur des réseaux de puces à ADN Affymétrique. Une ANOVA a été exécutée pour estimer les effets dus au lot des puces, à l'opérateur et aux groupes de population. Nous décrivons des méthodes de permutation pour obtenir des niveaux de signification empiriques dans des plan non balancés, en ajustant pour les

putational complexities in computing permutation p-values efficiently.

comparaisons multiples. Nous discutons également des complexités informatiques pour calculer efficacement les permutations des valeurs-p.

**Session 4G**                      **Monday, May 31 • Lundi 31 mai, 17:00 – 18:30**                      **Z-330**

**NSERC Workshop: Tips to Prepare your Next Discovery Grant Application**

**Atelier du CRSNG : Conseils pour l'élaboration de votre prochaine demande de subvention à la découverte**

NSERC representatives and Grant Selection Committee members will present an interactive overview of the peer review process, inform you of the latest changes at NSERC, give useful advice for the preparation of your next NSERC application and answer your questions on the functioning of grant selection committees.

The workshop is open to all researchers. It is particularly helpful for new faculty members and researchers likely to apply (or re-apply) in the fall.

Des employés du CRSNG et des membres des comités de sélection des subventions du CRSNG donneront une présentation en mode interactif sur le processus d'évaluation par les pairs, vous renseigneront sur les derniers changements opérés au CRSNG, vous donneront des conseils utiles pour l'élaboration de votre prochaine demande de subvention au CRSNG et répondront à vos questions sur le fonctionnement des comités de sélection des subventions.

Tous les chercheurs peuvent participer à l'atelier, qui sera particulièrement utile aux professeurs et aux chercheurs récemment embauchés qui comptent présenter une demande (ou une nouvelle demande) à l'automne.

**Session 5**                      **Tuesday, June 1 • Mardi 1<sup>er</sup> juin, 08:30 – 10:00**                      **B-2285**

**Gold Medal Address**

**Allocution de la récipiendaire de la médaille d'or**

[MS-87]

**Understanding Associations: Implications for the Design and Analysis of Longitudinal Surveys**

**Comprendre les associations : Implications pour la planification et l'analyse d'enquêtes longitudinales**

Mary THOMPSON, *University of Waterloo*

I will discuss the place of longitudinal surveys in the collection of methods we bring to bear on understanding associations and discerning causal processes in the social sciences. The evaluation of interventions can sometimes take advantage of so-called "natural experiments." I will show how decisions concerning the sampling scheme, and the use of the design in estimation, can assist the experimental analogy. Illustrations will include studies on the role of media violence in promoting aggressive behavior, and the evaluation of tobacco control measures.

Je discuterai de la place des sondages longitudinaux à travers les méthodes que nous appliquons pour comprendre les associations et discerner les processus causals dans les sciences sociales. L'évaluation des interventions peut parfois tirer profit des expériences naturelles. Je montrerai comment les décisions au sujet du plan d'échantillonnage et l'utilisation du plan dans l'estimation peuvent aider l'analogie expérimentale. Les illustrations incluront des études sur l'influence de la violence dans les médias sur les comportements agressifs et l'évaluation des mesures de contrôle du tabac.

**Session 6A**                      **Tuesday, June 1 • Mardi 1<sup>er</sup> juin, 10:30 – 12:00**                      **AA-1140**

**Special Invited Session of the Biostatistics Section**

**Allocution de l'invité d'honneur du Groupe de biostatistique**

[MS-88]

**Key Features in the Design and Analysis of DNA Microarray Studies****Caractéristiques importantes dans la planification et l'analyse d'expériences microréseaux d'ADN**Richard SIMON, *National Cancer Institute, Rockville, MD*

I will present my perspectives on the design and analysis of DNA microarray studies. I will review evaluations of design approaches including dye swapping, use of a common reference, technical replicates and sample pooling. I will discuss the relative merits of several proposed approaches to class comparison, that is, finding genes that are differentially expressed in different tissues or conditions. I will present a simple but useful approach to planning sample size for class comparison studies. Finally, I will discuss approaches to class prediction in the setting where the number of candidate predictors ( $p$ ) greatly exceeds the number of cases ( $n$ ).

Je présenterai mes perspectives sur le design et l'analyse d'études microréseaux d'ADN. Je passerai en revue les évaluations des approches de design incluant la permutation des colorants, l'utilisation d'une référence commune, les répliques techniques et l'association d'échantillons. Je discuterai des mérites relatifs de plusieurs approches proposées pour classer les comparaisons, c.-à-d. trouver des gènes qui s'expriment différemment dans différents tissus ou conditions. Je présenterai une approche simple mais utile pour planifier la taille de l'échantillon pour des études de comparaison de classes. Finalement, je discuterai des approches pour prédire les classes dans la situation où le nombre de candidats prédicteurs ( $p$ ) excède considérablement le nombre de cas ( $n$ ).

**Session 6B****Tuesday, June 1 • Mardi 1<sup>er</sup> juin, 10:30 – 12:00****AA-1360****Analysis on Complex Surveys****L'analyse de données d'enquêtes complexes**

[MS-89]

**Analysis of Complex Sample Survey Data: An Appraisal****Analyses de données d'enquêtes complexes : une évaluation**J. N. K. RAO, *Carleton University*

In practice, data obtained from complex sample surveys often involve clustering, stratification, post-stratification and unequal sampling weights. As a result, application of standard methods to survey data without taking account of design features can lead to erroneous inferences: underestimation of standard errors of parameter estimates, inflated test levels and erroneous model diagnostics. New methods that take proper account of the complexity of survey data have been proposed. In this talk, I will provide an appraisal of some of the proposed methods. I will also present a unified approach based on resampling estimating functions.

En pratique, les données obtenues à partir d'enquêtes par échantillonnage complexes impliquent souvent la classification, la stratification, la post-stratification et des poids d'échantillonnage inégaux. Par conséquent, l'application des méthodes standards pour des données d'enquêtes sans tenir compte des caractéristiques du plan peut mener à des inférences inexactes : sous-estimation de l'écart type des estimés, gonflement du niveau des tests et diagnostics de modèle incorrects. De nouvelles méthodes qui tiennent compte de manière appropriée de la complexité des données d'enquêtes ont été proposées. Dans cette présentation, j'évaluerai certaines des méthodes proposées. Je présenterai également une approche unifiée basée sur des fonctions d'estimation de rééchantillonnage.

**[MS-90]****Analysis of Categorical Data from Complex Sample Surveys Using Inverse Sampling****Analyse de données catégoriques provenant de sondages complexes en utilisant l'échantillonnage inverse**Emmanuel BENHIN, *Statistics Canada/Statistique Canada*, J. N. K. RAO, *Carleton University*

Analysis of complex survey categorical data using classical statistical methods without taking into account the complex nature of the data may lead to asymptotically invalid statistical inferences. Several methods have been developed that account for the survey design, but these methods require additional information such as survey weights, design effects or cluster identification for micro data. An alternative approach is by undoing the complex data structures so that standard methods can be applied to data files generated by inverse sampling. We propose a combined estimating equation approach to analyze such data files in the context of categorical survey data.

L'analyse des données catégoriques de sondages complexes en utilisant des méthodes statistiques classiques sans tenir compte de la nature complexe des données peut mener à des inférences statistiques asymptotiquement invalides. Plusieurs méthodes qui tiennent compte du plan du sondage ont été développées, mais celles-ci exigent de l'information additionnelle telle que les poids d'échantillonnage, les effets du plan ou l'identification de grappes pour des micro données. Une approche alternative est de défaire les structures complexes des données de sorte que les méthodes standards peuvent être appliquées aux fichiers de données générés par échantillonnage inverse. Nous proposons une approche d'équations d'estimation combinée pour analyser de tels fichiers de données dans le contexte des données de sondage catégoriques.

**[MS-91]****Multi-level Modelling of Complex Survey Data****Modélisation multi-niveaux de données de sondages complexes**Danny PFEFFERMANN, *Hebrew University & University of Southampton*

I will discuss and illustrate the use of three different approaches for multi-level modeling (MLM) of complex survey data. Design based inference controls the bias under informative sampling but is basically restricted to point estimation. A more general approach is to extract the MLM holding for the sample from the model holding for the population and the sample selection probabilities. It is implemented conveniently by use of Bayesian methods. When the purpose of the analysis is the prediction of small area means, a third approach is to fit an MLM to the sample with no reference to the population model.

Je vais discuter et illustrer l'utilisation de trois approches différentes pour la modélisation à multi-niveaux (MMN) des données de sondages complexes. L'inférence basée sur la randomisation contrôle le biais sous l'échantillonnage informatif, mais elle est restreinte à l'estimation ponctuelle. Une approche plus générale est d'extraire la forme de la MMN pour l'échantillon à partir de la forme du modèle de la population et aux probabilités d'échantillonnage. Cette approche est mise en application aisément au moyen de méthodes bayésiennes. Quand le but de l'analyse est la prédiction de moyennes sur de petits domaines, une troisième approche est d'ajuster une MMN à l'échantillon sans référence au modèle de la population.

**Session 6C****Tuesday, June 1 • Mardi 1<sup>er</sup> juin, 10:30 – 12:00****Z-310****Applications of Wavelets to Statistical Problems****Applications des ondelettes à des problèmes statistiques**



**[MS-92]****How Statistics Can Help You See in the Dark!****Comment la statistique peut-elle vous aider à voir dans le noir ?**

Guy NASON, *University of Bristol*, Piotr FRYZLEWICZ, *Imperial College*

Haar–Fisz transforms are a new class of multi-scale variable transforms that can Gaussianize and variance-stabilize intensity sequences. Haar–Fisz transforms for Poisson intensity data perform better than Anscombe and are also usually better than specialist Bayesian algorithms. Haar–Fisz is also extremely fast and simple. We explain the Haar–Fisz transform, demonstrate its performance and demonstrate some low-light image enhancement applications.

Les transformations de Haar–Fisz forment une nouvelle classe de transformations de variables multi-échelles qui peuvent normaliser et stabiliser la variance de séquences d'intensité. Les transformations de Haar–Fisz pour des données d'intensité de Poisson performant mieux que Anscombe et sont habituellement meilleures que les algorithmes bayésiens spécialisés. La transformation de Haar–Fisz est également très rapide et simple. Nous expliquons la transformation de Haar–Fisz, nous démontrons sa performance et nous montrons quelques applications en perfectionnement d'image à basse-lumière.

**[MS-93]****Wavelet Kernel Penalized Estimation for Random Design Regression****Noyaux autoreproduisants à base d'ondelettes et débruitage de signaux sur plans déterministes non équadistants**

Anestis ANTONIADIS, *Université Joseph Fourier, Grenoble*, U. AMATO, *Consiglio Nazionale delle Ricerche, Napoli*, M. PENSKY, *University of Central Florida*

Consider the regression problem  $y_i = f(x_i) + e_i$ ,  $i = 1, \dots, n$ , where  $f$  is an unknown compactly supported regression function to be estimated,  $x_i$ 's are univariate design points, not necessarily equispaced, and the  $e_i$ 's are errors assumed to be independent centered random variables with an unknown common variance. In this work, we consider and study a class of nonseparable wavelet estimators for the nonparametric regression problem using a penalized least-squares approach with specific penalties. For this, we focus on wavelet kernels derived via Reproducing Kernel Hilbert Spaces (RKHS) which provide a unified context for regularization. Such estimations issues are well-known in the literature and have been studied by several authors in the general nonparametric setting of the smoothing spline framework. The resulting penalties will be shown equivalent to some Besov semi-norms. The resulting estimators are easy to compute, and have excellent properties in terms of prediction accuracy and interpretability.

Considérons le problème de régression  $y_i = f(x_i) + e_i$ ,  $i = 1, \dots, n$ , où  $f$  est une fonction inconnue à support compact, les  $x_i$  sont des points déterministes non nécessairement équadistants et les  $e_i$  sont des erreurs aléatoires indépendantes centrées et de même variance inconnues. Dans ce travail, nous proposons et étudions une classe d'estimateurs de la fonction inconnue  $f$  fondés sur des procédures de régularisation dérivées de noyaux hilbertiens auto-reproduisants à base d'ondelettes. Les pénalités résultantes apparaissent comme des semi-normes d'espace de Besov. Les estimateurs obtenus sont facilement calculables, interprétables et possèdent d'excellentes propriétés.

**[MS-94]****Some Applications of Wavelets in Functional Linear Models Analysis****Quelques utilisations des ondelettes dans l'analyse de modèles linéaires fonctionnels**R. OGDEN, *Columbia University*

In functional data analysis, collected data are regarded as continuous functions of another variable (e.g., spectra, images). It is common in such situations to represent functional data in terms of an orthogonal basis and apply multivariate statistical techniques. Using a wavelet basis allows not only the ability to represent non-smooth aspects of functional data but, owing to their inherent multiresolution structure, also allow analysis across multiple scales. The use of wavelet methods in functional linear models analysis is demonstrated using data from a psychological study on judgment of the collinearity of line segments as a function of angle.

Dans l'analyse de données fonctionnelles, les données collectées sont considérées comme des fonctions continues d'une autre variable (par exemple, spectres, images). Il est commun dans de telles situations de représenter les données fonctionnelles en fonction d'une base orthogonale et d'appliquer des techniques statistiques multivariées. L'utilisation de la base d'ondelettes permet non seulement de représenter des aspects non lisses des données fonctionnelles dus à la structure multi-résolution inhérente, mais permet également l'analyse à plusieurs résolutions. L'utilisation des méthodes d'ondelettes dans l'analyse de modèles linéaires fonctionnels est démontrée en utilisant des données d'une étude psychologique sur le jugement de la colinéarité de segments de lignes en fonction de l'angle.

**Session 6D****Tuesday, June 1 • Mardi 1<sup>er</sup> juin, 10:30 – 12:00****Z-330****Monte Carlo Methods: Theory and Applications****Méthodes de Monte-Carlo : Théorie et applications****[MS-95]****Asymptotic Properties of Monte Carlo Estimators for Diffusions****Propriétés asymptotiques des estimateurs pour les diffusions**Marcel RINDISBACHER, *University of Toronto*, J. DETEMPLE, *Boston University*, R. GARCIA, *Université de Montréal*

This paper studies the limit distributions of Monte Carlo estimators of diffusion processes. Expected approximation errors are characterized and used to construct second-order bias corrected estimators. Such bias correction eliminates the size distortion of asymptotic confidence intervals and allows to examine the relative efficiency of estimators. The variance-stabilizing transformation is again found to increase the speed of convergence. For comparison we also study the Milstein scheme. We derive new convergence results for this scheme. Our results are illustrated in the context of a dynamic portfolio choice problem and of simulated-based estimation of diffusion processes.

Cette présentation étudie les distributions limites des estimateurs Monte Carlo des processus de diffusion. Les erreurs d'approximation espérées sont caractérisées et utilisées pour construire des estimateurs corrigés pour le biais du deuxième ordre. Une telle correction du biais élimine la déformation de taille des intervalles de confiance asymptotiques et permet d'examiner l'efficacité relative des estimateurs. La transformation de stabilisation de la variance augmente à nouveau la vitesse de convergence. Pour des fins de comparaisons, nous étudions également le schéma de Milstein. Nous dérivons de nouveaux résultats de convergence pour ce schéma. Nos résultats sont illustrés dans le contexte d'un problème de sélection dynamique de portefeuille et dans le contexte de l'estimation simulée de processus de diffusion.

**[MS-96]****Decomposition Bounds for MCMC Convergence Rates****Limites de décomposition pour les taux de convergence de MCMC**Neal MADRAS, *York University*, John Wai Kong YUEN, *Brock University*

The rate of convergence of a Markov chain to its equilibrium distribution has been much studied in recent years. This talk describes methods that analyze convergence rates by decomposing a Markov chain into smaller pieces. The idea is that if the chain equilibrates rapidly on each piece, and if the chain moves from piece to piece efficiently, then the entire chain equilibrates rapidly. In some examples we obtain lower and upper bounds on the convergence rate that are reasonably close to one another. Such examples shed light on the issue of when one form of decomposition works better than another.

Le taux de convergence d'une chaîne de Markov vers sa distribution d'équilibre a beaucoup été étudié ces dernières années. Cette présentation décrit les méthodes qui analysent des taux de convergence en décomposant une chaîne de Markov en petits morceaux. L'idée est que si la chaîne tend vers l'équilibre rapidement sur chaque morceau et si la chaîne se déplace de morceaux en morceaux efficacement, alors la chaîne entière tend vers l'équilibre rapidement. Nous obtenons des limites inférieures et supérieures sur le taux de convergence qui sont raisonnablement près l'une de l'autre pour quelques exemples. Ceux-ci permettent de mieux comprendre lorsqu'une forme de décomposition fonctionne mieux que d'autres.

**[MS-97]****New Selection in a Branching and Interacting Particle Systems****Nouvelle sélection pour des systèmes de particules en interaction**Ivan GENTIL, *Université Paris Dauphine*, Bruno RÉMILLARD, *HEC Montréal*

In a work written in collaboration with Bruno Rémillard we are interested in improving the selection in an interacting particle system. These particle systems are used to approximate the optimal filter in nonlinear filtering problems. We show the first results leading to a convergence in  $L^2$  for the so-called "stratification" selection. We illustrate this work with simulations which illustrate that we can surely improve the results.

Dans un travail écrit en collaboration avec Bruno Rémillard nous nous intéressons à améliorer la sélection dans un système de particules en interaction. Ces systèmes de particules sont utilisés pour approcher le filtre optimal dans des problèmes de filtrages non linéaires.

Nous montrons les premiers résultats permettant d'obtenir un théorème de convergence  $L^2$  pour la sélection dites de « stratification ». Nous illustrons ce travail avec des simulations, celles-ci montrant que l'on peut sûrement obtenir des résultats plus satisfaisants.

**Session 6E****Tuesday, June 1 • Mardi 1<sup>er</sup> juin, 10:30 – 12:00****Z-317****Survey Sampling I****Méthodologie d'enquête I****[MS-98]****Issues with High Vacancy Areas and Remote Areas in the Canadian Labour Force Survey Sample Redesign****Problèmes avec les secteurs à faible densité et les régions éloignées dans la révision du design de l'Enquête sur la population active**Edward CHEN, J. LINDEYER & G. LAFLAMME, *Statistics Canada/Statistique Canada*

Canadian Labour Force Survey is the largest household survey in Canada to provide monthly estimates of employment, unemployment and labour

L'Enquête sur la population active canadienne est la plus grande enquête de ménages au Canada qui fournit des estimations mensuelles d'indicateurs d'emploi, de chômage et

market indicators. The coverage of the survey includes all geographic areas in Canada. Included in some provinces are remote areas in the vast north and high vacancy areas where data collection costs are high.

This paper describes a study undertaken to improve the current LFS design in the high vacancy areas and remote areas with the aim to reduce collection costs. 2001 Census and the ongoing LFS data quality information are used to evaluate and develop the new sample design.

du marché du travail. La couverture du sondage comprend toutes les régions géographiques du Canada. Inclus dans certaines provinces sont les régions éloignées du grand nord et les secteurs à faible densité, où les coûts de collecte de données sont élevés. Cette présentation décrit une étude entreprise pour améliorer le design actuel du LFS dans les secteurs à faible densité et les régions éloignées dans le but de réduire les coûts de collecte. Le recensement de 2001 et l'information continue des données du LFS sont utilisés pour évaluer et développer le nouveau plan d'échantillonnage.

### [MS-99]

#### **Person-level Sampling in the Canadian Nutrition Survey**

#### **Sélection des personnes dans l'enquête canadienne sur la nutrition**

Fritz PIERRE & Yves BÉLAND, *Statistics Canada/Statistique Canada*

The major objective of the Cycle 2.2 of the Canadian Community Health Survey (CCHS) is to estimate the distribution of usual dietary intake using a 24-hour dietary recall, with repeated measures. As well, data on food insecurity, anthropometric measurements, health status, health determinants and socio-demographic characteristics will also be collected. The survey will cover persons of all ages living in private dwellings in the 10 provinces (~ 36,000 interviews) in order to produce intake distributions for 15 key age-sex domains of interest. This paper will describe in depth the person-level sampling strategy, one main aspect of the sample design.

Le principal objectif du cycle 2.2 de l'Enquête sur la santé dans les collectivités canadiennes (ESCC) est d'estimer pour 15 domaines d'intérêt les distributions de consommation pour différents éléments nutritifs; l'information sera recueillie via un journal alimentaire de 24 heures avec seconds rappels chez les Canadiens de tout âge résidant dans des logements privés dans les 10 provinces (~ 36000 interviews). On collectera également des données sur la précarité alimentaire, les mesures anthropométriques, l'état de santé, les déterminants de la santé et les caractéristiques sociodémographiques. Cet article décrit la stratégie de la sélection des personnes, un aspect important du plan déchantillonnage.

### [MS-100]

#### **The Weighting Strategy for Cycle 2.1 of the Canadian Community Health Survey**

#### **Pondération du cycle 2.1 de l'Enquête sur la santé dans les collectivités canadiennes**

Marco GRENIER & François BRISEBOIS, *Statistics Canada/Statistique Canada*

The regional component of the Canadian Community Health Survey is conducted by Statistics Canada with a total sample of 130,000 respondents in order to provide region-level data for 133 health region. For the weighting of cycle 2.1, in addition to dealing with the complexities that were present in cycle 1.1, such as the use of two survey frames, new challenges arose. One example is the integration of a study measuring the collection mode effect (in person vs. telephone) on estimates. The strategy developed to produce the survey weights will be presented with emphasis placed on the new challenges.

La composante régionale de l'Enquête sur la santé dans les collectivités canadiennes est menée par Statistique Canada auprès d'environ 130 000 répondants et vise à produire des données à l'échelle de 133 régions socio-sanitaires. Pour la pondération du cycle 2.1, en plus d'avoir à composer avec les mêmes complexités qu'au cycle 1.1 telles que l'utilisation de deux bases de sondage, de nouveaux défis se sont présentés. Mentionnons, entre autres, l'intégration à l'enquête d'une étude visant à déterminer l'effet du mode de collecte (personne vs téléphone) sur les estimations. La stratégie développée pour la pondération sera présentée en mettant l'accent sur les défis rencontrés.

**[MS-101]****Estimating Lifetime Risk****Estimation du risque à vie**

Julie HORROCKS, *University of Guelph*, Mieko YOSHIHAMA & Kristin DUNKLE, *University of Michigan*, Saori KAMANO, *National Institute of Population and Social Security Research, Tokyo*

Early age of sexual initiation in women is associated with a higher lifetime risk of intimate partner violence. However the extent to which this increased risk is due to longer exposure is unclear. We discuss this issue in the context of a World Health Organization (WHO) study of intimate partner abuse in Yokohama, Japan, a stratified, multi-stage sample of 1,371 women between the ages of 18 and 49. We investigate methods for estimating lifetime risk from survey data, and compare methods for disentangling covariate effects from exposure length effects, using Cox proportional hazards models, graphical summaries and other techniques.

Le jeune âge de l'initiation sexuelle chez les femmes est associé à un risque à vie plus élevé de violence par un partenaire intime. Cependant le fait que cette hausse du risque est due à une plus longue exposition n'est pas clair. Nous discutons de ce problème dans le contexte d'une étude de l'Organisation mondiale de la santé (OMS) sur l'abus des partenaires intimes à Yokohama au Japon sur un groupe stratifié à plusieurs degrés de 1371 femmes entre 18 et 49 ans. Nous étudions des méthodes pour estimer le risque sur la durée de vie des données de sondage et comparons des méthodes pour démêler les effets des covariables et l'effet du temps d'exposition en utilisant les modèles de taux de panne proportionnels de Cox, des sommaires graphiques et d'autres techniques.

**[MS-102]****Statistical Problems with Inflation Measures in Transition Countries****Problèmes statistiques avec les mesures d'inflation dans les pays en transition**

Ante ROZGA, *University of Split/Université de Split*

Fixed basket of goods and services becomes less representative as consumers respond to price changes and new choices. This is connected with 'new economy,' globalization and transition countries. Sources of bias are introduction of new products and quality change. The results are overestimated Laspeyres price indices. GDP and other economic series could be underestimated, leading to wrong decisions by authorities. These problems are solved using hedonistic indices, expensive for transition countries. The share of 'gray economy' in GDP is higher than in EU. In Croatia, situation is better because there were elements of market economy before transition.

Le panier fixe des biens et des services devient de moins en moins représentatif alors que les consommateurs répondent aux changements de prix et aux nouveaux choix. Ceci est relié à la nouvelle économie, la globalisation et les pays en transition. Les sources de biais sont l'introduction de nouveaux produits et de changement de la qualité. Les résultats sont des indices de prix de Laspeyres surestimés. Le PIB et d'autres séries économiques ont pu être sous-estimés, menant à de fausses décisions par les autorités. Ces problèmes sont résolus en utilisant des index hédonistes dispendieux pour des pays en transition. La portion de « l'économie grise » dans le PIB est plus élevée que dans l'UE. En Croatie, la situation est meilleure parce qu'il y avait des éléments d'économie de marché avant la transition.

**Session 6F****Tuesday, June 1 • Mardi 1<sup>er</sup> juin, 10:30 – 12:00****AA-1355**

Biostatistics I

Biostatistique I

**[MS-103]****A Bayesian Advantage in Setting Passing Scores: The PPOp Curve  
Un avantage bayésien pour introduire des scores de passage : la courbe PPOp**

Howard WAINER, *National Board of Medical Examiners, Philadelphia*, Xiaohui WANG, *University of North Carolina*, William P. SKORUPSKI, *University of Massachusetts*, Eric T. BRADLOW, *University of Pennsylvania*

In this talk, we demonstrate an interesting use of the posterior distributions (and corresponding posterior samples of proficiency) that are yielded by fitting a fully Bayesian test scoring model to a complex assessment. Specifically, we examine the efficacy of the test in combination with the specific passing score that was chosen through expert judgment, or, in general, any external a priori criterion. This demonstration provides evidence of how this tool can be helpful in a truly life-and-death situation.

Dans cette présentation, nous illustrons une utilisation intéressante des distributions *a posteriori* (et des échantillons *a posteriori* de compétence correspondants) qui sont données en ajustant un modèle de note de test à une évaluation complexe entièrement bayésien. Spécifiquement, nous examinons l'efficacité du test en combinaison avec les notes de passage spécifiques provenant de jugements d'experts ou en général, n'importe quel critère *a priori* externe. Cette démonstration fournit un exemple sur la manière dont cet outil peut être utile dans une véritable situation de vie ou de mort.

**[MS-104]****Using Random Datasets and Automatic Marking in a Web Supported Biostatistics Module  
Utilisation de jeux de données aléatoires et notation automatique dans un module de biostatistique sur internet**

Paul COREY, Malcolm KOO, Raymond TAM & Andrew WHITE, *University of Toronto*

Using ASP, MySQL and SAS, a Web module was developed to support a graduate medical statistics course. Students, on line, receive random datasets generated using SAS. Their analysis, using a statistical package of choice, is tested in a multiple-choice format.

Using feedback from online quizzes, students take online timed exams using random datasets. Results from quizzes and exam are automatically marked and made available.

This module can be used to support a course in statistics or assess statistical proficiency as prerequisite for a course or program.

Technical and pedagogic issues along with the results of a satisfaction survey are discussed.

En utilisant ASP, MySQL et SAS, un module Web a été développé dans le cadre d'un cours gradué de statistique médicale. Les étudiants reçoivent, en ligne, des jeux de données aléatoires générés par SAS. En utilisant un package statistique de leur choix, les analyses des étudiants sont testées dans un format de choix-multiples. En utilisant le feedback des quiz effectués sur Internet, les étudiants passent des examens en ligne chronométrés en utilisant des jeux de données aléatoires. Les quiz et les examens sont notés automatiquement et les résultats sont rendus disponibles. Ce module peut être utilisé dans le cadre d'un cours de statistique ou pour évaluer la compétence statistique comme préalable à un cours ou à un programme. Des problèmes techniques et pédagogiques ainsi que les résultats d'une enquête de satisfaction sont discutés.

**[MS-105]****Analysis of Developmental Trajectories and Binary Longitudinal Data  
Analyses de trajectoires de développement mental et données longitudinales binaires**

Victor NAWA & K. S. BROWN, *University of Waterloo*

Developmental trajectories can be used to study how behavior changes over time or age. For example, in studying smoking onset, different patterns of onset may be identified and related to charac-

La trajectoire de développement mental peut être utilisée pour étudier comment le comportement change dans le temps ou avec l'âge. Par exemple, pour étudier le fait de commencer à fumer, différents modèles de commencement

teristics of the individuals. Recently methods designed to identify homogeneous clusters of individuals with similar developmental trajectories have been proposed. We propose using the EM algorithm to classify subjects to different developmental trajectories and to examine the role of covariates in distinguishing these groups. We also look at the question of starting values and propose some methods of dealing with missing values.

peuvent être identifiés et liés aux caractéristiques des individus. Récemment, des méthodes ont été proposées pour identifier les grappes homogènes d'individus avec des trajectoires de développement mental semblables. Nous proposons d'utiliser l'algorithme EM pour classer des sujets par rapport à différentes trajectoires de développement mental et pour examiner le rôle des covariables pour distinguer ces groupes. Nous regardons également la question des valeurs de commencement et proposons quelques méthodes pour traiter des valeurs manquantes.

### [MS-106]

#### **Detecting Conjunctions in Images**

#### **Détection des conjonctions dans les images**

Moh'd ALODAT, *McGill University*

A conjunction is defined as the occurrence of two or more events at the same location in space. The events are where a smooth image of a test statistic exceeds a fixed threshold. The conjunction is the intersection of these random excursion sets. We use the conjunction to detect brain activation at a common location from fMRI (functional magnetic resonance imaging). In this abstract we use the spatial extent of the conjunction as a test statistic and find its approximate p-value using random field theory.

Une conjonction est définie comme l'occurrence de deux événements ou plus au même endroit dans l'espace. Les événements sont là où une image lisse d'une statistique de test excède un seuil fixe. La conjonction est l'intersection de ces ensembles aléatoires d'excursion. Nous utilisons la conjonction pour détecter l'activation du cerveau à un endroit commun avec des images IRMF (imagerie par résonance magnétique fonctionnelle). Dans cette présentation, nous utilisons le volume spatial de la conjonction comme statistique de test et trouvons sa valeur-p approximative en utilisant la théorie des champs aléatoires.

### [MS-107]

#### **Estimating Sensitivity & Specificity for Binocular Data: Application to Ophthalmology**

#### **Estimation de la sensibilité et de la spécificité pour des données binoculaires : applications en ophtalmologie**

Alexander DE LEON, *University of Calgary*, C. RUDNISKY, *University of Alberta*, G. SINGH, *University of Calgary*

Binocular data typically arise in ophthalmology, where pairs of eyes are screened, through some diagnostic test, for the presence or absence of certain abnormalities or pathologies. Treating the eyes as independent and adopting the usual approach in estimating sensitivity and specificity of the test ignores the correlation between the eyes, and may consequently yield incorrect estimates. We propose a likelihood-based method of estimating sensitivity and specificity using a parametric model for paired binary outcomes that accounts for inter-eye and inter-pathology correlations. We present an analysis of data from a study on diabetic retinopathy to illustrate the method.

Les données binoculaires sont fréquentes en ophtalmologie, où des paires de yeux sont examinés par un certain test de diagnostic pour détecter la présence ou l'absence de certaines anomalies ou pathologies. Traiter les yeux de manière indépendante et adopter une approche habituelle pour estimer la sensibilité et la spécificité du test ne tient pas compte de la corrélation entre les yeux, et par conséquent, peut donner de mauvaises estimations. Nous proposons une méthode basée sur la vraisemblance pour estimer la sensibilité et la spécificité en utilisant un modèle paramétrique pour des résultats binaires appariés qui tient compte de la corrélation inter-œil et inter-pathologie. Nous présentons une analyse des données d'une étude sur rétinopathie diabétique pour illustrer la méthode.

**Session 7A**      **Tuesday, June 1 • Mardi 1<sup>er</sup> juin, 13:30 – 15:00**      **AA-1140**

Isobel Loutit Invited Address on Business and Industrial Statistics  
 Allocution Isobel-Loutit en statistique industrielle et en gestion

[MS-108]

**Graph Mining****Exploitation de graphique**Daryl PREGIBON, *Google, Inc., New York*, Corinna CORTES & Chris VOLINSKY, *Google, Inc. & AT&T Labs*

Transactional data that occurs in telecommunications, financial, and retail applications can be represented as a graph. The size of such graphs can be very large so that mining such data poses significant technical challenges. We discuss our experience in mining large graphs paying special attention to the dynamic nature of the underlying applications, namely that the data presents itself not as a static data set but rather as a continuous data stream. We introduce a definition of a dynamic graph that has served us well in representing telecommunications data. We illustrate the ideas with examples from toll fraud detection.

Les données transactionnelles que l'on retrouve dans les télécommunications, la finance et dans les applications dans le commerce de détail peuvent être représentées par un graphe. La taille de tels graphes peut être très grande de sorte que l'analyse de telles données pose des défis techniques significatifs. Nous discutons de notre expérience dans l'exploitation de grands graphes en prêtant une attention particulière à la nature dynamique des applications sous jacentes, à savoir que les données ne se présentent pas comme un jeu de données statique mais plutôt comme un flux continu de données. Nous introduisons une définition d'un graphe dynamique qui nous a bien servi pour représenter des données de télécommunications. Nous illustrons les idées par des exemples de détection de fraude de péage.

**Session 7B**      **Tuesday, June 1 • Mardi 1<sup>er</sup> juin, 13:30 – 15:00**      **Z-310**

Shape Analysis of 3D Brain Mapping Data  
 Analyse de forme de données de cartographie cérébrale 3D

[MS-109]

**Cortical Thickness from MRI****Épaisseur du cortex à partir d'images par résonance magnétique**Jason LERCH, Alex ZIJDENBOS & Alan C. EVANS, *McGill University*

The inner and outer surfaces of the human cerebral cortex are extracted using deformable models. The thickness of the cortex is measured using a distance metric between the two polyhedra at every point on the surface. Linear statistical models are computed at each vertex of the surface to quantify differences in cortical anatomy across groups or against neurological/psychological variables. The results are viewed on an average surface model and significance defined by using Random Field Theory or False Discovery Rate to control for multiple comparisons. Results from a series of normal and pathological populations will be presented.

Les surfaces internes et externes du cortex cérébral humain sont extraites en utilisant des modèles déformables. L'épaisseur du cortex est mesurée en utilisant une métrique de distance entre les deux polyèdres à chaque point sur la surface. Des modèles statistiques linéaires sont calculés à chaque sommet de la surface pour mesurer les différences dans l'anatomie corticale à travers des groupes ou par rapport à des variables neurologiques ou psychologiques. Les résultats sont observés sur un modèle de surface moyenne et la signification statistique est définie en utilisant la théorie des champs aléatoires ou le taux de faux positifs pour contrôler pour les comparaisons multiples. Des résultats d'une série de populations normales et pathologiques seront présentés.



**[MS-110]****A Structural Strategy for Morphometry of the Cerebral Cortex  
Une stratégie structurelle pour la morphométrie du cortex cérébral**

Jean-François MANGIN, D. RIVIÈRE & A. CACHIA, *Service Hospitalier Frédéric Joliot, Orsay*, D. L. COLLINS & Alan C. EVANS, *McGill University*, J. REGIS, *CHU La Timone, Marseille*

Advances in neuroimaging have led to an increasing recognition that certain neuroanatomical structures may be preferentially modified by particular cognitive skills or diseases. Most of the approaches dedicated to automatic morphometry rely on a point-by-point strategy based on warping each brain toward a reference coordinate system. During this talk, I will describe an alternative object-based strategy dedicated to the cortex. This strategy relies on an artificial neuroanatomist performing automatic recognition of the main cortical sulci and parcellation of the cortical surface into gyral patches. This pattern recognition system is designed in the Markovian Random Fields framework.

Les progrès de la neuroimagerie ont montré que certaines structures neuroanatomiques peuvent être modifiées par des talents cognitifs particuliers ou par certaines pathologies. La plupart des approches dédiées à une morphométrie automatique sont fondées sur une stratégie point par point qui consiste à placer les cerveaux dans un système de coordonnées de référence. Durant cet exposé, je présenterai une stratégie alternative qui consiste à décomposer le cortex en entités élémentaires. Cette stratégie repose sur un neuroanatomiste artificiel qui reconnaît automatiquement les principaux sillons et parcellise la surface corticale en gyri. Ce système de reconnaissance des formes est développé dans le cadre des champs de Markov.

**[MS-111]****Mathematical/Computational Challenges in Brain Mapping  
Défis mathématiques et informatiques dans la cartographie du cerveau**

Paul THOMPSON & Arthur W. TOGA, *University of California, Los Angeles*

Analyzing large image databases, powerful mathematical algorithms can now detect disease-specific patterns of brain structure and function. We describe statistical atlases that measure how the brain varies across age and gender, across time, in health and disease, and in large human populations ( $N > 1000$  MRIs). Modeling brain structures as 3D curves and surfaces, we define flows, metrics, covariant PDEs, and statistical fields on these manifolds, encoding anatomical differences across subjects or groups. We use this reference information to detect and visualize brain changes over time in Alzheimer's disease, development and schizophrenia, patterns of medication response, and genetic influences on brain structure.

Lorsque nous analysons de grandes bases de données d'images, les algorithmes mathématiques puissants peuvent maintenant détecter des types de structures et de fonctions du cerveau spécifiques à certaines maladies. Nous décrivons les atlas statistiques qui mesurent comment le cerveau change à travers l'âge et le sexe, à travers le temps, tant dans des cas en santé que malades et ce dans de grandes populations humaines ( $N > 1000$  IRM). En modélisant les structures du cerveau avec des courbes en 3D et des surfaces, nous définissons des flux, des métriques, des EDP covariants et des champs statistiques sur ces variétés, codant des différences anatomiques à travers les sujets ou les groupes. Nous utilisons cette information de référence pour détecter et visualiser les changements dans le cerveau avec le temps pour la maladie d'Alzheimer, pour le développement et la schizophrénie, pour les types de réponse aux médicaments et pour l'influence génétique sur la structure du cerveau.

**Session 7C****Tuesday, June 1 • Mardi 1<sup>er</sup> juin, 13:30 – 15:00****Z-330**

Multi-level Modelling

Modélisation multi-niveau

**[MS-112]****Multilevel Multiprocess Modelling of Partnership Transitions and Fertility in Britain  
Modélisation multi-états multi-niveaux de la transition d'union et de la fertilité en Angleterre**Fiona STEELE, Constantinos KALLIS, Harvey GOLDSTEIN & Heather JOSHI, *University of London*

We describe a general framework for the analysis of correlated event histories, motivated by a study of partnership transitions and fertility among British women. A multilevel multistate competing risks model is used to analyze transitions from marriage and unmarried cohabitation. These partnership transitions are modeled jointly with fertility using a simultaneous equation model. Modeling the two processes simultaneously allows for the possibility that partnership transitions may depend on outcomes of the fertility process and partnership stability (as measured by the hazard of a partnership transition) might affect fertility decisions; we also allow for unobserved woman-level characteristics that affect both processes.

Nous décrivons un cadre général pour l'analyse de l'histoire d'événements corrélées, motivé par une étude sur la transition d'union et la fertilité des femmes britanniques. Un modèle de concurrence de risque multi-états multi-niveaux est utilisé pour analyser les transitions à partir du mariage et de l'union libre. Ces transitions d'union sont modélisées conjointement avec la fertilité en utilisant un modèle d'équations simultanées. La modélisation des deux processus tient compte simultanément de la possibilité que les transitions d'union peuvent dépendre des résultats du processus de fertilité et que la stabilité de l'union (mesurée par le risque d'une transition d'union) pourrait affecter des décisions de fertilité; nous tenons compte également de caractéristiques non observées au niveau des femmes qui affectent les deux processus.

**[MS-113]****Some Challenging Issues in the Analysis of the National Longitudinal Survey of Children and Youth  
Quelques problèmes intéressants dans l'analyse du sondage longitudinal des enfants et de la jeunesse**Renjun MA, *University of New Brunswick*

Recently, we have studied obesity, self-esteem, literacy, smoking, alcohol and drug use among Canadian children and adolescents based on the National Longitudinal Survey of Children and Youth (NLSCY). These issues often involve analysis of spatially and temporally correlated categorical data. In this talk, we will discuss formulation of certain policy related issues into appropriate statistical questions. Some open problems might also be discussed.

Récemment, nous avons étudié l'obésité, l'estime de soi, le niveau d'instruction, le tabagisme, l'utilisation d'alcool et de drogues parmi les enfants et les adolescents canadiens basé sur le sondage longitudinal national des enfants et de la jeunesse (NLSCY). Ces problèmes impliquent souvent l'analyse de données catégoriques temporellement et spatialement corrélées. Dans cette présentation, nous discuterons de la formulation des problèmes de politique publique en questions statistiques appropriées. Quelques problèmes non résolus pourraient également être discutés.

**[MS-114]****Accounting for Survey Design in Multilevel Models: An Empirical Study  
Tenir compte du plan d'échantillonnage dans les modèles multi-niveaux : une étude empirique**James BROWN & Nyovani MADISE, *University of Southampton*, David STEEL, *University of Wollongong, Australia*

Secondary data analysts are now aware of the impact a complex survey design can have when fitting linear models and computer packages now allow them to 'control' for the impact of weighting, clustering, and stratification. A problem occurs when the impact of clustering is of substan-

Les analystes de données se rendent maintenant compte de l'impact que peut avoir un plan d'échantillonnage complexe lorsqu'ils tentent d'ajuster des modèles linéaires et les logiciels informatiques leur permettent maintenant de « contrôler » pour l'impact de la pondération, du regroupement et de la stratification. Un problème se produit quand l'impact du

tive interest and the preferred method of analysis is multilevel modeling. Pfeiffermann et al. (1998) proposes an approach to probability weighting with iterative generalized least squares. The paper investigates this approach using a simulation study and compares the performance with a model-based approach to account for the design, and a standard weighting approach in STATA.

regroupement est d'un intérêt substantiel et que la méthode d'analyse voulue est la modélisation multi-niveaux. Pfeiffermann et al. (1998) ont proposé une approche à la pondération de la probabilité avec des moindres carrés généralisés itératifs. L'article étudie cette approche par une étude de simulation et compare la performance par une approche modéliste pour tenir compte du plan et également avec une approche pondérée standard dans STATA.

## Session 7D      Tuesday, June 1 • Mardi 1<sup>er</sup> juin, 13:30 – 15:00      AA-1360

### Nonlinear Time Series Chroniques non linéaires

[MS-115]

#### Dynamic Integration of Time- and State-domain Methods for Volatility Estimation

#### Méthode d'intégration dynamique des domaines du temps et de l'état pour l'estimation de la volatilité

Jianqing FAN, *Princeton University*, Jiancheng JIANG, *Peking University*

This paper address the issues of dynamically integrating the information from time- and state-domain for volatility prediction. We first approach the problem by proposing two Bayesian methods, and then propose to estimate the volatility via dynamically integrating information from both the time and the state domains. The estimators from both domains are optimally combined based on a data driven weighting strategy. Extensive simulations are conducted to demonstrate that the newly proposed procedure outperforms some popular ones such as the RiskMetrics and the historical simulation approaches, among others. Empirical studies endorse convincingly our integration method.

Cette présentation traite des questions d'intégration dynamique de l'information du temps et de l'état pour la prévision de la volatilité. Tout d'abord, nous approchons le problème en proposant deux méthodes bayésiennes, nous proposons ensuite d'estimer la volatilité par l'intermédiaire de l'intégration dynamique de l'information des domaines de temps et de l'état. Les estimateurs des deux domaines sont combinés de façon optimale basés sur une stratégie pondérée obtenue à l'aide des données. Des simulations intensives sont effectuées pour démontrer que la procédure proposée surpasse certaines méthode populaires tels que RiskMetrics et les approches de simulations historiques entre autres. Les études empiriques approuvent d'une manière convaincante notre méthode d'intégration.

[MS-116]

#### Statistical Process Control for Nonlinear Time Series

#### Processus de contrôle statistique pour des séries chronologiques non linéaires

Wolfgang SCHMID & Stefan SCHIPPER, *Europa-Universität Viadrina, Frankfurt*

Non-linear time series are widely applied in finance because they explain the time-varying behavior of an asset. The risk is measured by the conditional variance (volatility). An investor is highly interested to detect changes in the volatility as soon as possible. In this talk we consider sequential procedures for the surveillance of the risk process. Several control charts for the variance of a univariate GARCH process are proposed. Based on the av-

Les séries chronologiques non linéaires sont largement utilisées en finance parce qu'elles expliquent le comportement variable dans le temps d'un titre. Le risque est mesuré par la variance conditionnelle (la volatilité). Un investisseur est fortement intéressé de détecter des changements dans la volatilité aussitôt que possible. Dans cette présentation, nous considérons des procédures séquentielles pour la surveillance du processus de risque. Plusieurs diagrammes de contrôle pour la variance d'un processus GARCH uni-

erage run length these charts are compared with each other. The problem of monitoring the covariance matrix of a multivariate GARCH process is discussed as well.

varié sont proposés. Basé sur le temps de course moyen, ces diagrammes sont comparés entre eux. Le problème de surveillance de la matrice de covariance d'un processus GARCH multivarié est également discuté.

**[MS-117]**

**A Bootstrap Test of Conditional Heteroscedasticity in Nonlinear Time Series**

**Un test bootstrap pour l'hétéroscédasticité pour des séries chronologiques non linéaires**

Min CHEN, *Chinese Academy of Sciences, Beijing/Académie des sciences de Chine, Beijing*, Gemai CHEN, *University of Calgary*

In this talk we propose an efficient bootstrap test of conditional heteroscedasticity in nonlinear autoregressive time series. The test statistics are of the Kolmogorov–Smirnov and Cramér–von Mises type. An asymptotic distribution theory is developed, and the power of our proposed test is studied through simulation for finite samples and under normal and nonnormal distributions. Data Analysis using real data is also illustrated.

Dans cette présentation, nous proposons un test bootstrap efficace pour l'hétéroscédasticité conditionnelle dans les séries chronologiques auto-régressives non linéaires. Les statistiques des tests sont du type Kolmogorov-Smirnov et Cramer-von Mises. La distribution asymptotique est obtenue et la puissance du test proposé est étudiée par des simulations pour les échantillons finis et sous des distributions normales et non normales. Une analyse de données à partir de vraies données est également présentée.

**Session 7E**

**Tuesday, June 1 • Mardi 1<sup>er</sup> juin, 13:30 – 15:00**

**AA-1355**

**Survey Sampling II**

**Méthodologie d'enquête II**

**[MS-118]**

**Collection follow-up score function and response bias**

**Fonction de score de suivi de collecte et biais de réponse**

Hansheng XIE & Sanping CHEN, *Statistics Canada/Statistique Canada*

Collection follow-up score functions based on weighted revenue are currently used in many business surveys. The objective is to maximize the coverage of the survey in terms of the primary variable of interest—total revenue, under the constraint of a limited follow-up budget. Such a strategy inevitably introduces a response bias toward high-revenue units in the sample, whose effect is amplified through donor imputation during data editing. This study examines various approaches of alleviating and correcting the response bias problem in the use of collection follow-up score function. These approaches will be illustrated and compared using simulated processes.

Des fonctions de score de suivi de collecte basées sur le revenu pondéré sont actuellement utilisées dans plusieurs sondages sur les entreprises. L'objectif est de maximiser la couverture du sondage en terme de la variable d'intérêt principal — revenu total, sous la contrainte d'un budget limité de suivi. Une telle stratégie présente inévitablement un biais de réponse vers les unités à revenus élevés dans l'échantillon, dont l'effet est amplifié par l'imputation de donneurs durant l'édition des données. Cette étude examine diverses approches pour alléger et corriger le problème de biais de réponse dans l'utilisation de la fonction de score de suivi de collecte. Ces approches seront illustrées et comparées en utilisant des processus simulés.

**[MS-119]****Importance of a good imputation method for the education questions in the LFS****Importance d'une bonne méthode d'imputation aux questions d'éducation dans le cadre de l'EPA**

Caroline PELLETIER, *Statistics Canada/Statistique Canada*

The Labour Force Survey (LFS) like any investigation must face the treatment of non response. The study deals with the non response for the questions on education used to derive the highest level of education reached by a person. It is an auxiliary information used in the construction of the imputation classes of the hot-deck system developed for the principal component of the investigation. The method suggested will improve quality of the derived variable and will contribute to improve the imputation of missing data in the principal component while treating bias remaining in the data currently attributed manually by experienced coders.

L'Enquête sur la population active (EPA), comme toute enquête, doit faire face au traitement de la non-réponse. L'étude porte sur le traitement de la non-réponse pour les questions d'éducation servant à dériver le plus haut niveau d'éducation atteint par une personne. Il s'agit d'une information auxiliaire utilisée dans la formation des classes d'imputation du système hot-deck développé pour la composante principale de l'Enquête. La méthode proposée améliorera la qualité de la variable dérivée, contribuera à améliorer l'imputation de données manquantes dans la composante principale tout en remédiant au biais subsistant dans les données actuellement imputées manuellement par des codeurs expérimentés.

**[MS-120]****Use of Randomized Response Techniques in Dealing with Social Desirability Response Bias****L'utilisation de techniques de randomisation des réponses pour traiter le biais dans les réponses dû à la désirabilité sociale**

Sat GUPTA, Bill THORNTON & Sunil SINGHAL, *University of Southern Maine*, Javid SHABBIR, *Quad-I-Azam University, Islamabad*

Randomized response techniques have been presented as a means to overcome social desirability response bias on sensitive questions in personal surveys involving health-related attitudes and behaviors (alcohol/drug use, sexual practices). We discuss the relative effectiveness of a partial and a full randomized response technique. Compared to benchmark results from an anonymous survey, and results using a "bogus pipeline technique," the partial randomization technique proved more effective for sensitive items and no difference was apparent on nonsensitive items. In addition to SDB, a potential gender response bias was apparent on certain items (e.g., number of sexual partners, physical abuse).

Des techniques de réponse randomisées ont été présentées comme un moyen de surmonter le biais dans les réponses dû à la désirabilité sociale pour des questions sensibles dans des sondages personnels portant sur certaines attitudes et comportements relatifs à la santé (utilisation d'alcool, de drogue, pratiques sexuelles). Nous discutons de l'efficacité relative de techniques partiellement et complètement randomisées. Comparativement aux résultats de référence d'un sondage anonyme, et à ceux en utilisant une « technique de fausse direction », la technique de randomisation partielle s'est montrée plus efficace pour des items sensibles et montre aucune différence évidente pour les items non sensibles. En plus du BDS, un biais potentiel dans les réponses dû au sexe était apparent sur certains items (par exemple, nombre de partenaires sexuels, abus physique).

**[MS-121]****Graphical Analysis of Complex Survey Data****Analyse graphique de données provenant d'enquêtes à plan complexe**

Martin PANTEL & Yves LAFORTUNE, *Statistics Canada/Statistique Canada*

Graphical representation has long been recognized as an essential tool for exploratory data analysis in many fields of statistics. This tool is of-

L'analyse exploratoire de données à l'aide de représentation graphique est bien implantée dans plusieurs domaines de la statistique. Son utilisation est parfois négligée dans

ten neglected when applied to surveys with complex designs due to complications brought about by the large number of observations and to differing number of units represented by each observation. This presentation will provide an overview of some graphical methods that can be used despite these complications. Some modifications to well-known methods, to further adapt them to survey data, will also be proposed.

le cadre d'enquêtes à plan complexe à cause des difficultés apportées par le nombre élevé de données et, dans le cas des sondages, au nombre différent d'éléments représentés par chaque observation. La présente discussion donnera un aperçu de quelques méthodes permettant d'utiliser les méthodes graphiques populaires malgré ces difficultés. Quelques modifications à des méthodes bien connues seront également proposées afin de les rendre mieux adaptées au monde des enquêtes.

**Session 7F**                      **Tuesday, June 1 • Mardi 1<sup>er</sup> juin, 13:30 – 15:00**                      **Z-317**

**Environmental Statistics**  
**Statistique environnementale**

[MS-122]

**Does Socio-Demographic Status Influence the Effect of Ambient Air Pollution on Hospitalisation for Respiratory Conditions? Results from a Time Series Study in 10 Canadian Cities**

**Est-ce que le statut socio-démographique influence l'effet de la pollution atmosphérique ambiante sur l'hospitalisation pour des problèmes respiratoires ? Résultats d'une étude de séries chronologiques dans dix villes canadiennes**

Sabit ÇAKMAK, *Health Canada/Santé Canada*, Robert DALES, *University of Ottawa/Université d'Ottawa*

Social status influences respiratory morbidity but the mechanisms are not well understood. To determine if sociodemographics influence the susceptibility to ambient air pollution, we determined the association between daily respiratory hospitalizations and daily concentrations of ambient air pollution in ten large Canadian cities. Daily time-series analyses were employed and results were adjusted for climate variables. Results were then stratified by age, gender, and neighbourhood family education and income. Our results suggest that older males and those within less educated and lower income families may be more vulnerable to air pollutants as reflected by hospitalization for respiratory hospitalization.

Le statut social influence la mortalité due à des problèmes respiratoires mais les mécanismes ne sont pas bien compris. Pour déterminer si les conditions socio-démographiques influencent la susceptibilité à la pollution atmosphérique ambiante, nous déterminons l'association entre les hospitalisations quotidiennes dues à des problèmes respiratoires et les concentrations quotidiennes de pollution atmosphérique ambiante dans dix grandes villes canadiennes. Des analyses quotidiennes de séries chronologiques sont utilisées et des résultats ont été ajustés pour les variables climatiques. Les résultats sont ensuite stratifiés par âge, sexe et niveau d'éducation et revenu familial du voisinage. Nos résultats suggèrent que les hommes plus âgés et ceux venant de familles moins instruites et avec un niveau de revenu inférieur peuvent être plus vulnérables aux polluants atmosphériques comme le montre les hospitalisations pour problèmes respiratoires.

[MS-123]

**Evaluating the Performance of Regional Photochemical Ozone Modeling System**

**Évaluation de la performance de systèmes photochimiques de modélisation régionale de l'ozone**

Zhong LIU, Dou YIPING, Howard CHANG, Nhu LE & Jim ZIDEK, *University of British Columbia*

There are many physical models simulating ozone concentration levels with input of observed meteorological data. One interesting question is how

Il y a beaucoup de modèles physiques qui simulent les niveaux de concentration en ozone à partir de données météorologiques observées. Une question intéressante est jus-

well these simulated data predict the true levels. In this paper, we assess the performance of the AQM (Air Quality Model) model. First, we will review ideas and techniques for evaluation suggested by researchers in both atmospheric sciences and statistics. Then these techniques are applied in the case study.

The data used in the paper is from 1995 EPA's AIRS (Aerometric Information Retrieval System) database. Results will be discussed.

qu'à quel point ces données simulées prédisent-elles les vrais niveaux. Dans cette présentation, nous évaluons la performance du modèle AQM (Air Quality Model). Premièrement, nous passerons en revue les idées et les techniques d'évaluation suggérées par des chercheurs en science atmosphérique et en statistique. Ensuite ces techniques sont appliquées à une étude de cas. Les données utilisées proviennent de la base de données AIRS (Aerometric Information Retrieval System) de l'EPA en 1995. Nous discuterons également des résultats.

[MS-124]

**Logistic Analysis of Space-Time Data Investigating Potential Climatic Risk Factors of Waterborne Enteric Illness**

**Analyse logistique de données espace-temps pour investiguer des facteurs de risque potentiels pour les maladies entériques d'origine hydrique**

Dan GILLIS, *University of Guelph*

This talk will outline a statistical investigation of potential climatic, agricultural and water source and treatment risk factors for waterborne enteric disease. The study will focus on a case control hospitalization data set from 1992 through 1998, including various climatic (precipitation, temperature), agricultural (livestock density, soil treatment) and water source/treatment variables. Discussion will focus on methodologies used, problems encountered, and also summarize findings to date. Part of the talk will focus on different methods for interpolating the space-time weather data.

Cette présentation décrira une recherche statistique sur des facteurs de risque potentiels climatiques, agricoles, de la source et du traitement de l'eau pour les maladies entériques d'origine hydrique. L'étude se concentrera sur un jeu de données cas-témoin d'hospitalisations de 1992 à 1998, incluant diverses variables climatiques (précipitation, température), agricoles (densité des stocks vivants, traitement du sol) et sur les sources et le traitement de l'eau. La discussion se concentrera sur les méthodologies utilisées et les problèmes encourus et nous résumerons les résultats obtenus jusqu'à maintenant. Une partie de la présentation se concentrera sur différentes méthodes pour interpoler les données météorologiques espace-temps.

[MS-125]

**Time Series Analyses for the Seasonality of Foodborne Diseases and Temperature**

**Analyses de séries chronologiques pour la saisonnalité des maladies d'origine alimentaire et la température**

Manon FLEURY, *Health Canada/Santé Canada*, John HOLT, *University of Guelph*, Dominique CHARRON, *Health Canada/Santé Canada*, Brian ALLEN, *University of Guelph*, Abdel MAAROUF, *Environment Canada/Environnement Canada*

In 2002, generalized additive models, the most common method used for time series analyses were shown to over-estimate standard errors of the relative risks. This research paper explores different types of time series analytical approaches using both generalized additive models and generalized linear models with regression splines. Both methods are becoming widely used for environmental time series because they permit flexible adjustments for nonlinear confounding effects of time

En 2002, il a été montré que les modèles additifs généralisés, la méthode la plus utilisée pour des analyses de séries chronologiques, surestime l'écart type du risque relatif. Cette recherche explore différents types d'approches analytiques de séries chronologiques en utilisant des modèles additifs généralisés et des modèles linéaires généralisés avec des splines comme régresseurs. Les deux méthodes deviennent de plus en plus utilisées pour l'étude de séries chronologiques environnementales parce qu'elles permettent un ajustement flexible pour des effets non linéaires parasites de ten-

trends, seasonality and weather variables. The data analyzed in this paper looks at the potential shift in seasonal patterns of foodborne illness and temperature in Canada.

dances temporelles, pour la saisonnalité et pour les variables climatiques. Les données analysées dans cette recherche permettent d'étudier la variation potentielle dans les tendances saisonnières des maladies d'origine alimentaire et de la température au Canada.

[MS-126]

**Efficiency of Study Designs to Detect Interaction Effects Between Genetic and Environmental Factors on an Additive Risk Scale**

**Efficacité de plans d'échantillonnage pour détecter un effet d'interaction entre des facteurs génétiques et environnementaux sur une échelle additive**

Alexandre BUREAU, *University of Lethbridge*

Understanding the role that interactions between multiple genes and between genes and environmental exposures play in complex diseases is a current challenge. I consider the efficiency of prospective studies and case-control studies nested within prospective cohorts to estimate interaction on an additive risk scale, a choice with a biological basis. In both cases, the sample can be stratified on covariates under study. Under a range of scenarios for the relationships between genes, environmental exposures and disease status, a nested case-control sample frequency-matched on a risk factor provides high efficiency to estimate interaction between that risk factor and genetic variants.

Comprendre le rôle que les interactions entre de multiples gènes et entre gènes et expositions environnementales jouent dans les maladies complexes est un défi actuel. Je considère l'efficacité d'études prospectives et cas-témoin imbriquées à l'intérieur de cohortes pour estimer l'interaction sur une échelle additive, un choix biologiquement justifiable. Dans les deux cas, l'échantillon peut être stratifié par des covariables à l'étude. Sous une gamme de scénarios pour les relations entre gènes, expositions environnementales et statut cas/témoin, un échantillon cas-témoin imbriqué avec égales fréquences d'un facteur de risque offre une efficacité élevée pour estimer l'interaction entre ce facteur de risque et des variants génétiques.

**Session 8A**

**Tuesday, June 1 • Mardi 1<sup>er</sup> juin, 15:30 – 17:00**

**Z-310**

**The Canadian Initiative in Social Statistics and Cross-disciplinary Collaboration**

**L'initiative canadienne en matière de statistiques sociales et la collaboration interdisciplinaire**

[MS-127]

**Pushing the Limits: Using Statistics with Varying Amounts of Expertise**

**Pousser les limites : utiliser la statistique avec des niveaux d'expertise variables**

Michael SMITH & Steven RYTINA, *McGill University*

Social scientists using quantitative data have to make practical decisions with respect to procedures. They are often confronted with different models of appropriate practice, where what constitutes the best single practice is not always apparent. Examples would be: the adjustment of standard errors estimated from data collected using clustering and stratification; responses to multicollinearity; dealing with inequality measures (e.g., the Gini, Theil, and Atkinson coefficients); the ap-

Les sociologues qui utilisent des données quantitatives doivent prendre des décisions pratiques par rapport à des procédures. Ils sont souvent confrontés à différents modèles de la pratique appropriée, où ce qui constitue la meilleure pratique n'est pas toujours évident. Les exemples seraient : l'ajustement des erreurs types estimées à partir des données collectées à partir de plans avec grappes et stratification, les réponses à la multicollinéarité, traiter des mesures d'inégalité (par exemple les coefficients de Gini, de Theil, et d'Atkinson), l'utilisation appropriée des procédures d'inférence lors



propriate use of inferential procedures when using Census data or historical series; and the desirability of Bayesian procedures. We use several of these issues to illustrate the dilemmas that social scientists confront.

de l'utilisation de données de recensement ou de séries historiques et la désirabilité des procédures bayésiennes. Nous utilisons plusieurs de ces problèmes pour illustrer les dilemmes que les sociologues confrontent.

**[MS-128]**

**Sharing Statistical and Social Science Methodologies: Mixed Blessings  
Échange de méthodologies statistiques et de sciences sociales : approbation mixte**

Patricia NEWCOMBE-WELCH, *University of Waterloo*

The Research Data Centre (RDC) program is a result of the Canadian Initiative on Social Statistics which was intended to increase the quality of research which could be conducted by Canadian Social Scientists. In addition to providing Statistics Canada master data file access to researchers with approved proposals, the RDC program also provides a setting in which research initiatives could be conducted jointly between social scientists and statisticians, the latter being interested in methodological developments and improvements to the statistical procedures used by the former. This talk addresses potential areas for such collaboration, and also some of the associated impediments.

Le programme de Centres de données de recherche (CDR) est le résultat de l'Initiative canadienne en statistiques sociales dont l'objectif était d'augmenter la qualité des recherches qui peuvent être conduites par les chercheurs canadiens de sciences humaines. En plus de fournir un accès au fichier de données principal de Statistique Canada aux chercheurs dont la proposition est approuvée, le programme du CDR fournit également un environnement dans lequel les initiatives de recherches peuvent être conduites conjointement entre humanistes et statisticiens, ces derniers étant intéressés par les développements méthodologiques et les améliorations aux procédures statistiques utilisées par les premiers. Cette présentation aborde des sujets propices à une telle collaboration et également certains des irritants y étant associés.

**[MS-129]**

**Growth Trajectories in Childhood and Adolescent Overweight and Obesity in Canada  
Courbes de croissance de l'embonpoint et de l'obésité chez les enfants et les adolescents au Canada**

Renjun MA & Jon Douglas WILLMS, *University of New Brunswick*

There is growing concern over the overweight and obesity problem among children and adolescents worldwide. Recent cross-sectional studies have shown a progressive increase in the prevalence of overweight and obesity among Canadian children and adolescents. In our current study, we investigate the growth trajectories in childhood and adolescent obesity in Canada based on Canada's National Longitudinal Survey of Children and Youth (NLSCY). The use of longitudinal growth models have the potential to discern when children are most vulnerable to becoming overweight and obese, and assess the strength of both time-varying and time-invariant factors that contribute to this problem.

Il y a une préoccupation grandissante par rapport aux problèmes de l'embonpoint et d'obésité chez les enfants et les adolescents partout dans le monde. Des études transversales récentes ont montré une augmentation progressive de la prévalence de l'embonpoint et de l'obésité chez les enfants et les adolescents canadiens. Dans la présente étude, nous étudions la courbe de croissance de l'obésité à l'enfance et à l'adolescence au Canada basé sur l'Enquête longitudinale nationale sur les enfants et les jeunes du Canada (ELNEJ). L'utilisation des modèles de croissance longitudinaux ont le potentiel de distinguer les moments où les enfants sont les plus vulnérables à avoir un excès de poids ou à devenir obèses et d'évaluer la force des facteurs qui varient ou non dans le temps et qui contribuent à ce problème.

**Session 8B                      Tuesday, June 1 • Mardi 1<sup>er</sup> juin, 15:30 – 17:00                      AA-1140****Statistical Analysis of Human Growth Data****Analyse statistique de données sur la croissance humaine****[MS-130]****A Web-based System For Monitoring the Growth of Children****Un système basé sur Internet pour suivre la croissance des enfants**Michael HERMANUSSEN, *Aschauhof, Altenhof, Germany*, Christof MEIGEN, *Universität Leipzig*

National child preventive care programs are expensive; an alternative web-based surveillance program ([www.willi-will-wachsen.de](http://www.willi-will-wachsen.de)) is shown. Internet is an ideal tool for an interactive health care system. Accessibility of the web has rapidly increased (43% of private German households in 2002, particularly adolescents, and persons with children). Height, weight and various data of a child's growth can easily be provided by parents or non-medical personnel and then analyzed online. Interactive web-based monitoring programs are simple and cheap clinical decision support systems for augmenting the efficacy of traditional preventive care. Future programs may include child-web interactions, e.g., by evaluating virtual play games.

Les programmes nationaux de soins préventifs aux enfants sont dispendieux ; un programme de surveillance alternatif basé sur Internet ([www.willi-will-wachsen.de](http://www.willi-will-wachsen.de)) est présenté. L'Internet est un outil idéal pour un système interactif de soins de santé. L'accessibilité d'Internet a rapidement augmenté (43% des ménages privés allemands en 2002, en particulier les adolescents et les personnes avec des enfants). La taille, le poids et les diverses données de la croissance d'un enfant peuvent facilement être fournis par les parents ou le personnel non médical et alors être analysés en ligne. Les programmes de contrôle interactifs basés sur Internet sont des systèmes d'aide aux décisions cliniques simples et bon marché pour augmenter l'efficacité des soins préventifs traditionnels. Les programmes futurs pourront inclure des interactions entre Internet et l'enfant, par exemple en évaluant les jeux virtuels.

**[MS-131]****“Transmuting” Women Into Men: Galton’s Family Data on Human Stature****Transformer les femmes en hommes : Les données de Galton sur la taille**James HANLEY, *McGill University*

The first two regression lines, and the first correlations, were calculated by Francis Galton, in his work on heredity in sweet-peas and in humans. When (what would today be termed) ‘regressing’ the heights of adult children on those of their parents, Galton had to deal with the (statistical) complication that men are generally taller than women—but without modern-day statistical tools such as multiple regression and partial correlation. This talk presents Galton's original family data on stature. These data are used to assess how sharp, compared with modern-day tools, his methods were for dealing with this complication.

Les deux premières droites de régression, et les premières corrélations ont été calculées par Francis Galton dans son travail sur l'hérédité des pois sucrés et chez l'homme. Quand Galton a (dans ce qui aujourd'hui se nommerait) « régressé » la taille des enfants adultes sur celle de leurs parents, il a dû traiter la complication (statistique) que les hommes sont généralement plus grands que les femmes, mais sans outils statistiques modernes tels que la régression multiple et la corrélation partielle. Cette présentation présente les données originales de Galton sur la taille. Ces données sont utilisées pour évaluer la précision impressionnante de ses méthodes, comparativement aux outils modernes, pour traiter de cette complication.

[MS-132]

**Human Growth Data: Multiresolution Challenges and Phase/Amplitude Variation****Données de croissance humaine : défis multi-résolutions et variation de phase et d'amplitude**James RAMSAY, *McGill University*

Recent evidence, especially by Michael Hermanussen, points to growth features on several time scales: namely hours, days, weeks and years. This multiresolution aspect of growth raises many new questions about the regulation of growth, and calls for new statistical tools more oriented to the modeling of change.

Much of the interpersonal variation in growth is in the timing of events such as the pubertal growth spurt, called phase variation. Charts used by pediatricians assume the growth of children varies only in terms of amplitude, and consequently they offer a distorted image of what constitutes “normal” growth.

L'évidence récente, particulièrement celle provenant de Michael Hermanussen, suggère que les caractéristiques de la croissance évoluent sur plusieurs échelles de temps : à savoir heures, jours, semaines et années. Cet aspect multi-résolution de la croissance soulève beaucoup de nouvelles questions au sujet de la régularisation de la croissance et nécessite de nouveaux outils statistiques davantage orientés sur la modélisation du changement. La grande partie de la variation interpersonnelle de la croissance est dans la synchronisation des événements tels que la croissance pubertaire, appelée la variation de phase. Les diagrammes utilisés par les pédiatres supposent que la croissance des enfants varie seulement en termes d'amplitude, et par conséquent ils offrent une image distorsionnée de ce qui constitue la croissance « normale ».

**Session 8C****Tuesday, June 1 • Mardi 1<sup>er</sup> juin, 15:30 – 17:00****AA-1355****Empirical Likelihood Methodology****Méthodes empirico-vraisemblantistes**

[MS-133]

**Weighted Empirical Likelihood****Vraisemblance empirique pondérée**Jian-Jian REN, *University of Central Florida*

Weighted empirical likelihood is a newly developed likelihood method which can be used to construct tests and confidence intervals for various types of incomplete data, such as right censored data, doubly censored data, interval censored data and partly interval-censored data. In this talk, we give a rather general theorem for constructing Weighted Empirical Likelihood Ratio Confidence Intervals (WELRCI) for  $\theta_0$ , which includes the mean, survival probabilities, median, quantiles, M-statistic, and trimmed mean as special cases. Also, a general theoretical coverage accuracy equation for WELRCI is established under a unified framework for various types of censored data.

La vraisemblance empirique pondérée est une méthode de vraisemblance nouvellement développée qui peut être utilisée pour construire des tests et des intervalles de confiance pour différents types de données incomplètes, telles que des données censurées à droite, doublement censurées, censurées par intervalle et des données partiellement censurées par intervalles. Dans cette présentation, nous présentons un théorème général pour construire des intervalles de confiance du rapport de vraisemblance pondéré (ICRVP) pour  $\theta_0$  qui inclut la moyenne, les probabilités de survie, la médiane, les quantiles, les statistiques-M et la moyenne tronquée comme cas spéciaux. De plus, une équation théorique générale de la précision de couverture ICRVP est établie dans un cadre unifié pour différents types de données censurées.

**[MS-134]****Empirical Likelihood Analysis of a Semiparametric Mixture Model, with Application to QTL Mapping**  
**Analyse de vraisemblance empirique d'un modèle de mélange semi-paramétrique avec des applications au mapping de locus quantitatif**Jason FINE, *University of Wisconsin-Madison*, Fei ZOU, *The University of North Carolina at Chapel Hill*

We propose semiparametric alternatives to traditional parametric mixture models for QTL interval mapping. The log ratio of the component densities satisfies a linear model, with unspecified baseline density. It is demonstrated that constrained empirical likelihood has irregularity under the null hypothesis that the two densities are equal. Partial empirical likelihood gives unconstrained parameter estimators which are consistent and asymptotically normal, regardless of the null. The partial likelihood has interesting connections to conditional likelihood involving nuisance parameters. The partial likelihood estimator is more efficient than the estimator based on known nuisance parameters. The methods are applied to a rat study of breast cancer genes.

Nous proposons des alternatives semi-paramétriques aux modèles de mélange paramétriques traditionnels pour le mapping d'intervalle de LQ. Le rapport du log des densités des composantes satisfait le modèle linéaire avec une densité de base non spécifiée. Il est démontré que la vraisemblance empirique contrainte est irrégulière sous l'hypothèse nulle que les deux densités sont égales. La vraisemblance empirique partielle donne des estimateurs sans contrainte des paramètres qui sont convergents et asymptotiquement normaux, indépendamment de l'hypothèse nulle. L'estimateur de vraisemblance partielle a des liens intéressants avec la vraisemblance conditionnelle impliquant des paramètres de nuisance. L'estimateur de vraisemblance partielle est plus efficace que les estimateurs basés sur des paramètres de nuisance connus. Les méthodes sont appliquées à une étude sur les gènes du cancer du sein sur des rats.

**[MS-135]****Some Finite Sample Results on Empirical Likelihood**  
**Quelques résultats d'échantillons finis portant sur la vraisemblance empirique**Min TSAO, *University of Victoria*

The finite sample coverage properties of empirical likelihood ratio (ELR) confidence regions are in general difficult to obtain since the underlying distribution is unknown and the ELR does not have an analytic expression. In this talk, I will discuss one finite sample coverage property: bounds on the coverage probabilities of the ELR confidence regions based on estimating equations. I will discuss the mixture nature of the finite sample distribution of the ELR, which gives rise to the bounds. I will also discuss implications of the bounds on the empirical likelihood inference.

Les propriétés de couverture pour échantillon fini des régions de confiance du rapport de vraisemblance empirique (RVE) sont en général difficiles à obtenir parce que la distribution sous-jacente est inconnue et parce que le RVE n'a pas d'expression analytique. Dans cette présentation, je discuterai d'une propriété de couverture d'échantillon fini : des bornes sur les probabilités de couverture des régions de confiance du RVE basées sur des équations d'estimation. Je discuterai de la nature de mélange de la distribution d'échantillon fini du RVE qui mène aux bornes. Je discuterai également des implications des bornes sur l'inférence de vraisemblance empirique.

**Session 8D****Tuesday, June 1 • Mardi 1<sup>er</sup> juin, 15:30 – 17:00****Z-330****Teaching Statistics to Non-statisticians: Spreading the Good Word to Achieve Recognition for our Profession****L'enseignement de la statistique à des non-statisticiens : prêcher la bonne parole pour faire reconnaître notre profession**

**[MS-136]****Teaching Statistics to Future Managers: What Should Be the Main Expected Goal?****L'enseignement de la statistique à de futurs gestionnaires : quel doit être le principal objectif à atteindre ?**François BELLAVANCE, *HEC Montréal*

A course in statistics is mandatory in most university programs. This course is often the only one taken by students in their program. It is therefore important to efficiently use this unique opportunity to promote this precious decision making tool and convince future managers of the value added by the expertise and the work of statisticians. Hence, we need to go beyond introducing the basic statistical concepts and show real and relatively complex cases where the work of statisticians made a difference in the decision making process.

Dans la majorité des programmes universitaires diplômants, il y a généralement au moins un cours de statistique obligatoire. Il est aussi dans la plupart des cas le seul cours de statistique pris par les étudiants dans leur programme. Il est donc important de bien utiliser cette vitrine pour promouvoir ce précieux outil d'aide à la décision et convaincre les futurs gestionnaires de la valeur ajoutée de l'expertise et du travail des statisticiens. Ainsi, il faut aller au-delà de la simple introduction des notions de base et donner des exemples concrets et relativement complexes où le travail de statisticiens a fait une différence importante dans la prise de décision.

**[MS-137]****Challenges Regarding the Use of Data Analysis Software in the Teaching of Statistics****Les défis posés par l'utilisation de logiciels d'analyse de données dans l'enseignement de la statistique**Jean HARDY, *Services Conseils Hardy Inc., Québec*

The increasing sophistication of data analysis software systems, including their more friendly user-interfaces, did not completely succeed in smoothing their learning curve. We think that data analysis, as a key component of research methodologies relying on a statistical software system, deserves more than the smallest share in statistical training. This trend is present in the training of both future statisticians and other professionals using statistics. This presentation will review the long-term effects of this situation and present alternative solutions.

La sophistication croissante et l'amélioration de l'interface-utilisateur des principaux logiciels d'analyse de données n'en fait pas pour autant des outils simples d'accès. Or, il nous semble que l'analyse de données, comme facette de la méthodologie de recherche intimement liée à un logiciel statistique, n'occupe qu'une portion congrue de l'enseignement de la statistique. Cette tendance semble manifeste tant dans la formation des futurs statisticiens que dans celle des professionnels faisant usage de la statistique. Quels sont les effets à long terme de cette approche et comment peut-on en contrer les aspects néfastes ?

**[MS-138]****Teaching Statistics in Companies and Research Centers: Replacing the Toolbox by a Better Understanding of Concepts****L'enseignement de la statistique en entreprise et dans les centres de recherche : de la boîte à outils à la maîtrise des concepts**Natalie RODRIGUE, *Creascience Inc., Montréal*

Most people working in private or public research centers attended at least one course focused on statistical methods during their education. With several years of work experience, many consider attending a training session on this topic a useful refresher especially to validate and improve their research methods. Such an audience is clearly attracted toward applied sessions. Many actually

La plupart des personnes qui oeuvrent au sein d'une unité de recherche publique ou privée ont déjà été en contact avec la statistique par le biais d'au moins un cours universitaire. Après plusieurs années de travail, bon nombre jugent souhaitable de mettre à jour leurs connaissances dans ce domaine notamment pour valider et améliorer leurs pratiques de recherche. Clairement, ce type d'audience recherche une forme d'enseignement plus pragmatique. Pour beaucoup, la

consider statistics as a toolbox in which they know many tools but without knowing exactly when and why to use them. This talk will focus on the needs of this audience and the solutions that can be offered.

statistique constitue une boîte à outils dont ils connaissent le contenu sans toutefois toujours posséder le mode d'emploi, c'est-à-dire une compréhension des concepts sous-jacents. Cette présentation proposera une description des besoins de ce type d'utilisateur de la statistique et les solutions qu'il est possible de leur proposer.

## Session 8E

Tuesday, June 1 • Mardi 1<sup>er</sup> juin, 15:30 – 17:00

Z-220

### Survey Sampling III Méthodologie d'enquête III

[MS-139]

#### Calibration Estimators for Quantiles Estimateurs de calage pour les quantiles

Torsten HARMS & Pierre DUCHESNE, *Université de Montréal*

In survey sampling, calibration estimators are of widespread use. However their application is usually limited to the estimation of means or totals. A generalization of the calibration approach will be presented, where the measure of interest as well as the known covariates of the population are all quantiles. An analytic representation of the calibrated weights, as well as the resulting variance estimators and confidence intervals of the resulting estimator will be given. A small Monte Carlo study will compare this new estimator with other known estimators for quantiles that make use of auxiliary information such as difference-, ratio- or model-based approaches.

En échantillonnage, on utilise souvent des estimateurs de calage. Généralement leur application est limitée à l'estimation de moyennes ou de totaux. Nous allons présenter une généralisation de l'estimateur de calage où la mesure d'intérêt et les variables auxiliaires connues sont toutes des quantiles. Une représentation analytique des poids de calage et aussi de la variance de cet estimateur va être présentée. Dans une étude Monte-Carlo, nous allons comparer ce nouvel estimateur avec d'autres estimateurs pour quantiles qui utilisent de l'information auxiliaire.

[MS-140]

#### Estimation of Variance of a Generalized Regression Predictor Estimation de la variance d'un estimateur par la régression généralisé

Raghunath ARNAB, *University of Durban-Westville*

The generalized regression (greg) predictor is used for estimation of a finite population total when the study variable is well related to the auxiliary variable. In this paper we have derived the lower bound of the variances of the estimators of the variances of the greg predictors belonging to a certain class of estimators under a superpopulation model. The proposed optimal estimator can not be used in practice since it involves unknown parameters. Hence, some alternative variance estimators are also proposed. Simulation studies reveal that the proposed alternative estimators fare better than the exiting alternatives proposed by Särndal(1982).

L'estimateur par la régression généralisé (greg) est utilisé pour l'estimation du total d'une population finie quand la variable à l'étude est bien liée à la variable auxiliaire. Dans cette présentation, nous avons dérivé la borne inférieure de la variance des estimateurs de la variance des estimateurs greg appartenant à une certaine classe d'estimateurs sous un modèle de super-population. L'estimateur optimal proposé ne peut pas être utilisé en pratique puisqu'il implique des paramètres inconnus. Par conséquent, nous proposons également quelques estimateurs alternatifs de la variance. Les études de simulation montrent que les estimateurs alternatifs proposés font mieux que les alternatives proposées par Särndal (1982).

**[MS-141]****Replicate Weights and Disclosure Control  
Réplication des poids et contrôle de la divulgation**Wilson LU & Randy R. SITTER, *Simon Fraser University*

Statistical agencies routinely release micro-data for public use with stratum and/or psu indicator suppressed for confidentiality reasons. For the purpose of variance estimation pseudo-psu indicators are sometimes produced for use in linearization methods or replication weights for use in resampling methods. If care is not taken these can be used to (partially) reconstruct the stratum and/or psu indicators and thus inadvertently break confidentiality. In this talk we will demonstrate the dangers and adapt algorithms used from scheduling theory and elsewhere to attempt to reduce this danger.

Les agences gouvernementales de statistique mettent régulièrement à la disposition du public des microdonnées dont les indicateurs de strate et/ou de up (psu) ont été omis pour des raisons de confidentialité. Dans le but d'estimer la variance d'estimateurs par la méthode de linéarisation, des indicateurs de pseudo-up sont parfois produits, alors que pour utiliser les méthodes de rééchantillonnage, on crée des poids de répliques. Si on ne fait pas attention, ceux-ci peuvent être utilisés pour (partiellement) reconstruire la strate et/ou l'indicateur de up et, de ce fait, briser la confidentialité sans le faire exprès. Dans cette présentation, nous démontrons les dangers associés à ce problème et adaptons des algorithmes de la théorie de la planification d'horaires et de d'autres théories afin de réduire ces dangers.

**[MS-142]****Hierarchical Bayes Small Area Estimation with Estimated Sampling Variances  
Estimation de la variance échantillonnale dans un petit domaine par un modèle hiérarchique bayésien**Yong YOU, *Statistics Canada/Statistique Canada*

We consider the situation where the sampling variances are estimated by direct estimators in small area estimation. A hierarchical Bayes (HB) model is constructed for the direct survey estimators and the sampling variances estimators. We also consider the EBLUP approach. We compare the HB and EBLUP methods through empirical studies. Our results have shown that the proposed HB estimators perform better than the EBLUP estimators in terms of taking account of the uncertainty of estimating the sampling variances, especially when the sample sizes are small. When the sample sizes are large, both the HB and EBLUP estimators perform equally well.

Nous considérons la situation où les variances d'échantillonnage sont estimées par des estimateurs directs dans l'estimation sur de petits domaines. Un modèle hiérarchique bayésien (HB) est construit pour les estimateurs de sondage directs et les estimateurs de variance échantillonnale. Nous considérons également l'approche EBLUP. Nous comparons les méthodes HB et EBLUP par des études expérimentales. Nos résultats ont montré que les estimateurs HB proposés performant mieux que les estimateurs EBLUP en tenant compte de l'incertitude dans l'estimation de la variance échantillonnale, particulièrement quand la taille de l'échantillon est petite. Quand les tailles échantillonnales sont grandes, les estimateurs HB et EBLUP performant aussi bien.

**[MS-143]****Repair of Two-phase Calibration Methodology in Survey Sampling  
Réparation de la méthodologie de calibration en deux phases en échantillonnage**Sarjinder SINGH, *Saint Cloud State University*

A question has been answered raised by the Deville and Särndal (1992) calibration approach to the eminent survey statisticians working at the U.S. Bureau of Census, Statistics Canada, Australian Bureau of

Une question soulevée par l'approche par calibration de Deville et Särndal (1992) aux éminents statisticiens de sondage travaillant au bureau du recensement des États-Unis, à Statistique Canada, au Bureau australien des statistiques et de

Statistics, and their consultants. This paper is especially designed to repair the work of Hidioglou and Särndal (1995, 1998), and hence that of Singh (2000), and is based on Golden Jubilee Year 2003 celebration by Singh (2003) of the linear regression estimator owed to Hansen, Hurwitz and Madow (1953). It is shown that chain regression estimator is unique, and is based on submitted JSM-2004 presentation.

leurs consultants, a été résolue. Cet article est particulièrement conçu pour réparer le travail de Hidioglou et Särndal (1995, 1998) et par conséquent celui de Singh (2000), et est basé sur la célébration du jubilé d'or en 2003 par Singh (2003) de l'estimateur de régression linéaire dû à Hansen, Hurwitz et Madow (1953). Il est montré que l'estimateur de régression en chaîne est unique, et est basé sur une présentation soumise à JSM-2004.

**Session 8F****Tuesday, June 1 • Mardi 1<sup>er</sup> juin, 15:30 – 17:00****Z-317****Design and Analysis of Health Studies****Planification et analyse d'études sanitaires****[MS-144]****Analysis of Binary Data in  $2 \times 2$  Crossover Designs: A Kappa Based Approach****Analyse de données binaires dans des designs croisés  $2 \times 2$  : une approche basée sur la mesure Kappa**Jie HUANG, *Novartis Pharmaceuticals, East Hanover, NJ*, Ming-Chun HSU, *MedFocus*

When outcomes are binary data, testing of treatment difference in  $2 \times 2$  crossover designs is carried out by comparing discordance of the paired data from period 1 and period 2 (Nagelkerke, 1986). This method and the mixed effect model with a conditional approach (Diggle et al., 1994) have a flaw that ignores the concordance data from two periods. Kappa is a measure of agreement between two raters that introduces a correction for chance-expected agreement (Cohen, 1960). This research utilizes the idea of Kappa measurement to evaluate treatment difference in  $2 \times 2$  crossover designs. The result is also compared to the conditional approach.

Quand les résultats sont des données binaires, tester une différence de traitement dans des plans croisés  $2 \times 2$  s'effectue en comparant la discordance des données appariées de la période 1 à celles de la période 2 (Nagelkerke, 1986). Cette méthode ainsi que le modèle à effets mixtes avec une approche conditionnelle (Diggle et autres, 1994) ont une faille qui ignore la concordance des données aux deux périodes. Le Kappa est une mesure d'accord entre deux juges qui introduit une correction pour l'accord chance-espéré (Cohen 1960). Cette recherche utilise l'idée de la mesure Kappa pour évaluer la différence de traitement dans des plans croisés  $2 \times 2$ . Le résultat est également comparé à l'approche conditionnelle.

**[MS-145]****Evaluation of Sample Size in Thorough QT/QTc Study****Évaluation de la taille échantillonnale dans une étude QT/QTc complète**Jie ZHANG, *Novartis Pharmaceutical, East Hanover, NJ*

The combination of QT/QTc interval prolongation and documented cases of TdP with the use of a drug has resulted in regulatory actions, including withdrawal from the market or denial of marketing authorization. Recent drafted ICH (E14) guideline recommends that new drugs should conduct an ECG trial which dedicates to evaluate their cardiac repolarization effect (thorough QT/QTc study). In this paper, I will summarize sample sizes of some

La combinaison de la prolongation d'intervalle de QT/QTc et des cas documentés de TdP avec l'utilisation d'un médicament a mené à des actions régulatrices, incluant le retrait du marché ou de l'autorisation de vente. La référence récente ICH (E14) recommande que les nouveaux médicaments devraient être sujets à une étude ECG consacrée à évaluer leurs effets de repolarisation cardiaque (étude QT/QTc complète). Dans cette présentation, je récapitulerai les tailles échantionales de telles études récentes et leur puissances post-hoc. Je



recently conducted such studies and their post-hoc powers. I will present simulation results on sample size when the maximal mean change of QTc is the primary endpoint to assess the central tendency.

présenterai des résultats de simulations sur la taille échantillonnale quand le changement moyen maximal de QTc est le point final primaire pour évaluer la tendance centrale.

#### [MS-146]

##### **Simultaneous Confidence Intervals for Umbrella Alternative Intervalles de confiance simultanés pour une alternative parapluie**

Jianan PENG, *Acadia University*, Charles Chu-In LEE, *Memorial University of Newfoundland*, Lin LIU, *University of California, San Diego*

In some applications, treatment effect usually increases with the increase of dosage up to a certain level, then makes a downturn due to side effects. This type of up-then-down response pattern is called an umbrella ordering. The turning point is called the peak of the umbrella ordering. There are several procedures to obtain simultaneous confidence intervals for umbrella contrasts of normal means in the literature. In this talk, we consider simultaneous confidence lower bound for the mean difference between the peak and minimum of the lowest dosage and the highest dosage. One algorithm to derive the lower bound is provided.

Dans quelques applications, l'effet de traitement augmente habituellement avec l'augmentation du dosage jusqu'à un certain niveau et baisse ensuite en raison des effets secondaires. Ce type de réponse qui augmente pour ensuite descendre s'appelle une disposition en parapluie. Le point de changement s'appelle la pointe de la disposition en parapluie. Il y a plusieurs procédures pour obtenir des intervalles de confiance simultanés pour des contrastes de parapluie de moyennes normales dans la littérature. Dans cette présentation, nous considérons la borne inférieure de confiance simultanée pour la différence moyenne entre la pointe et le minimum du plus bas et du plus élevé dosage. Un algorithme pour dériver la borne inférieure est fourni.

#### [MS-147]

##### **Composite Outcomes in Cluster-randomized Clinical Trials: Experiences from the Community Hypertension Assessment Trial (CHAT)**

##### **Issues multiples dans des études cliniques à grappes randomisées : expérience de l'étude sur l'évaluation de l'hypertension dans la communauté (CHAT)**

Lehana THABANE, *McMaster University*, K. O'ROURKE, *Ottawa Health Research Institute*, L. CHAMBERS, *University of Ottawa/Université d'Ottawa*, J. KACZOROWSKI, T. KARWALAJTYS & C. LEVITT, *McMaster University*, W. HOGG, *University of Ottawa/Université d'Ottawa*, L. DOLOVICH & C. BELLEY, *McMaster University*

This presentation examines the use of composite endpoint(s) in a clinical trial. We use our experiences from the CHAT study to examine the challenges of choosing composite endpoints. The CHAT study is a paired cluster-randomized trial to assess the effectiveness of community pharmacy blood pressure (BP) sessions, with feedback to patients and physicians, in the management of BP among seniors. It relies on chart reviews as the primary source of data, which leads to several methodological challenges that include choice of primary endpoint that is valid, reliable, sensitive to change, has good statistical properties including efficiency, and is clinically meaningful.

Cette présentation examine l'utilisation de point(s) final(aux) multiple(s) dans une épreuve clinique. Nous utilisons notre expérience de l'étude CHAT pour examiner les défis de sélection des points finaux multiples. L'étude CHAT est une épreuve à grappes randomisées appareillées pour évaluer l'efficacité des tests de tension artérielle (TA) en pharmacie, avec commentaires aux patients et aux médecins, dans la gestion de la TA parmi les personnes âgées. Elle se fonde sur la revue de dossiers médicaux comme source primaire de données ce qui mène à plusieurs défis méthodologiques incluant le choix d'un point final primaire valide, fiable, sensible au changement, qui a de bonnes propriétés statistiques incluant l'efficacité et qui fait du sens médicalement.

**[MS-148]****Designing Dose Response Experiments for Microbial Pathogens  
Design d'expériences dose-réponse pour des microbes pathogènes**William ROSS & Jagadish RANGREJ, *Health Canada/Santé Canada*

Quantal dose response relationships are a fundamental tool for managing risks to human health from the ingestion of microbial pathogens. Theoretical models have been developed. Importantly, unlike the logistic model, these models do not have a point of symmetry. Due to the serious consequences of ingesting these pathogens human clinical studies are impossible. However, with increasing understanding of the mechanism of infection, animal models are now possible. A statistical analysis of the problem of choosing design points for the purpose of estimating model parameters is provided. The analysis touches on optimality, the use of external information and adaptive designs.

Les relations dose-réponse quantiques sont un outil fondamental pour contrôler les risques pour la santé humaine de l'ingestion de microbes pathogènes. Des modèles théoriques ont été développés. D'une manière importante, contrairement au modèle logistique, ces modèles n'ont pas de point de symétrie. En raison des conséquences graves dues à l'ingestion de ces microbes pathogènes, des études cliniques sur des humains sont impossibles. Cependant, avec l'amélioration de la compréhension du mécanisme d'infection, les modèles animaux sont maintenant possibles. Une analyse statistique du problème de sélection des points de design pour estimer les paramètres du modèle est fournie. L'analyse touche l'optimalité, l'utilisation d'information externe et les design adaptatifs.

**Session 9A1      Wednesday, June 2 • Mercredi 2 juin, 08:30 – 09:15      AA-1140**

Pierre Robillard Award Address

Présentation de la lauréate du Prix Pierre-Robillard

**Session 9A2      Wednesday, June 2 • Mercredi 2 juin, 09:15 – 10:00      AA-1140**

Canadian Journal of Statistics Award Address

Allocution des récipiendaires du Prix de La revue canadienne de statistique

**Session 9B      Wednesday, June 2 • Mercredi 2 juin, 08:30 – 10:00      AA-1360**

Statistics and Science

La statistique et la science

**[MS-149]****Climate Change: Where on Earth Are We Going?****Les changements climatiques : Mais où donc la Terre s'en va-t-elle ?**Lawrence MYSAK, *McGill University*

The very long-term climate regimes during the 4.6 billion year history of earth will be reviewed. Concepts of "icehouses" and "greenhouses" will be defined and climate variability from the 100-kyr ice age cycles to the shorter centennial fluctuations will be discussed. Climate and atmospheric carbon dioxide concentration changes during the last 1000 years will be shown along with scenarios of future climate and environmental changes. Finally, an answer to the question asked in the title will be proposed—but there could be surprises!

Les régimes climatiques à très long terme des 4,6 milliards d'années d'histoire de la terre seront passés en revue. Les concepts des dépôts de glace et des serres seront définis et la variabilité climatique des cycles glaciaires de période de 100k années aux fluctuations plus courtes centennales seront discutés. Les changements climatiques et les changements dans la concentration atmosphérique du dioxyde de carbone pendant les 1000 dernières années seront discutés avec des scénarios sur les changements climatiques et environnementaux futurs. Finalement, une réponse à la question posée dans le titre sera proposée, mais il pourrait y avoir des surprises !

**[MS-150]****Political Arithmetic: Social Science, Scientific Revolution, and Political Founding**  
**Le calcul politique : sciences sociales, la révolution scientifique et fondations politiques**William ALLEN, *Michigan State University*

If we define social science by what “social scientists” do, it appears amorphous and unsystematic. We debate the systems and approaches of social science, and their methods range from the merely quantitative to the merely narrative. Our question becomes, what is the relation between the scientific revolution, legislation, and social science. Alexander Hamilton demonstrated that it is rather elements of bias than limitations in “the sciences of morals and politics” that confound men. Montesquieu justified the application of strict measurement in the formulation of laws or policies. Therefore social science is precisely the application of strict calculations to variable subjects.

Si nous définissons les sciences sociales par ce que les chercheurs en « sciences sociales » font, elles semblent amorphes et non systématiques. Nous discutons des systèmes et des approches de la science sociale et leurs méthodes qui s’étendent du simplement quantitatif au simplement narratif. La question d’intérêt est alors : quelle est la relation entre la révolution scientifique, la législation et la science sociale. Alexandre Hamilton a démontré que c’est plutôt des éléments de biais que des limitations dans les « sciences morales et politiques » qui confondent les hommes. Montesquieu a justifié l’application de la mesure stricte dans la formulation des lois ou des politiques. Par conséquent, les sciences sociales sont précisément l’application de calculs stricts à différents sujets.

**[MS-151]****Statistical Sleuthing in Walkerton**  
**Investigation statistique à Walkerton**Murray MCQUIGGE, *Public Health Consulting & University of Western Ontario*

This talk will examine the Walkerton waterborne disaster of E.coli O157:H7 of May 2000. The analysis of the data from this outbreak will be of interest to statisticians, government, industry, and academia persons from the point of view of: (1) The epidemiology of the outbreak. (2) Surveillance data. (3) Surveillance analysis. (4) New and future developments in Canadian surveillance initiatives and technology.

Cette présentation examinera le désastre d’Eucoli O157 :H7 dans l’eau de Walkerton en mai 2000. L’analyse des données de cette épidémie sera d’intérêt pour les statisticiens, le gouvernement, l’industrie et les personnes du milieu universitaire du point de vue de : (1) L’épidémiologie de l’épidémie. (2) Données de surveillance. (3) Analyse de surveillance. (4) Nouveaux et futurs développements dans les initiatives et la technologie canadienne de surveillance.

**Session 9C                      Wednesday, June 2 • Mercredi 2 juin, 08:30 – 10:00                      Z-310****New Researchers and New Directions for Research in Biostatistics**  
**Nouveaux chercheurs et nouvelles perspectives de recherche en biostatistique****[MS-152]****New Challenges from Smaller Microarrays**  
**Nouveaux défis provenant de microréseaux plus petits**Timothy RAMSAY, *University of Ottawa/Université d’Ottawa*, George DOUGLAS & Carole YAU, *Health Canada/Santé Canada*

New technology is now becoming available for researchers to print their own microarrays. These personalized arrays will be designed not to test all

De nouvelles technologies sont maintenant disponibles pour permettre aux chercheurs d’imprimer leurs propres microréseaux. Ces réseaux personnalisés seront conçus non pas

of an organisms' genes, but only those that are expected to show a significant change. Much of the current methodology for analyzing microarray data depends on the assumption that most genes will not change expression in response to an experimental treatment, and is therefore not appropriate for these small custom arrays. This talk is designed primarily to outline the problems that need to be solved in this wide-open area of research.

pour tester tous les gènes d'un organisme, mais seulement ceux qui sont soupçonnés de montrer un changement significatif. La plupart des méthodologies actuelles pour analyser des microréseaux dépendent de l'hypothèse que la majorité des gènes ne changeront pas leur expression en réponse à un traitement expérimental, et ainsi ne sont pas appropriées pour ces petits microréseaux personnalisés. Cette présentation est conçue principalement pour souligner le problème qui doit être résolu dans ce domaine de recherche largement ouvert.

### [MS-153]

#### Connections Between Stagewise Algorithms and the LASSO

#### Connexions entre quelques algorithmes infinitésimaux et le LASSO

Jonathan TAYLOR, Rob TIBSHIRANI, Trevor HASTIE & Guenther WALTHER, *Stanford University*

The Lasso (Tibshirani, 1996) is a method for regularizing least squares regression via L1 constraints. The LAR (Least angle regression) algorithm of (Efron et al., 2003) provides an efficient method for computing the entire sequence of Lasso solutions. In the process, the LAR algorithm also provides a conceptual link between the Lasso and Forward Stagewise regression. The latter strategy is an important component in adaptive regression procedures like boosting, and hence this link helps us understand how boosting works.

In this talk we give a sequential criterion that is optimized by Forward Stagewise regression: it features a sequential minimum L1 arc-length penalty. We also characterize problems for which the coefficient curves for Lasso are monotone as a function of the L1 norm; this is the situation where all three procedures (LAR, Lasso, and Forward Stagewise) coincide.

Tibshirani (1996) a introduit le LASSO dans la contexte de la régression linéaire avec une contrainte de type L1 sur les coefficients. Plus récemment, Efron et al (2003) ont introduit l'algorithme LAR, un algorithme efficace pour calculer toutes les solutions LASSO. En même temps, l'algorithme LAR a produit un lien entre le LASSO et des algorithmes infinitésimaux "forward stagewise", notamment "boosting" de moindres carrés.

On discute des liens entre ces algorithmes et on décrit des conditions sous lesquelles les algorithmes donnent des solutions identiques. On décrit aussi un critère pour lequel l'algorithme "forward stagewise" est optimal.

### [MS-154]

#### Multivariate Analysis of Microarray Expression Data

#### Analyse multivariée de données d'expression microréseaux

Rafal KUSTRA, *University of Toronto*

Classical multivariate techniques, like Fisher discriminant analysis, have been applied to microarray data with some success. However, the dimensionality challenge, where the number of correlated measurements far exceeds the number of observations ( $p \gg n$ ), restricts the applicability of such techniques to either a small number of highly pre-filtered genes, or to "super-gene" data that results

Les techniques multivariées classiques, comme l'analyse discriminante de Fisher, ont été appliquées aux données microréseaux avec un certain succès. Cependant, le défi de la dimension, où le nombre de mesures corrélées dépasse de loin le nombre d'observations ( $p \gg n$ ), limite la mise en œuvre de telles techniques à un nombre restreint de gènes fortement préfiltrés, ou à des données de « super gène » qui résultent d'une action de réduction de la dimension telle

from a dimensionality reduction activity, such as clustering or PCA. In this talk we show how appropriate penalization can be applied to multivariate analysis of microarray data without any prior data reduction. We will present an efficient computational method and show results on real microarray data.

que la classification ou l'ACP. Dans cette présentation, nous montrons comment une pénalisation appropriée peut être appliquée à l'analyse multivariée de données microréseaux sans aucune réduction préalable des données. Nous présenterons une méthode de calcul efficace et montrerons des résultats sur de vraies données microréseaux.

## Session 9D      Wednesday, June 2 • Mercredi 2 juin, 08:30 – 10:00      Z-330

### Non-traditional Survey Methods Méthodes d'enquête non traditionnelles

[MS-155]

#### Sampling of Images Collected by Digital Camera for Estimating the Level of Cracking on the Quebec Roadways

#### Échantillonnage d'images captées par caméra vidéo pour l'estimation de l'étendue de la fissuration sur les chaussées du Québec

Diane LEROUX, *Ministère des Transports du Québec*

The state of cracks on the roadways is an important indicator for the persons in charge of managing the road network at the Ministry for Transportation of Quebec (MTQ). The annual gathering of this information on more than 13 000 km of roadway represents a considerable task if it's not automated. While waiting for reliable fully automated machines, the MTQ developed a sampling method to analyse the cracking process by video images collected by a digital camera mounted on an asculation truck. The presentation will relate to the development of the method as well as its validation.

L'état de la fissuration des chaussées est un indicateur important pour les responsables de la gestion du réseau routier du ministère des Transports du Québec (MTQ). La cueillette annuelle de cette information sur plus de 13 000 km de chaussée représente un travail considérable s'il n'est pas automatisé. Dans l'attente d'appareils fiables complètement automatisés, le MTQ a élaboré une méthode d'analyse de la fissuration par échantillonnage au moyen d'images vidéo captées par une caméra numérique montée sur un camion d'auscultation. La présentation portera sur le développement de la méthode ainsi que sur les validations dont elle a fait l'objet.

[MS-156]

#### Modeling and Increase of Use of the Tax Data for the Unified Enterprise Survey

#### Modélisation et utilisation accrue des données fiscales dans le cadre de l'Enquête unifiée sur les entreprises

Éric PELLETIER, *Statistics Canada/Statistique Canada*

An agreement with the Canadian Customs & Revenue Agency allows Statistics Canada to use tax data coming from the businesses. In 2002, a project was launched with the objective to use more tax data in order to reduce the response burden of businesses, reduce the cost of survey programs and to improve the data quality. So, for reference year 2003, the Unified Enterprise Survey (UES) will use tax data as replacement data for about 50% of the surveyed simple-single establishments This

Une entente avec l'Agence des douanes et du revenu du Canada permet à Statistique Canada d'utiliser les données fiscales provenant des entreprises. En 2002, un projet a été mis de l'avant ayant pour objectif d'utiliser davantage les données fiscales dans le but de réduire le fardeau de réponse des entreprises, de réduire le coût des enquêtes et d'améliorer la qualité des données. Ainsi, pour l'année de référence 2003, l'Enquête unifiée sur les entreprises (EUE) utilisera des données fiscales comme données de remplacement pour environ 50% des établissements simples échantillonnés. Cette pré-

presentation will describe the issues and the difficulties at the estimation stage when using survey data and tax data.

sensation traitera des difficultés et des défis reliés à l'estimation impliquant des données d'enquêtes et des données fiscales.

**[MS-157]**

**Research on our Data Collection Activities; a Way to Better Understand and Improve Our Work**

**La recherche sur nos activités de collecte ; une façon de mieux comprendre et d'améliorer notre travail**

Tony LABILLOIS, *Statistics Canada/Statistique Canada*, Judy LEE, *Statistique Canada/Statistique Canada*

Efficiency of data collection was hard to quantify and management information used to be limited. With Computer Assisted Interviewing and the standardization of software, a wealth of information is now available. After how many attempts should we stop calling? What is the best time to call? How will this change improve efficiency? These are examples of what we can find with simple quantitative methods. A group is exploring data and developing tools to help collection managers to better understand and plan their activities. This presentation will give an overview of the objectives of the group and describe some results.

L'efficacité des activités de collecte de données a traditionnellement été difficile à quantifier et l'information de gestion pour améliorer la collecte était difficile à produire de façon compréhensive, actuelle et normalisée pour un grand nombre d'enquêtes. Avec l'arrivée de l'interview assistée par ordinateur et la normalisation du logiciel de collecte, une montagne d'information sur la collecte est maintenant disponible. Le système enregistre les activités des intervieweurs pendant le cycle de collecte. Après combien de tentatives devrait-on cesser d'appeler? Quel est le meilleur moment pour appeler? De quel manière ce changement améliorera-t-il l'efficacité? Ces questions ne sont que quelques exemples de toutes celles auxquelles on peut vouloir répondre avec des méthodes d'analyse exploratoire appliquées à ces données. Un groupe de Statistique Canada se penche sur ces données et développe des outils pour aider les gestionnaires de collecte à mieux comprendre et planifier leurs activités. Cette présentation donnera un aperçu des objectifs et du mandat de ce groupe et décrira certains exemples de résultats trouvés en utilisant les nouveaux prototypes d'outils pour l'analyse.

**Session 9E      Wednesday, June 2 • Mercredi 2 juin, 08:30 – 10:00      AA-1355**

**Stochastic Processes and Their Applications**

**Processus stochastiques et leurs applications**

**[MS-158]**

**Entropy and Semi-Markov Processes**

**Entropie et processus semi-markoviens**

Nikolaos LIMNIOS, *Université de Technologie de Compiègne*, Valérie GIRARDIN, *Université de Caen*

Entropy and Markov processes are linked since the first version of the asymptotic equipartition property stated by Shannon in 1948 for Markov chains. We define explicitly the entropy rate for semi-Markov processes.

Among a given collection of functions satisfying constraints, selecting the one with the maximum entropy is equivalent to adding the least amount of information possible to the problem considered.

L'entropie et les processus markoviens sont liés depuis la première version de la propriété d'équipartition asymptotique formulée par Shannon en 1948 pour des chaînes de Markov. Nous donnons une définition du taux d'entropie explicite pour les processus semi-markoviens. En choisissant dans une famille de fonctions réalisant des contraintes celle d'entropie maximum on ajoute le moins possible d'information au problème considéré. La définition d'un taux d'entropie explicite pour les processus permet de leur appliquer

The definition of an explicit entropy rate for processes allows one to extend to them the maximum entropy method. We study different problems for Markov and semi-Markov processes, illustrated in reliability, queuing theory, seismology, ...

les méthodes de maximum d'entropie. Nous étudions différents problèmes pour les processus markoviens et semi-markoviens, avec des illustrations en fiabilité, files d'attente, analyse sismologique...

**[MS-159]**

**Dependence Ordering of Markov Processes and Queueing Networks**

**Dépendance d'ordre de processus de Markov et de réseaux de files d'attentes**

Rafal KULIK, *University of Ottawa/Université d'Ottawa*, Hans DADUNA & Cornelia SAUER, *Universität Hamburg*, Ryszard SZEKLI, *Uniwersytet Wrocławski*

We characterize dependence orderings of Markov processes on partially ordered state space. A characterization is given in terms of infinitesimal generators. This is counterpart to a totally ordered case which is given in Hu and Pan (2000). As a special case it allows us to obtain correlation-in-time inequality for Markov processes. We applied our results for a comparison of a transient behavior of standard Jackson networks and Jackson networks which are subject to breakdowns and repairs.

Nous caractérisons la dépendance d'ordre des processus de Markov sur un espace partiellement ordonné. Une caractérisation est donnée en termes de générateurs infinitésimaux. C'est un contre-partie au cas totalement ordonné qui est donné dans Hu et Pan (2000). Comme un cas spécial, il nous permet d'obtenir l'inégalité de la corrélation dans le temps pour des processus de Markov. Nous avons appliqué nos résultats pour une comparaison d'un comportement passager des réseaux standard de Jackson et des réseaux de Jackson qui sont sujets aux pannes et aux réparations.

**[MS-160]**

**Construction of New Classes of Pseudonormal Markovian Stochastic Processes**

**Construction de nouvelles classes de processus stochastiques markoviens pseudo-normaux**

Jerzy FILUS, *Oakton Community College, DesPlaines, IL*, Lidia Z. FILUS, *Northeastern Illinois University*

Classes of new pseudonormal Markov stochastic processes with a discrete and with a continuous time are constructed. Among the obtained classes of stochastic processes a new generalization of the Wiener process is specified. Two parallel descriptions of that generalization, including the one by parabolic type partial differential equations, are obtained. In association with the above, a new extension of the notion of martingale is proposed. Also the problem of symmetry is raised. An analysis of nonsymmetric cases yields a variety of new trends or drifts of the considered stochastic processes that are or may be defined.

Des classes de nouveaux processus stochastiques markoviens pseudo-normaux à temps discret et continu sont construites. Parmi les classes de processus stochastiques obtenues, une nouvelle généralisation du processus de Wiener est spécifiée. Deux descriptions parallèles de cette généralisation, y compris celle par le type parabolique d'équations aux dérivées partielles, sont obtenues. Conjointement avec ce qui précède, nous proposons une nouvelle extension de la notion de martingale. Nous soulevons également le problème de la symétrie. Une analyse des cas non symétriques donne une variété de nouvelles tendances ou dérives des processus stochastiques considérés qui sont, ou peuvent être définies.

**[MS-161]****Weak Convergence Approach to Compound Poisson Processes Perturbed by Diffusion  
Convergence faible pour des processus de Poisson composés perturbés par un facteur de diffusion**Joykrishna SARKAR & Arusharka SEN, *Concordia University*

We consider a sequence of compound Poisson (CP) risk processes (RP) which converges weakly to the standard Wiener process as the number of claims goes to infinity. By adding this sequence to the classical risk process we obtain a sequence of RP which are themselves CP and converge weakly to the classical process perturbed by diffusion. We then show how the joint density function of the surplus immediately prior to ruin and other random variables, the expected discounted penalty function for the latter process can easily be obtained from the corresponding results for the classical process.

Nous considérons une séquence de processus de risque (PR) de Poisson composés (PC) qui converge faiblement vers un processus standard de Wiener quand le nombre de réclamations tend vers l'infini. En ajoutant cette séquence au processus classique de risque nous obtenons une séquence de PR qui sont eux-mêmes PC et qui convergent faiblement vers le processus classique perturbé par un facteur de diffusion. Nous montrons alors comment la fonction de densité conjointe du surplus immédiatement avant la ruine et d'autres variables aléatoires, la fonction escomptée de pénalité prévue pour le dernier processus, peut facilement être obtenue à partir des résultats correspondants pour le processus classique.

**[MS-162]****A Test for “Long-Memory” Processes  
Un test pour les processus à « longue mémoire »**David THOMSON, *Queen's University*

This talk describes a multitaper test to distinguish long-memory processes from ones that are nonstationary. A frequency-dependent test statistic is described which is the largest canonical correlation between the multitaper eigencoefficients averaged over data blocks offset in time. I examine several of the processes that have been described in the statistics literature as “long-memory” and find that persistent features generally have simpler explanations.

Cette présentation décrit un test de type « multitaper » pour distinguer les processus à longue mémoire de ceux qui sont non stationnaires. Une statistique de test dépendante de la fréquence est décrite comme la plus grande corrélation canonique entre la moyenne des valeurs propres « multitaper » moyennées sur des blocs de données décalées dans le temps. J'examine plusieurs des processus qui sont décrits dans la littérature statistique comme étant de type « longue mémoire » et constate que les caractéristiques persistantes ont généralement des explications plus simples.

**[MS-163]****On the Rate of Convergence of the Dependent Bootstrapped Means  
Sur le taux de convergence des moyennes bootstrap dépendantes**Andrei VOLODIN, *University of Regina*

The main focus of the present investigation is to obtain the convergence rates in the form of Baum–Katz/Erdos/Hsu–Robbins/Spitzer type complete convergence result (i.e., a strong law of large numbers type result) and one sided exact convergence rate (i.e., a law of the iterated logarithm type result) for the dependent bootstrapped means from a general sequence of random variables.

Le principal intérêt de la présente recherche est d'obtenir les taux de convergence sous la forme de type convergence complète du résultat de Baum–Katz/Erdos/Hsu–Robbins/Spitzer (c.-à-d., un résultat de type loi forte des grands nombres) et d'un taux de convergence exact unilatéral (c.-à-d., un résultat de type loi des logarithmes itérés) pour les moyennes bootstrap dépendantes provenant d'une suite générale de variables aléatoires.



**Session 9F      Wednesday, June 2 • Mercredi 2 juin, 08:30 – 10:00      Z-317****Survival Analysis II****Analyse de durées de vie II****[MS-164]****Bounds on Self-consistent Estimates of Survival with Multivariate Interval Censored Data****Bornes sur les estimations autoconvergentes de la survie avec des données censurées par intervalle multivariées**Xuecheng LIU & Alain C. VANDAL, *McGill University*, Robert GENTLEMAN, *Harvard University*

Nonparametric (NP) maximum likelihood estimates (MLE) of survival can be represented by discrete probability vectors. Such NPMLEs are always self-consistent (SC). We present lower and upper bounds for all SC estimate of survival given multivariate interval censored data (MICD). We present an interesting application of these bounds: with MICD, the survival NPMLE is not unique and the maximizers form a polytope well-defined by linear inequalities. We use our results on SC estimates to produce upper bounds on the size of this so-called nonuniqueness polytope.

Les estimations du maximum de vraisemblance (EMV) non paramétriques (NP) de la survie peuvent être représentées par des vecteurs de probabilité discrets. De tels EMVNP sont toujours autoconvergentes (AC). Nous présentons les bornes inférieures et supérieures pour toutes les estimations AC de la survie avec des données censurées par intervalle multivariées (DCIM). Nous présentons une application intéressante de ces bornes : avec des DCIM, L'EMVNP de la survie n'est pas unique et le maximum forme un polytope bien défini par des inégalités linéaires. Nous utilisons nos résultats sur des estimations AC pour obtenir des bornes supérieures sur la taille de ce polytope non unique.

**[MS-165]****Prediction of Recurrent Events****Prédiction d'événements récurrents**Marc FREDETTE & Jerry LAWLESS, *University of Waterloo*

In many applications there are certain types of events occurring repeatedly over time. It is often of interest to predict the future number of occurrences of these events. However, such a prediction problem often goes unrecognized and is treated as an estimation problem, an oversight that may lead to serious bias.

I will present useful prediction models for recurrent events. These events are modeled using flexible nonhomogeneous Poisson processes and a possible heterogeneity among the individuals is assessed using random effects. Also, I will present effective calibration techniques which provide prediction intervals with coverage probabilities close to the desired level.

Il est fréquent de rencontrer des situations où des événements se produisent plusieurs fois dans le temps. Il est intéressant de prédire le nombre d'occurrences de ces événements dans le futur. Malheureusement, un tel problème de prédiction est souvent traité comme un problème d'estimation, une négligence qui peut créer de sérieux biais.

Nous étudierons des modèles pour événements récurrents qui fournissent d'excellentes prédictions. L'occurrence de ces événements est modélisée avec des processus de Poisson non homogènes. De plus, ces modèles permettent une possible hétérogénéité entre les individus grâce à l'incorporation d'effets aléatoires. Nous discuterons également de méthodes de calibrage efficaces donnant des intervalles de prédiction avec des probabilités de couverture près du seuil désiré.

**[MS-166]****Issues in Testing and Model Misspecification Involving Multiple Event Data****Problèmes des méthodes de comparaison et de l'énoncé erroné d'un modèle d'analyse des délais d'apparition d'un évènement multiple**Jean-Marie BOHER & Richard COOK, *University of Waterloo*

Randomized clinical trials are frequently designed in settings where patients experience multiple clinical events over the course of follow-up. Methods for dealing with this type of data include techniques based on marginal models for multivariate failure time data (Wei, Lin, and Weissfeld, 1989) methods based on conditional models for event times (Prentice, Williams, and Peterson, 1981) and methods based on marginal rate functions (Andersen and Gill, 1982). In this talk we review assumptions underlying these different approaches, critically evaluate their appropriateness for recurrent event data, and derive asymptotic properties of estimators of treatment effect. The frequency properties of these tests are assessed for finite samples based on simulation studies. An application to an asthma data set is used to illustrate the interpretation of the tests.

De nombreux essais cliniques randomisés se déroulent dans un contexte où les patients développent plusieurs évènements cliniques en cours d'étude. Plusieurs méthodes d'analyse co-existent, des modèles basées sur une approche marginale (Wei, Lin, Weissfeld, 1989), conditionnelle (Prentice, Williams, Peterson, 1981) ou un modèle des taux d'incidence des fréquences marginales (Andersen, Gill, 1982). Une revue critique des hypothèses qui sous tendent les différents modèles et l'énoncé des propriétés asymptotiques des estimateurs de l'effet du traitement sont réalisés. Les propriétés de fréquence des tests pour échantillons finis sont évaluées par simulations. Un jeu de données sur l'asthme illustre l'interprétation des tests.

**[MS-167]****Bivariate Truncated Failure Time Data****Données de temps de bris tronquées bivariées**Rinku SUTRADHAR & Richard COOK, *University of Waterloo*

Incomplete failure time data often arise when observations are only available if they are in pre-specified truncation intervals. Our focus is on problems involving bivariate truncated failure time data where interest lies in the degree of association between paired failure times. We present a likelihood approach based on a conditional formulation using piecewise constant baseline hazard functions. Results from simulation studies are reported. The proposed methods are illustrated using a motivating data set on age of first hospitalization for mental illness for siblings where these hospitalizations were restricted to lie between 1926 and 1943.

Les données de temps de bris incomplètes surgissent souvent quand les observations sont seulement disponibles si elles sont dans des intervalles de troncation pré-spécifiés. Nous nous attardons aux problèmes impliquant des données de temps de bris tronquées bivariées où l'intérêt est le degré d'association entre les moments appariés de bris. Nous présentons une approche de vraisemblance basée sur une formulation conditionnelle en utilisant des fonctions de taux de panne de base qui sont constantes par morceaux. Des résultats d'études de simulations sont rapportés. Les méthodes proposées sont illustrées en utilisant un jeu de données sur l'âge de la première hospitalisation pour une maladie mentale pour des enfants de même progéniture où les hospitalisations sont limitées entre 1926 et 1943.

**Session 10A      Wednesday, June 2 • Mercredi 2 juin, 10:30 – 12:00      AA-1360**Statistical Hydrology  
Hydrologie statistique

**[MS-168]****On the Contribution of Statistics to Water Resources Management for Production  
Sur l'apport de la statistique à la gestion des ressources en eau à des fins de production hydroélectrique**Vincent FORTIN, *Institut de recherche d'Hydro-Québec*

Water represents a major energy resource for Quebec's economy. Approximately 95 % of the energy produced in Quebec comes from hydroelectricity. Under the effect of natural and socio-economic factors this resource is becoming rare and expensive. The goal of this introductory talk is to present the big issues of the hydroelectric production planning. In particular, we will discuss the contribution of statistical modeling to the decision-making process involved in the planning of production strategies and the development of new hydraulic sites. The presentations which will follow will provide precise examples of applications of statistics to water management.

L'eau constitue une ressource énergétique majeure pour l'économie du Québec. L'énergie produite au Québec est environ à 95% d'origine hydroélectrique. Sous l'effet de facteurs naturels et socio-économiques cette ressource devient rare et onéreuse. Cette communication introductive a pour objet de présenter les grands enjeux de la planification de la production hydroélectrique. Nous discuterons en particulier de l'apport de la modélisation statistique pour la prise de décision relative à l'établissement des stratégies de productions et à l'aménagement de nouveaux ouvrages hydrauliques. Les conférences qui suivront nous donnerons des exemples précis d'application de la statistique à la gestion des ressources en eau.

**[MS-169]****Bayesian Modelisation of Hydrographs  
Modélisation bayésienne d'hydrogrammes**James MERLEAU & Jean-François ANGERS, *Université de Montréal*, Luc PERREAULT, *Institut de recherche d'Hydro-Québec*

Generating synthetic hydrographs which reproduce characteristics of observed ones is a standing problem in hydrology. The difficulty of building a hydrograph generator stems from the varying periodicity of salient features from one year to another. Registration of hydrographs, according to important characteristics such as the spring flood, can be used to obtain a standard hydrograph. A Bayesian non-parametric model based on spline functions can then be used to generate hydrographs. From the posterior density on the spline coefficients, it is possible to generate sets of coefficients that define synthetic hydrographs which exhibit the features encountered in observed hydrographs.

Générer des hydrogrammes synthétiques qui reproduisent les caractéristiques de ceux observés est un problème important en hydrologie. La difficulté de construire un générateur d'hydrogramme provient de la périodicité variable des caractéristiques typiques d'une année à l'autre. L'enregistrement des hydrogrammes selon des caractéristiques importantes telles que la crue printanière, peut être utilisé pour obtenir un hydrogramme standardisé. Un modèle bayésien non paramétrique basé sur des fonctions de splines peut alors être utilisé pour générer des hydrogrammes. De la densité postérieure sur les coefficients des splines, il est possible de produire des ensembles de coefficients définissant les hydrogrammes synthétiques qui montrent les caractéristiques des hydrogrammes observés.

**[MS-170]****Seasonal Hydrological Forecast: Some Statistical Models Currently Used at Hydro-Québec  
Prévision hydrologique saisonnière : quelques modèles statistiques employés à Hydro-Québec**Luc PERREAULT, Vincent FORTIN & Jocelyn GAUDET, *Institut de recherche d'Hydro-Québec*, Oli SVEINSSON, *Columbia University*

We present three types of models used at Hydro-Québec to predict the flood volume 12 months, 6 months and 1 month in advance. The first type con-

Nous présentons trois types de modèles utilisés à Hydro-Québec pour prévoir les volumes de crue 12 mois, 6 mois et 1 mois à l'avance. Le premier type considère la persistance

siders the persistence observed in the annual inflow series. It consists of hidden Markov chains models. The second type attempts to develop the relationship which exists between flood volumes and certain climatic indicators (temperature on the surface of the oceans, pressure on the sea level, etc). The last type of model takes into account ground measures (covered nival and basic flow).

observée dans les séries d'apports annuels. Il s'agit de modèles à base de chaînes de Markov cachées. Le second type s'attache à valoriser la relation qui existe entre les volumes de crue et certains indicateurs climatiques (température à la surface des océans, pression au niveau de la mer, etc.). Enfin, le dernier type de modèles prend en compte des mesures de terrain (couvert nival et le débit de base).

### [MS-171]

#### What is the Future of Copulas in Hydrology?

#### Quel futur pour les copules en hydrologie ?

Anne-Catherine FAVRE & Salaheddine EL ADLOUNI, *Institut national de la recherche scientifique—Eau, Terre & Environnement*, Luc PERREAULT, *Institut de recherche d'Hydro-Québec*

For several applications in hydrology we need to use multivariate statistical distributions. Classical approaches are limited since they allow to use only a restricted number of distributions. Moreover the marginal distributions need to be the same. An alternate approach is the use of copulas. Copulas are distributions with uniform marginals. Copulas allow to model the dependence structure independently from the marginal distributions. This approach has been frequently used in finance but only rarely in hydrology. This talk presents several applications of copulas in hydrology and their future in this area.

Plusieurs applications en hydrologie nécessitent l'utilisation de lois statistiques multivariées. Les approches utilisées classiquement sont limitées car elles ne permettent d'utiliser qu'un nombre restreint de distributions. De plus les lois marginales doivent être similaires. Une approche alternative réside dans l'utilisation de copules. Les copules sont des fonctions de répartition dont les marges sont uniformes. Les copules permettent de modéliser la structure de dépendance indépendamment des distributions marginales. Cette approche a beaucoup été utilisée en finance, mais de manière très marginale en hydrologie. Cette conférence présente plusieurs applications des copules en hydrologie et leur avenir dans ce domaine.

## Session 10B      Wednesday, June 2 • Mercredi 2 juin, 10:30 – 12:00      AA-1140

### Empirical Likelihood Methods in Survey Sampling

### Méthodes empirico-vraisemblantistes en méthodologie d'enquête

### [MS-172]

#### Empirical Likelihood Methods for Sample Survey Data: An Overview

#### Méthodes de vraisemblance empirique pour des données de sondage : un survol

J. N. K. RAO, *Carleton University*

The use of empirical likelihood (EL) in sample surveys dates back to Hartley and Rao (1968). In this talk, I will present an overview of the developments in EL methods for sample survey data. Topics covered will include EL estimation using auxiliary population information, dual frame surveys and data with imputation for missing values. EL confidence intervals will be studied. Issues related to pseudo-EL for general sampling designs will also be discussed.

L'utilisation de la vraisemblance empirique (VE) dans les sondages remonte à Hartley et Rao (1968). Dans cette présentation, je ferai un survol des développements dans les méthodes de VE pour des données de sondage. Les sujets couverts incluront l'estimation de VE en utilisant de l'information auxiliaire sur la population, les sondages à deux bases de sondages et l'imputation pour des valeurs manquantes. Les intervalles de confiance de VE seront étudiés. Des problèmes reliés à la pseudo-VE pour les plans d'échantillonnage généraux seront également discutés.

[MS-173]

**Confidence Intervals for the Mean of a Population Containing Many Zero Values Under Unequal Probability Sampling****Intervalles de confiance pour la moyenne d'une population contenant plusieurs zéros sous un plan d'échantillonnage à probabilités inégales**Jiahua CHEN, *University of Waterloo*

Consider a population that contains many zero values and we wish to have a confidence interval for the population mean. Due to the skewness of the population distribution, the traditional central limit theorem based confidence interval has poor coverage probabilities.

In Kvanli, Shen and Deng (1998) and Chen, Chen and Rao (2003), the problem is investigated both parametrically and nonparametrically when samples are nearly independent and identically distributed.

In this talk, we consider the problem of constructing a confidence interval based on unequal probability sampling plans. We show that both parametric and nonparametric likelihood approaches are useful.

Soit une population qui contient beaucoup de valeurs zéro et nous souhaitons avoir un intervalle de confiance pour la moyenne de la population. En raison de l'asymétrie de la distribution de la population, l'intervalle de confiance traditionnel basée sur le théorème limite centrale a de faibles probabilités de couverture. Dans Kvanli, Shen et Deng (1998) et Chen, Chen et Rao (2003), le problème est étudié paramétriquement et non paramétriquement quand les échantillons sont presque indépendants et identiquement distribués. Dans cette présentation, nous considérons le problème de construction d'un intervalle de confiance basé sur des plans d'échantillonnage à probabilités inégales. Nous montrons que les approches de vraisemblance paramétriques et non paramétriques sont utiles.

[MS-174]

**Some Algorithmic Aspects of the Empirical Likelihood Method in Survey Sampling****Quelques aspects algorithmiques de la méthode de vraisemblance empirique en échantillonnage**Changbao WU, *University of Waterloo*

Recent developments in the empirical likelihood (EL) method for survey sampling have attracted attention from survey statisticians. Practical considerations for using the method depend largely on the availability of simple and efficient algorithms for computing the maximum EL estimators and/or the EL confidence intervals. In this talk, I will briefly describe the modified Newton–Raphson procedure of Chen, Sitter and Wu (2002) for nonstratified sampling designs, and then show that under suitable reformulation the algorithm can also be used to handle stratified sampling, the most commonly used design in survey practice. Several other computational issues will also be discussed.

Les développements récents dans la méthode de vraisemblance empirique (VE) en échantillonnage ont attiré l'attention des statisticiens qui travaillent en échantillonnage. Les considérations pratiques pour utiliser la méthode dépendent en grande partie de la disponibilité d'algorithmes simples et efficaces pour calculer les estimateurs de vraisemblance empirique maximale et/ou les intervalles de confiance basés sur VE. Dans cette présentation, je décrirai brièvement la procédure de Newton-Raphson modifiée de Chen, Sitter et Wu (2002) pour des plans d'échantillonnages non stratifiés, et ensuite prouver que sous une reformulation appropriée, l'algorithme peut également être utilisé pour manipuler des plans stratifiés, qui sont les plans d'échantillonnage les plus utilisés en pratique. Plusieurs autres problèmes computationnels seront également discutés.

**Session 10C****Wednesday, June 2 • Mercredi 2 juin, 10:30 – 12:00****Z-310**

Parametric Survival Analysis

Analyse paramétrique de durées de vie

**[MS-175]****Multivariate Accelerated Failure Time Regression Analysis  
Analyse de régression multivariée de temps de défaut accéléré**Jerry LAWLESS, *University of Waterloo, W. HE, Samuel Lunenfeld Research Institute*

This talk considers the analysis of multivariate survival times when the main interest concerns the effects of covariates on the marginal distributions. Parametric accelerated failure time or location-scale models are flexible, considerably easier to use than semiparametric analogues, and possess desirable efficiency and robustness properties. We demonstrate these features and illustrate them on bivariate data from a clinical trial on vision loss for persons with diabetes.

Cette présentation considère l'analyse des temps de survie multivariés où l'intérêt principal concerne les effets des covariables sur les distributions marginales. Les modèles paramétriques de temps accéléré ou les modèles localisation-échelle sont flexibles, considérablement plus faciles à utiliser que leurs analogues semi-paramétriques, et possèdent des propriétés souhaitables d'efficacité et de robustesse. Nous démontrons ces caractéristiques et nous les illustrons sur des données bivariées d'un essai clinique sur la perte de vision pour des personnes atteintes du diabète.

**[MS-176]****Over 50 Years of Log-normal Model use in Breast Cancer Investigations  
Plus de 50 ans d'utilisation du modèle log-normal dans les recherches sur le cancer du sein**Judy-Anne CHAPMAN, *Women's College Hospital & University of Waterloo*

Clinical investigations standardly utilize the Cox model although there may be substantive evidence of nonproportional hazards. Further, the usual step-wise selection of factors may incorrectly ascribe prominent factors. An overview of the literature shows a rationale for why a log-normal survival analysis model should be considered, at least in breast cancer investigations, and to provide evidence of why all subsets factor selection should be considered more routinely, as an adjunct to step-wise factor selection. The simple alternative of all subsets log-normal survival analysis will be demonstrated for several cohorts of breast cancer patients.

Les recherches cliniques utilisent de façon standard le modèle de Cox bien que des éléments du domaine d'application puissent ne pas supporter l'hypothèse de taux de panne proportionnels. De plus, la sélection pas-à-pas habituelle des facteurs peut attribuer incorrectement des facteurs dominants. Une revue de la littérature illustre pourquoi un modèle d'analyse de survie log-normal devrait être considéré, au moins dans les recherches sur le cancer du sein, et suggère que la sélection via la méthode de tous les sous-ensembles de facteurs devrait être considérée de façon plus routinière comme complément à la sélection pas-à-pas. L'alternative simple de l'analyse de survie log-normale via l'exploration de tous les sous-ensembles sera démontrée pour plusieurs cohortes de patientes atteintes du cancer du sein.

**[MS-177]****Evidence of Applicability of the Log-normal to Other Cancers, and Clinical Usefulness  
Évidence de l'applicabilité du modèle log-normal à d'autres types de cancer et utilité clinique**Patricia TAI, *University of Saskatchewan*

The log-normal model is useful for predicting long-term survival rates of a prospective series of 871 cervical cancer patients. For those who died with disease, Phase 1 used the minimum chi-square test to assess the goodness of fit of survival times to a log-normal distribution. Phase 2 used the maximum likelihood method to estimate the long-term survival rates from short-term follow-up data. Af-

Le modèle log-normal est utile pour prévoir des taux de survie à long terme d'une série prospective de 871 patients atteints d'un cancer cervical. Pour ceux qui sont morts de la maladie, la phase 1 utilise le test du khi-deux minimal pour évaluer l'adéquation des temps de survie à une distribution log-normale. La phase 2 utilise la méthode du maximum de vraisemblance pour estimer les taux de survie à long terme à partir des données de suivi à court terme. Après une période

ter a five-year period of diagnosis, the 12-year CSSR was predicted to be 90 percent at the end of two-year follow-up, and was validated by the Kaplan-Meier method. The model is also useful for other cancer sites.

diagnostique de cinq ans, la prédiction du CSSR à 12 ans était 90 pourcent à la fin de 2 ans de suivi et a été validée par la méthode de Kaplan-Meier. Le modèle est également utile pour d'autres types de cancer.

## Session 10D      Wednesday, June 2 • Mercredi 2 juin, 10:30 – 12:00      Z-330

### Case Study II: Anorexia Nervosa

### Étude de cas II : Anorexie mentale

[MS-178]

#### Presentation of the Data Set

#### Présentation du jeu de données

Peggy NG, *York University*

AN patients were assessed within two weeks of their admission to hospital. Interviews with their parents took place in a private setting on the hospital ward. The interviews with the control participants (parents and their daughters) took place either at their home or at a laboratory at the university, whichever was more convenient. At the point of initial contact, all participants were told that the purpose of the study was to investigate relationships among exercising, dieting, and certain personality factors in young women and their families. The parents were told that they would be asked about their daughter's physical activity history as a means of validating the information obtained from the child. Girls under 16 years of age provided verbal assent and then written consent from at least one parent. All parents participating in the study provided their own informed written consent. During a structured interview by Davis or her trained research assistant, general demographic information and the physical activity data were collected. Following that, participants were asked to complete the questionnaire measures. Each session lasted approximately an hour to an hour and a half.

Serial assessments of physical activity levels in adolescent patients with AN and their parents were obtained.

1) Investigate the different developmental trajectories of physical activity in the patient (AN) and control groups.

2) Do Female Adolescents with Anorexia Nervosa (AN) with a propensity to lose weight reflect a family transmission where activity levels are high and

Les patientes anorexiques ont été évaluées dans les deux premières semaines de leur séjour à l'hôpital. Leurs parents ont été interviewés dans un salon privé de l'hôpital. Les interviews avec les membres du groupe de contrôle (parents et filles adolescentes) se sont déroulés soit au domicile familial, soit dans un laboratoire de l'université, suivant ce qui était le plus commode.

Lors du contact initial, les chercheurs ont expliqué à tous les participants que l'objectif de l'étude était d'étudier les relations entre l'exercice physique, le régime alimentaire et certains facteurs de la personnalité chez les jeunes femmes et dans leurs familles. Ils ont précisé aux parents qu'ils leur demanderaient des détails sur l'activité physique de leur fille afin de valider les informations obtenues auprès des jeunes filles. Les filles de moins de 16 ans ont donné leur consentement verbal et fourni le consentement écrit d'un parent au minimum. Tous les parents qui ont personnellement participé à l'étude ont donné leur propre consentement informé par écrit.

Au cours d'entretiens structurés, Davis et ses assistants de recherche qualifiés ont collecté des informations démographiques générales et des données sur l'activité physique des adolescentes. À la suite de cette séance, ils ont demandé aux participants de compléter les mesures de questionnaire. Chaque séance a duré d'une heure à une heure et demie.

Les chercheurs ont ainsi obtenu des évaluations en série des niveaux d'activité physique chez les adolescentes anorexiques et chez leurs parents.

1) Explorez les trajectoires de développement de l'activité physique dans les groupes de patients (AN) et de contrôle.

2) Les adolescentes souffrant d'anorexie mentale (AN) et ayant une propension à perdre du poids reflètent un schéma de transmission familiale dans lequel les niveaux d'activité

over-consumption is avoided or discouraged. Does this behavioural profile increase risk for AN?

physique sont élevés et la surconsommation est évitée ou découragée. Ce profil de comportement augmente-t-il le risque d'anorexie AN ?

## Session 10E      Wednesday, June 2 • Mercredi 2 juin, 10:30 – 12:00      AA-1355

### Statistical Inference

#### L'inférence

[MS-179]

#### Conservative Priors for Hierarchical Models

#### Densités *a priori* conservatrices pour les modèles hiérarchiques

Shahadut HOSSAIN & Paul G. GUSTAFSON, *University of British Columbia*

Hierarchical models are suitable to model many real life phenomena, where data arise in a nested fashion. Estimating a phenomenon with hierarchical models can be viewed as a smoothing problem. While summarizing such phenomena via hierarchical model we do not want to undersmooth them because, in most practical applications undersmoothing is a more serious type of error than oversmoothing.

In order to guard against undersmoothing we have proposed a conservative prior for the variance component in Bayesian hierarchical models. We have investigated the performance of the proposed prior in normal-normal hierarchical model, and non-parametric regression curve estimation problem through simulation studies.

Les modèles hiérarchiques conviennent pour modéliser beaucoup de phénomènes de la réalité où les données surgissent de manière imbriquée. Estimer un phénomène avec des modèles hiérarchiques peut être vu comme un problème de lissage. Tout en résumant de tels phénomènes par des modèles hiérarchiques, nous ne voulons pas les sous-lisser parce que dans la plupart des applications pratiques, le sous-lissage est un type d'erreur plus sérieux que le sur-lissage.

Pour se prémunir du sous-lissage, nous avons proposé une densité *a priori* conservatrice pour la variance dans les modèles hiérarchiques bayésiens. Nous avons étudié la performance de la densité *a priori* proposée dans le modèle hiérarchique normal-normal et le problème d'estimation de la courbe de régression non paramétrique par des études de simulations.

[MS-180]

#### A Robust Fit for Generalised Additive Models

#### Un ajustement robuste pour des modèles additifs généralisés

Matias SALIBIAN-BARRERA & Azadeh ALIMADAD, *Carleton University*

Generalized Additive Models are a powerful modeling tool that extends nonparametric regression to Generalized Linear Models. It is easy to show that the backfitting algorithm used with these GAM models can be highly sensitive to atypical observations. As in classical linear regression, this lack of robustness may result in estimated relationships that fail to fit “good” observations well because they try to accommodate outliers. In this talk I will briefly introduce a robust fit for GAM models that is not severely affected by atypical observations in the data and that can be used to identify outliers in these models.

Les modèles additifs généralisés (MAG) sont des outils de modélisation puissants qui prolongent la régression non paramétrique aux modèles linéaires généralisés. Il est facile de prouver que l'algorithme de back-fitting utilisé par ces modèles MAG peut être extrêmement sensible aux observations atypiques. Comme en régression linéaire classique, ce manque de robustesse peut avoir comme conséquence que les relations estimées n'ajustent pas bien les « bonnes » observations parce qu'elles s'adaptent à des observations aberrantes. Dans cette présentation, je présenterai brièvement un ajustement robuste pour des modèles MAG qui n'est pas sévèrement affecté par des observations atypiques dans les données et qui peut être utilisé pour identifier des données aberrantes dans ces modèles.



**[MS-181]****Fitting Nonlinear Mixed Effects Models Robustly  
Ajustement robuste de modèles non linéaires à effets mixtes**Debbie DUPUIS, *University of Western Ontario*, Joanna MILLS-FLEMMING, *Université de Genève*

Nonlinear mixed effects (NME) models are used for inference in a number of applications. In this talk, we focus specifically on earthquake risk and insurance, and a detailed example of the prediction of peak horizontal accelerations is discussed. Relevant NME models are examined and the consequences of misspecifying assumptions on the error term are quantified. Well-known earthquake data is reanalyzed. A robust fit of the usual model, using recently developed robust weighted maximum likelihood estimators, is presented.

Les modèles non linéaires à effets mixtes (NEM) sont utilisés pour faire de l'inférence dans un certain nombre d'applications. Dans cette présentation, nous nous concentrons spécifiquement sur le risque des tremblements de terre et les problèmes d'assurance. Nous discutons également d'un exemple détaillé de la prévision des accélérations horizontales maximales. Des modèles NEM appropriés sont examinés et les conséquences des hypothèses mal spécifiées sur le terme d'erreur sont quantifiées. Des données bien connues sur les tremblements de terre sont réanalysées. Un ajustement robuste du modèle habituel à l'aide des estimateurs pondérés du maximum de vraisemblance robustes récemment développés est présenté.

**[MS-182]****On Decomposing the Watson Efficiency of OLS in a Partitioned Weakly Singular Linear Model  
Sur la décomposition de l'efficacité de Watson des MCO dans un modèle linéaire partitionné faiblement singulier**Ka Lok CHU, *McGill University*, J. ISOTALO & S. PUNTANEN, *University of Tampere, Finland/Université de Tampere, Finlande*, G. P. H. STYAN, *McGill University*

We consider the estimation of regression coefficients in a partitioned weakly singular linear model and focus on questions related to the Watson efficiency of the ordinary least squares (OLS) estimator of a subset of the parameters with respect to the best linear unbiased estimator. Certain submodels are also considered. The conditions under which the Watson efficiency in the full model splits into a function of some other Watson efficiencies is given special attention.

Nous considérons l'estimation des coefficients de régression dans un modèle linéaire partitionné faiblement singulier et nous nous concentrons sur des questions liées à l'efficacité de Watson de l'estimateur des moindres carrés ordinaire (MCO) d'un sous-ensemble de paramètres par rapport au meilleur estimateur linéaire sans biais. Certains sous-modèles sont également considérés. On porte une attention particulière aux conditions sous lesquelles l'efficacité de Watson dans les modèles complets se sépare en une fonction de d'autres efficacités de Watson.

**[MS-183]****Predictive Analysis of Stress-strength Reliability Models  
Analyse prédictive du modèle de fiabilité contrainte-force**Thu PHAM-GIA & N. TURKKAN, *Université de Moncton*

Stress-strength reliability, in its static form has the expression:  $P(X > Y)$ , where  $Y$ , the Stress and  $X$ , the Strength, are independent random variables with known distributions. In a Bayesian context, we now suppose that tests will be performed on  $X$ , and on  $Y$ , separately, resulting in posterior distributions for both, and hence in posterior system reli-

Le modèle de fiabilité « Contrainte-Force », sous sa forme simple est :  $P(X > Y)$ , où  $Y$ , la contrainte, et  $X$ , la force, sont des variables aléatoires indépendantes avec des distributions connues. Dans le contexte bayésien, on suppose que des tests seront appliqués à  $X$ , et à  $Y$ , séparément, conduisant à la fiabilité *a posteriori* du système,  $P(X^* > Y^*)$ .

Nous utiliserons la distribution prédictive conjointe des ré-

bility  $P(X^* > Y^*)$ .

We will use the joint predictive distribution of the possible outcomes of these tests to give a probability value to this posterior reliability. General results will be presented, and three specific cases will be studied in detail, with supporting graphs and charts.

sultats possibles de ces tests pour donner une probabilité à cette fiabilité *a posteriori*. Des résultats généraux seront présentés et trois cas spécifiques seront étudiés en détail.

**Session 10F**      **Wednesday, June 2 • Mercredi 2 juin, 10:30 – 12:00**      **Z-317**  
**Biostatistics II**  
**Biostatistique II**

**[MS-184]**

**The Evaluation of Information Retrieval Systems with Applications to Drug Discovery**

**Un paradigme de filtrage séquentiel pour détecter des items rares dans une grande base de données avec des application en découverte de médicaments**

Mu ZHU & Fuping HUANG, *University of Waterloo*

In the new information-based economy, the need to retrieve a small number of relevant items from a large database arises very often. The average precision is perhaps the most widely-used criterion for evaluating retrieval systems. A difficulty with the average precision is that it can make retrieval systems that behave very differently look very similar. Using a decision-theoretic argument, we introduce the notion of “optimal stopping recall” (OSR) and use it to evaluate a number of algorithms for drug discovery. We also briefly discuss potential connections between this work and the empirical measurement of human patience.

Dans la nouvelle économie de l’information, le besoin de rechercher un nombre restreint d’items appropriés dans une grande base de données se fait sentir très souvent. La précision moyenne est peut-être le critère le plus utilisé pour évaluer un système d’extraction de données. Une des difficultés de la précision moyenne est que des systèmes d’extraction de données qui fonctionnent très différemment obtiennent des résultats semblables. En utilisant un argument basé sur la théorie de la décision, nous introduisons la notion “d’arrêt de rappel optimal” (ARO) et l’utilisons afin d’évaluer plusieurs algorithmes pour la découverte de médicaments. Nous discutons également brièvement des connexions potentielles entre ces travaux et les mesures expérimentales de la patience humaine.

**[MS-185]**

**Using Multilevel Modeling to Study the Patient Care Delivery System**

**Modélisation multi-niveaux pour étudier le système de soins de santé**

Xiaoqiang LI, *University of Toronto*

Health science research often concerns problems that have a hierarchical structure, for example, when patients are nested within units and units are nested within hospitals. Without accounting for the possible clustering of effects within hospital units, the conclusions of the study could be invalid. It will generally cause inflation of type 1 errors which means more significant findings than it should be will be reported. An example of multiple membership classification is presented to illustrate how to handle complex hierarchical structure in patient care delivery system.

La recherche en science de la santé porte souvent sur des problèmes qui ont une structure hiérarchique, par exemple, quand des patients sont rassemblés dans des unités qui elles à leur tour sont regroupées dans les hôpitaux. Si on ne tient pas compte de l’effet possible des groupes d’unités à l’intérieur d’un hôpital, les résultats de l’étude pourraient être erronés. Ceci causera généralement une augmentation de l’erreur de type 1, ce qui implique que des résultats plus significatifs qu’en réalité seront rapportés. Un exemple de classification d’appartenance multiple est présenté pour illustrer comment manipuler la structure hiérarchique complexe dans les systèmes de soins de santé.

**[MS-186]****Data Mining Electronic Medical Records****Forage de données de dossiers médicaux électroniques**

Cathlin MCNALLY, Lehana THABANE & Anne HOLBROOK, *McMaster University*, Kathy GAEBEL, *St. Joseph's Healthcare, Hamilton*

This presentation will discuss the application of data-mining techniques using a large electronic medical record (EMR) database (COMPETE I). In COMPETE I, each family practice used the same EMR software, which stores patient demographics, diagnostics, and therapies. The database contains 36 tables and over 180 variables. Key variables include diseases, medications, and laboratory results. We applied the CART (classification and regression trees) technique to develop predictive models and study associations between outcomes and prognostic factors. These patterns and models are intended to pilot a prototype pharmacosurveillance system. The presentation will also discuss the challenges involved in preparing EMRs for data mining.

Cette présentation discutera de l'application des techniques de forage de données en utilisant une grande base de données (COMPETE I) de dossiers médicaux électroniques (DME). Dans COMPETE I, chaque clinique a utilisé le même logiciel de DME, qui stocke l'information sur la démographie, le diagnostic et les thérapies des patients. La base de données contient 36 tableaux et plus de 180 variables. Les variables principales incluent les maladies, la médication et les résultats de laboratoires. Nous avons appliqué la technique CART (arbres de classification et de régression) pour développer des modèles prédictifs et pour étudier les associations entre les résultats et les facteurs pronostiques. Ces modèles sont destinés à suivre un système prototype de pharmacosurveillance. La présentation discutera également des défis impliqués lors de la préparation des données DME pour le forage de données.

**[MS-187]****A Covariate-based Coefficient of Source Dependence for Capture-recapture Models Applied to Epidemiologic Studies****Un coefficient de dépendance de source basé sur les covariables pour des modèles de capture-libération appliqués aux études épidémiologiques**

Elizabeth TURNER & Alain C. VANDAL, *McGill University*

In capture-recapture modeling it is necessary to account for possible dependence between sources. Under the assumption that sources are conditionally independent given the covariates, we propose a coefficient of source dependence (CSD) for every set of sources. The CSDs are formed by two components: the distribution of the covariates in each source and the population distribution of the covariates. The first is estimated empirically and the second, along with the unknown population size, by including the CSDs as parameters into what we term an augmented marginal log-linear model. We describe advantages and applications of this technique.

Dans le contexte des modèles de capture-libération il est nécessaire de tenir compte d'une possible dépendance entre les sources. En supposant que les sources sont indépendantes conditionnellement aux valeurs de covariables, on propose un coefficient de dépendance de sources (CDS) pour chaque ensemble de sources. Les CDS sont formés de deux parties : la loi conditionnelle des covariables dans chaque source et la loi inconditionnelle des covariables. La première est estimée empiriquement et la seconde, la taille de la population étant inconnue, en incorporant les CDS en tant que paramètres dans ce qu'on appelle un modèle log-linéaire marginal augmenté. On décrit les avantages et les applications de notre technique.

**[MS-188]****Segregation of Quantitative Traits with Covariates  
Ségrégation de traits quantitatifs avec des covariables**J. LOREDO-OSTI, N. ROSLIN, & K. MORGAN, *McGill University*, A. BONNARDEAUX, *Université de Montréal*

There are genetic studies in which quantitative traits are often dichotomized into affected and unaffected after having adjusted them for confounding risk factors to perform segregation analysis. This leads to a loss in power. Segregation analysis with a major quantitative trait is an analysis of mixtures, constrained by the Mendelian rules of inheritance. We present an extended segregation analysis to take simultaneously a major locus and covariates to quantitative trait variation in independent families. The method is used to analyze a data set on hypercalciuria in French-Canadian families who were ascertained through a proband who had calcium oxalate kidney stones.

Il y a des études génétiques dans lesquelles des traits quantitatifs sont souvent dichotomisés comme affectés et non affectés après les avoir ajustés pour les facteurs de risque parasites dans le but de faire des analyses de ségrégation. Ceci mène à une perte de puissance. L'analyse de ségrégation avec un trait quantitatif majeur est une analyse de mélanges, contrainte par les règles Mendéliennes de l'hérédité. Nous présentons un prolongement de l'analyse de ségrégation pour prendre simultanément un locus important et des covariables à la variation quantitative de trait dans des familles indépendantes. La méthode est utilisée pour analyser un jeu de données sur l'hypercalciurie dans les familles canadiennes françaises qui ont été identifiées à travers un proposant qui a eu des pierres rein d'oxalate de calcium aux reins.

**Session 11A      Wednesday, June 2 • Mercredi 2 juin, 13:30 – 15:00      Z-310****Health Surveys****Les enquêtes dans le domaine de la santé****[MS-189]****Socioeconomic Status and Nonfatal Injuries Among Adolescents: Variations Across SES and Injury Measures****Statut socio-économique et les blessures non fatales chez les adolescents : variations à travers les mesures de SSE et les mesures de blessures**Elizabeth POTTER, *University of Ottawa/Université d'Ottawa*, Kathy N. SPEECHLEY, *Child Health Research Institute, London, ON*, John J. KOVAL, *University of Western Ontario*, Iris A. GUTMANIS, *Southwest Region Health Information Partnership, London, ON*, M. Karen CAMPBELL, *University of Western Ontario*, Douglas MANUEL, *Institute for Clinical Evaluative Sciences, Toronto*

This study explored whether the observed relationship between socioeconomic status (SES) and non-fatal injuries among adolescents was dependent on the measure of SES chosen (among 5 possible measures) or the type of injury considered. The analysis was based on the 1996-1997 National Population Health Survey (NPHS). Among males, a clear positive gradient with injury was observed for only 1 SES indicator, with total and recreation injuries. Among females, a positive relationship between SES and injuries was most apparent for 1 SES measure in relation to total and recreation injuries. Issues arising from the analysis of NPHS data will be discussed.

Cette étude explore si la relation observée entre le statut socio-économique (SSE) et les blessures non fatales parmi des adolescents dépend de la mesure de SSE choisie (parmi 5 mesures possibles) ou le type de blessure considéré. L'analyse est basée sur le sondage national sur la santé de la population de 1996 -1997 (SNSP). Parmi les hommes, nous avons observé un gradient positif clair avec blessures pour seulement un indicateur de SSE, avec toutes les blessures et les blessures récréatives. Parmi les femmes, une relation positive entre le SSE et les blessures était plus évidente pour une mesure de SSE par rapport à toutes les blessures et les blessures récréatives. Des problèmes résultant de l'analyse des données du SNSP seront discutés.

**[MS-190]****The Linkage of Household Survey Data with Administrative Records — A Mental Health Services Example from Ontario****Le jumelage de données de sondage sur les ménages avec des dossiers administratifs : un exemple sur les services de santé mentale en Ontario**Anne RHODES, *St. Michael's Hospital, Toronto*

This study inspected the finding that persons at a high distress level report more mental health service use than recorded in their health care records. The study sample,  $N = 36,892$ , 12 years and older, was drawn from the 1996/97 Ontario Health Survey. Respondents were individually linked to their administrative health care records in the past year. Almost 2/3 of those who were depressed in the past year were currently highly distressed. Differential reporting of use for highly distressed persons remained despite examination of alternative hypotheses. This study accentuates concerns about inferences made with self-reported mental health service use.

Cette étude investigate la constatation que les personnes à un niveau élevé de détresse font une plus grande utilisation des services de santé mentale que ce qui est enregistré dans leur dossier médical. L'échantillon de l'étude de taille  $N = 36892$ , de personnes de 12 ans et plus a été tiré du sondage sur la santé en Ontario de 1996-1997. Les répondants ont été individuellement associés à leur dossier médical administratif des dernières années. Presque le 2/3 de ceux qui ont été en dépression dans les années précédentes sont présentement fortement en détresse psychologique. Le rapport différentiel de l'utilisation pour les personnes fortement en détresse est demeuré, en dépit de l'examen d'hypothèses alternatives. Cette étude accentue les préoccupations concernant les inférences faites avec l'information auto-rapportée sur l'utilisation des services de santé.

**Session 11B      Wednesday, June 2 • Mercredi 2 juin, 13:30 – 15:00      AA-1360****Non-iid Sampling Data****Données d'échantillonnage non iid****[MS-191]****About Efficient Estimation and Valid and Powerful Testing in Quantitative Linear Models with Autocorrelated Errors****A propos de l'efficacité d'estimateurs et de la validité et la puissance de tests dans les modèles linéaires quantitatifs avec erreurs autocorrélées**Gülhan ALPARGU, *University of Massachusetts*, Pierre DUTILLEUL, *McGill University*

Questions of efficiency of slope estimators and validity and power of testing procedures are crucial in the analysis of quantitative linear models with autocorrelated errors. That is particularly true in the stepwise procedure of selection of an explanatory variable when the regressor already in the model is of different nature. In this context, we have studied the efficiency of six slope estimators, together with the validity and power of seven testing procedures. Theoretical results completed by simulations will be presented and discussed in relation to the nature of regressors, the sample size and the autocorrelation structure.

Les questions d'efficacité des estimateurs de pentes et de validité et puissance des procédures de test sont cruciales dans l'analyse des modèles linéaires quantitatifs avec erreurs autocorrélées. Cela est particulièrement vrai dans la procédure pas à pas de sélection d'une variable explicative lorsque le régresseur déjà présent dans le modèle est de nature différente. Dans ce contexte, nous avons étudié l'efficacité de six estimateurs de pentes ainsi que la validité et la puissance de sept procédures de test. Des résultats théoriques complétés par des simulations seront présentés et discutés en relation avec la nature des régresseurs, la taille de l'échantillon et la structure d'autocorrélation.

**[MS-192]****To Be or not to Be i.i.d.? A Very Good Question for a Sample of Data****Être ou ne pas être « i.i.d. » ? Une très bonne question pour un échantillon de données**Pierre DUTILLEUL, *McGill University*

Sample data can be ‘non-i.i.d.’ in different ways, through autocorrelation or lack of independence (see first ‘i.’) and parameters of the statistical distribution that are not identical for all observational units (see ‘i.d.’). Data transformation, modified estimation and modified testing are among the solutions proposed to overcome the problem when the ‘i.i.d.’ condition is recommended in statistical analysis. Some of the solutions are model-based, whereas the others are model-free. I will illustrate my literature review with examples, focusing on own experience over the years. I will also present some recent results.

Des données d’échantillonnage peuvent être non-i.i.d. de différentes manières, par la présence d’autocorrélation ou manque d’indépendance (voir le premier i. dans ‘i.i.d.’) et par des paramètres de la distribution statistique qui ne sont pas identiques pour toutes les unités observationnelles (voir ‘i.d.’). La transformation des données, l’estimation modifiée et le test modifié figurent parmi les solutions proposées pour surmonter le problème lorsque la condition i.i.d. est recommandée dans l’analyse statistique. Certaines solutions sont basées sur des modèles, d’autres pas. J’illustrerai ma revue de littérature par des exemples, en me concentrant sur ma propre expérience à travers les années. Je présenterai également des résultats récents.

**[MS-193]****Assessing Relationships Between Multivariate Spatial Data at Multiple Scales****L’évaluation de relations multiéchelles avec données spatiales multivariées**Bernard PELLETIER, *McGill University*

In the linear model of coregionalization,  $p$  spatial processes are decomposed into a mean component plus  $S$  random components, with same basic semivariogram functions for each process. Its fitting through auto- and cross-semivariograms generates sill matrix estimates used to predict the  $S$  spatial components of each process and assess their scale-specific relationships. A lack of stationarity at first order can affect that fitting. In addition, factorial cokriging predictors suffer from a smoothing effect that affects the covariance structure among predicted spatial components. I will present two new predictors of spatial components and a procedure to model the mean of processes.

Dans le modèle linéaire de corégionalisation,  $p$  processus spatiaux sont décomposés en une moyenne et  $S$  composantes aléatoires, avec les mêmes fonctions de base du semivariogramme pour chaque processus. Son ajustement aux semivariogrammes (autos et croisés) génère des matrices de paliers estimés, utilisées dans la prédiction des composantes spatiales de chaque processus et l’évaluation de leurs relations à chaque échelle. Un manque de stationarité du premier ordre peut affecter cet ajustement. De plus, les prédicteurs corrigés souffrent d’un effet de lissage qui affecte la structure de covariance entre les composantes spatiales prédites. Je présenterai deux nouveaux prédicteurs des composantes spatiales et une procédure de modélisation de la moyenne des processus.

**[MS-194]****Accounting for Genetic and Environmental Sources of Correlation in QTL Mapping Studies****Tenir compte des sources de corrélations génétiques et environnementales dans les études de mapping du LTQ**Hans-Peter PIEPHO & K. EMRICH, *Universität Hohenheim*

Plant breeders have an increasing interest in biometrical methods for mapping quantitative trait loci (QTL) of economically relevant traits. QTL mapping projects typically involve series of trials from multiple environments (locations, years). The re-

Les horticulteurs ont un intérêt croissant envers les méthodes biométriques pour tracer les locus de trait quantitatif (LTQ) des traits économiquement intéressants. Les projets de mapping de LTQ comportent habituellement une série d’épreuves dans de multiples environnements (endroits, an-

sulting data usually display marked genotype-by-environment interaction. QTL position and effects may be estimated by regression of phenotypic data on genetic marker data. This paper will present a mixed modeling approach to QTL mapping, which is exemplified using different multi-environment datasets. The importance of accounting for genetic as well as environmental sources of correlation is emphasized.

nées). Les données résultantes montrent habituellement une interaction génotype-environnement. La position et les effets du LTQ peuvent être estimés par la régression de données phénotypiques sur des données de marqueurs génétiques. Cette présentation propose une approche modéliste mixte au mapping de LTQ, qui est exemplifié en utilisant différents jeux de données multi-environnement. L'importance de tenir compte des sources de corrélation génétiques aussi bien qu'environnementales est soulignée.

## **Session 11C      Wednesday, June 2 • Mercredi 2 juin, 13:30 – 15:00      AA-1140**

### **Analysis of Longitudinal Data Analyse de données longitudinales**

[MS-195]

#### **Transition Models for Multivariate Longitudinal Binary Data Modèles de transition pour des données binaires longitudinales multivariées**

Richard COOK & Leilei ZENG, *University of Waterloo*

Interest often lies in tracking how two stochastic processes evolve and in such settings the scientific focus is on the associations between transition occurrences. We give estimating equations for joint estimation and inference of transitional models for multivariate longitudinal binary data based on GEE2 (Zhao and Prentice, 1990) and alternating logistic regression (Carey et al., 1993). We show that this method results in consistent estimates of the transition and association parameters, and can lead to improved efficiency. Application to a longitudinal study in the intensive care illustrates these methods. Extensions for multivariate and multinomial longitudinal data are indicated.

Il est souvent d'intérêt de déterminer comment deux processus stochastiques évoluent et dans de telles situations, le défi scientifique porte sur les associations entre les occurrences de transition. Nous donnons des équations d'estimation pour l'estimation conjointe et l'inférence des modèles transitoires pour des données binaires longitudinales multivariées basées sur les EEG2 (Zhao et Prentice, 1990) et la régression logistique alternante (Carey et collaborateurs, 1993). Nous montrons que cette méthode donne des estimations convergentes des paramètres de transition et d'association et peuvent mener à l'amélioration de l'efficacité. L'application à une étude longitudinale dans les soins intensifs illustre ces méthodes. Nous indiquons également les prolongements pour des données longitudinales multivariées et multinomiales.

[MS-196]

#### **Backfitting and Local Likelihood Methods for Nonparametric Mixed-effects Models with Longitudinal Data Backfitting et méthodes de vraisemblance locale pour des modèles non paramétriques à effets mixtes avec des données longitudinales**

Hulin WU, *University of Rochester*, Jeong-Gun PARK, *Frontier Science and Technology Research Foundation, Inc.*

For longitudinal data analyses, we consider a nonparametric mixed-effects model that characterizes both population effect and random effects as nonparametric functions. We propose a novel approach to use the local likelihood concept and a backfitting algorithm to combine the local polynomial regression techniques and the linear mixed-effects (LME) model methods for efficiently estimating

Pour des analyses de données longitudinales, nous considérons un modèle non paramétrique à effets mixtes qui exprime l'effet de la population et les effets aléatoires sous forme de fonctions non paramétriques. Nous proposons une nouvelle approche pour utiliser le concept de vraisemblance locale et un algorithme de « backfitting » pour combiner les techniques de régression polynomiale locale et le modèle linéaire à effets mixtes (LME) pour estimer efficacement la

both population mean and individual curve functions. The asymptotic properties of the proposed estimators are established. The simulation studies show that our approach performs better than the existing methods. We illustrate our estimation methods with an application to an AIDS clinical study.

moyenne de la population et les courbes individuelles. Les propriétés asymptotiques des estimateurs proposés sont établies. Les études de simulation montrent que notre approche performe mieux que les méthodes existantes. Nous illustrons nos méthodes d'estimation avec une application à une étude clinique sur le SIDA.

**[MS-197]**

**Robust Estimation of Parameter-driven Models for Longitudinal Count Data**

**Estimation robuste de modèles guidés par des paramètres pour des données de dénombrement longitudinales**

Rachel MacKay ALTMAN & Brian LEROUX, *University of Washington*

In this talk, we propose a general class of parameter-driven (latent variable) models for longitudinal count data. This class includes the generalized linear mixed model, hierarchical generalized linear model, and the hidden Markov model. We consider the interpretation of these models, and suggest guidelines for model selection within this class. We also discuss a parameter estimation method which yields estimates of the regression coefficients that are both efficient and robust to misspecification of the latent process. We apply these ideas to the analysis of a multiple sclerosis data set.

Dans cette présentation nous proposons une classe générale de modèles guidés par des paramètres (variable latente) pour des données de dénombrement longitudinales. Cette classe inclut le modèle linéaire mixte généralisé, le modèle linéaire généralisé hiérarchique et le modèle caché de Markov. Nous considérons l'interprétation de ces modèles et faisons des suggestions pour la sélection de modèles à l'intérieur de cette classe. Nous discutons également d'une méthode d'estimation des paramètres qui donne des estimations des coefficients de régression efficaces et robustes à la mauvaise spécification du processus latent. Nous appliquons ces idées à l'analyse d'un jeu de données sur la sclérose en plaques.

**Session 11D      Wednesday, June 2 • Mercredi 2 juin, 13:30 – 15:00**

**Z-330**

**Statistics and the Physical Sciences**

**La statistique et les sciences physiques**

**[MS-198]**

**Fundamental Modeling of Chemical Processes—Statistical Challenges and Approaches**

**Modélisation fondamentale de processus chimiques, défis et approches statistiques**

Kim MCAULEY, *Queen's University*

Chemical engineers develop mathematical models, based on their knowledge of physical and chemical phenomena that occur in laboratory and industrial-scale chemical reactors. I will describe how the complex differential and algebraic equation models are developed, and how they are used. The talk will focus on why it is often very difficult to estimate the model parameters from experimental data. I will introduce techniques for overcoming parameter estimation problems, and methods for designing additional experiments aimed at improving the estimates of the most important model parameters. Polymerization reactor examples will be used to illustrate the techniques.

Les ingénieurs chimiques développent des modèles mathématiques basés sur leur connaissance des phénomènes physiques et chimiques qui se produisent en laboratoire et dans les réacteurs chimiques à échelle industrielle. Je décrirai comment les modèles d'équations différentielles et algébriques complexes sont développés et comment ils sont utilisés. La présentation indiquera pourquoi il est souvent très difficile d'estimer les paramètres des modèles à partir des données expérimentales. Je vais introduire des techniques pour surmonter les problèmes d'estimation de paramètres et des méthodes pour concevoir de nouvelles expériences visant à améliorer les estimateurs des paramètres les plus importants du modèle. Des exemples du réacteur de polymérisation seront utilisés pour illustrer les techniques.



**[MS-199]****Bayesian Analysis of RR Lyrae Distances and Kinematics****Analyse bayésienne des distances et cinématiques des étoiles de RR Lyrae**

William JEFFERYS & Thomas R. JEFFERYS, *University of Texas at Austin*, Thomas G. BARNES, *McDonald Observatory*

We have developed a hierarchical Bayes model to analyze the distances, luminosities, and kinematics of RR Lyrae stars. Our model relates these parameters to the observed proper motions, radial velocities, apparent luminosities and metallicities of the stellar sample. We use a Metropolis-within-Gibbs sampler to draw an MCMC sample from the full posterior distribution of the parameters (including latent variables), and draw inferences on the quantities of interest in the usual way. We are testing our model with a small database from the literature and will eventually apply it to a new large database from the European HIPPARCOS satellite.

Nous avons développé un modèle de Bayes hiérarchique pour analyser les distances, la luminosité et la cinématique des étoiles de RR Lyrae. Notre modèle relie ces paramètres aux mouvements propres observés, aux vitesses radiales, aux luminosités apparentes et aux métallicités de l'échantillon stellaire. Nous utilisons un échantillonneur Metropolis-dans-Gibbs pour tirer un échantillon MCMC de la distribution *a posteriori* des paramètres (y compris les variables latentes), et nous faisons de l'inférence sur les quantités d'intérêt de la manière habituelle. Nous testons notre modèle avec une petite base de données provenant de la littérature et nous l'appliquerons éventuellement à une vaste nouvelle base de données provenant du satellite européen HIPPARCOS.

**[MS-200]****Multivariate Deconvolution and Applications****Déconvolution multivariée et applications**

Peter KIM, *University of Guelph*

Let  $Z$  be a random matrix and suppose  $Z = GX$ , where the random matrix  $X$  denotes the true random measurement and the random matrix  $G$  is an independent error matrix. The multivariate deconvolution problem involves the estimation of the density function of  $X$  from a random sample from  $Z$ . Sharp minimax results for the multivariate deconvolution problem are possible and have important applications to medical imaging and physical chemistry. A class of imaging problems dealing with single photon emission tomography (SPECT) is discussed along with a problem in texture analysis.

Soit  $Z$  une matrice aléatoire et posons  $Z = GX$  où la matrice aléatoire  $X$  dénote la véritable mesure aléatoire et la matrice aléatoire  $G$  est une matrice d'erreurs indépendantes. Le problème multivarié de déconvolution comporte l'estimation de la fonction de densité de  $X$  à partir d'un échantillon aléatoire de  $Z$ . Il est possible d'obtenir des résultats minimax précis pour le problème multivarié de déconvolution et ils ont des applications importantes en imagerie médicale et en chimie physique. Nous discutons d'une classe de problèmes en imagerie traitant de la tomographie d'émission de photons simples (SPECT) et également d'un problème en analyse de textures.

**Session 11E****Wednesday, June 2 • Mercredi 2 juin, 13:30 – 15:00****Z-317****Testing Procedures****Procédures de test**

**[MS-201]****Power Similarity of Two Groups of Tests for the Exponential Distribution  
Similarité des puissances de deux groupes de tests pour la distribution exponentielle**Richard LOCKHART & Michael A. STEPHENS, *Simon Fraser University*

Two groups of tests for the exponential distribution, with apparently very different motivations, have been observed to have very similar powers. The tests are the straightforward tests based on the empirical distribution function when the scale parameter must be estimated, and the tests known as Total-time-on-test statistics (using similar test statistics), arising from reliability considerations. Consideration of the Gaussian processes on which they depend shows the statistics to be very close on the null hypothesis and for contiguous alternatives. For other alternatives, Monte Carlo studies also demonstrate closeness, and this explains the similar powers.

Nous avons observé que deux groupes de tests pour la distribution exponentielle, basés sur des motivations apparemment très différentes, ont des puissances très semblables. Les tests sont ceux basés directement sur la fonction de répartition empirique quand le paramètre d'échelle doit être estimé, et les tests connus sous le nom de temps-total-sur-test (utilisant des statistiques de test semblables) inspirés de considérations de fiabilité. En considérant les processus gaussiens sur lesquels ils dépendent, on note que les statistiques sont très proches pour l'hypothèse nulle et pour des alternatives contiguës. Pour d'autres alternatives, les simulations de Monte Carlo démontrent également la proximité et ceci explique les puissances semblables.

**[MS-202]****Clustering of Goodness-of-fit Tests  
Groupement des tests d'adéquation**Steve OVERDUIN & Michael A. STEPHENS, *Simon Fraser University*

Many goodness-of-fit tests, say for the normal or for the exponential distribution, appear to do the same job — rejecting nearly the same samples for the same size test, and having very similar powers. We report a clustering procedure to attempt to put similar tests together, with the intention of picking for use those tests which appear to have the most natural motivation, and also to suggest combinations of tests which will overall have greater power.

Plusieurs tests d'adéquation, disons pour la distribution normale ou exponentielle, semblent faire le même travail ; ils rejettent presque les mêmes échantillons pour le même niveau du test et ils ont des puissances très semblables. Nous rapportons une procédure de classement pour tenter de regrouper les tests semblables, avec l'intention d'utiliser les tests qui semblent avoir la motivation la plus naturelle et afin de suggérer également des combinaisons de tests qui auront une plus grande puissance.

**[MS-203]****Multiple Testing in a Hierarchical Framework  
Tests multiples dans un contexte bayésien**Aurélien LABBE & Mary E. THOMPSON, *University of Waterloo*

We study the problem of testing the equality of two sample means under a Bayesian framework, when the number of tests performed is large. We propose to use the posterior probability of the one-sided alternative as a measure of significance for each test. We show that under some conditions on the model this probability can be seen as an approximation of a frequentist p-value. It can then be used in some standard multiple testing procedures developed un-

Nous étudions le problème des tests multiples dans un contexte bayésien. Nous proposons d'utiliser la probabilité *a posteriori* de l'hypothèse alternative comme valeur seuil pour chaque test. Nous montrons que sous certaines conditions du modèle, cette probabilité peut s'écrire comme une approximation d'une valeur-p. Nous pouvons donc utiliser une probabilité combinant l'information *a priori* et échantillonnale pour des procédures usuelles de tests multiples, développées dans un contexte fréquentiste. Nous appliquons

der a frequentist framework, with the advantage of working with a probability combining the prior and the observed information. An application will be seen in microarray-data analysis.

ces résultats à des données de microréseaux.

**[MS-204]**

**Accuracy of Different Methods when Testing a Binomial Proportion  
Précisions de différentes méthodes pour tester une proportion binomiale**

Ana-Maria STAICU & Nancy REID, *University of Toronto*

In this talk we will emphasize the importance of using higher order approximations as an alternative to the commonly used first order approximations (Wald, score, likelihood ratio). The results are tested on some examples of interest in epidemiologic studies. We present the basic framework for applying third order approximations, with specific reference to the exponential model. We then focus on the case of binomials, which is important in analyzing pure-count cohort data, or case-control data. The accuracy of the p-values obtained by first and higher order approximations is investigated and the tests that give more precise results are shown.

Dans cette présentation, nous soulignons l'importance d'utiliser des approximations d'ordre supérieur comme alternative aux approximations du premier ordre généralement utilisées (Wald, score, rapport de vraisemblance). Les résultats sont testés sur quelques exemples d'intérêt en épidémiologie. Nous présentons le cadre de base pour appliquer les approximations du troisième ordre avec le modèle exponentiel comme référence spécifique. Nous nous concentrons alors sur le cas binomial qui est important dans l'analyse de données de cohortes de dénombrement pur ou de données cas-témoin. La précision des valeurs-p obtenues à partir des approximations du premier ordre et d'ordres supérieurs est étudiée et nous indiquons les tests qui donnent les résultats les plus précis.

**[MS-205]**

**Point Estimation Following a Group Sequential Test  
Estimation ponctuelle suivant un test séquentiel par groupes**

Keyue DING, *Queen's University*

Repeated significance tests not only increase the overall type I error, but also introduce bias in estimating the unknown primary parameter. In this paper, a bias adjusted estimate of the drift parameter of a Brownian motion is proposed conditioning on the stopping time. Monte Carlo studies show that it outperforms other existing estimates in terms of overall mean squared error.

Les tests d'hypothèses répétés augmentent non seulement l'erreur globale de type I, mais introduisent également un biais dans l'estimation du paramètre inconnu primaire. Dans cette présentation, un estimateur du paramètre de dérive d'un mouvement Brownien ajusté pour le biais est proposé conditionnellement au temps d'arrêt. Des études de Monte-Carlo montrent qu'il surpasse d'autres estimateurs existants en fonction de l'erreur quadratique moyenne.

**[MS-206]**

**Preliminary Test Estimation in Inverse Gaussian Data  
Estimation préliminaire pour des données gaussiennes inverses**

Debaraj SEN & Yogendra CHAUBEY, *Concordia University*

This proposal deals with some inference problems related with inverse Gaussian models. We consider preliminary test estimation for the mean in the one sample problem. We also adapt an estimator from the normal case for the mean of an inverse Gaus-

Cette présentation traite de quelques problèmes d'inférence reliés aux modèles gaussiens inversés. Nous considérons l'estimation préliminaire pour la moyenne dans le problème d'un seul échantillon. Nous adaptons également un estimateur du cas normal pour celui de la moyenne d'une distribu-

sian distribution. The properties of the new estimator will be investigated using numerical studies. Efficiency of the new estimators are analyzed with respect to the sample mean.

tion gaussienne inversée. Les propriétés du nouvel estimateur seront étudiées en utilisant des études numériques. L'efficacité des nouveaux estimateurs dans le cas de la moyenne échantillonnale est analysée.

**Session 11F      Wednesday, June 2 • Mercredi 2 juin, 13:30 – 15:00      AA-1355**

**Epidemiology  
Épidémiologie**

[MS-207]

**Modeling Patient Preferences for Colorectal Cancer (CRC) Screening Strategies**

**Modélisation des préférences des patients pour des stratégies de dépistage du cancer colorectal (CCR)**

Lehana THABANE, D. A. MARSHALL & J. K. MARSHALL, *McMaster University*, F. R. JOHNSON, *Research Triangle Institute, NC*, K. A. PHILLIPS, *University of California, San Francisco*, G. FOSTER, *McMaster University*, B. J. O'BRIEN, *McMaster University*

Regular screening is often recommended to reduce mortality from colorectal cancer (CRC). CRC screening is at suboptimal levels in the United States and is currently under consideration in Canada. One approach to increase screening rates is to have better understanding of the desirable test features, and to develop and promote tests accordingly.

This presentation will provide the results of a study using a discrete choice experiment to measure patient preferences for CRC screening attributes and estimate the trade-offs between the levels of attributes using a sample of 1074 patients aged 40–60 years from an Ontario primary care network.

Le dépistage régulier est souvent recommandé pour réduire la mortalité dû au cancer colorectal (CCR). Le dépistage du CCR est à un niveau sous optimal aux États-Unis et est actuellement à l'étude au Canada. Une approche pour augmenter le taux de dépistage est d'avoir une meilleure compréhension des caractéristiques désirables des tests, et de développer et favoriser les tests en conséquence.

Cette présentation fournira les résultats d'une étude qui utilise une expérience à choix discrets pour mesurer les préférences des patients pour les caractéristiques de dépistage du CRC et pour estimer les différences entre les niveaux des caractéristiques en utilisant un groupe de 1074 patients âgés de 40 à 60 ans d'un réseau de soin primaire d'Ontario.

[MS-208]

**Stratified and Randomized Play-the-winner Rule for Designs in Clinical Trials**

**Règle miser-sur-le-gagnant stratifiée et randomisée pour des designs d'études cliniques**

Yuanyuan LIANG & Keumhee Chough CARRIÈRE, *University of Alberta*

A new simpler allocation rule is proposed for treatment assignments in sequential clinical trials. The stratified and randomized play-the-winner rule (SRPWR) is an extension of randomized play-the-winner rule (RPWR) to more than two treatments. It is also a modified rule for the case where the success probabilities of a treatment depend on both treatments and known confounders (for example, patient's age, gender and disease status). On the average, the implementation of the SRPWR in a clinical trial tends to assign more patients to the

Nous proposons une nouvelle règle plus simple pour l'attribution des traitements dans les épreuves cliniques séquentielles. La règle miser-sur-le-gagnant stratifiée et randomisée (SRPWR) est une prolongation de la règle miser-sur-le-gagnant randomisée (RPWR) à plus de deux traitements. C'est également une règle modifiée dans le cas où les probabilités de succès d'un traitement dépendent des deux traitements et d'un facteur parasite connu (par exemple l'âge, le sexe et le statut de la maladie). En moyenne, la réalisation du SRPWR dans une épreuve clinique tend à attribuer plus de patients au meilleur traitement. Elle élimine également le

better treatment. It also eliminates the selection bias and allows delayed responses to treatments from patients.

biais de sélection et permet des réponses retardées aux traitements des patients.

**[MS-209]**

**Bayesian Synthesis of Information in a Complex Evidence Structure: HIV Surveillance in England and Wales, 2001**

**Analyse bayésienne de données à structure complexe : surveillance du VIH en Angleterre et au pays de Galles, 2001**

Aicha GOUBAR & A. E. ADES, *Medical Research Council—Health Services Research Collaboration, Bristol, D.*  
DE ANGELIS, *Medical Research Council—Biostatistics Unit, Cambridge, O. N.* GILL, *Health Protection Agency—Communicable Disease Surveillance Centre, London*

Routine HIV surveillance systems in the UK generate different types of data from several distinct surveys: HIV prevalence, prevalence of undiagnosed HIV, risk-group size, proportion diagnosed, and risk-group composition of diagnosed individuals. We constructed models with 67 parameters to simultaneously estimate HIV prevalence, risk-group size and proportion diagnosed in 9 HIV risk-groups in 3 regions, from 76 available data-points from year 2001, using Bayesian MCMC. Hierarchical structures shared information across risk-groups and regions, and methods were developed to incorporate 'biased' data. This new approach to synthesis of surveillance information allows us to assess the consistency of multiple sources of evidence.

Les systèmes de surveillance routinier du VIH en Angleterre produisent différents types de données à partir de plusieurs sondages distincts : la prévalence du VIH, la prévalence de VIH non diagnostiqué, la taille du groupe à risque, la proportion diagnostiquée et la composition du groupe à risque d'individus diagnostiqués. Nous avons construit des modèles avec 67 paramètres pour estimer simultanément la prévalence du VIH, la taille du groupe à risque et la proportion diagnostiqués dans 9 groupes à risque de 3 régions, à partir de 76 données ponctuelles de l'année 2001 en utilisant des MCMC bayésiennes. Des structures hiérarchiques qui partagent de l'information à travers les groupes à risque et les régions ainsi que des méthodes ont été développées pour incorporer des données biaisées. Cette nouvelle approche pour l'analyse d'information de surveillance nous permet d'évaluer l'uniformité de plusieurs sources de données.

**[MS-210]**

**Assessing Sensitivity to Unobserved Confounding in Observational Studies: A Bayesian Approach**

**Évaluation de la sensibilité des facteurs parasites non observés dans des études d'observation : une approche bayésienne**

Lawrence MCCANDLESS & Paul G. GUSTAFSON, *University of British Columbia, A. R. LEVY, University of British Columbia & St. Paul's Hospital*

Assessing sensitivity to unobserved confounding in observational studies traditionally involves sensitivity analysis (SA) techniques using a range of bias parameters obtained from external information. Bayesian methods offer an innovative approach to SA by incorporating uncertainty about bias parameters into the analysis as prior information. We consider a Bayesian approach to SA in the context of logistic regression models for a dichotomous exposure, outcome and unobserved confounder. Using simulation techniques, Bayesian SA is contrasted to traditional SA. We present an application

L'évaluation de la sensibilité dû aux facteurs parasites non observés dans des études d'observation implique traditionnellement des techniques d'analyse de sensibilité (AS) en utilisant une gamme de paramètres de biais obtenus à partir d'informations externes. Les méthodes bayésiennes offrent une approche innovatrice à l'AS en incorporant l'incertitude par rapport aux paramètres de biais comme de l'information a priori. Nous considérons une approche bayésienne à l'AS dans le contexte des modèles de régression logistiques pour des variables explicatives, des résultats et des facteurs parasites non observés dichotomiques. En utilisant des techniques de simulation, l'AS bayésienne est contrastée par rap-

assessing the effects of highly active antiretroviral therapy on triglyceride levels in persons infected with HIV.

port à l'AS traditionnelle. Nous présentons une application évaluant les effets de la thérapie antirétrovirale fortement active aux niveaux de la tryglycérider chez les personnes atteintes du VIH.

**Session 11G**      **Wednesday, June 2 • Mercredi 2 juin, 13:30 – 15:00**      **Z-240**

Panel Discussion on the Mathematical Sciences Institutes and Programs: As a Statistician, Should I be Concerned?

Débat sur les instituts et les programmes en sciences mathématiques : En tant que statisticien, suis-je concerné ?

Christian LÉGER, *Centre de recherches mathématiques*, Thomas SALISBURY, *Fields Institute*, Manfred TRUMMER, *Pims*, Arvind GUPTA, *MITACS*, Jamie STAFFORD, *NPCDS*

The Director or Deputy Director of CRM, Fields Institute, Pims, MITACS, and NPCDS take a look at the current relationship between the Mathematical Sciences Institutes and Programs and the statistical community and explore ways to expand them. Each organisation will briefly describe its purpose, past statistical activities that they have done as well as look at the future opportunities for interaction and activities with the statistical community. A discussion with the audience will follow. We hope that this session will provide good ideas to improve the relationship between the institutes and the statistical community for the mutual benefit of both.

Le directeur ou le directeur adjoint du CRM, du Fields Institute, du Pims, de MITACS et du PNSDC examinent les liens entre les instituts et les programmes en sciences mathématiques et la communauté statistique et explorent des avenues pour les augmenter. Chaque organisation décrit brièvement sa raison d'être, ses activités statistiques passées, ainsi que les opportunités pour de nouvelles interactions et activités conjointes avec la communauté statistique. Une discussion avec le public suivra. Nous espérons que cette séance générera de bonnes idées afin d'améliorer les liens unissant les instituts et la communauté statistique pour le bénéfice mutuel de tous.

**Session 12A**      **Wednesday, June 2 • Mercredi 2 juin, 15:30 – 17:00**      **AA-1140**

Address of the 2004 CRM-SSC Prize Winner  
Allocution du lauréat du Prix CRM-SSC 2004

**Session 12B**      **Wednesday, June 2 • Mercredi 2 juin, 15:30 – 17:00**      **Z-330**

Nonparametric and Bayesian Methods  
Méthodologie bayésienne et non paramétrique

[MS-211]

**Noninformative Priors Based on Asymptotic Likelihood Methods**

**Densités *a priori* non informatives basées sur des méthodes de vraisemblance asymptotiques**

Xiaobin YUAN & D. A. S. FRASER, *University of Toronto*

Many Bayesian analyses are performed with noninformative priors. Box & Tiao (1973) used a “data-translated likelihood” to motivate the use of uniform priors and introduced “approximated data-translated likelihood” to motivate Jeffreys’s gen-

Plusieurs analyses bayésiennes sont faites avec des loi *a priori* non informatives. Box et Tiao (1973) ont utilisé une fonction de « vraisemblance traduite par les données » pour motiver l'utilisation des densités *a priori* uniformes et ont introduit « la fonction de vraisemblance approximative tra-

eral rule. Recently developed asymptotic likelihood methods show that a general model can be approximated by a location model to third-order in a ‘first derivative’ neighborhood of a data point. This location reparametrization can then be used as a metric for uniform priors. We focus on constructing priors for a scalar parameter of interest in the presence of nuisance parameters.

duite par les données » pour motiver la règle générale de Jeffreys. Les méthodes de vraisemblance asymptotiques récemment développées montrent qu’on peut faire l’approximation d’un modèle général par un modèle de position au troisième degré dans un voisinage de la « première dérivée » d’une donnée. Cette reparamétrisation de la position peut alors être utilisée comme une métrique pour les densités *a priori* uniformes. Nous nous concentrons sur la construction de densités *a priori* pour un paramètre d’intérêt scalaire en présence de paramètres de nuisance.

### [MS-212]

#### A New Class of Prior Distributions in Bayesian Nonparametric Statistics

#### Une nouvelle classe de distributions *a priori* en statistique bayésienne non paramétrique

Raluca BALAN, *University of Ottawa/Université d’Ottawa*

We consider a new class of nonparametric prior distributions on the real line which satisfies the Markov property. This class includes the neutral to the right (NR) priors and it was recently proved to be “closed” in the Bayesian sense. Despite its analytical complexity, what makes a Markov prior distribution more appealing than a NR prior is the following fact: in the case of a Markov prior distribution, the posterior probability that a new data point falls beyond  $t$  given a sample, depends on the exact values of those observations that are greater than  $t$ , not only on their number.

On considère une nouvelle classe de lois *a priori* sur la ligne droite qui satisfait la propriété de Markov. Cette classe inclut les lois *a priori* neutres à droite (ND) et est « fermée » dans le sens bayésien. Malgré sa complexité analytique, une loi *a priori* de type markovien est plus attrayante qu’une loi ND à cause de la caractérisation suivante : dans le cas d’une loi *a priori* markovienne, la probabilité *a posteriori* qu’une nouvelle observation soit supérieure à une valeur  $t$ , ayant observé un échantillon de données, dépend des valeurs exactes de ces données qui sont plus grandes que  $t$ , et pas seulement de leur nombre.

### [MS-213]

#### The Wilcoxon Signed-rank Test for Cluster Correlated Data

#### Le test de Wilcoxon des rangs signés pour données corrélées en grappes

Denis LAROCQUE, *HEC Montréal*

In this paper, we adapt the Wilcoxon signed-rank test to the case of cluster correlated data. A simple modification of the estimator of the asymptotic variance is sufficient to obtain a valid asymptotic procedure. We derive the asymptotic null distribution of the statistic. A simulation study is performed in order to investigate the finite sample performance of the test. The results show that the performance of the test is very good for all designs and distributions considered when compared to competitors based on signs and on the overall average.

Dans cette présentation, nous adaptons le test des rangs signés de Wilcoxon au cas de données corrélées en grappes. Une simple modification de l’estimateur de la variance asymptotique est suffisante pour obtenir une procédure valide. Nous obtenons la loi asymptotique de la statistique. La performance du test pour petits échantillons est étudiée à l’aide d’une simulation. Les résultats montrent que la performance du test est très bonne, comparativement à des tests basés sur les signes et sur la moyenne globale, dans tous les cas considérés.

**[MS-214]****On Blest's Correlation Coefficient****À propos du coefficient de corrélation de Blest**Jean-François PLANTE, *University of British Columbia*, Christian GENEST, *Université Laval*

Blest (2000) proposed a nonparametric coefficient of correlation penalizing discrepancies in the first ranks. Such a statistic cannot fulfill Scarsini's (1984) axioms, but may yet provide improved power in testing independence when compared to Spearman's  $\rho$  and Kendall's  $\tau$ . Limiting distribution of Blest's coefficient and of an improved variant are presented. Their corresponding population parameters are calculated for different models of importance including the normal paradigm. The talk is based on results published in the CJS in 2003.

Blest (2000) propose un coefficient de corrélation non paramétrique punissant plus sévèrement les désordres qui surviennent dans les premiers rangs. Une telle statistique ne peut respecter tous les axiomes de Scarsini (1984), ce qui n'empêche guère le coefficient de Blest (et une variante proposée) d'offrir une performance supérieure au  $\rho$  de Spearman et au  $\tau$  de Kendall en tant que statistique pour tester l'indépendance. La loi limite des coefficients proposés est présentée. Leurs paramètres de population sont calculés pour certains modèles d'importance incluant le paradigme normal. Ces résultats ont fait l'objet d'un article dans la RCS en 2003.

**[MS-215]****Clustering for Functional Biomechanical Data Using a Model-based Distance Measure****Classification de données fonctionnelles biomécaniques en utilisant une mesure de distance basée sur un modèle**Jeffrey BAKAL, Glen TAKAHARA, J. Terry SMITH & Joan STEVENSON, *Queen's University*

We consider a method for clustering smooth functional data, which is a hybrid of the model- and distance-based approaches. Pre-processing reduces dimensionality while preserving the functional nature of the data. We register the curves, permitting extraction of pertinent point properties, then model the curves using basis functions. In the clustering phase, we use a dissimilarity derived from an approximation to the joint probability of membership, computed using an estimated multivariate normal mixture. The resulting dendrogram, which depicts the distance between subclusters as their probability of disjunction, is an attractive feature. Advantages of this method are displayed using functional biomechanical data.

Nous considérons une méthode pour grouper des données fonctionnelles lisses qui est un hybride des approches basées sur les modèles et sur les distances. Le pré-traitement réduit la dimension tout en préservant la nature fonctionnelle des données. Nous faisons coïncider les courbes permettant l'extraction des propriétés ponctuelles pertinentes, ensuite nous modélisons les courbes en utilisant des fonctions d'une base. Dans la phase de classification, nous utilisons une dissimilarité dérivée d'une approximation de la probabilité conjointe d'appartenance calculée en utilisant un mélange normal multivarié estimé. Le dendrogramme résultant, qui montre la distance entre les sous-groupes comme leur probabilité de disjonction, est une caractéristique attrayante. Les avantages de cette méthode sont montrés en utilisant des données biomécaniques fonctionnelles.

**[MS-216]****Tests for Duration Clustering and Diagnostic Checking of ACD Models Using Kernel-based Spectral Density Estimators****Tests pour effets ACD et évaluation de l'adéquation des modèles ACD avec des estimateurs à noyau de la densité spectrale**Pierre DUCHESNE, *Université de Montréal*, Maria PACURAR, *HEC Montréal*

ACD models have been widely used to model financial data. We propose two classes of tests for ACD effects and one class for the adequacy. The

Les modèles ACD sont utilisés pour modéliser des données financières. Nous proposons deux classes de tests pour tester les effets ACD et une classe pour l'ajustement. Les tests d'ef-



tests of ACD effects of the first class are obtained by comparing a kernel-based spectral density estimator and the spectral density under the null hypothesis of no ACD effects. The second class exploits the one-sided nature of the alternative hypothesis. Tests for the adequacy are obtained by comparing a kernel-based spectral density estimator of the residuals and the null hypothesis of adequacy. Asymptotic distributions are obtained. Simulations and an application with financial data are presented.

fets ACD de la première classe sont obtenus en comparant un estimateur à noyau de la densité spectrale et la densité spectrale sous l'hypothèse nulle. La seconde classe exploite la nature univariée de l'hypothèse alternative. Les tests d'ajustement sont obtenus en comparant un estimateur à noyau de la densité spectrale des résidus et l'hypothèse nulle d'ajustement. Les distributions asymptotiques sont obtenues. Des simulations et une application avec des données financières sont présentées.

**Session 12C**      **Wednesday, June 2 • Mercredi 2 juin, 15:30 – 17:00**      **AA-1360**  
**Estimation**  
**Estimation**

[MS-217]

**Estimating the Parameters of Some Ecological Systems**

**Estimation des paramètres de certains systèmes écologiques**

Sévérien NKURUNZIZA & Sorana FRODA, *Université du Québec à Montréal*

We consider the system of ordinary differential equations of Lotka–Volterra. In the Froda–Colavita method, the evolution of predator-prey interactions is modeled as the trajectory of the deterministic system plus random error. In this procedure, the authors suppose that the random errors of prey and predator are independent. Moreover, they suppose that for each population the error is an iid process. We propose a flexible methodology, which takes into account the possible dependence between the errors. We suppose that the errors follow an Ornstein–Uhlenbeck process. That allows to improve the estimation and to have a better prediction model.

On s'intéresse à la modélisation stochastique et à l'estimation des paramètres du système d'équations différentielles de Lotka–Volterra. Dans la méthode proposée par Froda–Colavita, on modélise l'évolution du couple proie-prédateur observé comme étant la trajectoire du système déterministe à laquelle s'ajoute une erreur aléatoire. Dans ladite procédure, on suppose que les erreurs qui sont reliées à la proie et au prédateur sont indépendantes. D'autre part, on suppose que chacune de ces erreurs est un processus indépendant, identiquement distribué. Nous proposons une approche plus flexible qui tient compte de la dépendance entre les erreurs. Pour ce faire, nous supposons que ces erreurs suivent un processus d'Ornstein-Uhlenbeck. Ces hypothèses plus générales permettent d'améliorer l'estimation et de construire un meilleur modèle de prédiction.

[MS-218]

**Estimating a Bounded Location Parameter: Recent Results**

**Estimation d'un paramètre de position borné : résultats récents**

Éric MARCHAND, *University of New Brunswick*

For estimating a location parameter  $\theta$  of a symmetric density  $f(x - \theta)$ , where  $\theta$  is constrained to an interval  $[a, b]$ , we review some recent results giving simple conditions for which the Bayes estimator with respect to the uniform prior on the boundary  $\{a, b\}$ : (i) dominates the MLE under squared-

Nous présentons des résultats récents pour l'estimation, sous la perte quadratique, d'un paramètre de position  $\theta$  borné à un intervalle  $[a, b]$ . Pour des densités symétriques  $f(x - \theta)$ , nous décrivons des conditions simples pour lesquelles l'estimateur de Bayes par rapport à la loi *a priori* uniforme sur la frontière  $\{a, b\}$  : (i) domine l'estimateur à vraisemblance

error loss; (ii) is minimax under squared-error loss (work in (i) with François Perron, and work in (ii) with Bill Strawderman).

maximum ; (ii) est minimax (travail en (i) avec François Perron, et en (ii) avec Bill Strawderman).

### [MS-219]

#### Estimation of a cdf and its Derivative

#### Estimation de la fonction de répartition et de sa dérivée

Mohammed HADDOU & François PERRON, *Université de Montréal*

We consider the problem of estimating a cdf  $F$  by an estimator  $G$  based on a sample of size  $n$ . As  $n$  increases, we want  $G$  to be smooth on the regions where  $F$  is absolutely continuous and we want to obtain uniform convergence as well. Our approach uses splines with  $m$  nodes and the nodes depend on the sample. We show that our results hold if  $m$  tends to infinity as  $n$  tends to infinity. When  $F$  is absolutely continuous we show that the derivative of  $G$  tends to the derivative of  $F$  almost surely if  $m$  and  $n/m$  tend to infinity as  $n$  tends to infinity. We give some applications.

On considère le problème de l'estimation d'une f.r.  $F$  par un estimateur  $G$  basé sur un échantillon de taille  $n$ . On désire que  $G$  soit lisse dans les régions où  $F$  est absolument continue et on souhaite obtenir une convergence uniforme. Pour cela, nous utilisons des splines à  $m$  noeuds ( $m$  dépendant de  $n$ .) Nous montrons que l'on obtient les résultats souhaités si  $m$  et  $n$  tendent vers l'infini. Lorsque  $F$  est absolument continue, on montre que la dérivée de  $G$  tend vers la dérivée de  $F$  presque sûrement quand  $m$  et  $n/m$  tendent vers l'infini. Nous donnons quelques applications.

### [MS-220]

#### Construction of Optimizing Distributions with Applications in Estimation

#### Construction de distributions d'optimisation avec des applications en estimation

Saumen MANDAL, *University of Manitoba*

There are many problems in statistics, which require the calculation of one or more optimizing distributions or measures. Examples include design, estimation, image processing and stratified sampling problems. First, we cite some possible estimation problems and their properties in connection with optimal design theory. Then we consider the problem of determining maximum likelihood estimates under an hypothesis of marginal homogeneity for data in a square contingency table. This is an example of an optimization problem with respect to variables which should be nonnegative and satisfy several linear constraints. We consider a class of multiplicative algorithms for solving this problem.

Il y a beaucoup de problèmes en statistique qui nécessitent le calcul d'une ou plusieurs distributions ou mesures d'optimisation. Les exemples incluent les problèmes de design, d'estimation, de traitement d'image et d'échantillonnage stratifié. Premièrement, nous citons quelques problèmes d'estimation possibles et leurs propriétés en liaison avec la théorie de design optimal. Ensuite, nous considérons le problème de déterminer des estimations du maximum de vraisemblance sous une hypothèse d'homogénéité marginale pour des données d'un tableau de contingence carré. Ceci est un exemple d'un problème d'optimisation car les variables doivent être non négatives et doivent satisfaire plusieurs contraintes linéaires. Nous considérons une classe d'algorithmes multiplicatifs pour résoudre ce problème.

**[MS-221]****Multivariate Generalization of R-estimators Based on Data Depth****Généralisation multivariée des estimateurs R basée sur la profondeur des données**Shoja'eddin CHENOURI, *University of Waterloo*

The univariate R-estimators have been extensively explored. They were originally derived from rank-tests of hypothesis specifying the value of the center of a symmetric distribution. The univariate median and the Hodges–Lehmann estimator are two special cases in the class of R-estimators. In this talk, we shall develop multivariate versions of the R-estimators and the Hodges–Lehmann estimate using the concept of data depth. A data depth is a measure of how central a given point is with respect to a multivariate distribution or data cloud. We shall further discuss about robustness and the asymptotic distribution of the multivariate Hodges–Lehmann estimator.

Les estimateurs R univariés ont été grandement étudiés. À l'origine, ils ont été dérivés à partir de tests de rang pour des hypothèses sur le centre d'une distribution symétrique. La médiane univariée et l'estimateur Hodges-Lehmann sont deux cas spéciaux dans la classe des estimateurs R. Dans cette présentation, nous développerons des versions multivariées des estimateurs R et l'estimateur de Hodges-Lehmann en utilisant le concept de la profondeur des données. La profondeur des données est une mesure déterminant dans quelle mesure un point est central par rapport à une distribution multivariée ou par rapport à un nuage de points. De plus, nous discuterons de la robustesse et de la distribution asymptotique de l'estimateur multivarié de Hodges-Lehmann.

**[MS-222]****Comparisons of Estimation Procedures for Nonlinear Multilevel Models****Comparaisons des procédures d'estimation pour des modèles multi-niveaux non linéaires**Ali Reza FOTOUHI, *University College of the Fraser Valley*

We introduce General Multilevel Models and discuss the estimation procedures that may be used to fit multilevel mixed models. We apply the proposed procedures to three-level binary data generated in a simulation study. We compare the procedures by two criteria, bias and efficiency. We find that the estimates of the fixed effects and variance components are substantially and significantly biased using Longford's Approximation and Goldstein's Generalized Least Squares approaches by two software packages VARCL and ML3. We show that the Markov Chain Monte Carlo and Nonparametric Maximum Likelihood approaches perform much better than VARCL and ML3.

Nous présentons les modèles multi-niveaux généraux et discutons des procédures d'estimation qui peuvent être utilisées pour ajuster les modèles mixtes multi-niveaux. Nous appliquons les procédures proposées à des données binaires à trois niveaux générées par une étude de simulation. Nous comparons les procédures par deux critères, le biais et l'efficacité. Nous constatons que les estimés des effets fixes et des composantes de variance sont substantiellement et significativement biaisés en utilisant l'approximation de Longford et l'approche des moindres carrés généralisés de Goldstein par deux logiciels VARCL et ML3. Nous démontrons que les approches des chaînes de Markov Monte Carlo et du maximum de vraisemblance non paramétrique performant mieux que le VARCL et le ML3.

**Session 12D****Wednesday, June 2 • Mercredi 2 juin, 15:30 – 17:00****Z-317****Experimental Design and Related Issues****Planification d'expérience et sujets connexes**

**[MS-223]****Statistical Engineering: A New Algorithm for Reducing Variation in Manufacturing Processes**  
**L'ingénierie statistique : un nouvel algorithme pour réduire la variation dans les processus de**  
**manufacturation**Stefan STEINER & R. Jock MACKAY, *University of Waterloo*

Statistical Engineering is a new problem-solving algorithm and quality improvement system. The algorithm is useful for addressing chronic excess variation in high volume manufacturing. The algorithm is structured to be as proscriptive as possible focusing on “how to” solve problems.

The goal of this talk is to provide an overview of the Statistical Engineering algorithm. The focus of the seminar will be on the seven possible variation reduction approaches. Consideration of these different approaches drives the choices in the algorithm. The ideas will be illustrated with numerous examples from our consulting work.

L'ingénierie statistique est un nouvel algorithme de résolution de problèmes et un système d'amélioration de la qualité. L'algorithme est utile pour considérer la variation excessive chronique dans la fabrication à haut volume. L'algorithme est structuré de manière à être le plus proscriptif possible et se concentre sur la manière de résoudre des problèmes. Le but de cette présentation est de fournir une vue d'ensemble de l'algorithme d'ingénierie statistique. La conférence se concentre sur les sept approches possibles pour réduire la variation. Diverses considérations de ces différentes approches motivent les différents choix faits dans l'algorithme. Les idées seront illustrées avec de nombreux exemples de nos expériences en consultation.

**[MS-224]****Analysis of Factorial Experiments with Concurrent Measures Using Jackknife**  
**Analyse d'expériences à plan factoriel avec des mesures simultanées en utilisant le jackknife**Asokan MULAYATH VARIYATH, Bovas ABRAHAMN & Jiahua CHEN, *University of Waterloo*

Factorial experiments with concurrent measures as responses are common in industrial applications. The existing analysis methods often regard concurrent measures as sole response variables without replicates. Consequently, no degrees of freedom are left for error variance estimation in these methods. In reality, concurrent measures are obtained from replicated primary response variables. Precious information is hence wasted. In this paper, we suggest a jackknife based approach on the replicated primary responses to provide an estimate of error variance of concurrent measures. We compare the proposed method with two existing methods and investigate the consistency based on simulation studies.

Les expériences à plan factoriel avec des mesures simultanées comme réponses sont communes dans les applications industrielles. Les méthodes d'analyse existantes considèrent souvent les mesures simultanées comme des variables réponses uniques sans répliques. En conséquence, aucun degré de liberté n'est laissé pour l'estimation de la variance de l'erreur dans ces méthodes. En réalité, les mesures simultanées sont obtenues à partir des variables réponses primaires répliquées. De l'information précieuse est par conséquent gaspillée. Dans cet article, nous suggérons une approche basée sur le jackknife des réponses primaires répliquées pour fournir une estimation de la variance de l'erreur des mesures simultanées. Nous comparons la méthode proposée à deux méthodes existantes et étudions la convergence basée sur des études de simulation.

**[MS-225]****Eigenvalues in Optimal Design: Information Matrix vs Information Metric**  
**Les valeurs propres dans les plans optimaux : matrice d'information vs métrique d'information**Mikelis BICKIS, *University of Saskatchewan*

The statistical concept of covariance is mathematically equivalent to an inner product. While an inner product is conventionally represented by a sym-

La covariance, concept statistique, est équivalente d'un point de vue mathématique à un produit intérieur qu'on représente habituellement par une matrice symétrique. Il est plus na-

metric matrix, it is more natural to think of it as a metric. In the theory of optimal designs, traditional criteria for optimality are defined in terms of the eigenvalues of the information matrix. This matrix however, is more properly understood as a metric on the parameter space. The eigenvalues of the matrix do not have a natural meaning since they are not invariant under reparametrization. Consequently, eigenvalue-based criteria should be limited to models having natural parameters.

turel, pourtant, de le considérer comme une métrique. Dans la théorie des plans optimaux, les critères d'optimalité traditionnels sont basés sur les valeurs propres de la matrice d'information. Plus exactement, cette matrice induit une métrique sur l'espace des paramètres. Les valeurs propres de la matrice n'ont pas une signification naturelle car elles ne sont pas invariantes sous une reparamétrisation. En conséquence, on doit limiter l'utilisation de tels critères aux modèles dont les paramètres ont une interprétation naturelle.

#### [MS-226]

### Robust Designs for Approximate Generalized Linear Models

#### Designs robustes pour des modèles linéaires généralisés approximatifs

Adeniyi ADEWALE & Douglas P. WIENS, *University of Alberta*

In generalized linear model, the response,  $Y$ , follows a distribution from the exponential family of distributions. Mean response is usually a nonlinear function of the linear predictor,  $\eta = z'(x)\beta$ , termed the link function. Following the expository work of Box and Draper (1959) on designing for (linear) regression experiments, we construct designs with an eye on the approximate nature of the assumed generalized linear model. We obtain an expression for the asymptotic bias due to model misspecification. Rather than obtaining designs maximizing just the information matrix, we investigate design criteria that also incorporate the bias engendered by the model misspecification.

Dans le modèle linéaire généralisé, la réponse  $Y$  suit une distribution de la famille exponentielle. La réponse moyenne est habituellement une fonction non linéaire du prédicteur linéaire,  $\eta = z'(x)\beta$ , nommée la fonction de lien. Suivant le travail de Box et Draper (1959) sur la conception d'expériences de régression (linéaires), nous construisons des designs en gardant à l'oeil la nature approximative du modèle linéaire généralisé supposé. Nous obtenons une expression pour le biais asymptotique dû à la spécification erronée du modèle. Plutôt que d'obtenir un design qui maximise uniquement la matrice d'information, nous étudions des critères de plans d'expérience qui incorporent également le biais engendré par la spécification erronée du modèle.

#### [MS-227]

### Optimal Designs for Linear Haar-Wavelet Models

#### Designs optimaux pour les modèles linéaires d'ondelettes de Haar

Yongge TIAN & A. M. HERZBERG, *Queen's University*

Optimal experimental designs for linear Haar-wavelet models are presented in this talk. It is shown that the eigenvalues of the covariance matrix of the best linear unbiased estimate of the unknown parameters in a linear Haar-wavelet model can be represented in analytical form. From this result, the A-, D-, E-, G-, Q- and weighted G-optimal designs for the model can be constructed explicitly. The equivalences among these six types of optimal design are also presented.

Des designs expérimentaux optimaux pour les modèles linéaires d'ondelettes de Haar sont présentés dans cette présentation. Nous montrons que les valeurs propres de la matrice de covariance du meilleur estimateur linéaire sans biais des paramètres inconnus dans un modèle linéaire d'ondelettes de Haar peuvent être représentées sous forme analytique. De ce résultat, les design A-, D-, E-, G-, Q- et G pondérés optimaux pour le modèle peuvent être construits explicitement. Les équivalences parmi ces six types de designs optimaux sont également présentées.

**[MS-228]****Correction of an Erroneous Assertion of Margolin et al.****Correction d'une affirmation erronée de Margolin et co-auteurs**G. Arthur MIHRAM, *Princeton, NJ*, Danielle MIHRAM, *University of Southern California*

Margolin (& co-author) quite erroneously faulted [JASA 73:512, 1978] the Principium(-a) of Seeding [cf. SIMULATION. STATISTICAL FOUNDATIONS AND METHODOLOGY, Academic Press, 1972; then J. STATIST. COMPUT. SIMUL. 3:29, 1974], asserting that its application in designed experimentation with dynamic stochastic algorithmic models (i.e., simulations) would not measure “experimental error.” This paper again reveals that rigorous adherence to the Principia not only provides for the requisite experimental error but also ensures that measures of random-block effects can be secured in such experiments. We review statistical terms: random sample, experimental error, blocking, treatments, variance-components, ‘time series records’ in such similar experimentation.

Margolin (et co-auteurs) ont tout à fait manqué [JASA 73 :512, 1978] au principe de seeding [ cf. SIMULATION; STATISTICAL FOUNDATIONS & METHODOLOGY, Academic Press, 1972 ; puis J. STATIST. COMPUT. SIMUL. 3 :29, 1974], affirmant que son application dans des plans d'expériences avec des modèles algorithmiques stochastiques dynamiques (c.-à-d., simulations) ne mesurerait pas l'erreur expérimentale. Cet article indique encore que l'adhérence rigoureuse au Principe donne non seulement l'erreur expérimentale requise, mais s'assure également que les mesures d'effets de blocs aléatoires peuvent être fixées dans de telles expériences. Nous passons en revue les termes statistiques : échantillon aléatoire, erreur expérimentale, blocage, traitements, composantes de variance, séries chronologiques dans une telle expérimentation.

**Session 12E      Wednesday, June 2 • Mercredi 2 juin, 15:30 – 17:00      AA-1355****Various Topics****Thèmes divers****[MS-229]****Information Criteria and Change Point Problems****Critères d'information et problème du point de changement**Jianmin PAN & Jiahua CHEN, *University of Waterloo*

Information criteria are useful in model selection problems. Model complexity is a key in commonly used information criteria such as AIC and BIC. Existing methods lack some desirable properties when applied to the change point problem. A new criterion is proposed which modifies the existing ones. The new criterion is consistent and the resulting test statistic has a simple limiting distribution. We study the properties of change point's estimator under the alternative, and derive its best convergence rate and limiting distribution. Further, simulation results indicate that the new criterion has better power in detecting a change compared to the SIC.

Les critères d'information sont utiles dans les problèmes de sélection de modèles. La complexité du modèle est un élément clef dans les critères d'information habituellement utilisés tels que AIC et BIC. Les méthodes existantes ne satisfont pas certaines propriétés souhaitables une fois appliquées au problème du point de changement. Nous proposons un nouveau critère qui modifie ceux existants. Le nouveau critère est convergent et la statistique du test résultante a une distribution limite simple. Nous étudions les propriétés de l'estimateur du point de changement sous l'alternative et dérivons son meilleur taux de convergence et sa distribution limite. De plus, des résultats de simulations indiquent que le nouveau critère a une meilleure puissance pour détecter un changement que le SIC.

**[MS-230]****A Hierarchical Latent Class Model for Comparing Individual and Composite Tests  
Modèle hiérarchique de classe latente pour comparer des tests individuels et composés**Henrik STRYHN, *University of Prince Edward Island*, Keith E. BAPTISTE, *University of Liverpool*

Latent class models have become a popular tool to evaluate several diagnostic tests on the same samples, when no test is considered a gold standard. In this work, we developed hierarchical latent class models to compare test characteristics of culturing a cow's four quarter milk samples and composite (pooled) samples for penicillin-resistant staphylococci bacteria; the data comprised 1816 cows in 23 herds. Correlation of (quarter level) latent variables within cows was modeled by a beta-binomial and a Dirichlet-multinomial distribution, respectively, for dichotomous and trichotomous outcomes. Within-herd correlation was modeled by Gaussian random effects. The models were analyzed by likelihood-based methods.

Les modèles de classe latente sont devenus un outil populaire pour évaluer plusieurs tests de diagnostic sur les mêmes échantillons quand aucun test n'est considéré comme la référence. Dans ce travail, nous avons développé des modèles hiérarchiques de classe latente pour comparer des caractéristiques de tests de la culture de quatre échantillons de lait d'une vache et les échantillons composés (combiné) pour les bactéries de staphylocoques résistantes à la pénicilline ; les données comportent 1816 vaches dans 23 troupeaux. La corrélation des variables latentes (le niveau du quart) pour une vache est modélisée par une distribution bêta-binomiale et Dirichlet-multinomiale pour des résultats dichotomiques et trichotomiques respectivement. La corrélation à l'intérieur d'un troupeau est modélisée par des effets aléatoires gaussiens. Les modèles sont analysés par des méthodes de vraisemblance.

**[MS-231]****Descriptors of Large-scale Structures in Random Sets  
Descripteurs de la structure à grande échelle dans des ensembles aléatoires**Jeffrey PICKA, *University of New Brunswick*

Realizations of random sets have been analyzed through estimates of the moments of their realizations, or through statistics based on Minkowski measures. These statistics measure local features and uniformly average these measurements over an entire realization, ignoring large-scale structure. From models used in physics, new statistics determined by large scale structure can be developed. Laws of large numbers exist for these statistics, and random-walk-based estimates on many small samples can suggest which of these laws may be used to model the physical properties of large specimens of composite materials.

Les réalisations d'ensembles aléatoires ont été analysées par des estimations du moment de leurs réalisations ou par des statistiques basées sur la mesure de Minkowski. Ces statistiques mesurent les caractéristiques locales et font la moyenne de ces mesures uniformément sur une réalisation entière en ignorant la structure à grande échelle. À partir des modèles utilisés en physique, de nouvelles statistiques déterminées par la structure à grande échelle peuvent être développées. Des lois des grands nombres existent pour ces statistiques et les estimations basées sur les marches aléatoires sur plusieurs petits échantillons peuvent suggérer lesquelles de ces lois peuvent être utilisées pour modéliser les propriétés physiques de grands spécimens de matières composites.

**[MS-232]****Some Aspects of the First Principal Dimension of a Random Variable  
Quelques aspects de la première dimension principale d'une variable aléatoire**Carles CUADRAS & Daniel CUADRAS, *Universitat de Barcelona*

A random variable with continuous distribution can be expanded as a series of principal components of a related Bernoulli stochastic process. Some prop-

Une variable aléatoire avec une distribution continue peut être exprimée comme une série de composantes principales d'un processus stochastique de Bernoulli relié. Quelques

erties of the first principal component, which may characterize the distribution of the variable, and an inequality concerning a function and his derivative are obtained. The logistic distribution has special interest, as the first principal component is the cumulative distribution function. This property may justify the logit transformation. Stochastic dependence between variables is also studied.

propriétés de la première composante principale, qui peut caractériser la distribution de la variable, ainsi qu'une inégalité au sujet d'une fonction et sa dérivée sont obtenues. La distribution logistique est particulièrement intéressante car la première composante principale est sa fonction de répartition. Cette propriété peut justifier la transformation logit. La dépendance stochastique entre les variables est également étudiée.

## Session 12F      Wednesday, June 2 • Mercredi 2 juin, 15:30 – 17:00      Z-310

### Biostatistics III Biostatistique III

[MS-233]

#### **Nonparametric Estimation of Reliability for Semi-Markov Processus Based on $N$ Histories** **Estimation non paramétrique de la fiabilité de processus semi-markoviens basée sur $N$ histoires**

Nikolaos LIMNIOS, *Université de Technologie de Compiègne*, Brahim OUHBI, *École Nationale Supérieure d'Arts et Métiers, Meknès, Maroc*

We will present results on the estimation of ergodic finite semi-Markov processes from data observed by considering  $N$  independent censored sample paths, with application in survival and reliability analysis. Using the maximum likelihood function, nonparametric estimators are derived for: the hazard rate of transitions between states, the semi-Markov kernel, the Markov renewal matrix, the semi-Markov transition matrix, and the survival function. As the number of observed histories becomes large, strong consistency and asymptotic normality of the estimators are proved.

Nous présentons des résultats sur l'estimation de processus semi-markoviens finis ergodiques à partir d'observations de  $N$  trajectoires censurées indépendantes, avec des applications en analyse de survie et fiabilité. Nous déduisons de la fonction de maximum de vraisemblance des estimateurs non paramétriques pour le taux de hasard de transition entre états, le noyau semi-markovien, la matrice de renouvellement markovienne, la matrice de transition semi-markovienne et la fonction de survie. Nous montrons la convergence forte et la normalité asymptotique de ces estimateurs lorsque le nombre d'histoires tend vers l'infini.

[MS-234]

#### **Parametric and Nonparametric Estimation of Current Status Data with a Nonsusceptible Fraction** **Estimation paramétrique et non paramétrique de données d'état actuel avec une fraction non susceptible**

Jin QIAN & Richard COOK, *University of Waterloo*

We discuss the design and analysis of some recently completed surgical trials and the information available to compare seroconversion rates when blood tests are only conducted at the time of discharge. This type of design provides current status data on the times of seroconversion, but interest lies in whether or not patients will seroconvert. We describe an EM algorithm to estimate the "non-susceptible" fraction representing patients who will not seroconvert, as well as a nonparametric estimation of the distribution of the time to seroconver-

Nous discutons du plan et de l'analyse de quelques essais chirurgicaux récemment réalisés et de l'information disponible pour comparer des taux de séroconversion quand les analyses sanguines sont uniquement effectuées au moment du congé de l'hôpital. Ce type de plan fournit des données sur l'état actuel des temps de séroconversion, mais l'intérêt consiste à savoir si les patients seront séroconvertis. Nous décrivons un algorithme EM pour estimer la fraction "non susceptible" représentant les patients qui ne seront pas séroconvertis, ainsi qu'une estimation non paramétrique de la distribution du temps à la séroconversion. Des graphiques de



sion. Likelihood plots are also examined to gain insight into the nature and extent of information available.

vraisemblance sont également examinés afin de mieux comprendre la nature et l'ampleur de l'information disponible.

**[MS-235]**

**Testing for Ordered Group Effects in Bivariate Dichotomous and Continuous Outcomes  
Tester l'effet d'ordre de groupe dans des résultats bivariés dichotomiques et continus**

Chul PARK & Patrick J. FARREL, *Carleton University*

An ordering effect on bivariate responses consisting of binary and continuous variables is studied using conditional logistic models. These models are formulated for several treatment groups that reflect the ordering effect through the different dosage levels. We first develop iterative procedures for obtaining maximum likelihood estimates constrained by hypotheses, and then derive a likelihood ratio test for detecting the ordering effect. This testing procedure is applied to a real data set to investigate the effect of a color additive on the development of RE tumors. Differences between joint and marginal analyses are also illustrated.

Un effet de l'ordre sur des réponses bivariées combinant des variables binaires et continues est étudié en utilisant les modèles logistiques conditionnels. Ces modèles sont formulés pour plusieurs groupes de traitement qui reflètent l'effet d'ordre à travers les différents niveaux de dosage. Nous développons tout d'abord des procédures itératives pour obtenir des estimations du maximum de vraisemblance contraints par des hypothèses, et dérivons ensuite un test du rapport de vraisemblance pour détecter l'effet d'ordre. Cette méthode de test est appliquée à un vrai jeu de données pour étudier l'effet d'un additif colorant sur le développement des tumeurs RE. Des différences entre les analyses conjointes et marginales sont également illustrées.

**[MS-236]**

**Bootstrap Method for Classification Using Random Effects Model  
Méthode bootstrap pour la classification en utilisant des modèles à effets aléatoires**

Narasimha PRASAD, *University of Alberta*

Often in clinical studies, measurements are taken on the same individuals over time for diagnostic purposes. In such situations epidemiologists use random effects models to estimate the expected number of measurements exceeding a pre-specified threshold value and also the probability of no exceedance. Recently, random effects models are also used in classifying a subject either to "diseased" or "not" given the results of repeated diagnostic test when no gold standard test is available. In this talk I will consider bootstrap methods to evaluate predictors as diagnostic tools for classification and to study properties of diagnostic tests through sensitivity and specificity parameters.

Souvent dans les études cliniques, des mesures sont prises sur les mêmes individus dans le temps pour des fins diagnostiques. Dans de telles situations, les épidémiologistes utilisent des modèles à effets aléatoires pour estimer le nombre espéré de mesures qui excèdent une valeur-seuil pré-spécifiée et également la probabilité de ne pas l'excéder. Récemment, les modèles à effets aléatoires sont également utilisés pour classifier un sujet comme étant malade ou non, à partir des résultats de tests de diagnostic répétés quand aucun test standard bien établi est disponible. Dans cette présentation, je considérerai les méthodes bootstrap pour évaluer des prédicteurs en tant qu'outils de diagnostic pour la classification et pour étudier les propriétés des tests de diagnostic à travers les paramètres de sensibilité et de spécificité.

# Author Index • Index des auteurs

ABDOUS, Belkacem	84	BROWN, James	107
ABRAHAMN, Bovas	157	BROWN, K. S.	103
ADES, A. E.	150	BULL, Shelley	94
ADEWALE, Adeniyi	158	BUREAU, Alexandre	113
ALIMADAD, Azadeh	137	CACHIA, A.	106
ALLDREDGE, Mathew	79	CAKMAK, Sabit	111
ALLEN, Brian	112	CAMACHO, Fernando	58
ALLEN, William	124	CAMPBELL, David	82
ALODAT, Moh'd	104	CAMPBELL, M. Karen	141
ALPARGU, Gülhan	142	CARRIÈRE, Keumhee Chough	149
ALTMAN, Rachel MacKay	145	CHAMBERS, L.	122
AMATO, U.	98	CHAMBERS, Raymond	76
AMISS, Scott	65	CHANG, Howard	111
ANGERS, Jean-François	132	CHAPMAN, Judy-Anne	59, 135
ANTONIADIS, Anestis	98	CHAREKA, Patrick	80
APPADOO, Srimantoorao	64	CHARLAND, Kim	87
ARNAB, Raghunath	119	CHARRON, Dominique	112
BABINEAU, Denise	83	CHAUBEY, Yogendra	82, 148
BADESCU, Alex	72	CHEN, Edward	100
BAKAL, Jeffrey	153	CHEN, Gemai	63, 109
BALAN, Raluca	152	CHEN, Jiahua	93, 134, 157, 159
BALAN, Raluca M.	75	CHEN, Min	109
BALASOORIYA, Uditha	83	CHEN, Sanping	109
BAPTISTE, Keith E.	160	CHEN, Zengjing	72
BARNES, Thomas G.	146	CHENOURI, Shoja'eddin	156
BÉDARD, Mylène	65	CHOI, Yun Hee	84
BÉLAND, Yves	101	CHU, Ka Lok	138
BELLAVANCE, François	118	COLLINS, D. L.	106
BELLEY, C.	122	CONNOLLY, Jennifer	67
BELLHOUSE, David	60	COOK, Richard	131, 144, 161
BENHIN, Emmanuel	97	COOK, Richard J.	74
BICKIS, Mikelis	157	COREY, Paul	103
BINDER, David	76	CORTES, Corinna	105
BINGHAM, Derek	60	CUADRAS, Carles	160
BINYAVANGA, Kamanzi	92	CUADRAS, Daniel	160
BOHER, Jean-Marie	131	DABROWSKI, André	93
BONNARDEAUX, A.	141	DADUNA, Hans	128
BOROVKOV, K.	89	DALES, Robert	111
BOUHADDIoui, Chafik	81	DAVIES, Katherine	83
BRADLOW, Eric T.	103	DE ANGELIS, D.	150
BRANT, Rollin	78	DE LA FUENTE, Marta	66
BRILLINGER, David	61	DE LA FUENTE, Soledad	66
BRIOLLAIS, L.	94	DE LEON, Alexander	104
BRISEBOIS, François	101	DENG, Dianliang	92

- DETEMPLE, J. .... 99  
 DEVLIN, B. .... 66  
 DI CRESCENZO, Antonio .... 89  
 DING, Keyue .... 148  
 DOLOVICH, L. .... 122  
 DOUGLAS, George .... 124  
 DU, Meng .... 71  
 DUCHESNE, Pierre .... 119, 153  
 DUFOUR, Jean-Marie .... 80, 81  
 DUGAS, Charles .... 77  
 DUNKLE, Kristin .... 102  
 DUPUIS, Debbie .... 138  
 DURRETT, Rick .... 70  
 DUTILLEUL, Pierre .... 142, 143  
  
 EL ADLOUNI, Salaheddine .... 133  
 EL BIDANI, Omar .... 137  
 ELIAS, Diana .... 66  
 EMRICH, K. .... 143  
 EVANS, Alan C. .... 105, 106  
 EVRA, Rose Carline .... 137  
  
 FAN, Jianqing .... 108  
 FANG, Miao .... 71  
 FARHAT, Abdeljelil .... 80  
 FARREL, Patrick J. .... 162  
 FAVRE, Anne-Catherine .... 133  
 FENG, Zeny .... 93  
 FERNANDEZ, Favia .... 66  
 FILUS, Jerzy .... 128  
 FILUS, Lidia Z. .... 128  
 FINE, Jason .... 117  
 FLEURY, Manon .... 112  
 FORTIN, Vincent .... 132  
 FOSTER, G. .... 149  
 FOTOUHI, Ali Reza .... 75, 156  
 FRANSSSEN, Edmee .... 71  
 FRASER, D. A. S. .... 151  
 FREDETTE, Marc .... 130  
 FRODA, Sorana .... 154  
 FRYZLEWICZ, Piotr .... 98  
  
 GAEBEL, Kathy .... 140  
 GARCIA, R. .... 99  
 GARRIDO, José .... 76  
 GATARAYIHA, Jean-Philippe .... 137  
 GAUDET, Jocelyn .... 132  
 GENEST, Christian .... 153  
 GENTIL, Ivan .... 100  
 GENTLEMAN, Robert .... 130  
  
 GILL, O. N. .... 150  
 GILLIS, Dan .... 112  
 GILMOUR, Graeme .... 88  
 GIORNO, V. .... 89  
 GIRARDIN, Valérie .... 127  
 GOLDSTEIN, Harvey .... 107  
 GOUBAR, Aicha .... 150  
 GOYAL, S. K. .... 82  
 GREEN, J. .... 94  
 GREENWOOD, Celia .... 94, 96  
 GRENIER, Marco .... 101  
 GRIFFITHS, Robert .... 70  
 GU, Hong .... 82  
 GUO, Meijie .... 71  
 GUO, Tong .... 86  
 GUPTA, Arvind .... 151  
 GUPTA, Sat .... 110  
 GUSTAFSON, Paul G. .... 137, 150  
 GUTMANIS, Iris A. .... 141  
  
 HADDOU, Mohammed .... 155  
 HAMM, Jeremy .... 62  
 HANLEY, James .... 115  
 HARDY, Jean .... 118  
 HARMS, Torsten .... 119  
 HASTIE, Trevor .... 125  
 HE, W. .... 94, 135  
 HERMANUSSEN, Michael .... 115  
 HERNÁN, Miguel .... 87  
 HERZBERG, A. M. .... 158  
 HOGG, W. .... 122  
 HOLBROOK, Anne .... 140  
 HOLT, John .... 112  
 HORROCKS, Julie .... 102  
 HOSSAIN, Shahadut .... 137  
 HSU, Ming-Chun .... 121  
 HU, Feifang .... 85  
 HUANG, Fuping .... 139  
 HUANG, Jie .... 121  
 HUNT, Daniel .... 64  
 HUYER, Sophia .... 68  
  
 IORIO, Maria De .... 70  
 ISHWARAN, H. .... 73  
 ISMAILA, Afisi .... 71  
 ISOTALO, J. .... 138  
  
 JEFFERYS, Thomas R. .... 146  
 JEFFERYS, William .... 146  
 JIANG, Jiancheng .... 108

- JIN, Zi ..... 63  
 JOHNSON, F. R. .... 149  
 JOSHI, Heather ..... 107
- KACZOROWSKI, J. .... 122  
 KALLIS, Constantinos ..... 107  
 KAMANO, Saori ..... 102  
 KARWALAJTYS, T. .... 122  
 KHAN, Bashir ..... 91  
 KHAN, Shahjahan ..... 90  
 KIM, Hyang Mi ..... 91  
 KIM, Peter ..... 146  
 KOO, Malcolm ..... 103  
 KOPCIUK, Karen ..... 94  
 KOVAL, John J. .... 141  
 KULIK, Rafal ..... 128  
 KULPERGER, Reg ..... 72  
 KUSTRA, Rafal ..... 125
- LABBE, Aurélie ..... 147  
 LABILLOIS, Tony ..... 127  
 LAFAYE DE MICHEAUX, Pierre ..... 91  
 LAFLAMME, G. .... 100  
 LAFORTUNE, Yves ..... 110  
 LAKHAL CHAIEB, Lajmi ..... 84  
 LAN, Yongxu ..... 71  
 LAROCHE, Danielle ..... 88  
 LAROCQUE, Denis ..... 152  
 LAWLESS, Jerry ..... 130, 135  
 LE, Nhu ..... 85, 111  
 LEE, Charles Chu-In ..... 122  
 LEE, Judy ..... 127  
 LÉGER, Christian ..... 91, 151  
 LERCH, Jason ..... 105  
 LEROUX, Brian ..... 145  
 LEROUX, Diane ..... 126  
 LEVITT, C. .... 122  
 LEVY, A. R. .... 150  
 LI, Haitao ..... 73  
 LI, Xiaoqiang ..... 139  
 LIANG, Yuanyuan ..... 149  
 LIEVESLEY, Denise ..... 67  
 LIMNIOS, Nikolaos ..... 127, 161  
 LINDEYER, J. .... 100  
 LIU, Jinnan ..... 71  
 LIU, Lin ..... 122  
 LIU, Xuecheng ..... 130  
 LIU, Zhong ..... 111  
 LOCKHART, Richard ..... 147  
 LOEPPKY, Jason ..... 60
- LOREDO-OSTI, J. .... 141  
 LOW, C.-K. .... 83  
 LU, Wilson ..... 120  
 LUNAZZI, Helena ..... 66
- MA, Renjun ..... 107, 114  
 MAAROUF, Abdel ..... 112  
 MACKAY, R. Jock ..... 157  
 MADISE, Nyovani ..... 107  
 MADRAS, Neal ..... 100  
 MANDAL, Saumen ..... 155  
 MANGIN, Jean-François ..... 106  
 MANUEL, Douglas ..... 141  
 MARCHAND, Éric ..... 154  
 MARSHALL, D. A. .... 149  
 MARSHALL, J. K. .... 149  
 MATARISE, Florance ..... 80  
 MATTHEWS, David E. .... 84  
 MCAULEY, Kim ..... 145  
 MCCANDLESS, Lawrence ..... 150  
 MCCLEAN, Kathy ..... 87  
 MCCLAUGHLIN, J. .... 94  
 MCLEOD, Ian ..... 65  
 MCNALLY, Cathlin ..... 140  
 MCQUIGGE, Murray ..... 124  
 MCRAE, Ken ..... 58  
 MEIGEN, Christof ..... 115  
 MÉLARD, Guy ..... 81  
 MERLEAU, James ..... 132  
 MIHRAM, Danielle ..... 159  
 MIHRAM, G. Arthur ..... 159  
 MILLS-FLEMMING, Joanna ..... 138  
 M'LAN, Cyr Emile ..... 94  
 MOGHTADERI, Azadeh ..... 62  
 MONETTE, Manon ..... 87  
 MOORE, Ian ..... 63  
 MOORHEAD, Philip ..... 61  
 MORGAN, K. .... 141  
 MULAYATH VARIYATH, Asokan ..... 157  
 MYSAK, Lawrence ..... 123
- NASON, Guy ..... 98  
 NAWA, Victor ..... 103  
 NEBEBE, F. .... 82  
 NEUPANE, Binod ..... 71  
 NEWCOMBE-WELCH, Patricia ..... 114  
 NG, Peggy ..... 136  
 NIELSEN, Rasmus ..... 70  
 NKUNDUMUKIZA, Eraste ..... 71  
 NKURUNZIZA, Sévérien ..... 154

- NOBILE, A. G. . . . . 89  
NOVIKOV, Alexander . . . . . 89
- O'BRIEN, B. J. . . . . 149  
OGDEN, R. . . . . 99  
O'ROURKE, K. . . . . 122  
OUHBI, Brahim . . . . . 161  
OVERDUIN, Steve . . . . . 147
- PACURAR, Maria . . . . . 153  
PAN, Jianmin . . . . . 159  
PANTEL, Martin . . . . . 110  
PARK, Chul . . . . . 162  
PARK, Jeong-Gun . . . . . 144  
PARKHOMENKO, E. . . . . 94  
PELLETIER, Bernard . . . . . 143  
PELLETIER, Caroline . . . . . 110  
PELLETIER, Éric . . . . . 126  
PENG, Jianan . . . . . 122  
PENSKY, M. . . . . 98  
PERREAULT, Luc . . . . . 132, 133  
PERRON, François . . . . . 155  
PFEFFERMANN, Danny . . . . . 97  
PHAM-GIA, Thu . . . . . 138  
PHILLIPS, K. A. . . . . 149  
PICKA, Jeffrey . . . . . 160  
PIEPHO, Hans-Peter . . . . . 143  
PIERRE, Fritz . . . . . 101  
PLANTE, Jean-François . . . . . 153  
PLATT, Robert . . . . . 74, 86  
POLLOCK, Kenneth . . . . . 79  
POTTER, Elizabeth . . . . . 141  
PRADEL, Roger . . . . . 79  
PRASAD, Narasimha . . . . . 162  
PREGIBON, Daryl . . . . . 105  
PUNTANEN, S. . . . . 138
- QIAN, Jin . . . . . 161
- RAI, Shesh . . . . . 64  
RAMSAY, James . . . . . 116  
RAMSAY, Timothy . . . . . 124  
RANGREJ, Jagadish . . . . . 123  
RAO, J. N. K. . . . . 96, 97, 133  
REGIS, J. . . . . 106  
REID, Nancy . . . . . 148  
RÉMILLARD, Bruno . . . . . 100  
REN, Jian-Jian . . . . . 116  
RHODES, Anne . . . . . 142  
RICCIARDI, L. M. . . . . 89
- RINALDO, A. . . . . 66  
RINDISBACHER, Marcel . . . . . 99  
RIVEST, Louis-Paul . . . . . 84  
RIVIÈRE, D. . . . . 106  
RODRIGUE, Natalie . . . . . 118  
ROEDER, Kathryn . . . . . 66  
ROSLIN,, N. . . . . 141  
ROSS, William . . . . . 123  
ROUTLEDGE, Rick . . . . . 77  
ROY, Jean-Denis . . . . . 77  
ROY, Laurent . . . . . 87  
ROY, Roch . . . . . 81  
ROZGA, Ante . . . . . 102  
RUDNISKY, C. . . . . 104  
RYTINA, Steven . . . . . 113
- SAIDI, Abdessamad . . . . . 81  
SAINT-PIERRE, Richard . . . . . 69  
SAINT-PIERRE, Yves . . . . . 88  
SALEH, A. K. Md. E. . . . . 91  
SALIBIAN-BARRERA, Matias . . . . . 137  
SALISBURY, Thomas . . . . . 151  
SANCHEZ LEZAMA, Ana Paola . . . . . 71  
SARKAR, Joykrishna . . . . . 129  
SAUER, Cornelia . . . . . 128  
SAULNIER, Michel . . . . . 69  
SCHIOPU-KRATINA, Ioana . . . . . 75  
SCHIPPER, Stefan . . . . . 108  
SCHMID, Wolfgang . . . . . 108  
SCHWARZ, Carl . . . . . 58, 62, 79  
SEN, Arusharka . . . . . 129  
SEN, Debaraj . . . . . 148  
SHABBIR, Javid . . . . . 110  
SHI, Peilin . . . . . 60  
SIMON, Richard . . . . . 96  
SIMONS, Ted . . . . . 79  
SINGH, G. . . . . 104  
SINGH, Sarjinder . . . . . 120  
SINGHAL, Sunil . . . . . 110  
SITTER, Randy R. . . . . 60, 120  
SKORUPSKI, William P. . . . . 103  
SLOBODIANIK, Nikolai . . . . . 71  
SMITH, J. Terry . . . . . 153  
SMITH, Michael . . . . . 113  
SNEDDON, Gary . . . . . 93  
SPEECHLEY, Kathy N. . . . . 141  
STAFFORD, Jamie . . . . . 151  
STAIUCU, Ana-Maria . . . . . 148  
STEEL, David . . . . . 107

- STEEL, Sarel J. . . . . 92  
 STEELE, Fiona . . . . . 107  
 STEINER, Stefan . . . . . 157  
 STEPHENS, Michael A. . . . . 147  
 STEVENSON, Joan . . . . . 153  
 STRYHN, Henrik . . . . . 160  
 STYAN, G. P. H. . . . . 138  
 SUN, Tao . . . . . 71  
 SUN, Ye . . . . . 71  
 SUTRADHAR, Brajendra C. . . . . 93  
 SUTRADHAR, Rinku . . . . . 131  
 SVEINSSON, Oli . . . . . 132  
 SZEKLI, Ryszard . . . . . 128  
  
 TAI, Patricia . . . . . 135  
 TAKAHARA, Glen . . . . . 62, 153  
 TAM, Raymond . . . . . 103  
 TAN, Ye . . . . . 74  
 TAYLOR, Jonathan . . . . . 125  
 THABANE, Lehana . . . . . 122, 140, 149  
 THAVANESWARAN, A. . . . . 64  
 THOMPSON, Mary . . . . . 95  
 THOMPSON, Mary E. . . . . 93, 147  
 THOMPSON, Paul . . . . . 106  
 THOMSON, David . . . . . 62, 129  
 THOMSON, David J. . . . . 63, 65  
 THORNTON, Bill . . . . . 110  
 TIAN, Yongge . . . . . 158  
 TIBSHIRANI, Rob . . . . . 78, 125  
 TOGA, Arthur W. . . . . 106  
 TOIBER TEMIN, Noemi . . . . . 71  
 TRABELSI, Abdelwahed . . . . . 80  
 TRUMMER, Manfred . . . . . 151  
 TSAO, Min . . . . . 117  
 TURKKAN, N. . . . . 138  
 TURNER, Elizabeth . . . . . 140  
 TURNER, Rolf . . . . . 80  
  
 URRUTIA, María Inés . . . . . 66  
  
 VANDAL, Alain C. . . . . 130, 140  
 VAN EEDEN, Constance . . . . . 86  
 VÉGIARD, Sylvain . . . . . 57, 68  
 VOLINSKY, Chris . . . . . 105  
 VOLODIN, Andrei . . . . . 129  
  
 WAINER, Howard . . . . . 103  
 WALTHER, Guenther . . . . . 125  
 WANG, Jing . . . . . 71  
 WANG, Liqun . . . . . 90  
  
 WANG, Rui . . . . . 71  
 WANG, Shufang . . . . . 71  
 WANG, Xiaohui . . . . . 103  
 WANG, Xikui . . . . . 71, 72  
 WANG, Yan . . . . . 72  
 WANG, Zilin . . . . . 60  
 WASSERMAN, L. . . . . 66  
 WELLS, Martin . . . . . 73  
 WHITE, Andrew . . . . . 103  
 WIENS, Douglas P. . . . . 158  
 WILLIAMS, Trish . . . . . 67  
 WILLIAMSON, Ian . . . . . 62  
 WILLMS, Jon Douglas . . . . . 114  
 WORSLEY, Keith . . . . . 59  
 WU, Changbao . . . . . 134  
 WU, Hulin . . . . . 144  
 WU, Lang . . . . . 75  
  
 XIE, Hansheng . . . . . 109  
  
 YAN, Joel . . . . . 88  
 YAUK, Carole . . . . . 124  
 YE, Jane . . . . . 60  
 YI, Yanqing . . . . . 71  
 YIPING, Dou . . . . . 111  
 YORK, Thomas . . . . . 70  
 YOSHIHAMA, Mieko . . . . . 102  
 YOU, Yong . . . . . 120  
 YU, Long . . . . . 73  
 YUAN, Xiaobin . . . . . 151  
 YUEN, John Wai Kong . . . . . 100  
  
 ZAREPOUR, Mahmoud . . . . . 73  
 ZENG, Leilei . . . . . 74, 144  
 ZHANG, Jie . . . . . 121  
 ZHANG, Lingyun . . . . . 63  
 ZHANG, Ying . . . . . 65  
 ZHOU, Julie . . . . . 60  
 ZHOU, Xian . . . . . 71  
 ZHU, Mu . . . . . 139  
 ZIDEK, Jim . . . . . 85, 111  
 ZIJDENBOS, Alex . . . . . 105  
 ZOU, Fei . . . . . 117

# Author Index (per session)

## Index des auteurs (par session)

ABDOUS, Belkacem .....	3F	BRIOLLAIS, L. ....	4F
ABRAHAMN, Bovas .....	12D	BRISEBOIS, François .....	6E
ADES, A. E. ....	11F	BROWN, James .....	7D
ADEWALE, Adeniyi .....	12D	BROWN, K. S. ....	6F
ALIMADAD, Azadeh .....	10E	BULL, Shelley .....	4F
ALLDREDGE, Mathew .....	3D	BUREAU, Alexandre .....	7F
ALLEN, Brian .....	7F		
ALLEN, William .....	9B	CACHIA, A. ....	7B
ALODAT, Moh'd .....	6F	CAKMAK, Sabit .....	7F
ALPARGU, Gülhan .....	11B	CAMACHO, Fernando .....	0A
ALTMAN, Rachel MacKay .....	11C	CAMPBELL, David .....	3E
AMATO, U. ....	6C	CAMPBELL, M. Karen .....	11A
AMISS, Scott .....	0B	CARRIÈRE, Keumhee Chough .....	11F
ANGERS, Jean-François .....	10A	CHAMBERS, L. ....	8F
ANTONIADIS, Anestis .....	6C	CHAMBERS, Raymond .....	3A
APPADOO, Srimantoorao .....	0B	CHANG, Howard .....	7F
ARNAB, Raghunath .....	8E	CHAPMAN, Judy-Anne .....	0A, 10C
		CHAREKA, Patrick .....	3E
BABINEAU, Denise .....	3F	CHARLAND, Kim .....	4C
BADESCU, Alex .....	2E	CHARRON, Dominique .....	7F
BAKAL, Jeffrey .....	12B	CHAUBEY, Yogendra .....	3E, 11E
BALAN, Raluca .....	12B	CHEN, Edward .....	6E
BALAN, Raluca M. ....	2F	CHEN, Gemai .....	0B, 7D
BALASOORIYA, Uditha .....	3F	CHEN, Jiahua .....	4F, 10B, 12D, 12E
BAPTISTE, Keith E. ....	12E	CHEN, Min .....	7D
BARNES, Thomas G. ....	11D	CHEN, Sanping .....	7E
BÉDARD, Mylène .....	0B	CHEN, Zengjing .....	2E
BÉLAND, Yves .....	6E	CHENOURI, Shoja'eddin .....	12C
BELLAVANCE, François .....	8D	CHOI, Yun Hee .....	3F
BELLEY, C. ....	8F	CHU, Ka Lok .....	10E
BELLHOUSE, David .....	0B	COLLINS, D. L. ....	7B
BENHIN, Emmanuel .....	6B	CONNOLLY, Jennifer .....	2A
BICKIS, Mikelis .....	12D	COOK, Richard .....	9F, 11C, 12F
BINDER, David .....	3A	COOK, Richard J. ....	2F
BINGHAM, Derek .....	0B	COREY, Paul .....	6F
BINYAVANGA, Kamanzi .....	4E	CORTES, Corinna .....	7A
BOHER, Jean-Marie .....	9F	CUADRAS, Carles .....	12E
BONNARDEAUX, A. ....	10F	CUADRAS, Daniel .....	12E
BOROVKOV, K. ....	4D		
BOUHADDIOUI, Chafik .....	3E	DABROWSKI, André .....	4F
BRADLOW, Eric T. ....	6F	DADUNA, Hans .....	9E
BRANT, Rollin .....	3C	DALES, Robert .....	7F
BRILLINGER, David .....	0B	DAVIES, Katherine .....	3F

- DE ANGELIS, D. .... 11F  
 DE LA FUENTE, Marta ..... 0B  
 DE LA FUENTE, Soledad ..... 0B  
 DE LEON, Alexander ..... 6F  
 DENG, Dianliang ..... 4E  
 DETEMPLE, J. .... 6D  
 DEVLIN, B. .... 1  
 DI CRESCENZO, Antonio ..... 4D  
 DING, Keyue ..... 11E  
 DOLOVICH, L. .... 8F  
 DOUGLAS, George ..... 9C  
 DU, Meng ..... 2D  
 DUCHESNE, Pierre ..... 8E, 12B  
 DUFOUR, Jean-Marie ..... 3E  
 DUGAS, Charles ..... 3B  
 DUNKLE, Kristin ..... 6E  
 DUPUIS, Debbie ..... 10E  
 DURRETT, Rick ..... 2C  
 DUTILLEUL, Pierre ..... 11B  
  
 EL ADLOUNI, Salaheddine ..... 10A  
 EL BIDANI, Omar ..... 10D  
 ELIAS, Diana ..... 0B  
 EMRICH, K. .... 11B  
 EVANS, Alan C. .... 7B  
 EVRA, Rose Carline ..... 10D  
  
 FAN, Jianqing ..... 7D  
 FANG, Miao ..... 2D  
 FARHAT, Abdeljelil ..... 3E  
 FARREL, Patrick J. .... 12F  
 FAVRE, Anne-Catherine ..... 10A  
 FENG, Zeny ..... 4F  
 FERNANDEZ, Favia ..... 0B  
 FILUS, Jerzy ..... 9E  
 FILUS, Lidia Z. .... 9E  
 FINE, Jason ..... 8C  
 FLEURY, Manon ..... 7F  
 FORTIN, Vincent ..... 10A  
 FOSTER, G. .... 11F  
 FOTOUHI, Ali Reza ..... 2F, 12C  
 FRANSSEN, Edmee ..... 2D  
 FRASER, D. A. S. .... 12B  
 FREDETTE, Marc ..... 9F  
 FRODA, Sorana ..... 12C  
 FRYZLEWICZ, Piotr ..... 6C  
  
 GAEBEL, Kathy ..... 10F  
 GARCIA, R. .... 6D  
 GARRIDO, José ..... 3B  
  
 GATARAYIHA, Jean-Philippe ..... 10D  
 GAUDET, Jocelyn ..... 10A  
 GENEST, Christian ..... 12B  
 GENTIL, Ivan ..... 6D  
 GENTLEMAN, Robert ..... 9F  
 GILL, O. N. .... 11F  
 GILLIS, Dan ..... 7F  
 GILMOUR, Graeme ..... 4C  
 GIORNO, V. .... 4D  
 GIRARDIN, Valérie ..... 9E  
 GOLDSTEIN, Harvey ..... 7D  
 GOUBAR, Aicha ..... 11F  
 GOYAL, S. K. .... 3E  
 GREEN, J. .... 4F  
 GREENWOOD, Celia ..... 4F, 6A  
 GRENIER, Marco ..... 6E  
 GRIFFITHS, Robert ..... 2C  
 GU, Hong ..... 3E  
 GUO, Meijie ..... 2D  
 GUO, Tong ..... 4B  
 GUPTA, Arvind ..... 11G  
 GUPTA, Sat ..... 7E  
 GUSTAFSON, Paul G. .... 10E, 11F  
 GUTMANIS, Iris A. .... 11A  
  
 HADDOU, Mohammed ..... 12C  
 HAMM, Jeremy ..... 0B  
 HANLEY, James ..... 8B  
 HARDY, Jean ..... 8D  
 HARMS, Torsten ..... 8E  
 HASTIE, Trevor ..... 9C  
 HE, W. .... 4F, 10C  
 HERMANUSSEN, Michael ..... 8B  
 HERNÁN, Miguel ..... 4B  
 HERZBERG, A. M. .... 12D  
 HOGG, W. .... 8F  
 HOLBROOK, Anne ..... 10F  
 HOLT, John ..... 7F  
 HORROCKS, Julie ..... 6E  
 HOSSAIN, Shahadut ..... 10E  
 HSU, Ming-Chun ..... 8F  
 HU, Feifang ..... 4A  
 HUANG, Fuping ..... 10F  
 HUANG, Jie ..... 8F  
 HUNT, Daniel ..... 0B  
 HUYER, Sophia ..... 2A  
  
 IORIO, Maria De ..... 2C  
 ISHWARAN, H. .... 2E  
 ISMAILA, Afisi ..... 2D



- ISOTALO, J. .... 10E
- JEFFERYS, Thomas R. .... 11D
- JEFFERYS, William .... 11D
- JIANG, Jiancheng .... 7D
- JIN, Zi .... 0B
- JOHNSON, F. R. .... 11F
- JOSHI, Heather .... 7D
- KACZOROWSKI, J. .... 8F
- KALLIS, Constantinos .... 7D
- KAMANO, Saori .... 6E
- KARWALAJTYS, T. .... 8F
- KHAN, Bashir .... 4E
- KHAN, Shahjahan .... 4E
- KIM, Hyang Mi .... 4E
- KIM, Peter .... 11D
- KOO, Malcolm .... 6F
- KOPCIUK, Karen .... 4F
- KOVAL, John J. .... 11A
- KULIK, Rafal .... 9E
- KULPERGER, Reg .... 2E
- KUSTRA, Rafal .... 9C
- LABBE, Aurélie .... 11E
- LABILLOIS, Tony .... 9D
- LAFAYE DE MICHEAUX, Pierre .... 4E
- LAFLAMME, G. .... 6E
- LAFORTUNE, Yves .... 7E
- LAKHAL CHAIEB, Lajmi .... 3F
- LAN, Yongxu .... 2D
- LAROCHE, Danielle .... 4C
- LAROCQUE, Denis .... 12B
- LAWLESS, Jerry .... 9F, 10C
- LE, Nhu .... 4A, 7F
- LEE, Charles Chu-In .... 8F
- LEE, Judy .... 9D
- LÉGER, Christian .... 4E, 11G
- LERCH, Jason .... 7B
- LEROUX, Brian .... 11C
- LEROUX, Diane .... 9D
- LEVITT, C. .... 8F
- LEVY, A. R. .... 11F
- LI, Haitao .... 2E
- LI, Xiaoqiang .... 10F
- LIANG, Yuanyuan .... 11F
- LIEVESLEY, Denise .... 2A
- LIMNIOS, Nikolaos .... 9E, 12F
- LINDEYER, J. .... 6E
- LIU, Jinnan .... 2D
- LIU, Lin .... 8F
- LIU, Xuecheng .... 9F
- LIU, Zhong .... 7F
- LOCKHART, Richard .... 11E
- LOEPPKY, Jason .... 0B
- LOREDO-OSTI, J. .... 10F
- LOW, C.-K. .... 3F
- LU, Wilson .... 8E
- LUNAZZI, Helena .... 0B
- MA, Renjun .... 7D, 8A
- MAAROUF, Abdel .... 7F
- MACKEY, R. Jock .... 12D
- MADISE, Nyovani .... 7D
- MADRAS, Neal .... 6D
- MANDAL, Saumen .... 12C
- MANGIN, Jean-François .... 7B
- MANUEL, Douglas .... 11A
- MARCHAND, Éric .... 12C
- MARSHALL, D. A. .... 11F
- MARSHALL, J. K. .... 11F
- MATARISE, Florance .... 3E
- MATTHEWS, David E. .... 3F
- MCAULEY, Kim .... 11D
- MCCANDLESS, Lawrence .... 11F
- MCCLEAN, Kathy .... 4C
- MCLAUGHLIN, J. .... 4F
- MCLEOD, Ian .... 0B
- MCNALLY, Cathlin .... 10F
- MCQUIGGE, Murray .... 9B
- MCRAE, Ken .... 0A
- MEIGEN, Christof .... 8B
- MÉLARD, Guy .... 3E
- MERLEAU, James .... 10A
- MIHRAM, Danielle .... 12D
- MIHRAM, G. Arthur .... 12D
- MILLS-FLEMMING, Joanna .... 10E
- M'LAN, Cyr Emile .... 4F
- MOGHTADERI, Azadeh .... 0B
- MONETTE, Manon .... 4C
- MOORE, Ian .... 0B
- MOORHEAD, Philip .... 0B
- MORGAN, K. .... 10F
- MULAYATH VARIYATH, Asokan .... 12D
- MYSAK, Lawrence .... 9B
- NASON, Guy .... 6C
- NAWA, Victor .... 6F
- NEBEBE, F. .... 3E
- NEUPANE, Binod .... 2D

- NEWCOMBE-WELCH, Patricia ..... 8A  
 NG, Peggy ..... 10D  
 NIELSEN, Rasmus ..... 2C  
 NKUNDUMUKIZA, Eraste ..... 2D  
 NKURUNZIZA, Sévérien ..... 12C  
 NOBILE, A. G. .... 4D  
 NOVIKOV, Alexander ..... 4D  
  
 O'BRIEN, B. J. .... 11F  
 OGDEN, R. .... 6C  
 O'ROURKE, K. .... 8F  
 OUHBI, Brahim ..... 12F  
 OVERDUIN, Steve ..... 11E  
  
 PACURAR, Maria ..... 12B  
 PAN, Jianmin ..... 12E  
 PANTEL, Martin ..... 7E  
 PARK, Chul ..... 12F  
 PARK, Jeong-Gun ..... 11C  
 PARKHOMENKO, E. .... 4F  
 PELLETIER, Bernard ..... 11B  
 PELLETIER, Caroline ..... 7E  
 PELLETIER, Éric ..... 9D  
 PENG, Jianan ..... 8F  
 PENSKY, M. .... 6C  
 PERREAULT, Luc ..... 10A  
 PERRON, François ..... 12C  
 PFEFFERMANN, Danny ..... 6B  
 PHAM-GIA, Thu ..... 10E  
 PHILLIPS, K. A. .... 11F  
 PICKA, Jeffrey ..... 12E  
 PIEPHO, Hans-Peter ..... 11B  
 PIERRE, Fritz ..... 6E  
 PLANTE, Jean-François ..... 12B  
 PLATT, Robert ..... 2F, 4B  
 POLLOCK, Kenneth ..... 3D  
 POTTER, Elizabeth ..... 11A  
 PRADEL, Roger ..... 3D  
 PRASAD, Narasimha ..... 12F  
 PREGIBON, Daryl ..... 7A  
 PUNTANEN, S. .... 10E  
  
 QIAN, Jin ..... 12F  
  
 RAI, Shesh ..... 0B  
 RAMSAY, James ..... 8B  
 RAMSAY, Timothy ..... 9C  
 RANGREJ, Jagadish ..... 8F  
 RAO, J. N. K. .... 6B, 10B  
 REGIS, J. .... 7B  
  
 REID, Nancy ..... 11E  
 RÉMILLARD, Bruno ..... 6D  
 REN, Jian-Jian ..... 8C  
 RHODES, Anne ..... 11A  
 RICCIARDI, L. M. .... 4D  
 RINALDO, A. .... 1  
 RINDISBACHER, Marcel ..... 6D  
 RIVEST, Louis-Paul ..... 3F  
 RIVIÈRE, D. .... 7B  
 RODRIGUE, Natalie ..... 8D  
 ROEDER, Kathryn ..... 1  
 ROSLIN, N. .... 10F  
 ROSS, William ..... 8F  
 ROUTLEDGE, Rick ..... 3C  
 ROY, Jean-Denis ..... 3B  
 ROY, Laurent ..... 4C  
 ROY, Roch ..... 3E  
 ROZGA, Ante ..... 6E  
 RUDNISKY, C. .... 6F  
 RYTINA, Steven ..... 8A  
  
 SAIDI, Abdessamad ..... 3E  
 SAINT-PIERRE, Richard ..... 2B  
 SAINT-PIERRE, Yves ..... 4C  
 SALEH, A. K. Md. E. .... 4E  
 SALIBIAN-BARRERA, Matias ..... 10E  
 SALISBURY, Thomas ..... 11G  
 SANCHEZ LEZAMA, Ana Paola ..... 2D  
 SARKAR, Joykrishna ..... 9E  
 SAUER, Cornelia ..... 9E  
 SAULNIER, Michel ..... 2B  
 SCHIOPU-KRATINA, Ioana ..... 2F  
 SCHIPPER, Stefan ..... 7D  
 SCHMID, Wolfgang ..... 7D  
 SCHWARZ, Carl ..... 0A, 0B, 3D  
 SEN, Arusharka ..... 9E  
 SEN, Debaraj ..... 11E  
 SHABBIR, Javid ..... 7E  
 SHI, Peilin ..... 0B  
 SIMON, Richard ..... 6A  
 SIMONS, Ted ..... 3D  
 SINGH, G. .... 6F  
 SINGH, Sarjinder ..... 8E  
 SINGHAL, Sunil ..... 7E  
 SITTER, Randy R. .... 0B, 8E  
 SKORUPSKI, William P. .... 6F  
 SLOBODIANIK, Nikolai ..... 2D  
 SMITH, J. Terry ..... 12B  
 SMITH, Michael ..... 8A

- SNEDDON, Gary ..... 4F  
 SPEECHLEY, Kathy N. .... 11A  
 STAFFORD, Jamie ..... 11G  
 STAICU, Ana-Maria ..... 11E  
 STEEL, David ..... 7D  
 STEEL, Sarel J. .... 4E  
 STEELE, Fiona ..... 7D  
 STEINER, Stefan ..... 12D  
 STEPHENS, Michael A. .... 11E  
 STEVENSON, Joan ..... 12B  
 STRYHN, Henrik ..... 12E  
 STYAN, G. P. H. .... 10E  
 SUN, Tao ..... 2D  
 SUN, Ye ..... 2D  
 SUTRADHAR, Brajendra C. .... 4F  
 SUTRADHAR, Rinku ..... 9F  
 SVEINSSON, Oli ..... 10A  
 SZEKLI, Ryszard ..... 9E  
  
 TAI, Patricia ..... 10C  
 TAKAHARA, Glen ..... 0B, 12B  
 TAM, Raymond ..... 6F  
 TAN, Ye ..... 2F  
 TAYLOR, Jonathan ..... 9C  
 THABANE, Lehana ..... 8F, 10F, 11F  
 THAVANESWARAN, A. .... 0B  
 THOMPSON, Mary E. .... 4F, 5, 11E  
 THOMPSON, Paul ..... 7B  
 THOMSON, David J. .... 0B, 9E  
 THORNTON, Bill ..... 7E  
 TIAN, Yongge ..... 12D  
 TIBSHIRANI, Rob ..... 3C, 9C  
 TOGA, Arthur W. .... 7B  
 TOIBER TEMIN, Noemi ..... 2D  
 TRABELSI, Abdelwahed ..... 3E  
 TRUMMER, Manfred ..... 11G  
 TSAO, Min ..... 8C  
 TURKKAN, N. .... 10E  
 TURNER, Elizabeth ..... 10F  
 TURNER, Rolf ..... 3E  
  
 URRUTIA, María Inés ..... 0B  
  
 VANDAL, Alain C. .... 9F, 10F  
 VAN EEDEN, Constance ..... 4A  
 VÉGIARD, Sylvain ..... 0A, 2B  
 VOLINSKY, Chris ..... 7A  
 VOLODIN, Andrei ..... 9E  
  
 WAINER, Howard ..... 6F  
  
 WALTHER, Guenther ..... 9C  
 WANG, Jing ..... 2D  
 WANG, Liqun ..... 4D  
 WANG, Rui ..... 2D  
 WANG, Shufang ..... 2D  
 WANG, Xiaohui ..... 6F  
 WANG, Xikui ..... 2E  
 WANG, Yan ..... 2E  
 WANG, Zilin ..... 0B  
 WASSERMAN, L. .... 1  
 WELLS, Martin ..... 2E  
 WHITE, Andrew ..... 6F  
 WIENS, Douglas P. .... 12D  
 WILLIAMS, Trish ..... 2A  
 WILLIAMSON, Ian ..... 0B  
 WILLMS, Jon Douglas ..... 8A  
 WORSLEY, Keith ..... 0B  
 WU, Changbao ..... 10B  
 WU, Hulin ..... 11C  
 WU, Lang ..... 2F  
  
 XIE, Hansheng ..... 7E  
  
 YAN, Joel ..... 4C  
 YAU, Carole ..... 9C  
 YE, Jane ..... 0B  
 YI, Yanqing ..... 2E  
 YIPING, Dou ..... 7F  
 YORK, Thomas ..... 2C  
 YOSHIHAMA, Mieko ..... 6E  
 YOU, Yong ..... 8E  
 YU, Long ..... 2E  
 YUAN, Xiaobin ..... 12B  
 YUEN, John Wai Kong ..... 6D  
  
 ZAREPOUR, Mahmoud ..... 2E  
 ZENG, Leilei ..... 2F, 11C  
 ZHANG, Jie ..... 8F  
 ZHANG, Lingyun ..... 0B  
 ZHANG, Ying ..... 0B  
 ZHOU, Julie ..... 0B  
 ZHOU, Xian ..... 2D  
 ZHU, Mu ..... 10F  
 ZIDEK, Jim ..... 4A, 7F  
 ZIJDENBOS, Alex ..... 7B  
 ZOU, Fei ..... 8C