

SSC 2007 Meeting
Memorial University
St. John's, Newfoundland,
June 10-13



Congrès 2007 de la SSC
Université Memorial
St-Jean, Terre-Neuve,
10 au 13 juin



Table of Contents • Table des matières

Welcome • Bienvenue	4
Sponsors • Commanditaires	5
Organizers • Organisateurs	6
Exhibitors • Exposants	8
General Information • Information générale	9
Committees and Meetings • Comités et réunions	15
List of Sessions • Liste des sessions	19
Program • Programme	24
Abstracts • Résumés	59
Author Index • Index des auteurs	157
Author Index (per session) • Index des auteurs (par session)	162

Welcome • Bienvenue

Welcome to Memorial University, the host for the 35th annual meeting of the society. Those of you have also attended the 1990 SSC annual meeting, it is a great joy to see you back here at Memorial.

Memorial is one of the great higher learning institutions in the country. Starting with 57 students in September 1925, today Memorial has become the largest Atlantic university with more than 17,000 students. This university has awarded so far more than 50,000 degrees, most to residents of the province. The university has a beautiful campus with scenic views in all directions including the fabulous view of signal hill, one of the historical land marks in the province of Newfoundland and Labrador. We hope that you enjoy the scientific sessions of the meeting and all social events that will take place over the next few days in the beautiful environment of this great host university.

We also welcome you to St. John's, the oldest city of North America, known as the City of Legends, where you can view icebergs, watch whales, and experience Newfoundland's unique culture. Also, the Cape Spear, the most easterly point of North America is not far from the city, where one can experience the unique beauty of sunrise. We hope that you have planned for an extended stay in St. John's following the meeting to enjoy these and other unique land marks and heritages such as signal hill and fishing villages.

Enjoy.

Bienvenue à l'Université Memorial, qui accueille le 35^e congrès annuel de la Société. Pour ceux d'entre vous qui avez également participé au Congrès annuel de la SSC en 1990, c'est un grand plaisir de vous revoir à Memorial.

Memorial est l'un des grands établissements d'enseignement supérieur du Canada. Fondée en septembre 1925 avec 57 étudiants, Memorial est aujourd'hui la plus grande université des provinces atlantiques, avec plus de 17 000 étudiants. À ce jour, elle a décerné plus de 50 000 diplômes, pour la plupart à des résidents de la province. L'université se vante d'un magnifique campus et de vues panoramiques, dont une superbe de la colline Signal, haut lieu de l'histoire de la province de Terre-Neuve-et-Labrador. Nous espérons que vous profiterez pleinement des séances scientifiques du congrès, ainsi que de tous les événements sociaux qui se dérouleront au cours des prochains jours dans l'environnement agréable de cette belle université hôte.

Nous vous souhaitons également la bienvenue à St. John's, la plus vieille ville d'Amérique du nord, « Ville des légendes », où vous pourrez admirer les icebergs, observer les baleines et apprendre à connaître la culture unique de Terre-Neuve. De plus, du cap Spear, point le plus à l'est de toute l'Amérique du nord, situé non loin de la ville, les lève-tôt pourront admirer de magnifiques levers de soleil. Nous espérons que vous avez prévu de prolonger votre séjour à St. John's après le congrès pour profiter de tout ce que la région a à offrir, y compris la colline Signal et les nombreux villages de pêcheurs.

Bienvenue !

Sponsors • Commanditaires

SSC 2007 thanks the following sponsors for their extremely generous contributions for the success of this meeting:

Le Congrès SSC 2007 remercie les commanditaires suivants pour leurs très généreuses contributions qui ont mené au succès du congrès :

- Faculty of Science, Memorial University
- Department of Mathematics and Statistics, Memorial University
- Atlantic Association for Research in the Mathematical Sciences (AARMS)
- Centre de recherches mathématiques (CRM)
- The Fields Institute for Research in Mathematical Sciences (FIELDS)
- Pacific Institute for the Mathematical Sciences (PIMS)
- Mathematics of Information Technology and Complex Systems (MITACS)
- MINITAB
- SAS Institute, Inc.
- PEARSON Education Canada
- THOMSON NELSON

Organizers • Organisateurs

Local Arrangements Committee • Comité organisateur local

Chair • Président: Brajendra Sutradhar

To begin the conference work, the first thing that I looked for was a nice local attraction for the preparation of the conference web site. I thank Mr. Ed Kirby and his colleagues at the Newfoundland and Labrador Department of Tourism for their co-operation in the selection of the iceberg picture that you have seen on the web site and elsewhere. The construction of the conference web page begun with one graduate and another senior undergraduate students, but the major part of the final web page was done with the help of our departmental systems administrator Dwayne J. Hart. I sincerely thank all of them for their excellent work. Catherine Cox did almost all French translation for the web page. A very big thank is due to her. Thanks goes to Christian Léger too for helps with translation.

The organization of various local events would not have been possible without the work of my statistics colleagues. Special thanks go to Alwell Oyet, J Concepción Loredó-Osti, and Zhao Zhi Fan for their works as committee members. I also wish to express my sincere thanks to all of our secretaries: Rosalind English, Leonce Morrissey, Jackie Costello, and Jennifer Bishop for helping me in organizing committee meetings at various levels and for their special helps during the on-site registration. Thanks go to our graduate students too for their volunteered works during the registration and various cultural events.

Our departmental system administrators Dwayne J. Hart and Tom Boland were in charge of setting up the internet facilities and running the computers for paper presentations. The committee is extending special thanks to them for their wonderful works. Todd Constantine and Calvin Rose helped us with all class room supports. Thanks are also due to them.

The program with abstracts was produced using a database at CRM under the supervision of Christian Léger. Our special thanks go to Christian Léger and staff at CRM, in particular, André Montpetit, and Daniel Ouimet.

Félix Labrecque-Synnot, a student from Université de Montréal, translated most abstracts under the supervision of Hélène Bérard (Statistics Canada), Chair of the Bilingualism Committee. Members of the Bilingualism Committee have reviewed the abstracts while Christian Léger did final revisions.

Beginning at a leisurely pace well over two years ago and lately in a much more intense fashion, I have worked closely with Brian Mallard of Conference Services at the Memorial University. He has taken care of the room bookings and provided me many necessary contact information which made my work easier.

Chartwells Catering services was responsible for all food, including the banquet, the welcoming reception, the barbecue dinner and various committees' breakfasts, lunches and dinners. I thank Thelma Careen catering manager at Memorial University and the executive chef Tim Hills for their help especially in local food selection. Thanks are also due to Tim Hills and his staff for preparation of the food.

I also wish to thank Mykila Sherren for her graphic art work, Peggy Chafe and staff of the printing services, and Dean Hogan for helping us with the preparation of signs and banners.

Finally, I am thankful to Mr. Thaddeus Dreher, President of the Newfoundland and Labrador Multicultural Council for putting the multi-cultural music and dance program on Monday, June 11 evening, as a part of our Newfoundland Soiree.

Lorsque j'ai commencé à préparer le congrès, la première chose à trouver était une attraction locale pour en illustrer le site Web. Je tiens à remercier M. Ed Kirby et ses collègues du ministère du Tourisme de Terre-Neuve-et-Labrador qui m'ont aidé à sélectionner l'image de l'iceberg que vous avez vue sur le site Web et ailleurs. La construction de la page Web du Congrès a été lancée par un étudiant gradué et un autre étudiant en fin de premier cycle, puis finalisée pour la plupart avec l'aide de l'administrateur du réseau de notre département, Dwayne J. Hart. Je les remercie sincèrement pour leur excellent travail. Catherine Cox s'est chargée d'une grande partie de la traduction en français de la page Web - un grand merci. Mes remerciements aussi à Christian Léger, qui a aidé à la traduction.

L'organisation des divers événements locaux aurait été impossible sans l'aide de mes collègues en statistique. Je tiens à remercier tout particulièrement Alwell Oyet, J Concepción Loredó-Osti et Zhao Zhi Fan, mes co-membres du comité d'organisation. Je souhaite également exprimer ma sincère gratitude à tous nos secrétaires, Rosalind English, Leonce Morrissey, Jackie Costello et Jennifer Bishop, qui m'ont aidé à organiser des réunions de comité de divers niveaux et les inscriptions en ligne. Un grand merci aussi à nos étudiants gradués qui travailleront bénévolement pendant l'inscription et lors de divers événements culturels.

Les administrateurs du réseau de notre département, Dwayne J. Hart et Tom Boland, se sont chargés de mettre en place le service Internet et les ordinateurs requis pour les exposés. Le comité tient à les remercier tout particulièrement pour leurs travaux. Todd Constantine et Calvin Rose nous ont aidés à installer toutes les salles de classe – un grand merci aussi.

Le programme et les résumés ont été compilés à l'aide d'une base de données du CRM sous la supervision de Christian Léger. Nos sincères remerciements à Christian Léger et le personnel du CRM, notamment André Montpetit et Daniel Ouimet.

Félix Labrecque-Synnot, un étudiant de l'Université de Montréal, a traduit la plupart des résumés sous la supervision d'Hélène Bérard (Statistique Canada), présidente du Comité du bilinguisme. Des membres du Comité du bilinguisme ont relu les résumés, Christian Léger se chargeant de la révision finale.

Depuis un début tranquille il y a plus de deux ans, puis plus récemment de façon bien plus intense, j'ai travaillé en étroite collaboration avec Brian Mallard des Services de conférences à l'Université Memorial. Il s'est chargé des réservations de salles et m'a fourni beaucoup d'informations de contact qui m'ont facilité la tâche.

Les services de restauration Chartwells Catering ont organisé tous les repas, y compris le banquet, la réception de bienvenue, la grillade et divers déjeuners, dîners et soupers de comités. Je tiens à remercier Thelma Careen, directrice de la restauration de l'université, ainsi que le premier chef Tim Hills, pour leur aide, notamment en ce qui concerne la sélection de plats locaux. Un grand merci également à Tim Hills et à son personnel pour la préparation des repas.

Je souhaite remercier Mykila Sherren pour son travail d'art graphique, Peggy Chafe et le personnel des services d'impression, et Dean Hogan pour son aide dans la préparation de la signalisation et des bannières.

Enfin, je suis reconnaissant envers M. Thaddeus Dreher, président du Conseil multiculturel de Terre-Neuve-et-Labrador, d'avoir organisé le programme de musique et danses multiculturelles pour la soirée du lundi 11 juin, dans le cadre de notre Soirée Terre-Neuve.

Program Committee • Comité du programme

Chair • Président: John Braun, *The University of Western Ontario*

Members • Membres:

Biostatistics Section • Groupe de biostatistique

Paul Gustafson, *University of British Columbia*

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

Gemai Chen, *University of Calgary*

Survey Methods Section • Groupe de méthodologie d'enquête

Milorad Kovacevic, *Statistics Canada*

Probability Section • Groupe de probabilité

Andrei Volodin, *University of Regina*

Exhibitors • Exposants

Four exhibitors: Pearson Education Canada, Thomson Nelson, Minitab, and SAS Institute will exhibit their books and softwares to the delegates during all three days of the meeting. Their booths are located in the atrium on the first floor of the Inco Innovation Centre (IIC-AF1).

Quatre exposants, Pearson Education Canada, Thomson Nelson, Minitab et SAS Institute, présenteront leurs livres et logiciels aux participants pendant les trois jours du congrès. Leurs kiosques sont situés dans l'atrium au rez-de-chaussée du Centre d'innovation Inco (IIC-AF1).

General Information • Information générale

On-site registration will take place in the atrium on the first floor of the Inco Innovation Centre (IIC-AF1) from 14:00 to 22:00 on Saturday, June 9; from 11:00 to 18:00 on Sunday, June 10; from 8:00 to 15:00 on Monday, June 11; and from 8:00 to 11:00 on both Tuesday, June 12 and Wednesday, June 13. Whether you are pre-registered or registering on site, you will receive your conference bag along with the registration package from this on-site registration location.

Le bureau d'inscription sur site sera ouvert dans l'atrium au rez-de-chaussée du Centre d'innovation Inco (IIC-AF1) de 14 heures à 22 heures samedi 9 juin, de 11 heures à 18 heures dimanche 10 juin, de 8 heures à 15 heures lundi 11 juin et de 8 heures à 11 heures mardi 12 juin mercredi 13 juin. Que vous soyez pré-inscrit ou que vous ayez attendu pour vous inscrire sur site, vous recevrez votre sac de congrès et votre trousse d'inscription à ce bureau.

Transportation and Parking • Transport et stationnement

For those staying in the motel or one of the 2 conference hotels, there will be a conference bus from Monday, June 11 to Wednesday, June 13, leaving at 7:45 a.m. every morning from your hotel/motel to the Inco Innovation Centre (IIC) (main meeting place) located in the centre of the university campus. On Monday evening, there will be a bus available at 9:45 p.m., right after the cultural show, to take you to your hotel from the Arts and Administration (AA) building. Also, on Tuesday evening a bus will be available at 9:45 p.m. following the banquet to take you to your hotel from the R. Gushue Hall.

For those renting a car and who would like to use the university parking facilities free of cost, you can park your car in the parking lot 15 during the conference from 8:30am to 5:00pm on everyday from Sunday, June 10 to Wednesday, June 13. The IIC is only a 3 minutes walk away from the parking lot.

Pour ceux d'entre vous qui logez au motel ou à l'un des deux hôtels du congrès, un autobus de congrès quittera l'hôtel/motel chaque matin à 7 heures 45, du lundi 11 juin au mercredi 13 juin, pour vous conduire au Centre d'innovation Inco (IIC) (lieu principal du congrès), qui est situé au centre du campus universitaire. Lundi soir, un autobus repartira pour les hôtels à 21 heures 45, à la fin de la soirée culturelle, au départ du bâtiment des Arts et Administration (AA). De même, mardi soir un autobus repartira pour les hôtels à 21 heures 45, à la fin du banquet, au départ du pavillon R. Gushue.

Pour ceux d'entre vous qui avez loué une voiture et qui souhaitez stationner gratuitement sur le campus, vous pouvez laisser votre véhicule au parc de stationnement 15 de 8 heures 30 à 17 heures pendant la durée du congrès, du dimanche 10 juin au mercredi 13 juin. Le centre IIC est à trois minutes de marche du parc de stationnement.

Directions • Comment vous retrouver

There are campus and city maps in every conference bag so that you may easily find your way around the campus and the city.

For those staying in the Hatcher House Residence, it would be a 3 minutes walk from the residence to the IIC.

For those staying in the motel or one of the 2 conference hotels, and would like to take a taxi to go the meeting place, the taxi ride should be short and pleasant for an approximate fare between 10 and 15 dollars. The IIC of the university is only 5 minutes away by taxi from Holiday Inn or Super 8 Motel, and 7 minutes away from Ramada St. John's.

On a good sunny day walking from your motel/hotel to the university can be a good experience. If you are staying in Holiday Inn and would like to walk to the university, just take Prince Philip Drive from the traffic intersection in front of the hotel and walk until you reach the overhead food court at the University Centre. The IIC then will be on your left. If you would like to walk from the Super 8 motel to the university, you will take Higgins line and walk toward the university. Take left on Allandale Road and then right on Prince Philip Drive until you reach the overhead food court at the University Centre. The IIC then will be on your left.

For those staying in Ramada St. John's and who would like to walk from the hotel to the university, you will take Kenmount Road and walk toward the university. Kenmount Road turns to Fresh Water Road after an over-bridge. Walk on Fresh Water Road for a short while until you meet Elizabeth Avenue on your left. Then walk on this Elizabeth Avenue until you reach the Arts and Administration (A) building of the university on your left. The IIC is located just next to the Arts and Administration building.

Vous trouverez dans votre sac de congrès un plan du campus et de la ville qui vous aidera à retrouver votre chemin facilement.

Pour ceux d'entre vous qui logez à la résidence universitaire Hatcher House, celle-ci est à trois minutes de marche du centre IIC.

Pour ceux d'entre vous qui logez au motel ou dans l'un des deux hôtels du congrès et qui souhaitez prendre un taxi pour l'université, la route est agréable et devrait vous coûter de 10 à 15 dollars. En taxi, le centre IIC est à cinq minutes de route de l'hôtel Holiday Inn ou du motel Super 8, et à sept minutes de l'hôtel Ramada St. John's.

S'il fait beau, il peut être agréable de faire le chemin du motel/hôtel au campus à pied. Si vous logez au Holiday Inn, empruntez Prince Philip Drive depuis l'intersection devant l'hôtel et continuez jusqu'à l'espace-restauration au Centre universitaire. Vous verrez alors le centre IIC sur votre gauche. À partir du motel Super 8, prenez Higgins Line en direction de l'université. Tournez à gauche sur Allandale Road puis à droite sur Prince Philip Drive jusqu'à rejoindre l'espace-restauration du Centre universitaire. Vous verrez alors le centre IIC sur votre gauche.

Pour ceux d'entre vous qui logez à l'hôtel Ramada St. John's, empruntez Kenmount Road en direction de l'université. Kenmount Road devient Fresh Water Road après un passage supérieur. Continuez un moment sur Fresh Water Road jusqu'à découvrir Elizabeth Avenue sur votre gauche. Empruntez alors Elizabeth Avenue jusqu'au bâtiment des Arts et Administration (A), qui sera sur votre gauche. Le centre IIC est situé tout à côté de celui-ci.

Rooms • Salles

The meeting will take place in the Inco Innovation Centre (IIC), located in the centre of the campus of the university. This IIC facility is surrounded by the Arts and Administration (AA), Science (SN) and the Queen Elizabeth II Library (L) buildings. The on-site registration, all coffee breaks and displays by exhibitors will take place in the atrium on the first floor of the IIC. The workshops and scientific sessions will take place in IIC-2001, AA-1043, AA-1046, SN-2109, SN-2105, SN-2041, and SN-1103.

The Board meeting on Sunday will be held in Junior Common Room (JCR) at R. Gushue Hall (RGH). All other committee meetings held over either breakfast or lunch, as well as the Wednesday Board meeting, will be held in either JCR or Dining Room (DR) at R. Gushue Hall.

Le Congrès se tiendra dans le Centre d'innovation Inco (IIC), situé au centre du campus universitaire. Le centre IIC est entouré des bâtiments des Arts et Administration (AA) et des Sciences (SN) et de la bibliothèque Queen Elizabeth II (L). L'inscription sur site, les pauses café et les kiosques des exposants seront situés dans l'atrium du rez-de-chaussée du centre IIC. Les ateliers et les séances scientifiques auront lieu dans les salles IIC-2001, AA-1043, AA-1046, SN-2109, SN-2105, SN-2041 et SN-1103.

Dimanche, la réunion du Conseil d'administration aura lieu dans le Junior Common Room (JCR) du pavillon R. Gushue (RGH). Les autres réunions de comités, qui se tiendront à l'heure du déjeuner ou du dîner, et la réunion du Conseil d'administration du mercredi, se tiendront dans le JCR ou dans la salle à manger (DR) du pavillon R. Gushue.

Workshops • Ateliers

All three workshops organized by the Biostatistics, Survey Methods, and Business and Industrial Statistics sections of the SSC will take place on Sunday, June 10, from 9:00 to 17:00. The Biostatistics workshop will be held in IIC-2001, whereas the workshops by the Survey Methods and the Business and Industrial statistics sections will take place in AA-1043 and AA-1046, respectively.

Les ateliers organisés par le Groupe de biostatistique, le Groupe des méthodes d'enquête et le Groupe de statistique industrielle et de gestion de la SSC se tiendront dimanche 10 juin de 9 heures à 17 heures. L'atelier de biostatistique aura lieu dans la salle IIC-2001, celui du groupe des méthodes d'enquête se tiendra dans la salle AA-1043 et celui de statistique industrielle et de gestion, dans la salle AA-1046.

Poster Session • Séance d'affichage

The poster session will be held on Sunday, June 10 from 14:00 to 17:00 in the Atrium on the first floor of the Inco Innovation centre (IIC-AF1).

La séance d'affiches se tiendra dimanche 10 juin de 14 heures à 17 heures dans l'atrium du rez-de-chaussée du centre Inco (IIC-AF1).

NSERC News Session and Workshop • Séance de nouvelles et atelier offerts par le CRSNG

Monday, June 11 • Lundi 11 juin
12:45–13:30 (IIC-2001) & 17:00–18:00 (AA-1043)

An NSERC news session will take place on Monday, June 11 from 12:45 to 13:30 in IIC-2001. Serge Villemure, Director, Chemistry, Engineering and Mathematical Sciences at NSERC will make a presentation and lead a discussion on (1) the review of grant selection committees (GSC) structure, (2) international review of discovery grants program, and (3) discovery grants competitions results for 2007.

Also, there will be an NSERC workshop on the same day from 17:00 to 18:00 in AA-1043. NSERC representatives and statistics grant selection committee members will lead this workshop. The news session and the workshop are open to all interested researchers. New faculty members are especially welcome!

Une séance d'information CRSNG aura lieu lundi 11 juin, de 12 heures 45 à 13 heures 30 dans la salle IIC-2001. Serge Villemure, le directeur de Chimie, Ingénierie et Sciences mathématiques du CRSNG, fera une présentation et animera une discussion sur : (1) la restructuration des comités de sélection des subventions de recherche, (2) la révision internationale du programme de subventions à la découverte, et (3) les résultats des concours de subventions à la découverte pour 2007.

Un atelier CRSNG aura également lieu le même jour de 17 heures à 18 heures dans la salle AA-1043. Des représentants du CRSNG et des membres du comité de sélection des subventions de recherche en statistiques l'animeront. La séance d'information et l'atelier sont ouverts à tous les chercheurs intéressés. Les nouveaux membres du corps professoral sont tout particulièrement les bienvenu(e)s !

Welcome Reception • Réception de bienvenue

The welcoming reception will take place at the R. Gushue Hall on Sunday, June 10 from 18:30 to 20:30. The reception will begin with traditional Newfoundland instrumental music by Dave Painting and Graham Wells, two local talented musicians. All conference attendees and companions are warmly invited to come and enjoy a social time with old and new friends in the statistical community. A cash bar will be available for the whole 2 hours duration. Your reception coupon given in your conference bag will allow you to have 2 beers/soft-drinks/juices free from the cash bar.

La réception de bienvenue aura lieu au pavillon R. Gushue dimanche 10 juin de 18 heures 30 à 20 heures 30. Elle débutera par une sélection d'airs instrumentaux traditionnels de Terre-Neuve interprétés par Dave Painting et Graham Wells, deux talentueux musiciens locaux. Les participants au Congrès et personnes accompagnantes sont tous chaleureusement invités à venir profiter de la soirée, retrouver de vieux amis et s'en faire de nouveaux dans la communauté statistique. Un bar payant sera ouvert pendant les deux heures de la réception. Le coupon de réception, que vous trouverez dans votre sac de congrès, vous donnera droit à deux bières, boissons gazeuses ou jus de fruits gratuits au bar.

Food Service on Campus • Restauration sur le campus

You can purchase your lunch from the overhead food court in the University Centre which is open for lunch all through out the year. This is conveniently located at the north side of the IIC with only a parking lot between the two centres.

The second option to have your lunch would be to go to one of the restaurants such as Pasta Plus Café or Quintanas de la Plaza in the Churchill Square, which would be a 5 minutes walk from the IIC. To locate this square, walk toward the intersection of Elizabeth Avenue and Allandale Road. You will then find the square on Elizabeth Avenue next to the North Atlantic gas station. The North Atlantic gas station itself houses a Subway restaurant which you may like to keep in your list as a quick lunch option.

À midi, vous pouvez acheter à manger dans l'espace-restauration du Centre universitaire, qui est ouvert toute l'année en milieu de journée. Le centre est bien situé, au nord du centre IIC, séparé de celui-ci par un parc de stationnement.

La deuxième possibilité est de dîner dans l'un des restaurants, comme Pasta Plus Café ou Quintanas de la Plaza, au centre commercial Churchill Square, à cinq minutes à pied du centre IIC. Dirigez-vous vers l'intersection d'Elizabeth Avenue et d'Allandale Road. Le centre est situé sur Elizabeth Avenue à côté de la station-service North Atlantic. La station-service elle-même contient un restaurant Subway, où vous pouvez également vous procurer un dîner sur le pouce.

Barbecue/Newfoundland Soiree • Grillade/Soirée Terre-Neuve

A special evening is planned for Monday, June 11. This event will begin with a barbecue dinner, followed by a multi-cultural music and dance program. If you have registered for this special evening, you are given 2 coupons in your conference bag, one to be used for the barbecue and the other for the multi-cultural program.

The barbecue will take place from 18:00 to 19:15 in and around the Dining Room at R. Gushue Hall. Your entertainment of the evening will begin with 'Gulliver's Spree' featuring Dave Painting and Tom Boland singing traditional Newfoundland Folk Songs. You will be expected to present your barbecue coupon for your choice of meal. An indoor cash bar will be available from 18:00. Your barbecue coupon will also allow you to have one beer/soft-drink/juice free from the cash bar.

The barbecue dinner will be followed by a multi-cultural music and dance program to be held in the Reid Theater at the Arts and Administration Building from 19:30 to 21:30. It takes approximately 5 minutes to walk from the R. Gushue Hall to the Reid Theater. You will be expected to have your seat in the Theater by 19:25 latest. You then sit back, relax and enjoy the program featuring Newfoundland and Labrador Multicultural Council singing and dancing.

Une soirée spéciale est prévue lundi 11 juin. La réception débutera par une grillade, qui sera suivie d'un programme de musique et de danses multiculturelles. Si vous vous êtes inscrit pour cette soirée spéciale, vous aurez reçu deux coupons dans votre trousse du congrès, l'un pour la grillade et l'autre pour le programme multiculturel.

La grillade se tiendra de 18 heures à 19 heures 15 dans la salle à manger du pavillon R. Gushue et aux alentours. Le divertissement débutera avec « Gulliver's Spree », une collection de chansons traditionnelles de Terre-Neuve interprétées par Dave Painting et Tom Boland. Pour être servi, vous devrez présenter votre coupon de grillade. Un bar payant sera ouvert à partir de 18 heures à l'intérieur du bâtiment. Votre coupon de grillade vous donnera également droit à une bière, une boisson gazeuse ou un jus de fruits gratuit au bar.

La grillade sera suivie d'un programme de musique et de danses multiculturelles qui se tiendra au théâtre Reid, dans le bâtiment des Arts et Administration de 19 heures 30 à 21 heures 30. Il faut environ cinq minutes pour vous rendre à pied du pavillon R. Gushue au théâtre Reid. Vous êtes prié de rejoindre votre fauteuil dans le théâtre à 19 heures 25 au plus tard. Ensuite, vous n'avez plus qu'à vous détendre et à apprécier le programme, qui inclut des chansons et des danses du Conseil multiculturel de Terre-Neuve-et-Labrador.

Women in Statistics Reception • Réception pour les femmes en statistique

The Canadian Section of the Caucus for Women in Statistics and the SSC Committee on Women in Statistics cordially invite all interested men and women to attend an informal get-together over dinner in Giovanni Cabot's restaurant, 10 Kings Road at 20:30 on Sunday, June 10. Students are especially welcome.

La section canadienne du Caucus for Women in Statistics et le Comité sur les femmes en statistique de la SSC invitent cordialement tou(te)s les intéressé(e)s à participer à une rencontre informelle autour d'un souper au restaurant Giovanni Cabot's, 10 Kings Road, à 20 heures 30, dimanche 10 juin. Les étudiant(e)s surtout sont les bienvenu(e)s !

Banquet

The conference banquet will take place in the Dining Room at R. Gushue Hall on Tuesday evening, June 12. All participants who did not decline their banquet ticket when they registered are cordially invited to attend the banquet. Banquet coupons are distributed with the registration package and you will be expected to present your coupon on the dinner table for your choice of meal. A cash bar will be available from 19:00. Dinner will be served at 19:30. It is expected that several entertainers/musicians will entertain you through-out the banquet evening.

The Society awards will be presented at the end of the evening.

Le banquet du congrès se tiendra dans la salle à manger du pavillon R. Gushue mardi soir, 12 juin. Les participants qui n'ont pas décliné le billet de banquet lors de leur inscription sont cordialement invités à y participer. Lors du repas, vous êtes prié de placer votre coupon de banquet, qui vous a été remis avec votre trousse d'inscription, sur la table avant d'être servi. Un bar payant sera ouvert à partir de 19 heures. Le souper sera servi à 19 heures 30. Des artistes et musiciens vous divertiront tout au long de la soirée.

Les prix de la Société seront présentés en fin de soirée.

Internet Access • Accès Internet

Computers with internet access are available in three different locations: Rooms HH-3030 & HH-3056 [HH-Henrietta Harvey] in the Mathematics building; On the main lobby at the Queen Elizabeth II Library (L) building; and Rooms C-2003 & C-2004 in the Chemistry Physics building. Your conference bag contains a page with a user id and password. The HH cum Mathematics building is located next to the Science building toward the Elizabeth Avenue, whereas the Chemistry-Physics (C) building is also located next to the Science building but toward the Prince Philip Drive. The Queen Elizabeth II Library (L) is located next to the IIC.

Des ordinateurs avec accès Internet sont disponibles dans trois lieux différents : dans les salles HH-3030 et HH-3056 [HH-Henrietta Harvey] du bâtiment de Mathématiques ; dans le hall principal de la bibliothèque Queen Elizabeth II (L) ; et dans les salles C-2003 et C-2004 du bâtiment de Chimie-Physique. Vous trouverez dans votre sac du congrès un feuillet qui indique votre ID d'utilisateur et votre mot de passe. Le bâtiment HH (Mathématiques) est situé à côté du bâtiment des Sciences près d'Elizabeth Avenue, le bâtiment de Chimie-Physique se situant de l'autre côté du bâtiment des Sciences, dans la direction de Prince Philip Drive. La bibliothèque Queen Elizabeth II (L) est située à côté du centre IIC.

Committees and Meetings

Comités et réunions

Calendar Listing • Horaire journalier

Saturday June 9 • Samedi 9 juin

18:00–22:00 Executive Committee • *Comité exécutif* Trinity Board Room
Holiday Inn

Sunday, June 10 • Dimanche 10 juin

9:00–11:00 Finance Committee • *Comité des finances* Junior Common Room
R. Gushue Hall

10:30–14:30 Heads Meeting • *Rencontre des directeurs de départements* SN-2105

11:00–12:00 Publications Committee • *Comité des publications* Dining Room
R. Gushue Hall

12:00–16:00 Board of Directors • *Conseil d'administration* Junior Common Room
R. Gushue Hall

Monday, June 11 • Lundi 11 juin

7:00–8:00 Biostatistics Section (Section Officers) • *Groupe de biostatistique (directeurs du groupe)* Junior Common Room
R. Gushue Hall

Business and Industrial Statistics Section (Section Officers) • Dining Room
Groupe de statistique industrielle et de gestion (directeurs du groupe) R. Gushue Hall

Survey Methods Section (Section Officers) • *Groupe des méthodes d'enquête (directeurs du groupe)* Junior Common Room
R. Gushue Hall

Probability Section (Section Officers) • *Groupe de probabilité (directeurs du groupe)* Dining Room
R. Gushue Hall

12:15–13:15 Accreditation Appeals Committee • *Comité des appels d'accréditation* Dining Room
R. Gushue Hall

CJS Editorial Board • *Comité de rédaction de la RCS* Junior Common Room
R. Gushue Hall

Public Relations Committee • *Comité des relations publiques* Dining Room
R. Gushue Hall

13:30–15:00 Accredited Members General Meeting (open) • *Assemblée générale des membres accrédités (public)* Junior Common Room
R. Gushue Hall

- 17:00–18:00 Business and Industrial Statistics Section (General Meeting) • SN-2105
Groupe de statistique industrielle et de gestion (assemblée générale)
- Survey Methods Section (General Meeting) • *Groupe des méthodes d'enquête (assemblée générale)* SN-2109
- Biostatistics Section (General Meeting) • *Groupe de biostatistique (assemblée générale)* IIC-2001
- Probability Section (General Meeting) • *Groupe de probabilité (assemblée générale)* AA-1046

Tuesday, June 12 • Mardi 12 juin

- 7:00–8:00 Research Committee • *Comité de la recherche* Junior Common Room
 R. Gushue Hall
- Accreditation Committee • *Comité sur l'accréditation* Dining Room
 R. Gushue Hall
- Professional Development Committee • *Comité sur le perfectionnement professionnel* Junior Common Room
 R. Gushue Hall
- Working Group on the Position of CJS Managing Editor • East-Side Mario's
Groupe de travail pour le poste de rédacteur gérant de la RCS Holiday Inn
- 12:15–13:15 NPCDS • *PNSDC* Junior Common Room
 R. Gushue Hall
- Bilingualism Committee • *Comité du bilinguisme* Dining Room
 R. Gushue Hall
- Committee on Women in Statistics • *Comité sur les femmes en statistique* Junior Common Room
 R. Gushue Hall
- Statistical Education Committee • *Comité d'éducation en statistique* Dining Room
 R. Gushue Hall
- 17:00–18:15 SSC General Meeting • *Assemblée générale de la SSC* IIC-2001

Wednesday, June 13 • Mercredi 13 juin

- 7:00–8:00 Biostatistics Section (Section Officers) • *Groupe de biostatistique (directeurs du groupe)* Junior Common Room
 R. Gushue Hall
- Business and Industrial Statistics Section (Section Officers) • Dining Room
Groupe de statistique industrielle et de gestion (directeurs du groupe) R. Gushue Hall
- Survey Methods Section (Section Officers) • *Groupe des méthodes d'enquête (directeurs du groupe)* Junior Common Room
 R. Gushue Hall
- Probability Section (Section Officers) • *Groupe de probabilité (directeurs du groupe)* Dining Room
 R. Gushue Hall
- 12:15–13:15 Liaison Dining Room
 R. Gushue Hall
- Program Committee • *Comité du programme* Junior Common Room
 R. Gushue Hall
- 17:30–19:30 Board of Directors • *Conseil d'administration* Junior Common Room
 R. Gushue Hall

19:30 Executive Committee • *Comité exécutif* Junior Common Room
R. Gushue Hall

Alphabetical Listing • Liste alphabétique

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

Monday, June 11 • lundi 11 juin, 12:15–13:15, Dining Room, R. Gushue Hall

Accreditation Committee • *Comité sur l'accréditation*

Tuesday, June 12 • mardi 12 juin, 7:00–8:00, Dining Room, R. Gushue Hall

Accredited Members General Meeting (open) • *Assemblée générale des membres accrédités (public)*

Monday, June 11 • lundi 11 juin, 13:30–15:00, Junior Common Room, R. Gushue Hall

Bilingualism Committee • *Comité du bilinguisme*

Tuesday, June 12 • mardi 12 juin, 12:15–13:15, Dining Room, R. Gushue Hall

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

Monday, June 11 • lundi 11 juin, 17:00–18:00, IIC-2001

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

Monday, June 11 • lundi 11 juin, 7:00–8:00, Junior Common Room, R. Gushue Hall

Wednesday, June 13 • mercredi 13 juin, 7:00–8:00, Junior Common Room, R. Gushue Hall

Board of Directors • *Conseil d'administration*

Sunday, June 10 • dimanche 10 juin, 12:00–16:00, Junior Common Room, R. Gushue Hall

Wednesday June 13 • mercredi 13 juin, 17:30–19:30, Junior Common Room, R. Gushue Hall

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

Monday, June 11 • lundi 11 juin, 17:00–18:00, SN-2105

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

Monday, June 11 • lundi 11 juin, 7:00–8:00, Dining Room, R. Gushue Hall

Wednesday, June 13 • mercredi 13 juin, 7:00–8:00, Dining Room, R. Gushue Hall

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

Monday, June 11 • lundi 11 juin, 12:15–13:15, Junior Common Room, R. Gushue Hall

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

Tuesday, June 12 • mardi 12 juin, 12:15–13:15, Junior Common Room, R. Gushue Hall

Executive Committee • *Comité exécutif*

Saturday June 9 • samedi 9 juin, 18:00–22:00, Holiday Inn, Trinity Board Room

Wednesday June 13 • mercredi 13 juin, 19:30, Junior Common Room, R. Gushue Hall

Finance Committee • *Comité des finances*

Sunday, June 10 • dimanche 10 juin, 9:00–11:00, Junior Common Room, R. Gushue Hall

Heads Meeting • *Rencontre des directeurs de départements*

Sunday, June 10 • dimanche 10 juin, 10:30–14:30, SN-2105

Liaison

Wednesday, June 13 • mercredi 13 juin, 12:15–13:15, Dining Room, R. Gushue Hall

NPCDS • *PNSDC*

Tuesday, June 12 • mardi 12 juin, 12:30–13:30, Junior Common Room, R. Gushue Hall

Probability Section (General Meeting) • Groupe de probabilité (assemblée générale)

Monday, June 11 • lundi 11 juin, 17:00–18:00, AA-1046

Probability Section (Section Officers) • Groupe de probabilité (directeurs du groupe)

Monday, June 11 • lundi 11 juin, 7:00–8:00, Dining Room, R. Gushue Hall

Wednesday, June 13 • mercredi 13 juin, 7:00–8:00, Dining Room, R. Gushue Hall

Professional Development Committee • Comité sur le perfectionnement professionnel

Tuesday, June 12 • mardi 12 juin, 7:00–8:00, Junior Common Room, R. Gushue Hall

Program Committee • Comité du programme

Wednesday, June 13 • mercredi 13 juin, 12:15–13:15, Junior Common Room, R. Gushue Hall

Public Relations Committee • Comité des relations public

Monday, June 11 • lundi 11 juin, 12:15–13:15, Dining Room, R. Gushue Hall

Publications Committee • Comité des publications

Sunday, June 10 • dimanche 10 juin, 11:00–12:00, Dining Room, R. Gushue Hall

Research Committee • Comité de la recherche

Tuesday, June 12 • mardi 12 juin, 7:00–8:00, Junior Common Room, R. Gushue Hall

SSC General Meeting • Assemblée générale de la SSC

Tuesday, June 12 • mardi 12 juin, 17:00–18:15, IIC-2001

Statistical Education Committee • Comité d'éducation en statistique

Tuesday, June 12 • mardi 12 juin, 12:15–13:15, Dining Room, R. Gushue Hall

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

Monday, June 11 • lundi 11 juin, 17:00–18:00, SN-2109

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

Monday, June 11 • lundi 11 juin, 7:00–8:00, Junior Common Room, R. Gushue Hall

Wednesday, June 13 • mercredi 13 juin, 7:00–8:00, Junior Common Room, R. Gushue Hall

Working Group on the Position of CJS Managing Editor • Groupe de travail pour le poste de rédacteur-gérant de la RCS

Tuesday, June 12 • mardi 12 juin, 7:00–8:00, East Side Mario's, Holiday Inn

List of Sessions • Liste des sessions

Sunday June 10 • Dimanche 10 juin

14:00–17:00

IIC-AF1 Session 0 Contributed Poster Session
Séance d'affichage

Monday June 11 • Lundi 11 juin

08:30–09:00

IIC-2001 Session 01A Inaugural Session
Séance inaugurale

09:00–10:00

IIC-2001 Session 01B SSC Presidential Invited Address
Allocution de l'invité de la présidente de la SSC

10:30–12:00

IIC-2001 Session 02A Diagnostics for the Retrospective Problem in Sequential Population Analysis of Exploited Fish Stocks

Diagnostiques pour le problème rétrospectif dans l'analyse séquentielle des populations pour l'exploitation des stocks de poissons

AA-1043 Session 02B Design and Analysis of Experiments for Complex Physical Systems

Planification et analyse d'expériences pour des systèmes physiques complexes

SN-2109 Session 02C Generalized Quasilikelihood Inference

Inférence par quasi-vraisemblance généralisée

AA-1046 Session 02D Recent Developments in Multivariate Survival Data Analysis Methodology

Développements récents dans la méthodologie de l'analyse de données de survie multivariées

SN-2105 Session 02E Survey Analysis

Analyse d'enquêtes

SN-2041 Session 02F Issues in Clinical Trials

Problématiques dans les essais cliniques

10:30–11:45

SN-1103 Session 02G Statistical Theory I
Théorie statistique I

12:45–13:30

IIC-2001 News from NSERC
Nouvelles du CRSNG

13:45–14:45

IIC-2001 Session 03A *Canadian Journal of Statistics Award Address*
Allocution du récipiendaire du Prix de la Revue canadienne de statistique

13:45–15:00

SN-2109 Session 03B Case Study I: Demographic Methods & Case Study II: Growth of Atlantic Herring
Étude de cas I : Méthodes démographiques & Étude de cas II : La croissance du hareng de l'Atlantique

13:30–15:00

SN-1103 Session 03C Mathematical Finance and Stochastic Analysis I
Finance mathématique et analyse stochastique I

13:45–14:45

AA-1043 Session 03D Controversies over Fish Stocks
Controverses concernant les stocks de poissons

13:30–15:00

AA-1046 Session 03E Random Effects and Multi-level Modelling
Effets aléatoires et modélisation multi-niveaux

SN-2105 Session 03F Survey Analysis and Operations
Analyse d'enquêtes et opérations

15:30–17:00

AA-1046 Session 04A Strategic Directions in Household Surveys Development
Directions stratégiques dans le développement des enquêtes sur les ménages

AA-1043 Session 04B Mathematical Methods for Studying Patient Wait Times
Méthodes mathématiques pour étudier les temps d'attente des patients

IIC-2001 Session 04C Reliability Models and Their Applications
Modèles de fiabilité et leurs applications

SN-2105 Session 04D Markov Chain Monte Carlo and Quasi Monte Carlo – A Productive Marriage
Chaînes de Markov Monte Carlo et quasi Monte Carlo – Un mariage productif

SN-2109 Session 04E Case Study II: Growth of Atlantic Herring
Étude de cas II : La croissance du hareng de l'Atlantique

SN-2041 Session 04F Statistics and the Environment
La statistique et l'environnement

SN-1103 Session 04G Statistical Theory II
Théorie statistique II

17:00–18:00

AA-1043 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 12 • Mardi 12 juin**08:45–10:00**

IIC-2001 Session 05A Gold Medal Address
Allocution du récipiendaire de la médaille d'or

10:30–12:00

SN-2109 Session 06A Remedies for Sample Survey Outliers
Remèdes pour les valeurs aberrantes en échantillonnage

AA-1046 Session 06B Mathematical Finance and Stochastic Analysis II
Finance mathématique et analyse stochastique II

AA-1043 Session 06C Statistical Methods for Correlated Data
Méthodes statistiques pour données corrélées

10:30–11:30

IIC-2001 Session 06D Isobel Loutit Invited Address on Business and Industrial Statistics
Allocution Isobel Loutit sur la statistique industrielle et de gestion

10:30–12:00

SN-2105 Session 06E Forestry
Foresterie

10:30–11:45

SN-2041 Session 06F Statistical Genomics
Génomique statistique

13:30–14:30

AA-1043 Session 07A Special Invited Session of the Probability Section
Allocution de l'invité d'honneur du groupe de probabilité

13:30–15:00

AA-1046 Session 07B Statistical Genetics
Génétique statistique

IIC-2001 Session 07C Statistical Graphics and Image Analysis
Graphiques statistiques et analyse d'image

SN-2105 Session 07D Statistics Applied to Hydrology and Climatology
Statistique appliquée à l'hydrologie et à la climatologie

SN-2109 Session 07E The National Institute for Complex Data Structures: Our Community in Scientific Collaborations
l'Institut national pour les structures de données complexes : Notre communauté dans des collaborations scientifiques

SN-2041 Session 07F Longitudinal Data and Panel Counts
Données longitudinales et dénombrement de panels

15:30–17:00

AA-1043 Session 08A Survey Design and Analysis
Planification et analyse d'enquêtes

15:30–16:30

SN-2041 Session 08B Stochastic Partial Differential Equations and their Applications
Equations aux dérivées partielles stochastiques et leurs applications

- IIC-2001 Session 08C Inaugural Invited Lecture of the National Institute for Complex Data Structures
Allocution inaugurale de l'invité d'honneur de l'Institut national sur les structures de données complexes
- SN-2109 Session 08D Confidence Intervals and Regions
Intervalles et régions de confiance
- AA-1046 Session 08E Graphical Methods and Data Visualization
Méthodes graphiques et visualisation de données

15:30–17:00

- SN-2105 Session 08F Business and Industrial Statistics
Statistique industrielle et de gestion

Wednesday June 13 • Mercredi 13 juin**08:30–9:30**

- IIC-2001 Session 09A Pierre Robillard Award Address
Allocution du récipiendaire du prix Pierre-Robillard

08:30–10:00

- AA-1043 Session 09B Genetic Epidemiology
Épidémiologie génétique

08:30–9:30

- AA-1046 Session 09C Data with Extra Zeros Applied in Ecology
Données avec un surplus de zéros et applications en écologie

08:30–10:30

- SN-2109 Session 09D Latest Developments in Small Area Estimation, Theory and Practice
Derniers développements pour l'estimation de petites régions : théorie et aspects pratiques

08:30–10:00

- SN-2105 Session 09E Nonparametric Curve Estimation
Estimation non paramétrique de courbes

08:30–9:45

- SN-2041 Session 09F Stochastic Processes and Their Applications
Processus stochastiques et leurs applications

10:30–12:00

- SN-2109 Session 10A Special Invited Session sponsored by the Survey Methods Section
Allocution de l'invité d'honneur du groupe de méthodologie d'enquête
- IIC-2001 Session 10B Statistical Methods for Causal Inference
Méthodes statistiques pour l'inférence causale
- SN-2041 Session 10C Matrices and Statistics I
Matrices et statistique I
- SN-2105 Session 10D Point Processes
Processus ponctuels
- AA-1043 Session 10E Biostatistical Methods
Méthodes biostatistiques

10:30–11:45

AA-1046 Session 10F Statistical Theory III
Théorie statistique III

13:30–14:30

IIC-2001 Session 11A Address of the winner of the Centre de Recherches Mathématiques/Statistical Society of Canada Prize
Allocution du récipiendaire du prix Centre de recherches mathématiques/Société statistique du Canada

15:00–16:30

AA-1043 Session 12A Biostatistics
Biostatistique

IIC-2001 Session 12B Applications of Empirical Likelihood in Survey Sampling
Applications de la vraisemblance empirique en échantillonnage

SN-2105 Session 12C Matrices and Statistics II
Matrices et statistique II

SN-2109 Session 12D Robust Statistics
Statistiques robustes

AA-1046 Session 12E Biostatistical Modelling
Modélisation biostatistique

SN-2041 Session 12F Time Series
Séries chronologiques

Program • Programme

Dimanche 10 juin • Sunday, June 10

14:00–17:00 Session 0

IIC-AF1

Contributed Poster Session

Séance d'affichage

Contributed Paper Session • Séance de communications libres

14:00 William HARPER, *Otterbein College*, Ted ESCHENBACH, *TGE Consulting*

Development of Exact Confidence Intervals for the Ratio of Two Poisson Oil Spill Rates • Développement d'intervalles de confiance exacts pour le rapport de taux de déversement d'hydrocarbures modélisés par des variables poissoniennes [MS-1]

14:00 Qinggang WANG & John KOVAL, *University of Western Ontario*

Best Alpha Level of Significance and Selection Statistics in Backward Stepwise Logistic Regression • Meilleur seuil de signification alpha et statistiques de sélection dans la régression logistique pas-à-pas en arrière [MS-2]

14:00 Yingru (Celia) HUANG, Ying ZHANG & Paul CABILIO, *Acadia University*

Analysis of the Blanding's Turtle's Growth • Analyse de la croissance de la tortue de Blanding [MS-3]

14:00 Sameer PARPIA, John KOVAL & Allan DONNER, *University of Western Ontario*

Evaluation and comparison of confidence interval construction methods for the kappa statistic in the presence of rater bias • Évaluation et comparaison de méthodes de construction d'intervalles de confiance pour la statistique kappa en présence de biais dû au juge [MS-4]

14:00 Qihao XIE, *Bombardier Aerospace*, Narayanaswamy BALAKRISHNAN, *McMaster University*, Debasis KUNDU, *India Institute of Technology*

Exact Inference For A Simple Step-stress Model From the Exponential Distribution Under Time Constraint • Inférence exacte pour un modèle simple à contraintes échelonnées à partir de la distribution exponentielle sous contraintes de temps [MS-5]

14:00 Jeffrey DOWDEN, *Memorial University*, Noel CADIGAN, *Fisheries and Oceans Canada*

Improved Estimation and Forecasts of Stock Maturities using Generalized Linear Mixed Effects Models • Estimation et prévision améliorée de l'échéance des actions utilisant des modèles linéaires généralisés à effets mixtes [MS-6]

14:00 Paramjit GILL, *UBC Okanagan*, Charmaine DEAN, *Simon Fraser University*, Bob PROSSER & Bruce CARLETON, *University of British Columbia*

Spatio-temporal Trends for Treated Asthma in the Canadian Province of British Columbia • Tendances spatio-temporelles pour l'asthme traité dans la province canadienne de Colombie-Britannique [MS-7]

14:00 Matthew DAVIS, *University of Calgary*, David FLETCHER, *University of Otago*

Confidence Intervals for the Mean of a Zero-inflated Distribution • Intervalles de confiance pour la moyenne d'une distribution gonflée à zéro [MS-8]

14:00 Matthew DAVIS, *University of Calgary*

Robust Estimates of the Correlation Parameter for the Bivariate Normal with Extensions to Elliptical Distributions • Estimés robustes du paramètre de corrélation de la normale bivariée et extension aux distribution elliptiques [MS-9]

14:00 Vaneeta GROVER, *Dalhousie University*, David COLE, *University of Toronto*, David HAMILTON, *Dalhousie University*

Hardy-Weinberg Disequilibrium (HWD) Analysis in Candidate Gene Association Studies of Heterogeneous Disorders: Solving the Problem of the Multiplicative Model • Analyse du déséquilibre d'Hardy-Weinberg dans les études d'association du gène candidat de désordres hétérogènes : résoudre le problème du modèle multiplicatif [MS-10]

14:00 Joseph AMUAH & Michael ELIASZIW, *University of Calgary*

Using MLE for the Concurrent Assessment of Interrater and Intrarater Reliability and a Comparison to the Method of Moments Approach • Utilisation du maximum de vraisemblance pour l'évaluation simultanée de la fiabilité inter- et intra-juge et comparaison avec l'approche de la méthode des moments [MS-11]

14:00 Xin MA, *University of Kentucky*

Extending Education Statistics with CART • Étendre les statistiques d'éducation au moyen d'arbres de classification et de régression [MS-12]

14:00 Patrick BROWN, *Cancer Care Ontario/Action Cancer Ontario*, Sabrina Russo & David COOMES, *University of Cambridge*, Mark REES, *University of Sheffield*

Growth/survival Trade-offs in Tropical Forests using Bayesian Hierarchical Models • Utilisation des modèles hiérarchiques bayésiens pour estimer la relation entre la croissance et la survie des espèces d'arbres dans les forêts tropicales [MS-13]

Lundi 11 juin • Monday, June 11

08:30–09:00 Session 01A

IIC-2001

Inaugural Session

Séance inaugurale

Special Session • Conférence spéciale

Special Guest Speakers • Conférenciers invités de marque

Organizer and Chair • Responsable et président: Brajendra SUTRADHAR, *Memorial University*

08:30 Dr. Axel MEISEN, President, *Memorial University*

08:40 Dr. Eddy CAMPBELL, Vice president Academic, *Memorial University*

09:00–10:00 Session 01B

IIC-2001

SSC Presidential Invited Address

Allocution de l'invité de la présidente de la SSC

Special Session • Conférence spéciale

Organizer and Chair • Responsable et présidente: Charmaine DEAN, *Simon Fraser University*

09:00 Norman BRESLOW, *University of Washington*

The Case-control Study: Origins and Modern Conception • L'étude cas-témoins : origines et conception moderne [MS-14]

10:00–10:30 Pause-café • Coffee Break

IIC-AF1

10:30–12:00 Session 02A

IIC-2001

Diagnostics for the Retrospective Problem in Sequential Population Analysis of Exploited Fish Stocks

Diagnostiques pour le problème rétrospectif dans l'analyse séquentielle des populations pour l'exploitation des stocks de poissons

Invited Paper Session • Présentations sur invitation

Session sponsored by the *Biostatistics Section*

Session parrainée par le *Groupe de biostatistique*

Chair • Président: Carl SCHWARZ, *Simon Fraser University*

Organizers • Responsables: Noel CADIGAN, *Fisheries and Oceans Canada*

10:30 Larry JACOBSON & Chris LEGAULT, *Northeast Fisheries Science Center*, Bob MOHN, *Fisheries and Oceans Canada*, Al SEAVER, *Northeast Fisheries Science Center*

Statistical Properties of the rho Statistic for Retrospective Bias – A Simulation Analysis • Propriétés statistiques de la statistique rho pour le biais rétrospectif – une analyse par simulation [MS-15]

11:15 Noel CADIGAN, *Fisheries and Oceans Canada*

Diagnostics for detecting the source of retrospective patterns • Diagnostiques détectant la source de tendances rétrospectives [MS-16]

10:30–12:00 Session 02B**AA-1043**

Design and Analysis of Experiments for Complex Physical Systems

Planification et analyse d'expériences pour des systèmes physiques complexes

Invited Paper Session • Présentations sur invitation

Session sponsored by the *Business and Industrial Statistics Section*Session parrainée par le *Groupe de statistique industrielle et de gestion*Organizer and Chair • Responsable et président: Derek BINGHAM, *Simon Fraser University*10:30 Dave HIGDON, *Los Alamos National Laboratory*, Derek BINGHAM, *Simon Fraser University*, Brian WILLIAMS, *Los Alamos National Laboratory*

Gaussian Process Models for High Dimensional Spaces • Modèles de processus gaussiens pour espaces de dimension élevée [MS-17]

11:00 Vijay NAIR, *University of Michigan*, Derek BINGHAM, *Simon Fraser University*

Modeling and Analysis of Dispersion Effects in Robust Design Experiments • Modélisation et analyse d'effets de dispersion dans des plans d'expériences robustes [MS-18]

11:30 Crystal LINKLETTER, *Simon Fraser University*

Latent Socio-spatial Process Model for Social Networks • Modèle de processus socio-spatial latent pour réseaux sociaux [MS-19]

10:30–12:00 Session 02C**SN-2109**

Generalized Quasilikelihood Inference

Inférence par quasi-vraisemblance généralisée

Invited Paper Session • Présentations sur invitation

Organizer and Chair • Responsable et président: Pat FARRELL, *Carleton University*10:30 Wasimul BARI & Brajendra SUTRADHAR, *Memorial University*

A Fully Standardized Mallows Type Quasi-likelihood Inference for Generalized Linear Models • Inférence standardisée de type de Mallows par quasi-vraisemblance pour modèles linéaires généralisés [MS-20]

11:00 Milorad KOVACEVIC, *Statistics Canada/Statistique Canada*, Brajendra SUTRADHAR, *Memorial University*

Analysis of Incomplete Longitudinal Survey Binary Data by Conditional Weighted Generalized Quasi-likelihood • Analyse par quasi-vraisemblance conditionnelle pondérée généralisée de données binaires incomplètes tirées d'enquêtes longitudinales [MS-21]

11:30 Katrina ROGERS-STEWART, *Carleton University*, Brajendra SUTRADHAR, *Memorial University*, Patrick FARRELL, *Carleton University*

Generalized Quasilikelihood Estimation of a Nonlinear Dynamic Model for Longitudinal Binary Data • Estimation par quasi-vraisemblance généralisée d'un modèle dynamique non-linéaire pour des données binaires longitudinales [MS-22]

10:30–12:00 Session 02D**AA-1046**

Recent Developments in Multivariate Survival Data Analysis Methodology

Développements récents dans la méthodologie de l'analyse de données de survie multivariées

Invited Paper Session • Présentations sur invitation

Organizer and Chair • Responsable et président: Wenqing HE, *University of Western Ontario*10:30 (Tony) Jianguo SUN & Manhua CHEN, *University of Missouri*, Xingwei TONG, *Beijing Normal University*

Regression Analysis of Multivariate Interval-censored Failure Time Data • Analyse de régression de données multivariées de temps de défaillance censurées par intervalles [MS-23]

11:00 Jianwen CAI & Sangwook KANG, *University of North Carolina at Chapel Hill*

Marginal Hazards Regression for Retrospective Studies within Cohort with Correlated Failure Time Data • Régression pour la fonction de risque marginale dans des études rétrospectives intra-cohortes avec temps de défaillances corrélés [MS-24]

11:30 Charmaine DEAN, *Simon Fraser University*, Farouk NATHOO, *University of Victoria*

Spatial Multi-state Transitional Models for Longitudinal Event Data • Modèles spatiaux transitionnels à états multiples pour données d'évènements longitudinales [MS-25]

10:30–12:00 Session 02E**SN-2105**

Survey Analysis

Analyse d'enquêtes

Contributed Paper Session • Séance de communications libres

Chair • Présidente: Mary MARCH, *Statistics Canada/Statistique Canada*10:30 Thomas LOUGHIN, *Simon Fraser University Surrey*, Christopher BILDER, *University of Nebraska*

Modeling Association Between Two or More Multiple-response Categorical Variables • Modéliser l'association entre plusieurs variables catégorielles à réponses multiples [MS-26]

10:45 Serge GODBOUT, *Statistics Canada/Statistique Canada*

What is the Relationship between Average Weekly Earnings and Average Monthly Earnings? A Latent Variable Approach • Quelle est la relation entre les gains hebdomadaires moyens et la paie mensuelle moyenne? Une approche par variable latente [MS-27]

11:00 Ivan CARRILLO GARCIA & Changbao WU, *University of Waterloo*

The GEE Methodology for Longitudinal Surveys with Missing Responses • La méthode des équations estimatrices généralisées pour les enquêtes longitudinales avec valeurs manquantes [MS-28]

11:15 Norberto PANTOJA-GALICIA, *University of Waterloo*

Using Local Likelihood Density Estimation for Bias Correction at and near a Boundary • Utilisation d'estimation de densité par vraisemblance locale afin de corriger le biais à et autour d'une frontière [MS-29]

11:30 Owen PHILLIPS & Avi C. SINGH, *Statistics Canada/Statistique Canada*

A Justification and Improvement of Woodruff's Method for Quantile Estimation from Survey Data • Une justification et amélioration de la méthode de Woodruff pour l'estimation de quantiles à partir de données d'enquêtes [MS-30]

11:45 Avi SINGH & Mike TAM, *Statistics Canada/Statistique Canada*

A Generalization of Fieller's Method for Improved Variance and Interval Estimation of Ratio of Means • Une généralisation de la méthode de Fieller pour une estimation améliorée de la variance et des intervalles de confiance pour le rapport de moyennes [MS-31]

10:30–12:00 Session 02F

SN-2041

Issues in Clinical Trials

Problématiques dans les essais cliniques

Contributed Paper Session • Séance de communications libres

Chair • Présidente: Alison GIBBS, *University of Toronto*

10:30 Lihui ZHAO & Joan HU, *Simon Fraser University*, Stephen LAGAKOS, *Harvard University*

Statistical Monitoring of Clinical Trials with Multivariate Response or Multiple Arms Using Repeated Confidence Bands • Suivi statistique d'essais cliniques avec réponse multivariée ou à plusieurs groupes en utilisant des bandes de confiance répétées [MS-32]

10:45 Cheng Lin YE & Lehana THABANE, *McMaster University*

How to Deal with Patients Who Withdraw Data from Randomized Controlled Trials: Individual versus Public Rights in Scientific Research • Comment réagir face aux patients retirant leurs données d'essais cliniques randomisés : droits individuels et droits publics en recherche scientifique [MS-33]

11:00 Ying ZHANG & Paul CABILIO, *Acadia University*, Maja GRUBISIC, *University of Southern California*, Femida GWADRY-SRIDHAR, *University of Western Ontario*

Modeling Medication Compliance • Modéliser l'observance à un traitement [MS-34]

11:15 Rinku SUTRADHAR, *University of Toronto*, Samuel Lunenfeld Research Institute, Richard Cook, *University of Waterloo*

Clustered Progressive Multi-state Processes Under Incomplete Observation • Processus progressif en grappes à états multiples sous observation incomplète [MS-35]

11:30 Ji (Emmy) CHENG, Lehana THABANE, Jack SUN & Richard WHITLOCK, *McMaster University*

Dealing with Zero Events – Comparing Different Strategies Used in Meta-analysis • Traiter les études sans événements – comparaisons de différentes stratégies utilisées en méta-analyse [MS-36]

11:45 Jinhui MA, *McMaster University*

Comparisons of Bayesian and Classical Methods in Analysis of Cluster Randomized Trial: The CHAT Trial • Comparaisons de méthodes bayésiennes et classiques dans l'analyse d'essais cliniques randomisés par grappes : l'étude CHAT [MS-37]

10:30–11:45 Session 02G

SN-1103

Statistical Theory I

Théorie statistique I

Contributed Paper Session • Séance de communications libres

Chair • Président: Reg KULPERGER, *University of Western Ontario*

10:30 Bashir KHAN, *Saint Mary's University*

Stein-type Estimation of the Mean: The Behrens-Fisher Problem • Estimation de type Stein de la moyenne : le problème de Behrens-Fisher [MS-38]

10:45 Ahmad ZGHOUL, *Qatar University*

Order Statistics and Records from a Family of J -shaped Distributions • Statistiques d'ordre et records d'une famille de distributions en forme de J [MS-39]

11:00 S. Ejaz AHMED, *University of Windsor*, Andrei VOLODIN, *University of Regina*

Asymptotic Expansion of the Coverage Probability of James-Stein Estimators • Développement asymptotique de la probabilité de couverture des estimateurs de James-Stein [MS-40]

11:15 Amir PAYANDEH, *University of New Brunswick*, Éric MARCHAND, *Université de Sherbrooke*

Estimation of a Bounded Location Parameter for p -variate spherical distributions • Estimation d'un paramètre de position multidimensionnel dont la norme est bornée [MS-41]

11:30 Serge PROVOST, *University of Western Ontario*

On the Distribution of Functions of Elliptically Contoured Random Vectors • Sur la distribution de fonctions de vecteurs aléatoires appartenant à la classe des lois elliptiques [MS-42]

12:00–13:30 Dîner • Lunch

12:45–13:30 Nouvelles du CRSNG • News from NSERC

IIC-2001

Special Session • Conférence spéciale

Organizer and Chair • Responsable et président: François BELLAVANCE, *HÉC Montréal*

12:45 Serge VILLEMURE, *NSERC/CRSNG*

News from NSERC • Nouvelles du CRSNG [MS-43]

13:45–14:45 Session 03A

IIC-2001

Canadian Journal of Statistics Award Address

Allocution du récipiendaire du Prix de la *Revue canadienne de statistique*

Special Session • Conférence spéciale

Organizer and Chair • Responsable et présidente: Mary LESPERANCE, *University of Victoria*

13:45–15:00 Session 03B

SN-2109

Case Study I: Demographic Methods & Case Study II: Growth of Atlantic Herring

Étude de cas I : Méthodes démographiques & Étude de cas II : La croissance du hareng de l'Atlantique

Case Study • Étude de cas

Chair • Président: Georges MONETTE, *York University*

Organizers • Responsables: Peggy NG, *York University*

13:45 Georges MONETTE, *York University*

Introduction of the Demographic Methods Data Set • Présentation du jeu de données sur les méthodes démographiques [MS-44]

13:55 Andrea Soo, Zhulin HE, Pik Tan WONG & George HE, *University of Calgary*

14:10 John WHEELER, *Department of Fisheries and Oceans*

Introduction of the Growth of Atlantic Herring Data Set • Présentation du jeu de données sur la croissance du hareng de l'Atlantique [MS-45]

14:20 Chun-Po Steve FAN, Gerald LEBOVIC & Ying QI, *University of Toronto*

14:35 Shahedul Ahsan KHAN, *University of Waterloo*

13:30–15:00 Session 03C

SN-1103

Mathematical Finance and Stochastic Analysis I

Finance mathématique et analyse stochastique I

Invited Paper Session • Présentations sur invitation

Session sponsored by the *Probability Section*

Session parrainée par le *Groupe de probabilité*

Organizer and Chair • Responsable et président: Martin BARLOW, *University of British Columbia*

13:30 Matt DAVISON, TYSON WHITEHEAD & Mark REESOR, *University of Western Ontario*

A Bias Correction Technique for Monte Carlo Pricing of Early Exercise Options • Une technique de correction du biais pour la fixation des prix d'options bermudiennes par Monte-Carlo [MS-46]

14:00 Rogemar MAMON, *University of Western Ontario*, Tak Kuen SIU, *Herriot-Watt University, Edinburgh*, Christina ERLWEIN, *Brunel University, London*

Valuation of CDS in a Regime-switching Market • Évaluation du swap de défaut de crédit dans un marché avec changement de régimes [MS-47]

14:30 Tom HURD, *McMaster University*

Hybrid Models of Credit Risk – The Best of both Worlds? • Modèles hybrides du risque de crédit – le meilleur des deux mondes ? [MS-48]

13:45–14:45 Session 03D

AA-1043

Controversies over Fish Stocks

Controverses concernant les stocks de poissons

Special Late Breaking Session • Conférence spéciale dernière nouvelle

Session sponsored by the / Session parrainée par *Atlantic Association for Research in the Mathematical Sciences (AARMS)*

Organizer and Chair • Responsable et président: Noel CADIGAN, *Fisheries and Oceans Canada*

13:45 Peter SHELTON, *Department of Fisheries and Oceans*

From Science-based to Ad Hoc Fisheries Management Off the East Coast of Canada • D'une gestion des pêches scientifique à une gestion ad hoc sur la côte Est du Canada [MS-49]

13:30–15:00 Session 03E**AA-1046**

Random Effects and Multi-level Modelling

Effets aléatoires et modélisation multi-niveaux

Contributed Paper Session • Séance de communications libres

Chair • Président: Doug WOOLFORD, *University of Toronto*13:30 Denis LAROCQUE, *HEC Montréal*, Jaakko NEVALAINEN & Hannu OJA, *University of Tampere*

One-Sample Location Tests for Multilevel Data • Tests de position pour données multi-niveaux [MS-50]

13:45 Marie-Eve BEAUCHAMP, Robert PLATT & James HANLEY, *McGill University*

Performance of Estimation Methods for Generalized Linear Mixed Models for Binary Outcome Data with a Low Number of Occurrences • Performance de méthodes d'estimation pour modèles linéaires généralisés mixtes pour données dont la variable réponse est binaire avec peu d'événements positifs [MS-51]

14:00 Jiajia ZHANG, *Memorial University*, Yingwei PENG, *Queen's University*

Identifiability of Mixture Cure Frailty Model • Identifiabilité du modèle de mélange de fragilité d'un traitement [MS-52]

14:15 Depeng JIANG & Debra PEPLER, *York University*

A Generalized Mixed Modeling Approach for Causal Inference in Intervention Study • Une approche par modélisation mixte généralisée pour l'inférence causale dans une étude d'intervention [MS-53]

14:30 Rashid AHMED & K. Stephen BROWN, *University of Waterloo*

Use of Markov Transition Model to Analyze Longitudinal Data on Youth Smoking • Utilisation d'un modèle de Markov transitionnel dans l'analyse de données longitudinales sur les jeunes fumeurs [MS-54]

14:45 John NEUHAUS & Charles McCULLOCH, *University of California, San Francisco*

An Assessment of Bias due to Misspecification of the Random Effects Distribution in Generalized Linear Mixed Models • Une évaluation du biais résultant d'une erreur de spécification de la distribution des effets aléatoires dans les modèles linéaires généralisés mixtes [MS-55]

13:30–15:00 Session 03F**SN-2105**

Survey Analysis and Operations

Analyse d'enquêtes et opérations

Contributed Paper Session • Séance de communications libres

Chair • Président: Chris WILD, *University of Auckland*13:30 Carlos LEON, *Statistics Canada/Statistique Canada*

Reporting Response Rates in Characteristic Surveys • Taux de réponse pour les enquêtes sur les caractéristiques [MS-56]

- 13:45 Limei SUN & Laurent ROY, *Statistics Canada/Statistique Canada*
New Collection Methodology in the 2006 Census of Population • Nouvelle méthodologie de collecte pour le Recensement de l'an 2006 [MS-57]
- 14:00 Beatrice CHAPMAN & Sarah FRANKLIN, *Statistics Canada/Statistique Canada*
Sampling and Retaining Respondents for the National Longitudinal Survey of Children and Youth (NLSCY) • Échantillonnage et retenue de répondants pour l'Enquête longitudinale nationale sur les enfants et les jeunes (ELNEJ) [MS-58]
- 14:15 François BRISEBOIS, Julie GIRARD & Hansheng XIE, *Statistics Canada/Statistique Canada*
Challenges Surrounding the Use of Tax Data in the Development of Quarterly Indicators for Various Service Industries • Défis entourant l'utilisation des données fiscales dans le développement d'indicateurs trimestriels pour diverses industries de services [MS-59]
- 14:30 Mary MARCH, *Statistics Canada/Statistique Canada*
Evaluating the Quality of Establishment Survey Frames • Évaluer la qualité des bases de sondage dans les enquêtes auprès des établissements [MS-60]
- 14:45 Cathlin SARAFIN & Steven THOMAS, *Statistics Canada/Statistique Canada*
Weighting of Cumulated Samples for the Canadian Community Health Survey • Pondération des échantillons cumulés pour l'Enquête sur la santé dans les collectivités canadiennes [MS-61]

15:00–15:30 Pause-café • Coffee Break**IIC-AF1****15:30–17:00 Session 04A****AA-1046****Strategic Directions in Household Surveys Development****Directions stratégiques dans le développement des enquêtes sur les ménages**

Invited Paper Session • Présentations sur invitation

Session sponsored by the *Survey Methods Section*Session parrainée par le *Groupe des méthodes d'enquête*Organizer and Chair • Responsable et président: Edward CHEN, *Statistics Canada/Statistique Canada*

- 15:30 Jack GAMBINO, Jean-Louis TAMBAY & Guy LAFLAMME, *Statistics Canada/Statistique Canada*
Statistics Canada's New Household Survey Strategy • La nouvelle stratégie de Statistique Canada pour les enquêtes sur les ménages [MS-62]
- 16:00 Nicola BENNETT, *Office for National Statistics, United Kingdom*
Towards a Single Continuous Population Survey for the UK • Vers une seule enquête de population continue pour le Royaume-Uni [MS-63]
- 16:30 Alfredo NAVARRO, *US Census Bureau*
American Community Survey: Statistical Issues and Challenges • L'enquête américaine sur les communautés : défis et perspectives statistiques [MS-64]

15:30–17:00 Session 04B**AA-1043****Mathematical Methods for Studying Patient Wait Times****Méthodes mathématiques pour étudier les temps d'attente des patients**

Invited Paper Session • Présentations sur invitation

Session sponsored by the *Biostatistics Section*Session parrainée par le *Groupe de biostatistique*Chair • Présidente: James ADCOCK, *University of Western Ontario*Organizer • Responsable: Wendy LOU, *University of Toronto*15:30 Michael CARTER, *University of Toronto*

Using Operations Research to Manage Wait Lists • Utilisation de la recherche opérationnelle dans la gestion de files d'attentes [MS-65]

16:00 Wendy LOU, *University of Toronto*

Aggregation Methods for Health Performance Indicators of Wait Times • Méthodes d'agrégation pour indicateurs de performance des temps d'attentes dans le système de santé [MS-66]

16:30 Boris SOBOLEV, *University of British Columbia*

Methods for Studying Events on Surgical Wait Lists • Méthodes pour l'étude d'événements sur des listes d'attentes chirurgicales [MS-67]

15:30–17:00 Session 04C**IIC-2001****Reliability Models and Their Applications****Modèles de fiabilité et leurs applications**

Invited Paper Session • Présentations sur invitation

Session sponsored by the *Business and Industrial Statistics Section*Session parrainée par le *Groupe de statistique industrielle et de gestion*Organizer and Chair • Responsable et président: Thierry DUCHESNE, *Université Laval*15:30 George WHITMORE, *McGill University*, Mei-Ling Ting LEE, *Ohio State University*

Uncoupling Longitudinal Data in Threshold Regression • Découpler des données longitudinales en régression de seuil [MS-68]

16:00 Jerry LAWLESS, *University of Waterloo*

Reliability and Warranty Claims for Systems with Age and Usage Scales • Fiabilité et réclamations de garanties pour des systèmes à échelles d'âge et d'utilisation [MS-69]

16:30 Andrew JARDINE & Dragan BANJEVIC, *University of Toronto*

Condition Based Maintenance: Estimating the Remaining Useful Life of an Asset • Entretien conditionnel : estimer la vie utile d'un actif [MS-70]

15:30–17:00 Session 04D**SN-2105**

Markov Chain Monte Carlo and Quasi Monte Carlo – A Productive Marriage
 Chaînes de Markov Monte Carlo et quasi Monte Carlo – Un mariage productif

Invited Paper Session • Présentations sur invitation

Organizer and Chair • Responsable et président: Radu CRAIU, *University of Toronto*

15:30 Art OWEN & Seth TRIBBLE, *Stanford University*

QMC for MCMC • QMC pour MCMC [MS-71]

16:15 Christiane LEMIEUX, *University of Waterloo*, Radu CRAIU, *University of Toronto*

Acceleration of the Multiple-Try Metropolis Algorithm using Antithetic and Stratified Sampling •
 Accélération de l'algorithme de Metropolis à essais multiples via l'échantillonnage antithétique et stratifié [MS-72]

15:30–17:00 Session 04E**SN-2109**

Case Study II: Growth of Atlantic Herring

Étude de cas II : La croissance du hareng de l'Atlantique

Organizer and Chair • Responsable et présidente: Alison GIBBS, *University of Toronto*

15:30 John WHEELER, *Department of Fisheries and Oceans*

Introduction to the Data Set • Présentation du jeu de données [MS-73]

15:40 Xiao Bo (Alice) LI, *University of Saskatchewan*

15:55 Oksanna FILL, Lindsay JACKS, Hui WANG & C.L. YE, *McMaster University*

16:10 Hao WANG, *University of New Brunswick*

16:25 Shun-Fu CHEN, Michael Anthony ROTONDI & Julia TALEBAN, *University of Western Ontario*

16:40 Baojiang CHEN, Xiaoqin XIONG & Yan YUAN, *University of Waterloo*

15:30–17:00 Session 04F**SN-2041**

Statistics and the Environment

La statistique et l'environnement

Contributed Paper Session • Séance de communications libres

Chair • Présidente: Jennifer ASIMIT, *Samuel Lunenfeld Research Institute*

15:30 Amanda HALLADAY, *Dalhousie University*

Parameter Estimation using Monte Carlo Methods and the EM Algorithm • Estimation de paramètres utilisant des méthodes de Monte-Carlo et l'algorithme EM [MS-74]

15:45 Carolyn HUSTON & Carl SCHWARZ, *Simon Fraser University*, Jake SCHWEIGERT, *Department of Fisheries and Oceans*

Finding the Fish: Straying Trends Between B.C. Pacific Herring (*Clupea pallasii*) Populations •
 Trouver le poisson : orientations de populations de hareng (*Clupea pallasii*) au large de la Colombie-Britannique [MS-75]

16:00 Anne Michele MILLAR, *Mount Saint Vincent University*

Bootstrap Inference for Ranking and Selection: Applications to Forest Genetics • Inférence par bootstrap pour le classement et la sélection et applications en génétique forestière [MS-76]

16:15 Paul NGUYEN, *University of Western Ontario*, Linh NGUYEN, *Trent University*

An R^2 alternative to measure the accuracy of species distribution models • Une mesure de précision de modèles de distribution des espèces alternative au R^2 [MS-77]

16:30 Daniel GILLIS & Stephanie DIXON, *University of Guelph*

Climate Change and Gastrointestinal Illness in Alberta – A Spatial Temporal Analysis using Conditionally Autoregressive Models • Changement climatique et maladies gastro-intestinales en Alberta – une analyse spatiale et temporelle utilisant des modèles conditionnellement autorégressifs [MS-78]

15:30–17:00 Session 04G

SN-1103

Statistical Theory II

Théorie statistique II

Contributed Paper Session • Séance de communications libres

Chair • Président: Hyung-Tae HA, *Kyungwon University*

15:30 Jeffrey PICKA & Mingxia DENG, *University of New Brunswick*

Classification of Random Patterns • Classification de formes aléatoires [MS-79]

15:45 Éric MARCHAND, *Université de Sherbrooke*

Balanced-type Loss Functions: Review and Recent Results • Fonctions de perte “équilibrées” : revue et résultats récents [MS-80]

16:00 David VAUGHAN, *Wilfrid Laurier University*

Parameter Estimation in the Skew Generalized Secant Hyperbolic Distribution • Estimation des paramètres de la distribution sécante hyperbolique généralisée asymétrique [MS-81]

16:15 Alexandre LEBLANC, *University of Manitoba*

Chung-Smirnov Property for Bernstein Estimators of Distribution Functions • Propriété de Chung-Smirnov et estimateurs de Bernstein d’une fonction de répartition [MS-82]

16:30 Ali GHODSI & Michael BIGGS, *University of Waterloo*, Michael BOWLING *University of Alberta*, Dana WILKINSON, *University of Waterloo*, David WINGATE, *University of Michigan*

Scalable Action Respecting Embedding (ARE) • Emboîtement respectant les actions à échelles multiples [MS-83]

16:45 Chris WILD, Alastair SCOTT, Alan LEE & Yannan JIANG, *University of Auckland*

Maximum Likelihood Estimation under Three-phase Sampling • Estimation par maximum de vraisemblance sous un plan d’échantillonnage à trois degrés [MS-84]

17:00–18:00 Atelier du CRSNG: Conseils pour l’élaboration de votre prochaine demande de subvention à la découverte • NSERC Workshop: Tips to Prepare your Next Discovery Grant Application

AA-1043

Mardi 12 juin • Tuesday, June 12

08:45–10:00 Session 05A

IIC-2001

Gold Medal Address

Allocution du récipiendaire de la médaille d'or

Special Session • Conférence spéciale

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

08:45 Chris FIELD, *Dalhousie University*

Modelling Biological Data: Several Vignettes • Modélisation de données biologiques : plusieurs vignettes [MS-85]

10:00–10:30 Pause-café • Coffee Break

IIC-AF1

10:30–12:00 Session 06A

SN-2109

Remedies for Sample Survey Outliers

Remèdes pour les valeurs aberrantes en échantillonnage

Invited Paper Session • Présentations sur invitation

Session sponsored by the *Survey Methods Section*

Session parrainée par le *Groupe des méthodes d'enquête*

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

10:30 Jean-François BEAUMONT, *Statistics Canada/Statistique Canada*, Louis-Paul RIVEST, *Université Laval*

A Weight Smoothing Method for Dealing with Stratum Jumpers in Business Surveys • Une méthode de lissage des poids pour faire face aux sauteuses de strate dans les enquêtes auprès des entreprises [MS-86]

11:15 Mary MULRY & Roxanne FELDPAUSCH, *U.S. Census Bureau*

Treating Influential Values in a Monthly Retail Trade Survey • Traiter les valeurs influentes dans une enquête mensuelle sur le commerce de détail [MS-87]

10:30–12:00 Session 06B

AA-1046

Mathematical Finance and Stochastic Analysis II

Finance mathématique et analyse stochastique II

Invited Paper Session • Présentations sur invitation

Session sponsored by the *Probability Section*

Session parrainée par le *Groupe de probabilité*

Organizer and Chair • Responsable et président: Robert ELLIOTT, *University of Calgary*

10:30 Abel CADENILLAS & LUZ SOTOMAYOR, *University of Alberta*

Solutions of Consumption-investment Problems in Financial Markets with Regime Switching • Solutions aux problèmes de consommation-investissement dans des marchés financiers avec changement de régime [MS-88]

11:00 Cody HYNDMAN, *Concordia University*

Pricing Derivatives in Affine Term Structure Models • Fixation des prix de dérivées dans des modèles du terme de structure affine [MS-89]

11:30 Anatoliy SWISHCHUK, *University of Calgary*

Modelling and Pricing of Variance Swaps for Multi-Factor Stochastic Volatilities with Delay • Modélisation et fixation des prix de swaps de variance pour volatilités stochastiques à plusieurs facteurs avec délai [MS-90]

10:30–12:00 Session 06C

AA-1043

Statistical Methods for Correlated Data

Méthodes statistiques pour données corrélées

Invited Paper Session • Présentations sur invitation

Session sponsored by the *Committee on Women in Statistics* and the *Caucus for Women in Statistics*
Session parrainée par le *Comité des femmes en statistique* et le *Caucus des femmes en statistique*

Organizer and Chair • Responsable et présidente: Tulay KORU-SENGUL, *McMaster University*

10:30 Gerarda DARLINGTON, *University of Guelph*, Allan DONNER, *University of Western Ontario*

Meta-analysis of Cluster Randomization Trials for Binary Outcomes • Méta-analyse d'essais randomisés par grappes pour réponse binaire [MS-91]

11:00 Joanna FLEMMING, *Dalhousie University*

Tribute to a Champion of the Oceans • Hommage à un champion des océans [MS-92]

11:30 Stephanie DIXON & Gerarda DARLINGTON, *University of Guelph*

Modeling Family-based Genetic Case-control Study Designs in the Presence of a Continuous Time Variable • Modélisation de plans d'études cas-témoins génétiques basées sur la famille pour une variable continue [MS-93]

10:30–11:30 Session 06D

IIC-2001

Isobel Loutit Invited Address on Business and Industrial Statistics

Allocution Isobel Loutit sur la statistique industrielle et de gestion

Special Session • Conférence spéciale

Session sponsored by the *Business and Industrial Statistics Section*

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

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

10:30 Bovas ABRAHAM, *University of Waterloo*

Statistics in Business and Industry: A Historical Perspective and Ideas for Implementation • La statistique dans l'industrie : une perspective historique et des idées de mise en œuvre [MS-94]

10:30–12:00 Session 06E**SN-2105**

Forestry

Foresterie

Invited Paper Session • Présentations sur invitation

Session sponsored by the *National Institute for Complex Data Structures*Session parrainée par l'*Institut national pour les structures de données complexes*Organizer and Chair • Responsable et président: Reg KULPERGER, *University of Western Ontario*10:30 Cordy TYMSTRA, *Alberta Sustainable Resource Development*

Fire Simulation using the Prometheus Model • Simulation d'incendies à l'aide du modèle Prométhée [MS-95]

11:00 Douglas WOOLFORD, *University of Toronto*

Stochastic Lighting-fire Risk Models: Past, Present and Future • Modèles stochastiques du risque de feux causés par la foudre : passé, présent et futur [MS-96]

11:30 Jason MYERS, *University of Toronto*

Analysing Robustness of Forest Operational Plans • Analyser la robustesse des plans opérationnels pour les forêts [MS-97]

10:30–11:45 Session 06F**SN-2041**

Statistical Genomics

Génomique statistique

Contributed Paper Session • Séance de communications libres

Chair • Président: Duncan MURDOCH, *University of Western Ontario*10:30 Ahmed HOSSAIN, Joseph BEYENE & Andrew WILLAN, *University of Toronto and SickKids Research Institute*

An Approximate Likelihood Ratio Test for Testing Means of Generalized Logistic Distribution of Type III and Implications for Detecting Differential Expression in Microarrays • Un test du rapport de vraisemblance approximatif pour la moyenne d'une distribution logistique généralisée de type III et ses implications dans la détection d'expressions différentielles dans les micropuces [MS-98]

10:45 Yaling YIN, Miķelis BICKIS & Chris SOTEROS, *University of Saskatchewan*

A Further Note on Controlling the False Discovery Rate • Une note additionnelle sur le contrôle du taux de fausses découvertes [MS-99]

11:00 Shaheena BASHIR & Angelo CANTY, *McMaster University*

Bootstrapping of Gene Expression Data • Utilisation du bootstrap avec des données d'expression génétique [MS-100]

11:15 Jennifer ASIMIT & Shelley BULL, *Samuel Lunenfeld Research Institute, University of Toronto*

A Scan Statistic for Identifying Chromosomal Regions of Significant Association • Une statistique de balayage pour identifier les régions chromosomales d'association significative [MS-101]

11:30 Julia BRETTSCHEIDER, *Queen's University*

Statistical Quality Assessment for High-dimensional Genomic Measurements • Évaluation de la qualité statistique en présence de mesures génomiques de dimension élevée [MS-102]

12:00–13:30 Dîner • Lunch

13:30–14:30 Session 07A

AA-1043

Special Invited Session of the Probability Section

Allocution de l'invité d'honneur du groupe de probabilité

Special Session • Conférence spéciale

Session sponsored by the *Probability Section*

Session parrainée par le *Groupe de probabilité*

Organizer and Chair • Responsable et président: Murray BURKE, *University of Calgary*

13:30 Robert ELLIOTT, *University of Calgary*, Tak Kuen SIU, *Heriot Watt University*, Hailiang YANG, *University of Hong Kong*

Insurance Claims Modulated by a Hidden Marked Point Process • Primes d'assurance modulées par un processus ponctuel marqué caché [MS-103]

13:30–15:00 Session 07B

AA-1046

Statistical Genetics

Génétique statistique

Invited Paper Session • Présentations sur invitation

Chair • Président: Paul PENG, *Memorial University*

Organizers • Responsables: Paul PENG & J. C. LOREDO-OSTI, *Memorial University*

13:30 Laurent BRIOLLAIS & Yunhee CHOI, *Samuel Lunenfeld Research Institute*, Karen KOPCIUK, *Alberta Cancer Board*, Tariq HASAN, *University of New Brunswick*

Methodological Issues in Penetrance Estimation Studies • Problèmes méthodologiques dans les études d'estimation de pénétrance [MS-104]

14:00 Jinko GRAHAM & Linnea DUKE, *Simon Fraser University*

A Graphical Tool for Exploring Gene-environment Interaction in Case-parent Trios • Un outil graphique pour explorer l'interaction gène-environnement dans les trios cas-parent [MS-105]

14:30 Michael NEWTON, *University of Wisconsin*

Network-based Measurement of Gene-set Enrichment • Mesure basée sur les réseaux pour l'enrichissement d'un ensemble de gènes [MS-106]

13:30–15:00 Session 07C**IIC-2001**

Statistical Graphics and Image Analysis
Graphiques statistiques et analyse d'image

Invited Paper Session • Présentations sur invitation

Session sponsored by the *Business and Industrial Statistics Section*
Session parrainée par le *Groupe de statistique industrielle et de gestion*

Chair • Président: Alan WELSH, *Australian National University*

Organizers • Responsables: Peihua QIU, *University of Minnesota*

13:30 Jia LI, *Pennsylvania State University*

Real-Time Computerized Annotation of Pictures via D2-Clustering • Annotation informatisée d'images en temps réel par la classification D2 [MS-107]

14:00 Christine MUELLER, *University of Kassel*, Tim GARLIPP, *University of Oldenburg*

Detection of Circular Shapes of Fungi and Bacteria Colonies • Détermination de formes circulaires des colonies cryptogamiques et bactériennes [MS-108]

14:30 Peihua QIU, *University of Minnesota*

Nonparametric Estimation of the Point Spread Function for Blind Image Deblurring • Estimation nonparamétrique de la fonction de dispersion des points en correction d'images flées [MS-109]

13:30–15:00 Session 07D**SN-2105**

Statistics Applied to Hydrology and Climatology
Statistique appliquée à l'hydrologie et à la climatologie

Invited Paper Session • Présentations sur invitation

Organizer and Chair • Responsable et présidente: Anne-Catherine FAVRE, *Institut national de la recherche scientifique*

13:30 Éric PARENT & Jacques BERNIER, *MORSE AgroParisTech*, Vincent FORTIN, *Météorologie Canada*

Risk Assessment under Climatic Change : An Impossible Mission? • Évaluer le risque en situation de changement climatique : mission impossible ? [MS-110]

14:00 Jean-François ANGERS, *Université de Montréal*

Statistician's Role in Climate Change Research • Le rôle d'un statisticien dans la recherche en changement climatique [MS-111]

14:30 Luc PERREAULT, *Institut de recherche d'Hydro-Québec*, Jocelyn GAUDET, *Hydro-Québec*, James MERLEAU, *Institut de recherche d'Hydro-Québec*

Assessing Hydrological Forecasts at Hydro-Québec • Évaluation des prévisions hydrologiques à Hydro-Québec [MS-112]

13:30–15:00 Session 07E**SN-2109**

The National Institute for Complex Data Structures: Our Community in Scientific Collaborations

l'Institut national pour les structures de données complexes : Notre communauté dans des collaborations scientifiques

Session sponsored by the *National Institute for Complex Data Structures*

Session parrainée par l'*Institut national pour les structures de données complexes*

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

13:30 Alina DYACHENKO, Antonio CIAMPI & Russell STEELE, *McGill University*

Disease Models using Latent Classes in the Presence of Missing Data: The Example of Delirium • Modèles de maladies utilisant des classes latentes en présence de valeurs manquantes : l'exemple du délire [MS-113]

13:30 Antonio CIAMPI, Imad KHOURY & Godfried TOUSSAINT, *McGill University*, Sadora ANTONIANO, *IIMAS-UNAM, Mexico*, Carl MURIE & Robert NADON, *McGill University*

Proximity-graphs-based Clustering of Micro-array Probes • Classification de sondes de puces à ADN basée sur des graphes de proximité [MS-114]

13:30 Crystal LINKLETTER, *Simon Fraser University*

Variable Selection for Gaussian Process Models in Computer Experiments • Sélection de modèle pour modèles de processus gaussiens dans des expériences informatiques [MS-115]

13:30 Xiaofei SHI & Rafal KUSTRA, *University of Toronto*

Dimension Reduction of Genome-wide SNP Data by Sparse Principal Components Analysis • Réduction de dimension de données SNP sur tout le génome par analyse en composantes principales creuse [MS-116]

13:30 Nicholas MITSAKAKIS & Rafal KUSTRA, *University of Toronto*

Graphical Models for Microarray Expression Data based on Low-order Partial Correlations • Modèles graphiques pour données d'expression de micro-puces basés sur les corrélations partielles de bas ordre [MS-117]

13:30 Elizabete ALMEIDA, Michael DOWD & Joanna MILLS-FLEMMING, *Dalhousie University*

An Analysis of Bacterioplankton Abundance Levels using a Cyclical State Space Model • Une analyse des niveaux d'abondance de bactérioplancton utilisant un modèle d'espace des états cyclique [MS-118]

13:30 Wilson LU, *Acadia University*, Randy R. SITTER, *Simon Fraser University*

Disclosure Risk and Variance Estimation • Risque de divulgation et estimation de variance [MS-119]

13:30–15:00 Session 07F**SN-2041**

Longitudinal Data and Panel Counts

Données longitudinales et dénombrement de panels

Contributed Paper Session • Séance de communications libres

Chair • Présidente: Julie HORROCKS, *University of Guelph*

- 13:30 M. Tariqul HASAN, *University of New Brunswick*, Gary SNEDDON, *Memorial University*
Zero-inflated Poisson Regression for Longitudinal Data • Régression de Poisson gonflée à zéro pour données longitudinales [MS-120]
- 13:45 Xiaoqin XIONG & Joel DUBIN, *University of Waterloo*
Association Study of Longitudinal data Measured at Distinctive Timepoints • Étude d'association de données longitudinales mesurées à des temps distincts [MS-121]
- 14:00 Manish MADAN & Brajendra SUTRADHAR, *Memorial University*
Quasilikelihood Inferences in Gamma AR (1) Models for Longitudinal Data • Inférences basées sur la quasi-vraisemblance avec des modèles Gamma de type AR (1) pour des données longitudinales [MS-122]
- 14:15 Adefowoqe ODUEYUNGBO, Lehana THABANE & Maureen MARKLE-REID, *McMaster University*
Intention to Treat Analysis and the MCMC Method for Imputing Missing Values in Longitudinal Studies: An Example Using Data from a Proactive Nursing Health Promotion Intervention Study in Frail Older Home Clients • Analyses d'intention de traitement et méthode MCMC pour imputer les valeurs manquantes dans une étude longitudinale : un exemple provenant d'une étude proactive sur des personnes âgées de santé fragile vivant en résidence [MS-123]
- 14:30 Xingqiu ZHAO & Narayanaswamy BALAKRISHNAN, *McMaster University*
Multi-sample Comparison of Point Processes with Panel Count Data • Comparaison d'échantillons multiples de processus ponctuels avec données longitudinales de dénombrement [MS-124]

15:00–15:30 Pause-café • Coffee Break**IIC-AF1****15:30–17:00 Session 08A****AA-1043**

Survey Design and Analysis

Planification et analyse d'enquêtes

Contributed Paper Session • Séance de communications libres

Chair • Président: Pierre LAVALLÉE, *Statistics Canada/Statistique Canada*

- 15:30 ÉRIC LANGLET, *Statistics Canada/Statistique Canada*
Sample Size Determination for the Aboriginal Post-censal Surveys • Détermination de taille d'échantillon pour les enquêtes postcensitaires autochtones [MS-125]
- 15:45 Lori STRATYCHUK & Jean DUMAIS, *Statistics Canada/Statistique Canada*
Ontario Survey on the Prevalence of Hypertension in Adults: Practical Aspects in Sample Allocation and Collection • Enquête sur la prévalence de l'hypertension chez les Ontariens adultes : aspects pratiques sur l'allocation et la collecte d'un échantillon. [MS-126]
- 16:00 Louis-Paul RIVEST, *Université Laval*, Michel FERLAND, *Statistics Canada/Statistique Canada*
The Lavallée Hidiroglou Algorithm for Stratification in Business Surveys: A Summary and some new Developments • L'algorithme de Lavallée Hidiroglou pour la stratification dans les enquêtes entreprises : bilan et nouvelles avancées [MS-127]
- 16:15 Rebecca MORRISON, Sarah-Maude DION & Suzelle GIROUX, *Statistics Canada/Statistique Canada*

Sample Design of the Canadian Health Measures Survey • Plan d'échantillonnage de l'Enquête canadienne sur les mesures de la santé [MS-128]

16:30 Lenka MACH & Abdelnasser SAÏDI, *Statistics Canada/Statistique Canada*, Rob PETTAPEICE, *Statistics Canada/Statistique Canada & University of Waterloo*

A Study of the Properties of the Rao-Wu Bootstrap Variance Estimator: What Happens when Assumptions do not Hold? • Une étude des propriétés de l'estimateur Rao-Wu de la variance bootstrap lorsque certaines hypothèses ne sont pas vérifiées [MS-129]

15:30–16:30 Session 08B

SN-2041

Stochastic Partial Differential Equations and their Applications

Equations aux dérivées partielles stochastiques et leurs applications

Invited Paper Session • Présentations sur invitation

Session sponsored by the *Probability Section*

Session parrainée par le *Groupe de probabilité*

Chair • Président: Hao YU, *University of Western Ontario*

Organizers • Responsables: Hao YU, *University of Western Ontario* & Shui FENG, *McMaster University*

15:30 Michael KOURITZIN, *University of Alberta*

On Particle Representation of Markov Processes • Sur la représentation particulière des processus de Markov [MS-130]

16:00 Feng-Yu WANG, *Beijing Normal University*

Harnack Inequality for Stochastic Porous Medium Equation and Applications • Inégalité d'Harnack pour l'équation stochastique de médium poreux et applications [MS-131]

15:30–16:30 Session 08C

IIC-2001

Inaugural Invited Lecture of the National Institute for Complex Data Structures

Allocution inaugurale de l'invité d'honneur de l'Institut national sur les structures de données complexes

Special Session • Conférence spéciale

Session sponsored by the *National Institute for Complex Data Structures*

Session parrainée par l'*Institut national pour les structures de données complexes*

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

15:30 Andrew LAWSON & Bo CAI, *University of South Carolina*

Latent Structure in Space-time Disease Maps: Links to Environmental Hazard Events • Structures latentes dans les cartes spatio-temporelles des maladies : liens aux événements présentant un risque pour l'environnement [MS-132]

15:30–16:30 Session 08D**SN-2109**Confidence Intervals and Regions
Intervalles et régions de confiance

Invited Paper Session • Présentations sur invitation

Organizer and Chair • Responsable et président: ÉRIC MARCHAND, *Université de Sherbrooke*15:30 William STRAWDERMAN, *Rutgers University*, ÉRIC MARCHAND, *Université de Sherbrooke*

Frequentist Coverage of Bayes Credible Regions • Taux de couverture fréquentiste des régions de confiance bayésienne [MS-133]

16:00 Dominique FOURDRINIER, *Université de Rouen*

Estimating the Confidence of a Confidence Region • Estimation du degré de confiance d'une région de confiance [MS-134]

15:30–16:30 Session 08E**AA-1046**Graphical Methods and Data Visualization
Méthodes graphiques et visualisation de données

Contributed Paper Session • Séance de communications libres

Chair • Président: Yogendra CHAUBEY, *Concordia University*15:30 Linglong KONG & Ivan MIZERA, *University of Alberta*

Bivariate Quantile as Tukey Depth Contours: A Connection to Directional Quantile Regression • Courbes de niveaux pour la profondeur de Tukey à l'aide de quantiles bivariés : un lien avec la régression quantile directionnelle [MS-135]

15:45 Sohee KANG, *University of Toronto*, Laurent BRIOLLAIS, *Samuel Lunenfeld Research Institute, Mount Sinai Hospital*

Application of Bayesian Graphical Model to Study the Genetic Susceptibility to Breast Cancer • Application des modèles graphiques bayésiens pour l'étude de la susceptibilité génétique au cancer du sein [MS-136]

16:00 Matthew DUBINS, *York University*

William Playfair and Bias in Statistical Graphics • William Playfair et le biais dans les graphiques statistiques [MS-137]

16:15 Mikēlis BICKIS & Eric NEUFELD, *University of Saskatchewan*

Visualizing Inequity • La visualisation d'iniquité [MS-138]

15:30–17:00 Session 08F**SN-2105**Business and Industrial Statistics
Statistique industrielle et de gestion

Contributed Paper Session • Séance de communications libres

Chair • Président: Gerarda DARLINGTON, *University of Guelph*

15:30 Fateh CHEBANA & Taha OUARDA, *Institut National de la Recherche Scientifique*

Homogeneity testing in multivariate regional frequency analysis • Test d'homogénéité en analyse fréquentielle régionale multivariée [MS-139]

15:45 K Raneeth THIAGARAJAH, *Illinois State University*

Adjusted Score Test for the Homogeneity of Gamma Populations • Test de cotes ajusté pour l'homogénéité de populations Gamma [MS-140]

16:00 Min YU, *BC Hydro*

Electricity Time of Use Rates Impact Analysis in BC • Analyse de l'impact des taux de temps d'utilisation de l'électricité en Colombie-Britannique [MS-141]

16:15 Amanda LAFONTAINE & Klass WINSTON, *Ipsos-Reid*

The Impact Of Product Complexity on Pricing Research • L'impact de la complexité des produits sur la recherche en fixation des prix [MS-142]

16:30 Félix LABRECQUE-SYNNOTT & Jean-François ANGERS, *Université de Montréal*, Anne-Catherine FAVRE & Michel SLIVITZKY, *Institut national de la recherche scientifique*

Bimodal Hydrological Data Modeling • Modélisation de données hydrologiques bimodales [MS-143]

Mercredi 13 juin • Wednesday, June 13

08:30–09:30 Session 09A

IIC-2001

Pierre Robillard Award Address

Allocution du récipiendaire du prix Pierre-Robillard

Special Session • Conférence spéciale

Organizer and Chair • Responsable et président: Ed Susko, *Dalhousie University*

08:30–10:00 Session 09B

AA-1043

Genetic Epidemiology

Épidémiologie génétique

Invited Paper Session • Présentations sur invitation

Session sponsored by the *Biostatistics Section*Session parrainée par le *Groupe de biostatistique*Organizer and Chair • Responsable et présidente: Karen KOPCIUK, *University of Calgary*08:30 Dan NICOLAE, *University of Chicago*

Efficient Information Extraction in Genome-wide Association Studies • Extraction efficace de l'information dans les analyses portant sur tout le génome [MS-144]

09:00 Lei SUN, *University of Toronto*

The Multiplicity Problem in Genome-wide Association (GWA) Studies • Le problème de multiplicité dans les études sur tout le génome [MS-145]

09:30 Joanna BIERNACKA & Heather CORDELL, *Newcastle University*

A Composite Likelihood Approach for Identification of Genetic Variants Potentially Causally Associated with a Disease • Une approche basée sur la vraisemblance composite pour identifier les variantes génétiques présentant potentiellement un lien causal avec une maladie [MS-146]

08:30–09:30 Session 09C

AA-1046

Data with Extra Zeros Applied in Ecology

Données avec un surplus de zéros et applications en écologie

Invited Paper Session • Présentations sur invitation

This session is dedicated to the memory of Ram Myers (1952–2007). Ram was a world-renowned fisheries scientist who pioneered the use of sophisticated statistical methodology to analyse the state of the world fisheries.

Cette séance est dédiée à la mémoire de Ram Myers (1952–2007). Ram était un scientifique halieutique reconnu mondialement qui a été un pionnier dans l'utilisation de méthodes statistiques sophistiquées pour l'analyse de l'état des pêcheries dans le monde.

Organizer and Chair • Responsable et présidente: Joanna MILLS FLEMMING, *Dalhousie University*

08:30 Alan WELSH, *Australian National University*, Simon BARRY, *CSIRO Mathematics and Information Systems*

Generalized Additive Models for Zero Inflated Counted Data • Modèles additifs généralisés pour données de dénombrement gonflés à zéro [MS-147]

09:00 Eva CANTONI, *University of Geneva*, Joanna MILLS-FLEMMING, *Dalhousie University*, Alan WELSH, *Australian National University*

Zero-inflated Clustered Count Data • Données de comptage en grappes avec sur-représentation de zéros [MS-148]

08:30–10:30 Session 09D

SN-2109

Latest Developments in Small Area Estimation, Theory and Practice

Derniers développements pour l'estimation de petites régions : théorie et aspects pratiques

Invited Paper Session • Présentations sur invitation

Session sponsored by the *Survey Methods Section*

Session parrainée par le *Groupe des méthodes d'enquête*

Chair • Président: Jack GAMBINO, *Statistics Canada/Statistique Canada*

Organizers • Responsables: Susana RUBIN BLEUER, *Statistics Canada/Statistique Canada*

08:30 Parthasarathi LAHIRI, *University of Maryland, College Park*, Snigdhanu CHATTERJEE, *University of Minnesota*

Estimation of Mean Squared Prediction Error of an Empirical Best Predictor • Estimation de l'erreur quadratique pour une prévision empirique optimale [MS-149]

08:50 Narasimha PRASAD & Subhash LELE, *University of Alberta*

Small Area Prediction using Convex Combination of Parametric and Non-parametric Models • Prévision pour petits domaines en utilisant une combinaison convexe de modèles paramétriques et non paramétriques [MS-150]

09:10 Susana RUBIN BLEUER, Serge GODBOUT & Yves MORIN, *Statistics Canada/Statistique Canada*

Evaluation of Small Domain Estimators for the Survey of Employment, Payrolls and Hours • Évaluation d'estimateurs pour petits domaines pour l'Enquête sur l'emploi, la rémunération et les heures de travail [MS-151]

11:30 Discussant • Modérateur

J.N.K.RAO, *Carleton University*

08:30–10:00 Session 09E

SN-2105

Nonparametric Curve Estimation

Estimation non paramétrique de courbes

Contributed Paper Session • Séance de communications libres

Chair • Président: Bruce SMITH, *Dalhousie University*

08:30 Hyung-Tae (Kiefer) HA, *Kyungwon University*

Differentiated Logdensity Approximants • Approximants différenciés de la logdensité [MS-152]

08:45 Xuewen LU & Murray D. BURKE, *University of Calgary*

Nonparametric Estimation of Linear Functionals of a Bivariate Distribution under Univariate Censoring • Estimation non paramétrique de fonctions linéaires d'une distribution bivariée sous censure univariée [MS-153]

09:00 Jie SUN, *University of Calgary*

Polynomial Spline Estimation of Partially Linear Single-index Proportional Hazards Regression Models • Estimation par splines polynomiales de modèles de régression des risques proportionnels partiellement linéaires [MS-154]

09:15 David TOLUSSO & Richard COOK, *University of Waterloo*

Estimation of Mean Functions for Interval-censored Recurrent Event Data with Terminal Events • Estimation de la fonction moyenne pour des événements récurrents censurés par intervalles en présence d'événements terminaux [MS-155]

09:30 Jason NIELSEN & Charmaine DEAN, *Simon Fraser University*

Mixed Nonhomogeneous Poisson Process Spline Models for the Analysis of Recurrent Event Panel Data • Modèle spline de processus de Poisson non-homogène mixte pour l'analyse de données longitudinales d'événements récurrents [MS-156]

09:45 Yogendra CHAUBEY, *Concordia University*

Non-parametric Density Estimation for Length Biased Data • Estimation de densités non paramétriques pour des données biaisées par la longueur [MS-157]

08:30–09:45 Session 09F

SN-2041

Stochastic Processes and Their Applications

Processus stochastiques et leurs applications

Contributed Paper Session • Séance de communications libres

Chair • Présidente: Amanda HALLADAY, *Dalhousie University*

08:30 Katherine DAVIES, *University of Western Ontario*

Statistical Behaviour of a Spatial Renewal Process • Comportement statistique d'un processus de renouvellement spatial [MS-158]

08:45 Djilali AIT AODIA, *Université de Sherbrooke*, Paolo BULLA & Pietro MULIERE, *Bocconi University, Milan*

Reinforced Urn Processes Indexed by Binary Tree • Processus d'urnes renforcés indexés par un arbre binaire [MS-159]

09:00 Tingting GOU & Duncan MURDOCH, *University of Western Ontario*

Simulation of Extremes of Stochastic Differential Equations • Simulation des extrêmes d'équations différentielles stochastiques [MS-160]

09:15 Xiaowen ZHOU, *Concordia University*

Almost Sure Local Extinction for Super-Brownian Motion with Lebesgue Initial Measure • Extinction locale presque sûre pour le mouvement super-Brownien avec mesure initiale de Lebesgue [MS-161]

09:30 Veeresh GADAG & J. KUMAR, *Memorial University*, U. BALASOORIYA & C.K. LOW, *Nanyang Technological University, Singapore*

On some Properties of a Class of Generalized Gamma Distributions • Quelques propriétés d'une classe de distributions Gamma généralisées [MS-162]

10:00–10:30 Pause-café • Coffee Break

IIC-AF1

10:30–12:00 Session 10A

SN-2109

Special Invited Session sponsored by 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*

Session parrainée par le *Groupe des méthodes d'enquête*

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

10:30 Robert SMITH, Carolyn CAHILL & Michael BORDT, *Statistics Canada/Statistique Canada*

Environmental Surveys and Indicators – Conceptual and Methodological Challenges • Enquêtes et indicateurs environnementaux – Défis conceptuels et méthodologiques [MS-163]

11:30 Discussant • Modératrice

Julie HASS *Statistics Norway*

10:30–12:00 Session 10B

IIC-2001

Statistical Methods for Causal Inference

Méthodes statistiques pour l'inférence causale

Invited Paper Session • Présentations sur invitation

Session sponsored by the *Biostatistics Section*

Session parrainée par le *Groupe de biostatistique*

Chair • Président: John KOVAL, *University of Western Ontario*

Organizer • Responsable: Robert PLATT, *McGill University*

10:30 Erica MOODIE & David A STEPHENS, *McGill University*

Quantifying Dose-response for a Continuous Treatment in the Presence of Non-compliance or Confounding • Évaluation de la réponse à la dose reçue pour un traitement en présence de conformité partielle [MS-164]

- 11:00 Robert PLATT, J.A. “Chris” DELANEY & Geneviève LEFEBVRE, *McGill University*
Variable Selection in Marginal Structural Models • Sélection de variable dans les modèles marginaux structuraux [MS-165]
- 11:30 Thomas TENHAVE, Jennifer FAERBER & Marshall JOFFE, *University of Pennsylvania*, Robert GALLOP, *Westchester University*
Causal Analyses of Post-randomization Effect Modification in Clinical Trials • Analyses causales de modification post-randomisation des effets en essais cliniques [MS-166]

10:30–12:00 Session 10C**SN-2041**

Matrices and Statistics I

Matrices et statistique I

Invited Paper Session • Présentations sur invitation

Organizer and Chair • Responsable et président: George STYAN, *McGill University*

- 10:30 Richard BRUALDI, *University of Wisconsin – Madison*
Ordering (0,1)-matrices, Contingency Tables, and Stochastic Matrices • Ordonner des matrices (0,1), des tableaux de contingence et des matrices stochastiques [MS-167]
- 11:00 Jeffrey HUNTER, *Massey University, Auckland, New Zealand*
Coupling and Mixing Times in Markov Chains • Couplage et temps de stationnarité des chaînes de Markov [MS-168]
- 11:30 Simo PUNTANEN, *University of Tampere, Finland*, Oskar Maria BAKSALARY, *Adam Mickiewicz University, Poland*, George P.H. STYAN, *McGill University*
Some Comments on the Authorship Matrix in Bibliometrics • Quelques commentaires sur la matrice de paternité littéraire en bibliométrie [MS-169]

10:30–12:00 Session 10D**SN-2105**

Point Processes

Processus ponctuels

Invited Paper Session • Présentations sur invitation

Chair • Présidente: Katherine DAVIES, *University of Western Ontario*Organizers • Responsables: Rolf TURNER, *University of New Brunswick*

- 10:30 Gail IVANOFF, *University of Ottawa*, Ely MERZBACH, *Bar Ilan University*, Mathieu PLANTE, *none*
A Compensator Characterization of Point Processes on Partially Ordered Metric Spaces • Une caractérisation compensatrice de processus ponctuels sur des espaces métriques partiellement ordonnés [MS-170]
- 11:15 James RAMSAY, *McGill University*
The Interplay between Event Times and Event Characteristics • L'interaction entre temps d'événements et caractéristiques d'événements [MS-171]

10:30–12:00 Session 10E**AA-1043**

Biostatistical Methods

Méthodes biostatistiques

Contributed Paper Session • Séance de communications libres

Chair • Président: Paul GUSTAFSON, *University of British Columbia*10:30 Jianan PENG, *Acadia University*, Chun-In LEE, *Memorial University*, Karolyn DAVIS, *Carleton University*

Stepwise Confidence Intervals for Monotone Dose-response Studies • Intervalles de confiance pas-à-pas pour les études de doses/réponses monotones [MS-172]

10:45 Yan YUAN, *University of Waterloo*

Estimation of Prediction Loss with Censored Survival Data • Estimation de la perte de prévision en présence de données de survie censurées [MS-173]

11:00 Yuanyuan LIANG & Kimmie Chough CARRIERE, *University of Alberta*

Optimal Crossover Designs: Investigating the Role of Baseline Measurements • Plans alternés optimaux : éclairer le rôle des mesures de base [MS-174]

11:15 Long Yang WU, *University of Waterloo*

A Bayesian Algorithm to Detect Genomic DNA Copy Number Changes in Array CGH Data • Un algorithme bayésien pour détecter les changements du nombre de copies d'ADN génomique dans des données de puce CGH [MS-175]

11:30 Yanqing YI, *University of Manitoba*

The Efficiency of Estimation for Response Adaptive Designs • L'efficacité de l'estimation sous des plans adaptatifs à la réponse [MS-176]

11:45 Weiwei XIONG, *McMaster University*

Central Venous Line-related Thromboembolism in Children with Lymphoreticular Malignancy • Maladie thromboembolique liée à l'emploi d'un cathéter central veineux chez les enfants atteints de lymphoréticulopathie [MS-177]

10:30–11:45 Session 10F**AA-1046**

Statistical Theory III

Théorie statistique III

Contributed Paper Session • Séance de communications libres

Chair • Président: Angelo CANTY, *McMaster University*10:30 Simon GUILLOTTE & François PERRON, *Université de Montréal*

Bayes, MCMC and Copulas • Bayes, MCMC et Copules [MS-178]

10:45 Devon LIN, *Simon Fraser University*, Wilson LU, *Acadia University*, Randy SITTER, *Simon Fraser University*

Fast approximately balanced bootstrap without construction • Bootstrap rapide approximativement équilibré sans construction [MS-179]

- 11:00 Marco POLLANEN, *Trent University*, Mei-Ling HUANG & Wai Kong YUEN, *Brock University*
An Efficient Randomized Quasi-Monte Carlo Algorithm for the Pareto Distribution • Un algorithme quasi-Monte-Carlo randomisé efficace pour la distribution Pareto [MS-180]
- 11:15 Marie-Hélène TOUPIN, Mhamed MESFIOUI & Jean-François QUESSY, *Université du Québec à Trois-Rivières*
Goodness-of-fit Tests for Copulas based on Spearman's Dependence Function • Tests d'adéquation pour les copules basés sur la fonction de dépendance de Spearman [MS-181]
- 11:30 Ali GHODSI, *University of Waterloo*, Dale SCHUURMANS, *University of Alberta*, Finnegan SOUTHEY, *Google Inc.*
Regularized greedy importance sampling • Échantillonnage d'importance régularisé avare [MS-182]

12:00–13:30 Dîner • Lunch**13:30–14:30 Session 11A****IIC-2001**

Address of the winner of the Centre de Recherches Mathématiques/Statistical Society of Canada Prize

Allocution du récipiendaire du prix Centre de recherches mathématiques/Société statistique du Canada

Special Session • Conférence spéciale

Organizer and Chair • Responsable et présidente: Charmaine DEAN, *Simon Fraser University*

14:30–15:00 Pause-café • Coffee Break**IIC-AF1****15:00–16:30 Session 12A****AA-1043**

Biostatistics

Biostatistique

Contributed Paper Session • Séance de communications libres

Chair • Président: Paramjit GILL, *UBC-Okanagan*

- 15:00 Taraneh ABARIN & Liqun WANG, *University of Manitoba*

Second-order Least Squares Estimation of Censored Regression Models • Estimation de modèles de régression censurés par moindres carrés du second ordre [MS-183]

- 15:15 Paul McNICHOLAS & Brendan MURPHY, *Trinity College Dublin*

Parsimonious Gaussian Mixture Models • Modèles de mélanges gaussiens parcimonieux [MS-184]

- 15:30 Peter SONG, *University of Waterloo*

Will Quadratic Inference Function (QIF) Take Over Generalized Estimating Equations (GEE) in Longitudinal Data Analysis? • Les fonctions d'inférence quadratiques remplaceront-elles les équations d'estimations généralisées pour les analyses de données longitudinales? [MS-185]

- 15:45 Baojiang CHEN, Grace YI & Richard COOK, *University of Waterloo*

Inverse Probability Weighted GEE with Auxiliary Progressive Processes • EEG pondérées par probabilités inverses avec processus progressifs auxiliaires [MS-186]

16:00 Qian ZHOU, *University of Waterloo*, Huazhen LIN, *Sichuan University, China*, Peter X.-K. SONG, *University of Waterloo*

Comparison of Kernel Estimating Equations with Global and Local Variance Structures in Time-varying Coefficient Models • Comparaisons d'équations estimatrices de noyaux à structures de variances globale et locale pour des modèles à coefficients variant avec le temps [MS-187]

15:00–16:30 Session 12B

IIC-2001

Applications of Empirical Likelihood in Survey Sampling

Applications de la vraisemblance empirique en échantillonnage

Invited Paper Session • Présentations sur invitation

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

15:00 Jae-kwang KIM, *Yonsei University*

Calibration Estimation using Empirical Likelihood in Unequal Probability Sampling • Estimation de la calibration utilisant la vraisemblance empirique dans le contexte de l'échantillonnage à probabilités inégales [MS-188]

15:30 Song CHEN, *Iowa State University*

Efficient Semiparametric Estimation with Surrogate Outcome Data • Estimation semi-paramétrique efficace en présence de données de substitution [MS-189]

16:00 Suojin WANG, *Texas A&M University*, Jiahua CHEN, *University of British Columbia*, Changbao WU, *University of Waterloo*

Empirical Likelihood Inferences under Measurement Errors with Complex Survey Data • Inférences par vraisemblance empirique avec des erreurs de mesures en présence de données d'enquête complexes [MS-190]

15:00–16:30 Session 12C

SN-2105

Matrices and Statistics II

Matrices et statistique II

Invited Paper Session • Présentations sur invitation

Chair • Président: Simo PUNTANEN, *University of Tampere, Finland*

Organizer • Responsable: George STYAN, *McGill University*

15:00 Hans Joachim WERNER, *University of Bonn*

On the Linear Aggregation Problem in the General Gauss-Markov Model • Sur le problème d'agrégation linéaire dans un modèle général de Gauss-Markov [MS-191]

15:30 S. Ejaz AHMED, *University of Windsor*, Kjell DOKSUM, *University of Wisconsin*, Shakhawat HOSSAIN, *University of Windsor*, Jinghong You, *University of North Carolina*

Delineating the Effect of Mis-specification in Partially Linear Models • Effets d'une mauvaise spécification des modèles partiellement linéaires [MS-192]

16:00 George STYAN, *McGill University*, Simo PUNTANEN, *University of Tampere, Finland*

An Illustrated Philatelic Introduction to “Matrices and Statistics” • Une introduction philatélique illustrée aux “matrices et à la statistique” [MS-193]

15:00–16:30 Session 12D

SN-2109

Robust Statistics

Statistiques robustes

Invited Paper Session • Présentations sur invitation

Organizer and Chair • Responsable et président: Matias SALIBIAN-BARRERA, *University of British Columbia*

15:00 Ivan MIZERA & Linglong KONG, *University of Alberta*

Bivariate Quantile Regression via Conditional Halfspace Depth Contours • Régression des quantiles bivariés au moyen de contours de profondeur de demi-espaces conditionnels [MS-194]

15:30 Julie ZHOU & Beiyan OU, *University of Victoria*

Robustness in designs for field experiments • Robustesse en planification d’expériences sur le terrain [MS-195]

16:00 David TYLER, *Rutgers University*

Exploring Multivariate Data via Multiple Scatter Matrices • Exploration de données multivariées au moyen de multiples matrices de dispersion [MS-196]

15:00–16:30 Session 12E

AA-1046

Biostatistical Modelling

Modélisation biostatistique

Contributed Paper Session • Séance de communications libres

Chair • Président: Patrick BROWN, *University of Toronto*

15:00 Rob DEARDON, *University of Guelph*

How much Data Does it Take to Parameterize an Epidemic Model? • Combien d’observations sont nécessaires à la paramétrisation d’un modèle d’épidémie ? [MS-197]

15:15 Eleanor PULLENAYEGUM, *McMaster University*, Cath LAM & Cedric MANLHIOT, *Hospital for Sick Children*, Brian FELDMAN, *Hospital for Sick Children and University of Toronto*

Sensitivity Analysis for Marginal Structural Models: Checking the Inverse Probability-of-treatment Weights • Analyse de sensibilité pour modèles structuraux marginaux : vérifier la pondération par probabilités inverses [MS-198]

15:30 Aurélie LABBE, *Université Laval*

Latent Class Models for Pedigree Data • Modèles avec classes latentes pour des données familiales [MS-199]

15:45 Giles HOOKER, *Cornell University*

Experimental Design for Qualitative Fitting • Plan d’expérience pour ajustement de modèles qualitatifs [MS-200]

16:00 Tobie SURETTE & Tobie SURETTE, *Department of Fisheries and Oceans*, Jacques ALLARD, *Université de Moncton*

A Two-component Regression Mixture Model with Covariate-dependent Weights for Bivariate Morphometric Data • Un mélange de régressions à deux composantes ayant des poids variables appliqué à des données morphométriques bivariées [MS-201]

16:15 Yun Hee CHOI & Laurent BRIOLLAIS, *Samuel Lunenfeld Research Institute*

Modelling Multistage Sampling of Familial Data Using a Composite Likelihood Method • Modélisation de l'échantillonnage à plusieurs niveaux de données familiales en utilisant une méthode de vraisemblance composite [MS-202]

15:00–16:30 Session 12F

SN-2041

Time Series

Séries chronologiques

Contributed Paper Session • Séance de communications libres

Chair • Président: Bashir KHAN, *St. Mary's University*

15:00 Mohammad ROKNOSSADATI & Mahmoud ZAREPOUR, *University of Ottawa*

Multivariate Autoregression of Order One with Infinite Variance • Autorégression multivariée d'ordre un avec variance infinie [MS-203]

15:15 Nalini MAHALINGAM & Alwell OYET, *Memorial University*

Score Test for Homogeneity of Variances in Nonparametric Regression with Autocorrelated Errors via Wavelets • Un test de cotes pour l'homogénéité des variances en régression non paramétrique avec erreurs autocorrélées au moyen d'ondelettes [MS-204]

15:30 Chafik BOUHADDIOUL, *United Arab Emirates University*, Jean-Marie DUFOUR, *Université de Montréal*

Tests of Causality Between two Multivariate Infinite-order Autoregressive Series • Test de causalité entre deux séries vectorielles autorégressives d'ordre infini [MS-205]

15:45 David THOMSON, *Queen's University*

Robust Multitaper Spectrum Estimates • Estimateurs de spectres robustes à filtes multiples [MS-206]

16:00 Azadeh MOGHTADERI, David J THOMSON & Glen TAKAHARA, *Queen's University*

Dynamic Spectrum Estimation for Nonstationary Time Series • Estimation spectrale dynamique pour séries chronologiques non stationnaires [MS-207]



NSERC NEWS SESSION
Monday, June 11, 2007
Location: IIC-2001
12:45 - 1:30 p.m.

Serge Villemure, Director, Chemistry, Engineering and Mathematical Sciences, will present and requests participant's input on the following topics:

Grant Selection Committees (GSC) Structure Review: The current structure of NSERC's discipline-based GSCs has been in place for more than 30 years. NSERC has recently undertaken a review of such structure and would like to update the community on the review's status as well as gather input on the process so far, and on future directions.

International Review of Discovery Grants Program: NSERC is commissioning an International Review to assess the merit of its approach to supporting research in the natural sciences and engineering, and the extent to which its Discovery Grants Program fosters and supports research excellence. NSERC will present the context of this study as well as gather feedback from the research community on the initial steps of the study.

Discovery Grants Competitions Results 2007: NSERC will present a brief report on the most recent Discovery Grants competition.

NOUVELLES DU CRSNG
Lundi, le 11 juin, 2007
Lieu : IIC-2001
12h45 - 13h30

Serge Villemure, directeur, Division de la chimie, du génie et des sciences mathématiques, couvrira les sujets suivants et recueillera vos commentaires :

Revue de la structure des Comités de sélection (CSS) : La structure actuelle de nos CSS est en place depuis plus de 30 ans. Le CRSNG a récemment entrepris une revue de cette structure et aimerait donner une mise à jour à la communauté sur cette revue, ainsi que recueillir les commentaires de celle-ci sur le processus en place, et sur ses orientations futures.

Revue internationale du Programme de subventions à la découverte : Le CRSNG a entrepris une revue internationale pour évaluer le mérite de son approche au soutien à la recherche en sciences naturelles et en génie, ainsi que le degré auquel le Programme de subventions à la découverte favorise et soutient l'excellence en recherche. Le CRSNG vous décrira le contexte de cette étude et recueillera les commentaires de la communauté de recherche sur les étapes initiales de celle-ci.

Résultats des concours 2007 des subventions à la découverte : Le CRSNG vous présentera un bref rapport sur le plus récent concours de subventions à la découverte.



**NSERC Workshop
TIPS TO PREPARE YOUR NEXT
DISCOVERY GRANT APPLICATION**

Monday, June 11, 2007

Location: AA-1043

5:00 - 6:00 p.m.

NSERC representatives and Grant Selection Committee members invite you to attend an interactive workshop with useful advice for the preparation of your next NSERC application, statistics from the 2007 competition and answer your questions on the functioning of the Statistical Sciences Grant Selection Committee.

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.

**Atelier offert par le CRSNG
CONSEILS POUR L'ÉLABORATION
DE VOTRE PROCHAINE DEMANDE
DE SUBVENTION À LA DÉCOUVERTE**

Lundi, le 11 juin, 2007

Lieu :: AA-1043

17h00 - 18h00

Des employés du CRSNG et des membres des comités de sélection des subventions vous invitent à un atelier en mode interactif avec des conseils utiles pour l'élaboration de votre prochaine demande de subvention au CRSNG, les statistiques pour le concours 2007 et répondront à vos questions sur le fonctionnement sur le comité de sélection des subventions en statistiques.

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.

Abstracts • Résumés

Session 0 **Dimanche 10 juin • Sunday, June 10, 14:00–17:00**

IIC-AF1

Contributed Poster Session

Séance d'affichage

[MS-1]

Development of Exact Confidence Intervals for the Ratio of Two Poisson Oil Spill Rates

Développement d'intervalles de confiance exacts pour le rapport de taux de déversement d'hydrocarbures modélisés par des variables poissoniennes

William HARPER, *Otterbein College*, Ted ESCHENBACH, *TGE Consulting*

An extensive study of offshore oil spills in the Gulf of Mexico with extensions to the northern border of Alaska involved the comparison of two Poisson oil spill rates. This paper develops Excel functions that compute exact confidence intervals for the ratio of the two Poisson means. Both the statistical and Excel VBA work are detailed with tips for creating Excel functions. These functions along with needed exact binomial Excel functions are available free on the web at <http://www.otterbein.edu/home/fac/WLLVHRPR/>. The functions are illustrated with actual Gulf of Mexico oil spill data.

Une étude approfondie de déversements d'hydrocarbures dans le Golfe du Mexique avec extensions à la frontière nord de l'Alaska impliquait la comparaison de deux taux poissoniens de déversements d'hydrocarbures. Cet article développe des fonctions Excel calculant des intervalles de confiance exacts pour le rapport de deux moyennes poissoniennes. Le travail statistique et informatique est détaillé, avec des conseils pour la création de fonctions Excel. Ces fonctions, ainsi que les fonctions binomiales requises, sont disponibles gratuitement sur le web à <http://www.otterbein.edu/home/fac/WLLVHRPR/>. Les fonctions sont illustrées avec des données réelles de déversement d'hydrocarbures dans le Golfe du Mexique.

[MS-2]

Best Alpha Level of Significance and Selection Statistics in Backward Stepwise Logistic Regression

Meilleur seuil de signification alpha et statistiques de sélection dans la régression logistique pas-à-pas en arrière

Qinggang WANG & John KOVAL, *University of Western Ontario*

SAS/IML programs were written to provide Monte Carlo simulations to determine the best alpha level for the Chi-square stopping criterion and three selection criteria: Log-likelihood ratio statistic (LR), Score statistic (SC) and Wald's statistic (WD). Performance was evaluated using Efron's estimated true error rate of prediction. In the multivariate

Des programmes SAS/IML ont été écrits afin de fournir des simulations de Monte-Carlo pour déterminer le meilleur seuil alpha pour le critère d'arrêt chi-deux et trois critères de sélection : la statistique du rapport de log-vraisemblances (LR), la statistique de score (SC) et la statistique de Wald (WD). La performance fut évaluée en utilisant l'estimateur d'Efron du taux réel d'erreur de prévision. Dans le cas nor-

normal case, the best alpha decreased significantly with the number of predictor variables for LR and SC. In the multivariate binary case, it decreased significantly with both the number and correlation of predictor variables for three selection criteria. An overall recommendation is that the LR or SC with alpha between 0.20 and 0.40 should be used in backward stepwise logistic regression.

mal multivarié, le meilleur alpha diminuait significativement avec le nombre de variables prédictives pour LR et SC. Dans le cas multivarié binaire, il diminuait significativement avec le nombre et la corrélation des variables prédictives pour les trois critères de sélection. Une recommandation globale est d'utiliser le LR ou SC avec un alpha compris entre 0.20 et 0.40 dans la régression logistique pas-à-pas en arrière.

[MS-3]

Analysis of the Blanding's Turtle's Growth

Analyse de la croissance de la tortue de Blanding

Yingru (Celia) HUANG, Ying ZHANG & Paul CABILLO, *Acadia University*

Do you think a Blanding's turtle's different scutes (sections on the shell) grow at the same rate? Nobody has attempted to answer this question before. We will cover this and also explore pairwise homogeneity. We use scute width as a growth measurement to model Blanding's turtles' growth curves by fitting exponential and polynomial models. For residual diagnostic check of the growth models, Ljung-Box Tests and Kendall's Tau Tests are implemented. In order to determine how temperature affects Blanding's turtles' growth, several statistical modeling techniques will be discussed, including Non-Linear Mixed Effect Model and Linear Mixed Effect Model.

Pensez-vous que les différentes sections de la carapace de la tortue de Blanding croissent à la même vitesse ? Personne n'a tenté de répondre à cette question auparavant. Nous allons couvrir ce problème et explorer l'homogénéité deux-à-deux. Nous utilisons la largeur des sections de la carapace comme mesure de croissance pour modéliser les courbes de croissance de tortues de Blanding en ajustant des modèles exponentiels et polynomiaux. Des tests de Ljung-Box et du Tau de Kendall sont mis en oeuvre afin d'effectuer un diagnostic à partir des résidus de ces modèles. Afin de déterminer comment la température affecte la croissance de tortues de Blanding, nous discutons de plusieurs techniques de modélisation statistique, incluant le modèle non-linéaire à effets mixtes et le modèle linéaire à effets mixtes.

[MS-4]

Evaluation and comparison of confidence interval construction methods for the kappa statistic in the presence of rater bias

Évaluation et comparaison de méthodes de construction d'intervalles de confiance pour la statistique kappa en présence de biais dû au juge

Sameer PARPIA, John KOVAL & Allan DONNER, *University of Western Ontario*

The use of the kappa to measure interrater agreement in reliability studies is widely used. It is reasonable to assume that observers do not have homogenous rates in their judgements. Confidence intervals which account for violation of homogeneity have not been developed. This research compares the robustness of two methods, Goodness-of-fit and Modified Wald methods for various degrees of heterogeneity and sample sizes. Evaluation was done in terms of bias and coverage. Both methods show satisfactory results when rater bias is below 10. However, as rater bias increases, the Modified Wald method using Cohen's kappa performs better.

L'utilisation du kappa de Cohen pour mesurer l'accord inter-juges dans les études de fiabilité est largement répandue. Il est raisonnable de supposer que les observateurs n'ont pas de taux homogènes dans leurs jugements. Des intervalles de confiance prenant en considération le non respect de l'hypothèse usuelle d'homogénéité n'ont pas été développés. Cette recherche compare la robustesse de deux méthodes, la méthode de l'adéquation de l'ajustement et la méthode de Wald modifiée, pour différents degrés d'hétérogénéité et différentes tailles échantillonales. L'évaluation est faite en termes de biais et de couverture. Les deux méthodes donnent des résultats satisfaisants lorsque le biais des juges est inférieur à 10. Cependant, alors que le biais des juges augmente,

la méthode de Wald modifiée utilisant le kappa de Cohen performe mieux.

[MS-5]

Exact Inference For A Simple Step-stress Model From the Exponential Distribution Under Time Constraint
Inférence exacte pour un modèle simple à contraintes échelonnées à partir de la distribution exponentielle sous contraintes de temps

Qihao XIE, *Bombardier Aerospace*, Narayanaswamy BALAKRISHNAN, *McMaster University*, Debasis KUNDU, *India Institute of Technology*

In reliability and life-testing experiments, the researcher is often interested in the effects of extreme or varying stress factors such as temperature, voltage and load on the lifetimes of experimental units. Step-stress test, which is a special class of accelerated life-tests, allows the experimenter to increase the stress levels at fixed times during the experiment in order to obtain information on the parameters of the life distributions more quickly than under normal operating conditions. In this paper, we consider the simple step-stress model from the exponential distribution when there is time constraint on the duration of the experiment. We derive the maximum likelihood estimators (MLEs) of the parameters assuming a cumulative exposure model with lifetimes being exponentially distributed. The exact distributions of the MLEs of parameters are obtained through the use of conditional moment generating functions. We also derive confidence intervals for the parameters using these exact distributions, asymptotic distributions of the MLEs and the parametric bootstrap methods, and assess their performance through a Monte Carlo simulation study. Finally, we present two examples to illustrate all the methods of inference discussed here.

Dans les expériences de fiabilité et de durée de vie, le chercheur est souvent intéressé aux effets de facteur de stress extrêmes ou variables tels la température, le voltage, et la charge sur la durée de vie d'unités expérimentales. Les tests à contraintes échelonnées, qui forment une classe spéciale de tests de durée de vie accélérés, permettent à l'expérimentateur d'augmenter le niveau de stress à des temps fixes durant l'expérience de façon à obtenir de l'information sur les paramètres des distributions des temps de survie plus rapidement que sous des conditions normales d'opération. Dans cet article, nous considérons le modèle simple à contraintes échelonnées obtenu à partir d'une distribution exponentielle lorsqu'il y a des contraintes de temps sur la durée de l'expérience. Nous obtenons les estimateurs du maximum de vraisemblance (EMVs) des paramètres en supposant un modèle d'exposition cumulative avec temps de survie exponentiels. Les distributions exactes des EMVs des paramètres sont obtenus en utilisant les fonctions génératrices des moments conditionnelles. Nous obtenons également des intervalles de confiance pour les paramètres en utilisant ces distributions exactes, les distributions asymptotiques des EMVs et le bootstrap paramétrique, et évaluons leur performance au moyen d'une étude de simulation de Monte-Carlo. Finalement, nous présentons deux exemples afin d'illustrer les méthodes d'inférences mentionnées ici.

[MS-6]

Improved Estimation and Forecasts of Stock Maturities using Generalized Linear Mixed Effects Models
Estimation et prévision améliorée de l'échéance des actions utilisant des modèles linéaires généralisés à effets mixtes

Jeffrey DOWDEN, *Memorial University*, Noel CADIGAN, *Fisheries and Oceans Canada*

The mature component of a fish stock is usually referred to as the spawning stock biomass (SSB). Good estimates of the proportions mature-at-age (maturities) are required for good estimates of SSB. A common model used to estimate maturities is logistic regression; however, there are problems with this approach. Often the annual trends in co-

On réfère généralement à la composante mature d'un stock de poissons comme le stock de géniteurs. De bons estimés pour les proportions de poissons matures-à-cet-âge (maturités) sont requis pour de bons estimés du stock de géniteurs. Un modèle fréquemment utilisé pour estimer les maturités est la régression logistique ; cependant, il y a certains problèmes avec cette approche. Souvent les tendances an-

hort maturities are fairly smooth. We explore the autocorrelation structure in cohort maturities using a generalized linear model with mixed effects to improve estimates and forecasts of maturities. We examine two important commercial species in the Northwest Atlantic, Atlantic cod and American plaice.

nelles dans les maturités d'une cohorte sont raisonnablement lisses. Nous explorons la structure d'autocorrélation des maturités d'une cohorte en utilisant un modèle linéaire généralisé à effets mixtes pour améliorer les estimés et prévisions des maturités. Nous examinons deux espèces commercialement importantes dans l'Atlantique Nord-Ouest, la morue de l'Atlantique et la plie américaine.

[MS-7]

Spatio-temporal Trends for Treated Asthma in the Canadian Province of British Columbia

Tendances spatio-temporelles pour l'asthme traité dans la province canadienne de Colombie-Britannique

Paramjit GILL, *UBC Okanagan*, Charmaine DEAN, *Simon Fraser University*, Bob PROSSER & Bruce CARLETON, *University of British Columbia*

We study the spatial variation and identify areas of greatest risk with respect to asthma-related outcomes in BC. We analyse the magnitude of variation and assess the role of contributing factors at the local health area (LHA) level. Key goals are to provide LHA-specific estimates of the rates and to provide estimates of covariate effects. Ecological regression methods are used to relate geographical variations in treated asthma rates to the geographical distribution of suspected risk factors. Asthma prevalence rates at the LHA level are compared using log-linear Poisson Markov random field models to accommodate local spatial correlation.

Nous étudions la variation spatiale et identifions les régions de plus grand risque relativement aux résultats liés à l'asthme en Colombie-Britannique. Nous analysons l'ordre de grandeur de la variation et évaluons le rôle des facteurs contribuant au niveau des régions de santé locale (LHA). Les objectifs clés sont de fournir des estimés spécifiques aux LHA des taux et de fournir des estimés des effets des covariables. Des méthodes de régression écologiques sont utilisées pour mettre en lien les variations géographiques dans les taux d'asthme traité et la distribution géographique des facteurs de risque supposés. Les taux de prévalence d'asthme au niveau LHA sont comparés en utilisant des modèles de champs aléatoires de Markov log-linéaires de Poisson pour accommoder la corrélation spatiale.

[MS-8]

Confidence Intervals for the Mean of a Zero-inflated Distribution

Intervalle de confiance pour la moyenne d'une distribution gonflée à zéro

Matthew DAVIS, *University of Calgary*, David FLETCHER, *University of Otago*

For many research studies in ecology and fisheries, the data are skewed and contain a high proportion of zeros. This can lead to using a model involving a zero-inflated distribution. We compare different methods for calculating confidence intervals of the mean of such a distribution when the sample size is not large enough for a normal-based interval to be reliable. Our focus is on two distributions commonly used in these application areas, a truncated negative binomial distribution, and the log-normal distribution. The former is applicable for abundance, the later for biomass.

Pour plusieurs recherches en écologie et en pêche, les données sont asymétriques et contiennent une proportion élevée de zéros. Ceci peut mener à l'utilisation d'un modèle impliquant une distribution gonflée à zéro. Nous comparons différentes méthodes de calcul d'intervalles de confiance pour la moyenne de telles distributions lorsque l'échantillon est trop petit pour qu'un intervalle basé sur l'approximation normale soit fiable. Nous nous concentrons sur deux distributions souvent utilisées dans ces domaines d'application, une distribution binomiale négative tronquée et la distribution lognormale. La première est applicable à l'abondance, la seconde à la biomasse.

[MS-9]**Robust Estimates of the Correlation Parameter for the Bivariate Normal with Extensions to Elliptical Distributions****Estimés robustes du paramètre de corrélation de la normale bivariée et extension aux distribution elliptiques**Matthew DAVIS, *University of Calgary*

When there are outliers present in a bivariate normal sample the estimation of the correlation with Pearson's product moment correlation coefficient can be extremely biased. Here a collection of non-parametric based estimates of the correlation parameter will be introduced and their bias in the presence of outliers will be found to be much less than that of the standard estimate. Some of these estimates will be shown to be appropriate to estimate the association parameter in the larger family of elliptical distributions.

Lorsque des observations aberrantes sont présentes dans un échantillon normal bivarié, l'estimation de la corrélation avec le coefficient de corrélation des moments mixtes de Pearson peut être extrêmement biaisée. Ici, une collection d'estimés du coefficient de corrélation basés sur une approche non paramétrique seront présentés et leur biais en présence d'observations aberrantes sera démontrée de loin inférieure à celle de l'estimé standard. Il sera démontré que certains de ces estimés sont appropriés afin d'estimer le paramètre d'association dans la famille plus large des distributions elliptiques.

[MS-10]**Hardy-Weinberg Disequilibrium (HWD) Analysis in Candidate Gene Association Studies of Heterogeneous Disorders: Solving the Problem of the Multiplicative Model****Analyse du déséquilibre d'Hardy-Weinberg dans les études d'association du gène candidat de désordres hétérogènes : résoudre le problème du modèle multiplicatif**Vaneeta GROVER, *Dalhousie University*, David COLE, *University of Toronto*, David HAMILTON, *Dalhousie University*

The information resident in genotyping of first-degree relatives (siblings) has been investigated. Expressions for the genotypic probabilities and the HWD coefficient in siblings and probands have been obtained. Two different scenarios, when the disease status of a sibling is positive and when it is negative, were considered leading to different results. In particular, for the multiplicative model the HWD coefficient was non zero for probands when the sibs were not affected. The results suggest that study of both probands and siblings may avoid a limitation inherent in the HWD model for case-control disease association testing in complex, heterogeneous disorders.

L'information résidant dans le génotypage de la parenté de premier degré (frères et soeurs) a été étudiée. Des expressions pour les probabilités génotypiques et pour le coefficient du déséquilibre d'Hardy-Weinberg des frères et soeurs et des sujets ont été obtenues. Deux scénarios différents, lorsque le statut de maladie d'un frère (ou d'une soeur) est positif et lorsqu'il est négatif, ont été considérés, menant à des résultats différents. En particulier, pour le modèle multiplicatif le coefficient HWD était non nul pour les sujets lorsque les frères et les soeurs n'étaient pas affectés. Les résultats montrent qu'une étude des sujets et de leurs frères et soeurs peut contourner une limite inhérente au modèle HWD pour des tests cas-témoins de l'association de la maladie dans le cas de désordres complexes et hétérogènes.

[MS-11]**Using MLE for the Concurrent Assessment of Interrater and Intrarater Reliability and a Comparison to the Method of Moments Approach****Utilisation du maximum de vraisemblance pour l'évaluation simultanée de la fiabilité inter- et intra-juge et comparaison avec l'approche de la méthode des moments**Joseph AMUAH & Michael ELIASZIW, *University of Calgary*

For continuous measurements, agreement is often estimated by the intraclass correlation coefficient (ICC). For balanced designs, concurrent estimation

Pour des mesures continues, l'accord est souvent mesuré par le coefficient de corrélation intraclass (CCI). Pour des plans équilibrés, l'estimation simultanée du CCI a été développée

of ICC has been developed using the Method of Moments (MoM) Approach. This study provides another approach via the Maximum Likelihood Estimation (MLE). Estimation of agreement and associated confidence intervals are undertaken. Using Monte Carlo simulation methods, the MLE approach is compared to the MoM approach using properties of estimators such as unbiasedness and consistency. Using small, moderate and large samples, the associated 95% confidence interval estimates are also assessed and compared using their coverage probabilities, shapes and mean interval widths.

[MS-12]

Extending Education Statistics with CART

Étendre les statistiques d'éducation au moyen d'arbres de classification et de régression

Xin MA, *University of Kentucky*

This paper explores the recently developed data mining technique of CART (classification and regression tree) as it extends the analytical power of education statistics in addressing research questions in education for which traditional statistical techniques show great limitations. Using data from the Longitudinal Study of American Youth (LSAY), this paper showcases how CART can provide unique insights into the data which are hidden by way of traditional statistical techniques. This paper also demonstrates that the combination of CART with other advanced education statistical techniques can substantially broaden research questions in education that researchers are now capable of asking and addressing.

en utilisant la méthode des moments (MM). Cette étude fournit une autre approche au moyen de l'estimation par maximum de vraisemblance (MV). L'estimation de l'accord est considérée et des intervalles de confiance associés sont obtenus. En utilisant des méthodes de simulation de Monte-Carlo, les approches MV et MM sont comparées, focalisant sur les propriétés des estimateurs telles que l'absence de biais et la convergence. En utilisant des échantillons de petite, moyenne et grande tailles, les intervalles de confiance de niveau 95% résultants sont évalués et comparés sur la base de leur probabilité de couverture, de leur forme et de leur largeur moyenne.

Cet article explore la technique d'exploration de données récemment développée CART (arbres de classification et régression) car elle étend la puissance analytique des statistiques d'éducation à des questions de recherche en éducation pour lesquelles les techniques statistiques traditionnelles avaient de grandes limitations. En utilisant des données de la Longitudinal Study of American Youth (LSAY), cet article présente comment CART peut fournir des éclaircissements uniques relatifs aux données et qui ne peuvent être détectées par les techniques statistiques traditionnelles. Cet article démontre aussi que la combinaison de CART avec d'autres techniques avancées de statistique en éducation peut substantiellement élargir les questions de recherche en éducation que les chercheurs sont maintenant capables de poser et de traiter.

[MS-13]

Growth/survival Trade-offs in Tropical Forests using Bayesian Hierarchical Models

Utilisation des modèles hiérarchiques bayésiens pour estimer la relation entre la croissance et la survie des espèces d'arbres dans les forêts tropicales

Patrick BROWN, *Cancer Care Ontario/Action Cancer Ontario*, Sabrina Russo & David COOMES, *University of Cambridge*, Mark REES, *University of Sheffield*

Tropical forests are characterised by the presence of a great diversity of species, each with different growth rates and average survival times. One ecological hypothesis of "live fast, die young" proposes that species face a growth/survival trade-off. To quantify the relationship between growth and survival, models for growth and survival of indi-

Les forêts tropicales sont caractérisées par une grande diversité d'espèces, chacune ayant des taux de croissance et temps de survie moyen différents. L'hypothèse écologique de "vivre rapidement, mourir jeune" suggère que les espèces doivent contrebalancer la croissance et la survie. Pour évaluer la relation entre la croissance et la survie, des modèles pour la croissance et la survie individuelle des arbres ont été

vidual trees are developed and modified to include species-level random effects. The models are applied to large datasets of census of tree diameter in tropical forest plots, with the size of the datasets and the complexity of the models being a computational challenge.

développés et modifiés pour inclure des effets aléatoires au niveau des espèces. Ces modèles ont été ensuite appliqués aux grandes bases de données représentant un recensement du diamètre des arbres de parcelles en forêt tropicale ; avec la taille des bases de données et la complexité des modèles effectués ce calcul devient un véritable défi.

Session 01A**Lundi 11 juin • Monday, June 11, 08:30–09:00****IIC-2001**

Inaugural Session
Séance inaugurale

Session 01B**Lundi 11 juin • Monday, June 11, 09:00–10:00****IIC-2001**

SSC Presidential Invited Address
Allocution de l'invité de la présidente de la SSC

[MS-14]

The Case-control Study: Origins and Modern Conception
L'étude cas-témoins : origines et conception moderne

Norman BRESLOW, *University of Washington*

Nineteenth and 20th century contributions to the case-control study are briefly described. Case-control and case-cohort designs are increasingly used today to limit ascertainment of expensive genomic data to the most informative participants in a cohort study. Published reports, however, are often based only on complete data records for cases and controls, thus ignoring substantial amounts of information. Analysis methods developed during the past 2+ decades, many now implemented in standard computer packages, may substantially improve conventional methods for some problems. Horwitz-Thompson estimation in semiparametric models, with sampling weights adjusted via calibration or estimation to improve efficiency, offers one robust approach.

Les contributions du 19^e et 20^e siècles à l'étude cas-témoins sont brièvement décrites. Les plans cas-témoins et cas-cohorte sont de plus en plus utilisés de nos jours pour limiter l'étude de données génomiques dispendieuses aux participants les plus informatifs d'une étude de cohorte. Les rapports publiés, cependant, sont souvent uniquement basés sur les données complètes pour les cas et les témoins, ignorant ainsi une quantité substantielle d'information. Des méthodes d'analyses développées aux cours des décennies récentes, dont plusieurs sont maintenant implémentées dans des programmes informatiques standards, peuvent substantiellement améliorer les méthodes conventionnelles face à certains problèmes. L'estimation Horwitz-Thompson de modèles semi paramétriques, avec pondération échantillonnale ajustée par calibration ou estimation pour améliorer l'efficacité, offre une approche robuste.

Session 02A Lundi 11 juin • Monday, June 11, 10:30–12:00 IIC-2001

Diagnostics for the Retrospective Problem in Sequential Population Analysis of Exploited Fish Stocks

Diagnostiques pour le problème rétrospectif dans l'analyse séquentielle des populations pour l'exploitation des stocks de poissons

[MS-15]

Statistical Properties of the rho Statistic for Retrospective Bias – A Simulation Analysis

Propriétés statistiques de la statistique rho pour le biais rétrospectif – une analyse par simulation

LARRY JACOBSON & CHRIS LEGAULT, *Northeast Fisheries Science Center*, BOB MOHN, *Fisheries and Oceans Canada*, AL SEAVER, *Northeast Fisheries Science Center*

Retrospective bias in fish stock assessment affects important estimates of recent fishing mortality and biomass. Simulations were used describe the statistical distribution and probability of Type I & II errors for “Mohn’s rho”, a statistic sometimes used to measure retrospective patterns. The statistical distribution of the statistic must be determined via simulation on a case-by-case basis.

Le biais rétrospectif dans l'évaluation des stocks de poissons affecte des estimés importants de la mortalité par pêche et de la biomasse. Des simulations ont été utilisées afin de décrire la distribution d'échantillonnage et les probabilités d'erreurs de type I et II pour le coefficient “rho de Mohn”, une statistique parfois utilisée afin de mesurer le biais rétrospectif. La distribution de la statistique doit être déterminée au moyen de simulations sur une base de cas par cas.

[MS-16]

Diagnostics for detecting the source of retrospective patterns

Diagnostiques détectant la source de tendances rétrospectives

Noel CADIGAN, *Fisheries and Oceans Canada*

Retrospective patterns arising from fitting a fish stock assessment model indicate a deviation from some model assumption about population processes or what the input data represent. Retrospective problems can be so severe that the model is considered to be too unreliable for stock assessment. There are many possible sources of retrospective patterns, and it is difficult in practice to determine which are more likely. We use local influence diagnostics to investigate whether small changes or perturbations to model inputs can remove or reduce retrospective patterns. These diagnostics may help determine more likely causes. This will be examined with simulated data.

Les profils rétrospectifs survenant suite à l'ajustement d'un modèle d'évaluation des stocks de poisson indiquent une déviation par rapport aux hypothèses du modèle sur les processus de la population ou sur ce que représentent les données saisies. Les problèmes rétrospectifs peuvent être si sévères que le modèle peut être considéré comme trop peu fiable pour l'évaluation des stocks. Il y a plusieurs sources possibles de profils rétrospectifs, et, en pratique, il est difficile de déterminer lesquelles sont les plus vraisemblables. Nous devons utiliser des diagnostics d'influence locale afin de déterminer si de petits changements ou perturbations aux intrants du modèle peuvent retirer ou réduire les profils rétrospectifs. Ces diagnostics peuvent aider à déterminer les causes les plus vraisemblables. Des données simulées illustreront cette technique.

Session 02B**Lundi 11 juin • Monday, June 11, 10:30–12:00****AA-1043**

Design and Analysis of Experiments for Complex Physical Systems

Planification et analyse d'expériences pour des systèmes physiques complexes

[MS-17]**Gaussian Process Models for High Dimensional Spaces****Modèles de processus gaussiens pour espaces de dimension élevée**Dave HIGDON, *Los Alamos National Laboratory*, Derek BINGHAM, *Simon Fraser University*, Brian WILLIAMS, *Los Alamos National Laboratory*

Gaussian process models have proven to be very useful in modeling computer simulation output. This is because many computer codes are essentially noiseless and respond very smoothly to changes in input settings. In a typical setting, the simulation output is a function of a p -dimensional input vector x . When the number of simulations carried out to train the GP model are much over 1000, standard modeling and fitting approaches become computationally burdensome. In this talk, we describe strategies we've found useful for fitting such models (with large p) when the number of simulation runs is large.

Les modèles de processus gaussiens se sont montrés très utiles dans la modélisation des résultats de simulation informatique. Ceci est dû au fait que plusieurs codes informatiques sont essentiellement sans bruit et ont une réponse très lisse aux changements dans les entrées. Dans une situation typique, le résultat de simulation est fonction d'un vecteur d'entrées x de dimension p . Lorsque le nombre de simulations effectuées pour entraîner le modèle PG est beaucoup plus que 1000, les approches usuelles de modélisation et d'ajustement deviennent lourdes en termes de calculs. Dans cet exposé, nous décrivons des stratégies que nous avons trouvées utiles dans l'ajustement de tels modèles (avec p grand) lorsque le nombre de simulations effectuées est grand.

[MS-18]**Modeling and Analysis of Dispersion Effects in Robust Design Experiments****Modélisation et analyse d'effets de dispersion dans des plans d'expériences robustes**Vijay NAIR, *University of Michigan*, Derek BINGHAM, *Simon Fraser University*

There is a growing literature on the analysis of dispersion effects, motivated primarily by robust design studies. We consider direct modeling of the response and show that the method has a serious deficiency. The dispersion effects measured by control-by-noise interactions can be estimated only up to scaling constants that are usually unknown. This has important implications for identifying important effects, for parameter design optimization, and for the use of combined-array designs. We discuss some alternative approaches for addressing them. The consequences of the scaling issue for mean-variance modeling and related approaches will also be briefly discussed.

Il y a une littérature croissante sur l'analyse des effets de dispersion, motivée principalement par les études à plans robustes. Nous considérons la modélisation directe de la réponse et montrons que la méthode a une lacune sérieuse. Les effets de dispersion mesurés par les interactions contrôle-bruit ne peuvent être estimés qu'à une constante d'échelle (généralement inconnue) près. Ceci a des implications importantes sur l'identification d'effets importants, l'optimisation paramétrique du plan, et l'utilisation de plans à tableaux combinés. Nous discutons de certaines approches alternatives pour les traiter. Les conséquences de la question de l'échelle pour la modélisation moyenne-variance et les approches liées seront également brièvement présentées.

[MS-19]

Latent Socio-spatial Process Model for Social Networks
Modèle de processus socio-spatial latent pour réseaux sociaux

Crystal LINKLETTER, *Simon Fraser University*

Bioterrorism and new epidemics, such as SARS, have increased interest in computer codes that simulate disease spread. Such simulators incorporate a social network structure. Developing a stochastic social network model to emulate such programs is challenging; social connections exhibit unique dependencies, such as clustering and transitive behavior, heuristically described as “a friend of a friend is a friend”. Here, we present a new socio-spatial process model that smooths the relationship between covariates/inputs and sample network connections, allowing inference for connection probabilities and generation of likely social networks. This is an important step toward modelling disease transmission.

Le bioterrorisme et les nouvelles épidémies, telles le SRAS, ont mené à un intérêt accru pour les programmes simulant la propagation des maladies. De tels simulateurs incorporent une structure de réseaux sociaux. Développer un modèle stochastique de réseaux sociaux pour émuler de tels programmes est difficile ; les liens sociaux ont des dépendances uniques, telles le regroupement et un comportement transitif, décrit heuristiquement par “les amis de mes amis sont mes amis”. Ici, nous présentons un nouveau modèle de processus socio-spatial lissant la relation entre les covariables/entrées et les connections du réseau échantillonnal, permettant une inférence sur les probabilités de connexion et la génération de réseaux sociaux vraisemblables. Ceci est une étape importante dans la modélisation de la transmission des maladies.

Session 02C**Lundi 11 juin • Monday, June 11, 10:30–12:00****SN-2109**

Generalized Quasilielihood Inference
Inférence par quasi-vraisemblance généralisée

[MS-20]

A Fully Standardized Mallow’s Type Quasi-likelihood Inference for Generalized Linear Models
Inférence standardisée de type de Mallow par quasi-vraisemblance pour modèles linéaires généralisés

Wasimul BARI & Brajendra SUTRADHAR, *Memorial University*

It is well known that one or more outlying points in the data may adversely affect the consistency of the quasi-likelihood (QL) or the likelihood estimators for the regression effects. Similar to the QL approach, the existing outliers resistant Mallow’s type QL (MQL) estimation approach may also produce biased regression estimators. As a remedy, by using a fully standardized score function in the MQL estimating equation, in this paper, we demonstrate that the fully standardized MQL (FSMQL) estimators are almost unbiased ensuring its higher consistency performance. Both count and binary responses subject to one or more outliers are used in the study. The small sample as well as asymptotic results for the competitive estimators are discussed.

Il est bien connu qu’une ou plusieurs valeurs aberrantes parmi les données peuvent affecter négativement la convergence d’estimateur de vraisemblance ou quasi-vraisemblance (QV) pour les effets de régression. Similairement à l’approche QV, l’approche de la quasi-vraisemblance de Mallow (QVM) résistante aux valeurs aberrantes peut aussi produire des estimateurs de régression biaisés. Nous pouvons pallier à ce problème en utilisant une fonction de score complètement standardisée dans l’équation estimatrice de la QVM. Nous démontrons dans cet article que les estimateurs de QVM complètement standardisés (QVMCS) sont presque sans biais, assurant une meilleure convergence. Des données binaires et de dénombrement sujettes à au moins une observation aberrante sont utilisées dans cette étude. Nous discutons des résultats asymptotiques et pour petits échantillons des estimateurs compétitifs.

[MS-21]**Analysis of Incomplete Longitudinal Survey Binary Data by Conditional Weighted Generalized Quasi-likelihood****Analyse par quasi-vraisemblance conditionnelle pondérée généralisée de données binaires incomplètes tirées d'enquêtes longitudinales**Milorad KOVACEVIC, *Statistics Canada/Statistique Canada*, Brajendra SUTRADHAR, *Memorial University*

We propose a conditional weighted generalized quasi-likelihood (CWGQL) approach based on an unbiased estimating equation. It yields consistent estimates for the mean regression effects when modeling longitudinal survey binary data where some data are missing at random. This approach takes the correlation structure of the repeated data into account which makes it an efficient approach in estimating the regression effects.

A simulation study shows that the CWGQL outperforms the existing inverse probability weighted (INPW) GEE approach with respect to both consistency and efficiency. We illustrated the studied methods by analyzing data from the Survey of Labour and Income Dynamics.

Nous proposons l'approche de quasi-vraisemblance conditionnelle pondérée généralisée (QVCPG) fondée sur une équation d'estimation sans biais. Cette approche produit des estimations convergentes pour les paramètres moyennes de modèles de régression pour l'analyse de données binaires d'enquêtes longitudinales où certaines données sont manquantes au hasard. L'approche QVCPG tient aussi compte de la structure de corrélation des données répétées, ce qui en fait une méthode efficace pour l'estimation des paramètres de régression. Une étude par simulation a montré que la méthode QVCPG est plus performante que l'approche EEGP fondée sur l'INPP aussi bien du point de vue de la convergence que de l'efficacité. Nous avons illustré cette étude en analysant des données de l'Enquête sur la dynamique du travail et du revenu.

[MS-22]**Generalized Quasilielihood Estimation of a Nonlinear Dynamic Model for Longitudinal Binary Data****Estimation par quasi-vraisemblance généralisée d'un modèle dynamique non-linéaire pour des données binaires longitudinales**Katrina ROGERS-STEWART, *Carleton University*, Brajendra SUTRADHAR, *Memorial University*, Patrick FARRELL, *Carleton University*

In the longitudinal set-up, it is common to collect repeated binary responses along with a set of time-dependent multidimensional covariates from a large number of independent individuals. Farrell and Sutradhar (2006) have proposed a nonlinear dynamic logistic model for such data that permits full ranges for the correlations between the responses. In the present research, we propose, implement, and evaluate a generalized quasi-likelihood estimation approach (see Sutradhar 2003) for the model. We compare this approach with a procedure based on maximum likelihood, which is known to produce estimators that are fully efficient.

Dans le cadre d'une étude longitudinale, il est fréquent que l'on recueille des réponses binaires répétées ainsi qu'un ensemble de covariables multivariées dépendantes du temps à partir d'un grand nombre d'individus indépendants. Farrell et Sutradhar (2006) ont proposé un modèle dynamique non linéaire logistique pour de telles données qui permet l'étendue complète des corrélations possibles entre les réponses. Dans la présente recherche, nous développons et évaluons une approche d'estimation par quasi-vraisemblance généralisée (voir Sutradhar 2003) pour ce modèle. Nous comparons cette approche à une procédure basée sur le maximum de vraisemblance menant à des estimateurs que nous savons efficaces.

Session 02D Lundi 11 juin • Monday, June 11, 10:30–12:00 AA-1046

Recent Developments in Multivariate Survival Data Analysis Methodology
Développements récents dans la méthodologie de l'analyse de données de survie multivariées

[MS-23]

Regression Analysis of Multivariate Interval-censored Failure Time Data

Analyse de régression de données multivariées de temps de défaillance censurées par intervalles

(Tony) Jianguo SUN & Manhua CHEN, *University of Missouri*, Xingwei TONG, *Beijing Normal University*

A number of authors have discussed regression analysis of multivariate or correlated failure time data when only right-censoring is present. This talk will consider the same problem but when one has to deal with interval-censoring, which means that the failure time of interest is observed or known only to belong to certain intervals rather than exactly or to be right-censored. Several models and their corresponding inference procedures will be discussed.

Plusieurs auteurs ont discuté de l'analyse de régression de données de temps de défaillance multivariées ou corrélées en présence de censure à droite. Cet exposé considère le même problème, mais en présence de données censurées par intervalles, ce qui veut dire que nous observons uniquement que le temps de défaillance d'intérêt appartient à un certain intervalle plutôt que sa valeur exacte. Nous discutons de plusieurs modèles et des procédures d'inférence correspondantes.

[MS-24]

Marginal Hazards Regression for Retrospective Studies within Cohort with Correlated Failure Time Data

Régression pour la fonction de risque marginale dans des études rétrospectives intra-cohortes avec temps de défaillances corrélés

Jianwen CAI & Sangwook KANG, *University of North Carolina at Chapel Hill*

A retrospective dental study was conducted to evaluate the degree to which pulpal involvement affects tooth survival. Due to the clustering of teeth, the survival times within each subject could be correlated and thus the conventional method for the case-control studies cannot be directly applied. We propose a marginal model approach for this type of correlated case-control within cohort data. Weighted estimating equations are proposed for parameter estimation. Asymptotic properties of the proposed estimators are investigated and their finite sample properties are assessed via simulations studies. The proposed method is applied to the aforementioned dental study.

Une étude dentaire rétrospective a été menée afin d'évaluer le degré auquel l'implication pulpaire affecte la survie dentaire. À cause du regroupement des dents, les temps de survie pour chaque sujet peuvent être corrélés et la méthode conventionnelle pour des études cas-témoins ne peut donc pas être directement appliquée. Nous proposons une approche par modèle marginal pour ce type de données corrélées cas-témoins à l'intérieur d'une cohorte. Des équations estimatrices pondérées sont proposées pour l'estimation des paramètres. Les propriétés asymptotiques des estimateurs sont étudiées, et leurs propriétés avec un échantillon fini sont évaluées au moyen de simulation. La méthode proposée est appliquée à l'étude dentaire mentionnée plus haut.

[MS-25]

Spatial Multi-state Transitional Models for Longitudinal Event Data

Modèles spatiaux transitionnels à états multiples pour données d'événements longitudinales

Charmaine DEAN, *Simon Fraser University*, Farouk NATHOO, *University of Victoria*

Multistate transitional models are useful for analysis in longitudinal studies where at any point in time, individuals may be said to occupy one of a

Les modèles transitionnels à états multiples sont utiles pour les analyses en études longitudinales où à chaque point dans le temps, les individus peuvent être dans un état parmi un en-

discrete set of states and interest centers on the transition process between states. We develop a hierarchical Bayesian model for the analysis of such longitudinal data when the processes corresponding to different subjects may be spatially correlated. Here, joint modelling of both spatial correlation as well as correlation between different transition rates is required and a multivariate spatial approach is employed. The methodology is applied to a Quebec study of revascularization intervention.

semble discret d'états possibles, et où notre intérêt est centré sur le processus de transition entre les états. Nous développons un modèle hiérarchique bayésien pour l'analyse de telles données longitudinales lorsque les processus correspondant aux différents sujets peuvent être spatialement corrélés. Ici, la modélisation conjointe de la corrélation spatiale et de la corrélation entre les différents taux de transition est requise et une approche spatiale multivariée est employée. Cette méthodologie est appliquée à une étude d'intervention en revascularisation au Québec.

Session 02E**Lundi 11 juin • Monday, June 11, 10:30–12:00****SN-2105**

Survey Analysis

Analyse d'enquêtes

[MS-26]

Modeling Association Between Two or More Multiple-response Categorical Variables**Modéliser l'association entre plusieurs variables catégorielles à réponses multiples**Thomas LOUGHIN, *Simon Fraser University Surrey*, Christopher BILDER, *University of Nebraska*

Multiple-response categorical variables (MRCVs) are common in surveys where respondents are instructed to “mark all that apply” from a list of items. They result in binary vectors. Questions about associations among MRCVs often relate to marginal associations. Fitting loglinear models to account for all possible outcomes of the binary vector is often impossible. Marginal generalized log-linear models are proposed that allow the associations between items of different MRCVs to vary across levels of the MRCVs in structured ways, while reducing the number of parameters to be estimated relative to a full loglinear model.

Les variables catégorielles à réponses multiples (VCRM) sont fréquentes dans les enquêtes où les répondants ont pour instruction de “cocher tous les items pertinents” à partir d'une liste d'items. Elles donnent des vecteurs binaires. Des interrogations sur les associations entre VCRM sont souvent liées aux associations marginales. Il est souvent impossible d'ajuster des modèles log-linéaires de façon à tenir compte de tous les résultats possibles du vecteur binaire. Nous proposons des modèles log-linéaires marginaux généralisés, qui permettent aux associations entre items de différentes VCRM de varier à travers les niveaux des VCRM de façon structurée, tout en réduisant le nombre de paramètres à estimer relativement à un modèle log-linéaire complet.

[MS-27]

What is the Relationship between Average Weekly Earnings and Average Monthly Earnings? A Latent Variable Approach**Quelle est la relation entre les gains hebdomadaires moyens et la paie mensuelle moyenne? Une approche par variable latente**Serge GODBOUT, *Statistics Canada/Statistique Canada*

In the Survey of Employment, Payrolls and Hours, the estimation of average weekly earnings is done using a regression estimator with average monthly earnings x as an auxiliary variable. The model currently used, a regression line between y and x , includes an intercept. The choice to use an intercept was based on empirical results from observed data.

Dans l'Enquête sur l'emploi, la rémunération et les heures de travail, l'estimation des gains hebdomadaires moyens se fait à l'aide d'un estimateur par régression utilisant la paie mensuelle moyenne x comme donnée auxiliaire. Le modèle actuellement utilisé a été choisi à partir de résultats empiriques sur les données observées. À partir de leur définition, nous poserons des hypothèses sur les variables x et y basées sur la

We propose a hypothesis on the distribution of x and y based on their definition and the latent variable z . We will calculate the conditional density function of the observed variables. Based on a simulation, we will evaluate the performance of different models.

présence d'une variable latente z . Nous développerons la distribution de densité conditionnelle entre les variables observées. Nous comparons la performance de différents modèles à partir de résultats de simulations.

[MS-28]

The GEE Methodology for Longitudinal Surveys with Missing Responses

La méthode des équations estimatrices généralisées pour les enquêtes longitudinales avec valeurs manquantes

Ivan CARRILLO GARCIA & Changbao WU, *University of Waterloo*

We present results on two critical aspects of the Generalized Estimating Equation (GEE) approach, consistency and variance of the GEE estimators, for analyzing longitudinal surveys under complex designs and with missing responses. We discuss the use of the hot-deck and the nearest neighbour imputation methods and in each case we address the consistency and the variance of the estimators. We contrast this approach with the reweighting method proposed by Robins *et al.*, for which the missing responses are handled through direct modelling of the missing probabilities. Data from the National Longitudinal Survey of Children and Youth is used for illustration.

Nous présentons des résultats par rapport à deux aspects critiques de l'approche des équations estimatrices généralisées (EEG), la convergence et la variance des estimateurs EEG, dans l'analyse d'enquêtes longitudinales sous des plans complexes avec des valeurs manquantes. Nous discutons l'utilisation des méthodes d'imputation du "hot deck" et du plus proche voisin, traitant dans chaque cas de la convergence et de la variance des estimateurs. Nous comparons cette approche avec la méthode de repondération proposée par Robins *et al.*, dans laquelle les valeurs manquantes sont traitées par une modélisation directe des probabilités manquantes. Les données de l'Enquête longitudinale nationale sur les enfants et les jeunes sont utilisées pour des fins d'illustration.

[MS-29]

Using Local Likelihood Density Estimation for Bias Correction at and near a Boundary

Utilisation d'estimation de densité par vraisemblance locale afin de corriger le biais à et autour d'une frontière

Norberto PANTOJA-GALICIA, *University of Waterloo*

The intuitive appeal of kernel density estimation makes this methodology an attractive candidate to obtain nonparametric estimates of a probability density function. In spite of the advantages presented by standard kernel density estimation techniques, however, a problem of increased bias at and near the boundary is present in certain scenarios. In this talk we will present a local likelihood estimator suitable for a density which is believed to have a value greater than zero at and near a boundary. Applications using complex survey data will be illustrated.

L'attrait intuitif de l'estimation de densité par noyaux fait de cette méthode un candidat attrayant pour l'obtention d'estimés non paramétriques d'une fonction de densité. Malgré les avantages des techniques usuelles d'estimation de densité par noyaux, un problème de biais accru à et autour d'une frontière se présente dans certains scénarios. Dans cet exposé, nous présentons un estimateur de vraisemblance locale convenable pour une densité supposée non nulle à et autour d'une frontière. Des applications utilisant des données d'enquête complexe seront illustrées.

[MS-30]**A Justification and Improvement of Woodruff's Method for Quantile Estimation from Survey Data
Une justification et amélioration de la méthode de Woodruff pour l'estimation de quantiles à partir de données d'enquêtes**Owen PHILLIPS & Avi C. SINGH, *Statistics Canada/Statistique Canada*

Woodruff's (1952) method for interval estimation of a quantile provides good coverage properties of CI for all quantiles even for those corresponding to low or high proportions. This is surprising because the CI for a proportion based on a normal pivotal is known to misbehave for low or high values of the true proportion. Using the estimating function (EF) confidence distribution approach, we provide a justification and give an alternative principled approach to variance estimation. Empirical results will be presented.

La méthode de Woodruff (1952) pour l'estimation par intervalles d'un quantile a de bonnes propriétés de couverture pour tous les quantiles, même ceux correspondant à des proportions faibles ou élevées. Ceci est surprenant parce que l'IC pour une proportion basée sur un pivot normal a un mauvais comportement pour des valeurs faibles ou élevées de la proportion réelle. En utilisant l'approche des fonctions estimatrices (FE) de la distribution de confiance, nous apportons une justification et donnons une approche suivant un principe alternatif pour l'estimation de la variance. Des résultats empiriques seront présentés.

[MS-31]**A Generalization of Fieller's Method for Improved Variance and Interval Estimation of Ratio of Means
Une généralisation de la méthode de Fieller pour une estimation améliorée de la variance et des intervalles de confiance pour le rapport de moyennes**Avi SINGH & Mike TAM, *Statistics Canada/Statistique Canada*

The usual Wald method based on the Taylor linearization for interval estimation of ratio of means of variables y and x is known to perform poorly for moderate or small sample sizes due to skewness in the distribution of the ratio of sample means. An improved confidence interval (CI) was proposed by Fieller (1940) using a normal pivotal based on the estimating function (EF) defined by the difference between the sample mean of y and the unknown ratio times the sample mean of x . In this paper, we propose improved interval and variance estimators using the EF-Confidence distribution approach.

La méthode de Wald usuelle pour l'estimation par intervalles d'un rapport de moyennes de variables x et y , basée sur la linéarisation de Taylor, performe mal pour des échantillons de taille faible ou moyenne à cause de l'asymétrie de la distribution du rapport de moyennes d'échantillons. Fieller (1940) a proposé un intervalle de confiance (IC) amélioré utilisant un pivot normal basé sur une fonction estimatrice (FE) définie par la différence entre la moyenne de l'échantillon de y et le rapport inconnu multiplié par la moyenne de l'échantillon de x . Dans cet article, nous proposons des estimateurs d'intervalle et de variance améliorés utilisant l'approche de la distribution FE-confiance.

Session 02F**Lundi 11 juin • Monday, June 11, 10:30–12:00****SN-2041**

Issues in Clinical Trials

Problématiques dans les essais cliniques

[MS-32]**Statistical Monitoring of Clinical Trials with Multivariate Response or Multiple Arms Using Repeated Confidence Bands****Suivi statistique d'essais cliniques avec réponse multivariée ou à plusieurs groupes en utilisant des bandes de confiance répétées**Lihui ZHAO & Joan HU, *Simon Fraser University*, Stephen LAGAKOS, *Harvard University*

Clinical trials with multivariate response or multiple arms have become increasingly common because of their potential efficiency and cost saving. Interim analyses of such studies are often guided by parametric assumptions for the underlying probability models. There are situations where it is not clear at the outset how the responses differ among the treatment groups and what kinds of differences are clinically meaningful. More flexible designs and monitoring procedures are therefore desirable. In this talk, we extend the repeated confidence bands approach (Hu and Lagakos, 1999) to studies with multivariate target function. We use a recent AIDS clinical trial to illustrate how to apply the multivariate repeated confidence bands (MRCB) approach in practice. This is a joint work with X. Joan Hu and Stephen W. Lagakos.

Les essais cliniques avec réponses multivariées ou à plusieurs groupes sont de plus en plus utilisés grâce à leur efficacité et à l'économie potentielle qu'ils représentent. Les analyses intermédiaires de telles études sont souvent guidées par des hypothèses paramétriques sur les modèles de probabilité sous-jacents. Il y a des situations où, au départ, il est impossible de voir clairement comment les réponses diffèrent d'un groupe à l'autre, ni quelles sortes de différences seront cliniquement significatives. Des plans et procédures de suivi plus flexibles sont donc désirables. Dans cet exposé, nous généralisons l'approche par intervalles de confiance répétées (Hu and Lagakos, 1999) à des études avec une fonction cible multivariée. Nous utilisons un essai clinique récent sur le SIDA pour illustrer comment appliquer l'approche des intervalles de confiance répétées multivariées (BCRM) en pratique. Ce travail fut réalisé conjointement avec X. Joan Hu et Stephen W. Lagakos.

[MS-33]**How to Deal with Patients Who Withdraw Data from Randomized Controlled Trials: Individual versus Public Rights in Scientific Research****Comment réagir face aux patients retirant leurs données d'essais cliniques randomisés : droits individuels et droits publics en recherche scientifique**Cheng Lin YE & Lehana THABANE, *McMaster University*

The issue of individual versus public rights in research has been widely debated by ethicists and legal scholars. However, the issue of how to deal with patients who withdraw their data from a randomized controlled clinical trial has received little or no attention. This presentation will discuss the implications of withdrawing data on study integrity and ways to handle the problem. We look for guidance from research ethics guidelines, published literature, legislations and suggestions from research ethics boards and professional organizations.

La question des droits individuels versus les droits publics en recherche a suscité de vifs débats chez les éthiciens et les érudits de la loi. Cependant, la question du traitement de patient retirant leurs données d'un essai clinique randomisé a reçu peu, voire pas du tout, d'attention. Cette présentation discutera des implications de retirer des données sur l'intégrité de l'étude et des façons de traiter ce problème. Nous cherchons conseil auprès des principes d'éthique de recherche, de la littérature publiée, de législations et de suggestions de comités d'éthiques et organisations professionnelles pour la recherche.

[MS-34]**Modeling Medication Compliance
Modéliser l'observance à un traitement**

Ying ZHANG & Paul CABILLO, *Acadia University*, Maja GRUBISIC, *University of Southern California*, Femida GWADRY-SRIDHAR, *University of Western Ontario*

Distributions of measurements of medication compliance are often skewed and clustered due to both the nature of the disease for which medication is being taken, and to the subjective nature of human decisions. Often, continuous compliance measures are dichotomized, thereby reducing their power. We propose a mixed distribution framework so as to model medication compliance using a flexible censored data analysis approach. For demonstration purposes, we use a data set from a clinical intervention trial which comprises 134 cases of congestive heart failure. Our approach can be generalized to other health research outcomes.

Les distributions des mesures d'observance à un traitement sont souvent asymétriques et en grappes à cause de la nature de la maladie nécessitant le traitement et de la nature subjective des décisions humaines. Souvent, les mesures continues d'observance sont dichotomisées, réduisant ainsi leur puissance. Nous proposons une structure de distribution mixte pour modéliser l'observance du traitement utilisant une approche flexible d'analyse de données censurées. Pour fins de démonstration, nous utilisons un jeu de données d'une étude d'intervention clinique comprenant 134 cas d'insuffisance cardiaque congestive. Notre approche peut être généralisée à d'autres résultats de recherches en santé.

[MS-35]**Clustered Progressive Multi-state Processes Under Incomplete Observation
Processus progressif en grappes à états multiples sous observation incomplète**

Rinku SUTRADHAR, *University of Toronto*, Samuel Lunenfeld Research Institute, Richard COOK, *University of Waterloo*

A conditionally Markov time non-homogeneous model is described for the analysis of clustered progressive multi-state processes under panel observation. Random effects are assumed to be a subject-specific vector arising from a continuous multivariate distribution. The variance-covariance matrix of the random effects distribution accommodates the degree of heterogeneity in the transition rates between subjects, as well as the correlation in the transition rates between different states. A Monte Carlo EM approach for estimating the parameters is developed. We apply our methods to a dataset of patients with psoriatic arthritis and characterize the course of damage in the joints of the hands.

Un modèle conditionnellement Markovien non homogène dans le temps est décrit pour l'analyse de processus progressifs en grappes à états multiples sous observations longitudinales. Nous supposons que les effets aléatoires sont un vecteur spécifique à chaque sujet provenant d'une distribution continue multivariée. La matrice de variance-covariance de la distribution des effets aléatoires s'adapte au degré d'hétérogénéité dans les taux de transition entre les sujets, ainsi que la corrélation dans les taux de transition entre les différents états. Une approche EM Monte-Carlo pour estimer les paramètres est développée. Nous appliquons nos méthodes à un jeu de données de patients affectés par des rhumatismes psoriasiques et caractérisons l'évolution des dégâts dans les articulations des mains.

[MS-36]**Dealing with Zero Events – Comparing Different Strategies Used in Meta-analysis
Traiter les études sans événements – comparaisons de différentes stratégies utilisées en méta-analyse**

Ji (Emmy) CHENG, Lehana THABANE, Jack SUN & Richard WHITLOCK, *McMaster University*

Dealing with zero-event studies is a common problem in meta-analysis. Though different strategies were proposed, there is no agreement in the litera-

Traiter des études sans événements est un problème courant en méta-analyse. Bien que différentes stratégies aient été proposées, il n'y a pas de consensus dans la littéra-

ture on which method works best. In this presentation, we discuss the pros and cons of different methods such as Mantel-Haenszel, Peto, Inverse Variance and Bayesian approach in handling this problem. We apply the methods in meta-analysis of surgical randomized controlled trials of the use of ASA versus no-ASA to reduce the rates of re-operations. The results show that the performance of the methods depends partly on the number of studies with zero events in either or both arms.

ture sur une méthode plus efficace. Dans cette présentation, nous discutons des avantages et inconvénients de différentes méthodes pour traiter ce problème, telles les approches de Mantel-Haenszel, de Peto, d'inverse de la variance et bayésienne. Nous appliquons ces méthodes à une méta-analyse d'essais cliniques randomisés chirurgicaux de l'utilisation contre la non utilisation de l'ASA pour réduire les taux de ré-opération. Les résultats montrent que la performance des méthodes dépend en partie du nombre d'études sans événements.

[MS-37]

Comparisons of Bayesian and Classical Methods in Analysis of Cluster Randomized Trial: The CHAT Trial
Comparaisons de méthodes bayésiennes et classiques dans l'analyse d'essais cliniques randomisés par grappes : l'étude CHAT

Jinhui MA, *McMaster University*

Cluster randomized trials are increasingly being used in health science research. However, the efficiency and consistency using different analytical methods have received little attention. Based on a pragmatic trial of community pharmacy blood pressure clinics linked with family medicine practices to improve the monitoring and management of high blood pressure among older adult, this presentation will discuss some traditional classical methods and Bayesian approach to analyze data from cluster randomized trials. We will also compare the results from the various analytical methods.

Les essais cliniques randomisés par grappes sont de plus en plus utilisés dans la recherche en santé. Cependant, peu d'attention a été portée à l'efficacité et à la convergence lorsque différentes méthodes analytiques sont utilisées dans ce contexte. En se basant sur un essai clinique pragmatique sur la pression sanguine mesurée en pharmacie relié à des pratiques de médecine familiale pour améliorer la détection, et la gestion, de la haute pression chez les adultes plus âgés, cette présentation discutera de quelques méthodes classiques traditionnelles et de l'approche bayésienne pour analyser des données d'essais cliniques randomisés par grappes. Nous allons aussi comparer les résultats obtenus à partir des différentes méthodes analytiques.

Session 02G

Lundi 11 juin • Monday, June 11, 10:30–11:45

SN-1103

Statistical Theory I
Théorie statistique I

[MS-38]

Stein-type Estimation of the Mean: The Behrens-Fisher Problem
Estimation de type Stein de la moyenne : le problème de Behrens-Fisher

Bashir KHAN, *Saint Mary's University*

Testing the equality of the means of two normal populations when the two variances are unequal and unknown is called the Behrens-Fisher problem. For the two-sample case, let the p variate vectors X_{ij} be normally distributed with mean μ_i and covariance matrix Σ_i , where $i = 1, 2$ and $j = 1, 2, \dots, N_i$. Our problem is to estimate μ_1 when it is suspected that $\mu_1 = \mu_2$ may hold as-

Tester l'égalité des moyennes de deux populations normales lorsque les deux variances sont inégales et inconnues est appelé le problème de Behrens-Fisher. Pour le cas à deux échantillons, soit le vecteur de dimension p X_{ij} normalement distribué avec moyenne μ_i et covariance Σ_i , où $i = 1, 2$ et $j = 1, 2, \dots, N_i$. Notre problème consiste à estimer μ_1 lorsqu'on soupçonne que l'hypothèse $\mu_1 = \mu_2$ est valide, en supposant les deux matrices de covariance inégales et incon-

suming that the two covariance matrices are unequal and unknown. Although numerous procedures have been provided to solve this problem, we follow the Scheffe's procedure. We propose five estimators namely the unrestricted, restricted, preliminary test, Stein-type and the positive-rule for the estimation of μ_1 . The quadratic biases, mean squares error matrices, and the quadratic risk functions of the estimators are derived and their performances are compared via quadratic bias and risk function criteria.

nues. Malgré le fait que plusieurs procédures aient été proposées pour résoudre ce problème, nous utilisons la procédure de Scheffé. Nous proposons cinq estimateurs, soient les estimateurs non restreint, restreint, de test préliminaire, de type Stein et la règle positive pour l'estimation de μ_1 . Les biais quadratiques, matrices d'erreurs quadratiques moyennes, et les fonctions de risque quadratique des estimateurs sont obtenus et leurs performances comparées selon les critères du biais quadratique et de la fonction de risque.

[MS-39]

Order Statistics and Records from a Family of J -shaped Distributions

Statistiques d'ordre et records d'une famille de distributions en forme de J

Ahmad ZGHOUL, *Qatar University*

In this paper we study order statistics and record values from a family of J -shaped distributions. Expressions for moments and product moments of order statistics and records are given. Numerical computations for moments, product moments, and correlations of records are displayed. Some recurrence relations and bounds for moments of order statistics and records are derived. Based on a sample of lower record values, the maximum likelihood estimator for the shape parameter is obtained.

Dans cet article nous étudions les statistiques d'ordre et valeurs record d'une famille de distributions en forme de J . Des expressions pour les moments et moments mixtes des statistiques d'ordre et records sont données. Des calculs numériques pour les moments, moments mixtes et corrélations des records sont présentés. Quelques relations de récurrences et bornes inférieures pour les moments de statistiques d'ordres et records sont obtenues. À partir d'un échantillon de valeurs record inférieures, l'estimateur du maximum de vraisemblance pour le paramètre de forme est obtenu.

[MS-40]

Asymptotic Expansion of the Coverage Probability of James-Stein Estimators

Développement asymptotique de la probabilité de couverture des estimateurs de James-Stein

S. Ejaz AHMED, *University of Windsor*, Andrei VOLODIN, *University of Regina*

We provide a new approach to the asymptotic expansion construction of the coverage probability of the confidence sets re-centered at the James and Stein and its positive-part Stein estimators. The coverage probability of these confidence sets depends on the non-centrality parameter as in the case of risks of these estimators. The new approach allows to obtain the asymptotics analysis of the coverage probabilities for the two cases, namely when non-centrality parameter tends to zero or infinity. For both cases we provide a simple approximation of the coverage probabilities. Some graphical and tabular results are provided.

Nous présentons une nouvelle approche à la construction par développement asymptotique de la probabilité de couverture de régions de confiance recentrés aux estimateurs James et Stein et à l'estimateur de partie positive de Stein. La probabilité de couverture de ces régions de confiance dépend du paramètre de décentralité comme dans le cas du risque de ces estimateurs. La nouvelle approche permet d'obtenir les analyses asymptotiques des probabilités de couverture dans les deux cas, c'est-à-dire lorsque le paramètre de décentralité tend vers zéro ou l'infini. Pour les deux cas, nous présentons une approximation simple des probabilités de couverture. Quelques résultats graphiques et tabulaires sont donnés.

[MS-41]**Estimation of a Bounded Location Parameter for p -variate spherical distributions
Estimation d'un paramètre de position multidimensionnel dont la norme est bornée**Amir PAYANDEH, *University of New Brunswick*, Éric MARCHAND, *Université de Sherbrooke*

We consider the problem of estimating, under squared error, the location parameter of a p -variate spherical distribution where the parameter is bounded to a ball. After briefly reviewing related literature concerned primarily with providing dominating estimators of usual procedures obtained from the criteria of unbiasedness and maximum likelihood for instance, we describe recent results which focus on the frequentist performance of estimators based on the linear minimax criterion. Our results also apply and are illustrated in the univariate case of a constraint to an interval.

Nous exposons des résultats récents en ce qui concerne l'estimation ponctuelle d'un paramètre de position multidimensionnel sous la perte quadratique lorsque la norme de ce paramètre est bornée supérieurement. Dans un premier temps, nous passons en revue quelques résultats de dominance où des procédures basées sur les critères du sans-biais ou de la vraisemblance maximale sont améliorées. Dans un second temps, nous décrivons des résultats récents sur la performance fréquentiste d'estimateurs basés sur le critère du minimax linéaire. Nous illustrons aussi les résultats dans le cas univarié où le paramètre de position est borné à un intervalle.

[MS-42]**On the Distribution of Functions of Elliptically Contoured Random Vectors
Sur la distribution de fonctions de vecteurs aléatoires appartenant à la classe des lois elliptiques**Serge PROVOST, *University of Western Ontario*

First, a representation of the density function of noncentered elliptically contoured random vectors is obtained from an alternative decomposition of such vectors. It is then shown that the expectations of certain functions of these vectors can be determined from their Gaussian counterparts. In particular, representations of the moments, moment-generating function and distribution function of quadratic forms in elliptically contoured vectors are provided. A numerical example illustrates the results.

On obtient d'abord une représentation de la fonction de densité de vecteurs aléatoires décentrés appartenant à la classe des lois elliptiques à partir d'une décomposition non standard de tels vecteurs. On démontre ensuite que les espérances de certaines fonctions de ces vecteurs peuvent être exprimées en termes de résultats connexes s'appliquant à des vecteurs gaussiens. En particulier, on obtient des représentations des moments, de la fonction génératrice des moments et de la fonction de répartition de formes quadratiques dont les vecteurs possèdent des distributions elliptiques. Un exemple numérique sera également présenté.

[MS-43]**News from NSERC
Nouvelles du CRSNG**Serge VILLEMURE, *NSERC/CRSNG*

I will present and requests participant's input on the following topics:

Grant Selection Committees (GSC) Structure Review: The current structure of NSERC's discipline-based GSCs has been in place for more than 30 years. NSERC has recently undertaken a review of such structure and would like to update the community on the review's status as well as gather input on the process so far, and on future directions.

Je couvrirai les sujets suivants et recueillerai vos commentaires :

Revue de la structure des Comités de sélection (CSS) : La structure actuelle de nos CSS est en place depuis plus de 30 ans. Le CRSNG a récemment entrepris une revue de cette structure et aimerait donner une mise à jour à la communauté sur cette revue, ainsi que recueillir les commentaires de celle-ci sur le processus en place, et sur ses orientations futures.

International Review of Discovery Grants Program: NSERC is commissioning an International Review to assess the merit of its approach to supporting research in the natural sciences and engineering, and the extent to which its Discovery Grants Program fosters and supports research excellence. NSERC will present the context of this study as well as gather feedback from the research community on the initial steps of the study.

Discovery Grants Competitions Results 2007: NSERC will present a brief report on the most recent Discovery Grants competition.

Revue internationale du Programme de subventions à la découverte : Le CRSNG a entrepris une revue internationale pour évaluer le mérite de son approche au soutien à la recherche en sciences naturelles et en génie, ainsi que le degré auquel le Programme de subventions à la découverte favorise et soutient l'excellence en recherche. Le CRSNG vous décrira le contexte de cette étude et recueillera les commentaires de la communauté de recherche sur les étapes initiales de celle-ci.

Résultats des concours 2007 des subventions à la découverte : Le CRSNG vous présentera un bref rapport sur le plus récent concours de subventions à la découverte.

Session 03A Lundi 11 juin • Monday, June 11, 13:45–14:45 IIC-2001

Canadian Journal of Statistics Award Address

Allocution du récipiendaire du Prix de la *Revue canadienne de statistique*

Session 03B Lundi 11 juin • Monday, June 11, 13:45–15:00 SN-2109

Case Study I: Demographic Methods & Case Study II: Growth of Atlantic Herring

Étude de cas I : Méthodes démographiques & Étude de cas II : La croissance du hareng de l'Atlantique

[MS-44]

Introduction of the Demographic Methods Data Set

Présentation du jeu de données sur les méthodes démographiques

Georges MONETTE, *York University*

[MS-45]

Introduction of the Growth of Atlantic Herring Data Set

Présentation du jeu de données sur la croissance du hareng de l'Atlantique

John WHEELER, *Department of Fisheries and Oceans*

Atlantic herring is fished commercially in the coastal waters of Newfoundland. Management measures for herring supply include a minimum allowable fish size which was set in the 1970's. Recently, fishers have encountered increased percentages of under-sized herring and contend that the minimum allowable fish size should be reduced. The data for this case study are measurements from annual samples from the herring fishery. The primary objectives are: (1) to determine if the growth rates of Atlantic herring have changed, and (2) to examine whether there are spatial differences in these changes. This case study was provided by John Wheeler of Fisheries and Oceans Canada.

Le hareng de l'Atlantique est pêché commercialement dans les eaux côtières de Terre-neuve. Les mesures de gestion des stocks de hareng incluent une taille minimale permise pour les prises de poissons qui a été fixée dans les années 70. Récemment, les pêcheurs ont rencontré des pourcentages de plus en plus grands de hareng sous la taille minimale et proposent de réduire la taille minimale permise. Les données pour cette étude de cas sont des mesures d'échantillons annuels effectuées lors de la pêche du hareng. Les objectifs principaux sont : (1) déterminer si les taux de croissance du hareng de l'Atlantique ont changé, et (2) examiner si il existe des différences spatiales dans ces changements. Cette étude de cas fut fournie par John Wheeler de Pêches et Océans Canada.

Session 03C**Lundi 11 juin • Monday, June 11, 13:30–15:00****SN-1103****Mathematical Finance and Stochastic Analysis I****Finance mathématique et analyse stochastique I****[MS-46]****A Bias Correction Technique for Monte Carlo Pricing of Early Exercise Options****Une technique de correction du biais pour la fixation des prix d'options bermudiennes par Monte-Carlo**Matt DAVISON, Tyson WHITEHEAD & Mark REESOR, *University of Western Ontario*

We present a method for reducing bias in Monte Carlo price estimates for Bermudan-style options. At each exercise opportunity we assume there is an unbiased estimator of the claim value at the next exercise point. We justify an approximation to the distribution of this statistic to derive an expression, which can be estimated in the simulation context, for the bias. It is easy to use this bias estimate to compute bias-reduced estimators of the claim value. We show that our method provides significant improvements over the vanilla stochastic mesh technique and that it is more computationally efficient than nonparametric bootstrapping.

Nous présentons une méthode afin de réduire le biais dans les estimés de prix d'options bermudiennes par Monte-Carlo. À chaque opportunité d'exercice, nous supposons qu'il existe un estimateur sans biais de la valeur du titre au prochain point d'exercice. Nous justifions une approximation de la distribution de cette statistique pour obtenir une expression, pouvant être estimée dans un contexte de simulation, pour le biais. Il est facile d'utiliser cet estimé du biais pour calculer des estimateurs à biais réduits de la valeur du titre. Nous montrons que notre méthode apporte des améliorations significatives sur la technique de maillage stochastique usuelle et qu'elle est plus efficace en terme de puissance de calcul que le bootstrap non paramétrique.

[MS-47]**Valuation of CDS in a Regime-switching Market****Évaluation du swap de défaut de crédit dans un marché avec changement de régimes**Rogemar MAMON, *University of Western Ontario*, Tak Kuen SIU, *Herriot-Watt University, Edinburgh*, Christina ERLWEIN, *Brunel University, London*

The pricing of credit default swap (CDS) under a structural framework is considered. We assume the interest rate process, the appreciation rate and volatility of the firm's asset values are switching regimes. The switching dynamics is governed by a continuous time Markov chain. The Esscher transform is utilised to find an equivalent martingale measure for this valuation problem in the incomplete market setting. A system of coupled PDEs satisfied by the default probabilities is derived. A numerical example for the case of a 2-state Markov chain is provided to investigate the impact of regime-switching to CDS swap rate.

La fixation des prix du swap de défaut de crédit (SDC) sous un cadre structurel est considérée. Nous supposons que le processus de taux d'intérêt, le taux d'appréciation et la volatilité de la valeur de l'actif de la firme sont des processus à plusieurs régimes. La dynamique des changements de régime est déterminée par une chaîne de Markov à temps continu. La transformée d'Esscher est utilisée pour trouver une mesure martingale équivalente pour ce problème d'évaluation dans le contexte d'un marché incomplet. Un système d'équations aux dérivées partielles couplées satisfait par les probabilités de défaut est obtenu. Un exemple numérique pour le cas d'une chaîne de Markov à deux états est fourni afin d'étudier l'impact du changement de régime sur le taux de swap du SDC.

[MS-48]

Hybrid Models of Credit Risk – The Best of both Worlds?**Modèles hybrides du risque de crédit – le meilleur des deux mondes ?**Tom HURD, *McMaster University*

The classic problem of credit risk is the pricing of the bonds issued by medium and large firms, and the valuation of their associated derivative securities. Two competing approaches to credit risk, namely the structural models and the intensity-based models, can be reconciled in various ways. In this talk, we shall review the basic requirements for modelling credit risk, and examine the pros and cons of various established methods. In the end we will arrive at some simple hybrid models that combine the attractive features of both approaches.

Le problème classique du risque de crédit est la fixation des prix d'obligations émises par des firmes de moyenne ou grande taille, et l'évaluation de leurs titres dérivés associés. Deux approches en compétition au risque de crédit, les modèles structurels et les modèles basés sur les intensités, peuvent être réconciliés de différentes façons. Dans cet exposé, nous faisons la revue des conditions nécessaires de base pour la modélisation du risque de crédit, et examinons le pour et le contre des différentes méthodes usuelles. Nous arrivons finalement à des modèles hybrides simples combinant les caractéristiques attrayantes des deux approches.

Session 03D**Lundi 11 juin • Monday, June 11, 13:45–14:45****AA-1043****Controversies over Fish Stocks****Controverses concernant les stocks de poissons**

[MS-49]

From Science-based to Ad Hoc Fisheries Management Off the East Coast of Canada**D'une gestion des pêches scientifique à une gestion ad hoc sur la côte Est du Canada**Peter SHELTON, *Department of Fisheries and Oceans*

The link between science and decision-making for fisheries off Canada's east coast has weakened over the last two decades and fisheries management has become increasingly *ad hoc*. Reasons include the demand for a large degree of flexibility in the decision-making process and continuing low credibility of scientific knowledge following disastrous collapses. Quantitative advice on the management of these fisheries has improved but is under-utilized. Increasing emphasis on resource ownership is bringing into question the central role of publicly funded science to ensure public good. Direct market-based feedback through product eco-labeling may provide the incentive to fish sustainably in the future.

Le lien entre la science et la prise de décision pour les pêches au large de la côte Est du Canada s'est affaibli au cours des deux dernières décennies et la gestion des pêches est devenue de plus en plus *ad hoc*. Ceci est dû, entre autres, à la demande d'un grand degré de flexibilité dans le processus de prise de décision et à la faible crédibilité persistante de la connaissance scientifique suite à des effondrements désastreux. Les conseils de nature quantitative dans la gestion des pêches sont de plus en plus tenus en compte mais demeurent sous-utilisés. L'emphase plus grande sur la possession de ressources met en question le rôle central de la science financée par les fonds publics afin d'assurer le bien commun. Une rétroaction directe basée sur le marché au moyen de l'éco-étiquetage des produits pourrait inciter à promouvoir une pêche durable dans l'avenir.

Session 03E**Lundi 11 juin • Monday, June 11, 13:30–15:00****AA-1046****Random Effects and Multi-level Modelling****Effets aléatoires et modélisation multi-niveaux****[MS-50]****One-Sample Location Tests for Multilevel Data****Tests de position pour données multi-niveaux**DENIS LAROCQUE, *HEC Montréal*, JAAKKO NEVALAINEN & HANNU OJA, *University of Tampere*

In this paper, we consider testing the location parameter with multilevel (or hierarchical) data. A general family of weighted test statistics are introduced. This family includes extensions to the case of multilevel data of familiar procedures like the t , the sign and the Wilcoxon signed-rank tests. An investigation of the relative merits of selected members of the family of tests is achieved theoretically by deriving their relative asymptotic efficiency (ARE) and empirically via a simulation study. Explicit formulas for optimal weights and a discussion of the impact of omitting a level are provided for 2 and 3-level data.

Dans cet article, nous considérons le problème de tester un paramètre de position avec des données multi-niveaux (ou hiérarchiques). Une famille de tests pondérés est proposée. Elle contient des généralisations au cas de données multi-niveaux de procédures familières telles le test t , le test du signe et le test de Wilcoxon des rangs signés. Les performances des tests sont étudiées théoriquement (efficacité asymptotique de Pitman) et empiriquement (simulation). Des formules explicites pour les poids optimaux dans le cas de données à 2 et 3 niveaux sont présentées. On discute également de l'impact d'omettre un niveau sur l'analyse.

[MS-51]**Performance of Estimation Methods for Generalized Linear Mixed Models for Binary Outcome Data with a Low Number of Occurrences****Performance de méthodes d'estimation pour modèles linéaires généralisés mixtes pour données dont la variable réponse est binaire avec peu d'événements positifs**MARIE-EVE BEAUCHAMP, ROBERT PLATT & JAMES HANLEY, *McGill University*

Penalized quasi-likelihood (PQL) is among the most commonly used estimation method for generalized linear mixed models. Some authors have recommended using adaptive Gauss-Hermite quadrature (AGHQ) because it is expected to produce less biased parameter estimates. We compared via simulations the performance of PQL and AGHQ for several settings of binary outcome data with a low number of occurrences. Different types of failure were observed: convergence problems, parameter estimates and standard errors outside of realistic ranges, and biased parameter estimates. We attempted to characterize the circumstances in which the estimation procedures failed.

La quasi-vraisemblance pénalisée (PQL) est parmi les méthodes d'estimation les plus utilisées dans la pratique pour les modèles linéaires généralisés mixtes. Certains auteurs ont recommandé d'utiliser plutôt la quadrature de Gauss-Hermite adaptative (AGHQ) étant donné qu'elle est supposée produire des estimés des paramètres moins biaisés. Nous avons comparé à l'aide de jeux de données simulées la performance des méthodes PQL et AGHQ dans plusieurs situations de données avec variable réponse binaire avec peu d'événements positifs. Différents types d'échec ont été observés : problèmes de convergence, estimés des paramètres et erreurs types irréalistes, et estimés des paramètres biaisés. Nous avons tenté de caractériser les circonstances dans lesquelles les méthodes d'estimation ont échoué.

[MS-52]

Identifiability of Mixture Cure Frailty Model**Identifiabilité du modèle de mélange de fragilité d'un traitement**Jiajia ZHANG, *Memorial University*, Yingwei PENG, *Queen's University*

We consider a mixture cure frailty model that generalizes the general mixture cure model by adding a frailty term in the latency distribution. In this talk we investigate the identifiability of the mixture cure frailty model. We show that the mixture cure frailty model with covariates in the cure fraction and the latency component is identifiable if the cure fraction is modeled by a non-constant function. These results can be extended to the general mixture cure model and the frailty model easily. We also investigate the identifiability by simulations, and demonstrate that a mixture cure model that may practically still lack of identifiability even though the model's identifiability is proved theoretically. Thus, caution must be paid when using the mixture cure frailty model.

Nous considérons un modèle de mélange de fragilité d'un traitement qui généralise le modèle de mélange de traitement général en ajoutant un terme de fragilité dans la distribution latente. Dans cet exposé, nous étudions l'identifiabilité du modèle de mélange de fragilité d'un traitement. Nous montrons que le modèle de mélange de fragilité d'un traitement avec covariables dans la fraction du traitement et la composante latente est identifiable si la fraction du traitement est modélisée par une fonction non constante. Ces résultats peuvent être étendus au modèle de mélange de traitement général et au modèle de fragilité aisément. Nous étudions aussi l'identifiabilité au moyen de simulations, et démontrons qu'un modèle de mélange de traitement peut pratiquement être non identifiable alors que l'identifiabilité du modèle est prouvée théoriquement. Il faut donc faire preuve de prudence en utilisant le modèle de mélange de fragilité d'un traitement.

[MS-53]

A Generalized Mixed Modeling Approach for Causal Inference in Intervention Study**Une approche par modélisation mixte généralisée pour l'inférence causale dans une étude d'intervention**Depeng JIANG & Debra PEPLER, *York University*

This study proposed a piece-wise growth model and generalized mixed modeling approach to make causal inference in longitudinal data set formed when an intervention administered to produce change in individual outcome variables. With the proposed approach, we can examine intra-individual and inter-individual trajectories for both immediate and long-term changes following the intervention program. These methods were applied to data from a quasi-experimental school-based bullying prevention program, which was implemented in three elementary schools in Toronto over a three-year period, to show how the generalized mixed modeling approach can have on the conclusions draw from real-world evaluation studies.

Cette étude propose une approche par modèle de croissance par morceaux et un modèle mixte généralisé pour faire de l'inférence causale sur un ensemble de données longitudinales formé lorsqu'une intervention est administrée afin de provoquer un changement des les variables réponses individuelles. Avec l'approche proposée, nous pouvons examiner les trajectoires intra-individu et inter-individus pour les changements immédiats et à long terme suivant le programme d'intervention. Ces méthodes furent appliquées à des données provenant d'un programme quasi-expérimental de prévention de l'intimidation à l'école implémenté dans trois écoles primaires de Toronto sur une période de trois ans. Nous montrons comment l'approche par modélisation mixte généralisée peut avoir un impact sur les conclusions tirées d'études réelles.

[MS-54]**Use of Markov Transition Model to Analyze Longitudinal Data on Youth Smoking
Utilisation d'un modèle de Markov transitionnel dans l'analyse de données longitudinales sur les jeunes fumeurs**Rashid AHMED & K. Stephen BROWN, *University of Waterloo*

We examined a progressive three-state model that incorporates a random effect distribution to analyze longitudinal data from a large community smoking prevention trial which includes missing data.

A pattern mixture model is proposed to deal with such binary longitudinal data. In this approach, subjects are first divided into groups depending on their missing data pattern and then variables based on these groups are used as model covariates. The initial estimates are obtained using the Expectation-Maximization (EM) algorithm with weights combined with the Quasi-Newton method to maximize the complete data likelihood in the maximization steps. Finally we used the MCMC method to test the effect of different random effect distributions on parameter estimates.

Nous examinons un modèle progressif à trois états qui incorpore une distribution d'effet aléatoire pour analyser des données longitudinales d'un grand essai de prévention du tabagisme incluant des données manquantes.

Un modèle de mélange de tendances est proposé afin de traiter de telles données binaires longitudinales. Dans cette approche, les sujets sont d'abord divisés en groupes selon la tendance de leurs données manquantes ; des variables basées sur ces groupes sont ensuite utilisées comme covariables de modèle. Les estimés initiaux sont obtenus en utilisant l'algorithme de maximisation de l'espérance (EM) avec poids combinés avec la méthode Quasi-Newton pour maximiser la vraisemblance des données complètes lors des étapes de maximisation. Finalement, nous avons utilisé une méthode MCMC pour tester l'effet de différentes distributions de l'effet aléatoire sur les estimés des paramètres.

[MS-55]**An Assessment of Bias due to Misspecification of the Random Effects Distribution in Generalized Linear Mixed Models****Une évaluation du biais résultant d'une erreur de spécification de la distribution des effets aléatoires dans les modèles linéaires généralisés mixtes**John NEUHAUS & Charles McCULLOCH, *University of California, San Francisco*

Data analysis using generalized linear mixed models typically specifies a parametric distribution for the random effects. The consequences of misspecifying this distribution is subject to debate; some authors suggest that large biases can arise while others show that there will typically be little bias for the parameters of interest. Since closed form formulae for bias typically do not exist, we develop numerical methods to evaluate bias and present extensive numerical assessments of its magnitude. These assessments indicate that random effects distribution misspecification typically produces little bias when estimating slope coefficients but may yield biased intercepts and variance components estimators.

L'analyse de données utilisant des modèles linéaires généralisés mixtes nécessite typiquement la spécification d'une distribution paramétrique pour les effets aléatoires. Les conséquences d'une erreur de spécification de cette distribution portent à controverse : certains auteurs suggèrent que des biais de taille importante peuvent survenir tandis que d'autres auteurs ont montré qu'il y a typiquement un faible biais concernant les paramètres en question. Comme une expression analytique pour ce biais n'existe pas, nous avons développé des méthodes numériques destinées à évaluer le biais et son ordre de grandeur. Ces évaluations indiquent qu'une erreur de spécification de la distribution des effets aléatoires produit généralement un faible biais lors de l'estimation des coefficients de pente, mais peut produire des estimateurs biaisés pour le terme constant et les composantes de la variance.

Session 03F**Lundi 11 juin • Monday, June 11, 13:30–15:00****SN-2105****Survey Analysis and Operations**
Analyse d'enquêtes et opérations**[MS-56]****Reporting Response Rates in Characteristic Surveys**
Taux de réponse pour les enquêtes sur les caractéristiquesCarlos LEON, *Statistics Canada/Statistique Canada*

The Unified Enterprise Survey integrates several annual business surveys. In order to reduce costs and respondent burden, administrative data is used for some sample units. Starting in 2006 units within this sub-sample have received a questionnaire covering only the variables that are not available on the administrative records: thus, a mixed collection mode is used for these units. As Statistics Canada's current Guidelines for reporting non-response do not explicitly cover mixed collection surveys, we have developed an extension that allows us to handle the cases when the survey units are made up of a collection component and an administrative component.

L'enquête unifiée auprès des entreprises (EUE) compte plusieurs enquêtes annuelles. Pour plusieurs d'entre elles nous utilisons des données administratives pour remplacer un certain nombre d'unités dans l'échantillon de base afin de réduire les coûts de collecte et le fardeau de réponse. À partir de l'année 2006 les unités de ce sous-échantillon reçoivent un questionnaire portant uniquement sur les variables non disponibles sur les fichiers administratifs : pour chacune de ces unités nous avons donc un mode de collecte mixte. Les normes et lignes directrices pour la déclaration des taux de non-réponse de Statistique Canada ne couvrant pas encore explicitement les enquêtes avec un mode de collecte mixte, nous avons développé une extension qui permet de traiter le cas où les unités sont faites d'une composante collecte et d'une composante provenant de données administratives.

[MS-57]**New Collection Methodology in the 2006 Census of Population**
Nouvelle méthodologie de collecte pour le Recensement de l'an 2006Limei SUN & Laurent ROY, *Statistics Canada/Statistique Canada*

Several new collection methods were implemented in the 2006 Census of Population. For the first time, a block canvass operation was performed to update the Address Register, which made it feasible to mail the questionnaires through Canada Post to 70% of households across Canada. Completed questionnaires were mailed back to a central data processing centre where they were scanned and subjected to automated editing for completeness. Failed-edit-follow-ups were done from three Census Help Line Centres. The 2006 Census also provided the option for respondents to complete their questionnaires on the Internet. These new methods and some results will be discussed.

Plusieurs nouvelles méthodes de collecte ont été mises en place dans le Recensement de l'an 2006. Pour la première fois, une opération de porte à porte par blocs a été effectuée afin de mettre à jour le registre des adresses, permettant ainsi de poster les questionnaires directement à 70% des foyers au Canada, en passant par Postes Canada. Une fois remplis, les questionnaires étaient postés à un centre de traitement des données où ils étaient analysés et soumis à un processus de vérification automatisé afin de vérifier que les questionnaires étaient complétés. Le suivi des questionnaires n'ayant pu être vérifiés de façon automatique était fait à partir de trois centres d'Assistance téléphonique du recensement. Le Recensement de l'an 2006 présentait aussi de nouvelles options permettant aux répondants de compléter leurs questionnaires au moyen d'Internet. Nous allons discuter de ces nouvelles méthodes et de quelques résultats.

[MS-58]**Sampling and Retaining Respondents for the National Longitudinal Survey of Children and Youth (NLSCY)
Échantillonnage et retenue de répondants pour l'Enquête longitudinale nationale sur les enfants et les jeunes (ELNEJ)**Beatrice CHAPMAN & Sarah FRANKLIN, *Statistics Canada/Statistique Canada*

With increased demand for data on development of children and youth, the objectives of the National Longitudinal Survey of Children and Youth (NLSCY) and the sampling methods have evolved. Focus will be on techniques to produce adequate sample sizes and maintain high response rates over time. Methods to be discussed are: frame selection to meet specific data requirements, augmenting the sample to allow users to produce domain-specific cross-sectional estimates, switching to a non-monotonic approach (i.e., contacting sample units with a past episode of non-response) and introducing flags to concentrate interviewer efforts on cases with a poorer response history.

Suite à une demande accrue de données sur le développement des enfants et des jeunes, les objectifs de l'Enquête longitudinale nationale sur les enfants et les jeunes (ELNEJ) et les méthodes d'échantillonnage ont évolué. Cette présentation est centrée sur les techniques produisant des tailles d'échantillons adéquates et maintenant des hauts taux de réponse à travers le temps. Les méthodes présentées sont : la sélection d'une base de sondage afin de répondre à certains critères, l'augmentation de l'échantillon de façon à permettre aux utilisateurs de produire des estimations transversales pour des domaines spécifiques, le passage à une approche non-monotone (c.-à-d., contacter les unités échantillonnées avec un épisode passé de non-réponse) et l'introduction de variables indicatrices pour concentrer l'effort des interviewers sur les cas ayant un historique de réponse plus faible.

[MS-59]**Challenges Surrounding the Use of Tax Data in the Development of Quarterly Indicators for Various Service Industries****Défis entourant l'utilisation des données fiscales dans le développement d'indicateurs trimestriels pour diverses industries de services**François BRISEBOIS, Julie GIRARD & Hansheng XIE, *Statistics Canada/Statistique Canada*

The “Quarterly Services Indicators” is a new project at Statistics Canada that aims at tracking the economic productivity on a quarterly basis for more than forty services industries. Administrative data, namely Goods and Services Tax (GST) data, constitute the foundation of this project. This is complemented by a small survey of complex structured businesses in order to provide an accurate industrial and provincial profile of the revenue in the country. The presentation will give an overview of the methodological challenges encountered during the development stage of this project, especially in the development of quality measures.

Les “Indicateurs trimestriels de service” sont le sujet d'un nouveau projet à Statistique Canada qui vise à suivre la productivité économique sur une base trimestrielle pour plus de quarante industries oeuvrant dans le domaine du service. Des données administratives, plus spécifiquement les données de Taxe sur les produits et services (TPS), constituent la fondation de ce projet. Le tout est complété par une petite enquête d'entreprises à structures complexes, et ce afin d'être en mesure de fournir un portrait exact de la distribution des revenus à l'échelle provinciale et selon les différents secteurs industrielles. La présentation donnera un aperçu des défis méthodologiques rencontrés durant le développement de ce projet, plus particulièrement en ce qui a trait au développement de mesures de qualité.

[MS-60]**Evaluating the Quality of Establishment Survey Frames****Évaluer la qualité des bases de sondage dans les enquêtes auprès des établissements**Mary MARCH, *Statistics Canada/Statistique Canada*

Most Statistics Canada establishment surveys make use of one or more list frames such as the Business Register or other lists. Information on the list frames is sometimes enhanced by adding information from an area frame or perhaps telephone or e-mail directories, or from another survey's data. Frames are continuously being improved as the variety of types of enterprise surveys grows. This paper will look at quality issues associated with enterprise survey frames and consider strategies for dealing with them. The quality assessment will address the six components of quality defined by the agency's quality management framework.

La plupart des enquêtes auprès des établissements de Statistique Canada font usage d'au moins une base de sondage de type liste, telle le Registre des entreprises. L'information des listes est parfois améliorée en ajoutant de l'information d'une base de sondage de type aréolaire, ou peut-être de répertoires téléphoniques ou d'adresses de courrier électronique, ou des données d'une autre enquête. Les bases de sondage sont continuellement améliorées, alors que la variété de types d'enquête auprès des entreprises s'accroît. Cet article porte sur les problèmes de qualité associés aux bases de sondage des enquêtes auprès des entreprises et considère des stratégies pour les traiter. L'évaluation de la qualité répondra à six composantes de qualité définies par le cadre de gestion de la qualité de l'agence.

[MS-61]**Weighting of Cumulated Samples for the Canadian Community Health Survey****Pondération des échantillons cumulés pour l'Enquête sur la santé dans les collectivités canadiennes**Cathlin SARAFIN & Steven THOMAS, *Statistics Canada/Statistique Canada*

The Canadian Community Health Survey (CCHS) has recently changed to continuous collection where data will be collected for different samples for the duration of the survey. Similar to a rolling sample, the main advantage of such an approach will be the ability to cumulate sampled units over different periods of time depending on the detail required for estimation. The challenge with this design is calculating weights that properly reflect the population of interest. This paper will discuss some of the different methods available for such a design and will highlight the methodology chosen for the CCHS.

L'Enquête sur la santé dans les collectivités canadiennes (ESCC) est récemment passée à une collecte continue où les données seront recueillies pour différents échantillons pendant toute la durée de l'enquête. Comme dans le cas d'un échantillon successif, l'avantage principal d'une telle méthode est la capacité de cumuler les unités échantillonnées pour différentes périodes de temps selon le niveau de détail requis pour les estimations. Le défi de ce plan est de s'assurer que la pondération reflète correctement la population d'intérêt. Ce document discutera des différentes méthodes disponibles pour un tel plan de sondage et mettra l'accent sur la méthodologie choisie pour l'ESCC.

Session 04A Lundi 11 juin • Monday, June 11, 15:30–17:00 AA-1046**Strategic Directions in Household Surveys Development****Directions stratégiques dans le développement des enquêtes sur les ménages****[MS-62]****Statistics Canada's New Household Survey Strategy****La nouvelle stratégie de Statistique Canada pour les enquêtes sur les ménages**Jack GAMBINO, Jean-Louis TAMBAY & Guy LAFLAMME, *Statistics Canada/Statistique Canada*

We first review Statistics Canada's current approach to household surveys. Statistics Canada has conducted most household surveys using an area frame with the Labour Force Survey design as the key component. With increased use of telephone interviewing, technological advances and creation of an Address Register, alternatives to the current approach have become more feasible. Consequently, there have been recent investigations of such alternatives. In the second part of the presentation, we describe a new approach to household surveys, featuring increased integration, common core content, a master sample and use of a list frame, along with some technical issues that arise.

Nous faisons d'abord une brève revue de l'approche courante de Statistique Canada pour les enquêtes auprès des ménages. La plupart des enquêtes auprès des ménages menées par Statistique Canada utilisent une base aréolaire pour laquelle le plan de l'Enquête sur la population active est une composante clé. Grâce à une utilisation accrue des entrevues téléphoniques, aux avancées technologiques, et à la création d'un registre des adresses, des alternatives à l'approche courante deviennent plus réalisables. Conséquemment, de telles alternatives ont récemment été étudiées. Dans la deuxième partie de cette présentation, nous décrivons une nouvelle approche pour les enquêtes auprès des ménages, utilisant une intégration accrue, un contenu central commun, un échantillon principal et l'usage d'une base de sondage de type liste, ainsi que certaines questions techniques soulevées par cette nouvelle approche.

[MS-63]**Towards a Single Continuous Population Survey for the UK****Vers une seule enquête de population continue pour le Royaume-Uni**Nicola BENNETT, *Office for National Statistics, United Kingdom*

The Office for National Statistics (ONS) plans to integrate the Government household surveys on which it leads into a single continuous population survey to achieve better value for money from its survey program and to deliver a range of new statistical outputs to meet increasing demand for information. A comprehensive integration of the survey process is proposed: an integrated field force of interviewers administering a common modular questionnaire, an unclustered sample and fieldwork design, more integrated processing system, production of outputs from a single common source and a range of new annual outputs on a rolling quarterly basis.

Le Bureau des statistiques nationales (ONS) planifie intégrer les enquêtes gouvernementales sur les ménages qu'il dirige dans une seule enquête continue sur la population afin d'atteindre une meilleure optimisation des ressources de son programme d'enquête et de fournir de nouveaux produits statistiques afin de répondre à une demande d'information croissante. Une intégration compréhensive du processus d'enquête est proposée : une équipe d'interviewers intégrée administrant un questionnaire commun modulaire, un plan d'échantillonnage et de travail sur le terrain non en grappes, un système de traitement plus intégré, une production de résultats à partir d'une seule source commune et une variété de nouveaux produits annuels sur une base trimestrielle.

[MS-64]**American Community Survey: Statistical Issues and Challenges
L'enquête américaine sur les communautés : défis et perspectives statistiques**Alfredo NAVARRO, *US Census Bureau*

The American Community Survey is the Census Bureau's alternative to replace the decennial long form in the 2010 Census. New ventures are usually accompanied by challenges. Census long form data are heavily used for grants allocation, model development by transportation planners, and in a variety of economic and social applications. The paper will describe the ACS statistical methodology for the production of multi-year estimates and basic plans for data releases. The switch to continuous measurement and the production of multiple sets of estimates result in a whole of new issues regarding the use and interpretation of these estimates. This paper addresses some of those issues from a data user perspective.

L'enquête américaine sur les communautés est l'alternative du Bureau de recensement visant à remplacer le questionnaire long décennal dans le recensement de l'an 2010. De nouvelles initiatives sont souvent accompagnées de nouveaux défis. Les données du questionnaire long du recensement sont grandement utilisées pour l'attribution de subventions, le développement de modèles par les planificateurs des transports, et dans une variété d'applications économiques et sociales. Cet article décrit la méthodologie statistique de l'ACS pour la production d'estimations sur plusieurs années et des plans de base pour la communication des données. La transition vers la production de mesures en continue et la production de multiples ensembles d'estimations résulte en un ensemble de nouvelles questions relatives à l'utilisation et à l'interprétation de ces estimations. Cet article répond à certaines de ces questions avec la perspective d'un utilisateur des données.

Session 04B**Lundi 11 juin • Monday, June 11, 15:30–17:00****AA-1043****Mathematical Methods for Studying Patient Wait Times****Méthodes mathématiques pour étudier les temps d'attente des patients****[MS-65]****Using Operations Research to Manage Wait Lists
Utilisation de la recherche opérationnelle dans la gestion de files d'attentes**Michael CARTER, *University of Toronto*

Long wait lists are high on the political agenda recently. The medical community has been struggling with ways to deal with the issues. Typically we do not know who is waiting; we simply read anecdotal stories of long delays. Although funding restrictions over the past decade have certainly contributed to the problem, we believe that there are major issues with resource allocation. It should be possible to significantly increase capacity through better allocation procedures.

We present several practical examples of the use of operations research methods to help understand the demand patterns for medical treatments and to improve resource utilization.

Les longues files d'attentes sont devenues une priorité politique récemment. La communauté médicale cherche des façons de traiter ce problème. Typiquement, nous ne savons pas qui attend : nous lisons simplement des histoires anecdotiques de longs délais. Bien que des restrictions budgétaires au cours de la dernière décennie ont certainement contribué au problème, nous croyons qu'il y a des problèmes majeurs dans la répartition des ressources. Il devrait être possible d'augmenter significativement la capacité au moyen de meilleures procédures de répartition.

Nous présentons plusieurs exemples pratiques d'utilisation de méthodes de recherche opérationnelle pour aider à la compréhension des tendances de demande pour les traitements médicaux et pour améliorer l'utilisation des ressources.

[MS-66]**Aggregation Methods for Health Performance Indicators of Wait Times****Méthodes d'agrégation pour indicateurs de performance des temps d'attentes dans le système de santé**

Wendy Lou, *University of Toronto*

A wealth of data on the state of our public health care system has been collected over the past few years to assess system performance. It is currently difficult to compare instantaneous performance indicators across regions and over time, due to underlying differences and temporal variations in the states of local health systems and the populations they serve. An aggregated approach over a well-defined duration may better reflect system performance overall. A statistical framework for aggregating and comparing indicators, including patient wait times, will be proposed, and examples will be given using data compiled by CIHI and Statistics Canada.

Une grande quantité d'informations sur l'état de notre système de santé public a été amassée au cours des dernières années afin d'évaluer sa performance. Il est présentement difficile de comparer des indicateurs de performance instantanés entre différentes régions à travers le temps à cause des différences sous-jacentes et des variations temporelles dans l'état des systèmes de santé locaux et des populations qu'ils desservent. Une approche agrégée sur une durée bien définie peut mieux refléter la performance globale du système. Un cadre statistique pour l'agrégation et la comparaison d'indicateurs, incluant le temps d'attente des patients, sera proposé, et des exemples seront donnés à partir de données compilées par l'ICIS et Statistique Canada.

[MS-67]**Methods for Studying Events on Surgical Wait Lists****Méthodes pour l'étude d'événements sur des listes d'attentes chirurgicales**

Boris SOBOLEV, *University of British Columbia*

Studies on access to health care are often interested in determining the relationship between delay and risk of adverse events while waiting. When access to surgery is rationed by wait lists, patients requiring the operation are subject to competing events of surgery, pre-operative death, or unplanned admission. Methodologically, measuring the probability of adverse events on wait lists is similar to quantifying the risk of an event during follow-up in a population exposed to competing risks. In this paper, we consider one measure for summarizing wait-list outcomes – cumulative incidence function, as well as a regression model using its pseudo-values.

Les études sur l'accès à des soins de santé se sont souvent intéressées à déterminer la relation entre les délais et le risque d'événements néfastes lors de l'attente. Lorsque l'accès aux chirurgies est rationné par des files d'attentes, les patients nécessitant une opération sont sujet aux événements compétitifs de la chirurgie, de la mort pré-opérative et de l'admission non planifiée. Méthodologiquement, mesurer les probabilités d'événements néfastes sur les files d'attentes est similaire à quantifier le risque d'un événement lors du suivi dans une population exposée à des risques compétitifs. Dans cet article, nous considérons une mesure permettant de résumer les résultats de la file d'attente – fonction d'incidence cumulative, ainsi qu'un modèle de régression utilisant ses pseudo-valeurs.

Session 04C**Lundi 11 juin • Monday, June 11, 15:30–17:00****IIC-2001****Reliability Models and Their Applications****Modèles de fiabilité et leurs applications****[MS-68]****Uncoupling Longitudinal Data in Threshold Regression****Découpler des données longitudinales en régression de seuil**George WHITMORE, *McGill University*, Mei-Ling Ting LEE, *Ohio State University*

A popular reliability model is based on the concept that an item degrades according to a stochastic process and failure occurs when the process first reaches a critical threshold. Longitudinal data in this context refer to sequential observations on the degradation of the item, its covariates, or both, prior to failure. This regression data structure, called threshold regression, is common in practical survival and reliability settings. Breaking longitudinal series into single records is one strategy for handling these data. This study looks at the validity of this uncoupling procedure. Practical cases illustrate the uncoupling conditions and the threshold regression technique.

Un modèle de fiabilité populaire est basé sur le concept qu'un objet se dégrade selon un processus stochastique et que la panne survient lorsque le processus atteint pour la première fois un seuil critique. Les données longitudinales dans ce contexte font référence aux observations séquentielles sur la dégradation de l'objet et/ou ses covariables avant la panne. Cette structure de régression, appelée régression de seuil, est courante dans le contexte pratique de la fiabilité et de la survie. Diviser la série longitudinale en simples records est une stratégie pour traiter de telles données. Cette étude porte sur la validité de cette procédure de découplage. Des cas pratiques illustrent les conditions de découplage et la technique de régression de seuil.

[MS-69]**Reliability and Warranty Claims for Systems with Age and Usage Scales****Fiabilité et réclamations de garanties pour des systèmes à échelles d'âge et d'utilisation**Jerry LAWLESS, *University of Waterloo*

The reliability of equipment is often a function of both age and usage. This talk will discuss models for failure when this is the case, along with some of the challenges of fitting such models to observational data. In particular, incomplete data problems that arise with warranty claims will be considered. Proposed methodology will be illustrated on some car warranty data.

La fiabilité d'équipement est souvent une fonction de son âge et de son utilisation. Cet exposé portera sur des modèles pour la panne dans ce cas, ainsi que certains des défis dans l'ajustement de ces modèles aux données. En particulier, les problèmes de données incomplètes présents dans les réclamations de garanties seront considérés. La méthodologie proposée sera illustrée sur des données de garantie sur des automobiles.

[MS-70]**Condition Based Maintenance: Estimating the Remaining Useful Life of an Asset****Entretien conditionnel : estimer la vie utile d'un actif**Andrew JARDINE & Dragan BANJEVIC, *University of Toronto*

An estimate of remaining useful life (RUL) of equipment is useful to reliability specialists to decide to do preventive maintenance or delay it due to production requirements. It's usually assumed that in later life the hazard function is increasing,

Un estimé de la vie résiduelle utile (VRU) d'un équipement est utile aux spécialistes de la fiabilité dans leur décision d'effectuer un entretien préventif ou de le remettre à plus tard afin de répondre à des besoins de production. Il est généralement supposé que le taux de panne est croissant et que

and the expected RUL is decreasing. It's been noticed in case studies that the standard deviation of RUL decreases, but that the coefficient of variation increases. We have proved this is true under general conditions and that the limiting distribution of RUL is exponential, with relatively high variability. Aspects of including condition monitoring in prediction of RUL will be discussed.

la VRU espérée est décroissante. Des études de cas ont permis d'observer que l'écart type de la VRU décroît, mais que son coefficient de variation augmente. Nous avons prouvé que cela est vrai sous des conditions générales et que la distribution limite de la VRU est exponentielle, avec une variabilité relativement élevée. Nous discutons également de certains aspects de l'inclusion de la surveillance de l'état de l'équipement dans la prévision de la VRU.

Session 04D Lundi 11 juin • Monday, June 11, 15:30–17:00 SN-2105

Markov Chain Monte Carlo and Quasi Monte Carlo – A Productive Marriage Chaînes de Markov Monte Carlo et quasi Monte Carlo – Un mariage productif

[MS-71]

QMC for MCMC QMC pour MCMC

Art OWEN & Seth TRIBBLE, *Stanford University*

Monte Carlo (MC) methods are used in many fields. Recent decades have seen many improvements in MC methods. Quasi-Monte Carlo (QMC) methods have developed that make MC more accurate. Markov chain Monte Carlo (MCMC) methods make MC more widely applicable. Recently it has been shown that QMC constructions that are “completely uniformly distributed”(CUD) or even weakly CUD, are suitable for use in MCMC applications. This talk reviews that work, gives sufficient conditions for a sequence to be (W)CUD and looks at some numerical examples.

Les méthodes Monte-Carlo (MC) sont utilisées dans plusieurs domaines. Plusieurs améliorations ont été apportées aux méthodes MC dans les décennies récentes. Des méthodes quasi Monte-Carlo (QMC) ont été développées rendant les méthodes MC encore plus précises. Les méthodes de chaînes de Markov Monte-Carlo (MCMC) rendent les méthodes MC plus largement applicable. Récemment, il a été démontré que les constructions QMC qui sont “complètement uniformément distribuées” (CUD), ou même faiblement CUD, peuvent être utilisées dans des applications MCMC. Cet exposé passe en revue ce travail, donne des conditions suffisantes pour qu'une suite soit (F)CUD et présente quelques exemples numériques.

[MS-72]

Acceleration of the Multiple-Try Metropolis Algorithm using Antithetic and Stratified Sampling Accélération de l'algorithme de Metropolis à essais multiples via l'échantillonnage antithétique et stratifié

Christiane LEMIEUX, *University of Waterloo*, Radu CRAIU, *University of Toronto*

The Multiple-Try Metropolis is a recent extension of the Metropolis algorithm in which the next state of the chain is selected among a set of proposals. In this talk, we propose a modification of this algorithm which allows the use of correlated proposals, particularly antithetic and stratified proposals. The method is particularly useful for random walk Metropolis in high dimensional spaces and can be used easily when the proposal distribution is Gaussian. We explore the use of quasi-Monte Carlo methods to generate highly stratified samples. A

L'algorithme de Metropolis à essais multiples est une récente extension de l'algorithme de Metropolis dans laquelle le prochain état de la chaîne est choisi parmi un ensemble de candidats (ou essais). Dans cet exposé, nous proposons une modification de cet algorithme qui permet l'utilisation de candidats corrélés. Nous étudions, en particulier, des corrélations induites par des échantillons antithétiques et stratifiés. Pour produire des échantillons hautement stratifiés, nous utilisons des méthodes quasi Monte-Carlo. La modification proposée est particulièrement utile pour la marche aléatoire de Metropolis dans des espaces en haute dimension et peut être ap-

series of examples is presented to evaluate the potential of the method.

pliquée aisément lorsque la distribution candidate est gaussienne. Des exemples sont donnés afin d'illustrer le potentiel de la méthode.

Session 04E

Lundi 11 juin • Monday, June 11, 15:30–17:00

SN-2109

Case Study II: Growth of Atlantic Herring

Étude de cas II : La croissance du hareng de l'Atlantique

[MS-73]

Introduction to the Data Set

Présentation du jeu de données

John WHEELER, *Department of Fisheries and Oceans*

Atlantic herring is fished commercially in the coastal waters of Newfoundland. Management measures for herring supply include a minimum allowable fish size which was set in the 1970's. Recently, fishers have encountered increased percentages of under-sized herring and contend that the minimum allowable fish size should be reduced. The data for this case study are measurements from annual samples from the herring fishery. The primary objectives are: (1) to determine if the growth rates of Atlantic herring have changed, and (2) to examine whether there are spatial differences in these changes. This case study was provided by John Wheeler of Fisheries and Oceans Canada.

Le hareng de l'Atlantique est pêché commercialement dans les eaux côtières de Terre-neuve. Les mesures de gestion des stocks de hareng incluent une taille minimale permise pour les prises de poissons qui a été fixée dans les années 70. Récemment, les pêcheurs ont rencontré des pourcentages de plus en plus grands de hareng sous la taille minimale et proposent de réduire la taille minimale permise. Les données pour cette étude de cas sont des mesures d'échantillons annuels effectuées lors de la pêche du hareng. Les objectifs principaux sont : (1) déterminer si les taux de croissance du hareng de l'Atlantique ont changé, et (2) examiner si il existe des différences spatiales dans ces changements. Cette étude de cas fut fournie par John Wheeler de Pêches et Océans Canada.

Session 04F

Lundi 11 juin • Monday, June 11, 15:30–17:00

SN-2041

Statistics and the Environment

La statistique et l'environnement

[MS-74]

Parameter Estimation using Monte Carlo Methods and the EM Algorithm

Estimation de paramètres utilisant des méthodes de Monte-Carlo et l'algorithme EM

Amanda HALLADAY, *Dalhousie University*

Parameter estimation in linear state-space models has been widely examined when the errors are assumed to be Normally distributed. In such cases, the well-known Kalman filter can be applied. However, there has been little research exploring cases where the errors do not satisfy the assumptions of Normality. In this talk, I will explore the use of Monte Carlo methods and the EM algorithm to estimate parameters in the non-Gaussian and non-

L'estimation des paramètres de modèles linéaires à espaces d'états lorsque les erreurs sont supposées gaussiennes a été largement étudiée. Dans ce cas, le célèbre filtre de Kalman peut être appliqué. Cependant, il y a eu peu de recherches concernant les cas où l'hypothèse de la normalité des erreurs n'est pas vérifiée. Dans cet exposé, j'explore l'utilisation de méthodes de Monte-Carlo et l'algorithme EM afin d'estimer les paramètres dans un cadre non-gaussien et non-linéaire. À des fins d'illustration, cette méthode sera appliquée à des

linear setting. For illustration, I will apply this method on some tracking data of Leatherback turtles.

données de pistage de tortues luth.

[MS-75]

**Finding the Fish: Straying Trends Between B.C. Pacific Herring (*Clupea pallasii*) Populations
Trouver le poisson : orientations de populations de hareng (*Clupea pallasii*) au large de la Colombie-Britannique**

Carolyn HUSTON & Carl SCHWARZ, *Simon Fraser University*, Jake SCHWEIGERT, *Department of Fisheries and Oceans*

Herring tag data is available starting from 1936 in British Columbia. We estimate the herring stray rates between five geographically distinct populations defined by the Department of Fisheries and Oceans (DFO) using a multi-state capture-recapture population model. Based on deviances from these straying estimates, we will identify years where there have been noticeable quantitative changes in herring stray behaviour. Hopefully these changes can be associated with recorded environmental indicators. Special consideration will be given to how to treat missing data resulting from both fisheries closures and missed tagging events.

Des données d'étiquetage de harengs sont disponibles à partir de 1936 pour la Colombie Britannique. Nous estimons les taux de dispersion des harengs entre cinq populations géographiquement distinctes définies par le ministère des Pêches et des Océans (MPO) en utilisant un modèle de population capture-recapture à multi-états. À partir des déviations de ces estimés de dispersion, nous allons identifier les années où il y a eu un changement quantitatif perceptible dans le comportement des harengs en termes d'orientation. Nous espérons que ces changements pourront être associés aux indicateurs environnementaux mesurés. Une considération spéciale sera donnée au traitement de données provenant de fermetures de pêcheries et d'événements d'étiquetage manqués.

[MS-76]

**Bootstrap Inference for Ranking and Selection: Applications to Forest Genetics
Inférence par bootstrap pour le classement et la sélection et applications en génétique forestière**

Anne Michele MILLAR, *Mount Saint Vincent University*

In tree improvement programs individual parent trees are ranked on their estimated genetic value for maximizing the height of future offspring. Groups can be ranked in terms of genetic diversity – lack of diversity may lead reduced resistance to pests or disease. Individual and group criteria can then be combined in a selection process. Bootstrap provides inference for such a process including the estimation of the individual ranks, the acceptability of the group selection, and the potential gain in height.

Dans les programmes d'amélioration génétique forestière, les arbres parentaux sont classés par leur valeur génétique estimée, pour maximiser la hauteur des futures progénitures. Les groupes peuvent aussi être classés en terme de diversité génétique – une diversité réduite pourrait conduire à une résistance moindre aux ravageurs et maladies. Les critères de groupe et individuels peuvent être ensuite combinés pour la sélection. Le bootstrap permet l'inférence dans ce processus, incluant l'estimation des rangs individuels, l'acceptabilité de la sélection de groupe et les gains potentiels en hauteur.

[MS-77]**An R^2 alternative to measure the accuracy of species distribution models
Une mesure de précision de modèles de distribution des espèces alternative au R^2** Paul NGUYEN, *University of Western Ontario*, Linh NGUYEN, *Trent University*

The coefficient of determination (R^2) is a summary statistic for predictive accuracy in regression models. However, R^2 has been criticized as it may not be the most intuitive measure of accuracy. An alternative measure of accuracy that has recently been applied to ecology is the area under the receiver operating characteristic (ROC) curve. A nonparametric estimator of accuracy for least squares regression is described with similar interpretation to the area under the ROC curve for binary classification problems. The proposed accuracy estimator has several advantages over R^2 to warrant its routine use as an auxiliary analysis to least squares regression.

Le coefficient de détermination (R^2) est une statistique sommaire pour la prévision prédictive de modèles de régression. Cependant, le R^2 a été critiqué car il peut ne pas être la mesure de précision la plus intuitive. Une mesure de précision alternative récemment appliquée en écologie est l'aire sous la courbe de fonction d'efficacité de l'observateur (courbe ROC). Un estimateur non paramétrique de la précision pour la régression par moindres carrés ayant une interprétation similaire à l'aire sous la courbe ROC pour problèmes de classification binaire est décrit. L'estimateur de précision proposé a plusieurs avantages par rapport au R^2 justifiant son utilisation régulière comme analyse auxiliaire à la régression par moindres carrés.

[MS-78]**Climate Change and Gastrointestinal Illness in Alberta – A Spatial Temporal Analysis using Conditionally Autoregressive Models****Changement climatique et maladies gastro-intestinales en Alberta – une analyse spatiale et temporelle utilisant des modèles conditionnellement autorégressifs**Daniel GILLIS & Stephanie DIXON, *University of Guelph*

The IPCC has indicated that changes to the world climate, including shifts in the incidence and severity of extreme weather are a result of human activities. To adapt to these changes, an understanding of the effect on human activity are needed. Following several waterborne disease outbreaks, there is specific interest in identifying relationships between climate and human health. Using Alberta health and climate data, this talk will highlight expected changes to incidence rates of gastrointestinal illness based on several climate change scenarios. Incidence rates are obtained using conditionally autoregressive models that adjust for spatial and temporal correlations within the data.

Le GIEC a indiqué que les changements du climat mondial, incluant des décalages dans l'incidence et la sévérité de conditions atmosphériques extrêmes, sont un résultat des activités humaines. Pour pouvoir s'adapter à ces changements, une compréhension de leur effet sur l'activité humaine est nécessaire. Suite à l'éclosion de plusieurs maladies hydriques, il y a un intérêt spécifique dans l'identification des relations entre le climat et la santé humaine. En utilisant des données de santé et du climat en Alberta, cet exposé mettra en lumière les changements attendus des taux d'incidence des maladies gastro-intestinales en se basant sur plusieurs scénarios de changement climatique. Les taux d'incidence sont obtenus en utilisant des modèles conditionnellement autorégressifs s'ajustant aux corrélations spatiales et temporelles dans les données.

Session 04G**Lundi 11 juin • Monday, June 11, 15:30–17:00****SN-1103****Statistical Theory II
Théorie statistique II****[MS-79]****Classification of Random Patterns****Classification de formes aléatoires**Jeffrey PICKA & Mingxia DENG, *University of New Brunswick*

Statistical methods are required to distinguish between realizations of different pattern-forming processes. In most cases classical tests cannot be developed, but statistical learning methods can be used if many independent realizations of the patterns are available. By developing classifiers based on innovative descriptive statistics, realizations of germ-grain algorithms can be classified with error rates comparable with or better than the usual type I error probabilities used in tests.

Des méthodes statistiques sont requises afin de distinguer entre les réalisations de différentes formes aléatoires. Dans la plupart des cas, des tests classiques ne peuvent être développés, mais des méthodes d'apprentissage statistique peuvent être utilisées si plusieurs réalisations indépendantes des formes sont disponibles. En développant des classificateurs basés sur des statistiques descriptives innovatrices, les réalisations d'algorithmes germe-grain peuvent être classifiées avec des taux d'erreurs comparables ou meilleurs que les probabilités d'erreur de type I utilisés dans les tests.

[MS-80]**Balanced-type Loss Functions: Review and Recent Results****Fonctions de perte “équilibrées” : revue et résultats récents**Éric MARCHAND, *Université de Sherbrooke*

Balanced loss functions combine two criteria for assessing the precision of an estimate, namely proximity to a target estimator (e.g., least-squares) and proximity to an unknown parameter. Key historical contributions along with recent results are reviewed in this talk.

Les fonctions de perte dites “équilibrées” combinent deux critères pour mesurer la précision d'un estimateur, soient les proximités à un estimateur cible (ex. moindres carrés) et à un paramètre inconnu. Nous passons en revue des propriétés de ces fonctions de perte et présentons des résultats récents.

[MS-81]**Parameter Estimation in the Skew Generalized Secant Hyperbolic Distribution****Estimation des paramètres de la distribution sécante hyperbolique généralisée asymétrique**David VAUGHAN, *Wilfrid Laurier University*

The generalized secant hyperbolic distribution (GSH) has proven to be an effective model of symmetric error distributions which exhibit non-normal kurtosis. However, the logarithm of financial returns, often assumed to have a normal distribution, has been shown by extensive empirical studies to have skewness that cannot be ignored. The skew generalized secant hyperbolic (SGSH) family of distributions is a natural extension of the GSH family, with all moments finite and able to model both thin and thick tailed symmetric or skewed distributions. In this talk, parameter estimation in the SGSH family is illustrated with examples.

La distribution sécante hyperbolique généralisée (SHG) s'est montrée être un modèle efficace de distributions d'erreur symétriques à kurtose non-normale. Cependant, des études empiriques approfondies ont démontré que le logarithme de rentabilité financière, souvent supposé normal, a une asymétrie ne pouvant être ignorée. La famille de distributions sécante hyperbolique généralisée asymétrique (SHGA) est une généralisation naturelle de la famille SHG, dont tous les moments sont finis et pouvant modéliser des distributions à ailes relevées ou non, symétriques ou asymétriques. Dans cet exposé, l'estimation paramétrique dans la famille SHGA est illustrée à l'aide d'exemples.

[MS-82]

Chung-Smirnov Property for Bernstein Estimators of Distribution Functions
Propriété de Chung-Smirnov et estimateurs de Bernstein d'une fonction de répartition

Alexandre LEBLANC, *University of Manitoba*

In this talk, we will discuss the fact that the Chung-Smirnov property holds for Bernstein estimators of distribution functions under different conditions on the underlying distribution to be estimated. We will also discuss general results that characterize the closeness between these Bernstein estimators and the empirical distribution function in terms of the modulus of oscillation of the uniform empirical process.

Durant cette présentation, nous examinerons les estimateurs d'une fonction de répartition basés sur l'utilisation des polynômes de Bernstein. Plus particulièrement, nous mettrons l'accent sur le fait que ces estimateurs possèdent la propriété de Chung-Smirnov sous divers scénarios quant à la régularité de la fonction de répartition sous-jacente. Nous présenterons aussi un résultat général permettant de quantifier l'écart entre les estimateurs de Bernstein et la fonction de répartition empirique en termes du module d'oscillation du processus empirique uniforme.

[MS-83]

Scalable Action Respecting Embedding (ARE)
Emboîtement respectant les actions à échelles multiples

Ali GHODSI & Michael BIGGS, *University of Waterloo*, Michael BOWLING *University of Alberta*, Dana WILKINSON, *University of Waterloo*, David WINGATE, *University of Michigan*

ARE is a non-linear dimensionality reduction technique for embedding high-dimensional observation trajectories from interaction with a complex system. The core of ARE is a semi-definite optimization with constraints that actions are distance-preserving in the resulting embedding. Unfortunately, this formulation is limited to relatively small datasets. This paper shows how to solve large problems of this type by a conjugate gradient optimization in the primal variables of the ARE formulation. Although this optimization is not convex, we show how to compute an excellent seed near the global minimum by introducing an Isomap variation that incorporates actions.

L'emboîtement respectant les actions (ERA) est une technique non-linéaire de réduction de dimension pour emboîter des trajectoires d'observations de dimension élevée d'interactions avec un système complexe. Le coeur d'ERA est une optimisation semi-définie sous la contrainte que les actions préservent les distances dans l'emboîtement résultant. Malheureusement, cette formulation est limitée à des jeux de données relativement petits. Cet article montre comment résoudre de grands problèmes de ce type par optimisation du gradient conjugué dans les variables primaires de la formulation d'ERA. Malgré que cette optimisation ne soit pas convexe, nous montrons comment calculer une excellente solution initiale près du minimum global en introduisant une variante basée sur des cartes isométriques incorporant les actions.

[MS-84]

Maximum Likelihood Estimation under Three-phase Sampling
Estimation par maximum de vraisemblance sous un plan d'échantillonnage à trois degrés

Chris WILD, Alastair SCOTT, Alan LEE & Yannan JIANG, *University of Auckland*

Semiparametric maximum likelihood estimation is developed for data from three-phase case-control sampling in a way that unifies the work by Scott and Wild (1997, *Biometrika*) on two-phase sampling with a prospective first phase, and Breslow and Holubkov (1997, *JRSSB*) with a case-control

L'estimation par maximum de vraisemblance semi-paramétrique est développée pour des données d'échantillonnage de cas-témoins à trois degrés d'une façon unifiant les travaux de Scott et Wild (1997, *Biometrika*) sur l'échantillonnage à deux degrés avec un premier degré prospectif, et de Breslow et Holubkov (1997, *JRSSB*) avec un premier degré cas-

first phase. Two-phase studies permit cost savings by only obtaining information on covariates that are particularly expensive or particularly invasive for comparatively small subsamples of study subjects. Three-phase sampling extends these capabilities. We also discuss the semiparametric efficiency of semiparametric maximum likelihood for this class of problems.

NSERC representatives and Grant Selection Committee members invite you to attend an interactive workshop with useful advice for the preparation of your next NSERC application, statistics from the 2007 competition and answer your questions on the functioning of the Statistical Sciences Grant Selection Committee.

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.

témoin. Les études à deux degrés permettent des économies en recueillant l'information sur les covariables particulièrement dispendieuses ou invasives pour un sous-échantillon comparativement petit de sujets. L'échantillonnage à trois degrés étend ces possibilités. Nous discutons aussi de l'efficacité semi-paramétrique du maximum de vraisemblance semi-paramétrique pour cette classe de problèmes.

Des employés du CRSNG et des membres des comités de sélection des subventions vous invitent à un atelier en mode interactif avec des conseils utiles pour l'élaboration de votre prochaine demande de subvention au CRSNG, les statistiques pour le concours 2007 et répondront à vos questions sur le fonctionnement sur le comité de sélection des subventions en statistiques.

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 05A Mardi 12 juin • Tuesday, June 12, 08:45–10:00 IIC-2001

Gold Medal Address

Allocution du récipiendaire de la médaille d'or

[MS-85]

Modelling Biological Data: Several Vignettes

Modélisation de données biologiques : plusieurs vignettes

Chris FIELD, *Dalhousie University*

We will develop the theme of Biology as a rich source of interesting problems where the data is noisy and often multivariate. We will focus on several diverse data sets and related scientific questions coming from Biology and discuss the issues of model building and inference for each case. We will emphasize that the appropriate modelling requires careful thought and usually provides methodological challenges. The data sets considered include bird count data, tracking data from leatherback turtles, and amino acid sequence data from primates and cetaceans.

Nous développons le thème de la biologie comme source riche de problèmes intéressants où les données comportent beaucoup de bruit et sont souvent multivariées. Nous nous concentrons sur plusieurs jeux de données de maladies et sur les questions scientifiques liées provenant de la biologie et nous discutons des questions de construction de modèle et d'inférence pour chaque cas. Nous soulignons qu'une modélisation appropriée requiert une réflexion soignée et est souvent source de défis méthodologiques. Les jeux de données considérés incluent des données de dénombrement d'oiseaux, des données de pistage de tortues luth, et des données de séquences d'acides aminés chez les primates et les cétacés.

Session 06A**Mardi 12 juin • Tuesday, June 12, 10:30–12:00****SN-2109****Remedies for Sample Survey Outliers****Remèdes pour les valeurs aberrantes en échantillonnage****[MS-86]****A Weight Smoothing Method for Dealing with Stratum Jumpers in Business Surveys****Une méthode de lissage des poids pour faire face aux sauteuses de strate dans les enquêtes auprès des entreprises**Jean-François BEAUMONT, *Statistics Canada/Statistique Canada*, Louis-Paul RIVEST, *Université Laval*

The problem of stratum jumpers in stratified business surveys is mainly due to inaccurate size information at the time of sampling, which may result in assigning a large sampling weight to a large size unit. Such units may be quite influential and substantially inflate the variance of the estimates. They can be viewed as being outliers with respect to the implicit model used at the sampling stage. Therefore, standard robust estimation techniques can be used to handle this problem. Instead, we will consider an alternative weight smoothing method and illustrate its application using data from the Workplace and Employee Survey.

Le problème des “sauteuses” de strate dans les enquêtes stratifiées auprès des entreprises est principalement dû à des informations imparfaites sur la taille de l’entreprise au moment de l’échantillonnage, ce qui peut conduire à assigner un grand poids de sondage à une unité de grande taille. De telles unités peuvent avoir une grande influence et augmenter significativement la variance des estimations. Elles peuvent être vues comme étant des données aberrantes par rapport au modèle implicite utilisé à l’étape d’échantillonnage. Par conséquent, les techniques usuelles d’estimation robuste peuvent être utilisées pour tenir compte de ce problème. Au lieu de ces techniques, nous considérerons une méthode alternative de lissage des poids et illustrerons son application au moyen des données de l’Enquête sur le milieu de travail et les employés.

[MS-87]**Treating Influential Values in a Monthly Retail Trade Survey****Traiter les valeurs influentes dans une enquête mensuelle sur le commerce de détail**Mary MULRY & ROXANNE FELDPAUSCH, *U.S. Census Bureau*

Influential observations occur infrequently in establishment surveys but are problematic when they do appear. An observation is considered influential if it is true and its weighted contribution has an excessive effect on the estimated total. The goal of the study has been to propose methodologies that use such an observation but in a manner that allows a more appropriate impact on the estimated total. In particular, the study has applied Winsorization (Clarke 1995; Chambers, Kocic, Smith, and Cruddas 2000) and weighted M -estimation (Beaumont and Alavi 2004, Beaumont 2004) to estimate monthly revenue from the U.S. Monthly Retail Trade Survey.

Dans les enquêtes sur les établissements, les observations influentes sont rares, mais problématiques lorsque présentes. Une observation est considérée comme influente lorsqu’elle est véridique et que sa contribution pondérée a un effet excessif sur le total estimé. Le but de cette étude était de proposer des méthodologies utilisant une telle observation mais de manière à ce qu’elle ait un impact plus approprié sur le total estimé. En particulier, l’étude a appliqué la Winsorisation (Clarke 1995; Chambers, Kocic, Smith, et Cruddas 2000) et la M -estimation pondérée (Beaumont et Alavi 2004, Beaumont 2004) pour estimer le revenu moyen à partir de l’enquête américaine mensuelle sur le commerce de détail.

Session 06B**Mardi 12 juin • Tuesday, June 12, 10:30–12:00****AA-1046****Mathematical Finance and Stochastic Analysis II****Finance mathématique et analyse stochastique II****[MS-88]****Solutions of Consumption-investment Problems in Financial Markets with Regime Switching****Solutions aux problèmes de consommation-investissement dans des marchés financiers avec changement de régime**Abel CADENILLAS & LUZ SOTOMAYOR, *University of Alberta*

We consider a consumption and investment problem where the market presents multiple regimes. An investor taking decisions continuously in time selects a consumption-investment policy to maximize his expected total discounted utility of consumption until bankruptcy. The market coefficients and the investor's utility of consumption are dependent on the regime of the financial market, which is modeled by a finite-state continuous-time Markov chain. We obtain explicit optimal consumption and investment policies for specific HARA utility functions. We also show make an economic analysis of the solution.

Nous considérons un problème de consommation et d'investissement où le marché présente de multiples régimes. Un investisseur prenant des décisions de façon continue dans le temps choisit une politique de consommation-investissement dans le but de maximiser son utilité de consommation escomptée espérée totale jusqu'à la ruine. Les coefficients du marché et l'utilité de consommation de l'investisseur dépendent du régime du marché financier, ce qui est modélisé par une chaîne de Markov à temps continu ayant un nombre fini d'états. Nous obtenons, de façon explicite, des politiques optimales de consommation et d'investissement pour des fonctions d'utilité HARA spécifiques. Nous faisons également une analyse économique de la solution.

[MS-89]**Pricing Derivatives in Affine Term Structure Models****Fixation des prix de dérivées dans des modèles du terme de structure affine**Cody HYNDMAN, *Concordia University*

We consider the valuation of term structure derivatives in the context of affine term structure models. Extending our previous characterization of the zero coupon bond price as a forward-backward stochastic differential equation (FBSDE) to general transforms, combined with simulation methods for solving FBSDEs and transform inversion methods, we are able to provide numerical methods for calculating the prices of more complicated term structure derivatives. The pricing of swaptions in the context of affine term structure models presents an interesting example where our method seems to provide some advantages.

Nous considérons l'évaluation des dérivées du terme de structure dans le contexte de modèles du terme de structure affine. En généralisant notre caractérisation précédente du prix d'obligation à zéro coupon comme une équation différentielle stochastique avant-arrière (EDSAA) aux transformées généralisées, et en utilisant des méthodes pour résoudre des EDSAA et des méthodes d'inversion de transformées, nous pouvons obtenir des méthodes numériques pour calculer les prix de dérivées de terme de structure plus compliquées. La fixation des prix d'options swap dans le contexte de modèles du terme de structure affine est un exemple intéressant où notre méthode semble fournir quelques avantages.

[MS-90]

Modelling and Pricing of Variance Swaps for Multi-Factor Stochastic Volatilities with Delay
Modélisation et fixation des prix de swaps de variance pour volatilités stochastiques à plusieurs facteurs avec délai

Anatoliy SWISHCHUK, *University of Calgary*

Variance swaps for financial markets with underlying asset and multi-factor stochastic volatilities with delay are modelled and priced in this paper. We obtain some analytical closed forms for the expectation and variance of the realized continuously sampled variances for multi-factor (one-, two- and three-factor) stochastic volatilities with delay. As applications, we provide numerical examples using the S&P60 Canada Index (1998-2002) to price variance swaps with delay for all these models.

Les swaps de variance pour marchés financiers avec actif sous-jacent et volatilités stochastiques à plusieurs facteurs avec délai sont modélisés et leur prix sont fixés dans cet article. Nous obtenons quelques formes analytiques pour l'espérance et la variance des variances réalisées à échantillonnage continu pour des volatilités stochastiques avec délai à plusieurs (un, deux et trois) facteurs. Comme applications, nous fournissons des exemples numériques utilisant l'indice S&P60 Canada (1998-2002) pour déterminer le prix de swap de variance avec délai pour tous ces modèles.

Session 06C**Mardi 12 juin • Tuesday, June 12, 10:30–12:00****AA-1043**

Statistical Methods for Correlated Data

Méthodes statistiques pour données corrélées

[MS-91]

Meta-analysis of Cluster Randomization Trials for Binary Outcomes
Méta-analyse d'essais randomisés par grappes pour réponse binaire

Gerarda DARLINGTON, *University of Guelph*, Allan DONNER, *University of Western Ontario*

Cluster randomization trials are those that randomize intact social units of individuals, rather than individuals themselves, to different intervention groups. Results from several similar such studies may then be combined to test the overall effect of an intervention. Five possible approaches for meta-analysis of cluster randomization trials having a binary outcome are compared. Two modifications of the Mantel-Haenszel statistic, a general inverse variance approach and a procedure based on the Woolf statistic are considered. For simulation scenarios considered, an adjusted Mantel-Haenszel procedure provides a valid test with the greatest power slightly outperforming the general inverse variance approach.

Les essais randomisés par grappes sont ceux qui randomisent des unités sociales d'individus intacts, plutôt que les individus eux-mêmes, entre différents groupes d'intervention. Les résultats de telles études similaires peuvent alors être combinés afin de tester l'effet global d'une intervention. Cinq approches possibles pour la méta-analyse d'essais randomisés par grappes avec réponse binaire sont comparées. Deux modifications à la statistique Mantel-Haenszel, une approche de variance de l'inverse généralisée et une procédure basée sur la statistique de Woolf sont considérées. Pour les scénarios de simulation considérés, une procédure de Mantel-Haenszel ajustée fournit un test valide avec la plus grande puissance, légèrement plus performant qu'avec l'approche de variance de l'inverse généralisée.

[MS-92]**Tribute to a Champion of the Oceans
Hommage à un champion des océans**Joanna FLEMMING, *Dalhousie University*

Dr. Ransom Myers is referred to by many as a “Champion of the Oceans” for his research devoted to understanding the changes in marine biodiversity since the advent of industrial fishing. Ram was recently named by Fortune magazine as one of the World’s “Top 10 To Watch” for his impact in this area. In this talk I will present some “statistical issues” related to modeling correlated data that Ram (and others) believes should be better addressed by our discipline. These ideas have been captured in discussions with Ram over the past several years.

[Ram was recently diagnosed with brain cancer and died on March 27.]

Dr. Ransom Myers est décrit par plusieurs comme un “champion des océans” pour sa recherche sur les changements de la biodiversité marine depuis l’arrivée de la pêche industrielle. Ram a récemment été nommé par la revue Fortune l’une des “10 personnes à surveiller” pour son impact dans ce domaine. Dans cet exposé, je vais présenter certaines questions statistiques liées à la modélisation de données corrélées auxquelles Ram (et d’autres) croit que notre discipline devrait mieux répondre. Ces idées proviennent de discussions avec Ram au cours des dernières années.

[Ram a récemment été diagnostiqué avec un cancer cérébral et est décédé le 27 mars dernier.]

[MS-93]**Modeling Family-based Genetic Case-control Study Designs in the Presence of a Continuous Time Variable
Modélisation de plans d’études cas-témoins génétiques basés sur la famille pour une variable continue**Stephanie DIXON & Gerarda DARLINGTON, *University of Guelph*

Family-based genetic case-control designs are valuable for studies of complex diseases with low prevalence. Generally these studies have binary response with observed candidate genes and environmental factors that potentially contribute to disease status. In addition, correlation among family members must be considered. When a continuous time dependent variable, such as time of diagnosis, is present, survival analysis methods for correlated observations can be used. A literature review of current modelling strategies will be presented. Further, modifications to the partial likelihood will be described in an attempt to incorporate the ascertainment method of case-control data.

Les plans cas-témoins génétiques basés sur les familles sont utiles pour les études de maladies complexes à faible prévalence. Généralement, ces études ont une réponse binaire avec des gènes candidats et des facteurs environnementaux observés contribuant potentiellement au statut du sujet. De plus, la corrélation entre les membres d’une famille doit être considérée. Lorsqu’une variable dépendante continue, tel le temps de diagnostic, est présente dans l’étude, les méthodes d’analyse de survie pour observation corrélées peuvent être utilisées. Une revue de littérature des stratégies de modélisation sera présentée. De plus, des modifications à la fonction de vraisemblance partielle seront décrites dans une tentative d’incorporer la méthode d’identification des données cas-témoins.

Session 06D**Mardi 12 juin • Tuesday, June 12, 10:30–11:30****IIC-2001**

Isobel Loutit Invited Address on Business and Industrial Statistics
 Allocution Isobel Loutit sur la statistique industrielle et de gestion

[MS-94]

Statistics in Business and Industry: A Historical Perspective and Ideas for Implementation
La statistique dans l'industrie : une perspective historique et des idées de mise en œuvre

BOVAS ABRAHAM, *University of Waterloo*

We start with a brief outline of the early developments in business and industrial statistics. Statisticians, over the years, have devised many tools to be utilized for general business and industrial improvement. However, there is a wide gap between the available tools and what are practiced in business and industry. Thus it is important for statisticians to direct serious attention to bridging this gap if statistics is to be relevant in business and industry and to the society at large. In the latter part of the presentation we will look at ideas for implementing statistics in business and industry.

Nous commençons par un bref aperçu des premiers développements dans les statistiques industrielles et de gestion. Les statisticiens, au fil des ans, ont développé plusieurs outils utilisés pour des améliorations générales en industrie et en gestion. Cependant, il y a un grand écart entre les outils disponibles et ce qui est mis en pratique dans l'industrie et le monde des affaires. Il est donc important pour les statisticiens de concentrer leurs efforts afin de combler cet écart pour que la statistique soit pertinente dans l'industrie et la gestion, et pour la société en général. Dans la partie finale de cette présentation, nous allons regarder des idées relatives à la mise en œuvre de statistiques en industrie et en gestion.

Session 06E**Mardi 12 juin • Tuesday, June 12, 10:30–12:00****SN-2105**

Forestry
 Foresterie

[MS-95]

Fire Simulation using the Prometheus Model
Simulation d'incendies à l'aide du modèle Prométhée

Cordy TYMSTRA, *Alberta Sustainable Resource Development*

In 1971, the combined efforts of a statistician in United States, and a fire researcher in Canada, resulted in the development of the first computer model to simulate fire growth. Since then, many advancements have been made with valuable contributions from mathematicians and statisticians. Fire management agencies across Canada currently use a wave propagation model called Prometheus. This is a spatially explicit, deterministic fire simulation model. The vertex propagation results in complex behaviours (loops and knots) that are challenging to resolve. There is also interest to incorporate stochasticity in the model to account for uncertainty and variability in the fire environment.

En 1971, les efforts combinés d'un statisticien aux États-Unis et d'un chercheur étudiant les feux au Canada ont mené au développement du premier modèle informatique simulant la croissance des feux. Depuis, plusieurs avancées ont été faites grâce à des contributions de mathématiciens et de statisticiens. Les agences de gestion des feux au Canada utilisent présentement un modèle de propagation des ondes appelé Prométhée. Il s'agit d'un modèle de simulation des feux déterministe et spatialement explicite. La propagation des vertex a pour résultat des comportements complexes (boucles et noeuds) qui sont difficiles à résoudre. Il y a aussi un intérêt pour l'incorporation de stochasticité dans le modèle afin de tenir compte de l'incertitude et de la variabilité dans l'environnement du feu.

[MS-96]**Stochastic Lighting-fire Risk Models: Past, Present and Future****Modèles stochastiques du risque de feux causés par la foudre : passé, présent et futur**Douglas WOOLFORD, *University of Toronto*

Accurately assessing the risk of lightning ignited fires is of great interest in the forest management community. Investigations into this problem have been furthered by the advent of nationally implemented fire danger rating systems, technological advances in lightning detection systems, and increases in computing power. Emphasizing the problem of assessing lightning-ignition risk for Ontario, an overview of early models, models currently in development, and potential future directions for this research is discussed.

Évaluer précisément le risque de feux de forêts dus à la foudre est d'un grand intérêt pour la communauté de la gestion des forêts. Les investigations dans ce domaine ont été aidées par la mise en oeuvre de systèmes d'évaluations nationaux du danger de feu, par les avancées technologiques en détection de la foudre, et par l'augmentation de la puissance informatique. En mettant l'accent sur le problème d'évaluation de risque de déclenchement de feux dus à la foudre en Ontario, nous passons en revue les premiers modèles, et discutons des modèles en cours de développement, et des orientations futures possibles dans ce domaine de recherche.

[MS-97]**Analysing Robustness of Forest Operational Plans****Analyser la robustesse des plans opérationnels pour les forêts**Jason MYERS, *University of Toronto*

Production in the wood supply chain is complicated by the uncertain environment, markets, and lack of good inventory information. The goals of our research are to understand sources of uncertainty, quantify their effect on forest operations and plans, quantify the value of improved information, model robustness of decision models, and explore implications for Hierarchical Planning in Forest Management. This paper describes a methodology to assess the value of better information in a typical planning environment, by developing a simulation/optimization model that mimics the operational planning process, including the decision maker's ability to adapt to new information through recourse decisions.

La production dans la chaîne logistique du bois est compliquée par l'environnement et les marchés incertains, et par le manque d'information sur l'inventaire. Les buts de notre recherche sont de comprendre les sources d'incertitude, de quantifier leur effet sur les opérations et plans forestiers, de quantifier l'utilité d'une meilleure information, de modéliser la robustesse des modèles de décision, et d'explorer les implications pour la planification hiérarchique en gestion des forêts. Cet article décrit une méthodologie permettant d'évaluer l'utilité d'une meilleure information dans un environnement de planification typique, en développant un modèle de simulation/optimisation imitant le processus de planification opérationnelle, incluant la capacité du décideur à s'adapter à de nouvelles informations au moyen d'actions de recours.

Session 06F**Mardi 12 juin • Tuesday, June 12, 10:30–11:45****SN-2041**

Statistical Genomics

Génomique statistique

[MS-98]**An Approximate Likelihood Ratio Test for Testing Means of Generalized Logistic Distribution of Type III and Implications for Detecting Differential Expression in Microarrays****Un test du rapport de vraisemblance approximatif pour la moyenne d'une distribution logistique généralisée de type III et ses implications dans la détection d'expressions différentielles dans les micropuces**Ahmed HOSSAIN, Joseph BEYENE & Andrew WILLAN, *University of Toronto and SickKids Research Institute*

Most available popular methods in the study of microarray data analysis adjust the normal distribution assumption and are based on Wald statistic. Due to the quality of the statistic used to describe the difference among groups of data, there are situations when these methods are inefficient, especially when expression levels are skewed. We recommend solution for violations of the normality assumption is to use a Generalized Logistic Distribution of Type II (GLDII) model and provide an Approximate Likelihood Ratio Test (ALRT) to test the means of the GLDII. The performance of the method is demonstrated in a simulation study.

La plupart des méthodes disponibles pour l'étude de données de micropuces présupposent une distribution normale et sont basées sur la statistique de Wald. À cause de la qualité de la statistique utilisée pour décrire la différence entre les groupes de données, il y a des situations où ces méthodes sont inefficaces, particulièrement lorsque les niveaux d'expression sont asymétriques. Nous recommandons, comme solution au non respect de l'hypothèse de normalité, l'utilisation du modèle de la distribution logistique généralisée de type 2 (DLG2) et présentons un test du rapport de vraisemblance approximatif pour tester les moyennes de la DLG2. La performance de la méthode est analysée à l'aide d'une étude de simulation.

[MS-99]**A Further Note on Controlling the False Discovery Rate****Une note additionnelle sur le contrôle du taux de fausses découvertes**Yaling YIN, Miķelis BICKIS & Chris SOTEROS, *University of Saskatchewan*

Storey has proposed a procedure and shown numerically that it gains more power than the Benjamini-Hochberg procedure while controlling the same false discovery rate FDR. However, when the number of hypotheses is small and the proportion of true null hypotheses is large, we prove that the reverse result is possible. When two procedures produce the same FDR, the difference between their powers increases as the distance between the null and alternative distributions decreases. We also show by simulation that, modifying the Benjamini-Hochberg approach to incorporate an estimate of the proportion of true null hypotheses, gives a procedure with superior power.

Storey a proposé une procédure et montré numériquement qu'elle gagne plus de puissance que la procédure Benjamini-Hochberg tout en contrôlant le même taux de fausse découverte (TFD). Cependant, lorsque le nombre d'hypothèses est petit et que la proportion d'hypothèses nulles vraies est grande, nous prouvons que le résultat inverse est possible. Lorsque deux procédures produisent le même TFD, la différence entre leurs puissances augmente lorsque la distance entre l'hypothèse nulle et l'alternative diminue. Nous montrons aussi par simulation que, en modifiant l'approche Benjamini-Hochberg pour incorporer un estimé de la proportion de vraies hypothèses nulles, nous obtenons une procédure de puissance supérieure.

[MS-100]**Bootstrapping of Gene Expression Data****Utilisation du bootstrap avec des données d'expression génétique**Shaheena BASHIR & Angelo CANTY, *McMaster University*

Gene microarrays study is a challenging field. The purpose is to identify the differentially expressed genes to identify the causes of some disease. The challenges occur due to a number of reasons; gene expressions for a very large number of genes are available as compared to very small sample size, and possibility of high level of non-normal random noises. Permutation tests are useful in this context. In this talk, we will discuss how permutation tests based on bootstrapping can be helpful in identifying differentially expressed genes. The results are based on the investigation of the T1D in mice model.

L'étude de puces à ADN est un domaine stimulant. Le but est d'identifier les gènes exprimés différemment pour identifier la cause d'une certaine maladie. Les défis proviennent de plusieurs raisons ; les expressions génétiques sont disponibles pour un très grand nombre de gènes comparativement à une très faible taille échantillonnale, et il y a possibilité d'un haut niveau de bruit aléatoire non-normal. Les tests de permutation sont utiles dans ce contexte. Dans cet exposé, nous allons discuter de l'utilité des tests de permutation basés sur le bootstrap dans l'identification de gènes à expression différentielle. Les résultats sont basés sur une investigation du modèle du T1D chez la souris.

[MS-101]**A Scan Statistic for Identifying Chromosomal Regions of Significant Association****Une statistique de balayage pour identifier les régions chromosomales d'association significative**Jennifer ASIMIT & Shelley BULL, *Samuel Lunenfeld Research Institute, University of Toronto*

We propose a scan statistic for identifying clusters of genetic markers exhibiting association between gene expression (GE) and DNA copy number (CN) in breast tumour microarrays. At each genetic marker, as a measure of the association between GE and CN, we regress GE on CN. In the development of the scan statistic, a marked Poisson process along each chromosome is considered. The occurrence of a genetic marker on a chromosome is the event of interest, with the mark being the associated regression coefficient. Regions identified as clusters of significant associations are hypothesized to harbour genes involved in breast cancer progression.

Nous proposons une statistique de balayage pour identifier les groupes de marqueurs génétiques démontrant une association entre l'expression génétique (EG) et le nombre de copies (NC) d'ADN dans les micropuces étudiant le cancer du sein. À chaque marqueur génétique, comme mesure d'association entre EG et NC, nous régressons EG sur NC. Dans le développement de la statistique de balayage, un processus de Poisson marqué est considéré pour chaque chromosome. L'occurrence d'un marqueur génétique sur un chromosome est l'événement d'intérêt, alors que la marque est le coefficient de régression associé. Nous faisons l'hypothèse que les régions identifiées comme groupe d'associations significatives contiennent des gènes impliqués dans la progression du cancer du sein.

[MS-102]**Statistical Quality Assessment for High-dimensional Genomic Measurements****Évaluation de la qualité statistique en présence de mesures génomiques de dimension élevée**Julia BRETTSCHEIDER, *Queen's University*

Given correlated measurements of tens of thousands of different variables with a few replicates each, how can one estimate their precision and accuracy? This question is motivated by new molecular assay technologies that have opened up new

Comment estimer la précision et la justesse de mesures corrélées de dizaines de milliers de variables différentes, chacune ayant à peine quelques mesures répétées ? Les nouvelles technologies de dosage moléculaire ont ouvert de nouvelles voies à la recherche scientifique, et soulèvent cette

avenues of research for scientists. A weak point of such research is the reproductibility. I introduce statistical methods to numerically and spatially assess microarray data quality in absence of “gold-standards”. The methods are based on residuals obtained from a robust fit a linear model that reflects the measurement process. Performance of the quality assessment methods will be demonstrated on a variety of datasets.

question. Un point faible d’une telle recherche est sa reproductibilité. Je présente des méthodes statistiques permettant d’évaluer la qualité de données de micropuces en l’absence de standards de référence. Ces méthodes sont basées sur des résidus obtenus en ajustant un modèle linéaire robuste tenant compte du processus de mesure. La performance des méthodes d’évaluation de la qualité sera illustrée à l’aide d’une variété de jeux de données.

Session 07A**Mardi 12 juin • Tuesday, June 12, 13:30–14:30****AA-1043**

Special Invited Session of the Probability Section

Allocution de l’invité d’honneur du groupe de probabilité

[MS-103]

Insurance Claims Modulated by a Hidden Marked Point Process**Primes d’assurance modulées par un processus ponctuel marqué caché**Robert ELLIOTT, *University of Calgary*, Tak Kuen SIU, *Heriot Watt University*, Hailiang YANG, *University of Hong Kong*

Recently there has been considerable interest in the applications of Markov-modulated compound Poisson models in ruin theory. A Markov-modulated compound Poisson model can incorporate the impact of the states of an economy on the surplus process. However, the Markov chain in the model is not observable in practice. In this paper, we develop a way for filtering and smoothing the hidden states of the economy underlying the Markov-modulated compound Poisson model. We first suppose that the stochastic intensity of the compound Poisson observation process switches over time according to a continuous-time finite-state hidden Markov chain. Then we consider the case when both the stochastic intensity and the distribution of the claim sizes of the compound Poisson process depend on the Markov chain. The filter and smoother for the hidden states provide us with an appropriate method to select or estimate a risk model in the mean-squared-error sense. We also discuss the estimation of the unknown parameters in the two Markov-modulated compound Poisson models using the robust filter-based and smoother-based EM algorithms.

Récemment, un intérêt considérable s’est développé autour des modèles de Poisson composé modulé par un processus de Markov et de leurs applications dans la théorie de la ruine. Un modèle de Poisson composé modulé par un processus de Markov peut incorporer l’impact des différents états d’une économie sur le processus de surplus. Cependant, en pratique la chaîne de Markov dans ce modèle n’est pas observable. Dans cet article, nous développons un moyen de filtrer et de lisser les états cachés de l’économie sous-jacent au modèle de Poisson composé modulé par un processus de Markov. Dans un premier temps nous supposons que l’intensité stochastique du processus d’observation de Poisson composé change au cours du temps en suivant une chaîne de Markov cachée à espace d’état fini et à temps continu. Nous considérons alors le cas où à la fois l’intensité stochastique et la distribution des valeurs de la demande du processus de Poisson composé dépendent de la chaîne de Markov. Le filtre et le lisseur pour les états cachés nous permettent en utilisant une méthode appropriée de sélectionner et d’estimer un modèle de risque au sens des moindres carrés. Nous discutons aussi du problème d’estimation des paramètres inconnus dans les modèles de Poisson composés modulés par un processus de Markov en utilisant le filtre et le lisseur combinés avec des algorithmes EM.

Session 07B Mardi 12 juin • Tuesday, June 12, 13:30–15:00 AA-1046**Statistical Genetics
Génétique statistique****[MS-104]****Methodological Issues in Penetrance Estimation Studies
Problèmes méthodologiques dans les études d'estimation de pénétrance**

Laurent BRIOLLAIS & Yunhee CHOI, *Samuel Lunenfeld Research Institute*, Karen KOPCIUK, *Alberta Cancer Board*, Tariq HASAN, *University of New Brunswick*

Characterization of disease genes through the estimation of a cumulative risk function (i.e. the penetrance) is particularly important for the genetic counseling of individuals who are identified as carriers but poses important statistical challenges. In this presentation, we will give an overview of these challenges, including the choice of an efficient design, the modeling of correlated familial data, and the potential biases due to missing data and competing risks. We will discuss these problems in light of previous penetrance studies and our own analysis on breast and colorectal cancer data from Ontario and Newfoundland.

La caractérisation de gènes de maladie à partir de l'estimation de la fonction de risque cumulé (c.à.d. la pénétrance) est particulièrement importante pour le conseil génétique des individus qui sont identifiés comme porteurs mais pose des défis statistiques importants. Dans cette présentation, nous donnerons un aperçu de ces défis qui englobent le choix d'un plan expérimental efficace, la modélisation de données familiales corrélées et les biais potentiels dus aux données manquantes et risques compétitifs. Nous discuterons ces problèmes à la lumière d'études de pénétrance antérieures et de nos propres analyses sur des données de cancer du sein et colorectal de l'Ontario et de Terre-Neuve.

[MS-105]**A Graphical Tool for Exploring Gene-environment Interaction in Case-parent Trios
Un outil graphique pour explorer l'interaction gène-environnement dans les trios cas-parent**

Jinko GRAHAM & Linnea DUKE, *Simon Fraser University*

We propose a simple smoother for exploring statistical interaction (GxA) between a SNP (G) and a nongenetic risk factor (A) in case-parent trios. The smoother can be used as a diagnostic tool for checking the presence of GxA after conducting a genetic association test, such as the TDT, that does not account for interaction. Alternately, if an interaction model is fit, the smoother can be used to check model adequacy. We condition on parental genotypes to avoid the false impression of interaction when the assumed G-A independence of a case-only analysis is violated. Some possibilities for extension are raised.

Nous proposons un lisseur simple pour l'exploration de l'interaction statistique (GxA) entre un SNP (G) et un facteur de risque non génétique (A) dans des trios cas-parents. Le lisseur peut être utilisé comme outil diagnostique vérifiant la présence de GxA après avoir conduit un test d'association génétique, tel l'étude de déséquilibre de transmission, qui ne tient pas compte de l'interaction. De façon alternative, si un modèle d'interaction est utilisé, le lisseur peut être utilisé pour vérifier l'adéquation du modèle. Nous conditionnons sur les génotypes parentaux afin d'éviter l'impression fautive d'une interaction lorsque l'indépendance G-A d'une analyse considérant les cas uniquement n'est pas respectée. Quelques possibilités de généralisations sont formulées.

[MS-106]**Network-based Measurement of Gene-set Enrichment****Mesure basée sur les réseaux pour l'enrichissement d'un ensemble de gènes**Michael NEWTON, *University of Wisconsin*

A problem in expression data analysis is how to incorporate exogenous biological information organized in networks of gene sets. The gene ontology project (GO) provides the best example of such networks. On a given set of genes one can ask if the proportion of expression-altered genes is unusual compared to that same proportion in a random set of genes. To address limitations of the one-set-at-a-time-versus-the-whole-system approach, I will present a solution that gauges the significance of one set's enrichment in terms of random subsets of that set's parents in the GO graph. Empirical and theoretical evidence support this strategy.

Un problème dans l'analyse de données d'expression est l'incorporation d'information biologique exogène organisée dans des réseaux d'ensembles de gènes. Le projet d'ontologie des gènes (OG) constitue le meilleur exemple de tels réseaux. Il est possible de se demander si la proportion de gènes à expression altérée dans un ensemble donné de gènes est inhabituelle comparativement à cette proportion pour une ensemble aléatoire de gènes. Pour répondre aux limitations de l'approche "un-ensemble-à-la-fois-contre-tout-le-système", je vais présenter une solution qui jauge la signification de l'enrichissement d'un ensemble en termes de sous-ensembles aléatoires des parents de cet ensemble dans le graphe OG. Des justifications empiriques et théoriques supportent cette stratégie.

Session 07C**Mardi 12 juin • Tuesday, June 12, 13:30–15:00****IIC-2001****Statistical Graphics and Image Analysis****Graphiques statistiques et analyse d'image****[MS-107]****Real-Time Computerized Annotation of Pictures via D2-Clustering****Annotation informatisée d'images en temps réel par la classification D2**Jia LI, *Pennsylvania State University*

Automated annotation of digital pictures is a highly challenging technology with potentially significant applications. In this work, an image is represented by a discrete distribution whose support contains a variable number of arbitrary vectors, each describing one segmented region. A new clustering algorithm, namely D2-clustering (D2 stands for discrete distribution), is developed to minimize the sum of Mallows distance between sample points and their corresponding cluster centroids. A mixture model is proposed based on D2-clustering and a distance preserving mapping to vector spaces. The ALIPR system is built using these methods and has achieved real-time annotation with good accuracy.

L'annotation automatisée d'images digitales est une technologie présentant plusieurs défis et ayant des applications potentiellement significatives. Dans ce travail, une image est représentée par une distribution discrète dont le support contient un nombre variable de vecteurs arbitraires, décrivant chacun une région fragmentée. Un nouvel algorithme de classification, nommé classification D2 (pour distribution discrète), est développé afin de minimiser la somme des distance de Mallow entre les point échantillonnage et les centroïdes des groupes correspondants. Un modèle de mélanges basé sur la classification D2 et une fonction vers les espaces vectoriels préservant les distances. Le système ALIPR est construit en utilisant ces méthodes et permet d'obtenir une annotation en temps réel avec une bonne précision.

[MS-108]**Detection of Circular Shapes of Fungi and Bacteria Colonies****Détermination de formes circulaires des colonies cryptogamiques et bactériennes**

Christine MUELLER, *University of Kassel*, Tim GARLIPP, *University of Oldenburg*

The forms of bacteria and fungi colonies are often more or less regular disks with different radii. We present here a method to detect the circles of these colonies from photos given as pixel arrays of grey values. The first step is to detect edge candidates, i.e. pixel positions which are close to the borders of the colonies. This is done by a method of Qiu (1997) based on rotational difference kernel estimators and a multiple testing procedure. In a second step, the circles are detected from the edge candidates with a method based on redescending M-estimators.

Les colonies de bactéries et de champignons ont souvent la forme de disque plus ou moins régulier à rayons variables. Nous présentons ici une méthode afin de détecter les cercles de ces colonies à partir de photographies présentées sous la forme d'un tableau de pixels prenant différents tons de gris. La première étape est de détecter les candidats possibles pour les contours, c.-à-d. les pixels près des frontières des colonies. Ceci est fait avec la méthode de Qiu (1997) basée sur les estimateurs à noyaux de différence rotationnelle et une procédure pour des tests multiples. Dans la seconde étape, les cercles sont détectés à partir des contours possibles à l'aide d'une méthode basée sur des M-estimateurs redescendants.

[MS-109]**Nonparametric Estimation of the Point Spread Function for Blind Image Deblurring****Estimation nonparamétrique de la fonction de dispersion des points en correction d'images flées**

Peihua QIU, *University of Minnesota*

Observed images are usually blurred versions of the true images. The major purpose of image deblurring is to restore the original image from its blurred version. A blurred image can be described by the convolution of the original image with a point spread function (psf) that characterizes the blurring mechanism. Thus, one essential problem for image deblurring is to estimate the psf, which turns out to be a challenging task. Most existing image deblurring procedures assume that either the psf is completely known or it has a parametric form. Motivated by some image applications, we suggest a method for estimating the psf nonparametrically in certain cases. Both theoretical and numerical studies show that it works well in applications.

Les images observées sont souvent des versions brouillées des images réelles. Le principal but du débrouillage d'image est de restaurer l'image originale à partir de l'image brouillée. Une image brouillée peut être décrite comme une convolution de l'image originale avec une fonction de dispersion des points (fdp) caractérisant le brouillage. L'estimation de la fdp est donc un problème crucial mais difficile en débrouillage d'image. La plupart des procédures de débrouillage d'image existantes supposent que la fdp est entièrement connue ou qu'elle est de forme paramétrique. Motivés par quelques applications, nous suggérons une méthode afin d'étudier la fdp non paramétriquement dans certains cas. Des études théoriques et numériques démontrent que cette approche fonctionne bien pour ces applications.

Session 07D**Mardi 12 juin • Tuesday, June 12, 13:30–15:00****SN-2105**

Statistics Applied to Hydrology and Climatology

Statistique appliquée à l'hydrologie et à la climatologie

[MS-110]**Risk Assessment under Climatic Change : An Impossible Mission?****Évaluer le risque en situation de changement climatique : mission impossible ?**Éric PARENT & Jacques BERNIER, *MORSE AgroParisTech*, Vincent FORTIN, *Météorologie Canada*

Traditional design values for engineering works rely on the notion of “return period”. A return period needs a stationary framework to get a rigorous theoretical meaning: under changing climatic conditions other alternative notions must be found to cope with risk assessment and other models are to be developed to work with non stationary geophysical phenomena. Taking the Senegal river annual hydrological series as an example of non-stationary behaviour, we report our first investigations on a shifting level model and study how to give a design value equivalent to a 100 year return period for drought risk assessment.

Les valeurs de projet de l'ingénierie traditionnelle s'appuient sur la notion de “période de retour”. La “période de retour” n'a de sens que dans un contexte stationnaire : en situation de changement climatique, il faut chercher d'autres outils à substituer à la “période de retour” et d'autres modèles pour relâcher les hypothèses iid d'occurrence des phénomènes géophysiques ! Sur un exemple d'apports annuels du fleuve Sénégal, nous rapportons les résultats de nos premières investigations sur un modèle à changements aléatoires de niveaux et nous étudions comment donner un équivalent à une “période de retour” centennale pour le risque de sécheresse.

[MS-111]**Statistician's Role in Climate Change Research****Le rôle d'un statisticien dans la recherche en changement climatique**Jean-François ANGERS, *Université de Montréal*

During this talk, I will speak about the role I've had to play as a statistician, in various research projects on climate change and environment. I met the principal instigator of these projects while on sabbatical leave at Ouranos (Consortium on Regional Climatology and Adaptation to Climate Change). During this presentation, I will briefly discuss the complexity of the different research projects. However, I will mostly focus on the important aspects of these various projects where a statistician can play an important part in the investigation. Finally, I will talk about the communication problems between the different groups of researchers.

Dans cette présentation, je parlerai du rôle que j'ai joué en tant que statisticien dans différents projets de recherche sur les changements climatiques et l'environnement. J'ai rencontré les responsables de ces projets lors d'une année sabbatique passée à Ouranos (Consortium sur la climatologie régionale et l'adaptation aux changements climatiques). Durant cette présentation, je parlerai brièvement de la complexité des projets de recherche concernés. Par contre, j'insisterai davantage sur des aspects importants de ces différents projets où les statisticiens peuvent être conduits à jouer un rôle important. Finalement, j'aborderai aussi les problèmes de communication entre les différents groupes de scientifiques.

[MS-112]**Assessing Hydrological Forecasts at Hydro-Québec
Évaluation des prévisions hydrologiques à Hydro-Québec**

LUC PERREAULT, *Institut de recherche d'Hydro-Québec*, Jocelyn GAUDET, *Hydro-Québec*, James MERLEAU, *Institut de recherche d'Hydro-Québec*

Hydrological forecasting is central for hydropower production and risk management related to extreme events. In hydrology and climatology, deterministic forecasting predominates, but Hydro-Québec distinguishes itself because it produces probabilistic hydrometeorological forecasts since the early 80's. But how can we evaluate a probabilistic forecast? A quality control system has been developed and recently implemented to assess the hydrological forecasts produced daily for more than 90 basins managed by Hydro-Québec. In this presentation, we present the strategies and the scoring rules considered to measure the performance of the hydrological forecasting system.

La prévision hydrologique constitue la clé d'une gestion efficace de la production hydroélectrique ainsi que du contrôle actif des risques liés aux événements extrêmes. Dans un monde où la prévision déterministe prédomine, Hydro-Québec se distingue car elle produit des prévisions probabilistes depuis plus de 20 ans. Toutefois, ce n'est que tout récemment qu'elle s'est dotée d'un système de contrôle de qualité adapté aux prévisions probabilistes qu'elle émet quotidiennement pour plus de 90 bassins versants. Comment quantifier la qualité d'une telle prévision? Dans cette communication, nous présentons les stratégies d'évaluation et les scores adoptés par Hydro-Québec.

Session 07E**Mardi 12 juin • Tuesday, June 12, 13:30–15:00****SN-2109****The National Institute for Complex Data Structures: Our Community in Scientific Collaborations**

l'Institut national pour les structures de données complexes : Notre communauté dans des collaborations scientifiques

[MS-113]**Disease Models using Latent Classes in the Presence of Missing Data: The Example of Delirium
Modèles de maladies utilisant des classes latentes en présence de valeurs manquantes : l'exemple du délire**

Alina DYACHENKO, Antonio CIAMPI & Russell STEELE, *McGill University*

We describe disease experience through a model based on latent classes and a hidden Markov chain; we apply these concepts to delirium data. The number of latent classes is determined using the BIC and the AIC. Missing data are handled through a pragmatic multiple imputation approach: we create multiple imputed versions of the data, calculate the AIC and BIC criteria on each of the imputed versions, and choose the number of latent classes by minimizing the average of the AIC and BIC over the imputation complete data sets. Our empirical approach will be compared with a fully Bayesian one.

Nous décrivons l'expérience de la maladie au moyen d'un modèle basé sur les classes latentes et une chaîne de Markov cachée; nous appliquons ces concepts à des données de délire. Le nombre de classes latentes est déterminé en utilisant le BIC et l'AIC. Les données manquantes sont traitées avec une approche d'imputation multiple pragmatique: nous créons plusieurs versions imputées des données manquantes, calculons les critères AIC et BIC pour chaque version imputée, et choisissons le nombre de classes latentes en minimisant la moyenne de l'AIC et du BIC. Notre approche empirique sera comparée à une approche complètement bayésienne.

[MS-114]**Proximity-graphs-based Clustering of Micro-array Probes
Classification de sondes de puces à ADN basée sur des graphes de proximité**

Antonio CIAMPI, Imad KHOURY & Godfried TOUSSAINT, *McGill University*, Sadora ANTONIANO, *IIMAS-UNAM, Mexico*, Carl MURIE & Robert NADON, *McGill University*

Clustering algorithms are often criticised for suggesting artefactual structures. We demonstrate with an example from micro-array analysis that proximity graphs provide visual evidence in favour or against the existence of clusters. In calibration, we attempt to describe the relationship between intensity and concentration. If there are natural clusters in the space of the probes, a locally linear model can be developed by attaching a distinct calibration equation to each cluster. One may hope that this is better than one linear model. With the use of proximity graphs for data visualization, we conclude that there are no clusters in our data.

Une critique souvent faite aux algorithmes de classification est le fait que les structures qu'ils suggèrent soient des artefacts. Nous démontrons à l'aide d'un exemple d'analyse de données de puces à ADN que les graphes de proximité apportent de l'évidence visuelle pour ou contre l'existence de grappes. Lors de la calibration, nous tentons de décrire la relation entre l'intensité et la concentration. S'il y a des grappes naturelles dans l'espace des sondes, un modèle localement linéaire peut être développé en associant une équation de calibration distincte à chaque grappe. On peut espérer que cette approche soit meilleure qu'un seul modèle linéaire. En utilisant les graphes de proximité pour visualiser nos données, nous concluons qu'il n'y a pas de grappes présentes dans celles-ci.

[MS-115]**Variable Selection for Gaussian Process Models in Computer Experiments
Sélection de modèle pour modèles de processus gaussiens dans des expériences informatiques**

Crystal LINKLETTER, *Simon Fraser University*

Simulation of complex phenomena often requires a large number of inputs and is computationally expensive. Identifying the inputs that most impact the system can be a critical step. Here, we introduce a new, simple method for identifying active factors in Gaussian process models for computer screening experiments. The posterior distribution of an inert factor is used as a reference distribution against which the importance of true experimental factors can be assessed.

La simulation de phénomènes complexes nécessite souvent un grand nombre d'entrées et beaucoup d'opérations. L'identification des entrées ayant le plus grand impact sur le système peut être une étape cruciale. Nous introduisons ici une nouvelle méthode simple pour identifier les facteurs actifs dans des modèles de processus gaussiens pour expériences informatiques de sélection. La distribution a posteriori d'un facteur inerte est utilisée comme distribution de référence contre laquelle l'importance des facteurs expérimentaux réels peut être confrontée.

[MS-116]**Dimension Reduction of Genome-wide SNP Data by Sparse Principal Components Analysis
Réduction de dimension de données SNP sur tout le génome par analyse en composantes principales creuse**

Xiaofei SHI & Rafal KUSTRA, *University of Toronto*

Principal component analysis (PCA) is widely used in dimensionality reduction. To make the results from PCA more interpretable, Zhou *et al.* (2004) developed sparse principal component analysis (SPCA) using the lasso to produce modified principal components with sparse loadings. Here we implement SPCA in a large colorectal cancer project

L'analyse en composantes principales (ACP) est fréquemment utilisée en réduction de dimension. Afin de rendre les résultats de l'ACP plus interprétables, Zhou *et al.* (2004) ont développé l'analyse en composantes principales creuse (ACPC) utilisant la méthode du lasso pour produire des composantes principales modifiées ayant une matrice de pondération creuse. Nous utilisons ici l'ACPC dans un pro-

(ARCTIC) in order to construct a prediction model using over 600k single nucleotide polymorphisms (SNPs). The subset of SNPs included for Sparse PCA since is based univariate p -values for testing the association with the disease using a sparse version of a supervised PCA framework described by Bair *et al.*, 2006. We will build and compare the prediction models constructed on real genomic data using Sparse PCA and other techniques. Area under curve (AUC) will be used for comparison.

jet d'étude du cancer colorectal de grande envergure (ARCTIC) dans le but de construire un modèle de prévision utilisant plus de 600k SNPs. Le sous-ensemble de SNPs inclus pour l'ACPC est basé sur les valeurs- p univariées des tests d'association avec la maladie utilisant une version creuse du cadre d'ACP supervisée décrit par Bair *et al.*, 2006. Nous allons construire et comparer les modèles de prévision à partir de données génomiques réelles basés sur l'ACPC et d'autres méthodes. L'aire sous la courbe sera utilisée pour comparer les modèles.

[MS-117]

Graphical Models for Microarray Expression Data based on Low-order Partial Correlations

Modèles graphiques pour données d'expression de micro-puces basés sur les corrélations partielles de bas ordre

Nicholas MITSAKAKIS & Rafal KUSTRA, *University of Toronto*

Gaussian Graphical Models (GGM) are statistical tools to analyze the joint distribution of continuous random vectors, by representing conditional (in-)dependencies between the variates in concentration graphs. Such graphs can be used in genomic studies in order to construct gene association networks from gene expression data. Since the number of variables (genes) is usually much larger than the sample size, estimation of, and inference on, full partial correlations between the variables is challenging. Instead, graphs based on partial correlations of low order have been proven to provide a good approximation to concentration graphs. Here we investigate the empirical null distribution of partial correlations of low but increasing order and how this can be used to infer full conditional dependencies between the variables. We utilize our findings in the implementation of a novel algorithmic approach for covariance selection and graph construction. We apply and evaluate our method on synthetic and real biological data.

Les modèles graphiques gaussiens sont des outils statistiques utilisés afin d'analyser la distribution conjointe de vecteurs aléatoires continus, en représentant les (in-)dépendances conditionnelles entre les variables dans des graphes de concentration. De tels graphes peuvent être utilisés dans des études génomiques afin de construire des réseaux d'associations entre les gènes à partir de données d'expression génétique. Puisque le nombre de variables (les gènes) est habituellement beaucoup plus grand que la taille de l'échantillon, l'estimation des corrélations partielles complètes (et les inférences sur ces corrélations) sera difficile. Il a été démontré que des graphes basés sur les corrélations partielles de bas ordre fournissent une bonne approximation des graphes de concentration. Ici, nous étudions la distribution nulle empirique des corrélations partiellement d'ordre bas mais croissant, et comment ceci peut être utilisé afin d'inférer les dépendances conditionnelles complètes entre les variables. Nous utilisons nos résultats dans la mise en oeuvre d'une nouvelle approche algorithmique à la sélection de covariance et à la construction de graphes. Nous appliquons et évaluons notre méthode à des données biologiques réelles et synthétiques.

[MS-118]

An Analysis of Bacterioplankton Abundance Levels using a Cyclical State Space Model

Une analyse des niveaux d'abondance de bactérioplancton utilisant un modèle d'espace des états cyclique

Elizabeth ALMEIDA, Michael DOWD & Joanna MILLS-FLEMMING, *Dalhousie University*

Abundance levels for various types of plankton and other biogeochemical variables were recorded between 1992 and 2005 from the Bedford Basin, NS, Canada as part of an on going oceanographic mon-

Dans le cadre d'un programme d'observation océanographique, les niveaux d'abondance de divers types de plancton, ainsi que d'autres variables biogéochimiques, furent enregistrés dans le bassin Bedford entre 1992 et 2005. Les résul-

itoring program. Results are presented from time series analysis of these data. In particular, a cyclical state space model is developed and applied to these data. This allows for the seasonal cycle of the abundance levels to be adaptive, thereby allowing us to examine interannual changes in the seasonal cycle (in terms of timing and level). Parameter estimation for this state space model is also considered.

[MS-119]

Disclosure Risk and Variance Estimation

Risque de divulgation et estimation de variance

Wilson LU, *Acadia University*, Randy R. SITTER, *Simon Fraser University*

Protecting respondents from disclosure of their identity in publicly released survey data is of practical concern to many government agencies. Methods for doing so include suppression of cluster and stratum identifiers and altering or swapping record values between respondents. Unfortunately, stratum and cluster identifiers are usually needed for variance estimation using linearization and for replication methods as resampling is typically done on first-stage sampling units within strata. One might feel that releasing a set of replicate weights that also have stratum and cluster identifiers suppressed might circumvent this problem to some extent, especially using some random resampling such as the bootstrap. In this article, we first demonstrate that by viewing the replicate weights as observations in a high dimensional space one can easily use clustering algorithms to reconstruct the cluster identifiers irrespective of the resampling method even if the resampling weights are randomly altered. We then propose a fast algorithm for swapping cluster and strata identifiers of ultimate units before creating replicate weights without significantly impacting resulting variance estimates of characteristics of interest. The methods are illustrated by application to publicly released data from the National Health and Nutrition Examination Surveys, where such disclosure issues are extremely important.

tats d'analyse en séries chronologiques de ces données sont présentés. En particulier, un modèle d'espace des états cyclique est développé et appliqué à ces données. Ceci permet au cycle saisonnier des niveaux d'abondance d'être adaptatif, nous permettant ainsi d'examiner les changements interannuels dans le cycle saisonnier (en termes de temps et de niveau). L'estimation paramétrique pour ce modèle d'espace des états est aussi considérée.

La protection des répondants contre la divulgation de leur identité dans les enquêtes rendues publiques est d'intérêt pratique pour plusieurs agences gouvernementales. Des méthodes à cette fin comprennent la suppression des identificateurs de strate et de grappe, et l'altération ou la permutation des valeurs enregistrées entre les répondants. Malheureusement, les identificateurs de strate et de grappe sont nécessaires pour l'estimation de la variance par linéarisation et pour les méthodes de réplication, puisque le rééchantillonnage est souvent fait sur les unités échantillonnales de premier ordre à l'intérieur des strates. On pourrait croire que rendre public un ensemble de poids répliqués à identificateurs de strate et de grappes également supprimés pourrait contourner ce problème jusqu'à un certain degré, surtout en utilisant une méthode de rééchantillonnage aléatoire telle le bootstrap. Dans cet article, nous démontrons d'abord qu'en considérant les poids répliqués comme des observations dans un espace de dimension élevée, on peut facilement utiliser des algorithmes de classification afin de reconstruire les identificateurs de grappes indépendamment de la méthode de rééchantillonnage, même si les poids de rééchantillonnage sont altérés aléatoirement. Nous proposons alors un algorithme rapide pour permuter les identificateurs de grappe et de strate des unités finales avant de créer des poids répliqués sans causer d'impact significatif sur les estimés de la variance des caractéristiques d'intérêt. Les méthodes sont illustrées à l'aide d'une application à des données rendues publiques provenant des "National Health and Nutrition Examination Surveys", où les questions de divulgation des données sont extrêmement importantes.

Session 07F**Mardi 12 juin • Tuesday, June 12, 13:30–15:00****SN-2041****Longitudinal Data and Panel Counts****Données longitudinales et dénombrement de panels****[MS-120]****Zero-inflated Poisson Regression for Longitudinal Data****Régression de Poisson gonflée à zéro pour données longitudinales**M. Tariqul HASAN, *University of New Brunswick*, Gary SNEDDON, *Memorial University*

Zero-inflated Poisson (ZIP) regression models can be used to account for excess zeros in count data. We propose an extension of the ZIP model that is appropriate for longitudinal data. Our extension includes a nonstationary, observation-driven correlation structure. We will discuss estimation of the model parameters and the model's application to the analysis of health care utilization data.

Les modèles de régression de Poisson gonflés à zéro (ZIP) peuvent être utilisés pour tenir compte d'un excès de zéros dans des données de dénombrement. Nous proposons une extension du modèle ZIP appropriée pour des données longitudinales. Notre extension inclut une structure de corrélation non stationnaire basée sur les observations. Nous allons discuter de l'estimation des paramètres du modèle et de l'application du modèle à l'analyse de données d'utilisation de soins de santé.

[MS-121]**Association Study of Longitudinal data Measured at Distinctive Timepoints****Étude d'association de données longitudinales mesurées à des temps distincts**Xiaoqin XIONG & Joel DUBIN, *University of Waterloo*

For longitudinal data where the response and predictors are measured at different times, we need to do some preprocessing on the data (for example, smoothing) before applying any traditional longitudinal modeling. We propose a nonparametric smoothing method by defining equally spaced bins of time for response and predictors. We apply this binning method on a longitudinal hemodialysis study to find the association between a health event (e.g., infection) and a protein marker for inflammation. After smoothing, Zero-inflated Poisson (ZIP) models with random effects are fitted to the data. Properties of the model parameters estimated as a result of smoothing step will be discussed.

Pour des données longitudinales où la variable réponse et les prédicteurs sont mesurés à des temps différents, nous devons faire un pré-traitement des données (par exemple, un lissage) avant d'appliquer une quelconque méthode traditionnelle de modélisation longitudinale. Nous proposons une méthode de lissage non paramétrique en définissant des regroupements de temps équidistants pour la réponse et les prédicteurs. Nous appliquons cette méthode de regroupement à une étude longitudinale d'hémodialyse pour trouver l'association entre un événement (par exemple, une infection) et un marqueur protéinique pour l'inflammation. Après le lissage, des modèles de Poisson gonflés à zéro avec effets aléatoires sont ajustés aux données. Nous discuterons des propriétés des paramètres du modèle estimé suite à l'étape du lissage.

[MS-122]**Quasilielihood Inferences in Gamma AR (1) Models for Longitudinal Data****Inférences basées sur la quasi-vraisemblance avec des modèles Gamma de type AR (1) pour des données longitudinales**Manish MADAN & Brajendra SUTRADHAR, *Memorial University*

The time series analysis of gamma data (exponential being a special case) is common in many economic or engineering researches. As opposed to

L'analyse des séries chronologiques de données de loi gamma (la loi exponentielle constituant un cas particulier) est répandue dans plusieurs recherches en économie et en

the time series set up, in this paper, we have considered a longitudinal set up and studied an AR (1) type longitudinal gamma model consisting of a regression vector, a scale, and a longitudinal correlation parameter. Through a simulation study it is shown that the generalized quasilielihood (GQL) estimation approach provides consistent and highly efficient estimates of these parameters.

géné. Par opposition à l'implantation par série chronologique, nous considérons dans cette article une implantation longitudinale et avons étudié un modèle longitudinal de loi gamma de type AR(1) consistant en un vecteur de régression, une échelle, et un paramètre de corrélation longitudinale. Au moyen d'une étude de simulation, nous montrons que l'approche d'estimation par quasi-vraisemblance généralisée (QVG) donne des estimés convergents et hautement efficaces de ces paramètres.

[MS-123]

Intention to Treat Analysis and the MCMC Method for Imputing Missing Values in Longitudinal Studies: An Example Using Data from a Proactive Nursing Health Promotion Intervention Study in Frail Older Home Clients

Analyses d'intention de traitement et méthode MCMC pour imputer les valeurs manquantes dans une étude longitudinale : un exemple provenant d'une étude proactive sur des personnes âgées de santé fragile vivant en résidence

Adefowope ODUEYUNGBO, Lehana THABANE & Maureen MARKLE-REID, *McMaster University*

The validity of a statistical test is compromised when all observations from a study are not available for analysis. Missing values due to attrition/loss to follow-up constitute a major problem in the statistical analysis of data from longitudinal studies. We review some of the methods available for imputing data where missing values are assumed to be missing-at-random or missing-completely-at-random. An illustrative example of the MCMC method of imputation is given using data from a nursing intervention study. We also consider a multivariate analysis of variance model as opposed to univariate tests.

La validité d'un test statistique est compromise lorsque toutes les observation d'une étude ne sont pas disponible pour l'analyse. Les valeurs manquantes dues à l'attrition/perte au suivi constituent un problème majeur dans l'analyse statistique de données d'études longitudinales. Nous passons en revue quelques-unes des méthodes disponibles pour imputer les données lorsque les valeurs manquantes sont supposées manquantes-aléatoirement ou manquantes-complètement-aléatoirement. Un exemple illustrant la méthode MCMC d'imputation est donné en utilisant des données d'une étude d'intervention en sciences infirmières. Nous considérons également une analyse multivariée d'un modèle de variance par opposition aux tests univariés.

[MS-124]

Multi-sample Comparison of Point Processes with Panel Count Data

Comparaison d'échantillons multiples de processus ponctuels avec données longitudinales de dénombrement

Xingqiu ZHAO & Narayanaswamy BALAKRISHNAN, *McMaster University*

This paper considers the problem of multi-sample nonparametric comparison of point processes with panel count data, which arise naturally when recurrent events are considered. Such data frequently occur in medical follow-up studies, for example. For the problem considered, we construct a class of nonparametric test statistics based on the integrated weighted differences between the estimated mean functions of the point processes. The asymptotic distributions of the proposed statistics are derived and their finite-sample properties are exam-

Cet article considère le problème de la comparaison non paramétrique à échantillons multiples de processus ponctuels avec données longitudinales de dénombrement, qui survient naturellement dans un contexte d'événements récurrents. De telles données se retrouvent fréquemment dans des études de suivi médical, par exemple. Pour le problème considéré, nous construisons une classe de statistiques de test non paramétriques basées sur les différences pondérées intégrées entre les fonctions moyennes estimées des processus ponctuels. Les distributions asymptotiques des statistiques proposées sont obtenues et leurs propriétés, dans un contexte

ined through Monte Carlo simulations. A set of panel count data from a cancer study is analyzed and presented as an illustrative example.

d'échantillon fini, sont examinées par simulations de Monte-Carlo. Un ensemble de données de dénombrement longitudinales provenant d'une étude sur le cancer est analysé et présenté comme exemple illustratif.

Session 08A Mardi 12 juin • Tuesday, June 12, 15:30–17:00

AA-1043

Survey Design and Analysis

Planification et analyse d'enquêtes

[MS-125]

Sample Size Determination for the Aboriginal Post-censal Surveys

Détermination de taille d'échantillon pour les enquêtes postcensitaires autochtones

Éric LANGLET, *Statistics Canada/Statistique Canada*

The post-censal surveys draw their samples from answers to the Census long form, which is administered to approximately one in five households in Canada (2B component) except in remote areas where it is administered to all households (2D component). The Census long form sample constitutes the first phase sample. A second phase stratified random sample is then selected according to variables observed in the first phase. Sample sizes are derived by approximating this sampling plan and by using an optimal allocation between the 2B and 2D strata. Several adjustments are then performed on the resulting sample sizes.

Les enquêtes post-censitaires tirent leurs échantillons à partir des réponses fournies au questionnaire long du recensement, administré à environ un ménage sur cinq au Canada (partie 2B) sauf dans les régions éloignées où tous les ménages doivent le compléter (partie 2D). Ceci constitue l'échantillon de première phase. Un échantillon stratifié de deuxième phase est ensuite tiré selon les caractéristiques observées à la première phase. Afin d'estimer les tailles d'échantillon requises, une approximation du plan est utilisée pour ensuite effectuer une allocation optimale entre les strates 2B et 2D. De nombreux ajustements sont par la suite effectués sur les tailles obtenues.

[MS-126]

Ontario Survey on the Prevalence of Hypertension in Adults: Practical Aspects in Sample Allocation and Collection

Enquête sur la prévalence de l'hypertension chez les Ontariens adultes : aspects pratiques sur l'allocation et la collecte d'un échantillon.

Lori STRATYCHUK & Jean DUMAIS, *Statistics Canada/Statistique Canada*

In 2005, the Heart and Stroke Foundation of Ontario (HSFO) launched a study of the prevalence of hypertension in Ontario. The HSFO desired results publishable for four ethnic groups. Additional constraints included the HSFO's budget and desired final sample size, the mandatory inclusion of certain geographic areas, as well as the anticipated non-response rates. This paper concentrates on the sample allocation and data collection aspects of the survey.

En 2005, la Fondation des maladies du coeur d'Ontario (FMCO) a lancé une étude sur la prévalence de l'hypertension en Ontario. La FMCO voulait des résultats publiables pour quatre groupes ethniques. Des contraintes additionnelles incluaient le budget de la FMCO et la taille échantillonnale finale désirée, l'inclusion obligatoire de certaines régions géographiques, et les taux de non-réponse anticipés. Cet article se concentre sur les aspects de la répartition de l'échantillon et de la collecte de données.

[MS-127]**The Lavallée Hidiroglou Algorithm for Stratification in Business Surveys: A Summary and some new Developments****L'algorithme de Lavallée Hidiroglou pour la stratification dans les enquêtes entreprises : bilan et nouvelles avancées**

Louis-Paul RIVEST, *Université Laval*, Michel FERLAND, *Statistics Canada/Statistique Canada*

When implementing the Lavallée & Hidiroglou stratification algorithm one uses a measure of the size of each unit in the population. The algorithm gives the boundaries for the strata and a target sample size in each stratum needed to obtain a predetermined level of accuracy. Often the size measure for the calculations differs from the true size observed in the survey. This happens when some businesses go bankrupt in the time interval between the measure of their size and the survey. Non-response can jeopardize the accuracy of the survey estimates. This talk suggests generalizations of the Lavallée & Hidiroglou algorithm that incorporate a death probability and a non-response rate in the construction of the strata. These new developments will be illustrated with numerical examples.

Pour mettre en oeuvre l'algorithme de Lavallée-Hidiroglou on dispose d'une mesure de la taille pour chaque unité de la population ; l'algorithme donne des bornes pour les strates et les tailles d'échantillon à sélectionner dans chaque strate pour obtenir la précision désirée. Souvent la mesure de taille utilisée dans les calculs diffère de la vraie mesure observée dans l'enquête. C'est le cas lorsque des entreprises font faillite entre le moment de la mesure de leur taille et le début de l'enquête. De plus la non-réponse compromet parfois l'atteinte des objectifs de précision. Cette présentation suggérera des généralisations de l'algorithme de Lavallée et Hidiroglou qui incorporent une probabilité de mourir et une probabilité de non-réponse dans la construction des strates. Ces nouveaux développements seront illustrés à l'aide d'exemples numériques.

[MS-128]**Sample Design of the Canadian Health Measures Survey****Plan d'échantillonnage de l'Enquête canadienne sur les mesures de la santé**

Rebecca MORRISON, Sarah-Maude DION & Suzelle GIROUX, *Statistics Canada/Statistique Canada*

Statistics Canada, in partnership with Health Canada and the Public Health Agency of Canada, is planning to begin the collection of the Canadian Health Measures Survey (CHMS) in the spring of 2007. The CHMS aims to overcome limitations of self-reported data by directly measuring health indicators, and physical activity and fitness levels from a nationally representative sample of 5,000 Canadians, aged 6 to 79. A multi-stage sample design was developed to meet the objectives and logistics of the survey. The presentation will provide an overview of the CHMS with emphasis on the sampling plan.

Statistique Canada, en partenariat avec Santé Canada et l'Agence de santé publique du Canada, planifie débuter la collecte de l'Enquête canadienne sur les mesures de la santé (ECMS) au printemps 2007. L'ECMS a pour but de surmonter les limites des données autodéclarées en mesurant directement des indicateurs de santé, l'activité physique et le niveau de condition physique d'un échantillon national représentatif de 5 000 canadiens âgés de 6 à 79 ans. Un plan d'échantillonnage à plusieurs degrés a été développé pour répondre aux objectifs et à la logistique de l'enquête. La présentation donnera une vue d'ensemble de l'ECMS en mettant l'accent sur le plan d'échantillonnage.

[MS-129]**A Study of the Properties of the Rao-Wu Bootstrap Variance Estimator: What Happens when Assumptions do not Hold?****Une étude des propriétés de l'estimateur Rao-Wu de la variance bootstrap lorsque certaines hypothèses ne sont pas vérifiées**

Lenka MACH & Abdelnasser SAÏDI, *Statistics Canada/Statistique Canada*, Rob PETTAPEICE, *Statistics Canada/Statistique Canada & University of Waterloo*

The Rao-Wu bootstrap variance estimation method is frequently used at Statistics Canada. It is simple to implement, the same formula applies to all point estimators, including non-smooth statistics, and it enables the analysts to use the design-based approach. The method assumes Primary Sampling Units are selected with replacement or the first-stage sampling fractions are negligible. We examine the properties of the Rao-Wu bootstrap variance estimator, for a two-stage design and a variety of scenarios: different sampling methods, varying sampling fractions, several variables and point estimators. We also compare these properties with the ones of the analytical variance estimators.

La méthode d'estimation de la variance bootstrap Rao-Wu est fréquemment utilisée à Statistique Canada. La méthode est simple à mettre en place, la même formule s'applique à tous les estimateurs ponctuels incluant les statistiques non lisses et elle permet aux analystes d'utiliser l'approche selon le plan d'enquête. La méthode suppose que les unités primaires d'échantillonnage sont sélectionnées avec remise ou que les fractions d'échantillonnage au premier degré sont négligeables. Nous examinons les propriétés de l'estimateur de la variance bootstrap de Rao-Wu pour un plan à deux degrés et sous une diversité de scénarios : différentes méthodes d'échantillonnage, différentes fractions de sondage, plusieurs variables et estimateurs ponctuels. Nous comparons aussi ces propriétés à celles des estimateurs analytiques de la variance.

Session 08B**Mardi 12 juin • Tuesday, June 12, 15:30–16:30****SN-2041****Stochastic Partial Differential Equations and their Applications****Equations aux dérivées partielles stochastiques et leurs applications****[MS-130]****On Particle Representation of Markov Processes****Sur la représentation particulière des processus de Markov**

Michael KOURITZIN, *University of Alberta*

Kurtz and Ocone introduced the filtered martingale problem to transfer uniqueness from martingale problems to the corresponding filter and related spdes. Later, Donnelly and Kurtz transferred uniqueness of particle systems to measure-valued processes, using their lookdown process that views measure-valued processes as mappings of particle systems. Kurtz generalized these ideas by formulating the conditional martingale problem as well as associated Markov mapping techniques. In this talk, I will attempt to summarize these important contributions and explain how a main result behind the conditional martingale problem can be extended, simplified and applied to recover uniqueness for certain stochastic equations.

Kurtz et Ocone ont présenté le problème des martingales filtrées pour transférer l'unicité des problèmes de martingales au filtre correspondant et aux "spdes" qui leur sont reliés. Donnelly et Kurtz ont ensuite transféré l'unicité des systèmes à particules aux processus à valeurs dans un espace de mesures, en utilisant leur processus "lookdown" qui considère ces derniers comme une transformation des systèmes à particules. Kurtz a généralisé ces idées en formulant le problème des martingales conditionnelles ainsi que les techniques de transformation markoviennes associées. Dans cet exposé, je vais tenter de résumer ces contributions importantes et d'expliquer comment un important résultat sur le problème des martingales conditionnelles peut être généralisé, simplifié et appliqué afin d'obtenir l'unicité dans le cas de certaines équations stochastiques.

[MS-131]

Harnack Inequality for Stochastic Porous Medium Equation and Applications
Inégalité d'Harnack pour l'équation stochastique de médium poreux et applications

Feng-Yu WANG, *Beijing Normal University*

By using the coupling method and Girsanov transformations, the dimension-free Harnack inequality was established for the solution to a class of stochastic generalized porous media equations. As applications, the strong Feller property as well as estimates of the transition density are derived.

En utilisant la méthode de couplage et les transformations de Girsanov, nous établissons l'inégalité de Harnack indépendante de la dimension pour la solution d'une classe d'équations stochastiques généralisées pour médium poreux. Comme application, la propriété forte de Feller et les estimés de la densité de transition sont obtenus.

Session 08C**Mardi 12 juin • Tuesday, June 12, 15:30–16:30****IIC-2001**

Inaugural Invited Lecture of the National Institute for Complex Data Structures

Allocution inaugurale de l'invité d'honneur de l'Institut national sur les structures de données complexes

[MS-132]

Latent Structure in Space-time Disease Maps: Links to Environmental Hazard Events
Structures latentes dans les cartes spatio-temporelles des maladies : liens aux événements présentant un risque pour l'environnement

Andrew LAWSON & Bo CAI, *University of South Carolina*

Disease map variation over time can be characterized by models with spatial, temporal and spatio-temporal components. These models fall into a Bayesian Hierarchical framework and software is available for posterior sampling. It is relevant to consider an extension of this where it is believed that a common underlying factor or factors contributes to the ST disease variation. In this talk I will discuss this model and its implementation, including issues such as component labeling, identifiability and the effect of having fixed or stochastic number of components. An example of linkage to ST environmental pollution profiles will be considered.

Les variations dans le temps des cartes des maladies peuvent être caractérisées par des modèles à composantes spatiale, temporelle, et spatio-temporelle. Ces modèles s'inscrivent dans un cadre hiérarchique bayésien et des logiciels sont disponibles pour l'échantillonnage a posteriori. Il est pertinent de considérer une extension de ceci au cas où nous croyons qu'un ou des facteurs sous-jacents contribuent à la variation spatio-temporelle de la maladie. Dans cet exposé, je vais discuter de ce modèle et de sa mise en oeuvre, incluant des questions telles que l'étiquetage des composantes, l'identifiabilité et l'effet d'avoir un nombre fixe ou stochastique de composantes. Un exemple de lien aux profils ST de pollution environnementale sera considéré.

Session 08D**Mardi 12 juin • Tuesday, June 12, 15:30–16:30****SN-2109**

Confidence Intervals and Regions

Intervalles et régions de confiance

[MS-133]**Frequentist Coverage of Bayes Credible Regions****Taux de couverture fréquentiste des régions de confiance bayésienne**William STRAWDERMAN, *Rutgers University*, Éric MARCHAND, *Université de Sherbrooke*

For estimating a positive normal mean, Zhang and Woodroffe (2003) as well as Roe and Woodroffe (2000) investigate $100(1 - \alpha)\%$ HPD credible sets associated with priors obtained as the truncation of noninformative priors onto the restricted parameter space. Namely, they establish the attractive lower bound of $(1 - \alpha)/(1 + \alpha)$ for the frequentist coverage probability of these procedures. In this work, we establish that the lower bound of $(1 - \alpha)/(1 + \alpha)$ is applicable for a substantially more general setting with underlying distributional symmetry, and present various illustrations and related properties. Investigations of non-symmetric models are carried out and similar results are obtained.

Pour estimer une moyenne normale positive, Zhang et Woodroffe (2003), ainsi que Roe et Woodroffe (2000) étudient $100(1 - \alpha)\%$ des ensembles de plus haute densité a posteriori associés aux a priori obtenus comme la troncation d'a priori non informatifs sur l'espace paramétrique restreint. En particulier, ils trouvent la borne inférieure de $(1 - \alpha)/(1 + \alpha)$ pour la probabilité de couverture fréquentiste de ces procédures. Dans ce travail, nous montrons que la borne inférieure $(1 - \alpha)/(1 + \alpha)$ est applicable dans un cadre substantiellement plus général avec symétrie des distributions sous-jacentes, et nous présentons diverses illustrations et propriétés liées. En étudiant des modèles non-symétriques, nous obtenons des résultats similaires.

[MS-134]**Estimating the Confidence of a Confidence Region****Estimation du degré de confiance d'une région de confiance**Dominique FOURDRINIER, *Université de Rouen*

For a random vector X with spherically symmetric distribution around an unknown location parameter, we consider a confidence region $C(X)$ of this parameter, with fixed confidence coefficient. It is known that, for observed values x from X , this confidence coefficient is inadequate to assess the true confidence in the corresponding region $C(x)$ and that it is needed to take into account the observation x . Through formal proofs, we yield estimates of this confidence better than the confidence coefficient, with no recourse to simulations.

Pour un vecteur aléatoire X de loi à symétrie sphérique autour d'un paramètre de position inconnu, nous considérons une région de confiance $C(X)$ de ce paramètre, de niveau de confiance fixé. On sait que, pour des réalisations x de X , ce niveau est insuffisant pour apprécier le vrai degré de confiance en la région $C(x)$ à disposition et qu'il est nécessaire de prendre en compte l'observation x . Nous mettons en évidence, de manière formelle, des estimations de ce degré de confiance meilleures que le niveau de confiance fixé, sans avoir recours à des simulations.

Session 08E**Mardi 12 juin • Tuesday, June 12, 15:30–16:30****AA-1046**

Graphical Methods and Data Visualization

Méthodes graphiques et visualisation de données

[MS-135]**Bivariate Quantile as Tukey Depth Contours: A Connection to Directional Quantile Regression****Courbes de niveaux pour la profondeur de Tukey à l'aide de quantiles bivariés : un lien avec la régression quantile directionnelle**Linglong KONG & Ivan MIZERA, *University of Alberta*

Halfspace (Tukey) depth is a popular concept in multivariate data analysis and data mining, with applications as diverse as forest fires synthetic plots, bivariate growth charts and so on. We will present a characterization of halfspace depth through directional quantile regression and discuss properties of depth contours that are of interest for their data-analytic interpretation. If time permits, some more properties of halfspace depth contours will be considered as well.

La profondeur de demi-espace (de Tukey) est un concept populaire en analyse de données multivariées et en exploration de données, avec des applications aussi diverses que les graphes synthétiques de feux de forêts, les tableaux de croissance bivariés, et ainsi de suite. Nous allons présenter une caractérisation de la profondeur de demi-espace au moyen de la régression quantile directionnelle et discuter des propriétés des courbes de niveaux pour la profondeur et de leur interprétation en analyse de données. Si le temps le permet, quelques propriétés additionnelles des courbes de niveaux pour la profondeur de demi-espace seront également considérées.

[MS-136]**Application of Bayesian Graphical Model to Study the Genetic Susceptibility to Breast Cancer****Application des modèles graphiques bayésiens pour l'étude de la susceptibilité génétique au cancer du sein**Sohee KANG, *University of Toronto*, Laurent BRIOLLAIS, *Samuel Lunenfeld Research Institute, Mount Sinai Hospital*

There have been an increased research interest and activity for graphical modeling in various statistical fields. In genetic studies, the advances in modern DNA technology combined with the rapid development of computing facilities have made genetic analyses using graphical modeling feasible. For example, graphical models have been successfully applied in genetic association studies to identify patterns of allelic association between markers (Thomas and Camp, 2004). Verzilli *et al* (2006) used discrete graphical models to study genotype-phenotype associations that are scalable to genome wide association studies (GWAS).

We employed the Bayesian graphical modeling framework to seek for genetic association between Single Nucleotide Polymorphisms (SNPs) from the DNA repair pathway and the breast cancer phenotype. The decomposable graphical models for discrete data with Bayesian and non-Bayesian perspectives are studied. The impact of different set-

Il y a eu un intérêt de recherche croissant et une activité grandissante pour la modélisation graphique dans des domaines variés de la statistique. Dans les études génétiques, les avancées sur les technologies modernes de l'ADN associées au développement rapide des moyens de calcul ont rendu possible les études génétiques utilisant la modélisation graphique. Par exemple, les modèles graphiques ont été appliqués pour identifier des structures d'association allélique entre marqueurs (Thomas et Camp, 2004). Verzilli *et al.* (2006) ont utilisé les modèles graphiques discrets pour étudier les associations de type génotype-phénotype qui peuvent être évaluées dans les études d'association sur l'ensemble du génome (GWAS). Nous avons utilisé le cadre de la modélisation graphique bayésienne pour rechercher l'association génétique entre les polymorphismes d'une seule paire de base (SNPs) dans le réseau de réparation de l'ADN et le phénotype du cancer du sein. Les modèles graphiques décomposables pour données discrètes sont étudiés dans une perspective bayésienne et non bayésienne. L'impact de différentes distributions a priori, tels que

tings of priors, such as uninformative and informative, and the effect of hyper-parameters on priors are compared along with non-Bayesian graphical modeling method. We also present different estimation procedures based on either MCMC (Markov chain Monte Carlo) or RJMCMC (Reversible Jump Markov Chain Monte Carlo) (Green 1995). The final graphs obtained from both methods are compared with a discussion about their relative performance.

non informatif ou informatif, et les effets des hyper paramètres reliés sont comparés, de même que des méthodes de modélisation graphique non-bayésiennes. Nous présentons également différentes procédures d'estimation basées sur MCMC et RJMCMC (Green 1995). Les graphes finaux obtenus à partir des deux méthodes sont comparés selon une discussion à propos de leur performance relative.

[MS-137]

William Playfair and Bias in Statistical Graphics

William Playfair et le biais dans les graphiques statistiques

Matthew DUBINS, *York University*

Today statistical graphics are a widely accepted means of conveying massive amounts of quantitative information in a few moments. This climate stands in contrast to what was considered acceptable around 200 years ago. This presentation will look at the graphical products of one of Europe's earliest proponents of visualizing quantitative information, William Playfair. Each graphical example shall be evaluated according to how well it summarizes and conveys the data, and how much bias has been introduced by its author. Through analyzing each example, principles relevant to the creation of good statistical graphics today shall be made clear.

Aujourd'hui, les graphiques statistiques sont un moyen largement accepté de transmettre des quantités massives d'informations en quelques instants. Ce climat contraste beaucoup avec ce qui était considéré acceptable il y a environ 200 ans. Cette présentation portera sur les produits graphiques d'un des premiers européens à proposer de visualiser l'information quantitative, William Playfair. Chaque exemple graphique sera évalué d'après sa capacité à bien résumer et communiquer l'information, et d'après le biais introduit par son auteur. L'analyse de ces exemples nous révélera des principes relatifs à la création de bons graphiques statistiques utilisables de nos jours.

[MS-138]

Visualizing Inequity

La visualisation d'iniquité

Mikēlis BICKIS & Eric NEUFELD, *University of Saskatchewan*

Investigations of unfair discrimination in employee salaries typically involve a diverse population with many confounding variables. Assessments of inequity have usually used some kind of regression analysis to adjust the variable of interest for the confounders. Sensible interpretation of such analyses requires understanding the causal links between the confounders, which can be modeled by an acyclic directed graph. We present a scheme for representing such relationships to aid in grasping concepts of direct, systemic, and spurious discrimination, and in visualizing the effects of various interventions.

Des enquêtes sur la disparité injuste des salaires impliquent habituellement une population diverse et beaucoup de variables confondues. L'évaluation de l'iniquité utilise habituellement une analyse de régression quelconque pour ajuster la variable d'intérêt aux variables confondues. Une interprétation raisonnable d'une telle analyse dépend d'une compréhension des liens causaux entre les variables, que l'on peut modéliser à l'aide d'un graphe acyclique orienté. Nous présentons un système de représentation pour de telles relations pour aider à comprendre les concepts de la discrimination directe, systémique, et factice, et pour faciliter la visualisation des effets de diverses interventions.

Session 08F**Mardi 12 juin • Tuesday, June 12, 15:30–17:00****SN-2105**

Business and Industrial Statistics
Statistique industrielle et de gestion

[MS-139]

Homogeneity testing in multivariate regional frequency analysis
Test d'homogénéité en analyse fréquentielle régionale multivariée

Fateh CHEBANA & Taha OUARDA, *Institut National de la Recherche Scientifique*

Droughts, floods and storms are complex hydrological events that are described with multivariate characteristics such as peak, duration and total volume. It is important to jointly consider all these characteristics when carrying out a multivariate regional frequency analysis. The object of this communication is to extend the univariate homogeneity test to such context. Multivariate L-moments are used to define the statistics and general copula models to describe the dependence. Based on several types of simulated regions, results show the power of the proposed multivariate homogeneity test to detect heterogeneity on the whole structure of the model and on the marginal distributions.

Les sécheresses, inondations et tempêtes sont des événements hydrologiques complexes qui sont décrits avec des caractéristiques multivariées telles que la pointe, la durée et le volume total. Il est important de considérer conjointement toutes ces caractéristiques en effectuant une analyse fréquentielle régionale multivariée. L'objet de cette communication est d'étendre le test d'homogénéité univarié à un tel contexte. Les L-moments multivariés sont employés pour définir les statistiques et des modèles généraux de copule pour décrire la dépendance. Basés sur plusieurs types de régions simulées, les résultats montrent la puissance du test proposé à détecter l'hétérogénéité sur la structure entière du modèle et sur les distributions marginales.

[MS-140]

Adjusted Score Test for the Homogeneity of Gamma Populations
Test de cotes ajusté pour l'homogénéité de populations Gamma

K Ranee THIAGARAJAH, *Illinois State University*

Over the last few decades, the gamma distribution has arisen as one of the important distributions to model life-testing situations. A wide variety of lifetime data fits adequately to gamma distribution because of the flexibility in choice of the shape and scale parameters. In this paper, we consider the problem of testing homogeneity of gamma populations using methods of combining two independent statistics, namely two likelihood ratio statistics, two score statistics, and two adjusted score statistics. The performance of these procedures is examined by Monte Carlo simulations. In general, the method of combining two independent adjusted score statistics performs well relative to other two methods.

Au cours des dernières décennies, la distribution gamma s'est révélée l'une des distributions importantes dans la modélisation des durées de vie. La distribution gamma peut être ajustée adéquatement à une grande variété de données de temps de survie en raison de la flexibilité dans le choix de ses paramètres de forme et d'échelle. Dans cet article, nous considérons le problème de tester l'homogénéité de populations gamma en utilisant des méthodes combinant deux statistiques indépendantes, soit deux statistiques de rapport de vraisemblances, deux statistiques de score, et deux statistiques de score ajusté. La performance de ces procédures est examinée grâce à des simulations de Monte-Carlo. En général, la méthode combinant deux statistiques de score ajusté indépendantes performe bien relativement aux deux autres méthodes.

[MS-141]**Electricity Time of Use Rates Impact Analysis in BC****Analyse de l'impact des taux de temps d'utilisation de l'électricité en Colombie-Britannique**Min YU, *BC Hydro*

In 2006, BC Hydro launched an electricity conservation time of use (TOU) project, called the conservation research initiative (CRI), in three areas to examine how TOU rates can help individual British Columbians to reduce peak demand and/or shift highest consumption period to off-peak hours. One control rate and five trial rates are applied to 2000 households who participate in the CRI project and hourly consumption information are recorded for individual participants.

GLM models are fitted to compare the impacts of different rate in each area. Also Logistic regression model helps to determine whether TOU rates shift peak demand to off-peak hours.

En 2006, Hydro-BC a lancé un projet de temps d'utilisation (TDU) de conservation de l'électricité, nommé initiative de recherche en conservation (IRC), dans trois régions afin d'examiner comment les taux de TDU peuvent aider les Britanno-Colombiens à réduire la demande de pointe et/ou transférer la période de plus haute consommation hors des heures de pointe. Un taux de contrôle et cinq taux expérimentaux sont appliqués à 2000 ménages qui participent au projet IRC et l'information de consommation horaire est enregistrée pour chaque participant.

Des modèles GLM sont ajustés afin de comparer les impacts de différents taux dans chaque région. Le modèle de régression logistique aide aussi à déterminer si les taux de TDU ont permis de transférer les demandes à des heures hors de la période de pointe.

[MS-142]**The Impact Of Product Complexity on Pricing Research****L'impact de la complexité des produits sur la recherche en fixation des prix**Amanda LAFONTAINE & Klass WINSTON, *Ipsos-Reid*

Within market research, a common technique for determining optimal price, as well as the level of price sensitivity, is the Van Westendorp technique. For a major retailer, this technique was applied to a number of products ranging from simple items to complicated leading edge ideas. In doing so, it was discovered that many of the results were inconsistent, in particular for more complicated or higher priced products. In this talk, I will describe the techniques employed to determine possible sources for these inconsistencies. Furthermore an attempt will be made to better understand the relationship between price and perceived value.

À l'intérieur d'une étude de marché, une technique commune pour déterminer le prix optimal, ainsi que le niveau de sensibilité aux prix, est la technique de Van Westendorp. Pour un détaillant majeur, cette technique fut appliquée à plusieurs produits allant d'items simples à des idées complexes d'avant-garde. Ce faisant, plusieurs des résultats se sont avérés non cohérents, et ce, en particulier pour les produits plus complexes ou à prix plus élevé. Dans cet exposé, je décris les techniques employées pour déterminer les sources possibles de ces incohérences. De plus, une tentative afin de mieux comprendre la relation entre le prix et la valeur perçue est faite.

[MS-143]**Bimodal Hydrological Data Modeling****Modélisation de données hydrologiques bimodales**Félix LABRECQUE-SYNNOTT & Jean-François ANGERS, *Université de Montréal*, Anne-Catherine FAVRE & Michel SLIVITZKY, *Institut national de la recherche scientifique*

We are interested in multimodal data modeling using mixture models and hidden Markov chains. We also consider these models in the presence of an

Nous nous intéressons à la modélisation de données bimodales à l'aide de modèles de mélange et de chaînes de Markov cachées. Nous considérons également des modèles de

explanatory variable, which we suppose to be correlated to the implicit state of the mixture model or Markov chain. We present and analyse a precipitation dataset. In particular, we are interested in detecting possible trends or abrupt changes in the data. Sequential and moving-window analysis are carried out to this end.

régression linéaire généralisés en présence d'une variable explicative, que nous supposons corrélée avec la variable implicite d'état du modèle de mélange ou de la chaîne de Markov. Nous présentons et analysons un jeu de données de précipitations. En particulier, nous souhaitons déceler de possibles tendances et perturbations dans les données. Des analyses séquentielles et à fenêtres mobiles sont effectuées à cette fin.

Session 09A **Mercredi 13 juin • Wednesday, June 13, 08:30–09:30** **IIC-2001**

Pierre Robillard Award Address

Allocution du récipiendaire du prix Pierre-Robillard

Session 09B **Mercredi 13 juin • Wednesday, June 13, 08:30–10:00** **AA-1043**

Genetic Epidemiology

Épidémiologie génétique

[MS-144]

Efficient Information Extraction in Genome-wide Association Studies

Extraction efficace de l'information dans les analyses portant sur tout le génome

Dan NICOLAE, *University of Chicago*

Genome-wide association studies aim to detect DNA variation modifying disease risk in humans. The challenges in analyzing the data generated by these studies come from the complexity of the genetic data and of the disease-causing mechanisms. I will discuss methods for measuring and extracting information available on all known genetic variation using the genotyped SNPs. I will also describe how the genetic association test, TUNA, that results from efficiently interrogating untyped variation, can be extended to quantitative traits and gene-gene interaction analyses.

Les études d'association sur tout le génome visent à détecter les variations d'ADN modifiant le risque de maladie chez l'humain. Les défis dans l'analyse des données générées par ces études proviennent de la complexité des données génétiques et des mécanismes causant les maladies. Je vais discuter de méthodes pour mesurer et extraire l'information disponible sur toute la variation génétique connue en utilisant les SNPs genotypés. Je vais aussi décrire comment le test d'association génétique TUNA, qui résulte d'un questionnement efficace de la variation non typée, peut être étendu à des analyses de traits quantitatifs et d'interactions entre deux gènes.

[MS-145]

The Multiplicity Problem in Genome-wide Association (GWA) Studies

Le problème de multiplicité dans les études sur tout le génome

Lei SUN, *University of Toronto*

Due to the recent advance in high-throughput genotyping technology, many current disease mapping studies are moving towards the Genome-Wide Association (GWA) study design. One of the main methodological challenges in the analyses of GWA studies is the inherent problem of high-dimensional hypothesis testing, because hundreds of thousands

Grâce aux avancées récentes en technologie de génotypage à haut débit, plusieurs études de cartographie des maladies se dirigent vers le plan d'analyse d'association sur tout le génome (ATG). Un des principaux défis méthodologiques des analyses ATG est le problème intrinsèque aux tests d'hypothèses à hautes dimensions, car des centaines de milliers de marqueurs génétiques sont étudiés simultanément.

or more genetic markers are investigated simultaneously. I will discuss some of the False Discovery Rate (FDR) -based approaches, focusing on three characteristics of GWA studies: data are massive and correlated, signals are sparse, and strength of the signals is weak.

Je vais expliquer certaines approches basées sur le taux de fausse découverte, en me concentrant sur trois caractéristiques des études ATG : les données sont massives et corrélées, les signaux sont rares et de faible intensité.

[MS-146]

A Composite Likelihood Approach for Identification of Genetic Variants Potentially Causally Associated with a Disease

Une approche basée sur la vraisemblance composite pour identifier les variantes génétiques présentant potentiellement un lien causal avec une maladie

Joanna BIERNACKA & Heather CORDELL, *Newcastle University*

“Direct” genetic association arises when a “causal” variant alters disease susceptibility. In families, genetic association leads to linkage (co-segregation of genes) in regions surrounding associated genes. Direct association with a sole causal variant in the region should explain the observed linkage. Biernacka and Cordell (2006) used a likelihood approach to test the null hypothesis that a given candidate is the only causal variant in a region, or is “fully” associated with the sole causal variant, and can therefore explain the linkage. Here we present a composite likelihood method that models information from several candidate variants and can substantially increase power.

Une association génétique “directe” survient lorsqu’une variante “causale” altère la susceptibilité à une maladie. Dans les familles, l’association génétique mène à une liaison (co-ségrégation de gènes) dans les régions entourant les gènes associés. L’association directe avec une seule variante causale devrait expliquer la liaison observée. Biernacka et Cordell (2006) ont utilisé une approche basée sur la vraisemblance pour tester l’hypothèse nulle qu’un candidat donné est la seule variante causale dans une région, ou est “complètement” associé à la seule variante causale, et peut donc expliquer la liaison. Nous présentons ici une méthode de vraisemblance composite qui modélise l’information de plusieurs variantes candidates et peut substantiellement augmenter la puissance du test.

Session 09C Mercredi 13 juin • Wednesday, June 13, 08:30–09:30 AA-1046

Data with Extra Zeros Applied in Ecology

Données avec un surplus de zéros et applications en écologie

[MS-147]

Generalized Additive Models for Zero Inflated Counted Data

Modèles additifs généralisés pour données de dénombrement gonflés à zéro

Alan WELSH, *Australian National University*, Simon BARRY, *CSIRO Mathematics and Information Systems*

We consider a flexible method for modelling zero inflated count data. We first model the association between presence and the available covariates and second, model the relationship between abundance and the covariates, conditional on presence. We compare this approach to that of Lambert (1992, *Technometrics*). We then incorporate the use of Generalized Additive Models (GAM) in the modelling steps. We explore the link and variance functions needed for the use of standard GAM software

Nous considérons une méthode flexible pour modéliser des données de dénombrement gonflés à zéro. Nous modélisons d’abord l’association entre la présence et les covariables disponibles et, ensuite, modélisons la relation entre l’abondance et les covariables, conditionnellement à la présence. Nous comparons cette approche à celle de Lambert (1992, *Technometrics*). Nous incorporons alors l’utilisation de modèles additifs généralisés (GAM) dans les étapes de modélisation. Nous explorons les fonctions de lien et de variance requises pour l’utilisation de programmes standard GAM avec des

with zero inflated data. We then fit the model to data on stem counts of *Eucalyptus mannifera* in a region of South East Australia.

données gonflées à zéro. Nous ajustons alors le modèle à des données de dénombrement de tiges d'*Eucalyptus mannifera* dans une région de l'Australie du sud est.

[MS-148]

Zero-inflated Clustered Count Data

Données de comptage en grappes avec sur-représentation de zéros

Eva CANTONI, *University of Geneva*, Joanna MILLS-FLEMMING, *Dalhousie University*, Alan WELSH, *Australian National University*

We will consider situations where count data not only present an excess of zero, but also where they come in clusters. We will define a conditional (hurdle or two-part) model where we additionally introduce random effects on both parts of the model to account for the cluster structure. We will show how we estimate the model by maximum likelihood, via a Laplace approximation. An application on the estimation of sharks abundance will illustrate our proposed methodology. If time permits, we will discuss few issues related to prediction with mixed models. This is joint work with J. Mills-Flemming and A. Welsh.

Pour cet exposé on considèrera des données de comptage qui non seulement présentent une sur-représentation de zéros, mais aussi qui sont structurées en grappes. On définira un modèle conditionnel de type “haie” ou “deux parties” dans lequel on introduira des effets aléatoires pour tenir compte de la structure en grappes. Nous écrirons la vraisemblance du modèle, que nous maximiserons via une approximation de Laplace. La méthodologie sera illustrée sur un jeu de données visant à décrire l'abondance de requins. Si le temps le permet, on discutera de quelques thèmes liés à la prévision dans les modèles à effets mixtes. Ceci est un travail en collaboration avec J. Mills-Flemming et A. Welsh.

Session 09D Mercredi 13 juin • Wednesday, June 13, 08:30–10:30 SN-2109

Latest Developments in Small Area Estimation, Theory and Practice

Derniers développements pour l'estimation de petites régions : théorie et aspects pratiques

[MS-149]

Estimation of Mean Squared Prediction Error of an Empirical Best Predictor

Estimation de l'erreur quadratique pour une prévision empirique optimale

Parthasarathi LAHIRI, *University of Maryland, College Park*, Snigdhanu CHATTERJEE, *University of Minnesota*

In this paper, we retain the main attraction of the parametric bootstrap, i.e. its simplicity, and provide a unified parametric bootstrap MSPE estimator that provides the second-order unbiased property for a wide variety of small area models. Unlike other existing parametric bootstrap methods, our method achieves the second-order unbiasedness property via appropriate generation of the resamples and not through a complex bias correction procedure. Through a Monte Carlo simulation study, we demonstrate the superiority of our method over the existing methods.

Dans cet article, nous retenons le principal attrait du bootstrap paramétrique, c.-à-d. sa simplicité, et présentons un estimateur de l'EQPM par bootstrap paramétrique unifié sans biais de second ordre pour une grande variété de modèles de petits domaines. Contrairement aux autres méthodes existantes par bootstrap paramétrique, la nôtre est sans biais de second ordre grâce à une génération appropriée des rééchantillons et non au moyen d'une procédure complexe de correction de biais. Nous démontrons la supériorité de notre méthode sur les autres méthodes existantes au moyen d'une étude de simulation de Monte-Carlo.

[MS-150]**Small Area Prediction using Convex Combination of Parametric and Non-parametric Models
Prévision pour petits domaines en utilisant une combinaison convexe de modèles paramétriques et non paramétriques**Narasimha PRASAD & Subhash LELE, *University of Alberta*

Survey practitioners, to obtain small area estimates, have used parametric models to “borrow-strength” from similar areas. Such methods heavily depend on the underlying model. To remedy this problem, non-parametric techniques have been adopted. An approach based on non-parametric regression is robust against model specification and performs well when the number of small areas is very large. However, for moderate to large number of small areas a parametric model-based approach performs better even if the assumed parametric model is not the correct one. We propose a compromise approach that considers a convex combination of a parametric and a non-parametric model.

Les planificateurs de sondages, afin d’obtenir des estimés pour petits domaines, ont utilisé des modèles paramétriques afin de capitaliser sur l’information d’aires similaires. De telles méthodes dépendent fortement du modèle sous-jacent. Pour pallier à ce problème, des techniques non-paramétriques ont été adoptées. Une approche basée sur la régression non paramétrique est robuste à la spécification du modèle et performe bien même lorsque le nombre de petits domaines est très grand. Cependant, pour un nombre grand ou modéré de petits domaines, une approche basée sur un modèle paramétrique est plus performante même si le modèle paramétrique supposé est incorrect. Nous proposons un compromis considérant une combinaison convexe d’un modèle paramétrique et d’un non paramétrique.

[MS-151]**Evaluation of Small Domain Estimators for the Survey of Employment, Payrolls and Hours
Évaluation d’estimateurs pour petits domaines pour l’Enquête sur l’emploi, la rémunération et les heures de travail**Susana RUBIN BLEUER, Serge GODBOUT & Yves MORIN, *Statistics Canada/Statistique Canada*

The Survey of Employment, Payrolls and Hours provides monthly estimates of payrolls, employment, paid hours and earnings. The proposed new design uses survey and administrative data to produce generalized regression (GREG) estimators for average weekly earnings, which are approximately unbiased and with controlled variability for pre-determined strata. As in most surveys, there are many domains of interest with small sample size for which the GREG estimators are inefficient. In this paper, we extend the pseudo-eblop method developed by You and Rao to the case of unequal error variances, and we study the relative performance of several cross-sectional small domain estimators.

L’Enquête sur l’emploi, la rémunération et les heures de travail fournit des estimations mensuelles sur la rémunération, l’emploi, les heures payées et le revenu. Le plan nouvellement proposé utilise des données administratives et d’enquêtes afin de produire des estimateurs de régression généralisée (GREG) pour le revenu hebdomadaire moyen, qui sont approximativement sans biais et à variabilité contrôlée pour des strates pré-déterminées. Comme dans la plupart des enquêtes, il y a plusieurs domaines d’intérêts à faible taille d’échantillon pour lesquels les estimateurs GREG sont peu efficaces. Dans cet article, nous généralisons la méthode pseudo-eblop développée par You et Rao au cas d’erreurs à variances inégales, et nous étudions la performance relative de plusieurs estimateurs pour des données transversales pour de petits domaines.

Session 09E Mercredi 13 juin • Wednesday, June 13, 08:30–10:00 SN-2105**Nonparametric Curve Estimation
Estimation non paramétrique de courbes****[MS-152]****Differentiated Logdensity Approximants
Approximants différenciés de la logdensité**Hyung-Tae (Kiefer) HA, *Kyungwon University*

An alternative density approximation methodology according to which the first derivative of the logarithm of the density approximant is expressed as a ratio of polynomials, is being proposed. The polynomial coefficients of the numerator and denominator of the rational function are determined by solving the system of linear equations which results from matching the moments of the approximant to those of the distribution being approximated. The results will be compared in specific applications to those obtained from the saddlepoint technique and methodologies based on orthogonal polynomials and Edgeworth expansions.

Une méthodologie alternative pour l'approximation de la densité selon laquelle la première dérivée du logarithme de l'approximant de la densité est exprimée comme un rapport de polynômes est proposée. Les coefficients polynomiaux du numérateur et du dénominateur de la fonction rationnelle sont déterminés en résolvant le système d'équations linéaires obtenu en posant les moments de l'approximant égaux aux moments de la distribution approximée. Les résultats seront comparés, pour des applications spécifiques, à ceux obtenus avec la technique de point de selle et les méthodologies basées sur les polynômes orthogonaux et les développements en séries d'Edgeworth.

[MS-153]**Nonparametric Estimation of Linear Functionals of a Bivariate Distribution under Univariate Censoring
Estimation non paramétrique de fonctions linéaires d'une distribution bivariée sous censure univariée**Xuewen LU & Murray D. BURKE, *University of Calgary*

In the presence of univariate censoring, a class of nonparametric estimators is proposed for linear functionals of a bivariate distribution of paired failure times. The estimators are shown to be asymptotically normal. An adjusted empirical log-likelihood is developed and proved to follow a chi-square distribution asymptotically. Two types of confidence intervals, based on the normal approximation method and the empirical likelihood method respectively, are constructed to make inference about the linear functionals. Their performance is evaluated in a real example and several simulation studies. The results indicate the empirical likelihood method works well compared to the normal approximation method.

En présence de censure univariée, une classe d'estimateurs non paramétriques est proposée pour des fonctionnelles linéaires d'une distribution bivariée de temps de survie pairés. Nous démontrons que les estimateurs sont asymptotiquement normaux. Nous développons une log-vraisemblance empirique ajustée et prouvons qu'elle suit une distribution asymptotiquement chi-deux. Deux types d'intervalles de confiance, basés respectivement sur l'approximation normale et sur la vraisemblance empirique, sont construits afin de faire de l'inférence au sujet des fonctionnelles. Leur performance est évaluée à l'aide d'un exemple réel et de plusieurs études de simulation. Les résultats indiquent que la méthode par vraisemblance empirique fonctionne bien comparativement à la méthode par approximation normale.

[MS-154]**Polynomial Spline Estimation of Partially Linear Single-index Proportional Hazards Regression Models
Estimation par splines polynomiales de modèles de régression des risques proportionnels partiellement linéaires**Jie SUN, *University of Calgary*

The Cox proportional hazards model usually assumes linearity of the log hazard function on covariate variables. In this paper, I will consider an alternative way to model both possible linearity and nonlinearity of the covariate effects to the log hazard in the proportional hazards model with a set of covariates. I propose a partially linear single-index proportional hazards regression model and apply a polynomial spline smoothing method to model the structured nonparametric single-index component, in order to reduce the dimensionality of the covariates and to obtain efficient estimates of the covariates' effects. Monte Carlo simulation studies are presented.

Le modèle des risques proportionnels de Cox suppose la linéarité du log de la fonction de risque sur les covariables. Dans cet article, je considère une approche alternative permettant de modéliser à la fois la linéarité et une éventuelle non linéarité des effets des covariables sur le log de la fonction de risque dans le modèle des risques proportionnels. Je propose un modèle de régression partiellement linéaire à indice simple des risques proportionnels et applique une méthode de lissage par splines polynomiales à la partie structurée non paramétrique de façon à réduire la dimension des covariables et à obtenir des estimateurs efficaces de leurs effets. Des études de simulation de Monte-Carlo sont présentées.

[MS-155]**Estimation of Mean Functions for Interval-censored Recurrent Event Data with Terminal Events
Estimation de la fonction moyenne pour des événements récurrents censurés par intervalles en présence d'événements terminaux**David TOLUSSO & Richard COOK, *University of Waterloo*

When subjects may experience recurrent events, interest often lies in the mean function reporting the cumulative expected number of events over time. Exact times of events are unknown when subjects are only seen at periodic assessment times. Methods of estimating the mean function are considered using a Poisson assumption as well as an estimator based on state occupancy probabilities when states are defined based on the cumulative number of events. Local likelihood methods feature prominently in both approaches. Terminal events will also be considered. Data from a recent study of patients with metastatic cancer will be used for illustration.

Lorsque les sujets peuvent subir des événements récurrents, la fonction moyenne rapportant le nombre espéré total d'événements à travers le temps est souvent d'intérêt. Les temps exacts des événements sont inconnus lorsque les sujets sont uniquement vus à des temps d'évaluation périodiques. Nous considérons des méthodes d'estimation de la fonction moyenne utilisant une hypothèse de Poisson, ainsi qu'un estimateur basé sur les probabilités d'états lorsque la définition des états est basée sur le nombre cumulatif d'événements. Les méthodes de vraisemblance locales occupent une place centrale dans les deux approches. Nous considérons également les événements terminaux. Des données d'une étude récente de patients atteints de cancer métastatique seront utilisées à des fins illustratives.

[MS-156]**Mixed Nonhomogeneous Poisson Process Spline Models for the Analysis of Recurrent Event Panel Data
Modèle spline de processus de Poisson non-homogène mixte pour l'analyse de données longitudinales d'événements récurrents**Jason NIELSEN & Charmaine DEAN, *Simon Fraser University*

A flexible semi-parametric model for analyzing longitudinal panel count data is presented. The model assumes that the counts for each subject are generated by a nonhomogeneous Poisson process with a smooth intensity function. Such smooth intensities are modeled with adaptive splines and allow for the incorporation of time varying covariates. Both random and discrete mixtures of intensities are considered to account for complex correlation structures, heterogeneity and hidden subpopulations common to this type of data. An estimating equation approach to inference requiring only low moment assumptions is developed and the approach is illustrated on an entomological example.

Un modèle semi-paramétrique flexible pour données de dénombrement longitudinales est présenté. Le modèle suppose que les comptes pour chaque sujet sont générés par un processus de Poisson non homogène avec une fonction d'intensité lisse. De telles intensités lisses sont modélisées avec des splines adaptatives et permettent d'incorporer des covariables variant avec le temps. Des mélanges discrets et aléatoires d'intensités sont considérés pour tenir compte des structures de corrélation complexes, de l'hétérogénéité et de sous-populations cachées communes à ce type de données. L'inférence est faite à l'aide d'équations estimatrices nécessitant uniquement des présupposés sur les premiers moments. Cette méthode est illustrée sur un exemple en entomologie.

[MS-157]**Non-parametric Density Estimation for Length Biased Data
Estimation de densités non paramétriques pour des données biaisées par la longueur**Yogendra CHAUBEY, *Concordia University*

Length-biased and censored data may appear when analyzing times of duration as well as in applications to special sampling procedures in forestry. For smooth estimation of the unbiased density and/or distribution in these situations, the usual symmetric kernels may not be appropriate. While the data in the former application may be censored, in the latter case it may not be. Here we propose a smooth estimator of the survival function in both the cases using non-negative kernels. Applications and large sample properties of the resulting estimators of density hazard and mean residual life functions are investigated.

Des données biaisées par la longueur et censurées peuvent survenir dans l'analyse de temps de durée ainsi que dans des applications de procédures d'échantillonnage spéciales en foresterie. Dans ces situations, les noyaux symétriques usuels peuvent ne pas être appropriés pour une estimation lisse et sans biais de la densité et/ou de la distribution. Alors que, dans le premier cas, les données peuvent être censurées, il est possible qu'elles ne le soient pas dans le second. Nous proposons ici un estimateur lisse de la fonction de survie pour ces deux cas en utilisant des noyaux non-négatifs. Des applications, ainsi que les propriétés lorsque les échantillons sont grands des estimateurs résultants pour les fonctions de risque et d'espérance de vie résiduelle sont étudiées.

Session 09F Mercredi 13 juin • Wednesday, June 13, 08:30–09:45 SN-2041**Stochastic Processes and Their Applications****Processus stochastiques et leurs applications****[MS-158]****Statistical Behaviour of a Spatial Renewal Process****Comportement statistique d'un processus de renouvellement spatial**Katherine DAVIES, *University of Western Ontario*

The one-dimensional renewal process has been extensively researched, and its importance in applications is well-understood. Recently, a multi-dimensional analogue of the renewal process has been proposed by Ivanoff and Merzbach (2005). The purpose of the current talk is to discuss the two-dimensional version of this process in terms of its statistical behaviour. With an algorithm developed to simulate such a process, various statistical techniques can be investigated. Primarily, this talk will focus on parametric inference for various models via simulation and real data sets. Moreover, the use of such a process to model the spread of forest fires will be explored.

Le processus de renouvellement unidimensionnel a été étudié en profondeur et son importance et ses applications sont bien comprises. Récemment, un analogue multidimensionnel de processus de renouvellement a été proposé par Ivanoff et Merzbach (2005). Le but de cet exposé est de discuter de la version bidimensionnelle de ce processus en termes de comportement statistique. Avec un algorithme développé afin de simuler un tel processus, des techniques statistiques variées peuvent être étudiées. Cet exposé sera principalement centré sur l'inférence paramétrique pour des modèles variés au moyen de simulations et de jeux de données réels. En outre, l'utilisation d'un tel processus afin de modéliser la propagation de feux de forêts sera explorée.

[MS-159]**Reinforced Urn Processes Indexed by Binary Tree****Processus d'urnes renforcés indexés par un arbre binaire**Djilali AIT AODIA, *Université de Sherbrooke*, Paolo BULLA & Pietro MULIERE, *Bocconi University, Milan*

A class of models In the vast literature devoted to urn models (see Johanson and Kotz "Urn Model and Their Applications" (1977), as a general reference) is the Pólya urn which was introduced by Eggenberger and Pólya in 1923. Muliere, P, Secchi, S, and Walker S.G have initiated the study of class of a processes called reinforced random process indexed by tree (see, Muliere *et al* "Partially Exchangeable Process Indexed by the Vertices of a k-tree constructed via reinforcement" 2005) .This process describes the evolution a system of urns in n draws associated with a tree structure. We are interested in a particular case of infinity binary tree, in fact this work deals mainly with two classical models of binary trees . We present the Muliere-Secchi-Walker construction of reinforced random processes indexed by a binary tree and we study same property of this process.

Une classe de modèles dans la vaste littérature dédiée aux modèles d'urnes (voir Johanson et Kotz "Urn Model and Their Application" (1977), comme référence générale) est l'urne de Pólya, présentée par Eggenberger et Pólya in 1923. Muliere, P, Secchi, S, et Walker S.G ont initié l'étude d'une classe de processus appelée processus aléatoires renforcés indexés par un arbre (voir, Muliere *et al.* "Partially Exchangeable Process Indexed by the Vertices of a k-tree constructed via reinforcement", 2005). Ce processus décrit l'évolution d'un système d'urnes en n tirages associés à une structure d'arbre. Nous sommes intéressés au cas particulier de l'arbre binaire infini, en fait ce travail traite principalement de deux modèles classiques d'arbres binaires. Nous présentons la construction Muliere-Secchi-Walker de processus aléatoires renforcés indexés par arbres binaires et étudions la même propriété de ce processus.

[MS-160]**Simulation of Extremes of Stochastic Differential Equations
Simulation des extrêmes d'équations différentielles stochastiques**Tingting GOU & Duncan MURDOCH, *University of Western Ontario*

In this paper we consider the problem of simulating the times where the extreme values occur of a Brownian Bridge or other diffusion. Our recursive algorithm simulates highs, lows, the time of their occurrence, times of barrier crossings etc. The method is capable of simulating these quantities simultaneously from their joint distributions. We demonstrate by example that this method can be thousands of times faster than the Euler method for equivalent accuracy.

Dans cet article nous considérons le problème de la simulation de temps où les valeurs extrêmes d'un pont brownien ou d'une autre diffusion sont réalisées. Notre algorithme récursif simule les hauts, les bas, leur temps de réalisation, les temps de franchissement de barrière, etc. La méthode est capable de simuler ces quantités simultanément à partir de leurs distributions conjointes. Nous démontrons au moyen d'un exemple que cette méthode peut être des milliers de fois plus rapide que la méthode d'Euler, pour une précision équivalente.

[MS-161]**Almost Sure Local Extinction for Super-Brownian Motion with Lebesgue Initial Measure
Extinction locale presque sûre pour le mouvement super-Brownien avec mesure initiale de Lebesgue**Xiaowen ZHOU, *Concordia University*

Let X be a d -dimensional $(1+\beta)$ -super-Brownian motion starting at Lebesgue measure. Let (A_t) be closed balls in R^d centered at 0 with radiuses $g(t)$ increasing in t . Let T be the first time such that $X_t(A_t) = 0$ for all $t > T$. We find some sharp conditions under which either $P\{T < \infty\} = 1$ or $P\{T = \infty\} = 1$. The conditions are in terms of integral tests on $g(t)$.

Soit X , un mouvement super-brownien- $(1+\beta)$ de dimension d débutant à la mesure de Lebesgue. Soit (A_t) , des boules fermées dans R^d centrées à 0, de rayons $g(t)$ croissant en t . Soit T , le premier temps où $X_t(A_t) = 0$ pour tout $t > T$. Nous trouvons des conditions ne pouvant être relâchées sous lesquelles nous avons soit $P\{T < \infty\} = 1$ ou $P\{T = \infty\} = 1$. Les conditions prennent la forme de tests d'intégrales sur $g(t)$.

[MS-162]**On some Properties of a Class of Generalized Gamma Distributions
Quelques propriétés d'une classe de distributions Gamma généralisées**Veeresh GADAG & J. KUMAR, *Memorial University*, U. BALASOORIYA & C.K. LOW, *Nanyang Technological University, Singapore*

We introduce a class of continuous distributions called class of Generalized-Gamma distributions, which includes Gamma distribution as a particular case. We show that the new class is a class of geometrically infinitely divisible distributions. We obtain a characterization property of the distributions in this class. Using the relation between the distributions in this new class and the stable distributions, we provide a method of generating the distributions in the new class. We introduce a first order autoregressive process through the distributions in this new class and study some properties.

Nous présentons une classe de distributions continues appelée la classe de distributions gamma généralisées, qui inclut la distribution gamma comme cas particulier. Nous montrons que cette nouvelle classe est une classe de distributions géométriquement infiniment divisibles. Nous obtenons une propriété caractérisant les distributions de cette classe. En utilisant la relation entre les distributions de cette nouvelle classe et les distributions stables, nous donnons une méthode pour générer les distributions de la nouvelle classe. Nous présentons un processus autorégressif de premier ordre à travers les distributions de la nouvelle classe et étudions certaines de ses propriétés.

Session 10A Mercredi 13 juin • Wednesday, June 13, 10:30–12:00 SN-2109

Special Invited Session sponsored by the Survey Methods Section
 Allocution de l'invité d'honneur du groupe de méthodologie d'enquête

[MS-163]

Environmental Surveys and Indicators – Conceptual and Methodological Challenges
Enquêtes et indicateurs environnementaux – Défis conceptuels et méthodologiques

Robert SMITH, Carolyn CAHILL & Michael BORDT, *Statistics Canada/Statistique Canada*

Statistics Canada has embarked on an expansion of its environmental program that will see a doubling of the number of surveys and an increase in derived statistical products such as indicators and accounts. The agency will begin regular collection of environmental information directly from households. It is also proposing the collection of pollution emissions data from businesses. New indicators of environmental sustainability are already produced jointly with Environment Canada and Health Canada. This work requires new thinking about the concepts and methods that underpin environment statistics. The presentation will explore some of the more novel aspects of this new thinking.

Récemment, Statistique Canada s'est lancé dans l'agrandissement de son programme environnemental où nous y verrons le double du nombre d'enquêtes et une augmentation des produits statistiques dérivés tel que les indicateurs et comptes. L'agence commencera la collecte de données environnementales auprès des ménages. Il est également proposé de faire une incursion dans le domaine de la collecte des données en termes d'émissions de pollution provenant d'entreprises. De concert avec Environnement Canada et Santé Canada, de nouveaux indicateurs portant sur la durabilité environnementale sont déjà en voie de production. Tout ce travail requiert une nouvelle façon de penser en termes de concepts et méthodes en vue d'étayer les statistiques environnementales. La présentation étudiera certains des aspects les plus inédits de cette nouvelle façon de penser.

Session 10B Mercredi 13 juin • Wednesday, June 13, 10:30–12:00 IIC-2001

Statistical Methods for Causal Inference
 Méthodes statistiques pour l'inférence causale

[MS-164]

Quantifying Dose-response for a Continuous Treatment in the Presence of Non-compliance or Confounding
Évaluation de la réponse à la dose reçue pour un traitement en présence de conformité partielle

Erica MOODIE & David A STEPHENS, *McGill University*

In a longitudinal study of dose-response, non-compliance compromises the estimation of the true effect of treatment. Standard regression methods cannot remove the bias introduced by patient-selection of dose. Using an approach based on the Generalized Propensity Score (GPS) [Hirano and Imbens, 2004], it is possible to construct a balancing score to provide a more meaningful (and less biased) estimation procedure for the true effect of dose. We apply the GPS methodology using on a novel formulation of the treatment density to the Monitored Occlusion Treatment of Amblyopia Study (MOTAS), which investigated the dose-response relationship between occlusion and visual acuity.

Dans une étude longitudinale dose-réponse, la non-observance compromet l'estimation de l'effet réel du traitement. Les méthodes de régression ordinaires ne peuvent pas éliminer le biais introduit par la sélection de la dose par le patient. En utilisant une approche basée sur le score de propension généralisé (SPG) [Hirano et Imbens, 2004], il est possible de construire un score d'équilibrage qui nous donne une procédure d'estimation plus sensée (et moins biaisée) de l'effet réel de la dose. Nous appliquons la méthodologie SPG en utilisant une nouvelle formulation de la densité du traitement à l'étude du traitement de l'amblyopie à occlusion surveillée, qui étudie la relation dose-réponse entre l'occlusion et l'activité visuelle.

[MS-165]**Variable Selection in Marginal Structural Models****Sélection de variable dans les modèles marginaux structuraux**Robert PLATT, J.A. “Chris” DELANEY & Geneviève LEFEBVRE, *McGill University*

Like many approaches to causal inference in observational data, marginal structural models (MSMs) require a model for the distribution of exposure conditional on confounders. We explore MSMs across a range of scenarios, including one reflecting a large pharmacoepidemiologic study. We show that including non-confounding predictors of exposure in treatment models can lead to biased and highly variable estimators in small to moderate samples. Using variables related only to the outcome is less biased. Estimates using all predictors of exposure appear asymptotically unbiased, but convergence is slow. We discuss these results in the context of other results on treatment models.

Comme plusieurs approches à l'inférence causale à partir de données observationnelles, les modèles marginaux structuraux (MMSs) nécessitent un modèle pour la distribution de l'exposition conditionnelle aux facteurs parasites. Nous explorons les MMSs à travers une série de scénarios, dont un reflétant une grande étude pharmacoépidémiologique. Nous montrons que l'inclusion de prédicteurs non-parasites de l'exposition dans les modèles du traitement peut mener à des estimateurs biaisés et hautement variables dans des échantillons de petite à moyenne taille. L'utilisation de variables liées uniquement à la réponse réduit ce biais. Les estimés utilisant tous les prédicteurs de l'exposition semblent asymptotiquement sans biais, mais la convergence est lente. Nous discutons de ces résultats parmi d'autres sur les modèles de traitement.

[MS-166]**Causal Analyses of Post-randomization Effect Modification in Clinical Trials****Analyses causales de modification post-randomisation des effets en essais cliniques**Thomas TENHAVE, Jennifer FAERBER & Marshall JOFFE, *University of Pennsylvania*, Robert GALLOP, *Westchester University*

We present a linear rank preserving model (RPM) approach for analyzing post-randomization interactions in randomized studies. Unlike standard interaction analyses, our approach does not assume that the post-randomization effect modifier is also randomly assigned to individuals in addition to the randomized baseline intervention (i.e., sequential ignorability). The *G*-estimation procedure for the proposed RPM represents an extension of the work on direct effects of randomized intervention effects by Robins and Greenland (1994) and Ten Have *et al.* (2007). Comparisons are made with principal stratification in the context of two suicide/depression intervention studies.

Nous présentons une approche par modèle linéaire préservant les rangs (MPR) pour l'analyse d'interactions post-randomisation en études randomisées. Contrairement aux analyses d'interactions habituelles, notre approche ne suppose pas que le modificateur d'effet post-randomisation soit aussi attribué aléatoirement aux individus en plus de l'intervention de base randomisée (c.-à-d. l'ignorabilité séquentielle). La procédure de *G*-estimation pour le MPR proposée représente une généralisation des travaux de Robins et Greenland (1994) et Ten Have *et al.* (2007) sur les effets directs d'effets d'intervention randomisés. Des comparaisons sont faites avec la stratification principale dans le contexte de deux études d'intervention en suicide/dépression.

Session 10C Mercredi 13 juin • Wednesday, June 13, 10:30–12:00 SN-2041**Matrices and Statistics I****Matrices et statistique I****[MS-167]****Ordering (0,1)-matrices, Contingency Tables, and Stochastic Matrices****Ordonner des matrices (0,1), des tableaux de contingence et des matrices stochastiques**Richard BRUALDI, *University of Wisconsin – Madison*

There are several ways in which one might (partially) order the matrices belonging to a class of matrices (defined e.g. by specifying the total sum of all the entries or the sum of entries in each row and column). Some of these use the majorization (a.k.a. dominance) order. Another way is to generalize the classical Bruhat (partial) order of permutations of $1, 2, \dots, n$. After a brief review, we shall discuss this new order for the matrices in the title and evaluate its implications. (Talk partially based on joint work with Suk-Geun Hwang and Louis Deaett.)

Il y a plusieurs façons possible d'ordonner (partiellement) les matrices d'une classe de matrices (définie par exemple en spécifiant la somme totale de tous les éléments, ou la somme des éléments de chaque rangée et colonne). Certaines utilisent l'ordre de majorisation (c.-à-d. dominance). Une autre façon est de généraliser l'ordre Bruhat (partiel) classique des permutations de $1, 2, \dots, n$. Après une brève revue, nous allons discuter de ce nouvel ordre pour les matrices du titre et évaluer ses implications. (Exposé en partie basé sur un travail en collaboration avec Suk-Geun Hwang et Louis Deaett.)

[MS-168]**Coupling and Mixing Times in Markov Chains****Couplage et temps de stationnarité des chaînes de Markov**Jeffrey HUNTER, *Massey University, Auckland, New Zealand*

We explore the properties of the time to coupling and the time to mixing in Markov chains. In particular, we compare the expected time to coupling with the expected time to mixing (as introduced by the author in “Mixing times with applications to perturbed Markov chains”, *Linear Algebra Appl.* (417, 108–123 (2006).) The two-state cases and three state cases are examined in detail.

Nous explorons les propriétés du temps de couplage et de stationnarité des chaînes de Markov. En particulier, nous comparons le temps de couplage moyen avec le temps de stationnarité moyen (tel qu'introduit par l'auteur dans “Mixing times with applications to perturbed Markov chains”, *Linear Algebra Appl.* (417, 108–123 (2006).) Les cas des chaînes à deux et à trois états sont examinés en détail.

[MS-169]**Some Comments on the Authorship Matrix in Bibliometrics****Quelques commentaires sur la matrice de paternité littéraire en bibliométrie**Simo PUNTANEN, *University of Tampere, Finland*, Oskar Maria BAKSALARY, *Adam Mickiewicz University, Poland*, George P.H. STYAN, *McGill University*

Associated with a bibliography by a particular author (with several coauthors), we define an “authorship matrix” $A = a_{ij}$, where $a_{ij} = 1$ if bibliographic entry number i is written with coauthor number j , and $a_{ij} = 0$ otherwise. Following work by Oskar Maria Baksalary and George P. H. Styan [*Linear Algebra Appl.* 410 (2005), 3–53], we interpret the

Associée à une bibliographie par un auteur donné (avec plusieurs coauteurs), nous définissons une “matrice de paternité littéraire” $A = a_{ij}$, où $a_{ij} = 1$ si l'entrée bibliographique numéro i est écrite avec le coauteur j , et $a_{ij} = 0$ sinon. Suivant les travaux d'Oskar Maria Baksalary et George P. H. Styan [*Linear Algebra Appl.* 410 (2005), 3–53], nous interprétons les valeurs propres et vecteurs propres de $A'A$, avec une ré-

eigenvalues and eigenvectors of $A'A$, with particular reference to the authorship matrices for Jerzy K. Baksalary (1944–2005), Sujit Kumar Mitra (1932–2004) and Milton Sobel (1919–2002).

férence particulière aux matrices de paternité littéraire pour Jerzy K. Baksalary (1944–2005), Sujit Kumar Mitra (1932–2004) et Milton Sobel (1919–2002).

Session 10D Mercredi 13 juin • Wednesday, June 13, 10:30–12:00 SN-2105

Point Processes

Processus ponctuels

[MS-170]

A Compensator Characterization of Point Processes on Partially Ordered Metric Spaces

Une caractérisation compensatrice de processus ponctuels sur des espaces métriques partiellement ordonnés

Gail IVANOFF, *University of Ottawa*, Ely MERZBACH, *Bar Ilan University*, Mathieu PLANTE, *none*

The compensator (the integral of the stochastic intensity) of a point process on the real line is of fundamental importance since it characterizes the law of the process. How to extend this classical result of Jean Jacod to point processes on more general partially ordered sets has been a longstanding open problem. Specifically, how does one define the compensator of a point process on a general partially ordered set in such a way that the compensator exists, is unique, and characterizes the law of the process? We resolve this problem in two steps. First, when the point process is totally ordered, we define the compensator as a family of one-parameter compensators. We prove that this family is unique and characterizes the finite dimensional distributions of the point process. Second, this result can then be applied to an arbitrary point process via an embedding of the process into a totally ordered point process on a larger space. We present some examples, including the multiparameter partial sum process, single line point processes, multiparameter renewal processes, and obtain a new characterization of the two-parameter Poisson process.

Le compensateur (l'intégrale de l'intensité stochastique) d'un processus ponctuel sur les réels est d'importance fondamentale puisqu'il caractérise la loi du processus. Comment généraliser ce résultat classique de Jean Jacod aux processus ponctuels sur des ensembles partiellement ordonnés plus généraux a longtemps été un problème ouvert. Spécifiquement, comment définir le compensateur d'un processus ponctuel sur un ensemble partiellement ordonné général de façon à ce que le compensateur existe, soit unique et caractérise la loi du processus? Nous résolvons ce problème en deux étapes. Premièrement, lorsque le processus ponctuel est totalement ordonné, nous définissons le compensateur comme une famille de compensateurs à un paramètre. Nous prouvons que cette famille est unique et caractérise les distributions à dimension finie du processus ponctuel. Deuxièmement, ce résultat peut ensuite être appliqué à un processus ponctuel arbitraire en emboîtant le processus dans un processus ponctuel totalement ordonné sur un espace plus grand. Nous présentons quelques exemples, incluant le processus de sommes partielles multiparamétriques, les processus ponctuels sur une seule ligne, les processus de renouvellement multiparamétriques, et obtenons une nouvelle caractérisation du processus de Poisson à deux paramètres.

[MS-171]

The Interplay between Event Times and Event Characteristics

L'interaction entre temps d'événements et caractéristiques d'événements

James RAMSAY, *McGill University*

Event times, or point processes can involve all three types of time: long term unbounded time, short term elapsed time, and circular or seasonal time. These three roles for time can jointly vary in com-

Les temps d'événements, ou processus ponctuels, peuvent impliquer trois types de temps : temps à long terme non borné, temps écoulé à court terme, et temps circulaire ou saisonnier. Ces trois types de temps peuvent varier conjointe-

plex ways. Events often have measurable characteristics, and these can depend on the three time types as well as on covariates external to the process. We illustrate MPP concepts with some of our favorite data, and show how functional data analysis methodology can be adapted to describe interesting features. MPP's look to us like a large real-world problem waiting for new approaches.

ment de façons complexes. Les événements ont souvent des caractéristiques mesurables, et celles-ci peuvent dépendre des trois types de temps ainsi que de covariables externes au processus. Nous illustrons des concepts MPP à l'aide de certaines de nos données favorites, et montrons comment la méthodologie d'analyse de données fonctionnelles peut être adaptée afin de décrire des caractéristiques intéressantes. Les MPPs nous semblent un grand problème du monde réel attendant le développement de nouvelles approches.

Session 10E Mercredi 13 juin • Wednesday, June 13, 10:30–12:00 AA-1043

Biostatistical Methods Méthodes biostatistiques

[MS-172]

Stepwise Confidence Intervals for Monotone Dose-response Studies Intervalles de confiance pas-à-pas pour les études de doses/réponses monotones

Jianan PENG, *Acadia University*, Chun-In LEE, *Memorial University*, Karolyn DAVIS, *Carleton University*

In dose-response studies, one of the most important issues is the identification of the minimum effective dose (MED), where the MED is defined as the lowest dose such that the mean response is better than the mean response of a zero-dose control by a clinically significant difference. Usually the dose-response curves are monotonic. Various authors have proposed step-down test procedures based on contrasts among the sample means to find the MED. In this paper, we improve Marcus and Peritz's method (1976, *Journal of Royal Statistical Society, Series B*, Vol 38, 157–165) and combine Hsu and Berger's DR method (1999, *Journal of the American Statistical Association*, Vol 94, 468–482) to construct the lower confidence bound for the difference between the mean response of any non-zero dose level and that of the control under the monotonicity assumption to identify the MED. The proposed method is illustrated by numerical examples and simulation studies on power comparisons are presented.

En études dose-réponse, l'un des problèmes les plus importants est l'identification de la dose effective minimale (DEM), où la DEM est définie comme la plus petite dose telle que la réponse moyenne est meilleure que la réponse moyenne d'un contrôle à dose nulle par une différence cliniquement significative. Les courbes dose-réponses sont habituellement monotones. Plusieurs auteurs ont proposé des procédures de tests descendants basés sur les contrastes des moyennes échantillonnales afin de trouver la DEM. Dans cet article, nous améliorons la méthode de Marcus et Peritz (1976, *Journal of Royal Statistical Society, Series B*, Vol 38, 157–165) et combinons la méthode DR de Hsu et Berger (1999, *Journal of the American Statistical Association*, Vol 94, 468–482), afin de construire la borne inférieure de l'intervalle de confiance pour la différence entre la réponse moyenne de n'importe quel niveau de dose non nulle et celle du contrôle sous l'hypothèse de monotonie pour identifier la DEM. La méthode proposée est illustrée par des exemples numériques et des études de simulation comparant les puissances sont présentées.

[MS-173]

Estimation of Prediction Loss with Censored Survival Data**Estimation de la perte de prévision en présence de données de survie censurées**Yan YUAN, *University of Waterloo*

Measures of prediction power for regression models are typically based on the expected loss of a predictor. For survival data, the absolute error loss is often used; the corresponding optimal predictor, median survival time, is of frequent interest. However, the actual loss of a censored observation is unknown and so estimation of the expected loss using validation data or cross-validation is problematic. We give an alternative way of estimating the absolute error loss when censoring is present.

Les mesures de la puissance prédictive de modèles de régression sont habituellement basées sur la perte espérée d'un prédicteur. Pour des données de survie, la perte de type erreur absolue est souvent utilisée ; le prédicteur optimal correspondant, le temps de survie médian, est d'intérêt fréquent. Cependant, la perte réelle d'une observation censurée est inconnue, et l'estimation de la perte espérée utilisant des données de validation ou la validation croisée est problématique. Nous donnons une façon alternative d'estimer la perte de type erreur absolue lorsque des données sont censurées.

[MS-174]

Optimal Crossover Designs: Investigating the Role of Baseline Measurements**Plans alternés optimaux : éclairer le rôle des mesures de base**Yuanyuan LIANG & Kimmie Chough CARRIERE, *University of Alberta*

Generally, crossover designs lose the practical appeal when required to allow more than the traditional first-order carryover effects. This is when utilizing the baseline measurements clearly saves the design efficiency. We explore the effect of baseline measurements on optimal design results. We strongly recommend to use the baseline measurements as they improve the efficiency by 2-3 folds. We also find that there is a dramatic reduction in variability for estimating the direct treatment effect contrast when extending two-period designs to three-period or four-period designs. Such a reduction is much more pronounced when baseline measurements are incorporated in each period.

Les plan alternés perdent généralement leur attrait pratique lorsqu'il est nécessaire de permettre plus que les effets résiduels de traitement de premier ordre traditionnels. C'est le cas lorsque l'utilisation des mesures de base permet clairement une plus grande efficacité du plan. Nous explorons les effets des mesures de base sur les résultats du plan optimal. Nous recommandons fortement d'utiliser les mesures de base car elle peuvent augmenter l'efficacité par un facteur de 2 ou 3. Nous trouvons également qu'il y a une réduction dramatique de variabilité dans l'estimation des contrastes directs de l'effet traitement lorsque les plans à deux périodes sont étendus à trois ou quatre périodes. Cette réduction est beaucoup plus prononcée lorsque les mesures de base sont incorporées à chaque période.

[MS-175]

A Bayesian Algorithm to Detect Genomic DNA Copy Number Changes in Array CGH Data**Un algorithme bayésien pour détecter les changements du nombre de copies d'ADN génomique dans des données de puce CGH**Long Yang WU, *University of Waterloo*

ArrayCGH is a high-throughput technology to generate genomic DNA copy number profiles. It plays an important role in cancer research and diagnosis (Feuk et al., *Nature Review Genetics* 7, 85–97). We present a Bayesian algorithm to detect DNA copy number changes in arrayCGH. The algorithm implements a flexible mean-variance shift model for

La puce CGH est une technologie à haut débit permettant de générer des profils du nombre de copies d'ADN génomique. Elle joue un rôle important dans la recherche et le diagnostic du cancer (Feuk et al., *Nature Review Genetics* 7, 85–97). Nous présentons un algorithme bayésien pour détecter les changements du nombre de copies d'ADN dans des données de puce CGH. L'algorithme met en oeuvre un modèle de dé-

chromosomal segments with different copy numbers. A peeling procedure equipped with Bayes factor was employed to estimate segments boundaries. Compared with many current methods, it has the advantages of ranking chromosomal segments and identifying influential observations. We illustrate the algorithm with real data analysis and simulation studies.

calage flexible de la moyenne et variance pour les segments chromosomaux avec différents nombres de copies. Une procédure de décollage munie de facteurs de Bayes a été employée pour estimer les frontières des segments. Comparativement à plusieurs méthodes disponibles, elle a l'avantage de classer les segments chromosomaux et d'identifier les observations influentes. Nous illustrons l'algorithme avec une analyse de données réelles et des études de simulation.

[MS-176]

**The Efficiency of Estimation for Response Adaptive Designs
L'efficacité de l'estimation sous des plans adaptatifs à la réponse**

Yanqing Yi, *University of Manitoba*

This paper discusses the efficiency of estimation for a general class of response adaptive designs. The lower bound of exponential rate of consistent estimators of parameters is established and it is shown that under some assumptions the exponential rate of the maximum likelihood estimator attains the lower bound for response adaptive trials with dichotomous responses. Under certain conditions, it is proved that the maximum estimator of treatment effects is asymptotic efficient in Bahadur's sense for a general class of response adaptive trials.

Cet article discute de l'efficacité de l'estimation pour une classe de plans adaptatifs à la réponse. La borne inférieure du taux exponentiel d'estimateurs convergents de paramètres est établie et il est montré que, sous certaines suppositions, le taux exponentiel de l'estimateur du maximum de vraisemblance atteint la borne inférieure pour des essais adaptatifs à la réponse avec réponses dichotomiques. Sous certaines conditions, il est prouvé que l'estimateur du maximum des effets du traitement est asymptotiquement efficace au sens de Bahadur pour une classe générale d'essais adaptatifs à la réponse.

[MS-177]

**Central Venous Line-related Thromboembolism in Children with Lymphoreticular Malignancy
Maladie thromboembolique liée à l'emploi d'un cathéter central veineux chez les enfants atteints de lymphoréticulopathie**

Weiwei XIONG, *McMaster University*

Children with acute lymphoblastic leukemia (ALL) are at high risk for thromboembolism (TE). Central venous line (CVL) is a reported risk factor for TE in children. This presentation is based on a non-interventional retrospective study of 150 children with lymphoreticular malignancy treated since 1990. We evaluate the incidence and prevalence of CVL-related thrombosis and infections. We also determine whether CVL type, vein location, site of insertion and insertion technique are associated the incidence of TE.

Les enfants atteints de leucémie aiguë lymphoblastique (LAL) sont à haut risque pour la thromboembolie (TE). L'utilisation d'un cathéter central veineux (CCV) est un facteur de risque observé pour la TE chez les enfants. Cette présentation est basée sur une étude rétrospective non-interventionnelle de 150 enfants atteints de lymphoréticulopathie traités depuis 1990. Nous évaluons l'incidence et la prévalence des thromboses et infections liées au CCV. Nous déterminons aussi si le type de CCV, l'emplacement de la veine, le site d'insertion et la technique d'insertion sont associés à l'incidence de la TE.

Session 10F Mercredi 13 juin • Wednesday, June 13, 10:30–11:45 AA-1046**Statistical Theory III
Théorie statistique III****[MS-178]****Bayes, MCMC and Copulas****Bayes, MCMC et Copules**SIMON GUILLOTTE & FRANÇOIS PERRON, *Université de Montréal*

Copulas are functions linking the marginal distributions together to form the joint distribution of a random vector. They hold the dependence structure between the individual variates. Copulas are becoming widely used by the applied sciences community, and in the past decade, they have also become the standard tool used in financial modeling. It is thus of actual interest to construct efficient estimators. We propose a Bayesian approach in a non-parametric setup. Some kernel type estimators exist in the literature and are known to behave well asymptotically, although not necessarily for finite samples. On the other hand, the estimator we propose has better behavior for such samples and becomes an interesting candidate to use in practice. The drawback of our method resides in the complexity of the computations required. We use high performance computing techniques which I shall also describe.

Nous proposons une approche Bayésienne pour l'estimation non paramétrique de copules. Ce sujet est très actuel, étant donné le nombre croissant d'applications des copules pour modéliser la dépendance et ce, dans des contextes très variés. Quelques autres méthodes non paramétriques d'estimation font présentement partie de la littérature, cependant, bien que les estimateurs proposés se comportent convenablement de manière asymptotique, ils possèdent certains défauts pour des échantillons de tailles finies. L'estimateur que nous proposons ne possède pas ces défauts et, en ce sens, il devient un candidat intéressant à utiliser en pratique. Le prix à payer pour la méthode que nous proposons se situe dans la complexité des calculs pour y parvenir. L'aspect computationnel est donc non-négligeable, nous faisons appel à des techniques de calcul à haute performance, notamment au parallélisme sur des ordinateurs à multiples processeurs.

[MS-179]**Fast approximately balanced bootstrap without construction****Bootstrap rapide approximativement équilibré sans construction**DEVON LIN, *Simon Fraser University*, WILSON LU, *Acadia University*, RANDY SITTER, *Simon Fraser University*

We develop second-order balanced and approximately balanced resamples that retain the nice properties demonstrated by Graham *et al.* (1990) but require essentially no design construction.

Nous développons des rééchantillons approximativement équilibrés de second ordre qui retiennent les bonnes propriétés démontrées par Graham *et al.* (1990) mais ne nécessitent essentiellement aucune construction du plan.

[MS-180]**An Efficient Randomized Quasi-Monte Carlo Algorithm for the Pareto Distribution****Un algorithme quasi-Monte-Carlo randomisé efficace pour la distribution Pareto**MARCO POLLANEN, *Trent University*, MEI-LING HUANG & WAI KONG YUEN, *Brock University*

In this talk we present a new randomized quasi-Monte Carlo method for estimating the mean and variance of the Pareto distribution. In many Monte Carlo simulations, there are some stability prob-

Dans cet exposé nous présentons une nouvelle méthode randomisée quasi Monte-Carlo pour estimer la moyenne et la variance de la distribution Pareto. Dans plusieurs simulations Monte-Carlo, il y a des problème de stabilité lors de

lems for estimating the population Pareto variance by using the sample variance. We propose a randomized quasi-random number generator to generate Pareto random samples, such that the sample mean and sample variance estimators gain more efficiency. This technique can also be applied to estimates of other thick-tailed distributions.

l'estimation de la variance d'une population Pareto en utilisant la variance échantillonnale. Nous proposons un générateur de nombres quasi-aléatoires randomisé pour générer des échantillons aléatoires de distribution Pareto permettant une amélioration de l'efficacité de la moyenne échantillonnale et de la variance échantillonnale. Cette technique peut aussi être appliquée à d'autres distributions à ailes relevées.

[MS-181]

Goodness-of-fit Tests for Copulas based on Spearman's Dependence Function

Tests d'adéquation pour les copules basés sur la fonction de dépendance de Spearman

Marie-Hélène TOUPIN, Mhamed MESFIOUI & Jean-François QUESSY, *Université du Québec à Trois-Rivières*

In this talk, a new family of semi-parametric goodness-of-fit tests for copula models will be introduced. First, Spearman's dependence function will be defined and computed for many copula models. Then, an entirely non-parametric estimation of this function is proposed and the weak convergence of an associated empirical process is established. This result enables to build consistent Cramér-von Mises type goodness-of-fit statistics. It will be shown by simulation that the latter often lead to powerful tests. An illustrative example on hydrological data will close the presentation.

Dans cette présentation, une nouvelle famille de tests semi-paramétriques d'adéquation pour des modèles de copules sera introduite. En premier lieu, la fonction de dépendance de Spearman sera définie et calculée pour plusieurs modèles de copules. Ensuite, une estimation entièrement non-paramétrique de cette fonction sera proposée et la convergence faible d'un processus empirique associé sera établie. Ce résultat permet de construire des statistiques d'adéquation de type Cramér-von Mises convergentes. On montrera par simulations que ces statistiques conduisent souvent à des tests puissants. Une illustration sur un jeu de données hydrologiques clôturera l'exposé.

[MS-182]

Regularized greedy importance sampling

Échantillonnage d'importance régularisé avare

Ali GHODSI, *University of Waterloo*, Dale SCHUURMANS, *University of Alberta*, Finnegan SOUTHEY, *Google Inc.*

We present a new method for conducting Monte Carlo inference which combines explicit search with importance sampling. The idea is to reduce the variance of importance sampling by searching for significant points in the target distribution. Previous work has demonstrated the feasibility of implementing this method and proved that the technique is unbiased in both discrete and continuous domains. In this work we demonstrate our procedure on a few simple inference tasks and show that it can improve the inference quality of standard MCMC methods, including Gibbs sampling, Metropolis sampling, and Hybrid Monte Carlo.

Nous présentons une nouvelle méthode pour effectuer de l'inférence par Monte-Carlo combinant la recherche explicite avec l'échantillonnage d'importance. L'idée est de réduire la variance de l'échantillonnage d'importance en cherchant des points importants dans la distribution cible. Des travaux précédents ont démontré la faisabilité de cette méthode et prouvé que la technique est sans biais dans des domaines discrets et continus. Dans ce travail, nous appliquons notre procédure à quelques tâches d'inférence simple et montrons qu'elle peut améliorer la qualité de l'inférence des méthodes MCMC standard, incluant l'échantillonneur de Gibbs, l'échantillonnage Metropolis, et la méthode Monte-Carlo hybride.

Session 11A Mercredi 13 juin • Wednesday, June 13, 13:30–14:30 IIC-2001

Address of the winner of the Centre de Recherches Mathématiques/Statistical Society of Canada Prize

Allocution du récipiendaire du prix Centre de recherches mathématiques/Société statistique du Canada

Session 12A Mercredi 13 juin • Wednesday, June 13, 15:00–16:30 AA-1043

Biostatistics
Biostatistique

[MS-183]

Second-order Least Squares Estimation of Censored Regression Models**Estimation de modèles de régression censurés par moindres carrés du second ordre**

Taraneh ABARIN & Liqun WANG, *University of Manitoba*

We propose the second-order least squares estimation which is an extension of the ordinary least squares method, for censored regression models when the error term has a general parametric distribution (not necessarily normal). We prove strong consistency and asymptotic normality of the estimator under some regularity conditions, propose a computationally simpler estimator which is consistent and asymptotically normal under the same regularity conditions. Simulation studies compare the finite sample behavior of the estimator with the MLE.

Nous proposons l'estimation par moindres carrés du second ordre, une extension de la méthode des moindres carrés ordinaires, pour des modèles de régression censurés où le terme d'erreur a une distribution paramétrique générale (pas nécessairement normale). Nous prouvons la convergence forte et la normalité asymptotique de l'estimateur sous certaines conditions de régularité. De plus, nous proposons un estimateur plus simple à calculer qui est aussi convergent et asymptotiquement normal sous les mêmes conditions de régularité. Des études de simulation comparent le comportement de cet estimateur et de l'EMV dans un contexte d'échantillon fini.

[MS-184]

Parsimonious Gaussian Mixture Models**Modèles de mélanges gaussiens parcimonieux**

Paul McNICHOLAS & Brendan MURPHY, *Trinity College Dublin*

Parsimonious Gaussian mixture models (PGMMs) are developed using a latent Gaussian model. These models provide a united modeling framework which includes the mixture of probabilistic principal component analyzers and mixture of factor of analyzers models as special cases. This family of PGMMs, which are based on the mixtures of factor analyzers, includes parsimonious models that have not previously been developed. Maximum likelihood estimates for the parameters in these models are found using an AECM algorithm. These models are applied to the datasets from the areas of food science and bio-informatics and they are shown to give excellent results.

Les modèles de mélanges gaussiens parcimonieux (MMGP) sont développés en utilisant un modèle gaussien latent. Ces modèles fournissent une structure de modélisation unie incluant le mélange d'analyseurs en composantes principales probabilistes et le mélange de modèles d'analyseurs de facteurs comme cas spéciaux. Cette famille de MMGP, qui sont basés sur les mélanges d'analyseurs de facteurs, inclut des modèles parcimonieux n'ayant pas déjà été développés. Les estimateurs du maximum de vraisemblance pour les paramètres de ces modèles sont trouvés au moyen d'un algorithme AECM. Ces modèles sont appliqués à des jeux de données de bromatologie et de bioinformatique et nous montrons qu'ils donnent d'excellents résultats.

[MS-185]**Will Quadratic Inference Function (QIF) Take Over Generalized Estimating Equations (GEE) in Longitudinal Data Analysis?****Les fonctions d'inférence quadratiques remplaceront-elles les équations d'estimations généralisées pour les analyses de données longitudinales ?**Peter SONG, *University of Waterloo*

In this talk, I will present some recent results on the quadratic inference function (QIF), which is a method of statistical inference proposed originally by Qu *et al.* (2000) in marginal models for longitudinal data. In particular, I will show some comparisons of the QIF with the currently popular generalized estimating equations (GEE) method. Our SAS Macro QIF will be introduced in the presentation to carry out related numerical calculations involved in our comparison study.

Dans cet exposé, je vais présenter certains résultats récents sur la fonction d'inférence quadratique (FIQ), qui est une méthode d'inférence statistique originellement proposée par Qu *et al.* (2000) pour des modèles marginaux pour données longitudinales. En particulier, je vais montrer quelques comparaisons de la FIQ avec la méthode populaire des équations estimatrices généralisées (EEG). Notre macro SAS FIQ sera introduite dans la présentation pour effectuer les calculs numériques impliqués dans notre étude comparative.

[MS-186]**Inverse Probability Weighted GEE with Auxiliary Progressive Processes
EEG pondérées par probabilités inverses avec processus progressifs auxiliaires**Baojiang CHEN, Grace YI & Richard COOK, *University of Waterloo*

Missing data arise commonly in longitudinal studies and present a major challenge for statistical analysis. We describe methods for the analysis of incomplete longitudinal data when an auxiliary response is available that follows a progressive model. Parameters of the missing data process are estimated based on such an auxiliary process using the Expectation-Maximization (EM) algorithm. Weighted estimating equations are introduced for the response of interest to estimate the covariate effects. Simulation results indicate that this method works well for missing at random and missing not at random mechanisms.

Les données manquantes surviennent fréquemment dans les études longitudinales et présentent un défi majeur à l'analyse statistique. Nous décrivons des méthodes pour l'analyse de données longitudinales incomplètes lorsqu'une réponse auxiliaire suivant un modèle progressif est disponible. Les paramètres du processus des données manquantes sont estimés à partir du processus auxiliaire en utilisant l'algorithme de maximisation de l'espérance (algorithme EM). Des équations estimatrices pondérées sont présentées pour la réponse d'intérêt afin d'estimer les effets des covariables. Des résultats de simulation indiquent que cette méthode fonctionne bien en présence de données manquantes aléatoirement et non aléatoirement.

[MS-187]**Comparison of Kernel Estimating Equations with Global and Local Variance Structures in Time-varying Coefficient Models****Comparaisons d'équations estimatrices de noyaux à structures de variances globale et locale pour des modèles à coefficients variant avec le temps**Qian ZHOU, *University of Waterloo*, Huazhen LIN, *Sichuan University, China*, Peter X.-K. SONG, *University of Waterloo*

This presentation focuses on two kernel-based estimation methods in time-varying coefficient models for longitudinal data. In the formulation of local kernel GEE, one method proposed by Carroll and

Cette présentation porte sur deux méthodes d'estimation par noyaux de modèles à coefficient variant selon le temps pour données longitudinales. Dans la formulation de l'EEG du noyau local, une méthode proposée par Carroll et Lin (2000)

Lin (2000) uses a global variance structure, while the other method proposed by Chen and Jin (2005) uses a local variance structure. A comparison of these two methods will be illustrated via simulation studies. Our conclusions from this investigation would provide useful guidelines to practitioners for their choice between these two kernel GEE methods.

utilise une structure de variance globale, alors que la méthode proposée par Chen et Jin (2005) utilise une structure de variance locale. Une comparaison de ces méthodes sera illustrée à l'aide d'études de simulation. Nos conclusions suite à cette investigation devraient fournir des lignes directrices aux praticiens dans leur choix entre ces deux méthodes d'EEG de noyaux.

Session 12B Mercredi 13 juin • Wednesday, June 13, 15:00–16:30 IIC-2001

Applications of Empirical Likelihood in Survey Sampling Applications de la vraisemblance empirique en échantillonnage

[MS-188]

Calibration Estimation using Empirical Likelihood in Unequal Probability Sampling Estimation de la calibration utilisant la vraisemblance empirique dans le contexte de l'échantillonnage à probabilités inégales

Jae-kwang KIM, *Yonsei University*

Calibration estimation, which can be roughly described as a method of adjusting the original design weights to incorporate the known population totals of the auxiliary variables, has become very popular in sample surveys. Under simple random sampling, Chen and Qin (1993) suggested that the calibration estimator maximizing the constrained empirical likelihood can make efficient use of the auxiliary variables. We extend the result to unequal probability sampling. Asymptotic properties of the proposed calibration estimator are discussed. Extensions of the proposed method are discussed and results from a limited simulation study are presented.

L'estimation par calibration, qui peut être décrite comme une méthode d'ajustement des poids originaux afin d'incorporer les totaux de populations connus de variables auxiliaires, est devenue très populaire dans les enquêtes par sondages. Sous l'échantillonnage aléatoire simple, Chen et Qin (1993) ont suggéré que l'estimateur par calibration maximisant la vraisemblance empirique contrainte peut faire un usage efficace des variables auxiliaires. Nous étendons ce résultat à l'échantillonnage à probabilités inégales. Nous discutons des propriétés asymptotiques de l'estimateur de calibration proposé et de généralisations de la méthode proposée. Nous présentons des résultats d'une étude de simulation limitée.

[MS-189]

Efficient Semiparametric Estimation with Surrogate Outcome Data Estimation semi-paramétrique efficace en présence de données de substitution

Song CHEN, *Iowa State University*

We consider parameter estimation where some variable is missing in some observations; however, a surrogate variable of the missing variable is always available. We propose using the empirical likelihood to obtain weights based on estimating functions for the covariate and the surrogate. These weights are then used to weigh the original estimating equations for parameter estimation. The proposed method has the robustness property of

Nous considérons l'estimation paramétrique lorsqu'une variable n'est pas disponible pour certaines observation, mais qu'une variable de substitution de la variable manquante est toujours disponible. Nous proposons l'utilisation de la vraisemblance empirique afin d'obtenir des poids basés sur les fonctions estimatrices pour la covariable et la variable de substitution. Ces poids sont ensuite utilisés pour pondérer les équations estimatrices originales pour l'estimation paramétrique. La méthode proposée a la propriété de mener à

giving consistent parameter estimates if either the propensity score or the conditional estimating function is correctly specified. We apply the method to a dataset from the 1996 U.S. presidential election.

des estimateurs convergents si le score de propension ou la fonction estimatrice conditionnelle sont bien spécifiés. Nous appliquons la méthode à un jeu de données de l'élection présidentielle américaine de 1996.

[MS-190]

Empirical Likelihood Inferences under Measurement Errors with Complex Survey Data

Inférences par vraisemblance empirique avec des erreurs de mesures en présence de données d'enquête complexes

Suojin WANG, *Texas A&M University*, Jiahua CHEN, *Univeristy of British Columbia*, Changbao WU, *University of Waterloo*

In this talk, we demonstrate that the general concept of empirical likelihood (EL) can be effectively used for inferences under measurement errors when the sample data are collected through a complex survey design. We consider two samples where one is measured with error and the other without error, and the two samples are either nested or independent. Several estimation techniques, including calibration and a two-sample density ratio model, are used to bring models on measurement errors into the EL inferences. Some simulation results on the estimation of population mean, the distribution function and some other parameters of interest are reported.

Dans cette présentation, nous démontrons que le concept général de vraisemblance empirique (VE) peut être effectivement utilisé pour des inférences en présence d'erreurs de mesures lorsque les données sont recueillies selon un plan de sondage complexe. Nous considérons deux échantillons, un présentant des erreurs de mesures et l'autre sans erreur, et les deux échantillons sont soit emboîtés, soit indépendants. Plusieurs techniques d'estimation, incluant la calibration et un modèle de rapport de densités de deux échantillons, sont utilisés pour introduire ces modèles sur les erreurs de mesures dans les inférences VE. Des résultats de simulation sur l'estimation de la moyenne de population, de la fonction de répartition et de d'autres paramètres d'intérêt sont rapportés.

Session 12C Mercredi 13 juin • Wednesday, June 13, 15:00–16:30 SN-2105

Matrices and Statistics II

Matrices et statistique II

[MS-191]

On the Linear Aggregation Problem in the General Gauss-Markov Model

Sur le problème d'agrégation linéaire dans un modèle général de Gauss-Markov

Hans Joachim WERNER, *University of Bonn*

We consider the linear aggregation problem in the general possibly singular Gauss-Markov model. For the true underlying micro relations, which explain the micro behavior of the individuals, no restrictive rank conditions are assumed. We introduce several estimators for certain linear transformations of the systematic part of the corresponding macro relations and discuss their properties. [This research includes some joint work with Fikri Akdeniz (Cukurova University, Adana, Turkey).]

Le problème d'agrégation linéaire est considéré dans le modèle général de Gauss-Markov avec possibilité de singularités. Pour les micro-relations réelles sous-jacentes, qui expliquent le micro-comportement des individus, on ne suppose aucune condition restrictive sur le rang. Plusieurs estimateurs sont présentés pour certaines transformations linéaires de la partie systématique des macro-relations correspondantes et leur propriétés sont discutées. [Fikri Akdeniz, Université Cukurova, Adana, Turquie, a collaboré à cette recherche.]

[MS-192]**Delineating the Effect of Mis-specification in Partially Linear Models
Effets d'une mauvaise spécification des modèles partiellement linéaires**

S. Ejaz AHMED, *University of Windsor*, Kjell DOKSUM, *University of Wisconsin*, Shakhawat HOSSAIN, *University of Windsor*, Jinghong YOU, *University of North Carolina*

A partially linear model estimation problem is considered when the vector of coefficients in the linear part can be partitioned. In this situation, inference about restricted model may benefit from moving the estimate for the full model in the direction of the estimate without the nuisance variables (Steinian shrinkage). The relative performance of the estimators are examined. We also consider a LASSO type estimator for partially linear models and give a Monte Carlo simulation comparison of shrinkage, pretest and LASSO estimators. The comparison shows that shrinkage method performs better than LASSO when the number of restriction on parameter space is large.

Un problème d'estimation d'un modèle partiellement linéaire est considéré lorsque le vecteur des coefficients de la composante linéaire peut être partitionné. Dans cette situation, l'inférence sur le modèle restreint peut bénéficier d'un déplacement de l'estimé pour le modèle complet dans la direction de l'estimé sans les variables de nuisance (rétrécissement de type Stein). La performance relative des estimateurs est examinée. Nous considérons également un estimateur de type LASSO pour modèles partiellement linéaires et donnons une comparaison par Monte-Carlo des estimateurs par rétrécissement et de type prétest et LASSO. La comparaison montre que la méthode par rétrécissement performe mieux que LASSO lorsque le nombre de restrictions sur l'espace paramétrique est grand.

[MS-193]**An Illustrated Philatelic Introduction to “Matrices and Statistics”
Une introduction philatélique illustrée aux “matrices et à la statistique”**

George STYAN, *McGill University*, Simo PUNTANEN, *University of Tampere, Finland*

In this talk we identify scholars who have contributed to both matrices and statistics, and who in addition have been honoured with a postage stamp. While statistics has a very long history, matrices and postage stamps are less than 200 years old: the first postage stamp, the “Penny Black”, was issued by Great Britain in May 1840, and Arthur Cayley (1821–1895) introduced matrices in 1857. The scholars we have identified include: Giovanni Domenico Cassini (1625–1712), Pafnuty Lvovich Chebyshev (1821–1894), Leonhard Euler (1707–1783), Carl Friedrich Gauss (1777–1855), Pierre Simon Laplace (1749–1827), and John William Strutt, 3rd Baron Rayleigh (1842–1919).

Dans cette présentation nous identifions les érudits qui ont contribué à la fois aux matrices et à la statistique, et qui de plus ont été honorés par un timbre poste. Alors que la statistique a une longue histoire, les matrices et les timbres poste ont moins de 200 ans : le premier timbre poste, le “Penny Black”, fut émis par la Grande Bretagne en mai 1840, et Arthur Cayley (1821–1895) a introduit les matrices en 1857. Les savants que nous avons identifiés incluent : Giovanni Domenico Cassini (1625–1712), Pafnuty Lvovich Chebyshev (1821–1894), Leonhard Euler (1707–1783), Carl Friedrich Gauss (1777–1855), Pierre Simon Laplace (1749–1827), et John William Strutt, 3e Baron Rayleigh (1842–1919).

Session 12D Mercredi 13 juin • Wednesday, June 13, 15:00–16:30 SN-2109**Robust Statistics****Statistiques robustes****[MS-194]****Bivariate Quantile Regression via Conditional Halfspace Depth Contours****Régression des quantiles bivariés au moyen de contours de profondeur de demi-espaces conditionnels**Ivan MIZERA & Linglong KONG, *University of Alberta*

For bivariate data, the contours of the location half-space (Tukey) depth provide a possible extension of univariate quantiles that enjoys several appealing intuitive and formal properties, like interpretability, affine equivariance, and robustness. However, it may not be that clear how to obtain such contours in a nonparametric regression setting, conditionally on covariates – a question arising in the context of the construction of bivariate biometric growth charts. We approach this problem through a characterization of depth contours as the envelope of directional quantile regression lines, and discuss nonparametric regression approaches capable of carrying this strategy.

Pour des données bivariées, les courbes de niveaux de la profondeur demi-espace (de Tukey) de position sont une généralisation possible des quantiles univariés bénéficiant de plusieurs propriétés formelles et intuitivement désirables, telles l'interprétabilité, l'équivariance affine, et la robustesse. Cependant, l'obtention de ces courbes de niveau dans un contexte de régression non paramétrique conditionnelle aux covariables n'est pas claire ; cette question survient dans le contexte de la construction de courbes de croissance biométriques bivariées. Nous approchons ce problème au moyen d'une caractérisation des courbes de niveau pour la profondeur comme l'enveloppe de droites de régression quantile directionnelle, et discutons des approches de régression non paramétrique pouvant utiliser cette stratégie.

[MS-195]**Robustness in designs for field experiments****Robustesse en planification d'expériences sur le terrain**Julie ZHOU & Beiyan OU, *University of Victoria*

Optimal designs for field experiments are useful in planning agricultural experiments, environmental studies, etc. Optimal designs depend on spatial correlation structures of field plots and have been studied for various spatial processes. However, in practice, the correlation structures are never known exactly, so there is a need to study optimal designs which are robust against misspecification of spatial correlation structures. In this talk, various neighborhoods of a covariance matrix are introduced and discussed, and minimax design criteria are proposed. Examples are given to compare some robust designs with optimal designs.

Les plans optimaux pour expériences sur le terrain sont utiles dans la planification d'expériences agricoles, d'études environnementales, etc. Les plans optimaux dépendent des structures de corrélations spatiales des parcelles et ont été étudiés pour des processus spatiaux variés. Cependant, en pratique, les structures de corrélation ne sont jamais exactement connues ; il est donc nécessaire d'étudier les plans optimaux robustes à une mauvaise spécification des structures de corrélations spatiales. Dans cet exposé, plusieurs voisinages d'une matrice de covariance sont présentés et détaillés, et le critère de planification minimax est proposé. Des plans robustes sont comparés à des plans optimaux à l'aide d'exemples.

[MS-196]**Exploring Multivariate Data via Multiple Scatter Matrices****Exploration de données multivariées au moyen de multiples matrices de dispersion**David TYLER, *Rutgers University*

New multivariate methods based on comparing different estimates of multivariate scatter are introduced. The eigenvalue-eigenvector decomposition of one relative to another generates an affine invariant coordinate system for the multivariate data. By plotting the data with respect to this new invariant coordinate system, various data structures can be revealed. Under certain independent component analysis models, the invariant coordinates correspond to the independent components. When the data comes from a mixture of elliptical distributions, a subset of the invariant coordinates correspond to Fisher's linear discriminant subspace, even though the class identification of the data points are unknown.

De nouvelles méthodes multivariées basées sur la comparaison de différents estimés de la dispersion sont présentées. La décomposition valeurs propres-vecteurs propres de l'une relative à l'autre génère un système de coordonnées affines invariantes pour les données multivariées. En traçant le graphe des données par rapport à ce nouveau système de coordonnées invariantes, diverses structures des données peuvent être révélées. Sous certains modèles d'analyse en composantes indépendantes, les coordonnées invariantes correspondent aux composantes indépendantes. Lorsque les données proviennent d'un mélange de distributions elliptiques, un sous-ensemble des coordonnées invariantes correspond au sous-espace discriminant linéaire de Fisher, même si les données elles-mêmes n'ont pas été classifiées.

Session 12E Mercredi 13 juin • Wednesday, June 13, 15:00–16:30 AA-1046**Biostatistical Modelling****Modélisation biostatistique****[MS-197]****How much Data Does it Take to Parameterize an Epidemic Model?****Combien d'observations sont nécessaires à la paramétrisation d'un modèle d'épidémie ?**Rob DEARDON, *University of Guelph*

In a system as complicated as that defined by an infectious disease epidemic, models are necessary to understand what drives the system and, thus, how it might be controlled. However, one of the biggest problems with regards understanding infectious diseases is in data collection, which is both a difficult, and an expensive, process. Questions remain about how much, and what type of, data need to be collected to reliably inform model parameterization and conclusions drawn. Here, we use the example of the UK 2001 foot-and-mouth epidemic, along with simulated data, to discuss about such topics.

Dans un système aussi compliqué que celui défini par une épidémie d'une maladie infectieuse, des modèles sont nécessaires afin de comprendre ce qui dirige le système et, donc, comment il peut être contrôlé. Cependant, un des principaux problèmes quant à la compréhension des maladies infectieuses réside dans la collecte des données, qui est un processus à la fois difficile et dispendieux. Des questions persistent quant à savoir combien, et quel type de données sont nécessaires afin de s'assurer de la fiabilité de la paramétrisation du modèle et des conclusions tirées. Ici, nous utilisons l'exemple de l'épidémie de fièvre aphteuse de 2001 au Royaume-Uni, ainsi que des données simulées, afin de discuter de ces sujets.

[MS-198]**Sensitivity Analysis for Marginal Structural Models: Checking the Inverse Probability-of-treatment Weights
Analyse de sensibilité pour modèles structureaux marginaux : vérifier la pondération par probabilités inverses**

Eleanor PULLENAYEGUM, *McMaster University*, Cath LAM & Cedric MANLHIOT, *Hospital for Sick Children*, Brian FELDMAN, *Hospital for Sick Children and University of Toronto*

Marginal structural models (MSMs) are an increasingly popular framework in which observational data is re-weighted in order to draw causal inferences. The ability of an MSM to accomplish this rests crucially on treatment being unconfounded by covariates in the re-weighted dataset. This presentation discusses how this important property of the weights can be evaluated when both the treatment and the covariates are time-dependent. Often treatment effect estimates will be sensitive to the choice of weights. A framework within which to explore reasons for this sensitivity and to select a good weighting scheme is suggested. The methods are illustrated using a study of juvenile dermatomyositis.

Les modèles marginaux structureaux (MMSs) sont une structure de plus en plus populaire dans laquelle des données observationnelles sont repondérées afin d'effectuer de l'inférence causale. Cette capacité d'un MMS repose cruciallement sur la non confusion de l'effet du traitement et des covariables dans le jeu de données repondéré. Cette présentation discute de l'évaluation de cette propriété importante des poids lorsque les poids et le traitement sont tous deux dépendants du temps. Souvent, les estimés de l'effet traitement seront sensibles au choix des poids. Une structure dans laquelle explorer les raisons de cette sensibilité et choisir un bon arrangement des poids est suggérée. Les méthodes sont illustrées utilisant une étude sur la dermatomyosite juvénile.

[MS-199]**Latent Class Models for Pedigree Data
Modèles avec classes latentes pour des données familiales**

Aurélie LABBE, *Université Laval*

Considering clinical data arising from a collection of pedigrees, the general objective of our research is to develop methods to identify homogeneous disease sub-types based on multivariate disease measurements. Our approach is to develop statistical models describing the multivariate symptoms of subjects in families as a function of latent homogeneous disease classes. We extended latent class analysis to allow dependence between the latent disease class status of relatives within nuclear families. An EM algorithm maximizes the likelihood and a cross-validation approach selects the optimal model. An application of our approach to disease gene mapping by linkage analysis will be presented.

Considérant des données cliniques provenant d'une collection de familles, l'objectif général de notre recherche est de développer des méthodes pour identifier des sous-types homogènes de maladies. Notre approche consiste à développer un modèle statistique décrivant les symptômes multivariés de sujets dans des familles comme une fonction de classes latentes homogènes. Nous avons étendu les modèles de classes latentes existants de façon à permettre une dépendance entre les classes des individus d'une même famille. Un algorithme EM maximise la vraisemblance et une approche de validation croisée sélectionne le modèle optimal. Une application à la localisation de gènes par analyse de liaison sera présentée.

[MS-200]**Experimental Design for Qualitative Fitting
Plan d'expérience pour ajustement de modèles qualitatifs**Giles HOOKER, *Cornell University*

This talk explores experimental design for models which mimic observed data in meaningful ways, but which exhibit poor quantitative performance. This type of problem arises frequently for models described by nonlinear dynamics and is encountered in ecology and neurobiological applications. Frequently, we are interested in aspects of a system which may be defined in terms of system parameters in this miss-specified model. Moreover, design for such systems are frequently given in terms of time-varying inputs into a system. I develop a mathematical formulation and propose a pseudo-MCMC approach to designing sequential experiments. The ideas are illustrated with examples from neuro-dynamics.

Cet exposé explore les plans expérimentaux pour modèles imitant les données observées de façon significative, mais présentant une faible performance quantitative. Ce type de problème est souvent rencontré avec des modèles décrits par une dynamique non linéaire et est rencontré dans des applications en écologie et en neurobiologie. Fréquemment, nous sommes intéressés par des aspects d'un système pouvant être défini en termes de paramètres du système dans ce modèle mal spécifié. En outre, les plans pour ces modèles sont souvent donnés en termes d'intrants dans un système variant en fonction du temps. Je développe une formulation mathématique et propose une approche pseudo-MCMC pour planifier les expériences séquentielles. Ces idées sont illustrées à l'aide d'exemples en neuro-dynamique.

[MS-201]**A Two-component Regression Mixture Model with Covariate-dependent Weights for Bivariate Morphometric Data****Un mélange de régressions à deux composantes ayant des poids variables appliqué à des données morphométriques bivariées**Tobie SURETTE & Tobie SURETTE, *Department of Fisheries and Oceans*, Jacques ALLARD, *Université de Moncton*

Morphometric data from living organisms are of biological interest and are widely available in the field of fisheries. Often, relationships between morphometric variables characterize groups of organisms, such as species, sex and maturity phase. Datasets in many cases typically contain two or more unidentified components. We propose a regression mixture model with covariate-dependent weights to treat the two-component case. Using snow crab (*Chionoecetes opilio*) data, we apply the model to classify, and determine the relative growth rates according to maturity phase, as well as the size-at-sexual maturity. We compare the performance of the proposed model and a few alternative models.

Les données morphométriques sont d'intérêt biologique et sont largement disponibles dans le domaine de la pêche. Souvent, les relations entre variables morphométriques sont utilisées pour caractériser des groupes d'organismes tel que l'espèce, le sexe ou la phase de maturité. Dans certains cas, un ensemble de données morphométriques peut contenir deux composantes non-identifiées ou plus. Nous proposons un mélange de régressions à deux composantes ayant des poids qui varient en fonction d'une variable indépendante. Ce modèle est appliqué à des données de crabe des neiges (*Chionoecetes opilio*) pour obtenir ensuite la classification et la croissance relative en fonction de la phase de maturité ainsi que la taille à maturité sexuelle. Nous comparons la performance du modèle proposé avec un petit ensemble de modèles alternatifs.

[MS-202]

Modelling Multistage Sampling of Familial Data Using a Composite Likelihood Method
Modélisation de l'échantillonnage à plusieurs niveaux de données familiales en utilisant une méthode de vraisemblance composite

Yun Hee CHOI & Laurent BRIOLLAIS, *Samuel Lunenfeld Research Institute*

In genetic studies, multistage sampling permits the allocation of resources to families that are most informative for a given objective while allowing population-based inference. In this study, we make inference about the cumulative risk of breast cancer in BRCA mutation carriers (penetrance) from data collected in three family registries following a multistage design. We investigate and compare various likelihood methods to analyze these data, in particular a composite likelihood method previously proposed for multivariate survival data, both in terms of bias and efficiency of the penetrance estimator.

Dans les études génétiques, l'échantillonnage à plusieurs niveaux permet d'allouer les ressources aux familles les plus informatives pour un objectif donné tout en rendant possible une inférence à la population entière. Dans cette étude, nous faisons une inférence sur le risque cumulé de cancer du sein chez les porteurs d'une mutation de BRCA (la pénétrance) à partir de données collectées dans trois registres familiaux en suivant un plan d'échantillonnage à plusieurs niveaux. Nous étudions et comparons plusieurs méthodes de vraisemblance, en particulier une méthode de vraisemblance composite proposée auparavant pour l'analyse de données de survie multivariées, à la fois en termes de biais et d'efficacité de l'estimateur de la pénétrance.

Session 12F Mercredi 13 juin • Wednesday, June 13, 15:00–16:30 SN-2041

Time Series
Séries chronologiques

[MS-203]

Multivariate Autoregression of Order One with Infinite Variance
Autorégression multivariée d'ordre un avec variance infinie

Mohammad ROKNOSSADATI & Mahmoud ZAREPOUR, *University of Ottawa*

We consider the limiting behavior of a vector autoregressive model of order one (VAR(1)) with i.i.d. innovations vector with dependent components in the domain of attraction of a multivariate stable law with possibly different indices of stability. It is shown that in some cases the ordinary least squares (OLS) estimates are inconsistent. It is also revealed that certain M-estimates, with some regularity conditions, as an appropriate alternative, not only resolve inconsistency of the OLS estimates, but also give higher consistency rates in all cases.

Nous considérons le comportement limite d'un modèle vectoriel autorégressif d'ordre un (VAR(1)) avec vecteur d'innovations i.i.d. dont les composantes sont dépendantes et appartenant au domaine d'attraction d'une loi stable multivariée avec indices de stabilité potentiellement différents. Il est démontré que, dans certains cas, les estimateurs par moindres carrés ordinaires (MCO) ne sont pas convergents. Il est aussi montré que certains estimateurs-M, sous certaines conditions de régularité, forment une alternative appropriée en résolvant les problèmes de non-convergence des estimateurs MCO tout en atteignant de meilleurs taux de convergence dans tous les cas.

[MS-204]**Score Test for Homogeneity of Variances in Nonparametric Regression with Autocorrelated Errors via Wavelets****Un test de cotes pour l'homogénéité des variances en régression non paramétrique avec erreurs autocorrélées au moyen d'ondelettes**Nalini MAHALINGAM & Alwell OYET, *Memorial University*

We discuss Neyman's partial score test for homogeneity of variances in nonparametric models. We find that the discrete wavelet transform approach used by Cai, Hurvich and Tsai (1998) does not lead to a consistent estimate of the response function which in turn affects the score statistic. We follow the approach of Oyet and Sutradhar (2003) to examine the performance of the score test for homogeneity when the data arises from a nonparametric model with correlated errors. In the context of a time series, the data can be treated as arising from a time series model that is nonstationary in the mean and variance.

Nous discutons du test de score partiel de Neyman pour l'homogénéité des variances de modèles non paramétriques. Nous trouvons que l'approche par transformation par ondelettes discrètes utilisée par Cai, Hurvich et Tsai (1998) ne nous amène pas à un estimé convergent de la fonction réponse, ce qui affecte alors la statistique de score. Nous suivons l'approche de Oyet et Sutradhar (2003) pour examiner la performance du test de score pour l'homogénéité lorsque les données proviennent d'un modèle non paramétrique avec erreurs corrélées. Dans le contexte d'une série chronologique, les données peuvent être traitées comme provenant d'un modèle de séries chronologiques non stationnaire en moyenne et en variance.

[MS-205]**Tests of Causality Between two Multivariate Infinite-order Autoregressive Series****Test de causalité entre deux séries vectorielles autorégressives d'ordre infini**Chafik BOUHADDIoui, *United Arab Emirates University*, Jean-Marie DUFOUR, *Université de Montréal*

In this paper, we propose simple tests for non-causality at various horizons which can be implemented only through linear regression methods and do not involve the use of artificial simulations. This will be done, in particular, by introducing multiple horizon infinite-order vector autoregressions which can be approximated by a multiple horizon finite-order vector autoregressions. Under some regular conditions, we study the estimation of parameters obtained from the approximation of the infinite-order autoregression by a finite autoregression. We study the testing of non-causality at various horizons for infinite-order stationary processes. An asymptotic power of the test will be considered. Finally, a small simulation study and a real application will be presented.

Dans cet article, nous proposons des tests simples de non-causalité à différents horizons qui peuvent être mis en oeuvre par des méthodes de régression linéaires qui ne requièrent pas de méthodes artificielles de simulation. Nous introduisons les modèles autorégressifs multivariés d'ordre infini à différents horizons qui seront approximés par des modèles autorégressifs multivariés d'ordre fini à différents horizons. Nous étudions l'estimation des paramètres obtenus par cette approximation. Ainsi, nous proposons un test de non-causalité à différents horizons pour des processus autorégressifs stationnaires d'ordre infini. La puissance asymptotique du test sera aussi considérée. Une étude de simulation et une application réelle seront présentées.

[MS-206]**Robust Multitaper Spectrum Estimates
Estimateurs de spectres robustes à filtres multiples**David THOMSON, *Queen's University*

A problem in seismology is analyzing the Earth's "hum," or continuously excited normal modes. A major statistical problem is that the amplitude is tiny, about 3 angstroms (\AA) per second; there are many outliers, about 30 magnitude-4 earthquakes per day; and one needs at least 30 days of data to resolve the frequencies. This requires much better robust spectrum estimates.

A multitaper spectrum estimate using long data segments provides the necessary frequency resolution, but is not robust. Robustness can be incorporated by using the complex demodulation form of multitaper estimates with adaptive censoring.

Un problème en séismologie est l'analyse du "bourdonnement" de la Terre, ou de ses modes normaux continuellement excités. Un problème statistique majeur est que l'amplitude est minuscule, environ 3 angströms (\AA) par seconde; il y a beaucoup de valeurs aberrantes, environ 30 tremblements de terre de magnitude 4 par jour; et au moins 30 jours de données sont nécessaires pour résoudre les fréquences. Ceci nécessite de meilleurs estimateurs de spectre robustes.

Un estimateur de spectre à filtres multiples utilisant de longs segments de données nous apporte la résolution de fréquences nécessaire, mais n'est pas robuste. La robustesse peut être incorporée en utilisant la forme de démodulation complexe des estimés à filtres multiples avec censure adaptative.

[MS-207]**Dynamic Spectrum Estimation for Nonstationary Time Series
Estimation spectrale dynamique pour séries chronologiques non stationnaires**Azadeh MOGHADERI, David J THOMSON & Glen TAKAHARA, *Queen's University*

The notion of the spectral density function of a stationary stochastic process can be extended to nonstationary processes by the dynamic spectrum. The NSQI estimator of the dynamic spectrum was proposed by Thomson (2000). We apply the NSQI estimator to a simulated nonstationary AR(2) process. Simulation results reveal that "sudden" changes in the dynamic spectrum in certain regions cannot be captured by the NSQI estimator. We introduce an NSQI-like expansion of the dynamic spectrum in Chebyshev polynomials, which gives superior results. We propose estimators of the Chebyshev expansion coefficients, and test their bias, consistency, and variance properties numerically and theoretically.

La notion de la fonction de densité spectrale d'un processus stochastique stationnaire peut être généralisée aux processus non stationnaires au moyen du spectre dynamique. L'estimateur NSQI du spectre dynamique fut proposé par Thomson (2000). Nous appliquons l'estimateur NSQI à un processus non stationnaire AR(2) simulé. Les résultats de simulation révèlent que des changements "soudains" du spectre dynamique dans certaines régions ne peuvent être retenus par l'estimateur NSQI. Nous présentons un développement du spectre dynamique en fonction des polynômes de Chebyshev, similaire au NSQI, qui donne de meilleurs résultats. Nous proposons des estimateurs des coefficients de Chebyshev dans le développement, et testons leur biais, leur convergence, et les propriétés de leur variance d'un point de vue numérique et théorique.

The translation team did its best. Please forgive us for our misinterpretations, we are neither professional translators nor specialists in all areas of statistics.

L'équipe de traduction a fait de son mieux. Excusez-nous pour les erreurs de traduction, nous ne sommes ni des traducteurs professionnels ni des spécialistes dans tous les domaines de la statistique.

Author Index • Index des auteurs

- ABARIN, Taraneh 145
 ABRAHAM, Bovas 103
 AHMED, Rashid 84
 AHMED, S. Ejaz 77, 149
 AIT AODIA, Djilali 134
 ALLARD, Jacques 153
 ALMEIDA, Elizabete 114
 AMUAH, Joseph 63
 ANGERS, Jean-François 111, 126
 ANTONIANO, Sadora 113
 ASIMIT, Jennifer 106
 ATHALE, Uma 142

 BAKSALARY, Oskar Maria 138
 BALAKRISHNAN, Narayanaswamy 61, 117
 BALASOORIYA, U. 135
 BANJEVIC, Dragan 91
 BARI, Wasimul 68
 BARRY, Simon 128
 BASHIR, Shaheena 106
 BEAUCHAMP, Marie-Eve 82
 BEAUMONT, Jean-François 99
 BENNETT, Nicola 88
 BERNIER, Jacques 111
 BEYENE, Joseph 105
 BICKIS, Mikélis 105, 124
 BIERNACKA, Joanna 128
 BIGGS, Michael 97
 BILDER, Christopher 71
 BINGHAM, Derek 67
 BORDT, Michael 136
 BOUHADDIUI, Chafik 155
 BOWLING, Michael 97
 BRESLOW, Norman 65
 BRETTSCHEIDER, Julia 106
 BRIOLLAIS, Laurent 108, 123, 154
 BRISEBOIS, François 86
 BROWN, K. Stephen 84
 BROWN, Patrick 64
 BRUALDI, Richard 138
 BULL, Shelley 106
 BULLA, Paolo 134

 BUREAU, Alexandre 152
 BURKE, Murray D. 131

 CABILIO, Paul 60, 75
 CADENILLAS, Abel 100
 CADIGAN, Noel 61, 66
 CAHILL, Carolyn 136
 CAI, Bo 121
 CAI, Jianwen 70
 CAMPBELL, Dr. Eddy 65
 CANTONI, Eva 129
 CANTY, Angelo 106
 CARLETON, Bruce 62
 CARRIERE, Kimmie Chough 141
 CARRILLO GARCIA, Ivan 72
 CARTER, Michael 89
 CHAMBERS, L 76
 CHAPMAN, Beatrice 86
 CHATTERJEE, Snigdhanu 129
 CHAUBEY, Yogendra 133
 CHEBANA, Fateh 125
 CHEN, Baojiang 146
 CHEN, Jiahua 148
 CHEN, Manhua 70
 CHEN, Song 147
 CHENG, Ji (Emmy) 75
 CHOI, Yun Hee 154
 CHOI, Yunhee 108
 CIAMPI, Antonio 112, 113
 COLE, David 63
 COOK, Richard 75, 132, 146
 COOMES, David 64
 CORDELL, Heather 128
 CRAIU, Radu 92
 CROTEAU, Jordie 152

 DARLINGTON, Gerarda 101, 102
 DAVIES, Katherine 134
 DAVIS, Karolyn 140
 DAVIS, Matthew 62, 63
 DAVISON, Matt 80
 DEAN, Charmaine 62, 70, 133

DEARDON, Rob	151	GRUBISIC, Maja	75
DELANEY, J.A. "Chris"	137	GUILLOTTE, Simon	143
DENG, Mingxia	96	GWADRY-SRIDHAR, Femida	75
DION, Sarah-Maude	119	HA, Hyung-Tae (Kiefer)	131
DIXON, Stephanie	95, 102	HALLADAY, Amanda	93
DOKSUM, Kjell	149	HAMILTON, David	63
DOLOVICH, L	76	HANLEY, James	82
DONNER, Allan	60, 101	HARPER, William	59
DOWD, Michael	114	HASAN, M. Tariqul	116
DOWDEN, Jeffrey	61	HASAN, Tariq	108
DUBIN, Joel	116	HASS, Julie	136
DUBINS, Matthew	124	HIGDON, Dave	67
DUFOUR, Jean-Marie	155	HOOKE, Giles	153
DUKE, Linnea	108	HOSSAIN, Ahmed	105
DUMAIS, Jean	118	HOSSAIN, Shakhawat	149
DYACHENKO, Alina	112	HU, Joan	74
ELIASZIW, Michael	63	HUANG, Mei-Ling	143
ELLIOTT, Robert	107	HUANG, Yingru (Celia)	60
ERLWEIN, Christina	80	HUNTER, Jeffrey	138
ESCHENBACH, Ted	59	HURD, Tom	81
FAERBER, Jennifer	137	HUSTON, Carolyn	94
FARRELL, Patrick	69	HYNDMAN, Cody	100
FAVRE, Anne-Catherine	126	IVANOFF, Gail	139
FELDMAN, Brian	152	JACOBSON, Larry	66
FELDPAUSCH, Roxanne	99	JARDINE, Andrew	91
FERLAND, Michel	119	JIANG, Depeng	83
FIELD, Chris	98	JIANG, Yannan	97
FLEMMING, Joanna	102	JOFFE, Marshall	137
FLETCHER, David	62	KACZOROWSKI, J	76
FORTIN, Vincent	111	KANG, Sangwook	70
FOURDRINIER, Dominique	122	KANG, Sohee	123
FOX, Carlotta	139	KHAN, Bashir	76
FRANKLIN, Sarah	86	KHOURY, Imad	113
GADAG, Veeresh	135	KIM, Jae-kwang	147
GALLOP, Robert	137	KONG, Linglong	123, 150
GAMBINO, Jack	88	KOPCIUK, Karen	108
GARLIPP, Tim	110	KOURITZIN, Michael	120
GAUDET, Jocelyn	112	KOVACEVIC, Milorad	69
GHODSI, Ali	97, 144	KOVAL, John	59, 60
GILL, Paramjit	62	KUMAR, J	135
GILLIS, Daniel	95	KUNDU, Debasis	61
GIRARD, Julie	86	KUSTRA, Rafal	113, 114
GIROUX, Suzelle	119	LABBE, Aurélie	152
GODBOUT, Serge	71, 130	LABRECQUE-SYNNOTT, Félix	126
GOU, Tingting	135	LAFLAMME, Guy	88
GRAHAM, Jinko	108	LAFONTAINE, Amanda	126
GROVER, Vaneeta	63		

LAGAKOS, Stephen	74	MIZERA, Ivan	123, 150
LAHIRI, Parthasarathi	129	MOGHADERI, Azadeh	156
LAM, Cath	152	MOHN, Bob	66
LANGLET, Éric	118	MONETTE, Georges	79
LAROCQUE, Denis	82	MOODIE, Erica	136
LAWLESS, Jerry	91	MORIN, Yves	130
LAWSON, Andrew	121	MORRISON, Rebecca	119
LEBLANC, Alexandre	97	MUELLER, Christine	110
LEE, Alan	97	MULIERE, Pietro	134
LEE, Chun-In	140	MULRY, Mary	99
LEE, Mei-Ling Ting	91	MURDOCH, Duncan	135
LEFEBVRE, Geneviève	137	MURIE, Carl	113
LEGAULT, Chris	66	MURPHY, Brendan	145
LELE, Subhash	130	MYERS, Jason	104
LEMIEUX, Christiane	92	NADON, Robert	113
LEON, Carlos	85	NAIR, Vijay	67
LEUENG, Denis	147	NATHOO, Farouk	70
LI, Jia	109	NAVARRO, Alfredo	89
LIANG, Yuanyuan	141	NEUFELD, Eric	124
LIN, Devon	143	NEUHAUS, John	84
LIN, Huazhen	146	NEVALAINEN, Jaakko	82
LINKLETTER, Crystal	68, 113	NEWTON, Michael	109
LOU, Wendy	90	NGUYEN, Linh	95
LOUGHIN, Thomas	71	NGUYEN, Paul	95
LOW, C.K.	135	NICOLAE, Dan	127
LU, Wilson	115, 143	NIELSEN, Jason	133
LU, Xuewen	131	ODUEYUNGBO, Adefowope	117
MA, Jinhui	76	OJA, Hannu	82
MA, Xin	64	OROURKE, K	76
MACH, Lenka	120	Ou, Beiyuan	150
MADAN, Manish	116	OUARDA, Taha	125
MAHALINGAM, Nalini	155	OWEN, Art	92
MALLICK, Taslim	69	OYET, Alwell	155
MAMON, Rogemar	80	PANTOJA-GALICIA, Norberto	72
MANLHIOT, Cedric	152	PARENT, Éric	111
MARCH, Mary	87	PARPIA, Sameer	60
MARCHAND, Éric	78, 96, 122	PAYANDEH, Amir	78
MARKLE-REID, Maureen	117	PENG, Jianan	140
McCULLOCH, Charles	84	PENG, Yingwei	83
McNICHOLAS, Paul	145	PEPLER, Debra	83
MEISEN, Dr. Axel	65	PERREAULT, Luc	112
MÉRETTE, Chantal	152	PERRON, François	143
MERLEAU, James	112	PETTAPIECE, Rob	120
MERZBACH, Ely	139	PHILLIPS, Owen	73
MESFIOUI, Mhamed	144	PICKA, Jeffrey	96
MILLAR, Anne Michele	94	PLANTE, Mathieu	139
MILLS-FLEMMING, Joanna	114, 129	PLATT, Robert	82, 137
MITSAKAKIS, Nicholas	114		

POLLANEN, Marco	143	SUN, Jie	132
PRASAD, Narasimha	130	SUN, Lei	127
PROSSER, Bob	62	SUN, Limei	85
PROVOST, Serge	78	SURETTE, Tobie	153
PULLENAYEGUM, Eleanor	152	SUTRADHAR, Brajendra	68, 69, 116
PUNTANEN, Simo	138, 149	SUTRADHAR, Rinku	75
		SWISHCHUK, Anatoliy	101
QIN, Jing	147		
QIU, Peihua	110	TAKAHARA, Glen	156
QUESSY, Jean-François	144	TAM, Mike	73
		TAMBAY, Jean-Louis	88
RAMSAY, James	139	TENHAVE, Thomas	137
RAO, J.N.K.	130	THABANE, Lehana	74–76, 117, 142
REES, Mark	64	THIAGARAJAH, K Ranee	125
REESOR, Mark	80	THOMAS, Steven	87
RICHARDS, Evelyn	104	THOMPSON, Mary	72
RIVEST, Louis-Paul	99, 119	THOMSON, David	156
ROGERS-STEWART, Katrina	69	THOMSON, David J	156
ROKNOSSADATI, Mohammad	154	TOLUSSO, David	132
ROY, Laurent	85	TONG, Xingwei	70
RUBIN BLEUER, Susana	130	TOUPIN, Marie-Hélène	144
RUSSO, Sabrina	64	TOUSSAINT, Godfried	113
		TRIBBLE, Seth	92
SAÏDI, Abdelnasser	120	TYLER, David	151
SARAFIN, Cathlin	87	TYMSTRA, Cordy	103
SCHUURMANS, Dale	144		
SCHWARZ, Carl	94	VAUGHAN, David	96
SCHWEIGERT, Jake	94	VILLEMURE, Serge	78
SCOTT, Alastair	97	VOLODIN, Andrei	77
SEAVER, Al	66		
SHELTON, Peter	81	WANG, Feng-Yu	121
SHI, Xiaofei	113	WANG, Liqun	145
SINGH, Avi C.	73	WANG, Qinggang	59
SITTER, Randy R.	115, 143	WANG, Suojin	148
SIU, Tak Kuen	80, 107	WELSH, Alan	128, 129
SLIVITZKY, Michel	126	WERNER, Hans Joachim	148
SMITH, Robert	136	WHEELER, John	79, 93
SNEDDON, Gary	116	WHITEHEAD, Tyson	80
SOBOLEV, Boris	90	WHITLOCK, Richard	75
SONG, Peter X.-K.	146	WHITMORE, George	91
SOTEROS, Chris	105	WILD, Chris	97
SOTOMAYOR, Luz	100	WILKINSON, Dana	97
SOUTHEY, Finnegan	144	WILLAN, Andrew	105
STEELE, Russell	112	WILLIAMS, Brian	67
STEPHENS, David A	136	WINGATE, David	97
STRATYCHUK, Lori	118	WINSTON, Klass	126
STRAWDERMAN, William	122	WOOLFORD, Douglas	104
STYAN, George P.H.	138, 149	WU, Changbao	72, 148
SUN, (Tony) Jianguo	70	WU, Long Yang	141
SUN, Jack	75		
		XIE, Hansheng	86

XIE, Qihao	61
XIONG, Weiwei	142
XIONG, Xiaoqin	116
YANG, Hailiang	107
YE, Cheng Lin	74
YI, Grace	146
YI, Yanqing	142
YIN, Yaling	105
YOU, Jinghong	149
YU, Min	126
YUAN, Yan	141
YUEN, Wai Kong	143
ZAREPOUR, Mahmoud	154
ZGHOUL, Ahmad	77
ZHANG, Jiajia	83
ZHANG, Ying	60, 75
ZHAO, Lihui	74
ZHAO, Xingqiu	117
ZHOU, Julie	150
ZHOU, Qian	146
ZHOU, Xiaowen	135

Author Index (per session)

Index des auteurs (par session)

ABARIN, Taraneh	12A	BRUALDI, Richard	10C
ABRAHAM, Bovas	06D	BULL, Shelley	06F
AHMED, Rashid	03E	BULLA, Paolo	09F
AHMED, S. Ejaz	02G, 12C	BUREAU, Alexandre	12E
AIT AODIA, Djilali	09F	BURKE, Murray D.	09E
ALLARD, Jacques	12E		
ALMEIDA, Elizabete	07E	CABILIO, Paul	0, 02F
AMUAH, Joseph	0	CADENILLAS, Abel	06B
ANGERS, Jean-François	07D, 08F	CADIGAN, Noel	0, 02A
ANTONIANO, Sadora	07E	CAHILL, Carolyn	10A
ASIMIT, Jennifer	06F	CAI, Bo	08C
ATHALE, Uma	10E	CAI, Jianwen	02D
		CAMPBELL, Dr. Eddy	01A
BAKSALARY, Oskar Maria	10C	CANTONI, Eva	09C
BALAKRISHNAN, Narayanaswamy	0, 07F	CANTY, Angelo	06F
BALASOORIYA, U.	09F	CARLETON, Bruce	0
BANJEVIC, Dragan	04C	CARRIERE, Kimmie Chough	10E
BARI, Wasimul	02C	CARRILLO GARCIA, Ivan	02E
BARRY, Simon	09C	CARTER, Michael	04B
BASHIR, Shaheena	06F	CHAMBERS, L	02F
BEAUCHAMP, Marie-Eve	03E	CHAPMAN, Beatrice	03F
BEAUMONT, Jean-François	06A	CHATTERJEE, Snigdhasu	09D
BENNETT, Nicola	04A	CHAUBEY, Yogendra	09E
BERNIER, Jacques	07D	CHEBANA, Fateh	08F
BEYENE, Joseph	06F	CHEN, Baojiang	04E, 12A
BICKIS, Mikelis	06F, 08E	CHEN, Jiahua	12B
BIERNACKA, Joanna	09B	CHEN, Manhua	02D
BIGGS, Michael	04G	CHEN, Shun-Fu	04E
BILDER, Christopher	02E	CHEN, Song	12B
BINGHAM, Derek	02B	CHENG, Ji (Emmy)	02F
BORDT, Michael	10A	CHOI, Yun Hee	12E
BOUHADDIUI, Chafik	12F	CHOI, Yunhee	07B
BOWLING, Michael	04G	CIAMPI, Antonio	07E
BRESLOW, Norman	01B	COLE, David	0
BRETTSCHNEIDER, Julia	06F	COOK, Richard	02F, 09E, 12A
BRIOLLAIS, Laurent	07B, 08E, 12E	COOMES, David	0
BRISEBOIS, François	03F	CORDELL, Heather	09B
BROWN, K. Stephen	03E	CRAIU, Radu	04D
BROWN, Patrick	0	CROTEAU, Jordie	12E

DARLINGTON, Gerarda	06C	GILL, Paramjit	0
DAVIES, Katherine	09F	GILLIS, Daniel	04F
DAVIS, Karolyn	10E	GIRARD, Julie	03F
DAVIS, Matthew	0	GIROUX, Suzelle	08A
DAVISON, Matt	03C	GODBOUT, Serge	02E, 09D
DEAN, Charmaine	0, 02D, 09E	GOU, Tingting	09F
DEARDON, Rob	12E	GRAHAM, Jinko	07B
DELANEY, J.A. "Chris"	10B	GROVER, Vaneeta	0
DENG, Mingxia	04G	GRUBISIC, Maja	02F
DION, Sarah-Maude	08A	GUILLOTTE, Simon	10F
DIXON, Stephanie	04F, 06C	GWADRY-SRIDHAR, Femida	02F
DOKSUM, Kjell	12C		
DOLOVICH, L	02F	HA, Hyung-Tae (Kiefer)	09E
DONNER, Allan	0, 06C	HALLADAY, Amanda	04F
DOWD, Michael	07E	HAMILTON, David	0
DOWDEN, Jeffrey	0	HANLEY, James	03E
DUBIN, Joel	07F	HARPER, William	0
DUBINS, Matthew	08E	HASAN, M. Tariqul	07F
DUFOUR, Jean-Marie	12F	HASAN, Tariq	07B
DUKE, Linnea	07B	HASS, Julie	10A
DUMAIS, Jean	08A	HE, George	03B
DYACHENKO, Alina	07E	HE, Zhulin	03B
		HIGDON, Dave	02B
ELIASZIW, Michael	0	HOOKE, Giles	12E
ELLIOTT, Robert	07A	HOSSAIN, Ahmed	06F
ERLWEIN, Christina	03C	HOSSAIN, Shakhawat	12C
ESCHENBACH, Ted	0	HU, Joan	02F
		HUANG, Mei-Ling	10F
FAERBER, Jennifer	10B	HUANG, Yingru (Celia)	0
FAN, Chun-Po Steve	03B	HUNTER, Jeffrey	10C
FARRELL, Patrick	02C	HURD, Tom	03C
FAVRE, Anne-Catherine	08F	HUSTON, Carolyn	04F
FELDMAN, Brian	12E	HYNDMAN, Cody	06B
FELDPAUSCH, Roxanne	06A		
FERLAND, Michel	08A	IVANOFF, Gail	10D
FIELD, Chris	05A		
FILL, Oksanna	04E	JACKS, Lindsay	04E
FLEMMING, Joanna	06C	JACOBSON, Larry	02A
FLETCHER, David	0	JARDINE, Andrew	04C
FORTIN, Vincent	07D	JIANG, Depeng	03E
FOURDRINIER, Dominique	08D	JIANG, Yannan	04G
FOX, Carlotta	10D	JOFFE, Marshall	10B
FRANKLIN, Sarah	03F		
		KACZOROWSKI, J	02F
GADAG, Veeresh	09F	KANG, Sangwook	02D
GALLOP, Robert	10B	KANG, Sohee	08E
GAMBINO, Jack	04A	KHAN, Bashir	02G
GARLIPP, Tim	07C	KHAN, Shahedul Ahsan	03B
GAUDET, Jocelyn	07D	KHOURY, Imad	07E
GHODSI, Ali	04G, 10F	KIM, Jae-kwang	12B

KONG, Linglong	08E, 12D	MANLHIOT, Cedric	12E
KOPCIUK, Karen	07B	MARCH, Mary	03F
KOURITZIN, Michael	08B	MARCHAND, Éric	02G, 04G, 08D
KOVACEVIC, Milorad	02C	MARKLE-REID, Maureen	07F
KOVAL, John	0	McCULLOCH, Charles	03E
KUMAR, J	09F	McNICHOLAS, Paul	12A
KUNDU, Debasis	0	MEISEN, Dr. Axel	01A
KUSTRA, Rafal	07E	MÉRETTE, Chantal	12E
LABBE, Aurélie	12E	MERLEAU, James	07D
LABRECQUE-SYNNOTT, Félix	08F	MERZBACH, Ely	10D
LAFLAMME, Guy	04A	MESFIOUI, Mhamed	10F
LAFONTAINE, Amanda	08F	MILLAR, Anne Michele	04F
LAGAKOS, Stephen	02F	MILLS-FLEMMING, Joanna	07E, 09C
LAHIRI, Parthasarathi	09D	MITSAKAKIS, Nicholas	07E
LAM, Cath	12E	MIZERA, Ivan	08E, 12D
LANGLET, Éric	08A	MOGHTADERI, Azadeh	12F
LAROCQUE, Denis	03E	MOHN, Bob	02A
LAWLESS, Jerry	04C	MONETTE, Georges	03B
LAWSON, Andrew	08C	MOODIE, Erica	10B
LEBLANC, Alexandre	04G	MORIN, Yves	09D
LEBOVIC, Gerald	03B	MORRISON, Rebecca	08A
LEE, Alan	04G	MUELLER, Christine	07C
LEE, Chun-In	10E	MULIERE, Pietro	09F
LEE, Mei-Ling Ting	04C	MULRY, Mary	06A
LEFEBVRE, Geneviève	10B	MURDOCH, Duncan	09F
LEGAULT, Chris	02A	MURIE, Carl	07E
LELE, Subhash	09D	MURPHY, Brendan	12A
LEMIEUX, Christiane	04D	MYERS, Jason	06E
LEON, Carlos	03F	NADON, Robert	07E
LEUENG, Denis	12B	NAIR, Vijay	02B
LI, Jia	07C	NATHOO, Farouk	02D
LI, Xiao Bo (Alice)	04E	NAVARRO, Alfredo	04A
LIANG, Yuanyuan	10E	NEUFELD, Eric	08E
LIN, Devon	10F	NEUHAUS, John	03E
LIN, Huazhen	12A	NEVALAINEN, Jaakko	03E
LINKLETTER, Crystal	02B, 07E	NEWTON, Michael	07B
Lou, Wendy	04B	NGUYEN, Linh	04F
LOUGHIN, Thomas	02E	NGUYEN, Paul	04F
Low, C.K.	09F	NICOLAE, Dan	09B
LU, Wilson	07E, 10F	NIELSEN, Jason	09E
LU, Xuewen	09E	ODUEYUNGBO, Adefowope	07F
MA, Jinhui	02F	OJA, Hannu	03E
MA, Xin	0	OROURKE, K	02F
MACH, Lenka	08A	Ou, Beiyang	12D
MADAN, Manish	07F	OUARDA, Taha	08F
MAHALINGAM, Nalini	12F	OWEN, Art	04D
MALLICK, Taslim	02C	OYET, Alwell	12F
MAMON, Rogemar	03C		

PANTOJA-GALICIA, Norberto	02E	SIU, Tak Kuen	03C, 07A
PARENT, Éric	07D	SLIVITZKY, Michel	08F
PARPIA, Sameer	0	SMITH, Robert	10A
PAYANDEH, Amir	02G	SNEDDON, Gary	07F
PENG, Jianan	10E	SOBOLEV, Boris	04B
PENG, Yingwei	03E	SONG, Peter X.-K.	12A
PEPLER, Debra	03E	Soo, Andrea	03B
PERREAULT, Luc	07D	SOTEROS, Chris	06F
PERRON, François	10F	SOTOMAYOR, Luz	06B
PETTAPIECE, Rob	08A	SOUTHEY, Finnegan	10F
PHILLIPS, Owen	02E	STEELE, Russell	07E
PICKA, Jeffrey	04G	STEPHENS, David A	10B
PLANTE, Mathieu	10D	STRATYCHUK, Lori	08A
PLATT, Robert	03E, 10B	STRAWDERMAN, William	08D
POLLANEN, Marco	10F	STYAN, George P.H.	10C, 12C
PRASAD, Narasimha	09D	SUN, (Tony) Jianguo	02D
PROSSER, Bob	0	SUN, Jack	02F
PROVOST, Serge	02G	SUN, Jie	09E
PULLENAYEGUM, Eleanor	12E	SUN, Lei	09B
PUNTANEN, Simo	10C, 12C	SUN, Limei	03F
QI, Ying	03B	SURETTE, Tobie	12E
QIN, Jing	12B	SUTRADHAR, Brajendra	02C, 07F
QIU, Peihua	07C	SUTRADHAR, Rinku	02F
QUESSY, Jean-François	10F	SWISHCHUK, Anatoliy	06B
RAMSAY, James	10D	TAKAHARA, Glen	12F
RAO, J.N.K.	09D	TALEBAN, Julia	04E
REES, Mark	0	TAM, Mike	02E
REESOR, Mark	03C	TAMBAY, Jean-Louis	04A
RICHARDS, Evelyn	06E	TENHAVE, Thomas	10B
RIVEST, Louis-Paul	06A, 08A	THABANE, Lehana	02F, 07F, 10E
ROGERS-STEWART, Katrina	02C	THIAGARAJAH, K Ranee	08F
ROKNOSSADATI, Mohammad	12F	THOMAS, Steven	03F
ROTONDI, Michael Anthony	04E	THOMPSON, Mary	02E
ROY, Laurent	03F	THOMSON, David	12F
RUBIN BLEUER, Susana	09D	THOMSON, David J	12F
RUSSO, Sabrina	0	TOLUSSO, David	09E
SAÏDI, Abdelnasser	08A	TONG, Xingwei	02D
SARAFIN, Cathlin	03F	TOUPIN, Marie-Hélène	10F
SCHUURMANS, Dale	10F	TOUSSAINT, Godfried	07E
SCHWARZ, Carl	04F	TRIBBLE, Seth	04D
SCHWEIGERT, Jake	04F	TYLER, David	12D
SCOTT, Alastair	04G	TYMSTRA, Cordy	06E
SEAVER, Al	02A	VAUGHAN, David	04G
SHELTON, Peter	03D	VILLEMURE, Serge	CRSNG
SHI, Xiaofei	07E	VOLODIN, Andrei	02G
SINGH, Avi C.	02E	WANG, Feng-Yu	08B
SITTER, Randy R.	07E, 10F	WANG, Hao	04E

WANG, Hui	04E
WANG, Liqun	12A
WANG, Qinggang	0
WANG, Suojin	12B
WELSH, Alan	09C
WERNER, Hans Joachim	12C
WHEELER, John	03B, 04E
WHITEHEAD, Tyson	03C
WHITLOCK, Richard	02F
WHITMORE, George	04C
WILD, Chris	04G
WILKINSON, Dana	04G
WILLAN, Andrew	06F
WILLIAMS, Brian	02B
WINGATE, David	04G
WINSTON, Klass	08F
WONG, Pik Tan	03B
WOOLFORD, Douglas	06E
WU, Changbao	02E, 12B
WU, Long Yang	10E
XIE, Hansheng	03F
XIE, Qihao	0
XIONG, Weiwei	10E
XIONG, Xiaoqin	04E, 07F
YANG, Hailiang	07A
YE, Cheng Lin	02F, 04E
YI, Grace	12A
YI, Yanqing	10E
YIN, Yaling	06F
YOU, Jinghong	12C
YU, Min	08F
YUAN, Yan	04E, 10E
YUEN, Wai Kong	10F
ZAREPOUR, Mahmoud	12F
ZGHOUL, Ahmad	02G
ZHANG, Jiajia	03E
ZHANG, Ying	0, 02F
ZHAO, Lihui	02F
ZHAO, Xingqiu	07F
ZHOU, Julie	12D
ZHOU, Qian	12A
ZHOU, Xiaowen	09F



Save 20% on all SAS® Publishing products while at the conference!

Visit the SAS® Publishing Booth to learn more about...



New Books and New Upcoming Titles

SAS® Press – Expert tips and tools for SAS professionals

SAS® Documentation – Convenient access via the Web, help, or hard copy

SAS® Learning Edition – Innovative and comprehensive learning version of SAS®

SAS Publishing can help you enhance your SAS experience with a wide range of resources, from publications to software to training. Stop by the booth, place your order and receive a 20% discount on your purchase. SAS staff will be on hand to answer any questions and to assist you with your order.

**THE
POWER
TO KNOW®**

Enjoy the Conference!

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and all other countries. ® indicates USA registration. Copyright © 2007. All rights reserved. 427761US.0107



Student learning begins in the classroom. What happens outside of class makes the difference!

With MyStatLab, students can assess their understanding of principles, develop a personalized study plan to work on areas of weakness, and master skills through limitless practice with a variety of learning tools.

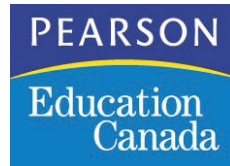
MyStatLab courses include: tutorial exercises with guided solutions and sample problems, Java applets, statistical software and other resources designed specifically to help students succeed in statistics.

To learn more, go to: www.mystatlab.com

Visit our online catalogue for information on other Pearson products!
www.pearsoned.ca/highered



THANK YOU FOR YOUR SUPPORT



MERCI POUR VOTRE SOUTIEN