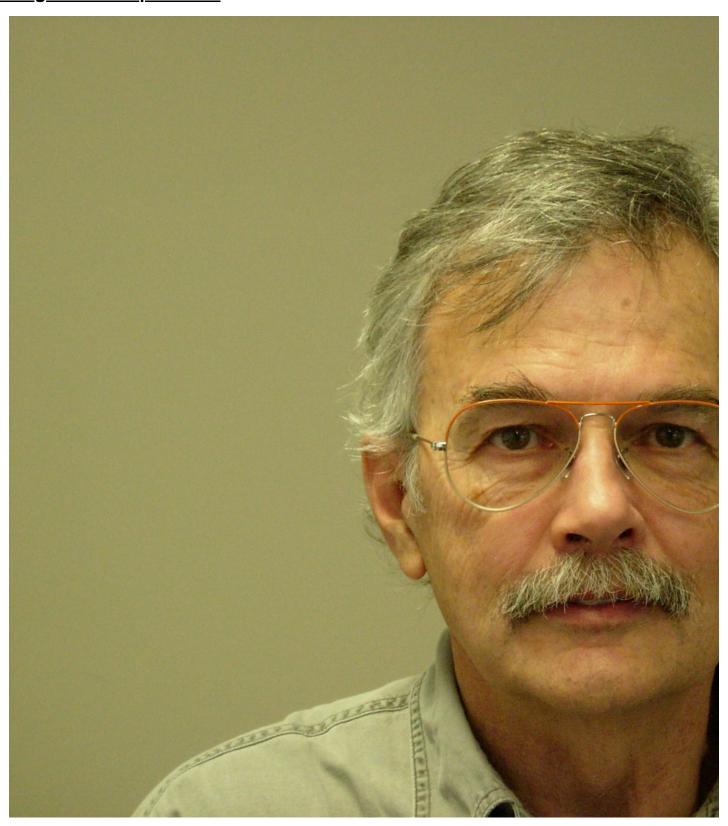
Message from the president



A year has come and gone, and my term as president of the Society will soon conclude.

The last several months have flown by, as we have all been dramatically affected by the Covid-19 crisis, juggling activities and schedules to stay afloat. My sincerest hope is that you and your families are well, and coping with the situation as best as possible.

It was very unfortunate that we had to cancel our 2020 annual meeting, which was to have been held at Carleton University. A number of you contacted me to express your support in my having had to make the difficult decision to cancel. I appreciated your thoughts, but to be honest, this was among the easiest decisions I have ever made.

I was terribly saddened, however, that hard work of the local organizing committee, led by **Shirley Mills** and **Yiqiang Zhao**, the program committee, led by **J. C. Loredo-Osti**, our executive assistant **Michele Benoit**, and our administrative assistant **Marie-Pierre Nantel**, was for naught. These individuals, and a host of others, put in an

incredible amount of work in organizing what would have been a spectacular meeting, and I give my sincerest thanks to all involved.

All was not lost. There were a number of conference activities which moved online.

We had two superb online presentations—the <u>presidential invited address</u> by **David Dohono**, and the <u>Pierre Robillard Award</u> address by **Shixio Zhang**. These talks were recorded and made available for later viewing.

A number of other conference related activities were moved in entirety to an online format:

- The Canadian Statistics Student Conference, always a feature of our annual meeting, is organized entirely by students—this year led by co-chairs **Steve Ferreira Guerra** and **Thai-Son Tang**.
- The Case Studies in Data Analysis Competition, organized by Pingzhao Hu.
- A number of workshops, including those organized by the Biostatistics Section, the Section on Business and Industrial Statistics, the Section on Data Science and Analytics, and a forthcoming workshop of the Survey Methods Section.

Details on all of the 2020 online conference events can be found here.

I would like to thank all of the individuals involved in organizing and contributing to these activities. You have given the SSC an online presence—one that we can build on—a silver lining to an otherwise very dark cloud.

A highlight of our yearly conference is the presentation of awards at our annual meeting banquet. The 2020 award presentations will be made together with the 2021 award presentations at the 2021 annual meeting, to be held at Memorial University of Newfoundland, and the 2020 award talks, apart from the Robillard address which was streamed, will be included in the 2021 program.

July 1 is a time of transition. I would like to thank all outgoing board and committee members, and welcome those of you who will be taking up these many roles. Without your volunteer efforts, we would cease to function as a Society.

Special thanks to **Changbao Wu**, who is completing his term as meetings coordinator, and welcome to **Nadia Ghazzali**, who joins us in that role. A personal thank you to **Robert Platt**, who finishes his term as past president, and who has given me such able guidance during the past year, and a very special thank you to our executive secretary, **Liwellyn Armstrong**, who keeps not only the president, but all other executive and board members, informed and on track.

Welcome to Grace Yi, who joins us as president elect, and to Wendy Lou, who transitions to president. We are in extremely capable hands.

I look forward to seeing you at SSC 2021 in St. John's.

Bruce Smith

Results of the SSC 2020 Case Studies in Data Analysis Competition

The Case Studies in Data Analysis Poster Competition was held online on June 1, 2020. This year we have designed two case studies:

Case Study 1: Predicting hourly electricity demand in Ontario

Case Study 2: Predicting podcast popularity in iTunes

There were seven teams participated in the Case Study 1, and four teams participated in the Case Study 2. In total, there were 31 students from nine universities who have participated in the competition. We have invited eight judges to evaluate all the teams' work. Each of the teams' posters was carefully assessed and evaluated by four judges during May 26–May 31. A consensus meeting of the judges was held to make the final ranking of the teams' work on June 1, 2020. The award committee selected two winning teams for each case study. One honourable mention was also selected for the Case Study 1. Each of the winning teams will receive an award with a value of \$750 from SSC. The Canada Energy Regulator has kindly supported one of the awards for the Case Study 1. Every student in the competition will receive a certificate of participation.

The competition results are as follows:

Case Study 1

Judges

Dr. Jose Ribas Fernandes, Canada Energy Regulator

Dr. Ehsan Karim, University of British Columbia

Dr. Nasri Bouchra, University of Montréal

Carolyn Augusta, University of Saskatchewan

Winner 1:

Poster Title: Smoothers, harmonic autoregression, and gradient boosting for seasonal time series: Predicting hourly electricity demand in Ontario

University: University of Alberta

Team Members: Xilai Fu, Danyi Liu, Ree Yang, Yijia Yu

Mentor: Matthew Pietrosanu



From left to right: Xilai Fu, Danyi Liu, Ree Yang, Yijia Yu.

Winner 2 (The Canada Energy Regulator (CER) Award):

Poster Title: On-line Bayesian prediction of hourly electricity demand for the residential sector in Ontario

University: McGill University

Team Members: Victoire Michal, Paritosh Kumar Roy, Dirk Douwes-Schultz

Mentor: Dr. Alexandra M. Schmidt





From top (left) to bottom: Dirk Douwes-Schultz, Victoire Michal, Paritosh Kumar Roy

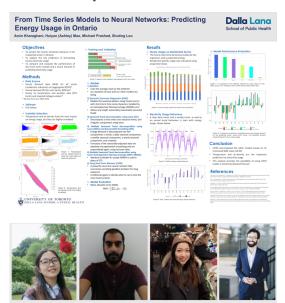
Honourable Mention:

Poster Title: From time series models to neural networks: Predicting energy usage in Ontario

University: University of Toronto

Team Members: Amin Kharaghani, Huiyan (Ashley) Mao, Michael Prashad, Shuting Lou

Mentor: Dr. Wendy Lou



Case Study 2

Judges

Dr. Kathryn Morrison, McGill University

Dr. Derek Bingham, Simon Fraser University

Dr. Xuan Li, University of Minnesota

Dr. Chel Hee Lee, University of Calgary

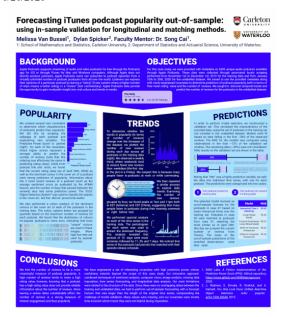
Winner 1:

Poster Title: Forecasting iTunes podcast popularity out-of-sample: Using in-sample validation for longitudinal and matching method

University: Carleton University and University of Waterloo

Team Members: Melissa Van Bussel (Carleton), Dylan Spicker (Waterloo)

Mentor: Dr. Song Cai





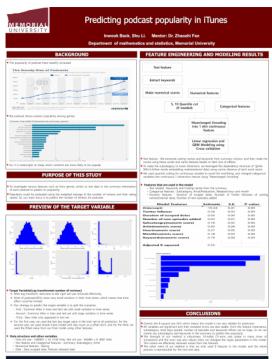
From left to right: Melissa Van Bussel, Dylan Spicker

Winner 2:

Poster Title: Predicting podcast popularity in iTunes
University: Memorial University of Newfoundland

Team Members: Inwook Back, Shu Li

Mentor: Dr. Zhaozhi Fan





The award committee would like to acknowledge all the teams and their mentors who put up great efforts in the competition! Thanks also go to **Dr. Kathryn**Morrison for her excellent work in preparing for Case Study 1, and to the team in Canada Energy Regulator (**Dr. José Ribas Fernandes**, **Dr. Ryan Hum, Mantaj**Hundal, Lukas Hansen, Michael Nadew, Matthew Hansen) for their excellent work in preparing for Case Study 2 and for providing an award to support Case Study 1.

We also thank all the judges for their hard work in making valuable decisions about selecting the winners and honourable mentions in the competition. Last but not all the least, we would also like to take this opportunity to thank Michelle Benoit, Marie-Pierre Nantel, Dr. Changbao Wu, and Dr. Angelo Canty for their efforts, help, and support in preparing for the competition.

The award committee: Dr. Ehsan Karim, Dr. Chel Hee Lee; Dr. Kathryn Morrison, and Dr. Pingzhao Hu (Chair)

Winners of The Canadian Journal of Statistics Award 2020



The Canadian Journal of Statistics Award is presented each year by the Statistical Society of Canada to the author(s) of an article published in the journal, in recognition of the outstanding quality of the methodological innovation and presentation. This year's winner is the article entitled "Doubly sparse regression incorporating graphical structure among predictors" (Volume 47, no. 4, pp. 729–747) by Matthew Stephenson, R. Ayesha Ali, and Gerarda A. Darlington.

When biological data is associated with a complex system, then it is advantageous to exploit the structure of the system to predict the biological response. The paper presents a novel approach, doubly sparse regression incorporating graphical structure (DSRIG), that first models the underlying structure of the variables in the system, and then leverages this structural information to improve prediction of the response. This model is highly flexible and has excellent predictive abilities compared to other methods, particularly when only a fraction of the variables are related to the response. It can help identify relevant predictors as well as potential predictors that may warrant further scrutiny for understanding the response. DSRIG is valid in both high-dimensional settings, where the number of variables exceeds the sample size, and in settings in which some predictors are highly correlated with one another.



Matthew Stephenson received his MSc and PhD degrees in Statistics from the University of Guelph in 2014 and 2019, respectively. His areas of research interest include statistical machine learning, regularized regression, biostatistics, and bioinformatics.



R. Ayesha Ali completed her PhD in Statistics at the University of Washington in 2002 and is now an associate professor in the Department of Mathematics and Statistics at the University of Guelph. Her research involves statistical methods for complex high dimensional systems across diverse domains. Specific applications include biostatistics, animal health, and ecology.



Gerarda A. Darlington completed her PhD in Statistics at the University of Waterloo and is now a full professor in the Department of Mathematics and Statistics at the University of Guelph. Her research interests include statistical methods for correlated observations, methods for epidemiologic studies, and the design and analysis of cluster randomized trials. She has been honoured as one of Guelph's Women of Distinction for her role as a mentor to women in STEM fields; she is also the recipient of the University of Guelph's 2018 John Bell Award in recognition of her outstanding contributions to university education.

The citation for the award reads:

"The article entitled 'Doubly sparse regression incorporating graphical structure among predictors' by Matthew Stephenson, R. Ayesha Ali, and Gerarda A. Darlington is recognized for an excellent presentation of impressive methodological development and application in machine learning."

Angelo Canty was primarily responsible for producing this material.

W. John Braun, recipient of the SSC Award for Impact of Applied and Collaborative Work 2020



The recipient of the 2020 Statistical Society of Canada Award for Impact of Applied and Collaborative Work is **W. John Braun**, professor in the Department of Computer Science, Mathematics, Physics and Statistics, University of British Columbia Okanagan. The award recognizes outstanding contributions by a member of the SSC in collaborative research and applied work, the importance of which derives primarily from its relatively recent impact on a subject area outside of the statistical sciences, on an area of application, or on an organization.

Born in Winnipeg, John Braun studied applied mathematics at the University of Calgary where he earned his BSc and MSc degrees. In 1992 he completed his PhD in Statistics at the University of Western Ontario under the supervision of **Reg Kulperger**. His research involved the study of point processes and other stochastic processes and, in his thesis, he developed numerical techniques to study parameter identifiability questions for an interacting particle system that had been proposed as a model for tumour growth.

John has held positions at a number of universities, starting with the University of Winnipeg, where he began a collaboration with a psychophysicist in developing point process models for human reaction time. He spent most of 1999 at the Australian National University, working with the late **Peter Hall** and several of his postdoes on a variety of smoothing and bootstrapping problems. In 2000 he accepted a position at Western University where he attained the rank of full professor and where for five years, he chaired the Statistics Graduate Program. While at Western University he worked with a number of collaborators both there and at the University of Toronto to develop an interacting particle system-based wildfire growth model. With some enterprising PhD students, he also studied lightning-based wildfire ignitions and storm cell evolution using point process techniques.

In 2014 he became head of the Department of Computer Science, Mathematics, Physics and Statistics at UBC Okanagan. The following year John began serving the Canadian Statistical Sciences Institute (CANSSI) as deputy director and co-leading the Rocky Mountain Data Science Centre, a joint health sciences training initiative of the University of Calgary and UBC under the sponsorship of CANSSI.

John's research in statistics has often been motivated by scientific problems coming from psychology, biology, medicine, engineering, and physics. His methodological research is concerned with smoothing and inference techniques as they apply to data visualization and process monitoring. John's research on stochastic modelling and statistical inferencing for large-scale spatial-temporal data sets has led to numerous research projects for a diverse range of applications. John collaborated with the Institute for Catastrophic Loss Reduction (ICLR), the Canadian Forest Service, Ontario Ministry of Natural Resources & Forestry and Alberta Wildfire to study wildland fire regime modelling. He has been instrumental in providing statistical approaches to wildfire modelling and management. His work on Prometheus, the wildland fire growth and spread model, has produced variants that can provide probabilistic forecasts for a particular fire in real time.

In addition to his research contributions, John has impacted the promotion and advancement of the R language for statistical computing. He authored and coauthored two monographs on using R, which are in their 2nd and 3rd editions. These materials are excellent resources for statistical education, training and research. In addition, John has submitted six R packages to CRAN which have been widely used by researchers, students, engineers, and practitioners.

The excellence of John's collaborative work was recognized by his receipt of *The Canadian Journal of Statistics* Award in 2006 and also by his students receiving a presentation award (2015) and a poster award (2009) at the annual meetings of the Statistical Society of Canada.

The citation for the award reads:

"To W. John Braun, for his outstanding contributions to advancing statistical computing technology and statistical modelling for environmental safety and wellness."

Carl Schwarz was primarily responsible for producing this material.

Leah Smith, recipient of Lise Manchester Award 2020



The 2020 Lise Manchester Award is conferred upon **Leah Smith** of the Canadian Cancer Society (CCS). This award is given every other year by the Statistical Society of Canada to commemorate the late **Dr. Lise Manchester**'s abiding interest in using statistical methods to study matters of relevance to society. The award recognizes excellence in statistical research that helps guide public policy in Canada.

Leah is currently senior manager for surveillance at the CCS, chairs the Canadian Cancer Statistics Advisory Committee and is a member of the Canadian Council of Cancer Registries and the Pan-Canadian Cervical Cancer Screening Network. She has conducted many local and national print, radio, and TV interviews.

Following her BSc in Psychology from Memorial University of Newfoundland, Leah was trained in epidemiology. She received an MSc in 2010 from Queen's University where her thesis work addressed the use and safety of the quadrivalent human papilloma virus (HPV) vaccine in Grade 8 girls, and a PhD in 2014 from McGill University for her thesis "The Impact of Human Papillomavirus Vaccination on Adolescent Health Outcomes: An Application of the Regression Discontinuity Design." Following a postdoctoral fellowship at Queen's University's Department of Public Health Sciences and Centre for Health Services and Policy she joined the Canadian Cancer Society as an epidemiologist in 2015.

The award is for her work, published in the Canadian Medical Association Journal (2015), Pediatrics (2015) and the International Journal of Epidemiology (2017) that addressed the real-world impacts and possible unintended consequences of HPV vaccination campaigns and programs. These issues cannot be answered by randomized controlled trials and nonexperimental studies using standard tools are likely to be biased. They therefore must be addressed using rigorous methods that are capable of measuring causal effects. In her first article Leah asked whether HPV immunization programs or the receipt of the vaccine itself affect sexual behaviour among teenage girls. She combined a novel and complex quasi-experimental method, regression discontinuity, with high-quality administrative health data and a sound understanding of the policy context. Under appropriate conditions this method provides unbiased estimates and is not subject to confounding in the same way that standard tools are. Leah's was the first study to plausibly estimate the causal effect of HPV vaccination on risky sexual behaviours. She found that neither the implementation of Ontario's HPV program nor receipt of the vaccine itself lead to changes in rates of pregnancy or sexually transmitted infections.

These results were and still are important for policymakers, health professionals, parents, and young adults. Informing parents and the wider public regarding the scientific evidence of the costs and benefits of vaccines for young people continues to be an important health policy challenge. Her *CMAJ* article received widespread media attention, including in the CBC, the Globe and Mail and the Wall Street Journal. Leah's second article, a clinically important and policy relevant evaluation of the effect of the Ontario HPV vaccine policy, and of the vaccine itself, on cervical dysplasia and anogenital warts in adolescent girls reached a large clinical audience and received significant media attention. Her third article, in a top epidemiology journal with a strong methodological bent, guides other researchers to responsibly and productively use the regression discontinuity design.

The citation for the award reads:

"To Leah Smith, for her influential work on the positive effects, and the lack of negative effects, of Ontario's HPV vaccination program; and for showing how the use of rigorous statistical methods, quasi-experimental settings and population-based provincial administrative data can make significant contributions to improving health and health care systems in Canada and elsewhere."

James Hanley was primarily responsible for producing this material.

Erica Moodie, recipient of the CRM-SSC Prize in Statistics 2020



The CRM-SSC Prize in Statistics is awarded annually by the Centre de recherches mathématiques (CRM) and the Statistical Society of Canada (SSC) in recognition of outstanding research carried out primarily in Canada by a statistician during the first fifteen years after completing a PhD. The 2020 recipient of this prize is **Erica**

Moodie, William Dawson scholar and professor in the Department of Epidemiology, Biostatistics, and Occupational Health at McGill University.

Erica was born and raised in Winnipeg. Her interest in science, shared by her sister Zoe who is also a biostatistician, was fostered by their parents, zoologist **Ric Moodie** and biostatistician **Patricia F. Moodie**. After studying mathematics and statistics at the University of Winnipeg (BA, 2000), Erica specialized in epidemiology at the University of Cambridge (MPhil, 2001) and in biostatistics at the University of Washington (MSc, 2004; PhD, 2006). She joined McGill as an assistant professor in 2006 and was promoted to the rank of associate professor in 2012. She was the director of the Biostatistics Graduate Programs from 2012–2019.

In her thesis, written under the supervision of **Thomas Richardson**, Erica studied inference for optimal dynamic treatment regimes (DTR). DTRs are sequential decision-making strategies that define rules for optimal allocation of resources. Over the past 15 years Erica has established herself as a world expert in this area. With two books and over 130 peer-reviewed research papers to date, she has become one of the most prominent biostatisticians of her generation. Her papers have appeared in *Biometrics*, *Biometrika*, *JASA*, *Statistics in Medicine* and many other top-tier international journals.

Through her research, Erica has made a fundamental shift in the way that DTR estimators are used and viewed. Many traditional approaches to DTR estimation are opaque. In a series of papers on this topic, she has proposed a new form of regression-based estimation which ensures interpretability and accessibility to end-users in a broad range of fields. This work, beginning with postdoctoral researcher **Michael Wallace** (now at Waterloo) and continued with several other trainees, includes a suite of model selection, diagnostic, and validation tools, which were notably absent in nearly all alternative methods. Her approach, implemented in R, has raised the bar for all subsequent DTR analyses.

Erica's work in dynamic decision-making sits within the more general area of causal inference, where interest lies in determining the un-confounded effect of specific treatment factors on the outcome. With postdoctoral fellow **Olli Saarela** (University of Toronto) and McGill collaborator **David Stephens**, she developed new insights into Bayesian causal procedures that bridge two traditionally distinct areas of statistics. With her PhD student **Mireille Schnitzer** (Université de Montréal) she extended the use of targeted maximum likelihood estimation to more general settings. With PhD students **Ryan Kyle** (plotly) and **Nabila Parveen** (Health Canada) she considered extensions of, and applications for, measurement error corrections in settings involving systematic undercounting or covariate dependent errors.

Today, Erica's work lies at the heart of DTR research. In 2013, she coauthored with **Bibhas Chakraborty** the first textbook on statistical aspects of DTRs. This bestseller, published by Springer, was followed in 2016 by an authoritative collection of works describing the current tools available to implement DTR analysis, which Erica coedited with Michael Kosorok.

In contributions to epidemiology with her colleagues Marina Klein, Joe Cox, and numerous PhD students, Erica has also applied causal inference techniques to investigate questions relating to, e.g., the interplay between food insecurity and depression, and determining individual- and community-level factors associated with prescription and illicit drug use in an HIV-HCV coinfected population. In total, Erica has supervised 15 MSc and 18 PhD students, as well as 7 postdoctoral fellows.

Erica has demonstrated exceptional leadership at the national and international level. She has been an associate editor for *Biometrics* (since 2013) and served in the same capacity for *JASA—Theory & Methods* (2014–2019). Moreover, she was the scientific program chair of the 2017 SSC Annual Meeting, and she was an associate director of CANSSI. She was also co-chair of the Causal Inference group in the STRATOS initiative of the International Society for Clinical Biostatistics.

In recognition of her contributions, Erica was listed as one of the top 50 researchers worldwide in causal inference (*The American Statistician*, 2017) and she won the McGill Principal's Prize for Outstanding Emerging Researchers in 2018. She became an elected member of the International Statistical Institute in 2015 and was a recipient of Excellence in Mentoring Awards by her department in 2017 and 2019. She is grateful to her husband and two boys for laughter, support, and keeping her active.

The citation for the award reads:

"To Erica Moodie, for her outstanding contributions to biostatistics, notably in causal inference, precision medicine, and dynamic treatment regimes, and her influential contributions to substantive areas of application such as HIV and mental health."

Christian Genest was primarily responsible for producing this material.

Paul Gustafson, recipient of SSC Gold Medal 2020



This year's recipient of the Gold Medal of the Statistical Society of Canada is **Paul Gustafson**. This prestigious award is bestowed upon a person who has made outstanding contributions to statistics, or to probability, either to mathematical developments or in applied work. It is intended to honour current leaders in their field.

Paul was born in Birmingham, U.K. in 1968, albeit to Saskatchewanian parents. In 1978 his family relocated to Prince Rupert, BC where Paul completed high school. He enjoyed his BSc (Mathematics) and MSc (Statistics) at UBC so much that, after completing his PhD at Carnegie Mellon University in 1994, he returned to the campus, first as a postdoctoral fellow and then as an assistant professor.

During his MSc studies Paul became rather enamoured with Bayes' theorem. He pursued this interest in his PhD thesis, written under the supervision of Larry Wasserman, in which he proposed formal schemes to assess the sensitivity of posterior inferences to the choice of prior distribution. Bayesian methods have remained front and center in Paul's work ever since. His record demonstrates he is a wonderful scientist as well as an exceptional statistical scientist, with landmark contributions to understanding what inference is or is not possible in complex settings due to design, missing or mismeasured data, and the other ills plaguing real studies. For instance, his

2005 Statistical Science paper on model expansion advocates use of a scientifically "honest" model rather than conveniently changing the model to suit inferential limitations. Currently he is interested in Bayesian approaches within the domains of causal inference, evidence synthesis, measurement error and partial identification.

Paul's publications remarkably range from deep theoretical insights to important applied work, often in the same paper. Much of his research program is inspired by biostatistical and epidemiological applications, and he frequently collaborates with health scientists on a range of medical conditions. His 2014 *American Journal of Epidemiology* collaborative article on multiple sclerosis was selected by the editors as one of the best in field that year.

Paul's two books, Measurement Error and Misclassification in Statistics and Epidemiology: Impacts and Bayesian Adjustments (2004) and Bayesian Inference in Partially Identified Models: Exploring the Limits of Limited Data (2015) survey a large portion of his research contributions and disseminates it to a broader audience. Both books reflect Paul's interest in understanding what combinations of modelling assumptions and prior information will yield inferences that are usefully narrow, rather than uselessly wide.

At UBC, Paul has been an important mentor. He has supervised or co-supervised 13 PhD students, 24 MSc students and four postdoctoral fellows; many have gone on to academic positions. He has also taught at all levels. His considerable service to the department includes two stints as acting head and the current headship since 2019. Paul also played a pivotal role as founding co-director of UBC's Master of Data Science program, which welcomed its first cohort of students in 2016 and is now heavily subscribed.

Beyond UBC, Paul has taken on a variety of editorial and service commitments, nationally and internationally. He was editor of *The Canadian Journal of Statistics* (2007–2009) and in 2014 he became the inaugural special editor for statistical methods at the journal, *Epidemiology*, a role he still holds. Paul is a former chair of NSERC's GSC 14 committee for grants in the statistical sciences and he was the program chair for SSC 2015 held at Dalhousie University.

Paul's work has previously been recognized with the CRM-SSC Prize in 2008 and an ASA Fellowship in 2011. Additionally, he is a two-time recipient of an NSERC Discovery Grant Accelerator Supplement.

Paul is enduringly grateful for support and foolishness-filtering from his wife, public health physician extraordinaire and occasional collaborator, **Reka Gustafson**. Whenever possible Paul and Reka can be spotted on their bicycles, either locally or in far-off lands.

For better or worse, Paul and Reka seem to have passed on their fondness for west coast life and UBC education to the next generation. One son has just completed an engineering degree, the other is midway through a Computer Science and Physics program and their daughter is about to embark on her UBC adventure this fall.

The citation for the award reads:

"To Paul Gustafson, in recognition of his seminal contributions to the foundations of inference, particularly Bayesian robustness, sensitivity analysis, and model identifiability; for exceptional advances to recognize the limitations of observational studies due to mismeasurement and other inherent obstacles, along with effective statistical tools to counter those limitations; for fundamental contributions to methodology for biomedical applications and epidemiology; and for his outstanding record of mentorship and service to the statistical science community in Canada and internationally."

Will Welch was primarily responsible for producing this material.

Karen Kopciuk, recipient of the Distinguished Service Award 2020



Karen Kopciuk is the recipient of the 2020 Distinguished Service Award from the Statistical Society of Canada (SSC). This award honours an individual who has played an important and substantial role in fostering the growth and success of the Canadian Statistical Sciences community through leadership in the SSC.

Karen completed a BSc in Psychology (1992) and a BSc in Statistics (1994) from the University of Calgary. From the University of Waterloo, she received an MMath in Statistics (1995) and a PhD in Statistics and a Certificate in University Teaching (2001). In 2003, after a MITACS-funded postdoc at the Tanenbaum-Lunenfeld Research Institute of Mount Sinai Hospital in Toronto, Karen returned to Calgary to join the Alberta Cancer Board as a research scientist.

Karen also holds adjunct appointments at the University of Calgary in the departments of Mathematics & Statistics, Oncology, and Community Health Sciences. Within her Alberta Health Services department Karen leads the Data Support and Statistical Support units, planning for computational, statistical, and personnel needs as well hiring and supervising staff for these units. She has supervised and co-supervised students at undergraduate, graduate, and postdoctoral levels, in diverse areas including biology, cancer biology, epidemiology, computer science, biostatistics, and statistics.

As a long-time member of the SSC Karen has served with distinction on numerous committees including as a member of the SSC election and the fundraising committees, as *Liaison* associate editor, and as secretary and president of the Biostatistics Section. She was faculty co-chair of the inaugural 2013 Canadian Statistics Student Conference which is now in its 8th year. She is currently serving her second term as regional representative for Alberta, British Columbia, and the Yukon on the SSC board of directors. Karen's involvement has enabled the region's ideas and concerns to be raised at a national level, creating initiatives and numerous opportunities that especially benefit statistical communities in the West.

Karen has been an active organizer of workshops and invited sessions at the SSC annual meetings including several career sessions for students. She was local arrangements co-chair for the 2019 SSC Annual Meeting at the University of Calgary and initiated the Advanced Biostatistics, Statistics and Actuarial Science Graduate Scholarship there, based on the SSC gift to the UC departments of Mathematics & Statistics and Community Health Sciences.

Beginning in 2015 Karen initiated and organized the biannual workshops on Advanced Statistical Methods at the University of Calgary Biostatistics Centre (UCBC). Through her national and international network of researchers and collaborators, eminent statisticians have visited the University of Calgary and given one- or two-day workshops on cutting-edge statistical topics and applications. These workshops have been a boon to University of Calgary students and to students joining online, providing unique educational, leadership and team-building opportunities beyond the classroom. Statisticians, data scientists, and analysts from both government agencies and industry have provided updates to their knowledge and expertise. These workshops put local statisticians and the profession on the map and have substantially raised UCBC's profile in the local, national, and international statistics community.

Another high-impact statistical community engagement initiative is the founding of the Rocky Mountain Data Science Network (RMDSN) in 2017, a joint venture between the University of Calgary, led by **Karen** and **Rob Deardon**, and the University of British Columbia-Okanagan, led by **John Braun**. The RMDSN brings together researchers, clinicians, health scientists, and members from several health organizations and is a designated Health Science Collaborating Centre (HSCC) of the Canadian Statistical Sciences Institute (CANSSI). As one of 11 such HSCCs in Canada, RMDCN aims to promote interdisciplinary collaborations that advance research through the development and application of innovative statistical methodologies that meet the research needs of the health community. An important component of this initiative is the training of biostatistics graduate students. Karen is leading the internship-training program of RMDSN, for which she has developed training guidelines and marketing materials, identified appropriate projects and solicited applications from students. To date this program has funded 10 students in various placements in Calgary and Kelowna where they have obtained real-world experience in the art and science of data analysis.

The citation for the award reads:

"To Karen Kopciuk, for exemplary service for over 15 years to the Biostatistics Section, the SSC board and several SSC committees; for her efforts to support students and trainees through career events and for co-founding the Canadian Statistics Student Conference; and for her efforts to support members of the statistical community through activities from the Rocky Mountain Data Science Network, a CANSSI Health Sciences Collaborating Centre."

Recap: SSC 2020 meeting online



Société statistique du Canada



Saturday, May 30, 2020:

Canadian Statistics Student Conference

Monday, June 1, 2020:

SSC 2020 Presidential Invited Address: David L. Donoho, (Stanford University), A Statistical Perspective on Covid-19 Testing (YouTube)

Abstract:

Today's Covid-19 public health crisis has become an economic crisis, in which many national GDPs are suffering historically unprecedented peacetime declines. Underlying these simultaneous crises and various nations' responses are many questions for which the statistician's perspective ought to/could have/should have/might still improve the science and inform national policy. I will discuss some of the important tests for Covid-19, some of the underlying technology, and some of the areas where statistics can contribute either improved understanding or improved technology.

I am not an expert!

Wednesday, June 3, 2020:

SSC 2020 Pierre Robillard Award Address: Dr. Shixiao Zhang, (Fred Hutchinson Cancer Research Center) Multiply Robust Inference for Causal Inference with Missing Data (YouTube)

Abstract:

Pretest-posttest studies are commonly used for assessing the effect of a treatment. We propose an empirical-likelihood-based approach to both testing and estimation of the average treatment effect and the chance of a longer survival in non-randomized pretest-posttest studies where the posttest outcomes are subject to missingness. The proposed empirical likelihood ratio test and the estimation procedure are multiply robust in the sense that multiple working models are allowed for the propensity score of treatment assignment, the missingness probability and the outcome regression, and the validity of the test and the estimation requires only a certain combination of those multiple working models to be correctly specified.

Thursday, June 4, 2020:

SSC Data Science and Analytics Section Online Workshop: Analysis of Unstructured Text Data with Prof. Nathan Taback (University of Toronto) and Prof. Dave Campbell (Carleton University)

(YouTube) Part 1 will cover data acquisition, cleaning, tokenization, and exploration

(YouTube) Part 2 will cover topic models for clustering documents and converting words into low dimensional numeric vectors.

Friday, June 5, 2020:

SSC Biostatistics Online Workshop: Intro to Epidemic Modelling with Prof. Rob Deardon (University of Calgary)

(YouTube) Session 1

(YouTube) <u>Session 2</u>

Saturday, June 6, 2020:

SSC Business and Industrial Statistics Online Workshop: Interactive Exploratory Data Visualization in R with Prof. Wayne Oldford, (University of Waterloo)

A workshop introduction to loon and related packages

More details here.

Wednesday, June 17, 2020

SSC Board Meeting beginning at 2:00 p.m. Eastern Time

Thursday, June 25, 2020

SSC Annual General Meeting beginning at 2:30 p.m. Eastern Time

CANSSI news



Canadian Statistical Sciences Institute Institut canadien des sciences statistiques

Data • Discoveries • Decisions Données • Découvertes • Décisions

CANSSI AGM and Town Hall on NSERC Discovery Institute Support Program

This year, we'll be having our first ever <u>virtual AGM</u> on Thursday, June 18, 2020. It will take place online at 9:00 PDT (12:00 EDT) and will be followed by a **Town Hall Meeting** at 10:00 PDT (13:00 PDT).

All CANSSI member institutions are allowed to send one member representative to vote at the AGM, but anyone is welcome to attend.

The Town Hall session is also open to anyone. Statisticians are encouraged to attend to learn about the CANSSI proposal to the NSERC Discovery Institute Support Program. NSERC has opened the competition for institute funding. For the first time, CANSSI will apply in a head-to-head competition with the mathematics institutes (PIMS, BIRS, Fields, CRM, AARMS) as well as the Canadian Institute for Theoretical Astrophysics (CITA).

Learn more about the Annual General Meeting and the Town Hall here.

Hope to see you there!

CANSSI Strategic Retreat Report—Planning for the DIS Proposal

CANSSI held a strategic retreat from March 6–8, 2020. Participants from across Canada attended. The goal of the retreat was to bring together a cross-section of the Canadian statistics and inferential data science community to consider CANSSI current and future operations in preparation for the NSERC Discovery Institutes Support (DIS) competition. The intention is to submit a proposal for CANSSI that has a strong grassroots foundation in the Canadian statistics and inferential data science community. The retreat was a key step in fulfilling this goal. The full report (available online) is intended to capture the ideas, advice, questions, and issues that arise out of the discussions in the retreat.



CANSSI Strategic Retreat, March 2020

CANSSI Rapid Response Program—COVID-19

CANSSI is requesting proposals to the CANSSI Rapid Response Program (CRRP) that address issues related to COVID-19. Learn more about the program and how to apply here.

Some exciting projects that have already been approved include:

- Estimating the Number of Hidden COVID-19 Cases led by Laura Cowen, Junling Ma, and Pauline van den Driessche at the University of Victoria
- Perfectionnement de l'estimation des matrices de contacts sociaux pour la modélisation de l'épidémie de COVID-19 au Canada led by Alexandre Bureau, Mélanie Drolet, and Marc Brisson at Université Laval
- Host Genetics for SARS-CoV-2 Severity and Infection: Power and Control led by Lloyd Elliott at Simon Fraser University
- Statistical Methods for COVID-19 Mortality Forecasting at the Small-Area Level led by Jeffrey Rosenthal and Patrick Brown at the University of Toronto

Read more about these projects on our website.

New CANSSI Collaborative Research Team Projects for 2020

Two new Collaborative Research Teams (CRT) kicked off their projects in 2020.

- Addressing Spatial and Computational Issues in Integrated Analysis of Modern Ecological Data: The team leaders are Laura Cowen, University of Victoria; Simon Bonner, Western University; and Saman Muthukumarana, University of Manitoba.
- Modern Techniques for Survey Sampling and Complex Data: The team leaders are David Haziza, University of Montreal and Changbao Wu, University of Montreal.

Learn more about all of our current CRTs here.

Distinguished Visitor Program Deadline-July 31, 2020

Looking to have a leading international statistical scientist at your university? CANSSI has a program to help fund that. Your distinguished visitor should plan to give a 1-hour lecture for a general audience in addition to a scientific talk, and should meet with researchers and graduate students during the visit. Find out how to apply here.

A newly launched COVID-19 Canada website

In the midst of the global outbreak with over 300,000 worldwide death cases of COVID-19, Canada has reported 79,101 confirmed cases of the novel coronavirus (COVID-19) as of May 19, 2020, in which the severity differs from region to region. To provide a timely view and understanding of the evolving pandemic in Canada, a research team at the University of Western Ontario developed a real-time interactive web-based platform, https://covid-19-canada.uwo.ca/. The platform reports the real-time evolvement of COVID-19 in different regions of Canada by providing informative visualization from multiple angles and statistical analysis using different models. The focus is on Ontario, Alberta, British Columbia, and Quebec, the four provinces in Canada that have the most serious situations. Our objectives are to provide timely summarized displays of the development of COVID-19 in Canada to offer real-time evidence-based findings and to create an interactive platform of up-to-date information on COVID-19 in Canada.

The website primarily consists of three sections: DATA VISUALIZATION, STATISTICAL ANALYSIS, and RESEARCH MANUSCRIPTS, which highlight real-time tracking of the development of COVID-19 with visualized graphs and forecast future trends with applications of different predictive models. By providing research-based statistical analysis, we shed the light on the epidemiological characteristics of COVID-19. In addition, the section PUBLIC DATA SOURCES offers quick access to a collection of data sources for researchers and industry professionals to conduct their own research on COVID-19. SELECTED NEWS and STOP VIRUS provide news updates and links for self-assessment tools in the local region as well as preventive measures suggested by the government. READER'S CORNER was created for website visitors to share opinions or feedback.

With the support from the Faculty of Science at the University of Western Ontario, the platform was developed by the research group, co-led by professors **Grace Y.** Yi and **Wenqing He**, which includes **Dan Liu**, **Yuan Du**, **Yasin Khadem Charvadeh**, **Jingyu Cui**, **Li-Pang Chen**, **Kaida Cai**, **Qihuang Zhang**, **Gansen Deng**, and **Joy He**. It was officially launched on May 8, 2020.

Grace Y. Yi, PhD

Professor

Canada Research Chair in Data Science (Tier 1)

Department of Statistical and Actuarial Sciences

Department of Computer Science

University of Western Ontario

Statistical Society of Canada contributors list 2019



Société Statistical statistique Society du Canada of Canada

Dear SSC members and friends,

Thank you for your generous contributions to the Statistical Society of Canada for 2019. This list celebrates each of you and your support of the mission of the SSC.

"Friends of the SSC," who have given \$200 or more, are indicated by an asterisk (*).

Acadia University

Nassr Saleh Al-Maflehi

Mayer Alvo

Joe Apaloo

Llwellyn M. Armstrong*

Eshetu G. Atenafu

John D Baser

François Bellavance*

Nadeene Blanchard-Martin

Cynthia J. Bocci

Angelo J. Canty

Hugh Chipman*

Gabriela Cohen Freue

Philip Christopher Darling

Donald A. Dawson

Joel A. Dubin

Bruce Dunham

Bradley W. Frankland

Jack Gambino*

Eric Germain

Kilani Ghoudi

Alison L. Gibbs*

Paul Gustafson*

X. Joan Hu

Cody Hyndman

Jack D. Kalbfleisch

Jerald F. Lawless

Christian Léger*

Kuan Liu

Richard A Lockhart*

Tom M. Loughin

Kenneth B. McRae

Joanna Mills Flemming

Shirley E. Mills*

Johanna G. Neslehova

Albert John Petkau*

David W. Plumstead

Josh W. Pohlkamp-Hartt

Harris Quach

Jean-Francois Quessy

Nancy Reid*

John R. Rutherford*

Anna Theresa Santiago

Dena L Schanzer

Lei Sun

Thai-Son Tang

Mary E. Thompson*

David J. Thomson*

Jean Vaillancourt*

Xu (Sunny) Wang

Kenneth Laurence Weldon

David R. Wilson

Hubert Wong

Changbao Wu*

Lang Wu

Ying Zhang*

And 47 anonymous donors

The Statistical Society of Canada is grateful for a generous memorial donation by the family of the late Marc Moore.

Online short course on Bayesian nonparametrics

Nonparametric Bayesian data analysis Peter Müeller UT Austin, USA

July 6-8, 2020 2:00-4:00 p.m. (EST)

This online short course is jointly organized by the CRM-StatLab and the graduate program of Biostatistics, McGill University.

Free registration. Details will soon be posted on the CRM website.

Abstract:

All models are wrong, but some are useful. Many statisticians know and appreciate G.E.P. Box's comment on statistical modelling. Often the choice of the final inference model is a compromise of an accurate representation of the experimental conditions, a preference for parsimony and the need for a practicable implementation. The competing goals are not always honestly spelled out, and the resulting uncertainties are not fully described. Over the last 20 years a powerful inference approach that allows to mitigate some of these limitations has become increasingly popular. Bayesian nonparametric (BNP) inference allows to acknowledge uncertainty about an assumed sampling model while maintaining a practically feasible inference approach. We could take this feature as a pragmatic characterization of BNP as flexible prior probability models that generalize traditional models by allowing for positive prior probability models on infinite dimensional parameter spaces. A typical application of BNP is to density estimation.

In this short course we review some of the popular models, including Dirichlet process (DP) models, Polya tree models, DP mixtures and dependent DP (DDP) models. We will review some of the general modelling principles, including species sampling models, stick breaking priors, product partition models for random partition and normalized random measures with independent increments. We will briefly discuss some of the main computational algorithms and available software. The discussion will be illustrated by applications to problems in biostatistics and bioinformatics.

Topics covered:

- Definition of BNP and introduction
- Density estimation: Dirichlet process (DP), Stick breaking, DP mixtures, DP clustering, DP mixtures: posterior simulation, Polya trees (PT)
- Regression: BNP survival regression, Dependent DP (DDP), Anova DDP, Weighted mixture of DP, Kernel stick breaking process, Gaussian process priors
- Hierarchical priors: Hierarchical DP, Nested DP
- Mixed effects models: Random effects distributions multiple subpopulations & classification

Target audience & prerequisites: Anyone with an appreciation for data analysis, and basic knowledge of Bayesian inference. At the level of, for example, Hoff (2009), A First Course in Bayesian Statistical Models. Or any other basic text in Bayesian inference.

References: Items #2-4 are free online, #1 is probably available as free PDF from your library (same for Hoff (2009))

- 1. The course will follow the book: Müller, P., Quintana, F., Jara, A., and Hanson, T. (2015), Bayesian Nonparametric Data Analysis, Springer.
- 2. Maybe read (before the course) P. Müller and R. Mitra (2013), "Bayesian Nonparametric Inference—Why and How," Bayesian Analysis, 8, 269-302.
- 3. Lecture notes of a similar course: Müller, P. and Rodriguez, A., (2012) "Nonparametric Bayesian Inference," IMS Lecture Notes, free at https://projecteuclid.org/euclid.cbms/1362163742
- 4. Excellent notes by Peter Orbantz, at http://www.gatsby.ucl.ac.uk/~porbanz/papers/porbanz BNP draft.pdf

Independent Summer Statistics Community: Beyond coursework!



Knowing that COVID-19 would possibly bring disappointing disruptions to our students' summer plans, professors Liza Bolton, Nathalie Moon, and Nathan Taback, along with Megan Whitehead-Douglas—mentorship and internship coordinator in the Department of Statistical Sciences, came together to introduce an exciting initiative to support undergraduate students: The Independent Summer Statistics Community (ISSC).

At the core of the ISSC are curated resources about portfolio building, career preparation, and personal development, plus a platform on which to share, discuss, and support each other.

Some of the benefits to students of joining the ISSC:

- DataFest! An online version of the popular annual ASA DataFest @ UofT will be run in early June, and ISSC members have exclusive access to register.
- **Professional portfolio**. From building a website or blog to learning how to use GitHub or make a Shiny app—ISSC supports members in developing their personal brand!
- Remote working skills. Practising collaborating remotely with popular tools like Slack; building the online working skills that will matter to interviewers.
- Networking and community. Forging connections with fellow students, invited speakers and participating faculty. It is these kinds of connections that often lead to future opportunities.
- Curated resources. The ISSC team has carefully gathered a range of useful resources that are shared on Slack or as one of the six topics in the weekly 6σ Sunday newsletter.

Since it was announced in late April, over 700 statistics students have registered for the ISSC. Over 100 students joined the live online launch event on May 1, while others tuned in later for the recording. For the past several weeks, registrations for the University of Toronto's first online iteration of the ASA DataFest (website: https://datafestuoft.github.io/) have been pouring in, with approximately 300 ISSCers eager to put their statistical modelling and communication skills into practice. In preparation for DataFest, ISSC members have been actively participating in a number of virtual events, including weekly TidyTuesday meetups and a workshop on Gathering Data from Messy Sources, with Rohan Alexander, incoming faculty member at the Department of Statistical Sciences. Coming up, the ISSC team is looking forward to offering ISSCers the opportunity to complete the CliftonStrengths assessment and engage with alumni at career panels.

"It is an extremely rewarding experience to connect with and mentor students 'beyond coursework,'" said Liza Bolton; "in a setting where the only assessments are the students' own evaluations of their progress towards their goals." Nathalie Moon added, "Students pick their own pace and interests, and this creates space to share the types of tools, skills and ideas that often can't be the focus of undergraduate classes."

The Department of Statistical Sciences is looking forward to seeing how the ISSC continues to develop over the course of the summer, and the lasting impact the community will carry forward for students.

News from the committee on membership



Société Statistical statistique Society du Canada of Canada

Most of the people I studied with now work as "data scientists," yet few would define themselves as statisticians in the classical sense of the term. The <u>Data Science and Analytics section</u> is thus a welcome addition to the SSC family and I am happy to report that already 120 members have joined it, making it the second most popular SSC section after biostatistics. Thanks are due to the initiators and to the inaugural committee for putting the section on its feet.

The committee on membership has been looking into ways of improving members experience and ways of convincing members, particularly students, to renew their membership. We will be updating the list of benefits on the website to make access to the *Canadian Journal of Statistics*easier. In 2014, the strategic planning committee circulated the results of a vast survey of our members who listed the "intangible benefits" as being prime factors for membership: "Respondents most valued the current SSC benefits of advocacy on their behalf, outreach in the education sector, and access to a large network of colleagues." (*Liaison*, vol. 28, no. 4, p. 35). We want to make these intangible benefits of the SSC more concrete to newcomers by including examples of leadership throughout the years and testimonials on our website. To this effect, we would be glad if you would agree to share your experience as a member with us using this form; comments and testimonials are welcome.

Léo Belzile

Chair, committee on membership

News from HEC Montréal



A great deal has happened since we last provided news from HEC in *Liaison* and are worthy of mention. Two of our colleagues are new chairholders: **Aurélie Labbe** is the inaugural recipient of the <u>FRQ-IVADO Chair in Data Science</u> and <u>Laurent Charlin received a CIFAR-AI Chair</u>. **Bruno Rémillard** was made Honorary Professor by the board following his retirement. **François Bellavance** started his first mandate as <u>Director of Academic Programs</u> and is now weathering the storm caused by the pandemic with his usual dedication and skills. Among other recent news worthy of mention:

- Congratulations to Luc Villandré and David Ardia for their funded <u>IVADO Covid-19 initiatives</u>. I encourage you to have a look at the <u>COVID-19 Data Hub</u>, which aims to provide worldwide users with a unified data hub.
- Maxime Larocque, an MSc student in the business intelligence program supervised by Jean-François Plante and Michel Adès (UQAM), is the recipient of
 the Best Master Thesis Award for his work "Bagged Genetic Agorithms for Objective Variable Selection."
- Jian Tang received an Amazon Faculty Research Award and funding for collaborative work from Microsoft-Mila for his project "Towards Combining Statistical Relational Learning and Graph Neural Networks for Reasoning."
- Claudie Ratté-Fortin is the recipient of an IVADO postdoctoral scholarship for a project on machine learning for modelling extreme events.

The Department of Decision Sciences continues to expand: **Juliana Schulz**, **David Ardia**, and myself joined last summer and we are happy to welcome **Dena Firoozi** as assistant professor within our ranks starting June 1. Dena received her PhD in Electrical Engineering from McGill and is currently postdoctoral fellow at the University of Toronto. She is specialized in optimal control, estimation theory and mean-field games with applications in financial engineering. Congratulations also to **David Ardia** for his promotion (with tenure) to associate professor. Information about other new hirings will be communicated in an upcoming issue.

A warm welcome to another new faculty member!



Our department is delighted to announce the addition of yet another brilliant colleague to our faculty. We look forward to working with Vianey, starting this July, and warmly welcome her to our team. Additional hiring news will be reported in the next issue of *Liaison*.



<u>Vianey Leos Barajas</u> will be joining the Department of Statistical Sciences and The School of the Environment as an assistant professor in July 2020. Vianey primarily works in the area of statistical ecology. She has had the opportunity to collaborate with researchers all over the world analyzing data in fascinating areas such as shark movements and snake tracking. Most recently, she began working in environmental statistics, with a focus on spatial modelling of disease data. Vianey looks forward to pursuing collaborative projects with researchers from both departments at the University of Toronto and creating an ecological and environmental statistics group.

By Radu Craiu

News from Université Laval



The School of Actuarial Science at Université Laval is proud to welcome **Thai Nguyen** as assistant professor as of February 2020. Thai Nguyen obtained a PhD in Applied Mathematics at the University of Rouen in 2014. As a postdoctoral fellow, he then took part in the research and teaching activities of the Institute of Insurance Science at the Universität Ulm. His research focuses on stochastic calculus and their applications in finance, economics and insurance. Within the context of financial mathematics, he is particularly interested in approximate hedging under transaction costs, indifference pricing, expected utility maximization under market impact, stochastic control theory with risk constraints and non-concave optimization. Welcome Thai!



Thai Nguyen

The Department of Mathematics and Statistics of Université Laval congratulates the student **Ouassima Benouari** for her nomination as recipient of the SFds-ISQ-SSC Scholarship for the quality of her academic record and her master's thesis, entitled "Variable Selection Criterion for Mixed Conditional Logistic Regression Models when the Random Effects Structure is Unknown."

The Statistical Consulting Service (SCS) announces the hiring of **Laurence Desbois-Bédard** as of March 30. Laurence obtained a master's degree in statistics from Université Laval in 2017, under the supervision of **Anne-Sophie Charest** and **Louis-Paul Rivest**. She worked on the generation of synthetic data for continuous variables using various copula-based methods.



Laurence Desbois-Bédard

<u>Q & A with Jesse Gronsbell, assistant professor, Department of Statistical Sciences at the University of Toronto</u>



With a PhD in Biostatistics from Harvard University, where she was supported by a Ruth L. Kirschstein Predoctoral Individual National Research Service Award from the National Institutes of Health, **Jesse Gronsbell**'s research has a focus on the development of statistical methods for electronic health records and mobile health data. Her work has been recognized with a Young Researchers Award from the International Society of Nonparametric Statistics and a Gertrude Cox Scholarship honourable mention from the American Statistical Association.

U of T Statistical Sciences: Welcome to the StatsSci team and congrats on your new role! What got you interested in U of T and the Stats Department?

Thanks so much! I am humbled to be joining the Stats Department and can't wait to get started. For me, the decision to come to U of T was an easy one. Within the department and across the university, researchers are doing groundbreaking work at the intersection of medicine, statistics, and machine learning. This unique mix of technical talent and strength in clinical research makes U of T an ideal place for a biostatistician like myself. Also, personally and politically, I am really excited to move to Toronto.

Can you tell us a bit more about your background?

It's ironic that I entered academia because I wasn't sure if college was right for me. When I graduated high school, I got a job as a line chef at a Mexican restaurant because I love to cook. Unfortunately, I was quickly fired after destroying an industrial-sized batch of guacamole. Since my chef dreams were crushed (along with a ton of avocados), I decided to give college a try.

I took a discrete mathematics course in my sophomore year that led me to major in math. While I enjoyed theoretical classes, I was drawn to how math could be used to formalize, provide structure to, and ultimately solve real-world problems—particularly those in high-impact fields like medicine. This interest led me to take courses in biostatistics and pursue a PhD in the field.

I was fortunate to do my graduate work under the supervision of Dr. Tianxi Cai at Harvard, then spend a year as a postdoc at Stanford with Dr. Lu Tian. Currently, I'm a data scientist at Alphabet's Verily Life Sciences on the mental health research and development team led by Dr. Menachem Fromer. After having a taste of life in industry and with the encouragement of these amazingly supportive mentors, I am happy to have found my home back in academia at U of T.

Moving from Boston to Toronto in the middle of all the current restrictions—I can't imagine how stressful that might be! What are you most looking forward to about life in Toronto?

It's difficult to digest the disruption that this pandemic has caused. The restrictions complicating my physical move are a small problem to have right now. I am beyond lucky to have the opportunity to pursue my dream job during this time and grateful to have something to look forward to!

Academically, I am eager to fully re-engage with my research after being in industry for the last few years. Personally, I'm looking forward to the adventure of living in such a large and diverse new city.

What do you love about statistics?

John Tukey said it well, "The best thing about being a statistician is that you get to play in everyone's backyard." In my short career, I have had the opportunity to apply statistics to baseball pitch classification, ballistic missile defense, and disease diagnosis. It is exciting to have a skill that is applicable to many different fields and in turn allows you to continuously learn something you otherwise never would have encountered. I think it's really fun to work with experts in other domains to figure out the right problem to solve and the appropriate statistical framework to do it.

What courses will you be teaching this fall?

I found out yesterday I will be teaching STA2112H, Mathematical Statistics I. Nancy Reid will be teaching the second half of the course in the spring so I am excited to collaborate with her.

Can you tell us about your current research interests?

I develop statistical and machine learning methods to understand the digital phenotype—the health information that is left behind in the digital data we generate in our everyday lives. Most of my work is centered on routinely collected clinical data in the form of electronic health records (EHR) and more recently on mobile health data captured from a smartphone. Both of these data sources hold great potential in furthering our understanding of health and disease outside of traditional clinical studies. However, the data is captured in situ and as a result, is of lower quality than what we would gather in a controlled study. This often prohibits the use of standard analysis methods and paves the way for statisticians like myself to develop new methods to better accommodate the data.

Like you said, your research has focused on electronic health records data. Does EHR data have potential to shape our understanding of COVID-19?

That is a great question and I think the honest answer is that researchers are still figuring it out. It is important to remember that scientifically, COVID-19 is new, and the knowledge base of the virus is just beginning to be built. This limits the kinds of statistical analyses that can be conducted and the conclusions that can be drawn.

That being said, one of the great advantages of EHR data is that it allows for rapid consolidation of large amounts of clinical data from affected countries. In fact, The Consortium for Clinical Characterization of COVID-19 by EHR (4CE) has recently published an article that investigates data covering over 27,000 COVID-19 cases from 96 hospitals in the US, France, Italy, Germany, and Singapore. The team was able to wrangle an enormous amount of data and summarize findings in under two weeks. This impressive effort demonstrates the power of EHR data to enable studies of evolving diseases like COVID-19. However, these large-scale analyses are most effective in complementing and potentially validating traditional clinical research studies due to inherent limitations of EHR data, such as only capturing patients who are "sick enough" to go to the hospital and lacking information on important environmental factors.

What about smartphone data?

At the population level, I'm excited by public health monitoring apps. In particular, the HowWeFeel (HWF) app allows users to report demographic and symptom information in as little as 30 seconds, as well as see how people near to them are feeling. Aggregated user data is helping to find new hotspots of the virus and to identify populations who may be at higher risk of getting sick. And as an added plus, HWF will donate a meal to people in need through Feeding America with each new download.

Things are more challenging at the individual level. With the recent rise in telemedicine, there is a lot of interest in developing markers from the smartphone that can be used to capture health status between visits, particularly in mental health. However, there isn't a one-to-one mapping between the markers that were useful pre-pandemic and those that are useful during the pandemic. For instance, daily GPS data can be transformed to identify periods with semantic meaning such as "home", "work", and "commute". A change in someone's routine can indicate a shift in mental health status. But with most people spending most of their time at home, many of the markers that we once relied on now have almost no variability. It will be interesting to see the developments in this space in the coming months.

What do you hope to accomplish in your tenure at U of T?

There's a lot I want to accomplish, but two things immediately come to mind.

For one, I'm eager to bring my experience in industry to my research, teaching, and mentoring. One of the biggest differences between academia and industry is the speed with which a statistics or machine learning problem is translated from paper to production. I hope to develop a research program that effectively bridges the theoretical and the practical, and to do the same in my teaching. As a student I was unsure about the decision between industry and academia, so I can also offer some guidance to students at U of T.

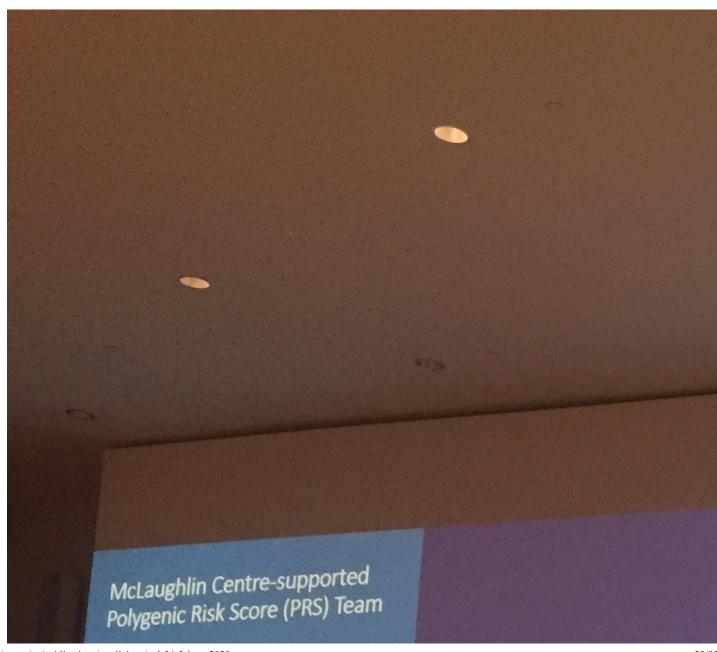
The second goal I have is personal. I am the first in my family to receive a college degree and my path to this point hasn't been easy. It is important for students to know that there is no single path to success and I think it's necessary for faculty to speak to that. My experiences have made me passionate about encouraging the traditionally underrepresented to pursue higher education. Last year, I spoke at the annual ACM Richard Tapia Celebration of Diversity in Computing Conference that brings together students and professionals in computing from all backgrounds and ethnicities. I am now using some of my spare time during the stay-at-home order to review proposals for next year's doctoral consortium. I'm looking forward to engaging in (or even developing) programs like these as a faculty member at U of T.

Do you have some fun facts about yourself you could share? What do you do in your free time to unwind? Any interesting hobbies?

I moved a lot as a kid and spent a number of years in Fargo, North Dakota. While I was there, I took up competitive figure skating and still enjoy hitting the ice from time to time. I have also been distance running for 15 years and hate to go a day without getting in a few miles. And as I mentioned, I love to cook—just not on the line.

© 2020 U of T Statistical Sciences

Team accelerator grant in the amount of \$300,000 renewed!



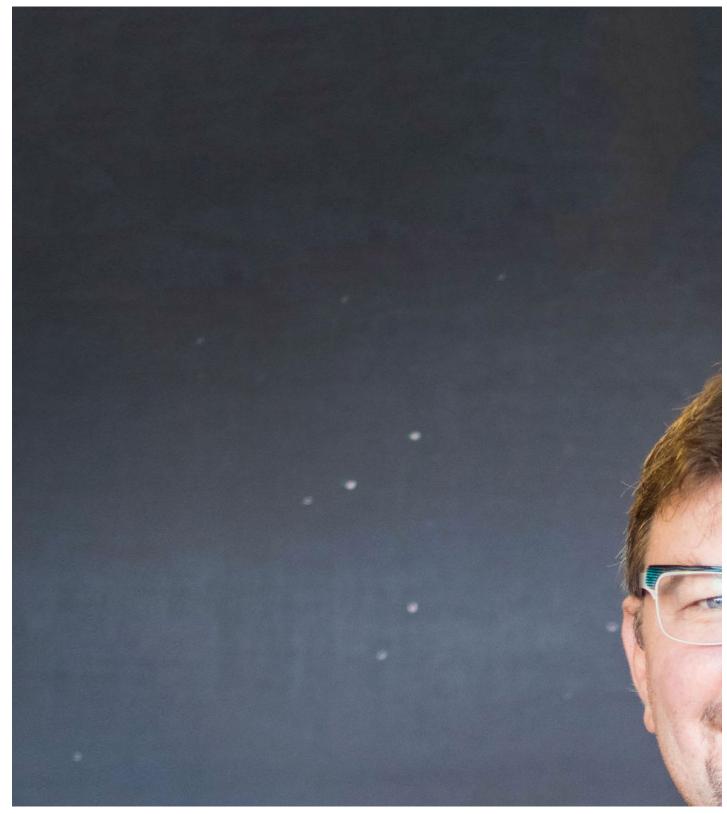


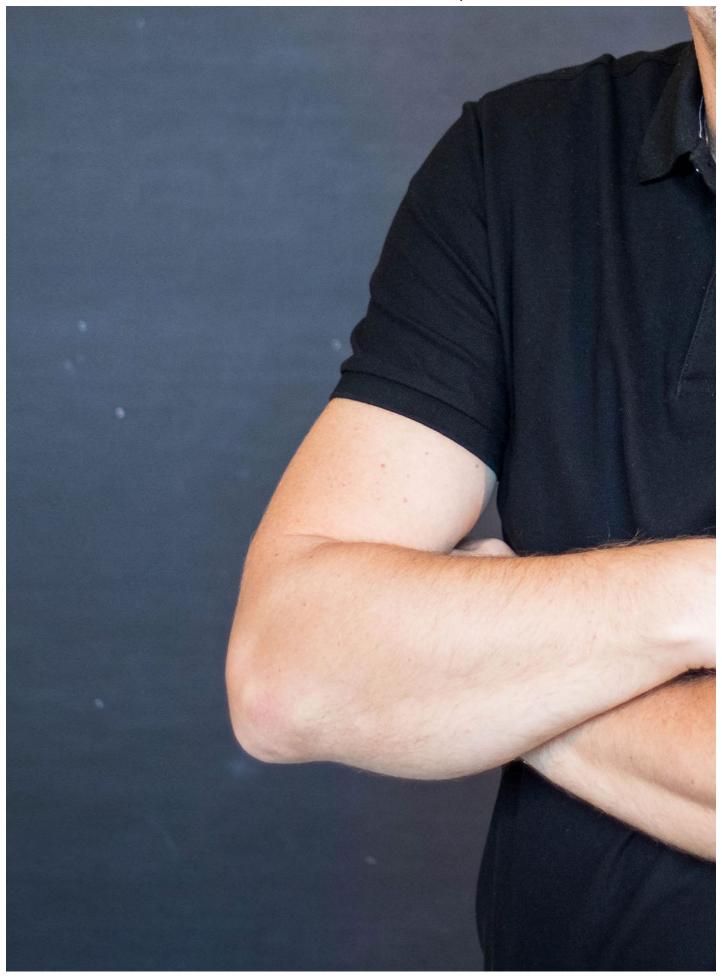
We are proud to announce that a team accelerator grant, co-lead by our Associate Chair **Lei Sun** and statistical faculty members, **Linbo Wang**, **Lisa Strug**, **Dehan Kong**, and **Qiang Sun** has been renewed for another year, in the amount of \$300,000, by the Accelerator Grants in Genomic Medicine 2020 Program of the <u>U of T McLaughlin</u> Centre.

The team grant develops "A unified polygenic risk score analytics platform for risk prediction, causal inference and discovery." The research was recently highlighted by the <u>CANSSI Ontario Research Day</u> on March 5, 2020.

In addition to collaborators from DLSPH, HSC, MSH, and CAMH, the team would like to sincerely thank the McLaughlin Centre for supporting research in developing statistical methodologies that advance genomic medicine.

Radu Craiu named Fellow of the Institute of Mathematical Statistics





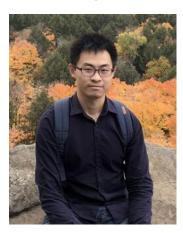
Radu Craiu, chair of the Department of Statistical Sciences, has been named Fellow of the Institute of Mathematical Statistics (IMS). He received the prestigious award for his fundamental contributions to the methodology and theory of Markov chain Monte Carlo, for diverse investigations into statistical uses of copulas, for significant insights into statistical genetics and false discovery rates, for editorship of several research journals, for mentorship of numerous research students, and for leadership in the unprecedented expansion of his department's research compliment.

Other significant awards he has received include a CRM-SSC Award (2016) and elected member of the International Statistical Institute (2015).

To be named an IMS fellow is an impressive honour. Each person nominated for a fellowship is assessed by a committee of peers.

Professor Craiu would like to thank the institute. He is delighted to receive this honour and to join the ranks of other IMS fellows.

Shixiao Zhang, winner of the SSC 2020 Pierre Robillard Award



Shixiao Zhang is this year's winner of the <u>Pierre Robillard Award</u> of the Statistical Society of Canada (SSC) which recognizes the best PhD thesis in probability or statistics defended at a Canadian university in a given year. His thesis, entitled "Multiply Robust Empirical Likelihood Inference for Missing Data and Causal Inference Problems," was written while Shixiao was a doctoral student in the Department of Statistics and Actuarial Science at the University of Waterloo under the supervision of **Dr. Peisong Han** and **Dr. Changbao Wu**.

Currently, Shixiao is a postdoctoral research fellow at the Fred Hutchinson Cancer Research Center in Seattle, Washington. He joined the Fred Hutchinson Cancer Research Center in July 2019 after the completion of his PhD studies at the University of Waterloo.

His current research interests are broadly focused on missing data, causal inference and personalized medicine. During his PhD and postdoctoral studies, his main focus has been in applying important techniques such as empirical likelihood and calibration methods to improve estimation efficiency and robustness against possible model mis-specification. More recently, at the Fred Hutchinson Cancer Research Center, he is working to develop individualized treatment rules (ITRs) using statistical learning techniques such as support vector machines. In particular, he is focused on finding the optimal treatment recommendation that maximizes the treatment benefit for various patients based on their individual characteristics.

Previously, Shixiao earned his BSc in Statistics and Mathematics from the University of Hong Kong in January 2014 and an MA in Statistics from York University in July 2015.

About the award

The criteria used in selecting the winner of the Pierre Robillard Award includes the originality of ideas and techniques, the possible applications and their treatment, and the potential impact of the work. The award is named in memory of **Professor Pierre Robillard**, an outstanding dynamic young statistician at the Université de Montréal, whose untimely death in 1975 cut short what promised to be a highly distinguished career.

The citation for the award reads:

"To Shixiao Zhang, for the thesis entitled 'Multiply Robust Empirical Likelihood Inference for Missing Data and Causal Inference Problems'."

The award was announced on the SSC website on May 19, 2020. Dr. Shixiao Zhang will present his award winning address on June 3, 2020 through a virtual platform

Five Canadian researchers named IMS Fellows

Each year, the Institute of Mathematical Statistics (IMS) selects a class of fellows who have demonstrated excellence in research or leadership in statistics or probability. This high distinction was awarded this year to 35 individuals, i.e., approximately 1% of the institute's active membership. Five members of the 2020 class of fellows are affiliated to Canadian universities, namely probabilist **Louigi Addario-Berry** (McGill U.) and statisticians **V. Radu Craiu** (U. of Toronto), **Charmaine B. Dean** (U. of Waterloo), **Johanna G. Nešlehová** (McGill U.), and **Grace Y. Yi** (Western U.). For a summary of their accomplishments and a complete list of the 2020 IMS Fellows, see https://imstat.org/2020/05/17/congratulations-to-the-2020-ims-fellows/

Charmaine Dean named IMS Fellow

WATER LOO



FACULTY OF MATHEMATICS Department of Statistics and Actuarial Science

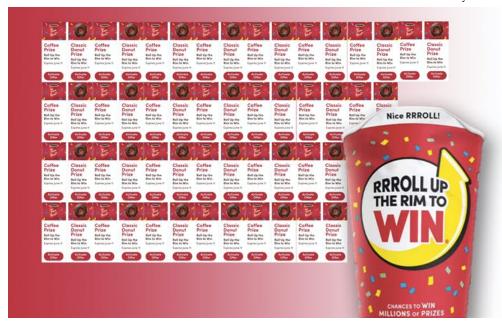


Charmaine Dean, vice president, Research and professor in the Department of Statistics and Actuarial Science, University of Waterloo, has been named Fellow of the Institute of Mathematical Statistics (IMS). Dr. Dean received the award for her scientifically important contributions to the analysis of count data, disease mapping, spatio-temporal data and more; for her outstanding leadership to the statistical profession, her record of mentorship and for her enormous work in keeping statistics visible at the center of science.

Each fellow nominee is assessed by a committee of their peers for the award. In 2020, after reviewing 73 nominations, 35 were selected for fellowship. Created in 1935, the Institute of Mathematical Statistics is a member organization that fosters the development and dissemination of the theory and applications of statistics and probability. The IMS has 3,500 active members throughout the world. Approximately 12% of the current IMS membership has earned the status of fellowship.

MEDIA CONTACT | Elyse Gustafson

Here's how I cracked Roll up the Rim and won (almost) every time





A statistics professor used his expertise in calculating probabilities to come up with a 98 winning percentage for Tim Hortons popular Roll up the Rim contest.

In 2003, geological statistician **Mohan Srivastava** cracked a Canadian lottery game. After getting a scratch-off ticket as a gift, he spotted a flaw in the game. Putting his probability skills to work, he figured out how to identify winning tickets prior to purchase. It didn't make him rich—"beating the game wasn't worth my time," he said—but Srivastava's experience showed it's very hard to make a truly random game.

Tim Hortons launched its annual Roll up the Rim promotion in mid-March. Due to the coronavirus pandemic, the contest's iconic physical cups were replaced by a digital game on the company's loyalty app. This change had major implications for how prizes were awarded. It also created an opportunity to shift the odds heavily in your favour.

Under the old format, the mechanics of the game were simple. Tim Hortons would print millions of promotional cups and about one out of every six cups had a prize under its paper rim. If the company expected to sell 180 million coffees during the contest period, then they'd print 180 million promotional cups and make about 30 million of them winners. Stores would then sell the cups until supplies ran out.

This meant that while Tim Hortons couldn't perfectly predict how many coffees they'd sell during the contest period, the odds remained the same. If sales were unexpectedly high, the cups would simply run out sooner. If sales were low, the game would last a little longer. Either way, the number of prizes and the number of cups given out—the numerator and denominator of our probability equation—were fixed.

"Digital rolls" changed the odds

This year, the digital game changed all that. Instead of physical cups, players earned "digital rolls" with their purchases. These could be rolled at any time by tapping a virtual coffee cup on the app.

Unlike the physical cups however, Tim Hortons couldn't control how many digital rolls were given out. It depended on their actual sales during the four-week contest period. The number of prizes (the numerator) was still fixed, but the number of entries (the denominator) was out of their hands. They had to find a new way to distribute the prizes.

The solution, as I wrote about before the start of the contest, meant Roll up the Rim became a bit like a slot machine.

Changes to Roll up the Rim contest made the game more like a slot machine

Each prize—from a coffee to a car—was assigned a "winning timeframe," a short period of time during which it could be won. If you were the first player to tap your app during that timeframe, you won the prize. Tim Hortons could then distribute all the advertised prizes across the contest period and every single one—at least in theory—could be won.

With some of these winning timeframes as short as 0.1 seconds though, it was certainly possible some would pass with no one playing a roll. Anticipating this, Tim Hortons included a rule where any unclaimed prizes were rolled over to another day of the contest, right up until the final day. A prize would only go unclaimed if it

wasn't won by the digital roll deadline at midnight on April 21. Although the in-store contest ended April 7, players had an additional two weeks to use their digital rolls.

Prizes rolled over

This new game format meant your odds of winning depended not only on how many digital rolls you earned, but also how many others were playing. If sales were low, then fewer people would play—and more of those winning timeframes would pass without the prize being claimed. Those unclaimed prizes would roll over—and keep rolling over—until the contest deadline.

It's not known whether social isolation had an impact on the number of people who played Roll up the Rim this year.

We don't know for sure because Tim Hortons hasn't announced its sales figures during the contest, but it's reasonable to assume that <u>fewer coffees have been sold during the coronavirus pandemic lockdown</u>. Lower sales could theoretically increase the odds of winning a prize through the app, but only if you knew when to play.

As a statistics professor, it seemed appropriate to conduct a small experiment. I knew there was little chance of winning a big prize—over 99% of the contest's prizes are coffee or doughnuts—but I was statistically curious.

There were other factors to consider. How randomly were the prizes distributed? Were they more likely to be won during the start of the contest? What times of day did people play? Was I more likely to win in the middle of the night? There was no way to know for sure, but I made some guesses and got to work.

An avalanche of unclaimed prizes

Taking advantage of the various ways to earn bonus rolls (such as owning a reusable cup or making a purchase through the app), I racked up 96 entries. I then waited until the very last day of the contest to play them. If my theory was correct, based on statistical models I had calculated before playing, there would be an avalanche of unclaimed prizes waiting to be won.

Waking up at 5 a.m.—I wanted to minimize the chance I was playing at the same time as someone else—I was finally ready to roll.

On my first play, I won a free coffee. A good start, but maybe I'd just gotten lucky.

Author Michael Wallace won 67 coffees and 27 doughnuts within a few minutes by calculating the best time to play Roll up the Rim.

I played again. Another coffee. Then a doughnut. Then another coffee. It took nearly 15 minutes to play my 96 rolls.

I lost twice.

I didn't win a car or a TV or even a \$25 Tim Hortons gift card, but I didn't expect to. I won 67 coffees and 27 doughnuts—my statistical models predicted I'd win 66 and 28.

With 94 wins out of 96 entries, I had a 98% success rate. Before the contest, Tim Hortons estimated a player's chances were just 1-in-9, or about 11%. By finding predictable elements in a seemingly random process, I was able to dramatically shift those odds. I'll be telling my probability students about this for years.

Like Mohan Srivastava, I knew this wasn't a get-rich-quick scheme. Your odds of landing a big prize—even when almost every roll is a winner—are still incredibly small.

The contest rules suggest the value of my prizes if I buy the most expensive coffees and doughnuts possible is about \$500. Not a bad return on investment, but I'm going to see if Tim Hortons can donate the prizes to my local hospital instead.

By Michael Wallace, Assistant Professor, Department of Statistics and Actuarial Science, <u>University of Waterloo</u>

This article is republished from <u>The Conversation</u> under a Creative Commons license. Read the <u>original article</u>.

The story was also picked up by <u>Macleans</u>, <u>CTV Kitchener</u>, and many Metroland Media Group outlets across Ontario. Wallace was also interviewed on the <u>Lynda Steele show</u> on CKNW 980 Vancouver.

Credit: (Photo Illustration/The Conversation), CC BY-NC

<u>The Department of Statistics and Actuarial Science congratulates the department's newest University Research Chair</u>

WATERLOO



FACULTY OF MATHEMATICS Department of Statistics and Actuarial Science



The Department of Statistics and Actuarial Science at the University of Waterloo has appointed Professor Alexander Schied as University Research Chair.

Schied has written more than 60 papers focused on risk management, modelling and optimization in finance and economics, robustness and model uncertainty, and issues arising from market microstructure and price impact. Organizers of major conferences around the world including World Congress of the Bachelier Finance Society, General Advanced Mathematical Methods in Finance (AMaMeF) and Swissquote Conference, Society for Industrial and Applied Mathematics (SIAM) Conference on Financial Mathematics, and the Conference on Stochastic Processes and Their Applications have invited Schied to deliver a plenary lecture.

"Professor Schied is very deserving of the honour of being named as a University Research Chair," said **Stefan Steiner**, chair, Department of Statistics and Actuarial Science. "He has displayed excellence in the field of probability theory and stochastic analysis with applications to mathematical finance and economics."

In 2002, along with **Hans Föllmer**, Schied wrote the book *Stochastic Finance: An Introduction in Discrete Time*. The book was republished in 2011 and 2016. He has also been a member of several editorial boards and co-editor of *Finance and Stochastic*, as well as a member of several committees for conferences around the world.

Each year the university grants faculty members to a seven-year term as University Research Chair to recognize exceptional achievement and pre-eminence in a particular field of knowledge. The appointee can decide between a teaching reduction of one course per year or an annual stipend.

Richard Cook appointed University Professor

WATERLOO



FACULTY OF MATHEMATICS Department of Statistics and Actuarial Science

The University of Waterloo appointed Professor Richard Cook a University Professor in recognition of his outstanding research contributions in the field of biostatistics. Waterloo's designation of University Professor acknowledges exceptional scholarly achievement and international pre-eminence.

Cook is a Faculty of Mathematics research chair and a former Tier 1 Canada Research Chair in Statistical Methods for Health Research. A faculty member in the Department of Statistics and Actuarial Science for over 25 years, Cook holds cross-appointments to the School of Public Health and Health Systems in the Faculty of Applied Health Sciences and the Faculty of Health Science at McMaster University.

"Professor Cook is truly one of the world's current leading biostatisticians," said **Stefan Steiner**, chair of the Department of Statistics and Actuarial Science. "He has made groundbreaking contributions to many areas of biostatistics, including the analysis of life history data, statistical methods for missing and incomplete data, analysis of longitudinal data, and in the design and analysis of clinical and epidemiological studies."

Cook has coauthored more than 300 peer-reviewed scientific research papers and coauthored with **Jerry Lawless** two fundamental texts in biostatistics: *Multistate Models for the Analysis of Life History Data* and *The Statistical Analysis of Recurrent Events*. While Cook has made numerous advancements in statistical theory, his research is heavily motivated by real-world problems in public health and medicine. Through his network of collaborators, his work has advanced understanding of disease processes and treatment in the fields of rheumatology, transfusion medicine, and oncology.

He has received numerous awards and accolades. In addition to receiving an Outstanding Performance Award five times during his career at the University of Waterloo, Cook

- delivered the Scandinavian Journal of Statistics Keynote Lecture at the Nordic-Baltic Meeting of the International Biometrics Society (2019);
- received the SSC Gold Medal for his outstanding contributions to the field of statistics (2017);
- was named the 2016 Armitage Lecturer at the University of Cambridge;
- was named Fellow of the American Statistical Association (2008); and
- received the CRM-SSC Prize from the Centre de researches mathématiques (CRM) and the Statistical Society of Canada (SSC) for research accomplishments and excellence during the first 15 years following a doctoral degree (2007).

Cook is also known for his supervision and mentoring of graduate students and research fellows. In 2017 he received the Waterloo Award of Excellence in Graduate Supervision. He has supervised 23 doctoral students, four of whom were winners of the Pierre Robillard Award for the best Canadian PhD thesis in statistics or probability in their year of graduation. He has also supervised 11 postdoctoral fellows. Many of these former students and fellows are now themselves leading biostatistical researchers in universities and industry across Canada and internationally.

