Volume 51 • Issue 3

The members' magazine of the Institute of Mathematical Statistics

IMS Bulletin

April/May 2022

CONTENTS

Read it online: imstat.org/news

Carver Award: Xuming He

We are pleased to announce that **Xuming He** is the 2022 IMS Carver Medal recipient. Xuming He is the H.C. Carver Collegiate Professor of Statistics at the University of Michigan, and President-Elect (2021–2023) of the International Statistical Institute (ISI). He was selected for the Carver Medal, an award for service to the IMS, for "his decades-long contributions to the IMS in multiple capacities, including Editor of the *IMS Bulletin*, IMS Council member, Committee Chairs, and conference program co-chairs; and for his strong and conscientious leadership in a wide range of other professional services."

2022 Carver Award winner, Xuming He

Xuming He earned his PhD in statistics (1989) from the University of Illinois at Urbana-Champaign. He has made fundamental research contributions to statistics in the areas of quantile regression, semi-parametric and nonparametric models, robust statistics, asymptotic theory, data depth, subgroup analysis, data-adaptive inference, resampling methods, and statistical applications to concussion research, climate studies, bioinformatics, and educational testing. He has published over 160 papers in leading international journals. He has mentored and graduated more than 30 PhD students, who have taken positions in academia and various industry sectors worldwide.

Xuming has received numerous awards, including the Founders Award from the American Statistical Association (ASA, 2021), Distinguished Faculty Achievement Award from the University of Michigan (2021), International Chinese Statistical Association Distinguished Achievement Award (2015), ISI Service Award (2013), and IMS Medallion Lecture (2007), among many others. He is an elected member of ISI, and a Fellow of IMS, ASA, and the American Association for the Advancement of Science (AAAS). Xuming has made exceptional contributions to the statistical community, including serving as the co-editor of *JASA Theory & Methods* (2012–14), Chair of the Program Committee for the 59th ISI World Statistics Congress (2013), Chair of the ASA Committee on Meetings (2012–16), 2010 JSM Program Chair, and Program Director of Statistics at the US National Science Foundation (2003–05).

On receiving the news of his award, Xuming said, "As a lifetime member of the IMS, I really appreciate the opportunities that I have to work with so many of my colleagues from around the world to serve the IMS and our profession. I believe that if everyone gives a little bit of their time to the common good we can achieve a lot. The Carver Medal is especially dear to my heart; Harry Carver was an influential mathematical statistician in Michigan and I am really humbled by the honor."

The Carver Medal was created by the IMS in honor of Harry C. Carver, Founding Editor of the *Annals of Mathematical Statistics* and one of the founders of the IMS. The medal will be awarded to Xuming He at the 2023 JSM in Toronto.

IMS Bulletin

Volume 51 • Issue 3
April/May 2022
ISSN 1544-1881

Contact information

IMS Bulletin Editor: Tati Howell
bulletin@imstat.org

Managing Editor: Bob Keener

Contributing Editors: Radu Craiu, Anirban DasGupta, Yoram Gat, Ruobin Gong, David Hand, Takis Konstantopoulos, Xiao-Li Meng and Kavita Ramanan

Find us online:

w https://imstat.org/news
https://www.facebook.com/IMSTATI
https://twitter.com/InstMathStat

IMS Dues and Subscriptions Office

Contact the IMS regarding your dues, membership, subscriptions, orders or change of address:

t 877-557-4674 [*toll-free in USA*]
t +1 216 295 2340 [*international*]
f +1 216 295 5661
e dues.subs@imstat.org

IMS Business Office
Executive Director, Elyse Gustafson

Contact the IMS regarding any other matter, including advertising, copyright permission, offprint orders, copyright transfer, societal matters, meetings, fellows nominations and content of publications:

t 877-557-4674 [*toll-free in USA*]
t +1 216 295 2340 [*international*]
f +1 216 295 5661
e erg@imstat.org

Executive Committee

IMS Members' News

Royal Statistical Society Awards: Guy Medals

The UK Royal Statistical Society has announced its medals and prizes for 2022.

The **Guy Medal in Gold** is awarded to **Nancy Reid** for her pioneering work on higher-order approximate inference which provides a foundational basis for optimal information extraction from data, and has wide-ranging impact on the practice of data analysis. She has made numerous deep mathematical and methodological contributions to the theory of likelihood inference, in particular on the tangent exponential model and directional inference. She is known for her extensive service to the worldwide statistical community, and as a role model for future generations of statisticians. [*See also the COPSS DAAL announcement on page 4.*]

The **Guy Medal in Silver** is awarded to **Paul Fearnhead** for his numerous outstanding contributions to statistics, particularly in Bayesian computational statistics, changepoint methods, population genetics and inference for continuous-time stochastic processes.

The **Guy Medal in Bronze** is awarded to **Rajen Shah** for his pioneering research on large-scale statistical learning, including important contributions to goodness-of-fit testing, interaction search and conditional independence testing. His work is characterised by innovative methodological development, precise analysis and practical relevance, as evidenced in his five *JRSSB* papers.

The **Barnett Award** is awarded to **Ruth King**, who is internationally known for her work in statistical ecology, and in particular bringing to prominence in this field, the use of Bayesian statistics. She has worked with different types of data—notably capture-recapture and ring-recovery—as well as others such as count data and telemetry data. In addition, she has worked with various associated models, including integrated population models, state-space and hidden Markov-type models applied to ecological data. She has published more than 70 papers and co-authored two books. She has also played important roles in the National Centre for Statistical Ecology, including as deputy director, and has championed environmental and ecological statistics nationally and internationally.

The **David Cox Research Prize** is awarded to **Dominik Rothenhäusler** for his ground-breaking work on causality, robustness and replicability with wide-ranging implications on many practical problems.

Read more at https://rss.org.uk/news-publication/news-publications/2022/general-news/announcing-our-honours-recipients-for-2022/.

American Association for the Advancement of Science Fellows

The American Association for the Advancement of Science has announced its Class of 2021 Fellows. Among the 564 AAAS Fellows were eight in the Section on Statistics, five of whom (highlighted) are IMS members: **Veerabhadran Baladandayuthapani**, University of Michigan; **Valen Earl Johnson**, Texas A&M University; **Mingyao Li**, University of Pennsylvania; **Tapabrata (Taps) Maiti**, Michigan State University; **Sarah Nusser**, Iowa State University; **Annie Qu**, University of California, Irvine; **Nalini Ravishanker**, University of Connecticut; and **Luke Tierney**, University of Iowa.

Election as a Fellow of AAAS honors those whose efforts on behalf of the advancement of science or its applications in service to society have distinguished them among their peers and colleagues.

The complete list is at https://www.aaas.org/page/2021-fellows.

IMS Members' News

SIAM Early Career Prize in Data Science

Weijie Su, assistant professor of statistics at the University of Pennsylvania, has received the inaugural Early Career Prize in Data Science awarded by the Society for Industrial and Applied Mathematics (SIAM). It recognizes an outstanding early career researcher in the mathematics of data science (https://www.siam.org/prizes-recognition/activity-group-prizes/detail/siag-data-early-career-prize). Weijie will give a talk at the 2022 SIAM conference on Mathematics of Data Science (MDS22: September 26–30, 2022, in San Diego). His citation is for "outstanding contributions to the theoretical and computational foundations of data science."

Steven Moen wins ASA Student Travel Award

IMS student member **Steven T. Moen**, PhD student at the University of Wisconsin, won the ASA John J. Bartko Scholarship Award to attend the Conference on Statistical Practice (which was virtual this year, in February). Steven decided to study economics in college but especially the models and systems used in the field, which led him to statistics. His focus is on financial statistics and studying the theoretical aspects of time-dependent data. Currently, Moen is collaborating with Russell Green, a senior economist at the International Monetary Fund, on a paper about the market microstructure of India's sovereign bond market, which could help prevent another currency crisis.

O'Malley appointed to new role at TDI

James O'Malley was recently appointed to the Peggy Y. Thomson Professorship in the Evaluative Clinical Sciences; he is the first statistician in this role. O'Malley is a professor in The Dartmouth Institute for Health Policy & Clinical Practice (TDI) and department of biomedical data science in the Dartmouth Geisel School of Medicine. He will use the award to advance cutting-edge research at the intersection of statistics and health services research. O'Malley's research work includes social network analysis, hierarchical modeling, causal inference, and Bayesian analysis with problems of interest often being motivated by compelling and complex issues in health services and health policy.

Lawrence D. Brown PhD Student Awards

Apply now!

Are you a PhD student? Apply by May 1, 2022, for the 2023 Lawrence D. Brown PhD Student Award. This annual travel award will be given to three PhD students, who will present their research at a special invited session during the 2023 IMS Annual Meeting (at the JSM in Toronto, August 5–10, 2023).

https://imstat.org/ims-awards/ims-lawrence-d-brown-ph-d-student-award/

NOMINATIONS OPEN FOR 2023 INTERNATIONAL PRIZE

The International Prize in Statistics is awarded every two years to an individual or team *"for major achievements using statistics to advance science, technology, and human welfare."* Previously awarded to Sir David Cox (2017), Brad Efron (2019) and Nan Laird (2021). Nominate the 2023 winner: https://statprize.org/nominations.cfm. Deadline August 15.

= access published papers online

IMS Journals and Publications

Annals of Statistics: Enno Mammen, Lan Wang
https://imstat.org/aos
https://projecteuclid.org/aos

Annals of Applied Statistics: Ji Zhu
https://imstat.org/aoas
https://projecteuclid.org/aoas

Annals of Probability: Christophe Garban, Alice Guionnet
https://imstat.org/aop
https://projecteuclid.org/aop

Annals of Applied Probability: Kavita Ramanan, Qiman Shao: https://imstat.org/aap
https://projecteuclid.org/aoap

Statistical Science: Sonia Petrone
https://imstat.org/sts
https://projecteuclid.org/ss

IMS Collections
https://projecteuclid.org/imsc

IMS Monographs and *IMS Textbooks:* Mark Handcock
https://www.imstat.org/journals-and-publications/ims-monographs/

IMS Co-sponsored Journals and Publications

Electronic Journal of Statistics: Grace Yi & Gang Li
https://imstat.org/ejs
https://projecteuclid.org/ejs

Electronic Journal of Probability: Bénédicte Haas
https://projecteuclid.org/euclid.ejp

Electronic Communications in Probability: Siva Athreya
https://projecteuclid.org/euclid.ecp

Journal of Computational and Graphical Statistics: Galin Jones, Faming Liang https://www.amstat.org/ASA/Publications/Journals.aspx
log into members' area at imstat.org

Probability Surveys: Mikhail Lifshits
https://imstat.org/ps
https://www.i-journals.org/ps/

Statistics Surveys: Marloes Maathuis
https://imstat.org/ss
https://projecteuclid.org/euclid.ssu

IMS-Supported Journals

ALEA: Latin American Journal of Probability and Statistics: Daniel Remenik
http://alea.impa.br/english

Annales de l'Institut Henri Poincaré (B): Giambattista Giacomin, Yueyun Hu
https://imstat.org/aihp
https://projecteuclid.org/aihp

Bayesian Analysis: Mark Steel
https://projecteuclid.org/euclid.ba

Bernoulli: Davy Paindaveine
https://www.bernoulli-society.org/
https://projecteuclid.org/bj

Brazilian Journal of Probability and Statistics: Mário de Castro
https://imstat.org/bjps
https://projecteuclid.org/bjps

IMS-Affiliated Journals

Observational Studies: Nandita Mitra
https://https://obs.pennpress.org/

Probability and Mathematical Statistics: Krzysztof Bogdan, Krzysztof Dębicki
http://www.math.uni.wroc.pl/~pms/

Stochastic Systems: Shane Henderson
https://pubsonline.informs.org/journal/stsy

Nancy Reid: 2022 COPSS Distinguished Lecturer

Rebecca W. Doerge, Carnegie Mellon University, is the COPSS Distinguished Achievement Award Committee Chair. Rebecca, together with Christian Genest and Erica E.M. Moodie, McGill University, writes:

The Committee of Presidents of Statistical Societies (COPSS) has selected **Nancy Reid**, OC, FRS, FRSC, Professor in the Department of Statistical Sciences at the University of Toronto, to give the Distinguished Achievement Award Lectureship at the 2022 Joint Statistical Meetings in Washington, DC. Her talk is tantalizingly entitled *"Likelihood and its Discontents."*

Nancy Reid's work has had major impact in the development of statistical theory. She has made unique contributions to the linking of modern themes and traditional concepts in statistical science. As noted by one of her supporters, "she has contributed fundamental and path-breaking work in a wide range of statistical problems, including nonparametric estimation for survival data, applications of differential geometry to statistics, conditional inference, profile and composite likelihood methods, higher order asymptotics, connections between Bayes and frequentist methods, and... the list goes on." Reid's work is wide-ranging and she has shown a striking aptitude for focusing on problems with a high practical impact and reward. The clarity of her writing and her attention to detail have also enhanced her lifetime interest in bringing statistical thinking to non-specialists.

Reid studied at the University of Waterloo (BMath 1974), the University of British Columbia (MSc 1976), Stanford University (PhD 1979), and Imperial College, London (PDF 1980). She joined the University of Toronto in 1986 from the University of British Columbia. She has held several leadership roles in statistical science, including Editor or Associate Editor for several leading journals, Chair of the Department (1997–2002), President of the Institute of Mathematical Statistics (1996–97), Vice President of the International Statistical Institute (1999–2001), President of the Statistical Society of Canada (2004–05), and Director of the Canadian Statistical Sciences Institute (2015–19).

Reid's early research on bivariate influence functions and functional expansions provided theoretical and practical tools for the analysis of censored data. Her 1987 paper with the late Sir David Cox on orthogonal parameters and approximate conditional inference, read to the Royal Statistical Society, has been very influential. Her work has led to new approximation techniques and to a deeper understanding of the foundations of statistical inference. The author of numerous books and papers, she maintains an active research profile focused in part on the investigation of the relationship between significance functions and Bayesian posterior distributions, and generalized fiducial inference and inferential models.

Nancy Reid

Among her many awards, Reid was the first woman to receive the COPSS Presidents' Award (1992), the first recipient of the Canadian Mathematical Society's Krieger–Nelson Prize (1995), IMS Wald Lecturer in 2000, and the 2009 Gold Medalist of the Statistical Society of Canada. In 2016, the Royal Statistical Society awarded her the Guy Medal in Silver. [*Reid has just been awarded the Guy Medal in Gold: see page 2.*]

Reid is a Fellow of the American Statistical Association, the Institute of Mathematical Statistics, and the American Association for the Advancement of Science. She is a corresponding Fellow of the Royal Society of Edinburgh, a Foreign Associate of the US National Academy of Sciences and, in 2015, she was appointed Officer of the Order of Canada. Her authoritative contributions to the theory of statistical inference, her commitment to excellence in statistical applications, and her outstanding service to the community make her a most apt recipient of the COPSS Distinguished Achievement Award and Lectureship.

COPSS Leadership Academy welcomes eight new members

The COPSS Leadership Academy recognizes early-career statistical scientists who show evidence of and potential for leadership and who will help shape and strengthen the field. Eight new members of the Academy have been announced. They are:

Xi Chen, New York University: *For notable contributions to statistical inference for distributed, online, and high-dimensional data, to stochastic optimization, and to statistical applications in business domains, for outstanding educational efforts to the next generation of business leaders, and for significant industrial impacts.*

Natalie Dean, Emory University: *For ground-breaking, high-impact work in the development of innovative study designs and analyses for evaluating novel*

Continues on **page 5**

COPSS Awards continued

COPSS Leadership Academy continued from page 4

vaccines, and for wide-reaching public engagement and thought leadership during the COVID-19 pandemic.

Davina Durgana, Oxford University and Walk Free: *In recognition of outstanding service through pioneering scientific research and policy communication to end human trafficking, and advocacy efforts to encourage youth and women in STEM fields across the world.*

Philip Ernst, Rice University: *For significant contributions of extraordinary merit to applied probability and mathematical statistics, particularly the resolution of the longstanding conjecture of Yules' nonsense correlation, outstanding teaching and leadership.*

Pierre Jacob, ESSEC Business School: *For path-breaking contributions to Monte Carlo algorithms and Bayesian statistics, and for exemplary dedication as teacher and mentor.*

Kristian Lum, Twitter: *For strong contributions to the study of algorithmic fairness and ethics in statistics and data science; for high-quality, high-impact collaboration and advocacy in social justice; and for leadership in maximizing our profession's openness and support of all members.*

Lester Mackay, Microsoft Research: *For significant contributions to the theory and practice of statistical machine learning.*

Betsy Ogburn, John Hopkins Bloomberg School of Public Health: *For creative methodological innovations in causal and network analysis; for contributions to education; for generous service to the profession and society, including leadership in addressing the COVID-19 pandemic.*

2022 Elizabeth L. Scott Award and Lectureship: Madhu Mazumdar

The COPSS 2022 Elizabeth L. Scott Award and Lectureship recipient is **Madhu Mazumdar**, Director of the Institute for Healthcare Delivery Science at the Mount Sinai Health System and is a Professor of Biostatistics at the Center of Biostatistics, Department of Population Health Science and Policy.

Madhu Mazumdar developed methodologies for detecting publication bias in meta-analysis; for adjusting selection bias in clinical trials; for allowing interim looks at data in clinical trials comparing diagnostic tests; for developing and validating quality of life questionnaire; and for estimating misclassification rate of responders when oncologic response criteria were changed. Her collaborative research resulted in personalized treatment regimen for various cancer types and orthopedic surgeries. Her work also changed guidelines for staging cancer and practice guidelines for use of anesthetics. She developed innovative interdisciplinary educational and research programs, in collaboration with various clinical departments that increased productivity through grantsmanship and augmented clinical revenue through improved care delivery.

Dr. Mazumdar's Scott lecture is titled *"Biostatistical methods and team science: Generating evidence for optimization of clinical practice"*. She will speak on innovative statistical applications to catalyze healthcare delivery, addressing two specific challenges facing US healthcare: 1) how to choose patients for knee-replacement surgery who will benefit most in terms of their quality of life and what is the cost-effectiveness of this procedure? and 2) how to improve quality of cancer care through modelling of incurred cost? Dr. Mazumdar will highlight the critical role of statistical methods in answering these questions and will illustrate how collaborations, guided with principles of team science, provide opportunities for practicing leadership, embracing diversity, managing conflict, and sharing credit.

IMS Statement about Ukraine

The IMS strongly condemns the unprovoked Russian aggression towards Ukraine and its people, leading to an unjust war that is creating a deplorable humanitarian crisis and causing deaths, injuries, and displacement of innocent civilians, including colleagues and their families in the Ukrainian mathematical, statistical, and scientific communities.

The IMS is committed not only to pursuing mathematical, statistical, and scientific knowledge, but also to improving its human side. We stand with all IMS members, especially those from Ukraine and Russia, who are trying to make the world a better, safer, and peaceful place, in ways big and small.

Furthermore, the IMS fully endorses the following joint statement by the National Academies of G7 States and published on 2 March 2022:

"The unprovoked attack against Ukraine, a democratic and independent country, is a blatant violation of international law and of core values of humanity. The Russian invasion is an assault on the fundamental principles of freedom, democracy and self-determination, which provide the basis for academic freedom and opportunities for scientific exchange and cooperation.

In this dark hour, our thoughts and deepest sympathy are with the people of Ukraine. We are determined to support the National Academy of Sciences of Ukraine. We stand in solidarity with the scientific community and the scientists in Ukraine.

We acknowledge the Russian scientists and citizens who are ashamed of this attack and speak out against the war.

We call on the Russian leadership to immediately cease all military action against Ukraine and put an end to this war."

Special IMS Lecture Previews

Rietz Lecture: Hans-Georg Müller

Hans-Georg Müller is Professor of Statistics at the University of California, Davis. While in high school, he received the First Prize in the National Mathematics Competition of Germany (Bundeswettbewerb Mathematik) in 1974 and 1975, followed by a PhD in Mathematics from the University of Ulm in Germany and a MD from the University of Heidelberg. After deliberating whether to pursue a career in medicine or statistics, he accepted an appointment as Hochschulassistent (non-tenure track Assistant Professor) in Biostatistics at the University of Marburg in 1984 and was appointed as Associate Professor with tenure at the University of Erlangen-Nürnberg in 1987. He joined the faculty at UC Davis in 1988 as Associate Professor, was promoted to Professor in 1990 and to Distinguished Professor in 2007. His most cherished part of the job is to mentor PhD students and postdocs; 34 students have completed their PhD under his supervision. He has his best ideas while hiking in the wilderness of Northern California and scrambling in the Sierra Nevada mountain range.

He has served as chair of the Department of Statistics, founding chair of the Graduate Program in Biostatistics at UC Davis and co-editor of *Statistica Sinica*. Over the years he has been engaged in various biomedical research consortia, aiming to quantify brain and neurocognitive development from brain imaging data, to sequence the wheat genome, and to model aging, longevity and human mortality, with support from NIH, NSF and the Bill and Melinda Gates Foundation. His contributions to statistics include early work on smoothing methods for nonparametric regression and density estimation, methodology for growth curves and change-points, and semiparametric and structured modeling in regression. He then devoted major efforts to develop the theoretical and methodological foundations for functional data analysis, notably functional principal component analysis, empirical dynamics, time warping and functional regression, and also to build a bridge between functional and longitudinal data analysis, aligning these two areas. In his recent research he is developing concepts and methods for the emerging field of statistical analysis and inference for random objects, including distributional data analysis and transport regression. He gave an IMS Medallion Lecture in 2007, received the Senior Noether Scholar Award for Nonparametric Statistics in 2016; he is an elected member of ISI, and a fellow of IMS, ASA and the American Association for the Advancement of Science.

Hans will give this Rietz Lecture at the IMS Annual Meeting in London, June 27–30, 2022.

The Emerging Field of Random Objects and Metric Statistics

Random Objects are random variables that take values in a separable metric space, where vector operations such as addition and scalar multiplication are generally not available. Examples that are practically relevant for data analysis include samples of distributions, covariance matrices and covariance surfaces, networks and trees. With the advent of ever more advanced data recording and processing tools, such complex data are increasingly encountered, while available statistical methodology has been lacking. A special case of interest are random objects that take values in distribution spaces equipped with the (sliced) Wasserstein or the Fisher–Rao metric.

The emerging new field for the analysis of such random objects can be characterized as metric statistics. It is an area with many open and challenging problems, due to the absence of the vector space structure that underpins classical and high-dimensional statistics. Its emphasis on random objects and metric methods differentiates it from object-oriented data analysis, which was the theme of a 2010 Workshop at the Statistical and Applied Mathematical Sciences Institute (SAMSI) in North Carolina. This workshop was jointly organized by Steve Marron, Hans-Georg Müller and Jane-Ling Wang and brought the need for further research in this general area to the fore, stimulating many of the subsequent developments.

Techniques that have been developed for complex data in vector spaces, notably functional data, where the random elements are usually assumed to be square integrable and smooth random functions in a Hilbert space, are generally not applicable for random objects due to the lack of linearity. However, these techniques provide guidance of what notions to aim for (there are also some limited scenarios where functional data analysis can be more directly applied for the analysis of random objects after a linearization step). Accordingly, in the statistical analysis of random objects one aims at notions of means (barycenters, generally defined as Fréchet means), variances (Fréchet variance), conditional means and ensuing regression models (conditional Fréchet means), visualization and

Continues on **page 7**

inference (two- and multi-sample tests).

A baseline scenario in metric statistics is the situation where the available information is limited to the distances between pairs of random objects. This scenario motivates the classical Fréchet mean and more recent extensions to conditional Fréchet means, with local, global and penalized versions of Fréchet regression providing specific implementations for the case of Euclidean predictors. Asymptotic properties can be obtained through empirical process theory. The study of Fréchet regression and more generally regression for random objects is far from complete and the presentation will include some of our recent results along these lines.

When random objects reside in geodesic metric spaces one can do more. Geodesic metric spaces of special interest include distribution spaces with Wasserstein or Fisher–Rao metrics. In these and other geodesic spaces the geodesics that connect random objects define geodesic transports that are optimal transports in the special case of the Wasserstein space. We recently introduced a transport algebra that gives rise to an intrinsic transport regression. Some other recent developments will also be discussed, including notions of depth and visualization for random objects.

This presentation is based on joint research with Satarupa Bhattacharjee, Yaqing Chen, Xiongtao Dai, Paromita Dubey, Jianing Fan, Álvaro Gajardo, Zhenhua Lin, Alexander Petersen and Changbo Zhu. Relevant preprints include arXiv:2006.13548, arXiv:2105.05439.

Neyman Lecture: Heping Zhang

Heping Zhang is Susan Dwight Bliss Professor of Biostatistics, Professor of Child Study, and Professor of Statistics and Data Science at Yale University. He has published over 350 research articles and monographs in theory, methodology, and applications of statistics. He is particularly interested in biomedical research including epidemiology, genetics, child and women health, mental health, and substance use. He directs the Collaborative Center for Statistics in Science that coordinates major national research networks to understand the etiology of pregnancy outcomes and to evaluate treatment effectiveness for infertility.

Zhang is a fellow of the American Statistical Association and a fellow of the Institute of Mathematical Statistics. He was named the 2008 Myrto Lefkopoulou distinguished lecturer by Harvard School of Public Health and a 2011 IMS Medallion Lecturer. Dr. Zhang was the founding Editor-in-Chief of *Statistics and Its Interface* and is the past coordinating Editor of the *Journal of the American Statistical Association.*

Heping Zhang's Neyman Lecture will be given at the IMS Annual Meeting in London, June 27–30, 2022.

Genes, Brain, and Us

Many human conditions, including cognition, are complex and depend on both genetic and environmental factors. After the completion of the Human Genome Project, genome-wide association studies have associated genetic markers such as single-nucleotide polymorphisms with many human conditions and diseases. Despite the progress, it remains difficult to identify genes and environmental factors for complex diseases—the so-called geneticist's nightmare. Furthermore, although the impact of these discoveries on human health is not shock and awe, "drugs with support from human genetic studies for related effects succeed from phase 1 trials to final approval twice as often as those without such evidence." Therefore, it is important and promising, while challenging, to identify genetic variants for complex human health-related conditions.

Many of us have devoted a tremendous amount of effort, or even our entire careers, to develop statistical theory and methods to meet this challenge. This talk is not intended to provide a comprehensive review of the massive progress of related methods and discoveries. Instead, I will focus on some of the work that many of my students assisted me in, over the past several years.

The first area is the identification of super-variants. A super-variant is a set of alleles in multiple loci of human genome although unlike the loci in a gene, contributing loci to a super-variant can be anywhere in the genome. The concept of super-variant follows a common practice in genetic studies by the means of collapsing a set of variants,

Continues on **page 8**

Neyman Lecture: Heping Zhang *continued*

specifically single nucleotide polymorphisms. The novelty and challenge lie in how to find, replicate, interpret, and eventually make use of the super-variants. Our work has been mainly based on the use of tree- and forest-based methods, and a data analytic flow that we proposed in 2007, which in retrospect resembles the spirit of "deep learning" that Hinton coined in 2006.

The second area is our progress in conducting statistical inference for high dimensional and structured data objects. Not only do such data objects more and more commonly appear in imaging genetic studies, but also in other areas of data science including artificial intelligence. They do not belong to a Euclidean space for which most of the statistical theory and methods such as the distribution function are developed. How do we analyze data objects in non-Euclidean spaces?

I will highlight the concepts and important properties of ball covariance and divergence, which, respectively, act as a measure of dependence between two random objects in two possibly different Banach spaces, and as a measure of difference between two probability measures in separable Banach spaces. As an example, I will demonstrate how the ball divergence was applied to study genetics of brain volumetric phenotypes.

To develop the concept of the distribution function in a metric space, we introduce a class of quasi-distribution functions, or metric distribution functions. We lay the foundation for the use of metric distribution functions by establishing the critical correspondence theorem, the Glivenko–Cantelli property, and the Donsker property for the metric distribution functions in metric spaces. The randomness of data objects can then be assessed by the distribution of metric between random object and a fixed location. Finally, the metric distribution function can play a similar role to the classic distribution function statistical inference including homogeneity test and mutual independence test for non-Euclidean random objects.

In summary, I hope to present the clear and strong motivation from the genetic and neuroimaging studies for the development of our statistical methodology and theory, share intuition and/or solid theory foundation, and demonstrate the broad and great utility of the methods to be touched on. Importantly, I hope to underscore the importance of interplay among applications, methodology, and theory in statistical research.

Schramm Lecture: Russell Lyons

Russell Lyons obtained his PhD in harmonic analysis at the University of Michigan in 1983. Within a few years, some coincidences led him to switch to probability, where he has been happily ever since. Lyons is James H. Rudy Professor of Mathematics and Adjunct Professor of Statistics at Indiana University. His primary area of research is discrete probability and its connections to other areas of mathematics, including ergodic theory, geometric group theory, analysis, geometry, and combinatorics. He is also very interested in the teaching of statistics and has done some research in statistics. Lyons was a Sloan Foundation Fellow, a Visiting Miller Research Professor, an IMS Medallion Lecturer, an Invited Speaker at the International Congress of Mathematicians, and a Simons Foundation Fellow, and gave an Hour Address at the Joint Mathematics Meetings. He is a Fellow of the American Mathematical Society.

This IMS/Bernoulli Society Schramm Lecture will be given at the IMS London meeting in June.

Monotonicity in Continuous-Time Random Walk

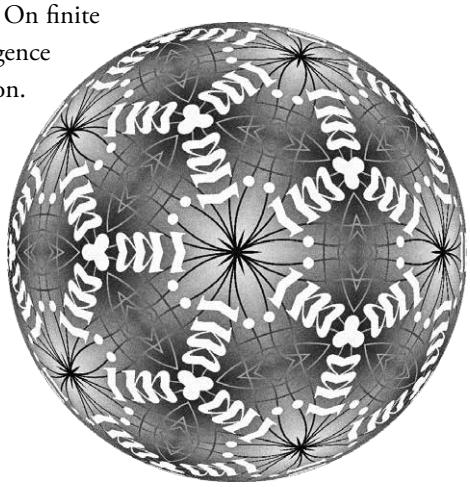
Suppose that $r_n > 0$ for all non-negative integers n. This determines a symmetric birth-and-death chain with transition rates r_n from n to $n+1$ and from $n+1$ to n. That is, there are Poisson processes with rates r_n such that when the Markov chain is at n, it moves to $n-1$ or $n+1$ at the next time that one of the corresponding Poisson processes has an event. Consider how long it takes the chain, starting from 0, to reach n for the first time. Is this stochastically decreasing in the set of rates? In other words, if any rate is increased, will this tend to make the time to reach n decrease? The fact that transitions occur faster suggests that the answer is yes, but the subtlety is that transitions are more likely in each direction where the rate is increased, not only in the direction that increases the location. Nevertheless, the first intuition is correct: the time does decrease stochastically. This follows from a result of Karlin and McGregor (1957).

This is but one example of a kind of monotonicity in the rates for (variable-speed) continuous-time random walk on graphs. We will survey results, counterexamples, and open questions. We will give general ideas of proofs, but avoid technicalities.

Suppose that G is a graph with positive rates on its edges, used for Poisson processes: When a walker is at a vertex, it moves to a neighbor at the time of the next event that occurs for the corresponding incident edges. The transition probability of being back at the starting vertex at time t is well known to be decreasing in t. Now, increasing t by a factor of c is equivalent to fixing t and increasing all rates by a factor of c. But what if we fix t and increase only one rate? Then this return probability *need not* be decreasing. However, Benjamini and Schramm (2005) proved that when G is finite, if we average the return probabilities over all starting vertices, then we do get a smaller number.

Both this result and that above on birth-and-death chains depend on a spectral comparison for Laplacian matrices. For return probabilities, there have been many extensions, including to infinite graphs. Suppose that Γ is a group generated by some finite subset, S. Let G be the corresponding Cayley graph, meaning that the vertices of G are the elements of Γ and the edges are $\{x, xs\}$ for $x \in \Gamma$ and $s \in S$. Let $r_s > 0$ for $s \in S$. If the rate on every edge of the form $\{x, xs\}$ is r_s, then the return probability at time t is the same for all $x \in \Gamma$. Thus, when Γ is finite, there is no need to average. One of the extensions alluded to is that we do have monotonicity for all Cayley graphs, without averaging, even for infinite Cayley graphs. In many cases, we also have monotonicity for the *expected* return probability on a Cayley graph when the edge rates form a random field whose law is invariant under left group multiplication, and we are comparing two random fields where one has edge rates at least those of the other. But whether this same phenomenon holds for *all* Cayley graphs is an open question due to Fontes and Mathieu (2006), even on regular trees.

Most of our talk will be devoted to other types of monotonicity on Cayley graphs. On infinite graphs, we may ask about the limiting linear rate of escape, i.e., the limit of the distance divided by the time. Does this increase when the rates are increased? On finite graphs, we may ask about the convergence to the stationary (uniform) distribution. Does this happen faster when the rates are increased? It turns out that both questions have surprising answers. The figure, based on the IMS logo, has a symmetry group where monotonicity occurs in all expected ways. This is new work of the speaker with Graham White.



Medallion Lecture: Rina Foygel Barber

Erielle Bakkum Photography

Rina Foygel Barber is the Louis Block Professor of Statistics at the University of Chicago. Her research focuses on problems in distribution-free inference, multiple testing, sparse and low-rank methods, and medical imaging. She received an M.S. in Mathematics in 2009 and a PhD in Statistics in 2012 from University of Chicago, and was a postdoctoral research fellow at Stanford University in 2012–13 before joining the faculty at University of Chicago in 2014. She is the recipient of a Sloan Research Fellowship (2016), an NSF CAREER award (2017), the Tweedie New Researcher Award (2017), the Peter Gavin Hall IMS Early Career Prize (2020), and the COPSS Presidents' Award (2020).

Rina's Medallion Lecture will be given at the IMS London meeting in June: see https://www.imsannualmeeting-london2022.com/

Distribution-free prediction: exchangeability and beyond

Distribution-free prediction is a recently developed field in statistics that seeks to provide predictive inference for the output of any estimation algorithm, without requiring assumptions on the data distribution or on the underlying algorithm. Given a regression algorithm that produces a fitted model $\hat{\mu}(x)$ to predict $Y \mid X = x$, the goal is to construct a prediction interval that is valid (i.e., has a $1-\alpha$ coverage probability) regardless of the algorithm we used for fitting $\hat{\mu}$, and regardless of whether the model $\hat{\mu}$ actually fits well to the data distribution. Conformal prediction, pioneered by Vladimir Vovk and collaborators beginning in the late 1990s, provides exactly this type of guarantee. It can be paired with any model fitting algorithm to provide distribution-free predictive coverage, as long as the training and test data are drawn from the same distribution.

In this talk, I will describe two recent extensions to the conformal prediction framework. First, I will present the jackknife+ and CV+ procedures. Traditionally, cross-validation based prediction intervals (in particular, leave-one-out cross-validation, which is also known as the jackknife) generally provide the right coverage level empirically, but theoretical guarantees for these methods require an assumption of algorithmic stability (namely, the predicted value $\hat{\mu}(X_{n+1})$ is stable to perturbations of the training data, e.g., removing one training point). In our work, we propose a modification to this method called the jackknife+, or more generally K-fold CV+ as a modification to K-fold cross-validation, which is closely related to the cross-conformal predictors of Vovk and collaborators. Interestingly, the jackknife+ is guaranteed to provide $1-2\alpha$ coverage in the worst case (rather than the target level $1-\alpha$), but under stability conditions, both jackknife and jackknife+ result in $\approx 1-\alpha$ coverage. This method offers a compromise between conformal prediction, which may be computationally infeasible in many large-scale settings, and hold-out set methods (i.e., split conformal prediction), which are highly computationally efficient but result in wider prediction intervals due to data splitting.

Like conformal prediction, the jackknife+ relies on the assumption that the training and test data are drawn from the same distribution—that is, the training and test data points are exchangeable. Moreover, both methods assume that the regression algorithm used for fitting $\hat{\mu}$, while arbitrary, must treat the training data points symmetrically. In some applications, both of these conditions may be too restrictive—we may suspect that our data is not exchangeable due to phenomena such as distribution drift, and moreover we may also wish to fit models $\hat{\mu}$ using regression methods that do not treat data points symmetrically in order to correct for this potential drift. The second extension I will describe is a new framework for non-exchangeable conformal methods (including non-exchangeable versions of conformal prediction, split conformal prediction, and the jackknife+), where both of these assumptions are relaxed. First, a small randomization step in the method allows for regression algorithms that are not symmetric in the training data points $i = 1, \ldots, n$, with no resulting loss of coverage if the data points are indeed exchangeable. Next, if in fact the data points are not exchangeable (e.g., due to distribution drift over time), then placing weights on the n training data points before running the method enables the final prediction interval to be robust to these changes in the distribution. In combination, these two new properties allow for conformal type methods that can be deployed far beyond the exchangeable regime, using non-symmetric algorithms for more accurate estimation in distribution drift type settings, and providing robust predictive validity guarantees in the non-exchangeable setting as well.

This work is joint with Emmanuel Candès, Aaditya Ramdas, and Ryan Tibshirani.

Medallion Lecture: Dylan Small

Dylan Small received his PhD in Statistics in 2002 from Stanford University with Tze Leung Lai as his thesis advisor. Dylan is the Universal Furniture Professor of Statistics and Data Science in the Wharton School of the University of Pennsylvania and is currently the chair of the Department of Statistics and Data Science. His research focuses on causal inference and applications of statistics to public health and public policy. He was the founding editor of the journal *Observational Studies*. Dylan has advised 28 PhD students on their dissertations, and has mentored several undergraduates and postdoctoral fellows on research.

Dylan Small's Medallion Lecture will be given at JSM2022, August 6–11, 2022, in Washington DC.

Protocols for Observational Studies: Methods and Open Problems

For learning about the causal effect of a treatment, a randomized controlled trial (RCT) is considered the gold standard. However, randomizing treatment is sometimes unethical or infeasible, and instead an observational study may be conducted. Three sources[1] of the strength of a properly designed RCT are:

(1) randomization – by randomly assigning people to treatment or control, an RCT creates a fair comparison between the two groups;
(2) identical processes – a properly designed RCT takes pains to apply to the treatment and control group all other processes in the same way such as adjuvant therapies, follow-up and outcome assessment; and
(3) protocol – a properly designed RCT is driven by a strong protocol with pre-specified hypotheses about pre-specified outcomes and a pre-specified analysis.

The first source of strength, randomization, is unique to an RCT. But the second and third sources of strength have nothing to do with randomization and can be made part of a well-designed observational study. This talk is about making the third source, a protocol, a part of an observational study. We will illustrate the value of protocols for observational studies in three applications – the effect of playing high school on later life mental functioning, the effect of police seizing a gun when arresting a domestic violence suspect on future domestic violence and the effect of mountaintop mining on health – and discuss methodologies for observational study protocols and open problems.

Investigators may be concerned that writing down too strict a protocol will limit their ability to learn from the data through exploratory data analysis and to make study design choices that maximize power. The power of a study depends on features of the study design, including how much a test statistic emphasizes tail observations, what to do about outliers, how different subgroups are incorporated into the analysis and how different outcomes are incorporated into the analysis. The best choices for these features often depend on aspects of the population about which one is uncertain before obtaining data. Ideally one would like to adapt the study design to the data but if there is no protocol which prespecifies the analysis, the researcher's biases (conscious or subconscious) can influence the results. As the psychologist Fred Emery said, "Instead of constantly adapting to change, why not change to be adaptive?"

An adaptive protocol seeks to allow for adapting to the data while protecting against researcher bias. We will discuss three strategies for designing adaptive protocols:

(i) specifying up front the multiple test statistics, subgroups, outcomes etc. that will be considered and using a multiple testing procedure that controls an appropriate error rate (e.g., family-wise Type I error rate or false discovery rate);
(ii) splitting the sample, using part of the sample to make design choices and the other part of the sample for analysis using these design choices; and
(iii) looking at secondary aspects of the data to make design choices.

While many of the considerations for designing a protocol for an RCT and an observational study are the same, there are differences. We will discuss some of these differences, such as how power comparisons between various multiple testing procedures differs between RCTs and observational studies.

[1] Moses, L. E. (1995). Measuring effects without randomized trials? Options, problems, challenges. *Medical Care*, AS8–AS14.

IMS Lawrence Brown PhD Student Award Lecture: Rungang Han

Rungang Han is currently a postdoctoral associate in the Department of Statistical Science at Duke University, advised by Professor David Dunson and Professor Anru Zhang. Prior to that, he obtained his PhD in Statistics in 2021 from the University of Wisconsin-Madison, under the supervision of Professor Anru Zhang. He received a BS in Statistics in 2017 from Zhejiang University, China. He is interested in methodology and theory in high-dimensional statistics, Bayesian statistics, machine learning and non-convex optimization. His recent interest focuses on large-scale statistical matrix/tensor inference.

Rungang Han is one of the three winners of the Lawrence D. Brown PhD Student Awards, and will be presenting this lecture in a special session at the 2022 IMS Annual Meeting in London, UK, June 27–30, 2022. See https://www.imsannualmeeting-london2022.com/special-sessions for more information about the award session.

Exact Clustering in Tensor Block Model: Statistical Optimality and Computational Limit

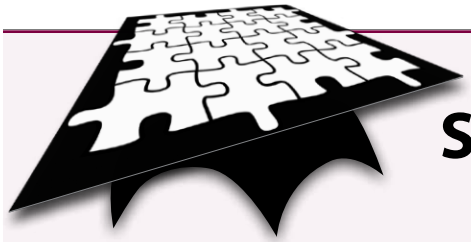
In recent years, the analysis of tensors or high-order arrays has emerged as an active topic in statistics, applied mathematics, machine learning, and data science. Datasets in the form of tensors arise from various scientific applications, such as collaborative filtering, neuroscience, hyper-spectral imaging, longitudinal data analysis, and more. Despite many recent celebrated results in statistical tensor data analysis such as tensor regression, tensor completion and tensor PCA, another important model named multi-way tensor block model has not been well studied yet. This model naturally generalizes the classic multivariate clustering models to multi-dimensional scenarios where one assumes the indexes along each dimension can be partitioned into several clusters.

While clustering analysis is prevalent in discovering heterogeneous patterns in usual multivariate data, it has unique challenges when the data is organized as a multi-dimensional array. Several extensions to high-order clustering have been developed in recent years and they typically fall into two types with different limits. On one hand, the spectral or MLE-based methods that can achieve statistical optimality are usually computationally intractable; On the other hand, polynomial-time efficient algorithms that aim at surrogate objectives (such as convex relaxation) usually suffer from the statistical sub-optimality. The tradeoff between statistical optimality and computational efficiency usually comes up in modeling low-rank tensors as it has fairly complicated algebraic structures.

In this work, we design a two-stage non-convex polynomial-time efficient procedure for the task of high-order clustering under tensor block models. At the first stage, we proposed a High-order Spectral Clustering (HSC) to initialize the clusters estimation; at the second stage, we develop an iterative scheme called High-order Lloyd (HLloyd) to further refine the clusterings. The HSC algorithm can be taken as a tensor-version spectral method and with power iterations; while the HLloyd algorithm can be seen as a high-order extension of Lloyd algorithm for 1-dimensional K-means clustering to order-d clustering. We then provide the non-asymptotic statistical properties of these two algorithms and establish the clustering error rates for the full procedure. Our proof technique features a new perturbation analysis on low-rank tensor estimation without any singular value gap assumptions.

Apart from the newly proposed algorithm and its theoretical guarantee, we discover an intriguing interplay between statistical optimality and computational efficiency of high-order clustering. Specifically, we introduce a signal-to-noise ratio (SNR) for block models which quantifies the minimum gaps between block means. This notion completely characterizes the hardness of the high-order clustering in tensor block model, which can be divided into three regions. In the strong SNR region, our method can achieve exact clustering in polynomial time; In the weak SNR region, we develop a mini-max lower bound to show that no algorithm succeeds in the task of high-order clustering; In the modest SNR region, we show that the problem is statistical possible while computational infeasible. That is, computing any estimate that achieves exact clustering is as hard as solving a version of hyper-graphic planted clique detection problem which is conjectured to be polynomial-time unsolvable. In conclusion, our theory reveals the phase transition property for multi-way clustering under tensor block model, which is further supported by the numeric experiments.

This is joint work with Yuetian Luo, Dr. Miaoyan Wang and Dr. Anru Zhang.



Student Puzzle Corner 38

Student Puzzle Editor Anirban DasGupta returns to consideration of statistical problems in this issue. The problem falls in the class of irregular problems. He says, "Certainly all of you have seen inference problems about uniform distributions with one or more unknown endpoints. That is one of the simplest irregular inference problems."

A tiger moves around a circular home territory of an unknown radius ρ, the circle being centered at a known point (x_0, y_0). Paw prints of the tiger have been detected at points (x_j, y_j), $j = 1, 2, \ldots, n$.

(a) Write a model for the problem.

(b) Find the MLE $\hat{\rho}$ of ρ under your model.

(c) Find, if it is possible, the exact distribution of the MLE in part (b).

(d) Find the asymptotic distribution of this MLE, i.e., find sequences a_n, b_n and a nondegenerate distribution G such that the distribution of $a_n(\hat{\rho} - b_n)$ converges in law to G.

Deadline: May 1, 2022

Student members of IMS are invited to submit solutions to bulletin@imstat.org (with subject "Student Puzzle Corner").

The names of student members who submit correct solutions, and the answer, will be published in the issue following the deadline.

The Puzzle Editor is Anirban DasGupta. His decision is final.

Solution to Puzzle 37

A reminder of the puzzle: Guest Puzzler Stanislav Volkov explains his problem about a speeding random walk:

For $n = 1, 2, \ldots$ let $X_n = \pm n$ with equal probabilities and assume that X_ns are independent.

Define a "speeding" random walk by $S_0 = 0$ and $S_n = X_1 + X_2 + \ldots + X_n$ (note that if X_n were ± 1 then it would be the ordinary, simple, symmetric random walk). A possible path for this random walk, starting from 0, would be 0, 1, 3, 0, 4, 9, 15, 8, ….

The Question:
Is the random walk S_n recurrent or transient (in the usual sense)?

We will show that $\sum_n \mathbb{P}(S_n = 0) < \infty$. Then, by the Borel–Cantelli lemma $S_n = 0$ only for finitely many ns a.s.. Next, note that if $\mathbb{P}(S_n = a \text{ i.o.}) > 0$ for some $a \in \mathbb{Z}_+$, then $\mathbb{P}(S_n = a - 1 \text{ i.o.}) > 0$ as well, since

$$\mathbb{P}(S_{n+2} = a - 1 \mid S_n = a) = \mathbb{P}(X_{n+1} = n + 1, X_{n+2} = -(n + 2) \mid S_n = a)$$
$$= \mathbb{P}(X_{n+1} = n + 1, X_{n+2} = -(n + 2)) = \tfrac{1}{4}.$$

By induction we get a contradiction with the fact that $\mathbb{P}(S_n = 0 \text{ i.o.}) = 0$. The argument for $a \in \mathbb{Z}_-$ is similar, so we conclude that $\mathbb{P}(S_n = a \text{ i.o.}) = 0$ for all $a \in \mathbb{Z}$ and thus S_n is transient in the sense that $|S_n| \to \infty$ a.s..

Now it remains to estimate $\mathbb{P}(S_n = 0) = A_n/2^n$ where A_n is the number of ways to put plus and minus signs in the sequence 1, 2, 3, …, n so that the result will be zero:

$$\pm 1 \pm 2 \pm \ldots \pm n = 0.$$

From the parity argument it follows that $A_n = 0$ unless $n \bmod 4 = 0$ or 1; for $n = 0, 1, 2, \ldots$ we can also easily compute $A = (1, 0, 0, 2, 2, 0, 0, 8, 14, 0, 0, \ldots)$. An internet search, e.g., indicates that this is a sequence A063865, see [1]. The asymptotic formula for A_n was given in [2]: $A_n \sim \sqrt{6/\pi} \cdot 2^n \cdot n^{-3/2}$. Hence

$$\sum_n \mathbb{P}(S_n = 0) \sim \sum_n \frac{\sqrt{6/\pi}}{n^{3/2}} < \infty \text{ as required.}$$

[1] The On-Line Encyclopedia of Integer Sequences, https://oeis.org/A063865
[2] Sullivan, Blair D. On a conjecture of Andrica and Tomescu. *J. Integer Seq.* 16 (2013), no. 3, Article 13.3.1, 6 pp.

Obituary: David Cox

1924–2022

Sir David Roxbee Cox died suddenly at his home in Oxford on January 18, 2022, leaving Joyce, his patient and supportive wife of almost 75 years, and their four children. David was so engaged and energetic in research that his death came as a shock.

David was born on July 15, 1924, in Birmingham, England, and held academic positions at University of Cambridge, Birkbeck College London, Imperial College London, and Nuffield College Oxford. He spent 15 months in the United States in 1955–56, visiting the University of North Carolina, Princeton, and UC Berkeley. During that time, he presented a Special Invited Lecture to the IMS, which became his highly influential 1958 paper, "Some problems connected with statistical inference" (46)[1] [*Numbered references refer to David's publication list at Nuffield College*[1].]

His seemingly boundless energy and enthusiasm, his brilliant mind, and a great deal of hard work, gave the world a remarkable oeuvre of contributions to all areas of statistical science and many of its applications. It is impossible in a short piece to do justice to these, and more detailed accounts will no doubt appear in other venues. A short article in the previous issue[2] highlights his key awards, although special note must be taken of the Royal Society's Copley Medal (2010), its highest honour. Previous winners include Gauss, Fisher, Bohr, and Einstein. Here we sketch some of his contributions to stochastic processes, foundations of inference, and statistical theory.

The doubly-stochastic Poisson process, also known as the Cox process, appeared in a remarkable paper in 1955 read to the Royal Statistical Society (36). The paper also includes a precursor to generalized linear models, the use (without comment) of the confidence distribution, attention to graphical displays, variance components, model fitting, model assessment, and so much more. In the discussion, Pearson commented, "it might have been a better policy to have narrowed his field of discussion and provided more illustrative material", but other comments were more astute — those of Bartlett on doubly stochastic processes are especially illuminating. David's grounding in stochastic processes suffused and influenced his work throughout his life. His insightful analysis of dependence in large data sets (361) is discussed by modelling the accumulation of data, and its variability, as a stochastic process with potentially long-range dependence. In his *Statistical Science* interview[3] he noted that his work on the proportional hazards model was directly informed by his background in stochastic processes.

David's work on inference brought clarity to a subject whose foundations were fragmented, sometimes incomprehensible, and occasionally badly flawed. Such was his modesty that he attributed many of the key ideas in his more philosophical papers and books on statistical inference to Fisher. The abstract of his masterfully lucid 1958 paper (46) notes: "It consists of some general comments, few of them new, about statistical inference. Parts of the paper are controversial; these are not put forward in any dogmatic spirit." In spite of these disclaimers, the paper is a landmark in the development of the foundations of inference. It covers many aspects of current relevance, including formal discussion of confidence distributions, but is best known

Sir David Cox in 2014 at his 90th Birthday Symposium. Photo courtesy of Nuffield College Oxford.

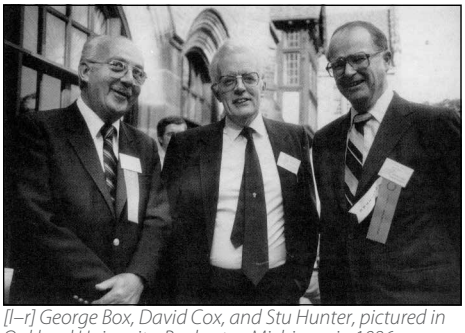
for its convincing demonstration of the need for appropriate conditioning in order to ensure scientifically relevant conclusions from statistical inference. This led directly to a long and important philosophical discussion, initiated by Birnbaum, on the role of the likelihood principle in inference and the interplay between frequentist and Bayesian inference. The paper also revealed that conditional inference is usually incompatible with ideas of optimality that remain popular today. The question of where to limit the conditioning is a challenging one, discussed for example in (226, §2.4). In the simplest setting, an arbitrarily granular choice renders each individual uninformative about others, while too coarse a conditioning typically yields conclusions irrelevant to the question at hand. When there are many nuisance parameters the appropriate conditional formulation becomes particularly elusive, although the conceptual argument for distinguishing samples of varying degrees of information remains compelling. A first attempt appeared in (46) and subsequent work sought to achieve the appropriate conditioning approximately.

In (139) he used an approximating curved exponential family to derive what

Continues on **page 15**

he called a local ancillary statistic, and obtained an approximation to the distribution of the maximum likelihood estimator, conditional on this statistic. Several other papers in the same issue of Biometrika tackled related problems, and the so-called p^*-approximation emerged as a common thread. David's interest in this was not focussed on the impressive numerical accuracy of the higher-order approximations, but on the implications of their structure for the foundations of inference. He refused to be dazzled by intricate mathematics or clever computation, unless it was demonstrably effective for solving what he might call "real problems". His pair of books with Barndorff-Nielsen (188, 226) contain a great deal of challenging mathematical detail, but are also full of statistical insight and enlightening examples.

David said that none of his books were written to be textbooks, although the very influential *Theoretical Statistics* (113) with David Hinkley is an exception. The emphasis on concepts of statistical inference and their relevance for applications, along with the parallel de-emphasis on mathematical details, distinguishes it from most books on statistical inference or mathematical statistics. It places likelihood and sufficiency at the centre of the theory of statistics, and may be the first text to clarify the distinction between significance testing,



[l–r] George Box, David Cox, and Stu Hunter, pictured in Oakland University, Rochester, Michigan, in 1986.

as developed by Fisher, and Neyman and Pearson's approach to hypothesis testing, treating both in considerable detail. Every potential principle of statistical inference is first explained, and then challenged, so effectively that the book can seem a collection of counter-examples. This is consistent with David's firm belief that the foundations and methods of statistical inference must be continually challenged and evaluated against their utility for applications, a point made strongly in his 2006 book (315), and again in (363). His writing on statistical significance and p-values seemed to need repeating for each new generation; a modern and concise account was published in 2020 (378).

David's contributions flowed smoothly between foundations of inference, theoretical analysis, development of methodology, and applications. He himself did not view these aspects as separate. A prominent example of this coalescence of ideas was his development of logistic regression (48, 49, 98), a topic so ingrained in modern statistical training that the ingenuity in its conception can be easily overlooked. A key aspect in the development was to specify sufficient statistics for the regression coefficients that coincide with those of a normal-theory linear model. The logistic construction emerges as the unique model for binary data that produces such unification, and an elegant theory of conditional inference then ensues, evading maximum likelihood fitting. This work is just one of his many unifying accomplishments; it reduces in the simplest special case to Fisher's (1935)[4] conditional analysis of the 2×2 contingency table, and leads in (90, 98) to the observation that all exponential-family responses can be treated in essentially the same way. A more flexible version, allowing one to renounce these simple sufficient statistics, was proposed by Nelder and Wedderburn (1972)[5] in the form of generalized linear models.

David was best-known for his 1972 paper (106) on the proportional hazards model. Its impact was both fundamental and immediate. It is ranked 16th in *Nature*'s list of most cited papers of all time in all fields, and was cited in the awards of the inaugural International Prize in Statistics (2016), the BBVA Foundation Frontiers of Knowledge Award (2017) and the Kettering Award (1990) for the most outstanding recent contribution to the diagnosis or treatment of cancer. As was characteristic of all his work, the elegant ease with which the results seemed to materialise partially masked the fundamental leaps involved in their inception, and the remarkable command of intuition, insight and technique that brought them to fruition. Once again, the motivating applications were the basis for foundational development, and partial likelihood emerged in 1975, encompassing conditional[6] and marginal[7] likelihood analyses.

His work remained current, and sometimes ahead of its time, hindered by prevailing computational considerations. In 1975 he gave an early elucidation of post-selection inference (115), demonstrating serious loss of inferential guarantees that sometimes arise when the research question to be studied is selected in the light of the data, and establishing the theoretical properties of sample-splitting in the simplest example. A key idea presented in passing in (90), further elaborated in (114, 193), resurfaced when he gave a totally new perspective on the sparse high-dimensional regression problems routinely encountered in genomics research. If multiple low-dimensional models are compatible with the data, his view was that one should aim to

Continues on **page 16**

David Cox, 1924–2022

Continued from previous page

report them all, rather than a single model effective for prediction. This underpins the development of confidence sets of models in (371, 374).

David's influence on science and statistical science was extraordinary, and his work will repay careful study for many years to come. His death leaves science much the poorer, without his keen judgment and unfailing curiosity; without his capacity to set the course of advancing knowledge with a single decisive contribution. Those fortunate enough to have crossed his path, professors and students alike, will remember a modest gentleman, keenly interested in everything scientific, thoughtful and perhaps a little bit shy. Until he stood up to deliver his talk. Then, one had a glimpse of his formidable intellectual energy and creativity, and remembered that talk for a very long time.

We miss him.

Written by Heather Battey, Imperial College London, and Nancy Reid, University of Toronto

References

1 Numbered references refer to the publication list at Nuffield College: https://www.nuff.ox.ac.uk/Users/Cox/Publications.html

2 *IMS Bulletin* (2022), 51, 1

3 Reid, N. (1994). *Statist. Sci.*, 9, 439–455.

4 Fisher, R.A. (1935). *J. R. Statist. Soc.*, 98, pp. 39–54.

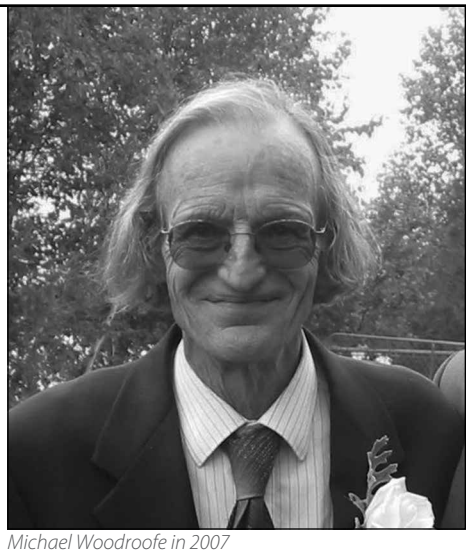
5 Nelder, J. and Wedderburn, R. (1972). *J. R. Statist. Soc. A*, 135, pp. 370–384.

6 Bartlett, M. S. (1936). *Proc. Roy. Soc. A*, 154, pp. 124–137.

7 Fraser, M. S. (1968). *The Structure of Inference*. John Wiley & Sons.

Obituary: Michael Woodroofe

1940–2022



Michael Woodroofe in 2007

Michael Barrett Woodroofe, IMS Fellow and former Leonard J. Savage Collegiate Professor of Statistics at the University of Michigan, passed away on February 22, 2022 at the age of 81.

Michael was born on March 17, 1940 in Corvallis, Oregon, and grew up in the small town of Athena in northeast Oregon. His mother was an English teacher and his father mainly farmed, although he also taught science at high school for a while. Michael graduated from McEwen High School in 1958, then, encouraged by his parents, continued his studies at Stanford University, earning a B.S. in Mathematics in 1962. Returning to his home state, he completed a PhD in Mathematics in 3 years at the University of Oregon.

After his PhD, Michael spent a year as a Research Associate at Stanford University, followed by two years as an Assistant Professor of Statistics at Carnegie Mellon University, before joining the faculty at the University of Michigan in 1968 as an Assistant Professor of Mathematics. One year later he helped found the Department of Statistics at Michigan, joining it as an Associate Professor. He was promoted to full professor in 1973, and remained with the Michigan Statistics Department for the rest of his career, apart from a few year-long visits away. The visits away from Michigan consisted of a Visiting Associate Professorship of Statistics at Columbia University, 1970–71, spending time with Herbert Robbins; a Visiting Professorship of Mathematics at the Massachusetts Institute of Technology, 1976{77, enjoying the company of Herman Chernoff; and a position as Professor II of Statistics at Rutgers University, 1983–84, where he had several close friends, especially Harold Sachrowitz who shared a keen interest with him in baseball. At Michigan, he served as department chair from 1977 until 1983. In 1994 he was awarded the college's highest faculty honor, an endowed professorship, to become the Leonard J. Savage Collegiate Professor of Statistics.

Michael was an eminent mathematical statistician working near the interface of probability and statistics. Early in his career he made seminal contributions to renewal theory in sequential analysis, including applications to repeated significance tests and fixed-width confidence intervals. His paper in the late 70's on the one-armed bandit with a covariate may be the first article on contextual bandits, now an active

area of inquiry in machine learning. Later on, motivated by applications to astronomy, he began working on nonparametric function estimation with shape constraints, with notable contributions to global risk, consistency, the behavior of estimators near boundaries, degrees of freedom, and the performance of different bootstrap methods. Later in his career, he had an ongoing interest in central limit theory for additive functionals of Markov chains and stationary processes, with papers giving general conditions for martingale approximations, developing quenched, conditional, and local limit theorems, and proving invariance principles for stationary random fields.

Over the course of his career Michael supervised the dissertation research of more than 40 PhD students, many of whom went on to have productive research careers of their own. He also served as a mentor for many young faculty members at Michigan, always happy to listen to their ideas and offer valuable suggestions. Michael was also active on editorial boards, most notably serving as the Editor of the *Annals of Statistics* from 1992–94, the last years it had a single editor.

In 1974 Michael married Frances Smock, an elementary school teacher. They had three children, Russ, Carolyn, and Blake, and they have all prospered. Russ is now an Associate Professor of Mathematics at the University of Primorska in Slovenia, Carolyn is a chemist in the Chemistry and Synthesis Center in the NIH's National Heart, Lung, and Blood Institute in Maryland, and Blake runs a small construction company in Ypsilanti, Michigan. Michael enjoyed spending the summer time with his family at a cottage on a lake in northern lower Michigan. He liked to swim and mountain bike there, and found he could also work quite productively without the distraction of internet access. At age 8, Michael's youngest son Blake developed a keen interest in hockey, and Michael took an adult class to learn more about the game. Finding he liked it, he became an avid hockey player, later playing on a club team with Blake.

Michael retired in 2009, but continued to collaborate with coauthors until around 2017. During this time his health gradually declined, and hearing loss that began in midlife became increasingly severe. In the last years of his life he suffered from dementia, spending most of his final year in an assisted living unit in Palm Bay, Florida.

Michael was not only an exceptional statistician but also a wonderful human being who touched the lives of many. He leaves behind a great legacy and will be sorely missed. He is survived by Fran, Russ, Carolyn, and Blake, and 4 grandchildren.

Written by Robert Keener, University of Michigan, with help from friends.

Obituary: Yuan Shih Chow

1924–2022

Yuan Shih Chow, Professor Emeritus of Statistics at Columbia University, passed away on March 3, 2022 in his hometown in Hubei, China. He suffered a stroke earlier and never recovered. He was 97 years old. He is survived by three daughters, Letitia, Eunice and Grace, two sons, Patrick and Nelson, seven grand children and four great-grand children.

Y.S. Chow was born on September 1, 1924, in Hubei province in China. As the eldest child of a local leader, he grew up in war and turmoil. His hometown was on the front line during most of the Japanese invasion and occupation of China. When he was 16, he journeyed, mostly on foot, to Chongqing (Chung-King), the wartime Chinese capital, to finish his high school education. He entered Zhejiang University (then known as National Che Kiang University) in 1944 as a freshman majoring mathematics. In early 1949, after graduating from college, he took a teaching position with Department of Mathematics at the National Taiwan University (NTU) in Taipei. He left NTU in 1954 to pursue a PhD in mathematics with a full scholarship from the University of Illinois at Urbana-Champaign (UIUC). There, he completed his PhD thesis in 1958 on martingale theory under the supervision of J.L. Doob. After a one-year postdoctoral research under A.H. Taub in aerodynamics at UIUC, he joined the IBM Watson Lab, then a joint venture between IBM and Columbia University, as a member of research staff. He started teaching as an adjunct faculty member for Department of Mathematical Statistics at Columbia University. He left Watson Lab in 1961 to become a full time faculty at Columbia. He joined Purdue University in 1962 as Associate Professor of Mathematics and Statistics and was promoted to full professor in 1965. He returned to Columbia University in 1968 and stayed until his retirement in 1993. At different

Yuan Shih Chow, 1924–2022

Continued from previous page

times, he was a visiting professor at the University of California at Berkeley, University of Heidelberg, Germany and the National Central University, Taiwan. He served as Director of the Institute of Mathematics at Academia Sinica, Taiwan and Director of the Center of Applied Statistics at Nankai University, Tianjin.

Y.S. Chow was a strong advocate for mathematics, statistics and computer science. While serving as Director of the Institute of Mathematics at Academia Sinica, he founded the *Bulletin of the Institute of Mathematics, Academia Sinica*. He was key in the development of computer science and technology in Taiwan, and was instrumental in establishing the Institute of Statistical Sciences, Academia Sinica.

Y.S. Chow was a superb mentor and kind advisor to students and young people in their career development. His PhD students include William F. Stout, Robert James (Jim) Tomkins, Bruce M. Brown, Chao Agnes Hsiung, K.K. Gordon Lan, Kai Fun Yu, Adam T. Martinsek, Cun-Hui Zhang, Moshe Olshanski and Xinghong Xue. He also influenced and guided many people in pursuing their academic careers in statistics and related disciplines, including Pei-Er Lin, Deng-Yuan Huang, Chien-Fu Jeff Wu, Ching-Zong Wei, Regina Y. Liu, Zukang Zheng, Mai Zhou, Chuanshu Ji, Yingyi Qian, Yusheng Zheng, Minggao Gu, Zhiliang Ying, Zhaohai Li, Victor H. de la Pena, Cheng-der Fuh, Steven Kou, Weibiao Wu, Ching-Kang Ing and Haipeng Xing, among others.

Y.S. Chow was a fellow of the Institute of Mathematical Statistics, an elected member of the International Statistical Institute and an elected member of the Academia Sinica.

Y.S. Chow was an influential and prominent scholar in probability theory and statistics. His books include *Great Expectations: The Theory of Optimal Stopping* (1971), co-authored with Herbert Robbins and David Siegmund, *Probability Theory* (1978) co-authored with Henry Teicher, and *Topics on Random Walks* (1995). His works in martingale theory have been widely recognized and used in statistics and related disciplines. The Hájek–Rényi–Chow inequality, an important extension of Kolmogorov's maximum inequality on sums of independent variables and Doob's maximum inequality on martingales, can be used to provide a direct proof of the strong law of large numbers, bypassing the Kronecker lemma. In the theory of martingale convergence on directed/partially ordered sets, the Vitali–Chow criterion unified the classical Lebesgue and Ward's differential calculus under a martingale structure. His work on martingale local convergence provided a highly useful tool in sequential analysis, time series analysis, adaptive control and optimization and analysis of machine learning algorithms. In a series of papers in the 1960s, Y.S. Chow and Herbert Robbins jointly laid the mathematical foundation for the theory of optimal stopping and set the stage for the further development of the subject. In sequential estimation, Chow and Robbins, in a landmark paper (Chow and Robbins, 1965), developed a construction of fixed-width confidence intervals for a Gaussian mean without knowing the variance, and later on Chow and his students developed sophisticated theory on the boundedness of the regret in such problems.

Yuan Shih (Y.S.) Chow

After his retirement from Columbia University, Y.S. Chow devoted much of his energy towards philanthropic activities. With his own money, he established two college educational funds in his hometown: one for high-achieving students and the other for economically disadvantaged students. He also funded a high school mathematics competition in his hometown which is known as the Yuan Shih Chow Mathematics Competition Prize.

For additional stories about Y.S. Chow, see an interview in *Statistical Science* (Ying and Zhang, 2006).

Written by Zhiliang Ying, Columbia University, and Cun-Hui Zhang, Rutgers University

Written by Witten

So long, and thanks for all the tips

Daniela Witten writes: I've never been shy about asking for advice. As a first-year grad student, I'd beg more senior students for tips on passing quals and writing a dissertation. And as junior faculty, I asked senior colleagues how to manage my time, prioritize projects, advise students, and ultimately get tenure. *Advice is free, and by nature I'm a hoarder.*

But time passed, as it does. And one day I noticed something funny. People had started to ask me for advice. *Me!!* A person who knows absolutely nothing about anything!! Me, who always asks other people for advice! How bizarre!!

And that's when I realized that *we are all just mucking our way through*. All of the advice that I had pursued for the previous decade was just a bunch of flawed attempts by imperfect humans to make sense of a world besieged by chaos. In other words, *advice is free, and you get what you pay for.* I'm not saying that advice is useless. On the contrary, it is often extremely valuable. It provides a window into someone else's life. It's a reflection of what worked for them, and what didn't. But it is not a mirror into your life, nor is it a crystal ball that predicts your future (if only those existed!).

I've since learned that when receiving advice, there's no need to hoard. It's best to follow the KonMari method (https://konmari.com/about-the-konmari-method/): I keep the advice that speaks to my heart and sparks joy. Otherwise, I *let it go*.

So, without further ado, I'm going to dedicate this column to my top 10 tips for PhD students. If you're a PhD student, then I hope that some of this advice speaks to your heart. And if it doesn't, then you know what to do with it!

10) **It's not about the destination, it's about the journey.** Yes, some of your PhD classmates derived their first minimax bounds *in utero*. Good for them! We all have our own paths, and it's not a race.

9) **Give it your best shot.** Once you decide to do something — whether it's going to grad school, pursuing a summer internship, carrying out a research project, or writing a column for the *IMS Bulletin* — commit to it 100%.

8) **If at first you don't succeed…** We all deal with rejection: rejected papers, unsuccessful grant applications, and more. Anyone who says otherwise is either new to academia, or lying. Pick yourself up and try again, and always celebrate your successes 100 times more than you mourn your failures.

7) **Choose your PhD research advisor wisely.** This is the single most important professional relationship of your PhD, and potentially of your entire career. Are they approachable and supportive? Will they help you achieve your potential? Talk to their current and past students, but remember that what works for you might not work for someone else, and vice-versa.

6) **Round out your skill set.** A successful career in statistics requires not only solid technical skills, but also strong written and oral communication skills. Your overall career success will be limited by the weakest of those three sets of skills. Work to improve your weakest skill set. Take an improv class or try your hand at creative writing… the possibilities are endless.

5) **Play to your strengths.** Are you incredibly talented at deriving concentration inequalities, consulting for non-statisticians, teaching undergrad service courses, collaborating as part of a large interdisciplinary team, *and* programming? If you answered "yes", then please stop showing off. Most of us humans have strengths in just one or two of those areas. Your PhD is an opportunity to discover those strengths, and to identify a dissertation project — and, ultimately, a career path — that makes the most of them.

4) **Find balance.** No matter how much you love your work, it will never love you back. And you may never again have a schedule that is as flexible as during your PhD. So, get your work done during the week and enjoy your weekends. Backpack through Europe, become an extremely accomplished knitter, or dive deep into the Netflix archive. Carpe diem.

3) **Hold your cohort close and your enemies clos…** — that's not right. What I meant to say is, be friends with your PhD cohort. These people will be your primary professional contacts throughout your career. And remember, there are no enemies. Everyone's just trying to pass quals.

2) **Give yourself some grace.** You're doing something spectacularly challenging that most people would never even consider attempting. You're absolutely amazing!!!! Yay for you!!!! And always remember that *done is better than perfect*.

1) **Have fun.** You don't need to love every minute of your PhD, and I promise that you won't! But make sure that, on balance, you like what you're doing. Otherwise, you may achieve your career goals (a tenure-track job at a great university? a well-paid job at the company you've always wanted to work at?), only to discover that you were chasing someone else's dream.

That's it for now! Reach out to me @daniela_witten on Twitter to let me know what joy-sparking tips for PhD students make your top 10 list.

Ruobin Gong interviews Claire McKay Bowen

Winner of the 2021 COPSS Leadership Academy and author of *Protecting Your Privacy in a Data-Driven World*, **Claire McKay Bowen** [*pictured left*], Principal Research Associate at the Urban Institute, talks with our contributing editor **Ruobin Gong** [*right*] about data privacy, public policy research, and data science education.

RG: *Claire, thank you for taking the time to speak with me. As a statistician, what made you choose a career in public policy research?*

CB: One of the big drivers for my career paths and decisions is what are the ways I can make the most impact. This is why I worked in the government initially, did internships at Los Alamos National Lab and at the Census Bureau, and went back to Los Alamos for my postdoc. To make an impact also goes into my research. From early on, I know I'm application-driven. I majored in math and physics, because math is the language of science, but I didn't want to be a theoretician. I want to do applications in science. Through the winding path, I realized that I really like the analytical part about science. That led to statistics, where I seek for the most interesting challenges where I can make the most impact.

RG: *What are some things that you enjoy most about working at the Urban Institute?*

CB: Part of why I enjoy working at Urban (https://www.urban.org/) is knowing that I'm making a difference. My work goes a step further than just the research. It translates theory into practice. Urban Institute is a non-profit, non-partisan public policy research institution. We provide evidence-based research on public policy issues, mostly in the United States and sometimes internationally. While we're not an advocacy group, it is exciting to think that we are trying to make the society better, and life for the better.

On a day-to-day basis, my work involves a lot of writing. It is funny because, as a child, I hated writing. I went into math and science because I wouldn't have to write! The joke is on me, as the bulk of my time now is spent on writing – not just blogs and communication pieces, but also grant proposals and full research papers.

RG: *Writing is the vehicle of communication. Whatever path we choose, it seems like we can't get away from it!*

CB: Exactly. Through my undergraduate studies in physics, I got into science communication. How do I get people interested in science and understand what's going on in science? That has also driven my current career, which is trying to communicate my research in privacy because there's just so few of us who work in that space. If we don't help try to communicate what's going on in privacy, we're not going to have the right voices at the table when important decisions are made, such as how do people get access to data amidst privacy concerns.

RG: *You are a prolific writer and communicator in the privacy sphere. Your book discusses the technical challenges and also its historical, legal, and ethical aspects. In your view, how can statisticians go beyond the technical work that we do, to have our own voice, and to shape the public discourse on data privacy and other modern challenges?*

CB: We should be communicating issues and ideas from data privacy to students in their early education. I talked about this in my book, that most students don't learn about data privacy until they're in graduate school. Even then, it's mostly in computer science programs, except for a few statistics programs. However, a lot of discussions in the privacy community are about applications to public datasets. For instance, the new differentially private census data is at the center of much debate. What many people don't realize is that the census data have been always altered even prior to the 2020 Census. Much of the debate suffers from miscommunication: people not knowing what's going on. With that said, in the past, the Census Bureau did disclose data when they shouldn't have, which contributed to a mistrust between the community and the government. This is also why the Census Bureau now takes such a strong stance on privacy. These things should have already been discussed early on in a data science education.

Another thing is, the statistics community has not been effective in advocating for our own work. The COPSS Presidents' Award (https://community.amstat.org/copss/awards/presidents) is our most prestigious award, for somebody whose career makes a significant contribution to the field. But outside of statistics, rarely have people heard of the COPSS award, whereas everybody knows what the Fields medal is. Xiao-Li Meng once said to me: "money talks". The Fields medal was historically given with a cash prize. It was not a big amount — really small in comparison to the Nobel Prize, which awards a million dollars. But, it was still a significant amount, especially back then. People probably thought, "Oh wow,

it must be prestigious." We now have the new Rousseeuw Prize for Statistics (https://www.rousseeuwprize.org/) that comes with a million dollars. I think that'll help quite a bit to elevate statistics.

RG: *What advice do you have for our aspiring data scientists who are considering a career in research and analytics in public policy and nonprofit?*

CB: When people ask me what skills they should work on, I tell them it is communication. When we hire somebody at Urban, we have a minimum threshold for coding and for their statistical background. We don't expect them to know privacy – we can teach them. But the one thing that I cannot teach as easily on the job is how to communicate with the audience, both verbally and in writing. When we interview people, we ask for a code sample and a writing sample. We ask questions such as, "Can you explain your research?" and, "How would you pitch it to a funder?". As a non-profit, we must be somewhat entrepreneurial in seeking funding, both through grant applications and through talking to potential funders and showcasing the kind of work that we do. So another important thing is to show passion and excitement for the work you do, which is also not easy to replicate sometimes. You want to convince other people to be excited about your work, and you should feel excited about it too.

RG: *What are some exciting things that you are currently working on?*

CB: I'll tell you a technical one and a fun one. We're working with the Internal Revenue Service (IRS) to expand access to the taxpayer dataset (https://www.urban.org/research/publication/synthetic-supplemental-public-use-file-low-income-information-return-data-methodology-utility-and-privacy-implications). Taxpayer datasets are very sensitive information, but if made accessible, the potential public policy impact is huge, because we could do more targeted decision-making, such as for stimulus packages or certain tax policy laws.

We're constructing a synthetic dataset for public release, by balancing privacy implications with the usefulness of certain analyses. For example, if an economist wants to run a regression kink design analysis, could there be a way to apply their code on the public dataset, which would have the same structure as the confidential one?

What we're developing is a validation server (https://www.urban.org/research/publication/privacy-preserving-validation-server-prototype, https://arxiv.org/abs/2110.12055) which has access to the confidential data. Submitting the code through the validation server returns a noisy answer. This will be an automated system. The Census Bureau had what's called the synthetic data server through Cornell University, with two synthetic datasets and the corresponding confidential dataset. However, the queries are manually reviewed for whether the answer is noisy enough. This could take a long time if the demand exceeds staff time. Similarly, the IRS has only so many staff to maintain researcher access. An automated system would speed it up. Another motivation is that, to access many of these sensitive official datasets, you have to be a U.S. citizen. This eliminates a lot of researchers. The clearance process also takes a long time. I've gone through multiple of them and it doesn't get easier.

The fun project I am working on with a colleague is called Data4Kids (https://www.urban.org/data4kids-virtually-teaching-kids-about-data-science). It is a toolkit to help teach kids data visualization and data science. When the pandemic happened, a lot of teachers struggled to teach kids virtually. We decided to make a toolkit that could be used both virtually and in-person.

We decided to do themed data stories. We pick a dataset and modify it for different grade levels: elementary school, middle school, and high school. The data also has a messy version and a clean version. The data story is a full lifecycle. We ask: *What is your data question? How do you collect the data? Is it messy? How do you analyze and visualize the data?* The ending talks about data privacy, ethics and equity. We ask the students: *Are you represented in this data? Would the answers from the data apply to you? Would it apply to your family and your neighbors?* This way we get them to think when they hear something on the news: "Does that apply to me?"

RG: *It's wonderful to hear you talk about this. It can be difficult to get the students, especially kids, started with thinking about data equity, because the subject is so vast. But you have distilled it into a couple of simple questions that are also fundamental. "Does it apply to me?" I think this is beautiful.*

CB: When people say, "Can we really teach data science to a third grader?" I think the answer is yes. We just have to frame it differently.

RG: *Thank you, Claire!*

Read more about Claire at https://clairemckaybowen.com/ and meet the new COPSS Leadership Academy members on page 4.

Recent papers: Two Open-Access journals

Stochastic Systems

Focusing on the interface of applied probability and operations research, *Stochastic Systems* is the flagship journal of the INFORMS Applied Probability Society and is published through a cooperative agreement between INFORMS/APS and IMS. This open-access journal seeks to publish high-quality papers that substantively contribute to the modeling, analysis, and control of stochastic systems. The contribution may lie in the formulation of new mathematical models, in the development of new mathematical methods, or in the innovative application of existing methods. A partial list of applications domains that are germane to this journal include: service operations; logistics, transportation and communications networks (including the Internet); computer systems; finance and risk management; manufacturing operations and supply chains; and revenue management.
Read it at **https://pubsonline.informs.org/toc/stsy/current**

Volume 12, Issue 1, March 2022

Probability Surveys

Probability Surveys is a peer-reviewed electronic journal which publishes survey articles in theoretical and applied probability. The style of articles may range from reviews of recent research to graduate textbook exposition. Articles may be broad or narrow in scope. The essential requirements are a well specified topic and target audience, together with clear exposition. The journal is sponsored by the Institute of Mathematical Statistics and by the Bernoulli Society. *Probability Surveys* is an Open Access journal. The full text of each article published is freely available to all readers.
Read it at **https://projecteuclid.org/journals/probability-surveys/current**

Volume 19, 2022 (to date)

Volume 18, 2021

IMS journals are Plan S compliant

Are you wondering if your article is **Plan S compliant** (https://www.coalition-s.org/) when submitting to our journals? We are pleased to report, all our journals are compliant: that's *Annals of Applied Probability, Annals of Applied Statistics, Annals of Probability, Annals of Statistics* and *Statistical Science.* Articles will be printed with a CC-BY license, you can see the Plan S copyright form at https://imstat.org/wp-content/uploads/2022/03/copyrightTA-cOAlition-S.pdf.

In addition, IMS-supported journals are Plan S compliant and use the same copyright form, these include: *Annales de l'Institut Henri Poincaré, Bernoulli,* and *Brazilian Journal of Probability and Statistics.*

The IMS co-sponsored Gold Open Access journals have used a CC-BY 4.0 license since they were founded and are also Plan S compliant, these include: *Electronic Communications in Probability, Electronic Journal of Probability, Electronic Journal of Statistics, Journal of Computational and Graphical Statistics, Probability Surveys,* and *Statistics Surveys.*

If you use the journal tracker tool and our journals aren't showing up yet, they will soon. In the meantime, please know we have met the requirements. In fact, the IMS journals have always been ahead of the curve, we have allowed our authors to self-archive with no embargo for over 15 years.

If you have any concerns or questions, please email the IMS Executive Director, Elyse Gustafson at erg@imstat.org.

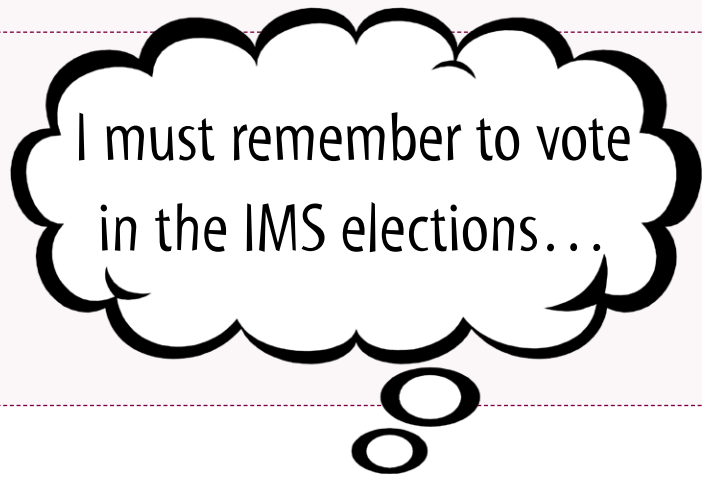
Plan S
Making full & immediate Open Access a reality

Plan S is an initiative for Open Access publishing that was launched in September 2018. The plan is supported by cOAlition S (https://www.coalition-s.org/), an international consortium of research funding and performing organisations. Plan S requires that, from 2021, scientific publications that result from research funded by public grants must be published in compliant Open Access journals or platforms.

2022 IMS Elections: Have *you* voted?

Have you decided who you would like to represent you on the IMS Council, and voted in the IMS elections? In case you haven't, here's a refresher. The nominee for President-Elect is **Michael Kosorok**. The 10 candidates standing for the five available places on Council are **Siva Athreya, Rina Foygel Barber, Gilles Blanchard, Alexandra Carpentier, Sayan Mukherjee, Sofia Charlotta Olhede, Debashis Paul, Judith Rousseau, Ryan Tibshirani, and Harrison H. Zhou**. Read about them at **https://imstat.org/elections/candidates/**

Elections close May 13, 2022. Vote online at https://www.imstat.org/portal/voting/8/step1



IMS meetings around the world

Joint Statistical Meetings

2022 Joint Statistical Meetings
August 6–11, 2022. Washington DC
w https://ww2.amstat.org/meetings/jsm/2022/
Registration and housing open May 2, 2022.
Early registration deadline May 31, 2022.

AUGUST 6–11
2022
WASHINGTON, DC
JSM
STATISTICS: A FOUNDATION FOR INNOVATION

JSM dates for 2023–2026

IMS Annual Meeting @ JSM 2023	JSM 2024	IMS Annual Meeting @ JSM 2025	JSM 2026
August 5–10, 2023	August 3–8, 2024	August 2–7, 2025	August 1–6, 2026
Toronto, Canada	Portland, Oregon, USA	Nashville, TN, USA	Boston, MA, USA

22nd Meeting of New Researchers in Statistics and Probability NEW
August 3–6, 2022
George Mason University, Fairfax, VA (close to Washington DC)
w https://imstat.org/nrc2022/
The IMS New Researchers Group excited to resume this in-person conference, taking place as usual immediately before JSM (which is in Washington DC, August 6–11, 2022).
All new faculty, or soon-to-be faculty, in statistics, probability and data science (i.e. you received your PhD in or since 2015, or expect to by the end of this year) are encouraged to apply **by April 15**: see website for form.

South East Probability Conference NEW
May 31–June 1, 2022, at Duke University *and* August 8–9, UNC Chapel Hill
w https://services.math.duke.edu/~rtd/SEPC2022/SEPC2022.html
There will be *TWO* 2022 Southeastern Probability Conferences this year, both IN-PERSON meetings. Speakers at the May 31–June 1 SEPC (Duke University, Durham, NC): Erik Bates (Wisconsin), Juraj Foldes (Virginia), Konstantin Matetski (Columbia), Scott McKinley (Tulane), Evita Nestoridi (Princeton), Mariana Olvera-Cravioto (UNC), and Marrianna Russkikh (MIT).

The second SEPC in August at UNC Chapel Hill features Julia Gaudio (Northwestern, IE and MS), Jiaoyang Huang (NYU), Hye-Won Kang (U of Maryland), Annie Katsevich (NYU), Nicholas Lancier (Arizona State), Elizaveta Rebrova (Princeton ORFE), Sebastien Roch (Wisconsin), and Philippe Sosoe (Cornell).

For the second meeting, graduate students and postdocs can apply for support. Details forthcoming.

2022 IMS Annual Meeting
June 27–30, 2022. London, UK
w www.imsannualmeeting-london2022.com
The 2022 IMS Annual Meeting will be held in London immediately before COLT, with extra one-day workshop planned [*see announcement, right*] between the two meetings. **Registration for the IMS–COLT workshop is included in your fee.** Program and Local Chair: Qiwei Yao.

2022 IMS–COLT Joint Workshop
July 1, 2022. London, UK
w https://bguedj.github.io/colt-ims-2022.github.io/
The 2022 IMS Annual Meeting [*see left*] will be immediately followed by the first IMS–COLT joint workshop, a one-day meeting in a hybrid format (on-site in central London, and online), linking the IMS and COLT communities of researchers. (COLT is the annual Conference on Learning Theory, and will take place in 2022 immediately after this IMS–COLT workshop day.) **If you're registered for the IMS Annual Meeting, this meeting is included.**

At a glance: forthcoming IMS Annual Meeting and JSM dates

2022

IMS Annual Meeting: London, UK, **June 27–30, 2022**

JSM: Washington DC, **August 6–11, 2022**

2023

IMS Annual Meeting @ JSM: Toronto, **August 5–10, 2023**

2024

IMS Annual Meeting/ 11th World Congress: Bochum, Germany, **August 12–16, 2024**

JSM: Portland, OR, **August 3–8, 2024**

2025

IMS Annual Meeting @ JSM: Nashville, TN, USA, **August 2–7, 2025**

2026

IMS Annual Meeting: TBD

JSM: Boston, MA, **August 1–6, 2026**

More IMS meetings

New Advances in Statistics and Data Science
May 24–26, 2022, in Honolulu, Hawaii NEW

w http://zke.fas.harvard.edu/HawaiiConference/Main.html

As Statistics and Data Science continue to evolve in the big data era, we face new challenges that call for novel interdisciplinary research combining statistics, machine learning, econometrics, computer science, and biology. This 3-day conference brings together leading researchers from these various fields. They will offer new perspectives and methodologies, and address the opportunities and challenges of 21st century Data Science.

Registration opens on April 15, 2022. The registration fee is $250 (regular) and $100 (student) plus service charge (about $5). Speakers and attendants will fund themselves for the travel and lodging expenses.

The conference will be held at the Hyatt Regency Waikiki Beach Resort and Spa (2424 Kalakaua Avenue, Honolulu, Hawaii, United States, 96815).

BNP13: the 13th Conference on Bayesian Nonparametrics
October 24–28, 2022 in Puerto Varas, Chile

w https://midas.mat.uc.cl/bnp13/

This meeting aims to gather in Chile leading experts in this already consolidated and rapidly expanding field for a full week of plenary, invited, contributed and poster talks, reflecting the many and varied aspects of the theoretical, methodological, and applied work in the discipline. We also expect to attract many young researchers to the event, especially those residing in South America and in Chile. The event features three keynote lectures (free topic), invited and contributed sessions, spanning all the wide spectrum of theory, methodology and applications of Bayesian methods from the nonparametric viewpoint: asymptotics, advanced computation, robustness, inference in high-dimensional settings, deep learning, machine learning, uncertainty quantification, clustering and applications.

2022 IMS International Conference on Statistics and Data Science (ICSDS)
December 13–16, 2022
Florence, Italy

w https://sites.google.com/view/icsds2022

The first IMS International Conference on Statistics and Data Science (ICSDS) is to be held in Florence, Italy, from December 13–16, 2022. The objective of ICSDS is to bring together researchers in statistics and data science from academia, industry and government in a stimulating setting to exchange ideas on the developments in modern statistics, machine learning, and broadly defined theory, methods and applications in data science. The conference will consist of plenary sessions, and about 50 invited, contributed and poster sessions. **Young researchers are particularly encouraged to participate**, with a portion of the invited sessions designated for them.

JOINT IMS–COLT WORKSHOP: JULY 1, 2022, LONDON
w https://bguedj.github.io/colt-ims-2022.github.io/

For the first time, the IMS Annual Meeting and COLT are co-located in London in 2022, with the IMS Annual Meeting happening on June 27–30 and COLT on July 2–5. There will be a joint one-day workshop between the two conferences, on July 1, to offer a chance to the COLT and IMS audiences to mingle and engage on topics related to statistics and learning theory.

All registered attendees to the IMS Annual Meeting or to COLT are automatically registered for this joint workshop. Join us in London!

The speakers are Emmanuel Candès, Nati Srebro and Vladimir Vovk for tutorials and talks, on the topics of conformal prediction and mathematics of deep learning. The workshop will close with a social event to give IMS and COLT participants a chance to mingle.

Organizers: Benjamin Guedj, Inria (France) and UCL (UK), Peter Grünwald, CWI and Leiden University (Netherlands), Ciara Pike-Burke, Imperial College London (UK), and Susan Murphy, Harvard (USA).

This event is supported by IMS and COLT.

More IMS meetings around the world

Asia-Pacific Seminar in Probability and Statistics: Ongoing and online

w https://sites.google.com/view/apsps/home

The Asia-Pacific Seminar in Probability and Statistics (APSPS) is a monthly online seminar, broadcast on a mid-month Wednesday via Zoom. The seminar series was created as a permanent forum for good research in the field. Topics include: probabilistic models for natural phenomena, stochastic processes and statistical inference, statistical problems in high-dimensional spaces, asymptotic methods, statistical theory of diversity. The organizers — Sanjay Chaudhuri (NUS, Singapore), Mark Holmes (University of Melbourne, currently at UBC, Vancouver), Estate Khmaladze (VUW, Wellington; Chair), Krishanu Maulik (ISI, Kolkata), Spiro Penev (UNSW, Sydney), Masanobu Taniguchi (Waseda University, Tokyo), Lijiang Yang (Tsinghua University, Beijing), and Nakahiro Yoshida (University of Tokyo) — seek an emphasis on novelty, beauty, and clarity. Presentations are intended to be accessible to good postgraduate students in probability and mathematical statistics.

If you are interested in receiving email announcements about the next speakers, send an email to any of the Board members listed above.

Statistics in the Big Data Era

June 1–3, 2022

UC Berkeley, CA, USA

w https://simons.berkeley.edu/workshops/statistics-big-data-era

This conference is focused on the changing role and nature of the discipline of statistics in the time of a data deluge in many applications, and increasing success of artificial intelligence at performing many data analysis tasks. The conference aims to bring together experts in statistical methodology and theory for complex and big data with researchers focused on a range of applications, from genomics to social networks, and to provide opportunities for new researchers to learn about both emerging methods and applications. The conference will also be an occasion to celebrate Professor Peter Bickel's 80th birthday.

2024 IMS annual meeting

Bernoulli–IMS 11th World Congress in Probability and Statistics

August 12–16, 2024

Ruhr-University Bochum, Germany

w TBC

ENAR/IMS Spring Meeting

March 22–25, 2023.

Nashville, TN, USA

w https://enar.org/

Stochastic Networks

June 20–24, 2022. Cornell Univ., Ithaca, NY

w https://sites.northwestern.edu/snc2022/

Stochastic networks is a multifaceted area of research concerned with the modeling, stability, control, performance, approximation, and design of stochastic networks. It gives rise to challenging and subtle mathematical problems, whose solution often requires a combination of ideas and techniques from several branches of mathematics, including probability theory, stochastic processes, analysis, optimization, algorithms, combinatorics, and graph theory. Research in this area is strongly motivated by applications in diverse domains, ranging from telecommunications and manufacturing to service operations, biological and social networks, revenue management, and health care. The conference series, initiated in 1987 and held biennially, is a major forum for researchers to learn of the latest developments and new research directions in stochastic networks.

2022 WNAR/IMS/JR Annual Meeting NEW

VIRTUAL. June 10–15, 2022

w https://wnarofibs.wildapricot.org/WNAR2022

Due to the ongoing COVID-19 pandemic, the 2022 WNAR/IMS/JR meeting will be held virtually. **Registration is open.** The program features Short Courses (June 10-11) on Veridical Data Analysis and Clinical Development of New Drugs; Invited Oral Sessions, Contributed Oral Sessions, and Student Paper Sessions (June 13–15); and the Presidential Invited Address (June 14). Nicholas Jewell, UC Berkeley and London School of Hygiene and Tropical Medicine will be the presidential invited speaker. The title and abstract of his talk will be announced shortly. The abstract submission deadline has passed.

One World ABC Seminar: Ongoing and online

w https://warwick.ac.uk/fac/sci/statistics/news/upcoming-seminars/abcworldseminar

The One World Approximate Bayesian Computation (ABC) Seminars are fortnightly seminars that take place via Zoom on Thursdays at 11:30am, UK time. The idea is to gather members and disseminate results and innovation during these weeks and months under lockdown. Register to receive the webinar link via email. The organizers welcome proposals for future talks. This webinar is part of the larger One World seminar initiative [*see below*].

One World Probability Seminar (OWPS): Ongoing and online

w https://www.owprobability.org/one-world-probability-seminar/future-seminars

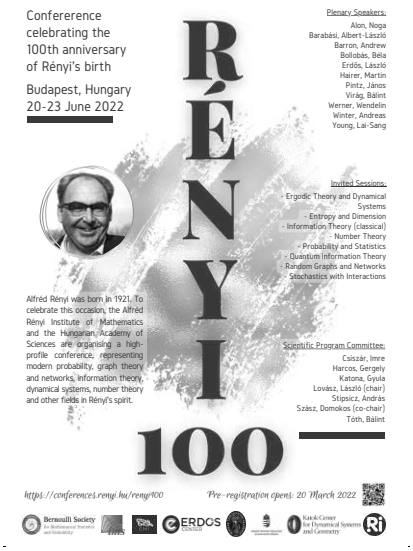
Thursdays, 14:00 UTC/GMT. Please subscribe to the mailing list for updates about the upcoming seminars and other events: https://www.owprobability.org/mailing-list

Conference celebrating 100th anniversary of Rényi's birth
June 20–23, 2022
Budapest, Hungary
w https://conferences.renyi.hu/renyi100

Alfréd Rényi was born on the 20th of March 1921. To celebrate this occasion, the Alfréd Rényi Institute of Mathematics and the Hungarian Academy of Sciences are organising a high-profile conference, representing modern probability, graph theory and networks, information theory, dynamical systems, number theory and other fields in Rényi's spirit, i.e., not respecting strictly the borders between these and possibly other areas of pure and applied mathematics.

The conference will take place **June 20–23, 2022** (having been postponed from 2021 due to the pandemic) in Budapest, Hungary.

We invite posters to be presented during the conference. The deadline for all submissions is **8 May 2022**. Students can apply for a waiver of the registration fee: see https://conferences.renyi.hu/renyi100/registration-fee.



Other meetings around the world

7th International BFF (Bayesian, Fiducial, and Frequentist) conference
May 2–4, 2022
University of Toronto, Canada
w https://gravitypull.swoogo.com/BFF7

The 7th International BFF conference will be hosted by the Department of Statistical Sciences at the University of Toronto on May 2-4.

The Bayesian, Fiducial, and Frequentist (BFF) community began in 2014 as a means to facilitate scientific exchange among statisticians and scholars in related fields that develop new methodologies with in mind the foundational principles of statistical inference. The community encourages and promotes research activities to bridge foundations for statistical inferences, to facilitate objective and replicable scientific learning, and to develop analytic and computing methodologies for data analysis.

Over the last eight years, BFF meetings have served as a venue for researchers and practitioners to share ideas, update research progress, highlight important open problems in both theory and implementation, and most importantly discuss future directions of such research.

Previous iterations of the BFF have taken place at Duke University (BFF6), University of Michigan (BFF5), Harvard University (BFF4), Rutgers University (BFF3), and East China Normal University (BFF1 & BFF2).

This edition of the conference will focus is on methodological, computational and ethical principles of data science and is dedicated to the memory of our late colleague, Donald A. S. Fraser.

Registration and additional details about the event are available on the conference website: https://gravitypull.swoogo.com/BFF7

The conference will also host a virtual poster session on May 4th. Contributions are accepted until April 10.

Bocconi Summer School in Advanced Statistics and Probability: *Random structures and combinational statistics*
July 11–21, 2022. Lake Como, Italy
w https://bocconi2022.lakecomoschool.org/

The Summer School is designed for PhD students, and possibly brilliant MSc graduates (or final-year students) interested in pursuing doctoral studies in Statistics, Probability, Data Science, Machine Learning, Computer Science, Applied Mathematics, and related areas. **The School particularly welcomes applications from ethnic and gender groups that are under-represented in the field of statistics.**

The instructors are Luc Devroye (McGill University, Montreal) and Gabor Lugosi (Universitat Pompeu Fabra, Barcelona).

2022 Women in Statistics and Data Science (WSDS) conference
October 6–8, 2022. St. Louis, USA
w https://ww2.amstat.org/meetings/wsds/2022/submitanabstract.cfm

Abstract submission deadline: April 15.

Employment Opportunities

Australia: Sydney
The University of Sydney
Lecturer/Senior Lecturer in Data Science
https://jobs.imstat.org/job//61779982

China: Hong Kong
The University of Hong Kong
Tenure-Track Professor/Associate Professor/Assistant Professor in Business Analytics (several posts)
https://jobs.imstat.org/job//62033043

China: Shenzhen
The Chinese University of Hong Kong, Shenzhen
Faculty Openings in Statistics
https://jobs.imstat.org/job//61400990

Denmark: Aarhus
Technical University of Denmark, Department of Mechanical Engineering
Postdoctoral positions in Mathematics and Statistics
https://jobs.imstat.org/job//61917994

Germany: Bremerhaven
Biosciences, Deep Sea Ecology and Technology
PhD Position Analysis of ecological time-series using mathematical Tools (m/f/d)
https://jobs.imstat.org/job//62239641

United Kingdom: Coventry
University of Warwick
Research Fellow
https://jobs.imstat.org/job//62245078

United Kingdom: Leeds
University of Leeds
Professor of Statistics or Data Analytics
https://jobs.imstat.org/job//61759587

United States: Tempe, AZ
Arizona State University, School of Mathematical & Statistical Sciences
Lecturer Positions in Mathematical Sciences
https://jobs.imstat.org/job//61876224

United States: Tempe, AZ
Arizona State University, School of Mathematical & Statistical Sciences
Presidential Postdoctoral Fellowship in Mathematical Biology
https://jobs.imstat.org/job//62388302

United States: Berkeley, CA
University of California Berkeley
Lecturer, Department of Mathematics, College of Letters and Science
https://jobs.imstat.org/job//62116956

United States: La Jolla, CA
University of California San Diego
Assistant Professor (Ladder-rank): Broad Area search in Data Science (HDSI)
https://jobs.imstat.org/job//61394031

United States: Fruitland, ID
Treasure Valley Classical Academy
Mathematics and Science Teacher
https://jobs.imstat.org/job//60806737

United States: Bloomington, IN
IU School of Public Health
Chair of the Department of Kinesiology
https://jobs.imstat.org/job//61396763

United States: Baton Rouge, LA
Louisiana State University
Department Head/Chair of Experimental Statistics, Tenured
https://jobs.imstat.org/job//61550041

United States: Lincoln, NE
University of Nebraska Lincoln, Department of Statistics, IANR
Open Rank Professor, Statistical Geneticist
https://jobs.imstat.org/job//61779707

United States: Buffalo, NY
University at Buffalo, SUNY
Chair, Department of Biostatistics
https://jobs.imstat.org/job//61795562

United States: Memphis, TN
The University of Memphis, School of Public Health
Research Assistant/Associate Professor in Biostatistics
https://jobs.imstat.org/job//61876772

United States: Seattle, WA
Starbucks Coffee Company
Data Scientist Sr
https://jobs.imstat.org/job//62012168

International Calendar of Statistical Events

IMS meetings are highlighted in maroon with the ims logo, and new or updated entries have the NEW or UPDATED symbol. Please submit your meeting details and any corrections to Elyse Gustafson: ims@imstat.org

At the time of writing, some meetings are known to be POSTPONED or canceled. Where new dates are known, they are included here. Some meetings, marked ONLINE, are offering a virtual format. Please check meeting websites for updates.

Online and Ongoing

NEW ONLINE ims **Asia-Pacific Seminar in Probability and Statistics** w https://sites.google.com/view/apsps/home

ONLINE ims **COPSS–NISS COVID-19 Data Science Webinar series** w https://www.niss.org/copss-niss-covid-19-data-science-webinar-series

ONLINE ims **One World ABC Seminar** w https://warwick.ac.uk/fac/sci/statistics/news/upcoming-seminars/abcworldseminar

ONLINE ims **One World Probability Seminar** w https://www.owprobability.org/one-world-probability-seminar

ONLINE ims **One World YoungStatS Webinar series** w https://youngstats.github.io/categories/webinars/

ONLINE **Video series:** *The Philosophy of Data Science* w https://www.podofasclepius.com/philosophy-of-data-science

April 2022

April 25–29: Nicosia, Cyprus. **BNP Networking Event** w TBC

May 2022

May 12–18: Erice, Italy. **Stochastic Methods in Game Theory** w https://sites.google.com/view/erice-smgt2020/the-workshop

May 17–29: Istanbul. CODIT 2022 w https://codit2022.com/

NEW ims May 24–26: Honolulu, Hawaii. New Advances in Statistics and Data Science w http://zke.fas.harvard.edu/HawaiiConference/Main.html

NEW ims May 31–June 1: Duke University, USA. South East Probability Conference (1 of 2) w https://services.math.duke.edu/~rtd/SEPC2022/SEPC2022.html

June 2022

ims June 1–3: Berkeley, USA. **Statistics in the Big Data Era** w https://simons.berkeley.edu/workshops/statistics-big-data-era

NEW ims June 10–15: Online. 2022 WNAR/IMS/JR Annual Meeting w https://wnarofibs.wildapricot.org/WNAR2022

June 13–17: Charlottesville, VA, USA. **Ethical Challenges of AI in Biomedicine** w http://innovation.lab.virginia.edu/

June 14–18: Paphos, Cyprus. **International Symposium on Nonparametric Statistics, ISNPS** w http://cyprusconferences.org/isnps2022/

NEW ims June 20–23: Budapest, Hungary. **Rényi 100: celebrating 100 years since Rényi's birth** w https://conferences.renyi.hu/renyi100

June 20–23: Timisoara, Romania. **23rd Conference of the Romanian Society of Probability and Statistics** w https://spsr.ase.ro/conferinta-nationala-spsr/

June 25–July 1: Montreal, Canada. **ISBA World Meeting 2022** w https://isbawebmaster.github.io/ISBA2022

ims June 27–30: London, UK. **IMS Annual Meeting** w www.imsannualmeeting-london2022.com

Image: freepik.com/wwalakte

ims June 27–July 1: Wuhan, China. **Stochastic Processes and their Applications** w http://spa2022.whu.edu.cn

June 27–July 1: Darwin, Australia. **Joint Southern Statistical Meetings (JSSM2022)** w https://statsoc.org.au/event-3529236

Continues on page 29

International Calendar *continued*

July 2022

July 1: London, UK. **IMS–COLT one-day workshop** (between IMS meeting and COLT meeting)
w https://bguedj.github.io/colt-ims-2022.github.io/

July 10–15: Riga, Latvia. **XXXI International Biometric Conference (IBC2022)** **w** www.biometricsociety.org/meetings/conferences

POSTPONED ~~July 18–22: Moscow, Russia. **33rd European Meeting of Statisticians** **w** https://ems2022.org/~~

August 2022

August 2–13: Campinas, Brazil. **São Paulo School of advanced science on singular stochastic partial differential equations and their applications**
w https://www.ime.unicamp.br/spas2022//

NEW August 3–6: George Mason University, VA, USA. 22nd Meeting of New Researchers in Statistics and Probability **w** https://imstat.org/nrc2022/

August 6–11: Washington DC, USA. **JSM 2022**
w https://ww2.amstat.org/meetings/jsm/2022/

NEW August 8–9: UNC Chapel Hill, USA. South East Probability Conference (2 of 2) **w** https://services.math.duke.edu/~rtd/SEPC2022/SEPC2022.html

August 21–25: Newcastle, UK. **International Conference for Clinical Biostatistics** **w** http://www.iscb.info/

September 2022

September 7–10: UC Santa Cruz, CA, USA. **O'Bayes 2022** **w** https://obayes.soe.ucsc.edu/

September 12–15: Aberdeen, UK. **RSS International Conference**
w https://rss.org.uk/training-events/conference2022/

October 2022

October 24–28: Puerto Varas, Chile. **BNP13: Bayesian Nonparametrics** **w** https://midas.mat.uc.cl/bnp13/

December 2022

December 13–16: Florence, Italy. **IMS International Conference on Statistics and Data Science (ICSDS)**
w https://sites.google.com/view/icsds2022

December 18–20: Hong Kong. **ICSA International Conference**
w https://www.icsa.org/12th-icsa-international-conference-december-18-20-2022/

January 2023

January dates TBC (postponed from January 2022): Melbourne, Australia. **IMS Asia Pacific Rim Meeting (IMS-APRM2021)** **w** http://ims-aprm2021.com/

July 2023

UPDATED July 15–20: Ottawa, Canada. **ISI World Statistics Congress** **w** https://www.isi2023.org/

August 2023

August 5–10: Toronto, Canada. **IMS Annual Meeting at JSM 2023** **w** http://www.amstat.org/ASA/Meetings/Joint-Statistical-Meetings.aspx

July 2024

Dates TBC: Venice, Italy. **ISBA World Meeting 2024**
w https://bayesian.org/2024-world-meeting/

August 2024

August 3–8: Portland, OR, USA. **JSM 2024** **w** http://www.amstat.org/ASA/Meetings/Joint-Statistical-Meetings.aspx

August 12–16: Bochum, Germany. **Bernoulli/IMS World Congress in Probability and Statistics** **w** TBC

August 2025

August 2–7: Nashville, TN, USA. **IMS Annual Meeting at JSM 2025** **w** http://www.amstat.org/ASA/Meetings/Joint-Statistical-Meetings.aspx

Membership and Subscription Information: 2022

Journals

The scientific journals of the Institute of Mathematical Statistics are *The Annals of Statistics*, *The Annals of Probability*, *The Annals of Applied Statistics*, *The Annals of Applied Probability*, and *Statistical Science*. The *IMS Bulletin* is the news organ of the Institute.

Individual Memberships

Each individual member receives the *IMS Bulletin* (print and/or electronic) and may elect to receive one or more of the five scientific journals. Members pay annual dues of $105. An additional $120 is added to the dues of members for each scientific journal selected ($77 for *Stat Sci*). **Reduced membership dues** are available to full-time students, new graduates, permanent residents of countries designated by the IMS Council, and retired members.

Individual and General Subscriptions

Subscriptions are available on a calendar-year basis. **Individual subscriptions** are for the personal use of the subscriber and must be in the name of, paid directly by, and mailed to an individual. Individual subscriptions for 2022 are available to *The Annals of Applied Probability* ($235), *The Annals of Applied Statistics* ($235), *The Annals of Probability* ($235), *The Annals of Statistics* ($235), *Statistical Science* ($192), and *IMS Bulletin* ($115). **General subscriptions** are for libraries, institutions, and any multiple-readership use. Institutional subscriptions for 2022 are available to *The Annals of Applied Probability*, *The Annals of Applied Statistics*, *The Annals of Probability*, and *The Annals of Statistics* (each title $536 online only / $649 print+online), *Statistical Science* ($307/$370), and *IMS Bulletin* ($153 print). Airmail delivery is no longer offered.

IMS Bulletin

The *IMS Bulletin* publishes articles and news of interest to IMS members and to statisticians and probabilists in general, as well as details of IMS meetings and an international calendar of statistical events. Views and opinions in editorials and articles are not to be understood as official expressions of the Institute's policy unless so stated; publication does not necessarily imply endorsement in any way of the opinions expressed therein, and the *IMS Bulletin* and its publisher do not accept any responsibility for them. The *IMS Bulletin* is copyrighted and authors of individual articles may be asked to sign a copyright transfer to the IMS before publication.

The *IMS Bulletin* (ISSN 1544-1881) is published eight times per year, in January/February, March, April/May, June/July, August, September, October/November and December, by the Institute of Mathematical Statistics, 9760 Smith Rd, Waite Hill, Ohio 44094, USA. Periodicals postage paid at Cleveland, Ohio, and at additional mailing offices. Postmaster: Send address changes to 9760 Smith Rd, Waite Hill, Ohio 44094, USA or dues.subs@imstat.org. Printed by The Sheridan Press, 450 Fame Avenue, Hanover, PA 17331, USA.

Information for Advertisers

General information: The *IMS Bulletin* and webpages are the official news organs of the Institute of Mathematical Statistics. The *IMS Bulletin*, established in 1972, is published 8 times per year. Print circulation is around 3,500 paper copies, and it is also free online in PDF format at https://www.imstat.org/ims-bulletin-archive/, posted online about two weeks before mailout (average downloads over 8,000). Subscription to the *IMS Bulletin* costs $115: call 877-557-4674 (US toll-free) or +1 216 295 2340 (international), or email dues.subs@imstat.org. The IMS website, https://imstat.org, established in 1996, receives over 30,000 visits per month.

Advertising job vacancies: A single 60-day online job posting costs just $329.00. We will also include the basic information about your job ad (position title, location, company name, job function and a link to the full ad) in the *IMS Bulletin* at no extra charge. See https://jobs.imstat.org

Advertising meetings, workshops and conferences: Meeting announcements here and on the IMS website at https://imstat.org/meetings-calendar/ are free. Submit your meeting details at https://www.imstat.org/ims-meeting-form/

Rates and requirements for display advertising: Display advertising allows for placement of camera-ready ads for journals, books, software, etc. A camera-ready ad should be sent as a grayscale PDF (min. 300dpi, with all fonts embedded). Email your advert to Elyse Gustafson ims@imstat.org or see https://imstat.org/advertise

	Dimensions: width x height	Rate
1/3 page *horizontal*	4.93" wide x 4.0" high (125.5 x 102 mm)	$285
1/3 page *vertical*	2.39" wide x 9.42" high (60.7 x 239.1 mm)	$285
1/2 page *horizontal*	7.5" wide x 4.7" high (190.5 x 119.4 mm)	$355
1/2 page *vertical*	3.67" wide x 9.42" high (93.1 x 239.1 mm)	$355
Full page (to edge, including ⅛" bleed)	8.75" wide x 11.25" high (222 mm x 286 mm)	$480
Full page (within usual Bulletin margins)	7.5" wide x 9.42" high (190.5 mm x 239.1 mm)	$480

Deadlines and mailing dates for *IMS Bulletin*

Issue	Deadline	Online by	Mailed
1: January/February	**December 1**	December 15	January 1
2: March	**February 1**	February 15	March 1
3: April/May	**March 15**	April 1	April 15
4: June/July	**May 1**	May 15	June 1
5: August	**July 1**	July 15	August 1
6: September	**August 15**	September 1	September 15
7: Oct/Nov	**September 15**	October 1	October 15
8: December	**November 1**	November 15	December 1

Volume 51 · Issue 3

the next issue is

June/July 2022

Read IMS Bulletin articles online at https://imstat.org/news

DEADLINES for submissions

May 1, then **July 1**

Please see inside the back cover for subscription details and information for advertisers, including all our deadlines and requirements

Journal alerts

For email alerts when new IMS journal issues are released, sign up at **https://imstat.org/portal/login**

The purpose of the Institute is to foster the development and dissemination of the theory and applications of statistics and probability

IMS: Organized September 12, 1935