

# IMS Bulletin

#### April/May 2023

#### **CONTENTS**

#### 1 International Prize in Statistics: C.R. Rao

2–3 **Members' news:** Jonathan Niles-Weed, Alain-Sol Sznitman, Yong Zeng, Jun Zhu, Yulia Gel, Edsel Peña, Kanti Mardia, Catherine Calder, Marek Kimmel, Ping Ma, Wendy Martinez, Peihua Qiu, Hui Zou

#### 4–7 COPSS Awards

- 8 Written by Witten: Tenure is not the hardest thing
- 9 Spiegelman Award
- 10 XL-Files: ChatGPT–First Contact
- 12 **Obituary:** Don Ylvisaker
- 14 **Recent papers:** *Probability Surveys* and *Stochastic Systems*
- 15 **IMS Lecture Previews:** Bin Yu, Wing Hung Wong, Massimiliano Gubinelli
- 17 IMS New Researchers Conference
- 18 Student Puzzle 44
- 19 IMS Council candidates
- 26 New ACM-IMS Journal of Data Science
- 27 Meetings (including online)
- 32 Employment Opportunities
- 32 Calendar of Meetings
- 35 Information for Advertisers

Read it online: imstat.org/news

ΠŢ

# **International Prize in Statistics**

**C.R. Rao**, an Indian professor whose work over 75 years ago continues to exert a profound influence on science, has been awarded the 2023 International Prize in Statistics.

In his remarkable 1945 paper published in the *Bulletin of the Calcutta Mathematical Society*, Calyampudi Radhakrishna (C.R.) Rao demonstrated three fundamental results that paved the way for the modern field of Statistics and provided statistical tools heavily used in science today. The first, now known as the Cramér–Rao lower bound, provides a means for knowing when a method for estimating a quantity is as good as any



method can be. The second result, named the Rao–Blackwell Theorem (because it was discovered independently by eminent statistician David Blackwell), provides a means for transforming an estimate into a better – in fact, an optimal – estimate. Together, these results form a foundation on which much of Statistics is built. And the third result provided insights that pioneered a whole new interdisciplinary field that has now flourished as "information geometry." Combined, these results help scientists more efficiently extract information from data.

Information geometry has recently been used to aid the understanding and optimization of Higgs boson measurements at the Large Hadron Collider (LHC), the world's largest and most powerful particle accelerator. It has also found applications in recent research on radars and antennas and contributed significantly to advancements in artificial intelligence, data science, signal processing, shape classification and image segregation.

The Rao–Blackwell process has been applied to stereology, particle filtering, and computational econometrics among others, while the Cramér–Rao lower bound is of great importance in such diverse fields as signal processing, spectroscopy, radar systems, multiple image radiography, risk analysis and quantum physics.

"In awarding this prize, we celebrate the monumental work by C.R. Rao that not only revolutionized statistical thinking in its time but also continues to exert enormous influence on human understanding of science across a wide spectrum of disciplines," said Guy Nason, Chair of the International Prize in Statistics Foundation.

The International Prize in Statistics is awarded every two years by a collaboration among five leading international statistics organizations. The prize recognizes a major

# **IMS** Bulletin

Volume 52 • Issue 3 April/May 2023 ISSN 1544-1881

### **Contact information**

IMS Bulletin Editor: Tati Howell bulletin@imstat.org

#### UPDATED Managing Editor: Dan Nordman

Contributing Editors: Radu Craiu, Anirban DasGupta, Ruobin Gong, David Hand, Takis Konstantopoulos, Xiao-Li Meng and Daniela Witten

#### 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**

President:	Peter Bühlmann president@imstat.org
President-Elect:	Michael Kosorok president-elect@imstat.org
Past President:	Krzysztof (Chris) Burdzy president-past@imstat.org
Treasurer:	Jiashun Jin jiashun@stat.cmu.edu
Program Secretary:	Annie Qu aqu2@uci.edu
Executive Secretary:	Edsel Peña pena@stat.sc.edu

# **IMS Members' News**

#### Jonathan Niles-Weed receives 2023 IMS Tweedie New Researcher Award

The IMS Tweedie New Researcher Award provides funds for travel to present the Tweedie New Researcher Invited Lecture at the IMS New Researchers Conference. The recipient of the 2023 Tweedie New Researcher Award is **Jonathan Niles-Weed**.

Dr. Niles-Weed is an Assistant Professor of Mathematics and Data Science at the Courant Institute of Mathematical Sciences and Center for Data Science at New York University. He was selected by the IMS Committee on Travel Awards: "For driving the development of statistical optimal transport and expanding the boundaries of statistics to problems with a geometric



structure by drawing connections with probability, analysis, and computation."

Apply by April 28 to attend the New Researchers Conference, if you received your PhD in or since 2016, or expect to by the end of 2023. More details can be found on the conference website: https://sites.google.com/site/linbowangpku/nrc-2023.

#### Alain-Sol Sznitman elected to German Academy

IMS Fellow **Alain-Sol Sznitman**, ETH Zurich, has been elected to the German National Academy of Sciences, Leopoldina. See www.leopoldina.org.

Professor Sznitman has been full Professor at the ETH Zurich, Department of Mathematics, since 1991. Born on December 13, 1955 in Paris, he studied at the Ecole Normale Supérieure in Paris. Following his studies he became Attaché de Recherche at CNRS in 1979, and worked at the Université Pierre et Marie Curie. In 1983 he received the Doctorat d'Etat, became Chargé de Recherche, and then spent two years at the Courant Institute in New York, returning to Paris in September 1985. In 1987 he was appointed associate professor at the Courant Institute, and in 1990, full professor. Honors include an Alfred P. Sloane Fellowship (1989); the Rollo Davidson Prize (1991); plenary lectures at the First European Congress of Mathematics in Paris (1992) and the International Congress of Mathematicians in Berlin (1998); the Line and Michel Loève Prize (1999); membership of the Academia Europaea (2008), the Inaugural Class of AMS Fellows (2012) the IMU Circle (2014). In 2022 he was the recipient of the Blaise Pascal Medal in Mathematics and became member of the European Academy of Sciences.

Alain-Sol Sznitman's research concerns probability theory, with a special focus on problems connected with physics, in particular problems of random media. He is the author of several scientific publications and member of several scientific committees.

#### **NSF Appoints New Statistics Program Directors**

**Yong Zeng**, University of Missouri–Kansas City, has become a permanent program director and **Jun Zhu**, University of Wisconsin–Madison, became a rotator program director of the Division of Mathematical Sciences in the National Science Foundation Directorate of Physical and Mathematical Sciences. They join **Yulia Gel** of The University of Texas at Dallas and **Edsel Peña** of the University of South Carolina–Columbia, who are in their second and third years, respectively, as rotator program directors of the statistics program.

IMS Bulletin • 3

#### Kanti Mardia in UK Honours List

IMS member and Fellow **Kanti Mardia** has been made an Officer of the Order of the British Empire (OBE). A senior research professor at the University of Leeds, UK, Mardia is a distinguished Indian statistician who is known for his pioneering work on multivariate analysis, directional data, and statistical measures of shape. His book *Multivariate Analysis* is well known among statistics students. Mardia is the founder of the Royal Statistical Society's Mardia Prize, awarded for interdisciplinary research. Among his many honors, Mardia received the Guy Medal in Silver by the RSS and the Wilks Memorial Award by the American Statistical Association. He was also awarded the IISA Lifetime Achievement Award in 2019.

Kanti Mardia was appointed Officer of the Order of the British Empire (OBE) in the 2023 New Year Honours for services to statistical science.

#### American Association for the Advancement of Science elects Statistics Section Fellows

In October, the American Association for the Advancement of Science (AAAS) council elected nearly 505 members as fellows, and a virtual induction ceremony took place on February 13. Election as a fellow honors members 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 eight new fellows of AAAS Section U (Statistics), six of whom are IMS Fellows or members, are:

Catherine A. Calder, The University of Texas at Austin Jacqueline M. Hughes-Oliver, North Carolina State University Marek Kimmel, Rice University Ping Ma, University of Georgia Wendy L. Martinez, US Census Bureau Dionne L. Price, US Food and Drug Administration Peihua Qiu, University of Florida Hui Zou, University of Minnesota

#### International Prize in Statistics: continued from cover

achievement by an individual or team in the statistics field, particularly an achievement of powerful and original ideas that has led to practical applications and breakthroughs in other disciplines.

Rao will receive the prize, which comes with an \$80,000 award, this July at the biennial International Statistical Institute World Statistics Congress, which will be held in Ottawa, Ontario, Canada.

The first International Prize in Statistics was awarded in 2017 to **David R. Cox** for the development of the Cox proportional hazards model, which allows researchers to investigate patient survival rates in complex studies. **Bradley Efron** received the award in 2019 for a statistical method known as the bootstrap, a clever computational method for assessing uncertainty in applied statistics. **Nan Laird** received the award in 2021 for the development of powerful methods that have made possible the analysis of complex longitudinal studies.

More information about the life and work of C.R. Rao can be found in many places, including: https://mathshistory.st-andrews.ac.uk/Biographies/Rao/, https://www.psa. gov.in/article/prof-calyampudi-radhakrishna-rao/347, and https://en.wikipedia.org/wiki/C.\_R.\_Rao.

### m = 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 Mhttps://projecteuclid.org/imsc

IMS Monographs and IMS Textbooks: Mark Handcock https://www.imstat.org/journals-andpublications/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 Mhttps://projecteuclid.org/euclid.ejp

Electronic Communications in Probability: Siva Athreya

Mhttps://projecteuclid.org/euclid.ecp

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

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

Statistics Surveys: Yingying Fan 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 Mhttps://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 Journa

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

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

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



### **COPSS Awards 2023**

### **Bin Yu: Distinguished Achievement Award and Lectureship**

**Citation:** For fundamental contributions to information theory, statistical and machine learning methodology; interdisciplinary research in fields such as genomics, neuroscience, remote sensing and document summarization; and for outstanding dedication to professional service, leadership, and mentoring of students and young scholars.

Haiyan Huang, Kathryn Roeder, and Jianwen Cai write: The Committee of Presidents of Statistical Societies (COPSS) has selected Bin Yu from the University at California (UC) at Berkeley for the 2023 Distinguished Achievement Award and Lectureship (DAAL). Formerly known as the R.A. Fisher Award and Lectureship, the DAAL recognizes meritorious achievement and scholarship in statistical science and recognizes the highly significant impact of statistical methods on scientific investigations. She will deliver the DAAL Lecture at the 2023 Joint Statistical Meetings (JSM) in Toronto on veridical data science.

Yu's current research focuses on practice, algorithm, and theory of statistical machine learning, interpretable machine learning, and causal inference. Her group is engaged in interdisciplinary research with scientists from genomics, neuroscience, and precision medicine. She and her group have developed the PCS framework for veridical data science towards responsible, reliable, and transparent data analysis and decision-making. PCS stands for predictability, computability and stability. It unifies, streamlines, and expands on ideas and best practices of machine learning and statistics to uncover and address a hidden universe of uncertainties well beyond sample–sample uncertainty in a data science life cycle.

In the past, she jointly developed a highly cited spatially adaptive wavelet image denoising method and a low-complexity low-delay perceptually lossless audio coder that was incorporated in Bose wireless speakers. She also co-developed a fast and well-validated Arctic cloud detection algorithm. Her collaborative paper in 2011 with the Gallant Lab at Berkeley on movie reconstruction from fMRI brain signals received extensive and intensive coverages by numerous media outlets, including The Economist, Forbes, Der Spiegel, Daily Mail, New Scientist and Massachusetts Institute of Technology (MIT) Technology Review. This work was named one of the best 50 inventions in 2011 by Time magazine. She and collaborators mapped a cell's destiny in Drosophila via stability-driven NMF, and used the PCS framework to stress-test or internally validate clinical decision rules used in the ER. Previously, she pioneered Vapnik-Chervonenkis (VC) type theory needed for asymptotic analysis of time series and spatio-temporal processes. She made fundamental contributions to information theory and statistics through work on minimum description length (MDL) and entropy

estimation. Recently, she and her collaborators developed iterative random forests (iRF), X-learner for heterogeneous treatment effect estimation in causal inference, hierarchical shrinkage (HS) decision trees, and Fast and interpretable greedy trees (FIGS).

She is the Class of 1936 Second Chair in the College of Letters and Science and Chancellor's Distinguished Professor, Departments of Statistics and Electrical Engineering & Computer Sciences, and Center for Computational Biology, at the University of California at Berkeley. She obtained her BS in Mathematics from Peking University and her MS and PhD in Statistics from UC Berkeley. She was an Assistant Professor at University of Wisconsin-Madison, Visiting Assistant Professor at Yale University, a member of the Technical Staff at Lucent Bell Labs, and a Miller Research Professor at Berkeley. She was a Visiting Faculty at MIT, ETH, Poincare Institute, Peking University, INRIA-Paris, Fields Institute at the University of Toronto, Newton Institute at Cambridge University, and the Flatiron Institute in New York City. She has also served as Chair of the Department of Statistics at UC Berkeley and had a crucial role in envisioning the intellectual and organizational vision for the Division of Computing, Data Science, and Society (CDSS) at UC Berkeley as a faculty advisory committee member.

Yu is a Member of the US National Academy of Sciences and the American Academy of Arts and Sciences. She was President of the IMS in 2013–14, a Guggenheim Fellow, Tukey Memorial Lecturer of the Bernoulli Society, and IMS Rietz Lecturer. In 2018, Yu was awarded the Elizabeth L. Scott Award by COPSS for principled leadership in the international scientific community; for commitment and actions towards diversity, equity, and inclusion; for consistently mentoring and encouraging women students and new researchers in statistics and data science; and for scientific contributions to statistical and machine learning methodology at the highest scholarly level. She holds an Honorary Doctorate from The University of Lausanne, and served on the inaugural scientific advisory board of the UK Turing Institute of Data Science and AI. She is serving on the editorial board of PNAS and as a senior advisor at Simons Institute for the Theory of Computing at Berkeley.

Bin Yu will also give this year's IMS Wald Memorial Lectures at JSM in 2023 [*see preview on page 15*].

IMS Bulletin • 5



Bin Yu: Distinguished Achievement Award & Lectureship



Michael Kosorok: George W. Snedecor Award



Karen Bandeen-Roche: F.N. David Awara

### Michael Kosorok: COPSS George W. Snedecor Award

**Citation:** For foundational, creative, and original contributions to mathematical statistics; for methodological developments in empirical processes and machine learning; for advancement of precision health; and for mentoring of students, postdocs, and junior faculty.

Sebastien Haneuse and Eric Laber write: Michael Kosorok is the consummate biostatistics methodologist. His work is creative, rigorous, and motivated by urgent and impactful problems in biomedicine. He has authored more than 200 peer-reviewed manuscripts appearing in top-tier journals and conference proceedings. In 2008 he published his monograph Introduction to Empirical Processes and Semiparametric Inference, which quickly became the canonical introduction to the area. Shortly thereafter, Kosorok focused his research efforts on artificial intelligence and precision medicine. He was among the first to provide rigorous theoretical results for machine learning methods for estimation of optimal treatment regimes including some of the earliest applications of Q-learning and direct search estimation. His paper on outcome weighted learning (which has been cited nearly 750 times) began long and fruitful lines of research on direct-search estimation and revealed important connections between optimal treatment regimes and classification.

In addition to his prolific publication record, Kosorok has shaped the field through his service and mentoring. He served as the head of the Biostatistics Department at UNC from 2006–20, chair of the COPSS Presidents' Awards committee, and is currently President-elect of the IMS. He has mentored more than 50 PhD students, many of whom now hold prominent positions at academic institutions or in industry.

In the publication cited in the award (Nguyen et al, 2020), Kosorok along with his co-authors, developed a novel direct-search estimator of the optimal regime under a 2×2 crossover design. Such crossover designs are common in pilot testing, rare diseases, and other settings in which recruiting a large pool of participants is difficult. Nevertheless, prior to this publication, there were no direct search estimators for estimating an optimal treatment regime under such a design. The proposed method accounts for carryover effects, uses a convex relaxation for computational efficiency, and is Fisher consistent.

This publication is an illustration of Kosorok's research modus operandi. He identifies an important practical problem, develops a novel methodological approach, and then provides a rigorous and complete description of the method's operating characteristics.

### Karen Bandeen-Roche: COPSS F.N. David Award

**Citation:** For outstanding leadership and service in the biostatistics and statistics community, for her leadership in statistical education, and for her achievements in biostatistical research, particularly in the field of aging research and frailty.

Brian Caffo, Ingo Ruczinski and Catherine Crespi write: For the past 15 years Karen Bandeen-Roche has served as the Hurley Dorrier Professor and Chair of the Department of Biostatistics at the Johns Hopkins Bloomberg School of Public Health. She is a highly established biostatistician with signature leadership in research,

#### the past education and administration.

A biostatistical generalist in her methods research, Karen has published nearly 300 peer-reviewed manuscripts in a wide range of topics and journals. She has contributed greatly to latent variable model methodology and led methods work in the medical and

# **COPSS Awards** continued

**Karen Bandeen-Roche, COPSS F.N. David Award** continued epidemiology literature. In particular, Karen has been a tireless leader in the promotion of biostatistical thinking in general, especially in the field of aging. Her leadership roles at the Johns Hopkins Center on Aging and Health have guided an organization that is the size of a large department of statistics. Karen is a Fellow of both the American Statistical Association (ASA) and the Gerontological Society of America, recognizing her contributions as a bridge researcher who is helping to close the gap between aging research and biostatistics.

Karen has been an active participant and played leadership roles in many of the major biostatistical and statistical organizations. Of particular relevance are her service as chair and executive committee member of the International Biometric Society (IBS) Eastern North American Region (ENAR), chair and executive committee member of the Caucus on Academic Representatives of the ASA, chair of the Biometrics section of the ASA, and chair of the Biostatistical Methods and Research Design (BMRD) National Institutes of Health (NIH) study section, which represents the primary funding outlet for methodological, theoretical and general applied biostatistical research. Karen also has been elected to the IBS Executive Board, which allows her to shape the vision and activities of the International Biometric Society. Since 2020, Karen also serves on the Board of Directors for the National Institute of Statistical Sciences.

In education, Karen's leadership has expanded biostatistical instruction including onsite, hybrid and online teaching at all levels. Under her leadership the department expanded its teaching role outside of the university, having a large presence in online open education, including Massive Open Online Courses (MOOCs). At the same time Karen remains a dedicated and passionate in-person educator herself and is regularly recognized as such. She has won the Bloomberg School's highest teaching award, the "Golden Apple Award for Excellence in Teaching", an annual award given to only four faculty members among the entire teaching faculty (offering more than 600 courses annually in the Bloomberg School of Public Health). In addition, she is a three-time recipient of the "Advising, Mentoring, and Teaching Recognition Award", an award given annually by the students of the Johns Hopkins Bloomberg School of Public Health in recognition of outstanding educational contributions.

On July 1, 2023 Karen will step down as the chair of the Johns Hopkins Department of Biostatistics, handing over a department in excellent financial shape and co-ranked as number 1 in the nation according to US News and World Report.

# Meet the 2023 COPSS Emerging Leader Awardees

Maya Sternberg, the COPSS Secretary/Treasurer, introduces the eight new COPSS Emerging Leaders. In creating this award, COPSS recognizes the increasingly important role that early-career statistical scientists are playing in shaping the future of the discipline. The award is designed both to call attention to the efforts of these individuals and to provide a mechanism for them to share their vision for the field with each other and the statistical community. The selection criteria for the award are intentionally broad and designed to highlight contributions in areas not traditionally recognized by other early-career awards in the statistical sciences. The awardees are:



#### Peng Ding

University of California, Berkeley For outstanding contributions to the foundations and applications of causal inference, for both randomized experiments and observational studies, with emphasis on settings with high-dimensional covariates and complex structures.

Peng Ding

#### **COPSS Emerging Leader Award** continued



Edgar Dobriban

#### Edgar Dobriban

University of Pennsylvania

For fundamental contributions to the development of random matrix theory-based statistical methods, theory for analyzing massive datasets, uncertainty quantification in machine learning, including parallel analysis for principal component analysis, distributed statistical learning, scalable inference via random projections; for innovative methods for COVID-19 pooled testing using hypergraph factorization; for outstanding mentoring.

#### Jingyi Jessica Li

University of California, Los Angeles

For innovative and disruptive research at the junction of statistics and biology, especially in statistical genomics. For advocacy of the importance of statistical rigor in the biomedical science community. For outreach efforts and commitment to improve the diversity in quantitative research.



Gongjun Xu

### University of Michigan

For making breakthroughs on challenging problems in the behavioral sciences, for significant theoretical and methodological contributions to latent variable models, high-dimensional inference, survival analysis, and for outstanding editorial services and leadership.

#### Lorin Crawford

Yates Coley

Gongjun Xu

#### Microsoft Corporation & Brown University

Kaiser Permanente Washington Health Research Institute

For path-breaking research combining theory and methods of mathematics, statistics and computing to generate new knowledge and insight about the genetic basis of disease, and for exceptional mentoring of students from multiple scientific disciplines.

For impactful statistical contributions in the areas of ethical clinical prediction model development and learning health systems science; for significant leadership and advocacy to advance

justice, equity, diversity, and inclusion in the profession and practice of statistics.



Yates Coley

#### Avi Feller

University of California, Berkeley

For groundbreaking research in causal inference and program evaluation; for bridging statistics, public policy, and education research; and for commitment to building a more inclusive field.



Veronika Rockova

#### Veronika Rockova

#### University of Chicago

For groundbreaking contributions to theory and methodology at the intersection of Bayesian and frequentist statistics, for outstanding editorial service to the profession, and for excellence in the advising and supervision of doctoral students



Jingyi Jessica L



Lorin Crawford



Avi Feller

### Written by Witten: Tenure is not the hardest thing

On a cloudy February day in Seattle in early 2014, I was feeling pretty good. My tenure process had made it through all but the final stage of approval (which is typically a formality), my husband and I were expecting our first baby in just over five weeks, and I was building a beautiful life surrounded by friends in a city I loved. I had never faced a challenge I couldn't overcome, and—although, as a statistician, I should have known better than to extrapolate—I foresaw only unicorns and rainbows in my future.

Then, suddenly, I did not feel good. Quite the opposite: I felt excruciating pain. I tried to check myself in to Labor & Delivery, where a male OB turned me away after cheerily informing me that this was all a normal part of the third trimester of pregnancy, and that I should return to see him in five weeks, but not a minute before, when it was my time to have a baby.

I listened to that doctor (who, I realized later, had himself never experienced labor!) like the diligent student that I have always been, and continued to go to work for the next two days, occasionally sending understated texts to my husband ("1... CANNOT BELIEVE... THE NEXT... 5 WEEKS... ARE GOING TO FEEL THIS WAY") and literally crying during a meeting with a grad student about a paper revision\*. Finally, 48 hours in, my husband dragged me sobbing to the car and drove me back to Labor & Delivery, where that same doctor said "You're having a baby!" with no apparent memory of our interaction just 48 hours earlier.

So, a few hours later I found myself lying on a hospital bed, immobile from the midsection down, after a major surgery that involved cutting through five layers of abdominal wall (skin, subcutaneous tissue, fascia, muscle, peritoneum) and then slicing open my uterus to extract a baby the size of a large melon. Baby was in the NICU, and I was two floors away because—and you truly cannot make this stuff up—the hospital bed could not fit in the NICU. So there I lay, staring at the ceiling, until a nurse finally handed me my cell phone. And there it was: the official email notification that the university had approved my tenure. I swiped it away without opening it to stare at the photo of my baby that my husband had texted from the NICU.

During the ensuing nine years, I have had two more children, and have found that a research career is actually quite different from parenting. Some of the major differences are summarized in Table 1.

	Career in research	Parenting
Sleep	I occasionally stay up late for a deadline, or sleep poorly due to stress.	Four months of absolutely no sleep at all with my first two children. My third child (almost 4 years old) still wakes up overnight. I hope that when my youngest goes to college I'll no longer be woken up overnight by a child, but by then I'll also be too old to sleep well.
Credit for ideas	For the most part, if I publish quickly, then I get credit for my ideas.	None whatsoever. I can tell a tantruming toddler that they need a snack, and they won't listen. When I finally succeed in shoving a handful of crackers and a cup of milk down their throat, they will stop crying, but they definitely won't remember it was my idea in the first place.
Major surgery	No.	Three times.
Unwanted bodily fluids	No.	Pee and poop and spit-up, often, especially during the early years (but the trauma endures). Occasional vomit for the foreseeable future. Snot often and ongoing.
Payment for my time	Yes.	No, at least in the US, where parenting is treated like an ill-advised and extremely expensive hobby, with no guarantee of paid parental leave, or even access to healthcare, much less financial assistance during the next 18 years and beyond.
Other recog- nition	Major awards from professional societies, praise from my colleagues, appreciation from my students, etc.	Occasional hugs from kids.
Judgment from others	Yes, from reviewers, editors, hiring com- mittees, award committees, colleagues, and students.	Yes, from literally everyone with a pulse.

#### Table 1: Several key ways in which a career in research is different from parenting.

April/May · 2023



Quite honestly, I am amazed that for all of our society's so-called progress in science and medicine, the process of having a baby still looks like this. The birthing parent's life is literally put at risk—in 2021, the US had 32.9 maternal deaths per 100,000 live births, with much worse outcomes for non-Hispanic Black women (69.9 deaths per 100,000 births)\*\*. Treating the chance of death associated with each live birth as an independent event (which I know I shouldn't, but please bear with me—pun intended—as I'm trying to make a point here), this means that I had a 0.1% chance of dying in childbirth, and my Black counterpart a 0.2% chance. As a general policy, if I'm going to do something that carries with it a 0.1% chance of death, then at a minimum I expect honor, recognition, payment, and—truly, I don't think this is too much to ask—a Lifetime documentary about my bravery and fortitude (and certainly not a hefty hospital bill!!). I also firmly believe that every parent deserves a full-sized statue in their honor, and that single parents and those without complete financial security deserve several.

The process of becoming a parent has made me a better researcher, because I've learned the importance of patience: much like getting shoes on my three-year-old's feet, sometimes the research process just can't be rushed. And it's made me a better research mentor, not because my grad students are like small children, but because they are not: they are functional adult humans who are fully capable of making good choices, with or without my guidance, and who absolutely never poop in my office.

Becoming a parent has also made the non-research aspects of my job feel easier: navigating interpersonal conflicts at work is much easier than navigating them at home. This is because all of my colleagues have fully-developed prefrontal cortexes\*\*\*.

Now that my youngest child is almost four, I'm emerging from the deep woods of parenting small children, and I can reflect back on these past nine years. No, getting tenure is not the hardest thing I've ever done. It also is not the most rewarding. After all, my publications will never love me back.

But, I will also never need to help them poop.

Daniela Witten lives in Seattle with her family. She would like any future parents reading this column to know that baby poop is really not that bad. It's toddler poop that you need to watch out for.

- \* The revision was going fine, but I was not. Sorry again, Shizhe Chen.
- \*\* https://www.cdc.gov/nchs/data/hestat/maternal-mortality/2021/maternal-mortalityrates-2021.htm
- \*\*\* https://en.wikipedia.org/wiki/Prefrontal\_cortex and the minutes for my department's recent faculty meetings.

### Mortimer Spiegelman Award nominations

The Applied Public Health Statistics Section of the American Public Health Association

(APHA) presents the Mortimer Spiegelman Award annually to honor a statistician below the



age of 40 in the calendar year of the award who has made outstanding contributions to health statistics, especially public health statistics. The award, established in 1970, is presented annually at the APHA meeting. The award serves to: honor the outstanding achievements of both the recipient and Spiegelman; encourage further involvement in public health by the finest young statisticians; and increase awareness of APHA and the Applied Public Health Statistics Section in the academic statistical community.

The Spiegelman Award recipient must be a health statistician who has made outstanding contributions to statistical methodology and its applications in public health (broadly defined).

The award is open to early-career investigators regardless of race, gender, sexual orientation, nationality or citizenship. Specifically, candidates must either be under age 40 throughout the award calendar year; or have obtained a terminal degree in statistics or a statistics-related field in the last 10 years. (For extenuating circumstances, the committee will make exceptions to the eligibility requirements: see https://www. spiegelmanaward.org/ for more information.)

Nominations are due by **June 1**. Check website for nomination requirements.

# XL-Files: ChatGPT — First Contact



#### Xiao-Li Meng chats about ChatGPT:

By now, it's worth having a chat if you know someone who *hasn't* chatted with or about ChatGPT. Everyone I've talked to about it has been either impressed or depressed, sometimes concomitantly, with a few wanting to suppress the whole thing, almost contemptuously.

My first try at ChatGPT (based on GPT-3.5) was prompted by the need of having a fundraising strategy for *Harvard Data Science Review*, as I continue my role as its Editor-in-Chief after taking a sabbatical leave. I provided the simple instruction: "Provide a fundraising strategy for Harvard Data Science Review" on Feb. 25, 2023, and it immediately produced a list of eight steps:

"To develop a successful fundraising strategy for the Harvard Data Science Review, here are some potential steps that could be taken:

- Identify potential donors: Start by identifying potential donors who have an interest in data science and are likely to support the mission and goals of the journal. This could include alumni, foundations, corporations, and individuals.
- 2. Develop a case for support: Develop a clear and compelling case for why the journal is important and why it deserves financial support. Highlight the journal's impact on the field of data science, its unique contributions, and its plans for the future.

- 3. Establish giving levels: Create giving levels with corresponding benefits to encourage donors to contribute at different levels. This could include naming opportunities, special recognition, or invitations to events.
- 4. Host events: Host events to bring potential donors together and build relationships with them. This could include a launch event for a new issue, a donor appreciation event, or a panel discussion on a topic related to data science.
- Leverage social media: Use social media to promote the journal and its fundraising efforts. Share success stories, progress updates, and behindthe-scenes content to keep donors engaged and informed.
- 6. Apply for grants: Research and apply for relevant grants and funding opportunities that align with the mission and goals of the journal.
- 7. Seek corporate sponsorships: Reach out to companies in the data science field and seek sponsorships for the journal. This could include corporate underwriting for specific issues or ongoing sponsorship for the journal as a whole.
- 8. Develop a planned giving program: Develop a planned giving program to encourage donors to include the Harvard Data Science Review in their estate plans. This can help to secure long-term support for the journal.

By implementing these strategies, the Harvard Data Science Review can build relationships with potential donors and secure the financial support it needs to continue its important work in advancing the field of data science."

There are no million-dollar ideas here that I'm unfamiliar with, but I'm unaware of any human strategists could provide (prior to ChatGPT) such a list in an instant, regardless of their worth. In fact, my one-line instruction would be considered too laconic for a human employee. This is a disruptive technology that will replace some human jobs and do much better, just as calculators made collectors' items out of abacuses. The question is not when or how, since it's happening in real time, but to what extent.

Of course, I trust that few of you are as unwise as I am to get myself into needing a fundraising plan. But I am sure we all can use tools like ChatGPT to help our research, as many of us have already tried. Having burdened myself with projects that date back to last century (yes, I am still working on a paper initiated in 1999), my ultimate wish would be to dictate a research idea to ChatGPT while pouring myself a glass, and by the time I sit down, a draft would be ready for some personalization (with the help of a dosage of Chinglish).

Don't worry, I am neither wishing nor suggesting that ChatGPT would or could replace me intellectually. Being an educator means that I am always in the business of creating better minds than mine, and hence being replaced is a feature, not a bug. However, I do not wish to be replaced-at least, not yet-because we are living in an exciting but dark (digital) age, with almost all of us in the deep dark when it comes to the science behind ChatGPT and other similar products. The absence of information makes it so that we can never be sure if such an engineering masterstroke would not inflict a societal massive stroke. It is an unlikely event, just as a stroke does not happen to most people. But our collective and enduring mental exercise is in order, as a preventive measure, just as habitual healthy physical and mental lifestyles can downshift the risk of stroke.

Indeed, my limited interaction with

ChatGPT gave me a taste of wonder and worry, even just within the academic research enterprise, where the stakes are low (as in, "In academia, the politics are so high because the stakes are so low," a corollary of Sayre's law). Because I lead a research course on astrostatistics, I am a coauthor on an article to provide statistical maxims for an astrophysics journal, at the invitation of its editor. Writing as a committee is never easy, and after many rounds of Zooming, we settled on eight maxims, all variations of Box's theme, "All models are wrong, but some are more useful," leading to eight sections of the article:

- 1. Data Collection: All data have stories behind them, but some stories are mis-told.
- 2. Processing: All data are messy, but some are more easily cleaned.
- 3. *Modeling:* All models are a simplification, but some are more justified.
- *4. Assumptions: All assumptions are fallible, but some are more credible.*
- 5. *Methods:* All methods have their purpose, but some are more versatile.
- 6. Checking: All models require assumptions, but some assumptions are more easily checked.
- 7. Computation: All computations are vulnerable to error, but some are more resilient.
- 8. Interpretation: All results are subject to interpretation, but some interpretations are less contrived.

Subconsciously suspecting that the list is too clever for its own sake, I assigned the draft to the students of "The Art and Practice of Teaching and Communicating Statistics", a required PhD course that I have co-taught for almost two decades, where we engage students in many forms of communications, such as leading TA (teaching assistant) sessions, holding office hours, presenting research talks, and writing and responding to review reports. With all author-identifying information removed (including some references to my own work), the students were asked to provide a critical and constructive review report, as well as an editorial recommendation to the journal editor. I received four insightful reports, with a collective recommendation of rejection and resubmission. A principal reason for rejection was that the list of eight defeats its very purpose: there were too many boxes to deliver the intended content as pithily and pungently as Box did. I never enjoyed a happier moment of rejection: my suspicion was confirmed, and my students' judgments were clearly sounder than mine—after all, I played a non-trivial role in coming up with the list of eight.

Of course, I'd be even happier if ChatGPT could provide an IKEA style of packaging, hence another terse instruction to ChatGPT to pack the eight into four or five. Instantly, it delivers:

"Data: Data have stories and can be messy, but some are more reliable than others. Modeling: Models are simplifications, and assumptions are fallible, but some are more justifiable and credible.

**Methods:** Methods have purposes and can be versatile, but some are more appropriate for certain situations.

**Results:** Results are subject to interpretation and errors, but some interpretations are less contrived and more accurate."

Not perfect, since it missed point 7, which is critical but often overlooked. However, the rhythmic line on modeling inspired me: why not create a poem for the list? Though like many self-conscious teenagers (living in a faraway land and time) I poured my love into self-pleasing poems, my midlife crisis is still too much in its early stage to need Calliope's soothing voice. ChatGPT, on the other hand, wrote a birthday poem that brought tears to the eyes of a colleague's mother, even though he told her in the subject line that the poem was from an artificial heart. Asking ChatGPT to provide a summative poem then is a no-brainer.

Speaking of no brain, however, I must digress. One great use of ChatGPT for me is to check on my Chinglish (though I hope the volume of Chinglish I supply is insufficient for ChatGPT to populate it). When I asked ChatGPT if the penultimate sentence of the preceding paragraph was in proper English, it changed my original phrase "teared" to "brought tears to." That of course is fully expected. Unexpectedly, however, it also commented, "The use of the phrase 'artificial heart' adds a unique and interesting touch to the sentence, suggesting that the poem may have been especially heartfelt or emotional despite being written by an artificial intelligence."

Now who can distinguish that from human intelligence? It understood-for all practical purposes and intentions of the term-my pun. More importantly, it was an unprompted comment, which generates an impression of a being capable of spontaneous reflections to emotional stimuli. Most critically, especially if you think I'm being overly sentimental (you are not wrong) it is the sentiment that the ChatGPT's comment generated for me, a human being. It is a pristine appreciation of my love for pun untainted by any motivation, real or perceived (on my part), unless I let my cynicism take over: ChatGPT is using me as a guinea pig. But either way, I was, or would be, engaged emotionally by a machine.

And *that* is something that is worth chatting about. But this XL-Files already is XL in size. I therefore hope you are sufficiently intrigued to indulge my tease, and wait for the poem composed by ChatGPT, until the next XL-Files: **Tenure by GPT-***n*: **Make it or Fake it**.

# **OBITUARY: Donald Ylvisaker** 1933–2022

Professor Don Ylvisaker passed away peacefully on March 20, 2022. Don played a pivotal role in establishing the UCLA Department of Statistics. As head of the Division of Statistics within the UCLA Mathematics Department, he was instrumental in shaping the personality of the emerging department as one focused on not exclusively on theory, but also on teaching, consulting, and computation. Don was a Fellow of the IMS and the ASA.

Don was born in Minneapolis in 1933, and received a BA in Mathematics and Economics from Concordia College, an MA in Mathematics from the University of Nebraska, and a PhD in Statistics from Stanford. His dissertation was On Time Series Analysis and Reproducing Kernel Hilbert Spaces, with advisor Emanuel Parzen. He joined UCLA in 1968. Although he retired in 1996, he was extremely active in the department that he helped found, and, for the next ten years, rarely missed a faculty meeting or an opportunity to have coffee with his colleagues and discuss statistics, campus politics, or poker. He was known as a great teller of stories, which were often about courtroom statistical arguments, lottery oddities, games, and professional magicians.

Don's research placed him in the center of some of the most important issues of our time, including the use of adjustments in the US Census, counting homeless people, and, indirectly at least, the O.J. Simpson trial. For many years he was a statistical consultant to the California Lottery, and advised the Lottery on a great number of issues, including ensuring random outcomes for some of their more elaborate games.

Don began his career as a mathematical statistician. At a time when computer resources were scarce and rudimentary, he specialized in stochastic processes for modeling complex phenomena. While most research in statistics deals with noisy data gathered by either observational studies or by design of experiment, a natural connection of his expertise to the complexity study in computer science gradually emerged. Don became one among the few visionary statisticians to focus on error-free data generated by computer models.

In 1984, Don presented an IMS Special Invited Lecture and published a paper entitled "Prediction and Design" (Annals of Statistics 1987). This is a subject in which he had long been interested and had made substantial contributions. The paper provided a broad framework, G-MAP (Gaussian, Markov Associated Process), for connecting a wide range of problems concerning finite observation of a stochastic process, thus presenting a fresh perspective to bridge Gaussian Markov random field, Kriging in geostatistics, infinite-dimensional estimation under reproducing kernel Hilbert space, model robust design, time series sampling, etc. His papers on design and analysis of computer experiments are among the most highly cited in the area.

Don's writing is elegant and precise, warm, and candid. A sense of the personality of the legendary Don Ylvisaker was captured in a transcribed record of a lively dialogue with his long-term collaborator Jerry Sacks in "After 50+ Years in Statistics, An Exchange" (*Statistical Science* 2012). This dialogue portrays how Don's research transformed from mathematics-driven statistics to data-driven statistics in response to the rapidly evolving world and several critical moments, including the unexpected loss of Jack Kiefer, founder of optimal design.

The current UCLA Department of



Nils Donald "Don" Ylvisaker

Statistics was created on the foundation of Don's view of statistics. Since 1986, the Division of Statistics in the Department of Mathematics had some autonomy in constructing its undergraduate and graduate courses, but not many resources were available for a graduate program and for statistics research. It is fair to say that Mathematics saw the importance of statistics service teaching, but it was more reluctant to prioritize statistics research, especially the data and computing centered statistics that Don envisioned. Thus, hiring statisticians who could help the Division grow grew increasingly problematic. Don became convinced that statistics at UCLA could only flourish outside the Department of Mathematics.

In 1986 the Dean of Social Sciences proposed a program to overhaul graduate teaching and research in Quantitative Methods in the Social Sciences. He also created a campus-wide committee, which included Don from the Division of Statistics, to advise in the construction of the program. Don saw an opportunity to promote a far-reaching reorganization of statistics. He made sure the persons hired into the social sciences program were data- and computing-oriented, and that Continued from previous page

the ultimate goal of the reorganization was a standalone Department of Statistics. In order to grow UCLA Statistics, he also realized it was necessary to build on the foundation of undergraduate statistics teaching. Thus, a great deal of the initial effort of the interdisciplinary group that grew out of the social science initiative focused on modernizing lower-division statistics teaching. Don was instrumental in obtaining large NSF grants to reorganize and computerize statistics teaching and research. Especially after the 1994 retirement wave in the Division of Statistics, a number of lecturers were hired into semi-permanent positions and into full participating membership in the statistics group. After more than 25 years, most of them are still here. The data emphasis in Don's philosophy was evident in the establishment of a Statistical Consulting Center, in searching contacts with outside

companies and organizations in the Los Angeles area, and in the initial moves to create a professional Master's degree (all of this in an up-to-then non-existing department). It was also a natural outcome of Don's view of statistics that the initial members of the new Department have degrees in a variety of disciplines, not just in statistics or mathematics. Data are everywhere. It goes without saying that subsequent developments have shown that Don's view anticipated world-wide evolution in the discipline of statistics, even to the extent that some statistics departments now want to have "data science" in their name.

Don's influence and personality have had a deep and lasting effect on the Department. He was a strong advocate and supporter of DataFest, and served, without fail, as a judge for the annual event, even after moving to Santa Barbara. In honor of his commitment and support, the Best Insight Award has been renamed the Don Ylvisaker Best Insight Award. Although Don's retirement was 26 years ago, he kept close tabs on department news, communicated frequently with his colleagues, and felt to many as if he was still part of the department. His loss is deeply felt and his contributions long-lasting.

To keep Don's memory alive in the department, the department will be giving an annual award, the Donald Ylvisaker Award for the Best Practice of Statistics, to the graduate student whose practice of statistics best reflects Don's sense that the truth of Statistics is best expressed in real problems.

Written by Jan de Leeuw, Rob Gould, and Ker-Chau Li, Department of Statistics, University of California–Los Angeles



# Recent papers: Two open-access journals

# **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.

The Editor-in-Chief is Mikhail Lifshits, St.Petersburg State University.

Read it at https://projecteuclid.org/journals/probability-surveys/current

#### Volume 20, 2023 (to date)

Tensor- and spinor-valued random fields with applications to continuum physics and cosmology	ANATOLIY MALYARENKO, MARTIN OSTOJA-STARZEWSKI; 1-86
On homogeneous and oscillating random walks on the integers.	JULIEN BRÉMONT; 87-112
The method of stochastic characteristics for linear second-order hypoelliptic equations.	JURAJ FÖLDES, DAVID P. HERZOG; 113-169
A review of exact results for fluctuation formulas in random matrix theory.	PETER J. FORRESTER; 170-225
Probabilistic representations of fragmentation equations	MADALINA DEACONU, ANTOINE LEJAY; 226-290
Proof methods in random matrix theory.	MICHAEL FLEERMANN, WERNER KIRSCH; 291–381

# Stochastic Systems

*Stochastic Systems* is the flagship journal of the INFORMS Applied Probability Society. The journal publishes high-quality papers that substantively contribute to the modeling, analysis, and control of stochastic systems. A paper's contribution may lie in the formulation of new mathematical models, the development of new mathematical or computational methods, the innovative application of existing methods, or in the opening of new application domains.

Relative to application-focused journals, *Stochastic Systems* concentrates on on how applied probability plays a significant, and not just supporting role in this field. Relative to other applied probability outlets, *Stochastic Systems* focuses exclusively on operations research content.

Stochastic Systems provides open access to all of its content on the principle that making research freely available to the public supports a greater global exchange of knowledge. Articles are published under the Creative Commons CC-BY: Attribution license. Authors retain copyright and reuse rights as defined at https://creativecommons.org/licenses/by/4.0/legalcode.

The Editor-in-Chief is Shane Henderson, Cornell University.

Read it at https://pubsonline.informs.org/toc/stsy/current

#### Volume 13, Issue 1, March 2023

The Join-the-Shortest-Queue System in the Halfin-Whitt Regime: Rates of Convergence to the	Diffusion Limit
Generalized Max-Weight Policies in Stochastic Matching	. MATTHIEU JONCKHEERE, PASCAL MOYAL, CLAUDIA RAMÍREZ, NAHUEL SOPRANO-LOTO; 40–58
Convergence Rates of Epsilon-Greedy Global Optimization Under Radial Basis Function Interpol	ation
Large-Time Behavior of Finite-State Mean-Field Systems With Multiclasses	
Asymptotically Optimal Inventory Control for Assemble-to-Order Systems	

# **IMS Lecture Previews**

### **Bin Yu: Abraham Wald Lectures**

Bin Yu is Chancellor's Distinguished Professor and Class of 1936 Second Chair in Statistics, EECS, and Computational Biology at UC Berkeley. Her research focuses on statistical machine learning practice and theory and interdisciplinary data problems in neuroscience, genomics, and precision medicine. She and her team developed in context iterative random forests (iRF), hierarchical shrinkage (HS) for decision trees, Fast Interpretable Greedy-Tree Sums (FIGS), stability-driven NMF (staNMF), and adaptive wavelet distillation (AWD) from deep learning models. She is a member of the National Academy of Sciences and American Academy of Arts and Sciences. She was a Guggenheim Fellow, Tukey Memorial Lecturer of the Bernoulli Society, IMS Rietz Lecturer, and COPSS E. L. Scott Awardee. She holds an Honorary Doctorate from The University of Lausanne. She served on the inaugural scientific advisory board of the UK Turing Institute of Data Science and Al and is currently on the editorial board of *PNAS*.

These two IMS Wald Lectures will be given at the **Joint Statistical Meetings** in Toronto, August 5–10, 2023.

#### Wald Lecture 1:

#### Seeking Boolean interactions in biomedicine and proofs

Thresholding or Boolean behaviors of biomolecules underlie many biological processes. Decision-trees capture such behaviors and tree-based methods such random forests have been shown to succeed in predictive tasks in genomics and medicine.

In this talk, we use UK Biobank data and a stable version of the random forests, iterative random forests (iRF), to recommend gene and gene–gene interactions that have data evidence for driving a heart disease, Hypertrophic Cardiomyopathy (HCM).

Four out of the five recommendations are shown to be causal for HCM in follow-up gene-silencing experiments. This and other empirical successes of iRF motivate a theoretical investigation of its tractable version under a new local sparse and spiky (LSS) model where the regression function is a linear combination of Boolean interactions of features. The tractable version of iRF is shown to be model selection consistent under this new model, under the conditions of feature independence and non-overlap of interactions.

#### Wald Lecture 2: Sparse dictionary learning and deep learni

#### Sparse dictionary learning and deep learning in practice and theory

Sparse dictionary learning has a long history and produces wavelet-like filters when fed with natural image patches, corresponding to the VI primary visual cortex of the human brain. Wavelets as local Fourier Transforms are interpretable in physical sciences and beyond. In this talk, we will first describe adaptive wavelet distillation (AWD) to turn black-box deep learning models interpretable in cosmology and cellular biology problems while improving predictive performance. Then we present theoretical results that, under simple sparse dictionary models, gradient descent in auto-encoder fitting converges to one point on a manifold of global minima, and which minimum depends on the batch size. In particular, we show that when using a small batch-size as in stochastic gradient descent (SGD) a qualitatively different type of "feature selection" occurs.

[Editor's note: Bin Yu will also be delivering the COPSS Distinguished Achievement Award Lecture at JSM this year: see the article on page 4.]





### Wing Hung Wong: Grace Wahba Lecture

Professor Wing Hung Wong serves on the faculty of Stanford University where he is currently Professor of Statistics, Professor of Biomedical Data Science, and Stephen R. Pierce Family Goldman Sachs Professorship in Science & Human Health. Before joining Stanford in 2004, he held teaching positions at the University of Chicago, The Chinese University of Hong Kong, UCLA and Harvard University. He chaired the Stanford Department of Statistics from 2009 to 2012. Professor Wong was the winner of the COPSS Presidents' Award (1993) and the COPSS Distinguished Achievement Award (2021). He was elected to the US National Academy of Sciences in 2009 and the Academia Sinica in 2010. He was a founding member of the Hong Kong Academy of Sciences in 2015.

His past research contributions include:

 mathematical statistics, where he clarified the large sample properties, in general parameter spaces, of likelihood functions and of sieve maximum likelihood estimates;



- 2. Bayesian statistics, where he introduced sampling-based algorithms into Bayesian computational inference; and
- 3. computational biology, where he developed tools for the analysis of microarrays and sequencing data and applied them to study gene regulatory systems.

Technologies from his group have led to the formation of several companies in the space of genomics data analysis and genomic medicine.

His current research interests include the use of deep neural networks in mainstream statistical areas such as density estimation, clustering, Bayesian analysis and causal inference; and the development of novel genomics technologies based on semiconductors. In the latter, his group recently created a platform for 2D control of droplets that will enable complex cell biology experiments to be performed independently but in a massively parallel and programmable manner. He believes that the design and analysis of such experiments will present new opportunities for statistical research.

This will be the second IMS Grace Wahba Lecture (the first was given at the IMS annual meeting in London last year by Michael Jordan). It will be given at the **Joint Statistical Meetings** in Toronto, August 5–10, 2023.

#### Causal inference by encoding generative modeling

We consider the problem of inferring the causal effect of an exposure variable X on an outcome variable Y. Besides X and Y, a high-dimensional covariate V is also measured. It is assumed that confounding variables that may cause bias in the desired causal inference are low-dimensional features of V. Under this assumption, we propose an encoding generative modeling (EGM) approach for the estimation of the average dose response function, a function that captures, in an average sense, the dependency of Y on X when confounders were hold fixed.

We show that EGM provides a framework for us to develop deep learning-based estimates for the structural equations that describe the causal relations among the variables. We will present numerical and theoretical evidence to demonstrate the effectiveness of our approach.



### **Massimiliano Gubinelli: Medallion Lecture**

Massimiliano Gubinelli received a PhD in theoretical physics in 2003 from the University of Pisa. He held positions at the Universities of Pisa, Paris XI, Paris Dauphine and has been Hausdorff Chair at the University of Bonn. He is currently the Wallis professor of Mathematics at the University of Oxford. His interests lie at the crossroads of analysis, probability and mathematical physics. In particular he contributed to develop novel ideas in the analysis of non-linear stochastic ordinary and partial differential equations describing various physical phenomena linked to universality in statistical mechanics and to the Euclidean approach to quantum field theories. He delivered an ICM invited lecture in 2019 and the Lévy lecture at SPA 2019 and at the IMS World Congress in 2020.

This Medallion lecture will be delivered at the **43rd Conference on Stochastic Processes and their Applications (SPA)** in Lisbon, Portugal, July 24–28, 2023. Meeting details: https://www.spa2023.org/

#### Stochastic analysis of rough random fields

In recent years there have been steady advances in our understanding of a broad class of non-Gaussian distributional random fields that can be described via local partial differential equations involving universal source of randomness like a white noise. This line of research extends the fundamentals ideas of Ito beyond the case of diffusion processes and involve a deep interplay between analysis, algebra, probability and eventually geometry. The applications go from the description of the macroscopic universal fluctuations in weakly nonlinear microscopic random systems to the construction of Euclidean quantum field theories below their critical dimension. More generally, these ideas suggests new points of view in the analysis of such random fields via a natural coupling to some Gaussian source of randomness. The partial differential equations involved are usually ill-posed from a classical point of view and in order to tackle them in a fully non-perturbative way one has to develop a fine understanding on how the fluctuations on a wide range of scales are coupled by the



non-linearities and how they influence the large scale behavior of solutions. Resolving these challenges also relies on the deep ideas of renormalization theory originating from physics, providing at the same time a framework where to study and implement those ideas in a fully rigorous non-perturbative way.

In this talk I will review some of the progress in these topics and highlight the many open problems and opportunities for new mathematics.

### **IMS New Researchers Conference**

The IMS NRC 2023 will be held at the University of Toronto from August 2–5, 2023, in its usual spot right before JSM. The organizers have several panels planned, from publishing to funding, mentoring, promotion, and so on. We seek to provide participants with a



broad array of topics covering areas where a young researcher may face potential challenges, and hold a discussion surrounding these, all the while providing opportunities to ask questions of senior researchers on these topics. We also hope young folks form connections with each other through the NRC that can last as a solid long-term network in their careers.

Invited speakers and panelists: Peter Bühlmann, ETH Zurich; Jianqing Fan, Princeton University; Paul Fearnhead, Lancaster University; Chao Gao, University of Chicago (IMS Tweedie New Researcher Award Recipient 2021); Yulia Gel, The University of Texas at Dallas; Xuming He, University of Michigan; Jonathan Niles-Weed, New York University (IMS Tweedie New Researcher Award Recipient 2023); Annie Qu, University of California Irvine; Nancy Reid, University of Toronto; Dylan Small, University of Pennsylvania; Michael Stein, Rutgers University; Lan Wang, University of Miami; and Leilei Zeng, University of Waterloo.

The application deadline for the conference is April 24, 2023. We urge all young researchers within the eligibility criteria (received a PhD in or after 2016, or expects to receive one by the end of 2023) to apply. More details can be found on the conference website: https://sites.google.com/site/linbowangpku/nrc-2023



Student Puzzle Editor Anirban DasGupta poses another two problems, and says, "For the first time in our problem corner, we will consider a problem that is hard to solve analytically, and we will ask you to write an algorithm to solve the problem as accurately as you can by simulation. The second one is a traditional problem." Send us your solution, to either or both. Puzzle 44.1.

As you were making coffee in your kitchen on waking up, you noticed that a huge centipede has curled itself up on your kitchen floor and its head and tail are joined together. Petrified by it, you took a coffee mug with a circular mouth and tried to cover the centipede. Assume that the centipede is eight inches long. A coffee mug of what internal diameter can cover the centipede Deadline: May 7, 2023 completely? Give the smallest diameter suggested by your algorithm. Send your algorithm with your answer.

#### **Puzzle 44.2**.

Suppose  $X_{1'}X_{2'}$  ... are i.i.d. standard Cauchy. For  $n \ge 2$ , denote by  $T_n$  the proportion of pairs  $i < j \le n$  such that  $X_i + X_i \ge 0$ .

> (i) Does  $T_n$  converge in probability to some constant c? If so, to what? (ii) Does there exist a sequence  $\{a_n\}$  and a nontrivial distribution G such that  $a_n(T_n - c) \stackrel{\mathcal{L}}{\Rightarrow} G$ ? If so, identify  $\{a_n\}$  and G.

#### 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 to either or both of these puzzles, 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 43

Thank you to Alberto Bordino (University of Warwick), Soham Bonnerjee (University of Chicago), and Abhinandan Dalal (University of Pennsylvania) for sending their solutions. Anirban DasGupta explains:

#### Puzzle 43.1 (see the reminder, right)

If you take a large value of n, and plot the pairs  $(k, \overline{X}_k)$  for k = 1, 2, ..., n for various sets of simulated standard Cauchy variables, you should see that the paths oscillate and that they revisit neighborhoods of a given level b within the range of the path. We cannot show many such plots here due to reasons of space. We can frame this rigorously. We can show by a simple calculation that given any real b, and  $0 < \epsilon \le 1$ ,  $\sum_{n=1}^{n} P(|\bar{X}_n - b| \le \epsilon) = \infty$ . All we need to use is that for each  $n, \bar{X}_n$  is distributed as a standard Cauchy. This is sufficient to guarantee that with probability one, b is an accumulation point for the sequence  $\{\overline{X}_n\}$ . If we choose b = 0, then we can conclude that with probability one,  $\limsup |\overline{X}_n - e^n| = \infty$ . You will notice that  $e^n$  is nothing special; we can choose basically any sequence going to  $\infty$ .

#### Puzzle 43.2

We will use the standard result that a unique Bayes rule with a finite Bayes risk is admissible. The uniqueness of a Bayes rule follows from strict convexity of  $L(\theta, a) = (\theta - a)^2$  in *a*; here  $\theta$  stands for *N*. Consider now the proper prior with mass function  $\frac{c}{N^4}$ , N = 1, 2, ... Denoting the digamma function by  $\psi(x)$ , the posterior mean E(N|X) equals

$$\delta_B(X) = \frac{-4\psi^{(3)}(X)}{\psi^{(4)}(X)}.$$

 $\delta_B(X)$  has a quadratically bounded risk function and so a finite Bayes risk for the prior  $\frac{c}{N^4}$ . Thus  $\frac{-4\psi^{(3)}(X)}{\psi^{(4)}(X)}$ is one admissible estimator of N.

Puzzle 43.1: (i) Simulate 1,000 standard Cauchy variables, and plot  $\bar{X}_{p}$ against *n* for n = 1, 2, ...,1000, where  $\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$ . (ii) Briefly discuss what do you see in this plot that is interesting.

(iii) What is your guess for  $\limsup_{n\to\infty} |\bar{X}_n - e^n|,$ and justify why.

Puzzle 43.2: Suppose X has a discrete uniform distribution on the set {1, 2, ..., N}, where  $N \ge 1$  is an unknown integer-valued parameter. Construct an explicit admissible estimator of N under squared error loss and provide a proof of your estimator's admissibility.

# **2023 IMS Elections: The candidates**

It's time to think about who you would like to represent you on the IMS Council, and vote in the IMS elections. This year, the nominee for President-Elect is Tony Cai. There are 10 candidates standing for the five available places on Council: Sourav Chatterjee, Gábor Lugosi, Marina Meila, Andrew B. Nobel, Roberto Imbuzeiro Oliveira, Jasjeet S. Sekhon, Weijie Su, Caroline Uhler, Huixia Judy Wang, and Linda Zhao. Read about them here or on the IMS website, at **https://imstat.org/elections/candidates/. Voting closes June 24, 2023.** 

# **President-Elect Nominee**

### <u>Tony Cai</u>

Daniel H. Silberberg Professor of Statistics and Data Science, The Wharton School; Professor, Applied Mathematics and Computational Science Graduate Group; and Associate Scholar, Department of Biostatistics, Epidemiology, and Bioinformatics, Perelman School of Medicine; Department of Statistics and Data Science, University of Pennsylvania



#### http://www-stat.wharton.upenn.edu/~tcai/

#### **Education**

BS Mathematics, 1986, Hangzhou University, China MS Applied Mathematics, 1989, Shanghai Jiao Tong University, China MA Mathematics, 1992, State University of New York at Buffalo, USA PhD Statistics, 1996, Cornell University, USA

#### **Research Interests**

Statistical machine learning High-dimensional statistics Large-scale inference Statistical decision theory Functional data analysis Nonparametric function estimation Applications to genomics and financial engineering

#### **Previous Service to the Profession**

Co-Editor, Annals of Statistics, 2010–12 Associate Editor, JRSSB (2014–18), JASA(2005–2010), Annals of Statistics (2004–09), Statistica Sinica (2005–11), Statistics Surveys (2006-09)Board of Advisors, Institute for Mathematical and Statistical Innovation, 2020-2024 IMS Council, 2020–23 COPSS Distinguished Achievement Award & Lecture Committee, 2020-23 Program Co-Chair, IMS Annual Meeting 2017 and 2007 Bernoulli Society Publications Committee, 2011–14 COPSS Presidents' Award Committee, 2009–12, 2014; Chair, 2012 IMS Committee on Special Lectures 2008-12 IMS Committee on Nominations 2008-09 Co-Chair, Scientific Committee, IMS-China International Conference, Weihai, China, 2009 IMS Committee to Select Editors 2000-03

#### **Brief Statement**

It is an honor to be nominated for the IMS Presidency. The importance of statistics and data science cannot be overstated in today's data-driven world. As a leading scholarly society in statistics, IMS plays a crucial role in advancing the discipline through its numerous initiatives and activities. My vision is to continue the outreach efforts to industry and diverse geographic regions, expand the membership base to foster a more inclusive and collaborative community, enhance fundraising efforts to support new researchers, and expand the activities in data science and other areas related to data-driven AI.

# **Council Nominees** There are 10 candidates, for five places on IMS Council.

### <u>Sourav Chatterjee</u>

Professor of Mathematics and Statistics, Department of Mathematics and Department of Statistics, Stanford University



#### https://souravchatterjee.su.domains/

#### Education

- Bachelor of Statistics, 2000, Indian Statistical Institute, Kolkata, India
- Master of Statistics, 2002, Indian Statistical Institute, Kolkata, India
- PhD in Statistics, 2005, Stanford University, USA

#### **Research Interests**

Probability theory Mathematical statistics Mathematical physics

#### **Previous Service to the Profession**

- Member of the Scientific Research Board of the American Institute of Mathematics, 2022–2025 Member of the Scientific Advisory Committee of the Simons Laufer Mathematical Sciences Institute (formerly Mathematical
- Sciences Research Institute), 2022–2026
- Associate editor for the *Proceedings of the London Mathematical Society*, 2023 onwards
- Associate editor for the *Annals of Applied Probability*, 2022 onwards Associate editor for *Communications in Mathematical Physics*, 2019 onwards
- Associate editor for Sankhya, Series A, 2012–2015
- Associate editor for *Probability Theory and Related Fields*, 2011–2015
- Associate editor for the Annals of Probability, 2009-2014
- Associate editor for the *Annales de l'Institut Henri Poincare (B)*, 2008–2013
- Organized many conferences and workshops and served on IMS committees

#### **Brief Statement**

It is a great privilege and honor to be nominated for election to the IMS council. The IMS plays a very important role in supporting mathematical statistics and probability throughout the world, and I will do my best to help advance its causes. I am particularly interested in maintaining the high quality of the IMS journals, and promoting IMS initiatives for increasing diversity and outreach to underrepresented communities.

### <u>Gábor Lugosi</u>

ICREA Research Professor, Department of Economics and Business, Pompeu Fabra University, Barcelona, Spain



#### http://www.econ.upf.edu/~lugosi/

#### Education

PhD in Electrical Engineering, 1991, Hungarian Academy of Sciences, HungaryMS in Electrical Engineering, 1987, Technical University of Budapest, Hungary

#### **Research Interests**

Theory of machine learning Combinatorial statistics Inequalities in probability Random graphs and random structures Information theory

#### **Previous Service to the Profession**

Editor and co-founder of *Mathematical Statistics and Learning* (2017–) Associate editor of *Annals of Applied Probability* (2016–2021) Associate editor of *Probability Theory and Related Fields* (2015–) Action editor of *Journal of Machine Learning Research* (2005–) Member of the editorial board of *Machine Learning Journal* (2006–2020) Associate editor of *TEST* (2002–) Associate editor of *ESAIM: Probability and Statistics* (2005–) Associate editor of *IEEE Transactions on Information Theory* (1999–2002)

Associate editor of *Statistics & Decisions* (2002–2011) Associate editor of *Scandinavian Journal of Statistic*s (2010–2015)

Member of the editorial board of *Foundations and Trends in* Machine Learning (2007– )

#### **Brief Statement**

It is a great honor to be nominated to the IMS Council. Both probability and statistics are booming largely thanks to the new challenges coming from data science, machine learning, bioinformatics, and other fields. IMS faces the challenge of keeping a leading role in this increasingly competitive environment and promoting rigorous thinking. I would be honored to help maintain this leadership.

### <u>Marina Meila</u>

Professor, Department of Statistics, University of Washington

https://sites.stat.washington.edu/mmp/

#### **Education**

- MS in Electrical Engineering, 1985, Polytechnic Institute of Bucharest, Romania
- PhD in Electrical Engineering and Computer Science, 1999, Massachusetts Institute of Technology, USA

#### **Research Interests**

#### Geometric data analysis

Ranked data and preferences, with applications to peer review Statistics and data science for molecules and materials Theoretical foundations of clustering, robust identifiability and distribution free guarantees Networks, graphs, graphical models

Scalable accurate statistical inference

#### **Previous Service to the Profession**

- Institute for Pure and Applied Mathematics (IPAM), Scientific Advisory Board, since 2023
- Chair, Association for Uncertainty in AI (AUAI) since 2022
- Program co-chair for: the International Conference on Machine Learning, ICML 2021, Uncertainty in AI Conference, UAI 2013, Conference for AI and Statistics, AISTATS 2007; past AE for *JMLR, IEEE TPAMI, JASA-TAS*
- Organizer of several Long Programs: IPAM 2016, IPAM 2019, and IMSI "Data driven material informatics: statistical methods and mathematical analysis" upcoming March 2024
- Workshop leader for "Fostering a research culture" at Kinnaird College for Women, Pakistan, 2012.

#### **Brief Statement**

These are exciting times of growth and transformation in our field. IMS has a vital role in bringing high quality statistics to bear on day-to-day data practices. I am honored to be nominated for the IMS council and I hope that in this role I can support our community in expanding the scope of statistics to respond to new types of data challenges, and our field's societal impacts.

Continues on page 22



# Council Nominees continued

### <u>Andrew B. Nobel</u>

Paul Ziff Distinguished Professor, Department of Statistics and Operations Research, University of North Carolina, Chapel Hill



https://nobel.web.unc.edu/

#### **Education**

PhD 1992 Stanford University MS 1988 Stanford University BS 1985 Cornell University

#### **Research Interests**

Machine Learning Statistical Genomics Learning from Dynamical Systems Analysis of Networks

#### **Previous Service to the Profession**

Associate Editor: IEEE Transactions on Information Theory, Electronic Journal of Statistics, Annals of Statistics, Journal of the Royal Statistical Society, Series B.

#### Service to IMS:

Program Secretary/Executive Committee, 2003–2006 Equality and Diversity (2020–22) Publications (2015–18) Nominations (2008–09) Special Lectures (2003–09) Selection of Administrative Officers (2011–13)

#### **Brief Statement**

Now, as ever, times are changing. For the IMS the emergence of data science as a stand-alone discipline, the ascendancy of artificial intelligence, and fundamental shifts in the fields of statistics and probability provide challenges and opportunities. The IMS needs to respond to the changing demographics and research interests of its existing members, while targeting new members, especially younger researchers, whose interests may not fall neatly into the categories around which statistics and probability have historically been organized. I am honored to be nominated for the IMS Council, and am eager to play a role in addressing these issues.

### <u>Roberto Imbuzeiro Oliveira</u>

Professor, Instituto de Matemática Pura e Aplicada (IMPA)



https://sites.google.com/view/roboliv

#### Education

BS Mathematics, Pontifícia Universidade Católica do Rio de Janeiro (PUC-Rio), 1999

MS Mathematics, Instituto de Matemática Pura e Aplicada (IMPA), 2000

PhD Mathematics, Courant Institute of Mathematical Sciences, New York University, 2004

#### **Research Interests**

Probability and connections with Statistics, Optimization and other topics

Concentration inequalities and high dimensional probability Markov processes and interacting particle systems Random graphs and other discrete objects

#### **Previous Service to the Profession**

#### Current:

Associate editor for ALEA (Latin American Journal of Probability and Mathematical Statistics) Associate editor for Operations Research Senior Program Committee member for the Conference on Learning Theory (COLT) in 2022 and 2023 Member of the board of directors of "Sociedade Brasileira de Matemática" (Brazilian Mathematical Society-SBM) Secretary of the "Unión Matemática de Latinoamerica y Caribe" (Mathematical Union for Latin America and the Caribbean-UMALCA) Deputy editor-in-chief of the Fronteiras da Matemática collection of "Sociedade Brasileira de Matemática" (Brazilian Math Society) Past: Member of the local organizing committee for the International Congress of Mathematicians in Rio de Janeiro (ICM 2018) Member of the IMS Nominations Committee, 2020–2021 Editor in Chief of ALEA, 2019–2021

Deputy Head of Evaluation of Graduate Programs in Mathematics and Statistics for Brazilian funding agency (CAPES), 2018–2021 Co-organizer of the Thematic Program on Graph Theory at IMPA, 2018

Co-organizer of numerous other scientific events

#### **Brief Statement**

I am honored to be nominated for the IMS Council election. Over the years, the IMS has promoted statistical research and education of the highest quality. These goals remain absolutely central, but our day and age bring their own pressing challenges: making Statistics and Data Science more global, more broadly understood, and more aware of its ethical and societal aspects. The Council is an ideal forum for proposing initiatives to address these issues. In particular, it should strive to support researchers from all groups and backgrounds; promote educational and outreach activities; and encourage interdisciplinary interactions.

### <u> Jasjeet S. Sekhon</u>

Eugene Meyer Professor, Statistics and Data Science, Yale University



#### https://www.jsekhon.com/

#### Education

PhD Cornell University, 1999 BA(Hons) University of British Columbia, 1993

#### **Research Interests**

Causal Inference Machine Learning Applied Statistics Data Science Industrial Applications

#### **Previous Service to the Profession**

Founding co-editor of *Journal of Causal Inference* with Judea Pearl, Maya Petersen, and Mark van der Laan
Associate Editor: *Journal of the American Statistical Association T&M* (2014–2017)
Associate Editor: *Statistics and Public Policy* (2014–2018)

#### **Brief Statement**

Being a member of the IMS is a privilege, and it is a great honor to be nominated for the IMS Council. It is an exciting and dynamic time for the field of statistics, and the IMS is critical in providing opportunities for members of our field to grow as researchers and educators, for the development of interdisciplinary research between statistics and related fields, and for supporting students and new researchers in statistics. I have worked in many different fields of statistics in both academic and industrial settings, which brings a differentiating perspective that may be valuable for the IMS community to further embrace and integrate with the data science revolution that is impacting almost all fields of human knowledge.

### <u>Weijie Su</u>

Associate Professor of Statistics and Data Science, Department of Statistics and Data Science, Wharton School, University of Pennsylvania



http://www-stat.wharton.upenn.edu/~suw/

#### Education

BS in Mathematics, 2011, Peking University PhD in Statistics, 2016, Stanford University

#### **Research Interests**

High-Dimensional Statistics Privacy-Preserving Data Analysis Peer Review Deep Learning Theory Optimization

#### **Previous Service to the Profession**

Associate Editor for *Journal of American Statistical Society (JASA)* (2023–present), *NeurIPS* (2021, 2022, 2023), *ICML* (2023) Awards Committees: IMS Peter Hall Prize (2023), ICSA Student Paper Award (2021), Student Paper Award of ASA Section on Statistical Learning and Data Science (2019) Co-organizer of several workshops

# **Council Nominees** continued

#### **Brief Statement**

It is a great honor to be nominated for the council position. IMS has been successfully providing valuable opportunities for generations of statisticians and probabilists, and I am grateful to have benefitted from these resources. As we are entering a new era of exciting AI developments, I am confident that IMS has much to offer in facilitating interdisciplinary research in statistics, data science, and AI. If elected, I am committed to finding ways to contribute to IMS's mission of supporting young statisticians to thrive in these areas.

### <u>Caroline</u> Uhler

Full Professor at MIT, and Co-Director of the Eric and Wendy Schmidt Center at the Broad Institute, Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology (MIT) and Broad Institute



#### https://www.carolineuhler.com/

#### Education

PhD in Statistics, 2011, UC Berkeley

#### **Research Interests**

Causal inference Probabilistic graphical models Representation learning and generative modeling Applications to genomics

#### **Previous Service to the Profession**

IMS representative, Joint Committee on Women in the Mathematical Sciences (2017–20)

- Council Member (elected), International Statistical Institute (2021–25)
- Elected member, Scientific Advisory Board of the Gladstone Institutes (2023)
- General and Program Chair of inaugural CLeaR Conference (2022) Advisor, Science Advisory Board, Chan Zuckerberg Initiative (2022)
- Judge for the Regeneron Science Talent Search (2015–21)

- International Conference on Machine Learning (ICML) tutorial co-chair (2021)
- International Conference on Machine Learning (ICML) sponsorship co-chair (2020)
- Damon Runyon Quantitative Biology Fellowship Award Selection Committee (2020–present)
- Associate editor, SIAM Journal on Mathematics of Data Science (2020–22)
- Associate editor, SIAM Journal on Applied Algebra and Geometry (2019–22)
- Associate editor, Harvard Data Science Review (2019-22)
- Program Director, SIAM Activity Group on Algebraic Geometry (2018–20)

#### **Brief Statement**

Given the data explosion that we are witnessing in many fields, the combination of statistics and computation is becoming ever more critical. Building on my experience working at the intersection of mathematical statistics, machine learning, and genomics, I would be excited to contribute towards strengthening the links between the IMS and the broader data science community. The IMS plays a critical role in developing and nurturing the current and next generation of mathematical statisticians world-wide. I would be truly honored to serve as an IMS council member and promote a diverse and inclusive community.

### Huixia Judy Wang

Professor and Chair, Department of Statistics, The George Washington University

#### https://blogs.gwu.edu/judywang/

#### Education

PhD Statistics, 2006, University of Illinois, Urbana–Champaign, IL, USA

MS Statistics, 2002, Fudan University, Shanghai, China BS Statistics, 1999, Fudan University, Shanghai, China

#### **Research Interests**

Extreme value theory and applications



Bioinformatics and biostatistics Longitudinal data analysis Nonparametric (semiparametric) regression Quantile Regression Spatial data analysis Survival analysis

#### **Previous Service to the Profession**

#### Service to IMS:

- IMS Program Chair for 2023 Joint Statistical Meetings, 2022–23 (Chair)
- Joint Bernoulli Society/IMS Publications Management Committee, 2021–24 (Member)

IMS Committee on Nominations, 2018–19, 2020–21 (Member) Editorial services:

- Associate Editor for The Annals of Statistics, 2013-present
- Associate Editor for the *Journal of the American Statistical Association (Theory and Methods)*, 2011–present
- Associate Editor for Bernoulli, 2022-present
- Associate Editor for Stat, 2015–21
- Associate Editor for the Review Sections of the *Journal of the American Statistical Association* and *The American Statistician*, 2013–16
- Other professional services:
- Program Director, Division of Mathematical Sciences, US National Science Foundation, 2018–22
- Committee of Presidents of Statistical Societies (COPSS), 2019–21 (Secretary and Treasurer)

International Chinese Statistical Association (ICSA), Program Committee for ICSA Applied Statistics Symposium, 2021 (Co-Chair)

- ICSA, Award Committee, 2018–20 (Member), 2020–21 (Chair)
- International Indian Statistical Association (IISA), Student Paper Competition Committee for IISA International Conference on Statistics, 2017 (Member)
- ASA, Section on Nonparametric Statistics, 2014 (Secretary), 2015 (Treasurer)

ASA, Section on Nonparametric Statistics, Student Paper Competition Committee, 2015 (Member)

- Eastern North American Region (ENAR), Organizing Committee for Junior Researcher's Workshop, 2012 (Member)
- ICSA, Program Committee, 2012 (Member)
- International Conference of Robust Statistics (ICORS), Organizing Committee, 2012 (Member)

ENAR, Program Committee for ENAR Spring Meeting, 2011
(Member)
ASA, Local Organizing Committee, Women in Statistics
Conference, 2014 (Member)
ICSA, Program Chair for 2011 JSM, 2011
ASA, General Methodology Program Chair for 2010 JSM, 2010

#### **Brief Statement**

I am honored and humbled to be nominated for the IMS Council. Serving on editorial boards and publication, program, and awards committees has been an incredibly rewarding experience. I am excited to have the opportunity to continue contributing to the IMS community by upholding the high standards of the IMS journals, promoting diversity, supporting the next generation of researchers, and fostering interdisciplinary collaboration between probability, statistics, and other fields. With the rapidly changing landscape of data science and statistics, IMS plays a critical role in advancing the field. I am eager to contribute to its mission of supporting the community and creating broader impacts.

Linda Zhao

Professor of Statistics and Data Science, Department of Statistics and Data Science, The Wharton School, University of Pennsylvania



https://statistics.wharton.upenn.edu/profile/lzhao/

#### Education

PhD, 1993 from Cornell University

#### **Research Interests**

Network data Data driven decisions in Business Crowd sourcing Post model selection inference Methods in model free settings Non-parametric empirical Bayes

# **Council Nominees** continued

#### **Previous Service to the Profession**

IMS Fellows committee, 2021–
ASA Noether Awards Committee, 2021–
Fundraiser and committee member (2019) launching IMS Lawrence D. Brown PhD Student Award
CWS (Caucus for Women in Statistics), IMS Liaison, Chair of Public Relations Committee, 2016–
Co-organizer of WiDS@Penn conference, 2021– (Women in Data Science)
Organizer of DSL, 2019– (Data Science Live workshop at Penn)
ICSA various committees
ENAR 2013 IMS Program Chair, Orlando
Former Associate editors for JSPI, Statistica Sinica and Journal of Asia Business Studies

#### **Brief Statement**

As a proud lifetime member of IMS, I am deeply honored to be nominated for the IMS council. In the new era of data science, I strongly believe that IMS should lead the way in solving data problems in science, social science, medicine, business, and government, while providing fundamental theoretical foundations for data science. I would like to continue to contribute to strengthening and broadening IMS's role in 1) bridging statistics, data science, and related fields, and fostering interdisciplinary research, 2) creating more career development opportunities, especially for students, as well as young people, women, and underrepresented researchers, and 3) developing modern statistics and data science pedagogies for the changing landscape of our field.

Elections close June 24, 2023. Vote online at https://www.imstat.org/portal/voting/8/step1

I must remember to vote in the IMS elections...

# **New ACM–IMS Journal of Data Science**

ACM/IMS *Journal of Data Science (JDS)* is a joint journal of the Association of Computing Machinery (ACM) and the IMS, publishing high-impact research from all areas of data science, across foundations, applications and systems. The scope of the journal is multi-disciplinary and broad, spanning statistics, machine learning, computer systems, and the societal implications of data science. *JDS* accepts original papers as well as novel surveys that summarize and organize critical subject areas.

The inaugural issue of the new journal is online now. Volume 1, issue 1 of *JDS* contains papers on "Batched Neural Bandits" (Quanquan Gu, Amin Karbasi, Khashayar Khosravi, Vahab Mirrokni, and Dongruo Zhou);



"Identification and semiparametric efficiency theory of non-ignorable missing data with a shadow variable" (Wang Miao, Lan Liu, Yilin Li, Eric Tchetgen Tchetgen, and Zhi Geng); "Record Fusion via Inference and Data Augmentation" (Alireza Heidari, George Michalopoulos, Shrinu Kushagra, Ihab F. Ilyas, and Theodoros Rekatsinas); "Optimistic Rates: A Unifying Theory for Interpolation Learning and Regularization in Linear Regression" (Lijia Zhou, Frederic Koehler, Danica J. Sutherland, and Nathan Srebro); "DNBP: Differentiable Nonparametric Belief Propagation" (Anthony Opipari, Jana Pavlasek, Chao Chen, Shoutian Wang, Karthik Desingh, and Odest Chadwicke Jenkins); and "Data Management for ML-based Analytics and Beyond" (Daniel Kang, John Guibas, Peter Bailis, Tatsunori Hashimoto, Yi Sun, and Matei Zaharia).

The Editors-in-chief are Jelena Bradic, Stratos Idreos, and John Lafferty.

Read the papers, and find out how to submit your paper, at http://jds.acm.org/

# IMS meetings around the world

**JSM 2026** 

August 1-6, 2026

Boston, MA, USA

### Joint Statistical Meetings

#### 2023 Joint Statistical Meetings August 5–10, 2023 in Toronto

w https://ww2.amstat.org/meetings/jsm/2023/ The IMS Program Chair is Huixia Judy Wang, George Washington University.

.....

#### JSM dates for 2024-2026

JSM 2024 August 3–8, 2024 Portland, Oregon, USA IMS Annual Meeting @ JSM 2025 August 2–7, 2025 Nashville, TN, USA

NEW

#### The 8th Workshop on Biostatistics and Bioinformatics May 5–7, 2023 Atlanta, USA

w math.gsu.edu/yichuan/2023Workshop/ Biostatistics and Bioinformatics have been playing very important roles in scientific research fields in recent years. The goal of the 8th workshop is to stimulate research and to foster the interaction of researchers in the research areas. The keynote speaker is **Nilanjan Chatterjee**, Bloomberg Distinguished Professor of Biostatistics and Medicine at the Johns Hopkins University Bloomberg School of Public Health and Johns Hopkins School of Medicine

#### 43rd Conference on Stochastic Processes and their Applications July 24–28, 2023 in Lisbon, Portugal

w https://www.spa2023.org/ Featuring talks by Louigi Addario-Berry (Schramm lecture), Riddhipratim Basu, René Carmona (Doob lecture), Jean-Dominique Deuschel, Massimiliano Gubinelli (Medallion lecture), Martina Hofmanova, Richard Kenyon (Medallion lecture), Gesine Reinert, Makiko Sasada, Sylvia Serfaty (Medallion lecture), and Horng-Tzer Yau (Lévy lecture). Doeblin & Itô prize lecturers to be announced.



#### The Annual ASA/IMS Spring Research Conference (SRC) on Statistics in Industry and Technology May 24–26, 2023 Banff Centre, Alberta, Canada

w https://sites.google.com/view/src2023/home

A three-day summit of talks and activities, the SRC is the annual meeting of the Section on Physical and Engineering Statistics (SPES) of the American Statistical Association (ASA), co-sponsored by the IMS. Although historically emphasizing industrial statistics, design of experiments, quality and reliability, the meeting increasingly emphasizes modern methods on statistics, machine learning and high performance computing in statistical methodology, with diverse applications encompassing all areas of applied sciences. The conference is intended to stimulate interactions among statisticians, researchers in the application areas, and industrial practitioners.

### IMS New Researchers Conference

#### (23rd Meeting of New Researchers in Statistics and Probability) August 2–5, 2023

#### University of Toronto, Canada

w https://sites.google.com/site/linbowangpku/nrc-2023 The purpose of the conference is to promote interaction and networking among new researchers in statistics and probability. Anyone who has received a PhD in or after 2016, or expects to by the end of 2023, is eligible to attend, though participation is by invitation only. Participants from under-represented groups are especially encouraged to apply. The meeting will take place immediately before JSM Toronto. Apply via the website for a poster or speed-talk session.

See the list of speakers on the website, or on page 17.

### At a glance:

forthcoming IMS Annual Meeting and JSM dates

### 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

UPDATED

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



IMS Annual Meeting: TBD

JSM: Boston, MA, August 1–6, 2026

# **More IMS meetings**

# The 2023 Workshop of Statistical Network Analysis and Beyond (SNAB 2023)

#### June 15–17, 2023 (immediately before WNAR/IMS meeting) Anchorage, Alaska, USA

#### w https://www.snab2023.org/

Over the span of three days, this workshop aims to unite researchers in the field of network science and related disciplines, providing an avenue for the exchange of innovative ideas and recent findings. The workshop will encompass a wide range of topics, ranging from statistical network modeling to more extensive fields such as tensor modeling, social/economic network analysis, deep learning, and text analysis. The event will feature approximately 12 research talks each day, alongside a mixer with the poster competition on the first day and a banquet on the second day.

Contact: Please contact tianxili@virginia.edu or riczw@ rams.colostate.edu for any questions regarding the workshop.

UPDATED

#### WNAR 2023 June 18–21, 2023 Anchorage, Alaska, USA

#### w https://wnar.org/wnar2023

Scientific program chair: Audrey Hendricks, University of Colorado Denver. IMS program chair: Hua Zhou, UCLA. Local organizer: Jiaqi Huang, Alaska Department of Fish and Game. Chair of Student Award Committee: Charlotte Gard, New Mexico State University.

The 2022 WNAR/IMS meeting will be held in Anchorage, Alaska from Sunday, June 18 through Wednesday, June 21, 2023. The conference will be held at Hilton Anchorage in downtown Anchorage. It immediately follows the 2023 Workshop on Statistical Network Analysis and Beyond (SNAB2023), see the announcement above.

Registration is open (\$450 for IMS members). See the link at https://wnar.org/event-5048851/Registration

Program highlights include short courses, the WNAR President's invited presentation, invited and contributed oral sessions, student paper sessions, a welcome reception and a banquet.

Please contact the WNAR/IMS local organizer Jiaqi Huang (jiaqi.huang@alaska.gov), WNAR program chairs Audrey Hendricks and Wen (Rick) Zhou (wnarprogramchair@gmail.com), or IMS program chair Hua Zhou (huazhou@ucla.edu) for more information.

#### Joint Conference on Statistics and Data Science in China July 11–13, 2023. Peking University, Beijing, China



w https://www.stat-center.pku.edu.cn/en/Events/conference/1364567.htm This is the inaugural Joint Conference on Statistics and Data Science, initiated and sponsored by four Chinese Statistical Societies (Applied Statistics Association, Chinese Probability and Statistics Society, National Industrial Statistics Teaching and Research Association, and Chinese Business Statistics Association) and IMS-China. The conference aims to promote academic exchanges in the field of statistics and data science, cultivate a good academic environment, lead statistical innovations, and promote interfaces among academia and industry.

#### 21st INFORMS/Applied Probability Society meeting June 28–30, 2023. Nancy, France

w https://informs-aps2023.event.univ-lorraine.fr/

Featuring an IMS Medallion Lecture by Sylvie Méléard and the Marcel Neuts Lecture by Beatrice Meini; other Plenary Lectures by Frédéric Chazal, Sean Meyn and Amy R. Ward. Tutorial speakers are Paul Embrechts and Sarah Penington. Registration is open. The deadline for early-bird registration rates is May 1.

#### Statistical Foundations of Data Science and their Applications: A conference in celebration of Jianqing Fan's 60th birthday May 8–10, 2023. Princeton University, New Jersey, USA

w http://fan60.princeton.edu/

The conference will bring together collaborators and leading researchers in statistics and data science to provide an excellent forum for scientific communications and promote collaborations among researchers in statistics and data science. The program covers a wide range of topics presenting recent developments and the state of the art in a variety of modern research topics on statistics and data science as well as their applications.

#### Southeastern Probability Conference (two in 2023) May 15–16, 2023 at Duke University, USA, *and* August 14–15, 2023 at University of Virginia, Charlottesville

w https://services.math.duke.edu/~rtd/

May SEPC organized by Rick Durrett with Shankar Bhamidi (UNC), Michael Damron (Georgia Tech), and Evita Nestoridi (Stonybrook). Speakers: Amol Aggarwal (Columbia), Morgane Austern (Harvard), Emma Bailey (CUNY), Dor Elboim (Princeton, working with Allan Sly), David Harper (Georgia Tech, working with Michael Damron), Nina Holden (NYU), Chris Janjigian (Purdue), Philippe Sosoe (Cornell), and Yuxin Zhou (Northwestern, working with Auffinger). Graduate students and postdocs can apply for 10–12 grants of up to \$500 to partially support the cost of their attendance: see website, apply ASAP.

#### ICSA 2023 China Conference June 30–July 3, 2023 Chengdu, Sichuan, China

w https://maths.swjtu.edu.cn/english/ICSA\_2023\_China\_ Conference/Conference\_Introduction.htm The 2023 ICSA China Conference will be held at Chengdu, Sichuan, China from June 30 to July 3, 2023. It is co-organized by the Southwest Jiaotong University. The conference venue is Jinniu Hotel (http://www.jnhotel. com/).

The theme of this conference is "Data Science with Applications to Big Data Analysis and AI", in recognition of the big data era.

The executive and organizing committees have been working diligently to put together a strong and comprehensive program, including keynote lectures, invited sessions, poster sessions, junior researcher award session, and exciting social events. Our scientific program reflects recent challenges in statistics, business statistics, and biostatistics, which are related to the big data analysis. The conference will provide great opportunities for learning, networking and collaborations. Participants will share the thoughts and ideas with conference guests, and receive inspirations from old research ideas and develop new ones.

#### 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, Mark Holmes, Estate Khmaladze (chair), Krishanu Maulik, Spiro Penev, Masanobu Taniguchi, Lijiang Yang, and Nakahiro Yoshida—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.

### 

2024 ENAR/IMS Spring Meeting March 10–13, 2024 Baltimore, MD, USA w https://enar.org/meetings/future.cfm

#### IMS–APRM in Melbourne, Australia NEW DATES: January 4–7, 2024

w https://ims-aprm2024.com/ IMS-APRM will provide an excellent forum for scientific communications and collaborations for researchers in Asia and the Pacific Rim, and promote communications and collaborations between the researchers in this area and those from other parts of the world. The program covers a wide range of topics in statistics and probability, presenting recent developments and the state of the art in a variety of modern research topics and in applications. The conference organizers are now accepting session proposals (deadline for submission is April 16, 2023).



#### 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 **monthly** seminars that take place via Zoom on Thursdays, typically 9.30am or 1.30pm [UK time]. 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/ Thursdays, 14:00 UTC/GMT [resuming in September]. Please subscribe to the mailing list for updates about the upcoming seminars and other events: https://www.owprobability.org/mailing-list

#### WNAR2024

held in conjunction with the 2024 Graybill Conference June 9–12, 2024 Fort Collins, Colorado, USA w https://wnar.org/ Meetings The 2024 meeting of the Western North American Region of The International Biometric Society will be held joint with the 2024 Graybill Conference.

#### IMS annual meeting

Bernoulli–IMS 11th World Congress in Probability and Statistics August 12–16, 2024 Ruhr-University Bochum, Germany

w https://www.bernoulliims-worldcongress2024. org/ UPDATED The Institute of

Mathematical Statistics Annual Meeting will be held at the 11th World Congress.

# Other meetings and events around the world

#### 2023 International Indian Statistical Association (IISA) Conference June 1–4, 2023 in Colorado School of Mines, USA

#### w https://www.intindstat.org/conference2023/

The annual conference of the International Indian Statistical Association (IISA) will take place at the Colorado School of Mines, Golden, Colorado, USA, a beautiful Denver suburb at the foot of the Rocky Mountains, from June 1 to 4, 2023. At the IISA-2023 conference, there will be five plenary talks, twelve special invited talks, three short courses taught by well-known researchers, and many invited and contributed talks and posters covering all areas of statistics, biostatistics, probability, machine learning, and related topics. Many of the participants can stay on campus for a nominal cost. If sponsorship funds come through, the expenses of several participants would be partially reimbursed. IISA welcomes the submission of invited session proposals at https://www.intindstat.org/conference2023/ invitedSessionProposal Portals for contributed talk abstract submission and registration will open shortly.

#### Colloquium in honor of Lutz Dümbgen for his 60th birthday 12 May, 2023 University of Bern, Switzerland

w https://www.imsv.unibe.ch/research/ talks/birthday\_lutz/index\_eng.html Speakers: Günther Walther (Stanford), Kaspar Rufibach (Basel), Angelika Rohde (Freiburg im Breisgau), Richard Samworth (Cambridge). There is no conference fee but registration is mandatory.

#### Foundations of Causal Graphical Models and Structure Discovery

#### May 15–19, 2023

#### Texas A&M University, College Station, USA

w https://web.stat.tamu. edu/~yni/cbms The Department of Statistics at Texas A&M University will host this five-day CBMS conference. Causal discovery

will host this live-day CDWS conference. Causal discovery is becoming increasingly popular in both statistics and machine learning. It is a tool to generate causal hypotheses and bring new insights into existing association-based methods. Kun Zhang from the departments of philosophy and machine learning at Carnegie Mellon University and Mohamed bin Zayed University of Artificial Intelligence will deliver 10 lectures on causal discovery. Tentatively, he will cover representations and usage of causal models, how causality is different from and connected to association, recent machine learning methods for causal discovery, and why and how the causal perspective helps in several learning tasks.

The conference registration fee is \$25, and lunch will be provided.

Visit the conference website to register by April 30.

Two interesting conferences will be held in the following week of our CBMS conferences: the American Causal Inference Conference in Austin, TX (https://sci-info.org/ annual-meeting/) and the Center for Approximation and Mathematical Data Analytics Conference in College Station, TX (https://sites.google.com/tamu.edu/camdaconference/).

#### Panel on the Role of Biostatistics in an Increasingly Big Data/Data Science World

#### May 3, 2023, from 12:00–13:30 EST (Online)

https://www.niss.org/events/panel-role-biostatistics-increasingly-bigdatadata-science-world

What is the role of biostatistics in an increasingly big data/data science focused world. There is so much attention these days on machine learning and AI in biomedical research. What are the opportunities for synergies between biostatistics and other data science disciplines? Don't miss this exciting (and free) panel discussion!

Panelists: Xihong Lin, Jeff Goldsmith, Lance A. Waller and Yu Shyr with moderator David Benkeser.

#### QUARTO Short Course

#### Friday, April 7, 2023, from 12:00–13:30 EST (Online)

#### https://www.niss.org/events/gsn-quarto-short-course

The NISS Graduate Student Network is hosting a new short course on Quarto. The instructor will be Dr. Mine Çetinkaya-Rundel from the Department of Statistical Science at Duke University. Quarto is the next generation of R Markdown for publishing, including dynamic and static documents and multi-lingual programming language support. With Quarto you can create documents, books, presentations, blogs or other online resources. This workshop will introduce you to authoring Quarto documents using RStudio Visual Editor, which provides a user interface for editing all of Pandoc markdown. You will also get a brief look into different formats (such as slides), publishing Quarto documents to QuartoPub, and writing scientific manuscripts with Quarto.

#### Random Structures, Applied Probability and Computation LMS (London Mathematical Society) Research School on Probability

#### June 26–30, 2023 University of Liverpool, UK

w https://sites.google.com/view/Ims-school-liverpool2023/ home

In recent years, the study and understanding of random phenomena, processes and structures have been of great interest among mathematicians and modern probability theorists. These studies often lead to novel connections across the other areas of mathematics and real-world applications (e.g., in the physical, biological, engineering and computer sciences). The research school is intended for **PhD students and early-career researchers interested in probability theory, statistics, and applications**. Our goal is to present, via short courses, selected areas of Probability Theory aiming at the exploration of their connections and encouraging young researchers to think and work outside the box.

Featuring Plenary talks by Igor Kortchemski (CNRS, École polytechnique, France), Remco van der Hofstad (Technische Universiteit Eindhoven, Netherlands), and Aleksandar Mijatovic (University of Warwick); and taught courses on Computational methods for Lévy models - Søren Asmussen (Aarhus Universitet, Denmark), Selfsimilar Markov trees - Jean Bertoin (Universität Zürich, Switzerland), and Pólya urns and other reinforced processes - Cécile Mailler (University of Bath).

Present a poster on your research: register at the website by April 24.

The organizers are Ehsan Azmoodeh, Gabriel Hernán Berzunza Ojeda, Linglong Yuan and Takis Konstantopoulos. If you have any questions, please send an email to lms.liverpool2023@gmail.com.

The London Mathematical Society Research Schools provide training for research students in all contemporary areas of mathematics. Students and post-docs from both the UK and abroad can meet a number of leading international experts in the topic as well as other early career researchers working in related areas.

The LMS Research Schools take place in the UK and support participation of research students from both the UK and abroad. The lecturers are expected to be international leaders in their field. The LMS Research Schools are often partially funded by the Heilbronn Institute for Mathematical Research.

#### Royal Statistical Society Discussion Meeting: Estimating means of bounded random variables by betting May 23, 2023. London, UK

https://rss.org.uk/training-events/events/key-events/discussion-papers/ Join us on May 23rd at 5pm (BST) to hear the authors present their Series B discussion paper at a discussion meeting at RSS Offices in London and online. The event is open to everyone and free to attend. There will be ample networking opportunities at Errol Street with receptions both before and after the meeting. Meeting details and the preprint can be found here: https://rss.org.uk/training-events/events/key-events/discussion-papers/. Registration is essential. We encourage comments on the paper following the presentation by the authors. This can then be written up in 400 words for publication in the journal. Contact j.shorten@rss.org.uk for full details and to reserve a speaking slot of up to 5 minutes. Alternatively, you can decide on the day to make a comment. We very much hope to see you there.

#### UCLA Synthetic Data Workshop April 13–14, 2023 at UCLA Faculty Club, Los Angeles, USA

https://www.niss.org/events/ucla-synthetic-data-workshop We are pleased to announce the UCLA synthetic data workshop. This two-day workshop is hosted by the UCLA Department of Statistics and co-sponsored by NISS, IDRE and UCLA-Amazon Science Hub. The workshop is held at the UCLA Faculty Club. Synthetic data generation is a rapidly growing and highly disciplinary area that draws much attention from both academia and industry. For the development of algorithmic procedures for fraud detection and spam identification, as well as for the construction of AI-driven models in industries like manufacturing and supply chain management, synthetic data has become a valuable resource. The advantages of synthetic data include cost savings, increased speed, agility, increased intelligence, and cutting-edge privacy. According to the Gartner report, synthetic data will overshadow real data in training machine/deep learning models by 2030. Additionally, the MIT Technology Review named synthetic data as one of its top 10 game-changing innovations for 2022 earlier this year. Therefore, it is believed that synthetic data generation will be an indispensable part of the next-generation machine learning workflow. Despite numerous successful applications of synthetic data, its scientific foundation, e.g., the tradeoff among fidelity, utility, and privacy, is still missing. Additionally, industrial standards for generating and utilizing synthetic data are not fully developed. Furthermore, the privacy law concerning about synthetic data has not been fully developed. Therefore, this workshop is to form a community of synthetic data researchers (from statistics, machine learning and mathematics), policymakers, and industrial partners, and bring them together to collaborate on the development of the theory, methodology, and algorithms needed to produce synthetic benchmark datasets and algorithms. Updates to Speakers, Program and Poster schedules available at the workshop website.

# **Employment Opportunities**

#### **Hong Kong**

**City University of Hong Kong** Dean of School of Data Science https://jobs.imstat.org/job//68553039

#### Switzerland: Lausanne

Ecole Polytechnique Federale de Lausanne PhD https://jobs.imstat.org/job//52027160

#### Switzerland: Lausanne

**Ecole Polytechnique Federale de Lausanne** (EPFL), Mathematics Dept. Postdoctoral Researcher

https://jobs.imstat.org/job//68226979

#### United States: Irvine, CA

UCI Assistant/ Associate/ Full Professor In-Residence https://jobs.imstat.org/job//68275019

#### **United States: Fort Collins, CO**

**Colorado State University** Research Associate II and III https://jobs.imstat.org/job//68432467

#### United States: Washington, DC CIA

Science, Technology & Weapons Analyst https://jobs.imstat.org/job//65521691

United States: Washington, DC CIA DA Fellowship Program https://jobs.imstat.org/job//65521688

#### **United States: Chicago, IL**

**University of Chicago** Preceptor in Data Science https://jobs.imstat.org/job//68077382

#### **United States: Durham, NC**

NCSSM Instructor of Mathematics https://jobs.imstat.org/job//68332195

#### United States: South Orange, NJ Seton Hall University

Assistant Professor in Applied Mathematics https://jobs.imstat.org/job//68077031

#### **United States: New York, NY**

Success Academy Charter Schools Associate, 3-12 Mathematics Content Design https://jobs.imstat.org/job//68467898

#### **United States: Buffalo, NY**

University at Buffalo Chair, Department of Biostatistics https://jobs.imstat.org/job//68244166

#### **United States: College Station, TX**

Texas A&M University, Department of Statistics Research Data Scientist https://jobs.imstat.org/job//68346623

#### **United States: Fairfax, VA**

# George Mason University, Department of Statistics.

Tenure-track and renewable-term, non-tenure-track Assistant/Associate Professor https://jobs.imstat.org/job//68586895

# **International Calendar of Statistical Events**

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

#### **Online and Ongoing**

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

Webinar series w https://www.niss.org/copss-niss-covid-19-datascience-webinar-series

**ONLINE** One World ABC Seminar w https://warwick.ac.uk/fac/sci/statistics/news/upcomingseminars/abcworldseminar

**ONLINE** One World Probability Seminar

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

**ONLINE** 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 2023

April 4–6: Livingstone, Zambia. 2023 Zambia Conference: Better Lives for 2030 w https://www.zamstats.gov.zm/2023-iaos-isiconference/

::: Advertise current job opportunities for only \$355 for 60 days ::: See https://jobs.imstat.org for details :::

**MONLINE** April 7: Online 12:00–13:30 EST. QUARTO Short Course w https://www.niss.org/events/gsn-quarto-shortcourse

April 13–14: Los Angeles, USA. UCLA Synthetic Data Workshop w https://www.niss.org/events/ucla-synthetic-dataworkshop

#### May 2023

With the second second

May 5-7: Atlanta, USA. 8th Workshop on Biostatistics and Bioinformatics w https://math.gsu.edu/ yichuan/2023Workshop/

Wims May 8-10: Princeton University, NJ, USA. Statistical Foundations of Data Science and their Applications w http://fan60.princeton.edu/

12 May: Bern, Switzerland. Colloquium in honor of Lutz Dümbgen for his 60th birthday w https://www.imsv.unibe.ch/ research/talks/birthday\_lutz/index\_eng.html

*Lims* May 15–16: Duke University, NC, USA. Southeastern Probability Conference I w https://services.math.duke.edu/~rtd

May 15–19: College Station, USA. Foundations of Causal Graphical Models and Structure Discovery w https://web.stat. tamu. edu/~yni/cbms

May 23: London, UK. Royal Statistical Society Discussion Meeting: Estimating means of bounded random variables by betting w https://rss.org.uk/training-events/events/key-events/ discussion-papers/

May 23–26: St. Louis, Missouri, USA. Symposium on Data Science and Statistics w https://ww2.amstat.org/meetings/sdss/2023/ index.cfm

May 24–26: Banff, Canada. ASA/IMS Spring Research Conference (SRC) on Statistics in Industry and Technology w https://sites.google.com/view/src2023/home

May 30–June 3: State College, PA, USA. USCOTS2023: United States Conference on Teaching Statistics w https://www.causeweb.org/cause/uscots/uscots23

#### June 2023

June 1–4: Colorado School of Mines, USA. 2023 International Indian Statistical Association (IISA) Conference w https://www.intindstat.org/conference2023/

June 6–9: West Lafayette, USA. 10th International Purdue Symposium on Statistics w https://www.stat.purdue.edu/ symp2023/index.html

**HYBRID** June 6–9: Heraklion, Greece, and online. ASMDA 2023 Conference and Demographics 2023 Workshop w http:// www.asmda.es/

June 12–16: Esch-sur-Alzette, Luxembourg. International Conference on Malliavin Calculus and Related Topics w https:// math.uni.lu/icmcrt/

**Market Statistical Network Analysis and Beyond** (SNAB 2023) **w** https://www.snab2023.org/

June 18–21: Anchorage, USA. WNAR2023 w https://wnar. org/wnar2023

VEW June 26–30: Liverpool, UK. Random Structures, Applied Probability and Computation (*LMS Research School on Probability*) w https://sites.google.com/view/lms-school-liverpool2023/home

*ums* June 28–30: Nancy, France. 21st INFORMS/APS Meeting w https://informs-aps2023.event.univ-lorraine.fr/

**Conference w** https://maths.swjtu.edu.cn/english/ICSA\_2023 China\_Conference w https://maths.swjtu.edu.cn/english/ICSA\_2023\_ China\_Conference/Conference\_Introduction.htm

#### **July 2023**

July 3–7: Warsaw, Poland. 34th European Meeting of Statisticians (EMS) w https://ems2023.org

July 4-6: Nottingham, UK. Modelling in Industrial Maintenance and Reliability w ima.org.uk/20581/12th-mimar/

July 7–10: Thessaloniki, Greece. International Workshop on Applied Probability (IWAP2023) w http://iwap2020.web.auth.gr/

July 10–14: São Paolo, Brazil. 16th CLAPEM: Latin American Congress of Probability & Mathematical Statistics w https://www.ime.usp.br/~16clapem/

### International Calendar continued

#### July 2023 continued

July 11–13: Peking University, Beijing, China. Joint Conference on Statistics and Data Science in China w https:// www.stat-center.pku.edu.cn/en/Events/conference/1364567.htm

July 15–20: Ottawa, Canada. ISI World Statistics Congress w https://www.isi2023.org/conferences/ottawa-2023/

Lims July 24–28: Lisbon, Portugal. 43rd Conference on Stochastic Processes and their Applications w https://www.spa2023.org/

#### August 2023

Lims August 2–5: Toronto, Canada. IMS New Researchers Conference w sites.google.com/site/linbowangpku/nrc-2023

August 5–10: Toronto, Canada. IMS Annual Meeting at JSM 2023 w https://ww2.amstat.org/meetings/jsm/2023/

*Lims* August 14–15: University of Virginia, USA. Southeastern Probability Conference II w https://services.math.duke.edu/~rtd

August 20–25: Tokyo, Japan. ICIAM2023: 10th International Congress on Industrial and Applied Mathematics w https:// iciam2023.org/

### September 2023

September [TBC]: Ljubljana, Slovenia. 23rd European Young Statisticians Meeting w https://sites.google.com/view/eysm2023

### December 2023

**Conference on Statistics and Data Science (ICSDS) w** https:// sites.google.com/view/icsds2023

#### January 2024

**IPDATED** January 4–7 (postponed from January 2021): Melbourne, Australia. IMS Asia Pacific Rim Meeting (IMS-APRM2024) w http://ims-aprm2024.com/

#### March 2024

March 10–13: Baltimore, USA. 2024 ENAR/IMS Spring Meeting w http://www.enar.org/meetings/future.cfm

#### June 2024

June 9–12: Fort Collins, Colorado, USA. WNAR2024, joint with Graybill Conference w https://wnar.org/meetings

#### **July 2024**

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

July 7–14: Sydney, Australia. 15th International Congress on Mathematics Education w https://icme15.com/home

#### August 2024

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

Congress in Probability and Statistics w TBC

### August 2025

JSM 2025 w http://www.amstat.org/ASA/Meetings/

#### August 2026

**Lims** August 1–6: Boston, MA, USA. **JSM 2026 w** http://www. amstat.org/ASA/Meetings/Joint-Statistical-Meetings.aspx

Are we missing something? If you know of any statistics or probability meetings which aren't listed here, please let us know. You can email the details to Elyse Gustafson at ims@imstat.org, or you can submit the details yourself at https://www.imstat.org/ ims-meeting-form/ We'll list them here in the Bulletin, and on the IMS website too, at imstat.org/meetings-calendar/

#### Membership and Subscription Information: 2023

#### 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 \$130 is added to the dues of members for each scientific journal selected (\$87 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 2023 are available to *The Annals of Applied Probability* (\$245), *The Annals of Applied Statistics* (\$245), *The Annals of Probability* (\$245), *The Annals of Statistics* (\$245), *Statistical Science* (\$202), and *IMS Bulletin* (\$115). **General subscriptions** are for libraries, institutions, and any multiple-readership use. Institutional subscriptions for 2023 are available to *The Annals of Applied Probability*, *The Annals of Applied Statistics*, *The Annals of Probability*, and *The Annals of Statistics* (each title \$563 online only/\$707 print+online), *Statistical Science* (\$324/\$403), and *IMS Bulletin* (\$167 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. Copyright © 2023 by the Institute of Mathematical Statistics. 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 **\$355.00**. 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
<sup>1</sup> /3 page horizontal	4.93" wide x 4.0" high (125.5 x 102 mm)	\$305
1/3 page vertical	2.39" wide x 9.42" high (60.7 x 239.1 mm)	\$305
1/2 page horizontal	7.5" wide x 4.7" high (190.5 x 119.4 mm)	\$380
1/2 page vertical	3.67" wide x 9.42" high (93.1 x 239.1 mm)	\$380
Full page (to edge, including 1/8" bleed)	8.75" wide x 11.25" high (222 mm x 286 mm)	\$520
Full page (within usual Bulletin margins)	7.5" wide x 9.42" high (190.5 mm x 239.1 mm)	\$520

#### 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

# the **next June/July** 2023

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

# DEADLINES 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

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

64 (print) 6 (online)

# Ann. statist Feb 2023 THE ANNALS of **STATISTICS**

AN OFFICIAL JOURNAL OF THE INSTITUTE OF MATHEMATICAL STATISTICS

#### Articles

Grouped variable selection with discrete optimization: Computational and statistical perspectives . . HUSSEIN HAZIMEH, RAHUL MAZUMDER AND PETER RADCHENKO 1 Characterizing the SLOPE trade-off: A variational perspective and the Donoho-Tanner limit...... ZHIQI BU, JASON M. KLUSOWSKI, CYNTHIA RUSH AND WEIJIE J. SU 33 Optimal estimation of high-dimensional Gaussian location mixtures NATALIE DOSS, YIHONG WU, PENGKUN YANG AND HARRISON H. ZHOU 62 High-dimensional latent panel quantile regression with an application to asset pricing ALEXANDRE BELLONI, MINGLI CHEN, OSCAR HERNAN MADRID PADILLA AND ZIXUAN (KEVIN) WANG 96 ScreeNOT: Exact MSE-optimal singular value thresholding in correlated noise David Donoho, MATAN GAVISH AND ELAD ROMANOV 122 Testing for outliers with conformal p-values . . . STEPHEN BATES, EMMANUEL CANDÈS, LIHUA LEI, YANIV ROMANO AND MATTEO SESIA 149 Sharp global convergence guarantees for iterative nonconvex optimization with random ......KABIR ALADIN CHANDRASEKHER, ASHWIN PANANJADY AND CHRISTOS THRAMPOULIDIS 179 Breaking the winner's curse in Mendelian randomization: Rerandomized inverse variance weighted estimator.....XINWEI MA, JINGSHEN WANG AND CHONG WU 211 233 A new approach to tests and confidence bands for distribution functions LUTZ DÜMBGEN AND JON A. WELLNER 260 Conditional predictive inference for stable algorithms LUKAS STEINBERGER AND HANNES LEEB 290 Variable selection, monotone likelihood ratio and group sparsity...CRISTINA BUTUCEA, ENNO MAMMEN, MOHAMED NDAOUD AND ALEXANDRE B. TSYBAKOV 312 Nested Markov properties for acyclic directed mixed graphs THOMAS S. RICHARDSON, ROBIN J. EVANS, JAMES M. ROBINS AND ILYA SHPITSER 334

Uniform consistency in nonparametric mixture models BRYON ARAGAM AND RUIYI YANG 362

Vol. 51, No. 1-February 2023