# IMS Bulletin



#### March 2020

### **CONTENTS**

- 1 World Congress in Seoul
- 2 **Members' news:** Alan Welsh, Jay Bartroff, Larry Goldstein, Gary Lorden, Xiao-Li Meng, David Allison
- 4 Ruobin Gong: Beyond Enigmatic
- 6 **Radu's Rides:** A True Embarrassment of Riches
- 8 **Previews:** Roger Koenker, Paul Rosenbaum, Nicolas Curien
- 11 International Data Science in Schools Project
- 12 Student Puzzle Corner; New Researchers Conference
- 13 **Recent papers:** Electronic Journal of Statistics; Statistics Surveys
- 14 **Obituary:** Colin Blyth, Janos Galambos, Jonathan Cryer
- 16 Meetings
- 23 Employment Opportunities
- 24 International Calendar of Meetings
- 27 Information for Advertisers

Read it online: imstat.org/news

# **World Congress in Seoul**

The Bernoulli–IMS 10th World Congress in Probability and Statistics will be held at Seoul National University in Seoul, South Korea, from August 17th to 21st, 2020.



This quadrennial joint meeting is of steadily increasing importance to the scientific vitality of the fields of statistics and probability and their applications. The Congress in Seoul will have 14 (plenary) named lectures, one public lecture, 40 invited sessions featuring 120 speakers, as well as contributed talks and poster sessions.

Immediately before the Congress, on August 15–16, a two-day **Young Researchers Meeting** will also be held at Seoul National University. The first day focuses on Data Science, and the second features presentations and discussions on career development. (Registration deadline is 18 July, but if you're registered for the World Congress, there's no additional fee.)

**Registration, hotel information and abstract submission** to the World Congress are now open at the website https://www.wc2020.org/. The abstract submission dead-line is March 31, and the early-bird registration deadline is May 31.

Also on the Congress website, you will find more information about the named lectures and invited sessions, the Young Researchers Meeting, accommodation, transportation, and visas.

As the largest city in South Korea, dynamic Seoul is a bewitching mix of ancient and modern structures, packaged in a surprisingly compact metropolis that has earned it the designation of a UNESCO City of Design. The nation's capital has a cutting-edge cityscape of glass, steel and futuristic skyscrapers, which tower over traditional wooden houses with tiled roofs and mazes of cobbled alleys in distinct, village-like districts.

As Program Chair Siva Athreya, and Local Organizing Committee Chair Hee-Seok Oh, say, *"We look forward to seeing you in Seoul!"* 





# **IMS** Bulletin

Volume 49 • Issue 2 March 2020 ISSN 1544-1881

### **Contact information**

IMS Bulletin Editor: Tati Howell bulletin@imstat.org

UPDATED Managing Editor: Bob Keener

Contributing Editors: Radu Craiu, Anirban DasGupta, Yoram Gat, David Hand, Takis Konstantopoulos, Xiao-Li Meng, Regina Nuzzo, Kavita Ramanan and Terry Speed

UPDATED w https://imstat.org/news

f 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:          | Susan Murphy<br>president@imstat.org      |
|---------------------|---|
| President-Elect:    | Regina Liu<br>president-elect@imstat.org  |
| Past President:     | Xiao-Li Meng<br>president-past@imstat.org |
| Treasurer:          | Zhengjun Zhang<br>zjz@stat.wisc.edu       |
| Program Secretary:  | Ming Yuan<br>ming.yuan@columbia.edu       |
| Executive Secretary | :Edsel Peña<br>pena@stat.sc.edu           |

# **IMS Members' News**

#### Alan Welsh receives Australian Academy of Science's Hannan Medal

Professor Alan Welsh, Australian National University, has been recognised by the Australian Academy of Science for his highly original and insightful contributions to the advancement of statistics. Professor Welsh was elected a Fellow of the Australian Academy (FAA) in 2007 (and an IMS Fellow in 1990).

Alan Welsh has developed useful new methodology, derived the properties of these and other methods and clarified relationships between different statistical methods, all in a particularly wide variety of problems. He has developed innovative new models for count data with many zeros and compositional data, including for longitudinal and clustered forms of these data. He has made important contributions to inference, robustness, the bootstrap and model selection for mixed models. His research on applications of smoothing methods to clustered data demonstrated that remarkable improvements can be achieved by taking proper and careful account of the dependence structure when constructing a smoother. Professor Welsh contributed to resolving how to do maximum likelihood estimation for sample survey data and, in ecological survey analysis, he made especially important contributions to distance sampling and occupancy modelling. All this work, and more, has the characteristic of theoretical depth combined with substantial practical relevance.

The Hannan Medal, awarded by the Australian Academy of Science, recognizes research in any of the fields of statistical science, pure mathematics, applied mathematics and computational mathematics and is made in one of those three areas in turn at two-yearly intervals. It honors the contribution to time series analysis of the late Professor Edward J. Hannan, FAA, Professor of Statistics at the Research School of Social Sciences of the Australian National University.

The Hannan award is made for research carried out mainly in Australia. Work carried out through a candidate's entire career is taken into consideration but special weight is given to recent research.

You can watch a video about Alan Welsh's award at https://www.science.org.au/ opportunities-scientists/recognition/honorific-awards/honorific-awardees/2019-awardees



A still from the Australian Academy of Science video interview with Alan Welsh

# **More Members' News**

#### Jay Bartroff, Larry Goldstein and Gary Lorden investigate baseball home run statistics

Members Jay Bartroff, Larry Goldstein and Gary Lorden were chosen by the Commissioner of Major League Baseball (MLB) to be the statisticians on a committee with physicists,

engineers, and baseball experts studying MLB's recent surge in home runs. The committee studied vast amounts of game data, performed laboratory tests on balls, visited the Rawlings manufacturing facility where MLB's baseballs are manufactured, and made recommendations to MLB and Rawlings for future monitoring, testing, and storage of baseballs. You can read a summary of the committee and its work at: https://www.mlb.com/press-release/ major-league-baseball-receives-report-onincreased-home-run-rate-278136100.



There is also a commentary on the committee's conclusions at https://www.mlb.com/ news/mlb-report-on-baseballs-home-run-rates-c278120310

#### Xiao-Li Meng interview at ITHAKA's Next Wave Conference

The *Scholarly Kitchen* (https://scholarlykitchen.sspnet.org/) is a moderated and independent blog of the Society for Scholarly Publishing that brings together differing opinions, commentary, and ideas, and presents them openly.

In a recent post, Roger C. Schonfeld reported: "At ITHAKA's Next Wave conference, leaders from academic societies, libraries, publishers, and other organizations come together to focus on important strategic issues challenging higher education, specifically in the United States. Each year at this event, ITHAKA President, Kevin Guthrie, interviews someone who is at the leading edge, working to change organizations, systems or practices. On December 4, 2019, Kevin talked with Xiao-Li Meng, Professor of Statistics at Harvard University, about the increasingly central role data science is playing in research and teach-

Data science creates a new common language for all of us. It creates a platform for us to talk with each other about shared interest and that's good for humankind. – Xiao-Li Meng ing, the changes he sees on the horizon and is helping to foster, and the impetus behind the recently launched open access *Harvard Data Science Review* where he is Editor-in-Chief."

You can read the interview at https:// scholarlykitchen.sspnet.org/2020/01/16/ data-science-revolution-interview-xiao-limeng/ or watch the video at https://vimeo. com/385032627.

#### **Correction: David Allison**

In the last issue we incorrectly located David Allison, who had given a testimony for the National Academies, at the University of Alabama, Birmingham. In fact, David is now Dean and Distinguished Professor at the Indiana University School of Public Health–Bloomington.

#### IMS Journals and Publications

- Annals of Statistics: Ming Yuan, Richard Samworth http://imstat.org/aos @http://projecteuclid.org/aos
- Annals of Applied Statistics: Karen Kafadar http://imstat.org/aoas ©http://projecteuclid.org/aoas
- Annals of Probability: Amir Dembo http://imstat.org/aop ©http://projecteuclid.org/aop
- Annals of Applied Probability: Francois Delarue, Peter Friz http://imstat.org/aap @http://projecteuclid.org/aoap
- Statistical Science: Sonia Petrone http://imstat.org/sts @http://projecteuclid.org/ss
- IMS Collections Mhttp://projecteuclid.org/imsc
- IMS Monographs and IMS Textbooks: Nancy Reid https://www.imstat.org/journals-andpublications/ims-monographs/

#### IMS Co-sponsored Journals and Publications

- *Electronic Journal of Statistics*: Domenico Marinucci http://imstat.org/ejs @ http://projecteuclid.org/ejs
- Electronic Journal of Probability: Andreas Kyprianou © https://projecteuclid.org/euclid.ejp
- Electronic Communications in Probability: Giambattista Giacomin
- Mhttps://projecteuclid.org/euclid.ecp
- Journal of Computational and Graphical Statistics: Tyler McCormick http://www.amstat.org/publications/jcgs
- Dilog into members' area at imstat.org Statistics Surveys: David Banks
- http://imstat.org/ss http://projecteuclid.org/ssu
- Probability Surveys: Ben Hambly http://imstat.org/ps Mhttp://www.i-journals.org/ps/

#### IMS-Supported Journals

- ALEA: Latin American Journal of Probability and Statistics: Roberto Imbuzeiro Oliveira Mhttp://alea.impa.br/english
- Annales de l'Institut Henri Poincaré (B): Gregory Miermont, Christophe Sabot http://imstat.org/aihp @ http://projecteuclid.org/aihp
- Bayesian Analysis: Michele Guindani Mhttps://projecteuclid.org/euclid.ba
- Bernoulli: Mark Podolskij, Markus Reiß http://www.bernoulli-society.org/ ©1 http://projecteuclid.org/bj
- Brazilian Journal of Probability and Statistics: Enrico Colosimo http://imstat.org/bjps @http://projecteuclid.org/bjps

#### IMS-Affiliated Journal

- Observational Studies: Dylan Small Mhttps://obsstudies.org/
- Probability and Mathematical Statistics: K. Bogdan, M. Musiela, J. Rosiński, W. Szczotka, & W.A. Woyczyński © http://www.math.uni.wroc.pl/~pms
- Stochastic Systems: Shane Henderson Mhttps://pubsonline.informs.org/journal/stsy

# Now, Your Information is Beyond Enigmatic



Ruobin Gong is Assistant Professor of Statistics at Rutgers University. Her research interests lie at the theoretical foundations of Bayesian and generalized Bayesian methodologies, statistical modeling, inference, and computation with differentially private data, and ethical implications of aspects of modern data science. Ruobin received her PhD in statistics from Harvard University in 2018. She is currently an associate editor of the *Harvard Data Science Review*.

We've all been there: your airline company emails you and asks your opinion about their service. Still bummed out about last summer when they almost lost your luggage, you take up the invitation hoping to rant about it, only to find out three pages into the survey they want to know how old you are and what you do for a living.

Deep down inside, you know that telling the truth may eventually help the airline company understand your needs as a budding academic. In the summer months, we fly from conference to conference living out of a suitcase, and parting ways with it for even a single day will throw our talks, meetings, and travel plans into havoc. But still, that doesn't feel reason enough to warrant an all-out confession. Even if telling the truth may potentially do you (and your fellow academics) a big favor, it's not worth the risk to expose your most personal information. After all, who can guarantee what I put into the survey will be used solely for the betterment of my or others' experience, and never against my rights - to anonymity, confidentiality, and privacy? So, you put down some made-up age and occupation in the answer box, and quickly move on.

When it comes to surveys, the feeling of distrust that a respondent harbors against the surveyor corrodes the quality of the data. To refuse participation biases the survey at the sampling stage, and erroneous answers further harms data informativeness. In marketing surveys, faults like this at most render a company clueless about the true opinions of its customers. Yet grievous consequences await when distrust imbues surveys of more substantial significance. In just about a month, the 2020 U.S. decennial Census, the most comprehensive enumeration of the living population of America, will hit the ground running. As mandated by the U.S. Constitution, data obtained through the Census supply crucial factual evidence to economic and policy decisions. The Census serves as the basis for the apportionment of House seats, as well as the allocation of federal funding and resources (Sullivan, 2020). It is a massive and serious statistical undertaking.

The Census differs from all other surveys in one important aspect. By design, it should cover 100% of its target population, that is every single person living in the United States at the time it takes place. Therefore, when it comes to Census data releases, privacy protection carries insurmountable significance due to the sheer number of respondents involved. The 2020 Census made a revolutionary step forward by endorsing a new, and formal, standard for privacy protection, called differential privacy (Dwork et al., 2006; Abowd, 2018). Differential privacy draws a sharp distinction from the heuristic approach to privacy protection that traditional methods typically follow, such as full and partial suppression of data tables and swapping of individuals. It supplies a mathematical definition on what is meant by the *privacy* of data releases, which doubles as a metric to quantify the amount of privacy the data release gives away at most.

How does differential privacy work? Suppose that you're filling out the Census questionnaire, and let's denote your *true* answer as  $x_t$ . You are reminded of the risk of a hypothetical privacy breach, and contemplate whether to put down instead a *fabricated* answer, say  $x_f$ . Your answer, together with billions of others', constitute the enormous Census database  $\mathcal{D}$ , which takes the value  $\mathcal{D}(x_t)$  if you supplied the true answer, or  $\mathcal{D}(x_f)$  if the fabricated one. (Let's say the others' answers, whether true or fabricated, are identical in  $\mathcal{D}(x_t)$  and  $\mathcal{D}(x_f)$ .) Finally, the Census Bureau releases the database summary generated by a probabilistic algorithm, based on the observed (and confidential) database: S = S ( $\mathcal{D}$ ).

Suppose an ill-minded hacker is eyeing the Census data, hoping to learn about your information. Looking at the released summary S, the hacker needs to discern between two possibilities: that the data you contributed was true (H) or fabricated (  $\overline{H}$  ). If the algorithm that generated S is sufficiently private, the information contained in S (expressed in terms of probabilities) that can sufficiently discern H from  $\overline{H}$  is limited. Precisely speaking, S is  $\epsilon$ -differentially private if the log ratio of its probabilities evaluated under either hypothesis (i.e. their respective likelihoods) is bounded within the  $\epsilon$ -neighborhood around zero:

$$\log \frac{P(S \mid H)}{P(S \mid \bar{H})} \in [-\epsilon, \epsilon], \qquad [1]$$

and that such is true for every respondent (including you) who contributes to the Census database D. The  $\epsilon$  here, called the privacy loss budget, controls the extent to which we are willing to tolerate discernibility among hypotheses, or leak of information. The smaller the  $\epsilon$ , the more stringent the bound becomes, and the less informative S is relative to the pair of hypotheses *H* and  $\bar{H}$ .

The above quantity looks like an incredibly simple, if not overly simple, metric to quantify the so-called "information" in S regarding H versus  $\bar{H}$  . But do take it seriously. The failure to maintain this log probability ratio at a small magnitude by its encrypted messages was the Achilles' heel of the Naval Enigma and the Tunny machines, a deadly giveaway that led to their heroic breaking by the genius scientists at Bletchley Park during World War II (McGrayne, 2011; Zabell, 2012, 2015). In I.J. Good's account of Alan Turing's statistical contribution during the war (Good, 1979), he defined the "weight of evidence" concerning a hypothesis H as against  $\overline{H}$  provided by evidence S, written as  $W(H/\bar{H}:S)$ , a quantity that works out to be precisely the log probability ratio in [1]. For cipher machines such as the Enigma and the Tunny, S stands for the encrypted messages, and H,  $\bar{H}$  are hypotheses concerning the different configurations of the cipher wheels. If a configuration hypothesis receives from S a disproportionately large weight of evidence relative to other hypotheses, there is reason to think that it may be the correct configuration. Turing called one unit of the log probability ratio in [1] a natural ban, which is equal to 4.34 units of deciban (ten times the base 10 logarithm of the probability ratio), the "smallest change in weight of evidence that is directly perceptible to human intuition" (Good, 1979, p394). Carrying over this calculation to the privacy context, a privacy loss budget  $\epsilon$  set at or less than  $1/4.34 \approx 0.23$  makes the hypotheses regarding the truthfulness of your input data probabilistically indiscernible, based on the differentially private release S. It would be fair to say, then, that the privacy algorithm behind the released

summary *S* encrypts your personal information securely, a job much better done than the Enigma machine. In other words, your information is now *"beyond Enigmatic"*!

Differential privacy brings clarity to the meaning of privacy through a formal and verifiable definition, setting a rigorous standard for implementation, investigation and future improvement. It merits other benefits from a technical point of view. Data releases compliant with differential privacy are resistant to post-processing, and behave nicely under the compounding of multiple sources; see Dwork et al. (2014) for details and Wood et al. (2018) for an approachable introduction. Differential privacy further permits the transparent dissemination of the privacy algorithm without compromising the privacy guarantee, drawing an analogy with public-key encryption. For statisticians, this means that the data curator is free to publicize the inner specification of the privacy mechanism (as in the case of US Census Bureau, 2020), paving ways for statistical methods to account for its effect, and to maintain inferential validity based on private releases (Gong, 2019).

Just as no probabilistic promise is ever definitive, no privacy is absolute if we simultaneously demand to learn useful information. With differential privacy, however, the tradeoff between privacy and information is put in concrete terms. In the 2020 Census, we in America will collectively pay an epsilon price in privacy, in exchange for a large body of useful knowledge about this country we live in, and the people who live in it with us. The Census operationalizes democracy and equality through enabling fair and data-driven allocation of resources. A transparent and effective framework for privacy protection is yet another reason to overcome mistrust, and to actively and honestly participate. Time will tell whether the price we pay is money well spent.

#### References

- Abowd, J.M. (2018). The US Census Bureau Adopts Differential Privacy. In Proc. 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, 2867–2867. ACM.
- Dwork, C., McSherry, F., Nissim, K., and Smith, A. (2006). Calibrating Noise to Sensitivity in Private Data Analysis. In *Theory of Cryptography Conference*, 265–284. Springer.
- Dwork, C. and Roth, A. (2014). The Algorithmic Foundations of Differential Privacy. In *Foundations and Trends in Theoretical Computer Science*, **9**(3–4): 211–407.
- Gong, R. (2019). Exact inference with approximate computation for differentially private data via perturbations. arXiv:1909.12237
- Good, I. J. (1979). Studies in the History of Probability and Statistics XXXVII: A. M. Turing's Statistical Work in World War II. *Biometrika*, **66**(2): 393–396.
- McGrayne, S. B. (2011). The Theory That Would Not Die: How Bayes' Rule Cracked the Enigma Code, Hunted Down Russian Submarines, and Emerged Triumphant from Two Centuries of Controversy. Yale University Press.
- Sullivan, T. A. (2020). Coming to our Census: How social statistics underpin our democracy (and republic). *Harvard Data Science Review*, 2(1). https://hdsr.mitpress.mit.edu/ pub/lq1cbvkv.
- US Census Bureau (2020). 2010 Demonstration Data Products. https://www.census.gov/programssurveys/decennial-census/2020-census/planningmanagement/2020-census-data-products/2010demonstration-data-products.html. Accessed on Feb 02, 2020.
- Wood, A., Nissim, K., et al. (2018).
  Differential Privacy: A Primer for a Non-Technical Audience. Vanderbilt Journal of Entertainment & Technology Law 21(1): 209.
- Zabell, S. (2012). Commentary on Alan M. Turing: The Applications of Probability to Cryptography. *Cryptologia*, **36**(3):191–214.
- Zabell, S. (2015). Statistics at Bletchley Park. In Reeds, J. A., Diffie, W., and Field, J., eds, Breaking Teleprinter Ciphers at Bletchley Park: An edition of I.J. Good, D. Michie and G. Timms: General report on Tunny with emphasis on statistical methods (1945), pages lxxv-ci. John Wiley & Sons.

# Radu's Rides: A True Embarrassment of Riches

Radu V. Craiu, our newest Contributing Editor, is Professor and Chair of Statistical Sciences at the University of Toronto. He studied Mathematics at the University of Bucharest (BS 1995, MS 1996), and received a PhD from the Department of Statistics at The University of Chicago in 2001. His main research interests are in computational methods in statistics, especially, Markov chain Monte Carlo algorithms (MCMC), Bayesian inference, copula models, model selection procedures and statistical genetics. He is currently Associate Editor for the *Harvard Data Science Review, Journal of Computational and Graphical Statistics, The Canadian Journal of Statistics* and *STAT - The ISI's Journal for the Rapid Dissemination of Statistics Research.* He received the 2016 CRM–SSC prize and is an Elected Member of the International Statistical Institute.



e live in interesting times. The initial excitement produced by the successful application of machine learning algorithms in speech and image recognition has blossomed into a veritable scientific revolution. No one remembers or cares whether the writing was on the wall, as by now it's everywhere. The abundance of ideas and, naturally, related papers is impressive and with the power of 2020 hindsight we now realize that it has taken the machine learning community and its adjacent disciplines by surprise.

The NeurIPS conference grew in just a few years from fewer than 800 participants to more than 8000, with over 3000 registration tickets being sold in 38 seconds. Other top conferences in ML (ICML) or artificial intelligence (AISTATS) followed suit. None of these organisations were prepared to handle the immense volume that was precipitously lavished upon them. The acceptance rates for conference papers have plummeted, while social media is abuzz with communications from frustrated researchers who see their papers summarily rejected after reviews of questionable quality. Even in the Good Old Times of Yore, we all experienced the occasional lack of understanding from our peer reviews, but the new order has crystallized the notion that, simply put, there aren't enough peers for all these reviews!

Talking to statisticians, applied

mathematicians or non-ML computer scientists, one hears a lot of grumbling about transfer of resources from their disciplines and neglect of fundamental research that until yesterday supported the type of models that machine learners had built their successes on. Students migrate in droves from mathematics, statistics and other computer science sub-disciplines to machine learning, while many young researchers feel that the fastest way to glory is not through painstaking theoretical work, but rather algorithmic manipulations that lead to instances of successful predictions. It is unfortunate that the many who chose this path fail to see the intellectual leaders of the ML field who are working in a very different mode, one in which human intelligence and careful thinking, backed by principled inference, fuels the artificial intelligence. In this intense race towards a rather vaguely defined goal, the discipline most under-served-or, worse, harmed-is ML itself. One does not need to dig deep to find out that the consequences are hurting the community at all levels.

On the student side, I was recently astounded to learn from Andreas Madsen's article in *Medium*, "Becoming an Independent Researcher and getting published in ICLR with spotlight," that a young undergraduate aspiring to enter a good PhD program in ML must have already published two papers in the proceedings of some of the top conferences of the field. In other words, one must produce a significant part of a PhD thesis *before* entering a PhD program! The incredibly high hoops one must jump through in order to enter the academic world will likely deter many talented people from pursuing this path. In turn, this lack of formal training that reinforces a principled approach to scientific investigation threatens to dilute the impact of future methodological contributions.

On the faculty side, especially the pre-tenure kind which makes up the largest part of the ML academic community, the avalanche of papers published daily on arXiv, bioRxiv, etc., makes it nearly impossible to keep up with the literature. (A search on Google Scholar for "machine learning arxiv" yields 10,500 results, just in 2020.) "Nobody reads nor cites, just writes" is being repeated to me ad nauseam. The contents of prestigious ML conference proceedings, once reliable sources of good papers, are increasingly contested due to the lack of reviewers. In a recent Science interview, the famous statistician and prominent debunker John Ioannidis warns that the two tenets of academic publishing, credit and responsibility, are in serious jeopardy because of publication inflation. This may be merely anecdotal for tenured faculty, but it can be life-altering and health-damaging for pre-tenure researchers.

On the more practical aspects of *Continues on* page 7

research applications in industry, there is no doubt that the initial successes of ML were spectacular. Carefully designed algorithms have been able to win at Go, translate between languages, predict the next word I am about to type and, by and large, produce a general sense that humanity is about to win at life or, at least, reinvent life itself (see Yuval Noah Harari's 2018 book, 21 Lessons for the 21st Century). Since those initial successes, the well-known adage, "When you have a large hammer, every problem looks like a nail" has come to describe the modus operandi for much of the industry-related data-driven research. Problems associated with moderate data volumes - usually handled with a careful statistical analysis - are now lumped together with problems that benefit from millions, or billions, of data points and are handled using similar algorithms. The potential for spectacular/costly errors is growing fast and one wonders where it will strike first (not if...).

Is there a single solution to all this? Not really. If we have to start somewhere, maybe a call to the ML community-at-large to turn their attention to foundational issues is a good start. If taken seriously, a wider effort to set research on principled legs will come accompanied by a shrinkage of propensity for publication, an increase in the standard of proof and more self-restraint in grandeur claims. The Data Science ecosystem would benefit from similar efforts in adjacent disciplines. Statisticians started this a while ago, and continue to re-examine foundational issues in various forms and forums, a good example being the BFF series of conferences (with its next stop in Toronto: see poster [right] and more info at http://www.fields.utoronto. ca/activities/19-20/BFF7). We look forward to hosting and learning from statisticians, probabilists, computer scientists and

philosophers who will discuss the different statistical paradigms— Bayesian, frequentist and fiducial—for conducting sound data science. And we promise to *not* put out any conference proceedings!



# **Previews of Special IMS Lectures**

# **IMS Medallion Lecture: Roger Koenker**



Roger Koenker

Roger Koenker was born in 1947 in North Dakota, graduated from Grinnell College in 1969 and received his PhD from the University of Michigan in 1974. He began his academic career at the University of Illinois at Urbana-Champaign in 1974. From 1976–83, he was a Member of the Technical Staff in the Department of Economics of Bell Laboratories. He returned to the University of Illinois in 1983, where he was Professor of Economics and Statistics until 2018. Since 2018 he has been Honorary Professor of Economics at University College London. Much of his research has focused on quantile regression methods, which were introduced in joint work with Gib Bassett in the late 1970s. He received the Emanuel and Carol Parzen Prize for Statistical Innovation in 2010. Since 2010 his research has diversified, including work on shape-constrained density estimation and total variation regression smoothing. His most recent work has focused on nonparametric maximum likelihood methods for mixture models.

Roger will give this Medallion Lecture at the Joint Statistical Meetings in Philadelphia in August.

#### Some Unlikely Likelihoods

Statistics as a discipline has oscillated between the poles of parametric and distributional inference from its infancy. Before 1900, emphasis was placed on estimating means and medians without much concern about what underlying distributions might justify such estimates. Karl Pearson felt compelled to expand upon the dominant model of the (Gaussian) law of errors with his family of (momentary) densities. Fisher pointed out that inference about such densities should be based on their parameters. Modern statistical theory, having extracted most of the blood from the parametric turnip, has once again turned back toward models of distributions and nonparametrics. And yet, Fisher's beloved method of maximum likelihood for parametric models has proven to be a vital tool for nonparametric data analysis.

Maximum likelihood has been an especially fruitful approach for shape-constrained density estimation, as foretold by Grenander. Curiously, attempts to extend log-concave density estimation to algebraic tailed distributions by maximum likelihood were unmanageable, and modified objectives based on Rényi entropies have proven to be more convenient. Closely related nonparametric maximum likelihood methods are also effective for mixture models, as suggested initially by Robbins. The nonparametric MLE of Kiefer and Wolfowitz now plays an important role in empirical Bayes compound decision theory. Random coefficient binary response models provide a novel illustration. Relying on modern convex optimization, these methods share the advantage that they are free of pesky tuning parameters that plague other nonparametric estimation methods. Many important questions remain open, notably, the asymptotic behavior of profile likelihood estimators.

Finally, there may be some remarks that attempt to clarify the role of likelihood in the general formulation of quantile regression, noting that, in effect, appropriately weighted estimation of the quantile regression process may be regarded as a scheme for maximum likelihood estimation of the entire conditional distribution. Again, there are many interesting open problems including those of computational implementation and adaptive efficiency.

# **IMS Medallion Lecture: Paul Rosenbaum**

Paul R. Rosenbaum is the Robert G. Putzel Professor in the Department of Statistics at the Wharton School of the University of Pennsylvania, where he has worked since 1986. Prior to this, he worked at the US Environmental Protection Agency, the University of Wisconsin–Madison, and the Educational Testing Service. He is the author of three books: *Observational Studies* (Springer, 2nd edition, 2002), *Design of Observational Studies* (Springer, 2010), and *Observation and Experiment: An Introduction to Causal Inference* (Harvard University Press, 2017). Paul received his PhD in Statistics from Harvard University in 1980, with Donald B. Rubin and Arthur Dempster as thesis advisors. He delivered the Fisher Lecture and received the R. A. Fisher Award in 2019, and the George W. Snedecor Award in 2003, both from COPSS.



Paul Rosenbaum

Paul's Medallion Lecture will be given at the Joint Statistical Meetings in Philadelphia in August.

#### **Replication and Evidence Factors in Observational Studies**

Observational studies are often biased by the failure to adjust for a covariate that was not measured. A series of studies may replicate an association because the bias that produced this association has been replicated, not because a treatment effect has been demonstrated. If a limited sample size is not the major problem in an observational study, then an increase in sample size is not the solution. To be of value, a replication should remove, or reduce, or at least vary a potential source of bias that resulted in uncertainty in earlier studies.

Having defined the goal of replication in this way, we may ask: Can one observational study replicate itself? Can it provide two statistically independent tests of one hypothesis about treatment effects such that the two tests are susceptible to different unobserved biases? Can the sensitivity analyses for these two tests be combined using meta-analytic techniques as if they came from unrelated studies, despite using the same data twice? Can such a combination provide stronger evidence that an association is an effect caused by the treatment, not a bias in who was selected for treatment? When this is possible, the study is said to possess two evidence factors. A study has two evidence factors if it permits two (essentially) statistically independent analyses using the same data that are affected by different types of unmeasured biases. More specifically, the sensitivity analyses for the two factors must be capable of combination as if they came from different unrelated studies, despite using the data twice. This latter condition is in some ways stronger than statistical independence, in other ways weaker.

The talk is divided into three parts:

- (i) a brief, largely conceptual discussion of replication in observational studies;
- (ii) a longer, more technical discussion with results about and practical examples of evidence factors,
- (iii) consideration of algorithmic aspects of building study designs with evidence factors.

#### Some References about Replication and Evidence Factors

- Rosenbaum, P.R. (2001). Replicating effects and biases. *American Statistician*, 55(3), 223–227.
- Rosenbaum, P.R. (2010). Evidence factors in observational studies. *Biometrika*, 97(2), 333-345.
- Rosenbaum, P.R. (2011). Some approximate evidence factors in observational studies. *Journal of the American Statistical Association*, 106(493),285–295.
- Zhang, K., Small, D. S., Lorch, S., Srinivas, S., Rosenbaum, P.R. (2011). Using split samples and evidence factors in an observational study of neonatal outcomes. *Journal of the American Statistical Association*, 106(494), 511–524.
- Zubizarreta, J. R., Neuman, M., Silber, J. H., Rosenbaum, P. R. (2012). Contrasting evidence within and between institutions that provide treatment in an observational study of alternate forms of anesthesia. *Journal of the American Statistical Association*, 107(499), 901–915.
- Rosenbaum, P.R. (2015). How to see more in observational studies: Some new quasi-experimental devices. *Annual Review of Statistics and Its Application*, 2, 21–48.

Rosenbaum, P.R. (2017). The general structure of evidence factors in observational studies. *Statistical Science*, 32(4), 514–530.

- Karmakar, B., Small, D. S., Rosenbaum, P. R. (2019). Using approximation algorithms to build evidence factors and related designs for observational studies. *Journal of Computational and Graphical Statistics*, 28(3), 698–709.
- Karmakar, B., Small, D. S., Rosenbaum, P. R. (2020). Using evidence factors to clarify exposure biomarkers. *American Journal* of *Epidemiology*, to appear.

# IMS/Bernoulli Society Doob Lecture: Nicolas Curien



Nicolas Curien

Nicolas Curien is Professor, since 2014, in the Mathematics Institute at Université Paris-Saclay, France. He specializes in probability theory and, more precisely, in random two-dimensional geometry. He received his PhD at Orsay (now Paris-Saclay) in 2011 and then spent two years at Sorbonne University as a CNRS researcher. His main focus is on the study of random discrete structures (such as random trees, random planar maps or graphs) and their scaling limits. He was awarded the Rollo-Davidson prize in 2015 and was a lecturer at the Saint-Flour summer school in 2019.

Nicolas will give this Doob Lecture at the World Congress in Seoul, in August.

#### Random planar maps and Doob's h-transformation

The *h*-transformation in the sense of Doob is now classic in probability theory and consists in twisting the transition probabilities p(x,y) of a Markov process using a harmonic function h(x) as follows

$$q(x,y):=\frac{h(y)}{h(x)}p(x,y).$$

It can be used, for example, to quickly define processes conditioned on event of zero probability, such as conditioning a Brownian motion to stay positive or conditioning an oscillating random walk to stay positive (Bertoin & Doney, 1992). On the other hand, enumeration of planar maps (graphs drawn on the plane) is a well-established theory in combinatorics and in recent years it became apparent that some functions appearing in the asymptotic enumeration of maps (with a boundary) turned out to be harmonic functions for certain random walks on Z. Using Markovian explorations of random planar maps known as the *peeling processes* (Angel 2003, Budd 2015), those functions can be used via *h*-transformations to define random planar maps conditioned on events of zero probability such as being infinite, having About the Doob Lecture

The Doob lecture, given in memory of Joseph Leo "Joe" Doob (1910–2004), is generously supported by the *Illinois Journal of Mathematics*, and is a joint IMS–Bernoulli Society lecture in World Congress years (those divisible by 4); in other years the lecture is given at the Stochastic Processes and Applications conference.

Joe Doob, who was President of the IMS in 1950, was awarded the 1979 US National Medal of Science by Jimmy Carter, for "his work on probability and mathematical statistics, characterized by novel and fruitful ideas of a general character that opened new fields of study which [...] now are acclaimed worldwide." Some of the areas of mathematics to which Joe Doob made important contributions are: complex function theory, ergodic theory, martingale theory, mathematical statistics, Markov processes, the general theory of stochastic processes, and probabilistic potential theory. He wrote the 1953 classic, *Stochastic Processes*, and two others. His other honors included being elected a member of the US National Academy

an infinite general or simple boundary or being of "hyperbolic" type.

In this talk, we will survey some of the recent developments in the construction and the study of infinite random planar maps using the peeling process and its connections with random walks.



of Sciences and of the American Academy of the Arts and Sciences. He was President of the American Mathematical Society and was awarded the AMS Career Prize in 1984. His obituary in *The Times* noted that as well as his work, Joe would also be remembered "for his personal qualities: bluff kindness, good humour and wit, and a complete lack of personal ambition, combined with gritty dedication to a highly technical field of study over a working lifetime of 70-odd years." *Joe Doob [on the right] receives the National* 

Joe Doob [on the right] receives the National Medal of Science from US President Jimmy Carter

# **International Data Science in Schools Project**

Nick Fisher, University of Sydney & ValueMetrics Australia, is coordinator of IDSSP, an international cooperative venture to create a framework for teaching Data Science in high schools, and to teach their teachers how to deliver the curriculum. He explains: What *is* Data Science? Twenty years ago, the term 'Data Science' hardly existed at all, although there is strong evidence that the subject itself has been in existence for over half a century (*cf.* David Donoho's excellent article "50 Years of Data Science" in the *Journal of Computational and Graphical Statistics*, **26**:4, 745-766).

These days, it is impossible to avoid. If there was ever a STEM bandwagon, this is it. University departments and schools are being re-badged, all manner of Data Science courses for students

at almost any stage of learning have sprung up, a myriad journals and websites have sprouted, and jobs advertising Data Science positions vastly outnumber the amount of suitably qualified people entering the workforce. In the US alone, there is evidence of a shortfall in the hundreds of thousands.

Yet there does not appear to be wide-

spread common understanding of what the

term actually means, let alone what might be taught in a course purporting to be a general introduction to the subject.

So what is Data Science, when should we learn it, and who's going to teach it?

At its simplest, Data Science is a life skill—the science of learning from data—and something that every child needs to know about to help them cope with the vagaries of life.

So, there is a pressing need to equip school students with this skill. And of course, that means there is a similarly pressing need to teach teachers how to teach it.

The International Data Science in Schools Project (IDSSP: www.idssp.org) was set up early in 2018 to tackle this issue.

IDSSP is an international collaborative activity involving leading computer scientists, statistical scientists, curriculum experts and teachers from Australia, Canada, England, Germany, Holland, New Zealand and the USA and supported by national and international professional societies, groups and companies.

The project has two objectives.

Firstly, to ensure that school students acquire a sufficient understanding and appreciation of how data can be acquired and used to make decisions so that they can make informed judgments in their daily lives, as students and then as adults. In particular, we envisage future generations of lawyers, journalists, historians, and many others, leaving school with a basic understanding of how to work with data to make decisions in the presence of uncertainty, and how to interpret quantitative information presented to them in the course of their professional and personal activities.

Secondly, it aims to instil in more scientifically able school students sufficient interest and enthusiasm for Data Science that they will seek to pursue tertiary studies with a view to making a career in the area.

In both cases, we want to teach people how to learn from data.

Our goal is to provide the content for a pre-calculus course in Data Science that is fun to learn and fun to teach. A total of some 240 hours of instruction is envisaged. As a parallel development we aim to devise a program will enable teachers from a wide variety of

> backgrounds—either mathematics teachers or from any other discipline that involves data—to learn to present such a course well.

Phase I of work has now been completed and made freely available. This was an entirely voluntary effort. Two curriculum frameworks (see http://www.idssp.org/ pages/framework.html)have been created

to support development of a pre-calculus course in Data Science that is rigorous, engaging and accessible to all students, and a joy to teach.

It is envisaged the material will be used not just in schools, but also as a valuable source of information for Data Science courses in community colleges and universities and for private study.

Now we are pursuing Phase 2:

- to develop the resources to support courses based on the curriculum frameworks; and
- to devise and implement a course aimed at prospective teachers of Data Science.

The deliverables would include:

- Excellent course materials and resources to support delivery of pre-calculus *Introductory Data Science* courses in a variety of modes, so that it would be fun to teach Data Science and fun to learn it.
- A course and assessment process to teach teachers from a variety of backgrounds how to teach Data Science well.

As with Phase 1, all materials will be made publicly available.

It is intended that Phase 2 be carried out professionally: contributors will be recompensed for their time, professional production of high quality materials, and with contracted project management.

Interested to find out more? Take a look at the website www. idssp.org or email idssp.info@gmail.com.

Our goal is to provide the content for a pre-calculus course in Data Science that is fun to learn and fun to teach.

# **Student Puzzle Corner 28**

The problem framed this time is at least partially a classic problem in geometry. You can find a lot in the literature about where this general problem arises in numerous fields of application. Some previous exposure to spherical geometry would probably be helpful, particularly for part (e). Here is the problem; try to do as many parts as you can.

a) Suppose P,Q are two points on the n-dimensional unit sphere  $S^{n-1} = \{x \in \mathbb{R}^n : ||x|| = 1\}$ . Find the density of the Euclidean distance PQ between PQ. Draw a nice picture of simulations of the PQ distance in dimensions n = 2, 3, 5, 10, 25.

b) Find the expectation of the Euclidean distance PQ between P,Q.

- c) Is there something very interesting about the fourth moment of the distance PQ when the Deadline: April 25, 2020 dimension *n* is even? What is it?
  - d) Derive an asymptotic expansion for the expected distance PQ of part (b). Explain intuitively why the leading term in the asymptotic expansion is what it is.

e) Suppose n = 3, and that you have to travel the shortest path from P to Q along the surface of  $S^{n-1} = S^2$ . What would be the expected distance travelled? Is this answer larger than the answer to part (b)?

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

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

The Puzzle Editor's decision is final.

### Solution to Puzzle 27

12 · IMS Bulletin

Contributing Editor Anirban DasGupta writes on the previous problem, which was about the German tank problem:

Denote  $X_{(n)}, X_{(1)}$  by Y, X respectively. By using the tailsum formula, one has  $E(Y) = \theta + N - c(n, N)$ , where

$$c(n,N) = \frac{N-n}{n+1}.$$

Thus, by symmetry, we get a linear unbiased estimate aW + b, where a = a(n) > 1, b = -1.

This is indeed the UMVUE of N by the Lehmann–Scheffé theorem, provided the parameters  $\theta$ , *N* are arbitrary, i.e.,  $(\theta + 1, N) \in \mathcal{N} \times \mathcal{N}$ . The joint p.m.f. of (X, Y) can be written too, from which the second moment of W follows.



Thanks to Andrej Srakar, PhD student in Mathematical Statistics at the University of Ljubljana, Slovenia [pictured left], for sending in a solution.

#### 22nd Meeting of New Researchers in Statistics and Probability

The 22nd IMS New Researchers Conference will be held in Philadelphia, PA, during its usual slot right before JSM. The purpose of the conference is to promote interaction and networking among new researchers in statistics and probability and to provide them with valuable insights from leaders in the field. Hosted by the University of Pennsylvania (UPenn), it will take place from July 29–August 1, 2020.

The meeting is graciously supported by the National Science Foundation. The selected applicants will be asked to pay a registration fee to confirm their participation. However, contingent on the availability of funds, support to defray travel and housing costs will be offered, though we invite participants to seek partial funding of their own.

Invited speakers include IMS President Susan Murphy, Harvard University, and IMS President-Elect Regina Liu, Rutgers University.

Applications are invited now, from anyone who received a PhD in or after 2015, or expects to by the end of 2020. The number of participants is limited. The deadline for receipt of applications is March 22, 2020. By April 6, applicants will be notified of whether they have been selected to attend. Women and minorities are especially encouraged to apply. For more information, please see http://groups.imstat.org/newresearchers/conferences/nrc.html

# Recent papers: two co-sponsored journals

# **Electronic Journal of Statistics**

The *Electronic Journal of Statistics* (*EJS*) publishes research articles and short notes in theoretical, computational and applied statistics. The journal is open access. Articles are refereed and are held to the same standard as articles in other IMS journals. Articles become publicly available shortly after they are accepted. Read it at **https://projecteuclid.org/euclid.ejs** 

### Volume 14: Number 1, 2020

| On the predictive potential of kernel principal components  | BEN JONES, ANDREAS ARTEMIOU, AND BING LI; 1 - 23                                  |
|---|---|
| Monotone least squares and isotonic quantiles   | ALEXANDRE MÖSCHING AND LUTZ DÜMBGEN; 24 - 49                                      |
| Non-parametric adaptive estimation of order 1 Sobol indices in stochastic models, with an application       | n to Epidemiology GWENAËLLE CASTELLAN, ANTHONY COUSIEN, VIET CHI TRAN; 50 - 81    |
| Model-based clustering with envelopes   |   |
| Nonparametric false discovery rate control for identifying simultaneous signals                             |   |
| Efficient estimation in expectile regression using envelope models  |   |
| Estimation of linear projections of non-sparse coefficients in high-dimensional regression                  |   |
| Perspective maximum likelihood-type estimation via proximal decomposition                                   |   |
| Bayesian variance estimation in the Gaussian sequence model with partial information on the means           | GIANLUCA FINOCCHIO AND JOHANNES SCHMIDT-HIEBER; 239 - 271                         |
| Assessing prediction error at interpolation and extrapolation points.                                       | ASSAF RABINOWICZ AND SAHARON ROSSET; 272 - 301                                    |
| Asymptotics and optimal bandwidth for nonparametric estimation of density level sets $\ldots \ldots \ldots$ | WANLI QIAO; 302 - 344   |
| Sparse equisigned PCA: Algorithms and performance bounds in the noisy rank-1 setting                        | ARVIND PRASADAN, RAJ RAO NADAKUDITI, AND DEBASHIS PAUL; 345 - 385                 |
| Parseval inequalities and lower bounds for variance-based sensitivity indices.                              |   |
| Recovery of simultaneous low rank and two-way sparse coefficient matrices, a nonconvex approach             |   |
| On polyhedral estimation of signals via indirect observations   | ANATOLI JUDITSKY AND ARKADI NEMIROVSKI; 458 - 502                                 |
| Gaussian field on the symmetric group: Prediction and learning  | RANÇOIS BACHOC, BAPTISTE BROTO, FABRICE GAMBOA, AND JEAN-MICHEL LOUBES; 503 - 546 |
| Drift estimation for stochastic reaction-diffusion systems  | GREGOR PASEMANN AND WILHELM STANNAT; 547 - 579                                    |
| On the Letac-Massam conjecture and existence of high dimensional Bayes estimators for graphical m           | odels EMANUEL BEN-DAVID AND BALA RAJARATNAM; 580 - 604                            |
| Generalised cepstral models for the spectrum of vector time series  |   |
| Statistical convergence of the EM algorithm on Gaussian mixture models                                      |   |
| Nonparametric confidence intervals for conditional quantiles with large-dimensional covariates              | LAURENT GARDES; 661 - 701   |
| Detection of sparse positive dependence   | ERY ARIAS-CASTRO, RONG HUANG, AND NICOLAS VERZELEN; 702 - 730                     |
| Profile likelihood bi-clustering  | CHERYL FLYNN AND PATRICK PERRY; 731 - 768   |
| Estimation of a semiparametric transformation model: A novel approach based on least squares minim          | nization  |
| The bias of isotonic regression   | RAN DAI, HYEBIN SONG, RINA FOYGEL BARBER, AND GARVESH RASKUTTI; 801 - 834         |
| Modal clustering asymptotics with applications to bandwidth selection                                       | ALESSANDRO CASA, JOSÉ E. CHACÓN, AND GIOVANNA MENARDI; 835 - 856                  |
| On a Metropolis-Hastings importance sampling estimator  | DANIEL RUDOLF AND BJÖRN SPRUNGK; 857 - 889  |
| Reduction problems and deformation approaches to nonstationary covariance functions over spheres            | EMILIO PORCU, RACHID SENOUSSI, ENNER MENDOZA, MORENO BEVILACQUA; 890 - 916        |

# **Statistics Surveys**

*Statistics Surveys* publishes survey articles in theoretical, computational, and applied statistics. The style may range from reviews of recent research to graduate textbook exposition. The essential requirements are a well specified topic and target audience, together with clear exposition. Read it at **https://projecteuclid.org/euclid.ssu** 

### Volume 14, 2020 (to date)

| Estimating the size of a hidden finite set: Large-sample behavior of estimators.                              | SI CHENG, DANIEL J. ECK, AND FORREST W. CRAWFORD; 1 - 31       |
|---|--|
| Flexible, boundary adapted, nonparametric methods for the estimation of univariate piecewise-smooth functions | UMBERTO AMATO, ANESTIS ANTONIADIS, AND ITALIA DE FEIS; 32 - 70 |

# OBITUARY: Colin R. Blyth

CANADIAN STATISTICIAN COLIN ROSS BLYTH died on August 22, 2019, at the age of 96. Born in Guelph, Ontario, on October 24, 1922, he studied at Queen's University, Kingston (BA, 1944), the University of Toronto (MA, 1946), and the University of California at Berkeley (PhD, 1950), where he was Erich Lehmann's first PhD student. During his career, he held positions at the University of Illinois at Urbana-Champaign (1950–74) and at Queen's University (1971–87) in addition to being a statistical consultant for the Illinois State Geological Survey (1952–55).

Colin Blyth contributed to the development of classical mathematical statistics with over 30 research articles in mathematics, statistics, and geology journals. His thesis was concerned with minimax decision procedures. He was the first to show that the average of a random sample of Gaussian variables is admissible and the method he used to prove this result bears his name. During his career, he wrote papers on topics such as Stirling's approximation, Simpson's paradox, Cramér–Rao type inequalities, convolutions of Cauchy distributions, Neyman shortest unbiased confidence intervals, the relative efficiency of tests, as well as hypothesis estimation and acceptability profiles for two-by-two contingency tables.

At Urbana-Champaign, he supervised five PhD students (Glen Meeden, Wayne Nelson, Raman N. Pillai, Robert G. Staudte and Madanlal T. Wasan), who had successful careers and gave him over 30 academic descendants.

In recognition of his contributions to the profession, including as an associate editor for *JASA* (1967–71), Colin was made a fellow of the Institute of Mathematical Statistics (1974) and the American Statistical Association (1975). He was also granted membership to Pi Mu Epsilon and to Sigma Xi, both in 1949.

In his retirement, Colin combined his passions for music and languages to write *Gaelic Names of Pipe Tunes* in 1994; he also edited *Sullivan Ross, Volume 1: A Restored Edition*, published in 2010, which provides a unique window on the (bagpipe and violin) music of rural Ontario from 1850



Photograph of Colin R. Blyth taken while he was attending the "Struwwelpeter Reconsidered" conference held at the University of Minnesota, November 9–11, 1995.

to 1900. Moreover, he composed many poems (e.g., "Kate O'Shanter," published in *Scottish Field* in 1993 under the penname "Seanair," Gaelic for Grandfather) and wrote verse translations of mid-19th century German children's classics: *Struwwelpeter Tales of Hoffmann* (1995), *Struwwelpeter 2000* (2000) and *Max & Moritz 2000* (2006).

Colin is survived by his wife of 64 years, Valerie Thompson, and their children, Mary Alice Snetsinger (Rob), Georgina Roche, Colin M. (Trish), Heather (Rob Smith), Alec (Lisa), and Donald, and by nine grandchildren. He will be sorely missed. *Christian Genest, McGill University* 



Janos Galambos

# **OBITUARY: Janos Galambos** 1940–2019

PROFESSOR JANOS GALAMBOS, a well respected and influential mathematician, passed away on December 19, 2019 at the age of 79. Galambos is survived by his wife, Éva Galambos.

Janos Galambos was born in Zirc, Hungary, on September 1, 1940. He entered Eötvös Loránd University in Budapest in 1958 and graduated with a PhD in 1963, under the supervision of Alfréd Rényi.

He was Assistant Professor at Eötvös Loránd University from 1964–65, a Lecturer at the University of Ghana, Legon, from 1965–69, and a Lecturer at the University of Ibadan, Nigeria, from 1969–70. In 1970, he joined the faculty at Temple University, Philadelphia,

#### Janos Galambos, 1940–2019

Continued from previous page

and remained there until his retirement in 2012.

Galambos was a devoted researcher whose work spanned a variety of topics in mathematics. He contributed to the theoretical development and applications of probability, number theory, and statistics.

His dissertation dealt with probabilistic inequalities, and he continued in this direction for many years. He mainly worked on Bonferroni-type inequalities - inequalities that are valid in any probability space—and used them as a unifying tool to solve problems in different branches of mathematics. His first paper in the area focused on graph-sieve inequalities. These inequalities originated in number theory, and they were later applied to probability. They turned out to be very important in the derivation of limit theorems in combinatorics, extremes, random subsets, etc.: see, for example, "The role of graph theory in some sieve arguments of probability theory," In: Graph Theory and Applications (ed.: Y. Alavi et al), Lecture Notes in Math., Springer, 303 (1972).

During his stay at the University of Ghana, Galambos attended seminar lectures given by Sir A. Oppenheim on representation of real numbers by infinite series. He became interested in probabilistic aspects of these series and developed a general theory which culminated in the monograph, *Representations of Real Numbers by Infinite Series*, Springer-Verlag, *Lecture Notes in Mathematics Series*, 1976. This monograph ends with a collection of open problems that have inspired the work of many mathematicians since.

Galambos developed a probabilistic framework to study arithmetical functions

arising in number theory. He was able to obtain simple, probabilistic proofs of well-known theorems of Erdős, Kubilius, and Delange on the asymptotic distribution of arithmetical functions. Another major contribution was the extension of a theorem of Bakštys dealing with the limiting distribution of strongly multiplicative functions, "On the distribution of strongly multiplicative functions," *Bulletin of the London Mathematical Society* **3** (1971).

His work on multiplicative functions via probability theory motivated the study of a general probabilistic theory of infinite products of random variables. Later in his career he worked on the development of this theory and used it to study the limiting distributions of multiplicative functions and to develop characterization results based on products of random variables.

In statistics, he made seminal contributions to characterization using properties of order statistics and extreme value theory. Galambos's contribution to extremes has been important not only for its theoretical development, but also for its impact on changing the attitude of the mathematical community towards the subject. His book on extremes, The Asymptotic Theory of Extreme Order Statistics (Wiley, New York, 1978), was the first book on the subject written for mathematicians. The book was translated into Russian in1984, and its second edition was translated into Chinese in 2001.Galambos started his long sequel of papers on extremes by characterizing the limit distribution for the maximum(under proper normalization) of a wide class of dependent random variables. He also contributed to the development of the theory of extremes for i.i.d. random vectors.

Galambos was influential in many other aspects of statistics. His joint work with H.A. David is considered to be one of the fundamental papers on the theory of concomitants: "The Asymptotic Theory of Concomitants of Order Statistics," (with H.A. David), *Journal of Applied Probability* **11** (1974). He developed nonparametric tests for extreme value distributions and was instrumental in the development of a methodology for the proper use of extreme value theory in applications.

Galambos wrote over 130 papers and 8 books, and he was also an editor for several books. The importance of his work has been recognized throughout the world: he was a frequent speaker at international conferences and traveled widely as a guest of universities and scientific institutions. He was an elected member of the Hungarian Academy of Science, the International Statistical Institute and the Spanish Royal Academy of Engineers, and a Fellow of the Institute of Mathematical Statistics.

Professor Janos Galambos will be remembered around the world, not only for his profound work in mathematics, but also for his role as an inspiring teacher and mentor.

Italo Simonelli, Duke Kunshan University, Kunshan, China

You can read more about the work of Professor Janos Galambos in a longer *Laudatio* that Italo Simonelli wrote: it's online at http://ac.inf.elte.hu/ Vol\_034\_2011/005.pdf

# **OBITUARY: Jonathan Cryer** 1939–2019

PROFESSOR EMERITUS JONATHAN D. CRYER passed away peacefully at his home in Iowa City on December 11, 2019, after a long battle with ampullary cancer.

Jon, who was born February 10, 1939, in Toledo, Ohio, earned his PhD in Statistics in 1966 from North Carolina State University, Durham, where he worked with M. R. Leadbetter on a topic about level crossing by normal processes, with applications to reliability. Jon was hired by the University of Iowa (UI) in 1966, joining Allen Craig, Lloyd Knowler, Bob Hogg, Byron Cosby, James Hickman, John Birch, and Tim Robertson in what was then the newly created Department of Statistics, now known as the Department of Statistics and Actuarial Science. He retired from the UI in 2001.

Jon made profound contributions to statistical education. Back in the mid-1970s, Bob Hogg, H.D. Hoover and Jon Cryer organized a two-week conference for mathematicians who were teaching statistics in small liberal arts colleges. These colleagues did not have appropriate statistical training and often relied on recommendations from professional societies in designing their own statistical courses. The conference organized by Jon and his colleagues introduced to these isolated statistics professors the importance of data analysis, robustness and using computers in teaching statistics - cutting-edge ideas at that time. Their emphasis on using computers was especially forward-looking, given that computer programming required cumbersome card punching and long turnaround times from main-frame computers. It was a highly successful conference. One of the participants, Sister Paschel Hocum, wrote Jon a two-page thank you letter dated July 30, 1977, in which she said, "You and

your colleagues awakened in me feelings for statistics that I have long cherished in mathematics. It is fun, it is exciting, it is creative. Those feelings will have to be an asset when I teach Statistics second semester."

Jon was active in organizing and speaking at similar workshops for improving statistical education, from teaching AP statistics to business statistics in business schools, including the annual week-long STATS Workshop from 1993 to 1996, as well as the 1998 and 1999 Making Statistics More Effective in Schools of Business conferences.

Jon co-authored two well-known textbooks: Time Series Analysis (first edition published in 1986, and second edition in 2008, with Kung-Sik Chan being the second author) and Statistics for Business: Data Analysis and Modeling in 1991 (with Robert B. Miller). As Jon was an early advocate for the use of computers in statistical education, in the first edition of his Time Series Analysis, he integrated theory, applications and computing using Minitab to do data analysis and model fitting. His business statistics textbook took an innovative approach with a heavy emphasis on processes and quality, reflecting his broader view of statistics as "the study of processes and what [one] can learn from observing them." It emphasizes data analysis (the second edition contains a disk collecting over a hundred real data sets), experimental and survey design, while minimizing the use of formal probability theory and hypothesis testing. In a way, his books are harbingers of the trend that "books have begun illustrating statistical ideas using more and more data, whether real or realistic." 1 In addition, Jon co-authored (with Barbara F. Ryan and Brian L. Joiner) the Minitab handbook, a classic that is now in its sixth edition.



Jon Cryer in 2011

The importance placed by Jon on real data in statistical education is well illustrated by a public flier he sent to former students of his Applied Time Series class and other interested parties, on December 8, 1975. It advertised a "demonstration of various real time, continuous time, [and] time series techniques based on equipment provided by the Department of Physics. [...] With the aid of the white noise generators, sine wave generators, filters, oscilloscopes, and a Hewlett Packard 3580A Spectrum Analyzer, we will be illustrating certain concepts of time series analysis."

Jon was elected a Fellow of the American Statistical Association in 1996, in recognition of his service to the profession and his contributions to statistical education. He received a University of Iowa collegiate teaching award in 2000.

Jon was a gifted musician. He sang in his church choir his entire 53 years in Iowa City. In retirement, he played euphonium and string bass with several local musical groups in Iowa City. Jon was a passionate teacher, a gifted musician and a gentle man, who will be greatly missed.

Kung-Sik Chan, University of Iowa

<sup>1</sup> Cryer, J. D. (2002). "A Review of the Lessons Learned at the Conferences on Making Statistics More Effective in Schools and Business." In *ICOTS6: Proceedings of the 6th International Conference on Teaching Statistics*, Cape Town, South Africa.

# IMS meetings around the world

## Joint Statistical Meetings: 2020–2025

IMS sponsored meeting

#### **JSM 2020**

#### August 1–6, 2020. Philadelphia, PA, USA.

w http://ww2.amstat.org/meetings/jsm/2020/

JSM (the Joint Statistical Meetings) is the largest gathering of statisticians and data scientists held in North America. It is also

one of the broadest, with topics ranging from statistical applications to methodology and theory to the expanding boundaries of statistics, such as analytics and data science. JSM also offers a unique opportunity for statisticians in academia, industry, and government to exchange ideas and explore opportunities for collaboration.

.....

#### IMS sponsored meetings: JSM dates for 2020-2024

| IMS Annual Meeting | 2022 Joint Statistical | IMS Annual Meeting  |
|--------------------|------------------------|---------------------|
| @ JSM 2021         | Meetings               | @ JSM 2023          |
| August 7–12, 2021, | August 6–11, 2022      | August 5–10, 2023   |
| Seattle, WA        | Washington DC          | Toronto, ON, Canada |

#### IMS sponsored meeting

Bernoulli/IMS 10th World Congress in Probability and Statistics August 17–21, 2020. Seoul, South Korea w http://www.wc2020.org **REGISTRATION, HOUSING AND ABSTRACT** 

SUBMISSION ARE OPEN

Program chair: Siva Athreya; Local chair: Hee-Seok Oh. The 10th World Congress in Probability and Statistics (WC2020), jointly organized by the Bernoulli Society and IMS, will be hosted by Seoul National University. We are expecting to attract more than 900 experts from over 40 countries.

This upcoming World Congress will take a comprehensive look at the latest developments in statistics and probability as well as the current trends emerging from all associated fields. A special lecture series will document a variety of modern research topics with in-depth uses and applications of these disciplines as they relate to science, industrial innovation, and society as a whole.

#### IMS sponsored meeting

#### WNAR/IMS/JR 2020 Meeting June 14–17, 2020. Anchorage, Alaska.

#### w http://www.wnar.org/page-18098

The 2020 WNAR/IMS/JR (Japanese Region) meeting will be held in Anchorage, Alaska from Sunday, June 14 through Wednesday, June 17, 2020. The conference will be held at Hilton Anchorage in downtown Anchorage. Join us during the season of the midnight sun in Alaska's largest city, nestled between the Chugach Mountains and Cook Inlet. Anchorage is a place rich in culture, history, natural wonders, and wildlife. Explore more about Anchorage and the surrounding sites. More conference details will be provided as they become available.



@ JSM 2025 August 2-7, 2025

#### IMS sponsored meeting

**JSM 2024** 

August 3-8, 2024

Portland, Oregon

2020 IMS/ASA Spring Research Conference May 20-22, 2020

w https://files.oakland.edu/users/qu/web/

SRC2020.html The 2020 Spring Research Conference

may be available.

# **IMS Annual Meeting**

## NEW

### **Oakland University, Rochester, MI, USA**

(SRC) of the IMS and the ASA's Section on Physical and Engineering Sciences (SPES) features keynote speakers Alfred O. Hero (University of Michigan), Roshan Joseph (Georgia Institute of Technology), Agus Sudjianto (Wells Fargo) and Jeff Wu (Georgia Institute of Technology). We invite you to submit contributed papers and posters. Partial funding for students

# Nashville, TN, USA

#### IMS Annual Meeting @

JSM: Seattle, August 7-12, 2021

## 2022

2021

**IMS Annual Meeting:** TBC

JSM: Washington DC, August 6-11, 2022

### 2023

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

## 2024

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

JSM: Portland, OR, August 3-8, 2024

## At a glance:

forthcoming IMS Annual Meeting and ISM dates

### 2020

JSM: Philadelphia, August 1-6, 2020

#### IMS Annual Meeting/

**10th World Congress:** Seoul, South Korea, August 17-21, 2020

NEN

# More IMS meetings around the world

#### IMS-sponsored meeting

#### FODS-2020: ACM–IMS Foundations of Data Science Conference October 18–20, 2020. Seattle, WA, USA

#### w https://fods.acm.org

The Association for Computing Machinery (ACM) and the Institute of Mathematical Statistics (IMS) have come together to launch a conference series on the Foundations of Data Science. Our inaugural event, the ACM–IMS Interdisciplinary Summit on the Foundations of Data Science, took place in San Francisco in 2019. Starting in 2020 we will have an annual conference with refereed conference proceedings. This interdisciplinary event will bring together researchers and practitioners to address foundational data science challenges in prediction, inference, fairness, ethics and the future of data science.

Submission deadline: April 13, 2020; notification: July 15, 2020; camera-ready: August 1, 2020.

#### IMS co-sponsored meeting

### Seminar on Stochastic Processes 2020 March 4–7, 2020. Michigan State University, Lansing, MI, USA

w https://stt.natsci.msu.edu/events/ssp2020/

As well as informal presentations by conference participants, there will be plenary talks by five invited speakers: Martin Barlow (Kai Lai Chung Lecturer), Ioana Dumitriu, Martina Hofmanová, Firas Rassoul-Agha and Samy Tindel. The main conference will be held on March 5–7. On March 4, there will be a special set of tutorial lectures and discussions targeted at early-career researchers, given by René Carmona. Further information on funding and accommodations, and more details about the conference, including the online registration form, is available at the conference website above.

#### IMS co-sponsored meeting

#### Statistics in the Big Data Era May 27–29, 2020. Berkeley, CA, USA

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

#### NEW

#### IMS co-sponsored meeting Southeast Probability Conference May 11–12, 2020 Duke University, USA

w https://services.math.duke.edu/~rtd/SEPC2020/SEPC2020.html This small regional conference has been held at the Duke mathematics department seven times in the last nine years. The seven speakers are: Erik Bates (Berkeley), Julia Gaudio (MIT), Jack Hanson (City College), Jeff Kuan (Texas A&M), Oanh Nguyen (Princeton), Soledad Villar (NYU) and Matt Wascher (Wisconsin).

There will be a reception/poster session on Monday night. The meeting is supported by NSF grant DMS 2011385. Graduate students and postdocs can apply for partial support.

Like the Midwest and Northeast Probability Conferences, our goal to serve the probabilists in our region — though unlike those meetings, which attract a large number of participants from a wide area, ours is a small meeting, with a leisurely schedule that allows ample opportunity for speakers and other participants to interact. Graduate students and postdocs have a chance to learn about recent developments from experts. A poster session at the end of the first day will allow young researchers to present their work. Slides of the main lectures will be made available after the meeting on the conference web page to more broadly disseminate their content.

#### IMS co-sponsored meeting

### Mathematical Statistics and Learning June 2–5, 2020. Barcelona, Spain.

w https://www.msl2020.org

The meeting aims to bring together leading experts from diverse areas of mathematical statistics and machine learning who are interested in the mathematical foundations of our fields. The common theme of the meeting is modelling and statistical analysis of data from large complex systems, which leads to high-dimensional and structured problems. There will be four special morning lectures: Francis Bach (INRIA), Liza Levina (Michigan), Luc Devroye (McGill), and Judith Rousseau (Oxford).

#### IMS sponsored meeting

#### Bernoulli–IMS 11th World Congress in Probability and Statistics (including the 2024 IMS Annual Meeting)

#### August 12–16, 2024. Ruhr-University Bochum, Germany w TBC

The Bernoulli–IMS World Congress in Probability and Statistics is held every four years. Details to follow, but for now, please save the date!



#### IMS sponsored meeting ENAR dates, 2020–2022 March 22–25, 2020: in Nashville, TN

w www.enar.org/meetings/future.cfm The 2020 ENAR/IMS meeting will be in Nashville (and the following year in Baltimore, and then Houston in 2022). Featuring a *Fostering Diversity in Biostatistics* workshop, connecting underrepresented minority students interested in biostatistics with professional biostatisticians in academia, government and industry.

#### IMS sponsored meetings

| ENAR/IMS 2021     | ENAR/IMS 2022     |
|-------------------|-------------------|
| March 14–17, 2021 | March 27–30, 2022 |
| Baltimore, MD     | Houston, TX       |

#### IMS co-sponsored meeting

Second Workshop: Emerging Data Science Methods for Complex Biomedical and Cyber Data March 26–27, 2020 Augusta, USA

w https://www.augusta.edu/mcg/dphs/workshop2/index.php The workshop features several eminent speakers who are pioneers in the fields of data science including big data analytics. The workshop aims to foster the collaborative research between data science/ statistics and other disciplinary science for the purpose of meeting the very hardest and most important data and model-driven scientific challenges, and for enhancing the much needed skills of the next generation workforce. We have witnessed the explosion of complex and big data from various disciplines, social media, cyber traffic, and environment surrounding us in the recent decade. Data scientists and statisticians are blessed with such variety of data that they have never seen before, yet are also facing many challenges because of the complexity and massiveness of such data. The goal of this workshop fits into the society's demand of fostering collaborative research between data science/statistics and other disciplines in science for the purpose of meeting the very hardest and most important data and model-driven scientific challenges. The Workshop participants will learn statistical and data science methods to handle the enormously complex biomedical and cyber science data, and help them develop analytical thinking, statistical reasoning, communication skills and creativity.

#### IMS co-sponsored meeting

#### Frontier Probability Days May 8–10, 2020. Las Vegas, Nevada, USA

w http://lechen.faculty.unlv.edu/FPD20/ Frontier Probability Days 2020 (FPD'20) is a regional workshop, taking place at the University of Nevada, Las Vegas. Its purpose is to bring together mathematicians, both regionally and globally, who have an interest in probability and its applications. FPD aims to complement other regional conferences in Probability that are held annually elsewhere in the US.

If you would like to participate and/ or speak at the conference, please fill out a registration form on or before **April 19**, **2020**. Registration is required but is free. To be considered for **financial support**, fill out a registration form by March 22: see the website for information.

NEW

### The 8th Workshop on Biostatistics and Bioinformatics

May 8–10, 2020. Atlanta, GA, USA w https://math.gsu.edu/yichuan/ 2020Workshop/

Biostatistics and Bioinformatics have played important roles in scientific research fields in recent years. The workshop will provide the opportunity for faculty and graduate students to meet top researchers, identify important directions for future research and facilitate research collaborations. The keynote speaker is **Nilanjan Chatterjee** (Johns Hopkins University Bloomberg School of Public Health and Johns Hopkins School of Medicine). There will be invited talks by distinguished researchers, and a poster session by young researchers and graduate students. Partial travel awards available.

#### IMS co-sponsored meeting

#### Conference on New Developments in Probability May 18–20, 2020 New Orleans, USA

New Oriedits, USA

w https://sse.tulane.edu/math/conferences-workshops/cndp2020 The Conference on New Developments in Probability (CNDP) 2020 is the second in a conference series hosted by Women in Probability. These conferences are intended to run every four years and are devoted to current significant research in probability theory, which this year will include topics in interacting particle systems, stochastic dynamical systems, machine learning, information theory, diffusions in degenerate geometries, mathematical physics, and mathematical biology. There will also be a special pedagogical lecture by Robin Pemantle on teaching applied probability at the graduate level.

#### IMS sponsored meeting

#### IMS Asia Pacific Rim Meeting 2021 January 5–8, 2021 Melbourne, Australia

#### w http://ims-aprm2021.com/

The sixth meeting of the Institute of Mathematical Statistics Asia Pacific Rim Meeting (IMS-APRM) will provide an excellent worldwide forum for scientific communications and collaborations for researchers in Asia and the Pacific Rim, and promote collaborations between researchers in this area and other parts of the world.



# Other meetings and events around the world

NEW

NEW

#### Reproducibility in Data Science: Statistical Methods and Applications

#### July 6–17, 2020. Lake Como, Italy

w http://bocconi2020.lakecomoschool.org

The Bocconi Summer School in Advanced Statistics and Probability, on "Reproducibility in Data Science: statistical methods and applications", will take place July 6–17, 2020, at Villa del Grumello, Lake Como, Italy. The 2020 Edition of the school is organized by Bocconi University, Milan, in collaboration with the University of Oxford and Imperial College London. The instructor is **Chiara Sabatti** (Stanford University, USA), with tutorials given by Stephen Bates (Stanford University, USA) and Matteo Sesia (Stanford University, USA).

The two-week summer school is open to all interested researchers, but is especially designed for PhD students. The deadline for applications is April 7, 2020.

For more information, contact BBS.statistics@unibocconi.it, or sonia.petrone@unibocconi.it

#### 76th Annual Deming Conference on Applied Statistics December 7–11, 2020 Atlantic City, USA

#### w https://demingconference.org

The 76th Annual Deming Conference on Applied Statistics will be held from December 7 to 9, 2020, followed by three parallel twoday short courses on December 10 and 11, at the state-of-the-art Tropicana Casino and Resort, Havana Tower, Atlantic City, NJ.

The purpose of the conference is to provide a learning experience on recent developments in statistical methodologies in biopharmaceutical applications. The conference is composed of twelve three-hour tutorials on current topics in applied biopharmaceutical statistic and FDA regulations, two one-hour distinguished keynotes on Monday and Tuesday, and a special session on Wednesday. The books, on which these sessions are based, are available for sale at an approximately 40% discount. Attendees will receive program proceedings of the presentations. There will also be poster sessions. The conference is sponsored by the ASA Biopharmaceutical Section. Walter Young has chaired this conference for 50 consecutive years.

The full program, as well as a downloadable printed version, will be available on our website by June 1st, and online registration will open in August.

For more information about the conference, please email Din Chen dinchen@email.unc.edu or visit the website.

#### NOTE DATES: this conference was previously announced

*as July 8–11, 2021, but the correct dates are July 15–18, 2021.* STATISTICS 2021 CANADA: 6th Canadian Conference in Applied Statistics

#### July 15–18, 2021. Montreal, Quebec, Canada

w http://www.concordia.ca/artsci/events/statistics-2021.html The Department of Mathematics & Statistics and Department of Supply Chain & Business Technology Management of Concordia University will be hosting this conference, which is dedicated to all areas of statistical sciences. In addition to the traditional theoretical and applied areas, interdisciplinary research would be encouraged and promoted including major theme areas of Applied Statistics, Big Data, Bioinformatics, Biostatistics, Computational Statistics, Data Mining, Data Science, Demography, Econometrics, Health Statistics, Model Selection and Validation, Social Statistics, Supply Chain, Survival Analysis, and Stochastic Methods in Ecology, Finance and Engineering.

# Dimensionality Reduction and Inference in High-Dimensional Time Series

#### June 2–3, 2020

#### Maastricht, The Netherlands

#### w http://sbe.maastrichtuniversity.nl/hdts2020/

This two-day workshop, at Maastricht University, The Netherlands, aims to provide a platform for exchanging and discussing the latest developments in econometrics and statistics on topics related to dimensionality reduction and inference in high-dimensional time series, including (but not limited to) issues related to post-selection inference, statistical learning, penalized regression methods, factor models and common features.

# 11th IMA International Conference on Modelling in Industrial Maintenance and Reliability (MIMAR) July 14–16, 2020

#### Nottingham, UK

#### w https://tinyurl.com/IMAMIMAR

The 11th International Conference on Modelling in Industrial Maintenance and Reliability (MIMAR) is the premier maintenance and reliability modelling conference in the UK and builds upon a very successful series of previous conferences. It is an excellent international forum for disseminating information on the state-ofthe-art research, theories and practices in maintenance and reliability modelling and offers a platform for connecting researchers and practitioners from around the world.

NEW

NEW

2020 ASA Biopharmaceutical Section Regulatory-Industry Statistics Workshop (Biopharm 2020) September 23–25, 2020 Rockville, MD, USA

w https://ww2.amstat.org/meetings/ biop/2020/

Biopharm 2020 is sponsored by the ASA Biopharmaceutical Section in cooperation with the FDA Statistical Association. The conference lasts two days, with invited sessions co-chaired by statisticians from industry, academia, and the FDA. In addition, short courses on related topics are offered the day prior to the workshop. The workshop draws hundreds of attendees annually and is the most relevant conference for statistical practitioners in the biopharmaceutical arena. So, mark your calendar for **June 11**, when **Early Bird Registration** begins.

*Poster and roundtable proposal submission is open!* As a poster presenter or roundtable discussion leader, you have a valuable opportunity to share and refine your ideas with the best in the field, including statistical practitioners in the biopharmaceutical arena from industry, academia, and the FDA. Roundtable discussion topic proposals are due March 27. Poster proposals are due April 15.

#### BFF7: The 7th Bayes, Fiducial and Frequentist Statistics Conference May 6–8, 2020. Toronto, Canada

w http://www.fields.utoronto.ca/ activities/19-20/BFF7

The Statistical Sciences Department at the University of Toronto is hosting BFF7, the 7th Bayes, Fiducial and Frequentist Statistics Conference.

(See the poster on page 7.)

### ISBA 2020: World Meeting of the International Society for Bayesian Analysis June 28–July 3, 2020

### Kunming, China

#### w https://bayesian.org/isba2020-home/

The ISBA World Meeting is the biennial flagship meeting of the International Society of Bayesian Analysis. ISBA 2020 will feature Foundational Lectures on Bayesian Statistics, Keynote Talks, Invited Talks, as well as Contributed Talks and Posters covering the latest advancement in Bayesian theory, methodology, computation, and applications. Short courses on various topics in Bayesian statistics will be offered on Sunday June 28, 2020, and the main conference is from June 29, 2020 to July 3, 2020.

Foundational Lectures will be given by: Alicia Carriquiry (Iowa State University), Wesley Johnson (University of California Irvine), Xiao-Li Meng (Harvard University), and Fernando Quintana (Pontificia Universidad Católica de Chile). The Keynote Lecturers are Ismael Castillo (Sorbonne University), Leonhard Held (University of Zurich), Amy Herring (Duke University), and Suchi Saria (Johns Hopkins University).

The Bruno de Finetti Lecturer is James Berger (Duke University), with discussants Luis Pericchi (University of Puerto Rico) and Igor Pruenster (Bocconi University).

The Susie Bayarri Lecturer is Tamara Broderick (Massachusetts Institute of Technology), with discussants Sylvia Richardson (University of Cambridge) and David Dunson (Duke University).

#### SLDS 2020 May 27–29, 2020 Irvine, USA

NEW

#### w https://asaslds.github.io/SLDS2020/

SLDS conferences are bi-annual conferences organized by ASA's section on Statistical Learning and Data Science. SLDS 2020 (asaslds.github.io/SLDS2020/) is again a joint event with the Section on Nonparametric Statistics of ASA, which will be held in Irvine, California on May 27–29, 2020.

SLDS conferences bring together the vibrant community at the interface of Statistics and Data Science, with a lot of networking and learning opportunities. SLDS 2020 will also feature keynote talks from Deepak Agarwal (LinkedIn), Regina Liu (Rutgers) and Jane-Ling Wang (UC Davis). There will be more than 47 invited sessions organized by the SLDS 2020 program committee. There will also be Invited short courses (May 26, 2020) on Deep Learning, AI and Precision Medicine, offered by Annie Qu (UC Irvine), Xiao Wang (Purdue), Edgar Dobriban (Penn), Haoda Fu (Eli Lilly). Seats are limited. Lunch panels will offer opportunities for discussing challenges and opportunities faced by statisticians in light of data science.

Registration and Hotel reservation with a group rate are both open. We reserved a limited block of rooms at the conference venue, Hyatt Regency Newport Beach. We look forward to seeing you all in Irvine!

#### 40th International Symposium on Forecasting July 5–8, 2020 Rio de Janeiro, Brazil

w https://isf.forecasters.org/ The International Symposium on Forecasting (ISF) is the premier forecasting conference, attracting the world's leading forecasting researchers, practitioners, and students. Through a combination of keynote speaker presentations, academic sessions, workshops, and social programs, the ISF provides many excellent opportunities for networking, learning, and fun.

### SMS: Discrete Probability, Physics and Algorithms June 29–July 10, 2020 Montreal, Canada

whttp://www.msri.org/summer\_schools/925 Probability theory, statistics as well as mathematical physics have increasingly been used in computer science. The goal of this school is to provide a unique opportunity for graduate students and young researchers to developed multi-disciplinary skills in a rapidly evolving area of mathematics. Topics will fall into two main categories: problems related to spin glasses and problems related to random algorithms. Tutorials will cover spin glasses, constraint satisfiability, randomized algorithms, Monte-Carlo Markov chains and high-dimensional statistics, sparse and random graphs, computational complexity, estimation and approximation algorithms.



#### Applied Statistics 2020 (AS2020) September 20–23, 2020 Ribno (Bled), Slovenia

NEW

NEW

NEW

w http://conferences.nib.si/AS2020 The conference, organized in Ribno in the vicinity of beautiful Lake Bled, will provide an opportunity for statistics researchers, data scientists and analysts, and other professionals from various statistical and related fields to come together, present their research, and learn from each other. Crossdiscipline and applied paper submissions are especially welcome. A three-day program consists of invited paper presentations, contributed paper sections from diverse topics, and starts with a workshop. Selected papers will be published in Advances in Methodology and Statistics, a peer-reviewed journal of the Statistical Society of Slovenia. Invited Speakers: Jan Beyersmann, Ulm University, Germany; Anuška Ferligoj, University of Ljubljana, Slovenia; Richard De Veaux, Williams College, USA. Papers from diverse areas of statistics and methodology are appreciated: biostatistics and bioinformatics; data collection; data mining; data science; design of experiments; econometrics; mathematical statistics; measurement; modeling and simulation; network analysis; sampling techniques; social science methodology; statistical applications; statistics education; other areas of statistics

#### 2nd International Conference on Statistics: Theory and Applications (ICSTA'20) August 19–21, 2020 Prague, Czech Republic

w https://icsta.net/ Topics themes for ICSTA'20 include,

but are not limited, to: Applied statistics, Big data, Bioinformatics, Data mining, Medical statistics, Social statistics, Statistical software, Statistics and the environment, Statistical signal processing, Statistical methodology, Time-series Analysis

### 7th Workshop on Stochastic Methods in Game Theory July 1–8, 2020 Erice, Italy

w https://sites.google.com/view/ericesmgt2020/the-workshop

The goal of this workshop is to examine some recent developments of the interaction between stochastics and game theory. The focus will be on game theoretic models that heavily use stochastic tools and on stochastic methods that find relevant applications in game theory.

#### 7th IMA Conference on Numerical Linear Algebra and Optimization June 24–26, 2020. Birmingham, UK

w https://ima.org.uk/12530/7th-imaconference-on-numerical-linear-algebraand-optimization/

The success of modern codes for large-scale optimization is heavily dependent on the use of effective tools of numerical linear algebra. On the other hand, many problems in numerical linear algebra lead to linear, nonlinear or semidefinite optimization problems. The purpose of the conference is to bring together researchers from both communities and to find and communicate points and topics of common interest. This Conference has been organised in cooperation with the Society for Industrial and Applied Mathematics (SIAM). Conference topics include any subject that could be of interest to both communities, such as: direct and iterative methods for large sparse linear systems; eigenvalue computation and optimization; large-scale nonlinear and semidefinite programming; effect of round-off errors, stopping criteria, embedded iterative procedures; optimization issues for matrix polynomials; fast matrix computations; compressed/sparse sensing; PDE-constrained optimization; distributed computing and optimization; applications and real time optimization. See website for invited speakers and info.

NEW

# **Employment Opportunities around the world**

Australia: Sydney, NSW

University of Sydney Postdoctoral Research Associate, Mathematics and Statistics https://jobs.imstat.org/job//52408010

#### Australia: Sydney, NSW

The University of Sydney, School of Mathematics and Statistics Lecturer in Statistics https://jobs.imstat.org/job//52403482

#### Australia: Sydney, NSW

**The University of Sydney** Lecturer in Financial Mathematics https://jobs.imstat.org/job//52672090

#### Canada: Waterloo, ON

University of Waterloo Lecturer https://jobs.imstat.org/job//53064146

#### Canada: Waterloo, ON

University of Waterloo Lecturer https://jobs.imstat.org/job//53064147

#### Hong Kong: Kowloon

**City University of Hong Kong** Assistant or Associate Professor in Business Statistics https://jobs.imstat.org/job//52919397

#### Singapore:

Department of Statistics and Applied Probability, National University of Singapore Faculty https://jobs.imstat.org/job//51462934

#### Switzerland: Lausanne

**Ecole Polytechnique Fédérale de Lausanne** Postdoctoral Researcher at the Institute of Mathematics, EPFL https://jobs.imstat.org/job//52027331

#### Switzerland: Lausanne

**Ecole Polytechnique Fédérale de Lausanne** PhD Position in Statistics at EPFL https://jobs.imstat.org/job//52027160

#### United States: Berkeley, CA

**University of California, Berkeley** Assistant Teaching Professor https://jobs.imstat.org/job//50921303

#### United States: Berkeley, CA

**University of California, Berkeley** Assistant Teaching Professor https://jobs.imstat.org/job//50921304

#### United States: Berkeley, CA

University of California, Berkeley Lecturer - Department of Statistics https://jobs.imstat.org/job//53159515

United States: Berkeley, CA Department of Statistics Research Training Group Postdoctoral Scholar https://jobs.imstat.org/job//52847466

#### United States: Santa Cruz, CA University of California Santa Cruz

Assistant Teaching Professor of Statistics https://jobs.imstat.org/job//52706012

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

**Colorado State University, Department of Statistics** Department Chair https://jobs.imstat.org/job//52422328

#### United States: New Haven, CT Yale School of Public Health

Investigator Track Position in Biostatistics https://jobs.imstat.org/job//52371914

#### **United States: Grinnell, IA**

Grinnell College, Department of Mathematics and Statistics Assistant Professor of Statistics 2 Year Term (start Fall 2020) https://jobs.imstat.org/job//52997753

#### United States: Chicago, IL

University of Chicago Instructional Professor https://jobs.imstat.org/job//52998195

#### United States: Portland, OR

**Portland State University** Assistant Professor https://jobs.imstat.org/job//52730067

#### United States: Philadelphia, PA

**Temple University, Fox School of Business** Non-Tenure Track and Adjunct Faculty Positions in Statistical Science, Data Science, and Business Analytics https://jobs.imstat.org/job//51924976

### United States: Philadelphia, PA

**Temple University, Fox School of Business** Tenure Track and Tenured Positions in Statistical Science https://jobs.imstat.org/job//51924925

#### **United States: Arlington, VA**

**Biocomplexity Institute and Initiative** Research Faculty Positions in Statistical Sciences (SDAD) - Biocomplexity https://jobs.imstat.org/job//52182623

#### **United States: Seattle, WA**

# University of Washington Department of Biostatistics

Open Rank Faculty Position https://jobs.imstat.org/job//52935364

# **International Calendar of Statistical Events**

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

#### March 2020

March 4–7: Lansing, MI, USA. Seminar on Stochastic Processes 2020 w https://stt.natsci.msu.edu/events/ssp2020/

March 16–18: Knoxville, TN, USA. NIMBioS/SCMB Investigative Workshop on Quantitative Education in Life Science Graduate Programs w http://www.nimbios.org/workshops/WS\_quantedu

March 22–25: Nashville, TN, USA. ENAR Spring Meeting w http://www.enar.org/meetings/future.cfm

March 26: London, UK. 6th IMA Conference on Mathematics in Defence and Security w https://ima.org.uk/12970/6th-imaconference-on-mathematics-in-defence-and-security/

Methods for Complex Biomedical and Cyber Data w https:// www.augusta.edu/mcg/dphs/workshop2/index.php

### April 2020

April 6–9: Warsaw, Poland. Statistical Methods and Artificial Intelligence w https://sites.google.com/view/iwsmai

April 15–17: Geneva, Switzerland. Workshop on Statistical Data Editing w https://reg.unog.ch/event/31130/

April 26–28: Gainesville, FL, USA. Conference on Applied Statistics in Agriculture and Natural Resources w https:// conference.ifas.ufl.edu/applied-stats/

#### May 2020

May 6–8: Toronto, Canada. BFF7: The 7th Bayes, Fiducial and Frequentist Statistics Conference w http://www.fields. utoronto.ca/activities/19-20/BFF7

May 8–10: Las Vegas, NV, USA. Frontier Probability Days **w** http://lechen.faculty.unlv.edu/FPD20/

May 8–10: Atlanta, GA, USA. 8th Workshop on Biostatistics and Bioinformatics **w** https://math.gsu.edu/yichuan/2020Workshop/ May 11–12: Duke University, USA. Southeast Probability Conference w https://services.math.duke.edu/~rtd/ SEPC2020/SEPC2020.html

May 18–20: New Orleans, USA. Conference on New Developments in Probability w https://sse.tulane.edu/math/conferences-workshops/cndp2020

May 20–22: Rochester, MI, USA. 2020 IMS/ASA Spring Research Conference w https://files.oakland.edu/users/qu/ web/SRC2020.html

May 27–29: Berkeley, CA, USA. Statistics in the Big Data Era w https://simons.berkeley.edu/workshops/statistics-big-data-era

May 27–29: Irvine, USA. SLDS 2020 w https://asaslds. github.io/SLDS2020/

May 31–June 3: Ottawa, ON, Canada. 2020 SSC Annual Meeting w https://ssc.ca/en/meetings/2020-annual-meeting-ottawa

#### June 2020

June 1–26: Vancouver, BC, Canada. 2020 PIMS-CRM Probability Summer School whttp://www.math.ubc.ca/Links/ssprob20/

Reduction and Inference in High-Dimensional Time Series w http://sbe.maastrichtuniversity.nl/hdts2020/

June 2–5: Barcelona, Spain. Mathematical Statistics and Learning w https://www.msl2020.org

June 2–5: Barcelona, Spain. 6th Stochastic Modeling Techniques and Data Analysis International Conference (SMTDA2020). Also featuring Demographics 2020 Workshop w www.smtda.net

June 3–6: Pittsburgh, PA, USA. Symposium on Data Science and Statistics w https://ww2.amstat.org/meetings/sdss/2020/

June 14–17: Anchorage, Alaska, USA. WNAR/IMS/JR Meeting w https://www.wnar.org/page-18098

June 15–18: New Orleans, LA, USA. Sixth International Conference on Establishment Statistics (ICES-VI) w http://ww2.amstat.org/meetings/ices/2020/ June 15–18: Thessaloniki, Greece. IWAP 2020 (10th International Workshop on Applied Probability) w http://iwap2020.web.auth.gr

June 15–19: Paphos, Cyprus. International Symposium on Nonparametric Statistics 2020 w http://cyprusconferences.org/ isnps2020/

June 17–19: Paris Orsay, France. Mixtures, Hidden Markov Models and Clustering w https://www.math.u-psud.fr/~mhc2020/

WNAR/IMS/JR (Japanese Region) meeting w http://wnar.org/ page-18098

June 22–26: Sydney, Australia. International Statistical Ecology Conference (ISEC2020) w http://www.isec2020.org/

June 24–26: Birmingham, UK. 7th IMA Conference on Numerical Linear Algebra and Optimization w https://ima.org. uk/12530/7th-ima-conference-on-numerical-linear-algebra-andoptimization/

June 24–27: Brno, Czech Republic. Fifth International Workshop on Functional and Operatorial Statistics (IWFOS 2020) w https://iwfos2020.sci.muni.cz/

June 28–July 3: Kunming, China. ISBA 2020: World Meeting of the International Society for Bayesian Analysis w https://bayesian.org/isba2020-home/

June 29–July 3: Nový Smokovec, Slovakia. LinStat 2020 w https://linstat2020.science.upjs.sk/

June 29–July 10: Montreal, Canada. SMS: Discrete Probability, Physics and Algorithms w http://www.msri.org/ summer\_schools/925

#### **July 2020**

July 1–8: Erice, Italy. 7th Workshop on Stochastic Methods in Game Theory w https://sites.google.com/view/erice-smgt2020/ the-workshop

July 5–8: Rio de Janeiro, Brazil. 40th International Symposium on Forecasting w https://isf.forecasters.org/

July 5–11: Portoroz, Slovenia. 8th European Congress of Mathematics. w http://www.8ecm.si/

July 6–10: Gold Coast, QLD, Australia. 2020 Australian and New Zealand Statistical Conference w https://anzsc2020.com.au

July 6–10: Seoul, South Korea. 30th International Biometric Conference (IBC2020) w https://www.biometricsociety. org/2018/07/ibc-2020-seoul-preview/

July 6–17: Lake Como, Italy. Reproducibility in Data Science: Statistical Methods and Applications w http:// bocconi2020.lakecomoschool.org

Conference on Modelling in Industrial Maintenance and Reliability (MIMAR) w https://tinyurl.com/IMAMIMAR

#### August 2020

**Lims** August 1–6: Philadelphia, PA, USA. **JSM 2020 w** http://www. amstat.org/ASA/Meetings/Joint-Statistical-Meetings.aspx

Ims August 17–21: Seoul, Korea. Bernoulli/IMS World Congress in Probability and Statistics w http://www.wc2020.org

August 19–21: Prague, Czech Republic. 2nd International Conference on Statistics: Theory and Applications (ICSTA'20) w https://icsta.net/

August 23–27: Krakow, Poland. 41st Annual Conference of the ISCB w www.iscb2020.info

#### September 2020

September 9–11: Manchester, UK. 2nd IMA Conference on the Mathematics of Robotics w https://ima.org.uk/11468/imaconference-on-mathematics-of-robotics/

September 20–23: Ribno (Bled), Slovenia. Applied Statistics 2020 (AS2020) w http://conferences.nib.si/AS2020

September 23–25: Rockville, MD, USA. 2020 ASA Biopharmaceutical Section Regulatory-Industry Statistics Workshop (Biopharm 2020) w https://ww2.amstat.org/meetings/ biop/2020/

#### October 2020

October 1–3: Pittsburgh, PA, USA. Women in Statistics and Data Science Conference w https://ww2.amstat.org/meetings/ wsds/2020/

Foundations of Data Science Conference w https://fods.acm.org

# International Calendar continued

### November 2020

November 4–6: Utrecht, The Netherlands. Big Data Meets Survey Science (BigSurv20) w https://www.bigsurv20.org/

### December 2020

December 7–11: Atlantic City, USA. 76th Annual Deming Conference on Applied Statistics w https://demingconference.org

December 15–17: Manipal, India. 28th International Workshop on Matrices and Statistics (IWMS 2020) w https://carams.in/events/ international-workshop-on-matrices-and-statistics/

### January 2021

*Lims* January 5–8: Melbourne, Australia. IMS Asia Pacific Rim Meeting (IMS-APRM2021) w http://ims-aprm2021.com/

### March 2021

*w* http://www.enar.org/meetings/future.cfm

### July 2021

July 11–15: The Hague, The Netherlands. 63rd ISI World Statistics Congress 2021 w http://www.isi2021.org/

July 15–18: Montreal, Canada. Statistics 2021 Canada w https://www.concordia.ca/artsci/events/statistics-2021.html

### August 2021

August 7–12: Seattle, WA, USA. IMS Annual Meeting at JSM 2021 w http://www.amstat.org/ASA/Meetings/Joint-Statistical-Meetings.aspx

### March 2022

March 27–30: Houston, TX, USA. ENAR Spring Meeting w http://www.enar.org/meetings/future.cfm

#### June 2022

June 27–July 1: Darwin, Australia. Joint Southern Statistical Meetings 2022 (JSSM2022) w https://statsoc.org.au/event-3529236

### **July 2022**

*ims* July/August [exact dates TBC]: London, UK. IMS Annual Meeting w TBC

July 10–15: Riga, Latvia. XXXI International Biometric Conference (IBC 2022) w https://www.biometricsociety.org/meetings-events/ ibcs/

### August 2022

w http://www.amstat.org/ASA/Meetings/Joint-Statistical-Meetings.aspx

### August 2023

at JSM 2023 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 erg@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

#### 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 \$110 is added to the dues of members for each scientific journal selected (\$70 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 2020 are available to *The Annals of Applied Probability* (\$225), *The Annals of Applied Statistics* (\$225), *The Annals of Probability* (\$225), *The Annals of Statistics* (\$225), *Statistical Science* (\$185), and *IMS Bulletin* (\$115). **General subscriptions** are for libraries, institutions, and any multiple-readership use. Institutional subscriptions for 2020 are available to *The Annals of Applied Probability*, *The Annals of Applied Statistics*, *The Annals of Probability*, and *The Annals of Statistics* (each title \$525 online only/\$618 print+online), *Statistical Science* (\$301/\$352), and *IMS Bulletin* (\$146 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, 3163 Somerset Dr, Cleveland, Ohio 44122, USA. Periodicals postage paid at Cleveland, Ohio, and at additional mailing offices. Postmaster: Send address changes to 3163 Somerset Dr, Shaker Heights, OH 44122 USA or dues.subs@imstat.org. Copyright © 2020 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 \$315.00. 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 http://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 announcement 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/EPS with all fonts embedded. Email your advert to Audrey Weiss, IMS Advertising Coordinator admin@imstat.org or see https://imstat.org/advertise

|  | Dimensions: width x height                    | Rate  |
|--|---|-------|
| ⅓page  | 4.9" wide x 4" high (125 x 102 mm)            | \$275 |
| ½ page   | 7.5" wide x 4" high (190 x 102 mm)            | \$345 |
| ⅔page  | 4.9" wide x 8" high (125 x 203 mm)            | \$405 |
| Full page (to edge,<br>including 1⁄8" bleed)     | 8.75" wide x 11.25" high<br>(222 mm x 286 mm) | \$465 |
| Full page (within usual <i>Bulletin</i> margins) | 7.5" wide x 9.42" high<br>(190 mm x 239 mm)   | \$465 |

#### **Deadlines and Mail Dates for IMS Bulletin**

| lssu | le               | 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 1     | August 15   | September 1 |
| 7:   | Oct/Nov          | September 15 | October 1   | October 15  |
| 8:   | December         | November 1   | November 15 | December 1  |

# <u>the</u> March 2020

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

# **DEADLINES** submissions March 15, then May 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

# MIDS. DIDECTORISE FEBRUARY 2020 THE ANNALS of**STATISTICS**

AN OFFICIAL JOURNAL OF THE INSTITUTE OF MATHEMATICAL STATISTICS

#### Articles

| The phase transition for the existence of the maximum likelihood estimate in       high-dimensional logistic regression EMMANUEL J. CANDÈS AND PRAGYA SUR       2'         Rerandomization in 2 <sup>K</sup> factorial experiments       XINRAN LI, PENG DING AND DONALD B. RUBIN       4'         Sparse SIR: Optimal rates and adaptive estimation   | 7<br>3<br>1<br>3      |
|--|-----------------------|
| Rerandomization in 2 <sup>K</sup> factorial experiments       XINRAN LI, PENG DING AND DONALD B. RUBIN       43         Sparse SIR: Optimal rates and adaptive estimation       Kal TAN, LEI SHI       AND ZHOU YU       64         Sparse SIR: Optimal rates and adaptive estimation       Sparse SIR: Optimal rates and adaptive estimation       Kal TAN, LEI SHI       AND ZHOU YU       64         Robust sparse covariance estimation by thresholding Tyler's M-estimator       JOHN GOES, GILAD LERMAN AND BOAZ NADLER       80         Model assisted variable clustering: Minimax-optimal recovery and algorithms       FLORENTINA BUNEA, CHRISTOPHE GIRAUD, XI LUO, MARTIN ROYER       81         New G-formula for the sequential causal effect and blip effect of treatment in sequential       causal inference       XIAOQIN WANG AND LI YIN       133         Envelope-based sparse partial least squares       GUANGYU ZHU AND ZHIHUA SU       16       0       0       111         Adaptive risk bounds in univariate total variation denoising and trend filtering       ADITYANAND GUNTUBOYINA, DONOVAN LIEU, SAPXASACHI CHATTERJEE       202         Spectral and matrix factorization methods for consistent community detection in       multi-layer networks       203         Statistical inference for model parameters in stochastic gradient descent       XI CHEN, JASON D. LEE, XIN T. TONG AND YICHEN ZHANG       253         Bootstrap confidence regions based on M-estimators under nonstandard conditions       STEPHEN M. S. LEE AND P                            | 3<br>1<br>5<br>1<br>1 |
| XINRAN LI, PENG DING AND DONALD B. RUBIN<br>XINRAN LI, PENG DING AND DONALD B. RUBIN<br>AND ZHOU YU<br>Robust sparse covariance estimation by thresholding Tyler's M-estimator<br>JOHN GOES, GILAD LERMAN AND BOAZ NADLER<br>Model assisted variable clustering: Minimax-optimal recovery and algorithms<br>FLORENTINA BUNEA, CHRISTOPHE GIRAUD, XI LUO, MARTIN ROYER<br>AND NICOLAS VERZELEN<br>New <i>G</i> -formula for the sequential causal effect and blip effect of treatment in sequential<br>causal inference   | 3<br>4<br>5<br>1<br>8 |
| Sparse SIR: Optimal rates and adaptive estimation       KAI TAN, LEI SHI<br>AND ZHOU YU       64         Robust sparse covariance estimation by thresholding Tyler's M-estimator<br>JOHN GOES, GILAD LERMAN AND BOAZ NADLER       86         Model assisted variable clustering: Minimax-optimal recovery and algorithms<br>FLORENTINA BUNEA, CHRISTOPHE GIRAUD, XI LUO, MARTIN ROYER<br>AND NICOLAS VERZELEN       81         New G-formula for the sequential causal effect and blip effect of treatment in sequential<br>causal inference       11         Senvelope-based sparse partial least squares       GUANGYU ZHU AND ZHIHUA SU       16         Optimal rates for community estimation in the weighted stochastic block model       18         Adaptive risk bounds in univariate total variation denoising and trend filtering<br>ADITYANAND GUNTUBOYINA, DONOVAN LIEU, SABYASACHI CHATTERIEE       200         Spectral and matrix factorization methods for consistent community detection in<br>multi-layer networks       201         Statistical inference for model parameters in stochastic gradient descent<br>XI CHEN, JASON D. LEE, XIN T. TONG AND YICHEN ZHANG       205         Bootstrap confidence regions based on M-estimators under nonstandard conditions<br>DIMITRIS BERTSIMAS AND BART VAN PARYS       300         Testing for principal component directions under weak identifiability<br>DAVY PAINADAUENEL, JULEN REMY AND THOMAS VERDEBOUT       324   | 1<br>5<br>1           |
| Robust sparse covariance estimation by thresholding Tyler's M-estimator<br>JOHN GOES, GILAD LERMAN AND BOAZ NADLER       86         Model assisted variable clustering: Minimax-optimal recovery and algorithms<br>FLORENTINA BUNEA, CHRISTOPHE GIRAUD, XI LUO, MARTIN ROYER<br>AND NICOLAS VERZELEN       81         New G-formula for the sequential causal effect and blip effect of treatment in sequential<br>causal inference       81         Envelope-based sparse partial least squares       GUANQYU ZHU AND ZHIHUA SU       16         Optimal rates for community estimation in the weighted stochastic block model<br>MIN XU, VARUN JOG AND PO-LING LOH       182         Adaptive risk bounds in univariate total variation denoising and trend filtering<br>ADITYANAND GUNTUBOYINA, DONOVAN LIEU, SABYASACHI CHATTERJEE       202         Spectral and matrix factorization methods for consistent community detection in<br>multi-layer networks       SUBHADEEP PAUL AND YUGUO CHEN       230         Statistical inference for model parameters in stochastic gradient descent<br>XI CHEN, JASON D. LEE, XIN T. TONG AND YICHEN ZHANG       25         Bootstrap confidence regions based on M-estimators under nonstandard conditions<br>STEPHEN M. S. LEE AND PUYUDI YANG       274         Sparse high-dimensional regression: Exact scalable algorithms and phase transitions<br>DIMITRIS BERTSIMAS AND BART VAN PARYS       300         Testing for principal component directions under weak identifiability<br>DAVY PAINDAVEINE, JULIEN REMY AND THOMAS VERDEBOUT       324 | 5                     |
| Model assisted variable clustering: Minimax-optimal recovery and algorithms<br>FLORENTINA BUNEA, CHRISTOPHE GIRAUD, XI LUO, MARTIN ROYER<br>AND NICOLAS VERZELEN 11:<br>New G-formula for the sequential causal effect and blip effect of treatment in sequential<br>causal inference  | 3                     |
| FLORENTINA BUNEA, CHRISTOPHE GIRAUD, XI LUO, MARTIN ROVER<br>AND NICOLAS VERZELEN       111         New G-formula for the sequential causal effect and blip effect of treatment in sequential<br>causal inference       113         Envelope-based sparse partial least squares       XIAOQIN WANG AND LI YIN       133         Envelope-based sparse partial least squares       GUANGYU ZHU AND ZHIHUA SU       161         Optimal rates for community estimation in the weighted stochastic block model<br>MIN XU, VARUN JOG AND DO-LING LOH       182         Adaptive risk bounds in univariate total variation denoising and trend filtering<br>ADITYANAND GUNTUBOYINA, DONOVAN LIEU, SABYASACHI CHATTERJEE<br>AND BODHISATTVA SEN       200         Spectral and matrix factorization methods for consistent community detection in<br>multi-layer networks       SUBHADEEP PAUL AND YUGUO CHEN       230         Statistical inference for model parameters in stochastic gradient descent<br>XI CHEN, JASON D. LEE, XIN T. TONG AND YICHEN ZHANG       251         Bootstrap confidence regions based on M-estimators under nonstandard conditions<br>STEPHEN M. S. LEE AND PUYUDI YANG       274         Sparse high-dimensional regression: Exact scalable algorithms and phase transitions<br>DIMITRIS BERTSIMAS AND BART VAN PARYS       300         Testing for principal component directions under weak identifiability<br>DAVY PAINDAVEINE, JULIEN REMY AND THOMAS VERDEBOUT       324   | 3                     |
| New G-formula for the sequential causal effect and blip effect of treatment in sequential causal inference       133         Envelope-based sparse partial least squares       GUANGYU ZHU AND ZHIHUA SU       161         Optimal rates for community estimation in the weighted stochastic block model       MIN XU, VARUN JOG AND PO-LING LOH       182         Adaptive risk bounds in univariate total variation denoising and trend filtering       ADITYANAND GUNTUBOYINA, DONOVAN LIEU, SABYASACHI CHATTERJEE       203         Spectral and matrix factorization methods for consistent community detection in       multi-layer networks       204         Statistical inference for model parameters in stochastic gradient descent       XI CHEN, JASON D. LEE, XIN T. TONG AND YICHEN ZHANG       205         Bootstrap confidence regions based on M-estimators under nonstandard conditions       STEPHEN M. S. LEE AND PUYUDI YANG       277         Sparse high-dimensional regression: Exact scalable algorithms and phase transitions       DIMITRIS BERTSIMAS AND BART VAN PARYS       300         Testing for principal component directions under weak identifiability       DAVY PAINDAVEINE, JULEN REMY AND THOMAS VERDEBOUT       324   | 3                     |
| Envelope-based sparse partial least squares  | 3                     |
| Optimal rates for community estimation in the weighted stochastic block model<br>MIN XU, VARUN JOG AND PO-LING LOH       18:         Adaptive risk bounds in univariate total variation denoising and trend filtering<br>ADITYANAND GUNTUBOYINA, DONOVAN LIEU, SABYASACHI CHATTERJEE<br>AND BODHISATTVA SEN       20:         Spectral and matrix factorization methods for consistent community detection in<br>multi-layer networksSUBHADEEP PAUL AND YUGUO CHEN       230:         Statistical inference for model parameters in stochastic gradient descent<br>XI CHEN, JASON D. LEE, XIN T. TONG AND YICHEN ZHANG       25:         Bootstrap confidence regions based on M-estimators under nonstandard conditions<br>STEPHEN M. S. LEE AND PUYUDI YANG       274         Sparse high-dimensional regression: Exact scalable algorithms and phase transitions<br>DIMITRIS BERTSIMAS AND BART VAN PARYS       300         Testing for principal component directions under weak identifiability<br>DAVY PAINDAVEINE, JULIEN REMY AND THOMAS VERDEBOUT       324   | 3                     |
| Adaptive risk bounds in univariate total variation denoising and trend filtering       ADITYANAND GUNTUBOYINA, DONOVAN LIEU, SABYASACHI CHATTERJEE         AND BODHISATTVA SEN       20:         Spectral and matrix factorization methods for consistent community detection in       20:         Substrain and matrix factorization methods for consistent community detection in       20:         Statistical inference for model parameters in stochastic gradient descent       23:         Statistical inference for model parameters in stochastic gradient descent       25:         Bootstrap confidence regions based on M-estimators under nonstandard conditions       27:         Sparse high-dimensional regression: Exact scalable algorithms and phase transitions       20:         DIMITRIS BERTSIMAS AND BART VAN PARYS       30:         Testing for principal component directions under weak identifiability       32:         DAVY PAINDAVEINE, JULIEN REMY AND THOMAS VERDEBOUT       32:   |                       |
| Spectral and matrix factorization methods for consistent community detection in multi-layer networks   | j –                   |
| Statistical inference for model parameters in stochastic gradient descent       X1 CHEN, JASON D. LEE, X1N T. TONG AND YICHEN ZHANG       251         Bootstrap confidence regions based on M-estimators under nonstandard conditions       STEPHEN M. S. LEE AND PUYUDI YANG       274         Sparse high-dimensional regression: Exact scalable algorithms and phase transitions       DIMITRIS BERTSIMAS AND BART VAN PARYS       300         Testing for principal component directions under weak identifiability       DAVY PAINDAVEINE, JULIEN REMY AND THOMAS VERDEBOUT       324   | )                     |
| Bootstrap confidence regions based on M-estimators under nonstandard conditions<br>STEPHEN M. S. LEE AND PUYUDI YANG 274<br>Sparse high-dimensional regression: Exact scalable algorithms and phase transitions<br>DIMITRIS BERTSIMAS AND BART VAN PARYS 300<br>Testing for principal component directions under weak identifiability<br>DAVY PAINDAVEINE, JULIEN REMY AND THOMAS VERDEBOUT 324  |                       |
| Sparse high-dimensional regression: Exact scalable algorithms and phase transitions<br>DIMITRIS BERTSIMAS AND BART VAN PARYS 300<br>Testing for principal component directions under weak identifiability<br>DAVY PAINDAVEINE, JULIEN REMY AND THOMAS VERDEBOUT 32/  | ł                     |
| Testing for principal component directions under weak identifiability<br>DAVY PAINDAVEINE, JULIEN REMY AND THOMAS VERDEBOUT 324  | )                     |
|  | ł                     |
| The multi-armed bandit problem: An efficient nonparametric   |                       |
| solution   | )                     |
| Concentration and consistency results for canonical and curved exponential-family<br>models of random graphs MICHAEL SCHWEINBERGER   |                       |
| The summarical bootstrap   | 7                     |
| The numerical bootshap   |                       |
| Changliang Zou, Guanghui Wang and Runze Li 413   | ;                     |
| Uniformity valid confidence intervals post-model-selection<br>FRANÇOIS BACHOC, DAVID PREINERSTORFER AND LUKAS STEINBERGER 440  | )                     |
| Efficient estimation of linear functionals of principal components<br>VLADIMIR KOLTCHINSKII, MATTHIAS LÖFFLER AND RICHARD NICKL 464  | ,                     |

Vol. 48, No. 1-February 2020