American Statistical Association
Section on Statistical Learning and Data Science

SLDS Newsletter

Spring 2026
Editor: Sandra Safo

Table of Contents

SLDS Officers

Jun Yan	2026 Chair
Jaime Speiser	Chair Elect
Jing Lei	Past Chair
Xiaoyue Maggie	Secretary/Treasurer
Jiwei Zhao	Program Chair
Jingyi Jessica Li	Program Chair Elect
David Choi	CoS Representative
Sandra Safo	CoS Representative
Mladen Kolar	CoS Representative
Xiaoke Zhang	Publication Officer
Boxiang Wang	Webinar Coordinator

SLDS Website,
Published newsletters are archived under Newsletters under Helpful Documents.

SLDS 2026 Conference Website

Message from the Chair

It is a great honor to serve as Chair of the SLDS Section, one of the largest sections within ASA, with close to 3,000 members. Our size reflects the central role that statistical learning and data science now play across science, industry, and society, and it carries corresponding responsibility. As ASA's focal point for statistical learning and data science, our section advances research, convenes forums for discussion, and builds partnerships that promote principled statistical methods in data science. Stepping into this role, I am mindful that I am building on the work of dedicated leaders and volunteers before me, and I look forward to advancing our section's work together.

Strong leadership transitions are essential to the stability and continued growth of our section. I would like to express sincere appreciation to our outgoing officers—George Michailidis (Past Chair); Zhao Ren (Secretary/Treasurer); Irina Gaynonova (Program Chair); Jaime Speiser (COS Representative); Giles Hooker (Publication Officer); and Zhihua Sophia Su (Webinar Coordinator)—for their dedicated service and contributions. In particular, Zhihua Su led our webinar series for five years with distinction; it has become one of the most consistent and well-recognized ASA webinar programs, held monthly and carefully archived on our YouTube channel. I warmly welcome our new officers—including Jaime Speiser, transitioning from COS Representative to Chair Elect; Xiaoyue Maggie Niu (Secretary/Treasurer); Jingyi Jessica Li (Program Chair Elect); David Choi (COS Representative); Xiaoke Zhang (Publication Officer); and Boxiang Wang (Webinar Coordinator). Our section's vitality depends on sustained volunteer commitment and thoughtful succession planning. I encourage members to consider serving in future leadership roles.

Our Section's activities reflect sustained engagement across venues. At JSM 2025, the Section

sponsored five invited sessions (including the Leo Breiman Award Session), six topic-contributed sessions, more than twenty contributed paper sessions, and three professional development short courses, demonstrating our continued presence in ASA's flagship meeting. In parallel, the 2024 SLDS Conference, our first biennial meeting since the pandemic disrupted the conference cycle, was successfully held in Newport Beach, California, under the leadership of George Michailidis, reestablishing this flagship event and convening a broad community across academia and industry. These efforts are complemented by the dedicated work of our award committees. I thank the Leo Breiman Award Committee, chaired by Ming Yuan, and the Student Paper Award Committee, chaired by Yang Ning and Runze Li, for their careful review processes and thoughtful selections that recognize outstanding contributions and support the continued development of our field.

The SLDS 2026 Conference will be held in New York City on November 1–3 under the theme Inference and Intelligence. I am pleased to co-chair the organizing committee with our Past Chair Jing Lei. We thank our local organizers, Yang Feng and Wen Zhou, for their careful preparation and leadership. Negotiating hotel arrangements, coordinating audiovisual support, and managing an exceptionally strong set of invited session proposals—far more than we originally planned—have required sustained effort and sound judgment. We are continuing discussions with the hotel to secure additional space. I encourage members to attend, engage actively in discussion, and invite colleagues and students to participate. Strong participation will reflect the vitality of SLDS and its central role in advancing principled statistical methods across the data science landscape.

Sustained engagement requires strong communication, and we are therefore reviving the SLDS Newsletter as a regular platform for connection and transparency. I previously served as the inaugural editor of the Lifetime Data Science (LiDS) Newsletter and saw firsthand how a consistently published newsletter sustains momentum and strengthens a section's identity. I am grateful to Sandra Safo, our COS Representative, for volunteering to lead this effort. Through this newsletter, we will report on section activities, highlight member achievements, and share perspectives on emerging developments in statistical learning and data science. I encourage members to contribute news, ideas, and short pieces.

As statistical learning and data science continue to evolve rapidly, our section must remain attentive to problems of genuine importance. These include strengthening the statistical foundations of artificial intelligence, advancing trustworthy and responsible machine learning, developing scalable methods for large and complex data, contributing to emerging areas such as quantum computing, and advancing data-driven solutions to societal challenges in health, climate, and public policy. We are strengthening collaboration with the ASA interest group Stats Up AI and expanding engagement with industry partners to ensure that principled statistical thinking remains central in both research and practice. Through our conference, webinar series, sponsored sessions, and this newsletter, SLDS provides forums and resources to support these efforts. I encourage members to leverage these platforms and contribute to shaping the future of our field

Jun Yan
Chair 2026

Professor
Department of Statistics
University of Connecticut

SLDS 2026 Conference

The SLDS 2026 Conference will take place November 1–3, 2026 at the New York Marriott Hotel at the Brooklyn Bridge in Brooklyn, NY (SLDS 2026 Website). As the section's flagship meeting, SLDS brings together researchers and practitioners from academia, industry, and government to exchange ideas at the forefront of statistical learning and artificial intelligence.

The conference theme, "Inference and Intelligence," highlights the dynamic interplay between principled statistical reasoning and modern algorithmic developments. The program will span a wide range of topics, including artificial intelligence, big data analytics, causal inference, deep learning, graphical models, high-dimensional statistics, learning theory, machine learning, model selection, network analysis, spatiotemporal modeling, and text and image analytics. Applications will reflect ongoing work across the health, social, and engineering sciences, as well as data-driven challenges arising in the technology, finance, and pharmaceutical, and sports.

The conference will feature keynote presentations from

- David Banks (Duke University)
- Dean Foster (Amazon)
- David Rosenberg (Bloomberg)
- Bin Yu (University of California, Berkeley)
- Tian Zheng (Columbia University)

Their participation highlights the conference's commitment to fostering cross-sector collaboration and

connecting methodological developments with real-world implementation.

Preparations for the SLDS 2026 Conference are advancing steadily through the coordinated efforts of multiple committees. The Program Committee, chaired by Jiwei Zhao and Wen Zhou, has received over 100 invited session proposals and is working closely with the Local Committee to accommodate this strong response. The Local Committee, chaired by Yang Feng and Wen Zhou, is actively engaging with the conference hotel to secure additional space to support the expanded program. The Short Course Committee, chaired by Giles Hooker, has lined up short courses described in the following section. The Student/Early Career Paper Awards are being organized under the leadership of Nathaniel Sean O'Connell. The conference website, maintained by Yan Li, continues to provide timely updates and essential information. We also acknowledge Zhao Ren, for maintaining excellent records that support the planning process. These collective efforts reflect the strong commitment of SLDS volunteers to delivering a high-quality conference.The conference offers short courses, parallel invited sessions, a poster session, and a banquet.

A special issue associated with SLDS 2026 is being planned in the Journal of Statistical Planning and Inference (JSPI). Further details will be announced on the conference website as they become available.

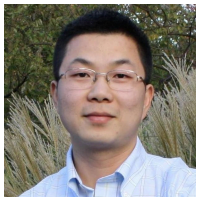
The Fundraising Committee is seeking sponsors for this conference. Sponsoring SLDS 2026 offers a unique opportunity for your organization to engage with talented professionals and students passionate about statistics and data science. Your sponsorship will provide direct access to emerging leaders in the field while supporting efforts to increase diversity and inclusion. By aligning with SLDS, you can:

- Enhance your brand visibility among the statistical learning and data science community.
- Connect with professionals and students from a wide array of academic programs and industries.
- Promote your commitment to education, diversity, and the future of data science.
- Recruit top-tier talent through exclusive access to participants during networking events.

SLDS is offering different tiers of sponsorship to meet your organization's goals and budget. If you are interested in sponsoring or learning more, please reach out to Jaime Speiser at jaime.speiser@wfusm.edu.

Building on strong program development and broad community interest, SLDS 2026 is positioned to be a highly engaging and well-attended meeting. Registration will open on June 1. The combination of a rigorous scientific program, active participation across sectors, and the continued growth of the SLDS community reflects the central role of statistical learning and data science in modern research and practice. Hosted in New York City—one of the world's leading centers for academia, industry, and innovation—the conference also offers a dynamic setting for professional exchange and collaboration. We look forward to welcoming the community to SLDS 2026.

Call for submissions: Student, Postdoc, and Early Career Researcher Paper Awards



Jing Lei
Co-Chair, SLDS 2026 Conference
Past Chair 2026

Professor
Department of Statistics & Data Science
Carnegie Mellon University

SLDS Short Courses

As part of our commitment to continuous learning, SLDS is pleased to announce the following short courses scheduled for this year. These sessions are designed to provide high-impact, practical training on emerging tools and methodologies shaping our field today. Topics currently in development include:

Bin Yu: **Veridical Data Science**
George Michailedes: **Optimization for statistics**
Yang Feng: **Transfer Learning**
Hongtu Zhu, Xiao Wang and Runpeng Dai: **Deep Learning Methods in Advanced Statistical Problems**
Yuxin Chen and Yuting Wei: **Statistical and Algorithmic Foundations of Diffusion Models**
Ivan Diaz, Kara Rudolf, Nick Williams: **Beyond the ATE**

Keep an eye on your inbox and our official website for the dates, full schedule and instructor bios. We look forward to seeing you there!

Giles Hooker
Chair, 2026 Short Course Committee
Professor
Department of Statistics and Data Science
University of Pennsylvania

JSM 2026 Program

The 2026 Joint Statistical Meetings (JSM) will be held in Boston, Massachusetts, August 1-6, 2026, with the theme "Communities in Action: Advancing Society".

SLDS is excited to sponsor the five invited sessions, 11 topic-contributed sessions, and 24 contributed sessions. In addition, as the co-sponsor, SLDS will be sponsoring 30 invited sessions and 17 topic-contributed sessions.

The five invited sessions are:

1. Statistical Frontiers in Large Language Models, Organizer: Will Wei Sun.
2. High-dimensional Statistics meet Deep earning, Organizer: Soufiane Hayou.
3. Generative AI in Modern Statistical Learning, Organizer: Kaizheng Wang.
4. Statistical Innovations for Heterogeneous and High-Dimensional Data, Organizer: Xuan Bi.
5. Integrating data from external sources: new methods for observational studies and randomized trials, Organizer: Peisong Han.

The eleven topic-contributed sessions are:

1. Extending Inference from Fixed Samples to Anytime Valid Guarantees, Organizer: Margaret Gamalo.
2. Recent Advances in Causal Discovery and Graphical Structure Learning, Organizer: Wei Jin.
3. Leveraging Diverse Data Sources: Transfer Learning, Data Integration, and Distributed Approaches, Organizer: Di Wang.
4. Statistical Innovations in Harnessing Large Language Models and EHR Data for Rare Disease Research, Organizer: Kimberly Greco and Cathy Shyr.
5. Integrating Large Language Models into Statistical Methodology: Opportunities and Challenges, Organizer: Edgar Dobriban.
6. Graph-Driven Statistical Learning for Complex-Structured Data, Organizer: Yinqiu He.
7. AI, Machine Learning, and Modern Statistical Methods for Complex Biomedical Data, Organizer: Gang Li.
8. Adaptive treatment designs for optimal medical care, Organizer: Justin Weltz.
9. Advances in Reinforcement Learning and Adaptive Experimentation, Organizer: Yongyi Guo.
10. Do you *really* know what's important? Organizer: Cynthia Rudin.
11. SLDS Student Paper Awards, Organizer: Yang Ning and Runze Li.

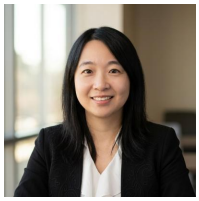
The breadth and depth of these sessions reflect the continued growth of statistical learning and data science within the ASA and the strong engagement of the SLDS community. We look forward to an active and well-attended program in Boston.

Jiwei Zhao
Program Chair 2026

Associate Professor
Department of Biostatistics & Medical Informatics
Department of Statistics
University of Wisconsin-Madison

Treasurer's Report

The Statistical Learning & Data Science Section generated a total income of $16,620.93 in 2025, primarily driven by member dues ($12,093.00) and net shares from conferences/proceedings ($4,527.93). Total expenses for the year amounted to $16,861.09, with the largest costs being awards and plaques ($7,000.00), honorariums ($4,875.00), and food and beverage expenses ($4,136.19). Overall, the section ended the year with a balance of $108,501.



Xiaoyue (Maggie) Niu
Treasurer/Secretary 2026-2027

Associate Professor
Director of the Statistical Consulting Center
Department of Statistics
Pennsylvania State University

Leo Brieman Award

The Leo Breiman Award recognizes outstanding contributions at the interface of statistical learning and computation, honoring the legacy of Leo Breiman, whose work on classification and regression trees, ensemble methods, and random forests helped bridge statistics and computer science. The award is given biennially to senior and junior scholars whose theoretical or methodological contributions have had substantial and sustained impact on both the field and its practical applications. The award includes a monetary prize, a plaque, and an invited lecture at the Joint Statistical Meetings in the year following the award announcement.

This year, the **Senior Award** goes to **Tony Cai** (University of Pennsylvania), for "deep theoretical insight paired with transformative methodological advances that have shaped both the foundations and the practice of contemporary statistics and machine learning. "

Dr. Tony Cai received his PhD from Cornell University in 1996. He joined the University of Pennsylvania in 2006 and is currently the Daniel H. Silberberg Professor and Professor of Statistics and Data Science. His research interests include high-dimensional inference, large-scale multiple testing,

nonparametric function estimation, functional data analysis, and statistical decision theory, with applications to compressed sensing, medical imaging and microarray data analysis.

Dr. Cai received the Committee of Presidents of Statistical Societies (COPSS) Presidents' Award in 2008, the Noether Distinguished Scholar Award from the American Statistical Association in 2023 and the Frontiers of Science Award at the 2023 International Congress of Basic Science. He was elected as a Fellow of the Institute of Mathematical Statistics in 2006, the President of International Chinese Statistical Association in 2016, and a Fellow of the American Association for the Advancement of Science in 2024.

The **Junior Award** is shared by **Yuxi Chen** (University of Pennsylvania), for "outstanding research at the intersection of statistics, data science, and machine learning, advancing the theoretical and algorithmic foundations of these fields while significantly expanding their frontiers," and **Anru Zhang** (Duke University), for "outstanding method and theory research in ML and high dimensional statistics, and impactful applications to real-world challenges, particularly in the field of biomedical informatics."

Dr. Yuxin Chen is currently an associate professor of statistics and data science and of electrical and systems engineering at the University of Pennsylvania. Before joining UPenn, he was an assistant professor of electrical and computer engineering at Princeton University. He completed his Ph.D. in Electrical Engineering at Stanford University and was also a postdoc scholar at Stanford Statistics. His current research interests include high-dimensional statistics, nonconvex optimization, and machine learning theory. He has received the Alfred P. Sloan Research Fellowship, the SIAM Activity Group on Imaging Science Best Paper Prize, the ICCM Best Paper Award (gold medal), and was selected as a finalist for the Best Paper Prize for Young Researchers in Continuous Optimization. He has also received the Princeton Graduate Mentoring Award.

Dr. Anru Zhang is a primary faculty member jointly appointed in the Department of Biostatistics & Bioinformatics and the Department of Computer Science at Duke University. He is also the Eugene Anson Stead, Jr. M.D. Associate Professor and serves as Associate Chair for Research in the Department of Biostatistics & Bioinformatics. His current research interests include generative models, biomedical data science, tensor learning, and high-dimensional statistics. He received the COPSS Emerging Leader Award, IMS Tweedie Award, ASA Gottfried E. Noether Junior Award, AMIA Data Science Outstanding Paper Award, and an NSF CAREER Award. Two of his PhD students have received the IMS Lawrence D. Brown Award. He currently serves as an associate editor for the Annals of Statistics, the Journal of the American Statistical Association (Theory & Methods and Applications & Case Studies), Statistica Sinica, ASA Discoveries, and Statistics and Its Interface. His research is currently supported by two NIH R01 grants (one as sole PI and one as MPI with clinical investigators).

We congratulate the awardees on this well-deserved recognition!

We thank the Leo Breiman Award Committee for their dedicated service to SLDS in reviewing the applications and selecting the awardees. The committee members were Ming Yuan (Chair, Columbia University), Liza Levina (University of Michigan), Alessandro Rinaldo (University of Texas at Austin), Mladen Kolar (University of Southern California), and Jing Lei (Carnegie Mellon University).

Ming Yuan
Chair, Breiman Award Committee

Professor of Statistics
Columbia University

JSM 2026 Student Paper Award

The annual JSM SLDS Student Paper Competition has become a highly sought-after opportunity for students to showcase their research in statistical learning and data science. This year, the competition attracted over 90 submissions, reflecting both the growing interest in the field and the strong engagement of the student community. The submissions covered a wide range of topics, highlighting methodological innovation as well as impactful applications

The Student Paper Award Committee selected winners through a careful and thorough review process. The six Student Paper Award recipients are (in alphabetical order):

- Young Hyun Cho, Purdue University, "Beyond Data Splitting: Full-Data Conformal Prediction by Differential Privacy"
- Patrick Kramer, Carnegie Mellon University, "Causal Inference with High-Dimensional Treatments"
- Ziqi Liu, Carnegie Mellon University, "Representation Learning with Blockwise Missingness and Signal Heterogeneity"
- Nan Lu, Harvard University; Chinese Academy of Sciences, "Contextual Online Uncertainty-Aware Preference Learning for Human Feedback"

- Zhiyu Xu, Columbia University, "Latency-Response Theory Model: Evaluating Large Language Models via Response Accuracy and Chain-of-Thought Length"
- Wenjin Zhang, Columbia University, "Discrete Causal Representation Learning"

The six awardees are invited to present their papers in a 2026 JSM topic-contributed session titled "SLDS Student Paper Awards" (1549).

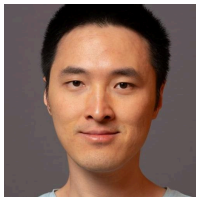
In addition, the committee selected two honorable mentions (in alphabetical order):

- Sophia Lu, Stanford University, "Likelihood-Free Adaptive Bayesian Inference via Nonparametric Distribution Matching"
- Gözde Sert, Texas A&M University, "Bayesian semiparametric causal inference: Targeted doubly robust estimation of treatment effects"

The large number of submissions required a correspondingly broad pool of judges:

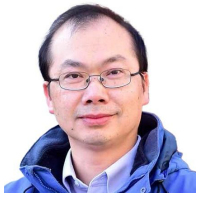
- Jeongyoun Ahn, Korea Advanced Institute of Science & Technology
- Jesus Arroyo, Texas A&M University
- Hengrui Cai, University of California, Irvine
- Xiaowu Dai, University of California, Los Angeles
- Paromita Dubey, University of Southern California
- Ethan Fang, Duke University
- Yuqi Gu, Columbia University
- Richard Guo, University of Michigan
- Ning Hao, University of Arizona
- Yinqiu He, University of Wisconsin–Madison
- Rebecca Killick, Lancaster University
- Mladen Kolar, University of Southern California
- Dehan Kong, University of Toronto
- Siddhesh Kulkarni, Bristol Myers Squibb
- Can Le, University of California, Davis
- Rongjie Liu, University of Georgia
- Xiaoqian Liu, University of California, Riverside
- Meimei Liu, Virginia Tech
- Rong Ma, Harvard University
- Qing Mai, Florida State University
- Yang Ni, Texas A&M University
- Ashley Petersen, University of Minnesota
- Zhao Ren, University of Pittsburgh
- Weining Shen, University of California, Irvine
- Hai Shu, New York University
- Wei Sun, Purdue University
- Boxiang Wang, University of Iowa
- Miaoyan Wang, University of Wisconsin–Madison
- Yuexi Wang, University of Illinois Urbana-Champaign
- Zhenke Wu, University of Michigan
- Ganggang Xu, University of Miami
- Lingzhou Xue, Pennsylvania State University
- Fei Xue, Purdue University
- Dana Yang, Cornell University
- Shihao Yang, Georgia Institute of Technology
- Zhuoran Yang, Yale University
- Guo Yu, University of California, Santa Barbara
- Mengxin Yu, Washington University in St. Louis
- Yubai Yuan, Pennsylvania State University
- Anru Zhang, Duke University
- Emma Zhang, Emory University
- Ruizhi Zhang, University of Georgia
- Yuan Zhang, Ohio State University
- Zhigen Zhao, Temple University
- Wenxin Zhou, University of Illinois Chicago

We thank all committee members for their substantial service and careful deliberation, which ensured a fair and rigorous evaluation process and made it possible to recognize these outstanding contributions.



Yang Ning
Co-Chair, Student Award Committee

Associate Professor
Department of Statistics & Data Science
Cornell University



Runze Li
Co-Chair, Student Award Committee

Eberly Family Chair Professor
Department of Statistics
Pennsylvania State University

SLDS Webinar Series

The SLDS Webinar Series continues to provide a consistent and well-attended platform for sharing advances in statistical learning and data science. The series maintains a regular schedule and features speakers from academia, industry, and government, with topics spanning methodology, computation, and impactful applications.

We are also pleased to announce our next two speakers. Dr. Anru Zhang (Duke) and Dr. Hadley Wickham (Posit) are scheduled to present on April 23 and May 28, respectively, both at 2:00 p.m. U.S. Eastern Time. Dr. Anru Zhang is the recipient of the 2026 Leo Breiman Award (Junior Category) and is an Associate Professor with joint appointments in the Departments of Biostatistics & Bioinformatics and Computer Science at Duke University. Dr. Hadley Wickham is Chief Scientist at Posit, where he leads the tidyverse team, and he received the 2019 COPSS Presidents' Award.

Please mark these dates on your calendar. More details will be shared with members as they become available.

Here is a list of recent and 2025 Webinars:

- 2026-03-26 — Lihua Lei (Stanford University, Graduate School of Business) — Compound Selection Decisions: An Almost SURE Approach.
- 2026-02-26 — Zhuoran Yang (Yale University, Statistics & Data Science) — How Do Transformers Learn to Implement Algorithms?
- 2026-01-26 — Glen Wright Colopy (Polaris Data Labs) — The AI Train Has Not Left the Station. But You Do Need to Step on Board.
- 2025-12-11 — Harrison Zhou (Yale University, Statistics & Data Science) — From Score Estimation to Sampling
- 2025-12-03 — Jiwei Zhao (University of Wisconsin–Madison) — Estimation and Inference in Unsupervised Domain Adaptation.
- 2025-11-05 — Jiancong Xiao and Xiang Li (University of Pennsylvania, Postdoctoral Researchers) — LLM Training and Coding for Statistical Learning and Data Science.
- 2025-09-23 — Liza Levina (University of Michigan, Statistics) — Towards Interpretable and Trustworthy Network-Assisted Prediction.
- 2025-08-27 — Haoda Fu (Amgen, Exploratory Biostatistics) — Open Research Problems for Statistical Learning and Data Science.
- 2025-07-31 — Po-Ling Loh (University of Cambridge) — Advice for Early-Career Researchers.
- 2025-07-01 — Giles Hooker (University of Pennsylvania, Statistics & Data Science) — Trees and V's: Inference for Ensemble Models.
- 2025-06-04 — Susan Murphy (Harvard University, Statistics & Computer Science) — Reinforcement learning in the Dyadic Setting.
- 2025-05-08 — Peter Song (University of Michigan SPH, Biostatistics) — Functional Accelerometer Data Analysis via Mixed Integer Optimization.
- 2025-04-02 — Irina Gaynanova (University of Michigan SPH, Biostatistics) — Distributional Learning for Wearable Data.
- 2025-02-26 — Adel Javanmard (USC Marshall, Data Sciences & Operations) — Learning from Aggregated Responses: Improving Model Utility Under Privacy Constraints.
- 2025-01-22 — Song Mei (UC Berkeley, Statistics & EECS) — Revisiting Neural Network Approximation Theory in the Age of Generative AI.

Links for future registrations can be found on our EventBrite, and links for past recordings can also be found on our Section's YouTube Channel. You may also search for "Eventbrite ASA SLDS" or "YouTube ASA SLDS" to find them directly. Thank you for subscribing to our YouTube channel.

I would like to express my sincere thanks to my predecessor, Dr. Zhihua Sophia Su, whose leadership established a strong foundation for the series. The structure, speaker pipeline, and archive of recorded talks that we benefit from today are the result of her careful and sustained effort. I am so much beyond grateful for the opportunity to build on this foundation and to continue developing the series.

Looking ahead, I welcome suggestions for future speakers and topics. We encourage contributions that represent the breadth of our field, including methodological developments, interdisciplinary work, and applications in practice. If you are interested in presenting or would like to nominate a speaker, please contact me at boxiang-wang@uiowa.edu.

Boxiang Wang
Webinar Chair

Associate Professor
Director of Graduate Studies
Department of Statistics and Actuarial Science
University of Iowa