



E-Tidings Newsletter

SCASA Events and News



Volume 14, Issue 3

March, 2026



PAGES 2 -3 SCASA 45th Annual Applied Statistics Workshop

PAGE 4 Traveling Course

PAGE 5 AI Tools for Data Science & Stats Webinar Series

PAGE 6 Dr. Normalcurvesaurus, Ph.D. Presents



SCASA's 45th Applied Statistics Workshop * Friday, May 8, 2026



Enju Liu

MD, PhD

Senior Biostatistician, Institutional Centers
for Clinical and Translational Research
(ICCTR)

Assistant Professor of Pediatrics, Harvard
Medical School

SCASA CORDIALLY INVITES YOU TO THE 45TH ANNUAL APPLIED STATISTICS WORKSHOP

DATE: Friday, May 8, 2026

TIME: 9:00 AM to 3:00 PM

PLACE: on Zoom

INSTRUCTOR: Dr. Enju Liu

TOPIC: UNDERSTANDING COLLIDER BIAS IN BIOSTATISTICAL ANALYSIS

COST: \$5-\$20

REGISTRATION: [Link](#)

Continues on the next page.

Continued from the previous page.

INSTRUCTOR'S BIO: Dr. Enju Liu is an Assistant Professor in Pediatrics at Harvard Medical School. She has a broad background in Medicine, Biostatistics and Epidemiology with specific training in nutrition and statistical methods. She earned her PhD in Nutritional Epidemiology and has extensive experience in design, analysis and conduct of observational studies and randomized controlled trials. She collaborates with investigators at Boston Children's Hospital, across the Harvard Community as well as outside Harvard on the international level. She currently serves as the Associate Director for the Harvard Catalyst Biostatistics Program. Her research has focused on preventive and therapeutic effects of dietary factors, nutritional supplements, and other interventions on child and maternal health outcomes. Her research activities have resulted in over 120 publications in peer-reviewed scientific journals, including the New England Journal of Medicine, Cell, and the American Journal of Clinical Nutrition.

COURSE DESCRIPTION: In biostatistical analysis, adjustment for potential confounders is a common practice to determine the effect of an exposure or treatment on health outcomes. However, little attention has been given to the potential distortion of the association between exposure and outcome caused by collider bias. A collider refers to a variable that is caused by both the exposure and the outcome or risk factors of the outcome. It is often overlooked but can introduce bias into data analysis, leading to erroneous conclusions. Various statistical approaches commonly employed to control potential confounding, such as restriction, stratification, or adjustment for the collider in regression models, can inadvertently introduce collider bias. This short course aims to equip researchers, biostatisticians, and data analysts in the fields of public health or clinical research with valuable knowledge to increase their awareness of collider bias. Throughout the course, participants will learn how to identify and interpret causal diagrams and directed acyclic graphs (DAGs), which serve as powerful tools for assessing and understanding collider bias. Real-world examples and case studies will be utilized to illustrate the potential impact of collider bias on study results, emphasizing the importance of accurately addressing this bias in data analysis and decision-making.



The Traveling Course is coming to SCASA



DAY/TIME: TBD, in September-October, 2026

PRESENTATION MODE: Half-day, online

COURSE TITLE: Tree-Based Machine Learning Methods for Prediction and Variable Selection

INSTRUCTORS: Drs. Hemant Ishwaran and Min Lu, University of Miami

ABSTRACT: Tree-based machine learning methods offer several benefits in data analysis, including non-linearity, robustness, scalability and handling mixed data types. This course emphasizes practical learning with hands-on code examples and result interpretations, which is essential for understanding and applying these techniques. Based on the widely popular R package "randomForesSRC", we will present methods for computing predicted outcomes, variable importance indices and other inference estimates. In addition, we will introduce a new model-independent variable selection method, called the rule-based variable priority, and present its implementation using the R package "varPro". For all these analyses, we will cover different types of outcomes including continuous, categorical, multivariate, survival and competing risk outcomes. Utilizing real-world datasets from medicine and public health, topics in these analyses will provide hands-on code, working examples and result interpretations. We will provide additional code for visualizing model results and constructing coefficient tables for interpretation, and address scenarios such as imbalanced classes, unsupervised problems, fast implementation on big data and protection of confidential data.

Thu, Apr 2, 11:00 AM PDT



**AI Tools
for Data Science & Stats**

Webinar Series

**The AI Train Has Not Left the Station.
(But Get On Board.)**



Check out this webinar series hosted by our friends at San Francisco Bay Area Chapter!



Hey folks,

We're kicking off the Spring webinar series on AI tools for data scientists and statisticians.

TITLE: The AI Train Has Not Left the Station. (But Get On Board.)

DATE AND TIME: THURSDAY, APRIL 2, 2026, 2:00 TO 3:00 PM EASTERN TIME

REGISTRATION LINK: [HTTPS://WWW.LINKEDIN.COM/EVENTS/7440847551005650946/](https://www.linkedin.com/events/7440847551005650946/)

ABSTRACT: AI is simultaneously overhyped and underhyped (Schrödinger's hype-cat, if you will...and in a blackbox too!). It's overhyped because we're drowning in AI slop, confidently wrong answers, and hallucinated citations. Also, some folks are using chatbots as a substitute for thinking. (Admittedly, it's largely the same crowd that will substitute anything for thinking, so that's not AI's fault.)

It's also underhyped for data scientists and statisticians, specifically. Because these tools are opening up huge opportunities for us - AI's knocking over major barriers in what we can do and how long it takes to do it. It's high signal + high noise, but the profession can't afford to wait for the dust to settle before jumping in.

I'd like to cut through the noise and get you comfortable with trying some AI tools at the ground level. We'll cover straightforward ways to get real value from AI for both repetitive and open-ended work, and we'll review the "AI blooper reel" - showing where these tools fail - so you develop a feel for when to trust them and when not to.

This webinar will stay high-level and non-technical. If you're looking for AI-assisted coding, check out the Claude Code webinar on April 9.

Claude Code, Cowork, & Chat - Thursday, April 9, 2026 - [Registration Link](#)

Want to stay in the loop? Sign up for the mailing list to get updates and post-webinar materials: [Link](#)

Hope to see you there!

Glen Wright Colopy
DPhil Oxon
Host | The Data & Science Podcast
Head of Data Science | Polaris Data Labs

DR. NORMALCURVESAURUS, PH.D., PRESENTS

