Bernard G. Greenberg Distinguished Lectures

2022 Greenberg Guest Lecturer

Professor Andrew Gelman, PhD

Department of Statistics and Department of Political Science, Columbia University

All the ways that Bayes can go wrong

Lecture 1: Thursday, May 12, 10:00–11:00 AM Eastern - LINK
This will be a remote / virtual event, if desired gather together in 133 Rosenau Hall
https://unc.zoom.us/j/94209312948?pwd=dmtteWVDd0UzME5VL2tmRXVKTWF6Zz09
Meeting ID: 942 0931 2948
Passcode: 817083

From sampling and causal inference to policy analysis: Interactions and the challenges of generalization

Lecture 2: Thursday, May 12, 2022, 2:00–3:00 PM Eastern – LINK
This will be a remote / virtual event, if desired gather together in 133 Rosenau Hall
https://unc.zoom.us/j/96082871462?pwd=ZW5Xd0tENUV1aXUvd2hSTFFiN2hFQT09
Meeting ID: 960 8287 1462
Passcode: 768739

Statistical workflow

Lecture 3: Friday, May 13, 10:00–11:00 AM Eastern – LINK
This will be a remote / virtual event, if desired gather together in 133 Rosenau Hall
https://unc.zoom.us/j/91862747023?pwd=Vnk4Y2VRVXhxWXZXQWw3RmNDSXJVQT09
Meeting ID: 918 6274 7023
Passcode: 506611

Professor Gelman's biography and each lecture abstract are on following pages.

The Bernard G. Greenberg Distinguished Lecture Series honors the first chair of the UNC Biostatistics Department, Dr. Bernard G. Greenberg, who later served with distinction as dean of the School of Public Health from 1972 to 1982. The recipient of the Greenberg Distinguished Lecturer Award is selected the Biostatistics faculty based on the quality and public health impact of nominees' research.

Lecturer Biography

Andrew Gelman is a professor of statistics and political science at Columbia University. He has received the Outstanding Statistical Application award three times from the American Statistical Association, the award for best article published in the American Political Science Review, and the Council of Presidents of Statistical Societies award for outstanding contributions by a person under the age of 40. His books include Bayesian Data Analysis (with John Carlin, Hal Stern, David Dunson, Aki Vehtari, and Don Rubin), Teaching Statistics: A Bag of Tricks (with Deb Nolan), Data Analysis Using Regression and Multilevel/Hierarchical Models (with Jennifer Hill), Red State, Blue State, Rich State, Poor State: Why Americans Vote the Way They Do (with David Park, Boris Shor, and Jeronimo Cortina), A Quantitative Tour of the Social Sciences (co-edited with Jeronimo Cortina), and Regression and Other Stories (with Jennifer Hill and Aki Vehtari).

Andrew has done research on a wide range of topics, including: why it is rational to vote; why campaign polls are so variable when elections are so predictable; why redistricting is good for democracy; reversals of death sentences; police stops in New York City, the statistical challenges of estimating small effects; the probability that your vote will be decisive; seats and votes in Congress; social network structure; arsenic in Bangladesh; radon in your basement; toxicology; medical imaging; and methods in surveys, experimental design, statistical inference, computation, and graphics.

Lecture 1. All the ways that Bayes can go wrong

Probability theory is false. Weak priors give strong and implausible posteriors. If you could give me your subjective prior I wouldn't need Bayesian inference. The best predictive model averaging is non-Bayesian. There will always be a need to improve our models. Nonetheless, we still find Bayesian inference to be useful. How can we make the best use of Bayesian methods in light of all their flaws?

Lecture 2. From sampling and causal inference to policy analysis: Interactions and the challenges of generalization

The three central challenges of statistics are generalizing from sample to population, generalizing from control to treated group, and generalizing from observed data to underlying constructs of interest. These are associated with separate problems of sampling, causal inference, and measurement, but in real decision problems all three issues arise. We discuss the way in which varying treatment effects (interactions) bring sampling concerns into causal inference, along with the real challenges of applying this insight into real problems. We consider applications in medical studies, A/B testing, social science research, and policy analysis.

Lecture 3. Statistical Workflow

Statistical modeling has three steps: model building, inference, and model checking, followed by possible improvements to the model and new data that allow the cycle to continue. But we have recently become aware of many other steps of statistical workflow, including simulated-data experimentation, model exploration and understanding, and visualizing models in relation to each other. Tools such as data graphics, sensitivity analysis, and predictive model evaluation can be used within the context of a topology of models, so that data analysis is a process akin to scientific exploration. We discuss these ideas of dynamic workflow along with the seemingly opposed idea that statistics is the science of defaults. We need to expand our idea of what data analysis is, in order to make the best use of all the new techniques being developed in statistical modeling and computation.