

Sailing Through Data: Discoveries and Mirages

Emmanuel Candès, Stanford University

6:45 pm – 8pm, 05/22/2019, Golden Gate, 1195 Bordeaux Dr, Sunnyvale, CA

Video conference in Bellevue WA (Mt Rainier) and Beijing (北京, K2-F6 碧桃春)

For a long time, science has operated as follows: a scientific theory can only be empirically tested, and only after it has been advanced. Predictions are deduced from the theory and compared with the results of decisive experiments so that they can be falsified or corroborated. This principle formulated by Karl Popper and operationalized by Ronald Fisher has guided the development of scientific research and statistics for nearly a century. We have, however, entered a new world where large data sets are available prior to the formulation of scientific theories. Researchers mine these data relentlessly in search of new discoveries and it has been observed that we have run into the problem of irreproducibility. Consider the April 23, 2013 Nature editorial: “Over the past year, Nature has published a string of articles that highlight failures in the reliability and reproducibility of published research.” The field of Statistics needs to re-invent itself to adapt to the new reality where scientific hypotheses/theories are generated by data snooping. We will make the case that statistical science is taking on this great challenge and discuss exciting achievements. In particular, we will introduce the method of knockoffs, which reliably selects which of the many potentially explanatory variables of interest (e.g. the absence or not of a mutation) are indeed truly associated with the response under study (e.g. the risk of getting a specific form of cancer). We will see that this line of work opens up new perspectives in applied probability, in scientific computing, and even in deep learning.

Bio: Emmanuel Candès is the Barnum-Simons Chair in Mathematics and Statistics, and professor of Electrical Engineering (by courtesy) at Stanford University, where he currently chairs the Department of Statistics. Emmanuel’s work lies at the interface of mathematics, statistics, information theory, signal processing and scientific computing. Candès graduated from the Ecole Polytechnique in 1993 with a degree in science and engineering, and received his Ph.D. in Statistics from Stanford University in 1998. He received the 2006 Alan T. Waterman Award, perhaps the highest honor bestowed by the National Science Foundation. He is a member of the National Academy of Sciences and the American Academy of Arts and Sciences. Candès has been named a 2017 MacArthurFellow. He has given 80 plenary lectures at major international conferences, not only in mathematics and statistics but in many other areas as well including biomedical imaging and solid-state physics.