UCLA College | Physical Sciences Statistics

UCLA Fielding School of Public Health Biostatistics

2023 De Leeuw Seminar

Inference-based Monte Carlo Integrations: A Likelihood or Bayesian Paradox?

By Xiao-Li Meng
Department of Statistics, Harvard University

ABSTRACT: This talk is based on Uehara and Meng (2023), which has the same title, and the following abstract "Monte Carlo integrations can be improved when there is reliable auxiliary information, such as mathematical bounds or numerical approximations. Whereas the Bayesian paradigm seems to provide a perfectly principled approach for incorporating such prior information, attempts to do so have led to the seemingly paradoxical discovery of useless likelihood functions in the context of computing normalizing constants(Wasserman, 2013).We explain that the crux of the paradox is not with the likelihood theory, but rather incorrect specifications of the model parameters with the Monte Carlo data, using the modeling-what-we-ignore idea of Kong et al.(2003). However, there is a real Bayesian paradox: if we were able to perform Bayesian Monte Carlo exactly, then we would not have needed the method in the first place, because the original integration problem requires less computation. We then show that there is a practical resolution to this paradox, using the profile likelihood obtained in Kong et al. (2006), and that this approximation is second-order valid asymptotically. We also investigate a computationally more efficient approximation via an artificial likelihood of Geyer (1994). This artificial likelihood approach is only first-order valid, but there is a computationally trivial adjustment to render its second-order validity. We demonstrate empirically the efficiency of these artificial likelihood-based Bayesian estimators, compared to the usual design-based Monte Carlo estimators, such as bridge sampling estimators. The effectiveness of the artificial Bayesian strategies also provides both a demonstration and food for thoughts on seeking optimal trade-offs between statistical efficiency and computational efficiency, a widely open problem at the core of the foundation of data science. A simple example via the symmetrized importance sampling estimator demonstrates the mathematical possibility of unlimited gains by such a trade-off strategy."

About the Speaker

Xiao-Li Meng is a renowned statistician who is the Founding Editor-in-Chief of Harvard Data Science Review and the Whipple V. N. Jones Professor of Statistics. He was named the best statistician under the age of 40 by Committee of Presidents of Statistical Societies (COPSS) in 2001 and is the recipient of numerous awards and honors for his more than 150 publications in at least a dozen theoretical and methodological areas. He is the author of "The XL-Files," a thought-provoking and entertaining column in the Institute of Mathematical Statistics (IMS) Bulletin, and his interests range from the theoretical foundations of statistical inferences to statistical methods and computation. He received his BS in mathematics from Fudan University in 1982 and his PhD in statistics from Harvard in 1990. Meng was on the faculty of the University of Chicago from 1991 to 2001 before returning to Harvard, where he served as the Chair of the Department of Statistics (2004–2012) and the Dean of Graduate School of Arts and Sciences (2012–2017). In 2020, he was elected to the American Academy of Arts and Sciences.

Wednesday April 12, 2023, 3:30pm-4:30pm

Luskin Conference Center, Legacy Room