

中央研究院統計科學研究所

學術演講

講題：Advancing Statistical Foundations in AI: From Predictive Modeling to Generative Data Science

演講人：Dr. Chi-Hua Wang (王啟樺 博士)

Trustworthy AI Lab, Department of Statistics & Data Science, UCLA

時間：2025-02-05 (Wed.) 10:30-12:00

地點：Auditorium, B1F, Institute of Statistical Science; The tea reception will be held at 10:10.

備註：Online live streaming through Cisco Webex will be available.

Abstract

At the intersection of statistics and artificial intelligence lies a profound opportunity to push the boundaries of both theory and application. In this talk, I will share my research journey in advancing the statistical foundations of AI, structured into three interconnected parts, each addressing critical challenges in Predictive AI and Generative AI.

Part 1 explores my work on Dynamic Pricing, a cornerstone of Predictive AI. By developing adaptive pricing models grounded in bandit-based frameworks, always valid inference, and high-dimensional regularization, I tackle the exploration-exploitation trade-off inherent in dynamic decision-making. These models enable statistically rigorous, privacy-aware, and real-time applications in industries like e-commerce and advertising, demonstrating how statistical methods can drive impactful results in complex markets.

Part 2 focuses on Privacy Auditing, bridging the domains of Predictive AI and Generative AI. This research leverages statistical hypothesis testing to design data-driven frameworks that quantify and mitigate privacy risks, including membership inference attacks and data copying in generative models. By combining theoretical guarantees, such as differential privacy, with practical evaluations, I aim to provide actionable tools that align privacy preservation with analytical utility, addressing one of the most pressing concerns in modern AI.

Part 3 delves into Generative Data Science, where I focus entirely on Generative AI. My work prioritizes the synthesis and evaluation of synthetic tabular data, emphasizing statistical fidelity and machine learning utility. By establishing robust evaluation metrics and exploring domain-specific applications in finance, retail and advertising, I aim to unlock the potential of synthetic data as a secure and powerful tool for collaborative analytics.

These three parts form a cohesive vision that integrates statistical rigor with practical solutions, advancing both Predictive AI and Generative AI. Through this talk, I seek to demonstrate how a solid statistical foundation can not only address theoretical challenges but also enable transformative applications, laying the groundwork for the next generation of AI innovations.



中央研究院

統計科學研究所