

## 趙民德教授講座

主講者: 張源俊 (Yuan-chin Ivan Chang) 教授



張教授的主要研究領域包含了序列分析(Sequential Analysis)、廣義線性模型( Generalized Linear Model), 教育統計(Educational Statistics)、機器/統計學習( Machine/Statistical Learning)、長期追蹤資料分析(Longitudinal Data Analysis)等。主題看似不多，但其實都不相近。這反應了他研究的選擇與態度。張教授有興趣的問題與涉獵領域十分豐富，其中沉潛多年當屬音樂。除了是古典音樂的深度愛好者，他也曾參與合唱團，學習/演奏長笛，音樂素養不在話下。在數學/統計科普推廣上，他一直不遺餘力，從調音與數學，到孔子談網路教育，以及新近的咖啡、音樂與學術品味快閃演講，樂此不疲。如此豐富的生活，他仍擔任了不同時期 Sequential Analysis Journal, Journal of Japanese Society of Computational Statistics 的 Associated Editor; Statistica Sinica, JCSA 的 Members of Chief-Editors, Managing Editor 等職務。在正式或非正式的場合，他常與其他研究者，特別是年輕一輩的研究者交換研究想法與新方向。提攜鼓勵新一代學者上，他投注了相當的心力。早在 AI 還在常溫階段的 2004 年，他就推動了 Statistical and Machine Learning Workshop; 資料科學統計合作社自然也與他關係密切... 林林總總，未完待續...

## 趙民德教授講座

主講者: 張源俊 (Yuan-chin Ivan Chang) 教授

中央研究院統計科學研究所

### **Distributed sequential estimation procedures – Something old is new again.**

#### **Abstract**

Data collected from distributed sources or sites commonly have different distributions or contaminated observations. Active learning procedures allow us to assess data when recruiting new data into model building. Thus, combining several active learning procedures together is a promising idea, even when the collected data set is contaminated. In this talk, we will briefly describe how we can apply sequential analysis in Statistics, a concept raised by Wald (1945) during WWII, to modern machine learning such that we cannot only save both the sampling and computing costs, and stay robust while dealing with the contaminated data scenarios commonly found analyzing big data sets, and a demonstration with Python is available through Github at <https://github.com/zhuojianc/dsep>.