

國立中央大學

統計研究所

學術演講

主 講 人：林長鑾 教授（國立中興大學應用數學系）

講 題：**Forward stepwise random forest analysis for experimental designs**

時 間：**110年10月05日（星期二）上午11：00 ~ 12：00**

地 點：**中央大學鴻經館M-116室**

茶 會：**上午 10：30 ~ 11：00** 地 點：**鴻經館 M510 室**

ABSTRACT

In experimental designs, it is usually assumed that the data follow normal distributions and the models have linear structures. In practice, experimenters may encounter different types of responses and be uncertain about model structures. If this is the case, traditional methods, such as the ANOVA and regression, are not suitable for data analysis and model selection. We introduce the random forest analysis, which is a powerful machine learning method capable of analyzing numerical and categorical data with complicated model structures. To perform model selection and factor identification with the random forest method, we propose a forward stepwise algorithm and develop Python and R codes based on minimizing the OOB error. Six examples including simulation and case studies are provided. We compare the performance of the proposed method and some frequently used analysis methods. Results show that the forward stepwise random forest analysis, in general, has a high power for identifying active factors and selects models that have high prediction accuracy.

◎敬請張貼

歡迎參加◎