

國立中正大學數學系

暨應用數學碩士班、統計科學碩士班

學 術 演 講

**Overlapping group screening for detection of gene-environment interactions
with application to TCGA high-dimensional survival genomic data**

王价輝教授

Prof. Jie-Huei Wang

逢甲大學統計學系

Department of Statistics, Feng Chia University

Abstract

Background: In the context of biomedical and epidemiological research, gene-environment (G-E) interaction is of great significance to the etiology and progression of many complex diseases. In high-dimensional genetic data, two general models, marginal and joint models, are proposed to identify important interaction factors. Most existing approaches for identifying G-E interactions are limited owing to the lack of robustness to outliers/contamination in response and predictor data. In particular, right-censored survival outcomes make the associated feature screening even challenging. In this article, we utilize the overlapping group screening (OGS) approach to select important G-E interactions related to clinical survival outcomes by incorporating the gene pathway information under a joint modeling framework. **Results:** Simulation studies under various scenarios are carried out to compare the performances of our proposed method with some commonly used methods. In the real data applications, we use our proposed method to identify G-E interactions related to the clinical survival outcomes of patients with head and neck squamous cell carcinoma, and esophageal carcinoma in The Cancer Genome Atlas clinical survival genetic data, and further establish corresponding survival prediction models. Both simulation and real data studies show that our method performs well and outperforms existing methods in the G-E interaction selection, effect estimation, and survival prediction accuracy. **Conclusions:** The OGS approach is useful for selecting important environmental factors, genes and G-E interactions in the ultra-high dimensional feature space. The prediction ability of OGS with the Lasso penalty is better than existing methods. The same idea of the OGS approach can apply to other outcome models, such as the proportional odds survival time model, the logistic regression model for binary outcomes, and the multinomial logistic regression model for multi-class outcomes.

Keywords: Gene-environment interaction, Joint model, Lasso, Overlapping group screening, Survival prediction, TCGA

Joint work with Kang-Hsin Wang and Yi-Hau Chen

日 期:112 年 2 月 22 日(星期三) 16:10~17:00

地 點: 本校數學館 527 教室 (嘉義縣民雄鄉大學路 168 號)

茶 會: 15:30~16:00 數學館四樓 409 室舉行

歡迎參加 敬請公佈