國立中央大學統計研究所

學術演講

主 講 人:王价輝 博士(中央研究院 統計科學研究所)

講 題: Overlapping Group Screening for Detection of Gene-gene Interactions:

Application to Gene Expression Profiles with Survival Trait

時 間:107年10月09日(星期二) 上午11:00 ~ 12:00

地 點:中央大學鴻經館M429室

茶 會: 上午 10:30 ~ 11:00 地 點: 鴻經館 510 室

ABSTRACT

Background: The development of a disease is a complex process that may result from joint effects of multiple genes. In this article, we propose the overlapping group screening (OGS) approach to determining active genes and gene-gene interactions incorporating prior pathway information. The OGS method is developed to overcome the challenges in genome-wide data analysis that the number of the genes and gene-gene interactions is far greater than the sample size, and the pathways generally overlap with one another. The OGS method is further proposed for patients' survival prediction based on gene expression data.

Results: Simulation studies demonstrate that the performance of the OGS approach in identifying the true main and interaction effects is good and the survival prediction accuracy of OGS with the Lasso penalty is better than the ordinary Lasso method. In real data analysis, we identify several significant genes and/or epistasis interactions that are associated with clinical survival outcomes of diffuse large B-cell lymphoma (DLBCL) and non-small-cell lung cancer (NSCLC) by utilizing prior pathway information from the KEGG pathway and the GO biological process databases, respectively.

Conclusions: The OGS approach is useful for selecting important genes and epistasis interactions in the ultra-high dimensional feature space. The prediction ability of OGS with the Lasso penalty is better than existing methods. The OGS approach is generally applicable to various types of outcome data (quantitative, qualitative, censored event time data) and regression models (e.g. linear, logistic, and Cox's regression models).

Keywords: gene-gene interaction, Lasso, overlapping group, survival prediction

◎敬請張貼