

Colloquium

Influence Analysis on Linear Regression for Symbolic Interval Data

主講人：黃郁芬 教授
國立中正大學數學系

時 間：106 年 5 月 17 日(三) 14：30

地 點：應用數學系多媒體教室（理 408 室）

摘 要：

Nowadays, with the advent of computers, data sets become inevitably large than before. This brings the difficulty in performing standard statistical analysis. Hence, such huge data sets might be aggregated in some fashion and the resulting summary data may be represented by lists, intervals, histograms and the like, which are called symbolic data. Linear regression is one of the most popular and useful tools to analyze the data for studying the relationship between a response variable and its explanatory variable(s). During the fitting process, observations that are suspicious can greatly influence the results of the analysis. Therefore, detection of such influential points becomes an essential task. Many literatures have studied for the influence analysis in linear regression for classical data. However, to our knowledge, a study in the influence analysis on regression for symbolic data has not been explored in the literature. In this paper, we develop three sample versions of the influence function motivated by Hampel (1974) to identify suspicious concepts which cause seriously adverse effects on the linear regression analysis results for symbolic interval data. Also we illustrate these proposed methods with simulation studies and real data examples.