

Abstract No:I-3-2

# **Model selection based on resampling approaches for cluster longitudinal data with missingness in outcomes**

沈仲維<sup>1</sup>

<sup>1</sup> 國立中正大學

## **Abstract**

In medical and health studies, longitudinal and cluster longitudinal data are often collected, where the response variable of interest is observed repeatedly over time and along with a set of covariates. Model selection becomes an active research topic but has not been explored largely due to the complex correlation structure of the data set. To address this important issue, in this paper, we concentrate on model selection of cluster longitudinal data especially when data are subject to missingness. Motivated from the expected weighted quadratic loss of a given model, data perturbation and bootstrapping methods are used to estimate the loss and then the model that has the smallest expected loss is selected as the best model. To justify the proposed model selection method, we provide various numerical assessments and a real application regarding the asthma data set is also analyzed for illustration.

Keyword: bootstrap, data perturbation, generalized estimating equations, variable selection