Novel Empirical Likelihood Method for the Cumulative Hazard Ratio under Stratified Cox Models

Yichuan Zhao[1], Dazhi Zhao[2,*]

Department of Mathematics and Statistics, Georgia State University

ABSTRACT

Evaluating the treatment effect is a crucial topic in clinical studies. Nowadays, the ratio of cumulative hazards is often applied to accomplish this task, especially when those hazards may be nonproportional. The stratified Cox proportional hazards model, as an important extension of the classical Cox model, has the ability to flexibly handle nonproportional hazards. In this article, we propose a novel empirical likelihood method to construct the confidence interval for cumulative hazard ratio under the stratified Cox model. The large sample properties of the proposed profile empirical likelihood ratio statistic are investigated, and the finite sample properties of the empirical likelihood-based estimators under some different situations are explored in simulation studies. The proposed method was finally applied to perform statistical analysis on a real-world dataset on the survival experience of patients with heart failure.

Keywords: Cumulative hazard ratio; Empirical likelihood; Nonproportionality; Stratified Cox model; Survival analysis

GUEST: Graphical Models for Ultrahigh-Dimensional and Error-Prone Data by the Boosting Algorithm

Li-Pang Chen, Hui-Shan Tsao

Department of Statistics, National Chengchi University

ABSTRACT

In bioinformatics studies, understanding the network structure of gene expression variables is one of the main interests. In the framework of data science, graphical models have been widely used to characterize the dependence structure among multivariate random variables. However, the gene expression data possibly suffer from ultrahigh-dimensionality and measurement error, which make the detection of network structure challenging. The other important application of gene expression variables is to provide information to classify subjects into various tumors or diseases. In supervised learning, while linear discriminant analysis is a commonly used approach, the conventional implementation is limited in precisely measured variables and computation of their inverse covariance matrix, which is known as the precision matrix. To tackle those challenges, we introduce a new method called GUEST, which refers to Graphical models for Ultrahigh-dimensional and Error-prone data by the booSTing algorithm. The estimation strategy includes measurement error correction in high-dimensional variables under various distributions and then applies the boosting algorithm to identify the network structure and estimate the precision matrix. When the precision matrix is estimated, it can be used to construct the linear discriminant function and improve the accuracy of the classification.

Keywords: Feature screening; measurement error; network structure; statistical learning

A Likelihood Approach for Data Integration Involving Missing Data and Misclassified Variables

Zheng Yu[1], **Hua Shen[2*]**

Department of Mathematics and Statistics, University of Calgary

ABSTRACT

Integrating information from probability and nonprobability samples for unbiased inference can encounter both missing data and measurement errors, distorting population estimates. We develop a likelihood-based method that addresses missing outcomes and misclassified covariates in a probability sample, as well as dual misclassification of both outcome and covariate in a nonprobability sample. Simulation studies show substantial bias reduction over naive methods, and a real-world example underscores the method's practical advantages for more reliable inference.

Keywords: Data Integration, Measurement Error, Nonprobability Sample