

The L_q -norm learning for ultrahigh-dimensional survival data: an integrative framework

Xuerong Chen

Southwestern University of Finance and Economics
E-mail: chenxuerong@swufe.edu.cn

Abstract: In the era of precision medicine, survival outcome data with high-throughput predictors are routinely collected from many biomedical studies. Models with an exceedingly large number of covariates are either infeasible to fit or likely to incur low predictability because of overfitting. Variable screening is key in identifying and removing irrelevant attributes. Recent years have seen a surge in screening methods, but most of them rely on some particular modeling assumptions. Motivated by a study on detecting gene signatures for multiple myeloma patients' survival, we propose a model-free L_q -norm learning procedure, which includes the well-known Cramér–von Mises and Kolmogorov criteria as two special cases. The work provides an integrative framework for detecting predictors with various levels of impact, such as short- or long-term impact, on censored outcome data. The framework naturally leads to a scheme which combines results from different q 's to reduce false negatives, an aspect often overlooked by the current literature. We show that our method possesses sure screening properties. The utility of the proposal is confirmed with simulation studies and an analysis of the multiple myeloma study.