

A Semiparametric Model with Application to Analysis of Lifetime Data

Ming-Yen Cheng
National Taiwan University, Taiwan

Wenyang Zhang
University of Kent, UK

Traditional parametric modelling assumes the samples come from a specific family of distributions, and estimates the unknown parameters based on the samples. Maximum likelihood estimation is most powerful in that case. In reality, we don't know which family of distributions the samples come from, and model misspecification can result in seriously biased estimators. Nonparametric modelling makes no assumptions on the distribution; however, it would pay a price on the variance side since there are more unknowns to be estimated. Even worse, when the dimension of the covariates is large, nonparametric modelling has the "curse of dimensionality" problem. A promising idea is to relax the restrictions imposed by parametric models to make model specification more flexible. We develop a hybrid of parametric and nonparametric models which has a wide range of applications. It copes with multiple covariates and adapts to dynamic structural changes well. The associated estimation problem is solved by a simple and effective method. The proposed estimator of the parametric part has root-n convergence rate, and the estimator of the nonparametric part enjoys an oracle property. Model selection is discussed. A simulation study demonstrates the performance of the proposed methods. Finally, the proposed model is used to analyse the infant mortality data of China.

[Ming-Yen Cheng, Department of Mathematics, National Taiwan University, Taipei 106, Taiwan; E-MAIL: cheng@math.ntu.edu.tw]