Ann Inst Stat Math (2010) 62:835–853
DOI 10.1007/s10463-008-0192-2

Nonparametric analysis of doubly truncated data

Pao-sheng Shen

Received: 24 April 2007 / Revised: 5 November 2007 / Published online: 26 August 2008

Abstract One of the principal goals of the quasar investigations is to study luminosity evolution. A convenient one-parameter model for luminosity says that the expected log luminosity, T^*, increases linearly as $\theta_0 \cdot \log(1+Z^*)$, and $T^*(\theta_0) = T^* - \theta_0 \cdot \log(1+Z^*)$ is independent of Z^*, where Z^* is the redshift of a quasar and θ_0 is the true value of evolution parameter. Due to experimental constraints, the distribution of T^* is doubly truncated to an interval (U^*, V^*) depending on Z^*, i.e., a quadruple (T^*, Z^*, U^*, V^*) is observable only when $U^* \le T^* \le V^*$. Under the one-parameter model, $T^*(\theta_0)$ is independent of $(U^*(\theta_0), V^*(\theta_0))$, where $U^*(\theta_0) = U^* - \theta_0 \cdot \log(1+Z^*)$ and $V^*(\theta_0) = V^* - \theta_0 \cdot \log(1+Z^*)$. Under this assumption, the nonparametric maximum likelihood estimate (NPMLE) of the hazard function of $T^*(\theta_0)$ (denoted by $\hat{\mathbf{h}}$) was developed by Efron and Petrosian (J Am Stat Assoc 94:824–834, 1999). In this note, we present an alternative derivation of $\hat{\mathbf{h}}$. Besides, the NPMLE of distribution function of $T^*(\theta_0)$, $\hat{F}$, will be derived through an inverse-probability-weighted (IPW) approach. Based on Theorem 3.1 of Van der Laan (1996), we prove the consistency and asymptotic normality of the NPMLE $\hat{F}$ under certain condition. For testing the null hypothesis $H_{\theta_0} : T^*(\theta_0) = T^* - \theta_0 \cdot \log(1+Z^*)$ is independent of Z^*, (Efron and Petrosian in J Am Stat Assoc 94:824–834, 1999). proposed a truncated version of the Kendall's tau statistic. However, when T^* is exponential distributed, the testing procedure is futile. To circumvent this difficulty, a modified testing procedure is proposed. Simulations show that the proposed test works adequately for moderate sample size.

Keywords Double truncation · Nonparametric MLE · Kendall's tau