AISM (2007) 59: 617–654
DOI 10.1007/s10463-006-0067-3

Sam Efromovich

Optimal nonparametric estimation of the density of regression errors with finite support

Received: 30 March 2005 / Revised: 26 December 2005 /
Published online: 25 July 2006

Abstract Knowledge of the probability distribution of error in a regression problem plays an important role in verification of an assumed regression model, making inference about predictions, finding optimal regression estimates, suggesting confidence bands and goodness of fit tests as well as in many other issues of the regression analysis. This article is devoted to an optimal estimation of the error probability density in a general heteroscedastic regression model with possibly dependent predictors and regression errors. Neither the design density nor regression function nor scale function is assumed to be known, but they are suppose to be differentiable and an estimated error density is suppose to have a finite support and to be at least twice differentiable. Under this assumption the article proves, for the first time in the literature, that it is possible to estimate the regression error density with the accuracy of an oracle that knows "true" underlying regression errors. Real and simulated examples illustrate importance of the error density estimation as well as the suggested oracle methodology and the method of estimation.

Keywords and phrases Adaptation · Error depending on predictor · Heteroscedasticity · Minimax · Pinsker oracle