

Kernel-Density Estimation with Missing Data and Auxiliary Variables

Suzanne R. Dubnicka

Kansas State University

E-Mail: dubnicka@ksu.edu

Abstract: In most parametric statistical analyses, knowledge of the distribution of the response variable, or of the errors, is important. As this distribution is not typically known with certainty, one might initially construct a histogram or estimate the density of the variable of interest to gain insight regarding the distribution and its characteristics. However, when the response variable is incomplete, a histogram will only provide a representation of the distribution of the observed data. In the AIDS Clinical Trial Study protocol 175, interest lies in the difference in CD4 counts from baseline to final follow-up, but CD4 counts collected at final follow-up were incomplete. Therefore, we propose methods for estimating the density of an incomplete response variable when auxiliary data are available. Properties of the proposed density estimator will be presented. Simulation studies verify these properties and show that the proposed density estimator performs well.