Application of Nonparametric Quantile Regression to Body Mass Index Percentile Curves from Survey Data

Yan Li[1,2]*
[1]Joint Program in Survey Methodology, University of Maryland, College Park, 20742,
yli@survey.umd.edu

Barry I. Graubard[2]
[2]Biostatistics Branch, National Cancer Institute, Bethesda MD 20892

Edward L. Korn[3]
[3]Biometric Research Branch, National Cancer Institute, Bethesda, MD 20892,

Increasing rates of overweight among children in the U.S. stimulated interest in obtaining national percentile curves of body size to serve as a benchmark in assessing growth development in clinical and population settings. In 2000, the U.S. Centers for Disease Control and Prevention (CDC) developed conditional percentile curves for Body mass index (BMI) for ages 2-20 years. The 2000 CDC BMI-for-age curves are partially parametric and only partially incorporated the survey sample weights in the curve estimation. As a result, they may not fully reflect the underlying pattern of BMI-for-age in the population. This motivated us to develop a nonparametric double-kernel-based method and automatic bandwidth selection procedure. We include sample weights in the bandwidth selection, conduct median correction to reduce small sample smoothing bias, and rescale the bandwidth to make it scale-invariant. Using this procedure we re-estimate the national percentile BMI-for-age curves and the prevalence of high-BMI children in the U.S.

Key Words: Bandwidth selection, Conditional percentile, Kernel estimator, Local linear regression, National Health and Nutrition Examination Survey, National Health Examination Survey.