A Spline-based Lack-of-Fit Test for Linear Models

Chin-Shang Li
Department of Public Health Sciences, Division of Biostatistics,
University of California, Davis

Abstract

An unspecified smooth function is used to depict the effect of a continuous covariate in a partially linear model. The smooth function is approximated by a linear combination of fixed-knot cubic *B*-splines with a second-order difference penalty on the adjacent cubic *B*-spline coefficients in order to prevent over-fitting. The penalized least-squares method is used to estimate the model parameters. Based on the resulting penalized least-squares estimator, two Wald-type spline-based test statistics are constructed to test the null hypothesis of the linearity of the smooth function in which the partially linear model becomes a linear model. It can be shown that the distribution of the first test statistic asymptotically is the distribution of a linear combination of independent chi-squared random variables, each with one degree of freedom, when the number of knots is fixed and under the null hypothesis. In practice, to determine a value of the smoothing parameter, the expected value of the first test statistic under the null hypothesis is set equal to a specified value. When the number of knots is fixed, the limiting null distribution of the second test statistic is a chi-squared distribution with $q + 2$ degrees of freedom, where q is the number of knots used in the cubic *B*-splines. The power performances of the two proposed tests are studied with simulations. The practical use of the proposed methodology is illustrated with a real-life data set.