

Abstract

Regression is a commonly used method to explore the relationship between a dependent variable Y and one or more explanatory variables X . In ordinary regression, the focus is on estimation the conditional mean of the response variable Y given $X=x$. In contrast, quantile regression(Koenker and Bassett 1978) is a method of estimating the relationship between X and the conditional quantiles of Y given $X=x$. Hence, quantile regression provides a more complete description of the conditional distribution of Y given $X=x$ if all quantiles are of the interest. Since quantile regression is not covered in the undergraduate statistics curriculum in NUS, this project requires me to pick up an operational understanding of 3 different quantile regression methods: the asymmetric L1 estimation method of Koenker and Bassett, the local linear method and the kernel density estimation method and implement them to analyze a set of environmental chemical exposure data collected from a biomonitoring study. An imputation-based procedure is also explored when, for the sake of saving costs, individuals are pooled together and only their aggregate exposure is known but not their individual exposures.