Iterative method for the reducing the impact of outlying data points: ensuring data completeness

Svetlana Jesiļevska
Senior officer
06.09.1986, 28
Central statistical bureau of Latvia
www.csb.gov.lv
Riga, Latvia
mozir@inbox.lv

Abstract

Data editing is essential to check the survey data for possible data problems. Outlying data values are frequently encountered in sample surveys. Consequently, in working with data, the correctness of the reported values must be verified, and if a reported value constitutes an outlier, its appropriate treatment needs to be considered. In this paper, the Iterative method for the reducing the impact of outlying data points that can cope with incomplete data is proposed. Data quality dimension completeness is not mentioned among quality criteria defined by Eurostat, nevertheless this dimension is very important. In this paper, the author justifies the importance of data completeness dimension as well as proposes the method that aims to improve statistical data completeness. The novelty of the Iterative method for the reducing the impact of outliers is the following: an iterative approach for determining the outlying data points is proposed; outliers are determined considering the impact of complex factors; estimation of weight coefficients of the outliers and estimation of the total measurement error of the non-linear factor’s model is carried out.