Investigating Outliers Detection Methods for the Iranian Manufacturing Establishment Survey Data

Rezaei Ghahroodi Zahra*
Statistical Research and Training Center, Tehran, Iran z_rezaee@sbu.ac.ir

Baghfalaki, Taban
Shahid Beheshti University, Tehran, Iran tbaghfalaki@yahoo.com

Ganjali, Mojtaba
Shahid Beheshti University, Tehran, Iran m-ganjali@sbu.ac.ir

The role and importance of the industrial sector in the economic development specify the necessity of having accurate and timely data for exact planning. As outliers data in establishment surveys are common due to the structure of the economy, the evaluation of survey data by identifying and investigating outliers prior to the release of data is necessary. In this paper the practical application of different robust multivariate outlier detection methods based on the Mahalanobis distance with BACON algorithm, minimum volume ellipsoid (MVE) estimator, minimum covariance determinant (MCD) estimator, Stahel-Donoho estimator is presented. Also some univariate outlier detection methods such as Hadi and Simonoff (1993) method, using some regression models, are presented. These methods are illustrated using a real data set on Iranian Manufacturing Establishment Survey (IMES). These data are collected each year by the Statistical Center of Iran using sampling weights. In this paper it is demonstrated that the use of different robust outlier detection methods (multivariate and univariate), in a number of manufacturing industries, leads to the same results.

Key Words: robust multivariate outlier detection; sampling weight; Winsorization; Mahalanobis distance