Monitoring Photochemical Pollutants based on Symbolic Interval-Valued Data Analysis

Liang-Ching Lin*
Department of Statistics, National Cheng Kung University, Tainan, Taiwan.
Meihui Guo
Department of Applied Mathematics, National Sun Yat-sen University, Kaohsiung, Taiwan
Sangyeol Lee
Department of Statistics, Seoul National University, Seoul, 08826, South Korea.

Abstract

This study considers monitoring photochemical pollutants for anomaly detection based on symbolic interval-valued data analysis. For this task, we construct control charts based on the principal component scores of symbolic interval-valued data. Herein, the symbolic interval-valued data are assumed to follow a normal distribution, and an approximate expectation formula of order statistics from the normal distribution is used in the univariate case to estimate the mean and variance via the method of moments. In addition, we consider the bivariate case wherein we use the maximum likelihood estimator calculated from the likelihood function derived under a bivariate copula. We also establish the procedures for the statistical control chart based on the univariate and bivariate interval-valued variables, and the procedures are potentially extendable to higher dimensional cases. Monte Carlo simulations and real data analysis using photochemical pollutants confirm the validity of the proposed method. The results particularly show the superiority over the conventional method that uses the averages to identify the date on which the abnormal maximum occurred.

Keywords: Control chart; Monitoring photochemical pollutants; Symbolic principal component analysis; Symbolic data analysis