

Distribution and robustness of a distance-based multivariate coefficient of variation

AERTS Stéphanie, HEC-ULg, University of Liège, stephanie.aerts@ulg.ac.be
Joint work with HAESBROECK Gentiane and RUWET Christel

BSS 2014, November 6-7

When one wants to compare the homogeneity of a characteristic in several populations that have different means, the advocated statistic is the univariate coefficient of variation. However, in the multivariate setting, comparing marginal coefficients may be inconclusive. Therefore, several extensions that summarize multivariate relative dispersion in one single index have been proposed in the literature (see Albert & Zhang, 2010, for a review).

In this talk, focus is on a particular extension, due to Voinov & Nikulin (1996), based on the Mahalanobis distance between the mean and the origin of the design space. Some arguments are outlined for justifying this choice. Then, properties of its sample version under elliptical symmetry are discussed. Under normality, this estimator is shown to be biased at finite samples. In order to overcome this drawback, two bias corrections are proposed and compared by means of simulations.

Moreover, the empirical estimator also suffers from a lack of robustness, which is illustrated by means of influence functions. A robust counterpart based on the Minimum Covariance Determinant estimator is advocated.

References

- Aerts, S., Haesbroeck, G. and Ruwet, C. (2014a) Multivariate coefficients of variation: comparison and influence functions. *Submitted for publication*.
- Aerts, S., Haesbroeck, G. and Ruwet, C. (2014b) Distribution under elliptical symmetry of a distance-based multivariate coefficient of variation. *Submitted for publication*.
- Albert, A. and Zhang, L. (2010). A novel definition of the multivariate coefficient of variation. *Biometrical Journal*, 52, 667–675.
- Voinov V.G. and Nikulin M.S. (1996). *Unbiased estimators and their applications*. Vol. 2, multivariate case. Dordrecht: Kluwer.