

To impute or not to? A multivariate goodness-of-fit testing perspective

Danijel Aleksić

University of Belgrade, Faculty of Organizational Sciences, Faculty of Mathematics
e-mail: danijel.aleksic@fon.bg.ac.rs

Bojana Milošević

University of Belgrade, Faculty of Mathematics
e-mail: bojana.milosevic@matf.bg.ac.rs

Abstract. A multivariate normality assumption is a crucial for validity of many methods of statistical inference. Therefore, there are many proposed statistical tests for testing the mentioned assumption. However, all of the currently available tests are suitable for complete samples. When the data are not complete, i.e. some of the values are missing, one needs to adapt the existing methodology to overcome this issue. Here, we consider several approaches for usage of BHEP test for testing the multivariate normality in the context of incomplete datasets with various missingness mechanisms. We explore behavior of each of them for large sample sizes, i.e. asymptotically, as well as for small sample sizes in an extensive empirical study.

Keywords: missing data, empirical characteristic function, weighted L^2 test

References

- [1] **L. Baringhaus, N. Henze.** A consistent test for multivariate normality based on the empirical characteristic function. *Metrika*, 1988, pages 339–348.
- [2] **D. Alekić, M. Cuparić, B. Milošević.** Non-degenerate U-statistics for data missing completely at random with application to testing independence. *Stat*, 2023, 12 (1).