Proceedings of the 4th International Conference on Statistics: Theory and Applications (ICSTA'22)
Prague, Czech Republic – July 28- 30, 2022
Paper No. 157
DOI: 10.11159/icsta22.157

On Consistent Hypothesis Testing In General Hilbert Spaces

Daniel Gaigall[1]
[1]Mathematical Institute-
University of Koblenz and Landau, Campus Koblenz
Universitätsstraße 1, 56070 Koblenz, Germany
gaigall@uni-koblenz.de

Extended Abstract

Inference on the basis of high-dimensional and functional data are two topics which are discussed frequently in the current statistical literature. A possibility to include both topics in a single approach is working on a very general space for the underlying observations, such as a separable Hilbert space. We propose a general method for consistently hypothesis testing on the basis of random variables with values in separable Hilbert spaces. We avoid concerns with the curse of dimensionality due to a projection idea. We apply well-known test statistics from nonparametric inference to the projected data and integrate over all projections from a specific set and with respect to suitable probability measures. In contrast to classical methods, which are applicable for real-valued random variables or random vectors of dimensions lower than the sample size, the tests can be applied to random vectors of dimensions larger than the sample size or even to functional and high-dimensional data. In general, resampling procedures such as bootstrap or permutation are suitable to determine critical values. The idea can be extended to the case of incomplete observations. Moreover, we develop an efficient algorithm for implementing the method.

Examples are given for testing goodness-of-fit in a one-sample situation in [1] or for testing marginal homogeneity on the basis of a paired sample in [2]. Here, the test statistics in use can be seen as generalizations of the well-known Cramér-von-Mises test statistics in the one-sample and two-samples case. The treatment of other testing problems is possible as well. By using the theory of U-statistics, for instance, asymptotic null distributions of the test statistics are obtained as the sample size tends to infinity. Standard continuity assumptions ensure the asymptotic exactness of the tests under the null hypothesis and that the tests detect any alternative in the limit. Simulation studies demonstrate size and power of the tests in the finite sample case, confirm the theoretical findings, and are used for the comparison with concurring procedures.

A possible application of the general approach is inference for stock market returns, also in high data frequencies. In the field of empirical finance, statistical inference of stock market prices usually takes place on the basis of related log-returns as data. In the classical models for stock prices, i.e., the exponential Lévy model, Black-Scholes model, and Merton model, properties such as independence and stationarity of the increments ensure an independent and identically structure of the data. Specific trends during certain periods of the stock price processes can cause complications in this regard. In fact, our approach can compensate those effects by the treatment of the log-returns as random vectors or even as functional data.

References

[1] M. Ditzhaus and D. Gaigall, “A consistent goodness-of-fit test for huge dimensional and functional data,” Journal of Nonparametric Statistics, vol. 30, no. 4, pp. 834-859, 2018.
[2] M. Ditzhaus and D. Gaigall, “Testing marginal homogeneity in Hilbert spaces with applications to stock market returns,” TEST, 2022.