Uniform-over-dimension asymptotic theory with an application to high-dimensional testing of locations

Subhajit Dutta
Indian Statistical Institute, Kolkata Centre

Abstract

Asymptotic methods for hypothesis testing in high-dimensional data usually require the dimension of the observations to increase to infinity, often with an additional relationship between the dimension (say, p) and the sample size (say, n). On the other hand, multivariate asymptotic testing methods are valid for fixed dimension only and their implementations typically require the sample size to be large compared to the dimension to yield desirable results. In practical scenarios, it is usually not possible to determine whether the dimension of the data conforms to the conditions required for the validity of the high-dimensional asymptotic methods for hypothesis testing, or whether the sample size is large enough compared to the dimension of the data. In this work, we first describe the notion of uniform-over-p convergences and subsequently, develop a uniform-over-dimension central limit theorem. An asymptotic test for the two-sample equality of locations is developed, which now holds uniformly over the dimension of the observations. Using simulated and real data, it is demonstrated that the proposed test exhibits better performance compared to several popular tests in the literature for high-dimensional data as well as the usual scaled two-sample tests for multivariate data, including the Hotelling's T^2 test for multivariate Gaussian data.