M. Rauf Ahmad
Uppsala University

U-tests for high-dimensional two-sample mean testing problems under non-normality and Behrens-Fisher setting

Test statistics for two-sample general linear hypothesis, particularly focusing on profile analysis and the usual multivariate hypothesis of the difference of mean vectors, are considered when the dimension, p, may exceed the sample sizes, n_1, n_2, the underlying distributions may not necessarily be normal, and the covariance matrices may also be unequal (the Behrens-Fisher problem), with the sample sizes simultaneously kept unequal. Under mild assumptions on the traces of the covariance matrices, the statistics are shown to be asymptotically chi-square distributed when $n_i, p \to \infty$. The tests are referred to as U-tests, since they are constructed as linear combinations of one- and two-sample first-order degenerate U-statistics, and consequently, their limit distributions are based on the asymptotic theory of degenerate U-statistics. Through extensive simulation studies, under a wide variety of parameter settings, the test statistics are shown to be accurate for both size control and power, even with sample sizes as moderate as 10 or 20, and any dimension. Practical applications of the test statistics are also illustrated using real data sets from different fields of study.