



Contribution ID: 26

Type: **not specified**

Combined tests for high-dimensional industrial data

Monday, 17 May 2021 15:40 (20 minutes)

A class of multivariate tests for the two-sample location problem with high-dimensional low sample size data and with complex dependence structure, that are increasingly common in industrial statistics, is described. The tests can be applied when the number of variables is much larger than the number of objects, and when the underlying population distributions are heavy-tailed or skewed.

The tests are based on combination of tests based on interpoint distances. It is proved that the tests are exact, unbiased and consistent. It is also shown that the tests are very powerful under normal, heavy-tailed and skewed distributions.

The tests can be applied also to fields other than industrial statistics, such as to case-control studies with high-dimensional low sample size data from medical imaging techniques (like magnetic resonance, computed tomography or X-ray radiography), chemometrics and microarray data (proteomics, transcriptomics). Moreover, the tests are very promising for designing new multivariate control charts.

Primary author: MAROZZI, Marco (Ca' Foscari University of Venice (Italy))

Presenter: MAROZZI, Marco (Ca' Foscari University of Venice (Italy))

Session Classification: Modelling / DoE for optimization

Track Classification: Data Science in Process Industries