Kernel Cumulants*

Zoltán Szabó†

Abstract

Mean and covariance are omnipresent in data science to capture (a limited form of) difference and independence of random quantities, with a wide range of successful applications. In this talk, I am going to present a nonlinear higher-order variant of these measures which (i) are able to characterize the equality of distributions and independence under mild conditions, (ii) are easy to estimate with minimal computational overhead compared to their degree one counterparts, (iii) achieve improved power in two-sample and independence testing as illustrated in environmental and traffic data analysis.

Paper: link.

*Data Science Seminar, University of York, UK (Apr. 30, 2025); abstract.

†This is joint work with Patric Bonnier and Harald Oberhauser.