Abstracts of the Ninth Workshop on Simulation (2018) 1-2

The Hellinger dependence measure and its root-n estimation

Gery Geenens, **Pierre Lafaye de Micheaux** and Spiridon Penev [1]

Testing for the independence between two or more random variables or vectors is a very active topic of statistical research; see for example [2] for a recent survey.

In his seminal paper [3], Rényi formulated 7 axioms that an ideal dependence measure should satisfy. However, these axioms are too strong for general use. In this work, we give a list of reasonable properties that any dependence measure between two random variables should satisfy.

This leads us to propose the following new dependence measure between two random variables X_1 and X_2:

$$\Delta(X_1, X_2) = 1 - \int\int_{\mathcal{I}} \sqrt{c_{12}}(u_1, u_2) du_1 du_2,$$

where $\mathcal{I} = [0, 1] \times [0, 1]$. This dependence measure is based on the Hellinger distance. It involves the square root of the joint copula density of X_1 and X_2.

We show how to obtain a root-n consistent estimator of our new Hellinger dependence measure by borrowing ideas from the k-nearest neighbours technique. This estimator is then used to build a new test of independence between two random variables. Empirical simulations show the good performance of our new approach against several competitors, such as [4]and [1].

References

[1] Ruth Heller, Yair Heller, and Malka Gorfine. A consistent multivariate test of association based on ranks of distances. *Biometrika*, 100(2):503–510, 2013.

[2] Julie Josse and Susan Holmes. Measuring multivariate association and beyond. *Stat. Surv.*, 10:132–167, 2016.

[3] A. Rényi. On measures of dependence. *Acta Math. Acad. Sci. Hungar.*, 10:441–451 (unbound insert), 1959.

[4] Gábor J. Székely, Maria L. Rizzo, and Nail K. Bakirov. Measuring and testing dependence by correlation of distances. *Ann. Statist.*, 35(6):2769–2794, 2007.

[1]School of Mathematics and Statistics, UNSW Sydney, NSW 2052, Australia, E-mail: lafaye@unsw.edu.au