Advances in goodness-of-fit tests based on energy statistics for right censored data

María Vidal-García[1], Rosa M. Crujeiras[1] and Wenceslao González-Manteiga[1]

[1]CITMAga, 15782 Santiago de Compostela, Spain. Department of Statistics, Mathematical Analysis and Optimization. Universidade de Santiago de Compostela (USC).

ABSTRACT

When proposing goodness-of-fit tests, a well-established approach from a nonparametric perspective is to use some kind of distance between functions (e.g. Kolmogorov-Smirnov test, Cramer von Mises). In that framework, using a weighted L2 distance between distributions leads to a particular measure which is closely connected with energy distance. Due to its interesting properties, a wide variety of energy distance-based tests have been proposed such as test of independence (Székely et al., 2007), test for comparing populations or goodness-of-fit test (Székely and Rizzo, 2017) among others.

In the context of Survival Analysis there are usually restrictions when observing the variable of interest in a sample, that is, the survival times. Therefore, collected data will combine complete observations and partial or censored observations, and specific tools are required to appropriately deal with this kind of samples. A particularly popular scenario is that of random right censoring. Previously mentioned energy distance-based tests can be adapted to this context as shown in Edelmann et al. (2022) for the test of independence.

In this talk, an energy distance-based goodness-of-fit test for right censored data will be presented. Implementations as a permutation test based on both complete and incomplete U-statistics (Janson, 1984) is discussed and their performance is assessed via simulations.

REFERENCES

Datta, S., Bandyopadhyay, D. and Satten, G. A. (2010) Inverse probability of censoring weighted U-statistics for right-censored data with an application to testing hypothesis. Scandinavian Journal of Statistics, 37(4), 680-700.

Edelmann, D., Welchowski, T. and Benner, A. (2022) A consistent version of distance covariance for right-censored survival data and its application in hypothesis testing. Biometrics, 78(3), 867-879.

Janson, S. (1984) The asymptotic distributions of incomplete U-statistics. Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete, 66(4), 495-505.

Székely, G. J. and Rizzo, M. L. (2007) Measuring and testing dependence by correlation of distances. Annals of Statistics, 35(6), 2769-2794.

Székely, G. J. and Rizzo, M. L. (2017) The energy of data. Annual Review of Statistics and Its Application, 4, 447-479.