Random projections and goodness of fit tests for multidimensional data

J.A. Cuesta-Albertos, cuestaj@unican.es

Universidad de Cantabria, Spain.

Abstract

Counterexamples showing two different multidimensional distributions with a common marginal are very well known. The usual way to construct those counterexamples starts by fixing the marginal and, then, constructing two different distributions sharing this marginal. In [1] a different point of view is taken. There, the authors begin by having two multidimensional distributions P and Q, and, then, they consider the following problem: Given a continuous probability measure μ (for instance, gaussian), which is the μ-measure of the vectors, h, which satisfy that the (one-dimensional) marginals of P and Q along the line determined by h coincide? The answer is 1 if $P = Q$ and 0 if P and Q are different. Two sample goodness-of-fit tests follow straightforward from this result. Moreover, it was shown in [2] that this result can be extended to cover some families of distributions, thus providing ways to construct goodness-of-fit tests to those families. In particular, there it is shown that a distribution is gaussian if and only if almost every (one-dimensional) projection is gaussian. In this talk I will comment those results and will present some applications to real data sets.

References

[1] Cuesta-Albertos, J.A., R. Fraiman and T. Ransford (2007). *A sharp form of the Cramér-Wold theorem.* To appear in J. Theoret. Probab.

[2] Cuesta-Albertos, J.A., E. del Barrio, R. Fraiman and C. Matrán (2007). *The random projection method in goodness of fit for functional data.* To appear in Comput. Statist. and Data Anal.