

Optimal nonparametric testing of Missing Completely At Random, and its connections to compatibility

Tom Berret

University of Warwick

Abstract:

Given a set of incomplete observations, we study the nonparametric problem of testing whether data are Missing Completely At Random (MCAR). Our first contribution is to characterise precisely the set of alternatives that can be distinguished from the MCAR null hypothesis. This reveals interesting and novel links to the theory of

Frechet classes (in particular, compatible distributions) and linear programming, and we leverage tools developed in these fields to propose MCAR tests that are consistent against all detectable alternatives. Moreover, we define a natural measure of ease of detectability (an incompatibility index), and exploit ideas from max-flow min-cut theory to prove that our tests achieve the optimal minimax separation rate according to this measure in certain cases.