

Abstract: “Testing Ising Models”

Given samples from an unknown multivariate distribution, is it possible to distinguish whether it is the product of its marginals versus being far from every product distribution? Similarly, is it possible to distinguish whether it equals a given distribution versus being far from it? These problems of testing independence and goodness-of-fit have received enormous attention in statistics, information theory, and theoretical computer science, with sample-optimal algorithms known in several interesting regimes of parameters. Unfortunately, it has also been understood that these problems become intractable in large dimensions, necessitating exponential sample complexity.

Motivated by exponential lower bounds for general distributions as well as the ubiquity of Markov Random Fields (MRFs) in modeling high-dimensional distributions, we initiate the study of distribution testing on structured multivariate distributions, particularly the prototypical example of MRFs: the Ising Model. We demonstrate that, in this structured setting, we can avoid the curse of dimensionality, obtaining sample and time efficient testers for independence and goodness-of-fit. Along the way, we develop new tools for establishing concentration of functions of the Ising model, using and extending Chatterjee's exchangeable pairs framework. In particular, we prove tighter concentration results for multi-linear functions of the Ising model in the high-temperature regime.

Preprint: <https://arxiv.org/abs/1612.03147>.