

# Sensitive $\chi^2$ testing via sampling tripartite 3-uniform hypergraphs

András Hubai<sup>a,\*</sup>, Tamás Róbert Mezei<sup>a</sup>, Ferenc Béres<sup>b</sup>,  
András Benczúr<sup>b,c</sup>, István Miklós<sup>a,b</sup>

<sup>a</sup>HUN-REN Rényi Institute, 1053 Budapest, Reáltanoda u. 13-15.  
Hungary; <sup>b</sup>HUN-REN SZTAKI, 1111 Budapest, Lágymányosi u. 11.  
Hungary; <sup>c</sup>Széchenyi University, 9026 Győr, Egyetem tér 1. Hungary  
\*hubaiandras@gmail.com

When assessing the independence of two categorical variables, the standard approach involves their bipartite interaction graphs transformed into contingency tables and analyzed using the  $\chi^2$  test. With three categorical variables, the interaction graph is a tripartite 3-uniform hypergraph, leading to a 3D contingency table. The  $\chi^2$  test remains applicable to assess the independence of any two variables of the trio through a 2D projection of the contingency table.

In this presentation, we introduce a more sensitive statistical test grounded in hypergraph theory. Specifically, we propose a hypergraph-based exact test that compares a  $\chi^2$  aggregation metric of the above interaction graph with a random sample of hypergraphs that share the same degree distribution.

In related research, the authors established the NP-hardness of sampling tripartite 3-uniform hypergraphs with prescribed degree distributions[1]. To address this challenge, we present a practical parallel tempering-based sampling method. We demonstrate, both on synthetic and real-world datasets, that the hypergraph-based exact  $\chi^2$  test consistently outperforms the conventional  $\chi^2$  test.

## References

- [1] arXiv:2308.13251 (2023)