Proceedings of Machine Learning Research vol 291:1–2, 2025 38th Annual Conference on Learning Theory

Testing (Conditional) Mutual Information

Jan Seyfried JAN.SEYFRIED@U.NUS.EDU
Centre for Quantum Technologies, National University of Singapore

Sayantan Sen SAYANTAN789@GMAIL.COM
Centre for Quantum Technologies, National University of Singapore

Marco Tomamichel MARCO.TOMAMICHEL@NUS.EDU.SG
Department of Electrical and Computer Engineering, Centre for Quantum Technologies, National University of Singapore

Editors: Nika Haghtalab and Ankur Moitra

We investigate the sample complexity of mutual information and conditional mutual information testing. For the conditional mutual information testing, given access to independent samples of a triplet of random variables (A, B, C) from an unknown distribution, we want to distinguish between two cases: (i) A and C are conditionally independent, i.e., $I(A\!:\!C|B) = 0$, and (ii) A and C are conditionally dependent, i.e., $I(A\!:\!C|B) \geq \varepsilon$ for some threshold $\varepsilon \in (0,1)$. We establish an upper bound on the number of samples required to distinguish between the two cases with high confidence, as a function of ε and the three alphabet sizes d_A, d_B, and d_C, respectively, where $d_A \geq d_C$,

$$\widetilde{O}\left(\max\left\{\min\left\{\frac{d_A^{\frac{3}{4}} d_B^{\frac{3}{4}} d_C^{\frac{1}{4}}}{\varepsilon}, \frac{d_A^{\frac{2}{3}} d_B^{\frac{2}{3}} d_C^{\frac{1}{3}}}{\varepsilon^{\frac{4}{3}}}\right\}, \frac{d_A^{\frac{1}{2}} d_B^{\frac{3}{4}} d_C^{\frac{1}{2}}}{\varepsilon}, \min\left\{\frac{d_A^{\frac{1}{4}} d_B^{\frac{7}{8}} d_C^{\frac{1}{4}}}{\varepsilon}, \frac{d_A^{\frac{2}{7}} d_B^{\frac{6}{7}} d_C^{\frac{2}{7}}}{\varepsilon^{\frac{8}{7}}}\right\}\right\}\right). \tag{1}$$

We further show tightness in the first two regimes. Work on the related problem of conditional independence testing in the ℓ_1 distance, Canonne et al. (2018), provides partial lower bounds which carry over to our setting. Together, they show that the intricate structure of our sample complexity is indeed necessary, and we further conjecture that the scaling in the different parameters is tight. For the special case of mutual information testing (when B is trivial), we establish that the resulting sample complexity from (1) is also optimal, up to polylogarithmic terms. Our approach reduces testing for conditional independence problem to equivalence testing in D_H^2 distance between P_{ABC} and $Q_{ABC} := P_{A|B}P_{C|B}P_B$, which can be translated to a polylogarithmic number of instances of equivalence tests in the ℓ_2 distance using a bucketing technique (see Diakonikolas and Kane (2016)), which we adapt to account for the additional structure of our problem. Since sampling from Q_{ABC} using samples from P_{ABC} is not directly possible, we further require a case distinction into two regimes, depending on whether $P_{B=b}$ is above or below a certain threshold. In the small regime, we can only simulate Q_{ABC} approximately, leading to weakly correlated and slightly biased samples. We present a new estimator for equivalence testing that can handle such correlated samples, which might be of independent interest. The lower bounds from Canonne et al. (2018) show that the scaling in d_B and ε we achieve in the small regime are optimal, suggesting that our method of sampling cannot be fundamentally improved. Our estimator is also able to recover the optimal bounds for equivalence testing in the ℓ_2 distance, as proved in Chan et al. (2014). [1]

1. Extended abstract. Full version appears as arXiv:2506.03894, v1.

Acknowledgments

The authors would like to thank the anonymous reviewers for their suggestions, which improved the presentation of the paper. JS would like to thank Christopher Chubb for discussions on an early version of this problem. The authors thank Josep Lumbreras for proof reading the draft. This project is supported by the National Research Foundation, Singapore through the National Quantum Office, hosted in A*STAR, under its Centre for Quantum Technologies Funding Initiative (S24Q2d0009).

References

Clément L Canonne, Ilias Diakonikolas, Daniel M Kane, and Alistair Stewart. Testing conditional independence of discrete distributions. In *Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing*, pages 735–748, 2018.

Siu-On Chan, Ilias Diakonikolas, Paul Valiant, and Gregory Valiant. *Optimal Algorithms for Testing Closeness of Discrete Distributions*, pages 1193–1203. 2014. doi: 10.1137/1.9781611973402.88. URL https://epubs.siam.org/doi/abs/10.1137/1.9781611973402.88.

Ilias Diakonikolas and Daniel M. Kane. A new approach for testing properties of discrete distributions. In *2016 IEEE 57th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 685–694, 2016. doi: 10.1109/FOCS.2016.78. URL https://ieeexplore.ieee.org/document/7782983. ISSN: 0272-5428.