MR4900163 62G10 62H20

Gao, Zhe (PRC-HEF-SMG); **Wang, Roulin** (PRC-ECNU-SS);
Wang, Xueqin (PRC-HEF-SMG); **Zhang, Heping** (1-YALE-SPH)

Studentized tests of independence: random-lifter approach. (English. English summary)

Ann. Statist. **53** (2025), *no.* 2, 703–723.

In this article, a new approach to the testing for independence between two random objects is proposed. So far, a variety of independence tests have been developed, but a challenge lies in the nonstandard nature of their asymptotic distributions under the null of independence. The asymptotic null distributions generally depend on second-order Wiener chaos, and obtaining p-values from such distributions is computationally demanding.

To mitigate computational difficulties, the authors propose the random-lifter method. Suppose that we are interested in testing independence between two random objects $\mathbf{X}$ and $\mathbf{Y}$. Also suppose that there is another random variable or a random lifter Z, which is independent of $(\mathbf{X}, \mathbf{Y})$. A key idea for the random-lifter method is that testing independence between $\mathbf{X}$ and $\mathbf{Y}$ is equivalent to testing independence between $\mathbf{X}$ and $(\mathbf{Y}, Z)$. Due to independence between $\mathbf{X}$ and $(\mathbf{Y}, Z)$, the new dependence measure involving nonparametric kernel weights of the random lifter Z takes the form of a degenerate U-statistic. The martingale central limit theorem applies to the random-lifter test statistic, which can be obtained through a suitable Studentization of this dependence measure, and as a consequence, the test statistic has the standard normal limit under the null. Accordingly, p-values can be obtained directly from the standard normal table, and it is unnecessary to compute percentiles of asymptotic null distributions.

Moreover, it is demonstrated that the random-lifter test statistic has a minimax property. Implementation details for this independence test including algorithms and bandwidth selection are also discussed. Monte Carlo results indicate that the null distribution of the test statistic can be well approximated by the standard normal distribution for sample sizes of 500 or more, and that power properties of the proposed test are comparable to those of existing independence tests. *Masayuki Hirukawa*

[References]

1. Albert, M., Laurent, B., Marrel, A. and Meynaoui, A. (2022). Adaptive test of independence based on HSIC measures. *Ann. Statist.* **50** 858–879. MR4404921 https://doi.org/10.1214/21-aos2129 MR4404921
2. Deb, N. and Sen, B. (2023). Multivariate rank-based distribution-free nonparametric testing using measure transportation. *J. Amer. Statist. Assoc.* **118** 192–207. MR4571116 https://doi.org/10.1080/01621459.2021.1923508 MR4571116
3. Fukumizu, K., Gretton, A., Sun, X. and Schölkopf, B. (2007). Kernel measures of conditional dependence. In *Advances in Neural Information Processing Systems* (J. Platt, D. Koller, Y. Singer and S. Roweis, eds.) **20**. Curran Associates, Red Hook.
4. Gao, H. and Shao, X. (2023). Two sample testing in high dimension via maximum mean discrepancy. *J. Mach. Learn. Res.* **24** Paper No. [304], 33 pp. MR4664741 MR4664741
5. Gao, L., Fan, Y., Lv, J. and Shao, Q.-M. (2021). Asymptotic distributions of high-dimensional distance correlation inference. *Ann. Statist.* **49** 1999–2020. MR4319239 https://doi.org/10.1214/20-aos2024 MR4319239
6. Gao, Z., Wang, R., Wang, X. and Zhang, H. (2025). Supplement to “Studentized Tests of Independence: Random-Lifter approach.” https://doi.org/10.1214/

24-AOS2478SUPPA, https://doi.org/10.1214/24-AOS2478SUPPB

7. Gretton, A., Bousquet, O., Smola, A. and Schölkopf, B. (2005). Measuring statistical dependence with Hilbert-Schmidt norms. In *Algorithmic Learning Theory. Lecture Notes in Computer Science* **3734** 63–77. Springer, Berlin. MR2255909 https://doi.org/10.1007/11564089_7 MR2255909
8. Gretton, A., Fukumizu, K., Teo, C., Song, L., Schölkopf, B. and Smola, A. (2007). A kernel statistical test of independence. *Adv. Neural Inf. Process. Syst.* **20**.
9. Gretton, A., Herbrich, R., Smola, A., Bousquet, O. and Schölkopf, B. (2005). Kernel methods for measuring independence. *J. Mach. Learn. Res.* **6** 2075–2129. MR2249882 MR2249882
10. Gretton, A., Smola, A., Bousquet, O., Herbrich, R., Belitski, A., Augath, M., Murayama, Y., Pauls, J., Schölkopf, B. et al. (2005). Kernel constrained covariance for dependence measurement. In *Proceedings of the Tenth International Workshop on Artificial Intelligence and Statistics* (R. G. Cowell and Z. Ghahramani, eds.). *Proceedings of Machine Learning Research* **R5** 112–119. PMLR. Reissued by PMLR on 30 March 2021.
11. Kim, I., Balakrishnan, S. and Wasserman, L. (2020). Robust multivariate nonparametric tests via projection averaging. *Ann. Statist.* **48** 3417–3441. MR4185814 https://doi.org/10.1214/19-AOS1936 MR4185814
12. Kim, I., Balakrishnan, S. and Wasserman, L. (2022). Minimax optimality of permutation tests. *Ann. Statist.* **50** 225–251. MR4382015 https://doi.org/10.1214/21-aos2103 MR4382015
13. Li, T. and Yuan, M. (2024). On the optimality of Gaussian kernel based nonparametric tests against smooth alternatives. *J. Mach. Learn. Res.* **25** 1–62. MR4829179
14. Lopez-Paz, D., Hennig, P. and Schölkopf, B. (2013). The randomized dependence coefficient. *Adv. Neural Inf. Process. Syst.* **26**.
15. Lyons, R. (2013). Distance covariance in metric spaces. *Ann. Probab.* **41** 3284–3305. MR3127883 https://doi.org/10.1214/12-AOP803 MR3127883
16. Moignard, V., Woodhouse, S., Haghverdi, L., Lilly, A. J., Tanaka, Y., Wilkinson, A. C., Buettner, F., Macaulay, I. C., Jawaid, W. et al. (2015). Decoding the regulatory network of early blood development from single-cell gene expression measurements. *Nat. Biotechnol.* **33** 269–276.
17. Pan, W., Wang, X., Zhang, H., Zhu, H. and Zhu, J. (2020). Ball covariance: A generic measure of dependence in Banach space. *J. Amer. Statist. Assoc.* **115** 307–317. MR4078465 https://doi.org/10.1080/01621459.2018.1543600 MR4078465
18. Pfister, N., Bühlmann, P., Schölkopf, B. and Peters, J. (2018). Kernel-based tests for joint independence. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **80** 5–31. MR3744710 https://doi.org/10.1111/rssb.12235 MR3744710
19. Rigollet, P. and Tsybakov, A. B. (2007). Linear and convex aggregation of density estimators. *Math. Methods Statist.* **16** 260–280. MR2356821 https://doi.org/10.3103/S1066530707030052 MR2356821
20. Schweizer, B. and Wolff, E. F. (1981). On nonparametric measures of dependence for random variables. *Ann. Statist.* **9** 879–885. MR0619291 MR0619291
21. Scott, D. W. (1979). On optimal and data-based histograms. *Biometrika* **66** 605–610. MR0556742 https://doi.org/10.1093/biomet/66.3.605 MR0556742
22. Shekhar, S., Kim, I. and Ramdas, A. (2023). A permutation-free kernel independence test. *J. Mach. Learn. Res.* **24** Paper No. [369], 68 pp. MR4690318 MR4690318
23. Shi, H., Drton, M. and Han, F. (2022). Distribution-free consistent independence tests via center-outward ranks and signs. *J. Amer. Statist. Assoc.* **117** 395–410. MR4399094 https://doi.org/10.1080/01621459.2020.1782223 MR4399094

24. Silverman, B. W. (2018). *Density Estimation for Statistics and Data Analysis. Monographs on Statistics and Applied Probability*. CRC Press, London. MR0848134 https://doi.org/10.1007/978-1-4899-3324-9 MR0848134
25. Sriperumbudur, B. K., Gretton, A., Fukumizu, K., Schölkopf, B. and Lanckriet, G. R. G. (2010). Hilbert space embeddings and metrics on probability measures. *J. Mach. Learn. Res.* **11** 1517–1561. MR2645460 MR2645460
26. Székely, G. J. and Rizzo, M. L. (2013). The distance correlation t-test of independence in high dimension. *J. Multivariate Anal.* **117** 193–213. MR3053543 https://doi.org/10.1016/j.jmva.2013.02.012 MR3053543
27. Székely, G. J., Rizzo, M. L. and Bakirov, N. K. (2007). Measuring and testing dependence by correlation of distances. *Ann. Statist.* **35** 2769–2794. MR2382665 https://doi.org/10.1214/009053607000000505 MR2382665
28. Wang, X., Zhu, J., Pan, W., Zhu, J. and Zhang, H. (2023). Nonparametric statistical inference via metric distribution function in metric spaces. *J. Amer. Statist. Assoc.* **119** 2772–2784. MR4833914
29. Wen, C., Wang, X., Gao, Z. and Jiang, Y. (2023). High-dimensional robust test of independence via Grothendieck's covariance. Technical report.
30. Zhang, K. (2019). BET on independence. *J. Amer. Statist. Assoc.* **114** 1620–1637. MR4047288 https://doi.org/10.1080/01621459.2018.1537921 MR4047288
31. Zhang, Q., Filippi, S., Gretton, A. and Sejdinovic, D. (2018). Large-scale kernel methods for independence testing. *Stat. Comput.* **28** 113–130. MR3741641 https://doi.org/10.1007/s11222-016-9721-7 MR3741641

Note: This list reflects references listed in the original paper as accurately as possible with no attempt to correct errors.