The Journal of Open Source Software

dit: a Python package for discrete information theory

Ryan G. James[1], **Christopher J. Ellison**[1], **and James P. Crutchfield**[1]

1 Complexity Sciences Center, Department of Physics, University of California at Davis

DOI: 10.21105/joss.00738

Software

- Review
- Repository
- Archive

Submitted: 27 April 2018
Published: 31 May 2018

Summary

dit("Dit: A Python Package for Discrete Information Theory. Available at: Https://Github.com/Dit/Dit" n.d.) is a Python package for the study of **d**iscrete **i**nformation **t**heory. Information theory is a mathematical framework for the study of quantifying, compressing, and communicating random variables (Cover and Thomas 2006)(MacKay 2003)(Yeung 2008). More recently, information theory has been utilized within the physical and social sciences to quantify how different components of a system interact. dit is primarily concerned with this aspect of the theory.

Information theory is a powerful extension to probability and statistics, quantifying dependencies among arbitrary random variables in a way tha tis consistent and comparable across systems and scales. Information theory was originally developed to quantify how quickly and reliably information could be transmitted across an arbitrary channel. The demands of modern, data-driven science have been coopting and extending these quantities and methods into unknown, multivariate settings where the interpretation and best practices are not known. For example, there are at least four reasonable multivariate generalizations of the mutual information, none of which inherit all the interpretations of the standard bivariate case. Which is best to use is context-dependent. dit implements a vast range of multivariate information measures in an effort to allow information practitioners to study how these various measures behave and interact in a variety of contexts. We hope that having all these measures and techniques implemented in one place will allow the development of robust techniques for the automated quantification of dependencies within a system and concrete interpretation of what those dependencies mean.

dit implements the vast majority of information measure defined in the literature, including entropies (Shannon(Cover and Thomas 2006), Renyi, Tsallis), multivariate mutual informations (co-information(Bell 2003)(McGill 1954), total correlation(Watanabe 1960), dual total correlation(Te Sun 1980)(Han 1975)(Abdallah and Plumbley 2012), CAEKL mutual information(Chan et al. 2015)), common informations (Gács-Körner(Gács and Körner 1973)(Tyagi, Narayan, and Gupta 2011), Wyner(Wyner 1975)(W. Liu, Xu, and Chen 2010), exact(Kumar, Li, and El Gamal 2014), functional, minimal sufficient statistic), and channel capacity(Cover and Thomas 2006). It includes methods of studying joint distributions including information diagrams, connected informations(Schneidman et al. 2003)(Amari 2001), marginal utility of information(Allen, Stacey, and Bar-Yam 2017), and the complexity profile(Y. Bar-Yam 2004). It also includes several more specialized modules including bounds on the secret key agreement rate(Maurer and Wolf 1997), partial information decomposition(Williams and Beer 2010), rate-distortion theory(Cover and Thomas 2006) & information bottleneck(Tishby, Pereira, and Bialek 2000), and others. Please see the dit homepage for a complete and up-to-date list.

Where possible, the implementations in dit support multivariate, conditional forms even if not defined that way in the literature. For example, dit implements the multivariate, conditional exact common information even though it was only defined for two variables.

References

Abdallah, S. A., and M. D. Plumbley. 2012. "A Measure of Statistical Complexity Based on Predictive Information with Application to Finite Spin Systems." *Physics Letters A* 376 (4). Elsevier:275–81. https://doi.org/10.1016/j.physleta.2011.10.066.

Allen, Benjamin, Blake C Stacey, and Yaneer Bar-Yam. 2017. "Multiscale Information Theory and the Marginal Utility of Information." *Entropy* 19 (6). Multidisciplinary Digital Publishing Institute:273. https://doi.org/10.3390/e19060273.

Amari, Shun-ichi. 2001. "Information Geometry on Hierarchy of Probability Distributions." *Information Theory, IEEE Transactions on* 47 (5). IEEE:1701–11. https://doi.org/10.1109/18.930911.

Bar-Yam, Y. 2004. "Multiscale Complexity/Entropy." *Advances in Complex Systems* 7 (01). World Scientific:47–63. https://doi.org/10.1142/S0219525904000068.

Bell, A. J. 2003. "The Co-Information Lattice." In *Proc. Fifth Intl. Workshop on Independent Component Analysis and Blind Signal Separation*, edited by S. Makino S. Amari A. Cichocki and N. Murata, ICA 2003:921–26. New York: Springer.

Chan, Chung, Ali Al-Bashabsheh, Javad B Ebrahimi, Tarik Kaced, and Tie Liu. 2015. "Multivariate Mutual Information Inspired by Secret-Key Agreement." *Proceedings of the IEEE* 103 (10). IEEE:1883–1913. https://doi.org/10.1109/JPROC.2015.2458316.

Cover, T. M., and J. A. Thomas. 2006. *Elements of Information Theory*. Second. New York: Wiley-Interscience. https://doi.org/10.1002/047174882X.

"Dit: A Python Package for Discrete Information Theory. Available at: Https://Github.com/Dit/Di n.d. Accessed April 27, 2018.

Gács, Peter, and János Körner. 1973. "Common Information Is Far Less Than Mutual Information." *Problems of Control and Information Theory* 2 (2):149–62.

Han, T. S. 1975. "Linear Dependence Structure of the Entropy Space." *Information and Control* 29. Elsevier:337–68. https://doi.org/10.1016/S0019-9958(75)80004-0.

Kumar, G. R., C. T. Li, and A. El Gamal. 2014. "Exact Common Information." In *Information Theory (Isit), 2014 Ieee International Symposium on*, 161–65. IEEE. https://doi.org/10.1109/ISIT.2014.6874815.

Liu, Wei, Ge Xu, and Biao Chen. 2010. "The Common Information of N Dependent Random Variables." In *Communication, Control, and Computing (Allerton), 2010 48th Annual Allerton Conference on*, 836–43. IEEE. https://doi.org/10.1109/ALLERTON.2010.5706995.

MacKay, D. J.C. 2003. *Information Theory, Inference and Learning Algorithms*. Cambridge university press.

Maurer, Ueli, and Stefan Wolf. 1997. "The Intrinsic Conditional Mutual Information and Perfect Secrecy." In *IEEE International Symposium on Information Theory*, 88–88. Citeseer. https://doi.org/10.1109/ISIT.1997.613003.

McGill, W. J. 1954. "Multivariate Information Transmission." *Psychometrika* 19 (2). Springer:97–116. https://doi.org/10.1109/TIT.1954.1057469.

Schneidman, E., S. Still, M. J. Berry, W. Bialek, and others. 2003. "Network Information and Connected Correlations." *Phys. Rev. Lett.* 91 (23). APS:238701. https://doi.org/10.1103/PhysRevLett.91.238701.

Te Sun, H. 1980. "Multiple Mutual Informations and Multiple Interactions in Frequency Data." *Information and Control* 46. Elsevier:26–45. https://doi.org/10.1016/S0019-9958(80)90478-7.

Tishby, Naftali, Fernando C Pereira, and William Bialek. 2000. "The Information Bottleneck Method." *arXiv Preprint Physics/0004057*.

Tyagi, H., P. Narayan, and P. Gupta. 2011. "When Is a Function Securely Computable?" *Information Theory, IEEE Transactions on* 57 (10). IEEE:6337–50. https://doi.org/10.1109/TIT.2011.2165807.

Watanabe, S. 1960. "Information Theoretical Analysis of Multivariate Correlation." *IBM Journal of Research and Development* 4 (1). IBM:66–82. https://doi.org/10.1147/rd.41.0066.

Williams, Paul L, and Randall D Beer. 2010. "Nonnegative Decomposition of Multivariate Information." *arXiv Preprint arXiv:1004.2515*.

Wyner, A. D. 1975. "The Common Information of Two Dependent Random Variables." *Information Theory, IEEE Transactions on* 21 (2). IEEE:163–79. https://doi.org/10.1109/TIT.1975.1055346.

Yeung, R. W. 2008. *Information Theory and Network Coding*. Springer. https://doi.org/10.1007/978-0-387-79234-7.