How Random is Random? A Topological Perspective

Zoltan Kocsis Marwan Fayed

Abstract

Non-randomness in seemingly random numbers can bias results in scientific simulation, gambling, and cryptography. Random number sequences produced by physical generators are expensive, irreproducible, and have a limited output rate [1]. Computational pseudo-random sequences have the appearance of being random despite being determined by mathematical rules [2]. The quality of individual sequences is assessed using statistical tests [3, 4]. However, there is a shortage of numerical measures for comparing pseudo-random sequences against each other [5]. Here, we show that this gap can be filled by using topological data analysis to extract spatial structure from any number sequence. We calculate the Persistent Homology [6, 7, 8] of known random sequences and find a specific common spatial structure present on every scale. Crucially, when evaluated against this structure, all pseudo-random sequences will differ at some scale. We use the scale at which the first departure occurs as a numerical indicator of quality. Results reveal that the inversive congruential generators EICG1 and EICG7 have detectable regularities in low dimensions, demonstrating that lattice results [9] are insufficient for ruling out other spatial structures. Our work establishes a link between topological data analysis and pseudo-random number generation, laying a foundation for future advancements in the reliability of random sampling and its applications.

*Both authors are with the University of Stirling, UK, {zak,mmf}@cs.stir.ac.uk .

References

[1] S. Pironio, A. Acín, S. Massar, A. Boyer de la Giroday, D. N. Matsukevich, P. Maunz, S. Olmschenk, D. Hayes, L. Luo, T. A. Manning, and C. Monroe. Random numbers certified by bell's theorem. *Nature*, 464(7291):1021–1024, apr 2010.

[2] Donald E. Knuth. *The Art of Computer Programming, Volume 2 (3rd Ed.): Seminumerical Algorithms.* Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 1997.

[3] Juan Soto. Statistical testing of random number generators. In *In: Proceedings of the 22nd National Information Systems Security Conference*, 1999.

[4] Lawrence E. Bassham, III, Andrew L. Rukhin, Juan Soto, James R. Nechvatal, Miles E. Smid, Elaine B. Barker, Stefan D. Leigh, Mark Levenson, Mark Vangel, David L. Banks, Nathanael Alan Heckert, James F. Dray, and San Vo. Sp 800-22 rev. 1a. a statistical test suite for random and pseudorandom number generators for cryptographic applications. Technical report, Gaithersburg, MD, United States, 2010.

[5] Peter Hellekalek. Inversive pseudorandom number generators: Concepts, results and links. In *Proceedings of the 1995 Winter Simulation Conference*, pages 255–262. IEEE Press, 1995.

[6] Peter Bubenik and Jonathan A. Scott. Categorification of persistent homology. *Discrete & Computational Geometry*, 51(3):600–627, 2014.

[7] Herbert Edelsbrunner and John Harer. *Persistent Homology - a survey*, volume 453, pages 257–282. American Mathematical Society, 2008.

[8] Andrew Tausz, Mikael Vejdemo-Johansson, and Henry Adams. JavaPlex: A research software package for persistent (co)homology. In Han Hong and Chee Yap, editors, *Proceedings of ICMS 2014*, Lecture Notes in Computer Science 8592, pages 129–136, 2014.

[9] Harald Niederreiter. On a new class of pseudorandom numbers for simulation methods. *Journal of Computational and Applied Mathematics*, 56(1):159 – 167, 1994.