UNCOVERING THE LINK BETWEEN DIMENSIONAL ANALYSIS AND CAUSALITY

Zach del Rosario[1], Lluís Jofre[2] and Gianluca Iaccarino[3]
1 *Olin College* zdr@olin.edu, 2 *Technical University of Catalonia* lluis.jofre@upc.edu, 3 *Stanford University* jops@stanford.edu

Statistical formulations of causality are designed to overcome *lurking variables* [Joi81; PB14]: important factors that are neglected by the experimenter or hidden by easily observable correlations. This focus is necessary to deal with the complexities of social and biological systems, where enormous complexity makes controlling all important factors impossible. Statistical methods are formulated to mitigate [RR83], or innoculate [BHH+78], an analysis against such lurking factors using a theory-independent approach of treatment assignment based on randomized tests. However, in the physical sciences, the ideas stemming from dimensional analysis enable a more physics-constrained approach.

The familiar Buckingham Pi theorem [Buc14] is the fundamental result of dimensional analysis, stating that any physical law involving measured quantities is necessarily a function of a smaller number of *dimensionless groups*. However, this fundamental result can be endowed with greater structure; a simple log-transform of input quantities leads to a vector subspace interpretation of dimensionless numbers [CRI16].

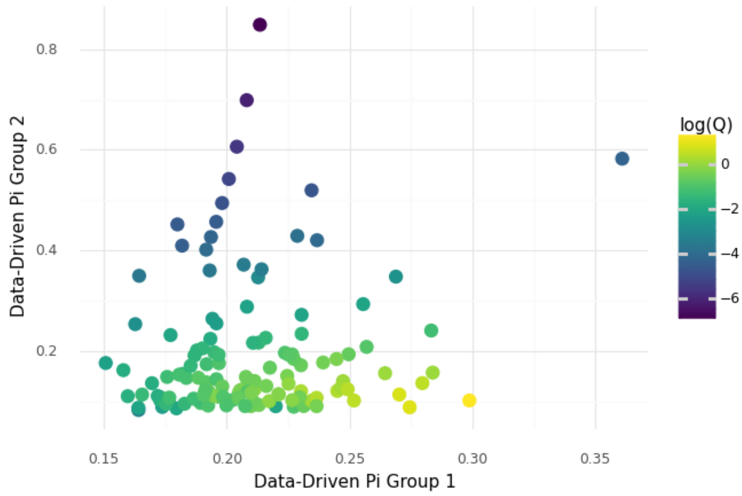

Figure 1. Data-driven dimensional analysis can identify relevant dimensionless numbers. Buckingham Pi predicts 12 dimensionless numbers for this case study; a data-driven analysis identifies two dimensionless numbers that are sufficient to describe the observed trends [JdI20].

The subspace formulation of the Buckingham Pi theorem has two important applications: (i) a formal analysis of lurking variables [RLI19]; and (ii) data-driven approaches to dimensional analysis (Fig. 1). In this talk, we review recent developments in dimensional analysis that provide a physics-constrained view of causality applied to the classical pipe flow experiments by Reynolds and a realistic dataset of particle-laden turbulent flow simulations subject to radiation.

References

[Buc14] Edgar Buckingham. "On physically similar systems; illustrations of the use of dimensional equations". In: *Phys. Rev.* 4 (4 Oct. 1914), pp. 345–376. DOI: 10.1103/PhysRev.4.345. URL: http://link.aps.org/doi/10.1103/PhysRev.4.345.

[BHH+78] George EP Box, William H Hunter, Stuart Hunter, et al. *Statistics for experimenters*. Vol. 664. John Wiley and sons New York, 1978.

[Joi81] Brian L. Joiner. "Lurking variables: Some examples". In: *The American Statistician* 35.4 (1981), pp. 227–233.

[RR83] Paul R Rosenbaum and Donald B Rubin. "The central role of the propensity score in observational studies for causal effects". In: *Biometrika* 70.1 (1983), pp. 41–55.

[PB14] Judea Pearl and Elias Bareinboim. "External validity: From do-calculus to transportability across populations". In: *Statistical Science* 29.4 (2014), pp. 579–595.

[CRI16] Paul G Constantine, Zachary del Rosario, and Gianluca Iaccarino. "Many physical laws are ridge functions". In: *arXiv preprint arXiv:1605.07974* (2016).

[RLI19] Zachary del Rosario, Minyong Lee, and Gianluca Iaccarino. "Lurking Variable Detection via Dimensional Analysis". In: *SIAM/ASA Journal on Uncertainty Quantification* 7.1 (2019), pp. 232–259. DOI: 10.1137/17M1155508. eprint: https://doi.org/10.1137/17M1155508.

[JdI20] Lluís Jofre, Zachary del Rosario, and Gianluca Iaccarino. "Data-driven dimensional analysis of heat transfer in irradiated particle-laden turbulent flow". In: *International Journal of Multiphase Flow* (Apr. 2020). DOI: 10.1016/j.ijmultiphaseflow.2019.103198.