Dissertation Abstract

When scientists make inferences from data, they are guided by theoretical virtues such as unity, testability, parsimony and paucity of free parameters. If challenged to defend those norms, scientists commonly appeal to "Ockham's razor", which recommends a preference for simplicity. But when exactly is one answer to a scientific question simpler than another? And what is the justification for systematically favoring simple conclusions? *Formal learning theory* is a mathematical laboratory for studying inductive method. In that framework, one justifies scientific norms by showing that they are necessary for efficiently arriving at the true answer to a scientific question. In Genin and Kelly [forthcoming], we generalize this framework and employ it to give a precise definition of simplicity, and a novel justification of Ockham's razor. We show that systematically preferring simple answers is a necessary condition for converging to the truth as directly as possible, even if the truth is complex. A major shortcoming of this earlier work is that it models input data deterministically, rather than as the result of a statistical process. Sober [2015] and others doubt that these learning-theoretic arguments can be made to apply to the probabilistic setting of real scientific inference. I resolve that worry in my dissertation, where I generalize the formal learning framework to apply literally to statistical inference as practiced in science and machine learning [Genin and Kelly, 2017].

The advances in my dissertation suggest new ways of approaching traditional problems in philosophy. For example, a major goal of twentieth-century philosophy of science was to show how science could make progress toward the truth even if, at any moment, our best theories are false. To that end, Popper [1962] and others tried to develop a theory of truthlikeness, hoping to prove that theories get closer to the truth over time. That program encountered several notable setbacks. I propose that the locus of investigation be shifted to scientific methods, rather than scientific theories. Say that a method is *progressive* if, no matter which theory is true, the objective chance that the method outputs the true theory is strictly increasing with sample size. In other words: the more data the scientist collects, the more likely their method is to output the true theory. Although progressiveness is not always feasible, it should be our regulative ideal. Say that a method is *α-progressive* if, no matter which theory is true, the chance that it outputs the true theory never decreases by more than α as the sample size grows. This property ensures that collecting more data cannot set your method back too badly. In my dissertation, I prove that, for typical problems, there exists an α-progressive method for every $\alpha > 0$. Furthermore, every α-progressive method must obey a probabilistic version of Ockham's razor. As a demonstration of the utility of the framework, I exhibit α-progressive methods for inferring the Markov equivalence class of a causal graph from observational data, where previously only pointwise-consistent methods were available.

Considerations of progressiveness suggest concrete modifications of standard statistical methods in order to make true findings more likely to be replicated. Suppose that a group of researchers propose recruiting 100 patients to investigate whether a new drug is better at treating migraine than a placebo. In their grant proposal, the researchers analyze their statistical method and conclude the following: if the new drug is better than placebo, the chance that their method detects the improvement is greater than 50%. The funding agency is satisfied. Soon after, the researchers publish a paper claiming to have discovered a promising new treatment. Now, suppose that a replication study is proposed with 150 patients. However, the *ex ante* analysis reveals that the objective chance of detecting an improvement over placebo, if one exists, has decreased to 40%. The chance of replicating successfully has gone down, even though the first study may well be correct, and yet the investigators propose performing a larger study! Such methods are not progressive, and ought to be avoided. This worry is not only theoretical. Chernick and Liu [2012] observe just this kind of regressive behavior in standard hypothesis tests of the binomial proportion. But they have only noticed the tip of the iceberg, for similar considerations attend all statistical inference methods. In my dissertation, I give general conditions under which progressive tests are guaranteed to exist.

References

M. Chernick and C. Liu. The saw-toothed behavior of power vs. sample size and software solutions. *The American Statistician*, 2012.

Konstantin Genin and Kevin T. Kelly. The topology of statistical verifiability. In *Proceedings of the Sixteenth Conference on Theoretical Aspects of Rationality and Knowledge (TARK)*, 2017.

Konstantin Genin and Kevin T. Kelly. Learning, theory choice, and belief revision. *Studia Logica*, forthcoming.

Karl R Popper. Conjectures and refutations: The growth of scientific knowledge. 1962.

Elliott Sober. *Ockham's Razors.* Cambridge University Press, 2015.