Realism, Empiricism, Instrumentalism –

Learning from the Law of Likelihood and AIC *

Elliott Sober

What is now called the "no-miracle argument" for scientific realism is inspired by a striking sentence of Hilary Putnam's (1975) – that scientific realism is "the only philosophy that doesn't make the success of science a miracle." Putnam used the term "miracle" to denote extreme improbability, presumably with no theological implication. Note that Putnam's sentence pits realism against all forms of anti-realism. To assess this argument, we need to get clear on what these isms mean. They often are taken to identify different interpretations that a single scientific theory might be given, but they also are used to characterize distinct generalizations about the goals of science. The no-miracles argument has been applied to both, but I'll first focus on the former. I'll discuss the goals version of the isms after that.

I propose to understand how these three isms apply to a given theory T as follows:

(REAL) T is true.

(EMP) T is empirically adequate.

(INST) T is predictively accurate.

The empiricist's proposition EMP means that T is true in what it says about observables (Van Fraassen 1980). The realist's REAL is bolder – it says that T is true in *everything* it says – about observables and unobservables alike. REAL entails EMP, but not conversely. You might think that REAL also entails INST. Later on, I'll explain why this isn't so. You also might think that instrumentalism says something stronger than INST – that T is neither true nor false. Instrumentalists used to say that theories are tools – they are like hammers, and hammers are neither true nor false. I think this formulation of instrumentalism is a mistake. The old-school

* English version of "Realismo, Empirismo, Strumentalismo. Ci oche possiamo imparare dalla legge di verosimiglianza e dall'AIC." in Raffaella Campaner e Carlo Gabbani (eds.), *Realismo e antirealism nelle scienze. Carocci editore* 2023, pp. 266-272.

formulation should be replaced by the idea that a good theory is one that is predictively accurate; such theories might be true, but they need not be.

A useful representation of the no-miracle argument, I suggest, can be constructed by using the Law of Likelihood, which says that

(LoL) Observation O favors hypothesis H_1 over hypothesis H_2 precisely when
$Pr(O \mid H_1) > Pr(O \mid H_2)$.

The LoL has been defended by Hacking (1965), Edwards (1972), Royall (1977), and Sober (2008b, 2015a).

The LoL applies to the no-miracles argument as follows. Suppose you check numerous predictions that theory T makes and observe that the predictions were all very accurate; call this observation "O." The no-miracles argument may look plausible since the following inequality seems to be true:

(I_1) $Pr(O \mid \text{T is true}) >> Pr(O \mid \text{T is false})$.

However, there is a problem with this representation of the argument. Although I_1 does bring in the realist interpretation of T, it has the defect of not considering the empiricist and instrumentalist interpretations of that theory. EMP does not entail that theory T is false, nor does INST.

Correcting this defect in I_1 leads to the following two inequalities:

(I_2) $Pr(O \mid \text{T is true}) >> Pr(O \mid \text{T is empirically adequate})$
(I_3) $Pr(O \mid \text{T is true}) >> Pr(O \mid \text{T is predictively accurate})$

I_2 compares realist and empiricist interpretations of T; I_3 compares realist and instrumentalist interpretations. I suggest that I_2 and I_3 are both mistaken. Unfortunately for the no-miracles argument, the likelihoods mentioned are equal.

One virtue of the LoL approach to the no-miracles argument is that it exposes a flaw in a central idea of empiricism – that we are cut off from knowing about unobservables even though there is no such cut-off with respect to our knowledge of observables. Van Fraassen (1980) formulates this thought by saying that observational evidence never obliges us to believe the claim that a theory is true if the theory makes claims about unobservables; in contrast, observations sometimes do oblige us to believe theories that are strictly about observables. This empiricist claim concerning the epistemic significance of the distinction between observables and unobservables dissolves when the dichotomous concept of belief is replaced by the three-place relation of favoring used in the LoL. The LoL is an inherently *contrastive* epistemic tool; its point is to *compare* theories with each other, not to evaluate them one at a time. Given this, there is a pleasing symmetry in how realist interpretations of competing theories and empiricist interpretations of those theories are related:

$\Pr(O \mid T_1 \text{ is true}) > \Pr(O \mid T_2 \text{ is true})$ if and only if
$\Pr(O \mid T_1 \text{ is empirically adequate}) > \Pr(O \mid T_2 \text{ is empirically adequate})$

From the point of view of the LoL, we are *not* cut off from evaluating theories that make claims about unobservables (Sober 2008a). The plural term "theories" in the last sentence is key. The fact that we can't see electrons doesn't show that it's impossible to test theories against each other that disagree about the existence or the characteristics of electrons.

It might be objected that my criticism of the no-miracles argument fails because the LoL is defective. One suggested defect is that the LoL is not Bayesian; it fails to take account of prior probabilities and it fails to use a defensible measure of the Bayesian notion of degree of confirmation; for the latter criticism, see Fitelson (2011). My reply is that there is no reason to think that the LoL makes sense only if it is Bayesian; however, I think that Bayesians should accept the LoL because of its relationship to the odds formulation of Bayes's theorem (for discussion, see Sober 2008b, 2015).

A different criticism of the LoL is that it conflicts with a (supposedly) better epistemology – namely, *inference to the best explanation* (IBE). If IBE were better, we should ask whether REAL, if true, would provide a better explanation of O than EMP and INST would do if they were true. However, IBE is only a slogan until a plausible theory of IBE is presented that clarifies what it means for one explanation to be better than another and also shows why the identified sense of better explanation is epistemically relevant to using observations to evaluate theories. Lipton (2004), Psillos (2007), and others have tried to satisfy this two-fold requirement. Suffice it to say that much work remains to be done here (Roche and Sober 2013, 2019; Cabrera 2017).

Turning now to empiricism's relationship to instrumentalism, I'll use the following example to argue that EMP and INST are different – empirical adequacy is not the same as predictive accuracy. Consider two vast fields of corn plants. You randomly sample a hundred plants from the first and find that the average height in your sample is 69 inches; you do the same for the second field and find that the sample average there is 65 inches. You then consider two models for how the two fields of corn are related:

(Null) the average height in field 1 – the average height in field 2 = 0
(Diff) the average height in field 1 – the average height in field 2 = α

The α in Diff is an adjustable parameter; it represents an existential quantifier ("there exists a number α such that …"). Diff is an infinite disjunction, where each disjunct assigns a different value to α.

The maximum likelihood estimate of α, given the observations I mentioned, is that $\alpha = 4$ inches. In other words, the disjunct in Diff that fits the data best is

Best(Diff) the average height in field 1 – the average height in field 2 = 4

Null makes a prediction about what you'll observe if you draw new samples of 100 plants from each field and so does Best(Diff). Which of these predictions should you expect to be more accurate? Notice that Diff fits the data perfectly, whereas Null does not; on the other hand, Null is more parsimonious than Diff if parsimony is measured by counting adjustable parameters.

Model builders in science have long recognized that they can fit the data at hand perfectly if they make their models sufficiently complicated. They also have found that complex models, when fitted to old data, often do very poorly in predicting new data. Complex models that do well in fitting but poorly in predicting are said to *over-fit* the data. This suggests that Null may be more predictively accurate than Diff. However, I think we can be very confident that Null is false and Diff is true. Surely these two vast fields of corn do not have exactly the same mean heights. Diff is empirically adequate, in that what it says about observables is true, but it may not be predictively accurate (when compared with Null).

The fact just mentioned about the lived experience of model-builders is not a brute fact; it has a mathematical explanation, which has been developed by statisticians working in the field of model selection theory (Burnham and Anderson 2002). An important starting point of this enterprise was Akaike's (1973) theorem and what later came to be called "AIC," the Akaike Information Criterion. Akaike's theorem states that

An unbiased estimate of model M's predictive accuracy = log[Pr(data | f(M)] – k.

Here f(M) is the best-fitting member of M, k is the number of adjustable parameters that M contains, and "log" denotes the natural logarithm. AIC is the criterion that uses the right-hand side of this equation to estimate of M's predictive accuracy. I separate theorem from criterion since unbiasedness is neither necessary nor sufficient for an estimator to be the best one to use. Applying AIC to the competition between Null and Diff may yield the result that Null has the better AIC score; whether this happens depends on the data.

Let's now consider realism, empiricism, and instrumentalism as claims about the goals of science, rather than as claims about how a given theory should be interpreted. This means we need to consider these three propositions:

(REAL*) The ultimate epistemic goal of science is to find theories that are true.

(EMP*) The ultimate epistemic goal of science is to find theories that are empirically adequate.

(INST*) The ultimate epistemic goal of science is to find theories that are predictively accurate.

Notice that these asterisked formulations all talk about "the" ultimate epistemic goal of science. I formulate them in this way because realists, empiricists, and instrumentalists typically defend global, monistic philosophies. I'm inclined to pluralism. Scientists sometimes have realist ambitions and sometimes they are instrumentalists; indeed, the same scientist may be a realist about one research problem and an instrumentalist about another. My comment here is not merely descriptive; seeking true theories and seeking theories that are predictively accurate are different scientific goals and both are legitimate. Instrumentalism is especially relevant in understanding the epistemology that is appropriate when models contain idealizations.

References

Akaike, H. (1973). "Information Theory as an Extension of the Maximum Likelihood Principle." In B. Petrov and F. Csaki (eds.), *Second International Symposium on Information Theory*. Budapest: Akademiai Kiado, pp. 267-281.

Burnham, K. and Anderson, D. (2002). *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach.* Springer-Verlag, 2nd edition.

Cabrera, F. (2017). "Can there be a Bayesian Explanationism? On the Prospects of a Productive Partnership." *Synthese* 194: 1245-1272.

Edwards, A. (1972). *Likelihood.* expanded 2nd edition Johns Hopkins University Press, 1992.

Fitelson, B. (2011). "Favoring, Likelihoodism, and Bayesianism." *Philosophy and Phenomenological Research* 83(3): 666-672.

Hacking, I. (1965). *The Logic of Statistical Inference.* Cambridge University Press.

Lipton, P. (2004). *Inference to the Best Explanation*. Routledge. expanded 2nd edition.

Psillos, S. (2007). "The Fine Structure of Inference to the Best Explanation." *Philosophy and Phenomenological Research* 74: 441-448.

Putnam, H. (1975). "What is Mathematical Truth?" in *Mathematics, Matter and Method.* Cambridge University Press.

Roche, W. and Sober, E. (2013). "Explanatoriness is Evidentially Irrelevant, or Inference to the Best Explanation meets Bayesian Confirmation Theory." *Analysis* 73: 659-668.

Roche, W. and Sober, E. (2019). "Inference to the Best Explanation and the Screening-Off Challenge." *Revista Teorema* 38(3): 121-142.

Royall, R. (1997). *Statistical Evidence – a Likelihood Paradigm*. Boca Raton, FL: Chapman and Hall.

Sober, E. (2008a). "Empiricism." In S. Psillos and M. Curd (eds.), *The Routledge Companion to Philosophy of Science*, pp. 129-138.

Sober, E. (2008b). *Evidence and Evolution – the Logic Behind the Science.* Cambridge University Press.

Sober, E. (2015a). *Ockham's Razors*. Cambridge University Press.

Sober, E. (2015b). "Two Cornell Realisms – Moral and Scientific." *Philosophical Studies* 172: 905-924.

Van Fraassen, B. (1980). *The Scientific Image.* Oxford University Press.