An application of Carnapian inductive logic to philosophy of statistics

Teddy Groves, University of Kent

Introduction

In my talk I claim that an argument in philosophy of statistics found in Gelman and Shalizi (2012) can be improved using Carnapian inductive logic.

Gelman and Shalizi argue against the 'conventional philosophy' of Bayesian statistics, which stipulates that statistical models should be chosen only on the basis of how well they represent knowledge, and only draw conclusions on the basis of knowledge that is represented probabilistically. According to Gelman and Shalizi, these stipulations rule out practically important ways of choosing and evaluating such models. They claim that a hypothetico-deductive philosophy does better justice to the practice of Bayesian statistics.

I argue that an approach informed by Carnapian inductive logic should play the role of hypothetico-deductivism in Gelman and Shalizi's arguments.

A Carnapian approach to choosing statistical models

Carnap sought to *explicate*—that is, usefully replace—everyday inductive discourse with statements of inductive logic. An inductive logic consists of a formal language together with a 'confirmation function' associating pairs of its sentences with positive real numbers. The focus of Carnap's research was to find formal conditions that inductive logics must satisfy in order to be satisfactory explications, and to identify which confirmation functions various combinations of such conditions pick out.

For example, *Johnson's sufficientness postulate* requires that the value assigned to pairs of sentences (h, e), where h describes the instantiation of some property and e describes a sample of observations, must depend only on the frequency of the property in the sample and the sample's size. Carnap argued that that this was a plausible condition of explication and showed that the confirmation functions satisfying it can be characterised by a single continuous parameter.

De Finetti's theorem identifies a correspondence between a class of inductive logics that interested Carnap and an important class of statistical models: any probability distribution over independent identically distributed random variables can be represented by a Carnapian inductive logic satisfying the condition *constant exchangeability*, and vice versa.

This correspondence suggests the following Carnapian approach to selecting statistical models: choose statistical models whose corresponding inductive logics are good explications. This is the approach that I believe should play the role of hypothetico-deductivism in Gelman and Shalizi's argument.

Advantages of the Carnapian approach

Agreement with Gelman and Shalizi's arguments

In line with Gelman and Shalizi's criticism of the first stipulation of the 'conventional philosophy', Carnap argued that inductive logics need not represent beliefs in order to be good explications. Analogously to their second argument, Carnap also favoured the use of methods for evaluating inductive logics that are not

themselves adequate . I argue that these inductive logical arguments are convincing and allow the Carnapian approach to accommodate Gelman and Shalizi's criticisms.

It is harder for hypothetico-deductivism to accommodate Gelman and Shalizi's criticisms, as the latter depend on pragmatic reasoning about how statistical models are used. This kind of reasoning is awkward to construe hypothetico-deductively, in contrast to the Carnapian approach, which is inherently pragmatic.

Technical fruitfulness

The Carnapian approach is technically fruitful, allowing statisticians to identify the prior distributions corresponding to natural qualitative assertions about the processes underlying their data. I rehearse an argument to this effect from Zabell (2011), present the correspondence between the Dirichlet family of probability distributions and the Johnson's sufficientness postulate as a case study and point to further work in the inductive logic literature that could be relevant to applied Bayesian statistics.

Permissiveness

A final advantage of the Carnapian approach over hypothetico-deductivism is that it is in harmony with Gelman and Shalizi's reluctance to articulate formally precise procedures for evaluating statistical models. The Carnapian approach offers a principled justification for this attitude: it says that such procedures should depend on which logics are good explications of everyday discourse and so fundamentally concern which form of expression it is preferable to adopt in a particular situation. In Carnap (1950), Carnap argues, in strikingly similar terms to Gelman and Shalizi, that philosophers should refrain from trying to give definitive answers to questions about which forms of expression to adopt.

On the other hand I claim, following arguments in Mayo (2013), that Gelman and Shalizi's underformulation of model-evaluation procedures is difficult to square with hypothetico-deductivism.

Addressing qualms about inductive logic

In the last section of my talk I address two reasons why statisticians might be cautious of the Carnapian interpretation of Gelman and Shalizi's argument, arguing that neither is well-founded.

The problem of induction

It might be thought that the Carnapian approach suffers particularly from the problem of induction due to its use of 'inductive' logic. An anti-inductive philosophy of statistics, according to which there is no need to reason inductively, might seem preferable.

I argue that the Carnapian approach has no greater problem of induction than its alternative. On the one hand, while statements of Carnapian inductive logic are just as difficult to justify as the everyday inductive statements that they explicate and Hume problematised, no additional difficulty arises from the formalisation. On the other hand, anti-inductivism must address the 'practical problem of induction', as set out in Salmon (1981), by providing rational grounds to be confident in their predictions without employing any kind of inductive reasoning. This task seems at least as difficult as overcoming the original problem.

One true inductive logic

Secondly, statisticians might be wary of the Carnapian approach on the grounds that it is associated with a discredited attempt to find one formal method that explicates all of scientific discourse.

To allay such concerns I present quotes from Carnap's work which demonstrate that he did not have this ambition, and indeed considered it impossible.

References

Carnap, R. (1950). Empiricism, Semantics and Ontology. *Revue Internationale de Philosophie*, 4:20–40.

Gelman, A. and Shalizi, C. R. (2012). Philosophy and the practice of bayesian statistics. *British Journal of Mathematical and Statistical Psychology*, 66(1):8–38.

Mayo, D. G. (2013). The error-statistical philosophy and the practice of bayesian statistics. *British Journal of Mathematical and Statistical Psychology*, 66(1):57–64.

Salmon, W. C. (1981). Rational prediction. *British Journal for the Philosophy of Science*, pages 115–125.

Zabell, S. (2011). Carnap and the logic of inductive inference. *Inductive Logic*, 10:265.