



Abducted by Bayesians?

Jan-Willem Romeijn¹

Department of Philosophy, University of Groningen, Netherlands



ARTICLE INFO

Article history:

Available online 3 July 2013

Keywords:

Abduction
Underdetermination
Bayesian inference
Hierarchical Bayesian modelling

ABSTRACT

This paper discusses the role of theoretical notions in making predictions and evaluating statistical models. The core idea of the paper is that such theoretical notions can be spelt out in terms of priors over statistical models, and that such priors can themselves be assigned probabilities. The discussion substantiates the claim that the use of theoretical notions may offer specific empirical advantages. Moreover, I argue that this use of theoretical notions explicates a particular kind of abductive inference. The paper thus contributes to the discussion over Bayesian models of abductive inference.

© 2013 Published by Elsevier B.V.

1. Introduction

In this section I introduce theoretical notions in a statistical context. Against this background I specify the two central claims of this paper.

1.1. Theoretical notions and empirical content

In what follows I adopt the view that the empirical content of a statistical hypothesis H is given by its likelihoods, that is, by the probabilities of data E conditional on the hypothesis, written as $P(E|H)$ (cf. Douven [6]). If two hypotheses H and H^* have identical likelihoods, that is,

$$P(E|H) = P(E|H^*)$$

for all possible observations E , then they have the same empirical content.² We will say that a distinction between hypotheses is based on a theoretical notion, or theoretical for short, if it distinguishes two hypotheses H and H^* while these hypotheses have identical likelihood functions. Furthermore, a statistical model is defined to be a collection of hypotheses, $\mathcal{H} = \{H_1, H_2, \dots, H_n\}$. We will say that a distinction between two models \mathcal{H} and \mathcal{H}^* is theoretical if the hypotheses in the model have pairwise identical likelihood functions, $P(E|H_j) = P(E|H_j^*)$ for $0 < j \leq n$ and for all E .

A central point in this paper is that models whose distinction is theoretical may still differ in empirical content, because of the priors we define over them. We will look at models \mathcal{H} and \mathcal{H}^* that differ theoretically in the sense specified above, but that are associated with different stories concerning the data generating system. Such stories motivate different priors over the models in question, and these priors again lead to a different empirical content for the two models. The data may be used to choose between the models in virtue of their association with different priors.

^{E-mail address:} j.w.romeijn@rug.nl.

¹ The author is simultaneously a research fellow at the Philosophy Department of the University of Johannesburg.

² If we assume the likelihood principle, according to which all evidence pertaining to a hypothesis is mediated by the likelihoods of that hypothesis, then the two hypotheses cannot be distinguished by any evidence. See Royall [32].

This approach to comparing statistical models has been around for a while, and indeed can be traced back to Gaifman [11].³ It is perfectly coherent to assign probabilities to probability assignments, as is routinely done in Bayesian statistics, and it is also coherent to add more layers of probabilistic analysis, assigning probabilities to the probability assignments over the ground level assignments, and so on. More recently, this idea has taken root in statistics, more precisely in hierarchical Bayesian modelling (cf. Gelman et al. [12], Gelman and Hill [13]). This approach compares models on their marginal likelihoods, in which the prior over the model is an explicit component. Henderson et al. [16] discuss some philosophical applications of hierarchical modelling.

1.2. Central claims of this paper

The upshot of this paper is that theoretical notions can play an active role in statistical inference, and that in particular cases we can tell apart theoretically distinct models by empirical means. Higher-order probabilities play a crucial role in capturing these theoretical notions and in making them empirical. These two claims answer two critical discussions on theoretical notions in science. The first of these is that theoretical notions can be understood as methodological tools, and should not be discarded as superfluous and non-empirical. This claim can be viewed a response to the theoretician's dilemma discussed in Hempel [15]. He argues that, if the aims of science are indeed empirical, there is no role for theoretical notions, since we can purge a scientific theory of such notions without losing any of its empirical content. Against this, I argue that certain uses of theoretical notions lead to more efficient inferences from the data.

The second claim is that the use of theoretical notions indicated above captures a particular kind of abductive inference. This is a reply to van Fraassen [37], who argues that abductive inference is probabilistically incoherent. To some extent I go along with van Fraassen's way of framing Bayesian abduction. I consider a set of statistical hypotheses, and take their empirical content to be given by their likelihoods. But van Fraassen then proposes to capture the role of theoretical notions, e.g., their explanatory force, by additional changes to the probability assignment over the hypotheses, after processing the Bayesian update. The use of theoretical notions thus leads to probabilistic incoherence. In response, the Bayesian model of abduction proposed in this paper uses theoretical notions to motivate priors over multiple models, which are then adapted by Bayesian conditioning. Because of the different priors, the impact of the data on the models is different, allowing us to tell apart models on the basis of theoretical considerations.

2. Bayesian models of abduction

This section discusses some earlier attempts to accommodate the use of theoretical notions in Bayesian inference, and thus to reconcile it with abduction. See, for instance, Day and Kincaid [3], Okasha [25], Salmon [33], Sober [35], McGrew [23], and Lipton [22].⁴ Following the structure of Bayes' rule, these attempts incorporate the explanatory or theoretical considerations in the prior probability, in the likelihoods, or in both.

2.1. Abduction by priors

One idea is to model abduction in a Bayesian framework by means of prior probabilities, namely by shifting prior weight to more explanatory hypotheses. However, as argued by Milne [24], if we want to capture the explanatory considerations in a Bayesian update, then we need to portray this head start of explanatory hypotheses as resulting from the impact of some sort of evidence. Milne then notes that the characteristics that make a hypothesis explanatory, like simplicity, aesthetic quality, and the like, will typically be carried by the hypothesis from the very beginning, because they are logically entailed by the hypothesis or because the hypothesis is constructed to have those characteristics. Therefore, any evidential impact of theoretical characteristics runs into the problem of logical omniscience, or in this case equivalently, the old evidence problem.⁵

There may be theoretical characteristics that are not analytic in this way, but that somehow rely on evidence or background knowledge. But as illustrated by the burglar story in Weisberg [38], it is still not clear that their impact can then be modelled by means of a prior probability. Say that you find your house in a mess, valuables are missing, and by way of explanation you imagine either of two things: there has been a burglar in the house, or alternatively, one burglar in your house was disturbed by another, then both were discovered by a policeman who chased them away and then took advantage of the situation. Weisberg argues that the first story is more explanatory, quite independently of how probable you find these stories to start with. At least in a subjective Bayesian framework, nothing forces us to align our prior probabilities to our judgement of the explanatory force of the hypothesis.

³ Considering the influence of Prof. Gaifman's work on the philosophy of science, and on the philosophy of statistics in particular, the connection of this paper to his work is a rather weak one. I regret that the philosophical appraisal of Gaifman and Snir's seminal paper on rich languages, which was planned for this special issue, is still under construction.

⁴ Douven [5], Tregear [36] and Weisberg [38] react to the criticisms of van Fraassen by showing that explanatory considerations can be modelled as rational changes of belief that do not comply to the Bayesian model. I think these reactions deserve separate attention, but for the moment I seek to maintain the Bayesian norms.

⁵ Several authors have proposed Bayesian models of learning such analytic truths; see Earman [7]. But there is certainly no consensus over these models. I will leave further discussion over them aside.

Download English Version:

<https://daneshyari.com/en/article/4662997>

Download Persian Version:

<https://daneshyari.com/article/4662997>

[Daneshyari.com](https://daneshyari.com)