Assessing Scientific Theories

Stephan Hartmann

Munich Center for Mathematical Philosophy
LMU Munich

California Institute of Technology
April 2016

The Problem

- Question: How are scientific theories assessed?
- Answer: Theories are assessed in the light of empirical data.
- However: This does not work too well in fundamental physics. In some cases one has to wait very long for an empirical confirmation (e.g. in the case of the Higgs), and in other cases it is not clear whether there will ever be empirical data (e.g. in the case of String Theory). Can theories such as String Theory be confirmed at all?
- More specifically: Are there "non-empirical" ways of assessing scientific theories? – Some people believe so.
- In this talk, I will assess one corresponding argument structure and explore how different theories of confirmation deal with it.
- It turns out that Bayesian Confirmation Theory provides an adequate framework to analyze this argument structure and that we therefore do not need a new methodology of science as, e.g., the physicists George Ellis and Joe Silk (2014) recently suggested.

The No Alternatives Argument (NAA)

Scientists often argue like this:

(P_1) Theory or hypothesis H satisfies several desirable conditions (e.g., it incorporates various fundamental principles, coheres with other theories,. . .)

(P_2) Despite a lot of effort, the scientific community has not yet found an alternative to H.

(C) Hence, we have *one* reason in support of H.

- This is the so-called No Alternatives Argument (NAA)
- Here are some examples of it. . .

Examples from Fundamental Physics

There are many examples of NAAs in fundamental physics, mainly because discriminating empirical evidence is hard to come by. Here are three:

1. The Higgs Mechanism
   This mechanism was invented in 1964, but its empirical confirmation had to wait until 2012. However, there was little doubt that the mechanism was in principle correct. There was no alternative. . .
2. String Theory
   This theory cannot (yet) be tested empirically. What speaks in its favor is that it resolves the conflict between QFT and GTR, (non straight forward) coherence arguments, and the NAA.
3. Cosmic Inflation
   This model (or class of models) enjoys a limited degree of empirical confirmation; trust in the theory crucially relies on the NAA.
   This is a nice example which can be used to study how empirical and non-empirical (NAA) confirmation can work together.

Questions

1. How good are no alternatives arguments?
2. Under what conditions do they work?

To address these questions, we analyze the NAA in the framework of the currently most popular theory of confirmation, viz. Bayesian Confirmation Theory.

Overview

1. Motivation
2. Theory Assessment
3. A Bayesian Analysis of the NAA
4. Two Follow Ups
5. Outlook

II. Theory Assessment

Assessing Arguments

- Scientists use arguments to support their theories, models and hypotheses.
- Traditionally, one distinguishes three types of arguments.
  - Deductive arguments such as: (P1) Copper conducts electricity, (P2) This is a piece of copper, (C) This piece of copper conducts electricity.
  - Inductive arguments such as: (P1) This piece of copper conducts electricity, (P2) That piece of copper conducts electricity,. . . (C) Copper conducts electricity.
  - Abductive arguments such as: (P) QM provides the best explanation why copper conducts electricity. (C) QM is true.
- Philosophers ask: How good are these arguments? What can we say about their validity?
- Further question: Are there other argument types that scientists use?
- The NAA is one such new argument type which is used in support of a given scientific theory. We ask under which conditions it is a good argument.

Confirmation and Corroboration

- Empirical data are relevant for the assessment of scientific theories and scientific theories should account for the empirical phenomena in their domain. After all, they are empirical theories, and not just pieces of mathematics.
- Empirical data E ("evidence") confirm or disconfirm a given theory H. This means that we have a good reason to belief in the truth of H and therefore a good reason to apply the theory in the future.
- But what does it exactly mean that E confirms H? How can this relation be explicated? To address this question, we will look at two confirmation theories developed by philosophers of science.
- Before, however, we consider Popper's falsificationism. Popper was an anti-inductivist: For him, a theory can never be confirmed. It can only be corroborated which – importantly – has no implications for our beliefs in the theory's expected future performance.

Theory Assessment I: Popper's Falsificationism

- According to naive falsificationism, a theory or hypothesis H is corroborated if an empirically testable prediction of H obtains. Otherwise it is falsified and should be rejected and replaced by an alternative theory.
- N.B.: More sophisticated versions of falsificationism have the same problem as naive falsificationism, and so I won't discuss them here.
- It is important to note that, according to falsificationism, a theory can only be corroborated empirically. Hence, a Popperian cannot make sense out of the NAA (at least not in a straight forward way). The NAA is not an acceptable way to justify a theory.

Theory Assessment II: The Hypothetico-Deductive Model

- According to the hypothetico-deductive model, a theory or hypothesis H is confirmed by a piece of evidence E iff E is predicted by H (i.e. if E is a deductive consequence of H) and if E is observed.
- Also here, a theory or hypothesis can only be confirmed empirically and it is hard to imagine how a defender of the hypothetico-deductive model can make sense out of the NAA.
- The model has a number of (other) well-known problems, e.g.
  1. The Tacking Problem: If E confirms H, then it also confirms $\mathrm{H} \wedge \mathrm{X}$. Note that X can be a completely irrelevant proposition. This is counter-intuitive.
  2. Degrees of confirmation: Some evidence confirms a theory or hypothesis more than other evidence. However, according to the hypothetico-deductive model, we can only make the qualitative inference that E confirms H (or not).

Theory Assessment III: Bayesian Confirmation Theory

- According to Bayesian Confirmation Theory, a theory or hypothesis H is confirmed by a piece of evidence E iff the observation of E raises the (subjective) probability of H.
- Scientists attach a degree of belief (= a probability) to a theory or hypothesis and change ("update") it in the light of new evidence.
- Reasons are provided why this is a rational procedure (e.g. Dutch Book arguments).
- What evidence? An observed instance of a law, testimony,. . .
- How should one update? Conditionalization (at least in many cases): The posterior probability of H (i.e. $P'(\mathrm{H})$) follows from the *prior probability* of H (i.e. $P(\mathrm{H})$), the *likelihood* of the evidence (i.e. $P(\mathrm{E}|\mathrm{H})$) and the *expectancy* of the evidence (i.e. $P(\mathrm{E})$):

Bayes Theorem

$$P'(\mathrm{H}) := P(\mathrm{H}|\mathrm{E}) = \frac{P(\mathrm{E}|\mathrm{H}) \cdot P(\mathrm{H})}{P(\mathrm{E})}$$

Discussion

- Bayesian Confirmation Theory can be applied in a straightforward way to empirical testing, i.e. to the case where a direct deductive or inductive consequence E of H is observed.
- Note that Bayesian Confirmation Theory accounts for the fact that some evidence confirms a hypothesis better than another piece of evidence. One way to measure the degree of confirmation is by using the difference measure $d(\mathrm{H},\mathrm{E}) := P(\mathrm{H}|\mathrm{E}) - P(\mathrm{H})$.
- What is more, it turns out that the Bayesian machinery is flexible enough to also model other ways of confirming theories.
- Hence I claim, contra Ellis and Silk, that we do not need a new methodology for science. Bayesian Confirmation Theory suffices to account for everything we need.
- We will now show how this works for the no alternatives argument.

III. A Bayesian Analysis of the NAA

Formalizing NAAs

- Let $\mathcal{D}$ be a set of data and $\mathcal{C}$ be a set of constraints. The community aims at a theory that accounts for $\mathcal{D}$ and satisfies $\mathcal{C}$.
- The hypothesis H accounts for $\mathcal{D}$ and satisfies $\mathcal{C}$.
- So far, no alternative hypothesis has been found that explains $\mathcal{D}$ and satisfies $\mathcal{C}$.

We ask: To what extend does this observation confirm H?

To address this question, we introduce two propositional variables:

1. T has two values, viz. T: The hypothesis H is true, and ¬T: The hypothesis H is not true.
2. F also has two values, viz. F: The scientific community has not yet found an alternative to H that accounts for $\mathcal{D}$ and satisfies $\mathcal{C}$, and ¬F: The scientific community has found an alternative to H that accounts for $\mathcal{D}$ and satisfies $\mathcal{C}$.

A First Bayesian Attempt

Goal: Show that F confirms T, i.e. that

$$d(\mathrm{T},\mathrm{F}) := P(\mathrm{T}|\mathrm{F}) - P(\mathrm{T}) > 0.$$

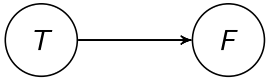
$d(\mathrm{T},\mathrm{F})$ is the difference measure of confirmation.
$P(\mathrm{T})$ is the prior probability of T, and $P(\mathrm{T}|\mathrm{F})$ is the posterior probability of T, i.e. the probability of T after having learned the evidence F.

Question: How can this be done?

- Wave our hands and claim that F is *obviously* positively relevant for T.
- But why should we accept this? And even if we do so, the Bayesian machinery does not add much or anything to the argument.

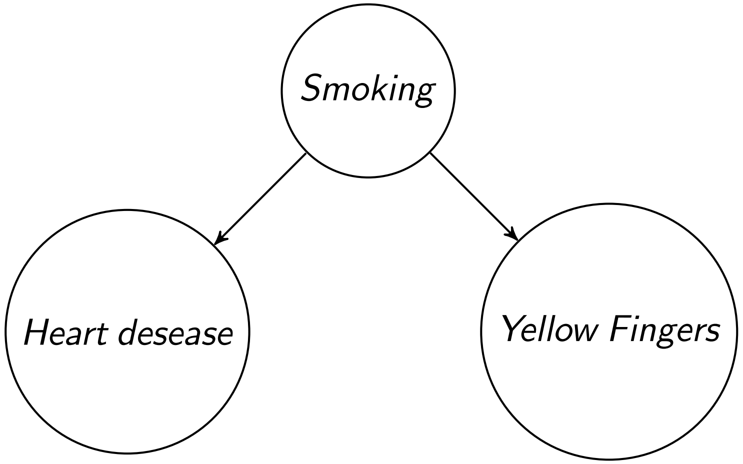
Another Worry

Here is a standard Bayesian Network representation that is used to model the relation between a hypothesis and a piece of evidence:



- On this account, there is a direct influence of the hypothesis on the evidence. (Example: T: All raven are black. F: This raven is black.)
- But no such relation holds in the present case: T does not deductively or inductively imply F.
- F is at best some kind of indirect evidence for T. It is an example of what we call non-empirical evidence.
- Hence our conjecture: There is a common cause variable Y that is responsible for the correlation between F and T. Technically speaking, Y screens off F and T.

Common Causes: An Illustration



Introducing Y

We introduce a third variable.

3. Y has N values, viz. Y_i: There are exactly i hypotheses which explain $\mathcal{D}$ and fulfill $\mathcal{C}$. (H is one of them.)

Note that N can be ∞.

- The alternative theories make different predictions and provide different explanations, and can therefore be individuated. Theories which make exactly the same (or almost the same) empirical predictions, are considered to be identical. Scientists have a good sense of what counts as a different theory, and what not.
- We claim that scientists have beliefs (supported by arguments) about the distribution of the Y_i.

Relations between F, T and Y

Repetition:

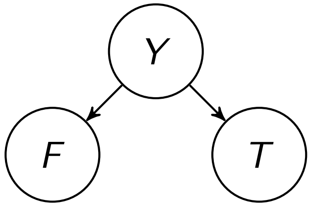
1. T has two values, viz. T: The hypothesis H is true, and ¬T: The hypothesis H is not true.
2. F also has two values, viz. F: The scientific community has not yet found an alternative to H that accounts for $\mathcal{D}$ and satisfies $\mathcal{C}$, and ¬F: The scientific community has found an alternative to H that accounts for $\mathcal{D}$ and satisfies $\mathcal{C}$.
3. Y has N values, viz. Y_i: There are exactly i hypotheses which account for $\mathcal{D}$ and fulfill $\mathcal{C}$. (H is one of them.)

A Bayesian Network Representation

T is conditionally independent of F given Y:

Conditional Independence

$$T \perp\!\!\!\perp F | Y$$



Once we know the value of Y, we won't learn anything new about the truth value of T if we learn the truth value of F.

The Prior Probabilities

To complete the Bayesian Network, we have to fix the values of $P(\mathrm{Y_i})$, $P(\mathrm{T}|\mathrm{Y_i})$ and $P(\mathrm{F}|\mathrm{Y_i})$ for all i.

1. The Priors y_i

$P(\mathrm{Y_i}) =: y_i$, with $0 \leq y_i < 1$.

This assignment reflects the fact that we do not know the number of alternative theories a priori.

2. The Conditional Probabilities t_i

The Conditional Probabilities

$P(\mathrm{T}|\mathrm{Y_i}) =: t_i$ are monotonically decreasing in i (and $t_1 = 1$).

This is plausible: The more alternative theories there are, the less sure we can be that H is true.

A natural choice is to apply the Principle of Indifference and to assume that $t_i = 1/i$, but for our purposes the weaker assumption is sufficient.

3. The Conditional Probabilities f_i

The Conditional Probabilities

$P(\mathrm{F}|\mathrm{Y_i}) =: f_i$ are monotonically decreasing in i (and $f_1 = 1$).

This is plausible: The more alternative theories there are, the less likely it is that scientists have not yet found one.

Calculating the Difference Measure $d(\mathrm{T}, \mathrm{F})$

Given our "common-cause" Bayesian Network, the following holds:

Lemma

$$d(\mathrm{T}, \mathrm{F}) = \frac{1}{2\,P(\mathrm{F})} \cdot \sum_{i \neq j=1}^{N} (f_i - f_j)\,(t_i - t_j)\, y_i\, y_j$$

Theorem

If f_i and t_i are monotonically decreasing in i and if there is at least one pair (i, j) with $j > i$ for which

(i) $y_i\, y_j > 0$,

(ii) $f_i > f_j$,

(iii) $t_i > t_j$,

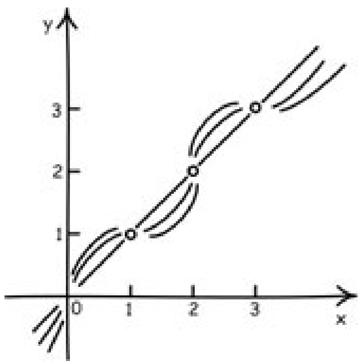
then $d(\mathrm{T}, \mathrm{F}) > 0$.

Discussion

- Note that the assumptions of the theorem are rather weak. If an agent assigns degrees of belief that satisfy them, then she will be rational to make the NAA.
- However, if someone believes, for example, that the number of alternatives has a fixed value, then F does not confirm T and the NAA has no pull. One could, for example, argue that the number of alternatives is infinite (i.e. that $y_\infty = 1$).
- Note, though, that scientists are often convinced that the number of alternative theories is rather small (without knowing the precise value). They are impressed by the difficulty to construct them. And this explains their conviction (supported by our analysis) that F confirms T.
- But is this line of thought convincing?

The Underdetermination Thesis

According to the Underdetermination Thesis, there is always an infinite number of alternative theories that is consistent with a given (finite or countably infinite) set of data $\mathcal{D}$. Scientific theories are underdetermined by empirical data.



The Underdetermination Thesis

- So should a rational agent set $y_\infty = 1$?
- The NAA-er has several options to respond:
  - (i) Argue that the constraints $\mathcal{C}$ reduce the number of alternatives to a finite number.
  - (ii) Argue that the Underdetermination Thesis only shows that $0 < y_\infty < 1$ and that there is at least one (finite) pair (i, j) for which the assumptions of the theorem hold.
  - (iii) Richard Dawid defends the so-called meta-inductive argument (MIA) according to which underdetermination is restricted if the theory in question belongs to a successful research program with many well-confirmed theories. In the case of String Theory, this is the QFT program.

    See R. Dawid: *String Theory and the Scientific Method*. Cambridge: Cambridge University Press 2013.

Two Other Worries

1. Often different scientists disagree upon what (i) the data $\mathcal{D}$ are the theory should account for, and (ii) which theoretical constraints $\mathcal{C}$ the theory should satisfy. (For example, a defender of Loop Quantum Gravity sets different constraints than a string theorist.)

   Clearly, a NAA is only convincing if one agrees on $\mathcal{C}$ and $\mathcal{D}$.
2. The difficulty of the problem (or the ability of the scientist) should be included in the Bayesian model. This can be done in a straightforward way (see the paper). The present results hold for a fixed value of the difficulty of the problem.

   Note, though, that one will need evidence for the claim that scientists would find an alternative if there were one.

What do the Probabilities Mean?

- We adopt the Bayesian framework according to which probabilities measure the (subjective) strength of belief of an agent in the truth of a certain hypothesis. This agent can be a scientist or, perhaps, the whole scientific community.
- It is debatable what that hypothesis actually is. Do scientists attach degrees of belief to the truth of theories *tout cours*? Or do they only attach degree of beliefs to statements like "The theory or model will do well for the next applications in its domain." Here it needs to be specified (and is an empirical question) what the domain of a theory or model is. This would be consistent with the claim (which Popper also holds) that all theories are (strictly speaking) false or hold only ceteris paribus.
- In any case, Bayesianism allows us to study the statics and dynamics of the beliefs that scientists hold.

IV. Two Follow Ups

1. Inference to the Best Explanation (IBE)

- Under which conditions is IBE justified?
- Replace F by F'. F' has two values, viz. F: The scientific community has not yet found a better explanation than H, and ¬F: The scientific community has found a better explanation than H.
- T and Y remain as before and the independence assumption (i.e. $T \perp\!\!\!\perp F'|Y$) holds.
- We can therefore proceed as before and study under which conditions IBE works, i.e. when (i) $P(\mathrm{T}|\mathrm{F}') > P(\mathrm{T})$ and (ii) $P(\mathrm{T}|\mathrm{F}') \approx 1$.
- Question: What about van Fraassen's "bad lot" argument?
- The answer depends on how good our reasons are that the true theory is amongst a given number of alternatives.
- Note the interesting difference between ordinary reasoning and scientific reasoning w.r.t. IBE.

2. The Existence of God

- Recall N. R. Hansson's argument against the existence of God: "So far I have not seen a good argument *for* the existence of God. This observation is, in turn, a good argument *against* the existence of God."
- We proceed as before and introduce three propositional variables:
  1. T'' has two values, viz. T": God exists, and ¬T": God does not exist.
  2. F'' also has two values, viz. F": I have not found a good argument for the existence of God, and ¬F": I have found a good argument for the existence of God.
  3. Y'' has N values, viz. Y_i": There are exactly i arguments for the existence of God.
- The independence assumption (i.e. $T'' \perp\!\!\!\perp F'' | Y''$) holds and the parameters f_i'', t_i'' and y_i'' are defined as before.

2. The Existence of God (Cont'd)

The Conditional Probabilities

$P(\mathrm{T''}|\mathrm{Y''_i}) =: t_i''$ are monotonically increasing in i.

Remember our (suitably adapted) lemma:

Lemma

$$d(\mathrm{T''}, \mathrm{F''}) = \frac{1}{2\,P(\mathrm{F''})} \cdot \sum_{i \neq j = 1}^{N} (f_i'' - f_j'')\,(t_i'' - t_j'')\,y_i''\,y_j''$$

Theorem

If f_i'' are monotonically decreasing in i and t_i'' are monotonically increasing in i and if there is at least one pair (i, j) with $j > i$ for which (i) $y_i''\,y_j'' > 0$, (ii) $f_i'' > f_j''$, and (iii) $t_i'' < t_j''$, then $d(\mathrm{T''}, \mathrm{F''}) < 0$.

V. Outlook

Outlook

- Science changed considerably over the last decades, and so did the methodology of science.
- Non-empirical ways of assessing scientific theories (such as the NAA) raise a number of interesting philosophical issues.
- Deductivist accounts of confirmation and corroboration cannot accommodate the NAA.
- Hence, if one does not want to dismiss this kind of reasoning, then an alternative to deductivism is needed.
- The NAA can be analyzed in the Bayesian framework and we have investigated under which conditions it is a good argument.
- Hence, we do not need a new methodology of science (as Ellis and Silk suggest): The Bayesian framework is flexible enough.
- Further indirect ways of assessing scientific theories, such as analogue simulations, can be analyzed similarly.
- Are there alternatives ways of modeling the NAA? – I tried to find one, but so far without any success. . .

Thanks for your attention!

- The talk is based on a joint paper with Richard Dawid (Stockholm) and Jan Sprenger (Tilburg).
- The paper appeared in *The British Journal for the Philosophy of Science* 66(1): 213-234 (2015).