Modeling in Philosophy of Science

Stephan Hartmann
London School of Economics

Some questions I am interested in

- How are scientific theories and models confirmed?
- How is evidence for or against a theory evaluated?
- How do different theories hang together?
- Can aspects of scientific theory change be explained philosophically?

I am not the first to address these questions…

Two traditions…

1. Normativism
   - Examples: Falsificationism, Bayesianism
   - Typically motivated on *a priori* grounds
2. Descriptivism
   - Examples: Kuhn, naturalized philosophies of science
   - Examine specific case studies that contribute to a better understanding of science

…and their problems

1. Normativism
   - challenged by insights from the history of science (e.g. Popper and the stability of normal science)
   - often “too far away” from real science
2. Descriptivism
   - not clear how generalizable insights or normative standards can be provided
Desiderata for a methodology of science

1. It should be normative and provide a defensible general account of scientific rationality.
2. It should provide a framework to illuminate “intuitively correct judgments in the history of science and explains the incorrectness of those judgments that seem clearly intuitively incorrect (and shed light on ‘grey cases’)” (John Worrall)

How can this goal be achieved?

- Mimic successful scientific methodology and...
- construct (philosophical) models in the framework of a (philosophical) theory.

I'll explain what I mean by this.

Outline

Motivation
I. Modeling in Science
II. Textbook Bayesianism
III. Bayesian Networks
IV. Example 1: Variety-of-Evidence Thesis
V. Example 2: Scientific Theory Change
VI. Methodological Reflections
VII. Conclusions

I. Modeling in Science
The ubiquity of models in science

- **Highschool**: Bohr model of the atom, model of the pendulum,…
- **University**: Standard Big Bang Model, Standard Model of particle physics,…
- **Even later**: Physicists like Lisa Randall (Harvard) use models to learn about the most fundamental question of the universe.

We observe a shift from theories to models in science, and this shift is reflected in the work of philosophers of science in the last 25 years.

Theories and models

- Scientists often use the words *theory* and *model* interchangeably.
- So how can one distinguish between theories and models?

Features of theories

- Examples: Newtonian Mechanics, Quantum Mechanics
- General and universal in scope
- Abstract
- Often difficult to solve (example: QCD)
- No idealizations involved (ideally…)

Features of models

- Examples: the model of a pendulum, the Bohr model of the atom, gas models, the MIT Bag model,…
- Specific and limited in scope
- Concrete
- Intuitive and visualizable
- Can be solved
- Involve idealizing assumptions
Two kinds of models

- Models of a theory
  - example: the treatment of the pendulum is a model of Newtonian Mechanics
  - the theory acts as a modeling framework (and not much follows from the theory w/o specifying a model)
- Phenomenological models
  - example: the MIT Bag model
  - there is no theory into which a model can be embedded.

Modeling in philosophy

- So maybe we should construct models in philosophy as well!
- Some authors do this already, see, e.g., the work of Brian Skyrms on the social contract and Clark Glymour et al.’s work on causal discovery.
- I will show that models are also valuable tools for understanding the methodology of science.

What is Bayesianism?

- Quantitative confirmation theory
- When (and how much) does a piece of evidence $E$ confirm a hypothesis $H$?
- Typically formulated in terms of probabilities = subjective degrees of belief
- Normative theory (see Dutch books)
- Textbooks: Howson & Urbach: Scientific Reasoning, and Earman: Bayes or Bust.
The mathematical machinery

- Confirmation = positive relevance between $H$ and $E$
- Start with a prior probability $P(H)$
- Updating-rule: $P_{\text{new}}(H) := P(H|E)$
- By mathematics ("Bayes' Theorem") we get:
  $$P_{\text{new}}(H) = P(E|H) \frac{P(H)}{P(E)}$$
- $H$ confirms $E$ iff $P_{\text{new}}(H) > P(H)$
- $H$ disconfirms $E$ iff $P_{\text{new}}(H) < P(H)$

Examples of Textbook Bayesianism

- Let's assume that $H$ entails $E$: $P(H|E) = 1$.
- Then 
  $$P_{\text{new}}(H) = P(H)/P(E)$$
- (i) It is easy to see that Bayesianism can account for the insight that surprising evidence (i.e. if $P(E)$ is small) confirms better.
- (ii) Let's assume we have different pieces of evidence: 
  $$P(E) = P(E_1, E_2, \ldots, E_n) = P(E_1) \cdot P(E_2|E_1) \cdot P(E_3|E_2, E_1) \cdots \cdot P(E_n|E_{n-1}, \ldots, E_1).$$
  Hence, less correlated pieces of evidence confirm better: the variety-of-evidence thesis.

We conclude

- Textbook Bayesianism explains very general features of science.
- However, it turns out to be too general as it does not take into account de facto constraints of the scientific practice (such as dependencies between measurement instruments). Taking them into account might lead to different conclusions.
- Moreover, Textbook Bayesianism does not have an account of what a scientific theory is.

Jon Dorling’s version

- Jon Dorling aims at reconstructing specific episodes in the history of science and fitting them into the Bayesian apparatus.
- To do so, one has to assign specific probability values to the hypotheses and likelihoods in question.
- This variant of Textbook Bayesianism is too specific.
**Upshot**

- We need a version of Bayesianism that is not too general (to connect to the practice of science), and not too specific (to gain some philosophical insight).
- It would also be nice to have an account that has a somewhat wider scope, i.e. an account that reaches beyond confirmation theory.

  *Modeling will do the job!*

---

**Modeling and Bayesianism**

- Models help bridging the gap between a general theory (here: Bayesianism) and the scientific practice.
- Bayesianism is taken to be a modeling framework (just like Newtonian Mechanics is).
- The models we will construct are *models of a theory*.
- These models will not be too specific, i.e. we will not assign specific numbers to the probability variables, but explain general features of the methodology of science.

---

**Glymour’s distinction**

(i) **Platonic program**: formulate necessary and sufficient conditions for $x$, come up with general and universal theories that aim explaining everything... and often end up explaining nothing!

(ii) **Euclidian program**: make idealized assumptions and explore the consequences of these assumptions... Just like scientists do it!

---

**Should we go for the Euclidian program in philosophy?**

- certainly not always, but sometimes
- pluralism of methods
III. Bayesian Networks

An example from medicine

- T: Patient has tuberculosis
- X: Positive X-ray

Given information:
- \( t := P(T) = 0.01 \)
- \( p := P(X|T) = 0.95 = 1 - P(\neg X|T) = 1 - \) rate of false negatives
- \( q := P(X|\neg T) = 0.02 = \) rate of false positives

What is \( P(T|X) \)? => Apply Bayes' Theorem

\[
P(T|X) = \frac{P(X|T) P(T)}{P(X|T) P(T) + P(X|\neg T) P(\neg T)} = \frac{p t}{p t + q (1-t)} = \frac{t}{t + (1-t) x} \text{ with } x := \frac{q}{p} = 0.32
\]

A more complicated (=realistic) scenario

Smoking

Visit to ...?

Lung Cancer

Tuberculosis

Dyspnoea

Positive X-Ray

P(S) = 0.23

\( P(B|S) = 0.3 \)

\( P(B|\neg S) = 0.05 \)

Parlance:
- “T causes X”
- “T directly influences X”
What is a Bayesian Network?

- A BN is a directed acyclic graph (dag, for short) with a probability distribution defined over it.
- BNs represent conditional independencies that hold between (sets of) variables, which helps reducing the number of probabilities that have to be fixed to specify the joint probability distribution.
- The Parental Markov Condition is implemented in the network.

What’s so good about BNs?

- Bayesian Networks are intuitive
- Specifying the joint probability distribution \( P(A_1, A_2, \ldots, A_n) \) over \( n \) binary variables requires the specification of \( 2^n - 1 \) probabilities.
- However, in a BN one only has to specify the probability of each node given its parents: \( P(A_i|\text{par}(A_i)) \).

Algorithms

- One can show that the joint probability
  \[
P(A_1, A_2, \ldots, A_n) = P(A_1|\text{par}(A_1)) \cdot P(A_1|\text{par}(A_1)) \cdots P(A_n|\text{par}(A_n))
  \]
- There are even more efficient algorithms to compute whatever probability one is interested in.

How to construct a Bayesian Network model

- Fix a set of variables.
- Specify the probabilistic independencies that hold among them.
- Construct a Bayesian Network and specify all \( P(A_i|\text{par}(A_i)) \).
- Calculate the requested (conditional or unconditional) probabilities.
The proof of the pudding is in the eating...

IV. Example 1: The Variety-of-Evidence Thesis

The question
- Does the variety-of-evidence thesis always hold?
- What happens, for example, if the measurement instruments (that provide us with the evidence for the hypothesis we test) are only partially reliable?

One reading of the thesis
- A hypothesis $H$ was positively tested with an instrument $I_1$.
- Now, a second test should be performed.
- Which scenario is to be preferred?
  - Scenario 1: use another instrument $I_2$ (i.e. the tests are independent)
  - Scenario 2: use the same instrument again (i.e. the tests are dependent)
- Variety-of-Evidence Thesis: Scenario 1 leads to a higher posterior probability and is to be preferred.

But is the thesis true? $\rightarrow$ construct a model!
The Basic Model

\[ P(\text{CON}) = .52 \]
\[ a = .2 \]
\[ a = .52 \]
\[ a = .9 \]

Single vs. Multiple Instruments

The Relative Strength of Confirmation

Use the theory of Bayesian Networks to calculate the posterior probability of the hypothesis for both cases!

To find out which procedure leads to more confirmation, calculate the difference

\[ \Delta P = P'(\text{HYP}|\text{REP}_1, \text{REP}_2) - P(\text{HYP}|\text{REP}_1, \text{REP}_2) \]

After some algebra, one obtains:

\[ \Delta P > 0 \iff 1 - 2(1 - a)(1 - \rho) > 0 \]
Interpretation

There are two conflicting considerations:
1. Independent test results from two instruments yield stronger confirmation than dependent test results from a single instrument.
2. Coherent test results obtained from a single instrument increase our confidence in the reliability of the instrument, which increases the degree of confirmation of the hypothesis.

The variety-of-evidence thesis challenged

Under certain conditions, test results from a single test instrument provide greater confirmation than test results from multiple independent instruments.

Upshot

- We have constructed and analyzed a toy model (which is still very far away from real science).
- The model shows that the variety-of-evidence thesis is not sacrosanct. It can fail, and it does fail under certain circumstances.
- The model is used as an exploratory tool. It suggests something, but the results need to be interpreted (ideally) model independently.
The problem

- Scientific theories change over time. What changes, and what remains?
- The philosophical debate so far has been too much focused on the very big question as to whether a rational reconstruction of theory change can be given.
- My aim is more modest: I would like to explain some aspects of scientific theory change.
- Here are two examples…

Intermezzo: What is a scientific theory?

- Textbook Bayesianism has no account of what a scientific theory is. This is one of its (many) shortcomings.
- I propose a Bayesian account.
Received views

- Syntactic view:
  - theories are linguistic entities
  - sets of assumptions (and their consequences)
- Semantic view:
  - theories are non-linguistic entities
  - realizations of an abstract formalism

The probabilistic view

- Theories are networks of interrelated models.
- Models \((M_i)\) are conjunctions of propositions that account for a specific phenomenon \(P_i\) (e.g. instantiations of laws). One model for each phenomenon. Conditional independence assumptions hold.
- There is a joint probability distribution over all propositional variables \(M_i, P_i\).
- From this, the posterior probability of the theory (given the phenomena) can be obtained.

Representing theories by Bayesian Networks

The empirical and the non-empirical

- What I just presented represents the empirical part of the theory.
- Additionally, a theory has a non-empirical (or heuristic) part that helps us to construct models (cf. theories as modeling frameworks), might be axiomatized, contains (non-probabilified) laws, etc.
(i) The stability of normal science

- In normal science, more and more applications of a theory are considered. The goal here, however, is not to test the theory. It is taken for granted and applied. It is rather the scientist who is tested.
- An immediate problem for the Bayesian:
  - $P(M_1, M_2, M_3) < P(M_1, M_2)$
  - so by adding more and more applications of a theory, the joint probability will sooner or later be below the threshold (if there is any). I refer to this as the conjunction problem.
- So can (the rationality of) normal science be defended in a Bayesian framework?

Coherence

- The conjunction problem can be (dis-)solved by taking evidence ($P_i$) into account. Then it is possible that $P(M_1, M_2, M_3|P_1, P_2, P_3) > P(M_1, M_2|P_1, P_2)$
- Question: When is this the case?
- Answer: There must be enough positive relevance among the models (and the evidence must confirm the models).
- One way to measure how relevant the models are for each other is by examining how well they fit together or cohere.

A theorem

Let $T = \{M_1, \ldots, M_n\}$ and $T' = T \cup \{M_{n+1}\}$. Each model $M_i$ is independently supported by a piece of evidence $E_i$. Let $E = \{E_1, \ldots, E_n\}$ and $E' = E \cup \{E_{n+1}\}$ and let the strength of the evidence be characterized by the likelihood quotient $x_i = P(E_i|M_i)/P(E_i|M_j)$ (for $i = 1, \ldots, n+1$). If (i) all models are supported with the same strength of evidence (i.e. $x_1 = x_2 = \ldots = x_{n+1}$), (ii) the prior probabilities of both theories are the same (i.e. $P(T) = P(T')$), and (iii) $T'$ is more coherent than $T$ according to the procedure developed in Bovens and Hartmann (2003), then the posterior probability of $T'$ is greater than the posterior probability of $T$, i.e. $P(T'|E) > P(T|E)$. (Proof omitted)

How does this theorem help?

- It is plausible that, within normal science, theories become more and more coherent.
- I also arguably plausible that the prior of the theory remains (more or less) constant if a model is added and that all models are (more or less) equally well supported.
- Then the theorem tells us that the posterior probability of the theory increases in the course of normal science. QED.
(ii) Why anomalies hurt so much

- Here I depart from Kuhn.
- Anomalies (such as the discovery of line spectra in the late 19th century) hurt the whole theory, and not just a part of it.
- As a consequence, the whole theory is (and, perhaps, has to be) given up (or at least its domain of applicability has to be restricted).
- What is the rationale for this?

Coherence and confirmation transmission

- In recent work, Dietrich & Moretti (2006) showed that sufficiently coherent sets of propositions transmit confirmation.
- I.e. if $E$ confirms one of the propositions of a theory (which is, remember, a highly coherent system), it also confirms any other proposition of the theory as well as the whole theory (i.e. the conjunction of all propositions).
- This is a rationale for indirect confirmation.

…and disconfirmation

- The same holds, of course, also for disconfirmation.
- So an anomaly (= there is no model for a specific phenomenon or all models fail to account for the phenomenon) hurts the whole theory in normal science (which is, again, highly coherent), and this can be fatal for it.

But what about theory change?

- Here more case studies are needed…
- Examine case studies and extract features that survive theory change, and those that do not, and try to explain this in a formal model.
- Current research:
  - formulate an account of intertheory relations that avoids the extremes - reductionism à la Nagel and an untamed pluralism (à la Cartwright and Dupré).
  - coherence and unification
VII. Naturalized Bayesianism

- Naturalized Bayesianism is a middle ground between
  - aprioristic Textbook Bayesianism, and
  - naturalized philosophies of science
- Bayesianism is a modeling framework, and
- (generalizations from) case studies provide the necessary “empirical” input (which I hope to get from historical studies).

VIII. Conclusions

- Philosophical models are heuristically important and help us explain features of the methodology of science.
- They differ from scientific models as we only have our intuitions to test them.
- There is a reflexive equilibrium between our intuitions and the consequences of the model.
- Though Naturalized Bayesianism provides good explanations, some features of science might resist a Bayesian explanation. We then have to find another account to explain them. Just like in science.

Thanks for your attention!