Knowledge, Reason and Action
PHIL2606

2nd section
Scientific Methodology

Dr Luca Moretti
Centre for Time
University of Sydney

luca.moretti@arts.usyd.edu.au
www.lucamoretti.org

Introduction: what is (scientific) methodology?

- The label "methodology" in philosophy identifies - very roughly - the discipline (a) that investigates whether there are methods to achieve knowledge and (b) that aims to provide a precise description of these methods.

  (Knowledge is usually replaced with less problematic surrogates, such as: justification, warrant, rational acceptability, confirmation, inductive support, and so on).

- Methodology is often conceived of as scientific methodology. The presupposition is that the only method to attain knowledge is the scientific one, and that any other method we might use in everyday life simply approximates to the scientific method.

- Methodology overlaps with both philosophy of science and epistemology.

- The scope of philosophy of science is however wider than the one of methodology, as it also encompasses the metaphysics of science (i.e. the analysis of central scientific concepts, like space, time, causation, etc.) and specific issues such as: scientific realism, theory underdetermination, theory incommensurability, etc.

- The relations between methodology and epistemology are more complex. Often methodology presupposes notions and findings proper to epistemology (for example, the notion of empiricism and the thesis that all knowledge is empirical). On the other hand, epistemology sometimes presupposes notions and findings proper to methodology (for example, the notion of inductive logic or the Bayes' theorem).

Introduction: examples of prominent methodologists

- Aristotle (384BC-322BC)

  He invented syllogistic logic (the ancestor of a branch of deductive logic called predicate logic) and formulated the first version of the principle of induction by enumeration).

- Francis Bacon (1561-1628)

  He formulated a version of what we can call the "experimental method" (a set of practical rules for deciding among rival hypotheses on the grounds of experimental evidence).

- John Stuart Mill (1806-1873)

  He formulated a more modern version of the "experimental method" for the specific purpose of deciding among rival hypotheses that postulate causal relationships between phenomena.

- Rudolf Carnap (1891-1970)

  He defined a formal system of inductive logic based on a mathematical account of the notion of probability. He also gave a quantitative (probabilistic) account of the notion of confirmation.

- Jaakko Hintikka (1929-)

  He is the founder of epistemic logic - a branch of deductive logic that deals with statements including expressions such as 'it is known that…' and 'it is believed that…'.

Introduction: what we will do in this course

- We will focus on the problem of providing a rationally acceptable and philosophically useful formulation of inductive logic.

More specifically:

- We will explore the general possibility of answering the traditional problem of induction by appealing to a system of inductive logic.
- We will examine qualitative and quantitative versions of inductive logic and we will try to evaluate whether they are acceptable and whether they explicate scientific methodology.
- We will consider important objections to the possibility of developing any adequate system of inductive logic and we will examine an alternative non-inductivist account of scientific methodology.

Introduction: plan of the course

- **Lectures 1&2.** Topic: **Inductive Logic and the Problem of Induction**

  Reading list:

  B. Skyrms, Choice and Chance, ch. 1;

  B. Skyrms, Choice and Chance, ch. 2.

- **Lectures 3&4.** Topic: **Qualitative Confirmation**

  Reading list:

  C. Hempel, ‘Studies in the Logic of Confirmation', in his Aspects of Scientific Explanation: and Other Essays in the Philosophy of Science;

  T. Grimes, ‘Truth, Content, and the Hypothetico-Deductive Method’. Philosophy of Science **57** (1990).

- **Lectures 5&6.** Topic: **Falsificationism against Inductive Logic.**

  Reading list:

  J. Ladyman, Understanding Philosophy of Science, ch. 3, ‘Falsificationism’;

  Sections from: I. Lakatos, ‘The methodology of scientific research programmes’ in I. Lakatos and A. Musgrave (eds.), Criticism and Growth of Knowledge.

- **Lectures 7&8.** Topic: **Quantitative Confirmation: Bayesianism**

  Reading list:

  David Papineau, ‘Confirmation', in A. C. Grayling, ed., Philosophy.

  (Additional material will be provide before the lectures).

Lecture 1
What is inductive logic?

Requested reading:
B. Skyrms, Choice and Chance, ch. 1

Relevance of inductive arguments

- Inductive arguments are used very often in everyday life and in science:

  Example 1: I go to Sweden. One day, I speak to 20 people and I find out that they all speak a very good English. I thus infer that the next person I will meet in Sweden will probably speak a very good English.

  Example 2: The general theory of relativity entails that:

  (a) gravity will bend the path of a light ray if the ray passes close to a massive body,

  (b) there are gravitational waves,

  (c) Mercury's orbit has certain (anomalous) features (precession of Mercury's perihelion).

  Scientists have verified many instances of (a), (b) and (c). From this, the have inferred that the general theory of relativity is probably true.

Difference between deductive and inductive logic

- Logic in general is the discipline that studies the strength of the evidential link between the premises and the conclusion of arguments.

- An argument is simply a list of declarative sentences (or statements) such that one sentence of the list is called conclusion and the others premises, and where the premises state reasons to support the claim made by the conclusion.

- A declarative sentence is any one that aims to represent a fact and that can be true or false.

  'Sydney is in Australia' is a declarative sentence.

  'Hey!' and 'How are you?' are not declarative sentences.

- Deductive logic aims to individuate all and only the arguments in which the conclusion is entailed by the premises. Namely, any argument such that if the premises are true, it is logically necessary that the conclusion is true. (This is the highest possible level of evidential support). All these arguments are called deductively valid.

- Inductive logic aims to individuate - roughly - all and only the arguments in which the conclusion is strongly supported by the premises. Namely, any argument such that if the premises are true, it is highly plausible or highly probable (but not logically necessary) that the conclusion is true.

- Any argument can be evaluated by determining (a) whether its premises are de facto true and (b) whether its premises support its conclusion. These two questions are independent.

  Logicians are not interested in (a), they are only interested in (b).

Strength of inductive arguments

- This is a deductively valid argument:

  I live on the Moon and my name is Luca, therefore, I live on the moon.

  This is a deductively invalid argument:

  (*) All 900.000 cats from Naples I have examined so far were in fact cat-robots, therefore, the next cat from Naples I will examine will be a cat-robot.

- All inductive arguments are deductively invalid, and are more or less inductively strong. The strength of an argument coincides with the evidential strength with which the conclusion of the argument is supported by its premises.

  Argument (*) is a strong inductive argument. For if its premise is true, its conclusion appears very plausible.

  The following is instead a weak inductive argument:

  I live on the moon and my name is Luca, therefore, the next cat from Naples I will examine will be a cat-robot.

  This is a even weaker inductive argument:

  All 900.000 cats from Naples I have examined so far were in fact cat-robots, therefore, the next cat from Naples I will examine will not be a cat-robot.

  We can hardly think of an inductive argument weaker than this:

  My name is Luca, therefore, my name is not Luca.

Types of inductive arguments

- A widespread misconception of logic says that deductively valid arguments proceed from the general to the specific and that inductively strong arguments proceed from the specific to the general. This is simply false. Consider these counterexamples:
- A deductively valid argument from general to general:

  All men are mortal, therefore, all men are mortal or British.
- A deductively valid argument from particular to particular:

  John Smith is Australian and Hegel was a philosopher. Therefore, Hegel was a philosopher.
- An inductively strong argument from general to general:

  All bodies on the earth obey Newton's laws. All planets obey Newton's laws. Therefore, all bodies obey in general Newton's laws.
- An inductively strong argument from general to particular:

  All African emeralds are green. All Asian emeralds are green. All Australian emeralds are green. Therefore, the first American emerald I will see will be green.
- An inductively strong argument from particular to particular:

  The pizza I had at Mario's was awful. The wine I drank at Mario's was terrible. The salad I ate at Mario's was really disgusting. The watermelon I had at Mario's was rotten. Therefore, the coffee I am going to drink at Mario's will not probably taste delicious.

Deduction, induction and information

- An essential feature of any deductively valid argument is the following:

  All information conveyed by the conclusion of any such argument is already included in its premises.

  This explains why any deductively valid argument is such that the truth of its premises guarantees the truth of its conclusion.

  This also explains why no deductively valid argument can - strictly speaking - provide us with fresh knowledge.

- An essential feature of any inductive argument is the following:

  At least part of the information conveyed by the conclusion of any such argument is not included in its premises.

  This is why the truth of the premises of any inductive strong argument cannot guarantee the truth of its conclusion.

  This also explains why all inductively strong arguments seem capable to provide us with fresh knowledge.

Distinguishing psychology from logic

- Consider again this argument:

  (a) All 900.000 cats from Naples I have examined so far were in fact cat-robots.

  Therefore:

  (b) The next cat from Naples I will examine will be a cat-robot.

  Inductive logic (and logic in general) does not study the mental process by means of which I arrive at having the belief (b) if I believe (a). This might be investigated by psychology.

  Inductive logic does not provide rules to obtain belief (b) from belief (a).

  Inductive logic gives rules to establish whether the belief (a) justifies the belief (b).

Possible types of inductive logic

- An inductive logic can be purely qualitative.

  We can think of such a logic as a set of rules for singling out all possible arguments in which the premises render the conclusion highly plausible and only these arguments.

  Different qualitative inductive logics will individuate alternative sets of all such arguments.

- An inductive logic can be comparative.

  We can think of such a logic as a set of rules for ordering all possible arguments according to their strength. This logic allows us to say, given any two arguments A and B, whether A is stronger than B or vice versa, or whether A and B are equally strong.

  Different comparative inductive logics will induce alternative orderings over the set of all possible arguments.

- Finally, an inductive logic can be quantitative.

  We can think of it as a set of rules for giving each argument a number, and only one, that represents its degree of strength. The number typically identifies the degree of probability of the argument's conclusion given the truth of its premises. This kind of probability is generally called inductive probability.

  Different quantitative inductive logics will give the same arguments alternative values of strength.

- Any inductive logic can be formal if the language in which the arguments are expressed is formalized (i.e. if there are precise rules for the formation and transformation of statements).

Inductive probability and epistemic probability

- Inductive probability is not a property of single statements but the probability of a statement given other statements (i.e. the property of the conclusion of an argument given its premises). Inductive probability is a relational property of statements.

- Apparently, we can think of the degree of probability of single statements independently of any argument. For example, of the statement:

  (*P*) In Sydney there is a person who speaks 40 different languages.

  If asked, many would say that the probability of *P* is very low.

- But also this kind of probability is in fact relational. For when we are to evaluate *P*'s probability, we should take into account all relevant evidence we have. (For instance, evidence about the linguistic abilities of the average person, about similar cases in history, etc.). Ideally, we should consider all evidence we have.

  We can think of this kind of probability as the probability of a statement given background evidence (of a person in a given time).

  Let us call this kind probability epistemic probability. The degree of epistemic probability of a statement always depends on specific background evidence and changes as the latter changes.

  The epistemic probability of a statement *S* given background evidence *K* is the inductive probability of the conclusion *S* of the argument with premises *K*.

- Epistemic probability is the one really relevant in methodology, as our evaluations of probability will be fully rational only if we consider all relevant information and so all information we have in a given time.

Lecture 2

The justification of inductive logic and the traditional problem of induction

Requested reading:

B. Skyrms, Choice and Chance, ch. 2

The justification of inductive logic

- Let us focus on quantitative inductive logic.

  As I have said, we can think of it as a set of rules for giving each argument of a language a value of strength, which represents the degree of probability of the argument's conclusion given the truth of its premises.

- Suppose we have actually defined one specific set IL of such logical rules. How can we justify the acceptance of IL?

  We should at least show that IL satisfies two conditions:

  (1) The probability assignments of IL accord well with common sense and scientific practice (for instance, in the sense that the arguments that are considered strong or week on an intuitive basis will receive a, respectively, high or low degree of probability).

  In other words, we should show that IL is nothing but a precise formulation (or reconstruction) of the intuitive inductive logic that underlies common sense and science.

  (2) IL is a reliable tool for grounding our expectations of what we do not know on what we do know.

  Both tasks are formidable!

- Notice that if IL satisfies both (1) and (2), we can explain why science and (to some extent) common sense are means of knowledge. (This is an example of how methodology ties up with epistemology and philosophy of science).

The rational justification of inductive logic and the traditional problem of induction

- Suppose we have an inductive logic IL that satisfies condition (1). (Namely, IL accords well with common sense and scientific practice).

  How can we show that IL also satisfies condition (2)? Namely, that IL is a reliable tool for grounding our expectations of what we do not know on what we do know?

  This problem coincides with the so-called "traditional (or classical) problem of induction", which is often described as the problem of providing a rational justification of induction.

- David Hume (1711-1776), in his *An Inquiry Concerning Human Understanding*, first raised this problem in full force; and he famously concluded that this problem cannot be solved.

  Hume interpreted the claim that IL is a reliable tool for grounding our expectations of what we do not know on what we do know in the specific sense that IL is a reliable tool for our predictions of the future. He argued that there is no rational way to show that IL is actually reliable for predictions.

  (Indeed, Hume didn't think of inductive logic as an articulated system of rules, such as IL. He just focused on some basic inductive procedures. His criticism can however be generalized to hit IL as a whole, no matter how IL is specified in detail).

- Hume's problem should carefully be distinguished from the one recently raised by Nelson Goodman (1906-1998) - which is often called the "new riddle of induction". Very roughly, Goodman has argued that, if induction by generalization works, it works "too well". As it justifies crazy generalizations which are obviously false.

Hume's objection (1)

- Hume's argument (i.e. a generalization of it) consists of two steps:

  Step (1), we set up a plausible criterion for the rational justification of our inductive logic IL.

  Step (2), we show that it is impossible to satisfy this criterion.

- Step (1)

  As we know, the epistemic probability of a statement *S* is the inductive probability of the argument that has *S* as conclusion and that embodies all available information in its premises.

  Let us call all inductive arguments that embody all available information in their premises, E-arguments.

  Consider now that if a statement *S* about the future has high epistemic probability (on the grounds of a strong E-argument), it is natural to predict that *S* will prove true. And, more generally, it is natural to expect more or less strongly that *S* will be true as *the* epistemic probability of *S* is, respectively, higher or lower.

  It is also quite natural to believe that strong E-arguments will give true conclusion most of the time. And, more generally, that stronger E-arguments will have true conclusion more often than weaker E-arguments.

  These considerations lead to the following criterion for the rational justification of IL:

  (RJ) IL is rationally justified if and only if it is shown that the E-arguments to which IL assigns high probability yield true conclusions from true premises most of the time, (and that the E-arguments to which IL assigns higher probability yield true conclusions from true premises more often than the arguments to which IL assigns lover probability).

Hume's objection (2)

- Step 2

Let us now show that the criterion (RJ) for the rational justification of IL cannot be satisfied.

(RJ) will be fulfilled if we show that the E-arguments to which IL assigns high probability yield true conclusions from true premises most of the time.

As the conclusions of several of these E-arguments are not yet verified, we should show that many (or most) of them will be verified in the future. We may try to show it by means of (1) a deductively valid meta-argument or (2) an inductively strong meta-argument (where a meta-argument is simply an argument about arguments).

But method (1) will not work.

We want to show that certain contingent statements will prove true in the future by using a deductive meta-argument. To achieve this result, our meta-argument must have contingent premises that we know to be true now. Such premises can concern only the past and the present, but not the future. But then, since all information conveyed by the conclusion of a deductively valid argument must be already included in its premises, no deductive meta-argument could ever establish any contingent truth about the future. No deductive meta-argument can show that contingent statements will prove true in the future.

Method (2) will not work either.

We want to show that most of the E-arguments to which IL assigns high probability will yield true conclusions from true premises, by using an inductively strong meta-argument that moves from the true premise asserting, among other things, that IL worked well in the past. But, since IL is our inductive logic, this strong meta-argument will be one of the E-arguments to which IL gives high probability and that we want to show to be reliable. We are just begging the question!

In conclusion, the criterion (RJ) for the rational justification of IL cannot be satisfied.

Four replies to Hume

- Philosophers have tried to answer Hume's challenge in at least four distinct ways. Precisely:

  (1) They have argued that IL can be rationally justified by appealing to the principle of the uniformity of nature.

  (2) They have insisted that the inductive justification of IL does not beg the question.

  (3) The have tried to provide a pragmatic justification (or vindication) of IL.

  (4) They have suggested that the traditional problem of induction should be "dissolved" as a non-problem rather than resolved.

- Let us examine each of these replies. Unfortunately, none of them appears successful.

The appeal to the principle of the uniformity of nature

- Something like a principle of the uniformity of nature would seem to underlie both scientific and common-sense judgments of inductive strength.
- This principle says that, roughly, nature is uniform in many respects and that, in particular, the future will resemble the past (for instance: material bodies have always been attracting one another, and they will always do it in the future).
- How could this principle rationally justify IL? Suppose, to simplify, that the correct formulation of the principle of the uniformity of nature is the following:

  (UN) If 10.000 objects of the same kind instantiate a given property, then all objects of that kind will always instantiate that property.

  (UN) would explain why certain E-arguments to which IL assigns high probability will actually yield true conclusions from true premises most of the time. Perhaps, arguments of this form:

  > (Among many other observations) property P has been observed in 9.000 objects of type O, therefore, the next object of type O will have the property P.

- But this ingenious reply to Hume is doomed to fail for at least two reasons:

  (a) To begin with, the task of giving an exact formulation of the principle of the uniformity of nature may prove impossible: how can we distinguish in advance between seeming regularities (i.e. mere coincidences) and substantive regularities (e.g. causal links)? To draw such a distinction we should plausibly use concepts and hypotheses embedded in the scientific theory of the universe, which is still in progress.

  (b) More importantly, suppose we give the principle of the uniformity of nature a definite formulation. How could we ever justify our belief in this principle? Clearly, we could not justify our belief by deductive arguments, and the appeal to inductive arguments would beg the question!

The inductive justification of inductive logic (1)

- Brian Skyrms (still alive but very old) has worked out an inductive argument to justify rationally IL.

- As Skyrms himself has acknowledged, this argument is eventually unsuccessful, but not because it begs the question. There is thus a sense in which Hume was wrong!

- Skyrms' argument exploits the fact that inductive arguments can be made at distinct hierarchical levels.

  The first level is that of inductive basic-arguments - that is, arguments about natural phenomena.

  The second level is that of inductive meta-arguments - that is, arguments about basic-arguments.

  The third level is that of inductive meta-meta-arguments - that is, arguments about meta-arguments.

  The fourth level is that of inductive meta-meta-meta-arguments - that is, arguments about meta-meta-arguments.

  And so on.

- According to Skyrms, the success of an inductively strong E-argument made at a given level can be justified by using an inductively strong E-argument made at the successive level.

  This system generates no vicious circularity, for there is no attempt to justify E-arguments made at a given level by assuming that the E-arguments made at that very level are already justified.

The inductive justification of inductive logic (2)

- Here is what Skyrms has in mind. Suppose I have successfully used 10 basic-E-arguments. I can try to justify the claim that my next basic-E-argument (the #11) will also be successful by this strong meta-E-argument:

(M1) (Among many other facts) 10 basic-E-arguments have been successful. Therefore, the basic-E-argument #11 will be successful too.

Notice that I cannot justify the claim that the meta-E-argument (M1) will be successful.

Suppose however that the basic-E-argument #11 proves actually successful in accordance with the prediction of (M1). I can then try to justify the claim that the basic-E-argument #12 will be successful by using a new meta-E-argument:

(M2) (Among many other facts) 11 basic-E-arguments have been successful. Therefore, the basic E-argument #12 will be successful too.

But, again, I cannot justify the claim that the meta-E-argument (M2) will be successful.

Suppose however I keep on using basic-E-arguments and meta-E-arguments in this way until I arrive at the following meta-E-argument:

(M11) (Among many other facts) 20 basic-E-arguments have been successful. Therefore, the basic-E-argument #21 will be successful too.

At this point, I have successfully used 10 meta-E-arguments, and I can try to justify the claim that the meta-E-argument (M11) will be successful by the following strong meta-meta-E-argument:

(MM1) (Among many other facts) 10 meta-E-arguments have been successful. Therefore, the meta-E-argument #11 (i.e. M11) will be successful.

I cannot justify the claim that the meta-meta-E-argument (MM1) will be successful. Yet, after successfully using a sufficient number of meta-meta-E-arguments, I can try to justify the last of them by applying to a meta-meta-meta-argument. This process will continue indefinitely.

Problems of the inductive justification of inductive logic (1)

- Skyrms' method is to the effect that E-arguments made at a given level can be rationally justified by E-arguments made at the successive level. The latter E-arguments can in turn be rationally justified by other E-arguments made at the successive level, and so on indefinitely.

  This method is also to the effect that, at any level *n*, no E-argument will be rationally justified if there is a superior level at which no E-argument is rationally justified. For this would disable necessary components of the inductive "mechanism" by means of which E-arguments at level *n* are credited with rational justification.

  The problem is that, in any given time, none of the E-arguments made at the top level is rationally justified.

  This seems to entail that no E-arguments made at any level will ever be rationally justified by appealing to Skyrms' method.

- Briefly, a problem of Skyrms' method seems to be that the inductive procedure by means which rational justification is credited to E-arguments will never be rationally justified.

Problems of the inductive justification of inductive logic (2)

- Skyrms has himself emphasized that a serious fault of his inductive justification of induction is that it can “justify” counterinductive logic as a reliable instrument for predictions.

  Counterinductive logic is based on the assumption that - roughly - the future will not resemble the past at all.

- Consider for instance a counterinductive logical system CIL such that:

  (1) it assigns a high degree of probability to any argument to which IL assigns a low degree of probability

  (2) it assigns a low degree of probability to any argument to which IL assigns a high degree of probability.

  So, for example, the following argument will be ranked as strong on CIL:

  > None of the 100 hungry grizzly bears I have examined so far was very convivial and friendly to me. Therefore, the next one will certainly be.

- CIL is simply a crazy logical system - it is strongly intuitive that CIL allows for completely unreliable predictions and that CIL cannot be justified rationally.

  Yet Skyrms’ method does allow us to “justify” CIL as well as IL (I am putting aside the objection to Skyrms’ method considered before).

- A sensible conclusion we should draw is that the kind of “justification” provided by Skyrms’ method, whatever it might be, is certainly not rational justification.

Problems of the inductive justification of inductive logic (3)

- This is the way in which CIL can be "justified". Suppose I have unsuccessfully used 10 basic-E-arguments of CIL. I can try to justify the claim that my basic-E-argument #11 of CIL will eventually be successful by this strong meta-E-argument of CIL:

  (M1) (Among many other facts) 10 basic-E-arguments of CIL have failed. Therefore, the basic-E-argument #11 will be successful.

  But I cannot justify the claim that the meta-E-argument (M1) of CIL will be successful.

  Suppose however that the basic-E-argument #11 proves unsuccessful, against the prediction of (M1). I can then try to justify the claim that the basic-E-argument #12 of CIL will be successful by using a new and stronger meta-E-argument of CIL:

  (M2) (Among many other facts) 11 basic-E-arguments of CIL have failed. Therefore, the basic-E-argument #12 of CIL will be successful.

  But, again, I cannot justify the claim that the meta-E-argument (M2) of CIL will succeed.

  Suppose however I keep on using basic-E-arguments and meta-E-arguments in this way until - after many failures - I arrive at the following meta-E-argument:

  (M11) (Among many other facts) 20 basic-E-arguments of CIL have failed. Therefore, the basic-E-argument #21 of CIL will be successful.

  At this point, I have unsuccessfully used 10 meta-E-arguments of CIL, and I can try to justify the claim that the meta-E-argument (M11) of CIL will be successful by the following strong meta-meta-E-argument:

  (MM1) (Among many other facts) 10 meta-E-arguments of CIL have failed. Therefore, the meta-E-argument #11 (i.e. M11) of CIL will be successful.

  I can continue this crazy process indefinitely!

The pragmatic vindication of inductive logic

- The pragmatic solution of the traditional problem of induction was first proposed in papers by Herbert Feigl (1902-1988) and Hans Reichenbach (1891-1953).

- Pragmatists believe that we cannot show that the system IL of common sense and scientific inductive logic is actually reliable for predictions. They believe - consequently - that we cannot rationally justify IL, in the strong sense of this expression.

  However, pragmatists contend that we can justify - or vindicate - IL in a weaker sense. Precisely, it would be possible to show that:

  (PV) If there exist any inductive logic X which is a reliable device for predictions (which is not guaranteed!), then IL is also a reliable device for predictions.

  Pragmatists believe that we had better accept IL. For IL might actually be successful and, given (PV), no other inductive logic has better chances to succeed than IL.

- Roughly, the pragmatist argument for showing (PV) runs as follows: Suppose there is an inductive logic X, different from IL, which is a reliable device for predictions. This means that (a) X has been reliable many times in the past and that (b) X will be reliable in the future. But then, given the truth of (a) and (b), the strong inductive E-argument of IL:

  (Among many other observations) X has been observed to be reliable many times in the past, therefore, X will be reliable in the future,

  would actually give a true prediction! In the same way, IL would successfully predict the success of each strong inductive E-argument of X. Thus, after all, if X is a reliable instrument for prediction, IL is also so. This would demonstrate (PV).

Why the pragmatic reply fails

- Unfortunately, the pragmatist argument, when made fully explicit, appears incorrect.

  Notice that the methodologically relevant interpretation of (PV) is the following:

  (PV*) If there exist any inductive logic X which is a reliable device for the predictions of natural phenomena in general, then IL is also a reliable device for the predictions of natural phenomena in general.

  The problem is that (PV*) is false. As the fact that IL would be reliable in predicting the success of X in predicting, in turn, the behavior of natural phenomena in general does not entail that IL would also be a reliable device for the predictions of natural phenomena in general.

  IL would be successful in making predictions only about a very limited range of natural phenomena (e.g. about some of our possible practices and their empirical success), but IL could not be used to predict correctly most natural phenomena.

  Consider finally that IL would be utterly useless if the inductive logic X, though existent, were unknown to us.

- In conclusion it is not so evident that - as pragmatists have argued - we had better accept IL. For, though IL might actually be successful, we cannot discard the possibility that an alternative inductive logic might have better chances than IL.

The dissolution of the traditional problem of induction

- Some philosophers have argued that no argument whatsoever is necessary to justify rationally a system of inductive logic IL that comply with common sense and scientific practice. This reply to Hume typically comes in one of these three forms:

  (1) What we are looking for when we try to justify rationally IL is the guarantee that the E-arguments that IL ranks as strong will always give us true conclusions from true premises. But this is absurd, as induction is not deduction! So, we should not seek to justify rationally IL.

  (2) Anyone who doubts the rationality of accepting IL does not understand the words she is using. For to be rational just entails accepting IL. Thus, there is no need to justify rationally IL.

  (3) Asking for the rational justification of IL means asking beyond the limits where justification makes sense. For it is impossible to justify rationally IL. Looking for a rational justification of IL makes simply no sense.

- But none of these replies is fully convincing:

  Argument (1) simply misrepresents the problem. Rationally justifying IL means showing that the strong E-arguments of IL give true conclusions from true premises just most of the time, and not always. There is no conflation between induction and deduction.

  Argument (3) is based on the undemonstrated assumption that it is impossible to justify rationally IL. (Many will have the feeling that an improved version of the pragmatic or of the inductive justification of induction might eventually succeed).

  Argument (2) is the most sophisticated but also dubious. One problem is that it seems to presupposes a form of cultural relativism, which will be rejected by those who believe rationality to be objective and trans-cultural. Another problem is that it presupposes a “static” conception of rationality. Many of our norms are vague, unreasonable and incoherent. We do not apply them mechanically, but rather interpret, criticise and improve them. Our conception of rationality seems to evolve through a process of self-criticism. To be rational does not just entail accepting IL.

Conclusions

- The traditional problem of induction is (or can be formulated as) that of showing that the inductive logical system IL, which accords well with common sense and scientific practice, is a reliable tool for predictions.
- Hume argued that to accomplish this task is impossible. For if we appeal to a deductive argument, the argument will prove deductively invalid, and if we appeal to an inductive argument, the argument will prove viciously circular.

- We have considered four possible replies to Hume:

  (1) The attempt to justify IL's predictive power by appealing to the principle of the regularity of nature. This attempt fails because we know neither how to formulate this principle nor how to justify it.

  (2) The attempt to justify IL's predictive power by an inductive procedure that distinguishes different levels of induction; this procedure does not beg the question. But this reply is ineffective because the overall inductive procedure will never be shown to be reliable, and because if this procedure worked out, it would also "justify" as a reliable tool for predictions a logical system which is incompatible with IL and that is utterly absurd.

  (3) The pragmatic invitation to "bet" on IL by arguing that, if any inductive logic is a reliable device for predictions, IL is also so. But this conditional is false - we can think of possible worlds in which IL is not reliable in predicting natural phenomena. The pragmatic "bet" does not seem justified.

  (4) The proposal to dismiss or dissolve the traditional problem of induction. But this proposal rests either on a misinterpretation of the problem, or on the dogmatic acceptance of a form of cultural relativism and on a "static" and false conception of rationality, or on the dogmatic assumption that the traditional problem of induction cannot possibly be solved. This proposal cannot be accepted.

- In conclusion, the traditional problem of induction is still open.

Knowledge, Reason and Action
PHIL2606

2nd section
Scientific Methodology

Dr Luca Moretti
Centre for Time
University of Sydney

luca.moretti@arts.usyd.edu.au
www.lucamoretti.org

Plan of the course

- **Lectures 1&2.** Topic: **Inductive Logic and the Problem of Induction**
  Reading list:
  B. Skyrms, *Choice and Chance*, ch. 1;
  B. Skyrms, *Choice and Chance*, ch. 2.

- **Lectures 3&4.** Topic: **Qualitative Confirmation**
  Reading list:
  C. Hempel, 'Studies in the Logic of Confirmation', in his *Aspects of Scientific Explanation: and Other Essays in the Philosophy of Science*;
  T. Grimes, 'Truth, Content, and the Hypothetico-Deductive Method'. *Philosophy of Science* **57** (1990).

- **Lectures 5&6.** Topic: **Falsificationism against Inductive Logic.**
  Reading list:
  J. Ladyman, *Understanding Philosophy of Science*, ch. 3, 'Falsificationism';
  Sections from: I. Lakatos, 'The methodology of scientific research programmes' in I. Lakatos and A. Musgrave (eds.), *Criticism and Growth of Knowledge*.

- **Lectures 7&8.** Topic: **Quantitative Confirmation: Bayesianism**
  Reading list:
  D. Papineau, 'Confirmation', in A. C. Grayling, ed., *Philosophy.*
  (Additional material will be provide before the lectures).

Confirmation and inductive logic

- Suppose you set up an inductive logic consisting of a set of rules for singling out all possible arguments in which the premises render the conclusion highly plausible and only these arguments.

  Suppose that these arguments are such that their conclusion are hypotheses or theories and their premises are reports of observation (or, more generally, evidential statements).

  Such a logic would be a logic of confirmation.

- There are two very general notions of confirmation:

  Absolute confirmation (presupposed above): evidential statement *E* confirms in an absolute sense hypothesis *H* if the truth of *E* makes *H* highly plausible (or plausible over a stipulated threshold of probability).

  Relative confirmation: *E* confirms in a relative sense *H* if the truth of *E* simply increases the plausibility of *H.*

- It is possible to define both logics of absolute confirmation and logics of relative confirmation.

  A logic of confirmation can be merely qualitative or comparative or quantitative. Furthermore, a logic of confirmation can be either formal or non-formal.

- The expression confirmation theory is often used as a synonym of logic of confirmation.

Hempel’s and Grimes’ logics of confirmation

- In these two lectures, we will be examining - among other things - the logics of confirmation defined by Hempel and by Grimes.

  Both logics are merely qualitative and deal with relative confirmation.

  Moreover, both logics are formal - that is, the statements that instantiate the relations of confirmation are expressed in a formalized language.

- What both Hempel and by Grimes aim at is defining a set of rule to establish whether an evidential statement confirms (i.e. makes more plausible) a hypothesis on the grounds of the mere logical structure of the evidential statement and the hypothesis. (The contents of evidential statements and hypotheses are irrelevant for applying these rules).

- Neither Hempel nor Grimes is interested in providing something like a rational justification of, respectively, his logical system.

  Hempel and Grimes are just interested in working out a rational reconstruction of the real confirmation practices of scientists.

- Neither Hempel nor Grimes has succeed (or has completely succeed) in this purpose. We can however learn a lot from their mistakes and from the problems they have uncovered.

Qualitative confirmation

Lecture 3
Hempel's contribution to the logic of confirmation

Requested reading:

C. Hempel, 'Studies in the Logic of Confirmation'.

The relevance of Hempel's work

- 'Studies in the logic of confirmation', by Carl Hempel (1905-1997), was first published in 1945.
- The specific logic of confirmation put forward in this paper is today considered unacceptable because - principally - counterintuitive in some respects and too narrow in the range of application. Yet Hempel's paper is still important today, as it sets out a general conceptual framework for the analysis of the notion of confirmation.
- In particular, the necessary conditions for any logic of confirmation put forward in this article are hotly debated still nowadays.
- Another reason of the importance of this paper rests on Hempel's discussion of a surprising confirmation paradox - the so called Paradox of the Ravens - which is still discussed in contemporary philosophy of science.

  (After Hempel's paper, many essays in methodology were dedicated to the analysis of paradoxes of confirmation).

Hempel's project

- Hempel focuses on empirical statements - that is, statements that can in principle be tested because it is possible to state in advance what experiential findings would constitute favorable or unfavorable evidence for them.

  Any evidence favorable to a given statement confirms that statement, and any evidence unfavorable to a statement disconfirms that statement. (Evidence is irrelevant to a statement if it neither confirms nor disconfirms that statement).

- Hempel's project is providing a formal logic of confirmation parallel to formal deductive logic. This logic of confirmation is conceived of as a set of rules for determining whether an empirical statement confirms or disconfirms another empirical statement on the basis of their mere logical form.

  Hempel believes that objectivity and impersonality - and so rationality - require independence from content and reliance on only formal structure.

- Hempel's logic of confirmation is only qualitative but is supposed to "pave the way" to more sophisticated, quantitative and comparative versions of it (which Hempel never produced).

Logic of confirmation and actual science

- The logic of confirmation is supposed to be a rational reconstruction of the practices of confirmation and disconfirmation of scientists.

  Precisely, such a rational reconstruction is supposed to expose the normative principles underlying these practices and somehow implicit in them.

- The logic of confirmation is an abstract model of (aspects of) the research behavior of scientists. As any abstract model, it must take into account of the characteristics of actual scientific procedure but it also contains idealized elements that cannot really be observed in the behavior of actual scientists. (For instance, actual scientists typically infringe many methodological rules, as they make mistakes and, for different reasons, not always behave in a fully rational way!).

- According to Hempel, it is possible to distinguish three phases in the scientific activity of testing a hypothesis *H* (not necessarily distinct in real science):

  (1) Acceptance of observation reports describing the results of suitable experiments or observations.

  (2) Ascertaining whether observation reports confirm, disconfirm or are irrelevant to *H*.

  (3) Deciding whether accepting or rejecting *H* or suspending judgment about *H* on the grounds of the strength of the evidential support of the observation reports and other epistemological factors (e.g. the degree of simplicity and coherence of the hypothesis, its explanatory power, whether total evidence also confirms the hypothesis etc.)

  The logic of confirmation is meant to provide a rational reconstruction of phase (2).

The formal language of the logic of confirmation

- For Hempel, confirmation is a relation between two statements: an observation report and a hypothesis.
- Confirmation is thus an intra-linguistic relation and not a relation between a statement (a hypothesis) and an extra-linguistic entity (an observation or a fact).

  This conception of confirmation allows to formulate the logic of confirmation in parallel to deductive logic, in which the relation of entailment is also intra-linguistic.

- From a formal point of view, the language of the logic of confirmation coincides with the one of standard predicate logic.

  It essentially includes: predicate constants (P, Q, R, …), individual constants (a, b, c, …), individual variables (x, y, z, …), the universal and the existential quantifiers ((…) and $\exists$), and the connective symbols of negation, conjunction, disjunction and material implication ($\sim$, &, $\vee$, $\supset$). [Notice that Hempel uses '•', instead of '&', for the the symbol of conjunction]

  For instance, if P means '…is fat' and a refers to John, P(a) means 'John is fat', while ~P(a) means 'John is not fat' (literally, 'it is not the case that John if fat').

  If P means '…is fat' and Q means '…is a man', $\exists x(Q(x) \,\&\, P(x))$ means 'there is a fat man' (literally: 'there is something x who is fat and is a man').

  If R means '… eats …', a refers to John and b refers to Bill, R(a,b) means 'John eats Bill'..

Pragmatic restrictions to the language of the logic of confirmation

- The language of the logic of confirmation can be an idealization of actual scientific language (and not, for instance, of political language), if the types of predicates and of individual constants used in this logic are subject to suitable restrictions.

- The evidence adduced in support of a scientific hypothesis typically consists in data accessible to direct observation. Where 'direct' is to be conceived of in a loose sense to allow observation instruments - like microscope and telescope - considered sufficiently reliable by the scientific community.

  Besides, scientific hypotheses can be about directly observable facts and properties (in the same loose sense) or about unobservable facts and properties.

- Hempel stipulates, therefore, that the language of the logic of confirmation includes:

  (observational language)

  \- observational predicates, which refer to directly observable properties and relations (such as '… burns with a yellow light' and '… is taller than …');

  \- observational individual constants that refer to directly observable objects (such as 'Saturn' and 'this cat');

  (non-observational language)

  \- non-observational predicates that refer to unobservable properties and relations (such as '… is a quark');

  \- non-observational individual constants that refer to unobservable objects (such as 'that black hole').

Hypotheses and observation reports

- A hypothesis is any correctly formulated statement of the language of the logic of confirmation. Hypotheses include typically - but not necessarily - quantifiers.

  For instance, if P means '… is a kangaroo' and Q means '…is an insect', the following is a hypothesis:

  $$(x)(P(x) \supset Q(x)),$$

  which means: 'All kangaroo are insects' (literally: 'for every *x*, if *x* is a kangaroo, *x* is an insect).

- An observation report is a logical conjunction of more elementary statements, called observation statements, each of which just asserts or denies an observation predicate of one or more observable individuals.

  Therefore, an observation report is any correctly formulated statement of the language of the logic of confirmation such that it can include only the logical connectives & and ~, observational predicates and observational individual constants. (No observation report can include quantifiers!)

  For instance, if P means '… is a kangaroo', R means '… eats …', a refers to John and b refers to Bill, the following is an observation report:

  $$P(a)\ \&\ {\sim}R(a,b),$$

  which means: 'John is a kangaroo and John does not eat Bill'.

- Notice that all observation reports are hypotheses, but the reverse is false.

Nicod's criterion of confirmation

- Before proposing his own conception of confirmation, Hempel examines two conception very popular at that time. The first coincides with the so called Nicod's criterion of confirmation.
- This criterion applies only to universal conditional; for instance, to a statement like this:

  (1) $(x)(\text{Swan}(x) \supset \text{White}(x))$

  Which means 'All swans are white' (literally: for everything x, if x is a swan, x is white).

- Nicod's criterion states that:

  (i) an observation report confirms a hypothesis with the form of a universal conditional if and only if the observation report asserts that an individual satisfies both the antecedent and the consequent of the hypothesis.

  (ii) an observation report disconfirms a hypothesis with the form of a universal conditional if and only if the observation report asserts that an individual satisfies the antecedent but not the consequent of the hypothesis.

  Thus, (1) is for instance confirmed by:

  Swan(a) & White(a) (i.e. 'a is a swan and a is white'),

  and disconfirmed by:

  Swan(b) & ∼White(b) (i.e. 'b is a swan and b is not white').

- Nicod's criterion can be extended to apply to conditionals including more than one universal quantifiers. For instance:

  (2) $(x)(y)\ ((\text{Cat}(x)\ \&\ \text{Dog}(y)) \supset \text{Hate}(x,y))$

  Which means 'All cats and dogs hate reciprocally'. (2) is confirmed by Cat(d) & Dog(f) & Hate(d,f) and disconfirmed by Cat(g) & Dog(h) & ∼Hate(g,h).

Shortcomings of Nicod's criterion (1)

- The first, obvious shortcoming of Nicod's criterion is that it applies just to universal conditional, while scientific hypotheses may have many different logical forms.

  Here are examples of scientific hypotheses that are not universal conditionals:

  - Meningitis is caused by a virus - formally: $\exists x(\text{Virus}(x) \,\&\, \text{Cause-of-meningitis}(x))$.
  - All light rays have the same speed - formally: $\exists x(y)(\text{Light-ray}(y) \supset \text{Speed}(x,y))$.

  We cannot apply Nicod's criterion to confirm either of these simple hypotheses!

- Nicod's criterion is defective in a subtler sense too.

  Consider two logically equivalent statements *H* and *H** (i.e. such that *H* entails *H**, and *H**entails *H*). *H* and *H** have the same content. If *H* and *H** are hypotheses, they should be seen as alternative formulations of the same hypothesis.

  It is strongly intuitive - and accepted by scientists - that if an observation report *E* confirms a hypothesis *H*, *E* must also confirm any hypothesis *H** logically equivalent *H.*

  The confirmation of a hypothesis should depend on only the content of the hypothesis and not on its formulation.

  Yet Nicod's criterion makes the confirmation of a hypothesis depend on also its formulation.

Shortcomings of Nicod's criterion (2)

- Here are two examples of how Nicod confirmation is content-dependent.

- Consider these logically equivalent statements: 'All ravens are black' and 'Whatever is not black is not a raven'. They correspond, respectively, to the following universal conditionals:

  (H) $(x)(\text{Raven}(x) \supset \text{Black}(x))$;

  (H^*) $(x)(\sim\text{Black}(x) \supset \sim\text{Raven}(x))$.

  Any evidence for H must be equal evidence for H^*, and *vice versa*.

  Consider however the following observation reports:

  (E) Raven(a) & Black(a);

  (E^*) ~Black(b) & ~Raven(b).

  On Nicod's criterion, E does confirm H but does not confirm or disconfirm H^*. On the other hand, on the same criterion, E^* does confirm H^* but does not confirm or disconfirm H.

- The fact that E^* should count as evidence for H may appear paradoxical (this is the famous paradox of the ravens). But Hempel is convinced that the impression of a paradoxical situation is just a psychological illusion depending on not understanding the actual content of H, which is identical to the one of H^*. Many philosophers do not agree with Hempel, and whether E^* should confirm H is still an open issue.

- The second example of how Nicod confirmation is content-dependent is less controversial. Consider this universal conditional:

  (H^{**}) $(x)((\text{Raven}(x) \,\&\, \sim\text{Black}(x)) \supset (\text{Raven}(x) \,\&\, \sim\text{Raven}(x)))$;

  H^{**} is logically equivalent to H and H^*. Yet in contrast with H and H^*, H^{**} cannot possibly be Nicod-confirmed by any observation report whatsoever. For no individual could ever satisfy the consequent of H^{**}, which is logically inconsistent.

The prediction criterion of confirmation

- The second conception of confirmation analyzed by Hempel coincides with the so called prediction criterion of confirmation.
- This criterion is based on the consideration that since hypotheses are generally used as devices for predictions, any verified prediction of a hypothesis should count as a confirmation of that hypothesis and any falsified prediction should count as a disconfirmation of it.

- (PREDICTION CRITERION OF CONFIRMATION) If *H* is a hypothesis and *E* an observation report, then:

  - *E* confirms *H* if *E* can be divided into exactly two observation statements *E1* and *E2* (the conjunction of which is equivalent to *H*) such that *E2* can be logically deduced from *E1* in conjunction with *H*, but not from *H* alone.

  - *E* disconfirms *H* if *E* logically contradicts *H*.

  - *E* is neutral to *H* if *E* neither confirms nor disconfirms *H*.

- Examples. Let *H* be the hypothesis that all metals, when heated, expand. Formally:

$$(x)((\text{Metal}(x) \mathbin{\&} \text{Heated}(x)) \supset \text{Expand}(x)).$$

The observation report Metal(a) & Heated(a) & Expand(a) confirms *H*. For, it can be divided into the two conjuncts Metal(a) & Heated(a) and Expand(a), such that Metal(a) & Heated(a) in conjunction with *H* entails Expand(a).

The observation report Metal(a) & Heated(a) & ~Expand(a) disconfirms *H*. For the former logically contradicts the latter.

Finally, the observation report Cat(b) & Dog(c) & Hate(a,c) is neutral to *H*.

Shortcoming of the prediction criterion of confirmation

- One problem with this criterion is that it is too narrow to serve as a general definition of confirmation, as it does not apply to hypotheses that refer to unobservable entities.
- Consider the hypothesis *H* asserting that if electrons are flowing in a piece of iron, the latter heats up; formally:

  $$(x)((\text{Iron}(x) \ \&\ \text{Electron-Flow}(x)) \supset \text{Heating-up}(x)).$$

  The problem is that there is no observation report that can be used to confirm or disconfirm *H* on the prediction criterion of confirmation.
- For instance, the statement:

  Iron(a) & Electron-Flow(a) & Heating-up(a)

  can be divided into the two parts Iron(a) & Electron-Flow(a) and Heating-up(a) such that *H* in conjunction with Iron(a) & Electron-Flow(a) entails Heating-up(a).

  But this does not fulfill the the prediction criterion of confirmation. For Electron-Flow(a) is no observation statement. So, Iron(a) & Electron-Flow(a) & Heating-up(a) is no observation report.
- It might be thought that there is some observation statement *S* that entails the statement Electron-Flow(a). If this were true, the conjunction

  Iron(a) & *S* & Heating-up(a)

  would count as an observation report.

  But this is false: no finite conjunction of observation statements logically entails (without auxiliary assumptions) any non-observational (or theoretical) statement. Observation statements can only make a theoretical statement more or less probable - they can only confirm or disconfirm it.

Hempel's conditions of adequacy for any definition of confirmation (1)

- A problem of both Nicod's criterion of confirmation and the prediction criterion of confirmation is that they are not applicable to every scientific hypothesis, so they are not suitable as general definitions of confirmation.
- A basic condition of adequacy for any (general) definition of confirmation is that it must be applicable to any scientific hypothesis independently of its logical form and of the terms and predicates it might include.
- Furthermore any definition of confirmation must be materially adequate - that is, it should be a good approximation to the notion of confirmation implicit in scientific procedures.
- Hempel believes that there are five further necessary conditions that any adequate definition of confirmation must satisfy. Precisely:

(1) EQUIVALENCE CONDITION. Whatever confirms (disconfirms) one of two logically equivalent hypotheses, also confirms (disconfirms) the other.

This condition is justified on the simple consideration that two logically equivalent hypotheses are in fact the very same hypotheses.

(2) ENTAILMENT CONDITION. Any hypothesis which is entailed by an observation report is confirmed by it.

Intuitively, a statement *A* confirms a statement *B* if the truth of *A* makes the truth of *B* more plausible. If *A* just entails *B*, the truth of *A* will verify *B*. So, what is described by the ENTAILMENT CONDITION is - intuitively - a limiting but sound case of confirmation.

Hempel's conditions of adequacy for any definition of confirmation (2)

(3) CONSEQUENCE CONDITION. If an observation report confirms every one of a class *K* of hypotheses, then it also confirms any sentence which is a logical consequence of *K*.

This condition seems justified on the consideration that any consequence of *K* is nothing but an assertion of all or part of the combined content of the original hypotheses in *K*. Thus, if that content is confirmed, it is intuitive that any consequence of *K* will be confirmed too.

(4) SPECIAL CONSEQUENCE CONDITION. If an observation report confirms a hypothesis *H*, then it also confirms every consequence of *H*.

This is just a special case of the CONSEQUENCE CONDITION. (Notice that the prediction criterion of confirmation violates this condition. For if an observation statement *E1* and a hypothesis *H* jointly entail another observation statement *E2,* it is not guaranteed that *E1* and any given consequence *H** of *H* will jointly entail *E2).*

(5) CONSISTENCY CONDITION. Every observation report is logically compatible with the class of all the hypotheses which it confirms.

This conditions has two important consequences:

(5a) No observation report confirms any hypothesis with which is not logically compatible.

(5b) No observation report confirms any hypotheses that contradict each other.

The consistency condition is counterintuitive

- The CONSISTENCY CONDITION entails that no observation report confirms any hypotheses that contradict each other. But this consequence is counterintuitive, and so is the CONSISTENCY CONDITION.

  Consider the following example: I see a person who lies on the floor without moving. This observation report, intuitively, confirms at least three hypotheses that contradict each other:

  - (a) that person is sleeping;

  - (b) that person has lost consciousness;

  - (c) that person is dead.

  Similar examples could easily be find in science.

  For example, the observation statement that material bodies attract each other appears to confirm both Newton's mechanics and the General Relativity, which contradict each other!

- Today, most methodologists reject the CONSISTENCY CONDITION.

  (It would be easy to show that rejecting the CONSISTENCY CONDITION involves rejecting both the CONSEQUENCE CONDITION and the SPECIAL CONSEQUENCE CONDITION).

- Today, most methodologists accept a weaker version of the CONSISTENCY CONDITION, which states that no observation report confirms two hypotheses such that one is the logical negation of the other.

  For instance, if I see a person who lies on the floor, this observation cannot confirm - intuitively - both these statements: 'that person is sleeping' and 'that person is not sleeping'.

Hempel on the converse consequence condition

- Hempel considers a possible further condition for the adequacy of the notion of confirmation:

  CONVERSE CONSEQUENCE CONDITION. If an observation report *E* confirms a hypothesis *H*, then *E* also confirms every hypothesis *H** that entails *H*.

- Although some methodologists accept the CONVERSE CONSEQUENCE CONDITION (for instance, the hypothetico-deductivists) Hempel does not.

  His reason is subtle: If we accept the CONVERSE CONSEQUENCE CONDITION together the ENTAILMENT CONDITION and the SPECIAL CONSEQUENCE CONDITION, we obtain the absurd result that any observational report confirms any hypothesis whatsoever!

- Let us show, for instance, that the observation report *E* = 'this pig does not fly' confirms the hypothesis *H* = 'Santa Claus will appear very soon in Sydney'.

  Proof. *E* entails itself. Therefore, by the ENTAILMENT CONDITION, the observation report *E* confirms the hypothesis *E*. But the hypothesis *E* is in turn entailed by the hypothesis *E* & *H*. Therefore, by the CONVERSE CONSEQUENCE CONDITION, the observation report *E* also confirms *E* & *H.* Finally, the hypothesis *E* & *H* entails the hypothesis *H.* Therefore, by the SPECIAL CONSEQUENCE CONDITION, the observation report *E* confirms the hypothesis *H*. QED.

- Against Hempel, it can be argued that this proof presupposes accepting the SPECIAL CONSEQUENCE CONDITION, which should instead be rejected together the CONSISTENCY CONDITION. Yet Hempel's proof can be re-formulated by using only the unproblematic ENTAILMENT CONDITION and the CONVERSE CONSEQUENCE CONDITION.

Hempel's definition of confirmation (1)

- As I have said, Hempel thinks of confirmation as a logical relation between statements just as the one of logical entailment.

  Whether *A* entails *B* does not depend on whether *A* is true or not true. Analogously, for Hempel, whether *A* confirms *B* does not depend on whether *A* has been verified or has not been verified.

  For Hempel, *A* confirms B if, intuitively, *A* would raise our confidence in *B*, if *A* were verified.

  For Hempel, whether *A* confirms *B* is a logical problem, while whether *A* is actually verified is a pragmatic problem, which is not investigated by the methodologist.

- Actual scientists are certainly interested in whether their observation reports are verified. But this belongs - according to Hempel - to the third phase of testing a hypothesis.

  While the second phase coincides with evaluating whether given observation reports confirm a hypothesis (in the specified sense), the third phase consists in deciding whether accepting or rejecting the hypothesis on the grounds of - among other elements - the verification of the confirming observation reports.

- Hempel's conception of confirmation is meant to be compatible with general fallibilism. Namely, the position (which Hempel does not reject) that no scientific statement can be verified or falsified in a definitive way.

  General fallibilism entails that there are no irrevocably accepted observation reports. For instance, our initially accepted observation reports may be rejected if the theories about the observation instruments we have used are shown to be incorrect or unreliable.

Hempel's definition of confirmation (2)

- Hempel's basic intuition is that an observation report confirms a hypothesis if the hypothesis would be true in case the universe included only the objects described in the observation report.
- For instance, consider again the hypothesis:

  (H) (x) (Raven(x) $\supset$ Black(x)).

  Consider now this observation report:

  (E) Raven(a) & Black(a) & ~Raven(b) & Black(b).

  The objects described in the observation report E are only two: a and b. E confirms the hypothesis H because H would be true if the universe included only a and b.

- To turn this intuition into a definition we need to appeal to the notion of the development of a hypothesis H for a finite class C of individuals. This notion (which could be defined precisely by a logical technique called recursion) roughly corresponds to the notion of what H would assert if the universe included only the objects of C.

  Consider the hypothesis H = (x) (P(x) & ~Q(x)) ('every object has the property P and does not have the property Q') and the class of objects C = {a, b}. The development of H for C is the statement: P(a) & ~Q(a) & P(b) & ~Q(b).

  Consider now the hypothesis H^* = (x) (P(x) v Q(x)) ('every object has the property P or the property Q') and, again, the class C = {a, b}. The development of H^* for C is the statement:

  ((P(a)&Q(a)) v(P(a)&~Q(a)) v (~P(a)&Q(a))) & ((P(b)&Q(b)) v (P(b)&~Q(b) v (~P(b)&Q(b))).

Hempel’s definition of confirmation (3)

- Hempel provides a two-step definition of confirmation (called satisfaction criterion of confirmation):
  (a) An observation report *E* directly confirms a hypothesis *H* if *E* entails the development of *H* for the class of objects described in *E*.
  (b) An observation report *E* confirms a hypothesis *H* if *H* is entailed by a class of sentences each of which is directly confirmed by *E*.

  The “detour” through (a) and (b) is necessary to make this definition of confirmation comply with all five necessary conditions of adequacy considered before. (That this is actually so has been shown by Hempel in another paper).

- The substance of Hempel’s definition is however contained in (a) (notice that if *E* directly confirms *H*, *E* confirms *H*). For (a) embodies the basic intuition that an observation report confirms a hypothesis if the hypothesis would be true in case the universe included only the objects described in the observation report.

  Consider again the hypothesis *H* = (x) (P(x) v Q(x)) and the observation report *E* = P(a) & Q(a) & P(b) & Q(b). The class of objects described in *E* is {a, b}. *E* (directly) confirms *H* because *E* entails the development of *H* for {a, b}, that is:
  ((P(a)&Q(a)) v (P(a)&~Q(a)) v (~P(a)&Q(a))) & ((P(b)&Q(b)) v (P(b)&~Q(b) v (~P(b)&Q(b))).

- Hempel also defines disconfirmation, neutrality, verification and falsification:
  - An observation report *E* disconfirms a hypothesis *H* if *E* confirms the logical negation of *H*.
  - An observation report *E* is neutral to a hypothesis *H* if *E* neither confirms nor disconfirms *H*.
  - An observation report *E* verifies a hypothesis *H* if *E* entails *H*.
  - An observation report *E* falsifies a hypothesis *H* if *E* is logically incompatible with *H*.

Shortcomings of Hempel's definition of confirmation

- The satisfaction criterion of confirmation (i.e. Hempel's definition of confirmation) is, some sense, a good approximation to the criterion of confirmation implicit in the procedures of actual science.

- Furthermore, an advantage of the satisfaction criterion of confirmation is that it applies to observational hypotheses of any logical form - this certainly represents an improvement over - for instance - Nicod's criterion.

- Hempel believes that his definition of confirmation is also appropriate because fulfills all five conditions of adequacy he himself has put forward. But this is instead a shortcoming of the satisfaction criterion of confirmation! For the CONSISTENCY CONDITION, and so the CONSEQUENCE CONDITION and the SPECIAL CONSEQUENCE CONDITION, are counterintuitive.

- Another shortcoming of Hempel's criterion is that it lacks in generality, as it can only apply to observational hypotheses (i.e. hypotheses expressed in observational language), while many interesting and important scientific hypotheses are expressed in a highly theoretical and so non-observational language (e.g. Quantum Mechanics and General Relativity).

- Finally, Hempel's criterion disregards the fact that scientific hypothesis are typically tested not in isolation but in conjunction with many other hypothesis and assumptions (i.e. auxiliary or background information). In this sense, Hempel's criterion is far from being a good approximation to the criterion of confirmation implicit in the procedures of actual science.

Knowledge, Reason and Action
PHIL2606

2nd section
Scientific Methodology

Dr Luca Moretti

Centre for Time

University of Sydney

luca.moretti@arts.usyd.edu.au

www.lucamoretti.org

Qualitative confirmation

Lecture 4
Grimes' formulation of hypothetico-deductivism

Requested reading:

T. Grimes, 'Truth, Content, and the Hypothetico-Deductive Method'.

Hypothetico-deductivism and Grimes' project

- Hypothetico-deductivism is the view that - roughly - hypotheses and theories are confirmed by their logical consequences. (Hempel's criterion of confirmation is not a form of hypothetico-deductivism, as an observation report need not be entailed by a hypothesis to Hempel-confirm that hypothesis).
- There are cases in which an observation statement not entailed by a hypothesis confirms that hypothesis (for instance, 'this raven is black' is not entailed by 'all ravens are black', but it is strongly intuitive that, in normal circumstances, the former confirms the latter).

  But this is not a problem for hypothetico-deductivists, as what they aim at is - typically - specifying sufficient but not necessary conditions for the confirmation of a hypothesis.
- An important advantage of hypothetico-deductivism over Hempel's criterion is that the former applies to both observational hypotheses and highly theoretical hypotheses (as long as they entail testable consequences).
- Hypothetico-deductivism satisfies the CONVERSE CONSEQUENCE CONDITION, which entails a paradoxical consequence. But this is no problem for hypothetico-deductivism, as it satisfies neither the SPECIAL CONSEQUENCE CONDITION nor the ENTAILMENT CONDITION, which are also necessary to obtain the paradox.
- Hypothetico-deductivism has however been argued to be untenable because precise formulations of this methodology prove materially inadequate (i.e. far from real science) and also plagued by novel absurd consequences.
- Grimes' purpose is that of providing a fully acceptable and precise formulation of hypothetico-deductivism.

The basic formulation of hypothetico-deductivism

- According to Grimes, hypothetico-deductivism can basically be expressed by the following criterion:

(HD) Observation statement *E* confirms hypothesis *H* if (1) *E* is verified, (2) *H* entails *E*, and (3) both *H* and *E* are contingent (i.e.neither logically true nor logically false).

Some pragmatic distinction between observation language (in which *E* is expressed) and theoretical language would seem to be presupposed by Grimes.

Why condition (1)? Grimes, in contrast with Hempel, requires *E* to be verified to confirm *H* because, plausibly, he believes that this implicit in the intuitive notion of confirmation. (Not all hypothetico-deductivists accept this requirement).

Why condition (3)? If (3) were omitted, (a) any hypothesis would be confirmed by any observational tautology, which is counterintuitive, and (b) any logically inconsistent hypothesis would be confirmed by any verified observation statement, which is absurd!

Proof of (a). Any statement *H* entails any tautological statement, and so any observational tautology *E* (e.g. ‘today is raining or today is not raining’). Since *E* is a tautology, *E* is necessarily true, and so is verified. Conditions (1) and (2) of (HD) are thus satisfied. If condition (3) were omitted, *E* would confirm *H*.

Proof of (b). Any logically inconsistent statement *H* (e.g. ‘the sun is a star and the sun is not a star’) entails any statement and so any verified observation statement *E*. Conditions (1) and (2) of (HD) are thus satisfied. If condition (3) were omitted, *E* would confirm *H*.

Problems of the basic formulation of hypothetico-deductivism (1)

- (HD) proves unsatisfactory because of, principally, the following three difficulties:

(Duhem's thesis). (HD) assumes that a single hypothesis *H* can entail an observation statement *E*. However, in actual science, single hypotheses rarely entail observation statements. Instead, typically, only when a hypothesis is embedded within a larger network of auxiliary hypotheses and assumptions, the logical conjunction of all of them will entail some testable consequence. (For instances, an astrophysical hypothesis about the composition of stars will entail observational consequences only in conjunction with hypotheses about light's propagation, about our observation instruments - e.g. telescopes - and in conjunction with statements that certify that conditions for the correct functioning of these observation instruments actually take place).

(Tacking by conjunction problem). If observation statement *E* confirms hypothesis *H* on (HD), then, given any statement *X* whatsoever such that the conjunction *H* & *X* is contingent, *E* also confirms *H* & *X* on (HD). (For instance, any observation confirming Newton mechanics also confirms the conjunction of Newton mechanics and the hypothesis you will be abducted by aliens, or that God exists!)

Proof. Suppose *E* confirms *H* on (HD) and consider any statement *X* such that *H* & *X* is contingent. As *E* confirms *H*, (1) *E* is verified. Furthermore, as *E* confirms *H*, *H* entails *E*; thus, (2) *H* & *X* entails *E* too. Finally, since *E* confirms *H*, (3) *E* is contingent, and so is *H* & *X* by assumption. Conditions (1), (2) and (3) for the (HD)-confirmation of *H* & *X* by *E* are satisfied. Therefore, *E* confirms *H* & *X* on (HD). QED.

Problems of the basic formulation of hypothetico-deductivism (2)

(Tacking by disjunction problem). Let *H* be any contingent hypothesis that entrails an observational consequence *E*. And let *Y* be any observation statement such that the disjunction *E* v *Y* is contingent. If *E* v *Y* is verified, *E* v *Y* confirms *H* on (HD).

Proof. Suppose that *H* entails the observational consequence *E*, and that *H* is contingent, and consider an observation statement *Y* such that *E* v *Y* is contingent. As *H* entails *E, H* also entails *E* v *Y.* If *E* v *Y* is verified, condition (1) is satisfied. Moreover, as *H* entails *E* v *Y*, condition (2) is fulfilled. Finally, as both *H* and *E* v *Y* are contingent by assumption, condition (3) is also satisfied. Therefore, *E* v *Y* confirms *H* on (HD). QED.

To appreciate the devastating effect of this problem, consider for instance a (contingent) theory *H* about the composition of the stars that entails an empirical statement *E* we are not able to test (because, say our present technology is not sufficiently powerful or sophisticated). Consider now the verified statement *Y* = 'I am in Sydney'. The disjunction *E* v *Y* is very plausibly contingent. Since *Y* is verified, *E* v *Y* is verified too. And the theory *H* about the composition of the stars is confirmed!

- Both tacking by conjunction and disjunction problems might perhaps be set up by modifying the standard notion of logical entailment (Grimes briefly considers an unsuccessful attempt in this direction).

  In any cases, a solution non-revisionist of our standard logic would surely be preferable!

Grimes on content and confirmation

- Grimes believes that a solution of all such problems will come easier if the intuitive idea underlying hypothetico-deductivism is clarified.

- According to Grimes, hypothetico-deductivism is prima facie grounded in two intuitions:

(a) The verification of a statement *E* confirms a hypothesis *H* insofar as it is an indication that the content of *H* - i.e. what *H* says about the world - may be true.

(b) The verification of a logical consequence of a hypothesis confirms the hypothesis because it is equivalent to the verification of part of the content of that hypothesis (if part of the content of a hypothesis turns out to be true, this is surely an indication that its whole content of that hypothesis may be true).

In conclusion, Grimes believe that the intuitive idea underlying hypothetico-deductivism is the following:

(c) A hypothesis is confirmed if part of its content is shown to be true.

Grimes' solution of the tacking by disjunction problem (1)

- The tacking by disjunction problem is this:

  If *H* is contingent and entrails *E*, and *Y* is a statement such that *E* v *Y* is contingent, If *E* v *Y* is verified, *E* v *Y* confirms *H* on (HD).

- According to Grimes, the tacking by disjunction problem is a serious problem for hypothetico-deductivism because, in many cases, *E* v *Y* is - intuitively - not part of the content of *H*. Consequently, the verification of *E* v *Y* should not confirm *H*.

  (Consider the case *H* is a geological theory and *Y* is the statement 'my name is Luca' As *H* appears to say nothing about me and my name, *E* v *Y* does not seem part of *H*'s content).

- For Grimes, the solution of this problem requires two steps. (1) Giving a general method to determine, among all logical consequences of any given hypotheses, those that identify the hypothesis' content - i.e. its narrow consequences. (2) Re-formulating hypothetico-deductivism such that only narrow consequences can confirm hypotheses.

STEP (1). <u>Grimes' definition of narrow consequence</u>.

Any statement not including quantifiers - and so any observation statement - can be expressed in Boolean normal form.

The formulation of a statement *E* in Boolean normal form is a statement logically equivalent to *E* that makes it explicit the conditions of truth of *E* in terms of a finite disjunction of conjunctions of the atomic statements of *E* or of their negations.

Consider the statement Pa ⊃ Qa. Its atomic statements are Pa and Qa. Its formulation in Boolean normal form is (Pa & Qa) v (Pa & ~Qa) v (~ Pa & ~ Qa).

The formulation of an atomic statement in Boolean normal form is the atomic statement itself.

Grimes' solution of the tacking by disjunction problem (2)

STEP (1). Grimes' definition of narrow consequence.

Given two statements *A* and *B*, *B* is a narrow consequence of *A* if and only If (a) both *A* and *B* are contingent, (b) *B* can be formulated in Boolean normal form and (c) *A* logically entails at least one of the disjuncts included in the formulation of *B* in Boolean normal form.

For instance, Pa v Qa is a narrow consequence of Pa & Qa. For (a) both statements are contingent, (b) Pa v Qa can be formulated in Boolean normal form, precisely: (Pa & Qa) v (Pa & ~Qa) v (~ Pa & Qa), (c) Pa & Qa logically entails the first of the three disjuncts included in the formulation of Pa v Qa in Boolean normal form.

Notice that if *B* is a narrow consequence of *A*, *B* is a logical consequence of *A*. But, if *B* is a logical consequence of *A*, *B* is not necessarily a narrow consequence of *A*.

In particular, it holds true that, though *A* v *Y* is a logical consequence of *A*, *A* v *Y* is not a narrow consequence of *A*.

Thus, if the observation statement *E* is a narrow consequence of the hypothesis *H*, it is in general false that, given the random statement *Y*, *E* v *Y* is also a narrow consequence of *H.*

Because of this specific feature, Grimes believes that the narrow consequences of hypotheses are suitable to express their contents.

STEP (2). Grimes' re-formulation of hypothetico-deductivism.

(HD1) Observation statement *E* confirms hypothesis *H* if (1) *E* is verified, (2) *E* is a narrow consequence of *H*, and (3) both *H* and *E* are contingent.

Grimes' "dissolution" of the tacking by conjunction problem

- It is very easy to show that if *E* is a narrow consequence of *H*, *E* is also a narrow consequence of *H* & *X*, where *X* is any random statement.

  It follows that also (HD1) is affected by the tacking by conjunction problem. That is:

  > If observation statement *E* confirms hypothesis *H* on (HD1), then, given any statement *X* whatsoever such that the conjunction *H* & *X* is contingent, *E* also confirms *H* & *X* on (HD1).

- Grimes does not think this is a serious problem. Grimes emphasizes that if *E* confirms *H* & *X*, from this it does not follow that *E* confirms the random statement *X* (this would really be a devastating consequence!).

  Grimes seems right: hypothetico-deductivism - and in particular (HD1) - does not satisfy the SPECIAL CONSEQUENCE CONDITION (i.e. if a hypothesis is confirmed, all its consequences are confirmed).

- An objection may be that the very fact that *E* confirms *H* & *X,* where X is a random statement, is in itself counterintuitive.

  But Grimes contends that is not true. This is his argument: If *E* is part of the content of the statement *H*, *E* is part of the content of any conjunctive statement of which *H* is a conjunct, and so of *H* & *X.* Hypothetico-deductivism is based on the intuition that a hypothesis is confirmed if part of its content is shown to be true. Thus, it is intuitive for the hypothetico-deductivist that *E* confirms the hypothesis *H* & *X* when *E* is shown to be true.

An unsuccessful attempt to cope with Duhem's problem

- To cope with Duhem's observation that a hypothesis typically entails testable consequences only if embedded within a larger network of auxiliary hypothesis, hypothetico-deductivists have re-formulated their confirmation criterion as follows:

  (HD2) Observation statement *E* confirms hypothesis *H* with respect to auxiliary hypotheses and assumptions *T* if (1) *E* is verified, (2) the conjunction *H* & *T* entails *E*, (3) both *H* & *T* and *E* are contingent, and (4) *T* alone does not entail *E*.

  Clause (4) is necessary to avoid the absurd result that any statement simply consistent with *T* be confirmed by *E.* The request is, instead, that only the hypotheses indispensable for the derivation of *E* through *T* can be confirmed by *E*.

- But (HD2) appears problematic for at least two reasons (emphasized by C. Glymour):

  - when scientists embed a hypothesis *H* within a larger network of auxiliary hypotheses *T* to obtain testable consequences, they often intend to test also (at least some of) the hypotheses deducible from *T*. In other words, they want to know how confirmation distributes over the different hypotheses of the network. (HD1) does not permit this kind of evaluations.

  - (HD2) entails the absurdity that given two randomly chosen, contingent statements *E* and *H*, if *E* is verified, *E* will in most cases confirm *H* with respect to an accepted auxiliary hypothesis.

  (In particular, the accepted auxiliary hypothesis is $H \supset E$, which is true if *E* is true. Glymour has shown that *E*, when verified, will always confirm *H* with respect to $H \supset E$ on (HD2), except when ~*E* entails *H*).

Grimes on the Duhem's problem

- Grimes believes that the his own formulation of hypothetico-deductivism, namely,

  (HD1) Observation statement *E* confirms hypothesis *H* if (1) *E* is verified, (2) *E* is a narrow consequence of *H*, and (3) both *H* and *E* are contingent,

  can also cope with Duhem's problem at least to some extent.

  (i) Grimes first stresses that although many hypotheses entail narrow observational consequences only in conjunction with auxiliary hypotheses, this does not hold true of any hypothesis. Therefore, the version of hypothetico-deductivism expressed by (HD1) can still find application in science. Notice that (HD1) provides sufficient but not necessary conditions for confirmation.

  (ii) Grimes also emphasizes that, if a hypothesis *H* can entail a narrow observational consequence *E* only in conjunction with auxiliary hypotheses *T*, (HD1) can still apply to establish whether the conjunctive theory *H* & *T* is confirmed as a whole by *E*.

  (iii) Finally Grimes argues that the confirmation of observational generalizations can be assessed on (HD1). For, given the generalization $(x)(P(x) \supset Q(x))$, Q(a) is a narrow consequence of the conjunction $P(a)$ & $(x)(P(x) \supset Q(x))$.

- But claims (i) and (iii) are objectionable. Against (i). It is dubious that there are interesting (scientific) hypotheses which individually entail narrow observational consequences. The best candidates would be observational generalizations, but none of them can entail narrow consequences.

  Against (iii). (HD1) allows us to state that the verification of Q(a) confirms $P(a)$ & $(x)(P(x) \supset Q(x))$, while we would like to say that the verification of P(a) & Q(a) confirms $(x)(P(x) \supset Q(x))$.

Concluding considerations about Grimes' work

- Grimes' "dissolution" of the tacking by conjunction problem may appear prima facie plausible.

- Yet Grimes' solution of the tacking by disjunction problem appears less convincing. Grimes is probably too rush in identifying the content of a hypothesis with the class of its narrow consequences. (He does not mention - explicitly - any such identification, but this seems to be what he has really in his mind).

  Two opposite difficulties can emerge: first, more sophisticated versions of the tacking by disjunction problem might still apply. Second, consequences that are, intuitively, part of a theory's content might not be included in the class of its narrow consequences.

- The Duhem thesis seems to constitute a real difficulty for hypothetico-deductivism. Unless a tenable formulation of hypothetico-deductivism that makes confirmation relative to background or auxiliary information is put forward, hypothetico-deductivism is doomed to explicate of only a negligible part and secondary aspects of the procedures of confirmation of real science.

Knowledge, Reason and Action
PHIL2606

2nd section
Scientific Methodology

Dr Luca Moretti
Centre for Time
University of Sydney

luca.moretti@arts.usyd.edu.au
www.lucamoretti.org

Plan of the course

- **Lectures 1&2.** Topic: **Inductive Logic and the Problem of Induction**
  Reading list:
  B. Skyrms, Choice and Chance, ch. 1;
  B. Skyrms, Choice and Chance, ch. 2.
- **Lectures 3&4.** Topic: **Qualitative Confirmation**
  Reading list:
  C. Hempel, 'Studies in the Logic of Confirmation', in his Aspects of Scientific Explanation: and Other Essays in the Philosophy of Science;
  T. Grimes, 'Truth, Content, and the Hypothetico-Deductive Method'. Philosophy of Science **57** (1990).
- **Lectures 5&6.** Topic: **Falsificationism against Inductive Logic.**
  Reading list:
  J. Ladyman, Understanding Philosophy of Science, ch. 3, 'Falsificationism';
  Sections from: I. Lakatos, 'The methodology of scientific research programmes' in I. Lakatos and A. Musgrave (eds.), Criticism and Growth of Knowledge.
- **Lectures 7&8.** Topic: **Quantitative Confirmation: Bayesianism**
  Reading list:
  D. Papineau, 'Confirmation', in A. C. Grayling, ed., Philosophy.
  (Additional material will be provide before the lectures).

Popper and Lakatos against inductivism

- A central assumption of most methodologists is that inductive logic is an essential component of scientific method. Karl Popper (1902-1994) has forcefully denied this assumption.
- Precisely, Popper's claims are that (1) it is impossible to justify rationally any inductive procedure, and that (2) this fact is not catastrophic, as scientific rationality is grounded in no form of inductive logic (according to Popper, scientists simply do not appeal to induction in their methodological decisions!).
- While claim (1) is not original (Popper has however raised a number quite sophisticated objections against quantitative versions of inductive logic), contention (2) is original and surprising.

  Popper has tried to develop a rational reconstruction of scientific methodology - alternative to the traditional, inductivist one - which he himself has named falsificationism.

  According to falsificationism, scientists do not confirm theories by experiments, and they do not prefer more confirmed theories to less confirmed theories. What scientists do is rather this: (very roughly) they try to falsify theories by experiments, and they stick to the theories not yet falsified while rejecting the ones already falsified.
- The way in which Popper's falsificationism copes with Duhem's problem (and related difficulties) is however objectionable as resting on arbitrary decisions of scientists.

  Imre Lakatos (1922-1974) has tried to make up for these difficulties by defining a sophisticated version of falsificationism, which Lakatos has named methodology of research programs.

Falsificationism against inductive logic

Lecture 5
Popper’s falsificationism

Requested reading:

J. Ladyman, Understanding Philosophy of Science, ch. 3, ‘Falsificationism’.

Induction and demarcation

- Popper believes that the two central problems of traditional scientific methodology are:

  (Problem of induction) Explaining how it is possible that experimental results support inductively hypotheses and theories.

  (Problem of demarcation) Determining what logical and methodological features demarcate science from pseudo-science (or simply from other disciplines).

- Popper dismisses the problem of induction by arguing that induction is impossible and that theories are not supported inductively - namely, no observation or experimental result can make any theory more probable or plausible.

  Popper rehabilitates Hume's classical argument and formulates precise objections against probabilistic theories of confirmation.

- Popper "resolves" problem of demarcation by arguing that what essentially characterizes scientific statements (i.e. hypotheses and theories) from a logical point of view is the possibility of falsifying them (rather than confirming them).

  Popper also argues that the method of science essentially coincides with a set of logical techniques for falsifying hypotheses and theories. Scientific method does not coincides with inductivism but with falsificationism.

Induction, the context of justification and the context of discovery

- Popper contends that much of the alleged relevance of inductive logic for scientific methodology depends on two mistakes:

  (1) Believing (as Newton, Bacon and Mill did) that scientific theories are invented by means of inductive inferences (e.g. by induction by generalization).

  (2) Believing that the process of inventing theories is part of scientific methodology.

  According to Popper, the truth is instead that:

  (1*) Most hypotheses and theories simply cannot be invented by inductive inferences. For, principally, we cannot inductively infer any description of theoretical entities by observations.

  (2*) Scientific methodology does not investigate the process by which scientists invent hypotheses - Popper does not believe that any method or logic of invention can actually be formulated.

  Scientific methodology is only concerned with how to justify (e.g. accepting or rejecting) theories. The way in which a theory is invented is irrelevant for the question of whether a theory is justified or not justified.

- For Popper, science can be investigated from the perspective of the context of justification and from the perspective of the context of discovery (or invention).

  Scientific methodology belongs to the context of justification.

  History, sociology and psychology of science may investigate the processes by which scientific theories have been invented; in such a case, these disciplines belong to the context of discovery.

Logical asymmetry between verification and falsification

- Popper emphasizes that scientific laws (represented by universal generalizations) cannot be verified by observation statements though they can be falsified by them.

  The same is true - more generally - of scientific hypotheses and theories: while they are not verifiable by their observational consequences, they can be falsified by the latter.

  - Suppose (*L*) $(x)(P(x) \supset Q(x))$ is a scientific law. As scientific laws typically range over a potentially infinite number of individuals, no finite conjunction of statements representing instances of *L* - like P(a) & Q(a) & P(b) & Q(b) - is logically equivalent to *L*. Therefore, the verification of no finite conjunction of statements representing instances of *L* can verify *L*.

  Yet the verification of this statement alone, P(a) & ~Q(a), falsifies *L*.

  - Suppose *T* is a scientific theory. As scientific theories have typically a potentially infinite class of observational consequences and typically make assertions about unobservable entities, no finite conjunction of observational consequences of *T* will be logically equivalent to *T*. Therefore, the verification of no finite conjunction of observational consequences of *T* can verify *T*.

  Yet, if *T* entails the observational consequence *E*, the mere verification of ~*E* produces - by modus tollens - the falsification of *T*.

- Since theories and generalizations cannot be verified by evidence, the temptation is to argue that they can at least be confirmed by evidence. But this raises the terrible problem of justifying induction.

- Popper contends that if scientific methodology focuses on falsification, rather than verification or confirmation, no problematic use of inductive logic is made necessary. For the procedures by which we falsify scientific laws and theories depend on only deductive logic.

When a hypothesis is scientific

- According to Popper, a hypothesis - and, more generally, a statement - is scientific if can in principle be falsified by some specific observation statements (but has not been falsified by it), where 'specific' means that the contents of the relevant observation statements must be described with precision and in advance.

  This criterion immediately excludes from the class of the scientific hypotheses, for instance:

  - metaphysical statements (e.g. 'God is omniscient').

  - tautologies and analytical statements (e.g. 'All females are not males').

  - vague and unclear statements (e.g. 'John Smith is almost bald', horoscopes).

- Popper calls the observation statements that can in principle falsify a hypothesis the potential falsifiers of that hypothesis.

- Popper believes that this simple criterion is strong enough to show that Freudian psychoanalysis is not a scientific theory, for its content is too vague to entail any definite class of potential falsifiers.

  On the other hand, Popper believes that most physical and chemical theories do satisfy his criterion, as they entail a very well defined class of potential falsifiers.

- Popper does not believe that non-scientific hypotheses are meaningless or useless for science. For example, in the history of science, scientists were often inspired by their religious or metaphysical belief to make scientific conjectures. (Think of non-falsifiable beliefs like 'nature is simple' or 'God made nature simple').

Falsificationist methodology (1)

- For Popper, science proceeds by developing and testing hypotheses with the specific purpose of trying to falsify them.

  If a hypothesis has been falsified, it must be abandoned. If a hypothesis has not been falsified, scientists should provisionally accept it, though this does not mean that the hypothesis has become more probable or plausible.

  That hypothesis must then be subjected to further and more stringent tests and ingenious attempts to falsify it.

  Any hypothesis that has survived some empirical test is said corroborated (not confirmed!) by evidence.

- Popper seems to believe that by eliminating all false hypotheses one after the other we come closer and closer to the true hypotheses.

- When a hypothesis *H* has been eliminated, it should be replaced by a new one *H** that includes all potential falsifiers of *H* (but not the actual falsifier!) and new potential falsifiers. In other words, the empirical content of *H** must be wider than that of *H*.

  This requirement prevents *H* from being simply "adjusted" with some ad hoc change.

  Suppose, for instance, that *H* is: 'The mass of any material body does not change with the body's speed'.

  H is falsified by the observation that the mass of particle *p* has increased with *p*'s speed. It would be ad hoc replacing *H* with *H**, which says: 'The mass of any material body, with the sole exception of particle *p,* does not change with the body's speed'.

  To be acceptable, *H** ought to allow for precise predictions of novel facts.

Falsificationist methodology (2)

- The falsificationist requirement that if H^* supersedes H, H^* must predict new facts can be made more precise:

  (1) the new facts should be facts of a new type (with respect to the facts predicted by H),

  (2) they should be improbable in the light of background knowledge.

  For instance: General Relativity superseded Special Relativity. A fact of a new type predicted by General Relativity is that gravity can influence the geometrical features of light rays - this prediction appeared very improbable (even weird!) in the light of background knowledge.

- According to Popper, though Marxism is a scientific theory (as it allows for precise predictions), Marxists are not scientists. For they have adjusted several times Marx's theory with ad hoc changes to elude falsifications. The adjusted theories gave no new prediction about improbable facts.

- Popper's falsificationism makes Bacon's conception of crucial experiments more sophisticated.

  According to Bacon, if T and T^* are two rival theories, scientists can choose between them by testing two incompatible predictions of, respectively, T and T^*. Suppose T predicts that E and T^* predicts that $\sim E$. If E is verified, scientists have to accept T and reject T^*. While, if $\sim E$ is verified, scientists have to accept T^* and reject T.

  According to Popper, the mere fact that E is verified is sufficient for the rejection of T^*, but not for the acceptance of T. T can be accepted if its empirical content is wider than T^* empirical content.

Problems of falsificationism (1)

(i) Existential statements are not falsifiable

A scientific generalization with form $(x)(P(x) \supset Q(x))$ typically ranges over an infinite number of individuals, so it cannot verified by any finite conjunction of particular statements about those individuals.

Analogously, it can be argued that a scientific existential hypothesis with form $\exists x(P(x))$ ranges over an infinite number of individuals; thus, it cannot be falsified by any finite conjunction of particular statements about those individuals.

A falsifier of $\exists x(P(x))$ would be an infinite conjunction of statements such that each of them asserts that one given individual of the infinite domain does not have the property P.

Therefore, if a scientific hypotheses has the logical form $\exists x(P(x))$, it cannot be considered scientific by falsificationism!

But this does not seem a real problem for falsificationism. For scientific existential hypotheses typically are circumstantiated claims.

Cosmological theories, for instance, do not state, simply, that there are black holes - they say where they are supposed to be or in which conditions we are supposed to expect that there is a black hole. In the same way, biochemical theories do not just state that the DNA exists - they say where (in which part of the cell) the DNA is.

Circumstantiated existential claims are, plausibly, falsifiable.

(ii) Probabilistic hypotheses are not falsifiable

This is a more serious problem. Some scientific hypotheses consist in claims about the probability of single events. For example H = 'the probability that these atoms of uranium 235 will decay in N years is M'. Suppose that M is less that 1, and that after N years no atom has decayed. Does this evidence falsify H? No, because M is less than 1. H is compatible with the possibility that after N years no atom has decayed. H is not scientific for falsificationism!

Problems of falsificationism (2)

(iii) Some general scientific principles are not falsifiable

Consider the Principle of Conservation of Energy, which asserts that energy can take different forms but it cannot be created or destroyed.

As a matter of logic, this principle cannot be falsified. For it does not describe all forms of energy there are in the universe. So, any apparent counterexample can be neutralized with the assumption that energy has turned into a new form still unknown to us.

(As a matter of fact, in the history of science, every time an experiment resulted in an apparent counterexample to this principle, to avoid its falsification, scientists rejected other parts of physics - even the very experimental results - or they posited new forms of energy).

Another unfalsifiable principle accepted by Descartes and Hertz is the one asserting that any phenomenon can be explained in terms of "mechanisms", that is, by forces that act by contact.

Principles like these are said by Lakatos to be syntactically unfalsifiable.

These principles are surely scientific principles, but they are not so according to falsificationism, because scientists cannot specify in advance their potential falsifiers.

(iv) Sometimes scientists ignore falsifications

Any important scientific theory has anomalies - that is experimental results at odds with the theory but "tolerated" by the scientists. (For example, the anomalous perihelion of Mercury incompatible with Newton's mechanics was known for ages and tolerated).

(Popper has tried to justify such a relaxed attitude of scientists by arguing that it is rational - or reasonable - to stick a partially successful theory in front of recalcitrant evidence until a better theory become available).

Problems of falsificationism (3)

(v) Scientists do have expectations about the future

It can hardly be denied that scientists believe in hypotheses and theories with different degrees of strength and that they do make evaluations about the plausibility of hypotheses on the grounds of empirical evidence. In other words, scientists make use of inductive procedures. Consequently, falsificationism does not seem to provide a truthful reconstruction of scientific methodology.

But the problem is even worse, because falsificationism itself presupposes some notion of inductive support. Remember that Popper requires that if a theory H^* supersedes another theory H, H^* must predict novel facts which are improbable in the light of background knowledge!

(vi) Falsifications depend in part on stipulations

Popper’s falsificationism embeds the epistemological doctrine called fallibilism. That is, the view that any accepted empirical statement may prove false (this is trivially true of all hypotheses).

Popper believes that any accepted observation statement can virtually be withdrawn if some of the hypotheses or assumptions involved in the process of acceptance are rejected.

Popper however believes that if the scientific community agrees that certain observational statements are to be accepted, these statements can be used to reject once and for all any theory incompatible with them.

This involves that in most (if not all) falsifications there is a risky, conventional element.

Problems of falsificationism (4)

(vii) Duhem's thesis

The Duhem thesis is a problem for falsificationism as prevents the exact localization of falsifications.

Suppose *H* is Newton's gravitational law. Given *H*, Newton predicted that Halley's comet had to return in 1758; let *E* be this prediction. Suppose *E* says that the comet will be visible in a specified area of the sky in a specified time of 1758. *E* was verified.

Suppose *E* is falsified. According to a crude version of falsificationism, the falsification of *E* entails the falsification of *H* via modus tollens. That is:

H entails *E*. *E* is false. Therefore, *H* is false.

But this view is too simplistic. For it is not true that *H* entails *E!*

To obtain *E*, we need to specify the initial conditions of the system - for instance: the mass of the comet and of the other bodies in the solar systems, their relative positions and velocities, etc. Let *S* be the statement specifying these values.

We should also assume that there are no disturbing factors (e.g. unknown planets or asteroids). Let *C* be the statement expressing these ceteris paribus conditions.

To obtain *E*, we also need to use other laws of Newton's dynamics. Let *L* be the conjunction of these auxiliary laws.

To derive that the comet will be visible, we need to apply the laws of optics relative to the lenses in the telescope we will be using for the observation. Let *O* express these laws.

Finally, we need to specify that, at the time of the observation, the lenses work properly (for instance, their shape does not change). Let call *P* the statement specifying it.

It is the conjunction *S*&*C*&*L*&*O*&*P*&*H* that entails *E.* If *E* is false, by modus tollens, we obtain:

S&*C*&*L*&*O*&*P*&*H* entails *E*. *E* is false. Therefore, *S*&C&*L*&*O*&*P*&*H* is false.

If *S*&*C*&*L*&*O*&*P*&*H* is false, at lest one of its conjuncts must be false. This does not mean that just *H* is false! *H* may well be true and, for instance, *L* or *O* may be false.

Concluding considerations on Popper's falsificationism

- The crucial problem of Popper's falsificationism is that it does not offer any realistic picture (though idealized) of scientific methodology. Scientists do use inductive procedures, they do accept unfalsifiable scientific principles, they even tolerate - in certain cases - falsified theories!

- Furthermore, if falsificationism were conceived of as a purely normative doctrine (what scientists should do, no matter of what they actually do), it would prove internally incoherent. As while the falsificationist claims that inductive reasoning is illicit, she makes use of induction to assess the acceptability of theories (new theories have to entail new improbable predictions).

- Quite ironically, the relevance of Popper's falsificationism rests more on the problems it uncovers than on the problems it resolves.

  It should be clear at this point that the Duhem thesis raises serious difficulties to any kind of scientific methodology in general, as it plagues both confirmation procedures (at least, when confirmation is qualitative) and falsification procedures.

  Another important point is the following: Popper implicitly admits that no (interesting) scientific hypothesis can definitively be disproved by observations, for observation statements are themselves fallible.

  This raises the suspicion that, in concrete science, most theories can be "resuscitated" after being "falsified" by experiments. But if no interesting theory never "dies" for real, science runs the risk of becoming a chaotic activity in which - as P. Feyerabend once said - 'everything goes'!

Knowledge, Reason and Action
PHIL2606

2nd section
Scientific Methodology

Dr Luca Moretti
Centre for Time
University of Sydney

luca.moretti@arts.usyd.edu.au
www.lucamoretti.org

Lecture 6
Lakatos’ methodology of scientific research programmes

Requested reading:

I. Lakatos, ‘Methodology of Scientific Research Programmes’

Lakatos' project

- Lakatos principally aims to provide a sophisticated formulation of falsificationism capable to lessen the conventionalist elements of Popper's falsificationism. These elements are - according to Lakatos - functional to settle problems generated by both fallibilism and the Duhem thesis.

  (Notice that Lakatos provides his own reading of Popper's falsificationism, which is often more complex and perhaps more incoherent than in Lakatos' picture).

- To allow for the definitive elimination of theories, Popper's falsificationism requires the introduction of methodological decisions that are too risky and arbitrary. For, to eliminate a theory on the grounds of incompatible evidence, scientists are asked to stipulate that auxiliary assumptions are true, while they may just be false. As a result, true or approximately true theories may be irrevocably rejected!

  To avoid methodological anarchism ('everything goes'), Popper's falsificationism appears to turn science into a largely arbitrary and subjective enterprise.

- Lakatos believes that falsificationism can re-formulated to the effect that the rejection of scientific theories does not rest on arbitrary stipulations but is instead empirically grounded and, to a large extent, rational and objective.

  This requires focusing not on theories (conceived of as static sets of statements) but, rather, on research programs (conceived of as dynamic sequences or progressions of theories).

Dogmatic falsificationism

- Lakatos believes that Popper's falsification - which Lakatos calls methodological falsificationism - is an improvement of a cruder position called dogmatic falsificationism.

- Dogmatic falsificationism requires that scientific theories be essentially fallible - and so falsifiable - but it retains the infallibility of the empirical basis. In other words, the claim is that, (1) while scientific hypothesis are falsifiable, observation statements are not. Observation statements are, once warrantedly asserted, incorrigibly true.

  Another conviction of the dogmatic falsificationist is that (2) single theories entail observation statements.

  From (1) and (2), the dogmatic falsificationist concludes that (3) (single) theories can conclusively be disproved (i.e. proved to be false) by observations.

- Lakatos emphasizes that (2) is incorrect because of Duhem's thesis.

  But also (1), according to Lakatos, is incorrect.

  First, in science, there is no sharp borderline between "theoretical" and "observation" statements. And it is psychologically proven that most of - if not all - our judgments of perception do presuppose some interpretation of what we see. Thus, if theories are fallible, plausibly, most observation statements are fallible too (except perhaps observation statements like: 'this is this' and 'that is that').

  Second, for the dogmatic falsificationist, observation statements should be justified by observable facts - i.e. by extra-linguistic entities. But this is a category mistake. For only a statement can justify another statement. Facts can perhaps cause the utterance of a statement, but they cannot justify it.

- In conclusion, according to Lakatos, dogmatic falsificationism is untenable; and it is simply incorrect to affirm that theories can conclusively be disproved by observations.

Methodological falsificationism and conventionalism

- For Lakatos, methodological falsificationism is nothing but a variant of the general perspective that in philosophy of science is called conventionalism.

  Very roughly, conventionalism says that hypotheses and theories that scientists find useful for different reasons are accepted by a conventional decision. When a theory is accepted, scientists will prevent its falsification by means of “logical stratagems” - i.e. mainly by ad hoc changes in the auxiliary assumptions employed to derive observation statements. Prominent conventionalists were Poincaré, Le Roy and Duhem himself.

- A problem of conventionalism is that it makes the correlation between accepted theories and true theories very thin, if not completely inexistent. (In what sense the accepted theories are, or may be, closer to the truth than the ones we have rejected?)

  Furthermore, on this view, it becomes unclear why empirical sciences are really empirical!

- Methodological falsificationism accounts for the empirical character of empirical sciences by imposing that empirical theories are to be rejected whenever in conflict with observation statements. Ad hoc changes in the auxiliary assumptions are forbidden or drastically limited.

- Methodological falsificationism is however a form of conventionalism because it is based on the conviction that observation statements - and not theories - are generally accepted or rejected by conventional agreement.

The four decisions that qualify methodological falsificationism (1)

- According to the methodological falsificationist, scientists can reject theories on the grounds of incompatible “observations” only after taking certain methodological decisions. Precisely, there are four types of decisions.

  Once taken, these decisions are irrevocable.

  (1) Scientists have to decide to select certain statements as “observation statements”. These statements must be singular and there must exist, at the time of the selection, some technique such that anyone who has learned it will be able to decide whether those statements are assertible or not.

  Decisions of this first kind involve assuming that the relevant technique is reliable - in particular, that the (interpretative) theories about the relevant observation instruments are correct.

  Scientists will assume this if repeated tests of the interpretative theories have given positive result.

  (2) By applying the appropriate technique, scientists have then to decide whether some of the selected “observation statements” are actually “true”.

  Decisions of this second type involve assuming that the conditions for the correct functioning of the observation instruments are satisfied. Scientists will assume this after repeatedly checking these conditions. Moreover, scientists may repeat the relevant observations and experiments several times.

  Decisions (1) and (2) allow scientists to establish the empirical basis that may “falsify” theories and hypotheses.

The four decisions that qualify methodological falsificationism (2)

(3) To make statistical hypotheses "falsifiable", scientists may decide to accept some conventional rejection rules that apply to them.

Statistical hypotheses make claims not about the probability of single events but about the probability of relative frequencies. A statistical law may have the following form:

(SL) The relative frequency (or proportion) of *A*-individuals that are *B*-individuals is *N*.

Suppose we want to test a law with the form of (SL). We examine a large class of *A*-individuals and we find out that the proportion of *A*-individuals that are *B*-individuals is different from *N*.

Does this falsify our law? No, because *N* is to be intended as the value of the limiting relative frequency. That this, the value to which the ratio $^{B}/_{A}$ tends when *A* tends to infinity and where *A* is the number of *A*-individuals and *B* is the number of *B* individuals. But any examined class of *A*-individuals will necessarily be finite!

(From a mathematical point of view, the fact that for any finite *A*, $^{B}/_{A}$ is different from *N* is compatible with the possibility that, if *A* tends to infinity, $^{B}/_{A}$ tends to *N*).

Scientists may however decide to consider the laws having the form of (SL) as "falsified" if - roughly - after examining a conventionally fixed number of *A*-individuals, the relative frequency of *A*-individuals that are also *B*-individuals is not yet *N*.

The four decisions that qualify methodological falsificationism (3)

(4) A set of scientific hypotheses $H_1, \ldots, H_n$ typically entails "observation statements" only in conjunction with statements specifying the initial conditions of the system and ceteris paribus clauses asserting that there are no other influencing (or disturbing) factors.

To locate accurately falsifications, before testing $H_1, \ldots, H_n$, scientists will decide that the statements specifying the initial conditions and the ceteris paribus clauses are "true". Consequently, any logical inconsistency with accepted "observation statement" will result - via modus tollens - in the "falsification" of the conjunction of $H_1, \ldots, H_n$.

Scientists will decide that the statements specifying the initial conditions are "true" only if repeated tests of them have given positive result.

Scientists will also assume initially that there are other influencing factors and they will test all these assumptions. If many of these hypotheses prove incorrect, scientists will decide that the ceteris paribus clauses are "true".

(Notice that, in Lakatos' picture, the interpretative theories and the assumptions about our observation instruments are not used to derive "observation statements" from theories. For "observation statements" are expressed in an objective form - i.e. 'there is a comet in the sky' rather than 'a comet is visible in the sky'. Interpretative theories and assumptions are however necessary, for Lakatos, to describe our "observations" in an objective form - i.e. to infer 'there is a comet in the sky' from 'a comet is visible in the sky'.)

- Although this is not part of methodological falsificationism, Lakatos suggests a fifth type of decision that seems coherent with Popper's view: scientists may decide that syntactically unfalsifiable principles can be "falsified" if are incompatible with a well-corroborated theory.

  (For instance, the principle that any phenomenon is explainable by forces that act by contact is "falsified" by a corroborated theory that assumes that forces act at a distance).

Problems of methodological falsificationism

- According to Lakatos, methodological falsificationism is affected by two general problems:

  (1) It asks scientists to take irrevocable decisions that may prove completely misleading.

  The methodological falsificationist wants to secure the possibility of rejecting some theories to allow for the possibility of progress (in Popper's sense), but the risk is that of rejecting the wrong theories!

  (2) Methodological falsificationism is not materially adequate. For scientists are very far from complying with its methodological rules. In particular:

  (2a) The history of science shows that most methodological decisions of the kinds identified by the methodological falsificationist have been re-discussed and revoked several times. Real scientific rationality is more complicated and much slower than in the picture of the methodological falsificationist.

  (2b) The methodological falsificationist thinks of empirical tests as, principally, a way to eliminate ("falsify") theories. But the history of science shows that they were often conceived of as means to corroborate theories. Despite Popper's opinion, corroboration seems to have a very central role in science.

  (2c) The methodological falsificationist describes scientific tests as two-cornered fights between one theory and experiments, while history of science suggests that tests typically are at least three-cornered fights between two (or more) rival theories and experiments.

- Lakatos' sophisticated version of methodological falsificationism is meant to make up for all these difficulties.

Sophisticated methodological falsificationism (1)

- A more plausible version of falsificationism should, to begin with, make the rules of theory acceptance and rejection more sophisticated. Lakatos proposes the following rules (which are indeed already considered in at least some of Popper's papers):

  (Theory acceptance) A theory is acceptable if and only if (1) has excess of empirical content (i.e. predicts new and improbable observable facts) over its predecessor and (2) some of this excess content is verified.

  (Theory rejection) A theory *T* is to be rejected if and only if another theory *T** has been proposed such that: (a) *T** has excess empirical content over *T*, (b) all unrefuted content of *T* is included in *T**, (c) some of the excess content of *T** is verified.

- These rules explain why corroboration is so important in science: if *T** is not corroborated by the verification of some of its new predictions, scientists cannot accept *T** and cannot reject its predecessor *T*.

  Notice that, on the above rejection rule, the "falsification" of *T* is neither sufficient nor necessary for the rejection of *T*.

  The "falsification" of *T* is not sufficient for its rejection because *T* cannot be rejected on the grounds of incompatible "observation statements" if no better theory *T** is also available (this might explain why scientific rationality is slow and why scientists often tolerate "falsifications").

  The "falsification" of *T* is not necessary for its rejection because *T* can be rejected if *T* is compatible with all "observation statements" but there exist a new theory *T** with additional and verified empirical content.

- These rules already cast some light on why, in concrete science, tests are typically at least three-cornered fights between two or more rival theories and experiments.

Sophisticated methodological falsificationism (2)

- To lessen the conventionalist component of falsificationism and to explain how methodological decisions are continuously re-discussed by scientists, the rules of acceptance and rejection should be applied to theories conceived of as including all auxiliary assumptions necessary to entail "observational consequences" (i.e. ceteris paribus conditions and initial conditions).

- Consider a series of theories T_1, T_2, T_3, … of this type, and suppose that each theory results from changing only the auxiliary assumptions of the previous theory in order to accommodate its empirical anomalies.

  A series of this type is said theoretically progressive if each theory in the series predicts some new and improbable fact not predicted by its predecessor.

  A theoretically progressive series is also empirically progressive if, for each theory in the series, some of the new predictions are verified.

  A series of theories which is both theoretically and empirically progressive is said progressive (otherwise is said degenerating).

  Notice that if a series of this type is progressive, it has been constructed in perfect accordance with the rules of theory acceptance and rejection formulated by Lakatos.

- Lakatos suggests that, to cope with "falsifications", scientists should construct a series of theories of this kind in which the auxiliary assumptions are continuously modified with the purpose of making the series progressive.

  Modifications are tolerable as non-ad hoc if they make the series at least theoretically progressive. Any series of theories which is not at least theoretically progressive should be considered pseudoscientific.

How sophisticated methodological falsificationism reduces the conventionalist component

- What is the effect of the procedure described by Lakatos?

(a) The decisions of the fourth kind (i.e. deciding that statements specifying initial conditions and ceteris paribus clauses are “true”) become completely useless.

(b) The decisions of the third kind (i.e. deciding rules for falsifying statistical hypotheses) are still necessary.

(c) The decisions of the first kind (i.e. deciding to select statements of certain types as “observation statements”) are still necessary. But these decisions are not particularly problematic!

(d) The decisions of the second type (i.e. deciding that certain “observation statements” are “true”) are still necessary.

Lakatos suggests that the arbitrariness of these decisions can be reduced by introducing “appeal procedures” (which are customary in real science) in which interpretative theories and assumptions that justify the problematic “observation statements” are tentatively replaced with alternatives that do not justify the same “observation statements”.

But this can only postpone the decisions of the second type, which are still necessary!

The methodology of research programs (1)

- Focusing on series of theories produces a better understanding of the situation in which rival theories based on incompatible principles compete with one another (e.g. Cartesian mechanics versus Newtonian mechanics, or Lorentz's electrodynamics versus Einstein's special theory of relativity).
- Consider two rival theories *A* and *B* of this type. Suppose *A* entails the "observation statement" *E*, and *B* entails the "observation statement" ~*E*. A crucial experiment is performed with the effect that *E* is accepted. *A* is corroborated while *B* is "falsified". The advocates of *B* will then replace *B* with B_1, in which some of the original auxiliary assumption are modified to cope with the "falsification". New "observation statements" are then derived from *A* and B_1. And now it is *A* that is "falsified" and replaced with A_1.

  This process will quickly lead to two rival series of theories A, A_1, A_2 , … and B, B_1, B_2 , … such each *A*-theory shares the central principles of the other *A*-theories, and each *B*-theory shares the central principles of the other *B*-theories.
- Lakatos calls series of theories of this kind research programs. Any research program complies with Lakatos' acceptance and rejection rules and it is additionally characterized by:

  (1) A negative heuristic (what scientists should not do), which specifies the central principles constituting the hard core of the program (e.g. the laws of Newtonian mechanics) and prevents scientists from directing modus tollens against them.

  (2) A positive heuristic (what scientists should do), which specifies how the central principles should be applied to concrete objects (i.e.how to construct suitable models of real objects). This typically comes in terms of a series of hints and suggestions concerning how to change and develop the protective belt of auxiliary assumptions (ceteris paribus clauses and initial conditions) that surround the hard core of the program to accommodate "falsifications".

The methodology of research programs (2)

- Finally, any research program complies with the rules for acceptance and rejection of research programs.
- A research program can be accepted if and only if the following two conditions are satisfied:

  (1) it is theoretically progressive;

  (2) it is at least intermittently empirically progressive.

  (Every new theory in the series of the research program must make new improbable predictions, and for at least some theories of the series some of the new predictions must be "verified").

  Any research program which does not satisfy both (1) and (2) is to be rejected.
- Conditions (2) shows, again, that corroboration has a central role in science.
- Notice also that since the hard-core of a research program typically includes syntactically unfalsifiable principles, the rule of rejection of research programs is a rule for the rejection of syntactically unfalsifiable principles!
- Furthermore, the existence of several historical cases of two or more research programs in competition suggests, again, that tests typically are at least three-cornered fights between two (or more) rival research programs and experiments.
- Another important result of the methodology of research programs is the explanation of why scientific rationality is, in certain cases, very slow.

  Let T_n be a theory of a research program RP that entails E. Suppose $\sim E$ is "verified". Advocates of RP will develop a series of successive variants T_{n+1}, T_{n+2}, T_{n+3}, … of T_n to cope with the recalcitrant evidence. Suppose however that, in the long run, none of the variants T_{n+1}, T_{n+2}, T_{n+3}, … proves empirically progressive. MP will eventually be rejected.

  Only retrospectively and after some time $\sim E$ will be perceived as refuting evidence for MP.

Concluding considerations on Lakatos' methodology (1)

- Lakatos' sophisticated version of falsificationism actually lessens the risky and arbitrary components of Popper's falsificationism.

  Furthermore, Lakatos' innovative approach, which focus on research programs, seems to offer an illuminating insight into the real methodological practices of scientists, and it explicates rationally some *prima facie* puzzling features of these practices (for instance, why scientists are sometimes slow in their decisions and why they seemingly tolerate "falsifications").

- Sophisticated methodological falsificationism is however plagued by some of the drawbacks that also affect Popper's conception of falsificationism.

  In particular, it can hardly be denied that scientists make use of inductive procedures, as they typically claim that experimental results make theories and hypotheses more or less probable.

  Lakatos has tried to explain the fact that theories appears in some cases positively supported by evidence by attributing a central methodological role to corroboration. The problem is that corroboration is not confirmation, as corroborated theories cannot be considered more plausible or probable than uncorroborated theories.

  Finally, Lakatos' falsificationism appears as internally incoherent as Popper's is. For, though Lakatos considers inductive reasoning illicit, Lakatos also claims that if a research program is acceptable, each new theory of the program has to entail new predictions which are improbable on background knowledge.

Concluding considerations on Lakatos' methodology (2)

- Sophisticated methodological falsificationism is also affected by a novel difficulty (exposed, for instance, in some of Feyerabend's papers).

  Lakatos thinks that science coincides with the activity of producing scientifically acceptable research programs. Where a research program is scientifically acceptable if and only if (1) it is theoretically progressive and (2) it is at least intermittently empirically progressive.

  Suppose so that the research program RP has been degenerating (i.e. not empirically progressive) for quite long time - namely, none of the successive theories of RP invented within that time have produced any "verified" new prediction (all new predictions have always proven "false").

  Scientists might then reject RP.

  The problem is that (2) does not specify for how long a scientist should wait before rejecting RP. If (2) is made more precise - e.g. 'the research program has to be (again) empirically progressive after at least 10 years' - this condition can legitimately be criticized as arbitrary (why just 10 and not 11 or 50 years?). Considering the level of conceptual and technological complexity of many sciences, it may be very hard to decide in advance the exact length of this period of time.

  On the other hand, if (2) is left as it is, this condition can legitimately be criticized as vacuous (no one could ever be accused of being irrational for obstinately sticking to a research program that has been degenerating for centuries!).

  But if no research program can be eliminated, isn't science an "anarchical" activity in which, after all, 'everything goes'? This seems weird. And how could we defend the quite plausible thesis that there has been scientific progress at least in the sense of eliminating false theories (for instance some hard cores of scientific research programs)?

Knowledge, Reason and Action
PHIL2606

2nd section
Scientific Methodology

Dr Luca Moretti
Centre for Time
University of Sydney

luca.moretti@arts.usyd.edu.au
www.lucamoretti.org

Plan of the course

- **Lectures 1&2.** Topic: **Inductive Logic and the Problem of Induction**
  Reading list:
  B. Skyrms, Choice and Chance, ch. 1;
  B. Skyrms, Choice and Chance, ch. 2.

- **Lectures 3&4.** Topic: **Qualitative Confirmation**
  Reading list:
  C. Hempel, 'Studies in the Logic of Confirmation', in his Aspects of Scientific Explanation: and Other Essays in the Philosophy of Science;
  T. Grimes, 'Truth, Content, and the Hypothetico-Deductive Method'. Philosophy of Science **57** (1990).

- **Lectures 5&6.** Topic: **Falsificationism against Inductive Logic.**
  Reading list:
  J. Ladyman, Understanding Philosophy of Science, ch. 3, 'Falsificationism';
  Sections from: I. Lakatos, 'The methodology of scientific research programmes' in I. Lakatos and A. Musgrave (eds.), Criticism and Growth of Knowledge.

- **Lectures 7&8.** Topic: **Quantitative Confirmation: Bayesianism**
  Reading list:
  D. Papineau, 'Confirmation', in A. C. Grayling, ed., Philosophy.
  (Additional material will be provide before the lectures).

Quantitative confirmation

Lectures 7 & 8
Bayesian confirmation theory

Requested reading:

D. Papineau, 'Confirmation'.

Additional readings:

W. C. Salmon (et al.). 1999. Introduction to philosophy of science. Hackett Publishing Company, pp. 66-103.

W. Talbott, 'Bayesian Epistemology'.

(http://plato.stanford.edu/entries/epistemology-bayesian/index.html).

Bayesian confirmation theory

- Bayesian confirmation theory, which can be seen as a quantitative form of inductive logic, is today extremely popular.
- This theory explicates the notion of confirmation by appealing to the laws of the probability calculus and, in particular, to the Bayes theorem, shown by Reverend Thomas Bayes (1701-1761). Bayesian confirmation theory did not emerge until the first formal axiomatizations of probability theory was given in the first half of the last century.
- The Bayesian interprets the notion of probability in a subjective and, typically, personal (or personalist) sense. Precisely, as indicating the degree of strength to which a subject believes in some state of affairs or proposition. This interpretation of probability was first proposed by Frank P. Ramsey (1903-1930) and Bruno DeFinetti (1906-1985).
- To make this interpretation plausible, the Bayesian appeals to a type of argument - called Dutch Book argument - to the effect that our degrees of beliefs have to comply with the axioms of the probability calculus.
- Bayesian confirmation theory is both extremely general and very precise and would seem to provide exceptionally accurate models of confirmation practices of actual science. Furthermore, Bayesian confirmation theory is not plagued by some of the difficulties that afflict other confirmation theories.
- Some drawbacks of Bayesian confirmation theory stem however from the subjective interpretation of probability.

The axioms of the probability calculus

- Consider a standard language for propositional logic. Any function P(…) that assigns each statement (atomic and complex) of this language a number is a function (or distribution) of probability if and only if satisfies the axioms of the probability calculus. These axioms can be formulated as follows:

  (A1) For every statement A, $0 \leq \mathrm{P}(A) \leq 1$.

  (A2) If A is necessarily true, $\mathrm{P}(A) = 1$.

  (A3) If A is necessarily false, $\mathrm{P}(A) = 0$.

  (A4) If A and B are mutually exclusive, $\mathrm{P}(A \vee B) = \mathrm{P}(A) + \mathrm{P}(B)$.

  (Notice that (A2) and (A4) do not forbid $\mathrm{P}(A) = 1$ and $\mathrm{P}(A) = 0$ when A is neither necessarily true nor necessarily false).

- We can additionally define the conditional probability of B given A (that is, the probability that B is true on the assumption that A is true), expressed by $\mathrm{P}(B|A)$, as follows:

  (A5) $\mathrm{P}(B|A) = \dfrac{\mathrm{P}(B \,\&\, A)}{\mathrm{P}(A)}$, with $\mathrm{P}(A) \neq 0$

- From (A1)-(A5) and the rules of propositional logic, the following theorems easily follow:

  (1) For every statement A, $\mathrm{P}(\sim A) = 1 - \mathrm{P}(A)$.

  (2) For every statements A and B, $\mathrm{P}(A \,\&\, B) = \mathrm{P}(A) + \mathrm{P}(B|A)$ (provided that $\mathrm{P}(A) \neq 0$).

  (3) If A entails B, $\mathrm{P}(A) \leq \mathrm{P}(\mathrm{B})$, and $\mathrm{P}(B|A) = 1$ (provided that $\mathrm{P}(A) \neq 0$).

  (Notice that (3) does not forbid $\mathrm{P}(B|A) = 1$ when A does not entail B).

Interpretations of probability

- The axioms of the probability provide formal constraints to the notion of probability, but they do not define completely this notion.

  Any interpretation of the term ‘probability’ - or of the symbol P(…) - that does satisfy the axioms of probability can be considered admissible, as formally adequate.

- There exist at least three important interpretations of P(…).

  Two of them - the frequency interpretation and the propensity interpretation - are called objective, as they make probability depend on only features of mind-independent reality.

  The third interpretation of P(…) - adopted by the Bayesians - is said subjective, as it defines probability in terms of epistemic features of the subject.

- In the last century, Rudolf Carnap (1891-1970) tried to develop a fourth interpretation of probability, which is commonly called the logical interpretation.

  According to it, the conditional probability of B given A is the degree of partial entailment to which A entails B.

  Carnap’s desire was - roughly - that of formulating a system of rules for establishing the value of the conditional probability of a statement given another statement on the mere grounds of the logical form of these statements.

  Carnap was able to define such rules for only a very limited artificial language, which could hardly be used to explicate the confirmation procedures of real science.

  Notice that none of the first three interpretations of P(…) allows us to establish the value of the conditional probability of a statement given another statement on the mere basis of their logical form (if not in very special cases).

The frequency interpretation (1)

- Consider an ordinary coin which is repeatedly flipped in the standard way. The following sequence of outcomes might be generated:

  H, T, H, H, T, H, T, H, T, T, T …

  This sequence of outcomes can be re-described in terms of the associated sequence of the relative frequencies with which the coin landed heads:

  1/1, 1/2, 2/3, 3/4, 3/5, 4/6, 4/7, 5/8, 5/9, 5/10, 5/11 …

  1/1 says that, so far, the coin has been flipped 1 time and has turned heads 1 time. 1/2 says that, so far, the coin has been flipped 2 times and has turned heads 1 time. 2/3 says that, so far, the coin has been flipped 3 times and has turned heads 2 times. And so on...

- The intuition the advocate of the frequency interpretation wants to account for is, roughly, this:

  (F) Saying that the probability that a given coin lands heads when flipped is 1/2 means saying that, if the coin were repeatedly flipped, the relative frequency with which it would land heads in the long run would approach to 1/2.

  To make (F) more precise, the advocate of the frequency interpretation reformulates it by appealing to the mathematical concept of limit:

  (F*) Saying that the probability that a given coin lands heads when flipped is 1/2 means saying that, as the number of the tosses of that coin tends to infinity (i.e. gets greater and greater), the relative frequency with which the coin lands heads tends to the limit 1/2.

  (Namely, given any number *e* however small, there is always some finite integer *N* such that for all numbers of tosses greater than *N*, the relative frequency with which the coin lands heads does not differ from 1/2 by more than *e*).

The frequency interpretation (2)

- In general, according to the frequency interpretation, the probability of a particular sort of repeatable event is to be identified with the limit of the sequence of the relative frequencies of that event.

- This kind of probability is objective, in the sense that the sequence of the occurrences of a given event is (in general) thought of as independent of our beliefs (the sequence would exist even if we did not exist).

- Notice that, on this interpretation of probability, it does not seem to make any sense to talk of probabilities of a single occurrence. Consequently, the advocates of the frequency interpretation have great difficulties in explaining what we mean with (apparently meaningful) statements like:

  ‘probably, tomorrow, it is going to rain’,

  ‘it is very improbable that John Smith will fail the final exam’.

- A possible drawback of the frequency interpretation is that it defines probability in terms of what would happen if something happened infinitely many times. Since, very plausibly, each physical object will exist for only a limited time, this interpretation defines probability in terms of non-existent and imaginary entities. Some philosophers might find this unacceptable. (However, advocates of the frequency interpretation can retort that science is full of similar idealizations!)

The propensity interpretation

- According to this interpretation - first clearly articulated by Karl Popper - probability is the degree of strength of the propensity (or tendency or chance) of each particular apparatus *X* to produce a given effect *Y* - for instance, the propensity of a flipped coin to land heads.

  On this interpretation, saying that the probability of a flipped fair coin to land heads is 1/2 just means saying that the degree of propensity of this apparatus to give this effect is 1/2.

- The propensity of a particular apparatus to produce an given effect is independent of our beliefs (the apparatus would have exactly the same degree of propensity even if we did not exist). In this sense, this kind of probability is objective.

- The propensity of a given apparatus to produce an given effect is not defined in the terms of any relative frequency but is thought of as a primitive property of the apparatus. (The claim is that the mechanism constituting the apparatus has a causal property that comes in degrees of strength quantifiable by the axioms of the probability calculus).

  On this interpretation, we can attribute a value of probability to a single event.

- Another advantage of this interpretation of probability over the frequency interpretation is that the former need not appeal to any counterfactual infinite sequences, or to the supposition that a given apparatus can exist for an infinite time.

- An objection to this interpretation of probability is that entails weird metaphysical consequences, as it leads us to conflate causes with effects.

  Let *B* be 'this coin lands heads' and let *A* be 'this coin is flipped'. $P(B|A)$ is, on this interpretation, the degree of the propensity of a given flipped coin to cause the effect of landing heads. According the probability calculus, if $P(B|A)$ is defined (i.e. has a determined value), $P(A|B)$ must be defined too. But how can we interpret $P(A|B)$? To say that this is the degree of propensity of a given coin that landed heads to cause the effect of being flipped is just conflating a cause with its effect!

The subjective interpretation

- According to this interpretation, the probability of a states of affair or of a statement *A* is the degree of strength to which a (ideally rational) person believes in *A*.

- Two central assumptions of this interpretation are: (1) belief comes in degrees and (2) degrees of belief comply or (more plausibly) should comply with the axioms of probability.

  Assumption (1) is *prima facie* plausible, at least if 'degree' is intended in the loosened sense of 'gradations'. For we all do have belief gradations of some sort. (Consider the difference in belief strength when you believe that 1+1=2 and when you believe that tomorrow it will rain).

  Assumption (2) is more controversial.

  It may be objected that, though we have belief gradations, we do not have precise numbers that represent such gradations.

  But the Bayesian can reply that associating precise numbers to belief gradations is nothing more than a heuristic device to achieve a clear-cut picture of scientific methodology.

- But how could you associate numbers to your belief gradations? A way to do it is connecting your degree of belief in a possible event *A* with the money you are ready to bet on *A*.

  Suppose your opponent bets $*N* on ~*A*, and that the winner takes all. If you are ready to bet on *A* up to $*N*, then you plausibly believe that the odds that *A* is true are 1/2. While, for instance, if you are ready to bet on *A* up to $ 2*N*, you believe the odds are 2/3. It appears very natural to identify your degree of belief in a statement *A* with the value *f* you assign to the (fair) odds to bet on *A*.

  There is a simple way to determine the value of *f*. If *M* is the amount of money you believe it is fair to bet on *A* when your opponent best *N* on ~*A*, then $f = M/(M + N)$. Let us also assume that $f = 1$ if and only if you already know that *A* is true.

Dutch Book arguments

- Although we have a method to assign numerical values to a subject's belief gradations, it may still be objected that these assignments will not necessarily comply with the axioms of probability - (i.e. that they they may be are probabilistically incoherent).

  For instance, a consequence of the axioms of probability is that if a statement *A* has probability equal to *n*, ~*A* must have probability equal to 1 – *n*. Yet it seems quite possible that a given subject may believe that the odds of *A* are equal to *n* and, at the same time, that the odds of ~*A* are not equal to 1 – *n.* Consequently, that subject's degrees of beliefs will not satisfy the probability calculus.

- The Bayesian agrees that this is de facto possible, as subjects are often - for different reasons - irrational.

  Yet the Bayesian argues that whenever a subject's degrees of belief violate the axioms of probability, the subject can be induced to make obviously irrational bets - precisely, bets with the consequence that she will lose in any case, no matter the outcome of the event about which the bet is made. The Bayesian's conclusion is that the degrees of belief of any rational subject should comply with the axioms of the probability.

- Any particular argument to show that a subject whose degrees of belief violate the axiom of probability will lose in any case is called a Dutch Book argument.

  Here is an example. Suppose you believe that the odds of an event *A* are 7/10 and that the odds of ~*A* are not 3/10, as the probability calculus would require, but 9/10.

  Given your beliefs, if I bet $3 on *A*, you will bet $7 on *A*, and if I bet $1 on ~*A* you will bet $9 on ~*A*. If *A* proves true, you will win $3 but lose $9. If ~*A* proves true, you will win $1 but lose $7. In any case, you will lose $6!

- It can be shown that no Dutch Book argument is possible when your degrees of belief respect the axioms of probability.

The Bayesian notion of confirmation

- The Bayesian accepts the subjective interpretation of probability.
- Intuitively, evidence E confirms hypothesis H if the truth of E would make H more credible. This intuition is formally explicated by the Bayesian definition of (incremental) confirmation:

  (C) E confirms H if and only if $P(H|E) > P(H)$.

  Namely, E confirms H if and only if the conditional probability of H given E is higher than the prior probability of H. (Notice that, the Bayesian, like Hempel, does not assume that E must be verified to confirm H).

  (C) can be formulated to make background knowledge K explicit:

  (C*) E confirms H given K if and only if $P(H|E \,\&\, K) > P(H|K)$

  That is, E confirms H on K if and only if the conditional probability of H given E and K is higher than the probability of H given K.

- (Incremental) disconfirmation is defined as follows:

  (D) E disconfirms H if and only if $P(H|E) < P(H)$.

- Finally, irrelevance is defined as follows:

  (I) E is irrelevant to H if and only if $P(H|E) = P(H)$.

- The definition of confirmation does not allow us to measure the degree of confirmation of H given E (if E confirms H, we just know that $P(H|E) > P(H)$). Two popular confirmation measures are the difference measure and the ratio measure. That is, respectively:

  $d(H, E) = P(H|E) - P(H)$ (i.e. the degree to which E confirms H is $P(H|E)$ minus $P(H)$)
  $r(H, E) = P(H|E) / P(H)$ (i.e. the degree to which E confirms H is $P(H|E)$ divided $P(H)$)

The Bayes theorem

- The Bayes theorem follows from the axioms of the probability calculus and the definition of conditional probability. Suppose *E* is an evidential statement and *H* a hypothesis. This is one of the typical formulations of the theorem:

  (BT) $$P(H|E) = \frac{P(H) \times P(E|H)}{P(E)}$$ (with $P(E) \neq 0$).

  $P(H|E)$ is the conditional probability of *H* given *E*.

  $P(E|H)$ is the conditional probability of *E* given *H* (also called likelihood of *H* given *E*).

  $P(H)$ is the prior probability of *H.*

  $P(E)$ is the prior probability of *E* (also called expectedness of *E*).

- The Bayes theorem allows us to calculate $P(H|E)$ when the values of the three parameters $P(H)$, $P(E|H)$ and $P(E)$ are known.

  Notice that, in this case, we can also evaluate whether $P(H|E)$ > or < or = $P(H)$ - that is, whether *E* confirms, disconfirms or is neutral to *H.*

- If *H* entails *E*, then $P(E|H) = 1$. Consequently, the Bayes theorem can be written in this simplified way:

  (BT*) $$P(H|E) = \frac{P(H)}{P(E)}$$ (with $P(E) \neq 0$).

- From (BT) and (BT*) it is clear that if $P(E) = 1$ or $P(H) = 1$ or $P(H) = 0$, *E* cannot confirm *H*.

Relevant features of the Bayesian notion of confirmation (1)

- According to the Bayesian,

  (C) E confirms H if and only if $P(H|E) > P(H)$.

  (C) satisfies several general and intuitive conditions of adequacy for confirmation. Here are some of its most relevant features:

1. (C) satisfies Hempel's equivalence condition (if E confirms H, and H is logically equivalent to H^*, E confirms H^*).
   For it can be shown that if $P(H|E) > P(H)$ and H is equivalent to H^*, $P(H^*|E) > P(H^*)$.
   The equivalence condition is intuitive and natural.

2. (C) satisfies Hempel's entailment condition (if E entails H, E confirms H).
   For it can be shown that if E entails H (and $P(E) \neq 0$ and $P(H) \neq 1$), $P(H|E) > P(H)$.
   The entailment condition is very intuitive, as it represents a limiting case of confirmation.
   (Notice that if E confirms H because E entails H, E verifies H. This is visible from the fact that, in this case, $P(H|E) = 1$).

3. (C) fulfils none of Hempel's further conditions.
   Notice that all those conditions are to some extent problematic.

4. (C) also satisfies basic hypothetico-deductivism (if H entails E, and H and E are both contingent, E confirms H).
   For it can be shown that if H entails E and $P(E) \neq 1$ and $P(H) \neq 0$, $P(H|E) > P(H)$. (Notice that $P(E) \neq 1$ and $P(H) \neq 0$, E and H must both be contingent).
   Basic hypothetico-deductivism is extremely intuitive.

Relevant features of the Bayesian notion of confirmation (2)

5. (C) allows E to confirm H even if H does not entail E (and if E does not entail H).

   For there are coherent distributions of probability such that $P(H|E) > P(H)$ when H does not entail E (and E does not entail H).

   This renders (C) extremely general and versatile.

6. (C) can apply to any statements E and H, no matter their logical form and whether they include theoretical or only observational terms. E will confirm H simply if there is a (coherent) probability distribution such that $P(H|E) > P(H)$.

   This renders (C), again, extremely general and versatile.

7. (C) accounts for the fact that evidence incompatible with a hypothesis falsifies the hypotheses.

   For it can be shown that, if E is logically incompatible with H (and $P(E) \neq 0$), then $P(H|E) = 0$.

9. (C) satisfies the intuitive condition for confirmation that the same evidence cannot confirm both a hypothesis and its logical negation (this is a weaker version of Hempel's consistency condition).

   For it can be proven that if $P(H|E) > P(H)$, it is impossible that $P(\sim H|E) > P(\sim H)$.

10. Finally, (C) explicates the widespread intuition that if a hypothesis is confirmed by an evidential consequence, the more improbable that consequence is, the more probable the hypothesis becomes.

    For it can be shown that if H entails both E and E^*, and $P(E) < P(E^*)$ (and $P(H) \neq 0$), then $P(H|E) > P(H|E^*)$.

How the Bayesian "dissolves" the tacking by conjunction problem

- Bayesian confirmation theory is affected by the tacking by conjunction problem.

  This follows from the fact that the Bayesian notion of confirmation satisfies basic hypothetico-deductivism. Namely, if H entails E, and $P(E) \neq 1$ and $P(H) \neq 0$, then $P(H|E) > P(H)$, so that E confirms H.

  The tacking by conjunction problem arises in this way:

  Suppose H entails E, and $P(E) \neq 1$ and $P(H) \neq 0$. In this case, E confirms H by basic hypothetico-deductivism.

  Consider now any statement X such that $P(H \& X) \neq 0$. As H entails E, $H \& X$ entails E too. Since, by assumption, $P(E) \neq 1$ and $P(H \& X) \neq 0$, by basic hypothetico-deductivism E also confirms $H \& X$.

  This may appear weird, as X is a random statement for the truth of which E may be completely irrelevant. (Suppose E is 'this body is faster than that body', H is Newton's kinematics, and X is 'my cat does not like any fish').

- But the Bayesian bites the bullet and argues, like Grimes, that it is instead intuitive that, if E confirms $H \& X$ because this straightforwardly follows from basic hypothetico-deductivism.

  The Bayesian also emphasizes that her notion of confirmation does not satisfy the special consequence condition (i.e. if A confirms B, and B entails C, A confirms C), so that from the fact that E confirms $H \& X$, it does not follow that E confirms X (which would be a disaster).

How the Bayesian resolves the tacking by disjunction problem

- Bayesian confirmation theory is affected by the tacking by disjunction problem.

  This follows, again, from the fact that the Bayesian notion of confirmation satisfies basic hypothetico-deductivism. The tacking by disjunction problem arises in this way:

  Suppose *H* entails *E,* and P(*E*) $\neq$ 1 and P(H) $\neq$ 0. In this case, *E* confirms *H* by basic hypothetico-deductivism.

  Consider now any statement *Y* such that P(*E* v *Y*) $\neq$ 1. As *H* entails *E, H* entails *E* v *Y.* Since P(*H*) $\neq$ 0 and P(*E* v *Y*) $\neq$ 1, by basic hypothetico-deductivism *E* v *Y* confirms also *H.*

- To solve this problem, the Bayesian first emphasizes that to evaluate properly the subjective probability of *H* given evidence, the evidence should properly be the total evidence one has in a given time. (As we have seen, Skyrms accepts this requirement too).

  A simplified version of the Bayesian's argument, then, goes on like this: if total evidence includes *E* v *Y*, this is so because it also includes (1) *E* or (2) *Y* or (3) *E* & *Y.* For we know that *E* v *Y* is true if we know that (1) or (2) or (3) is true. But then, it can be shown that *E* v *Y* has no weight for the process of the confirmation of *H*.

  Case (1). Total evidence is (*E* v *Y)* & *E*, which is logically equivalent to *E*. Therefore, to ascertain whether total evidence confirms *H*, just check whether P(*H*|*E*) > P(*H*).

  Case (2). Total evidence is (*E* v *Y)* & *Y*, which is logically equivalent to *Y*. To ascertain whether total evidence confirms *H*, just check whether P(*H*|*Y*) > P(*H*).

  Case (3). Total evidence is (*E* v *Y)* & (*E* & *Y*), which is logically equivalent to *E* & *Y*. To ascertain whether total evidence confirms *H*, just check whether P(*H*|*E* & *Y*) > P(*H*).

  In each case, total evidence may confirm, disconfirm or be neutral to *H*. For example, in (2), *Y* may confirms, disconfirms or be neutral to *H*. There is nothing paradoxical in all this!

How the Bayesian handles the Duhem problem

- The Duhem thesis says that only when a hypothesis is embedded within a larger network of auxiliary hypotheses and assumptions, the logical conjunction of all of them will entail testable consequences.

  A problem raised by this thesis to hypothetico-deductivism is that if a logical conjunction of hypotheses is confirmed by deductive evidence, it is unclear which of these hypotheses is actually confirmed by evidence (and, more generally, how degrees of conditional probability given evidence distribute over the different hypotheses of the conjunction).

- Bayesianism can handle this problem. Suppose the conjunction H_1 & H_2 & H_3 entails E.

  As we know, if $\mathrm{P}(E) \neq 1$ and $\mathrm{P}(H_1 \,\&\, H_2 \,\&\, H_3) \neq 0$, E does confirm H_1 & H_2 & H_3.

  Can I establish which of the conjuncts H_1, H_2, H_3 is confirmed (or disconfirmed) by E? Yes! I simply have to apply the Bayes theorem to each conjunct.

  For example, to ascertain whether E confirms H_2, I have to check whether $\mathrm{P}(H_2|E) > \mathrm{P}(H_2)$. Since, by Bayes' theorem,

$$\mathrm{P}(H_2|E) = \frac{\mathrm{P}(H_2) \times \mathrm{P}(E|H_2)}{\mathrm{P}(E)},$$

  all I have to do is to assess whether:

$$\frac{\mathrm{P}(H_2) \times \mathrm{P}(E|H_2)}{\mathrm{P}(E)} > \mathrm{P}(H_2).$$

  Naturally, I will be able to assess it - and to calculate the conditional probability H_2 given E, i.e. $\mathrm{P}(H_2|E)$ - if I have fixed the values of $\mathrm{P}(H_2)$, $\mathrm{P}(E|H_2)$ and $\mathrm{P}(E)$.

  It may turn out that E confirms, disconfirms or is neutral to H_2.

Advantages and disadvantages of Bayesianism (1)

- Bayesian confirmation theory, as a quantitative theory, allows for very precise evaluations and comparative evaluations about the confirmation of hypotheses and theories.
- The Bayesian notion of confirmation is extremely general so that it applies to any possible type of hypothesis.
- Furthermore, many features of the Bayesian notion of confirmation are very intuitive and make Bayesian confirmation theory versatile and, in different respects, close to real confirmation procedures of scientists.
- Finally, the formal apparatus at disposal of the Bayesian allows her to solve, or to handle, problems and paradoxes that afflict other theories of confirmation.

- Why, then, should not every one be a "happy Bayesian"?

  Because also Bayesian confirmation theory is affected by many problems, some of which are extremely technical. Here, let us consider only four general problems.

  The problem of rational justification

  First, notice that Bayesianism provides no general solution to the problem of justifying rationally inductive logic (though Bayesians have proposed some partial solution of the problem of justifying induction by enumeration or generalization).

  Consider a language on which a probability function that satisfies our intuitive judgments of inductive support is defined and select all pairs (X, Y) of statements for which the value $P(X|Y)$ is very high.

  Following Skyrms, we can say that this inductive logic is rationally justified if and only if for most pairs (X,Y), if X is true, Y is true too. But the Bayesian offers no reason to believe that this is actually the case.

Advantages and disadvantages of Bayesianism (2)

<u>The problem of the priors</u>

Although the problem of rational justification probably afflicts any inductivist methodology and, therefore, does not play against Bayesianism in particular, Bayesian methodology is plagued by a difficulty that not all inductivist methodologies have: the problem of the priors.

Bayesian confirmation theory rests on a subjective notion of probability. The Bayesian has thus no guarantee that distinct subjects S and S* will ascribe the same value to, for instance, the prior probability of a hypothesis *H*, or the prior probability of an evidential statement *E*.

But if S and S* give P(H) or P(E) divergent values, by employing Bayes' theorem S and S* will in general obtain divergent values of $P(H|E)$. And S and S* may also reach contrasting conclusions about whether *E* confirms *H* (indeed, this seems to happen quite often!)

No Dutch Book argument could force S and S* to converge on the same value of P(H) or of P(E). A Dutch Book argument can just persuade S to make her own assignments of probability coherent in themselves, and persuade S* to make her own assignments of probability coherent in themselves. The same language will admit, however, many alterative and incompatible distributions of probability that prove internally coherent.

Bayesians have replied that this is no serious problem. For subjects beginning with divergent prior probabilities will tend to converge in their final conditional probabilities, given a suitably long series of shared observations.

A problem is that the theorems that show this interesting result do not work for all possible initial values: there are initial degrees of prior probability which are consistent with the axioms of probability, but which will lead to no eventual convergence.

So, this reply is not fully convincing!

Advantages and disadvantages of Bayesianism (3)

The requirement of logical omniscience

The subjective interpretation of probability accepted by the Bayesian requires that our degrees of belief satisfy the axioms of the probability.

But this requirement can be met only if we are omniscient about deductive logic. For the probability calculus requires that the probability of any statement be no greater than any of its logical consequences. The effect is that, to make sure that our degrees of belief comply with the probability calculus, we should know all consequences of our beliefs (and of any conjunction of them). This is an unrealistic standard for human beings!

To cope with this difficulty, most Bayesians treat the assumption of logical omniscience as an ideal to which human beings can only more or less approximate.

The problem of old evidence

On the Bayesian account, if a statement E is true, then $P(E) = 1$. On the other hand, if $P(E) = 1$, it is impossible that $P(H|E) > P(H)$. That is, E cannot confirm any hypothesis.

These two facts raise a puzzle for Bayesian confirmation theory. Suppose E is an observation statement that has been known for some time - that is, E is old evidence. Since E is know to be true, $P(E) = 1$.

Let H be a scientific theory formulated to explain observations other than E.

It is however discovered that H implies E.

In scientific practice, the discovery that H implies E would typically be taken to confirm H. But, for the Bayesian, E cannot confirm H, as $P(E) = 1$!

Many solutions of this problem have been proposed, but none appears fully satisfactory.

THE END!