Phil. 4400
Notes #1: The problem of induction

I. Basic concepts:

The problem of induction:
- Philosophical problem concerning the justification of induction.
- Due to David Hume (1748).

Induction: A form of reasoning in which
a) the premises say something about a certain group of objects (typically, observed objects)
b) the conclusion generalizes from the premises: says the same thing about a wider class of objects, or about further objects of the same kind (typically, the unobserved objects of the same kind).

- *Examples:*

All observed ravens so far have been black.
So (probably) all ravens are black.

The sun has risen every day for the last 300 years.
So (probably) the sun will rise tomorrow.

Non-demonstrative (non-deductive) reasoning:
- Reasoning that is not deductive.
- A form of reasoning in which the premises are supposed to render the conclusion more probable (but not to entail the conclusion).

***Cogent* vs. *Valid* & *Confirm* vs. *Entail*:**
'Cogent' arguments have premises that *confirm* (render probable) their conclusions.
'Valid' arguments have premises that *entail* their conclusions.

The importance of induction:
- All scientific knowledge, and almost all knowledge depends on induction.
- The problem had a great influence on Popper and other philosophers of science.

Inductive skepticism:
Philosophical thesis that induction provides no justification for (*no reason* to believe) its conclusions.

II. An argument for inductive skepticism

1. There are (at most) 3 kinds of knowledge/justified belief:
   a. Observations
   b. A priori knowledge
   c. Conclusions based on induction
2. All inductive reasoning presupposes the "Inductive Principle" (a.k.a. the "uniformity principle"): "The course of nature is uniform", "The future will resemble the past", "Unobserved objects will probably be similar to observed objects"
3. So inductive conclusions are justified only if the IP is justified. (From 2.)
4. The IP cannot be justified:
   a. The IP cannot be directly observed.

b. The IP cannot be known a priori. (Contingency argument)
c. The IP cannot be justified by induction. (Circularity argument)
d. So the IP cannot be known/justified. (From 1, 4a, 4b, 4c.)

5. Therefore, no inductive conclusion is justified. (From 3, 4.)

Comments:

- (5) is *inductive skepticism*.
- Notice how radical (5) is. Conclusion is *not* merely: "Inductive conclusions are never 100% certain."

III. Preview of some responses

- Some reject (2). Claim that (2) begs the question / presupposes deductivism.
- Some reject (4b). (Russell)
- J.S. Mill rejects (4c).
- Some reject (1), claim that some other kind of inference has been overlooked. (Foster, Stove)
- Many have embraced the conclusion (5):
    - Hume
    - Popper

Phil. 4400
Notes #2: Popper & Falsificationism

Elements of Popper's theory:

I. Inductive skepticism

- Accepts Hume's argument for inductive skepticism (see previous lecture).
  - Principle of Induction is not analytic.
  - Principle of Induction not known by experience. Infinite regress problem.
  - No synthetic, a priori knowledge.
- Probabilism no help.
  - Just requires another principle of induction, e.g., "The future will *probably* resemble the past."
  - This faces the same difficulties.
- Hence, we must rely on *deductivism:* The view that all legitimate reasoning is deductive.

II. The deductive approach to science: Falsificationism

- A scientific theory is formulated. No logic of this. Distinguish: Logic vs. psychology
- Predictions are *deduced* from the theory.
- Predictions compared w/ observation:
  - If true, the theory is "corroborated."
  - If false, the theory is *falsified*.
- But corroborated theories *are not thereby supported*.

  "Nothing resembling inductive logic appears in the procedure here outlined. I never assume that we can argue from the truth of singular statements to the truth of theories. I never assume that by force of 'verified' conclusions, theories can be established as 'true', or even as merely 'probable'." (33)

III. The criterion of demarcation

- The problem of demarcation: How to distinguish scientific (empirical) theories from non-scientific theories.
- Answer: Scientific theories are *falsifiable*.
- Key point: The *asymmetry* between verification and falsification: A universal generalization can be falsified, but never verified.
- Examples:

  astrology, Marxism, sociobiology, psychoanalysis

IV. Objections

A. The Duhem-Quine Thesis:

An individual scientific theory entails no empirical predictions in isolation; auxiliary assumptions are always required.

• Corollary: A scientific theory can (consistently) be maintained come what may, provided one is willing to make sufficient adjustments elsewhere in one's system.
• Hence, the asymmetry Popper sees between falsification & verification is illusory.
• *Example:* How to refute the law of gravity?

B. Kuhn's Anomalies:

• There are *always* anomalies.
• A theory is rejected when *enough* anomalies accumulate, & there is a *better* theory. It is not "falsified".

C. The case of probabilistic theories:

• Probabilistic theories cannot be falsified.
• But some are scientific. E.g., quantum mechanics.
• Popper: introduces a "methodological rule" permitting them to be "regarded as" falsified.
• Stove's objection:
  - This approach would enable anyone to defend any logical claim (including inductivism).
  - More to the point: Popper can't change the facts of logic by fiat.
    1. Popper says only falsifiable theories are scientific.
    2. Popper also embraces deductivism.
    3. No probabilistic theory can be deductively falsified.
    4. So, on Popper's theory, no probabilistic theory is scientific.

  No "methodological rule" can change any of these facts.

D. The problem of pointlessness:

• On Popper's theory, there is *no reason* for believing any scientific theory. So what's the point?

Phil. 4400
Notes #3: Inference to the best explanation

Main idea:

- Induction depends upon inference to the best explanation. An inductive inference actually requires two steps:
  - *First*: An inference from observations to a hypothesis that provides the best explanation for those observations.
  - *Second*: An inference from that hypothesis to further predictions. (This step is deductive.)
  - Examples:

	Observation: This coin has come up heads 100 times in a row.
(inference to the best explanation)	
	Hypothesis: This coin has heads on both sides.
(deduction)	
	Prediction: This coin will continue to come up heads in the future.

	Observation: Bodies have always behaved gravitationally.
(inference to the best explanation)	
	Hypothesis: It is a law of nature that bodies behave gravitationally.
(deduction)	
	Prediction: Bodies will behave gravitationally in the future.

  - The *hypothesis* is justified because,
    a) unless there were some explanation, the *observation* would be highly improbable, and
    b) the hypothesis provides the best explanation.
  - *Induction is not a primitive form of inference*:
    Imagine that we somehow knew there were no laws of nature. Then would we be justified in thinking bodies will continue to behave gravitationally?

Skeptical objections:

- *Objection*: The observed regularity really doesn't require any explanation, because it is just as likely as every other possible sequence of events. (Example: Coin toss outcomes.)
  Reply: What matters is comparison of the probability of the observed regularity *on the alternative hypotheses* (not its probability compared to that of other possible observations).

- *Objections from alternative hypotheses:*

a) There is no relevant law; past regularity is purely due to chance.
   Problem:
   - This hypothesis is extremely improbable.

b) It is a law of nature that: (up until 2100 A.D., bodies behave gravitationally).
Problems:
- This creates a further mystery: What is so special about the year 2100?
- Actually, this hypothesis seems to be metaphysically impossible: the current time cannot be a causally relevant factor. (my point)

c) There is a law of nature that (bodies behave gravitationally), but the law ceases to exist in the year 2100 A.D.
Problems:
- This creates a further mystery: What is so special about the year 2100?
- Actually, this hypothesis seems to be metaphysically impossible: Laws of nature cannot stop existing. (my point)

d) It is a law that (in ϕ-circumstances, bodies behave gravitationally).
Where we define "ϕ-circumstances" in such a way that it applies to all the times when we have actually been observing bodies, but is unlikely to apply to other times.
Problem:
- This hypothesis gives a different explanation for different cases of gravitational behavior.
- Our explanation gives a *unified* explanation.
- Unified (and hence, simpler) explanations are more likely to be true.
    - It is improbable that you just happen to have always been observing during one of the ϕ-circumstances. I.e., if there were 5 million different causally relevant factors, it is improbable that you would happen to have been looking during exactly the times when one of the relevant circumstances held, unless those circumstances hold almost all the time.

Phil. 4400
Notes #4: Bayesianism

I. The pure mathematics of probability:

A. Measure theory:

- A measure on a set, S, is a function, M, which assigns positive real numbers to subsets of S, such that $M(\varnothing) = 0$, and such that, when A and B are disjoint subsets of S, $M(A \cup B) = M(A) + M(B)$. (Disjoint sets are sets that have no members in common.)
- Note how this captures the bare bones of the intuitive idea of 'measuring' the sizes of sets.

B. Probability theory as a branch of measure theory:

- A *probability measure* is a specific kind of measure: viz., a measure in which the measure of the total set is 1. In other words, a probability measure is a function, P, satisfying these axioms:

1. $P(T) = 1$	where T is the total set.
2. $P(A) \geq 0$	where A is any subset of T.
3. $P(A \cup B) = P(A) + P(B)$	whenever A and B are disjoint subsets of T.

In addition, an expression "|" can be defined in this way:

4. $P(A|B) = \dfrac{P(A \cap B)}{P(B)}$ (this will become useful later)

• A probability of *x* (in the *pure mathematics* sense) is just a number that is assigned to *x* by a probability measure.

II. Interpretations of probability:

- Probabilities are applied either to *events* or to *propositions*. Applied to propositions, there are 4 axioms:

1. $P(T) = 1$	where T is any tautology
2. $P(A) \geq 0$	where A is any proposition
3. $P(A \vee B) = P(A) + P(B)$	whenever A and B are mutually exclusive
4. $P(A \,\&\, B) = P(A) \times P(B|A)$	where "P(B|A)" is the probability of B given A

• Interpretations of probability:

I. Frequency interpretation (objective, physical)
II. Probabilities as objective, single-case propensities, or degrees of causal influence (objective, physical)
III. Rational degrees of belief. (subjective)
Related idea: fair betting odds. (Related: Dutch book arguments)
IV. A logical relation (like degrees of entailment). (Similar to III. Objective, non-physical.)
Related idea: probability of *x* as the measure of the set of possible worlds in which *x* is true.

• For our purposes (investigating the justification of induction), interpretations III and IV are the important ones.

III. Interesting results of probability theory:

• Probabilities of logically equivalent hypotheses are equal.
• If A≡B, then whatever confirms A confirms B, and vice versa.
  Bayesian definition of "confirmation":
  A confirms B iff P(B|A) > P(B).
• Bayes' theorem: Suppose H is a hypothesis, and E is a piece of evidence relevant to H. Then

5. $P(E \,\&\, H) = P(H \,\&\, E)$	obvious; logically equivalent prop's
6. $P(E) \times P(H\|E) = P(H) \times P(E\|H)$	applying axiom 4 to step 5
7. $P(H\|E) = \dfrac{P(H) \times P(E\|H)}{P(E)}$	6, division by P(E)

(7) is Bayes' Theorem. Another form of the theorem:

8. $$P(H|E) = \frac{P(H) \times P(E|H)}{P(E|H)\times P(H) + P(E|{\sim}H)\times P({\sim}H)}.$$

• Exercise for the student: prove (8).
• Note how Bayes' theorem explains aspects of scientific practice:
  a) A theory is confirmed by *predicted* observations. (P(E|H) is high.)
  b) A theory is better-confirmed by the less initially probable observations (or: observations that are less probable on the alternative hypotheses).
  c) Obviously, the higher the initial probability of H, the higher the probability of H in the light of observations.

• Fun aside: probability densities:
  • Suppose X is a random variable with real-number values between 2 and 3. What is the probability that $X = 2.3$ (exactly)? Answer: 0.
  • More useful concept: the probability density function, ρ:

$$\rho(x) = \lim_{e\to 0} \frac{P(x-e < X < x+e)}{2e}$$

  • Or: A probability density is a function that, when you integrate it over an interval, gives you the probability of the variable falling in that interval.
  • Note that the probability density is normally greater than 0 (and could be greater than 1).
  • Analogy: compare density of matter.

IV. The problem of induction: the Bayesian approach

- The Bayesian approach to the problem of induction: non-demonstrative reasoning is reasoning in accordance w/ the probability calculus. When learning *e*, one should alter one's degree of belief in *h* by "conditionalization".
  - *Conditionalization:* Upon learning new evidence *e*, you set your new degree of belief in *h* equal to what you previously estimated as the probability of *h* given *e*, i.e.: $P_{new}(h) = P_{old}(h|e)$.
- e *confirms* h iff $P(h|e) > P(h)$.
- Recall Bayes' Theorem:

$$P(h|e) = \frac{P(h) \times P(e|h)}{P(e)}.$$

- Apply this to Foster's proposal of inference to the best explanation. *h* is the hypothesis of natural necessity. *e* is the existence of some regularity. P(*e*) is low; P(*e*|*h*) is high. Thus, *e* confirms *h*.

V. A problem for Bayesianism:

What determines the *initial* probabilities?

- "Subjective" Bayesians: Many different sets of initial probabilities are equally good.
- "Objective" Bayesians: The probability calculus must be supplemented, to allow only one value for each initial probability. How?
  - A proposal: *the Principle of Indifference*: When there is no evidence relevant to which possibility is the case, assign each possibility the same probability.
  - A puzzle for the principle of indifference:

Problem: You know that Sue has traveled 100 miles, and it took her somewhere between 1 hour and 2 hours. That's all you know. What is the probability that her trip lasted between 1½ and 2 hours?

First answer: Applying the principle of indifference: 1 ½ - 2 hours is 50% of the total range of possible *durations*. Therefore, give it a 50% probability. Answer: ½.

Second answer: We know that Sue's average velocity during the trip was between 50 mph and 100 mph. Sue's trip lasted between 1½ and 2 hours iff her speed was between 50 mph and 67 mph. This is 1/3 of the total range of possible *velocities*. Answer: ⅓.

- A general statement of the problem:
  - Different *partitions* of (ways of dividing up) a given set of possibilities are possible.
  - Applying the Pr. of Indifference across different partitions yields inconsistent results.
  - Hence, the Pr. of Indiff. either is inconsistent, or requires a specification of a privileged way of partitioning.

Phil. 4400
Notes #5: David Stove's argument

I. Important concepts:

- *Inductive inference :* Two kinds:
  a) Inference from the frequency of a trait in a certain sample drawn from a larger population, to the frequency of that trait in the whole population.
  b) Inference from the premises of (a) to a prediction about the presence or absence of that trait in a particular unobserved individual.
- *Probabilistic independence :* A is probabilistically independent of B iff P(A|B) = P(A) (equivalently: P(B|A) = P(B)).
- *The skeptical thesis :* Two interpretations of it:
  a) Inductive inference has no cogency. P(H|E) = P(H), whenever E is inductive evidence, and H is an inductive conclusion.
  b) Alternatively: Separate observations are always probabilistically independent.
  Note: Thesis (a) is obviously probabilistically incoherent; (b), however, is not.
- *Proportional syllogism :* Everyone agrees that the following kind of inference is cogent:
  1. 99% of all A's are B.
  2. *x* is an A.
  3. ∴ *x* is B.
- *The law of large numbers :* This is a well-known theorem of probability:
  If the probability of an event E at each trial is *x*, then in a large number of trials, the frequency with which E occurs will almost certainly be close to *x*. (With increasing certainty as the number of trials increases.)
- *General strategy :* To show that an inductive inference can be reconstructed with proportional syllogisms & the law of large numbers; thereby showing that inductive inference is cogent.

II. The problem:

Assume the following:

- Pop is a population of 1 million ravens.
- S is a sample, from Pop, of 3000 ravens.
- 95% of the ravens in S are black.

To prove: It is highly probable that:

- Approximately 95% of the ravens in Pop are black.
- The next raven observed will be black.

General argument:

1. Almost all the 3000-fold samples of Pop are representative (no matter what the proportion of black ravens in Pop). (Arithmetical form of law of large numbers.)
2. Therefore, S is almost certainly representative. (From 1; proportional syllogism.)
3. The proportion of black ravens in S is 95%. (Given.)

4. Therefore, almost certainly, the proportion of black ravens in Pop is close to 95%. (From 2,3; deduction.)
5. Therefore (probably), the next observed raven from Pop will be black. (From 4; proportional syllogism.)

Further elaboration:
(1) Almost all the 3000-fold samples of Pop are representative:

- Best case: Suppose the proportion of black ravens in Pop = 100%. Then all samples are representative.
- Other best case: Suppose the proportion of black ravens in Pop = 0%. Then all samples are representative.
- Worst case: Suppose the proportion of black ravens in Pop = 50%. Even so, the vast majority of samples are representative:

  In general:

$$C(m,n) = \frac{m!}{n!(m-n)!}$$

  # of 3000-fold samples in Pop:

$$C(3000,\ 1\ million) = \frac{1{,}000{,}000!}{(3000!)(997{,}000!)} \approx 10^{8867.9}$$

  # of "representative" samples: We call a sample "representative" if it matches within 3% the frequency of black ravens in Pop. Hence, S will be representative iff it contains between 1410 and 1590 black ravens (and between 1590 and 1410 non-black ones). The number of samples containing 1410 black ravens and 1590 non-black ones is:

$$C(1410{,}500000) \times C(1590{,}500000) = \frac{500{,}000!}{(1410!)(498590!)} \times \frac{500{,}000!}{(1590!)(498410!)}$$

  This is not enough: We need the # of samples containing 1410 black ravens, + the # containing 1411 black ravens, + ... + the # containing 1590 black ravens. In other words:

$$\sum_{n=1410}^{1590} \frac{(500{,}000!)^2}{n!\,(500{,}000-n)!\,(3000-n)!\,(497{,}000+n)!} \approx 10^{8867.9-.00087}$$

  The proportion of representative samples, therefore, is:

$$\frac{10^{8867.9-.00087}}{10^{8867.9}} = 10^{-.00087} \approx 99.8\%$$

- Further important points:
  - The qualitative result holds for any population size, and for any sample size ≥ 3000; i.e., the sample will almost certainly be representative. (Statisticians figure out stuff like this.)
  - How does this relate to inference to the best explanation? In the "general argument" above:
    - (3) as the observation
    - (4) as the hypothesis
    - (5) as the prediction

Phil. 4400
Review of unit 1

Know what these things are:

Principle of induction
Kinds of inference:
- Demonstrative
- Non-demonstrative, incl.:
  - Induction
  - Inference to the best explanation
  - Proportional syllogism

Confirmation & cogent inferences
Inductive skepticism
Deductivism
Duhem-Quine thesis
Probability
- Axioms of probability (all 4)

Bayesianism
- Their interpretation of probability
- Definition of confirmation
- Bayes' Theorem & its epistemological significance
- How to update beliefs

Law of large numbers
Principle of Indifference
Problem of demarcation
- Popper's view of it

Know what these people thought about non-deductive inference:

Hume
Popper
Foster
Stove

Know these arguments:

Arg. for inductive skepticism
- Where (if anywhere) the above people would disagree with it.
- The 3 kinds of knowledge recognized by the arg.

Stove's objection to Popper
The Duhem-Quine objection to Popper
Foster's justification of induction
- The steps involved in justifying induction
- His criticism of alternative explanations

Stove's vindication of induction--
- The main steps in his argument & how they are justified

Main objection to Pr. of Indifference