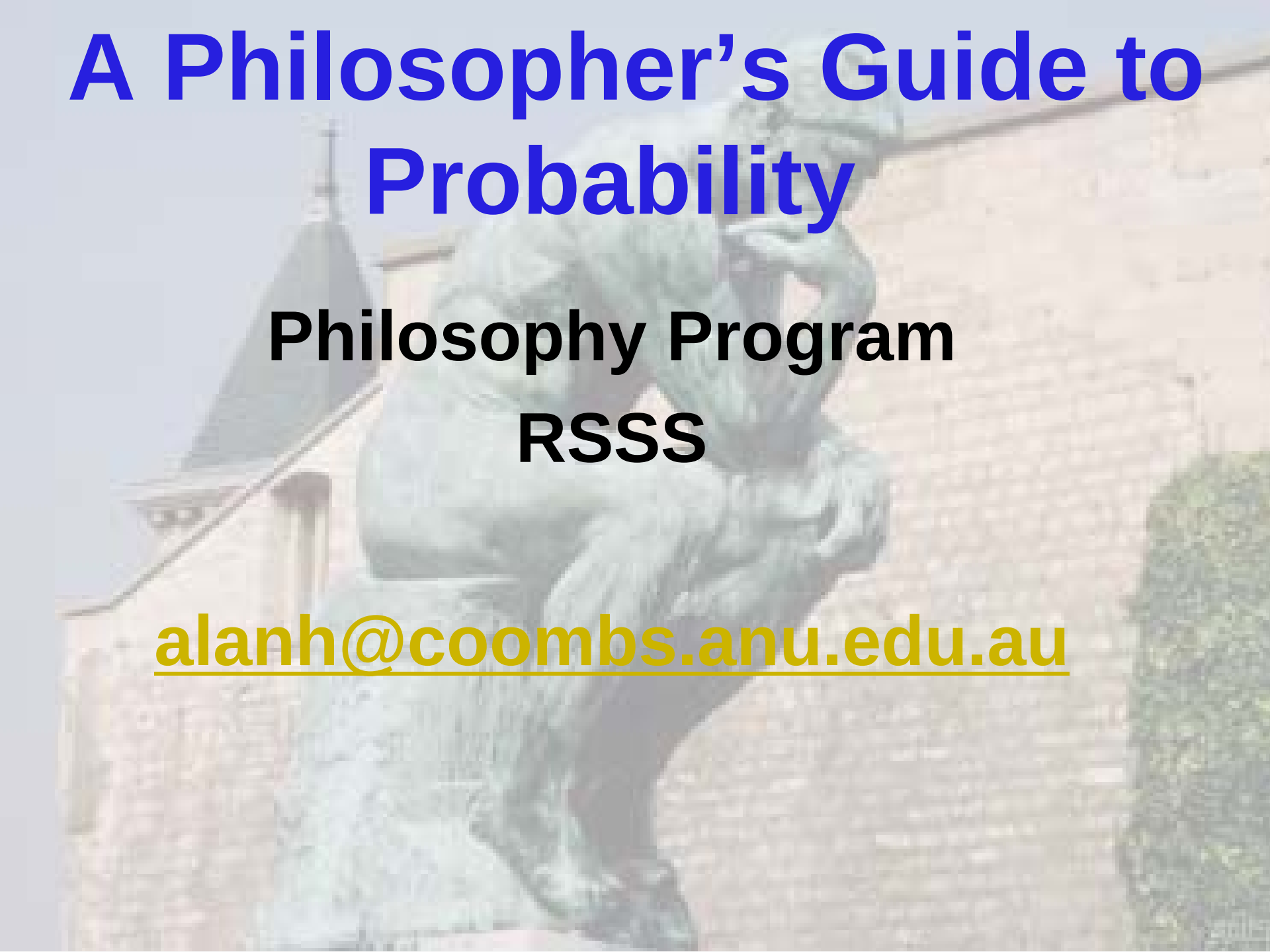
A Philosopher's Guide to Probability

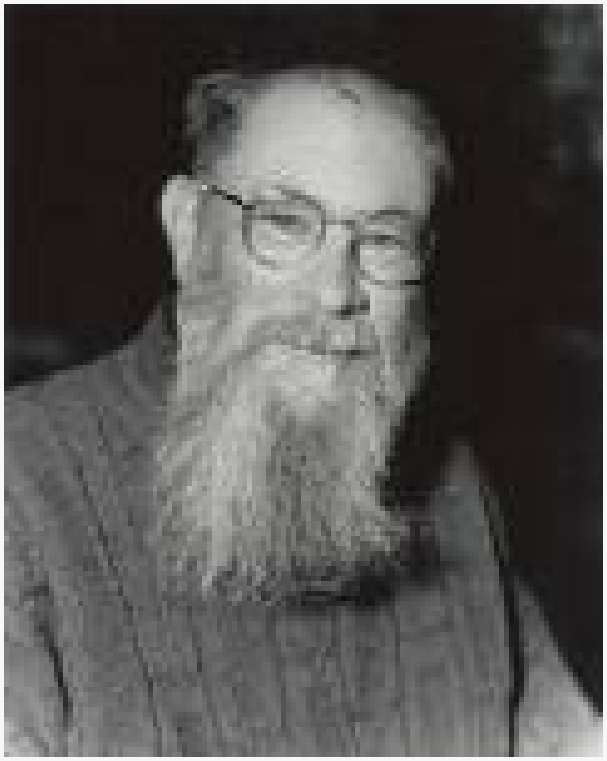


Philosophy Program

RSSS

alanh@coombs.anu.edu.au

David Lewis



Richard Jeffrey

Coming attractions

- Autobiographical prelude
- One minute history of probability
- Survey of the leading 'interpretations' of probability: classical, logical, frequentist, propensity, and subjectivist.
- Some recent developments, and my best bets for future avenues of research in the philosophical foundations of probability.

Autobiographical prelude

“What is probability?”

Autobiographical prelude

… but enough about me …

- “Probability is the very guide of life”

  - Bishop Butler

One-minute history of probability

- Epicurus, Lucretius
- Averroes
- Pascal-Fermat correspondence
- *Port-Royal Logic*
- 3 centuries of development
- Kolmogorov's axiomatization
- More recent hot topics: the end of this talk!

What is probability?

- Two ways of understanding this question:
  - How should probability theory be *formalized*?
  - What do statements of probability *mean*?

The formal theory of probability

- Probabilities are numerical values that are assigned to ‘events’, understood to be certain sets of possibilities belonging to some ‘universal set’ Ω (the set of all possible outcomes)

The formal theory of probability

- Probabilities conform to the following axioms:
  - $P(X) \geq 0$.
  - $P(\Omega) = 1$.
  - $P(X \text{ or } Y) = P(X) + P(Y)$ if X and Y cannot both occur.

The formal theory of probability

- The *conditional probability of X given Y* is given by the ratio of unconditional probabilities:

  $P(X \mid Y) = P(X \text{ and } Y)/P(Y)$, provided $P(Y) > 0$.

The formal theory of probability

- *Bayes' theorem*

  $P(A \mid B) = P(B \mid A).P(A)/P(B)$, if $P(B) > 0$

The formal theory of probability

- My rant against Bayes' theorem's mythic status:
  - "Bayesianism" is a triply curious name.
  - Bayes' theorem "describes the way in which one's beliefs about observing 'A' are updated by having observed 'B'." - Wikipedia

The ‘interpretations’ of probability

- The central *philosophical* question: what sorts of things are probabilities?
- The term ‘interpretation’ is misleading here.

Classical interpretation

- Suppose that our evidence does not discriminate among the members of some set of possibilities.
- Then the probability of an event is simply the fraction of the total number of possibilities in which the event occurs.

Classical interpretation

- Familiar from games of chance, probability puzzles.

Classical interpretation

- Problems:
  - Infinite cases
  - Contradictory results!
  - Bertrand's paradoxes

Logical interpretation

- Probability is an extension of logic.
- Valid arguments - e.g.

  p

  If *p* then *q*

  Therefore, q

Logical interpretation

- But an argument's premises may strongly support a conclusion without entailing it - e.g.

  the sun has risen every day in our experience

  Therefore,

  the sun will rise tomorrow.

- The strength of this support comes in degrees.

Logical interpretation

- Carnap called this degree of support the *logical probability* that an argument's conclusion is true, given that its premises are true.

Logical interpretation

- Problems:
  - Carnap's logical probabilities were purely based on syntax. But Goodman showed that inductive logic must be sensitive to the meanings of words, for syntactically parallel inferences can differ wildly in their inductive strength.

Logical interpretation

All observed snow is white.

Therefore,

All snow is white.

is an inductively strong argument.

However,

All observed snow is observed.

Therefore,

All snow is observed.

is inductively weak.

Frequency interpretations

- The probability of an outcome *A* in a reference class *B* is the proportion of occurrences of *A* within *B*.
- E.g., the probability of ‘6’ on a die that lands ‘6’ 3 times out of 10 tosses is, according to the frequentist, 3/10.

Frequency interpretations

- This is the dominant interpretation among scientists.
- It underpins classical (Neyman-Pearson) statistics.

Frequency interpretations

- Problems:
  - Problem of the single case.
  - Irrational-valued probabilities, like $1/\sqrt{2}$.

Frequency interpretations

- Some frequentists (notably Reichenbach 1949, and von Mises 1957) address this problem by considering infinite reference classes of hypothetical occurrences. Probabilities are then defined as limiting relative frequencies in suitable infinite sequences of trials.
- Coin example.

Frequency interpretations

- This creates new problems:
  - There is apparently no fact of the matter of how the coin in my pocket would have landed if it had been tossed indefinitely—it *could* yield any hypothetical limiting relative frequency that you like.
  - By suitably reordering the trials, the limiting relative frequency can be whatever you like.

Frequency interpretations

- The reference class problem.
  - Geelong vs Hawthorn: prior to the Grand Final
    - Geelong had won 22/23 = 0.96 of its games this year
    - Hawthorn had won 18/23 = 0.78

Propensity interpretations

- Like frequency interpretations, these regard probability as an objective feature of the world. Probability is thought of as a physical propensity, or disposition, or tendency of a system to produce given outcomes.

Propensity interpretations

- According to *single-case* propensity theories, propensities measure a system's tendencies to produce given outcomes.
- According to *long-run* propensity theories, propensities are tendencies to produce long-run outcome frequencies over repeated trials.

Propensity interpretations

- Problems:
  - Single-case propensities are untestable.
  - The long-run propensity view risks collapsing into frequentism.
  - The interpretation is uninformative.
  - How do we even know that propensities obey the probability calculus?

Subjectivist interpretations

- Probabilities are *degrees of belief*, or *credences* of appropriate agents.
- They presumably are not actual humans.
- Rather, they are ideally rational agents.

Subjectivist interpretations

- Betting interpretation of credences:

  your probability for the coin landing Heads is 1/2

  if and only if

  you are prepared to buy or sell for 50 cents a ticket that pays $1 if the coin lands Heads, nothing otherwise

  and so on.

Subjectivist interpretations

- Problems:
  - Like those of behaviourism, and of operationalism.
  - Your behavior in general, and your betting behavior in particular, is the result of your beliefs and desires working in tandem. How do we resolve these components?

Subjectivist interpretations

- Problems:
  - You may wish to misrepresent your true opinion.
  - You may particularly enjoy or abhor gambling.
  - Like a Zen master, you may lack a desire for worldly goods altogether.

Subjectivist interpretations

- Ramsey had a more sophisticated method: attributing utilities and probabilities to you on the basis of your preferences over various options, and gambles among them.
- If these preferences are well-behaved, then you can be ‘represented’ as maximizing expected utility according to a utility function and a probability function.

Subjectivist interpretations

- This is better than the betting interpretation, but its appeal to gambles, and to preferences more generally, is still problematic.
- For all we have seen, you can also be represented in *other*, radically different ways - e.g. Voodooism.

Subjectivist interpretations

Radical subjectivists such as de Finetti recognize no constraints on initial, or ‘prior’ subjective probabilities beyond their conformity to Kolmogorov’s axioms.

Updating

- Suppose that you initially have a probability function $P_{initial}$, and that you become certain of an event E (and of nothing more). What should be your new probability function P_{new}?

Updating

Conditionalization:

$$P_{new}(X) = P_{initial}(X \mid E), \text{ if } P_{initial}(E) > 0.$$

Updating

- Problems: conditionalization apparently does not allow for loss of certainties
  - Memory loss
  - Context-sensitive (‘indexical’) information

Updating

Problem:

Radical subjectivism has been charged with being too permissive. It apparently licenses credences that we would ordinarily regard as crazy.

Updating

Reply:

Appeal to *convergence theorems:*

In the long run, the effect of choosing one prior probability function rather than another is washed out: successive conditionalizations on the evidence will, with probability one, make a given agent eventually converge to the truth.

Updating

But as Keynes' quipped, "In the long run we shall all be dead."

Some recent developments

- Further constraints on credences
  - Alignment with classical probabilities: the principle of indifference
  - Alignment with logical probabilities
  - Alignment with frequencies: calibration
  - Alignment with objective chances (perhaps propensities)

Some recent developments

- Refinements of these interpretations:
  - Generalization of the classical interpretation to ‘maximum entropy’
  - Semantic approach to logical probabilities
  - ‘Randomness’ constraints in frequentism
  - ‘Best systems’ approaches to objective chances.

Some recent developments

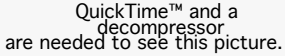
- Rivals to conditionalization?
- The ‘Sleeping Beauty problem’, and problems of ‘self-location’

Some recent developments

- Insights from the causal modeling literature (e.g. Bayesian nets, the ‘Causal Markov condition’)

Some recent developments

- Rival formal systems
  - Infinitesimal probabilities
  - Non-additive probabilities
  - Probability theories based on non-classical logics
  - Conditional probability as primitive


QuickTime™ and a
decompressor
are needed to see this picture.