Finite Frequency Theories of Probability

Philosophy of Probability
April 29th, 2013

Outline

Outline

Outline

Venn's Statement

Hájek summarizes finite frequency theories of probability by citing Venn:

> . . . *probability is nothing but that proportion*

Venn [1866].

That may seem like a precise definition, but you'll notice that discussions of "the" finite frequency theory often equivocate between several interpretations:

That may seem like a precise definition, but you'll notice that discussions of "the" finite frequency theory often equivocate between several interpretations:

- Probability is a measured proportion vs. Probability is a proportion in a wider population, some units of which have been observed.

That may seem like a precise definition, but you'll notice that discussions of "the" finite frequency theory often equivocate between several interpretations:

- Probability is a measured proportion vs. Probability is a proportion in a wider population, some units of which have been observed.
- Probability is an actual proportion vs. Probability is a hypothetical proportion

There are also a few subtle technical details that are often overlooked in discussing "the" finite frequency theory.

Venn's Theory Mathematized?

Suppose "probability *is* nothing but [a] proportion." The obvious way of turning Venn's definition into a mathematical theory of probability is as follows.

Venn's Theory Mathematized?

- Let the sample space Ω be some sample set of individuals of interest
  - E.g., All men between the ages of 20 and 65
- Let the set of events $\mathcal{A}$ be the power set $\mathcal{P}(\Omega)$

Venn's Theory Mathematized?

- Let the sample space Ω be some sample set of individuals of interest
  - E.g., All men between the ages of 20 and 65
- Let the set of events $\mathcal{A}$ be the power set $\mathcal{P}(\Omega)$
- So an event is identified with the set of individuals having some property
  - E.g. The event "getting lung cancer" is the set of men between the ages of 20 and 65 who have lung cancer.

Venn's Theory Mathematized?

- Let the sample space Ω be some sample set of individuals of interest
  - E.g., All men between the ages of 20 and 65
- Let the set of events $\mathcal{A}$ be the power set $\mathcal{P}(\Omega)$
- So an event is identified with the set of individuals having some property
  - E.g. The event "getting lung cancer" is the set of men between the ages of 20 and 65 who have lung cancer.
- The probability of an event A, then is just the proportion:

$$P(A) = \frac{|A|}{|\Omega|}$$

Is this the way Suppes formalizes "the" finite frequency theory?

Is this the way Suppes formalizes "the" finite frequency theory?
No.

For Suppes [2002], the finite frequency theory is formalized as follows:

- Let s be some finite sequence. For example:
  - Coin Flips: $s = \langle H, T, H, T, T, T, H \rangle$
  - Roll a die: $s = \langle 3, 3, 2, 1, 5, 3, 4, 6 \rangle$
  - Hair Color: $s = \langle \text{Black, Brown, Black, Blonde, Blonde, Red} \rangle$
- The range of a sequence is just its set of values:
  - Coin Flips: $\text{ran}(s) = \{H, T\}$
  - Roll a die: $\text{ran}(s) = \{1, 2, 3, 4, 5, 6\}$.
  - Hair Color: $\text{ran}(s) = \{\text{Black, Brown, Blonde, Red}\}$

- Let the sample space Ω be the range of a fixed sequence

- Let the sample space Ω be the range of a fixed sequence
- Let the set of events $\mathcal{A}$ be the power set of Ω.

Suppes' Formalization

- Let the sample space Ω be the range of a fixed sequence
- Let the set of events $\mathcal{A}$ be the power set of Ω.
- Define the probability of a set A to be the number of times it occurs in the sequence s.
- For example, if $s = \langle 3, 3, 2, 1, 5, 3, 4, 6 \rangle$, then

$$P(\{3\}) = \frac{3}{7}$$

- Moral: In Suppes [2002] formalization, the sample space Ω and the algebra are generated by an ordered object.

- Moral: In Suppes [2002] formalization, the sample space Ω and the algebra are generated by an ordered object.
- Note that, for finite sequences, reordering a sequence does not change an event's probability. For instance:
  - $P(\{3\}) = \frac{3}{7}$ if $s = \langle 3, 3, 2, 1, 5, 3, 4, 6 \rangle$,
  - $P(\{3\}) = \frac{3}{7}$ if $t = \langle 1, 2, 3, 3, 3, 4, 5, 6 \rangle$
  - And the two sequences can be obtained from each other by reordering the elements.

Three Distinctions:

- Probability is a measured proportion vs. Probability is a proportion in a wider population, some units of which have been observed.
- Probability is an actual proportion vs. Probability is a hypothetical proportion
- Algebra generated by a sequence vs. Sample space is an arbitrary unordered set.

Differences among Frequency Theories

As we discuss Hájek [1996] today, we should be careful to distinguish which interpretations of probability are undermined by each objection.

Salmon [1967] argues that an interpretation of probability ought to satisfy the following three requirements:

- **Admissibility:** An interpretation ought to satisfy probability axioms like Kolmogorov's
- **Ascertainability:** An interpretation ought to explain how probabilities can be measured
- **Applicability:** An interpretation ought to explain why probability is so use useful (especially in the sciences).

Hájek [1996] introduces a new criterion for an interpretation of probability:

> *'Probability', after all, is not just a technical term that one is free to define as one pleases. Rather, it is a concept whose analysis is answerable to our intuitions, a concept that has various associated platitudes (for example: "if X has probability greater than 0, then X can happen"). Thus, it is unlike terms like 'complete metric space' or 'Granger causation' or 'material conditional', for which there are stipulative definitions.*

In fact, Hájek [1996] argues that Salmon's three criteria are not sufficient for an adequate "philosophical" interpretation of probability.

Properties of Probability

[M]any quantities that have nothing to do with our intuitive notion of probability conform to Kolmogorov's axioms, and so in some sense provide an interpretation of them - think of mass, or length, or volume, which are clearly non-negative and additive, and which can be suitably normalized. . . . [T]his also shows that it is too glib to say that a satisfactory understanding of probability is provided as long as we find a concept of importance to science that conforms to the axioms-for that does not narrow down the field enough. In any case, the more strictly philosophical project of analysing our commonsensical concept would remain, much as the project of analysing our commonsensical concept of causation, say, would remain even if we had already done the job for the concept as it appears in science. [my emphasis]

[M]any quantities that have nothing to do with our intuitive notion of probability conform to Kolmogorov's axioms, and so in some sense provide an interpretation of them - think of mass, or length, or volume, which are clearly non-negative and additive, and which can be suitably normalized. . . . [T]his also shows that it is too glib to say that a satisfactory understanding of probability is provided as long as we find a concept of importance to science that conforms to the axioms-for that does not narrow down the field enough. In any case, the more strictly philosophical project of analysing our commonsensical concept would remain, much as the project of analysing our commonsensical concept of causation, say, would remain even if we had already done the job for the concept as it appears in science. [my emphasis]

Conforms to the axioms = Admissibility
Importance to Science = Applicability

Properties of Probability

[M]any quantities that have nothing to do with our intuitive notion of probability conform to Kolmogorov's axioms, and so in some sense provide an interpretation of them - think of mass, or length, or volume, which are clearly non-negative and additive, and which can be suitably normalized. . . . [T]his also shows that it is too glib to say that a satisfactory understanding of probability is provided as long as we find a concept of importance to science that conforms to the axioms-for that does not narrow down the field enough. In any case, the more strictly philosophical project of analysing our commonsensical concept would remain, much as the project of analysing our commonsensical concept of causation, say, would remain even if we had already done the job for the concept as it appears in science. [my emphasis]

Conforms to the axioms = Admissibility
Importance to Science = Applicability
If we assume important scientific quantities can be measured (i.e. are ascertainable) in some way, then the paragraph mentions all three of Salmon's criteria.

This argument raises two questions:

1. Why not abandon the intuitive concept of probability?
    - There are certain "intuitive notions" that we abandon for more scientific ones. For example, there are no longer any philosophers who analyze the "intuitive" concept of mass.
    - On the other hand, there is a ton of philosophy about the intuitive concepts of knowledge, causation, morality, law, ...
    - What makes a concept an object of philosophical investigation? (Its normativity?)

Question about this Argument

This argument raises two questions:

1. Why not abandon the intuitive concept of probability?
   - There are certain "intuitive notions" that we abandon for more scientific ones. For example, there are no longer any philosophers who analyze the "intuitive" concept of mass.
   - On the other hand, there is a ton of philosophy about the intuitive concepts of knowledge, causation, morality, law, ...
   - What makes a concept an object of philosophical investigation? (Its normativity?)
2. What are the other useful interpretations of the probability axioms that Hájek mentions?

Hajek argues there are many interpretations of probability that are useful in some way, so the applicability criterion is vague.

Hajek argues there are many interpretations of probability that are useful in some way, so the applicability criterion is vague.

What are his examples?

Hajek argues there are many interpretations of probability that are useful in some way, so the applicability criterion is vague.

What are his examples?

One is "Normalized length"

- Consider for a moment the property $L(x)$ which is interpreted as "length of an object x that is at most a meter long."
- If we interpret $\cup$ as placing two objects x and y end to end, then $L(x \cup y) = L(x) + L(y)$
- And clearly $0 \leq L(x) \leq 1$.

But what about the event structure?

- What is x^c?

But what about the event structure?

- What is x^c?
- Idea: The algebra contains a meter stick divided into disjoint parts.

But what about the event structure?

- What is x^c?
- Idea: The algebra contains a meter stick divided into disjoint parts.
- Length is useful, but why is $L(x)$ useful?

But what about the event structure?

- What is x^c?
- Idea: The algebra contains a meter stick divided into disjoint parts.
- Length is useful, but why is $L(x)$ useful?
- Does the event structure matter?

Break into small groups. I will assign each group four of the objections that are discussed in Suppes [2002] and Hájek [1996]. Answer the following questions.

Evaluating Finite Frequency Theories

- What is the conclusion of the argument? In particular,
  - Does the objection conclude that a finite frequency theory fails to meet one of Salmon's criteria of adequacy? If so, which criteria?
  - Is it an "intuition" based objection in the sense that it says the interpretation fails to make sense of the way that we use the word "probability"? If so, what is the intuition that is violated by a frequency theory?
  - If the objection is "intuition"-based, can it be turned into an argument that theory fails to meet one of Salmon's criteria?

Evaluating Finite Frequency Theories

- What is the conclusion of the argument? In particular,
  - Does the objection conclude that a finite frequency theory fails to meet one of Salmon's criteria of adequacy? If so, which criteria?
  - Is it an "intuition" based objection in the sense that it says the interpretation fails to make sense of the way that we use the word "probability"? If so, what is the intuition that is violated by a frequency theory?
  - If the objection is "intuition"-based, can it be turned into an argument that theory fails to meet one of Salmon's criteria?
- To which finite frequency theories (finite vs. infinite, actual vs. hypothetical, measured vs. unmeasured) is the objection directed? Which theories, if any, are not subject to the objection (or a similar one)?

References I

Hájek, A. (1996). Mises reduxRedux: fifteen arguments against finite frequentism. *Erkenntnis*, 45(2-3):209–227.

Salmon, W. C. (1967). *The Foundations of Scientific Inference*, volume 28. University of Pittsburgh Press.

Suppes, P. (2002). *Representation and Invariance of Scientific Structures*. CSLI Publications Stanford.

Venn, J. (1866). *The logic of chance: An essay on the foundations and province of the theory of probability, with especial reference to its application to moral and social science*. Macmillan.