# Lecture 10
# Popper's Propensity Theory; Hájek's Metatheory

Patrick Maher

Philosophy 517
Spring 2007

# Popper's propensity theory

## Introduction

- *One of the principal challenges confronting any objectivist theory of scientific knowledge is to provide a satisfactory understanding of physical probabilities. The earliest ideas here, known collectively as the frequency interpretation, have now been all but abandoned, and have been replaced by an equally diffuse set of proposals all calling themselves the propensity interpretation of probability.* (Miller 1994 p. 175)
- Popper is widely credited with introducing the propensity interpretation in the 1950s.
- We'll first look at what Popper meant by this and how he argued for the propensity interpretation.

## Definitions

- *The purely statistical or the frequency interpretation (I take these two designations as synonymous) . . . regards the statement*

$$p(a, b) = r$$

*as . . . asserting nothing but that the relative frequency of the event a in a sequence defined by the conditions b is equal to r . . . For example, "p(a, b) = 1/2" may mean "the relative frequency of tossing heads with a normal penny equals 1/2" (where a is getting heads upmost, and b is a sequence of tosses with a normal penny).* (1959 p. 26)
- *The propensity interpretation . . . differs from the purely statistical or frequency interpretation only in this—that it considers the probability as a characteristic property of the experimental arrangement rather than as a property of the sequence.* (1957 pp. 67–68)

# Popper's argument

**Probabilities of singular events** (1959 p. 29)

*From the point of view of the frequency interpretation, the probability of an* *<u>event of a certain kind</u>*—*such as obtaining a six with a particular die—can be* *<u>nothing but</u>* *the relative frequency of this kind of event in an extremely long (perhaps infinite) sequence of events. And if we speak of the probability of a* *<u>singular</u>* *event such as the probability of obtaining a six in the third throw made after nine o'clock this morning with this die, then, according to the purely statistical interpretation, we mean to say* *<u>only</u>* *that this third throw may be regarded as a member of a sequence of throws, and that, in its capacity as a member of this sequence, it shares in the probabilities of that sequence.*

### Objection to the frequency theory (1959 pp. 31–32)

*Let us assume that we have a loaded die, and that we have satisfied ourselves, after long sequences of experiments, that the probability of getting a six with this loaded die very nearly equals 1/4. Now consider a sequence b, say, consisting of throws with this loaded die, but including a few throws (two, or perhaps three) with a homogeneous and symmetrical die. Clearly, we shall have to say, with respect to each of these few throws with this correct die, that the probability of a six is 1/6 rather than 1/4, in spite of the fact that these throws are, according to our assumptions, members of a sequence of throws with the statistical frequency 1/4.*

*I believe that this simple objection is decisive.*

## Two sequences (1959 p. 33)

Popper introduces "*c*" as a name for a sequence of throws with the correct die. He continues:

*There are only two infinite, or very long, sequences in our case: the (actual) sequence b and the (virtual) sequence c. The throws in question belong to both of them. And our problem is this. Although they both belong to both of these sequences . . . we have no doubt whatever that in their case the proper, the true singular probability, is 1/6 rather than 1/4. Or in other words, although they belong to both sequences, we have no doubt that their singular probability is to be estimated as being equal to the frequency of the sequence c rather than b—simply because they are throws with a different (a correct) die, and because we estimate or conjecture that, in a sequence of throws with a correct die, the six will come up in 1/6 of the cases.*

## Admissible sequences (1959 p. 34)

- *All this means that the frequency theorist is forced to introduce a modification of his theory—apparently a very slight one. He will now say that an admissible sequence of events (a reference sequence, a "collective") must always be a sequence of repeated experiments.*
- *If this modification is introduced, then our problem is at once solved. For the sequence b will not be any longer an admissible reference sequence.*
- *Moreover, it seems that what I have here described as a "modification" only states explicitly an assumption which most frequency theorists (myself included) have always taken for granted.*
- *Yet, if we look more closely at this apparently slight modification, then we find that it amounts to a transition from the frequency interpretation to the propensity interpretation.*

# Critique of Popper

## Two kinds of frequency theory

Popper has a different conception of the frequency theory to the one I've been using. To prevent confusion, I'll use this terminology:

- *$f_X$ theory:* Explicatum is rf in some class or sequence of outcomes of $X$. (A frequency theory in my sense. The explicandum is $pp_X$.)
- *$f_S$ theory:* Explicatum is rf in a sequence $S$ of events, with no restriction on how these events were generated. (A frequency theory in Popper's sense. Explicandum not clearly identified.)

An $f_X$ theory is a "propensity interpretation" for Popper, and Popper himself accepts an $f_X$ theory.

> *The propensity interpretation may be presented as retaining the view that probabilities are conjectured or estimated statistical frequencies in long (actual or virtual) sequences.*
> (1959 p. 37)

## Von Mises' theory

Popper thought von Mises had an $f_S$ theory, but von Mises said:

- *The subject of probability theory is long sequences of experiments or observations repeated very often and under a set of invariable conditions.* (1964 p. 2)
- *We shall apply the term collective to a long sequence of identical observations or experiments ...* (1964 p. 10)
- *The probability of a 6 is a physical property of a given die ... Similarly, for a given pair of dice (including of course the total setup) the probability of a "double 6" is a characteristic property, a physical constant belonging to the experiment as a whole.* (1957 p. 14)

So von Mises is best interpreted as holding an $f_X$ theory.

## Claims about singular probabilities

- Popper distinguished two kinds of probabilities: (1959 p. 29)
  - *the probability of an event of a certain kind—such as obtaining a six with a particular die.*
  - *the probability of a singular event such as the probability of obtaining a six in the third throw made after nine o'clock this morning with this die.*
- His argument for the propensity interpretation was that it gives correct probabilities for singular events; he called these "singular probabilities."
- I claim that *all* pp's are relations between an experiment type and an outcome type. In that sense, I deny that there are "singular" pp's.

## Argument against singular pp's

1. *All pp's are relative to the experiment type.*
   - If $X =$ toss of this fair die, the pp of six is $1/6$.
   - If $X =$ toss of this die with exactly this force, in this position, etc., the pp of six is 0 or 1.
   - If $X =$ tossing a coin, choosing the fair die if the coin lands heads and the biased die otherwise, then the pp of six is $5/24$.
   - One event may be a result of all these experiment types.
2. *All pp's are relative to the outcome type.*
   - Suppose $X =$ toss of this fair die. The pp of six is $1/6$, the pp of an even number is $1/2$, but a particular event may instantiate both outcome types.
3. *Experiment and outcome type together determine pp.*
   - Popper assumes that all the tosses with the loaded die have the same "singular" probability, and likewise for all the tosses with the fair die. They don't vary across the "singular events," once $X$ and $O$ are fixed.

Therefore, pp's relate experiment and outcome types, not event tokens.

## What's wrong with $f_S$ theory

- Not what Popper said, since there are no singular pp's.
- The explicatum is rf in a sequence $S$ of events, with no restriction on how these events were generated. But when the elements of $S$ are not all outcomes of the same experiment, there is no corresponding explicandum.

## Example (Hájek 2003)

*I belong to the reference class consisting of myself, the Eiffel Tower, the southernmost sandcastle on Santa Monica Beach, and Mt Everest. Two of these four objects are less than 7 ft. tall . . . Yet it would be odd to say that my probability of being less than 7 ft. tall, relative to this reference class, is 1/2, even though it is perfectly acceptable (if uninteresting) to say that 1/2 of the objects in the reference class are less than 7 ft. tall.*

If the explicatum is restricted to cases where an explicandum exists, we get $f_X$ theory.

# Questions

1. What is meant by an $f_X$ theory, an $f_S$ theory, and the propensity interpretation of probability? How are these related to one another?
2. Did von Mises advocate what Popper calls the frequency interpretation of probability, or what he calls the propensity interpretation? Justify your answer.
3. Are there "singular" physical probabilities? Justify your answer.
4. Did Popper have a good argument against $f_S$ theory? If not, is there a better one? Explain.

# Hájek's metatheory

## Introduction

- As we've seen, Popper speaks of "interpretations of probability." This is common terminology.
- Hájek (2003) argues, correctly I think, that this terminology is misleading. I'll present his argument, then discuss his alternative.
- This is *metatheory of probability* because we're discussing theories of theories of probability.

## "Interpretation" is misleading (Hájek 2003, first paragraph)

- *Normally, we speak of interpreting a* <u>*formal system*</u>*, that is, attaching familiar meanings to the primitive terms in its axioms and theorems, usually with an eye to turning them into true statements about some subject of interest.*
- *However, there is no single formal system that is "probability," but rather a host of such systems. To be sure, Kolmogorov's axiomatization . . . has achieved the status of orthodoxy, and it is typically what philosophers have in mind when they think of "probability theory." Nevertheless, several of the leading "interpretations of probability" fail to satisfy all of Kolmogorov's axioms, yet they have not lost their title for that.*
- *Moreover, various other quantities that have nothing to do with probability do satisfy Kolmogorov's axioms, and thus are interpretations of it in a strict sense: normalized mass, length, area, volume . . . Nobody seriously considers these to be "interpretations of probability."*

### Hájek's alternative

*The so-called "interpretations of probability" would be better called "analyses of various concepts of probability," and "interpreting probability" is the task of providing such analyses.*

That is the second sentence of Hájek's article. You might think it deserves some explanation, but Hájek doesn't elaborate on it or refer to it again.

## Criteria of adequacy for interpretations of probability (sec. 2)

*What criteria are appropriate for assessing the cogency of a proposed interpretation of probability? Of course, an interpretation should be precise, unambiguous, and use well-understood primitives. But those are really prescriptions for good philosophizing generally; what do we want from our interpretations of* <u>*probability*</u>*, specifically? We begin by following Salmon (1966, 64), although we will raise some questions about his criteria, and propose some others. He writes:*

> <u>*Admissibility*</u>*.* [Must satisfy the probability calculus.]
>
> <u>*Ascertainability*</u>*. There [must] be some method by which, in principle at least, we can ascertain values of probabilities.*
>
> <u>*Applicability*</u>*. The force of this criterion is best expressed in Bishop Butler's famous aphorism, "Probability is the very guide of life."*

## Elaboration of the criteria (sec. 2)

- *There is no such thing as admissibility tout court, but rather admissibility with respect to this or that axiomatization.*
- *It is a little unclear in the ascertainability criterion just what "in principle" amounts to, though perhaps some latitude here is all to the good.*
- *Most of the work will be done by the applicability criterion. We must say more ... about what sort of a guide to life probability is supposed to be. Mass, length, area and volume are all useful concepts, and they are "guides to life" in various ways (think how critical distance judgments can be to survival); moreover, they are admissible and ascertainable, so presumably it is the applicability criterion that will rule them out. Perhaps it is best to think of applicability as a cluster of criteria, each of which is supposed to capture something of probability's distinctive conceptual roles; moreover, we should not require that all of them be met by a given interpretation.*

# Critique of Hájek

## Hájek's alternative terminology

*The so-called "interpretations of probability" would be better called "analyses of various concepts of probability," and "interpreting probability" is the task of providing such analyses.*

- What are the "various concepts of probability"?
  - Not concepts of ordinary language. If that was what Hájek meant, he'd surely discuss them in his article, and he doesn't.
  - What he does discuss are the different explicata (in my terminology). So the "various concepts of probability" must be the explicata.
- What are "analyses" of these "various concepts"?
  - Not evaluations or examinations of these explicata, since that isn't what interpretations of probability do.
  - Not explications of the explicata, which is senseless.

I conclude: Hájek isn't using words carefully here and has no clear account of what "interpretations of probability" really are.

## Better terminology

- The so-called "interpretations of probability" would be better called *explications of one of the ordinary language concepts of probability.*
- This would agree with Hájek if:
  - By "various concepts of probability" he meant the ordinary language concepts, ip and pp.
  - By "analyses" he meant "explications."
- I'd be happy if he meant that, but the evidence is against it.

## Hájek's method of evaluation

Hájek, following Salmon, evaluates "interpretations of probability" according to a list of criteria. But:

- Salmon and Hájek have no coherent rationale for deciding what criteria to use.

  E.g., on admissibility, Hájek says it might seem that it "goes without saying," then notes that it has no absolute meaning, and ends up suggesting that it could be violated if the interpretation "did a wonderful job of meeting the criteria of ascertainability and applicability."

- The criteria are vague; Salmon and Hájek have no principled way to make them precise.

  We've just seen that admissibility is vague; the others are worse.

- Different criteria are appropriate for different interpretations, but Salmon and Hájek have no principled basis for making this distinction.

## A better method of evaluation

- Since “interpretations of probability” are explications, the criteria for evaluating them are the ones that apply to any explication: the explicatum should be similar to the explicandum, exact, fruitful, and simple.
- Hájek’s criteria don’t need to be invoked but, insofar as they are appropriate, they follow from the fundamental criteria for an explication. (In fact, the criteria he mentions all follow from the criterion of similarity to the explicandum.) This gives a principled rationale for those criteria.
- The vagueness of Hájek’s criteria is avoided. Similarity to ip, in the respects necessary for the purposes that ip serves, is a fairly definite constraint; ditto for pp.
- The explicandum is not the same for all explications, hence the requirement of similarity to the explicandum means something different for different explications.

# Questions

5. What is wrong with the term "interpretations of probability"?
6. What should so-called "interpretations of probability" be called, according to Hájek? What, if anything, do you think Hájek meant by that? Justify your answer. What is Maher's alternative?
7. What method does Maher propose for evaluating theories of probability? Why do you think Hájek doesn't use this method? What is Hájek's alternative? What are the drawbacks of Hájek's method?

# References


- Hájek, Alan. 2003. "Interpretations of Probability," *The Stanford Encyclopedia of Philosophy*.
- Miller, David. 1994. *Critical Rationalism: A Restatement and Defense.* Open Court.
- Popper, Karl R. 1957. "The Propensity Interpretation of the Calculus of Probability, and the Quantum Theory," in S. Körner (ed.), *Observation and Interpretation*, 65–70.
- Popper, Karl R. 1959. "The Propensity Interpretation of Probability," *British Journal for the Philosophy of Science* 10, 25–42.
- Salmon, Wesley C. 1967. *The Foundations of Scientific Inference.* University of Pittsburgh Press. Cited by Hájek as 1966.
- von Mises, Richard. 1957. *Probability, Statistics and Truth.* 2nd English edition. George Allen & Unwin. Dover reprint.
- von Mises, Richard. 1964. *Mathematical Theory of Probability and Statistics.* Ed. Hilda Geiringer. Academic Press.