Lecture 9
Von Mises' Frequency Theory

Patrick Maher

Philosophy 517
Spring 2007

Introduction

- *Definition:* The **relative frequency** of an attribute in a class is the proportion of members of the class that have the attribute.
  - Example: If a coin is tossed 100 times, and lands heads 53 times, the relative frequency of heads in the class of these tosses is 0.53.
- Frequency theories are best understood as proposals for explicating pp using the concept of relative frequency.
- I'll present some simple versions of the frequency theory before turning to the classic theory of Richard von Mises.

Abbreviation: rf = relative frequency.

Observed rf

Definition

$f^o_X(O) = r$ iff, in the *observed* outcomes of X, O has an rf of r.

Example

If X = tossing a die, O = the die comes up six, and 10 tosses have been observed, of which two came up six, $f^o_X(O) = 1/5$.

The observed rf theory proposes f^o as explicatum for pp. Hurley (2006 p. 492) presents this as the frequency theory.

Some criticisms

1. f^o differs from person to person; pp doesn't.
2. f^o can be changed by making more observations; pp can't.

Actual rf

Definition

$f^a_X(O) = r$ iff, in *all actual* outcomes of X, O has an rf of r.

Example

In the preceding example, if the die is tossed 123 times altogether, and comes up six in 18 of those tosses, $f^a_X(O) = 18/123 = 0.1463$.

The actual rf theory proposes f^a as explicatum for pp.

Some criticisms

1. *f^a may not exist when pp does.*
   E.g., a coin that is never tossed.
2. *f^a may exist when pp doesn't.*
   E.g., $X =$ tossing a two-headed coin or a two-tailed coin.
3. *When both exist, their values often won't be the same.*
   E.g., a fair coin that is tossed exactly once.

Limiting rf

To avoid the divergence of explicandum and explicatum, frequency theorists often use an infinite reference class. *Problem:* Rf is not defined in an infinite class. *Solution:* Consider infinite sequences.

Definitions

- Let S be an infinite sequence of outcomes of X. Let $\text{rf}_n(O)$ be the rf of O in the first n elements of S. If $\text{rf}_n(O) \rightarrow r$ as $n \rightarrow \infty$, r is called *the limiting rf* of O in S.
- $f_X^\infty(O) = r$ iff repeating X infinitely would produce a sequence in which O has a limiting rf of r.

The limiting rf theory proposes f^∞ as explicatum for pp.

Von Mises' objection (1957 p. 23)

f^∞ can exist when pp doesn't.

E.g., a road has large stones every mile and small ones every tenth of a mile. X = observe the size of a stone, O = the stone is large. $f_X^\infty(O) = 1/10$ but $pp_X(O)$ doesn't exist.

Von Mises' theory

Von Mises modified the limiting rf theory by requiring that the sequence be *random*.

Definition (von Mises 1964 pp. 9–10)

A **place selection** is a rule for selecting a subsequence of a sequence in which the decision whether to retain the *n*th element does not depend on the value of that or any subsequent element of the sequence.

Examples

- Choosing every second element of the sequence.
- Choosing every element that follows one with outcome O.

Von Mises' idea: The random sequences are the ones in which the limiting rf of any attribute is the same in any subsequence obtained from the original sequence by a place selection.

Definition (von Mises 1964 p. 12)

A **collective** is an infinite sequence of outcomes in which each attribute has a limiting rf that is insensitive to place selections.

I.e., collectives are random sequences in which rf's converge.

Example

The observations along the road with stones isn't a collective because the limiting rf of "large" is sensitive to place selections.

Definition

$f_X^m(O) = r$ iff repeating X infinitely would produce a *collective* in which O has a limiting rf of r.

Von Mises proposed f^m as explicatum for pp.

Where we are going

- I will survey the criticisms that are often made of von Mises' theory and argue that none of them is a good criticism.
- At the end I'll give my own, different, reason for rejecting von Mises' theory.

Objections to the counterfactual

Definition

$f^j_X(O) = r$ iff there is a collective that would result if X were repeated infinitely and the limiting rf of O in that collective is r.

Jeffrey's objection (Jeffrey 1992 pp. 192–194)

- Jeffrey thought that von Mises' explicatum for pp was f^j.
- His objection was that, in cases such as tossing a fair coin, f^j doesn't exist because there isn't a collective that would result; many sequences could result.

My response

$f^m_X(O) = r$ does not imply that there is only one sequence that could be produced by repeating X infinitely. *Jeffrey's objection rests on an uncharitable interpretation of von Mises.*

Lewis's objection (Lewis 1986 p. 90)

There is no such thing as . . . the limiting frequency of heads, that would eventuate if some particular coin toss were somehow repeated forever. Rather there are . . . countless frequencies that might eventuate.

If this is right, f^m doesn't exist in most cases where pp does.

My response

1. It seems obviously correct that a casino will, over the course of a year or so, take in more money on roulette etc. than it pays out.
2. The pp that a casino will make a profit is less than 1. The pp, that a fair coin tossed forever will have a limiting rf of heads of $1/2$, equals 1.
3. Therefore, it is correct to say: If a fair coin were tossed forever, it would have a limiting rf of $1/2$. *Lewis is wrong.*

Questions

1. State the definition of f^o. Is this a good explicatum for pp? Justify your answer.
2. State the definition of f^a. Is this a good explicatum for pp? Justify your answer.
3. State the definition of f^∞. What was von Mises' objection to this as an explicatum for pp?
4. State the definitions of a *place selection*, a *collective*, and f^m.
5. What did Jeffrey take von Mises' explicatum for pp to be? What was his objection to this? Is that a good objection to von Mises' theory? Why, or why not?
6. What was Lewis's objection to f^m as an explicatum for pp? Is it a good objection? Justify your answer.

Ascertainability

- *The problem of ascertainability is the crucial problem facing the frequency theory.* (Salmon 1967 p. 84)
- It is "the very serious" objection to von Mises' theory. (Howson and Urbach 1993 p. 331)

The objection

- The values of f^m are limits of infinite sequences.
- We can only observe finitely many elements of a sequence and whatever we observe in a finite part of an infinite sequence implies nothing about the existence or value of a limit in the infinite sequence.
- Therefore, we can't know the values of f^m.

My response

1. The conclusion doesn't follow; we can have knowledge of limits by inductive reasoning.

   Example: Suppose a coin is tossed 1000 times with the rf of heads being calculated after each toss. If the rf deviates less and less from 1/2 as the number of tosses increases, then it is a reasonable inductive inference that, were the coin to be tossed forever under the same conditions, the rf of heads would approach a limit of 1/2. (von Mises 1957 pp. 14–15)

2. Observation of any finite number of outcomes also implies nothing about about the existence or value of a *pp*. Since a good explicatum must be similar to its explicandum, any adequate explicatum for pp must have this feature.

Reference class problem

This is raised against all versions of the frequency theory. "Many people consider it to be the most serious problem that frequentism faces." (Hájek 2006 p. 3)

The objection

- Suppose we want use the frequency theory to determine a probability for a particular event, Oa say.
- Let f be the frequency theory explicatum. It is assumed that we should take the probability of Oa to be $f_X(O)$ for some suitable X. The question is how to choose X, i.e., the "reference class."
- For natural choices of X, $f_X(O)$ may be unknown or not exist.
- Therefore, the frequency theory can't always provide probabilities for particular events. Yet we need such probabilities, e.g., for decision making.

Example (Venn 1888 pp. 222–223)

Let us assume, for example, that nine out of ten Englishmen are injured by residence in Madeira, but that nine out of ten consumptive persons are benefited by such a residence. These statistics, though fanciful, are conceivable and perfectly compatible. John Smith is a consumptive Englishman; are we to recommend a visit to Madeira in his case or not? In other words, what inferences are we to draw about the probability of his death? Both of the statistical tables apply to his case, but they would lead us to directly contradictory conclusions . . . Without further data, therefore, we can come to no decision.

My response

- What is being sought here is a probability for a proposition given the available evidence, i.e., an ip.
- Ip cannot in general be determined by pp, because often the required pp's either don't exist or are unknown.
- The frequency theory proposes an explicatum for pp, not ip.
- Thus the fact that the frequency theory can't determine ip's in all cases is no reason to think it fails in its purpose of explicating pp.

The objection reflects a failure to understand the purpose of frequency theories.

Reference sequence problem

In an infinite sequence of outcomes of tosses of a fair coin, there will be infinitely many heads and infinitely many tails. By reordering the sequence, the limit can be made to have any number in $[0, 1]$ as its limit, or to have no limit at all. Therefore, it is important how the sequence is ordered.

The objection (Hájek 2006 p. 9)

Such an ordering may apparently be extrinsic to the cases themselves, imposed on them from the outside. If there is no 'natural' ordering (whatever that may mean), or if there are multiple equally 'natural' orderings (whatever that may mean), the choice of ordering presumably is imposed by us. Subjectivism threatens, in virtue of the reference sequence problem (and perhaps also in the judgment of what is 'natural')—and I doubt that von Mises would have welcomed this commitment.

My response

- The specification of the ordering of the outcomes of X must be part of the definition of f_X^m. It is chosen to make f_X^m a good explicatum, in particular, similar to pp_X. This isn't a subjective motivation.
- Even if subjective factors did motivate the choice of ordering, that wouldn't make the explicatum a subjective concept. E.g., time order is an objective concept, even if it was chosen for subjective reasons.
- Hence, it is false that "subjectivism threatens."

The real problem

Carnap's second requirement for explications:

> *The characterization of the explicatum, that is, the rules of its use (for instance, in the form of a definition), is to be given in an* **exact** *form, so as to introduce the explicatum into a well-connected system of scientific concepts.* (Carnap 1950 p. 7)

The *real problem* with f^m is that it violates this condition.

- The definition of f^m appeals to what would happen if X were repeated infinitely often. This is a vague ordinary concept, not an exact one.
- Evidence of vagueness: For $X =$ tossing a fair coin, Lewis claimed the counterfactual is false, Howson and Urbach claim it is true, I used to think it is false but now think it is true.
- In fact, f^m seems less clear than pp; people who disagree about f^m don't disagree about pp. So here the explicatum is less clear than the explicandum!

Contrast with q

- In lecture 4 I specified an explicatum for pp, called q. This was done by stating postulates that relate q to an explicatum for ip, called p.
- As Carnap requires, that specification of q stated rules of use "in an exact form, so as to introduce the explicatum into a well-connected system of scientific concepts." There are no comparable rules for f^m.
- The explicit connections between q and p allowed us to give a precise account, using Bayes's theorem, of how the values of q are ascertained. Because von Mises lacked such explicit rules, he could only give vague appeals to intuitive induction.

Questions

7. Observation of finitely many outcomes of X doesn't imply anything about the existence or values of f^m_X. Why has this been thought to be an objection to f^m? Is it a good objection? Justify your answer to the latter question.
8. What is the reference class problem? Is it a serious difficulty for frequency theories of probability? Explain.
9. Is f^m a good explicatum for pp? Justify your answer.

References

- Hájek, Alan. 2006. "The Reference Class Problem is Your Problem Too." *Synthese*, forthcoming.
- Howson, Colin and Peter Urbach. 1993. *Scientific Reasoning: The Bayesian Approach*. 2nd ed. Open Court.
- Hurley, Patrick. 2006. *A Concise Introduction to Logic.* 9th ed. Wadsworth.
- Jeffrey, Richard. 1992. *Probability and the Art of Judgment.* Cambridge University Press.
- Lewis, David. 1986. *Philosophical Papers*, vol. 2. Oxford University Press.
- Salmon, Wesley C. 1967. *The Foundations of Scientific Inference*. University of Pittsburgh Press.
- Venn, John. 1888. *The Logic of Chance.* 3rd ed. Macmillan.
- 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.