Lecture 3
Explication of Inductive Probability

Patrick Maher

Philosophy 517
Spring 2007

Basic approach

The language $\mathcal{L}$

- Start by choosing a formalized language, like those studied in symbolic logic; call it $\mathcal{L}$. This will express the propositions whose inductive probability we want to explicate.
- The letters A, B, C denote sentences of $\mathcal{L}$.
- Assume $\mathcal{L}$ contains the usual truth-functional connectives: "$\sim$" for negation, "$\vee$" for disjunction, "." for conjunction, and "$\supset$" for material implication.

Examples

- "$\sim A \vee B$" says that either A is false or B is true (or both).
- "$\sim(A \vee B)$" says it is false that either A or B is true. It is equivalent to "$\sim A.\sim B$."
- "$(A.B) \supset C$" says that if A and B are true then C is true.

Classification of sentences

- Analytic: True merely due to the meanings of terms.
- Contradictory: False merely due to the meanings of terms.
- Consistent: Not contradictory.

Relations between sentences

- $A \Rightarrow B$ (A logically implies B): $A \supset B$ is analytic.
- $A \Leftrightarrow B$ (A is logically equivalent to B): $A \Rightarrow B$ and $B \Rightarrow A$.

Examples

- $A \vee {\sim}A$ is analytic.
- $A.{\sim}A$ is contradictory.
- If A is analytic then $B \Rightarrow A$, for any B.
- If A is contradictory then $A \Rightarrow B$, for any B.
- $A.B \Leftrightarrow B.A$.

Abbreviation: "ip" = inductive probability.

The function p

- Define a two-place function p which takes sentences of $\mathcal{L}$ as its arguments and has real numbers, denoted $p(A|B)$, as its values.
- The definition will consist in stating rules that determine the numbers $p(A|B)$ for all A and B in the domain of p.
- The function p will be our explicatum for ip, in the sense that $p(A|B)$ explicates the ip of A given B.
- This definition of p is chosen so as to make p a good explicatum, i.e., similar to ip, exact, fruitful, and simple.
- Similarity requires:
  - If $ip(A|B)$ has a numeric value then $p(A|B)$ has the same value.
  - If $ip(A|B) > ip(C|D)$ then $p(A|B) > p(C|D)$.
    (This can apply even if the ip's lack numeric values.)

Why sentences?

- A common choice is propositions or sets of possibilities.
- Motivation for sentences: Clearer and more fundamental; simpler than sets of possibilities.

Axioms of probability

Kolmogorov's axioms

The basic axioms of Kolmogorov (1933), adapted to sentences:

K1. $p(A) \geq 0$.

K2. $p(T) = 1$, if T is analytic.

K3. $p(A \vee B) = p(A) + p(B)$, if $A.B$ is contradictory.

Kolmogorov also adopted as a definition:

K4. $p(A|B) =_{df.} p(A.B)/p(B)$, provided $p(B) > 0$.

For sentences we also need:

K5. If $A \Leftrightarrow B$ then $p(A) = p(B)$.

Application of Kolmogorov's axioms to explication of ip

- *Method 1:* Define an unconditional probability function $p(A)$ that is an explicatum for the ip of A given no evidence; K1–K3 and K5 apply to it. Take $p(A|B)$, defined by K4, as the explicatum for $ip(A|B)$.
- *Problem:* This makes $p(A|B)$ undefined if $p(B) = 0$ but $ip(A|B)$ can exist even if $ip(B) = 0$.

  E.g., $ip(B|B) = 1$ and $ip(\sim B|B) = 0$.
- *Method 2:* Define both unconditional probability $p(\cdot)$ and conditional probability $p(\cdot|\cdot)$ directly, and relate them by the multiplication law:

  ML. $p(A.B) = p(A|B)p(B)$.
- *Problem:* Although this allows $p(A|B)$ to be defined when $p(B) = 0$, the axioms impose no constraints on the value of $p(A|B)$ when $p(B) = 0$. E.g., they allow $p(B|B) = -15$.

Axioms from Maher (2004)

We will require p to satisfy the following axioms for all sentences A, B, C, D of $\mathcal{L}$:

A1. $p(A|C) \geq 0$.

A2. $p(C|C) = 1$.

A3. $p(A|C) + p(\sim A|C) = 1$, provided C is consistent.

A4. $p(A.B|C) = p(A|C)\, p(B|A.C)$.

A5. If $A \Leftrightarrow B$ and $C \Leftrightarrow D$, then $p(A|C) = p(B|D)$.

- Here all probabilities are conditional. The ip of A given no evidence is explicated by $p(A|T)$, where T is analytic.
- We take $p(A|C)$ to be defined even if C is contradictory. *Motivation:* Fruitfulness.

Comparison

- A1 is K1 conditionalized.
- K2 conditionalized would be $p(T|C) = 1$, but that with the other axioms doesn't imply A2.
- K3 conditionalized would be

  $$p(A|C) + p(B|C) = p(A \vee B|C) \text{ if } A.B \text{ is contradictory.}$$

  A special case of that is:

  $$p(A|C) + p(\sim A|C) = p(A \vee \sim A|C).$$

  But $A \vee \sim A$ is analytic, so we can simplify to:

  $$p(A|C) + p(\sim A|C) = 1.$$

  We need to restrict this to consistent C, because the other axioms imply $p(A|C) = 1$ if C is contradictory. This gives A3.
- A4 is ML conditionalized.
- A5 is K5 conditionalized.

Justification for requiring p to satisfy A1–A5

- Quantitative inductive probabilities, where they exist, satisfy these axioms. So, since an explicatum should be similar to the explicandum, p should satisfy these axioms when the corresponding inductive probabilities have numerical values.
- The desideratum of fruitfulness justifies requiring that the same axioms hold even when the corresponding inductive probabilities lack numerical values.

Questions

1. What is the function p intended to do? What sorts of things does it have as its arguments and its values? What determines its values?
2. Describe two ways that Kolmogorov's axioms could be applied to the explication of inductive probability. For each, say whether it is adequate and, if not, why not.
3. State axioms A1–A5.
4. What is Maher's justification for requiring p to satisfy axioms A1–A5?

Propositions implied by A1–A5

P1

If $C \Rightarrow A$ then $p(A|C) = 1$.

Proof

$$\begin{aligned} p(A|C) &= p(A|C)p(C|C) && \text{by A2} \\ &= p(A|C)p(C|A.C) && \text{by A5 and } C \Rightarrow A \\ &= p(A.C|C) && \text{by A4} \\ &= p(C|C) && \text{by A5 and } C \Rightarrow A \\ &= 1 && \text{by A2} \end{aligned}$$

Corollaries

1. It T is analytic then $p(T|C) = 1$.
2. If C is contradictory then $p(A|C) = 1$.

P2

If C is consistent and $C \Rightarrow \sim A$ then $p(A|C) = 0$.

Proof

$$\begin{aligned} p(A|C) &= 1 - p(\sim A|C) && \text{by A3} \\ &= 1 - 1 && \text{by P1 and } C \supset \sim A \\ &= 0. \end{aligned}$$

P3

If C is consistent then $p(A|C) = p(A.B|C) + p(A.\sim B|C)$.

Proof

If $C \not\Rightarrow \sim A$, then $A.C$ is consistent and

$$\begin{aligned} p(A|C) &= p(A|C)[p(B|A.C) + p(\sim B|A.C)] && \text{by A3} \\ &= p(A.B|C) + p(A.\sim B|C) && \text{by A4.} \end{aligned}$$

If $C \Rightarrow \sim A$ then all quantities in P3 are zero, by P2.

P4

If C is consistent and $A.B$ is contradictory then

$$p(A \vee B|C) = p(A|C) + p(B|C).$$

Proof

$$\begin{aligned} p(A \vee B|C) &= p[(A \vee B).A|C] + p[(A \vee B).{\sim}A|C] && \text{by P3} \\ &= p(A|C) + p(B.{\sim}A|C) && \text{by A5} \\ &= p(A|C) + p(B|C) && \text{by A5 and } B \Rightarrow {\sim}A \end{aligned}$$

Significance

A1–A5 contain or imply conditionalized versions of all of K1–K5.

Unconditional	**Conditionalized**
K1	A1
K2	Corollary 1
K3	P4
ML	A4
K5	A5

Hence A1–A5 imply conditionalized versions of all the usual laws of probability.

Need more to define p

- Many different functions satisfy A1–A5.
- In fact, if C is consistent with both A and $\sim A$ then for any $r \in [0,1]$ there exists a function p that has $p(A|C) = r$ and satisfies A1–A5.
- The definition of p must be completed with further conditions.

Example

Let $C =$ a ball is either black or white, $A =$ it is white, $B =$ it is black. Neither $p(A|C)$ nor $p(B|C)$ is determined by the axioms. We could set $p(A|C) = 1/2$; then the axioms imply $p(B|C) = 1/2$.
Proof:

$$\begin{aligned} 1 &= p(C|C) && \text{by A2} \\ &= p(C.B|C) + p(C.{\sim}B|C) && \text{by P3} \\ &= p(B|C) + p(A|C) && \text{by A5} \\ &= p(B|C) + 1/2 && \text{by definition} \end{aligned}$$

Question

5. For each proposition below, prove that it follows from A1–A5 and the preceding propositions.

   P1. If $C \Rightarrow A$ then $p(A|C) = 1$.

   P2. If C is consistent and $C \Rightarrow \sim A$ then $p(A|C) = 0$.

   P3. If C is consistent then $p(A|C) = p(A.B|C) + p(A.{\sim}B|C)$.

   P4. If C is consistent and $A.B$ is inconsistent then
   $$p(A \vee B|C) = p(A|C) + p(B|C).$$

References

- Hájek, Alan. 2003. "What conditional probability could not be," *Synthese* 137, pp. 273–323. Available online with a uiuc connection. Argues that conditional probability should be taken as basic, not defined in terms of unconditional probability.
- Kolmogorov, A. N. 1933. *Grundbegriffe der Wahrscheinlichkeitrechnung*. Translated into English as *Foundations of the Theory of Probability*, Chelsea Publishing Company 1950. Available online. Kolmogorov's axioms are on p. 2 and his definition of conditional probability is on p. 6.
- Maher, Patrick. 2004. "Probability captures the logic of scientific confirmation," in *Contemporary Debates in the Philosophy of Science*, ed. Christopher R. Hitchcock, 69–93. Blackwell. On Google Books. A1–A5 are on pp. 70–71.