Lecture 11
Inductive Logic for Two Properties

Patrick Maher

Philosophy 517
Spring 2007

Introduction

The topic

- In Lecture 7 we saw:
  - A *family of properties* is a set of properties that belong to one modality (e.g., color), are mutually exclusive, and jointly exhaustive.
  - Carnap's explication of ip for an $\mathcal{L}$ whose primitive predicates denote the elements of one family of properties.
- It is important to be able to deal with properties from different modalities; e.g., "black raven" involves two modalities.
- Today I'll discuss the simplest case of this sort, where $\mathcal{L}$ contains just two primitive predicates but they denote properties from different modalities.

Notation and terminology

Predicates and partitions

- The two primitive predicates of $\mathcal{L}$ are denoted F_1^1 and F_1^2. Superscript indicates modality, subscript indicates property.

  Example: $F_1^1 = \text{raven}$, $F_1^2 = \text{black}$.
- F_2^i designates the complement of F_1^i.

  Example: $F_2^1 = \text{non-raven}$, $F_2^2 = \text{non-black}$.
- F_{lm} is the conjunction of F_l^1 and F_m^2.

  Example: $F_{11} = \text{black raven}$, $F_{12} = \text{non-black raven}$.
- A *partition of predicates* is a set of predicates that are mutually exclusive and jointly exhaustive. Notation:
  - $\mathcal{F}^1 = \{F_1^1, F_2^1\}$
  - $\mathcal{F}^2 = \{F_1^2, F_2^2\}$
  - $\mathcal{F}^{12} = \{F_{11}, F_{12}, F_{21}, F_{22}\}$.
- The partitions $\mathcal{F}^1$ and $\mathcal{F}^2$ each designate a family of properties (in Carnap's sense). $\mathcal{F}^{12}$ doesn't because it combines two modalities.

Sample descriptions

A *sample description with respect to partition* $\mathcal{F}$ is a conjunction of atomic sentences, each of which ascribes a predicate in $\mathcal{F}$ to a different individual.

Examples

Sample description	with respect to
$F_{11}a_1.F_{12}a_2$	$\mathcal{F}^{12}$
$F_1^1a_1.F_1^1a_2$	$\mathcal{F}^1$
$F_1^2a_1.F_2^2a_2$	$\mathcal{F}^2$

Notation:

- S: A sample description with respect to $\mathcal{F}^{12}$.
- S^1: The corresponding sample description with respect to $\mathcal{F}^1$.
- S^2: The corresponding sample description with respect to $\mathcal{F}^2$.

The problem

- We want conditions that will fix the value of $p(A|B)$, for all sentences A and B in $\mathcal{L}$, in such a way that p is a good explicatum for $ip(A|B)$.
- It suffices to fix the value of $p(F_{lm}a|S)$, for all sample descriptions S not containing a, where $l, m = 1, 2$.
- We'll now consider three methods for doing this.

MI: Method of independence

The method

$\mathcal{F}^1$ and $\mathcal{F}^2$ are independent in p and Carnap's $\lambda\gamma$ theorem applies to each of them separately. So:

$$\begin{aligned} p(F_{lm}a|S) &= p(F_l^1 a|S^1)\, p(F_m^2 a|S^2) \\ &= \frac{n_l^1 + \lambda\gamma_l^1}{n+\lambda}\, \frac{n_m^2 + \lambda\gamma_m^2}{n+\lambda}. \end{aligned}$$

Here:

$n =$ size of the sample described by S;

$n_l^i =$ number of individuals that S says have F_l^i;

$\gamma_l^i =$ the γ value for F_l^i.

Objection

Suppose S says $n/2$ individuals have F_{11} and $n/2$ have F_{22}. Then, using MI,

$$\begin{aligned} p(F_{12}a|S) &= \frac{n/2+\lambda\gamma_1^1}{n+\lambda}\,\frac{n/2+\lambda\gamma_2^2}{n+\lambda} \\ &\to \frac{1}{4} \text{ as } n\to\infty. \end{aligned}$$

But $ip(F_{12}a|S) \to 0$ in this example.

MC: Method of combination

The method

Carnap's $\lambda\gamma$ theorem applies to $\mathcal{F}^{12}$. Thus:

$$p(F_{lm}a|S) = \frac{n_{lm} + \lambda\gamma_{lm}}{n + \lambda},$$

where:

$$n_{lm} = \text{number of individuals that } S \text{ says have } F_{lm};$$

$$\gamma_{lm} = \text{the } \gamma \text{ value for } F_{lm}.$$

MC avoids the objection to MI

If S says $n/2$ individuals have F_{11} and $n/2$ have F_{22}, MC gives:

$$\begin{aligned} p(F_{12}a|S) &= \frac{\lambda\gamma_{12}}{n + \lambda} \\ &\to 0 \text{ as } n \to \infty. \end{aligned}$$

Objection

Using MC,

$$p(F_{12}b|F_{11}a.S) = \frac{n_{12} + \lambda\gamma_{12}}{n + 1 + \lambda} < \frac{n_{12} + \lambda\gamma_{12}}{n + \lambda} = p(F_{12}b|S).$$

But, under some circumstances,

$$ip(F_{12}b|F_{11}a.S) > ip(F_{12}b|S).$$

Example

Let F_1^1 = unicorn, F_1^2 = white. Given what we know, it is very improbable that any unicorns exist, but if a white unicorn were discovered, that would raise the probability that non-white unicorns also exist, and hence that an unobserved individual is a non-white unicorn.

MM: Mixed method

Definition

$\mathcal{F}^1$ and $\mathcal{F}^2$ are *statistically independent* if they are uncorrelated in the population, i.e., the proportion of individuals that have F_{11} is just the proportion that have F_1^1 times the proportion that have F_1^2.

Example

Let the individuals be days.

- If F_1^1 = Sunday, F_1^2 = rainy, then $\mathcal{F}^1$ and $\mathcal{F}^2$ are statistically independent.
- If F_1^1 = day in April, F_1^2 = rainy, then $\mathcal{F}^1$ and $\mathcal{F}^2$ are statistically dependent (I suppose).

Notation: "I" means $\mathcal{F}^1$ and $\mathcal{F}^2$ are statistically independent.

The method MM

MI is right when I is given and MC is right when $\sim I$ is given. So:

$$\begin{aligned} p(F_{lm}a|S.I) &= \frac{n_l^1 + \lambda\gamma_l^1}{n+\lambda}\,\frac{n_m^2 + \lambda\gamma_m^2}{n+\lambda} \\ p(F_{lm}a|S.{\sim}I) &= \frac{n_{lm} + \lambda\gamma_{lm}}{n+\lambda}. \end{aligned}$$

Also, $\gamma_{lm} = \gamma_l^1\gamma_m^2$ and $0 < p(I) < 1$.

By the law of total probability:

$$p(F_{lm}a|S) = p(F_{lm}a|S.I)\,p(I|S) + p(F_{lm}a|S.{\sim}I)\,p({\sim}I|S).$$

MM avoids the objection to MI

Suppose S says $n/2$ individuals have F_{11} and $n/2$ have F_{22}. With MM,

$$p(F_{12}a|S) = p(F_{12}a|S.I)\, p(I|S) + p(F_{12}a|S.{\sim}I)p({\sim}I|S).$$

But S becomes conclusive evidence for ${\sim}I$ as $n \to \infty$. Thus, as $n \to \infty$, $p(I|S) \to 0$ and

$$p(F_{12}a|S) \;\to\; p(F_{12}a|S.{\sim}I) = \frac{\lambda\gamma_{12}}{n+\lambda} \to 0.$$

MM avoids the objection to MC

With MM, if γ_1^1 is sufficiently small then

$$p(F_{12}b|F_{11}a) > p(F_{12}b).$$

Example

Let $\gamma_1^1 = 0.001$, $\gamma_1^2 = 0.1$, $\lambda = 2$, $p(I) = 1/2$. Then

$$p(F_{12}b|F_{11}a) = 0.1005 > 0.0009 = p(F_{12}b).$$

The left side is more than 100 times larger than the right!

(The appendix shows how these numbers are obtained.)

Questions

1. Describe method MI for explicating ip for two properties. Is it a good method? Justify your answer.
2. Describe method MC for explicating ip for two properties. Is it a good method? Justify your answer.
3. Describe method MM for explicating ip for two properties.
4. Is MM open to the objection against MI that was raised in class? Justify your answer.
5. Is MM open to the objection against MC that was raised in class? Justify your answer, without proofs.

AA: Carnap's axiom of analogy

Definition

Let ϕ_1 and ϕ_2 be predicates in $\mathcal{F}^{12}$.
$d(\phi_1, \phi_2)$ = the number of indices on which ϕ_1 and ϕ_2 differ.

Examples

$d(F_{11}, F_{12}) = 1; \quad d(F_{11}, F_{22}) = 2.$

Carnap's axiom of analogy (Carnap 1975, p. 320)

If $d(\phi_1, \phi_2) = 1$ and $d(\phi_1, \phi_3) = 2$ then

$$p(\phi_1 b | \phi_2 a.S) > p(\phi_1 b | \phi_3 a.S).$$

MM violates AA (Maher 2000 p. 72)

Example: If $n_{11} = n_{22} = 4$, $n_{12} = n_{21} = 0$, $\gamma_I^i = 1/2$, $\lambda = 2$, and $p(I) = 1/2$, then using MM:

$$p(F_{11}b|F_{12}a.S) = 0.394 < 0.407 = p(F_{11}b|F_{22}a.S).$$

Is MM defective? Or is AA not an appropriate requirement?

Analysis of the example (Maher 2000 p. 72)

$$p(F_{11}b|F_{12}a.S) = 0.394 < 0.407 = p(F_{11}b|F_{22}a.S)$$

$$p(F_{11}b|F_{12}a.S.I) = 0.248 > 0.207 = p(F_{11}b|F_{22}a.S.I)$$

$$p(F_{11}b|F_{12}a.S.{\sim}I) = 0.409 = p(F_{11}b|F_{22}a.S.{\sim}I)$$

$$p(I|F_{12}a.S) = 0.096 > 0.012 = p(I|F_{22}a.S)$$

So MM violates AA here because, given S:

1. I is more probable given $F_{12}a$ than given $F_{22}a$.
2. Raising the probability of I lowers the probability of $F_{11}b$; the pattern exhibited in S starts to look more like a coincidence.

Conclusion

- AA is based on the intuition that if ϕ_1 is more similar to ϕ_2 than to ϕ_3 then $ip(\phi_1 b|\phi_2 a.S) > ip(\phi_1 b|\phi_3 a.S)$.
- That intuition doesn't take account of the fact that learning $\phi_2 a$ or $\phi_3 a$ may also give us information about whether this similarity is statistically relevant.
- Therefore, we should reject AA.

Using MM has here given us a more sophisticated understanding of ip than we could have obtained by reasoning about that vague concept directly. That is the characteristic of a good explicatum! (Cf. lecture 2.)

AC: The axiom of convergence

This is another axiom endorsed by Carnap (1963, p. 976); basically, it says that probabilities should converge to observed relative frequencies in the long run. Here we'll focus on its application to $\mathcal{F}^{12}$, in which case it may be stated as follows.

The axiom

Let S_1, S_2, . . . , be an infinite sequence of sample descriptions with respect to $\mathcal{F}^{12}$, where each S_n is for a sample of size n and S_{n+1} entails S_n. Let n_{lm} denote the number of individuals that have F_{lm} according to S_n. Then, for a not involved in any of the S_n,

$$\lim_{n\to\infty} \left| p(F_{lm}a|S_n) - \frac{n_{lm}}{n} \right| = 0.$$

MI violates AC

The objection made earlier against MI was essentially that it violates AC. Restated in present terminology: Let S_n say that all n individuals are F_{11} or F_{22}, with (as near as possible) half being each. Then, using MI:

$$\lim_{n\to\infty}\left|p(F_{12}a|S_n)-\frac{n_{12}}{n}\right| = \lim_{n\to\infty}\left|\frac{n_1^1+\lambda\gamma_1^1}{n+\lambda}\,\frac{n_2^2+\lambda\gamma_2^2}{n+\lambda}-\frac{n_{12}}{n}\right|$$
$$=\left|\frac{1}{2}\,\frac{1}{2}-0\right|=\frac{1}{4}\neq 0.$$

MC satisfies AC

$$\lim_{n\to\infty}\left|p(F_{12}a|S_n)-\frac{n_{12}}{n}\right| = \lim_{n\to\infty}\left|\frac{n_{lm}+\lambda\gamma_{lm}}{n+\lambda}-\frac{n_{lm}}{n}\right|=0.$$

MM satisfies AC

- MM is a mixture of MI and MC.
- MC satisfies AC, as we've just seen.
- MI does not in general satisfy AC. When it doesn't,

$$\left| \frac{n_l^1}{n} \frac{n_m^2}{n} - \frac{n_{lm}}{n} \right| \not\to 0 \text{ as } n \to \infty.$$

  In that case, $p(I|S_n) \to 0$ as $n \to \infty$. But $p(I|S_n)$ is the weight on the MI component. Therefore, when the MI component doesn't satisfy AC, the weight on it becomes zero.
- Therefore, MM satisfies AC.

A different proof is given in Maher (2000).

Conclusion: So far, MM has survived all challenges!

Questions

6. State Carnap's axiom of analogy. Does MM satisfy it? Should an explicatum for ip satisfy it? Justify your answer to the latter question.
7. For each of the following, say whether it satisfies the axiom of convergence and prove that your answer is correct.
   (a) MI.
   (b) MC.
   (c) MM.

Appendix

Here I show how to obtain the numbers in the example that showed MM avoids the objection to MC. Let's start with:

Prior probability

$$
\begin{aligned}
p(F_{12}b) &= p(F_{12}b|I)\,p(I) + p(F_{12}b|{\sim}I)p({\sim}I) \\
&= \gamma_1^1\gamma_2^2 p(I) + \gamma_{12}p({\sim}I), \text{ by MM} \\
&= \gamma_1^1\gamma_2^2 p(I) + \gamma_1^1\gamma_2^2 p({\sim}I), \text{ by MM} \\
&= \gamma_1^1\gamma_2^2 \\
&= \gamma_1^1(1-\gamma_1^2) \\
&= (0.001)(1-0.1) \\
&= 0.0009.
\end{aligned}
$$

Before calculating the posterior probability we need:

Probability of I

$$
\begin{aligned}
p(I|F_{lm}a) &= \frac{p(F_{lm}a|I)p(I)}{p(F_{lm}a|I)p(I) + p(F_{lm}a|\sim I)p(\sim I)}, \text{ by Bayes' theorem} \\
&= \frac{\gamma_l^1 \gamma_m^2 \, p(I)}{\gamma_l^1 \gamma_m^2 \, p(I) + \gamma_{lm} \, p(\sim I)}, \text{ by MM} \\
&= \frac{p(I)}{p(I) + p(\sim I)}, \text{ by MM} \\
&= p(I).
\end{aligned}
$$

Finally, we calculate:

Posterior probability

$$
\begin{aligned}
& p(F_{12}b|F_{11}a) \\
&= p(F_{12}b|F_{11}a.I)\,p(I|F_{11}a) + p(F_{12}b|F_{11}a.{\sim}I)\,p({\sim}I|F_{11}a) \\
&= p(F_{12}b|F_{11}a.I)\,p(I) + p(F_{12}b|F_{11}a.{\sim}I)\,p({\sim}I), \\
&\quad \text{by the previous result} \\
&= \frac{1+\lambda\gamma_1^1}{1+\lambda}\,\frac{\lambda\gamma_2^2}{1+\lambda}\,p(I) + \frac{\lambda\gamma_{12}}{1+\lambda}\,p({\sim}I), \text{ by MM} \\
&= \frac{1+2(.001)}{1+2}\,\frac{2(.9)}{1+2}\,\frac{1}{2} + \frac{2(.0009)}{1+2}\,\frac{1}{2} \\
&= 0.1005.
\end{aligned}
$$

- Carnap, Rudolf. 1963. "Replies and systematic expositions," in Paul Arthur Schilpp (ed)., *The Philosophy of Rudolf Carnap*, 859–1013. Open Court.
- Carnap, Rudolf. 1975. "Notes on Probability and Induction," in Jaakko Hintikka (ed.), *Rudolf Carnap, Logical Empiricist*, 293–324. D. Reidel. The last section discusses inductive logic for two families of properties.

- Maher, Patrick. 1999. "Inductive Logic and the Ravens Paradox," *Philosophy of Science* 66, 50–70. Applied MC to the ravens paradox.
- Maher, Patrick. 2000. "Probabilities for Two Properties," *Erkenntnis* 52, 63–91. Derived MM from axioms; criticized alternatives.
- Maher, Patrick. 2001. "Probabilities for Multiple Properties: The Models of Hesse and Carnap and Kemeny," *Erkenntnis* 55, 183–216. Explored how to deal with more than two properties; did not find a satisfactory model.
- Maher, Patrick. 2004. "Probability Captures the Logic of Scientific Confirmation," in Christopher Hitchcock (ed.), *Contemporary Debates in the Philosophy of Science*, 69–93. Blackwell. First publication of mine to use the methodology of explication. Applied MM to the ravens and grue paradoxes.