Induction: an Afterthought

Uwe Saint-Mont

Funding: No specific funding was received for this work.

Potential competing interests: No potential competing interests to declare.

Abstract

The aim of Saint-Mont (2020) was to provide a rather general answer to Hume's problem. To this end, induction was treated within a straightforward formal paradigm, i.e., several connected tiers of abstraction.

This note reduces the latter argument to its very essence – a simple fourfold table. Moreover, it points out that a sound logical solution to the problem of induction boils down to proper accounting.

Uwe Saint-Mont[*]

Nordhausen University of Applied Sciences

[*] *Prof. U. Saint-Mont, Fachbereich Wirtschafts- und Sozialwissenschaften, Hochschule Nordhausen, Weinberghof 4, D-99734 Nordhausen; saint-mont@hs-nordhausen.de; ORCID: 0000-0001-6801-3658.*

Keywords: Hume's Problem; Induction; Circular Reasoning; Hierarchical Structures; Inference.

1. Introduction

Considering Hume's probem, the basic model introduced in Saint-Mont (2020) con- sists of two levels of abstraction, with the more general tier on top containing more information than the more specific tier at the bottom:

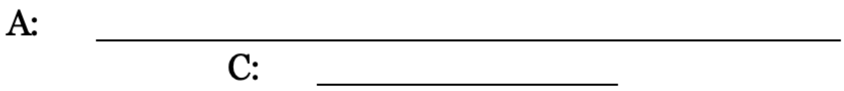


Fig. 1. Basic model with two tiers. *A* (more general), and *C* (less general).

Moving downwards, deduction skips some of the information. Moving upwards, induc- tion leaps from 'less to more'. Given several tiers, the latter article proceeds to figures 6 and 7, i.e., hierarchical and circular structures, in particular

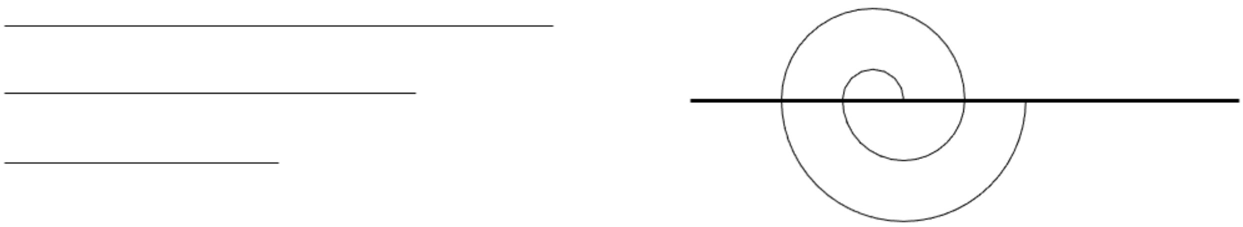
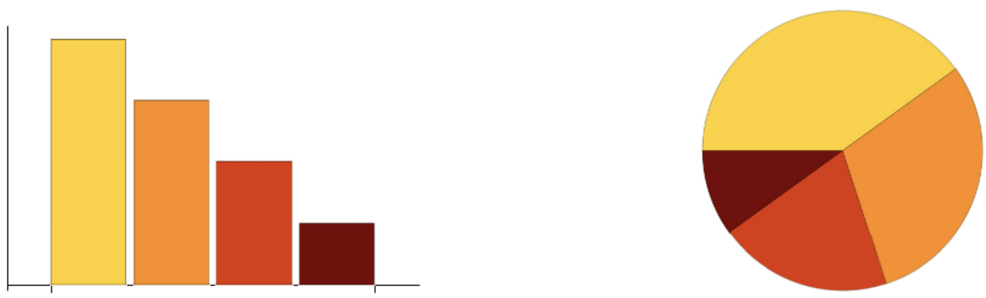


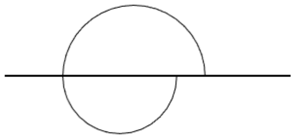
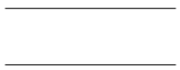
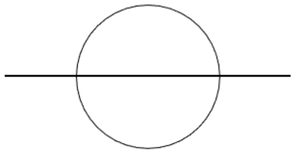
Fig. 2. A hierarchical and a partially recursive structure

2. A fourfold table

The crucial observation now is that in the above representations curvature is not asso- ciated with any kind of meaning. Therefore, the models in Fig. 2 are in fact equivalent – just as a bar chart and a pie chart visualizing the same numbers.



What is relevant, however, is the size $d(A, C)$ of an inductive step, i.e., the distance or difference between two successive tiers. Considering just two tiers, we obtain the following table:

1.	$d > 0$ $\Downarrow$	hierarchy 	spiral 
2.	$d = 0$ $\Updownarrow$	two lines of equal length 	circle 

The first line of the table says that if $d(A, C) > 0$, there are two lines of different length, which are equivalent to a spiral 'in curved terms'. Moreover, the arrow $\Downarrow$ indicates the step from 'more to less', i.e., the implication $A \Rightarrow C$.

The second line of the table considers the (degenerate) case that $d(A, C) = 0$. Thus the two lines have exactly the same length, meaning that A and C are equivalent ($A \Leftrightarrow C$). Moreover, 'in curved terms' the spiral becomes a circle.

Although rather elementary, the last table explains why Hume's idea that any inductive argument in favour of induction is supposed to be viciously circular is wrong. The problem of induction - being the inverse of deduction - is located on the first line. Worthless circular reasoning, however, is located on the second line.

Unfortunately, imprecise verbal reasoning may easily blur this fundamental distinction. For instance, the Stanford Encyclopedia of Philosophy writes on 'The Problem of In- duction' (Nov. 22, 2022) that "an argument for a principle may not presuppose the same principle (Non-circularity)". However, it is not clear if the latter sentence refers to the first or the second line: 'non-circularity' points toward an implication, whereas 'the same' rather indicates an equivalence.

If an argument is tautological, we are discussing an equivalence, and thus we find ourselves in the second line of the last table. If, on the other hand, we consider the problem of induction, where necessarily $d(A, C) > 0$, we are talking about the first line. Therefore, for a mathematician, the problem of induction typically boils down to giving some condition B, such that C and B imply A.

For instance, every square S is a rectangle R, thus $S \Rightarrow R$. Given a rectangle, the latter geometrical shape must have at least three sides of equal length in order to qualify as a square. Equivalently, as pointed out in the second footnote of Saint-Mont

(2020), the conditions that define a square are stronger (e.g., more numerous) than those describing a rectangle.

Given this perspective, all is well and good: A (closed) perfectly recursive circle is a nice model for an equivalence, and an (open), only partially recursive spiral is an equally fitting model for a hierarchical relation such as an implication. However, mixing up the lines or the four logical possibilities of the last table causes havoc, since identifying an inductive step with a perfectly circular argument means blending the logical functions $\Rightarrow$ and $\Leftrightarrow$. Solving the problem of induction thus becomes a hopeless endeavour, and in doing so, Hume established a conundrum that has confused generations of philosophers.

3. A (de)finite answer

The upshot of the ‘new riddle of induction’ consists in considering rather arbitrary generalizations. Therefore, in a nutshell, Goodman (1983) refers to the case $d(A, C) = \infty$. Given our point of view, he is correct in mistrusting such ‘unbounded’ inductive steps – e.g., the paradigmatic (notorious?) emerald changing its colour at any moment in time.

Thus we should consider the finite case $0 < d < \infty$ in more detail. The crucial insight essentially goes back to R.T. Cox (1946) and E.T. Jaynes (2003), who showed that any kind of inductive logic, proceeding from the particular to the general, and which there- fore has to consider degrees of certainty, is equivalent to probability theory. In other words, any ‘reasonable’ extension of classical logic leads to the axioms of probability theory (see Clayton and Waddington (2017), and Saint-Mont (2011), chap. 4.4, for more details). As a corollary, probability theory should give a general formal solution to Hume’s problem.

To this end, suppose there is some current state of knowledge $I(C)$ plus additional information $I(B)$, both finite, at your fingertips. Then the total information you should have is $I(C) + I(B|C) = I(B, C)$. The latter equation determines the unique and thus ‘correct distance’ between the more informative layer A, consisting of B and C, and the less informative layer consisting of C only, i.e., $d(A, C) = I(B, C) - I(C) = I(B|C)$.

With the usual definition of information $I(p) = \log(1/p)$, exponentiation gives an equivalent expression in terms of probabilities, see equation (2) in Saint-Mont (2020),

$$I(C) + I(B|C) = I(B, C) \Leftrightarrow p(C) \cdot p(B|C) = p(B, C). \quad (1)$$

The last equation remains true if one takes expectations, which yields the chain rule of entropy, $H(B, C) = H(C) + H(B|C)$, see Cover and Thomas (2006), pp. 17f, 22. Moreover, notice that the intersection $A = B \cap C$ is smaller than the sets it consists of. However, its corresponding amount of information $I(A) = -\log(p(A))$ is larger, and the above figures and tables should be interpreted in that way.

Actually, equation (1) is just a formal way of saying that it would be unreasonable if information – or equivalently, (probability) mass – either appeared out of nowhere or vanished into thin air. Rather, honest book-keeping must be sound (coherent) and not allow for such pathologies.

This is also the essence of the third axiom of probability theory, $p(B \cup C) = p(B) + p(C)$, if B and C are mutually exclusive events. Therefore, it is the foundation of equation (1), and it is also crucial in avoiding so-called ‘Dutch book arguments’, see Greenland (1998).

In sum, R.A. Fisher (1966), p. 4, was absolutely right when he concluded:

> *We may at once admit that any inference from the particular to the general must be attended with some degree of uncertainty, but this is not the same as to admit that such inference cannot be absolutely rigorous, for the nature and degree of the uncertainty may itself be capable of rigorous expression.*

> *In the theory of probability.. we have the classic example proving this possibility.. The mere fact that inductive inferences are uncertain can- not, therefore, be accepted as precluding perfectly rigorous and unequivocal inference.*

Information and probability theory give a logically sound answer to Hume's problem.

Statements and Declarations

Conflict of Interest: The author states that there is no conflict of interest.

Ethical Approval: not applicable

Funding: none

Availability of data and materials: not applicable

References

- Clayton, A.; and T. Waddington (2017). Bridging the intuition gap in Cox's theorem: A Jaynesian argument for universality. *International J. of Approximate Reasoning* **80**, 36-51.
- Cover, T.M.; and Thomas, J.A. (2006). Elements of Information Theory. (2nd ed.) New York: Wiley.
- Cox, R.T. (1946). Probability, Frequency, and Reasonable Expectation. *American J. of Physics* **14**, 1-13.
- Fisher, R.A. (1966). The Design of Experiments. (8th ed.). New York: Hafner Publish- ing Company.
- Goodman, N. (1983). Fact, fiction, and forecast. Harvard University Press.
- Greenland, S. (1998). Probability Logic and Probabilistic Induction. *Epidemiology*
- **9(3)**, 322-332.
- Jaynes, E.T. (2003). Probability Theory. The Logic of Science. Cambridge: Cambridge University Press.
- Saint-Mont, U. (2011). Statistik im Forschungsprozess. Heidelberg: Physika.
- Saint-Mont, U. (2020). Induction: a logical analysis. Foundations of Science https://doi.org/10.1007/s10699-020-09683-z