Jake Chandler

Department of Philosophy, University of Glasgow,
67-69 Oakfield Avenue, Glasgow G12 8QQ
✉ J.Chandler@philosophy.arts.gla.ac.uk

UNIVERSITY of GLASGOW

[2] Chance

0. Outline

1. **Probability and Inference: a primer (ctd)**
2. **‘Too improbable given chance alone’**

1. Probability and inference: a primer (ctd.)

- In the last lecture we saw that a probability function is a function from the elements of a certain set of propositions to the real numbers, such that:

  [P1] $0 \leq \Pr(P)$

  [P2] If P is a tautology, then $\Pr(P) = 1$

  [P3] If $P\&Q$ is a contradiction, then $\Pr(P \vee Q) = \Pr(P) + \Pr(Q)$
- We also provided the following definition of a conditional probability:

  [P4] if $\Pr(Q) > 0$, then $\Pr(P|Q) = \dfrac{\Pr(P \& Q)}{\Pr(Q)}$

1. Probability and inference: a primer (ctd.)

- Finally, we noted some standard terminology used in the context of scientific inference (where H denotes a hypothesis and E a body of evidence):
  - $\Pr(E|H) =_{def}$ the ‘likelihood’ of H,
  - $\Pr(E)$ and $\Pr(H) =_{def}$ the ‘prior probabilities’ of E and H, respectively,
  - $\Pr(H|E) =_{def}$ the ‘posterior probability’ of H.
- Now [P1] to [P4] above can be used to derive a very useful relationship between the posterior probability of a hypothesis, its prior probability and its likelihood…
- This relationship is known as *Bayes’ Theorem*…

1. Probability and inference: a primer

- There are a number of variants on this theorem:

  Bayes' Theorem (I): if Pr(*H*) > 0 and Pr(*E*) > 0, then

$$\Pr(H|E) = \frac{\Pr(E|H)\Pr(H)}{\Pr(E)}$$

  Or equivalently:

  Bayes' Theorem (II): if 1 > Pr(*H*) > 0, then

$$\Pr(H|E) = \frac{\Pr(E|H)\Pr(H)}{\Pr(E|H)\Pr(H) + \Pr(E|\neg H)\Pr(\neg H)}$$

- Bayes' Theorem essentially tells us how to revise our prior assessments of the probability of a hypothesis being true in the light of new data.

1. Probability and inference: a primer

- An example: screening for breast cancer.

  1% of women at age forty who participate in routine screening have breast cancer.

  80% of women with breast cancer will get positive mammographies.

  9.6% of women without breast cancer will also get positive mammographies.

  A woman in this age group had a positive mammography in a routine screening.

  What is the probability that she actually has breast cancer?

1. Probability and inference: a primer

- Let: *C* stand for breast cancer, *P* stand for a positive mammography.
- We know that:
  - $\Pr(C) = 0.01$, and hence $\Pr(\neg C) = 0.99$
  - $\Pr(P|C) = 0.8$
  - $\Pr(P|\neg C) = 0.096$
- We are after $\Pr(C|P)$. By *Bayes' Theorem (II)*:

$$\Pr(C|P) = \frac{\Pr(P|C)\Pr(C)}{\Pr(P|C)\Pr(C) + \Pr(P|\neg C)\Pr(\neg C)}$$

$$= \frac{0.8 \times 0.01}{0.8 \times 0.01 + 0.096 \times 0.99} \approx 0.078$$

1. Probability and inference: a primer

- The use of Bayes' theorem leaves us with a probability of a hypothesis being true given the evidence.

  Example: it is 99.88% likely that Mrs Jones doesn't have breast cancer given her negative test results.
- But it is commonplace - both in science and everyday life - to move from this kind of probability statement to a *categorical* rejection/acceptance.

  Example: we might safely assume from Mrs Jones' test results that she doesn't have breast cancer.
- To the extent that this kind of move is indeed legitimate (some indeed argue that it never *is* legitimate), what is the connection?

1. Probability and inference: a primer

- One possibility (amongst a fair few others):

  COMP: upon acquiring evidence E, accept hypothesis H iff $\Pr(H \mid E) > \Pr(H_i \mid E)$ for all competing hypotheses H_i and reject H otherwise (i.e. accept a hypothesis iff your data makes it more probable than its competitors, otherwise reject it).
- But there are some potential problems here, most notably the infamous *Lottery Paradox*.
- There is a *huge* literature on this. For further references, see Wheeler [2007] (advanced).
- With this in place, back to biology…

2. 'Too improbable given chance alone'

- In many discussions of the origins of the living world, the default hypothesis (invariably rejected) is that the organisms we observe were produced by chance alone.
- What people seem to have in mind when they talk of 'being produced by chance alone':

  An outcome O is the product of chance alone iff it was the product of a process such that O and each of its alternatives had an equal chance of being produced (e.g. the roll of a fair dice, the toss of a fair coin, etc…).
- Of course, given the complexity of the kinds of organisms that we currently observe, the space of possible alternatives is *vast* and the probability of the genesis of organisms as we know them conditional on chance alone is therefore vanishingly low.

2. 'Too improbable given chance alone'

- An *extremely common* kind of statement (echoed by Dawkins [1976], Dembski [1998] and many, many others).

  The probability of there currently existing organisms having the properties they do given chance alone is so low that these organisms cannot be the products of chance alone.

- I.e., where *O* stands for the fact that there currently exist organisms having the properties they do and *CH* stands for the hypothesis that these organisms were the product of chance alone:

  We observe that *O*. $\Pr(O \mid CH)$ is very low. Hence *CH* must be rejected.

2. 'Too improbable given chance alone'

- In spite of its huge popularity, however, the '$\Pr(O \mid CH) =$ low hence reject *CH*' argument doesn't work out…

Next lecture: 'Chance' (ctd)

- No reading.

Reference

- Dawkins R. [1976]: *The Selfish Gene*. Oxford: OUP.
- Dembski, W. [1998]: *The Design Inference: eliminating chance through small probabilities*. Cambridge: CUP.
- Wheeler, G. [2007]: 'A Review of the Lottery Paradox', in W. Harper and G. Wheeler (eds.) *Probability and Inference: Essays in the Honour of H. E. Kyburg Jnr*, London: King's College Publications, pp 1-31.