Bayesian Confirmation Theory, part 2

- According to BCT,
  1. scientists have certain degrees of beliefs about how the world is—which hypotheses are true and what evidence we will collect—and we can represent these degrees of belief, or credences, with a function $C$ from propositions to numbers in the interval $[0, 1]$.
  2. $C$ ought to be a probability function.
  3. scientists ought to update their credence function by conditionalizing on the evidence they acquire. That is, if they learn that $e$, then they ought to transition from their prior credence function $C(\cdot)$ to a posterior credence function $C^+(\cdot) = C(\cdot | e)$.

$$\begin{align*}
\text{prior} & \xrightarrow{e} \text{posterior} \\
C & \mapsto C^+ \\
C^+(\cdot) & \equiv C(\cdot | e)
\end{align*}$$

- We should do this for every piece of evidence we acquire. So, our credence function at any time $t$ should be our original credence function $C$ conditionalized upon all of our knowledge at time $t$, $k_t$.

$$C_t(\cdot) \equiv C(\cdot | k_t)$$

BCT and The Paradox of the Ravens

- Hempel’s paradox of the ravens is that, if we accept the following two principles,
  - Universal generalizations $(\forall x)(Fx \rightarrow Gx)$ are confirmed by their instances ($Fs$ that are $Gs$).
  - If a piece of evidence $e$ confirms a hypothesis $h$, and $h^*$ is logically equivalent to $h$, then $e$ confirms $h^*$ as well.

then we’re led to the absurd conclusion that the observation of a white shoe confirms the hypothesis that all ravens are black, since it is an instance of the universal generalization that all non-black things are non-ravens $((\forall x)(\neg Bx \rightarrow \neg Rx))$, which is equivalent to the universal generalization that all ravens are black $((\forall x)(Rx \rightarrow Bx))$.

- The Bayesian does accept Hempel’s second principle (we proved this in class last time). However, they do not accept the first. A universal generalization need not be confirmed by its instances.
A toy example: imagine that a scientist knows that there are 8,000 things in existence, and that 4 of these things are ravens, and suppose that they also know that half of the non-ravens are black while half of the non-ravens are non-black. They have split their opinion evenly between the hypothesis that all 4 of the ravens are black (All) and the hypothesis that 2 of the ravens are black and 2 of the ravens are non-black (Some).

\[
\begin{array}{c|c|c}
  & B & \neg B \\
\hline
R & 4 & 0 \\
\neg R & 3,998 & 3,998 \\
\end{array}
\quad
\begin{array}{c|c|c}
  & B & \neg B \\
\hline
R & 2 & 2 \\
\neg R & 3,998 & 3,998 \\
\end{array}
\]

And \( C(\text{All}) = C(\text{Some}) = 1/2 \).

Now, suppose that the scientist randomly selects an object \( a \) from all the things that there are in existence, and finds a non-black non-raven (\( \neg B \land \neg R \)). Both All and Some made this equally likely:

\[
C(\neg B \land \neg R \mid \text{All}) = C(\neg B \land \neg R \mid \text{Some}) = \frac{3,998}{8,000}
\]

Since

\[
C(\neg B \land \neg R) = C(\neg B \land \neg R \mid \text{All}) \cdot C(\text{All}) + C(\neg B \land \neg R \mid \text{Some}) \cdot C(\text{Some})
\]

\[
= \frac{3,998}{8,000} \cdot \frac{1}{2} + \frac{3,998}{8,000} \cdot \frac{1}{2}
\]

\[
= \frac{3,998}{8,000}
\]

the observation of a non-black non-raven should not affect the scientist’s credence that all ravens are black,

\[
C^+(\text{All}) = \frac{C(\neg B \land \neg R \mid \text{All})}{C(\neg B \land \neg R)} \cdot C(\text{All})
\]

\[
= \frac{3,998/8,000}{3,998/8,000} \cdot C(\text{All})
\]

\[
= C(\text{All})
\]

* So, according to BCT, hypotheses are not always confirmed by their instances.

On the other hand, if the scientist had randomly selected a black raven, then the hypothesis that all ravens are black would have been confirmed, since

\[
C(\text{Ra} \land B \mid \text{All}) = \frac{4}{8,000}
\]

while

\[
C(\text{Ra} \land B \mid \text{Some}) = \frac{2}{8,000}
\]
And since
\[ C(Ra \land Ba) = C(Ra \land Ba \mid All) \cdot C(All) + C(Ra \land Ba \mid Some) \cdot C(Some) \]
\[ = \frac{4}{8,000} \cdot \frac{1}{2} + \frac{2}{8,000} \cdot \frac{1}{2} \]
\[ = \frac{3}{8,000} \]
this means that
\[ C^+(All) = \frac{C(Ra \land Ba \mid All)}{C(Ra \land Ba)} \cdot C(All) \]
\[ = \frac{4/8,000}{3/8,000} \cdot \frac{1}{2} \]
\[ = \frac{4}{3} \cdot \frac{1}{2} \]
\[ = \frac{2}{3} \]

* So All receives a big boost of confirmation from the observation of a black raven.

Suppose, finally, that the scientist randomly selects a non-black thing (they know that it will be non-black) and finds that it is a non-raven (they didn’t know that it would be a non-raven).\(^1\) All make this slightly more likely than Some did,

\[ C(\neg Ra \mid \neg Ba \land All) = 1 \]
\[ C(\neg Ra \mid \neg Ba \land Some) = \frac{3,998}{4,000} \]

This means that
\[ C(\neg Ra \mid \neg Ba) = C(\neg Ra \mid \neg Ba \land All) \cdot C(All) + C(\neg Ra \mid \neg Ba \land Some) \cdot C(Some) \]
\[ = \frac{1}{2} + \frac{3,998}{4,000} \cdot \frac{1}{2} \]
\[ = \frac{3,999}{4,000} \]

Thus,
\[ C^+(All) = \frac{C(\neg Ra \mid \neg Ba \land All)}{C(\neg Ra \mid \neg Ba)} \cdot C(All) \]
\[ = \frac{1}{3,999/4,000} \cdot \frac{1}{2} \]
\[ = \frac{4,000}{3,999} \cdot \frac{1}{2} \]
\[ = \frac{4,000}{7,998} \]

* Thus, even though All gets confirmed by the observation of a non-black non-raven, the confirmation it receives is negligible.

---

\(^1\)Suppose also that the fact that a non-black thing was selected is independent of whether All or Some is correct, so that \(C(All \mid \neg Ba) = C(All)\) and \(C(Some \mid \neg Ba) = C(Some)\).
Problems with BCT

The Problem of the Priors

- Think about Goodman’s ‘New Riddle of Induction’: why do we conclude that all emeralds are green after observing \( N \) green emeralds, rather than concluding that all emeralds are grue—since all the green emeralds were also grue?

- What does the Bayesian say about this?

  - Another toy example: Suppose that a scientist knows that they will observe a randomly selected emerald each year from 2000 through the year 2100. In the year 2049, they have observed 50 green emeralds (\( e \)). Call the hypothesis that all emeralds are green \( \text{Green} \) and call the hypothesis that all emeralds are grue \( \text{Grue} \).

    Both of these hypotheses entail that the first 50 emeralds will be green, so \( C(e \mid \text{Green}) = 1 \) and \( C(e \mid \text{Grue}) = 1 \), but that means that the posterior ratio between \( C^+(\text{Green}) \) and \( C^+(\text{Grue}) \) must be the prior ratio between \( C(\text{Green}) \) and \( C(\text{Grue}) \):

    \[
    \frac{C^+(\text{Green})}{C^+(\text{Grue})} = \frac{\frac{C(e \mid \text{Green}) \cdot C(\text{Green})}{C(e)}}{\frac{C(e \mid \text{Grue}) \cdot C(\text{Grue})}{C(e)}} = \frac{C(e \mid \text{Green}) \cdot C(\text{Green})}{C(e \mid \text{Grue}) \cdot C(\text{Grue})} = \frac{C(\text{Green})}{C(\text{Grue})}
    \]

    - So, if \( \text{Green} \) is confirmed by \( e \), then so too must \( \text{Grue} \) be confirmed by \( e \), and by exactly the same factor. According to \( \text{BayesCT} \), if we end up thinking that \( \text{Green} \) is significantly more likely than \( \text{Grue} \), that must be because we started out thinking that \( \text{Green} \) was significantly more likely than \( \text{Grue} \).

- In general, \( \text{BayesCT} \) pushes a lot of the heavy-lifting back to the prior credence function. It appears that certain prior credence functions are more reasonable than others, but \( \text{BayesCT} \) doesn’t tell us how to distinguish the reasonable priors from the unreasonable ones.

- This is the problem of the priors—which priors should a scientist adopt?

  - Three ways to respond to the problem of the priors:
    1. Subjectivism—there’s no objectively correct priors that a scientist ought to adopt. If a scientist were to start out with gruesome priors, they would not be irrational to do so.
    2. Hyper Objectivism—there’s a unique rational prior that a scientist ought to start out with.
    3. Moderate Objectivism—some priors are rational and some are irrational, but there need not be a uniquely correct prior credence function.

    - For either kind of objectivism, the choice of objectively rational or reasonable priors will have to be justified in some way, but this looks very similar to the justificatory burden that we started out with, just situated in a probabilistic framework.
The Problem of Old Evidence

- Sometimes, theories can be confirmed by the fact that they retrodict evidence that we already knew about—old evidence.
  - For instance, it was known for a long time that the perihelion of mercury precessed faster than the Newtonian theory of gravitation predicted.\(^2\)
  - After Einstein formulated the General Theory of Relativity (GTR), it was discovered that the GTR could predict precisely the rate of the precession of the perihelion of mercury. This was taken to be a stunning confirmation of GTR.

- However, if a piece of evidence was known long before the theory was formulated, then the scientists should have already conditionalized on it, so they should already give it credence. So even if the hypothesis entails that evidence, so that \(C(e \mid h) = 1\), it won’t get any confirmation from the old evidence according to BCT, since

\[
C^+(h) = \frac{C(e \mid h)}{C(e)} \cdot C(h) \\
= \frac{1}{C(e)} \cdot C(h) \\
= \frac{1}{1} \cdot C(h) \\
= C(h)
\]

---

\(^2\)The perihelion of mercury is the point in mercury’s orbit when it is closest to the sun. This point precesses, or rotates around the sun, over time.