PIKSI-Logic
The Logic of Belief Revision

Northeastern University
Boston, MA

July 25, 2022

Motivating Credences

- We seem to have not just *beliefs*, but also *degrees of belief* or *credences*.
- **Example:** I am much more certain that the sun will rise tomorrow than that it won't rain today.

Bayesian Epistemology

There are three basic elements:

1. **Credences** are assigned to different propositions.
2. Credences are assumed to behave mathematically like probabilities. They are **subjective probabilities**.
3. **Bayesian conditionalization** tells you to update your credences in light of new evidence in a quantitatively exact way.

Credence or Subjective Probability

- A credence is something like a person's level of expectation for a proposition or a hypothesis or an event.
- We have credences for mundane propositions but also for more monumental propositions like scientific hypotheses.
- In order to appeal to credences, we need both (1) a **mathematical basis** for this concept and, (2) a **psychological basis** for this concept.

Mathematical Basis for Credence

Credences are subjective probabilities. Therefore, they conform to the probability axioms.

The Axioms of Probability

The mathematics of probability takes as basic two entities:

1. a set of outcomes, and
2. a function mapping the elements of that set to probabilities.

The set of outcomes is sometimes called the *outcome space*; the whole thing the *probability space*.

The Axioms of Probability

1. The probability function must map every outcome to a real number between zero and one.
2. The probability function must map an inevitable outcome (e.g., getting either heads or tails when one flips a coin) to one.
3. If two outcomes are mutually exclusive, meaning that they cannot both occur at the same time, then the probability of obtaining either one or the other is equal to the sum of the probabilities for the two outcomes occurring separately.

The Axioms of Probability

1. **Non-Negativity:** For every outcome E, P(E) ≥ 0.
2. **Normality:** For any inevitable outcome E, P(E) $= 1$.
3. **Additivity:** For mutually exclusive outcomes E and D, P(E $\vee$ D) $=$ P(E) $+$ P(D).

Psychological Basis for Credence

- A credence is a psychological property, like belief and desire.
- The psychological reality of credences presents more serious problems for the Bayesian. The existence of a complete set of precise numbers characterizing a level of expectation for all the various events seems implausible.

Credence or Subjective Probability

- The original response to this skeptical attitude was developed by Frank Ramsey (1931), who suggested that credences are closely connected to dispositions to make or accept certain bets.
- Others have maintained that credences are not dispositions to bet, but are rather psychological properties in their own right.

Bayesian Conditionalization

The Bayesian conditionalization says that, if your credence for some outcome D conditional on another outcome E is C(D|E), and if you learn that E has in fact occurred (and nothing else), you should set your unconditional credence for D—that is, C(D)—equal to C(D|E).

Bayesian Conditionalization:

$$C^{+}(D) = C(D|E)$$

Assuming **Probabilism**, this gets us:

$$P^{+}(D) = P(D|E)$$

Bayesian Conditionalization

- The Bayesian, then, postulates two mechanisms by means of which subjective probabilities may justifiably change:

  1. An *experiential process*, which has the effect of sending the subjective probability for some proposition to one.
  2. A *reasoning process*, governed by Bayesian conditionalization, that is triggered by the experiential process.

Outline | Degrees of Belief | Conditionalization | Bayes' Theorem | The Base Rate Fallacy | Jeffrey Conditionalization

Conditional Probability

The conditional probability of an outcome D given another outcome E is written P(D|E). Conditional probabilities are introduced by way of **the Ratio Formula:**

$$P(D|E) = \frac{P(D \cap E)}{P(E)}$$

(If P(E) is zero, then P(D|E) is undefined.)

Outline | Degrees of Belief | Conditionalization | Bayes' Theorem | The Base Rate Fallacy | Jeffrey Conditionalization

Conditional Probability

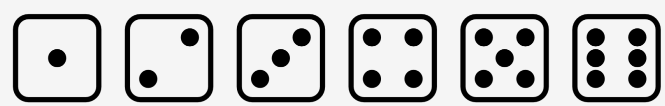
- To determine P(D|E):

  1. Restrict your view to the possible worlds in which the outcome E occurs. Imagine that these are the only possibilities.
  2. Then the probability of "D conditional on E" is the probability of D in this imaginary, restricted universe.

- **Example:** The probability of the die coming up *four* conditional on it coming up *even*.

Outline | Degrees of Belief | Conditionalization | Bayes' Theorem | The Base Rate Fallacy | Jeffrey Conditionalization

Conditional Probability



Conditional Probability

Conditional Probability

Order matters when it comes to conditional probability

Example 1: The probability of being a college student conditional on being a student at PIKSI-Logic.

Example 2: The probability of it being cloudy conditional on it raining.

Some History on Thomas Bayes:

- Was a Presbyterian Minister born in 1701.
- Only published two papers in his lifetime: one on theology, and the other a defence of Newton's calculus.
- Most famous paper was published posthumously: "An Essay towards Solving a Problem in the Doctrine of Chances".

From Conditional Probability to Bayes' Theorem

The Ratio Formula tells us that:

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}, \text{ if } P(B) \neq 0, \tag{1}$$

It also tells us that:

$$P(B \mid A) = \frac{P(A \cap B)}{P(A)}, \text{ if } P(A) \neq 0, \tag{2}$$

Solving for $P(A \cap B)$ in (2) and substituting this into (1) yields **Bayes' theorem**:

$$P(A \mid B) = \frac{P(B \mid A)P(A)}{P(B)}, \text{ if } P(B) \neq 0.$$

Bayes' Theorem

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

There are three different parts to this equation:

1. $\Pr(H|E)$=The probability of the hypothesis conditional on the evidence our experiment has provided us with.
2. $\Pr(H)$=The prior probability of the hypothesis before the evidence is in.
3. $\frac{\Pr(E|H)}{\Pr(E)}$=The force of the evidence.

Why Care About Bayes' Theorem?

- Bayes' Theorem shows us that Conditionalization encodes many truisms of scientific reasoning.
- Bayes' Theorem decomposes the conditional probability we are interested in into quantities that are accessible to us.

Bayes' Theorem (Expanded Formulation)

The Regular Formulation:

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

We can unpack $\Pr(E)$ **in the following way:**

$$\Pr(E) = \Pr(E \wedge H) + \Pr(E \wedge \neg H)$$

$$\Pr(E) = \Pr(E|H)\Pr(H) + \Pr(E|\neg H)\Pr(\neg H)$$

The Expanded Formulation:

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

Bayes' Theorem (Total Probability Formulation)

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

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

Example 1: Diagnostic Testing

Suppose that the false positive rate for a test that is meant to detect dementia is .05. And suppose that the true positive rate for this test is .8.

How confident should I be that I **don't** *have dementia after I have tested positive?*

Example 1: Diagnostic Testing

$$1)\ \Pr(P|D) = .05$$

$$2)\ \Pr(P|\neg D) = .8$$

$$3)\ \Pr(D) = .99$$

$$\frac{\Pr(\text{All False positives})}{\Pr(\text{All Positives})} = \frac{\Pr(P|D)\Pr(D)}{\Pr(P|D)\Pr(D) + \Pr(P|\neg D)\Pr(\neg D)}$$

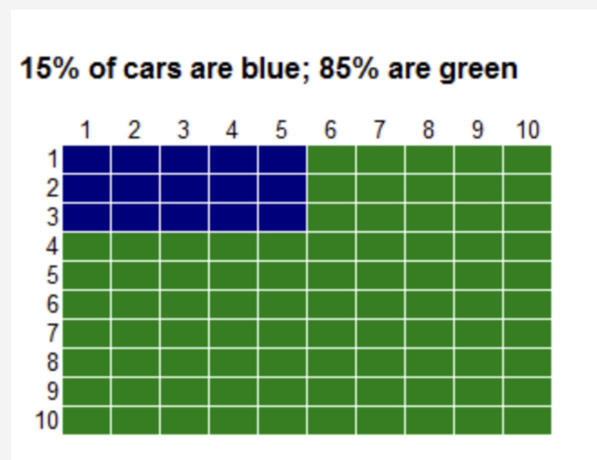
$$\frac{\Pr(\text{All False positives})}{\Pr(\text{All Positives})} = \frac{(0.05)*(0.99)}{[(0.05)*(0.99)] + [(0.8)*(0.01)]}$$

$$\frac{\Pr(\text{All False positives})}{\Pr(\text{All Positives})} = \frac{0.0495}{0.0495 + 0.008} = .86(\checkmark)$$

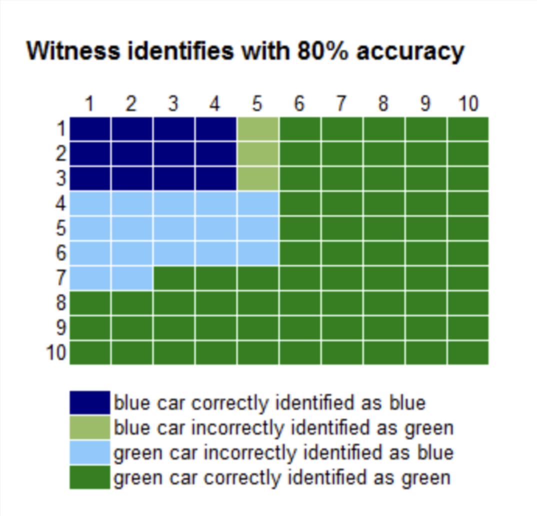
Example: Diagnostic Testing

- What allowed us to carry out the previous calculations is that we knew that the number of people in the population who have dementia is .01. And so we knew that the number of people in the population who *don't* have dementia is .99.
- D=.99
- This is the prior probability of not having dementia. Using this prior, we get the intuitively correct rate of false positives.
- ****An accurate test isn't always a predictive one****

Example 3: Taxi Cab



Example: Taxi Cab



Example: Taxi Cab

Problems: Bayes' Theorem

The probability that I have a coupon for Thai food in my mailbox is .4. The probability that I'll have Thai food for dinner given that there is a coupon in my mailbox is .9. The probability that I'll have Thai food given that I don't have a coupon is .3.

What is the probability that I have a coupon given that I have Thai food for dinner?

Problems: Bayes' Theorem

A Different Interpretation of Bayesianism

- In some contexts, Bayesianism is a theory that tells us that we ought to reason with *degrees of belief* rather than *simple belief*.
- However, in other contexts, Bayesianism is a theory that tells us that we ought to reason with *subjective probabilities* rather than *objective probabilities*.

Given that we've seen that Bayesianism provides an alternative theory of belief revision, it's interesting to note that it does this by means of providing an alternative theory of probability.

Frequentism vs. Bayesianism

Bayesianism: Probability is a matter of the degree of confidence that is warranted by the evidence.

Frequentism: Probability is a matter of the frequency with which certain events occur.

Frequentist Probability

- Probability is objective. Probabilities represent longterm frequencies of repeatable random experiments. For example, 'a coin has probability 1/2 of heads' means that the relative frequency of heads (number of heads out of number of flips) goes to 1/2 as the number of flips goes to infinity.
- $P(E|H)$
- Probability distributions range only over *data given a hypothesis*.

Bayesian Probability

- Probability is subjective. It is the agent's subjective degree of belief.
- $P(H|E) = \frac{P(E|H)P(H)}{P(E)}$.
- Bayesians take probability distributions to range not only over *data given a hypothesis*, but also over *hypotheses* (simpliciter) and *data* (simpliciter).

Actual Frequencies

Actual frequency is the number of times some event does occur.

Example: The number of times the coin actually lands heads divided by the total number of flips.

Hypothetical Frequencies

Hypothetical frequency is the number of times some event would occur if some procedure were repeated over a long period of time.

Example: The number of times the coin would land heads if you flipped it over and over for a long time, divided by the total number of hypothetical flips.

Some Advantages of the Frequentist Approach

- Gives an objective account of probability, i.e., avoids the subjectivism of Bayesianism.
- Seems intuitive in many cases (like where we flip a coin a hundred times).

Some Disadvantages of the Frequentist Approach

- The actual frequency interpretation of probability can't account for our intuitions about events that happen just once.
- Scientific theories don't seem to have frequencies.
- How often an event happens depends on what you compare it to. It depends on the reference class. This is called *the reference class problem*.

Some Advantages of the Bayesian Approach

- Can easily account for one-off events.
- Can account for the intuition that probabilities attach to scientific theories. We have degrees of belief in scientific theories.

Priors and the Replicability Crisis

- Something like priors—"initial assessments", informed guesses"—are widely used in other fields (e.g., actuaries, code breakers, psychologists).
- This suggests that we need the stabilizing presence of these priors in order to get any sort of plausible result at all.

Priors and the Replicability Crisis

"If a researcher shows me data that would only occur one time in twenty if geography didn't matter to hospital waiting times, then I'll become a firm believed in the "postcode lottery", because the ides was reasonably plausible to start with. But if a researcher shows me data that would only occur one time in a 1,000 if the position of Jupiter were irrelevant to British election results, I'll respond that this leaves the idea of a Jovian influence on the British voter only slightly less crazy than it always was."

Some Disadvantages of the Bayesian Approach

- Makes probabilities subjective. This seems unintuitive since probabilities attach to scientific theories, which are themselves supposed to be objective.

Jeffrey Conditionalization

The Bayesian conditionalization dictates that, if your credence for some outcome D conditional on another outcome E is C(D|E), and if you learn that E has in fact occurred (and you do not learn anything else), you should set your unconditional credence for D—that is, C(D)—equal to C(D|E).

Bayesian Conditionalization:

$$C^{+}(D) = C(D|E)$$

Assuming **Probabilism**, this gets us:

$$P^{+}(D) = P(D|E)$$

Jeffrey Conditionalization

- The Bayesian, then, postulates two mechanisms by means of which subjective probabilities may justifiably change:

  1. An *experiential process*, which has the effect of sending the subjective probability for some observable state of affairs to one.
  2. A *reasoning process*, governed by Bayesian conditionalization, that is triggered by the experiential process.

Jeffrey Conditionalization

- We experience the world in sometimes less than favorable conditions, as when we observe a tablecloth in dim lighting that may be green or blue or possibly even purple.
- Such examples led Jeffrey (1965) to propose a generalization of Conditionalization that takes our evidence to be a *partition* of propositions (a set of mutually exclusive and exhaustive propositions) with a set of values that sum to one:

Jeffrey Conditionalization

Jeffrey Conditionalization: When an observation bears directly on the probabilities over a partition $\{E_i\}$, changing them from $P(E_i)$ to $Q(E_i)$, the new probability for any proposition H should be:

$$Q(H) = \sum_i P(H \mid E_i)\, Q(E_i)\,, \text{if defined.}$$

Two Questions:

1. Should we regulate how experience figures into an update?
2. If so, how? Is there a way to formalize this?