Handout

Introduction and Probability Theory

Isaac Wilhelm

Introduction

Today's topic: introduction to probability, chance, credence, and propositional logic.

- Probability: roughly, how likely something is.
- Chance: roughly, how objectively likely something is.
- Credence: roughly, how confident we are in something.

Two faces of probability: chance and credence.

- Chance: probability in the world, rather than in our minds. Chances have to do with the way that the physical world evolves, not with how much we know.
  - The chance of a fair coin landing heads is $\frac{1}{2}$.
  - The chance of a six sided die coming up either 3 or 6 is $\frac{1}{3}$.
  - The chance of the thallium-208 decaying, in six minutes, is $\frac{3}{4}$.
  - According to some, the chance of Japan winning the world cup is equal to some specific value.
- Credence: probability in our minds, rather than in the world. Credences have to do with how much we know, not with the way that the physical

world evolves.

- The rational credence to have in a fair coin landing heads is $\frac{1}{2}$.
- The rational credence to have in a six sided die coming up either 3 or 6 is $\frac{1}{3}$.
- The rational credence to have in the thallium-208 decaying, in six minutes, is $\frac{3}{4}$.
- According to some, the rational credence to have in Japan winning the world cup is equal to some specific value (namely, the chance).

Probability Theory

Probability theory is a mathematical theory of probability. It consists of just two ingredients: a language to describe propositions; a function to assign probabilities to those propositions.

The first ingredient is a formal language L. In this course, L will be the language of propositional logic. So L consists of all sentences which can be formed from (i) sentence letters 'p', 'q', 'r', and so on, (ii) parentheses '(' and ')', and (iii) logical connectives '$\neg$', '$\wedge$', '$\vee$', '$\rightarrow$', and '$\leftrightarrow$'.

The second ingredient is a function 'Pr'. Basically, 'Pr' maps each sentence of L to a number between 0 and 1. Intuitively, if 'Pr' maps a sentence to a number, then that number represents the probability of that sentence being true. Or to put it another way: if X is a sentence of L, and X expresses some proposition, then '$Pr(X)$' represents the probability of that proposition obtaining.

For example, suppose we are about to flip a coin. Let 'p' be a sentence which represents the proposition "The coin lands heads." Suppose that $Pr(p) = \frac{1}{2}$. Then there is a $\frac{1}{2}$ probability—in other words, a 50% probability—that the coin will land heads. Similarly, note that if 'p' represents the proposition "The coin lands heads," then '$\neg p$' represents the proposition "The coin does not land heads." So suppose that $Pr(\neg p) = \frac{1}{2}$. Then there is a $\frac{1}{2}$ probability—in other words, a 50% probability—that the coin will not land heads; in other words, there is a $\frac{1}{2}$ probability that the coin will land tails.

The function 'Pr' satisfies two conditions. The first condition says that all tautologies[1] have probability 1—that is, 100% probability—of being true.

Tautology

For each tautology X of L, $Pr(X) = 1$.

For example, let 'p' be a sentence which represents the proposition "The coin lands heads." Then '$p \vee \neg p$' is a tautology, since it is always either true or false. So according to Tautology, $Pr(p \vee \neg p) = 1$. In other words, there is a 100% probability that the coin either lands heads or does not land heads. And that, of course, seems right. For obviously, the coin either will, or will not, land heads. So intuitively, the probability of the coin landing heads is 100%. And that is precisely what Tautology implies.

The second condition is based on a particular definition. For any sentences

[1] A tautology is a sentence which is always true.

X and Y, say that X and Y are 'mutually exclusive' just in case the following holds: if X is true, then Y is false.

Now for the second condition. Basically, it says that if two sentences are mutually exclusive, then the probability of at least one being true equals the sums of the probabilities of each being true.

> **Sum**
> For all mutually exclusive sentences X and Y of L, $Pr(X \vee Y) = Pr(X) + Pr(Y)$.

For example, suppose that we are about the role a fair, six-sided die. Let 'p' be a sentence which represents the proposition "The die lands on 1," and let 'q' be a sentence which represents the proposition "The die lands on 2." Since the die is fair, $Pr(p) = \frac{1}{6}$: in other words, since the die is fair, the probability of the die landing on 1 is one out of six. Similarly, $Pr(q) = \frac{1}{6}$: in other words, since the die is fair, the probability of the die landing on 2 is one out of six as well. In addition, note that the sentences 'p' and 'q' are mutually exclusive: for if one of those sentences is true – if the die lands on 1, say – then the other is false – the die does not land on 2, say. Therefore, according to Sum, $Pr(p \vee q) = Pr(p) + Pr(q) = \frac{1}{3}$: in other words, the probability of the die landing on *either* 1 *or* 2 is one out of three.

That, in short, is probability theory.[2] It is surprisingly simple. Basically, probability theory consists of (i) a formal language L of propositional logic,

[2] Or rather, that is the version of probability theory that we will discuss in this course. There are more mathematically complicated versions of probability theory which invoke the notions of sets and measures.

and (ii) a function 'Pr' from sentences of L to numbers between 0 and 1 which satisfies the conditions Tautology and Sum.

It turns out that this basic theory of probability can be used to derive a series of important results. Here are a few of them.

Theorem 1. *For all sentences X, $Pr(\neg X) = 1 - Pr(X)$.*

Theorem 2. *For all sentences X and Y, $Pr(X) = Pr(X \wedge Y) + Pr(X \wedge \neg Y)$.*

Theorem 3. *For all logically equivalent sentences X and Y, $Pr(X) = Pr(Y)$.*

Theorem 4 (law of total probability)**.** *Let X be a sentence. Let $Y_1, Y_2, \ldots, Y_n$ be sentences such that the following holds.*

(i) $Pr(Y_1 \vee Y_2 \vee \cdots \vee Y_n) = 1$.

(ii) For all i and j, Y_i and Y_j are mutually exclusive.

Then $Pr(X) = Pr(X \wedge Y_1) + Pr(X \wedge Y_2) + \cdots + Pr(X \wedge Y_n)$.

Before concluding, it is worth seeing how probability theory can be used to answer questions related to real-world phenomena. For example, consider three sports teams: team 1, team 2, and team 3. These teams are competing to win a championship. Exactly one of them will win. The likelihood of team 2 winning is equal to the likelihood of team 1 winning. Team 3 is eight times as likely to win as team 2. How likely is it that team 1 will win? How likely is it that team 2 will win? And how likely is it that team 3 will win?

To answer these questions, let 'a' be a sentence which represents the proposition that team 1 wins, let 'b' be a sentence which represents the proposition that team 2 wins, and let 'c' be a sentence which represents the proposition that team 3 wins. The term 'likelihood' should be taken to mean probability; so claims about the likelihood of this-or-that happening should be interpreted as claims about the values of a probability function 'Pr'. Let us now use this formalism, along with probability theory, to come up with some equations about the probabilities of various outcomes of the competition. That will help us answer the three questions listed above.

To start, note that since exactly one team will win, 'a' and '$b \vee c$' are mutually exclusive. For if team 1 wins, then neither team 2 nor team 3 will win. So if 'a' is true, then '$b \vee c$' is false. Therefore, by Sum, $Pr\big(a \vee (b \vee c)\big) = Pr(a) + Pr(b \vee c)$.

Similarly, since exactly one team will win, 'b' and 'c' are mutually exclusive too. For if team 2 wins, then team 3 will not win. So if 'b' is true, then 'c' is false. Therefore, by Sum again, $Pr(b \vee c) = Pr(b) + Pr(c)$.

In addition, note that since exactly one team will win, the probability that *at least one* team wins must be equal to 1. In other words, the following has probability 1 of obtaining: either team 1 wins, or team 2 wins, or team 3 wins. That is, $Pr\big(a \vee (b \vee c)\big) = 1$.

Since the likelihood of team 2 winning is equal to the likelihood of team 1 winning, $Pr(b) = Pr(a)$. This equation says that the probability of team 2 winning equals the probability of team 1 winning. And since team 3 is eight times as likely to win as team 2, $Pr(c) = 8Pr(b)$. This equation says that the probability of team 3 winning is eight times as great as the probability of team 2 winning.

Putting all these equations together, we get the following.

$$\begin{aligned}
1 &= Pr\big(a \vee (b \vee c)\big) \\
&= Pr(a) + Pr(b \vee c) \\
&= Pr(a) + Pr(b) + Pr(c) \\
&= Pr(a) + Pr(b) + 8Pr(b) \\
&= Pr(a) + 9Pr(b)
\end{aligned}$$

$$= Pr(a) + 9Pr(a)$$

$$= 10Pr(a)$$

where each equality in the above deduction uses one of the equations derived above. So $10Pr(a) = 1$. Therefore, $Pr(a) = .1$, so $Pr(b) = .1$ and $Pr(c) = .8$.