

Probability

Handout by Chris Piech

It is that time in the quarter (it is still Week 1) when we get to talk about probability. We are again going to build this up from first principles. We will heavily use the rules of counting that we learned earlier this week.

Event Spaces and Sample Spaces

A **sample space**, S , is the set of all possible outcomes of an experiment. For example:

1. Coin flip: $S = \{\text{Heads}, \text{Tails}\}$
2. Flipping two coins: $S = \{(H, H), (H, T), (T, H), (T, T)\}$
3. Roll of 6-sided die: $S = \{1, 2, 3, 4, 5, 6\}$
4. Number of emails in a day: $S = \{x \mid x \in \mathbb{Z}, x \geq 0\}$ (non-negative integers)
5. Number of Netflix hours in a day: $S = \{x \mid x \in \mathbb{R}, 0 \leq x \leq 24\}$

An **event space**, E , is some subset of S that we ascribe meaning to. In set notation, $E \subseteq S$.

1. Coin flip is heads: $E = \{\text{Heads}\}$
2. ≥ 1 head on 2 coin flips: $E = \{(H, H), (H, T), (T, H)\}$
3. Roll of die is 3 or less: $E = \{1, 2, 3\}$
4. Number of emails in a day ≤ 20 : $E = \{x \mid x \in \mathbb{Z}, 0 \leq x \leq 20\}$
5. “Wasted day” (≥ 5 Netflix hours): $E = \{x \mid x \in \mathbb{R}, 5 \leq x \leq 24\}$

We say that an event E occurs when the outcome of the experiment is one of the outcomes in E .

Probability

In the 20th century, people figured out one way to define what a probability is:

$$P(E) = \lim_{n \rightarrow \infty} \frac{n(E)}{n}$$

In English this reads: say you perform n trials of an experiment. The probability of a desired event E is the ratio of trials that result in an outcome in E to the number of trials performed (in the limit as your number of trials approaches infinity).

You can also give other meanings to the concept of a probability, however. One common meaning ascribed is that $P(E)$ is a measure of the chance of E occurring.

I often think of a probability in another way: I don’t know everything about the world. As a result I have to come up with a way of expressing my belief that E will happen given my limited knowledge. This interpretation acknowledges that there are two sources of probabilities: natural randomness and our own uncertainty.

Axioms of Probability

Here are some basic truths about probabilities:

Axiom 1: $0 \leq P(E) \leq 1$

Axiom 2: $P(S) = 1$

Axiom 3: If E and F mutually exclusive ($E \cap F = \emptyset$), then $P(E) + P(F) = P(E \cup F)$

You can convince yourself of the first axiom by thinking about the definition of probability above: when performing some number of trials of an actual experiment, it is not possible to get more occurrences of the event than there are trials (so probabilities are at most 1), and it is not possible to get less than 0 occurrences of the event (so probabilities are at least 0).

The second axiom makes intuitive sense as well: if your event space is the same as the sample space, then each trial must produce an outcome from the event space. Of course, this is just a restatement of the definition of the sample space; it is sort of like saying that the probability of you eating cake (event space) if you eat cake (sample space) is 1.

Provable Identities of Probability

Identity 1:

$$P(E^c) = 1 - P(E) \quad (= P(S) - P(E))$$

Identity 2:

$$\text{If } E \subseteq F, \text{ then } P(E) \leq P(F)$$

Identity 3:

$$P(E \cup F) = P(E) + P(F) - P(EF)$$

General Inclusion-Exclusion Identity:

$$P\left(\bigcup_{i=1}^n E_i\right) = \sum_{r=1}^n (-1)^{r+1} \sum_{i_1 < \dots < i_r} P(E_{i_1} E_{i_2} \dots E_{i_r})$$

This last rule is somewhat complicated, but the notation makes it look far worse than it is. What we are trying to find is the probability that any of a number of events happens. The outer sum loops over the possible sizes of event subsets (that is, first we look at all single events, then pairs of events, then subsets of events of size 3, etc.). The “-1” term tells you whether you add or subtract terms with that set size. The inner sum sums over all subsets of that size. The less-than signs ensure that you don’t count a subset of events twice, by requiring that the indices i_1, \dots, i_r are in ascending order.

Here’s how that looks for three events (E_1, E_2, E_3):

$$\begin{aligned} P(E_1 \cup E_2 \cup E_3) &= P(E_1) + P(E_2) + P(E_3) \\ &\quad - P(E_1 E_2) - P(E_1 E_3) - P(E_2 E_3) \\ &\quad + P(E_1 E_2 E_3) \end{aligned}$$

Equally Likely Outcomes

Some sample spaces have outcomes that are all equally likely. We like those sample spaces; they make it simple to compute probabilities. Examples of sample spaces with equally likely outcomes:

1. Coin flip: $S = \{Heads, Tails\}$
2. Flipping two coins: $S = \{(H, H), (H, T), (T, H), (T, T)\}$
3. Roll of 6-sided die: $S = \{1, 2, 3, 4, 5, 6\}$

Probability with equally likely outcomes: For a sample space S in which all outcomes are equally likely,

$$P(\text{Each outcome}) = \frac{1}{|S|}$$

and for any event $E \subseteq S$,

$$P(E) = \frac{\text{number of outcomes in } E}{\text{number of outcomes in } S} = \frac{|E|}{|S|}$$