All possible definitions of probability fall short of the actual practice.

(William Feller)

Conditional Probability

If two events are independent, the occurrence of one does not influence the probability of the other, but what if they are dependent? Let's consider an example.

Example: We flip two fair coins.

(a) I know the first lands on Heads. What is the probability both land on Heads?

When we flip two coins the possible outcomes are $\{TT, TH, HT, HH\}$, but once we know the first lands on Heads, the set of possible outcomes is reduced to $\{HT, HH\}$, both equally likely. Thus, the probability of HH is 1/2.

(b) I know at least one coin lands on Heads. What is the probability both land on Heads?

The set of possible outcomes is reduced to $\{TH, HT, HH\}$, equally likely, so the probability of HH is 1/3.

Remarks: Note that the probabilities vary depending on the extra information that we have. Also note that the extra information reduces the state space, and we compute probabilities in this smaller space.

To find probabilities once we know extra information means to **condition on** the extra information. More formally we define conditional probability as follows.

Definition: If A is an event of positive probability P(A) > 0, then the probability of B once we know A, or probability of B given A is defined as

$$P(B|A) = \frac{P(A \cap B)}{P(A)}.$$

We read P(B|A) as "probability of B given A".

Some important properties of conditional probability:

- 1. We assume P(A) > 0 because it makes no sense to condition on an event that has probability 0; furthermore, if P(A) = 0 the fraction in the definition is undefined!
- 2. Multiplying both sides of the definition by P(A) gives the Multiplication Formula

$$P(A \cap B) = P(A)P(B|A).$$

This formula will be very useful in computing two-step probabilities in the next lecture. Intuitively, for both A and B to occur, first A must occur, and then given that A occurs, B must occur too.

3. If A and B are independent, then

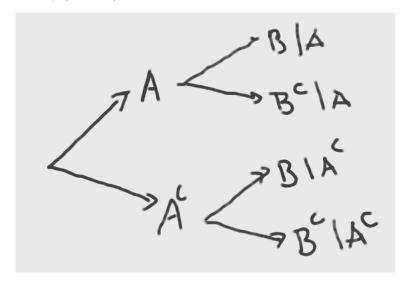
$$P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{P(A)P(B)}{P(A)} = P(B),$$

which makes sense since knowing A does not influence the probability of B in this case.

We can find the probability of some event B by considering the two possibilities of A occurring or A^c occurring (here A^c denotes the complement of the event A). Then

$$P(B) = P(B \cap A) + P(B \cap A^{c}) = P(A)P(B|A) + P(A^{c})P(B|A^{c}).$$

This is just a formal way to compute probabilities along the branches of a tree. We multiply horizontally in the tree (depth) and add vertically (breadth).



Examples:

1. Roll two dice. Let A = sum is 8, and B = first die lands on 3. Then

$$A = \{26, 35, 44, 53, 62\}, \quad B = \{31, 32, 33, 34, 35, 36\}.$$

Thus, P(A) = 5/36, P(B) = 1/6 and $P(A \cap B) = P(\text{pair } 35) = 1/36$

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{1/36}{1/6} = \frac{1}{6}$$
 and $P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{1/36}{5/36} = \frac{1}{5}$.

Another way to solve this problem, once we listed the outcomes in A and B, is to note that

- conditioning on B, we are left with 6 outcomes out of which 35 occurs with probability 1/6;
- conditioning on A, we are left with 5 outcomes out of which 35 occurs with probability 1/5.
- 2. Pick two cards from a deck of 52. Find the probability that both cards are spades.

Let A =first card is a spade and B =second card is a spade. Then we know P(A) = 13/52 and P(B) = 1/4. We also know that P(B|A) = 12/51 since conditioning on having picked a spade, there are 12 spades to pick from the remaining 51 cards. Therefore, using the multiplication formula:

$$P(A \cap B) = P(A)P(B|A) = \frac{13}{52} \times \frac{12}{51} = \frac{3}{51}.$$

Note that A and B are not independent since $P(A)P(B) = \frac{1}{16}$.

3. Based on past experience, 70% of students pass the midterm, out of which 80% pass the final. Only 40% of students failing the midterm pass the final. How many students pass the final?

Let M denote that the student passed the midterm and F that the student passed the final. We want to find the probability that a student (randomly chosen) passed the final, that is P(F). We know P(F|M) = .8 and $P(F|M^c) = .4$. We also know that 70% of students pass the midterm, i.e. P(M) = .7 Then

$$\begin{cases}
M, & (\text{prob. } 0.7) \begin{cases}
F, & (\text{prob. } 0.8) \\
F^c, & (\text{prob. } 0.2)
\end{cases} \\
M^c, & (\text{prob. } 0.3) \begin{cases}
F, & (\text{prob. } 0.4) \\
F^c, & (\text{prob. } 0.6)
\end{cases}$$

$$P(F) = P(F \cap M) + P(F \cap M^c) = P(M)P(F|M) + P(M^c)P(F|M^c) = (.7)(.8) + (.3)(.4) = .68$$

4. Five pennies are sitting on a table. One is a trick coin that has Heads on both sides, but the other four are normal. You pick up a penny at random and flip it four times, getting Heads each time. Given this, what is the probability you picked up the two-headed penny?

We have 2 cases: we pick a fair coin with probability 4/5 or the trick coin with probability 1/5.

$$\begin{cases} \text{ fair coin }, & (\text{prob. } \frac{4}{5}) \begin{cases} HHHH, & (\text{prob. } \frac{1}{16}) \\ \text{ something else }, & (\text{prob. } \frac{15}{16}) \end{cases} \\ \text{trick coin }, & (\text{prob. } \frac{1}{5}) \begin{cases} HHHH, & (\text{prob. } 1) \\ \text{ something else }, & (\text{prob. } 0) \end{cases}$$

Let T be the event that we picked the trick coin and let H denote the even that we get four H in four coin tosses. Then

$$P(T|H) = \frac{P(T \cap H)}{P(H)} = \frac{P(H|T)P(T)}{P(H|T)P(T) + P(H|T^c)P(T^c)} = \frac{1 \cdot \frac{1}{5}}{1 \cdot \frac{1}{5} + \frac{1}{16} \cdot \frac{4}{5}} = \frac{4}{5}$$