Lecture 10: Introduction to Probability Theory

Eric Vigoda
Discrete Mathematics for Computer Science

February 9, 2026

10 Introduction to Probability Theory

Now that we've gotten a good handle on the basics of combinatorics, which is largely about counting, we can now dive into probability theory.

10.1 Probability Distributions

Consider the following experiment: we roll a 6-sided die two times and look at the outcome. First off, the outcome is a sequence (x, y) where $x \in \{1, 2, 3, 4, 5, 6\}$ is the value of the first roll, and $y \in \{1, 2, 3, 4, 5, 6\}$ is the value of the second roll. The way the experiment was phrased – we roll the die two times – means the outcome is a sequence (x, y). If we instead said we roll two dice simultaneously, then the outcome would've been a set $\{x, y\}$ (a set of course is an unordered pair $\{x, y\}$) since there is no ordering for the dice.

Once again: roll a 6-sided die two times. The **sample space** is the set Ω of possible outcomes. In this example,

$$\Omega = \{(x, y) : x, y \in \{1, 2, 3, 4, 5, 6\}\}.$$

If we flipped a coin 3 times then the sample space would be:

$$\begin{aligned}\Omega &= \{b_1 b_2 b_3 : \text{for all } 1 \le i \le 3, b_i \in \{T, H\}\} \\ &= \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}.\end{aligned}$$

A **probability distribution** is defined by a sample space Ω, and a probability $\Pr[\omega]$ for each $\omega \in \Omega$ with the following two constraints:

1. **Non-negativity:** For all $\omega \in \Omega$, $0 \le \Pr[\omega] \le 1$;
2. **Sum to 1:** $\sum_{\omega \in \Omega} \Pr[\omega] = 1$.

(Note, in some classes or settings, you'll sometimes see $p(\omega)$ used in place of $\Pr[\omega]$.)

The simplest example of a probability distribution is the **uniform distribution**. In this case we set $\Pr[\omega] = 1/|\Omega|$ for every $\omega \in \Omega$. The first condition (non-negativity) is clearly satisfied. Also the second condition is satisfied since $\sum_{\omega \in \Omega} \Pr[\omega] = \sum_{\omega \in \Omega} \frac{1}{|\Omega|} = \frac{|\Omega|}{|\Omega|} = 1$.

In the above examples of rolling a die or flipping a coin, the natural probability distribution to consider is in fact the uniform distribution. Assuming a uniform distribution, then in the first example of rolling a die two times, the probability of seeing say $(1, 3)$ is $1/36$. Similarly, in the second example of flipping a coin 3 times, the probability of seeing HHT is $1/8$.

10.2 Events

An **event** is a subset of the sample space. For example, consider the first example of a rolling a die two times. An example event is the following. Let A be the set of pairs, and thus

$$A = \{(i, i) : i \in \{1, 2, 3, 4, 5, 6\}\}.$$

How do we compute the probability of an event, such as A? Well at this point we haven't even defined the probability of an event.

For $B \subseteq \Omega$, let

$$\Pr[B] = \sum_{\omega \in B} \Pr[\omega].$$

This definition of the probability of an event is well-defined because we assumed that we are given the probability $\Pr[\omega]$ for each $\omega \in \Omega$ in the definition of the prescribed probability distribution.

Going back to our earlier example of a rolling a die two times, and the event A being the set of pairs. Then, for the uniform distribution, we have the following:

$$\Pr[A] = \sum_{i=1}^{6} \Pr[(i,i)] = \sum_{i=1}^{6} \frac{1}{36} = 1/6.$$

Hence, in words, the probability that the two rolls have the same outcome is 1/6.

Now consider flipping a coin 3 times. What's the probability that we see two heads and 1 tail? First, let us redefine the sample space Ω as the following:

$$\Omega = \{b_1 b_2 b_3 : \text{for all } 1 \leq i \leq 3, b_i \in \{0,1\}\}.$$

Thus, we view a heads H as a 1, and tails T as a 0. This rephrasing allows us to compactly define the event of interest as follows:

$$A = \{b_1 b_2 b_3 \in \Omega : \sum_{i=1}^{3} b_i = 2\}.$$

Since $|A| = \binom{3}{2}$, then we have that the probability of this event is the following:

$$\Pr[\text{2 heads and 1 tail}] = \Pr[A] = \frac{\binom{3}{2}}{2^3} = \frac{3}{8}.$$

10.3 Examples

1. Roll a die two times. What is the probability that the sum of the dice is exactly 7?

   Answer: Consider the following event:

   $$A = \{(1,6), (2,5), (3,4), (4,3), (5,2), (6,1)\}.$$

   Then we can compute the probability as follows:

   $$\Pr[\text{sum of the dice is 7}] = \Pr[A] = \frac{|A|}{36} = \frac{6}{36} = 1/6.$$

2. Roll a die two times. What is the probability that at least one 6 appears?

   Let $A = \{(i,j) \in \Omega : i = 6 \text{ or } j = 6\}$. let's break it down as follows. Let B be the rolls where the first roll is a 6, thus $B = \{(6,j) \in \Omega : j \in \{1, \ldots, 6\}\}$. And let C be the rolls where the second roll is a 6, thus $C = \{(i,6) \in \Omega : i \in \{1, \ldots, 6\}\}$.

   By inclusion-exclusion,

   $$|A| = |B| + |C| - |B \cap C|.$$

   Thus,

   $$\Pr[\text{at least one 6 appearing}] = \Pr[A] = \Pr[B] + \Pr[C] - \Pr[B \cap C] = \frac{6}{36} + \frac{6}{36} - \frac{1}{36} = \frac{11}{36}.$$

3. Flip a coin 10 times. What's the probability that the number of heads and the number of tails are the same?

   Answer: In this case $|\Omega| = 2^{10}$. For simplicity consider H as 1 and T as 0. Then, consider the following event:

$$C = \{b_1 \dots b_{10} \in \Omega : \sum_{i=1}^{10} b_i = 5\}.$$

   Note, $|C| = \binom{10}{5}$, and therefore we have the following:

$$\Pr[\# \text{ of heads}=\# \text{ of tails}] = \Pr[C] = \frac{|C|}{2^{10}} = \frac{\binom{10}{5}}{2^{10}} \approx .246.$$

10.4 Poker hands

So far, all of our examples have had very small sample spaces. But the real power of probability comes when the sample space is too large to list explicitly. This is where combinatorics really matters. A classic example of this is poker, which we explored quite a bit in our combinatorics work.

A poker hand is a set of 5 cards drawn from a standard deck of 52 cards. Since order does not matter for the set of cards in our poker hand, thus we have that the sample space Ω satisfies

$$|\Omega| = \binom{52}{5}.$$

1. What is the probability that the poker hand has exactly one pair?

   Answer: Let A be the set of poker hands with exactly one pair, which means there is exactly one rank that appears twice, and all other ranks are distinct. We saw in HW 4 that

$$|A| = 13 \times \binom{4}{2} \times \binom{12}{3} \times 4^3.$$

   Therefore,

$$\Pr[\text{exactly 1 pair}] = \frac{13 \times \binom{4}{2} \times \binom{12}{3} \times 4^3}{\binom{52}{5}}.$$

   If we compute this using a calculator (which you will not be required to do in this class) we see that this probability is approximately 0.4226. That means we have roughly a 42% chance of getting exactly one pair.

2. What is the probability that the poker hand has all 5 red cards?

   Answer: Let A be the set of poker hands with all 5 cards being red. Then,

$$|A| = \binom{26}{5}.$$

   Therefore,

$$\Pr[\text{all cards red}] = \frac{\binom{26}{5}}{\binom{52}{5}} \approx .0253.$$

   Thus, the probability of getting all red cards is much smaller than the probability of exactly one pair.

10.5 Balls into Bins

Suppose we throw 100 labeled balls into 20 labeled bins where the balls are labeled $\{1, \ldots, 100\}$ and the bins are labeled $\{1, \ldots, 20\}$. Each ball is equally likely to land in each bin. We say the balls are independent and identically distributed, which we denote as iid. Let's break down precisely what iid means:

Independent The placement of the i-th ball into a bin has no dependence or influence on the placement of any other ball into a bin, so we say the placements are independent. More formally, we will later refer to this independence property as mutually independent.

Identically distributed: Each ball has the same probability distribution for its placement into a bin; in this case, it's the uniform distribution.

Let Ω denote all possible assignments of the 100 balls into 20 bins. More formally, an assignment is a function $\omega : \{1, \ldots, 100\} \to \{1, \ldots, 20\}$ which places each ball into one of the 20 bins. Since there are 20 choices for each ball, we have that

$$|\Omega| = 20^{100}.$$

Note that while many assignments ω may result in the same number of balls in each bin, these assignments are still treated as distinct outcomes in the probability space because the balls are labeled and thus are distinguishable from each other.

Since the balls are all uniformly distributed over all choices, then for each $\omega \in \Omega$,

$$\Pr[\omega] = \frac{1}{|\Omega|} = \frac{1}{20^{100}}.$$

Now we can ask some more interesting questions. What is the probability that bin 1 is empty? Let A denote the set of assignments $\omega \in \Omega$ where $\omega(i) \neq 1$ for all $1 \leq i \leq 100$. Since every ball has 19 remaining choices, we now see that:

$$|A| = 19^{100}.$$

Therefore,

$$\Pr[\text{Bin 1 is empty}] = \frac{19^{100}}{20^{100}} = \left(\frac{19}{20}\right)^{100}.$$

10.6 Birthdays

Now suppose we have 25 balls corresponding to 25 people, and we have 365 bins corresponding to the days of the (non-leap) year. What is the probability that at least two balls get assigned to the same bin? If each person had a birthday chosen uniformly at random then this corresponds to the probability that at least two people have the same birthday.

Let A denote the assignments where at least two balls get assigned to the same bin. Then $\overline{A}$ are the assignments where all balls get assigned to distinct bins. Then,

$$|A| = \binom{365}{25} \times 25! = \frac{365!}{345!}.$$

Also, note that

$$|\Omega| = 365^{25}.$$

Therefore,

$$\Pr[\geq 2 \text{ people have the same birthday}]1 - \Pr[\overline{A}] = 1 - \frac{365 \times (364) \times \cdots \times (346)}{365^{25}} \approx .5687.$$

Therefore, with ≥ 25 people, there is $> 50\%$ chance that 2 people have the same birthday.