

## 10.1: Probability vs. Probability Distribution

You understand probability- what is the likelihood of heads vs. tails in a coin flip? It's 50%, and I'm quite sure that even my drooling cat knows this! Let's go a step further- what is the probability of getting two heads when flipping two coins at once? To evaluate, divide the number of results you are interested in by the total number of possible outcomes:

Outcome 1	heads	heads
Outcome 2	heads	tails
Outcome 3	tails	heads
Outcome 4	tails	tails

We see four potential outcomes, and only one of them is the result we are interested in. Thus, the probability of heads-heads is  $\frac{1}{4} = 0.25 = 25\%$ . Note that this is the same as the product of the probability of heads on two sequential coin flips:  $0.25 = 0.5 \times 0.5$ . Likewise, the probability of observing an even 50/50 split for a 2 coin toss, i.e. either heads-tails or tails-heads, is 50%.

Now imagine that you're studying a gas in an enclosed chamber, the purpose of which is to understand the nature of the individual molecules. You are allowed to make one observation using a spectrometer that informs how many molecules are displaying the "heads" property. Since we suspect that the gas is composed of even-sided "coin" molecules we believe that exactly half the molecules should be "heads". The following table the odds that you will observe a 50/50 split, called  $P(50/50)$ . And while we established that the  $P(50/50)$  odds are 50% for two gas molecules, your intuition likely tells you that the probability for an exact even split drops as the population increases. This is in fact the case:

# coins	$P(50/50)$
2	50%
4	37.5%
100	7.96%
1000	2.52%
10,000	0.80%
100,000	0.255%
$N_A$	$\sim 0\%$

Based on the limiting behavior we can conclude that the probability of a perfect even split is 0% if observing an entire mole of "coin" molecules. We have come to the conclusion that the individual species **do not** have even sides, which is totally wrong!

The lesson here is that probability probably isn't helpful when simultaneously examining large populations of single molecules. The solution is to add some "wiggle room". For example, for 100 molecules there is only an  $\sim 8\%$  chance of observing exactly 50 heads and 50 tails. However, the odds of observing  $50 \pm 3$  heads, i.e. between  $47 \rightarrow 53$ , are  $\sim 52\%$ . Likewise, if we ping 1000 molecules the probability is  $\sim 51\%$  for observing  $500 \pm 10$  heads. Consequently, consideration of a spread in the results returns the correct microscopic description of an even-sided "coin" molecule. Let's add this information to the table:

# Molecules	$P(50/50)$	$\pm$ Spread (%)	$P(50/50 \pm \text{Spread})$
100	7.96%	3 (6.0%)	51.7%
1000	2.52%	10 (2.0%)	51.3%
10,000	0.80%	30 (0.60%)	50.2%
100,000	0.255%	105 (0.21%)	50.01%

**Table 10.2.** Incorporating a spread of potential observables yields higher probabilities.

We need a new mathematical expression for the probability that incorporates a spread. To do so, let's state that we are investigating property  $y$  and we wish to know the probability of observing it. Since  $y$  is observed over a very large population, we have to incorporate a spread in  $y$  just like in the above example. Mathematically we use the expression:  $P = P_{dens}(y) \cdot \Delta y$  to redefine the probability. Here,  $P$  is the probability we are familiar with,  $\Delta y$  is the spread, and  $P_{dens}(y)$  is a function that is like (but not exactly) a probability. We will call it the probability density, and we plot it in Figure 10.1 for the example of 10,000 "coin" molecules. Essentially it is a normalized histogram of the number of heads observed from a very, very large number of measurements. Also, "normalized" means that the area under the curve is 1.0. The curve is bell-shaped (also called "Gaussian") and is centered at 5,000 heads, which makes sense for 10,000 even-sided "coin" molecules. The fact that the area under the curve 1.0 means that there is a 100% chance that, when a measurement is made on a population of 10,000, the number of heads measured will be somewhere between 0 and 10,000. Of course that's true! The area from 49,970 to 50,030 is  $0.502 = 50.2\%$ , which means that if we ask the probability of measuring  $5,000 \pm 30$  heads the answer is  $\sim 50\%$ . Thus, when we ask the right questions, we regain the ability to discern the fact that the coin molecules have two even sides!

When we examine Table 10.2 further we see that the %spread drops as the population increases (fyi: the spread of 6% for 100 coins is calculated as  $\frac{3}{50} = 6\%$ ). This is akin to stating  $\Delta y \rightarrow \partial y$ , which makes our expression for probability:

$$P = P_{dens}(y) \cdot \Delta y \rightarrow P_{dens}(y) \cdot \partial y$$

We have discussed the properties of partials in previous chapters but we will review them here. First, a partial has units, and a probability has no units (it's like a fraction or a percentage). Consequently, we see that the probability density  $P_{dens}(y)$  has inverse units of the property  $y$  that is under investigation. Second, a partial is technically infinitely small, so the value of  $\partial y$  is 0. To understand the impact of this, let's switch from the game of heads and tails to study something more important, such as the probability that a gas phase molecule has a very specific kinetic energy  $E_i$ . The probability that a molecule has exactly  $E_i$  of kinetic energy is:  $P(E_i) = P_{dens}(E_i) \cdot \partial E = 0\%$ , which is not very useful. However, when we see a partial we know that it must be integrated to have any value. When we do so we can start making observations as to the nature of probability densities such as the following:  $\int_0^\infty P_{dens}(E) \cdot \partial E = 1.0$ . This is the property of normalization, which means that the probability of having some energy (or no energy at all) is 100%. In fact, all probability densities have to be normalized such that the following is always true when integrated over all possible values:

$$\int_{lower\ limit}^{upper\ limit} P_{dens}(y) \cdot \partial y = 1.0$$

**10.1.1 Averages.** Probabilities are calculated from integrating probability densities. There are some interesting tricks you can do with this, such as calculate the probability that a molecule is simultaneously moving up and sideways etc.; you can work some examples in problem set questions. What is more useful is calculating the average value of a property such as energy:  $\langle E \rangle$ . Probability distribution are both useful and often necessary for determining as much, and the correct calculus expression is:  $\langle E \rangle = \int_0^\infty E \cdot P_{dens}(E) \cdot \partial E$  where you integrate the thing you want the average of times its probability distribution over all possible values. In general, for any property  $y$ :

$$\langle y \rangle = \int_{lower\ limit}^{upper\ limit} y \cdot P_{dens}(y) \cdot \partial y$$

What is more interesting is that you can also determine what are called higher "moments", such as the 2<sup>nd</sup> moment (the average of the square) of property  $y$ :

$$\langle y^2 \rangle = \int_{lower\ limit}^{upper\ limit} y^2 \cdot P_{dens}(y) \cdot \partial y$$

You can even calculate averages using stipulations. For example, the velocity of a molecule in any direction can be between  $-\infty$  m/s and  $\infty$  m/s since it can go backwards. Thus, you can calculate things like the average velocity in the positive x-direction only by adjusting the limits of integration as so:

$$\langle v_x \rangle_+ = \int_{0\ m/s}^{\infty\ m/s} v_x \cdot P_{dens}(v_x) \cdot \partial v_x$$

where the "+" subscript means the average is conditional that the molecule is moving forwards.

Shown in Figure 10.2 are several graphs of common probability densities. One is the even “tophat” that describes coin tosses and the chances of winning a lottery- basically a very large number of games of chance. Also shown is the Gaussian, where the average  $\langle y \rangle$  is clearly at the top of the distribution. For a bell-shaped curve the average is coincident with the most probable  $y$ , which is labeled  $y_{mp}$ . Most probable values are defined to be at the top of the probability distribution, and can be found by solving the expression:

$$\frac{\partial P_{dens}(y_{mp})}{\partial y} = 0$$

for  $y_{mp}$ . This is the calculus expression for the maximum value of a function, i.e. the point where the curve is flat at the maximum. Last, we also see that the root-mean-square average value  $\sqrt{\langle y^2 \rangle}$  is near the top but is slightly higher than the average. In fact it is generally true that  $\sqrt{\langle y^2 \rangle}$  is greater than  $\langle y \rangle$ .

---

This page titled [10.1: Probability vs. Probability Distribution](#) is shared under a [CC BY-NC 4.0](#) license and was authored, remixed, and/or curated by [Preston Snee](#).