# The world of probability

**Probability** is the likelihood of observing a particular event.

* A number between 0 and 1
* When p = 0, the event never happens
* When p = 1, the event always happens

There are lots and lots (and lots and lots and lots) of rules that govern probability. However, today we are only concerned with one: the Law of Large Numbers.

The **Law of Large Numbers** states that the probability of an outcome is equal to its long run relative frequency.

In simpler terms, it says that the probability of an event is equal to the fraction of time we expect to observe that event over a very long time.

We know that the probability of flipping a coin and getting a heads is 1/2 because that's the proportion of time we expect to see a heads if we flip a coin for a very very very very very long time.

We could even calculate the probability of getting a red light at an intersection if we could record thousands of times that we passed the intersection and stopped or did not stop for a red light.

Why do we need the Law of Large Numbers right now? It allows us to substitute the word "probability" in the place of “proportion" in our statements about the Normal distribution.

We want to be able to talk about the probability of a data point with a Normal distribution being above or below a certain value.

### For example,

If healthy human body temperature has a Normal distribution, what is the probability of a single healthy human having a body temperature greater than 98.6 degrees F?

This is all great for data that has a Normal distribution; however, most data does not. Does that mean we can't figure out probabilities for those distributions?

# Enter the Central Limit Theorem

Does that mean we can't figure out probabilities for those distributions? - That depends a lot. There are sometimes some ways to calculate non-Normal probabilities, but it's much trickier. Data with a Normal distribution is much, much easier to deal with.

However, statistics comes to us with an amazing concept: **The Central Limit Theorem**.

The Central Limit Theorem tells us that, no matter what kind of distribution a sample of observations are drawn from, so long as the sample size is large enough (about 30-40) the sample mean will have a Normal distribution with a mean equal to the population mean.

A proof of the CLT is beyond the scope of Lambda School (see: calculus), but hopefully you will get a sense of how the CLT works by working through this exercise.

Open the coin flips and dice rolls app linked below:

<https://dansmyers.github.io/CentralLimitTheorem/>

Start by clicking on “Flipping a Coin” at the top.

Complete the table below using the app. You’ll change the “number of coins” and click “Single example trial” to complete the first empty column. Then you’ll repeat the process by generating a histogram of 10,000 trials (click the “Histogram of 10,000 trials) and fill in the last four columns.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Single example trial | Histogram of 10,000 trials | | | |
| # of coins | Fraction of heads | Mean | SD | Min | Max |
| 1 |  |  |  |  |  |
| 2 |  |  |  |  |  |
| 5 |  |  |  |  |  |
| 10 |  |  |  |  |  |
| 20 |  |  |  |  |  |
| 50 |  |  |  |  |  |
| 100 |  |  |  |  |  |

Describe the pattern you see as you increase the number of coins flipped. How does the sampling distribution of the sample proportion (the mean and SD generated from the histogram) change as you increase the number of coins flipped (which is akin to increasing the sample size).

Change the probability of a heads to 0.7. Repeat the process described above and fill in this table.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Single example trial | Histogram of 10,000 trials | | | |
| # of coins | Fraction of heads | Mean | SD | Min | Max |
| 1 |  |  |  |  |  |
| 2 |  |  |  |  |  |
| 5 |  |  |  |  |  |
| 10 |  |  |  |  |  |
| 20 |  |  |  |  |  |
| 50 |  |  |  |  |  |
| 100 |  |  |  |  |  |

Describe the pattern you see as you increase the number of coins flipped. How does the sampling distribution of the sample proportion (the mean and SD generated from the histogram) change as you increase the number of coins flipped (which is akin to increasing the sample size).

Now look at “Rolling a Dice”.

Complete the table below using the app. You’ll change the “number of dice” and click “Single example trial” to complete the first empty column. Then you’ll repeat the process by generating a histogram of 10,000 trials (click the “Histogram of 10,000 trials) and fill in the last four columns.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Single example trial | Histogram of 10,000 trials | | | |
| # of dice | Mean roll | Mean | SD | Min | Max |
| 1 |  |  |  |  |  |
| 2 |  |  |  |  |  |
| 5 |  |  |  |  |  |
| 10 |  |  |  |  |  |
| 20 |  |  |  |  |  |
| 50 |  |  |  |  |  |
| 100 |  |  |  |  |  |

Describe the pattern you see as you increase the number of dice rolled. How does the sampling distribution of the sample proportion (the mean and SD generated from the histogram) change as you increase the number of coins flipped (which is akin to increasing the sample size).