

## 4.1: Empirical Probability

One story about how probability theory was developed is that a gambler wanted to know when to bet more and when to bet less. He talked to a couple of friends of his that happened to be mathematicians. Their names were Pierre de Fermat and Blaise Pascal. Since then many other mathematicians have worked to develop probability theory.

Understanding probabilities are important in life. Examples of mundane questions that probability can answer for you are if you need to carry an umbrella or wear a heavy coat on a given day. More important questions that probability can help with are your chances that the car you are buying will need more maintenance, your chances of passing a class, your chances of winning the lottery, your chances of being in a car accident, and the chances that the U.S. will be attacked by terrorists. Most people do not have a very good understanding of probability, so they worry about being attacked by a terrorist but not about being in a car accident. The probability of being in a terrorist attack is much smaller than the probability of being in a car accident, thus it actually would make more sense to worry about driving. Also, the chance of you winning the lottery is very small, yet many people will spend the money on lottery tickets. Yet, if instead they saved the money that they spend on the lottery, they would have more money. In general, events that have a low probability (under 5%) are unlikely to occur. Whereas if an event has a high probability of happening (over 80%), then there is a good chance that the event will happen. This chapter will present some of the theory that you need to help make a determination of whether an event is likely to happen or not.

First you need some definitions.

### Definition 4.1.1

**Experiment:** an activity that has specific result that can occur, but it is unknown which results will occur.

### Definition 4.1.2

**Outcomes:** the result of an experiment.

### Definition 4.1.3

**Event:** a set of certain outcomes of an experiment that you want to have happen.

### Definition 4.1.4

**Sample Space:** collection of all possible outcomes of the experiment. Usually denoted as  $SS$ .

### Definition 4.1.5

**Event Space:** the set of outcomes that make up an event. The symbol is usually a capital letter.

Start with an experiment. Suppose that the experiment is rolling a die. The sample space is  $\{1, 2, 3, 4, 5, 6\}$ . The event that you want is to get a 6, and the event space is  $\{6\}$ . To do this, roll a die 10 times. When you do that, you get a 6 two times. Based on this experiment, the probability of getting a 6 is 2 out of 10 or  $1/5$ . To get more accuracy, repeat the experiment more times. It is easiest to put this in a table, where  $n$  represents the number of times the experiment is repeated. When you put the number of 6s found over the number of times you repeat the experiment, this is the relative frequency.

Table 4.1.1: Trials for Die Experiment

$n$	Number of 6s	Relative Frequency
10	2	0.2
50	6	0.12
100	18	0.18
500	81	0.162

$n$	Number of 6s	Relative Frequency
1000	163	0.163

Notice that as  $n$  increased, the relative frequency seems to approach a number. It looks like it is approaching 0.163. You can say that the probability of getting a 6 is approximately 0.163. If you want more accuracy, then increase  $n$  even more.

These probabilities are called **experimental probabilities** since they are found by actually doing the experiment. They come about from the relative frequencies and give an approximation of the true probability. The approximate probability of an event  $A$ ,  $P(A)$ , is

#### Definition 4.1.6

##### Experimental Probabilities

$$P(A) = \frac{\text{number of times } A \text{ occurs}}{\text{number of times the experiment was repeated}}$$

For the event of getting a 6, the probability would be  $\frac{163}{1000} = 0.163$ .

You must do experimental probabilities whenever it is not possible to calculate probabilities using other means. An example is if you want to find the probability that a family has 5 children, you would have to actually look at many families, and count how many have 5 children. Then you could calculate the probability. Another example is if you want to figure out if a die is fair. You would have to roll the die many times and count how often each side comes up. Make sure you repeat an experiment many times, because otherwise you will not be able to estimate the true probability. This is due to the law of large numbers.

#### Definition 4.1.7

**Law of large numbers:** as  $n$  increases, the relative frequency tends towards the actual probability value.

#### Note

Probability, relative frequency, percentage, and proportion are all different words for the same concept. Also, probabilities can be given as percentages, decimals, or fractions.

## Homework

### Exercise 4.1.1

- Example 4.1.2 contains the number of M&M's of each color that were found in a case (Madison, 2013). Find the probability of choosing each color based on this experiment.

Table 4.1.2: M&M Distribution

Blue	Brown	Green	Orange	Red	Yellow	Total
481	371	483	544	372	369	2620

- Eyeglassomatic manufactures eyeglasses for different retailers. They test to see how many defective lenses they made the time period of January 1 to March 31. Example 4.1.3 gives the defect and the number of defects. Find the probability of each defect type based on this data.

Table 4.1.3: Number of Defective Lenses

Defect type	Number of defects
Scratch	5865
Right shaped - small	4613
Flaked	1992

Defect type	Number of defects
Wrong axis	1838
Chamfer wrong	1596
Crazing, cracks	1546
Wrong shape	1485
Wrong PD	1398
Spots and bubbles	1371
Wrong height	1130
Right shape - big	1105
Lost in lab	976
Spots/bubble - intern	976

3. In Australia in 1995, of the 2907 indigenous people in prison 17 of them died. In that same year, of the 14501 non-indigenous people in prison 42 of them died ("Aboriginal deaths in," 2013). Find the probability that an indigenous person dies in prison and the probability that a non-indigenous person dies in prison. Compare these numbers and discuss what the numbers may mean.
4. A project conducted by the Australian Federal Office of Road Safety asked people many questions about their cars. One question was the reason that a person chooses a given car, and that data is in Example 4.1.4 ("Car preferences," 2013). Find the probability a person chooses a car for each of the given reasons.

Table 4.1.4: Reason for Choosing a Car

Safety	Reliability	Cost	Performance	Comfort	Looks
84	62	46	34	47	27

#### Answer

1.  $P(\text{blue}) = 0.184$ ,  $P(\text{brown}) = 0.142$ ,  $P(\text{green}) = 0.184$ ,  $P(\text{orange}) = 0.208$ ,  $P(\text{red}) = 0.142$ ,  $P(\text{yellow}) = 0.141$
3.  $P(\text{indigenous person dies}) = 0.0058$ ,  $P(\text{non-indigenous person dies}) = 0.0029$ , see solutions

This page titled [4.1: Empirical Probability](#) is shared under a [CC BY-SA 4.0](#) license and was authored, remixed, and/or curated by [Kathryn Kozak](#) via [source content](#) that was edited to the style and standards of the LibreTexts platform.

- [4.1: Empirical Probability](#) by [Kathryn Kozak](#) is licensed [CC BY-SA 4.0](#). Original source: <https://s3-us-west-2.amazonaws.com/oerfiles/statsusingtech2.pdf>.