Introduction to Probability

We frequently make estimates of the likelihood of an event happening. Such estimates might be based on:

- instinct/intuition
- past experiences
- logical deduction

We will use set theory to assign a measure to the likelihood (probability) of an event occurring in a way that agrees with our intuition. This will also add clarity to help us with deeper calculations.

Basic Definitions

Activities/phenomenons for which the outcome cannot be predicted with certainty are called **random experiments**. These can produce a variety of results called **outcomes**.

The set of all possible outcomes is called the **sample space** S. Each element of the sample space is a **sample point**.

An **event** is a subset of a sample space. It can consist of a single outcome (simple event), or more. The empty set and the entire sample space are also events.

Probability theory makes sense only in the context of activities that can be repeated or phenomena that can be observed multiple times. We call each observation or repetition of the experiment a **trial**.

Basic Definitions

Examples:

- Rolling a die and observing the number on the "up" face is an experiment with six possible outcomes: $\{1, 2, 3, 4, 5, 6\}$. Each roll is a trial.
- Asking a person if they intend to vote in the next election is an experiment with two possible outcomes: $\{$"yes", "no"$\}$. Each citizen asked is a trial.
- Measuring the amount of time it takes to find parking at the mall is an experiment with infinitely many possible outcomes (between 0 minutes to infinity).
- Rolling a die and seeing whether the number that comes up is even/odd is an experiment with two possible outcomes: $\{$"even", "odd"$\}$.

All these examples have some inherent **randomness**.

Making Predictions

We can make a prediction about how likely each possible outcome is. This is a measure of the likelihood (i.e. **the probability**) that the outcome will occur.

Example: If I roll a die, I would assume that each of the 6 numbers are equally likely to appear. I'd predict the probability of rolling a 5 to be 1/6.

Example: A baseball player's batting average in 2016 was 0.292. I would predict that in the 2017 season he will have a hit with probability 29%.

Example: If I toss a coin 10 times, I would *expect* that the total number of heads is going to be exactly 5 out of 10. It turns out that this only happens with probability $\approx 24.6\%$.

Empirical Probability

One way to **estimate the probability** of an outcome is by repeating the experiment, and counting how many times the outcome occurs. This is the **frequency** of the outcome.

The **relative frequency** of the outcome is the *proportion* of times that it occurs:

$$\text{relative frequency} = \frac{\text{frequency of the outcome}}{\text{total number of trials}}$$

This is called **empirical probability**, and is an estimate of the actual probability.

Empirical Probability

Example: A group of 300 voters were asked who they would vote for in the next election. The following table shows the results of the poll:

Candidate	**Frequency**	**Relative Frequency**
Melinda McNulty	120	
Mark Reckless	80	
Ian Lawless	100	
Total	300	

Estimate the probability that a voter chosen at random from the voting population will vote for any given candidate.

Empirical Probability

Estimate the probability that a voter chosen at random from the voting population will vote for any given candidate.

Candidate	**Frequency**	**Relative Frequency**
Melinda McNulty	120	$\frac{120}{300} = 0.4 = 40\%$
Mark Reckless	80	$\frac{80}{300} \approx 0.267 = 26.7\%$
Ian Lawless	100	$\frac{100}{300} \approx 0.333 = 33.3\%$
Total	300	100%

Empirical Probability

Example: A bored probability professor wants to verify that tossing a coin really does come up Heads 50% of the time, so he repeatedly tosses a coin and records the number of heads. The table below shows his results:

# trials	Freq. of Heads	Relative Frequency
4,040	2,048	$\frac{2048}{4040} = 0.5069$
10,000	5,067	$\frac{5067}{10000} = 0.5067$
12,000	6,019	$\frac{6019}{12000} = 0.5016$
24,000	12,012	$\frac{12012}{24000} = 0.5005$

The Law Of Large Numbers

The Law of large numbers: as more and more trials of an experiment are repeated, the relative frequency obtained approaches the actual probability.

If we have used (sound) logic to compute the probability for an outcome, and if we run our experiment "many" times, the relative frequency of the outcome will be "close" to the theoretical probability.

Note: This raises the question of how many trials are needed to get a good estimate. One needs to explore statistics in order to figure this out.

The Law Of Large Numbers

Example: A group of 300 voters were asked who they would vote for in the next election. The following table shows the results of the poll:

Candidate	Frequency	Relative Frequency
Melinda McNulty	120	0.40
Mark Reckless	80	0.267
Ian Lawless	100	0.333

Q: If we select a sample of 5,000 voters, how many people should we expect to vote for Melinda?

By the law of large numbers, the probability that a voter chooses Melinda is approximately 40%. Out of 5,000, we would expect $0.40 \cdot 5000 = 2000$ people to vote for Melinda.

The Law Of Large Numbers

Q: Is a pole of 300 individuals large enough to get a good estimate of voter preferences?

Obviously the larger the sample, the better the estimate for the preferences of the entire population, but also the more costly the polling process.

It turns out that most opinion polls tend to sample about 1,000 people. We'll learn what's so special about 1000!

The Sample Space

Recall that the **sample space** S for an experiment is the set of all possible outcomes. Each element (or outcome) is a **sample point**.

Example: Roll a six sided die and observe the number on the "up" face. What is the sample space?
$S = \{1, 2, 3, 4, 5, 6\}$.

Example: Roll a six sided die and observe whether the number on the uppermost face is even or odd. What is the sample space? $S = \{\text{even, odd}\}$.

Note: It is important to specify what is being observed/recorded in the experiment, as this determines your sample space!

The Sample Space

When determining the sample space S for an experiment, we need to make sure it has the following properties:

- Each element in S is a possible outcome.
- The set S covers all possible outcomes.
- No two outcomes can occur simultaneously.

Example: For rolling a die and observing the number on the "up" side, the following are NOT valid sample spaces:

$$\{0, 1, 2, 3, 4, 5, 6\}, \quad \{1, 3, 5, 6\},$$

$$\{1, 2, 3, 4, 5, 6, \text{“1 and 5”}\}$$

The Sample Space

Example: Determine the sample spaces for the following:

- Flip 2 coins and count the # of H. $S = \{0, 1, 2\}$
- Flip 2 coins and count the # of T. $S = \{0, 1, 2\}$
- Flip 2 of coins and observe how many times you get two Heads. $S = \{0, 1\}$
- Flip a pair of coins and observe how many times you get one H and one T. $S = \{0, 1\}$

The Sample Space

Example: Determine the sample spaces for the following:

- Roll a pair of dice and observe the minimum of the numbers on the "up" face. $S = \{1, 2, 3, 4, 5, 6\}$

- Roll a pair of dice and observe the sum of the numbers on the "up" face. $S = \{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$

Events

Recall that an **event** E is a subset of the sample space S. All subsets of S describe an event.

The event that "nothing happens" is $\emptyset$, and the event that "something happens" is S itself.

We say that an event **occured** if the outcome of a trail is an element of the event set.

Sometimes we can give a verbal description of E. The corresponding subset is comprised of all outcomes which fit the verbal description.

Events

Example: Roll a die and observe the number on the "up" face. The sample space is $S = \{1, 2, 3, 4, 5, 6\}$. Find the subsets corresponding to the following verbal description of the given events:

$E =$ the event that the number is even.
$E = \{2, 4, 6\}$

$F =$ the event that the number is larger than two.
$F = \{3, 4, 5, 6\}$

If we roll a 3, then F *has occurred*, but E *has not occurred*.

Properties of Probability

Let $S = \{e_1, e_2, \ldots, e_n\}$ be the sample space for an experiment, with outcomes $e_1, e_2, \ldots, e_n$.

1. For each outcome e_i in S, we have its corresponding probability $P(e_i)$, which is a number between 0 and 1:

$$0 \leq P(e_i) \leq 1$$

2. The probabilities of all the outcomes must add up to 1:

$$P(e_1) + P(e_2) + \cdots + P(e_n) = 1$$

3. The probability $P(E)$ of an event E is the sum of the probabilities of all the outcomes inside E.
4. $P(\emptyset) = 0$ and $P(S) = 1$.
5. Complement principle: $P(E) = 1 - P(E')$

Properties of Probability

Example: Students in a dorm were asked how many pairs of shoes they owned and everybody had between 1 and 7 pairs of shoes. $S = \{1, 2, 3, 4, 5, 6, 7\}$. Which of the following is a valid assignments of probabilities?

# Pairs	Probability
1	1/4
2	1/4
3	1/16
4	1/16
5	1/8
6	1/8
7	1/8

# Pairs	Probability
1	−1
2	1/2
3	1/2
4	1/4
5	1/4
6	1/4
7	1/4

# Pairs	Probability
1	1
2	1/4
3	1/16
4	1/16
5	1/8
6	1/8
7	1/8

The only valid box is Box 1. The second box has a negative probability of -1, while the sum of the probabilities in the third box is more than 1.

Properties of Probability

Example: A bag of marbles contains 5 Red, 10 Blue, and 20 Green. You choose a marble at random from the bag and observe the color of the marble. Assign probabilities to the outcomes in the sample space for this experiment.

There are 35 marbles total. The chance of drawing a Red is $\frac{5}{35}$. The chance of drawing a Blue is $\frac{10}{35}$. The chance of drawing a Green is $\frac{20}{35}$. Notice all probabilities are between 0 and 1 and they add up to 1.

Note: we prefer to use shorter notation. The sample space is $S = \{R, B, G\}$, and the probabilities are

$$P(R) = \frac{5}{35} \qquad P(B) = \frac{10}{35} \qquad P(G) = \frac{20}{35}$$

Properties of Probability

Recall that to calculate the probability of an event E, we sum the probabilities of the outcomes in the event.

Example: A bag of marbles contains 5 Red, 10 Blue, and 20 Green. You choose a marble at random from the bag and observe the color of the marble. What is $P(\text{not red})$?

Since we don't want red, we are interested in the event that the marble is blue or green, so $E = \{B, G\}$. Then

$$P(E) = P(B) + P(G) = \frac{10}{35} + \frac{20}{35} = \frac{30}{35}$$

Alternately, we can use the complement principle!

$$P(E) = 1 - P(E') = 1 - P(Red) = 1 - \frac{5}{35} = \frac{30}{35}$$

Properties of Probability

Example: The following table shows the probabilities of a randomly selected student from our Finite Math class of falling into one of the four categories.

	Athlete	**Non-Athlete**
Freshman	0.45	0.16
Sophomore	0.26	0.13

What is the probability of a randomly chosen student to be a Freshman and an Athlete? 0.45 or 45%

What is the probability of a randomly chosen student to be an Athlete? $0.45 + 0.26 = 0.71$

What is the probability of a randomly chosen student to be a Sophomore? $0.26 + 0.13 = 0.39$

Properties of Probability

Example: Let $S = \{1, 2, 3\}$, $E = \{1, 2\}$ and $F = \{2, 3\}$. Suppose that $P(E) = 0.7$ and $P(F) = 0.5$. Given that $P(2) = 0.2$, find $P(1)$ and $P(2)$.

Since $P(E) = P(1) + P(2)$, we find that

$$P(1) = 0.7 - 0.2 = 0.5$$

Similarly, $P(F) = P(2) + P(3)$, so

$$P(3) = 0.5 - 0.2 = 0.3$$

Note: In this case, we don't need to know about $P(2)$ to determine all 3 events! As an exercise, try to solve for $P(1), P(2), P(3)$ by only using $P(E) = 0.7$ and $P(F) = 0.5$.

Properties of Probability

Example: The students in a dorm were asked how many siblings they had. The following table shows the relative frequencies of the outcomes.

# Siblings	Rel. Freq.
0	.2
1	.2
2	.2
3	.1
4	.1
5	.1
6	.05
7	.02
8	.01
9	.01
10	.01

Properties of Probability

Find the probability that a student chosen at random from the dorm will have more than 3 siblings.

# Siblings	Rel. Freq.
0	.2
1	.2
2	.2
3	.1
4	.1
5	.1
6	.05
7	.02
8	.01
9	.01
10	.01

The event is $E = \{4, 5, \ldots, 10\}$.

We add the probabilities inside E:

$P(E) = 0.1 + 0.1 + 0.05 + 0.02 + 0.01 + 0.01 + 0.01 = 0.3$

Alternately, we could use the complement principle:

$$1 - P(E') = 1 - (0.2 + 0.2 + 0.2 + 0.1) = 1 - 0.7 = 0.3$$