A **set** is a collection of objects. The objects are called the **elements** of the set. If a set has finitely many elements, it is a **finite set**, otherwise it is an **infinite set**. If the number of elements in a set is not too many, we can just list them out. We usually use capital letters to name sets and use braces $\{\}$ to de-limit a set:

Example: $S = \{1, 4, 10\}$ is the set that contains three elements, namely the numbers 1, 4, and 10.

When the size of the set is too large or infinite, we just give a description of them without listing them. For example, the set of real numbers, the set of even integers, the set of all books written before the year 2000.

If two sets A and B have the same elements, we say that they are equal, and write $A = B$.

A **subset** of a set is a sub-collection of the set. For example, if $S = \{1, 4, 10\}$, then $A = \{1, 4\}$ is a *subset* of S since A is a sub-collection of S. (We say that every element of A is an element of S.) If A is a subset of S, we write $A \subseteq S$.

By definition, every set is a subset of itself. If A is a subset of S but $A \neq S$, we say that A is a **proper subset** of S. If A is a proper subset of S, we write $A \subset S$.

The set that contains no object is the **empty set**, and is denoted by the mathematical symbol $\varnothing$. The empty set is a subset of any set.

If A, B are sets, the **union** of the two sets, denoted by $A \cup B$, is the (most likely bigger) set whose elements are elements that are either in A or in B (or both).

If A, B are sets, the **intersection** of the two sets, denoted by $A \cap B$, is the (most likely smaller) set whose elements are elements that are in **both** sets.

Example: If $A = \{3, 6, 9, 11, 15, 17\}$, and $B = \{0, 3, 9, 10, 15\}$, then

$A \cup B = \{0, 3, 6, 9, 10, 11, 15, 17\}$

$A \cap B = \{3, 9, 15\}$

If a set A is finite, we define the **order** of the set, denoted by $|A|$, to be the number of elements in a set.

For example, for the above sets A, and B, since set A has 6 elements, so $|A| = 6$. Also, $|B| = 5$. $A \cup B$ has 8 elements, therefore $|A \cup B| = 8$, and $|A \cap B| = 3$. You may have noticed that, for the above example, the number of elements in the union of A and B is equal to the number of elements in A plus the number of elements in B, then subtract the number of elements in the intersection of A and B. This is actually always true for any finite sets:

If A, B are finite sets, then:

$|A \cup B| = |A| + |B| - |A \cap B|$

Notice that $|\varnothing| = 0$

Definition: We say that two sets A and B are **mutually exclusive** if $A \cap B = \varnothing$

Formula: If A and B are mutually exclusive, then:

$|A \cup B| = |A| + |B|$

Example: If you have three different pairs of shoes, and you have 2 different shirts, how many different ways can you dress yourself with a pair of shoe and a shirt?

You have six different ways to dress yourself. Notice that the number six is the product of 2 and 3. This is not a coincidence, but an example of the **multiplication principle**:

If there's m many ways to perform an experiment, and *corresponding to each of these ways* there's n many ways to perform another experiment, then there are $m \times n$ many different ways to perform the combined experiment.

The multiplication principle can be expanded to cases in which more than 2 experiments are being performed. If experiment 1 can be performed in m_1 many ways, and experiment 2 can be performed in m_2 many ways, ..., and experiment k can be performed in m_k many ways, then there are $m_1 m_2 \cdots m_k$ many different ways in which all the combined experiments may be performed.

Example: In how many different ways can one buy a phone, a radio, and a television of each if there are 10 different brands of phone, 6 different brands of radio, and 8 different brands of TV to choose from?

Using the counting principle, there are $10 \cdot 6 \cdot 8 = 480$ ways in which one of each can be bought.

Permutation

In how many different ways can 10 candidates be chosen to fill the position of president, vice president, and treasurer?

Since there are 10 candidates, we may choose any one of them to be the president, that is, there are 10 ways to choose the president. After a president is chosen, there are only 9 people left, so we have 9 different ways to choose for the vice president. After that, we have only 8 people left, and any one of them can be the treasurer, so we have 8 different ways to choose the treasurer. Using the multiplication principle, there are $10 \cdot 9 \cdot 8 = 720$ ways to assign these three positions from the 10 candidates.

Example: How many different *words* can we form by using the letters of the word **MONEY**? Note: By a *word* here we do not mean a word that is necessarily in the dictionary, just any combination of the alphabets. Examples of "words" would be *MNOEY*, *MEOYN*, *OENMY* ...etc.

Notice that all five of the letters in the word *money* are different. We may use any of the five letters to be the first letter of the word, so for the first letter of the word we have 5 choices. On the second letter of the word we have 4 choices left, on the third letter we have 3 choices left, then we have 2 choices for the fourth letter, then only 1 choice for the last letter. Using the counting principle we have $5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 5! = 120$ different words that can be formed.

In the previous examples **the order in which we choose the elements matter**. In the first example, if John is chosen as the President and Mary the Vice President, it would be different than if Mary is the President and John the Vice President, even if both John and Mary are used in either case. In the second example, we are always using the same five letters, but the ordering of the letters make different *words*.

An arrangement of a particular order is called a **permutation**. In a permutation of objects, the order at which the elements are chosen/written matters, meaning that even if the same elements are used or chosen, but if they are used or chosen in different order, we consider that to be two different arrangement.

If we have n many objects and r of them are taken at a time, then the **number of permutations** is denoted by nPr and is given by the formula:

The number of permutations (ordered arrangements) of n distinct objects taken r at a time is given by:

$$nPr = \frac{n!}{(n-r)!} = n(n-1)(n-2)\cdots(n-r+1)$$

Example: How many different 3-letter *words* can be formed from the letters of the word **mother**

There are 6 different letters in the word *mother*, and we are to choose 3 letters to form a word, therefore, we have ${}_6P_3 = \dfrac{6!}{3!} = 120$ words.

Combination

A class has 25 students, and 6 of them are to be chosen to serve as student representative, how many different groups can be formed?

Notice in this example that there is no distinction between any of the students in the group, as long as the student is in the group, he/she is a class representative. This means that the order in which we choose the group of 6 does **not** matter.

If we have n many objects, and we choose r from the group, **and the order does not matter**, then each of such choice is called a **combination**.

The number of combinations (order does not matter) of n objects taken r at a time is given by

$$nCr = \binom{n}{r} = \frac{n!}{r!(n-r)!}$$

The notation nCr and the notation $\binom{n}{r}$ are equivalent and both meant the same thing. Namely the number of combinations of n objects taken r at a time. Both notation are read **n choose r**.

How many different groups of 6 numbers can be chosen from the numbers 1 through 49?

We have 49 numbers, and we want to choose 6 of them.

$$\binom{49}{6} = \frac{49!}{6!(49-6)!} = \frac{49!}{6!43!} = \frac{49 \cdot 48 \cdot 47 \cdot 46 \cdot 45 \cdot 44}{6!} = 13983816$$

There are 13983816 many different ways to choose 6 numbers from 49 numbers.

Sometimes, we may not have all the objects in a collection to be distinct, but instead represent two or more kinds of the same object. In this case care needs to be taken when we want to find the number of combinations or permutations.

Example:

A class of students has 10 girls and 12 boys. How many different ways can we form a group of five if 3 girls and 2 boys must be chosen?

From the group of 10 girls, we must choose 3, so we have $\binom{10}{3} = 120$ ways to choose these girls. From the group of 12 boys we must choose 2, so we have $\binom{12}{2} = 66$ choices. Using the multiplication principle we have $120 \cdot 66 = 7920$ ways to choose the group.

Example:

There are 11 kinds of donuts, 7 kinds of bagels, and 9 different favors of soda. If you want to buy 6 donuts, 3 bagels, and 3 sodas, how many different ways can you buy?

Clearly the order in which you buy the items is unimportant. You have to choose 6 donuts out of 11 kinds, so there's $\binom{11}{6} = 462$ ways to choose the donuts. You choose 3 bagels out of 7, so there's $\binom{7}{3} = 35$ ways to choose the bagels, and $\binom{9}{3} = 84$ ways to choose the sodas. Using the multiplication principle we have $462 \cdot 35 \cdot 84 = 1358280$ ways to buy our food.

Example:

Three biologists, four economists, and seven mathematicians are to sit in a row to watch a movie. If those of the same profession insist to be seated together, how many different ways can these people be seated?

Ans: The order in which these people seat is what distingushes one seating from another, so the order matters. Since the people of the same professions are to be seated together, and there are three professions, we have ${}_3P_3 = 3! = 6$ different ways to arrange the people of the same profession to seat together. For each of this arrangement, the people who are of the same profession can seat anyway among themselves. There are three biologists, so there are ${}_3P_3 = 6$ many ways the biologists can seat among themselves. Similarly, there are ${}_4P_4 = 4! = 24$ ways for the economists to seat among themselves, and ${}_7P_7 = 7! = 5040$ ways for the mathematicians to seat among themselves. Using the multiplication principle, there are

$(6)\,[6 \cdot 24 \cdot 5040] = 4354560$

ways for the professionals to sit.

Example: How many different *words* can be formed from the letters of the word *mississippi* ?

Ans: If all of the letters are different, then the multiplication principle tells us there are 11! many *words*. However, since some of the letters are the same, we have to account for the *words* that look the same with the same letters from different positions. In particular, sine there are 4 i 's, the arrangement among the i 's are the same and cannot be counted as different. There are 4! many arrangements among the 4 i 's, so our result must divide by 4!. Similarly, there are 4 s 's, so there are 4! (same) arrangements among the 4 s 's, and there are 2 p 's, so there are 2! (same) arrangements among the 2 p 's. We need to divide by these over-counted arrangements, and the result is:

$$\frac{11!}{4! \cdot 4! \cdot 2!} = 34650$$

In general, if there are n many objects to be arranged, but among these objects, n_1 of one kind is identical, and n_2 of another kind is identical, n_3 of third kind is identical ..., then the number of different permutations is given by:

$$\frac{n!}{n_1! \cdot n_2! \cdot n_3! \cdots n_k!}$$

Probability

Using the multiplication principle, and the formulas for combination and permutation, we can solve probability problems. But first we need to define a probability function.

Definition: An **experiment** is a procedure that produces an outcome. For example, the toss of a coin is an experiment (and the outcome would be head or tail). The rolling of a die is an experiment, and the outcome is 1, 2, 3, 4, 5, or 6.

The **sample space** of an experiment is the set that contains all **possible outcomes** of the experiment

Example, for the experiment of rolling a die, the *sample space* is the set $\{1, 2, 3, 4, 5, 6\}$. For the experiment of tossing a coin, the sample space would be the set $\{head, tail\}$.

If a sample space has finitely many elements, it is called a finite sample space. If it has infinitely many elements, it is an infinite sample space.

An **event** is a subset of the sample space. A **simple event** is an event of the sample space that contains only one element. For example, in the experiment of coin tossing, the sample space is $\{1, 2, 3, 4, 5, 6\}$. An event could be $\{2, 4, 6\}$ (getting an even number on the toss), and a simple event could be $\{3\}$ (getting the number three on the toss).

The **probability of a simple event** $\{e\}$ is defined as a number $P(\{e\})$ that satisfies the following two properties:

a. $0 \leq P(\{e\}) \leq 1$. That is, the probability of the event is always between 0 and 1, inclusive.

Intuitively, the number $P(\{e\})$ is the probability that a simple event (one particular outcome) $\{e\}$ will happen when the experiment is performed. The closer to 1, the more likely the event will happen. The closer to 0, the less likely the event will happen. If $P(\{e\}) = 1$, then event $\{e\}$ will always happen if the experiment is performed, if $P(\{e\}) = 0$, event $\{e\}$ will never happen.

b. The sum of the probabilites over all the simple events of the sample space is one. That is, if $\{e_1\}, \{e_2\}, \{e_3\}, ..., \{e_n\}$ are all the simple events of the sample space, then

$$P(\{e_1\}) + P(\{e_2\}) + P(\{e_3\}) + ... + P(\{e_n\}) = 1$$

This statement says that, among all the possible outcomes of an experiment, one of them must happen when the experiment is performed.

If A is an event, then the **probability of A**, written $P(A)$, is defined as the sum of the probabilities of the simple events of A.

If every simple event in a sample space is **equally likely** to occur, (we say that the outcomes are **random**, or that the experiment produces **random** outcomes), then the probability of any one simple event is equal to $\frac{1}{n}$, where n is the number of possible outcomes in the sample space. In general, if the sample space has n elements, and each element is equally likely to occur, and if A is an event of the sample space, then

$$P(A) = \frac{|A|}{n}$$

where $|A|$ is the order of A (number of elements in A).

That is, the probability of an event happening in which every outcome is equally likely, is equal to the number of outcomes of the event divided by the total number of possible outcomes.

Example: If a fair die is rolled, what is the probability that the number is greater than 4?

The event A here is *a number greater than 4*. This is the set $\{5, 6\}$. This set has 2 elements. The sample space $\{1, 2, 3, 4, 5, 6\}$ has 6 elements. The probability is therefore $= \frac{2}{6} = \frac{1}{3}$.

Example: If a fair die is rolled, what is the probability that the number is a 7?

The event A here is getting the *number 7*. The event is the empty set $\varnothing$. Since $|\varnothing| = 0$, the probability is $\frac{0}{6} = 0$. This means that this particular event will never happen.

If a fair die is rolled, what is the probability that the number is less than 7? The event A here is *number less than 7*, this is the set $\{1,\ 2,\ 3,\ 4,\ 5,\ 6\}$. This set has 6 elements. The sample space also has 6 elements. The probability of event A is $P(A) = \frac{6}{6} = 1$. This means that this event is destined to happen on any try of the experiment.

In calculating probability using the above formula, it is important that each outcome in the sample space is **equally likely**. We will use the terms like *fair die, fair coin*, or *choosing a number in random* to mean that each outcome of the experiment is equally likely.

Some basic formulas in probability:

When we use the symbol $P(A \,\mathrm{and}\, B)$ we mean the probability that A *and* B

will occur at the same time, this is read **probability of A and B**. The symbol $P(A \,\mathrm{or}\, B)$ means the probability that either A *or* B (or both) will occur, and is read **probability of A or B**. Sometimes we use $P(A \cap B)$ to mean $P(A \,\mathrm{and}\, B)$, and we use $P(A \cup B)$ to mean $P(A \,\mathrm{or}\, B)$.

Remember the formula used to count the size of a set,

$|A \cup B| = |A| + |B| - |A \cap B|$,

we have an equivalent version in probability:

Suppose A and B are two events, then

$$P(A \,\mathrm{or}\, B) = P(A) + P(B) - P(A \,\mathrm{and}\, B)$$

This formula says that the probability that A or B will occur is equal to the sum of their probabilities subtract the probability that they would *both* happen.

Two events are **mutually exclusive** if they cannot occur simutaneously. For example, when rolling a die, the event of *rolling an even number* and *rolling a 3* are mutually exclusive, since 3 is not an even number. In the case when two events A and B are mutually exclusive, $P(A \,\mathrm{and}\, B) = 0$, therefore, the formula for the probability of A or B reduces to:

If A and B are mutually exclusive events, then

$$P(A \,\mathrm{or}\, B) = P(A) + P(B)$$

Law of the Complement:

If A is any event, then the probability that A **does not** occur is equal to

$$1 - P(A)$$

That is, the probability that an event does not occur is equal to 1 minus the probability that the event will occur.

Example: A letter is randomly chosen from the letters of the word **Jacobsen**, what is the probability that it is a vowel?

Ans: There are 8 letters in the word, so the size of the sample space S is 8, the desirable event is *vowel*, which is the set $\{a, o, e\}$, which has a size of 3, using the probability formula, the probility of this is:

$\frac{3}{8}$

Example: Two fair dice are rolled, what is the probability that the sum of the two faces is greater than 9?

Ans: When two dice are rolled, there are all together 36 possible outcomes

(why?). Of these 36 outcomes, there are 6 of them whose sum is greater than 9, $\{(5,5),(4,6),(6,4),(5,6),(6,5),(6,6)\}$, so the probability of the event is:
$\frac{6}{36} = \frac{1}{6}$

Example: If you are dealt two cards randomly from a 52 card deck, what is the probability that you would be dealt a pair of Aces?

We first find out in how many different ways can two cards be dealt. There are a total of 52 cards, and we choose any two from it, so we have $\binom{52}{2} = 1326$ different ways two cards can be dealt from a 52 card deck. There's only 4 Aces, and we need to choose two from them, so there's $\binom{4}{2} = 6$ ways two Aces can be dealt. Taking the ratio of the two numbers we see that there's $\frac{6}{1326} \approx 0.0045 \approx 0.45\%$ chance that you will be dealt a pair of Aces.

Example:

A group of lawyers consist of 5 women and 7 men. If 4 lawyers are chosen randomly to represent a client, what is the probability that all of them will be women?

There are total of 12 lawyers, and we choose 4 of them, so there are $\binom{12}{4} = 495$ choices. We must choose 4 women out of the five women, so there are $\binom{5}{4} = 5$ ways to do so. There is $\frac{5}{495} \approx 0.010 \approx 1\%$ chance.

Example: The California lottary requires you to pick 5 numbers from the numbers 1 to 56, then pick a *mega number* from the numbers 1 to 46. In order to win the grand price, you must match all five of the numbers and the mega number. If you buy one ticket, what is your chance of winning?

Ans: We first calculate the size of the sample space, i.e., how many different 5-number (from 56) and 1-number (from 46) combinations are possible. For this problem, the order in which you pick the numbers do not matter, so this is a combination problem. There are 56 numbers and you need to pick 5, there are $\binom{56}{5}$ many ways to do that. Then you need to pick one number out of 46, there are $\binom{46}{1}$ many ways to do that. Together, there are $\binom{56}{5}\binom{46}{1} = 175,711,536$ many possible outcomes in the sample space. Since you buy only one ticket, the size of the event (the numbers of your ticket) is 1, so your chance of winning is:

$$\frac{1}{175,711,536}.$$

Conditional Probability

Sometimes, two events are related in such a way that the occurance of one event changes the probability of the occurance of another event. For example, the probability that a car would skid when breaking is 0.01%. However, during snow, the chance of a car skiding becomes 1%. If we let A be the event of "car skiding", and we let B be the event of "snow", then $P(A) = 0.01\%$, but the probability of A has changed if B has already occured. We use the notation $P(A|B)$ to mean *the probability of A given that B has already occured* (usually read as "Probability of A given B"). This is called a conditional probability because the probability relies on the fact that the condition B to have occured. In this case here, $P(A|B) = 1\%$. That is, the probability of a car skid is 1% when it has already snowed.

Formula for **conditional probability**:

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

provided that $P(B) \neq 0$.

This formula says that, the probability of A happening given that B has already happened is given by the *probability of A and B* divided by the *probability of B.*

Example: A group has 5 men and 7 women. Among the 5 men there are 3 athletes, and among the 7 women there are 4 athletes. If an individual is picked at random from the group, and it is known that he is a man. What is the probability that he is an athlete?

Since we already know that the person chosen is a man, and there's 3 athletes out of the five, so the probability that the person is an athlete is 3/5 or 60%. However, let us do it using the formula above:

Let A be the event of "person is an athlete".

Let B be the event of "person is a man".

Then we are interested in $P(A|B)$. What is $P(A \cap B)$? There's three male athlete out of 12 people, therefore the probability that someone is a man *and* an athlete is $P(A \cap B) = \frac{3}{12} = \frac{1}{4}$ The probability that a randomly chosen person is a man is $\frac{5}{12}$.

Using the above formula, we have

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{1/4}{5/12} = \frac{3}{5}$$

Example: The probability that a student will pass the midterm is 60%, and the probability that the student will pass the final is 53%. The probability that the student will pass both tests is 45%. If a student has already passed the midterm, what is the probability the she will pass the final exam?

Let A be the event of "passing the final"

Let B be the event of "passing the midterm"

$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{.45}{.60} \approx .75$

The **Multiplication Rule**:

From the formula for conditional probability, If we solve for the term $P(A \cap B)$, we get

$$P(A \cap B) = P(B) \cdot P(A|B)$$

Two events are said to be **independent events** if the outcome of one does not affect the outcome of another. For example, the events "rain" and "getting a call from an advertising agent" are two independent events, since the outcome of one should have no bearing on the other.

Mathematically, we can see that if two events A and B are independent, then the probability of one given that the other has occured should not change. That is,

Definition: Two events A and B are *independent events* if $P(A|B) = P(A)$ and $P(B|A) = P(B)$.

That is, two events are independent if the probability of one is not changed by the occurance of another.

In the multiplication rule, if two events are independent, then $P(A|B) = P(A)$ and this gives us:

If two events A and B are independent, then

$$P(A \cap B) = P(A) \cdot P(B)$$

That is, if two events are independent, then the probability that they both happen is just the product of the probability of them happening individually.

Example:

What is the probability that no head appears in ten consecutive tosses of a fair coin?

It would not be easy to find this probability directly. Instead, we use the formula for independent events to calculate *the probability that all tails show up on the ten tosses.* The probability that it is a tail on the first toss is 0.5, the probability that it is tail on the second toss is 0.5. In fact, the probability of the toss having a tail on any toss is 0.5. Since each toss is independent of the other, the probability that all ten tosses are tail is $P = (0.5)^{10} \approx 0.1\%$.

Example:

Birthday Paradox In a room of 40 people, what is the probability that at least two people will have the same birthday?

It is not easy to find this probability directly. Instead, we first calculate the *probability that no two people will have the same birthday.*

We use a 365 day calender year. The first person can have any birthday. In order *not* to have the same birthday as the first person, the second person's birthday must not be the same day as the first person, the probability of this happening is $\frac{364}{365}$ (since there are only 364 days on which the second person can have a birthday that is different than the first one). The third person's birthday must not be the same as the first two, and the probability of that happening is $\frac{363}{365}$. Since the birthday of each person is independent, the probability that none of them will have the same birthday would be

$$P = \frac{365}{365} \cdot \frac{364}{365} \cdot \frac{363}{365} \cdots \frac{326}{365}$$

$$= \frac{365 \cdot 364 \cdots 326}{365^{40}} \approx 0.109 \approx 11\%$$

The probability that no two people would have the same birthday is about 11%. Using the law of the complement, this means the probability that at least two people would have the same birthday is $1 - 0.109 \approx 89\%$.