MAT220 Lecture 4: Mathematical Foundations of Probability

Contents

In Lectures 1–3, we studied the basic language of statistics, visual descriptions of data, and numerical summaries such as the mean, median, variance and standard deviation. We now move from descriptive statistics towards probability theory. However, before getting there, we will first need to introduce some mathematical concepts that are used heavily throughout Probability Theory. In this lecture, we will do exactly this, by introducing the set theory and counting methods that will be used throughout later portions of our course.

1 Set Theory

1.1 What is a set?

In order to talk about a group of things mathematically, we need a way to write them down all together. That is what a set is: it is simply an unordered list of items.

Definition: Set

A **set** is an unstructured collection of things.

We can imagine a set as a mathematical box in which to put things. There is no inherent structure to the box, except that it exists and contains the things that it does. In fact, *that's it*: a set is simply the mathematical way to collect things together in one place. Given how open-ended this notion is, we can form sets from basically any objects. For example:

- the set of all numbers greater than 2,
- the set of all people in the classroom who have brown hair,
- the set of all students who arrived by train today,
- the set of all penguins named Carl.

We call the objects inside a set the *elements* of a set, they are like the members of the set. In mathematical notation, we write sets using the curly brackets $\{$ and $\}$. For example the notation:

$$A = \{1, 2, 3, 4, 5, 6, 7\}$$

is how we write the set A whose elements are $1, 2, 3, 4, 5, 6$ and 7. To emphasize: it is best to think of sets as having no inherent order. Of course, we have to list the elements of a set in *some* sort of order, however this does not matter at all: a set is only specified by its elements. For instance, the set A can be equally written as:

$$A = \{3, 4, 1, 7, 2, 5, 6\}, \qquad \text{or} \qquad A = \{3, 1, 4, 5, 7, 6, 2\}.$$

If an object x happens to be an element of the set A, then we write

$$x \in A,$$

which is simply read "x is in A". If x were *not* an element of A, then we would write a slash through the above symbol, similar to the way we write $\neq$. The notation:

$$x \notin A$$

reads "x is not in A". For example, if $A = \{1, 3, 5, 7, 9\}$, then $3 \in A$ but $4 \notin A$.

1.2 Rules for sets

Sets are the fundamental building block of almost all other mathematical structures, since before ever talking about the structure that a collection of things may or may not have, we first need to collect them all into the same place. In that sense, the theory of sets can be seen as a foundation of modern mathematics. However, although we may think of sets as a "proto-structure", there are still some basic ground rules.

Basic rules for sets

Rule 1: You cannot write the same element twice.
Rule 2: If two sets have all of the same elements, then they are equal.
Rule 3: Sets can have other sets as elements, but a set cannot contain itself.
Rule 4: There is a special set with absolutely nothing in it, called the empty set, written $\emptyset$.
Rule 5: There is no universal set containing absolutely everything.

Rule 1 states that we only write elements of a set *once*. This means, for example, the two sets below are equal:

$$\{1, 1, 2\} = \{1, 2\}$$

that is, when writing out a set we do not write any duplicates. In other words, a set does not count frequency of its members, only their presence. The second rule means that

$$\{1, 2, 3\} = \{3, 2, 1\} = \{1, 3, 2\},$$

which is another way of saying that the ordering of the elements within a set does not matter. Rule 3 allows us to write things like

$$A = \{1, 2, \{3, 4\}\},$$

which is a set containing *three* elements: the numbers 1 and 2, and the set $\{3, 4\}$. But, it does not allow us to write something like $A = \{A\}$, because this would lead to self-referential paradoxes. In symbols, Rule 3 forces the relation $A \notin A$ for all sets A. The fourth rule introduces the empty set, written $\emptyset$, which is the unique set containing nothing at all. This special set can be seen as a sort of zero element within set theory; a formal way to represent nothing. It should be noted that according to Rule 3, the empty set is not the same as the set containing the empty set:

$$\emptyset \neq \{\emptyset\},$$

since the first has zero elements and the second set has one element: $\emptyset$ itself. Finally, Rule 5 says that there is no set containing absolutely every mathematical object. This rule is imposed for technical reasons: if we had a set U that contained absolutely everything, then this set *is itself a thing*, meaning that by definition it ought to contain itself. But, this would contradict Rule 3.

Exercise: Equality of sets

Determine whether the following sets are equal.

(a) $A = \{1, 3, 5, 7, 9\}$ and $B = \{9, 7, 5, 3, 1\}$.
(b) $A = \{13, 6, 7, 9, 2\}$ and $B = \{6, 2, 13, 9\}$.
(c) $A = \{2, 4, 6, 8\}$ and $B = \{2, 2, 2, 2, 4, 6, 8, 8\}$.
(d) $A = \{14\}$ and $B = \{7 + 7\}$.

Solution

(a) Yes. The two sets have exactly the same elements, so $A = B$.
(b) No. The set A contains 7, but the set B does not. Therefore $A \neq B$.
(c) Yes. We ignore repeated copies of elements, so $A = B$.
(d) Yes. Since $7 + 7 = 14$, the two sets contain exactly the same element. Therefore $A = B$.

1.3 Cardinality

Since a set has no inherent structure, there is not much to talk about. However, what we can do is count up the elements contained within a set to get a notion of size. This "size of a set" has a special name: it is called the *cardinality* of a set.

Definition: Cardinality

For a finite set A, the **cardinality** of A is the number of elements in A. It is written

$$|A| .$$

For example,

$$|\{a, b, c, d, e\}| = 5, \qquad |\emptyset| = 0.$$

1.4 Subsets

As the name may suggest, a subset of a set A is some kind of smaller collection of things taken from A. Formally, we can define a subset as follows.

Definition: Subset

For any two sets A and B, we say that A is a **subset** of B, written

$$A \subseteq B,$$

if every element of A is also an element of B.

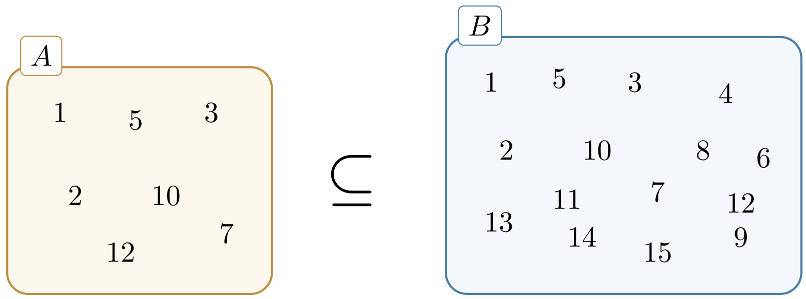
For example, the set

$$A = \{1, 2, 3, 5, 7, 10, 12\}$$

is a subset of the much larger set

$$B = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15\},$$

because every element contained in A also happens to be in B. In terms of structureless boxes, we can imagine a subset as follows.



The set A can be verified to be a subset of B by simply checking that each element of A is also present in B. Given that A has size 7, we can simply do this one-by-one.

There are two subset facts that are particularly important. For every set A, it is always the case that both

$$A \subseteq A \qquad \text{and} \qquad \emptyset \subseteq A.$$

The first statement says that every set is a subset of itself. This is always true, in some trivial sense: every element x in A is by default in A. The second statement above says that the empty set is a subset of every set, because there is no element of $\emptyset$ that could fail to be in A. This statement is often called *vacuously true*.

Exercise: Subsets

Suppose that

$$A = \{1, 2, 3, 4\}, \qquad B = \{2, 3, 5, 7, 9\},$$

and let S be the set of all prime numbers. Determine whether the following statements are true or false.

(a) $A \subseteq B$.
(b) $A \subseteq S$.
(c) $B \subseteq S$.
(d) $\{2, 3\} \subseteq B$.
(e) $A \subseteq A$.
(f) $\emptyset \subseteq A$.

Solution

(a) False. The elements 1 and 4 are in A, but they are not in B.
(b) False. The numbers 1 and 4 are not prime.
(c) False. The number 9 is in B, but 9 is not prime.
(d) True. Both 2 and 3 are elements of B, so $\{2, 3\} \subseteq B$.
(e) True. Every set is a subset of itself.
(f) True. The empty set is a subset of every set.

1.5 The powerset

Suppose that we have a set A, with n-many elements:

$$A = \{x_1, x_2, \ldots, x_n\}$$

We know from the previous section that we can form a subset of A by selecting some elements of A and putting them into a set, and then neglecting others if needs be. Starting from the bottom up, we know that for any element x in A, we can form subsets consisting of *only* this element: $\{x\} \subseteq A$. If we did this for every element of A, then we would create a lot of subsets of size 1:

$$\text{Subsets of size } 1 : \{x_1\}, \{x_2\}, \{x_3\}, \ldots, \{x_n\}.$$

We can also try to list out all of the subsets of size 2. This would consist of all pairs of elements:

$$\text{Subsets of size } 2 : \{x_1, x_2\}, \{x_2, x_3\}, \{x_1, x_3\}, \ldots.$$

There appears to be many more of these subsets, but in principle they can all be written down. In fact, we can repeat this process to create subsets of size $3, 4, 5$, and so on. Eventually this process will terminate, because A only has n elements, meaning that there is a unique subset of size n: A itself.

Hypothetically, after we were to complete this process, we would have written down *all* the subsets of A. To nobody's surprise, this collection can also be made into another set. We call this "set of all subsets of A" the *powerset* of A.

Definition: Powerset

The **powerset** of a set A, written $\mathcal{P}(A)$, is the set of all subsets of A.

For example, if

$$A = \{1, 2, 3\},$$

then

$$\mathcal{P}(A) = \big\{\emptyset, \{1\}, \{2\}, \{3\}, \{1, 2\}, \{1, 3\}, \{2, 3\}, \{1, 2, 3\}\big\}.$$

As you can see, the powerset seems to be much larger than the original. In the case above, the set $A = \{1, 2, 3\}$ has 3 elements, whereas its powerset has 8 elements. This is not a coincidence: it happens to be the case that $2^3 = 8$. Generally, we have the following rule.

Size of a powerset

If a finite set A has n elements, then

$$|\mathcal{P}(A)| = 2^n.$$

We will justify this rule later in the lecture.

2 Operations on Sets

2.1 Building new sets

There are a few ways to build new sets from old ones. Within the context of probability theory, we will focus on three main operations:

1. the complement of a set,
2. the intersection of sets, and
3. the union of sets.

As we will see in next lecture, these three operations will allow us to properly phrase (and therefore calculate) probabilities of events containing the words "not", "and" and "or". We will now review each of these operations briefly.

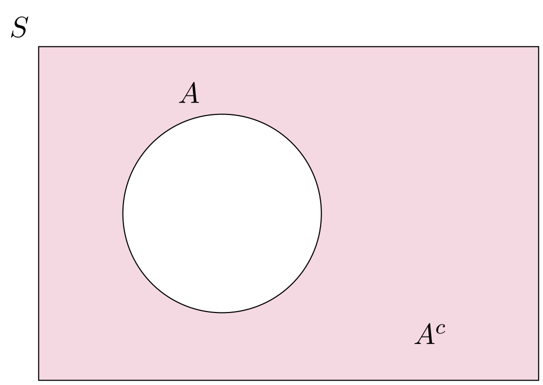
2.2 The Complement

The complement of A, sometimes written A^c, is simply the set of all things not in A. According to Rule 5, we cannot simply look at *all* the things not in A. Instead, we need to take these complements relative to some larger set S.

Definition: Complement

Let $A \subseteq S$. The **complement** of A in S is the set of all elements x in S that are not in A.

Visually, we can imagine the complement like the shaded area below:



For example, consider the set $S = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$. If we let A be the set of even numbers in S, i.e. $A = \{2, 4, 6, 8, 10\}$, then the complement of A in S will be all those numbers in S that are not in A. We can find these by writing out S and removing all the elements contained in A. Whatever is left will be the complement. Writing out S and removing the elements of A, we have

$$\{1, \cancel{2}, 3, \cancel{4}, 5, \cancel{6}, 7, \cancel{8}, 9, \cancel{10}\} \longrightarrow \{1, 3, 5, 7, 9\}.$$

Therefore the complement A^c equals the set $\{1, 3, 5, 7, 9\}$, i.e. the odd numbers in S.

2.3 The Intersection

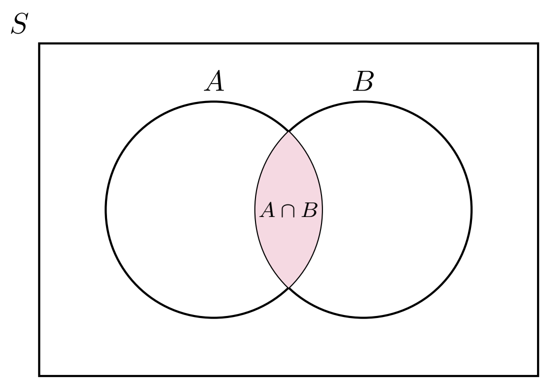
The intersection of A and B, sometimes written $A \cap B$, is the set of all things that are in A and B at the same time.

Definition: Intersection

The **intersection** of A and B is

$$A \cap B = \{x : x \in A \text{ and } x \in B\}.$$

Visually, the intersection of two sets might look like the shaded region below:



However, it should be noted that it's entirely possible for two sets to have an empty intersection. This would happen if the two circles in the diagram above did not overlap.

As an example, let $A = \{1, 2, 3, 4\}$ and $B = \{2, 3, 5, 7, 9\}$. For the intersection, we are looking for any elements that are in A and B at the same time. By inspection, we see that

$$A \cap B = \{2, 3\},$$

because 2 and 3 are the only elements that are in both sets.

2.4 The Union

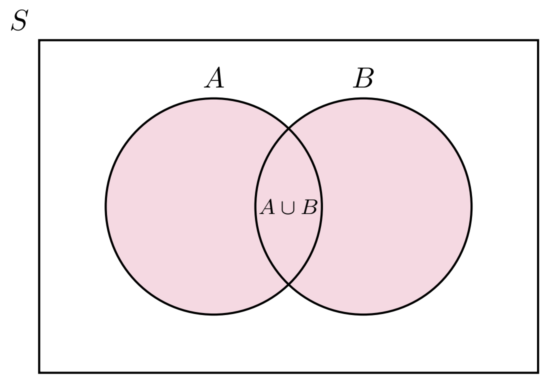
The union of A and B, sometimes written $A \cup B$, is the set of all things that are either in A or in B, or both. We can imagine this set as a set formed by joining together both A and B into one set.

Definition: Union

The **union** of A and B is

$$A \cup B = \{x : x \in A \text{ or } x \in B\}.$$

Visually, the union of two sets may look like the diagram below:



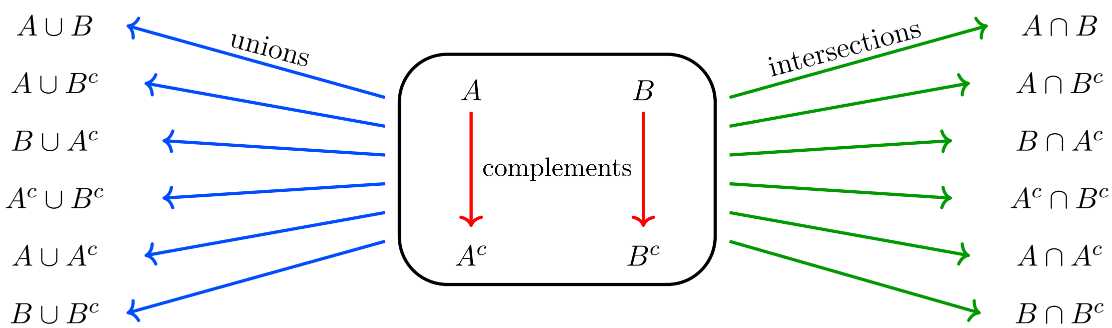
For example, suppose that $A = \{1, 2, 3\}$ and $B = \{2, 3, 4, 5\}$. Then the union will be the collection of all of these things together:

$$A \cup B = \{1, 2, 3, 4, 5\}.$$

Note that according to Rule 1 of Section 1.2, we **don't** write the elements 2 and 3 twice.

2.5 Combining operations

Using the three operations of complements, intersections and unions, we can build many other sets from a small amount of starting information. From just two sets A and B, both viewed as subsets of some larger set S, we can apply these three operations to create many new sets:



Exercise: Operations on sets

Let

$$S = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\},$$

and consider the three subsets

$$A = \{2, 4, 6, 8, 10\}, \qquad B = \{2, 3, 5, 7, 9\}, \qquad C = \{1, 3, 5, 7, 9\}.$$

Write out the following sets.

(a) A^c

(b) $A \cap B$

(c) $(A \cap B) \cup C$

(d) $A \cap C$

(e) $A \cap B \cap C$

Solution

(a) Since A is the set of even numbers in S, its complement is

$$A^c = \{1, 3, 5, 7, 9\}.$$

(b) The only element in both A and B is 2, so

$$A \cap B = \{2\}.$$

(c) Since $A \cap B = \{2\}$, we have

$$(A \cap B) \cup C = \{2\} \cup \{1, 3, 5, 7, 9\} = \{1, 2, 3, 5, 7, 9\}.$$

(d) There are no elements in both A and C, so

$$A \cap C = \emptyset.$$

(e) Since $A \cap C = \emptyset$, there cannot be anything in all three sets. Therefore

$$A \cap B \cap C = \emptyset.$$

2.6 Cardinality and operations

Set operations interact with cardinality in useful ways. Some basic facts are:

$$|A \cap B| \leq |A| \qquad \text{and} \qquad |A \cap B| \leq |B|.$$

This is because the intersection only contains elements that were already in both sets.

Similarly,

$$|A| \leq |A \cup B| \qquad \text{and} \qquad |B| \leq |A \cup B|.$$

This is because the union contains everything from A and everything from B.

However, when counting the size of $A \cup B$, we must avoid double-counting the overlap. This gives the formula

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

Example. Let

$$A = \{1, 2, 3, 4\}, \qquad B = \{3, 4, 5, 6, 7\}.$$

Then

$$|A| = 4, \qquad |B| = 5, \qquad A \cap B = \{3, 4\}.$$

So

$$|A \cap B| = 2.$$

Therefore

$$|A \cup B| = |A| + |B| - |A \cap B| = 4 + 5 - 2 = 7.$$

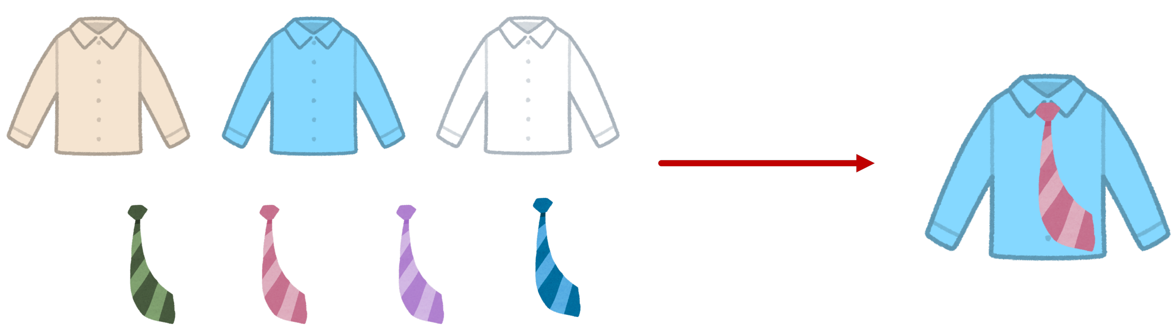
Indeed,

$$A \cup B = \{1, 2, 3, 4, 5, 6, 7\}.$$

3 Counting Configurations

3.1 The multiplication rule

Suppose that you have 3 shirts and 4 ties, and you want to wear one of each to go to work this morning. For example, one possible choice could be the following, the blue shirt and the red tie:



How many possible outfits can we make? Our problem of selecting an outfit can be broken down into two separate choices: selecting a shirt, and then selecting a tie independently. Picking a shirt first then offers us four possible outfits: the shirt plus any one of the four possible ties. Running through these choices for all the possible shirts, we end up with the following structured grid, which displays all possible configurations of shirt and tie.

As you can see, there are 12 total possible outfits, which is precisely the number of shirts (3) multiplied by the number of ties (4).

Instead of drawing pretty pictures of clothes, we can also convert the problem into mathematical language. If we label the three shirts by s_1, s_2, s_3, and we write the four ties as t_1, t_2, t_3, t_4, then an outfit would consist of a *pair* (s_i, t_j). The diagram above can be rewritten:

	t_1	t_2	t_3	t_4
s_1	(s_1, t_1)	(s_1, t_2)	(s_1, t_3)	(s_1, t_4)
s_2	(s_2, t_1)	(s_2, t_2)	(s_2, t_3)	(s_2, t_4)
s_3	(s_3, t_1)	(s_3, t_2)	(s_3, t_3)	(s_3, t_4)

$$3 \text{ shirt choices} \times 4 \text{ tie choices} = 12 \text{ total configurations.}$$

This grid-like structure of pairs is known as a *Cartesian product*. If we have two sets: $S = \{s_1, s_2, s_3\}$ and $T = \{t_1, t_2, t_3, t_4\}$ consisting of the shirts and ties, respectively, then we can "multiply" these two sets together to create the set of all pairs of the form (s_i, t_j). This new set, the Cartesian product of S and T, is written $S \times T$. What we have just discovered with our example of shirts and ties is that $|S \times T| = |S| \times |T|$. Within the context of counting configurations, we have the following general rule.

Multiplication Rule for Configurations

Suppose that a choice is made in stages. If the first stage has m options and the second stage has n options, then the total number of configurations is

$$m \times n.$$

For more than two stages, multiply the number of options at each stage.

Exercise: Counting configurations

You go to your favourite curry restaurant. They have a set deal where you choose:

- one of 5 starters,
- one of 15 main dishes,
- one of 5 desserts,
- one of 10 drinks.

How many unique meals can you create in total?

Solution

There are 5 choices for the starter, 15 choices for the main dish, 5 choices for the dessert, and 10 choices for the drink. By the multiplication rule, the total number of unique meals is

$$5 \times 15 \times 5 \times 10 = 3750.$$

So, there are 3750 unique meals in total. This means that you could visit every day for more than 10 years and never have the same meal twice.

4 Counting Permutations

In a moment we will introduce permutations and start counting them. However, in order to get there, we first need to introduce a new piece of notation.

Definition: Factorial

For a positive whole number n, the **factorial** of n is

$$n! = n(n-1)(n-2)\cdots 3 \cdot 2 \cdot 1.$$

We also define

$$0! = 1.$$

For example,

$$4! = 4 \times 3 \times 2 \times 1 = 24,$$

and

$$10! = 10 \times 9 \times 8 \times 7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1 = 3628800.$$

We stop at 1 because if we multiplied by 0, then every factorial would equal 0, which would be useless.

4.1 What is a permutation?

As we have seen in Section 1, a set of objects $A = \{a_1, a_2, \ldots, a_n\}$ does not care about the order that we write its elements in. However, we *can* decide to care about this, and write the elements of the set A out in different orders. To distinguish from the curly brackets of sets (which don't care about the order), we are going to write *ordered* lists using round brackets (and), so that, for example:

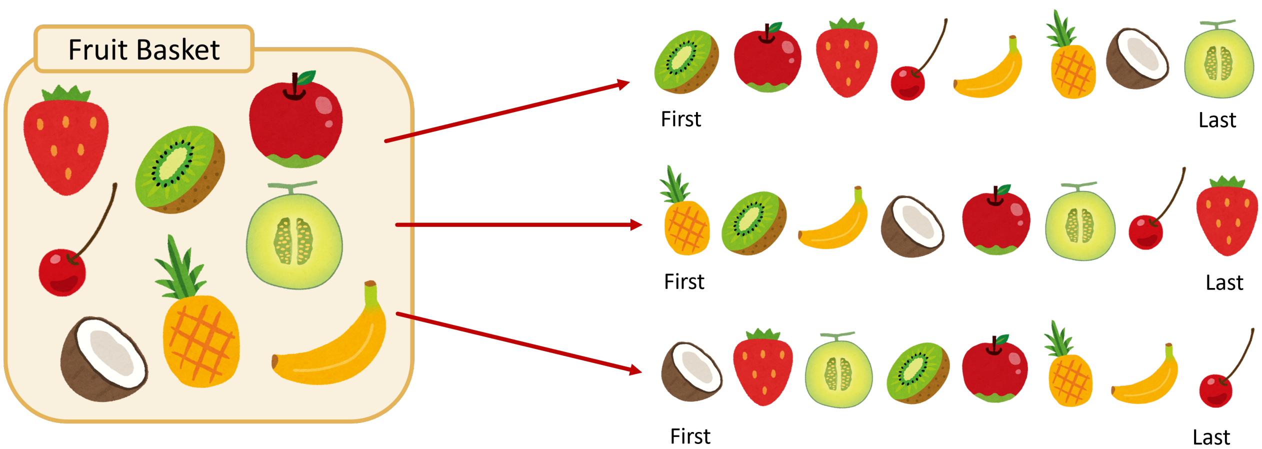
$$(a_1, a_2, \ldots, a_n) \qquad \text{and} \qquad (a_2, a_1, \ldots, a_n)$$

are considered different, even though they consist of the same n elements. Very simply, a permutation may be defined as follows.

Definition: Permutation

A **permutation** is an arrangement of objects where order matters.

As an example, suppose that we had a fruit basket consisting of 8 distinct pieces of fruit. If we were feeling particularly hungry, then we might want to eat every single piece of fruit in the basket. Since we have to eat them one at a time, there are many different possible orders. Three possible orders are listed below.



Three possible permutations of the 8 pieces of fruit.

4.2 Counting permutations

In order to try and understand how many permutations a given collection has, we can start with small numbers and see if we can notice a pattern. So, we start with $n = 1$ and keep increasing it:

$$n = 1: \quad 1 \text{ way},$$

$$n = 2: \quad 2 \times 1 = 2 \text{ ways},$$

$$n = 3: \quad 3 \times 2 \times 1 = 6 \text{ ways},$$

$$n = 4: \quad 4 \times 3 \times 2 \times 1 = 24 \text{ ways}.$$

Based on this, we see an obvious pattern emerging: we ought to use $n!$ to count the number of permutations of n objects. In fact, throughout this course we will use a slightly more general formula, which is given below.

General Rule for Permutations

The number of ways to arrange r objects in order from a collection of n distinct objects is

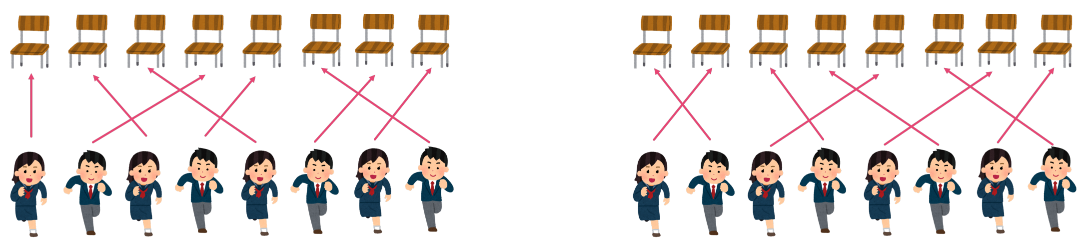
$$P_{n,r} = \frac{n!}{(n-r)!},$$

where n and r are whole numbers and $n \geq r$.

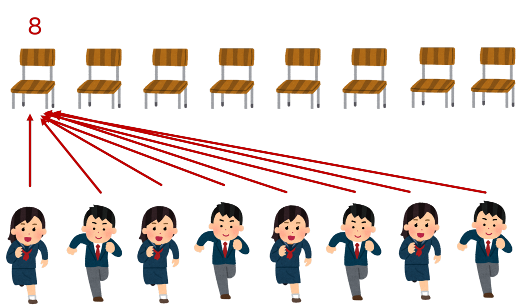
Upon first glance, this formula may appear to be quite random. So, we will now carefully justify it with an in-depth example.

4.3 Explaining the formulas using 8 students and 8 chairs

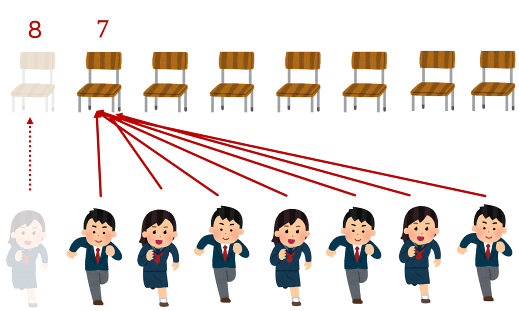
Suppose that we have a class with 8 students, and there are 8 chairs. We would like to create a "seating arrangement" by assigning the 8 students to the 8 chairs in the room. For example, the image below contains two possible seating arrangements.



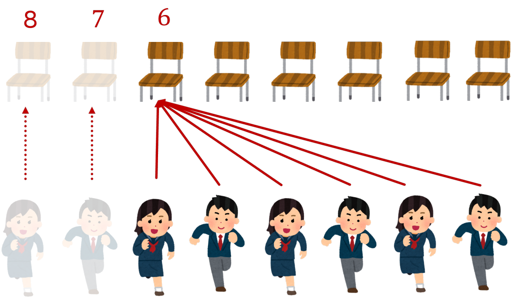
In effect, each assignment is the same as a permutation: we are taking the set of students and saying "you sit down first, you sit down second" and so on. So, the total number of seating arrangements ought to be $8! = 40320$. We will now explain this step-by-step.



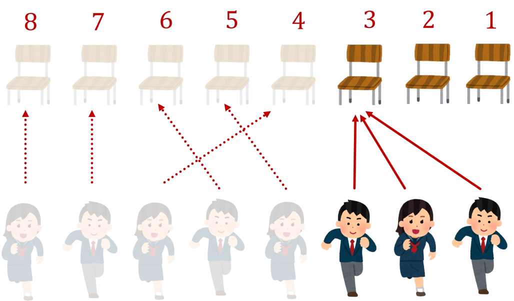
Step 1: The first chair. Since we start with 8 students, there are 8 possible choices for who can sit in the first chair.



Step 2: The second chair. When we try to repeat the process again for the second chair, we now only have 7 students and 7 empty chairs left. So, instead of there being 8 possible choices, there are only 7.



Step 3: The third chair. After we choose one of the 7 students to sit down in the second chair, we can then move on to the third. In this case, we have now already assigned *two* students, meaning that there are only 6 students who can possibly sit in the third chair. In other words, there are now 6 possible choices for the third chair.



Step 4: Continue the pattern. We may keep repeating this reasoning over and over again to get 5, then 4, then 3, then 2, and finally 1 possible choice for the remaining chairs.

In order to get the total number of possible configurations, we may use the multiplication rule for this 8-step process to conclude that the total number of seating arrangements is

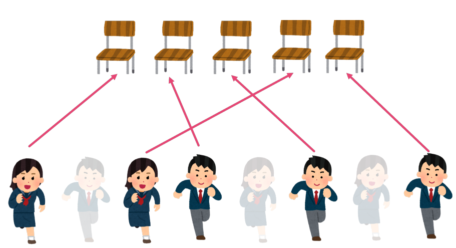
$$8 \times 7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1 = 8!.$$

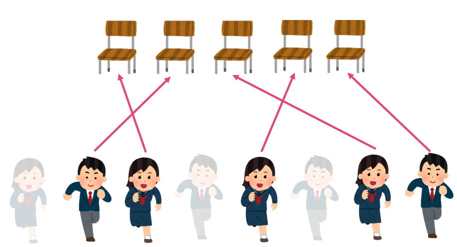
This recovers the factorial formula mentioned previously. But, what about the formula

$$P_{n,r} = \frac{n!}{(n-r)!}$$

mentioned previously? We will now explain this by modifying our setup.

Suppose now that we remove 3 chairs from the classroom, meaning that we only have 5 chairs remaining. So, if we try to sit 8 students down, then only 5 of them will be able to sit down and the other 3 will have to stand. Our next question is: what is the number of possible ways that we can sit 5 of the 8 students down? Two such arrangements are depicted below.





We can repeat the same reasoning as before, so we see that the first chair has 8 choices, the second chair has 7 choices and so on. However, since there are now only 5 possible chairs to fill, we must stop our procedure *early*, after 5 chairs. This means that our answer will not be 8!, but instead it will be given by the first 5 terms in the derivation from before:

$$P_{8,5} = 8 \times 7 \times 6 \times 5 \times 4.$$

Here $r = 5$, and we observe that we can write the same formula as:

$$P_{8,5} = \frac{8 \times 7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1}{3 \times 2 \times 1}.$$

This is the same formula as that mentioned previously. The idea here is that the numerator of $P_{n,r}$ tells us how to order n total objects, and the denominator $(n - r)!$ tells us how to order the $(n - r)$-many objects that we did not select. We divide by this number in order to remove them from the count. In the case of the chairs above, the $(n - r)!$ term corresponds to the 6 ways to order the 3 students who remain standing.

5 Counting Combinations

When discussing permutations, we cared about the order in which we arranged the students. That is, we lined up all the chairs in order and assigned different students by saying "you go first, you go second, . . .". We can imagine that there might be another way to do things: more generally, perhaps we could simply select r-many students and not care in which order they are placed in chairs. This would amount to simply choosing 5 students from the collection of 8. Two such choices are pictured below.

In fact, we can also count up "all the possible ways to choose r-many objects from n-many". These choices aren't permutations, because in this situation we don't care about their order. Instead, the choices are just picking some members of the group.

As a matter of fact, there is a formula for all the possible choices of r-many objects amongst a collection of n-many. In mathematics, we call all of these unordered choices "combinations", and we use the notation $C_{n,r}$ to denote the number of ways to choose r-many things from a collection of size n. The formula for $C_{n,r}$ is very similar to that of $P_{n,r}$, and it is given below.

General Rule for Combinations

The number of combinations of n objects taken r at a time is

$$C_{n,r} = \frac{n!}{r!(n - r)!},$$

where n and r are whole numbers and $n \geq r$.

Notice that here we have the formula for $P_{n,r}$ but we divide by an extra $r!$ term. Why do we do this? Well, in the case of $P_{n,r}$ we are asking for ordered selections, and the same choice of r-many objects can be ordered in $r!$ many different ways. However, for unordered combinations, we don't actually care about those $r!$ many things, so we divide the overall formula by $r!$ to remove all of those choices from our count.

To illustrate the formula, we may calculate the number of ways to select 5 students from the group of 8:

$$C_{8,5} = \frac{8 \cdot 7 \cdot 6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1}{(5 \cdot 4 \cdot 3 \cdot 2 \cdot 1)(3 \cdot 2 \cdot 1)}.$$

Of course, this can be put directly into a calculator. However, we can also observe that a factor of 5! exists on the top and bottom of this particular fraction. Therefore, we can cancel these to get:

$$C_{8,5} = \frac{8 \cdot 7 \cdot 6}{3 \cdot 2 \cdot 1} = \frac{336}{6} = 56.$$

So, there happen to be 56 ways to choose 5 students from 8 students.

5.1 Pascal's Triangle

As a matter of fact, the formulas $C_{n,r}$ happen to satisfy some interesting mathematical properties. We will start with two very simple observations that are worth remembering:

$$C_{n,0} = 1 \qquad \text{and} \qquad C_{n,n} = 1.$$

These equalities can be easily verified by inspecting the formulas algebraically. We have

$$C_{n,0} = \frac{n!}{0!(n-0)!} = \frac{n!}{0!n!} = 1,$$

because $0! = 1$, and similarly,

$$C_{n,n} = \frac{n!}{n!(n-n)!} = \frac{n!}{n!0!} = 1.$$

Heuristically speaking, there is a clear interpretation: there is exactly one way to select nothing from n objects, and there is also exactly one way to select everything.

Perhaps more interestingly, for $1 \leq r \leq n$, these $C_{n,r}$ formulas also satisfy the identity

$$C_{n,r-1} + C_{n,r} = C_{n+1,r},$$

which is commonly called *Pascal's Identity.* For the unconvinced, we may verify this algebraically.

Algebraic verification. We calculate:

$$C_{n,r-1} + C_{n,r} = \frac{n!}{(r-1)!(n-r+1)!} + \frac{n!}{r!(n-r)!}.$$

To add these fractions, use the common denominator

$$r!(n-r+1)!.$$

Then

$$\begin{aligned} C_{n,r-1} + C_{n,r} &= \frac{rn!}{r!(n-r+1)!} + \frac{(n-r+1)n!}{r!(n-r+1)!} \\ &= \frac{rn! + (n-r+1)n!}{r!(n-r+1)!} \\ &= \frac{(n+1)n!}{r!(n-r+1)!} \\ &= \frac{(n+1)!}{r!((n+1)-r)!} \\ &= C_{n+1,r}. \end{aligned}$$

There is also a simple counting explanation. To choose r objects from $n+1$ objects, focus on one special object. Either the special object is chosen, or it is not chosen.

- If the special object is chosen, then we still need to choose $r-1$ objects from the other n objects. This gives $C_{n,r-1}$ possibilities.
- If the special object is not chosen, then we need to choose all r objects from the other n objects. This gives $C_{n,r}$ possibilities.

Together, these two cases give

$$C_{n,r-1} + C_{n,r} = C_{n+1,r}.$$

For low values of n and r, we can arrange the values of $C_{n,r}$ into a triangle. At the top of the triangle is $C_{0,0} = 1$, the next row is $C_{1,0} = 1$ and $C_{1,1} = 1$. But, things start to get interesting once we add more rows. In the figure below, the values of $C_{n,r}$ are written in the form of a triangle.

$$\begin{array}{lccccccccccc}
\text{Row 0:} & & & & & & 1 & & & & & \\
\text{Row 1:} & & & & & 1 & & 1 & & & & \\
\text{Row 2:} & & & & 1 & & 2 & & 1 & & & \\
\text{Row 3:} & & & 1 & & 3 & & 3 & & 1 & & \\
\text{Row 4:} & & 1 & & 4 & & \color{red}{6} & & 4 & & 1 & \\
\text{Row 5:} & 1 & & 5 & & 10 & & 10 & & 5 & & 1 \\
\vdots & & & & & & \vdots & & & & &
\end{array}$$

$$\begin{array}{ccccccccccc}
 & & & & & C_{0,0} & & & & & \\
 & & & & C_{1,0} & & C_{1,1} & & & & \\
 & & & C_{2,0} & & C_{2,1} & & C_{2,2} & & & \\
 & & C_{3,0} & & C_{3,1} & & C_{3,2} & & C_{3,3} & & \\
 & C_{4,0} & & C_{4,1} & & \color{red}{C_{4,2}} & & C_{4,3} & & C_{4,4} & \\
C_{5,0} & & C_{5,1} & & C_{5,2} & & C_{5,3} & & C_{5,4} & & C_{5,5} \\
 & & & & & \vdots & & & & &
\end{array}$$

Observe that each interior number in the triangle is given by the sum of the two numbers above it, this is Pascal's Identity at work. For example, the "6" in Row 4 (coloured red) is given by the sum of the "3" and the "3" that are above it in Row 3. By applying this addition rule, we can form as many new rows of the triangle as we like.

Observe that each row is numbered by the "n" in $C_{n,r}$, and the entries for that row are given by running across the different values of r between 0 and n. This means that we don't need to calculate the value of $C_{n,r}$ directly, instead we can simply read off its value using the triangle.

5.2 Counting subsets with combinations

There is an important connection between combinations and powersets. Suppose A is a set with n elements. If we want to count the number of subsets of A with exactly r elements, then we are choosing r elements from n elements, and order does not matter. Therefore, the number of subsets of size r is

$$C_{n,r}.$$

Example. Let

$$A = \{a, b, c, d\}.$$

This set has size 4. We can count its subsets according to their size:

- $C_{4,0} = 1$ counts the single subset of size 0, namely $\emptyset$.
- $C_{4,1} = 4$ counts the subsets $\{a\}, \{b\}, \{c\}, \{d\}$.
- $C_{4,2} = 6$ counts the subsets $\{a, b\}, \{a, c\}, \{a, d\}, \{b, c\}, \{b, d\}, \{c, d\}$.
- $C_{4,3} = 4$ counts the subsets $\{a, b, c\}, \{a, b, d\}, \{a, c, d\}, \{b, c, d\}$.
- $C_{4,4} = 1$ counts the single subset of size 4, namely A itself.

Adding these together gives

$$1 + 4 + 6 + 4 + 1 = 16 = 2^4.$$

This agrees with the powerset rule:

$$|\mathcal{P}(A)| = 2^4 = 16.$$

This connection also explains why the rows of Pascal's triangle add to powers of 2.

$$\begin{array}{l c}
\text{Row 0:} & 1 = 1 \\
\text{Row 1:} & 1 + 1 = 2 \\
\text{Row 2:} & 1 + 2 + 1 = 4 \\
\text{Row 3:} & 1 + 3 + 3 + 1 = 8 \\
\text{Row 4:} & 1 + 4 + 6 + 4 + 1 = 16 \\
\text{Row 5:} & 1 + 5 + 10 + 10 + 5 + 1 = 32 \\
\vdots & \vdots \qquad \vdots
\end{array}$$