# 3 Combinations

In our discussion on permutations with indistinguishable elements, we arrived at a general formula by dividing the total number of permutations by the number of ways we could permute only the indistinguishable elements. We did so in order to avoid overcounting "identical" permutations.

If we extend the idea of indistinguishability to mean "indistinguishable up to ordering", then we obtain a new counting technique where we care only about the number of elements we take and not the order in which those elements are arranged. Instead of permutations of elements, we are taking **combinations** of elements.

## 3.1 Definition

We can define the notion of a combination formally, just like we defined permutations formally in the previous section. Both permutations and combinations rely on sets, so we require a little set theory knowledge. However, you've likely known what combinations were ever since you first learned about sets: combinations are just subsets in disguise.

**Definition 24** ($k$-combination)**.** Given a set $A$, a $k$-combination of $A$ is a size-$k$ subset of elements from $A$.

Since combinations are subsets, and since the arrangement of elements in a set doesn't matter, the arrangement of elements in combinations doesn't matter. This is the key distinction between permutations and combinations that we alluded to in the previous section: ordering matters for permutations, but not for combinations.

Further note that we always use the term ***k*-combination** when referring to a specific value or calculation. There is no such thing as a "combination of a set", since if we followed the same distinction between permutation and $k$-permutation, a "combination of a set" would just be the set itself. Thus, when we use the word "combination" on its own, we mean it in the broad, non-formal sense of selecting elements from a set without ordering.

**Example 25.** Suppose we have a set $A = \{1, 2, 3, 4\}$. All of the possible 2-combinations of $A$ are $\{1, 2\}$, $\{1, 3\}$, $\{1, 4\}$, $\{2, 3\}$, $\{2, 4\}$, and $\{3, 4\}$; exactly the same as all of the possible size-2 subsets of $A$.

It is incorrect for us to consider both $\{1, 2\}$ and $\{2, 1\}$ to be 2-combinations of $A$, since they both contain the same subset of elements from $A$. Therefore, we only count that particular 2-combination once.

Now that we know what combinations are, how can we count all possible $k$-combinations of an $n$-element set? This problem sounds very similar to our previous problem of counting all $k$-permutations of an $n$-element set; indeed, we can take almost the exact same approach we took when counting $k$-permutations. The only difference is that, since ordering doesn't matter with combinations, we need to include one additional term to guard against overcounting.

**Theorem 26.** *The number of $k$-combinations of a set with $n$ elements, where $0 \leq k \leq n$, is*

$$C(n, k) = \frac{n!}{k!(n-k)!}.$$

*Proof.* We can count the number of $k$-combinations of a set with $n$ elements by first calculating the number of $k$-permutations of the same set, and then dividing by the number of permutations of a $k$-element set. The division is necessary because the ordering of the $k$ elements does not matter. Thus, we have

$$\begin{aligned} C(n, k) &= \frac{P(n, k)}{P(k, k)} \\ &= \frac{n!/(n-k)!}{k!/(k-k)!} \\ &= \frac{n!}{k!(n-k)!}. \end{aligned}$$

□

Note that, as a consequence of Theorem 26, we have $C(n, 0) = 1$, $C(n, 1) = n$, and $C(n, n) = 1$ for all $n \geq 0$. The values $C(n, 0)$ and $C(n, n)$ should make sense from a set-theoretic standpoint, since for any $n$-element set $A$, there is only one zero-element subset ($\emptyset$) and only one $n$-element subset (the set $A$ itself).

*Remark.* You might have noticed that part of the previous proof looked similar to the proof of Theorem 22. This was not by coincidence; it's possible to prove Theorem 22 using $C(n, k)$ instead of $P(n, k)$. Try it!

Before we continue, it is worthwhile to point out an interesting symmetry property of combinations that may help us to solve some problems. In a $k$-combination, we take $k$ elements from an $n$-element set. However, this is no different from us taking $(n - k)$ elements and leaving them out of our final choice. From this observation, we get the aforementioned property.

**Theorem 27.** *For all natural numbers $n$ and $k$, where $0 \leq k \leq n$,*

$$C(n, k) = C(n, n - k).$$

*Proof.* Recall that $C(n, k) = \frac{n!}{k!(n-k)!}$. Substituting $(n - k)$ for $k$, we get

$$\begin{aligned} C(n, n-k) &= \frac{n!}{(n-k)!(n-(n-k))!} \\ &= \frac{n!}{(n-k)!k!}, \end{aligned}$$

and hence $C(n, k) = C(n, n - k)$. □

A common question to hear from students by this point is "how can we tell whether we need to calculate permutations or combinations in a problem?" It's certainly a reasonable question to ask, and you won't be faulted for wondering this yourself. Unfortunately, there is no surefire trick for determining which counting technique to use, apart from determining whether the problem statement emphasizes ordering of elements. Thus, if a question asked

> In how many ways can we choose 3 faculty members from a department of 10 faculty members?

then we would use combinations, since we care about only the number of faculty members. If, on the other hand, a question asked

> In how many ways can we line up 10 faculty members for a department photo?

then we would use permutations, since we care about both the number and the ordering of faculty members.

When in doubt, remember this mnemonic: **c**ombinations are for **c**hoosing some number of elements, and **p**ermutations are for **p**lacing those elements in a specific order.

**Example 28.** A group of five students are begging their discrete mathematics instructor to come up with examples that don't involve writing or scheduling exams. If the group chooses three representatives to talk with the instructor during office hours, how many possible combinations of representatives are there?

Suppose we call the students Alice, Bob, Carol, David, and Eve. If, for instance, Alice, Bob, and Carol attend the office hours, then that combination is no different than if Carol, Bob, and Alice attend; the representatives are the same. Altogether, we have the following subsets of three representatives each:

ABC, ABD, ABE, ACD, ACE, ADE, BCD, BCE, BDE, CDE.

In other terms, we have that $C(5, 3) = \frac{5!}{3!(5-3)!} = \frac{5!}{3!2!} = 10$.

**Example 29.** Recall that a byte consists of eight binary digits. Call a byte "balanced" if it contains an equal number of `0` bits and `1` bits.

How many balanced bytes exist? We can frame this problem in the following way. Since a balanced byte contains an equal number of `0`'s and `1`'s, we must have four occurrences of each bit within a balanced byte.

Thus, if we start with a blank byte with eight spaces, and we fill four of those spaces with `0`'s, then we are forced to fill the remaining four spaces with `1`'s.

In how many ways can we fill four spaces with `0`'s? This question is equivalent to asking in how many ways we can choose four spaces out of eight. This gives $C(8,4) = \frac{8!}{4!(8-4)!} = \frac{8!}{4!4!} = 70$, so there exist a total of 70 balanced bytes.

## 3.2 Combinations with Repetition

Just like with permutations, it is possible for us to calculate the number of $k$-combinations of a set when we are able to select elements from the set more than once. Unlike $k$-combinations without repetition, here we are able to select more copies of elements in our $k$-combination than those that are in the original set. Thus, a $k$-combination with repetition is not necessarily a subset of the original set.

How do we illustrate the process of taking a $k$-combination with repetition? Instead of us taking and replacing elements of the set to form the $k$-combination, we will work "in reverse" by writing out how many copies of each element our $k$-combination will contain. We represent this scenario in a uniquely American style: using the so-called **stars and bars** method. (Interestingly, the mathematician who popularized this method—William Feller—was born in Croatia, not America.)

With the stars and bars method, we partition our set into classes using bars (denoted $|$). Each class corresponds to a distinct element in the set. We then represent the number of copies of elements in that class to be included in our $k$-combination using the appropriate number of stars (denoted $\bigstar$).

**Example 30.** A student is enrolling in courses for the upcoming academic year. They plan to enrol in 10 courses. If the courses are to be selected from the set {CISC, MATH, PHYS, BIOL, CHEM}, then we can partition these five classes of course codes using four bars:

$$\underbrace{\quad}_{\text{CISC}} \,|\, \underbrace{\quad}_{\text{MATH}} \,|\, \underbrace{\quad}_{\text{PHYS}} \,|\, \underbrace{\quad}_{\text{BIOL}} \,|\, \underbrace{\quad}_{\text{CHEM}}$$

The student wants to take five CISC courses, three MATH courses, one PHYS course, and one CHEM course. If we denote one course by one star, our "stars and bars diagram" will look like the following:

$$\bigstar\bigstar\bigstar\bigstar\bigstar \,|\, \bigstar\bigstar\bigstar \,|\, \bigstar \,|\, \,|\, \bigstar$$

As the previous example illustrates, the number of elements in the set and the number of selections we make dictate the number of stars and bars that are available to us. If our set contains $i$ elements and we wish to make $j$ selections, then we will have $j$ stars and $(i-1)$ bars. As we also saw in the previous example, it is perfectly fine to have zero stars in a given partition; this just means we made no selections of elements from the corresponding class.

Using stars and bars in this way, it becomes evident that the number of ways to form a $k$-combination with repetition from a set of $n$ elements is exactly the same as the number of ways to arrange $k$ stars and $(n-1)$ bars in a row; that is, $(k+n-1)!$. However, since the stars and bars are indistinguishable, we must divide by both $k!$ and $(n-1)!$ to avoid overcounting.

**Theorem 31.** *The number of $k$-combinations of a set with $n$ elements, with repetition, is*

$$\frac{(k+n-1)!}{k!(n-1)!}.$$

*Proof.* Let $A = \{a_1, a_2, \ldots, a_n\}$. Consider a string of $(k+n-1)$ blank spaces

$$\underbrace{\sqcup\ \sqcup\ \sqcup\ \cdots\ \sqcup\ \sqcup\ \sqcup}_{k+n-1 \text{ times}}$$

and a set containing $k$ $\bigstar$s and $(n-1)$ $|$s. Each arrangement of $\bigstar$s and $|$s into the blank spaces constitutes a $k$-combination with repetition, with the number of $\bigstar$s between the start of the string and the first $|$ counting the number of selections of $a_1$, the number of $\bigstar$s between the first $|$ and second $|$ counting the number of selections of $a_2$, and so on. There is a total of $C(k+n-1, n-1)$ ways to place the $(n-1)$ $|$s into the blank spaces, and from this we force placement of the $k$ $\bigstar$s. Thus, there is a total of $C(k+n-1, n-1) = C(k+n-1, k) = \frac{(k+n-1)!}{k!(n-1)!}$ possible $k$-combinations of a set with $n$ elements where repetition is allowed. □

Returning to our course enrolment example, we see that if the student had no constraints on the 10 courses they wanted to take, then they would have a total of $C(10+5-1, 10) = C(14, 10) = C(14, 4) = 1001$ course combinations to choose from. We can use $k$-combinations with repetition to calculate many other interesting things; consider, for example, a computer algebra system that needs to find solutions to a given equation. The system could naïvely check every possible solution, but this can be slow. Using combinations, certain posibilities can be ruled out based on constraints or other conditions.

**Example 32.** Let $x$, $y$, and $z$ be natural numbers. How many solutions exist for the equation $x+y+z = 16$?

We can frame this problem as a combinatorial problem, since any solution to the given equation corresponds to a selection of 16 elements from a set of size 3 where we have $x$ elements from class 1, $y$ elements from class 2, and $z$ elements from class 3. Drawing a "stars and bars diagram", we have the following scenario:

$$\underbrace{\qquad}_{x} \,|\, \underbrace{\qquad}_{y} \,|\, \underbrace{\qquad}_{z}$$

Since we are calculating the number of 16-combinations of a set with 3 elements, we get that the total number of solutions to the equation is $C(16+3-1, 16) = C(18, 16) = C(18, 2) = 153$.

**Example 33.** Let $x$, $y$, and $z$ be natural numbers, this time with the constraints that $x \geq 2$, $y \geq 5$, and $z \geq 3$. How many solutions exist for the equation $x+y+z = 16$?

This example is very similar to the previous example, but with added constraints. Thus, we can follow the same procedure as before while keeping in mind that each of $x$, $y$, and $z$ must take on certain values; namely, we must have at least two elements from class 1, at least five elements from class 2, and at least three elements from class 3. Drawing a "stars and bars diagram", we have the following scenario:

$$\underbrace{\bigstar\bigstar}_{x} \,|\, \underbrace{\bigstar\bigstar\bigstar\bigstar\bigstar}_{y} \,|\, \underbrace{\bigstar\bigstar\bigstar}_{z}$$

Since ten stars are preassigned to the diagram, we must place the remaining six stars ourselves. This is equivalent to us calculating the number of 6-combinations of a set with 3 elements, which leads us to conclude that the total number of constrained solutions to the equation is $C(6+3-1, 6) = C(8, 6) = C(8, 2) = 28$.

# 4 Binomial Theorem

Let's now take a brief step back from counting and look at an algebraic problem: expanding binomials. As you likely learned in your first-year math classes (or even earlier), we can use the FOIL method—first, outer, inner, last—to expand the binomial $(x+y)^2$. This gives us the following result:

$$\begin{aligned}(x+y)^2 &= (x+y)(x+y) \\ &= x^2 + xy + xy + y^2 \\ &= x^2 + 2xy + y^2.\end{aligned}$$

Can we follow a similar technique for binomials with larger exponents? Of course; the FOIL method is just a special case of the distributive property of multiplication, which tells us that $a(b+c) = (ab+ac)$ for values

$a$, $b$, and $c$. To see how this works, let's consider $(x+y)^3$:

$$\begin{aligned}(x+y)^3 &= (x+y)(x+y)(x+y)\\ &= (x^2+2xy+y^2)(x+y)\\ &= x^3+x^2y+2x^2y+2xy^2+xy^2+y^3\\ &= x^3+3x^2y+3xy^2+y^3.\end{aligned}$$

Just for fun, let's also consider $(x+y)^4$ while we're at it:

$$\begin{aligned}(x+y)^4 &= (x+y)(x+y)(x+y)(x+y)\\ &= (x^3+3x^2y+3xy^2+y^3)(x+y)\\ &= x^4+x^3y+3x^3y+3x^2y^2+3x^2y^2+3xy^3+xy^3+y^4\\ &= x^4+4x^3y+6x^2y^2+4xy^3+y^4.\end{aligned}$$

By this point, you might notice a certain pattern is developing. Each term in the expansion of the binomial $(x+y)^n$ is of the form $ax^by^c$, where $b+c=n$ and where $a$ is some coefficient.

How can we calculate the value of the coefficient $a$ for some arbitrary term without writing the entire expansion? Let's begin by determining what this value $a$ is counting. We can write the general binomial $(x+y)^n$ as

$$(x+y)^n = \underbrace{(x+y)(x+y)\cdots(x+y)}_{n \text{ times}}$$

By the distributivity property of multiplication that we mentioned earlier, the expansion of this binomial will contain one term for each possible choice of $x$ and $y$, and we take $n$ choices. As an illustration, assume we are only choosing $x$. If we take $x$ a total of $n$ times, then we will add the term $x^n$ to our expansion. On the other hand, if we take $x$ a total of $n-4$ times, then we must take $y$ a total of 4 times in order to collect $n$ terms overall. Thus, we will add the term $x^{n-4}y^4$ to our expansion.

Generalizing this idea to us choosing $x$ a total of $b$ times and $y$ a total of $c$ times, where $b+c=n$, we see that the idea is equivalent to us calculating the number of ways we can choose $b$ occurrences of $x$ from $n$ binomials (equivalently, choosing $c$ occurrences of $y$). In other words, we're taking a $b$-combination from a set of binomials of size $n$ (equivalently, a $c$-combination), and therefore, the value $a$ is equal to $C(n,b)=C(n,c)$.

Before we continue, we will introduce a new notation used specifically in the context of binomials. We say that the **binomial coefficient** $\binom{n}{k}$ is the number of ways to choose $k$ elements from an $n$-element set, where $0 \le k \le n$. Sound familiar? It should; the binomial coefficient is exactly the same as a $k$-combination, but written using a different notation.

**Definition 34** (Binomial coefficient)**.** The binomial coefficient $\binom{n}{k}$, read as "$n$ choose $k$", is defined for $0 \le k \le n$ as

$$\binom{n}{k} = C(n,k) = \frac{n!}{k!(n-k)!}.$$

Now that we are familiar with binomial coefficients, we may use this notation to obtain the general form of a binomial expansion. We obtain the general form by way of the **binomial theorem**.

**Theorem 35** (Binomial theorem)**.** *Let $x$ and $y$ be variables, and let $n$ be a natural number. Then*

$$\begin{aligned}(x+y)^n &= \sum_{i=0}^{n}\binom{n}{i}x^{n-i}y^i\\ &= \binom{n}{0}x^n + \binom{n}{1}x^{n-1}y + \binom{n}{2}x^{n-2}y^2 + \cdots + \binom{n}{n-1}xy^{n-1} + \binom{n}{n}y^n.\end{aligned}$$

*Proof.* We prove by induction. Let $P(n)$ be the statement "$(x+y)^n = \sum_{i=0}^{n} \binom{n}{i} x^{n-i} y^i$".

When $n = 1$, we have $(x+y)^1 = x + y = \binom{1}{0}x^1 + \binom{1}{1}y^1 = \sum_{i=0}^{1} \binom{1}{i} x^{1-i} y^i$. Therefore, $P(1)$ is true.

Assume that $P(k)$ is true for some $k \in \mathbb{N}$. That is, assume that $(x+y)^k = \sum_{i=0}^{k} \binom{k}{i} x^{k-i} y^i$.

We now show that $P(k+1)$ is true. Multiply each side of the equation by $(x+y)$ to get

$$\begin{aligned}
(x+y)^{k+1} &= \left(\sum_{i=0}^{k} \binom{k}{i} x^{k-i} y^i\right)(x+y) \\
&= x\left(\sum_{i=0}^{k} \binom{k}{i} x^{k-i} y^i\right) + y\left(\sum_{i=0}^{k} \binom{k}{i} x^{k-i} y^i\right) \\
&= \sum_{i=0}^{k} \binom{k}{i} x^{k+1-i} y^i + \sum_{i=0}^{k} \binom{k}{i} x^{k-i} y^{i+1} \\
&= \binom{k}{0} x^{k+1} + \binom{k}{k} y^{k+1} + \sum_{i=1}^{k} \binom{k}{i} x^{k+1-i} y^i + \sum_{i=0}^{k-1} \binom{k}{i} x^{k-i} y^{i+1} \\
&= \binom{k}{0} x^{k+1} + \binom{k}{k} y^{k+1} + \sum_{i=1}^{k} \binom{k}{i} x^{k+1-i} y^i + \sum_{i=1}^{k} \binom{k}{i-1} x^{k-i} y^{i+1} \\
&= \binom{k}{0} x^{k+1} + \binom{k}{k} y^{k+1} + \sum_{i=1}^{k} \left(\binom{k}{i} + \binom{k}{i-1}\right) x^{k+1-i} y^i \\
&= \binom{k+1}{0} x^{k+1} + \binom{k+1}{k+1} y^{k+1} + \sum_{i=1}^{k} \binom{k+1}{i} x^{k+1-i} y^i \\
&= \sum_{i=0}^{k+1} \binom{k+1}{i} x^{k+1-i} y^i.
\end{aligned}$$

Therefore, $P(k+1)$ is true.

By the principle of mathematical induction, $P(n)$ is true for all $n \in \mathbb{N}$. □

*Remark.* It is possible to generalize binomial coefficients and the binomial theorem to polynomials with more than two terms, such as $(x+y+z)^n$. These generalizations are called multinomial coefficients and the multinomial theorem, respectively. As an exercise, think about how to formulate these generalizations.

Immediately from the statement of the binomial theorem, we get a variant of the theorem as a corollary.

**Corollary 36.** *Let $x$ be a variable and let $n$ be a natural number. Then*

$$(x+1)^n = \sum_{i=0}^{n} \binom{n}{i} x^i.$$

*Proof.* Follows from the binomial theorem when $y = 1$. □

In the proof of the binomial theorem, we require a particular identity that tells us something about the value of a binomial coefficient in terms of smaller binomial coefficients. Using this identity, which was named after the French mathematician Blaise Pascal, we are able to define binomial coefficients recursively, which is a great help in computational applications.

**Theorem 37** (Pascal's identity)**.** *For all $1 \leq k \leq n$,*

$$\binom{n}{k} + \binom{n}{k-1} = \binom{n+1}{k}.$$

*Proof.* By the definition of the binomial coefficient, we have

$$\begin{aligned}\binom{n}{k} + \binom{n}{k-1} &= \frac{n!}{k!(n-k)!} + \frac{n!}{(k-1)!(n-k+1)!} \\ &= \frac{n!}{k!(n-k+1)!} \cdot ((n-k+1)+k) \\ &= \frac{(n+1)!}{k!(n+1-k)!} = \binom{n+1}{k}.\end{aligned}$$

□

We can reason about Pascal's identity in the following way. Suppose $S$ is a set containing $n+1$ elements, and denote one "special" element as $a$. Let $T$ be the subset of $S$ not containing $a$. There are $\binom{n+1}{k}$ size-$k$ subsets of $S$, and these subsets either (i) do not contain $a$, but only contain $k$ elements from $T$, or (ii) contain both $a$ and $k-1$ elements from $T$. In scenario (i), there are $\binom{n}{k}$ possible subsets of $T$, and in scenario (ii), there are $\binom{n}{k-1}$ possible subsets of $T$, so altogether we have $\binom{n}{k} + \binom{n}{k-1}$ size-$k$ subsets of $S$.

With Pascal's identity, we can draw a beautiful triangular arrangement of binomial coefficients where the $k$th term in row $n+1$, $\binom{n+1}{k}$, is derived from the sum of the two terms $\binom{n}{k-1}$ and $\binom{n}{k}$ written directly above it. (Blank or nonexistent entries are taken to be zero.) This arrangement is called **Pascal's triangle**, in spite of the fact that it had been studied centuries before by other mathematicians. The first few rows of Pascal's triangle are as follows:

$$\begin{array}{ccccccccccccc} &&&&&& \binom{0}{0} &&&&&& \\ &&&&& \binom{1}{0} && \binom{1}{1} &&&&& \\ &&&& \binom{2}{0} && \binom{2}{1} && \binom{2}{2} &&&& \\ &&& \binom{3}{0} && \binom{3}{1} && \binom{3}{2} && \binom{3}{3} &&& \\ && \binom{4}{0} && \binom{4}{1} && \binom{4}{2} && \binom{4}{3} && \binom{4}{4} && \\ & \binom{5}{0} && \binom{5}{1} && \binom{5}{2} && \binom{5}{3} && \binom{5}{4} && \binom{5}{5} & \\ \binom{6}{0} && \binom{6}{1} && \binom{6}{2} && \binom{6}{3} && \binom{6}{4} && \binom{6}{5} && \binom{6}{6} \end{array} = \begin{array}{ccccccccccccc} &&&&&& 1 &&&&&& \\ &&&&& 1 && 1 &&&&& \\ &&&& 1 && 2 && 1 &&&& \\ &&& 1 && 3 && 3 && 1 &&& \\ && 1 && 4 && 6 && 4 && 1 && \\ & 1 && 5 && 10 && 10 && 5 && 1 & \\ 1 && 6 && 15 && 20 && 15 && 6 && 1 \end{array}$$

Aside from looking nice, Pascal's triangle reveals many hidden intricacies in the structure of and relationships between binomial coefficients. We present a few of these interesting results here without proof.

**Proposition 38.** *The following identities hold:*

1. $\binom{r}{k} = \binom{r}{r-k}$ *for all* $0 \leq k \leq r$ *(row symmetry);*
2. $\sum_{k=0}^{r} \binom{r}{k} = 2^r$ *for all* $r \geq 0$ *(row sum);*
3. $\sum_{k=0}^{r} \binom{r}{k}^2 = \binom{2r}{r}$ *for all* $r \geq 0$ *(row sum of squares);*
4. $\sum_{k=0}^{r} \binom{k}{c} = \binom{r+1}{c+1}$ *for all* $r, c \geq 0$ *(column sum).*

All things considered, what do binomial coefficients and the binomial theorem have to do with computing? For one, just as we saw in our discussion on combinations with repetition, the binomial theorem can speed up certain calculations in computer algebra systems. Multiplication is an intensive operation for a computer, and many multiplications at once can slow down a computation considerably. With the binomial theorem, we can expand certain expressions or calculate certain terms within an expression much more efficiently.