Chapter 2

Finite Counting

We have already seen examples of finite counting in our use of the sum and product rules and inclusion-exclusion formulae. Indeed, if we have some sets, then the sum rule or inclusion-exclusion formulae tell us how many ways there are to choose one item from these sets (that is, a single item is taken from the union of all the sets, so only one item is chosen in total). The product rule tells us instead how many ways there are to choose one item from each set (so the number of items chosen in total is equal to the number of sets).

Example. One society has 12 members, and another has 23 members; they have no members in common. How many ways are there to choose from the memberships

(i) a *single* representative for the societies?
(ii) a representative for each society?

Solution. Let A be the set of members of the first society, and B the set of members of the second society. Then $A \cup B$ is the set of members of either society, so the answer to (i) is $|A \cup B| = 12 + 23 = 35$. Choosing a representative from each society is the same as choosing a pair $(a, b) \in A \times B$, as a is then the representative for the first society and b is then the representative for the second. So the answer to (ii) is $|A \times B| = |A||B| = 12 \times 23 = 276$.[1] $\square$

More complicated counting arguments can be formulated by combining several choices of the above types; we will see several examples in this chapter. In particular, we will describe how many possible ways there are to choose *several* elements from a single set. This depends on the rules for the choices; in particular:

- Are we allowed to choose the same element more than once? If so, the number of possibilities will be higher than if not.

- Does the order in which the elements are chosen matter? That is, if we first choose a and then b, is this different from choosing first b and then a? If we consider these to be different, then there will be more possibilities than if we consider them to be the same.

Several of our examples in this chapter will consider probabilities of events. These are naturally linked to counting results because, if a random experiment has a finite number of

[1] For (ii), it may help to think like this: there are 12 choices for the first representative, and *for each* of these choices there are then 23 choices for the second representative, so in total there are 12 lots of 23 choices.

possible outcomes, and every outcome is equally likely, then the probability of an event E is

$$\mathbb{P}(E) = \frac{\text{number of outcomes for which } E \text{ occurs}}{\text{total number of outcomes}}.$$

We say that a selection is made uniformly at random if every outcome is equally likely. For example, rolling a fair standard 6-sided die selects an element of the set $\{1, 2, 3, 4, 5, 6\}$ uniformly at random.[2] Likewise, in the example above, if we choose uniformly at random a single representative for the societies, then each person has probability $\frac{1}{35}$ of being chosen.

2.1 Ordered Choice

We first consider the situation when the order in which items are chosen *does* matter. For this the following notation is very useful.[3]

Definition. For any natural number n, we define n factorial, denoted $n!$, by $n! = n \times (n-1) \times (n-2) \times \cdots \times 3 \times 2 \times 1$. We also define $0! = 1$.

So $0! = 1$, $1! = 1$, $2! = 2$, $3! = 6$, $4! = 24$, $5! = 120$, $6! = 720$, and so on.

Theorem 2.1. *Let r and n be non-negative integers, and let S be a set of size n.*

(i) The number of ways to choose r elements of S, if repetition is allowed and the order of choices matters, is n^r.
(ii) The number of ways to choose r elements of S, if the order of choices matters but repetition is forbidden, is $\frac{n!}{(n-r)!}$ if $r \leq n$ and zero if $r > n$.

Proof. For (i), note that there are n choices for the first element of the sequence. For each of these there are then n choices for the second element of the sequence, then n for the third, and so forth until the rth choice (the final one). So there are n^r possibilities in total.

This argument implicitly uses the product rule when multiplying the number of choices as in the last example. A formal argument might run as follows: note that the chosen elements form an ordered r-tuple $(x_1, \ldots, x_r)$ of r elements of S, where x_1 is the first element chosen, x_2 is the second element chosen, and so forth. Each such r-tuple arises from precisely one possible way to choose r elements of S, if repetition is allowed and the order of choices matters. Since S^r is the set of all such r-tuples $(x_1, \ldots, x_r)$, we deduce by the product rule that the number of possible ways to make the choices is $|S^r| = |S|^r = n^r$.

For (ii), observe that there are n choices for the first element of the sequence (any element of S), and for each of these there are $n-1$ choices for the second element (any element of S except that chosen for the first element). There are then $n-2$ choices for the third element of the sequence (any element of S except the two already chosen), $n-3$ choices for the fourth element, and so forth until the rth and final element, for which there

[2] All random selections considered in this course will be made uniformly at random; in future courses you will see examples of non-uniform random selections.

[3] Be careful not to mix up the mathematical and punctuational uses of the '!' symbol! For example, if you solve a problem and write 'the answer is 10!', do you mean 10 or 10 factorial?

are $n-(r-1)$ choices (any element of S except the $r-1$ previously chosen). So overall, provided $r \leq n$, the number of sequences is

$$n \cdot (n-1) \cdot (n-2) \cdot \ldots \cdot (n-r+1) = \frac{n!}{(n-r)!}.$$

Finally, if $r > n$ then, by the pigeonhole principle, however we make r choices from the n elements of S, there must be some element which is chosen at least twice. So there are no ways to choose r elements of S without repetition.

Again, the product rule is implicit in this argument. We could make this explicit by noting that each way to choose r elements corresponds to precisely one r-tuple

$$(y_1, \ldots, y_r) \in \{1, \ldots, n\} \times \{1, \ldots, n-1\} \times \{1, \ldots, n-2\} \times \cdots \times \{1, \ldots, n-r+1\}.$$

Indeed, we fix some order of the elements of S, and then the r-tuple $(y_1, \ldots y_r)$ corresponds to choosing the y_1th element of S, then the y_2th element of the $n-1$ remaining elements, then the y_3th element of the $n-2$ remaining elements, and so forth. □

Example. A bag contains 5 balls, labelled $1, 2, 3, 4$ and 5. I draw out two balls in turn from the bag.[4] What is the probability that the number on the second ball drawn is precisely one greater than the number on the first ball drawn, if:

(i) I replace the first ball before drawing the second?
(ii) I do not replace the first ball before the second is drawn?

Solution. Let x be the number on the first ball drawn and y be the number on the second ball drawn, so each outcome of drawing the balls is a pair (x, y). There are four outcomes of the selection which give the event described, namely $(1, 2), (2, 3), (3, 4)$ and $(4, 5)$, so it remains to calculate the total number of possible outcomes.

For (i) we are making two choices from the set $\{1, 2, 3, 4, 5\}$ with repetition allowed (because we replaced the first ball so it could be drawn again), and where order matters (because it matters which ball is drawn first). So by Theorem 2.1(i) there are $5^2 = 25$ possible outcomes, and so the probability is $4/25$.

For (ii) we are again making two choices from the set $\{1, 2, 3, 4, 5\}$ where order matters, but now repetition is forbidden (because we cannot draw the first ball again once it has been drawn). So by Theorem 2.1(ii) there are $5!/(5-2)! = 5 \times 4 = 20$ possible outcomes, and so the probability is $4/20 = 1/5$.[5] □

Definition. A <u>permutation</u> of a set S is an ordering of its elements.

That is, the permutations of a set of size r are the ordered r-tuples whose co-ordinates are precisely the elements of the set. For example, the permutations of $\{1, 2, 3\}$ are

$$(1, 2, 3), (1, 3, 2), (2, 1, 3), (2, 3, 1), (3, 1, 2), \text{ and } (3, 2, 1).$$

So there are $6 = 3!$ permutations of a set of three elements.

[4] Many questions of this type involve drawing balls from bags, or rolling dice, or dealing cards from a deck. Unless stated otherwise you should assume that each selection is uniformly random, that is, that each ball is equally likely to be drawn, each card is equally likely to be dealt, each face of the die is equally likely to come up, etc.

[5] In both parts of this exercise it is probably easier to argue by counting choices than to appeal to Theorem 2.1. However, in more complicated examples, where we are making a large number of choices as part of a larger argument, using Theorem 2.1 often simplifies the argument.

Corollary 2.2. *A set S of n elements has $n!$ permutations.*

Proof. The permutations of S are exactly the ordered sequences of n elements of S in which no element is repeated, and the number of these is $\frac{n!}{(n-n)!} = n!$ by Theorem 2.1(ii) applied with $r = n$. $\square$

Example. There are $n!$ ways for n people to line up in a queue, since each order is a permutation of the set of people.

We conclude this section with two more examples of ordered choice.

Example. How many anagrams are there of the word MATHS? In how many of these is 'T' immediately before 'H'?

Solution. Each anagram of MATHS is a permutation of the set $\{\mathrm{M}, \mathrm{A}, \mathrm{T}, \mathrm{H}, \mathrm{S}\}$, so by Corollary 2.2 there are $5! = 120$ anagrams. For the second part, note that the anagrams of MATHS in which 'T' is immediately before 'H' are precisely the permutations of the set $\{\mathrm{M}, \mathrm{A}, \mathrm{TH}, \mathrm{S}\}$ (i.e. we treat 'TH' as a single letter), and Corollary 2.2 tells us there are $4! = 24$ of these. $\square$

Example. I roll four standard dice. What is the probability that the numbers obtained are consecutive?

Solution. If we imagine rolling the dice in turn, then we are making four choices from $\{1, 2, 3, 4, 5, 6\}$ in which order matters and where repetition is allowed. So by Theorem 2.1(i) there are $6^4 = 1296$ possible outcomes. The outcomes for which the numbers obtained are consecutive are the permutations of $\{1, 2, 3, 4\}$, $\{2, 3, 4, 5\}$ and $\{3, 4, 5, 6\}$. Each of these sets has $4! = 24$ permutations by Corollary 2.2, so in total there are 72 such outcomes. So the probability is $72/1296 = 1/18$. $\square$

2.2 Unordered Choice

The common property of the examples of the previous section was that we cared about the order in which we made choices. This will often not be the case: for example, in the national lottery we do not care about the order in which the balls are drawn, only which balls are drawn. This is the distinguishing property of *unordered choice*, for which the next definition is crucial.[6]

Definition. Let n and r be non-negative integers with $0 \leq r \leq n$. Then the binomial coefficient of n and r, also called n choose r, and denoted $\binom{n}{r}$ is

$$\binom{n}{r} = \frac{n!}{r!(n-r)!} = \frac{n(n-1)(n-2)\ldots(n-r+1)}{r!}.$$

Note that

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

[6]Please use $\binom{n}{r}$ instead of ${}^nC_r$ or $C(n, r)$; the latter forms are not universally recognised.

For example,

$$\binom{10}{2} = \frac{10 \times 9}{2 \times 1} = 45.$$

$$\left(\text{Note this is much simpler than } \binom{10}{2} = \frac{10!}{2!\,8!} = \frac{3628800}{2 \cdot 40320} = \frac{3628800}{80640} = 45.\right)$$

$$\binom{14}{7} = \frac{14 \times 13 \times 12 \times 11 \times 10 \times 9 \times 8}{7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1} = \frac{14}{7 \times 2} \times \frac{10}{5} \times \frac{12}{6} \times \frac{8}{4} \times \frac{9}{3} \times 13 \times 11$$
$$= 1 \times 2 \times 2 \times 2 \times 3 \times 13 \times 11 = 3432.$$

(Note the rearrangement to make the calculation simpler, especially if you are not using a calculator.)

$$\binom{n}{n} = \binom{n}{0} = \frac{n!}{n!\,1!} = 1.$$

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

$$\binom{n}{n-2} = \binom{n}{2} = \frac{n \times (n-1)}{2 \times 1} = \frac{n(n-1)}{2}.$$

Theorem 2.3. *Let S be a set of size n, and let r be an integer with $0 \le r \le n$. Then there are $\binom{n}{r}$ ways to choose r members of S if the order of choice is irrelevant and repetition is forbidden. Equivalently, the number of subsets $R \subseteq S$ of size r is $\binom{n}{r}$.*

To see that these statements are equivalent, recall that a set may not have repeated elements, and the order of the elements is immaterial.

Proof. Consider a subset $R \subseteq S$ of size r . Then each permutation of R is an ordered sequence of r elements of S without repetition; furthermore permutations of different subsets R give different sequences. So

$$(\text{number of subsets } R \subseteq S \text{ of size } r) \times (\text{number of permutations of such an } R)$$
$$= (\text{number of ordered sequences of } r \text{ elements of } S\text{, without repetition}).$$

So, using Theorem 2.1 and Corollary 2.2, the number of subsets $R \subseteq S$ of size r is

$$\frac{(\text{\# of ordered sequences of } r \text{ elements of } S\text{, without repetition})}{(\text{\# of permutations of a set of size } r)} = \frac{n!/(n-r)!}{r!}$$
$$= \frac{n!}{r!(n-r)!} = \binom{n}{r}.$$

$\square$

Example. How many 5-card poker hands are there? How many contain two aces, a king, and two other cards (i.e. not an ace or king)? What is the probability that a random 5-card hand has this form?

Solution. The order of cards is irrelevant, and repetition is impossible (you can't have more than one of the same card), so the answer to the first part is the number of ways to choose 5 cards out of 52 without repetition and ignoring order, that is,

$$\binom{52}{5} = \frac{52 \times 51 \times 50 \times 49 \times 48}{5 \times 4 \times 3 \times 2 \times 1} = 2\,598\,960.$$

For the second part, there are $\binom{4}{2}$ ways to choose two of the four aces in the deck, and for each of these there are $\binom{4}{1}$ ways to choose one of the four kings in the deck. Then, whatever we have chosen so far, there are $\binom{44}{2}$ ways to choose two cards from the 44 cards which are not an ace or a king. So (by the product rule) in total the number of such hands is

$$\binom{4}{2}\binom{4}{1}\binom{44}{2} = 6 \times 4 \times 946 = 22\,704.$$

In particular, the probability that a random 5-card hand contains 2 aces, a king, and two other cards, is

$$\frac{22\,704}{2\,598\,960} \approx 0.008735.$$

□

Example. In the National Lottery we choose 6 numbers from the set $S = \{1, \ldots, 49\}$. The lottery draw then selects 6 numbers (i.e. a subset of S of size 6) uniformly at random. What is the probability of winning the jackpot by matching all six numbers? What is the probability of winning £25 by matching three numbers?[7]

Solution. The order in which the balls are selected is irrelevant, and a ball cannot be selected more than once. So the number of possibilities for the six numbers drawn is $\binom{49}{6}$. The probability that the outcome is the six numbers we chose is therefore

$$\frac{1}{\binom{49}{6}} = \frac{1}{\left(\frac{49\times48\times47\times46\times45\times44}{6\times5\times4\times3\times2\times1}\right)} = \frac{1}{13\,983\,816}.$$

For the second part, we count how many possible choices of six balls there are which match precisely three of the six balls we selected. For this, note that there are $\binom{6}{3}$ ways to choose three of our selected six balls, and for each of these choices there are $\binom{43}{3}$ ways to choose three of the 43 balls we did not select. So there are $\binom{6}{3}\binom{43}{3}$ possible outcomes of the draw which match exactly three of our selected balls, so the probability that this event occurs is

$$\frac{\binom{6}{3}\binom{43}{3}}{\binom{49}{6}} = \frac{20 \times 12\,341}{13\,983\,816} = \frac{246\,820}{13\,983\,816} \approx 0.01765 \qquad \text{(around 1 in 56.656).}$$

As an exercise, try calculating the probability of matching 5 numbers and the bonus ball.

□

Finally, we consider the case where repetition is allowed but we don't care about the order of choices.[8]

[7] In October 2015 the rules of the National Lottery were changed; the six numbers are now selected from the set $\{1, \ldots, 59\}$. As an exercise, calculate the effect this has on the winning probabilities in this example.

[8] Be very careful when using Theorem 2.4 to calculate probabilities, as these will often not be uniform. For example, if I roll two identical dice and ignore the order in which they are rolled, by Theorem 2.4 there are $\binom{6+2-1}{2} = 21$ distinct possible outcomes (exercise: list them), but these are not equally likely. For instance, I am twice as likely to get a three and a four than I am to get two sixes. For this reason, when we repeat random experiments with repetition allowed, such as rolling dice or drawing balls from a bag with replacement, we almost always count ordered outcomes and use Theorem 2.1. On the other hand, as we have already seen, when we repeat random experiments with repetition forbidden, such as dealing cards from a deck or drawing balls without replacement, we can either count outcomes with order using Theorem 2.1 or count outcomes without order using Theorem 2.3; the latter is usually simpler.

Theorem 2.4. *Let S be a set of size n. Then there are $\binom{n+r-1}{r}$ ways to choose r elements of S if repetition is allowed but the order of choosing is irrelevant.*

Proof. Suppose that $S = \{s_1, s_2, \dots, s_n\}$. Imagine $n + r - 1$ positions, represented by crosses. For example, for $n = 6$, $r = 4$ we have

$$\times \ \times \ \times \ \times \ \times \ \times \ \times \ \times \ \times$$

Choose $n - 1$ of these and circle them, for example.

$$\times \ \otimes \ \otimes \ \times \ \otimes \ \otimes \ \times \ \times \ \otimes$$

There are $\binom{n+r-1}{n-1} = \binom{n+r-1}{r}$ possibilities for which $n - 1$ crosses are circled. Crucially, each choice of $n - 1$ crosses corresponds to a unique way to choose r elements of S allowing repetition but ignoring order. Indeed, the number of $\times$ before the first $\otimes$ tells you how many times to choose s_1, the number of $\times$ between the first and second $\otimes$ tells you how many times to choose s_2, the number of $\times$ between the second and third $\otimes$ tells you how many times to choose s_3, and so forth, until the number of $\times$ after the final $\otimes$ tells you how many times to choose s_n. For instance, in the example above we would choose s_1 once, s_3 once and s_5 twice. Similarly, if we had circled

$$\otimes \ \otimes \ \otimes \ \times \ \times \ \times \ \times \ \otimes \ \otimes$$

then we would choose s_4 four times and not choose any other member of S at all, and if we had circled

$$\times \ \times \ \otimes \ \otimes \ \times \ \otimes \ \times \ \otimes \ \otimes$$

then we would choose s_1 twice, s_3 once and s_4 once. On the other hand, choosing s_1 once, s_3 once and s_6 twice would be obtained by circling[9]

$$\times \ \otimes \ \otimes \ \times \ \otimes \ \otimes \ \otimes \ \times \ \times.$$

□

Corollary 2.5. *For natural numbers r and n, the number of nonnegative integer solutions of $x_1 + x_2 + \cdots + x_n = r$ is $\binom{n+r-1}{r}$.*

The importance of this corollary is that it is the number of ways of distributing r identical objects among n people. So, for example, if I want to share 7 pound coins among 5 people then there are $\binom{5+7-1}{7} = \binom{11}{7} = 330$ possible ways to do this.

Proof. Each non-negative integer solution of the equation corresponds uniquely to a choice of r elements of $\{1, 2, \dots, n\}$, allowing repetition but ignoring order. Indeed, for each $1 \leq i \leq n$, the integer x_i tells you how many times to choose the element i. So by Theorem 2.4 the number of solutions is $\binom{n+r-1}{r}$.

Alternatively, you can argue directly in a similar way to the proof of Theorem 2.4: consider $n + r - 1$ positions, filled with the symbol $\times$. Then a choice of any $n - 1$ of the $\times$ to circle corresponds uniquely to a solution of the equation: the value of x_1 is the number of $\times$ before the first $\otimes$, the value of x_2 is the number of $\times$ between the first and second $\otimes$, and so forth.[10] □

[9]Take a moment to convince yourself that this is a genuine bijection. That is, for different choices of which crosses to circle we obtain different choices of members of S, and every possible way to choose r elements of s with repetition but ignoring order will be formed in this way.

[10]Exercise: fill in the details in this argument.

We can solve similar problems with different restrictions on $x_1, x_2, \dots, x_n$ similarly, as in the following example[11].

Example. How many solutions are there in *positive* integers to $x_1 + x_2 + x_3 = 101$?

Solution. Let $y_1 = x_1 - 1$, $y_2 = x_2 - 1$ and $y_3 = x_3 - 1$. Then $x_1 + x_2 + x_3 = 101$ is equivalent to $y_1 + y_2 + y_3 = 98$, and y_i is a non-negative integer if and only if x_i is a positive integer. So the number of positive integer solutions to $x_1 + x_2 + x_3 = 101$ is the number of non-negative integer solutions to $y_1 + y_2 + y_3 = 98$, which is $\binom{100}{98} = \binom{100}{2} = 4950$ by Corollary 2.5. $\square$

To summarise the results of Sections 2.1 and 2.2, the number of ways of choosing r elements from a set of size n is

	with order	ignoring order
repetition allowed	n^r	$\binom{n+r-1}{r}$
repetition forbidden	$\frac{n!}{(n-r)!}$	$\binom{n}{r}$

We conclude this section with two more advanced examples of how these results may be combined.

Example. I deal a hand of six cards from a standard 52-card deck. What is the probability that I get two cards of one suit and four cards of another suit? What is the probability that I get three cards of one suit and three cards of another suit?

Solution. First, note that there are $\binom{52}{6}$ possible 6-card hands. We first count the number of hands consisting of two cards of one suit and four of another. There are 4 possibilities for the suit from which we have four cards, and then $\binom{13}{4}$ possibilities for what these four cards are. Having chosen these there are then 3 possibilities for the suit of the remaining two cards, and then $\binom{13}{2}$ possibilities for what these two cards are. So there are $4 \cdot \binom{13}{4} \cdot 3 \cdot \binom{13}{2}$ such hands, and we conclude that the probability is

$$\frac{4 \cdot \binom{13}{4} \cdot 3 \cdot \binom{13}{2}}{\binom{52}{6}} \approx 0.0328727.$$

We next count the number of hands consisting of three cards of one suit and three of another. There are $\binom{4}{2}$ ways to choose which two suits the cards come from.[12] Having chosen these, there are $\binom{13}{3}$ ways to choose three cards from the first suit, and then $\binom{13}{3}$ ways to choose three cards from the second suit. So there are $\binom{4}{2} \cdot \binom{13}{3} \cdot \binom{13}{3}$ such hands, and we conclude that the probability is

$$\frac{\binom{4}{2} \cdot \binom{13}{3} \cdot \binom{13}{3}}{\binom{52}{6}} \approx 0.0241067.$$

$\square$

[11] See the problem sheets for further examples.

[12] Note carefully the difference between the solutions of the different parts of this example. The fundamental reason for this is that, for example, having two spades and four hearts is different from having two hearts and four spades, whilst having three hearts and three spades is the same as having three spades and three hearts.

Example. I roll five standard dice. What is the probability that I get at least three sixes?

Solution. If we imagine rolling the dice in order, there are 6^5 possible outcomes. In exactly one of these outcomes we get five sixes (the outcome where every die shows six). We next count the number of outcomes in which we get precisely four sixes. Note for this that there are five possibilities for which die does not show six, and there are five possibilities (1, 2, 3, 4 or 5) for the result of this die. So there are 25 outcomes in which we get precisely four sixes.[13] We can similarly count the number of outcomes in which we get precisely three sixes: there are $\binom{5}{2}$ possibilities for which two dice do not show six, and there are 5^2 possibilities for the values of these two dice, giving $\binom{5}{2}5^2 = 250$ outcomes in which we get exactly three sixes. So in total there are $1 + 25 + 250 = 276$ outcomes in which we get at least three sixes, so the probability of this event is $\frac{276}{6^5} = 0.03549.$ $\square$

Note in this example that we partitioned the outcomes satisfying an inequality (i.e. there are at least three sixes) into disjoint sets of outcomes satisfying equalities (i.e. the outcomes with exactly three sixes, those with exactly four sizes and those with exactly five sixes). We then summed the sizes of these sets (implicitly using the sum rule) to count those satisfying an inequality. In general this is the correct way to approach questions involving inequalities[14].

2.3 Mixed Choice

‘Mixed choice’ describes the situation when order ‘partly matters’. For example, in the following example the order of different letters matters (‘MISSISSIPPI’ is a different anagram from ‘IMSSISSIPPI’). However, swapping the positions of two of the ‘S’s, say, yields the same anagram.

Example. How many anagrams are there of ‘MISSISSIPPI’?

Solution. First, by Corollary 2.2 there are $11!$ possible orders for the elements of the set $\{\mathrm{M}, \mathrm{I}_1, \mathrm{S}_1, \mathrm{S}_2, \mathrm{I}_2, \mathrm{S}_3, \mathrm{S}_4, \mathrm{I}_3, \mathrm{P}_1, \mathrm{P}_2, \mathrm{I}_4\}$. Ignoring the subscripts, these orders give all the anagrams of ‘MISSISSIPPI’. However, each anagram of ‘MISSISSIPPI’ is formed $4!4!2!$ times in this manner: there are $4!$ possible arrangements of $\mathrm{I}_1, \mathrm{I}_2, \mathrm{I}_3$ and I_4, for each of these there are $4!$ possible arrangements of $\mathrm{S}_1, \mathrm{S}_2, \mathrm{S}_3$ and S_4, and for each of these there are $2! = 2$ possible arrangements of P_1 and P_2. So in total the number of anagrams of ‘MISSISSIPPI’ is $\frac{11!}{4!4!2!} = 34\,650$.

An alternative approach is to consider a sequence of 11 blank spaces, into which we will place the letters of MISSISSIPPI. First, we decide where we are going to put the four ‘S’s. There are 11 spaces to choose from, so there are $\binom{11}{4}$ possibilities to choose from (since we are choosing four of the eleven empty spaces). For any of these choices, there

[13] Note the difference here compared to the last example. This situation is ordered choice, so (for instance) there are five different possible ways to get four sixes and one five, as there are five choices for which die is the five. However, if we deal five cards from a deck and count unordered choices, then there is (for instance) only one way to get four kings and the ace of spades, since we are ignoring the order the cards come in. The essential thing is to be consistent throughout your solution in whether you are using ordered choice or unordered choice.

[14] Also keep in mind it may be easier to take the complement, that is, to count the outcomes that don’t satisfy the condition and subtract from the total. For example, if we wanted the probability of getting at least one six, it is easier to count the number of outcomes in which we get no sixes; subtracting this from the total number of possible outcomes gives the number of outcomes in which we get at least one six.

there are then seven empty spaces remaining, so there are $\binom{7}{4}$ possibilities for how to place the four 'I's among these. There are then three empty spaces, so $\binom{3}{2}$ choices for how to place the two 'P's. Finally, there is now only one empty space, so the remaining letter (the 'M') must be placed here - there is no choice. We conclude that the number of anagrams of 'MISSISSIPPI' is

$$\binom{11}{4}\binom{7}{4}\binom{3}{2} = \frac{11 \times 10 \times 9 \times 8}{4!} \cdot \frac{7 \times 6 \times 5 \times 4}{4!} \cdot \frac{3 \times 2}{2!} = \frac{11!}{4!4!2!} = 34\,650,$$

as before. □

Note that in the first method we considered the problem with order (i.e. treating all characters as different), then 'divided out the overcounting'; that is, we divided to reflect the fact that we have counted order where we didn't want to (e.g. the order of the four 'M's). On the other hand, in the second method we viewed the mixed choice as a series of separate unordered choices. In general both of these methods will work for mixed choice problems, but one may be significantly simpler than the other, depending on the problem.

There are many similar problems in which order 'partly matters' which can be tackled by similar arguments to those above. As an exercise, adapt the arguments given above to the following two problems.[15]

(i) In how many ways can 11 footballers be arranged in a 4-4-2 formation? That is, you need to choose 4 defenders, 4 midfielders and 2 forwards from the 11 players; the remaining player is the goalkeeper.
(ii) I have 11 distinct works of art, and I must give four to Alice, four to Bob, two to Charlotte and the remaining one to Dave. How many possible ways are there to achieve this distribution?

2.4 The Binomial Theorem and Consequences

In this section we will prove the Binomial Theorem and explore some of its corollaries. We begin with the following lemma.

Lemma 2.6. *For any integers r and n with $0 \leq r < n$ we have*

$$\binom{n}{r} + \binom{n}{r+1} = \binom{n+1}{r+1}.$$

[15] Also see the problem of counting paths and cycles in the Graph Theory chapter.

Proof. One approach is to write $\binom{n}{r} = \frac{n!}{r!\,(n-r)!}$ and similarly for the other terms, and argue as follows.

$$\begin{aligned}\binom{n}{r} + \binom{n}{r+1} &= \frac{n!}{r!\,(n-r)!} + \frac{n!}{(r+1)!\,(n-(r+1))!}\\ &= \frac{n!}{(r+1)!}\left(\frac{(r+1)}{(n-r)!} + \frac{1}{(n-r-1)!}\right)\\ &= \frac{n!}{(r+1)!\,(n-r)!}\big((r+1)+(n-r)\big)\\ &= \frac{n!(n+1)}{(r+1)!\,((n+1)-(r+1))!} = \binom{n+1}{r+1}.\end{aligned}$$

□

Proof. A more elegant argument uses the fact that $\binom{n}{r}$ is the number of ways to choose r elements from a set of size n. So

$$\begin{aligned}\binom{n+1}{r+1} &= \text{number of ways of choosing } (r+1) \text{ elements from } \{0,1,\ldots,n\}\\ &= \text{number of ways of choosing } (r+1) \text{ elements from } \{0,1,\ldots,n\} \text{ including } 0\\ &\quad + \text{number of ways of choosing } (r+1) \text{ elements from } \{0,1,\ldots,n\} \text{ excluding } 0\\ &= \text{number of ways of choosing } r \text{ elements from } \{1,\ldots,n\}\\ &\quad + \text{number of ways of choosing } (r+1) \text{ elements from } \{1,\ldots,n\}\\ &= \binom{n}{r} + \binom{n}{r+1}.\end{aligned}$$

In the above argument, choices are always unordered and without repetition. □

One consequence of Lemma 2.6 is that the binomial coefficients $\binom{n}{0}, \binom{n}{1}, \ldots, \binom{n}{n}$ are the numbers of the nth row of Pascal's triangle.

Theorem 2.7 (Binomial theorem)**.** *For any integer $n \geq 0$ and any $a, b \in \mathbb{R}$, we have*

$$(a+b)^n = \sum_{i=0}^{n} \binom{n}{i} a^i b^{n-i}.$$

This is why the numbers $\binom{n}{r}$ are called binomial coefficients: they are the coefficients in the binomial expansion $(a+b)^n$. As with the previous lemma, there are two principle ways to prove this theorem; the first is a nice 'combinatorial' argument using the fact that $\binom{n}{r}$ is the number of ways to choose r elements from a set of size n, ignoring order and without repetition, whilst the second uses the definition of $\binom{n}{r}$ as $\frac{n!}{r!\,(n-r)!}$ and a somewhat awkward induction argument.

Proof. (Combinatorial argument.) If we consider multiplying out $(a+b)^n$, it is easy to see that we must have

$$(a+b)^n = \sum_{i=0}^{n} t_i a^i b^{n-i}$$

for some $t_0, t_1, \dots, t_n \in \mathbb{N}$. (Indeed, when multiplying out, from each of the n brackets we take either the 'a' or the 'b' term, so their exponents always add up to n). There are n brackets we multiply out and the term $a^i b^{n-i}$ arises precisely when we choose the 'a' term from exactly i of the brackets (and thus the 'b' term from all the other $n-i$ brackets). There are $\binom{n}{i}$ ways of doing this. So we must have $t_i = \binom{n}{i}$. □

Proof. (By induction.) For each integer $n \geq 0$, let $P(n)$ be the statement that for any $a, b \in \mathbb{R}$, we have

$$(a+b)^n = \sum_{i=0}^{n} \binom{n}{i} a^i b^{n-i}.$$

Then $P(0)$ holds since for any $a, b \in \mathbb{R}$ we have $(a+b)^0 = 1 = a^0 = b^0 = \binom{n}{0}$. Furthermore,

$$\sum_{i=0}^{1} \binom{1}{i} a^i b^{1-i} = \binom{1}{0} b + \binom{1}{1} a = b + a = (a+b)^1.$$

So $P(1)$ holds.

Now suppose that $P(k)$ holds . Then

$$\begin{aligned}
(a+b)^{k+1} &= (a+b)(a+b)^k = (a+b)\left(\sum_{i=0}^{k} \binom{k}{i} a^i b^{k-i}\right) \\
&= \sum_{i=0}^{k} \binom{k}{i} a^{i+1} b^{k-i} + \sum_{i=0}^{k} \binom{k}{i} a^i b^{k+1-i} \\
&\overset{(\text{using } j=i+1)}{=} \sum_{j=1}^{k+1} \binom{k}{j-1} a^j b^{k+1-j} + \sum_{i=0}^{k} \binom{k}{i} a^i b^{k+1-i} \\
&= a^{k+1} + b^{k+1} + \sum_{i=1}^{k} \left(\binom{k}{i-1} + \binom{k}{i}\right) a^i b^{k+1-i} \\
&\overset{(\text{Lemma 2.6})}{=} \binom{k+1}{k+1} a^{k+1} + \binom{k+1}{0} b^{k+1} + \sum_{i=1}^{k} \binom{k+1}{i} a^i b^{k+1-i} \\
&= \sum_{i=0}^{k+1} \binom{k+1}{i} a^i b^{k+1-i}.
\end{aligned}$$

So $P(k+1)$ holds. So by the principle of mathematical induction, $P(n)$ holds for all n. □

Corollary 2.8. *For any integer $n \geq 0$ we have*

$$\sum_{i=0}^{n} \binom{n}{i} = 2^n.$$

Proof.

$$2^n = (1+1)^n = \sum_{i=0}^{n} \binom{n}{i} 1^i 1^{n-i} = \sum_{i=0}^{n} \binom{n}{i},$$

where the middle equality uses the Binomial Theorem with $a = b = 1$. □

Note that Corollary 2.8 gives another proof of Theorem 1.3, which stated that for any set X we have $|\mathscr{P}(X)| = 2^{|X|}$. That is, a set S of size n has 2^n subsets. Indeed,

$$\text{Total number of subsets of } S = \sum_{i=0}^{n} (\text{number of subsets of } S \text{ of size } i) = \sum_{i=0}^{n} \binom{n}{i} = 2^n,$$

where the final equality holds by Corollary 2.8.

Corollary 2.9. *For any integer $n \geq 1$ we have*

$$\binom{n}{0} - \binom{n}{1} + \binom{n}{2} - \binom{n}{3} + \cdots \pm \binom{n}{n} = 0,$$

where the $\pm$ at the end is $+$ if n is even and $-$ if n is odd.

Proof.

$$0 = (-1+1)^n = \sum_{i=0}^{n} \binom{n}{i} (-1)^i 1^{n-i} = \binom{n}{0} - \binom{n}{1} + \binom{n}{2} - \binom{n}{3} + \cdots \pm \binom{n}{n}.$$

□

Note that Corollary 2.9 does *not* hold for $n = 0$, as in this case we just have the term $\binom{n}{0} = (-1+1)^0 = 0^0 = 1$.

Corollary 2.10. *Let S be a set of size $n \geq 1$. Then S has 2^{n-1} subsets of even size and 2^{n-1} subsets of odd size.*

Proof. Corollary 2.9 says that any non-empty set S has the same number of subsets of even size as it does of odd size. Indeed if S has size n then the number of subsets of odd size is $\sum \binom{n}{r}$ where the sum ranges over all *odd* integers r between 0 and n, and the number of subsets of even size is $\sum \binom{n}{r}$ where the sum ranges over all *even* integers r between 0 and n; Corollary 2.9 says that these two sums are equal. Combining this with Theorem 1.3 gives the result. □

Again, this result does *not* hold for $n = 0$, since the empty set has one subset of even size (itself) and no subset of odd size.