The Binomial Theorem and Combinatorial Proofs

Shagnik Das

Introduction

As we live our lives, we are faced with innumerable decisions on a daily basis[1]. Can I get away with wearing these clothes again before I have to wash them? Which frozen pizza should I have for dinner tonight? What excuse should I use for skipping the gym today? While it may at times seem tiring to be faced with all these choices on a regular basis, it is this exercise of free will that separates us from the machines we live with.[2]

This multitude of choices, then, provides some measure of our own vitality — the more options we have, the more alive we truly are. What, then, could be more important than being able to count how many options one actually has?[4] As we have already seen during our first week of lectures, one of the main protagonists in these counting problems is the binomial coefficient, whose definition is given below.

Definition 1. *Given non-negative integers n and k, the binomial coefficient $\binom{n}{k}$ denotes the number of subsets of size k of a set of n elements. Equivalently, it is the number of unordered choices of k distinct elements from a set of n elements.*

However, the binomial coefficient leads a double life. Not only does it have the above definition, but also the formula below, which we proved in lecture.

Proposition 2. *For non-negative integers n and k,*

$$\binom{n}{k} = \frac{n^{\underline{k}}}{k!} = \frac{n(n-1)(n-2)\dots(n-k+1)}{k!} = \frac{n!}{k!(n-k)!}.$$

In this note, we will prove several more facts about this most fascinating of creatures.

The Binomial Theorem

The first of these facts explains the name given to these symbols. They are called the binomial coefficients because they appear naturally as coefficients in a sequence of very important polynomials.

Theorem 3 (The Binomial Theorem)**.** *Given real numbers[5] $x, y \in \mathbb{R}$ and a non-negative integer n,*

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}.$$

[1]Or so we would like to believe; the thought that we are merely going through the motions of some deterministic plot, with no real influence on the situations we find ourselves in, is rather depressing.

[2]Those that dully plod through their preprogrammed routines, mindlessly fulfilling the tasks they were designed to accomplish, with no sense of happiness[3]or accomplishment.

[3]That being said, our machines are also devoid of sadness. Whether it is then better to be a person or a machine is beyond the scope of this text.

[4]Many things, of course, but for the purposes of this narrative I shall pretend they do not exist.

[5]It is of no importance that our variables x and y be real here. The same theorem holds for two unknowns in any commutative ring.

In your homework you are asked to prove this theorem by induction. Here we will give a different proof.

Proof. The left-hand size represents the product of n copies of $x+y$:

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

We will now expand this expression, multiplying all the terms out together. Since there are n linear factors, each monomial will collect one variable from each factor, and hence will be of degree n. More specifically, each monomial will be of the form $x^k y^{n-k}$ for some $k \in \{0, 1, 2, \dots, n\}$ representing how many times x was chosen.

The coefficient of $x^k y^{n-k}$ is the number of ways of choosing x exactly k times. This is equivalent to choosing a subset of k of the n factors from which to choose x (with y being chosen from the rest). By definition, there are $\binom{n}{k}$ such subsets, and hence the coefficient of $x^k y^{n-k}$ is $\binom{n}{k}$.

Summing up over the different monomials, we find

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}. \qquad \square$$

Binomial Identities

While the Binomial Theorem is an algebraic statement, by substituting appropriate values for x and y, we obtain relations involving the binomial coefficients. Such relations are examples of *binomial identities*, and can often be used to simplify expressions involving several binomial coefficients. We provide some examples below.

Corollary 4. *The following relations all hold.*

(i) For all $n \geq 0$, $\sum_{k=0}^{n} \binom{n}{k} = 2^n$.

(ii) For all $n \geq 1$, $\sum_{k \text{ odd}} \binom{n}{k} = \sum_{k \text{ even}} \binom{n}{k}$.

(iii) For all $n \geq 0$, $\sum_{k=0}^{n} \binom{n}{k} 2^k = 3^n$.

Proof. These results follow almost directly from Theorem 3.

For (i), substitute $x = y = 1$. We then have

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

as desired.

For (ii), substitute $x = -1$ and $y = 1$. The left-hand side is $(-1+1)^n = 0^n = 0$. On the other hand, the monomial $x^k y^{n-k}$ evaluates to $(-1)^k$, which is -1 if k is odd and 1 if k is even. Thus

$$0 = \sum_{k \text{ even}} \binom{n}{k} - \sum_{k \text{ odd}} \binom{n}{k},$$

which rearranges to give the claimed relation.

Finally, for (iii), we instead substitute $x = 2$ and $y = 1$. $\square$

Combinatorial Proofs

The Binomial Theorem thus provides some very quick proofs of several binomial identities. However, it is far from the only way of proving such statements. A *combinatorial proof* of an identity is a proof obtained by interpreting the each side of the inequality as a way of enumerating some set. If they are enumerations of the same set, then by the principle of double-counting it follows that they must be equal. If they are different sets, but you can build a bijection between the two, then the bijection rule shows they must be equal.

Since the binomial coefficients are defined in terms of counting, identities involving these coefficients often lend themselves to combinatorial proofs. These proofs are usually preferable to analytic or algebraic approaches, because instead of just verifying that some equality is true, they provide some insight into *why* it is true. Moreover, once one has identified a bijection between two sets, restricting the bijection to certain subsets can often lead to several other identities.[6]

To illustrate the concept, we provide combinatorial proofs of the identities in Corollary 4.

Proof of Corollary 4(i). For (i), let S be a set of n elements, and count 2^S, which is the collection of all subsets of S. On the one hand, if we write $S = \{s_1, s_2, \ldots, s_n\}$, we can determine any subset $X \subseteq S$ by asking n questions: "is s_1 in X?", "is s_2 in X?", and so on until "is s_n in X?". Each question has two possible answers, and so by the product rule there are 2^n possible subsets.

On the other hand, we can classify the subsets based on their size: $\binom{S}{k}$ is the collection of subsets of S of size k. As the subsets can range in size from 0 to n, we have $2^S = \dot\bigcup_{k=0}^{n} \binom{S}{k}$. By the sum rule, $\left|2^S\right| = \sum_{k=0}^{n} \left|\binom{S}{k}\right|$. By definition, $\left|\binom{S}{k}\right| = \binom{n}{k}$, and so we have

$$2^n = \left|2^S\right| = \sum_{k=0}^{n} \binom{n}{k}. \qquad \square$$

Proof of Corollary 4(ii). Let S be a set of n elements, $S = \{s_1, s_2, \ldots, s_n\}$. Since $\binom{n}{k}$ counts the number of subsets of S of size k, $\sum_{k \text{ odd}} \binom{n}{k}$ counts the number of subsets of S of odd size. Similarly, $\sum_{k \text{ even}} \binom{n}{k}$ enumerates the subsets of S of even size. To prove the identity, we build a bijection $f : 2^X \to 2^X$.

Given $X \subseteq S$, define

$$f(X) = \begin{cases} X \cup \{s_n\} & \text{if } s_n \notin X \\ X \setminus \{s_n\} & \text{if } s_n \in X \end{cases}.$$

For every set X, we have $f(f(X)) = X$, since we either add and remove s_n to X, or remove and add back s_n to X. In particular, f is invertible, and so a bijection. Moreover, note that $f(X)$ changes the size of X by exactly one, and so f maps odd subsets to even subsets, and even subsets to odd subsets. Hence, restricting to the collection of odd-sized subsets, f gives a bijection to the collection of even-sized subsets. This shows that the two sides of the identity are indeed equal. $\square$

[6] Our proof of Corollary 4(ii) below is an example of this phenomenon.

Proof of Corollary 4(iii). Let S be a set of n elements. We know that $\binom{n}{k}$ denotes the number of subsets of S of size k. From (i), we know that 2^k denotes the number of subsets of a set of size k. Hence, by the product rule, $\binom{n}{k}2^k$ counts the number of ways of choosing a subset $B \subseteq S$ of size k, and then choosing a further subset $A \subseteq B$ (of arbitrary size). Taking a sum over all k between 0 and n then enumerates over all possible sizes of the subset B, and hence over all possible choices of B. Thus the left-hand side counts the number of ways of choosing sets A and B such that $A \subseteq B \subseteq S$.

Next we claim that 3^n counts the number of ways of partitioning S into three sets, so that $S = C \dot{\cup} D \dot{\cup} E$. Indeed, for every element $s_i \in S$, we ask which subset s_i belongs to. As there are three choices for each element (C, D or E), there are 3^n possible partitions.

To finish, we biject between the pairs (A, B) with $A \subseteq B \subseteq S$ and the triples (C, D, E) with $S = C \dot{\cup} D \dot{\cup} E$. We can define

$$f(A, B) = (A, B \setminus A, S \setminus B).$$

Since $A \subseteq B \subseteq S$, it follows that $A, B \setminus A$ and $S \setminus B$ are pairwise-disjoint. As $A \cup (B \setminus A) \cup (S \setminus B) = S$, it is indeed true that $f(A, B)$ gives a partition of S into three sets. To see that f is bjiective, we observe that its inverse is given by

$$f^{-1}(C, D, E) = (C, C \cup D).$$

The bijection shows that the number of pairs (A, B) is equal to the number of triples (C, D, E), and so $\sum_{k=0}^{n} \binom{n}{k} 2^k = 3^n$. □

Conclusion

Binomials coefficients are omnipresent in combinatorics, arising naturally in several contexts. While they can sometimes be difficult to work with directly, binomial identities often allow us to simplify expressions involving these coefficients, and hence it is useful to have a large collection of identities.[7]

The Binomial Theorem is a great source of identities, together with quick and short proofs of them. However, given that binomial coefficients are inherently related to enumerating sets, combinatorial proofs are often more natural, being easier to visualise and understand. Furthermore, they can lead to generalisations and further identities.

There are also several identities that do not follow from the Binomial Theorem, and you will meet a few on your homework. While it is again sometimes possible to prove these by using the formula for the binomial coefficients and slogging through some algebra, a combinatorial proof is usually preferable.

The above examples may have seemed rather mundane, with more work required for little reward. However, there are several examples in enumerative combinatorics of identities for which analytic proofs are known, but combinatorial proofs are desired for the extra insight they would bring. After all, mathematics is more about explaining *why* things are true rather than merely determining that they are.

[7] Could a shortage of these equalities be referred to as an identity crisis?