Newton Day 2017:
Newton's generalization of the binomial theorem

by Peter H. Mao for Margaret and Owen

December 25, 2017

Abstract

For Sir Isaac Newton's 375th birthday we consider his generalization of the binomial theorem. I'll give an overview of the counting techniques needed to calculate traditional binomial coefficients, which involve only integers, and then show how Newton extended the concept to rational numbers.

1 Introduction

Binomial coefficients, which underlie the binomial theorem, are either the simplest of all non-trivial counting problems or the most complicated of all trivial counting problems. Both the theorem and its vaunted coefficients are frequently encountered in physics, mathematics, and engineering. The coefficients are the number of (unordered) subsets of size k drawn from a set of size n. My favorite binomial coefficient question is, "How many k card hands are possible with an n card deck?"[1] When we talk about cards, it is obvious that n and k have to be integers – only magicians rip cards into pieces. Newton was a magician, in that his generalized binomial coefficients allow the number of cards in the deck to be a non-integer, though he still deals whole cards.

2 Counting

When we count, usually we are counting things. For example, the things might be fingers and toes, colored balls or beads, wins and losses, playing cards, or time. Depending on the question at hand, we might return our thing to the bag or pile from which they came, or we might not. The former case is counting with replacement, while the latter is counting without replacement.

2.1 Counting with replacement: n^k

An example of counting with replacement is the number of passwords of length k that can be generated from a set of n characters. Since we can choose the same characters an arbitrary number of times, in essence, we are returning the character to the set of available characters

[1] Maggie is *very* familiar with this question.

each time we choose. For simplicity, lets use the 26 capital letters of the English language as our set of characters.

1. How many one-letter passwords are there?

I think you know the answer – this would not provide much security, would it?

2. How many two-letter passwords are there?

We have 26 ways to choose the first letter and 26 ways to choose the second letter, so that makes 676 passwords.

3. How many three-letter passwords are there?

Is this getting repetitive yet? Twenty-six choices for the first character, 26 for the second and another 26 for the third, so we find that there are $26 \times 26 \times 26 = 26^3 = 17576$ three-character passwords.

4. How many 8-16 character passwords exist? If you were trying to hack into a computer, how long would it take you the check all of them?

I'll leave this last question as an exercise for you.

The essential point here is that for each character in the password, one considers the number of choices (which is always the same) and takes the product of the choices. We'll see that this general technique carries over to counting *without* replacement.

2.2 Counting without replacement: $\frac{n!}{(n-k)!}$

In counting without replacement, we don't return letters to the pile. Figuring out the number of passwords that you can make with a set of alphabet blocks or a set of Scrabble tiles are examples of counting without replacement. Let's see what happens to the number of passwords when we can't use any character more than once.

1. How many one-letter passwords are there?

Since we only draw from the pile once, the answer is the same as the counting *with* replacement case.

2. How many two-letter passwords are there?

Now we see something a bit more interesting. You have 26 choices for the first character, but only 25 for the second. That makes 650 passwords with unique letters, down from 675 for the previous case.

3. How many three-letter passwords are there?

Can you extend the technique from two to three characters? The answer is 15600, but I'll let you figure out how to get that number.

4. How many 26-letter passwords are there?

As in the previous cases, we have 26 choices for the first character, and one less for each subsequent character. Since we are using all 26 characters, there is only one choice for the last character of each password. As we've done before, the number of passwords is the product of the number of choices for each character: $26 \times 25 \times 24 \times \cdots \times 3 \times 2 \times 1$. The modern notation for this product, 26!, was introduced by the French mathematician Christian Kramp in 1808.[1] Read it as "26 factorial." The number of 26-letter passwords is enormous – just a shade more than 4×10^{26}, a number that easily exceeds the number of grains of sand in Santa Monica.

In general, the number of ways to order n things is $n!$, but if we are ordering k things from a larger set of n things, then the number of ways to do this is $n \times (n-1) \times \cdots \times (n-k+1)$. We can write this expression more compactly as $n!/(n-k)!$, which leads us to a different interpretation of the number-of-passwords question. The numerator, $n!$, is, as we've seen, the number of ways to order all n things, while the denominator, $(n-k)!$, is the number of ways to order the things that were *not* picked for the password. This happens to be an example of an important technique in counting – we can over-count the number of ways to do something, and then divide by the number of ways that are essentially the same to correct our counting.

To illustrate this point, let's consider a 4-character set $\{A, B, C, D\}$ and the number of 2-character passwords drawn from that set without replacement. Using the first technique introduced, you quickly arrive at $4 \times 3 = 12$. Using the newer technique, you would note that there are $4! = 24$ ways to order the set and $2! = 2$ ways to order the remaining characters, so we get $4!/2! = 12$ as you would expect. Concretely, in the second interpretation, we recognize that $ABCD$ and $ABDC$ are different orderings of our little set, but we treat both as equivalent 4-character representations of the two-character password AB.

Now, returning to our 26-character set of capital letters, I ask you:

4. How many 27-letter passwords are there?

Since we are constructing our passwords without replacement, we'll run out of choices for the last character of our password, so the number of such passwords is zero. Using the first interpretation, the last term in our product is 0, so the product is 0. Easy. Using the second interpretation, we write the answer as $26!/(26-27)! = 26!/(-1)!$, which we expect to equal 0. What does this tell us about $(-1)!$?

2.3 Further examples

2.3.1 Labeling a 6-sided die

Suppose I have a bunch of six-sided dice for which the manufacturer failed to label the sides. I want the sides uniquely labeled 1–6. How many ways can I label the sides of the dice?[2] Is this a "counting with replacement" or a "counting without replacement" problem? Are there any other constraints?

Standard six-sided dice are made with the constraint that the values on opposite sides always adds to 7. In Figure 1, Owen has kindly provided me with art work showing the rotations of a standard die. To make his life easy, I asked him to only show two sides for each rotation. The constraint on the numbering of standard dice means that there is always a vertex with 1, 2, and 3 adjacent to it (Figure 2). To my knowledge, however, there is no constraint on the ordering of those three sides. They may run clockwise (left-handed) or counter-clockwise (right-handed). Are the dice in your games predominantly left-handed or right-handed?

2.3.2 From beads on a string to beads on a key-ring

The examples here further illustrate the utility of over-counting and dividing to count the number of orderings of a set.

Suppose you have n distinctly colored beads and one glass bead of the same size, which you tie to the end of string. How many ways can you thread the n beads onto the string? On the left side of Figure 3, I've drawn two possible arrangements of 5 beads on a string.

[2]Disregard the orientation of the label on the face, whether it be dots or Arabic numerals.

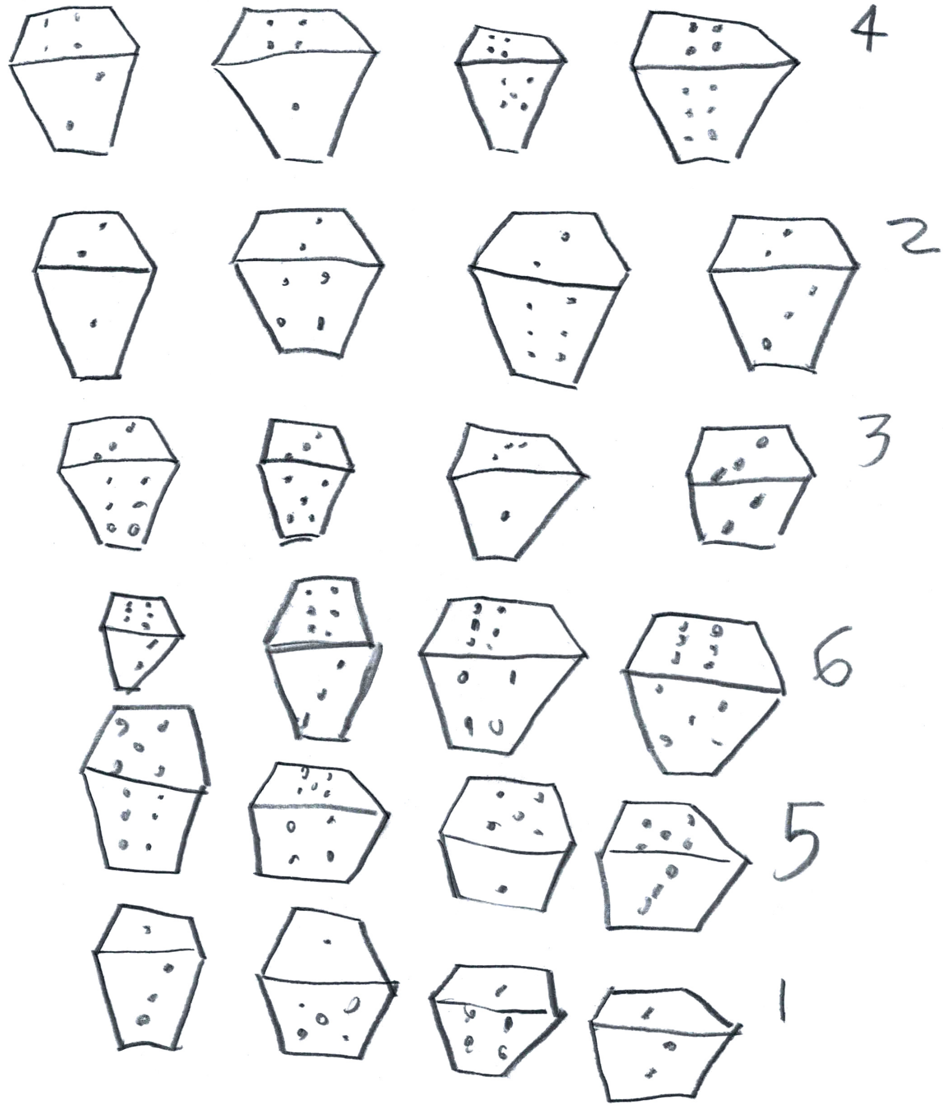

Figure 1: The 24 rotations of a standard die.

Now, tie the loose end of the string to the glass bead with the exact same knot, so you can't tell which bead went on first from which one went on last, as shown on the right side of Figure 3. How many distinct arrangements are there now?

Finally, take the beads off the string and thread them onto a split-wire key-ring (Figure 4). With no glass bead in the way, your beads are free to move on the key-ring, but not able to change their ordering. How many ways are there to arrange your n beads on a key-ring?

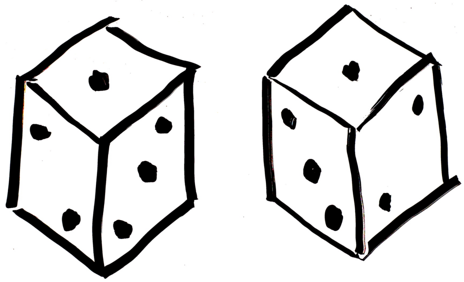

Figure 2: Right-handed and left-handed standard dice.

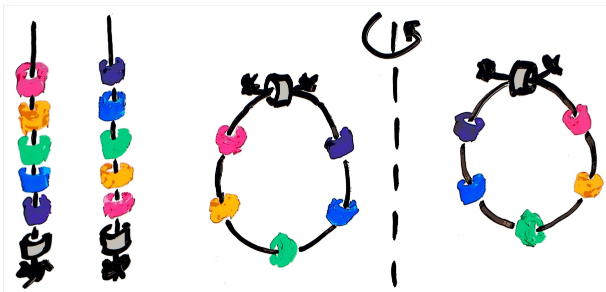

Figure 3: Beads on a string. By tying the loose end to the stopper bead, two different arrangements of the beads become the same.

2.4 Recap of counting

So far, we've only talked about counting ordered lists (passwords) of length k with the things in the lists drawn from a set of n things (capital letters). Obviously, k and n are non-negative integers. We might draw from the set with or without replacement, and we find the the number of unique lists that we can make is a product of the number of choices we have for each element of the list. We also found that we can arrive at the same answer by over-counting and then dividing by the number of equivalent representations.

The division trick does not always work. As a counter-example, consider counting the number of two-character English words that can be generated from the set $\{A, B, C, D\}$. The result is smaller than 12, might not divide 24, and would generate lots of arguments over that is and is not in the language. For our discussions here, the division trick is wonderful, but in the wider world, exercise caution. I'm sure there's a precise mathematical formulation for when it works (when there is a surjective homomorphic mapping?) but I haven't thought it through carefully.

Figure 4: Arrangements of 5 beads on a ring.

3 Intermission:[3] games of Newton Day

I always try to introduce a Newton-themed activity to the day's festivities. Newton's generalization of the binomial theorem really doesn't directly lend itself to any such activities, so I went with games loosely associated with counting: the fifteen puzzle, Rubik's Cube and Grime dice.

The fifteen puzzle (see Figure 5) is often associated with Sam Loyd, who offered $1000 for a solution to an unsolvable configuration of the puzzle. It is well known that an odd

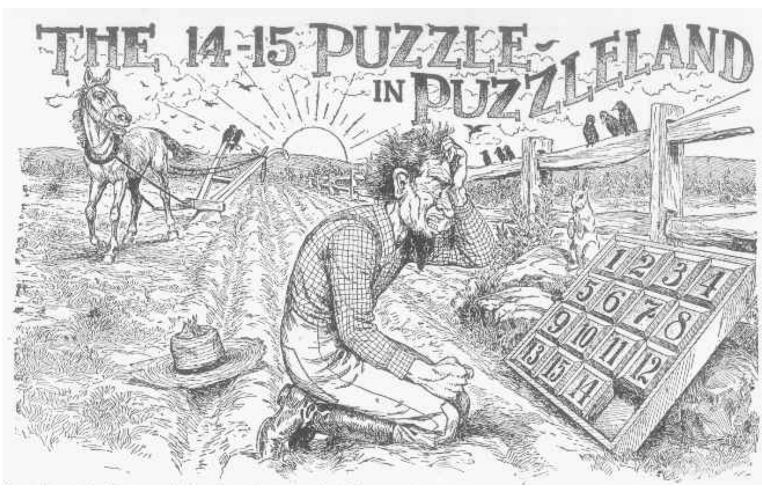

Figure 5: Sam Loyd's illustration of his unsolvable 15-puzzle (*Cyclopedia of 5000 Puzzles* (1914) by Sam Loyd, p. 235).

number of tile transpositions makes the puzzle unsolvable. Are all of the other ≈ 654 billion arrangements of the puzzle solvable? The answer was explored in full generality (for "non-bipartite graphs with at least nine vertices") by R. M. Wilson in 1974.[2]

The Rubik's Cube has been a cultural icon since its introduction in 1980. I was under the impression that the state of the cube with maximal distance from the solution was unknown, since the cube has over 43×10^{18} configurations. I recently found that this is a solved problem! In 2010, Tomas Rokicki, Herbert Kociemba, Morley Davidson, and John Dethridge

[3] I learned about "intermission" from Abhay Pasupathy's talk at the weekly Caltech physics colloquium in November, 2016.

used symmetry arguments and 35 CPU-years of computing time to determine that in the "half-turn metric," the maximum is 20 moves.[3] Four years later, Rokicki and Davidson used another 29 CPU-years of computing time to establish that in the "quarter-turn metric," the maximum number of moves is 26.[4] Kirk McKenzie introduced me to the $2 \times 2 \times 2$ cube when we were camping earlier this year. Oddly, the Rubik's version is marketed as a "2×2" cube. How many permutations does this have and how far can we get from the solution?

The last of this year's games is "Grime Dice."[5, 6] These are an invention of the British mathematician, James Grime, of Numberphile fame. They are an example of a nontransitive game, like "rock, paper, scissors." Grime Dice consists of 5 dice with nonstandard numberings such that each die has better-than-even odds of beating two other dice and worse-than-even odds of beating the other two. I couldn't get my hands on manufactured Grime Dice, so I got a big pile of blanks that we'll mark up ourselves and play.

4 The binomial theorem

How many k-card hands are possible from a deck of n cards? How many ways can I get k heads in n coin flips? These are the classic applications of the binomial theorem.

Let's consider the card problem first. There are n possibilities for the top card in the deck, then $n - 1$ possibilities for the next card, and so on. Is this the same as "counting without replacement" that we discussed before the Intermission? Almost, but not exactly – a card hand has no ordering, since the player is free to rearrange his or her cards at will. The number of ordered hands is $n!/(n-k)!$, but we have over-counted. Since our k-card hand can be rearranged in $k!$ ways, we arrive at the number of k-card hands by dividing the number of ordered hands by the number of rearrangements: $\frac{n!}{k!(n-k)!}$. This result is so common that it gets special notation:

$$\frac{n!}{k!(n-k)!} = \binom{n}{k} \tag{4.1}$$

and is read, "n choose k."

What about the coin flipping problem? It's exactly the same as the card problem. Make a correspondence between each coin flip with a specific card in the deck and deal the cards for which the coin lands "heads" to the card player. For example, I can relate the cards 10, J, Q, K, and A to the 1st, 2nd, 3rd, 4th, and 5th coin flips respectively (see Figure 6). The situation where the three heads occur on the 2nd, 4th, and 5th flips corresponds to the situation where I am dealt J, K, and A from the 5-card deck. Each unique 3-card hand from the 5-card deck corresponds to a unique sequence of 5 coin-flips, in which exactly 3 are heads, so the problems are isomorphic.

The word "binomial" suggests that this has something to do with two numbers. Indeed it does – all we've done so far is to get the coefficients of the binomial theorem. The original intent of the binomial theorem was to expand the expression

$$(t + h)^n = (t + h) \cdot (t + h) \cdots (t + h) \tag{4.2}$$

where t and h are numbers. After we multiply everything out, we'll have terms that look like $tthhtthhhttt\cdots$, $hhththtththtt\cdots$, and even $tttttttttttt\cdots$. The common feature among all the terms is that they all consist of exactly n symbols. Multiplication commutes, so we can rearrange each term so that the t's come before the h's, so for terms with k h's, we write them

Figure 6: Explicit correspondence between results of coin flips and cards in a deck. Each of the $\binom{5}{3}$ ways to get 3 heads corresponds to a unique 3-card hand from the 5 card deck.

as $t^{n-k}h^k$. How many terms do we have that look exactly the same? In other words, for each value of k ranging from 0 to n, how many ways were there to get exactly k h's? How many ways are there to get k heads in n coin flips? Our counting game solves the expansion of the binomial $(t+h)^n$:

$$(t+h)^n = \sum_{k=0}^{n} \binom{n}{k} t^{n-k} h^k \tag{4.3}$$

5 Newton's generalization of the binomial theorem

Now that you know the binomial theorem, we can finally get to the seed of this year's topic – Newton's generalization of the binomial theorem. He is said to have discovered this result during the Plague year, 1665, though the first communication of the result was in a letter to Leibniz written in 1676.[7] Admittedly, this is a bit of an oddball topic, since it likely ranks as Newton's least important contribution to mathematics. Historically, it may have been important in the development of calculus, but from the modern perspective, it is a largely irrelevant result.

Despite my disparaging remarks, let's see what Newton's generalization is all about. The binomial theorem tells us that the coefficients of the expansion of $(t+h)^n$ are of the form

$$\binom{n}{k} = \frac{n \times (n-1) \times \cdots \times (n-k+1)}{k!}. \tag{5.1}$$

This is, of course, equivalent to Equation 4.1, but with the $(n-k)!$ term in the denominator canceled out. Newton's generalization comes from the question of how to expand a binomial taken to a fractional power: $(t+h)^{m/n}$. Let $q = m/n$, then take the kth term in the expansion of $(t+h)^q$ to be

$$\binom{q}{k}_N = \frac{q \times (q-1) \times \cdots \times (q-k+1)}{k!}. \tag{5.2}$$

This looks a *lot* like Equation 5.1, but there's an important difference between these two equations. In Equation 5.1, k and n are both integers, so for $k > n$, there is always a term in the numerator that is equal to zero. Thus, the binomial coefficient is identically zero for all $k > n$. In Equation 5.2, if n does not divide m, then $q = m/n$ is *not* an integer, so the generalized binomial coefficient is nonzero for all values of k. Newton's generalization of the binomial theorem thus expands fractional powers of binomials as the infinite series:

$$(t+h)^q = \sum_{k=0}^{\infty} \binom{q}{k}_N t^{q-k} h^k. \tag{5.3}$$

Newton never formally proved this result.

I mentioned that $\binom{n}{k}$ is read as "n choose k." Can Newton's generalization be read the same way? Does it make sense to choose an integer subset of things from a non-integral set of things, like a deck of $51\frac{1}{2}$ cards? Can you flip a coin $5\frac{2}{3}$ times? One way to reconcile the physical interpretation of the binomial coefficients with Newton's generalization is to consider Newton's generalization as some sort of interpolation between the integer binomial coefficients. In this way, we find a tenuous philosophical thread between Newton's generalized binomial coefficients and Euler's gamma function, the "solution par excellence" to the interpolation of the factorial function, according to P. J. Davis.[8]

If you were to pick random numbers for t, h, and q and try to verify Equation 5.3, you would probably find that the series does not converge. One example that does work is calculating $\sqrt{2}$. First, cast this number into the form of a generalized binomial,

$$\sqrt{2} = 2^{\frac{1}{2}} = (1+1)^{\frac{1}{2}} \tag{5.4}$$

with $t = 1$, $h = 1$, and $q = \frac{1}{2}$. Being the identity of multiplication, 1 to any power is itself; therefore,

$$\begin{aligned}(1+1)^{\frac{1}{2}} &= \sum_{k=0}^{\infty} \binom{\frac{1}{2}}{k}_N \\ &= 1 + \frac{1}{2} - \frac{1}{8} + \frac{1}{16} - \frac{5}{128} + \frac{7}{256} - \cdots\end{aligned} \tag{5.5}$$

If you pull out a calculator, you'll find that this does indeed converge to $\sqrt{2}$, but it does so ever so slowly. If you use Newton-Raphson instead (by following the tangent of $x^2 - 2$ from your initial guess to the x-axis, you'll find that you do better after 1 step with a reasonable guess ($1.4^2 = 1.96$) than with 100 terms of the generalized binomial expansion, and by the 3rd step, you will have hit machine precision (10^{-16}). We'll do Newton-Raphson another year.

Equation 5.5 may not be an efficient way to calculate $\sqrt{2}$, but it's still a little special – $\sqrt{2}$ is famously irrational,[4] yet the right hand side of the equation consists of an infinite series of rational numbers.

6 Conclusion

The methods of counting described herein are only those needed to understand the coefficients of the binomial theorem. Since the goal this year was to get to Newton's generalization of the binomial theorem, I made no attempt to be complete with either counting or the binomial theorem itself. I hope that if you were not already good friends with the binomial theorem, that you now are, and that you find joy and pleasure in your future encounters with these numbers.

[4] Legend has it that the ancient Pythagoreans executed the discoverer of this property.

Figure 7: Owen and Maggie, December 24, 2017.

References

[1] https://en.wikipedia.org/wiki/Christian_Kramp

[2] Richard M. Wilson, "Graph Puzzles, Homotopy, and the Alternating Group," *J. Comb. Theory (B)* **16**, 86-96 (1974). (DOI: 10.1016/0095-8956(74)90098-7)

[3] www.cube20.org

[4] www.cube20.org/qtm

[5] http://grime.s3-website-eu-west-1.amazonaws.com/

[6] James Grime, "The Bizarre World of Nontransitive Dice: Games for Two or More Players," *College Math. J.*, **48** 1, 2-9 (Jan 2017). (DOI: 10.4169/college.math.j.48.1.2)

[7] Rob Iliffe, *Newton: A Very Short Introduction*, Oxford University Press, 2007.

[8] Philip J. Davis, "Leonhard Euler's integral: A historical profile of the gamma function," *Amer. Math. Monthly* **66** 10, 849-869 (Dec 1959). (DOI: 10.2307/2309786)