21128/15151 Counting Things

Emily Zhu

1 Counting Numerically

Counting is generally really difficult; there are many sets which cannot be counted in any nice way, and even if a set is nice to count, there are still many ways to accidentally undercount/overcount and relatively few ways to count correctly. Even so, that doesn't stop us from trying!

1.1 Fundamental Principles

Proposition 1 (Addition Principle)**.** If a finite set S is partitioned into $S_1, \ldots, S_k$, i.e. $\bigcup_{i=1}^k S_i = S$ and $S_i \cap S_j = \varnothing$ for $i \neq j$, then $|S| = \sum_{i=1}^k |S_i|$.

Informally, this states that if the thing we're trying to count can be completely split into some k nonoverlapping cases, then we can add the results we get for the cases to get what we're trying to count. For this, it is important to explicitly check that your cases actually cover everything and that they do not overlap. Often, these are pretty obvious, but nonzero work should be shown, especially in claiming that everything is attained. Technically, we should also be showing that our cases give us elements in the desired set, but we normally define the cases to do exactly that.

Proposition 2 (Multiplication Principle (informal))**.** If a set of n-tuples is created by picking coordinates in a process such that there are a_1 choices for the first coordinate, and given any choice of values of the first i coordinates, there are a_{i+1} choices for the $i+1^{st}$ coordinate, then the number of such n-tuples is $\prod_{i=1}^n a_i$.

Even more informally, this states that if the thing we're trying to count is found by some process in which each step always has the same number of choices (regardless of everything else), then we can multiply these numbers of choices to get what we're trying to count. It is extremely important to note that generates a tuple; often, we want to count the number of sets rather than tuples. We can work around this by treating these sets as having a specific ordering (which we get to decide), and then counting the number of tuples with that precise ordering.

Proposition 3 (Complementary Counting)**.** Given a finite set X and a subset $U \subseteq X$. Then $|U| = |X| - |X \setminus U|$.

Some times it is easier to count the complement than what you originally wanted—this is often the case when your original set is complicated or filled with many cases. A keyword which may suggest complementary counting is something like "at least" or "at most," where the complement has very few cases when looking at exact quantities. Of course, if the original "at least" only requires 2 exact cases when the complement requires like 4, attempting to use the complement is just much more work.

1.2 Checking Your Answer

After writing down a process, it can still be unclear if your process actually produces the set you wanted—perhaps it may have missed some elements or repeated some elements. (Formally, we want a bijection from the tuples found to whatever set we desire; typically we won't make any such bijection explicit.) It may be helpful to think about what the process produces and consider a sample element that we want to get from the set. Given this sample object, can you attain it somehow in your process? Can you attain it multiple ways? The latter is very commonly a problem which emerges when trying to count sets. Steps which say "pick anything" are prone to overcounting, especially if there is some other step with overlapping choices. Attempting to run your process may also help you to determine how to specify your process in your writeup.

Another common way to check your answer is to attempt to count your set in a different way—perhaps you can change the order in which you pick things, perhaps you can use complementary counting, or perhaps you can interpret your set differently (i.e. change the scenario you are in; I for one, like thinking about lattice paths).

Yet another way to check your answer (or actually to approach the problem) is to try a much smaller case of the same problem or a similar problem, so you can explicitly list out cases. This is often tedious and prone to other errors (maybe you made a different problem or you made a mistake listing out cases), but it can still be helpful in attempting to make your problem more concrete/approachable.

1.3 Lots of Problems

Example 1 (Warmup)**.** How many ways are there to label the 4 corners of a square with letters (i.e. A to Z) such that adjacent corners get different letters?

Thoughts. We begin labeling the top left corner—26 choices. We label the top right and bottom left—25 choices each. Now when we get to labeling the bottom right corner, we need it to be different from the top right and bottom left. At this point you may be tempted to say there are 24 choices left but if the top right and bottom left corner are the same, then there are actually 25 choices! We fix our argument to case on if corners are the same and then proceed to write it up...

Proof. We case on whether or not the top left and bottom right corners have the same label. Note that there are exactly two possibilities: either they have the same label or they have different labels. Since the second possibility is really the negation of the first possibility, one must happen, and clearly the corners cannot simultaneously have the same and different labels. Thus, we have a partition and can apply the addition principle.

- Case 1: Top left and bottom right corners have the same label. Our process is then:
  1. Pick the label for the top left corner. There are 26 choices.
  2. Pick the label for the bottom right corner. This is fixed to be the same as the top left, so we have 1 choice.
  3. Pick the label for the top right corner. Since this label must differ from that of the top left and bottom right, we omit that label and only have 25 choices.
  4. Pick the label for the bottom left corner. Analgously to the previous case, this label must different from that of the top left and bottom right, so there are 25 choices.

  By the multiplication principle, we have $26 \cdot 1 \cdot 25 \cdot 25$ ways in this case.

- Case 2: Top left and bottom right corners have different labels. Our process is then:
  1. Pick the label for the top left corner. There are 26 choices.
  2. Pick the label for the bottom right corner. It must differ from the top left corner, so we omit that label to get 25 choices.

3. Pick the label for the top right corner. It must differ from the top left and bottom right, so we omit those 2 labels to get 24 choices.
4. Pick the label for the bottom left corner. Again it differs from the top left and bottom right, so we have 24 choices.

By the multiplication principle, we have $26 \cdot 25 \cdot 24 \cdot 24$ ways in this case.

Combining the cases via the addition principle, we have a total of $26 \cdot 25^2 + 26 \cdot 25 \cdot 24^2 = 390,650$ ways to label the square. □

Example 2 (Two Suits)**.** How many hands of 4 cards from a standard 52-card deck contain cards of exactly 2 suits?

Thoughts. I actually screwed up many times when trying to count this. The main idea is that we want to pick which 2 suits appear and then pick the 4 cards from these suits. This suggest $\binom{4}{2}\binom{26}{4}$—but this does not work since it also includes cases where only 1 suit appears. You may then think that you just need to remove the number of ways to get 1 suit, but it turns out that this actually counts the number of ways to get cards of a single suit multiple times! Thus, it may be better to explicitly case on how many cards of each suit appear.

Proof. We case on whether there is 1 card of a suit and 3 cards of the other or 2 cards of both suits. Note that the suit with fewer cards must still have at least 1 card, and it cannot have more than 2 cards, since then it would have more than half of the total cards and so more cards than the other suit. Thus, we cover all cases. Clearly you cannot have different amounts of cards of a single suit in the same hand, so these cases are also disjoint. Thus, these form a partition and we can apply the addition principle.

- Case 1: 1 card of a suit and 3 cards of the other
  1. Pick the suit for the single card: $\binom{4}{1}$
  2. Pick the rank for the single card: $\binom{13}{1}$
  3. Pick the suit for the triple of cards: $\binom{3}{1}$ (note: we can't re-pick the first suit)
  4. Pick the ranks for the triple of cards: $\binom{13}{3}$

  By the multiplication principle, we have $4 \cdot 13 \cdot 3 \cdot \binom{13}{3}$ ways in the case.
- Case 2: 2 cards of each suit
  1. Pick the 2 suits (note that we do not care which order the suits are in): $\binom{4}{2}$
  2. Pick the 2 ranks from the suit first in alphabetical order: $\binom{13}{2}$
  3. Pick the 2 ranks from the remaining suit: $\binom{13}{2}$

  By the multiplication principle, we have $\binom{4}{2}\binom{13}{2}\binom{13}{2}$ ways in the case.

Combining via the addition principle, we have a total of $4 \cdot 13 \cdot 3 \cdot \binom{13}{3} + \binom{4}{2}\binom{13}{2}\binom{13}{2} = 81,120$ ways to pick a hand of 4 cards such that exactly 2 suits appear. □

We can also write an alternative proof using inclusion/exclusion and complementary counting which basically does what the original thoughts were.

Proof. (Inclusion/Exclusion + Complementary). We find the number of hands of 4 cards with exactly 2 suits by finding the number of hands of 4 cards such that at most 2 suits appear and then subtracting out the number of hands of 4 cards such that at most 1 suit appears. Note if at most 1 suit appears, then we have that exactly 1 suit appears.

- At most 2 suits: Define sets $S_1, \ldots, S_6$ to have hands of 4 cards contained among some set of 2 suits: in particular, we can let S_1 be hands of spades/hearts, S_2 of spades/clubs, S_3 of spades/diamond, S_4 of hearts/clubs, S_5 of hearts/diamonds, S_6 of clubs/diamonds. Note then that we want to find $\left|\bigcup_{i=1}^6 S_i\right|$ and so we proceed by inclusion/exclusion. First note that $|S_i| = \binom{26}{4}$ since we pick 4 cards from the 2 suits.

  Note that $S_i \cap S_j$ (for $i \neq j$) either is empty if there are no suits in common or is the set consisting of cards only from the single overlapping suit (if both suits were shared, the sets would be the same). In particular, there are $4 \times \binom{3}{2} = 12$ pairs of overlapping sets—4 ways to pick the suit they overlap in and $\binom{3}{2}$ ways to pick the other two suits for these sets. In this case $|S_i \cap S_j| = \binom{13}{4}$, the number of ways to pick 4 cards from a single suit.

  Similarly, $S_i \cap S_j \cap S_k$ for i, j, k different is only nonempty if all 3 have some suit in common. Note that it is not possible for the 3 sets to share 2 suits just as in the previous case. Then, there are $4 \times \binom{3}{3} = 4$ such triples of sets. Again, we have $|S_i \cap S_j \cap S_k| = \binom{13}{4}$, the number of ways to pick 4 cards from a single suit.

  We then apply Inclusion/Exclusion to get that

  $$\left|\bigcup_{i=1}^{6} S_i\right| = \sum_{i=1}^{6} |S_i| - \sum_{\{i,j\}\subseteq[6]} |S_i \cap S_j| + \sum_{\{i,j,k\}\subseteq[6]} |S_i \cap S_j \cap S_k| = 6\binom{26}{4} - 12\binom{13}{4} + 4\binom{13}{4}$$

- 1 suit: There are 4 ways to pick a suit, and then $\binom{13}{4}$ ways to pick 4 cards from that suit. By the multiplication principle, there are $4\binom{13}{4}$ ways total.

Using complementary counting, we have that there is a total of $6\binom{26}{4} - 8\binom{13}{4} - 4\binom{13}{4} = 6\binom{26}{4} - 12\binom{13}{4} = 81,120$ ways to pick a hand of 4 cards such that exactly 2 suits appear. □

Example 3. How many ways are there for me to distribute some amount of my 10 (essentially indistinguishable) Sinosaurus plushies to 4 distinguishable people? Note: I may want to keep some.

Remark 4. We can rephrase this as how many solutions are there to $x_1 + x_2 + x_3 + x_4 \leq 10$, where $x_i \in \mathbb{N}$ for $i \in [4]$?

Thoughts/Proof. We can add me in as another person who gets a nonnegative number of plushies, since I just get whatever is left. Then, our problem becomes how many solutions are there to $\sum_{i=1}^{5} x_i = 10$ where $x_i \in \mathbb{N}$ for $i \in [5]$. We now know how to solve this—there are $\binom{10+5-1}{5-1} = \binom{14}{4} = 1001$ ways to distribute plushies! □

Remark 5. The technique used to count the number of solutions to $\sum_{i=1}^{k} x_i = n$ is often called stars and bars (or balls and urns). This is since we are trying to arrange n indistinguishable stars and $k-1$ indistinguishable dividing bars in $n+k-1$ slots. It's also important to remember that the space before the first and after the last divider also account for x_1 and x_n respectively, so we only need $k-1$ dividers.

Example 4. Given 2 red dice and 3 green dice, how many ways are there such that exactly 3 dice rolls have the same value and the remaining two have different values? Note: I cannot tell the red dice apart, so getting a 5 and a 3 on the red dice is the same as getting a 3 and a 5.

Proof. We case on the number of red dice which are a part of the triple of the same value. Note that since there are 2 red dice, either none of them, 1 of them, or both of them are in that triple. These cases are clearly disjoint, so we have a partition.

- No red dice:
  1. Pick the value for the triple (these are all green dice): 6 ways
  2. Pick the values for the 2 red dice: $\binom{5}{2}$ ways (order does not matter)

  By multiplication principle we have $6\binom{5}{2}$ ways in this case.

- One red die:
  1. Pick the value for the triple (2 green, 1 red): 6 ways
  2. Pick the value for the remaining green dice: 5 ways
  3. Pick the value for the remaining red dice: 4 ways

  Note here that the order of values for the remaining dice matters since they are of different colors. By the multiplication principle we have $6 \cdot 5 \cdot 4$ ways in this case.

- Two red dice:
  1. Pick the value for the triple (2 red, 1 green): 6 ways
  2. Pick the values for the remaining green dice: $\binom{5}{2}$ ways (order does not matter)

  By multiplication principle, we have $6\binom{5}{2}$ ways in this case.

By the addition principle, we get a total of $6\binom{5}{2} + 6 \cdot 5 \cdot 4 + 6\binom{5}{2} = 360$ ways. □

Definition 6 (Lattice Path)**.** A lattice path is some sequence of moves each going one unit right or up; often we consider it from $(0,0)$ to (m,n) for $m,n \in \mathbb{N}$.

Example 5. How many lattice paths are there from $(0,0)$ to (m,n)?

Proof. Note that in every such lattice path, we have precisely m right moves and n up moves. Then, it is just a matter of deciding which of the $m+n$ total moves are right moves, since then the remaining moves are just up moves. Thus, there are $\binom{m+n}{m}$ paths. □

Example 6. How many lattice paths from $(0,0)$ to $(5,3)$ avoid $(1,1),(1,2),(2,1),(2,2)$?

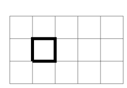


Figure 1: Avoid the darkened square

Proof. (Cases). Note that if we have both a right and an up in the first 3 moves, we intersect the square. Thus, the only ways to avoid the square are by going right 3 times or up 3 times for our first 3 moves. Clearly these do not overlap, so we have a partition of the desired paths.

- 3 rights: Then 2 rights and 3 ups remain for a total of $\binom{5}{2} = 10$ paths.
- 3 ups: Then 5 rights and 0 ups remain for a total of $\binom{5}{0} = 1$ path.

By the addition principle, we have a total of $10 + 1 = 11$ paths avoiding the square. □

Proof. (Complementary Counting). Note that it suffices to count the total number of paths and subtract the number of paths intersecting the square. There are $\binom{8}{3}$ total paths. Note that any path intersecting the square must either intersect $(1, 2)$ or $(2, 1)$—if it intersects $(1, 1)$, the next move brings us to one of these and if it intersects $(2, 2)$, we either came from an up (came from $(2, 1)$) or right (came from $(1, 2)$). Also no path can intersect both since that would require a left or down. This is then a partition.

- Reaches $(1, 2)$:
  1. Pick a path to get to $(1, 2)$: $\binom{3}{1}$ ways
  2. Pick a path to get from $(1, 2)$ to $(5, 3)$: $\binom{5}{1}$ ways (4 rights, 1 up)

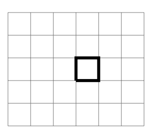
  By the multiplication principle, we have $3 \cdot 5$ ways in this case.

- Reaches $(2, 1)$:
  1. Pick a path to get to $(2, 1)$: $\binom{3}{1}$ ways
  2. Pick a path to get from $(2, 1)$ to $(5, 3)$: $\binom{5}{3}$ ways (3 rights, 2 ups)

  By the multiplication principle, we have $3\binom{5}{3}$ ways in this case.

By the addition principle, we have a total of $3 \cdot 5 + 3\binom{5}{3} = 45$ ways. We subtract this from the total number of paths to get $\binom{8}{3} - 3 \cdot 5 - 3\binom{5}{3} = 56 - 45 = 11$ paths avoiding the square. □

Remark 7. While in this case, the path avoiding the removed square was easy to directly case on, this is not generally the case. Consider running through the argument for:



Example 7. Find the number of lattice paths from $(0, 0)$ to (m, n) where you go left exactly once and the whole path remains within the rectangle formed by $(0, 0), (m, 0), (m, n), (0, n)$.

Proof. This ends up being a surprisingly simple process. We have a total of $m + n + 2$ moves, where $m + 2$ are horizontal moves and n are up's. We first pick which are the horizontal moves, of which there are $\binom{m+n+2}{m+2}$ ways. But then we must pick which of the horizontal moves is the left—it cannot be the first move, since then we'd go negative, and it cannot be the last move, since then we'd go too far right. However, any other moves keep us within the rectangle, so we have m choices for the left. By the multiplication principle, there are $m\binom{m+n+2}{m+2}$ ways. □

Example 8. How many functions are there from $[5] \to [8]$ which are not injective? What about from $[k] \to [n]$ in general, where $k \leq n$? (What happens if $k > n$?)

Thoughts. Please don't actually list them out. Explicitly trying to figure out where outputs match is also really difficult, but the "not" in the example should be quite suggestive...

Proof. We find the number of functions from $[k] \to [n]$ which are not injective by complementary counting. Note that there are a total of n^k functions—we choose one of n values for each of the k inputs. There are $n \cdot (n-1) \cdots (n-k+1) = \frac{n!}{(n-k)!}$ injective functions where we pick the output for each of the k inputs to be among the previously unpicked values. By complementary counting, we then get that there are $n^k - \frac{n!}{(n-k)!}$ not-injective functions.

In our specific case, that gives us $8^5 - \frac{8!}{3!} = 26048$ functions, which would've taken a very long time to list. ☐

Example 9. How many "words" can be formed from rearranging the letters of COMMITTEE?

Proof. We define a process to arrange the letters in the 9 positions as follows:

1. Pick 2 positions for the M's: $\binom{9}{2}$
2. Pick 2 positions for the T's: $\binom{7}{2}$
3. Pick 2 positions for the E's: $\binom{5}{2}$
4. Arrange the C, O, I within the remaining 3 positions: 3!

By the multiplication principle, we get that there are $\binom{9}{2}\binom{7}{2}\binom{5}{2}3! = \frac{9!}{2!2!2!} = 45360$ ways. ☐

Do you see where the second expression comes from? Also, if you happen to be me, the only "arrangement" that matters is the misspelling of "committees" as "commitees".

In counting, don't be afraid to try different approaches if your first approach seems tedious or seems to count the wrong thing. Do be very careful in constructing your processes though! Also remember that upon seeing solutions, they may seem very simple, but this does not mean that they are easy to find.

2 Counting in Two Ways

The main idea of counting in two ways is to show that two expressions are equal by exhibiting a set and showing that both sides count the size of this set. Note that by the uniqueness of the cardinality of a set, we can conclude that the two expressions are thus equal. It is also possible to have the expressions count two sets and then find a bijection between the sets—this may result in more work since you then need to define a function and prove that it is well-defined and bijective.

Perhaps the hardest part of counting in two ways is constructing the set; this is where you need to determine what your expressions can actually count and how they interact. Once you determine your set, you have likely figured out how your expressions count the set. In some sense, this reverses counting numerically, since you start with (abstract) sizes and try to build something of that size.

*Important: Please do **not** do algebraic manipulations to an expression when we ask you to prove it by counting in 2 ways! Also, "simple" manipulations should be fairly easy to reason about.*

2.1 Fundamental Expressions

When counting in two ways, it's very useful to have a sense of what the following can count/have prototypical examples of how to think about various expressions. This list is certainly nonexhaustive, and as you get practice, you should gain intuition on which of these may be easiest to work with. Note that in counting in two ways, you will often have a combination of these expressions, so it is good to understand how these sets interact.

- $\binom{n}{k}$: picking k things from n things
  - number of subsets of $[n]$ with size k
  - number of binary strings with k 1's (or k 0's)
  - number of lattice paths from $(0,0)$ to $(k, n-k)$ (see examples above)
- 2^n: making some binary choice n times (see a^n as well)
  - number of subsets of $[n]$
  - number of binary strings of length n
- a^n: deciding from a choices n times
  - number of functions $f : [n] \to [a]$
  - number of ways to color n distinguishable objects with a colors
  - number of ways to divide n things into a (possibly empty) groups
  - number of a-ary strings of length n (i.e. made of a symbols)
- $n!$: ordering n things
  - number of permutations of $[n]$ (or any other set)
- $\sum_{k=0}^{n}$: partitioning based on some k
  - k can represent the size of the sets in the partition
  - k can represent a special element in the sets (eg. the largest element; often you will pick fewer elements on one side for this case)
  - k can represent the number of elements we pick from some part of a partition
- n: pick 1 thing from n things
  - it can be helpful to interpret n as $\binom{n}{1}$ sometimes

There are other important identities (some of which we've alluded to in **Counting Numerically** such as $\left(\binom{n+k-1}{k-1}\right)$ which may also help you interpret the expressions in a more natural way. Perhaps one of the most important of those identities would be that $\binom{n}{k} = \binom{n}{n-k}$ so you can think about the things with that property (e.g. being in the set, being a 1) or without it. Also remember that order does not matter in multiplication, so rearranging the terms in the product may also help. Often expressions will be written in a way that looks aesthetically pleasing without regards to the process used to make them.

A good thing to think about is how many things you're picking and how many things you're picking from. From there, you also want to think about how these things interact—are they distinguishable? is it a single set? is a tuple of sets? a set of tuples? Or if you're working with a specific scenario, how exactly do these committees/cases interact? Do you have special elements? A lot of this intuition will come from practice, and the rest will come from just trying a lot of scenarios and attempting many matchings of "variables" to amounts of stuff getting picked.

2.2 Lots of Problems

(these are mostly sourced from *Mathematical Thinking: Problem-Solving and Proofs* by John. P. D'Angelo and Douglas B. West)
(I've also gotten lazy on showing intuition so you'll have to be at the session to hear about it!)

Example 10. By counting in 2 ways, show that $\sum_{k=1}^{n} 2^{k-1} = 2^n - 1$ for $n \in \mathbb{N}$.

Proof. Let S be the set of nonempty subsets of $[n]$. We prove that the LHS and RHS[1] count $|S|$. For the RHS, note that there are $2^n - 1$ nonempty subsets of $[n]$ since there are a total of 2^n subsets, and we must remove the one empty one.

For the LHS, let S_k be the set of subsets of $[n]$ with largest element k. Note that every nonempty subset of $[n]$ has a largest element which is between 1 and n, so $\bigcup_{k=1}^{n} S_k = [n]$ (technical note: this actually only shows that $\bigcup_{k=1}^{n} S_k \supseteq [n]$, but we had defined $S_k \subseteq S$, so $\bigcup_{k=1}^{n} S_k \subseteq S$). Furthermore note that the largest element of a set is unique, so $S_i \cap S_j = \varnothing$ for $i \neq j$. Thus, this defines a partition of S, so $\sum_{i=1}^{n} |S_k| = |S|$.

Now for a fixed k, we claim that $|S_k| = 2^{k-1}$. This k must be in any subset $[n]$ which is an element of S_k, and then we must decide which elements smaller than k are in the subset. There are $k-1$ elements smaller than k, each with a choice of being in or out of the set, so there are a total of 2^{k-1} subsets with k as the largest element.

Thus, we can conclude that $\sum_{i=1}^{k} 2^{k-1} = |S| = 2^n - 1$. □

Example 11. By counting in 2 ways, show that $\sum_{k=0}^{n} k\binom{n}{k} = n2^{n-1}$ for $n \in \mathbb{N}$. (This is also equal to $n\sum_{k=1}^{n}\binom{n-1}{k-1}$)

Proof. We count the number of committees of any size with a chairperson (who is in the committee) which can be formed from n distinguishable people in 2 ways. For the RHS, note that we can define a process by:

1. Pick a chairperson: n ways
2. Pick the rest of the committee; each person is either in or out of the committee and there are $n-1$ remaining people, so we have 2^{n-1} ways

[1]LHS = left hand side, RHS = right hand side

which gives us $n2^{n-1}$ ways by the multiplication principle.

For the LHS, we case on the number of people in the committee. Note that we must pick some number of people, and if k is the number of people in the committee, $0 \leq k \leq n$ since we have a nonnegative number of people in the committee and at most as many people as we have can be in the committee. Also the number of people in the committee cannot have 2 values, so these cases are distinct and form a partition.

Now for a fixed k, we claim that there are $k\binom{n}{k}$ ways to pick a committee of k people with a chairperson (who is in the committee). There are $\binom{n}{k}$ ways to pick the committee and k ways to pick the chairperson from the committee, which by the multiplication principle gives us $k\binom{n}{k}$ ways. By the addition principle, we thus have that $\sum_{k=0}^{n} k\binom{n}{k} = n2^{n-1}$. □

Example 12. By counting in 2 ways, show that $\sum_{i=0}^{k} \binom{m}{i}\binom{n}{k-i} = \binom{m+n}{k}$ for $m, n \in \mathbb{N}$.

This is also Example 4.2.51 in the Book.

Proof. We count the number of ways to pick k animals from $m + n$ distinguishable animals where m are birds and n are... not birds. For the RHS, note that we could simply pick k of the $m + n$ animals and be done and there $\binom{m+n}{k}$ ways to do so.

For the LHS, we case on the number of birds we pick. Note that we must pick some number of birds, and if i is the number of birds picked, clearly $0 \leq i \leq k$ (can only pick nonnegative number and can pick at most the total number of things we pick). Also we cannot pick different amounts of birds at the same time, so these cases are distinct and thus form a partition.

Now for a fixed i, note that the number of ways to pick k animals such that i are birds follows a 2-step process:

1. Pick the i birds: $\binom{m}{i}$ ways
2. Pick the remaining $k - i$ not-birds: $\binom{n}{k-i}$

which gives us $\binom{m}{i}\binom{n}{k-i}$ by the multiplication principle. By the addition principle, we thus have $\sum_{i=0}^{k} \binom{m}{i}\binom{n}{k-i} = \binom{n+m}{k}$. □

Example 13. By counting in 2 ways, show that $\sum_{k=-m}^{n} \binom{m+k}{r}\binom{n-k}{s} = \binom{m+n+1}{r+s+1}$ for $m, n, r, s \in \mathbb{N}$.

Proof. Let S be the set of subsets of $T = \{z \in \mathbb{Z} \mid -m \leq z \leq n\}$ of size $r + s + 1$. We prove that the RHS and LHS count $|S|$. First, $|T| = m + n + 1$. Then, for the RHS, note that the number of subsets of size $r + s + 1$ of a set of size $m + n + 1$ is simply $\binom{m+n+1}{r+s+1}$.

For the LHS, let S_k be the set of subsets of T of size $r+s+1$ such that the $r+1^{st}$ least element (i.e. the least element after removing the least element r times—there's really an induction argument hiding here) is k. Note that any subset of T of size $r + s + 1$ must have an $r + 1^{st}$ least element, and this element must be in T, so $-m \leq k \leq n$ and $\bigcup_{k=-m}^{n} S_k = S$. Furthermore, the $r + 1^{st}$ element is unique just as the least element is unique, so $S_i \cap S_j = \varnothing$ for $i \neq j$ and this indeed forms a partition, so by the addition principle, $\sum_{k=-m}^{n} |S_k| = |S|$.

Now for a fixed k, we find $|S_k|$. This is a 2-step process:

1. Pick the r elements less than k; we are picking from $\{z \in \mathbb{Z} \mid -m \leq z \leq k - 1\}$ which has size $(k - 1) + m + 1$ so there are $\binom{m+k}{r}$ ways
2. Pick the s elements greater than k; we are picking from $\{z \in \mathbb{Z} \mid k + 1 \leq z \leq n\}$ which has size $n - (k + 1) + 1$ so there are $\binom{n+k}{s}$ ways

so by the multiplication principle, there are a total of $\binom{m+k}{r}\binom{n+k}{s}$ ways. By the addition principle, we thus have $\sum_{k=-m}^{n}\binom{m+k}{r}\binom{n-k}{s} = |S| = \binom{m+n+1}{r+s+1}$. □

Remark 8. Note that for $S = \{z \in \mathbb{Z} \mid a \leq z \leq b\}$, $|S| = b - a + 1$ and not $b - a$. To convince yourself, define a bijection $f : [b - a + 1] \to S$ via $f(x) = a - 1 + x$. What happens when we plug in $f(b - a + 1)$? (Check that it's actually a bijection)

Example 14. By counting in 2 ways, show that $\binom{n}{k}\binom{n-k}{m} = \binom{n-m}{k}\binom{n}{m}$ for $n, k, m \in \mathbb{N}$.

Proof. Let $S = \{(A, B) \mid A, B \subseteq [n], A \cap B = \varnothing, |A| = k, |B| = m\}$. We show that the LHS and RHS count $|S|$. For the LHS, we have a 2-step process:

1. Pick A from $[n]$; since $|A| = k$, there are $\binom{n}{k}$ ways.
2. Pick B from $[n] \setminus A$ (which makes $A \cap B = \varnothing$); note that $|[n] \setminus A| = n - k$ and $|B| = m$, so there are $\binom{n-k}{m}$ ways.

By the multiplication principle, $|S| = \binom{n}{k}\binom{n-k}{m}$.

Now for the RHS, we also have a 2-step process:

1. Pick B from $[n]$; since $|B| = m$, there are $\binom{n}{m}$ ways.
2. Pick A from $[n] \setminus B$; note that $|[n] \setminus B| = n - m$ and $|A| = k$, so there are $\binom{n-m}{k}$ ways.

By the multiplication principle, we thus have $\binom{n}{k}\binom{n-k}{m} = |S| = \binom{n-m}{k}\binom{n}{m}$. □

Example 15. By counting in 2 ways, show that $\sum_{k=0}^{m}\binom{n}{k}\binom{m}{m-k}2^{n-k} = \sum_{l=0}^{n}\binom{n}{l}\binom{n+m-l}{m}$ for $m, n \in \mathbb{N}$. *Hint*[2]

Proof. Denote $[-n] := \{-z \mid z \in [n]\}$ and define

$$S = \{(A, B) \mid A \subseteq [-n] \cup [m], |A| = m, B \subseteq [-n], A \cap B = \varnothing\}$$

Intuitively, S is a set of pairs of subsets such that the first set has m elements and elements can be picked from $[m]$ or $[-n]$, and the second set (possibly empty) is just picked from whatever remains of $[-n]$. We show that the LHS and RHS count $|S|$.

For the RHS, let $S_l = \{(A, B) \in S \mid |B| = l\}$, i.e. the set of such pairs of subsets where we picked l elements of $[-n]$ for B. Note that we must pick some number of elements from $[-n]$, and we must have $0 \leq l \leq n$, since we are picking a nonnegative number of elements from at most n elements. We then have that $\bigcup_{l=0}^{n} S_l = S$. Furthermore, since the cardinality of a set is unique, $S_i \cap S_j = \varnothing$ for $i \neq j$. Thus, we have a partition and by the addition principle, $\sum_{l=0}^{n} |S_l| = |S|$.

For a fixed l, we claim that $|S_l| = \binom{n}{l}\binom{n+m-l}{m}$. This comes from a 2-step process:

1. Pick the l elements of $[-n]$ which belong to B: $\binom{n}{l}$ ways
2. Pick the m elements from the remaining elements of $[-n] \cup [m] \setminus B$; note that $|[-n] \cup [m] \setminus B| = n + m - l$ so there are $\binom{n+m-l}{m}$ ways

By the multiplication principle, there are $\binom{n}{l}\binom{n+m-l}{m}$ elements. So by the addition principle, $|S| = \sum_{l=0}^{n}\binom{n}{l}\binom{n+m-l}{m}$ and we're halfway there.

For the LHS, let $T_k = \{(A, B) \in S \mid |A \cap [-n]| = k\}$, i.e. the set of such pairs of subsets where we picked k elements of $[-n]$ for A. Note that we must have $0 \leq |A \cap [-n]| = k \leq |A| = m$, so

[2] Hint: pick 2 subsets

$\bigcup_{k=0}^{m} T_k = S$. Since cardinalities are unique, $T_i \cap T_j = \varnothing$ for $i \neq j$. Thus, this is a partition and by the addition principle, $\sum_{k=0}^{m} |T_k| = |S|$.

Then for a fixed k, we claim that $|T_k| = \binom{n}{k}\binom{m}{m-k}2^{n-k}$. This comes from a 3-step process:

1. Pick the k elements of $[-n]$ which belong to A: $\binom{n}{k}$ ways
2. Pick the $m-k$ remaining elements of A from $[m]$: $\binom{m}{m-k}$ ways
3. Pick a subset from the remaining elements of $[-n]$ for B: there are $n-k$ remaining elements of $[-n]$ and any of these can be in or out of B, giving 2^{n-k} ways

By the multiplication principle, there are $\binom{n}{k}\binom{m}{m-k}2^{n-k}$ ways, and by the addition principle we finally have $\sum_{k=0}^{m}\binom{n}{k}\binom{m}{m-k}2^{n-k} = |S| = \sum_{l=0}^{n}\binom{n}{l}\binom{n+m-l}{m}$. □

If you made it through all of this, congratulations! You are well on your way to becoming a counting master. ☺