Introduction to Proof (part 4): Counting Principles (by Evan Dummit, 2019, v. 1.00)

Contents

4 Counting Principles

Our goal in this chapter is to discuss a number of varied techniques for solving various problems related to counting and enumeration in finite sets. We begin by introducing basic counting principles, which lead us to study permutations, combinations, and binomial coefficients. We then discuss inclusion-exclusion and methods for counting with repetition, before discussing a useful result known as the pigeonhole principle. We close with an analysis of a number of moderately-related counting problems that both parallel and extend our earlier examples.

4.1 Enumeration Techniques

- We begin by showing how to employ some of our results on sets and cardinality to solve various kinds of counting problems.

4.1.1 Addition and Multiplication Principles

- Two fundamental counting principles are as follows:
  - ("Addition Principle") When choosing among n disjoint options labeled 1 through n, if option i has a_i possible outcomes for each $1 \le i \le n$, then the total number of possible outcomes is $a_1 + a_2 + \cdots + a_n$.
  - To illustrate the addition principle, if a restaurant offers 5 main courses with chicken, 6 main courses with beef, and 12 vegetarian main courses, then (presuming no course is counted twice) the total possible number of main courses is $5 + 6 + 12 = 23$.

- The addition principle can be justified using our results about cardinalities of unions of sets: if A_i corresponds to the set of outcomes of option i, then the union $A_1 \cup A_2 \cup \cdots \cup A_n$ corresponds to a single choice of one outcome from one of the A_i. Then because all of the different options are disjoint, the number of such choices is $\#(A_1 \cup A_2 \cup \cdots \cup A_n) = \#A_1 + \#A_2 + \cdots + \#A_n$.
- ("Multiplication Principle") When making a sequence of n independent choices, if step i has b_i possible outcomes for each $1 \le i \le n$, then the total number of possible collections of choices is $b_1 \cdot b_2 \cdot \cdots \cdot b_n$.
- To illustrate the multiplication principle, if a fair coin is tossed (2 possible outcomes) and then a fair 6-sided die is rolled (6 possible outcomes), the total number of possible results of flipping a coin and then rolling a die is $2 \cdot 6 = 12$.
- The multiplication principle can be justified using our results about cardinalities of Cartesian products: if B_i corresponds to the set of outcomes of choice i, then the elements of the Cartesian product $B_1 \times B_2 \times \cdots \times B_n$ correspond to ordered n-tuples of outcomes, one for each choice. The number of such n-tuples is $\#(B_1 \times B_2 \times \cdots \times B_n) = \#B_1 \cdot \#B_2 \cdot \cdots \cdot \#B_n$.

- By combining these principles appropriately, we can solve a variety of basic counting problems.

- Example: Determine the number of possible outcomes from rolling a 6-sided die 5 times in a row.
  - Each individual roll has 6 possible outcomes. Thus, by the multiplication principle, the number of possible sequences of 5 rolls is $6^5 = \boxed{7776}$.

- Example: An ice creamery offers 25 different flavors. Each order of ice cream may be served in either a sugar cone, a waffle cone, or a dish, and may have 2 or 3 scoops (which must be the same flavor). Also, any order may come with a cherry or nuts (or neither), but not both. How many different orders are possible?
  - We tabulate all of the possible choices separately.
  - First, we choose an ice cream flavor: there are 25 options.
  - Then we choose a sugar cone, waffle cone, or dish: there are 3 options.
  - Next we choose the number of scoops: there are 2 options.
  - Finally, we choose either a cherry, nuts, or neither: there are 3 options.
  - By the multiplication principle, the total number of possible orders is $25 \cdot 3 \cdot 2 \cdot 3 = \boxed{450}$.

- Example: In the Unicode family of character encodings, each character is represented by a string of n bits, each of which is either a 0 or 1 (where n depends on the particular implementation). If it is necessary to be able to encode at least 150,000 different characters, what is the smallest possible value of n that will suffice?
  - If we have a string of n bits each of which is 0 or 1, then by the multiplication principle the total number of possible strings is 2^n.
  - Thus, we want $2^n \ge 150000$. Taking logarithms, we need $n \ge \log_2(150000) \approx 17.194$, so the smallest integer value of n that will work is $n = \boxed{18}$.

- Example: Determine the number of subsets of the set $\{1, 2, \ldots, n\}$.
  - We may characterize a subset S of $\{1, 2, \ldots, n\}$ by listing, for each $k \in \{1, 2, \ldots, n\}$, whether $k \in S$ or $k \notin S$.
  - By the multiplication principle, the number of possible ways of making this sequence of n choices is $\boxed{2^n}$.

- Example: If $\#A = n$ and $\#B = m$, find the total number of functions $f : A \to B$.
  - If $A = \{a_1, a_2, \ldots, a_n\}$, then such a function is completely determined by the values $f(a_1)$, $f(a_2)$, ... , $f(a_n)$.
  - Since $\#B = m$, there are m possible choices for each of the n values $f(a_1)$, $f(a_2)$, ... , $f(a_n)$.
  - Since all such choices are allowed, the total number of functions is therefore $\boxed{m^n}$.

- Example: Find the number of positive integer divisors of 90000.

- Note that $90000 = 2^4 3^2 5^4$, so any positive integer divisor must have the form $2^a 3^b 5^c$ where $a \in \{0,1,2,3,4\}$, $b \in \{0,1,2\}$, and $c \in \{0,1,2,3,4\}$.
- On the other hand, every such integer is a divisor, and so since there are 5 choices for a, 3 for b, and 5 for c, there are $5 \cdot 3 \cdot 5 = \boxed{75}$ divisors in total.
- Remark: In the same way, one may see that $n = 2^{n_2} 3^{n_3} 5^{n_5} \cdots$ has a total of $(n_2+1)(n_3+1)(n_5+1)\cdots$ positive integer divisors.

- In many counting problems, we must break into several cases and tabulate possibilities separately.
- Example: At a car dealership, Brand X sells 11 different models of cars each of which comes in 20 different colors, while Brand Y sells 6 different models of cars each of which comes in 5 different colors. How many different possible car options (including brand, model, and color) can be purchased at the dealership?
  - If a Brand X car is purchased, there are 11 choices for the model and 20 choices for the color, so by the multiplication principle there are $11 \cdot 20 = 220$ possible options in this case.
  - If a Brand Y car is purchased, there are 6 choices for the model and 5 choices for the color, so by the multiplication principle there are $6 \cdot 5 = 30$ possible options in this case.
  - Since these two cases are disjoint, by the addition principle there are $220+30 = \boxed{250}$ possible car options in total.
- In other cases, we may use "complementary counting": count possibilities and then subtract ones that are not allowed to occur, or that have been double-counted.
  - More formally, we are applying the observation that if B is a subset of A, then $\#(A \backslash B) = \#A - \#B$.
- Example: A local United States telephone number has 7 digits and cannot start with 0, 1, or the three digits 555. How many such telephone numbers are possible?
  - The first digit has 8 possibilities (namely, the digits 2 through 9 inclusive) and the other six digits each have 10 possibilities. Thus, by the multiplication principle, there are $8 \cdot 10^6 = 8\,000\,000$ total telephone numbers.
  - However, we have included the numbers starting with 555: each of these has 10 choices for each of the last 4 digits, for a total of $10^4 = 10\,000$ telephone numbers.
  - Subtracting the disallowed numbers yields a total of $8\,000\,000 - 10\,000 = \boxed{7\,990\,000}$ local telephone numbers.
  - Remark: Another method is to count all 10^7 possible 7-digit numbers, and then subtract the 10^6 starting with 0, the 10^6 starting with 1, and the 10^4 starting with 555.

4.1.2 Permutations and Combinations

- Certain problem types involving rearrangements of distinct objects ("permutations"), or ways to select subsets of a particular size ("combinations"), arise frequently in counting problems.
- Example: Determine the number of permutations (i.e., ways to rearrange) the six letters ABCDEF.
  - There are 6 letters to be arranged into 6 locations.
  - For the first letter, there are 6 choices (any of ABCDEF).
  - For the second letter, there are only 5 choices (any letter except the one we have already chosen).
  - For the third letter, there are only 4 choices (any letter except the first two).
  - Continuing in this way, we see that there are 3 choices for the fourth letter, 2 choices for the fifth letter, and only 1 choice for the last letter.
  - By the multiplication principle, the total number of permutations is therefore $6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = \boxed{720}$.
- Example: A new company logo has four design elements, which must all be different colors chosen from red, orange, yellow, green, blue, and purple. How many different logos are possible?

- There are 6 possible colors. The first design element has 6 possible colors, the second has 5 possible colors (any of the 6 except the one already used), the third has 4 possible colors, and the fourth has 3 possible colors.
- Thus, the total number of logos is $6 \cdot 5 \cdot 4 \cdot 3 = \boxed{360}$.

- Both of the problems above are examples of computing <u>permutations</u>, where we choose k distinct items from a list of n possibilities, and where the order of our choices matters.
  - We can give a general formula for solving problems of this type in terms of factorials.

- <u>Definition</u>: If n is a positive integer, we define the number $n!$ (read "n factorial") as $n! = n \cdot (n-1) \cdot \cdots \cdot 2 \cdot 1$, the product of the positive integers from 1 to n inclusive. We also set $0! = 1$.
  - Some small values are $1! = 1$, $2! = 2$, $3! = 6$, $4! = 24$, $5! = 120$, and $6! = 720$.
  - The factorial function grows very fast: to 4 significant figures, we have $10! = 3.629 \cdot 10^6$, $100! = 9.333 \cdot 10^{157}$, and $1000! = 4.024 \cdot 10^{2567}$.
  - A useful approximation known as Stirling's formula[1] says that $n! \approx n^n e^{-n} \sqrt{2\pi n}$ for large n (in the sense that the ratio between the two quantities approaches 1 as n grows). In particular, $n!$ grows faster than any exponential function of the form A^n for any positive A.
  - The estimates from Stirling's formula are $10! \approx 3.599 \cdot 10^6$, $100! \approx 9.325 \cdot 10^{157}$, and $1000! \approx 4.024 \cdot 10^{2567}$, which are quite close even for small n.

- <u>Proposition</u> (Permutations): The number of ways of choosing k ordered items from a list of n distinct possibilities (where the order of the k items matters) is equal to $\dfrac{n!}{(n-k)!} = n \cdot (n-1) \cdot \cdots \cdot (n-k+1)$. In particular, the number of ways of rearranging n distinct items is $n!$.
  - <u>Remark</u>: The number of permutations of k elements chosen from a list of n is sometimes denoted $P(n,k)$, ${}_nP_k$ or $P\binom{n}{k}$.
  - <u>Proof</u>: There are n possibilities for the first item, $n-1$ for the second item (any possibility but the one already chosen), $n-2$ for the third item (any possibility but the two already chosen), ... , and $n-k+1$ possibilities for the kth item.
  - All such selections are valid, so the total number of possibilities is $n \cdot (n-1) \cdot \cdots \cdot (n-k+1)$ by the multiplication principle.
  - For the formula, notice that $n \cdot (n-1) \cdot \cdots \cdot (n-k+1) \cdot (n-k)! = n \cdot (n-1) \cdot \cdots \cdot (n-k+1) \cdot (n-k) \cdot \cdots \cdot 1 = n!$.
  - Thus, $n \cdot (n-1) \cdot \cdots \cdot (n-k+1) = \dfrac{n!}{(n-k)!}$.

- <u>Example</u>: A sports league has 31 teams in total. How many ways are there to choose 16 teams that make the playoffs, assuming that the ranking of the playoff teams matters and there are no ties?
  - We are choosing $k = 16$ teams from a list of $n = 31$, where the order matters. From our result on permutations, the total number of choices is $\boxed{\dfrac{31!}{15!}} = 31 \cdot 30 \cdot \cdots \cdot 16$.

- <u>Example</u>: If $\#A = n$, find the total number of bijections $f : A \to A$.
  - If $A = \{a_1, a_2, \ldots, a_n\}$, a function $f : A \to A$ is completely determined by the values $f(a_1)$, $f(a_2)$, ... , $f(a_n)$.
  - Since f must be a bijection, these n values are necessarily the elements $a_1, a_2, \ldots, a_n$, possibly rearranged.

[1] For completeness we outline a proof of Stirling's formula: the natural logarithm $\ln(n!) = \ln(1) + \ln(2) + \cdots + \ln(n)$ is a Riemann sum for the integral $\int_1^n \ln(x)\,dx$, which can be evaluated via integration by parts as $n\ln(n) - n + 1$. Using the trapezoidal rule to estimate the integral yields the approximation $\frac{1}{2}\ln(1) + \ln(2) + \ln(3) \cdots + \ln(n-1) + \frac{1}{2}\ln(n) \approx n\ln(n) - n + 1$. Rearranging and solving for $n!$ yields $n! \approx n^n e^{-n} C\sqrt{n}$. By a more careful analysis of the approximation error, the constant C can be computed as $\sqrt{2\pi}$, which yields Stirling's formula.

- Hence we see that the number of such bijections is the number of ways of permuting the n elements of A, which is $\boxed{n!}$.
- Remark: This interpretation of permutations as bijections is one reason we take $0! = 1$ (since there is one bijection from the empty set to itself, namely the empty relation).

- In certain other types of counting problems, the order of the list of the k items we choose from the list of n does not matter. We can also give a formula for counting in this way:

- Proposition (Combinations): The number of ways of choosing k unordered items from a list of n distinct possibilities (where the order of the k items does not matter) is equal to $\binom{n}{k} = {}_nC_k = \dfrac{n!}{k!(n-k)!} = \dfrac{n \cdot (n-1) \cdot \dots \cdot (n-k+1)}{k \cdot (k-1) \cdot \dots \cdot 1}$.
  - Remark: The symbols $\binom{n}{k}$ and ${}_nC_k$ are both typically read as "n choose k".
  - Proof: From our calculation above, we know that the number of ways to choose k ordered items from a list of n distinct possibilities is $\dfrac{n!}{(n-k)!}$.
  - If instead we want to count unordered lists, we can simply observe that for any unordered list, there are $k!$ ways to rearrange the k elements on the list.
  - Therefore, we have counted each unordered list $k!$ times, so the number of unordered lists is $\dfrac{1}{k!} \cdot \dfrac{n!}{(n-k)!} = \dfrac{n!}{k!(n-k)!}$, as claimed.

- In general, expanding the products of factorials is not the most efficient way to evaluate $\binom{n}{k}$.
  - Instead, using the formula $\binom{n}{k} = \dfrac{n \cdot (n-1) \cdot \dots \cdot (n-k+1)}{k \cdot (k-1) \cdot \dots \cdot 1}$ is typically fastest.
  - For example, computing $\binom{13}{4}$ as $\dfrac{13!}{4!9!}$ requires computing both $13!$ and $4!9!$, and then evaluating the quotient, which is rather cumbersome.
  - However, it is quite easy to compute $\binom{13}{4}$ by writing it as $\binom{13}{4} = \dfrac{13 \cdot 12 \cdot 11 \cdot 10}{4 \cdot 3 \cdot 2 \cdot 1} = 13 \cdot 11 \cdot 5 = 715$.

- Example: How many 3-element subsets of $\{1, 2, 3, 4, 5, 6, 7, 8, 9\}$ are there?
  - Since subsets are not ordered, we are simply counting the number of ways to choose 3 unordered elements from the given set of 9.
  - From our discussion of combinations, the number of such subsets is $\binom{9}{3} = \dfrac{9 \cdot 8 \cdot 7}{3 \cdot 2 \cdot 1} = \boxed{84}$.
  - Remark: More generally, the number of k-element subsets of $\{1, 2, \dots, n\}$ is $\binom{n}{k}$.

- Example: At a conference with 30 mathematicians, every pair of attendees shakes hands once. How many total handshakes occur?
  - Since pairs of people are not ordered, we are counting the number of ways to choose 2 attendees from a total of 30, which is $\binom{30}{2} = \dfrac{30 \cdot 29}{2 \cdot 1} = \boxed{435}$.

- Example: A pizza parlor offers 13 different possible toppings on a pizza. A pizza may have from 0 up to 3 different toppings. How many different pizza topping combinations are possible?
  - In general, there are $\binom{13}{k}$ possible pizzas that have exactly k toppings.
  - Thus, the number of pizzas with at most 3 toppings is $\binom{13}{0} + \binom{13}{1} + \binom{13}{2} + \binom{13}{3} = 1 + 13 + 78 + 286 = \boxed{378}$.

- Example: Determine the number of different full-house hands, consisting of 3 cards of one rank, and a pair of cards in another rank, that can be dealt from a standard 52-card deck.

- Note that there are 13 possible card ranks (A, 2-10, J, Q, K), and 4 cards of each rank (one in each of the four suits: hearts, clubs, spades, diamonds).
- First, there are 13 ways to choose the rank of the 3-of-a-kind, and then there are 12 ways to choose the rank of the pair.
- Once we have chosen the ranks, there are $\binom{4}{3} = 4$ ways to choose the three cards forming the 3-of-a-kind, and there are $\binom{4}{2} = 6$ ways to choose the two cards forming the pair.
- Thus, in total there are $13 \cdot 12 \cdot 4 \cdot 6 = \boxed{3744}$ possible full houses.

• Example: Determine the number of possible ways of permuting the letters in the word MISSISSIPPI.

- Since there are 11 letters, it might seem as if there are 11! permutations of the letters.
- However, not all of these permutations yield different words: for example, if we swap two of the Ss, the resulting words are the same.
- There are 4 Ss, 4 Is, 2 Ps, and 1 M, which we will arrange in that order.
- First, we place the 4 Ss: since there are 11 possible locations, there are $\binom{11}{4}$ ways to place them (since the 4 Ss are identical).
- Next we place the 4 Is: there are 7 remaining locations, so there are $\binom{7}{4}$ ways to place them.
- After this, there are 3 remaining locations in which we may place the 2 Ps, yielding $\binom{3}{2}$ choices. Finally, there is only 1 location for the M.
- In total, there are $\binom{11}{4} \cdot \binom{7}{4} \cdot \binom{3}{2} = 330 \cdot 35 \cdot 3 = \boxed{34650}$ ways of permuting the letters.
- Remark: Another way to perform the count is to determine the number of times each word shows up in the 11! permutations of the letters. Since there are 4! ways of permuting the 4 Ss among themselves, 4! ways of permuting the 4 Is, and 2! ways of permuting the 2 Ps, each word shows up $4! \cdot 4! \cdot 2!$ times. Thus, the number of different words is $\dfrac{11!}{4!4!2!} = 34650$.

• Example: Determine the number of possible ways of permuting the letters in the word BOSTONIANS that contain the word BOOS.

- The number of such permutations is the number of permutations of the six letters T, N, I, A, N, S and the string BOOS (which we can think of as being a single string).
- There are 2 Ns, and 1 each of T, I, A, S, and BOOS to arrange.
- First, we place the 2 Ns: since there are 7 possible locations, there are $\binom{7}{2}$ ways to place them. The remaining 5 strings can be permuted arbitrarily, so there are 5! ways to arrange them.
- In total, there are $\binom{7}{2} \cdot 5! = 42 \cdot 120 = \boxed{2520}$ ways of permuting the letters.
- Remark: As above, another way to perform the count is by observing that there are 7! ways to arrange the 7 given strings, but each arrangement is counted twice because of the two Ns, so there are only $7!/2 = 2520$ different arrangements.

4.1.3 Binomial Coefficients and the Binomial Theorem

• The numbers $\binom{n}{k}$ are called binomial coefficients because they arise as coefficients of binomial expansions:

• Theorem (Binomial Theorem): If n is a positive integer, then for any real numbers[2] x and y, $(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^{n-k} y^k = x^n + \binom{n}{1} x^{n-1} y + \binom{n}{2} x^{n-2} y^2 + \cdots + \binom{n}{n-1} x y^{n-1} + y^n$.

- Example: For $n = 4$, $(x+y)^4 = x^4 + 4x^3y + 6x^2y^2 + 4xy^3 + y^4 = \binom{4}{0}x^4 + \binom{4}{1}x^3y + \binom{4}{2}x^2y^2 + \binom{4}{3}xy^3 + \binom{4}{4}y^4$.
- Explicitly, this result says that in the expansion of $(x+y)^n$, the coefficient of $x^k y^{n-k}$ is equal to $\binom{n}{k}$.

[2] In fact this identity holds for any x and y such that $xy = yx$. In particular, it also holds for complex numbers, or (more generally) for elements of an arbitrary field, and also when x and y are general "indeterminate variables".

- Proof: Observe that in expanding the product $(x+y)\cdot(x+y)\cdot\dots\cdot(x+y)$, we may choose an x or a y from each of n terms. The term $x^k y^{n-k}$ will arise from products that choose exactly k terms equal to x: thus, from our discussion above, there are precisely $\binom{n}{k}$ such terms.
- Remark: There are various other proofs of the binomial theorem: another standard approach is to use induction on n.

- Example: Find the coefficient of x^8 in the expansion of $(2x-1)^{10}$.
  - By the binomial theorem, we have $(2x-1)^{10} = \displaystyle\sum_{k=0}^{10}\binom{10}{k}(2x)^{10-k}(-1)^k = \sum_{k=0}^{10}\binom{10}{k}2^{10-k}(-1)^k\cdot x^{10-k}$.
  - The coefficient of x^8 corresponds to the term with $k=2$, so the coefficient is $\displaystyle\binom{10}{2}2^{10-2}(-1)^2 = \frac{10\cdot 9}{2}\cdot 2^8\cdot 1 = \boxed{11520}$.

- The binomial theorem can also be extended to negative and non-integral exponents n, as first shown by Newton.
  - In general, if α is any real number, the "generalized" binomial expansion is

$$(x+y)^\alpha = \sum_{k=0}^{\infty}\binom{\alpha}{k}x^{n-k}y^k = x^\alpha + \alpha x^{\alpha-1}y + \frac{\alpha(\alpha-1)}{2}x^{\alpha-2}y^2 + \cdots$$

    where we interpret the binomial coefficient $\binom{\alpha}{k}$ as the polynomial $\dfrac{\alpha(\alpha-1)\cdots(\alpha-k+1)}{k!}$.
  - This expansion is valid whenever α, x, y are real numbers with $|x| > |y|$. When α is a nonnegative integer, it reduces to the usual binomial theorem we stated above, but for other α, the expansion is an infinite series (the requirement $|x| > |y|$ is necessary to ensure that the series converges for arbitrary α).

- Binomial coefficients show up in many different places and possess quite a large number of algebraic identities.
  - Many identities involving binomial coefficients can be established via direct manipulation of the binomial coefficients as quotients of factorials, or using algebraic statements like the binomial theorem.
  - However, many of these identities can also be proven very easily using bijections: the basic idea is to show that both expressions are counting the same quantity. Proofs of this nature, while sometimes difficult to come up with, can often shed more light on the underlying reason for the existence of the identity than a more direct algebraic proof.

- Proposition (Binomial Coefficient Identities): If n and k are integers with n positive and $0 \le k \le n$, the following are true:

  1. Reflection Identity: $\displaystyle\binom{n}{k} = \binom{n}{n-k}$.
     - Proof 1: We have $\displaystyle\binom{n}{k} = \frac{n!}{k!(n-k)!} = \frac{n!}{(n-k)!\cdot[(n-(n-k)]!} = \binom{n}{n-k}$.
     - Proof 2: Observe that $\displaystyle\binom{n}{k}$ is the number of ways of choosing a subset S of $\{1, 2, \dots, n\}$ having k elements.
     - Since any subset is uniquely determined by its complement S^c inside $\{1, 2, \dots, n\}$, the number of ways of choosing a set S with k elements is the same as the number of ways of choosing its complement S^c, which has $n-k$ elements.
     - This means $\displaystyle\binom{n}{k} = \binom{n}{n-k}$, as claimed.
  2. Step Identity: $\displaystyle\binom{n}{k} = \frac{n}{k}\cdot\binom{n-1}{k-1} = \frac{n-k+1}{k}\cdot\binom{n}{k-1}$.

- Proof: We have $\binom{n}{k} = \frac{n!}{k!(n-k)!} = \frac{n}{k} \cdot \frac{(n-1)!}{(k-1)!(n-k)!} = \frac{n}{k} \cdot \binom{n-1}{k-1}$. The other expression follows in the same way.

3. Recurrence Relation: $\binom{n}{k} = \binom{n-1}{k} + \binom{n-1}{k-1}$.

   - Proof 1: From the step identity we have $\binom{n-1}{k} + \binom{n-1}{k-1} = \frac{n-k}{n} \cdot \binom{n}{k} + \frac{k}{n} \cdot \binom{n}{k} = \binom{n}{k}$.
   - Proof 2: Consider selecting a subset of k elements from $\{1, 2, \ldots, n\}$.
   - If the set contains 1, then there are $\binom{n-1}{k-1}$ ways to select the remaining $k-1$ elements from $\{2, \ldots, n\}$, while if the set does not contain 1, then there are $\binom{n-1}{k}$ ways to select the k elements from $\{2, \ldots, n\}$.
   - Since these two possibilities are disjoint, the total number of ways of selecting a subset of k elements from $\{1, 2, \ldots, n\}$ is equal to the sum of these two quantities, whence $\binom{n}{k} = \binom{n-1}{k} + \binom{n-1}{k-1}$.

4. Sum Identity: $\binom{n}{0} + \binom{n}{1} + \binom{n}{2} + \cdots + \binom{n}{n} = \sum_{j=0}^{n} \binom{n}{j} = 2^n$.

   - Proof 1: Set $x = y = 1$ in the binomial theorem: this yields $2^n = (1+1)^n = \sum_{j=0}^{n} \binom{n}{j}$ as required.
   - Proof 2: Consider selecting a subset of $\{1, 2, \ldots, n\}$. There are $\binom{n}{j}$ such subsets having j elements, and so summing over all possible j shows that the total number of subsets of $\{1, 2, \ldots, n\}$ is $\binom{n}{0} + \binom{n}{1} + \binom{n}{2} + \cdots + \binom{n}{n}$. But as we saw before, the total number of subsets is also 2^n, so these quantities are equal.

5. Alternating Sum Identity: $\binom{n}{0} - \binom{n}{1} + \binom{n}{2} - \cdots + (-1)^n \binom{n}{n} = \sum_{j=0}^{n} (-1)^j \binom{n}{j} = 0$.

   - Proof 1: Set $x = 1$ and $y = -1$ in the binomial theorem: this yields $0^n = (1-1)^n = \sum_{j=0}^{n} (-1)^j \binom{n}{j}$ as required.
   - Proof 2: Observe that the sum $\binom{n}{0} + \binom{n}{2} + \binom{n}{4} + \cdots$ counts the total number of subsets of $\{1, 2, \ldots, n\}$ having an even number of elements, while $\binom{n}{1} + \binom{n}{3} + \binom{n}{5} + \cdots$ counts the total number of subsets of $\{1, 2, \ldots, n\}$ having an odd number of elements.
   - Now pair each subset of $\{1, 2, \ldots, n\}$ with the subset obtained by adding or removing 1 from it (as appropriate). Each subset is paired with another subset having either 1 more or 1 fewer element, and so each pair contains one subset with an even number of elements and an odd number of elements.
   - Hence the total number of subsets with an even number of elements is equal to the total number of subsets with an odd number of elements, so $\binom{n}{0} + \binom{n}{2} + \binom{n}{4} + \cdots = \binom{n}{1} + \binom{n}{3} + \binom{n}{5} + \cdots$. Rearranging then yields the identity.

6. "Hockey-Stick" Identity: For any $n \geq k$, $\binom{k}{k} + \binom{k+1}{k} + \cdots + \binom{n}{k} = \sum_{j=k}^{n} \binom{j}{k} = \binom{n+1}{k+1}$.

   - Proof 1: Fix k and use induction on n. The base case $n = k$ is trivial, since the identity reads $\binom{k+1}{k+1} = \binom{k}{k}$, which is true since both expressions are equal to 1.

- For the inductive step, suppose that $\binom{k}{k}+\binom{k+1}{k}+\cdots+\binom{n}{k}=\binom{n+1}{k+1}$. Then we have

$$\begin{aligned}\binom{k}{k}+\binom{k+1}{k}+\cdots+\binom{n}{k}+\binom{n+1}{k} &= \left[\binom{k}{k}+\binom{k+1}{k}+\cdots+\binom{n}{k}\right]+\binom{n+1}{k} \\ &= \binom{n+1}{k+1}+\binom{n+1}{k}=\binom{n+2}{k+1}\end{aligned}$$

by the recurrence relation (3) and the inductive hypothesis. This establishes the inductive step, so the result holds for all $n \geq k$ by induction.

- Proof 2: Consider selecting a subset of $k+1$ elements from $\{1,2,\ldots,n,n+1\}$.
- If the largest element is $j+1$, for any $k \leq j \leq n$, then there will be $\binom{j}{k}$ ways to choose the remaining k elements from $\{1,2,\ldots,j-1\}$.
- Since these possibilities are disjoint (the set has a unique largest element), the total number of ways of selecting the subset is $\binom{k}{k}+\binom{k+1}{k}+\cdots+\binom{n}{k}$, which is therefore equal to $\binom{n+1}{k+1}$.

- Owing to the recurrence relation $\binom{n}{k}=\binom{n-1}{k}+\binom{n-1}{k-1}$ for the binomial coefficients, we can list them in an array called Pascal's triangle, the first few rows of which are shown below:

Row 0:						1					
Row 1:					1		1				
Row 2:				1		2		1			
Row 3:			1		3		3		1		
Row 4:		1		4		6		4		1	
Row 5:	1		5		10		10		5		1

- The array is generated as follows: for each $n \geq 0$, the binomial coefficients $\binom{n}{k}$ with $0 \leq k \leq n$ are listed.
- The outer "diagonals" consist of 1s, and every other entry in row $n+1$ is equal to the sum of the two entries in the row above it (one above and to the left, the other above and to the right).
- There are many patterns in Pascal's triangle, many of which follow from the identities we have proven above. For example, the sum of the entries in the nth row is 2^n, which is equivalent to the sum identity (4). (Try to identify why the identity (6) is called the "hockey-stick identity"!)

4.1.4 Inclusion-Exclusion

- We now discuss a counting technique that will allow us to solve problems involving a number of overlapping categories.
  - Recall that we have shown that for any finite sets A and B, we have the intersection-union formula $\#(A\cup B)=\#A+\#B-\#(A\cap B)$.
  - We can extend this formula to three sets by noting that $A\cup B\cup C=A\cup(B\cup C)$: then by de Morgan's laws and the intersection-union formula, we have

$$\begin{aligned}\#(A\cup B\cup C) &= \#A+\#(B\cup C)-\#[A\cap(B\cup C)] \\ &= \#A+[\#B+\#C-\#(B\cap C)]-\#[(A\cap B)\cup(A\cap C)] \\ &= \#A+\#B+\#C-\#(B\cap C)-\#[(A\cap B)+\#(A\cap C)-\#(A\cap B\cap A\cap C)] \\ &= \#A+\#B+\#C-\#(A\cap B)-\#(A\cap C)-\#(B\cap C)+\#(A\cap B\cap C).\end{aligned}$$

* In other words, the cardinality of $A \cup B \cup C$ is the sum of the cardinalities of A, B, and C, minus the sum of the cardinalities of the three pairwise intersections $A \cap B$, $A \cap C$, and $B \cap C$, plus the cardinality of the overall intersection $A \cap B \cap C$.
* We can view this formula as giving successive "corrections" to the count of the elements of $A \cup B \cup C$ until every part of the union is counted exactly once: the terms $\#A + \#B + \#C$ count every element in $A \cup B \cup C$. But then each element in any two A, B, C is counted twice, so we subtract out each of those counts. But now each element in $A \cap B \cap C$ is counted a total of $3 - 3 = 0$ times, so we must add $\#(A \cap B \cap C)$ to obtain the correct count.
* We may extend this result to an arbitrary finite collection of sets, although describing the actual formula itself turns out to be somewhat complicated:

- Theorem (Inclusion-Exclusion): Suppose $A_1, A_2, \dots, A_n$ are any finite sets. Then

$$\begin{aligned}
\#\left[\bigcup_{i=1}^{n} A_i\right] &= \sum_{i=1}^{n} \#A_i - \sum_{1 \le i < j \le n} \#(A_i \cap A_j) + \sum_{1 \le i < j < k \le n} \#(A_i \cap A_j \cap A_k) - + \cdots + (-1)^{n-1} \#(A_1 \cap A_2 \cdots \cap A_n) \\
&= \sum_{k=1}^{n} (-1)^{k+1} \left[\sum_{1 \le i_1 < i_2 < \cdots < i_k \le n} \#(A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_k})\right] \\
&= \sum_{S \subseteq \{1,2,\dots,n\},\, S \neq \emptyset} (-1)^{\#S+1} \# \left[\bigcap_{j \in S} A_j\right].
\end{aligned}$$

  * In words, the cardinality of the union $\bigcup_{i=1}^{n} A_i$ is obtained by summing the cardinalities of the sets A_i, subtracting the sum of the cardinalities of the pairwise intersections $A_i \cap A_j$, adding the sum of the cardinalities of the threefold intersections $A_i \cap A_j \cap A_k$, and continuing (with signs alternating) until the overall intersection $A_1 \cap \cdots \cap A_n$ is obtained.
  * Proof: We induct on n. The base case $n = 1$ is trivial.
  * For the inductive step, suppose that the formula holds for any intersection of n sets, and consider any finite sets $A_1, A_2, \dots, A_n, A_{n+1}$.
  * If we write $A_1 \cup \cdots \cup A_n \cup A_{n+1} = [A_1 \cup \cdots \cup A_n] \cup A_{n+1}$ then by the intersection-union formula we see

$$\#[A_1 \cup \cdots \cup A_n \cup A_{n+1}] = \#[A_1 \cup \cdots \cup A_n] + \#A_{n+1} - \#[A_{n+1} \cap (A_1 \cup \cdots \cup A_n)]$$

    so it suffices to find the cardinalities of $A_1 \cup \cdots \cup A_n$ and $A_{n+1} \cap (A_1 \cup \cdots \cup A_n)$.
  * For the latter, by de Morgan's laws we have $(A_1 \cup A_2 \cup \cdots \cup A_n) \cap A_{n+1} = (A_1 \cap A_{n+1}) \cup (A_2 \cap A_{n+1}) \cup \cdots \cup (A_n \cap A_{n+1})$.
  * Let $B_i = A_i \cap A_{n+1}$, we note that if $S \subseteq \{1, 2, \dots, n\}$ is any nonempty collection of indices, then it follows immediately from the properties of intersections that $\cap_{j \in S} B_j = [\cap_{j \in S} A_j] \cap A_{n+1}$.
  * Then applying the induction hypothesis to $A_1 \cup \cdots \cup A_n$ and $B_1 \cup \cdots \cup B_n$, we see

$$\begin{aligned}
\#\left[\bigcup_{i=1}^{n} A_i\right] &= \#[A_1 \cup \cdots \cup A_n] + \#A_{n+1} - \#[B_1 \cup \cdots \cup B_n] \\
&= \sum_{i=1}^{n} \#A_i - \sum_{1 \le i < j \le n} \#(A_i \cap A_j) + \sum_{1 \le i < j < k \le n} \#(A_i \cap A_j \cap A_k) - \cdots + (-1)^{n-1} \#(A_1 \cap A_2 \cdots \cap A_n) \\
&\quad + \#A_{n+1} \\
&\quad - \sum_{i=1}^{n} \#(A_i \cap A_{n+1}) + \sum_{1 \le i < j \le n} \#(A_i \cap A_j \cap A_{n+1}) - \cdots + (-1)^{n} \#(A_1 \cap A_2 \cdots \cap A_n \cap A_{n+1}) \\
&= \sum_{i=1}^{n+1} \#A_i - \sum_{1 \le i < j \le n+1} \#(A_i \cap A_j) + \cdots + (-1)^{n} \#(A_1 \cap A_2 \cdots \cap A_{n+1})
\end{aligned}$$

    which is the desired formula.

- Although the actual formula for using inclusion-exclusion is quite cumbersome to write explicitly, it is much simpler to use in practice: we need only find the cardinalities of all of the (nontrivial) intersections among the sets A_i and sum them with appropriate signs.

- Example: Find the number of positive integers n, $1 \le n \le 2019$, that are divisible by 2 or 3.
  - We use inclusion-exclusion on the integers in the range $1 \le n \le 2019$, with A_1 the multiples of 2 and A_2 the multiples of 3. Then $A_1 \cap A_2$ is the set of integers that are multiples of both 2 and 3, which is to say, integers divisible by 6.
  - Then $\#A_1 = 1009$ since the elements of A_1 are $\{2, 4, 6, \dots, 2018\}$, of which there are $2018/2 = 1009$.
  - Likewise, $\#A_2 = 673$ since the elements of A_2 are $\{3, 6, 9, \dots, 2019\}$, of which there are $2019/3 = 673$.
  - Also, $\#(A_1 \cap A_2) =$ since the elements of A_2 are $\{6, 12, 18, \dots, 2016\}$, of which there are $2016/6 = 336$.
  - Thus by inclusion-exclusion we see $\#(A_1 \cup A_2) = \#A_1 + \#A_2 - \#(A_1 \cap A_2) = 1009 + 673 - 336 = \boxed{1346}$.

- In the example above, we can find the cardinalities of the sets without listing the elements directly by using the greatest integer function (also called the floor function) $\lfloor x \rfloor$, which is defined as the greatest integer $\le x$.
  - Explicitly: in the set $\{1, 2, \dots, n\}$ there will be exactly $\lfloor \frac{n}{m} \rfloor$ elements that are divisible by m, namely, $\{m, 2m, \dots, \lfloor \frac{n}{m} \rfloor m\}$.

- Example: Find the number of positive integers n, $1 \le n \le 2019$, that are divisible by 2 or 3 or 7.
  - We use inclusion-exclusion on the integers in the range $1 \le n \le 2019$, with A_1 the multiples of 2, A_2 the multiples of 3, and A_3 the multiples of 7.
  - Then $A_1 \cap A_2$ is the multiples of 6, $A_1 \cap A_3$ is the multiples of 14, $A_2 \cap A_3$ is the multiples of 21, and $A_1 \cap A_2 \cap A_3$ is the multiples of 42.
  - Then from the discussion above and the inclusion-exclusion formula, we obtain a total count of $\lfloor \frac{2019}{2} \rfloor + \lfloor \frac{2019}{3} \rfloor + \lfloor \frac{2019}{7} \rfloor - \lfloor \frac{2019}{6} \rfloor - \lfloor \frac{2019}{14} \rfloor - \lfloor \frac{2019}{21} \rfloor + \lfloor \frac{2019}{42} \rfloor$, which equals $1009 + 673 + 288 - 336 + 144 - 96 + 48 = \boxed{1442}$.

- Example: Find the number of seven-digit strings that contain at least five consecutive equal digits.
  - Let A_1 be the strings of the form $aaaaa\cdot\cdot$, let A_2 be the strings of the form $\cdot aaaaa \cdot$, and let A_3 be the strings of the form $\cdot\cdot aaaaa$.
  - Then the desired strings are the ones in $A_1 \cup A_2 \cup A_3$, whose cardinality we may calculate using inclusion-exclusion.
  - We have $\#A_1 = \#A_2 = \#A_3 = 10 \cdot 10^2$ since a and each unspecified digit may take any value.
  - For the intersections, $A_1 \cap A_2$ consists of strings of the form $aaaaaa\cdot$, $A_2 \cap A_3$ consists of strings of the form $\cdot aaaaaa$, and $A_1 \cap A_3$ and $A_1 \cap A_2 \cap A_3$ both consist of strings of the form $aaaaaaa$.
  - Thus $\#(A_1 \cap A_2) = \#(A_2 \cap A_3) = 10 \cdot 10$ and $\#(A_1 \cap A_3) = \#(A_1 \cap A_2 \cap A_3) = 10$.
  - So by inclusion-exclusion there are $3(10 \cdot 10^2) - 2(10 \cdot 10) - 10 + 10 = \boxed{2800}$ seven-digit strings containing at least five consecutive equal digits.

- In cases where all of the k-fold intersections have the same cardinality (which often occurs when there is some symmetry among the sets) we can give a more compact inclusion-exclusion formula.
  - Explicitly, since there are $\binom{n}{k}$ k-fold intersections and they all have the same cardinality as $A_1 \cap A_2 \cap \cdots \cap A_k$, the inclusion-exclusion formula reduces to
$$\# \left[\bigcup_{i=1}^{n} A_i \right] = \sum_{k=1}^{n} (-1)^{k+1} \binom{n}{k} \#(A_1 \cap A_2 \cap \cdots \cap A_k).$$

- Example: Find the number of n-digit strings that contain at least one zero digit.

- Let A_i be the strings where the ith digit is not zero. Then the desired strings are the ones in $A_1 \cup A_2 \cup \cdots \cup A_n$, whose cardinality we may calculate using inclusion-exclusion.
- For any (nonempty) subset $S \subseteq \{1, 2, \dots, n\}$ of indices, the strings in the intersection $\bigcap_{j \in S} A_j$ are simply the strings where the jth digit is zero for each $j \in S$, and the other digits may be any value.
- Then we can see that $\# \bigcap_{j \in S} A_j$ is equal to $10^{n-\#S}$, since there is 1 digit choice for each $j \in S$ and 10 digit choices for each $j \notin S$.
- Since all of the k-fold intersections have the same cardinality 10^{n-k}, by inclusion-exclusion the total number of desired strings is $\boxed{\displaystyle\sum_{k=1}^{n} (-1)^{k+1} \binom{n}{k} 10^{n-k}}$.
- We can simplify this sum using the binomial theorem: we know that $\displaystyle\sum_{k=0}^{n} (-1)^k \binom{n}{k} 10^{n-k} = (10-1)^n = 9^n$. The sum above has an extra factor of -1 and is missing the term for $k = 0$, so its value is $\boxed{10^n - 9^n}$.
- <u>Remark</u>: Of course, there is a much easier way to find that the total number of strings is $10^n - 9^n$, which is to observe that the complement consists of strings with all nonzero digits, of which there are clearly 9^n by the multiplication principle.

4.1.5 Counting With Repetition

- We have previously discussed the problem of counting the number of unordered collections of k distinct elements chosen from a set of n. Now we examine the very similar problem of counting the number of unordered collections of k not-necessarily-distinct elements chosen from a set of n.
  - In other words, we are asking for the total number of different unordered lists of k numbers chosen with replacement from $\{1, 2, 3, \dots, n\}$.
  - For example, if $n = 2$, then we are asking for the number of ways to select k numbers from $\{1, 2\}$. It is easy to see that there are $k + 1$ possible choices: we may choose i 1s and then $k - i$ 2s, for any $0 \le i \le k$.
  - If $n = 3$, we are selecting k numbers from $\{1, 2, 3\}$. If we select j 3s, for any $0 \le j \le k$, then this leaves $k - j$ selections from $\{1, 2\}$, and there are $k - j + 1$ ways to make these selections.
  - Then the total number of selections is $\sum_{j=0}^{k} (k - j + 1) = \dfrac{(k+1)(k+2)}{2} = \binom{k+2}{2}$ as can be seen by an easy induction or an application of the hockey-stick identity.
  - We can give a similar argument to reduce the case for $n = 4$ to a sum over the cases for $n = 3$. We are selecting k numbers from $\{1, 2, 3, 4\}$, so if we select j 4s, then this leaves $k - j$ selections from $\{1, 2, 3\}$. By the above analysis, there are $\binom{k-j+2}{2}$ ways to make these choices, so the total for $n = 4$ is thus $\sum_{j=0}^{k} \binom{k-j+2}{2} = \binom{k+2}{2} + \binom{k+1}{2} + \cdots + \binom{2}{2} = \binom{k+3}{3}$, again by the hockey-stick identity.
  - We may continue this pattern (and prove it using induction) to see that the total number of different unordered lists of k numbers chosen with replacement from $\{1, 2, 3, \dots, n\}$ is equal to $\binom{n+k-1}{n-1}$.
- The inductive argument we gave above is quite natural, but there is a much cleaner way to obtain this formula directly using bijections, which we now describe:
- <u>Theorem</u> ("Stars and Bars"): The unordered lists of k numbers chosen with replacement from $\{1, 2, 3, \dots, n\}$ are in bijection with arrangements of k stars and $n - 1$ bars in a line, and therefore the number of such lists is $\dbinom{n+k-1}{n-1}$. The bijection is obtained by viewing the bars as separators that divide the line into n regions, with the number of stars in region i corresponding to the number of times i appears in the list.
  - We can illustrate the bijection with $n = 3$ and $k = 4$: the arrangement $\star\star \mid \star \mid \star$ corresponds to the list $(1, 1, 2, 3)$, the arrangement $\mid \star\star \mid \star\star$ corresponds to the list $(2, 2, 3, 3)$, the arrangement $\star\star\star \mid\mid \star$ corresponds to the list $(1, 1, 1, 3)$, and the arrangement $\star\star\star\star \mid\mid$ corresponds to the list $(1, 1, 1, 1)$.
  - <u>Proof</u>: The proof is essentially trivial, since it is clear that the map described is a well-defined bijection: every arrangement corresponds to a unique list, and every list corresponds to a unique arrangement.

- We will remark that an unordered list is often called a <u>multiset</u>, and may be thought of as a set that allows repetition of elements.
  - The number of multisets with k elements chosen from $\{1, 2, \ldots, n\}$ is referred to as a <u>multiset coefficient</u>, and is sometimes denoted $\left(\!\!\binom{n}{k}\!\!\right)$, in analogy with the binomial coefficient $\binom{n}{k}$, which counts the number of sets with k elements chosen from $\{1, 2, \ldots, n\}$.
- <u>Example</u>: Find the number of unordered lists of 6 integers, allowing repetition, from the set $\{1, 2, 3, 4, 5\}$.
  - This is a "stars and bars" calculation with $k = 6$ stars and $n = 4$ bars. The stars are the 6 integers while the regions created by the 4 bars are the 5 possible values of the integers.
  - The total number of lists is therefore $\binom{n+k-1}{n-1} = \binom{9}{3} = \boxed{84}$.
- <u>Example</u>: Find the number of ways of placing 12 identical balls into 6 distinguishable boxes.
  - This is a "stars and bars" calculation with $k = 12$ stars and $n = 5$ bars. The stars are the 12 balls while the regions created by the 5 bars represent the 6 boxes.
  - The total number of ways is therefore $\binom{n+k-1}{n-1} = \binom{17}{4} = \boxed{2380}$.
- Here are some other variations on this counting problem that can be solved using the stars-and-bars approach.
- <u>Example</u>: Find the number of possible 7-digit phone numbers in which the digits are nonincreasing (i.e., each digit is less than or equal to the previous).
  - This is a "stars and bars" calculation with 7 stars and 9 bars. The stars are the 7 digits while the regions created by the 9 bars are the 10 possible values of the digits. The total is therefore $\boxed{\binom{16}{7} = 11440}$.
  - We will remark that it is also possible to solve this problem using bijections in a slightly different way: to the 7-tuple $(a_1, a_2 \ldots, a_7)$ of digits $0 \le a_1 \le a_2 \le \cdots \le a_7 \le 9$ we associate the 7-tuple $(a_1 + 1, a_2 + 2, \ldots, a_7 + 7)$, which has the property that $1 \le a_1 + 1 < a_2 + 2 < \cdots < a_7 + 7 \le 16$.
  - Then selecting the digits $(a_1, a_2, \ldots, a_7)$ is equivalent to selecting the numbers $(a_1 + 1, a_2 + 2, \ldots, a_7 + 7)$. But since these 7 numbers are integers between 1 and 16 inclusive in increasing order, there are $\binom{16}{7}$ ways to select them.
  - It is not hard to see that this correspondence is one-to-one, and that every such choice yields a valid selection $(a_1, a_2, \ldots, a_7)$. Hence the total, as above, is $\boxed{\binom{16}{7} = 11440}$.
- <u>Example</u>: Find the number of ordered triples (a, b, c) of nonnegative integers with $a + b + c = 19$.
  - Here, we can imagine arranging 19 stars and 2 bars in a line to create 3 regions. The integer a then corresponds to the number of stars in the first region, while b is the number of stars in the second region, and c is the number of stars in the third region.
  - It is easy to see that this correspondence is one-to-one, so we have a bijection between the ordered triples (a, b, c) and the arrangements of 19 stars and 2 bars in a line, and so the number of ordered triples is $\binom{21}{2} = \boxed{210}$.
  - <u>Remark</u>: In the same way, we can see that the number of ordered n-tuples of nonnegative integers $(a_1, a_2, \ldots, a_n)$ with $a_1 + a_2 + \cdots + a_n = k$ is equal to $\binom{n+k-1}{n-1}$.
- <u>Example</u>: Find the number of ordered triples (a, b, c) of positive integers with $a + b + c = 19$.
  - We can reduce this problem to the previous one by letting $a' = a - 1$, $b' = b - 1$, and $c' = c - 1$: then the condition that a, b, c are positive is equivalent to the condition that a', b', c' are nonnegative, while the condition $a + b + c = 19$ becomes $a' + b' + c' = 16$.
  - Then we can use stars-and-bars to enumerate the triples (a', b', c') with $a' + b' + c' = 16$ as above.
  - We conclude that the total number of triples (a', b', c'), and hence the total number of triples (a, b, c), is equal to $\binom{18}{2} = \boxed{153}$.

- Remark: In the same way, we can see that the number of ordered n-tuples of positive integers $(a_1, a_2, \ldots, a_n)$ with $a_1 + a_2 + \cdots + a_n = k$ is equal to $\binom{k-1}{n-1}$.

- Example: Find the number of 8-element subsets of $\{1, 2, 3, \ldots, 20\}$ that contain no consecutive elements.
  - We can view the 8 elements of the subset as being the bars, and the remaining 12 elements as being the stars.
  - Then the condition that no elements are consecutive is the same as saying that no two bars are consecutive, which in turn is the same as saying that each of the 7 regions on the inside contains at least one star.
  - If we now remove one star from each of the 7 interior regions, we obtain a bijection of the 8-element subsets of $\{1, 2, 3, \ldots, 20\}$ having no consecutive elements with the arrangements of 8 bars and 5 stars.
  - Hence there are precisely $\binom{13}{5} = \boxed{1287}$ such subsets.
  - Remark: In the same way, we can see that the number of k-element subsets of $\{1, 2, 3, \ldots, n\}$ that contain no consecutive elements is $\binom{n-k+1}{k}$, and more generally, the number of k-element subsets of $\{1, 2, 3, \ldots, n\}$ having no two elements differ by at most a is $\binom{n-ak+1}{k}$.

- Example: A pizza chain offers 15 different toppings, and any pizza may have from 0 up to 6 toppings, including duplicates (so one option is triple pepperoni, while another is double olives and quadruple peppers). How many different topping choices are possible?
  - Suppose the pizza has k toppings. Then the number of topping choices is a "stars and bars" calculation with k stars and $n = 14$ bars, so there are $\binom{14+k}{k}$ possible topping choices.
  - Since each $k = 0, 1, \ldots, 6$ is allowed, the total number of topping choices is $\binom{14}{0} + \binom{15}{1} + \cdots + \binom{20}{6} = \boxed{54264}$.

4.2 The Pigeonhole Principle

- We now establish several related facts about cardinality and finite sets that all fall under the umbrella of the so-called "pigeonhole principle". These results are very intuitively natural, but we can give formal proofs using the language we have developed about functions and sets.

4.2.1 Statements of the Pigeonhole Principle

- Proposition (Pigeonhole Principle): Suppose $m > n$. Then there exists no one-to-one function $f : \{1, 2, \ldots, m\} \to \{1, 2, \ldots, n\}$. More generally, if A and B are finite sets and $\#A > \#B$, then there exists no one-to-one function $f : A \to B$.
  - We often phrase this more intuitively as follows: suppose we have m pigeons and we place each pigeon into one of n holes. If $m > n$, then there must be at least one hole that has more than one pigeon. (This particular formulation is the reason for the name "pigeonhole principle".)
  - Proof: For the first statement, we show the result by contradiction.
  - If f is one-to-one, then f is a bijection between $\{1, 2, \ldots, m\}$ and $\mathrm{im}(f)$, and so $\#\mathrm{im}(f) = m$.
  - But since $\mathrm{im}(f)$ is a subset of the target set $\{1, 2, \ldots, n\}$, we also have $\#\mathrm{im}(f) \leq n$, and so $m \leq n$.
  - This contradicts the assumption that $m > n$, so there cannot exist any such function f.
  - The second statement follows simply by replacing $\{1, 2, \ldots, m\}$ with the set A and $\{1, 2, \ldots, n\}$ with the set B.

- Here are some other formulations of the pigeonhole principle.

- Proposition (Pigeonhole, Set Version): If S is a finite set with $\#S = m$, and $S = S_1 \cup S_2 \cup \cdots \cup S_n$ for some $m > n$, then $\#S_i > 1$ for at least one value of i.
  - Proof: Work by contradiction: if $\#S_i \leq 1$ for all i, then $\#S = \#(S_1 \cup S_2 \cup \cdots \cup S_n) \leq \#S_1 + \#S_2 + \cdots + \#S_n \leq n$, with the latter inequality following by inclusion-exclusion or induction.

- But this is a contradiction since $m > n$. Hence $\#S_i > 1$ for at least one value of i.
- Alternatively, we could deduce this formulation from the one we gave above by writing $S_i = \{x \in S : f(x) = i\}$, and then observing that $\#S_i > 1$ for some i is equivalent to saying that $f(x_1) = i = f(x_2)$ for two unequal values $x_1, x_2 \in S$, which in turn is the same as saying that f is not one-to-one.

- Proposition (Pigeonhole, Onto Version): If A and B are finite sets and $\#A < \#B$, then there exists no onto function $g : A \to B$.
  - The intuitive explanation here is that if we have more holes than pigeons, then at least one hole must not have a pigeon in it.
  - Proof: Suppose there did exist an onto function $g : A \to B$. For each $b \in B$, let $S_b = \{x \in A : g(x) = b\}$.
  - Then the sets S_b have union A by the assumption that g is onto, so by the set version of the pigeonhole principle above, at least one set, say g_c has cardinality larger than 1.
  - But this contradicts the assumption that g is a function, because then g would not be well-defined on the element c.

- We can also strengthen the pigeonhole principle as follows:

- Proposition (Average-Value Pigeonhole): If S is a finite set with $\#S = m$, and $S = S_1 \cup S_2 \cup \cdots \cup S_n$, then $\#S_i \geq \dfrac{m}{n}$ for at least one value of i. If S is infinite and $S = S_1 \cup S_2 \cup \cdots \cup S_n$, then at least one of the S_i must also be infinite.
  - The intuitive version is that if we place m pigeons into n holes, there must be (at least) one hole that has at least the average number $\dfrac{m}{n}$ of pigeons in it.
  - Proof: If $\#S = m$ and $\#S_i < \dfrac{m}{n}$ for all i, then $\#S = \#(S_1 \cup S_2 \cup \cdots \cup S_n) \leq \#S_1 + \#S_2 + \cdots + \#S_n < n \cdot \dfrac{m}{n} = m$, which contradicts the statement $\#S = m$.
  - The infinite version follows in the same way: if all of the S_i are finite, then by definition there exists a finite number N for which $\#S_i \leq N$ (namely, the maximum of all of the cardinalities).
  - Then we would have $\#S \leq \#S_1 + \cdots + \#S_n = n \cdot N$ which is finite, contradicting the assumption that S is infinite.

- By using the idea of the pigeonhole principle's proof we can establish the following very useful result about functions on finite sets of the same cardinality:

- Proposition (Maps on Same-Cardinality Sets): Suppose A and B are finite sets with $\#A = \#B$. Then a function $f : A \to B$ is one-to-one if and only if it is onto, if and only if it is a bijection.
  - Proof: Suppose $f : A \to B$ is one-to-one and $\#A = \#B$.
  - Then f is a bijection of A with $\mathrm{im}(f)$, so $\#\mathrm{im}(f) = \#A$.
  - But since $\#B = \#A$ and B is finite, the only possibility is to have $\mathrm{im}(f) = B$. Hence f is onto, as claimed.
  - Conversely, suppose $f : A \to B$ is onto. If we take $S_b = \{a \in A : f(a) = b\}$ for each $b \in B$, then the S_b are disjoint, $A = \cup_{b \in B} S_b$, and $\#S_b \geq 1$ for each $b \in B$ (since f is onto).
  - Then we can write $\#A = \#B \leq \#S_1 + \cdots + \#S_{\#B} \leq \#A$, meaning that we must have equality everywhere. This means $\#S_b = 1$ for each $b \in B$, and so f is one-to-one.
  - Hence f is one-to-one if and only if f is onto. This means either condition is equivalent to both, which is to say, either condition is equivalent to saying f is a bijection.

4.2.2 Examples of the Pigeonhole Principle

- Here are some problems that can be solved by using pigeonhole arguments:
- Example: A sock drawer contains 10 pairs of (identical) white socks, 8 pairs of blue socks, 3 pairs of black socks, and 1 pair of purple socks. What is the least number of socks that need to be taken out (without looking at them) in order to guarantee a matching pair?
  - If we think of the holes as the sock colors and the pigeons as the different socks being removed, then the pigeonhole principle says that if we have more pigeons than holes, then at least two pigeons are in the same hole.
  - Thus, if we draw 5 socks, we are guaranteed to have a matching pair, since there are only 4 possible colors.
- Example: Show that if 25 people are sitting in a room, then at least 3 of them must share the same birth month (e.g., October).
  - If the holes are the 12 birth months and the pigeons are the 25 people, then by the average-value pigeonhole principle, at least one month has at least $\dfrac{25}{12}$ people corresponding to it.
  - Since $\dfrac{25}{12} > 2$, there must be at least 3 people sharing the same birth month.
- Example: Show that if 51 elements from the set $\{1, 2, 3, \dots, 100\}$ are chosen, then at least one pair of the elements must sum to 101.
  - Observe that there are 50 pairs of elements summing to 101 are $\{1, 100\}$, $\{2, 99\}$, $\{3, 98\}$, ... , $\{50, 51\}$.
  - Thus, if we view the holes as the 50 pairs and the pigeons as the 51 elements being selected, then at least one hole must have 2 pigeons, which is to say, both elements of the pair are chosen.
  - But this means we obtain at least one pair of elements summing to 101, as claimed.
- Example: If a is any integer and m is a modulus, show that there must exist positive integers $p < q$ such that $a^p \equiv a^q \pmod{m}$.
  - Here, we want to look at the values $\{a^1, a^2, a^3, a^4, \dots\}$ modulo m.
  - Since there are only m residue classes modulo m and there are infinitely many different powers $a^1, a^2, a^3, a^4, \dots$, by the infinite version of the pigeonhole principle, there is some residue class containing infinitely powers.
  - In particular it has at least 2 powers a^p and a^q, so that $a^p \equiv a^q \pmod{m}$.
- Example: Show that if any five lattice points in the plane (i.e., points whose coordinates are both integers) are chosen, then at least one of the line segments joining one pair of these points has a lattice midpoint.
  - Since the midpoint of (a, b) and (c, d) is $(\frac{a+c}{2}, \frac{b+d}{2})$, the midpoint is a lattice point precisely when $a + c$ and $b + d$ are both even.
  - This is the same as saying that the midpoint is a lattice point precisely when the ordered pairs of residue classes $(\overline{a}, \overline{c})$ and $(\overline{b}, \overline{d})$ modulo 2 are equal.
  - Since there are only $2 \cdot 2 = 4$ possible ordered pairs of residue classes modulo 2, then if we have 5 such ordered pairs, by the pigeonhole principle some two of them must land in the same class. Then the midpoint of that segment is a lattice point, as required.
- Example: Show that if any 51 elements from the set $\{1, 2, 3, \dots, 100\}$ are chosen, then at least one of them must divide another one.
  - The idea is to find a way of partitioning the set into subsets that are totally ordered under divisibility: then if two elements are chosen in the same subset, one of them must divide the other.
  - One way to do this is to start with an odd integer and repeatedly double it: this gives the 50 sets $\{1, 2, 4, 8, \dots, 64\}$, $\{3, 6, 12, \dots, 96\}$, $\{5, 10, 20, \dots, 80\}$, ... , $\{99\}$.

- Hence by the pigeonhole principle, if we select 51 elements from $\{1, 2, 3, \ldots, 100\}$, at least two of them must land in the same of these 50 subsets, and then one of them will divide the other, as claimed.

- Example: Assume (somewhat contrary to reality) that friendship is a symmetric relation, and also that it is irreflexive, so that no one is friends with themself. Show that in any group of people, there must be some pair that have the same number of friends.
  - If there are n people, then each person can have between 0 and $n - 1$ friends, inclusive. This does not allow for applying the pigeonhole principle, since there are n possible numbers of friends and n people.
  - However, it is not actually possible to have both a person with 0 friends and a person with $n - 1$ friends: the person with 0 friends would be friends with nobody, while the person with $n - 1$ friends would be friends with everyone else.
  - Thus, in fact, there are at most $n - 1$ possible numbers of friends for any actual collection of n people. Thus by the pigeonhole principle, there are 2 people with the same number of friends.

- Example: In a group of 6 people, each pair of people is either acquainted or strangers. Show that either there are 3 mutual acquaintances or 3 mutual strangers in the group.
  - Choose any person A and consider their relation to the 5 remaining people in the group.
  - Since each of these 5 people is either an acquaintance or a stranger to A, by the pigeonhole principle, there must be at least 3 people who fall into the same category.
  - If these 3 are all acquantances, then consider their relation to one another: if any pair are acquaintances, then this pair and A form 3 mutual acquaintances. Otherwise, all three are strangers to one another, so they form a set of 3 mutual strangers.
  - The same logic applies if all 3 are strangers: either some pair of them are strangers in which case they and A are 3 mutual strangers, or all 3 are acquainted with one another, so they form a set of 3 mutual acquaintances.
  - Thus in all cases, there are either 3 mutual acquaintances or 3 mutual strangers in the group.
  - Remark: A group of 5 people need not have 3 mutual acquaintances or mutual strangers: if the five people are arranged in a circle and each person is acquainted with the two people next to them (but not the other two) then this arrangement has no set of 3 mutual acquaintances or 3 mutual strangers.
  - Remark: This type of problem falls into the area of combinatorial graph theory called Ramsey theory, which (broadly speaking) studies how large a set must be before a particular type of structure must necessarily exist.

- One (among several) more real-word application of the pigeonhole principle as well is the following:

- Example: Show that a lossless data compression algorithm cannot guarantee compression for all input data sets.
  - Suppose that each file is represented as a string of bits, and that the compression algorithm transforms every file into an output file that has fewer bits.
  - If we let A_N be the set of all files with at most N bits (note that A_N is finite, and in fact $A_N = 2^{N+1} - 1$ if we include the empty file), then if the compression algorithm never increases the size of an input file, it is a function $f : A_N \to A_N$.
  - The statement that the compression algorithm is lossless means that the original data set can always be recovered from its output, which is simply saying that f is one-to-one.
  - But now by our result on same-cardinality sets, this means that $f : A_N \to A_N$ is one-to-one, hence it is a bijection. Since this holds for every N, by an easy induction this means that f must map the files with exactly N bits to themselves, meaning that f cannot actually compress any file.
  - Remark: Another way of phrasing this result is that if a lossless data compression algorithm shortens any one file, then it must lengthen another one.

4.3 Other Examples of Counting Problems

- Using all of the techniques we have developed so far, we can solve a wide array of basic counting problems, of which we now discuss various examples.

4.3.1 Prime Powers Dividing Factorials

- Example: Find the largest power of 2 that divides 2019!.
  - An initial guess might be that since there are 1009 even terms in 2019! (namely, $2, 4, \dots, 2018$) that the power of 2 in the prime factorization of 2019! would be 2^{1009}.
  - However, each of the 504 terms $4, 8, \dots, 2016$ (namely, the terms divisible by 4) will actually contribute two factors of 2, so we must also count the additional 2^{504} that are contributed by these terms.
  - In a similar way we must also add the extra powers of 2 arising from the terms divisible by 8, 16, 32, ... , and 1024.
  - Because in general there are exactly $\lfloor \frac{2019}{k} \rfloor$ terms among $\{1, 2, \dots, 2019\}$ divisible by k, we can see that the total number of factors of 2 in 2019! is $\sum_{k=1}^{10} \lfloor \frac{2019}{2^k} \rfloor = \lfloor \frac{2019}{2} \rfloor + \lfloor \frac{2019}{4} \rfloor + \lfloor \frac{2019}{8} \rfloor + \cdots + \lfloor \frac{2019}{1024} \rfloor$.
  - Explicitly, this sum is $1009 + 504 + 252 + 126 + 63 + 31 + 15 + 7 + 3 + 1 = 2011$. (We will note in passing that each term is half of the previous one, rounded down.)
  - Hence we conclude that the largest power of 2 that divides 2019! is $\boxed{2^{2011}}$.
- By the same argument we can compute the largest power of p that divides $n!$:
- Proposition (Prime Powers in Factorials): If p is a prime, then the exponent of the largest power of p dividing $n!$ is $\sum_{k=1}^{\infty} \left\lfloor \frac{n}{p^k} \right\rfloor$.
  - Note that for $k > \log_p n$ the terms are 0, so the sum is actually finite for all n.
  - Proof: There are $\lfloor \frac{n}{p^k} \rfloor$ terms in $n!$ that are divisible by p^k.
  - Each term divisible by p contributes one factor of p, each term divisible by p^2 then contributes an additional factor of p, and so forth.
  - Hence the total number of factors of p is the sum $\sum_{k=1}^{\infty} \left\lfloor \frac{n}{p^k} \right\rfloor$ as claimed.
- There are also other several other formulas for this sum.
  - To start observe that the sum $\sum_{k=1}^{\infty} \frac{n}{p^k}$ without the floor function is a geometric series with sum $\frac{n}{p-1}$.
  - If we write $n = a_d p^d + a_{d-1} p^{d-1} + \cdots + a_1 p + a_0$ in base p, it is not hard to calculate that the difference between the geometric series and the original sum is equal to $\frac{a_0 + a_1 + \cdots + a_d}{p-1}$.
  - This yields a formula $\sum_{k=1}^{\infty} \left\lfloor \frac{n}{p^k} \right\rfloor = \frac{n - s_p(n)}{p-1}$ where $s_p(n)$ is the sum of the digits of n when written in base p.
  - These two formulas for the power of p dividing $n!$ are sometimes called de Polignac's formulas or Legendre's formulas.
- Example: Find the number of zeroes at the end of 2019! when it is written in base 10.
  - We count the number of factors of 2 and 5 appearing in 2019!.

- By the analysis above, the exponent of the greatest power of p dividing $n!$ is $\sum_{k=1}^{\infty} \lfloor \frac{n}{p^k} \rfloor$.
- The number of factors of 2 was computed above as 2011.
- The number of factors of 5 is $\lfloor \frac{2019}{5} \rfloor + \lfloor \frac{2019}{25} \rfloor + \lfloor \frac{2019}{125} \rfloor + \lfloor \frac{2019}{625} \rfloor = 403 + 80 + 16 + 3 = 502$.
- Since there are more 2s than 5s (which we could also have observed without actually counting them explicitly), the largest power of 10 is 10^{502}, so there are $\boxed{502}$ zeroes at the end of 2019!.

- We will also mention in passing that since we can compute the power of a prime p dividing any factorial, we can also apply it to find the power of p dividing a binomial coefficient $\binom{n}{k} = \frac{n!}{k!(n-k)!}$.

4.3.2 Derangements

- We now discuss the problem of counting permutations of a set in which no value is sent to its original place.
- Example (Derangements): Find the number of permutations of the set $\{1, 2, \dots, n\}$ in which no number appears in its original place. (Such permutations are called derangements.)
  - Remark: This problem is also known as the "hat check problem" since it can also be interpreted as follows: if n patrons check their hats at an establishment with a clerk who returns them in random order, in how many ways can none of the patrons receive their own hat back?
  - We use inclusion-exclusion to find this number D_n. If we let A_i be the set of permutations in which the number i appears in its original place, then the desired set is the complement of $A_1 \cup A_2 \cdots \cup A_n$.
  - For any (nonempty) subset $S \subseteq \{1, 2, \dots, n\}$ of indices, the permutations in the intersection $\bigcap_{j \in S} A_j$ are those in which the jth number of the permutation is j for each $j \in S$, and the remaining elements may have any value.
  - Thus, we see that $\# \bigcap_{j \in S} A_j$ is equal to $(n - \#S)!$, since the $\#S$ elements in S must be sent to themselves while the $n - \#S$ remaining elements may be permuted arbitrarily.
  - Since all of the k-fold intersections have the same cardinality $(n-k)!$, by inclusion-exclusion the cardinality of the union $A_1 \cup A_2 \cdots \cup A_n$ is $\sum_{k=1}^{n} (-1)^{k+1} \binom{n}{k} (n-k)!$.
  - Since $\binom{n}{k} = \frac{n!}{k!(n-k)!}$ we can simplify the sum as $\sum_{k=1}^{n} (-1)^{k+1} \frac{n!}{k!} = n! \cdot \left[\sum_{k=1}^{n} \frac{(-1)^{k+1}}{k!} \right]$.
  - Then the set of derangements has cardinality $n! - \#(A_1 \cup A_2 \cup \cdots \cup A_n)$, which can be simplified into the fairly elegant form $D_n = \boxed{n! \cdot \sum_{k=0}^{n} \frac{(-1)^k}{k!} = n! \cdot \left[\frac{1}{0!} - \frac{1}{1!} + \frac{1}{2!} - \frac{1}{3!} + - \cdots + (-1)^n \frac{1}{n!} \right]}$.

- Example: The number of derangements on $\{1, 2, 3, 4\}$ is $D_4 = 4! \cdot \left[\frac{1}{1} - \frac{1}{1} + \frac{1}{2} - \frac{1}{6} + \frac{1}{24} \right] = 9$.

- Using this formula, we can obtain a pleasantly simple way of computing the exact number of derangements even for large n.
  - Specifically, evaluating the Taylor series for e^x at $x = 1$ yields $e^{-1} = \frac{1}{0!} - \frac{1}{1!} + \cdots + (-1)^n \frac{1}{n!} + (-1)^{n+1} \frac{1}{(n+1)!} + \cdots$, so multiplying by $n!$ yields $\frac{n!}{e} = D_n + (-1)^{n+1} \frac{n!}{(n+1)!} + \cdots$.
  - Since the terms in the tail of the series alternate in sign and the first one is $\frac{n!}{(n+1)!} = \frac{1}{n+1}$, by standard properties of alternating series, the difference between the infinite series and the partial sum $n! \cdot \left[\frac{1}{0!} - \frac{1}{1!} + \frac{1}{2!} - \frac{1}{3!} + \cdots + (-1)^n \frac{1}{n!} \right]$ is less than the absolute value of the first term, which is $\frac{1}{n+1}$.
  - Thus, for $n \geq 2$, D_n is the nearest integer to $\frac{n!}{e}$. For example, since $\frac{8!}{e} = 14832.899\ldots$, we see that $D_8 = 14833$, as can be confirmed by using the formula.

4.3.3 Counting Equivalence Relations

- Next we analyze the problem of counting the number of equivalence relations on a set of cardinality n.
- Proposition: Let B_n be the number of distinct equivalence relations on a set with cardinality n. Then $B_{n+1} = \sum_{k=0}^{n} \binom{n}{k} B_k$.
    - Proof: As we have shown previously, an equivalence relation on a set is the same as a partition of the set.
    - It is clear that relabeling the elements of a set does not affect the partitions (we have an obvious bijection between them) so suppose the set of $n+1$ elements is $\{1, 2, \dots, n+1\}$.
    - Now imagine removing the subset S in the partition that contains $n+1$, and suppose $\#S = k+1$ for some $k \geq 0$.
    - Removing $S = S' \cup \{n+1\}$ from the partition of $\{1, 2, \dots, n+1\}$ will yield a partition of a subset of $\{1, 2, \dots, n\}$ consisting of the $n-k$ elements that remain after removing S' from $\{1, 2, \dots, n\}$. By definition, there are B_{n-k} such partitions.
    - Therefore, by summing over all possible subsets S' of $\{1, 2, \dots, n\}$, we see that $B_{n+1} = \sum_{S' \subseteq \{1,2,\dots,n\}} B_{n-\#S'}$.
    - Since there are $\binom{n}{n-k}$ possible subsets S' with $\#S' = k$, grouping the subsets together by cardinality yields the required formula $B_{n+1} = \sum_{k=0}^{n} \binom{n}{n-k} B_{n-k} = \sum_{k=0}^{n} \binom{n}{k} B_k$ upon reversing the summation.
- These numbers B_n from the proposition above are sometimes called the Bell numbers after the mathematician E.T. Bell, although (like many other named combinatorial objects) he was not the first to study them.
    - Since $B_0 = B_1 = 1$, we can use the recurrence relation to calculate the next few values of B_n:

n	0	1	2	3	4	5	6	7	8	9	10	11	12
B_n	1	1	2	5	15	52	203	877	4140	21147	115975	678570	4213597

    - Various other places the Bell numbers show up include as the number of possible rhyme schemes for an n-line poem, as the nth moment of the Poisson distribution with mean 1, and in the coefficient of x^n in the Taylor series at $x = 0$ for[3] the function $e^{e^x-1} = \sum_{n=0}^{\infty} \frac{B_n}{n!} x^n$.

4.3.4 The Multinomial Theorem

- We can generalize the binomial theorem to a "multinomial theorem" involving terms with more than 2 summands.
- Proposition (Multinomials): If n is a positive integer and $n_1, \dots, n_k$ are nonnegative integers with $n_1 + \cdots n_k = n$, then the coefficient of the term $a_1^{n_1} a_2^{n_2} \cdots a_k^{n_k}$ in the expansion of $(a_1 + a_2 + \cdots + a_k)^n$ is equal to the multinomial coefficient $\binom{n}{n_1, n_2, \dots, n_k} = \frac{n!}{n_1! n_2! \cdots n_k!}$. Then we obtain a multinomial expansion

$$(a_1 + a_2 + \cdots + a_k)^n = \sum_{n_1 + n_2 + \cdots + n_k = n} \binom{n}{n_1, n_2, \dots, n_k} a_1^{n_1} a_2^{n_2} \cdots a_k^{n_k}.$$

[3] In fact this can be proven using the recurrence relation we derived above, by observing that the derivative of $\sum_{n=0}^{\infty} \frac{B_n}{n!} x^n$ is $\sum_{n=0}^{\infty} \frac{B_{n+1}}{n!} x^n = \sum_{n=0}^{\infty} \sum_{k=0}^{n} \frac{B_k}{k!(n-k)!} x^n$ and observing that this is also the product of $e^x - 1 = \sum_{k=1}^{\infty} \frac{x^n}{n!}$ with $\sum_{n=0}^{\infty} \frac{B_n}{n!} x^n$ after rearrangement. Thus, both $\sum_{n=0}^{\infty} \frac{B_n}{n!} x^n$ and e^{e^x-1} satisfy the differential equation $y' = (e^x - 1)y$, so since they also satisfy the initial condition $y(0) = 1$, by the uniqueness of solutions to initial-value problems we get the claimed equality.

- Example: We have $(a+b+c)^3 = a^3+b^3+c^3+3(a^2b+ab^2+a^2c+ac^2+b^2c+bc^2)+6abc$, and indeed the coefficient of abc is $\binom{3}{1,1,1} = \dfrac{3!}{1!1!1!} = 6$ while the coefficient of ac^2 is $\binom{3}{1,0,2} = \dfrac{3!}{1!0!2!} = 3$.
- Proof: Imagine expanding the product $(a_1+a_2+\cdots+a_k)^n = (a_1+a_2+\cdots+a_k)\cdot\cdots\cdot(a_1+a_2+\cdots+a_k)$ using the distributive law.
- Each of the n terms contributes a single factor a_i to the product. So the number of times the term $a_1^{n_1}a_2^{n_2}\cdots a_k^{n_k}$ appears is the number of ways of selecting n_1 terms a_1, n_2 terms a_2, ... , and n_k terms a_k.
- Thus we wish to enumerate the total number of strings that contain n_1 terms a_1, n_2 terms a_2, ... , and n_k terms a_k.
- From our discussion of permutations of strings of letters, if we imagine all of the terms a_i as having a different color, then we would have $n!$ possible permutations, so there are $n!$ possible colored strings.
- Since each uncolored string has a total of $n_1!n_2!\cdots n_k!$ possible colorings, the total number of uncolored strings is $\dfrac{n!}{n_1!n_2!\cdots n_k!}$, as claimed.
- Remark: It is also possible to prove the multinomial theorem by induction on the number of variables k, using the binomial theorem in the inductive step.

- Example: Find the coefficient of $x^{20}y^{19}$ in $(x+2y+3)^{45}$.
  - From the multinomial theorem we see that the term with $x^{20}y^{19}$ in $(x+2y+3)^{45}$ will have $n_1 = 20$, $n_2 = 19$, hence $n_3 = 45-20-19 = 6$.
  - This term is then $\binom{45}{20,19,6}x^{20}(2y)^{19}3^6 = \dfrac{45!}{20!19!6!}2^{19}3^6 \cdot x^{20}y^{19}$, so the coefficient is $\boxed{\dfrac{45!}{20!19!6!}2^{19}3^6}$.
- Here are a few other properties of the multinomial coefficients and the multinomial expansion:
- Proposition (Multinomial Properties): Suppose $n_1,\dots,n_k$ are nonnegative integers with $n_1+\cdots+n_k = n$.
  1. The multinomial coefficient $\binom{n}{n_1,\dots,n_k} = \binom{n}{m_1,\dots,m_k}$ for any permutation $m_1,\dots m_k$ of $n_1,\dots,n_k$.
     - Proof: This follows immediately from $\binom{n}{n_1,\dots,n_k} = \dfrac{n!}{n_1!n_2!\cdots n_k!}$, since the denominators of both terms are the same, merely rearranged.
  2. The multinomial coefficient $\binom{n}{n_1,\dots,n_k} = \binom{n}{n_1}\binom{n-n_1}{n_2}\binom{n-n_1-n_2}{n_3}\cdots\binom{n-n_1-\cdots-n_{k-1}}{n_k}$.
     - Proof: This follows from the interpretation of the multinomial coefficient as the number of letter strings containing n_1 terms a_1, n_2 terms a_2, ... , and n_k terms a_k.
     - Equivalently, we could select the letters one type at a time: there are $\binom{n}{n_1}$ ways to choose the locations of the a_1s, then $\binom{n-n_1}{n_2}$ ways to choose the locations of the a_2s, and so forth, up to a total of $\binom{n-n_1-\cdots-n_{k-1}}{n_k}$ ways to select the a_ks.
     - Since the product counts the total number of strings, it equals the multinomial coefficient.
  3. The sum $\displaystyle\sum_{n_1+n_2+\cdots+n_k=n}\binom{n}{n_1,\dots,n_k}$ of all the multinomial coefficients is k^n.
     - Proof: Set $a_1 = a_2 = \cdots = a_k = 1$ in the multinomial theorem.
  4. The total number of different monomials appearing in the expansion of $(a_1+a_2+\cdots+a_k)^n$ is $\binom{n+k-1}{n}$.
     - Proof: This is a "stars and bars" calculation with n stars, representing the n terms in a product, and $k-1$ bars, corresponding to separators between the k possible subscripts on an a_i.
  5. The total number of ways of arranging n distinct objects into k distinct boxes, where n_i objects go into the ith box, is $\binom{n}{n_1,\dots,n_k}$.
     - Proof: This follows immediately from our interpretation of $\binom{n}{n_1,\dots,n_k}$ as the number of letter strings with n_i terms labeled i, obtained by listing the box number in which the objects are placed.

4.3.5 Stirling Numbers

- Next we study the so-called "Stirling numbers of the second kind", which count the number of equivalence relations on the set$\{1, 2, \ldots, n\}$ having exactly k equivalence classes.
  - Equivalently, these numbers count the number of ways of partitioning the set$\{1, 2, \ldots, n\}$ into exactly k unordered subsets.

- Proposition (Stirling Numbers): If $\left\{ n \atop k \right\}$ denotes the number of equivalence relations on the set$\{1, 2, \ldots, n\}$ having exactly k equivalence classes, then $\left\{ n \atop k \right\} = k \left\{ n-1 \atop k \right\} + \left\{ n-1 \atop k-1 \right\}$.
  - Remark: These numbers are called Stirling numbers of the second kind.
  - Proof: Suppose we have a partition of $\{1, 2, \ldots, n\}$ into exactly k subsets, and consider the subset containing n.
  - If n is a singleton, meaning that the set containing n is just $\{n\}$, then deleting $\{n\}$ from the partition yields a partition of $\{1, 2, \ldots, n-1\}$ into exactly $k-1$ subsets, and there are exactly $\left\{ n-1 \atop k-1 \right\}$ such partitions.
  - Otherwise, if n is not a singleton, then consider the partition obtained by deleting n: this yields a partition of $\{1, 2, \ldots, n-1\}$ into exactly k subsets, and there are exactly $\left\{ n-1 \atop k \right\}$ such partitions.
  - Then we may adjoin n to any one of these k subsets to produce a partition $\{1, 2, \ldots, n\}$ into exactly k subsets.
  - All of these cases are disjoint, so we obtain $\left\{ n \atop k \right\} = k \left\{ n-1 \atop k \right\} + \left\{ n-1 \atop k-1 \right\}$ as claimed.

- By using the easily calculated values $\left\{ 0 \atop 0 \right\} = 1$ and $\left\{ n \atop 0 \right\} = \left\{ 0 \atop n \right\} = 0$ for $n > 0$, we can calculate the values $\left\{ n \atop k \right\}$ recursively. Below is a small table of Stirling number values:

$n \backslash k$	0	1	2	3	4	5	6	7	8
0	1								
1	0	1							
2	0	1	1						
3	0	1	3	1					
4	0	1	7	10	1				
5	0	1	15	65	15	1			
6	0	1	31	350	140	21	1		
7	0	1	63	1701	1050	266	28	1	
8	0	1	127	7770	6951	2646	462	36	1

- The Stirling numbers of the second kind also obey a number of other properties and identities, much as the binomial coefficients do. Some of these are as follows:

- Proposition (Stirling Number Identities): If n and k are nonnegative integers, then the following hold:

  1. The number of onto functions from $\{1, 2, \ldots, n\}$ to $\{1, 2, \ldots, k\}$ is equal to $k! \left\{ n \atop k \right\}$.
     - Proof: If $f : \{1, 2, \ldots, n\} \to \{1, 2, \ldots, k\}$ is onto, consider the k sets $S_i = f^{-1}(\{i\}) = \{a \in \{1, 2, \ldots, n\} : f(a) = i\}$ for $i \in \{1, 2, \ldots, k\}$.
     - These sets give a partition of $\{1, 2, \ldots, n\}$ into k nonempty sets (each set is nonempty because f is onto).

- Conversely, any such list of nonempty subsets $S_1, S_2, \dots, S_k$ giving a partition of $\{1, 2, \dots, n\}$ yields a unique onto function $f : \{1, 2, \dots, n\} \to \{1, 2, \dots, k\}$.
- The number of such lists is $k!$ times the number of unordered partitions $\left\{ {n \atop k} \right\}$ of $\{1, 2, \dots, n\}$, so the number of onto functions from $\{1, 2, \dots, n\}$ to $\{1, 2, \dots, k\}$ is equal to $k! \left\{ {n \atop k} \right\}$, as claimed.

2. For any n, the nth Bell number is given by $B_n = \displaystyle\sum_{k=0}^{n} \left\{ {n \atop k} \right\}$.
    - Proof: For positive n, observe that any partition of $\{1, 2, \dots, n\}$ can have between 0 and n parts inclusive, and these cases are disjoint.
    - Since the total number of partitions of $\{1, 2, \dots, n\}$ is the Bell number B_n and the number of partitions into k parts is $\left\{ {n \atop k} \right\}$, the result follows immediately.
3. The Stirling number $\left\{ {n \atop k} \right\} = \dfrac{1}{k!} \displaystyle\sum_{j=0}^{k} \binom{k}{j} (-1)^{k-j} j^n$.
    - Proof: Using (1) it is sufficient to show that the total number of onto functions from $\{1, 2, \dots, n\}$ to $\{1, 2, \dots, k\}$ is equal to $\dfrac{1}{k!} \displaystyle\sum_{j=0}^{k} \binom{k}{j} (-1)^{k-j} j^n$.
    - To do this we first observe that the number of functions with image contained in any subset of $\{1, 2, \dots, k\}$ of cardinality a is simply n^a, and there are $\binom{n}{a}$ possible images of cardinality a.
    - We can then apply inclusion-exclusion, recursively, to compute the number of functions whose image has cardinality 1, 2, 3, ... , up to k (we omit the precise details).
    - The last calculation yields that there are $\displaystyle\sum_{j=0}^{k} \binom{k}{j} (-1)^{k-j} j^n$ onto functions from $\{1, 2, \dots, n\}$ to $\{1, 2, \dots, k\}$, as claimed.
4. For any n, k, we have $\left\{ {n+1 \atop k+1} \right\} = \displaystyle\sum_{j=k}^{n} \binom{n}{j} \left\{ {j \atop k} \right\} = \sum_{j=k}^{n} (k+1)^{n-j} \left\{ {j \atop k} \right\}$.
    - These results are both (in some sense) a version of the hockey-stick identity for the Stirling numbers.
    - Proof: Suppose that a partition of $\{1, 2, \dots, n+1\}$ with $k+1$ subsets is given.
    - For the first identity, consider deleting the subset containing $n+1$ and suppose it contains $n-j+1$ elements for some $k \le j \le n$.
    - Then there are $\binom{n}{n-j}$ ways to choose the remaining elements of S, and deleting S yields a partition of a subset of $\{1, 2, \dots, n\}$ having exactly j elements, of which there are $\left\{ {j \atop k} \right\}$ total possibilities.
    - In total, there are $\binom{n}{j} \left\{ {j \atop k} \right\}$ ways. Summing over all j yields the first identity.
    - For the second identity, write all of the elements in each subset in increasing order, and suppose that $j+1$ is the smallest element appearing at the beginning of any subset for some j with $k \le j \le n$.
    - If we imagine deleting all of the elements larger than j, we obtain a partition of $\{1, 2, \dots, j\}$ into exactly k subsets (each part of the original partition survives, except for the one that starts with $j+1$), and there are $\left\{ {j \atop k} \right\}$ such partitions.
    - If we now fill the remaining elements back into the original partition, then $j+1$ is added to a new part, and each the remaining $n-j$ elements can then be placed arbitrarily in any of the $k+1$ parts.
    - In total, there are $(k+1)^{n-j} \left\{ {j \atop k} \right\}$ ways. Summing over all j yields the second identity.

Well, you're at the end of my handout. Hope it was helpful.