Contents

1 Combinatorics **2**
- 1.1 Permutations . . . 2
  - 1.1.1 r-permutations . . . 3
- 1.2 Combinations . . . 3
- 1.3 Inclusion-Exclusion Principle . . . 4
- 1.4 Probability . . . 5
- 1.5 Conditional Probability . . . 7
- 1.6 k-nomial Theorem . . . 7

2 Propositional Logic **8**

3 Set Theory **9**

4 Big Oh-Notation **10**

5 Proofs **10**

1 Combinatorics

Suppose we have some task, such as making a sandwich. We begin with choosing a type of bread, which we have two options for. After that, we have to choose one type of condiment, of which there are 3 types. After choosing a condiment, we have 3 types of greens that we want to add to the sandwich. We can create a tree of these choices, and we see that the more options we add on, the larger the tree becomes. This provides us with the fundamental rule of combinatorics, the multiplication rule.

If E is some experiment that is conducted through k sequential steps, where every step s_i can be conducted in n_i different ways, the total number of ways that E can be conducted is

$$\prod_{i=1}^{k} n_i = n_1 \cdot n_2 \cdot \cdots \cdot n_k$$

Another example of this is the number of subsets of a set of n elements. For each element, there are two choices, either the element is in the subset or it is not in the subset. We can see that for a set of n elements, we have 2^n different subsets.

Let's look at another example. If we have to pick 3 projects total for a course, and we need to pick one from each category. If the first category has 20 projects, the second has 15, and the third has 40, by the multiplication rule, we have $20 \times 15 \times 40 = 12000$ different ways for him to pick his 3 projects. Note that if we didn't constrain him to picking one from every category, it would be different. Also note that we have an implicit constraint that no project is part of more than one category.

If instead he had to pick only 1 project, we can see that it's easy, he has $20 + 15 + 40 = 75$ different choices. This is the addition (sum) rule. Suppose that we have a goal G that can be reacher when any given one of the experiments E_i succeeds. If every E_i can be attained in $|E_i|$ ways, then the total number of ways in which G can happen is

$$\sum_{i=1}^{k} |E_i| = |E_1| + |E_2| + \cdots + |E_k|$$

If we want to generate bitstrings of length 6, and want to find how many of them take the form $1\,1\,x\,x\,1\,1$, we have 4, because in the middle we only have $2^2 = 4$ different options to fill the gap. Another question is how many begin with $1\,1$ or end with $1\,1$. In this case, both situations contain 2^4 strings. This tempts us to add them, but this overcounts the cases where both are true. We are effectively trying to compute the union of these two sets, and thus we have to remove the cases that are in the intersection of those two sets.

1.1 Permutations

To generate a permutation of a string, we simply swap the order of 0 or more characters of the string with each other. How many permutations of a string such as "machinery" are there? We can look at the new first character. We have 9 options for this slot, and for the second character we have only 8 options, because we have now used up one of the characters. Following this pattern, we are then given that there are 9! permutations of the string. Thus, a string of length n has $n!$ permutations. Note that the string "machinery" has no repeated characters.

How would we deal with a word like "puzzle"? The answer is definitely not 6!, but how do we get rid of the overcount? We have multiple ways to generate the same permutation, and to get rid of them, we can divide them out:

$$\frac{6!}{2!}$$

If he have a string like "scissor":

$$\frac{7!}{3!}$$

For "mississippi":

$$\frac{7!}{4!4!2!}$$

For "bookkeeper":

$$\frac{10!}{2!2!3!}$$

To generalize this, the total number of non-equivalent permutations of a string σ of letters of length n where there are n_a a's, n_b b's, all the way until n_z z's:

$$\frac{n!}{n_a! \cdot n_b! \cdot \ldots \cdot n_z!}$$

1.1.1 r-permutations

Permutations are best presented with strings, but r-permutations are best presented by sets. Here's an example. If we have 10 people, and we need to pick 3 people for a picture, and the order of the people matters. We can see that the number of ways to do this is $10 \times 9 \times 8$. We can write this:

$$\frac{10!}{(10-3)!}$$

If we have 8 books on a bookshelf, and we want to borrow 5 of them in a specific order, we have:

$$\frac{8!}{(8-5)!} = \frac{8!}{3!}$$

Generalizing this, if we have n options and we want to choose r of them in a specific order:

$$P(n,r) = \frac{n!}{(n-r)!}$$

1.2 Combinations

Earlier we talked about cases where the order of the people mattered. If instead of taking a photo of 3 people, if we have to form a PhD defense commitee of 3 people, does the order matter? We can see that it doesn't. We know that the answer can't be $P(10,3)$, but we know that it has to be less, as different cases that would be considered distinct would be collapsed into one case. We know that specifically, we can normalize this by dividing by 3!:

$$\frac{P(10,3)}{3!} = \frac{10!}{7!3!}$$

We are essentially asking how many subsets of 3 people can we retrieve out of a set of 10 people. This gives us the n choose r notation:

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

1.3 Inclusion-Exclusion Principle

The inclusion/exclusion principle generalizes the law of addition/subtraction. It allows us to calculate the cardinalities of unions:

$$A_1 \cup A_2 \cup \dots A_n$$

We have seen via example:

$$|A \cup B| = |A| + |B| - |A \cap B|$$

But we can generalize this principle to larger unions:

$$|A \cup B \cup C|$$

Let's do an example. If we have a password that needs to be between 4 to 6 letters long, with letters, digits, and 7 special characters (69 characters total). How many different passwords can we store? We can see that we have

$$69^4 + 69^5 + 69^6$$

different password possibilities. Here we see that there is no overlap between the set of 4-character passwords and 5-character passwords, likewise for 6-character passwords. Suppose we disallow reusing characters. We can still sum up our sets M_4, M_5, and M_6:

$$P(69, 4) + P(69, 5) + P(69, 6)$$

We see that the addition rule that we have been using, is just a specific case of the inclusion/exclusion rule, where the sets are finite and pairwise disjoint:

$$|A_1 \cup A_2 \cup \dots \cup A_n| = \sum_{i=1}^{n} |A_i|$$

Let's make this a little more complex. Alice likes passwords of length 6 that start with 'A', and Bob likes passwords of length 6 that end with 'B'. They use completely unique characters (no repeats). What is the total number of passwords that either Alice or Bob use.

Let's first start with the set of passwords that Alice can use, $|P_A|$, which is $P(68, 5)$. Similarly, $|P_B| = P(68, 5)$. We want to find the union of these two sets:

$$|P_A \cup P_B| = |P_A| + |P_B| - |P_A \cap P_B|$$

We now need to find $|P_A \cap P_B|$, which is $P(67, 4)$:

$$P(68, 5) + P(68, 5) - P(67, 4)$$

Let's do a number theory problem. How many numbers between 1 and 1000 are divisible by either 2 or 3?

We can take the set of even numbers and union it with the set of 3 divisble numbers, and subtract the set of numbers divisible by both:

$$\left\lfloor \frac{1000}{2} \right\rfloor + \left\lfloor \frac{1000}{3} \right\rfloor - \left\lfloor \frac{1000}{6} \right\rfloor = 667$$

If we now move to 3 sets:

$$|P \cup J \cup C| = |P| + |J| + |C| - (|P \cap J| + |P \cap C| + |J \cap C|) + (P \cap C \cap J)$$

If we want to find how many numbers between 1 and 1000 are divisible by 2,3, or 5, we can use the inclusion/exclusion principle for 3 variables:

$$\left\lfloor \frac{1000}{2} \right\rfloor + \left\lfloor \frac{1000}{3} \right\rfloor + \left\lfloor \frac{1000}{5} \right\rfloor - \left(\left\lfloor \frac{1000}{6} \right\rfloor + \left\lfloor \frac{1000}{10} \right\rfloor + \left\lfloor \frac{1000}{15} \right\rfloor \right) + \left\lfloor \frac{1000}{30} \right\rfloor$$

How can we do this for 4 sets? If we do it all out:

1. Add the sets
2. Subtract the pairwise intersections
3. Add the triple-intersections
4. Subtract the four-way intersection

We can generalize this to n variables, alternating adding and subtracting the n-way intersections.

1.4 Probability

The informal definition of probability is the number of ways that the event can happen divided by all the possibilities. However, this definition assumes that all possibilities are equally likely. For example, if we have a space of events that is imbalanced, if we draw two circles of the same radius at random points, the number of events contained in the two circles will not be the same. Thus we need a different way of formally defining probability.

We have a fair coin, and we toss it 3 times. What is the probability that we don't get any heads. We see that there are 8 total cases, and only 1 case in which we can get no heads, giving us a probability of $\frac{1}{8}$.

If we have two dice, and they are fair. What is the probability that we roll a total of 7? We see that there are 6 ways to get a total of 7, and there are a total of $6^2 = 36$ possibilities, giving us a probability of $\frac{1}{6}$. The probability that we can get 2 is 1/36, as we have only the case where both dice role 1.

Let's look at something more complicated, poker. In poker, we are interested in 5-card hands. Let's calculate the probability that we get a flush, which is 5 cards of the same suit. The denominator is the total number of 5-card hands, which is $\binom{52}{5}$ (Not $P(52, 5)$ because order does not matter). Now let's find how many 5-card hands are flushes. We have $\binom{4}{1}$ ways of picking the suit, and $\binom{13}{5}$ hands for each suit. This gives us a probability of

$$P = \frac{\binom{4}{1}\binom{13}{5}}{\binom{52}{5}}$$

which is roughly 0.2%.

Let's look at straights, 5 cards of consecutive rank. Aces can be either end of the straight. There can also be no wrap-arounds. What is the likelihood of getting a straight? We have to first pick one of 10 lower ends, and then pick a suit for that card, and then pick the other 4 cards from any suit:

$$\frac{10 \times \binom{4}{1}^5}{\binom{52}{5}}$$

Now what about straight flushes? Well we can choose our starting point, and pick a suit, and then there is only 1 suit to pick from:

$$\binom{10}{1}\binom{4}{1}$$

If we want to do the probability of getting a 4-of-a-kind:

$$\frac{\binom{13}{1}\binom{4}{4}\binom{48}{1}}{\binom{52}{5}}$$

What is the probability that a 3 card hand will have the same suit?

$$\frac{\binom{4}{1}\binom{13}{3}}{\binom{52}{3}}$$

We can use this to generalize the flush, letting C be the number of cards in a deck, N being the number of cards in a hand, and S being the number of suits:

$$\frac{\binom{S}{1}\binom{\frac{C}{S}}{N}}{\binom{C}{N}}$$

Let's move to character strings. If we have a set of 65 characters (uppercase, lowercase, 0-9, and 3 special characters). What is the probability that a string is of length 10 with no repeated characters.

We have that the number of length 10 strings with no repeated characters is $P(65, 10)$, and the total number of strings is 65^10:

$$\frac{P(65, 10)}{65^{10}}$$

What is the probability that we sample a length-10 string with at least one character that repeats?

$$\frac{65^{10} - P(65, 10)}{65^{10}}$$

What is the probability that we sample a length-66 string with no repeated characters?

This one is just 0, as we only have 65 characters.

Length-65 string with no repeated characters?

$$\frac{P(65, 65)}{65^{10}}$$

Moving back to cards, what is the probability that if we sample a card, it is a face card or a heart?

$$\frac{\binom{12}{1}}{\binom{52}{1}} + \frac{\binom{13}{1}}{\binom{52}{1}} - \frac{\binom{3}{1}}{\binom{52}{1}}$$

1.5 Conditional Probability

If we have an event A that occurs, how does it affect the probability of B? Well this depends on whether B is dependent or independent of A. If A and B are completely disjoint, then B will not see a change in its probability.

Suppose we have two dice. Let event A be that the sum of the dice is a multiple of 4, which has probability $\frac{9}{36} = \frac{1}{4}$. Let event B be that the first die comes up 3, which has probability $\frac{6}{36} = \frac{1}{6}$. What is the probability of A given B?

Intuitively, we see that we have only 6 outcomes, as B has occurred and it has 6 outcomes. Of these outcomes, only $(3, 1)$ and $(3, 5)$ will work, and therefore $P(A)$ is now $\frac{2}{6} = \frac{1}{3}$.

We can generalize conditional probability:

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

Two events are independent if

$$P(A|B) = P(A)$$

Two events are conditionally independent with respect to a third event if

$$P(A \cap B|C) = P(A|C)P(B|C)$$

This leads into Baye's Law:

$$P(A|B) = P(B|A)\frac{P(A)}{P(B)}$$

1.6 k-nomial Theorem

The binomial theorem states that

$$(x + y)^2 = x^2 + 2xy + y^2$$

$$(x + y)^3 = x^3 + 3x^2y + 3xy^2 + y^3$$

$$(x + y)^4 = x^4 + 4x^2y^2 + 6x^2y^2 + 4x^2y^2 + y^4$$

How can we generalize this pattern? If we have $(x + y)^5$, we have 2^5 terms, as we have 2 possible choices of variable for each term, and 5 terms that we are multiplying together. This gives us 32 terms. We now want to combine like terms, so how many of those terms are of the form x^2y^3? We can think of this like substrings, and we see that it is $\binom{5}{2}$ or $\binom{5}{3}$.

If we want to choose the coefficient of x^3y^4 in $(x + y)^7$, we have that it is $\frac{7!}{3!4!} = \binom{7}{3}$. Thus we can see how this generalizes. The coefficient of the $x^r y^{n-r}$ is the $\binom{n}{r}$. This leads us to the binomial theorem:

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

We have seen that

$$\binom{n}{r} = \binom{n}{n-r}$$

Here is another combinatorial identity:

$$(\forall n, r \in \mathbb{N}^{\geq 1}) \left[(r \leq n) \to \binom{n}{r} = \binom{n-1}{r-1} + \binom{n-1}{r} \right]$$

This one essentially says that the number of ways to pick r people from a set of n people is equal to the right side. The right says that we pick one of the people, and now are trying to fill up the rest of the slots, which is $\binom{n-1}{r-1}$. The second term is if we don't pick that person, and now we have $n-1$ people and still have r slots open.

This identity is the basis for Pascal's Triangle.

2 Propositional Logic

This is the most elementary kind of logic in computer science. Also knwn as Boolean Logic, as it was pioneered by George Boole. We begin wit the propositional symbol, a box/bit that can hold a value of true or false (1 or 0). They are denoted using lowercase english letters. There are 3 basic operations in boolean logic, conjunction (AND), disjunction (OR), and negation (NOT). The easiest one of these is the negation, which flips true values to false and false values to true. This is a unary operation, as it acts on only 1 value.

The conjunction, or logical AND, is represented by the $\wedge$. This is a binary operator, and returns 1 iff both arguments are 1. Note that the conjunction is commutative, that is, $p \wedge q = q \wedge p$.

We also have the disjunction, also known as OR. As long as one of the two arguments is true, then the result will be true.

We can then look at the XOR, or exclusive or, $p \oplus q$. This returns true iff the two inputs are different from each other. Essentially, this functions the same as the OR, except for the case where both are true. This is closer to the real world meaning of or (If someone says one or the other, they generally do not mean that both are an option).

If we look at the expression $p \vee (p \wedge q)$, we can actually see that this is just equivalent to p (from inspection of the truth table). This is known as the law of absorption. In general, if we have and expression of the form $p \vee (p \wedge \cdots)$, this is equivalent to p. In fact, we also see that expressions of the form $p \wedge (p \vee \cdots)$ are also equivalent to p.

The implication is another important operation, also known as the "if-then", $p \Rightarrow q$, also read as "p implies q". This is somewhat hard to explain. If we have an alien Gorslax studying Earth's birds, and he asks whether all Earth birds fly. They look at 3 birds, all of which fly, so he has positive evidence for the claim that all birds fly. Thus he has found positive evidence (this is the $T \implies T$ case). He then sees a rhino, which doesn't fly. This does not affect his hypothesis. This is the $F \implies F$ case. If he now sees a bat, which is not a bird but can fly. This does not affect his hypothesis. This is the $F \implies T$ case. However, if he sees an emu, which is a bird that cannot fly, this provides negative evidence for his hypothesis, also known as a counterexample. This is the $T \implies F$ case.

The bi-conditional is essentially the "if and only if", and is equivalent to the XNOR. $p \iff q$ returns True if both p and q are equivalent.

If we look at the statements

$$p \wedge (\neg p) \qquad p \vee (\neg p)$$

We see that the first one is a contradiction, in that it will always be 0, and the second is a tautology, in that it will always be 1.

We can also define De Morgan's law:

$$\neg(a \wedge b) = (\neg a) \vee (\neg b)$$

and the other law:

$$\neg(a \vee b) = (\neg a) \wedge (\neg b)$$

How can we prove equivalences? One way is to do it through truth tables. However, this can be tedious. Instead, we can chain laws of logical equivalence until the expressions are equivalent.

We now want to build some circuits. We want to build a circuit that adds arbitrarily large binary numbers. The rightmost operation will be performed by a Half-Adder, which will take in two bits of input, and return two bits, one for output and one as a carry bit. The next operation will take in 3 inputs (2 inputs and the carry bit from the half-adder), and return 2 outputs, one carry and one output. For this, we wil need a full-adder. For an n-bit adder, we need $n-1$ full adders and 1 half-adder.

To make the half-adder, we can just make the truth table, and we see that the result bit is given by the XOR of the two input bits, and the carry bit is the AND of the two bits.

We can actually simplify a 5-gate XOR:

$$x \oplus y \equiv (x \wedge (\neg y)) \vee ((\neg x) \wedge y) \equiv (x \vee y) \wedge (\neg(x \wedge y))$$

This XOR is made with only 4 gates, which overall is a large benefit. This gets us a 5-gate half-adder. However, we can actually use the AND inside the XOR to get the value for the outside AND, which means that we can actually build it with 4 gates total.

Now let's build the Full-Adder. We need it add 3 bits, and return 2 bits. We could do this with a truth table, but instead we can actually make it out of stacks of Half-Adders.

3 Set Theory

A set is defined as a collection of distinct objects. We use the notation $x \in S$ to state that x is an element of S. Generally, sets are unordered. There are some basic sets:

$$\mathbb{N} = \{0, 1, 2, 3, \dots\}$$

$$\mathbb{Z} = \{\dots, -3, -2, -1, 0, 1, 2, 3, \dots\}$$

$$\mathbb{Q} = \left\{\frac{a}{b}, a \in \mathbb{Z} \wedge b \in \mathbb{Z} \wedge b \neq 0\right\}$$

and of course $\mathbb{R}$.

B is a subset of A ($B \subset A$) iff

$$[\forall x \in U]\{(x \in A) \Rightarrow (x \in B)\}$$

A is a proper subset of B if A only contains elements of B. We can then define the union of two sets:

$$A \cup B = \{(x \in A) \vee (x \in B)\}$$

We can also define the intersection:

$$A \cap B = \{(x \in A) \wedge (x \in B)\}$$

Moving on to more complicated ones, we have the absolute complement:

$$A^C = \{(x \notin A)\} = \{(x \in U) \wedge (\neg(x \in A))\}$$

And the relative complement:

$$A - B = \{(x \in A) \wedge (x \notin B)\}$$

4 Big Oh-Notation

When talking about complexity, we drop the multiplicative constants, and we generally only consider the behavior as $n \to \infty$. We can define Big-Oh formally.

Let $f(n)$ and $g(n)$ be functions of n. We say that $f(n)$ is $\mathcal{O}(g(n))$ iff

$$(\exists n_0 \in \mathbb{N}, c \in \mathbb{R}^{>0})[(\forall n \geq n_0)[f(n) \leq c \cdot g(n)]]$$

n_0 allows us to look at the behavior past some point, and $c > 0$ tells us that the constants don't matter.

If we are told that we want to show that a function is $\mathcal{O}(n^2)$, such as $3n^2 - 4n + 100$. We just need to find a pair of values for n_0 and c such that the definition holds. In this case, $n_0 = 25$ and $c = 3$ work. Where did we get $c = 3$? Well we have the leading coefficient of the n^2 term.

5 Proofs

An integer n is called even iff there exists an integer k such that $n = 2k$. An integer is called odd iff it is not even. A corollary of this is that an integer n is odd iff there exists and integer k such that $n = 2k + 1$. This property is known as the parity of an integer. Note that from this definition, 0 is even.

If we have the statement "the sum of an odd and an even integer is odd", we have to try to prove it. We can do this in multiple ways, such as the verbal style, or the symbolic style. There is also the mix between those two.