

Counting & Probability Intro

Note 15
Note 16

Zeroth rule of counting: For sets A, B , if there exists a bijection $f : A \rightarrow B$, then $|A| = |B|$.

First rule of counting: When counting the number of ways to make a sequence of k choices, if there are n_1 options for the first choice, n_2 options for the second choice regardless of the first choice, etc., then there are a total of $n_1 \cdot n_2 \cdots n_k$ ways to make all k choices.

Second rule of counting: Let B be the set of unordered objects for counting a sequence of choices, and let A be the ordered set of such objects. If there exists an m -to-1 mapping $f : A \rightarrow B$, then $|A| = m \cdot |B|$.

These rules give the formulas for choosing k items out of n total items:

	with replacement	without replacement
order matters	n^k	$\frac{n!}{(n-k)!}$
order doesn't matter	$\binom{n+k-1}{k} = \binom{n+k-1}{n-1}$	$\frac{n!}{k!(n-k)!} = \binom{n}{k}$

The bottom-right entry is also known as *combinations*, and the bottom-left as *stars and bars*: to split k items into n categories, arrange k stars and $n-1$ bars separating the categories. Most counting problems involve a mix of all these methods; don't box yourself into using just one!

Probability space: a tuple $(\Omega, \mathbb{P})$, where the *sample space* Ω is the set of all outcomes ω , and the *probability function* $\mathbb{P}$ assigns a probability to each outcome, satisfying

$$0 \leq \mathbb{P}[\omega] \leq 1 \quad \text{and} \quad \sum_{\omega \in \Omega} \mathbb{P}[\omega] = 1.$$

Event: an event A is a subset of Ω , and $\mathbb{P}[A] = \sum_{\omega \in A} \mathbb{P}[\omega]$. The *complement* $\bar{A}$ consists of the outcomes not in A , and $\mathbb{P}[\bar{A}] = 1 - \mathbb{P}[A]$.

Uniform probability space: all outcomes are assigned the same probability, i.e. $\mathbb{P}[\omega] = \frac{1}{|\Omega|}$. Then for any event A , we have $\mathbb{P}[A] = \frac{|A|}{|\Omega|}$ — this is just counting!

1 Intro to Counting

Note 15

In the below questions, do not blindly apply the formulas on the previous page—try to re-derive them yourself! After you solve each part, also identify which of the squares in the table it corresponds to, and what the values of n and k are.

- How many 4-card hands are there in a standard 52-card deck? (Any rearrangement of the cards in your hand is counted as identical.)
- A US phone number consists of a country code (always “+1”), an area code (any 3 digits), a telephone prefix (any 3 digits), and a line number (any 4 digits). How many US phone numbers are there?

- (c) How many anagrams of the word “COUNT” are there? What about “BERKELEY”? (An *anagram* of a word is a rearrangement of its letters; “CTONU” and “TNUOC” are both anagrams of “COUNT”.)

The concept of “stars and bars” seems at first glance quite different from the idea of “order doesn’t matter, with replacement”. We’ll look at how these two concepts are actually the same.

- (d) Consider the problem of counting the number of ways of splitting ten \$1 bills among 4 people. How many ways of splitting the bills are there?

Compare this with the problem of counting the number of ways of arranging 10 numbers from the set $\{1, 2, 3, 4\}$, where order doesn’t matter and we have replacement. What is the relationship between the ways of counting these two scenarios? (*Hint*: Let each person be labeled by a digit from $\{1, 2, 3, 4\}$. Can you represent a split of the ten bills with a sequence of digits?)

Solution:

- (a) $\binom{52}{4}$. Order doesn’t matter here, since different arrangements of the same hand are identical. We do not allow replacement, since we are drawing cards from a deck without putting them back. Hence, since we have 52 cards total, and we’re picking 4 for our hand, there are $\binom{52}{4}$ possible hands.

This part corresponded to the “order doesn’t matter, without replacement” square in the table, with $n = 52$ and $k = 4$.

- (b) 10^{10} . The country code doesn’t change, so we have a total of 10 digits in the phone number that can change. These can each be any of the digits 0–9, so we have 10^{10} possible phone numbers.

Here, order matters (the sequence “12” is different from the sequence “21” in the phone number), and we have replacement (we can reuse digits).

This part corresponded to the “order matters, with replacement” square in the table, with $n = 10$ and $k = 10$.

- (c) There are $5!$ anagrams of the word “COUNT”, since we can arrange the 5 distinct letters with order and without replacement.

There are $\frac{8!}{3!}$ anagrams of the word “BERKELEY”. There are a total of $8!$ ways of arranging the letters, but this is an overcount: we have 3 E’s, and re-ordering the E’s by themselves does not change the resulting word. Since there are a total of $3!$ ways of arranging *just* the E’s amongst themselves, we must divide by this quantity to fix our overcounting.

This part corresponded to the “order matters, without replacement” square in the table, with $n = 5$ and $k = 5$ for the first word, and $n = 8$ and $k = 8$ for the second word (but we had to account for overcounting in the second case).

- (d) We can think of each person as a “bin” for the 10 bills to fall into. We can then think of this problem as arranging the 10 bills (thought of as “stars”) along with 3 dividers for the 4 bins (thought of as “bars”), in a line. For example, we could have the following split:

*** | ** | * | *****

(Note that this is just a rephrasing of the binary string correspondence with balls and bins used in the lecture notes.)

We have a total of 13 “objects” that we’re arranging here. Notice that once we choose the location of the 10 stars, the location of the bars are fixed (similarly, once we choose the location of the 3 bars, the location of the 10 stars are fixed). There are a total of $\binom{13}{10}$ ways of choosing the location of the 10

stars (or $\binom{13}{3}$ ways of choosing the location of the 3 bars), so this is the number of ways of arranging these objects in a line.

This means that there are $\binom{13}{10} = \binom{13}{3}$ ways of splitting ten \$1 bills among 4 people.

Trying to relate this to the problem of arranging 10 numbers from the set $\{1, 2, 3, 4\}$, suppose we label each person with a digit. A sequence of 10 of these digits then corresponds to a way of splitting the 10 bills, where each digit tells us which person the bill is given to. For example, the split from before could correspond to the following sequence:

$$(1, 1, 1, 2, 2, 3, 4, 4, 4, 4)$$

Here, we don't care about order, since the sequence $(1, 2)$ is the same as the sequence $(2, 1)$. In either case, person 1 and person 2 both get one bill each—the order in which the person receives the bill doesn't matter. We also allow replacement, since each person can receive multiple bills (their digit can appear multiple times in the sequence).

Put another way, these constraints essentially tell us that since order does not matter, we really only care about *count* of each digit. For stars and bars, each bin corresponds to a possible digit, and the number of stars in each bin corresponds to the number of times that digit appears in the sequence.

This part corresponded to the “order doesn't matter, with replacement” square in the table, with $n = 4$ and $k = 10$.

2 Halloweekend

Note 15
Note 16

It's Halloweekend in Berkeley! Let's use counting and probability to navigate through the festivities.

- A candy bowl has 4 pieces left: 2 chocolate bars and 2 boxes of raisins. One at a time, you grab pieces without looking and eat them until you get a chocolate bar. List the sample space Ω .
- Eight people enter a costume contest. The judges pick 3 finalists, then crown one of the finalists as the winner. How many ways can the finalists be chosen? How many ways can the finalists *and* winner be chosen?
- You have 10 identical tea lights for the 4 jack-o'-lanterns on your porch. Assuming there are no other constraints, how many ways can you distribute the tea lights among them? How many ways are there if every jack-o'-lantern needs to glow (to glow, a jack-o'-lantern needs at least one tea light)?
- While trick-or-treating, you knock on 3 doors. Each door has an equal chance of opening up or being unanswered (i.e. each outcome after 3 knocks is equally likely). What is the probability that exactly two doors are never opened?
- Four friends visit the pumpkin patch and each pick a pumpkin from 10 available varieties uniformly at random. What is the probability least two friends come home with the same variety?

[Hint: Use the complement of the event]

Solution:

- $\Omega = \{C, RC, RRC\}$, writing C for chocolate and R for raisins. The experiment stops at the first C , and with only 2 boxes of raisins the third grab is guaranteed to be chocolate.
- There are $\binom{8}{3} = 56$ ways to choose the finalists and $\binom{8}{3} \cdot 3 = 168$ ways to choose the finalists and winners. There are $8 \cdot 7 \cdot 6$ ordered selections of finalists, and each group is counted $3!$ times, so there are $\frac{8 \cdot 7 \cdot 6}{3!} = \binom{8}{3}$ groups; the winner is then one of the 3. Alternatively, choose the winner first: $8 \cdot \binom{7}{2} = 168$ — same thing.

- (c) $\binom{13}{3} = 286$ and $\binom{9}{3} = 84$. Each distribution corresponds to an arrangement of 10 stars and 3 bars. If every jack-o'-lantern must glow, place one tea light in each first; then we have an arrangement of 6 stars and 3 bars.
- (d) $\frac{3}{8}$. Writing O for an opened door and N otherwise, $\Omega = \{OOO, OON, ONO, ONN, NOO, NON, NNO, NNN\}$. Each outcome is equally likely (each has probability $1/|\Omega| = 1/8$). So we sum over the event: $\mathbb{P}[NNO] + \mathbb{P}[NON] + \mathbb{P}[ONN] = 3 \cdot 1/8 = 3/8$.
- (e) We consider the complement of the event, that none of the friends come home with the same variety. There are $10 \cdot 9 \cdot 8 \cdot 7 = 5040$ ways in which this happens. The probability of the complement occurring is thus $\frac{5040}{10^4}$. To find the probability of the original event, we take 1 minus this probability. So, the probability is

$$1 - \frac{5040}{10000} = 1 - \frac{63}{125} = \frac{62}{125}.$$

3 Picking Teams

Note 15

- (a) UC Berkeley is forming a new, n -person robotics team. There are $2n$ interested students: n mechanical engineers and n programmers. (Assume that no student is both a mechanical engineer and a programmer.)

Find the number of distinct n -person teams.

- (b) Suppose that all robotics teams must also name a team captain who is a mechanical engineer. Find the number of ways to pick an n -person team with a mechanical engineer as the captain.
- (c) The 25 students in CS 70 have a strange way of picking a study group. Each person is given a distinct number 1 through 25 based on random chance, and participants may form a group as long as the sum of their numbers is **at most** 162. (The TAs are afraid of the number 162 for some reason...)

How many different study groups are possible? That is, how many different subsets of students are there, which are allowable under the ranking rule? (**Hint:** $162 = [(\sum_{i=1}^{25} i) - 1]/2$)

Solution:

- (a) **Method 1:** $\binom{2n}{n}$ This is simply the number of ways you can pick n students to be on the team.

Method 2: $\sum_{k=0}^n \binom{n}{k}^2$

We begin by noticing that $\binom{n}{k}^2 = \binom{n}{k} \cdot \binom{n}{n-k}$. This product gives us the number of ways of picking k mechanical engineers and $n-k$ programmers to be on the team. We add up all the products involving anywhere from 0 mechanical engineers all the way to n . This gives us the total number of ways to pick a team of n students.

- (b) Take the same scenario as part (a) except now you have to choose a team captain who is also a mechanical engineer.

Method 1: $n \binom{2n-1}{n-1}$

This is the number of ways of picking the team captain multiplied with the number of ways of picking the rest of the team from the remaining students. The product gives the total number of teams with a mechanical engineer team captain.

Method 2: $\sum_{k=1}^n k \binom{n}{k}^2$

We begin similarly by noticing that $k \cdot \binom{n}{k}^2 = k \cdot \binom{n}{k} \cdot \binom{n}{n-k}$. Here as before we are picking k mechanical engineers and $n - k$ programmers to be on the team. However amongst the k mechanical engineers on the team, we choose one of them to be the team leader. We add up all the products involving anywhere from 1 mechanical engineers all the way to n . This gives us the total number of ways to pick a team of n students with a team captain who is a mechanical engineer.

- (c) Note that the sum of the first 25 positive integers is 325. We can construct a bijection between sets of numbers that have sum ≤ 162 and sum greater than 162, namely, for every set of numbers with sum ≤ 162 , the numbers not chosen form a set with sum greater than 162. Therefore, exactly half of all subsets of the 25 students have numbers that sum to ≤ 162 , or $(1/2)(2^{25}) = 2^{24}$ subsets.

4 Easter Eggs

Note 15
Note 16

You made the trek to Gateway for a Spring Break-themed homework party, and every attendee gets to leave with a party favor. You're given a bag with 20 chocolate eggs and 40 (empty) plastic eggs. You pick 5 eggs (uniformly) without replacement.

- What is the probability that the first egg you drew was a chocolate egg?
- What is the probability that the second egg you drew was a chocolate egg?
- What is the probability that the first and fifth eggs you drew were both empty plastic eggs?

Solution:

$$(a) \mathbb{P}[\text{chocolate egg}] = \frac{20}{60} = \frac{1}{3}.$$

- (b) We first want to count the number of configurations where the second egg is chocolate. There are two cases to consider: either the first egg is chocolate or it is plastic. If the first and second eggs are chocolate (there are $20 \cdot 19$ configurations of this), then there are $58 \cdot 57 \cdot 56$ ways to draw the last three eggs. If the first egg is plastic and the second egg is chocolate (there are $40 \cdot 20$ ways this happens), then there are still $58 \cdot 57 \cdot 56$ ways to draw the last three eggs. Dividing by the total number of ways to draw the eggs ($60 \cdot 59 \cdot 58 \cdot 57 \cdot 56$) and adding our casework, we get the probability is

$$\frac{(20 \cdot 19) \cdot 58 \cdot 57 \cdot 56 + (40 \cdot 20) \cdot 58 \cdot 57 \cdot 56}{60 \cdot 59 \cdot 58 \cdot 57 \cdot 56} = \frac{1}{3}.$$

Alternatively, we don't need to do casework. There are 20 ways to choose the chocolate egg that goes second. We can choose 4 from the rest of the 59 eggs to go in the other spots (where order matters). This gives a total count of $20 \cdot 59 \cdot 58 \cdot 57 \cdot 56$. There are $60 \cdot 59 \cdot 58 \cdot 57 \cdot 56$ orderings for the 5 eggs in total, so our probability is:

$$\frac{20 \cdot 59 \cdot 58 \cdot 57 \cdot 56}{60 \cdot 59 \cdot 58 \cdot 57 \cdot 56} = \frac{20}{60} = \frac{1}{3}.$$

Short calculation: By symmetry, this is the same probability as part (a), $1/3$. This is because we don't know what type of egg was picked on the first draw, so the distribution for the second egg is the same as that of the first.

- (c) By symmetry, since we don't know any information about the 2nd, 3rd, or 4th eggs, we have

$$\mathbb{P}[\text{5th egg} = \text{plastic} \cap \text{1st egg} = \text{plastic}] = \mathbb{P}[\text{2nd egg} = \text{plastic} \cap \text{1st egg} = \text{plastic}].$$

There are $40 \cdot 39$ ways to have the first two eggs be plastic (assuming order matters), and $58 \cdot 57 \cdot 56$ ways to choose the rest of the eggs. There are $60 \cdot 59 \cdot 58 \cdot 57 \cdot 56$ total ways of drawing the eggs, and

so the probability is

$$\frac{40 \cdot 39 \cdot 58 \cdot 57 \cdot 56}{60 \cdot 59 \cdot 58 \cdot 57 \cdot 56} = \frac{40 \cdot 39}{60 \cdot 59} = \frac{26}{59}.$$

5 Grids and Trees!

Note 15

Suppose we are given a grid, where one starts at the origin $(0,0)$. We want to eventually reach some coordinate (a,b) . On this grid, we are only allowed to move left, right, up, or down by increments of 1.

- (a) How many **shortest** paths are there that go from $(0,0)$ to (n,n) , for $n \geq 1$? (Hint: not all of the four possible moves may be useful.)
- (b) How many shortest paths are there that go from $(0,0)$ to $(n-1, n+1)$, for $n \geq 1$?

Now, consider shortest paths that meet the conditions where we can only visit points (x,y) where $y \leq x$. That is, the path cannot cross line $y = x$.

- (c) Compute the number of shortest paths from $(0,0)$ to (n,n) that cross $y = x$. (Hint: Let (i,i) be the first time the shortest path crosses the line $y = x$. Then the remaining path starts from $(i,i+1)$ and continues to (n,n) . If in the remainder of the path one exchanges y -direction moves with x -direction moves and vice versa, where does one end up?)

Solution: Let $(x,y) \rightarrow (x+1,y)$ be a move 'right' command, and $(x,y) \rightarrow (x,y+1)$ be a move 'up' command. The critical observation is that any shortest path must entirely be consisted only of move 'right' and 'up' commands.

- (a) There are $\binom{2n}{n}$ paths, as there are total number of $2n$ moves, and n of them must be move 'right' command, the rest of them must be the move 'up' command.
- (b) There are $\binom{2n}{n-1}$ paths as there are now $n-1$ move 'right' command. There are still $2n$ total moves to make.
- (c) We will argue that the set of all shortest paths that cross $y = x$ (we shall call these **invalid paths**), form a bijection with the set of all paths from $(0,0)$ to $(n-1, n+1)$. Suppose we have a path that crosses $y = x$. Once a path crosses $y = x$, we can flip the later portion of the path across the $y = x$ line. Let the first time an invalid path crosses $y = x$ be at (i,i) and arrives at $(i,i+1)$. Then if we do not flip the path, it will arrive at (n,n) by taking $n-i$ "right" commands, and $n-1-i$ "up" commands. If we flip these commands, it will go to $(i+(n-1-i), (i+1)+(n-i)) = (n-1, n+1)$. So all invalid paths map to a path from $(0,0)$ to $(n-1, n+1)$. Next we argue that all path from $(0,0)$ to $(n-1, n+1)$ maps to a invalid path. Paths from $(0,0)$ to $(n-1, n+1)$ must cross the line $y = x$, let it first cross the line at (i,i) and arrives $(i,i+1)$. Then it must take $(n-1)-i$ "right" commands, and $(n+1)-(i+1)$ "up" commands. We flip these commands and so we now have, $n-1-i$ "up" commands, $n-i$ "right" commands. Then the path will arrive $(i+(n-i), (i+1)+(n-1-i)) = (n,n)$ and this new path is considered as invalid path since it crosses $y = x$ at point (i,i) . So all paths from $(0,0)$ to $(n-1, n+1)$ can be mapped to an invalid paths.

So there is a bijective mapping between invalid paths and paths from $(0,0)$ to $(n-1, n+1)$. Hence, the number of invalid paths is $\binom{2n}{n-1}$.