Today.

More Counting.

Probability.

Sampling and counting.

First rule: $n_1 \times n_2 \cdots \times n_3$.

k Samples with replacement from n items: n^k.
Sample without replacement: $\frac{n!}{(n-k)!}$

Second rule: when order doesn't matter divide..when possible.

Sample without replacement and order doesn't matter: $\binom{n}{k} = \frac{n!}{(n-k)!k!}$.
"n choose k"

One-to-one rule: equal in number if one-to-one correspondence.

Sample with replacement and order doesn't matter: $\binom{k+n-1}{n-1}$.

Balls in bins.

"k Balls in n bins" $\equiv$ "k samples from n possibilities."

"indistinguishable balls" $\equiv$ "order doesn't matter"

"only one ball in each bin" $\equiv$ "without replacement"

5 balls into 10 bins
5 samples from 10 possibilities with replacement
Example: 5 digit phone numbers.
5 Balls/places choose from 10 bins/digits.

5 indistinguishable balls into 52 bins only one ball in each bin
5 samples from 52 possibilities without replacement
Example: Poker hands.
5 balls/cards into 52 bins/possible cards.

5 indistinguishable balls into 3 bins
5 samples from 3 possibilities with replacement and no order
Dividing 5 dollars among Alice, Bob and Eve.
5 dollars/balls choose from 3 people/bins.

Sum Rule

Two indistinguishable jokers in 54 card deck.
How many 5 card poker hands?
Sum rule: Can sum over disjoint sets.
No jokers "exclusive" or One Joker "exclusive" or Two Jokers

$$\binom{52}{5} + \binom{52}{4} + \binom{52}{3}.$$

Two distinguishable jokers in 54 card deck.
How many 5 card poker hands? Choose 4 cards plus one of 2 jokers!

$$\binom{52}{5} + 2 * \binom{52}{4} + \binom{52}{3}$$

Wait a minute! Same as choosing 5 cards from 54 or

$$\binom{54}{5}$$

Theorem: $\binom{54}{5} = \binom{52}{5} + 2 * \binom{52}{4} + \binom{52}{3}$.

Algebraic Proof: No need! Above is combinatorial proof. $\square$

Combinatorial Proofs.

Theorem: $\binom{n}{k} = \binom{n}{n-k}$

Proof: How many subsets of size k? $\binom{n}{k}$

How many subsets of size k?
Choose a subset of size $n-k$
and what's left out is a subset of size k.
Choosing a subset of size k is same
as choosing $n-k$ elements to not take.
$\implies \binom{n}{n-k}$ subsets of size k.

$\square$

Pascal's Triangle

$$\begin{array}{ccccccccc} &&&& 0 &&&& \\ &&& 1 && 1 &&& \\ && 1 && 2 && 1 && \\ & 1 && 3 && 3 && 1 & \\ 1 && 4 && 6 && 4 && 1 \end{array}$$

Row n: coefficients of $(1+x)^n = (1+x)(1+x)\cdots(1+x)$.

Foil (4 terms) on steroids:
2^n terms: choose 1 or x froom each factor of $(1+x)$.

Simplify: collect all terms corresponding to x^k.
Coefficient of x^k is $\binom{n}{k}$: choose k factors where x is in product.

$$\begin{array}{ccccccc} &&& \binom{0}{0} &&& \\ && \binom{1}{0} && \binom{1}{1} && \\ & \binom{2}{0} && \binom{2}{1} && \binom{2}{2} & \\ \binom{3}{0} && \binom{3}{1} && \binom{3}{2} && \binom{3}{3} \end{array}$$

Pascal's rule $\implies \binom{n+1}{k} = \binom{n}{k} + \binom{n}{k-1}$.

Combinatorial Proofs.

Theorem: $\binom{n+1}{k} = \binom{n}{k} + \binom{n}{k-1}$.
Proof: How many size k subsets of $n+1$? $\binom{n+1}{k}$.

How many size k subsets of $n+1$?
How many contain the first element?
Choose first element,
need to choose $k-1$ more from remaining n elements.
$\implies \binom{n}{k-1}$
How many don't contain the first element ?
Need to choose k elements from remaining n elts.
$\implies \binom{n}{k}$

So, $\binom{n}{k-1} + \binom{n}{k} = \binom{n+1}{k}$. $\square$

Combinatorial Proof.

Theorem: $\binom{n}{k} = \binom{n-1}{k-1} + \cdots + \binom{k-1}{k-1}$.
Proof: Consider size k subset where i is the first element chosen.

$$\{1,\ldots,i,\underline{\ldots,n}\}$$

Must choose $k-1$ elements from $n-i$ remaining elements.
$\implies \binom{n-i}{k-1}$ such subsets.

Note term $\binom{n-i}{k-1}$ corresponds to subsets where first item is i.

Do the terms correspond to disjoint Groups? Yes? No?
Any pair of subsets in different Groups have diferrent first items.
So Yes!!

Add their sizes up to get the total number of subsets of size k
which is also $\binom{n}{k}$. $\square$

Binomial Theorem: $x = 1$

Theorem: $2^n = \binom{n}{n} + \binom{n}{n-1} + \cdots + \binom{n}{0}$

Proof: How many subsets of $\{1,\ldots,n\}$?
Construct a subset with sequence of n choices:
element i **is in** or **is not** in the subset: 2 poss.
First rule of counting: $2 \times 2 \cdots \times 2 = 2^n$ subsets.

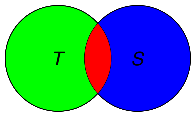
How many subsets of $\{1,\ldots,n\}$?
$\binom{n}{i}$ ways to choose i elts of $\{1,\ldots,n\}$.
Disjoint? Different size if in different group. So..Yes!.
Sum over i to get total number of subsets..which is also 2^n. $\square$

Simple Inclusion/Exclusion

Sum Rule: For disjoint sets S and T, $|S \cup T| = |S| + |T|$
Used to reason about all subsets
by adding number of subsets of size 1, 2, 3,...

Also reasoned about subsets that contained
or didn't contain an element. (E.g., first element, first i elements.)

Inclusion/Exclusion Rule:
For any S and T, $|S \cup T| = |S| + |T| - |S \cap T|$.



$S \cap T$

In T. $\implies |T|$
In S. $\implies + \ |S|$
Elements in $S \cap T$ are counted twice.
Subtract. $\implies -|S \cap T|$

$$|S \cup T| = |S| + |T| - |S \cap T|$$

Using Inclusion/Exclusion.

Inclusion/Exclusion Rule: For any S and T,
$|S \cup T| = |S| + |T| - |S \cap T|$.

Example: How many 10-digit phone numbers have 7 as their first or second digit?

S = phone numbers with 7 as first digit. $|S| = 10^9$

T = phone numbers with 7 as second digit. $|T| = 10^9$.

$S \cap T$ = phone numbers with 7 as first and second digit. $|S \cap T| = 10^8$.

Answer: $|S| + |T| - |S \cap T| = 10^9 + 10^9 - 10^8$.

Summary.

First Rule of counting: Objects from a sequence of choices:
n_i possibilitities for ith choice.
$n_1 \times n_2 \times \cdots \times n_k$ objects.

Second Rule of counting: If order does not matter.
Count with order. Divide by number of orderings/sorted object.
Typically: $\binom{n}{k}$.

Stars and Bars: Sample k objects with replacement from n.
Order doesn't matter.
Typically: $\binom{n+k-1}{k}$.

Inclusion/Exclusion: two sets of objects.
Add number of each subtract intersection of sets.
Sum Rule: If disjoint just add.

Combinatorial Proofs: Identity from counting same in two ways.
Pascal's Triangle Example: $\binom{n+1}{k} = \binom{n}{k-1} + \binom{n}{k}$.
RHS: Number of subsets of $n+1$ items size k.
LHS: $\binom{n}{k-1}$ counts subsets of $n+1$ items with first item.
$\binom{n}{k}$ counts subsets of $n+1$ items without first item.
Disjoint – so add!

CS70: On to probability.

Modeling Uncertainty: Probability Space

1. Key Points
2. Random Experiments
3. Probability Space

Key Points

- Uncertainty does not mean “nothing is known”
- How to best make decisions under uncertainty?
  - Buy stocks
  - Detect signals (transmitted bits, speech, images, radar, diseases, etc.)
  - Control systems (Internet, airplane, robots, self-driving cars, schedule surgeries in a hospital, etc.)
- How to best use ‘artificial’ uncertainty?
  - Play games of chance
  - Design randomized algorithms.
- Probability
  - Models knowledge about uncertainty
  - Discovers best way to use that knowledge in making decisions

The Magic of Probability

Uncertainty: vague, fuzzy, confusing, scary, hard to think about.

Probability: A precise, unambiguous, simple(!) way to think about uncertainty.

Uncertainty = Fear

Probability = Serenity

Our mission: help you discover the serenity of Probability, i.e., enable you to think clearly about uncertainty.

Your cost: focused attention and practice on examples and problems.

Random Experiment: Flip one Fair Coin

Flip a fair coin: (*One flips or tosses a coin*)

- Possible outcomes: Heads (H) and Tails (T) (*One flip yields either ‘heads’ or ‘tails’.*)
- Likelihoods: $H : 50\%$ and $T : 50\%$

Random Experiment: Flip one Fair Coin

Flip a fair coin:

What do we mean by the likelihood of tails is 50%?

Two interpretations:

- Single coin flip: 50% chance of ‘tails’ [subjectivist]

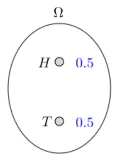
  Willingness to bet on the outcome of a single flip
- Many coin flips: About half yield ‘tails’ [frequentist]

  Makes sense for many flips
- Question: Why does the fraction of tails converge to the same value every time? Statistical Regularity! Deep!

Random Experiment: Flip one Fair Coin

Flip a fair coin: model

Physical Experiment



Probability Model

- The physical experiment is complex. (Shape, density, initial momentum and position, ...)
- The Probability model is simple:
  - A set Ω of outcomes: $\Omega = \{H, T\}$.
  - A probability assigned to each outcome: $Pr[H] = 0.5, Pr[T] = 0.5$.

Random Experiment: Flip one Unfair Coin

Flip an unfair (biased, loaded) coin:

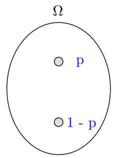
- Possible outcomes: Heads (H) and Tails (T)
- Likelihoods: $H : p \in (0, 1)$ and $T : 1 - p$
- Frequentist Interpretation:

  Flip many times $\Rightarrow$ Fraction $1 - p$ of tails
- Question: How can one figure out p? Flip many times
- Tautology? No: Statistical regularity!

Random Experiment: Flip one Unfair Coin

Flip an unfair (biased, loaded) coin: model

Physical Experiment



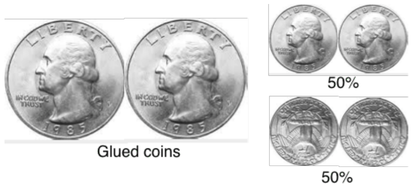
Probability Model

Flip Two Fair Coins

- Possible outcomes: $\{HH, HT, TH, TT\} \equiv \{H, T\}^2$.
- Note: $A \times B := \{(a, b) \mid a \in A, b \in B\}$ and $A^2 := A \times A$.
- Likelihoods: $1/4$ each.

Flip Glued Coins

Flips two coins glued together side by side:



- Possible outcomes: $\{HH, TT\}$.
- Likelihoods: $HH : 0.5, TT : 0.5$.
- Note: Coins are glued so that they show the same face.

Flip Glued Coins

Flips two coins glued together side by side:

- Possible outcomes: $\{HT, TH\}$.
- Likelihoods: $HT : 0.5, TH : 0.5$.
- Note: Coins are glued so that they show different faces.

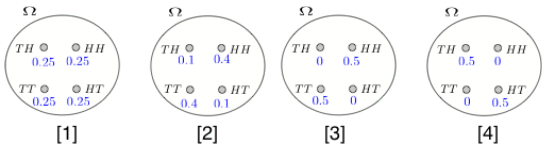
Flip two Attached Coins

Flips two coins attached by a spring:

- Possible outcomes: $\{HH, HT, TH, TT\}$.
- Likelihoods: $HH : 0.4, HT : 0.1, TH : 0.1, TT : 0.4$.
- Note: Coins are attached so that they tend to show the same face, unless the spring twists enough.

Flipping Two Coins

Here is a way to summarize the four random experiments:

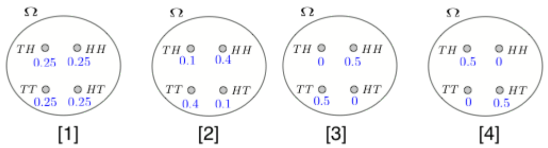


- Ω is the set of *possible* outcomes;
- Each outcome has a probability (likelihood);
- The probabilities are ≥ 0 and add up to 1;
- Fair coins: [1]; Glued coins: [3], [4];

  Spring-attached coins: [2];

Flipping Two Coins

Here is a way to summarize the four random experiments:



Important remarks:

- Each outcome describes the two coins.
- E.g., HT is one outcome of the experiment.
- It is wrong to think that the outcomes are $\{H,T\}$ and that one picks twice from that set.
- Indeed, this viewpoint misses the relationship between the two flips.
- Each $\omega \in \Omega$ describes one outcome of the complete experiment.
- Ω and the probabilities specify the random experiment.

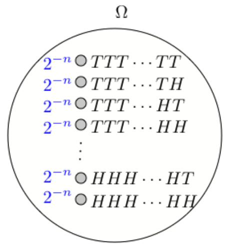
Flipping n times

Flip a fair coin n times (some $n \geq 1$):

- Possible outcomes: $\{TT\cdots T, TT\cdots H, \ldots, HH\cdots H\}$.

  Thus, 2^n possible outcomes.
- Note: $\{TT\cdots T, TT\cdots H, \ldots, HH\cdots H\} = \{H,T\}^n$.

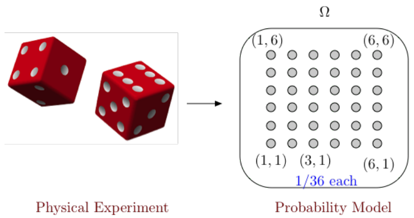
  $A^n := \{(a_1,\ldots,a_n) \mid a_1 \in A, \ldots, a_n \in A\}$. $|A^n| = |A|^n$.
- Likelihoods: $1/2^n$ each.



Roll two Dice

Roll a balanced 6-sided die twice:

- Possible outcomes: $\{1,2,3,4,5,6\}^2 = \{(a,b) \mid 1 \leq a,b \leq 6\}$.
- Likelihoods: $1/36$ for each.



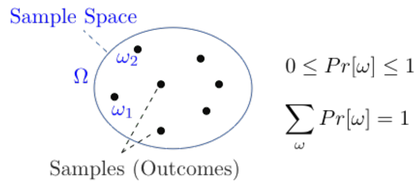
Probability Space.

1. A "random experiment":
   (a) Flip a biased coin;
   (b) Flip two fair coins;
   (c) Deal a poker hand.
2. A set of possible outcomes: Ω.
   (a) $\Omega = \{H,T\}$;
   (b) $\Omega = \{HH,HT,TH,TT\}$; $|\Omega| = 4$;
   (c) $\Omega = \{\ \underline{A\spadesuit\ A\diamondsuit\ A\clubsuit\ A\heartsuit\ K\spadesuit},\ \underline{A\spadesuit\ A\diamondsuit\ A\clubsuit\ A\heartsuit\ Q\spadesuit}, \ldots\}$
   $|\Omega| = \binom{52}{5}$.
3. Assign a probability to each outcome: $Pr: \Omega \to [0,1]$.
   (a) $Pr[H] = p, Pr[T] = 1-p$ for some $p \in [0,1]$
   (b) $Pr[HH] = Pr[HT] = Pr[TH] = Pr[TT] = \frac{1}{4}$
   (c) $Pr[\ \underline{A\spadesuit\ A\diamondsuit\ A\clubsuit\ A\heartsuit\ K\spadesuit}\ ] = \cdots = 1/\binom{52}{5}$

Probability Space: formalism.

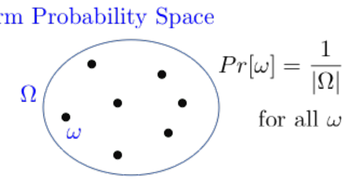
Ω is the **sample space.**
$\omega \in \Omega$ is a **sample point**. (Also called an **outcome**.)
Sample point ω has a probability $Pr[\omega]$ where

- $0 \leq Pr[\omega] \leq 1$;
- $\sum_{\omega\in\Omega} Pr[\omega] = 1$.



Probability Space: Formalism.

In a **uniform probability space** each outcome ω is equally probable: $Pr[\omega] = \frac{1}{|\Omega|}$ for all $\omega \in \Omega$.

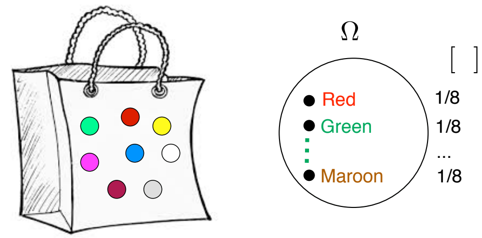


Examples:

- Flipping two fair coins, dealing a poker hand are uniform probability spaces.
- Flipping a biased coin is not a uniform probability space.

Probability Space: Formalism

Simplest physical model of a uniform probability space:



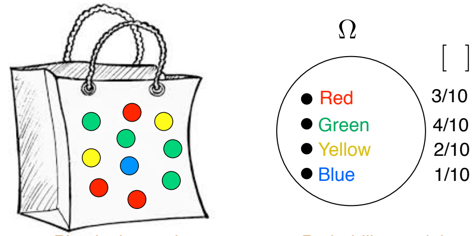
Physical experiment   Probability model

A bag of identical balls, except for their color (or a label). If the bag is well shaken, every ball is equally likely to be picked.

$$\Omega = \{\text{white, red, yellow, grey, purple, blue, maroon, green}\}$$
$$Pr[\text{blue}] = \frac{1}{8}.$$

Probability Space: Formalism

Simplest physical model of a non-uniform probability space:



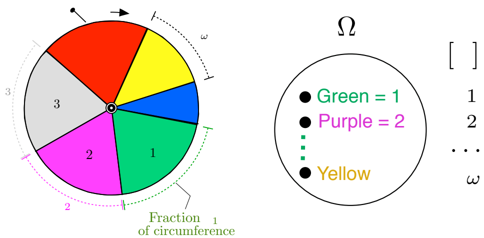
Physical experiment   Probability model

$$\Omega = \{\text{Red, Green, Yellow, Blue}\}$$
$$Pr[\text{Red}] = \frac{3}{10}, Pr[\text{Green}] = \frac{4}{10}, \text{ etc.}$$

Note: Probabilities are restricted to rational numbers: $\frac{N_k}{N}$.

Probability Space: Formalism

Physical model of a general non-uniform probability space:



Physical experiment   Probability model

The roulette wheel stops in sector ω with probability p_ω.

$$\Omega = \{1, 2, 3, \ldots, N\}, Pr[\omega] = p_\omega.$$

An important remark

- The random experiment selects one and only one outcome in Ω.
- For instance, when we flip a fair coin twice
  - $\Omega = \{HH, TH, HT, TT\}$
  - The experiment selects *one* of the elements of Ω.
- In this case, its wrong to think that $\Omega = \{H, T\}$ and that the experiment selects two outcomes.
- Why? Because this would not describe how the two coin flips are related to each other.
- For instance, say we glue the coins side-by-side so that they face up the same way. Then one gets HH or TT with probability 50% each. This is not captured by 'picking two outcomes.'

Lecture 15: Summary

Modeling Uncertainty: Probability Space

1. Random Experiment
2. Probability Space: $\Omega; Pr[\omega] \in [0,1]; \sum_\omega Pr[\omega] = 1$.
3. Uniform Probability Space: $Pr[\omega] = 1/|\Omega|$ for all $\omega \in \Omega$.