Probability
Announcements

• It is Friday. Story time!
• Sign up for sections by Sunday.
• Python Review Session today.
• We are going to make history today.
Sign up for Sections by Sunday noon
PSet FAQ

Q: Can I check my answers (many) times?
A: Yes. No penalty. We will know if you try all solutions.

Q: Will you just grade us on if the answers are correct?
A: No! You must explain your answer. Graded on style and correctness.

Q: How much explanation should we provide?
A: As detailed as the course reader worked examples

Q: Must I do all my work on the app?
A: No
**Enigma Machine**

**Two wires:** How many ways are there to place exactly two wires? Recall that wires are not considered distinct. Each letter can have at most one wire connected to it, thus you couldn’t have a wire connect ‘K’ to ‘L’ and another one connect ‘L’ to ‘X’

There are \( \binom{26}{2} \) ways to place the first wire and \( \binom{24}{2} \) ways to place the second wire. However, since the wires are indistinct, we have double counted every possibility. Because every possibility is counted twice we should divide by 2:

\[
\text{Total} = \frac{\binom{26}{2} \cdot \binom{24}{2}}{2} = 44,850
\]
PSet Status

Try and be here

About half the class has started
Python Review Session

Friday at 5pm PT with Ishira (online)

Find links, recordings, and setup here
Review
Counting Rules

Counting operations on $n$ objects

- Sort, order matters \{perms\}
  - Distinct: $n!$
  - Some Distinct: $\frac{n!}{n_1!n_2!\ldots}$

- Choose $k$ \{combinations\}
  - Distinct: $\binom{n}{k} = \frac{n!}{k!(n-k)!}$

- Put in $r$ buckets
  - Distinct: $r^n$
  - None Distinct: $\frac{(n + r - 1)!}{n!(r - 1)!}$
Counting Rules

Counting operations on \( n \) objects

- Sort, order matters \{perms\}
  \[ n! \]
- Some Distinct
  \[ \frac{n!}{n_1!n_2! \ldots} \]
- Choose \( k \) \{combinations\}
  \[ \binom{n}{k} = \frac{n!}{k!(n-k)!} \]
- Put in \( r \) buckets
  \[ r^n \]
- Distinct
  \[ (n + r - 1)! \]
- None Distinct
  \[ \frac{n!(r-1)!}{n!(r-1)!} \]
For a DNA tree we need to calculate the DNA distance between each pair of animals. How many calculations are needed?
Q: There are $n$ animals. How many distinct pairs of animals are there?
\[(\binom{n}{2})\]

(\binom{n}{2})
BRANCHED EVOLUTION
The genetic diversity in a tumour echoes Darwin’s Tree of Life.

Cancer starts with one cell mutating
End Review
Sample Space

- **Sample space**, $S$, is set of all possible outcomes of an experiment
  - Coin flip: $S = \{\text{Head, Tails}\}$
  - Flipping two coins: $S = \{\{H, H\}, \{H, T\}, \{T, H\}, \{T, T\}\}$
  - Roll of 6-sided die: $S = \{1, 2, 3, 4, 5, 6\}$
  - # emails in a day: $S = \{x \mid x \in \mathbb{Z}, x \geq 0\}$ (non-neg. ints)
  - YouTube hrs. in day: $S = \{x \mid x \in \mathbb{R}, 0 \leq x \leq 24\}$
Event Space

- **Event**, $E$, is some subset of $S$ \( \{E \subseteq S\} \)
  - Coin flip is heads: \( E = \{\text{Head}\} \)
  - $\geq 1$ head on 2 coin flips: \( E = \{\{H, H\}, \{H, T\}, \{T, H\}\} \)
  - Roll of die is 3 or less: \( E = \{1, 2, 3\} \)
  - # emails in a day $\leq 20$: \( E = \{x \mid x \in \mathbb{Z}, 0 \leq x \leq 20\} \)
  - Wasted day $\geq 5$ YT hrs.: \( E = \{x \mid x \in \mathbb{R}, 5 \leq x \leq 24\} \)

Note: When Ross uses: $\subset$, he really means: $\subseteq$
### Event Space

**Sample Space, S**

- **Coin flip**
  \[ S = \{\text{Heads, Tails}\} \]

- **Flipping two coins**
  \[ S = \{(H,H), (H,T), (T,H), (T,T)\} \]

- **Roll of 6-sided die**
  \[ S = \{1, 2, 3, 4, 5, 6\} \]

- **# emails in a day**
  \[ S = \{x \mid x \in \mathbb{Z}, x \geq 0\} \]

- **TikTok hours in a day**
  \[ S = \{x \mid x \in \mathbb{R}, 0 \leq x \leq 24\} \]

**Event, E**

- **Flip lands heads**
  \[ E = \{\text{Heads}\} \]

- **\geq 1 head on 2 coin flips**
  \[ E = \{(H,H), (H,T), (T,H)\} \]

- **Roll is 3 or less:**
  \[ E = \{1, 2, 3\} \]

- **Low email day (\leq 20 emails)**
  \[ E = \{x \mid x \in \mathbb{Z}, 0 \leq x \leq 20\} \]

- **Wasted day (\geq 5 TT hours):**
  \[ E = \{x \mid x \in \mathbb{R}, 5 \leq x \leq 24\} \]
What is a probability?
[suspense]
Number between 0 and 1
A number to which we ascribe meaning

\[ P(E) \]

* Our belief that an event \( E \) occurs
A number to which we ascribe meaning

* Our belief that an event $E$ occurs
What is a Probability?

\[ P(E) = \lim_{n \to \infty} \frac{n(E)}{n} \]
What is a Probability?

\[ P(E) = \lim_{n \to \infty} \frac{n(E)}{n} \]

\( n \) is the number of trails

The “event” \( E \) is that you hit the target

\( P\{E\} \approx \)
What is a Probability?

\[ P(E) = \lim_{n \to \infty} \frac{n(E)}{n} \]

The “event” \( E \) is that you hit the target.

\( n \) is the number of trails.

\[ P\{E\} \approx 0.00 \]
What is a Probability?

\[ P(E) = \lim_{n \to \infty} \frac{n(E)}{n} \]

The "event" \( E \) is that you hit the target.

\( n \) is the number of trails

\[ P\{E\} \approx 0.50 \]
What is a Probability?

\[ P(E) = \lim_{n \to \infty} \frac{n(E)}{n} \]

The “event” \( E \) is that you hit the target.

Hit: 2
Thrown: 3

\[ P\{E\} \approx 0.75 \]

\( n \) is the number of trails
What is a Probability?

\[ P(E) = \lim_{{n \to \infty}} \frac{n(E)}{n} \]

$n$ is the number of trials

Hit: 11
Thrown: 24

The “event” $E$ is that you hit the target

\[ P\{E\} \approx 0.46 \]
Probability from Analytic Solutions
Axioms of Probability

Recall: $S = \text{all possible outcomes. } E = \text{the event.}$

• Axiom 1: $0 \leq P(E) \leq 1$

• Axiom 2: $P(S) = 1$

• Axiom 3: If events $E$ and $F$ are mutually exclusive:

\[
P(E \cup F) = P(E) + P(F)
\]
Core Rules of Probability

Recall: $S = \text{all possible outcomes. } E = \text{the event.}$

- Axiom 1: $0 \leq P(E) \leq 1$

- Axiom 2: $P(S) = 1$

- Identity 3: $P(E^c) = 1 - P(E)$

Technically Identity 3 can be proved from the 3 axioms.
Special Case of Analytic Probability
Equally Likely Outcomes
Equally Likely Outcomes

Some sample spaces have equally likely outcomes.

- Coin flip: \( S = \{\text{Head, Tails}\} \)
- Flipping two coins: \( S = \{\{\text{H, H}\}, \{\text{H, T}\}, \{\text{T, H}\}, \{\text{T, T}\}\} \)
- Roll of 6-sided die: \( S = \{1, 2, 3, 4, 5, 6\} \)

If we have equally likely outcomes, then

\[
P(\text{Each outcome}) = \frac{1}{|S|}
\]

Therefore

\[
P(E) = \frac{\# \text{ outcomes in } E}{\# \text{ outcomes in } S} = \frac{|E|}{|S|} \quad \{\text{by Axiom 3}\}
\]
Not Everything is Equally Likely

• Play lottery.
  ▪ What is $P\{\text{Win}\}$?

• $S = \{\text{Lose, Win}\}$
• $E = \{\text{Win}\}$
• $P\{\text{Win}\} = \frac{|E|}{|S|} = \frac{1}{2} = 50\%$
Sum of Two Die = 7?

Roll two 6-sided dice. What is $P[\text{sum} = 7]$?

$S = \{ [1,1], [1,2], [1,3], [1,4], [1,5], [1,6], [2,1], [2,2], [2,3], [2,4], [2,5], [2,6], [3,1], [3,2], [3,3], [3,4], [3,5], [3,6], [4,1], [4,2], [4,3], [4,4], [4,5], [4,6], [5,1], [5,2], [5,3], [5,4], [5,5], [5,6], [6,1], [6,2], [6,3], [6,4], [6,5], [6,6] \}$
Sum of Two Die = 7?

Roll two 6-sided dice. What is probability the sum = 7?
Let \( E \) be the event that the sum is 7

\[
S = \{ [1,1], [1,2], [1,3], [1,4], [1,5], [1,6], [2,1], [2,2], [2,3], [2,4], [2,5], [2,6], [3,1], [3,2], [3,3], [3,4], [3,5], [3,6], [4,1], [4,2], [4,3], [4,4], [4,5], [4,6], [5,1], [5,2], [5,3], [5,4], [5,5], [5,6], [6,1], [6,2], [6,3], [6,4], [6,5], [6,6] \}
\]

\( E = \text{in blue} \)

\[
P(E) = \frac{|E|}{|S|} = \frac{6}{36} = 0.166
\]
Is it correct?

\[ P(E) = \frac{|E|}{|S|} = \frac{6}{36} = 0.166 \]
Sum of Two Die = 7?

$$P(E) = \frac{|E|}{|S|} = \frac{6}{36} = 0.166$$

```python
1    import random
2    from tqdm import tqdm
3
4    N_TRIALS = 10000000  # getting close to infinity
5    TARGET_SUM = 7       # do the two dice sum to 6?
6
7    def main():
8        n_events = 0
9    for i in tqdm(range(N_TRIALS)):
10       dice_total = run_experiment()
11       if dice_total == TARGET_SUM:
12          n_events += 1
13    pr_e = n_events / N_TRIALS
14    print(f"after {N_TRIALS} trials")
15    print(f"P(E) = {pr_e:.3f}", pr_e)
16
17    def run_experiment():
18        d_1 = roll_dice()
19        d_2 = roll_dice()
20        return d_1 + d_2
21
22    def roll_dice():
23        # give me a random dice roll
24        # alternatively random.randint(1, 7)
25        return random.choice([1,2,3,4,5,6])
26
27    if __name__ == '__main__':
28        # this starts the program in main
29        main()
```
Sum of Two Die = 2?

Roll two 6-sided dice. What is probability the sum = 7?

Let E be the event that the sum is 7

\[ S = \{ [1,1], [1,2], [1,3], [1,4], [1,5], [1,6], [2,1], [2,2], [2,3], [2,4], [2,5], [2,6], [3,1], [3,2], [3,3], [3,4], [3,5], [3,6], [4,1], [4,2], [4,3], [4,4], [4,5], [4,6], [5,1], [5,2], [5,3], [5,4], [5,5], [5,6], [6,1], [6,2], [6,3], [6,4], [6,5], [6,6] \} \]

E =
Sum of Two Die = 2?

Roll two 6-sided dice. What is probability the sum = 2?
Let E be the event that the sum is 2

\[ S = \{ [1,1], [1,2], [1,3], [1,4], [1,5], [1,6], [2,1], [2,2], [2,3], [2,4], [2,5], [2,6], [3,1], [3,2], [3,3], [3,4], [3,5], [3,6], [4,1], [4,2], [4,3], [4,4], [4,5], [4,6], [5,1], [5,2], [5,3], [5,4], [5,5], [5,6], [6,1], [6,2], [6,3], [6,4], [6,5], [6,6] \} \]

\[ E = \text{in red} \]

\[ P(E) = \frac{|E|}{|S|} = \frac{1}{36} = 0.027 \]
Other ways to make a Sample Space?

Each outcome

Value dice 1

Value dice 2
Sum of Two Die: Three options for the sample space

Think of the die as **distinct**

\[
\begin{bmatrix}
5 \\
\end{bmatrix},
\begin{bmatrix}
5 \\
\end{bmatrix}
\]

Think of the die as **indistinct**

\[
\left\{ \begin{array}{c}
5 \\
5
\end{array} \right\}
\]

Just look at the sum

\[
10
\]
Sum of Two Die = 7? Bug: Die are Indistinct

Roll two 6-sided dice. What is probability the sum = 7?
Let E be the event that the sum is 7

\[ S = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\} \]

Each outcome

Just look at the sum

\[ E = \text{in red} \]
\[ P(E) = \frac{|E|}{|S|} = \frac{1}{12} = 0.0833\ldots \]
Sum of Two Die: Three options for the sample space

Think of the die as **distinct**

Think of the die as **indistinct**

Just look at the sum

Value of a dice
Sum of Two Die = 7? Bug: Die are Indistinct

Roll two 6-sided dice. What is \( P(\text{sum} = 7) \)?

\[
S = \{\{1,1\}, \{1,2\}, \{1,3\}, \{1,4\}, \{1,5\}, \{1,6\}, \{2,2\}, \{2,3\}, \{2,4\}, \{2,5\}, \{2,6\}, \{3,3\}, \{3,4\}, \{3,5\}, \{3,6\}, \{4,4\}, \{4,5\}, \{4,6\}, \{5,5\}, \{5,6\}, \{6,6\}\}
\]

\[
E = \text{in blue}
\]

\[
P(E) = \frac{|E|}{|S|} = \frac{3}{20} = 0.15\text{?}
\]
Sum of Two Die: Three options for the sample space

Think of the die as **distinct**

\[
\begin{bmatrix}
5 \\
, \\
5
\end{bmatrix}
\]

Value of a dice

Think of the die as **indistinct**

\[
\{ 5, 5 \}
\]

Value of a dice

Just look at the sum

10

Think of the die as distinct

Just look at the sum

---

Piech, CS109, 2021
pigs and cows

- 4 cows and 3 pigs in a toy box. 3 drawn.
  - What is $P(1 \text{ cow and 2 pigs drawn})$?

Equally likely sample space? Thought experiment

4 cows

3 pigs
The Choice of Sample Space is Yours!

<table>
<thead>
<tr>
<th>Unordered</th>
<th>Distinct</th>
<th>Indistinct</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>{C₁, P₂, P₃}</td>
<td>{2 cows, 1 pig}</td>
</tr>
<tr>
<td></td>
<td>{C₁, C₂, C₃}</td>
<td>{3 cows}</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Ordered</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>[C₁, P₂, P₃]</td>
<td>[cow, pig, cow]</td>
</tr>
<tr>
<td></td>
<td>[C₁, C₂, C₃]</td>
<td>[cow, cow, cow]</td>
</tr>
</tbody>
</table>

Which choice will lead to equally likely outcomes?
4 pigs and 3 cows in a Bag. 3 drawn.

- What is \( P(1 \text{ pig and 2 cows drawn})? \)

- Ordered and Distinct:
  - Pick 3 ordered items: \( |S| = 7 \times 6 \times 5 = 210 \)
  - Pick pig as either 1st, 2nd, or 3rd item:
    \[
    |E| = \{4 \times 3 \times 2\} + \{3 \times 4 \times 2\} + \{3 \times 2 \times 4\} = 72
    \]
  - \( P(1 \text{ pig, 2 cows}) = \frac{72}{210} = \frac{12}{35} \)

- Unordered:
  - \( |S| = \binom{7}{3} = 35 \)
  - \( |E| = \binom{4}{1}\binom{3}{2} = 12 \)
  - \( P(1 \text{ pig, 2 cows}) = \frac{12}{35} \)
Make indistinct items **distinct** to get equally likely sample space outcomes

*You will need to use this “trick” with high probability*
• Consider 5 card poker hands.
  ▪ “straight” is 5 consecutive rank cards of any suit
  ▪ What is P(straight)?
Consider 5 card poker hands.

- “straight” is 5 consecutive rank cards of any suit
- What is $P(\text{straight})$?

\[
|S| = \binom{52}{5}
\]

\[
|E| = 10 \cdot \binom{4}{1}^5
\]

\[
P(\text{straight}) = \frac{|E|}{|S|} = \frac{10 \cdot \binom{4}{1}^5}{\binom{52}{5}} \approx 0.00394
\]
Straight Poker Hand

- Consider 5 card poker hands.
  - “straight” is 5 consecutive rank cards of any suit
  - “straight flush” is 5 consecutive rank cards of same suit
  - What is $P($straight, but not straight flush$)$?

\[
|S| = \binom{52}{5}
\]
\[
|E| = 10 \binom{4}{1}^5 - 10 \binom{4}{1}
\]

\[
P($straight$) = \frac{|E|}{|S|} = \frac{10 \binom{4}{1}^5 - 10 \binom{4}{1}}{\binom{52}{5}} \approx 0.00392
\]
When approaching an “equally likely probability” problem, start by defining sample spaces and event spaces.
Chip Defect Detection

- $n$ chips manufactured, 1 of which is defective.
- $k$ chips randomly selected from $n$ for testing.
  - What is $P\{\text{defective chip is in } k \text{ selected chips}\}$?

- $|S| = \binom{n}{k}$

- $|E| = \binom{1}{1} \left( \frac{n-1}{k-1} \right)$

- $P(\text{defective chip is in } k \text{ selected chips})$
  \[
  = \frac{\binom{1}{1} \left( \frac{n-1}{k-1} \right)}{\binom{n}{k}} = \frac{(n-1)!}{(k-1)!(n-k)!} = \frac{k}{n}
  \]
Target Revisited

Screen size = 800x800
Radius of target = 200

The dart is equally likely to land anywhere on the screen.

What is the probability of hitting the target?

\[
|S| = 800^2
|E| = \pi 200^2
\]

\[
p(E) = \frac{\pi \cdot 200^2}{800^2} \approx 0.1963
\]
Target Revisited

Screen size = 800x800
Radius of target = 200

The dart is equally likely to land anywhere on the screen.

What is the probability of hitting the target?

\[ |S| = 800^2 \]
\[ |E| = \pi 200^2 \]
\[ p(E) = \frac{\pi \cdot 200^2}{800^2} \approx 0.1963 \]
Let it find you.

SERENDIPITY

the effect by which one accidentally stumbles upon something truly wonderful, especially while looking for something entirely unrelated.
WHEN YOU MEET YOUR BEST FRIEND

Somewhere you didn't expect to.
Serendipity

• Say the population of Stanford is 17,000 people
  ▪ You are friends with ?
  ▪ Walk into a room, see 268 random people.
  ▪ What is the probability that you see someone you know?
  ▪ Assume you are equally likely to see each person at Stanford

The Travels and Adventures of SERENDIPITY
Many times it is easier to calculate $P(E^C)$.
Back to Axiom 3
Axioms of Probability

Recall: $S = \text{all possible outcomes. } E = \text{the event.}$

- **Axiom 1:** $0 \leq P(E) \leq 1$

- **Axiom 2:** $P(S) = 1$

- **Axiom 3:** If events $E$ and $F$ are mutually exclusive:

  $$P(E \cup F) = P(E) + P(F)$$
Mutually Exclusive Events

If events are mutually exclusive, probability of OR is simple:

\[ P(E \cup F) = P(E) + P(F) \]
If events are mutually exclusive, probability of OR is simple:

$$P(E \cup F) = \frac{7}{50} + \frac{4}{5} = \frac{11}{50}$$
Probability of "or"

\[ P(X_1 \cup X_2 \cup \cdots \cup X_n) = \sum_{i=1}^{n} P(X_i) \]

Wahoo! All my events are mutually exclusive.
If events are *mutually exclusive* probability of OR is easy!
\[ P(E^c) = 1 - P(E) ? \]

\[ P(E \cup E^c) = P(E) + P(E^c) \quad \text{Axiom 3. Since } E \text{ and } E^c \text{ are mutually exclusive} \]

\[ P(S) = P(E) + P(E^c) \quad \text{Since everything must either be in } E \text{ or } E^c \]

\[ 1 = P(E) + P(E^c) \quad \text{Axiom 2} \]

\[ P(E^c) = 1 - P(E) \quad \text{Rearrange} \]
Trailing the dovetail shuffle to it’s lair – Persi Diaconosis
Trailing the Dovetail Shuffle to Its Lair

By Dave Bayer and Persi Diaconis

Columbia University and Harvard University

We analyze the most commonly used method for shuffling cards. The main result is a simple expression for the chance of any arrangement after any number of shuffles. This is used to give sharp bounds on the approach to randomness: n log_n 2 shuffles are necessary and sufficient to mix up n cards.

See ingredients are the analysis of a card trick and the determination of the identities of a natural commutation subspecies in the symmetric group algebras.

1. Introduction. The dovetail shuffle is the most commonly used method of shuffling cards. Roughly, a deck of cards is cut about in half and then the two halves are riffled together. Figure 1 gives an example of a shuffle of a deck of 13 cards.

A mathematically precise model of shuffling was introduced by Gilbert and Shannon [see Gilbert (1955)] and independently by Reeds (1981). A deck of n cards is cut into two portions according to a binomial distribution; thus, the chance that k cards are cut off is \( \binom{n}{k} 2^{-n} \) for 0 ≤ k ≤ n. The two packets are then riffled together in such a way that cards drop from the left or right heaps with probability proportional to the number of cards in each heap. Thus, if there are A and B cards remaining in the left and right heaps, then the chance that the next card will drop from the left heap is A/(A+B). Such shuffles are easily described backwards: Each card has an equal and independent chance of being pulled back into the left or right heap. An inverse riffle shuffle is illustrated in Figure 2.

Experiments reported in Diaconis (1988) show that the Gilbert–Shannon–Reeds GSRD model is a good description of the way real people shuffle real cards. It is natural to ask how many times a deck must be shuffled to mix it up. In Section 3 we prove:

**Theorem 1.** If n cards are shuffled m times, then the chance that the deck is in an arrangement \( \pi \) is \( \left( \frac{2}{n} \right)^m \times \frac{r}{2^{2m}} \), where r is the number of rising sequences in \( \pi \).

Rising sequences are defined and illustrated in Section 2 through the analysis of a card trick. Section 3 develops several equivalent interpretations of the GSR distribution for riffle shuffles, including a geometric description as the motion of a point randomly dropped at random into the unit interval under the baker's transformation \( x \rightarrow 2x \mod 1 \). This leads to a proof of Theorem 1.

Section 3 also relates shuffling to some developments in algebra. A permutation \( \pi \) has a descent at \( i \) if \( \pi(i) > \pi(i + 1) \). A permutation \( \pi \) has \( r \) rising sequences if and only if \( \pi^{-1} \) has \( r - 1 \) descents. Let

\[
A_r = \sum_{\pi \text{ has } r \text{ descents}} \pi
\]

Theorem 2. An inverse riffle shuffle. (a) We begin with a sorted deck. (b) Each card is moved one way or the other uniformly at random. In "tail apart" a riffle shuffle and retrieve two packets. (c) The two packets are placed in sequence. (d) The two packets are still be identified in the shuffled deck; they are separated by a "descent" in the first ether. This shuffle is the inverse of the shuffle diagrammed in Figure 1.
What is the probability that in the $n$ shuffles seen since the start of time, yours is unique?

- $|S| = (52!)^n$
- $|E| = (52! - 1)^n$
- $P\{\text{no deck matching yours}\} = (52! - 1)^n / (52!)^n$

For $n = 10^{20}$,
- $P\{\text{deck matching yours}\} < 0.000000001$

* Assume 7 billion people have been shuffling cards once a second since cards were invented