terminology: ① probability ② experiment ③ sample space ④ event ⑤ conditional probability ⑥ independent

Generalized Permutations/Combinations

How many anagrams are there of **MISSISSIPPI**?

Method 1: Place groups of letters

Stage 1: place the **M**. $\binom{11}{1}$ choices.

Stage 2: place the **I**s. $\binom{10}{4}$ choices

Stage 3: place the **S**s. $\binom{6}{4}$ choices

Stage 4: place the **P**s. $\binom{2}{2}$ choices

One answer: $\binom{11}{1}\cdot\binom{10}{4}\cdot\binom{6}{4}\cdot\binom{2}{2}$

Method 2: Division rule

There are 11! ways to place the letters initially

There are 4! ways to scramble the I's, which doesn't change the permutation
$I_1SSI_2PPI_3SSI_4M$
$= I_3SSI_4PPI_2SSI_1M$

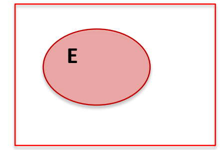
There are 4! ways to scramble the S's, which doesn't change the permutation

There are 2! ways to scramble the P's, which doesn't change the permutation

Equivalent Answer: $\frac{11!}{4!4!2!}$

Lec 23: Discrete Probability -- ANSWERS

- **Experiment**: Procedure that yields an outcome
- **Sample space**: Set of all possible outcomes
- **Event**: subset of the sample space



If S is a sample space of **equally likely** outcomes, the **probability of an event** E is

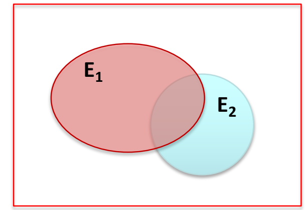
① $$p(E) = \frac{|E|}{|S|}$$

If S consists of **unequally likely** outcomes, then

② $$p(E) = \sum_{s\in E} p(s)$$

Fun facts:

- $0 \le p(E) \le 1$ ③
- $p(E) + p(\bar{E}) = 1$ ④
- $p(\bar{E}) = 1 - p(E)$ ⑤ ☆
- $p(E_1 \cup E_2) = p(E_1) + p(E_2) - p(E_1 \cap E_2)$



Example: 2 Dice

- Roll two dice. Find the probability that the sum is 7.

What could we use as our sample space, S?

Option 1: Let S = set of possible sums. |S| = 11

Option 2: Let S = set of unordered pairs. |S| = 21

Option 3: Let S = set of ordered pairs. |S| = 36

valid sample spaces, but not very **useful**, because outcomes are not equally likely

outcomes are equally likely

Which sample space is easiest to use, and why?

Option 3: S = set of ordered pairs is best

because that sample space contains ***equally likely outcomes***, which means that we can use the easy probability formula of p(E) = |E| / |S|.

Using the "easiest" sample space, find the probability that the sum is 7.

$S = \{(1,1) \dots (6,6)\}$

E = {(1,6), (2,5), (3,4), (4,3), (5,2), (6,1)}

p(E) = |E|/|S| = 6/36 = 1/6

Conditional Probability

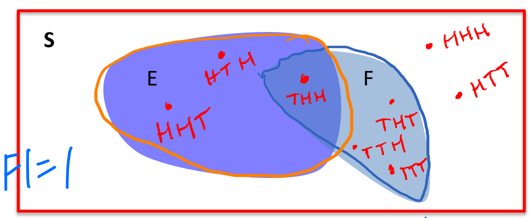
if S has equally likely outcomes

The **conditional probability** of event E given event F is

$$p(E\,|\,F) = \frac{p(E\cap F)}{p(F)} = \frac{|E\cap F|}{|F|}$$

$|S| = \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}$

$|E|=3, |F|=4, |E\cap F|=1$



Example: Flip a coin 3 times.

- E = total of two Heads
- F = first flip was Tails

Find p(E), p(F), p(E ∩ F), p(E|F), p(F|E)

$p(E) = \frac{|E|}{|S|} = \frac{3}{8}$

$p(F) = 4/8 = 1/2$

$p(E\cap F) = 1/8$

$p(E|F) = \frac{|E\cap F|}{|F|} = \frac{1}{4}$

$p(F|E) = \frac{|E\cap F|}{|E|} = \frac{1}{3}$

Conditional Probability Blitz (additional practice problems)

You roll a fair **red** and a **blue** 4-sided dice, and sum up their values.

- What is $\Pr[\text{sum} = 5]$?
- What is $\Pr[\text{sum} = 6]$?
- What is $\Pr[$**red die** rolls an even #$]$?

Event: sum=5: $E_5 = \{(1,4),(2,3),(3,2),(4,1)\}$, $|E_5| = 4$ $\Pr[E_5] = \frac{4}{16} = \frac{1}{4}$

Event: sum=6: $E_6 = \{(2,4),(3,3),(4,2)\}$, $|E_6| = 3$ $\Pr[E_6] = \frac{3}{16}$

Event: red even: $R = \{(2,1),(2,2),(2,3),(2,4),(4,1),(4,2),(4,3),(4,4)\}$,

$|R| = 8$ $\Pr[R] = \frac{8}{16} = \frac{1}{2}$

- What is $\Pr[\text{sum} = 5 \mid$ **red die** rolls an even #$]$?
- What is $\Pr[$**red die** rolls an even # $\mid \text{sum} = 5]$?

- What is $\Pr[\text{sum} = 6 \mid$ **red die** rolls an even #$]$?
- What is $\Pr[$**red die** rolls an even # $\mid \text{sum} = 6]$?

$R = \{(2,1),(2,2),(\mathbf{2},\mathbf{3}),(2,4),(\mathbf{4},\mathbf{1}),(4,2),(4,3),(4,4)\}$

$\Pr[E_5 \mid R] = \frac{2}{8} = \frac{1}{4}$

$E_5 = \{(1,4),(\mathbf{2},\mathbf{3}),(3,2),(\mathbf{4},\mathbf{1})\}$

$\Pr[R \mid E_5] = \frac{1}{2}$

$R = \{(2,1),(2,2),(2,3),(\mathbf{2},\mathbf{4}),(4,1),(\mathbf{4},\mathbf{2}),(4,3),(4,4)$

$\Pr[E_6 \mid R] = \frac{2}{8} = \frac{1}{4}$

$E_6 = \{(\mathbf{2},\mathbf{4}),(3,3),(\mathbf{4},\mathbf{2})\}$ $\Pr[R \mid E_6] = \frac{2}{3}$

Independence of Events

Events E and F are **independent** if and only if any/all of the following equivalent conditions hold:

$$\Pr(E|F) = \Pr(E)$$
$$\Pr(F|E) = \Pr(F)$$
$$\Pr(E \cap F) = \Pr(E) \cdot \Pr(F)$$

one is true ⇒ the other twos are true.

Are these independent?

- Roll two dice
  - E: the sum of the two dice is 5
  - F: the first die is a 1
  -

E = {(1,4), (2,3), (3,2), (4,1)},
F = {(1,1),(1,2),(1,3),(1,4),(1,5),(1,6)},
p(E) = 4/36 = 1/9, p(F) = 1/6
*p(E|F) = 1/6 ≠ p(E) **Not Independent**
*Or via: p($E \cap F$) = 1/36 ≠ 1/9 * 1/6 = p(E)p(F)

- Roll two dice
  - E: the sum of the two dice is 7
  - F: the first die is a 1

p(E) = 6/36 = 1/6
p(F) = 1/6
* p(E|F) = 1/6 = p(E) **Independent**
*Or via: p($E \cap F$) = 1/36 = 1/6 * 1/6 = p(E)p(F)

Birthday Problem

To simplify: assume 366 possible birthdays (including leap day), **all equally likely.**

How many people have to be in a room so that the probability that two people share a birthday is > 50%?

Solution: **Strategy**: compute p_n= probability that **all birthdays are distinct**.
Then $\Pr[\text{two people share a birthday}] = 1 - p_n$

- Sample space S: given n people (ordered), # ways to assign them all birthdays
  - *Product rule*: 366 choices for 1st person, 366 choices for 2nd person, etc. $|S| = 366^n$
- Event: given n people (ordered), # ways to assign them **distinct** birthdays.
  - *Product rule*: 366 choices for 1st person, 365 choices for 2nd person, 364 choices for 3rd person, etc. So $|E| = P(366, n)$

$$p_n = \frac{|E|}{|S|} = \frac{P(366,n)}{366^n}$$

P (No one shares the same birthday)

For what value of n is $1 - p_n > \frac{1}{2}$?

$\boxed{n = 23,\ 1 - p_n \approx 0.506}$ **Only 23 people!**