CSC380: Principles of Data Science

Probability 1

Xinchen Yu

Outline

- What is probability?

- Events

- Calculating probabilities

What is probability?

What is probability?

- Suppose I flip a coin, What is the probability it will come up heads? Most people say 50%, but why?

- "Nolan's new movie is coming out next weekend! There's a 100% chance you're going to love it."

Interpreting probabilities

Basically two different ways to interpret:
- Objective probability
  - based on logical analysis or long-run frequency of an event. It's derived from known facts, symmetry, or repeated experiments.

- Subjective probability
  - based on personal belief, opinion, or information about how likely an event is, especially when there's uncertainty or limited data.

Objective or Subjective?

It is a subjective probability: a belief based on their perception of how rare or meaningful the coincidence is, not a calculation based on statistical data.

Subjective probability

- Probabilities aren't in the world itself; they're in our knowledge/beliefs about the world.

- Can assign a probability to the truth of any statement that I have a degree of belief about.

We will focus on objective probability in this class.

Objective probability

- The probability of an event represents the long run proportion of the time the event occurs under repeated, controlled experimentation.
  - e.g. 0001110100111101000110
- Famous experiments in history on coin tosses

Experimenter	# Tosses	# Heads	Half # Tosses
De Morgan	4092	2048	2046
Buffon	4040	2048	2020
Feller	10000	4979	5000
Pearson	24000	12012	12000

Outcome, Event and Probability

Outcome

- Outcome is a **single** result/observation of a **random experiment**.

- Experiment 1: You flip a coin once.
  - The outcomes are: "Heads" or "Tails"

- Experiment 2: You roll a 6-sided die.
  - The outcomes are: 1, or 2, or 3, or 4, or 5, or 6

- Experiment 3: You tap "shuffle play" on your favorite Spotify playlist with 100 songs and the first song played.
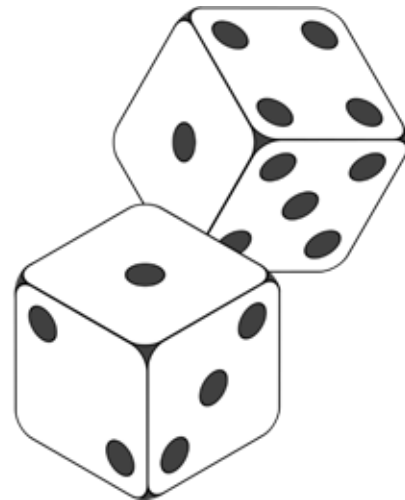 - The outcomes are: the 1st song in the list, or the 2nd song, ….

Suppose we roll <u>two fair dice</u>…

Suppose we roll <u>two fair dice</u>…

- ◆ What are the possible outcomes?
- ◆ What is the *probability* of the following:
  - ◆ rolling **even** numbers?
  - ◆ having two numbers sum to 6?
  - ◆ If one die rolls 1, the second die also rolling 1?

*…this is a **random process***.

How to formalize all these quantitatively?

The Sample Space

- The set of **all** possible outcomes of a random experiment is called the sample space, written as S.

- In math, the standard notation for a set is to write the individual members in curly braces:
    - S = {Outcome1, Outcome2, . . . , }

- Useful to visualize the sample space with an actual space.

The Sample Space
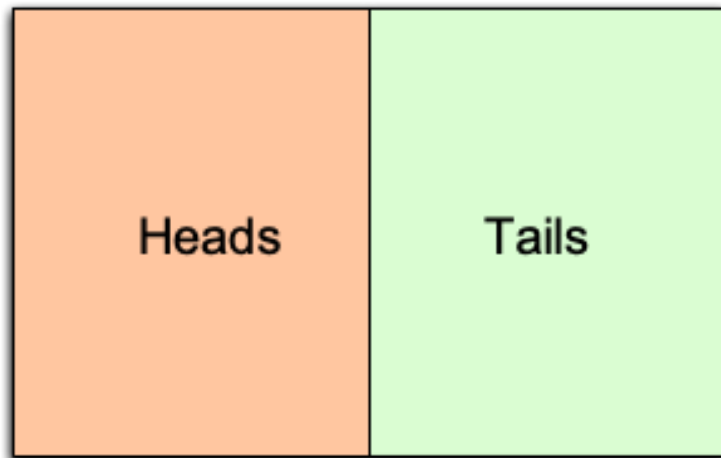

Figure: Visualization of a Sample Space

Examples of Sample Spaces

What's the sample space for a single coin flip?

- S = {Heads, Tails}

Examples of Sample Spaces

What is the sample space of rolling a die?

- S = {1, 2, 3, 4, 5, 6}

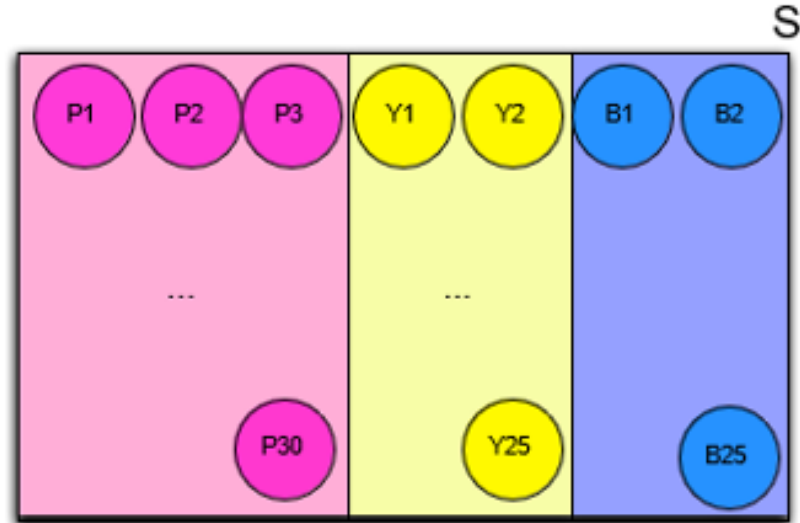|		
1	2	3
4	5	6

Examples of Sample Spaces

What is the sample space of drawing a ball out of a box containing 30 pink, 25 yellow, and 25 blue balls?

- S = {P1, P2, …, P30, Y1, …, Y25, B1, …, B25}

Examples of Sample Spaces

What's the sample space for...

- Randomly choosing a student from UA?
  - S = {Aarhus, Amaral, Balkan, . . . , Yao, Zielinski}

- Flipping two different coins?
  - S = {HH, HT, TH, TT}

- Flipping one coin twice?
  - S = {HH, HT, TH, TT}

- Observing the number of earthquakes in San Francisco in a particular year?
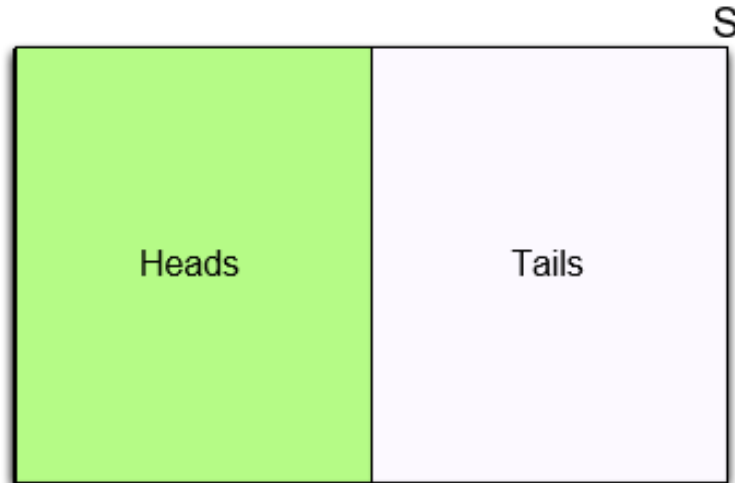 - S = {0, 1, 2, 3, . . . }

Events

- An event E is a **subset** of the sample space.

- An event E is a **set** of outcomes
    - When we make a particular observation, it is either "in" E or not.
    - Helpful to think about events as propositions (TRUE/FALSE): TRUE when the outcome is among the elements of the event set, and FALSE otherwise.
    - Is 4 in event E = {2, 4, 6}? ✅ YES → the **proposition is TRUE**
    - Is 4 in event F = {1, 3, 5}? ❌ NO → the **proposition is FALSE**

Examples of Events

What's the event set corresponding to the following propositions?
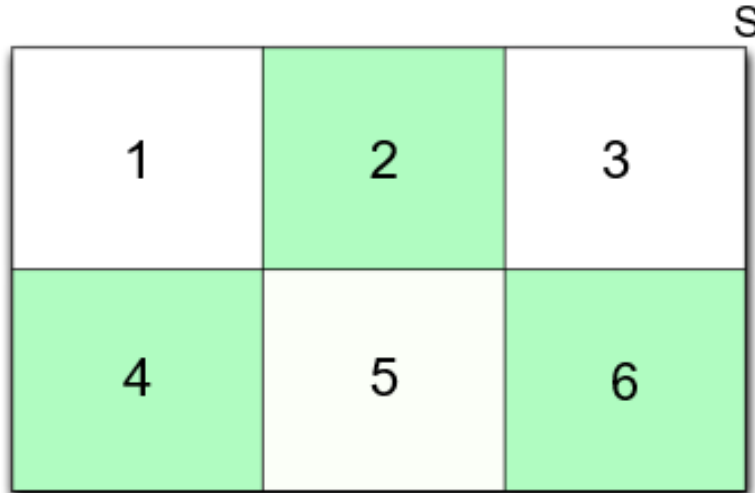
- "The coin comes up heads"
- E = {Heads}

Examples of Events

What's the event set corresponding to the following propositions?
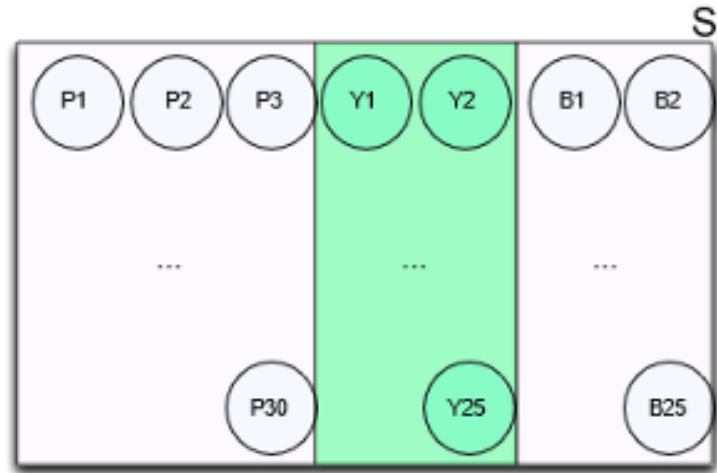
- "The die comes up an even number"
- E = {2, 4, 6}

Examples of Events

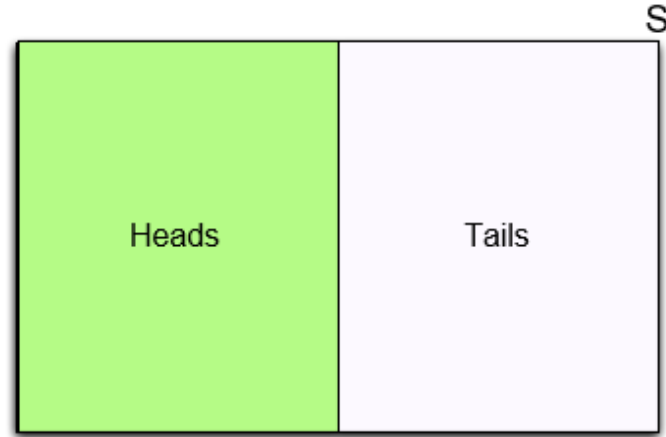What's the event set corresponding to the following propositions?

- "A yellow ball is chosen"
- E = {Y1, Y2, . . . , Y25}

Calculating Probabilities

Calculating probability

- We can think of the probability of an event E as its area, where S (sample space) always has a total area of 1.0

- So, the probability of E is the fraction of S that it takes up.
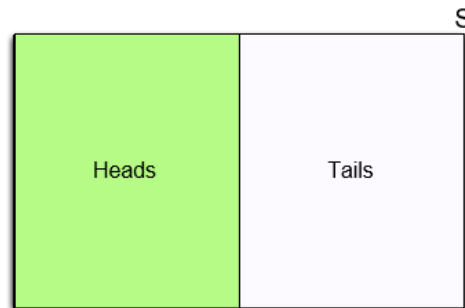

S

Heads | Tails

Calculating probability using symmetry

- If we have a sample space for which every outcome is equally likely (has the same area), then we can find event probabilities easily.

- Classic probability model to get the probability of an event E:

$$P(E) = \frac{\#\text{outcomes in } E}{\#\text{outcomes in } S}$$

S

| Heads | Tails |

Probability as Area

What is the probability of

$$P(E) = \frac{\#\text{outcomes in } E}{\#\text{outcomes in } S}$$

- Rolling a fair die and see an even number?
- E = {2,4,6}
- $P(E) = \dfrac{\#\{2,4,6\}}{\#\{1,2,3,4,5,6\}} = \dfrac{3}{6} = \dfrac{1}{2}$

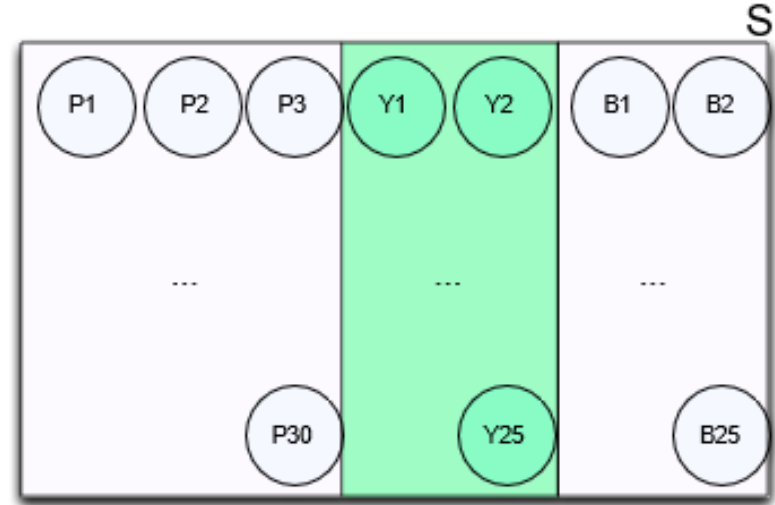| 1	2 P(2) = 1/6	3
4 P(4) = 1/6	5	6 P(6) = 1/6

P(S) = 1

P(Even) = P(2) + P(4) + P(6) = 3/6

Probability as Area

What is the probability of

- Selecting a yellow ball?
- E = {Y1, Y2, . . . , Y25}
- $P(E) = \frac{\# E}{\# S} = \frac{25}{30+25+25} = \frac{5}{16}$
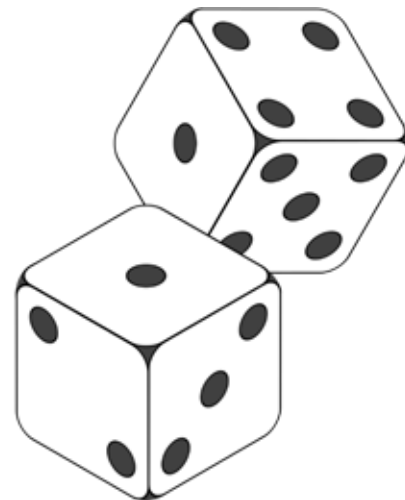

P(Yellow) = P(Y1) + ... + P(Y25) = 25 * (1/80)

- Suppose we throw two fair dice
  - What is the sample space S (space of all possible outcomes)?

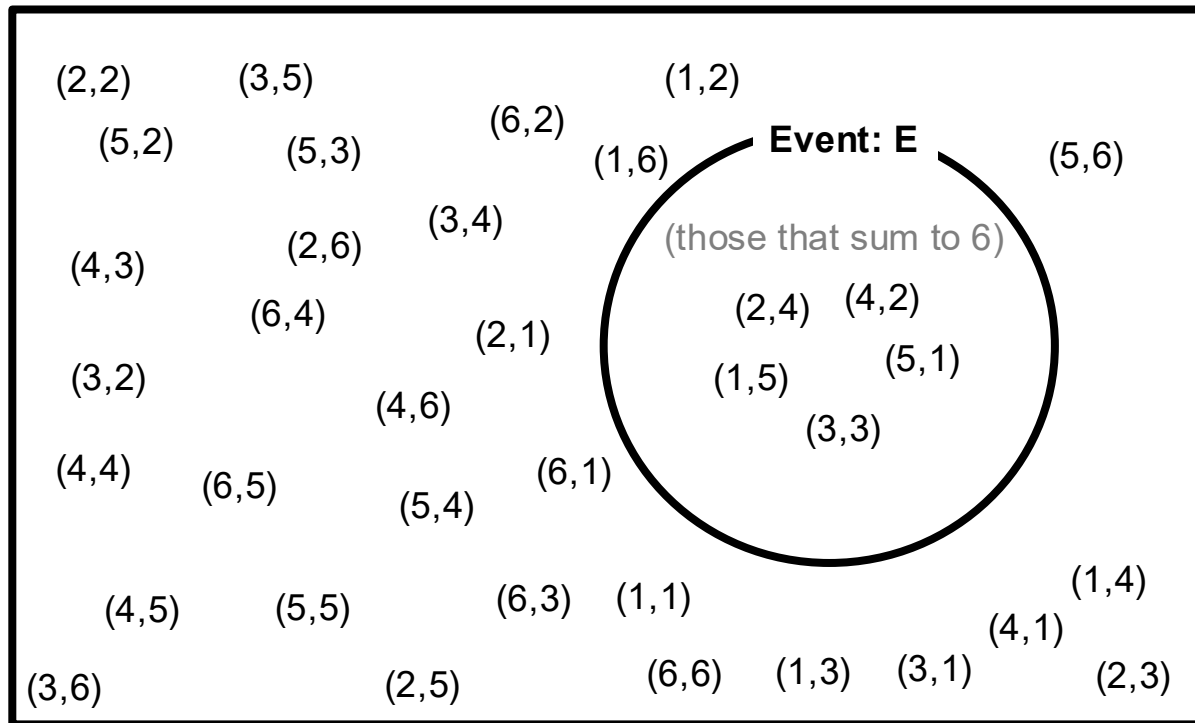 Event E: the two dice's outcomes sum to 6
  - What is the size of E?
  - What is the probability of E?

What is the probability of having two numbers sum to 6?

(2,2)   (3,5)   (1,2)

  (5,2)   (5,3)   (6,2)   **Event: E**   (5,6)

                        (1,6)

                (3,4)   (those that sum to 6)

  (4,3)   (2,6)

      (6,4)   (2,1)   (2,4) (4,2)

  (3,2)   (5,1)

                              (1,5)

          (4,6)   (3,3)

(4,4)   (6,5)   (6,1)

        (5,4)

  (4,5)   (5,5)   (6,3) (1,1)   (1,4)

                                      (4,1)

(3,6)   (2,5)   (6,6) (1,3) (3,1)   (2,3)

$$P(E) = \frac{\#\text{outcomes in } E}{\#\text{outcomes in } S}$$

S = $\{(a, b): a, b \in \{1, \ldots, 6\}\}$

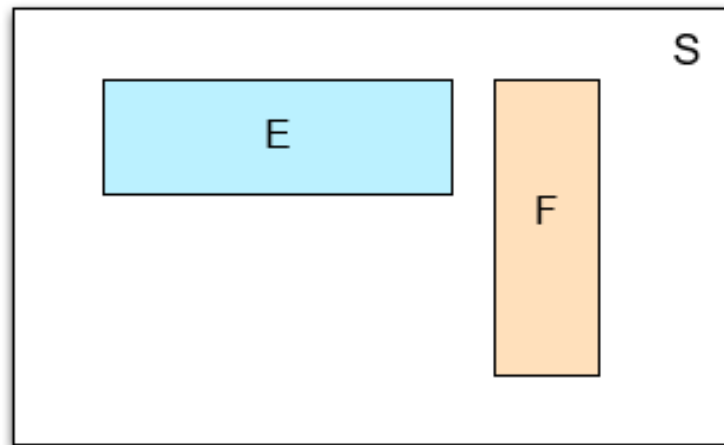Each outcome is equally likely

# of outcomes that sum to 6:
5

answer:
5/36 = 0.13888…

Disjoint events

- In general, breaking an event into *disjoint events* preserves the total probability

- **Disjoint**: E and F are *disjoint* if they cannot happen simultaneously,

- e.g. E = {2, 4}, F={1, 3, 5}

- In such cases,

$$P(E \text{ or } F) = P(E) + P(F)$$

Elementary events

- Notice that we can find the total probability of an event by breaking it into pieces and adding up the probabilities of the pieces:

    P(Even) = P(2) + P(4) + P(6) = 3/6

- These pieces are called 'elementary events': Events that correspond to exactly one outcome

- These elementary events are disjoint events

Partition

- We say that disjoint events $E_1, \ldots, E_n$ form a *partition* of E if any outcome in E lies in exactly one E_i

- e.g.
  - {Fr.} E_1, {Soph.} E_2 form a partition of {Lower division} E
  - {Fr.}, {Soph.} {J.} {Sen.} form a partition of S (sample space)

Freshmen	Sophomores	Juniors	Seniors

Probability as Area

For disjoint events E, F:
$$P(E \text{ or } F) = P(E) + P(F)$$

If disjoint events $E_1, \ldots, E_n$ forms a partition of E :
$$P(E) = P(E_1) + P(E_2) + \cdots + P(E_n)$$

If disjoint events $E_1, \ldots, E_n$ forms a partition of S , for event F (law of total probability):
$$P(F) = P(E_1, F) + P(E_2, F) + \cdots + P(E_n, F)$$

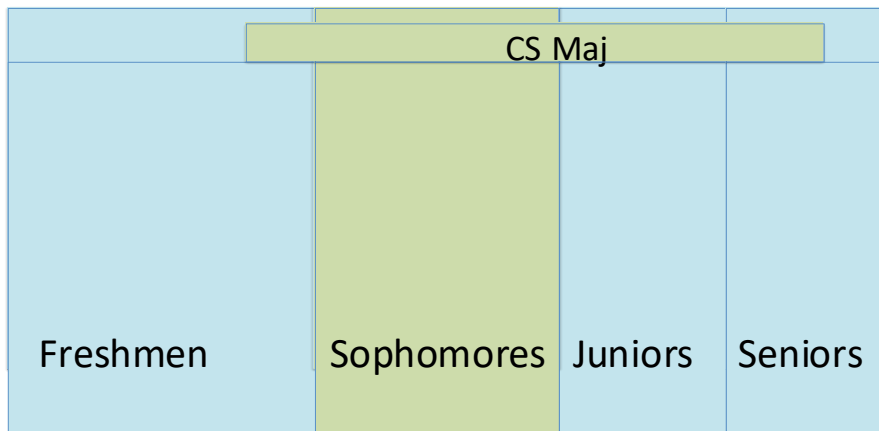Notation: $P(A, B)$ is a shorthand for $P(A \text{ and } B)$

Probability as Area

Examples

- $P(\text{CS}) = P(\text{Fr.}, \text{CS}) + P(\text{Soph.}, \text{CS}) + P(\text{J.}, \text{CS}) + P(\text{Sen.}, \text{CS})$

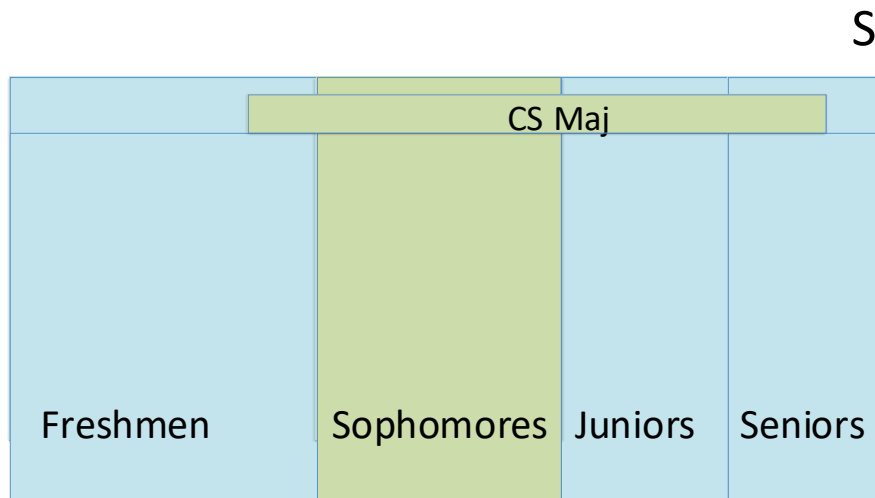 Notation: $P(A, B)$ is a shorthand for $P(A \text{ and } B)$

- $P(\text{Soph.}) = P(\text{CS}, \text{ Soph.}) + P(\text{nonCS}, \text{ Soph.})$
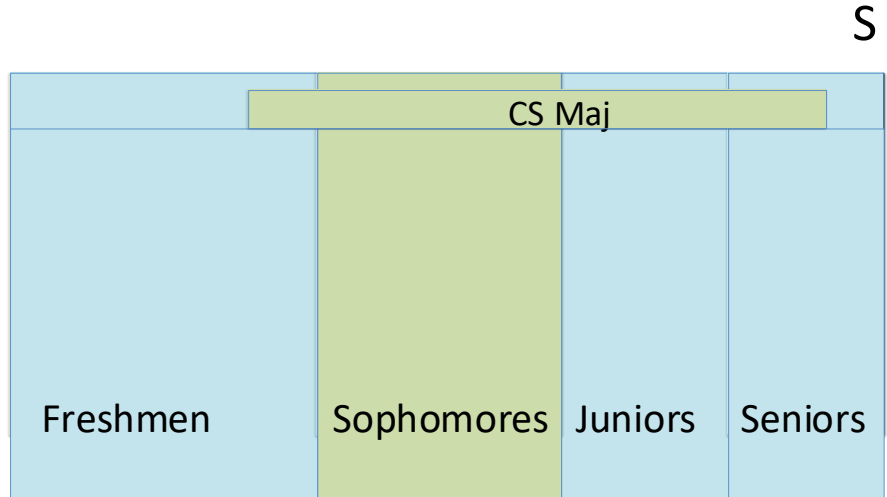
Probability as Area

If events E, F are non-disjoint, what is $P(E \text{ or } F)$?

- $P(\text{Soph.} \text{ or } \text{CS})$?
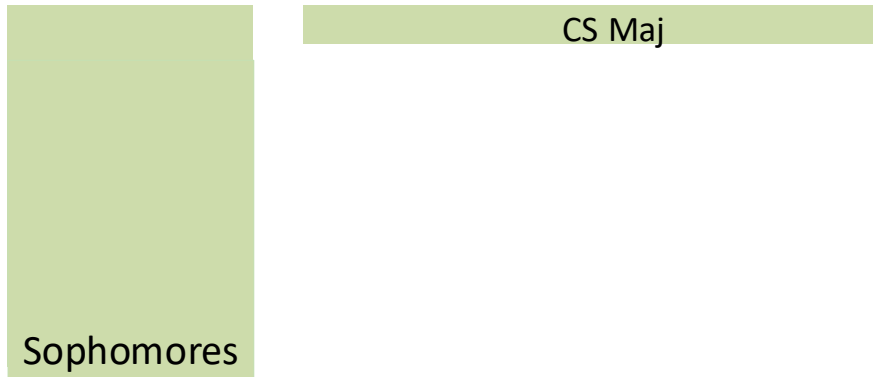- Note: events "Sophomore" and "CS Major" may overlap

Probability as Area

- E = { Soph OR CS }

- Is P(E) = P(Soph) + P(CS)?
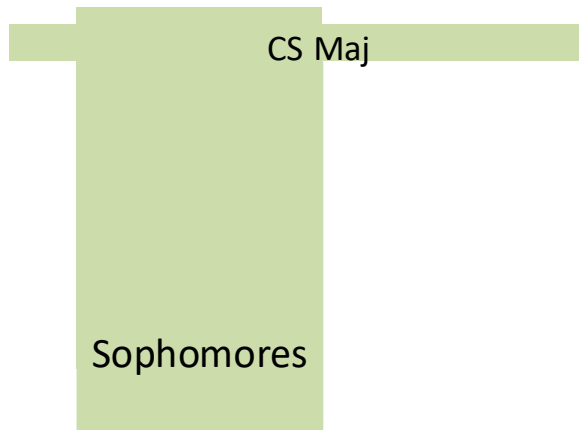 - No

- Which one is larger?
  - Let's see..

S

| | CS Maj | | |
| Freshmen | Sophomores | Juniors | Seniors |

Probability as Area

P(Soph) + P(CS) =
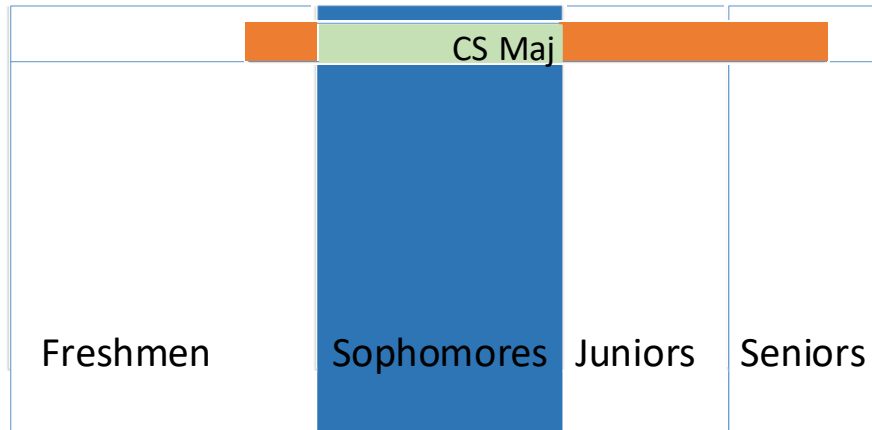
CS Maj

Sophomores

P(Soph or CS) =

CS Maj

Sophomores

Recap

- Outcome: a single observation

- Sample space S: the set of all possible outcomes

- Event: a set of outcomes, a subset of sample space

- Disjoint events:
  - cannot happen together
  - $P(E \text{ or } F) = P(E) + P(F)$

- Law of total probability: If disjoint events $E_1, \ldots, E_n$ forms a partition of S, for event F (law of total probability):

$$P(F) = P(E_1, F) + P(E_2, F) + \cdots + P(E_n, F)$$

Inclusion-Exclusion Principle

- P(Soph) = P(CS , Soph.) + P(Non-CS , Soph.)
- P(CS) = P(Fr. CS) + P(Soph. CS) + P(J. CS) + P(Sen. CS)

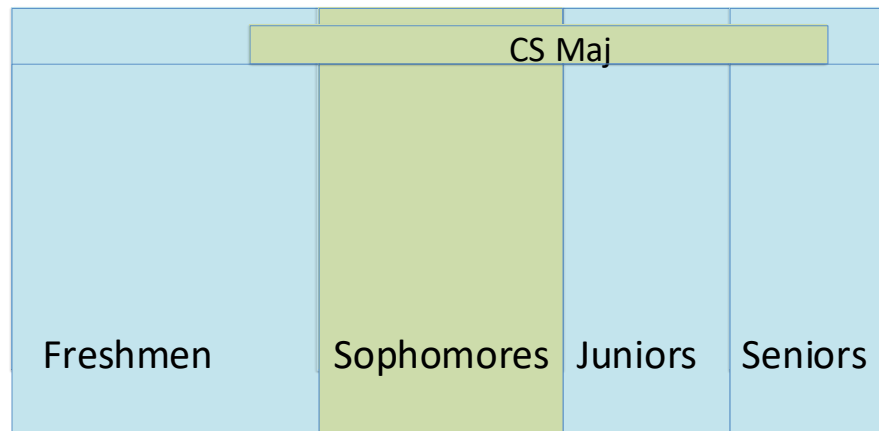- P (Soph or CS) = P(Soph) + P(CS) – P (Soph, CS)   Soph. CS is counted twice
  S

Inclusion-Exclusion Principle

Inclusion-Exclusion Principle For any events E and F,

$$P(E \text{ or } F) = P(E) + P(F) - P(E \text{ and } F)$$

Accounting for overlap
between E and F

S

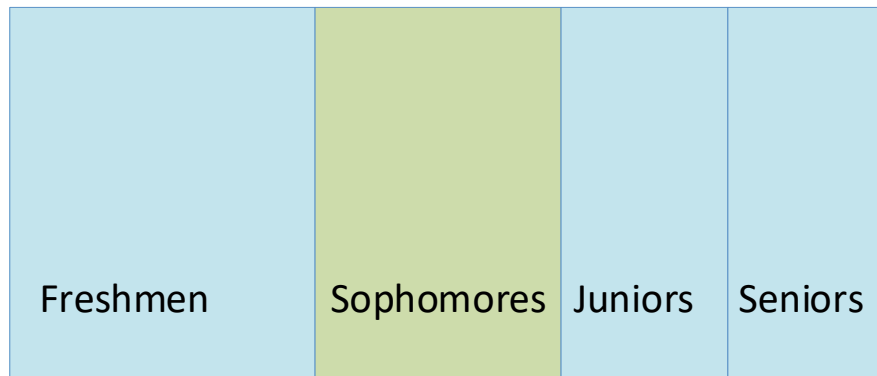# Complementary events

How would I find P(Non-Sophomore)? Two options:

- **Option 1**: List the non-sophomores and then count:
    - *P(Freshman) + P(Juniors) + P(Seniors)*
- **Option 2**: Use the fact that $P(S)$ = 1 and *subtract* instead;
    - *P(Non-Sophomore) = 1 − P(Sophomore)*

| Freshmen | Sophomores | Juniors | Seniors |

Set operations

Events and Set Theory

An event is a set of outcomes and a subset of sample space, so we use set theory to describe and combine them.

Two dice example:

E_1 : *First* die rolls 1 E_2 :*Second* die rolls 1

$E_1 = \{(1,1), (1,2), \ldots, (1,6)\}$ $E_2 = \{(1,1), (2,1), \ldots, (6,1)\}$

- Any die rolls 1: event $E_1 \, or \, E_2$

- Set operation: $E_1 \cup E_2$

Set operations

Two dice example:

E_1 : *First* die rolls 1 E_2 :*Second* die rolls 1

$E_1 = \{(1,1),(1,2),\ldots,(1,6)\}$ $E_2 = \{(1,1),(2,1),\ldots,(6,1)\}$

Operators on events:
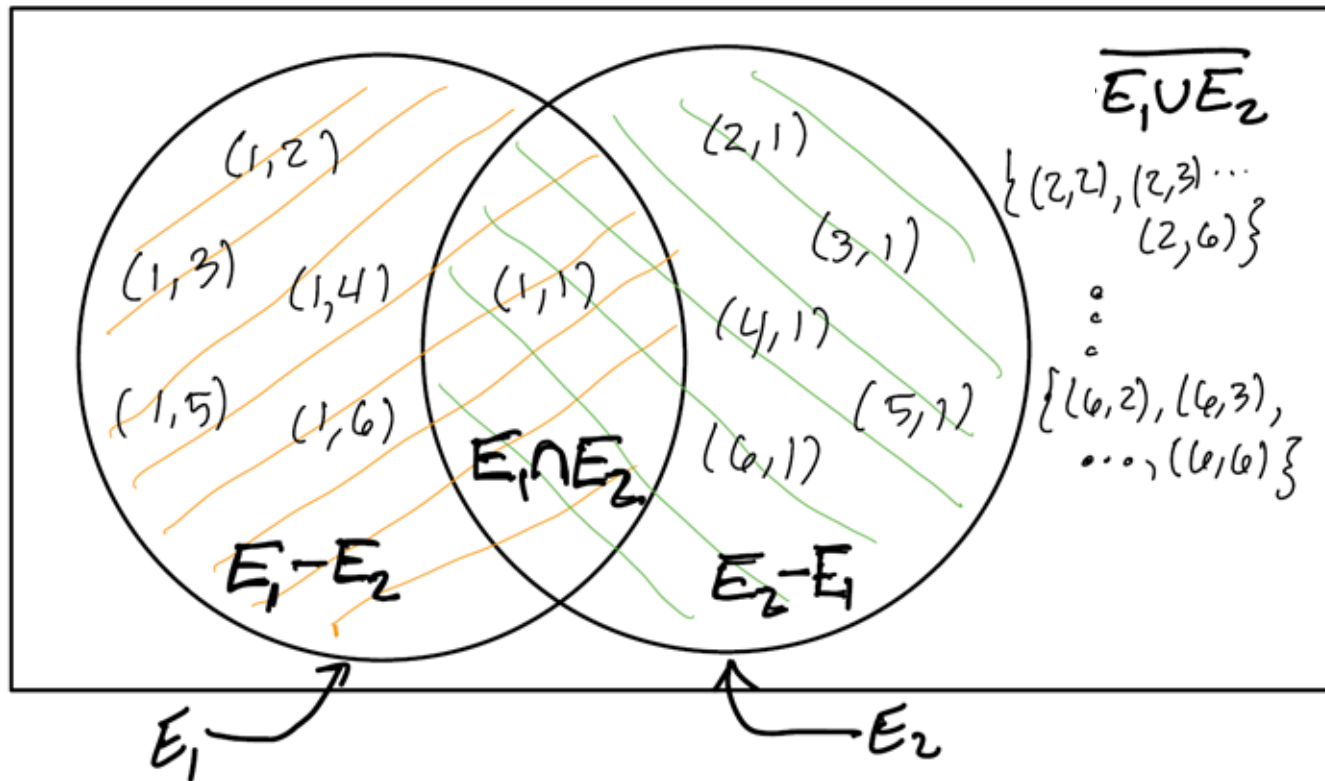
Operation	Value	Interpretation
$E_1 \cup E_2$	$\{(1,1),(1,2),\ldots,(1,6),(2,1),\ldots,(6,1)\}$	Any die rolls 1
$E_1 \cap E_2$	$\{(1,1)\}$	Both dice roll 1
$E_1 \setminus E_2$ $(= E_1 - E_2 := E_1 \cap E_2^c)$	$\{(1,2),(1,3),(1,4),(1,5),(1,6)\}$	Only the first die rolls 1
$\overline{E_1 \cup E_2}$ $(= (E_1 \cup E_2)^c)$	$\{(2,2),(2,3),\ldots,(2,6),(3,2),\ldots,(6,6)\}$	No die rolls 1

Can interpret these operations using a Venn diagram…

Set Theory: De Morgan Law

De Morgan Law 1 $(A \cup B)^C = A^C \cap B^C$
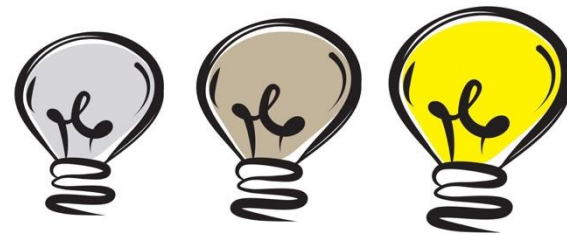
Example:
- A: I bring my cellphone
- B: I bring my laptop
- A^C: I don't bring my cellphone
- B^C: I don't bring my laptop

- $A \cup B$: I bring my cellphone or my laptop
- $(A \cup B)^C$: I bring neither my cellphone nor my laptop
- $A^C \cap B^C$: I didn't bring my cellphone & I didn't bring my laptop

Intersection / union over n events

- n lightbulbs
- E_i: i-th lightbulb is on

- How to describe the event that at least one lightbulb is on?
  - i.e. bulb 1 is on OR … OR bulb n is on
  $$E_1 \cup \cdots \cup E_n =: \cup_{i=1}^{n} E_i$$

- How to describe the event that all lightbulbs are on?
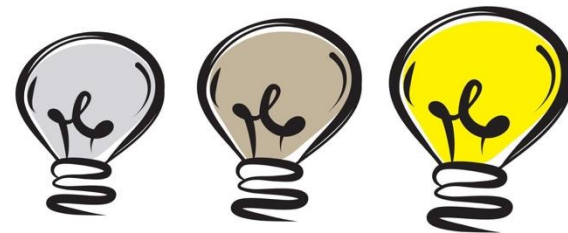 $$E_1 \cap \cdots \cap E_n = \cap_{i=1}^{n} E_i$$

- **De Morgan Laws:**

$$(E_1 \cup \cdots \cup E_n)^C = E_1^C \cap \cdots \cap E_n^C$$

Not (at least one bulb is on)   All bulbs are off

$$(E_1 \cap \cdots \cap E_n)^C = E_1^C \cup \cdots \cup E_n^C$$

Not (all bulbs are on)   At least one bulbs is off

Rules of Probability

Rules of probability

- To recap and summarize:

Rules of Probability

1. **Non-negativity:** All probabilities are between 0 and 1 (inclusive)

2. **Unity of the sample space:** $P(S) = 1$

3. **Complement Rule:** $P(E^C) = 1 - P(E)$

4. **Probability of Unions:**
   - *(a)* *In general,* $P(E \cup F) = P(E) + P(F) - P(E \cap F)$
   - *(b)* *If E and F are disjoint, then* $P(E \cup F) = P(E) + P(F)$

<u>Special case</u>

Assume each outcome is equally likely, and sample space is <u>finite</u>, then the probability of event is:

$$P(E) = \frac{|E|}{|S|}$$
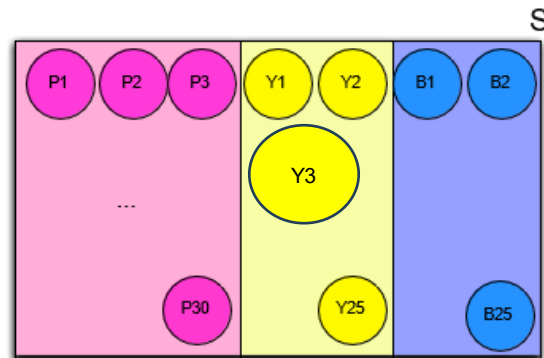
Number of elements in event set

Number of possible outcomes (e.g. 36)

This is called <u>classical probability model</u>

Rethinking the classical probability model

- Classical probability model assumes all outcomes equally likely
- When is this applicable?
  - *Fair* coin toss, *fair* dice throw, …
  - S = {P1, P2, …, P30, Y1, …, Y25, B1, …, B25}

- When is this assumption problematic?
  - *Unfair* coin toss (one side is heavier)
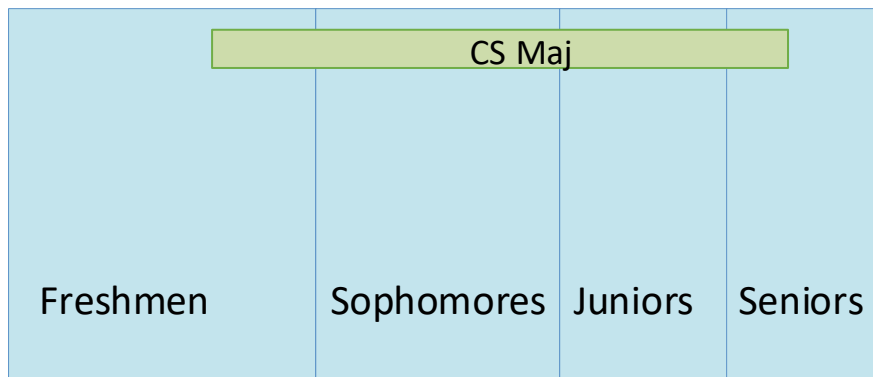 - A yellow ball is much larger

Law of Total Probability

Law of Total Probability

- We saw that:
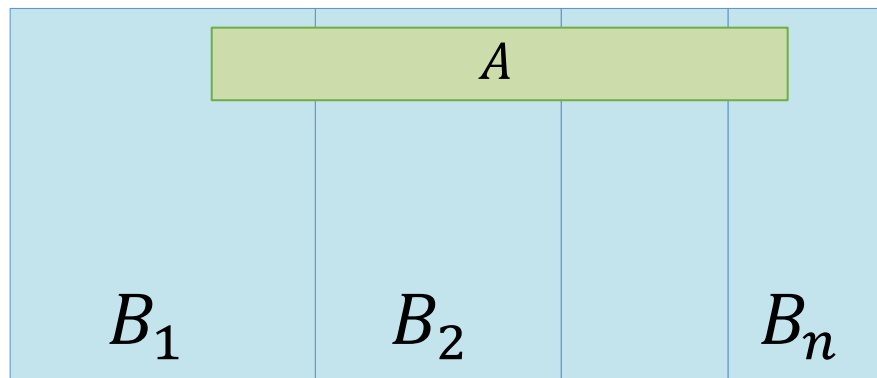$$P(\text{CS}) = P(\text{Fr.}, \text{CS}) + P(\text{Soph.}, \text{CS}) + P(\text{J.CS}) + P(\text{Sen.CS})$$

- Would the equality still be true if, say, we drop $P(\text{Sen.CS})$?
  - No: $\text{Fr.}, \text{Soph.}, J.$ no longer form a partition of {CS}

Law of Total Probability Suppose $B_1, \dots, B_n$ *form a partition* of the sample space S. Then,
$$P(A) = P(A, B_1) + \cdots + P(A, B_n)$$

- Recall notation: $P(A, B_1)$ is a shorthand for $P(A \cap B_1)$

Law of Total Probability: blood types

		Antigen B Absent	Antigen B Present	Marginal
Antigen A	Absent	0.44	0.10	0.54
	Present	0.42	0.04	0.46
	Marginal	0.86	0.14	1.00

- B, B^C form a partition of sample space S, so
- $P(A) = P(A, B) + P(A, B^C) = 0.04 + 0.42 = 0.46$
- Likewise,
- $P(B) = P(B, A) + P(B, A^C) = 0.04 + 0.10 = 0.14$

Summary: calculating probabilities

- If we know that all outcomes are equally likely, we can use

We will use combinatorics to do counting

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

Number of elements in event set

Number of possible outcomes (e.g. 36)

- If |E| is hard to calculate directly, we can try
    - the rules of probability
    - the Law of Total Probability, using an appropriate partition of sample space S