Computer Science

CSC380: Principles of Data Science

Probability 1

Chicheng Zhang

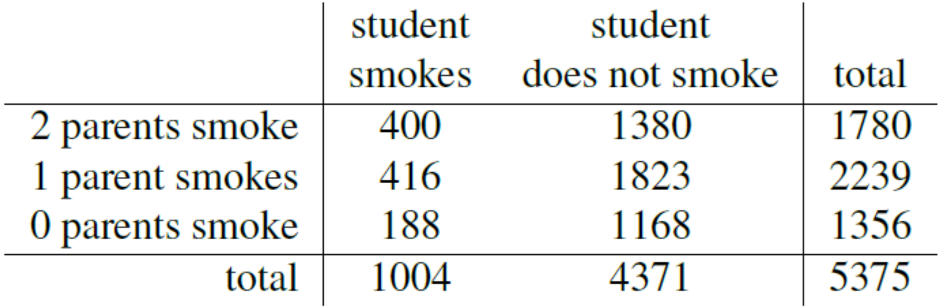
Announcements

- Readings will be given in about the size of one chapter of WJ book per lecture

- Please self-report all participation activities (in-class questions, Piazza answers, OH attendance) – thanks!

Motivation: analyzing data

Example Student smoking data



	student smokes	student does not smoke	total
2 parents smoke	400	1380	1780
1 parent smokes	416	1823	2239
0 parents smoke	188	1168	1356
total	1004	4371	5375

Q: are students with 2 parents smoking more likely to smoke (compared with general students)?

How to measure likeliness of outcomes?

To answer such questions, we will use the language of *probability*, and the concept of *conditional probability*.

Outline

- What is probability?
- Events
- Calculating probabilites
- Set Operations
- Law of Total Probability

What is probability?

What is probability?

- If I flip a coin, what is the probability it will come up heads?
- Most people say 1/2, but why is that?
- What is the probability that the coin will come up either heads or tails?

Principle of Symmetry

- The coin seems to be symmetric, so there’s no reason to think that one side is more likely than the other.

- What’s the probability of rolling a six with a six-sided die?
- What’s the probability of winning the lottery?
- What’s the probability of rain tomorrow?

Interpreting probabilities

- What does it mean to have a probability of 1/2?

- Basically two standard schools of thought on this:
  - Objective probability
  - Subjective probability

Objective Probability

- Probabilities are properties of the external world

- 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

Probability vs. Proportion

- Probability is the proportion of times the corresponding outcome would occur in many repeated trials of a phenomenon.

- Probability is long-term *relative frequency* or *proportion*.

Subjective Probability

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

- I have no reason to believe heads and tails have different probabilities, so I assign them both 1/2.

- Can assign a probability to the truth of any statement that I have a degree of belief about. E.g., Probability of
  - Raining tomorrow
  - Stock price going up this month

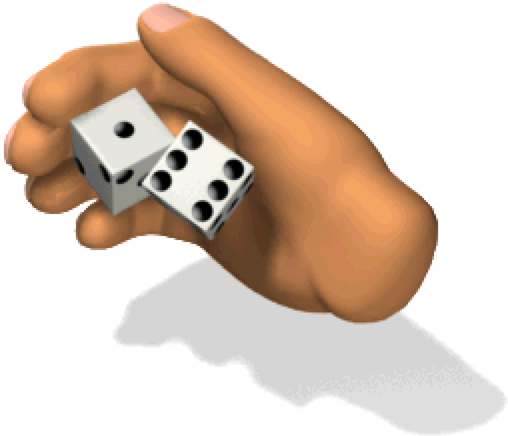
What is probability?

- We will focus on objective probability in the next few lectures
    - Can discuss subjective probability if we have time later in the course (Bayesian statistics)

Events

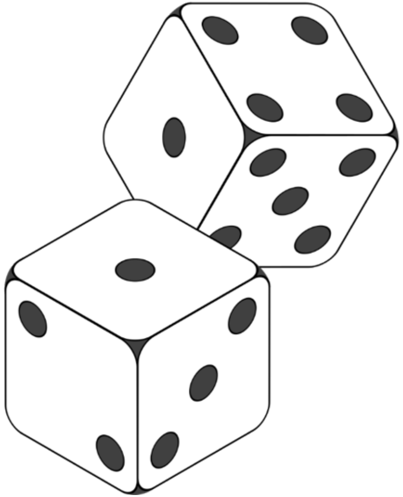
Events and Probability

Suppose we roll <u>two fair dice</u>…



Events and Probability

Suppose we roll <u>two fair dice</u>…



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

*…this is a **random process**.*

How to formalize all these quantitatively?

The Sample Space

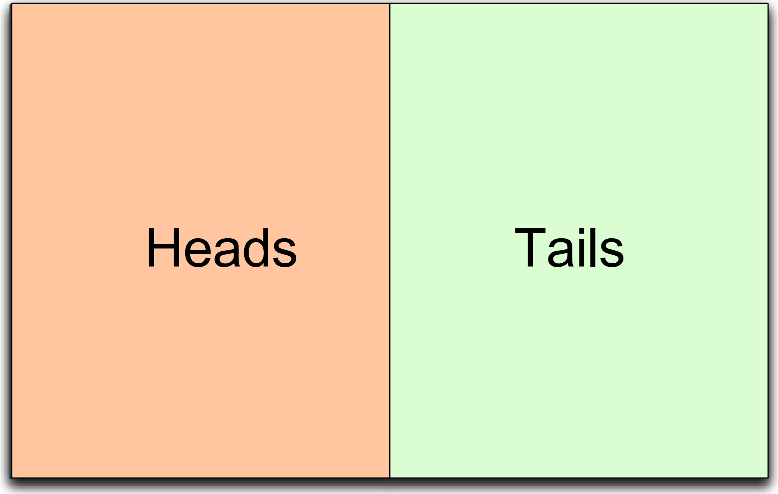
- Probability very closely tied to area, we use lots of spatial metaphors
- The set of all possible outcomes of a random experiment is called the sample space. Often written as S.
- In math, the standard notation for a set is to write the individual members in curly braces:
    - S = {Outcome1, Outcome2, . . . , }
- It's often 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}

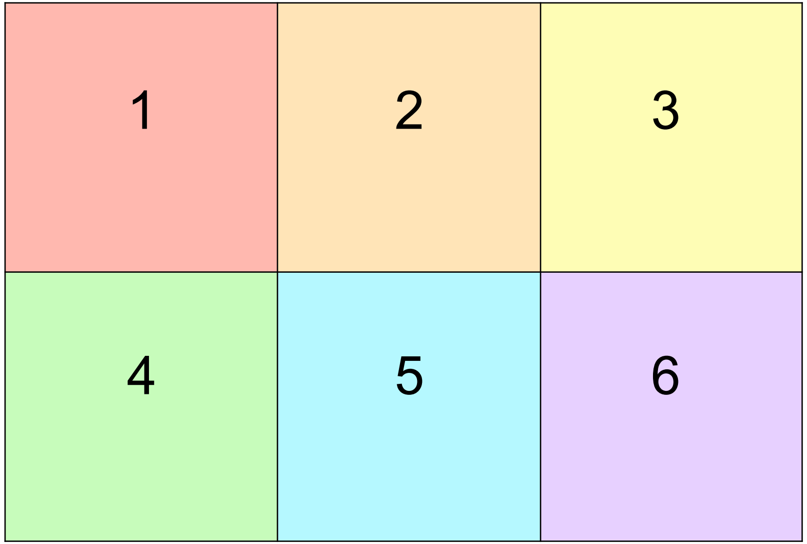


Heads

Tails

Examples of Sample Spaces

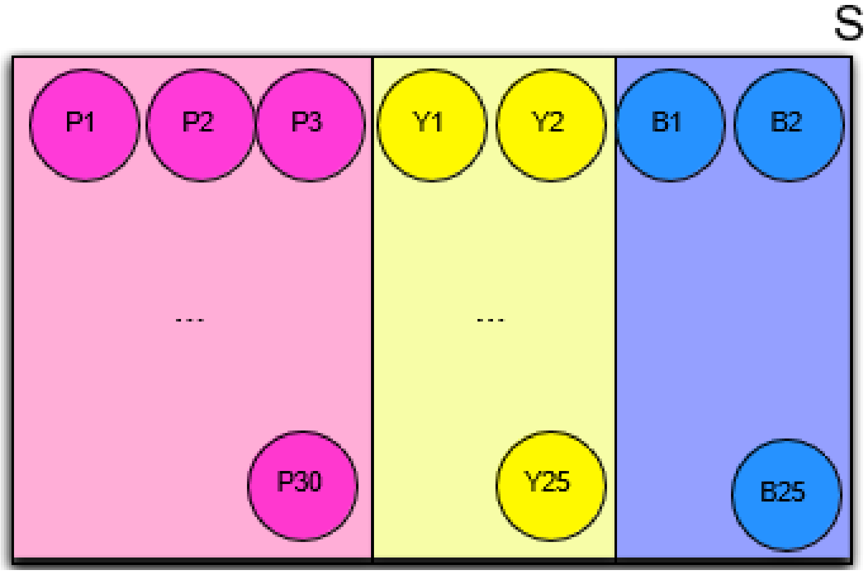
- What is the sample space of rolling a die?
- S = {1, 2, 3, 4, 5, 6}



1	2	3
4	5	6

Examples of Sample Spaces

- What is the sample space of drawing a ball out of an urn 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, . . . }

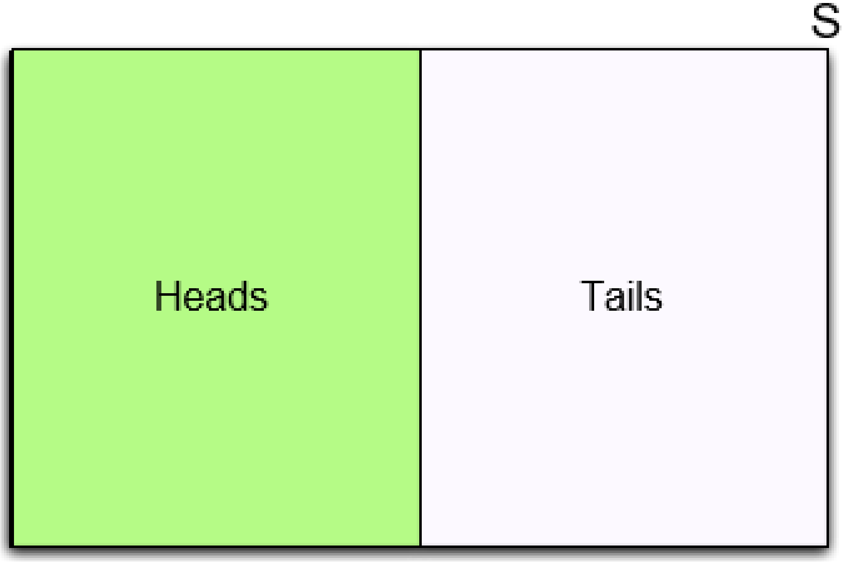
Events

- An event E is a subset of the sample space. When we make a particular observation, it is either “in” E or not.
- It is sometimes helpful to think about events as propositions (TRUE/FALSE statements).
- The proposition is TRUE when the outcome is among the elements of the event set, and FALSE otherwise.
- In other words, the event set contains exactly those outcomes which, if they occur, make the proposition TRUE.

Examples of Events

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

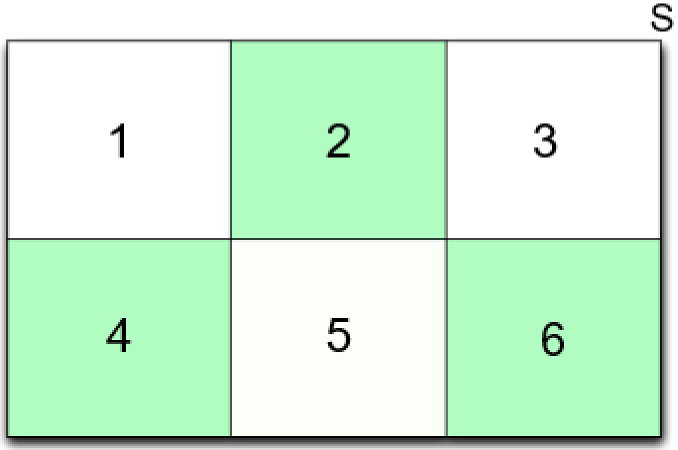
- "The coin comes up heads"
- E = {Heads}



Examples of Events

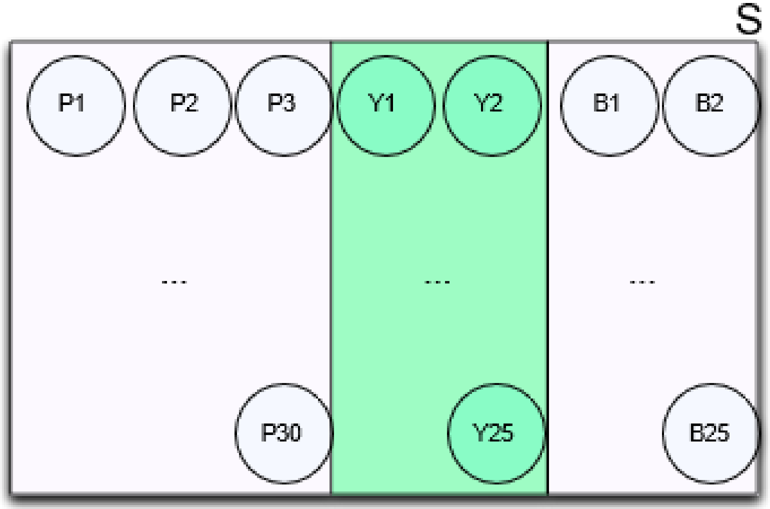
What’s the event set corresponding to the following propositions?

- “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}



Examples of Events

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

- "A sophomore is chosen"
  - E = {Alexandra E., Ana, . . . , Toby, Victoria}
- "There are more than 20 earthquakes"
  - E = {21, 22, 23, . . . }
- "I get exactly one heads"
  - E = {HT, TH}
- "I get at least one heads"
  - E = {HH, HT, TH}

Special events

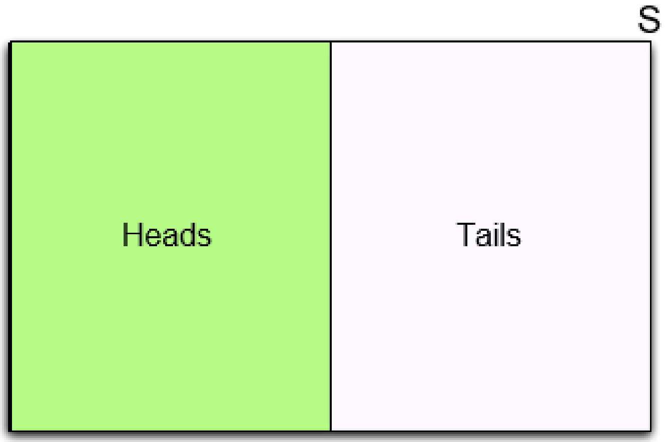
- The sample space S itself is an event
  - E.g. “there are at least zero earthquakes”
  - It is an event that always happens

- The empty set ∅ is also an event
  - It is an event that never happens
  - E.g. “the die comes up 7”

Calculating Probabilities

Calculating probability

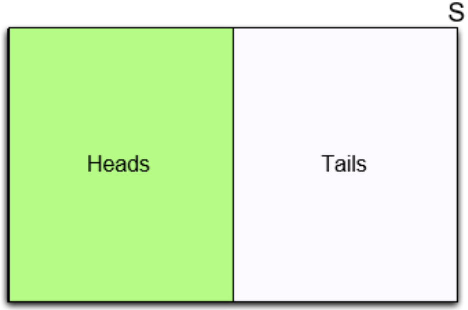
- We can think of the probability of an event E as its area, where S always has a total area of 1.0
- So, the probability of E is the fraction of S that it takes up.



Calculating probability using symmetry

- If we have a sample space for which the principle of symmetry applies (i.e., every outcome is equally likely), then we can find event probabilities easily.
- Oftentimes called the "classical probability model"
- Since every outcome has the same "area", we can just count:

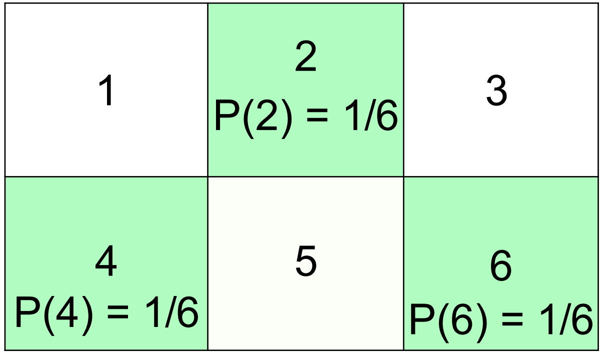
$$P(E) = \frac{\#\text{outcomes in } E}{\#\text{outcomes in } S}$$



Probability as Area

What is the probability of

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



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

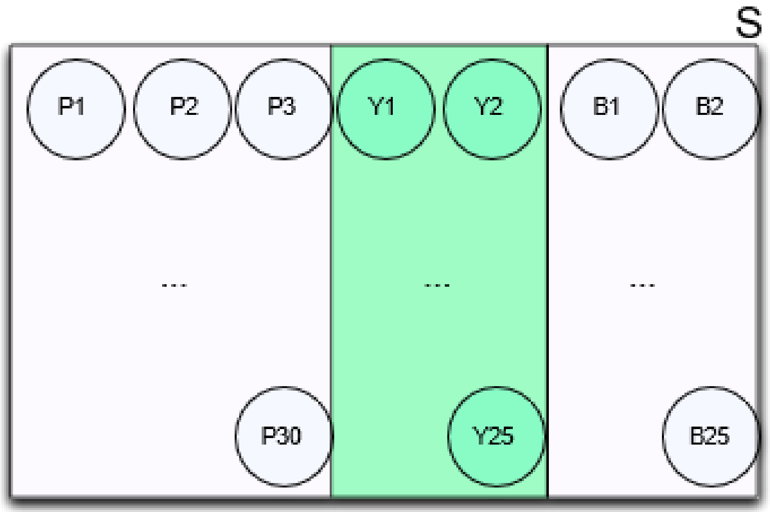
P(S) = 1

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

Probability as Area

What is the probability of

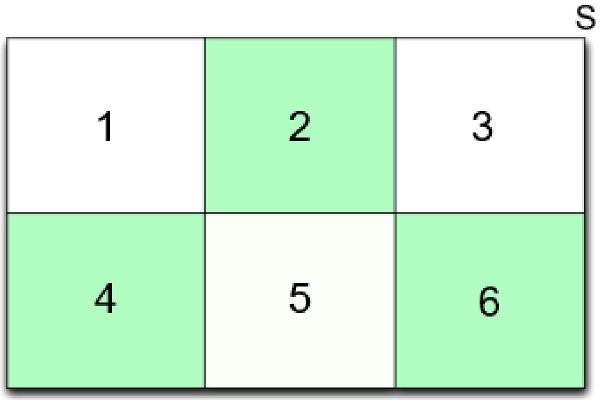
- Selecting a yellow ball?
- E = {Y1, Y2, . . . , Y25}
- $P(E) = \frac{\# E}{\# S} = \frac{25}{30+25+25} = \frac{5}{16}$



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

Recap

Sample space: set of all possible outcomes



Event

- “The die comes up an even number”
- E = {2, 4, 6}

Classical probability model:

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

In-class activity

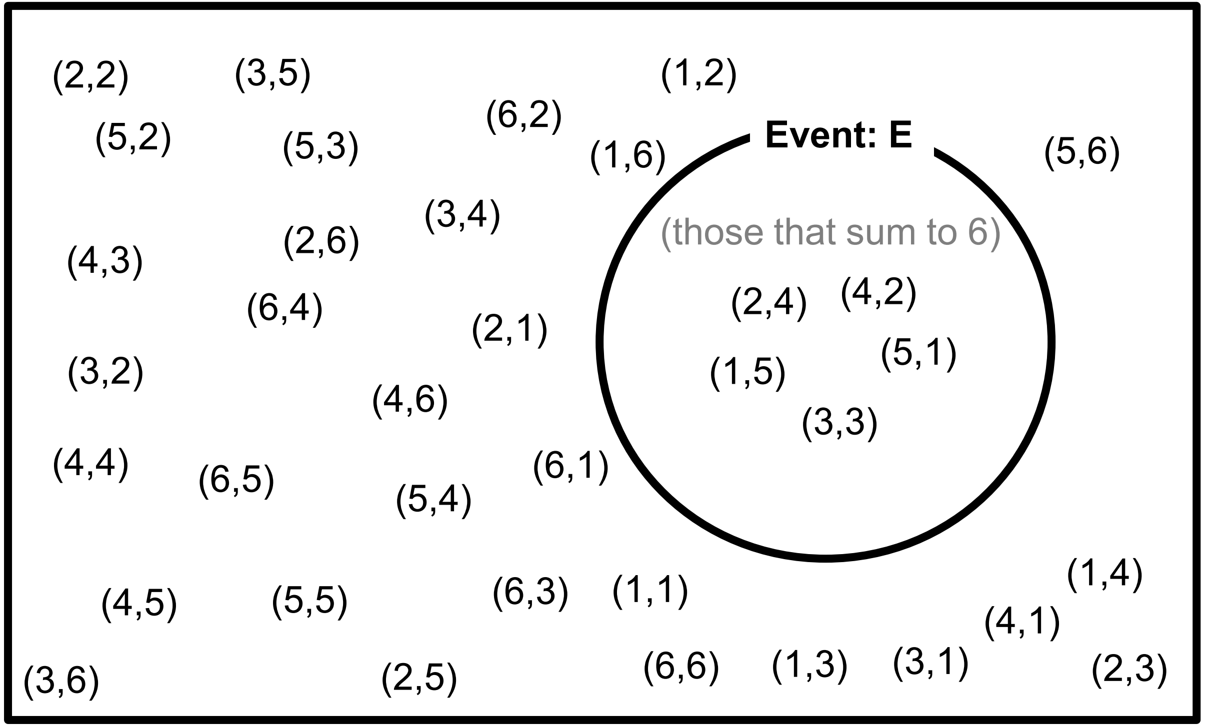
- Suppose we throw two fair dice
  - What is the sample space S (space of all possible outcomes)?
  - Hint: can use, e.g. (1,2) to represent that red die comes up 1 and blue die comes up 2

Event E: the two dice's outcomes sum to 6

- What is the size of E?
- What is the probability of E?

Random Events and Probability

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



(2,2) (3,5) (1,2)
(6,2)
(5,2) (5,3) (1,6)
Event: E (5,6)
(3,4)
(those that sum to 6)
(4,3) (2,6)
(2,4) (4,2)
(6,4)
(2,1)
(5,1)
(3,2) (1,5)
(4,6)
(3,3)
(4,4) (6,5) (6,1)
(5,4)
(1,4)
(6,3) (1,1)
(4,5) (5,5)
(4,1)
(3,6) (2,5) (6,6) (1,3) (3,1) (2,3)

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

Each outcome is equally likely

# of outcomes that sum to 6:
5

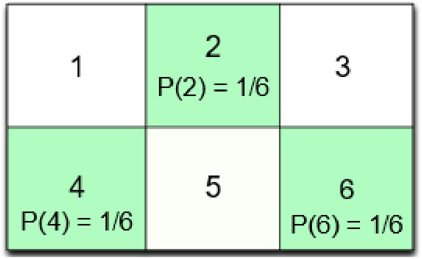
answer:
(1/36) * 5 = 0.13888…

Probability as Area

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

$$P(\text{Even}) = P(2) + P(4) + P(6)$$

$$P(\text{Yellow}) = P(Y1) + \cdot\cdot\cdot + P(Y25)$$



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

- These pieces are called ‘elementary events’
  - Events that correspond to exactly one outcome

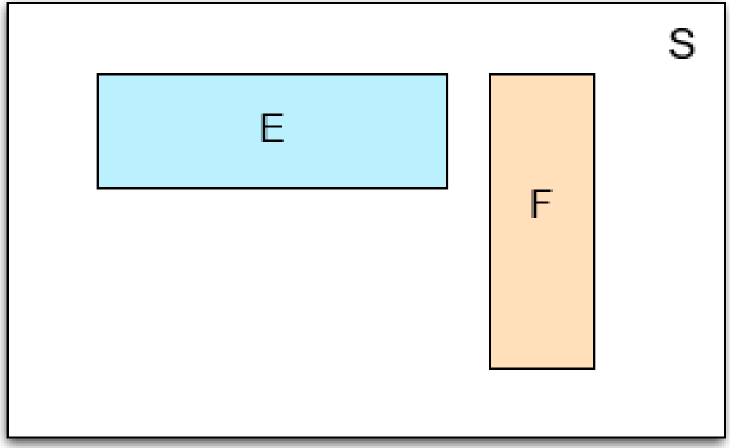
Probability as Area

- In general, breaking an event into *disjoint events* preserves the total probability
- E and F are said to be *disjoint* if they cannot happen simultaneously, e.g.

  E = {even numbers}, F={1,3,5}

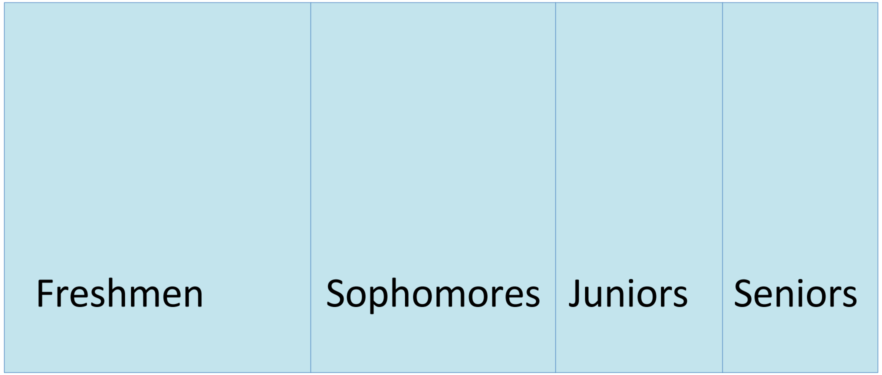
- In such cases,

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



Partition

- We say that 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.}, {Soph.} form a partition of {Lower division}
  - {Fr.}, {Soph.} {J.} {Sen.} form a partition of S



- In general, a partition of S do not leave any element out

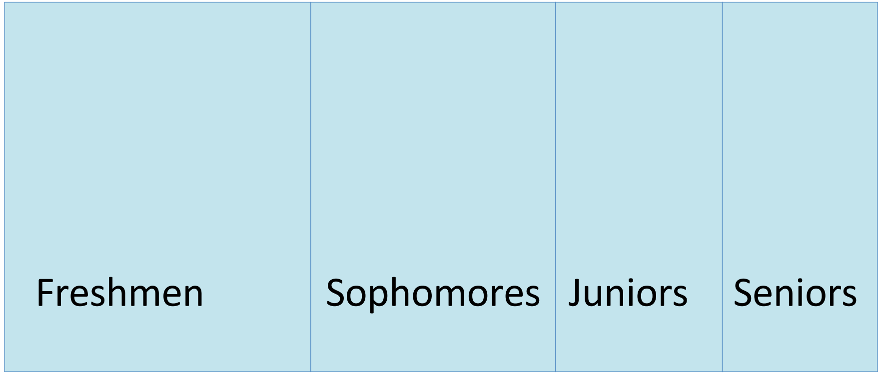
Probability as Area

- **Fact** For disjoint events E, F,

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

More generally, If $E_1, \dots, E_n$ forms a partition of E,

$$P(E) = P(E_1) + P(E_2) + \cdots + P(E_n)$$



Freshmen

Sophomores

Juniors

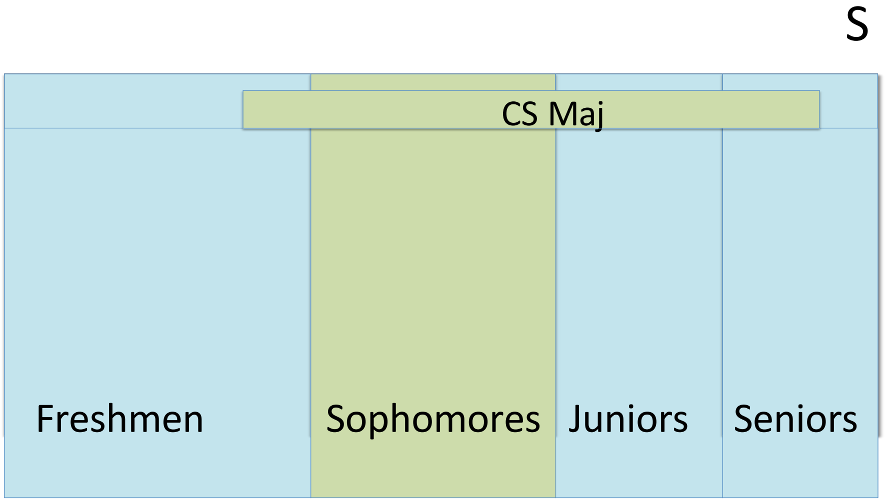
Seniors

Probability as Area

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

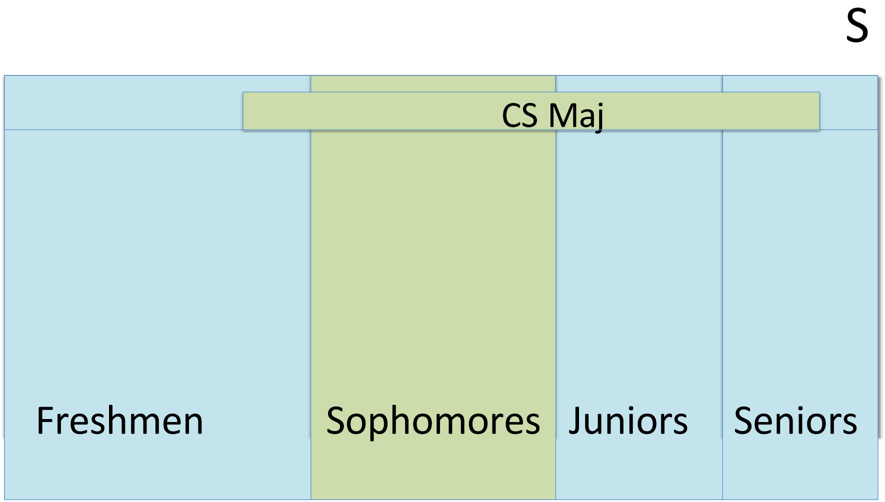
Notation: $P(A, B)$ is a shorthand for $P(A \text{ and } B)$

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



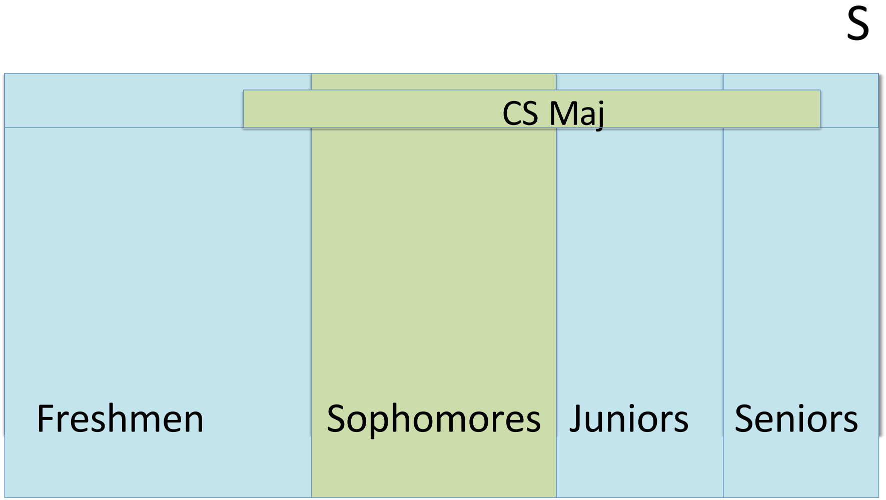
Probability as Area

- What about the probability of selecting a sophomore OR a CS major?
  - 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..



Probability as Area

P(Soph) + P(CS)

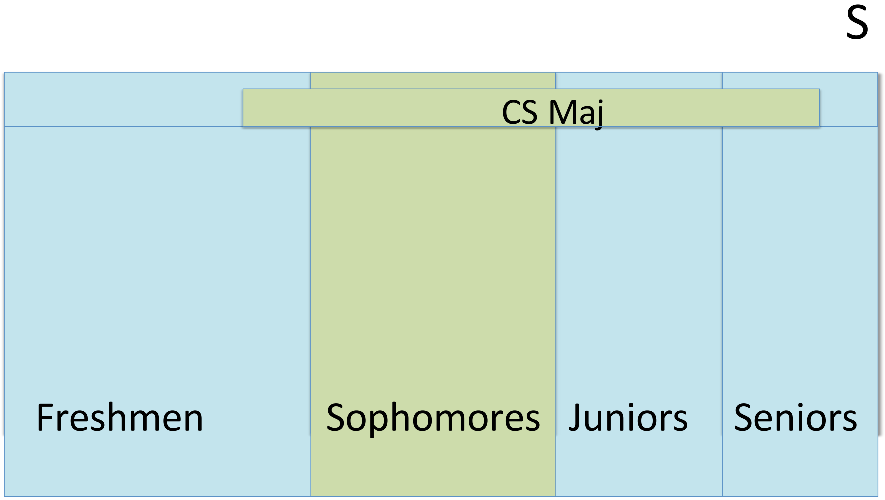
= P(Soph. CS) + P(Soph. Non-CS) +

P(Fr. CS) + P(Soph. CS) + P(J. CS) + P(Sen. CS)

Group Soph. CS is counted twice

So, P(Soph OR CS)

= P(Soph) + P(CS) - P(Soph. CS)

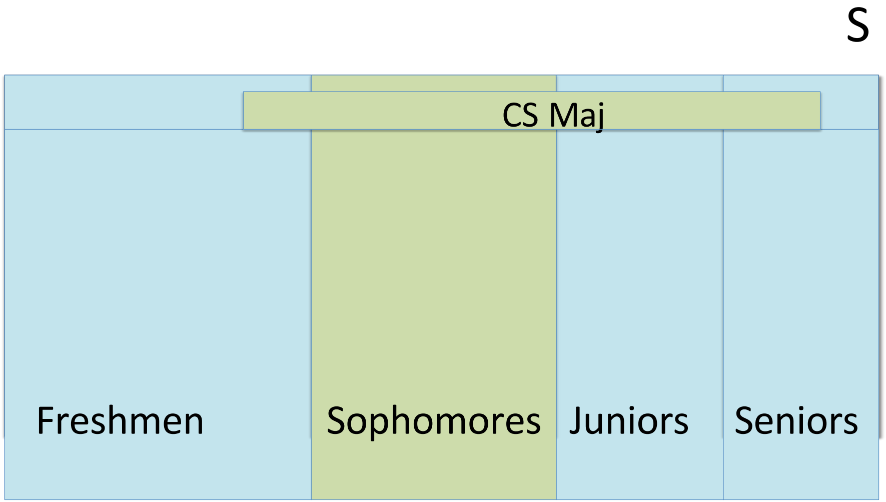


Inclusion-Exclusion Principle

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



Complementary events

How would I find *P*(Non-Sophomore)?

- Could just list the non-sophomores and then count, but we can use the fact that *P*(*S*) = 1 and *subtract* instead.
  - *P*(Non-Sophomore) = 1 − *P*(Sophomore)

Set operations

Set operations

Two dice example: *Suppose*

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

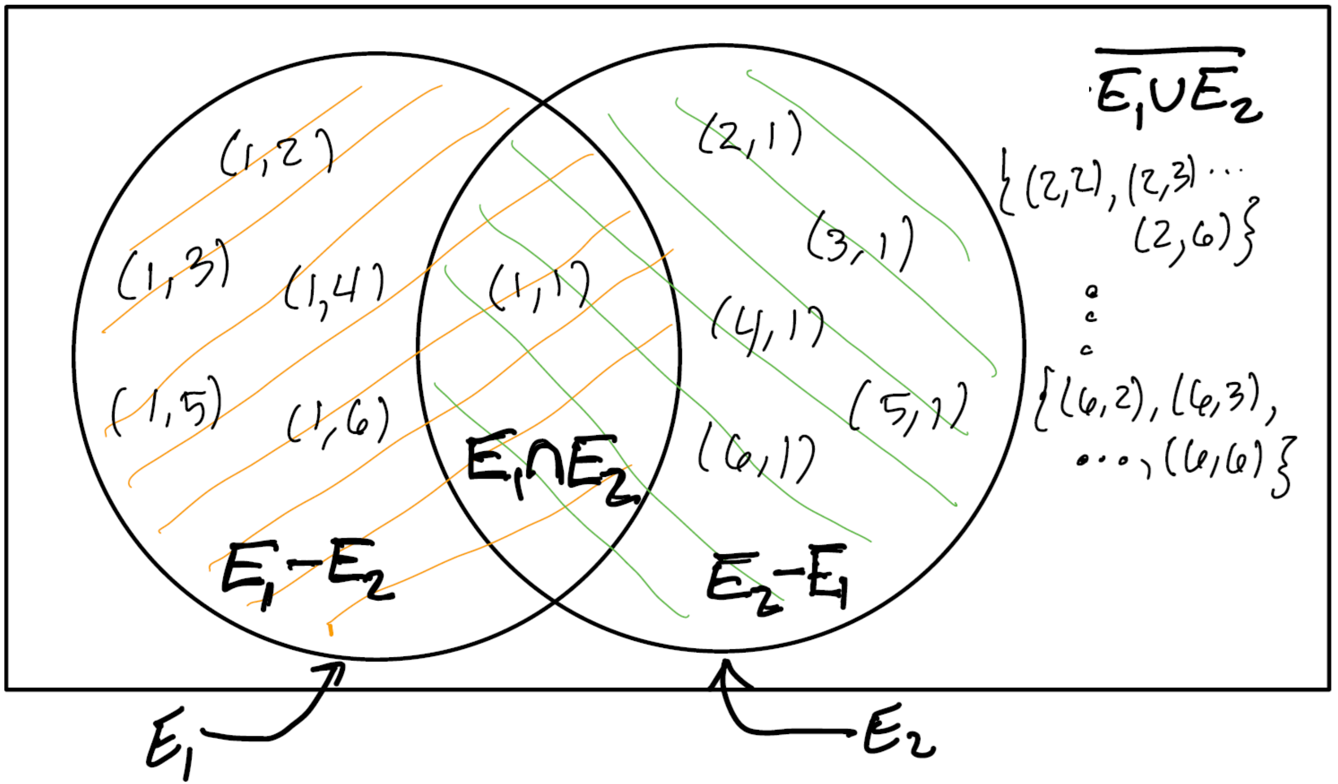
$$E_1 = \{(1,1),(1,2),\ldots,(1,6)\} \qquad E_2 = \{(1,1),(2,1),\ldots,(6,1)\}$$

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
$\boldsymbol{E_1 \setminus E_2}$ $(= E_1 - E_2 \ := E_1 \cap E_2^c)$	$\{(1,2),(1,3),(1,4),(1,5),(1,6)\}$	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

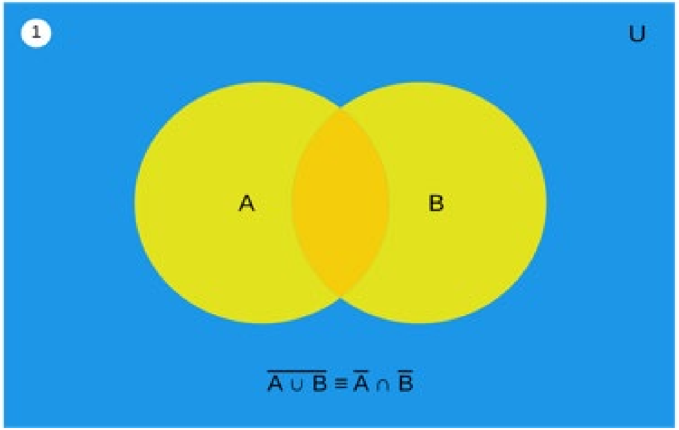
Set operations

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:

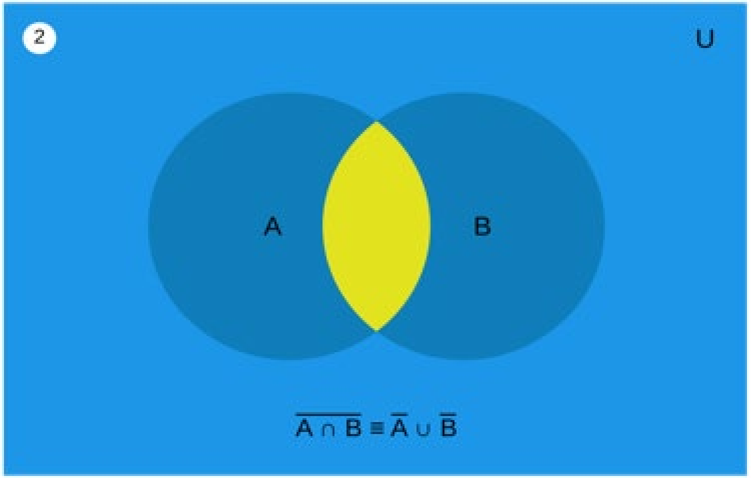
- A: I bring my cellphone
- B: I bring my laptop
- A^C: I don’t bring my cellphone
- B^C: I don’t bring my laptop

- $A \cup B$: 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

Set Theory: De Morgan Law

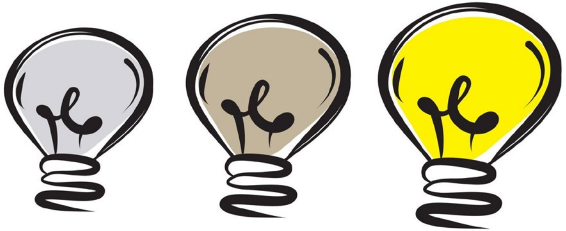
De Morgan Law 2 $(A \cap B)^C = A^C \cup B^C$

Ex: try to make sense of it using the same example above



- De Morgan Law generalizes to a collection of n events
  - But first, let's define some notations

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

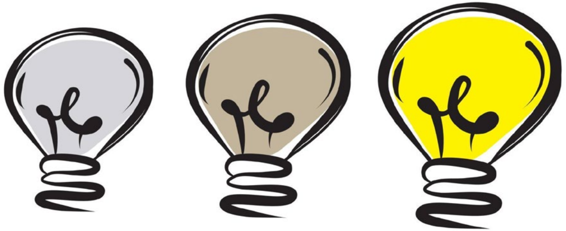
$$E_1 \cup \cdots \cup E_n =: \cup_{i=1}^{n} E_i$$

- 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 with n events

- **De Morgan Laws:**



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

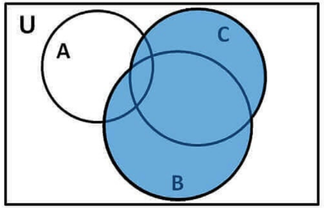
Not (at least one bulb is on) All bulbs are off

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

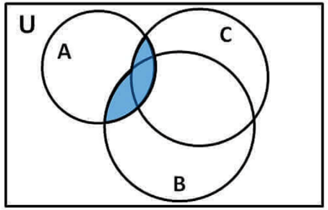
Not (all bulbs are on) At least one bulbs is off

Set operation: distributive law

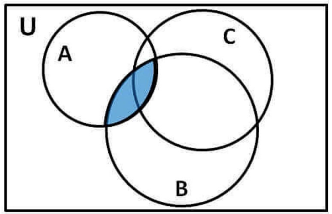
- Distributive law in arithmetics $a(x + y) = ax + ay$ carry over to sets
- **Distributive Law 1** $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$
- Justification by picture:



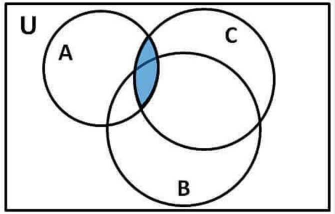
(B ∪ C)



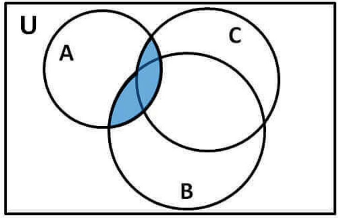
A ∩ (B ∪ C)



(A ∩ B)



(A ∩ B)



(A ∩ B) ∪ (A ∩ C)

Set operation: distributive law

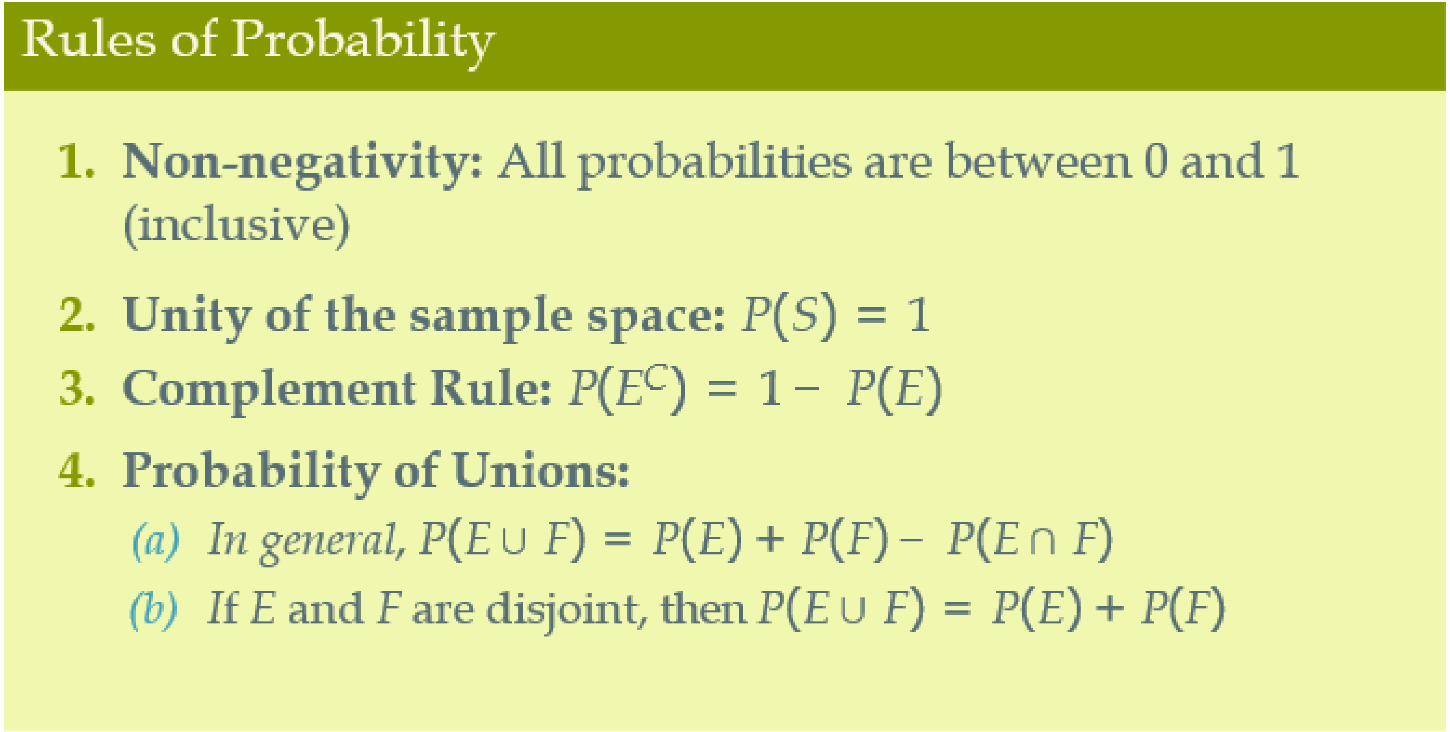
- **Distributive Law 2** $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$

- Can justify this by:
  - drawing a picture (like previous slide), or
  - proving it using Distributive Law 1 and De Morgan Law

Rules of Probability

Rules of probability

- To recap and summarize:



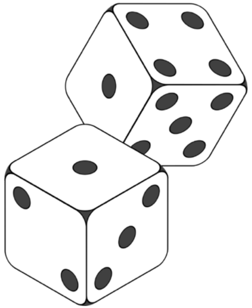
Rules of Probability

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 $P(E \cup F) = P(E) + P(F)$

Classical probability model

Special case

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



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

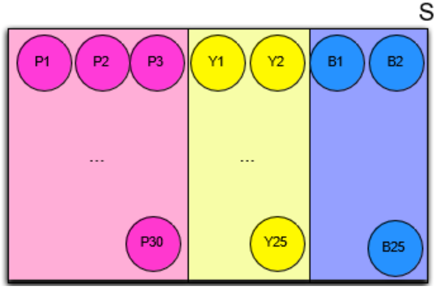
Number of elements in event set

Number of possible outcomes (e.g. 36)

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

Rethinking the classical probability model

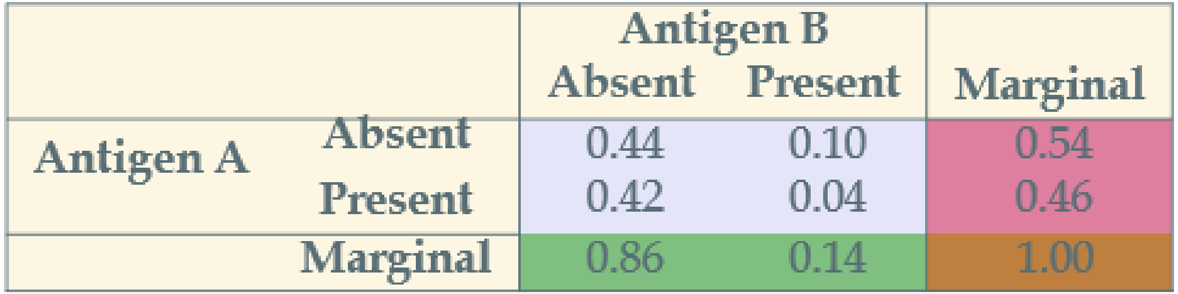
- Classical probability model assumes all outcomes are equally likely
- When is this applicable?
  - *Fair* coin toss, *fair* dice throw, …
  - In the urn example,
    S = {P1, P2, …, P30, Y1, …, Y25, B1, …, B25}



- When is this assumption problematic?
  - *Unfair* coin toss (e.g. one side of the coin is heavier)
  - In the urn example, S = {P, Y, B}
  - defining a good outcome space can sometimes simplify our reasoning

Exercise: Blood types

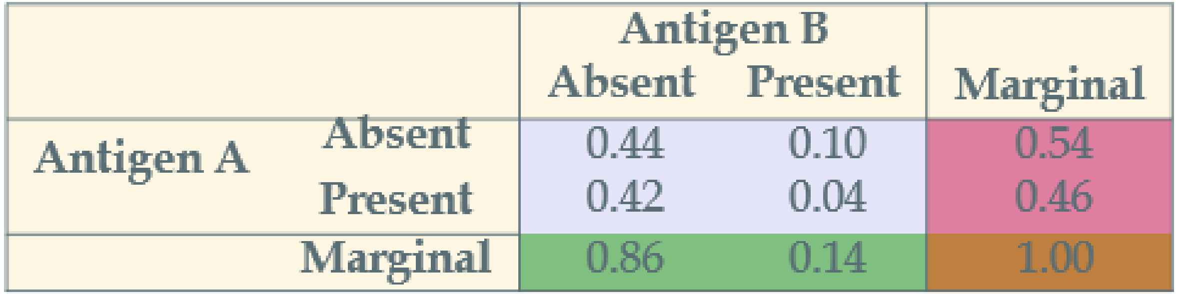
- Human blood is classified by the presence or absence of two antigens, called A and B. This gives rise to four types: O, A, B, and AB.



		Antigen B		
		Absent	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

Table: Probability Estimates for U.S. Blood Types

Exercise: Blood types

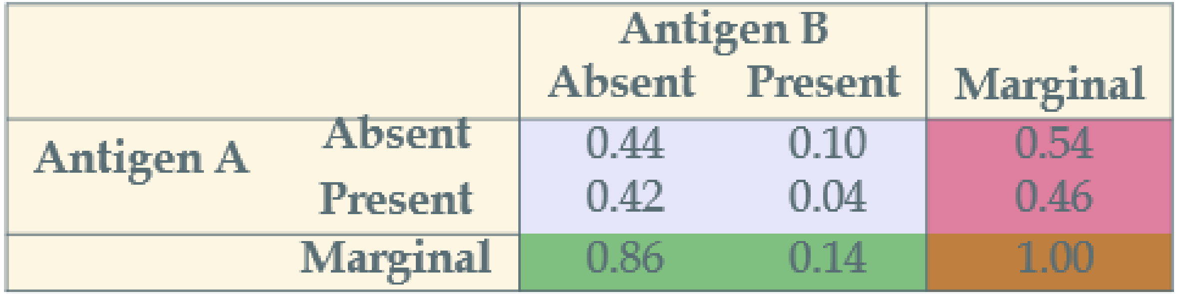


		Antigen B		
		Absent	Present	Marginal
Antigen A	Absent	0.44	0.10	0.54
	Present	0.42	0.04	0.46
	Marginal	0.86	0.14	1.00

Table: Probability Estimates for U.S. Blood Types

- If A is the event “presence of antigen A”, and B is the event “presence of antigen B”, what is:
  - $P(A \cap B)$? What is this event in words?
  - $P(A^C \cap B)$? What is this event in words?

Exercise: Blood types



		Antigen B		
		Absent	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

Table: Probability Estimates for U.S. Blood Types

- What is $P(A \cup B)$ in words? What is its numeric value?
- Can we rephrase this event?
- $A \cup B \ = \ (A^C \cap B^C)^C$, by De Morgan's Law
- So, using the Complement Rule:
- $P(A \cup B) = 1 - P(A^C \ \cap \ B^C \ )$, which in this case is easy to compute.

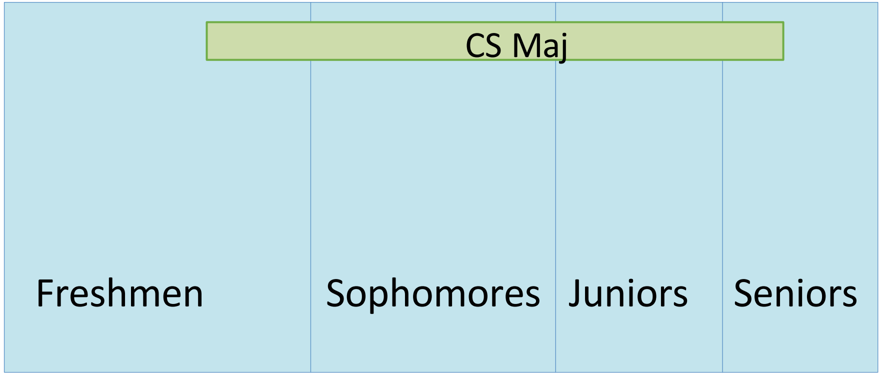
Law of Total Probability

Law of Total Probability

- We saw that:

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

- Is there a general rule behind this?

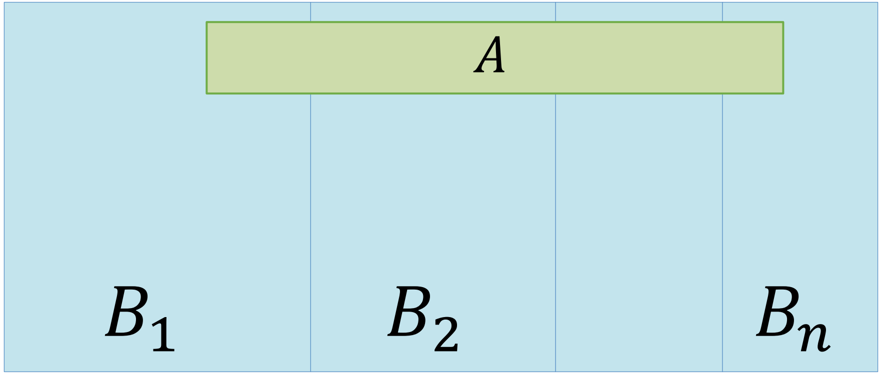


- Would the equality still be true if, say, we drop $P(\text{Sen}.\text{CS})$?
  - No – the three remaining events no longer form a partition of {CS}

Law of Total Probability

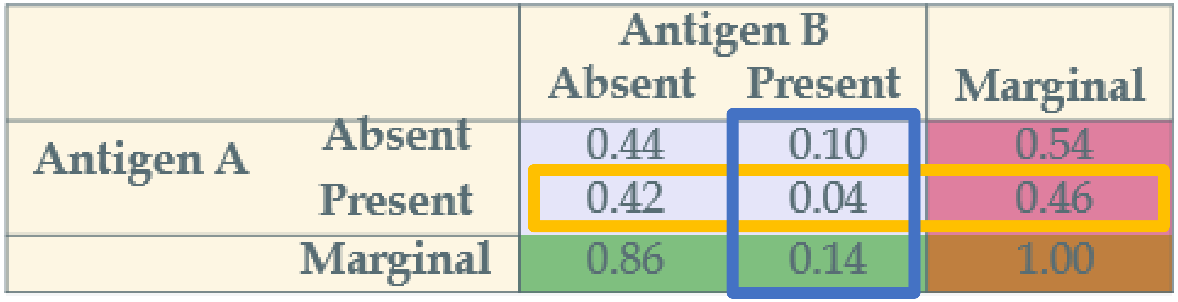
Law of Total Probability Suppose $B_1, \ldots, B_n$ *form a partition* of the sample space S. Then,

$$P(A) = P(A, B_1) + \cdots + P(A, B_n)$$



- Recall notation: $P(A, B_1)$ is a shorthand for $P(A \cap B_1)$
- Why? $A \cap B_1, \ldots, A \cap B_n$ form a partition of A

Law of Total Probability: blood types



		Antigen B		
		Absent	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

Table: Probability Estimates for U.S. Blood Types

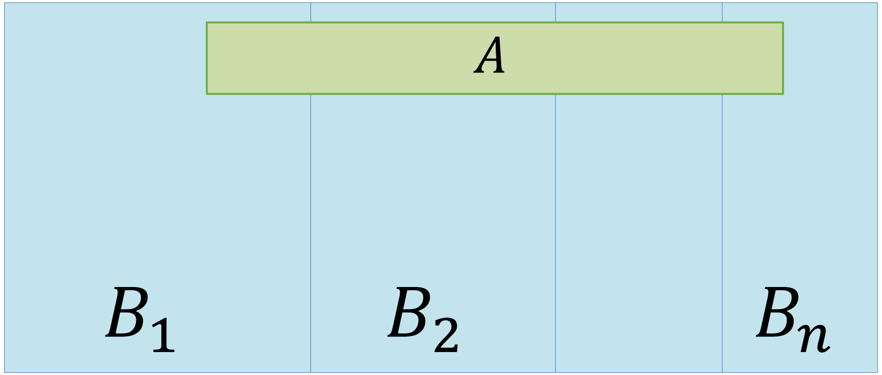
- B, B^C form a partition of sample space S, so
- $P(A) = P(A,B) + P(A,B^C) =$ 0.04+ 0.42 = 0.46
- Likewise,
- $P(B) = P(B,A) + P(B,A^C) =$ 0.04 + 0.10 = 0.14

Law of Total Probability: another example

Example Roll two fair dice. Let X be the outcome of the first die. Let Y be the sum of both dice. What is the probability that both dice sum to 6 (i.e., Y=6)?

$$p(Y=6)=\sum_{x=1}^{6} p(Y=6, X=x)$$

$\{X=1\}\ldots,\{X=6\}$ form a partition of sample space S



$$= p(Y=6, X=1) + p(Y=6, X=2) + \ldots + p(Y=6, X=6)$$
$$= \frac{1}{36} + \frac{1}{36} + \frac{1}{36} + \frac{1}{36} + \frac{1}{36} + 0 = \frac{5}{36}$$

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 using the rules of probability
- If this is still challenging, we can try using the Law of Total Probability, using an appropriate partition of sample space S

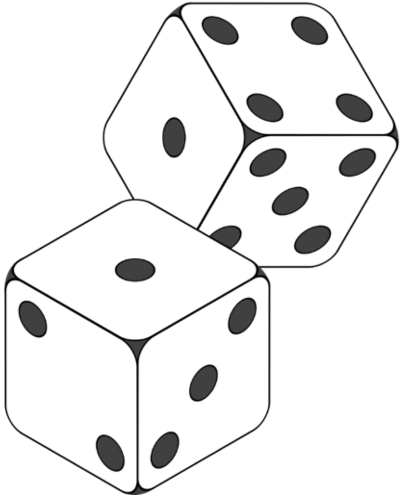
Backup

Random Events and Probability

Probability of a random event

$\approx$

Simulate the random process n times, the fraction of times this event happens



- How large should n be?
- Simulation results vary from trails?

Numpy Background

Numpy: numerical computing package

```
import numpy as np
np.random.randint(1,1+6,size=10)
=> array([5, 4, 1, 1, 1, 5, 5, 2, 4, 6])
```

randint(low,high,size)
: generate \`size\` random numbers in {low, low+1, …, high-1}

Numpy array

- Replaces python’s **list** in numpy.
- More numerical functionality
- It’s a ‘vector’ in mathematics.

```
a=np.array([1,2]); b=np.array([4,5])
a+b
⇒ np.array([5,7]) // elementwise addition
np.dot(a,b)
⇒ 14            // dot product
```

Random Events and Probability

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

```
import numpy as np
for n in [10,100,1_000,10_000,100_000]:
    res_dice1 = np.random.randint(1,6+1,size=n)
    res_dice2 = np.random.randint(1,6+1,size=n)
    res = [(res_dice1[i], res_dice2[i]) for i in range(len(res_dice1))]

    cnt = len(list(filter(lambda x: x[0] + x[1] == 6, res)))
    print("n=%6d, result: %.4f " % (n, cnt/n))
```

```
n=      10, result: 0.1000
n=     100, result: 0.1200
n=    1000, result: 0.1350
n=   10000, result: 0.1365
n=  100000, result: 0.1388
n= 1000000, result: 0.1385
```

```
n=      10, result: 0.1000
n=     100, result: 0.1900
n=    1000, result: 0.1540
n=   10000, result: 0.1366
n=  100000, result: 0.1371
n= 1000000, result: 0.1394
```

every time you run, you get a different result

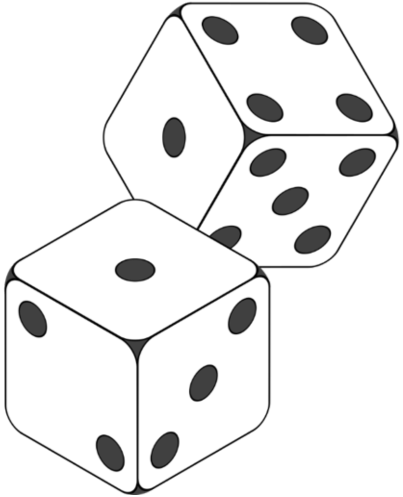
however, the number seems to **converge** to 0.138-0.139

There seems to be a precise value that it will converge to.. what is it?

Random Events and Probability

- **Theoretical probability** describes how likely an event is going to occur based on math.

- **Experimental probability** describes how frequently an event actually occured in an experiment.



Mathematics of Probability

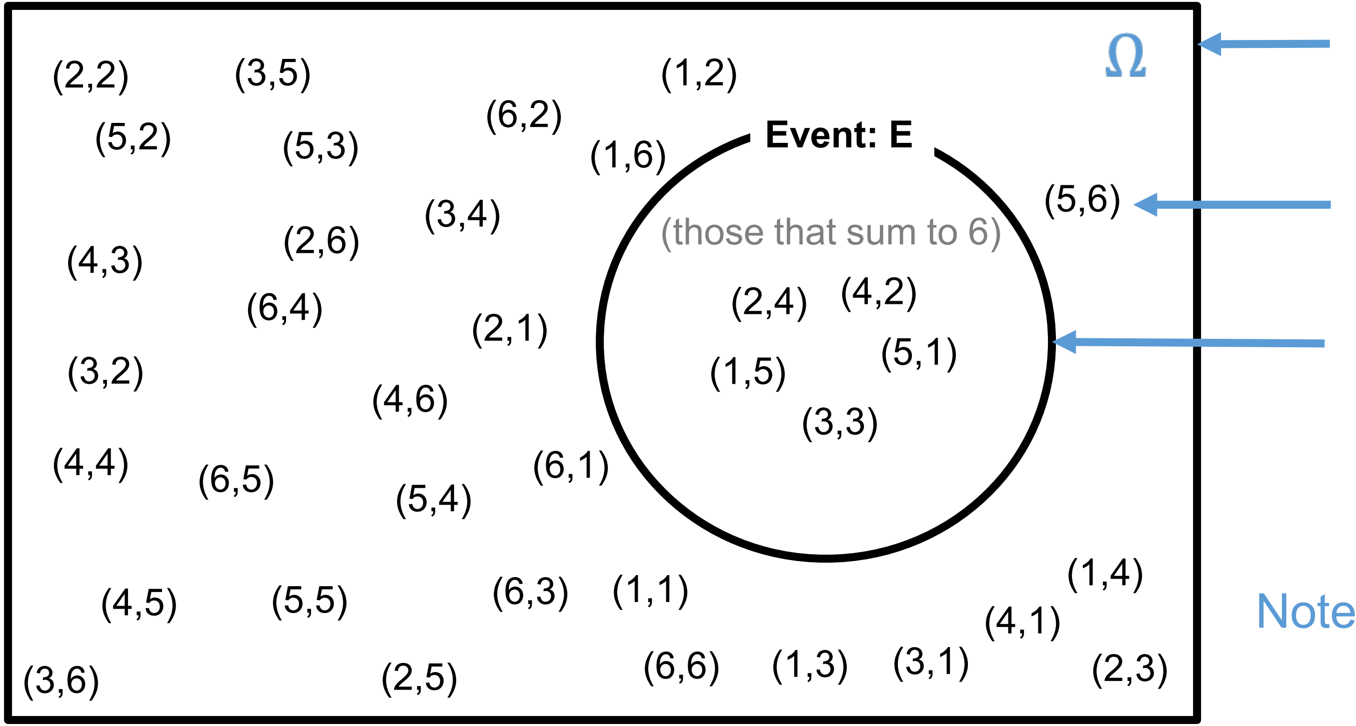
- **Probability** is a real-world phenomenon.

- But under what mathematical framework can we formulate **probability** so we can solve practical problems?
  - e.g., weather prediction, predicting the election outcome

- **<u>Disclaimer</u>**: not all mathematics correspond to real-world phenomenon (e.g., Banach–Tarski paradox). Fortunately, we will not talk about this in our lecture ☺

Random Events and Probability

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



Ω

(2,2) (3,5) (1,2) (6,2) (5,2) (5,3) (1,6) (3,4) (2,6) (4,3) (6,4) (2,1) (3,2) (4,6) (4,4) (6,5) (6,1) (5,4) (5,6) (4,5) (5,5) (6,3) (1,1) (1,4) (4,1) (3,6) (2,5) (6,6) (1,3) (3,1) (2,3)

Event: E

(those that sum to 6)

(2,4) (4,2) (1,5) (5,1) (3,3)

<u>Outcome space</u> consists of all possible outcomes

Each **<u>outcome</u>** $\omega \in \Omega$.

<u>Event</u> is a collection of outcomes and a subset of the outcome space

$E \subseteq \Omega$

Note: outcome = sample

Random Events and Probability

Some examples of events…

Q: how many such pairs? 9

- Both even numbers

$$E^{\text{even}} = \{(2,2),(2,4),\ldots,(6,4),(6,6)\}$$

- The sum of both dice is even,

$$E^{\text{sum even}} = \{(1,1),(1,3),(1,5),\ldots,(2,2),(2,4),\ldots\}$$

- The sum is greater than 12,

$$E^{\text{sum}>12} = \emptyset$$

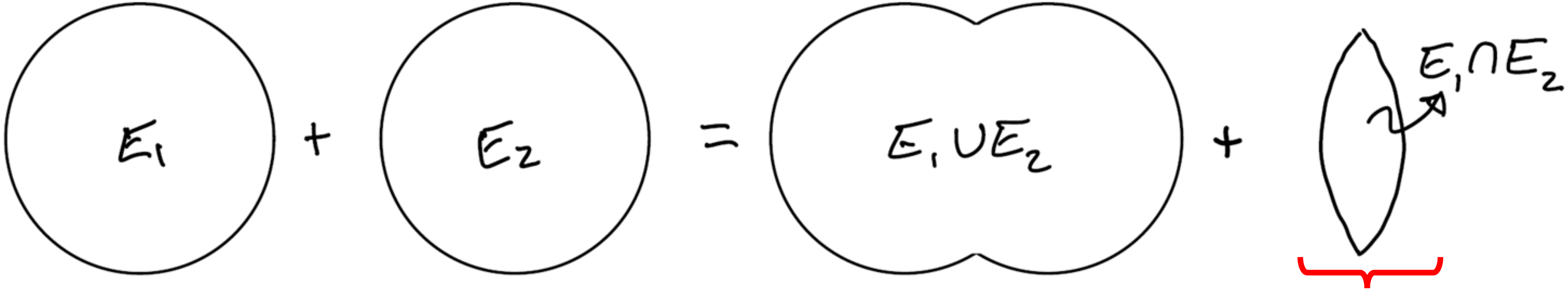
We can talk about impossible outcomes

Inclusion-exclusion Rule

Lemma: (inclusion-exclusion rule) *For <u>any</u> two events* E_1 *and* E_2,

$$P(E_1 \cup E_2) = P(E_1) + P(E_2) - P(E_1 \cap E_2)$$

Graphical Proof:



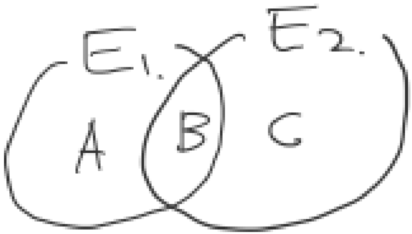
Subtract from both sides

Alternative Proof

Lemma: *For <u>any</u> two events* E_1 *and* E_2 *,*

$$P(E_1 \cup E_2) = P(E_1) + P(E_2) - P(E_1 \cap E_2)$$

Alternative proof:

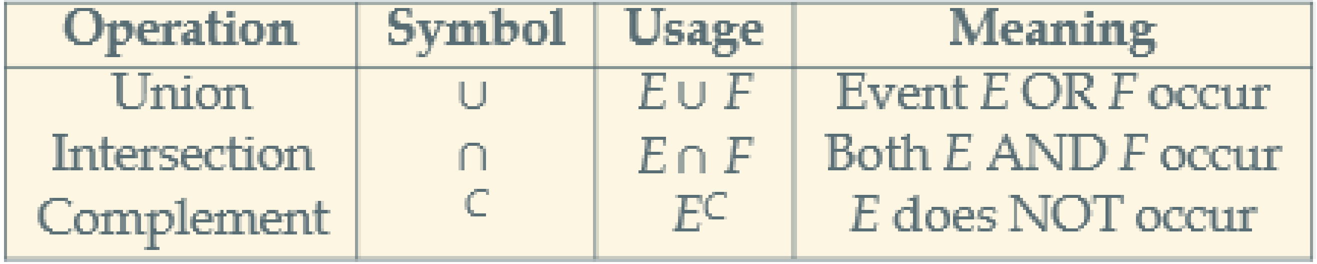


$A = E_1 - (E_1 \cap E_2)$
$B = E_1 \cap E_2$
$C = E_2 - (E_1 \cap E_2)$.

$$\begin{aligned}
&P(E_1 \cup E_2)\\
&= P(A \cup B \cup C)\\
&= P(A) + P(B) + P(C) \qquad \text{(by axiom 3)}\\
&= P(A) + P(B) + P(B) + P(C) - P(B)\\
&= P(A \cup B) + P(B \cup C) - P(B) \qquad \text{(by axiom 3)}
\end{aligned}$$

Set notations vs. Logic notations

- Rather than write out AND, OR and NOT all the time, we can use notation from set algebra.



Operation	Symbol	Usage	Meaning
Union	$\cup$	$E \cup F$	Event *E* OR *F* occur
Intersection	$\cap$	$E \cap F$	Both *E* AND *F* occur
Complement	^{C}	E^{C}	*E* does NOT occur

- Just like when we add or multiply two numbers we get back another number, if we take the union or intersection of two events, we get back a new event.
- The complement of an event is also an event (kind of like the negative of a number, or the evil twin of a person)

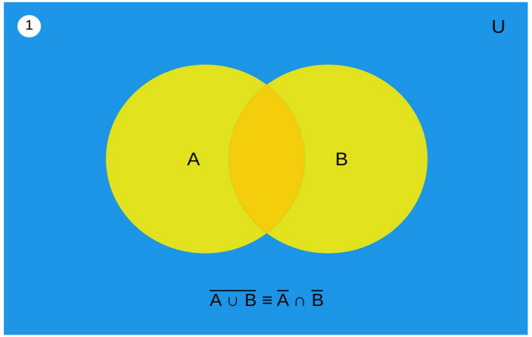
Set Theory: De Morgan Law

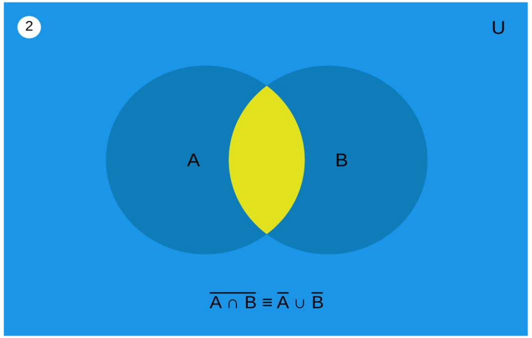
De Morgan Law 1 $(A \cup B)^C = A^C \cap B^C$

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





- $\neg\left(\bigcup_n A_n\right) = \bigcap_n \neg A_n$, $\neg\left(\bigcap_n A_n\right) = \bigcup_n \neg A_n$ DeMorgan

  Special case: $\neg(A \cup B) = \neg A \cap \neg B$

Notation: $\neg A \coloneqq A^c$

Random Events and Probability

But, what is probability, really?

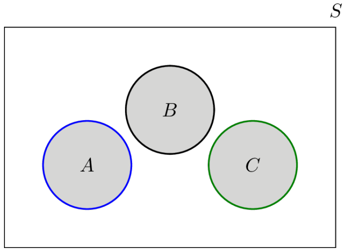
(e.g., can explain the probability of seeing an event when throwing two dice)

Mathematicians have found a set of conditions that 'makes sense'.

- Probability is a map P. ⇒ i.e., takes in an event, spits out a real value
- P must map events to a real value in interval [0,1].
- P is a (valid) **probability distribution** if it satisfies the following **axioms of probability**,

1. For any event E, $P(E) \geq 0$
2. $P(\Omega) = 1$
3. For any sequence of disjoint events $E_1, E_2, E_3, \ldots$

$$P\left(\bigcup_{i\geq 1} E_i\right) = \sum_{i\geq 1} P(E_i)$$



disjoint: intersection is empty

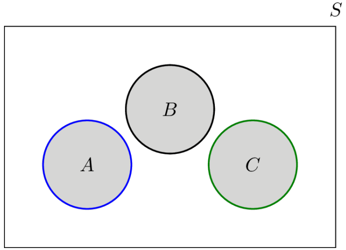
Random Events and Probability

- Many properties follows (i.e., can be proved mathematically)

$$\begin{aligned} \mathbb{P}(\emptyset) &= 0 \\ A \subset B &\implies \mathbb{P}(A) \leq \mathbb{P}(B) \\ 0 \leq \ \mathbb{P}(A) &\leq 1 \\ \mathbb{P}(A^c) &= 1 - \mathbb{P}(A) \\ A \bigcap B = \emptyset &\implies \mathbb{P}\left(A \bigcup B\right) = \mathbb{P}(A) + \mathbb{P}(B). \end{aligned}$$

E.g., throw a die. A= getting 1, B=getting an odd number

E.g., A= getting 1, B=getting 3 or 5



(I recommend that you maintain your own version of cheat sheet!)

Random Events and Probability

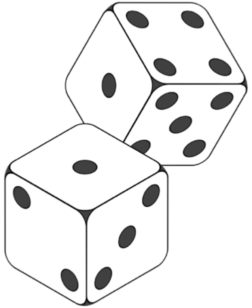
Special case

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

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

Number of elements in event set

Number of possible outcomes (36)



This is called uniform probability distribution

Q: What axiom we are using?
=> Axiom 3

(Fair) Dice Example: Probability that we roll even numbers,

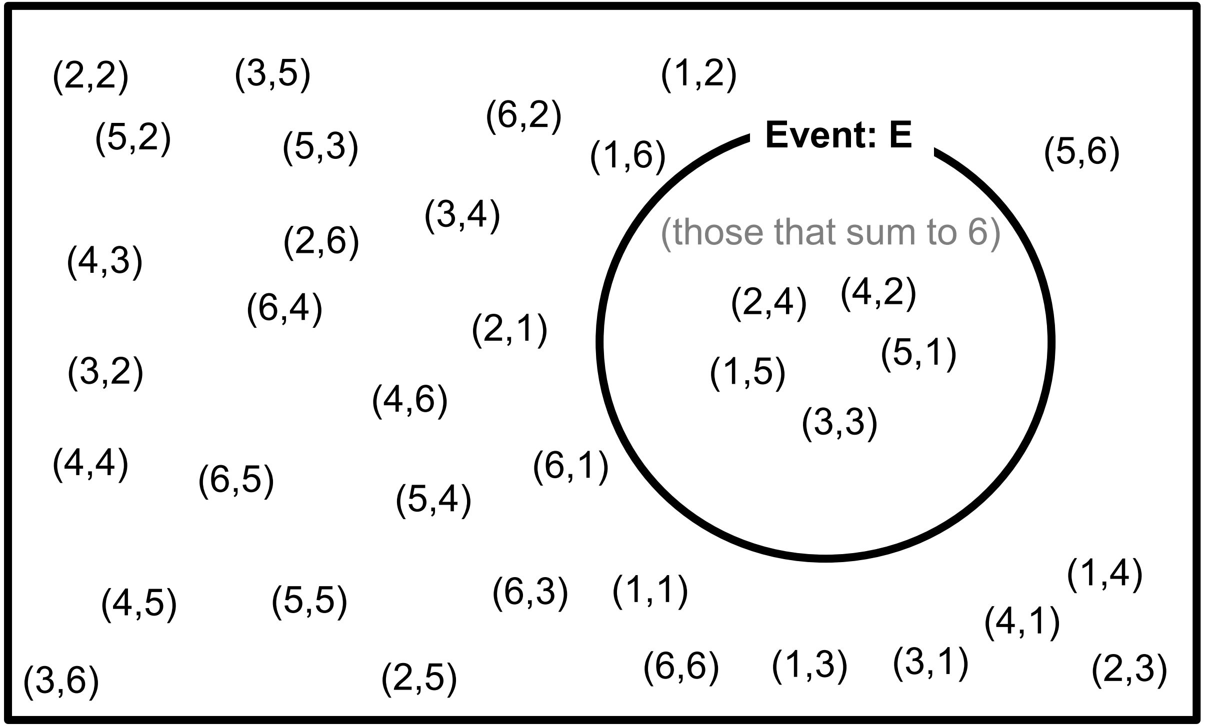
$$P((2,2) \cup (2,4) \cup \ldots \cup (6,6)) = P((2,2)) + P((2,4)) + \ldots + P((6,6))$$

$$= \frac{1}{36} + \frac{1}{36} + \ldots + \frac{1}{36} = \frac{9}{36}$$

9 Possible outcomes, each with equal probability of occurring

Random Events and Probability

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



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

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

(3,4)

(4,3) (2,6)

(6,4) (2,1)

(3,2)

(4,6)

(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)

Event: E

(those that sum to 6)

(2,4) (4,2)

(1,5) (5,1)

(3,3)

Each outcome is equally likely by the **independence**
(will learn this concept later)
=> 1/36

# of outcomes that sum to 6:
=> 5

answer:
(1/36) * 5 = 0.13888…

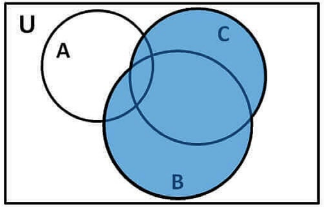
$$P(E) = \frac{|E|}{|\Omega|}$$

Set Theory: Distributive Law

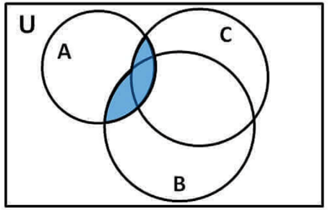
More results

- $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$ and $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$. // distributive law

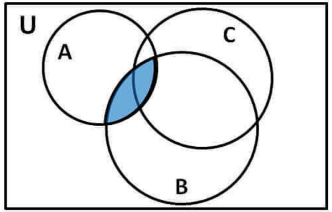
$$A \cap (\cup_i B_i) = \cup_i (A \cap B_i), \quad A \cup (\cap_i B_i) = \cap_i (A \cup B_i)$$



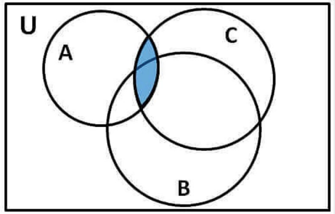
(B ∪ C)



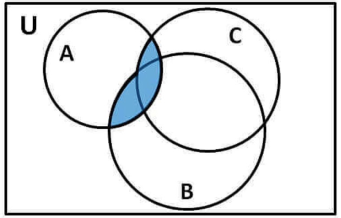
A ∩ (B ∪ C)



(A ∩ B)



(A ∩ B)



(A ∩ B) ∪ (A ∩ C)

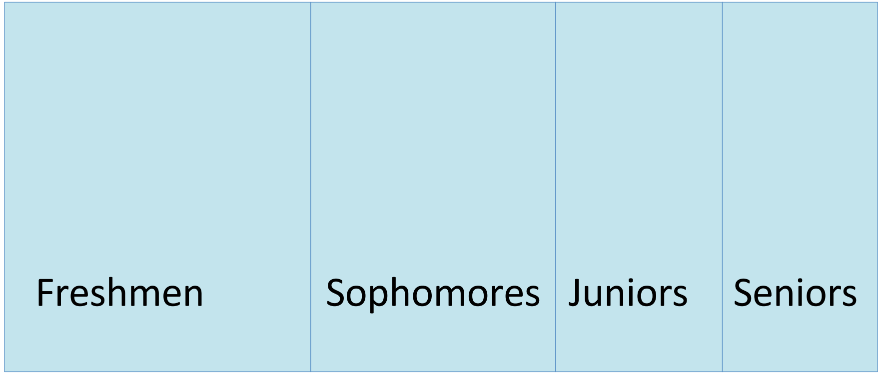
Probability as Area

- **Fact** For disjoint events E, F,

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

More generally, for pairwise disjoint events $E_1, \dots, E_n$,

$$P(E_1 \text{ or } E_2 \, .. \text{or } E_n) = P(E_1) + P(E_2) + \cdots + P(E_n)$$



Freshmen

Sophomores

Juniors

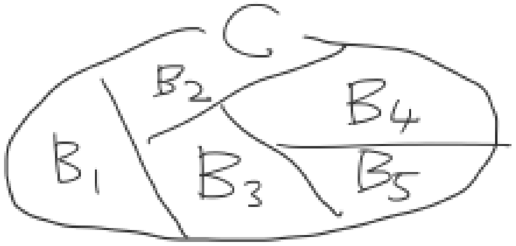
Seniors

Set Theory

Random Events and Probability

[Def] The set of events $\{B_i\}_{i=1}^n$ **partitions** outcome space $C \Leftrightarrow \cup_i B_i = C$ and $B_1, B_2, \ldots$ are disjoint.

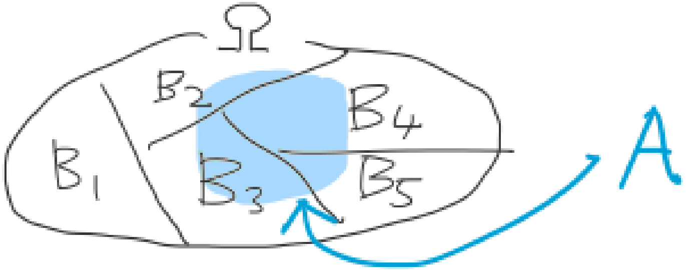
$$A = A \cap \Omega = A \cap (\cup_i B_i) = \cup_i (A \cap B_i)$$



Q: Why is this true?

A: Axiom 3 + distributive law!

Now, $\{A \cap B_i\}_{i=1}^n$ partitions A



Summary: calculating probabilities

- Most of the rules we learned is basically
  set theory + Rule 3b

So, here is a generic workflow for computing P(A).

1. Use set theory and slice and dice A into a manageable partition of A where P(each piece of partition) is easy to compute.
2. Apply Rule 3b.

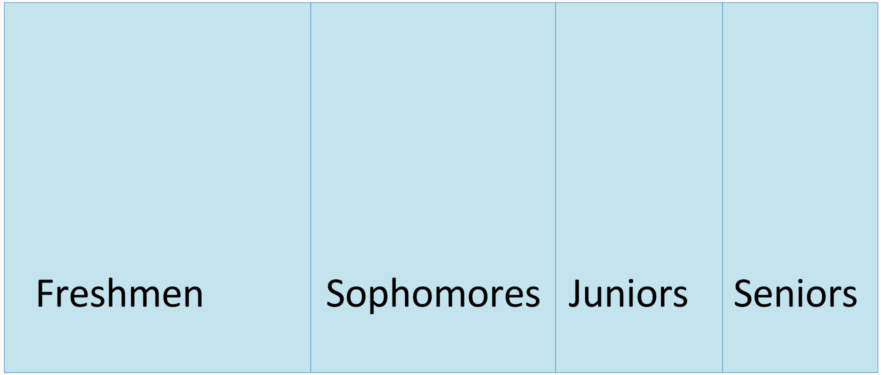
Distributive Law

- Similar to

$$a\left(\sum_{i=1}^{n} x_i\right) = \sum_{i}^{n} a x_i$$

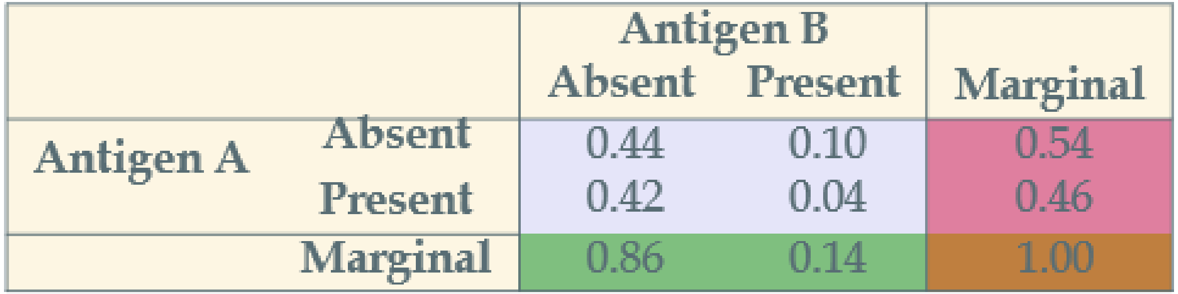
Partition

- We say that 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.}, {Soph.} form a partition of {Lower division}
  - {Lower division}, {Upper division} form a partition of S



- In general, any partition of S do not leave any element out

Exercise: Blood types



		Antigen B		
		Absent	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

Table: Probability Estimates for U.S. Blood Types

- If A is the event “presence of antigen A”, and B is the event “presence of antigen B”, what is:
  - $P(A)$?
  - $P(A \cap B)$? What is this event in words?
  - $P(A^C \cap B)$? What is this event in words?
  - $P(B^C)$? What is this event in words?