Computer Science

CSC380: Principles of Data Science

Probability 1

Xinchen Yu

Outline

- What is probability?
- Events
- Calculating probabilities
- Set Theory
- Law of Total Probability

What is probability?

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.

Outcome, Event and Probability

Outcome

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

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

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

- Example 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, ….

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.

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).
  - The proposition is 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**

Special events

- The sample space (S) includes all possible outcomes.
  - If an event E = S, then no matter what outcome occurs it's always in E.
  - e.g., E = {Heads, Tails}

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

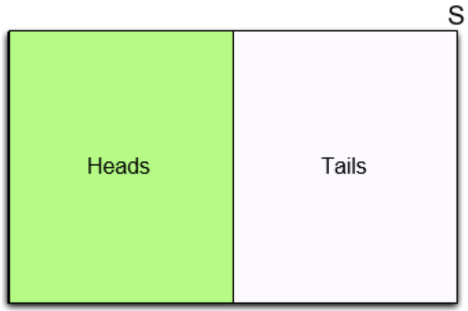
Calculating Probabilities

Calculating probability using symmetry

- If we have a sample space for which every outcome is equally likely, then we can find event probabilities easily.

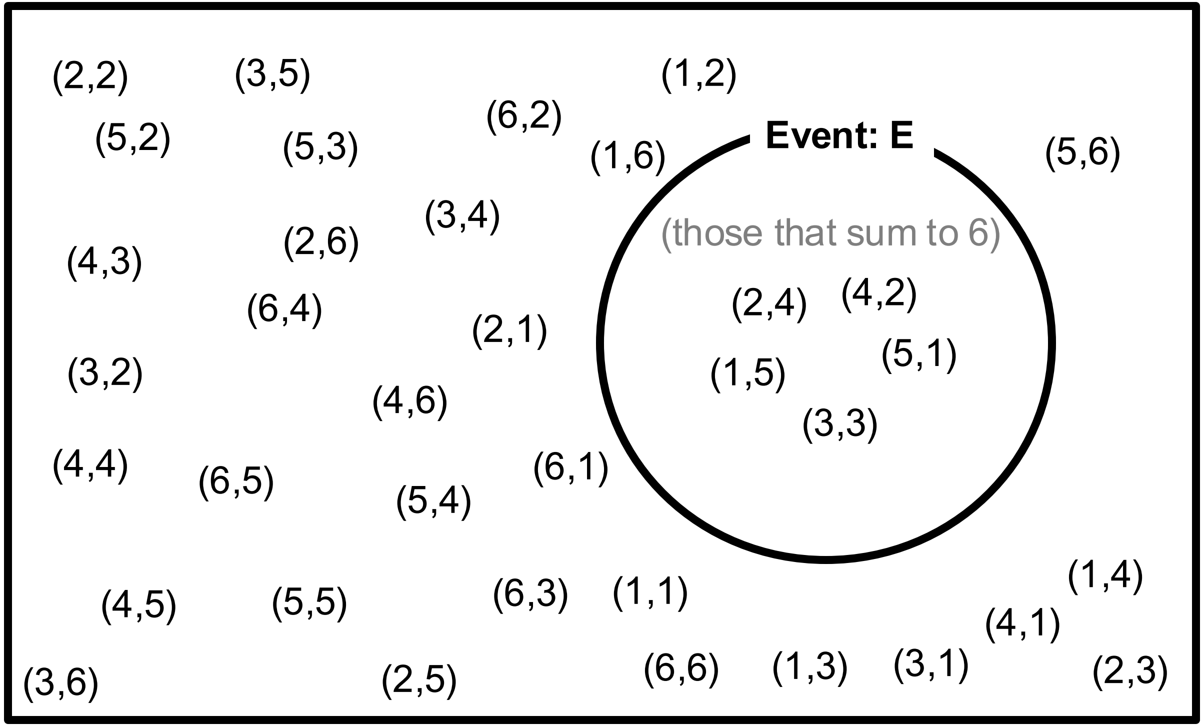
- Since every outcome has the same "area", we can just count:

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



Random Events and Probability

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



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

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

Each outcome is equally likely

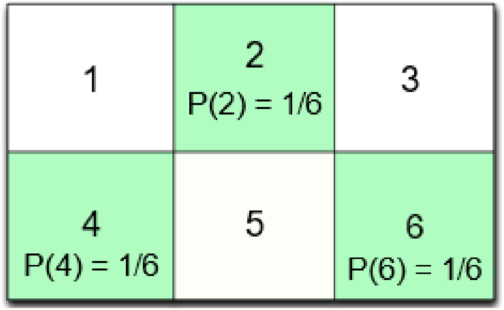
\# of outcomes that sum to 6:
5

answer:
5/36 = 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(Even) = P(2) + P(4) + P(6)



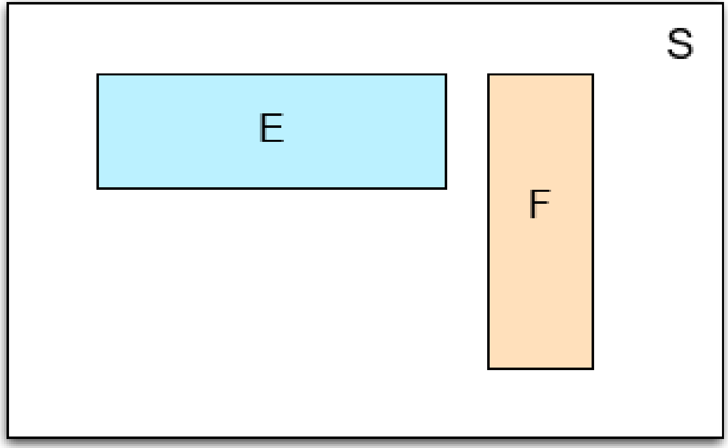
- These pieces are called ‘elementary events’
  - Events that correspond to exactly one outcome

Probability as Area

- In general, breaking an event into *disjoint events* preserves the total probability
- **Disjoint**: E and F are *disjoint* if they cannot happen simultaneously,
- e.g. E = {2, 4}, F={1, 3, 5}

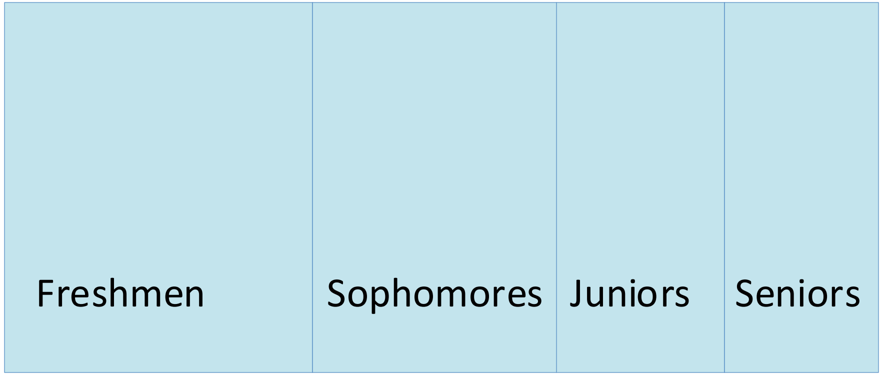
- In such cases,

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



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)



Probability as Area

For disjoint events E, F:

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

More generally, If disjoint events $E_1, \ldots, E_n$ forms a partition of :

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

For event F(law of total probability):

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

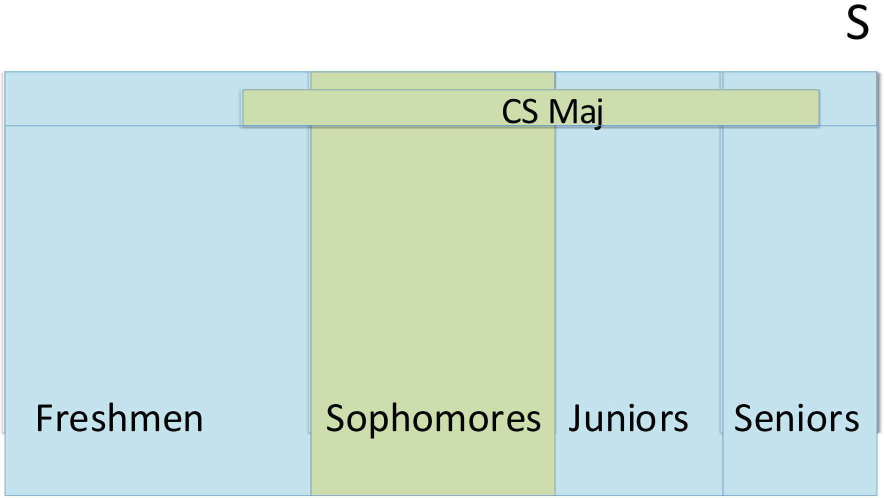
Notation: $P(A, B)$ is a shorthand for $P(A \text{ and } B)$

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.
  - $P(\text{Non-Sophomore}) = 1 - P(\text{Sophomore})$

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

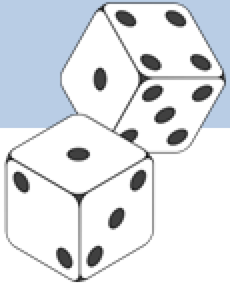
$$E_1 = \{(1,1),(1,2),\ldots,(1,6)\}$$

E_2 : *Second* die rolls 1

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

- 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

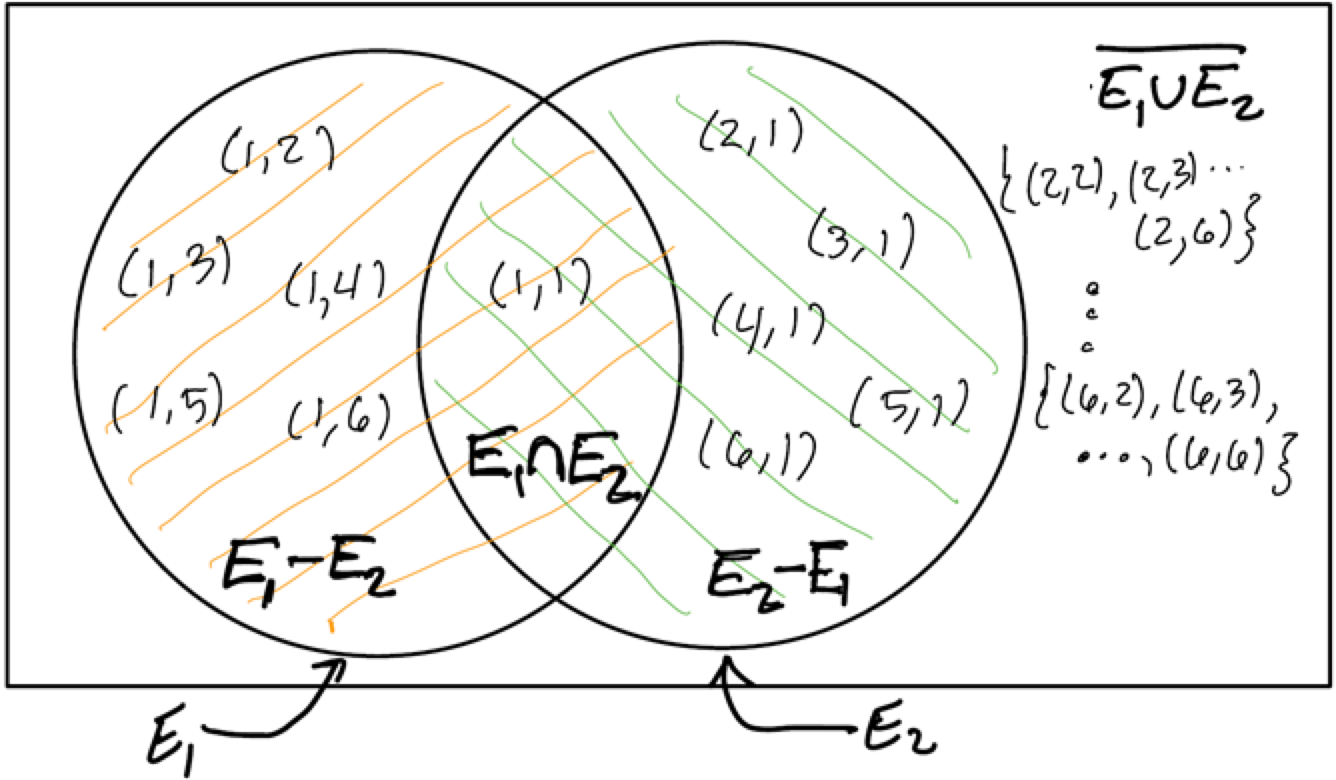
$E_1 = \{(1,1),(1,2),\ldots,(1,6)\}$ $E_2 = \{(1,1),(2,1),\ldots,(6,1)\}$

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

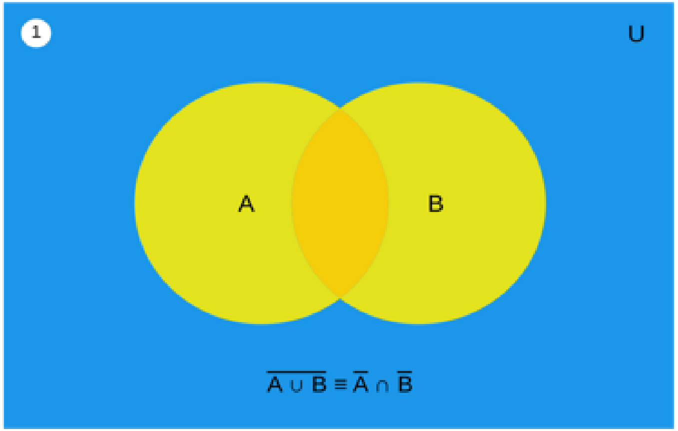
Set operations

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:

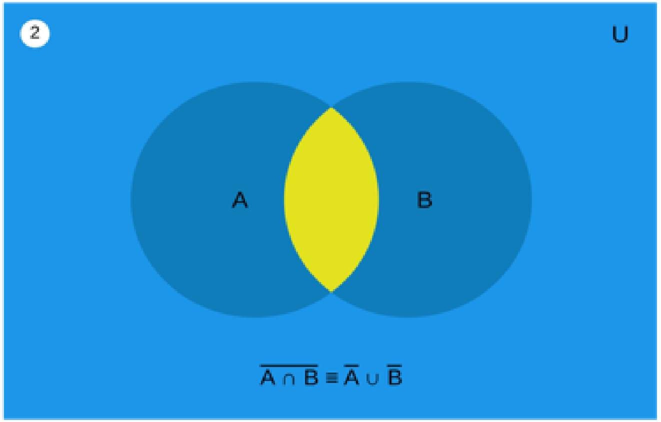
- A: I bring my cellphone
- B: I bring my laptop
- A^C: I don't bring my cellphone
- B^C: I don't bring my laptop

- $A \cup B$: 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 $(A \cap B)^C = A^C \cup B^C$

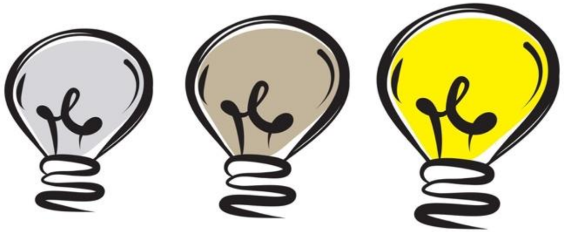
Ex: try to make sense of it using the same example above



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

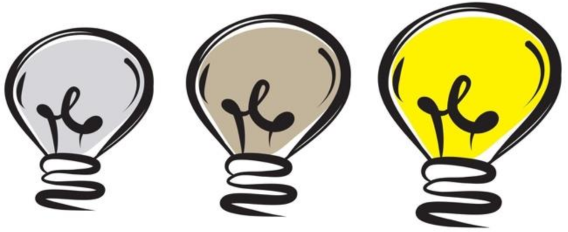
$$E_1 \cup \cdots \cup E_n =: \cup_{i=1}^{n} E_i$$

- 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:**



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

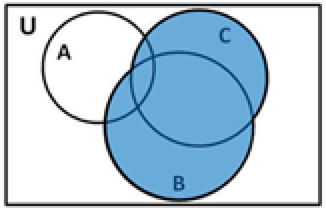
Not (at least one bulb is on) All bulbs are off

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

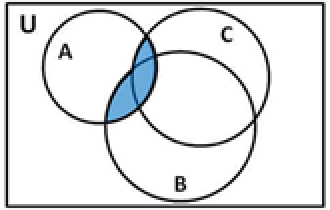
Not (all bulbs are on) At least one bulbs is off

Set operation: Distributive law

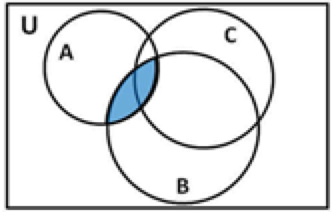
- Distributive law in arithmetics $a(x+y) = ax + ay$ carry over to sets
- **Distributive Law 1** $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$



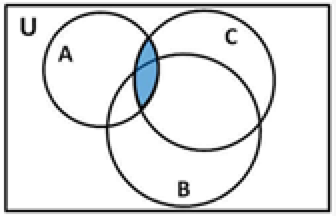
(B ∪ C)



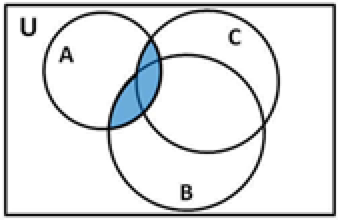
A ∩ (B ∪ C)



(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:

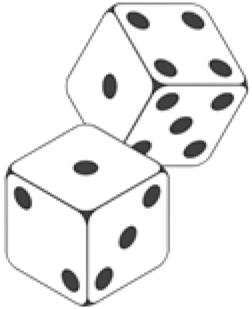
Rules of Probability

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 finite, then the probability of event is:



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

Number of elements in event set

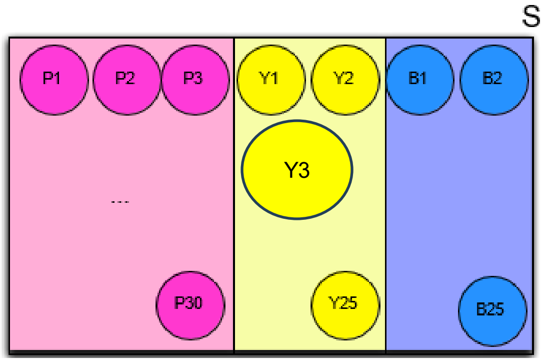
Number of possible outcomes (e.g. 36)

This is called classical probability model

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

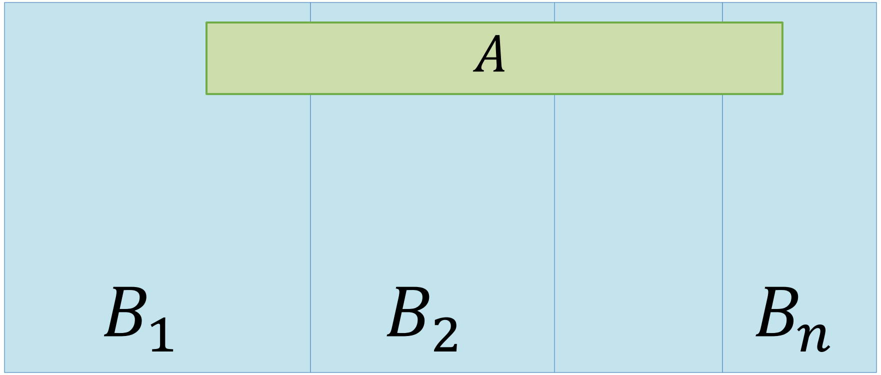


Law of Total Probability

Law of Total Probability

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

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