香港中文大學（深圳）
The Chinese University of Hong Kong, Shenzhen

DDA2001: Introduction to Data Science

Lecture 3: Elementary Probability Theory

Zicheng Wang

Announcement

- Our exams will cover conceptual and computational questions.
- 80% of computational questions are selected from a subsequently published problem set, with only the numbers changed.

Probability Theory

- A mathematical framework to model and quantify the uncertainty in real world.

- Example: I randomly select a student and ask for her/his birthday. I am not sure about her/his answer.

- Probability theory can help quantify the likelihood of each answer.

Consider a more complex situation…

- Many students registered for this class

- What is the likelihood/probability that at least two students have the same birthday?

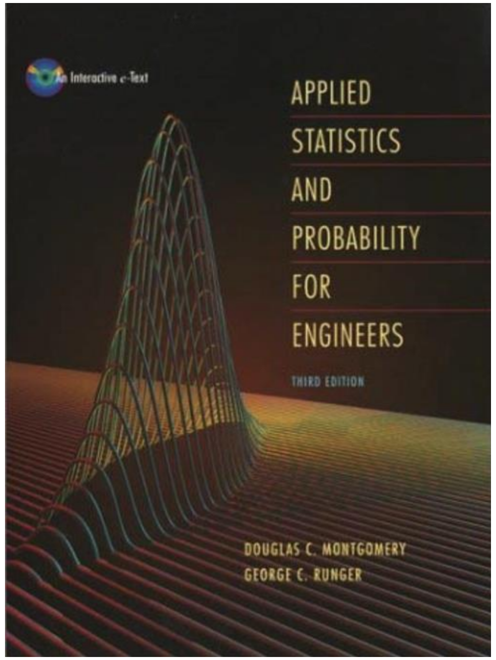
Probability and Statistics

- The basic problem that we study in probability is: Given a data generating process, what are the properties of the outcomes?

- The basic problem of statistical inference is the inverse of probability: Given the outcomes, what can we say about the process that generated the data?

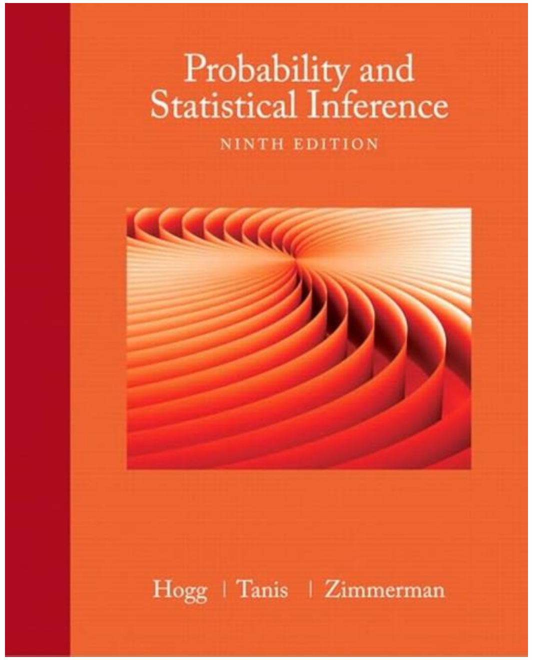
Suggested reading material:

- Applied Statistics and Probability for Engineers, Third Edition, Douglas C. Montgomery and George C. Runger.
- Chapter 2.1 – 2.3



Suggested reading material:

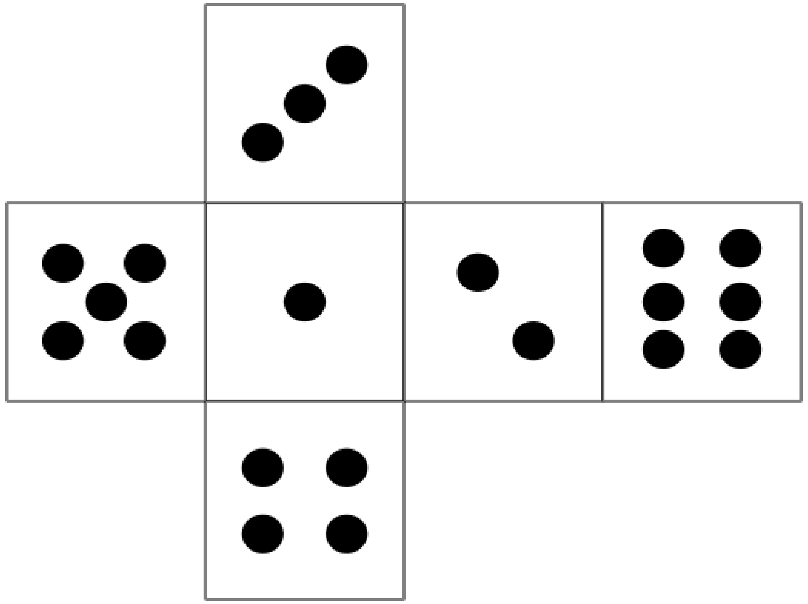
- Probability and Statistical Inference, 9th edition, Hogg, McKean and Craig.
- Chapter 1

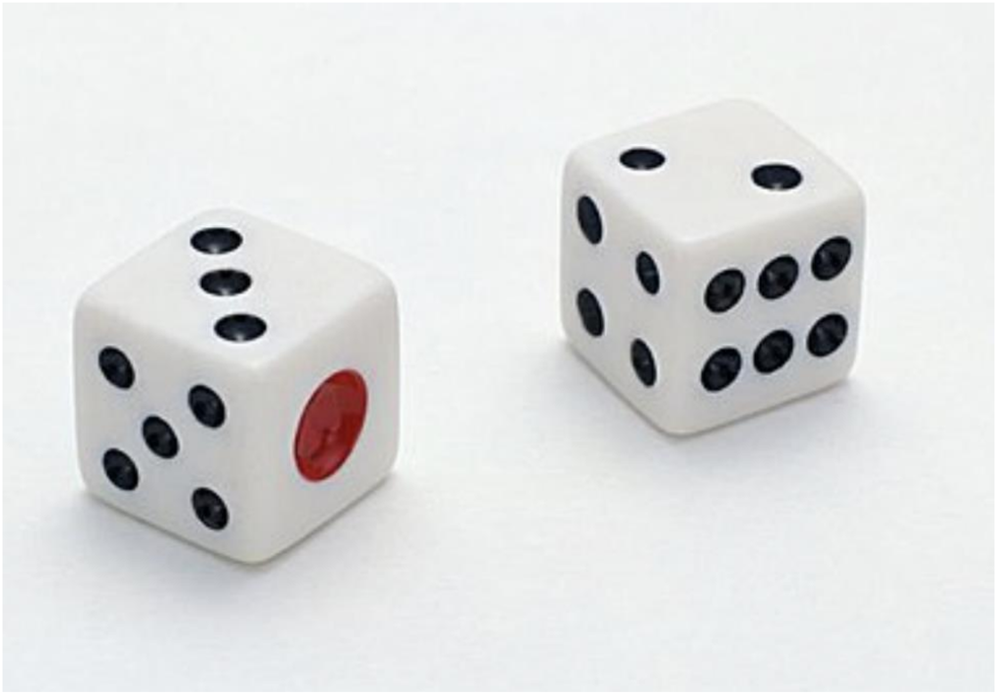


1. Definition of Probability

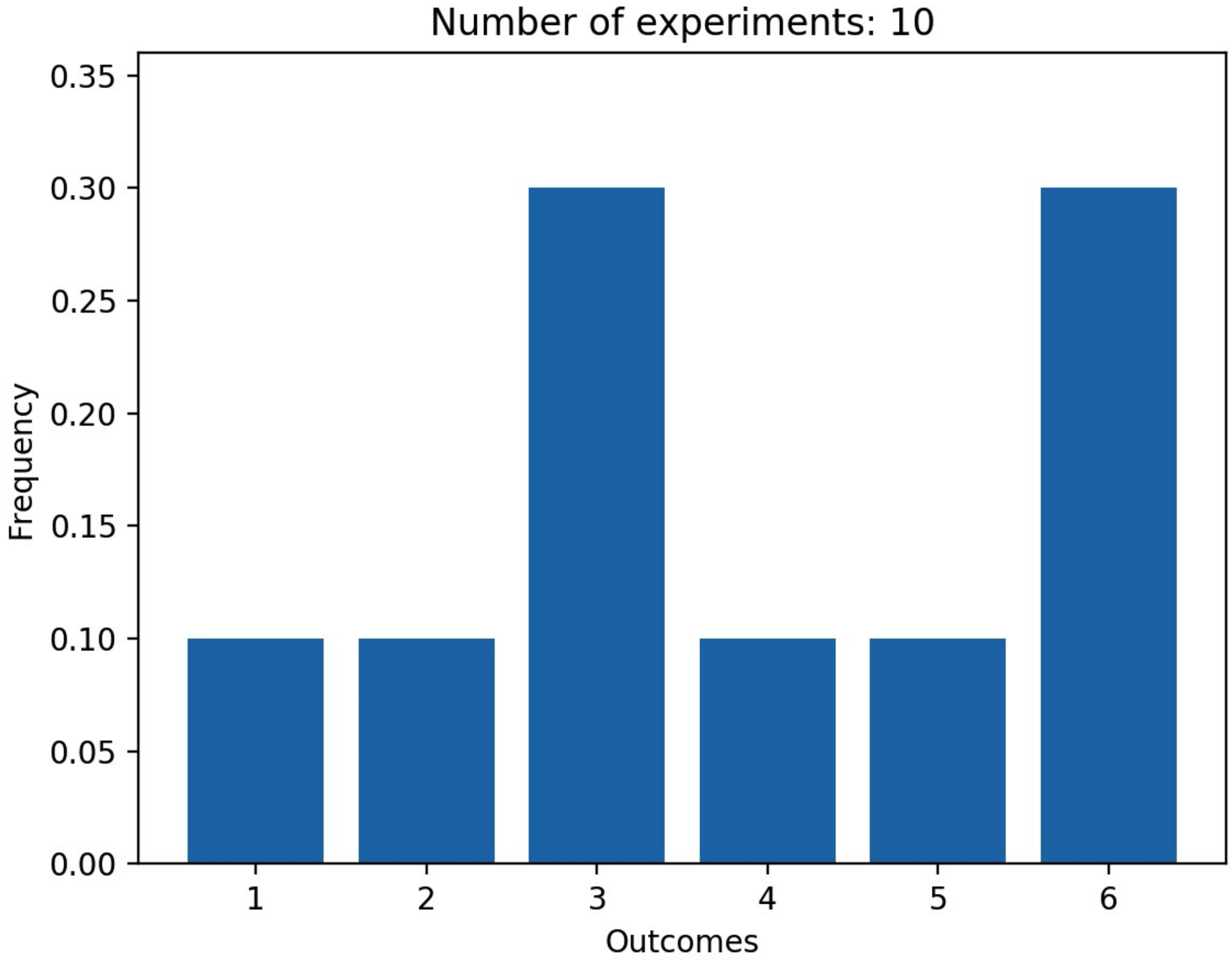
Given model, predict outcome.

- Suppose you roll the dice 10 times, how many number of 2 you will get? Guess a best number.

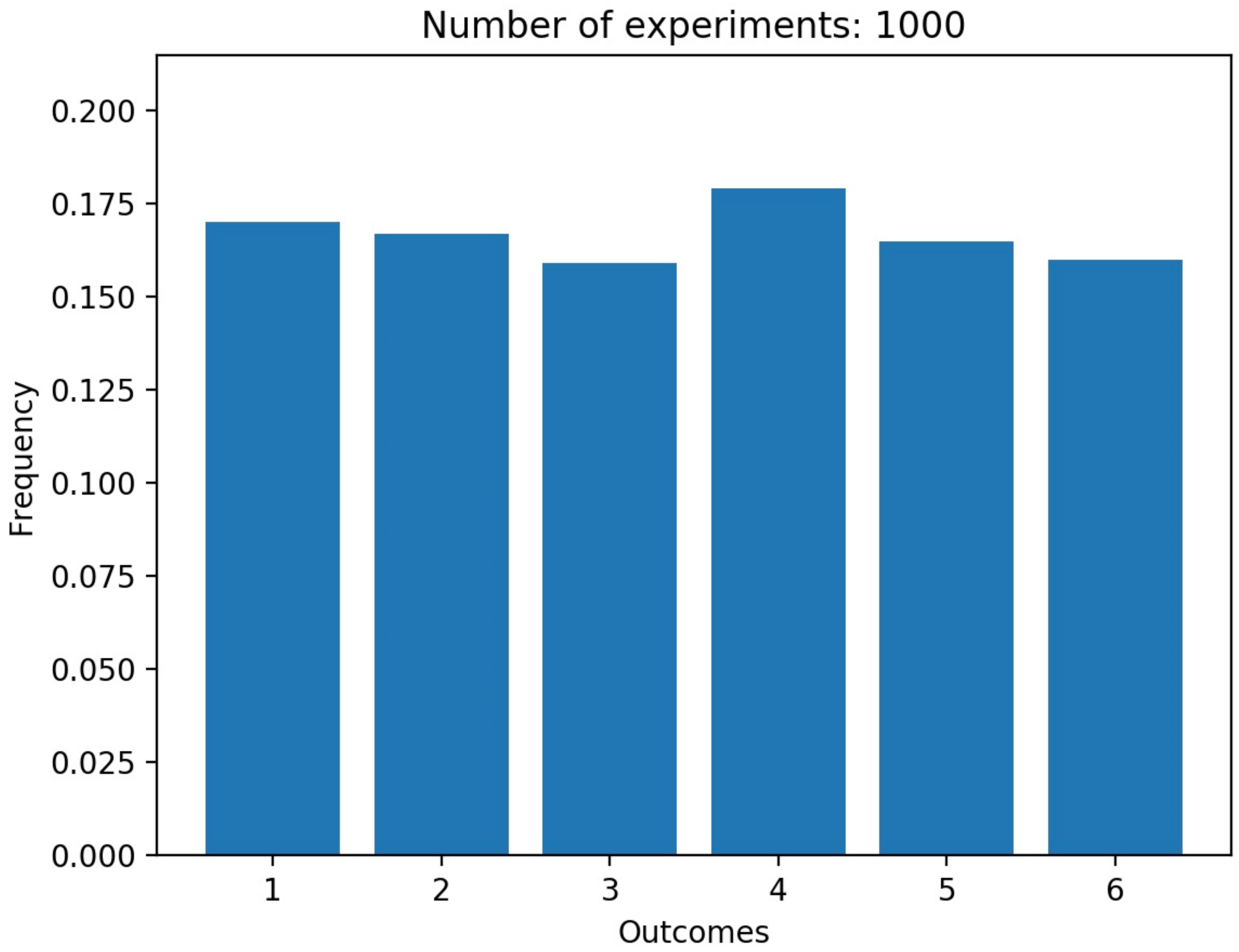




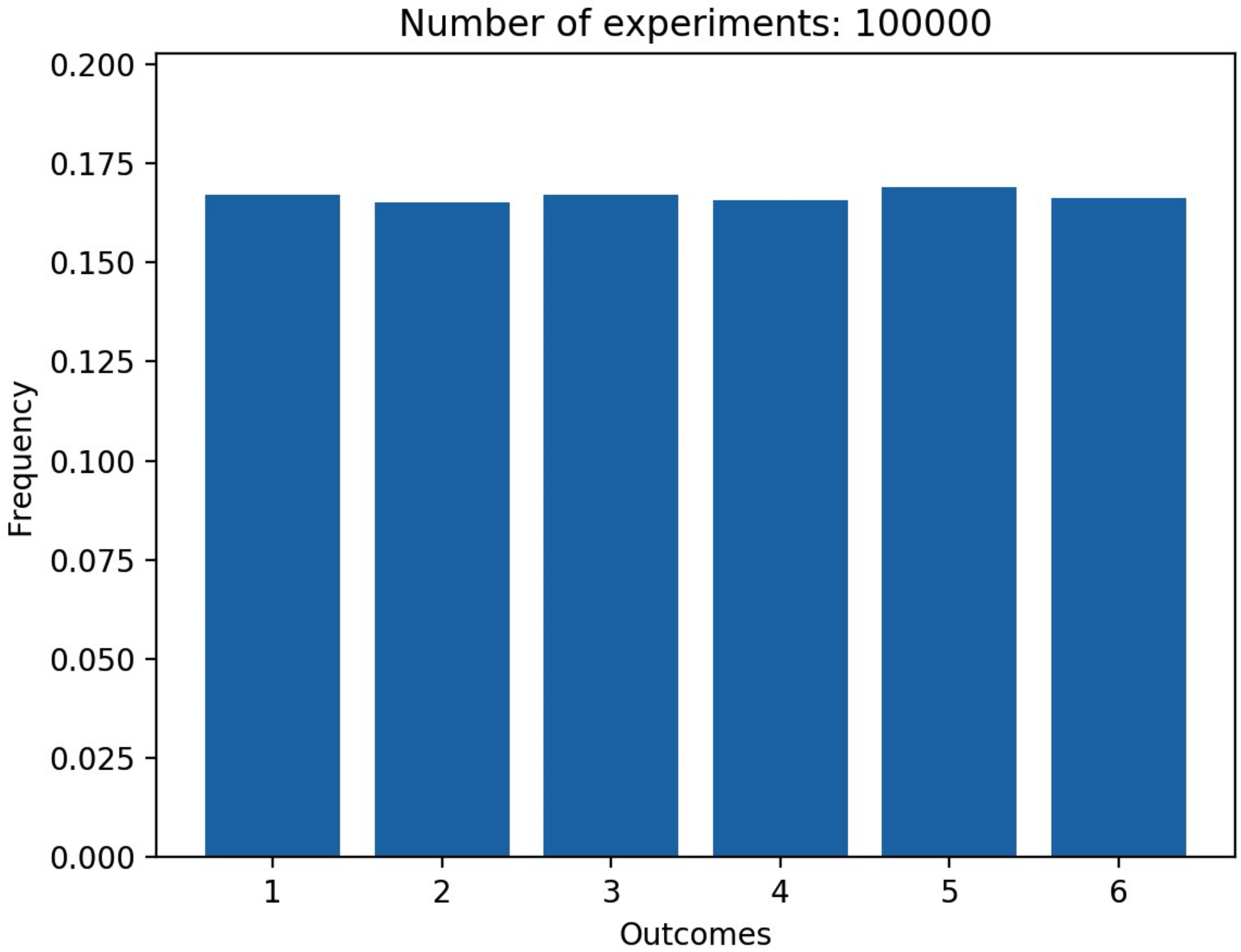
Roll a fair dice 10 times



Roll a fair dice 1000 times

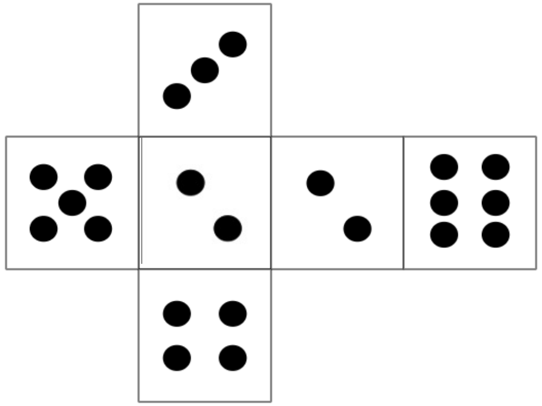


Roll a fair dice 100000 times



Example 2

- Experiment: roll an unfair dice 600000 times.



- How many number of 2 will you get?

What do they have in common?

- We are doing an experiment that can result in different outcomes.

- If repeating the experiment infinitely many times, the frequency of getting a certain outcome will converge to a certain number.

What is Probability?

- An experiment that can result in different outcomes, even though it is repeated in the same manner every time, is called a **random experiment**.

- Probability is used to quantify the likelihood, or chance, that an outcome of a random experiment will occur.

What is Probability?

- Random Experiment:

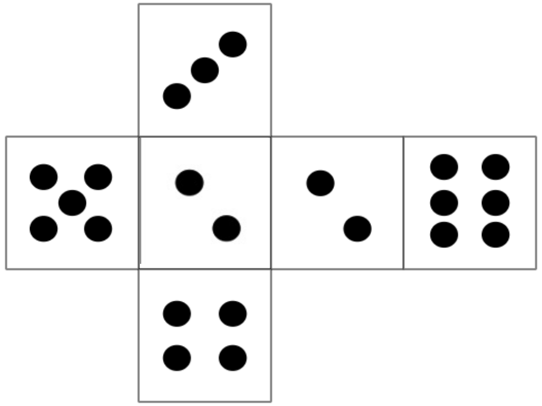
- Consider one possible outcome： ω
- The outcome ω happens with probability $P(\omega)$

- **It means:**
    - If we repeat such experiment **N** times
    - We observe **n** observations that the outcome is ω.
    - Then if N goes to infinity, $\boldsymbol{n/N}$ will approach $P(\omega)$.

Example 1

- Experiment: roll a fair dice once
- All possible outcomes: 1,2,3,4,5,6.
- What is the probability that the outcome is i, denoted by p_i, for i=1,2,3,4,5,6
- You roll N times, you observe N_i times where the outcome is i. When N goes to infinity, N_i/N will be p_i.
- As N_i/N approaches 1/6, $p_i = 1/6$

Example 2



- Experiment: roll an unfair dice
- What is the probability that the outcome is 2.
- You roll the dice N times, you observe N_2 times where the outcome is 2. When N goes to infinity, N_2/N will be p.
- As N_2/N approaches 2/6, p =1/3.

Questions

- Can a probability be negative?
- Can a probability be larger than 1?
- When rolling a fair dice, what is the probability of getting an outcome >3?

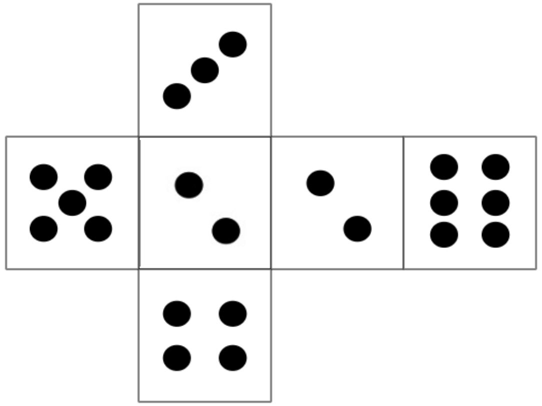
Terminologies

- **Random Experiment**: a repeatable procedure
- **Sample space**: set of all possible outcomes Ω.
- **Event**: a subset of the sample space.
- **Probability function**, P(ω): gives the probability for each outcome $\omega \in \Omega$
  - Probability is between 0 and 1
  - Total probability of all possible outcomes is 1.
  - If $A=\{\omega_1, \omega_2, \omega_3, \ldots\}$, $P(A)=P(\omega_1)+P(\omega_2)+P(\omega_3)+\cdots$

Example 1

- Experiment: roll a fair dice, report the number
- Sample space: S = {1,2,3,4,5,6}.
  - An outcome is a single result of a random experiment, incorporating the full information of your observation.
  - You can use letters, numbers, or other symbols to represent one outcome.
- Probability: P({i})= P(i) = 1/6, i=1,2,3,4,5,6.
  - We delete {}, if it contains a single outcome.
- Event A: the number is > 3: {4,5,6}
  - P(A) = P(4)+P(5)+P(6) = 3/6=1/2

Example 2



- Experiment: roll an unfair dice, report the number
- Sample space: S
- Probability: P(i) for $i \in S$.
- Event A: the number is < 4:
    - P(A)?

Properties

- Axioms of Probability

✓ $P(S) = 1$, S is the sample space
✓ $0 \leq P(A) \leq 1$ for any event A
✓ For any two events A and B, if they do not contain any common outcome, we have

$$P(A \cup B) = P(A) + P(B)$$

Example 3

The outcome of an experiment X is always either 1 or 2.

Suppose that $P(X=1)=1/4$. What is the value of $P(X=2)$?

2. Some Remarks

2. Some Remarks

Sample Space

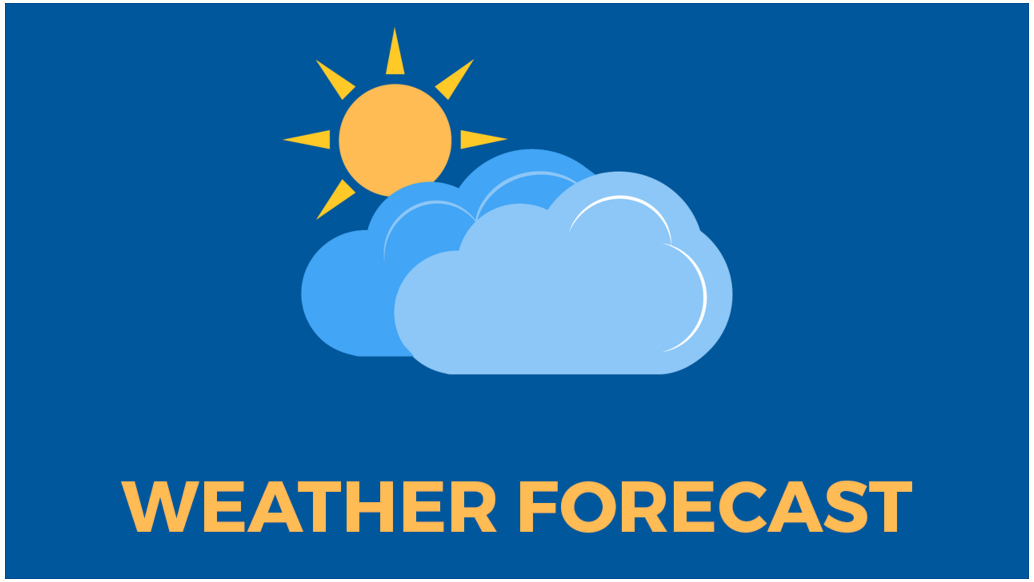
Example 4

- You bought 3 blind boxes which contain either Lord Voldemort or Harry Potter

  Describe the set of possible outcomes.

Example 5

- What will be the highest temperature tomorrow?



What’s the difference?

- Discrete or continuous: countable (listable) or not?

A sample space is **discrete** if it consists of a finite or countable infinite set of outcomes. A sample space is **continuous** if it contains an interval (either finite or infinite) of real numbers.

Exercise

- Which of the following are continuous?

1. The sum of numbers on a pair of two dice.
2. The possible sets of outcomes from flipping ten coins.
3. The possible sets of outcomes from flipping (countably) infinite coins.
4. The possible values of the temperature outside on any given day.
5. The possible times that a person arrives at a restaurant.

2. Some Remarks

Event

Events

- Events are sets:
  - ✓ Can describe in words
  - ✓ Can describe in notation
- Experiment: toss a coin 2 times.
- Event -- You get 1 or more heads

  = {HH, HT, TH}

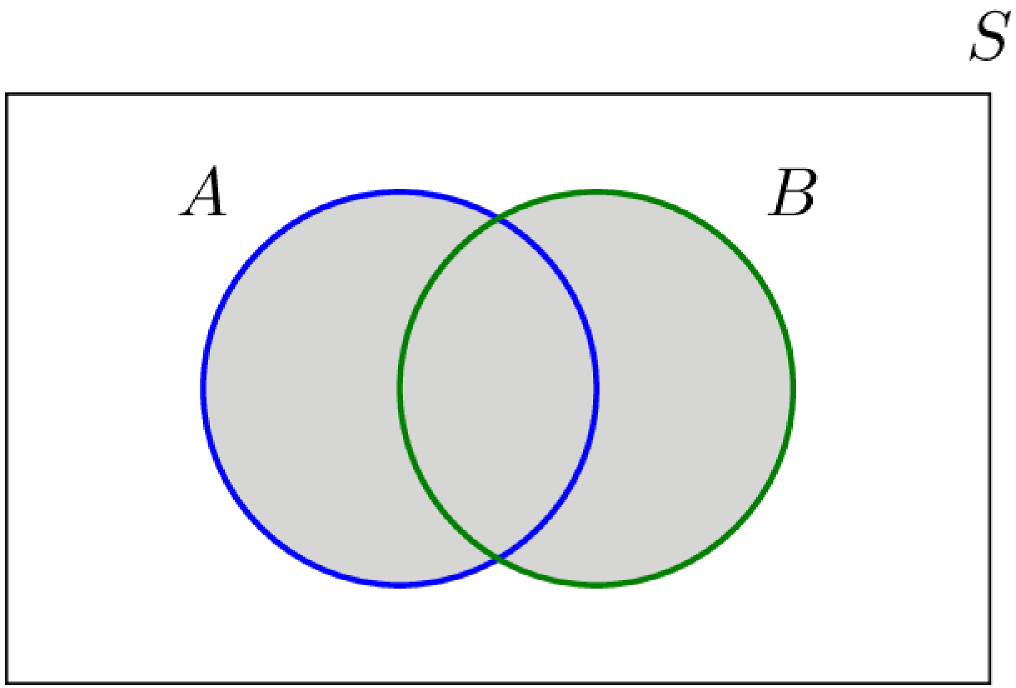
Set operations

- Events are sets, so we can use set operations

  - ✓ Unions
  - ✓ Intersections
  - ✓ Complements

Set operations - union

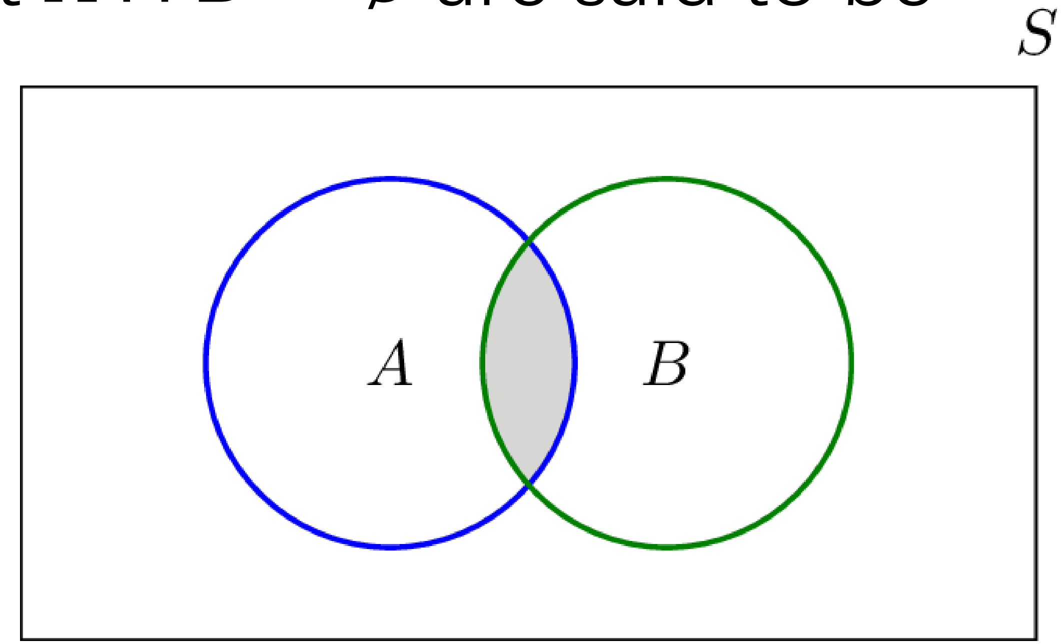
- The **union** of two events A and B is the event that consists of all outcomes that are contained in either A or B.
- We denote the union as (A or B) in words, and $A \cup B$ in notation



Venn diagrams

Set operations - intersection

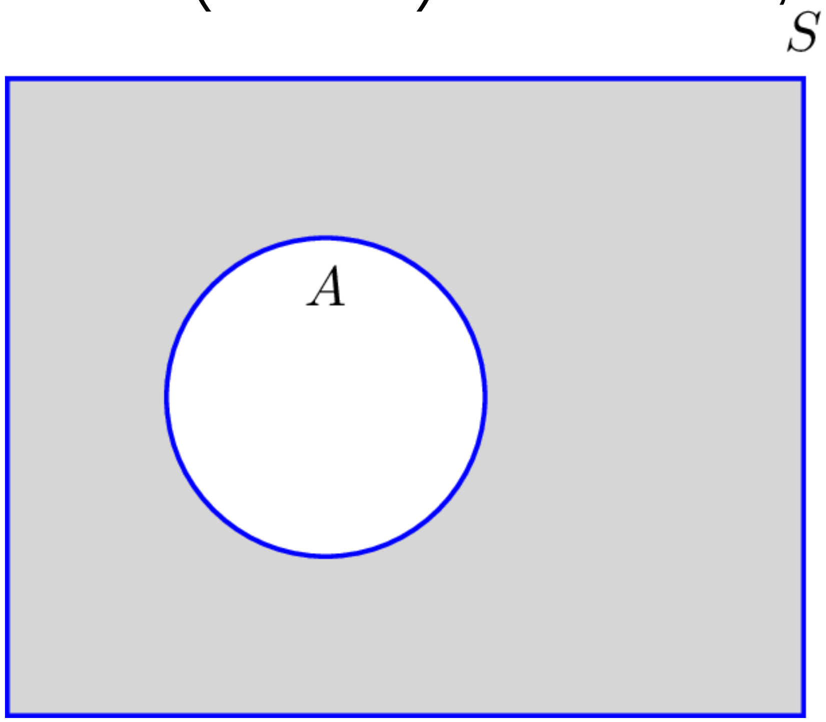
- The **intersection** of two events A and B is the event that consists of all outcomes that are contained in both A and B.
- We denote the intersection as (A and B) in words, and $A \cap B$ in notation
- Two events A and B, such that $A \cap B = \emptyset$ are said to be **mutually exclusive**.



Note: The intersection can be an empty set.

Set operations - complement

- The **complement** of an events A in a sample space S is the set of outcomes in the sample space that are not in the event A.
- We denote the complement as (not A) in words, and A' or A^c in notation



Note: The complement can also be an empty set, it can also be the whole sample space.

Example 6

- Toss a coin 2 times: sample space: {TT, HT, TH, HH}
- Event A: there is at least one Head: A = {HH, HT, TH}
- Event B: there is at least one Tail: B = {TT, HT, TH}

- Question:
  - What is the event A or B, namely ($\mathrm{A} \cup \mathrm{B}$)?
  - What is the event A and B, namely ($\mathrm{A} \cap \mathrm{B}$)?
  - What is the complement of event A?

Some useful properties

- The complement of the complement of an event is itself

$$(E')' = E$$

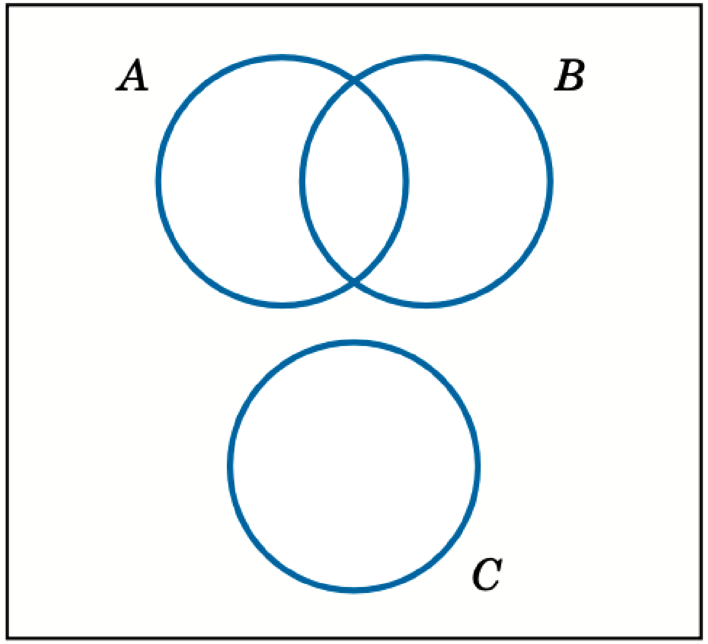
- The intersection and union does not depend on the order

$$A \cap B = B \cap A \qquad \text{and} \qquad A \cup B = B \cup A$$

- The complement of the union (intersection) is the intersection (union) of the complement

$$(A \cup B)' = A' \cap B' \qquad \text{and} \qquad (A \cap B)' = A' \cup B'$$

Exercise



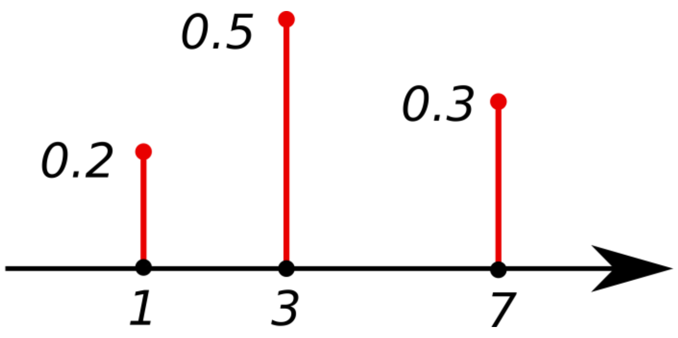
Shade the region that corresponds to each of the following events:

(a) A'
(b) $(A \cap B) \cup (A \cap B')$
(c) $(A \cap B) \cup C$
(d) $(B \cup C)'$
(e) $(A \cap B)' \cup C$

2. Some Remarks
Probability Function

Probability function

- Discrete:
  - ✓ Probability mass function.
  - ✓ P(ω): gives the probability for **each** outcome ω ∈ S



Continuous case will be defined later.

Calculation of Event Probability?

By the definition⋯

- First, you write down P(ω) for each ω ∈ S.

- Then, if A=$\{\omega_1, \omega_2, \omega_3, \ldots\}$, P(A)=P($\omega_1$)+P($\omega_2$)+P($\omega_3$) $+ \cdots$

Calculation of Event Probability?

By the definition⋯

- First, you write down P(ω) for each ω ∈ S.

Tables can really help in complicated examples!

- Then, if $A=\{\omega_1, \omega_2, \omega_3, \ldots\}$, $P(A)=P(\omega_1)+P(\omega_2)+P(\omega_3)+\cdots$

Example 7

A random experiment can result in one of the outcomes $\{a, b, c, d\}$ with probabilities 0.1, 0.3, 0.5, and 0.1, respectively. Let A denote the event $\{a, b\}$, B the event $\{b, c, d\}$, and C the event $\{d\}$. Then,

Event	a	b	c	d
Probability	0.1	0.3	0.5	.1

Example 7

A random experiment can result in one of the outcomes $\{a, b, c, d\}$ with probabilities 0.1, 0.3, 0.5, and 0.1, respectively. Let A denote the event $\{a, b\}$, B the event $\{b, c, d\}$, and C the event $\{d\}$. Then,

Event	a	b	c	d
Probability	0.1	0.3	0.5	.1

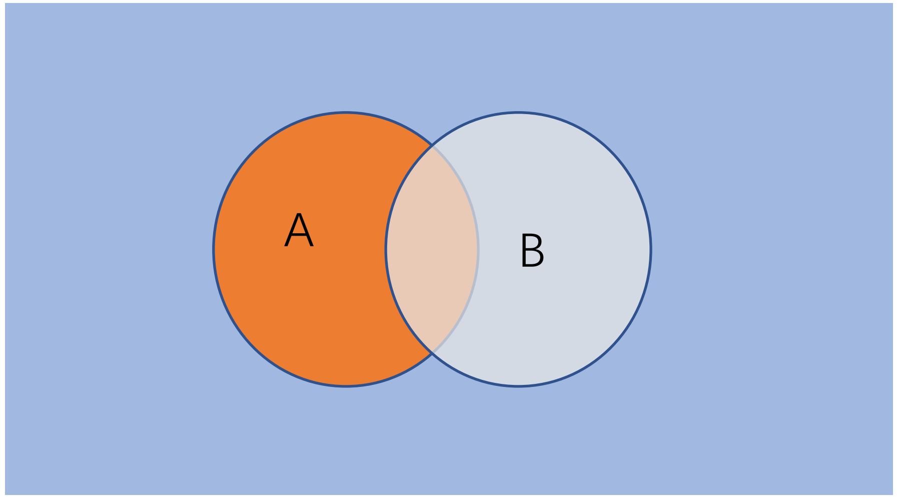
$$P(A) = 0.1 + 0.3 = 0.4$$

$$P(B) = 0.3 + 0.5 + 0.1 = 0.9$$

$$P(C) = 0.1$$

Calculation of Event Probability?

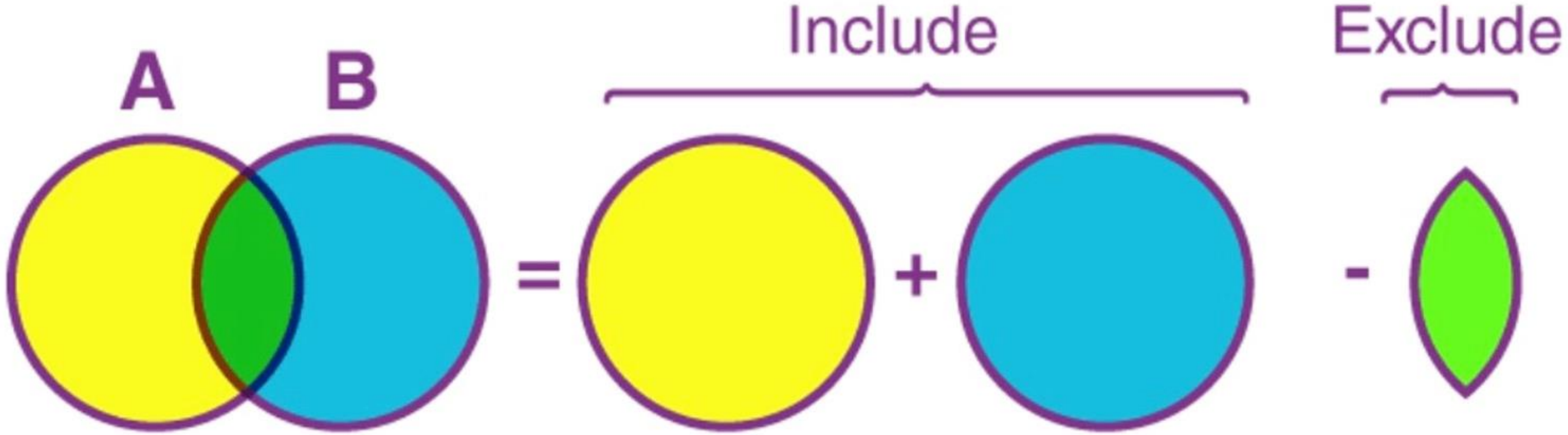
By event relationship…, say either A or B happen



How to calculate P(A or B) , namely $P(A \cup B)$?

Calculation of Event Probability?

Addition Rules

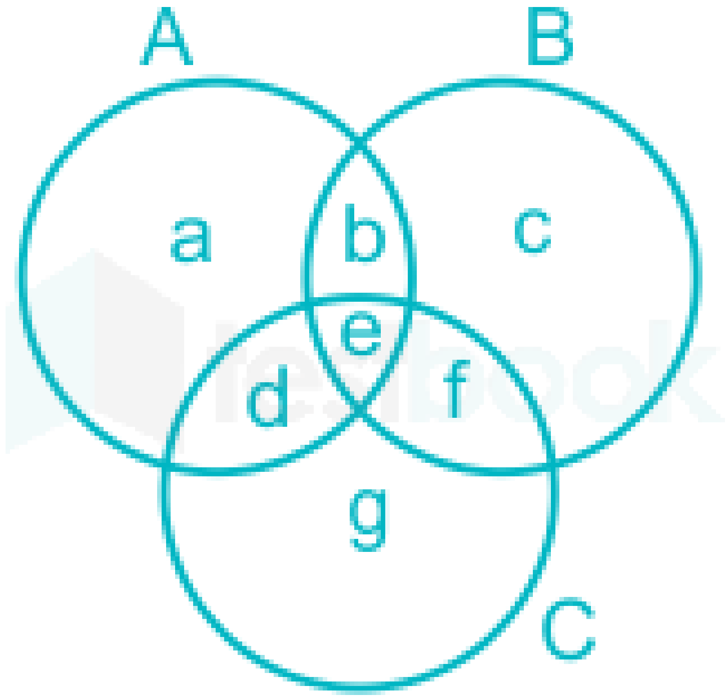


$$P(A \text{ or } B) = P(A) + P(B) - P(\ A \text{ and } B)$$

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

Addition Rules for more events?

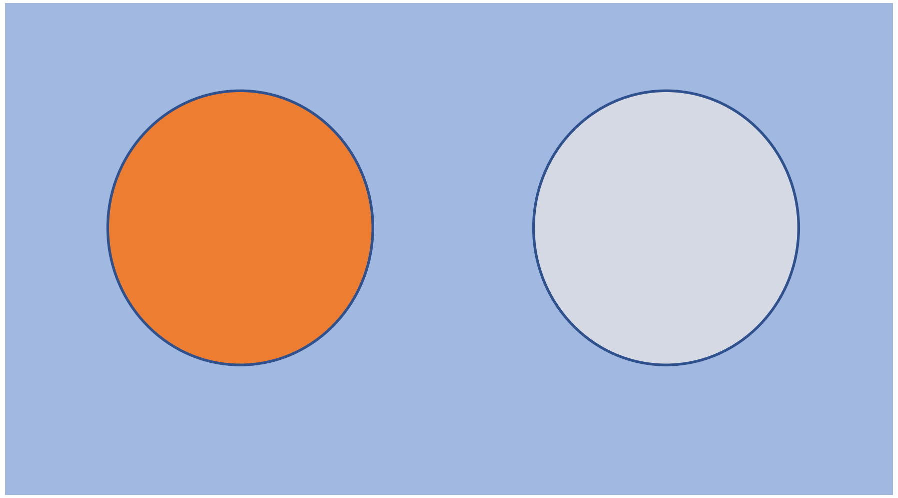
$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) \\ - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C)$$



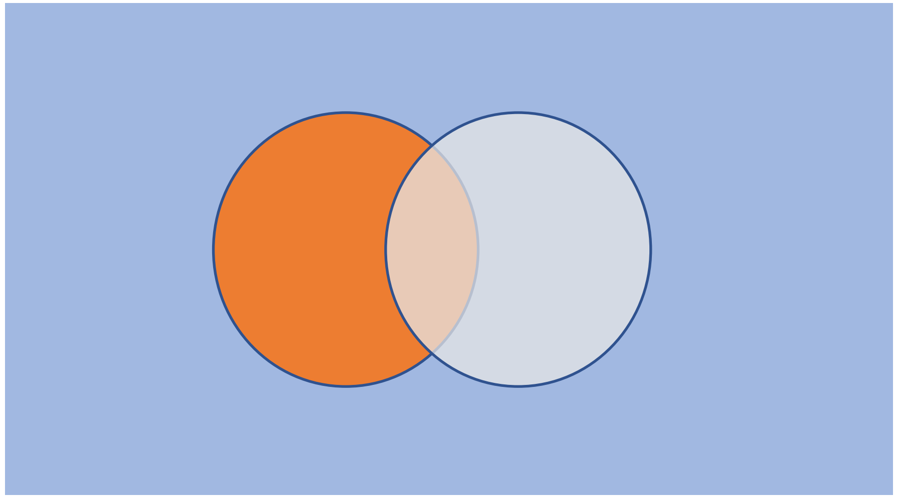
Focus on two events.

- $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

- Case: $P(A \cap B)=0$
- If A happens, B cannot happen



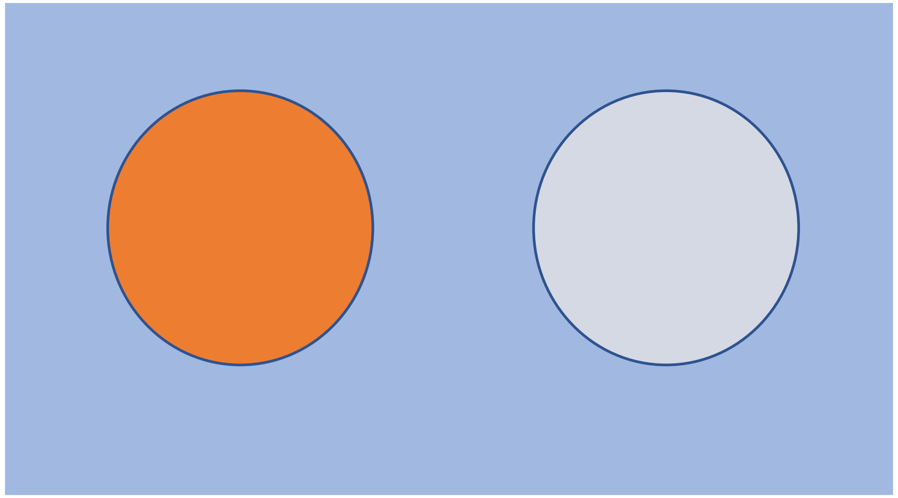
- Case: $P(A \cap B) \neq 0$
- If A happens, B may also happen



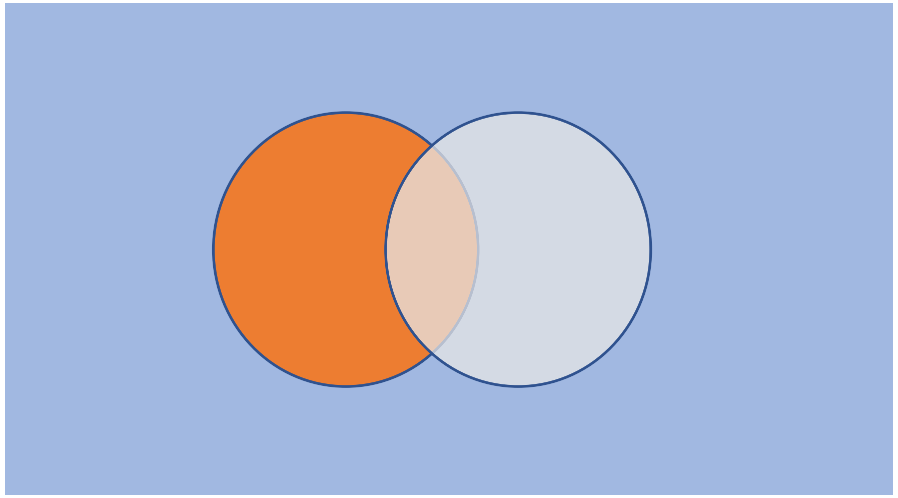
Focus on two events.

- $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

- Case: $P(A \cap B)=0$
- If A happens, B cannot happen
- A and B are **disjoint (mutually exclusive)**



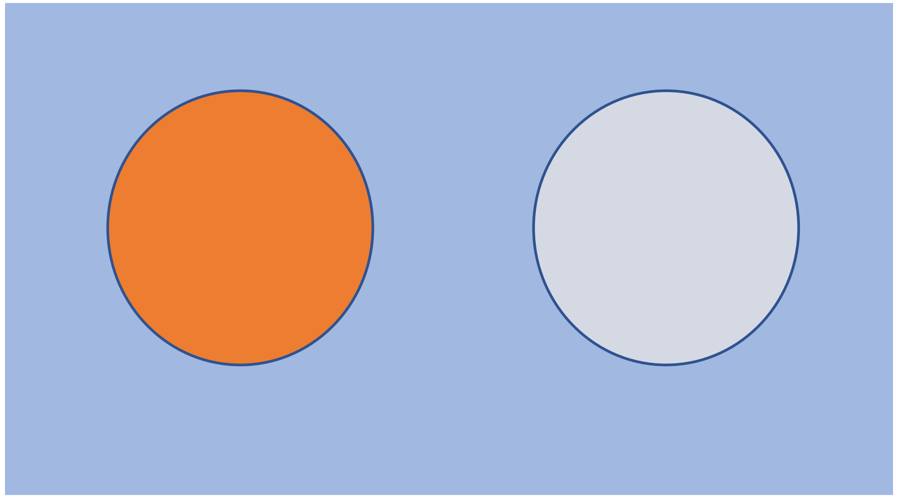
- Case: $P(A \cap B) \neq 0$
- If A happens, B may also happen
- A and B are **joint (not mutually exclusive)**



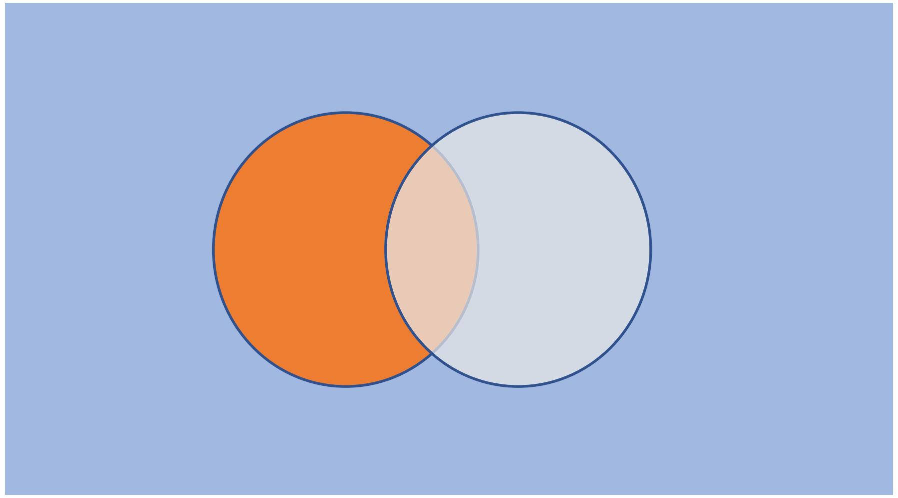
Focus on two events.

- $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

- Case: $P(A \cup B) = P(A) + P(B)$
- If A happens, B cannot happen
- A and B are **disjoint**



- Case: $P(A \cup B) < P(A) + P(B)$
- If A happens, B may also happen
- A and B are **joint**



Example 8

- Toss a coin 2 times: {TT, HT, TH, HH} equally likely

- H1: the first outcome is H
- T1: the first outcome is T
- H2: the second outcome is H

1 \ 2	T	H
T	TT	TH
H	HT	HH

Event	H1	T1	H1 or T1	H1 and T1
Probability	0.5	0.5	1	0

Event	H1	H2	H1 or H2	H1 and H2
Probability	0.5	0.5	0.75	0.25

Example 8

Event	H1	T1	H1 or T1	H1 and T1
Probability	0.5	0.5	1	0

Event	H1	H2	H1 or H2	H1 and H2
Probability	0.5	0.5	0.75	0.25

- H1: the first outcome is H
- T1: the first outcome is T
- H2: the second outcome is H

- $P(H1 \cup T1) = P(H1)+P(T1)$

  mutually exclusive

Example 8

Event	H1	T1	H1 or T1	H1 and T1
Probability	0.5	0.5	1	0

Event	H1	H2	H1 or H2	H1 and H2
Probability	0.5	0.5	0.75	0.25

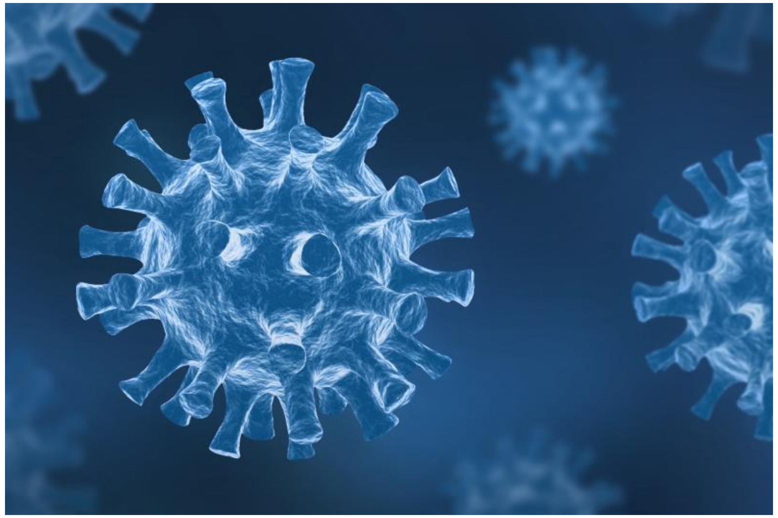
- H1: the first outcome is H
- T1: the first outcome is T
- H2: the second outcome is H

- $P(H1 \cup T1) = P(H1)+P(T1)$
- $P(H1 \cup H2) < P(H1)+P(H2)$

not mutually exclusive

Exercise: Joint or disjoint

A: Infected with COVID-19
B: Nucleic acid test is negative



Experiment: toss a coin 3 times.
A:“exactly 2 heads”
B:“exactly 2 tails