Discrete Probability

Lecture 4

Probability

- Probability theory provides a framework for quantifying uncertainty
- This is essential for drawing statistical inferences about the population beyond the sample data
- Vast majority of ML algorithms are probabilistic

- Intuitively, *probability* is a measure of likelihood for something to happen

- Numerically, probability ranges between 0 and 1:
  - probabilities near 0 correspond to unlikely events
  - probabilities near 1 correspond to almost certain events

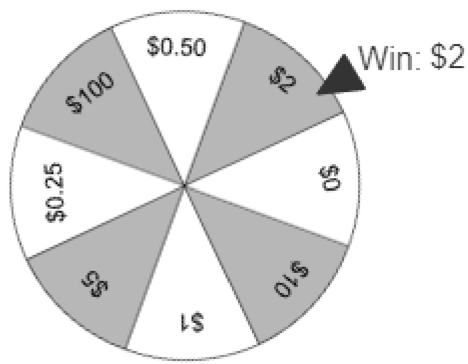
Terminology

- A *random process* (or a random *experiment*) is an action or process where an outcome is determined by chance
- The *sample space* is the set of all possible outcomes of a random process
- One *outcome* is one possible result of a random process
- (Discrete) *event* is a subset of a sample space: it is a single outcome or collection of outcomes
    - Events are typically denoted by A, B, or C

Example: random process

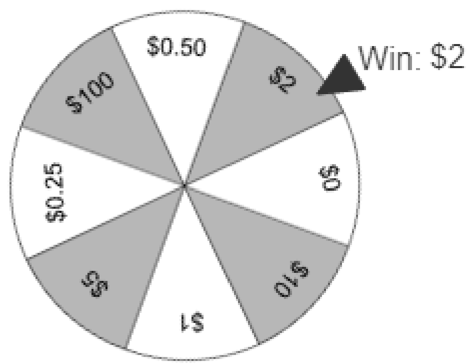
- A prize wheel has eight equally sized spaces. The wheel is spun and the amount won is the amount on the indicated winning space.

- One spin of the wheel is a *random process*.



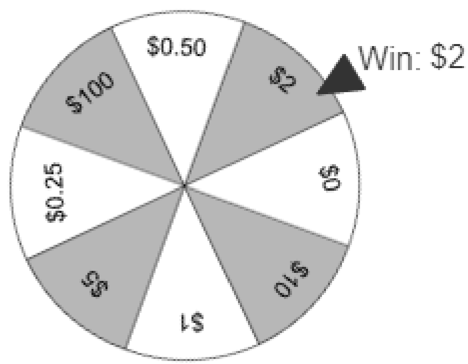
Example: sample space

- The possible prize amounts, or eight possible outcomes, are $0, $0.25, 0.50, $1, $2, $5, $10, and $100.
- Thus, the sample space is set {0, 0.25, 0.50, 1, 2, 5 10, 100}.



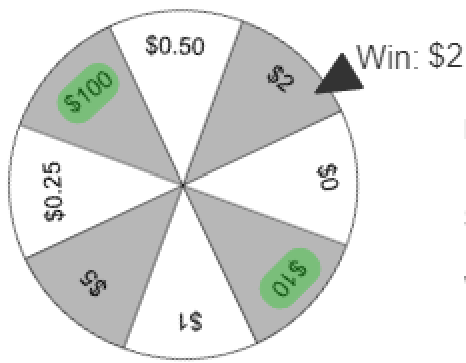
Example: single outcome

- Winning $2 after a single spin of the wheel is an example of an outcome



Example: event A

- Define event A as winning at least $10.
- Two outcomes from the sample space where event A is true: $10 and $100.
- Event A covers a subset {10, 100}

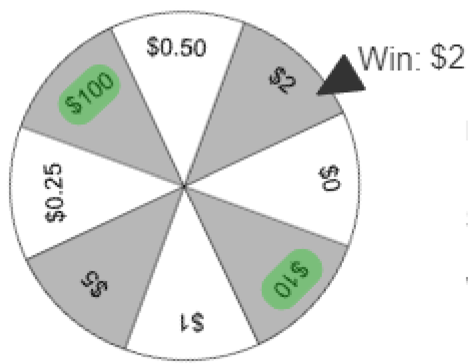


Example: probability of event A

- The probability of event A occurring on any spin is:

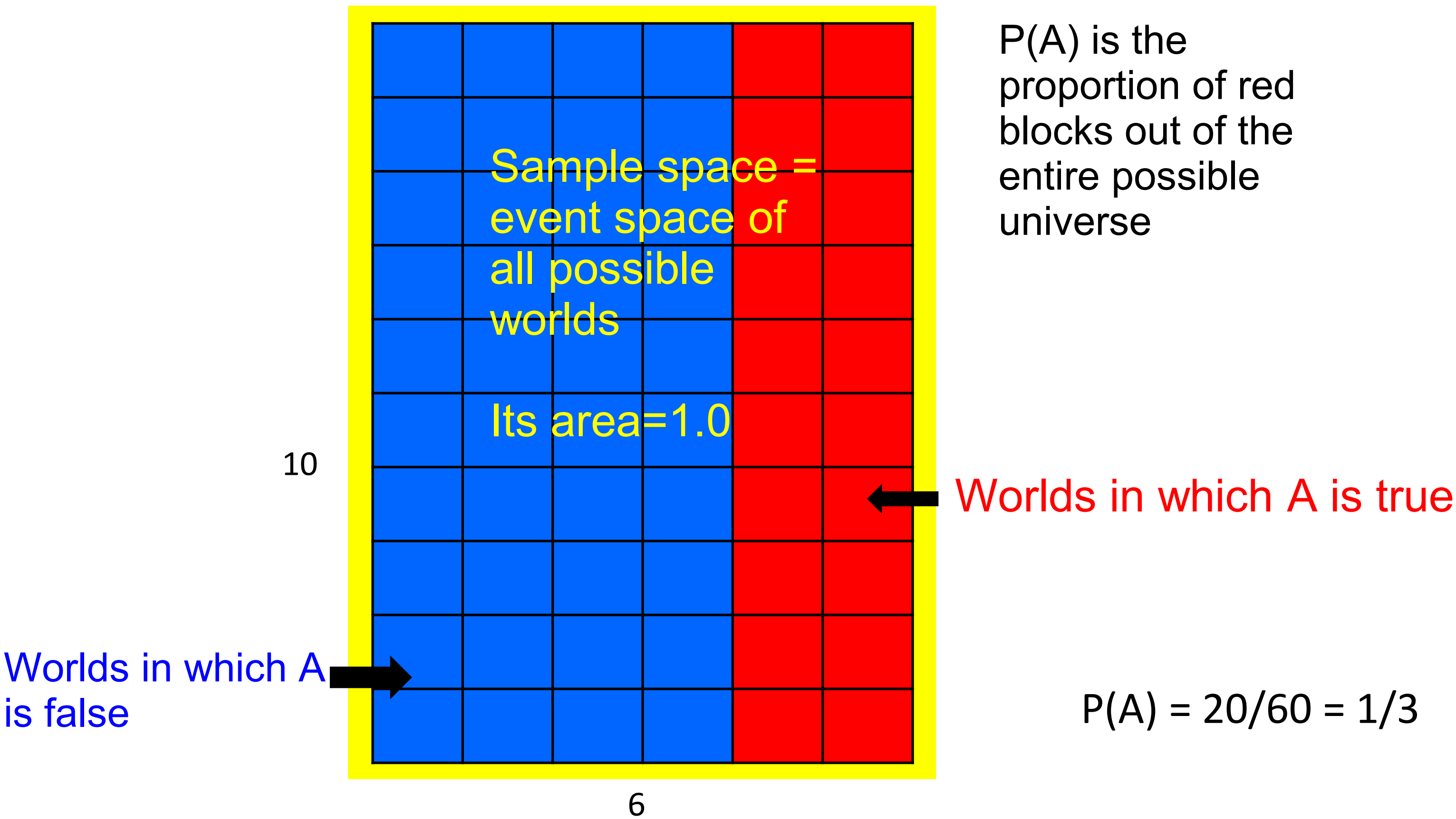
P(a) = $\frac{size(\{10, 100\})}{size(\{0, 0.25, 0.50, 1, 2, 5\ 10, 100\})} = \frac{2}{8} = 0.25$

- In practice that means that if the wheel is spun many, many times, the long-run proportion of times \$10 or \$100 is won is 25/100



Visualizing probability

- The probability of an event A, denoted P(A), is measured as the number of outcomes in A divided by the total number of equally likely outcomes in the entire sample space.



Terminology quiz

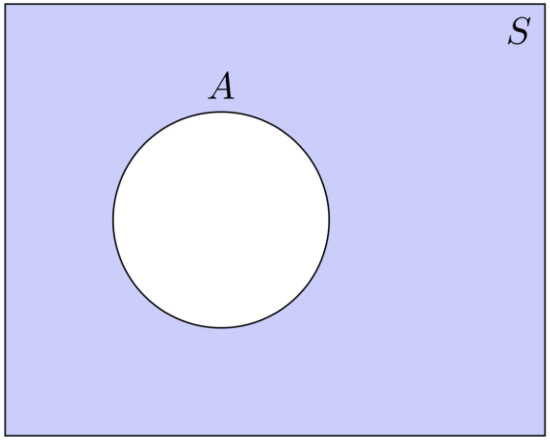
The following table gives the distribution of US household size (number of people in the household) based on the 2020 Current Population Survey.

Size	1	2	3	4	5	6	7+
Proportion	0.29	0.35	0.15	0.12	0.06	0.02	0.01

Derived events: complement of a set

We can define new events based on given events

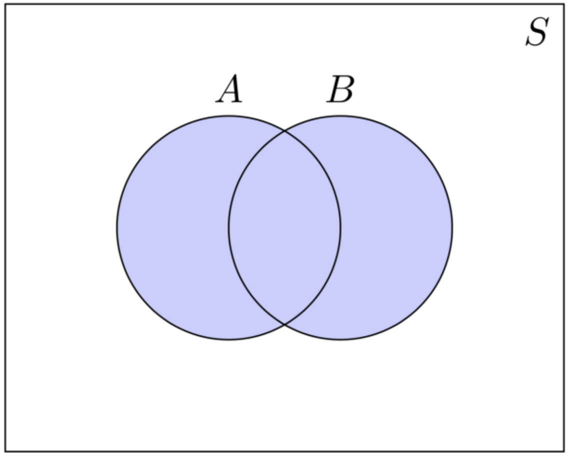
- The *complement* of event A, denoted ¬ A, is the event consisting of all outcomes in the sample space that are not in event A.



Derived events: union of two sets

We can define new events based on given events

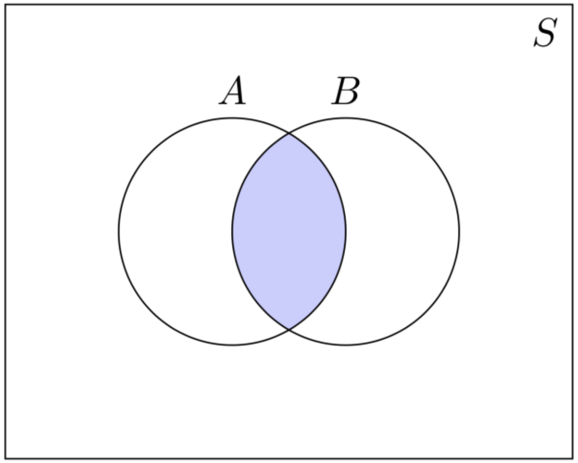
- The *union* of two events, A and B, denoted A U B, is the event consisting of all outcomes in A or B, **including** outcomes in both A and B.



Derived events: intersection of two sets

We can construct new events based on a given event A

- The *intersection of two events*, A and B, denoted $A \cap B$, is the event consisting of only the outcomes in **both** A and B.

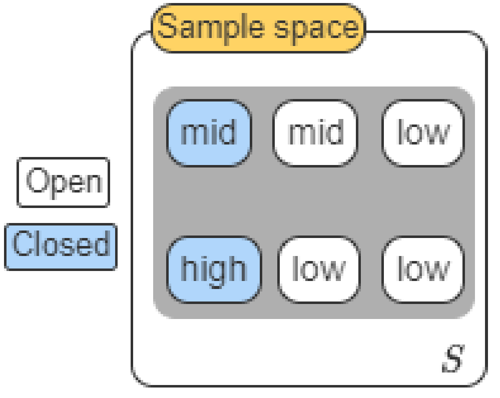


Derived events

We can construct new events based on a given event A

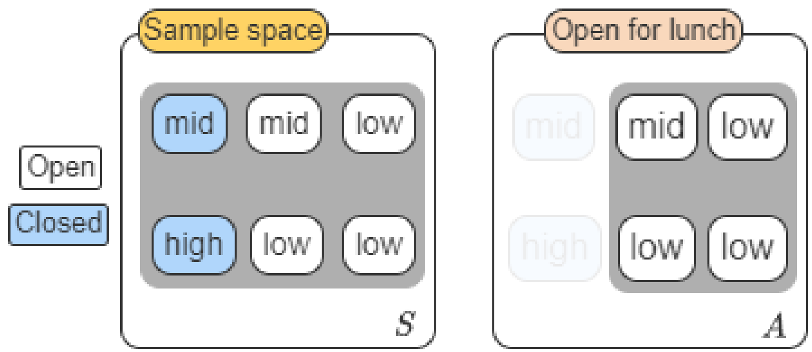
- The *complement* of event A, denoted ¬ A, is the event consisting of all outcomes in the sample space that are not in event A.
- The *union* of two events, A and B, denoted A U B, is the event consisting of all outcomes in A or B, **including** outcomes in both A and B.
- The *intersection of two events*, A and B, denoted A ∩ B, is the event consisting of only the outcomes in **both** A and B.

Example: random process and sample space



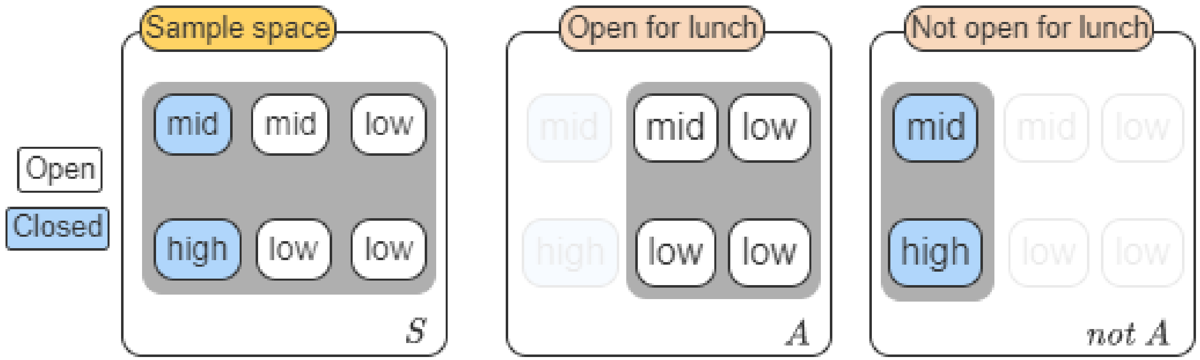
- One restaurant will be randomly selected from six options.
- Each restaurant has a price level (low, mid, or high) and is either open or closed for lunch.

Example: event A



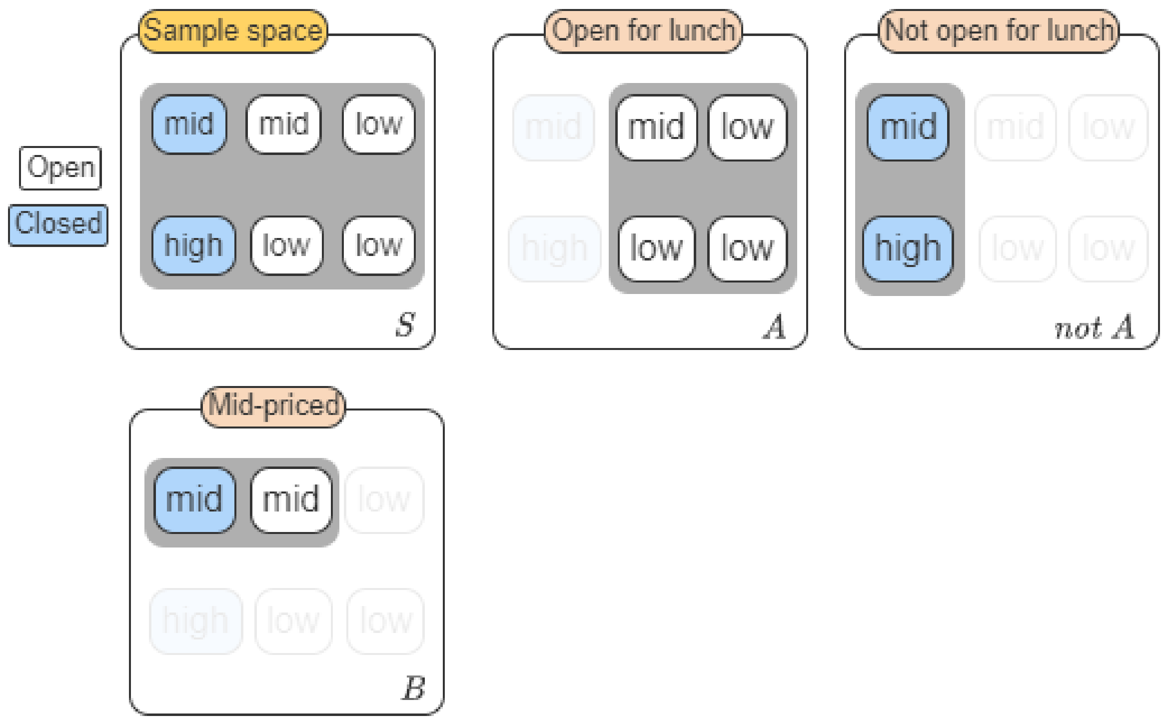
- We define event A as selecting a restaurant that is open for lunch.
- 4 restaurants, or outcomes in the sample space, are open for lunch and included in A.

Example: complement of event A



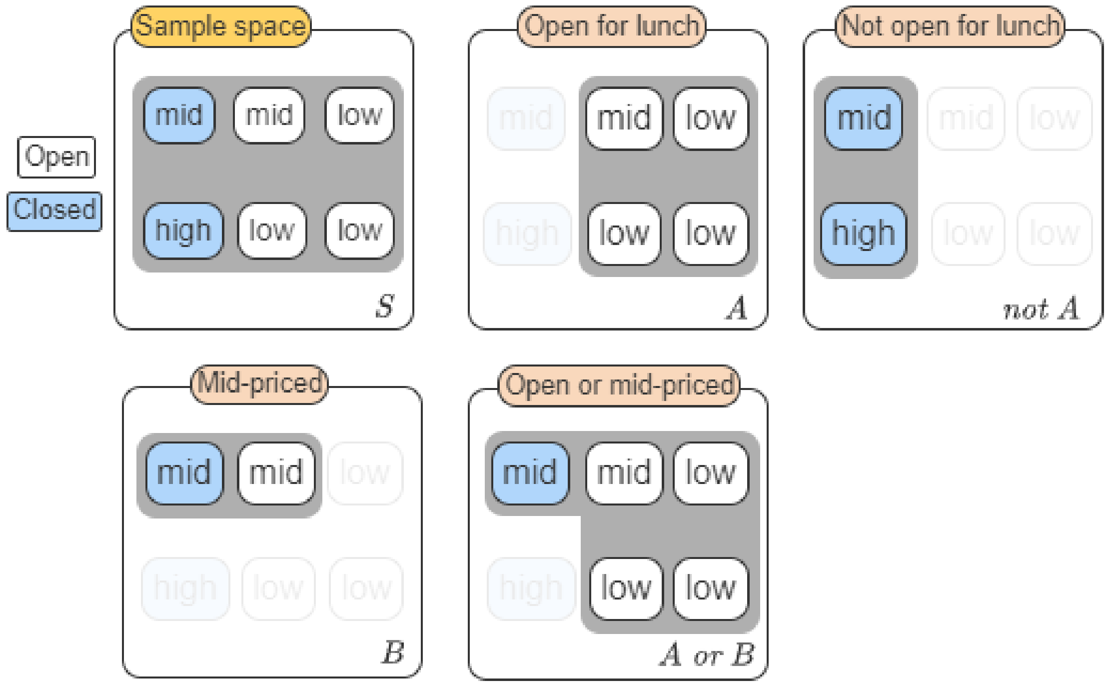
- The complement of event A is selecting a restaurant that is **not** open for lunch.
- The event ¬ A contains the 2 outcomes of the sample space which are not in A.

Example: event B



- We define event B as selecting a mid-priced restaurant.
- B covers 2 outcomes from the sample space.

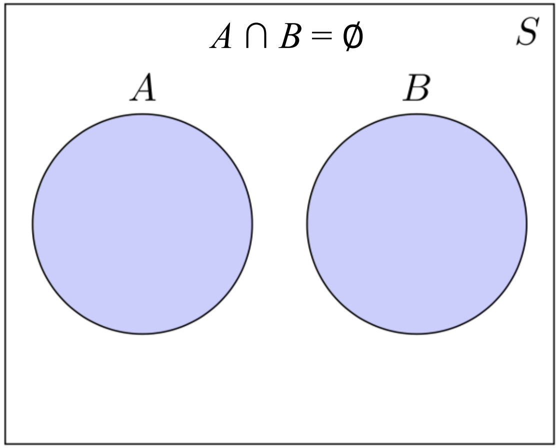
Example: event AUB



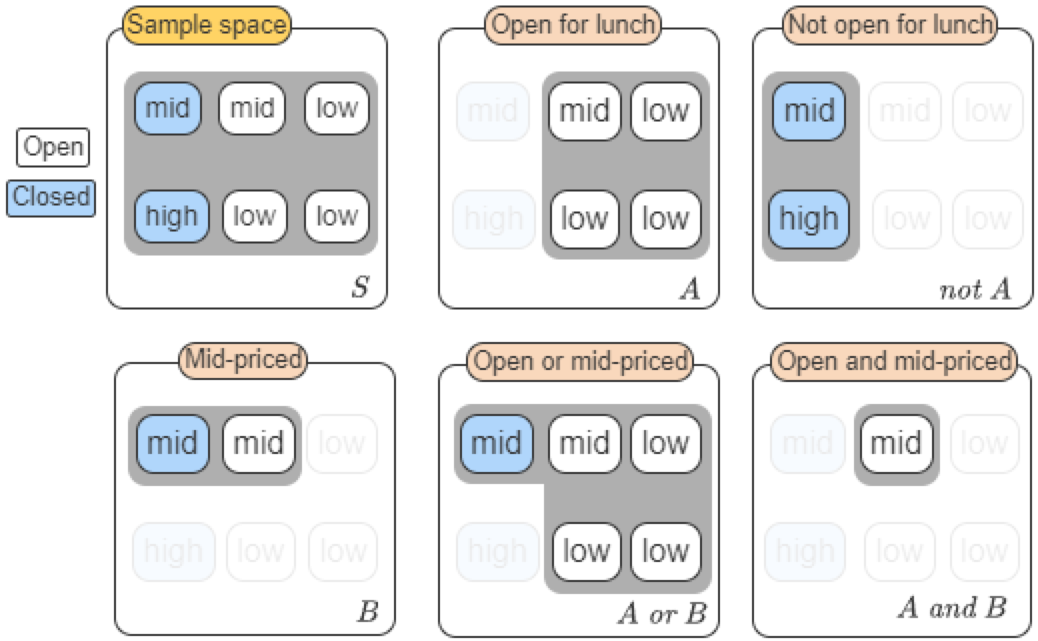
- The event AUB is selecting a restaurant that is either open for lunch or is mid-priced, including the one restaurant that is both open for lunch and mid-priced.

Disjoint (mutually exclusive) events

- The *intersection of two events*, A and B, denoted A ∩ B, is the event consisting of only the outcomes in **both** A and B.
- Two events, A and B, are considered *disjoint*, or mutually exclusive, if the two events have no outcomes in common.
  - In other words, if the event A ∩ B is empty or contains no outcomes.



Example: event A∩B



- The event A∩B is selecting a restaurant that is open for lunch **and** mid-priced.
- Because A∩B contains one outcome, events A and B are **not** disjoint.

Events quiz

Consider the process of randomly selecting a restaurant from the six options listed in the table below.

We define the events:

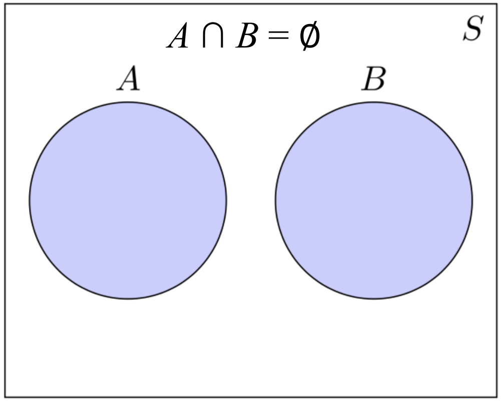
- A: randomly selecting a restaurant that is open for lunch
- B: randomly selecting a mid-priced restaurant
- C: randomly selecting a high-priced restaurant

Restaurant	1	2	3	4	5	6
Price	low	low	low	mid	mid	high
Lunch	open	open	open	open	closed	closed

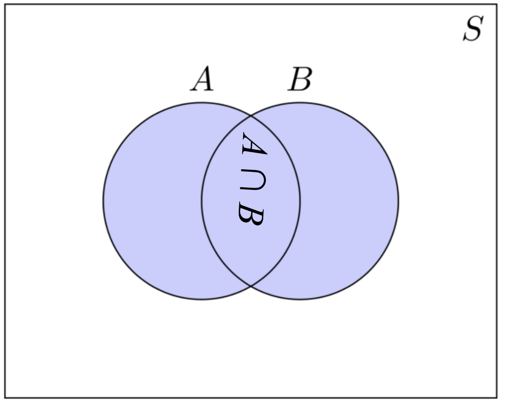
Axioms of probability

We accept without proof the following three axioms:

- The probability of any event is non-negative, P(A)≥0.
- The probability of the entire sample space is P(S)=1.
- If A and B are disjoint events, P(A U B)=P(A)+P(B).



Disjoint events

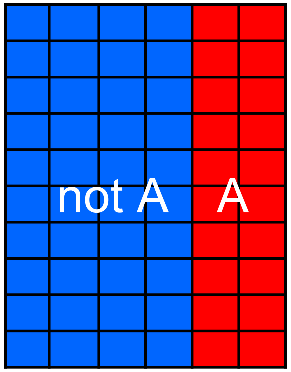


Non-disjoint events

Theorems 1. The complement rule

- The probability of the complement of event A can be found from the probability of event A:

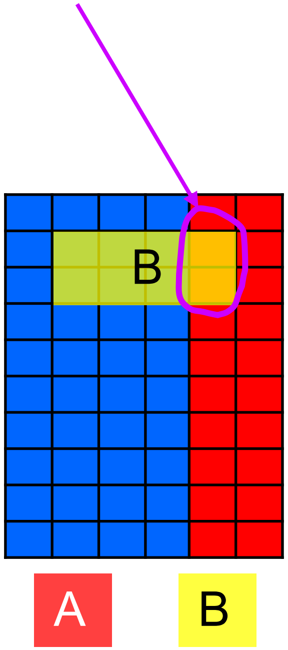
$$P(\text{not } A) = 1 - P(A).$$



Theorem 2. The addition rule (Union of two events)

- The probability of the union of any two events:

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



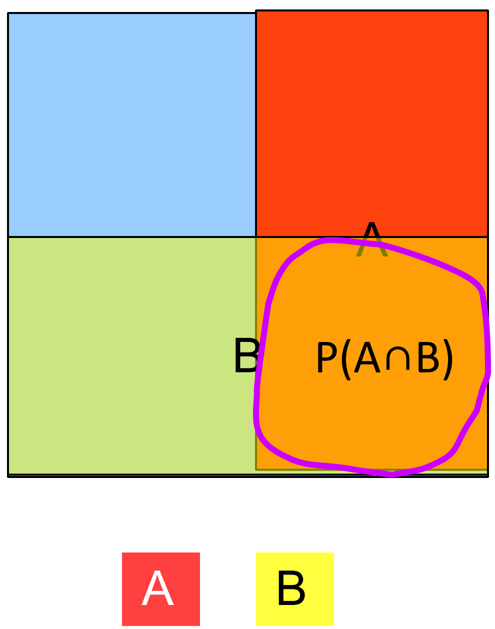
Independent events

- Two events are *independent* if the probability of one event does not affect the probability of the other.

Theorem 3. The multiplication rule (Intersection of two independent events)

- The probability of the intersection of independent events:

  $P(A \cap B) = P(A) * P(B)$.



Example:

$P(A) = ½$

$P(B) = ½$

$P(A \cap B) = ½ * ½ = ¼$

Exercise: data distribution

The following table gives the distribution of US household size (number of people in the household) based on the 2020 Current Population Survey.

Size	1	2	3	4	5	6	7+
Proportion	0.29	0.35	0.15	0.12	0.06	0.02	0.01

Exercise 1

Size	1	2	3	4	5	6	7+
Proportion	0.29	0.35	0.15	0.12	0.06	0.02	0.01

- Find the probability of randomly selecting a household with a size of more than 1

Exercise 1

Size	1	2	3	4	5	6	7+
Proportion	0.29	0.35	0.15	0.12	0.06	0.02	0.01

- Find the probability of randomly selecting a household with a size of more than 1

Event A: selection of household with size 1

P(not A) = 1 – 0.29 = 0.71

Exercise 2

Size	1	2	3	4	5	6	7+
Proportion	0.29	0.35	0.15	0.12	0.06	0.02	0.01

- Find the probability of randomly selecting a household with a size of 1 or more than 1

Exercise 2

Size	1	2	3	4	5	6	7+
Proportion	0.29	0.35	0.15	0.12	0.06	0.02	0.01

- Find the probability of randomly selecting a household with a size of 1 or more than 1

P(A or not A) = 1

Exercise 3

Size	1	2	3	4	5	6	7+
Proportion	0.29	0.35	0.15	0.12	0.06	0.02	0.01

- Find the probability of randomly selecting a household with a size of 5 or more, P(B).

Exercise 3

Size	1	2	3	4	5	6	7+
Proportion	0.29	0.35	0.15	0.12	0.06	0.02	0.01

- Find the probability of randomly selecting a household with a size of 5 or more, P(B).

$P(B)=0.06+0.02+0.01=0.09$

Exercise 4

Size	1	2	3	4	5	6	7+
Proportion	0.29	0.35	0.15	0.12	0.06	0.02	0.01

- Find the probability of randomly selecting a household with size 1 or 5 or more
- Event A: selection of household with size 1
- Event B: selection of household with size >= 5
- Find: P(A or B).

Exercise 4

Size	1	2	3	4	5	6	7+
Proportion	0.29	0.35	0.15	0.12	0.06	0.02	0.01

- Find the probability of randomly selecting a household with size 1 or 5 or more
- Event A: selection of household with size 1
- Event B: selection of household with size >= 5
- Find: P(A or B).

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

A and B are disjoint => $P(A \cap B) = 0$

$P(A) = 0.29$

$P(B) = 0.09$

$P(A \cup B) = 0.29 + 0.09 = 0.38$

Exercise 5

Size	1	2	3	4	5	6	7+
Proportion	0.29	0.35	0.15	0.12	0.06	0.02	0.01

- One household will be randomly selected from all households, and then a second household will be randomly selected from all households.
- Find the probability that both selected households are of size 1.

Exercise 5

Size	1	2	3	4	5	6	7+
Proportion	0.29	0.35	0.15	0.12	0.06	0.02	0.01

- One household will be randomly selected from all households, and then a second household will be randomly selected from all households (event .
- Find the probability that both selected households are of size 1.

C: selected household of size 1 in the first experiment

D: selected household of size 1 in the second experiment

Both events C and D are independent

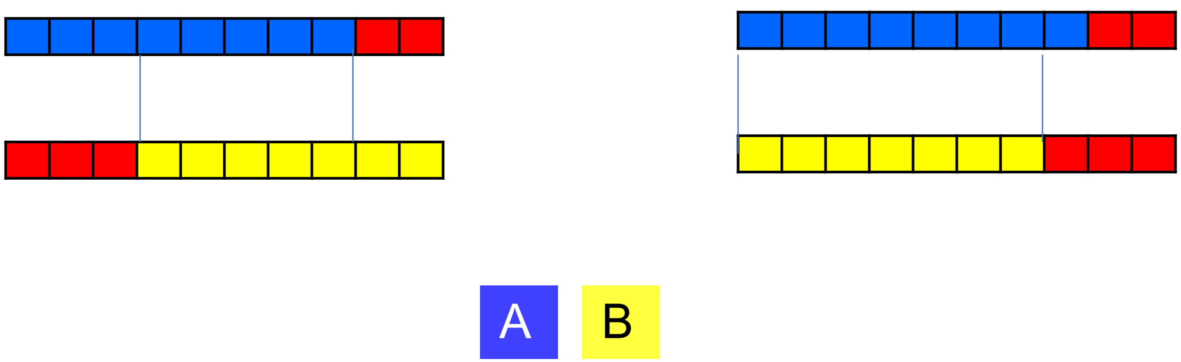
P(C and D) = 0.29 * 0.29 = 0.08

Stop and think exercise: intersection

- It is known that two events A and B in some probability space have probabilities 0.7 and 0.8.
- What is the minimal (maximum) possible probability of an event "A and B" ($A \cap B$)?

Stop and think exercise: intersection

- It is known that two events A and B in some probability space have probabilities 0.7 and 0.8.
- What is the minimal (maximum) possible probability of an event "A and B" (A∩B)?

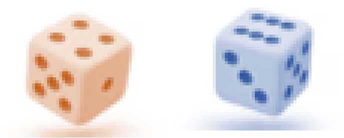


Answer: probability of intersection ranges from 0.5 to 0.7

General method for computing probability of an event

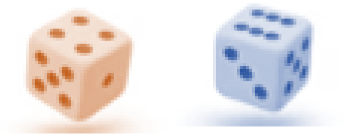
- Count total number of ALL possible outcomes (without additional information we assume that all events are equiprobable): N
- Count number of outcomes where event A is true: T (for True)
- To find P(A) divide T/N

Exercise 6: 2 dice



- We have 2 dice: a red die and a blue die
- If we throw 2 dice, what is the probability that the number on a red die is > than on the blue one?

Exercise 6: 2 dice



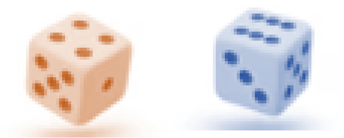
1. Count total number of ALL possible outcomes: N
2. Count number of outcomes where A is true: T
3. To find P(A) divide T/N

- We have 2 dice: a red die and a blue die
- If we throw 2 dice, what is the probability that the number on a red die is > than on the blue one?

We compute all possible outcomes by a rule of product (independent events):

N = 6x6 = 36 possible outcomes

Exercise 6: 2 dice



1. Count total number of ALL possible outcomes: N
2. Count number of outcomes where A is true: T
3. To find P(A) divide T/N

- We have 2 dice: a red die and a blue die
- If we throw 2 dice, what is the probability that the number on a red die is > than on the blue one?

We compute all possible outcomes by a rule of product (independent events):

N = 6x6 = 36 possible outcomes

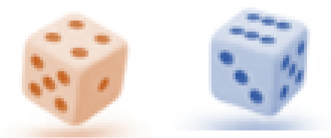
Now we need to count all outcomes where red number > blue number:

S=5+4+3+2+1=15

Enumerating:

Blue #	1	2	3	4	5	6
Red # > Blue	5	4	3	2	1	0

Exercise 6: 2 dice



1. Count total number of ALL possible outcomes: N
2. Count number of outcomes where A is true: T
3. To find P(A) divide T/N

- We have 2 dice: a red die and a blue die
- If we throw 2 dice, what is the probability that the number on a red die is > than on the blue one?

We compute all possible outcomes by a rule of product (independent events):

N = 6x6 = 36 possible outcomes

Now we need to count all outcomes where red number > blue number:

T=5+4+3+2+1=15

P(A) = 15/36

Enumerating:

Blue #	1	2	3	4	5	6
Red # > Blue	5	4	3	2	1	0

Exercise 7: most probable sum

- What is the most probable value of the sum of red and blue dice numbers?

	1	2	3	4	5	6
1						
2						
3						
4						
5						
6						

Exercise 7: most probable sum

- What is the most probable value of the sum of red and blue numbers?

- We again have 36 possible outcomes, we enumerate all of them, then compute the sum and see the most common sum

	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

Exercise 8: 2 boxes and 6 balls

Box 1

1	2

Box 2

3	4	5	6

We perform 2 selections: first we select a box, and then we select a ball. All the events are equiprobable

- What is the probability that we select a ball with number 2?
- What is the probability that we select a ball with number 6?
- What is the probability that we select an even number?

Exercise 8: 2 boxes and 6 balls

Box 1

1	2

Box 2

3	4	5	6

We perform 2 selections: first we select a box, and then we select a ball.
All the events are equiprobable

- What is the probability that we select a ball with number 2?
- What is the probability that we select a ball with number 6?
- What is the probability that we select an even number?

Events of selecting a box and then selecting a ball are not independent
To solve this, we draw the diagram

Exercise 8: 2 boxes and 6 balls

Box 1

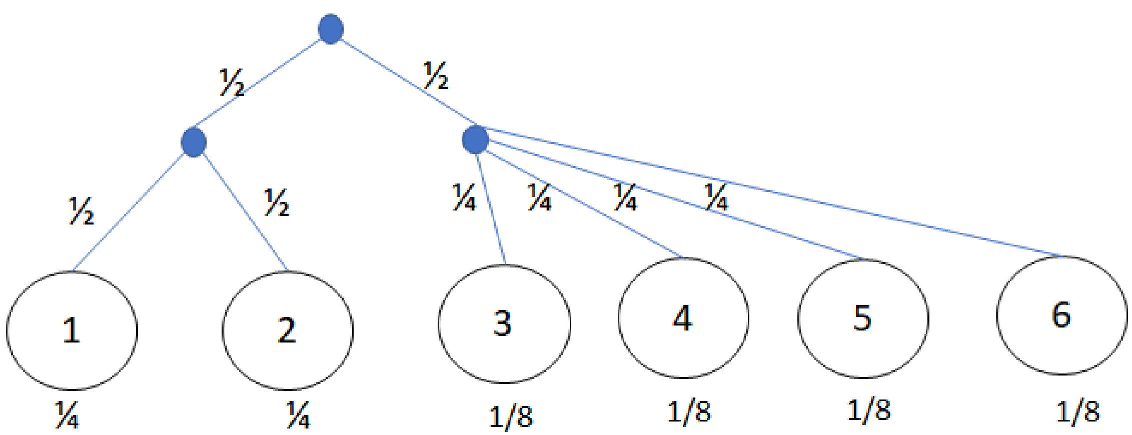
1	2

Box 2

3	4	5	6

We perform 2 selections: first we select a box, and then we select a ball. All the events are equiprobable

- What is the probability that we select a ball with number 2?
- What is the probability that we select a ball with number 6?
- What is the probability that we select an even number?



The probability of selecting each number is not equal
(depends on the selection of the box)

Exercise 8: 2 boxes and 6 balls

Box 1

1	2

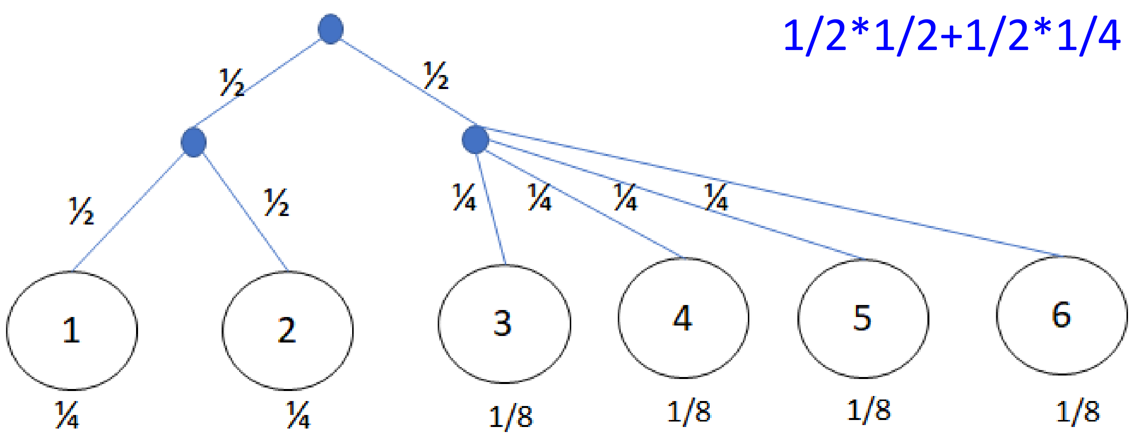
Box 2

3	4	5	6

We do 2 selections: first we select a box, and then we select a ball. All the events are equiprobable

- What is the probability that we select a ball with number 2? 1/4
- What is the probability that we select a ball with number 6? 1/8
- What is the probability that we select an even number?

1/2*1/2+1/2*1/4 + 1/2*1/4 = 1/2



Puzzle: life or death

King offers the following game to the prisoner.

- There are 2 boxes, 15 white and 15 black balls.
- The prisoner can distribute all balls between boxes as he likes, the only requirement is that no box is empty, and all balls are used.
- After that, the king picks one of the boxes, with probability 1/2 each, and then picks a random ball from the box (all with equal probabilities).
- If the ball is white, the prisoner is freed.
- Distribute the balls between boxes to maximize prisoner’s chances.

Puzzle: life or death

King offers the following game to the prisoner.

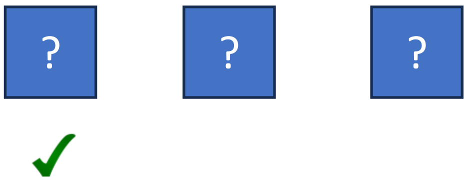
- There are 2 boxes, 15 white and 15 black balls.
- The prisoner can distribute all balls between boxes as he likes, the only requirement is that no box is empty, and all balls are used.
- After that, the king picks one of the boxes, with probability 1/2 each, and then picks a random ball from the box (all with equal probabilities).
- If the ball is white, the prisoner is freed.
- Distribute the balls between boxes to maximize prisoner's chances.

Box 1: 1 white ball

Box 2: 14 white + 15 black

Monty Hall three-door puzzle

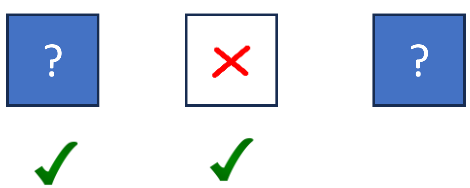
- Statistical game show. You are a contestant.
- You are asked to select one of three doors: the large prize is behind one of them and the other two doors are losers.



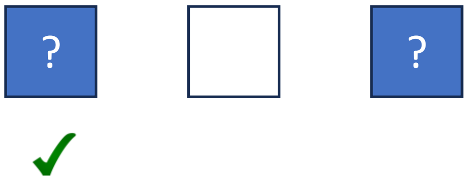
Monty Hall three-door puzzle

- Statistical game show. You are a contestant.
- You are asked to select one of three doors: the large prize is behind one of them and the other two doors are losers.

- Once you select a door, the show host, who knows what is behind each door, does the following:
  - First, whether or not you selected the winning door, he opens one of the other two doors that he knows is a losing door (selecting at random if both are losing doors).
  - Then he asks you whether you would like to switch doors.

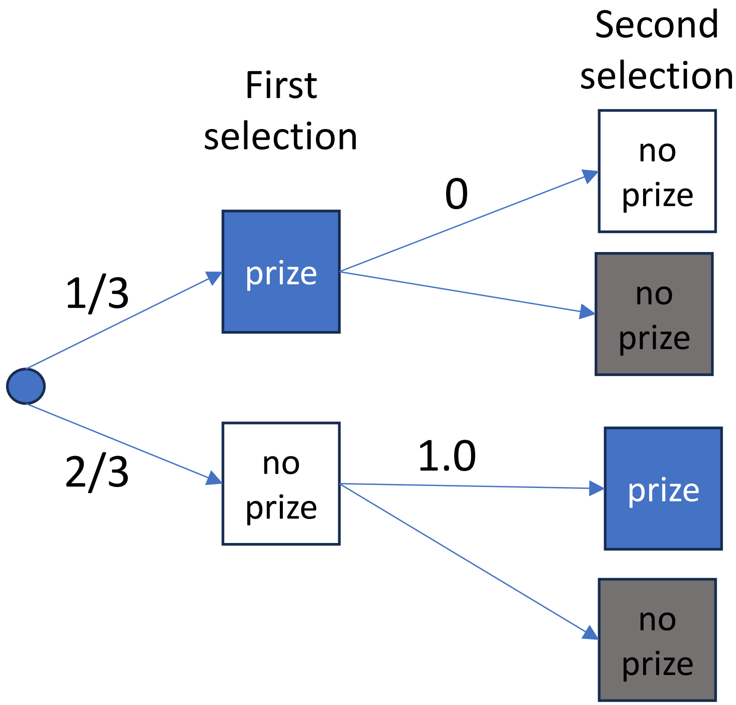


Monty Hall three-door puzzle

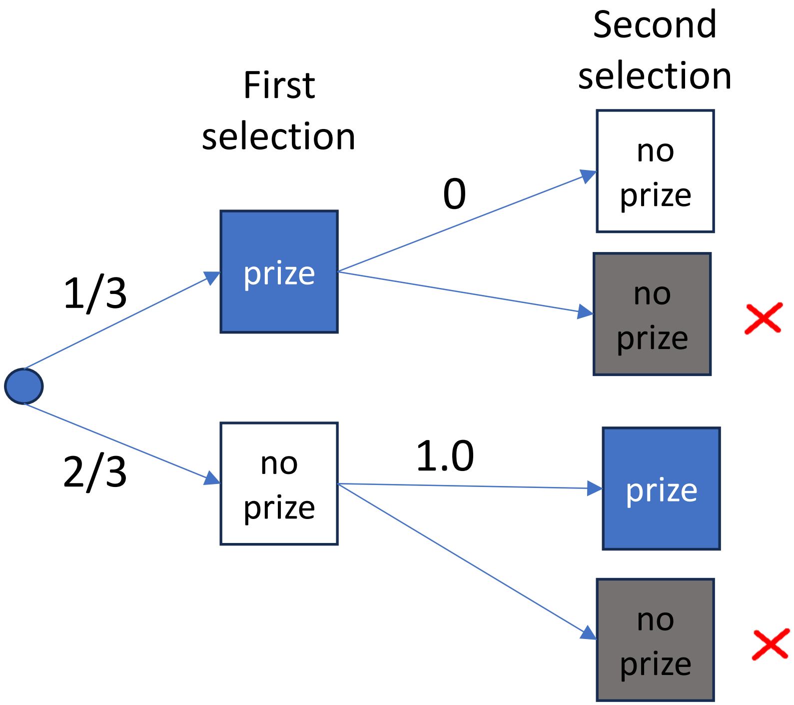


- Which strategy should you use? Should you change doors or keep your original selection, or does it not matter?
- What is the probability of winning in both cases?

Monty Hall three door puzzle: solution



Monty Hall three door puzzle: solution



- If you stop after the first selection, then probability of winning is 1/3
- If you make a change to your selection, then probability of winning is:

  1/3*0 OR 2/3*1 = 0 + 2/3 = 2/3.
- It makes sense to change the selected door: increases the probability of prize

We can count the number of outcomes using *combinatorics*

- Our next goal:
  - count number of outcomes **without enumerating them**
  - using formulas for selection