

# Probability

# Outlines

- Essential concepts in understanding probability
- Calculating probabilities using simple events
- Event relations and probability rules
- Independent events
- Conditional probabilities and Bayes' rule

# Introduction


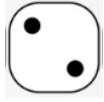
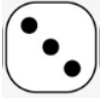
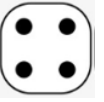


- The role of probability in statistics
  - Known population: describe the likelihood of a particular sample outcome
  - Unknown population: describe the properties of the population

# Concepts

- **Experiment** – the process by which an observation is obtained
- **Simple event** – the outcome observed on a single repetition of the experiment
- **Event** – a collection of simple events
- **Mutually exclusive events** – if one event occur, the others cannot.
- **Sample space** – a set of all possible simple events


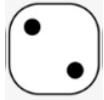

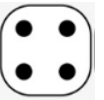


# Example 1

- Experiment: Roll the dice 100 times and observe the results.

- Simple events:  ,  ,  ,  ,  , 

- Event: even numbers are observed.

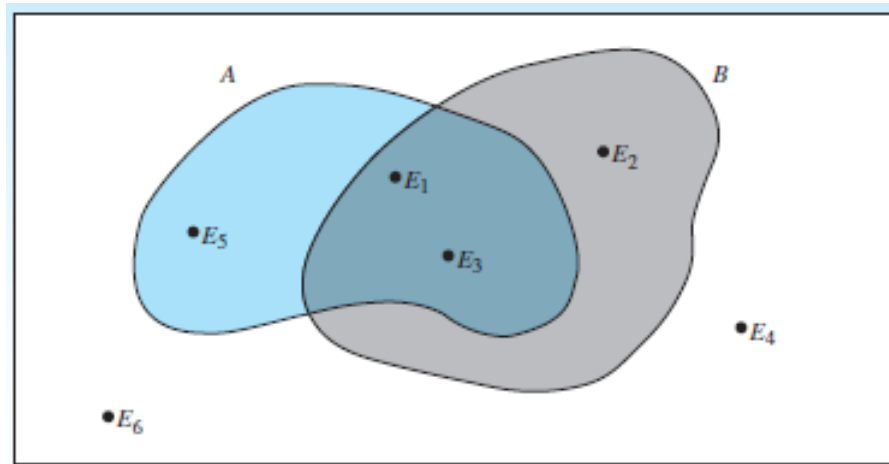
- Mutually exclusive events: all simple events are mutually exclusive.

- Sample space:  +  +  +  +  + 

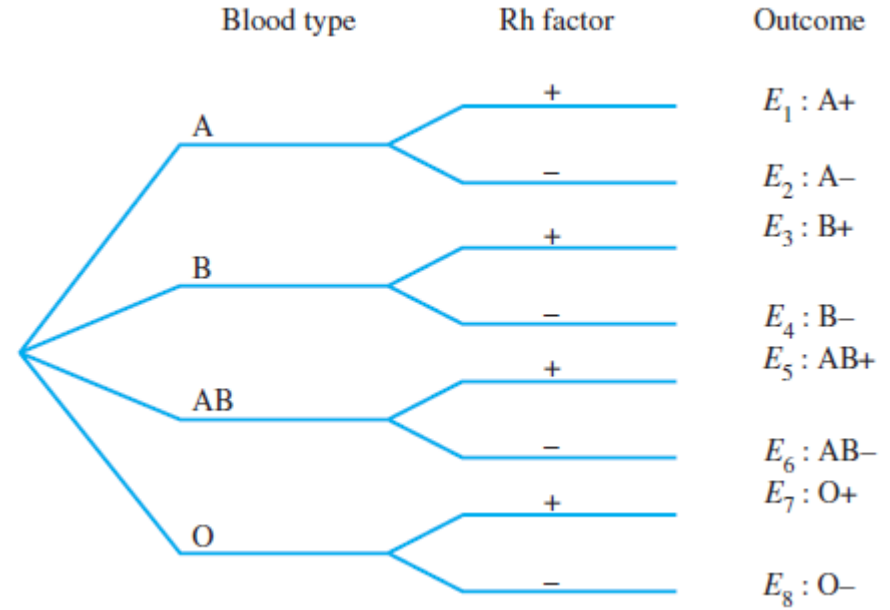
# Example 2

- **Experiment** – collect the age of 100 random males and 100 random females and put them in bins of U16, 17-50, 51-65, over 66
- **Simple events** – Assuming none of the 200 people was over 66. There was at least one observation of male and of female in each age group. Simple events are:
  - Male U16, Male 17-50, and Male 51-65
  - Female U16, Female 17-50, and Female 51-65
- **Events**
  - Event A: a person under 50 is picked.
  - Event B: a male is picked
  - Event C: a female is picked
- **Mutually exclusive events**
  - Events B and C are mutually exclusive.
  - Events A and B (or A and C) are not mutually exclusive.
- **Sample space** – comprised by all simple events

# Describing sample space



Venn diagram



Tree diagram

# Calculating probabilities using simple events

- Relative frequency,  $\frac{\text{Frequency}}{n}$
- Probability of an event A,  $P(A) = \lim_{n \rightarrow \infty} \frac{\text{Frequency}}{n}$
- It also equals the sum of probability of all simple events contained in A.
  - List all simple events in the sample space, i.e. the probability of all simple events considered MUST sum to 1.
  - Assign an appropriate probability for each simple event
  - Determine simple events resulting in the event of interest
  - Sum the probabilities of those simple events



# Example 3

- Event A: An observation of calcium between 400mg and 1000mg

- What are the simple events contained in A?

- What is the probability of event A?

<i>Calcium(mg)</i>	<i>Frequency</i>	<i>Relative Frequency</i>
(0 - 200]	65	0.09
(200 - 400]	174	0.24
(400 - 600]	178	0.24
(600 - 800]	123	0.17
(800 - 1000]	82	0.11
(1000 - 1200]	52	0.07
(1200 - 1400]	28	0.04
(1400 - 1600]	16	0.02
(1600 - 1800]	7	0.01
(1800 - 2000]	7	0.01
(2000 - 2200]	3	0.00
(2200 - 2400]	0	0.00
(2400 - 2600]	1	0.00
(2600 - 2800]	0	0.00
(2800 - 3000]	1	0.00
<b>Total</b>	<b>737</b>	<b>1.00</b>

## Example 2 – cont.

- **Experiment** – collect the age of 100 random males and 100 random females and put them in bins of U16, 17-50, and 51-65 (assuming no one above 65 was observed)
- **Events:**
  - Event A: a person under 50 is picked.
  - Event B: a male is picked
  - Event C: a female is picked
- **Questions:**
  - Draw a tree diagram of the sample space
  - What are the simple events contained in A, B, and C?
  - What is the probability of event A?

	Male	Female
<16	30	18
17-50	50	65
51-65	20	27
	<b>100</b>	<b>100</b>

# A review of useful counting rules

- Counting rules are helpful in identifying the number of simple events  $N$  in experiments, especially when  $N$  is *large*.

- **The mn Rule**

If an experiment is done in  $k$  stages with  $n_k$  ways to accomplish a stage  $k$ , the number of ways to accomplish the experiment, i.e. the number of simple events, is  $n_1 n_2 n_3 \dots n_k$ .

- **Examples:**

- Roll three 6-face dices, the total number of results is  $6 \times 6 \times 6 = 216$
- The total number possible combinations of male and female in 4 age groups are  $2 \times 4 = 8$ .
- There are 3 books A, B, C and 2 slots. The total number of ways to organize the books is  $3 \times 2 = 6$

# A review of useful counting rules

- **A counting rule for permutations (order of objects is important)**

The total number of ways to arrange  $n$  distinct objects, taking them  $r$  at a time is

$$P_r^n = \frac{n!}{(n-r)!} \quad \text{where } n! = n(n-1)(n-2) \dots (3)(2)(1)$$

- Examples:

- The total number of ways to arrange 5 different books is

$$P_5^5 = \frac{5!}{(5-5)!} = \frac{5!}{0!} = 5 * 4 * 3 * 2 * 1 = 120$$

- The total number of ways to select 5 people from 8 people (and order is important) is

$$P_5^8 = \frac{8!}{(8-5)!} = \frac{8!}{3!} = 8 * 7 * 6 * 5 * 4 = 6720$$

# A review of useful counting rules

- **A counting rule for combinations (order of objects is NOT important)**

The total number of ways to combine  $n$  distinct objects, taking them  $r$  at a time is

$$C_r^n = \frac{n!}{r!(n-r)!} \quad \text{where } n! = n(n-1)(n-2) \dots (3)(2)(1)$$

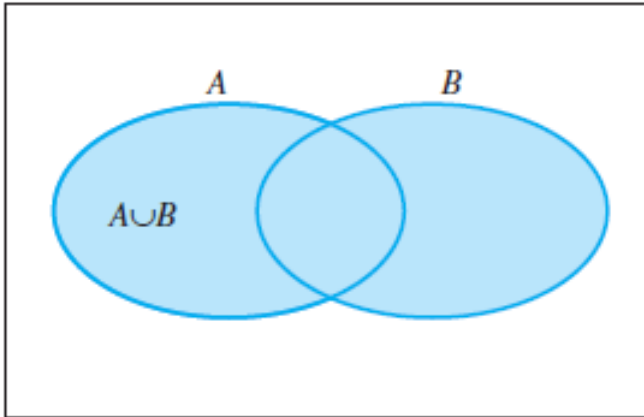
- Examples:

- The total number of ways to pick 3 books out of 5 different books is

$$C_3^5 = \frac{5!}{3!(5-3)!} = \frac{5!}{3!2!} = 10$$

- How many ways are there to pick 10 nurses out of 90 for a study in determining the attitudes of nurses toward various admin procedures? Is the order of selecting the nurses important?

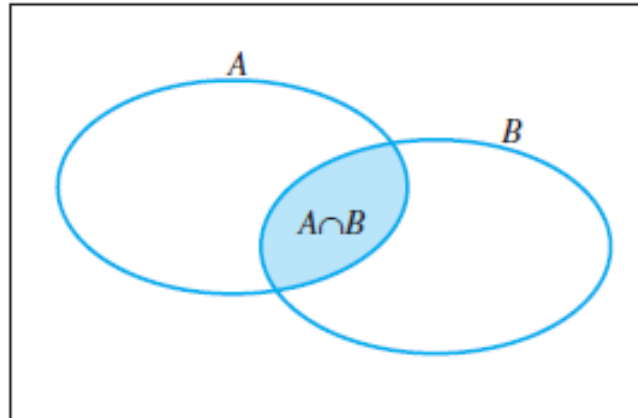
# Event Relations and Probability Rules



**Union of A and B:** either A or B or both occur

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

**The Addition Rule**

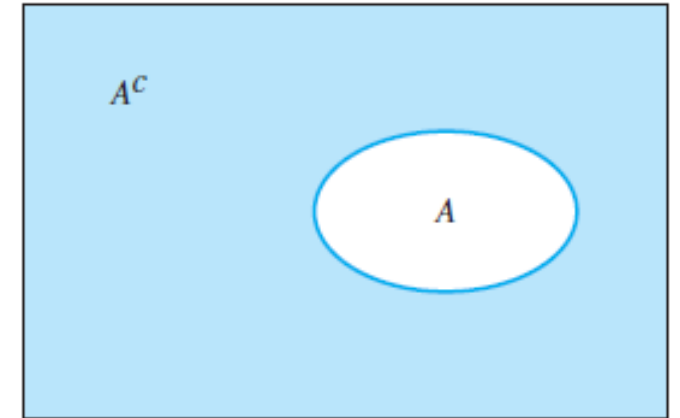


**Intersection of A and B:** both A and B occur

$$P(A \cap B) = P(A)P(B|A)$$

$$P(A \cap B) = P(B)P(A|B)$$

**The Multiplication Rule**



**Complement of A:** A does not occur

$$P(A^c) = 1 - P(A)$$

$$A \cup A^c = S$$

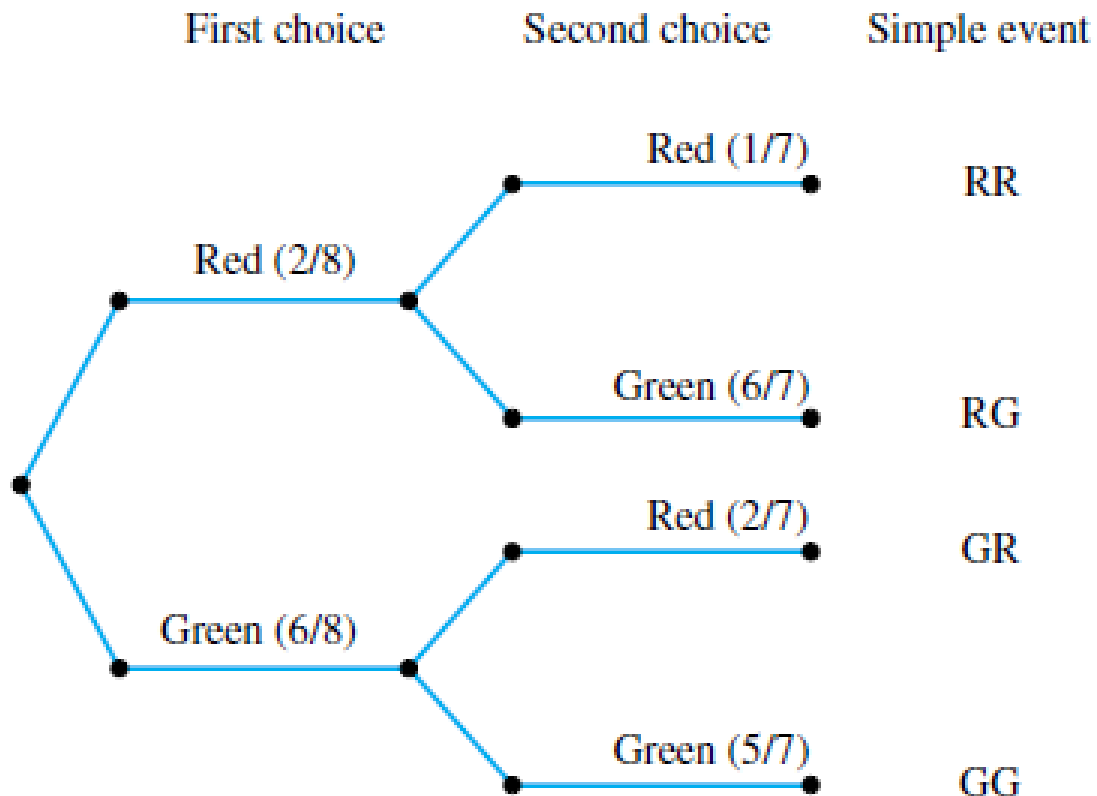
**The Rule for Complements**

# Example 4

- Toss 2 fair coins and record the outcomes. Below are the events of interest
  - A: Observe at least 1 head
  - B: Observe 2 different faces
- Simple events (can be from a tree diagram)
  - E1: HH,  $P(E1) = \frac{1}{4}$
  - E2: HT,  $P(E2) = \frac{1}{4}$
  - E3: TH,  $P(E3) = \frac{1}{4}$
  - E4: TT,  $P(E4) = \frac{1}{4}$
- $A = \{E1, E2, E3\}$ ,  $P(A) = \frac{3}{4}$
- $B = \{E2, E3\}$ ,  $P(B) = \frac{2}{4}$
- $A \cup B = \{E1, E2, E3\}$ ,  $P(A \cup B) = \frac{3}{4}$
- $A \cap B = \{E2, E3\}$ ,  $P(A \cap B) = \frac{1}{2}$
- $A^c = \{E4\}$ ,  $P(A^c) = \frac{1}{4}$

# Example 5

- There are 8 toys in a container – 2 red and 6 green. Pick random 2 toys.
- Event A: What is the probability of picking up 2 red toys?



$$P(A) = P(\text{Red on 1st})P(\text{Red on 2nd} \mid \text{Red on 1st})$$

$$P(A) = \binom{2}{8} \binom{1}{7} = 1/28$$



## Example 2 - cont.

- **Events:**

- Event A: a person under 50 is picked.
- Event B: a male is picked
- Event C: a female is picked

	Male	Female
<16	30	18
17-50	50	65
51-65	20	27
	<b>100</b>	<b>100</b>

- What is the probability of event A?
- What is the probability of event B?
- What is the probability of a male under 50 ( $A \cap B$ )?
- What is the probability of a person under 50 or a female ( $A \cup C$ )?
- What is the probability of a person over 50 ( $A^c$ )?

# Independent events

- Event A and event B are **independent** if and only if

$$P(A|B) = P(A) \quad \text{or} \quad P(A \cap B) = P(A)P(B)$$

- Extension of multiplication rules for three independent events

$$P(A \cap B \cap C) = P(A)P(B)P(C)$$

- **Example:** Roll 3 dices and observe the outcome. What is the probability of having 3 ?

# Checking independent events

- Roll a single dice and consider the following events
  - Event E: getting an even number
  - Event T: getting a number divisible by three
- Questions:
  - What is the probability of E?
  - What is the probability of getting an even number (Event E) if you are told that the number was also divisible by three (Event T)?
  - Does knowing that the number is divisible by 3 (Event T) change the probability that the number was even (Event E)?

**Are Event E and Event T independent?**

# Independent Events vs Mutual Exclusive Events

- Mutually exclusive events
  - Cannot both happen, e.g. head and tail cannot both happen in a coin toss
  - If A happened, B cannot happen,  $P(B|A) = 0$
  - Therefore mutually exclusive events are *dependent*.
  - $P(A \cap B) = 0, P(A \cup B) = P(A) + P(B)$
- Independent events
  - $P(A \cap B) = P(A)P(B), P(A \cup B) = P(A) + P(B) - P(A)P(B)$
- **Example:** What is the probability of drawing an ace and a 10 from a deck of 52 cards?

# Conditional Probabilities

- Conditional probability of an event B given that event A has occurred is

$$P(B|A) = \frac{P(A \cap B)}{P(A)} \quad \text{if } P(A) \neq 0$$

- Examples:

- What is the probability of a person <16 (Event  $B_1$ ) given that the person is a male (Event A)?
- What is the probability of a person a male (Event A) given that he is <16 (Event  $B_1$ )?
- Is  $P(A|B_1) = P(B_1|A)$ ?

	Male (A)	Female ( $A^c$ )	
<16 ( $B_1$ )	30	18	<b>48</b>
17-50 ( $B_2$ )	50	55	<b>105</b>
51-65 ( $B_3$ )	20	27	<b>47</b>
	<b>100</b>	<b>100</b>	<b>200</b>

	Male (A)	Female ( $A^c$ )	
<16 ( $B_1$ )	0.15	0.09	<b>0.24</b>
17-50 ( $B_2$ )	0.25	0.275	<b>0.525</b>
51-65 ( $B_3$ )	0.1	0.135	<b>0.235</b>
	<b>0.5</b>	<b>0.5</b>	<b>1</b>

# Bayes' Rule

- Bayes' rule of conditional probability

$$P(B_i|A) = \frac{P(B_i \cap A)}{P(A)} = \frac{P(B_i)P(A|B_i)}{\sum_{j=1}^k P(B_j)P(A|B_j)} \quad \text{for } i = 1, 2, \dots, k$$

- $B_1, \dots, B_j$  must be mutually exclusive and  $\sum_{j=1}^k P(B_j) = 1$
- Back to the example in the previous slide

$$P(B_1|A) = \frac{P(B_1 \cap A)}{P(B_1 \cap A) + P(B_2 \cap A) + P(B_3 \cap A)}$$

$$P(B_1|A) = \frac{P(B_1)P(A|B_1)}{P(B_1)P(A|B_1) + P(B_2)P(A|B_2) + P(B_3)P(A|B_3)}$$

$$P(B_1|A) = \frac{0.24 \cdot (30/48)}{0.24 \cdot (30/48) + 0.525 \cdot (50/105) + 0.235 \cdot (20/47)} = 0.3$$

	Male (A)	Female (A <sup>c</sup> )	
<16 (B <sub>1</sub> )	30	18	<b>48</b>
17-50 (B <sub>2</sub> )	50	55	<b>105</b>
51-65 (B <sub>3</sub> )	20	27	<b>47</b>
	<b>100</b>	<b>100</b>	<b>200</b>

	Male (A)	Female (A <sup>c</sup> )	
<16 (B <sub>1</sub> )	0.15	0.09	<b>0.24</b>
17-50 (B <sub>2</sub> )	0.25	0.275	<b>0.525</b>
51-65 (B <sub>3</sub> )	0.1	0.135	<b>0.235</b>
	<b>0.5</b>	<b>0.5</b>	<b>1</b>

# Bayes' Rule

$$P(B_i|A) = P(B_i) * \frac{P(A|B_i)}{P(A)} \quad \text{for } k = 1, 2, \dots, k$$

- $P(B_i)$  is **prior probability** – without knowledge of the condition A. Can be approximated as  $1/k$  if unknown.
- $P(B_i|A)$  is **posterior probability** – the updated version of the prior probability after observing information of the condition A in the sample.

# Bayes' Rule

$$P(B_i|A) = P(B_i) * \frac{P(A|B_i)}{P(A)}$$

- Example:
  - 60% of businesses that replaced their CEO last year has share price increased by >5%.
  - 35% of businesses that replaced their CEO last year doesn't have share price increased by >5%.
  - Last year data showed that the probability of share price increased by >5% is 4%.
  - What is the probability of a company's share price increased by >5% given it replaced the CEO?
- Solution hints
  - Event A: A CEO being replaced
  - Event B<sub>1</sub>: A business has share price increased by >5%.
  - What is the prior probability of a company having share price increased by >5%?
  - What is the probability of a CEO being replaced?
  - What is the probability of a CEO being replaced given that the share price increased by >5%?