Chapter 5: Probability

1

Terms

- **Random trial**: a process or experiment that has two or more possible **outcomes**
  - Flipping coin (random trial): heads or tails (possible outcomes)
  - Rolling die (random trial): 1, 2, 3, 4, 5, 6 (possible outcomes)
- **Event** (of interest): any potential subset of all possible outcomes
  - Flipping coin: heads
  - Rolling die: 3

2

Probability

- The **probability** of an event is the proportion of times the event would occur if we repeated a random trial over and over again under the same conditions
  - Probability ranges between zero and one
  - Pr[A] means "the probability of event A"
- Two events are **mutually exclusive** if they cannot occur at the same time
  - Pr[A and B] = 0

3

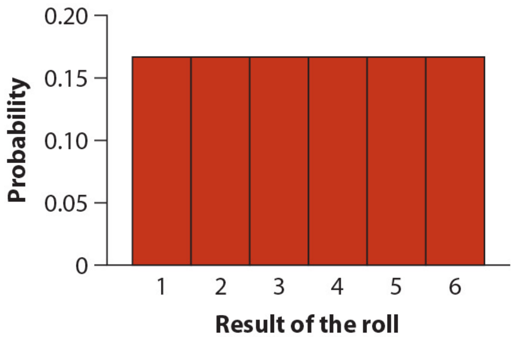
Probability

- The **probability distribution** is a list of the probabilities of all mutually exclusive outcomes of a random trial

4

Discrete probability distribution

- A discrete variable is measured in indivisible units
  - All categorical variables (e.g., present or absent) and many numerical variables (e.g., number of mates)



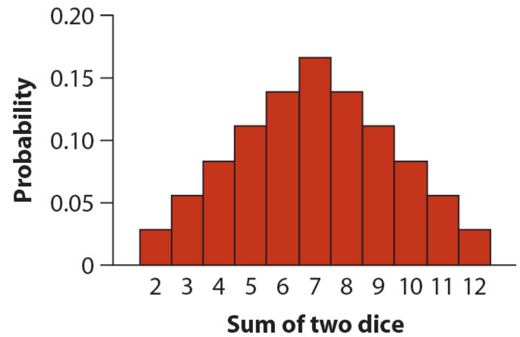


Fig 5.4-1

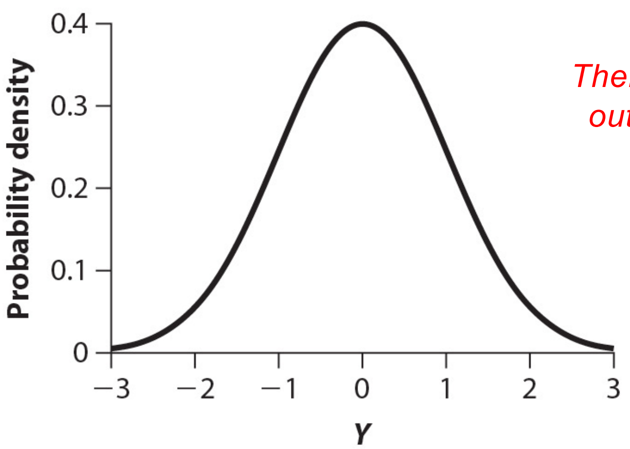
Probabilities sum to 1!

Fig 5.4-2

5

Continuous probability distribution

- Continuous variables can take on any real number value within some range



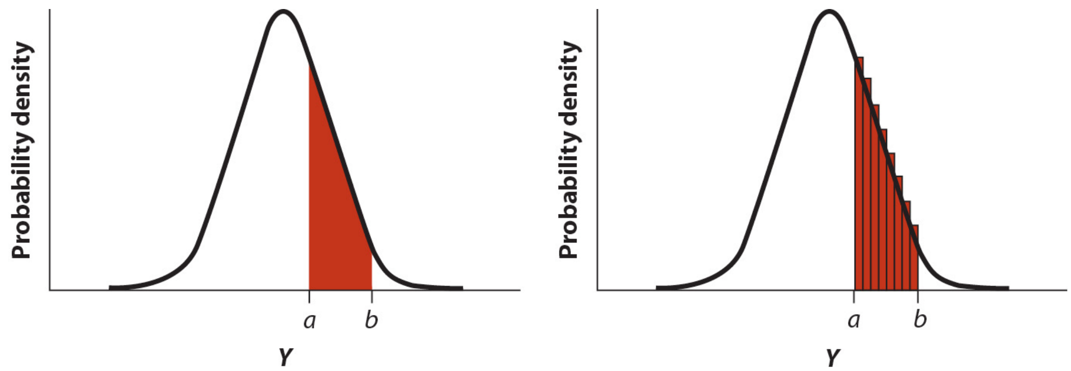
There are an infinite number of outcomes between -3 and 3!

Fig 5.4-3

6

Continuous probability distribution

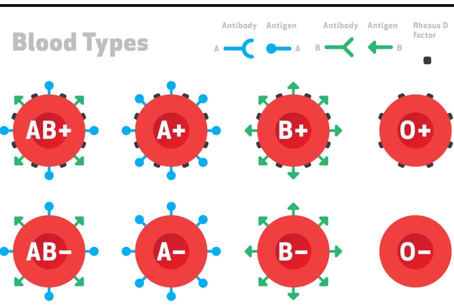
- Probability of Y being within some range is indicated by the area under the curve



Pr[a < Y < b]

Fig 5.4-4

7



https://www.youtube.com/watch_popup?v=ttjn1jVACk8

What is the probability of a person having type O?

TABLE 5.5-1 Probability that a randomly chosen American will have a given blood type. A, B, and O refer to ABO blood type, and "+" and "–" refer to Rh factor.

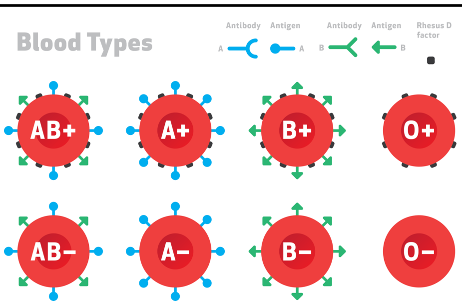
Blood type	Probability
O+	0.374
O–	0.066
A+	0.357
A–	0.063
B+	0.085
B–	0.015
AB+	0.034
AB–	0.006

8

Addition rule (either/or)

- The **addition rule:** if two events A and B are mutually exclusive then Pr[*A* or *B*] = Pr[*A*] + Pr[*B*]
- Can be extended to more than two events
- Probabilities of all possible mutually exclusive events sum to one
- The **probability of an event not occurring** is one minus the probability that it occurs
  - Pr[*not A*] = 1 – Pr[*A*]

9



https://www.youtube.com/watch_popup?v=ttjn1jVACk8

TABLE 5.5-1 Probability that a randomly chosen American will have a given blood type. A, B, and O refer to ABO blood type, and "+" and "–" refer to Rh factor.

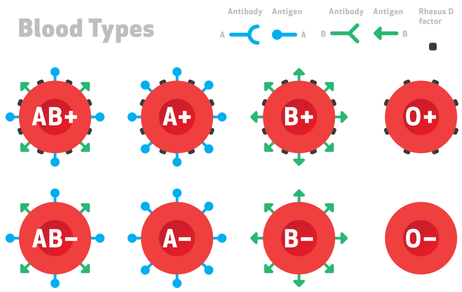
Blood type	Probability
O+	0.374
O–	0.066
A+	0.357
A–	0.063
B+	0.085
B–	0.015
AB+	0.034
AB–	0.006

What is the probability of a person having type O?

Pr[O+ or O-] = Pr[O+] + Pr[O-]

Pr[O+ or O-] = 0.374 + 0.066 = 0.44

10



https://www.youtube.com/watch_popup?v=ttjn1jVACk8

What is the probability of a person having anything but type O?

Pr[O+ or O-] = Pr[O+] + Pr[O-]

Pr[O+ or O-] = 0.374 + 0.066 = 0.44

Pr[not O] = 1 - 0.44 = 0.56

TABLE 5.5-1 Probability that a randomly chosen American will have a given blood type. A, B, and O refer to ABO blood type, and "+" and "–" refer to Rh factor.

Blood type	Probability
O+	0.374
O–	0.066
A+	0.357
A–	0.063
B+	0.085
B–	0.015
AB+	0.034
AB–	0.006

11

General addition rule (either/or)

- Not all events are mutually exclusive, so extra term is needed so you don't double-count outcomes
- Pr[*A* or *B*] = Pr[*A*] + Pr[*B*] – Pr[*A* and *B*]

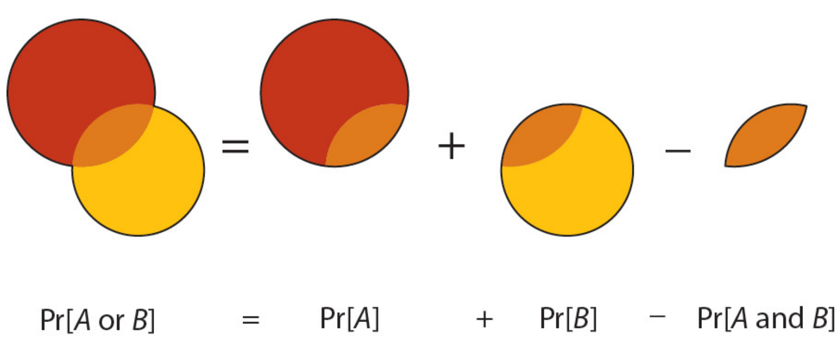


Fig 5.4-4

12

TABLE 5.5-1 Probability that a randomly chosen American will have a given blood type. A, B, and O refer to ABO blood type, and "+" and "–" refer to Rh factor.

Blood type	Probability
O+	0.374
O–	0.066
A+	0.357
A–	0.063
B+	0.085
B–	0.015
AB+	0.034
AB–	0.006

What is the probability of a person being type O or Rh + ?

Pr[O or +] = Pr[O] + Pr[+] – Pr[O and +]

0.374+0.066=0.440

0.374+0.357+0.085+0.034=0.850

0.374

0.440+0.850-0.374=0.916

13

14

Independent events

- Two events are **independent** if the occurrence of one does not inform us about the probability that the second will occur
  - e.g., two flips of a coin or roll of a die

15

Multiplication rule (and)

- If two events are independent then the probability that they both occur is the probability of the first event multiplied by the probability of the second event
- The **multiplication rule:** if two events A and B are independent then Pr[A and B] = Pr[A] x Pr[B]
- Can be extended to more than two events

16

Smoking and high blood pressure

- Both smoking and high blood pressure are associated with vascular diseases (e.g., strokes)
- Research has shown that smoking and high blood pressure are independent of each other
- In US: ~17% adults smoke and ~22% have high blood pressure

17

Smoking and high blood pressure

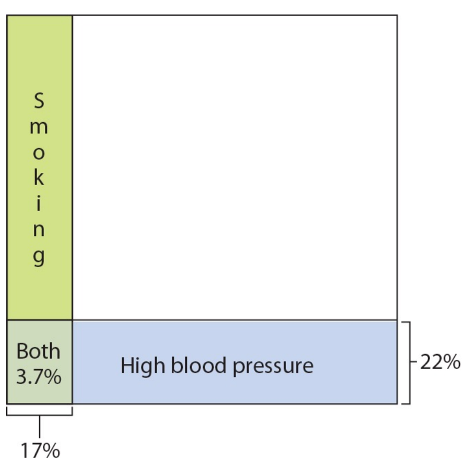


Fig 5.6-2

- What is the probability that a randomly chosen person is a smoker (S) and has high blood pressure (BP)?
- Pr[S and BP] = Pr[S] x Pr[BP]

0.17 x 0.22 = 0.037

18

19

Probability tree

- **Probability tree**: a diagram used to calculate the probabilities of combination of events from multiple random trials

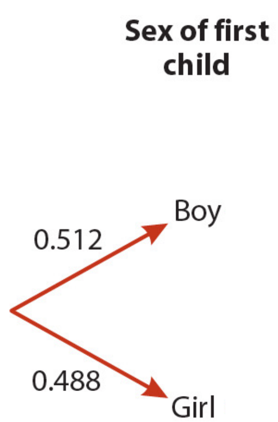
20

Ex 5.7: Sex and birth order

- Probability that a child is a boy: Pr[boy] = 0.512
- Thus, the probability that a child is a girl:
  Pr[girl] = 1 – Pr[boy] = 1 – 0.512 = 0.488
- What is the probability that a couple has two children, with one boy and one girl?

21

Ex 5.7: Sex and birth order



Note: total probability of all potential outcomes must sum to one

Fig 5.7-1

22

Ex 5.7: Sex and birth order

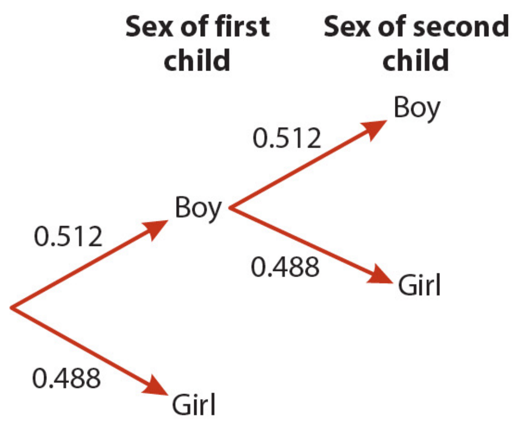


Fig 5.7-1

23

Ex 5.7: Sex and birth order

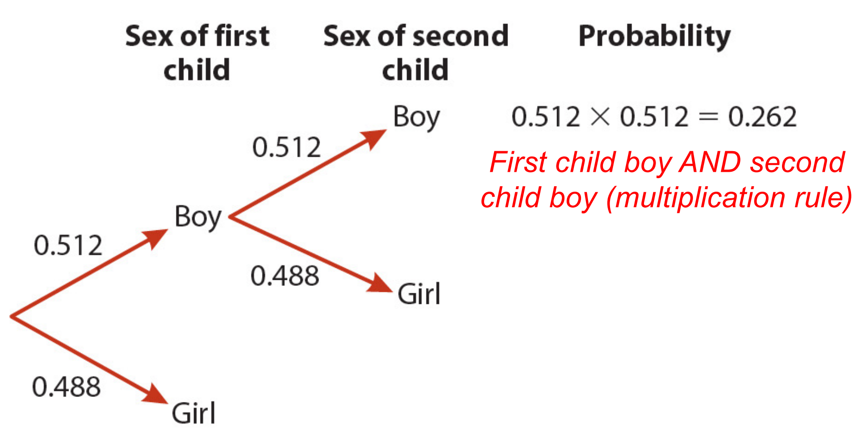


Fig 5.7-1

24

Ex 5.7: Sex and birth order

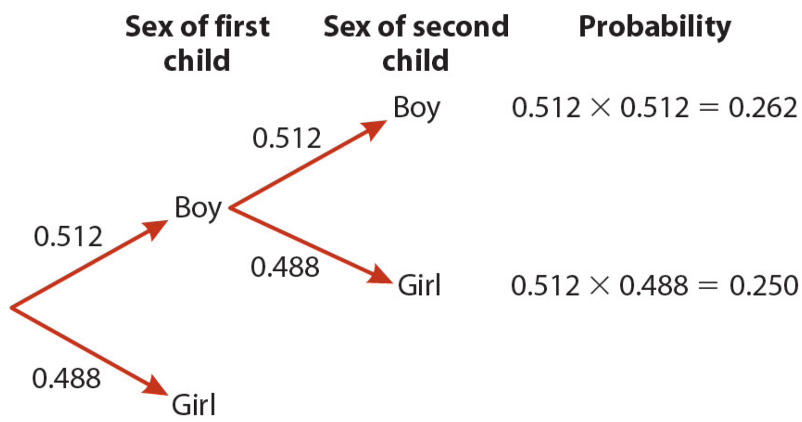


Fig 5.7-1

25

Ex 5.7: Sex and birth order

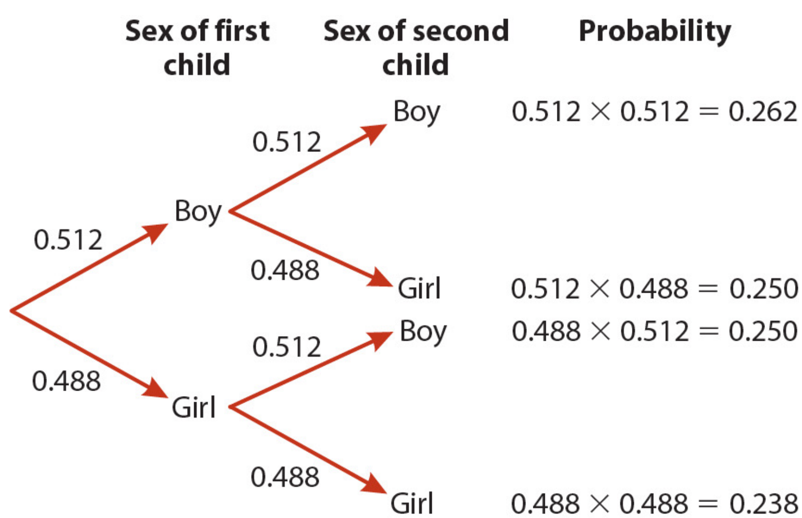


Fig 5.7-1

26

Ex 5.7: Sex and birth order

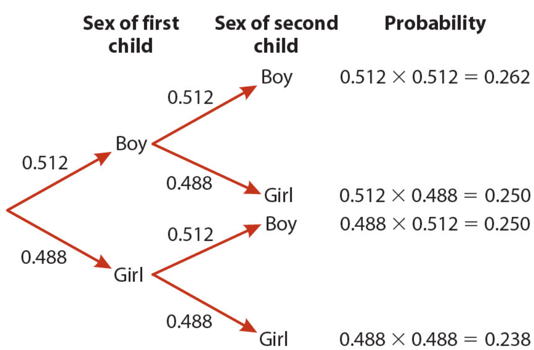


Fig 5.7-1

- What is the probability that a couple has two children, with one boy and one girl?
- Two possible paths:
  - 1: boy first, girl second
  - 2: girl first, boy second
- **Either** path results in the desired outcome
  - Addition rule
- Pr[boy and girl] = Pr[path1] + Pr[path2]
  Pr[boy and girl] = 0.25 + 0.25 = 0.5

27

Dependent events

- In many cases the probability of a particular event in the second trial depends on what happened in the first trial
  - i.e., the events are not independent (or dependent)
- *Nasonia* wasps
  - Parasitic (lays eggs on fly pupae)
  - Female can manipulate sex of offspring
  - If pupae previously parasitized: produce more sons
  - If pupae not previously parasitized: produce more daughters

28

Ex 5.8: Is this meat taken?

- *Nasonia* wasps
  - Parasitic (lays eggs on fly pupae)
  - Female can manipulate sex of offspring
  - If pupae previously parasitized: produce more sons
  - If pupae not previously parasitized: produce more daughters

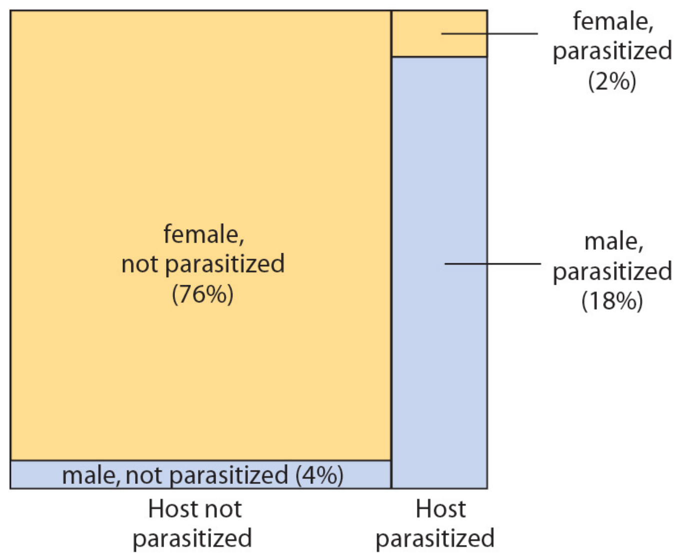


Fig 5.8-1

29

Ex 5.8: Is this meat taken?

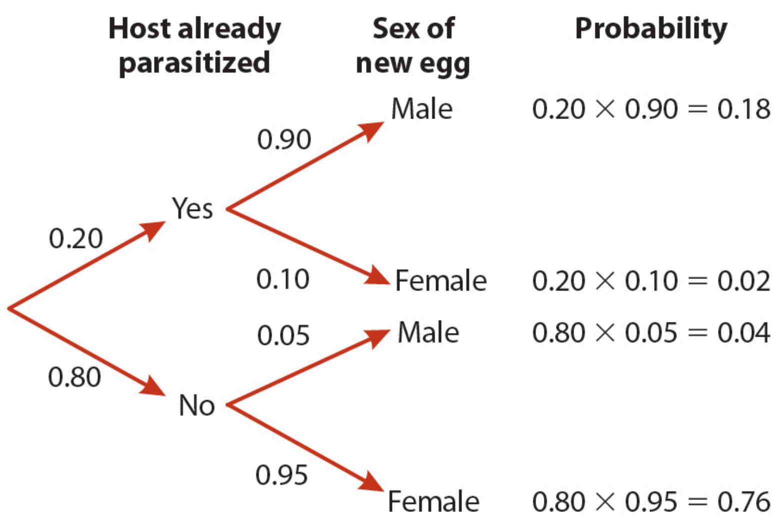


Fig 5.8-2

30

General multiplication rule

- The **general multiplication rule** finds the probability that both of two events occur, even if the two are dependent:

$$\Pr[A \; and \; B] = \Pr[A] \Pr[B|A]$$

31

Notes

- Skipping section 5.9

32