Physical Principles in Biology
Biology 3550
Fall 2016

# Lecture 5

# Introduction to Probability Theory

Wednesday, 31 August

©David P. Goldenberg
University of Utah
goldenberg@biology.utah.edu

# Probability: Some Definitions

- Outcomes – Possible results of a probabilistic experiment
  - For a coin toss: coin lands heads-up ($H$) or tails up ($T$)
  - For a roll of a six-sided die: The number of spots on the side that lands up (1, 2, 3, 4, 5 and 6)
  - Distinguished from "events", to be defined shortly
- Probability – A number with a possible value from 0 to 1, associated with a single outcome.
  - $p = 0$: Outcome will never occur.
  - $p = 1$: Outcome will always occur.
  - The sum of the probabilities of all possible outcomes of an experiment must equal 1.

# Two Interpretations of Probabilities

1. The frequentist interpretation
    - If the same experiment is repeated a large number, $N$, times, an outcome with probability $p$ will occur approximately $N \cdot p$ times.
    - “Law of large numbers”
    - Value of probabilities are defined by properties of the experiment.
2. The Bayesian interpretation
    - Quantity used to express (limited) knowledge or belief.
    - Probability can be updated using additional information.
    - Thomas Bayes (1702-1761): Equation for calculating revised probabilities.

# The Sample Space, $S$

- Set of all possible outcomes
  - For a single coin toss:
    $S = \{T, H\}$
    Curly braces are used to indicate sets.
  - For two independent coin tosses:
    $S = \{(H, H), (H, T), (T, H), (T, T)\}$
    Ordered pairs representing the results of the two tosses.
  - Are there other possible sample sets that could be defined for two coin tosses?
- The sample set must be complete, *i.e.* it must include all possible outcomes.
- The elements in the sample set must not overlap.
- Often most convenient to define sample set so that probabilities of all outcomes are equal. But, this isn't required.

# Clicker Question #1

How many outcomes are there in the (simplest) sample set for three coin tosses?

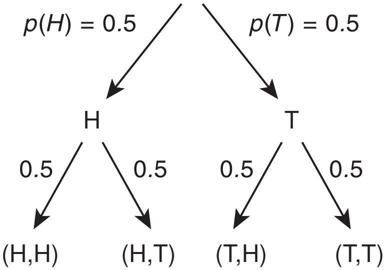
1. 3
2. 4
3. 6
4. 8
5. 10

# Events

- An event is a subset of the sample space.
- Some possible events defined for two coin tosses:
  - Two heads: $2H = \{(H, H)\}$
  - Two tails: $2T = \{(T, T)\}$
  - One heads and one tails: $1H1T = \{(H, T), (T, H)\}$
- The outcomes defined in the sample space are events, but additional events can usually be defined.
- Some other events that can be defined for two coin tosses:
  - One or more heads: $1^{+}H = \{(H, H), (H, T), (T, H)\}$
  - One or more tails: $1^{+}T = \{(H, T), (T, H), (T, T)\}$

# Calculating Probabilities: Sequential Sub-events

- Two coin tosses:




- Probabilities are multiplied

$$p\big((H,H)\big) = p(H)p(H) = 0.5 \times 0.5 = 0.25$$
$$p\big((H,T)\big) = p(H)p(T) = 0.5 \times 0.5 = 0.25$$
$$p\big((T,H)\big) = p(T)p(H) = 0.5 \times 0.5 = 0.25$$
$$p\big((T,T)\big) = p(T)p(T) = 0.5 \times 0.5 = 0.25$$

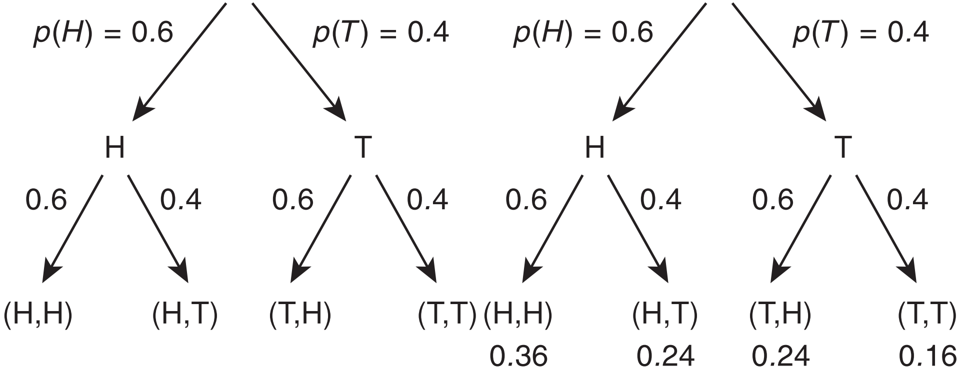
- Multiplication of probabilities is usually associated with “and”.

# Clicker Question #2

What is the probability of three heads in three coin tosses?

1. 0
2. 1/8
3. 1/3
4. 3/8
5. 3/4
6. 1

# Two Tosses of a Bad Coin




- Sum of probabilities:

$$p\big((H,H)\big) + p\big((H,T)\big) + p\big((T,H)\big) + p\big((T,T)\big) = 1$$

# Calculating Probabilities: Groups of Non-overlapping Outcomes or Events

- An event defined earlier for two coin tosses:

  One heads and tails: $1H1T = \{(H,T),(T,H)\}$

- Probability is calculated as a sum:

$$
\begin{aligned}
p(1H1T) &= p\big((H,T)\big) + p\big((H,T)\big) \\
&= p(H)p(T) + p(T)p(H) \\
&= 0.5 \times 0.5 + 0.5 \times 0.5 \\
&= 0.25 + 0.25 = 0.5 \qquad \text{(for a fair coin)}
\end{aligned}
$$

- This only works if the grouped events or outcomes do not overlap!
- Addition of probabilities is usually associated with "or".

# Another Example

- An event defined for two coin tosses:

  One or more heads: $1^{+}H = \{(H,H),(H,T),(T,H)\}$

- Calculation of probability:

$$p(1^{+}H) = p\big((H,H)\big) + p\big((H,T)\big) + p\big((H,T)\big)$$

$$= 0.25 + 0.25 + 0.25 = 0.75 \qquad \text{(for a fair coin)}$$

- Another way:
  - $1^{+}H$ includes all of the sample set, except $(T,T)$
  - For the entire sample set, the sum of probabilities is 1.

$$p(1^{+}H) = 1 - p\big((T,T)\big)$$

$$= 1 - 0.25 = 0.75 \qquad \text{(for a fair coin)}$$

# Clicker Question #3

What is the probability of exactly two heads in three coin tosses?

1. 0
2. 1/8
3. 1/4
4. 3/8
5. 1/2

# Clicker Question #4

A coin has been tossed 10 times and has landed heads-up each time. What is the probability that it will land heads-up the next time?

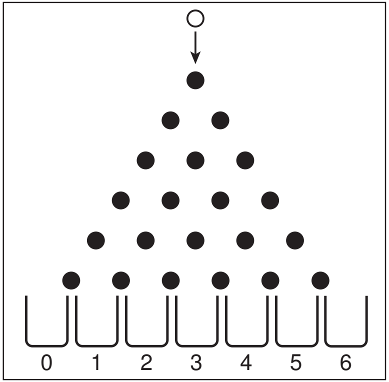
1. 0
2. Greater than zero but less than 1/2
3. 1/2
4. Greater than 1/2 but less than 1
   (The Bayesian interpretation.)
5. 1

# Clicker Question #5

The king has one sibling. What is the probability that the king's sibling is a sister?

1. 1/4
2. 1/3
3. 1/2
4. 2/3
5. 3/4

# A Six-row Plinko




- For $N$ plinko rows, there will be $N+1$ buckets for balls to land in.
- For convenience, buckets are numbered from 0 to $N$.
- How shall we define the sample set?

# Count The Paths to Reach a Given Bucket

- Define outcomes as all of the possible paths.
- Define events as final positions of ball, *i.e.* bucket numbers.
- How many possible paths are there?
  - For a 6-row plinko, each path involves 6 places to change direction.
  - The number of different paths is: $2 \times 2 \times 2 \times 2 \times 2 \times 2 = 2^6 = 64$
  - Each path has an equal probability, equal to 1/64
  - For an $n$-row plinko, the number of different paths is $2^n$, and the probability of each is $1/2^n = 2^{-n}$.