Physical Principles in Biology
Biology 3550
Fall 2018

Lecture 5:
Introduction to Probability

Wednesday, 29 August 2018

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

Announcements

- First Quiz: Friday, 31 Aug.
  25 min, second half of class.
- First problem set
  Posted on class web page; due Tuesday, 4 Sept.
- Discussion Sessions.
  - Tuesdays, 11:00 AM – noon. Life Science 102
  - Wednesdays, 8:30 – 9:30 AM. Gardner Common 4680
- Office hours:
  - Tuesdays: 9:30 - 11:00 AM
  - Wednesdays: 2:30 - 3:30 PM
  - Other times by appointment.
    Send me an e-mail message!

Clicker Question #1

Steve is very shy and withdrawn, invariably helpful but with very little interest in people or in the world of reality. A meek and tidy soul, he has a need for order and structure, and a passion for detail.

Is Steve more likely to be a librarian or a farmer?

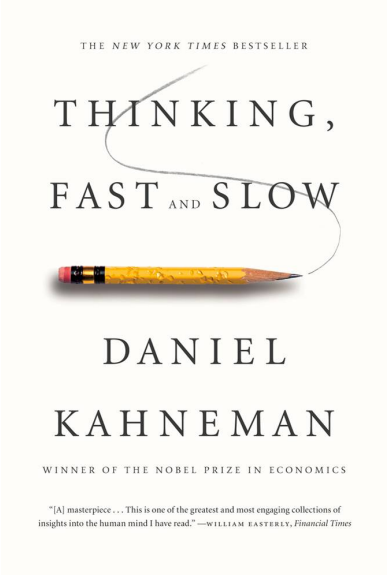
A) Librarian

B) Farmer

All answers count for now!

- Key information: There are about 15-times as many farmers than librarians in the US.
  From the Bureau of Labor Statistics Occupational Outlook Handbook:
  `https://www.bls.gov/ooh`

Thinking Fast and Slow



- Kahnemann is a psychologist who won the Nobel Prize in Economics for research on how people make decisions.
- With his collaborator, Amos Tversky, Kahnemann determined that people respond to questions or problems in two distinct ways:
  - Fast, instinctual way that is often right, but sometimes very wrong.
  - Slow, deliberate analysis.
- Highly recommended book!

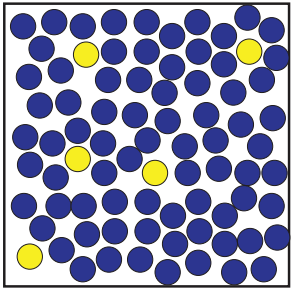
The Fast Way of Answering the Question

- Steve has the characteristics I associate with a librarian: Shy, with a need for order and structure.
- It makes more sense that Steve would be a librarian than a farmer.

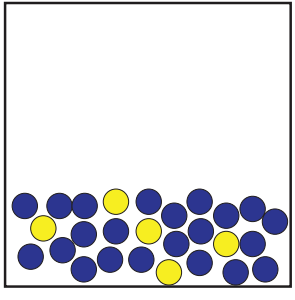
The Slow, Probabilistic Way of Answering the Question

- Steve is one of a large number of people in a population that are either librarians or farmers.
- Represent farmers and librarians as marbles in a box.

All farmers (blue) and librarians (yellow)



Shy, organized farmers and librarians



- If we pick a marble at random, are we more likely to pick a blue marble (farmer) or a yellow one (librarian)?

Probability: Some Definitions

- Outcomes – Possible results of a probabilistic process
  - 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.
  - What do we mean by this? What is implied?

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.
   - May not be able to determine probability from first principles or experiment.
   - Probability can be updated using additional information.
   - Thomas Bayes (1702-1761): Equation for calculating revised probabilities.
   - Somewhat controversial, but very important area in modern applications of probability and statistics.

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.
- The sum of the probabilities for all of the elements in the sample set must equal 1.

Clicker Question #2

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

A) 3

B) 4

C) 6

D) 8

E) 10

The Sample Set for Three Coin Tosses

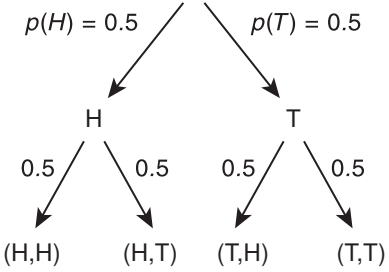
$$S = \{(H, H, H), (H, T, H), (H, H, T), (H, T, T)$$
$$(T, H, H), (T, T, H), (T, H, T), (T, T, T)\}$$

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 Trials

- 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 #3

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

A) 0

B) 1/8

C) 1/3

D) 3/8

E) 3/4

F) 1

The Probability of Three Heads in Three Coin Tosses

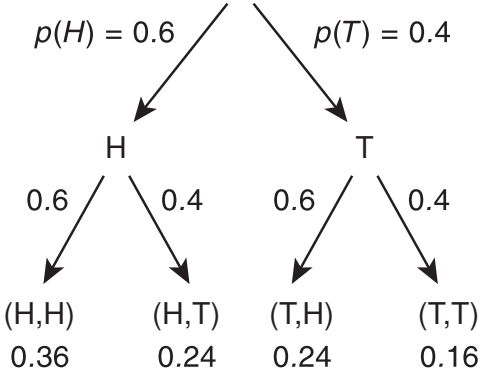
- Three sequential events

$$\xrightarrow{p=0.5} \mathrm{H} \xrightarrow{p=0.5} \mathrm{H} \xrightarrow{p=0.5} \mathrm{H}$$

$$p = \frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} = \frac{1}{8}$$

- Consider the sample set:
  - There are 8 elements in the sample set.
  - All outcomes in the sample set have equal probabilities (if the coin is fair).
  - For each outcome, $p = 1/8$.
  - Only one outcome has 3 heads, (H,H,H).

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 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:

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

- This only works for events if they are not overlapping!
- 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 #4

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

A) 0

B) 1/8

C) 1/4

D) 3/8

E) 1/2

The Probability of Exactly Two Heads in Three Tosses

- The sample set for 3 coin tosses:

$$S = (H, H, H), (H, T, H), (H, H, T), (H, T, T)$$
$$(T, H, H), (T, T, H), (T, H, T), (T, T, T)\}$$

How many outcomes have exactly 2 heads:

$$(H, T, H), (H, H, T), (T, H, H)$$

- Another way:
  - An outcome with exactly 2 heads has exactly 1 tails!
  - There are three places in the sequence to place the 1 tails.
  - Three outcomes have exactly 2 heads.

Clicker Question #5

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?

A) 0

B) Greater than zero but less than 1/2

C) 1/2

D) Greater than 1/2 but less than 1

E) 1