Lecture 11
Introduction to probability

Quantifying Uncertainty

- Statisticians rely on randomization to avoid bias when we gather data:
  randomly sampling subjects from a population in a survey
  randomly assigning subjects to treatments in an experiment

- Sampling brings <u>uncertainty</u> because, by chance, properties of the sample are not the same as the true values in the population

- Random sampling guarantees that the results we get from analyzing our data follow *the laws of mathematical probability*

Probability in our daily lives:

- Many things in life are the result of random phenomena
  - Weather, the stock market, automobile accidents, winning the lottery
- We often have to make decisions about uncertain events
  - Ex: Should I bring a raincoat with me to work today?
  - Ex: If I drive my car today, what is the risk I get into an accident?

The language of Probability

- In probability theory the real world is translated into “events” which are the outcomes of random phenomena

- An **event** is any possible outcome of a random phenomena or experiment

- **Probability of event A** – denoted mathematically as $P(A)$ **-** is a number between 0 and 1 that represents the likelihood of event A occurring

  Ex. the chance of flipping a coin and getting heads

  Ex. the chance of it raining this afternoon

Probability describes only what happens in the long run, assuming a large number of repetitions!

Example: Shuffling Songs

- Imagine that you have 1000 songs on your phone, each of them recorded exactly once. When you push the “shuffle” button, the phone plays a song at random from the list of 1000.

- What is the probability that the song that plays is your single favorite song?

$$P(The\ song\ that\ plays\ is\ my\ favorite\ song) = \frac{1}{1000}$$

- What is the probability that the song that plays if not your favorite song?

$$P(The\ song\ that\ plays\ is\ not\ my\ favorite\ song) = \frac{999}{1000}$$

Understanding Random Trials

- A **random trial** is a process or experiment that has a set of well – defined possible outcomes.

  - what makes it random is having more than one possible outcome

  - In our songs example, a random trial is defined as pushing the shuffle button once

  - The event we were interested in was "the song played was your favorite" which is one out of 1000 possible outcomes.

The sample space

- The **sample space** of an experiment or random phenomena is the collection of all possible outcomes
  - we will represent the sample space with $S = \{\ldots\}$

- In the songs example the sample space consists of all 1000 songs that could have been played

Another Example: Flipping A Coin

- Flipping a coin once is an example of **trial.** What are the possible outcomes that can occur for a single flip?

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

- Getting "heads" (or "tails") is an example of an event. We can think of events as subsets of the possible outcomes (i.e subsets of the sample space)

- In general, the probability of an event A is:

$$P(A) = \frac{\text{The number of ways it can happen}}{\text{Total number of outcomes}}$$

Coin Flip Continued…

- We will use n to denote the number of trials of an experiment or random phenomena.
- Suppose I flip a coin twice and observe the outcomes of both flips. How many trials have I performed?

$$n = 2$$

- What is the sample space of possible outcomes of two flips of a coin?

$$S = \{HH, HT, TH, TT\}$$

- What is the probability that both flips are tails?

$$P(\text{both tails}) = \frac{1}{4} = 0.25$$

Lets try another example

Suppose I roll a fair six-sided die:

What is the event we are interested in ?

$$Rolling\ a\ 6$$

What is the sample space?

$$S = \{1,2,3,4,5,6\}$$

what is the probability I roll a six?

$$P(\text{Rolling a 6}) = \frac{1}{6}$$

The rules of probability

1. $0 \leq P(A) \leq 1$ - the probability of event must be a value in [0,1]

2. $\sum_i P(A_i) = 1 = S$ - the probabilities of the sample space must add to 1

3. $P(A \text{ or } B) = P(A) + P(B)$ two events A and B are **disjoint** if they have no outcomes in common.

4. $P(A') = 1 - P(A)$ - the complement of the probability of A is 1 minus the probability of A

- the symbol A' means complement of A or "A not" and

Rule 3 explained

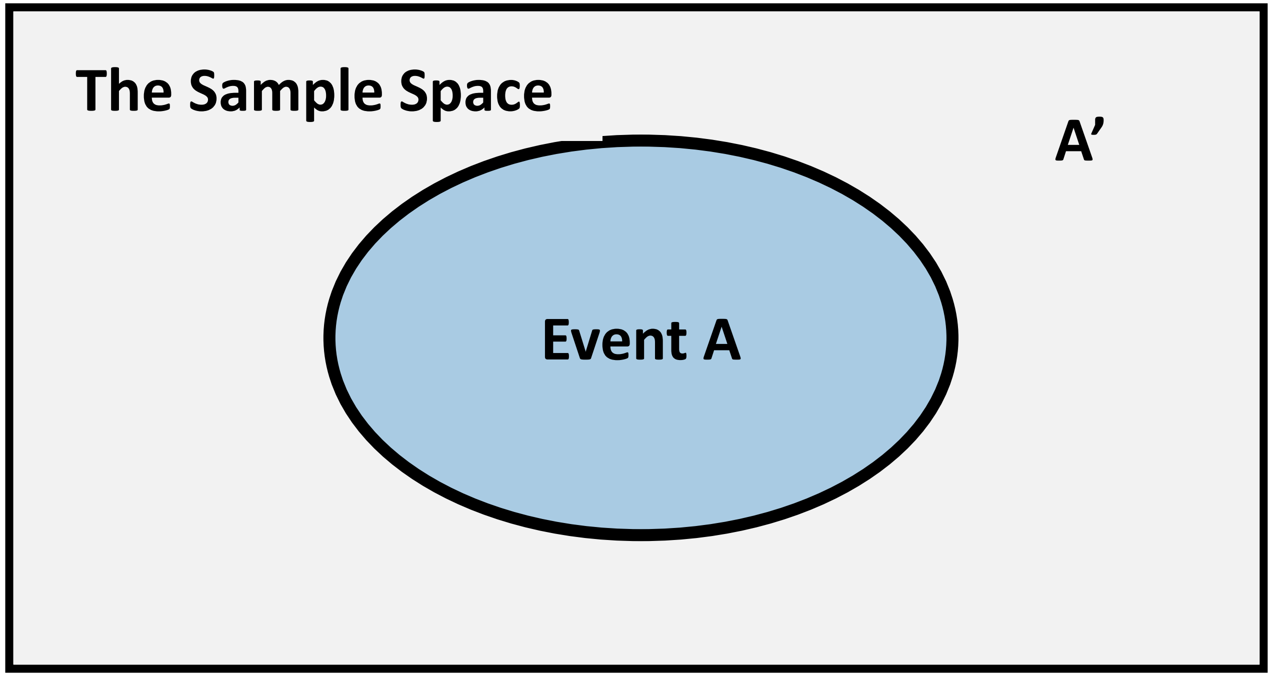
Ex. What is the probability we roll a fair six –sided die and get a 6?

$A = \text{roll a 6}$

$A' = \text{Not A} = \text{roll a value other than 6}$

$$P(A) = \frac{1}{6}$$

$$P(A') = 1 - P(A) = 1 - \frac{1}{6} = \frac{5}{6}$$

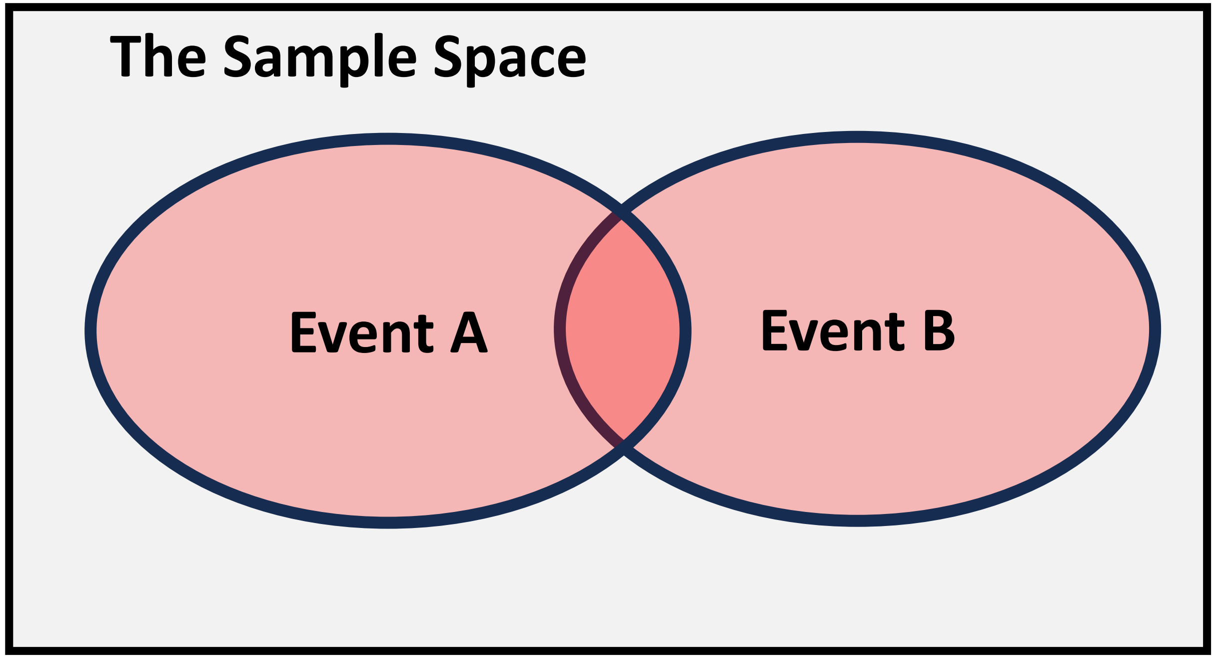


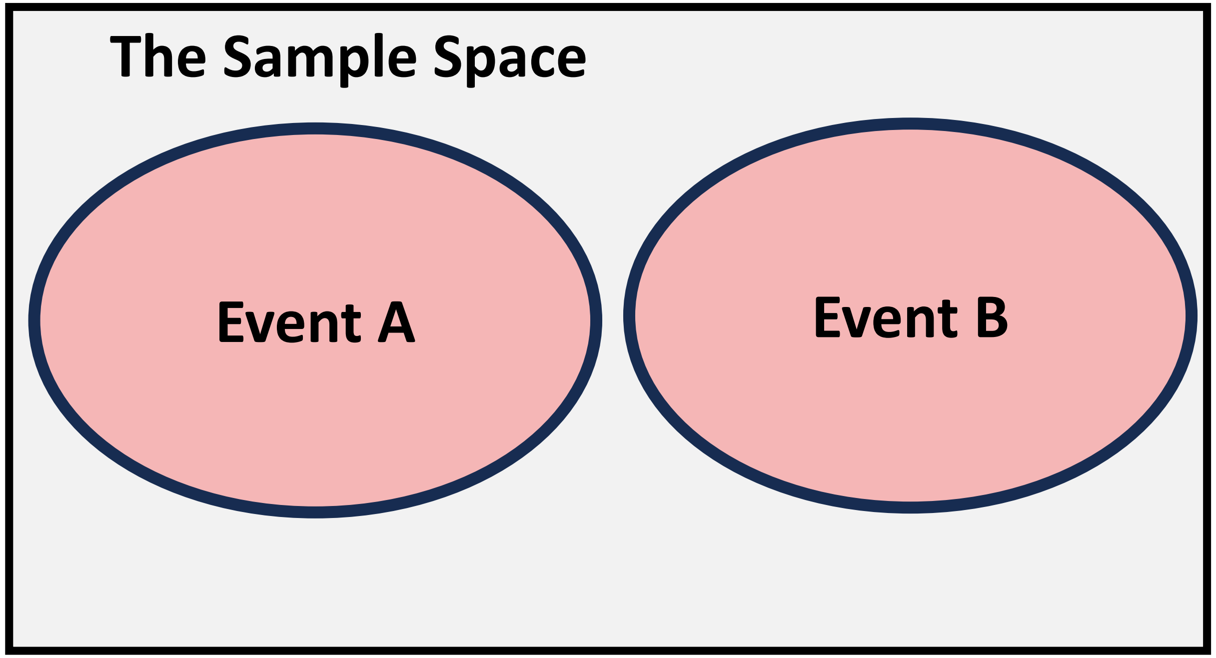
Unions of events

- The **union** of two events A and B is all the outcomes of both events: e.g the probability that either event A occurs or event B occurs

  - key word to watch out for to identify a union is the “or”

  - represented mathematically as $P(A \cup B)$ meaning $P(A\ or\ B)$



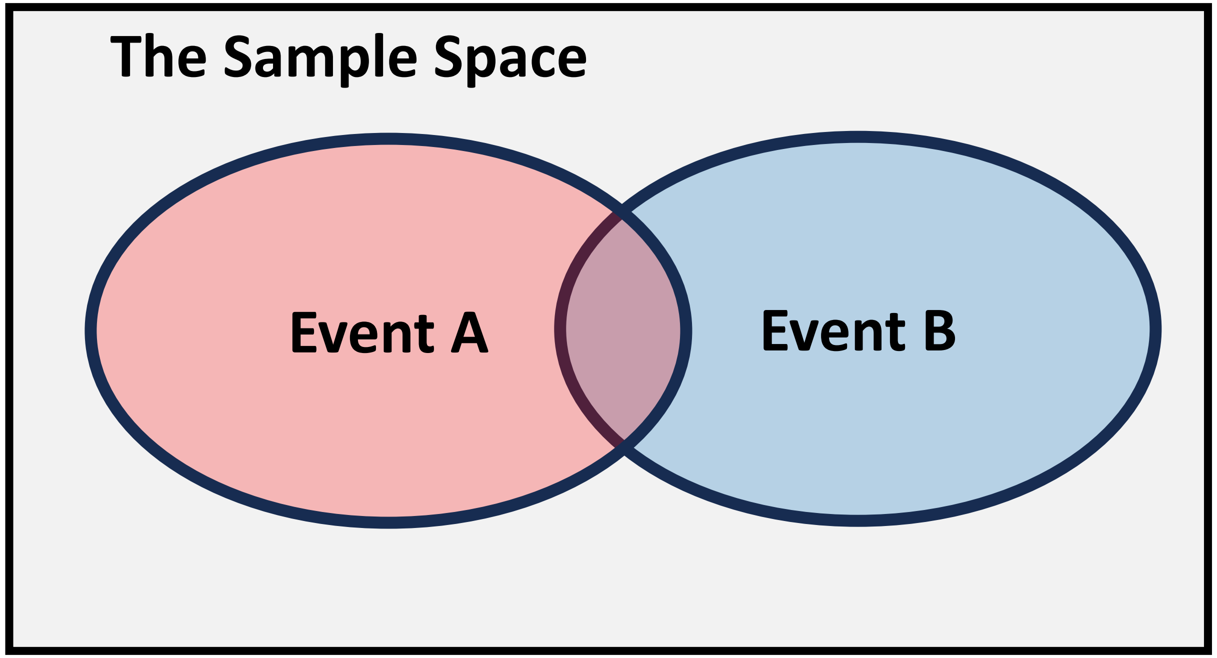


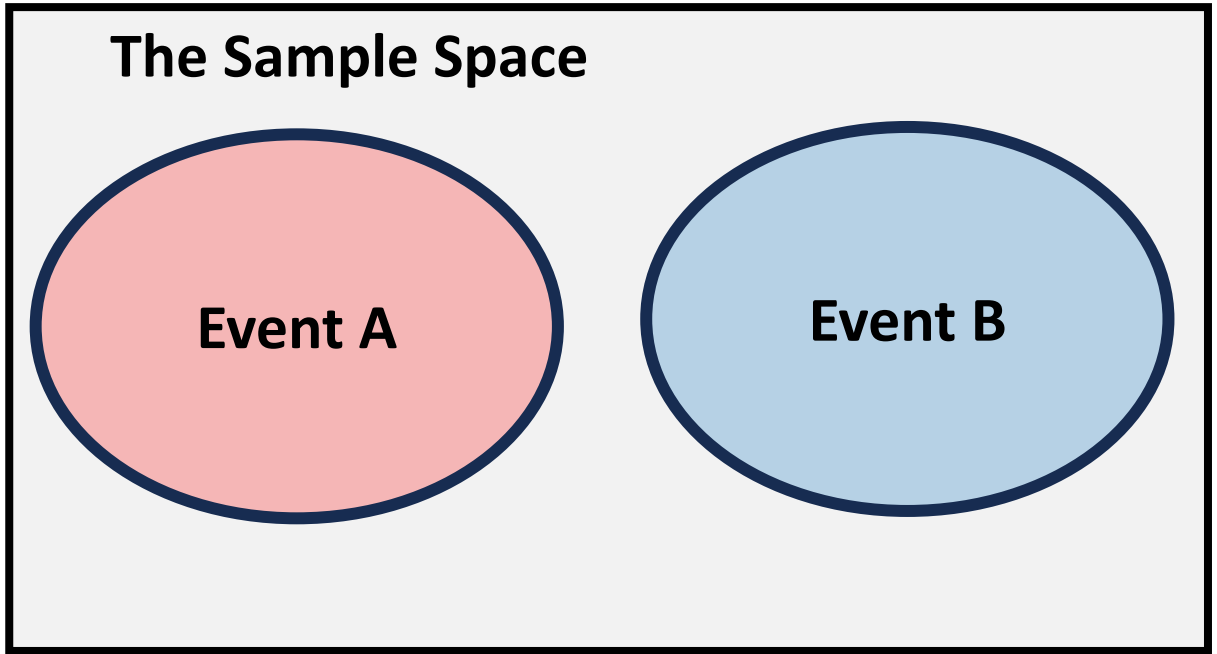
Intersections of Events

- The **intersection** of two events A or B is the set of events that A and B have in common: the probability that A and B occur

  - key word to watch out for to identify an intersection is the "and"

  - represented mathematically as $P(A \cap B)$ meaning $P(A\ and\ B)$





Where do probabilities come from?

Think back to flipping a fair coin a single time, what is the probability that we get “heads”?

- Our intuition tells us that it is ½ if we have an equal chance of getting heads or tails

- how can we be sure that this is the likelihood of getting heads?

Key concept:

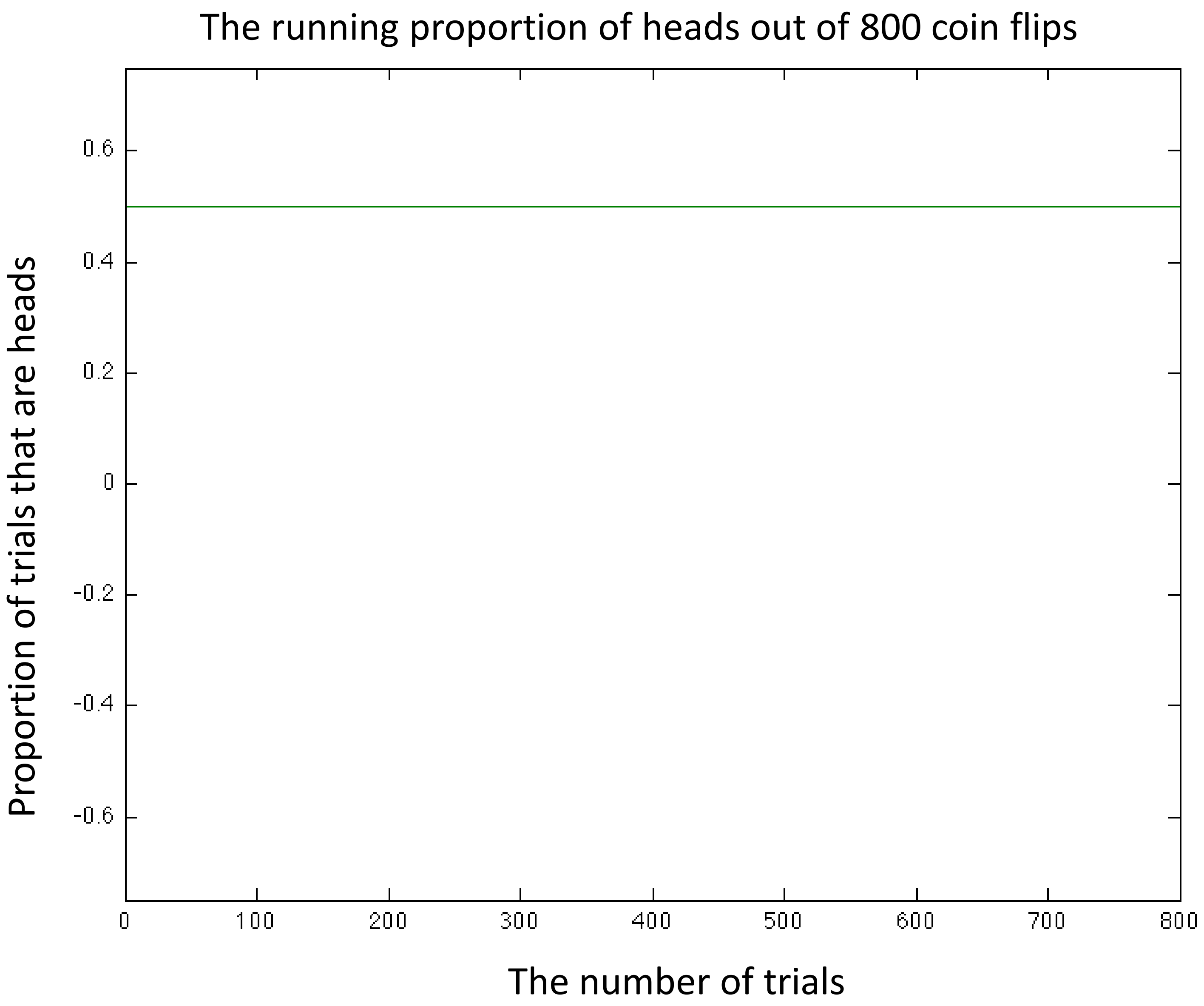
With random phenomena, the proportion of times that an event happens can be quite unpredictable in the short run, but very predictable in the long run

P<u>robabilities quantify the “long-run” behavior of random phenomena.</u>

The Law of Large Numbers

- In 1689, the Swiss mathematician Jacob Bernoulli proved that as the number of trials increases the proportion of occurrences of any given outcome approaches a particular number.

- **Law of Large Numbers** (LLN) states that the relative frequency of an event will tend to “approach” (in some sense) the probability of an event as the number of independent observations increases



Independent trials and finding probabilities

- To prove the LLN Bernoulli had to demonstrate that each flip of a coin was an independent trial

- Trials of a random phenomena are **independent trials** if the outcome of any one trial is not affected by the outcome of any other trial.

- In some cases, its reasonable to assume that the outcomes of an event are equally likely,

  - Ex. We assumed that flipping a coin was 50/50 because its reasonable to assume that the coin is equally balanced

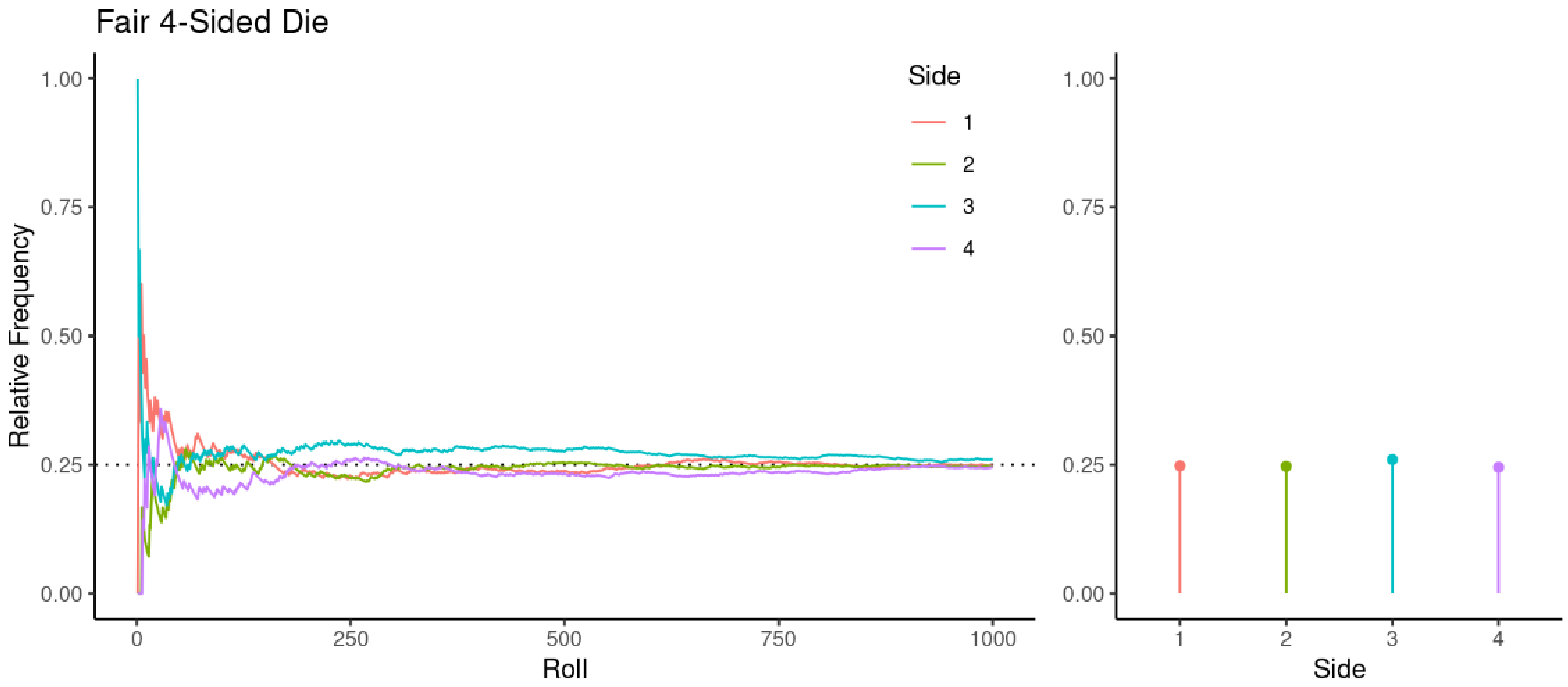
Finding probabilities cont.

- In other cases, we can use the long-run behavior of different outcomes as approximations of their probability
  - In other words, we can use the relative frequency as an estimate of the probability.

- Its important to note that these estimates depend on the size of n and will get better as n gets larger

The Law of Large Numbers

Ex.) Consider rolling a fair 4-sided die many times and looking at the distribution of the sides.



The Law of Large Numbers

Ex.) Consider rolling a loaded 4-sided die many times and looking at the distribution of the sides.