OI 3.1: Introduction to Probability

Probability

- Probability forms a foundation for statistics.
- You might already be familiar with many aspects of probability.
- Formalization of probability concepts is new for most.

Defining Probability

- A **trial/random process** gives rise to an outcome.
- The **support** $(\mathcal{S})$ is the set of all possible outcomes for a random process.
- An **event** is a collection of outcomes in the sample space.
- A **probability** is the chance or likelihood for an event.

Axioms (In words with no math symbols)

1) The probability of a particular event happening is always greater than or equal to 0.
2) The probability that any of the possible outcomes of a process (sample space) will occur is 1.
3) For a set of distinctly different events the probability of at least one occurring is the sum of probability of each event.

Axioms (In an elevator)

- The probability of *anything* (including everything and nothing) will always be between 0 and 1.
- Probabilities increase as you consider more outcomes.

Example 1

Suppose we flip a coin. We are interested in how likely it is to land on a head. Identify the following:

trial/random process

- *Flip a coin*

sample space

event of interest

Example 1

Suppose we flip a coin. We are interested in how likely it is to land on a head. Identify the following:

trial/random process

- *Flip a coin*

sample space

- $\{H, T\}$

event of interest

Example 1

Suppose we flip a coin. We are interested in how likely it is to land on a head. Identify the following:

trial/random process

- *Flip a coin*

sample space

- $\{H, T\}$

event of interest

- $\{H\}$

Example 2

Suppose we role a fair 6-sided die (d6). We are interested in how likely this d6 is to land on a 5 or a 6. Identify the following:

trial/random process

- *Role a d6*

sample space

event of interest

Example 2

Suppose we role a fair 6-sided die (d6). We are interested in how likely this d6 is to land on a 5 or a 6. Identify the following:

trial/random process

- *Role a d6*

sample space

- *{1, 2, 3, 4, 5, 6}*

event of interest

Example 2

Suppose we role a fair 6-sided die (d6). We are interested in how likely this d6 is to land on a 5 or a 6. Identify the following:

trial/random process

- *Role a d6*

sample space

- *{1, 2, 3, 4, 5, 6}*

event of interest

- *{5, 6}*

Example 3

Suppose we measure the height of an adult bison (in feet). We are interested in if a bison is between 4ft and 6ft. Identify the following:

trial/random process

- *Height of a bison in feet*

sample space

event of interest

Example 3

Suppose we measure the height of an adult bison (in feet). We are interested in if a bison is between 4ft and 6ft. Identify the following:

trial/random process

- *Height of a bison in feet*

sample space

- *[3ft, 7ft]*

event of interest

Example 3

Suppose we measure the height of an adult bison (in feet). We are interested in if a bison is between 4ft and 6ft. Identify the following:

trial/random process

- *Height of a bison in feet*

sample space

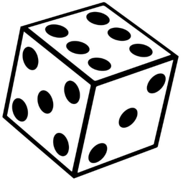
- *[3ft, 7ft]*

event of interest

- *[4ft, 6ft]*

How would we find these probabilities?

- Suppose we flip a fair sided coin. What is the probability we land on a head?
- Suppose we role a fair d6. What is the probability we land on a 5 or a 6?
- Suppose we measure the height of an adult bison (in feet). What is the probability the height is between 4ft and 6ft?



Different Types of Probabilities

It is helpful to distinguish between different types of probabilities.

- Theoretical Probabilities
  - *theory*: a system of ideas intended to explain something, especially one based on general principles independent of the thing to be explained.
- Empirical Probabilities

Different Types of Probabilities

It is helpful to distinguish between different types of probabilities.

- Theoretical Probabilities
  - *theory*: a system of ideas intended to explain something, especially one based on general principles independent of the thing to be explained.
- Empirical Probabilities
  - *empirical*: based on, concerned with, or verifiable by observation or experience

Theoretical Probability

- Based on a statistical, physical, or biological principal. It is a probability we would expect "in theory".
- We typically do not *know* this value because
  - experts are still studying it
  - not possible to record every value
- The theoretical probability of event A is typically denoted p

$$p := P(A)$$

Empirical Probabilities

- Sample probability
- Estimated based on observed data

Suppose

- $X =$ the number of times we observed the event
- $n =$ the number of trials

Then the empirical probability for our event is

$$\hat{p} = \frac{X}{n}$$

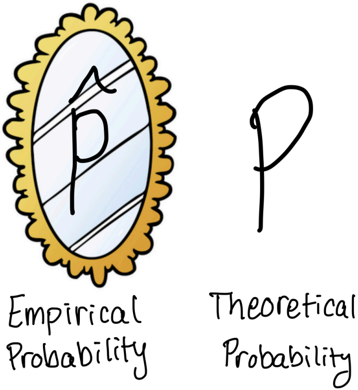
Empirical vs Theoretical Probability

Both types of probabilities are

- used all the time
- valid
- obey the axioms, and other probability rules and theorems we will go over.

Empirical vs Theoretical Probability

- Empirical probabilities are used to *estimate*. The we check this estimate with what we theorize as the *true (population) probabilities*.
- If the sample is good and the theory is true, then they should agree!



Fundamental Theorem of Statistics

Fundamental Theorem of Statistics

By the ***Fundamental Theorem of Statistics***, we expect that our empirical probabilities converge to the true probabilities as the sample size increases.

$$\hat{p} \to \text{True Probability}$$

...A moment to pause.

There are actually *three* probabilities statisticians care about:

- The empirical probability (comes from the sample)
- The theoretical probability (comes from our theory/guess)
- The truth (we do not usually know this)

...A moment to pause.

- The FTS connects says that as the sample size gets larger, the *empirical probability* will converge to the *true probability*.
- The FTS doesn't actually have anything to do with the *theoretical probability*.
- If the empirical probability is similar to our theoretical probability, then our theory was likely true.
- More on this later....

Fundamental Theorem of Statistics

- This theorem goes by many names: *Central Limit Theorem*, *Law of Large Numbers*, etc.
- Connects the different types of probabilities together.
- Many different versions of this theorem.
- Bigger samples are more likely to converge to the truth, but how big should it be?

Seeing the Theorem in Action

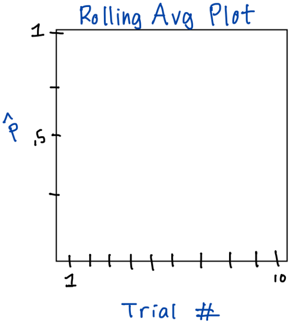
Suppose we flip a coin 10 times and obtain the following results:

$$H, H, T, T, T, T, H, T, T, T$$

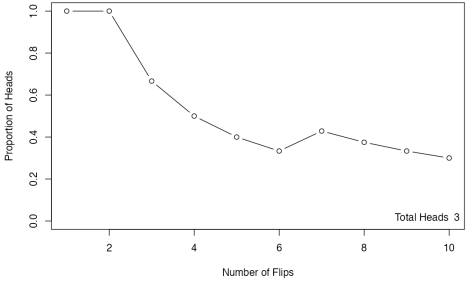
a) Fill out the following table.

b) Make a *rolling average* plot.

Trial #	$\hat{p}$
1	
2	
⋮	



Seeing the Theorem in Action



Computers can help us visualize this theorem. CLICK HERE

Theorem in Action

- The FTS works even as situations get more complicated!
- Suppose we want to know the probability of observing **three heads or less** if we flip 10 coins. We can still use this theorem!

Trial #	Trial Results	# Heads in Trial	$\hat{p}$
1	H, H, T, T, T, T, H, T, T, T	3	1/1
2	H, T, T, T, T, T, T, T, T, T	1	2/2
3	T, T, T, T, T, H, H, H, H, H	5	2/3

- CLICK HERE

The Moral

If we have enough data, eventually the *estimated* probability will converge to the *true* probability no matter how complicated.

More Probability Concepts

More on probability

- For now, we are going to concentrate on probability concepts and not so much on bridging the two different methods.
- We will typically use empirical probabilities with survey data.
- We will typically use theoretical probabilities with games, dice, cards, etc.
- The important thing to remember is that the following rules and concepts *always work.*
- So it does not matter if you are working with empirical probabilities, or are given information, the concepts we learn now will apply to *all facets of probability*.

Probability Concepts

- When working with random events we often want to know more than the probability of just one outcome in the sample space $\mathcal{S}$
- For example, when flipping 10 coins and counting the number of heads our sample space is $\mathcal{S} = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$. We could ask:
  - What is the probability of observing exactly 5 or exactly 6 heads?
  - What is the probability of observing anything except a 5 heads?
  - What is the probability of observing a number that is even or (strictly) larger than 3?

Probability Concepts

- We could answer the questions on the previous slide using the computer again, or flipping a lot of coins, or looking at every piece of data individually.
- However, sometimes we do not have the ability to do this.
- We can instead rely on the axioms!
- From the axioms we can find general mathematical rules that can let us skip simply counting occurrences.
- ...but first we need some key words.

Probability Concepts

For *a single random process*, we have the following definitions:

- **disjoint/mutually exclusive:** two outcomes, say A and B, that cannot both happen simultaneously.

$$P(A \text{ and } B) = 0$$

- **compliment**: All events in the sample space ($\mathcal{S}$) that are not in the event of interest. The complement of event A is denoted A^c, and

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

- **probability distribution:** map of disjoint outcomes and their associated probabilities.

Probability Concepts

We have seen probability distributions before, but they were in plots.

Suppose we flipping 10 coins and are interested in the number of heads. The distribution is below (rounded).

# of Heads	0	1	2	3	4	5	6	7	8	9	10
Probability	0.001	0.01	0.04	0.12	0.20	0.25	0.20	0.12	0.04	0.01	0.001

1) Are the events of observing exactly 5 or exactly 6 heads mutually exclusive?
2) What is the probability of observing anything except a 5?
3) What is the probability of observing a number that is even or (strictly) larger than 3?

General Addition Rule

If A and B are any two events, then the probability that at least one of them will occur is

$$P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B)$$

where $P(A \text{ and } B)$ is the probability that both events occur.

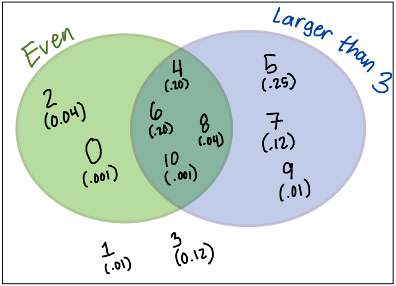
General Addition Rule

- The general addition rule formalizes relationships between multiple outcomes of random process.
- Note: If A and B are mutually exclusive then the rule simplifies to $P(A \text{ or } B) = P(A) + P(B) - 0$.

General Addition Rule

# of Heads	0	1	2	3	4	5	6	7	8	9	10
Probability	0.001	0.01	0.04	0.12	0.20	0.25	0.20	0.12	0.04	0.01	0.001

- Sometimes it is helpful to visualize this in a Venn diagram
- What is the probability of seeing a number that is even or (strictly) larger than 3?



Independence

For *two random processes*, we have the following definition:

- Two random processes are **independent** if knowing the outcome of one provides no useful information about the outcome of another.

$$P(A \text{ and } B) = P(A) \times P(B)$$

- If this does not hold the random processes are *dependent*

Practice Question

[OI 3.7] Swing voters. A Pew Research survey asked 2,373 randomly sampled registered voters their political affiliation (Republican, Democrat, or Independent) and whether or not they identify as swing voters. 35% of respondents identified as Independent, 23% identified as swing voters, and 11% identified as both.

a) Are being Independent and being a swing voter disjoint, i.e. mutually exclusive?
b) Draw a Venn diagram summarizing the variables and their associated probabilities.
c) What percent of voters are Independent but not swing voters?
d) What percent of voters are Independent or swing voters?
e) What percent of voters are neither Independent nor swing voters?
f) Is the event that someone is a swing voter independent of the event that someone is a political independent?

A Practice Question

[OI 3.2] Roulette wheel. The game of roulette involves spinning a wheel with 38 slots: 18 red, 18 black, and 2 green. A ball is spun onto the wheel and will eventually land in a slot, where each slot has an equal chance of capturing the ball.

a) You watch a roulette wheel spin 3 consecutive times and the ball lands on a red slot each time. What is the probability that the ball will land on a red slot on the next spin?

b) You watch a roulette wheel spin 300 consecutive times and the ball lands on a red slot each time. What is the probability that the ball will land on a red slot on the next spin?

c) Are you equally confident of your answers to parts (a) and (b)? Why or why not?

Practice Question

[OI 3.11] Educational attainment of couples for 2010 survey.

		Gender	
		Male	Female
	Less than 9th grade	0.07	0.13
	9th to 12th grade, no diploma	0.10	0.09
Highest	HS graduate (or equivalent)	0.30	0.20
education	Some college, no degree	0.22	0.24
attained	Associate's degree	0.06	0.08
	Bachelor's degree	0.16	0.17
	Graduate or professional degree	0.09	0.09
	Total	1.00	1.00

a) What is the probability that a randomly chosen man has at least a Bachelor's degree?

b) What is the probability that a randomly chosen woman has at least a Bachelor's degree?

c) What is the probability that a man and a woman getting married both have at least a Bachelor's degree? Note any assumptions you must make to answer this question.

d) If you made an assumption in part (c), do you think it was reasonable? If you didn't make an assumption, double check your earlier answer and then return to this part.