Starting Off:

1. What are the chances that you will go outside today?
2. What would those chances be if it is a weekend day?
3. What are the chances if it is it is raining really hard?
4. How did you go about coming up with your answer?

Foundation of Probabilities

Data Science Immersive

FLATIRON SCHOOL

Probability Background

- The goal of this unit is to make inferences about the population based on the sample.
  - If you want to predict the chance that a house will sell for more than a million dollars, you need to be able to calculate the probability, then update that probability based on some conditions.
- In order to learn about inference, we need to learn a few more things first. Inference requires an understanding of probability, random variables, and probability distributions. This lesson will focus on the first step, probability.
- Allows us to make informed decision
  - Should you make a living by purchasing lottery?
  - The Monte Carlo Fallacy
- Allows us to speak about uncertainty in an informed way
- **INTERVIEWS!**

Foundations of Probabilities

- Agenda today
  - Probability Notation
  - Set & Set Theory
  - Independent Probabilities
  - Conditional Probability

After today, you'll be able to...

- Translate an event of interest into common probability notation.
- Translate common probability notation into a phrase or sentence describing the event of interest.
- Compute and interpret set operations numerically and with diagrams.
- Determine if two events are independent, mutually exclusive or neither.
- Compute and interpret conditional events and conditional probabilities.

Probability Notation

Probability is the likelihood of an outcome. Before we can properly define probability, we must first define 'events.' It is helpful and convenient to denote the collection of events as a single letter rather than list all possible outcomes.

Event

a collection of outcomes, typically denoted by capital letters such as A, B, C, etc...

- Suppose we ask 30 students to record their eye color. We can define an event B to be blue eye color. In other words, let *B={blue eyes}*.

Probability Notation

Converting to Probability Notation

1. Identifying the outcome event of interest: {Getting a Tail when we toss a fair coin}.
2. Use a single letter or word to represent this outcome of interest: T={Getting a Tail when we toss a fair coin}, for instance.
3. State your interest in the probability of this outcome: P(T) which is read, "Probability of getting a Tail when we toss a fair coin."

Probability Notation

Now let's complicate things. Consider again tossing a fair coin. We stated P(T) was the probability we get a tail when we toss the coin. How would one write the probability statement if the outcome was getting two tails when the coin was tossed twice?

Applying the steps we get...

1. Identify the outcome of the event: *Getting a tail on both tosses of a fair coin.*
2. Use a single letter or word to represent this outcome of interest: *We can write this as T,T, where the first T represents the outcome of the first toss and the second T as the outcome of the second toss.*
3. State your interest in the probability of this outcome: *P(T,T) which is read, "Probability of getting a Tail on the first and second toss."*

*Note: Often the comma is eliminated and P(T,T) is written as P(TT).

Set Theory

Outcome Space

- The outcome space of a scenario is all the possible outcomes that can occur and is often denoted S. The outcome space may also be referred to as the sample space

- Consider the experiment where two fair six-sided dice are rolled and their face values recorded. Write down the outcome space.

- In probability theory, a set is denoted as a well-defined collection of **distinct** objects.

- Mathematically, you can define a set by S. If an element x belongs to a set S, then you'd write $x \in S$. On the other hand, if x does not belong to a set S, then you'd write $x \notin S$.

Set Theory

Consider the experiment where two fair six-sided dice are rolled and their face values recorded.

Directions: Write out the event in probability notation and then identify the outcome space.

Subset & Set Operations

Subset

Set T is a subset of set S if every element in set T is also in set S. The mathematical notation for a subset is $T \subseteq S$.

- T = {2,3,4}
- S = {1,2,3,4,5}
- $T \subseteq S$

Set Operations

Now that we know how to denote events, the next step is to use the set notation to represent set operations. Each operation will also be presented in a Venn Diagram. Set operations are important because they allow us to create a new event by manipulation of other events

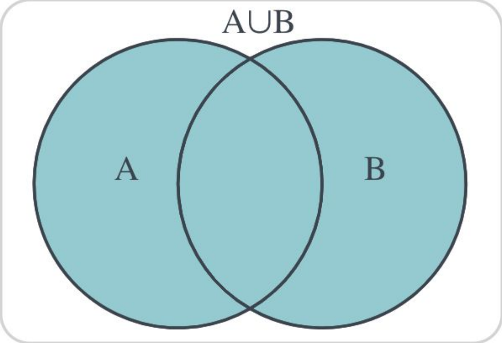
Set Operations

Union

The union of two events, A and B, contains all of the outcomes that are in A, B or both. In statistics, ‘or’ means at least one event occurs and therefore includes the event where both occur,

The union of A and B is denoted:

$$A \cup B$$



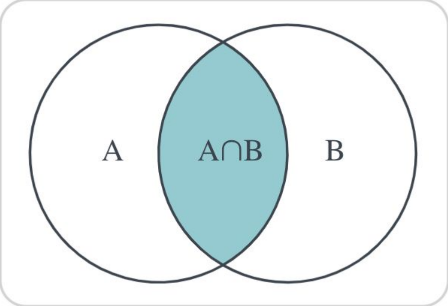
{outcomes in both A or B or both}

Intersection

The intersection of two events, A and B, contains all of the outcomes that are in both A and B.

The intersection is denoted by:

$$A \cap B$$



{outcomes in both A and B}

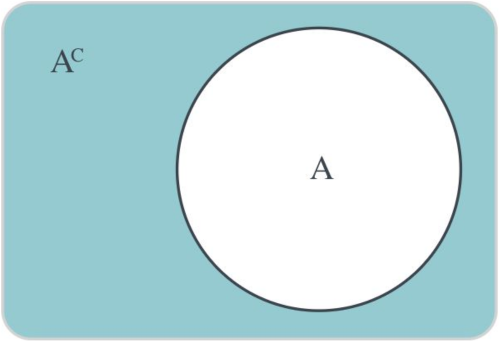
Set Operations

Complement

The complement of an event, A, contains all of the outcomes that are not in A.

The complement can be denoted as:

$$A^c, \bar{A}, \text{ or } A'$$



{outcomes not in *A*}

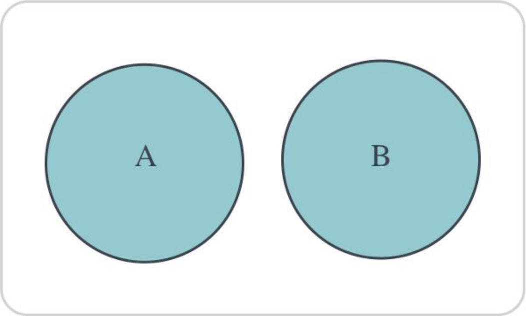
Mutually Exclusive

A and *B* are called mutually exclusive (or disjoint) if the occurrence of outcomes in *A* excludes the occurrence of outcomes in *B*.

There are no elements in $A \cap B$ and thus:

$$A \cap B = \varnothing$$

The empty set, denoted as $\varnothing$, is an event that contains no outcomes.



$$A \cap B = \varnothing$$

Set Operations

Suppose events A, B, and C are events of a particular scenario. Write the following using event notation.

1. At least one event occurs.
2. None of the events occur.
3. Only A occurs.

Set Operations

Let's go back to the example where we roll two fair six-sided die. Given the following events:

$$A = (3, 5)$$

$$B = \text{a 4 is rolled on the first die}$$

$$C = \text{a 5 is rolled on the second die}$$

$$D = \text{the sum of the dice is 7}$$

$$E = (7, 4)$$

Find $B \cap D$ and $B \cup D$

Practice:

- $B \cup C$
- $B \cap C$

Interpretations of Probability

Classical Interpretation of Probability

The probability that event E occurs is denoted by P(E). When all outcomes are equally likely, then:

$$P(E) = \frac{\textit{number of outcomes in } E}{\textit{number of possible outcomes}}$$

Subjective Probability

Subjective probability reflects personal belief which involves personal judgment, information, intuition, etc.

For example, what is P (you will get an A in a certain course)? Each student may have a different answer to the question.

Relative Frequency Concept of Probability (Empirical Approach)

If a particular outcome happens over a large number of events then the percentage of that outcome is close to the true probability.

For example, if we flip the given coin 10,000 times and observe 4555 heads and 5445 tails, then for that coin, P(H)=0.4555.

What is Probability?

Probability theory is the study on the frequency of a given event occurring in some context.

What is the probability of event *A* occurring?

- Event *A* is known as the event space, and all possible events are known as the sample space
- P(A) = A / (all possible events)
- For example, what is the probability of drawing an Ace in a deck of cards?

Probability Properties

Probability of an event

Probabilities will always be between (and including) 0 and 1. A probability of 0 means that the event is impossible. A probability of 1 means an event is guaranteed to happen.We denote the probability of event A as P(A).

$$0 \le P(A) \le 1$$

Probability of a complement

If A is an event, then the probability of A is equal to 1 minus the probability of the complement of A, A'.

$$P(A) = 1 - P(A') \text{ or } 1 = P(A) + P(A')$$

Probability Properties

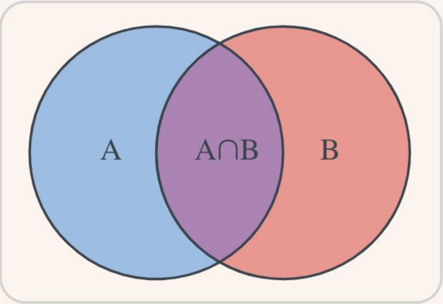
Probability of the empty set

If A and B are mutually exclusive, then $A \cap B = \varnothing$.
Therefore, $P(A \cap B) = 0$. This is important when we consider mutually exclusive (or disjoint) events.

Probability of the union of two events

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

If A and B are mutually exclusive, then $P(A) + P(B)$.



Probability Properties

Directions: Use the information given below to work out your answer to the following questions.

Given $P(A)=0.6$, $P(B)=0.5$, and $P(A \cap B)=0.2$.

1. Find $P(A')$.
2. Find $P(A \cap B')$.
3. Find $P(B \cap A')$.
4. Find $P(A \cup B)$.

Independent Probabilities

Independent Events
Two events, A and B, are considered independent events if the probability of A occurring is not changed based on any knowledge of the outcome of B.

Dependent Events
Two events are not independent, or dependent if the knowledge of the outcome of B changes the probability of A.

Independent Probabilities

Conditional Probability
The probability of one event occurring given that it is known that a second event has occurred. This is communicated using the symbol '$|$' which is read as "given."

For example, $P(A|B)$ is read as "Probability of A given B."

Computing Conditional Probability

The Probability of A given B:

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}$$

The Probability of B given A:

$$P(B \mid A) = \frac{P(B \cap A)}{P(A)}$$

Independent Probabilities

	Female	Male	Total
Undergraduate	3814	3428	7242
Graduate	2213	2787	5000
Total	6027	6215	12242

The two-way table below displays the World Campus enrollment from Fall 2015 in terms of level (undergraduate and graduate) and biological sex.

Choose one student from the sample, what is the probability that the student is a female?

If it is known that the student is a graduate student, what is the probability that the student is a female?

Independent Probabilities

Unless one is explicitly told that events are independent, one cannot simply assume that they are. However, for some events we naturally assume independence (e.g. a flip of a fair coin or the roll of a fair die).

We can check for independence of two events by showing that any ONE of the following is true. For any given probabilities for events A and B, the events are independent if:

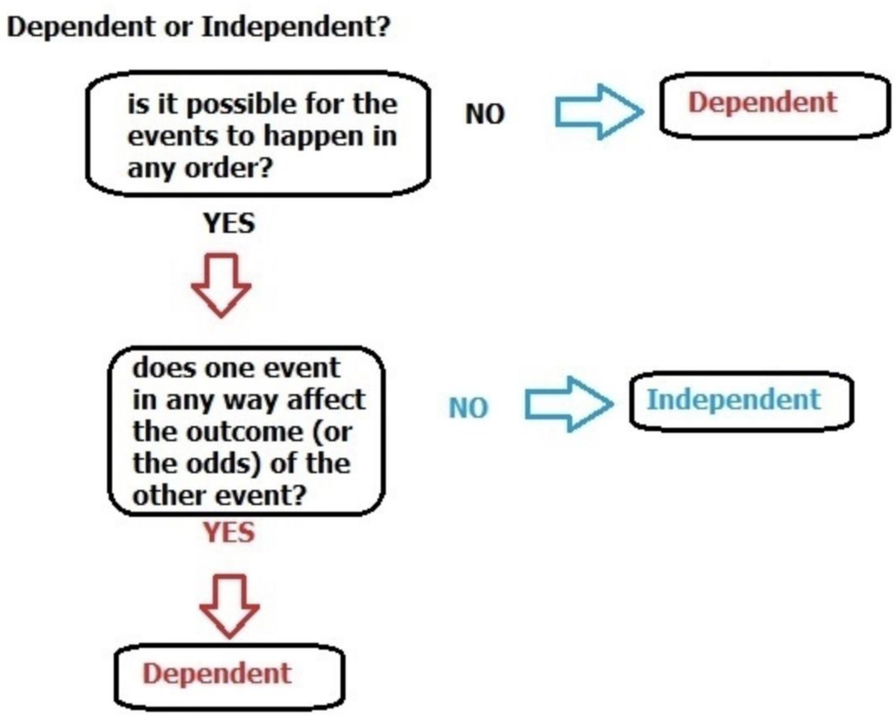
1. $P(A \cap B) = P(A) \cdot P(B)$
2. $P(A|B) = P(A)$

   or,

3. $P(B|A) = P(B)$

Independent Probabilities

- What are some examples of independent events?
- What are some examples of dependent events?



Independent Probabilities

You apply to grad schools at Harvard (H) and PSU (P). Probability accepted into Harvard is 0.3. Probability accepted into PSU is 0.6. The probability of being accepted to both is 0.25.

Noting the probabilities:

$P(H)=0.3$,

$P(P)=0.6$, and

$P(H \cap P)=.25$

Are events getting accepted into Harvard and into Penn State independent events?

Independent Probabilities

Independent vs. Mutually Exclusive

Students often confuse independent events and mutually exclusive events. The two terms mean very different things. Recall that if two events are mutually exclusive, they have no elements in common and thus cannot both happen at the same time. In fact, mutually exclusive events are dependent. If A and B are mutually exclusive events, then...

$$P(A \cap B) = 0 \neq P(A)P(B)$$

Independent Probabilities

Exxon will open a new gas station at a busy intersection in State College next year and feels that the probability that it will show a profit in its first year is 0.6 taking into consideration that Mobil may or may not open a gas station opposite to it. We know that if Mobil also opens a gas station opposite to it, the probability of a first year profit of the Exxon station will be 0.3 and the probability that Mobil will open the station opposite to it is 0.4.

1. What is the probability that both Exxon shows a profit in the first year and Mobil opens a gas station opposite to it?
2. What is the probability that either Exxon shows a profit in the first year or Mobil opens a gas station opposite to it?