Unit 5: Probability and Random Sampling

Case Study: UIUC Couse Data

- *We will consider our artificial UIUC course dataset to be a **population** and we will collect a **random sample** from this population both **with** and **without replacement.***
- *We will test whether the **expected proportion** of statistics classes in this sample is close to the **actual proportion** of statistics classes in this sample. Do we get closer as our sample size increases?*

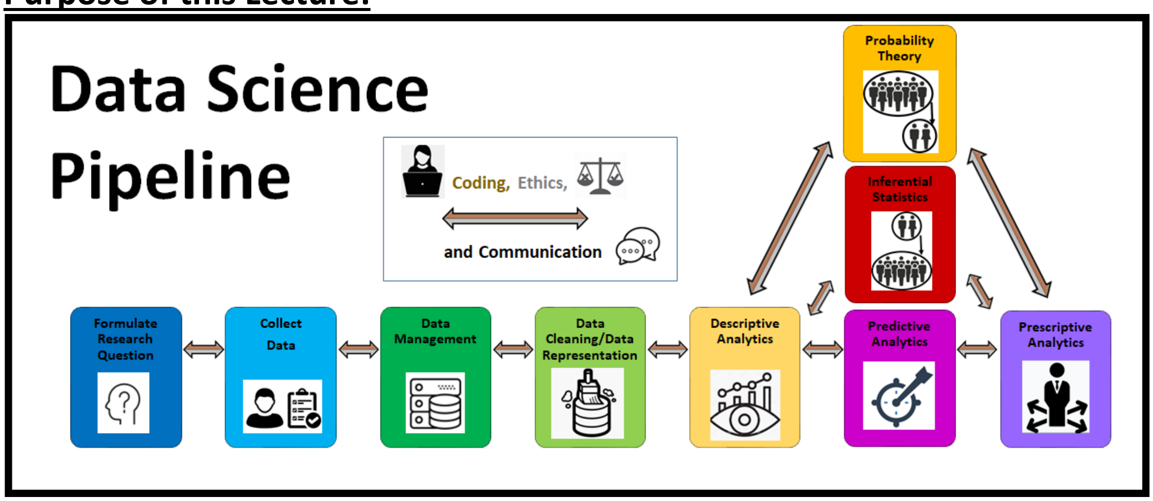
Case Study: Probability of Drawing a Full House

- *We will **calculate the probability** of drawing a full house from a deck of 52 playing cards.*
- *We will **approximate the probability** of drawing a full house from a deck of 52 playing cards by designing a **simulation**.*

Case Study: Coin Flips

- *We will **calculate the probability** of flipping exactly 6 tails after tossing a coin 10 times.*

Purpose of this Lecture:



In this lecture we will cover the following topics.

1. **How is probability used in the data science pipeline?**
   1.1. Relationship between inferential statistics and probability theory
2. **Experiment Related Definitions**
   2.1. Experiment
   2.2. Simple and Compound Events

Additional resources:

- Section 2.1 in Diez, Barr, and Cetinkaya-Rundel, (2015), *OpenIntro Statistics* https://www.openintro.org/download.php?file=os3&redirect=/stat/textbook/os3.php
- https://www.w3schools.com/python/python_for_loops.asp
- https://www.w3schools.com/python/python_operators.asp
- https://www.w3schools.com/python/ref_func_all.asp

---

1. How is probability used in the data science pipeline?

Definitions

Population: large body of data that is our target of interest

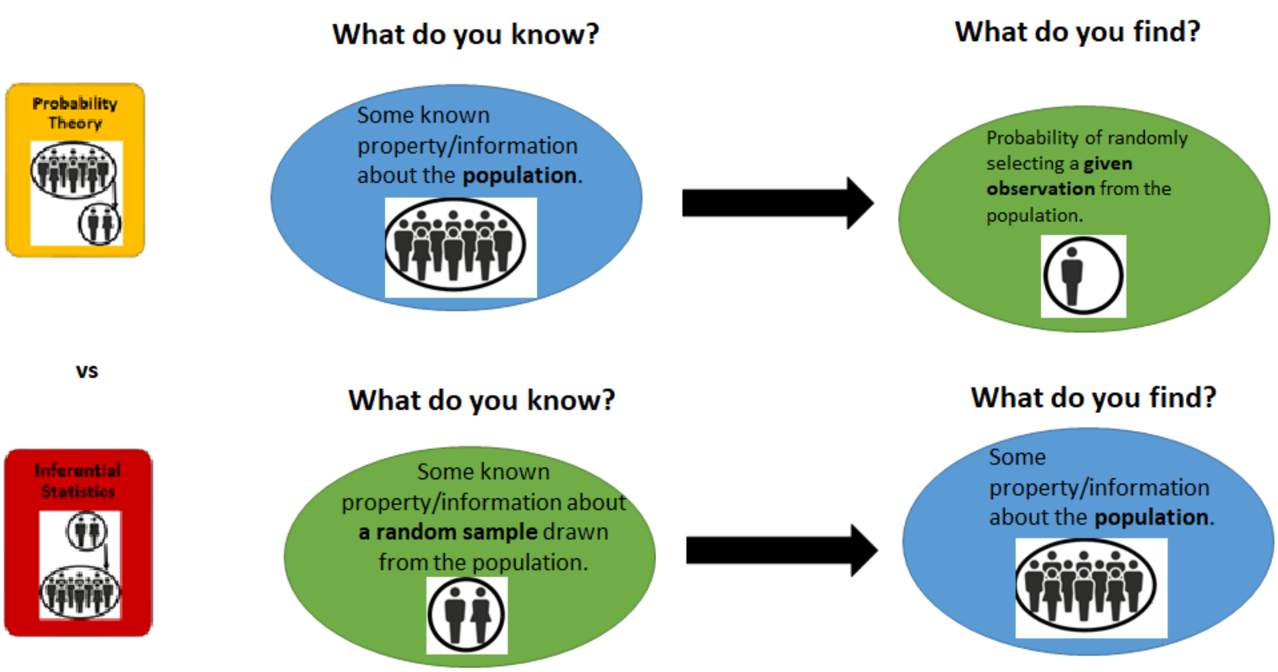
	course	section	enrolled
7	adv307	A	37
4	badm210	A	215
6	badm210	B	178
5	badm210	C	197
1	cs105	A	201
0	cs105	B	345
2	stat107	A	197
3	stat207	A	53

Sample: a subset of data selected from the population

	course	section	enrolled
7	adv307	A	37
4	badm210	A	215
6	badm210	B	178

Observation: a single "row" of data selected from the population (ie. sample of size n=1).

	course	section	enrolled
4	badm210	A	215



2. Experiment Related Definitions

An ___________________ is the process by which an **observation** is made. *(Most often we are interested in observations made from completely **uncontrollable situations.)***

Ex: Give an experiment you can perform with a single dice.

When an experiment is performed it can result in one or more outcomes, which are called ________________.

Ex: Give *some* examples of possible events that you could observe with this experiment.

A ______________ event can be decomposed into several events, while a ______________ event cannot.

The _____________________ (S) associated with an experiment is the set consisting of all possible simple events (ie. sample points).

Ex: Give the sample space associated with our single dice experiment.

3. Types of Random Sampling

<u>Type of Experiment</u>:

We can think of **collecting a random sample** (of n observations) from a population as conducting and observing which observations were in the sample as an experiment.

<u>Types of Random Sampling</u>:

- **Randomly Sampling** ___________________________ from a population means that observations can only be selected and put the sample _______________. Once we randomly select an observation, we do not "put the observation back" into the population.

  ****Go to Unit 5, section 3 Jupyter notebook for how collect a random sample (collected without replacement) of size n from a dataframe population.****

  Population

	course	section	enrolled
7	adv307	A	37
4	badm210	A	215
6	badm210	B	178
5	badm210	C	197
1	cs105	A	201
0	cs105	B	345
2	stat107	A	197
3	stat207	A	53

  Random Sample of Size n=5
  Drawn Without Replacement

	course	section	enrolled
1	cs105	A	201
4	badm210	A	215
2	stat107	A	197
3	stat207	A	53
5	badm210	C	197

- **Randomly Sampling** ____________________________ from a population means that observations CAN show up in the sample multiple times. Once we randomly select an observation, we "put it back into the population" where it *can* be randomly selected again.

  ****Go to Unit 5, section 3 Jupyter notebook for how collect a random sample (collected with replacement) of size n from a dataframe population.****

  Population

	course	section	enrolled
7	adv307	A	37
4	badm210	A	215
6	badm210	B	178
5	badm210	C	197
1	cs105	A	201
0	cs105	B	345
2	stat107	A	197
3	stat207	A	53

  Random Sample of Size n=5
  Drawn With Replacement

	course	section	enrolled
6	badm210	B	178
7	adv307	A	37
3	stat207	A	53
7	adv307	A	37
1	cs105	A	201

Ex: Suppose we conducted two experiments.

- Experiment 1: First we **randomly selected a sample of size n=2** from our population of 7 courses in the dataframe **with replacement**.
- Experiment 2: First we **randomly selected a sample of size n=2** from our population of 7 courses in the dataframe **without replacement**.

Which of the two **sample spaces** below corresponds to the random sampling experiment conducted a.) with replacement and b.) without replacement?

Population

	course	section	enrolled
0	cs105	B	345
1	cs105	A	201
2	stat107	A	197
3	stat207	A	53
4	badm210	A	215
5	badm210	C	197
6	badm210	B	178
7	adv307	A	37

Sample Space of Course Indices Selected in the **Random Sample** (of size n=2) that was Collected ___________ **Replacement**

{ 0, 0}	{ 2, 0}	{ 4, 0}	{ 6, 0}
{ 0, 1}	{ 2, 1}	{ 4, 1}	{ 6, 1}
{ 0, 2}	{ 2, 2}	{ 4, 2}	{ 6, 2}
{ 0, 3}	{ 2, 3}	{ 4, 3}	{ 6, 3}
{ 0, 4}	{ 2, 4}	{ 4, 4}	{ 6, 4}
{ 0, 5}	{ 2, 5}	{ 4, 5}	{ 6, 5}
{ 0, 6}	{ 2, 6}	{ 4, 6}	{ 6, 6}
{ 0, 7}	{ 2, 7}	{ 4, 7}	{ 6, 7}
{ 1, 0}	{ 3, 0}	{ 5, 0}	{ 7, 0}
{ 1, 1}	{ 3, 1}	{ 5, 1}	{ 7, 1}
{ 1, 2}	{ 3, 2}	{ 5, 2}	{ 7, 2}
{ 1, 3}	{ 3, 3}	{ 5, 3}	{ 7, 3}
{ 1, 4}	{ 3, 4}	{ 5, 4}	{ 7, 4}
{ 1, 5}	{ 3, 5}	{ 5, 5}	{ 7, 5}
{ 1, 6}	{ 3, 6}	{ 5, 6}	{ 7, 6}
{ 1, 7}	{ 3, 7}	{ 5, 7}	{ 7, 7}

Sample Space of Course Indices Selected in the **Random Sample** (of size n=2) that was Collected ___________ **Replacement**

	{ 2, 0}	{ 4, 0}	{ 6, 0}
{ 0, 1}	{ 2, 1}	{ 4, 1}	{ 6, 1}
{ 0, 2}		{ 4, 2}	{ 6, 2}
{ 0, 3}	{ 2, 3}	{ 4, 3}	{ 6, 3}
{ 0, 4}	{ 2, 4}		{ 6, 4}
{ 0, 5}	{ 2, 5}	{ 4, 5}	{ 6, 5}
{ 0, 6}	{ 2, 6}	{ 4, 6}	
{ 0, 7}	{ 2, 7}	{ 4, 7}	{ 6, 7}
{ 1, 0}	{ 3, 0}	{ 5, 0}	{ 7, 0}
	{ 3, 1}	{ 5, 1}	{ 7, 1}
{ 1, 2}	{ 3, 2}	{ 5, 2}	{ 7, 2}
{ 1, 3}		{ 5, 3}	{ 7, 3}
{ 1, 4}	{ 3, 4}	{ 5, 4}	{ 7, 4}
{ 1, 5}	{ 3, 5}		{ 7, 5}
{ 1, 6}	{ 3, 6}	{ 5, 6}	{ 7, 6}
{ 1, 7}	{ 3, 7}	{ 5, 7}	

4. Two Different Definitions of Probability

4.1 Bayesian Definition of Probability:

- **Probability of an event**: measure of one's belief in the occurrence of this event

	course	section	enrolled
0	cs105	B	345
1	cs105	A	201
2	stat107	A	197
3	stat207	A	53
4	badm210	A	215
5	badm210	C	197
6	badm210	B	178
7	adv307	A	37

4.2. Bayesian Definition of Probability:

- **Probability of an event**: the __________________ of times the outcome would occur if

  we observed the random process _____________________ of times.

Experiments with the Law of Large Numbers

Experiment 1: Let's say we collected a random sample of size **n=1000** *(with replacement).* ****Go to Unit 5, section 4.2 Jupyter notebook to conduct this experiment.****

a.) What proportion of classes in our random sample would you **expect** to be STAT207?

b.) What proportion of classes in our random sample were **actually** STAT207?

c.) What proportion of our sample would you **expect** to be a statistics class?

d.) What proportion of our sample were **actually** a statistics class?

Experiment 2: Let's say we collected a random sample of size **n=1,000,000** *(with replacement).* ****Go to Unit 5, section 4.2 Jupyter notebook to conduct this experiment.****

e.) What proportion of classes in our random sample would you **expect** to be STAT207?

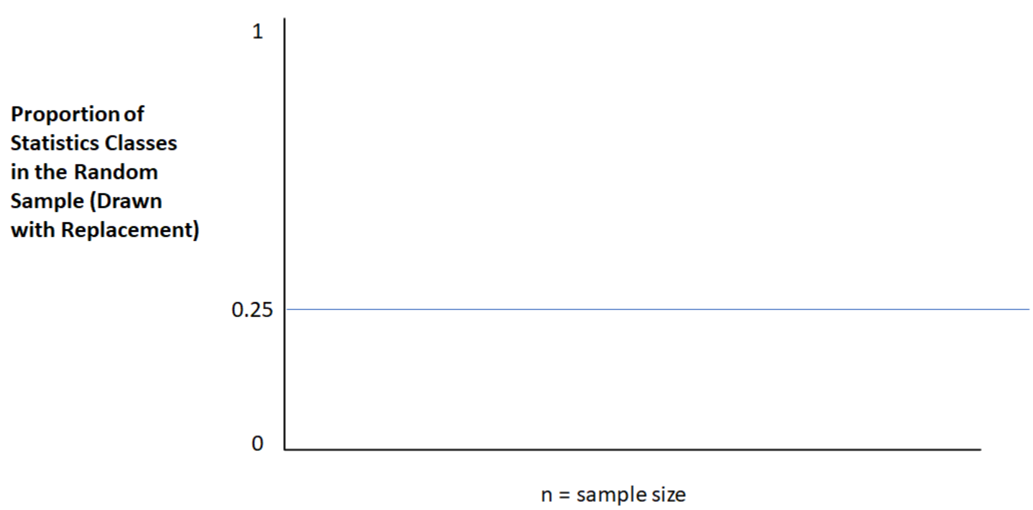
f.) What proportion of classes in our random sample were **actually** STAT207?

g.) What proportion of our sample would you **expect** to be a statistics class?

h.) What proportion of our sample were **actually** a statistics class?

4.3 Law of Large Numbers

As more observations are collected, the **proportion of actual occurrences** of a particular outcome converges to the **probability of that outcome**.



5. How to Calculate the Probability of Certain Types of Events

5.1. Event Type 1: How to Calculate the Probability of a Simple Event that Follows a Uniform Probability Model

Experiment 3: Let's say we collected a random sample of size **n=1,000,000** *(with replacement).* ****Go to Unit 5, section 4.2 Jupyter notebook to conduct this experiment.****

i.) What proportion of the random sample would we **expect** to be of any one type of class (ie. course and section combination)?

j.) What proportion of the random sample **actually** were of each class (ie. course and section combination)?

```
2    0.125526
5    0.125347
1    0.125267
7    0.125076
4    0.125046
3    0.124669
6    0.124576
0    0.124493
```

Definition:

In a **uniform probability model,** the probability of each simple event of an experiment is the same.

Probability Calculation Rule:

The probability of a simple event from that follows a uniform probability model is

_____________ (where n=# of simple events in sample space).

5.2. Event Type 2: How to Calculate the Probability of a Compound Event of Simple Events that Comes from a Uniform Probability Model

Probability Calculation Rule:

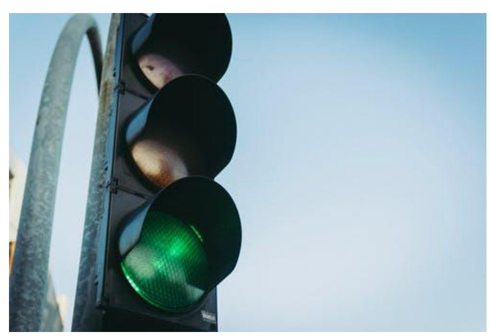
The probability of a compound event of simple events that follow a uniform probability model is _____________ (where n=# of simple events, k=# of simple events in which the compound event happens.).

Ex: What is the probability of randomly selecting a BADM class (course/section) from the population of courses shown to the right?

	course	section	enrolled
0	cs105	B	345
1	cs105	A	201
2	stat107	A	197
3	stat207	A	53
4	badm210	A	215
5	badm210	C	197
6	badm210	B	178
7	adv307	A	37

True or False
Suppose we flip a coin twice. The probability of getting at least one head is ¾.

True or False
Suppose we drive down a street with two traffic lights, where the probability of getting a green for each light is 0.6. The probability of getting at least one green on the road is ¾.



5.3. Event Type 3: How to Calculate the Probability of a Two Independent Events (or Two Observation Outcomes Sampling with Replacement)

Definitions

- Two events are **independent** of each other if the outcome of one event ______________ influence the outcome of the other event (and vice versa).

Mathematical Properties of Independent Events

If events A and B are __________________, then P(A and B) = __________________.

Properties of Random Sampling Two Observations from a Population With Replacement

- The outcomes of two observations randomly sampled from a population with replacement are _________________________.

Ex: What is the probability of randomly selecting two statistics courses from our dataset (population) *with replacement?*

	course	section	enrolled
0	cs105	B	345
1	cs105	A	201
2	stat107	A	197
3	stat207	A	53
4	badm210	A	215
5	badm210	C	197
6	badm210	B	178
7	adv307	A	37

{ 0, 0 }	{ 2, 0 }	{ 4, 0 }	{ 6, 0 }
{ 0, 1 }	{ 2, 1 }	{ 4, 1 }	{ 6, 1 }
{ 0, 2 }	{ 2, 2 }	{ 4, 2 }	{ 6, 2 }
{ 0, 3 }	{ 2, 3 }	{ 4, 3 }	{ 6, 3 }
{ 0, 4 }	{ 2, 4 }	{ 4, 4 }	{ 6, 4 }
{ 0, 5 }	{ 2, 5 }	{ 4, 5 }	{ 6, 5 }
{ 0, 6 }	{ 2, 6 }	{ 4, 6 }	{ 6, 6 }
{ 0, 7 }	{ 2, 7 }	{ 4, 7 }	{ 6, 7 }
{ 1, 0 }	{ 3, 0 }	{ 5, 0 }	{ 7, 0 }
{ 1, 1 }	{ 3, 1 }	{ 5, 1 }	{ 7, 1 }
{ 1, 2 }	{ 3, 2 }	{ 5, 2 }	{ 7, 2 }
{ 1, 3 }	{ 3, 3 }	{ 5, 3 }	{ 7, 3 }
{ 1, 4 }	{ 3, 4 }	{ 5, 4 }	{ 7, 4 }
{ 1, 5 }	{ 3, 5 }	{ 5, 5 }	{ 7, 5 }
{ 1, 6 }	{ 3, 6 }	{ 5, 6 }	{ 7, 6 }
{ 1, 7 }	{ 3, 7 }	{ 5, 7 }	{ 7, 7 }

Way 1: The sample space for this experiment is shown below.

- Each simple event in this sample space is equally likely, therefore the simple events here ____________________.
- k = number of simple events in sample space in which both are statistics courses = ___________.
- n = number of simple events in sample space = ___________.
- Probability = k/n = _________________.

	course	section	enrolled
0	cs105	B	345
1	cs105	A	201
2	stat107	A	197
3	stat207	A	53
4	badm210	A	215
5	badm210	C	197
6	badm210	B	178
7	adv307	A	37

Way 2: Properties of Independent Events

- Event 1: class 1 is a stats class
- Event 2: class 2 is a stats class
- Are they independent?

So thus we can use this mathematical property of independent events:

P(class 1 is a stats class AND class 2 is a stats class)

= P(class 1 is a stats class)* P(class 2 is a stats class)

=

Ex: Now Go to Unit 5, section 5.3 Jupyter notebook to conduct this experiment to approximate the probability of randomly selecting two statistics classes (when sampling with replacement).

Ex: Are the approximated probability and the actual probability close when using n=10000 trials?

5.4. Event Type 4: How to Calculate the Probability of a Two Dependent Events (or Two Observations Outcomes Sampling Without Replacement)

Definitions

- Two events are **dependent** of each other if the outcome of one event _____________ influence the outcome of the other event (and vice versa).

Mathematical Properties of Independent Events

If events A and B are __________________, then

- P(A and B) ≠ __________________
- P(A and B) = __________________
- P(A and B) = __________________

Properties of Random Sampling Two Observations from a SMALL Population Without Replacement

- The outcomes of two observations randomly sampled from a population with replacement are _________________________.

Ex: What is the probability of randomly selecting two statistics courses from our dataset (population) *without replacement?*

	course	section	enrolled
0	cs105	B	345
1	cs105	A	201
2	stat107	A	197
3	stat207	A	53
4	badm210	A	215
5	badm210	C	197
6	badm210	B	178
7	adv307	A	37

Way 1: Properties of Dependent Events

- Event 1: class 1 is a stats class
- Event 2: class 2 is a stats class
- Are they independent?

But we can use another mathematical property of dependent events:

P(class 1 is a stats class AND class 2 is a stats class)

= P(class 1 is a stats class)* P(class 2 is a stats class|class 1 is a stats class)

=

Sample Space (where Order Matters)

	{ 2, 0 }	{ 4, 0 }	{ 6, 0 }
{ 0, 1 }	{ 2, 1 }	{ 4, 1 }	{ 6, 1 }
{ 0, 2 }		{ 4, 2 }	{ 6, 2 }
{ 0, 3 }	{ 2, 3 }	{ 4, 3 }	{ 6, 3 }
{ 0, 4 }	{ 2, 4 }		{ 6, 4 }
{ 0, 5 }	{ 2, 5 }	{ 4, 5 }	{ 6, 5 }
{ 0, 6 }	{ 2, 6 }	{ 4, 6 }	
{ 0, 7 }	{ 2, 7 }	{ 4, 7 }	{ 6, 7 }
{ 1, 0 }	{ 3, 0 }	{ 5, 0 }	{ 7, 0 }
	{ 3, 1 }	{ 5, 1 }	{ 7, 1 }
{ 1, 2 }	{ 3, 2 }	{ 5, 2 }	{ 7, 2 }
{ 1, 3 }		{ 5, 3 }	{ 7, 3 }
{ 1, 4 }	{ 3, 4 }	{ 5, 4 }	{ 7, 4 }
{ 1, 5 }	{ 3, 5 }		{ 7, 5 }
{ 1, 6 }	{ 3, 6 }	{ 5, 6 }	{ 7, 6 }
{ 1, 7 }	{ 3, 7 }	{ 5, 7 }	

<u>Way 2</u>: If we assume that the *order in which we selected the two courses matters*, we can create the sample space of simple events for this experiment to the left.

- Each simple event in this sample space is equally likely, therefore the simple events here ___________________.
- k = number of simple events in sample space in which both are statistics courses = ___________.
- n = number of simple events in sample space = ___________.
- Probability = k/n = _________________.

Sample Space (where Order Doesn't Matter)

{ 0, 1 }			
{ 0, 2 }			
{ 0, 3 }	{ 2, 3 }		
{ 0, 4 }	{ 2, 4 }		
{ 0, 5 }	{ 2, 5 }	{ 4, 5 }	
{ 0, 6 }	{ 2, 6 }	{ 4, 6 }	
{ 0, 7 }	{ 2, 7 }	{ 4, 7 }	{ 6, 7 }
{ 1, 2 }			
{ 1, 3 }			
{ 1, 4 }	{ 3, 4 }		
{ 1, 5 }	{ 3, 5 }		
{ 1, 6 }	{ 3, 6 }	{ 5, 6 }	
{ 1, 7 }	{ 3, 7 }	{ 5, 7 }	

<u>Way 3</u>: If we assume that the *order in which we selected the two courses DOESN'T MATTER*, we can create the sample space of simple events for this experiment to the left.

- Each simple event in this sample space is equally likely, therefore the simple events here ___________________.
- k = number of simple events in sample space in which both are statistics courses = ___________.
- n = number of simple events in sample space = ___________.
- Probability = k/n = _________________.

Ex: Now Go to Unit 5, section 5.4 Jupyter notebook to conduct this experiment to approximate the probability of randomly selecting two statistics classes (when sampling with replacement).

Ex: Are the approximated probability and the actual probability close when using n=10000 trials?

6. Using Combinatorics Equations to Help Us Quickly Count Events

6.1 Event Counting Scenario #1:

n^k = **number of ways to select a set of k items WITH replacement from a population of size n and order/pattern/sequence matters.**

Ex: How many possible samples of classes (ie. course/section) of size k=2 can we create by randomly sampling with replacement from the population of classes (which has n=8 classes) where the order in which we collect the classes in the sample matters.

Population

	course	section	enrolled
0	cs105	B	345
1	cs105	A	201
2	stat107	A	197
3	stat207	A	53
4	badm210	A	215
5	badm210	C	197
6	badm210	B	178
7	adv307	A	37

Sample Space

{ 0, 0 }	{ 2, 0 }	{ 4, 0 }	{ 6, 0 }
{ 0, 1 }	{ 2, 1 }	{ 4, 1 }	{ 6, 1 }
{ 0, 2 }	{ 2, 2 }	{ 4, 2 }	{ 6, 2 }
{ 0, 3 }	{ 2, 3 }	{ 4, 3 }	{ 6, 3 }
{ 0, 4 }	{ 2, 4 }	{ 4, 4 }	{ 6, 4 }
{ 0, 5 }	{ 2, 5 }	{ 4, 5 }	{ 6, 5 }
{ 0, 6 }	{ 2, 6 }	{ 4, 6 }	{ 6, 6 }
{ 0, 7 }	{ 2, 7 }	{ 4, 7 }	{ 6, 7 }
{ 1, 0 }	{ 3, 0 }	{ 5, 0 }	{ 7, 0 }
{ 1, 1 }	{ 3, 1 }	{ 5, 1 }	{ 7, 1 }
{ 1, 2 }	{ 3, 2 }	{ 5, 2 }	{ 7, 2 }
{ 1, 3 }	{ 3, 3 }	{ 5, 3 }	{ 7, 3 }
{ 1, 4 }	{ 3, 4 }	{ 5, 4 }	{ 7, 4 }
{ 1, 5 }	{ 3, 5 }	{ 5, 5 }	{ 7, 5 }
{ 1, 6 }	{ 3, 6 }	{ 5, 6 }	{ 7, 6 }
{ 1, 7 }	{ 3, 7 }	{ 5, 7 }	{ 7, 7 }

6.2 Event Counting Scenario #2: PERMUTATIONS

$\frac{n!}{(n-k)!}$ = **number of ways to select a set of k items <u>WITHOUT replacement</u> from a population of size n and <u>order/pattern/sequence matters</u>.**

$n! =$

$5! =$

…

$0! =$

<u>**Ex:**</u> How many possible samples of classes (ie. course/section) of size k=2 can we create by randomly sampling <u>WITHOUT replacement</u> from the population of classes (which has n=8 classes) where the order in which we collect the classes in the sample matters.

Population

	course	section	enrolled
0	cs105	B	345
1	cs105	A	201
2	stat107	A	197
3	stat207	A	53
4	badm210	A	215
5	badm210	C	197
6	badm210	B	178
7	adv307	A	37

Sample Space

	{ 2, 0 }	{ 4, 0 }	{ 6, 0 }
{ 0, 1 }	{ 2, 1 }	{ 4, 1 }	{ 6, 1 }
{ 0, 2 }		{ 4, 2 }	{ 6, 2 }
{ 0, 3 }	{ 2, 3 }	{ 4, 3 }	{ 6, 3 }
{ 0, 4 }	{ 2, 4 }		{ 6, 4 }
{ 0, 5 }	{ 2, 5 }	{ 4, 5 }	{ 6, 5 }
{ 0, 6 }	{ 2, 6 }	{ 4, 6 }	
{ 0, 7 }	{ 2, 7 }	{ 4, 7 }	{ 6, 7 }
{ 1, 0 }	{ 3, 0 }	{ 5, 0 }	{ 7, 0 }
	{ 3, 1 }	{ 5, 1 }	{ 7, 1 }
{ 1, 2 }	{ 3, 2 }	{ 5, 2 }	{ 7, 2 }
{ 1, 3 }		{ 5, 3 }	{ 7, 3 }
{ 1, 4 }	{ 3, 4 }	{ 5, 4 }	{ 7, 4 }
{ 1, 5 }	{ 3, 5 }		{ 7, 5 }
{ 1, 6 }	{ 3, 6 }	{ 5, 6 }	{ 7, 6 }
{ 1, 7 }	{ 3, 7 }	{ 5, 7 }	

6.3 Event Counting Scenario #2: COMBINATIONS

$\binom{n}{k} = \frac{n!}{k!(n-k)!} =$ **number of ways to select a set of k items <u>WITHOUT replacement</u> from a population of size n and <u>order/pattern/sequence DOESN'T matter</u>.**

<u>Intuition:</u> $\frac{n!}{k!(n-k)!}$ is similar $\frac{n!}{(n-k)!}$, why do we need to divide by $k!$ when <u>order/pattern/sequence DOESN'T matter?</u>

- How many ways are there to order (ie. how many ways to permute) a set of k objects?

- The number of combinations= $\frac{total\ number\ of\ permutations}{total\ number\ of\ ways\ to\ permute\ a\ set\ of\ k\ objects} = \frac{\frac{n!}{(n-k)!}}{k!} = \frac{n!}{k!(n-k)!} = \binom{n}{k}$

<u>Ex:</u> How many possible samples of classes (ie. course/section) of size k=2 can we create by randomly sampling <u>WITHOUT replacement</u> from the population of classes (which has n=8 classes) where the order in which we collect the classes in the sample doesn't matter.

Population

	course	section	enrolled
0	cs105	B	345
1	cs105	A	201
2	stat107	A	197
3	stat207	A	53
4	badm210	A	215
5	badm210	C	197
6	badm210	B	178
7	adv307	A	37

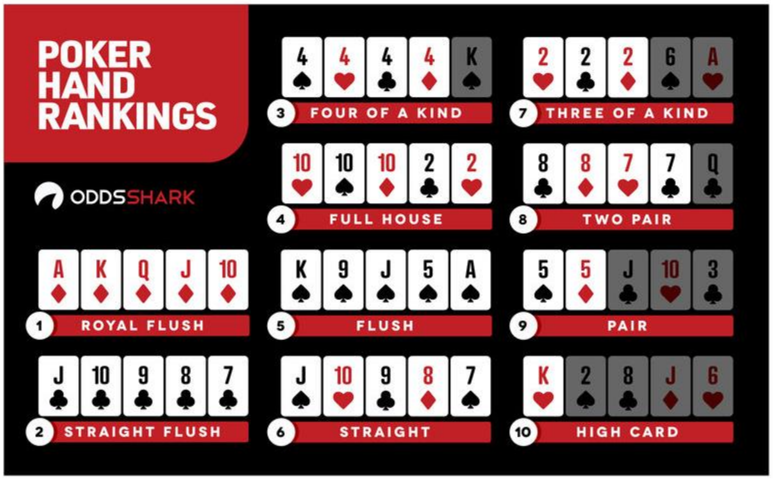
Sample Space

{ 0, 1 }			
{ 0, 2 }			
{ 0, 3 }	{ 2, 3 }		
{ 0, 4 }	{ 2, 4 }		
{ 0, 5 }	{ 2, 5 }	{ 4, 5 }	
{ 0, 6 }	{ 2, 6 }	{ 4, 6 }	
{ 0, 7 }	{ 2, 7 }	{ 4, 7 }	{ 6, 7 }
{ 1, 2 }			
{ 1, 3 }			
{ 1, 4 }	{ 3, 4 }		
{ 1, 5 }	{ 3, 5 }		
{ 1, 6 }	{ 3, 6 }	{ 5, 6 }	
{ 1, 7 }	{ 3, 7 }	{ 5, 7 }	

7. Using the Multiplication Rule to also Help Us Quickly Count Events

Multiplication rule: If an event is made up of a series of choices with k_1 ways to make the first choice, k_2 ways to make the second choice and so on, then the **total number of combinations of choices** is the product of the numbers of ways to make each of the individual choices (i.e., $k_1 \cdot k_2 \cdot k_3 \cdot \ldots$

8. <u>Case Study</u>: *Calculate* the probability of randomly drawing a full house from a standard 52 card deck of playing cards.



A **full house** is a sample of 5 cards in which:

- There are exactly **three** of one kind of **face** and
- there are exactly **two** of another kind of **face.**

(Image source: https://www.oddsshark.com/)

<u>**Describing the "Experiment" of Drawing 5 Cards:**</u>

1. **What is the "experiment" in this problem?**

2. **Give two examples of "simple events" that would be in the sample space of this experiment.**

3. **Does the "order" or "way" in which you randomly drew k=5 cards from the deck of n=52 cards matter when playing poker?**

4. **Are we randomly sampling cards with or without replacement from the deck?**

5. **Is the probability of any simple event in the sample space equally likely.**

6. Because the events in the sample space follow a **uniform probability model**, the probability of the compound event of simple events (ie. ______________________________) =
   - **DENOMINATOR = # of simple events** =

     ____

   - **NUMERATOR = # of simple events in which the compound event happens** =

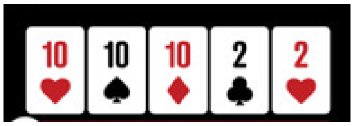
     ____

Breaking the problem down:

1. **DENOMINATOR = # of simple events** = number of combinations of k=5 card samples drawn from a deck without replacement of n=52 playing cards (where order of the draw doesn't matter) = __________________.

2. **NUMERATOR = # of simple events in which the full house happens**

=_______________

Ex: 

- **<u>Choice a</u>**: Which two distinct "face" cards will you choose?
  - "Face 1": designated for the three-suit card type
  - "Face 2": designated for the two-suit card type
- **<u>Choice b</u>**: Which three suits for "face 1" will you choose?
- **<u>Choice c</u>**: Which three suits for "face 1" will you choose?

<u>Using the Multiplication Rule</u>:

<u>How many options for choice a</u>?

- How many possible "ways" can we get k=two distinct face cards type out of a deck of n=13 types of faces where the "designation" for "face 1" card type and "face 2" card type matters.)

Possibilities for the Two Types of Faces in the Full House	
"Face 1" (for the three suits)	"Face 2" (for the two suits)
K	Q
K	J
K	10
K	9
K	8
K	7
K	6
K	5
K	4
K	3
K	2
K	A
Q	K
Q	J
Q	10
Q	9
Q	8
Q	7
Q	6
Q	5
Q	4
Q	3
Q	2
Q	A
...	...
...	...
A	K
A	Q
A	J
A	10
A	9
A	8
A	7
A	6
A	5
A	4
A	3
A	2

How many options for choice b?

- How many possible ways can we get k=3 distinct suits of the "face 1" card out of n=4 types of suits for "face 1"? (*Note that the "way"/"order" in which we select these three cards doesn't matter.)*

How many options for choice c?

- How many possible ways can we get k=2 distinct suits of the "face 2" card out of n=4 types of suits for "face 1"? (*Note that the "way"/"order" in which we select these three cards doesn't matter.)*

3. Putting it all together

Probability or randomly drawing a full house $=\frac{NUMERATOR}{DENOMINATOR} =$

9. <u>Case Study</u>: *Approximate* the probability of randomly drawing a full house from a standard 52 card deck of playing cards *with a simulation.*

<u>Ex</u>: Now Go to Unit 5, section 9 Jupyter notebook to *approximate the probability* of randomly selecting a full house from a standard 52 card playing deck using a simulation with n=50000 trials.

<u>Ex</u>: Are the approximated probability and the actual probability close when using n=50000 trials?

10. <u>Case Study</u>: We toss a coin 10 times. One possible sequence of heads and tails with 6 tails is THTTTHTTHH. How many possible sequences of heads and tails are there with exactly 6 tails?

<u>Describing the "Experiment" of Drawing 5 Cards:</u>

1. **What is the "experiment" in this problem?**

2. **Give two examples of "simple events" that would be in the sample space of this experiment.**
   HHHHHHHHHH

   TTTTTHTHHH

3. **Does the "way" in which k=6 positions from the n=10 possible coin flip positions in the sequence were selected to be a tails matter?**

4. **Is the probability of any simple event in the sample space equally likely?**

5. Because the events in the sample space follow a **uniform probability model**, the probability of the <u>compound event of simple events</u>
   <u>(ie.   )</u> =
   - **DENOMINATOR = # of simple events** =

     ____

   - **NUMERATOR = # of simple events in which the compound event happens** =

     ____

Breaking the problem down:

DENOMINATOR = # of ways to get a sequence of heads and tails of length 10 =

=n^k = **number of ways to select a set of *k=10* items <u>WITH replacement</u> from a population of size n=2 (Head/Tail) and <u>order/pattern/sequence matters</u>.**

NUMERATOR = # of ways to get a sequence of heads and tails of length 10, where 6 positions in the sequence are tails

$\binom{n}{k} = \frac{n!}{k!(n-k)!}$ = **number of ways to select a set of *k=6* positions <u>WITHOUT replacement</u> (to be the tails) from a population of size n=10 positions and <u>the way in which we selected these k=6 *positions* doesn't matter</u>.**

<u>Putting it all together</u>

Probability or randomly drawing a full house =$\frac{NUMERATOR}{DENOMINATOR}$ =

11. <u>Python Coding</u>:

*** for loops**

*** += operator**

*** if/then statements**

*** all() function**

See sections 5.3, 5.4 and 9 in the Unit 5 Jupyter notebook for information.