What do I know already? (Do not look up the answers to these questions. The purpose is to assess your current level of knowledge on these topics.)

A. In the context of probability, what does it mean to say that events A and B are dependent events? It means that the probability that B occurs is affected by the outcome of A or that the probability that A occurs is affected by the outcome of B.

B. A student randomly selects three marbles from a bag of marbles that contains 5 red marbles, 5 blue marbles, and 5 white marbles. Assume we are interested in the event that the student selects at least one white marble (one or more white marbles). What is the complement of that event? The event that ***none*** of the selected marbles is white.

C. If a bag of marbles has 5 red marbles, 5 blue marbles, and 5 white marbles, what is the probability that two marbles taken from the bag without replacement are both red?

P(2 red) = P(red) * P(red | first marble taken was red) = $\frac{\#\ red}{total} \cdot \frac{\#\ red\ left}{total\ left}$

= 5/15 * 4/14 = 0.095

D. Explain in your own words what information a discrete probability distribution provides: A discrete probability distribution provides the discrete list of possible outcomes for an experiment and the corresponding probabilities for each of those outcomes.

Learning Objectives: (Click the learning objectives below for a short clip on the topic.)

Know the Conditional Rule of Probability (1)
Use the Conditional Rule of Probability Without a Contingency Table (3)
Use the Conditional Rule of Probability With a Contingency Table (3)
Know the Multiplication Rule for Independent Events (1)
Use the Multiplication Rule for Independent Events (3)
Know the Multiplication Rule for Dependent Events (1)
Use the Multiplication Rule for Dependent Events (3)
Know the Approach to Calculate the Probability of At Least One of Something (1)
Calculate the Probability of At Least One of Something (3)
Contrast Discrete Random Variables and Continuous Random Variables (1)
Define a Probability Distribution of a Discrete Random Variable (1)
Describe the Characteristics of a Probability Distribution (2)

Exercises:

1. Use the data below from a study of mammograms conducted on women in their forties to estimate the probability that a randomly selected woman, in her forties, has cancer given that she has a positive exam result.

Cancer?	Positive Mammogram	Negative Mammogram	Totals
Cancer	32,608	10,870	43,478
No Cancer	267,298	2,689,224	2,956,522
Totals	299,906	2,700,094	3,000,000

$$P(cancer \mid positive) = \frac{\#\ of\ patients\ with\ Cancer \text{ in the } positive\ column}{total\ \#\ of\ positive\ results} = \frac{32,608}{299,906}$$

2. A survey of people from China, India, and Australia was conducted to determine the view of Japan in the Asia/Pacific region. Use the table below to estimate the probability that a randomly selected Indian holds a positive view of the Japanese.

	View of Japan			
Held by:	Negative	Neutral	Positive	Totals
Chinese	191	12	33	**236**
Indians	46	72	92	**210**
Australians	19	21	152	**192**
Totals	**256**	**105**	**277**	**638**

# of positive over total for the Indian row: 92 / 210 = 0.438

3. A computer programmer has written a simple program to generate passwords. The program creates password having the following structure: English Letter (capital or lowercase), English Letter (capital or lowercase), digit (0 - 9), digit (0 - 9), digit (0 - 9), One of the following symbols(!, #, $, %, or *). The letters cannot be repeated, but the digits can be repeated. How many unique passwords can this program randomly generate?
How many indicates a counting rule is needed. The fundamental counting rule can solve this problem.
Initially, there are 26 lowercase letters and 26 uppercase letters, so that is 52 total options for the first letter. The second letter cannot be the same as the first, so there are 51 options. There are ten digits available for each of the three digits, and there are five symbols to choose among.

52*51*10*10*10*5 = 13,260,000

4. Which of the following are discrete random variables? (select all that apply)
A. X is a variable representing the amount of time it takes a randomly selected student to complete this lab assignment.
B. Y is a variable representing the amount of typos found in a randomly selected page of a novel's first draft.
C. Z is a variable representing the number of people living in a randomly selected household in the U.S..
D. W is a variable that tracks the distance a randomly selected worker commutes to his/her job.
E. None of the above

5. The local government in a rural community tests water from wells in an effort to monitor contamination. To save time and money, the government mixes water from five wells and tests the mixture. If the mix tests positive for contamination, the local government has to test individual samples from each of the five wells to determine which wells are contaminated. If there is a 2.5% chance that a well has contaminated water, what is the probability that a randomly selected mix of five samples will test positive for contamination? (assume that the wells are independent of each other)

In order for the mix to test positive one or more of the five samples must be positive. Another way to say this is to say *at least one* of the five tests positive.

Note, the chance an individual sample is not contaminated is 1 - 0.025 = 0.975, and the chance all five are not contaminated is 0.975*0.975*0.975*0.975*0.975 because the wells are independent of each other.

$$P(\text{at least one is contaminated}) = 1 - P(\text{none of the five are contaminated}) = 1 - (0.975)^5$$

6. Use the data below from a study of mammograms conducted on women in their forties to estimate the probability that a randomly selected woman, in her forties, undergoing a mammogram, receives a positive exam result or has cancer.

Cancer?	Positive Mammogram	Negative Mammogram	Totals
Cancer	32,608	10,870	43,478
No Cancer	267,298	2,689,224	2,956,522
Totals	299,906	2,700,094	3,000,000

$$P(positive\ or\ cancer) = P(positive) + P(cancer) - P(positive \cap cancer)$$
$$\frac{\text{\# of positive}}{total} + \frac{\text{\# with cancer}}{total} - \frac{\text{\# with cancer that test positive}}{total} =$$
$$\frac{299,906}{3,000,000} + \frac{43,478}{3,000,000} - \frac{32,608}{3,000,000} \approx 0.104$$

7. Which of the following values could be the probability that an event occurs? (select all that apply)
A. 0.347 B. 0.741 C. 3.140 D. 1.07 E. -0.203 F. 1/3 G. 9/5

8. Does the following table meet the requirements of a probability distribution? If not, state why it is not a probability distribution. No, the probabilities do not sum to one.

X	P(X)
0	0.04
1	0.17
2	0.25
3	0.12
4	0.03

9. Daniel is rushing to pack for a trip. He reaches into his dryer at home to select two white undershirts. The dryer contains 12 of these white undershirts along with other items. Daniel doesn't know that his sister has left a pink lipstick in the pocket of her jeans, which has melted and stained three of Daniel's white undershirts. What is the probability that Daniel selects two white undershirts with lipstick stains when randomly selecting two shirts from the dryer without replacement?

$$P(2stained) = \frac{\#\text{ stained}}{total} \cdot \frac{\#\text{ stained } left}{total\ left} = \frac{3}{12} \cdot \frac{2}{11} \approx 0.045$$

10. Use the data below from a study of mammograms conducted on women in their forties to estimate the probability that a randomly selected forty-something woman undergoing a mammogram does not have cancer given that she has a negative exam result.

Cancer?	Positive Mammogram	Negative Mammogram	Totals
Cancer	32,608	10,870	43,478
No Cancer	267,298	2,689,224	2,956,522
Totals	299,906	2,700,094	3,000,000

$$P(no\ cancer \mid negative) = \frac{\#\ of\ \ patients\ without\ \ Cancer\text{ in the negative } column}{total\ \ \#\ of\ \ \text{negative } results}$$

$$= \frac{2,689,224}{2,700,094} \approx 0.996$$

11. A six-sided die has been loaded (corrupted) so that it does not land on each of its six sides equally. One turns up on the die 25% of the tosses; two turns up 20% of the tosses; three turns up 18% of the tosses, four turns up 16% of the tosses, and five turns up on the die 12% of the tosses. What percent of the time does the die turn up showing a six? 100% - 25% - 20% - 18% - 16% - 12% = **9%**

12. The following table is a probability distribution for X, where X represents the number of people who have visited Europe out of a random sample of six U.S. citizens. What is the probability that at least two randomly selected U.S. citizens out of six have visited Europe (use the probability distribution to answer)?
at least 2 means 2 or more... just add 0.0550 + 0.0055 + 0.0003 = **0.0608**

X	P(X)
0	0.6470
1	0.2922
2	0.0550
3	0.0055
4	0.0003
5	0.0000
6	0.0000

13. In the United States, 82% of people attending high school will graduate. If four people are randomly selected from the U.S. high school population, what is the probability that all four will graduate?

$$P(all\ \ 4\ \ graduate) = 0.82^4 \approx 0.452$$

14. Seventy-six percent of people who attend a high school in Florida graduate. If four students attending Florida high schools are randomly selected, what is the probability at least one of the four will graduate?

note: P(not graduating) = 1 - 0.76

$$P(\text{at least one graduates}) = 1 - P(\text{none of the four graduate}) = 1 - (0.24)^4$$

15. Does the following table meet the requirements of a probability distribution? If not, state why it is not a probability distribution. No, all probabilities must be between 0 and 1 inclusive.

X	P(X)
0	0.02
1	0.14
2	0.37
3	-0.12
4	0.41
5	0.18

16. A recent survey asked participants about their views on biomedical enhancements for humans. Almost three quarters of the participants (74%) believed brain enhancements were likely to be available in their lifetimes. Sixty-nine percent of those surveyed thought that brain chips designed to improve mental function would raise cause for concern. Fifty-five percent of those surveyed believed brain enhancements were likely to be available in their lifetimes and thought brain chips for improved mental function would raise cause for concern. Use the survey results to estimate the probability that a randomly selected person would find brain chips concerning given that they find the development of such enhancements likely to occur in their lifetime.

    P(brain chips concerning | likely to occur) = P(brain chips concerning *and* likely to occur) / P(likely to occur) = 0.55/0.74 ≈ 0.743

17. A store has a sale bin full of a popular shade of nail polish called "purple pride." If the bin contains 33 bottles of polish that are good and 3 bottles that are dried out, what is the probability that a person selects two of the bottles of polish for purchase at random and finds that both are dried out?

$$P(2\ dried\ \ out) = \frac{\#\ \text{dried}}{total} \cdot \frac{\#\ \text{dried}\ left}{total\ \ left} = \frac{3}{33+3} \cdot \frac{2}{33+2} \approx 0.005$$

18. True or False: The random variable X contains the number of bottles of Pepsi sold from the vending machines on campus each day. This random variable is an example of a continuous random variable. Since the number of bottles sold is a count, the variable is discrete. There cannot be a decimal number of bottles sold.