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. What is the maximum value possible for any probability (expressed as a decimal)? What is the minimum possible value? The maximum possible value is 1.00 (100%). The minimum possible value is 0 (0%).

B. If a coin is biased so that when tossed it will turn up heads 55% of the time, what does the *law of large numbers* tell us about the cumulative results of a very large number of coin tosses with this coin? Over the very long run (i.e. - after a very many number of tosses) the percentage of times the coin turns up heads will be very, very close to 55%. The more tosses that are made, the closer and closer this percentage of heads will get to 55%.

C. If an experiment has five possible outcomes, what must be true about the sum of probabilities for those five possible outcomes? The sum of probabilities must add up to 1.00 (100%) because all of the outcomes are accounted for (thus all of the probability is also accounted for).

D. In classical probability, the probability of an event is equal to the number of favorable outcomes divided by the total number of possible outcomes. What underlying assumption makes this approach to finding probability valid? The classical approach to probability assumes that all of the outcomes for an experiment are equally probable (equally likely).

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

Define Basic Probability (1)
Calculate Basic Probability (3)
Explain the Law of Large Numbers (2)
Discuss the Minimum and Maximum Likelihood of an Event (2)
Discuss the Sum of Probabilities for Collectively Exhaustive, Mutually Exclusive Events (2)
Use the Fundamental Counting Rule to Determine the Number of Possible Outcomes (3)
Define Combinations and the Factorial Operation (1)
Apply the Combinations Formula to Determine the Number of Possible Outcomes (3)
Contrast the Situations where It Is Appropriate to Use Combinations and Permutations (2)
Use the Addition Rule of Probability with and without a Contingency Table (3)
Use the Addition Rule of Probability for Mutually Exclusive Events (3)

Exercises:

1. There are three jobs to be filled in the Statistics lab: weekend supervisor, weekday supervisor, and weeknight supervisor. There are 20 applicants, and all of them are qualified to fill any of the three positions. Determine the number of ways the three positions can be filled with this set of applicants. 20 * 19 * 18 = 6840

2. An efficiency expert tracked the amount of time it took her to commute to work for an entire year. Her average commute time was 48 minutes. The standard deviation for the set of her commute times is 3.9 minutes. Would it be unusual for her commute to take more than 60 minutes? Why or why not? Yes, it would be unusual because a commute time of 60 minutes has a z score of 3.08. In other words, a 60 minute commute is more than three standard deviations above average.

3. Sports journalists analyzed a random sample of 177 penalty kicks taken in professional soccer. One hundred twenty-six of those kicks resulted in a goal scored. The other attempts were a miss. Use the results of the analysis to estimate the probability that a randomly chosen penalty kick results in a miss. P(miss) = (177-126)/177 = 0.288

4. Arsenic occurs naturally in the environment. As a result, the level of arsenic in some sources of water is dangerously high. A scientist in Florida claims that the water coming from his pipes has a mean arsenic level of 10 parts per million. The city took 35 random samples of his water to test the scientist's claim. The results of the water samples were summarized and expressed as a z score by using the claimed value of 10 parts per million as the mean. The resulting z score was 0.12. Based on the z score, does the scientist's claim appear to be plausible?

Yes, it seems plausible because the z score is not unusual at all under the assumption that the mean is 10 ppm. In fact, the positive z score indicates that the city's sample of water actually exhibited a higher amount of arsenic than 10 ppm.

5. Mary is filling small bags with candy for Halloween. Every small bag will have two pieces of candy in it. The two pieces of candy will be chosen from a large bag of Swedish Fish, Snickers, Butterfingers, Kit Kats, Twix, Skittles, and Starbursts. How many different pairings of candy can Mary create for the small bags? Assume that Mary will not fill a bag with two pieces of the same candy. For example, no bag would contain two Snickers bars. 7 C 2 = 21 There are 21 different pairings of these candies.

6. Tiger, a domestic shorthair cat, weighs 11.2 pounds. Crash, a Maine coon cat, weighs 21.0 pounds. The average weight for domestic shorthair cats is 9.0 pounds with a standard deviation of 2.1 pounds. The average weight for Maine coons is 17.5 pounds with a standard deviation of 3.8 pounds. Which cat, Tiger or Crash, is relatively heavier? For Tiger z = 1.05, for Crash z = 0.92, so Tiger is heavier for his peer group.

7. In one class, every student taking the final exam is timed. The times are recorded in minutes. What unit of measurement would the standard deviation of these times have? What unit for the variance?
   A. minutes for both
   B. minutes squared and minutes
   C. minutes squared for both
   D. minutes and minutes squared
   E. √minutes and minutes

8. A large survey of line cooks was recently conducted. Hourly wages for those line cooks had an average of $14.10. The median wage was $11.50, and the most common hourly salary range was reported to be between $9.25 and $10.25. Is the distribution likely left skewed, right skewed, or symmetric?

9. A credit card company requires all card members to choose an alpha-numeric pin in order to access account information. The pin consists of one letter chosen from the English alphabet followed by four digits (0 - 9). The digits can be repeated. How many different alpha-numeric pins can be formed using this approach? 26*10*10*10*10 = 260,000

10. 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 false positive exam result. P(false positive) = 267,298/3,000,000

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

11. Calculate the standard deviation for a set of 56 values that have the following summary values:

$\sum_{i=1}^{n} x_i = 39.209$ and $\sum_{i=1}^{n} x_i^2 = 28.094$ s = 0.108

12. While trying to solve a probability question on the exam, Jorge came up with the following expression as his solution: 0.348 + 0.112 - 0.462. He soon realized that he committed an error somewhere in his work. How did Jorge know his answer must be incorrect? The minimum a probability value can be is 0. The calculation above gives a negative result.

13. In a left-skewed distribution, the median is 14.9. Which of the following values could be the mean for the distribution? (select all that apply)

A. 12.3 B. 13.0 C. 15.4 D. 18.1 E. 16.2 F. 13.4

14. A surgeon randomly selects three patients who underwent knee surgery to participate in a new type of physical therapy. The surgeon plans to record the length of time it takes each of the three patients to recover full mobility. The surgeon plans to use the range as the measure of dispersion (variation) for the recovery times. Is the range an acceptable option under these circumstances? Why or why not? Yes, because the data set is very small. The surgeon is only going to have three recovery times.

15. Fast food should be fast. The manager of a local McDonald's claims that the average wait time at her restaurant is only 2.1 minutes. To test the store manager's claim, the owner of the store tracks the wait times of 40 randomly chosen orders. The collected data were summarized and converted into a z score by using the claimed value of 2.1 minutes as the population mean. The resulting z score was 3.17. Based on this result, does the manager's claim seem credible? No, it does not seem credible. The positive z score indicates the sample mean wait time was longer than the stated value of 2.1 minutes. Also, the z score is unusually large, so the sample data indicates a large difference between the sample mean and the claimed value of 2.1 minutes.

16. 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 has breast cancer or has a false positive exam result. P(Cancer or False Pos.) = P(cancer) + P(false Pos.) - P(cancer and false positive) = 43,478/3,000,000 + 267,298/3,000,000 - 0/3,000,000 = 0.104 *note: having breast cancer and a false positive mammogram are mutually exclusive events.

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

17. If there is a 0.46 probability that the textbook(s) for a randomly selected class on campus will cost between $100 and $200 inclusive and a 0.11 probability that the textbook(s) will cost more than $200. What is the probability that the textbook(s) will cost less than $100 for that class? 1 - 0.46 - 0.11 = 0.43 This problem relies on the fact that these events are mutually exclusive and exhaustive. This means the sum of their probabilities must add to one.