

Roulette Worksheet

Math 5

11.4.97

The purpose of this worksheet is to solidify the concepts of expected value and standard error, and to introduce the normal curve as a tool for approximating probabilities that would be painful to compute otherwise (even with the help of a calculator). It also introduces some mathematical terms for concepts we have already discussed in class. You do not need to worry about these terms. They are mentioned here so that you will be able to consult other texts, and so that you will know what people are talking about when they use mathematical terms to describe things that you already know.

Suppose you perform an experiment such as spinning a roulette wheel five times in a row and each time betting on the same column. A roulette wheel has 38 pockets. One is numbered 0, another is numbered 00, and the rest are numbered from 1 through 36. The croupier spins the wheel, and throws a ball onto the wheel. The ball is equally likely to land in any one of the 38 pockets. Before it lands, bets can be placed on the table (see diagram).

Let X be some result of the experiment that you are interested in. In this case, say X is the number of times you win in five spins of the wheel. In other words, the variable X represents the number of times you win by spinning the roulette wheel five times and each time betting on the same column.

Other results you may be interested in include the amount of money you win in five spins, the number of times the ball lands in a red pocket in five spins, the number of times the ball lands in the pocket labeled 00 in five spins, and so on. For now we will concentrate on the variable X as described above. You may wish to choose a different definition for X later. A variable that gives the outcome of an experiment is called a **random variable** (the mathematical definition is more technical, but this is the main idea).

You may wonder why the experiment consists of five spins of the wheel rather than just one. I claim that this is a natural way to view roulette. When you go to play roulette do you bet just once and then walk away, or do you play several times in a row? Repeating this experiment then simulates you coming back to the roulette table several times during the course of an evening. Or you could view repeating this experiment from the perspective of the casino. In this case repeating the experiment could represent many different people playing roulette over the course of an evening.

In the experiment we have described what are the possible values, x_1, x_2, x_3, x_4, x_5 , and x_6 , for X ? What is the probability that each of these values occurs? Let p_i be the probability that the value x_i occurs. Formally $p_i = P(X = x_i)$. Compute p_1, p_2, p_3, p_4, p_5 , and p_6 . An assignment of probabilities p_i to the possible outcomes x_i of an experiment is called a **probability density function**.

What is the box model for this experiment? In other words, if you were to translate this experiment into the language of boxes and tickets, what would be the different values you would find on tickets? How many tickets of each kind would there be? Note that so far the book has been considering one spin of the wheel to be the experiment, and they consider repeating the experiment some number of times. Here the experiment is spinning the wheel five times and you are repeating it once (for now).

What is the expected value for X ? If you were to repeat this experiment many times and each time compute X , the number of times you won, then take the average number of times you won from all these times, you would probably get close to the expected value for X . We write “the expected value for X ” as $E[X]$. There are two ways to compute $E[X]$. One is to take the average value of the tickets in the box. This is the method your book suggests. The other is to use the following formula:

If X is a random variable with possible values x_1, x_2, \dots, x_n and probability density function p_1, p_2, \dots, p_n , then the expected value of X is $E[X] = x_1p_1 + x_2p_2 + \dots + x_np_n$.

This definition is the one you will find in most books on probability. However I think your book’s definition is the most intuitive.

Compute $E[X]$ according to the definition in your book.

Compute $E[X]$ according to the definition above.

If you were to perform this experiment with an actual roulette wheel, the value you would get for X would probably not be exactly $E[X]$. Why not?

If you were to perform this experiment you would probably get a value for X which would be somewhere near $E[X]$. It would be $E[X]$ plus or minus some amount of chance error. If you were to perform this experiment (spinning the roulette wheel five times) over and over again you would probably see that 68% of the time you would win $E[X]$ times plus or minus SE dollars, where SE is the standard error of the experiment. 95% of the time you would win $E[X]$ plus or minus two times SE times. 99% of the time you would win $E[X]$ plus or minus three times SE times. What is SE for this experiment? If you were to repeat this experiment ten times (for a total of fifty spins), what would be the standard error?

To convince yourself that these theoretical numbers give realistic answers, you could either fly to Las Vegas and try the wheel yourself, or you could pull up data desk and have it simulate things for you. To do this do **Manip** → **Generate Random Numbers...** and fill in: **Generate 1 variable with 10 cases**, **#Bernoulli trials/ experiment = 5**, and **Prob(success) = .3** (note that .3 is approximately $\frac{12}{38}$). Next do **Plot** → **Histograms**. Choose **Plot Scale...** from the hyperview menu and set the bar width to 1 (if it's not set to 1 already). This will perform the experiment 10 times (for a total of fifty spins) and record the number of times you win for each experiment. Do this again for 10 experiments. Try increasing the number of experiments to 30 (do **Generate 1 variable with 30 cases**), then 100, then 1000, and finally 10,000. What do you notice?

The **Central Limit Theorem** says that when drawing at random with replacement from a box (such as using the box model for this roulette experiment), the probability histogram for the sum of the tickets (the total number of times you win at roulette) will follow the normal curve. This is true when the histogram is converted to standard units and the number of draws from the box (the number of repetitions of the experiment) is relatively large.

Choose one of your histograms where the number of trials is 1000. The area of each bar in the histogram gives the percentage of the time that the value it is above appeared as X in doing the experiment 1000 times. In this case the percentage of the time you won either four or five times out of five spins of the wheel in 1000 sets of five is the sum of the area in the bar above four and the area in the bar above five.

When the number of possible values for X is small, such as in this example, the percentage of time a certain result or set of results occurs is easy to compute. Just add each individual area. But what if there are many possible values for X ?

Change the experiment so that one trial consists of 50 spins of the wheel, rather than five. Simulate 1000 trials of this experiment by doing **Manip** → **Generate Random Numbers...** and fill in: **Generate 1 variable with 1000 cases, #Bernoulli trials/ experiment = 50**, and **Prob(success) = .3**. Next do **Plot** → **Histograms**. Choose **Plot Scale...** from the hyperview menu and set the bar width to 1.

To compute the percentage of the time you won between 10 and 30 times out of 50 in 1000 repetitions of the experiment, you can use the fact that this histogram approximately follows the normal curve, rather than having to figure out the areas of the bars above each of the numbers between 10 and 30. The answer you get will not be the exact percentage, but since the number of trials is large, it should be pretty close, according to the Central Limit Theorem.

First observe that the bar above 10 starts at 9.5 on the number line and ends at 10.5. Similarly the bar about 30 starts at 29.5 and ends at 30.5 (note that if you can't see a bar above 30, that means it's height is zero). So what you want is the area of the histogram between 9.5 and 30.5. To approximate this with the normal curve we must first convert to standard units.

How far is 9.5 away from the mean (the expected value of X) in terms of standard deviations? How far is 30.5?

What is the area under the standard normal curve between these two values?

The number you computed above is the percentage of time you would expect to see between 10 and 30 wins out of 50 spins on the roulette when if you repeated each set of 50 spins 1000 times. How would you figure out the probability of getting between 0 and 10 wins out of 50 spins of the wheel? The probability that you get more than 20?