

Algebra 2/Pre-Calculus

Probability Histograms and Area (Day 9, Statistics)

Name _____

In the last handout, we saw that standard deviations are a way of measuring how likely events are. In this handout, we will expand on this idea and see how it is connected to area.

1. Suppose we have an unfair coin that lands heads 55% of the time and lands tails 45% of the time. We flip the coin 25 times.

- a. Find the mean, variance, and standard deviation for the number of heads.

- b. Which numbers (of heads) are within one standard deviation of the mean?

- c. To answer the last question, we add and subtract 2.4875 from the mean:
 $13.75 - 2.4875 = 11.2625$ and $13.75 + 2.4875 = 16.2374$, so the values within one standard deviation of the mean are 12, 13, 14, 15, and 16. Which values are within two standard deviations of the mean?

- d. Which values are at least two standard deviations above the mean?

- e. Which values are at least three standard deviations below the mean?

Some answers a. mean = $25(.55) = 13.75$, variance = $25(.55)(.45) = 6.1875$, standard deviation = $\sqrt{25(.55)(.45)} = 2.4875$ c. 9, 10, 11, 12, 13, 14, 15, 16, 17, 18

d. 19, 20, 21, 22, 23, 24, 25 e. 0, 1, 2, 3, 4, 5, 6

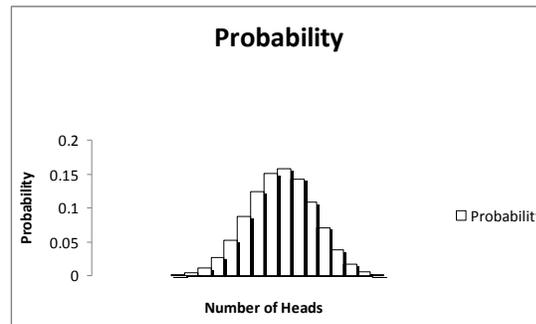
- e. Here's another way of approaching these problems: Calculate $(0.55h + 0.45t)^{25}$ using wolframalpha. Look at the expanded form. Can you find the probability of getting 14 heads? (Your answer should match the answer you got in part a.)
- f. Use the calculation from wolframalpha to find the probability of getting exactly 13 heads.
- g. Find the probability that the number of heads is within one standard deviation of the mean. (You will have to add up the probabilities of getting 12, 13, 14, 15, or 16 heads).
- h. You should have found that the probability was 0.6833. This means that 68.33% of the values were within one standard deviation of the mean. What percent of the data is within two standard deviations of the mean?

Answers a. $({}_{25}C_{14})(.55)^{14}(.45)^{11} = 0.1583$ b. 14 is closest to the mean (13.75) c. 0.1419 (a little lower than the probability of 14 heads) d. Probability is very close to zero (so a lot lower) e. Notice that one of the terms is $0.1583h^{14}t^{11}$. This means that the probability of getting 14 heads and 11 tails is 0.1583 f. 0.1511

g. 68.33% $(0.1236 + 0.1511 + 0.1583 + 0.1419 + 0.1084 = .6833)$

h. 95.68% $(.0266 + .0520 + .0867 + .1236 + .1511 + .1583 + .1419 + .1084 + .0701 + .0381 = .9568)$

3. Let's consider the histogram for this situation:



Each bar on the histogram represents a number of heads. The height of each bar is the probability of getting that number of heads. *Note:* Some of the bars are so short they don't even appear on the graph.

- Which number of heads has the highest bar?
- The bar for 14 heads is highest because 14 is the most likely outcome. (14 is the closest to the mean, 13.75). What is the height of this bar?
- The height of the bar is 0.1583 (because that is the probability of getting 14 heads). Shade in the bar for 14 heads on the graph. This bar is the shape of a rectangle. If this rectangle has a base of 1, what is its area?
- The area of this rectangle is also 0.1583. (Remember, the area of the rectangle is just base times height.) So the area of the bar for 14 heads is the same as the probability of getting 14 heads. This turns out to be a very important concept: Area = Probability!
- Shade in the bar for 13 heads and the bar for 14 heads. Find the total area of these two bars. (This will be the same as the probability of getting either 13 or 14 heads.)
- You should have found that the area for bars 13 and 14 was $.1511 + .1583 = .3094$. Shade in all of the bars that are within one standard deviation of the mean. What is the total area for these bars?
- Your answer for part **f** on this problem should have matched your answer for **2g**. What is the total area for all of the bars on the histogram? Explain.

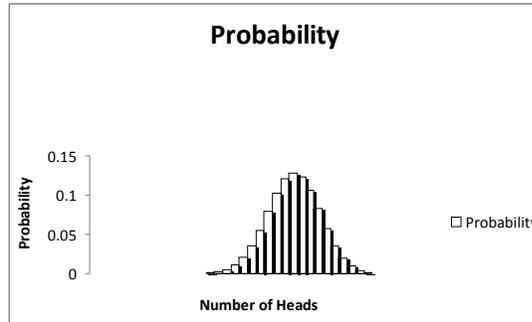
4. Suppose we have a bent coin that lands heads 60% of the time and lands tails 40% of the time. We flip it 40 times.
- Find the mean and standard deviation for the number of heads.
 - Which values are within one standard deviation of the mean?
 - What is the probability of getting exactly 24 heads? (You may use your calculator, but do not use wolframalpha.)
 - Woframalpha makes calculations like the last one easy. Use wolframalpha to find the probability of getting exactly 24 heads. Make sure your answer matches your answer for part **b**. *Hint:* What expression should we put into wolframalpha?
 - Here's the expression you should have put into wolframalpha: $(0.6h + 0.4t)^{40}$. Using wolframalpha and your calculator, find the probability of getting a value within one standard deviation of the mean. *Note:* The answer for **b** was 21, 22, 23, 24, 25, 26, and 27. You will need to add up the probabilities for each of these possibilities.

Some answers a. mean = $40(.6) = 24$, standard deviation = $\sqrt{40(.6)(.4)} = 3.0984$

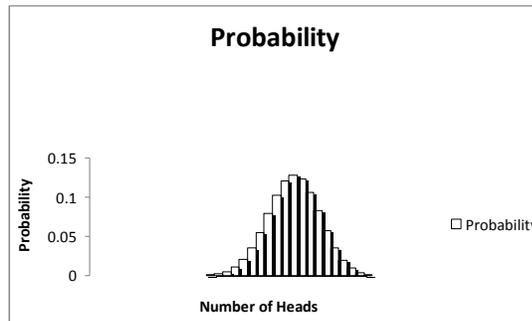
b. 21, 22, 23, 24, 25, 26, 27 c. $({}_{40}C_{24})(.6)^{24}(.4)^{16} = 0.1279$ d. $(.6h + .4t)^{40}$ has the term $0.1279h^{24}t^{16}$, so 0.1279. e. $.0792 + .1026 + .1204 + .1279 + .1228 + .1063 + .0827 = .7419$

4. (Problem 4, continued)

- f. The histogram for this situation is given below. Shade in all of the parts of the histogram that are within one standard deviation of the mean. What is the area of the shaded part?

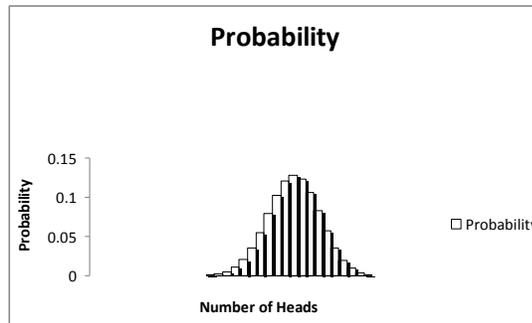


- g. On the last problem, you should have found that the area of the shaded part was 0.7419 (because that was the probability you found on in part e). Now shade in all of the values that within two standard deviations of the mean. What is the area of this shaded part?

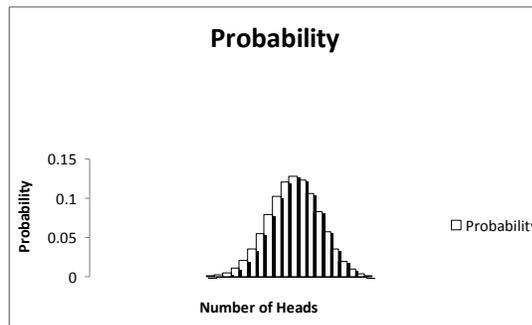


- i. On the last problem, you should have found that the area of the shaded part was 0.9655. (Here's an explanation of how you should have done this. The standard deviation is 3.0984, so two standard deviations would be $3.0984 \cdot 2 = 6.1968$. The mean is 24, so two standard deviations above would be $24 + 6.1968 = 30.1968$ and two standard deviations below would be $24 - 6.1968 = 17.8032$. Since the number of heads has to be a whole number, the possibilities are 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, and 30. Adding up all of these probabilities (which you got from wolframalpha), we get 0.9655.) This means that 96.55% of the time, the number of heads is within two standard deviations of the mean.

On the histogram below, shade in all of the values that are at least one standard deviation **above** the mean. Find the area of this shaded part. **Note:** After a certain point, the probabilities get so small as to be negligible. Don't worry about adding in the super tiny probabilities.

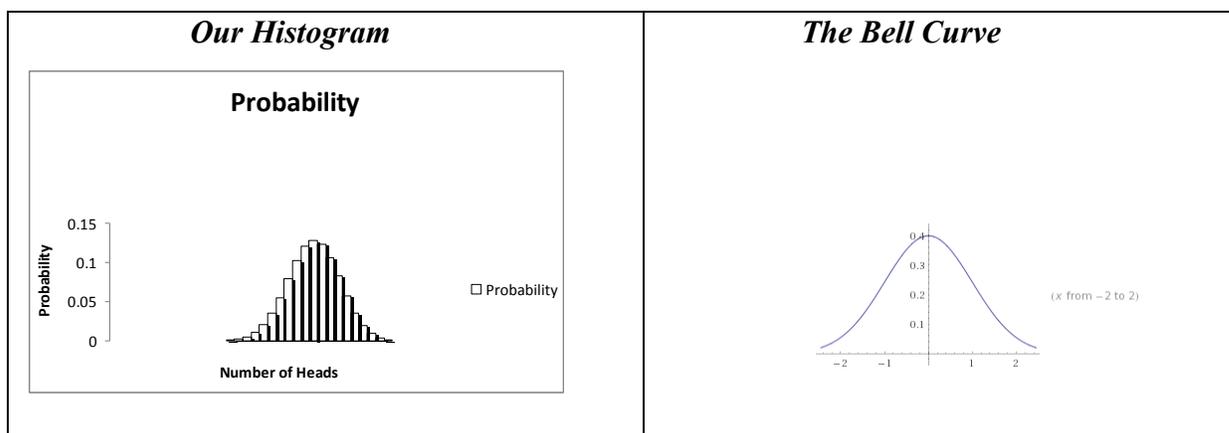


- j. In the last problem, you should have found that the area of the shaded part was about 0.1285. This means that 12.85% of the time, we will get a value that is at least one standard deviation above the mean. Now shade in the entire histogram. What is the area of the entire thing? **Hint:** You should be able to answer this question without doing any calculations!



We have established that we can find probabilities by finding areas on our histogram. Our next goal is to use the normal curve to estimate these probabilities. (This will eliminate some of the tedious addition we had to do on the last problem.)

5. Consider the histogram from the last problem and the normal curve (also known as the bell curve). Both are graph below.



- a. How does the shape of the histogram compare to the shape of the bell curve?

- b. Here's the equation for the bell curve: $y = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$. Notice that this equation involves the number e . The calculator has a button for e (right below the button for π). Press this button on the calculator. (You will need to use the 2nd feature, the same way you do π on the calculator.) You should get 2.718281828.

- c. You won't be required to memorize the equation for the bell curve, but you should be able to graph it on your calculator. Graph it now, using the following window: $-5 \leq x \leq 5$ and $0 \leq y \leq 0.5$. Does it look roughly like the histogram above?

- d. Sketch a picture of the graph you got from your calculator. Then shade in the part of the graph that is within one standard deviation of the mean. (Your picture should look a lot like the histogram in **4g**.)

- e. This area represents the probability that we are within one standard deviation of the on the bell curve. The calculator will find this area for us. First, graph the bell curve. Then, press 2^{nd} , CALC. (The CALC button is the second feature of the TRACE button, at the top of the calculator.) Choose 7: $\int f(x)dx$. The calculator will ask you for a lower limit. Press -1 and then press ENTER. (We use -1 because we are starting one standard deviation below the mean.) Now the calculator will ask you for an upper limit. Press 1 and then press ENTER.

If you've done everything right, the calculator will actually shade in this area for you! It will also give us the area at the bottom of the screen: 0.6827. This means that 68.27% of the bell curve is within one standard deviation of the mean.

- f. Using the same method from the last problem, determine what percent of the bell curve is within two standard deviations of the mean.

- g. What percent of the bell curve is within three standard deviations of the mean?

Some answers f. 0.9545 or about 95% g. 0.9973 or about 99.7%