Free Throws and the Binomial Distribution:

Who is the best free throw shooter on your high school basketball team? What is his season free throw percentage?

In this lesson, the variable p will represent the free throw percentage of your best free throw shooter. If the athlete is shooting with 90% accuracy then p = 0.9; thus p represents probability of a "success" (making a basket). What is the probability the player will make 3 shots out of 10? 25 shots out of 500? x shots out of n attempts?

Sound like a *binomial setting*? Let's check the 4 conditions:

1) If the events are independent then knowing the results of one free throw does not tell you anything about the next one. Are the events independent? Explain.

2) There are two outcomes for each event. True? Explain.

3) There are a fixed number of observations. True? Explain.

4) The probability of success is the same for each observation. True? Well, we will assume the value of p remains constant.

Recall how to find the probability the athlete will make 3 out of 10 shots if his free throw percentage is 90%.

- Using the binomial formula: $P(X=3)=\binom{10}{3}(.9)^3(.1)^7$
- Using the calculator: binompdf(10, 0.90, 3).

First you will create several different binomial probability distributions for the number of free throws made out of 10, using different values for the free throw percentage.

Step 0: Open a new Document and then a Lists & Spreadsheets page.

Step 1: In column A, make a list of the integers 1 to 10. A fast way to do this is to type =seq(w,w,1,10) in the gray box above cell A1. Name this list **x**. This is the random number of free throws made out of 10.

Step 2: In cell B1, type the formula =binompdf(10,0.90,a1). Then click and hold on this cell and arrow down so that all the cells b1 through b10 are highlighted. Press enter and the formula will fill down the column. Now name this list **bd90**.

Step 3: Repeat Step 2 in column C with the formula =binompdf(10,0.80,a1) and name this list **bd80**.

Step 4: Repeat Step 2 in columns D, E, and F with the probability in the binompdf with p = 0.70, 0.60, and 0.50 and naming the lists **bd70, bd60,** and **bd50** respectively.

Use the table you just created to answer questions 5 & 6.

5) What is the probability of making 3 free throws when the free throw percentage is 80%? What is the probability for making 8 free throws?

6) What is the probability of making at most 3 free throws out of 10 when the free throw percentage is 50%? Do this calculation also with the binomial formula to verify the answer in the table.

Step 5: Now you will create graphs for the binomial distributions you just made. Add a Data & Statistics page to your document.

Step 6: Arrow over to the horizontal axis, add the variable **x**, then arrow over to the vertical axis and add the variable **bd90**.

Step 7: Repeat Steps 5 and 6 for the distributions **bd80, bd70, bd60,** and **bd50.** If you like, you can put more than one graph on the same screen by choosing Page Layout in the Tools Menu (ctrl Home).

You should now have five graphs of the binomial probability distributions for these different values of p.

7) Describe the changes in the graphs you made as the probability decreases.

Step 8: Go back to the graph of **bd90**. From the Menu, choose Analyze, and then Plot Function. Enter normpdf(**x**,0.9, $\sqrt{10\cdot 0.9\cdot 0.1}$)

Recall the mean of the binomial distribution ($n\cdot p$) and the standard deviation of the binomial distribution ($\sqrt{n\cdot p(1-p)}$).

Step 9: Repeat Step 8 for the remaining 4 graphs on pages 1.3 through 1.6.

8) For which of the 5 graphs do you think the normal density curve **best** approximates the binomial distribution?

Step 10: Now open the file approx_binomial.tns.

Step 11: Page 1.1 shows a graph of the binomial distribution for the chosen values of p and n on the sliders. You may need to Zoom Data (Right click – Ctrl Menu) when you change the values of n or p.

Step 12: Page 1.2 shows a graph of the residuals for this binomial distribution and the normal curve using the chosen values for n and p to calculate the mean and standard deviation for the normal curve.

Page 1.3 shows the binomial distribution and the residuals for the normal approximation on the same page if you prefer to see them at the same time.

9) Experiment with many different values of n and p. You can click on a single point of the dot plot to read it’s residual coordinates.

If n is relatively small, how does the value of p influence how well the normal distribution fits the binomial distribution?

As n gets large, does the value of p influence how well the normal distribution fits the binomial distribution? Explain.

10) An established rule in statistics is that a normal distribution N($np, \sqrt{np(1-p)}$) approximates a binomial distribution for values of n and p such that $n \cdot p \geq 10$. Explain how the graph of the residuals helps you visualize why this rule makes sense. Use various values of n and p and then sketch the plots for at least two values that follow this established rule. Show the largest residual for each plot.

11) Set the value of n to 100 and the value of p to approximately 0.6. Back to our free throw example to explore a cumulatative case. Suppose the athlete shoots 100 free throws and has a 60% free throw percentage. What is the probability that the athlete will make **at least** 50 free throws?

Insert a new Calculator page to find $P(X \geq 50)$ using both binomcdf() and normalcdf().

How close is your normalcdf() approximation to the answer using binomcdf()?

Would you recommend using the cumulative normal distribution function to estimate this binomial probability? Why or why not?

Wrap It Up!

Under what conditions would you recommend the normal distribution be used to estimate binomial probabilities?