NAME: Gavin McRoy

OBJECTIVES: The purpose of this activity is to use simulations in JMP to approximate the sampling distribution of the sample **proportion**. Upon successful completion of this activity, you will be able to…

- Calculate the mean and standard error for a sampling distribution of a sample proportion given population information and the sample size,
- Use JMP to draw a random sample of a given size,
- Simulate the sampling distribution of the sample proportion
- Describe the changes that take place in the shape of the sampling distribution of the sample proportion as the sample size increases, and
- Find and interpret probabilities using the sampling distribution of the sample proportion.

Delectable Delights is a large consumer food manufacturer selling its products in retail stores nationwide. You have landed your first job after graduation from Clemson in their advertising division. Since you took statistics as a part of your coursework, you are often called upon to perform data analysis for the advertising division, as well as other divisions of the company. The problem in this lab will be revisited in later labs.

DIRECTIONS: Answer the following questions using complete sentences as though you were presenting your analysis to the employees of Delectable Delights. Please provide any appropriate output and/or screenshots from JMP. Instructions for creating several types of graphs or tables and statistics can be found on Canvas in the file **JMP Instructions.docx**. Paste your answers and any output into this document. Read carefully!

Sampling Distribution of the Sample Proportion (75 pts)

Delectable Delights is evaluating the vision benefits it offers to its employees. They wonder if they should add some additional benefits to assist with the purchase of contact lenses. According to the Vision Council of America (VCA), 11% of American adults wear contact lenses. Delectable Delights wonders if the percentage of its employees who wear contact lenses is about the same as the percentage reported by the VCA.

We will revisit this problem in future labs. For now, let's set the stage for this type of problem.

Part 1 (35 points)

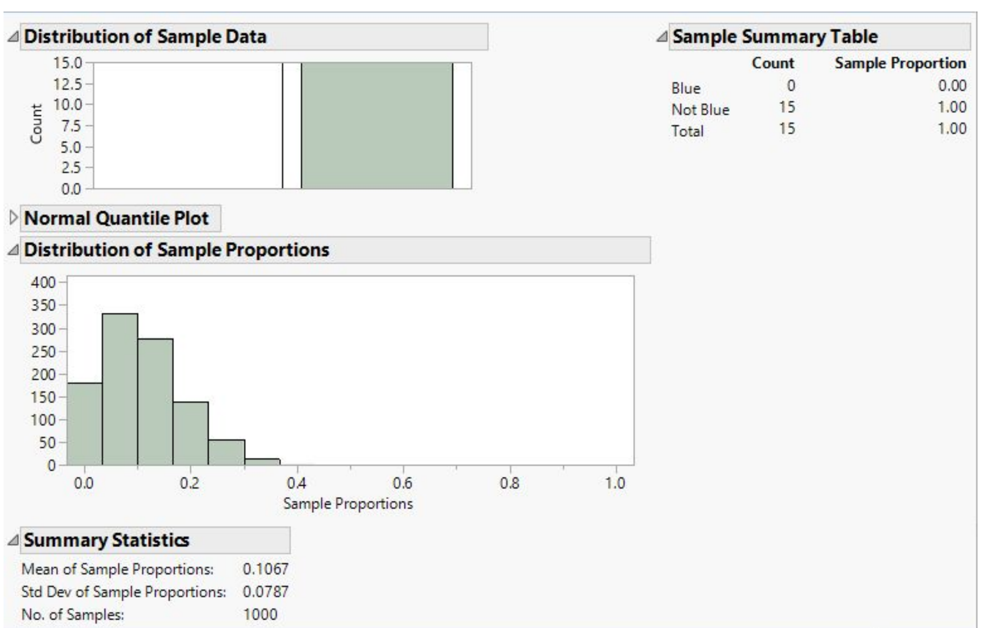
Many of the techniques used in this course require the sampling distribution of the sample statistic of interest to be approximately normally distributed. Let's explore what sample size we should select from the Delectable Delights employees so that the sampling distribution of the sample proportion of employees who wear contact lenses is approximately normally distributed. We will use the JMP applet found by selecting **Help >> Sample Data >> Teaching Scripts >> Interactive Teaching Modules >> Sampling Distribution of Sample Proportion**.

a) To estimate the sampling distribution for $\hat{p}$, we are going to draw random samples of size $n = 15$. Calculate the values of np and $n(1-p)$. Should you expect that the sampling distribution of the sample proportion for this combination of n and p to be approximately normally distributed? (5 pts)

15*.11 = 1.65 which is not greater than or equal to 5
15(1-.11) = 13.35 which is greater than 5. In order for it to be approximately normally distributed both 15*.11 and 15(1-.11) must be greater than 5 or equal to 5 which is not true. Therefore it's not normally distributed

b) Open the **Sampling Distribution of Sample Proportion** applet in JMP. In the **Population Characteristics** section, change the population proportion to 0.11 and the category name to Contacts. Under **Demo Characteristics**, change the sample size to 15 and the number of samples to 1000. Then click **Draw Additional Samples** to draw 1000 random samples of size n = 15 and to calculate the sample proportion that wear contacts for each of the 1000 samples.

Paste the **histogram** and **summary statistics** in the space below. Be sure to use the information for all 1000 samples, not just the most recent sample that is displayed at the top of the page. Notice that the Mean of the Sample Proportions is very close to the population proportion of 0.11. (10 pts)

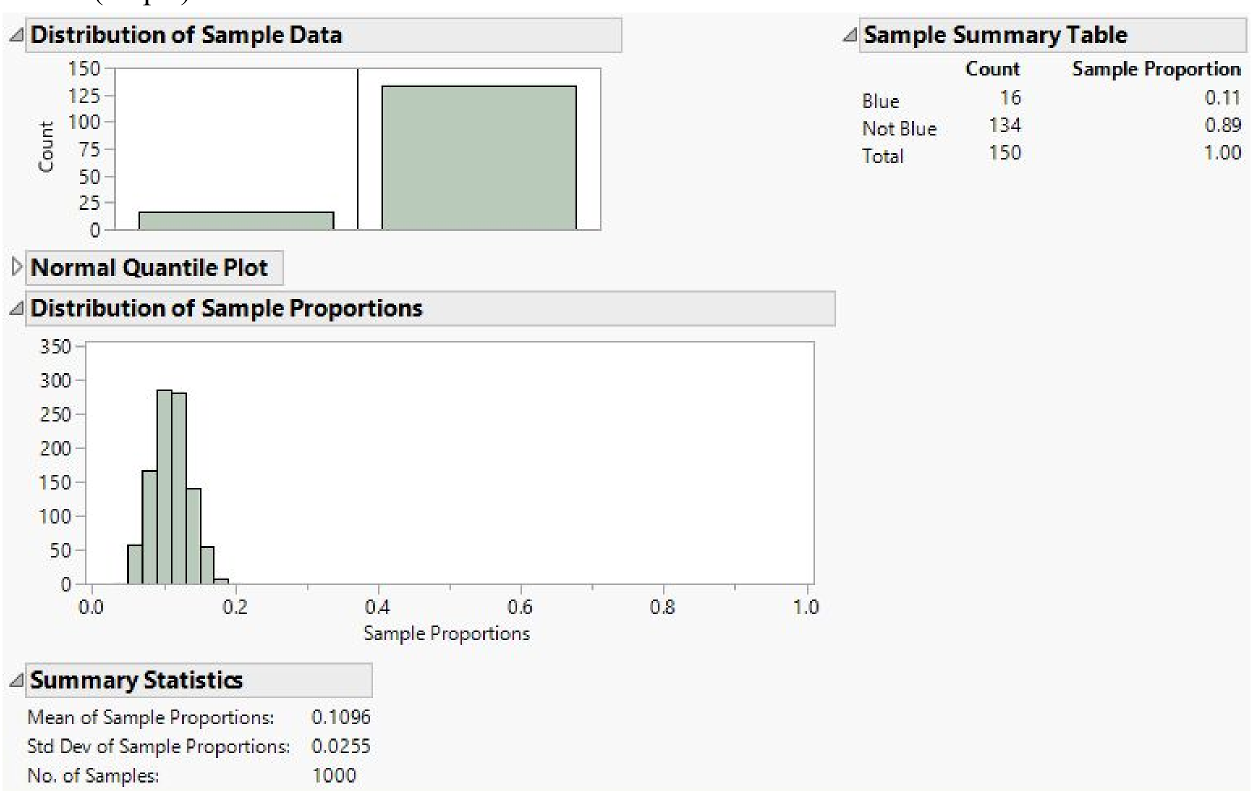


c) Now let's estimate the sampling distribution of $\hat{p}$ by drawing random samples of size $n = 150$. Calculate the values of np and $n(1-p)$. Should you expect that the sampling distribution of the sample proportion for this combination of n and p to be approximately normally distributed? (5 pts)

p = .11
150*.11 = 16.5 which is greater than 5
150(1-.11) = 133.5 which is greater than 5
In order to be approximately normally distributed both np and n(1-p) must be greater than or equal to 5. This is true therefore it is approximately normally distributed

d) Click **Reset Samples** in the JMP applet and repeat the process in Part b) to draw 1000 random samples of size $n = 150$ employees.

Paste the **histogram** and **summary statistics** for the sampling distribution in the space below. (10 pts)



e) Based on the results above, what sample size (15 or 150) should we select so that the sampling distribution of the sample proportion will be approximately normally distributed? (5 pts)

-We should chose 150 so that the sampling proportion will be approximately normally distributed

<u>Part 2</u> (40 points)

Regardless of your answer in Part e) above, suppose that you take ONE sample of 150 employees and calculate the proportion who wear contacts. Since $np = 16.5 \geq 5$ and $n(1-p) = 133.5 \geq 5$, we can consider the sampling distribution to be approximately normally distributed. We will then see if the proportion calculated is reasonably close to 0.11. For this problem, we will consider the sample results to be "reasonably close" to 0.11 if the sample proportion $\hat{p}$ is **within 2 standard deviations** from the true mean of the sampling distribution, which is the 0.11 (11%) given by the VCA.

a) Before you take your sample, calculate the mean and standard error for the sampling distribution of $\hat{p}$ for a sample of size 150 and proportion of success 0.11 (see page 124 of the Lecture Guide). Round the standard error to 3 decimal places. Notice that this value should be quite close to the simulated standard error (Std Dev of Sample Proportions) in Part d) above. (10 pts)

$\mu_{\hat{p}}$ = .11 since np>=5 and n(1-p) >=5

$\sigma_{\hat{p}}$ = sqrt(p(1-p)/n) = sqrt(.11(1-.11)/150) = .02555

b) Suppose that the following Sample Summary Table is generated for your single random sample of 150 employees.

Sample Summary Table	Count	Sample Proportion
	10	0.07
Contacts	140	0.93
Not Contacts	150	1.00
Total		

Is the sample proportion within 2 standard errors *below* or 2 standard errors *above* the mean of the sampling distribution you found in Part a)? (Hint: One way to answer this question is to find the z-score for the sample proportion using the formula on page 124 of your Lecture Guide. Is the z-score between -2 and +2?) (10 pts)

z = (.07 - .11) / sqrt(.11(1-.11) /150)
z = -1.56

The sample proportion is within 2 standard errors above or below of the sampling distribution

This question is similar to what you will do in Chapters 9 and 10, when we introduce *confidence intervals* as a method of making inference to a population based on a sample.

c) Suppose that you took a different random sample of 150 employees that resulted in a sample proportion $\hat{p} = 0.05$. What is the probability of getting this value of $\hat{p}$ or something smaller when the true population proportion is 0.11? In other words, find $P(\hat{p} < 0.05)$. Use probability notation and **show your work** to calculate a z-score by substituting values into the correct formula. You may round values to two decimal places, and you may find the final answer using either your calculator or the table. (10 pts)

z = (.05 - .11) / sqrt(.11(1-.11) /150)
z = -2.348

d) Using your calculation from Part c) above, fill in the blanks. (10 pts)

The probability of obtaining a sample proportion of 0.05 from a sample of size 150 when

The true population proportion is 0.11 is .0094. This is a low probability.
high or low?