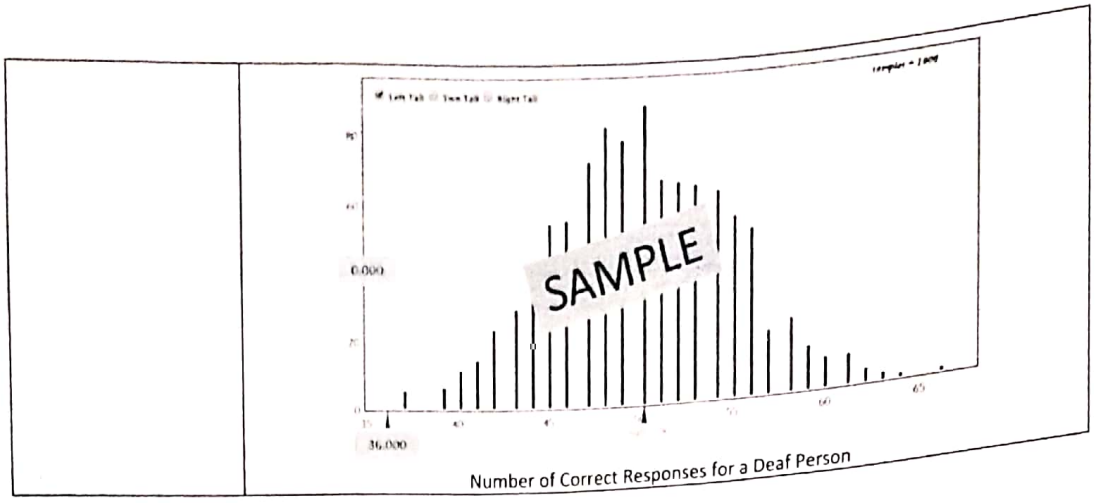


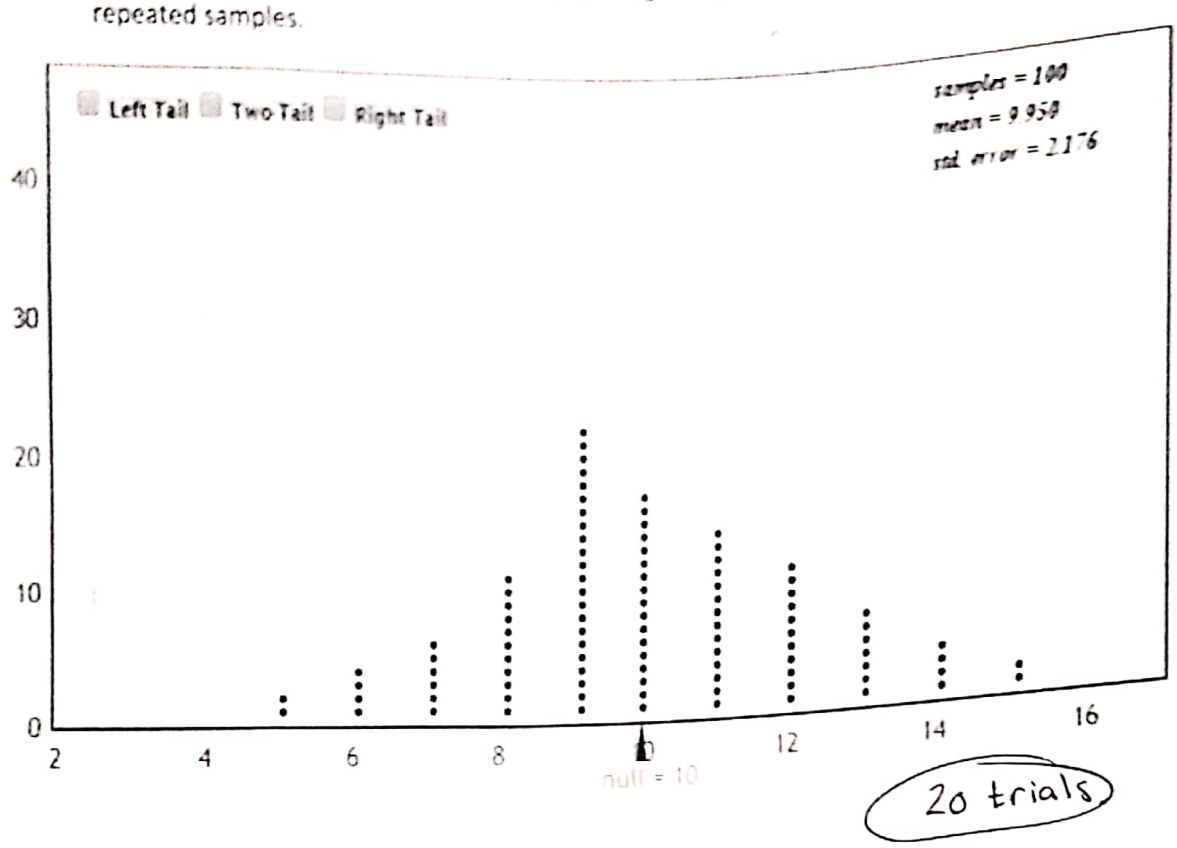
Questions:

9. If a person is truly deaf, how likely is it that they would obtain 36 or less correct out of 100 trials?

   NOT VERY LIKELY | SOMEWHAT LIKELY | LIKELY | VERY LIKELY

10. Does your investigation support or refute the notion that the subject under investigation is most likely lying about his ability to hear? Discuss.

To obtain repeated outcomes for the reference distribution, simple click the button associated with the number of repeated samples to generate. For example, to obtain 100 repeated samples, click Generate 100 Samples. Clicking additional times will generate additional repeated samples.



Recall the context of the deaf example, i.e. the simulation model was setup to mimic the outcomes of deaf people taking this forced-choice hearing test. Answer the following questions using the simulated reference distribution obtained above.

Questions

1. What does each dot on your graph represent? Discuss.

   A person taking the test, in this case they are simulated deaf people

2. If an individual is really deaf, what outcomes are we likely to observe?

   Results similar to that above. we could say anywhere 6-14 is likely deaf

3. What is the smallest value in your dotplot? 5

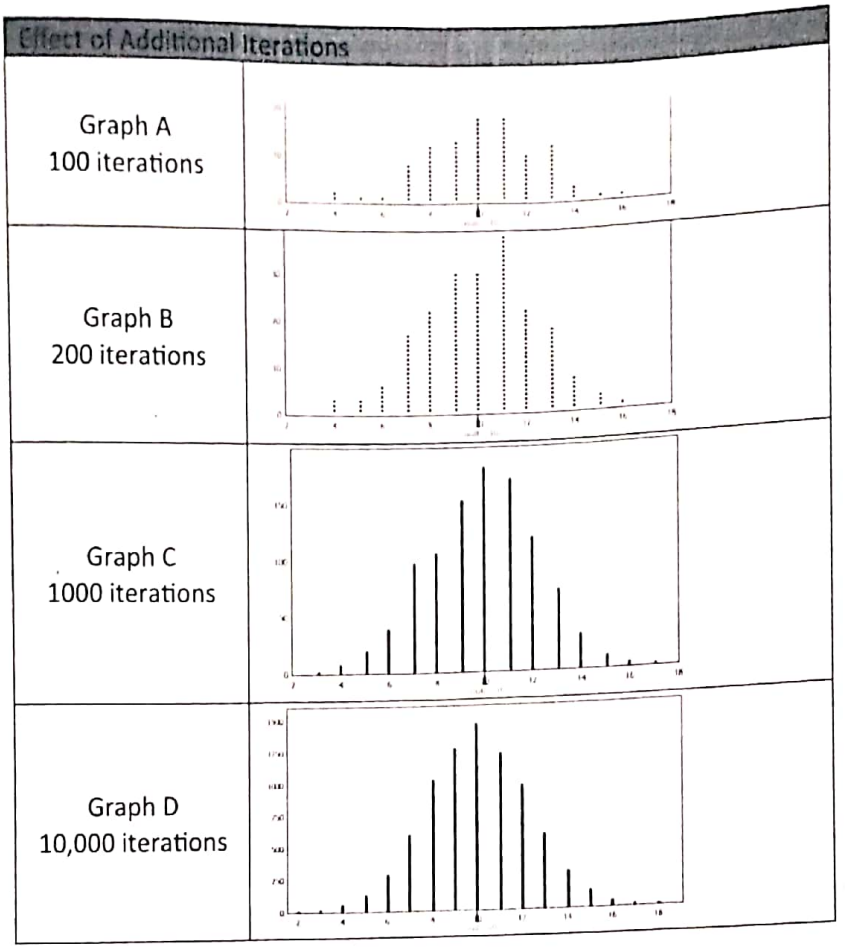
   a. Is it possible to obtain a value smaller than this if more simulations were carried out?

   Yes. It is unlikely that really low values will happen but it is still possible. The pyramid will grow but it will still be a pyramid. That's why we can't just use the threshold values to decide our range

b. If one additional simulation is done, is this outcome likely to be smaller, larger, or about the same as the previous smallest value? Discuss.

Larger because the smallest value is unlikely to occur - It will more likely be close to 10.

4. A statistician might argue that a plot based on 200 repeated iterations of the simulation is *better* than a plot based on 100 repeated iterations. Do you agree? Explain.



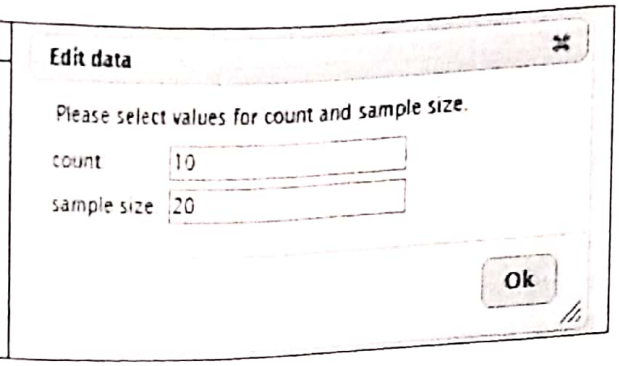
Questions

5. What are the similarities amongst these four graphs?

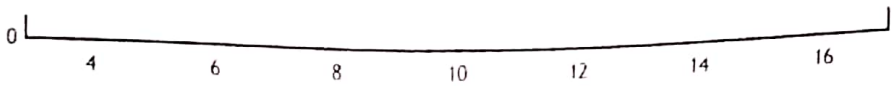
The general shape, median, mean

Step 1: Setting up the number line

To specify the appropriate number line, select Edit Data. In the Edit data window, you must specify values for count and sample size.

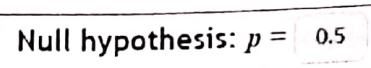
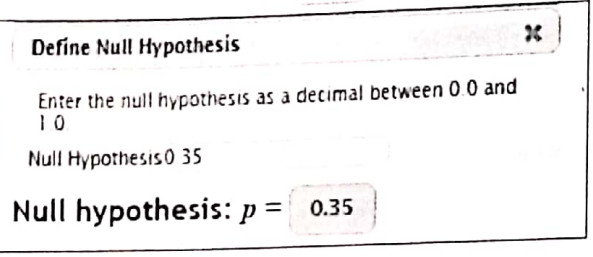
count = the expected value Sample size = the maximum value on the number line.	

Finally, you should change the scale on the number line from Proportion to Count. The number line for this example looks as follows.



Step 2: Specify the location of reference distribution

The 50/50 situation is the default location of the distribution on the number line. This can be changed by simply clicking on the Null Hypothesis: p = 0.50 box.

Default Setting	
Specify a different location for the reference distribution	

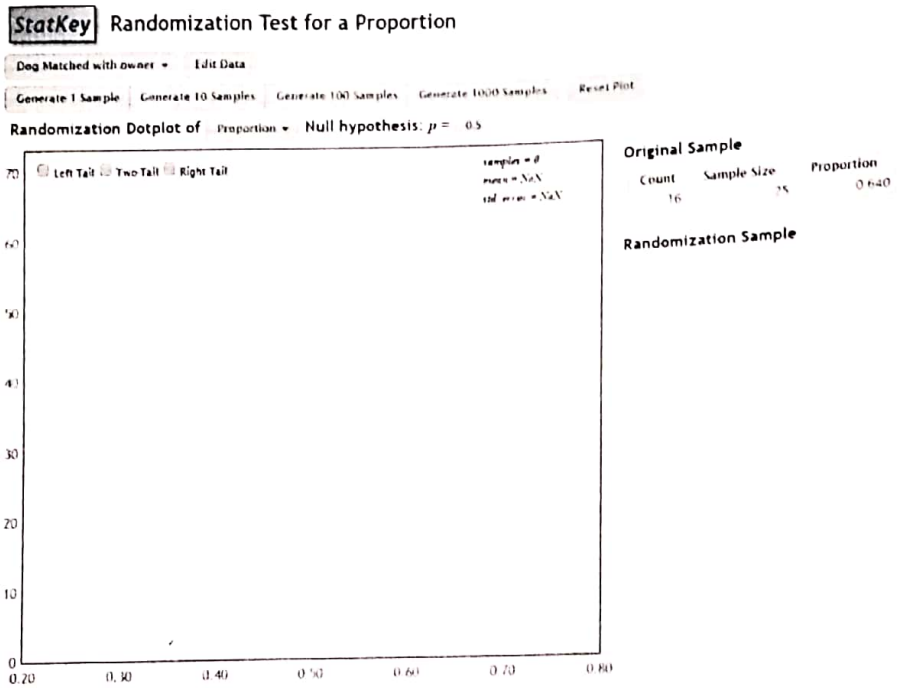
Step 3: Generate repeated outcomes

To save time and to gather more simulated results more quickly, software technologies can be used to simulate the outcomes obtained in class for the example in Handout #1. The results from the simulation model can then be used to help us understand what outcomes are likely (or unlikely) to occur. This simulation model will be constructed using the following parameters (i.e., under these conditions).

Necessary information for building a model	Setup for Deaf Example	StatKey Notation
• Number of completed trials	20	Denoted by Sample Size
• The likelihood or chance of obtaining a correct response	1 out of 2; $\frac{1}{2}=0.5$	Null hypothesis: p = 0.5

Setting up the Simulation using StatKey

App Link: http://www.lock5stat.com/StatKey/randomization_1_cat/randomization_1_cat.html



6. What would you consider an *unusual* outcome (i.e., outlier) to be in Graph A? How about Graph B?

~~A-4,13 B-4,~~

A - under 6, over 15 B - under 4, over 15

Developing a Rule for Identifying Outliers

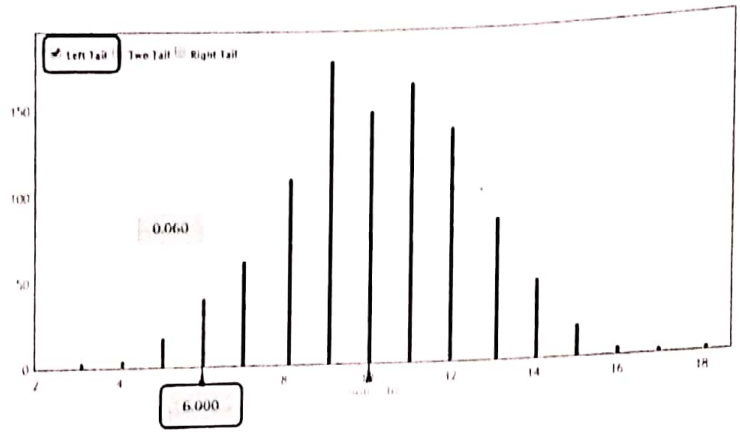
An **outlier rule** determines the location on the distribution for which you begin to believe observations are unlikely.

Comments:

- A statistician might argue that the determination of what constitutes an outlier should not be adversely affected by the number of iterations conducted in the simulation. In fact, if enough iterations of the simulation model are completed, eventually all possible values will appear on the dotplot.

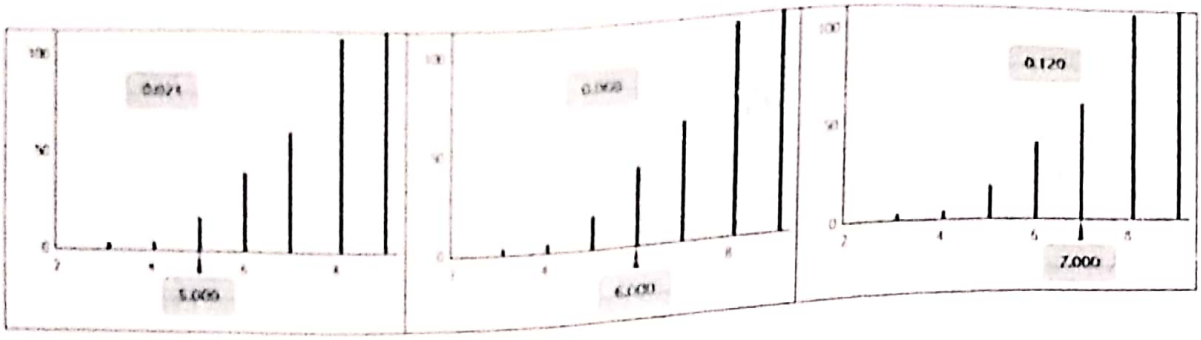
- An outlier rule is used to separate *likely* from *unlikely* observations. This is the main goal of obtaining the distribution for the number of correct responses. For example, a statistician might say that a value of **6 or less** would be a reasonable outlier rule. Such a rule can and should be applied consistently with any of the graphs shown above.

- The lower end of the reference distribution is of interest for the Deaf example, i.e. subject has ability to hear, but is thought to be lying about this ability. Thus, specify Left Tail in the top right corner of the graph. Next, in the cutoff box, specify 6.



Suppose a statistician decides to use 6 or less as the rule for determining when somebody is lying about their ability to hear. If this rule is used, then about 0.06 = 6% of the time we'd accuse a deaf person of lying about their ability to hear.

Cutoff = 5	Cutoff = 6	Cutoff = 7



Questions

Type 1,2 errors

7. Suppose instead of using 6 or less as the rule, 7 or less was used. Would this result in more or less deaf people being identified as lying? Would you advocate changing the rule to 7? Discuss.

More people Sweet spot is ~5%. So no I'd keep it at 6

8. Would an insurance company who fears insurance fraud rather use the rule of 6 or less or 7 or less? Discuss.

7. They'd catch more ~~###~~ people pretending to lie. unfortunately for deaf people, fortunately for the company, less deaf people would be fond as well.

Task

Set up a simulation in StatKey to evaluate the subject presented in the article by Pankratz, Fausti, and Peed titled "A Forced-Choice Technique to Evaluate Deafness in the Hysterical or Malingering Patient."
Source: Journal of Consulting and Clinical Psychology, 1975, Vol. 43, pg. 421-422.

In this study, the subject was asked to complete 100 trials (instead of 20 as used above). The sample size value should thus be set to 100. Obtain 1000 repeated sample for your simulation

Necessary information for building a model	Deaf Example
• Number of completed trials	100
• The likelihood or chance of obtaining a correct response *assuming* the subject cannot hear	1 out of 2; i.e., $\frac{1}{2}$
• Generate 1000 repeated samples	1000

Likely Outcomes from a Deaf Person	A SAMPLE PLOT IS PROVIDED HERE, DELETE THIS PLOT AND COPY THE GRAPH FROM YOUR SIMULATION.