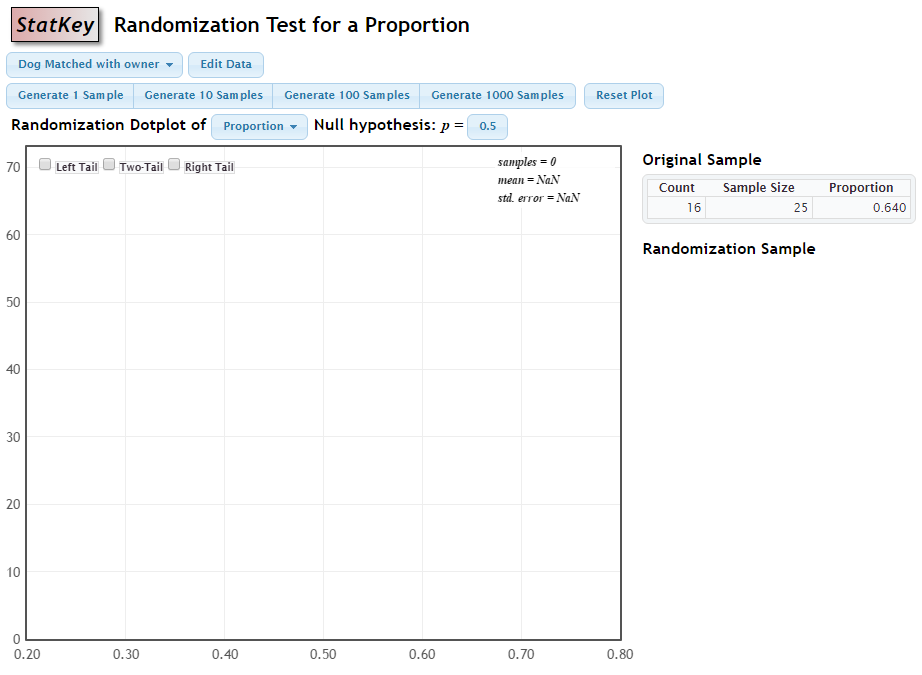
To save time and to gather more simulated results more quickly, software technologies can be used to simulate the outcomes obtained in class for the example in Handout #1. The results from the simulation model can then be used to help us understand what outcomes are likely (or unlikely) to occur. This 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; |  |

Setting up the Simulation using StatKey

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



Step 1: Setting up the number line

To specify the appropriate number line, select . 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  to . The number line for this example looks as follows.



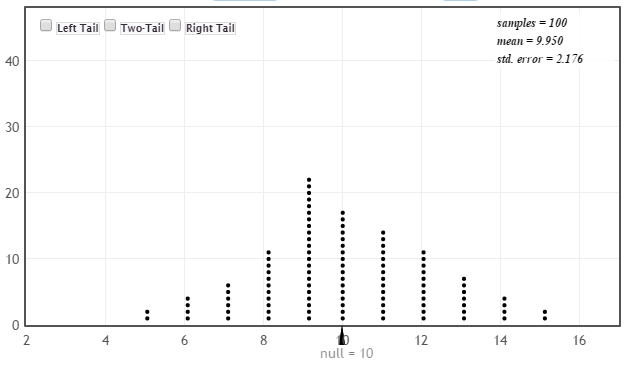
Step 2: Specify the location of reference distribution

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 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 . 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.
2. If an individual is really deaf, what outcomes are we likely to observe?
3. What is the smallest value in your dotplot? \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_
   1. Is it possible to obtain a value smaller than this if more simulations were carried out?
   2. If one additional simulation is done, is this outcome likely to be smaller, larger, or about the same as the previous smallest value? Discuss.
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.

|  |  |
| --- | --- |
| Effect of Additional Iterations | |
| Graph A 100 iterations |  |
| Graph B 200 iterations |  |
| Graph C 1000 iterations |  |
| Graph D 10,000 iterations |  |

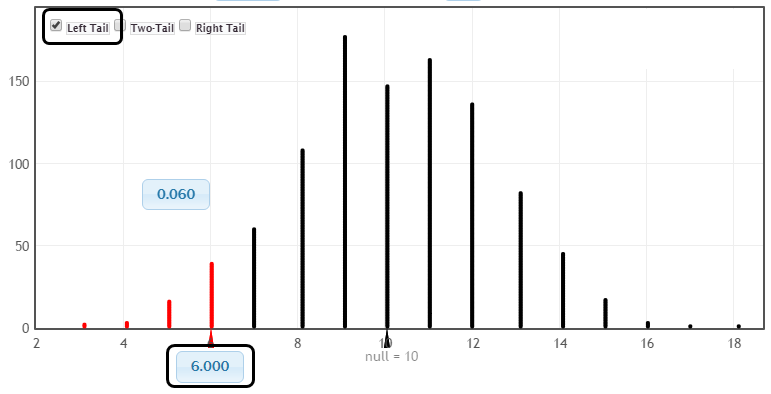
Questions

1. What are the similarities amongst these four graphs?
2. What would you consider an *unusual* outcome (i.e., outlier) to be in Graph A? How about Graph B?

|  |
| --- |
| 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

1. 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.
2. Would an insurance company who fears insurance fraud rather use the rule of 6 or less or 7 or less? Discuss.

Task

Set up a simulation in StatKeyto 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., |
| * 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.  Number of Correct Responses for a Deaf Person |

Questions:

1. 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

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