Data 8R
Summer 2017

Hypothesis Testing

Discussion 9: July 27, 2017

1 Terminology

Write down a definition, in your own words, for the following terms:

The Null Hypothesis

The Alternative Hypothesis

The Test Statistic

After we've defined our hypotheses, how do we go about testing them?

2 Create Some Hypotheses

Suppose that you're at the casino, playing dice (with a six-sided die). You suspect that the die is loaded - the dice rolls you see are abnormally high. Define a test statistic, null, and alternative hypotheses.

Null Hypothesis:

Alternative Hypothesis:

Test Statistic:

After trying your luck with the dice to no avail, you're back at work as a spearmint gum quality control specialist. You begin to notice that a lot of the gum has minor defects. You suspect that it might be due to more than chance. How do we go about testing this hypothesis?

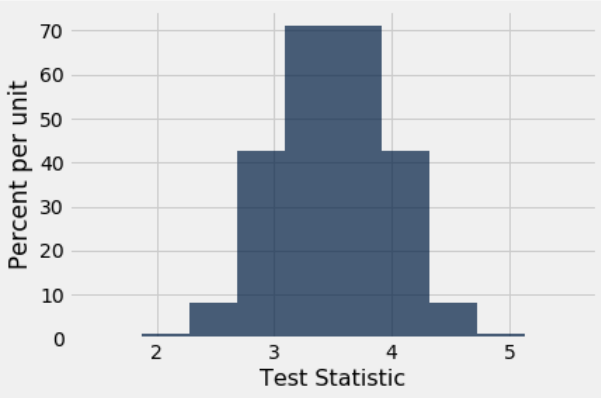
Null Hypothesis:

Alternative Hypothesis:

Test Statistic:

3 Evaluate the Hypotheses

After simulating the data for your dice rolls, you produce the following histogram:



If the mean of the dice rolls you observed was 3.923, what could you conclude from the histogram?

If the mean of the dice rolls you observed was instead 5.4, what could you concluded from the histogram?

Extra Optional question! Given that our data is stored in a table called **dice_data**, how can we create one simulation of the data, and then get our test statistic for that data? **dice_data** has 1 column named "Rolls". Suppose that we also have a table named **possible_dice_rolls** that contains all the possible dice rolls (1 through 6), which has the same column name.