Our "Coin" and the Language of Hypothesis Testing

Math 5 Crew

Department of Mathematics

Dartmouth College

The Null Hypothesis

- Question 1. Suppose our "coin" is fair. What do we expect to happen when we "flip" it?

The Null Hypothesis

- Question 1. Suppose our "coin" is fair. What do we expect to happen when we "flip" it?
- *Null Hypothesis*: Is usually some articulation of "Nothing is going on".

The Null Hypothesis

- Question 1. Suppose our "coin" is fair. What do we expect to happen when we "flip" it?
- *Null Hypothesis*: Is usually some articulation of "Nothing is going on".
- Our Example: "Heads" and "tails" are both equally likely.

The Alternate Hypothesis

- Question 1. Suppose our "coin" is fair. What do we expect to happen when we "flip" it?

The Alternate Hypothesis

- Question 1. Suppose our "coin" is fair. What do we expect to happen when we "flip" it?
- *Alternate Hypothesis*: Is usually some articulation of "Something is going on".

The Alternate Hypothesis

- Question 1. Suppose our "coin" is fair. What do we expect to happen when we "flip" it?
- *Alternate Hypothesis*: Is usually some articulation of "Something is going on".
- Our Example: Our "coin" is biased and either "heads" or "tails" is more likely.

Testing The Null Hypoothesis

- Question 2. Suppose we "flip" our "coin" 25 times. How might you decide whether or not our "coin" is biased?

Testing The Null Hypoothesis

- Question 2. Suppose we "flip" our "coin" 25 times. How might you decide whether or not our "coin" is biased?
- Find a numerical *parameter*, whose behavior you understand assuming the Null Hypothesis.

Testing The Null Hypoothesis

- Question 2. Suppose we "flip" our "coin" 25 times. How might you decide whether or not our "coin" is biased?
- Find a numerical *parameter*, whose behavior you understand assuming the Null Hypothesis.
- You intend to run an experiment which determines this number, and call this number a *test statistics*.

Testing The Null Hypoothesis

- Question 2. Suppose we "flip" our "coin" 25 times. How might you decide whether or not our "coin" is biased?
- Find a numerical *parameter*, whose behavior you understand assuming the Null Hypothesis.
- You intend to run an experiment which determines this number, and call this number a *test statistics*.
- Our Example: Number of "heads".

Reporting

- Question 2. Suppose we "flip" our "coin" 25 times. How might you decide whether or not our "coin" is biased?

Reporting

- Question 2. Suppose we "flip" our "coin" 25 times. How might you decide whether or not our "coin" is biased?
- **Before** performing your test you must determine for which values of this test statistic you are forced to *accept* the null hypothesis and for which values of this test statistic you will *reject the Null Hypothesis* and *accepet the Alternate Hypothesis*.

Reporting

- Question 2. Suppose we "flip" our "coin" 25 times. How might you decide whether or not our "coin" is biased?
- **Before** performing your test you must determine for which values of this test statistic you are forced to *accept* the null hypothesis and for which values of this test statistic you will *reject the Null Hypothesis* and *accepet the Alternate Hypothesis*.
- Our Example: Accept if 7 < Number Heads < 18, and reject otherwise.

Reporting

- Question 2. Suppose we "flip" our "coin" 25 times. How might you decide whether or not our "coin" is biased?

Reporting

- Question 2. Suppose we "flip" our "coin" 25 times. How might you decide whether or not our "coin" is biased?
- We call the values of the test statistic for which we will reject the Null Hypothesis the *Critical Region*

Reporting

- Question 2. Suppose we "flip" our "coin" 25 times. How might you decide whether or not our "coin" is biased?
- We call the values of the test statistic for which we will reject the Null Hypothesis the *Critical Region*
- Our Example: Accept if 7 < Number Heads < 18, and reject otherwise. Hence the critical region are the integers in $[0, 7]$ togehter with those in $[18, 25]$.

Deciding

- Question 3. Suppose the "coin" is fair. What is the chance that you declare the "coin" biased?

Deciding

- Question 3. Suppose the "coin" is fair. What is the chance that you declare the "coin" biased?
- How much risk are you willing to take that the Null hypothesis is correct but you reject it? .

Deciding

- Question 3. Suppose the "coin" is fair. What is the chance that you declare the "coin" biased?
- How much risk are you willing to take that the Null hypothesis is correct but you reject it? .
- Called the chance of a *type 1 error*.

Deciding

- Question 3. Suppose the "coin" is fair. What is the chance that you declare the "coin" biased?
- How much risk are you willing to take that the Null hypothesis is correct but you reject it? .
- Called the chance of a *type 1 error*.
- Our Example: Should be near 5 percent. Check it! .

Deciding

- Question 3. Suppose the "coin" is fair. What is the chance that you declare the "coin" biased?

Deciding

- Question 3. Suppose the "coin" is fair. What is the chance that you declare the "coin" biased?
- The risk of a type 1 error is called the *Significance Level* of the experiment..

Deciding

- Question 3. Suppose the "coin" is fair. What is the chance that you declare the "coin" biased?
- The risk of a type 1 error is called the *Significance Level* of the experiment..
- In order to assure yourself that you can call your results *statisitically significant* you must set you significance lelevel to be less than 5 percent.

Deciding

- Question 3. Suppose the "coin" is fair. What is the chance that you declare the "coin" biased?
- The risk of a type 1 error is called the *Significance Level* of the experiment..
- In order to assure yourself that you can call your results *statisitically significant* you must set you significance lelevel to be less than 5 percent.
- In order to assure yourself that you can call your results *highly significant* you must set you significance lelevel to be less than 1 percent.

Deciding

- How much risk are you willing to take that the Null hypothesis is false but you accept it

Deciding

- How much risk are you willing to take that the Null hypothesis is false but you accept it
- Called the risk of a *type 2 error*.

Deciding

- How much risk are you willing to take that the Null hypothesis is false but you accept it
- Called the risk of a *type 2 error*.
- Discussion: How might you approximate a type 2 error in our setting?!

Deciding

- How much risk are you willing to take that the Null hypothesis is false but you accept it

Deciding

- How much risk are you willing to take that the Null hypothesis is false but you accept it
- The *power* of a test relative to an assumption about the statistic's behavior under the alternate hypothesis is the chance that you correctly accept the alternate hypothesis under this assumption.

Deciding

- How much risk are you willing to take that the Null hypothesis is false but you accept it
- The*power* of a test relative to an assumption about the statistic's behavior under the alternate hypothesis is the chance that you correctly accept the alternate hypothesis under this assumption.
- Our Example: Assume the "die" has a 40 percent chance of coming up heads. What is the power of our test?

Is Hanover Water Yummy?

- Develop a hypothesis test in order to test whether Hanover tap water is worse tasting than bottled water.

Is Hanover Water Yummy?

- Develop a hypothesis test in order to test whether Hanover tap water is worse tasting than bottled water.
- Articulate in words:

Is Hanover Water Yummy?

- Develop a hypothesis test in order to test whether Hanover tap water is worse tasting than bottled water.
- Articulate in words:
- a null hypothesis, the alternate hypothesis, a parameter, what you will report for the various possible values of the statistic associated to your parameter, the chance you are taking of a type 1 error and (more difficult) the chance you are taking of a type 2 error,

Is Hanover Water Yummy?

- Develop a hypothesis test in order to test whether Hanover tap water is worse tasting than bottled water.
- Articulate in words:
- a null hypothesis, the alternate hypothesis, a parameter, what you will report for the various possible values of the statistic associated to your parameter, the chance you are taking of a type 1 error and (more difficult) the chance you are taking of a type 2 error,
- ...and the protocol of how you will run the experiment. For example can you make it double blind? Are there any obvious