Our "Coin" and the Language of Hypothesis Testing

Math 5 Crew

Department of Mathematics
Dartmouth College
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".
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.
Question 2. Suppose we "flip" our "coin" 25 times. How might you decide whether or not our "coin" is biased?
Testing The Null Hypothesis

• 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 Hypothesis

- 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.
- When you perform your test you will observe one of this parameter’s possible values. Call all this value your test statistic.
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.

When you perform your test you will observe one of this parameter’s possible values. Call all this value your test statistic.

Our Example: Number of "heads".
Question 2. Suppose we "flip" our "coin" 25 times. How might you decide whether or not our "coin" is biased?
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 the values of this test statistic for which you will accept the null hypothesis and the values for which you will reject the Null Hypothesis and accept the Alternate Hypothesis.
Question 2. Suppose we "flip" our "coin" 25 times. How might you decide whether or not our "coin" is biased?
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.
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: We will accept if $7 < \text{Number Heads} < 18$, and reject otherwise. Hence, the critical region is the collection of integers $K$ that satisfy either $0 \leq K \leq 7$ or $18 \leq K \leq 25$. 
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?
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 \textit{Significance Level} of the experiment.
- In order to assure yourself that you can call your results \textit{statistically significant}, you must set you significance level 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 statistically significant, you must set your significance level to be less than 5 percent.
- In order to assure yourself that you can call your results highly significant, you must set your significance level 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 the risk of 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?
- Relative to an assumption about the statistic’s behavior under the alternate hypothesis, the *power* of a hypothesis test is the probability 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?
- Relative to an assumption about the statistic’s behavior under the alternate hypothesis, the *power* of a hypothesis test is the probability that you correctly accept the alternate hypothesis under this assumption.
- Our Example: Assume the "coin" 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 or not Hanover tap water is worse tasting than bottled water.
Is Hanover Water Yummy?

• Develop a hypothesis test in order to test whether or not Hanover tap water is worse tasting than bottled water.

• Determine:
Is Hanover Water Yummy?

- Develop a hypothesis test in order to test whether or not Hanover tap water is worse tasting than bottled water.

- Determine:
  - a null hypothesis, the alternate hypothesis, a parameter, a test statistic, the critical region, the significance level, the power,
Is Hanover Water Yummy?

- Develop a hypothesis test in order to test whether or not Hanover tap water is worse tasting than bottled water.

- Determine:
  - a null hypothesis, the alternate hypothesis, a parameter, a test statistic, the critical region, the significance level, the power,
  - and the test’s protocol. For example, can you make it double blind? Are there any obvious confounding factors? What equipment and how much time will you need?