Hypotheses Test Procedures

MATH 2300

Sections 9.1 and 9.2

Is the claim wrong?

- An oil company representative claims that the average price for gasoline in Lubbock is \$2.30 per gallon.
- You think the average price is higher so you take a random sample and find that the mean is \$2.43 per gallon.
- Does this mean that the claim is wrong?

NOT NECESSARILY!

Hypothesis Testing

- A formal procedure in which we use sample data to test the plausibility of a hypothesis about:
  - The value of a parameter
  - The value of several parameters
  - An entire probability distribution
- We test a **null hypothesis** against an **alternative hypothesis**.

The Null Hypothesis

- The *null hypothesis* is denoted by $H_{0.}$
- Is a claim that is assumed to be true.
- Will be rejected only if the sample data provide substantial contradictory evidence.
  - Otherwise, continue to believe the plausibility of H_0
- The two possible conclusions of a test:
  - Reject H_0
  - Fail to reject H_0

The Alternative Hypothesis

- The *alternative hypothesis* is denoted as H_a.
- Is an assertion that is contradictory to H_0
- Often called the *researcher's hypothesis*
  - It is often a claim that a researcher would ultimately like to validate
- If we reject H_0, we do so in favor of H_a.

Analogy: Court Room

- One claim: defendant is not guilty
- Second claim: defendant ***is*** guilty
- In the US, we assume the defendant is not guilty and put the burden on the prosecution.

- Which is H_0?

Example: Mean

You are investigating the presence of radon in homes being built in a new development.

- If the mean level of radon is greater than 4 then send a warning to all home owners in the development.

Solutions:

HH_o: The mean level of radon for homes in the development is 4 or less ($\mu \leq 4$)

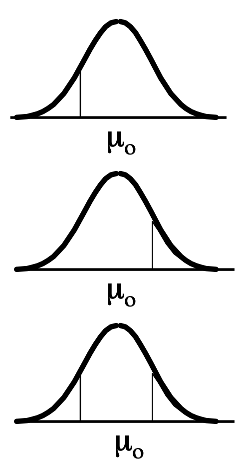
H_a: The mean level of radon for homes in the development is greater than 4 ($\mu > 4$)

Hypotheses for the Mean

Lower Tail Test: $H_0: \mu = \mu_o$, $H_a: \mu < \mu_o$

Upper Tail Test: $H_0: \mu = \mu_o$, $H_a: \mu > \mu_o$

Two Tail Test: $H_0: \mu = \mu_o$, $H_a: \mu \neq \mu_o$



Example: Mice

- It takes mice a mean time of 18 seconds to find their way through a maze.
- A researcher thinks that a loud noise will cause them to complete the maze more quickly.
- She measures how long 10 mice take to complete the maze with noise as a stimulus.
  - H_o:
  - H_a:

Testing Procedure

1. State H_0 and H_a.
2. Calculate the **test statistic**.
3. Determine the **critical value** to define the **rejection region**.
4. Reach conclusion about H_0.
5. State your conclusion in the context of your specific study.

The Test Statistic

- A function of the sample data.
- Provides a basis for testing a hypothesis.
- Measures compatibility between the data and the **null** hypothesis.
  - When H_o is true, we expect the data to closely agree with H_o.
  - When H_o is false, we expect the data to support H_a.

The Rejection Region

- The set of all test statistic values for which we would reject H_0.
  - Stated in terms of a **critical value**, which depends on assumptions about your data
- H_0 is rejected if and only if the observed/computed value of the test statistic falls in the rejection region

Example

- H_0: $\mu = 4$ H_a: $\mu > 4$
- Let the rejection region be all values of the test statistic greater than 1.645
- Would we reject H_0 for the following observed test statistics?
  - 1.25
  - 1.64
  - 1.66

How do we decide on a rejection region?

- We want to limit the probability of making errors.
- Type I Error
  - Rejecting H_0 when it is true.
  - Probability denoted by α.
- Type II Error
  - Not rejecting H_0 when it is false.
  - Probability denoted by β.

Distinguishing the Type of Error

Truth About the Population

Decision Based on Sample

	H_o True	H_a True
Reject H_o	Type I Error	Correct Decision
Fail to Reject H_o	Correct Decision	Type II Error

Alloy Example

- Test a new alloy to see if it has a higher tensile strength than that currently used.
- Type I Error
  - You find that the new alloy has a higher tensile strength when it really has an equal strength.
- Type II Error
  - You find that the new alloy has the same tensile strength as the old, when it really is higher.

Cancer Example

- A patient comes to be tested to see if he has a certain type of cancer. The initial diagnosis was that he does.
- Type I Error
  - You perform your diagnosis and conclude that he does not have cancer, when in reality, the initial diagnosis was correct.
- Type II Error
  - You perform your diagnosis and agree with the first diagnosis, when in reality, he does not have cancer.

Type I, Type II, or Correct?

- You perform a test and reject H_o.
- Further study is done and it is found that the value of μ is exactly as hypothesized.

Type I, Type II, or Correct?

- You perform a test and reject H_o.
- Further study is done and it is found that the value of μ is far different than hypothesized.

Type I, Type II, or Correct?

- You perform a test and do not reject H_o.
- Further study is done and it is found that the value of μ is far different than hypothesized.

Proposition

- An experiment and sample size are fixed with a chosen test statistic
- Then decreasing the size of the rejection region to obtain a smaller value of α results in a larger value of β for any particular parameter value included in H_a.

Choosing the Rejection Region

- Type I errors are typically more serious than Type II errors
- Approach is to select the largest value of α that can be tolerated.
- Choose a rejection region having that value of α rather than anything smaller.
- This value of α is called the **significance level** of the test.

Choosing the Rejection Region

- The *significance level*:
  - The maximum probability of rejecting H_o when it is true.
- Common values of α:
  - 0.01, 0.05, 0.10