HYPOTHESIS TESTING WITH ONE SAMPLE

MAT 152 – Statistical Methods I
Lecture 1
Instructor: Dustin Roten
Fall 2020

HYPOTHESIS TESTING

- **Confidence intervals** are one way to estimate a population parameter.
- A statistician can also make a **statistical inference** through **hypothesis testing**.
- Hypothesis testing consists of two contradictory **hypotheses** or statements, a data-oriented **decision**, and a well-formed **conclusion**.
  1. Setup two contradictory hypotheses.
  2. Collect sample data.
  3. Determine the correct distribution to perform the hypothesis test.
  4. Analyze the data.
  5. Decide on a correct hypothesis and write a meaningful conclusion.

HYPOTHESES

- Two hypotheses are required:

  - The **Null Hypothesis** (H_0): A statement of no difference between a sample mean/proportion and a population mean/proportion.

  - The **Alternative Hypothesis** (H_a): A claim that is contradictory to H_0.

- Since H_0 and H_a are contradictory, the evidence must be examined in order to make a correct statement. There are two outcomes:
  - Reject the null hypothesis H_0
  - Do not reject the null hypothesis.

EXAMPLES

- We want to test whether the mean GPA of students in American colleges is different from 2.0 (out of 4.0).
  - $H_0: \mu = 2$ (No change. The supposed mean is correct.)
  - $H_a: \mu \neq 2$ (The mean is, in fact, different from the supposed mean.)

- We want to test if college students take less than five years to graduate from college, on the average.
  - $H_0: \mu \geq 5$ (The contradictory statement to our hypothesis.)
  - $H_a: \mu < 5$ (The mean is, in fact, different from the supposed mean.)

- An article by U.S. News and World Report stated that 6.6% of U.S. students take advanced placement exams. Suppose we wish to test if the percentage is higher.
  - $H_0: p \leq 0.066$ (No change. The article is correct OR the percentage is lower.)
  - $H_a: p > 0.066$ (The percentage is, in fact, higher than the article stated.)

HYPOTHESES

H_0	H_a
$=$	$\neq, >, <$
$\geq$	$<$
$\leq$	$>$

If the null hypothesis claims that the mean/proportion is equal to some value, then the alternative hypothesis can have three different outcomes.
$H_0: \mu = 3$
$H_a: \mu \neq 3$ or $H_a: \mu > 3$ or $H_a: \mu < 3$

If the null hypothesis claims that the mean/proportion is greater than or equal to some value,
$H_0: \mu \geq 3$
$H_a: \mu < 3$

If the null hypothesis claims that the mean/proportion is less than or equal to some value,
$H_0: \mu \leq 3$
$H_a: \mu > 3$

ERRORS

- When a hypothesis test is performed, there are four possible outcomes.

1. The decision is not to reject H_0 when H_0 is true. ✓

2. The decision is to reject H_0 when H_0 is true. ✗
   (The null hypothesis is true but we reject it – Type I Error)

3. The decision is not to reject H_0 when H_0 is false. ✗
   (The null hypothesis is false but we do not reject it – Type II Error)

4. The decision is to reject H_0 when H_0 is false. ✓

TYPE I AND TYPE II ERRORS

- Type I Error
  - A true null hypothesis is rejected
  - The probability of this occurring is α. (The "leftover" from the confidence interval)
  - "We think that the status quo (H_0) is false, but it is true."

- Type II Error
  - Failure to reject a false null hypothesis
  - The probability of this occurring is β.
  - "We think that the status quo (H_0) is true, but it is true."

- The power of a test is given by $1 - \beta$.

EXAMPLE

A genetic lab claims to be able to increase the likelihood that a pregnancy will result in a boy being born. This claim must be tested.

H_0: The lab does not influence gender outcome.

H_a: The lab influences gender outcome. (The lab's claim)

Type I Error (α)

We believe that the lab influences gender outcomes, but it ACTUALLY doesn't.

Type II Error (β)

We believe that the lab does not influence the gender outcome, but it ACTUALLY does.

EXAMPLE

A certain experimental drug claims a cure rate of at least 75% for males with prostate cancer. Researchers wish to challenge this claim. What are the potential errors?

$H_0: p \geq 0.75$

$H_a: p < 0.75$ (Claim challenging the population statement/assumption)

Type I Error (α)

A cancer patient believes the cure rate is less than 75% when, in fact, it is 75% or higher.

Type II Error (β)

A cancer patient believes that the drug has a 75% cure rate when, in fact, it is less.

ASSUMPTIONS

A hypothesis test of a single population mean with a known σ. (z-test)

a simple random sample is taken from the population

the population is normally distributed OR the sample size is large

a normal distribution is used

A hypothesis test of a single population mean with an unknown σ. (t-test)

a simple random sample is taken from the population

the population is normally distributed OR the sample size is large

the population σ is estimated by the sample standard deviation

a Student t-distribution is used

A hypothesis test of a single population proportion

a simple random sample is taken from the population

the conditions for a binomial distribution must be met

the shape of the binomial distribution must be similar to a normal distribution

a binomial distribution is used

A QUICK REVIEW

- Two hypotheses are required for hypothesis testing:
  - The **Null Hypothesis** (H_0): A statement of no difference between a sample mean/proportion and a population mean/proportion.
  - The **Alternative Hypothesis** (H_a): A claim that is contradictory to H_0.
- There are two types of errors:
- **Type I Error**
  - A true null hypothesis is rejected
  - The probability of this occurring is α. (The "leftover" from the confidence interval)
  - "We think that the status quo (H_0) is false, but it is true."
- **Type II Error**
  - Failure to reject a false null hypothesis
  - The probability of this occurring is β.
  - "We think that the status quo (H_0) is true, but it is true."