Lecture 19
Hypothesis Testing

STAT 8010 Statistical Methods I
October 4, 2019

Whitney Huang
Clemson University

Hypothesis Testing

- **Hypothesis Testing**: A statistical procedure that use sample data to decide between two competing hypotheses about a population characteristic (e.g. μ)

- **Hypotheses**:
  - Null Hypothesis (H_0): An initial claim about a population characteristic
  - Alternative Hypothesis (H_a): The competing claim

- H_0 will be rejected in favor of H_a only if sample evidence strongly suggests that H_0 is false. Otherwise, H_0 will not be rejected. Therefore, the two possible decisions are:
  - Reject H_0 (and go with H_a)
  - Fail to Reject H_0

Courtroom Analogy

- In a criminal trial, we use the rule "innocent until proven guilty"
- Therefore, our hypotheses are:
  - H_0: Innocent
  - H_a: Guilty
- If we have strong evidence that the accused is not innocent, we reject H_0 (innocent) and conclude H_a (guilty)
- If we do not have enough evidence to say that the accused is guilty, we do not say that the accused is "innocent". Instead, we say that the accused is "not guilty"

Hypotheses

Be careful setting up hypotheses:

- A statistical hypothesis test is only capable of demonstrating strong support for the alternative hypothesis H_a (by rejecting the null hypothesis H_0)
- Failing to reject H_0 does not show strong support for the null hypothesis – **only a lack of strong evidence against H_0,** the null hypothesis

The 2×2 Decision Paradigm for Hypothesis Testing

True State	Decision	
	Reject H_0	Fail to reject H_0
H_0 is true	Type I error	Correct
H_0 is false	Correct	Type II error

Errors in Hypothesis Testing

- The probability of a type I error is denoted by α
- The probability of a type II error is denoted by β

CLEMSON UNIVERSITY

Test Statistics

- In a hypothesis test, our "evidence" comes in the form of a test statistic

- A test statistic incorporates a number of aspects of the sample: the **sample size**, the **point estimate**, the **standard deviation**, and the **hypothesized value**

- If we're conducting a hypothesis test about μ (assuming we don't know σ) we would use the following test statistic:

$$t^* = \frac{\bar{X} - \mu_0}{\frac{s}{\sqrt{n}}}$$

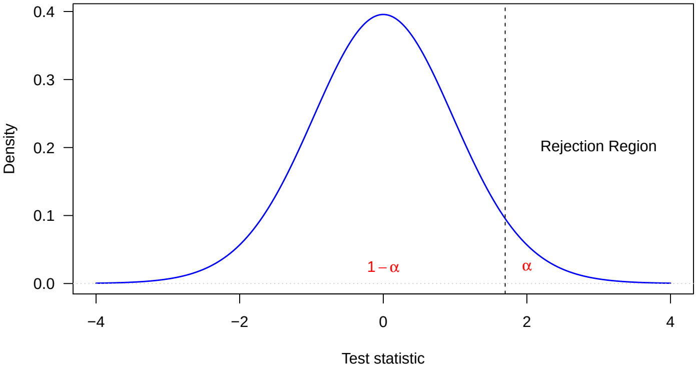
If $\mu = \mu_0$, we have $t^* \sim t(df = n - 1)$

Decision-Making: Rejection Region and P-Value Methods

- Decision based on t^*, H_a, and α, the **significant level**, that is pre-defined by the researcher
- Two approaches:
  - Rejection Region Method: reject H_0 if t^* is in the rejection region, otherwise fail to reject H_0
  - P-Value Method: reject H_0 if P-value is less than α, otherwise fail to reject H_0
- **Question:** How to determine the rejection region and how to compute P-value?

Rejection Region Method

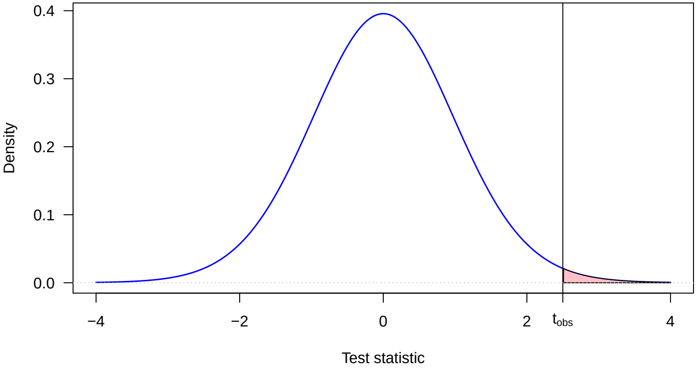
Let $H_0 : \mu = \mu_0$ vs. $H_a : \mu > \mu_0$ and $\alpha = 0.05$



Under the H_0, the test statistic $t^* = \frac{\bar{X} - \mu_0}{\frac{s}{\sqrt{n}}} \sim t(df = n - 1)$. The cutoff of the rejection region (=$t(0.05, n-1)$) can be found from a t-table

P-Value Method

Let $H_0 : \mu = \mu_0$ vs. $H_a : \mu > \mu_0$



P-value: the probability of getting a test statistic that is at least as extreme as the one we actually observed **if the null hypothesis is true** $\Rightarrow \mathbb{P}(t^* \geq t_{obs})$

Draw a Conclusion

Use the following "generic" conclusion:

> "We (do/do not) have enough statistical evidence to conclude that (H_a in words) at α% significant level."

- Reject $H_0 \Leftrightarrow$ do
- Fail to reject $H_0 \Leftrightarrow$ do not

Example (taken from The Cartoon Guide To Statistics)

New Age Granola Inc claims that average weight of its cereal boxes is 16 oz. The Genuine Grocery Corporation will send back a shipment if the average weight is any less.

Suppose Genuine Grocery Corporation takes a random sample of 49 boxes, weight each one, and compute the sample mean $\bar{X} = 15.90$ oz and sample standard deviation $s = 0.35$ oz.

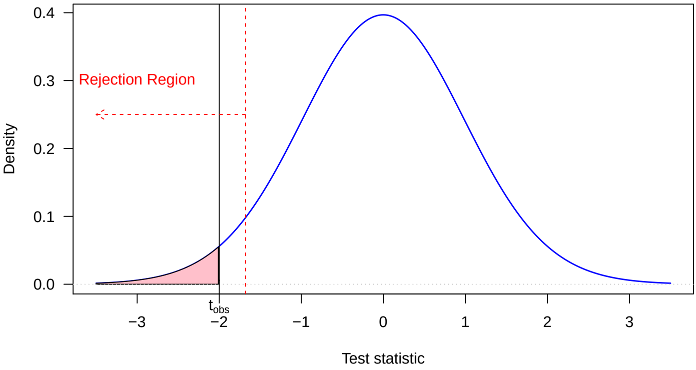
Perform a hypothesis test at 0.05 significant level to determine if they would reject H_0, and therefore, this shipment

Cereal Weight Example Cont'd

CLEMSON UNIVERSITY

1. $H_0 : \mu = 16$ vs. $H_a : \mu < 16$
2. Test Statistic: $t_{obs} = \frac{15.9-16}{0.35/\sqrt{49}} = -2$
3. Rejection Region Method: $-t(0.05, 48) = -1.68 \Rightarrow$ Rejection Region is $(-\infty, -1.68]$. Since t_{obs} is in rejection region, we reject H_0
4. P-Value Method: $\mathbb{P}(t^* \leq -2) = 0.0256 < \alpha = 0.05 \Rightarrow$ reject H_0
5. Draw a Conclusion: We do have enough statistical evidence to conclude that the average weight is less than 16 oz at 0.05% significant level

Cereal Weight Example Cont'd

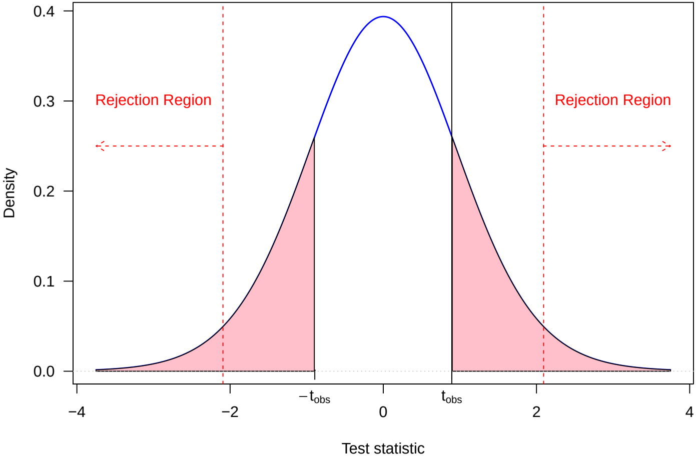


Example

A series of blood tests were run on a particular patient over five days. It is of interest to determine if the mean blood protein for this patient differs from 7.25, the value for healthy adults. Suppose the sample mean (n=20) is 7.35 and sample standard deviation is 0.5. Perform a hypothesis test using significance level of 0.05

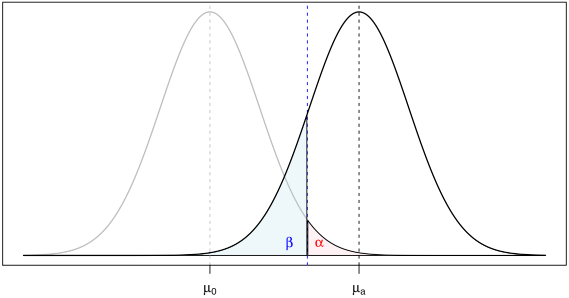
1. $H_0 : \mu = 7.25$ vs. $H_a : \mu \neq 7.25$
2. $t_{obs} = \frac{7.35-7.25}{0.5/\sqrt{20}} = 0.8944$
3. P-value: $2 \times \mathbb{P}(t^* \geq 0.8944) = 0.3823 > 0.05$
4. We do not have enough statistical evidence to conclude that the mean blood protein is different from 7.25 at 5% significant level

Example Cont'd



Type I & Type II Errors

- Type I error: $\mathbb{P}(\text{Reject } H_0 | H_0 \text{ is true}) = \alpha$
- Type II error: $\mathbb{P}(\text{Fail to reject } H_0 | H_0 \text{ is false}) = \beta$



$\alpha \downarrow \beta \uparrow$ and vice versa