Lecture 7
Hypothesis Testing: An Introduction

Topics

- Steps in a hypothesis test.
- Large sample tests for μ: two-tail.
- Large sample tests for μ: one-tail.
- Type I and Type II errors.

Section 7.4

Steps in a Hypothesis Test

Definitions

- In statistics, a *hypothesis* is an idea, an assumption, or a theory about the characteristics of one or more variables in one or more populations.
- A *hypothesis test* is a statistical procedure that involves formulating a hypothesis and using sample data to decide on the validity of the hypothesis.

Definitions (cont.)

- The *null hypothesis* is a statement about a parameter of the population. It is labeled H_0. The *alternative hypothesis* is a statement about a parameter of the population that is opposite to the null hypothesis. It is labeled H_A.
- A *test statistic* is a number that captures the information in the sample data. It is used to decide between the null and alternative hypotheses.

Definitions (cont.)

- The *significance level*, α, is the maximum probability tolerated for rejecting a true null hypothesis.
- The *p value* is the probability of a more extreme departure from the null hypothesis than the observed data.

Definitions (cont.)

- The *rejection region* is the range of values of the test statistic that will lead us to reject the null hypothesis. It is defined by the critical value. The area of the rejection region is α, the significance level.

Goals

- To understand the difference between estimation and testing.
- To learn the 5 steps involved in a hypothesis test.

Estimation Versus Testing

- Estimation is the process of providing a numerical value (point) or values (interval) for a population parameter based on information in a sample.
- Testing is a procedure for assessing whether sample data are consistent with statements (hypotheses) made about the population.

Steps

- **Step 1**: Set up the null and alternative hypotheses.
- **Step 2**: Define the test procedure (includes selecting a test statistic, an α level and a rejection region).
- **Step 3**: Collect the data and calculate the test statistic and the *p* value.
- **Step 4**: Decide whether to reject the null hypothesis.

Steps (cont.)

- **Step 5**: Interpret the results in the context of the problem.

Section 7.6

Large-Sample Test of the Mean: Two-Tail Tests

Definitions

- A *large-sample test of the mean* is conducted when the characteristic of interest is the population mean, μ, and either of these situations exists:
  - the population standard deviation, σ, is known (regardless of sample size)
  - the population standard deviation, σ, is unknown but $n \geq 30$

Definitions (cont.)

- A *two-tail test of the population mean* has these null and alternative hypotheses:
  - H_0: μ = [specified number]
  - H_A: μ ≠ [specified number]

Goal

- To apply the 5 steps of a hypothesis test to a two-tail test of a population mean, μ, based on information in a large sample ($n \geq 30$).

Example – Humerus Bones

- Humerus bones from the same species of animal tend to have approximately the same length-to-width ratios.
- For a particular species (Species A) the mean ratio is 8.5.
- Suppose 41 fossils of humerus bones of an unknown species were unearthed at an archeological site.

Example – Humerus Bones

- Suppose the length-to-width ratios of the 41 fossils were calculated.
- It is believed (hypothesized) that this site was not populated by Species A.

Example – Humerus Bones

- **Step 1**: Set up the null and alternative hypotheses.
- Let μ = the mean ratio for the site.
- H_0: was populated by Species A
- $H_0: \mu = 8.5$
- H_A: was not populated by Species A
- $H_A: \mu \neq 8.5$

Example – Humerus Bones

- **Step 2**: Define the test procedure.

- Consider as our test statistic $$Z = \frac{\bar{X} - \mu_0}{s/\sqrt{n}}$$ where $\mu_0 = 8.5$
- This measures the distance of between the mean ratio of the sample and the mean for Species A.
- We should reject H_0 for large values of $|Z|$.

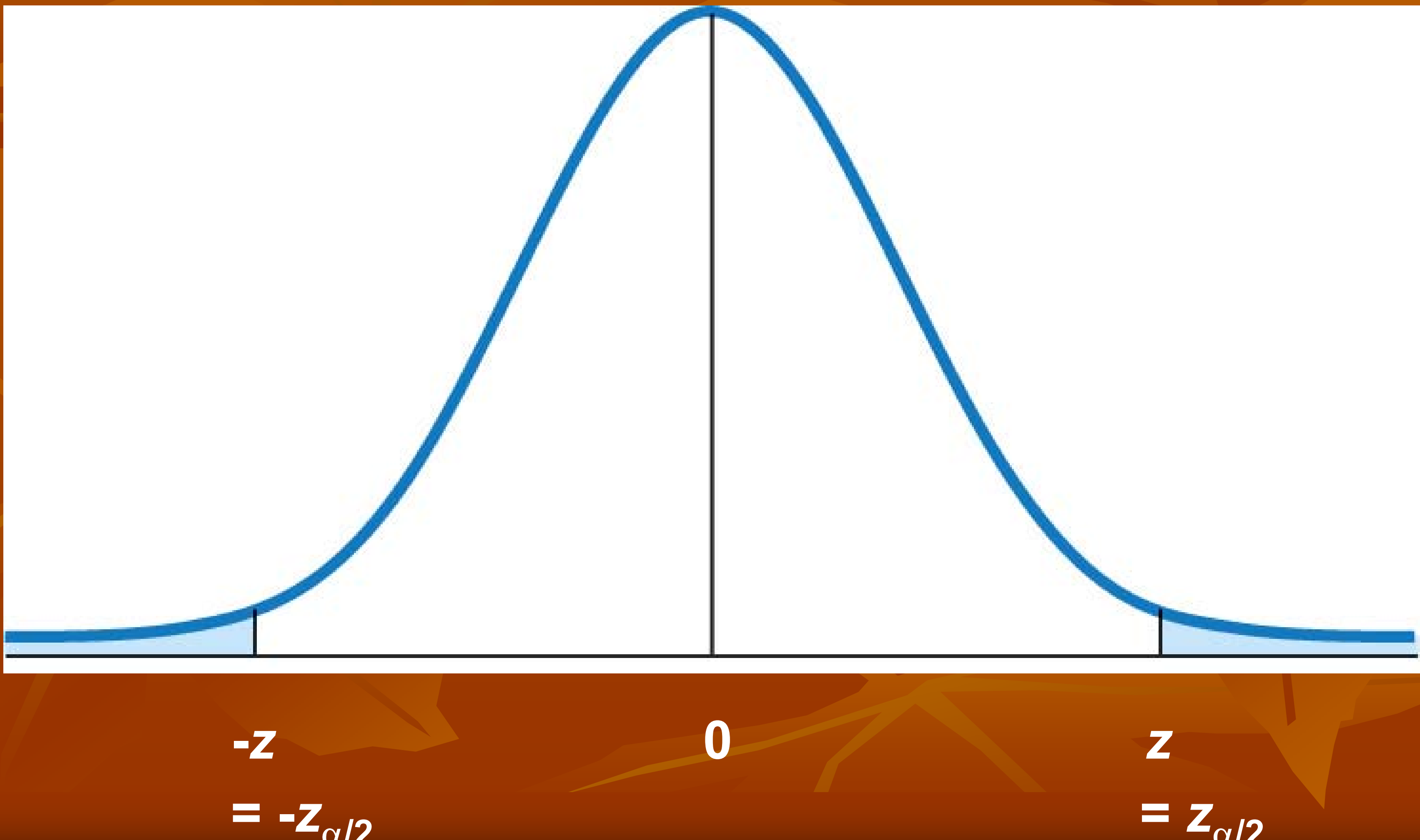
Example – Humerus Bones

- Consider as our rejection region $|Z| > z$.
- This means we reject H_0 if $Z > z$ or $Z < -z$.
- We need to choose z so that

$$\alpha = P(\text{reject a true } H_0)$$
$$= P(Z > z) + P(Z < -z), \text{ assuming } \mu = 8.5$$

- If H_0 is true then $Z \sim N(0,1)$.

Example – Humerus Bones



Example – Humerus Bones

- So we reject H_0 if $|Z| > z_{\alpha/2}$
- Equivalently, reject H_0 if $Z > z_{\alpha/2}$ or $Z < -z_{\alpha/2}$
- Let's choose $\alpha = 0.05$
- $z_{\alpha/2} = \mathrm{z}_{.025} = 1.96$
- Reject H_0 if $Z > 1.96$ or $Z < -1.96$

Example – Humerus Bones

- **Step 3**: Collect the data and calculate the test statistic and the *p* value.
- For the *n* = 41 ratios we have $\bar{X} = 9.258$ and *s* = 1.404
- $$Z = \frac{9.258 - 8.5}{1.404/\sqrt{41}} = 3.46$$

Example – Humerus Bones

- *p* value $= P(|Z| > |3.46|)$
  $= P(Z > 3.46) + P(Z < -3.46)$
  $= 2P(Z < -3.46)$
  $= 2(.0003)$
  $= 0.0006$

Example – Humerus Bones

- **Step 4**: Decide whether to reject the null hypothesis.
- $Z = 3.46$
- Rejection region: $Z > 1.96$ or $Z < -1.96$
- Reject $H_0: \mu = 8.5$
- Accept $H_A: \mu \neq 8.5$

Example – Humerus Bones

- **Step 5**: Interpret the results in the context of the problem.
- The fossils found at the site are not consistent with those that would come from Species A.
- Hence, the site was not populated by species A.

p Value Versus α

- The *p* value is the smallest value of α for which we could reject H_0.
- If *p* value < α then reject H_0.
- If *p* value ≥ α then do not reject H_0.
- A *p* value can be used in place of a rejection region.

Example – Humerus Bones

- $\alpha = 0.05$
- p value = 0.0006
- p value $< \alpha$
- Reject H_0: $\mu = 8.5$

Section 7.7

What Error Could You Be Making?

Definitions

- A *Type I error* is made when we reject the null hypothesis and the null hypothesis is actually true.
- A *Type II error* is made when we fail to reject the null hypothesis and the null hypothesis is actually false.
- The probability of making a Type I error is labeled α.
- The probability of making a Type II error is labeled β.

Goal

- To learn from our mistakes.

Truth versus Decision

	Truth	
Decision	H_0 true	H_A true
Believe H_0	Right choice	Type II error
Believe H_A	Type I error	Right choice

Example – Humerus Bones

- H_0: site <u>was</u> populated by Species A
- H_0: $\mu = 8.5$
- H_A: site <u>was</u> <u>not</u> populated by Species A
- H_A: $\mu \neq 8.5$
- Type I error = conclude the site <u>was</u> <u>not</u> populated by Species A when it actually <u>was</u>.

Example – Humerus Bones

- Type II error = conclude the site <u>was</u> populated by Species A when it actually <u>was</u> <u>not</u>.

Choosing α

- So far our hypothesis tests allow us to choose α.
- In many (but not all) applications a Type I error is more costly than a Type II error.
- As α decreases, β increases.
- As β decreases, α increases.
- Whichever error is the most costly is the probability of which you need to minimize.

Example – Humerus Bones

- Suppose it would be most embarrassing to claim that the site was not populated by Species A when it actually was.
- Make α small ($\alpha = 0.01$).
- Suppose it would cost you your academic career to say that the site was populated by species A when it actually was not.
- Make β small ($\alpha = 0.10$).

Reject H_0 Versus Don't Reject H_0

- Clearly, if we reject H_0 then we accept H_A.
- However, if we do not reject H_0 does this imply that we accept H_0?
- No. At least not in this class.
- Type I error = reject a true H_0
- P(Type I error) = α (known)
- Type II error = accept a false H_0
- P(Type II error) = β (unknown, for now)

Reject H_0 Versus Don't Reject H_0

- Since we know α we will allow ourselves the risk of rejecting H_0 when H_0 is true.
- However, since we do not know β we will not risk accepting H_0 when H_0 is false.
- If we do not reject H_0 we will conclude that there is insufficient evidence in the sample to say H_0 is false.

Section 7.8

Which Theory Should Go into the Null Hypothesis?

Definitions

- A *two-tail test* of the population mean has these null and alternative hypotheses:
  - H_0: $\mu = \mu_0$
  - H_A: $\mu \neq \mu_0$
- A *lower-tail test* of the population mean has these null and alternative hypotheses:
  - H_0: $\mu \geq \mu_0$
  - H_A: $\mu < \mu_0$

Definitions (cont.)

- An *upper-tail test* of the population mean has these null and alternative hypotheses:
  - H_0: $\mu \le \mu_0$
  - H_A: $\mu > \mu_0$

Two-Tail Tests

- H_0: $\mu = \mu_0$
- H_A: $\mu \neq \mu_0$
- H_A: $\mu > \mu_0$ or $\mu < \mu_0$
- In these cases we wish to show that the mean is not some specified value (research hypothesis).
- With this end, we assume the complement to be true (H_0) until we have sufficient evidence to reject H_0 (accept H_A).

Example – Humerus Bone

- μ = mean ratio for the site
- H_A: $\mu \neq 8.5$
- H_A: $\mu < 8.5$ or $\mu > 8.5$
- H_0: $\mu = 8.5$

Lower-Tail Tests

- H_0: $\mu \geq \mu_0$
- H_A: $\mu < \mu_0$
- In these cases we usually wish to show that the mean is less than some specified value (research hypothesis).
- With this end, we assume the complement to be true (H_0) until we have sufficient evidence to reject H_0 (accept H_A).

Example – Vending Machine

- μ = average amount of beverage dispensed into a 12 ounce cup
- You believe that you are being ripped off by this vending machine, i.e., $\mu < 12$ ounces.
- H_A: $\mu < 12$ ounces
- H_0: $\mu \geq 12$ ounces

Upper-Tail Tests

- H_0: $\mu \leq \mu_0$
- H_A: $\mu > \mu_0$
- In these cases we usually wish to show that the mean is greater than some specified value (research hypothesis).
- With this end, we assume the complement to be true (H_0) until we have sufficient evidence to reject H_0 (accept H_A).

Example – Grade Inflation

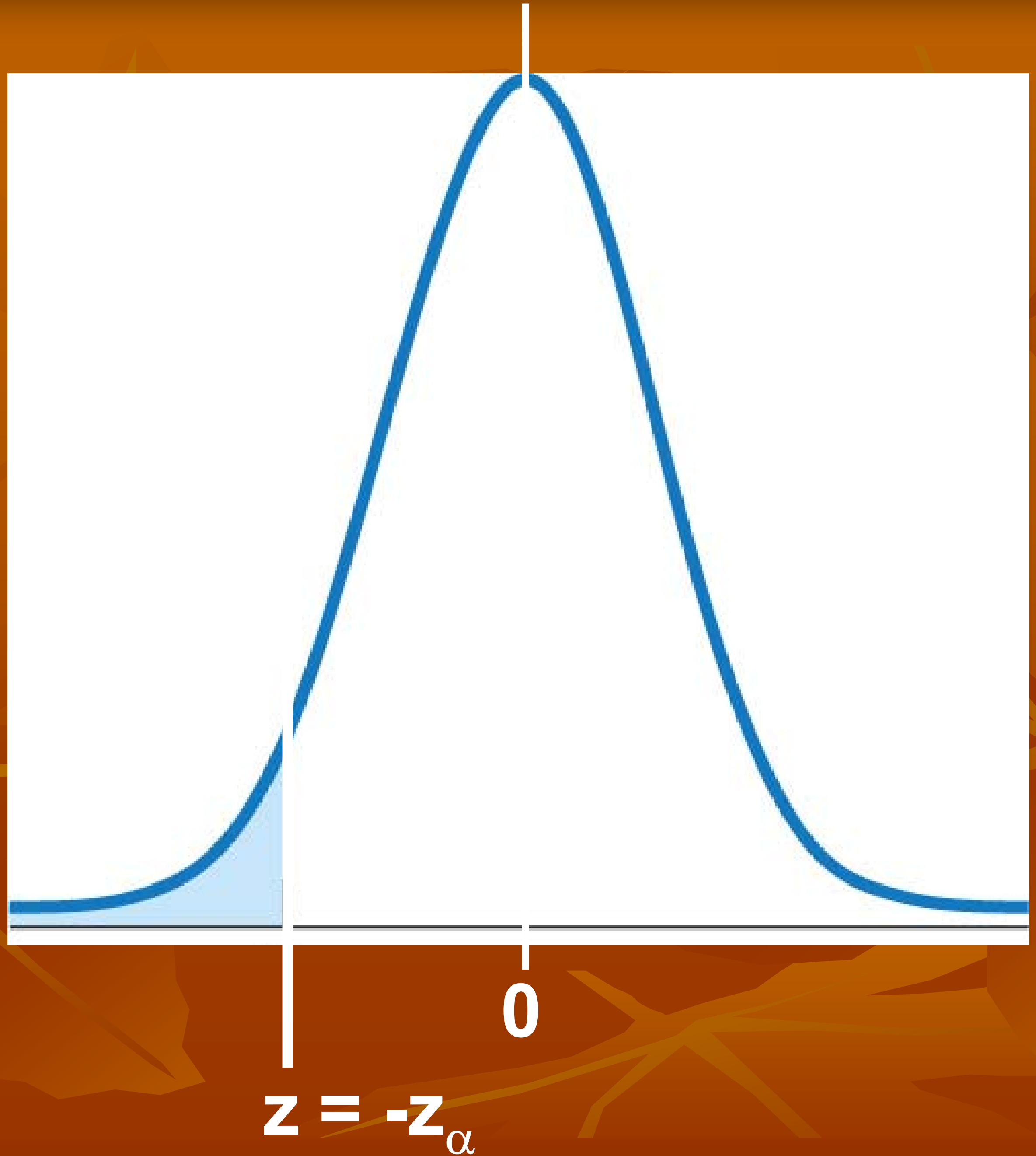
- μ = 2004-05 GPA at UCF
- The historical average is 2.7
- Suppose you wish to show that UCF is currently experiencing grade inflation, i.e., $\mu > 2.7$
- H_A: $\mu > 2.7$
- H_0: $\mu \leq 2.7$

Section 7.9

One-Tail Tests of the Mean: Large Sample

Lower-Tail Test

- H_0: $\mu \geq \mu_0$ versus H_A: $\mu < \mu_0$
- Test statistic: $Z = \dfrac{\bar{X} - \mu_0}{s/\sqrt{n}}$
- Rejection region: $Z < z$
- $\alpha = P(Z < z)$, given H_0 true



0
z = -z_α

Lower-Tail Test

- $z = -z_\alpha$
- Rejection region: $Z < -z_\alpha$
- *p* value = P(Z < observed Z)

Upper-Tail Test

- H_0: $\mu \leq \mu_0$ versus H_A: $\mu > \mu_0$
- Test statistic: $$Z = \frac{\bar{X} - \mu_0}{s/\sqrt{n}}$$
- Rejection region: $Z > z$
- $\alpha = P(Z > z)$, given H_0 true

0
z = z_α

Upper-Tail Test

- $z = z_{\alpha}$
- Rejection region: $Z > z_{\alpha}$
- *p* value = P(Z > observed Z)

Example – Grade Inflation

- μ = 2004-05 GPA at UCF
- The historical average is 2.7
- Suppose you wish to show that UCF is currently experiencing grade inflation.
- Suppose that from a sample of 64 students we had an average GPA of 2.85 with a standard deviation of 0.55.

Example – Grade Inflation

- Let the Type I error be 0.05.
- Step 1:
  - H_A: $\mu > 2.7$
  - H_0: $\mu \leq 2.7$
- Step 2:
  - Test statistic: $Z = \dfrac{\bar{X} - \mu_0}{s/\sqrt{n}}$
  - $\alpha = 0.05$
  - Rejection region: $Z > z_{0.05} = 1.645$

Example – Grade Inflation

- Step 3:
  - Test statistic: $$Z = \frac{2.85 - 2.7}{0.55/\sqrt{64}} = 2.18$$
  - *p* value = P(Z > 2.18) = 0.0146
- Step 4:
  - Reject H_0

Example – Grade Inflation

- Step 5:
  - The sample data indicates that we are currently experiencing grade inflation.