Hypothesis Testing

- Hypothesis testing can be used to determine whether a statement about the value of a population parameter should or should not be rejected.
- The null hypothesis, denoted by H_0, is a tentative assumption about a population parameter.
- The alternative hypothesis, denoted by H_a, is the opposite of what is stated in the null hypothesis.
- The hypothesis testing procedure uses data from a sample to test the two competing statements indicated by H_0 and H_a.

Developing Null and Alternative Hypotheses

- It is not always obvious how the null and alternative hypotheses should be formulated.
- Care must be taken to structure the hypotheses appropriately so that the test conclusion provides the information the researcher wants.
- The context of the situation is very important in determining how the hypotheses should be stated.
- In some cases it is easier to identify the alternative hypothesis first. In other cases the null is easier.
- Correct hypothesis formulation will take practice.

Developing Null and Alternative Hypotheses

- Alternative Hypothesis as a Research Hypothesis
  - Many applications of hypothesis testing involve an attempt to gather evidence in support of a research hypothesis.
  - In such cases, it is often best to begin with the alternative hypothesis and make it the conclusion that the researcher hopes to support.
  - The conclusion that the research hypothesis is true is made if the sample data provide sufficient evidence to show that the null hypothesis can be rejected.

Developing Null and Alternative Hypotheses

- Alternative Hypothesis as a Research Hypothesis
  - Example:
    A new teaching method is developed that is believed to be better than the current method.
  - Alternative Hypothesis:
    The new teaching method is better.
  - Null Hypothesis:
    The new method is no better than the old method.

Developing Null and Alternative Hypotheses

- Alternative Hypothesis as a Research Hypothesis
  - Example:
    A new sales force bonus plan is developed in an attempt to increase sales.
  - Alternative Hypothesis:
    The new bonus plan increase sales.
  - Null Hypothesis:
    The new bonus plan does not increase sales.

Developing Null and Alternative Hypotheses

- Alternative Hypothesis as a Research Hypothesis
  - Example:
    A new drug is developed with the goal of lowering blood pressure more than the existing drug.
  - Alternative Hypothesis:
    The new drug lowers blood pressure more than the existing drug.
  - Null Hypothesis:
    The new drug does not lower blood pressure more than the existing drug.

Developing Null and Alternative Hypotheses

- Null Hypothesis as an Assumption to be Challenged
  - We might begin with a belief or assumption that a statement about the value of a population parameter is true.
  - We then using a hypothesis test to challenge the assumption and determine if there is statistical evidence to conclude that the assumption is incorrect.
  - In these situations, it is helpful to develop the null hypothesis first.

Developing Null and Alternative Hypotheses

- Null Hypothesis as an Assumption to be Challenged
  - Example:
    The label on a soft drink bottle states that it contains 67.6 fluid ounces.
  - Null Hypothesis:
    The label is correct. $\mu \geq 67.6$ ounces.
  - Alternative Hypothesis:
    The label is incorrect. $\mu < 67.6$ ounces.

Summary of Forms for Null and Alternative Hypotheses about a Population Mean

- q The equality part of the hypotheses always appears in the null hypothesis.
- In general, a hypothesis test about the value of a population mean μ must take one of the following three forms (where μ_0 is the hypothesized value of the population mean).

$H_0: \mu \geq \mu_0$ $H_a: \mu < \mu_0$	$H_0: \mu \leq \mu_0$ $H_a: \mu > \mu_0$	$H_0: \mu = \mu_0$ $H_a: \mu \neq \mu_0$
One-tailed (lower-tail)	One-tailed (upper-tail)	Two-tailed

Null and Alternative Hypotheses

- Example: Metro EMS
  - A major west coast city provides one of the most comprehensive emergency medical services in the world. Operating in a multiple hospital system with approximately 20 mobile medical units, the service goal is to respond to medical emergencies with a mean time of 12 minutes or less.
  - The director of medical services wants to formulate a hypothesis test that could use a sample of emergency response times to determine whether or not the service goal of 12 minutes or less is being achieved.

Null and Alternative Hypotheses

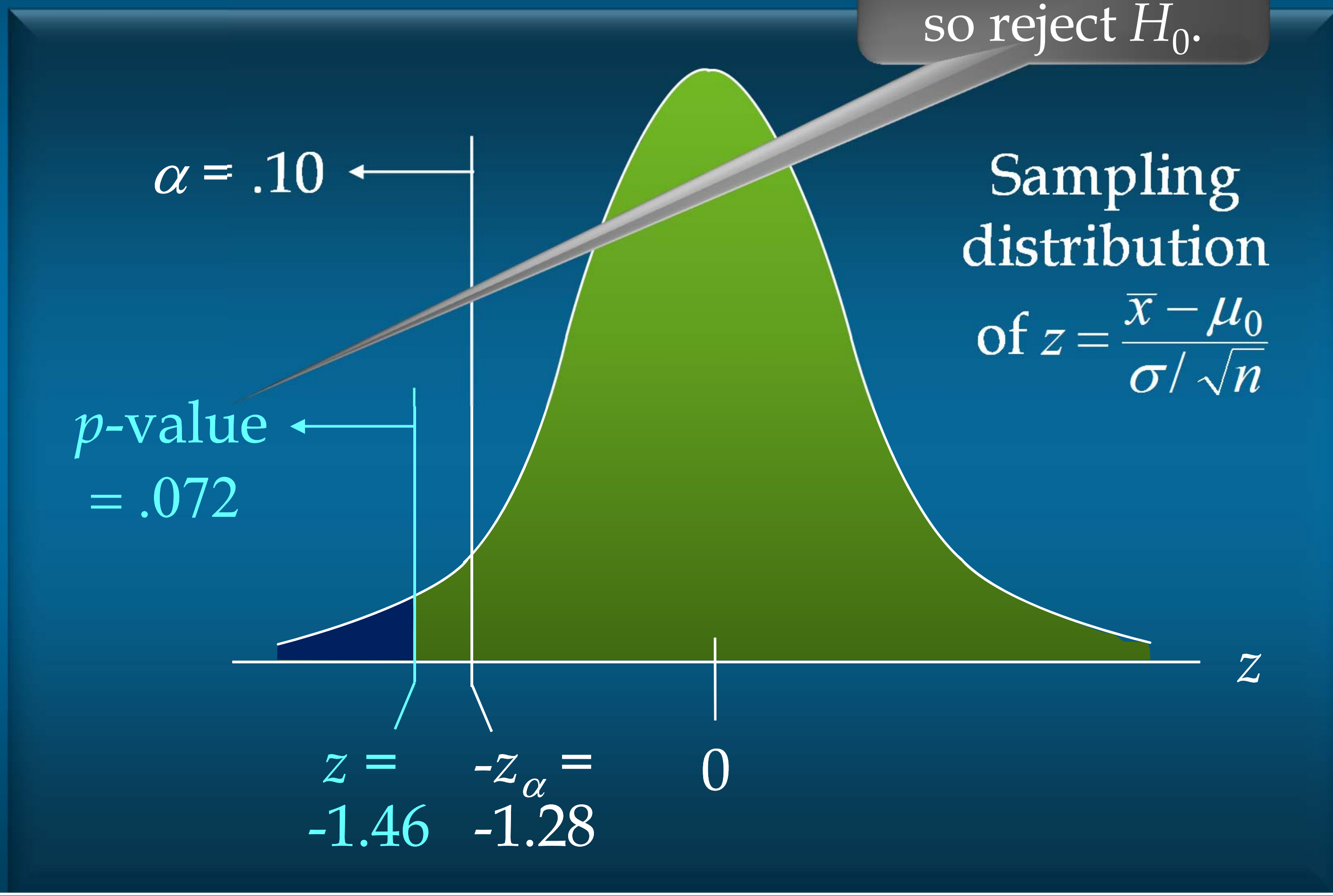
- H_0: $\mu \leq 12$ — The emergency service is meeting the response goal; no follow-up action is necessary.
- H_a: $\mu > 12$ — The emergency service is not meeting the response goal; appropriate follow-up action is necessary.

where: μ = mean response time for the population of medical emergency requests

Lower-Tailed Test About a Population Mean: σ Known

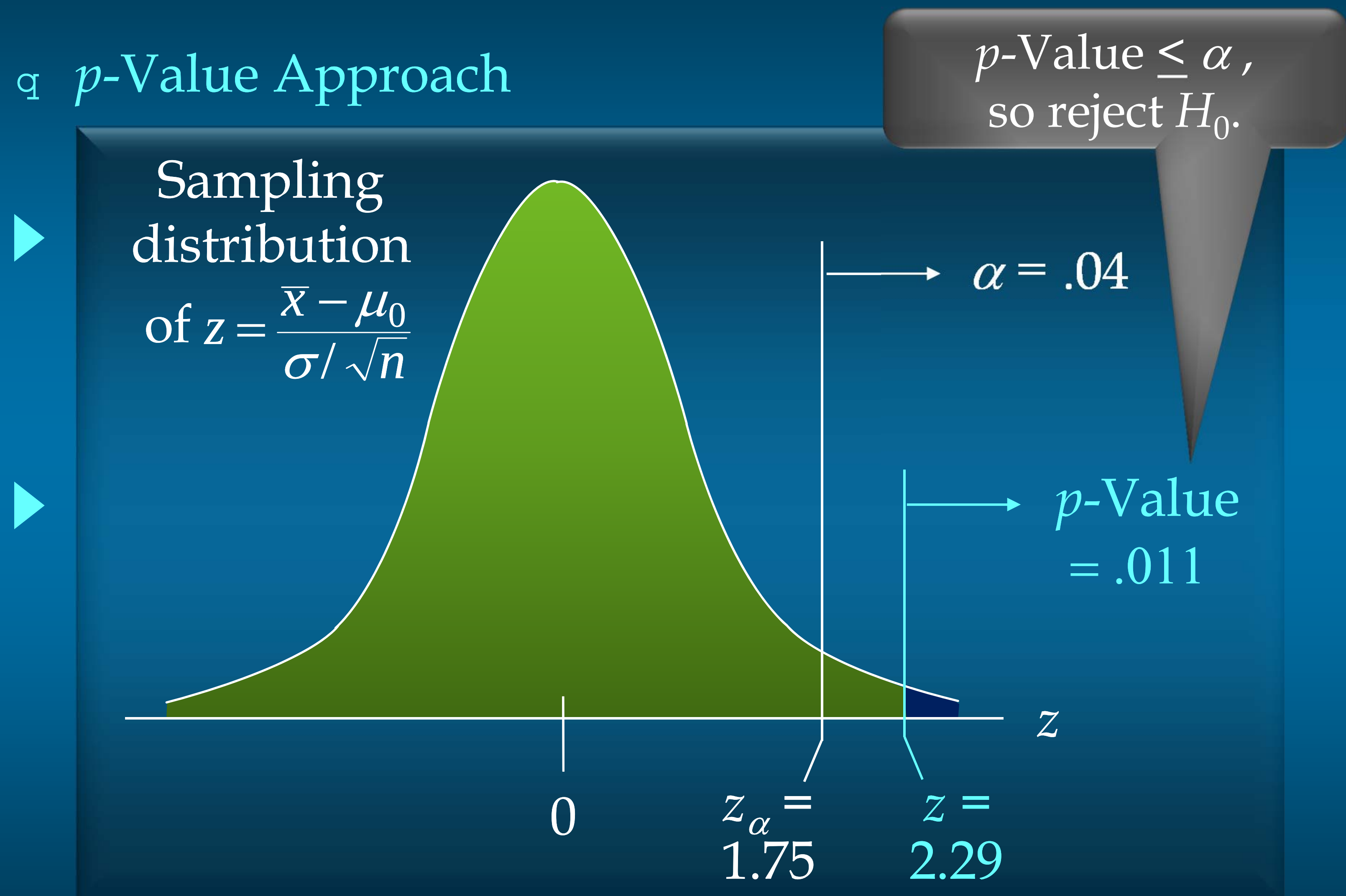
- *p*-Value Approach

p-Value $\leq \alpha$, so reject H_0.



Upper-Tailed Test About a Population Mean: σ Known

- *p*-Value Approach

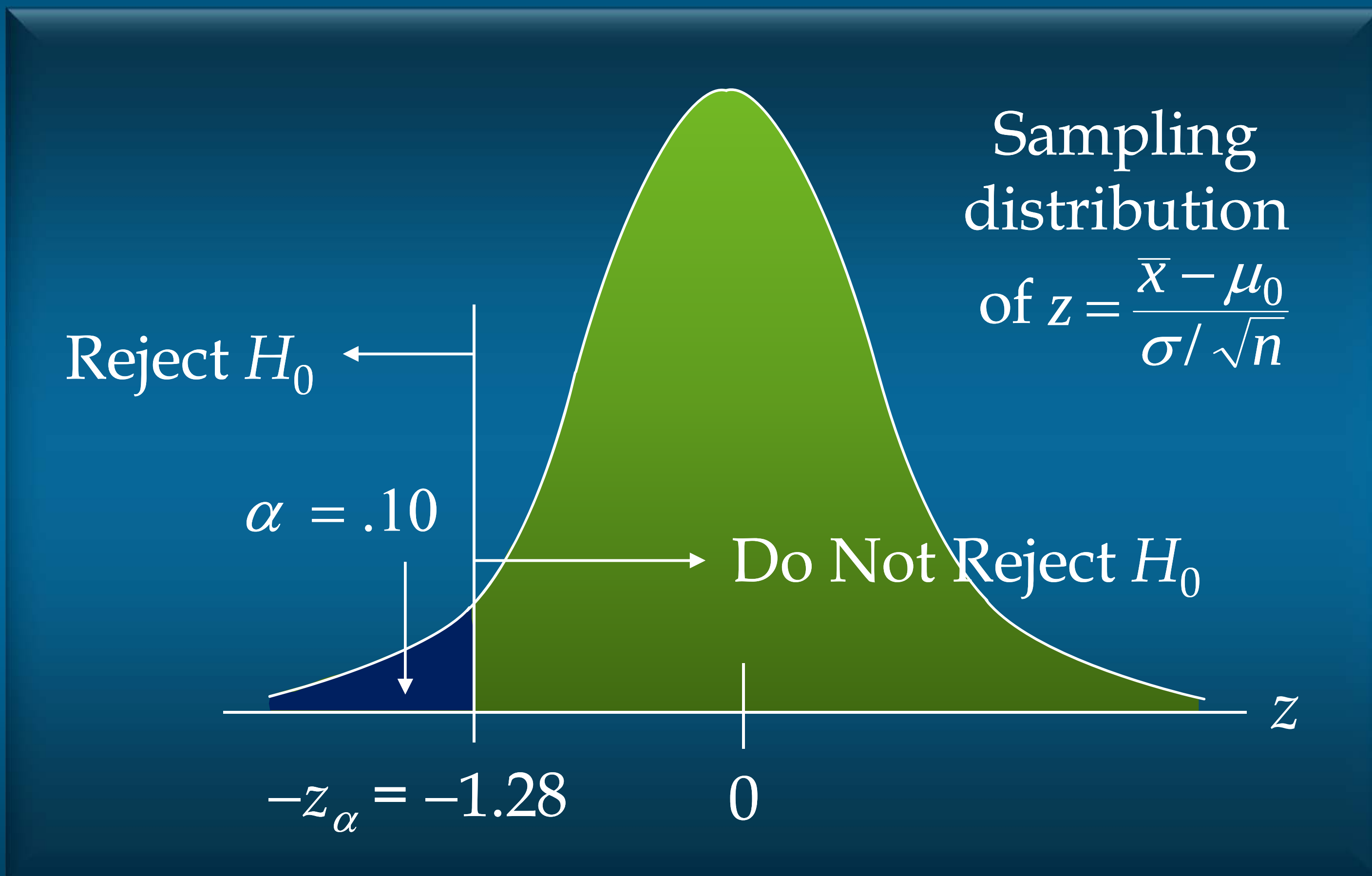


Critical Value Approach to One-Tailed Hypothesis Testing

- The test statistic z has a standard normal probability distribution.
- We can use the standard normal probability distribution table to find the z-value with an area of α in the lower (or upper) tail of the distribution.
- The value of the test statistic that established the boundary of the rejection region is called the <u>critical value</u> for the test.
- The rejection rule is:
  - Lower tail: Reject H_0 if $z \leq -z_\alpha$
  - Upper tail: Reject H_0 if $z \geq z_\alpha$

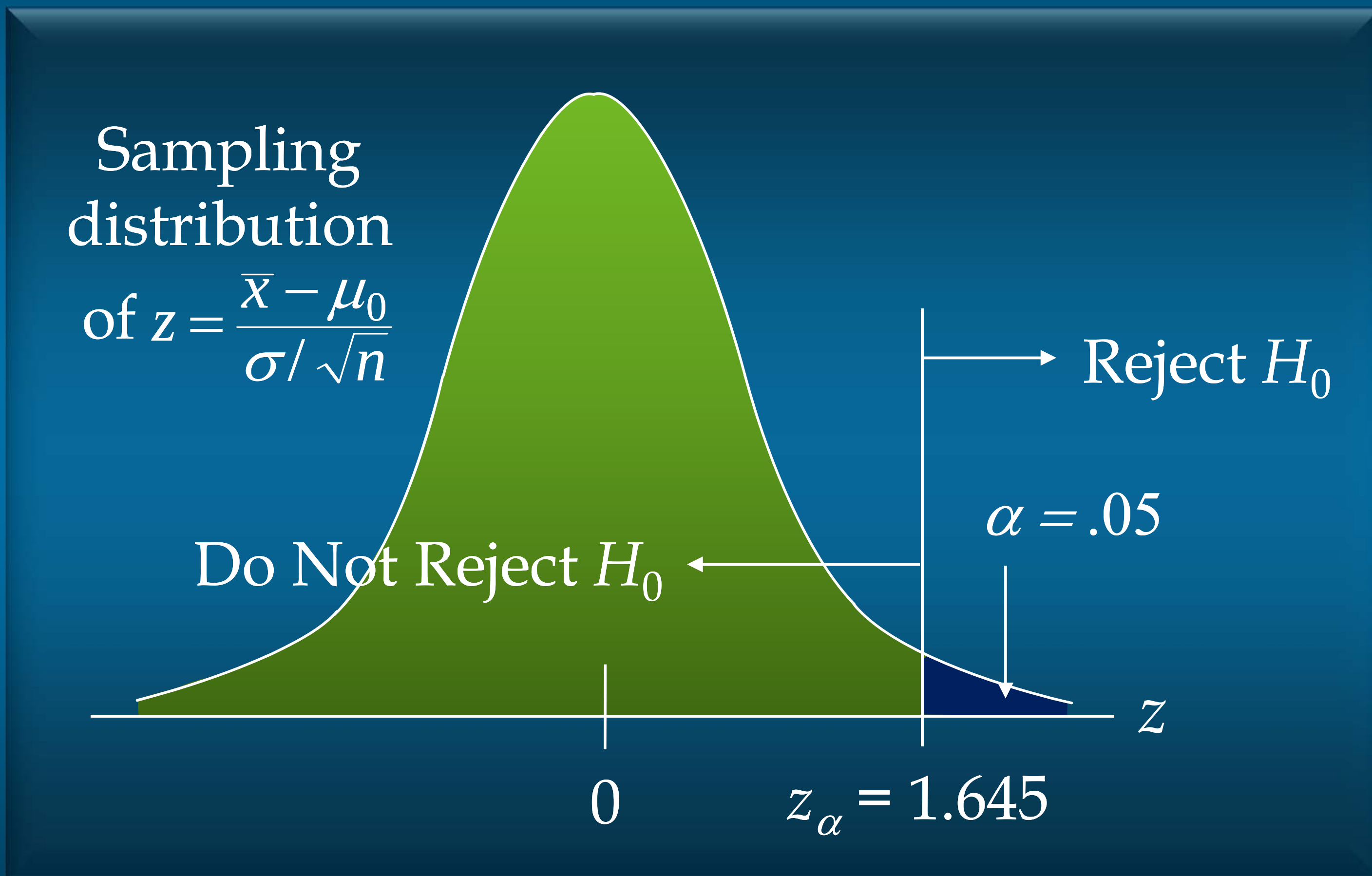
Lower-Tailed Test About a Population Mean: σ Known

- Critical Value Approach



Upper-Tailed Test About a Population Mean: σ Known

- Critical Value Approach



Steps of Hypothesis Testing

- Step 1. Develop the null and alternative hypotheses.
- Step 2. Specify the level of significance α.
- Step 3. Collect the sample data and compute the value of the test statistic.

<u>*p*-Value Approach</u>

- Step 4. Use the value of the test statistic to compute the *p*-value.
- Step 5. Reject H_0 if *p*-value $\leq \alpha$.

Steps of Hypothesis Testing

Critical Value Approach

- Step 4. Use the level of significance to determine the critical value and the rejection rule.
- Step 5. Use the value of the test statistic and the rejection rule to determine whether to reject H_0.

One-Tailed Tests About a Population Mean: σ Known

- Example: Metro EMS
  - The response times for a random sample of 40 medical emergencies were tabulated. The sample mean is 13.25 minutes. The population standard deviation is believed to be 3.2 minutes.
  - The EMS director wants to perform a hypothesis test, with a .05 level of significance, to determine whether the service goal of 12 minutes or less is being achieved.

One-Tailed Tests About a Population Mean: σ Known

- p -Value and Critical Value Approaches

1. Develop the hypotheses.

$$H_0: \mu \le 12$$
$$H_a: \mu > 12$$

2. Specify the level of significance.

α = .05

3. Compute the value of the test statistic.

$$z = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}} = \frac{13.25 - 12}{3.2 / \sqrt{40}} = 2.47$$

One-Tailed Tests About a Population Mean: σ Known

- p –Value Approach

4. Compute the p –value.

For z = 2.47, cumulative probability = .9932.

p–value = 1 – .9932 = .0068

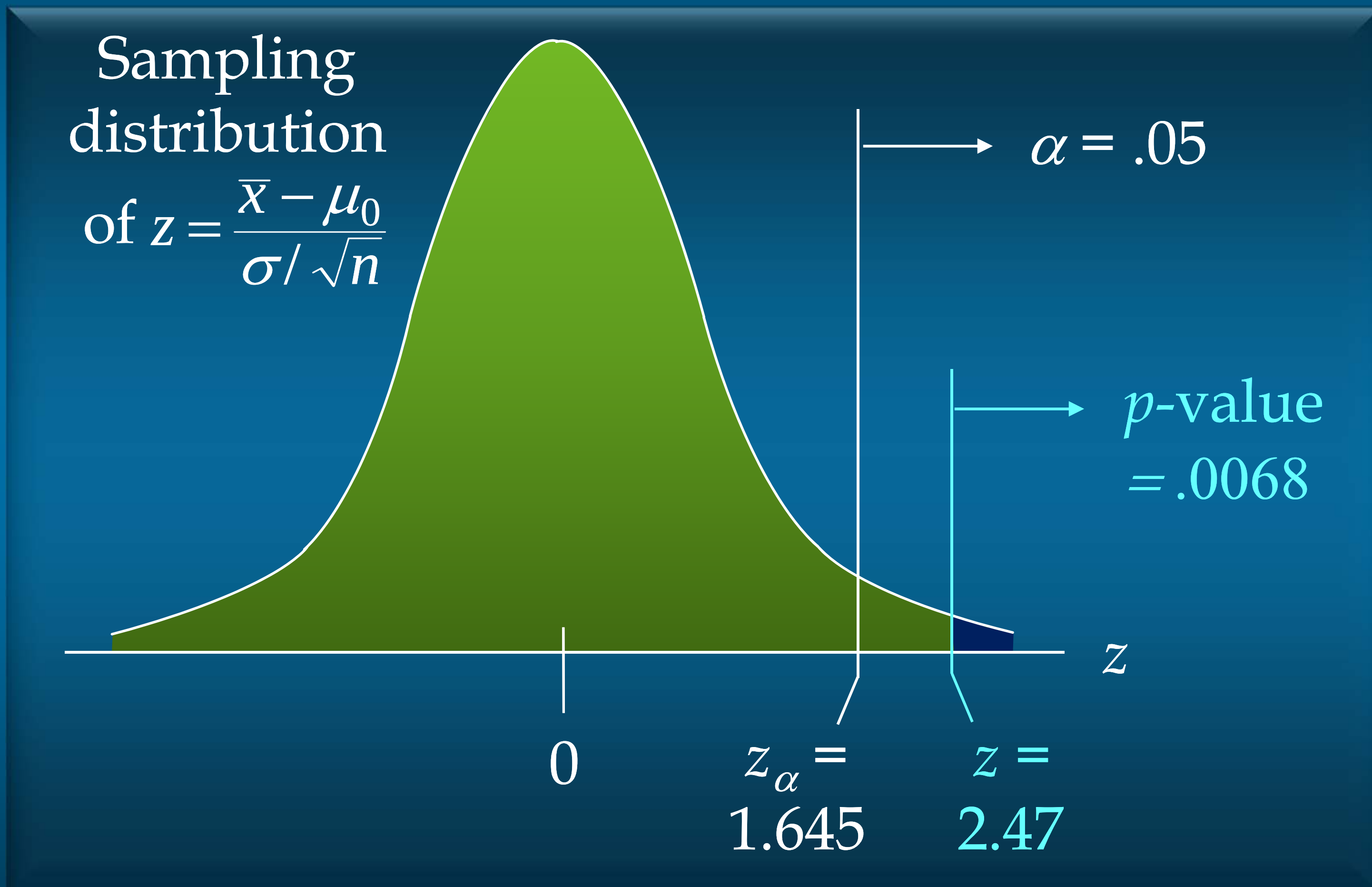
5. Determine whether to reject H_0.

Because p–value = .0068 $\leq$ α = .05, we reject H_0.

There is sufficient statistical evidence to infer that Metro EMS is not meeting the response goal of 12 minutes.

One-Tailed Tests About a Population Mean: σ Known

- p –Value Approach



One-Tailed Tests About a Population Mean: σ Known

- Critical Value Approach

4. Determine the critical value and rejection rule.

For $\alpha = .05$, $z_{.05} = 1.645$

Reject H_0 if $z \geq 1.645$

5. Determine whether to reject H_0.

Because $2.47 \geq 1.645$, we reject H_0.

There is sufficient statistical evidence to infer that Metro EMS is <u>not</u> meeting the response goal of 12 minutes.

p-Value Approach to Two-Tailed Hypothesis Testing

- Compute the p-value using the following three steps:
  1. Compute the value of the test statistic z.
  2. If z is in the upper tail ($z > 0$), compute the probability that z is greater than or equal to the value of the test statistic. If z is in the lower tail ($z < 0$), compute the probability that z is less than or equal to the value of the test statistic.
  3. Double the tail area obtained in step 2 to obtain the p –value.
- The rejection rule:

  Reject H_0 if the p-value $\leq \alpha$.

Critical Value Approach to Two-Tailed Hypothesis Testing

- The critical values will occur in both the lower and upper tails of the standard normal curve.
- Use the standard normal probability distribution table to find $z_{\alpha/2}$ (the z-value with an area of $\alpha/2$ in the upper tail of the distribution).
- The rejection rule is:

$$\text{Reject } H_0 \text{ if } z \le -z_{\alpha/2} \text{ or } z \ge z_{\alpha/2}.$$

Two-Tailed Tests About a Population Mean: σ Known

- Example: Glow Toothpaste

- The production line for Glow toothpaste is designed to fill tubes with a mean weight of 6 oz. Periodically, a sample of 30 tubes will be selected in order to check the filling process.
- Quality assurance procedures call for the continuation of the filling process if the sample results are consistent with the assumption that the mean filling weight for the population of toothpaste tubes is 6 oz.; otherwise the process will be adjusted.

Two-Tailed Tests About a Population Mean: σ Known

- Example: Glow Toothpaste
  - Assume that a sample of 30 toothpaste tubes provides a sample mean of 6.1 oz. The population standard deviation is believed to be 0.2 oz.
  - Perform a hypothesis test, at the .03 level of significance, to help determine whether the filling process should continue operating or be stopped and corrected.

Two-Tailed Tests About a Population Mean: σ Known

- *p* –Value and Critical Value Approaches

1. Determine the hypotheses.

$$H_0: \mu = 6$$
$$H_a: \mu \neq 6$$

2. Specify the level of significance.

α = .03

3. Compute the value of the test statistic.

$$z = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}} = \frac{6.1 - 6}{.2 / \sqrt{30}} = 2.74$$

Two-Tailed Tests About a Population Mean: σ Known

- p –Value Approach

  - 4. Compute the p –value.

    For z = 2.74, cumulative probability = .9969

    p–value = 2(1 – .9969) = .0062

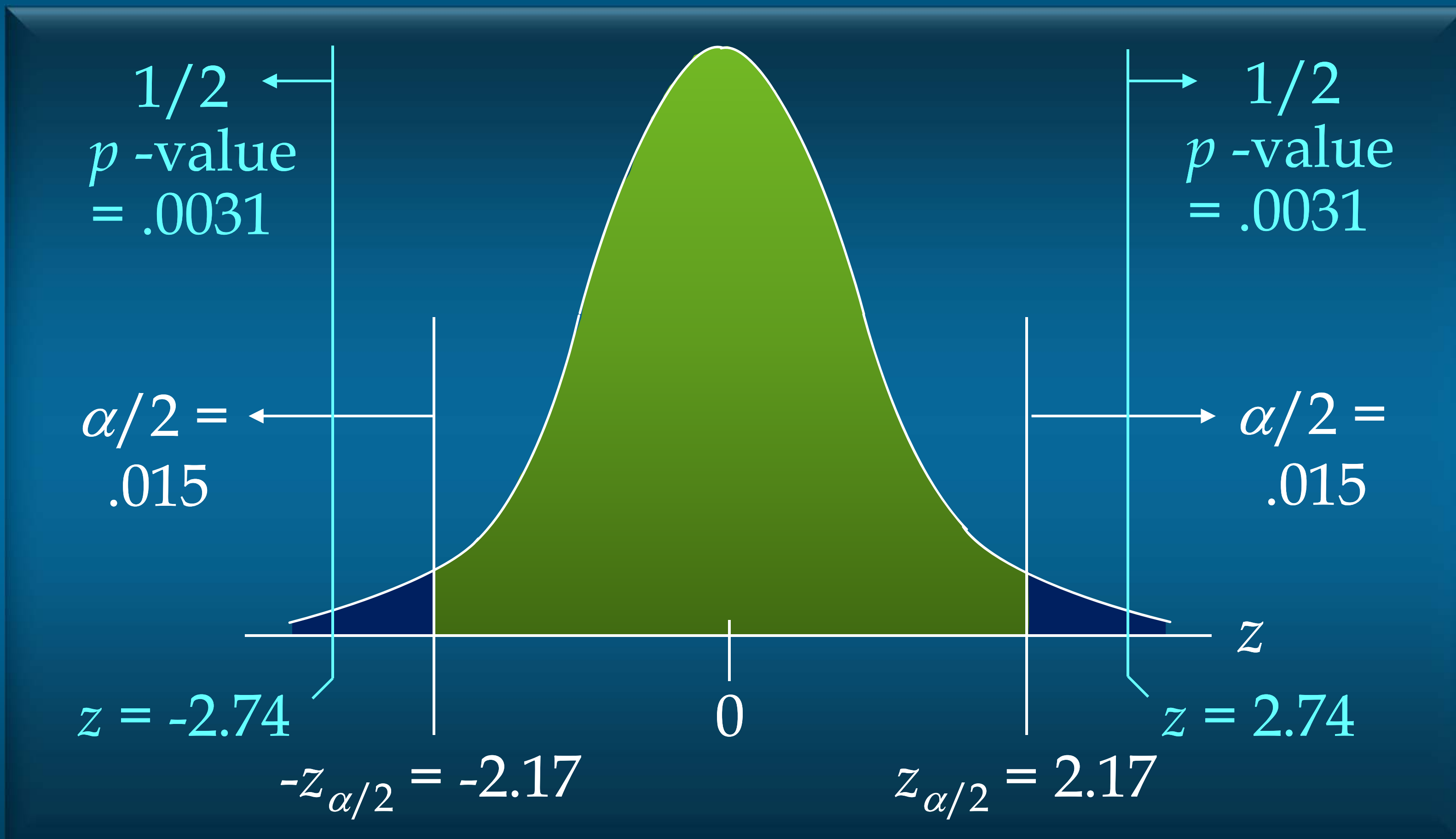
  - 5. Determine whether to reject H_0.

    Because p–value = .0062 $\leq \alpha$ = .03, we reject H_0.

    There is sufficient statistical evidence to infer that the alternative hypothesis is true (i.e. the mean filling weight is not 6 ounces).

Two-Tailed Tests About a Population Mean: σ Known

- *p*-Value Approach



Two-Tailed Tests About a Population Mean: σ Known

- Critical Value Approach

4. Determine the critical value and rejection rule.

For $\alpha/2 = .03/2 = .015$, $z_{.015} = 2.17$

Reject H_0 if $z \leq -2.17$ or $z \geq 2.17$

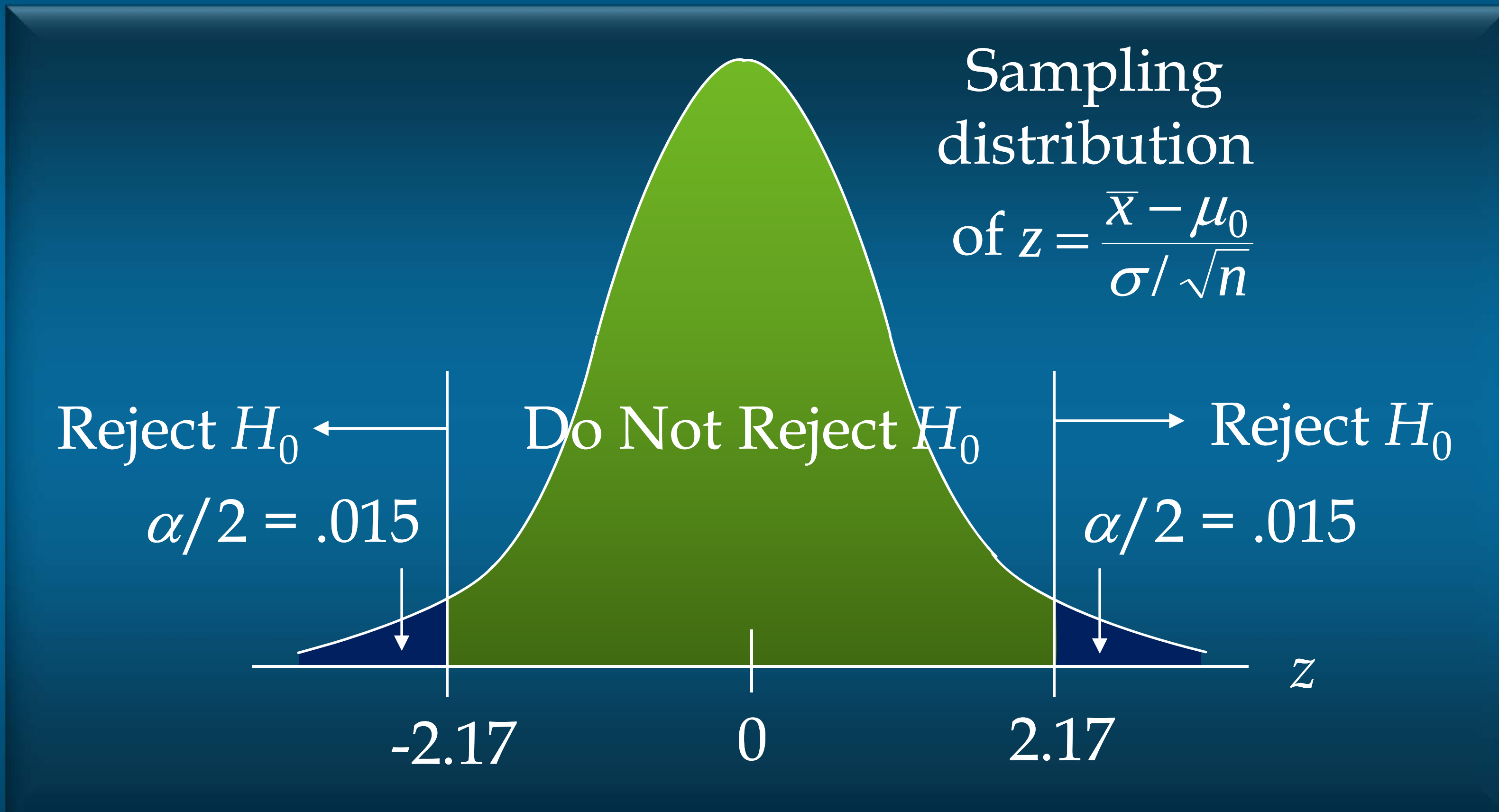
5. Determine whether to reject H_0.

Because $2.74 \geq 2.17$, we reject H_0.

There is sufficient statistical evidence to infer that the alternative hypothesis is true (i.e. the mean filling weight is not 6 ounces).

Two-Tailed Tests About a Population Mean: σ Known

- Critical Value Approach



Tests About a Population Mean: σ Unknown

- Test Statistic

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$$

This test statistic has a t distribution with n - 1 degrees of freedom.

Tests About a Population Mean: σ Unknown

- Rejection Rule: p -Value Approach

Reject H_0 if p –value $\leq \alpha$

- Rejection Rule: Critical Value Approach

$H_0\text{:}\ \ \mu \geq \mu_0$	Reject H_0 if $t \leq -t_\alpha$
$H_0\text{:}\ \ \mu \leq \mu_0$	Reject H_0 if $t \geq t_\alpha$
$H_0\text{:}\ \ \mu = \mu_0$	Reject H_0 if $t \leq -\ t_{\alpha/2}$ or $t \geq t_{\alpha/2}$

p -Values and the *t* Distribution

- The format of the *t* distribution table provided in most statistics textbooks does not have sufficient detail to determine the <u>exact</u> *p*-value for a hypothesis test.
- However, we can still use the *t* distribution table to identify a <u>range</u> for the *p*-value.
- An advantage of computer software packages is that the computer output will provide the *p*-value for the *t* distribution.

Example: Highway Patrol

- One-Tailed Test About a Population Mean: σ Unknown

- A State Highway Patrol periodically samples vehicle speeds at various locations on a particular roadway. The sample of vehicle speeds is used to test the hypothesis H_0: $\mu \leq 65$.

- The locations where H_0 is rejected are deemed the best locations for radar traps. At Location F, a sample of 64 vehicles shows a mean speed of 66.2 mph with a standard deviation of 4.2 mph. Use $\alpha = .05$ to test the hypothesis.

One-Tailed Test About a Population Mean: σ Unknown

- p –Value and Critical Value Approaches

1. Determine the hypotheses.

H_0: $\mu \leq 65$
H_a: $\mu > 65$

2. Specify the level of significance.

$\alpha = .05$

3. Compute the value of the test statistic.

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{66.2 - 65}{4.2/\sqrt{64}} = 2.286$$

One-Tailed Test About a Population Mean: σ Unknown

- *p* –Value Approach

4. Compute the *p* –value.

For $t = 2.286$, the p–value must be less than .025 (for $t = 1.998$) and greater than .01 (for $t = 2.387$).

$.01 < p\text{–value} < .025$

5. Determine whether to reject H_0.

Because p–value $\leq \alpha = .05$, we reject H_0.

We are at least 95% confident that the mean speed of vehicles at Location F is greater than 65 mph.

One-Tailed Test About a Population Mean: σ Unknown

- Critical Value Approach

4. Determine the critical value and rejection rule.

For $\alpha = .05$ and d.f. = 64 – 1 = 63, $t_{.05} = 1.669$

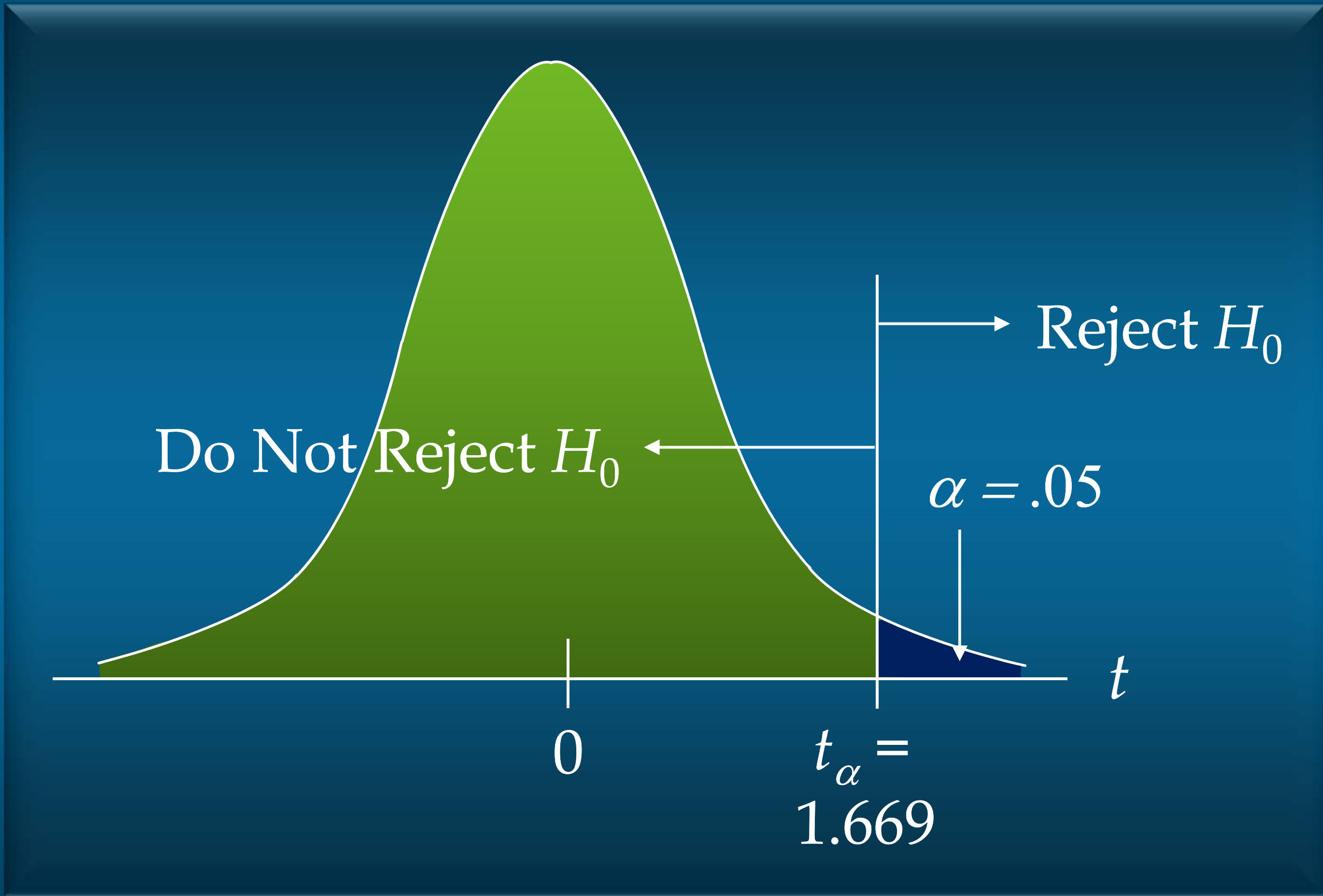
Reject H_0 if $t \geq 1.669$

5. Determine whether to reject H_0.

Because $2.286 \geq 1.669$, we reject H_0.

We are at least 95% confident that the mean speed of vehicles at Location F is greater than 65 mph. Location F is a good candidate for a radar trap.

One-Tailed Test About a Population Mean: σ Unknown



A Summary of Forms for Null and Alternative Hypotheses About a Population Proportion

- q The equality part of the hypotheses always appears in the null hypothesis.
- In general, a hypothesis test about the value of a population proportion p must take one of the following three forms (where p_0 is the hypothesized value of the population proportion).

$H_0: p \geq p_0$ $H_a: p < p_0$	$H_0: p \leq p_0$ $H_a: p > p_0$	$H_0: p = p_0$ $H_a: p \neq p_0$
One-tailed (lower tail)	One-tailed (upper tail)	Two-tailed

Tests About a Population Proportion

- Test Statistic

$$z = \frac{\bar{p} - p_0}{\sigma_{\bar{p}}}$$

where:

$$\sigma_{\bar{p}} = \sqrt{\frac{p_0(1 - p_0)}{n}}$$

assuming $np \geq 5$ and $n(1 - p) \geq 5$

Tests About a Population Proportion

- Rejection Rule: p –Value Approach

  Reject H_0 if p –value $\leq \alpha$

- Rejection Rule: Critical Value Approach

  - $H_0\!: \; p \leq p_0$ Reject H_0 if $z \geq z_\alpha$
  - $H_0\!: \; p \geq p_0$ Reject H_0 if $z \leq -z_\alpha$
  - $H_0\!: \; p = p_0$ Reject H_0 if $z \leq -z_{\alpha/2}$ or $z \geq z_{\alpha/2}$

Two-Tailed Test About a Population Proportion

- Example: National Safety Council (NSC)

  For a Christmas and New Year's week, the National Safety Council estimated that 500 people would be killed and 25,000 injured on the nation's roads. The NSC claimed that 50% of the accidents would be caused by drunk driving.

  A sample of 120 accidents showed that 67 were caused by drunk driving. Use these data to test the NSC's claim with $\alpha = .05$.

Two-Tailed Test About a Population Proportion

- p –Value and Critical Value Approaches

1. Determine the hypotheses.

$$H_0: p = .5$$
$$H_a: p \neq .5$$

2. Specify the level of significance.

$\alpha = .05$

3. Compute the value of the test statistic.

$$\sigma_{\bar{p}} = \sqrt{\frac{p_0(1-p_0)}{n}} = \sqrt{\frac{.5(1-.5)}{120}} = .045644$$

a common error is using $\bar{p}$ in this formula

$$z = \frac{\bar{p}-p_0}{\sigma_{\bar{p}}} = \frac{(67/120)-.5}{.045644} = 1.28$$

Two-Tailed Test About a Population Proportion

- p–Value Approach

4. Compute the p -value.

For z = 1.28, cumulative probability = .8997

p–value = 2(1 – .8997) = .2006

5. Determine whether to reject H_0.

Because p–value = .2006 > α = .05, we cannot reject H_0.

Two-Tailed Test About a Population Proportion

- Critical Value Approach

4. Determine the criticals value and rejection rule.

For $\alpha/2 = .05/2 = .025$, $z_{.025} = 1.96$

Reject H_0 if $z \leq -1.96$ or $z \geq 1.96$

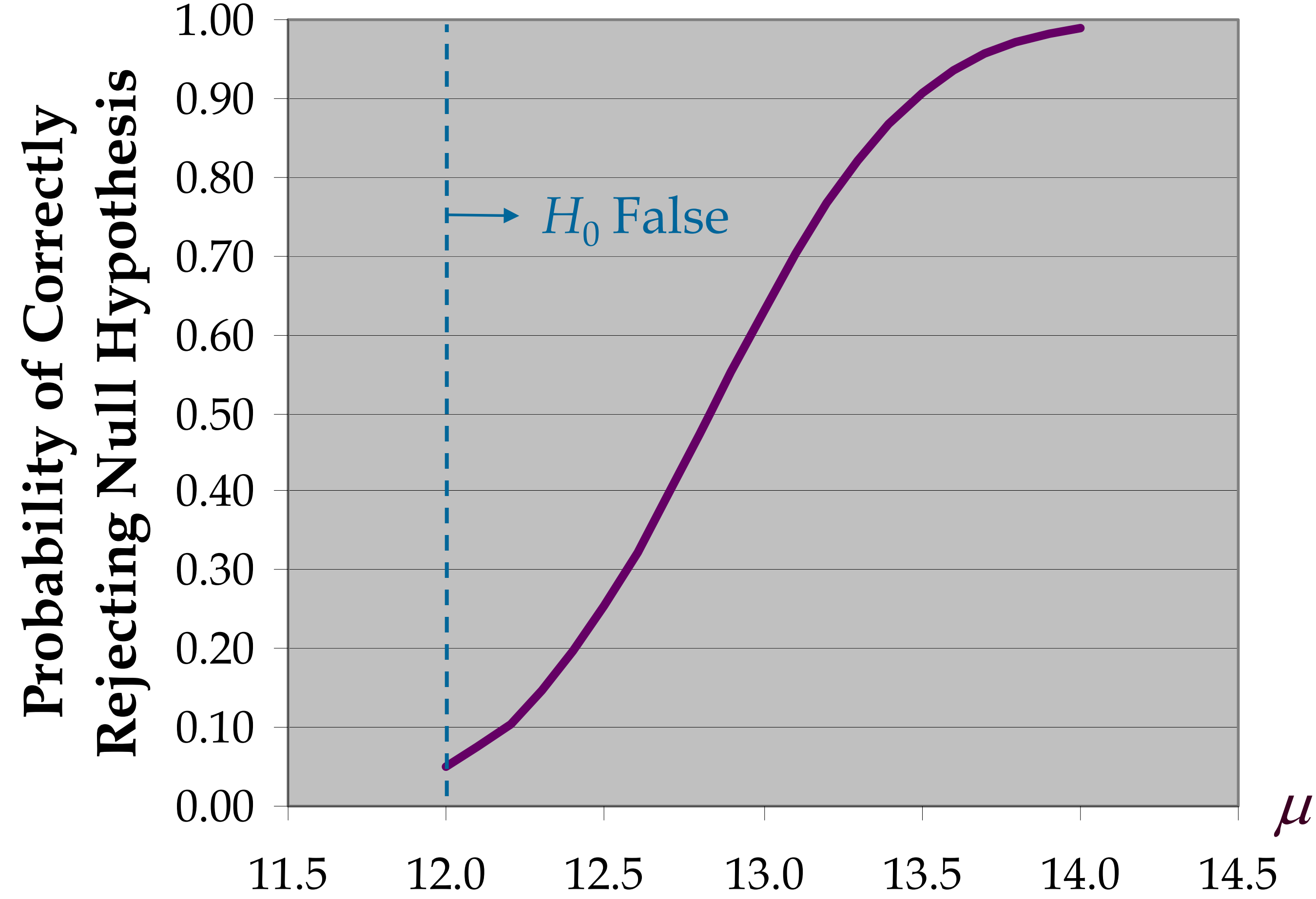
5. Determine whether to reject H_0.

Because 1.278 > -1.96 and < 1.96, we cannot reject H_0.

Power of the Test

- The probability of correctly rejecting H_0 when it is false is called the power of the test.
- For any particular value of μ, the power is $1 - \beta$.
- We can show graphically the power associated with each value of μ; such a graph is called a power curve. (See next slide.)

Power Curve

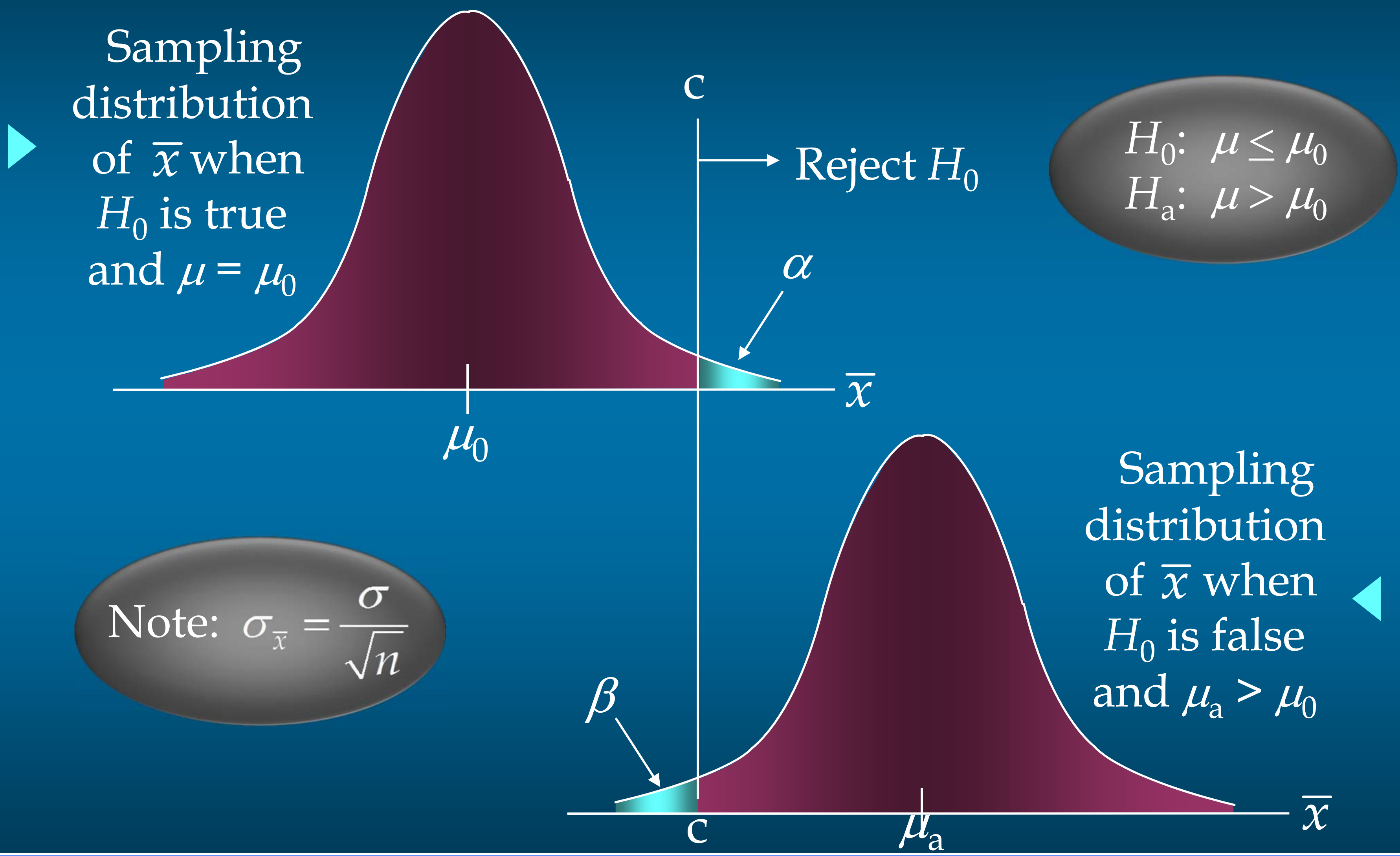


Probability of Correctly Rejecting Null Hypothesis
1.00
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
H_0 False
11.5
12.0
12.5
13.0
13.5
14.0
14.5
μ

Determining the Sample Size for a Hypothesis Test About a Population Mean

- q The specified level of significance determines the probability of making a Type I error.
- By controlling the sample size, the probability of making a Type II error is controlled.

Determining the Sample Size for a Hypothesis Test About a Population Mean



Determining the Sample Size for a Hypothesis Test About a Population Mean

$$n = \frac{(z_\alpha + z_\beta)^2 \sigma^2}{(\mu_0 - \mu_a)^2}$$

where

z_α = z value providing an area of α in the tail
z_β = z value providing an area of β in the tail
σ = population standard deviation
μ_0 = value of the population mean in H_0
μ_a = value of the population mean used for the Type II error

Note: In a two-tailed hypothesis test, use $z_{\alpha/2}$ not z_α

Determining the Sample Size for a Hypothesis Test About a Population Mean

- q Let's assume that the director of medical services makes the following statements about the allowable probabilities for the Type I and Type II errors:
  - •If the mean response time is $\mu = 12$ minutes, I am willing to risk an $\alpha = .05$ probability of rejecting H_0.
  - •If the mean response time is 0.75 minutes over the specification ($\mu = 12.75$), I am willing to risk a $\beta = .10$ probability of not rejecting H_0.

Determining the Sample Size for a Hypothesis Test About a Population Mean

$\alpha = .05, \ \beta = .10$

$z_\alpha = 1.645, \ z_\beta = 1.28$

$\mu_0 = 12, \ \mu_a = 12.75$

$\sigma = 3.2$

$$n = \frac{(z_\alpha + z_\beta)^2 \sigma^2}{(\mu_0 - \mu_a)^2} = \frac{(1.645 + 1.28)^2 (3.2)^2}{(12 - 12.75)^2} = 155.75 \approx 156$$

Relationship Among α, β, and n

- Once two of the three values are known, the other can be computed.
- For a given level of significance α, increasing the sample size n will reduce β.
- For a given sample size n, decreasing α will increase β, whereas increasing α will decrease b.