Hypothesis Testing

Part 2

STEP:

1) Two hypotheses: H_0 & H_1

2) ASSUME H_0 is TRUE

3) GOAL: determine if there is enough evidence to infer that H_1 is TRUE

4) Two possible decisions:

Reject H_0 in favor of H_1

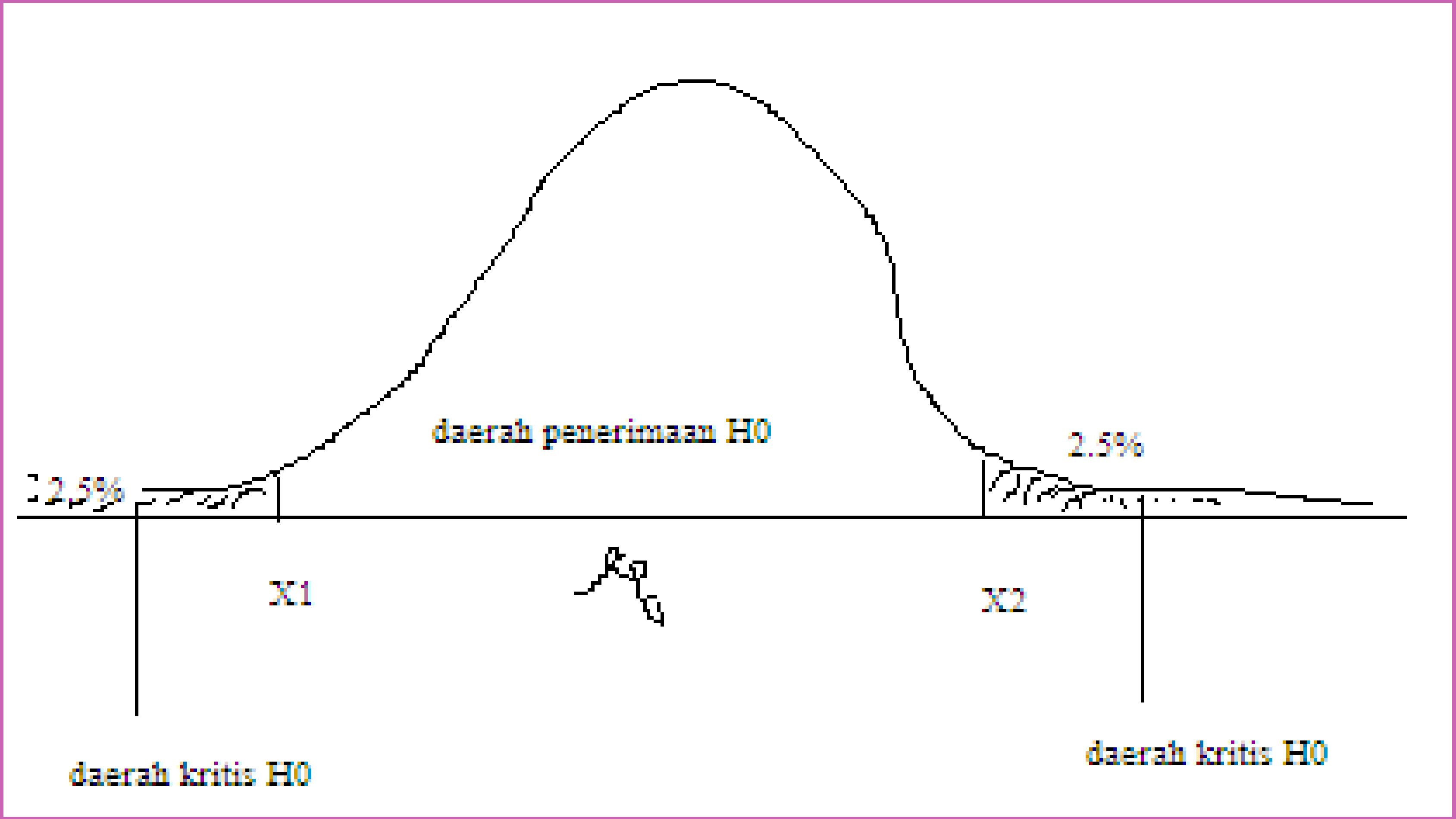
NOT Reject H_0 in favor of H_1

5) Two possible types of errors:

Type I: reject a true H_0 [P(TypeI)=α]

Type II: not reject a false H_0 [P(Type II)=β]

Significance Level



Ex $$H0 : \mu = \mu 0$$

Step on arrange hypothesis

i. Hypothesis : $a. H_0 : \theta = \theta_0$

$$H_1 : \theta \neq \theta_0$$

$$b. H_0 : \theta = \theta_0$$

$$H_1 : \theta > \theta_0$$

$$c. H_0 : \theta = \theta_0$$

$$H_1 : \theta < \theta_0$$

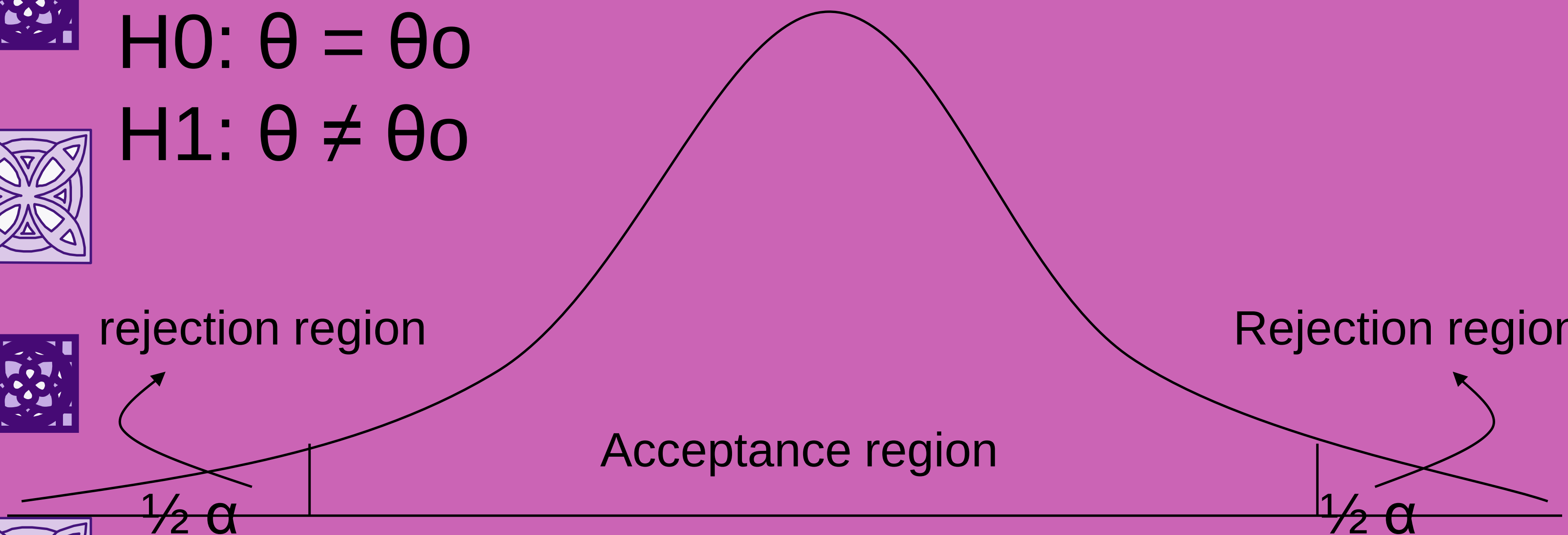
ii. Significance level

H1:

One of Learning methods better than the other

2 tailed

- H0: θ = θo
- H1: θ ≠ θo



iii. H0 accepted if

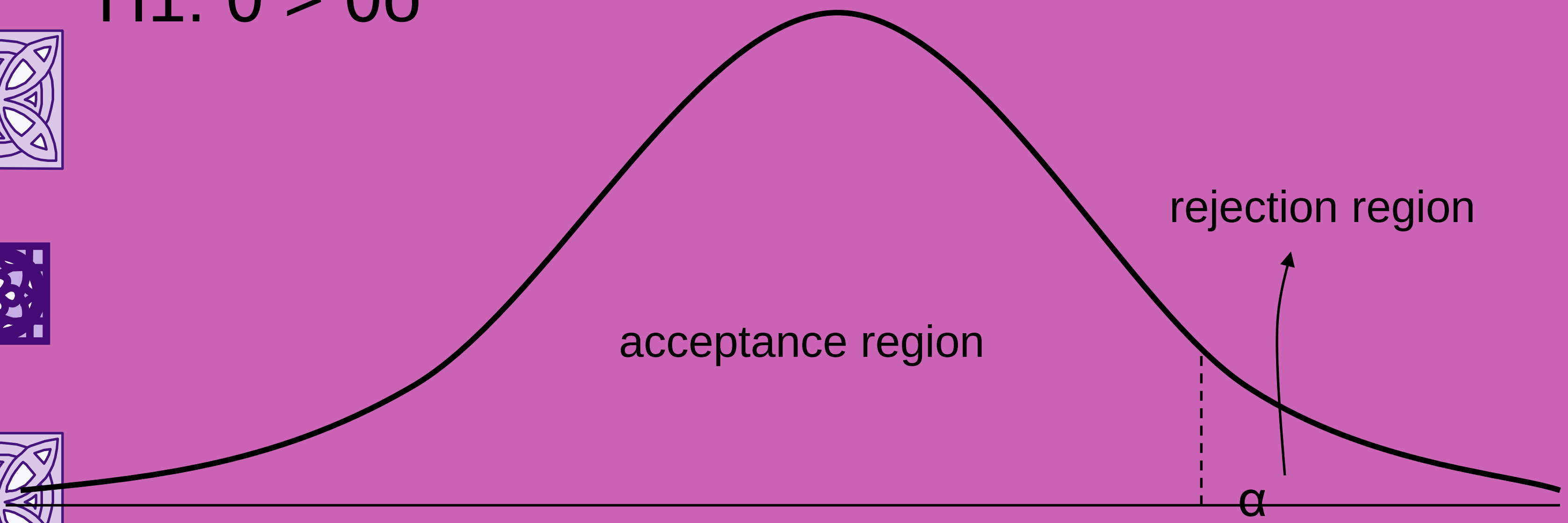
$$-Z_{1/2\alpha} < Z < Z_{1/2\ \alpha}$$

H1:

Learning Methods A better than B

1 tailed (right)

- H0: $\theta = \theta_o$
- H1: $\theta > \theta_o$



iii. H0 accepted if

$$z \leq z_{\alpha}$$

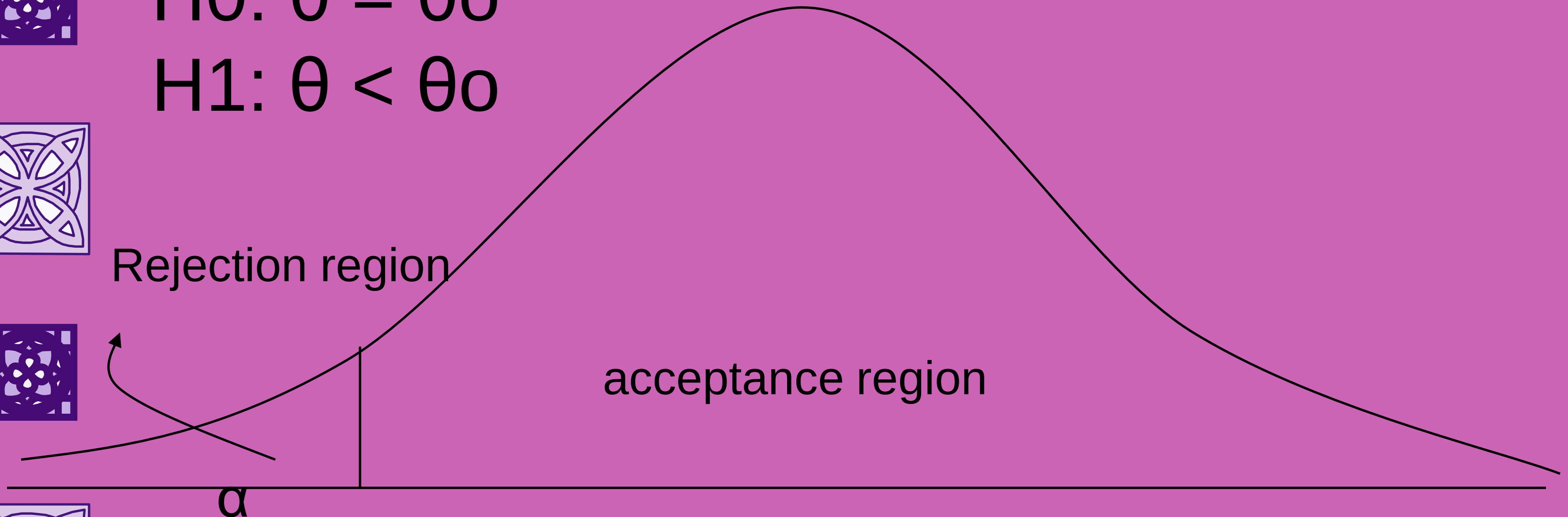
H1:

Learning Methods A worse thanB

1 tailed (left)

$H0: \theta = \theta o$

$H1: \theta < \theta o$



iii. H0 accepted if

$$z \geq -z_{\alpha}$$

iv. Calculation:

$$Z = \frac{\bar{X} - \theta_0}{\sigma / \sqrt{n}}$$

$$Z = \frac{\bar{X} - \theta_0}{s / \sqrt{n}} \text{ if } \sigma \text{ unknow}$$

Example 1

Will be tested if means of the heigth of chemistry class is 160 cm or different from that measure. If it take significance level 5% and sample random 100 students then it showed that means of height of the students is 163.5 cm with standard deviation 4.8 cm. Test the hypothesis!

Solution

i. Hypothesis: $H_0 : \mu = 160$

$H_1 : \mu \neq 160$

ii. Significance level :0.05

iii. Critical Region

$$H_0 \text{ rejected if } Z < -Z_{\frac{\alpha}{2}} \text{ or } Z > Z_{\frac{\alpha}{2}}$$

$$H_0 \text{ rejected if } Z < -1.96 \text{ or } Z > 1.96$$

iv. Calculation

$$Z = \frac{\bar{X} - \mu_0}{\sigma / \sqrt{n}} = \frac{163.5 - 160}{4.8 / \sqrt{100}} = 7.29$$

v. Conclution

Because Z=7.29>1.96 then H0 is rejected

p-Value

- The ***p-value*** of a test is the probability of observing a test statistic at least as extreme as the one computed given that the null hypothesis is true.

- In the case of our department store example, what is the ***probability*** of observing a sample mean ***at least as extreme*** as the one already observed (i.e. $\bar{x}$ = 178), given that the null hypothesis (H0: μ = 170) is true?

$$P(\bar{x} > 178) = P\left(\frac{\bar{x}-\mu}{\sigma/\sqrt{n}} > \frac{178-170}{65/\sqrt{400}}\right) = P(Z > 2.46) = .0069$$

p-value

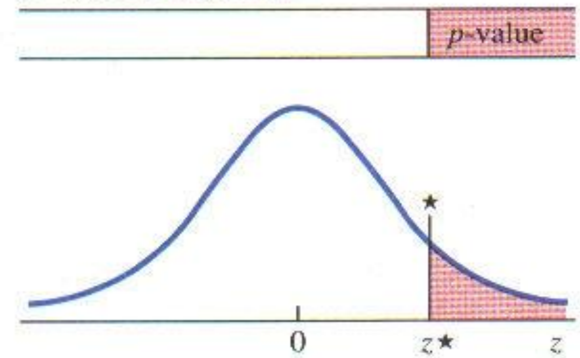
Table 8.6 Finding p-values

Case 1
H_a contains ">"
"Right tail"

p-value is the *area to the right of* $z^\star$

$p\text{-value} = P(z > z^\star)$

p-Value in Right Tail



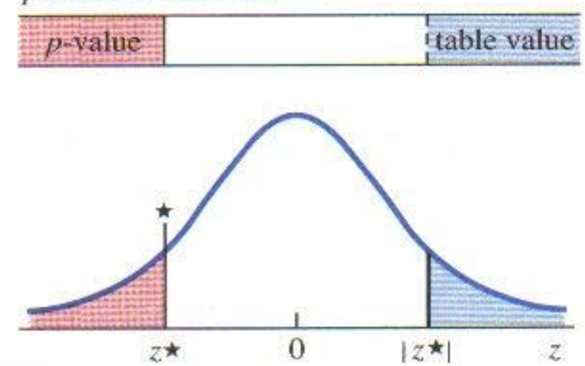
Case 2
H_a contains "<"
"Left tail"

p-value is the *area to the left of* $z^\star$

The area of the left tail is the same as the area in the right tail bounded by the positive $z^\star$, therefore

$p\text{-value} = P(z < z^\star) = P(z > |z^\star|)$

p-Value in Left Tail



Case 3
H_a contains "≠"
"Two-tailed"

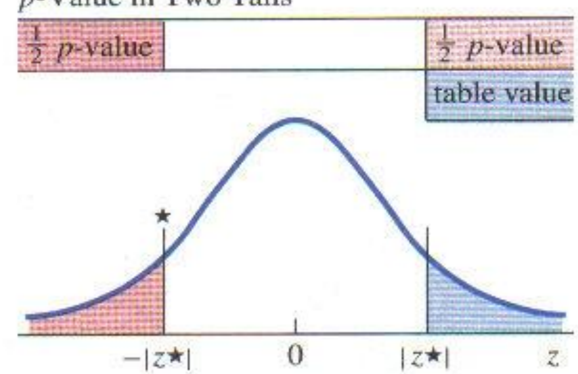
p-value is the *total area of both tails*

$p\text{-value} = P(z < -|z^\star|) + P(z > |z^\star|)$

Since both areas are equal, find the probability of one tail and double it.

Thus, $p\text{-value} = 2 \times P(z > |z^\star|)$

p-Value in Two Tails



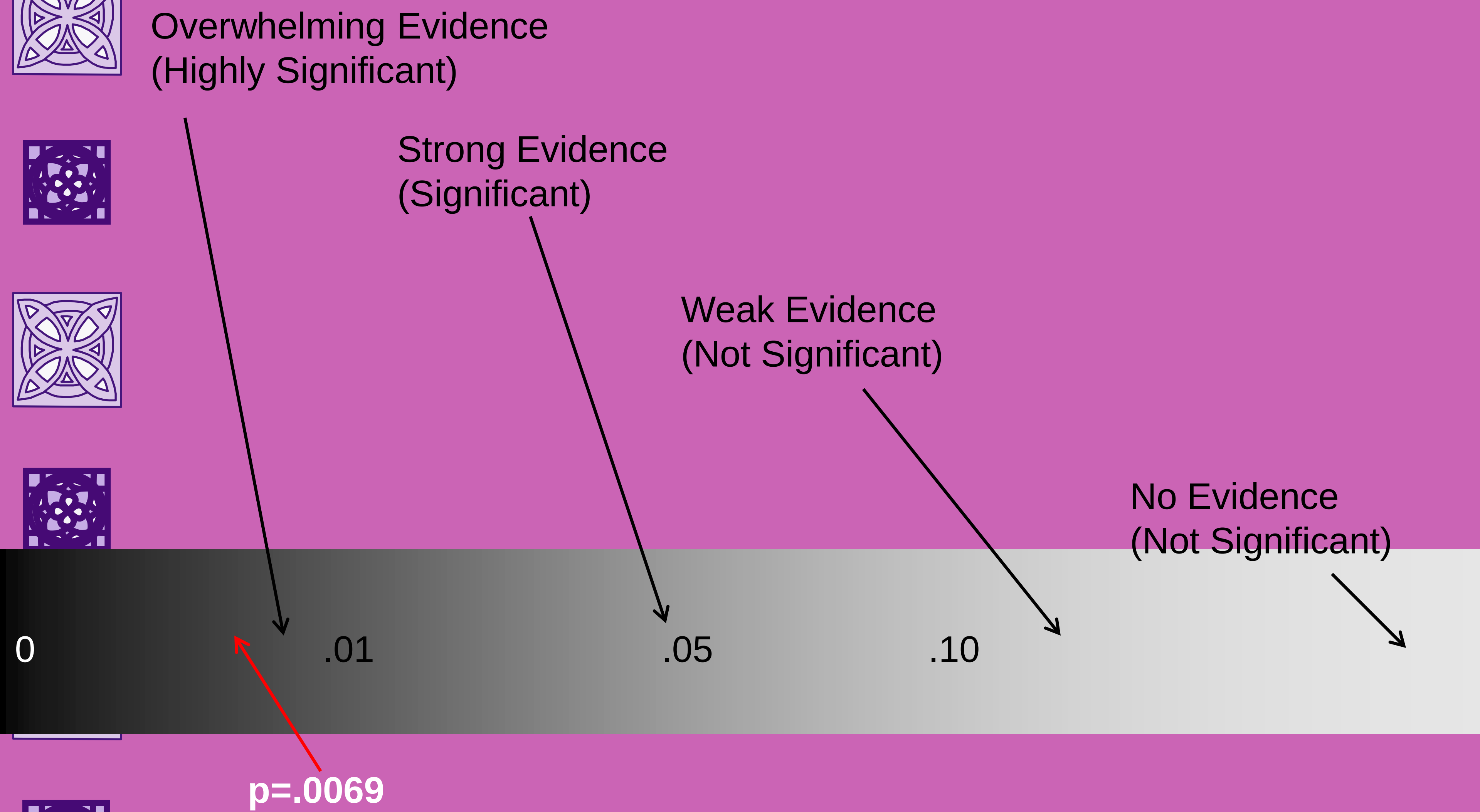
Interpreting the p-value…

The smaller the p-value, the more statistical evidence exists to support the alternative hypothesis.

- If the p-value is less than 1%, there is ***overwhelming evidence*** that supports the alternative hypothesis.
- If the p-value is between 1% and 5%, there is a ***strong evidence*** that supports the alternative hypothesis.
- If the p-value is between 5% and 10% there is a ***weak evidence*** that supports the alternative hypothesis.
- If the p-value exceeds 10%, there is ***no evidence*** that supports the alternative hypothesis.

We observe a p-value of .0069, hence there is overwhelming evidence to support H_1: $\mu > 170$.

Interpreting the p-value…



Interpreting the p-value…

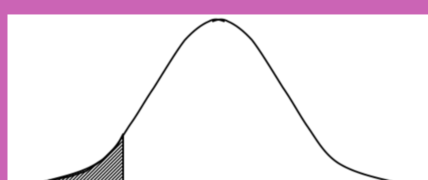
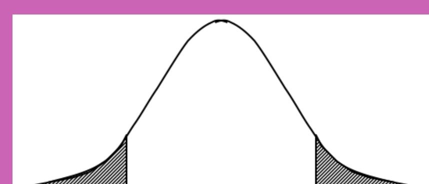
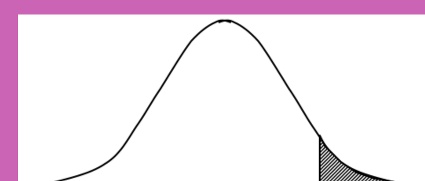
Compare the p-value with the selected value of the significance level: α

If the p-value is less than α , we judge the p-value to be small enough to reject the null hypothesis.

If the p-value is greater than α, we do not reject the null hypothesis.

Since p-value = .0069 < = .05, we reject H_0 in favor of H_1

Summary of One- and Two-Tail Tests…

One-Tail Test (left tail)	Two-Tail Test	One-Tail Test (right tail)
$H_0: \mu = \mu_0$ $H_1: \mu < \mu_0$	$H_0: \mu = \mu_0$ $H_1: \mu \neq \mu_0$	$H_0: \mu = \mu_0$ $H_1: \mu > \mu_0$
		

Large-Sample Test H_0:μ_1-μ_2=0 vs H_0:μ_1-μ_2>0

- ***H*$_0$: μ_1-μ_2 = 0 (No difference in population means**
- ***H*$_A$: μ_1-μ_2 > 0 (Population Mean 1 > Pop Mean 2)**

- $T.S.: z_{obs} = \dfrac{\overline{y}_1 - \overline{y}_2}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}}$
- $R.R.: z_{obs} \geq z_{\alpha}$
- $P-value: P(Z \geq z_{obs})$

- **Conclusion -** Reject H_0 if test statistic falls in rejection region, or equivalently the *P*-value is $\leq \alpha$

Example - Botox for Cervical Dystonia

- **Patients -** Individuals suffering from cervical dystonia
- **Response -** Tsui score of severity of cervical dystonia (higher scores are more severe) at week 8 of Tx
- **Research (alternative) hypothesis** - Botox A decreases mean Tsui score more than placebo
- **Groups -** Placebo (Group 1) and Botox A (Group 2)
- **Experimental (Sample) Results:**

$$\overline{y}_1 = 10.1 \quad s_1 = 3.6 \quad n_1 = 33$$

$$\overline{y}_2 = 7.7 \quad s_2 = 3.4 \quad n_2 = 35$$

Example - Botox for Cervical Dystonia

Test whether Botox A produces lower mean Tsui scores than placebo ($\alpha = 0.05$)

$$i.\quad H_0 : \mu_1 - \mu_2 = 0$$

$$H_A : \mu_1 - \mu_2 > 0$$

$$ii.\quad \alpha = 5\%$$

$$iii.\ z_{obs} = \frac{10.1 - 7.7}{\sqrt{\frac{(3.6)^2}{33} + \frac{(3.4)^2}{35}}} = \frac{2.4}{0.85} = 2.82$$

$$iv.\quad z_{obs} \geq z_{\alpha} = z_{.05} = 1.645$$

$$v.\quad P - val : P(Z \geq 2.82) = .0024$$

Conclusion: Botox A produces lower mean Tsui scores than placebo (since 2.82 > 1.645 and *P*-value < 0.05)

2-Sided Tests

- Many studies don't assume a direction with the difference μ_1-μ_2
- H_0: μ_1-μ_2 = 0 H_A: μ_1-$\mu_2 \neq 0$
- Test statistic is the same as before
- Decision Rule:
  - Conclude μ_1-μ_2 > 0 if $z_{obs} \geq z_{\alpha/2}$ (α=0.05 $\Rightarrow$ $z_{\alpha/2}$=1.96)
  - Conclude μ_1-μ_2 < 0 if $z_{obs} \geq -z_{\alpha/2}$ (α=0.05 $\Rightarrow$ $-z_{\alpha/2}$= -1.96)
  - Do not reject μ_1-μ_2 = 0 if $-z_{\alpha/2} \leq z_{obs} \leq z_{\alpha/2}$
- *P*-value: $2P(Z \geq |z_{obs}|)$

Power of a Test

- **Power -** Probability a test rejects H_0 (depends on $\mu_1 - \mu_2$)
  - H_0 True: Power = P(Type I error) = α
  - H_0 False: Power = 1-P(Type II error) = 1-β
- **Example:**
  - $H_0: \mu_1 - \mu_2 = 0 \quad H_A: \mu_1 - \mu_2 > 0$
  - $\sigma_1^2 = \sigma_2^2 = 25 \quad n_1 = n_2 = 25$
  - Decision Rule: Reject H_0 (at α=0.05 significance level) if:

$$z_{obs} = \frac{\overline{y}_1 - \overline{y}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} = \frac{\overline{y}_1 - \overline{y}_2}{\sqrt{2}} \geq 1.645 \quad \Rightarrow \quad \overline{y}_1 - \overline{y}_2 \geq 2.326$$