Hypothesis Testing: Significance

SECTION 4.3
- Significance level (4.3)
- Statistical conclusions (4.3)

p-value and $H_0$
- If the p-value is small, then a statistic as extreme as that observed would be unlikely if the null hypothesis were true, providing significant evidence against $H_0$
- The smaller the p-value, the stronger the evidence against the null hypothesis and in favor of the alternative

The smaller the p-value, the stronger the evidence against $H_0$.

Which of the following p-values gives the strongest evidence against $H_0$?

a) 0.005  
b) 0.1  
c) 0.32  
d) 0.56  
e) 0.94

Which of the following p-values gives the strongest evidence against $H_0$?

a) 0.22  
b) 0.45  
c) 0.03  
d) 0.8  
e) 0.71

Two different studies obtain two different p-values. Study A obtained a p-value of 0.002 and Study B obtained a p-value of 0.2. Which study obtained stronger evidence against the null hypothesis?

a) Study A  
b) Study B
Formal Decisions

- If the p-value is small:
  - the sample would be extreme if $H_0$ were true
  - the results are statistically significant
  - REJECT $H_0$
  - we have evidence for $H_a$
- If the p-value is not small:
  - the sample would not be too extreme if $H_0$ were true
  - the results are not statistically significant
  - DO NOT REJECT $H_0$
  - the test is inconclusive; either $H_0$ or $H_a$ may be true

Significance Level

- The significance level, $\alpha$, is the threshold below which the p-value is deemed small enough to reject the null hypothesis

$$p\text{-value} < \alpha \implies \text{Reject } H_0$$
$$p\text{-value} \geq \alpha \implies \text{Do not Reject } H_0$$

Statistical Significance

- $p\text{-value} < \alpha$
  - Results would be rare, if the null were true
  - Reject $H_0$
  - We have evidence that the alternative is true!
- $p\text{-value} \geq \alpha$
  - Results would not be rare, if the null were true
  - Do not reject $H_0$
  - We can make no conclusions either way

Statistical Conclusions

In a hypothesis test of

$$H_0: \mu = 10 \text{ vs } H_a: \mu < 10$$

the p-value is 0.002. With $\alpha = 0.05$, we conclude:

a) Reject $H_0$
b) Do not reject $H_0$
c) Reject $H_a$
d) Do not reject $H_a$
**Statistical Conclusions**

In a hypothesis test of

\[ H_0: \mu = 10 \text{ vs } H_a: \mu < 10 \]

the p-value is 0.002. With \( \alpha = 0.01 \), we conclude:

a) There is evidence that \( \mu = 10 \)
b) There is evidence that \( \mu < 10 \)
c) We have insufficient evidence to conclude anything

---

**Statistical Conclusions**

In a hypothesis test of

\[ H_0: \mu = 10 \text{ vs } H_a: \mu < 10 \]

the p-value is 0.21. With \( \alpha = 0.01 \), we conclude:

a) Reject \( H_0 \)
b) Do not reject \( H_0 \)
c) Reject \( H_a \)
d) Do not reject \( H_a \)

---

**Statistical Conclusions**

In a hypothesis test of

\[ H_0: \mu = 10 \text{ vs } H_a: \mu < 10 \]

the p-value is 0.21. With \( \alpha = 0.01 \), we conclude:

a) There is evidence that \( \mu = 10 \)
b) There is evidence that \( \mu < 10 \)
c) We have insufficient evidence to conclude anything

---

**Elephant Example**

\[ H_0: X \text{ is an elephant} \]
\[ H_a: X \text{ is not an elephant} \]

Would you conclude, if you get the following data?

- \( X \) walks on two legs
- \( X \) has four legs

---

**Never Accept \( H_0 \)**

- “Do not reject \( H_0 \)” is not the same as “accept \( H_0 \)”!
- Lack of evidence against \( H_0 \) is NOT the same as evidence for \( H_0 \)!

“For the logical fallacy of believing that a hypothesis has been proved to be true, merely because it is not contradicted by the available facts, has no more right to insinuate itself in statistical than in other kinds of scientific reasoning...” - Sir R. A. Fisher

---

**Red Wine and Weight Loss**

- Resveratrol, an ingredient in red wine and grapes, has been shown to promote weight loss in rodents, and has recently been investigated in primates (specifically, the Grey Mouse Lemur).
- A sample of lemurs had various measurements taken before and after receiving resveratrol supplementation for 4 weeks

Red Wine and Weight Loss

In the test to see if the mean resting metabolic rate is higher after treatment, the p-value is 0.013. Using $\alpha = 0.05$, is this difference statistically significant? (should we reject $H_0$: no difference?)

a) Yes  
b) No

Red Wine and Weight Loss

In the test to see if the mean body mass is lower after treatment, the p-value is 0.007. Using $\alpha = 0.05$, is this difference statistically significant? (should we reject $H_0$: no difference?)

a) Yes  
b) No

Red Wine and Weight Loss

In the test to see if locomotor activity changes after treatment, the p-value is 0.980. Using $\alpha = 0.05$, is this difference statistically significant? (should we reject $H_0$: no difference?)

a) Yes  
b) No

Red Wine and Weight Loss

In the test to see if the mean resting metabolic rate is higher after treatment, the p-value is 0.013. Using $\alpha = 0.05$, what can we conclude?

a) Mean resting metabolic rate is higher after resveratrol supplementation in lemurs  
b) Mean resting metabolic rate is not higher after resveratrol supplementation in lemurs  
c) Nothing

Red Wine and Weight Loss

In the test to see if the mean body mass is lower after treatment, the p-value is 0.007. Using $\alpha = 0.05$, what can we conclude?

a) Mean body mass is higher after resveratrol supplementation in lemurs  
b) Mean body mass is not higher after resveratrol supplementation in lemurs  
c) Nothing

Red Wine and Weight Loss

In the test to see if locomotor activity changes after treatment, the p-value is 0.980. Using $\alpha = 0.05$, what can we conclude?

a) Locomotor activity changes after resveratrol supplementation in lemurs  
b) Locomotor activity does not change after resveratrol supplementation in lemurs  
c) Nothing
Red Wine and Weight Loss

In the test to see if mean food intake changes after treatment, the p-value is 0.035.
Using \( \alpha = 0.10 \), is this difference statistically significant? (should we reject \( H_0: \) no difference?)

a) Yes  
b) No

Multiple Sclerosis and Sunlight

- It is believed that sunlight offers some protection against multiple sclerosis, but the reason is unknown
- Researchers randomly assigned mice to one of:
  - Control (nothing)
  - Vitamin D Supplements
  - UV Light
- All mice were injected with proteins known to induce a mouse form of MS, and they observed which mice got MS


Informal strength of evidence against \( H_0 \):

<table>
<thead>
<tr>
<th>Very Strong</th>
<th>Strong</th>
<th>Moderate</th>
<th>Some</th>
<th>Little</th>
</tr>
</thead>
<tbody>
<tr>
<td>1%</td>
<td>5%</td>
<td>10%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Formal decision of hypothesis test, based on \( \alpha = 0.05 \):

- Reject \( H_0 \) if the p-value is less than 0.05
- Do not reject \( H_0 \) if the p-value is greater than 0.05

In testing whether UV light provides protection against MS (UV light vs control group), the p-value is 0.002.
Multiple Sclerosis and Sunlight

- In testing whether Vitamin D provides protection against MS (Vitamin D vs control group), the p-value is 0.47.

Conclusions

- **p-value < α**
  - **Generic conclusion:** Reject $H_0$
  - **Conclusion in context:** We have (strong?) evidence that [fill in alternative hypothesis]

- **p-value ≥ α**
  - **Generic conclusion:** Do not reject $H_0$
  - **Conclusion in context:** We do not have enough evidence to conclude that [fill in alternative hypothesis]

To Do

- Read Section 4.3 (will cover errors next class)
- Do HW 4.3 (due Friday, 3/20)
- Enjoy spring break!!!