Course > Inference: Hypothesis Testing Overview > Hypothesis Testing Overview >
Hypothesis Testing: Making Conclusions

🔖 **Bookmark this page**

Hypothesis Testing: Making Conclusions

> **Learning Objective: Explain the logic behind and the process of hypotheses testing. In particular, explain what the p-value is and how it is used to draw conclusions.**

Hypothesis testing step 4: Making conclusions.

Since our conclusion is based on how small the p-value is, or in other words, how surprising our data are when H_o is true, it would be nice to have some kind of guideline or cutoff that will help determine how small the p-value must be, or how "rare" (unlikely) our data must be when H_o is true, for us to conclude that we have enough evidence to reject H_o.

This cutoff exists, and because it is so important, it has a special name. It is called the **significance level of the test** and is usually denoted by the Greek letter α. The most commonly used significance level is α = 0.05 (or 5%). This means that:

- if the p-value < α (usually 0.05), then the data we got is considered to be "rare (or surprising) enough" when H_o is true, and we say that the data provide significant evidence against H_o, so we reject H_o and accept H_a.

- if the p-value > α (usually 0.05), then our data are not considered to be "surprising enough" when H_o is true, and we say that our data do not provide enough evidence to reject H_o (or, equivalently, that the data do not provide enough evidence to accept H_a).

Important comment about wording.

Another common wording (mostly in scientific journals) is:

"The results are statistically significant" - when the p-value < α.

"The results are not statistically significant" - when the p-value > α.

Comments

1. Although the significance level provides a good guideline for drawing our conclusions, it should not be treated as an incontrovertible truth. There is a lot of room for personal interpretation. What if your p-value is 0.052? You might want to stick to the rules and say "0.052 > 0.05 and therefore I don't have enough evidence to reject H_o", but you might decide that 0.052 is small enough for you to believe that H_o should be rejected.

   It should be noted that scientific journals do consider 0.05 to be the cutoff point for which any p-value below the cutoff indicates enough evidence against H_o, and any p-value above it, *or even equal to it*, indicates there is not enough evidence against H_o.

2. It is important to draw your conclusions **in context**. It is **never enough** to say: *"p-value = ..., and therefore I have enough evidence to reject H_o at the .05 significance level."* You **should always add:** "... and conclude that ... (what it means in the context of the problem)".

3. Let's go back to the issue of the nature of the two types of conclusions that I can make.

   Either *I reject H_o and accept H_a (when the p-value is smaller than the significance level)* **or** *I cannot reject H_o (when the p-value is larger than the significance level).*

As we mentioned earlier, note that the second conclusion does not imply that I accept H_o, but just that I don't have enough evidence to reject it. Saying (by mistake) "I don't have enough evidence to reject H_o so I accept it" indicates that the data provide evidence that H_o is true, which is **not necessarily the case**. Consider the following slightly artificial yet effective example:

> ### Example

An employer claims to subscribe to an "equal opportunity" policy, not hiring men any more often than women for managerial positions. Is this credible? You're not sure, so you want to test the following **two hypotheses:**

- **H_o:** The proportion of male managers hired is 0.5

- **H_a:** The proportion of male managers hired is more than 0.5

Data: You choose at random three of the new managers who were hired in the last 5 years and find that all 3 are men.

Assessing Evidence: If the proportion of male managers hired is really 0.5 (H_o is true), then the probability that the random selection of three managers will yield three males is therefore 0.5 * 0.5 * 0.5 = 0.125. This is the p-value.

Conclusion: Using .05 as the significance level, you conclude that since the p-value = 0.125 > 0.05, the fact that the three randomly selected mangers were all males is not enough evidence to reject H_o. In other words, you do not have enough evidence to reject the employer's claim of subscribing to an equal opportunity policy.

However, **the data (all three selected are males) definitely does not provide evidence to accept the employer's claim (H_o).**

Scenario: Opinions on Gay Marriage

The following two hypotheses are tested:

- H_o: The proportion of U.S. adults who oppose gay marriage is roughly 50%.

- H_a: The proportion of U.S. adults who oppose gay marriage is above 50% (i.e., the majority oppose).

Suppose a survey was conducted in which a random sample of 1,100 U.S. adults was asked about their opinions about gay marriage, and based on the data, the p-value was found to be 0.002.

Comment: Throughout this activity use a 0.05 (5%) significance level (cutoff).

Learn By Doing

1/1 point (graded)
What does the fact that the p-value = 0.002 mean?

- ○ There is 0.002 probability of observing data like those observed.

- ○ There is 0.002 probability that 50% of U.S. adults oppose gay marriage.

- ● There is a probability of 0.002 (i.e., very unlikely) to observe data like those observed if the proportion of U.S. adults who oppose gay marriage were 50%. ✔

- ○ There is 0.998 probability that the majority of U.S. adults oppose gay marriage.

Answer
Correct:
Indeed, the p-value is the probability of observing data like those observed assuming that the null hypothesis, H_o, is true.

Submit

Learn By Doing

1/1 point (graded)

What can you conclude based on the p-value?

○ The data provide significant evidence that the proportion of U.S. adults who oppose gay marriage is 50%.

⦿ The data provide significant evidence that the majority of U.S. adults oppose gay marriage. ✔

○ The data do not provide enough evidence to conclude that the majority of U.S. adults oppose gay marriage.

○ The data provide evidence that H_a is more likely than H_o (i.e., it is more likely that the majority of U.S. adults oppose gay marriage).

Answer

Correct:

Indeed, since the p-value is small (less than 0.05) we have enough evidence to reject H_o and accept H_a, or in other words, to conclude that the majority of U.S. adults oppose gay marriage.

Submit

Learn By Doing

1/1 point (graded)

Say that the p-value was not given, but rather, the following conclusion was advertised: "The survey does not provide enough evidence to conclude that the majority of U.S. adults oppose gay marriage." Which of the following could have been the p-value that led to this conclusion?

⦿ 0.125 ✔

○ 1.96

○ 0.045

○   -1.96

Answer

Correct:

Indeed, when the p-value is not small (in this case, above 0.05) our conclusion is that we don't have enough evidence to reject H_o, or equivalently, that we don't have enough evidence to accept H_a.

[Submit]

Learn By Doing

1/1 point (graded)

When would you conclude that the data provide enough evidence that the proportion of U.S. adults who oppose gay marriage is 50%?

○   When the p-value is small (less than 0.05)

○   When the p-value is not small (above 0.05)

○   When exactly half the individuals in the sample oppose gay marriage and half support it

●   Never ✔

Answer

Correct:

Indeed, we never conclude that the data provide enough information to accept H_o. (Comment 3 right before this activity discusses this point).

[Submit]

Scenario: Annual Miles Driven

The following two hypotheses are tested:

- H_o: The average number of miles driven per year is 12,000.

- H_a: The average number of miles driven per year is less than 12,000.

In a survey, 1,600 randomly selected drivers were asked the number of miles they drive yearly. Based upon the results, the p-value = 0.068.

Comment: Throughout this activity use a 0.05 (5%) significance level.

Did I Get This

1/1 point (graded)
What does the fact that the p-value is 0.068 mean?

- ● There is a probability of 0.068 (reasonably likely) of observing data like those observed if the average number of miles driven per year is 12,000. ✔

- ○ There is a probability of 0.068 of observing data like those observed.

- ○ There is a probability of 0.068 that the average number of miles driven per year is 12,000.

- ○ There is a 0.932 probability that the average number of miles driven per year is less than 12,000.

Answer
Correct:
Indeed, the p-value is the probability of observing data like those observed assuming that the null hypothesis, H_o, is true.

Submit

Did I Get This

1/1 point (graded)
Based upon the p-value, which of the following conclusions can be made?

- ○ The data provide significant evidence that the average number of miles driven per year is less than 12,000.

- ● The data do not provide significant evidence that the average number of miles driven per year is less than 12,000. ✔

- ○ The data provide significant evidence that the average number of miles driven per year is 12,000.

○ The data provide significant evidence that the alternative hypothesis is more likely than the null hypothesis.

Answer
Correct:
Since the p-value is > 0.05, we cannot accept the alternative hypothesis. We could have also stated that we do not reject the null hypothesis (that the data does not provide significant evidence to reject that the average number of miles driven per year is 12,000).

Submit

Did I Get This

1/1 point (graded)
Suppose that the p-value was not given, but instead the following conclusion was made: data provides significant evidence that the average number of miles driven per year is less than 12,000. Which of the following could have been the p-value?

○ -1.96

● 0.032 ✔

○ 0.05

○ 0.068

Answer
Correct:
Indeed, the p-value must be less than 0.05 to provide enough evidence to reject the null hypothesis and accept the alternative.

Submit

Did I Get This

1/1 point (graded)
When would you conclude that the average number of miles driven per year is less than 12,000?

○ Never.

○ When the p-value is large (more than 0.05).

◉ When the p-value is small (less than 0.05). ✔

○ When the average number of miles is less than 12,000.

Answer
Correct: We can accept H_a if there is sufficient evidence, namely that the p-value < 0.05.

Submit

Open Learning Initiative ⬀