Confidence Intervals and Hypothesis Testing

Module 4c

Specifying the Null and Alternative Hypotheses

Learning Objectives

- Understanding the logic of hypothesis testing
- Identifying the Null and Alternative hypotheses
- Determine when to use one-tailed vs. two-tailed tests

Hypothesis Testing

We have been drawing inferences about μ using **confidence intervals**:
"There is a 95% chance that μ is between ____ and ____."

But what if we want to test a specific claim about μ?

"Is the mean income of Pittsburgh Registered Nurses over $40,000?"

Does our new integrated development environment affect programmer productivity?

In each case, we want to decide which of two possible hypotheses is true:

$H_1 : \mu > \$40,000$
$H_0 : \mu \leq \$40,000$

$H_1 : \mu \neq \mu_0$
$H_0 : \mu = \mu_0$

μ is some objective measure of productivity, and μ_0 is its historical average

H_1 : The hypothesis we are testing – also called the **alternative hypothesis**
H_0 : What we assume to be true in the absence of further evidence – also called the **null hypothesis**

Hypothesis Testing

We want to test the alternative hypothesis H_1: $\mu \neq 1000$ against the null hypothesis H_0: $\mu = 1000$.

Let us assume that we want to measure productivity in terms of lines of production-quality code written, and that historically we have achieved an average of $\mu_0 = 1000$ lines of code per day.

Generally, the alternative hypothesis H_1 indicates that there is an effect (e.g. significant increase or decrease in some quantity) while the null hypothesis H_0 indicates that there is no effect (e.g. the quantity has not changed significantly).

Our test will give one of two possible outcomes:

1. We can reject the null hypothesis, and thus the alternative hypothesis is true.
2. We cannot reject the null hypothesis. **This does not necessarily mean that the null is true!**

"We can conclude that $\mu \neq 1000$."

"We do not have sufficient evidence to conclude that $\mu \neq 1000$."

Hypothesis Testing

We want to test the alternative hypothesis H_1: $\mu \neq 1000$ against the null hypothesis H_0: $\mu = 1000$.

Let us assume that we want to measure productivity in terms of lines of production-quality code written, and that historically we have achieved an average of $\mu_0 = 1000$ lines of code per day.

Key idea: the sample evidence must **strongly** contradict the null hypothesis for us to reject it in favor of the alternative.

Our test will give one of two possible outcomes:

1. We can reject the null hypothesis, and thus the alternative hypothesis is true.

2. We cannot reject the null hypothesis. **This does not necessarily mean that the null is true!**

"We can conclude that $\mu \neq 1000$."

"We do not have sufficient evidence to conclude that $\mu \neq 1000$."

Tea Tasting Experiment

"A lady declares that by tasting a cup of tea made with milk she can discriminate whether the milk or tea infusion was first added to the cup." *The Design of Experiments* by Sir Ronald Fisher (1935)

- You're now going to the ISLE platform to answer the following questions:
  - Suppose she does **not** have the ability to discern which was added first and she just guesses. What's the probability, p, she'll guess correctly?
  - If we're skeptical of her ability to discern, what is the null hypothesis?
  - The alternative hypothesis?

Tea Tasting Experiment

"A lady declares that by tasting a cup of tea made with milk she can discriminate whether the milk or tea infusion was first added to the cup." *The Design of Experiments* by Sir Ronald Fisher (1935)

- Null Hypothesis: She can't distinguish which came first, the tea or milk ($p = 0.5$)
- Alternative Hypothesis: She can distinguish ($p > 0.5$)
- To test her discerning ability, 8 cups of tea are prepared, 4 with the tea added first and 4 with the milk added first.
- Decision criteria—We'll only reject the null hypothesis if the data from the test strongly supports the alternative hypothesis.
  - Suppose she gets all 8 right
  - P(guessing 8 out of 8 correctly)=.5*.5*…5= ≈ 0
- Very convincing evidence that she can discern which in reality she was able to do!

Hypothesis Testing Is <u>NOT</u> a Fair Race Between the Null and Alternative Hypotheses—Alternative Must Have Strong Support

Identifying H_1 and H_0

A statistical hypothesis is an assumption about some parameter of a population, such as the population mean μ or population proportion p.

The alternative hypothesis H_1 is some claim about a parameter that you want to demonstrate.

The null hypothesis H_0 is the assumption about this parameter that you must reject in order to show that H_1 is true.

Some examples of H_1 :

> "The community's average yearly expenditure on computing supplies is greater than $40."

> "Providing restaurant customers calorie information reduces their caloric consumption."

Specifying the Null Hypothesis

The <u>null hypothesis</u>, denoted by H_0, is expressed in one of three forms:

H_0: μ = specific value (often 0)

H_0: $\mu \leq$ specific value

H_0: $\mu \geq$ specific value

<u>Always</u> includes equality: =, ≤, or ≥

Specifying Alternative Hypothesis

The <u>alternative hypothesis</u>, denoted by *H_1 (or H_a)*, is the **opposite** of what is stated in the null hypothesis (null and alternative are mutually exclusive)

<u>Always</u> state in terms of a strict inequality (i.e., no "=")

- H_1 μ > value   if H_0: $\mu \leq$ value (one-tailed right)
- H_1 $\mu \neq$ value   if H_0: μ = value (two-tailed)
- H_1 μ < value   if H_0: $\mu \geq$ value (one-tailed left)

Tests of Understanding

For each scenario, state the null and alternative hypotheses:

- Suppose you are the purchasing agent for Boeing's 737 MAX electronic parts. Boeing's specifications for a particular device critical to the safe functioning of the 737 MAX requires that its mean life exceed 500 hours. A sample of 400 devices from the sole vendor of this part are tested. For this sample, average life was 510 hours with a standard deviation of 70 hours. Is this convincing evidence that the product meets specifications?
    - H_0:
    - H_1:
- A cereal manufacturer wants to test the performance of its filling machine. The machine is designed to discharge a mean of 12 ounces/box, and quality control wants to test for departures from that setting. The content of a sample of 100 boxes from the machine is weighted and the average fill is 12.1 ounces with a standard deviation of .5 ounces. Is this convincing evidence that the machine is not meeting the 12 ounce specification?
    - H_0:
    - H_1:
- A vocational training program claims that its graduates spend less time unemployed than individuals that only have a high school diploma. Suppose that on average the population of high school only grads are unemployed 25 days/year, whereas a sample of 100 graduates of the vocational program are on average unemployed 19 days/year. Does this data show that grads from this vocational program spend less time unemployed than high school only grads?
    - H_0:
    - H_1: