ECON120A: Hypothesis Testing

Week 10

Lapo Bini

UC San Diego

November 24, 2025

WHAT IS HYPOTHESIS TESTING?

- **Hypothesis testing** is a method of using **sample evidence** to evaluate a claim about a **population parameter** (μ). We formulate 2 hypotheses.
- The first one is the **null hypothesis** H_0, which is assumed to be true and represents the **"default belief"** (e.g., $H_0 : \mu = c$).
- The **alternative hypothesis** H_a is what we **seek evidence for**. It can be:
  - Left-sided: $H_a : \mu < c$
  - Right-sided: $H_a : \mu > c$
  - Two-sided: $H_a : \mu \neq c$
- How do we choose which one is right? We will use the sample mean $\bar{X}$. There are **two possible decisions**:
  - Reject H_0 (evidence supports H_a)
  - Fail to reject H_0 (evidence is not strong enough)

HOW DO WE CONDUCT A HYPOTHESIS TEST?

- **Step 1**: **State the hypotheses**. We must always have a null $\mathbf{H_0}$: μ **= c** vs. an alternative (left-, right- or two-sided).

- **Step 2**: Compute the **test statistic** Z, using the **null hypothesis**, which summarizes the evidence:

$$Z = \frac{\bar{X} - c}{\sigma/\sqrt{n}} \quad \text{(when } \sigma \text{ is known) or} \quad t = \frac{\bar{X} - c}{s/\sqrt{n}} \quad \text{(when } \sigma \text{ is unknown)}$$

- **Step 3**: Measure how unlikely the evidence is under H_0. **Two equivalent ways** to decide:
  1. **Critical value approach**: Reject H_0 when the test statistic enters the rejection region. Given different alternatives, the **rejection rules** are:

     RS: $Z > z_\alpha$ LS: $Z < -z_\alpha$ TS: $|Z| > z_{\alpha/2}$
  2. **p-value approach**: Compute the probability of observing a value at least as extreme as the test statistic Z (computed by assuming H_0 is true):

     RS: $p\text{-value} = P(z \geq Z \mid H_0)$ LS: $p\text{-value} = P(z \leq Z \mid H_0)$ TS: $p\text{-value} = 2P(z \geq |Z| \mid H_0)$

     The **rejection rule** is : reject H_0 if $p\text{-value} \leq \alpha$.

QUESTION 1

Q1: A report states that the mean number of texts that adults aged 18–24 send and receive daily is $\mu = 128$. We take a sample of thirty 25-34 year-olds to see whether their mean daily number of texts **differs** from 128. Suppose the sample showed a sample mean of 118.2 texts per day. Assume a population standard deviation of 33.17 texts per day.

(a) State the **null** and **alternative** hypotheses to test whether the mean for 25-34 year-olds differs from 128.

(b) Compute the **test statistic** (standardized Z).

(c) For significance level $\alpha = 0.05$, state the **critical values** and the **rejection rule** for this (two-sided) test.

(d) Using the test statistic and rejection rule, state your **conclusion** at the 5% significance level.

QUESTION 2

Q2: A production line operation is designed to fill cartons of laundry detergent to a **mean weight** of 32 ounces. Periodically, a sample of cartons is taken and weighed to determine whether **underfilling or overfilling** is occurring. If the data suggest underfilling or overfilling, the line will be shut down and adjusted.

- **Task 1:** Choose the appropriate **null** and **alternative hypotheses**:
  (a) $H_0 : \mu \leq 32, \quad H_a : \mu > 32$
  (b) $H_0 : \mu = 32, \quad H_a : \mu > 32$
  (c) $H_0 : \mu = 32, \quad H_a : \mu \neq 32$
  (d) $H_0 : \mu \geq 32, \quad H_a : \mu < 32$
- **Task 2:** Comment on the **conclusion and decision** when H_0 **cannot be rejected**.
- **Task 3:** Comment on the **conclusion and decision** when H_0 **can be rejected**.

QUESTION 3

Q3: The **average American** eats 9.5 pounds of chocolate annually. You are curious whether chocolate consumption is **higher** in Hershey, Pennsylvania. A sample of $n = 100$ individuals from the Hershey area has an average $\bar{x} = 10.05$ and volatility $s = 2.44$. Use $\alpha = 0.05$ to test whether mean annual chocolate consumption is higher in Hershey than in the U.S. overall.

(a) State the **null** and **alternative** hypotheses.

(b) Compute the **test statistic** for this (approximately) one-sample z (or large-sample t) test.

(c) Compute the corresponding **p-value**.

(d) State your **conclusion** at the 5% significance level.

QUESTION 4

Q4: The manager of a resort hotel claims that the **mean guest bill** for a weekend is **$500 or less**. An accountant believes the total charges have been **increasing** and will use a sample of future weekend guest bills to test the manager's claim.

- **Task 1:** Choose the appropriate **null** and **alternative hypotheses**:
  - (a) $H_0 : \mu \geq 500, \quad H_a : \mu < 500$
  - (b) $H_0 : \mu \leq 500, \quad H_a : \mu > 500$
  - (c) $H_0 : \mu = 500, \quad H_a : \mu \neq 500$
- **Task 2:** Comment on the **conclusion and decision** when H_0 **cannot be rejected**.
- **Task 3:** Comment on the **conclusion and decision** when H_0 **can be rejected**

QUESTION 5

Q5: Consider the hypothesis test $H_0: \ \mu \geq 20$, $H_a: \ \mu < 20$. A sample of $n = 50$ observations gives $\bar{x} = 19.1$. The population standard deviation is known to be $\sigma = 2$. Use $\alpha = 0.05$.

(a) Compute the **test statistic** (standardized Z).

(b) State the **critical value(s)** and the **rejection rule**. (Round critical value(s) to two decimal places)

(c) Compute the **p-value**. (Round to four decimal places.)

(d) Using $\alpha = 0.05$, state your **conclusion**.

QUESTION 6

Q6: An article in *Wired* reports that the mean daily screen time is 5 hours. You strongly believe that the mean screen time is **higher**. You take a random sample of $n = 144$ people: and $\bar{x} = 7.8$ hours, with standard deviation $s = 4.8$ hours. You test at a significance level $\alpha = 0.01$.

(a) State the **null** and **alternative hypotheses**.

(b) Compute the **test statistic**.

(c) Find the **critical value**.

(d) State whether you **reject** or **do not reject** H_0.

QUESTION 7

Q7: Consider the hypothesis test $H_0: \ \mu = 22$, $H_a: \ \mu \neq 22$. A sample of $n = 75$ observations gives $\bar{x} = 23$. The population standard deviation is $\sigma = 10$. Use $\alpha = 0.01$.

(a) Compute the **test statistic**.

(b) Find the **critical value**.

(c) Compute the **p-value**.

(d) State whether you **reject** or **do not reject** H_0.

QUESTION 8

Q8: Which of the following is an **appropriate alternative hypothesis** in a hypothesis test? Select the statement that **correctly describes** an **alternative hypothesis** about a **population parameter**. Explain why.

(A) The mean of a **sample** is equal to 50.

(B) The mean of a **population** is greater than 50.

(C) The mean of a **population** is equal to 50.

(D) The mean of a **sample** is greater than 50.

QUESTION 9

Q9: In testing $H_0 : \mu = 60$ vs. $H_a : \mu < 60$, the **p-value** is found to be 0.044 and the sample average is 56. Which of the following is the **correct interpretation** of the p-value? Select the option that correctly explains the **meaning of the p-value**.

(A) The probability of observing a sample mean equal to or larger than 56, when drawing a sample from a population with mean 60, is 0.044.

(B) The probability of rejecting the null when the null is true is 0.044.

(C) The probability that the population mean is smaller than 60 is 0.044.

(D) The probability of observing a sample mean as small as or even smaller than 56, when drawing a sample from a population with mean 60, is 0.044.

QUESTION 10

Q10: The IRS provides a toll-free help line. A report **claims** that callers must wait an **average of 12 minutes** on hold before talking to an employee. You decide to test $H_0: \ \mu = 12$ vs. $H_a: \ \mu < 12$. Assume the population standard deviation is $\sigma = 8$ minutes and that you will take a sample of $n = 49$. Before collecting data, you specify this **decision rule**:

$$\text{Reject } H_0 \text{ if the sample mean wait time } \bar{X} < 9.5 \text{ minutes.}$$

Task: Find the **significance level** α implied by this decision rule, that is

$$\alpha = P(\text{Reject } H_0 \mid H_0 \text{ is true}) = P\big(\bar{X} < 9.5 \mid \mu = 12\big).$$

QUESTION 11

Q11: Suppose the null hypothesis is **rejected** when a **one-sided** test is performed. If the test were instead conducted as a **two-sided** (with the same significance level and the same data), the null hypothesis would **always still be rejected**.

True or False? Briefly explain your reasoning.