Lecture: Hypothesis Testing

---

1 What is Hypothesis Testing?

Hypothesis testing is a **statistical method** used to make **decisions or inferences** about a population based on **sample data**.

- It helps us answer questions like:

"Is this new teaching method better than the old one?"
"Is the average lifetime of batteries really 500 hours?"

Key idea: You start with a claim (hypothesis) about a population and use **sample evidence** to test if it is likely true or not.

---

2 Components of Hypothesis Testing

1. **Null Hypothesis (H_0)**
   - The statement of **no effect, no difference, or equality**.
   - It's what we **assume true initially**.
   - Example: H_0: $\mu = 50$ (mean is 50)
2. **Alternative Hypothesis (H_1 or H_a)**
   - The statement we want to **test/prove**.
   - Indicates **effect, change, or difference**.
   - Example: H_1: $\mu \neq 50$ (mean is not 50)
3. **Significance Level (α)**
   - Probability of **rejecting H_0 when it is true** (Type I error).
   - Common values: 0.05 (5%), 0.01 (1%)
4. **Test Statistic**
   - A **numerical value** calculated from sample data to help decide whether to reject H_0.
   - Example: Z-test, t-test, chi-square test.

5. **Decision Rule**
    - Compare test statistic to **critical value(s)** or **p-value**.
    - Decide: **reject H_0** or **fail to reject H_0**.

---

3 Types of Hypothesis Testing

Based on the Tail

1. **Left-Tailed Test**
    - Tests if the parameter is **less than a certain value**.
    - Example: H_0: $\mu = 100$, H_1: $\mu < 100$
2. **Right-Tailed Test**
    - Tests if the parameter is **greater than a certain value**.
    - Example: H_0: $\mu = 100$, H_1: $\mu > 100$
3. **Two-Tailed Test**
    - Tests if the parameter is **different (either higher or lower)**.
    - Example: H_0: $\mu = 100$, H_1: $\mu \neq 100$

---

4 Z-Test

Definition

A **Z-test** is a hypothesis test where the **population standard deviation is known** and the sample size is **large (usually n ≥ 30)**.

Formula

$$Z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}}$$

Where:

- $\bar{x}$= sample mean

- μ = population mean (claimed)
- σ = population standard deviation
- n = sample size

Why We Use Z-Test

1. To **compare a sample mean with a known population mean**.
2. To check if the **difference is statistically significant**.
3. To make **inferences about the population**.

When to Use Z-Test

- Population standard deviation (σ) is **known**.
- Sample size is **large** ($n \geq 30$).
- Data is **normally distributed** (or approximately normal for large samples).
- For **testing population means or proportions**.

Types of Z-Tests

1. **One-Sample Z-Test**
   - Tests **one sample mean** against population mean.
2. **Two-Sample Z-Test**
   - Compares **two sample means**.
3. **Proportion Z-Test**
   - Tests a **sample proportion** against population proportion.

---

5 Steps for Hypothesis Testing Using Z-Test

1. **State H_0 and H_1**
2. **Choose significance level (α)**

3. **Calculate Z-statistic**
4. **Find critical Z-value(s)**
5. **Compare Z-statistic with critical value**
6. **Make decision and conclude**

---

1 Concept of Critical Value

The critical value is the Z-score that separates the rejection region (where we reject H_0) from the acceptance region (where we fail to reject H_0).

For a two-tailed test, the total significance level α is split equally in both tails:

$$\text{Each tail area} = \frac{\alpha}{2}$$

---

2 Two-Tailed Test at α = 0.05

- **Significance level: $\alpha = 0.05$**
- **Split between two tails: $\alpha/2 = 0.05/2 = 0.025$**

So, each tail has an area of 0.025.

---

3 Use Standard Normal Table (Z-Table)

- **We need Z such that the area to the left of -Z = 0.025 (left tail)**
- **Or equivalently, area to the left of +Z = 1 - 0.025 = 0.975 (right tail)**

Look up 0.975 in the cumulative Z-table → $Z \approx 1.96$

Hence, for a two-tailed test at α = 0.05, the critical Z-values are ±1.96.

---

4 Recap

$$\text{Two-tailed test critical Z} = \pm Z_{\alpha/2} = \pm Z_{0.025} = \pm 1.96$$

- **Left tail rejection: $Z < -1.96$**
- **Right tail rejection: $Z > +1.96$**
- **Middle region: $-1.96 \le Z \le 1.96$ → fail to reject H_0**

6 Example: Two-Tailed Z-Test

Problem:
The average IQ of students in a college is claimed to be 100. A sample of 36 students has a mean IQ of 104. Population $\sigma = 12$. Test at $\alpha = 0.05$.

Solution:

1. $H_0: \mu = 100$, $H_1: \mu \neq 100$
2. $Z = (104 - 100) / (12/\sqrt{36}) = 4 / 2 = 2$
3. Critical value for $\alpha=0.05$ (two-tailed) = ± 1.96
4. Decision: $Z = 2 > 1.96$ → Reject H_0
5. Conclusion: Mean IQ is **significantly different from 100**.

---

7 Summary

- **Hypothesis testing** helps in making decisions using sample data.
- **Types:** left-tailed, right-tailed, two-tailed.
- **Z-test** is used for known σ and large samples.
- Always state **H_0 and H_1**, calculate **Z**, and compare with **critical value**.
- It is widely used in **quality control, business, medicine, and social sciences**.

Hypothesis Testing Problems with Solutions

---

1 Left-Tailed Test

Problem:
A factory claims that the mean weight of sugar packets is **5 kg**. A sample of **36 packets** has a mean of **4.85 kg**. The population standard deviation is **0.5 kg**. Test at **5% significance level** if the mean weight is **less than 5 kg**.

Solution:

Step 1: Hypotheses

- H_0: $\mu = 5$
- H_1: $\mu < 5$

Step 2: Z-Statistic

$$Z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} = \frac{4.85 - 5}{0.5/\sqrt{36}} = \frac{-0.15}{0.0833} \approx -1.8$$

Step 3: Critical Value

- Left-tailed test at $\alpha = 0.05$ → $Z = -1.645$

Step 4: Decision

- $Z = -1.8 < -1.645$ → **Reject H_0**

Step 5: Conclusion
The mean weight of sugar packets is **significantly less than 5 kg**.

---

2 Right-Tailed Test

Problem:
A bottling machine claims to fill bottles with **1 liter**. A sample of **49 bottles** shows a mean fill of **1.04 liters**. Population standard deviation is **0.2 liters**. Test at **1% significance level** if the machine fills **more than 1 liter**.

Solution:

Step 1: Hypotheses

- H_0: $\mu = 1$

- H_1: μ > 1

Step 2: Z-Statistic

$$Z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} = \frac{1.04 - 1}{0.2/\sqrt{49}} = \frac{0.04}{0.02857} \approx 1.4$$

Step 3: Critical Value

- Right-tailed test at α = 0.01 → Z = 2.33

Step 4: Decision

- Z = 1.4 < 2.33 → **Fail to reject H_0**

Step 5: Conclusion
There is **no significant evidence** that the machine fills more than 1 liter.

---

3 Two-Tailed Test

Problem:
The average IQ of students is claimed to be **100**. A random sample of **36 students** gives a mean IQ of **104**. Population standard deviation is **12**. Test at **5% significance level** whether the mean IQ differs from 100.

Solution:

Step 1: Hypotheses

- H_0: μ = 100
- H_1: μ ≠ 100

Step 2: Z-Statistic

$$Z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} = \frac{104 - 100}{12/\sqrt{36}} = \frac{4}{2} = 2$$

Step 3: Critical Value

- Two-tailed test at α = 0.05 → Z = ±1.96

Step 4: Decision

- Z = 2 > 1.96 → **Reject H_0**

Step 5: Conclusion
The mean IQ is **significantly different from 100**.

Practice Questions: Hypothesis Testing (Z-Test)

---

1 Left-Tailed Test

1. A company claims that the average lifespan of its LED bulbs is **1,200 hours**. A sample of **36 bulbs** shows an average lifespan of **1,180 hours**. The population standard deviation is **60 hours**. Test at **5% significance level** if the bulbs last **less than claimed**.
2. The mean sugar content in a brand of soda is claimed to be **40 g per bottle**. A sample of **49 bottles** shows a mean of **39 g**. Population $\sigma = 2$ g. Test at **1% significance level** if the sugar content is **less than claimed**.

---

2 Right-Tailed Test

3. A factory claims that the average weight of rice bags is **50 kg**. A sample of **64 bags** shows a mean weight of **51 kg**. Population standard deviation is **4 kg**. Test at **5% significance level** if the bags weigh **more than claimed**.
4. A machine fills milk bottles with **1 liter**. A sample of **25 bottles** shows a mean fill of **1.02 liters**. Population $\sigma = 0.03$ liters. Test at **1% significance level** if the machine fills **more than 1 liter**.

---

3 Two-Tailed Test

5. The average IQ of students in a university is claimed to be **100**. A sample of **36 students** has a mean IQ of **105**. Population standard deviation is **12**. Test at **5% significance level** whether the mean IQ is **different from 100**.
6. A company claims that its battery lasts **500 hours on average**. A sample of **49 batteries** has a mean lifetime of **510 hours**. Population $\sigma = 20$ hours. Test at **1% significance level** whether the mean lifetime **differs from the claim**.

---

4 Extra Challenge

7. A manufacturer claims that **60%** of its products pass quality inspection. A sample of **200 products** shows **114 passed**. Test at **5% significance level** if the true proportion **differs from 60%**.

8. The average height of students in a college is **165 cm**. A sample of **36 students** has a mean height of **168 cm**. Population $\sigma = 6$ cm. Test at **5% significance level** whether the mean height **differs from 165 cm**.

Tips for Solving These Questions

1. Identify **H_0 and H_1**.

2. Determine the **type of test**: left, right, or two-tailed.

3. Compute the **Z-statistic** using:

$$Z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}}$$

4. Compare Z with **critical value** (or find p-value).

5. Make the **decision** and state the **conclusion in context**.