Testing of Hypothesis - Basics

1. Hypothesis:

A hypothesis is a testable statement or claim about a population parameter. It's an educated guess or a proposed explanation that can be investigated through data analysis.

2. Null Hypothesis (H_0):

- The null hypothesis is a statement of "no effect" or "no difference." It represents the status quo or the default assumption.
- Crucially, the null hypothesis *always* involves an equality sign (=). It states that there is no significant difference or relationship.

Example:

"The average weight of apples from this orchard is exactly 150 grams." (H_0: $\mu = 150$)

3. Alternative Hypothesis (H_1 or H_a):

- The alternative hypothesis contradicts the null hypothesis. It represents the researcher's belief or the effect they are trying to find.
- It uses inequality signs ($\neq$, $>$, $<$).

Example (continuing from above):

- "The average weight of apples from this orchard is not equal to 150 grams." (H_1: $\mu \neq 150$) (two-tailed)
- "The average weight of apples from this orchard is greater than 150 grams." (H_1: $\mu > 150$) (right-tailed)
- "The average weight of apples from this orchard is less than 150 grams." (H_1: $\mu < 150$) (left-tailed)

4. Type I and Type II Errors:

- **Type I Error (False Positive):**
  - This occurs when you reject the null hypothesis when it is actually true.
  - The probability of a Type I error is denoted by alpha (α).
- **Type II Error (False Negative):**
  - This occurs when you fail to reject the null hypothesis when it is false.
  - The probability of a Type II error is denoted by beta (β).

ACTION	**H_0 IS ACTUALLY**	...
	True	False
Do not reject H_0	Correct Outcome	Type II error
Reject H_0	Type I Error	Correct Outcome

5. Level of Significance (Alpha, α):

- The level of significance (alpha) is the probability of making a Type I error.
- It is a pre-determined threshold set by the researcher. Commonly used values are 0.05 (5%) and 0.01 (1%).
- A lower alpha value reduces the risk of a Type I error but increases the risk of a Type II error.

6. Power of the Test (1 - Beta, $1 - \beta$):

- The power of a test is the probability of correctly rejecting the null hypothesis when it is false.
- It is calculated as $1 - \beta$.
- A higher power means the test is more likely to detect a real effect.
- Factors that affect power include:
  - Sample size: Larger samples increase power.
  - Effect size: Larger effects are easier to detect.
  - Level of significance (alpha): Increasing alpha increases power (but also increases the risk of a Type I error).
  - Variability in the data. less variability increases power.

7. Test Statistic:

In statistical hypothesis testing, a test statistic is a crucial value calculated from sample data. It's used to determine whether to reject the null hypothesis.

Core Idea:

- A test statistic essentially summarizes the difference between what you observed in your data and what you would expect to observe if the null hypothesis were true.
- It measures how far your sample data deviates from the null hypothesis.
- By comparing the test statistic to a known probability distribution, you can determine the likelihood of obtaining your results if the null hypothesis were correct.
- The interpretation of a test statistic depends on its probability distribution. Common distributions include the standard normal (z), t, chi-square, and F distributions.

Common Examples:

- **Z-statistic:**
  - Used when the population standard deviation is known or with large sample sizes.
  - Measures how many standard deviations the sample mean is from the population mean.
- **t-statistic:**
  - Used when the population standard deviation is unknown and with smaller sample sizes.
  - Similar to the z-statistic but accounts for the uncertainty introduced by estimating the standard deviation.
- **Chi-square statistic:**
  - Used for categorical data to test for associations or goodness of fit.
- **F-statistic:**
  - Used in analysis of variance (ANOVA) to compare the variances of two or more groups.

8. Critical Region

- The critical region is the set of all values of the test statistic for which the null hypothesis will be rejected.
- Essentially, it's the area within a probability distribution where, if the test statistic falls, you conclude that the null hypothesis is unlikely to be true.

Relationship to Significance Level:

The size of the critical region is determined by the significance level (alpha, α).For example, if $\alpha = 0.05$, the critical region will encompass 5% of the distribution's area.

Critical Values:

The critical values are the boundaries of the critical region. They are the points that separate the region where you reject the null hypothesis from the region where you fail to reject it. These values are obtained from statistical tables (like the z-table or t-table) or statistical software.

Types of Tests and Critical Regions:

Two-tailed tests:

The critical region is divided into two parts, one in each tail of the distribution. Used when the alternative hypothesis states that the population parameter is "not equal to" a specific value.

One-tailed tests:

The critical region is located in only one tail of the distribution (either the right or the left). Used when the alternative hypothesis states that the population parameter is "greater than" or "less than" a specific value.

9. p-value

The p-value, or probability value, is the probability of obtaining test results at least as extreme as the results actually observed, assuming that the null hypothesis is correct.

In simpler terms, it tells you how likely it is that you would see the data you observed if the null hypothesis were true.

Key Interpretations:

Strength of Evidence:

A small p-value indicates strong evidence against the null hypothesis. It suggests that the observed results are unlikely to have occurred by chance alone. A large p-value indicates weak evidence against the null hypothesis. It suggests that the observed results could easily have occurred by chance.

Decision Making:

Researchers typically compare the p-value to a predetermined significance level (alpha, α), often 0.05.

If the p-value is less than or equal to α, the null hypothesis is rejected. This is often referred to as a "statistically significant" result.

If the p-value is greater than α, the null hypothesis is not rejected.

Important Considerations:

Not the Probability of the Null Hypothesis: It's crucial to understand that the p-value is not the probability that the null hypothesis is true. It's the probability of the observed data (or more extreme data) occurring, given that the null hypothesis is true.

Testing for single mean

Testing a single mean is a fundamental statistical procedure used to determine whether the average value of a population is equal to, greater than, or less than a specific hypothesized value. Here's a more detailed explanation:

Purpose:

- The primary goal is to compare a sample mean to a known or hypothesized population mean.
- This helps researchers determine if observed differences are statistically significant or simply due to random chance.

Key Components:

1.Hypotheses:

1. Null Hypothesis (H_o):
    1. States that there is no significant difference between the sample mean and the hypothesized population mean.
    2. Example: H_o: $\mu = \mu_o$ (where μ is the population mean and μ_o is the hypothesized mean).
2. Alternative Hypothesis (H_1 or H_a):
    1. Contradicts the null hypothesis, suggesting a significant difference.
    2. Can be one-tailed (directional) or two-tailed (non-directional).
    3. Examples:
        1. H_1: $\mu \neq \mu_o$ (two-tailed)
        2. H_1: $\mu > \mu_o$ (right-tailed)
        3. H_1: $\mu < \mu_o$ (left-tailed)

2. Which test to use:

Z-test:

Population standard deviation known and/or large sample size ($n \geq 30$).

t-test:

Population standard deviation unknown and small sample size ($n < 30$).

3. Test Statistic

Z-test (population standard deviation is known):

$$z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} \text{ (follows standard normal distribution)}$$

Z-test (population standard deviation is unknown & $n \geq 30$):

$$z = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} \text{ (follows standard normal distribution)}$$

t-test (population standard deviation is unknown & $n < 30$):

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} \text{ (follows student t distribution with degrees of freedom n-1)}$$

$\bar{x}$: sample mean, μ_0: Hypothesized mean, σ : population standard deviation,
s : Sample standard deviation, n : sample size

Note: The conclusion on whether or not to reject H_o can be drawn either using critical region or using p-value.

4. Critical Region:

Let the level of significance be α.

Note the notations z_α and $t_\alpha(m)$ are defined as:
$P(z > z_\alpha) = \alpha$ (Standard Normal Distribution – can be obtained from standard normal table)
$P(t > t_\alpha(m)) = \alpha$ (Student t Distribution with degrees of freedom m – can be obtained from t table)

In Z-test (z denotes the test statistic (both cases : standard deviation is known and unknown)):

H_1	Critical Region/ Region of rejection of H_o
H_1: $\mu \neq \mu_o$ (two-tailed)	$z < -z_{\alpha/2}$ or $z > z_{\alpha/2}$
H_1: $\mu > \mu_o$ (right-tailed)	$z > z_\alpha$
H_1: $\mu < \mu_o$ (left-tailed)	$z < -z_\alpha$

In t-test (t denotes the test statistic):

H_1	Critical Region/ Region of rejection of H_0
$H_1: \mu \neq \mu_0$ (two-tailed)	$t < -t_{\alpha/2}(n-1)$ or $t > t_{\alpha/2}(n-1)$
$H_1: \mu > \mu_0$ (right-tailed)	$t > t_\alpha(n-1)$
$H_1: \mu < \mu_0$ (left-tailed)	$t < -t_\alpha(n-1)$

5. p-value:

Let the level of significance be α.
If the p-value is less than or equal to α, the null hypothesis is rejected.
If the p-value is greater than α, the null hypothesis is not rejected.

In Z-test (z denotes the test statistic (both cases : standard deviation is known and unknown) and Z denotes a general standard normal random variable):

H_1	p-value
$H_1: \mu \neq \mu_0$ (two-tailed)	2* $P(Z < z)$ if z is negative 2* $P(Z > z)$ if z is positive
$H_1: \mu > \mu_0$ (right-tailed)	$P(Z > z)$
$H_1: \mu < \mu_0$ (left-tailed)	$P(Z < z)$

In t-test (t denotes the test statistic and T denotes a general student t random variable with degrees of freedom n-1):

H_1	p-value
$H_1: \mu \neq \mu_0$ (two-tailed)	2* $P(T < t)$ if z is negative 2* $P(T > t)$ if z is positive
$H_1: \mu > \mu_0$ (right-tailed)	$P(T > t)$
$H_1: \mu < \mu_0$ (left-tailed)	$P(T < t)$

Suppose a baker claims that his bread height is more than 15 cm, on the average. Several of his customers do not believe him. To persuade his customers that he is right, the baker decides to do a hypothesis test. He bakes 10 loaves of bread. The average height of the sample loaves is 17 cm. The baker knows from baking hundreds of loaves of bread that the **standard deviation** for the height is 0.5 cm.

Solution:

$H_0: \mu = 15$ ($\mu_0 = 15$, the hypothesized mean)

$H_1: \mu < 15$ (the claim of the customers)

← Left-tailed test

Given,

$n = 10$, $\bar{x} = 17$

$\sigma = 0.5 \Rightarrow \sigma$ is known $\Rightarrow$ z test.

The test statistic:

$$z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} = \frac{17-15}{0.5/\sqrt{10}} = 12.64$$

Critical Region:

$H_1: \mu < 15$ ∴ reject H_0 if $z < -z_\alpha$. Since α is not given, assume $\alpha = 0.05$.

From normal table, $-z_{0.05} = -1.645$. Clearly $z = 12.64 \nless -1.645$

∴ We fail to reject H_0.

Jeffrey, as an eight-year old, **established an average time of 16.43 seconds** for swimming the 25-yard freestyle, with a **standard deviation of 0.8 seconds**. His dad, Frank, thought that Jeffrey could swim the 25-yard freestyle faster by using goggles. Frank bought Jeffrey a new pair of expensive goggles and timed Jeffrey for **15 25-yard freestyle swims**. For the 15 swims, **Jeffrey's average time was 16 seconds. Frank thought that the goggles helped Jeffrey to swim faster than the 16.43 seconds.** Conduct a hypothesis test using a preset $\alpha = 0.05$. Assume that the swim times for the 25-yard freestyle are normal.

Solution:

$H_0: \mu = 16.43$ [No affect of goggle]

$H_1: \mu < 16.43$ [Goggles made the kid faster]

← Left-tailed test

Given,

$n = 15$, $\bar{x} = 16$

$\sigma = 0.8 \Rightarrow \sigma$ known $\Rightarrow$ z test.

Test statistic:

$$z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} = \frac{16 - 16.43}{0.8/\sqrt{15}} = -2.08$$

Critical region:

$H_1: \mu < 16.43 \Rightarrow$ reject H_0 if $z < -z_\alpha$. Given that $\alpha = 0.05$

$\therefore -z_\alpha = -1.645$

$z = -2.08 < -1.645 = -z_\alpha$.

$\therefore$ Reject H_0.

Using p-value:

$H_1: \mu < \mu_0 \Rightarrow$ p-value $= P(z < -2.08) = 0.0188 < 0.05 = \alpha$

Since p-value $< \alpha$, we reject H_0.

Statistics students believe that the average score on the first statistics test is 65. A statistics instructor thinks the average score is higher than 65. He samples ten statistics students and obtains the scores 65; 65; 70; 67; 66; 63; 63; 68; 72; 71. He performs a hypothesis test using a 5% level of significance. The data are from a normal distribution.

Solution:

$H_0: \mu = 65$ v/s $H_1: \mu > 65$

← Right-tailed test

Given,

$n = 10$, $\alpha = 0.05$

σ not known & $n < 30 \Rightarrow$ t test.

Test statistic:

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} \sim t(n-1)$$

$n = 10$, $\bar{x} = \frac{\Sigma x_i}{n}$, $s = \sqrt{\frac{\Sigma (x_i - \bar{x})^2}{n-1}}$

$\bar{x} = \frac{670}{10} = 67$ $s = \sqrt{\frac{92}{9}} = 3.19$

x	$x - \bar{x}$	$(x-\bar{x})^2$
65	-2	4
65	-2	4
70	3	9
67	0	0
66	-1	1
63	-4	16
63	-4	16
68	1	1
72	5	25
71	4	16
Sum: 670		92

$\Rightarrow \bar{x} = 67$

$$\therefore t = \frac{67 - 65}{3.19/\sqrt{10}} = \underline{\underline{1.98}}$$

Critical Region:

$H_1: \mu > 65 \Rightarrow$ Reject H_0 if $t > t_\alpha(9)$

Given $\alpha = 0.05$

$\therefore t_{0.05}^{(9)} = 1.833$

$t = 1.98 > 1.833 = t_{0.05}^{(9)}$

$\therefore$ Reject H_0.

P-value:

$H_1: \mu > 65 \Rightarrow$ p-value $= P(t > 1.98) < 0.05 = \alpha$.

$\therefore$ We reject H_0

If the claim of the instructor is an average score < 65,

$H_0: \mu = 65$
$H_1: \mu < 65$

test statistic: $t = 1.98$

Critical region: $t < -t_\alpha(9) = -1.833$

$t \nless 1.833$ $\therefore$ We fail to reject H_0 against $H_1: \mu < 65$.