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 (≠, >, <).

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 - β):

- 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:

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

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

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

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

$\Rightarrow \bar{x} = 67$

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

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 $H_1: \mu < 65$.

Practice Problems

- A random sample of 100 recorded in the united states during the past year showed an average life span of 71.8 years. Assuming a population standard deviation of 8.9 years, does this seem to indicate that the mean life span today is greater than 70 years? Use a 0.05 level of significance.

- The mean lifetime of a sample of 100 fluorescent light bulbs produced by a company is computed to be 1570 hours with a standard deviation of 120 hours. If μ is the mean lifetime of all the bulbs produced by the company, test the hypothesis $\mu = 1600$ hours against the alternative hypothesis $\mu \neq 1600$ hours, using the level of significance of 0.05. Find the P value of the test.

- A manufacturer of sports equipment has developed a new synthetic fishing line that the company claims has a mean breaking strength of 8 kilograms against the alternative that $\mu \neq 8$ kilograms if a random sample of 50 lines is tested and found to have a mean breaking strength of 7.8 kilograms. Use a 0.01 level of signinficance

Testing for single proportion

A proportion is often used to describe the fraction of a population that possesses a certain characteristic. For instance, "the proportion of voters who support candidate X" represents the number of voters who support the candidate divided by the total number of voters.

Purpose:

- To determine if a sample proportion significantly differs from a hypothesized population proportion.
- This helps researchers determine if observed differences are statistically significant or simply due to random chance.

Key Components:

1.Hypotheses:

1. Null Hypothesis (H_0):
   States that there is no significant difference between the sample proportion and the hypothesized population proportion.
   Example: H_0: $p = p_0$ (where p is the population proportion and p_0 is the hypothesized proportion).
2. Alternative Hypothesis (H_1 or H_a):
   Contradicts the null hypothesis, suggesting a significant difference.
   Can be one-tailed (directional) or two-tailed (non-directional).
   Examples:
   1. H_1: $p \neq p_0$ (two-tailed)
   2. H_1: $p > p_0$ (right-tailed)
   3. H_1: $p < p_0$ (left-tailed)

2. Which test to use:

Z-test:

Large sample size ($n \geq 30$).

t-test:

Small sample size ($n < 30$).

3. Test Statistic

$\hat{p}$: sample proportion, p_0: hypothesized proportion, $q_0 = 1 - p_0$, n : sample size

Z-test :

$$z = \frac{\hat{p} - p_0}{\sqrt{(p_0 q_0)/n}}$$ (follows standard normal distribution),

t-test:

$$t = \frac{\hat{p} - p_0}{\sqrt{(p_0 q_0)/n}}$$ (follows student t distribution with degrees of freedom n-1),

Note: The conclusion on whether or not to reject H_0 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):

H_1	Critical Region/ Region of rejection of H_0
H_1: $p \neq p_0$ (two-tailed)	$z < -z_{\alpha/2}$ or $z > z_{\alpha/2}$
H_1: $p > p_0$ (right-tailed)	$z > z_\alpha$
H_1: $p < p_0$ (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: $p \neq p_0$ (two-tailed)	$t < -t_{\alpha/2}(n-1)$ or $t > t_{\alpha/2}(n-1)$
H_1: $p > p_0$ (right-tailed)	$t > t_\alpha(n-1)$
H_1: $p < p_0$ (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 and Z denotes a general standard normal random variable):

H_1	p-value
H_1: $p \neq p_0$ (two-tailed)	2* $P(Z < z)$ if z is negative 2* $P(Z > z)$ if z is positive
H_1: $p > p_0$ (right-tailed)	$P(Z > z)$
H_1: $p < p_0$ (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: $p \neq p_0$ (two-tailed)	2* $P(T < t)$ if z is negative 2* $P(T > t)$ if z is positive
H_1: $p > p_0$ (right-tailed)	$P(T > t)$
H_1: $p < p_0$ (left-tailed)	$P(T < t)$

Problem 1:

A marketing expert for a pasta-making company believes that 40% of pasta lovers prefer lasagna. If 9 out of 20 pasta lovers choose lasagna over other pastas, what can be concluded about the expert's claim? Use a 0.05 level of significance.

Here we are interested in the proportion of people who prefers lasagna over other pastas, out of all pasta lovers.

H_0: $P = 0.4$

H_1: $P \neq 0.4$ $\qquad P_0 = 0.4 \quad \therefore q_0 = 1 - P_0 = 1 - 0.4 = 0.6$

$\alpha = 0.05$

$n = 20 < 30$ — $\therefore$ t test.

$\hat{p} = 9/20$

Test statistic, $t = \dfrac{\hat{p} - P_0}{\sqrt{\frac{P_0 q_0}{n}}} \sim t(n-1)$

$$= \frac{9/20 - 0.4}{\sqrt{\frac{0.4 \times 0.6}{20}}} = 0.4564$$

$H_1: p \neq 0.4$ $\therefore$ Critical region: $t < -t_{\alpha/2}(n-1)$ OR $t > t_{\alpha/2}(n-1)$

$$t_{0.025}(19) = 2.093$$

$t = 0.4564 \not> t_{\alpha/2}(n-1)$ $\quad t \not< -t_{\alpha/2}(n-1)$

$\therefore$ t does not belong to the critical region.

$\therefore$ We fail to reject H_0. $\therefore$ We don't have enough evidence to reject the claim that 40% prefers lasagna.

Problem A new radar device is being considered for a certain missile defense system. The system is checked by experimenting with aircraft in which a kill or a no kill is simulated. If, in 300 trials, 250 kills occur, accept or reject, at the 0.04 level of significance, the claim that the probability of a kill with the new system does not exceed the 0.8 probability of the existing device.

$\Rightarrow$

$H_0: p = 0.8$

v/s

$H_1: p < 0.8$

$p_0 = 0.8 \Rightarrow q_0 = 1 - p_0 = 0.2.$

$\alpha = 0.04$

$n = 300 \geq 30$ $\therefore$ Z test

$\hat{p} = \frac{250}{300} = 5/6$

Test statistic:

$$Z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0 q_0}{n}}} = \frac{0.8333 - 0.8}{\sqrt{\frac{0.8 \times 0.2}{300}}} = 1.4433$$

$H_1: P < 0.8$

$\therefore$ critical region: $z < -z_\alpha$

$z_{0.04} = 1.75$

$z = 1.4433 \nless -1.75$

$\therefore$ z does not lie on the critical region.

$\therefore$ We fail to reject H_0 //.

In a hospital 480 females and 520 male babies were born in a week. Do these figures conform the hypothesis that males and females are born in equal number?

Solution:

Let us take the hypothesis that the male and female babies are born in equal number. Then P=Q=1/2

Population proportion $P = 0.5$ (probability of male happening)

$Q = 0.5$ (probability of female happening)

- Let the null hypothesis be

$$H_o: P = 0.5 \ (P_0 = 0.5)$$

Here Proposition is male and female babies are born in equal number (i.e., P=Q=1/2) the alternative hypothesis be

$$H_1: P \neq 0.5 \quad \text{(two tailed test)}$$

- here total number of births be: $n = 480 + 520 = 1000$
- Actual number of female babies born=480

Actual probability of female happening from given data is

$$\text{Proportion of females born } p = \frac{\text{Actual number of female babies born}}{\text{total number of births}} = \frac{480}{1000} = 0.48$$

➢**Level of significance**:

at 5%

the interval of estimate is $(-1.96, 1.96)$

➢**Test statistic:**

$$z = \frac{p - P}{\left(\sqrt{\frac{PQ}{n}}\right)}$$

$$z = \frac{0.48 - 0.5}{\sqrt{\frac{(0.5)(0.5)}{1000}}} = -1.265$$

here statistic $z = -1.265 \in (-1.96, 1.96)$

Hence at 5% level of significance, we accept the null hypothesis that the males and females are born in equal proportions.

Practice Problems

- A survey was conducted among the citizens of a city to study their preference towards consumption of tea and coffee. Among 1000 randomly selected persons, it is found that 560 are tea-drinkers and the remaining are coffee-drinkers. Can we conclude at 1% level of significance from this information that both tea and coffee are equally preferred among the citizens in the city?

- A survey was conducted among the citizens of a city to study their preference towards consumption of tea and coffee. Among 1000 randomly selected persons, it is found that 560 are tea-drinkers and the remaining are coffee-drinkers. Can we conclude at 1% level of significance from this information that both tea and coffee are equally preferred among the citizens in the city?
- The manufacturer of a patent medicine claimed that it was 90% effective in relieving an allergy for a period of 8 hours. In a sample of 200 people who had the allergy, the medicine provided relief for 160 people. Determine whether the manufacturer's claim is legitimate by using 0.01 as the level of significance. Find the P value of the test.

Testing for comparison of proportions

Let's break down hypothesis testing for comparing proportions using z and t-tests, covering all those scenarios.

Purpose:

- The primary goal is to compare the difference of sample proportions to a known or hypothesized difference of population proportions of two different populations.
- 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 population means or a specific difference (d) is there between population means.

   Example: H_o: $p_1 - p_2 = 0$ ($p_1 = p_2$) or $p_1 - p_2 = d$

   2. Alternative Hypothesis (H_1 or H_a):

   Contradicts the null hypothesis, suggesting a significant difference.

   Can be one-tailed (directional) or two-tailed (non-directional).

   Examples:
   1. H_1: $p_1 - p_2 \neq 0$ ($p_1 \neq p_2$) or $p_1 - p_2 \neq d$ (two-tailed)
   2. H_1: $p_1 - p_2 > 0$ ($p_1 > p_2$) or $p_1 - p_2 > d$ (right-tailed)
   3. H_1: $p_1 - p_2 < 0$ ($p_1 < p_2$) or $p_1 - p_2 < d$ (left-tailed)

2. Which test to use:

Z-test

3. Test Statistic

$\widehat{p_1}$: sample proportion from population 1, $\widehat{p_2}$: sample proportion from population 2

d: Hypothesized difference between population proportions (is equal to 0 if H_o: $p_1 - p_2 = 0$ ($p_1 = p_2$))

n_1 : sample size from population1, n_2 : sample size from population2

x_1: number of favorable sample points out of sample from population 1

x_2: number of favorable sample points out of sample from population 2

$$z = \frac{\hat{p}_1 - \hat{p}_2 - d}{\sqrt{\hat{p}\hat{q}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \text{ (follows standard normal distribution)}$$

$$\hat{p} = \frac{x_1 + x_2}{n_1 + n_2} \qquad \hat{q} = 1 - \hat{p}$$

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 notation z_α is defined as:

$P(z > z_\alpha) = \alpha$ (Standard Normal Distribution – can be obtained from standard normal table)

H_1	Critical Region/ Region of rejection of H_o
H_1: $p_1 - p_2 \neq 0$ ($p_1 \neq p_2$) or $p_1 - p_2 \neq d$ (two-tailed)	$z < -z_{\alpha/2}$ or $z > z_{\alpha/2}$
H_1: $p_1 - p_2 > 0$ ($p_1 > p_2$) or $p_1 - p_2 > d$ (right-tailed)	$z > z_\alpha$
H_1: $p_1 - p_2 < 0$ ($p_1 < p_2$) or $p_1 - p_2 < d$ (left-tailed)	$z < -z_\alpha$

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 and Z denotes a general standard normal random variable):

H_1	P-value
H_1: $p_1 - p_2 \neq 0$ ($p_1 \neq p_2$) or $p_1 - p_2 \neq d$ (two-tailed)	2* $P(T < t)$ if z is negative 2* $P(T > t)$ if z is positive
H_1: $p_1 - p_2 > 0$ ($p_1 > p_2$) or $p_1 - p_2 > d$ (right-tailed)	$P(T > t)$
H_1: $p_1 - p_2 < 0$ ($p_1 < p_2$) or $p_1 - p_2 < d$ (left-tailed)	$P(T < t)$

A vote is to be taken among the residents of a town and the surrounding county to determine whether a proposed chemical plant should be constructed. The construction site is within the town limits, and for this reason many voters in the county believe that the proposal will pass because of the large proportion of town voters who favor the construction. To determine if there is a significant difference in the proportions of town voters and county voters favoring the proposal, a poll is taken. If 120 of 200 town voters favor the proposal and 240 of 500 county residents favor it, would you agree that the proportion of town voters favoring the proposal is higher than the proportion of county voters? Use an α = 0.05 level of significance.

Let p1 and p2 be the true proportions of voters in the town and county, respectively, favoring the proposal.

1. H_0: $p_1 = p_2$.
2. H_1: $p_1 > p_2$.
3. $\alpha = 0.05$.
4. Critical region: $z > 1.645$.

$$z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}\hat{q}(1/n_1 + 1/n_2)}}. \qquad \hat{p} = \frac{x_1 + x_2}{n_1 + n_2}$$

$$\hat{p}_1 = \frac{x_1}{n_1} = \frac{120}{200} = 0.60, \quad \hat{p}_2 = \frac{x_2}{n_2} = \frac{240}{500} = 0.48, \quad \text{and}$$
$$\hat{p} = \frac{x_1 + x_2}{n_1 + n_2} = \frac{120 + 240}{200 + 500} = 0.51.$$

$$z = \frac{0.60 - 0.48}{\sqrt{(0.51)(0.49)(1/200 + 1/500)}} = 2.9,$$
$$P = P(Z > 2.9) = 0.0019.$$

This is the p-value, which is less than alpha = 0.05. Thus H0 is rejected.

The critical region will be z > 1.645. 2.9>1.645. Thus the test statistic .ies on the critical region and hence we reject the null hypothesis.

Decision: Reject H0 and agree that the proportion of town voters favoring the proposal is higher than the proportion of county voters.

In a study on the fertility of married women conducted by Martin O'Connell and Carolyn C. Rogers for the Census Bureau in 1979, two groups of childless wives aged 25 to 29 were selected at random, and each was asked if she eventually planned to have a child. One group was selected from among wives married less than two years and the other from among wives married five years. Suppose that 240 of the 300 wives married less than two years planned to have children some day compared to 288 of the 400 wives married five years. Can we conclude that the proportion of wives married less than two years who planned to have children is significantly higher than the proportion of wives married five years?

⋆ Let n_1 and n_2 be the sample sizes for the two groups respectively.
It is given $n_1 = 300$; $n_2 = 400$

⋆ Let x_1 be the number of wives married less than two years who planned to have children in the sample. It is given $x_1 = 240$.

⋆ Let x_2 be the number of wives married five years who planned to have children in the sample. It is given $x_2 = 288$

- The null hypothesis is: $H_0 : p_1 = p_2$

- The alternative hypothesis is: $H_1 : p_1 > p_2$

The level of significance is $\alpha = 0.05$

⋆ Let p_1 be the true population proportion of wives married less than two years who planned to have children.

$$\hat{p}_1 = \frac{x_1}{n_1} = \frac{240}{300} = 0.8$$

⋆ Let p_2 be the true population proportion of wives married five years respectively.

$$\hat{p}_2 = \frac{x_2}{n_2} = \frac{288}{400} = 0.72$$

$$\hat{p} = \frac{x_1 + x_2}{n_1 + n_2} = \frac{240 + 288}{300 + 400} = 0.75$$

The test statistic is

$$z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

$$z = \frac{0.8 - 0.72}{\sqrt{0.75 \times (1 - 0.75)\left(\frac{1}{300} + \frac{1}{400}\right)}}$$

$$= \frac{0.08}{0.0329}$$

$$= 2.433$$

Decision?

Practice Problem

An urban community would like to show that the incidence of breast cancer is higher in their area than in a nearby rural area. (PCB levels were found to be higher in the soil of the urban community.) If it is found that 20 of 200 adult women in the urban community have breast cancer and 10 of 150 adult women in the rural community have breast cancer, can we conclude at the 0.05 level of significance that breast cancer is more prevalent in the urban community?

Testing for comparison of means

Let's break down hypothesis testing for comparing means using z and t-tests, covering all those scenarios.

Purpose:

- The primary goal is to compare the difference of sample means to a known or hypothesized difference of population means of two different populations.
- This helps researchers determine if observed differences are statistically significant or simply due to random chance.

Key Components:

1. Hypotheses:

1. Null Hypothesis (H_0):
   1. States that there is no significant difference between the population means or a specific difference (d) is there between population means.

   Example: H_0: $\mu_1 - \mu_2 = 0$ ($\mu_1 = \mu_2$) or $\mu_1 - \mu_2 = d$

   2. Alternative Hypothesis (H_1 or H_a):

   Contradicts the null hypothesis, suggesting a significant difference.

   Can be one-tailed (directional) or two-tailed (non-directional).

   Examples:

   1. H_1: $\mu_1 - \mu_2 \neq 0$ ($\mu_1 \neq \mu_2$) or $\mu_1 - \mu_2 \neq d$ (two-tailed)
   2. H_1: $\mu_1 - \mu_2 > 0$ ($\mu_1 > \mu_2$) or $\mu_1 - \mu_2 > d$ (right-tailed)
   3. H_1: $\mu_1 - \mu_2 < 0$ ($\mu_1 < \mu_2$) or $\mu_1 - \mu_2 < d$ (left-tailed)

2. Which test to use:

Z-test:

Population standard deviations are known and/or large sum of sample sizes ($n_1+n_2 \geq 30$).

t-test:

Population standard deviations are unknown and small sum of sample sizes ($n_1+n_2 < 30$).

3. Test Statistic

$\overline{x_1}$: sample mean from population 1, $\overline{x_2}$: sample mean from population 2

d: Hypothesized difference between population means (is equal to 0 if H_0: $\mu_1 - \mu_2 = 0$ ($\mu_1 = \mu_2$))

σ_1 : population standard deviation of population1, σ_2 : population standard deviation of population2,

s_1 : Sample standard deviation from population 1, s_2 : Sample standard deviation from population 2,

n_1 : sample size from population1, n_2 : sample size from population2

Z-test (population standard deviation is known):

$$z = \frac{\overline{x_1} - \overline{x_2}}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \quad or \quad \frac{\overline{x_1} - \overline{x_2} - d}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \text{ (follows standard normal distribution)}$$

Z-test (population standard deviation is unknown & $n_1+n_2 \geq 30$):

$$z = \frac{\overline{x_1} - \overline{x_2}}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \quad or \quad \frac{\overline{x_1} - \overline{x_2} - d}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \text{ (follows standard normal distribution)}$$

t-test (population standard deviation is unknown but having equal population variance ($\sigma_1^2 = \sigma_2^2$) & $n_1+n_2 < 30$):

$$t = \frac{\overline{x_1} - \overline{x_2}}{s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \quad or \quad \frac{\overline{x_1} - \overline{x_2} - d}{s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \text{ (follows student t distribution with degrees of freedom } m = n_1+n_2-2)$$

Here $s_p^2 = \frac{s_1^2(n_1-1)+s_2^2(n_2-1)}{n_1+n_2-2}$

t-test (population standard deviation is unknown but having unequal population variance ($\sigma_1^2 \neq \sigma_2^2$) & $n_1+n_2 < 30$):

$$t = \frac{\overline{x_1} - \overline{x_2}}{\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}} \quad or \quad \frac{\overline{x_1} - \overline{x_2} - d}{\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}} \text{ (follows student t distribution with degrees of freedom } m)$$

$$\text{Here } m = \frac{\left(\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}\right)^2}{\left(\frac{\left(s_1^2/n_1\right)^2}{(n_1-1)}+\frac{\left(s_2^2/n_2\right)^2}{(n_2-1)}\right)}$$

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):

H_1	Critical Region/ Region of rejection of H_o
$H_1: \mu_1 - \mu_2 \neq 0$ ($\mu_1 \neq \mu_2$) or $\mu_1 - \mu_2 \neq d$ (two-tailed)	$z < -z_{\alpha/2}$ or $z > z_{\alpha/2}$
$H_1: \mu_1 - \mu_2 > 0$ ($\mu_1 > \mu_2$) or $\mu_1 - \mu_2 > d$ (right-tailed)	$z > z_\alpha$
$H_1: \mu_1 - \mu_2 < 0$ ($\mu_1 < \mu_2$) or $\mu_1 - \mu_2 < d$ (left-tailed)	$z < -z_\alpha$

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

H_1	Critical Region/ Region of rejection of H_o
$H_1: \mu_1 - \mu_2 \neq 0$ ($\mu_1 \neq \mu_2$) or $\mu_1 - \mu_2 \neq d$ (two-tailed)	$t < -t_{\alpha/2}(m)$ or $t > t_{\alpha/2}(m)$
$H_1: \mu_1 - \mu_2 > 0$ ($\mu_1 > \mu_2$) or $\mu_1 - \mu_2 > d$ (right-tailed)	$t > t_\alpha(m)$
$H_1: \mu_1 - \mu_2 < 0$ ($\mu_1 < \mu_2$) or $\mu_1 - \mu_2 < d$ (left-tailed)	$t < -t_\alpha(m)$

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 and Z denotes a general standard normal random variable):

H_1	p-value
$H_1: \mu_1 - \mu_2 \neq 0$ ($\mu_1 \neq \mu_2$) or $\mu_1 - \mu_2 \neq d$ (two-tailed)	2* $P(Z < z)$ if z is negative 2* $P(Z > z)$ if z is positive
$H_1: \mu_1 - \mu_2 > 0$ ($\mu_1 > \mu_2$) or $\mu_1 - \mu_2 > d$ (right-tailed)	$P(Z > z)$
$H_1: \mu_1 - \mu_2 < 0$ ($\mu_1 < \mu_2$) or $\mu_1 - \mu_2 < d$ (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 m):

H_1	p-value
$H_1: \mu_1 - \mu_2 \neq 0$ ($\mu_1 \neq \mu_2$) or $\mu_1 - \mu_2 \neq d$ (two-tailed)	2* $P(T < t)$ if z is negative 2* $P(T > t)$ if z is positive
$H_1: \mu_1 - \mu_2 > 0$ ($\mu_1 > \mu_2$) or $\mu_1 - \mu_2 > d$ (right-tailed)	$P(T > t)$
$H_1: \mu_1 - \mu_2 < 0$ ($\mu_1 < \mu_2$) or $\mu_1 - \mu_2 < d$ (left-tailed)	$P(T < t)$

An examination was given to two classes consisting of 40 and 50 students, respectively. In the first class the mean grade was 74 with a standard deviation of 8, while in the second class the mean grade was 78 with a standard deviation 7. Is there a significant difference between the performance of the two classes at a level of significance of 0.05? What is the P value of the rest?

The research investigator is interested in studying whether there is a significant difference in the salaries of B.Tech grades in two metropolitan cities. A random sample of size 100 from Kolkata yields an average income of Rs. 20,150. Another random sample of 60 from Delhi results in an average income of Rs.20,250. If the variances of both the populations are given as ${\sigma_1}^2 = Rs. 40{,}000$ and ${\sigma_2}^2 = Rs. 32{,}400$ respectively.

A manufacturer claims that the average tensile strength of thread A exceeds the average tensile strength of thread B by at least 12 kilograms. To test this claim, 50 pieces of each type of thread were tested under similar conditions. Type A thread had an average tensile strength of 86.7 kilograms with a standard deviation of 6.28 kilograms, while type B thread had an average tensile strength of 77.8 kilograms with a standard deviation of 5.61 kilograms. Test the manufacturer's claim using a 0.05 level of significance.