* Hypothesis Testing :-

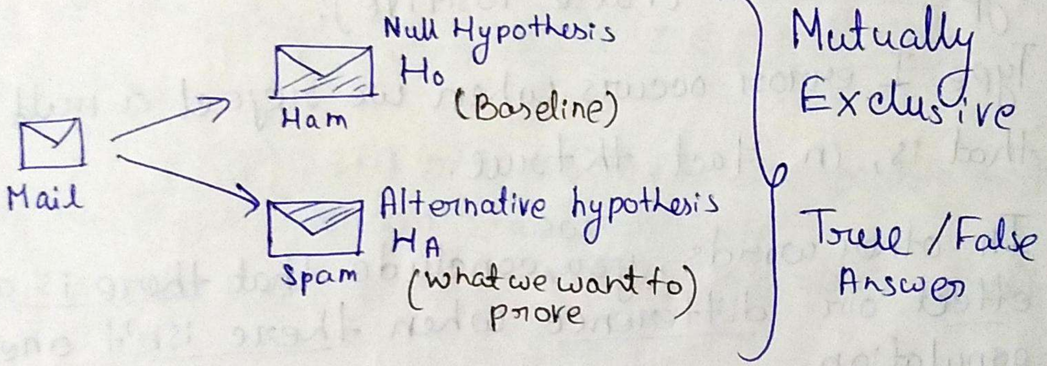
Hypothesis testing is a statistical method used to make inferences about population parameters based on sample data.

Initially, a tentative assumption is made about the parameter or distribution. (H_0)

Also an alternative statement is also stated (H_A or H_1).

If we get enough evidence against H_0, we choose H_A.

Eg:-



→ To determine the result of test :-

Plenty of evidence againts H_0 ⟶ Reject H_0 (and accept H_A)

* Type 1(I) & Type 2(II) error :

Decision	Reality: H_0 True (Not spam)	Reality: H_0 False (Spam)
Reject H_0 (Decide spam)	$\boxed{\text{Type I error}}$	Correct
Don't Reject H_0 (Decide not spam)	Correct	$\boxed{\text{Type II error}}$

These are the 2 different ways in which errors can occur when conducting hypothesis testing.

① Type I error (False Positive):

→ Type I error occurs when we reject a null hypothesis that is, in fact, ~~th~~ true.

→ In other words, ~~you~~ we conclude that there is a significant effect or difference when there isn't one in the population.

→ The probability of making a Type I error is denoted by α (alpha), and is also called significance level.

Common choices of α are 0.05 (5%) or 0.01 (1%).

Setting α lower, we reduce chance of Type I error by increases risk of Type II error.

(There is a trade off)

② Type II error :- (False Negetive)

→ Type II error occurs when we fail to reject a null hypothesis that is actually false.

→ In this case, we conclude that there is no significant effect or difference when there is one in the population.

→ The probability of making a Type II error is denoted by β(beta)

* P-Value, Critical Value and Significance Level:

① Significance level:-

It is often denoted as α (alpha), is a predefined threshold that researchers set before conducting a hypothesis test.

It represents the maximum acceptable risk of making a Type I error (False poisitive).

(Mostly $\alpha = 0.05$ or 0.01)

② P-Value:-

The p-value, on the other hand, is a statistical calculated from the sample data, representing the stregth of evidence againt the null hypothesis.

Researchers use the p-value to determine whether the null hypothesis should be rejected based on the chosen significance level.

→ While α is fixed and determined by the researchers, the p-value is data-dependent and can vary with different datasets.

③ Critical Value:-

A critical value, also known as a critical score or cutoff value, is a pre-determined ~~thes~~ threshold or specific value from a probability distribution that is used in statistical hypothesis testing. and corresponds ~~the~~ to the selected significance level (α).

It represents the point beyond which we would reject the null hypothesis.

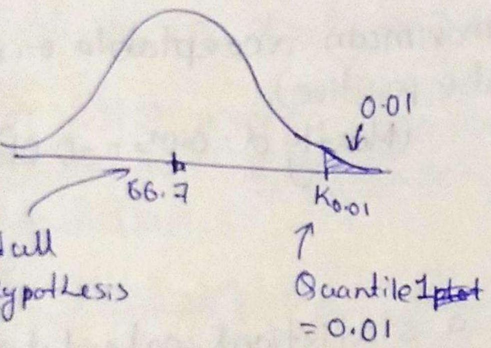
(It's usually found in statistical tables or can be calculated using software)

Eg:- $H_0: \mu = 66.7$ vs $H_0: \mu > 66.7$

$\alpha = 0.01$

Given, $n = 10$, $\sigma = 3$, $\bar{x} = 68.442$

$0.01 = P(\bar{X} > K_{0.01} \mid \mu = 66.7)$



If $\mu = 66.7$, $\bar{X} \sim N\left(66.7, \frac{3^2}{10}\right)$

$\boxed{K_{0.01} = 68.91}$

So, if $\bar{x} > 68.91$ reject H_0

→ Power of test:-

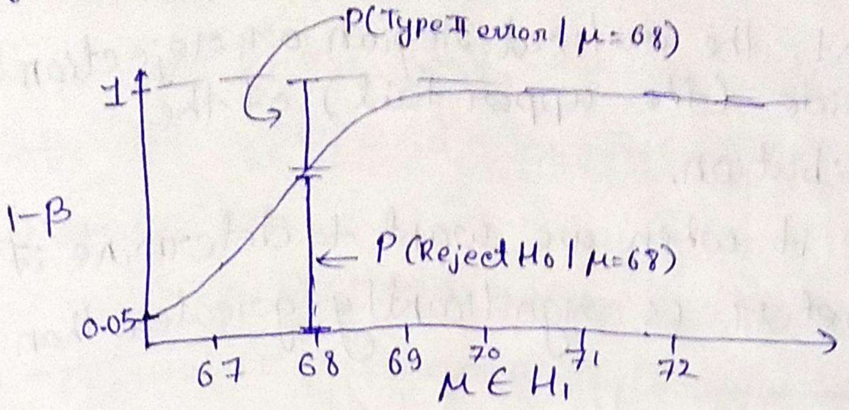
The power of test is a measure of the test's ability to detect a true effect or difference when it exist in the population.

Basically it quantifies the probability of correctly rejecting a false null hypothesis.

Type II error : $P(\text{Do not Reject } H_0 \mid \mu \in H_1) = \beta$

Power of test : $P(\text{Reject } H_0 \mid \mu \in H_1) = (1-\beta)$

Eg: Graph:



* Right-Tailed, Left-Tailed & Two-Tailed Tests:-

Note:-
Before starting a hypothesis test, we should ensure the reliability of data.

Reliability →
- Each sample has to be representive of the population.
- The data needs to be completely randomized.
- Sample size should be adequet. (Rule of thumb → 30 samples or more)

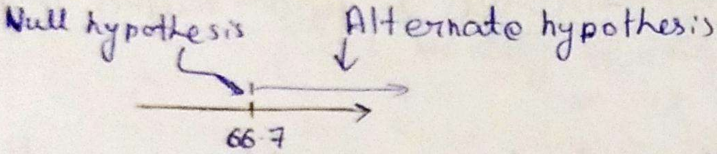
→ Right-tailed, Left-tailed and two-taild tests are the concepts that determins whether a test focuses on one side of distribution, the other side, or both sides.

• Right-Tailed Test:-

⇒ In a right-tailed test, the critical region or rejection is on the right side (the upper tail) of the probability distribution.

We use it when we want to determine if a population parameter is significantly greater than a specified value.

Eg:- $H_0 : \mu = 66.7$ vs $H_1 : \cancel{\mu =}\ \mu > 66.7$

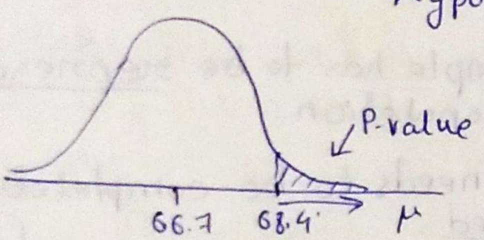


if $\bar{x} >> 66.7 \Rightarrow$ Reject Hypothesis

Type I error:
Determine $\mu > 66.7$, when μ did not change

Type II error:
Do not reject that $\mu = 66.7$, when in true $\mu > 66.7$

Eg:-



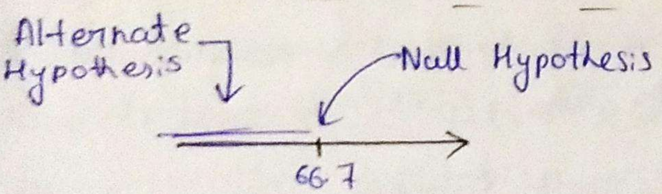
if p-value $< \alpha \Rightarrow$ reject H_0

• ~~Left~~ Left-Tailed Test:-

→ In a left-tailed test, the critical region or rejection is on the left side (the lower tail) of the probability distribution.

We use it determine if a population parameter is significantly less than a specified value.

Eg:- $H_0 : \mu = 66.7$ vs ~~$H_1 : \mu > 66.7$~~ $H_1 : \mu < 66.7$

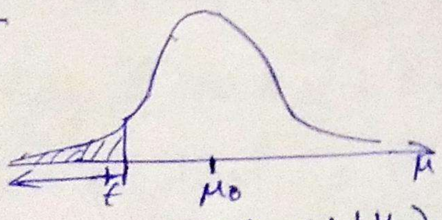


if $\bar{x} << 66.7 \Rightarrow$ Reject H_0

Type I error:
Determine $\mu < 66.7$, when population mean did not change.

Type II error:
Don't reject that $\mu = 66.7$. when in true $\mu < 66.7$

Eg:-



P-value : $P(T(X) < t \mid H_0)$

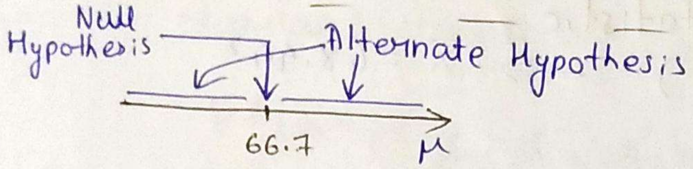
→ Two-Tailed Test :-

- In a Two-Tailed Test, the critical regian is split into both the upper and lower-tails of the probability distribution.

It is used when we want to determine if a population parameter is ~~split into both~~ significantly different from a specific value (not just greater or less)

- If the test statistic falls in either tail (greater than the upper critical value or smaller than the lower critical value) then we reject H_0 in favor of H_A.

Eg:- $H_0 : \mu = 66.7$ vs $H_1 : \mu \neq 66.7$



if $\bar{x} \gg 66.7$ or $\bar{x} \ll 66.7$ $\Rightarrow$ Reject H_0

Type I error:
Determine $\mu \neq 66.7$, when population mean did not change.

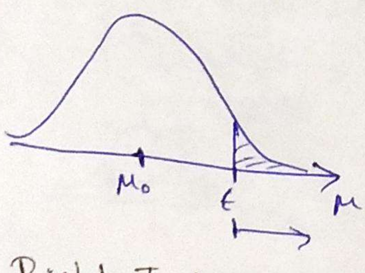
Type II error:
Don't reject that $\mu = 66.7$ when true $\mu \neq 66.7$

Eg:-

$T(X)$: Test Statistics

t : observed statistic

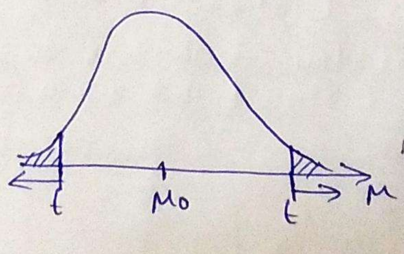
$H_0 : \mu = \mu_0$



Right-Tailed Test

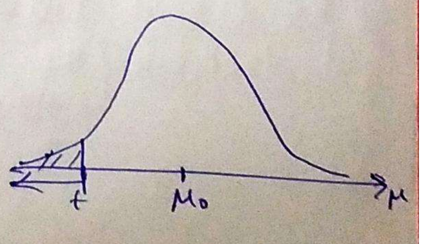
$P(T(X) > t \mid H_0)$

P-value



Two-Tailed Test

$P(|T(X) - \mu_0| > |t - \mu_0| \mid H_0)$



Left-Tailed Test

$P(T(X) < t \mid H_0)$

* Steps for Performing Hypothesis Testing :-

① State the hypotheses.

- Null hypothesis : the baseline → $H_0 : \mu = 66.7$
- Alternative hypothesis: the statement we want to prove → $H_1 : \mu > 66.7$

② Design the test.

- Decide the test statistic to work with. ⇒ $\overline{X}$
- Decide the significance level. ⇒ $\alpha = 0.05$

③ Compute the observed statistic (based on your sample) → $\overline{x} = 68.442$

④ Reach a conclusion:

- If the p-value is less than the significance level reject H_0

↳ $P(\overline{X} > 68.442 \mid \mu = 66.7) \overset{?}{<} 0.05$

*Chebyshev's Inequality :-

Chebyshev's inequality is a fundamental theorem in probability and statistics that provides a bound on the probability that a random variable takes values within a certain distance from its mean.

It is a useful tool for understanding the spread on variability of data, especially when the probability distribution of the random variable is unknown.

- Chebyshev's inequality states that, for any finite mean (μ) and a finite non-zero standard deviation (σ), the probability that the random variable deviates from its mean by more than K standard deviations is at most $1/K^2$, where $K>1$.

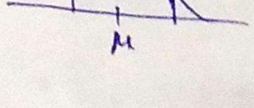
$$P(|X-\mu| \geq K\sigma) \leq \frac{1}{K^2}$$

- Further, It also upholds that $1-\left(\frac{1}{K^2}\right)$ of a distribution's values must be within, but not including, K, ~~σ away~~ Standard deviations away from the mean of the distribution.

$$P(\mu - K\sigma < x < \mu + K\sigma) > 1 - \frac{1}{K^2}$$

Eg:- $K=2$

$$P(\mu - 2\sigma < x < \mu + 2\sigma) > \underbrace{1 - \frac{1}{2^2}}_{\frac{3}{4} = 75\%}$$



$\frac{3}{4} = 75\%$

* Z-test & t-test :-

→ Z-test:-

Z-test is used when you have a Known population standard deviation (σ).

It is generally used for relatively large sample size (typically, $n > 30$). However, for smaller sample sizes, it can still be used if the data is normally distributed.

Formula: $$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}}$$

→ z-test follows a standard normal distribution. ($\mu = 0, \sigma = 1$)

→ T-test:-

The t-test is used when we don't know the population standard deviation (σ), and we estimate using the sample standard deviation (s).

It can be used for smaller sizes making it more versatile than the z-test.

Formula: $$t = \frac{\bar{X} - \mu}{s/\sqrt{n}}$$

→ The t-test follows t-distribution with n-1 degrees of freedom. (The t-distribution has heavier tails than the normal distribution & varies based on degrees of freedom)

Eg:-

$$T = \frac{\bar{X} - \mu}{s/\sqrt{10}} \sim t_9$$

$n = 10$

$\nu = 10 - 1$

Degrees of freedom

⇒ Variations in T-Tests :-

① **One-Sample T-Test :-**

Used to compare the mean of a single sample to a known population mean.

② **Independent Two-Sample T-Test :-**

Used to compare the means of two independent groups (Eg. treatment vs control) to determine if they are significantly different.

③ **Paired (Dependent) T-Test :-**

Used to compare means of paired or matched data points (Eg. before and after measurements) within the same group.

→ Independent Two-Sample t-Test : Two Tailed Test:-

An independent two-sample ~~t-test~~ t-test, also known as a two-sample t-test or an unpaired t-test, is a statistical difference b/w the means of two independent groups.

⇒ Assumptions :-

- Each person in both the samples are independent.
- All people in the sample from the two groups are different (Random sampling)
- Populations are approximately normally distributed.
- Populations having equal variance (Homoscedasticity).

Eg:-

$X \sim N(\mu_{US}, \sigma^2_{US})$ $\quad$ $Y \sim N(\mu_{Avg}, \sigma^2_{Avg})$

$$\bar{X} = \frac{1}{10}\sum_{i=1}^{10} x_i \quad ; \quad \bar{Y} = \frac{1}{9}\sum_{i=1}^{9} y_i$$

Then,

$$\bar{X} - \bar{Y} \sim N\left(\mu_{US} - \mu_{Arg}, \frac{\sigma_{US}^2}{10} + \frac{\sigma_{Arg}^2}{9}\right)$$

$$\Downarrow$$

$$\frac{(\bar{X}-\bar{Y}) - (\mu_{US} - \mu_{Arg})}{\sqrt{\frac{\sigma_{US}^2}{10} + \frac{\sigma_{Arg}^2}{9}}} \sim N(0,1)$$

When, we don't know, σ_{US} & σ_{Arg} we replace them with S_X^2 & S_Y^2

So,

$$T = \frac{(\bar{X}-\bar{Y}) - (\mu_{US} - \mu_{Arg})}{\sqrt{\frac{S_X^2}{10} + \frac{S_Y^2}{9}}} \sim t_\nu$$

$SE(\hat{P}_X - \hat{P}_Y)$

$$\nu(\text{Degrees of freedom}) = \frac{\left(\frac{S_X^2}{n_X} + \frac{S_Y^2}{n_Y}\right)^2}{\frac{\left(\frac{S_X^2}{n_X}\right)^2}{n_X - 1} + \frac{\left(\frac{S_Y^2}{n_Y}\right)^2}{n_Y - 1}}$$

→ Pooled standard deviation is used in the case of independent two-sample t-tests when we are comparing the means of two groups. It is used to estimate a common standard deviation for both groups.

$$S.D. = \sqrt{\frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}}$$

For computing common proportion we can estimate it by ~~using~~ pooling the samples.

Eg:- $\overset{p_1}{0.55} \times \overset{n_1}{1000} = 550$ voters,

$\overset{p_2}{0.58} \times \overset{n_2}{1500} = 870$ voters.

Then,
$$\frac{550+870}{1000+1500} = \frac{1420}{2500} = 56.8\%$$

So, we estimate $SE(\hat{p}_2 - \hat{p}_1) \doteq \sqrt{\frac{0.568(1-0.568)}{1000} + \frac{0.568(1-0.568)}{1500}}$

$= 0.02022$

→ Paired (Dependent) T-test:-

The paired sample t-test also sometimes called the dependent t-test, is a statistical procedure used to determine whether the mean difference between two sets of observations is zero.

In a paired sample t-test, each subject or entity is measured twice, resulting in pairs of observations.

It's common application would be when, we try to observe "how well a training program was?" or "how effective was a drug".

→ Assumptions:-

- The dependent variable must be continuous (interval/ratio)
- The observations are independent of one another.
- The dependent variable should be approximately normally distributed.
- The dependent variable should not contain any outliers.

→ Method:-

x	y
X_1	Y_1
X_2	Y_2
X_3	Y_3
$\vdots$	$\vdots$
X_n	Y_n

$x \rightarrow$ People ~~after~~ before a drug treatment

$y \rightarrow$ People after a drug treatment

(X_1, Y_1) are the same people before and after drug treatment.

Then,

$$\frac{(X_1-Y_1)+(X_2-Y_2)+(X_3-Y_3)+\dots+(X_n-Y_n)}{n}$$

$$\Rightarrow \bar{D} = \frac{D_1 + D_2 + D_3 + \dots + D_n}{n}$$

$\Rightarrow$ If X_i, Y_i are ~~gauss~~ gaussian $\Rightarrow D_i = X_i - Y_i$ is gaussian.

$$D_i \overset{i.i.d.}{\sim} N(\mu_D, \sigma_D^2)$$

Then,

$$\frac{\bar{D} - \mu_D}{\sigma_D/\sqrt{10}} \sim N(0, 1^2)$$

But σ_D is unknown

so, we consider s_D

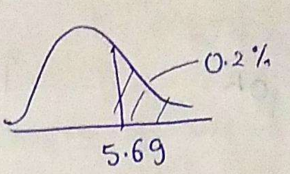
$$\sigma_D \rightarrow s_D = \sqrt{\frac{\sum_{i=1}^{n}(D_i - \bar{D})^2}{n-1}}$$

Then,

$$T = \frac{\bar{D} - \mu_D}{s_D/\sqrt{10}} \sim t_{10-1} = t_9$$

This will consider the p-value

Eg:- T=5.69



* Parametric & non-parametric test :-

→ Parametric test is a statistical test which assumes parameters and the distributions about the population are known.

It uses a mean value to measure the central tendency.

→ Non-parametric test does not make any assumptions and measures the central tendency with the median values.

→ Although non-parametric tests requires less assumptions than parametric tests.

Parametric tests are mor powerful than non-parametric tests.

Meaning parametric test gives a better guarantee when we reject a null hypothesis than in case for non-parametric test.

Properties	Parametric	Non - Parametric
Assumptions	Yes / many	No / Few
Central tendency values	Mean value	Median value
Correlation	Pearson	Spearman
Probability distribution	Normal	Arbitrary.
Population Knowledge	Requires	Does not require
Used for	Interval Data	Nominal Data
Applicability	Variables	Attributes & Variables
Eg:-	z-test , t-test	Kruskal-Wallis, Man-Whitney

	Parametric tests	Non-Parametric Tests
One sample →	sampt t-test	Wilcoxon test for one sample
Two dependent samples →	Paired sample t-test	Wilcoxon test
Two ~~ide~~ independent samples →	Unpaired sample t-test	Mann-Whitney U Test
More than two independent samples →	One factorial ANOVA	Kruskal-Wallis Test
More than two dependent samples →	Repeated Measures ANOVA	Friedman Test
Correlation between two samples. →	Pearson-Korrelation	Spearman-Korrelation

* Sign Test :-

→ The sign test is a non-parametric test is used to test whether or not groups are equally sized.

The sign test is used when dependent samples are ordered in pairs, where the bivariate random variables are mutually independent.

It is based on the direction of the plus and minus sign of the observation, and not on their numerical magnitude. It is a binomial sign test, with $p = 0.5$...

→ The sign test is considered a weaker test, because it tests the pair value below or above the median and it does not measure the pair difference.

→ Assumptions:-

- Data distributions :- The sign test is a non-parametric test (distribution free) so we do not assume that the data is normally distributed.

- Two sample :- Data should be from two samples. The population may differ for the ~~following~~ two samples.

- Dependent sample :- Dependent samples should be a paired sample or matched. Also known as 'before-after' sample.

Suppose we want to test the hypothesis that median (η) of a population has a specified value, say η_0, i.e.,

$$H_0 : \eta = \eta_0$$

Vs $H_1 : \eta \neq \eta_0$ (Two-tailed) or $H_1 : \eta > \eta_0$ (Right-tailed)

or $H_1 : \eta < \eta_0$ (Left-tailed)

→ Procedure:-

Let $X_1, X_2, X_3, \ldots, X_n$ be a random sample of size n from the given population with median $\eta = \eta_0$ (under H_0).

⇒ Subtract, η_0 from each X_i's and write

or Replace each observation with,

- exceeding η_0 ⇒ with plus (+) sign.
- less than η_0 → with minus (+) sign,
- equal to η_0 with zero (0).

→ By the definition of Median, we have

$$P(X > \text{Median}) = P(X < \text{Median}) = \frac{1}{2}$$

Thus, under H_0 ($\eta = \eta_0$):, we have

$$P(X > \eta_0) = P(X < \eta_0) = \frac{1}{2}$$

Hence, if H_0 is true, then the number of + signs should be approximately equals to the - signs.

If the difference in the number of plus (+) and minus (-) signs is due to chance variations (or fluctuations of sampling), then we fails to reject the H_0.

(Accept H_1)

Notations:- (After discarding zeros)

- T^+ = Number of positive sign
- T^- = Number of negative sign
- $T = \min(T^+, T^-)$
- $n = T^+ + T^-$

Main Procedure:-

The following steps are summarized:-

Step 1: Set the Hypothesis:-

Null Hypothesis:-	Alternative Hypothesis:-
$H_0 : \eta = \eta_0$	$H_1 : \eta < \eta_0$ or $H_1 : \eta > \eta_0$ or $H_1 : \eta \neq \eta_0$

Step 2: Compute T^+, T^-

Subtract η_0 from each observations
Discard zeros and hence
compute T^+, T^- values.

Step 3: Test statistics:

$$T = \min(T^+, T^-)$$

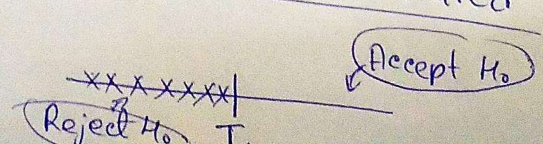
[for $n < 25$]

Step 4: Critical Region:

Define the critical region as $T \leq T_c$

Where, T_c is the critical value of T at given level of significance for one-tailed or two-tailed

[T_c is found from the table]



→ For Large samples, $n > 25$

Under H_0: $P(T > \eta_0) = P(T < \eta_0) = 0.5$

P = Probability of + signs $= 0.5 = \frac{1}{2}$
which is constant for each trial.

Hence, under H_0, the variable T has Binomial distribution with parameter n and $p = 0.5$, i.e., $T \sim Bino(n, 0.5)$

Therefore,

$$\text{Mean} = E(T) = np = n \times 0.5 = \frac{n}{2}$$

$$\sigma = \sqrt{var(T)} = \sqrt{npq} = \sqrt{\frac{n}{4}} = \frac{\sqrt{n}}{2} \qquad |\because var(T) = npq$$

So, the 3rd step of the procedure will change for large samples:-

→ Computing test statistics:-

If $T^+ < \frac{n}{2}$

$$Z = \frac{(T^+ + 0.5) - \frac{n}{2}}{\frac{1}{2}\sqrt{n}} \sim N(0,1)$$

If $T^+ > \frac{n}{2}$

$$Z = \frac{(T^+ - 0.5) - \frac{n}{2}}{\frac{1}{2}\sqrt{n}} \sim N(0,1)$$

⇒ Paired Sample Sign Test :-

It is a variation of sign test used in case of dependent or paired sample data

Let $(X_i, Y_i), i = 1, 2, \ldots, n$ be a collection of paired observations from two continuous population.

X	Y
⋮	⋮
n_1	n_2

To perform a paired test n_1 and n_2 must be $\boxed{n_1 = n_2}$

and $D_i = X_i - Y_i, i = 1, 2, \ldots, n$ be the paired differences.

Based on D_i we assign the signs.

Every step after it is identical to normal sign test.

Eg:-

X	Y	D	sign
4	5	-1	-
7	3	+4	+
9	9	0	~~0~~
1	4	-3	-
11	14	-3	-

Then, $T^+ = 1$

$T^- = 3$

$n = T^+ + T^- = 4$

$T = \min(T^+, T^-)$

$= 1$

Here, when $n \geq 10 \rightarrow$ Large sample

when $n < 10 \rightarrow$ small sample.

* Wilcoxon Signed-Rank Test:

→ The wilcoxon Signed - Rank test is, used in similar situations as the Mann-Whitney U-test.

The main difference is that the Mann-Whitney U-test tests two independent samples, whereas the Wilcox sign test tests two depent samples

→ The Wilcoxon signed rank test pools all differences, ranks them. and applies a negative sign to all the ranks where the difference between the two observations is negative. ~~and~~ This is called the signed rank.

→ The Wilcoxon test can be good alternative to the t-test when population means are not of interest; (Rather interest is the median).

→ The sign test is not very sensitive to the amount of variation of a data value from the conjectured value of the median. It uses only the information whether the value is above or below the median value.

To overcome it, a new test named as Wilcoxon Signed-Rank test is used, in addition to the signs, also uses the amount of variation of data values from the conjectured values and Hence is considered to be better than the sign test

Here the procedure is identical ~~as~~ for signed test the only changes are:

① Step 2: Computing T^+ & T^-

② Step 3: Test statistics (Only for large samples $n \geq 30$)

1) Step 2: Complete T^+, T^-

$\boxed{\text{Subtract } \eta_0 \text{ from each observation. Discard zeroes and hence compute } T^+, T^- \text{ values}}$

Eg:- $\boxed{\text{Median}=3}$

X	Y=X-Median	|Y|	Sign	Rank of |Y|
6	3	3	+	8
5	2	2	+	5.5
1	-2	2	−	5.5
2	-1	1	−	2
2	-1	1	−	2
5	2	2	+	5.5
7	4	4	+	10
5	2	2	+	5.5
3	0	0		
7	4	4	+	10
4	1	1	+	2
7	4	4	+	10

When, ~~we~~ we have multiple |Y| same value

Like,

→ 1, 1, 1 for $1^{st}, 2^{nd}, 3^{rd}$ rank, we assign these |Y| the avg ranks, $\frac{1+2+3}{3} = 2$

→ 2, 2, 2, 2 for $4^{th}, 5^{th}, 6^{th}, 7^{th}$, $\frac{4+5+6+7}{4} = 5.5$

$T^+ =$ ~~Pos.~~ Sum of positive Ranks

$T^- =$ sum of Negetive Ranks

$T^+ = 8 + 5.5 + 5.5 + 10 + 5.5 + 10 + 2 + 10$

$= 56.5$

$T^- = 5.5 + 2 + 2$

$= 9.5$

$T = \min(T^+, T^-) = 9.5$

---

2) Step 3: Test statistics:-

when, $n \geq 30$, we use the normal test

$$Z = \frac{T - \mu_T}{\sigma_T} \sim N(0,1)$$

where $T = \min(T^+, T^-)$

$$\mu_T = \frac{n(n+1)}{4}$$

$$\sigma_T = \sqrt{\frac{n(n+1)(2n+1)}{24}}$$

Then, check the value of observed z with the critical value of z derived from the question.

(Either given in question or find from z-table)

* Mann-Whitney U test :-

(Also called Mann-Whitney-Wilcoxon (MWU/MWW), Wilcoxon rank-sum test, or Wilcoxon-Mann-Whitney test)

It is a non-parametric test of the null hypothesis that, for randomly selected values X and Y from two populations, the probability of X being greater than Y is equal to the probability of Y being greater than X.

→ **Assumptions:-**

1. All the observations from both groups are independent of each other.
2. The responses are atleast ordinal.
3. Under the H_0, ~~both~~ the distribution of both ~~procedures~~ populations are identical.
4. The alternative H_1, the distributions are not identical.

Let $(X_1, X_2, \dots X_{n_1})$ and $(Y_1, Y_2, \dots, Y_{n_2})$ be independent random samples of size n_1 & n_2 from two populations with ~~p.d.f's~~ p.d.f.'s $f_1(x)$ & $f_2(y)$ respectively.

[We want to test if the samples have been drawn from the same population, that is $f_1(x) = f_2(x)$ for all x]

The Mann-Whitney test uses the sum of ranks (or rank-sum) for each sample.

→ Procedure :-

1. **Step 1:** (Similar to signed test)
   Set the Hypothesis

2. **Step 2:**
   Compute U_x, U_y

3. **Step 3:**
   Test statistics - U

4. **Step 4:** (Find from Mann-Whitney U table)
   Critical Region:

Changes :-

2. **Step 2:** Compute U_x, U_y

→ T_x : Sum of the ranks of the observation in X sample
→ T_y : " " " " " Y sample

→ $$\boxed{U_x = T_x - \frac{n_1(n_1+1)}{2}}$$ & → $$\boxed{U_y = T_y - \frac{n_2(n_2+1)}{2}}$$

→ n_1 = sample size of X
→ n_2 = " " " Y

→ $U = \min(U_x, U_y)$

The following equalities always satisfy:

→ $U_x + U_y = n_1 n_2$

→ $T_x + T_y = \frac{(n_1+n_2)(n_1+n_2-1)}{2}$

Eg:

X : 60 45 23 32

Y: 10 25 20 54 32 65 8

⇓

Sample : Y Y Y X Y X Y X Y X Y

Data : 8 10 20 23 25 32 32 45 54 60 65

Rank : 1 2 3 $\boxed{4}$ 5 $\boxed{6.5}$ 6.5 $\boxed{8}$ 9 $\boxed{10}$ 11

$T_x = 4+6.5+8+10$
$= 28.5$

$Ty = 1+2+3+5+6.5+9+11$
$= 37.5$

Then,

$U_x = T_x - \frac{n_1(n_1+1)}{2}$
$= 28.5 - \frac{4(4+1)}{2}$
$= 18.5$

$U_y = T_y - \frac{n_2(n_2+1)}{2}$
$= 37.5 - \frac{7(7+1)}{2}$
$= 9.5$

Therefore,

$U = \min(U_x, U_y)$
$= 9.5$

Now, compare it with critical value of U to see if we may accept H_0 or reject it.

When, $n_1 \leq 10$ & $n_2 \leq 10$, it is small sample.
otherwise large sample

③ Step 3: Compute the test statistics:-

$$z = \frac{U - \mu_U}{\sigma_U} \sim \mathcal{N}(0, 1)$$

where, $U = \min(U_x, U_y)$;

$$\mu_U = \frac{n_1 n_2}{2}$$

$$\sigma_U = \sqrt{\frac{n_1 n_2 (n_1 + n_2 + 1)}{12}}$$

Then, we use this z value to compare with critical value of z derived from the question.

* Bootstrap Hypothesis Test:-

It implements the concept of bootstraping:-

→ Elements of Hypothesis Test:-

① Specify H_0, H_A

② Choose Test-stat → Test-stat$_1 = |\bar{Y}_c - \bar{Y}_M|$ (mean)

→ Test-stat$_2 = |\hat{Med}_c - \hat{Med}_M|$ (median)

③ Determine distribution of test stat

④ Convert Test-stat → P-value.

$$\text{P-value} = \frac{\text{No. of Bootstrap-Test-stats} \geq \text{Observed Test-stat}}{\text{No. of Bootstraps}}$$

Reasons for using Bootstrap:-

→ Smaller sample size

→ Standard Error and /or sampling Distribution for test ~~stasti~~ statistic is difficult to work out

* Permutation Tests:-

A permutation test (also called re-randomized test or shuffle test) is an exact statistical hypothesis test making use of the proof by contradiction.

The method by which treatments are allocated to subjects in an experimental design is mirrored in the analysis of that design.

They are based on the idea of repeatedly shuffling or permuting the data to generate a null distribution under the assumption of no effect.

⇒ These tests are particularly useful when the assumptions of traditional parametric tests cannot be met or when the sample size is small.

⇒ Procedure :-

① Formulate the hypothesis :-

② Choose a test statistic :- (Mean diff, Median diff, etc)

④ Calculate the test statistic for each permutation.

⑤ Compare the observed Test statistic to the Permutation Distribution.

⑥ Calculate p-value

⇒ Permutation tests may be computationally intensive, especially with large datasets, as all possible permutations need to be considered.

---

* ANOVA :-

⇒ It stands for "Analysis of Variance" is an extremely important tool for analysis of data (both One way and Two way ANOVA is used).

It is ~~the~~ a statistical method to compare the population means of two or more groups by analyzing variance.

The variance would differ only when the means are significantly different.

→ ANOVA test is the way to find out if survey or experiments are significant.

We are testing groups to see if there's a difference between them.

→ One-way ANOVA :- (CRD → Completely Randomized Design)

⇛ **Purpose :**

Used when comparing means across more than two groups (treatments or levels) for a single independent variable.

Eg:- Comparing the mean scores of three different teaching methods.

→ It deals with one independent variable.

[Number of Independent Variables]

~~→ It has a one-dimensional design matrix. [Design Matrix]~~

→ It provides information about differences in means across different levels of single factor.

[Interpretation]

→ It compares three or more levels of one factor.

Basic Terms in ANOVA

① Experimental unit - The object on which a measurement (or measurements) is taken.

② Factor — It is an independent variable whose values are controlled and varied by the experimenter.

③ Level — It is the intensity setting of a factor

④ Treatment - It is a specific combination of factor levels.

Assumption for ANOVA :-

① Samples follow normal distribution.

② Samples have been selected randomly and independently.

③ Each group should have common variance

④ Data are independent. [one-way]

⑤ Additivity of variance : Total variance = Between variance + Within Variance

Note:-

→ $\text{ANOVA} = \dfrac{\text{Variance Between}}{\text{Variance Within}}$

→ Total variance = Variance Between + Variance Within

ANOVA > 1 [Reject H_0]

ANOVA $\leq$ 1 [Fail to Reject H_0]

→ Three Kinds of ~~variantions~~ variations in ANOVA :-

① Between Groups : Variation from one group to another

$$\sum n_i (\bar{x}_i - \bar{x})^2$$

② Within Groups : Variation among the observations of each specific group.

$$\sum\sum (x_{ij} - \bar{x}_i)^2$$

③ Total : Variations among all the observations (which is nothing but the sum of Between Groups & within Groups)

$$\sum\sum (x_{ij} - \bar{x})^2$$

Eg:-

Metods:	A	B	C
	10	8	9
	9	9	8
	8	10	7
	7.5	8	10
	8.5	8.5	9
	9	7	8
	10	9.5	7
	8	9	10
	8	7	9
	9	10	8
Mean:	8.7	8.6	8.5

Mean 8.6 ($\bar{\bar{y}}$)

Between Group Variation

$= 10(8.7-8.6)^2 + 10(8.6-8.6)^2 + 10(8.5-8.6)^2$

$= 0.2$

Within group variation:-

A : $(10-8.7)^2 + (9-8.7)^2 + \dots (9-8.7)^2 = 6.6$ ($\bar{y}_A$)

B : $(8-8.6)^2 + (9-8.6)^2 + \dots + (10-8.6)^2 = 10.9$ ($\bar{y}_B$)

C : $(9-8.5)^2 + (8-8.5)^2 + \dots + (8-8.5)^2 = 10.5$ ($\bar{y}_C$)

$6.6 + 10.9 + 10.5 = 28$

Method 1

$$\frac{\text{Variance Between}}{\text{Variance Within}} = \frac{0.2}{28} = \underline{0.0071} < 1$$

[Fail to reject H_0]

Method 2

$F_{stat} = \underline{0.0071}$

$$F_{critical} = \frac{\text{Numerator Degree of Freedom}}{\text{Denominator Degree of Freedom}}$$

→ Numerator Degree of Freedom = No. of samples − 1
$= 3 - 1 = \underline{2}$

→ Denominator Degree of Freedom $= \Sigma(n_j - 1) = n_T - K = 30 - 3$
$= \underline{27}$

$F_{critical} = F_{(2,27)} = 3.35$

$\therefore$ $F_{critical} > F_{stat}$; we fail to reject H_0.